

## **Final Report**

FDOT Contract No.: BDV31-977-29

UF Contract No.: 117252

# **Detection of Sinkholes or Anomalies Using Full Seismic Wave Fields: Phase II**

Principal Investigators: Michael C. McVay, Ph.D. (PI)  
Khiem Tran, Ph.D. (Co-PI)

Researchers: Scott Wasman, Ph.D.  
Brian Sullivan, M.S.  
Duminidu Siriwardane, M.S.

Department of Civil and Coastal Engineering  
Engineering School of Sustainable Infrastructure and Environment  
University of Florida  
P.O. Box 116580  
Gainesville, Florida 32611-6580

Developed for the



David Horhota, Ph.D., P.E., Project Manager

*October 2016*

## DISCLAIMER

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Florida Department of Transportation or the U.S. Department of Transportation.

Prepared in cooperation with the State of Florida Department of Transportation and the U.S. Department of Transportation.

## SI (MODERN METRIC) CONVERSION FACTORS (from FHWA)

### APPROXIMATE CONVERSIONS TO SI UNITS

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>LENGTH</b>				
<b>in</b>	inches	25.4	millimeters	mm
<b>ft</b>	feet	0.305	meters	m
<b>yd</b>	yards	0.914	meters	m
<b>mi</b>	miles	1.61	kilometers	km

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>AREA</b>				
<b>in<sup>2</sup></b>	square inches	645.2	square millimeters	mm <sup>2</sup>
<b>ft<sup>2</sup></b>	square feet	0.093	square meters	m <sup>2</sup>
<b>yd<sup>2</sup></b>	square yard	0.836	square meters	m <sup>2</sup>
<b>ac</b>	acres	0.405	hectares	ha
<b>mi<sup>2</sup></b>	square miles	2.59	square kilometers	km <sup>2</sup>

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>VOLUME</b>				
<b>fl oz</b>	fluid ounces	29.57	milliliters	mL
<b>gal</b>	gallons	3.785	liters	L
<b>ft<sup>3</sup></b>	cubic feet	0.028	cubic meters	m <sup>3</sup>
<b>yd<sup>3</sup></b>	cubic yards	0.765	cubic meters	m <sup>3</sup>

NOTE: volumes greater than 1000 L shall be shown in m<sup>3</sup>

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>MASS</b>				
<b>oz</b>	ounces	28.35	grams	g
<b>lb</b>	pounds	0.454	kilograms	kg
<b>T</b>	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>TEMPERATURE (exact degrees)</b>				
<b>°F</b>	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>ILLUMINATION</b>				
<b>fc</b>	foot-candles	10.76	lux	lx
<b>fl</b>	foot-Lamberts	3.426	candela/m <sup>2</sup>	cd/m <sup>2</sup>

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>FORCE and PRESSURE or STRESS</b>				
<b>Lbf</b>	poundforce	4.45	newtons	N
<b>kip</b>	kip force	1000	pounds	lbf
<b>lbf/in<sup>2</sup></b>	poundforce per square inch	6.89	kilopascals	kPa

**APPROXIMATE CONVERSIONS TO SI UNITS**

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>LENGTH</b>				
<b>mm</b>	millimeters	0.039	inches	in
<b>m</b>	meters	3.28	feet	ft
<b>m</b>	meters	1.09	yards	yd
<b>km</b>	kilometers	0.621	miles	mi

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>AREA</b>				
<b>mm<sup>2</sup></b>	square millimeters	0.0016	square inches	in <sup>2</sup>
<b>m<sup>2</sup></b>	square meters	10.764	square feet	ft <sup>2</sup>
<b>m<sup>2</sup></b>	square meters	1.195	square yards	yd <sup>2</sup>
<b>ha</b>	hectares	2.47	acres	ac
<b>km<sup>2</sup></b>	square kilometers	0.386	square miles	mi <sup>2</sup>

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>VOLUME</b>				
<b>mL</b>	milliliters	0.034	fluid ounces	fl oz
<b>L</b>	liters	0.264	gallons	gal
<b>m<sup>3</sup></b>	cubic meters	35.314	cubic feet	ft <sup>3</sup>
<b>m<sup>3</sup></b>	cubic meters	1.307	cubic yards	yd <sup>3</sup>

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>MASS</b>				
<b>g</b>	grams	0.035	ounces	oz
<b>kg</b>	kilograms	2.202	pounds	lb
<b>Mg (or "t")</b>	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>TEMPERATURE (exact degrees)</b>				
<b>°C</b>	Celsius	1.8C+32	Fahrenheit	°F

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>ILLUMINATION</b>				
<b>lx</b>	lux	0.0929	foot-candles	fc
<b>cd/m<sup>2</sup></b>	candela/m <sup>2</sup>	0.2919	foot-Lamberts	fl

SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>FORCE and PRESSURE or STRESS</b>				
<b>N</b>	newtons	0.225	poundforce	lbf
<b>kPa</b>	kilopascals	0.145	poundforce per square inch	lbf/in <sup>2</sup>

\*SI is the symbol for International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

## TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle <b>Detection of Sinkholes or Anomalies Using Full Seismic Wave Fields: Phase II</b>		5. Report Date <b>August 2016</b>	
		6. Performing Organization Code	
7. Author(s) <b>Michael C. McVay, Khiem T. Tran, Scott J. Wasman, Brian Sullivan, and Duminidu Siriwardane</b>		8. Performing Organization Report No.	
9. Performing Organization Name and Address <b>University of Florida – Dept. of Civil and Coastal Engineering Engineering School of Sustainable Infrastructure and Environment 365 Weil Hall – P.O. Box 116580 Gainesville, FL 32611-6580</b>		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. <b>BDV31-977-29</b>	
12. Sponsoring Agency Name and Address <b>Florida Department of Transportation 605 Suwannee Street, MS 30 Tallahassee, FL 32399</b>		13. Type of Report and Period Covered <b>Final Report 10/1/14 – 10/30/16</b>	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract <p>A new 2-D Full Waveform Inversion (FWI) software code was developed to characterize layering and anomalies beneath the ground surface using seismic testing. The software is capable of assessing the shear and compression wave velocities (<math>V_s</math> and <math>V_p</math>) for 1,200 cells (0.75 x 0.75 m resolution) in an 18 x 36 m (60 x 120 ft) region in 20 to 30 min on a standard laptop computer. The software, which includes a graphical software interface, allows the user to setup the test line (spacing, number of shots, and receiver), view the raw data, condition the data (filter, window, remove poor channels), preprocess the data (identify initial velocity profile), and analyze (invert) the conditioned data to obtain independent <math>V_s</math> and <math>V_p</math> results graphically or in file format (PDF).</p> <p>Next, a synthetic study was undertaken with the 2-D FWI software to investigate its capability of identifying 3-D voids (size, shape, and depth). Synthetic waveform data sets were generated using a 3-D finite difference model of a domain with a void of various sizes and depths, and surface-based test lines were both on the top of the void and away from the void. The synthetic waveform data sets were then analyzed by the 2-D FWI software as if they were collected from field experiments. The FWI results revealed that if the test line is on the top of a void, the maximum embedded depth at which the void can be detected is about three void diameters. In the case of test lines that are away from the void, the void can still be identified in inverted profile if the test line is near the void edge; however, the void becomes distorted and non-existent if the test line is at least one diameter from the void.</p> <p>The 2-D FWI software was then tested at five FDOT sites with unknown soil/rock layering and both known and unknown voids at time of seismic testing. Subsequent invasive testing (SPT and CPT) was performed near or at the anomaly for validation. A comparison of the seismic results with the SPT/CPT results revealed that FWI did an excellent job of identifying soil and rock layering (e.g., soil and limestone depths at Kanapaha and Newberry, FL), identifying unknown voids (Newberry and Tallahassee, FL), and the extent of existing voids (US 441 and Gainesville, FL). As with the synthetic study, field seismic results from test lines one diameter from the void did show void distortion, but it was not present further away.</p> <p>Finally, the software was demonstrated and used by FDOT field personnel on a test site. The seismic equipment used for this project included both a 24-channel coupled geophone array with sledge hammer point source, a new 24-channel land-streamer, and a propelled energy generator (PEG). The land streamer and PEG may be towed behind a vehicle and used to investigate the subsurface profiles beneath a roadway (e.g., US 441).</p>			
17. Key Words <b>Full Waveform Inversion, Sinkholes, Anomalies, Shear Wave, Compression Wave, Geophysical, Non-Destructive Testing, Inversion Analysis</b>		18. Distribution Statement <b>No restrictions.</b>	
19. Security Classif. (of this report) <b>Unclassified</b>	20. Security Classif. (of this page) <b>Unclassified</b>	21. No. of Pages <b>181</b>	22. Price

**Form DOT F 1700.7 (8-72)**

Reproduction of completed page authorized

## ACKNOWLEDGMENTS

The researchers would like to thank the Florida Department of Transportation (FDOT) for the financial support to carry out this research, and the State Materials Office in Gainesville, Florida, for providing access and assisting with many of the field tests (SPT, CPT, and seismic). This includes traffic control and CPT testing at US 441, assistance with seismic testing in Tallahassee, SPT and CPT testing at Kanapaha, and SPT testing at Newberry.

## EXECUTIVE SUMMARY

The focus of this research was the development of software for identification of soil and rock layering and anomalies (e.g., sinkholes, lenses, etc.) using nondestructive seismic surface wave testing. For the software to be readily adopted for site investigation, the following six tasks had to be accomplished: (1) Develop a “real time” field automatic algorithm for 2-D Full Waveform Inversion (FWI); (2) Code a software user interface that displayed all results visually, and was easily used by field technicians; (3) Investigate the algorithm’s resolution, i.e., detect the size and depth of an anomaly (e.g., void); (4) Identify the 3-D effects on the 2-D FWI analysis (location and distortion of anomaly) and how to overcome; (5) Test the developed software at 5 FDOT test sites; and (6) Provide training of FDOT personnel on software use at test sites.

Real-time independent assessment of shear (S) and compression (P) wave speeds in a 2-D plane (18 x 36 m with ~1,200 individual cells: 0.75 x 0.75 m) was identified in 30 minutes on a field computer. To achieve the latter, the code employed perfectly matched layers (PML) for boundary truncation (i.e., no wave reflection, limit size of domain), use of progressively increasing frequency source data along with virtual source and reciprocity of the wave fields (reduced number of calculations) to estimate gradient for the Gauss-Newton solution strategy (quadratic rate of convergence), and parallelization of the inversion algorithm. Generally, the analysis requires approximately 20 to 30 minutes on an eight-core laptop computer.

The new FWI software was written in C# (screens) and C++(routines, passing data, etc.) and was designed to enter data for setting up the analysis (dimensions of the domain, spacing of shots and geophones, number of shots, receivers, etc.), conditioning the data (filter, window, remove noise from signal, etc.), preprocess the data (identify initial shear wave velocity from top to bottom of domain), and analyze (invert) the filtered data to obtain shear (S) and compression

(P) wave velocities for each cell (domain) and associated error (residual). The results are displayed graphically for all cells as the inversion proceeds (lower frequency to higher frequency, i.e., coarse to fine resolution of domain) along with times for the analyses. It is expected that a line of seismic data will be analyzed (30 mins) while another line of data is collected, allowing for further testing if anomalies are identified.

To identify the size and depth of a void that may be detected by the FWI software (Task 3), synthetic surface wave data were generated for different void diameters and depths using a finite difference solution of 3-D wave equations. Subsequently, the synthetic surface wave data were inverted by the new (FWI) software to obtain the S- and P-wave velocities for each cell. A review of the results revealed that any void could be identified, if its embedment depth was no more than three times the void diameter.

To investigate the influence of void proximity to the analyzed 2-D line, the 3-D finite difference algorithm was again used to generate synthetic surface wave data, with the array of shots and receivers placed  $1/2$  diameter ( $1/2 D$ ) and one diameter ( $1 D$ ) (Task 4) from the voids. Analysis of the synthetic surface data using the FWI software showed that the void was still visible  $1/2 D$  away from the void, and the void however becomes distorted and non-existent if the test line is at least one diameter from the void. In case of an anomaly, it is recommended to test a new line located 1 diameter on each side of the analyzed line for clear identification of the void and its embedment.

Using the new FWI software, investigation of five Florida sites with known and unknown anomalies was undertaken. None of the sites had any prior information on soil and rock layering. In addition, no invasive testing (SPT and CPT) occurred until after anomalies were identified by seismic testing. A review of the invasive test results revealed that the seismic

analysis did an excellent job of identifying soil and rock layering (e.g., limestone depths at Kanapaha and Newberry), identifying unknown voids (Newberry and Tallahassee), and the extent of existing voids (US 441 and Gainesville retention pond). In addition, the field seismic test lines approximately one diameter from a known void were visible on the FWI scans, but were distorted. Testing on each side of the analyzed line improved the location of the void (i.e., depth and diameter).

Finally, training of FDOT personnel for both seismic testing and use of the new FWI software was undertaken on a FDOT-approved site. Based on their use and review, improvements were made with the input and output screens, and the user manual (appendix) of the software was developed.

# TABLE OF CONTENTS

	<u>page</u>
DISCLAIMER .....	ii
SI (MODERN METRIC) CONVERSION FACTORS (from FHWA).....	iii
TECHNICAL REPORT DOCUMENTATION PAGE .....	v
ACKNOWLEDGMENTS .....	vi
EXECUTIVE SUMMARY .....	vii
LIST OF FIGURES .....	xiii
INTRODUCTION .....	1
1.1 Background.....	1
1.2 Objective and Supporting Tasks .....	3
1.2.1 Task 1 – Development of a Fast and Automatic Algorithm of the 2-D Full Waveform Inversion on Synthetic Models .....	3
1.2.2 Task 2 – Develop Interface Software for the 2-D Full Waveform Inversion and Guidelines for Implementation .....	4
1.2.3 Task 3 – Investigation of Seismic Responses of Embedded Voids on Synthetic Models .....	6
1.2.4 Task 4 – Explore 3-D Effects in the 2-D Full Waveform Inversion on Synthetic Models .....	7
1.2.5 Task 5 – Test the Developed FWI Software on Full Scale Tests Sites .....	8
1.2.6 Task 6 – FDOT Training of Software and Preparation of Draft and Final Reports .....	9
DEVELOPMENT OF A FAST AND AUTOMATIC ALGORITHM OF 2-D FULL WAVEFORM INVERSION ON SYNTHETIC MODEL .....	10
2.1 Introduction.....	10
2.2 Development of a Fast FWI Algorithm .....	11
2.2.1 Perfectly Matched Layer (PML) .....	11
2.2.1.1 Background .....	11
2.2.1.2 Methodology and Implementation .....	13
2.2.1.3 Results .....	16
2.2.2 Parallel Computing.....	19
2.2.2.1 Background .....	19
2.2.2.2 Results .....	20
2.2.3 Windowing .....	22
2.2.3.1 Background .....	22
2.2.3.2 Methodology and Implementation .....	23
2.2.3.3 Results .....	24

2.2.4 Grid Reduction .....	25
2.2.4.1 Background .....	25
2.2.4.2 Methodology and Implementation .....	27
2.2.4.3 Results .....	29
2.3 Automation of the FWI Algorithm .....	31
2.3.1 Parameter setup and data conditioning .....	31
2.3.2 Initial Model .....	34
2.3.3 Inversion analysis .....	37
2.3.3.1 Methodology and Implementation .....	37
2.3.3.2 Results .....	40
2.4 Conclusion .....	43
DEVELOP INTERFACE SOFTWARE FOR THE 2-D FULL WAVEFORM INVERSION AND GUIDELINES FOR IMPLEMENTATION .....	44
3.1 Introduction.....	44
3.2 Summary of Software Development and Validation.....	45
3.3 Conclusion .....	45
INVESTIGATION OF SEISMIC RESPONSES OF EMBEDDED VOIDS ON SYNTHETIC MODELS.....	46
4.1 Introduction.....	46
4.2 Methodology.....	46
4.2.1 3-D Forward Modeling.....	49
4.2.2 3-D Wave Propagation Example .....	57
4.2.3 Energy Comparison .....	61
4.2.4 2-D inversion results of 3-D waveform data .....	66
4.3 Conclusion .....	68
EXPLORE 3-D EFFECTS IN THE 2-D FULL WAVEFORM INVERSION ON SYNTHETIC MODELS.....	72
5.1 Introduction.....	72
5.2 Methodology.....	73
5.3 Results.....	74
5.3.1 Distortion and Location of Off-center Void .....	80
5.4 Conclusion .....	80
TEST THE DEVELOPED FWI SOFTWARE ON FULL SCALE TEST SITES .....	88
6.1 Introduction.....	88
6.2 US 441 Site .....	89
6.3 Newberry Site .....	94
6.4 Gainesville Site .....	100
6.5 Tallahassee Site .....	104
6.6 Kanapaha Site .....	109
6.7 Conclusions.....	116

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.....	118
7.1 General.....	118
7.2 Development of a Fast and Automatic Algorithm of the 2-D Full Waveform Inversion on Synthetic Models.....	118
7.3 Development of Interface Software for the 2-D Full Waveform Inversion and Guidelines for Implementation .....	119
7.4 FWI’s Detection of Embedded Voids Directly Beneath 2-D Seismic Line .....	120
7.5 Identification of Voids Laterally Spaced from 2-D Seismic Line.....	121
7.6 Evaluating Developed FWI Software with Invasive Tests on Five FDOT Test Sites....	121
7.7 Recommendations for Further 3-D FWI Development.....	122
REFERENCES .....	124
FWI INVERSION SOFTWARE USER MANUAL .....	127

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
Figure 1.1 Snapshot of the FWI software .....	6
Figure 2.1 Wave reverberation .....	11
Figure 2.2 Wave propagation - grid expansion vs. original size .....	12
Figure 2.3 Computational time for grid expansion.....	13
Figure 2.4 Wave originating near boundary .....	14
Figure 2.5 PML layer (Kallivokas et al., 2013) .....	15
Figure 2.6 Parameter setting page.....	15
Figure 2.7 $ax$ vs. location .....	17
Figure 2.8 $bx$ vs. location .....	17
Figure 2.9 Wave image – no PML vs. PML.....	18
Figure 2.10 Wave plot – no PML vs. PML.....	19
Figure 2.11 CPU usage – parfor disabled .....	21
Figure 2.12 CPU usage – parfor enabled .....	22
Figure 2.13 Waveforms with noise.....	23
Figure 2.14 Windowing values .....	24
Figure 2.15 Windowed data.....	25
Figure 2.16 Computation time of grid sizes.....	26
Figure 2.17 Grids used in FWI .....	27
Figure 2.18 Grid reduction example .....	28
Figure 2.19 Inversion page .....	29
Figure 2.20 S-wave and P-wave velocities - no grid reduction .....	30
Figure 2.21 S-wave and P-wave velocities - grid reduced.....	31
Figure 2.22 Parameter page .....	32

Figure 2.23 Data import page .....	33
Figure 2.24 Good data after removing channels 1, 2, and 8 .....	33
Figure 2.25 Spectral imaging page .....	35
Figure 2.26 Linear profiles of several models .....	36
Figure 2.27 Initial model of S-wave velocity (m/s).....	36
Figure 2.28 Initial model of P-wave velocity (m/s).....	37
Figure 2.29 Comparison of observed and estimated data for 1 shot.....	41
Figure 2.30 S-wave and P-wave velocities (m/s) – data at 10 Hz .....	42
Figure 2.31 S-wave and P-wave velocities (m/s) – data at 20 Hz .....	42
Figure 4.1 Dual layer model .....	47
Figure 4.2 Void – 4.5 m deep .....	48
Figure 4.3 Void – 9 m deep .....	48
Figure 4.4 Void – 13.5 m deep .....	49
Figure 4.5 Staggered 3-D grid .....	51
Figure 4.6 $\partial\sigma\partial z$ surface plot.....	55
Figure 4.7 Perfectly matched layer added on to original domain .....	56
Figure 4.8 3-D wave propagation with and without the PML .....	58
Figure 4.9 3-D synthetic waveform data – shot located at center of the line 1 .....	59
Figure 4.10 3-D synthetic waveform data – shot located at end of the line 1 .....	60
Figure 4.11 2-D wavefield – shot centered in line 1 – line 1 vs. line 3 .....	60
Figure 4.12 2-D wavefield – shot at end of line 1 – line 1 vs. line 3.....	61
Figure 4.13 Vertical velocity at the receivers .....	62
Figure 4.14 Receiver comparison .....	62
Figure 4.15 Riemann sum setup.....	63
Figure 4.16 Riemann sum with magnitude of velocities .....	64

Figure 4.17 Ricker Wavelet .....	65
Figure 4.18 Percent change in energy for each void.....	66
Figure 4.19 Frequency range for the first inversion run .....	67
Figure 4.20 Frequency range for the second inversion run .....	67
Figure 4.21 Synthetic model of S-wave and P-wave velocities (m/s): (a) 4.5-m depth true mode; (b) initial model; (c) inverted model at 12 Hz; (d) inverted model at 22 Hz.....	69
Figure 4.22 Synthetic model of S-wave and P-wave velocities (m/s): (a) 9-m depth true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 22 Hz.....	70
Figure 4.23 Synthetic model of S-wave and P-wave velocities (m/s): (a) 13.5-m depth true model; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 22 Hz.....	71
Figure 5.1 Line adjacent to void .....	72
Figure 5.2 Vs and Vp in the Y-direction.....	73
Figure 5.3 3-D Vs model .....	75
Figure 5.4 3-D Vp model .....	75
Figure 5.5 Ricker Wavelet .....	76
Figure 5.6 Wavefield created from receiver line over a void .....	77
Figure 5.7 Frequency range for first iteration set .....	78
Figure 5.8 Frequency range for second iteration set.....	79
Figure 5.9 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 1 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz.....	82
Figure 5.10 Line 1 observed, estimated, and residual waveforms: (a) Shot 1 and (b) Shot 15 .....	83
Figure 5.11 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 2 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz. ....	84
Figure 5.12 Line 2 observed, estimated, and residual waveforms: (a) shot 1 and (b) shot 15 .....	85
Figure 5.13 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 3 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz. ....	86
Figure 5.14 Line 3 observed, estimated, and residual waveforms: (a) shot 1 and (b) shot 15 .....	87

Figure 6.1 US 441 Site: (a) land-streamer and propelled energy generator source, b) repaired sinkhole location .....	90
Figure 6.2 US 441 Site, line 1: comparison between observed and estimated data for shots at 0 and 16.5 m (0 ft and 55 ft) .....	92
Figure 6.3 US 441 site: S-wave and P-wave velocities: (a) line 1 on shoulder, (b) line 2 at middle of lane, and (c) CPT sounding .....	93
Figure 6.4 Newberry site: (a) photograph of the retention pond and (b) site survey map.....	97
Figure 6.5 Newberry, line Q: comparison between observed and estimated data for shots 1 and 12.....	98
Figure 6.6 Newberry site: S-wave and P-wave velocities: (a) Line P, (b) Line Q, (c) Line R, and (d) SPT blow counts.....	99
Figure 6.7 Gainesville site: (a) Test configuration, (b) Geophone array, and (c) Propelled energy generator source. ....	102
Figure 6.8 Gainesville site: comparison between observed and estimated data for the first and last shots. ....	103
Figure 6.9 Gainesville site: (a) S-wave and P-wave velocities and (b) SPT blow counts .....	104
Figure 6.10 Tallahassee site: test location and configuration .....	106
Figure 6.11 Tallahassee site: comparison between observed and estimated data for the middle and last shots .....	107
Figure 6.12 Tallahassee site: S-wave and P-wave velocities.....	108
Figure 6.13 Tallahassee site: elevations of three incident piles (32, 33 and 34). ....	109
Figure 6.14 Kanapaha site: test configuration .....	111
Figure 6.15 Kanapaha site, L1-EW: comparison between observed and estimated data for shots at 0 and 36 m (0 ft to 120 ft).....	112
Figure 6.16 Kanapaha site: S-wave and P-wave velocities (m/s): (a) to (j) are lines L1-EW to L10-EW, respectively. ....	113
Figure 6.17 Kanapaha site: 3-D view of S-wave and P-wave velocities (m/s): L1-EW to L5-EW. ....	115
Figure 6.18 Kanapaha site: 3-D view of S-wave and P-wave velocities (m/s): L6-EW to L10-EW.....	115
Figure 6.19 Kanapaha site: sample CPT sounding: (a) location and (b) CPT tip resistance. ....	116



# CHAPTER 1 INTRODUCTION

## 1.1 Background

Soil and rock layering variability, and presence of anomalies, voids, etc., is of major concern for both infrastructure (roads and bridges) and commercial and residential buildings. For instance, assessment of commercial and residential losses in Florida are given by the Florida Senate Committee on Banking and Insurance (Dec. 2010) report on sinkholes: “*Total property claims in Florida have increased from 2,360 in 2006 to 6,694 in 2010 for a total of 24,671 claims throughout that period. Total sinkhole property claims costs amounted to \$ 1.4 billion during this time.*” Pre-construction detection of sinkholes could have eliminated substantial portion of this cost, and assisted in remediation.

Since invasive testing, such as standard penetration test (SPT) and cone penetration test (CPT), identifies only small volumes of soil and rock (0.1%), soil/rock layering, anomalies (e.g., sinkhole) on a site, a site subsurface investigation should begin with a non-destructive testing (NDT). Then, having identified general profile layering and anomalies, invasive tests (CPT, SPT) could be conducted to obtain more detailed information. Several NDT approaches have been used to identify sinkholes. They include gravity, electrical resistivity (ER), ground penetrating radar (GPR), electromagnetic wave methods, and traditional seismic wave methods. However, all of these methods have limitations in identifying and quantifying variability (i.e., change due to water, material type, contaminant, or void).

Recently, work performed under the Florida Department of Transportation (FDOT) Research Project, BDK75-977-66, *Detection of Sinkholes or Anomalies Using Full Seismic Wave Fields*, established a new full waveform inversion (FWI) technique to improve the practice of sinkhole detection. The technique has been successfully applied to several synthetic and real

data sets for detection of embedded sinkholes/anomalies. Results from synthetic data sets show the capability of the FWI in characterization of embedded voids (air or water filled voids). Results from real data sets revealed that FWI does a good job of characterizing various site conditions which include an embedded concrete culvert, low-velocity anomalies, open chimneys, and naturally occurring embedded voids.

For the FWI to be a practical tool, a fast, graphical, software program needs to be developed for a laptop for field use. The software is expected to automatically run on raw field data, so that technicians can use it without significant training. The proposed software needs to complete an analysis (e.g., 120 ft line) in a reasonable time period (30 minutes) in order that the test lines can be used to decide locations for additional test lines (e.g., parallel, orthogonal, etc.) to ensure that as much information of layering, voids/anomalies, etc. in the field is recovered on one site visit. This also reduces unnecessary field testing and data processing efforts.

To achieve a “real-time” and automatic solution, a number of challenges associated with FWI inversion must be overcome. Specifically, the FWI is computationally expensive as it requires solving elastic wave equations thousands of times. The FWI code from the Phase I study (BDK75 977-66) took about 2 hours to analyze one test lines (24 phones, 25 shots) on a standard computer. To reduce the solution of about 30 minutes (the same as time to collect data for a line), advanced boundary conditions and solution convergence methods (e.g., gradient techniques) need to be implemented. In addition, variable grid dimensioning, temporal windowing, and parallel computing need to be implemented.

Finally, Full Waveform Inversion (FWI) employing 2-D data collection/analysis must be validated, i.e., location of soil/rock layer boundaries, size, depth and location of anomalies (e.g.,

sinkhole). This will require both analytical/numerical and field work. In the case of the field work, invasive testing (e.g., SPT and CPT) should be used to validate the FWI results.

## **1.2 Objective and Supporting Tasks**

The primary objective of this proposed work was to develop standalone software for 2-D FWI analysis of seismic waves for subsurface investigation (sinkhole detection, layering, top of rock, etc.). It was required that the software functions include: (1) importing geophone and measured shot information from a seismic data acquisition system (e.g., vibroscope), (2) filtering of raw data, (3) performing a 1-D scan to obtain initial velocity profiles, and (4) performing successive FWI analyses (higher frequencies) to obtain 2-D (S-wave, P-wave, and Poisson ratio) profiles of a site

Equally important, the work had to include substantiation of FWI's location (horizontally and vertical) of soil/rock layering and anomalies. In the case of anomalies (e.g., sinkhole), of interest is the size, depth and location of the void that may be detected. This work had to be accomplished both numerically, and in the field. In the case of the field testing, the FDOT identified multiple sites with known and unknown layering, anomalies, etc. for seismic testing. The FDOT (SMO) performed all invasive testing (SPT & CPT) at the sites.

### **1.2.1 Task 1 – Development of a Fast and Automatic Algorithm of the 2-D Full Waveform Inversion on Synthetic Models**

The goal of this task was to develop a fast and automatic algorithm of the FWI, which could be used to obtain information of subsurface profiles during field testing. For the testing/validation, synthetic data (e.g., surface velocity profiles) from numerical solutions (i.e., forward modeling) of embedded voids, representative of natural cavities in 2-D layered systems needed to be developed. The synthetic surface velocity data was then run through the FWI inversion software to predict wave velocities (S and P) on a cell by cell basis in the 2-D profile.

A comparison of the predicted (FWI) cell velocities was compared to the original synthetic values to validate the FWI algorithm. The focus of this task was to improve the accuracy of inverted results and reduce computer time for the real-time analysis of the FWI code.

In terms of the accuracy, to avoid local solutions of the deterministic inversion (Gauss-Newton method), appropriate initial velocity profiles (cell information to start the FWI) needed to be addressed in this proposed work. The work proposed to investigate the effectiveness of two approaches: 1) use of low frequency data on 1-D linear initial model, 2) use of global optimizations such as genetic algorithm and simulated annealing (Tran and Hiltunen, 2012a and b) to generate initial models.

In order to speed up the inversion, the following methods were employed :1) convolutional perfectly matched layers (Komatitsch and Martin, 2007; Giroux, 2012) for boundary truncation to reduce the sizes of the medium, and 2) an efficient technique (Sheen et al., 2006) using virtual sources and reciprocity of wave fields to calculate the gradient matrix. Also, different sized grids for the inversion, temporal windowing, and parallelizing computations were investigated. As the typical time of collecting seismic data is about 30 minutes for one line (24 geophones, 25 sledgehammer shots), the algorithm's full waveform inversion should be automatically completed in less than 30 minutes on a laptop computer for each test line. When successful, subsurface information from previous test lines could be used to determine locations of additional test lines to reduce the testing and data processing efforts.

### **1.2.2 Task 2 – Develop Interface Software for the 2-D Full Waveform Inversion and Guidelines for Implementation**

This task was to develop the user interface software and accompanying manuals for FWI use. The effort was to be particularly focused on developing a graphical user interface (GUI) that was user friendly, permitting seamless functionality between data input, analyses and output

of results. In addition, the GUI was developed for technician-level personnel and not a PhD level operator. For instance, a user was to have the ability to upload the raw data (from all shots and all receivers) from a hard drive, enter the spacing and number of geophones, the hammer locations, the strike sequence, set the parameters for the FWI analysis, and view the results. Figure 1.1 is an example mockup of the GUI for data input and conditioning, in which conditioned data is shown for one shot and one row of receivers. As with the new software, users had the ability to change conditioning parameters (i.e., filtering, windowing, etc.), remove bad channels (geophones) of recorded data, and check the quality of conditioned data for individual shots on the GUI before they are used in the analyses (or possibly obtain new data – i.e., “real time”). In addition, the GUI had to perform the FWI analyses (i.e., multiple frequency ranges, with different medium sizes) developed in Task 1. The software needed to subsequently produce subsurface profiles (P-wave and S-wave velocities), which the GUI was to show both on screen and in print.

The guideline for implementation of the software was documented in details (i.e., step by step) so that users can follow without significant training (Task 6). Included in the guideline was proper deployment and setup of test equipment, and test data analyses. An example of real data analyses was included for visual aids.

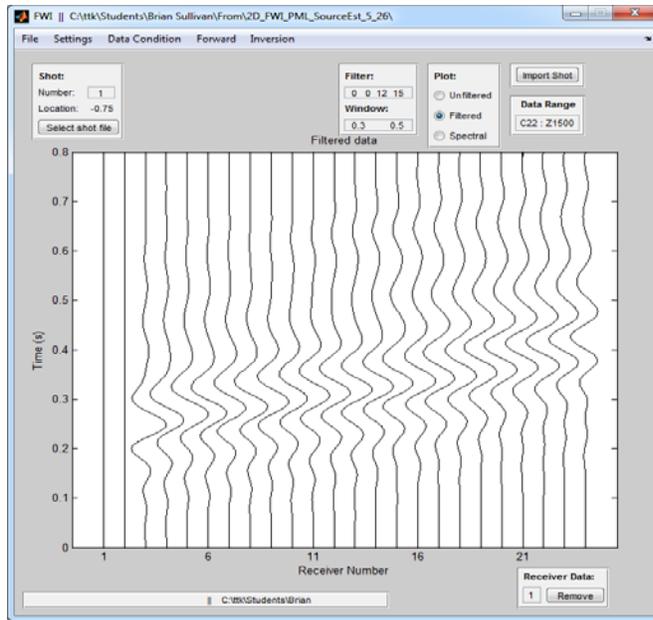


Figure 1.1 Snapshot of the FWI software

### 1.2.3 Task 3 – Investigation of Seismic Responses of Embedded Voids on Synthetic Models

This task was to investigate the sensitivity of the surface wave field in identifying voids (e.g., size, depth) if the source (e.g., hammer) and receivers (geophones) are over the void. The seismic response of an embedded void is a complex problem, as the seismic waves are reflected and refracted when they hit the void, and amount of reflected or refracted energy depends on the size and embedment of the void, material(s) filling inside of the void, and the frequency content of the wave field. A number of 3-D (finite void) models with and without embedded voids were to be used to generate surface wave fields using a finite-difference solution of 3-D wave equations. The wave fields with voids were to be compared against those without voids to investigate the changes of magnitudes and phases of the surface wave fields with respect to the voids. The change (sensitivity) of wave fields was to be examined as a function of an aspect ratio (depth of embedment/ size of void), the dominant frequency of the wave fields, and materials (air or water) filling the void. This study was to provide general ideas of what is the maximum aspect ratio (void size, depth, and wave field frequency content) in which the void can

possibly be identified, assuming the change of wave fields should be at least larger than ground noises (typically, a few percent of wave field). This study was to assist in interpreting inverted results of real data to determine whether identified anomalies are real or just artifacts (false anomalies) of inversion due to noises.

#### **1.2.4 Task 4 – Explore 3-D Effects in the 2-D Full Waveform Inversion on Synthetic Models**

This task was to investigate any distortion to FWI on identifying voids filled with air or water, if the source (e.g., hammer) and receivers (phones) are on a line adjacent to the void. Another task was to explore how much data is needed for locating the void characterized by the 2-D analysis adjacent to the void. A series of tests were conducted on synthetic models with 3-D voids filled by air or water and embedded at different depths and lateral extent. The models were designed to mimic real subsurface profiles at engineering scales (up to 30 m depth) with deep buried voids (> 10 m depth). Again, a finite-difference solution of 3-D wave equations was used to generate synthetic wave fields.

For a given model with a 3-D embedded void, several linear test lines were conducted uniformly on the surface in the vicinity of the void. Sensors and sources equally spaced were used for each test line. The data from each test line was then inverted individually by the algorithm developed in Task 1. It was expected that the void (zero S-wave velocity) could be seen as an anomaly (non-zero S-wave velocity) if the test line was close enough to the void, not necessary on top of the void. By comparing inverted profiles of individual lines, one can examine the location of each line relative to the void and its size and embedment depth. The algorithm developed in Task 1 may be modified if necessary to achieve the desired/required accuracy and resolution for the inversion process. For real field testing, if an anomaly is found at a test line, locations of other test lines can be determined to develop a clear understanding of the void/anomaly.

### **1.2.5 Task 5 – Test the Developed FWI Software on Full Scale Tests Sites**

The developed software was used at a minimum of 5 test sites with soil/rock layering and with embedded voids or sinkholes. The work was to focus on identifying the embedded voids in real time. As the FDOT has well documented a number of sites with histories of sinkholes (e.g., past activity), with non-destructive tests (GPR) and invasive tests (CPT and SPT), this work was to take advantage of the available information for verification of the developed test system and software on multiple sites in Florida. The test sites were identified as Newberry, Kanapaha, Keystone Heights, US441, and one additional site of the FDOT's choice.

For each site, it was expected that multiple testing lines would be used. Geophone array needed to be long enough (at least 30 m) for separation of P-wave and S-wave groups in the time domain at far field geophones. This separation would help to independently invert S-wave and P-wave velocities, which would assist in not only identifying soils, but voids and possibly the presence of water within the voids. Seismic energy was generated using both a sledgehammer and a propelled energy generator (PEG-40 Kg model) for comparison. The PEG can generate more consistent wave fields at a large frequency range from 5 Hz to 50 Hz than those generated by a sledgehammer, and thus it was expected to produce better inversion results.

All of the field collected data was analyzed using the software in real time. Data processing (filtering and windowing) was applied to the raw measured data before running inversions, and several inversion runs with different medium mesh sizes were used to optimize the computer time and view 2-D output. The characterized results were compared with known information of the sinkhole, or to results of independent invasive tests (SPT, CPT) to assess the capability of the FWI technique.

### **1.2.6 Task 6 – FDOT Training of Software and Preparation of Draft and Final Reports**

At the end of the software development, it was required that the contractors perform up to one week training with FDOT personnel on the use of the seismic system and software.

Following training in field, training on the functionality of the software was required. The training occurred at the State Materials Office in Gainesville, FL. A user manual was prepared with the final report that includes guidelines for implementation of the software, proper deployment and setup of test equipment, test procedure and data collection steps, and test data analyses.

## CHAPTER 2

### DEVELOPMENT OF A FAST AND AUTOMATIC ALGORITHM OF 2-D FULL WAVEFORM INVERSION ON SYNTHETIC MODEL

#### **2.1 Introduction**

The goal of this task is to develop a fast and automatic algorithm of the full waveform inversion (FWI), which will be used to obtain information of subsurface profiles during field testing. The target is to complete analysis for each test line (25 shots, 24 receivers) in a reasonable time period (30 minutes) in order that completed test lines can be used to decide locations of any addition test lines (e.g., parallel, orthogonal) to ensure that as much information of layering, voids/anomalies, etc. in the field is recovered on one site visit. This also reduces unnecessary field testing and data processing efforts. The focus of this task is to reduce computer time for the real-time analysis of the FWI code and reduce manual efforts during the analysis.

In order to speed up the inversion analysis, the following improvements have been implemented 1) convolutional perfectly matched layers for boundary truncation to reduce the sizes of the medium, 2) parallelizing computations, 3) temporal windowing to improve the convergence rate, and 4) various size grids for the inversion to optimize required computations. To automate the FWI analysis for minimum manual efforts, several improvements of the code have been implemented including visualized data conditioning, automated initial model, and automated analysis with estimated source signatures, an optimal step length, and stopping criteria.

## 2.2 Development of a Fast FWI Algorithm

### 2.2.1 Perfectly Matched Layer (PML)

#### 2.2.1.1 Background

Numerical analysis techniques have inherent problems at the boundary. Waves that encounter the boundary of the model without any conditions will rebound back into the medium. An example of these reverberations is shown in Figure 2.1.

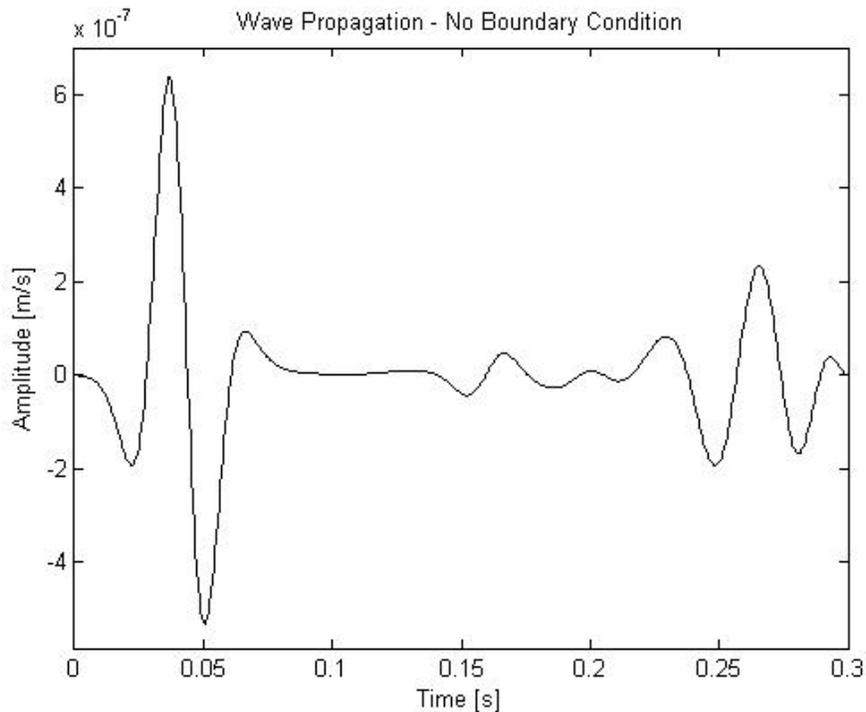


Figure 2.1 Wave reverberation

The wave in this figure travels in a homogeneous medium of a limited size. The main wave is between 0 s and 0.1 s, and waves reflected off the boundary are shown from 0.15 s to 0.3 s. These modeled reverberating waves can negatively impact the inversion (wave matching) because real data is measured in the infinite medium of the earth. To create an accurate model without implementing boundary conditions would mean that the area covered by the grid is large enough, so that either the waves do not reach the boundary or the waves reflected off the boundary do not impede the analysis within the time period being modeled. This is possible but

requires the grid size to be so large to the point where the computational time and amount of memory used is undesirable.

A practical application of this can show the difference in computation time and memory usage. For example, the area being modeled is 22.5 m deep with a length of 37.5 m (roughly 75 x 120 ft). With a vertical and horizontal grid spacing of 0.75 m (2.5 ft), the resulting grid is 30 x 50 in size. There are 25 shots starting at 0 and spaced 1.5 m (4.9 ft) apart and 24 receivers starting at 0.75 m (2.5 ft) spaced 1.5 m (4.9 ft) apart. The amount of time being modeled is 0.5 seconds with a time step of 0.0005 s for a total of 1000 time steps. The medium being used is homogeneous with a shear wave velocity of 200 m/s (656 ft/s), compression wave velocity of 400 m/s (1312 ft/s), Poisson ratio of 0.33, and a density of 1800 kg/m<sup>3</sup> (unit weight of 112 lb/ft<sup>3</sup>). The analysis shows that the grid must be extended by 100 grid points in each direction to provide enough space for the waves to propagate unbounded. This can be seen in Figure 2.2 below comparing the original grid size with a larger one.

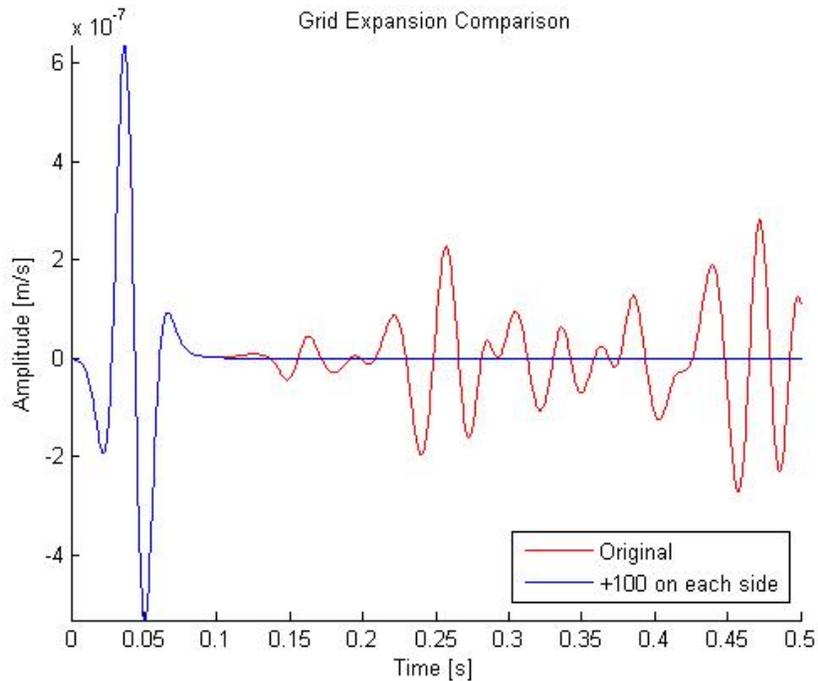


Figure 2.2 Wave propagation - grid expansion vs. original size

The original grid produces wave reflections while the increased grid size prevents the reflections for the given model time. The number of grid points is increased by more than 2000% for the 100 grid point expansion. The affect this has on the computation time can be seen in Figure 2.3.

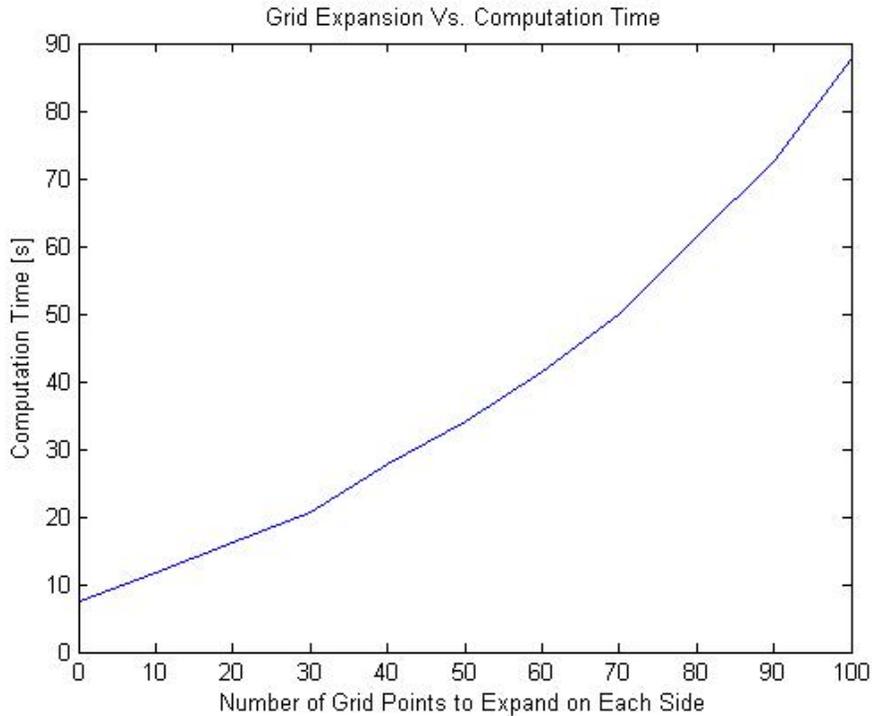


Figure 2.3 Computational time for grid expansion

The original grid has a computation time of 7 s (for 25 shots) while the required grid expansion of 100 points increases the computational time by more than an order of magnitude. This increase is unacceptable due to the numerous times this calculation is used. Grid expansion alone is not a viable option so boundary conditions must be implemented.

### 2.2.1.2 Methodology and Implementation

Increasing the grid size without implementing boundary conditions has been ruled out, so the next step is to implement a boundary condition that absorbs the wave and keeps both the computational time and memory usage to a minimum. There are many methods that have been

developed to act as an unbounded media. Examples include damping layers (Cerjan et al., 1985); (Sochacki et al., 1987), optimized conditions (Peng & Toksoz, 1995), the eigenvalue decomposition method (Donget al., 2005), and continued fraction absorbing conditions (Guddati & Lim, 2005). All of these methods excel in various circumstances but lack the desired wave absorption in several conditions. These can include any waves that impact the boundary at a small angle (Figure 2.4), low frequency waves, and specific wave types such as surface waves (Komatitsch & Martin, 2007).

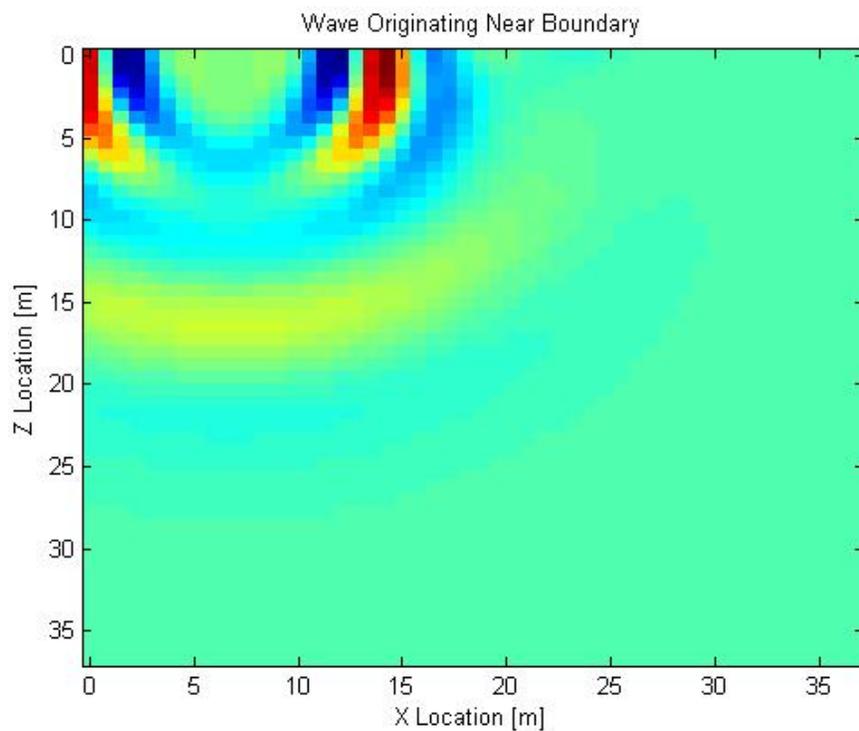


Figure 2.4 Wave originating near boundary

The chosen method for this application is derived from the Perfectly Matched Layer (PML) technique (Komatitsch & Martin, 2007), which successfully absorbs waves from any direction and any frequency. It adds additional grid points at each of the necessary boundaries to help absorb the incoming wave as shown in Figure 2.5. The amount of grid points in the

boundary, also known as the padding, can be adjusted in by using the pad input on the parameter page as seen in Figure 2.6.

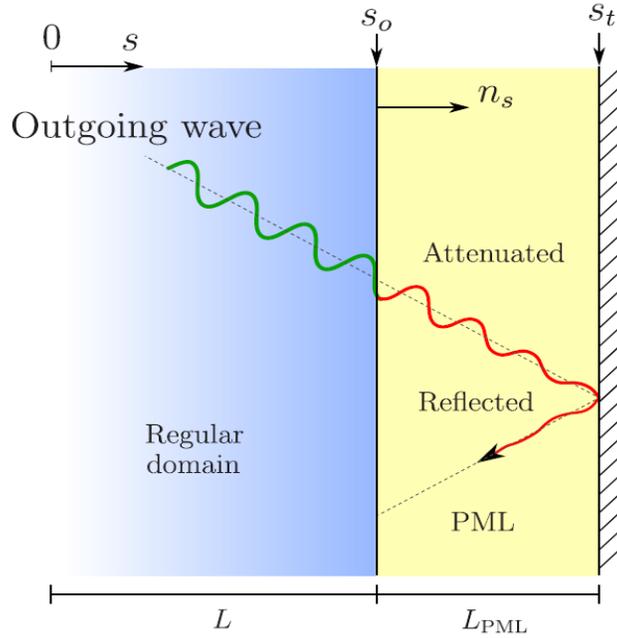


Figure 2.5 PML layer (Kallivokas et al., 2013)

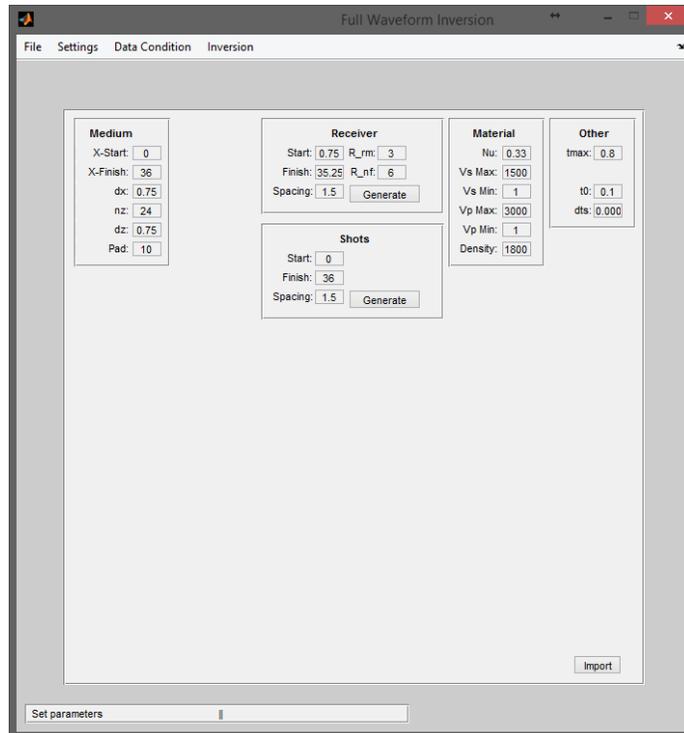


Figure 2.6 Parameter setting page

The PML technique being used is based on the strong form of the anisotropic elastic wave equation in time or domains as shown in Equations 2.1 or 2.2.

$$\rho \partial_t^2 s = \nabla \cdot (c: \nabla s), \quad \text{Eq. 2.1}$$

$$-\rho \omega^2 s = \nabla \cdot (c: \nabla s), \quad \text{Eq. 2.2}$$

where  $\rho$  is the density,  $s$  is the displacement vector,  $\omega$  is the angular frequency, and  $c$  is the elastic tensor (Lame's coefficients). It also uses a damping coefficient,  $d_x$ , which is set to zero inside of the domain and increases parabolically from zero to one in the PML. The final forms of the equations which are used directly in the computational code are represented by Equations 2.3 and 2.4.

$$\psi_x^n = b_x \psi_x^{n-1} + a_x (\partial_x)^{n+1/2}, \quad \text{Eq. 2.3}$$

$$\partial_{\bar{x}} = \frac{1}{\kappa_x} \partial_x + \psi_x, \quad \text{Eq. 2.4}$$

where  $\kappa_x$  is a set parameter,  $n$  represents the current time step,  $\psi_x$  is the memory variable,  $\partial_{\bar{x}}$  is the spatial differential (for both displacement and stress) that is being dampened, and  $a_x$  and  $b_x$  are defined by Equations 2.5 and 2.6 below (Komatitsch & Martin, 2007).

$$b_x = e^{-\left(\frac{d_x}{\kappa_x} + \alpha_x\right) \Delta t}, \quad \text{Eq. 2.5}$$

$$a_x = \frac{d_x}{\kappa_x (d_x + \kappa_x \alpha_x)} (b_x - 1), \quad \text{Eq. 2.6}$$

Using the example provided early,  $a_x$  and  $b_x$  values are presented by Figure 2.7 and Figure 2.8. These arrays are used in Equations 2.3 and 2.4 for every time step to help dampen the waves that encounter the perfectly matched layer.

### 2.2.1.3 Results

Once implemented, the perfectly matched layer provides significant improvements in computational time, memory usage, and model accuracy when compared to the basic grid

expansion technique discussed earlier. A comparison of wave fields is shown in Figure 2.9. The left column of this figure represents the original grid with an expansion of ten grid points to replicate the grid size used in the PML grid. Apparently, significant reflected signals are shown in the left column (No PML) and none in the right column (with PML) after  $t = 0.25$  seconds.

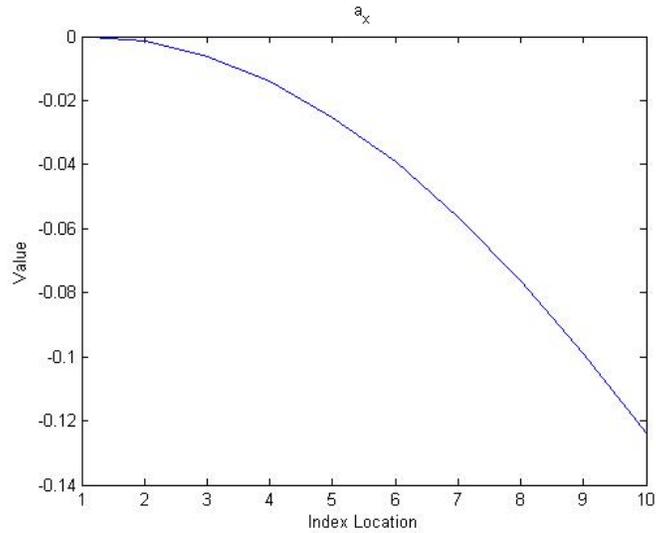


Figure 2.7  $a_x$  vs. location

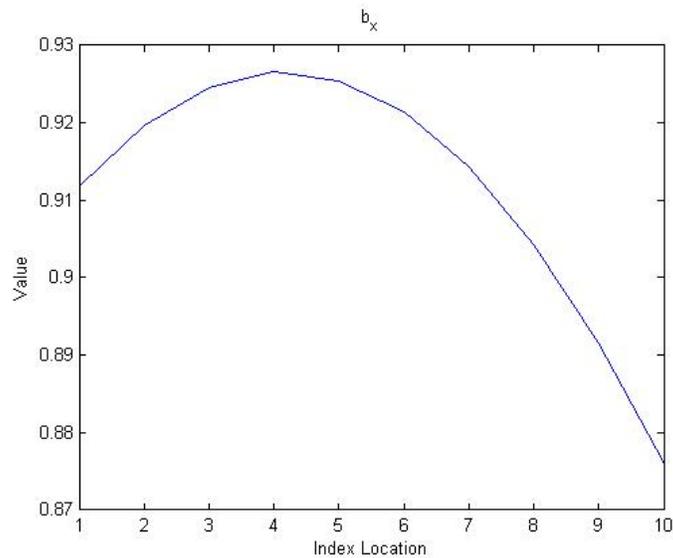


Figure 2.8  $b_x$  vs. location

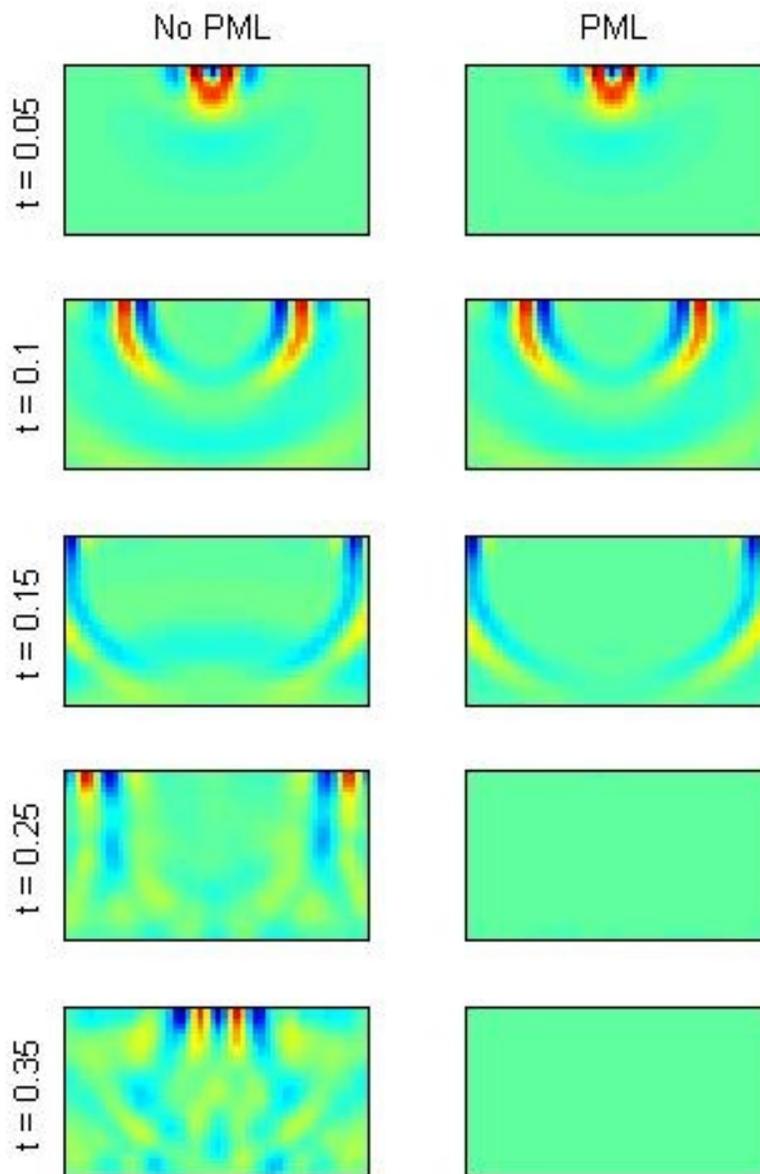


Figure 2.9 Wave image – no PML vs. PML

Figure 2.10 shows the waveforms at the top center of the grid for both the PML and non-PML analysis. The resulting waveform in the PML case has zero reflections within the required model time, resulting in significant improvement over its non-dampening counterpart.

Implementation of the PML does not only reduce the computer time for forward modeling, but

also accelerate the convergence rate of the inversion analysis (fewer iterations) and produce more accurate results. It is expected that about 30% of computer time is reduced.

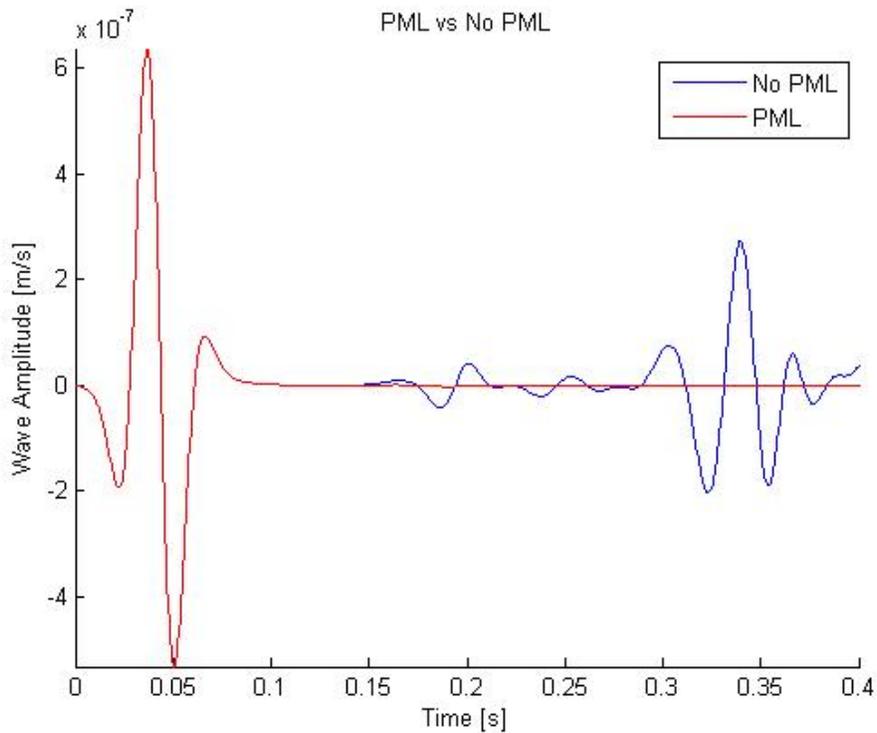


Figure 2.10 Wave plot – no PML vs. PML

## 2.2.2 Parallel Computing

### 2.2.2.1 Background

MATLAB has inherent problems when it comes to distributing the computational workload among all of the cores in the computer's CPU. As an example, seen below, MATLAB runs each loop in succession. This means that a one-minute loop being run eight times will take eight minutes.

```
for it = 1:8  
    ...  
end
```

The use of these types of loops is found throughout the inversion process. The parallel computing toolbox is used to help reduce these computation times.

The parallel computing toolbox decreases computation time by distributing the workload amongst the cores in the CPU (Moler, 2007). The “for” loops are no longer computed one at a time in succession. The option to divide the loop iterations amongst the cores is available by converting the “for” loops into “parfor” (parallel for) loops.

```
parfor it = 1:8
    ...
end
```

This allows for a large reduction in the computational time but can only be used in certain conditions. The most important one is that the loops cannot be based on values that were calculated in other loops. Below is an example of a “for” loop where the nth loop depends on values calculated from the previous loop (n-1).

```
parfor it = 2:4
    A(it) = A(it) + A(it-1);
end
```

This condition is satisfied for several of the “for” loops being used during the inversion process. These include loops over the range of shots and receivers.

### 2.2.2.2 Results

Implementing the “parfor” loops where possible greatly reduces the computational time due to the distributed load. This is clear when comparing the CPU usage before “parfor” implementation (Figure 2.11) to after its implemented (Figure 2.12). In this case, the CPU is the Intel Core™ i7-3770 CPU, which has eight cores. The second plots indicate larger amount of CPU usage than the first plots. This results in a decrease in the computational time of almost 200%.

The use of MATLAB’s parallel computing toolbox greatly decreases the computation time of the inversion process. This is done using the “parfor” loops in place of for loops where

possible. The example given indicates that all of the cores receive an increased workload, thus proving the increase in efficiency.

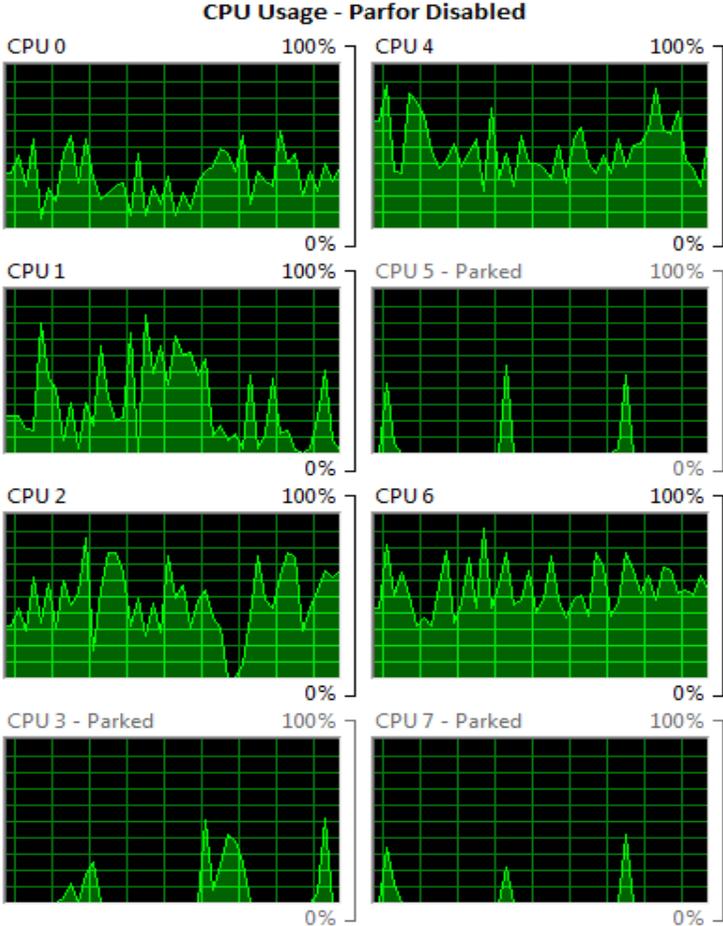


Figure 2.11 CPU usage – parfor disabled

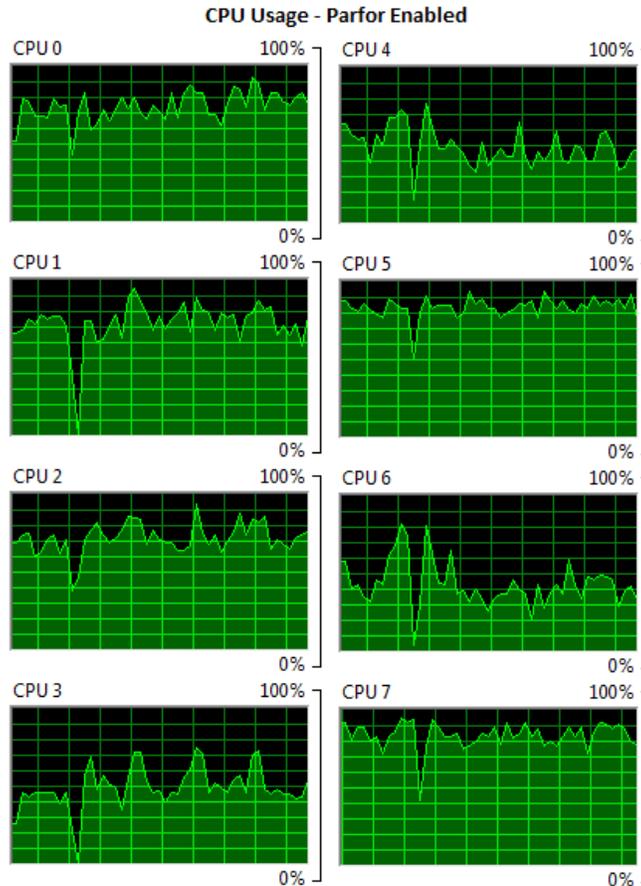


Figure 2.12 CPU usage – parfor enabled

## 2.2.3 Windowing

### 2.2.3.1 Background

The quality of data used in the full waveform inversion process varies based on the environment in which it is gathered. Synthetic data created based on computer models will have significantly less noise (some numerical noise) than data that was gathered in the field. In most cases, the data will be gathered in an environment where random noise will affect the data. This noise can vary from passing vehicles to a nearby construction site. An example can be seen below in Figure 2.13.

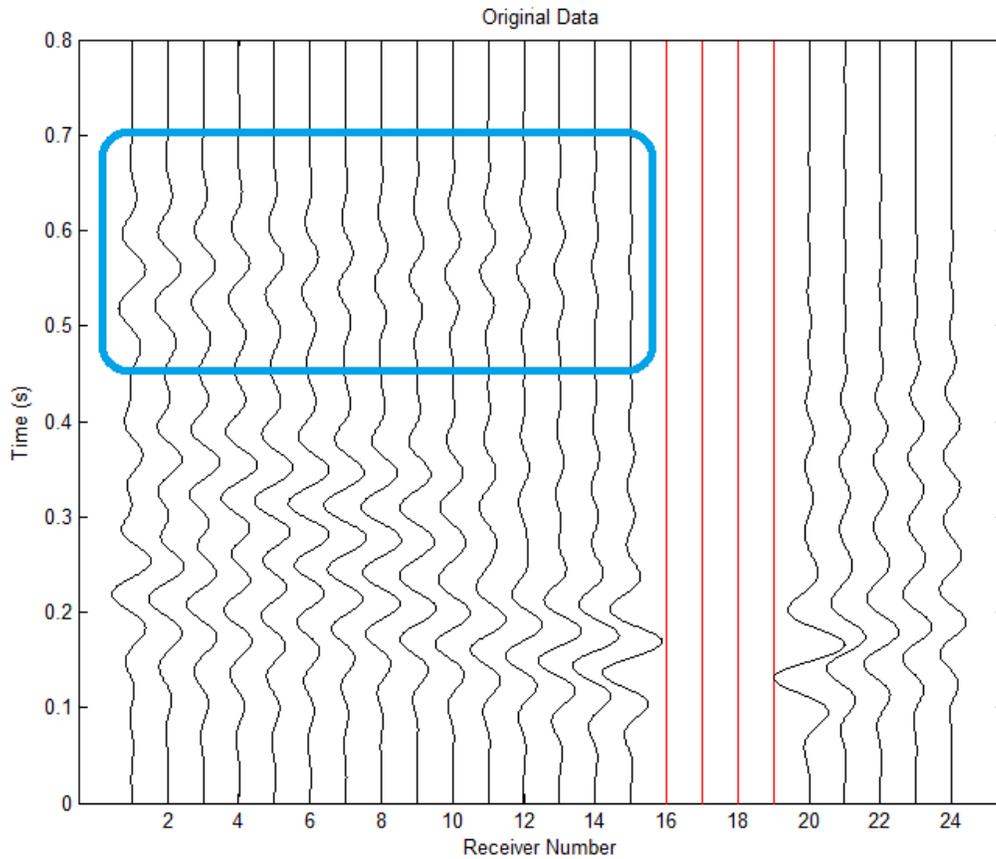


Figure 2.13 Waveforms with noise

The blue box outlines noise introduced to the data from the environment. The noise happens at the same time for several receivers; this pattern is different from that of seismic waves propagating from an active source. In this case, the noise came from traffic on a highway which will impact required computer time for inversion analysis and final results produced from the FWI.

### 2.2.3.2 Methodology and Implementation

The windowing function used in the FWI process reduces the noise individually in each receiver based on the data gathered for that receiver. This is done by finding the time location of the maximum magnitude in each receiver. The data is then set to zero based on two time values given by the user in seconds. The first time value determines how much data is accepted before the peak while the second is for the amount of data is accepted after the peak (Figure 2.14).

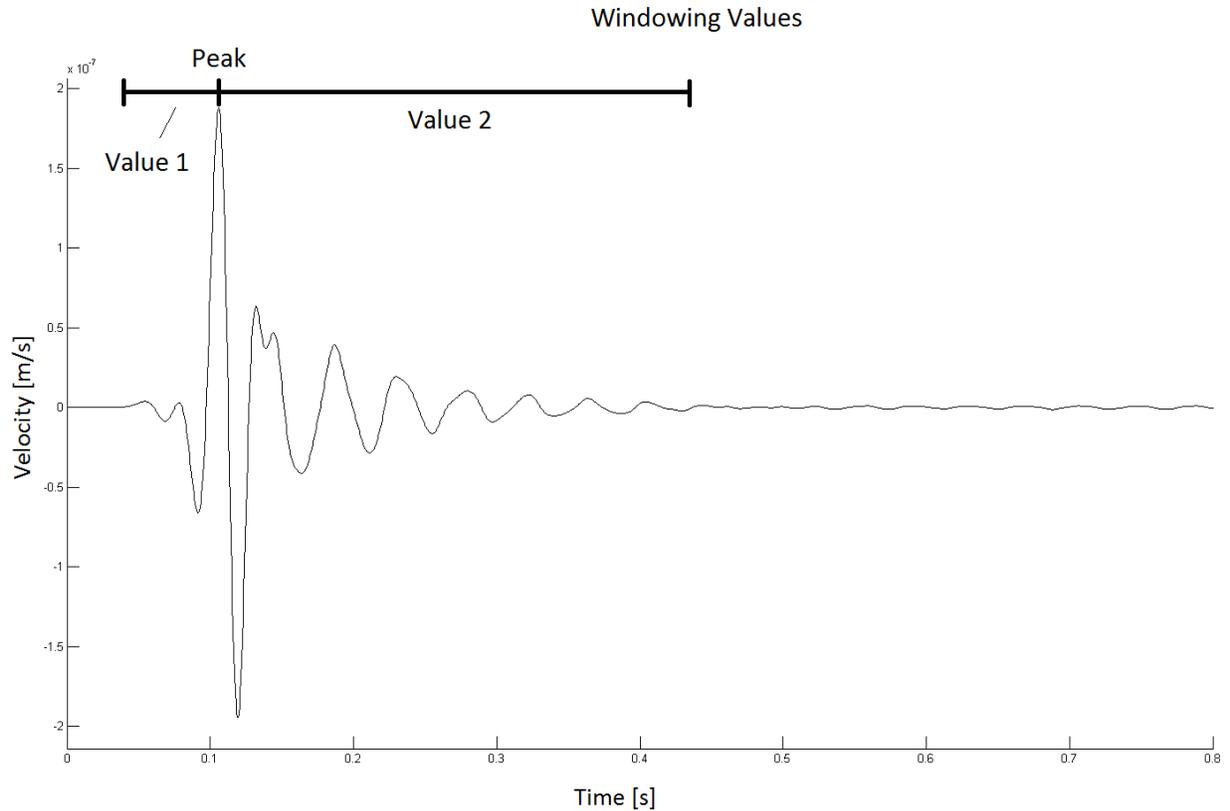


Figure 2.14 Windowing values

Any data outside of these time values will be gradually (tapering) set to zero. The windowing function is applied to the inputted data and the data created from the forward model during the inversion process. This allows for accurate comparison when calculating the residual between the two data sets.

### 2.2.3.3 Results

Shown in Figure 2.15 is the conditioned data by applying the windowing function on the measured data (Figure 2.13). The noise introduced by the environment is removed based on windowing values of 0.2 seconds for each. It was found in many cases that data windowing can improve the convergence rate (less iterations) because it is easier to match synthetic data with noise-reduced real data. It also provides more accurate inversion results.

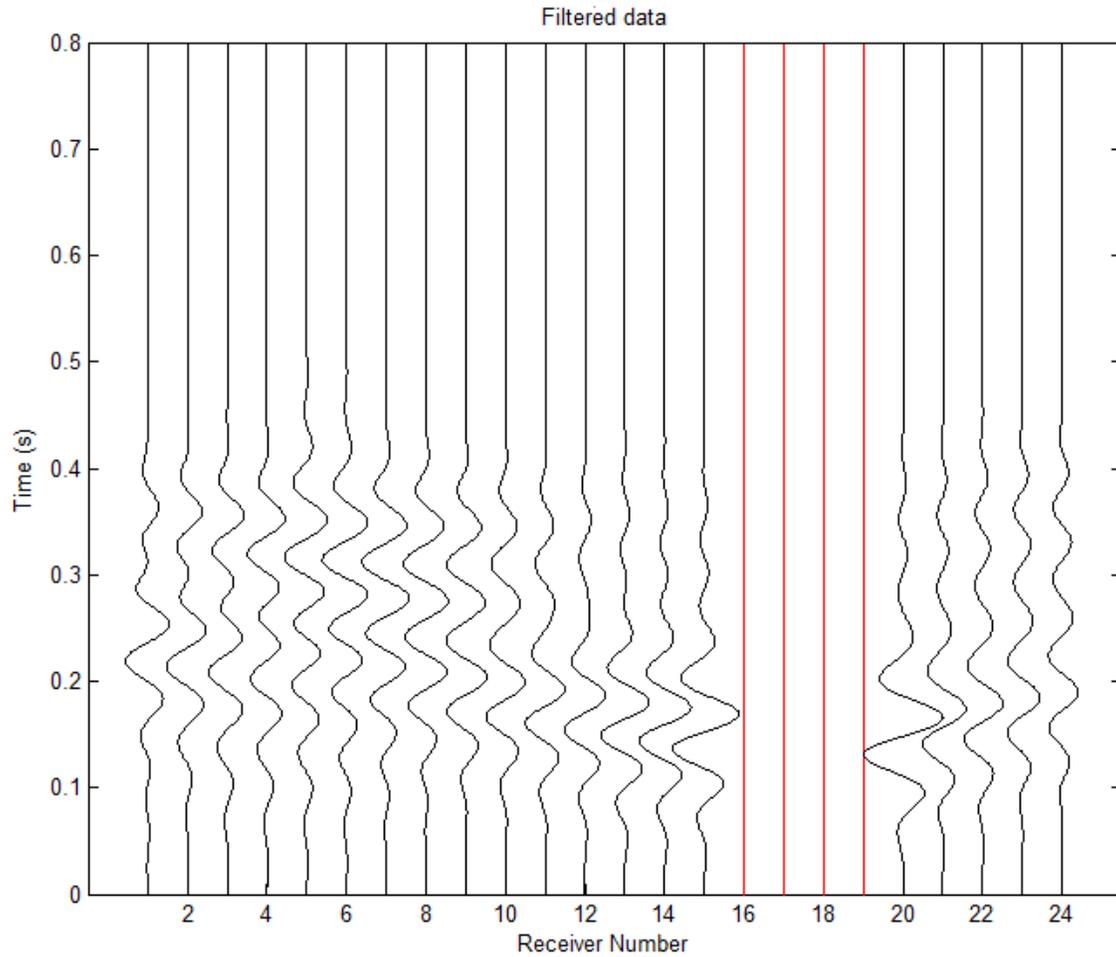


Figure 2.15 Windowed data

## 2.2.4 Grid Reduction

### 2.2.4.1 Background

All finite difference methods require the use of an array (matrix) of values to calculate the desired output. The time and memory necessary to compute these methods are directly related to the size of the array being used. This relationship can be seen in Figure 2.16 where the computational time of the forward elastic equations of a square grid is calculated.

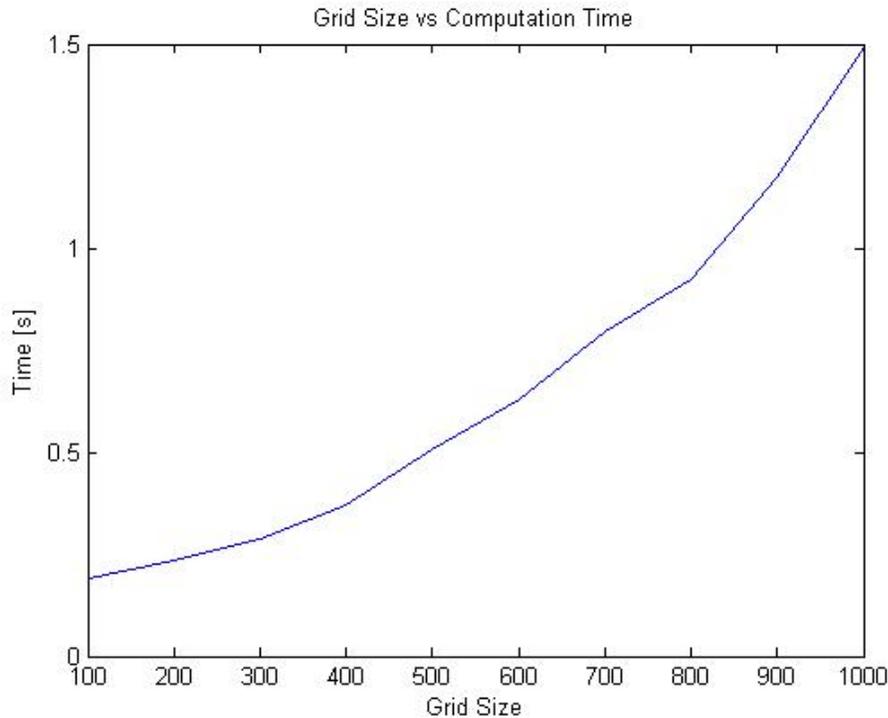


Figure 2.16 Computation time of grid sizes

In the case of the full waveform inversion, there are several grids (Figure 2.17) such as the particle velocities ( $V$ ,  $U$ ), stresses ( $T_{xx}$ ,  $T_{zz}$ ,  $T_{xz}$ ), and material properties including density ( $B1$ ,  $B2$ ) and Lamé's coefficients ( $L+2M$ ,  $M$ ). These are used in various stages of the process including the initial model generation, forward modeling, and inversion.

To reduce computer time, the grid reduction can be used at low frequency analysis runs as one of the following options:

- 1) Using a course mesh (larger grid spacing and fewer grid points) for both forward modelling and inversion. This may allow using only a portion of measured data, as modelled grid points may not be available at the receiver or shot locations.
- 2) Using a fine mesh for forward modelling and a coarse mesh for inversion. For example, a fine grid of  $0.75 \times 0.75$  m ( $2.5 \times 2.5$  ft) is used for forward modeling to

generate synthetic wave fields for all shots and receivers, and a coarse grid of 1.5 x 1.5 m (4.9 x 4.9 ft) is then used for inversion.

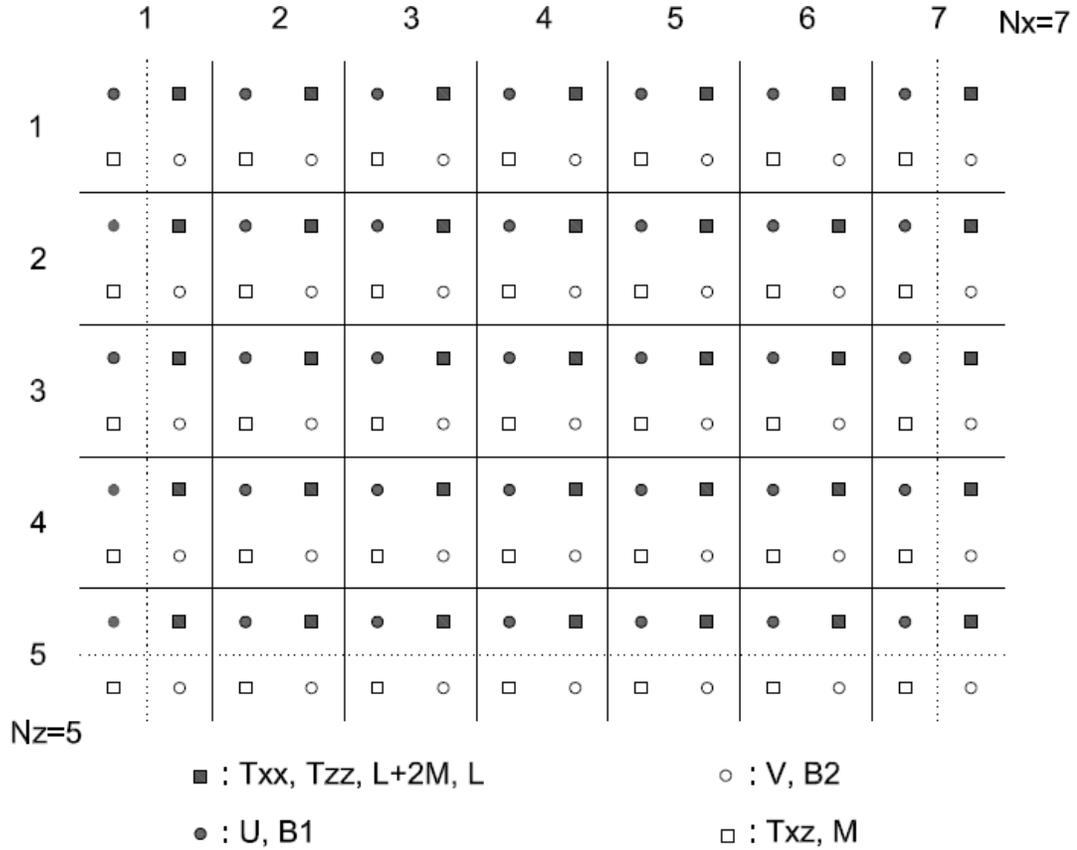


Figure 2.17 Grids used in FWI

### 2.2.4.2 Methodology and Implementation

For option 1, a larger grid spacing and fewer grid points are used, and appropriate receivers and shots for analysis. For option 2, a scaling function is implemented to exchange medium parameters and wave fields between fine and coarse grids. The scaling function allows for both reduction of the grid and restoration to the original size (fine mesh) based on the input parameters as:

$$[\text{output}] = \text{scale\_space}(\text{input}, S_x, S_z, N_z, N_x)$$

The input is the grid to be altered,  $S_x$  and  $S_z$  are the scaling factors in the x and z directions respectively,  $N_z$  and  $N_x$  are the original sizes of the grid, and output is the altered grid. The scaling factors must be greater than one for grid reduction (down scaling). Scaling a grid down in size means that the values for each new grid point are derived from the average of the grid points in that area. Figure 2.18 shows the use of the scale space function in a reduction scenario where  $S_x = 2$  and  $S_z = 3$ .

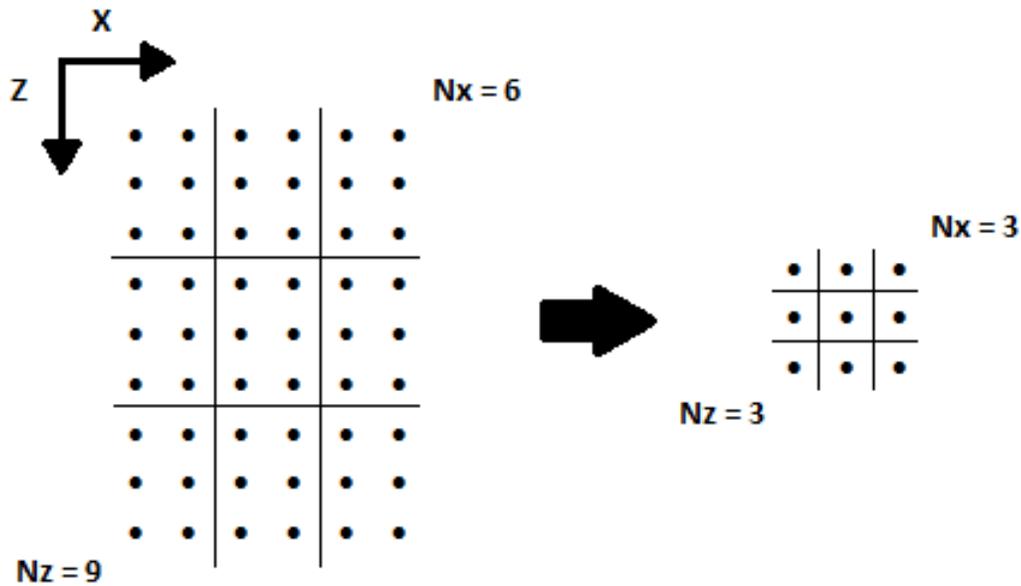


Figure 2.18 Grid reduction example

The segments separated by the lines in the left grid will be averaged to produce the values in the grid on the right. This is done automatically during the inversion process so all the user has to do is input the amount to scale down in the inversion page of the user interface (Task 2) as seen in Figure 2.19. With the input values  $S_x$  and  $S_z$  ( $> 1$  for grid reduction), the scale space function is used during the inversion process to update  $V_s$  and  $V_p$  on the coarse grid. To revert  $V_s$ ,  $V_p$  to the fine grid for forward modeling, the same scaling function is used by inputting the values  $1/2$  and  $1/3$  for  $S_x$  and  $S_z$ . Scaling up in size simply repeats the value for each grid point in the area.

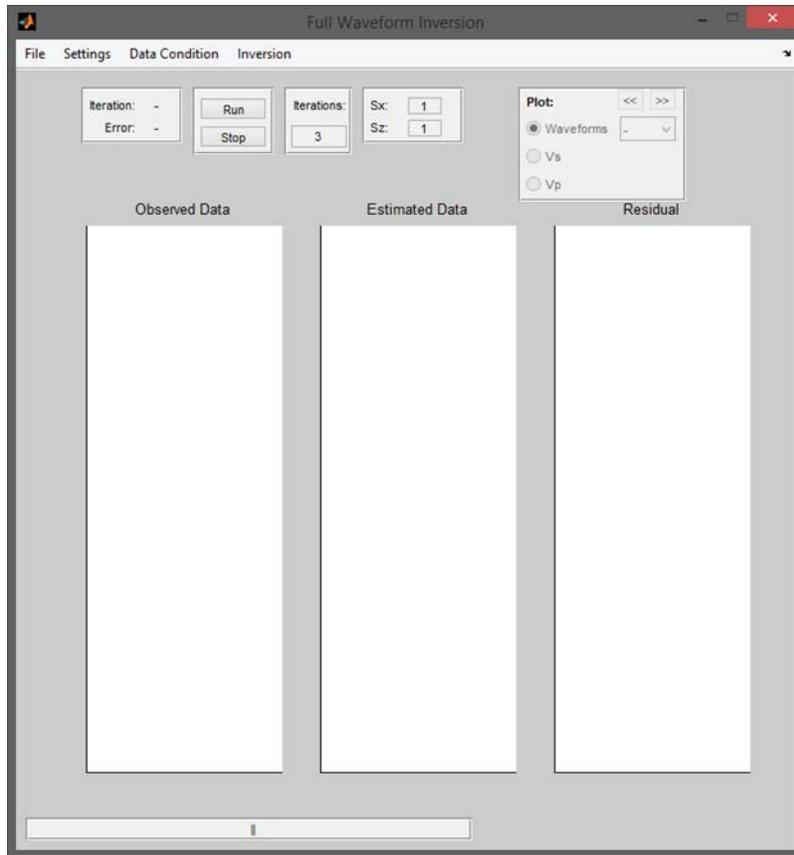


Figure 2.19 Inversion page

### 2.2.4.3 Results

The implementation of the scaling function results in an improvement in computational speed and a reduction in memory usage. Figures 2.20 and 2.21 show inverted results for one test line at the Newberry site with and without grid reduction. Low frequency data (less than 15 Hz with the central frequency of 10 Hz) were used for the analysis. Results in Figure 2.20 were produced by running ten iterations of the inversion with an average computation time of 185 seconds per iteration (about 30 minutes in total) on a standard computer. Implementation of PML, parallel computing, windowing presented above were successfully in achieving relatively quick results. The Vs profile reveals a shallow soil layer underlain by limestone, and an anomaly at distance of 18 m (60 ft).

Results in Figure 2.21 were generated using the same data and parameters, but with a grid reduction by a factor of 2 in each direction. Ten iterations were also performed with the average computational time per iteration of 86 seconds (about 14 minutes in total). This is about a 50% decrease in the time required to run the analysis while still providing the necessary information about the test site such as the soil and limestone layers and the anomaly. Using grid reduction will guarantee a field solution to obtain general information of the medium being tested. More detailed information can be achieved by further analysis at higher frequencies after field testing.

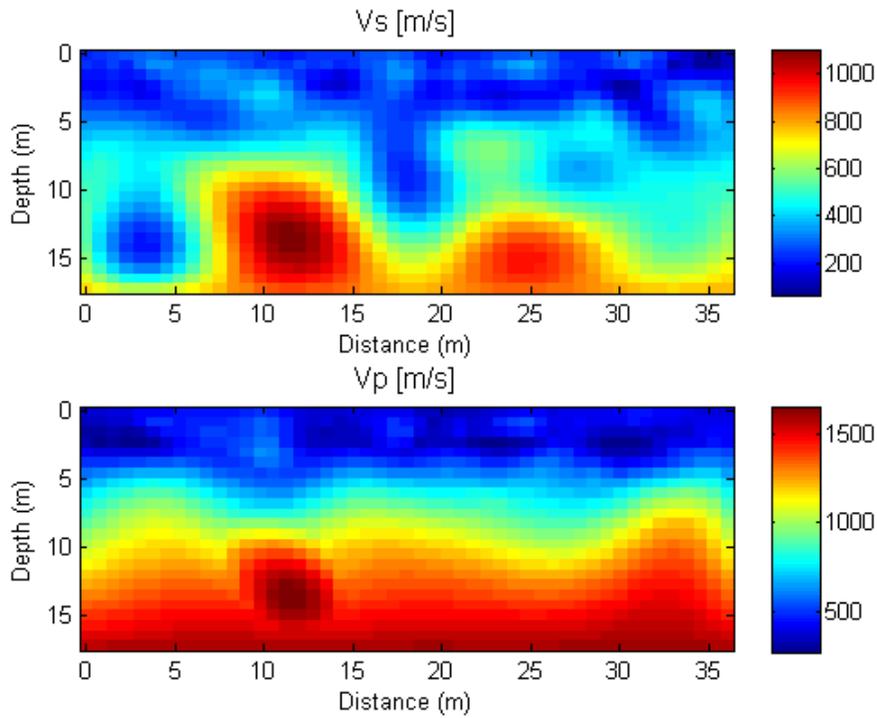


Figure 2.20 S-wave and P-wave velocities - no grid reduction

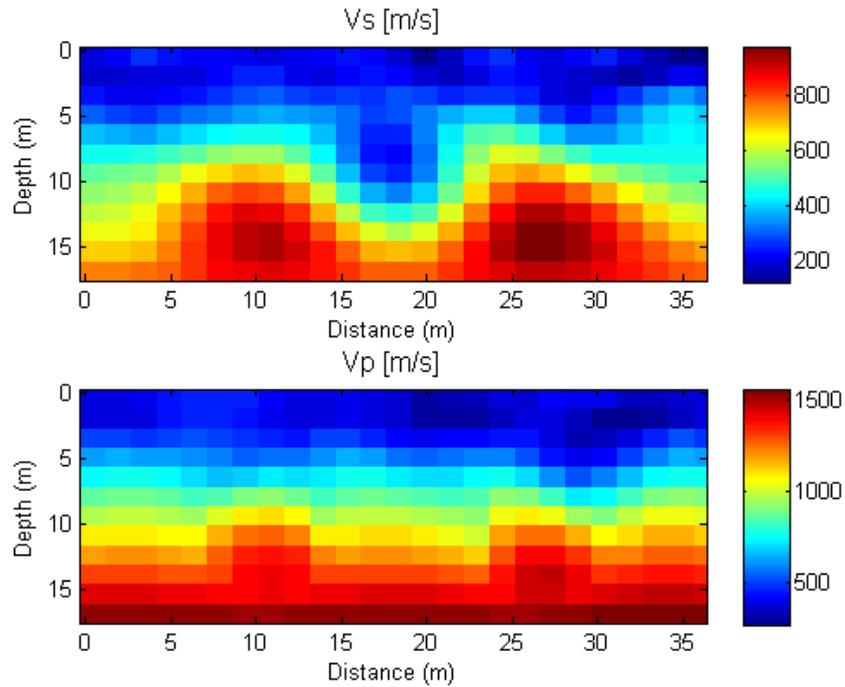


Figure 2.21 S-wave and P-wave velocities - grid reduced

### 2.3 Automation of the FWI Algorithm

Automation of the full waveform inversion program has been implemented through the graphical user interface (GUI), which will be reported in detail in Task 2. It is used herein only for demonstration of the automation process, which results in significant improvement in computation time and accuracy of the results. This is done in several steps starting with parameter setup and data conditioning. Once complete, the initial model is automatically generated using the user's parameters. Lastly, the inversion process is set to run without any input during the process from the user.

#### 2.3.1 Parameter setup and data conditioning

The first step is to input the parameters based on the test configuration. In this case, 24 receivers are used with a spacing of 1.5 m (5 ft). The shots are set between the receivers and at each end of the line for a total of 25. The typical material properties are used for minimum and maximum constraints while the time parameters are based on the data acquisition specifications

(delayed time, recorded time, sampling rate). The parameter page in Figure 2.22 shows the input parameters.

The next step in the process is to input the data by using the data import page (Figure 2.23). Data for all shots and receivers are automatically imported into the program. The data is then checked for any irregularities which are removed based on its impact. The ideal data set is one that has a consistent waveform pattern propagating from a source to all receivers. Shots with bad data are either removed completely (all receivers) or partially (only poor receivers) based on the severity of the abnormality. Figure 2.24 shows an example of removing data. Channels 1 and 2 are removed because of very large magnitudes (reduce near-field effect), and channel 8 is removed because the geophone was not working (almost no signal). The final step of data conditioning is to window and filter with appropriate input parameters such as windowing values (time) and frequency ranges.

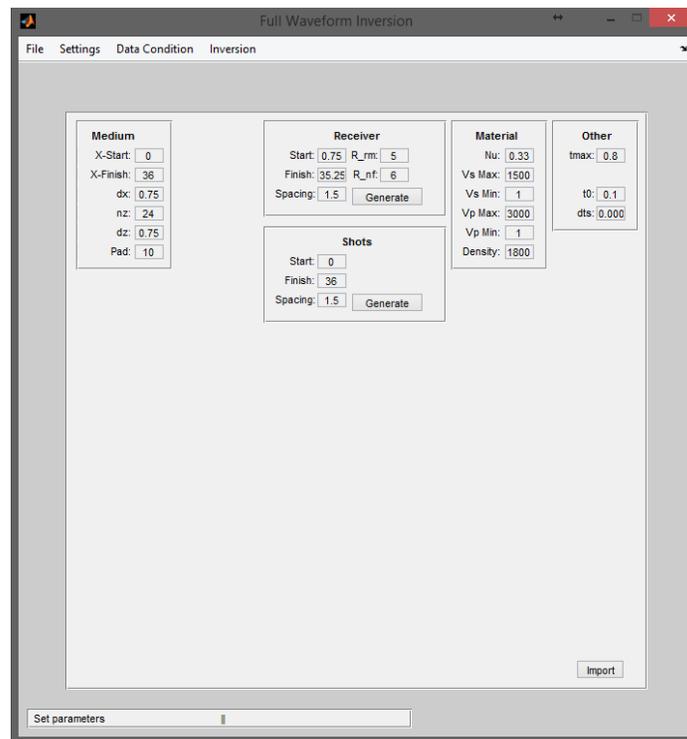


Figure 2.22 Parameter page

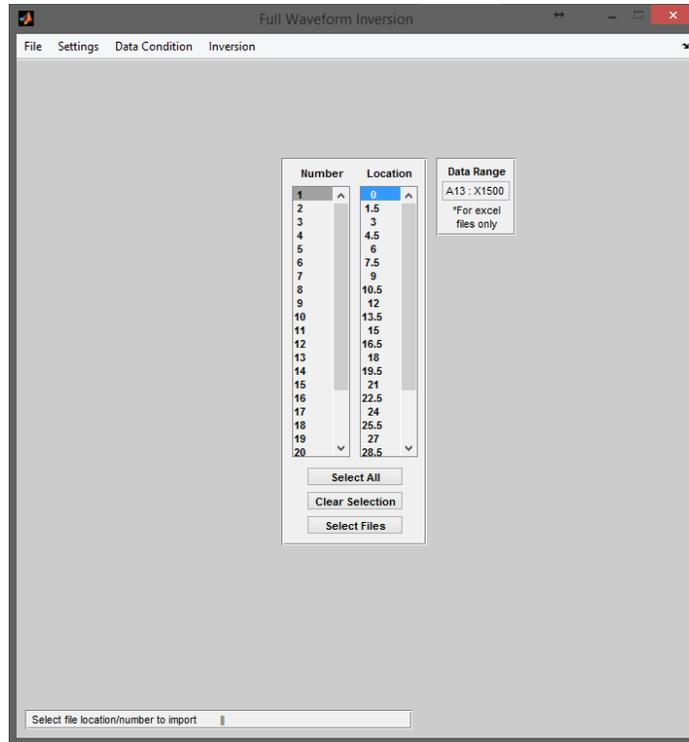


Figure 2.23 Data import page

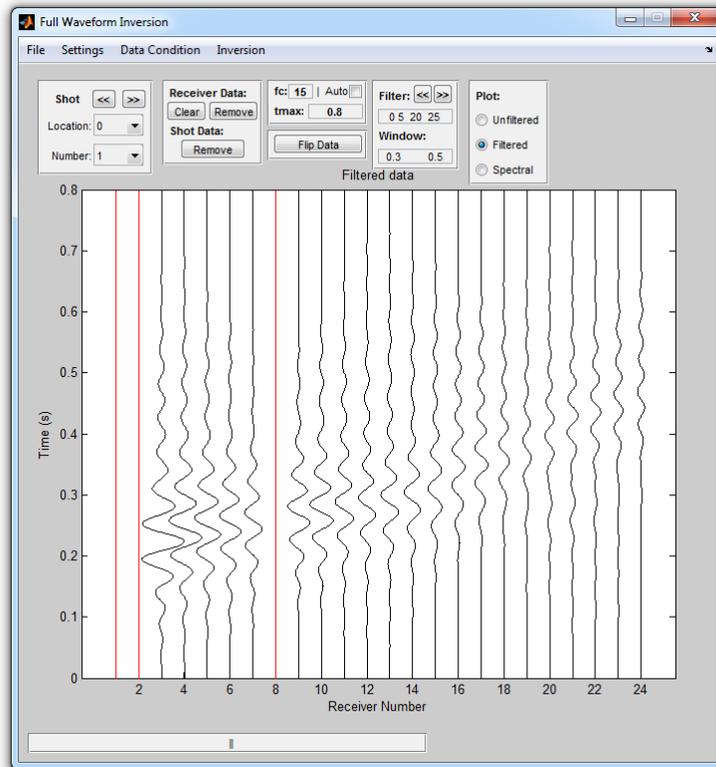


Figure 2.24 Good data after removing channels 1, 2, and 8

The FWI algorithm has been modified to handle the data removal. For example, if a channel is removed from measured data, all calculations related to the removed channel such as estimated data for comparison, gradient, Jacobian and Hessian matrices, and step length are modified accordingly. If a shot is removed, shot numbers and locations, shot order, sizes of data, etc. are also changed accordingly to make sure the waveform analysis is stable and accurate.

### **2.3.2 Initial Model**

For the deterministic FWI, a proper initial model is required to avoid the inversion being trapped in local minima. It must be close enough to the actual test site to allow for the inversion to find the correct solution (global minimum) while keeping the computational time required to an acceptable level. The initial model could be generated by using global inversion techniques, such as genetic algorithm or simulated annealing (Tran and Hiltunen, 2012a and b) on full waveforms. This approach likely produces a global solution but requires significant computer time. From initial study of synthetic models, (FDOT BDK-75-977-66) it has been found that if low frequency components are available, a 1-D linear initial model is usually good enough for inversion of a 2-D profile at engineering scales (less than 30 m (98 ft) depth) (Tran and McVay 2012). For simplicity and a quick solution, an estimate of the initial model is established via a spectral analysis of the measured data.

As an example, Figure 2.25 shows a normalized power spectrum obtained using the cylindrical beam-former technique for the measured data (Figure 2.24). Rayleigh wave velocity ( $V_R$ ) at high frequencies (40 to 50 Hz) is 200 m/s (656 ft/s), which is associated with the top layer; thus the S-wave velocity (slightly larger than  $V_R$ ) of the top layer is known. Rayleigh wave velocity ( $V_R$ ) at low frequencies (5 to 15 Hz) varies from 300 to 1000 m/s (984 to 3280 ft/s), which is the average velocity from the ground surface down to a depth of about one wavelength (including many layers). Thus S-wave velocity of the bottom layer is unknown.

Several 1-D linear models are searched. These S- wave velocity profiles are generated based on the known value at the top (200 m/s) and an assumed range (300 to 1000 m/s) at the bottom as shown in Figure 2.26. These models are then used to compute synthetic data and least squares errors. The model with the smallest error is used as the initial model. This is done automatically during inversion; the user only needs to input the top and bottom range of the S- wave velocity. The initial model is generated as shown in Figures 2.27 and 2.28 for S- wave and P-wave velocity, respectively.

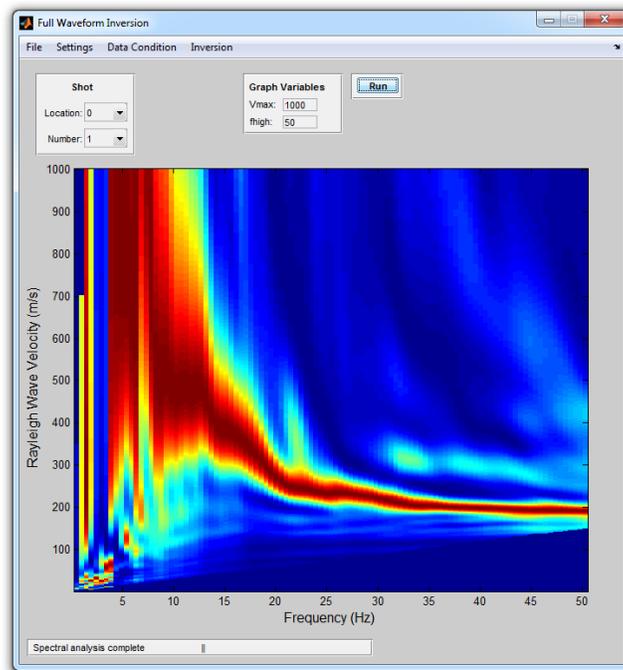


Figure 2.25 Spectral imaging page

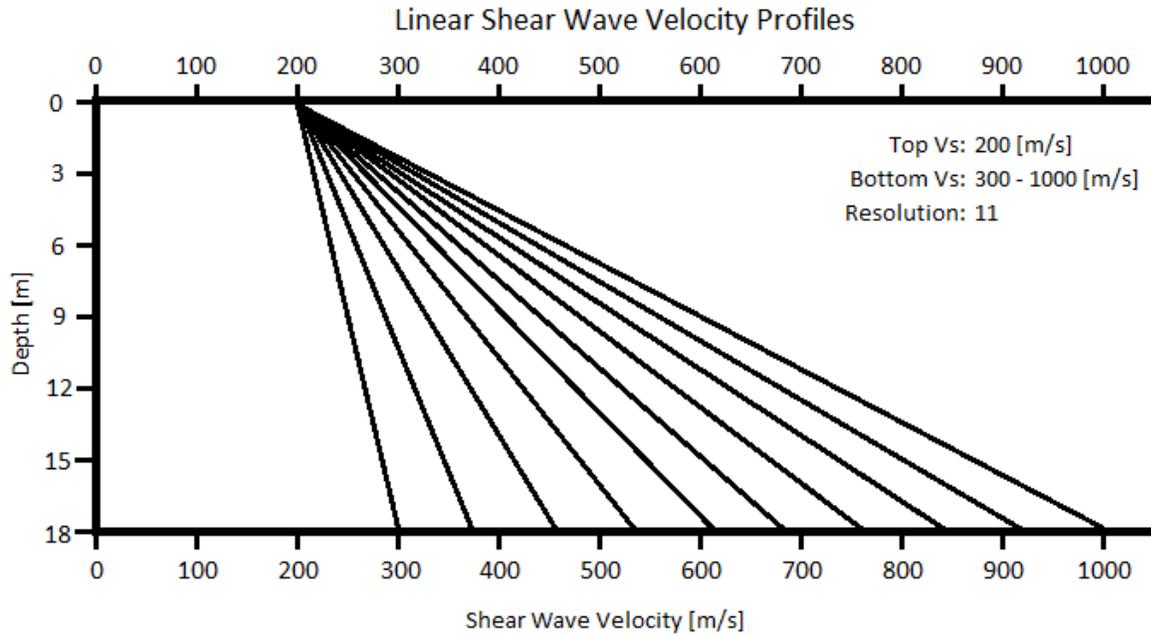


Figure 2.26 Linear profiles of several models

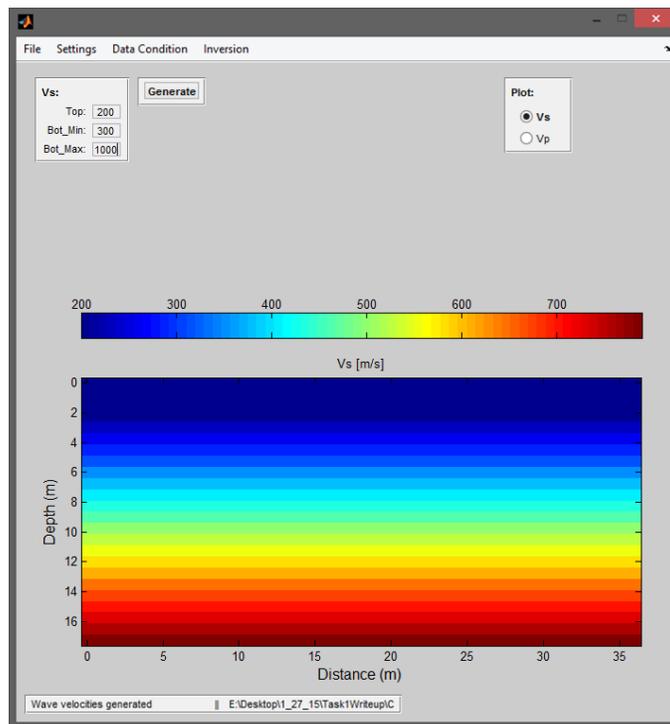


Figure 2.27 Initial model of S-wave velocity (m/s)

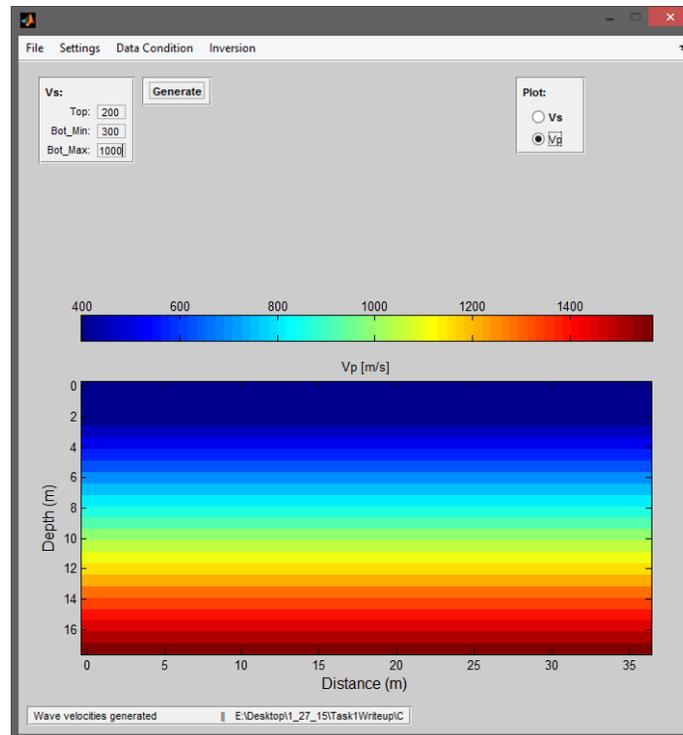


Figure 2.28 Initial model of P-wave velocity (m/s)

## 2.3.3 Inversion analysis

### 2.3.3.1 Methodology and Implementation

The full waveform inversion (FWI) technique (Tran et al., 2013, and Tran and McVay, 2012) has been developed in Phase I. It is a source-independent FWI, with observed and estimated wave fields that were convolved with appropriate reference traces to remove the influence of source signatures. However, the approach requires a manually careful selection of high-quality measured data near source locations for convolution; and thus preventing automated analysis. In addition, the high-quality near-field data may not be available in cases of highly variable shallow rock (multiple reflections), and this may lead to less accurate inverted results with near surface artifacts. To circumvent this issue, the FWI analysis has been modified and improved using estimated source signatures, which are determined by deconvolution of observed data with the Green's function. The benefit of using modified source signatures is that they can

act as a low-pass filter and address a significant fraction of the residuals between elastic estimated data and viscoelastic measured data (Groos et al., 2014).

Implementation of the modified FWI scheme involves the following steps:

- (1) Condition measured data and generate an initial model as presented in Sections 2.3.1 and 2.3.2**
- (2) Determine a source signature for each shot location**

The source estimation approach (Busch et al., 2012) developed for ground penetrating radar (GPR) data is employed for the seismic wave data. That is, in the time domain, the wave field data is a convolution of a source signature and the Green's function. In the frequency domain, it is equivalent to a multiplication:

$$\mathbf{F}(f, \mathbf{x}, \mathbf{m}) = \mathbf{G}(f, \mathbf{x}, \mathbf{m}) \cdot \mathbf{W}(f), \quad \text{Eq. 2.7}$$

where  $\mathbf{F}$ ,  $\mathbf{G}$ ,  $\mathbf{W}$  are the wave field, the Green's function and the source associated with a frequency  $f$ , respectively;  $\mathbf{m}$  is a model describing the parameters of the medium; and  $\mathbf{x}$  is space coordinates of a source and a receiver. With the model  $\mathbf{m}$ , the Green's function  $\mathbf{G}(f, \mathbf{x}, \mathbf{m})$  can be calculated by forward modelling with an assumed source  $\mathbf{W}(f)$  (e.g., Ricker wavelet). The estimated source is then obtained by deconvolution of measured data with the Green's function. Because the measured data consists of several offsets, the best-fit source wavelet  $\mathbf{W}_{\text{est}}(f)$  for each frequency  $f$  can be obtained by applying a least squares technique to solve the overdetermined system of equations. The inverse Fourier transform is then used to convert the estimated source from the frequency-domain to the time-domain.

- (3) Calculate residual between estimated and observed data for the  $i$ -th shot and  $j$ -th receiver:**

$$\Delta \mathbf{d}_{i,j} = \mathbf{F}_{i,j}(\mathbf{m}) - \mathbf{d}_{i,j}, \quad \text{Eq. 2.8}$$

where  $\mathbf{d}_{i,j}$  and  $\mathbf{F}_{i,j}(\mathbf{m})$  are the time-domain observed data and the estimated data associated with the model  $\mathbf{m}$  and the estimated source from step 2.

**(4) Calculate the least-squares error  $E(\mathbf{m})$ :**

$$E(\mathbf{m}) = \frac{1}{2} \Delta \mathbf{d}^t \Delta \mathbf{d}, \text{ where } \Delta \mathbf{d} = \{\Delta \mathbf{d}_{i,j}, i = 1, \dots, NS, j = 1 \dots NR\} \quad \text{Eq. 2.9}$$

where the superscript  $t$  denotes the matrix transpose.  $NS$  and  $NR$  are the numbers of shots and receivers, and  $\Delta \mathbf{d}$  is a column vector, which is the combination of residuals  $\Delta \mathbf{d}_{i,j}$  for all shots and receivers.  $\Delta \mathbf{d}$  is a column vector, which is the combination of residuals  $\Delta \mathbf{d}_{i,j}$  for all shots and receivers. If the number of time steps for each shot is  $NT$ , the size of  $\Delta \mathbf{d}$  is  $NT \times NS \times NR$ .

**(5) Calculate partial derivative of seismograms from the  $i$ -th shot and  $j$ -th receiver with respect to the  $p$ -th model parameter ( $m_p$ ):**

$$\mathbf{J}_{i,j} = \frac{\partial \mathbf{F}_{i,j}(\mathbf{m})}{\partial m_p}, i = 1 \dots NS, j = 1 \dots NR, p = 1 \dots M \quad \text{Eq. 2.10}$$

where  $M$  is the numbers of model parameters. If one calculates the whole matrix  $\mathbf{J}$ , its sizes will be  $NT \times NS \times NR$  rows and  $M$  columns. The partial derivative of seismograms with respect to each model parameter can be directly computed from the residual of two seismograms with and without perturbation of the model parameter. By perturbing individual model parameters, it requires  $(M + 1)$  forward modeling simulations for one shot or a total of  $NS \times (M + 1)$  simulations for the calculation of the matrix  $\mathbf{J}$ . Unfortunately, significant computer time is required for cases of a few thousand unknowns. To reduce the computations, the efficient technique of Sheen et al. (2006) of using virtual sources and reciprocity of wave fields, which requires only  $(NS + NR)$  simulations for the calculation of the matrix  $\mathbf{J}$  was employed. As an added benefit of the approach, it does not require storing the whole  $\mathbf{J}$  matrix when updating the model parameters. For details, see the work of Sheen et al. (2006).

**(6) Update model  $\mathbf{m}$  at iteration  $n+1$  from iteration  $n$  using the Tikhonov type regularization**

**(Tikhonov and Arsenin, 1977):**

$$\mathbf{m}^{n+1} = \mathbf{m}^n - \alpha^n [\mathbf{J}^t \mathbf{J} + \lambda_1 \mathbf{P}^t \mathbf{P} + \lambda_2 \mathbf{I}^t \mathbf{I}]^{-1} \mathbf{J}^t \Delta \mathbf{d}, \quad \text{Eq. 2.11}$$

Using the Gauss-Newton method for a full waveform inversion (ill-posed problem), regularization is important to maintain the optimization stability, especially for inversion of voids. In Eq. 2.11,  $\mathbf{I}$  is the identity matrix and  $\mathbf{P}$  is a matrix whose elements are either 1, -4, or 0, determined by a 2-D Laplacian operator. The choice of coefficients  $\lambda_1$  and  $\lambda_2$  between 0 and infinity is a compromise result. A larger value of  $\lambda_1$  provides more optimization stability but produces smoother inverted velocity models, which are not good for imaging the contrast in velocity between voids and soils. A larger value of  $\lambda_2$  provides less smooth inverted velocity but generates more artifacts. It was found that the values of 0.05 for  $\lambda_1$  and 0.0005 for  $\lambda_2$  are reasonable for this type of application. The optimal step length ( $\alpha^n$ ) is determined as:

$$\alpha^n \cong \frac{[\mathbf{J}^t \mathbf{g}^n]^t [\mathbf{F}(\mathbf{m}^n) - \mathbf{d}]}{[\mathbf{J}^t \mathbf{g}^n]^t [\mathbf{J}^t \mathbf{g}^n]}, \quad \text{Eq. 2.12}$$
$$\mathbf{g}^n = [\mathbf{J}^t \mathbf{J} + \lambda_1 \mathbf{P}^t \mathbf{P} + \lambda_2 \mathbf{I}^t \mathbf{I}]^{-1} \mathbf{J}^t [\mathbf{F}(\mathbf{m}^n) - \mathbf{d}].$$

**(7) Repeat steps 2 to 6 until the convergence is achieved.**

Using the optimal step length (Eq. 2.12), the inversion analysis is typically stopped after 10 to 15 iterations when the change of the least-squares error  $E(\mathbf{m})$  from one to the next iteration is less than 1%.

### **2.3.3.2 Results**

Two inversion runs were performed with central frequencies of 10 and 20 Hz, starting with the lower frequency run on the initial model shown in Figures 2.27 and 2.28. During the inversion, the medium of  $18 \times 36$  m ( $60 \times 120$  ft) was divided into cells of  $0.75 \times 0.75$  m ( $2.5 \times$

2.5 ft). S-wave and P-wave velocities of cells were updated simultaneously (no grid reduction). The first and second runs stopped respectively after 10 and 12 iterations, when the observed waveform data and the estimated waveform data were similar (Figure 2.29). Each run took about 30 minutes.

The inversion results are shown in Figures 2.30 and 2.31 for the two runs. It is observed that the final inverted S-wave velocity profile (Fig. 2.31) includes a shallow soil layer and a variable bedrock (S-wave velocity,  $V_s > 700$  m/s (2296 ft/s)) below about 6 m (20 ft) depth. It also shows an embedded anomaly at distance 18 m (60 ft) ( $V_s < 100$  m/s (328 ft/s)). The inverted P-wave profile (Fig. 2.31) is consistent with the S-wave profile, except the void that may be filled with water.

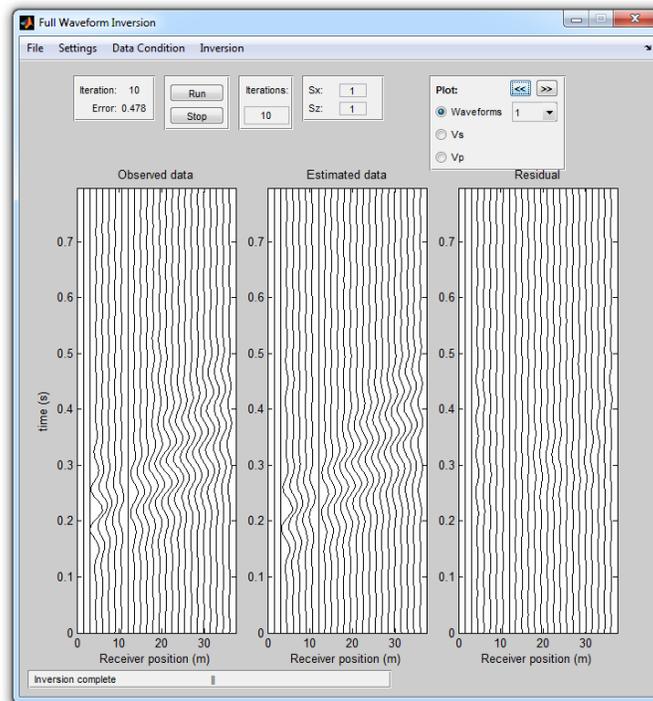


Figure 2.29 Comparison of observed and estimated data for 1 shot

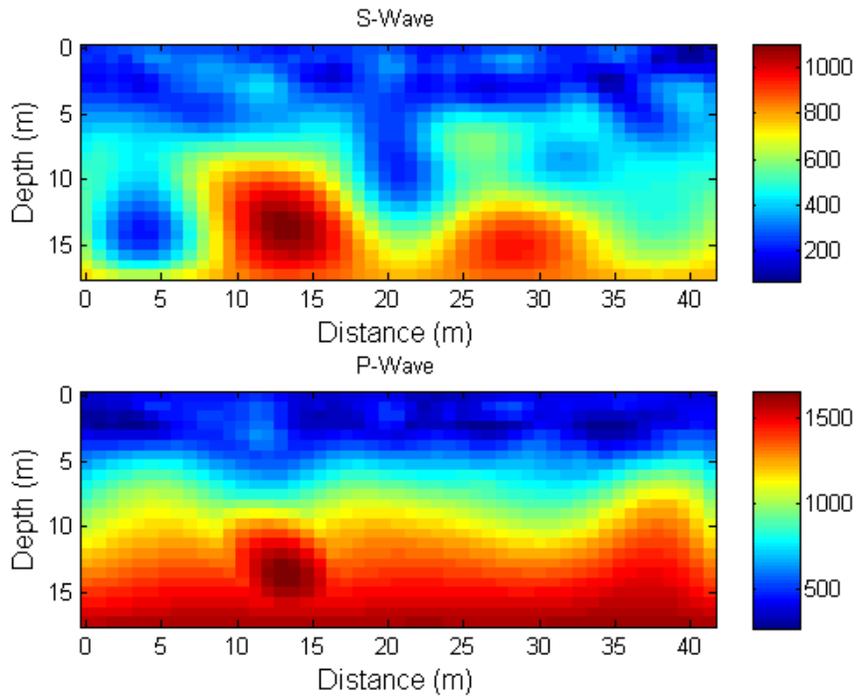


Figure 2.30 S-wave and P-wave velocities (m/s) – data at 10 Hz

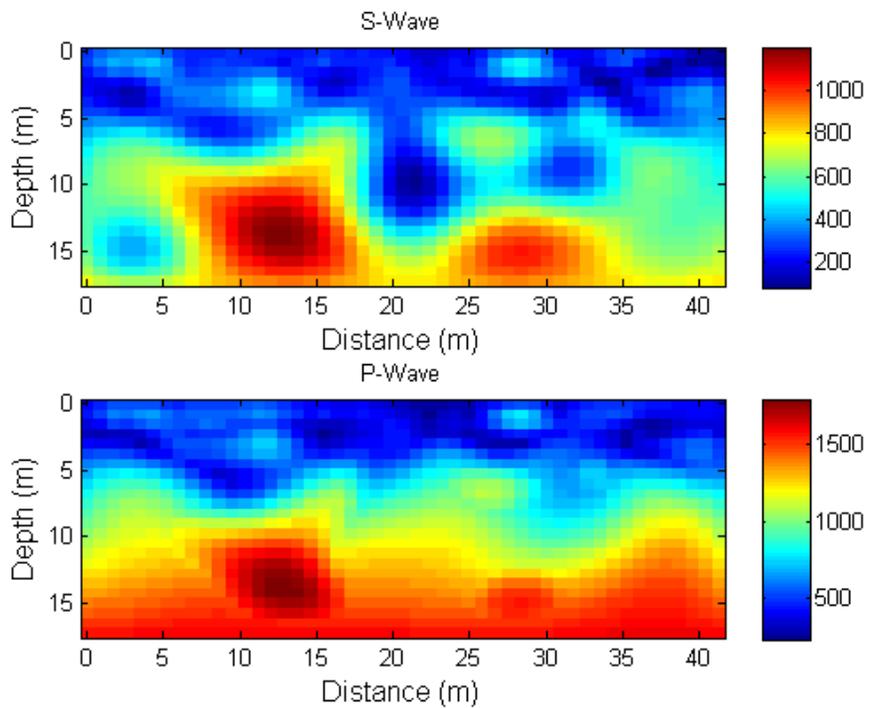


Figure 2.31 S-wave and P-wave velocities (m/s) – data at 20 Hz

## **2.4 Conclusion**

A fast and automatic algorithm of the full waveform inversion (FWI) for a field solution was developed. Convolutional perfectly matched layers, parallelizing computations, temporal windowing, and grid reduction have been implemented to reduce required computer time for waveform analysis. Visualized data conditioning, automated initial model, and automated analysis have also been implemented to reduce manual efforts during the analysis. The improved FWI algorithm can produce the field solution to obtain general information of the medium being tested within 20-30 minutes. More detailed information can be achieved by further analysis at higher frequencies after field testing.

## CHAPTER 3 DEVELOP INTERFACE SOFTWARE FOR THE 2-D FULL WAVEFORM INVERSION AND GUIDELINES FOR IMPLEMENTATION

### **3.1 Introduction**

The goal of this task is to develop a user interface software and user manual for FWI use. The effort is focused on developing the graphical user interface (GUI) for graphical input, analysis, and output. The GUI is developed for technician-level personnel that can operate in the field following basic training. For example, the user inputs the number of geophones, the spacing between the geophones, shot (source) locations, shot sequence, and the raw data collected in the field (receiver's data for all shots). The user is able to change the conditioning parameters (i.e., filtering, windowing, etc.), remove bad channels (bad receivers) of recorded data, and check the quality of the conditioned data for individual shots before they are used in the inversion analysis. The GUI performs the FWI analysis (i.e., multiple frequency ranges, different medium sizes) developed in Task 1. The software subsequently produces profiles of subsurface P-wave and S-wave velocities, which are displayed on the GUI. The user is able to save the conditioning parameters, conditioned data, wave velocities (P and S), and waveforms (error analysis), which can be opened in the GUI for future analysis and transferring analysis files.

This report includes the user manual for the FWI software and is accompanied by the executable file to run the software. A summary of the software development and validation follows.

### **3.2 Summary of Software Development and Validation**

The goal of developing a GUI for the software was for a user to upload shot files and obtain profiles of subsurface P-wave and S-wave in the field within 30 minutes. To accomplish this, the GUI was developed using the programming language C# (sharp) and the computational components of the software were written in C++ and Matlab DLLs (dynamic link libraries). As a result, the computation time in the inversion analysis of 13 shots was approximately 2.7 minutes per iteration or approximately 27 minutes for 10 iterations.

To guarantee that the code conversion to C++ and Matlab DLL was done correctly, without any losses or introduced error, each major computational part of the code had to be validated. The validation consisted of running a dataset (e.g., parameters, shot files) in the new code, then the original Matlab code and comparing the results. In each case, the absolute difference between the two results at each time step was calculated, and in each, there was zero or a negligible difference.

### **3.3 Conclusion**

A standalone, software program, with GUI, for the FWI of shot data has been developed. The program is capable of analyzing shot data and providing estimates of the subsurface P-wave and S-wave velocities in approximately 2.7 minutes per iteration. As a feature for error analysis, the program provides the measured, estimated, and residual waveforms per iteration. Furthermore, the inversion results can be saved and opened in the program, allowing for future analysis and transfer of analysis files. A user manual for the FWI software is included in the Appendix.

## CHAPTER 4 INVESTIGATION OF SEISMIC RESPONSES OF EMBEDDED VOIDS ON SYNTHETIC MODELS

### 4.1 Introduction

The Full Waveform Inversion (FWI) of seismic waves discussed in Task 1 has shown great potential for detection of anomalies such as voids. However, the ability of the FWI to detect buried anomalies is dependent on the depth and size of the void. Specifically, the void must be large enough that seismic wave will strike the void and will be reflected back (function of wave length) and have sufficient energy when it arrives at the ground surface that it will be detected (i.e., amplitude). Loss of signal due to damping is a function of material that the wave passes through and the distance the wave travels. This task was to determine the effect of varying depth on the ability of the program to detect voids. Specifically, a finite 3-D void was embedded at various depths, and waveform data was generated by a newly developed 3-D wave solution for an array of shots and receivers on top of the void (offline voids will be investigated in Task 4). Recorded waveforms with the inclusion of the void were compared with those from the model without the void to evaluate the sensitivity of void detection (energy change). Recorded 3-D waveforms with the inclusion of the void will also be analyzed to evaluate capability of the 2-D FWI in detection of 3-D voids.

### 4.2 Methodology

The study was conducted by generating synthetic data for a dual layer model (soil overlying rock) with several cases of a 4.5 m × 4.5 m × 4.5 m (15 ft × 15 ft × 15 ft) void at varying depths. The base model, Figure 4.1, is one that might be found in the field. It consists of a soil layer with a shear wave velocity of 200 m/s (656 ft/s) and a limestone layer that has a shear wave velocity of 700 m/s (2296 ft/s).

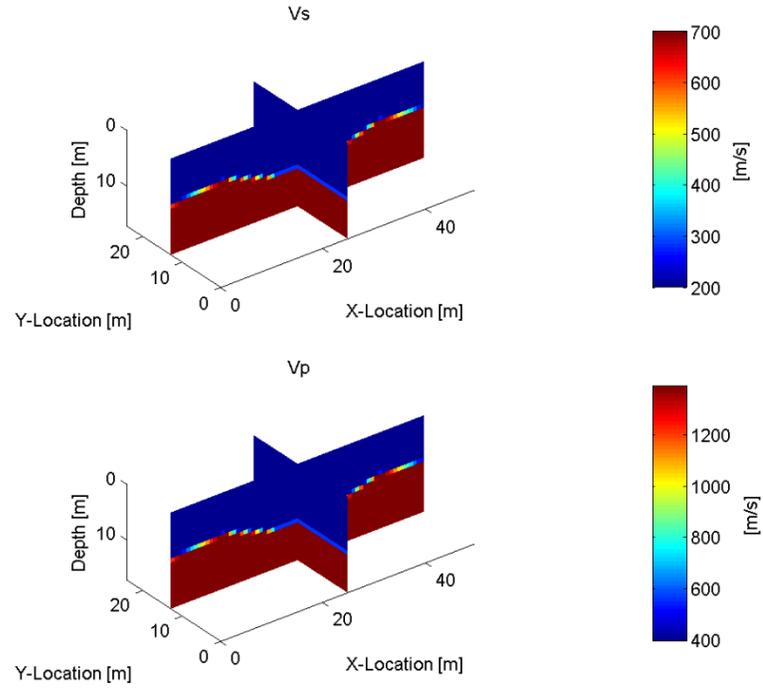


Figure 4.1 Dual layer model

The compression wave velocities are calculated using Equation 4.1,

$$V_p = \sqrt{\frac{2(1-\nu)}{1-2\nu}} V_s \quad \text{Eq. 4.1}$$

where  $\nu$ , Poisson's ratio, is set at 0.3 and  $V_s$  is the compression wave velocity in m/s.

The voids are located at depths of 4.5 m (Figure 4.2), 9 m (Figure 4.3), and 13.5 m (Figure 4.4) (15 ft, 30 ft, and 44 ft, respectively).

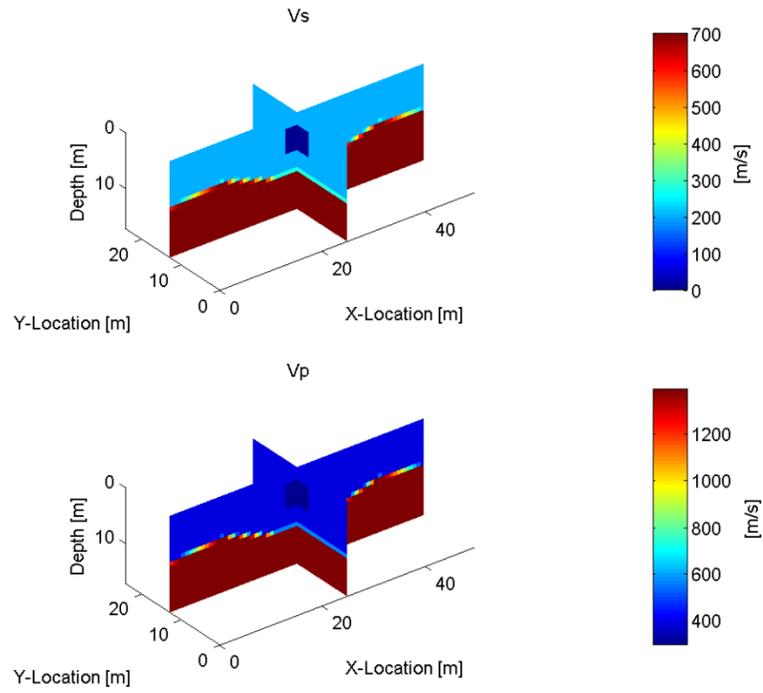


Figure 4.2 Void – 4.5 m deep

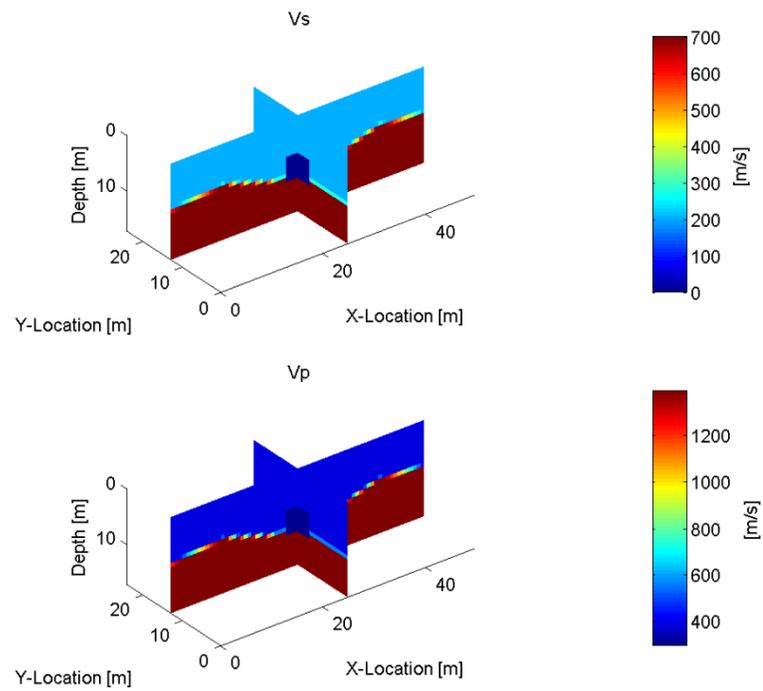


Figure 4.3 Void – 9 m deep

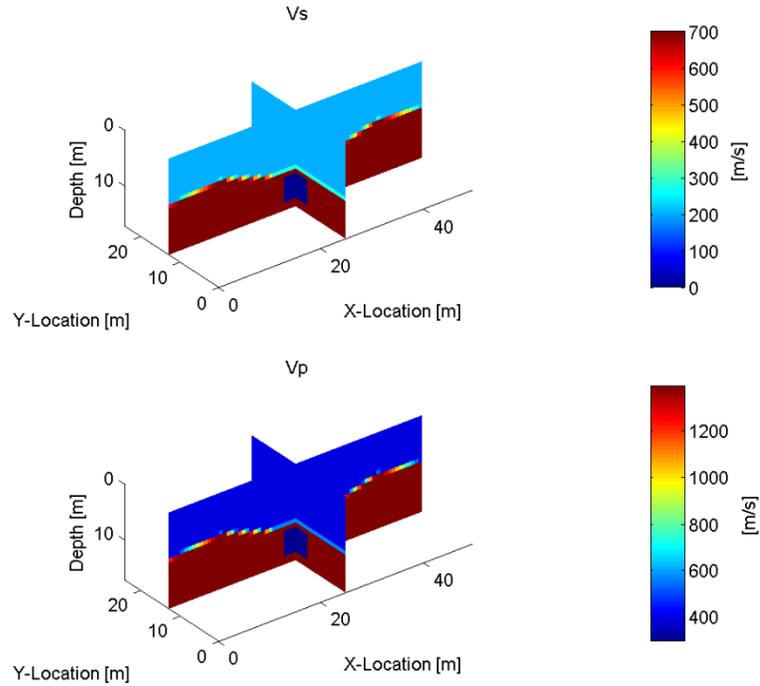


Figure 4.4 Void – 13.5 m deep

Each of these models is used to generate synthetic 3-D data to allow for accurate representation of the 3-D voids.

#### 4.2.1 3-D Forward Modeling

The synthetic 3-D data was created by using a set of first-order linear partial differential equations for isotropic materials (Equations 4.2 through 4.10). The first three of the elastodynamic equations governs particle velocity while the remaining equations govern the stress-strain tensors.

$$\dot{v}_x = \frac{1}{\rho} (\sigma_{xx,x} + \sigma_{xy,y} + \sigma_{xz,z}) + f_x \quad \text{Eq. 4.2}$$

$$\dot{v}_y = \frac{1}{\rho} (\sigma_{xy,x} + \sigma_{yy,y} + \sigma_{yz,z}) + f_y \quad \text{Eq. 4.3}$$

$$\dot{v}_z = \frac{1}{\rho} (\sigma_{xz,x} + \sigma_{yz,y} + \sigma_{zz,z}) + f_z \quad \text{Eq. 4.4}$$

$$\dot{\sigma}_{xx} = (\lambda + 2\mu)v_{x,x} + \lambda(v_{y,y} + v_{z,z}) \quad \text{Eq. 4.5}$$

$$\dot{\sigma}_{yy} = (\lambda + 2\mu)v_{y,y} + \lambda(v_{x,x} + v_{z,z}) \quad \text{Eq. 4.6}$$

$$\dot{\sigma}_{zz} = (\lambda + 2\mu)v_{z,z} + \lambda(v_{x,x} + v_{y,y}) \quad \text{Eq. 4.7}$$

$$\dot{\sigma}_{xy} = \mu(v_{x,y} + v_{y,x}) \quad \text{Eq. 4.8}$$

$$\dot{\sigma}_{xz} = \mu(v_{x,z} + v_{z,x}) \quad \text{Eq. 4.9}$$

$$\dot{\sigma}_{yz} = \mu(v_{y,z} + v_{z,y}) \quad \text{Eq. 4.10}$$

where,  $\mathbf{V}$ , with components:  $V_x, V_y, V_z$ , is the particle velocity vector;  $\mathbf{f}$ , with components  $f_x, f_y, f_z$ , is the body force vector;  $\boldsymbol{\sigma}$ , with components  $\sigma_{xx}, \sigma_{yy}, \sigma_{zz}, \sigma_{xy}, \sigma_{xz}, \sigma_{yz}$ , is the stress tensor;  $\rho$  is the mass density; and  $\mu$  and  $\lambda$  are Lamé's parameters. The subscript notation with a comma denotes a spatial derivative with respect to the coordinates  $x, y$ , and  $z$ , (e.g.,  $\sigma_{xz,x} = \frac{\partial \sigma_{xz}}{\partial x}$ ) and the over-dot denotes a time derivative.

The equations were implemented utilizing the velocity-stress staggered-grid finite difference technique in the time domain (Figure 4.5). The advantages of this technique include (i) source insertion can be expressed by velocity or stress; (ii) a stable and accurate representation for a planar free-surface boundary is easily implemented; (iii) the algorithm can be conveniently implemented on scalar, vector, or parallel computers; (iv) signal filtering and boundary truncation can be implemented with minimum effort; and (v) wavefields at multiple frequencies can be generated simultaneously in the time domain.

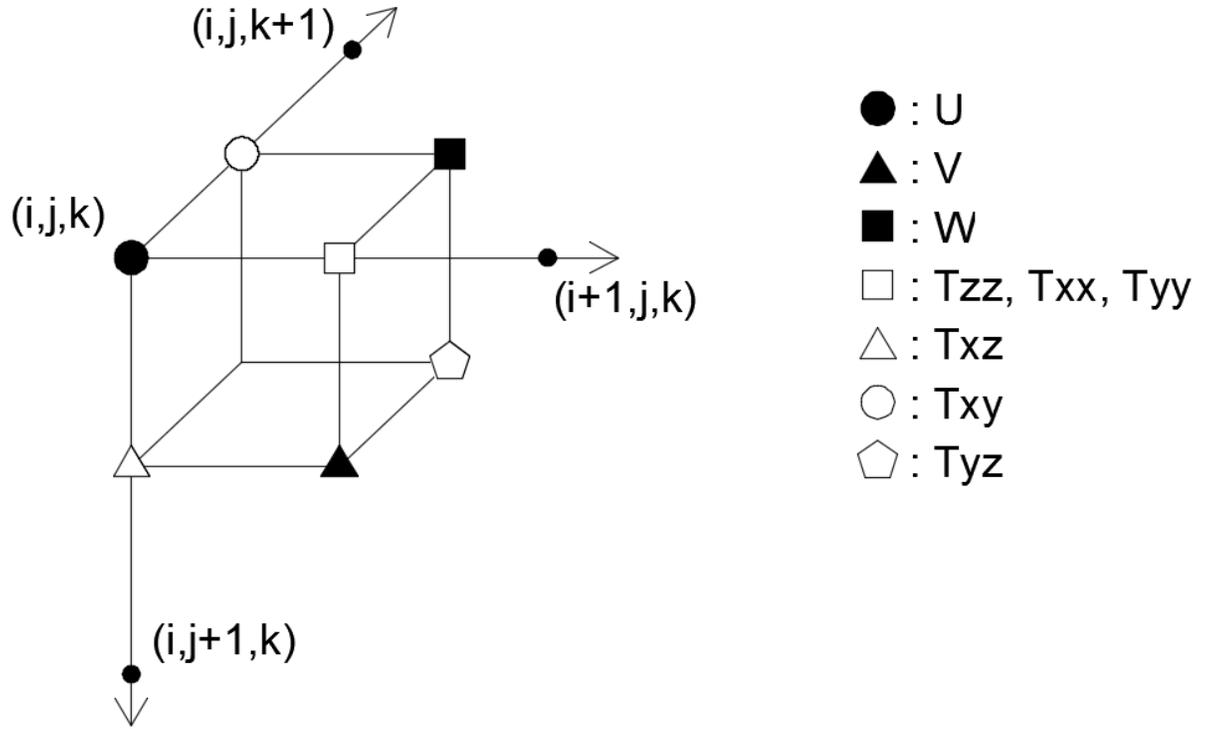


Figure 4.5 Staggered 3-D grid

With each time step the stress and velocities are updated to model the wave as accurately as possible. This is done by expressing Equations 4.2 through 4.10 in a finite difference form for both the velocities and stresses (Equations 4.11 through 4.19).

$$\begin{aligned}
 U_{i,j,k}^{n+\frac{1}{2}} = & U_{i,j,k}^{n-\frac{1}{2}} + B_{i,j,k} \frac{\Delta t}{\Delta x} \left( T_{xx}^n_{i+\frac{1}{2},j,k} - T_{xx}^n_{i-\frac{1}{2},j,k} \right) \\
 & + B_{i,j,k} \frac{\Delta t}{\Delta y} \left( T_{xy}^n_{i,j,k+\frac{1}{2}} - T_{xy}^n_{i,j,k-\frac{1}{2}} \right) \\
 & + B_{i,j,k} \frac{\Delta t}{\Delta z} \left( T_{xz}^n_{i,j+\frac{1}{2},k} - T_{xz}^n_{i,j-\frac{1}{2},k} \right)
 \end{aligned}
 \tag{Eq. 4.11}$$

$$\begin{aligned}
V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} &= V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n-\frac{1}{2}} + B_{i+\frac{1}{2},j+\frac{1}{2},k} \frac{\Delta t}{\Delta x} \left( T_{xz}^n_{i+1,j+\frac{1}{2},k} - T_{xz}^n_{i,j+\frac{1}{2},k} \right) \\
&\quad + B_{i+\frac{1}{2},j+\frac{1}{2},k} \frac{\Delta t}{\Delta y} \left( T_{yz}^n_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} - T_{yz}^n_{i+\frac{1}{2},j+\frac{1}{2},k-\frac{1}{2}} \right) \\
&\quad + B_{i+\frac{1}{2},j+\frac{1}{2},k} \frac{\Delta t}{\Delta z} \left( T_{zz}^n_{i+\frac{1}{2},j+1,k} - T_{zz}^n_{i+\frac{1}{2},j,k} \right)
\end{aligned} \tag{Eq. 4.12}$$

$$\begin{aligned}
W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} &= W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n-\frac{1}{2}} + B_{i+\frac{1}{2},j,k+\frac{1}{2}} \frac{\Delta t}{\Delta x} \left( T_{xy}^n_{i+1,j,k+\frac{1}{2}} - T_{xy}^n_{i,j,k+\frac{1}{2}} \right) \\
&\quad + B_{i+\frac{1}{2},j,k+\frac{1}{2}} \frac{\Delta t}{\Delta y} \left( T_{yy}^n_{i+\frac{1}{2},j,k+1} - T_{yy}^n_{i+\frac{1}{2},j,k} \right) \\
&\quad + B_{i+\frac{1}{2},j,k+\frac{1}{2}} \frac{\Delta t}{\Delta z} \left( T_{yz}^n_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} - T_{yz}^n_{i+\frac{1}{2},j-\frac{1}{2},k+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.13}$$

$$\begin{aligned}
T_{xx}^{n+1}_{i+\frac{1}{2},j,k} &= T_{xx}^n_{i+\frac{1}{2},j,k} + (L + 2M)_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta x} \left( U_{i+1,j,k}^{n+\frac{1}{2}} - U_{i,j,k}^{n+\frac{1}{2}} \right) \\
&\quad + L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta y} \left( W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i+\frac{1}{2},j,k-\frac{1}{2}}^{n+\frac{1}{2}} \right) \\
&\quad + L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta z} \left( V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} - V_{i+\frac{1}{2},j-\frac{1}{2},k}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.14}$$

$$\begin{aligned}
T_{yy}^{n+1}_{i+\frac{1}{2},j,k} &= T_{yy}^n_{i+\frac{1}{2},j,k} + (L + 2M)_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta y} \left( W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i+\frac{1}{2},j,k-\frac{1}{2}}^{n+\frac{1}{2}} \right) \\
&\quad + L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta x} \left( U_{i+1,j,k}^{n+\frac{1}{2}} - U_{i,j,k}^{n+\frac{1}{2}} \right) \\
&\quad + L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta z} \left( V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} - V_{i+\frac{1}{2},j-\frac{1}{2},k}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.15}$$

$$\begin{aligned}
T_{ZZ}^{n+1}{}_{i+\frac{1}{2},j,k} &= T_{ZZ}^n{}_{i+\frac{1}{2},j,k} + (L + 2M)_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta Z} \left( V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} - V_{i+\frac{1}{2},j-\frac{1}{2},k}^{n+\frac{1}{2}} \right) \\
&+ L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta x} \left( U_{i+1,j,k}^{n+\frac{1}{2}} - U_{i,j,k}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.16}$$

$$+ L_{i+\frac{1}{2},j,k} \frac{\Delta t}{\Delta y} \left( W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i+\frac{1}{2},j,k-\frac{1}{2}}^{n+\frac{1}{2}} \right)$$

$$\begin{aligned}
T_{xy}^{n+1}{}_{i,j,k+\frac{1}{2}} &= T_{xy}^n{}_{i,j,k+\frac{1}{2}} + M_{i,j,k+\frac{1}{2}} \frac{\Delta t}{\Delta y} \left( U_{i,j,k+1}^{n+\frac{1}{2}} - U_{i,j,k}^{n+\frac{1}{2}} \right) \\
&+ M_{i,j,k+\frac{1}{2}} \frac{\Delta t}{\Delta x} \left( W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i-\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.17}$$

$$\begin{aligned}
T_{xz}^{n+1}{}_{i,j+\frac{1}{2},k} &= T_{xz}^n{}_{i,j+\frac{1}{2},k} + M_{i,j+\frac{1}{2},k} \frac{\Delta t}{\Delta Z} \left( U_{i,j+1,k}^{n+\frac{1}{2}} - U_{i,j,k}^{n+\frac{1}{2}} \right) \\
&+ M_{i,j+\frac{1}{2},k} \frac{\Delta t}{\Delta x} \left( V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} - V_{i-\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.18}$$

$$\begin{aligned}
T_{yz}^{n+1}{}_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} &= T_{yz}^n{}_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} + M_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} \frac{\Delta t}{\Delta Z} \left( W_{i+\frac{1}{2},j+1,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i+\frac{1}{2},j,k+\frac{1}{2}}^{n+\frac{1}{2}} \right) \\
&+ M_{i+\frac{1}{2},j+\frac{1}{2},k+\frac{1}{2}} \frac{\Delta t}{\Delta y} \left( V_{i+\frac{1}{2},j+\frac{1}{2},k+1}^{n+\frac{1}{2}} - V_{i+\frac{1}{2},j+\frac{1}{2},k}^{n+\frac{1}{2}} \right)
\end{aligned} \tag{Eq. 4.19}$$

where  $i, j, k$  are the indicial location of the current point in the x, z, and y directions respectively;  $n$  represents the current time step while  $U, V$ , and  $W$  are the velocities in the x, z, and y directions.  $B$  is the inverse of the density at the location indicated by the indices and  $T$  represent the stresses in each direction.  $M$  and  $L$  are the Lamé's coefficients mentioned earlier in Equations 4.2 through 4.10. These equations are used for each time step based on the staggered grid to allow for waves to be modeled in time.

Special conditions are required at the boundaries of a modeled area when implementing finite difference equations for seismic waves. Specifically, any model without

boundary conditions will have boundaries that act like walls, or any incoming wave will be reflected off the boundary and propagate back through the domain. Obviously, any reflections do not accurately represent the field, i.e., an infinite medium. An ideal model will allow any wave to pass through the boundary without any reflections. The two conditions being used are a free surface condition and an absorbing condition.

The free surface condition, applied at the surface ( $z = 0$ ), occurs naturally in the field when the medium being modeled encounters open air. This allows for both the perpendicular normal stress ( $\sigma_{zz}$ ) and the parallel shear stresses ( $\sigma_{xz}$ ,  $\sigma_{yz}$ ) to be set to zero (Tran & McVay, 2012). This condition is critical when implementing the staggered grid due to the fact that numerically calculating  $\frac{\partial \sigma_{xz}}{\partial z}$ ,  $\frac{\partial \sigma_{yz}}{\partial z}$ , and  $\frac{\partial V_z}{\partial z}$  at the surface is difficult without the grid points for  $\sigma_{xz}$ ,  $\sigma_{yz}$ , and  $V_z$  above the surface.

The surface condition is carried out in two different ways. The first is done by substituting zero for  $\sigma_{xx}$  into Equation 4.7 and solving for  $\frac{\partial V_z}{\partial z}$ .

$$\frac{\partial V_z}{\partial z} = -\frac{\lambda}{\lambda + 2\mu} \left( \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} \right) \quad \text{Eq. 4.20}$$

Equation 4.20 is then put in terms of the staggered grid,

$$\begin{aligned} \left( \frac{\partial V}{\partial z} \right)_{i+\frac{1}{2},1,k}^{n+\frac{1}{2}} &= -\frac{L_{i+\frac{1}{2},1,k}}{(L+2M)_{i+\frac{1}{2},1,k}} \left( \frac{1}{dx} \left( U_{i+1,1,k}^{n+\frac{1}{2}} - U_{i,1,k}^{n+\frac{1}{2}} \right) \right. \\ &\quad \left. + \frac{1}{dy} \left( W_{i+\frac{1}{2},1,k+\frac{1}{2}}^{n+\frac{1}{2}} - W_{i+\frac{1}{2},1,k-\frac{1}{2}}^{n+\frac{1}{2}} \right) \right) \end{aligned} \quad \text{Eq. 4.21}$$

where the  $z$  coordinate remains at the surface ( $j = 1$ ). This is implemented when calculating  $\sigma_{xx}$  at the surface (Equation 4.14).

The second application of the surface condition is done by setting the remaining stresses in the condition ( $\sigma_{xz}$  and  $\sigma_{yz}$ ) to zero and determining the slope at the surface mathematically (Figure 4.6 & Equations 4.22 & 4.23).

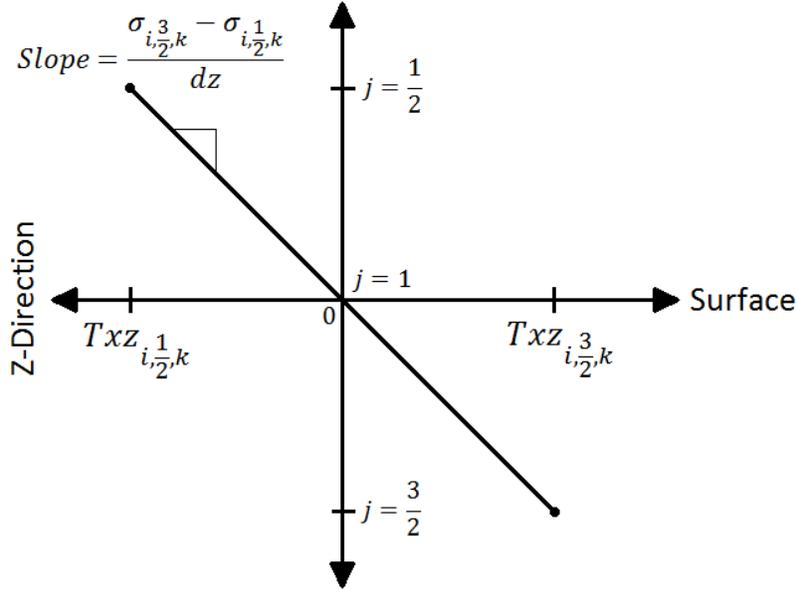


Figure 4.6  $\frac{\partial \sigma}{\partial z}$  surface plot

$$\begin{aligned} \left(\frac{\partial T_{xz}}{\partial z}\right)_{i,j,k}^n &= \frac{1}{dz} \left( T_{xz}_{i,3/2,k}^n - T_{xz}_{i,1/2,k}^n \right) = \frac{1}{dz} \left( T_{xz}_{i,3/2,k}^n - \left( -T_{xz}_{i,3/2,k}^n \right) \right) \\ &= \frac{2}{dz} T_{xz}_{i,3/2,k}^n \end{aligned} \quad \text{Eq. 4.22}$$

$$\begin{aligned} \left(\frac{\partial T_{yz}}{\partial z}\right)_{i+\frac{1}{2},j,k+\frac{1}{2}}^n &= \frac{1}{dz} \left( T_{yz}_{i+\frac{1}{2},3/2,k}^n - T_{yz}_{i+\frac{1}{2},1/2,k+\frac{1}{2}}^n \right) \\ &= \frac{1}{dz} \left( T_{yz}_{i+\frac{1}{2},3/2,k+\frac{1}{2}}^n - \left( -T_{yz}_{i+\frac{1}{2},3/2,k+\frac{1}{2}}^n \right) \right) = \frac{2}{dz} T_{yz}_{i+\frac{1}{2},3/2,k+\frac{1}{2}}^n \end{aligned} \quad \text{Eq. 4.23}$$

These equations are used when calculating both the velocity in x and y directions on the surface.

The remaining boundaries utilize an absorbing condition known as the perfectly matched layer (PML) which attenuates any waves near the boundary. The dampening effect is carried out

by extending the staggered grid normal to the surface requiring the boundary condition (Figure 4.7).

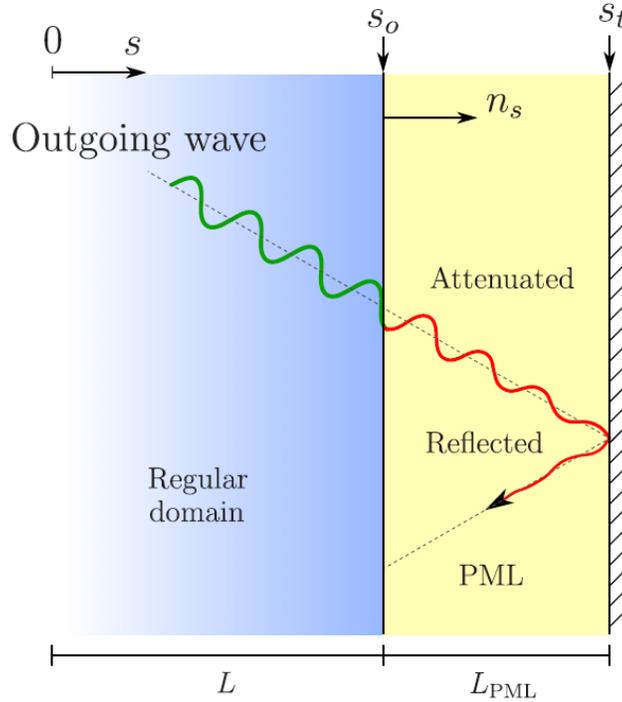


Figure 4.7 Perfectly matched layer added on to original domain

The  $S_o$  represents the original boundary while  $s_t$  is the boundary of the domain with the added grid points. Each grid point reduces the wave amplitude based on the values of  $a$  and  $b$  which are derived using several variables and constants.

$$b_x = e^{-\left(\frac{d_x}{k_x} + \alpha_x\right)\Delta t} \quad \text{Eq. 4.24}$$

$$a_x = \frac{d_x}{k_x(d_x + k_x\alpha_x)}(b_x - 1) \quad \text{Eq. 4.25}$$

where  $k_x$  is set to one to allow for the PML condition to replicate the classical PML coordinate transformation.  $\alpha_x$  is calculated based on the current grid point and the central frequency while  $d_x$  represents the damping profile such that  $d_x = 0$  inside the original domain and  $d_x > 0$  in the perfectly matched layer.  $a_x$  and  $b_x$  are implemented after the calculation of each stress and velocity derivative,  $\partial_x$ , by creating the memory variable  $\psi_x$  and applying it to Equation 4.27.

$$\psi_x^n = b_x \psi_x^{n-1} + a_x (\partial_x)^{n+\frac{1}{2}} \quad \text{Eq. 4.26}$$

$$\partial_{\bar{x}} = \frac{1}{k_x} \partial_x + \psi_x \quad \text{Eq. 4.27}$$

All interfering wave reflections are nullified when the boundary conditions discussed above are applied, allowing for proper modeling of infinite mediums such as those found in the field.

#### 4.2.2 3-D Wave Propagation Example

Both the surface and boundary modeling was evaluated with an example case of a homogeneous medium to verify that the waves behaved properly throughout the 3-D forward model. In this case, the model has a S-wave velocity of 200 m/s (656 ft/s) and a P-wave velocity that is generated from the S-wave velocity and a constant Poisson's ratio of 0.3 for the entire domain. The shot is located on the model surface.

The data was recorded for the whole medium at each time step to allow for 3-D images of the wave perturbing throughout the medium with and without the perfectly matched layer (PML) (Figure 4.8). The set of plots on the left indicate that the perfectly matched layer successfully dampens the wave impacting the boundaries while the plots on the right show significant reflected signals without the PML.

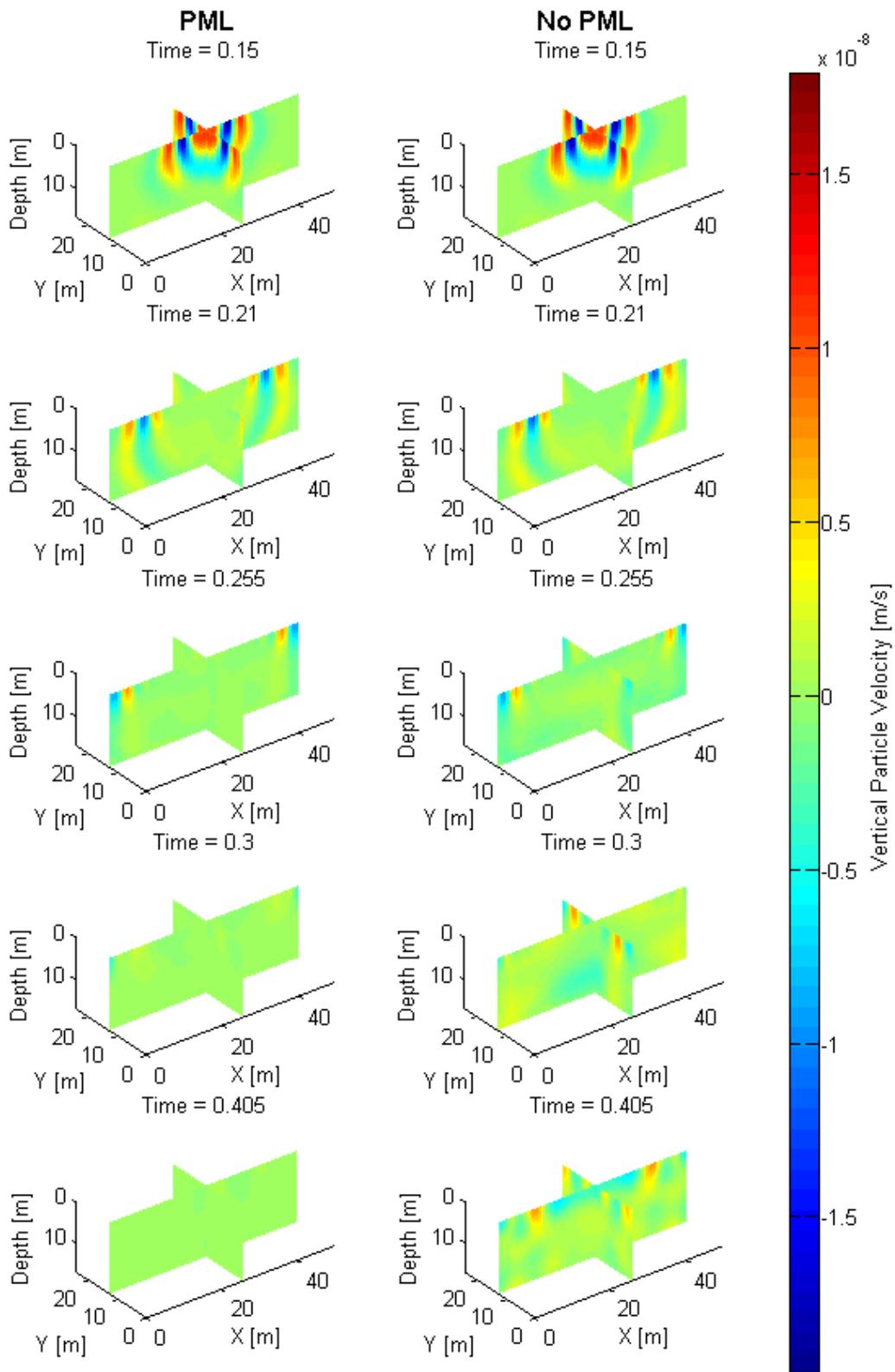


Figure 4.8 3-D wave propagation with and without the PML

Shown in Figure 4.9 and Figure 4.10 are the predicted surface waveform data for three receiver lines at 3 m separation in the y-direction due to two different shots at the end and center of the first receiver line. The 24 receiver points for each line are spaced 1.5 m apart along the x-axis (7.5 to 42 m (25 to 138 ft)). The resulting images indicate a steady progression of the wave throughout the medium with no reflections, which are expected in a homogeneous model.

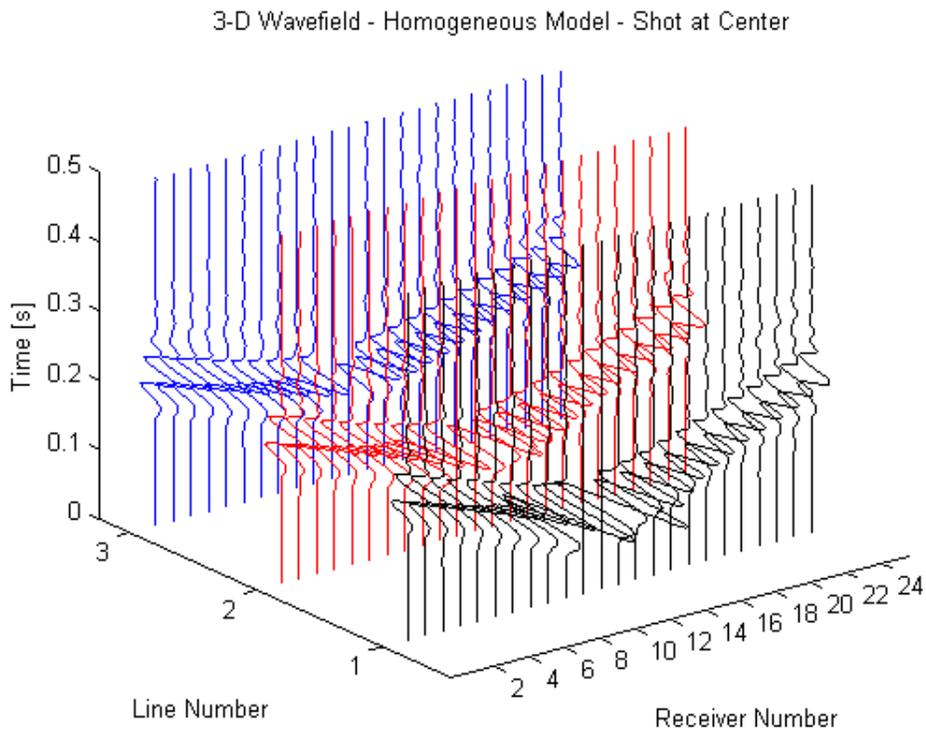


Figure 4.9 3-D synthetic waveform data – shot located at center of the line 1

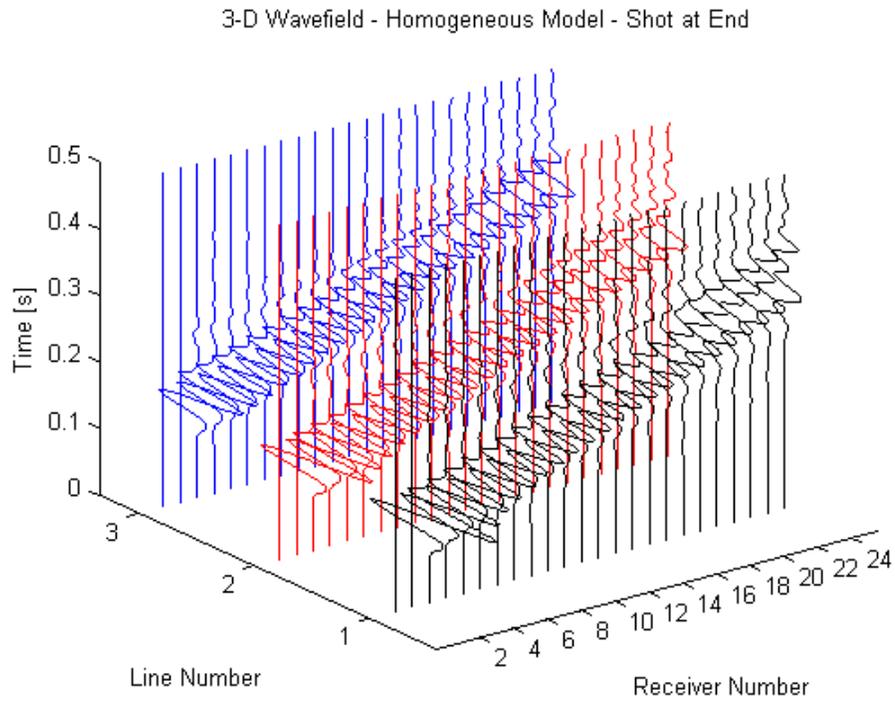


Figure 4.10 3-D synthetic waveform data – shot located at end of the line 1

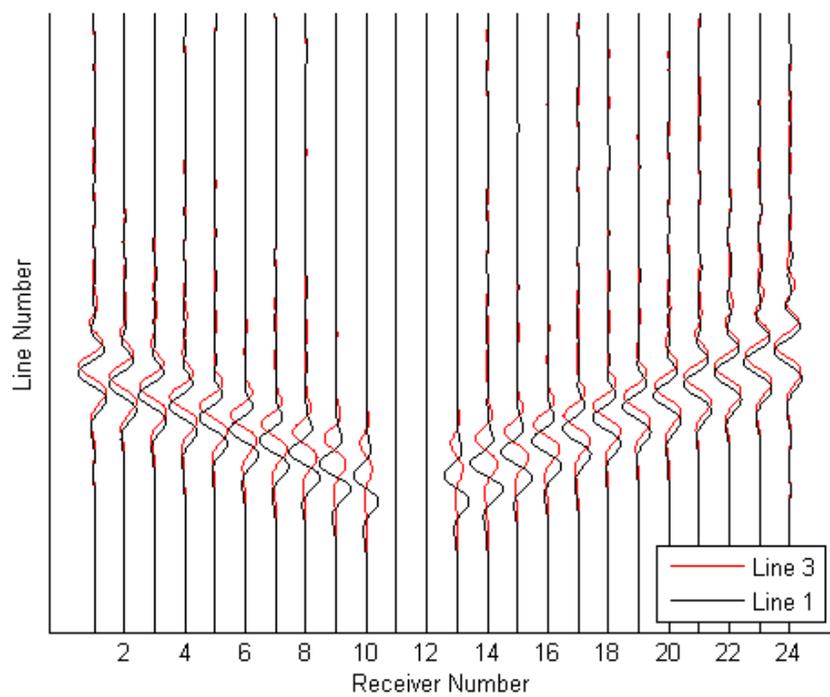


Figure 4.11 2-D wavefield – shot centered in line 1 – line 1 vs. line 3

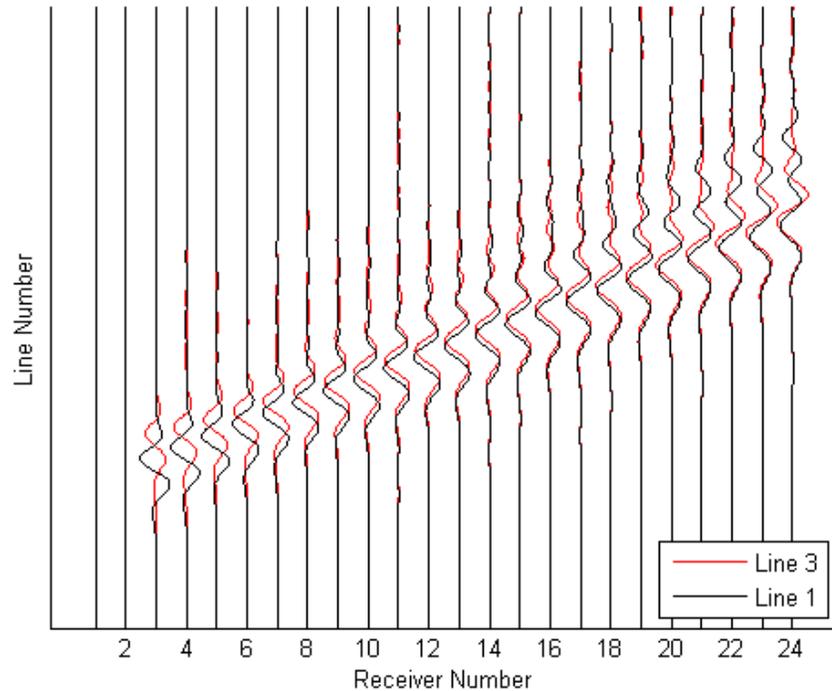


Figure 4.12 2-D wavefield – shot at end of line 1 – line 1 vs. line 3

In addition, both the first and third lines were overlaid to produce a 2-D image which makes the delayed wave arrival times viewable (Figures 4.11 and 4.12). Note, the receivers closest to the shot have been removed, and the data has been increased in magnitude (normalized), based on the closeness of the receiver to the shot. Receivers in the far field are magnified more, allowing for better viewing of the wave. Since the shots are perturbed on line 1, it is expected that the magnitude is greatest for line 1 with the other lines reduced in magnitude. There should also be a difference in arrival times when comparing the lines. This is evident when comparing lines 1 and 3 for different shots seen in the previous figures.

### 4.2.3 Energy Comparison

The model without the void and the three with the void were subsequently used in the 3-D forward model discussed earlier. The data from these models were gathered from 24 receiver points in a line on the surface (Figure 4.13).

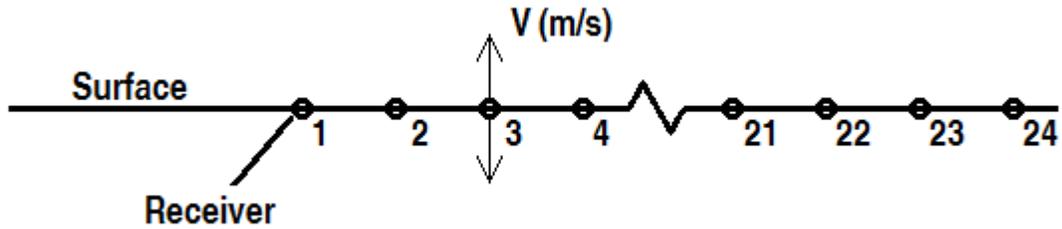


Figure 4.13 Vertical velocity at the receivers

The models with the voids are compared to the model without a void. This was done by calculating the percent change in energy of the wave in each of the receivers. It is expected that a model with voids will reflect wave energy back towards the surface. This can be seen when comparing a receiver above a void to one without a void (Figure 4.14).

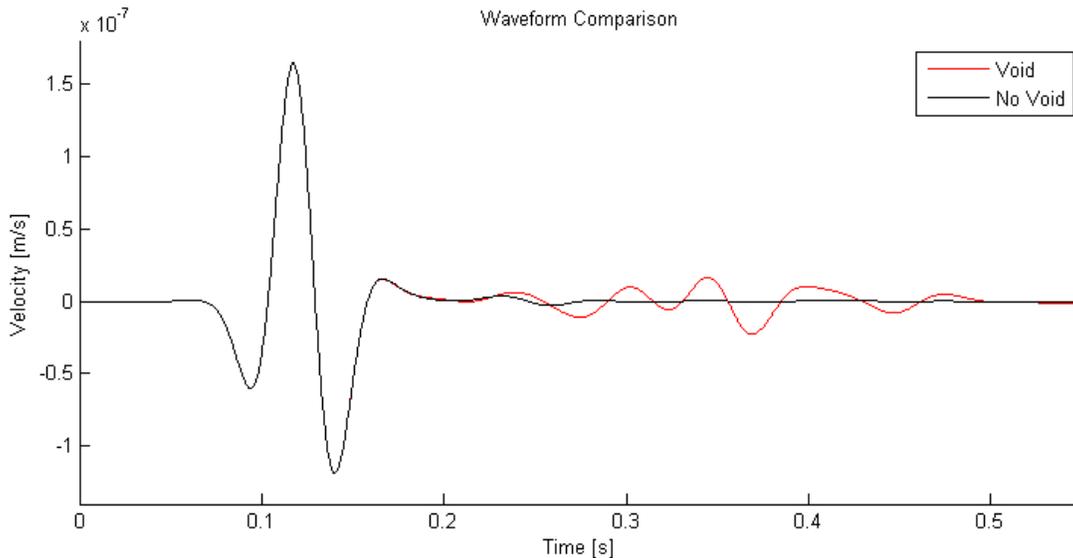


Figure 4.14 Receiver comparison

The two receivers show the same initial wave pattern but the receiver with the void indicates reflected waves in the later portion of the signal and more energy arriving at the surface (proportional to area under curve). Voids that are deeper will have wave amplitudes that are too dampened later in the signal with minimal impact on the energy arriving at the surface. It is possible to observe these effects by calculating the percent change in energy for voids at each depth.

The energy for each receiver can be calculated since each data point represents the vertical velocity of a mass. The kinetic energy of a mass is calculated using Equation 4.28,

$$E_k = \frac{1}{2}mV^2 \quad \text{Eq. 4.28}$$

where  $m$  is the mass and  $V$  is the velocity. The average energy for each receiver is used to account for all time steps. This is done using a Riemann sum shown in Figure 4.15.

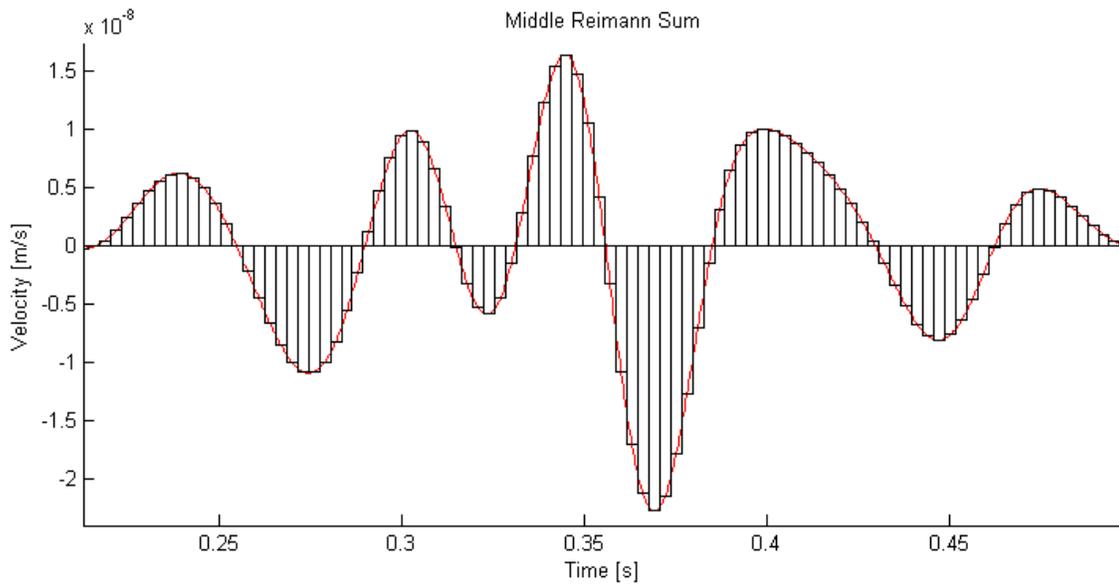


Figure 4.15 Riemann sum setup

The red line represents the actual data while the rectangles are the area taken up by multiplying each velocity by  $\Delta t$ . In this case figure the reflected wave is shown. Only the magnitudes of the velocities are taken into account when computing the Riemann sum to account for the total energy of the system (Figure 4.16 & Equation 4.29).

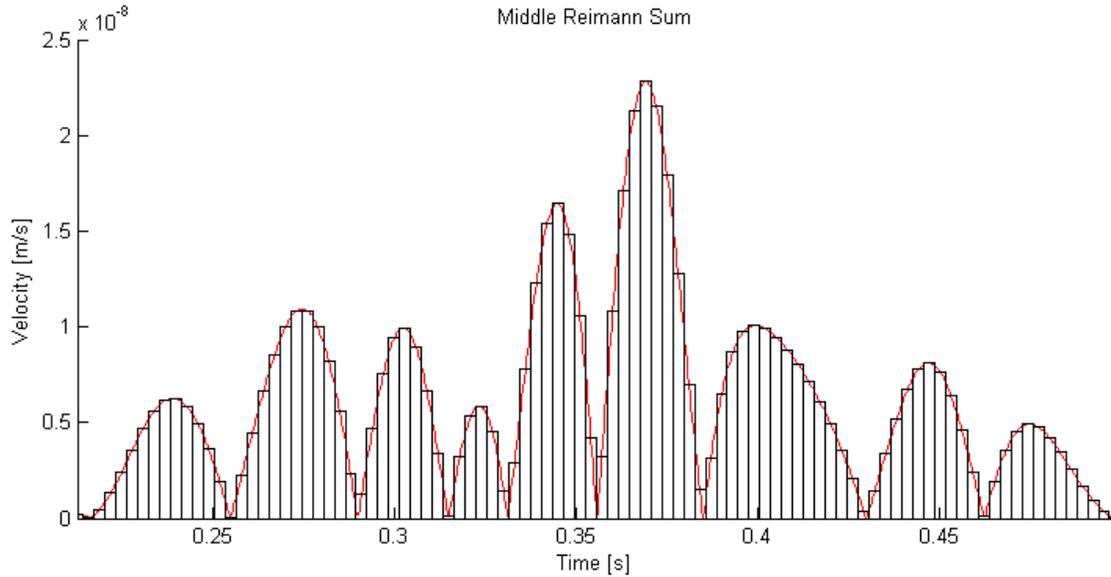


Figure 4.16 Riemann sum with magnitude of velocities

$$Average\ Energy = E_{Avg} = \frac{\Delta t \sum_{i=1}^{nt} \frac{1}{2} m V_i^2}{t_{max}} \quad Eq. 4.29$$

where  $\Delta t$  is the time step length,  $nt$  is the number of time steps, and  $t_{max}$  is the maximum time collected. Inputting that average energy equation into a percent change formula and reducing produces Equation 4.30,

$$Percent\ Change\ in\ Energy = \frac{E_{Avg_1} - E_{Avg_2}}{E_{Avg_1}} * 100 = \sum_{i=1}^{nt} \frac{V_{2,i}^2 - V_{1,i}^2}{V_{1,i}^2} * 100 \quad Eq. 4.30$$

where  $V_2$  represents the velocity with the void present and  $V_1$  is the velocity without it. Voids that are deep are expected to have minimal change in energy.

The 3-D models discussed earlier (Figures 4.3, 4.3, 4.4) are implemented into the 3-D forward model to produce data. The data is collected using a receiver line running along the surface centered in the  $y$  direction. The receiver points are located 1.5 m (5 ft) apart starting at 3.75 m (12 ft) and ending at 38.25 m (126 ft) for a total of 24 receivers. Twenty-nine shots were

produced along the same line starting at 0 m and ending at 42 m (138 ft) with spacing of 1.5 m (5 ft). The shot signature was a Ricker wavelet (Equation 4.31).

$$R(t) = [1 - 2\pi^2 f_c^2 (t - t_0)^2] * \exp[-\pi^2 f_c^2 (t - t_0)^2] \quad \text{Eq. 4.31}$$

where  $f_c$  is the frequency band center and  $t_0$  is the time shift for the wavelet. This produces the stress shown in Figure 4.17.

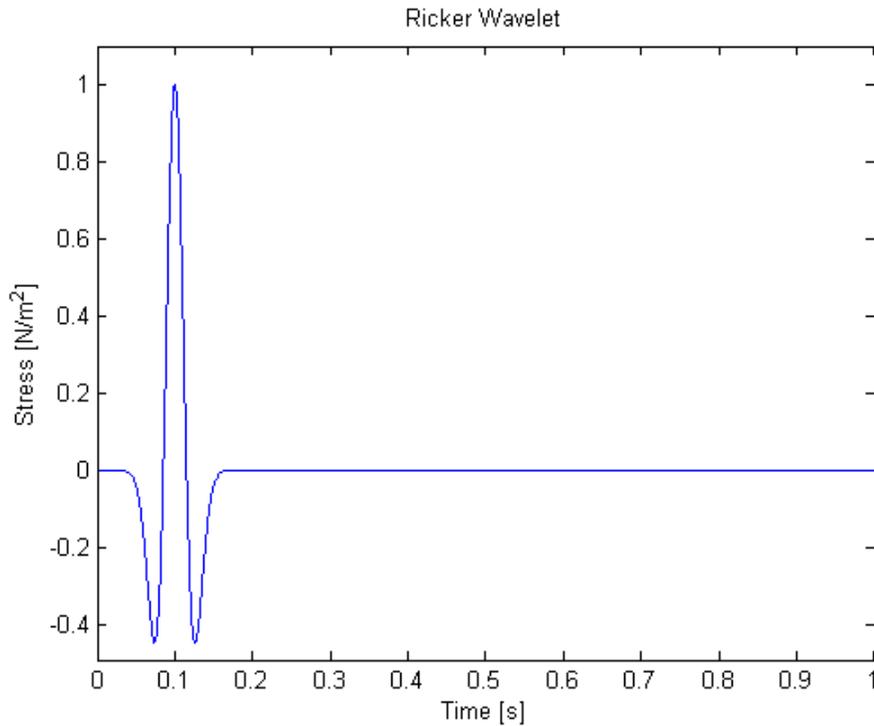


Figure 4.17 Ricker Wavelet

The Ricker wavelet was then used during the calculation of the vertical velocity (Equation 4.4) at the specified shot location. Two sets of data are generated with the central frequency set to 15 Hz and 20 Hz. Computing the percent change in energy using all shots and receivers for each void produces Figure 4.18.

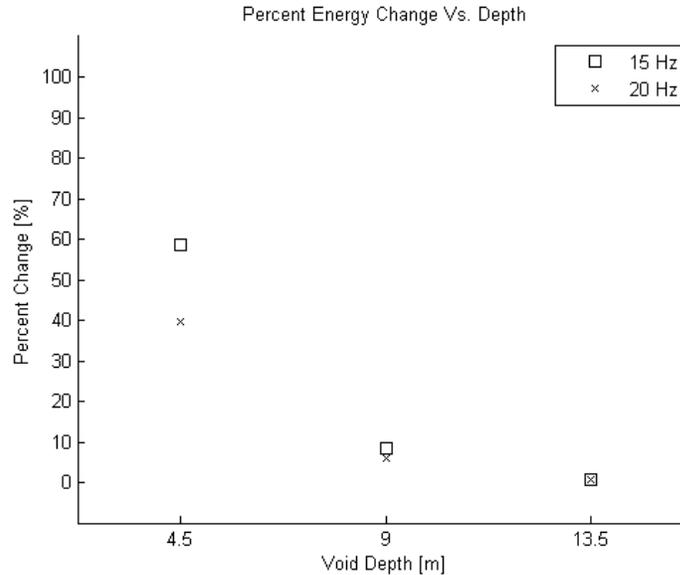


Figure 4.18 Percent change in energy for each void

Evident from Figure 4.18, the shallow void produces a percent change in energy over the no void case of approximately 60% for 15 Hz central frequency and 40% for 20 Hz. The energy change is around 8% for 9 m (30 ft) and 1% for 13.5 m (44 ft) of the void depth for both frequencies. Greater depths are not considered as the change of energy is small and may be less than the energy of surrounding noise (0.5 – 1 %), and thus the inverted results in cases of real data may not be credible.

#### 4.2.4 2-D inversion results of 3-D waveform data

Finally, the 3 synthetic surface data cases associated with the void models (Figures 4.3, 4.3, 4.4) were inverted by the 2-D FWI algorithm for comparison. The initial 2-D model for each case used S-wave velocity of 200 m/s (656 ft/s) at the surface and a linear increase to 600 m/s (1968 ft/s) at the bottom of the model (Figures 4.21b, 4.22b, and 4.23b). The inversion process for each line consists of two inversion runs. For the first inversion run, recorded waveform data is filtered through a frequency range of [0, 0, 12, 15] Hz (Figure 4.19) to produce

a central frequency of 12 Hz. Ten iterations were conducted for each data set to produce the results shown in Figures 4.21c, 4.22c, and 4.23c.

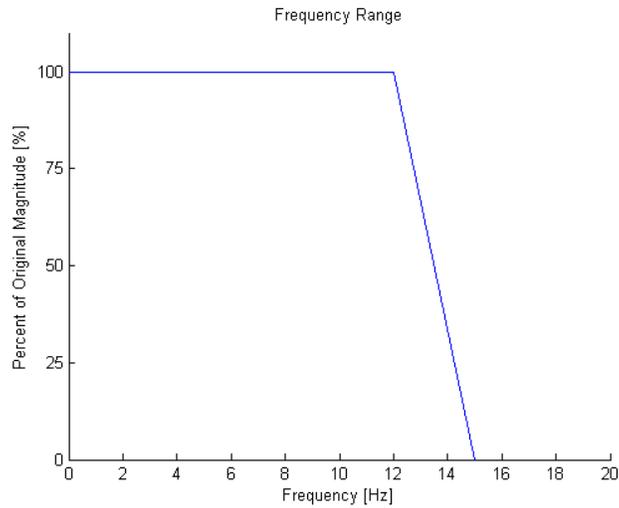


Figure 4.19 Frequency range for the first inversion run

For the second inversion run, the recorded waveform data was filtered through a frequency range of [10, 15, 25, 30] Hz (Figure 4.20) to produce a central frequency of 22 Hz. Each data set was run for 10 iterations for each data set to produce the results shown in Figures 4.21d, 4.22d, and 4.23d.

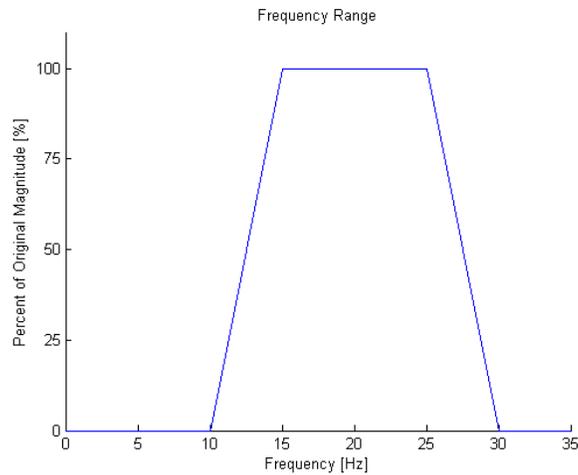


Figure 4.20 Frequency range for the second inversion run

For the void at 4.5-m depth (1 diameter), comparing the true model (Figure 4.21a) against the final inverted model (Figure 4.21d); the shallow void is generally identified in the S-wave

velocity profile. The void location (top of void at 4.5 m (15 ft)) and S-wave (close to 0) velocity are characterized. There exist some shallow artifacts, which are mostly due to the discrepancy between the true 3-D data and 2-D plane data used in the FWI. The P-wave velocity profile is also generally recovered.

For the void at 9-m (30-ft) depth (2 diameters), inverted results shown in Figure 4.22 clearly depict both the layers and the void. Soil layers are accurately characterized. The location, shape, and S-wave (close to 0) velocity of the void have been successfully identified in the S-wave velocity profile. The P-wave velocity profile is also generally recovered.

For the void at 13.5-m (44-ft) depth (3 diameters), inverted results shown in Figure 4.23 reveal an anomaly at the bottom of model in S-wave velocity profile; however, S-wave velocity of the anomaly is about 100 m/s (328 ft/s). This is due to the fact that the reflected energy from the void is relatively small (1%). Shallow artifacts are also observed due to the discrepancy of 2-D and 3-D data sets.

### **4.3 Conclusion**

The 2-D FWI ability to detect voids has been investigated using synthetic 3-D data. A finite 3-D void was embedded at various depths, and 3-D waveform data were generated using shots and receivers on top of the void. Recorded waveforms with the void were compared with those from the model without the void to evaluate the sensitivity of the void (energy change). It was found that if the void was embedded at a depth of more than 3 diameters, the sensitivity of the void is small (less than 1% energy change). For cases of real data, the energy change due to the void may be in the range of energy of surrounding noise (0.5 - 1 %), suggesting the interpreted results may not be credible. Recorded waveforms with the void were also be analyzed by the 2-D FWI, and results suggested the maximum embedded depth at which the void can be detected is about 3 void diameters.

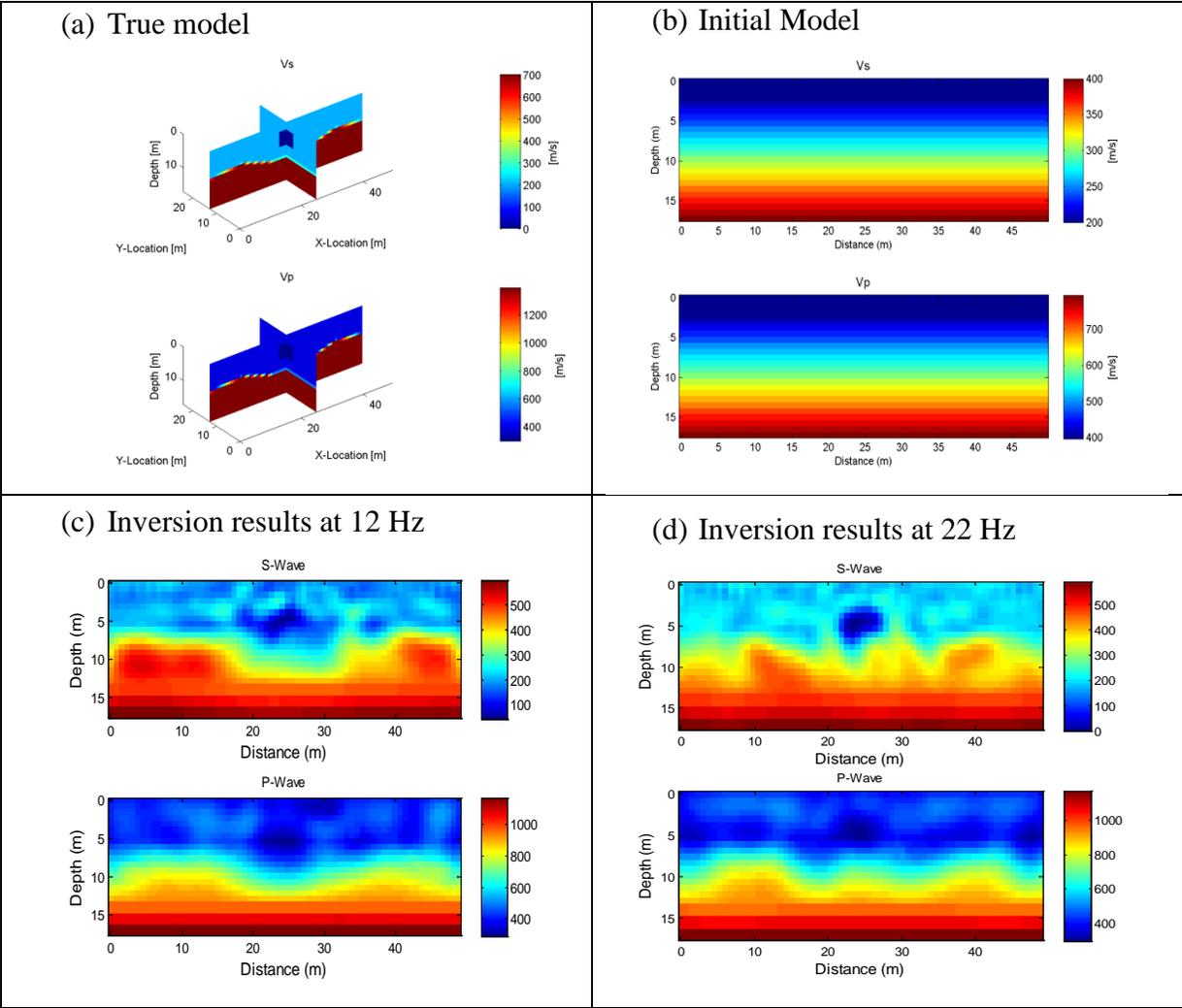


Figure 4.1 Synthetic model of S-wave and P-wave velocities (m/s): (a) 4.5-m depth true mode; (b) initial model; (c) inverted model at 12 Hz; (d) inverted model at 22 Hz.

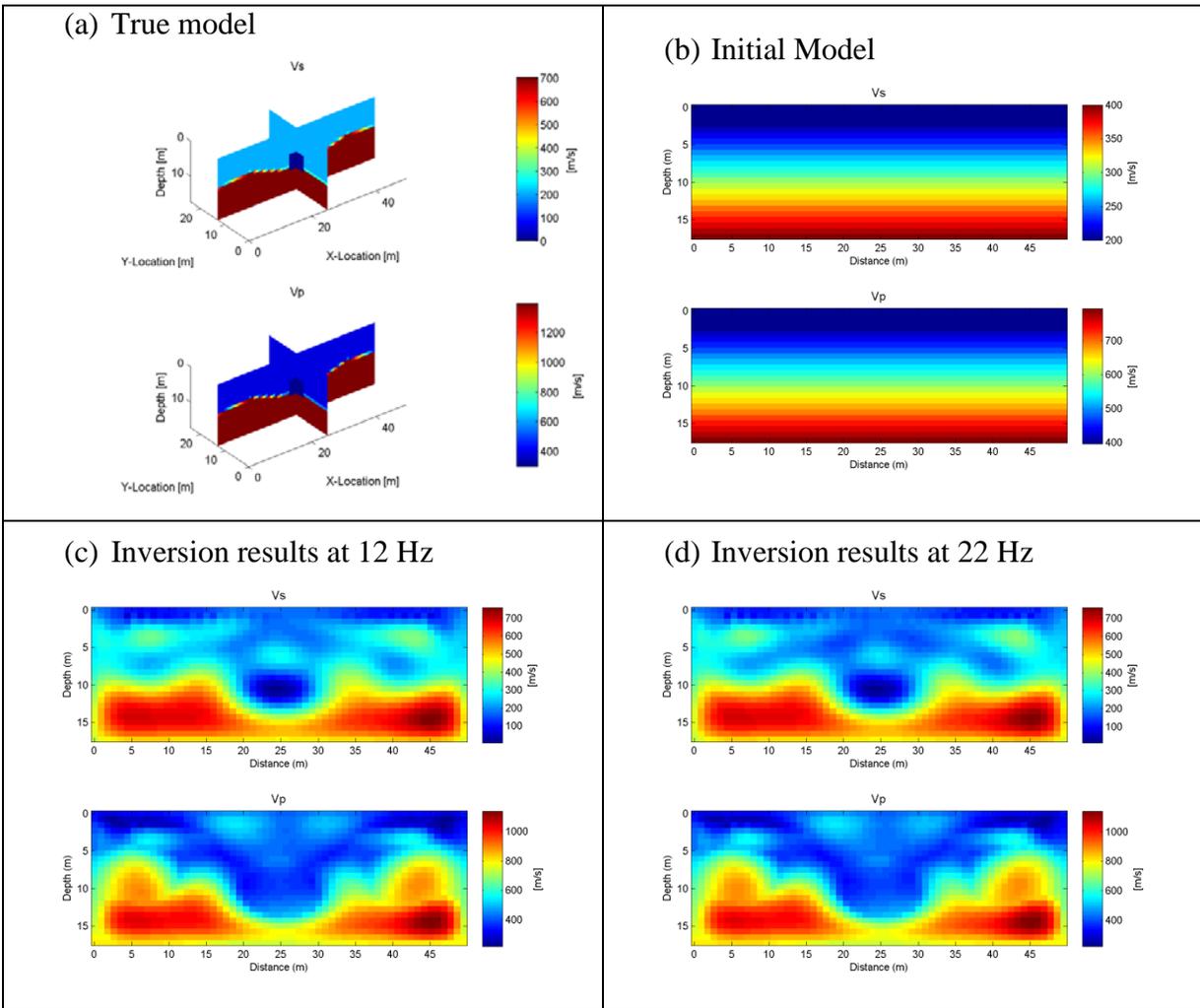


Figure 4.22 Synthetic model of S-wave and P-wave velocities (m/s): (a) 9-m depth true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 22 Hz.

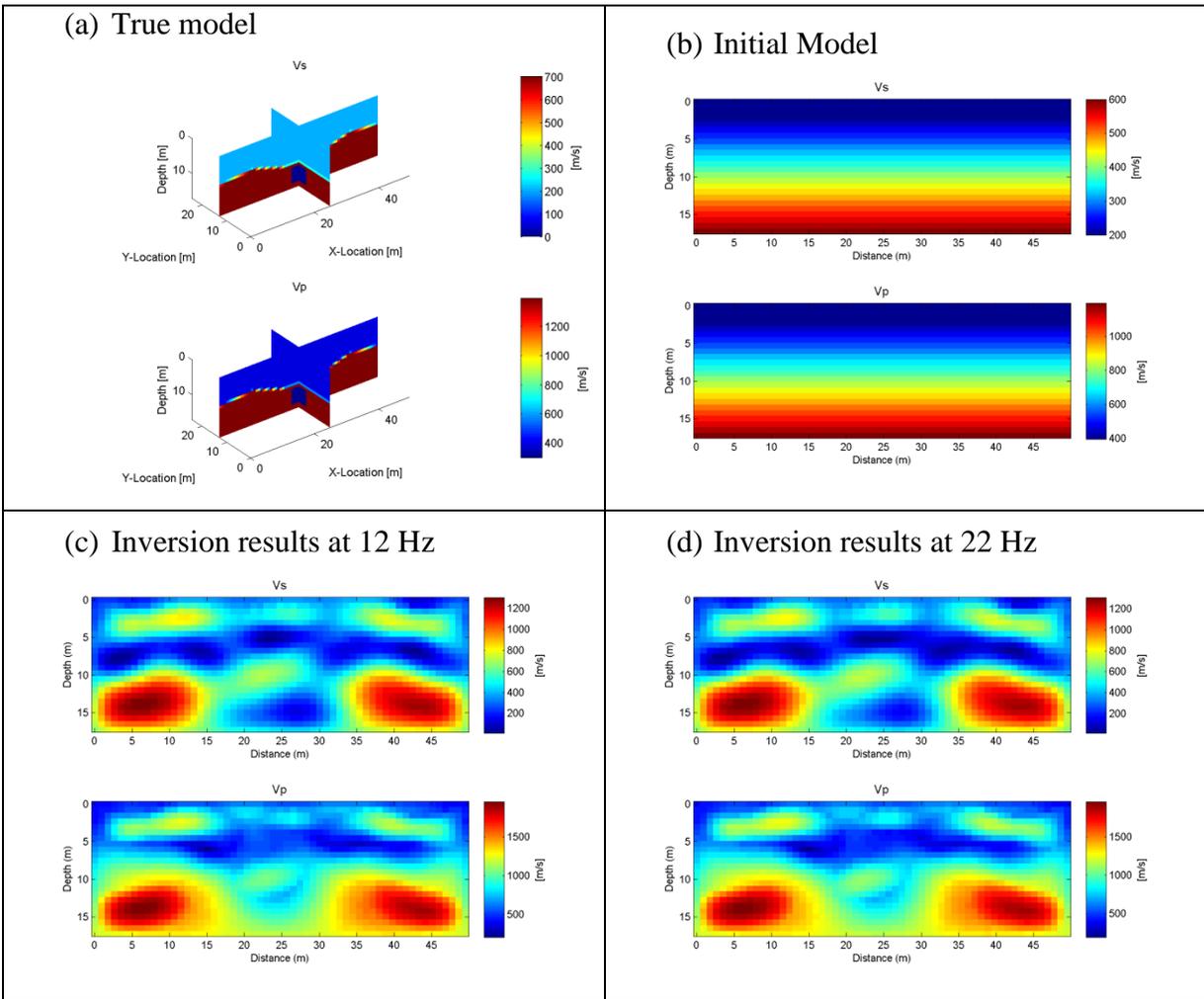


Figure 4.2 Synthetic model of S-wave and P-wave velocities (m/s): (a) 13.5-m depth true model; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 22 Hz.

CHAPTER 5  
EXPLORE 3-D EFFECTS IN THE 2-D FULL WAVEFORM INVERSION ON SYNTHETIC  
MODELS

**5.1 Introduction**

The 2-D inversion program presented in Task 1 was developed to analyze waveform data collected along a line of receivers and hammer strikes, resulting in a 2-D velocity image ( $V_s$  and  $V_p$ ) of the material beneath. In addition, for the analyses, it is assumed that the soil/rock material characteristic along the dashed line, Figure 5.1, is the same everywhere in the x-y plane with the material in the z-y (out of plane, Figure 5.1) plane assumed to vary.

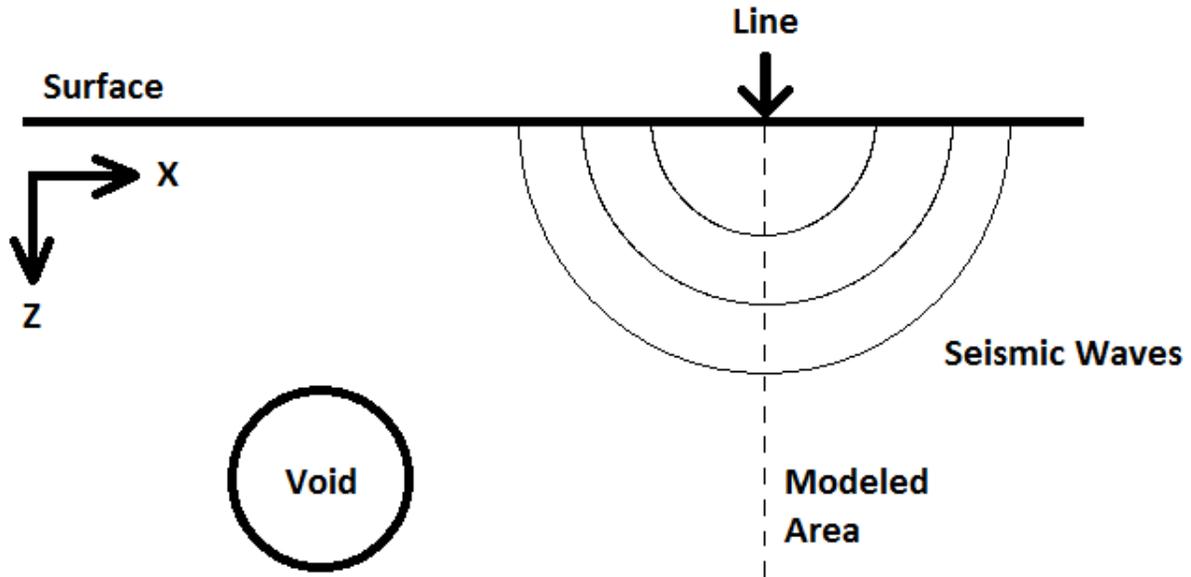


Figure 5.1 Line adjacent to void

Evident, if there exists a void away from the dashed line, Figure 5.1, in the x-y plane, a reflected signal will be generated and it may be collected in the receiver line and generate an inversion void. The effect these reflections have on the results is dependent on the size and depth of the void and its distance from the line. The goal of this task is to investigate the effects of these off-line voids.

## 5.2 Methodology

To investigate the off-line void effects, three synthetic lines were collected from a 3-D model and were inverted using the 2-D FWI program. The first line is directly above the void while the other two are offset three and six meters to allow for comparison (Figure 5.2).

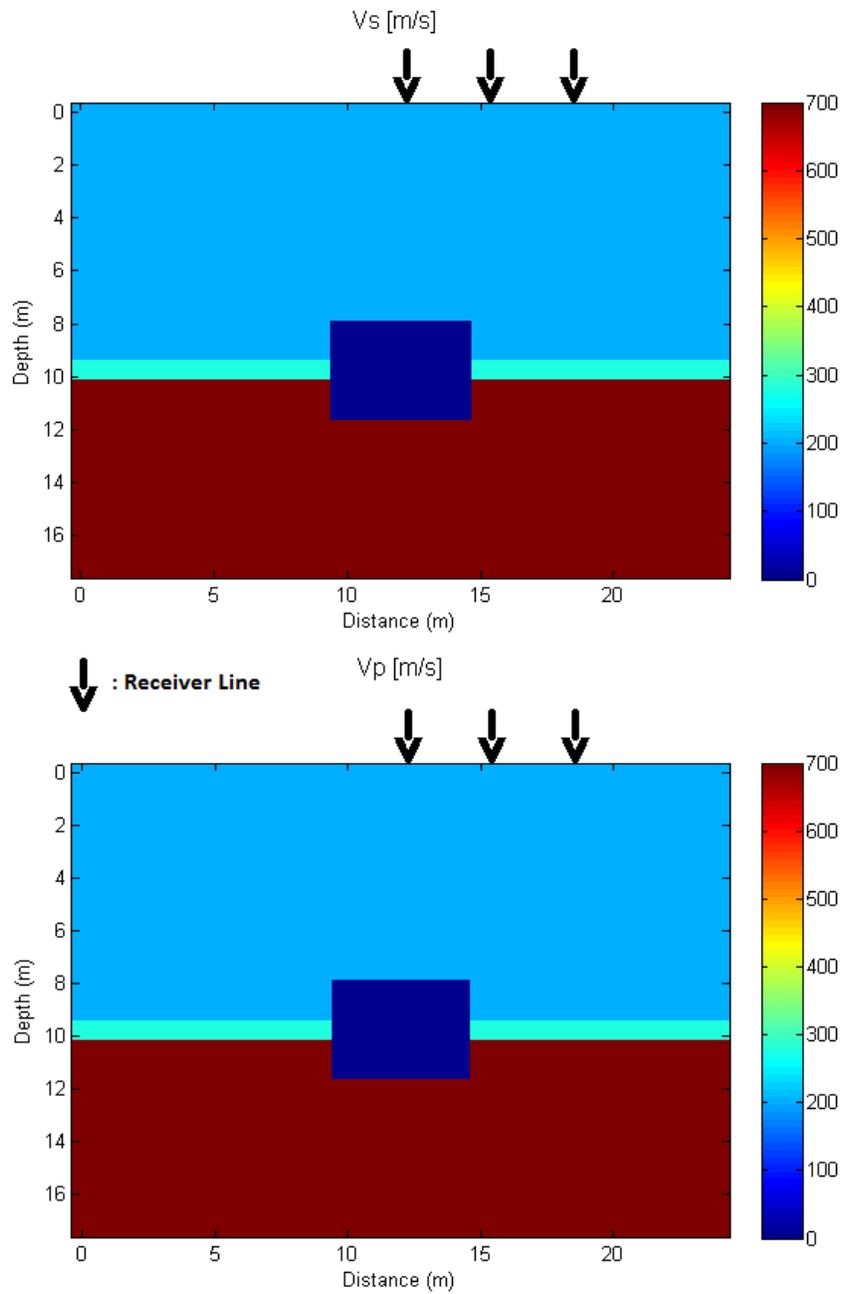


Figure 5.2  $V_s$  and  $V_p$  in the Y-direction

The data were generated using the solution of 3-D elastic wave equations. The synthetic data represents signals that would be collected in the field both directly above and adjacent to the void. The ideal outcome from the 2-D inversion is that the void is observable with the line directly above the void, but not for the other two lines, Figure 5.2.

Similar to Task 3, the synthetic 3-D data were created following the methodology presented in Section 4.2.1. The resulting equations and boundary conditions presented there are implemented into MATLAB to allow for modification of the parameters for each run and minimal computational time. With the model complete the two layer model with a void at the center can be used to generate data for testing.

### **5.3 Results**

A 3-D model, Figure 5.3 and Figure 5.4, is implemented in the 3-D forward model to produce the synthetic data required for 2-D inversion. The model consists of two layers with the following properties: top layer, soil,  $V_s = 200$  m/s (656 ft/s), 2<sup>nd</sup> layer, limestone,  $V_s = 700$  m/s, (2296 ft/s) with a void located at the center of the domain. The void,  $3.75$  m  $\times$   $5.25$  m  $\times$   $5.25$  m (12 ft  $\times$  17 ft  $\times$  17 ft), was characterized with a shear wave velocity ( $V_s$ ) of 0 m/s and a compression wave velocity ( $V_p$ ) of 300 m/s (984 ft/s). Note, the selected soil and rock layering was considered to be representative of Florida conditions.

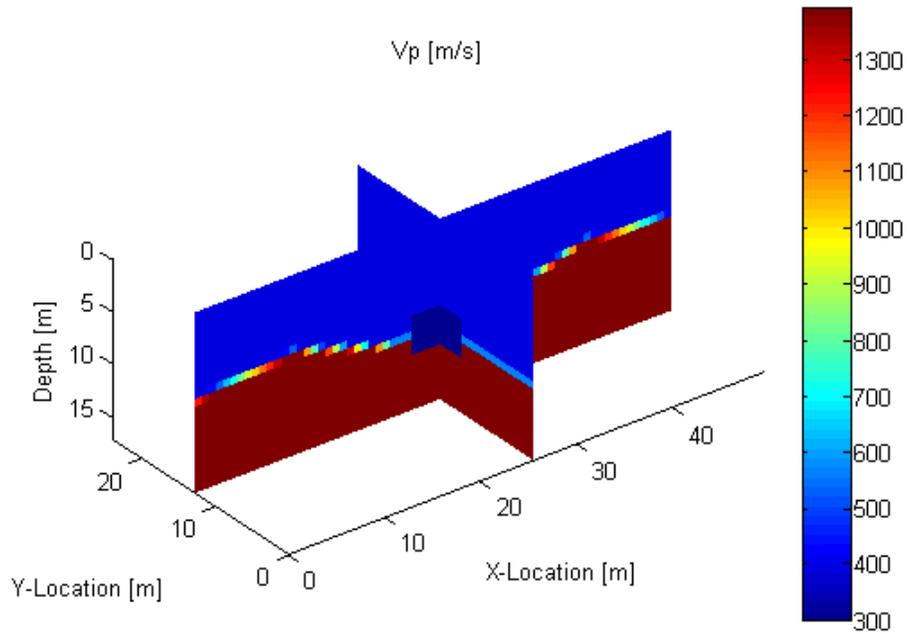


Figure 5.3 3-D  $V_s$  model

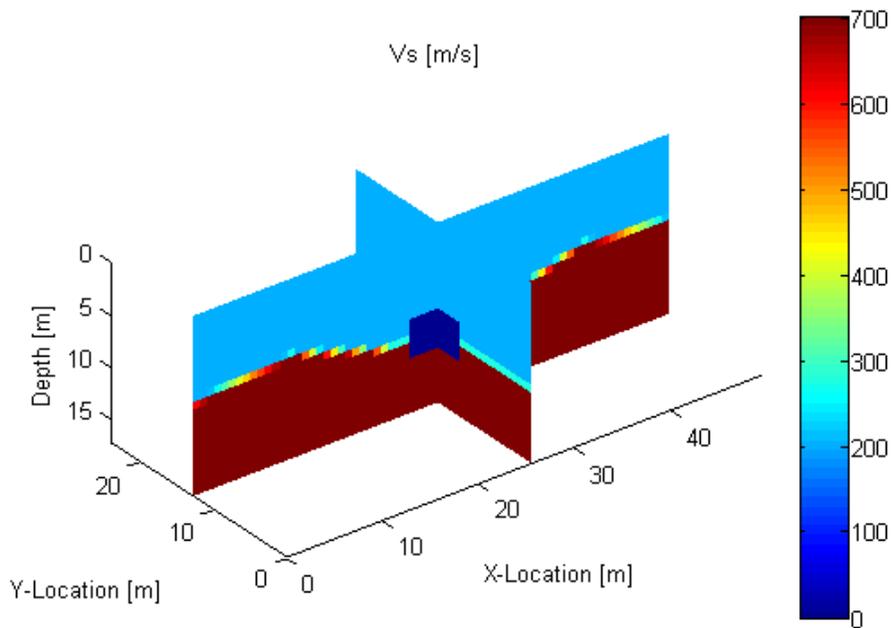


Figure 5.4 3-D  $V_p$  model

The three data collection lines, discussed in Section 5.2, are similarly set up with the exception of their position on the y-axis. The lines consist of 24 receivers spaced every 1.5 m (5 ft) along the x-axis from station 7.5 to 42 m (25 to 138 ft), with 29 shots at 1.5 m (5 ft) spacing

starting from 3.75 to 45.75 m (12 to 150 ft) on the ground surface. Line one is centered over the void while the two remaining lines are offset three and six meters in the y-direction.

Each shot is perturbed by altering the vertical normal stress,  $\sigma_{zz}$ , at the source. It is perturbed using a Ricker wavelet,

$$R(t) = [1 - 2\pi^2 f_c^2 (t - t_0)^2] * \exp[-\pi^2 f_c^2 (t - t_0)^2] \tag{1}$$

where  $f_c$  is the frequency band center and  $t_0$  is the time shift for the wavelet. This produces a stress shown in Figure 5.5.

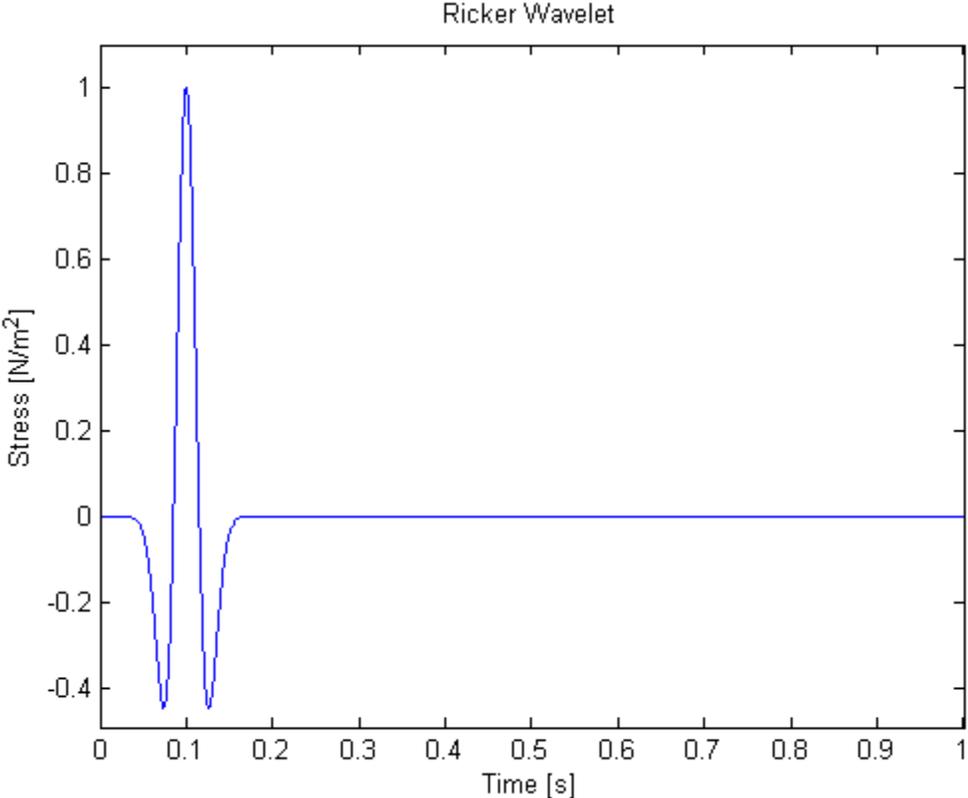


Figure 5.5 Ricker Wavelet

This is then applied to the specific grid point representing the calculation of the vertical velocity (Equation 3) at the source location. For each case, two sets of data are generated with the central frequency set to 15 Hz and 20 Hz with a time shift of 0.1 s. Figure 5.6 is an example of the wave field generated from a shot at the end of a receiver line over the void. The shot is

perturbed from the right end of the line meaning that the wave reflected off of the void is observable in the data collected from receivers earlier in the line.

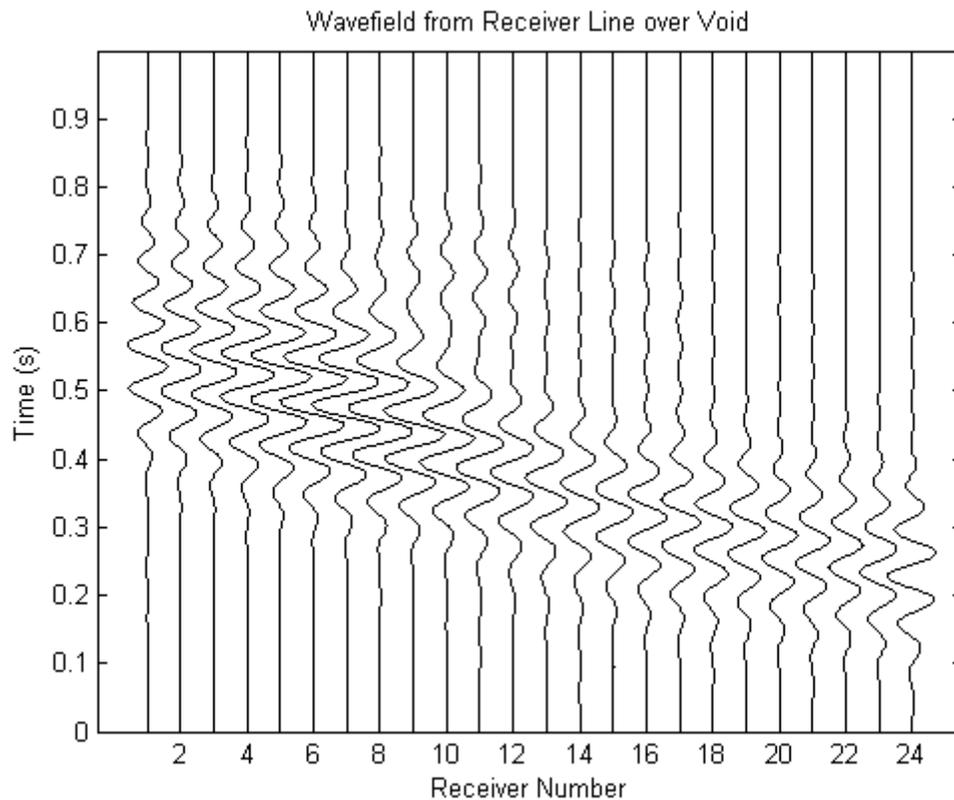


Figure 5.6 Wavefield created from receiver line over a void

The next step was to invert the data set using the 2-D algorithm for each line consisting of 29 shots with 24 receivers per shot for each data set. The initial model for each analysis was with a shear wave velocity of 200 m/s (656 ft/s) at the surface which linearly increased to 400 m/s (1312 ft/s) at the bottom of the model (Figures 5.9b, 5.11b, 5.13b).

The inversion process for each line consists of two sets of iterations. The first used the data generated with a central frequency of 12 Hz and a frequency range of [0, 0, 12, 15] Hz (Figure 5.7) to produce a central frequency of 12 Hz.

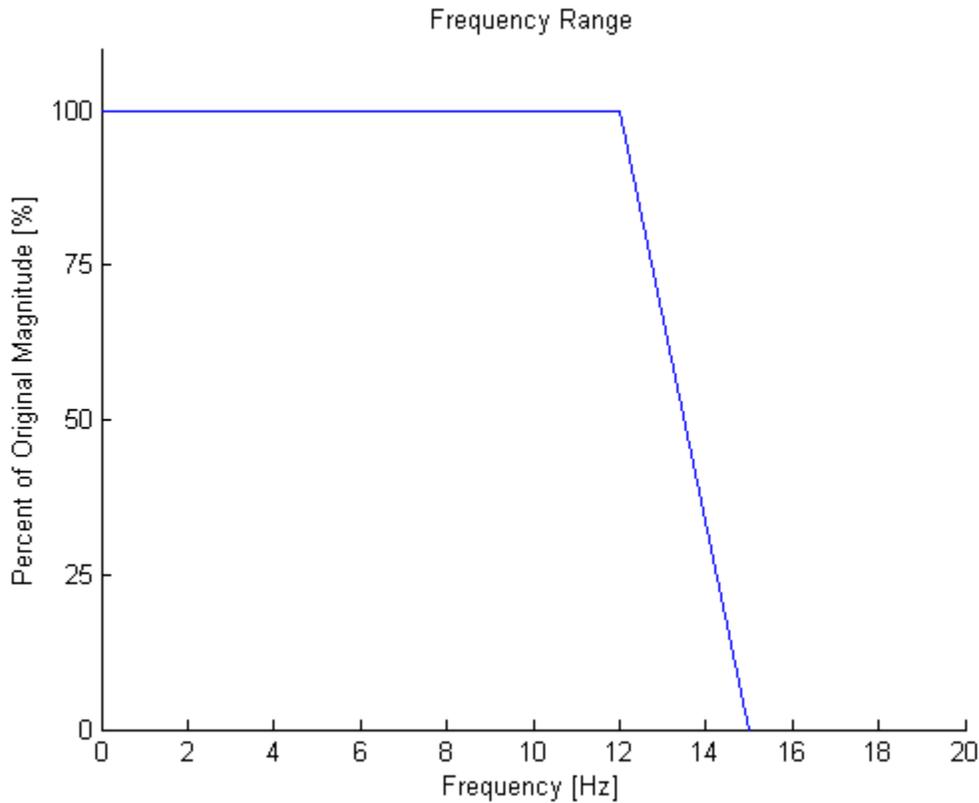


Figure 5.7 Frequency range for first iteration set

Note, the data was left mostly unaltered with a windowing range of 0.5 s both before and after the peak due to the lack of noise in the synthetic data. Each data set was run for ten iterations to produce the results found in Figures 5.9c, 5.11c, 5.13c.

The second set of iterations used the data generated with a central frequency of 20 Hz at a frequency range of [10 15 25 30] Hz (Figure 5.8) to produce a central frequency of 23 Hz. Once again, the data was windowed with a range of 0.5 s before and after the peak. Ten iterations were also conducted for each inversion to obtain a good waveform match for each line (Figures 5.10, 5.12, 5.14). The inversion results are shown in Figures 5.9d, 5.11d, 5.13d.

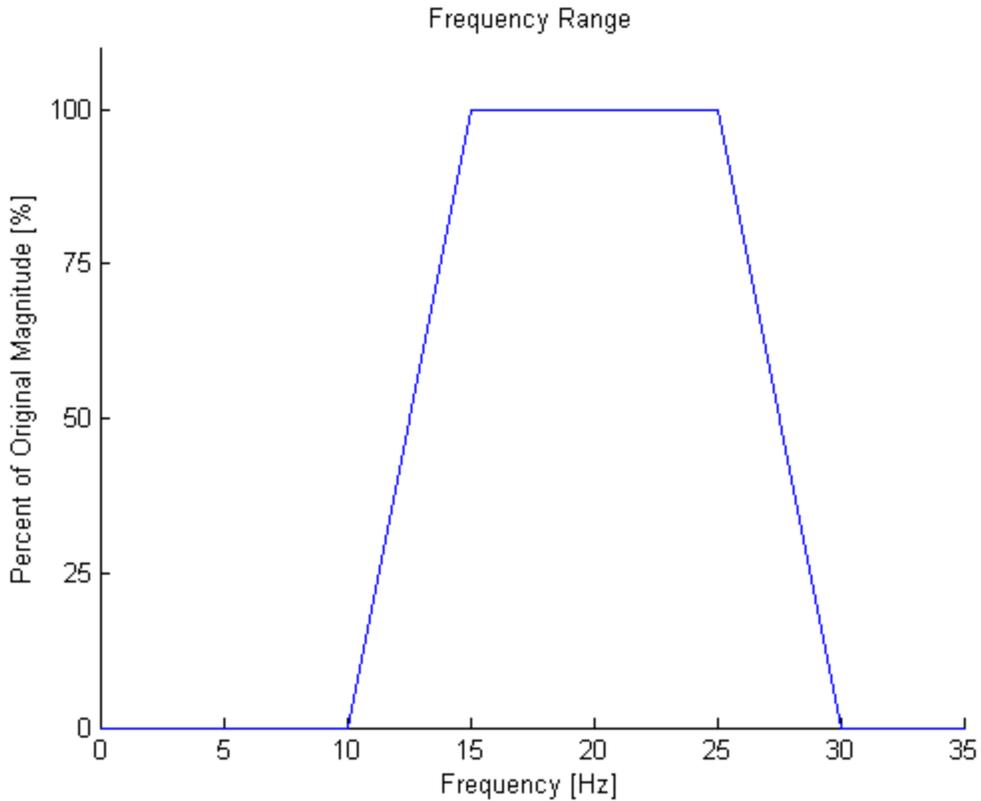


Figure 5.8 Frequency range for second iteration set

From a comparison of the true model for line 1 (Figure 5.9a) against the final inverted model (Figure 5.9d); it is evident that soil layers are accurately characterized. Especially the presence, location, shape, and S-wave (close to 0) velocity of the void have been successfully identified in the S-wave velocity profile. The P-wave velocity profile is also generally recovered. The second line, Figure 5.11, also clearly depicts both the layers and the void despite the fact that the void is not directly under the line. This is due to the close proximity of the void to the line (test line at the edge of the void). For cases of real experimental data, if inverted results show an anomaly with very low S-wave velocity ( $< 50$  m/s (164 ft/s)) at a shallow depth ( $< 3$  diameters) similar to those in Figure 5.9d or Figure 5.11d, it is mostly likely the anomaly is a void below or very close to the test line.

### **5.3.1 Distortion and Location of Off-center Void**

Of interest is if the seismic line which passes at least one diameter away from the void can identify the existence of a void. Shown in Figure 5.13 is the case of line 3, which is 1 diameter from the anomaly. Figure 5.13d successfully shows the overall correct wave velocities for the two layers. Evident, from Figure 5.13d,  $V_p$  and, to some extent, the  $V_s$  do show some signs of a void (6 m (20 ft) one diameter away from the center of the void). However, comparison of Figures 5.9d vs. 5.13d shows that the offline void seems to be distorted both vertically and horizontally. In addition, the figure shows that the inverted profile has smeared the low-velocity zone (Figure 5.13d). Specifically, the velocity ( $V_s$  and  $V_p$ ) of the void has increased whereas the surrounding soil has decreased. The smearing appears to be approximately  $\frac{1}{2}$  diameter to both sides of the actual void (Figure 5.9d). For cases of real experimental data, if inverted results show a smeared low-velocity zone similar to Figure 5.13d, it is recommended that additional test lines adjacent to both sides of the existing line be performed to verify the lateral extent of the void. For instance, if a line was obtained  $\frac{1}{2}$  diameter to the left in Figure 5.11d would be obtained, which clearly shows the lateral extent of void and no smearing. Verification of the synthetic 2-D and 3-D off-center simulation, distortion, etc. will be carried out experimentally in the field in Task 5.

### **5.4 Conclusion**

The application of the 2-D Full Waveform Inversion program discussed in Task 1 allows for only a slice of the ground to be viewed. With only a line of data gathered for each inversion, it is difficult to predict the effects of anomalies, such as a void, that are adjacent to the line. This task used a 3-D forward model to produce synthetic data with known ground properties. The synthetic data was collected at lines on the surface directly above the void and at the edge of void (half a diameter from the void center) and one diameter from the void. The inverted results

showed that (1) the void is well characterized if collected data is on top of the void (depth of void should be less than 3 diameters as discussed in Task 3, (2) the void, however, can be also identified in an inverted profile if the test line is near the void edge, and (3) the void effect is minimal if the test line is at least one diameter from the void.

The 3-D effects found in this task are assessed by real experimental data in Task 5 presented in the next chapter. Seismic data (multiple lines) from two sites (US 441 and Newberry) have been collected with identification of physical voids. The parallel seismic lines conducted both on top of voids and at various offsets allow assessing the 3-D effects. The data analyses are undertaken to identify the shift (distance from seismic line to center of void), shape and size of sinkhole voids, which are compared with CPT /SPT results at the sites.

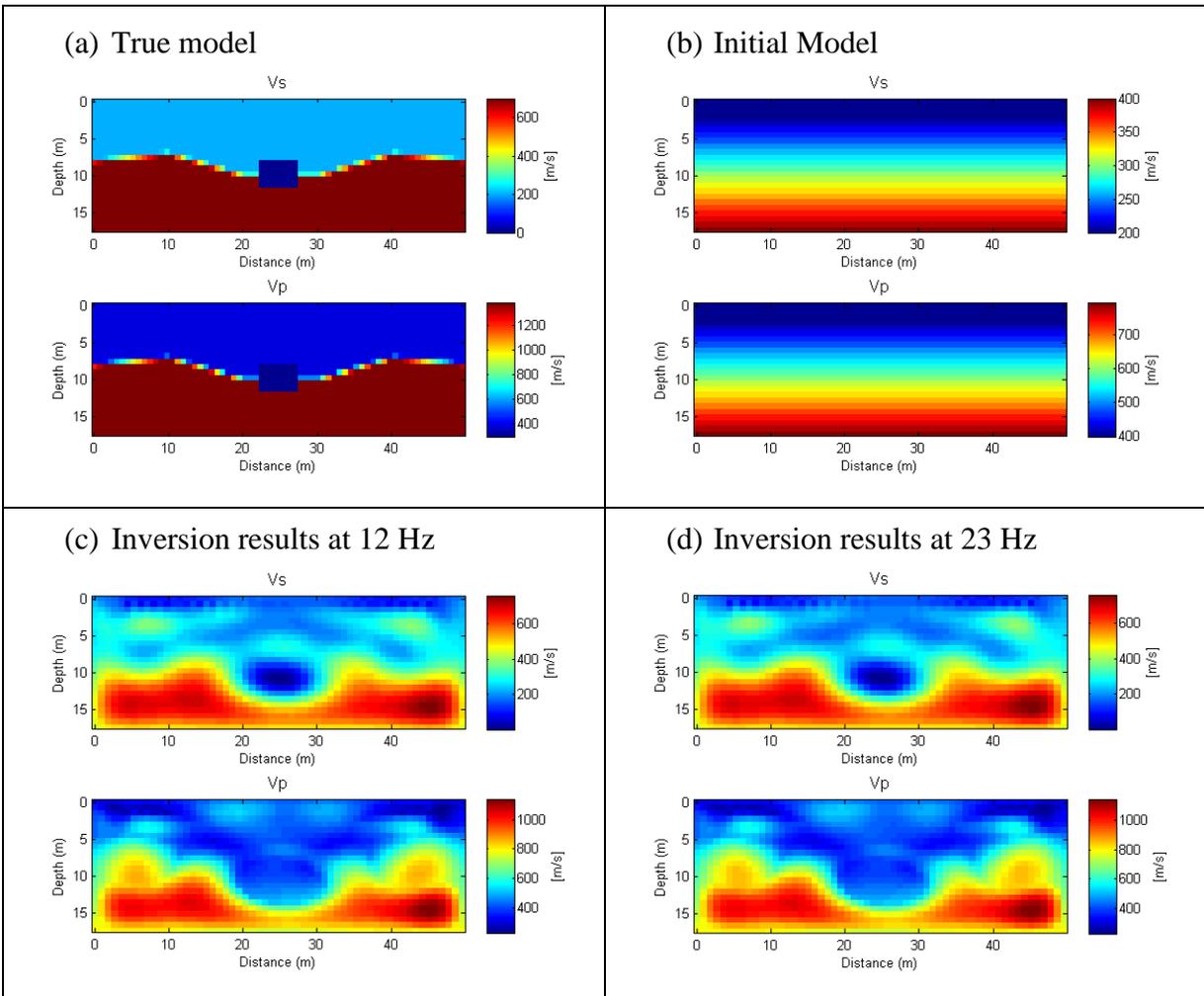


Figure 5.9 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 1 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz.

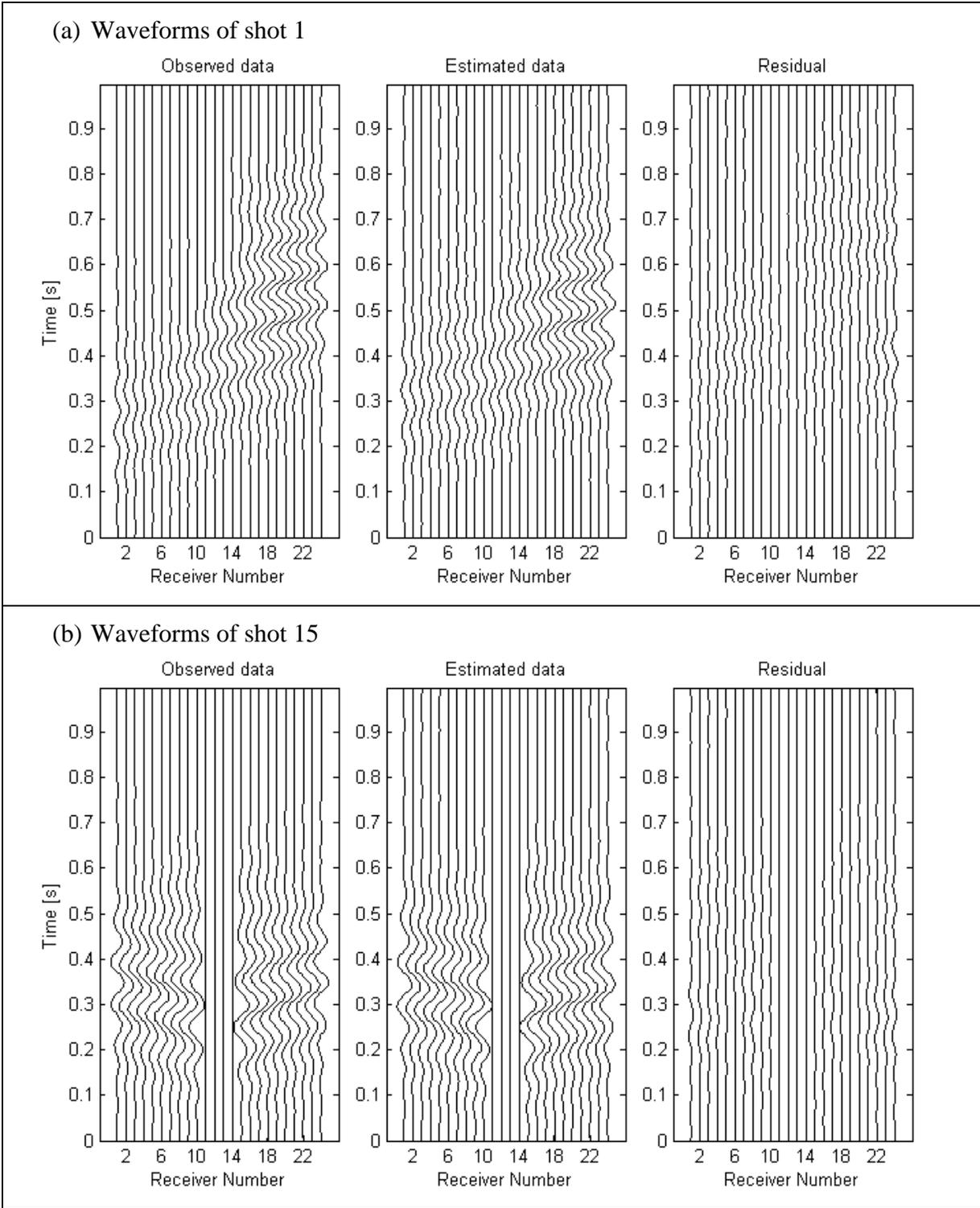


Figure 5.10 Line 1 observed, estimated, and residual waveforms: (a) Shot 1 and (b) Shot 15

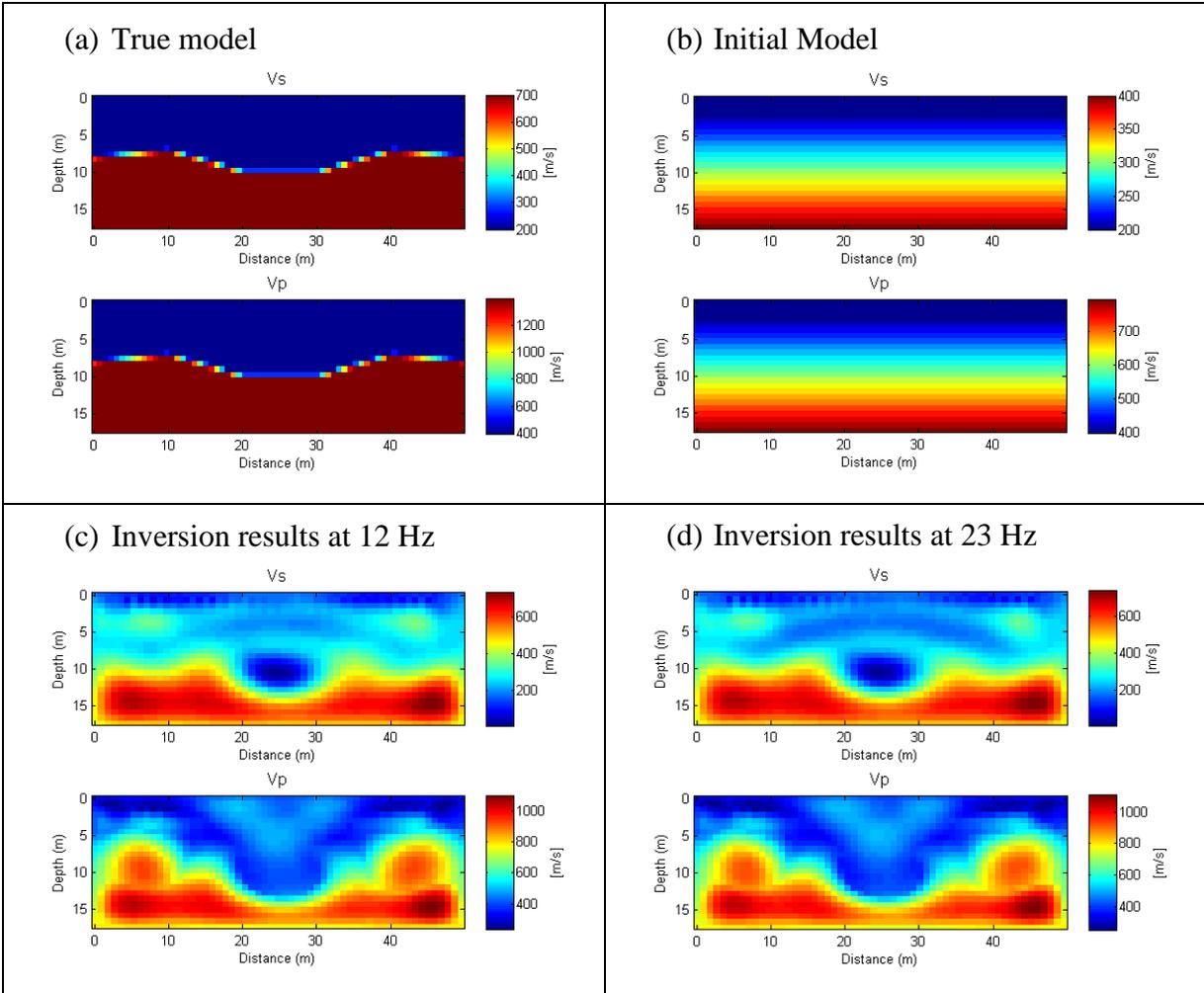


Figure 5.11 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 2 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz.

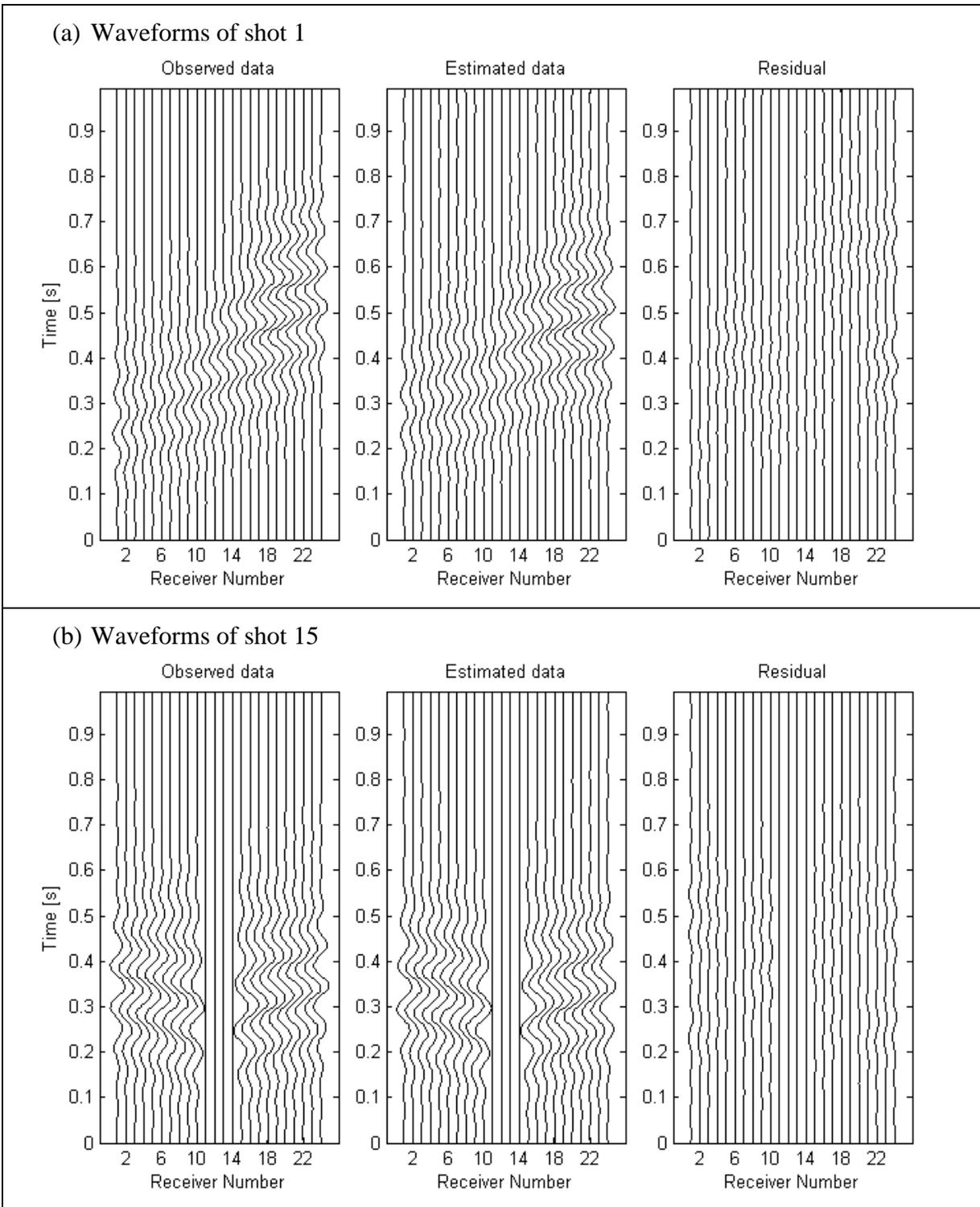


Figure 5.12 Line 2 observed, estimated, and residual waveforms: (a) shot 1 and (b) shot 15

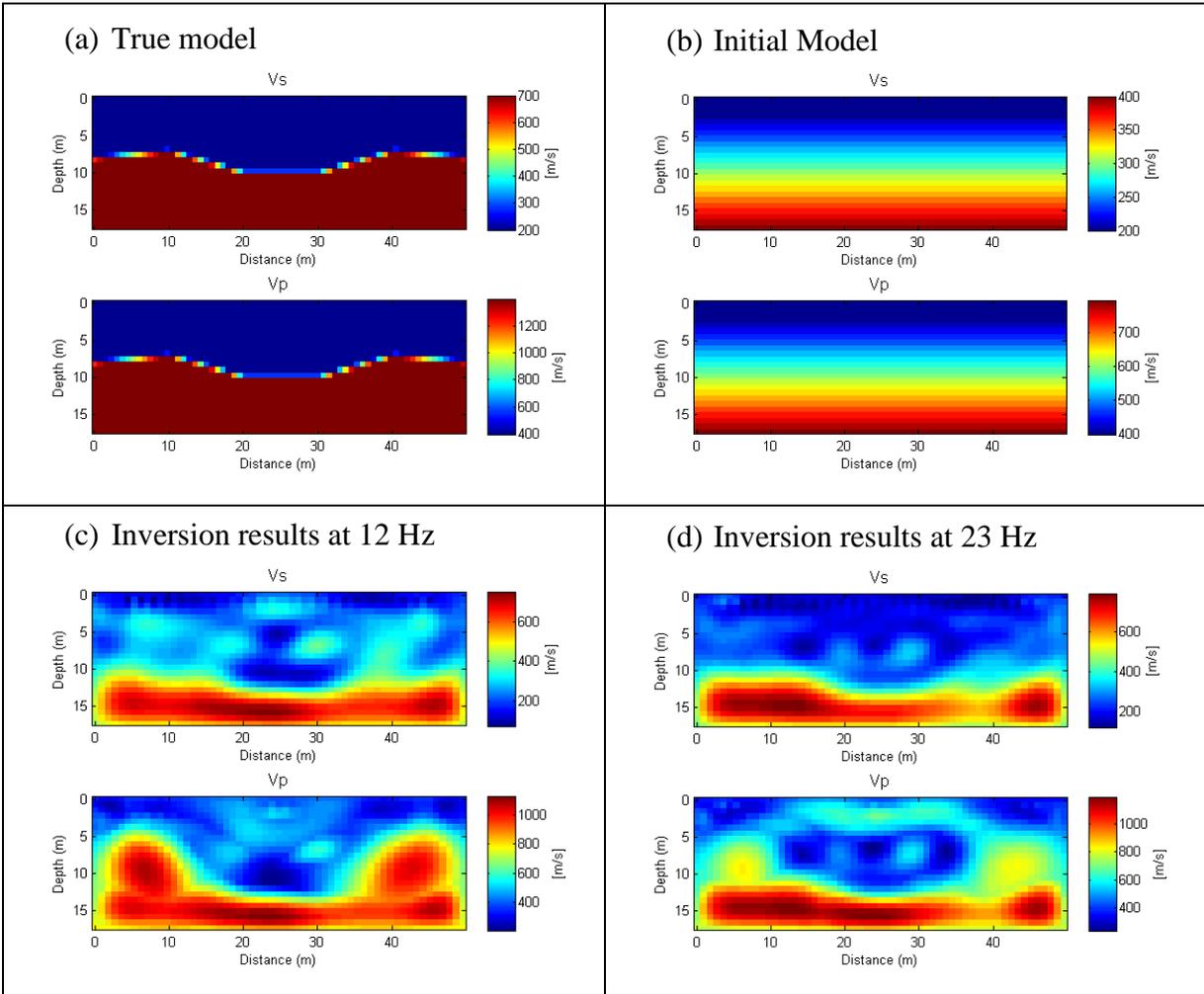


Figure 5.13 Synthetic model of S-wave and P-wave velocities (m/s): (a) line 3 true mode; (b) initial model; (c) inverted model at 12 Hz; and (d) inverted model at 23 Hz.

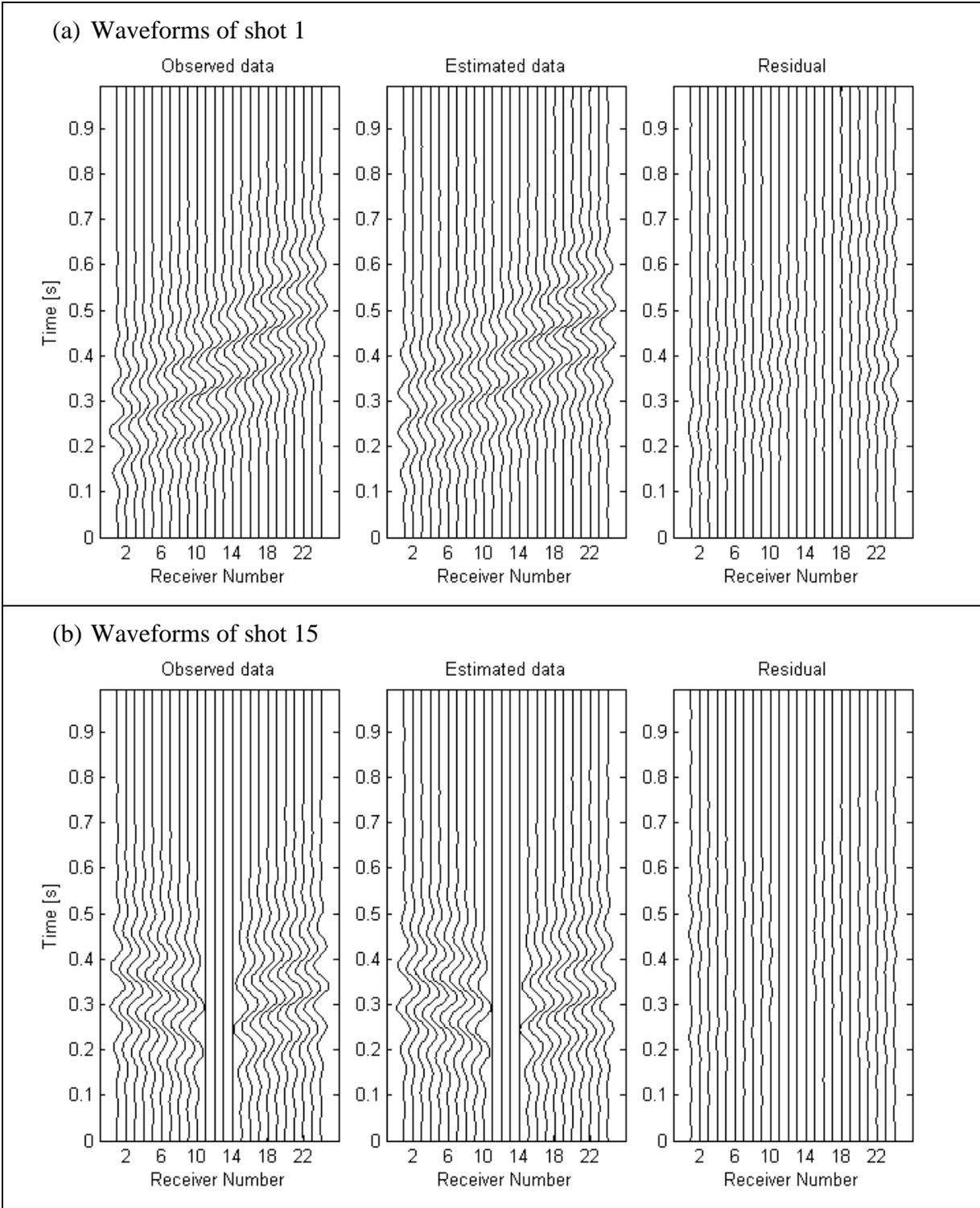


Figure 5.14 Line 3 observed, estimated, and residual waveforms: (a) shot 1 and (b) shot 15

## CHAPTER 6 TEST THE DEVELOPED FWI SOFTWARE ON FULL SCALE TEST SITES

### 6.1 Introduction

The goal of Task 5 is to test the developed FWI software (Task 2), and verify the 3-D effect of off-line voids (Task 4) on full scale field experiments. The software has been applied on 5 test sites with various subsurface conditions, including US 441, Newberry, Gainesville, Tallahassee, and Kanapaha sites in Florida. These sites were selected based on histories of sinkholes (e.g., past activity), available invasive tests (CPT and SPT), and results from the Phase I study.

For each test site, multiple seismic test lines were conducted. Test lines were at least 36 m (120 ft) long for separation of P-wave and S-wave groups in time domain at far-field geophones, which are important for extraction of both P-wave and S-wave velocities. Seismic energy was generated using either a sledgehammer (67 to 90 kN (15 to 20 lbs) or a propelled energy generator (PEG-40 Kg model)). Different sizes of sledgehammer, heights of drop for the PEG, and sizes of impact plates were tested to obtain good signals at a large frequency range from 5 to 50 Hz needed for waveform analysis.

All collected data were analyzed by the developed software. The waveform analysis includes (1) data conditioning by filtering, windowing and removing poor channels, (2) initial model generating, and (3) model updating (inversion) to obtain 2-D subsurface profiles (P-wave and S-wave velocities). All results presented were obtained within about 30 minutes on a standard desktop computer. The seismic results are compared to invasive tests (CPT and SPT) for assessment of the software and verification of off-line void effects. Field experiments and results from the five test sites are as follows.

## 6.2 US 441 Site

The test site is on US 441 Highway, in Marion County, Florida, USA. Seismic testing was conducted to assess a roadway segment that had a repaired sinkhole. The sinkhole opened in the highway in 2011 and was subsequently repaired by placement and compaction of sand and gravels to bring the roadway back to its original elevation. Unfortunately, the subsidence had continued, suggesting that the void was not completely filled. To investigate the size and extent of the anomaly, experimental data was collected on asphalt pavement using a new 24-channel land-streamer and a propelled energy generator (PEG) as shown in Figure 6.1a. The main advantage of using the land-streamer system is the elimination of the need of coupling the geophones to the roadway and the movement of the whole test system quickly for additional test lines. The land-streamer included 24-4.5 Hz vertical geophones equally spaced at 1.5 m (5 ft) spacing. The PEG was attached to a truck, which moved along the geophone array to the appropriate shot locations.

Two parallel test lines 2.5 m (8 ft) apart were conducted. The line 1 was on the road shoulder, and the line 2 was on the top of the sinkhole center as shown in Figure 6.1b. For each line, data was recorded by the land-streamer for 25 shots at 1.5 m (5 ft) spacing, for the total test length of 36 m (120 ft).

Data from both lines were analyzed by the FWI software. The analysis began with data conditioning by filtering, windowing and removing poor channels. Then the initial model was generated based on spectral analysis and the best model search as detailed in Task 1. The initial profile was selected as a linear increasing S-wave velocity from 250 m/s (820 ft/s) at the surface to 400 m/s (1312 ft/s) to a depth of 18 m (60 ft) over a length of 36 m (120 ft). The initial P-wave velocity for the domain was calculated from the S-wave velocities assuming that the initial Poisson's ratio throughout the domain was 0.3.

(a)



(b)



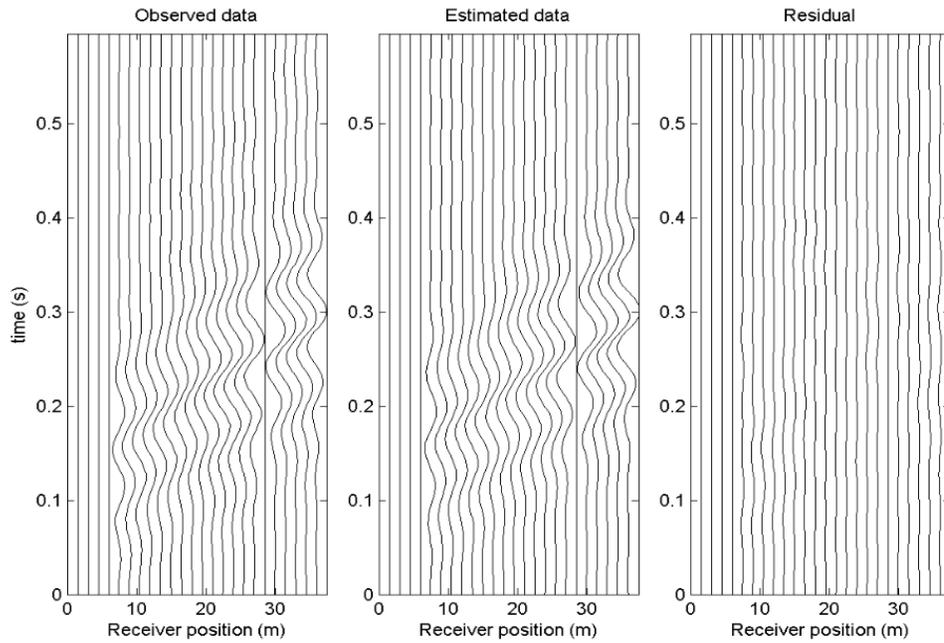
Figure 6.1 US 441 Site: (a) land-streamer and propelled energy generator source, b) repaired sinkhole location

For each test line, two separate inversion runs were performed on the filtered data sets using two central observed frequency ranges: 12 Hz and 16 Hz. Ricker wavelets having central frequencies of 12 and 16 Hz were used to estimate the active source (Propelled Energy Generator) for the forward modeling of each analysis. The first run began with the lower frequency range (central frequency of 12 Hz) using the linearly increasing velocity initial model, and the subsequent run for the central frequency of 16 Hz was completed using the inverted result of the lower frequency as the initial model. During inversion, both S-wave and P-wave velocities of all cells were updated independently, and each run was stopped after 20 iterations.

As an example, the observed and estimated waveforms and residuals (difference between observed and estimated) associated with the final inverted model are shown in Figure 6.2 for the shots at distance 0 and 16.5 m (0 ft and 55 ft) of test line 1 (on shoulder). Apparently, the observed and estimated waveforms are very similar, and the residuals are small for entire test length.

Seismic results of Lines 1 and 2 are shown in Figure 6.3a and 6.3b, respectively. The results of the 2 lines are very similar. Both include a soft soil layer at shallow depths from 0 to 6 m (0 to 20 ft), followed by a stiffer layer with embedded low-velocity zones. For test line 2 on the top of sinkhole center (Figure 6.3 b), there exists a very low-velocity anomaly ( $V_s < 50$  m/s (164 ft/s)) at the repaired sinkhole location (distance 16 m (52 ft)), suggesting the void may still exist or it is filled by soft raveled soils. The estimated size of the void is about 3 m (10 ft) diameter.

(a) Shot 1 at station 0 m (0 ft)



(b) Shot 12 at station 16.5 m (55 ft)

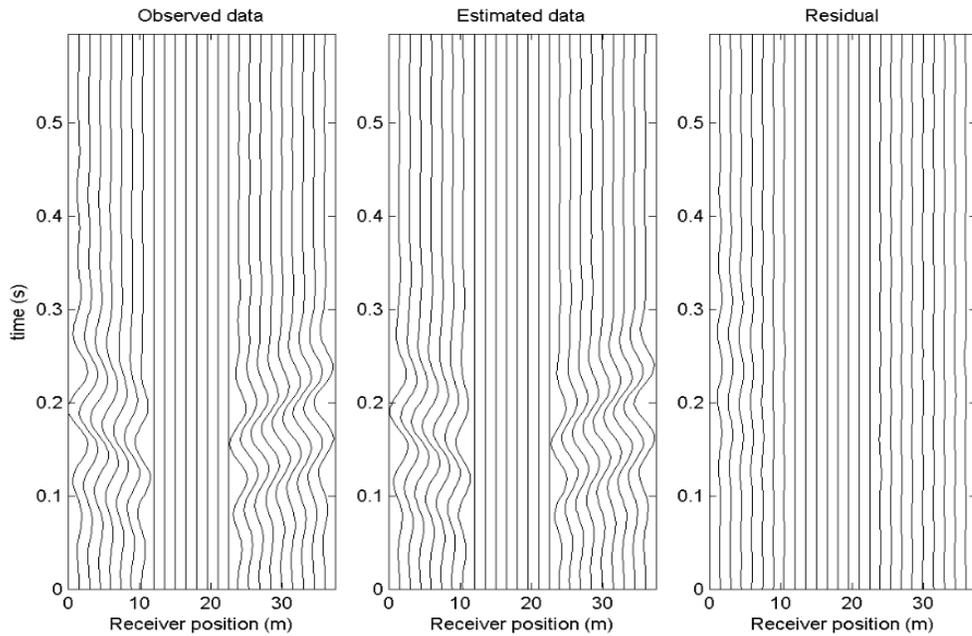


Figure 6.2 US 441 Site, line 1: comparison between observed and estimated data for shots at 0 and 16.5 m (0 ft and 55 ft)

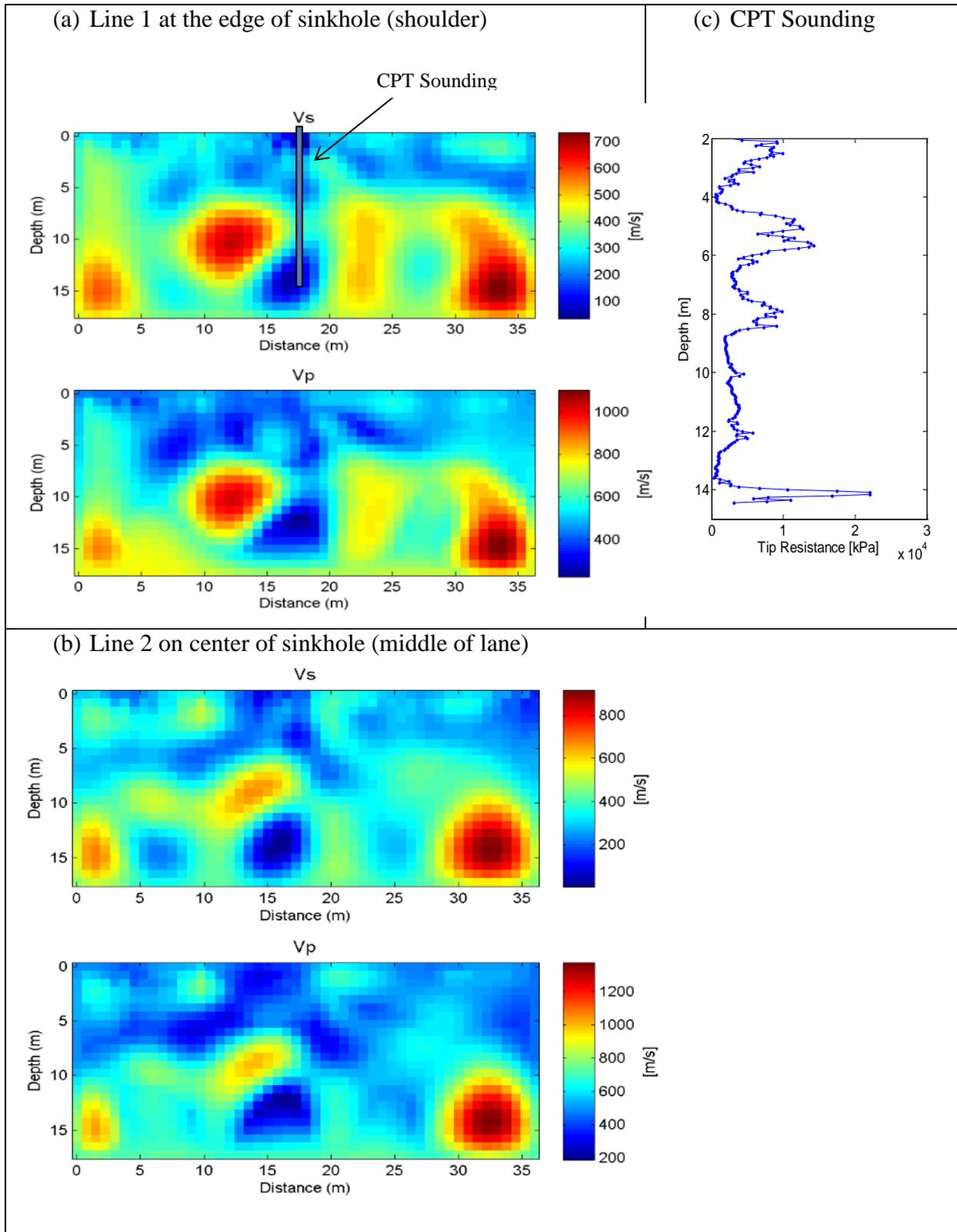


Figure 6.3 US 441 site: S-wave and P-wave velocities: (a) line 1 on shoulder, (b) line 2 at middle of lane, and (c) CPT sounding

Line 1 is 2.5 m (8 ft) from Line 2, and about one-diameter from the center of the void. Based the synthetic study on 3-D effect of off-line void, the influence of the void on the seismic result is not significant. Line 1 result (Figure 6.3 a) represents the true soil profile underneath the test line. Shown in Figure 6.3 c is the CPT sounding at Line 1, on the road shoulder. The cone tip resistance seems consistent with the seismic results, including a stiff zone from 5 to 9-m depth (17 to 30-ft), underlain by softer zone from 9 to 14-m depth (30 to 47-ft). However, the soft zone in the seismic result seems to be deeper and bigger at depth from 10 to 16 m (33 to 53 ft). This is believed to be due to 3-D effect of the void and the weathered limestone, and the velocity gradient in the characterized profile (smoothness generated by the FWI). For instance, both an off center void and seismic line were modeled as plane strain (i.e., 2D) resulting in distorted anomalies; however with 3D limestone (e.g., pinnacles) and voids, seismic waves may also be refracted and reflected resulting in slower arrival times resulting in deeper or translated anomalies

### **6.3 Newberry Site**

This site is a retention pond (Figure 6.4 a) in Newberry, Florida. The site consists of medium dense, fine sand and silt underlain by highly variable limestone; the top of limestone varies from 2 to 10 m (7 to 33 m) in depth. The site was divided into 25 north-south survey lines equally spaced a distance of 3 m (10 ft) apart and labeled A through Y (Figure 6.b a). Sixteen test lines (10 in Phase I, and 6 in Phase II) were conducted along lines K to T for both the north and south portions. Each line was conducted using a linear array of 24 4.5-Hz vertical geophones, and 25 shots at 1.5 m (5 ft) spacing for the total length of 36 m (120 ft). The seismic energy was created by a 16-pound sledgehammer (Phase I) or the propelled energy generator (Phase II).

The acquired seismic data from all 16 test lines have been analyzed by the developed FWI software. Results are similar to what were obtained in Phase I study. One void was found in Line Q. To verify the effect of the off-line void (Task 4), results from 3 lines P, Q, R are presented here.

For the analysis of the acquired data, the initial model was again established via a spectral analysis of the measured data. A linear increasing S-wave velocity from 200 m/s (656 ft/s) at the surface to 600 m/s (1968 ft/s) to a depth of 15 m (50 ft) over a length of 36 m (120 ft) was selected. The initial P-wave velocity of the model was calculated from the S-wave velocity profile, assuming that the initial Poisson's ratio of the whole domain was 0.25. The Poisson's ratio of 0.25 was selected mostly because of expected shallow bedrock (lower Poisson's ratio than that of soil) at the site. A value of 0.3 is recommended for sites with no prior subsurface information. The mass density of all material was kept constant at 1800 kg/m<sup>3</sup> (112 lb/ft<sup>3</sup>) for inversion. The medium of 15 x 36 m (50 x 120 ft) was divided into about 1000 cells of 0.75 x 0.75 m (2.5 x 2.5 ft). During inversion, S-wave and P-wave velocities of cells were updated independently, and the analysis was stopped after 20 iterations when the estimated and measured data are similar.

Shown in Figure 6.5 are the observed, estimated waveforms and residuals associated with the final inverted model for Line Q, shots 1 and 12. Apparently, the observed and estimated waveforms are very similar. The residuals are small for entire test length except at receivers near the sources, which may be attributed to near field effects.

Final seismic results are shown in Figure 6.6a, 6.6b, and 6.6c for Lines P, Q, and R, respectively. A consistent pattern is found for subsurface profiles at the 3 lines. They all include soft soils at shallow depths from 0 to 3 m (0 to 10 ft), highly variable limestone with S-wave

velocity more than 600 m/s (1968 ft/s) at the bottom, shallow limestone at distances of 12 m (40 ft) and 23 m (77 ft), and low-velocity anomalies near the middle of the array. Line Q (Figure 6.6 b) shows an embedded void at distance 18 m (60 ft) [S-wave velocity less than 50 m/s (164 ft/s)], which is confirmed by the SPT test (Figure 6.6d). The void diameter is about 3 m (10 ft). However, the predicted depth (6 to 9 m) is deeper than the real depth of the void (3.5 to 6.5 m), this is mostly attributed to the discrepancy between the estimated waveform data (plane strain) and the measured data (3-D/non-plane strain). The limestone pinnacles at distances 12 and 23 m made the assumed plane strain condition less accurate. Also, the limestone pinnacles (3D) altered the reflected/refracted seismic wave travel time to the void, resulting in less accurate results. As a consequence, a complete 3-D FWI analysis is required to characterize this 3-D challenging subsurface condition.

Lines P and R are 3 m (10 ft), or about one void diameter, away from the void center (line Q) on each side. Apparently, the void at line Q does not show up on line P and R results (Figure 6.6a and 6.6c). The results agree with the synthetic study on 3-D effect of off-line voids (Task 4). That is, the influence of the off-line void is minimal if the test line is one diameter or more away from the void.

(a) Photograph of the retention pond



(b) Site survey map

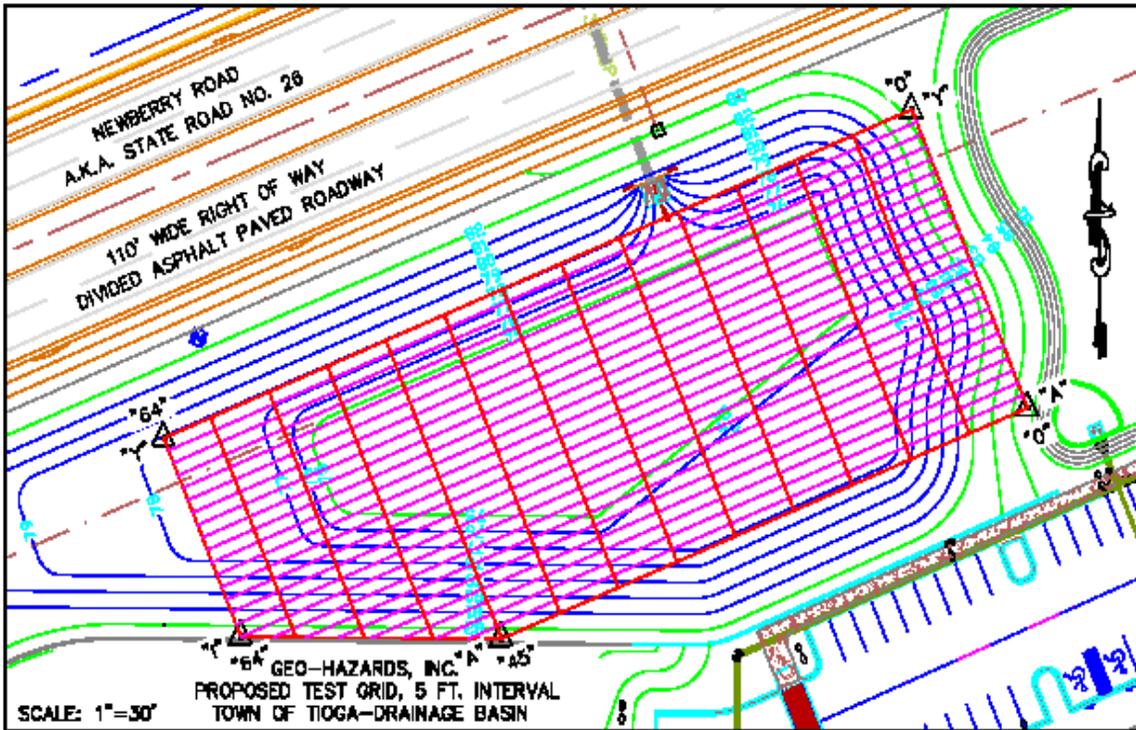


Figure 6.4 Newberry site: (a) photograph of the retention pond and (b) site survey map

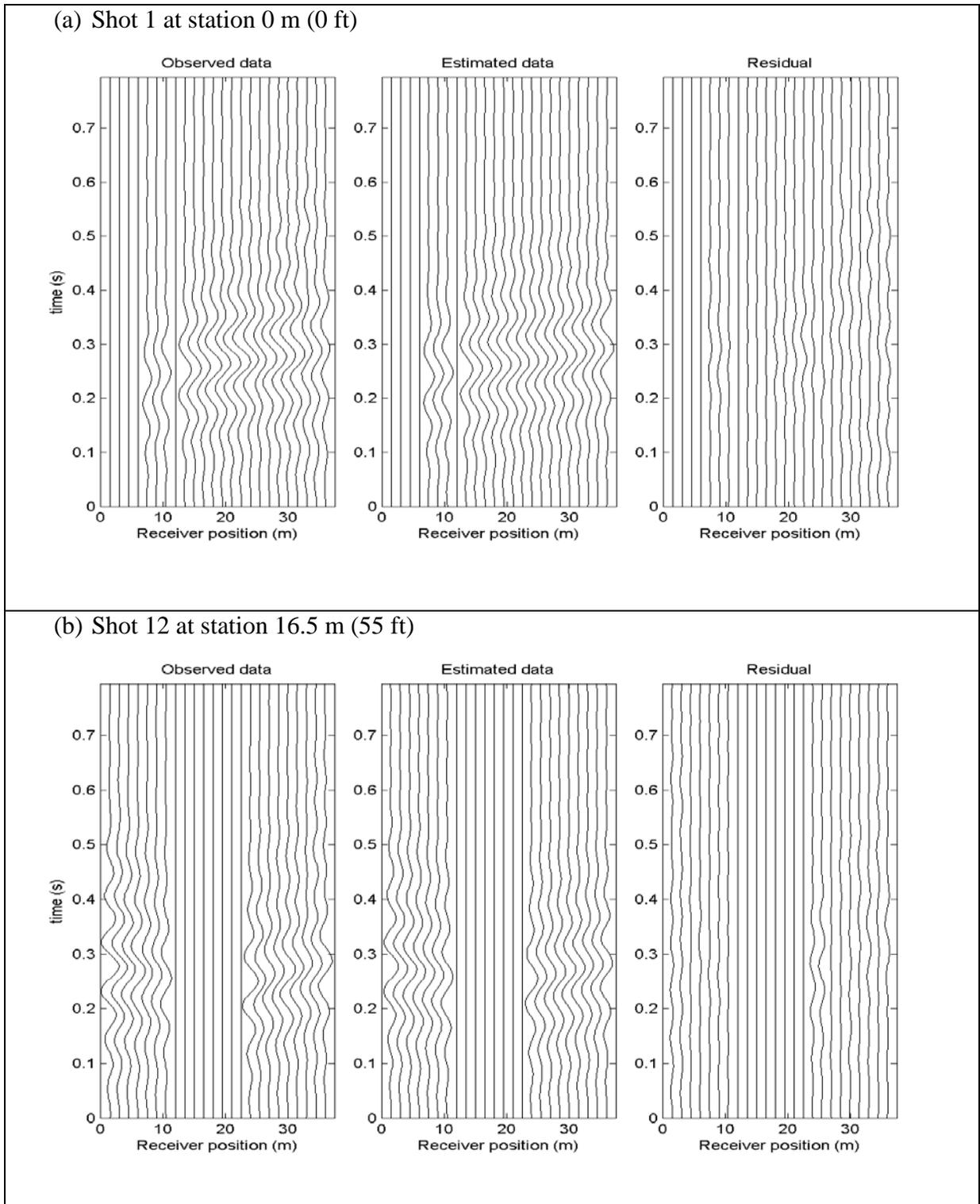


Figure 6.5 Newberry, line Q: comparison between observed and estimated data for shots 1 and 12.

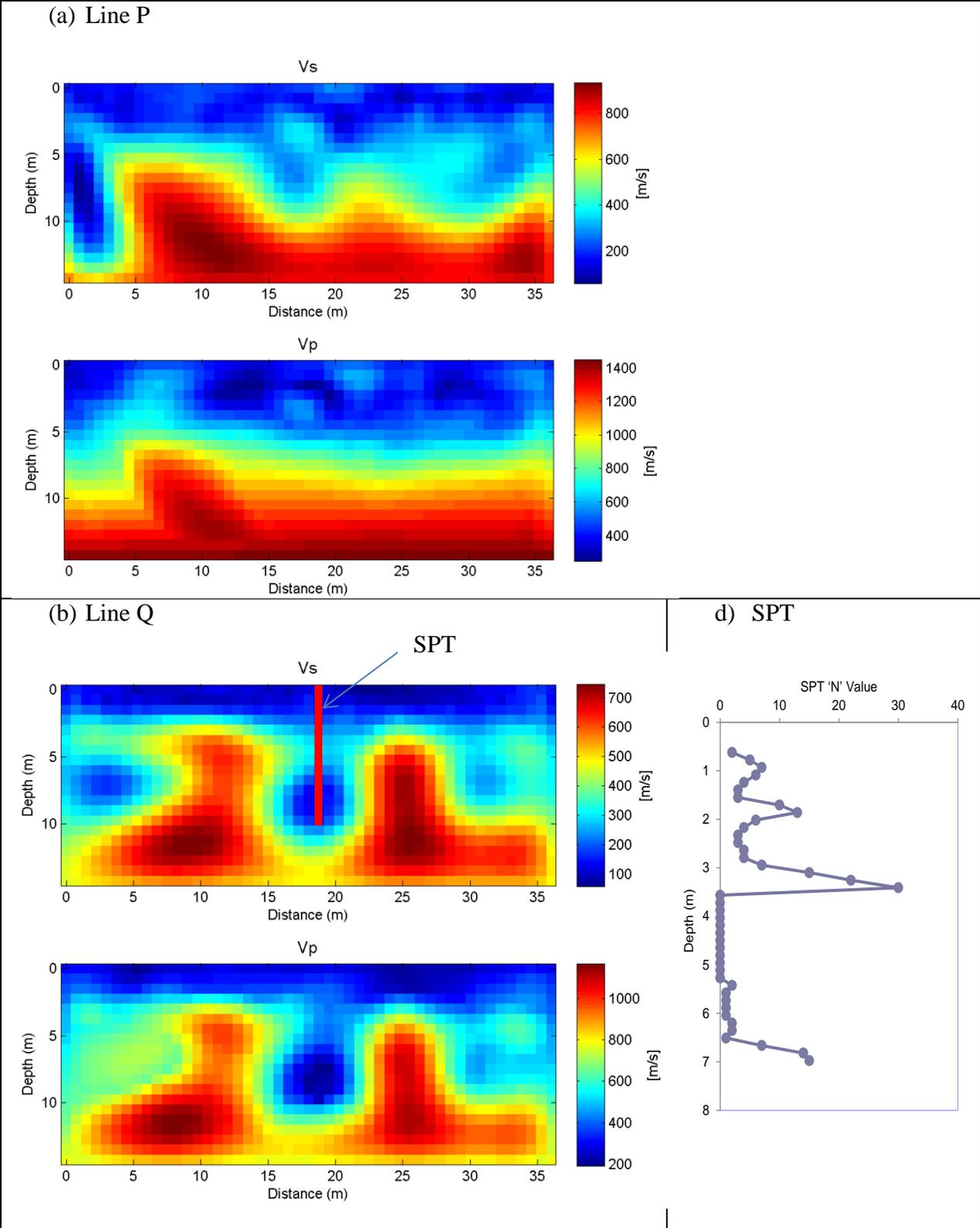


Figure 6.6 Newberry site: S-wave and P-wave velocities: (a) Line P, (b) Line Q, (c) Line R, and (d) SPT blow counts

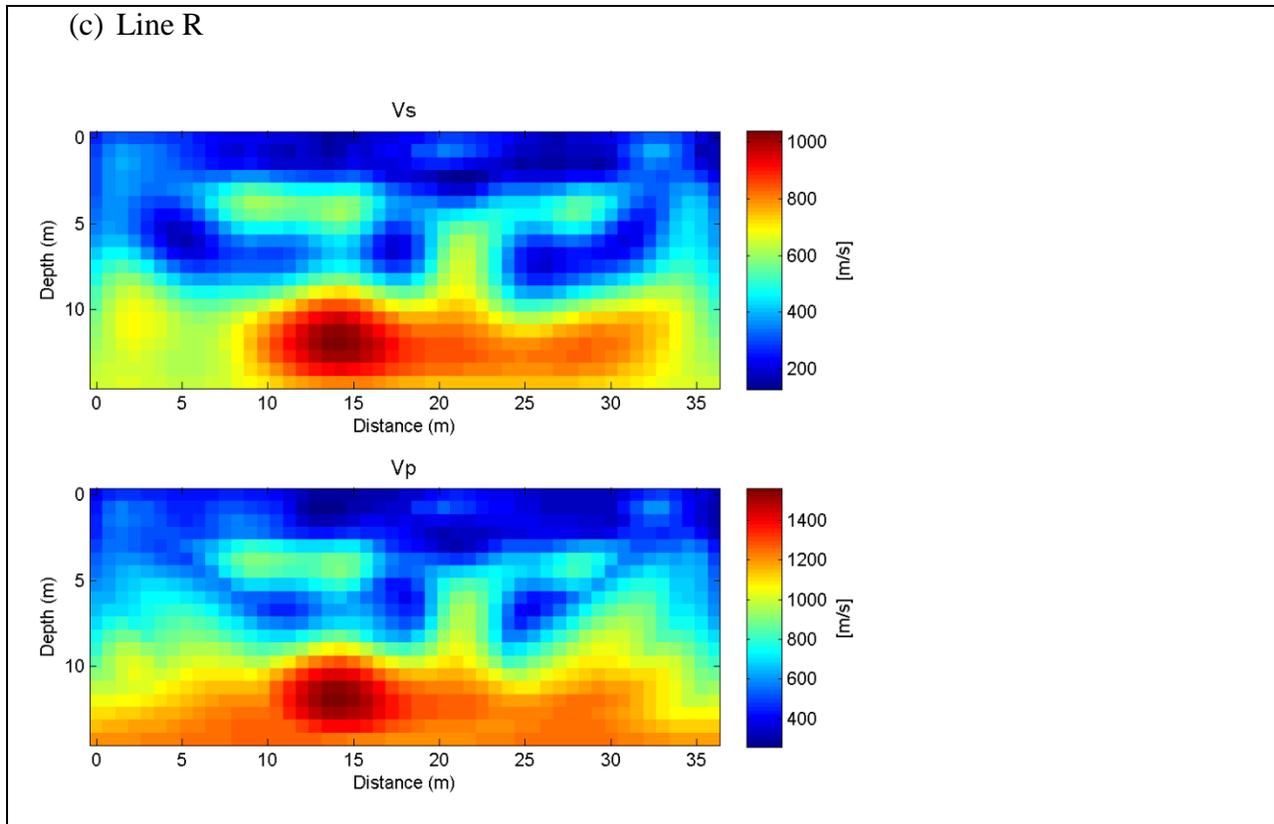


Figure 6.6 Newberry site: S-wave and P-wave velocities: (a) Line P, (b) Line Q, (c) Line R, and (d) SPT blow counts

#### 6.4 Gainesville Site

The test site is a dry retention pond in Gainesville, Florida. Four parallel test lines at 3 m apart (10 ft) were conducted. Each line was conducted using a linear array of 24 4.5-Hz vertical geophones at 1.5 m (5 ft) spacing, and 13 shots at 3 m (10 ft) spacing, for the total length of 36 m (120 ft) as shown in Figure 6.7. The seismic energy was created by the propelled energy generator.

Seismic data of all four test lines were analyzed by the FWI software, and no void was found underneath the test lines. Results from one test line (4 ft away from the open chimney as shown Figure 6.7 b) with an embedded low-velocity anomaly are presented here.

For the waveform analysis, two separate inversion runs were performed on the filtered data sets with central frequencies of 15 Hz and 20 Hz, beginning with the lower frequency range.

1-D linear increasing velocity initial model was generated by the software using the spectral analysis of the measured data. The medium of 18 x 36 m (60 x 120 ft) was divided into 1,152 cells of 0.75 x 0.75 m (2.5 x 2.5 ft). During inversion, both S-wave and P-wave velocities of all cells were updated independently, and each run was stopped after 20 iterations, when the change of the least-squares error became small (less than 1% from one to the next iteration).

The observed waveforms, the estimated waveforms, and residuals associated with the final inverted model are shown in Figure 6.8 for shots at stations 0 and 36 m (0 ft and 120 ft), respectively. The observed and estimated waveforms generally match. Large residuals of a few channels are mostly due to the reflected signals from the offline chimney, which are not modelled by the 2-D forward modelling.

Inverted results are shown in Figure 6.9a. The  $V_s$  profile includes a soft soil layer at shallow depths from 0 to 6 m (0 to 20 ft) followed by a stiffer soil layer. A valley of low-velocity soils is found at a distance of 23 m (77 ft) near the chimney. The  $V_p$  profile is consistent with the  $V_s$  profile. Shown in Figure 6.9b are the SPT blow counts at distance 23 m (77 ft) on the test line. The SPT results seem consistent with the inverted seismic results. Low SPT 'N' values for looser soils from 0 to 9-m depth (0 to-30 ft), high 'N' values for denser soils below 9-m (30-ft) depth, and a mild reversal of 'N' values from depth of 14 to 18 m (47 to 60 ft) associated with a low-velocity zones of  $V_s$  at the bottom of the medium (Figure 6.9a).

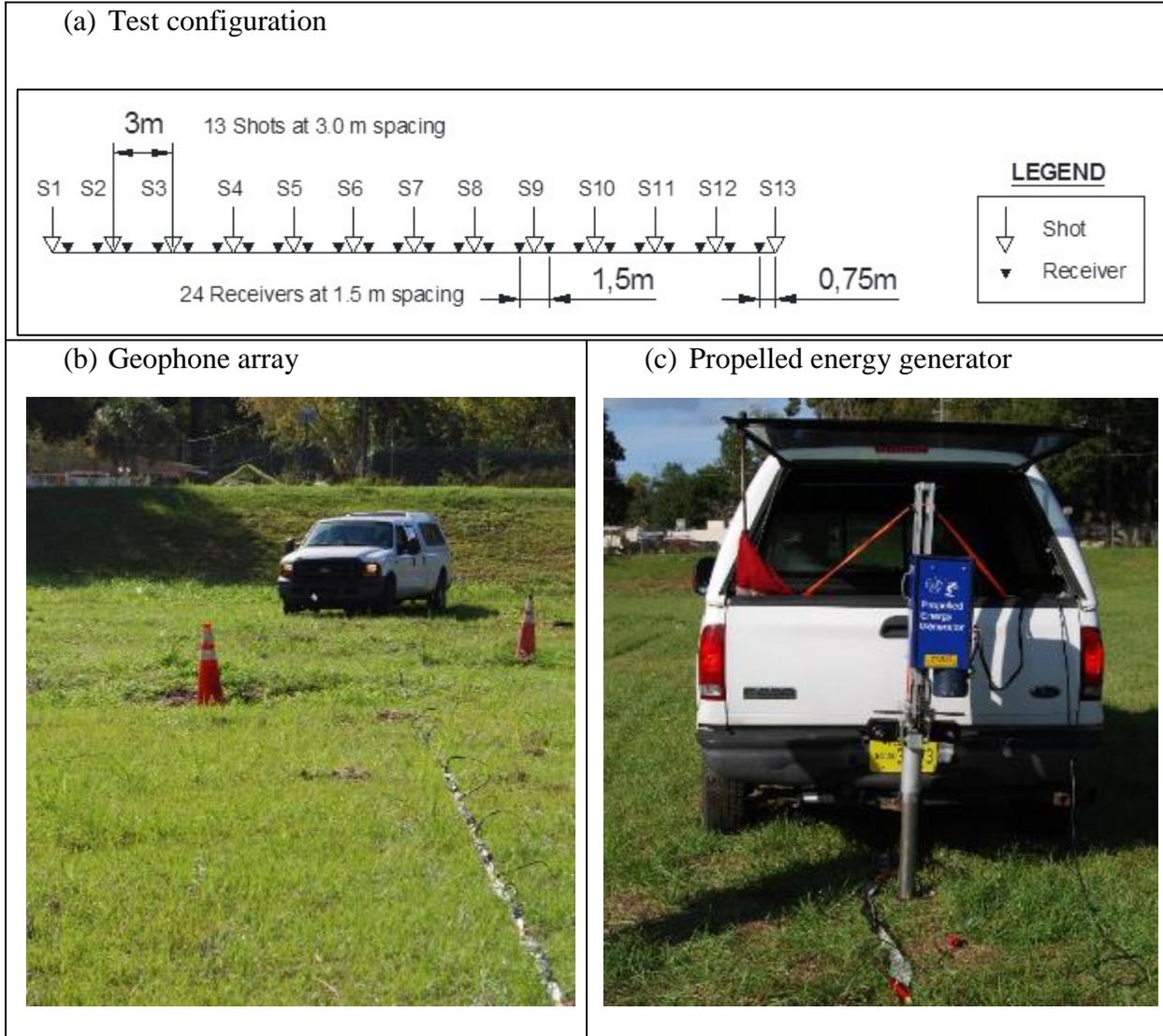
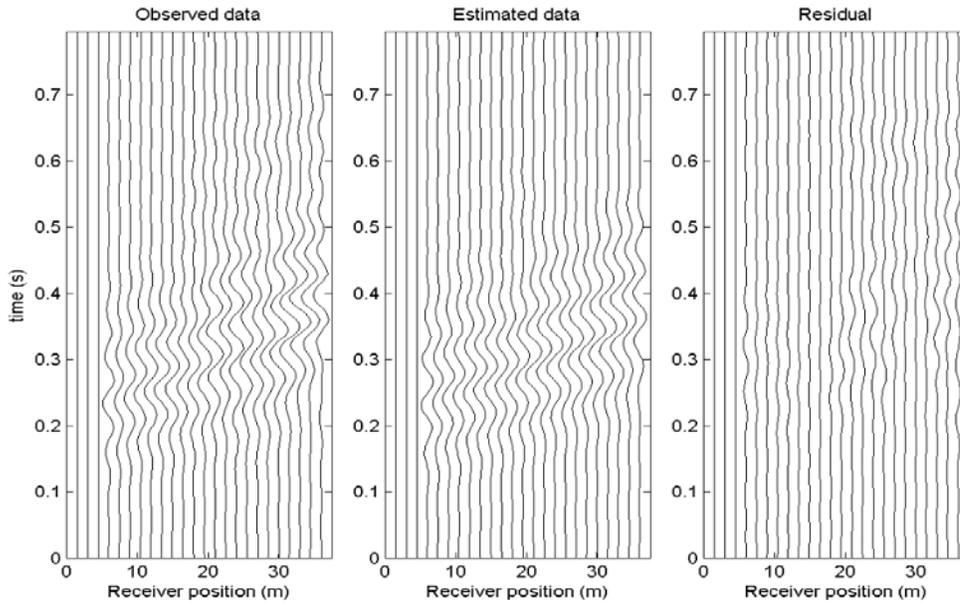


Figure 6.7 Gainesville site: (a) Test configuration, (b) Geophone array, and (c) Propelled energy generator source.

(a) Shot 1 at station 0 m (0 ft)



(b) Shot 13 at station 36 m (120 ft)

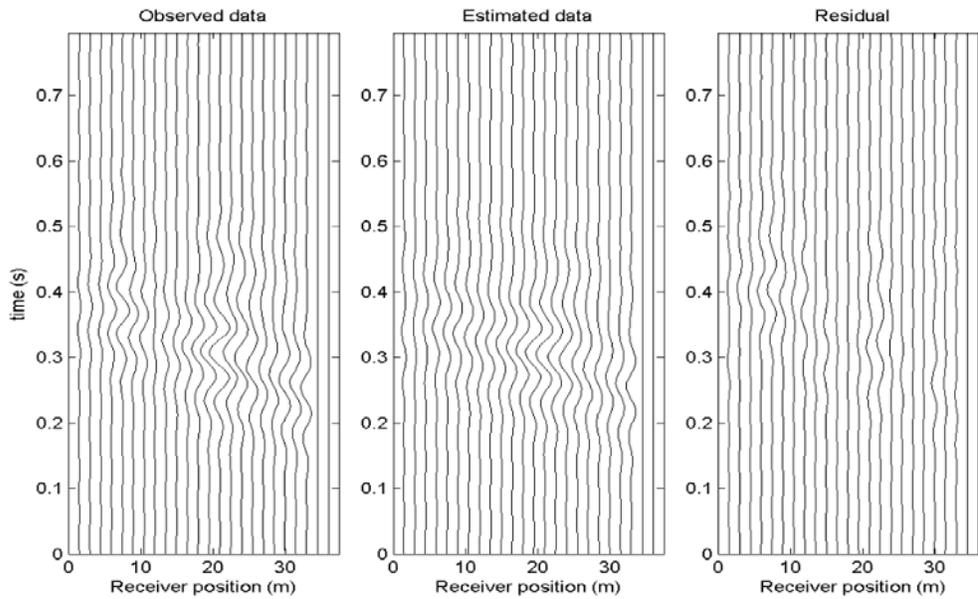


Figure 6.8 Gainesville site: comparison between observed and estimated data for the first and last shots.

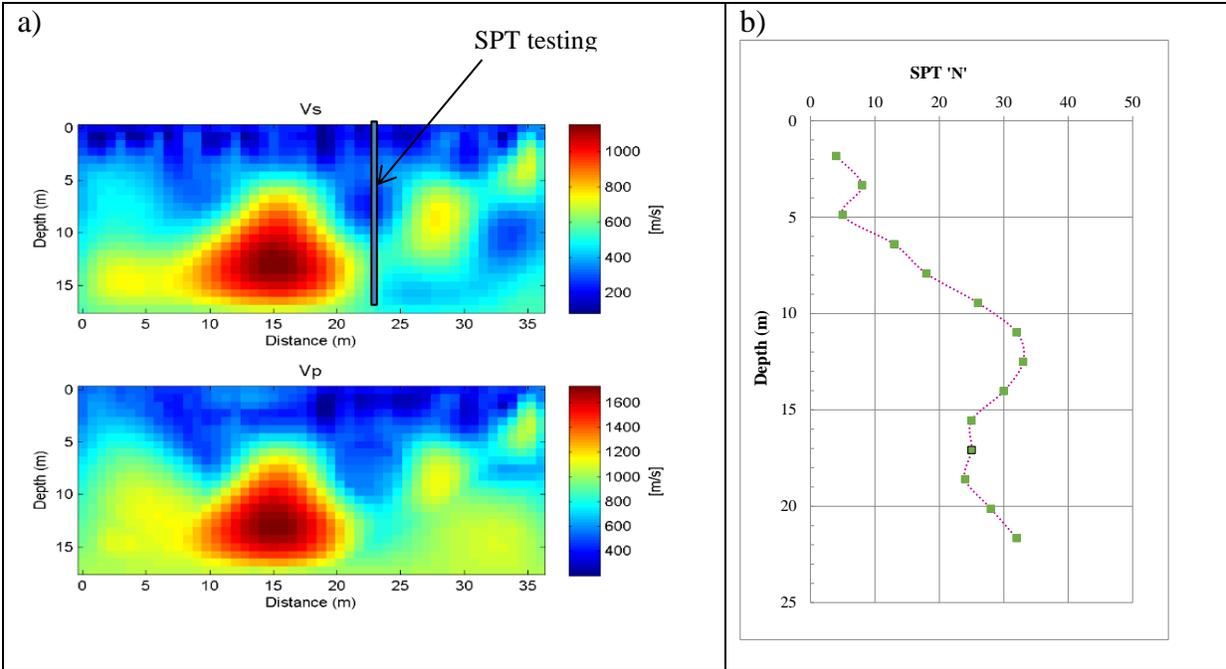


Figure 6.9 Gainesville site: (a) S-wave and P-wave velocities and (b) SPT blow counts

### 6.5 Tallahassee Site

The test site is on SR 263 or Capital Circle Rd in Tallahassee, Florida. The seismic survey was conducted at the location of pile refusal/punch through, which happened during pile driving at the site (Figure 6.10a). The survey was centered on the H-pile identified in Figure 6.10a as Pile 34A (one of the two piles for proposed sheet pile wall section 34). Three H-piles (Figure 6.10a) were driven to depths of from 14 to 15 m (45 to 50 ft). During the driving, the resistance increased significantly at 35ft, exhibiting signs of practical refusal. Once the pile “punched through”, heated and pressurized steam was released (Fig. 6.10b) with enough force to lodge clay materials between the pile hammer's helmet and H-pile.

To explore the subsurface condition for a possible explanation of the incidents, one test line was conducted using a linear array of 24 4.5-Hz vertical geophones and 29 shots at 1.5 m (5 ft) spacing, for the total length of 42 m (140 ft) as shown in Figure 6.10c. The seismic energy was created by a sledgehammer. Due to the challenging ground surface condition with very soft

soil or mud and bad weather (raining), part of collected data (10 shots) was not good enough for waveform analysis. Only data from 19 shots were used for analysis presented herein.

Collected data were analyzed by the FWI software. Two inversion runs were performed on the filtered data sets with central frequencies of 12 Hz and 20 Hz, beginning with the lower frequency range. One-dimensional linear increasing velocity initial model was generated by the software using the spectral analysis of the measured data. The medium of 21 x 36 m (70 x 120 ft) was divided into 1344 cells of 0.75 x 0.75 m (2.5 x 2.5 ft). Both S-wave and P-wave velocities of all cells were updated independently during inversion, and each run was stopped after 20 iterations, when the when the estimated and measured data are similar (Fig. 6.11) and the change of the least-squares error became small (usually less than 1% from one to the next iteration).

Inverted results are shown in Figure 6.12. The  $V_s$  profile includes a soft soil layer at shallow depths from 0 to 12 m (0 to 40 ft), a stiff soil layer at depths from 12 to 16 m (40 to 53 ft) at left and right of the profile, followed by a softer layer at the bottom of the medium. If the stiff medium with weaker void beneath existed at location of the piles, then the pile would exhibit refusal from 14 to 15-m (45 to 50-ft) depth (Figure 6.13). Subsequently, if the piles punched through or even broke up the stiff material, it may have been forced downward and underlying void may have moved upward expelling both gas and possibly steam/soil at ground surface. Evident, there exists a very low-velocity zone ( $V_s < 50$  m/s (167 ft/s) at the locations of incident piles 33 and 34 at distance from 15 to 23 m (50 to 77 ft). It is noted that incident pile 32 is not on the seismic line, the pile is about 2 m off the seismic line. The  $V_p$  profile is consistent with the  $V_s$  profile, including a stiff layer at depth from 12 m to 16 m (40 to 53 ft) at left and

right with a low-velocity zone ( $V_p$  of about 300 m/s (1000 ft/s)) at the incident location. The  $V_p$  of about 300 m/s (1000 ft/s) suggests that the void is filled by air, not water.

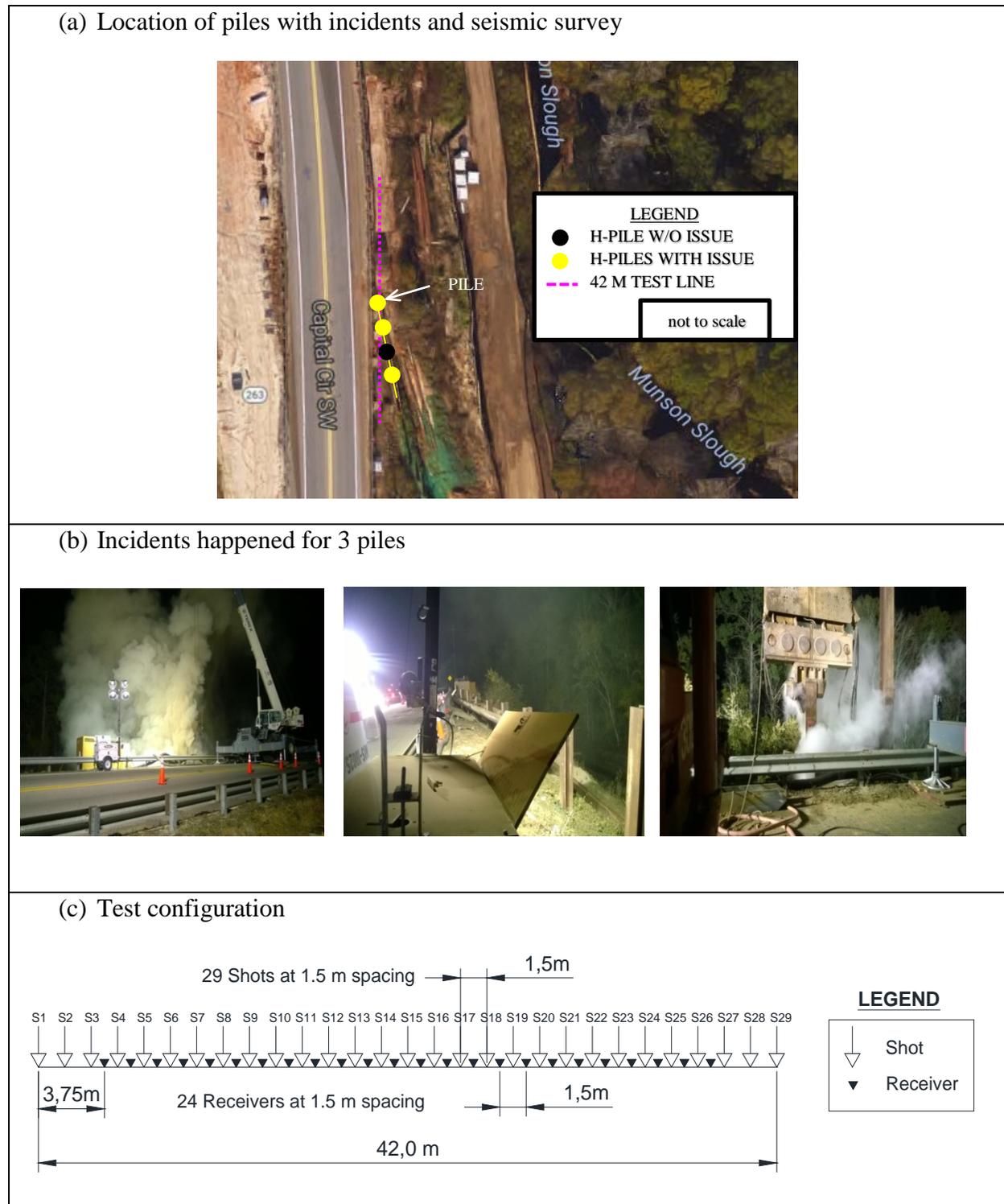


Figure 6.10 Tallahassee site: test location and configuration

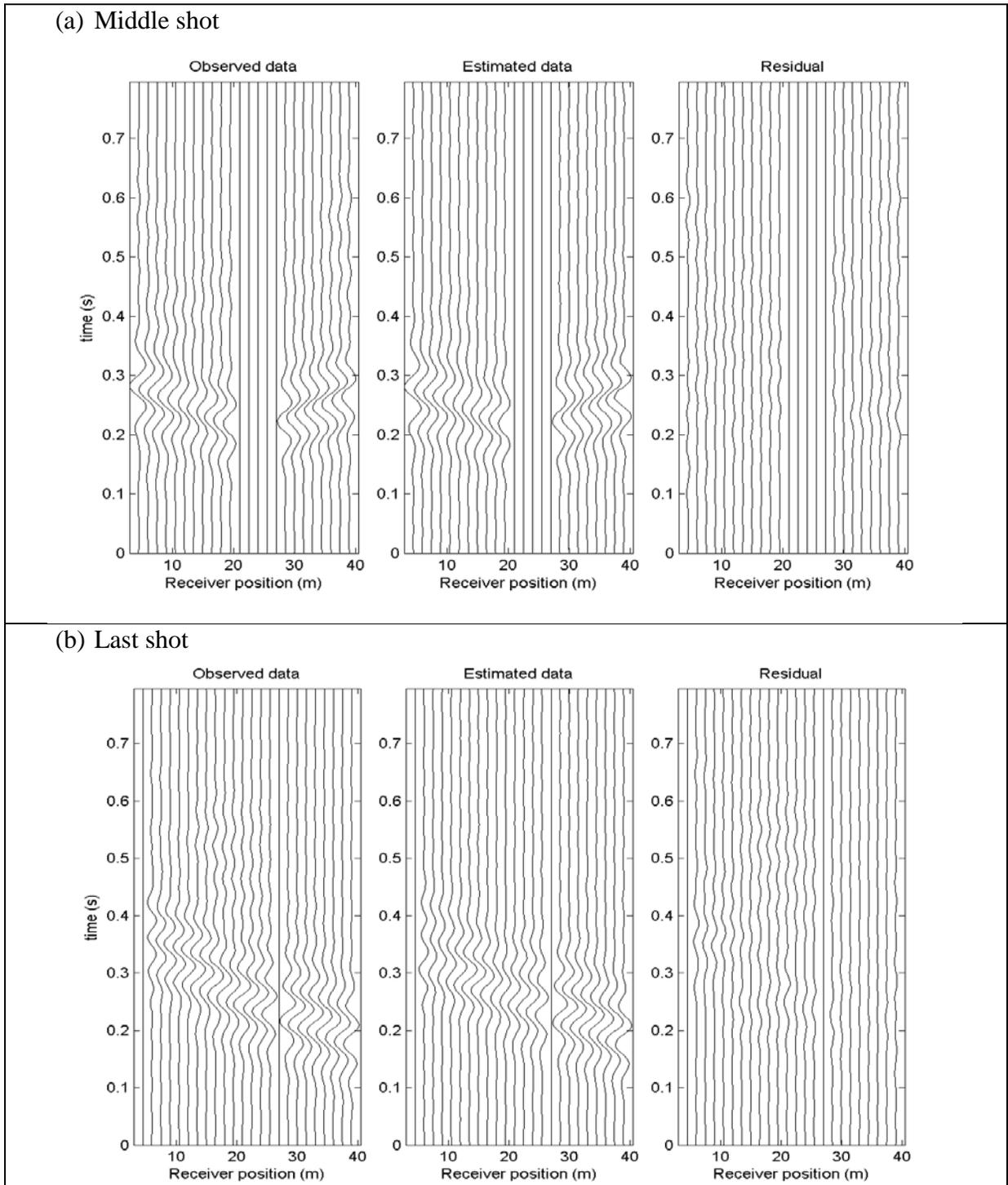


Figure 6.11 Tallahassee site: comparison between observed and estimated data for the middle and last shots

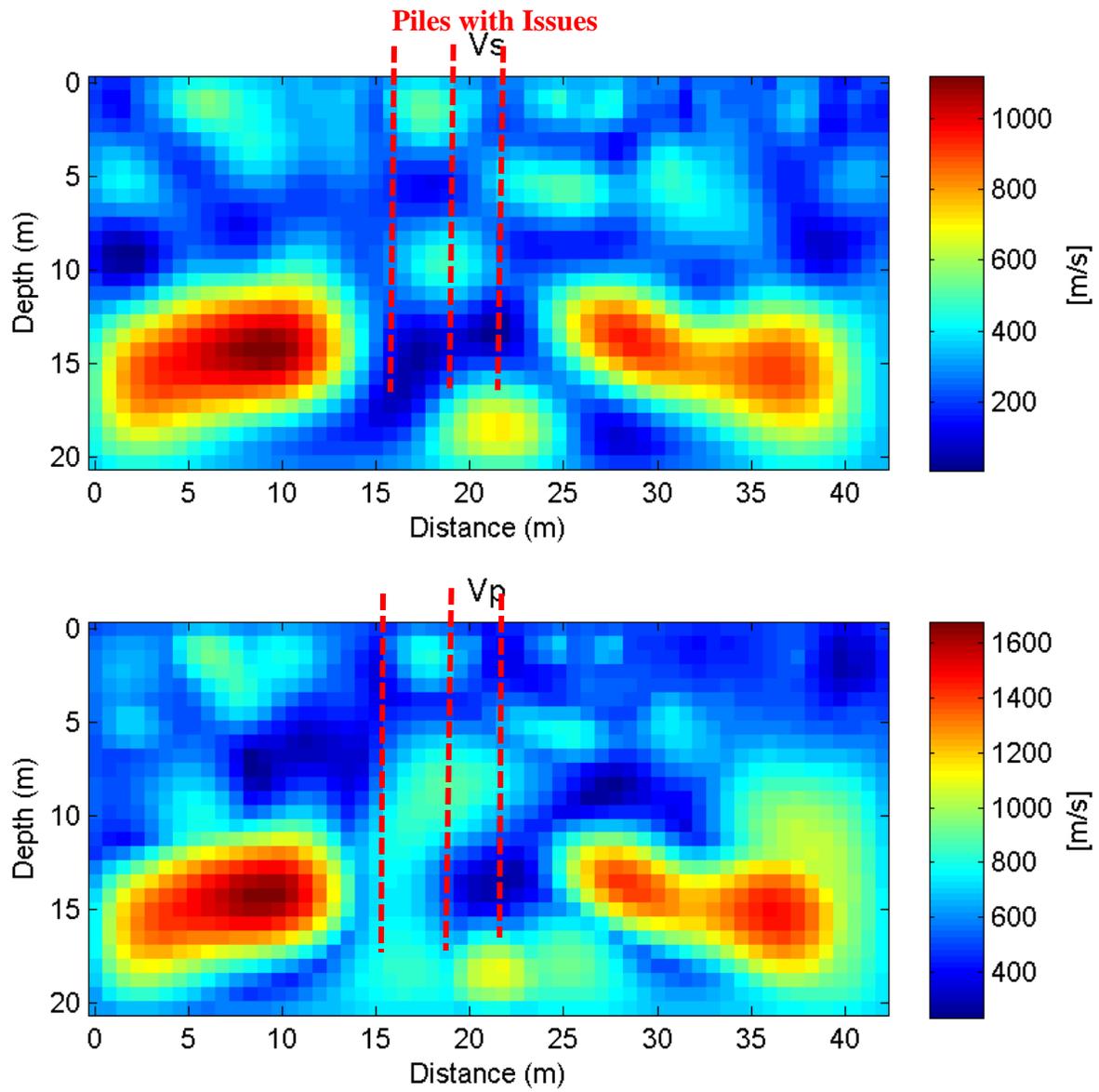


Figure 6.12 Tallahassee site: S-wave and P-wave velocities

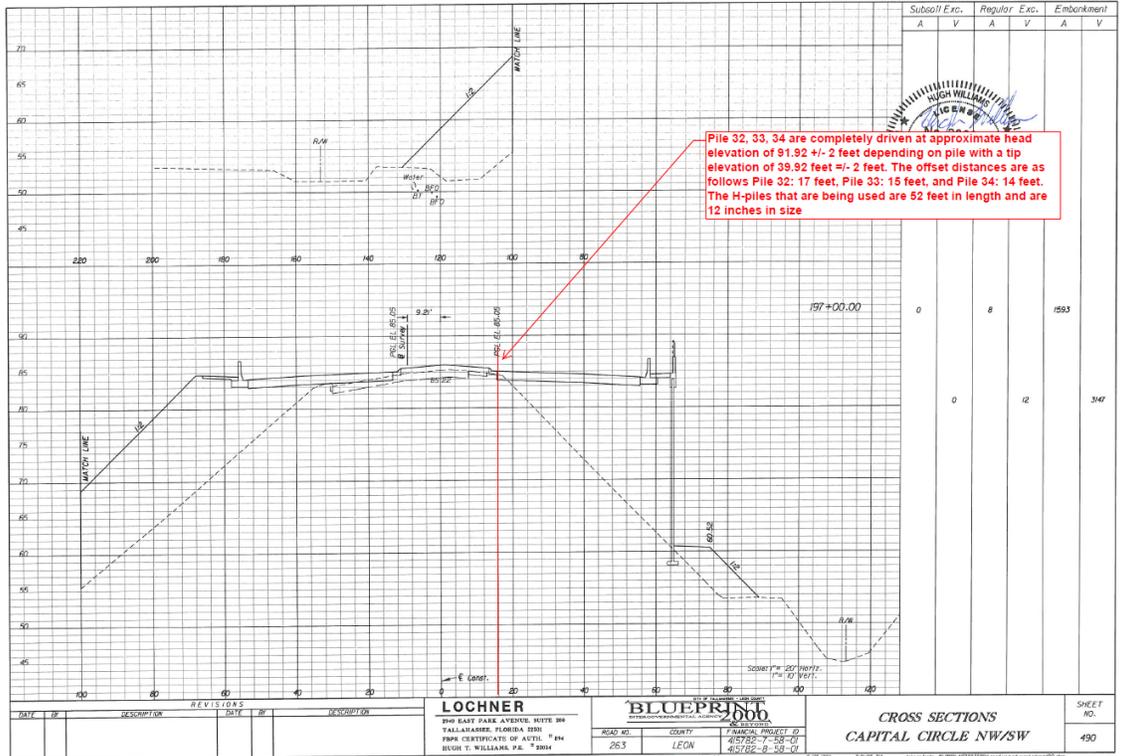


Figure 6.13 Tallahassee site: elevations of three incident piles (32, 33 and 34).

### 6.6 Kanapaha Site

The test site is in Kanapaha area of Gainesville at a FDOT maintenance storage area. . It was divided into 10 parallel east-west survey lines equally spaced 3.0 m (10 ft) apart. The lines were labeled L1-EW through L10-EW from East to West across the site. The first five lines were analyzed using 25 shots, which were 1.5 m (5 ft) apart for a length of 36 m (120 ft) ( Figure 6.14 a), while the other five lines were analyzed using 29 shots, which were 1.5 m (5 ft) apart for a length of 42 m (140 ft) (Figure 6.14 b). For all lines the station 0 m located at the eastern end of each line and 24 receivers at 1.5 m (5 ft) spacing were used. Data from all 10 lines were analyzed by the FWI software. Although no void was identified underneath any of the test lines, results are presented herein to demonstrate the FWI software capability to characterize highly variable soil and rock layers.

For data analysis of each line, two inversion runs were performed on the filtered data sets with central frequencies of 15 Hz and 19 Hz. Ricker wavelets having central frequencies of 15 and 19 Hz were used to estimate the active source (PEG) for the forward modeling of each analysis. The first run began with the lower frequency range (central frequency of 15 Hz) using the 1-D linear increasing velocity initial model, and the second run for the central frequency of 19 Hz was completed using the inverted result of the first run as the initial model. During inversion, both S-wave and P-wave velocities of all 0.75 x 0.75 m (2.5 x 2.5 ft) cells were updated independently, and each run was stopped after 20 iterations.

As an example, the observed and estimated waveform data, and residuals associated with the final inverted model of the first line (L1-EW) are shown horizontally in Figure 6.15 for the shots at stations 0 and 36 m (0 ft and 120 ft) respectively. It is evident that the observed and estimated data are similar across the entire range of shot-receiver offsets, except a few far-field channels. This could be due to measured signals reflected from off-line high-velocity objects (rock pinnacles) that cannot be modelled by 2-D plane strain forward simulation.

Inverted results from all 10 test lines are shown in Fig. 6.16. From the Vs profiles, a consistent pattern is observed for the subsurface profiles along the 10 lines. They all include soft soils at shallow depths from 0 to 10 m (0 to 32 ft), and highly variable limestone with Vs more than 600 m/s (2000 ft/s) at the bottom; the top of limestone layer varies from 8 to 10-m depth (23 to 32-ft). Vp profiles are consistent with the Vs profiles. For better demonstration of the site subsurface variation, 3-D views of Vs and Vp profiles are shown in Figure 6.17 and Figure 6.18. They clearly show a soft soil layer, underlain by a highly variable limestone layer.

A number of CPT tests were conducted at the site, and the top of the limestone layer was encountered at depths of 6 to 9 m (20 to 30 ft). Shown in Figure 6.19 is a sample of the CPT results on line L6-EW. Apparently, the CPT tip resistance is consistent with the seismic results. Both show a stiff layer at about 8 m (24 ft) in depth.

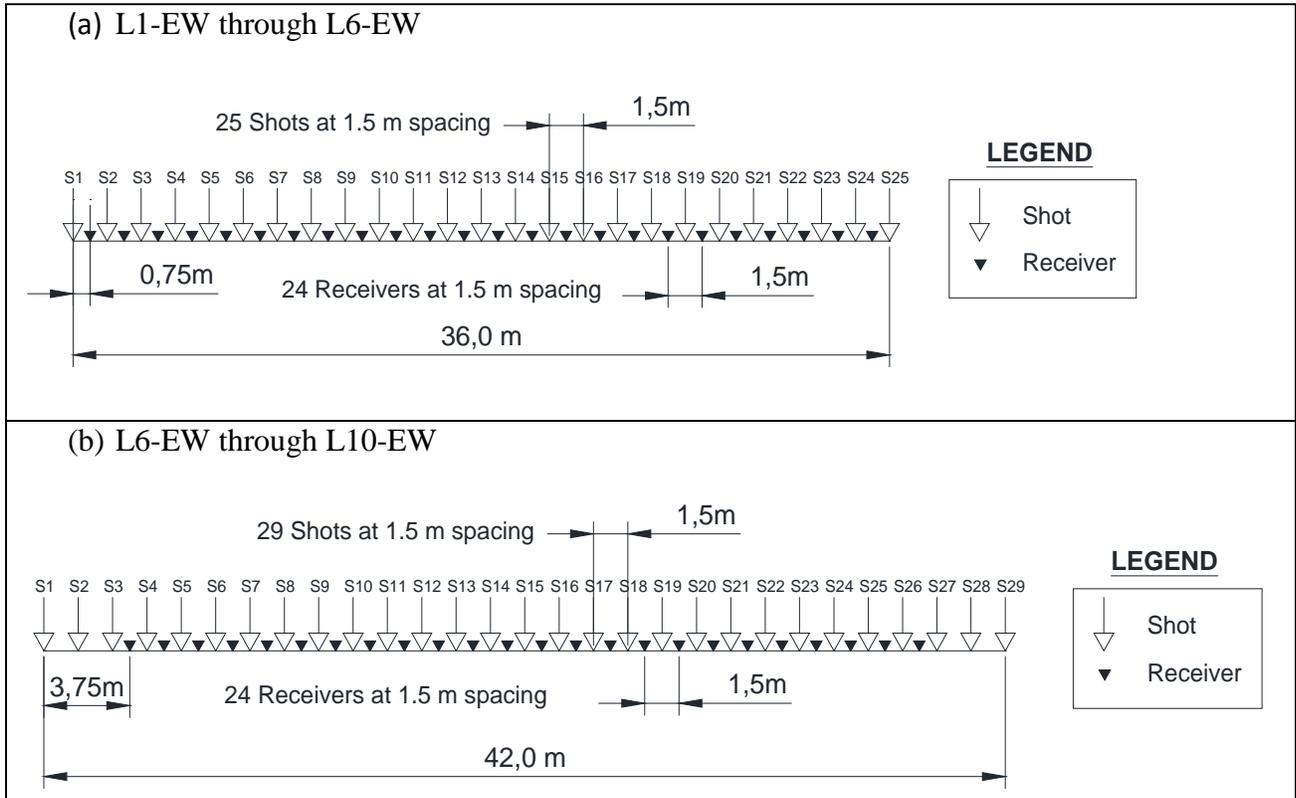
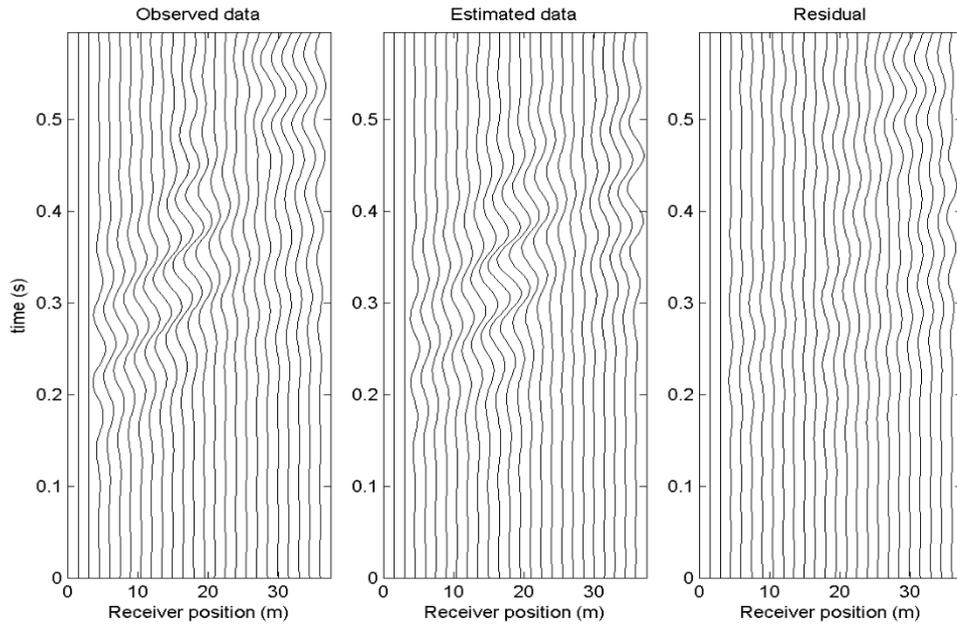


Figure 6.14 Kanapaha site: test configuration

(a) Shot 1 at station 0 m (0 ft)



(b) Shot 25 at station 36 m (120 ft)

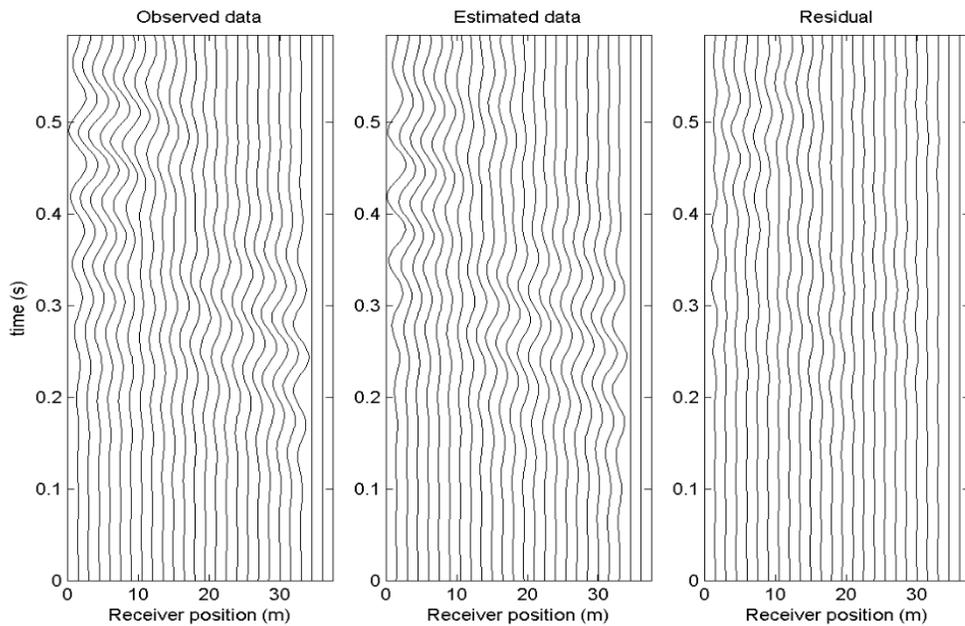


Figure 6.3 Kanapaha site, L1-EW: comparison between observed and estimated data for shots at 0 and 36 m (0 ft to 120 ft)

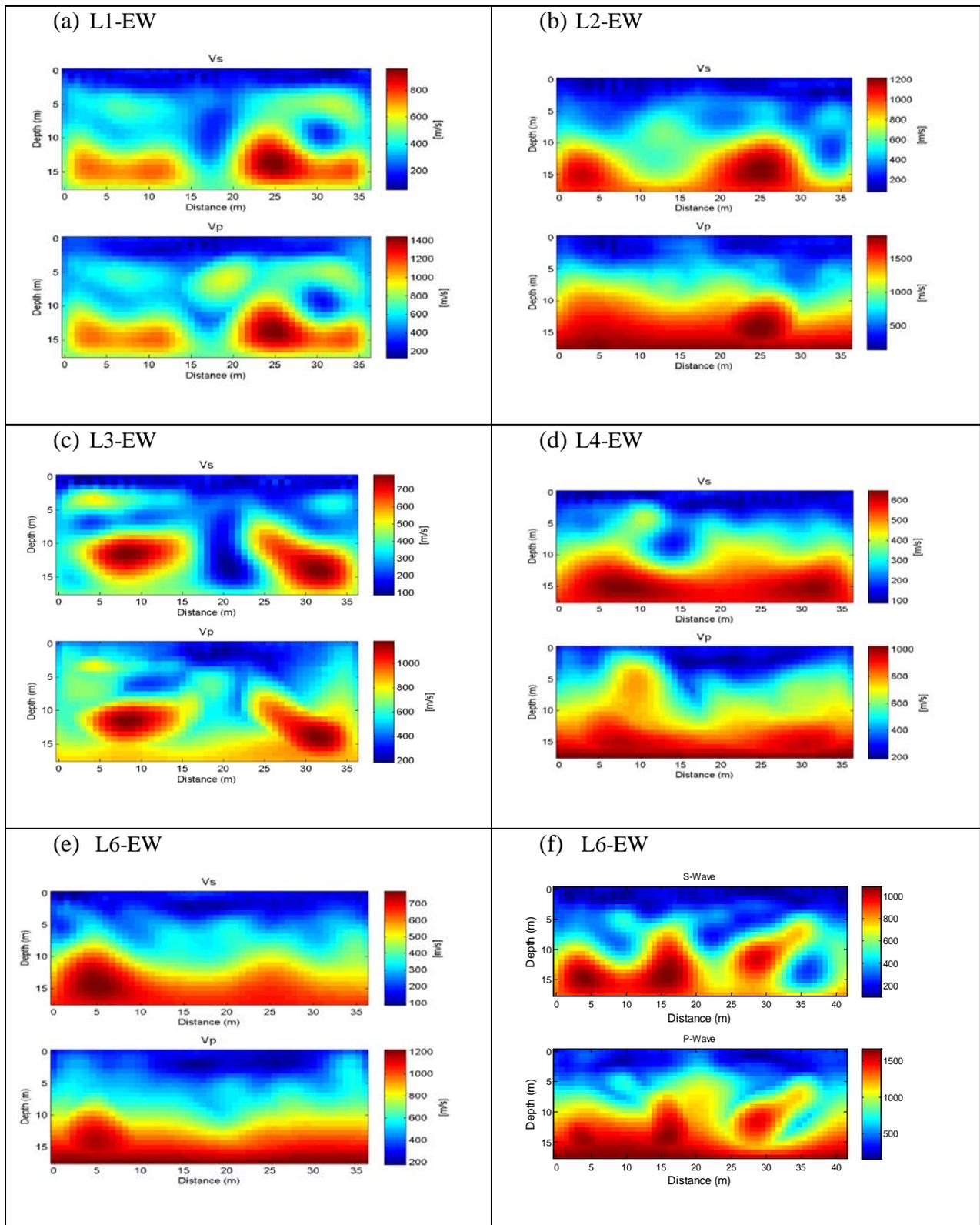


Figure 6.16 Kanapaha site: S-wave and P-wave velocities (m/s): (a) to (j) are lines L1-EW to L10-EW, respectively.

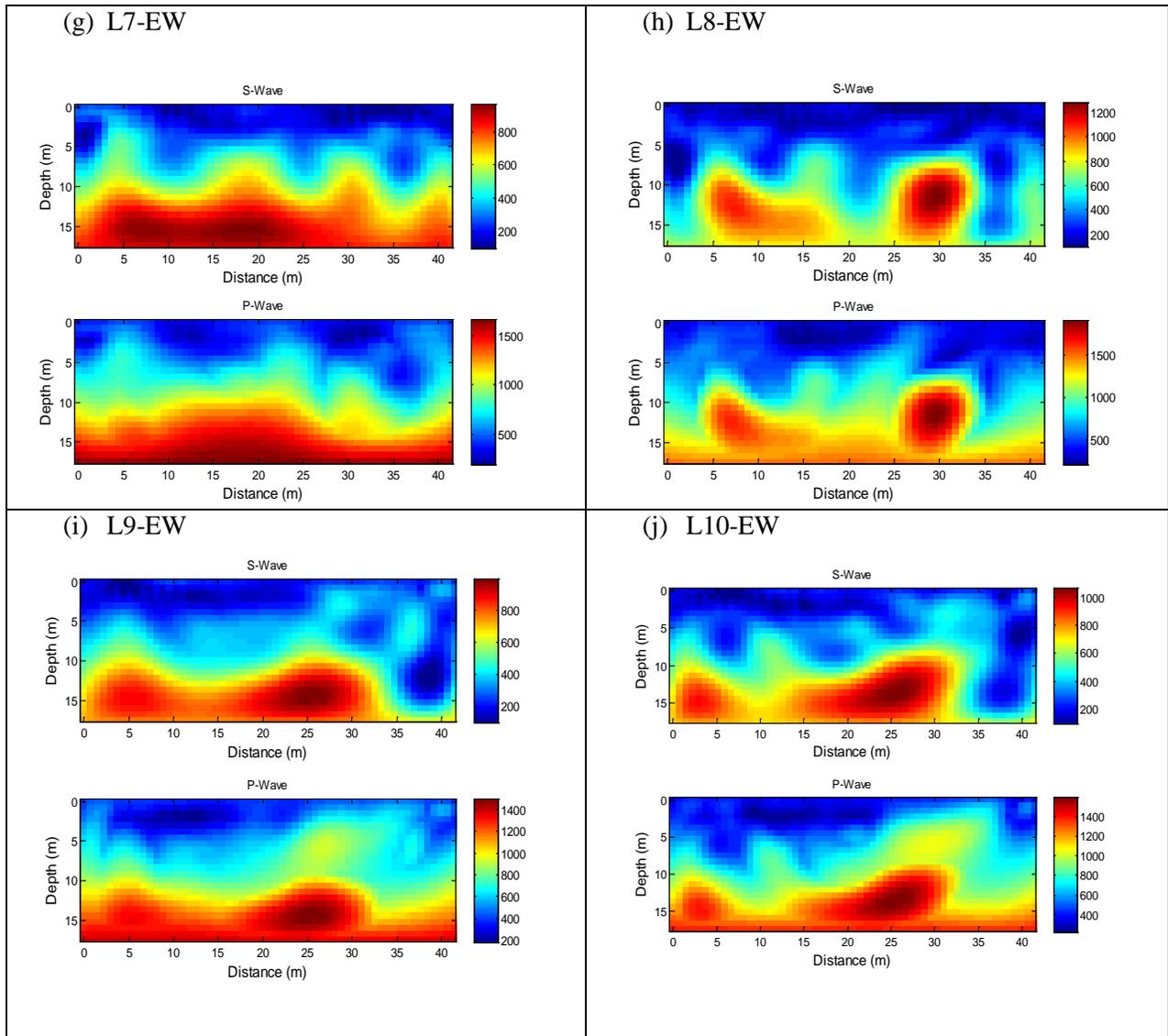


Figure 6.16 Kanapaha site: S-wave and P-wave velocities (m/s): (a) to (j) are lines L1-EW to L10-EW, respectively.

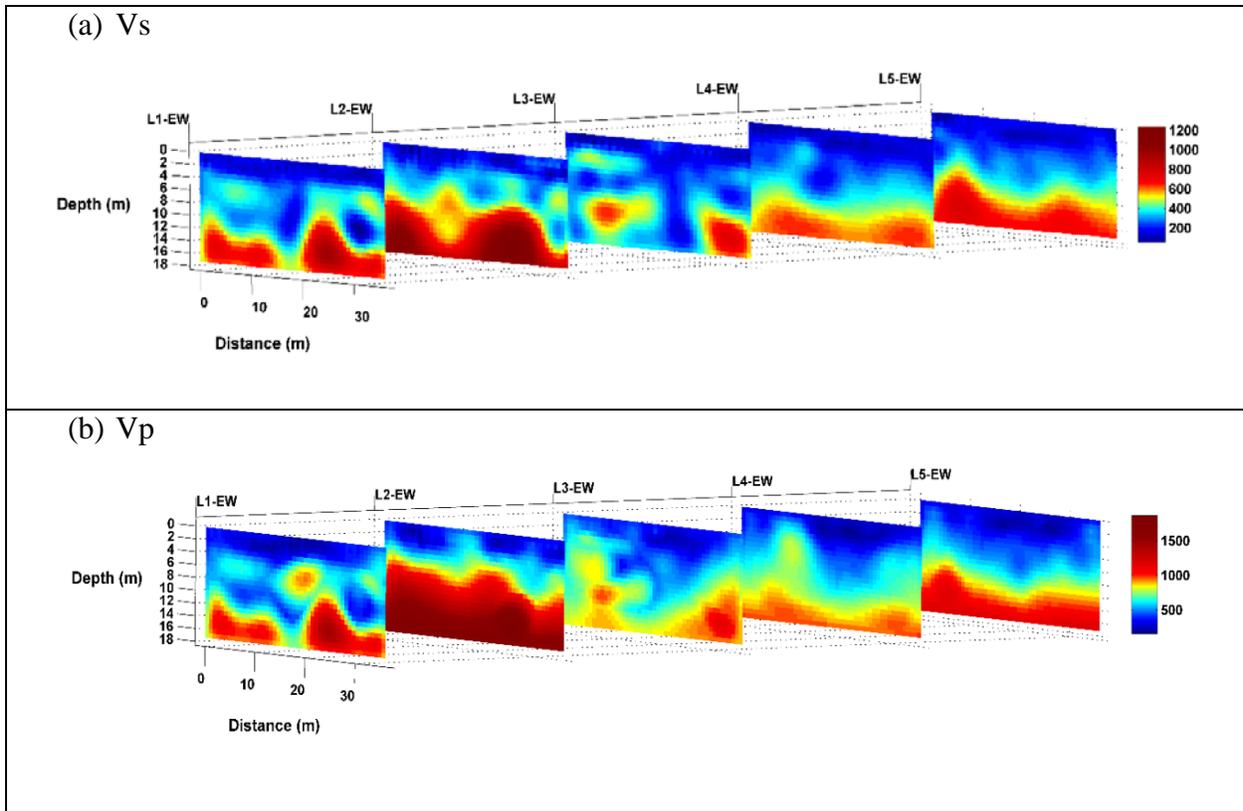


Figure 6.4 Kanapaha site: 3-D view of S-wave and P-wave velocities (m/s): L1-EW to L5-EW.

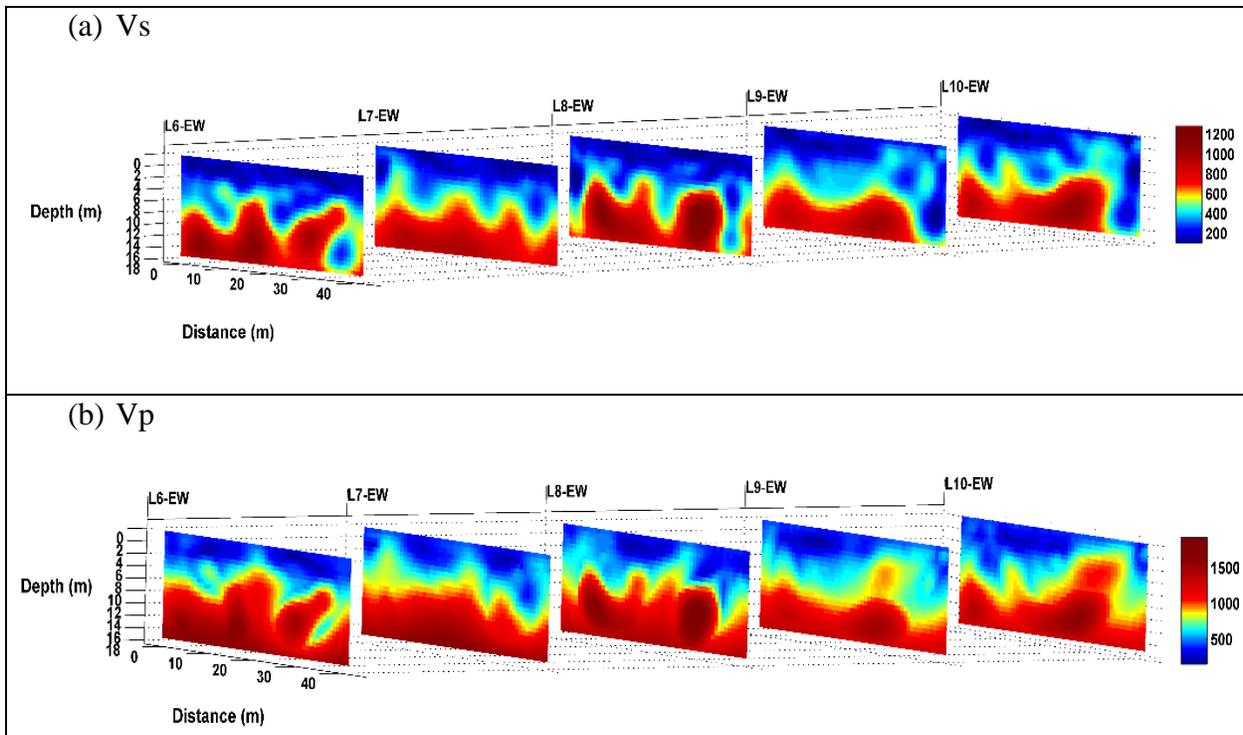


Figure 6.5 Kanapaha site: 3-D view of S-wave and P-wave velocities (m/s): L6-EW to L10-EW.

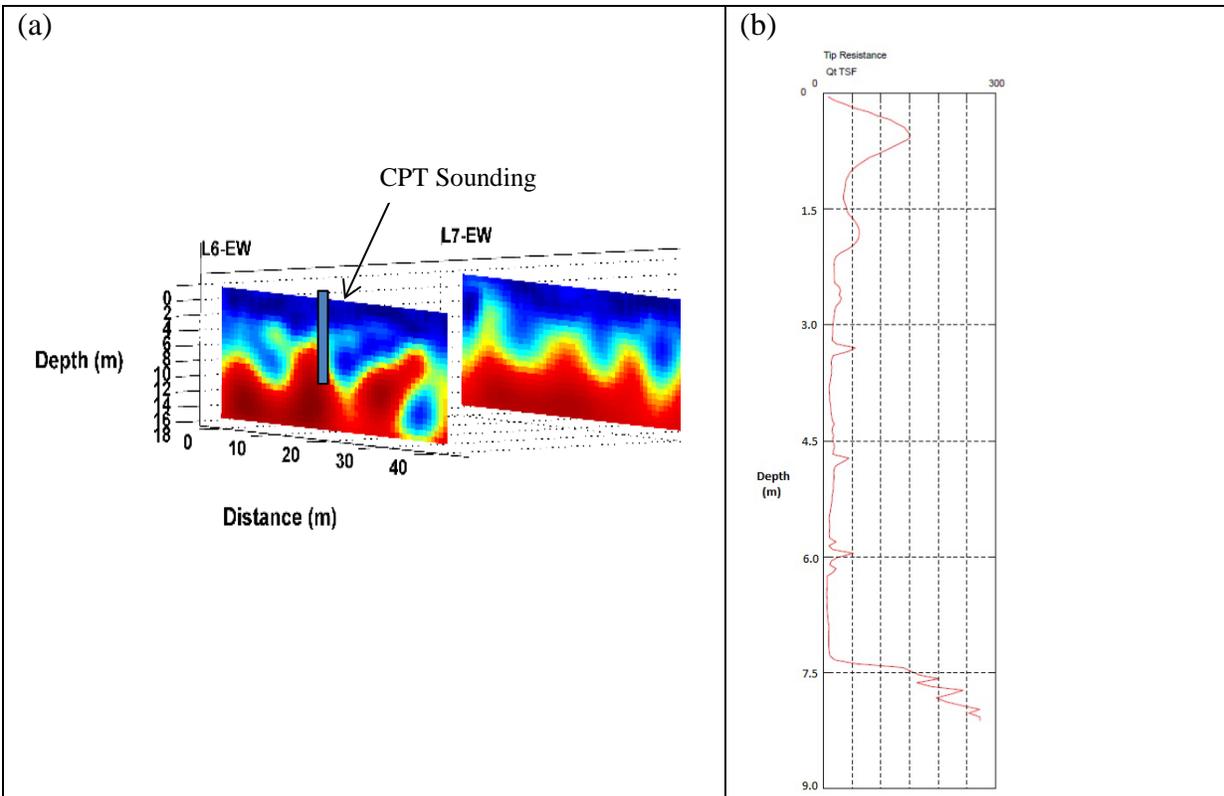


Figure 6.19 Kanapaha site: sample CPT sounding: (a) location and (b) CPT tip resistance.

## 6.7 Conclusions

The developed FWI software was used on data collected from 5 test sites in Florida with various subsurface conditions, including US 441, Newberry, Gainesville, Tallahassee, and Kanapaha sites. For each test site, multiple seismic test lines were conducted for at least 37 m (120 ft) long, using a 24-channel seismic system. Seismic energy was generated using either a sledgehammer 67 to 89 kN (15 to 20 lbs) or a PEG (PEG-40 Kg model). Seismic results were obtained by the software in about 30 minutes for each test line on a standard desktop computer.

Seismic results show that the FWI software was able to identify variability including anomalies within soil and limestone. The difference in exact depth, (e.g., 9 to 14m vs. estimated depth of 11 to 16m - US 441) may be attributed to the 3D nature of both the soil/rock layering (e.g., pinnacled limestone), i.e., seismic waves may be refracted and reflected resulting in slower arrival times resulting in deeper or translated anomalies vs. the 2D model. However, the

SPT/CPT results did confirm the general layering of soil and rock and the existence of the anomalies. This is a critical first step in using geophysical methods for site investigations to improve the site characterization with future improvements being 3D analyses.

## CHAPTER 7 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

### **7.1 General**

The focus of this research was the development of seismic testing software and procedures for identifying soil/rock layering and associated variability (horizontal & vertical), and anomalies (e.g., sinkholes) along a 2-D plane below the ground surface. This is very important when designing/constructing/repairing foundations for roads, bridges, etc. Since the analysis technique is new (i.e., Full Waveform Inversion, FWI), the work entailed: 1) developing an algorithm and graphical user interface that could process and analyze 2-D seismic data for a 60ft x 120ft region in 20 to 30 mins on a laptop computer; 2) identifying the resolution of both the layering and anomalies using seismic data both directly over or alongside anomaly/sinkhole; 3) validating the software results through field testing with known/unknown anomalies and layering through numerous invasive testing (SPT/CPT) on 5 FDOT test sites; and 4) training (demonstration, field testing, etc.) of FDOT personnel. A discussion of each follows.

### **7.2 Development of a Fast and Automatic Algorithm of the 2-D Full Waveform Inversion on Synthetic Models**

It was extremely important that a fast and automatic algorithm of Full Waveform Inversion be developed for the seismic 2-D analysis. For instance, FDOT analysis of a sinkhole under a highway must involve lane closures, which due to cost and time disruptions must occur rapidly. A maximum time of 30 minutes was identified as “real time” subsurface analysis, i.e., collect 2-D line of seismic data and analyze it prior to the collection of second line to assist in identifying where additional lines are needed. For required resolution, approximately 1200 cells

with independent assessment of shear (S) and compression (P) wave velocities are required. Therefore, a means to optimize this software is needed to accomplish this goal.

In order to accomplish this, the enhanced algorithm employed: 1) perfectly matched stiffness boundaries, i.e., no wave reflection – this limit size of domain which must analyzed; 2) use of progressive increasing frequency source data along with virtual source and reciprocity of the wave fields to reduce the number of calculations required to estimate the gradient; 3) use of a Gauss-Newton solution strategy- ensuring quadratic rate of convergence; and 4) parallelization of the inversion algorithm to run on multiple cores. All resulted in a run time of approximately 20 to 30 minutes on an 8 core conventional laptop computer.

### **7.3 Development of Interface Software for the 2-D Full Waveform Inversion and Guidelines for Implementation**

The goal of this task was to develop a user interface software and user manual for FWI analysis. The software was developed through the following two tasks: (1) transition of the FWI Matlab code (section 7.2) into C++ and Matlab dynamic link library (dll) format, and (2) developing the graphical user interface (GUI) with new code written in C# (sharp) format.

The GUI written in C# involves a set of screens representing the work flow for seismic testing and subsequent analysis/viewing. A screen setting up the 2-D plane (lateral and vertical depth), along with spacing of the seismic shots and receivers (geophones) is first. Next the data collected for the test line (i.e., shots and receivers) is brought into the software and subsequently conditioned (filter, window – remove noise from signal, etc.). Then the conditioned data is processed (e.g., determine initial velocity profiles from top to bottom) and analyzed, i.e., inverted to obtain independent assessment of shear (S) and compression (P) wave velocities for the 1200 to 1500 cells. The analyses occur over multiple frequencies (i.e., coarse to fine resolution of the velocities). Also displayed for each frequency analysis is the least-squares error. For instance,

the program provides the measured, estimated, and residual (error: measured – estimated) waveforms per iteration. Currently, the program is capable of analyzing shot data and providing estimates of the subsurface P-wave and S-wave velocities in approximately 2.7 minutes per iteration.

Finally, the inversion results can be saved and opened in the program, allowing for future analysis and transfer of analysis files. Additionally, the waveforms, and P-wave and S-wave profiles can be saved as images or pdf for use in reports. A user manual for the FWI software is included in the Appendix.

#### **7.4 FWI's Detection of Embedded Voids Directly Beneath 2-D Seismic Line**

Of great interest was the sensitivity of FWI detection of voids below a 2-D seismic line on the ground surface. A numerical study was performed using variable void sizes and depths to ensure negligible noise influences the final results. This entailed developing a 3-D Finite Difference grid entailing multiple layers (soil and rock) as well as variable size voids and representative depths (1 diameter, 2-D, 3-D, etc.). Subsequently each grid was subject to a point dynamic load (e.g., sledge hammer) and the velocity signals along a straight line on the surface. The recorded waveforms with the inclusion of the void were compared with those from the model without the void to evaluate the sensitivity of the void (energy change). It was found that if the void is embedded at a depth more than 3 diameters, the sensitivity of the void is small (less than 1% energy change). For cases of real data, the energy change due to the void may be in the range of the surrounding noise's energy (0.5 - 1 %), suggesting the interpreted results may not be credible. Recorded waveforms with the void were also be analyzed by the 2-D FWI, and the results suggested the maximum embedded depth that a void can be detected is about 3 void diameters.

## **7.5 Identification of Voids Laterally Spaced from 2-D Seismic Line**

The goal of this task was to investigate any distortion to the FWI on identifying voids filled with air or water, if the source and receivers are on a line adjacent to a void. The FWI software only analyzes a 2-D plane (e.g., 60ft deep and 120ft long) beneath the ground surface. With only a line of data gathered for each inversion, it is difficult to predict the effects of anomalies such as a void that are adjacent to the line. The work again used the 3-D Finite Difference algorithm to produce synthetic data along a 2-D line on the surface for the FWI analysis. The synthetic data was collected at lines on the surface directly above the void as well as at the edge of void (half a diameter from the void center) and multiple diameters from the void. The inverted results showed that: 1) the void is well characterized if the test line is on top of the void (depth of void should be less than 3 diameters, as outlined in Section 7.4), and 2) the void becomes distorted and non-existent if the test line is at least one diameter from the void.

## **7.6 Evaluating Developed FWI Software with Invasive Tests on Five FDOT Test Sites**

The new FWI software was tested on data collected at 5 sites in Florida with various subsurface conditions. The sites were: US 441, Newberry, Gainesville, Tallahassee, and Kanapaha sites. For each test site, multiple seismic test lines were conducted for at least 37 m (120 ft) long, using a 24-channel seismic system. Seismic energy was generated using either a sledgehammer 67 to 89 kN (15 to 20 lbs) or a PEG (PEG-40 Kg model). The sites had both known and unknown anomalies and varying soil/rock layering. Based on each of the seismic surveys, SPT borings or CPT soundings were performed adjacent or near seismic anomalies.

A review of the invasive test results revealed that the seismic analysis did an excellent job of identifying soil & rock layering (examples are top of limestone depths at Kanapaha and Newberry), identification of unknown voids (Newberry), and the extent of existing voids (examples are US 441 and Gainesville retention pond). However, predicted depths of voids were

deeper than their actual depths (examples are Newberry and US441), mostly due to 3-D effect that cannot be completely addressed by 2-D analyses. In addition, the field seismic test lines approximately one diameter from a known void were visible on the FWI scans, but were distorted. Subsequently testing on each side of the analyzed line resulted in improved identification of the void's location (i.e., depth and diameter).

Finally, training of FDOT personnel for both seismic testing and the use of the new FWI software was provided. Based on their use and review, improvements with the input and output screens and development of a user manual (appendix) were accomplished.

### **7.7 Recommendations for Further 3-D FWI Development**

The field experimental results presented in Chapter 6 showed the usefulness of the 2-D FWI method in locating embedded voids and characterizing variable soil/rock layers. However, the 2-D FWI method requires seismic data to be acquired right on the top of a void, and thus multiple test lines are usually needed because of the unknown location of the void.

Also off-line voids may be distorted due to 3-D effects. To overcome the limitations of the 2-D approach, it is recommended to develop a 3-D FWI method. Specifically, seismic wave fields will be acquired using sensors and sources located in uniform 2-D grids on the ground surface, and then inverted for the extraction of 3-D subsurface wave velocity profiles ( $V_s$  and  $V_p$ ).

The 3-D waveform analysis aims to increase accuracy and resolution of the resulting wave speed velocity profiles, and minimizing field testing effort. The potential advantages of the 3-D approach include *(i)* embedded voids could be well detected, as signals reflected and refracted from the voids at different angles are utilized in 3-D analyses; *(ii)* off-line voids, which are distorted by the 2-D analysis because of the 3-D effect, will be readily identified in the 3-D analysis since the 3-D analysis models both in- and out-of-plane waves; and *(iii)* the 3-D analysis requires less field testing efforts to detect a void than the 2-D analysis. While the 2-D FWI

approach requires seismic data to be acquired right on the top the void, the 3-D approach only requires data to be acquired in the vicinity of the void, since signals reflected and refracted from the void at different angles can be recorded by a 2-D grid of sensors on the ground surface.

## REFERENCES

- Bohlen, T. (2002). Parallel 3-D viscoelastic finite difference seismic modelling. *Computers & Geosciences*, 887-899.
- Brenders, A. J. and Pratt, R. G. (2007). Full waveform tomography for lithospheric imaging: Results from a blind test in a realistic crustal model: *Geophysical Journal International*; 168(1): 133-151.
- Busch, S., Van der Kruk, J., Bikowski, J., and Vereecken, H. (2012). Quantitative conductivity and permittivity estimation using full-waveform inversion of on-ground GPR data. *Geophysics*, 77(6): H79–H91.
- Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 705-708.
- Dong, L., She, D., Guan, L., & Ma, Z. (2005). An eigenvalue decomposition method to construct absorbing boundary conditions for acoustic and elastic wave equations. *Journal of Geophysics and Engineering*, 2(3):192-198.
- Florida Senate, Committee on Banking and Insurance. (2010). *Issues Relating to Sinkhole Insurance* (Interim Report 2011-104, Dec 2010).
- Giroux, B. (2012). Performance of convolutional perfectly matched layers for pseudospectral time domain poroviscoelastic schemes. *Computers & Geosciences*, 45: 149-160.
- Groos, L., Schäfer, M., Forbriger, T., and Bohlen, T., (2014). The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves. *Geophysics*, 79(6): R247–R261.
- Guddati, M., and Lim, K. (2005). Continued fraction absorbing boundary conditions for convex polygonal domains. *Numerical Methods in Engineering*, 66(6):949-977.

- Kallivokas, L.F., Fathi, A., Kucukcoban, S., Stokoe, K.H., Bielak, J., Ghattas, O. (2013). Site characterization using full waveform inversion. *Soil Dynamics and Earthquake Engineering*; 47:62–82.
- Komatitsch, D., and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 155-167.
- McVay, M., and Tran, K. (2013). *Detection of Sinkholes or Anomalies Using Full Seismic Wave Fields* (Florida Dept. of Transportation Research Report BDK75-977-66). University of Florida, Gainesville, FL.
- Moler, C. (2007). Parallel MATLAB: Multiple Processors and Multiple Cores. Retrieved from MathWorks: <http://www.mathworks.com/company/newsletters/articles/parallel-matlab-multiple-processors-and-multiple-cores.html>.
- Peng, C., and Toksoz, M. (1995). An optimal absorbing boundary condition for elastic wave modeling. *Geophysics*, 60(1):296-301.
- Sheen, D.H., Tuncay, K., Baag, C. E. and Ortoleva, P. J. (2006). Time domain Gauss–Newton seismic waveform inversion in elastic media: *Geophysical Journal International*, 167: 1373-1384.
- Sochacki, J., Kubichek, R., George, J., Fletcher, W., & Smithson, S. (1987). Absorbing Boundary Conditions and Surface Waves. *Geophysics*, 60-71.
- Tikhonov, A. N., and Arsenin, V.Y., (1977). *Solutions of ill-posed problems*. V. H. Winston & Sons, Washington, D.C.
- Tran, K. T., and Hiltunen, D. R. (2012a). One-dimensional inversion of full waveform using genetic algorithm: *Journal of Environmental and Engineering Geophysics*, 17: 197-213.

Tran, K. T., and Hiltunen, D. R. (2012b). Two-dimensional inversion of full waveform using simulated annealing, *Journal of Geotechnical and Geoenvironmental Engineering*, 138(9): 1075-1090.

Tran, K.T., McVay, M., Faraone, M., and Horhota, D. (2013). Sinkhole Detection Using 2-D Full Seismic Waveform Tomography, *Geophysics*, 78 (5), R175–R183.

Tran, K. T., and McVay, M. (2012). Site Characterization Using Gauss-Newton Inversion of 2-D Full Seismic Waveform in Time Domain, *Soil Dynamics and Earthquake Engineering*; 43: 16-24.

APPENDIX  
FWI INVERSION SOFTWARE USER MANUAL

# Full Waveform Inversion Program User Manual

## 1.0 Introduction

Welcome to FWI. This program applies a full waveform inversion process to produce shear and compression wave velocities of the medium. It is used in large scale applications such as void detection under roadways. The application includes several features such as:

- Modifiable parameters
- Easily import and modify data
- Generate an initial model
- Run the inversion

## 2.0 Parameters

The first aspect to the program is the parameters. These parameters are based on various conditions such as the receivers, shots, and material properties. Table 1 lists and describes each parameter that can be changed.

Model	
nz	- Number of grid points in z direction
dx	- Spacing of grid points in x direction
dz	- Spacing of grid points in z direction
Pad	- Number of grid points in boundary pad layer
Material properties	
Nu	- Nu of material
Vs Max	- Maximum shear wave velocity of material
Vs Min	- Minimum shear wave velocity of material
Vp Max	- Maximum compression wave velocity of material
Vp Min	- Minimum compression wave velocity of material
Receiver Properties	
Start	- Physical start location of receivers
Finish	- Physical end location of receivers
Spacing	- Spacing between receivers
R_rm	- Remove receivers that are within this distance from shot location
R_nf	- Distance from shot that is near field
Shot Properties	
Start	- Physical start location of shots
Finish	- Physical end location of shots

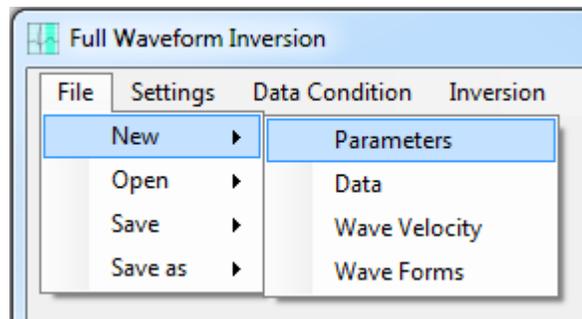
Spacing	-	Spacing between shots
Time Properties		
tmax	-	Maximum time
t0	-	Time to impulse
dts	-	Length of time interval

Table 1

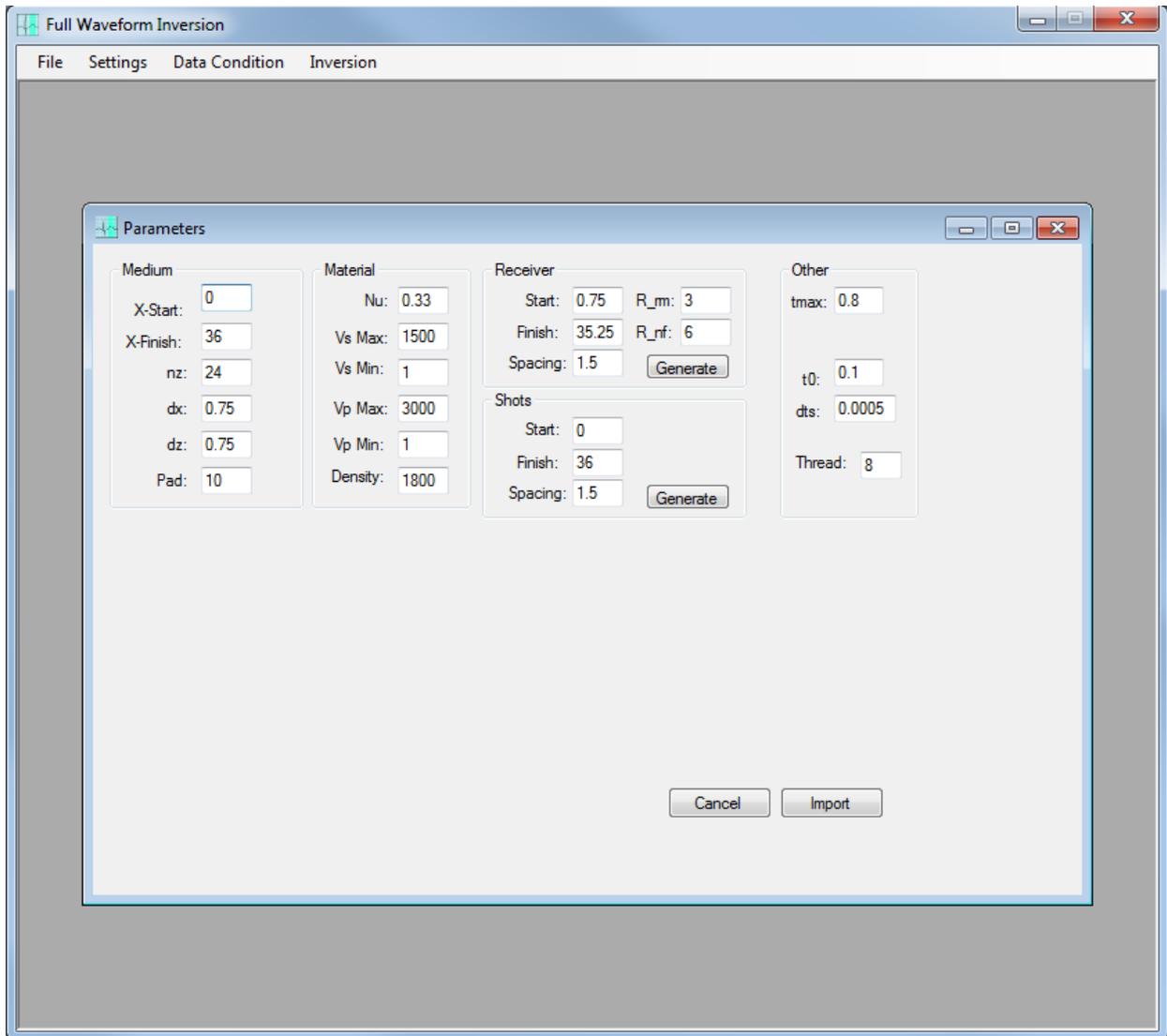
To start, import the desired parameters based on the data gathered in the field. There are two ways to do this.

## 2.1 New: Parameters

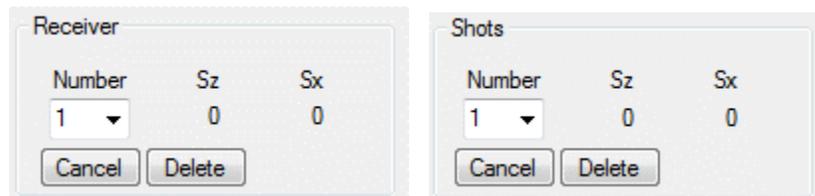
The first option for importing parameters into the program is to create new parameters for data analysis. To do this, click on File>New>Parameters as seen below.



The following screen will come up.



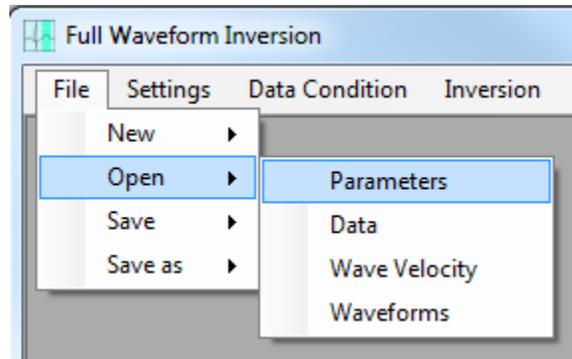
Input parameters based on gathered data by selecting the value to the right of the parameter name. The receiver and shot locations are produced by creating an array. The option to remove specific locations in these arrays is available by clicking “Generate”. This produces a new section with the number in the array, x location, and z location as seen below.



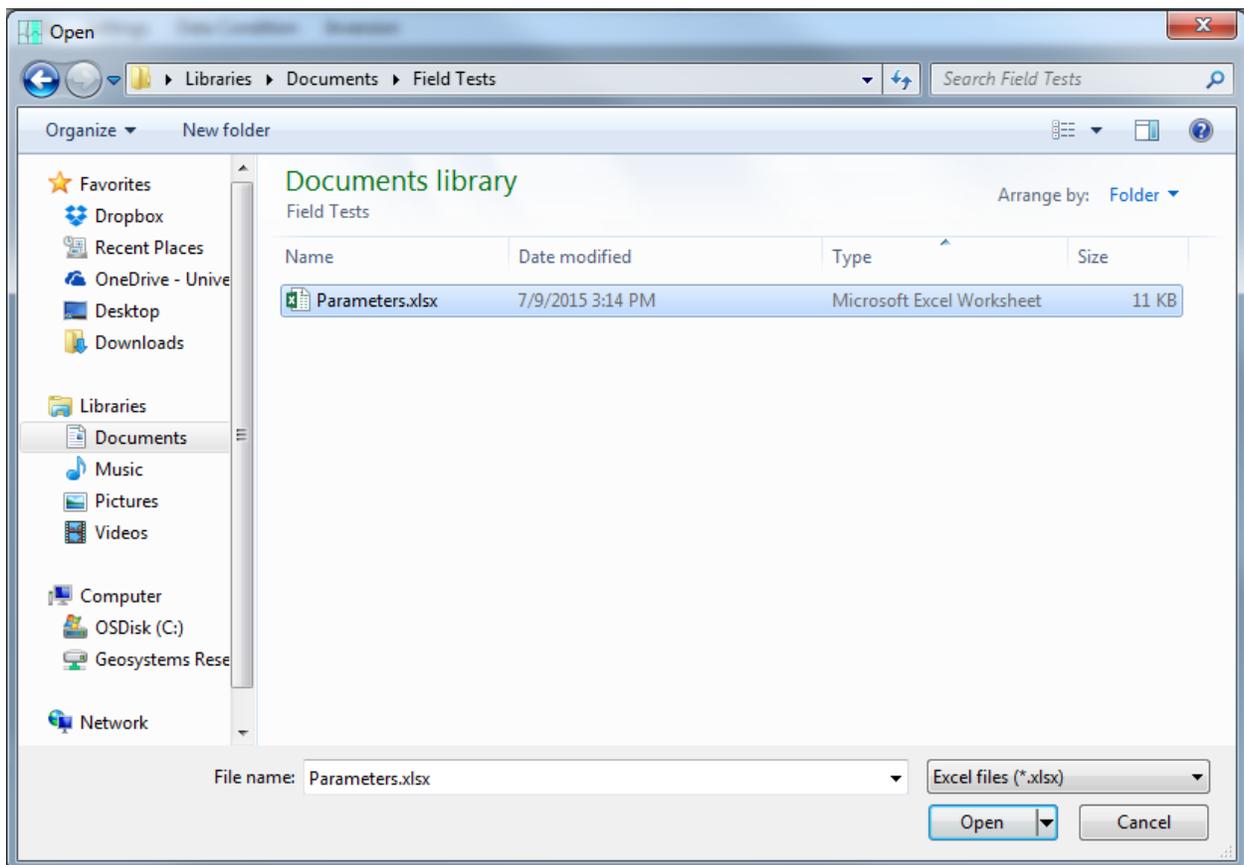
Select the shot or receiver number and delete it if desired. Cancel will return the program back to the original parameter page. Clicking “Import” will load the parameter values into the program. The status bar in the bottom left will say “Parameters imported” if they are successfully imported.

## 2.2 Open: Parameters

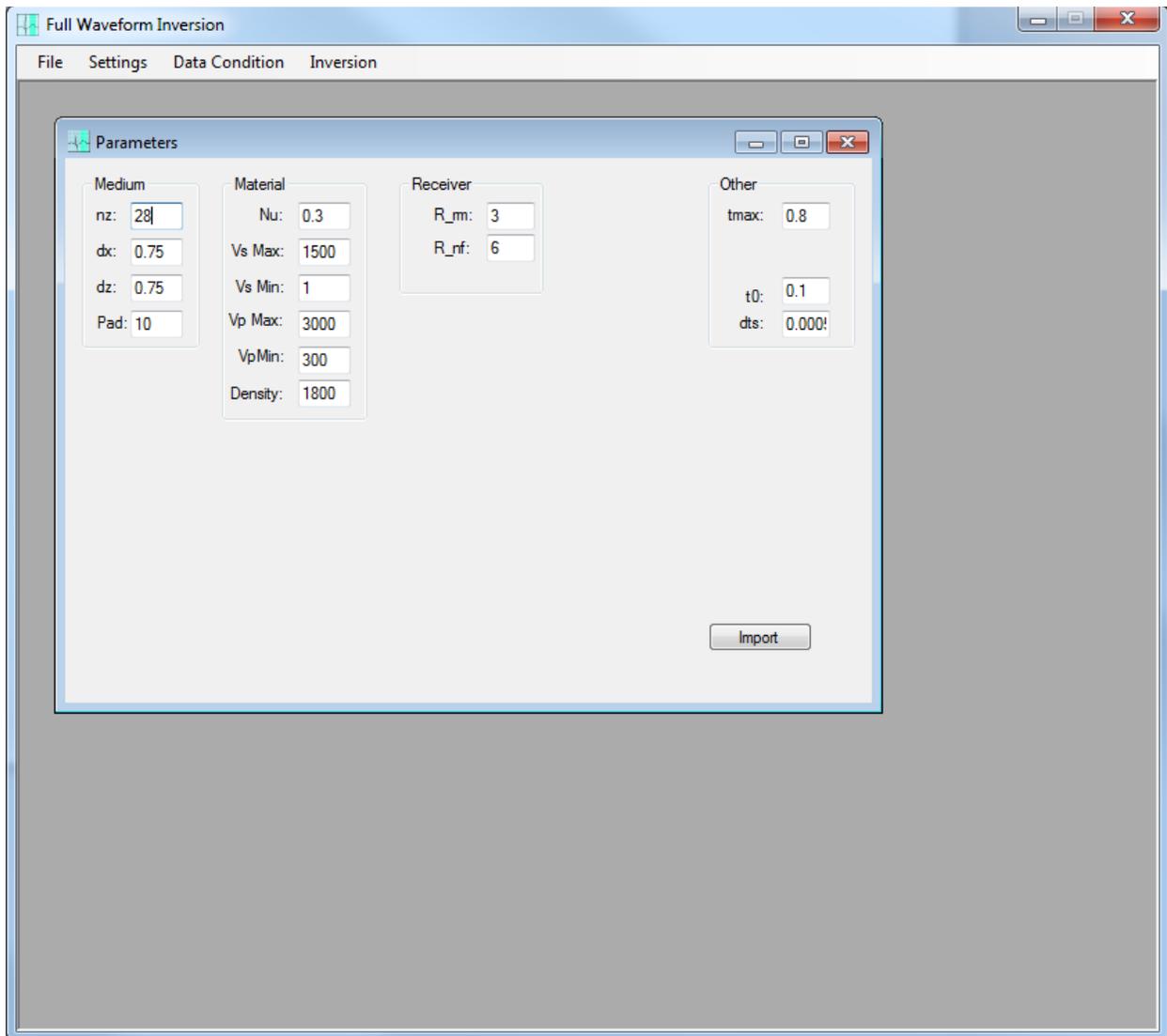
The second option for importing parameters is to open an already existent parameter file by clicking File>Open>Parameters as seen below.



This will produce the following screen.



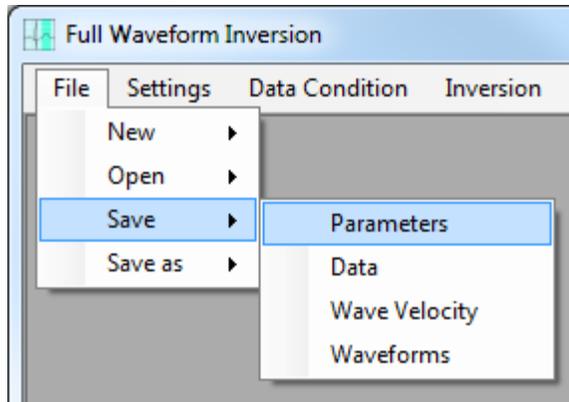
Navigate to the desired location of the parameter file and select the parameter file to open it. This will bring the program to the following page.



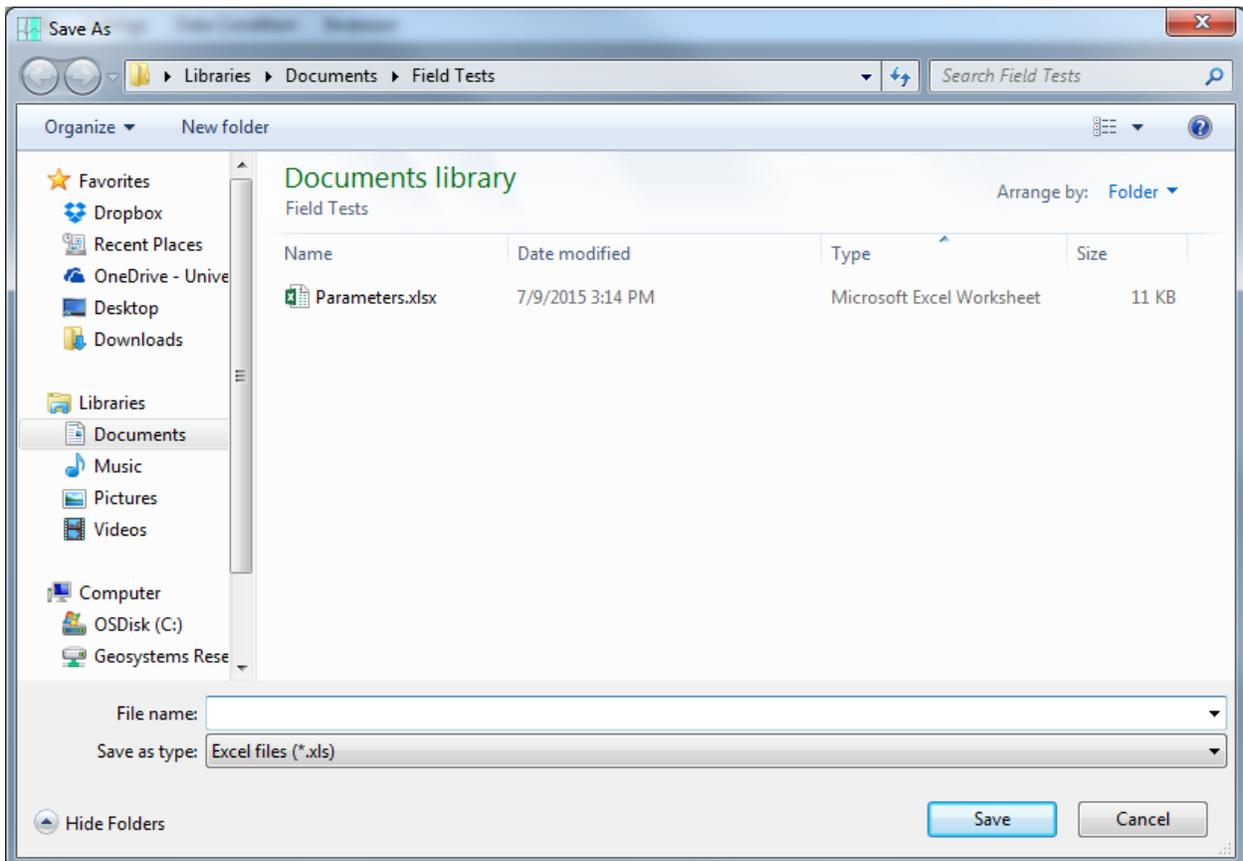
The status bar in the bottom left of the screen will say “Parameters imported” if the parameter file was successfully import. If needed, modify the current values and reimport them by selecting “Import” in the bottom right of the window.

### 2.3 Save: Parameters

To create a parameter file that can later be imported select File>Save>Parameters.



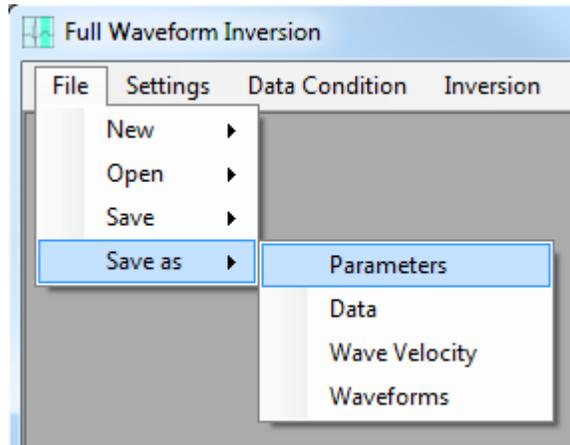
If a parameter file save location has already been selected or a parameter file has been imported, the save function will save the currently imported parameters over that name and location. If none have been selected then the following dialogue box will come up.



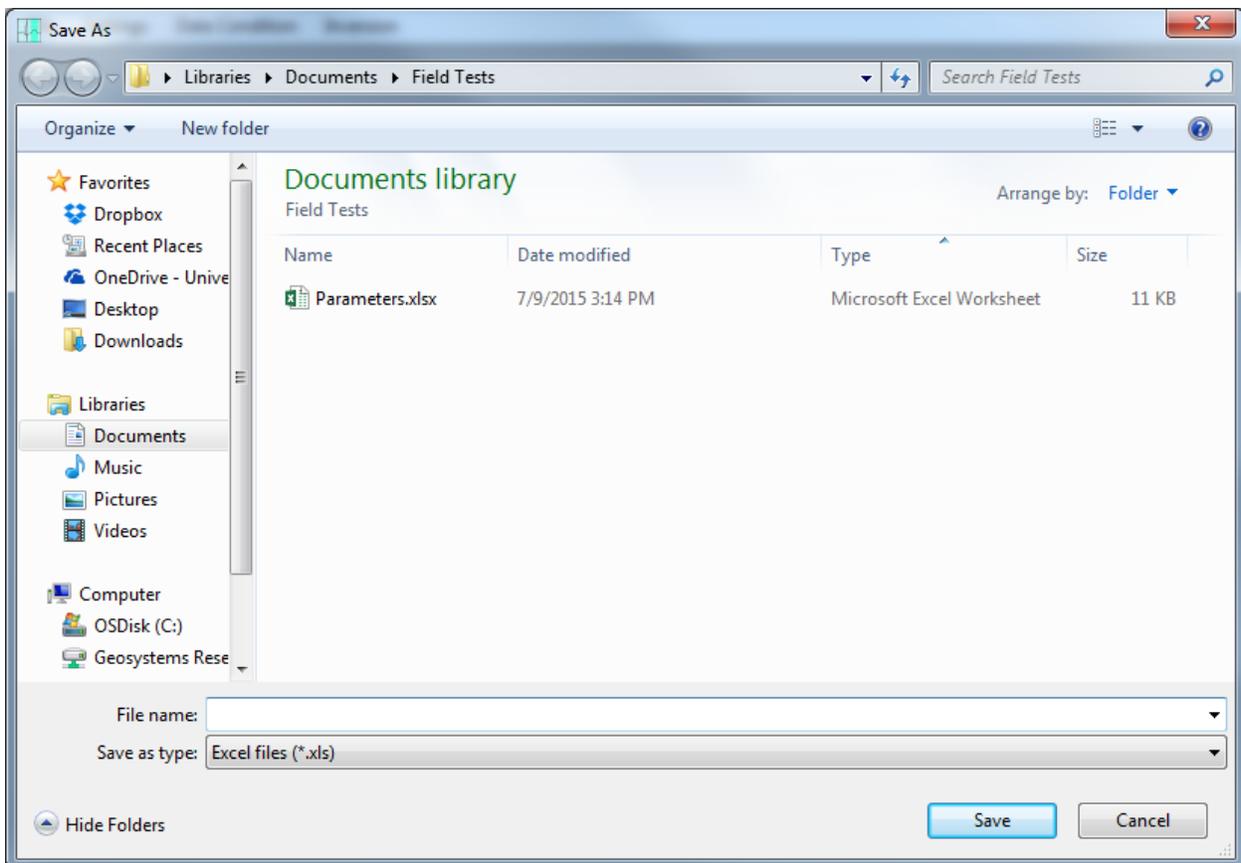
Navigate to the desired location and input the file name. Select “Save” to save to the name and location.

## 2.4 Save as: Parameters

The option to save the parameters to a new file location is available by selecting File>Save As>Parameters.



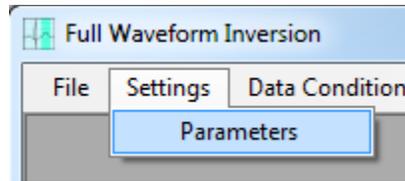
The following dialogue box is produced to allow for selection of a location and file name for the parameters.



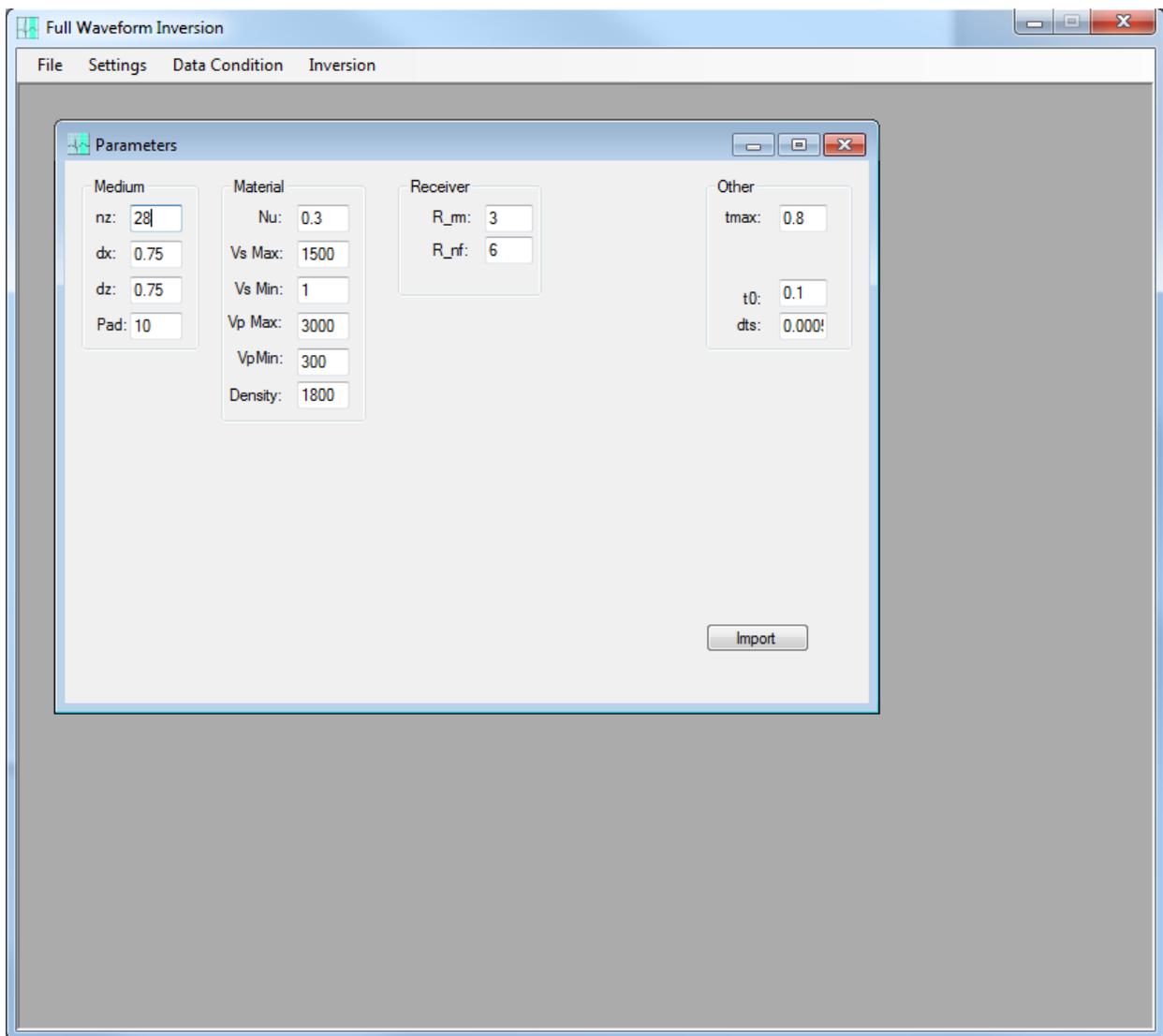
Select the desired location and name and click “Save”.

## 2.5 Change Parameters

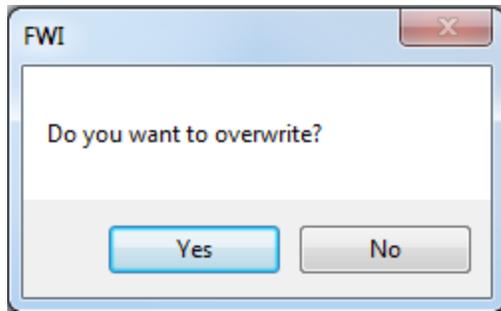
There is the option to change the parameters that have been imported into the program. This can be done by clicking Setting>Parameters.



Once selected, the following page comes up.



Modify the desired values and select import. The program will ask for permission to overwrite the existing variables within the program.



Select “Yes” to overwrite them and “No” to cancel. The status bar will say “Variables overwritten” if successful.

### 3.0 Data

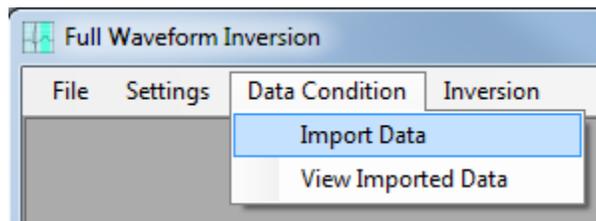
The data that is imported is used to verify the forward model being adjusted in the program. The data for each shot must be imported and modified so that accurate results are produced.

#### 3.1 Importing Data

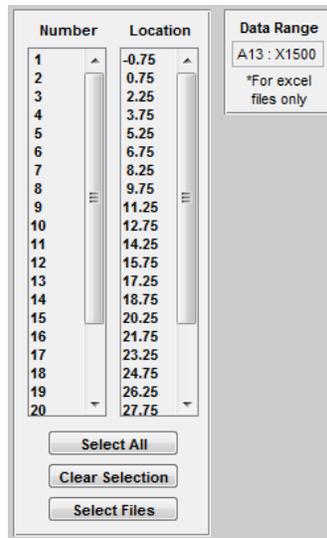
There are two options when importing data sets. Import the data files created for each shot or import an already saved data file.

##### 3.1.1 Import Shot Files

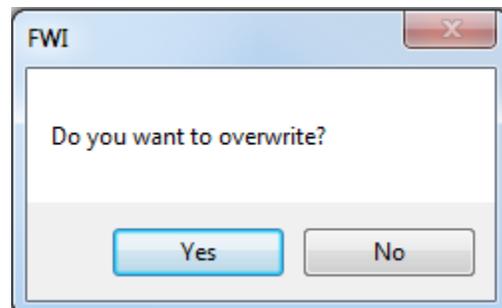
To import a shot file select Data Condition>Import Data.



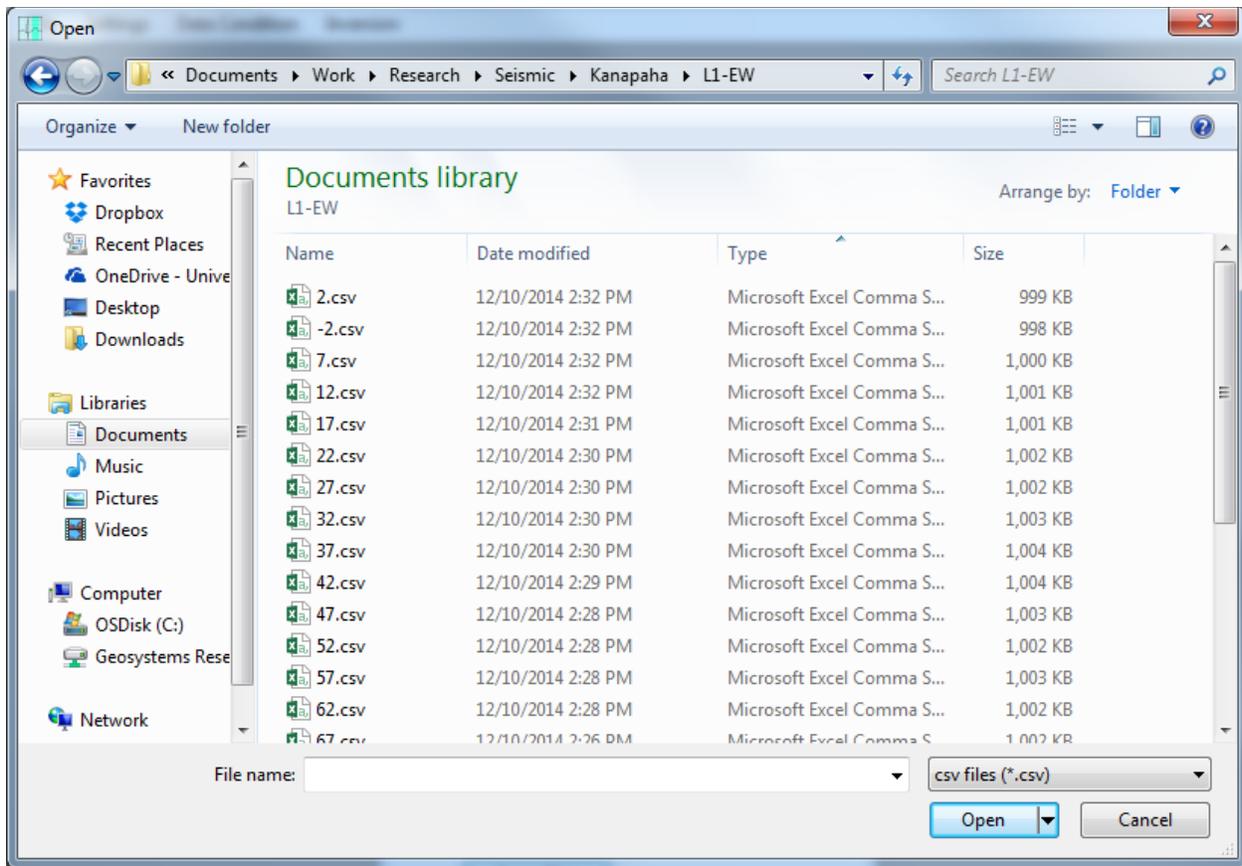
The follow controls will come up.



Select the shot information to be used by clicking, ctrl+clicking, or shift+clicking on the shot number/locations to be imported. Use the “Select All” button to select all of the shot locations and the “Clear Selection” button to deselect the locations. If the data is being imported from excel files indicate the data range in the file using the data range box on the right side of the window. Lastly, click “Select Files” to produce a file selection window. If data is already imported for a selected location, a popup window will appear asking to overwrite the existing data.



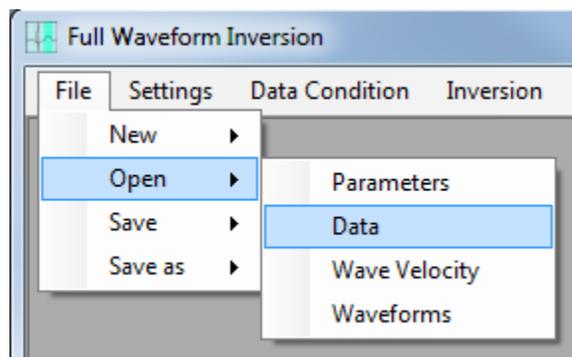
Select “No” to cancel and “Yes” to overwrite the data in those locations. A file selection window will appear if there were no conflicting locations or “Yes” is selected.



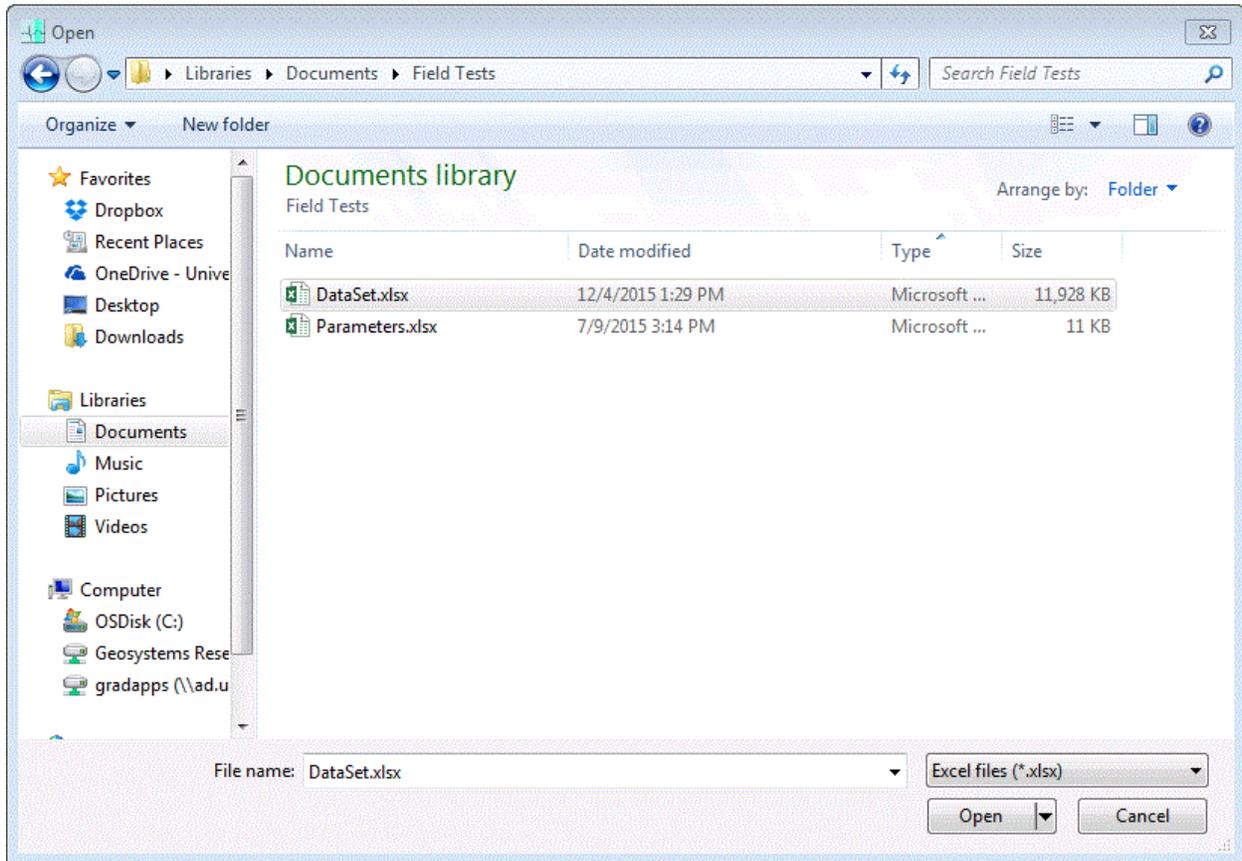
Select multiple files by clicking, ctrl+clicking, or shift+clicking the desired files. The files must be named in numerical order according to their location. In the example above, the names are based on their number in the array. They can also be named using their numerical location. As long as the files are named in order, the files being imported don't have to be selected in order. The program will match them with their appropriate location. Once the files are chosen, click "Open" to import the shots into the program. The imported data viewing window will appear (See section 3.3).

### 3.1.2 Open: Data

The second option is to load an already existing data set by clicking File>Open>Data.



If there is already data imported the program will ask for permission to overwrite the data. This produces a file selection window as seen below.



The data is imported once the file is selected. The imported data window is shown if the data is successfully imported (See section 3.3). All of the shots imported into the program can be modified here. Further details on modifying data are shown in section 3.2.

## 3.2 Modifying Data

Modifying data allows for improved results. The data can be altered in several ways including frequency filtering, time windowing, and removing bad receivers.

### 3.2.1 Frequency Filtering

The various frequencies contained in the data can be filtered out to allow for noise to be removed. This is done using four values in an array. Figure 1 shows a plot with the dampening value as a function of the filtering frequency.

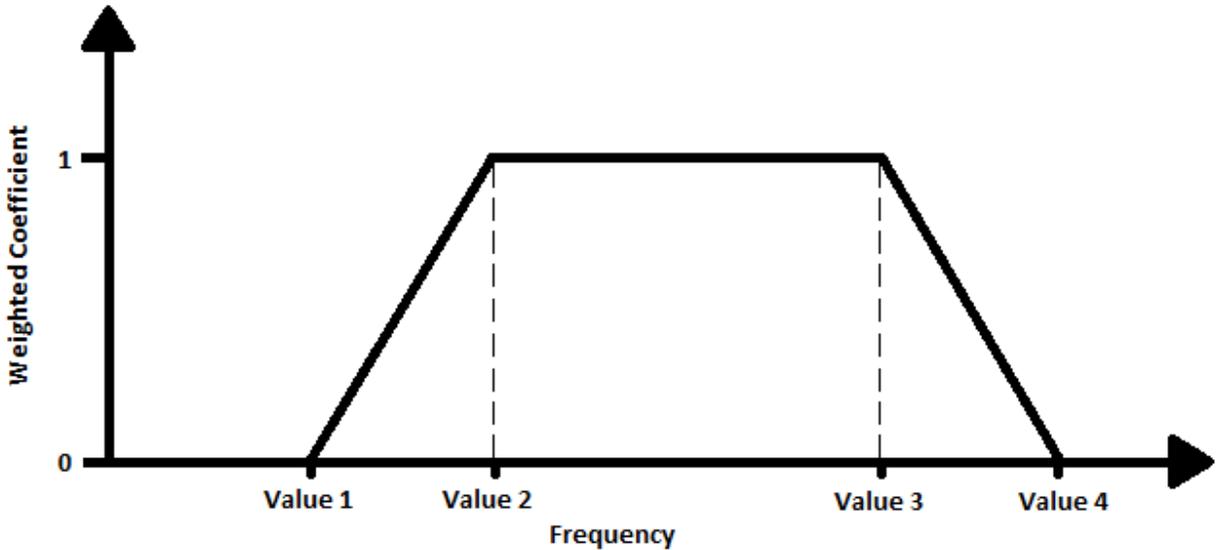


Figure 6

Any frequencies below Value 1 and above Value 4 will be removed. Frequencies between Value 1 and Value 2 are decreased by a factor between 0 and 1 based on its position between the two values. Any frequencies between Value 2 and Value 3 are unweighted. Frequencies between Value 3 and Value 4 are weighted based on their position between the two. These filtering values can be adjusted in the data viewing window shown when looking at the imported data (See section 3.3). The filter section will look like this on the viewing window:



Change the filtering frequencies by entering four values with at least one space between each. The other option is to use the arrows which will create an array that has values spaced five apart and increase or decrease based on the direction selected.

### 3.2.2 Time Windowing

Each receiver collects data for the same amount of time. The time in which the main wave isn't perturbing may contain noise that will throw off the results. The time windowing option allows for the data to be reduced to a desired amount of time both before and after the largest magnitude of the wave. This is done by changing two values in an array, the first of which is the time used before the wave while the second is length of time after the wave that the data is still used.

Figure 2 shows two sets of data from the same receiver.

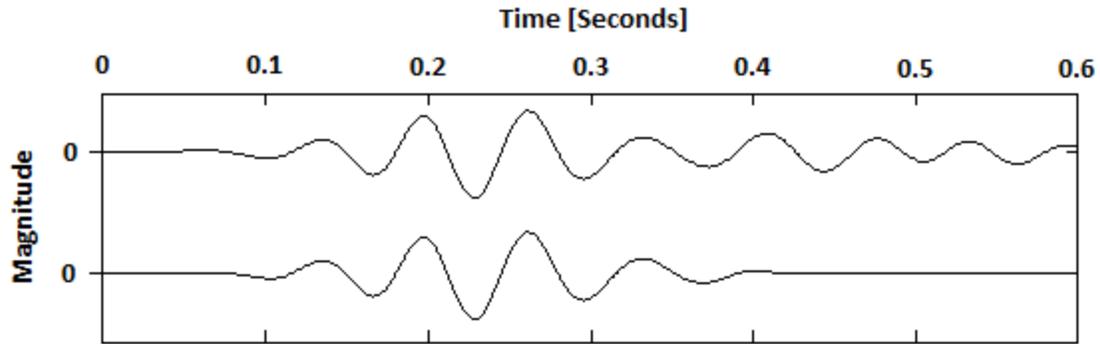
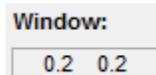


Figure 7

The top plot is data without any windowing. There is a noise towards the end of the data which will negatively affect results. Through the use of windowing, the noise can be cut out. The lower plot used windowing values of 0.2 and 0.2. Similarly to the filtering array, the windowing array can be adjusted in the imported data window (See section 3.3). On each of these pages the windowing section will look like this:



Adjust the windowing time by putting in two time values on any of data viewing windows with at least one space between them.

### 3.2.3 Maximum Time

The amount of time that is used from the gathered data can be adjusted to reduce the computation time and improve results. Figure 3 shows the application of reducing the maximum time on the same set of data.

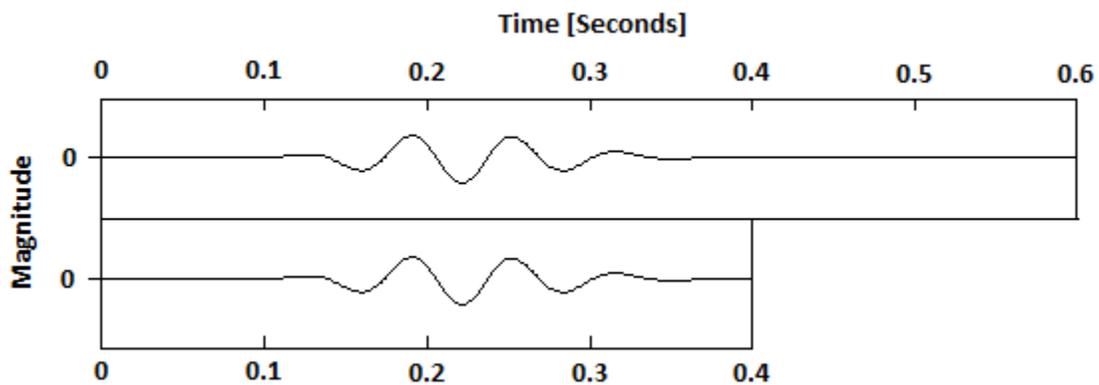


Figure 8

The original data, found on top, had a total time of 0.6 seconds. As seen in the figure, there isn't any signal after 0.4 seconds. Therefore, the data can be reduced by shortening the time to 0.4 as seen in the bottom set of data. This can be done in the program by viewing the imported data (See section 3.3). The option to alter the time is seen below:

tmax:

Alter the value in the box to change the maximum time. Reducing the time will remove the data at any point that is past that length of time. If the time is increased the program will input zeros for the values that are added on.

### 3.2.4 Removing Receivers

A bad receiver can result from various problems in the data collection process. The bad set of data can be removed to nullify the affect it has on the final results. This can be done in the imported data window in the box seen below.



Select the bad receiver in the graph by click on it. Use ctrl+click or shift+click to select multiple receiver plots. Any receivers to that are selected have bold plot lines. Press escape or push the "Clear" button to deselect the any receivers. To remove the selected receivers press delete or push the "Remove" button.

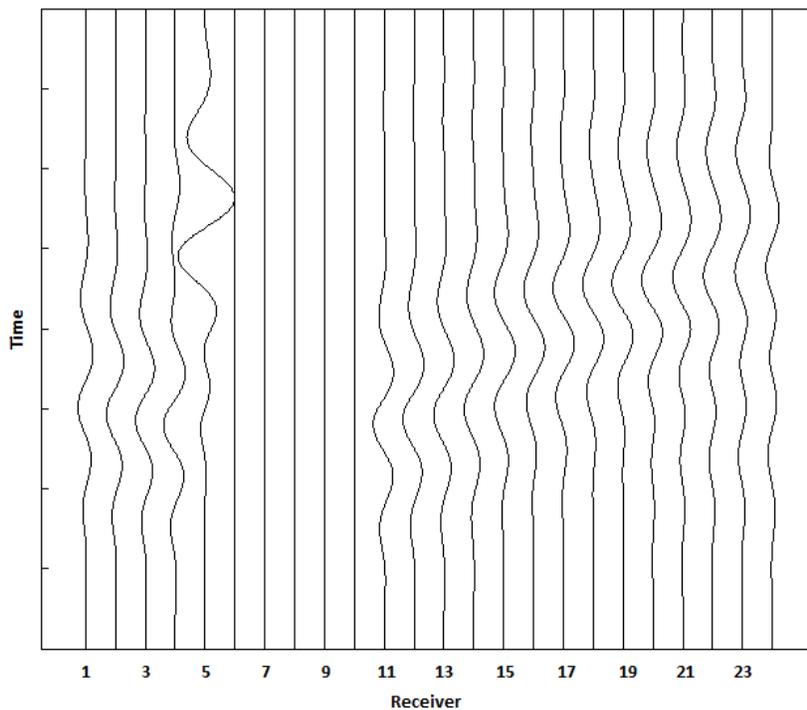


Figure 9

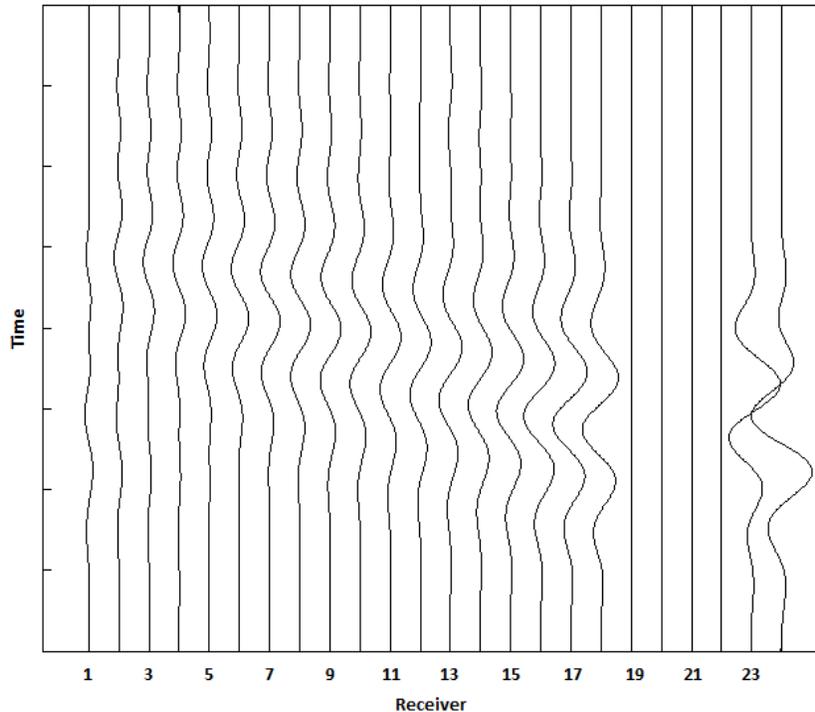
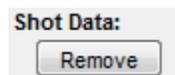


Figure 10

In Figure 4, receiver 5 contains bad data (mismatch the peak) and needs to be removed. The data in Figure 5 has several bad receivers. Receiver 24 is a bad receiver due to the mismatch in the location of maximums when compared to receiver 23. Receiver 23 is good due to the symmetry with receiver 18 across the shot location. Receiver one is offset quite a lot from the surrounding receivers. It is extremely important to remove all poor channels before analysis, as these channels will control the least-squares error and produce artifacts (false anomalies).

### 3.2.5 Removing a Shot

Sometimes there are too many bad receivers in a data set to use it. The option to remove a shot can be found in the imported data menu (See section 3.3). It is found below the remove receiver option shown below.



If it is decided that a shot needs to be removed simply navigate to that shot using the location/number popup menus or the arrows above those menus. Then click remove under “Shot Data:” If the data for that location needs to be imported again use the import shot option (See section 3.1.1).

### 3.2.6 Flip the Data

Sometimes the data is backwards with respect to the receiver number due to equipment setup (geophones are connected to seismograph in reverse order, or reverse shot locations). This can

throw off results produced by the program. To fix the data simply click the “Flip Data” button found on any data viewing window. An example of data that needs to be flipped is found in Figure 6.

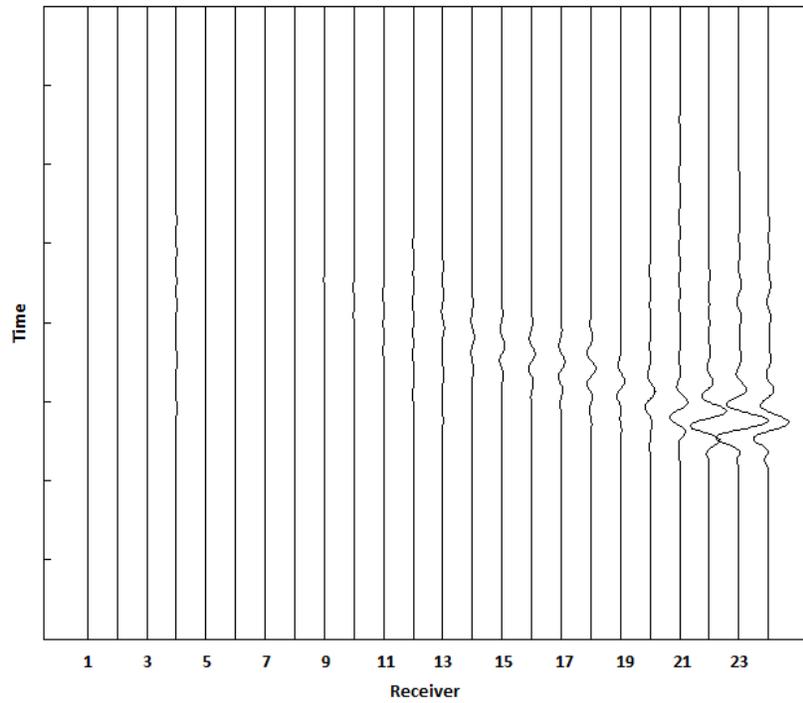


Figure 11

This data is from a shot whose location is near the first receiver. The current data shows that the shot is near the last receiver. Flipping the data produces the correct orientation shown in Figure 7.

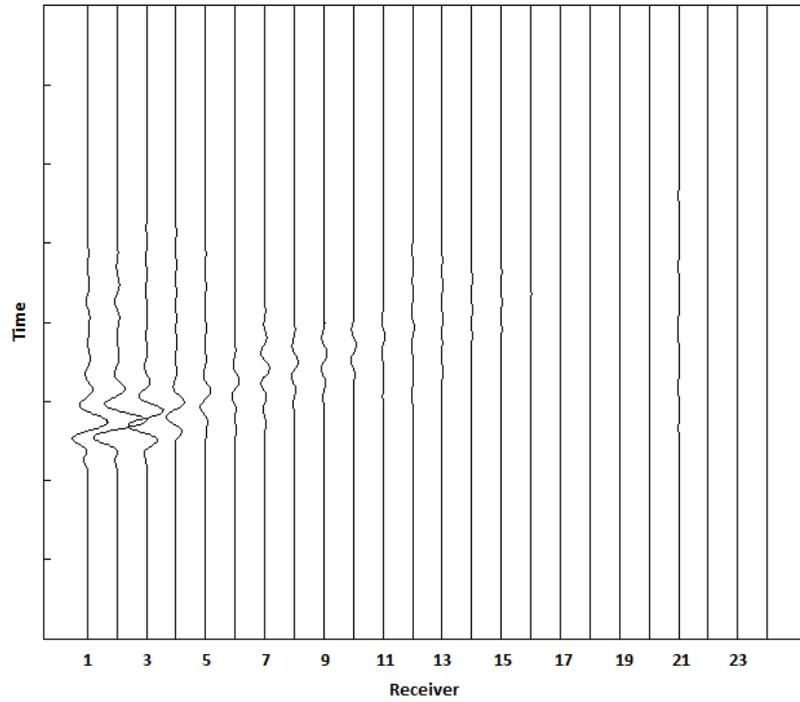


Figure 12

### 3.2.7 Changing Central Frequency

The central frequency is set based on the data that is imported into the program. To change the central frequency, navigate to the data viewing page (See section 3.3). Set the plot to the spectral setting in order to view the spectral analysis of the shot data. Pick a central frequency that lines up vertically with the peaks of the spectral analysis plot. An example can be seen in Figure 8.

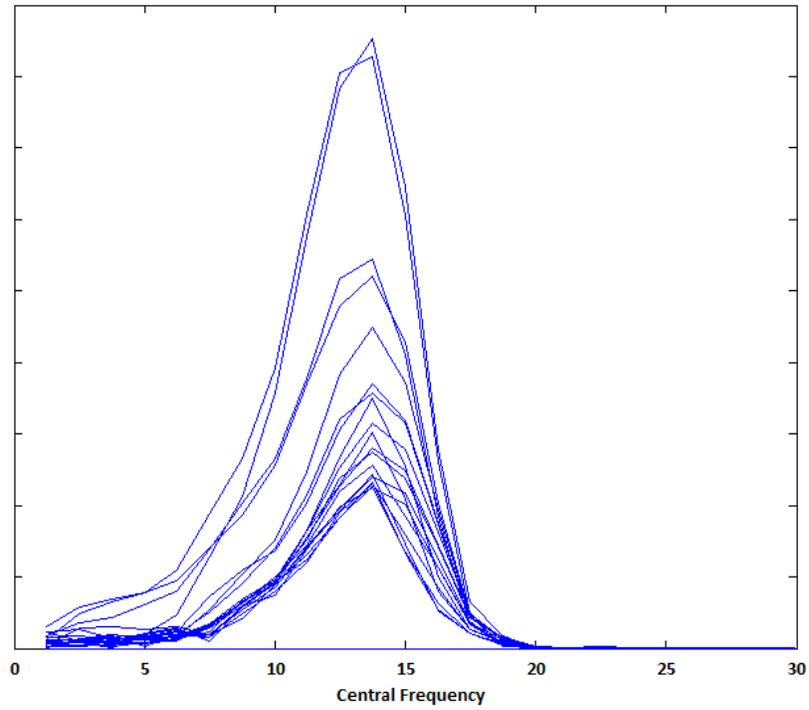


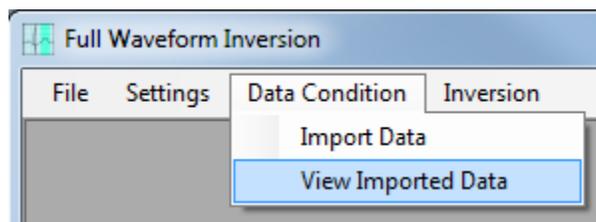
Figure 13

The central frequency for this data set should be set to 14. This is done by modifying the value found in the data viewing windows as seen below.

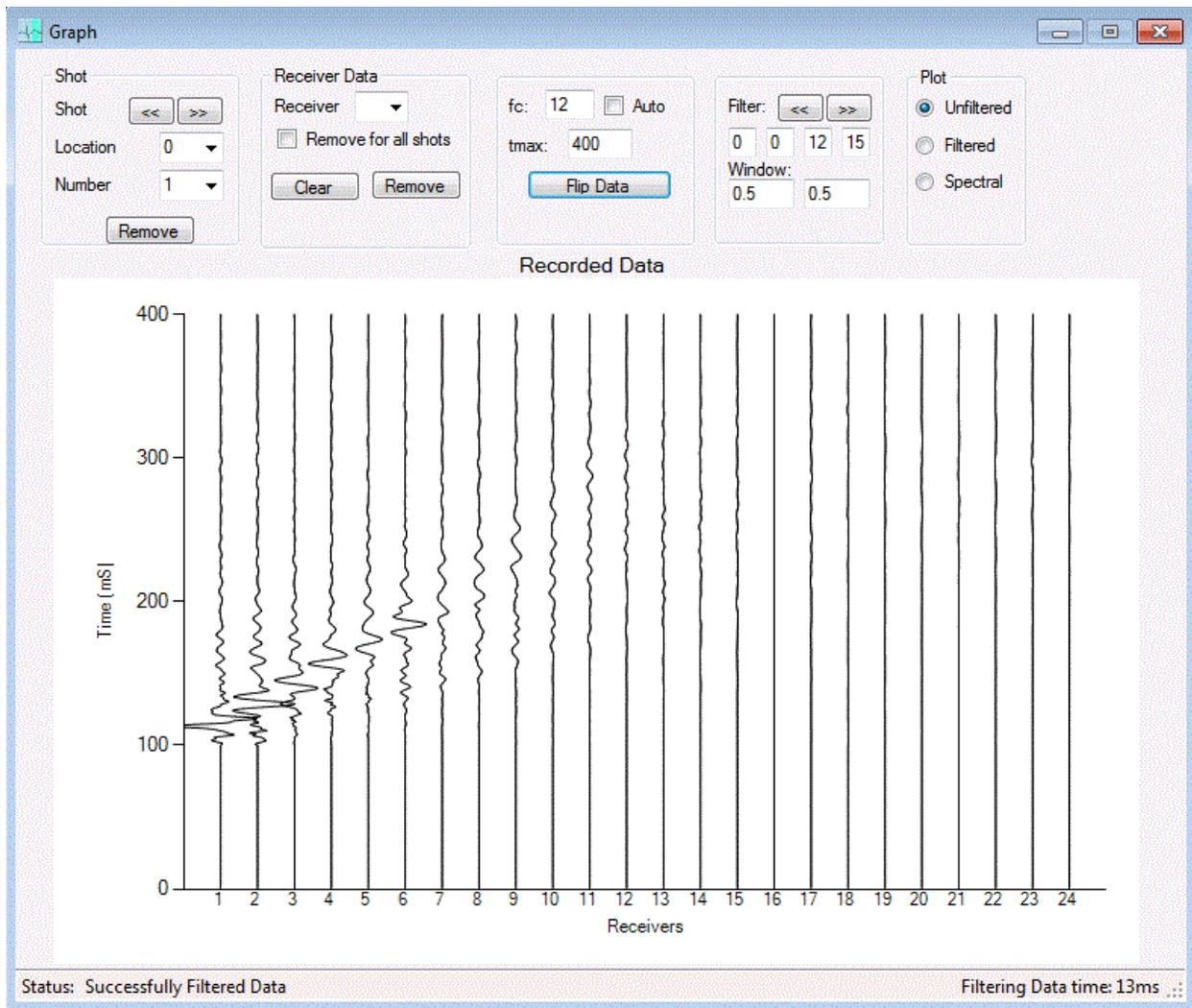


### 3.3 Viewing Imported Data

To view the data that has been imported into the program select Data Condition>View Imported Data.



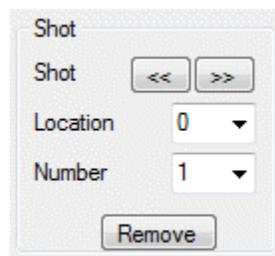
This will produce a window to view the data.



There are multiple data plots and the option to modify variables such as the central frequency, filtering, windowing, and remove any bad receivers (See section 3.2). Any changes made are automatically saved within the program.

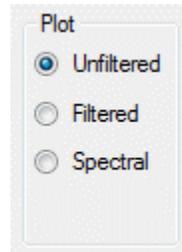
### 3.3.1 Change Shot Being Viewed

To change the shot data in the plot, use either of the dropdown menus or the arrows located in the shot box in the data viewing window.



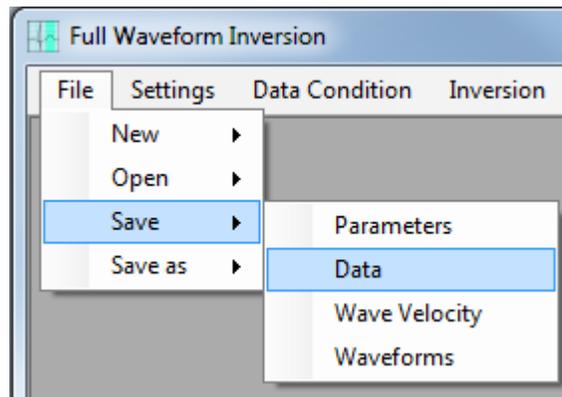
### 3.3.2 Change Plot

There are three plots that allow data to be viewed. The first, unfiltered, is the data without any filtering, windowing, or local receivers removed. Second is the filtered data which includes filtering, windowing, and removal of local receivers based on the R\_rm parameter. Last is the spectral plot which completes a spectral analysis on the shot data being viewed. To change the plot select the plot point next to the desired plot title in the “Plot” box found in each of the data viewing windows.

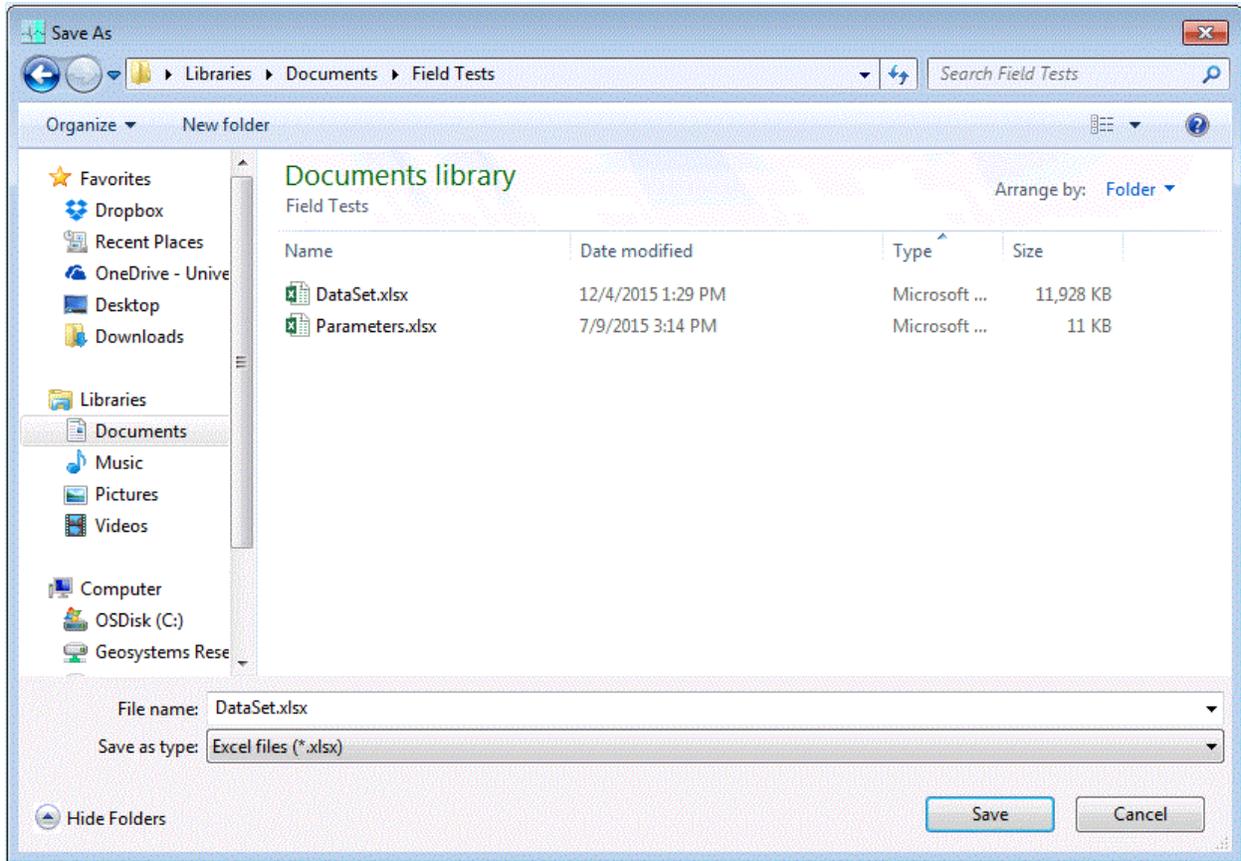


### 3.4 Save: Data

To create a data file that can later be imported select File>Save>Data.



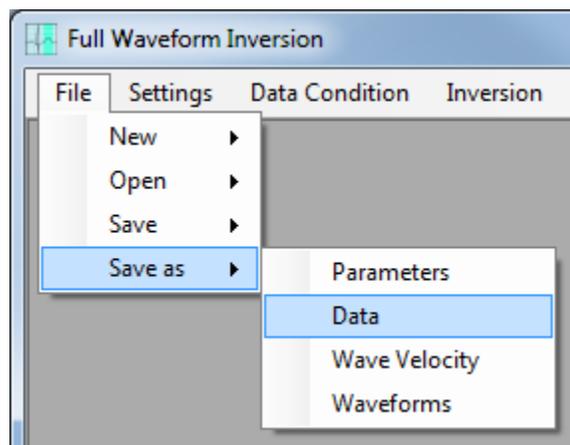
If a data file save location has already been selected or a data file has been imported, the save function will save the currently imported data over that name and location. If none have been selected then the following dialogue box will come up.



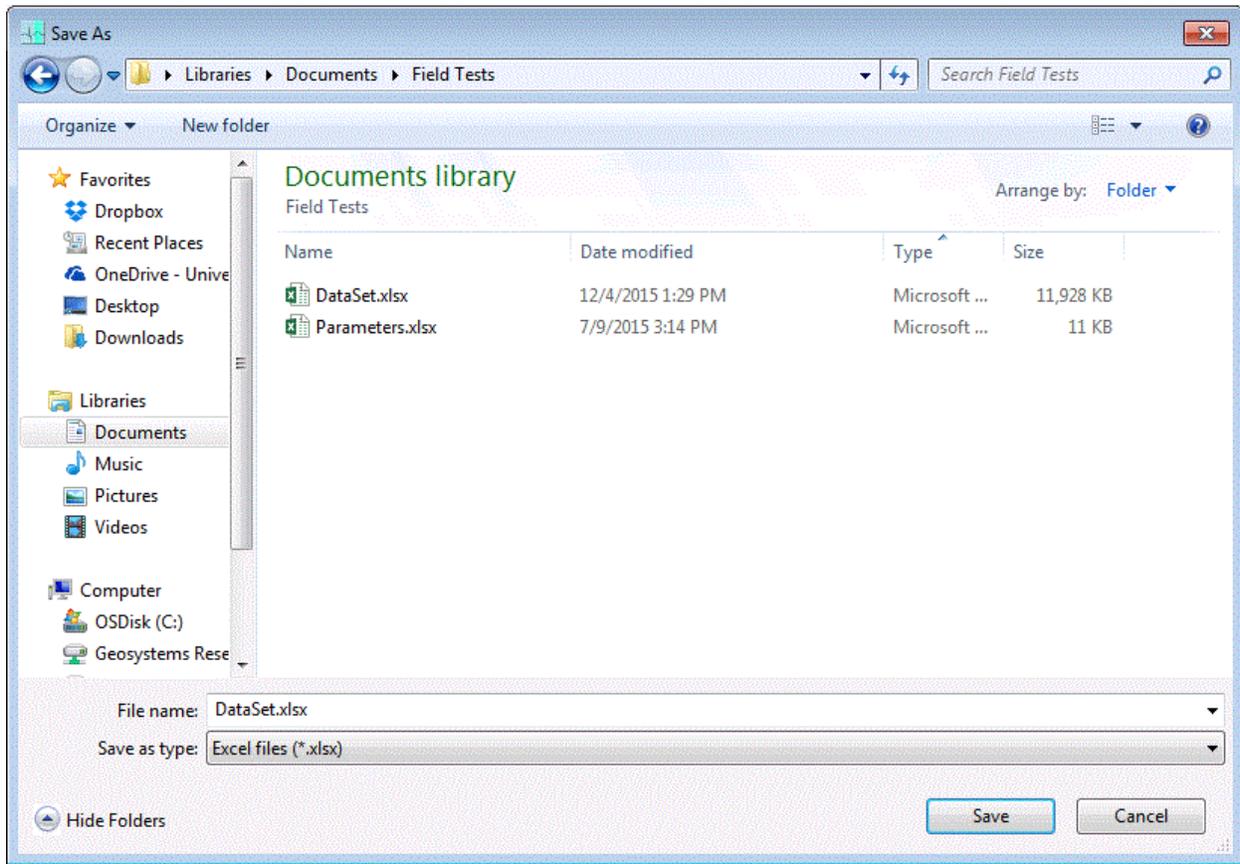
Navigate to the desired location and input the file name. Select “Save” to save to the name and location.

### 3.5 Save as: Data

There is also the option to save the data to a new file location using File>Save As>Data.



The following dialogue box is produced to allow for selection of a location and file name for the data.



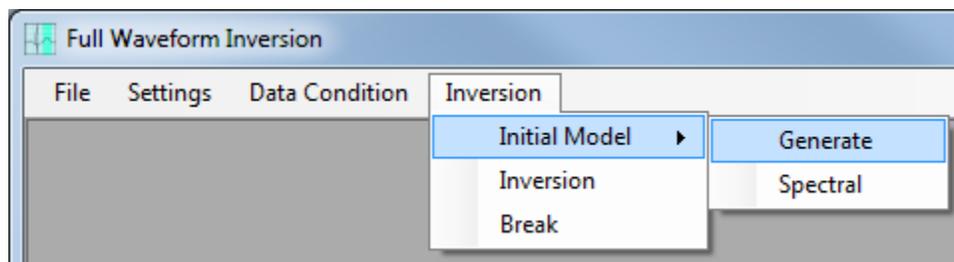
Select the desired location and name and select “Save”.

## 4.0 Wave Velocities

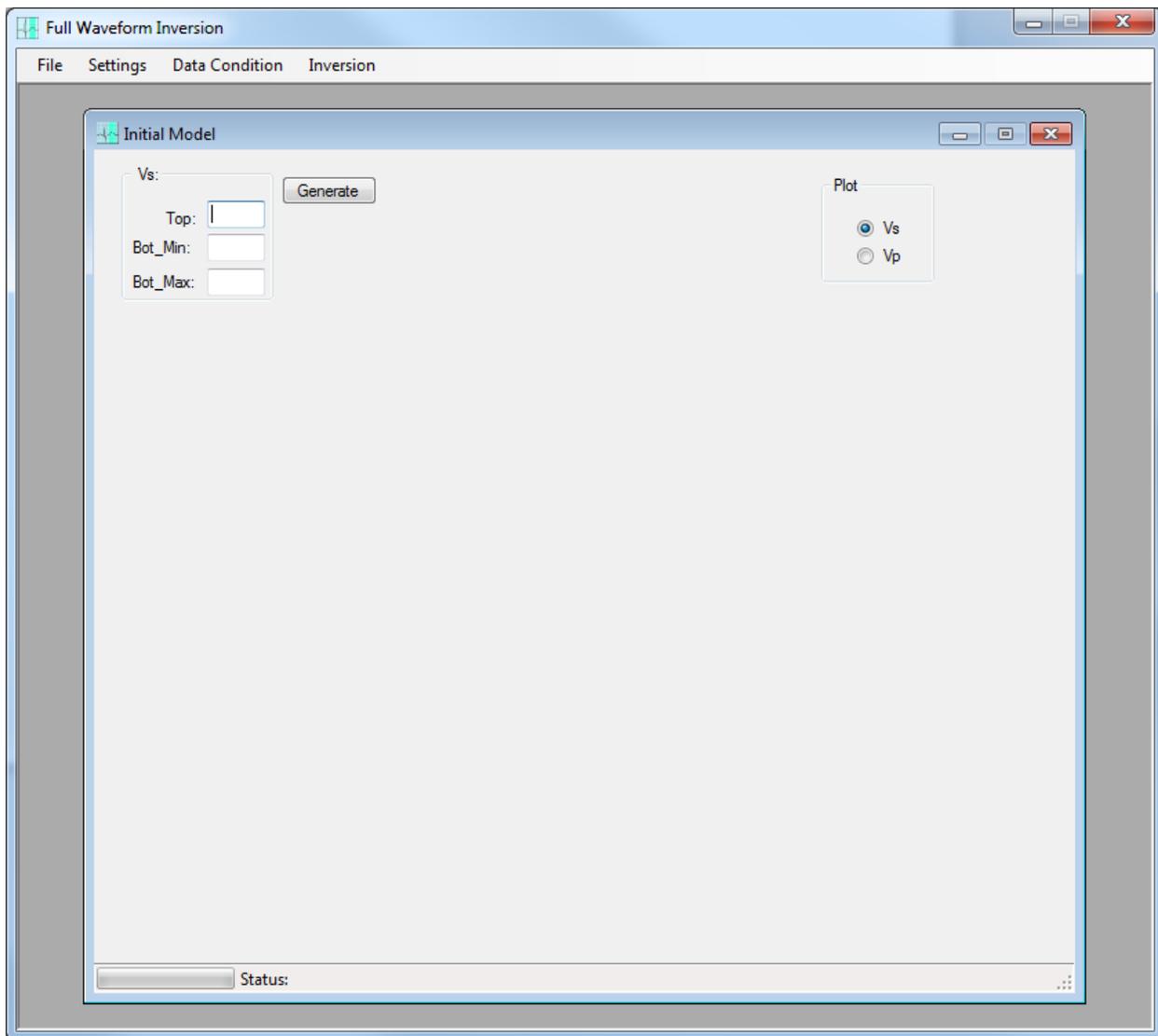
The wave velocities of the medium are represented using an array of values where each value pertains to a location within the volume. An initial array/model is required for the inversion process. This model can be generated based on the imported data or imported from a previously created wave velocity file.

### 4.1.1 Generate Wave Velocities

Navigate to the wave velocity generation page by selecting Inversion>Initial Model>Generate.



This will bring up the generation controls.

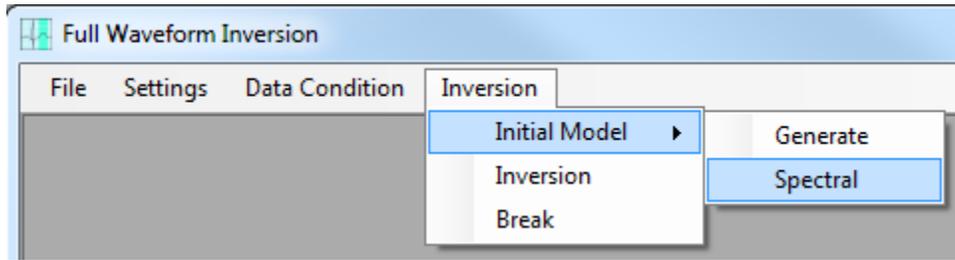


To generate the model input the shear wave properties of the surface and bottom of the volume being modeled. These properties can be determined by viewing the spectral image (See section 4.1.2). Once the values have been entered, select the “Generate” button. The status bar will say “Generating...” while the code is running. Once the models have been created a plot will appear based on the plot type selected in the box shown below.

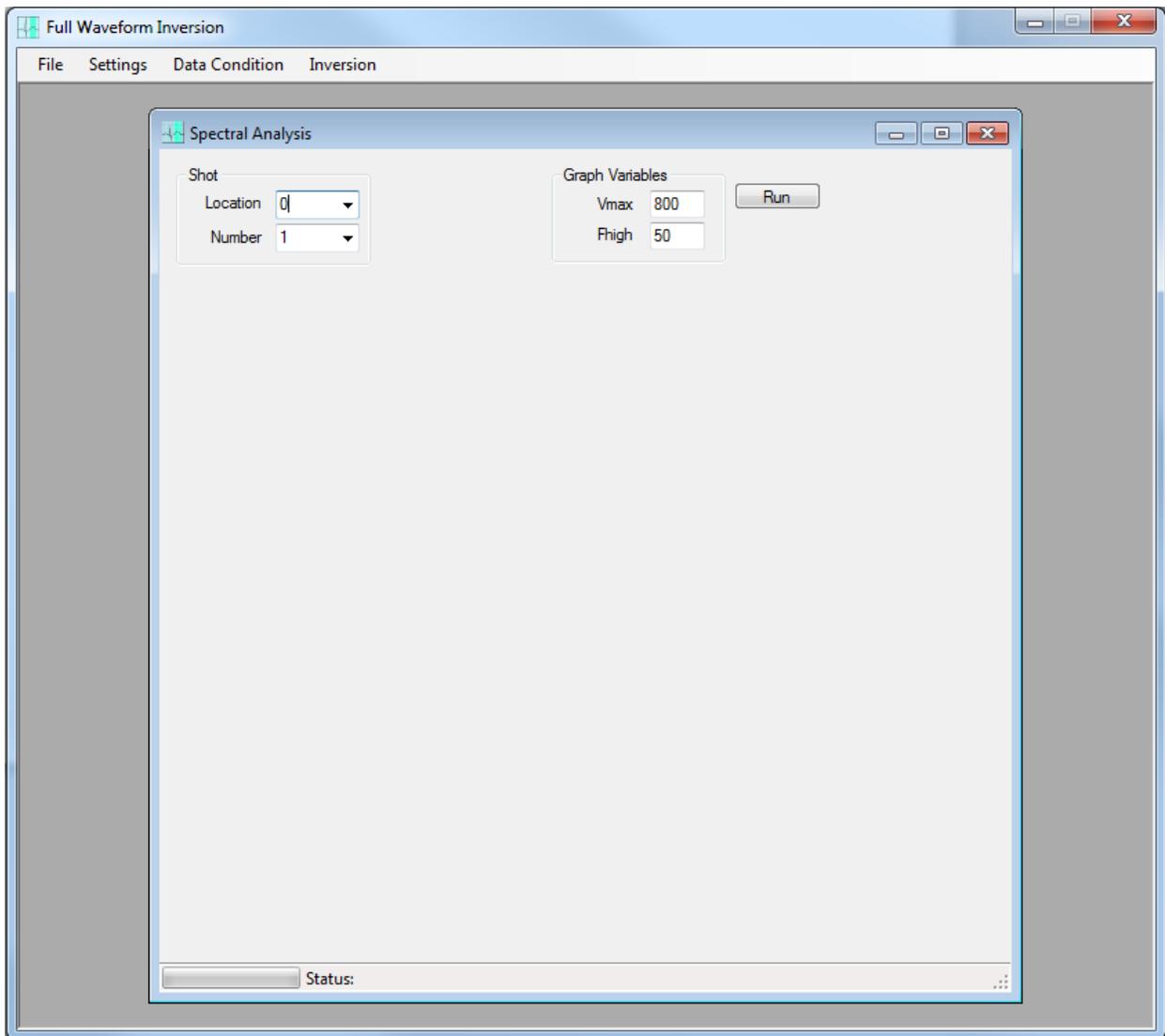


#### 4.1.2 Spectral Image

The spectral image is used to determine the initial properties used in the initial model generation. Select Inversion>Initial Model>Spectral to open the spectral image window.



This will bring up the window with the location and number of the shots imported.



Select the desired shot to be analyzed by using the drop down menu seen below.

Shot

Location 0

Number 1

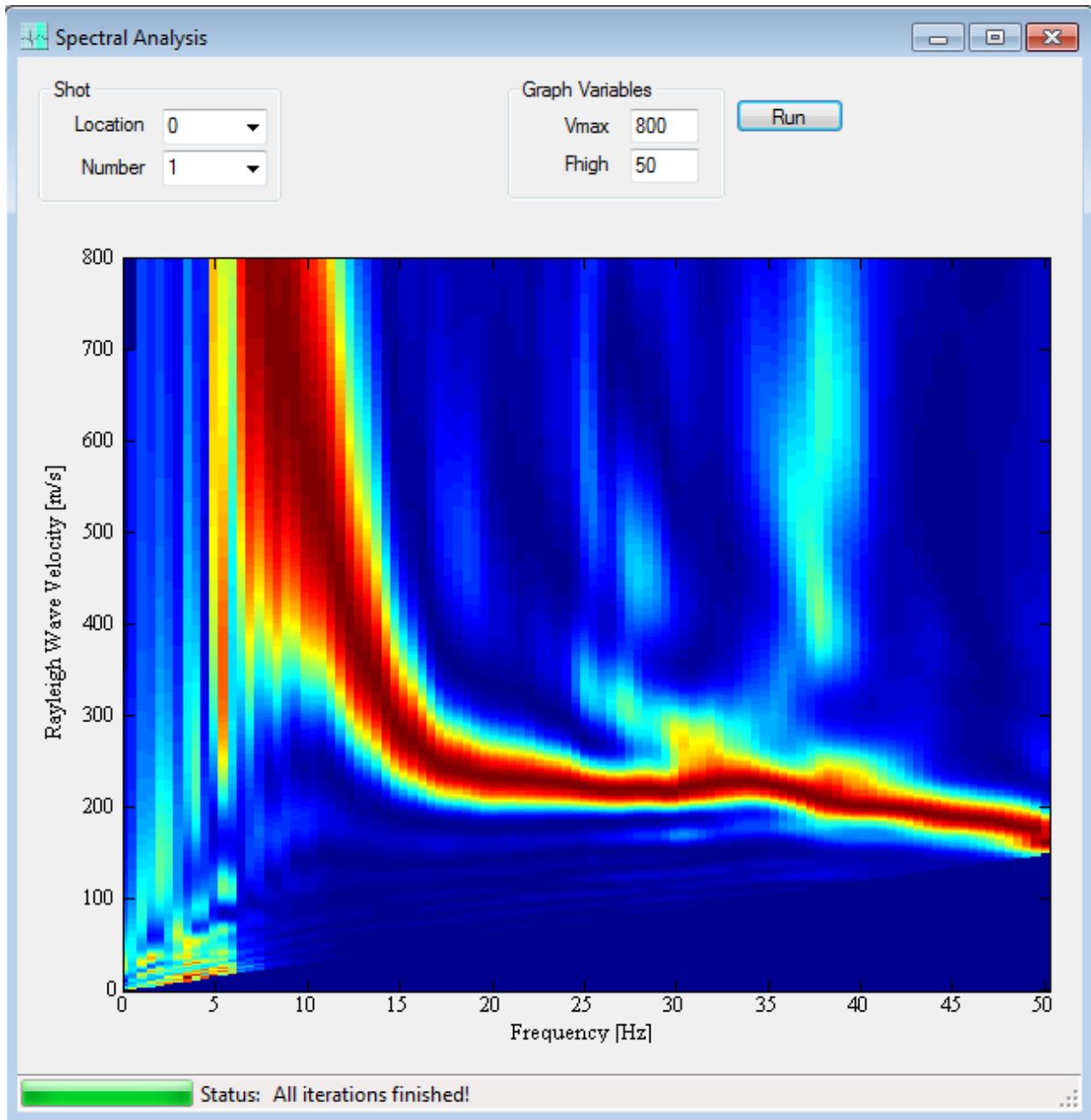
Select the plot properties such as maximum wave velocity and frequency by inputting values into the box shown.

Graph Variables

Vmax 800

Fhigh 50

Select the “Run” button to run the spectral analysis and produce an image. An example image can be found below.



In this case, the values used for model generation shown in section 4.1.1 are as follows.

Vs:

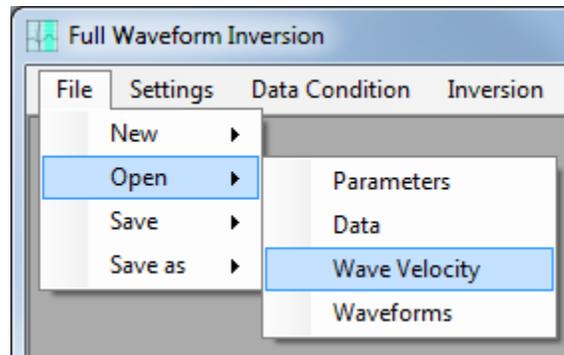
Top:	230
Bot_Min:	230
Bot_Max:	800

This is because the higher values, indicated by red, level out at around 230 m/s. This value is used for the surface shear wave velocity. The beginning of the good data in the image is around

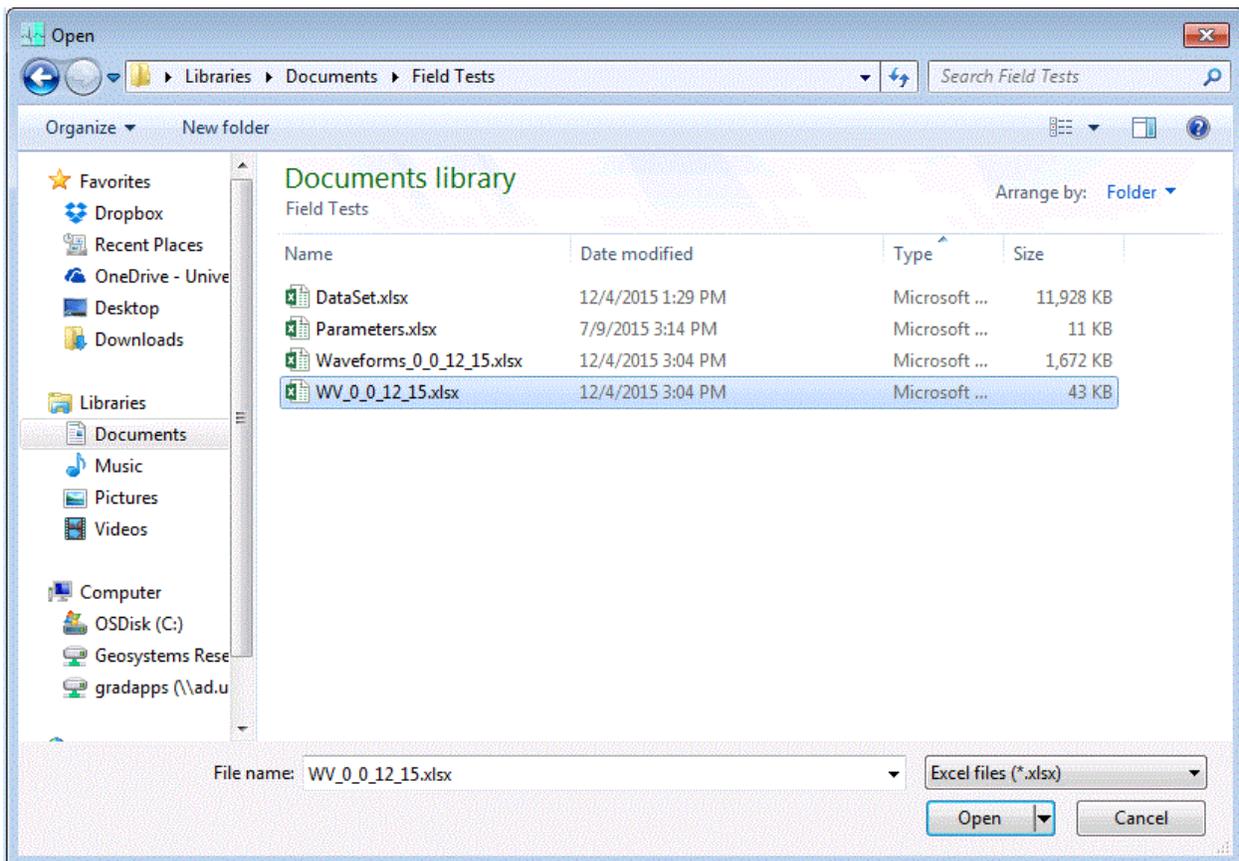
10 Hz at a maximum of around 800 m/s which will be used for the maximum value at the bottom of the model. The minimum at the bottom of the model is based on the lower shear wave velocities of the material being modeled.

#### 4.2 Open: Wave Velocities

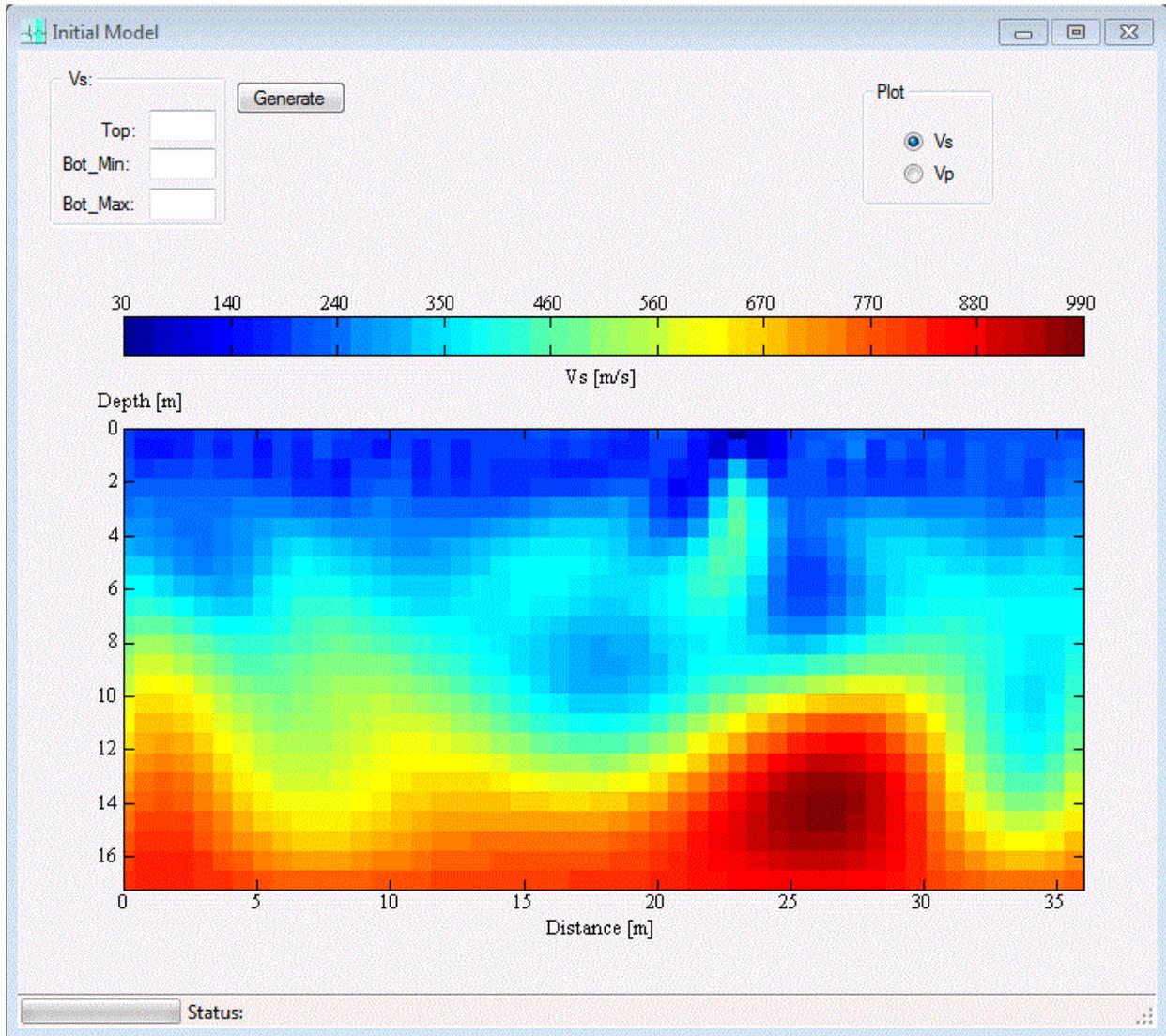
The second option for creating the initial model is to import it from a previously saved file. This can be done by selecting File>Open>Wave Velocity.



This will open a file selection window.

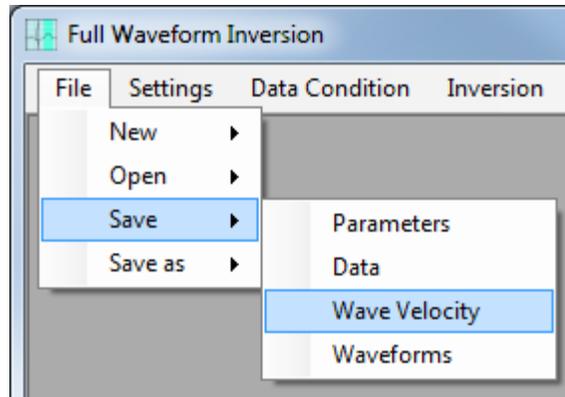


Navigate to the desired file and select “Open”. This will open the generate window (See section 4.1.1) and show the wave velocities.

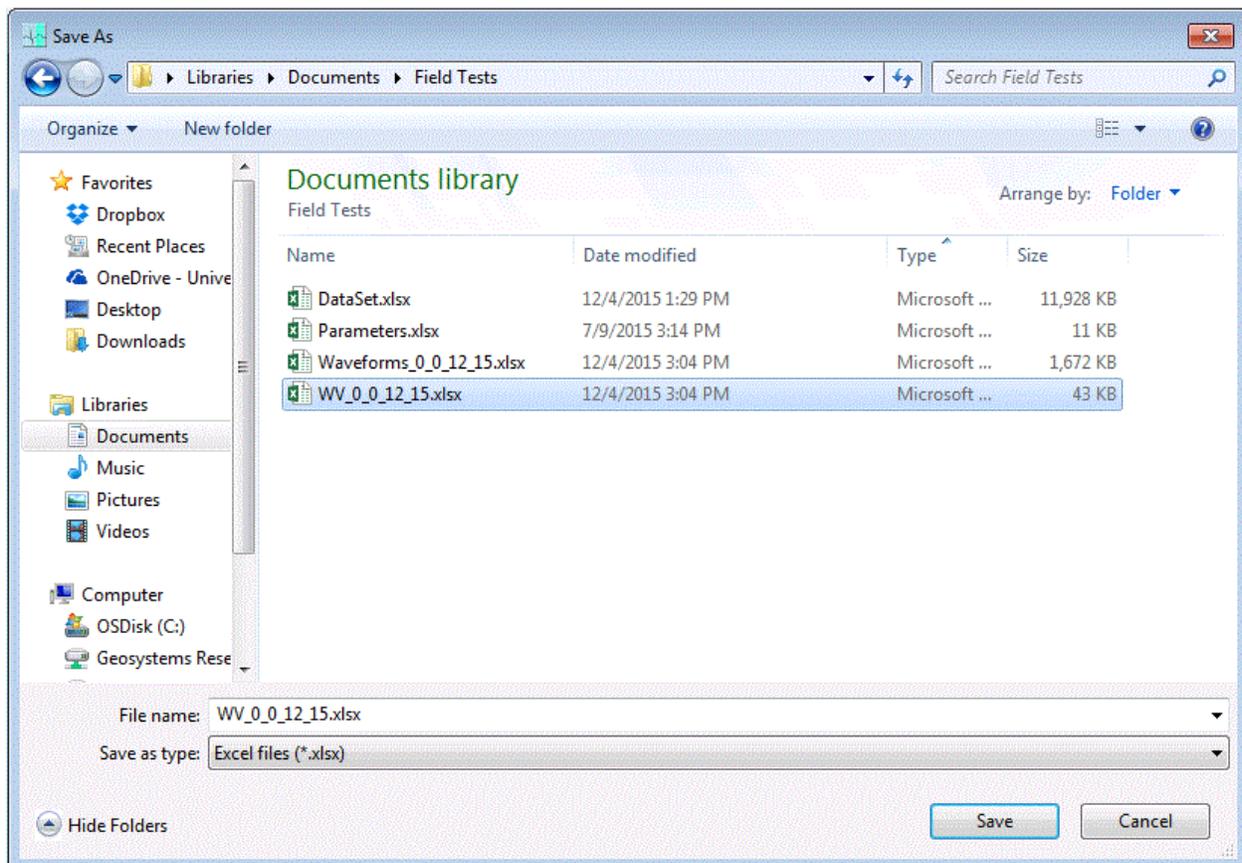


### 4.3 Save: Wave Velocity

To create a wave velocity file that can later be imported select File>Save>Wave Velocity.



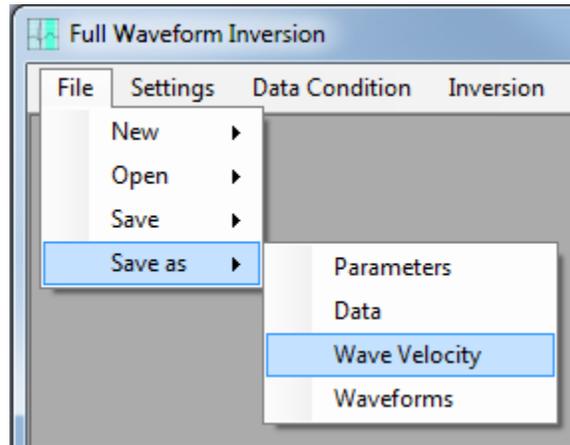
If a wave velocity file save location has already been selected or a wave velocity file has been imported, the save function will save the current wave velocity over that name and location. If none have been selected then the following dialogue box will come up.



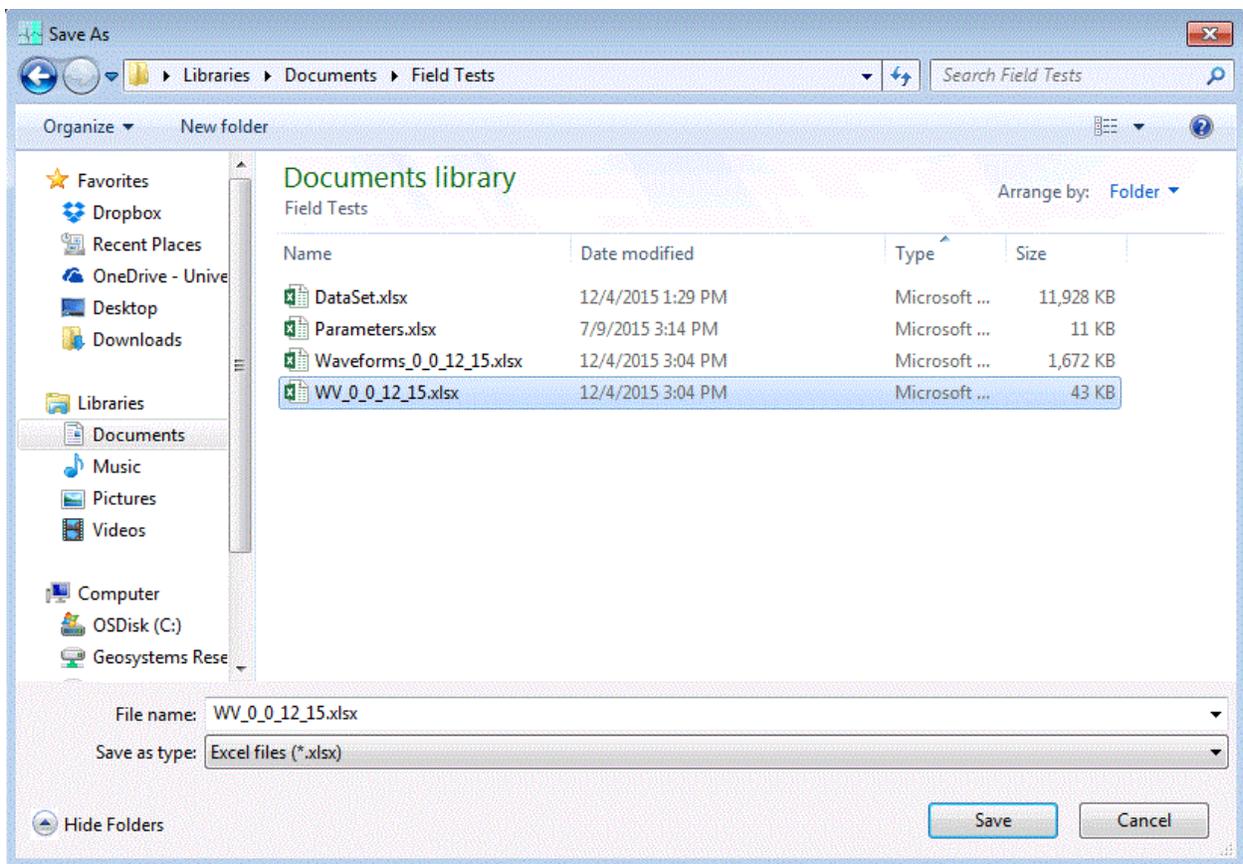
Navigate to the desired location and input the file name. Select “Save” to save to the name and location.

#### 4.4 Save as: Wave Velocity

The last option dealing with wave velocity files is to save the wave velocities to a new file location using File>Save As>Wave Velocity.



The following dialogue box is produced to allow for selection of a location and file name for the wave velocities.



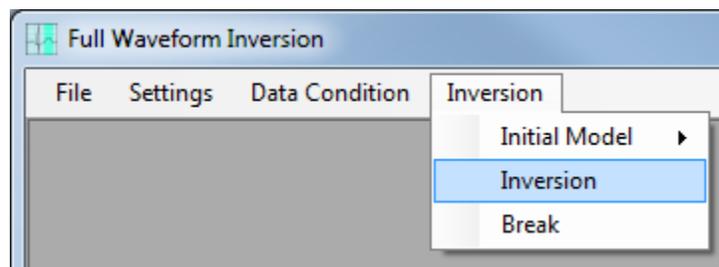
Select the desired location and name and select “Save”.

## 5.0 Inversion

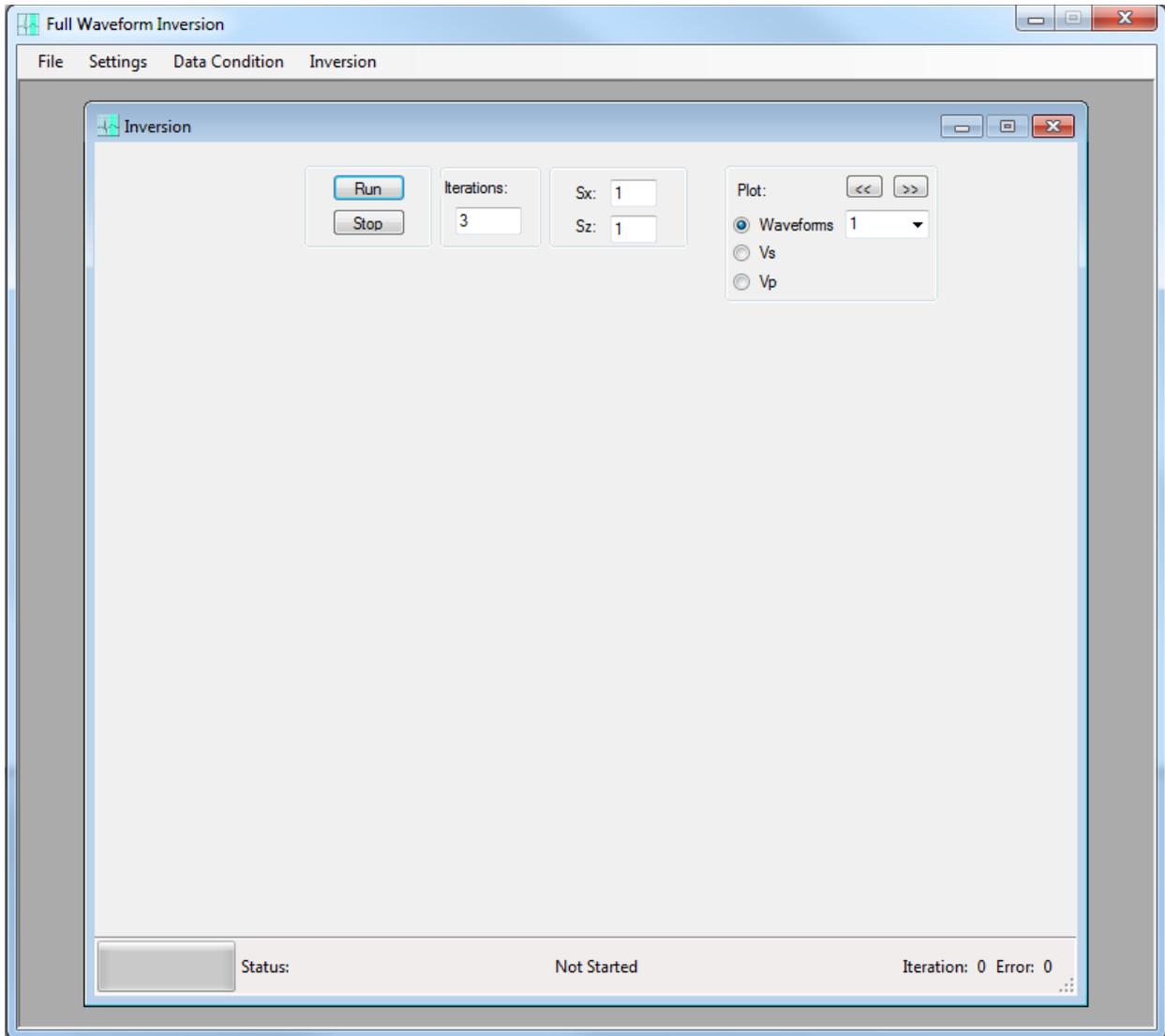
The inversion process combines all previous aspects of the user interface including the parameters, imported data, and wave velocities. It improves the wave velocity models by comparing synthetic waves perturbed in the models to data gathered in the field. Then it adjusts the model based on the difference in the synthetic data and actual data. This process is repeated for number of iterations desired. Updated wave velocity models and waveform comparison data is created once the inversion is complete.

### 5.1 Running Inversion

To begin the inversion process select Inversion>Inversion from the menu bar.



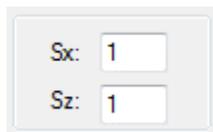
This opens the inversion window shown below.



Adjust the number of iterations to the desired amount using the edit box.

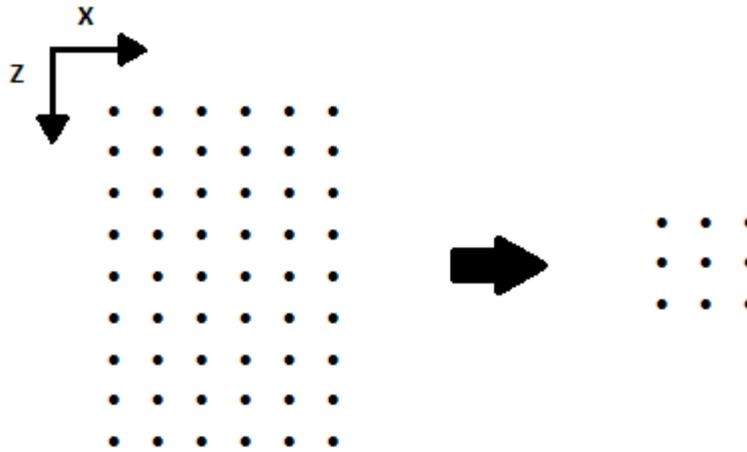


Input the Sx and Sz variables by changing the values in the box.



These values represent the amount that the space is compressed during the inversion process. Inputting one for both means it will remain the same. A two in the x direction means that every

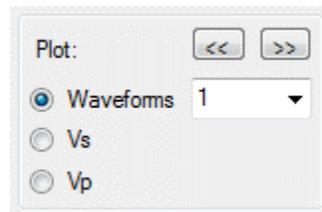
two grid points in the x direction will be averaged to create one. Likewise, a two in the z direction will compress the two points in the z direction to one. A visual example can be seen below:



In this case,  $S_x = 2$  and  $S_z = 3$ . This is because the x direction goes from six grid points to three grid points while the z direction goes from 9 grid points to 3 grid points. This reduction in the size of the model will reduce the resolution but allow for shorter computational times. It should be used only in the field for a quick solution if needed. To run the inversion process select “Run”. While the inversion is running, the program will update the box shown below with the previous iteration and error associated with that iteration. It will also update the status bar with the iteration number of the iteration currently running and the waveform plots.

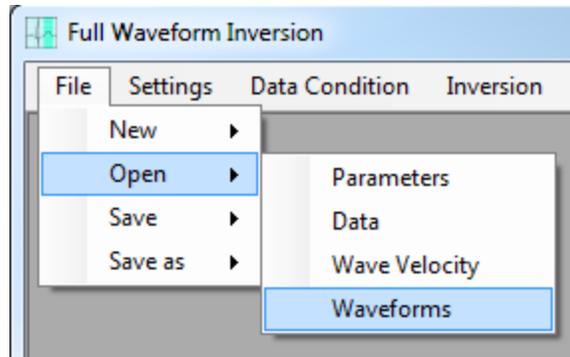
Iteration: 3 Error: 0.9293

Select “Stop” to stop the process after the current iteration is complete. This will keep the wave velocities and waveforms of the current iteration. Once the inversion is complete, the option to view the waveforms of each imported shot and the wave velocities is available by using the controls in the box seen below.

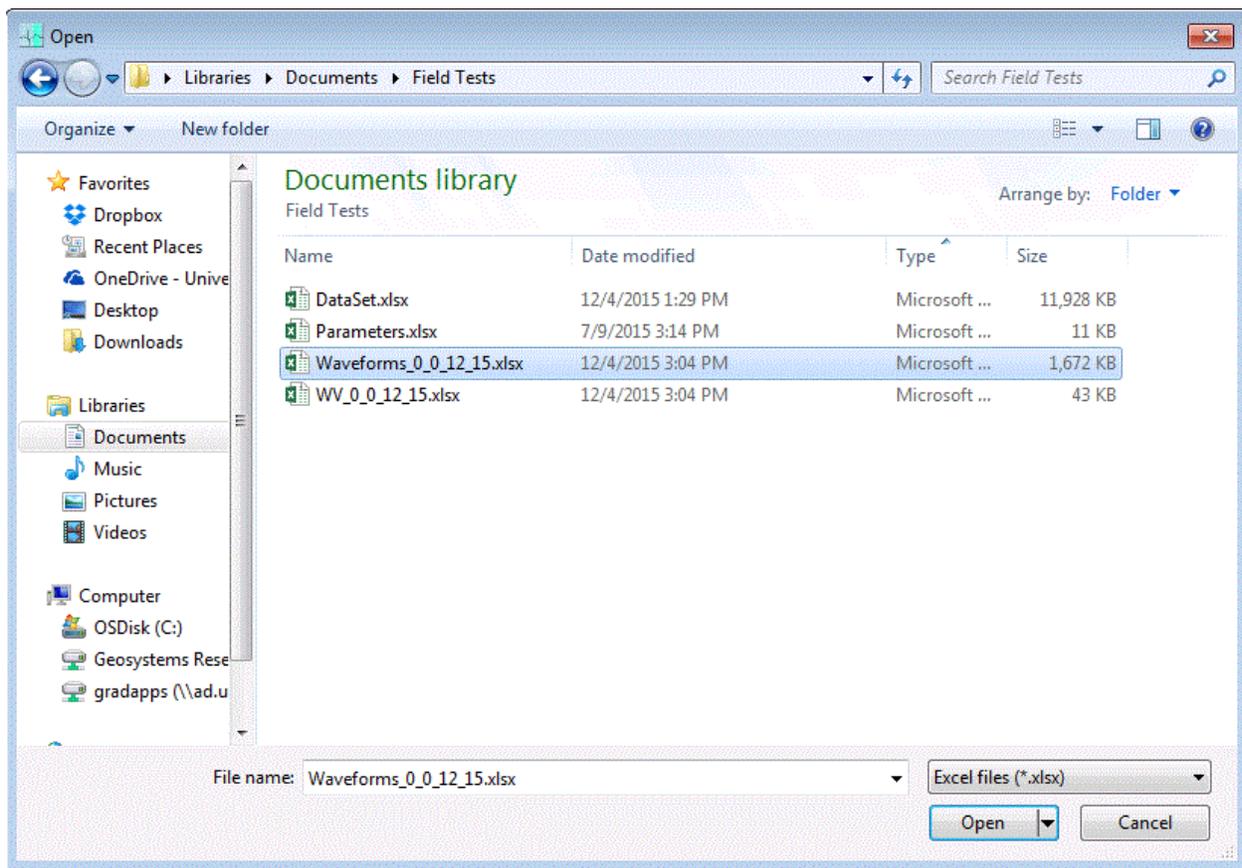


## 5.2 Open: Waveforms

The waveforms seen in the inversion window (See section 5.1) can be imported from a previously saved waveform file. This can be done by selecting File>Open>Waveforms.



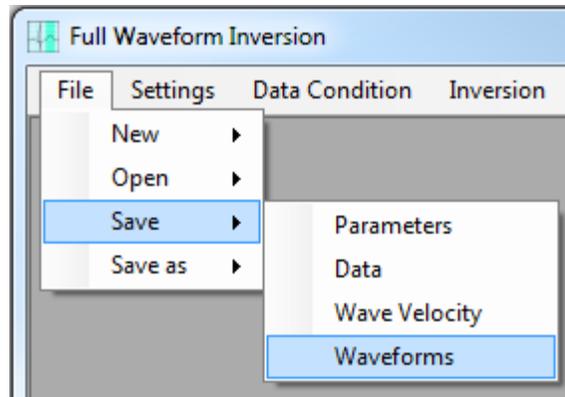
This will open up a file selection window.



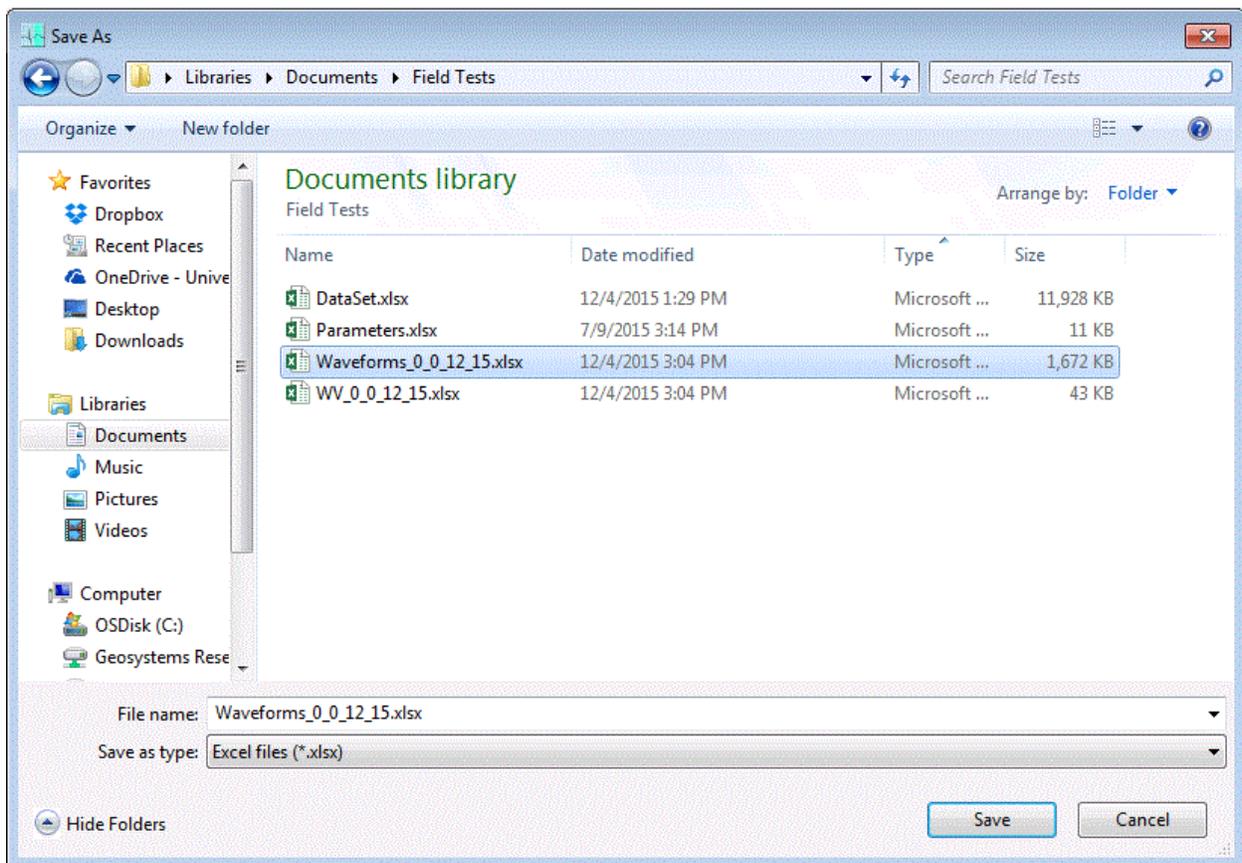
Choose the waveform file to be imported and select “Open” to import it.

### 5.3 Save: Waveforms

To create a waveform file that can later be imported select File>Save>Waveforms.



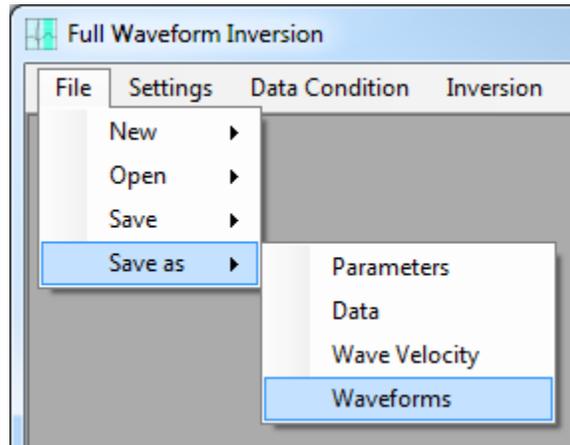
If a waveform file save location has already been selected or a waveform file has been imported, the save function will save the currently imported waveforms over that name and location. If none have been selected then the following dialogue box will come up.



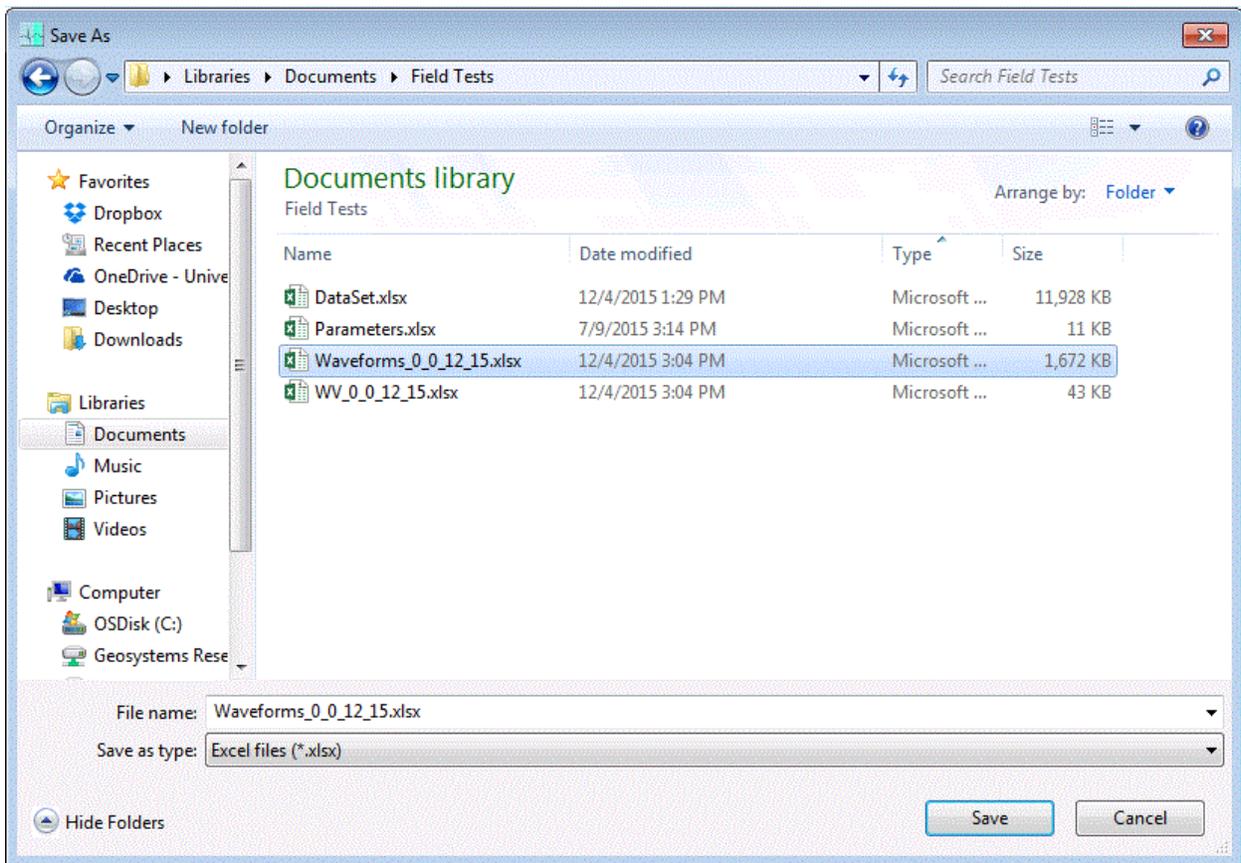
Navigate to the desired location and input the file name. Select “Save” to save to the name and location.

#### 5.4 Save as: Waveforms

The last option dealing with waveforms is to save the waveforms to a new file location using File>Save As>Waveforms.



The following dialogue box is produced to allow for selection of a location and file name for the waveforms.



Select the desired location and name and select “Save”.