

Important Notice

This copy may be used only for the purposes of research and private study, and any use of the copy for a purpose other than research or private study may require the authorization of the copyright owner of the work in question. Responsibility regarding questions of copyright that may arise in the use of this copy is assumed by the recipient.

UNIVERSITY OF CALGARY

Refining the amplitude of inverse scattering series internal multiple predictions

by

Andrew Scott Iverson

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN GEOLOGY AND GEOPHYSICS

CALGARY, ALBERTA

MAY, 2020

© Andrew Scott Iverson 2020

Abstract

Internal multiples are a form of noise in seismic data that degrades seismic image processing, leading to incorrect interpretations of the data. One method with minimal assumptions to attenuate internal multiples uses the inverse scattering series. Two key improvements with the method are developed. First, a new tool is established referred to as the downward generator space. This space used in combination with higher order terms and 2D adaptive subtraction improves the accuracy of the predicted internal multiples. Second, a non-stationary search limiting parameter is utilized in the time-offset domain allowing for increased flexibility in algorithm implementation. These new tools are applied to both synthetic and real data examples displaying the improvement in the prediction and attenuation of internal multiples.

Acknowledgements

I begin by thanking my family for their support, and specifically my partner and wife Steph Iverson for her support while I pursued my Masters. I am certain through all her enquiring and listening that we both learned more about this topic than we ever expected too. Graduate studies is something I had always wanted to undertake, and it would not have been possible without her.

I would like to thank my supervisors Dr. Kris Innanen and Dr. Daniel Trad. The topic that I studied was one I had significant interest in, but without their direction I would have struggled to pursue. Kris guided me through the method of inverse scattering giving me the knowledge and understanding to take on this topic. With Daniel's ability to communicate and meet with me I could understand the details of the method and how to program the algorithm both successfully and efficiently. After every meeting I had with either Kris or Daniel I always found myself reinvigorated to try and test new ideas and concepts due to the enthusiasm they always showed.

I would also like to thank all CREWES staff and students for the ideas, advice and support that I received specifically Scott Keating and Dennis Ellison. Lastly, I would like to thank NSERC CGS M, SEG Foundation Scholarship, Queen Elizabeth II Graduate (Master's) Scholarships, Alberta Graduate Student Scholarships, SEG Travel Grant and CREWES for the financial support I received during my studies, and Devon Energy for the donated dataset.

Dedication

*To my Dad,
Steve Randall Iverson*

Table of Contents

Abstract	ii
Acknowledgements	iii
Dedication	iv
Table of Contents	v
List of Figures and Illustrations	vii
List of Symbols, Abbreviations and Nomenclature	xii
CHAPTER ONE: BACKGROUND INFORMATION AND THEORY	1
1.1 The seismic method and multiples.....	1
1.2 Attenuation methods for seismic multiples	3
1.3 The Inverse Scattering Series and derivation of algorithms	4
1.3.1 The forward and inverse scattering series	5
1.3.2 Reduction to 1.5D offset-time and 1D time domain	11
1.3.3 Reduction to 1.5D tau-p domain	12
1.3.4 Reduction to 1D Pseudo-depth domain	13
1.4 Analytic example in 1D	14
1.5 Short vs. long path multiples	18
1.6 Adaptive subtraction	20
1.6.1 Autoconvolution of a wavelet	21
1.6.2 Theory.....	22
1.7 Thesis Overview	23
CHAPTER TWO: THE DOWNWARD GENERATOR SPACE	25
2.1 Numerical example with 1.D synthetic data	25
2.1.1 Modeling parameters	25
2.1.2 Internal multiple prediction	26
2.2 Downward Generator Space	27
2.3 Adaptive subtraction in 2D	29
2.3.1 Theory.....	30
2.3.2 Application to 1D numerical example.....	30
2.4 Pseudo-depth variant scalar	32
2.4.1 Scalar calculation.....	34
2.5 Numerical example with synthetic data in 1.5D	36
2.5.1 Modeling Parameters.....	37
2.5.2 internal multiple prediction	37
2.5.3 Downward Generator Space with a spatial dimension.....	39
2.6 Conclusions.....	41
CHAPTER THREE: HIGHER ORDER TERMS	42
3.1 Residual amplitude mismatch.....	42
3.2 Second order internal multiples	42
3.3 Inverse scattering series terms	45
3.3.1 Implementation of higher order terms	46
3.4 Numerical example with 1D synthetic data	47
3.5 Numerical example in 1D with a destructively interfered primary	51

3.6 Conclusions.....	55
CHAPTER FOUR: OFFSET-TIME DOMAIN ALGORITHM	56
4.1 Non-stationary seismic data.....	56
4.2 Non-stationary epsilon.....	56
4.3 Implementing the time domain algorithm	57
4.3.1 Masking operators and pseudocode.....	58
4.3.2 Differences between time and frequency domain	62
4.4 Numerical example with 1.5D synthetic data.....	63
4.4.1 Internal multiple prediction	64
4.5 Numerical example with 1.5D synthetic data and irregular spatial sampling	67
4.5.1 Random sampling numerical test	68
4.5.2 Orthogonal 3D survey geometry numerical test.....	71
4.6 Conclusions.....	76
CHAPTER FIVE: DEVON SYNTHETIC DATA.....	77
5.1 Donated dataset.....	77
5.2 Vertical Seismic Profile	78
5.3 Synthetic data modeling.....	80
5.3.1 Modeling algorithm	80
5.3.2 Synthetic VSP.....	81
5.4 Internal multiple prediction	87
5.4.1 Error quantification	88
5.5 Conclusions.....	96
CHAPTER SIX: DEVON REAL DATA	98
6.1 Application to recorded data.....	98
6.2 Recorded VSP.....	98
6.2.1 Processing.....	99
6.2.2 Corridor stacks.....	101
6.2.3 Internal multiple prediction	103
6.3 3D PSTM stacked seismic data	107
6.3.1 Well tie	108
6.3.2 Internal multiple prediction	111
6.3.3 Crossline internal multiple prediction	115
6.4 3D PSTM prestack seismic data	118
6.4.1 Time offset internal multiple prediction.....	120
6.5 Conclusions.....	122
CHAPTER SEVEN: CONCLUSIONS	123
7.1 Conclusions.....	123
7.2 Recommendations for future work	124
REFERENCES	126

List of Figures and Illustrations

Figure 1.1 a) Primary events for a three-layer model plus half-space, b) First order internal multiples for the three-layer model plus half-space, c) First order surface multiples, d) Short and long path multiples	2
Figure 1.2 Modified from (Weglein et al., 1997) displaying various scattering options from $G0V1G0V1G0V1G0$ where only far right example corresponds to an internal multiple	9
Figure 1.3 Schematic displaying how a multiple can be approximated with a combination of primaries through a convolution (*) and correlation (x).....	10
Figure 1.4 Tuning wedge for 3 layer model from (Widess, 1973)	18
Figure 1.5 Tuning wedge model for internal multiples Internal multiple prediction from ISS for tuning wedge model with variable epsilon.....	19
Figure 1.6 a) 40 Hz Ricker wavelet and autoconvolution of the wavelet, b) amplitude spectrum of the 40 Hz Ricker wavelet and autoconvolution of the wavelet.....	21
Figure 2.1 Velocity and depth model used for the 1D prediction.....	25
Figure 2.2 Reflectivity series for primaries, internal multiples and resulting seismic trace.....	26
Figure 2.3 Input seismic trace and 1D internal multiple prediction	27
Figure 2.4 Downward generator space displaying individual internal multiples. Prediction time is on the horizontal axis and downward generator time on the vertical axis	28
Figure 2.5 Internal multiple prediction with 1D adaptive subtraction.....	31
Figure 2.6 Internal multiple prediction with 2D adaptive subtraction.....	32
Figure 2.7 a) Summing over wavenumber then depth, b) proposed order of operations alteration summing over depth then wavenumber with scalar applied to give scaled prediction	33
Figure 2.8 Transmission loss scalar for internal multiple prediction.....	35
Figure 2.9 a) Trace with both scaled and unscaled predictions, b) isolating two multiples M323 and M313.....	36
Figure 2.10 Geologic model displaying velocities and depths used.....	37
Figure 2.11 (Left) Synthetic seismic shot record (Right) 1.5D tau-p multiple prediction.....	38
Figure 2.12 a) three dimensional view of the 3D DGS, b) a slice through zero offset similar to standard 2D DGS c). One DGS slice displaying the first order and higher order internal multiples, d) A deeper DGS slice displaying only higher order internal multiples.	40

Figure 3.1 Displaying the algorithm computing second order multiples with convolution (*) and correlation (x).....	43
Figure 3.2 Reflectivity series and trace for primaries, first and second order multiples	44
Figure 3.3 2D downward generator space displaying individual internal multiples, a) b3 term with first order multiples displayed with red box. b) b5PPI higher order multiples. c) b5PIP artifacts from the prediction	48
Figure 3.4 2D downward generator space displaying individual internal multiples after including higher order terms	49
Figure 3.5 Internal multiple prediction with higher order terms and 2D adaptive subtraction	50
Figure 3.6 Velocity and depth model with additional layer used for 1D prediction.	51
Figure 3.7 a) Primaries and multiples trace b) Primaries only trace highlighting location of missing primary due to internal multiples	52
Figure 3.8 a) Internal multiple prediction with 1D adaptive subtraction b) Internal multiple prediction with 2D adaptive subtraction and higher order terms.....	53
Figure 3.9 Trace with internal multiple attenuation using a) the 1D adaptive subtraction Trace with internal multiple attenuation and b) 2D adaptive subtraction and higher order terms..	54
Figure 4.1, adapted from (Innanen, 2015), visual display of the calculation of convolutions and correlations for a given time $t(j)$ through matrix multiplication, with the mask matrix applied to the convolution matrix M_C	58
Figure 4.2, adapted from (Innanen, 2015), displays the mask matrix applied to the convolution matrix where the shaded region is set to zero and the bounds are determined such that the lower-higher-lower criteria is met for the given $t(j)$	59
Figure 4.3 Display of the prediction algorithm for the (x, t) case for a given time $t(j)$ calculated for all offset with the convolution and correlation matrices, where the mask matrix is applied to each convolution matrix in the block matrix M_C . this is applied to the (x, t) trace (s) which has been stacked into a single column vector. Adapted from (Innanen, 2015).	60
Figure 4.4 Pseudo code displaying the implementation of the (x, t) case with the use of 2D convolution functions for a stationary epsilon. The mask is applied in a similar manner where all values are set to zero given that they are either below the calculation time (it) and epsilon number of samples above (it).	61
Figure 4.5 Pseudo code displaying the implementation of the (x, t) case with the use of 2D convolution functions for a nonstationary epsilon in both time and space dimensions.....	62

Figure 4.6 Visual display for the mask matrix applied to the data (s) prior to the 2D convolution function, where the bounds are determined such that the lower-higher-lower criteria is met for the given $t(j)$ for a) epsilon varying in time and constant with offset b) An example of epsilon varying in time and with offset with two epsilon values with a harsh cutoff	62
Figure 4.7 a) offset-time multiple prediction with epsilon=30, b) offset-time multiple prediction with epsilon=30	64
Figure 4.8 a) offset-time multiple prediction with spatially varying epsilon, b) epsilon schedule used for prediction with harsh cutoff	65
Figure 4.9 a) offset-time multiple prediction with spatially varying epsilon, b) epsilon schedule used for prediction with linear taper	66
Figure 4.10 (Left) offset-time multiple prediction with spatially varying epsilon (Right) epsilon schedule used for prediction varying in both offset and time	67
Figure 4.11 Velocity model with interface depths indicated by red lines	68
Figure 4.12 Synthetic data with total trace count of a) 256, c) 128 and e) 64. Internal multiple prediction for the given input data with trace count b) 256, d) 128 and f) 64	69
Figure 4.13 a) Positive offset synthetic, b) Internal multiple prediction	70
Figure 4.14 a) Velocity model, b) synthetic, c) time offset prediction	71
Figure 4.15 a) Source (red) and receiver (blue) geometry with rays for a single cdp, b) Offset and azimuth distribution for the displayed cdp	72
Figure 4.16 a) Input seismic trace from 3D seismic geometry, b) Internal multiple prediction...	73
Figure 4.17 a) Input seismic trace from 3D seismic geometry with blank traces, b) Internal multiple prediction	74
Figure 4.18 a) Input seismic trace from 3D seismic geometry with 2D Spline interpolation, b) Internal multiple prediction, c) Input seismic trace from 3D seismic geometry, d) Internal multiple prediction with the original traces displayed.....	75
Figure 5.1 Well logs including gamma ray, spontaneous potential, resistivity, neutron porosity, sonic and density.....	77
Figure 5.2 VSP with upgoing and down going energy and internal multiples	79
Figure 5.3 a) Recorded VSP with up and downgoing waves, b) Synthetic VSP with up and downgoing waves.....	81
Figure 5.4 a) calculated one way travel time and picked one way travel time, b) Smoothed difference between the two one way travel times	82

Figure 5.5 a) Synthetic VSP with first break pick displayed in blue with upgoing events, b) flattened synthetic VSP with first break pick displayed in blue with upgoing events.....	83
Figure 5.6 Full stack (primaries and multiples) and zero depth trace (primaries and multiples) .	85
Figure 5.7 Outside corridor stack (primaries) and zero depth trace (primaries and multiples)....	86
Figure 5.8 The 2D Downward generator space for the zero depth trace prediction.....	87
Figure 5.9 Internal multiples trace in red and internal multiple prediction in blue.	88
Figure 5.10 a) Multiples trace in red and internal multiple prediction in blue taken as absolute value of the multiples trace and difference with the multiples trace, b) a 50 point moving average window is used to smooth the traces c) a 50 point moving average window is applied to the input data trace for reference to the internal multiple trace.....	90
Figure 5.11 Displaying the 2D downward generator space for the synthetic VSP after stacking	91
Figure 5.12 Multiples trace and internal multiple prediction from 2D adaptive subtraction.	91
Figure 5.13 Absolute value of multiple trace, and difference with the two predictions with rolling average window of 50.	92
Figure 5.14 outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black.	94
Figure 5.15 Outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black. a) b3 subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction	95
Figure 5.16 Using 37 Hz Ricker wavelet displaying outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black for a) b3 subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction	96
Figure 6.1 Figure from (Cova, et al., 2018) a) source locations in Easting and Northing, b) Receiver location in depth and in line with the sources.....	99
Figure 6.2 Recorded VSP after processing with upgoing wave	100
Figure 6.3 a) Frequency content differences between full stack (Primaries and multiples) to outside corridor stack (primaries only) and b) using a decibel scale	101
Figure 6.4 Full stack (Primaries and multiples) and outside corridor stack (primaries only)	102
Figure 6.5 Internal multiple prediction and internal multiples trace	103
Figure 6.6 Downward generator space after stacking for 2D adaptive subtraction.....	104

Figure 6.7 Internal multiple prediction with 2D adaptive subtraction and internal multiples trace.....	105
Figure 6.8 Outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black for a) b3 subtraction, b) 1D adaptive subtraction and c) 2D adaptive subtraction	106
Figure 6.9 EBCDIC header displaying acquisition parameters and processing	107
Figure 6.10 Frequency spectrum of PSTM Stack.....	108
Figure 6.11 Well tie to 3D PSTM stacked data with a) synthetic in blue and stacked data in red, b) well tie over the time window of 1550–1900ms and c) cross correlation window with a max coefficient of 0.665	109
Figure 6.12 a) Synthetic outside corridor stack (Primaries) and trace from 3D PSTM Stack, b) Synthetic zero depth trace (Primaries and multiples) and trace from 3D PSTM Stack .	110
Figure 6.13 Internal multiple prediction and trace from 3D PSTM Stack.....	112
Figure 6.14 Downward generator space after stacking for 2D adaptive subtraction.....	113
Figure 6.15 Outside corridor stack (primaries) in blue, PSTM trace (primaries and multiples) in red and PSTM trace after internal multiple attenuation in black for a) b3 subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction.....	114
Figure 6.16 Crossline displaying 3D PSTM stack data a) input data, b) internal multiple prediction and c) internal multiple prediction with 2D adaptive subtraction	116
Figure 6.17 a) Crossline through PSTM stack. b) Crossline through PSTM stack after internal multiple attenuation, with red ovals highlighting significant areas of change due to internal multiple attenuation.	117
Figure 6.18 Well tie cross correlation a) before and b) after internal multiple attenuation.....	118
Figure 6.19 a) offset (ft) and azimuth acquisition distribution for the given CMP, b) seismic gather with the displayed acquisition before 5D interpolation	119
Figure 6.20 a) offset (ft) and azimuth acquisition distribution for the given CMP after 5D interpolation, b) seismic gather with the displayed acquisition after 5D interpolation	120
Figure 6.21 a) seismic gather after 5D interpolation, b) Time offset internal multiple prediction of seismic gather after 5D interpolation	121

List of Symbols, Abbreviations and Nomenclature

Symbol	Definition
<i>AVO</i>	Amplitude Variation with Offset
α	Compressional velocity
b_1	Recorded wavefield
b_n	Internal multiple prediction term n
cmp	Common mid point
c_0	Reference velocity
<i>CREWES</i>	Consortium for Research in Elastic Wave Exploration Seismology
<i>DGS</i>	Downward generator space
<i>EPSI</i>	Estimating Primaries by Sparse Inversion
δ	Delta function
ε	Search limiting parameter
ϵ	Depth below free surface
<i>FWI</i>	Full Waveform Inversion
<i>f-k</i>	Frequency-wavenumber
<i>f</i>	Filter
<i>G</i>	Greens Function
G_0	Reference greens function
G_s	Scattered field
<i>H</i>	Heaviside step function
<i>IM</i>	Internal Multiple
<i>ISS</i>	Inverse Scattering Series
<i>i</i>	Complex number
<i>k</i>	Fourier conjugate variable
<i>L</i>	Differential operator
L_0	Reference Differential operator
<i>L1</i>	L1 norm
<i>L2</i>	L2 norm
<i>MCG</i>	Multiple Contributing Gather
<i>P</i>	Primary
ρ	Density
ϕ	Minimization term
q_x	Vertical wavenumber
<i>R</i>	Reflection coefficient
<i>r</i>	Location
r_s	Source location
<i>T</i>	Transmission Coefficient
<i>tau-p</i>	Intercept time-ray parameter
τ	Time
<i>VSP</i>	Vertical Seismic profile
<i>V</i>	Perturbation operator
ω	Angular Frequency
z_n	Depths of layer n

Chapter One: Background Information and Theory

1.1 The seismic method and multiples

Seismic waves can be utilized to obtain subsurface information and properties. This begins with an active source such as dynamite or a seismic vibrator to generate the seismic waves. When the waves travel into the subsurface and cross an interface with varying elastic properties a portion of the energy is reflected, and the majority is transmitted. By recording and processing the reflections an image of the subsurface is created which is used to identify structure, faulting, fracturing and elastic properties. The reflections from the interface will ideally have amplitudes representative of the contrasting layer properties with no additional noise. The reflections can be used to estimate the layer properties through inversion methods such as Full Waveform Inversion (FWI), pre-stack joint or post stack inversion, amplitude variation with offset (AVO) analysis and other amplitude dependent techniques.

When recording land seismic data there are various types of noise sources that degrade the image. Random noise sources for land seismic data can include wind, vehicles, machinery, animals, and numerous other sources on the surface. Other land seismic specific issues often involve the near surface due to topography changes, relatively unconsolidated material, and the potential for heterogeneity. An ongoing issue in seismic imaging and interpretation is the removal of unwanted noise in the data. For this project the type of noise targeted for removal is due to multiple reflections in the subsurface specifically long path internal multiples.

When seismic waves reflect off multiple interfaces or boundaries below the surface these can obstruct the desired primary reflections (Iverson, 2014; Hilterman, et al., 2018; Zhang, et al., 2018). This degrades the seismic image which can lead to errors in interpretation of the data. The naming convention used to describe specific primaries and multiples is displayed (Figure 1.1). Multiples can be grouped into different categories, there are surface, internal, short and long path multiples. Surface multiples have at least one of the downward reflections occurring at the surface where internal multiples have all the multiple reflections occurring in the subsurface. Long path multiples arrive as a separate event whereas short path multiples occur so quickly after the primary they impact the wavelet shape (SEG wiki, 2019). First order multiples will have the three reflections where higher order multiples have an increased number of reflections in the subsurface (e.g. 2nd order have 5 reflections, 3rd order have 7 reflections ...).

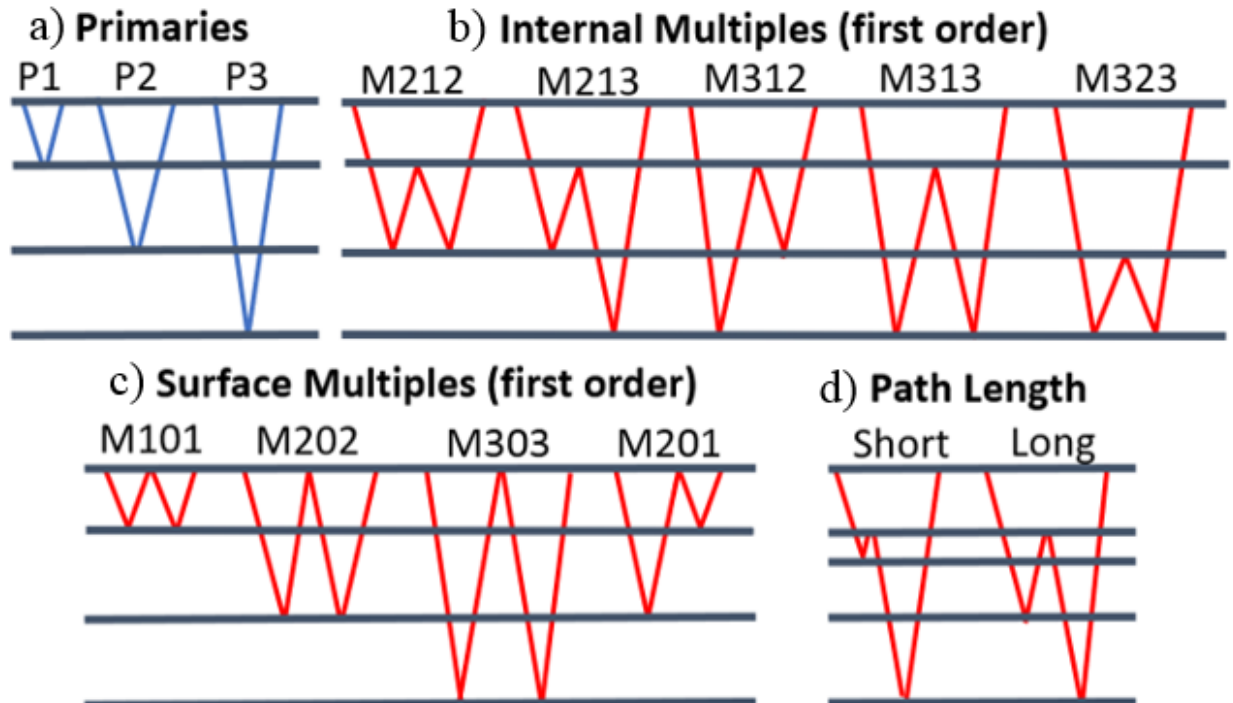


Figure 1.1 a) Primary events for a three-layer model plus half-space, b) First order internal multiples for the three-layer model plus half-space, c) First order surface multiples, d) Short and long path multiples

1.2 Attenuation methods for seismic multiples

Numerous methods have been utilized to attenuate multiples. These methods can be grouped into the three broad categories of deconvolution, filtering and wavefield prediction (Xiao, et al., 2003).

Deconvolution uses the periodic nature of multiples to develop an operator to attenuate them, with several assumptions such as the data is zero offset, the horizons are flat and there is no lateral variation (Xiao, et al., 2003). This is done by creating a deconvolution filter which includes the multiples to be predicted. Which multiples are predicted and attenuated is controlled through the choice of the prediction distance and filter length to attenuate short or long path multiples, equivalently referred to as short and long path reverberations (Peacock & Treitel, 1969). As the prediction distance decreases high frequency noise can be introduced so the parameter must be chosen purposefully (Peacock & Treitel, 1969). Various domains have also been implemented to attempt to overcome the zero offset requirement and apply to gathers (Perez & Henley, 2000). This is done by transforming the data into various domains where the data's periodic nature is more constant with offset such as the radial domain or after the application of normal move out to flatten the data (Perez & Henley, 2000).

Filtering methods rely on separating the primary and multiple events by transforming the data into various domains including f-k, tau-p or RADON (Kabir & Marfurt, 1999; Xiao, et al., 2003). Once transformed a mute is created in the chosen domain which ideally will only remove the multiples. Filtering can be successful given sufficient moveout differences between primary

and multiple reflections. When the moveout difference is small, the multiple energy is difficult to isolate from the primary. These methods will only attenuate long path multiples.

There are a number of wave equation based techniques (Verschuur, et al., 1992; Berkhout & Verschuur, 1997). This project implements the inverse scattering series (Weglein, et al., 1997) which falls under the wavefield prediction category and is fully data driven. These methods use some form of the wave equation to create a dataset which only includes multiples. An alternative method which uses similar mathematical techniques but requires an input subsurface model generally taken to be the migrated section uses this model to predict the resulting internal multiples (Pica & Delmas, 2008). Partial versions have been implemented of wavefield prediction methods such as XIMP (Wu & Dragoset, 2011) which require subsurface information inputs such as multiple generator horizons. A separate approach directly estimates primaries by sparse inversion (EPSI) using full waveform inversion (Groenestijn & Verschuur, 2009). This method attenuates multiples by inverting directly for primaries using the data which contains both primaries and multiples.

1.3 The Inverse Scattering Series and derivation of algorithms

The inverse scattering series method of multiple attenuation was developed by Weglein et al in the 1990's. The method predicts internal multiples using the recorded seismic data with no prior subsurface information requirements. By creating a prediction of the internal multiples this is used to directly remove them from the data. In theory what is desired is an algorithm which predicts the multiples exactly so the prediction can be subtracted directly from the data. The inverse scattering series correctly predicts the time the multiples occur, with small errors in the

predicted amplitudes. The amplitude errors have been recently improved upon with internal multiple elimination algorithms (Zou & Weglein, 2015; Zou, et al., 2016). In practice multiple attenuation is completed in two steps, first the internal multiples prediction is made followed by an adaptive subtraction step to remove them from the data. Any amplitude errors between the predicted and actual multiples in the data attempt to be rectified by the process of adaptive subtraction (Guitton & Verschuur, 2004; Keating, et al., 2015).

The issues addressed in this thesis involve the amplitudes of the predicted events and artifacts that can arise from the prediction. As the method is data driven and utilizes the data itself to make the prediction, any losses present in the data will be used in the prediction. How the amplitude errors arise in the inverse scattering series method is discussed. These amplitude errors can vary significantly at similar temporal locations making the designing of a filter through adaptive subtraction difficult. This leads to difficulty in using the prediction and trying to analyze if the algorithm is possibly damaging the primaries in the data. There can be difficulties with implementation of the method including parameterization of the search limiting parameter epsilon. Recently the algorithm has been applied in various domains with increased success (Sun & Innanen, 2014; Sun & Innanen, 2016).

1.3.1 The forward and inverse scattering series

The scattering series describes the relationship between the physical properties of an actual and reference medium and the impulse response of that reference and actual medium (Weglein, et al., 1997). First forward scattering will be shown and discussed. Scattering begins by describing the response in terms of a reference medium and a perturbation from this medium.

This will be done for a single frequency and assuming the source signature has been deconvolved (Weglein, et al., 1997) where,

$$LG = -\delta(r - r_s) \quad (1.1)$$

L is the differential operator for wave propagation in the actual medium G is the wavefield and r is the location relative to r_s . The equation for the reference medium is given by

$$L_0G_0 = -\delta(r - r_s) \quad (1.2)$$

Where, L_0 is the differential operator for wave propagation in a reference medium and G_0 function. For example, an acoustic wave equation with the source term being an impulsive point source (Weglein, et al., 1997) from Equation 1.2 gives,

$$\left(\frac{\omega^2}{\kappa_0} + \nabla \cdot \left(\frac{1}{\rho_0} \nabla \right) \right) G_0 = -\delta(r - r_s) \quad (1.3)$$

From Equation 1.1 and 1.2 these can be combined to give the perturbation operator V defined as,

$$V = L - L_0 \quad (1.4)$$

And the scattered field as,

$$G_s = G - G_0 \quad (1.5)$$

With the perturbation operator and scattered field using the Lippmann-Schwinger equation, which is the foundation for scattering theory gives,

$$G_s = G - G_0 = G_0VG \quad (1.6)$$

Which now relates the reference and actual wavefields through a perturbation operator. Then rearranging Equation 1.6 gives

$$G(1 - G_0V) = G_0 \quad (1.7)$$

And then solving for G

$$G = G_0 \frac{1}{1 - G_0 V} \quad (1.8)$$

Recognizing this is an infinite geometric series that can equivalently be expressed as the following

$$G = G_0(1 + G_0 V + G_0 V G_0 V + \dots) \quad (1.9)$$

Putting this back into Equation 1.6 gives the forward scattering series as a function of only V and G_0

$$G_s = G - G_0 = G_0 V G_0 + G_0 V G_0 V G_0 + \dots \quad (1.10)$$

Where G_s is equivalent to the data recorded on the surface now noted as D .

We now have an equation that can describe through perturbations and a reference wavefield the recorded seismic data. What we would like is an equation that will use the recorded data to solve for certain perturbation subsets. Inverse scattering series takes the recorded and reference wavefield to give the perturbation operator (Weglein, et al., 1997). The forward infinite series (Equation 1.10) for D can be equivalently written using the same geometric series reduction

$$D = G_0 V G_0 + G_0 V G_0 V G_0 + \dots = \frac{G_0 V G_0}{1 - V G_0} \quad (1.11)$$

Rearranging gives

$$D(1 - V G_0) = G_0 V G_0 \quad (1.12)$$

and isolating for $V G_0$ gives

$$D = G_0 V G_0 + D V G_0 \quad (1.13)$$

And solving for $V G_0$ gives

$$V G_0 = \frac{D}{D+G_0} \quad (1.14)$$

And finally rearranging and expanding in the form of a geometric series gives

$$V = \frac{1}{G_0} \left(\frac{D}{G_0} \frac{1}{1 - \left(\frac{D}{G_0}\right)} \right) = \frac{1}{G_0} \left(\frac{D}{G_0} \left(1 - \frac{D}{G_0} + \left(\frac{D}{G_0}\right)^2 - \dots \right) \right) \quad (1.15)$$

Giving the inverse series for the perturbation operator

$$V = \frac{1}{G_0} \left(\frac{D}{G_0} - \left(\frac{D}{G_0}\right)^2 + \left(\frac{D}{G_0}\right)^3 - \dots \right) = V_1 + V_2 + V_3 + \dots = \sum_{n=1}^{\infty} V_n \quad (1.16)$$

Now by inserting this back into the forward series (Equation 1.10) gives

$$D = G_0 \sum_{n=1}^{\infty} V_n G_0 + G_0 \sum_{n=1}^{\infty} V_n G_0 \sum_{n=1}^{\infty} V_n G_0 + \dots \quad (1.17)$$

Then expanding out the series in V and collecting like order terms gives,

$$\begin{aligned} D &= G_0 V_1 G_0 \\ 0 &= G_0 V_2 G_0 + G_0 V_1 G_0 V_1 G_0 \\ 0 &= G_0 V_3 G_0 + G_0 V_2 G_0 V_1 G_0 + G_0 V_1 G_0 V_2 G_0 + G_0 V_1 G_0 V_1 G_0 V_1 G_0 \end{aligned} \quad (1.18)$$

From (Weglein, et al., 1997) it was stated that for the 3rd order component only one of the terms will correspond to internal multiples ($G_0 V_1 G_0 V_1 G_0 V_1 G_0$) as this is the response of a wavefield with three perturbations. This equation can predict internal multiples from a recorded wavefield but at this stage it will also predict events that are not internal multiples. This leads to the selection of a subset of this to ensure that the events predicted are only internal multiples (Figure 1.2). This is accomplished by the selection of sub events that are initially deeper in the subsurface than shallower, than deeper referred to as the lower-higher-lower criteria (Weglein, et al., 1997).

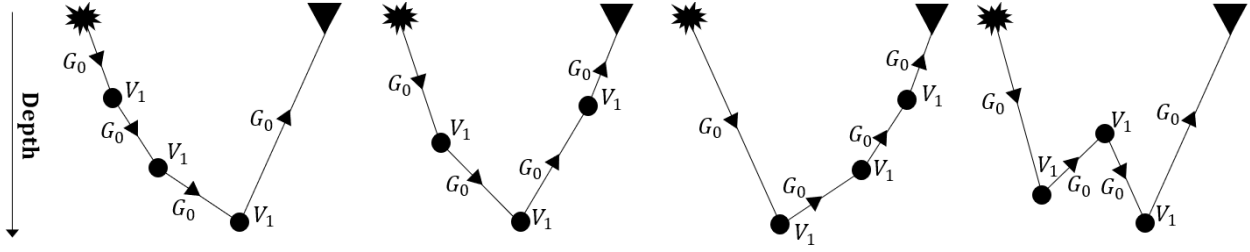


Figure 1.2 Modified from (Weglein et al., 1997) displaying various scattering options from $G_0V_1G_0V_1G_0V_1G_0$ where only far right example corresponds to an internal multiple

The selection of the subevent which will predict internal multiples from the data is carried out through the integration limits in the combination of the three perturbations. We now have a representation of internal multiples that can be solved for using only the input data, giving the equation for internal multiple prediction (Weglein, et al., 1997). The main assumptions of the algorithm are that surface multiples have been removed and that the data only contains primaries and internal multiples. The only additional input required is a term epsilon which is used to account for the bandwidth of the data used for prediction.

$$\begin{aligned}
 & b_3(k_g, k_s, \omega) \\
 &= \frac{1}{(2\pi)^2} \iint_{-\infty}^{\infty} dk_1 e^{-iq_1(\epsilon_g - \epsilon_s)} dk_2 e^{iq_2(\epsilon_g - \epsilon_s)} \int_{-\infty}^{\infty} dz_1 e^{i(q_g + q_1)z_1} b_1(k_g, -k_1, z_1) \\
 & \quad \times \int_{-\infty}^{z_1 - \epsilon} dz_2 e^{-i(q_1 + q_2)z_2} b_1(k_1, -k_2, z_2) \int_{z_2 + \epsilon}^{\infty} dz_3 e^{i(q_2 + q_s)z_3} b_1(k_2, -k_s, z_3), \quad (1.19)
 \end{aligned}$$

Where in Equation 1.19,

$$q_x = \frac{\omega}{c_0} \sqrt{1 - \frac{k_x^2 c_0^2}{\omega^2}}, \quad (1.20)$$

b_3 is the internal multiple prediction, b_1 is the prepared input data, q_x is the vertical wavenumber and ϵ is the depth below free surface of the source (s) and receiver (g), k is the Fourier conjugate variable, z_1, z_2 and z_3 are the depths chosen to satisfy lower-higher-lower relationship, c_0 is the

reference velocity and ε is the search limiting parameter. Epsilon sets a limit on the distance the multiple must have traveled to prevent the method from predicting multiples within the wavelength of a single wavelet.

Equation 1.19 predicts internal multiples through multiplication in the Fourier domain. In time this is equivalent to a combination of convolutions and correlations of these specific events that satisfy the location criteria. It is shown schematically how two deeper events can be convolved relative to a shallower event which is correlated to mimic the equivalent internal multiple (Figure 1.3).

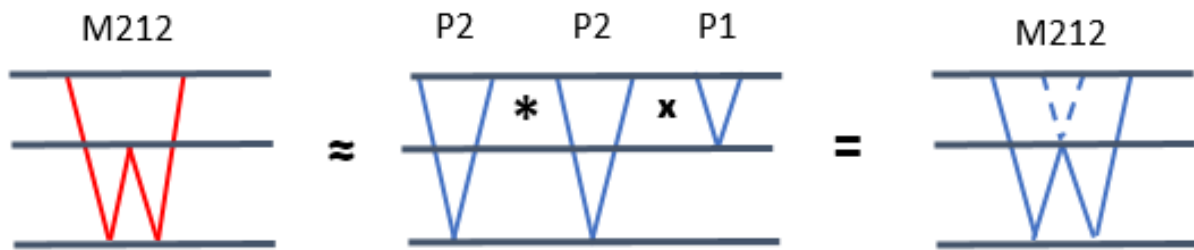


Figure 1.3 Schematic displaying how a multiple can be approximated with a combination of primaries through a convolution (*) and correlation (x)

Equation 1.19 with three perturbations gives a prediction for all orders of internal multiples. This is because the algorithm does not distinguish between primaries and multiples and will use internal multiples from the input to predict higher order internal multiples. The data with internal multiples fully attenuated is given by (Weglein, et al., 1997),

$$D^{IM}(k_g, k_s, \omega) = (-2iq_s)^{-1} \sum_{n=0}^{\infty} b_{2n+1}(k_g, k_s, q_g + q_s) \quad (1.21)$$

This is an infinite series which in practice is commonly truncated to only use the first term in the series (Equation 1.19). A recursive relationship was developed to solve for the higher

order terms in the series (Weglein, et al., 1997). There has been work recently showing the use of higher order terms in the series (Liang, et al., 2013), and this will be explored. The input data and method can be altered so that the procedure is carried out in other domains which has shown increased prediction accuracy (Sun & Innanen, 2014; Innanen, 2017). The derivation of these domains is shown.

1.3.2 Reduction to 1.5D offset-time and 1D time domain

The time offset is a domain which has been recently evaluated (Innanen, 2015b; Innanen, 2017). Equation 1.19 can be simplified and reduced to a 1.5D domain by assuming a $v(z)$ medium. This is accomplished by assuming that the source and receiver wavenumbers are equivalent.

$$k_g = k_s, \quad (1.22)$$

This assumption alters the vertical wavenumber from equation (1.17) to give the following

$$q_g + q_s = 2q_g = k_z, \quad (1.23)$$

Giving the 1.5D Version of the algorithm in frequency and pseudo-depth

$$b_3(k_g, \omega) = \int_{-\infty}^{\infty} dz_1 e^{ik_z z_1} b_1(k_g, z_1) \int_{-\infty}^{z_1 - \varepsilon} dz_2 e^{-ik_z z_2} b_1(k_g, z_2) \\ \times \int_{z_2 + \varepsilon}^{\infty} dz_3 e^{ik_z z_3} b_1(k_g, z_3), \quad (1.24)$$

The first step is to replace b_1 in terms of pseudo depth (z) with S_1 in terms of time (t) and letting $S_1(k_g, t)$ be the Fourier transform of $s_1(x, t)$ over the spatial dimension (Innanen, 2015a) gives,

$$b_3(k_g, \omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} S_1(k_g, t) \int_{-\infty}^{t - \varepsilon} dt_2 e^{i\omega t_2} S_1(k_g, t_2) \int_{t_2 + \varepsilon}^{\infty} dt'' e^{i\omega t''} S_1(k_g, t''), \quad (1.25)$$

For this the output domain is (k_g, ω) as the convolutions and correlations are applied through multiplication in the frequency domain. This can be equivalently rewritten so that the

convolutions and correlations are performed in the time domain. Giving the (k_g, t) version of the algorithm equation.

$$b_3(k_g, t) = \int_{-\infty}^{\infty} dt' S_1(k_g, t' - t) \int_{t'-(t-\varepsilon)}^{t'-\varepsilon} dt'' S_1(k_g, t' - t'') S_1(k_g, t''), \quad (1.26)$$

Then by noting the remaining spatial convolutions applied in frequency can also be written in time gives the offset-time (x, t) version of the algorithm (Innanen, 2017).

$$B_3(x, t) = \int_{-\infty}^{\infty} dx' \int_{-\infty}^{\infty} dt' s_1(x - x', t' - t) \int_{-\infty}^{\infty} dx'' \int_{t'-(t-\varepsilon)}^{t'-\varepsilon} dt'' s_1(x' - x'', t' - t'') s_1(x'', t''), \quad (1.27)$$

Then by assuming there is no spatial component in Equation 1.26

$$k_g = 0, \quad (1.28)$$

The algorithm can be written and reduced to a 1D time version.

$$B_3(t) = \int_{-\infty}^{\infty} dt' s_1(t' - t) \int_{t'-(t-\varepsilon)}^{t'-\varepsilon} dt'' s_1(t' - t'') s_1(t''), \quad (1.29)$$

The inverse scattering series can now be implemented in either time or time and space depending on the required number of dimensions. With the benefit of this being the data input and prediction output domains are identical.

1.3.3 Reduction to 1.5D tau-p domain

The tau-p domain has also been shown to have notable improvements (Sun & Innanen, 2014). Beginning with the previously reduced 1.5D version of the algorithm in frequency and pseudodepth (Equation 1.24) can be written in the tau-p domain as demonstrated in Coates & Weglein (1996) giving

$$b_3(p_g, \omega) = \int_{-\infty}^{\infty} d\tau_1 e^{i\omega\tau_1} b_1(p_g, \tau_1) \int_{-\infty}^{\tau_1-\varepsilon} d\tau_2 e^{-i\omega\tau_2} b_1(p_g, \tau_2) \int_{\tau_2+\varepsilon}^{\infty} d\tau_3 e^{i\omega\tau_3} b_1(p_g, \tau_3) \quad (1.30)$$

In the tau-p domain the data must be prepared to be input into the algorithm outlined below (Sun & Innanen, 2014). For the 1.5D version the input data is prepared by first transforming to the tau-p domain.

$$d(x_g, t) \xrightarrow{\tau p} D(p_g, \tau) \quad (1.31)$$

Then 1D Fourier transformed over τ

$$D(p_g, \tau) \xrightarrow{Fp} D_1(p_g, \omega) \quad (1.32)$$

Then scaled by $-2iq_s$

$$B_1(p_g, \omega) = -2iq_s D_1(p_g, \omega) \quad (1.33)$$

Applying the inverse Fourier transform over ω to give the prepared data for the algorithm.

$$B_1(p_g, \omega) \xrightarrow{iFp} b_1(p_g, \tau) \quad (1.34)$$

1.3.4 Reduction to 1D Pseudo-depth domain

Equation 1.24 can also be reduced to a 1D prediction algorithm (Eaid, et al., 2016). This 1D version will be used to introduce how the method calculates multiples. The 1D version of the algorithm assumes that there is no spatial dimension.

$$k_g = k_s = 0, \quad (1.35)$$

Then q_g reduces to

$$q_g = 2 \frac{\omega}{c_0}, \quad (1.36)$$

The original 2D equation (Equation 1.19) then reduces to the following 1D equation

$$b_3(\omega) = \int_{-\infty}^{\infty} dz_1 e^{i2\frac{\omega}{c_0}z_1} b_1(z_1) \int_{-\infty}^{z_1-\varepsilon} dz_2 e^{-i2\frac{\omega}{c_0}z_2} b_1(z_2) \int_{z_2+\varepsilon}^{\infty} dz_3 e^{i2\frac{\omega}{c_0}z_3} b_1(z_3), \quad (1.37)$$

The data preparation for the 1D version of the algorithm is displayed. First by first Fourier transforming the input data

$$d(t) \xrightarrow{F} D(\omega), \quad (1.38)$$

Then using the follow change of variables from frequency to wavenumber

$$k_z = \frac{2\omega}{c_0}, \quad (1.39)$$

Then Inverse Fourier transformed to pseudo depth

$$D(k_z) \xrightarrow{iF} b_1(z), \quad (1.40)$$

Where the pseudo-depth variable z is

$$z = \frac{c_0 t}{2}, \quad (1.41)$$

When implementing the method numerically the 1D version of the algorithm can be further simplified using a Heaviside step function (Eaid, et al., 2016). This has not altered the effectiveness of the equation as no new assumptions have been made, it simply reduces computational expense.

$$b_3(\omega) = \int_{-\infty}^{\infty} dz_1 e^{-i2\frac{\omega}{c_0}z_1} b_1(z_1) \left[\int_{z_1+\varepsilon}^{\infty} dz_2 e^{i2\frac{\omega}{c_0}z_2} b_1(z_2) \right]^2, \quad (1.42)$$

1.4 Analytic example in 1D

A 1D analytic example is completed to demonstrate how the method predicts internal multiples from an input data set. Using a geologic model that has two layers plus a half space the resulting seismic trace will contain two primary events and a multiple train with the first

order multiple referred to as M212 (Figure 1.1). Giving the following amplitudes for the events in the seismic trace.

$$\begin{aligned}
P_1 &= R_1 \\
P_2 &= T_{01}R_2 T_{10} \\
M_{212} &= T_{01}R_2(-R_1) R_2T_{10},
\end{aligned} \tag{1.43}$$

Where R is the zero-offset reflection coefficient with the subscript referencing the interface number, T is the transmission coefficient and P and M are the primary and multiple events. With corresponding traveltimes,

$$\begin{aligned}
t_1 &= 2 \frac{z_1}{v_1} \\
t_2 &= 2 \frac{z_1}{v_1} + 2 \frac{z_2 - z_1}{v_2} \\
t_{212} &= 2 \frac{z_1}{v_1} + 4 \frac{z_2 - z_1}{v_2},
\end{aligned} \tag{1.44}$$

The input data b_1 is given as follows

$$b_1(z) = P_1\delta(z - z_1) + P_2\delta(z - z_2) + M_{212}\delta(z - z_{212}) + \text{Higher Order IM} \tag{1.45}$$

Inserting this into the innermost integral of Equation 1.42 where \hat{z} is used to denote the integration variable and distinguish from z for the pseudo-depths of the layers.

$$\begin{aligned}
I_1(\hat{z}_1) &= \int_{\hat{z}_1 + \varepsilon}^{\infty} d\hat{z}_2 e^{i2\frac{\omega}{c_0}\hat{z}_2} [P_1\delta(\hat{z}_2 - z_1) + P_2\delta(\hat{z}_2 - z_2) + M_{212}\delta(\hat{z}_2 - z_{212}) \\
&\quad + \text{Higher Order IM}],
\end{aligned} \tag{1.46}$$

Solving the innermost integral gives the following result

$$= \begin{cases} P_1 e^{i2\frac{\omega}{c_0}z_1}, & z_1 > \hat{z}_1 + \varepsilon \\ 0, & z_1 < \hat{z}_1 + \varepsilon \end{cases} + \begin{cases} P_2 e^{i2\frac{\omega}{c_0}z_2}, & z_2 > \hat{z}_1 + \varepsilon \\ 0, & z_2 < \hat{z}_1 + \varepsilon \end{cases} + \begin{cases} M_{212} e^{i2\frac{\omega}{c_0}z_{212}}, & z_{212} > \hat{z}_1 + \varepsilon \\ 0, & z_{212} < \hat{z}_1 + \varepsilon \end{cases} + \dots \tag{1.47}$$

Can also be written as followed with the use of Heaviside step function

$$\begin{aligned}
I_1(\hat{z}_1) &= P_1 e^{i2\frac{\omega}{c_0}z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] + P_2 e^{i2\frac{\omega}{c_0}z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] \\
&\quad + M_{212} e^{i2\frac{\omega}{c_0}z_{212}} H[z_{212} - (\hat{z}_1 + \varepsilon)], \tag{1.48}
\end{aligned}$$

Squaring Equation 1.48 will give the result to be used in the next integral from Equation 1.39

$$\begin{aligned}
I_2(\hat{z}_1) &= P_1^2 e^{i2\frac{\omega}{c_0}2z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] + 2P_1P_2 e^{i2\frac{\omega}{c_0}(z_2+z_1)} H[z_1 - (\hat{z}_1 + \varepsilon)] \\
&\quad + P_2^2 e^{i2\frac{\omega}{c_0}2z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] + \dots, \tag{1.49}
\end{aligned}$$

Inserting Equation 1.49 into the outermost integral in Equation 1.42 integral gives

$$\begin{aligned}
b_3(\omega) &= \int_{-\infty}^{\infty} d\hat{z}_1 e^{-i2\frac{\omega}{c_0}\hat{z}_1} [P_1\delta(\hat{z}_1 - z_1) + P_2\delta(\hat{z}_1 - z_2) + M_{212}\delta(\hat{z}_1 - z_{212}) \\
&\quad + \text{Higher Order IM}] \times [P_1^2 e^{i2\frac{\omega}{c_0}2z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] \\
&\quad + 2P_1P_2 e^{i2\frac{\omega}{c_0}(z_2+z_1)} H[z_1 - (\hat{z}_1 + \varepsilon)] + P_2^2 e^{i2\frac{\omega}{c_0}2z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] + \dots], \tag{1.50}
\end{aligned}$$

Truncating this to only include the primary events

$$\begin{aligned}
b_3(\omega) &= \int_{-\infty}^{\infty} d\hat{z}_1 e^{-i2\frac{\omega}{c_0}\hat{z}_1} [P_1\delta(\hat{z}_1 - z_1) + P_2\delta(\hat{z}_1 - z_2)] \times \left[P_1^2 e^{i2\frac{\omega}{c_0}2z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] + \right. \\
&\quad \left. 2P_1P_2 e^{i2\frac{\omega}{c_0}(z_2+z_1)} H[z_1 - (\hat{z}_1 + \varepsilon)] + P_2^2 e^{i2\frac{\omega}{c_0}2z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] \right], \tag{1.51}
\end{aligned}$$

And rearranging for both primaries

$$\begin{aligned}
b_3(\omega) &= \int_{-\infty}^{\infty} d\hat{z}_1 e^{-i2\frac{\omega}{c_0}\hat{z}_1} P_1\delta(\hat{z}_1 - z_1) \times \left[P_1^2 e^{i2\frac{\omega}{c_0}2z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] + \right. \\
&\quad \left. 2P_1P_2 e^{i2\frac{\omega}{c_0}(z_2+z_1)} H[z_1 - (\hat{z}_1 + \varepsilon)] + P_2^2 e^{i2\frac{\omega}{c_0}2z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] \right] + \\
&\quad \int_{-\infty}^{\infty} d\hat{z}_1 e^{-i2\frac{\omega}{c_0}\hat{z}_1} P_2\delta(\hat{z}_1 - z_2) \times \left[P_1^2 e^{i2\frac{\omega}{c_0}2z_1} H[z_1 - (\hat{z}_1 + \varepsilon)] + \right. \\
&\quad \left. 2P_1P_2 e^{i2\frac{\omega}{c_0}(z_2+z_1)} H[z_1 - (\hat{z}_1 + \varepsilon)] + P_2^2 e^{i2\frac{\omega}{c_0}2z_2} H[z_2 - (\hat{z}_1 + \varepsilon)] \right], \tag{1.52}
\end{aligned}$$

Then solving gives the following result

$$\begin{aligned}
b_3(\omega) = & P_1^3 e^{i2\frac{\omega}{c_0}z_1} H[z_1 - (z_1 + \varepsilon)] + 2P_1^2 P_2 e^{i2\frac{\omega}{c_0}z_2} H[z_1 - (z_1 + \varepsilon)] \\
& + P_1 P_2^2 e^{i2\frac{\omega}{c_0}(2z_2 - z_1)} H[z_2 - (z_1 + \varepsilon)] + P_1^2 P_2 e^{i2\frac{\omega}{c_0}(2z_1 - z_2)} H[z_1 - (z_2 + \varepsilon)] \\
& + 2P_1 P_2^2 e^{i2\frac{\omega}{c_0}z_1} H[z_1 - (z_2 + \varepsilon)] + P_2^3 e^{i2\frac{\omega}{c_0}z_2} H[z_2 - (z_2 + \varepsilon)], \tag{1.53}
\end{aligned}$$

Now applying the Heaviside step functions as shown

$$\begin{cases} 0, & H[z_1 - (z_1 + \varepsilon)] \text{ for all } \varepsilon \\ 0, & H[z_1 - (z_2 + \varepsilon)] \text{ for all } \varepsilon \\ 0, & H[z_2 - (z_2 + \varepsilon)] \text{ for all } \varepsilon \\ 1, & H[z_2 - (z_1 + \varepsilon)] \text{ for } \varepsilon < (z_2 - z_1) \end{cases} \tag{1.54}$$

Results in the final truncated solution in the Fourier domain

$$b_3(\omega) = P_1 P_2^2 e^{i2\frac{\omega}{c_0}(2z_2 - z_1)}, \tag{1.55}$$

Then applying the inverse Fourier transform gives the internal multiple prediction.

$$b_3(t) = R_1 T_{01} R_2 T_{10} T_{01} R_2 T_{10} \delta(t - (2t_2 - t_1)), \tag{1.56}$$

Comparing this back to the multiple defined in Equation 1.43 and 1.44 the inverse scattering series has predicted the time of the multiple exactly. The amplitude prediction is off by the transmission coefficients from both the downgoing (T_{01}) and upgoing (T_{10}) transmission coefficients across the downward generating horizon (interface 1). The error in the amplitude arises from the outermost integral and is a function of the “generator” depth \hat{z}_1 in the analytic example, or z_1 from Equation 1.42. This is due to the P_2 event having traveled through the downward generator. Even given this analytic example with infinite bandwidth and no noise there are still a small amplitude errors present in the prediction. This transmission loss error has been previously shown (Zhang & Shaw, 2010; Pan & Innanen, 2013; Ramirez, 2013).

How the method predicts internal multiples was displayed for P-wave only data. Considerable effort has recently been undertaken to predict and attenuate converted wave internal multiples. Converted waves must be handled differently as great care must be taken when converting both P and S wave events to pseudo depth (Sun, et al., 2017a). Both the preparation and prediction has been displayed successfully (Sun, et al., 2017b).

1.5 Short vs. long path multiples

The inverse scattering series prediction algorithm can only predict long path multiples. If short path multiples present in the data this would need to be resolved with a different approach. To predict internal multiples from the data the events need to be sufficiently separated in time to be combined. This can be thought similarly to the tuning effect that displays how two seismic reflections interact as the time between events varies (Widess, 1973). This tuning thickness and the seismic response are often displayed with a wedge model (Figure 1.4).

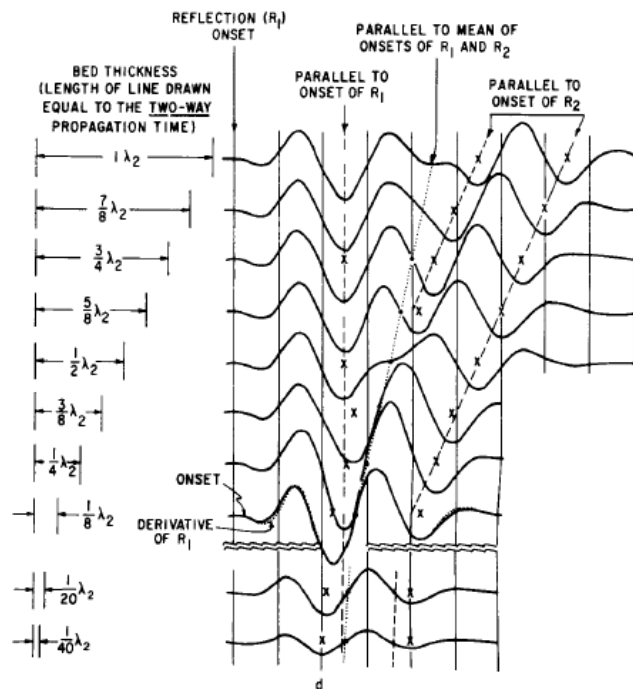


Figure 1.4 Tuning wedge for 3 layer model from (Widess, 1973)

Like the tuning wedge, a model is created with varying thickness to test the prediction algorithm for short vs long path multiples. This is done using a zero offset assumption with velocities displayed and convolved with a 40Hz Ricker wavelet with the primaries and internal multiples. Using the 1D inverse scattering series equation this model is solved for various layer thickness using the noted values of epsilon (Figure 1.5).

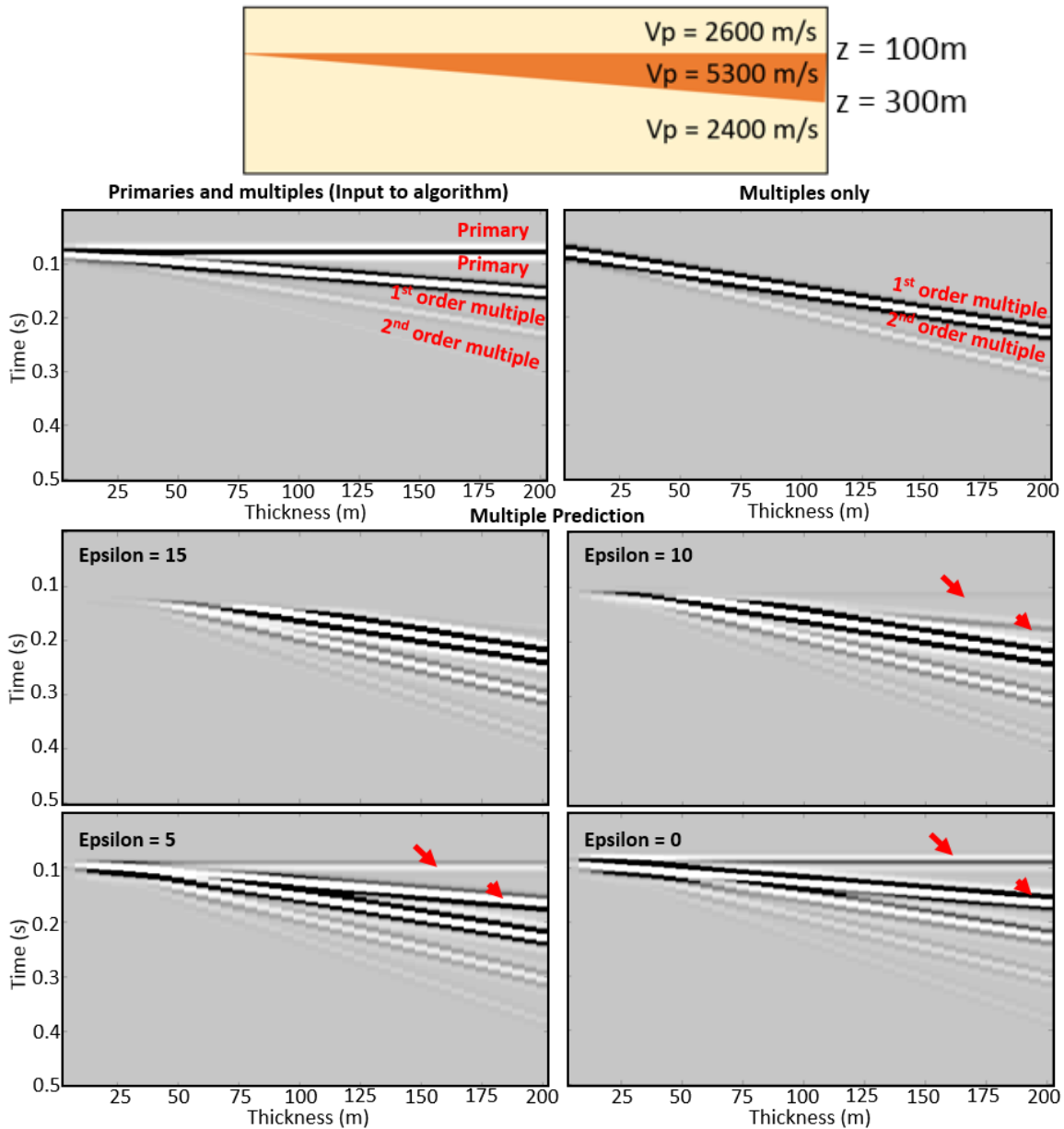


Figure 1.5 Tuning wedge model for internal multiples Internal multiple prediction from ISS for tuning wedge model with variable epsilon

The image with epsilon equal to zero exaggerates the issue when epsilon becomes small and the primary reflection begins to predict itself. A more reasonable epsilon value such as 15 displays how the internal multiples can be predicted without predicting primaries. With the larger epsilon once the two layers become sufficiently close and the multiples transitions from long to short path the multiple can no longer be predicted. Through the choice epsilon an explicit definition is made for what is termed a short path or long path multiple. As any sub events which are separated by a distance greater than epsilon are defined as long path where any sub events less than epsilon are considered short path.

Even in this simple model the algorithm is unable to predict all internal multiples when the distance between the primaries becomes sufficiently small. This then raises the question of how to judge the success of the algorithm. It may not be reasonable for the algorithm to predict and remove all multiples in a dataset. When using an increasingly complex model there will be an increased opportunity for internal multiples which consist of both short and long path multiples in quick succession of interfering.

1.6 Adaptive subtraction

It has been shown how small amplitude differences exist between the prediction and the actual internal multiples due to the downward generator. In practice with real data there can also be other sources which impact the prediction. If there are changes to the wavelet or nonstationary features this can cause a prediction with an amplitude mismatch. Ideally the internal multiples would be kinematically separated from the primaries to build the filter, but this will generally not be the case. Due to amplitude mismatches a filter is built to match the

predicted internal multiples to the data. This is a critical step as we want to correct for differences between the predicted and real internal multiples without damaging the primaries. Another cause for the requirement of a filter even for an ideal noise free synthetic is due to the way the method operates on bandlimited data show as the issue of the autoconvolution of a wavelet.

1.6.1 Autoconvolution of a wavelet

The main operations in the inverse scattering series prediction algorithm is convolution and correlation. To understand the impact of these operations a simple test is completed by autoconvolving a 40Hz (Figure 1.6).

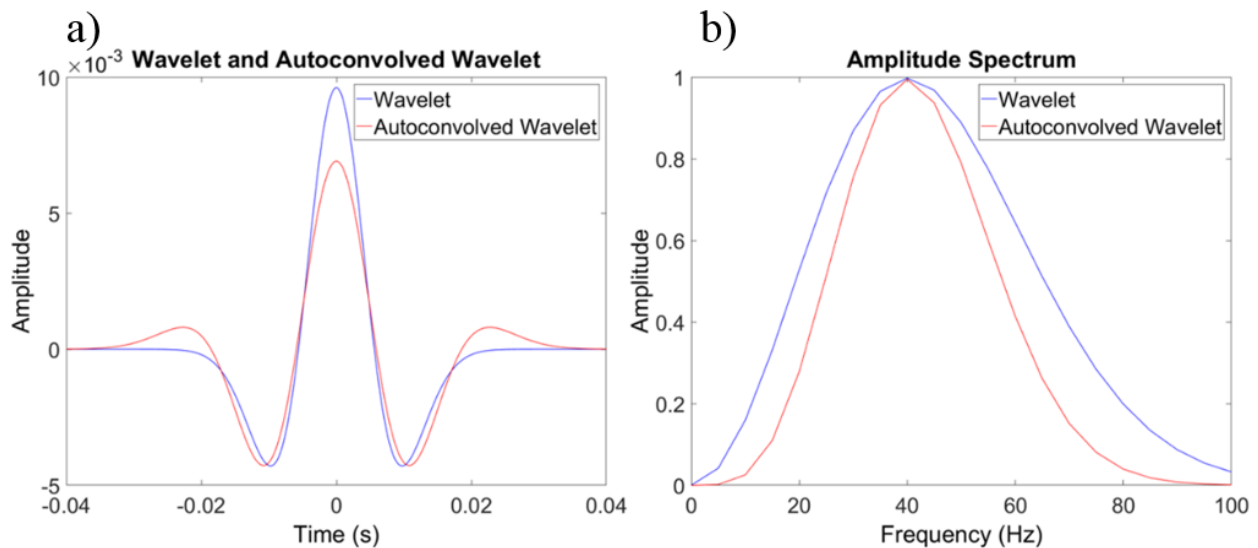


Figure 1.6 a) 40 Hz Ricker wavelet and autoconvolution of the wavelet, b) amplitude spectrum of the 40 Hz Ricker wavelet and autoconvolution of the wavelet

The result in the time domain is both the slight decrease in amplitude of the wavelet and the addition of the sidelobes. In the amplitude spectrum, the dominant frequency has remained the same for the prediction but there are frequencies lost at both the high and low end of the spectrum.

1.6.2 Theory

Adaptive subtraction is an approach to solve for a filter which will match two datasets by minimize a function. An adaptive subtraction algorithm has been developed and has been previously tested on internal multiple attenuation (Keating, et al., 2015). The tools utilized for this project are provided by CREWES. In one-dimension adaptive subtraction is carried out by some form of the following,

$$a = d - m * f = d - Mf \quad (1.57)$$

Where a is the minimizing term, d is the input data, m is the predicted multiples and f is the filter. It is shown either by convolving with the filter or through matrix multiplication where M is the internal multiple prediction as a convolution matrix. How you minimize a can be completed using various normalizations. This choice of minimization function will have an impact on the results and the effectiveness of the subtraction (Keating, et al., 2015). Once choice of normalization is the L2 norm given by

$$\|a\|_2 = \|d - Mf\|_2 = \sum_{n=1}^N a_n^2 \quad (1.58)$$

This is minimizing the square root of the sum of the squares. The filter can be solved with the least squares equation,

$$f = (M^T M)^{-1} M^T d \quad (1.59)$$

Another choice of normalization is the L1 norm and is given by

$$\|a\|_1 = \|d - Mf\|_1 = \sum_{n=1}^N |a_n| \quad (1.60)$$

For any of the minimization terms selected the filter design must be careful about how aggressive and nonstationary the filter can be. This is largely controlled through choice of filter length (Keating, et al., 2015). Since the goal of the objective function is to find a minimum, then

there can be an issue of the filter also minimizing primary energy. For synthetics this can be easier to manage as it is often known which events are primaries and multiples, but in practice this is not known. It was displayed how using a hybrid norm of L1 and L2 assisted the internal multiple prediction and allowed for the subtraction of unwanted multiples from a trace and reduced the impact to primaries in the data (Keating, et al., 2015).

1.7 Thesis Overview

Two techniques with the inverse scattering series internal multiple attenuation algorithm are introduced, developed and discussed. First, a proposed new dimension to the algorithm is developed. The higher order terms in the series are also implemented. These higher order terms were found through a different line of reasoning than the original mathematically rigorous proof. Also developed in this work is the importance of how the higher order terms are implemented.

Second, a domain for the algorithm is also explored in time and offset which was recently developed (Innanen, 2015b). The impact of this domain to the multidimensional problem with a nonstationary epsilon is explored for the first time and shown to be advantageous. Irregular spatial sampling and the impact on internal multiple prediction in this time-offset domain is also explored.

These changes to the implementation of the method are then brought together on both synthetic and real data. The workflow to combine these additions to the method to produce an optimal prediction to attenuate internal multiples in a systematic way is outlined. Much of the work that is shown in this thesis has been presented and shown in other reports and conference

abstracts. The chapters 1-4 were largely described by CREWES reports (Iverson & Innanen, 2017; Iverson, et al., 2017; Iverson, et al., 2018e) and Geoconvention abstracts (Iverson, et al., 2018c; Iverson, et al., 2019) and seg abstracts (Iverson, et al., 2018b). Chapters 5 and 6 were shown in CREWES reports (Iverson, et al., 2018d; Iverson, et al., 2018a).

Chapter Two: The Downward Generator Space

2.1 Numerical example with 1.D synthetic data

The 1D version of the inverse scattering series internal multiple prediction algorithm is implemented on a simple geologic model where the velocities and depths are displayed below assuming a constant density (Figure 2.1). Model parameters were chosen to produce internal multiples that have high amplitudes due to the large impedance contrasts and occur in distinct locations from the primary events. The 1D model was also chosen to strictly evaluate the zero offset case without any additional impacts from amplitude variations with offset. The goal is to evaluate the accuracy of the ISS prediction on a simple 1D model.

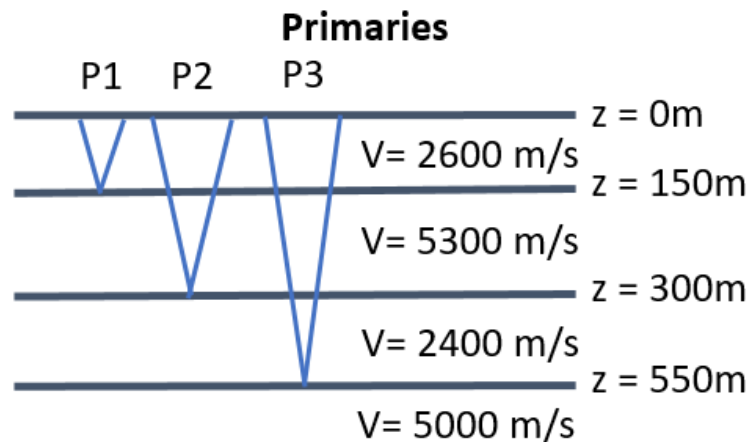


Figure 2.1 Velocity and depth model used for the 1D prediction.

2.1.1 Modeling parameters

The primaries, first and second order internal multiples were computed using zero offset reflection and transmission coefficients to create a reflectivity series with sample rate 0.0001s, then convolved with a 40Hz Ricker wavelet and resampled to 0.002s to create the seismic trace (Figure 2.2).

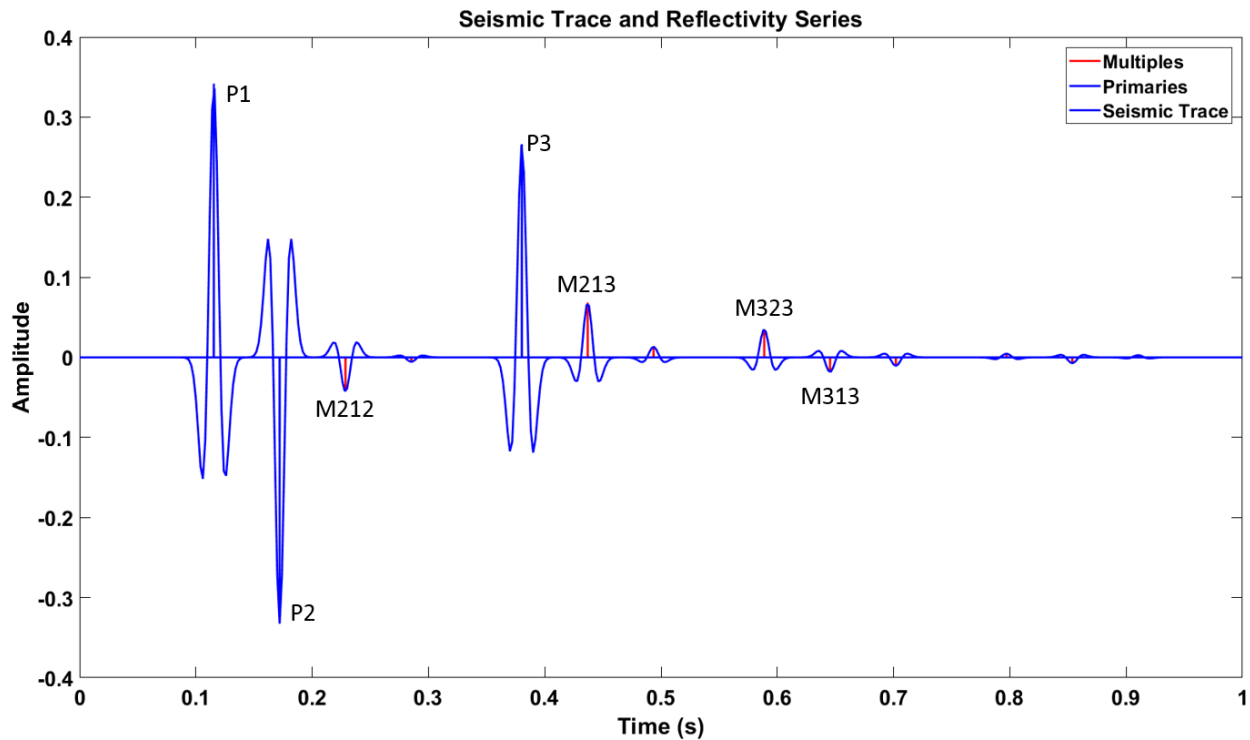


Figure 2.2 Reflectivity series for primaries, internal multiples and resulting seismic trace

2.1.2 Internal multiple prediction

The internal multiple prediction is carried out on the synthetic trace and the results of the prediction are displayed in Figure 2.3. This prediction was completed using an epsilon value of 15 due to the bandwidth and sampling rate of the data. At a minimum the output from the prediction requires a global scalar so the prediction has approximately the same amplitude as the input trace. In this example the global scalar was calculated by matching the peak amplitudes of the internal multiple (M212). This event was chosen to give a reasonable value for the global scalar. In practice it would not be known which events are primaries and multiples, but an estimate for a single scalar would still be required.

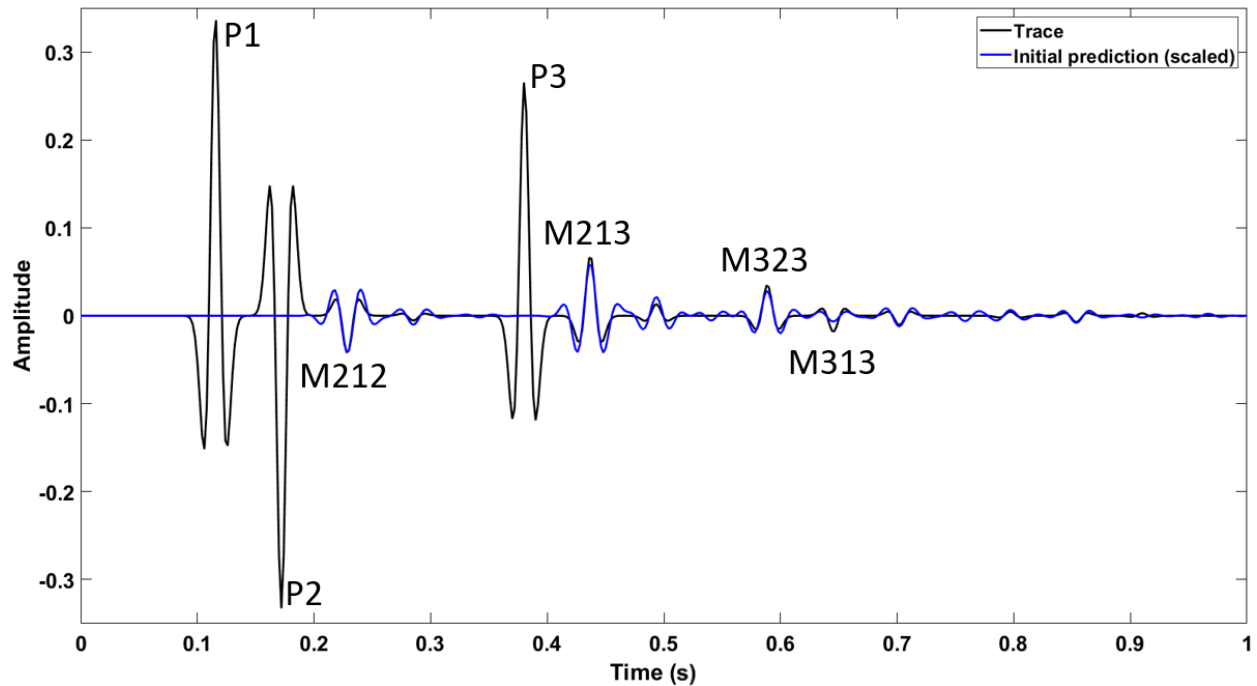


Figure 2.3 Input seismic trace and 1D internal multiple prediction

All internal multiples from the input data have been predicted correctly in time and the primaries have not been predicted. The amplitudes for the predictions are reasonably matched to those of the input trace. The wavelets on the prediction appear to have been altered relative to the input trace and now contain additional sidelobes, due to the previously noted autocorrelation of band limited data.

2.2 Downward Generator Space

In practice the algorithm is implemented numerically by solving each frequency slice for all possible pseudo-depth locations which obey the lower-higher-lower criteria. From Equation 1.42 the pseudo-depth location of the outermost integral varies at the location of the downward generator. The analytical example (Equations 1.43 - 1.56) displayed how the predicted amplitudes will be in error by the transmission loss across the downward generator. What is

proposed is to store the prediction from Equation 1.42 for the given downward generator giving the output prediction with two variables. To achieve the standard output the prediction is summed over all downward generators (z_1). Then taking this two-dimensional output and inverse Fourier transform giving over the frequency direction gives

$$b_3(z_1, \omega) \xrightarrow{iF} b_3(z_1, t) . \quad (2.1)$$

This has been termed the downward generating space. This space is displayed with time on the horizontal-axis and pseudodepth on the vertical-axis. This space which is a function of pseudodepth and time was created for the previous prediction and is displayed (Figure 2.4). This downward generator space is similar to the space created in the XIMP process referred to as multiple contributing gather (MCG) (Wu & Dragoset, 2011).

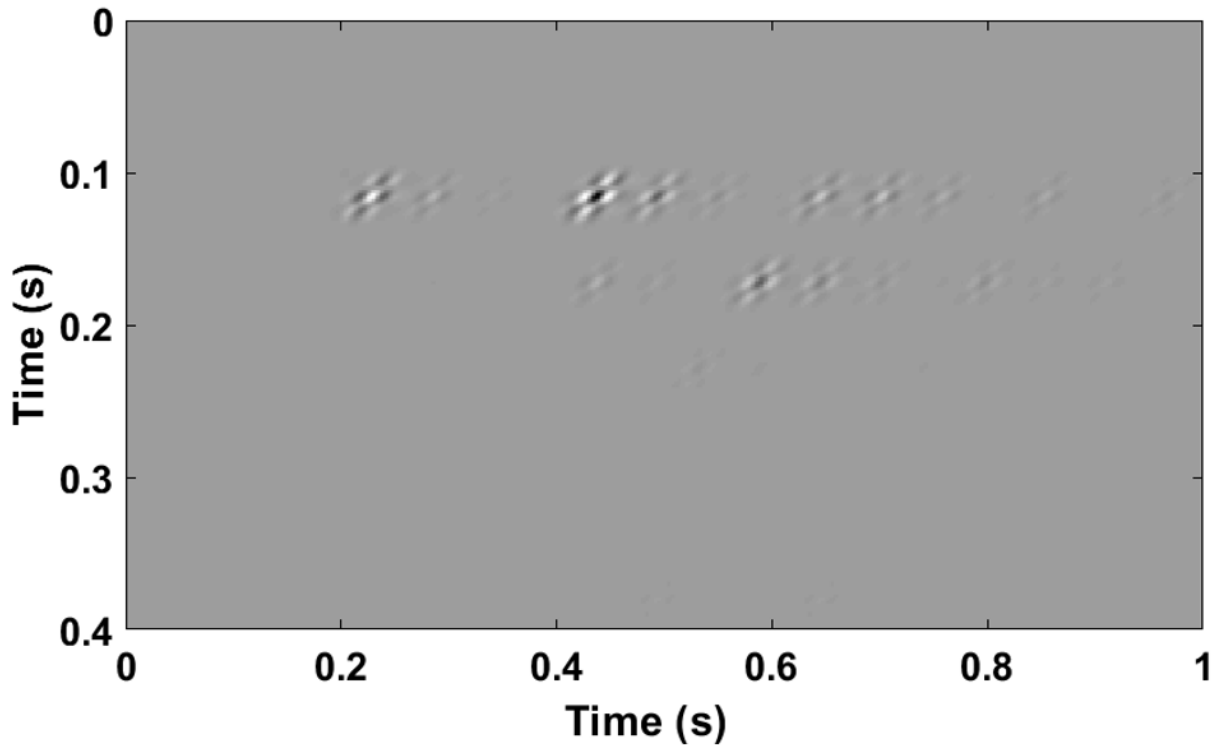


Figure 2.4 Downward generator space displaying individual internal multiples. Prediction time is on the horizontal axis and downward generator time on the vertical axis

The predicted internal multiples can be viewed in this uncollapsed space. Every two-dimensional wavelet like event is an individual internal multiple. For this input model there are only four first order multiples so the majority of the events are higher order internal multiples. Also displayed is the resulting multiple train from each downward generator. This space shows how at a given prediction time (all pseudo-depths at one time slice) there can be several internal multiples. If the adaptive subtraction and matching is completed after the data has been summed over all downward generators then that information is lost, and these amplitude issues can no longer be rectified with a single filter.

This space allows for two avenues to correct the amplitudes of the predict internal multiples. The first is to apply a 2D adaptive subtraction in this downward generator space (Keating, et al., 2015). Since the vertical axis is the dimension which causes amplitude errors this should be a more natural space for the adaptive subtraction to create the filter and minimize error. This 2D adaptive subtraction is similar to the iterative adaptive subtraction used for each multiple model from a selected generator horizon from XIMP (Wu & Dragoset, 2011). The second is to apply a scalar that varies with the vertical downward generator axis to account for the amplitude issues that arise such as transmission losses. This is similar to the multiple elimination algorithm which calculates the transmission losses from the input data alone within the algorithm (Ramirez & Weglein, 2008; Zou & Weglein, 2015).

2.3 Adaptive subtraction in 2D

To create the standard prediction the entire length in the pseudo depth dimension is stacked. The entire downward generator space could be input into the adaptive subtraction, but

in practice the space is partially stacked over the downward generator dimension. This is done for two reasons. First is to reduce the computational expense by reducing the number of traces, the second is to at a minimum stack over the width of the wavelet and to reduce the likelihood of overfitting. There may be an epsilon relationship here which could assist in choosing a stack size.

2.3.1 Theory

The standard approach takes the 1D trace and prediction with some type of energy minimization (Chapter 1.6). In the approach proposed here instead of solving for a single filter, now multiple filters are solved for over the downward generator direction and minimize the difference between the sum of the combination of these to the single dimension initial data trace.

$$\phi(f) = \| b_1(t) - b_3(z_1, t)f \|_p^p, \quad (2.2)$$

Where f is the filter and p is a chosen parameter. To prevent overfitting in the downward generator dimension a penalty term is added and minimize the following

$$\phi(f) = \| b_1(t) - b_3(z_1, t)f \|_p^p + \phi_R, \quad (2.3)$$

Where ϕ_R is a term which penalizes variation in the vertical axis. This gives a minimization function which will create a series of 1D filters. Arranging these 1D filters into a matrix, we can use matrix multiplication with the DGS to produce the adaptive subtracted multiple trace.

2.3.2 Application to 1D numerical example

Using the previous one-dimensional numerical example and prediction both the standard approach of adaptive subtraction on the 1D predicted trace and the 2D space is implemented with

the 2D adaptive subtraction. First the standard approach with adaptive subtraction is displayed (Figure 2.5).

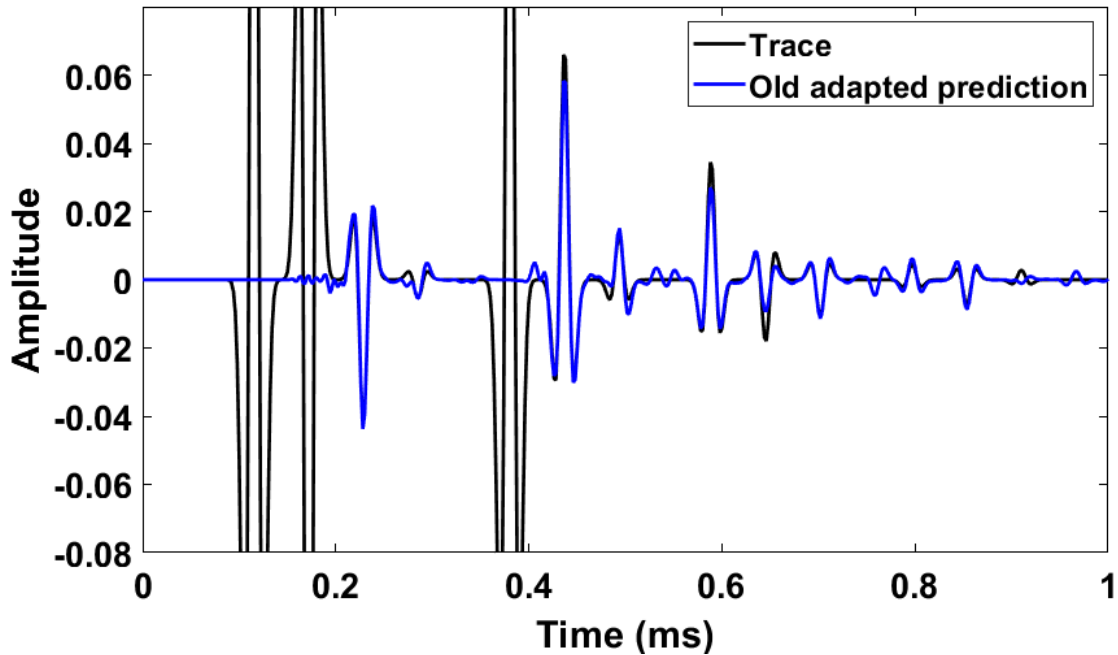


Figure 2.5 Internal multiple prediction with 1D adaptive subtraction

In general, the filter has been able to adjust the amplitudes and a reasonable prediction is made. There is still an amplitude mismatch which occurs at approximately 0.65 seconds. This is then put into the 2D adaptive subtraction and the result is displayed (Figure 2.6). This 2D trace is partially stacked along the row direction to give 10 traces.

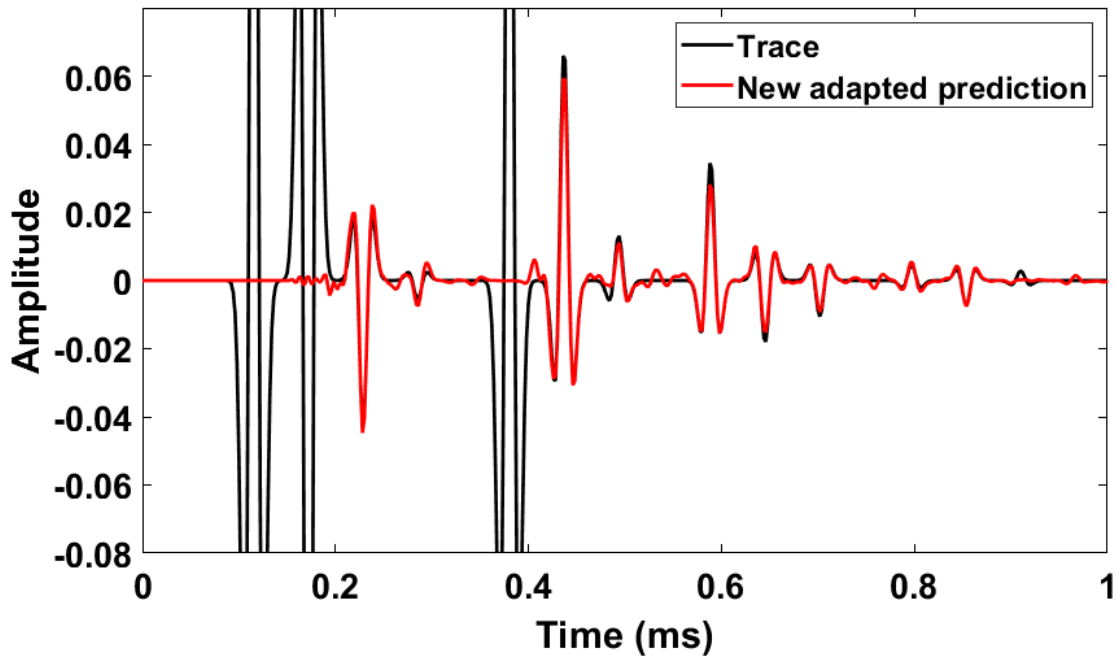


Figure 2.6 Internal multiple prediction with 2D adaptive subtraction

It is displayed how utilizing the 2D downward generator space to carry out the adaptive subtraction can improve the prediction. Adaptive subtraction in this space can attempt to account for the transmission losses due to the downward generator which the original algorithm does not account for. The method also remains data driven as there are still no subsurface information requirements.

2.4 Pseudo-depth variant scalar

Another option to adjust the amplitudes of the prediction is to create and apply a scalar which will correct for the transmission losses. This can be done without using the downward generator space by making a slight addition to the standard algorithm. A change to the order of operations for the numerical algorithm of the internal multiple prediction is done. This produces a more natural location to alter the equation and account for the transmission loss at the

generator. This is introduced into the last step where the prediction at a given depth z is multiplied by the scalar $\varphi(z)$. The pseudo code for how this is implemented is shown (

Figure 2.7).

<pre> a) for ik = kz1:kzmax { lpos = exp(i*kzPos(ik)*z) lneg = exp(-i*kzPos(ik)*z) intPos = b1z*lpos; intNeg = b1z*lneg; for iz = z1:zmax { inner = sum(intPos(iz+ε:zmax)) pred(ik) = pred(ik) + intNeg(iz)*inner*inner } } </pre>	<pre> b) for iz = z1:zmax { for ik = kz1:kzmax { lpos = exp(i*kzPos(ik)*z) lneg = exp(-i*kzPos(ik)*z) intPos = b1z*lpos; intNeg = b1z*lneg; inner = sum(intPos(iz+ε:zmax)) pred(iz, ik) = intNeg(iz)*inner*inner } pred(iz, :) = pred(iz, :)*φ(iz) } spred(ik) = sum(pred(iz, ik)) </pre>
---	--

Figure 2.7 a) Summing over wavenumber then depth, b) proposed order of operations alteration summing over depth then wavenumber with scalar applied to give scaled prediction

Another option is to use the downward generator space and simply apply the scalar to all times for a given pseudo-depth. Then to obtain the final prediction sum over all pseudo-depths. This additional scalar can be displayed in either the frequency domain as

$$sb_3(z_1, \omega) = \varphi(z_1) \times b_3(z_1, \omega), \quad (2.4)$$

Where sb_3 is the scaled version of b_3 . Or applied in the time domain after inverse Fourier transforming gives

$$sb_3(z_1, t) = \varphi(z_1) \times b_3(z_1, t), \quad (2.5)$$

2.4.1 Scalar calculation

One concern that immediately arises from applying a scalar to the prediction is that some subsurface information must be assumed. As the scalar is a precalculated depth dependent value to account for losses. One of the key benefits of the original inverse scattering series multiple attenuation method is that it is data driven. If some prior knowledge of the subsurface is now assumed, then this may be considered too significant an alteration to the original method. The internal multiple elimination algorithm has a similar goal to this and uses the data to predict the attenuation factor (Ramirez & Weglein, 2008; Zou & Weglein, 2015). If there is noise in the data or a data driven attenuation factor calculation is struggling, then a pre-calculated scalar may still assist. This can also be calculated to include any other losses such as those due to geometric spreading or attenuation. The goal of the addition of the scalar is to attempt to correctly calculate the amplitudes with the utilization of additional physics to reduce the load on adaptive subtraction.

For the previous geologic model the scalar is calculated by using the velocities and depths from the model and calculating the transmission loss with depth. In practice this could be implemented with a sonic and density log to calculate the transmission loss. For this example, the first and last values were extrapolated to zero depth and final depth and the intermediate values were linearly interpolated (Figure 2.8). A block model was not used so that there to prevent a step change at the location of the downward generator as this would significantly impact the amplitude across the width of the wavelet.

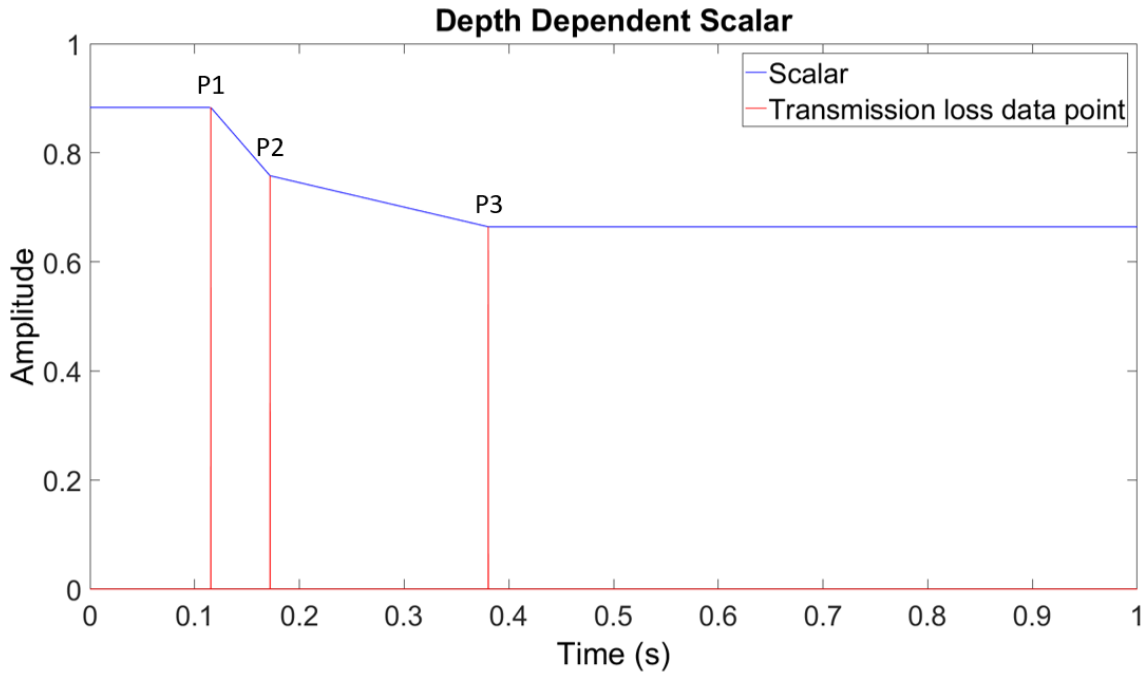


Figure 2.8 Transmission loss scalar for internal multiple prediction.

The result from both the original prediction and the scalar applied prediction is displayed

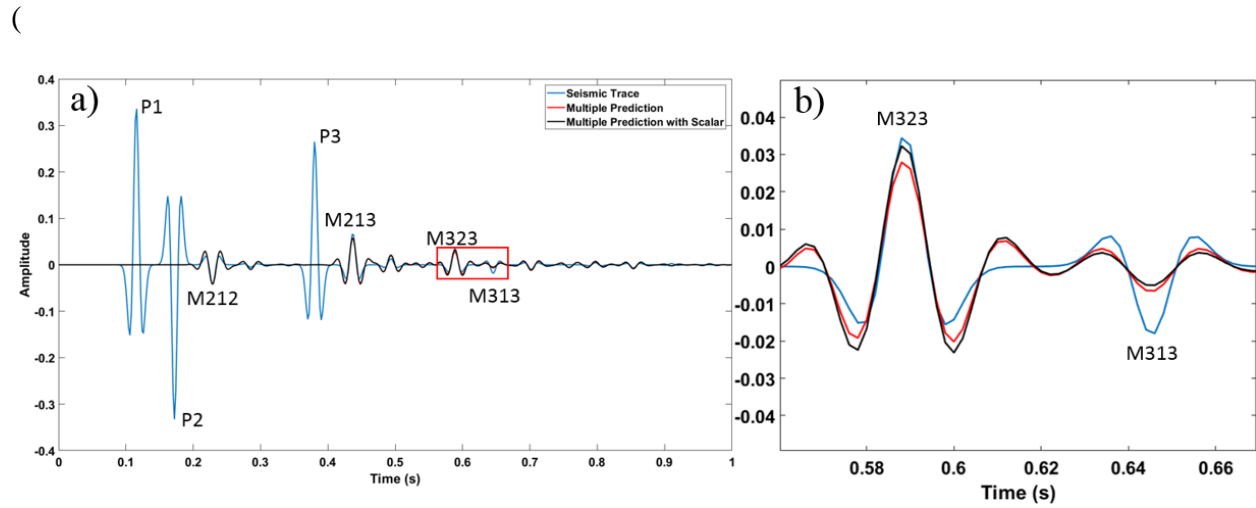


Figure 2.9). Similarly, to the previous example both have a single global scalar applied to compare the differences. This global scalar was computed by matching the amplitudes of the first multiple peak (M212). Both versions of the method have accurately predicted all multiples in the data. There exist small variations throughout the trace but in general both predictions are

comparable. The minimal differences between the two reflect the small adjustment made to the amplitude of the prediction. The largest impact of the depth dependent scalar should be located at the internal multiple M323.

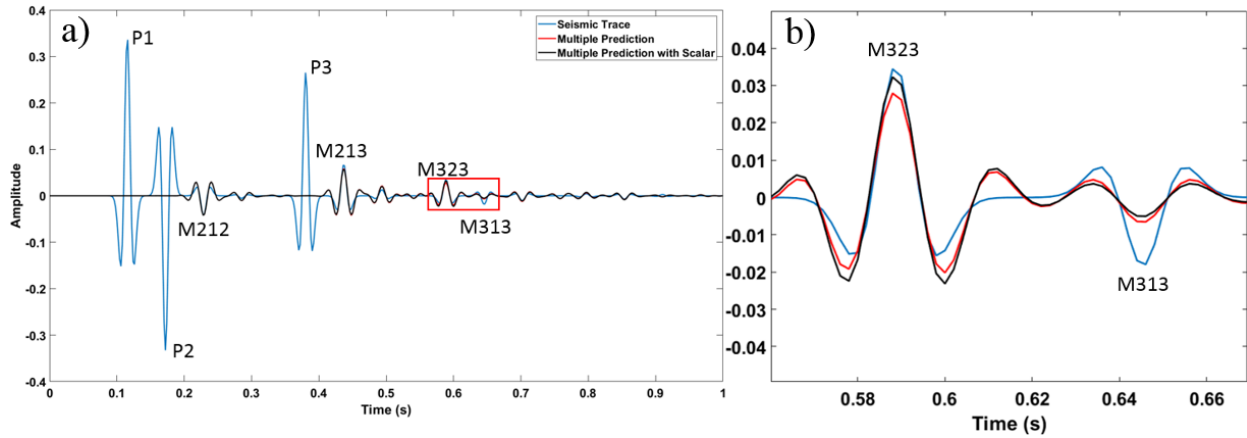


Figure 2.9 a) Trace with both scaled and unscaled predictions, b) isolating two multiples M323 and M313

This internal multiple will have transmission effects due to the second interface not previously accounted for. In

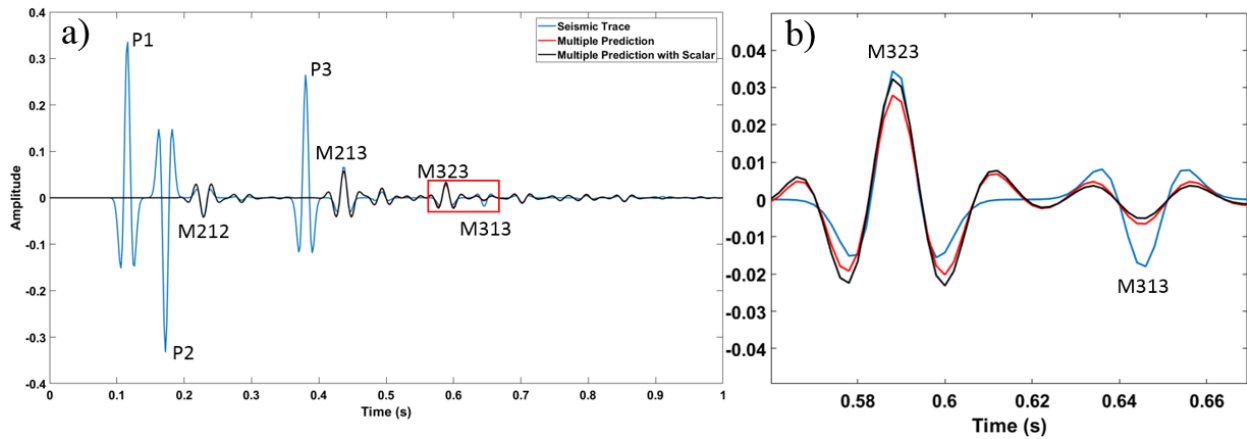


Figure 2.9 the peak amplitude of the multiple prediction for the scaled version has better accounted for extra transmission terms relative to the unscaled version. In

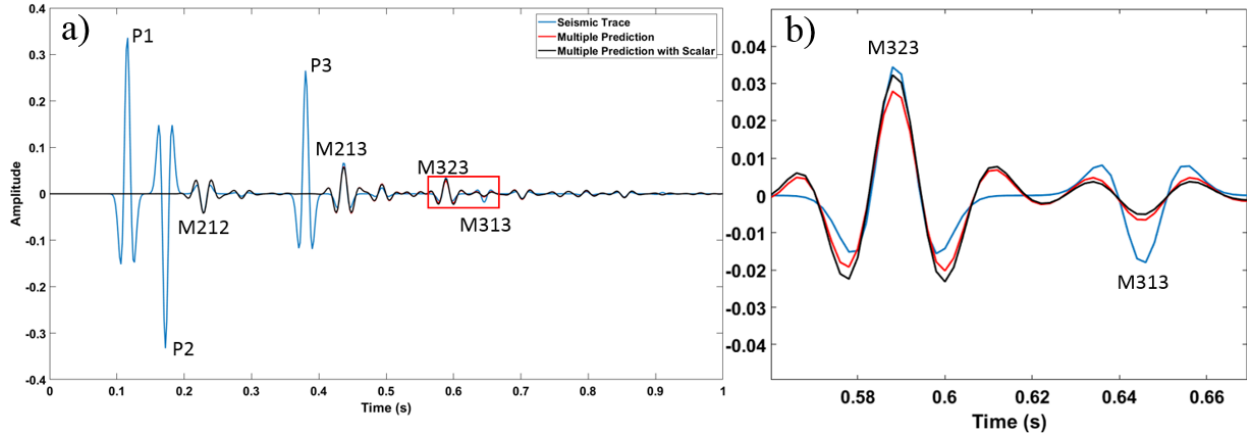


Figure 2.9 there is a slightly lower amplitude prediction for the multiple M313. Thus, this depth dependent scalar has improved the prediction for one multiple and been detrimental to another. Compared to the previous example with 2D adaptive subtraction (Figure 2.6) which could overcome this issue. This variation in success is due to higher order multiples which will be scrutinized in Chapter 3.

2.5 Numerical example with synthetic data in 1.5D

The previous examples all used zero offset traces. While this is a useful step both to understand the problem theoretically and could be applied to post stack or zero offset data there is a need to apply the method prestack. The downward generator space is displayed with a spatial component by creating data which has both time and offset. Displayed is the velocity and depth model, again assuming constant density with the primaries and internal multiple shown (Figure 2.10).

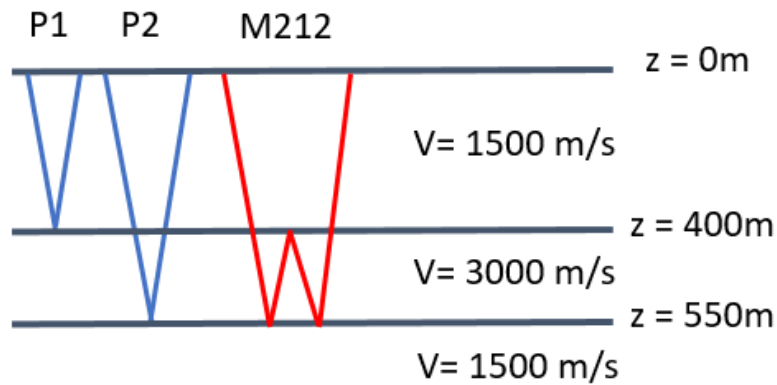


Figure 2.10 Geologic model displaying velocities and depths used

2.5.1 Modeling Parameters

Using finite difference modeling a shot record was created (afd_shotrec in CREWES MATLAB toolbox). Finite difference allowed for the modeling of all orders of multiples within the recorded window. The 2D model was spatially sampled every 10m and a temporal sample rate of 0.002s, with a grid that is 512x256 samples. The seismic shot record was created by convolving the result with a 30Hz Ricker wavelet.

2.5.2 internal multiple prediction

The 1.5D tau-p version of the algorithm (Equation 1.30) will be used to predict the internal multiples. The input data must be prepared prior to being put in the internal multiple prediction algorithm. From the CREWES Toolbox the first step involved applying a surface mute to ensure there are no erroneous values near the shot location. Also, in the tau-p preparation function in the CREWES toolbox, is a spatial cosine taper that is applied to minimize artifacts at the edges of the data and prediction (Sun & Innanen, 2014). The data is then transformed into the tau-p domain and a scale factor from Weglein ISS theory is applied. Due to the medium only varying in the vertical direction the prediction will only be completed on the

positive slowness values. The resulting model and internal multiple prediction using an epsilon value of 30 is displayed (Figure 2.11).

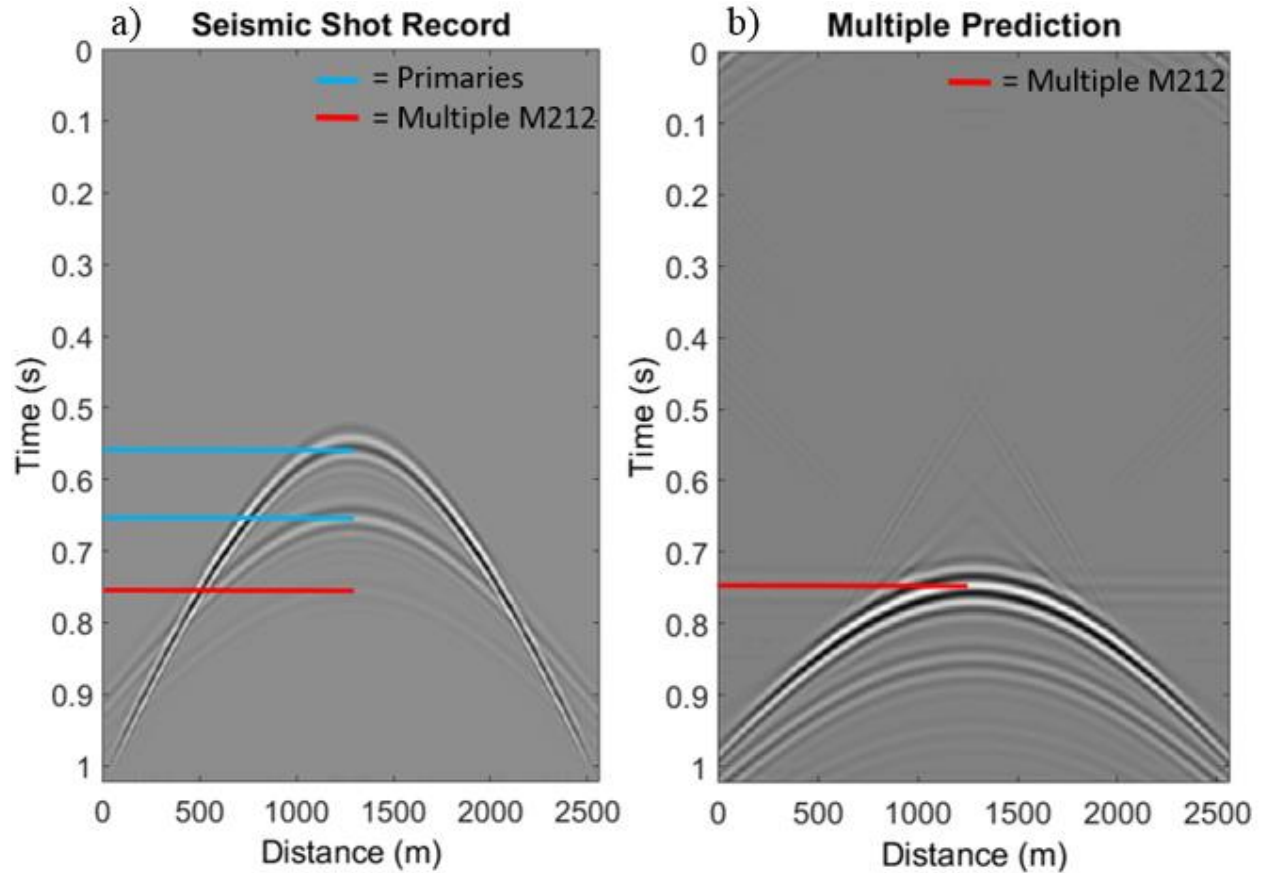


Figure 2.11 (Left) Synthetic seismic shot record (Right) 1.5D tau-p multiple prediction

The internal multiples have been accurately predicted in the data set. The chosen epsilon value appears to have been sufficiently small to allow for the prediction of the multiple but not so small as to predict energy from the primaries.

2.5.3 Downward Generator Space with a spatial dimension

The 1D problem had a two-dimensional DGS, now the 1.5D problem with time and offset has a three-dimensional DGS. This has been completed by storing the individual slices and since

there is a time and space component this must be stored in a 3rd dimension. This is displayed for all slices (Figure 2.12). In practice, when adaptively subtracting or applying a scalar it may be computationally beneficial to store the data into stacks, as the size of the matrix goes from (Nt by Nx) to (Nt by Nt by Nx). Displayed in the figure is the new dimension added to this 1.5D problem.

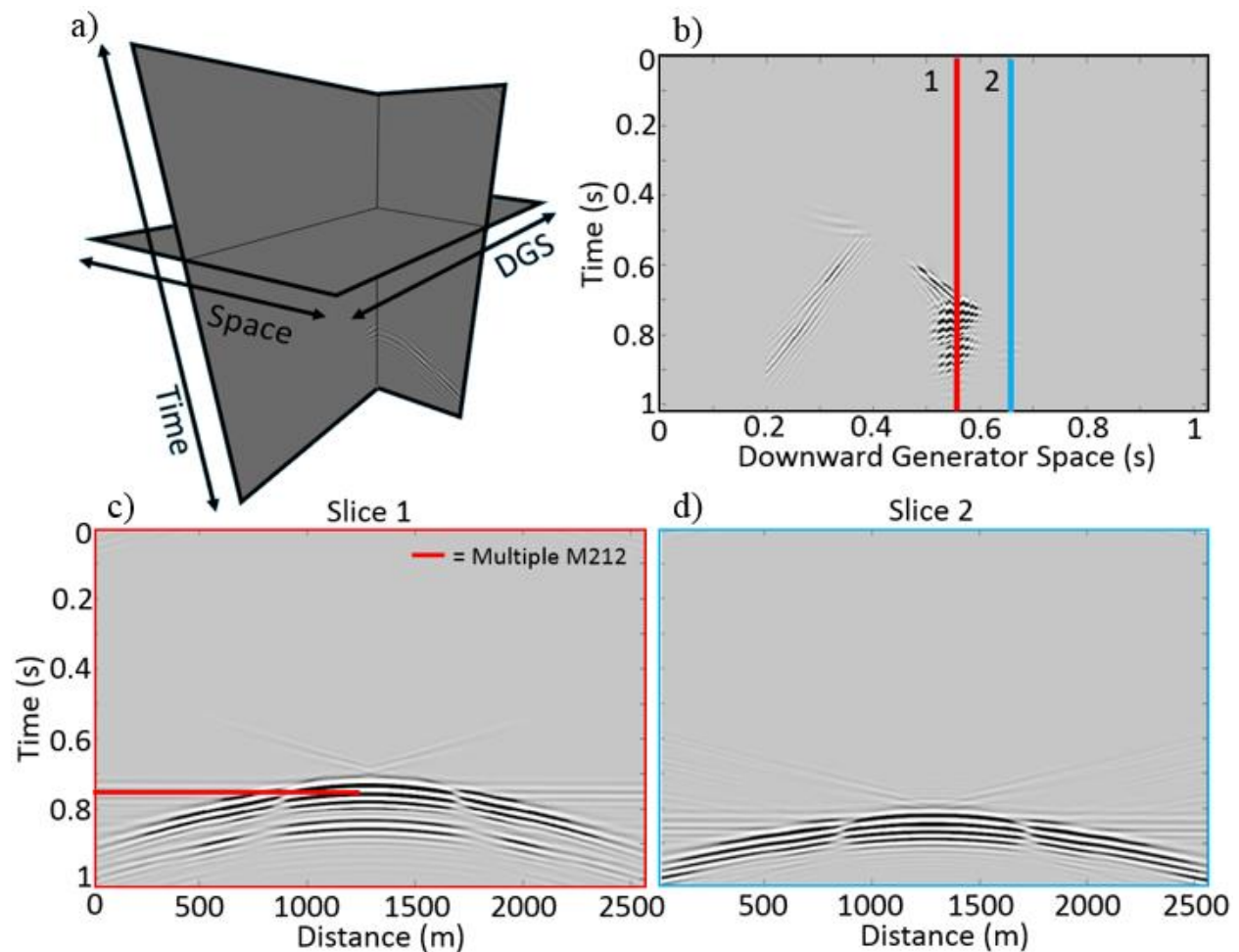


Figure 2.12 a) three dimensional view of the 3D DGS, b) a slice through zero offset similar to standard 2D DGS c). One DGS slice displaying the first order and higher order internal multiples, d) A deeper DGS slice displaying only higher order internal multiples

Looking at the single zero offset slice this is comparable to the 1D case. The individual multiples can be identified and the generator which caused the given multiple train is also identifiable. It can be seen here how the at the pseudo-time of the first primary all multiples can be seen. Then at the second primary pseudo-time the second order multiple is displayed. Looking at the single slice it does not appear to be identical to the final standard prediction. Again this must be summed over all pseudo-times to create the prediction wavelet.

2.6 Conclusions

This new dimension has been displayed for both the 1D and 1.5D case, with the goal of giving increased flexibility to correct amplitude errors that can not be rectified after the prediction is collapsed over this new dimension. This new space gives numerous new insights into the way the prediction algorithm functions. Both, from an increased understanding of the prediction and the ability to correct amplitude errors, either through some pre-calculated scalar or using a multidimensional adaptive subtraction. The addition of an extra dimension could also lead to the overpredicting of the data and the removal of primary energy. Constraints must be placed with the adaptive subtraction to force the algorithm to be smoothly varying over the new dimension. As real data propagates there are the effects of transmission loss, geometric spreading, attenuation, mode conversions all of these may affect the amplitudes at different points in the seismic trace. Through processing the data many of these effects are attempted to be corrected for and other noise sources are attempted to be removed. If there are any residual amplitude errors with the processed data, this space may also allow for the correction of these. This may assist in adding a new dimension to better fit the amplitudes of the resulting prediction.

Chapter Three: Higher Order Terms

3.1 Residual amplitude mismatch

Chapter two demonstrated how errors in the predicted amplitudes of internal multiples arise due to the downward generating horizon. Two methods were displayed to account for these errors. Using either the downward generator space and a 2D adaptive subtraction or a pre-calculated scalar. The two approaches improved the prediction, but amplitude mismatches remain compared with the input data. The cause of the residual amplitude mismatch is higher order internal multiples. The standard approach only uses the first term from the inverse scattering series. To further assist the prediction of higher order internal multiples, higher order terms from the inverse scattering series are investigated.

3.2 Second order internal multiples

The first term in the inverse scattering series takes all events from the input data to predict internal multiples. When internal multiples are present in the input data these events are also used in the algorithm. When internal multiples are applied in the algorithm these combine in the same manner as primaries through convolutions and correlations to predict higher order internal multiples. These higher order multiple predictions must still obey the lower-higher-lower relationship. This is another benefit of the algorithm as even the first term in the series uses the input data to predict all orders of multiples. The problem that arises is the second order multiples are generally overpredicted (Zhang & Shaw, 2010). This further complicates the issues with the predicted amplitudes as first order multiples are generally underpredicted and the second order multiples are overpredicted (Zhang & Shaw, 2010). These second order and higher internal

multiples are overpredicted because a single multiple can be predicted from several combinations of sub events (Figure 3.1).

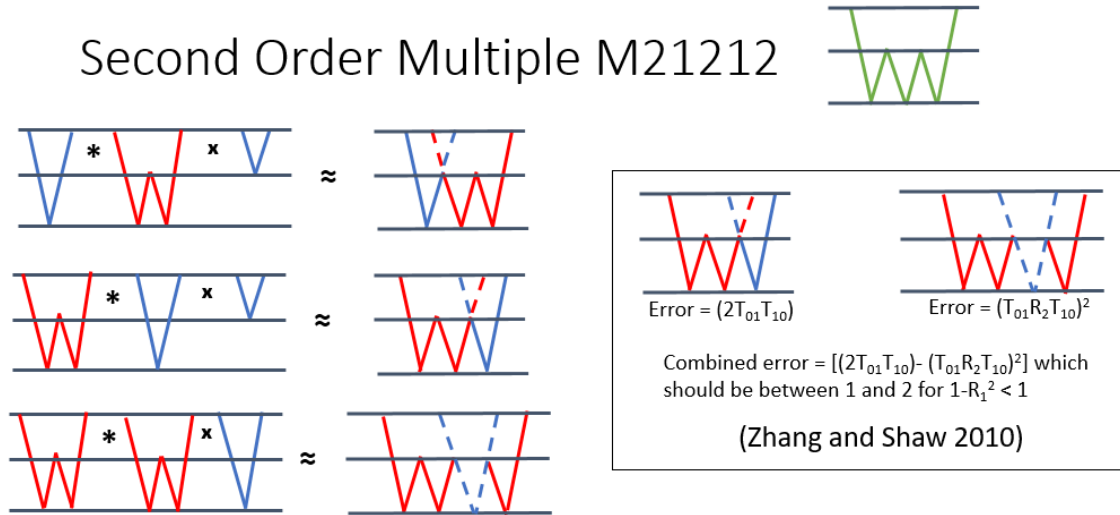


Figure 3.1 Displaying the algorithm computing second order multiples with convolution (*) and correlation (x)

The 1D analytic prediction in chapter one (Equations 1.43 - 1.56) was truncated to only include the primaries, which were combined to create the internal multiple prediction for the multiple M212. Therefore, only the single internal multiple prediction was solved for and not the infinite number of higher order multiples. Artifacts have also been shown to arise when using a multiple as the higher event in the lower-higher-lower criteria (Liang, et al., 2013; Ma & Weglein, 2015).

Displayed in Figure 3.2 is the seismic trace using the geologic model from chapter two (Figure 2.2) with both first and second order multiples modeled. At the location of the

underpredicted M313 multiple from

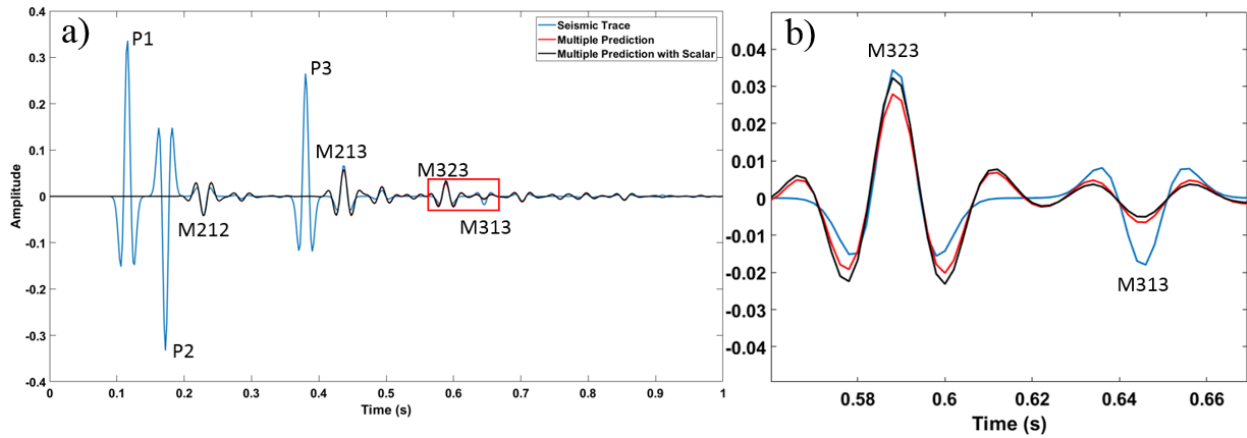


Figure 2.9 there is a second order multiple at this temporal location with opposite polarity compared to the first order multiple.

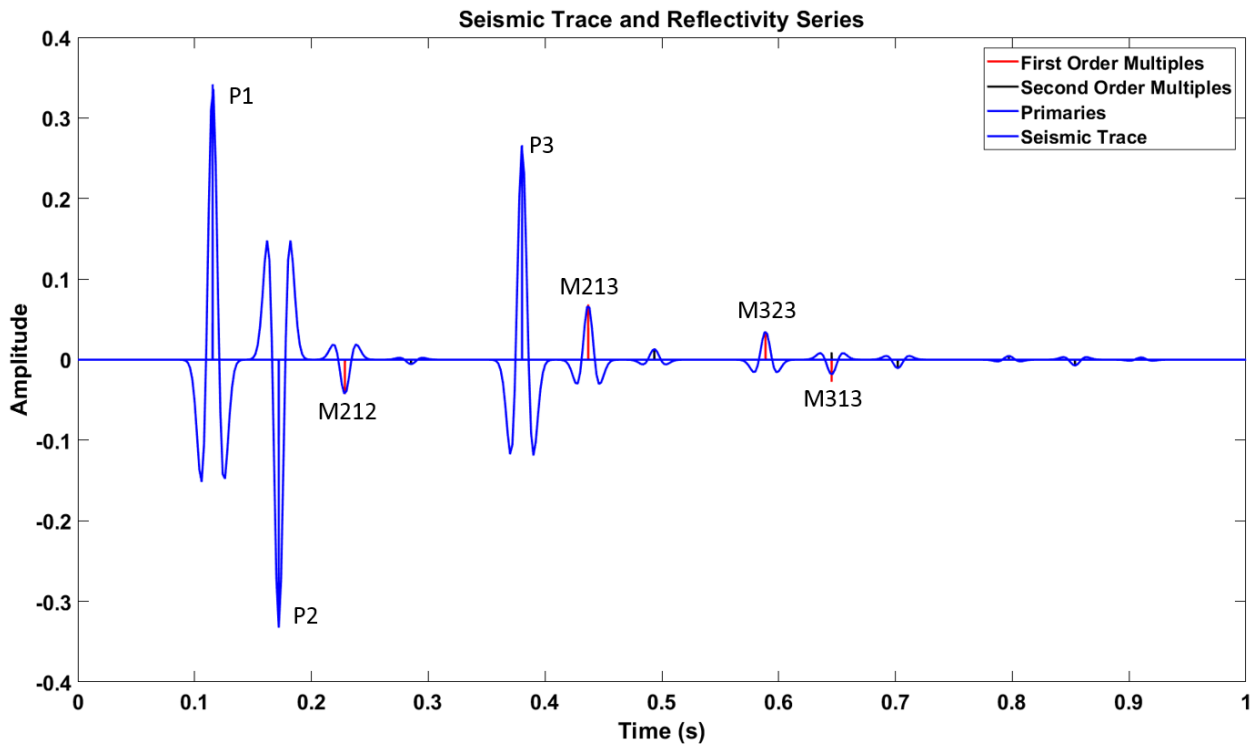


Figure 3.2 Reflectivity series and trace for primaries, first and second order multiples

In general, the amplitudes of the higher order internal multiples are small relative to the primaries and first order multiples. The amplitude correction necessary to adjust the prediction

is more complex than a global scalar, even if it varies with pseudo-depth. At the location of the multiple 313 there is a negative polarity first order multiple and positive polarity second order multiple, and this would require two different scalars at the same temporal location. Using internal multiples from the input data in the algorithm will also be impacted by transmission losses. The pseudo-depth varying scalar impacts this prediction as it is also applied to the multiples in the trace, though the scalar was designed from the primary transmission loss. Similarly, the 2D adaptive subtraction will attempt to account for the amplitude variations between first order and higher order internal multiples. The issue that is shown is that at a single temporal location there may be multiple scalars required due to coincident first order and higher order internal multiples.

Since the higher order multiples require a different amplitude adjustment as they have been overpredicted, what is required is a means to decouple the higher order multiples from the first order multiples. By isolating the higher order internal multiple predictions, the duplicate predictions of the same event can be removed.

3.3 Inverse scattering series terms

The original algorithm was derived by selecting a subset of the inverse scattering series which obeys the lower-higher-lower criteria. The issue due to these overpredicted higher order multiples is addressed with higher order terms in the inverse scattering series. The original algorithm calculates b_3 from the input data b_1 used in all portions of the algorithm (Equation 3.1).

$$b_3(\omega) = \int_{-\infty}^{\infty} dz_1 e^{i2\frac{\omega}{c_0}z_1} b_1(z_1) \int_{-\infty}^{z_1-\varepsilon} dz_2 e^{-i2\frac{\omega}{c_0}z_2} b_1(z_2) \int_{z_2+\varepsilon}^{\infty} dz_3 e^{i2\frac{\omega}{c_0}z_3} b_1(z_3), \quad (3.1)$$

The higher order terms are calculated by taking the multiple prediction output (b_3) and using this as one of the inputs into the algorithm (Liang, et al., 2013). The first higher order term to be used is referred to as b5PPI since it uses the two primary traces and then the internal multiple trace (Equation 3.2) (Liang, et al., 2013).

$$b_5^{PPI}(\omega) = \int_{-\infty}^{\infty} dz_1 e^{i2\frac{\omega}{c_0}z_1} b_1(z_1) \int_{-\infty}^{z_1-\varepsilon} dz_2 e^{-i2\frac{\omega}{c_0}z_2} b_1(z_2) \int_{z_2+\varepsilon}^{\infty} dz_3 e^{i2\frac{\omega}{c_0}z_3} b_3(z_3), \quad (3.2)$$

It should be noted that the primary trace contains both primaries and internal multiples. The internal multiple prediction from the first term in the series (b_3) must be accurate and should only contain internal multiples. This is critical to ensure that only higher order multiples are predicted with the higher order terms. By having one of the input traces being the predicted internal multiples there are theoretically no primaries in this trace. The importance of how the primary and internal multiple traces (b_1 and b_3) are combined and scaled is discussed.

There is also the potential for pure artifacts to arise in the prediction when only the first term is used. To remove these artifacts the term referred to as b5PIP (Equation 3.3) (Liang, et al., 2013) is also implemented. By using a similar approach but with the internal multiple trace used in a different portion of the algorithm these artifacts can be attenuated (Liang, et al., 2013).

$$b_5^{PIP}(\omega) = \int_{-\infty}^{\infty} dz_1 e^{i2\frac{\omega}{c_0}z_1} b_1(z_1) \int_{-\infty}^{z_1-\varepsilon} dz_2 e^{-i2\frac{\omega}{c_0}z_2} b_3(z_2) \int_{z_2+\varepsilon}^{\infty} dz_3 e^{i2\frac{\omega}{c_0}z_3} b_1(z_3), \quad (3.3)$$

3.3.1 Implementation of higher order terms

The higher order terms require the output from the original algorithm (b_3). A critical step to successful implementation of the higher order terms (b_5^{PPI} and b_5^{PIP}) is the scaling of the prediction prior to being used as input for the higher order terms. It has been noted that the term b_5^{PPI} required a scalar of 2 to account for both the b_5^{PPI} and b_5^{IPP} (Liang, et al., 2013). It will be displayed for these tests that an initial scalar of one is used successfully. This appears to be due to the attempting to remove just one of the additional predictions. The amplitude corrections discussed in chapter two can be used to assist the higher order predictions. If the higher order terms are to be applied successfully then an accurate estimate of the internal multiple trace must be known including the amplitudes of the internal multiples. Some options to correct the b_3 trace is to use a single global scalar, use standard 1D or 2D adaptive subtraction or the pseudo-depth varying scalar or the internal multiple elimination algorithm. One advantage with using adaptive subtraction is that it will both correct the amplitudes and adjust the frequency content due to convolution process adding the additional sidelobes. If the sidelobes are not accounted for then this effect will become compounded as the b_3 trace is used again in the algorithm. Once the terms are calculated the final internal multiple prediction can be made by summing the terms (Liang, et al., 2013),

$$IM(t) = b_3(t) + b_5^{PPI}(t) + b_5^{PIP}(t) \quad (3.4)$$

This can be similarly done if in the downward generator space.

3.4 Numerical example with 1D synthetic data

Using the 1D synthetic trace (Figure 3.2), the internal multiple prediction is made, and the downward generator space is used to display the resulting prediction of the b_3 and the two b_5 higher order terms (Figure 3.3).

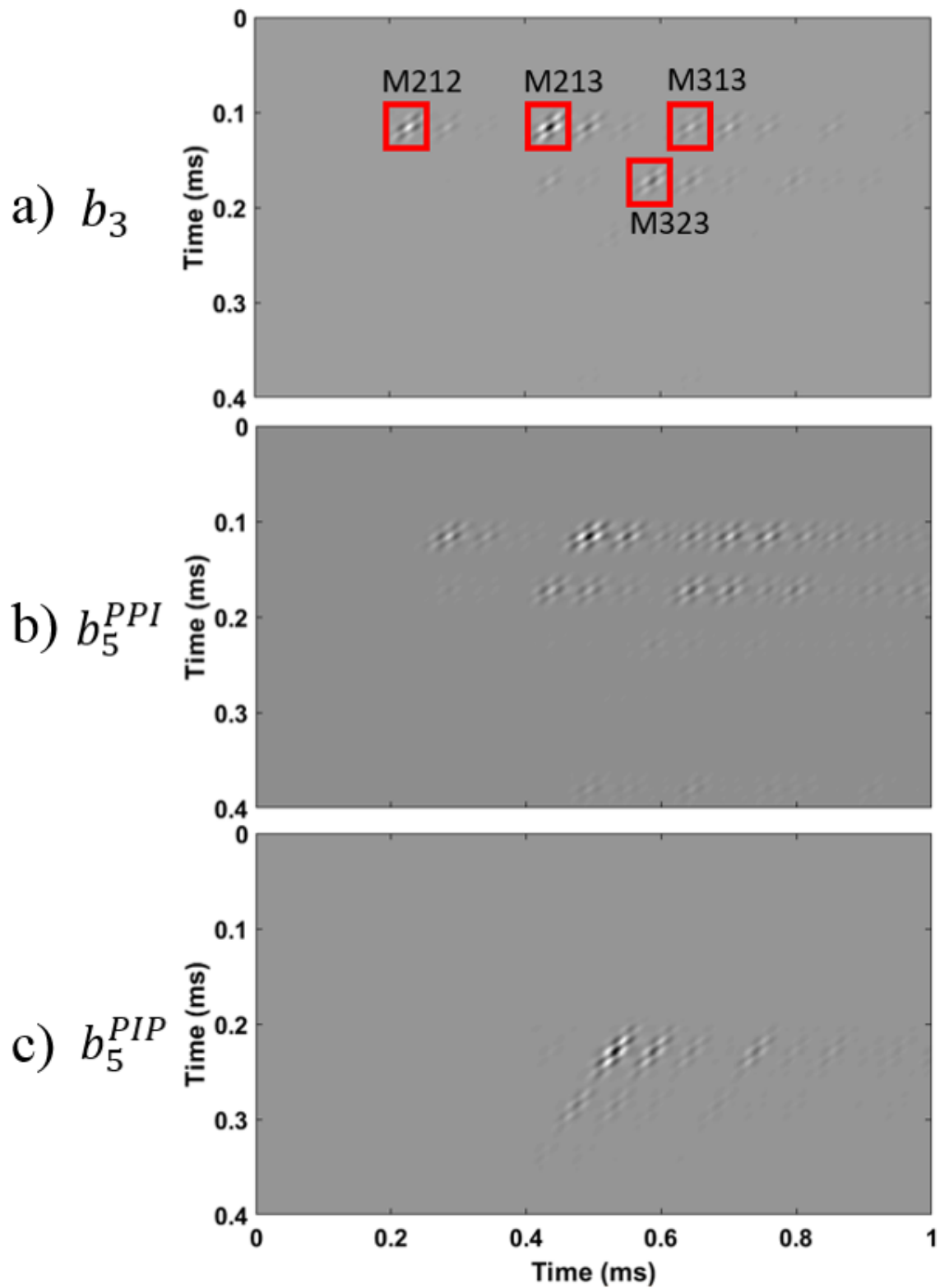


Figure 3.3 2D downward generator space displaying individual internal multiples, a) b_3 term with first order multiples displayed with red box. b) b_5^{PPI} higher order multiples. c) b_5^{PIP} artifacts from the prediction

The b_3 space is shown with the four first order multiples outlined. It shows that the first order internal multiples can be separated from both the higher order multiples and the events that are purely artifacts in the prediction. Recalling that the final 1D trace is created by summing over all the rows, this highlights how numerous internal multiples can contribute at a single point in time. There is also the option to use all three spaces to apply various scaling as required. If there were any problems with scaling of the initial prediction of b_3 prior to being put into higher order algorithm, then this could still be adjusted in this space.

These spaces are summed (Equation 3.4) to create one 2D downward generator space to be used for the prediction (Figure 3.4). By summing all these terms in the 2D space the result is a downward generator space which has been corrected for the overpredictions of higher order terms and pure artifacts have been removed.

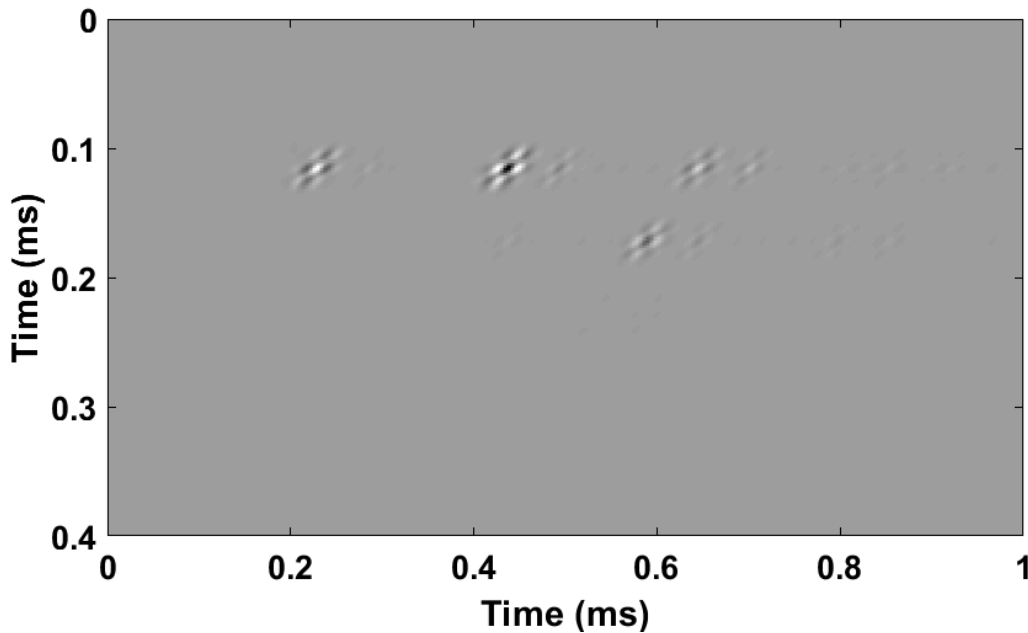


Figure 3.4 2D downward generator space displaying individual internal multiples after including higher order terms

This 2D space can now be used as previously shown with either the scalar or the 2D adaptive subtraction to obtain a final prediction. By using the 2D adaptive subtraction the result is displayed (Figure 3.5).

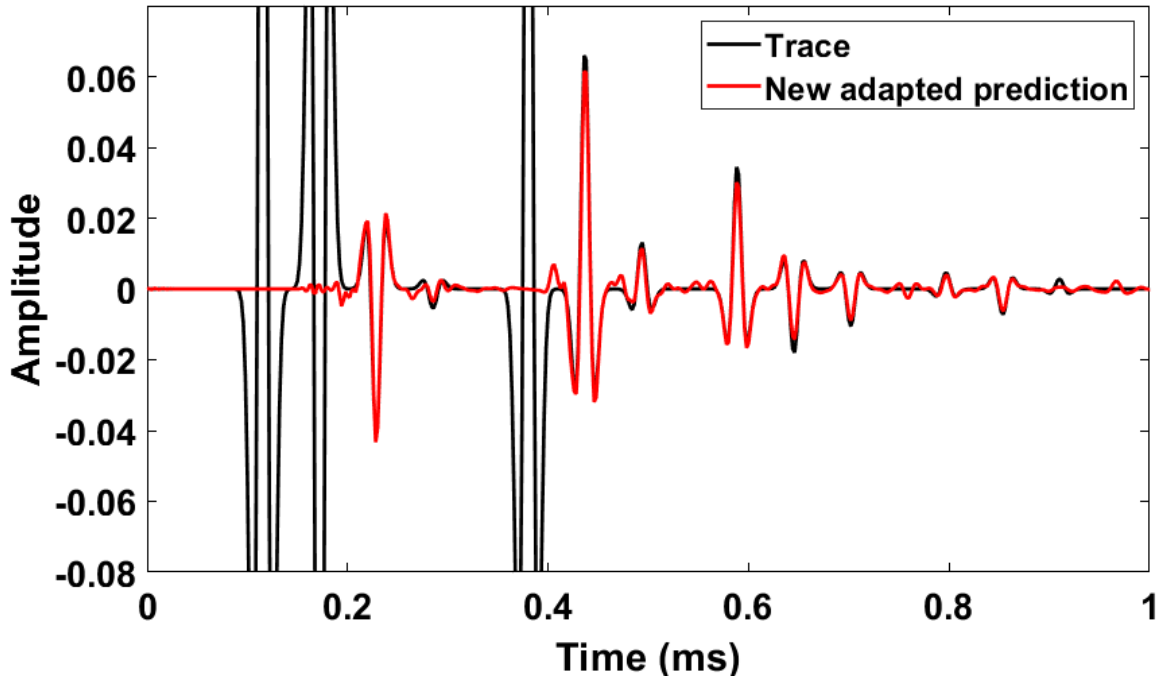


Figure 3.5 Internal multiple prediction with higher order terms and 2D adaptive subtraction

By combining both higher order terms and the 2D space a significant improvement in the prediction is made. Previously to correct these amplitudes a significantly non-stationary or harsh adaptive subtraction would be required. Through the combination of both the 2D space and higher order terms these multiples can now be more accurately attenuated. Ideally the method would not require adaptive subtraction and the amplitudes of all internal multiples would be predicted exactly from the original algorithm. In practice what these alterations have done through the DGS and the higher order terms is given the adaptive subtraction an improved

opportunity to be successful as it is attempting to give both correct the amplitudes first and give a more natural space to calculate the filters.

3.5 Numerical example in 1D with a destructively interfered primary

To display the uplift of the additional terms and why the adjustments to the predicted amplitudes is required, a subtle variant of the 1D example is displayed. This is done by adding an additional layer to the previous model with the new model displayed (Figure 3.6).

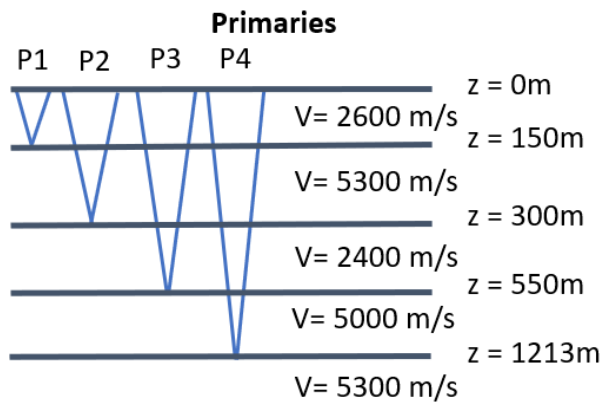


Figure 3.6 Velocity and depth model with additional layer used for 1D prediction.

The addition of the fourth layer was done to create a fourth primary which occurs at the same time with opposite polarity of an internal multiple (Figure 3.7). The result is an input trace which has minimal evidence of the fourth primary as it has destructively interfered with the multiple.

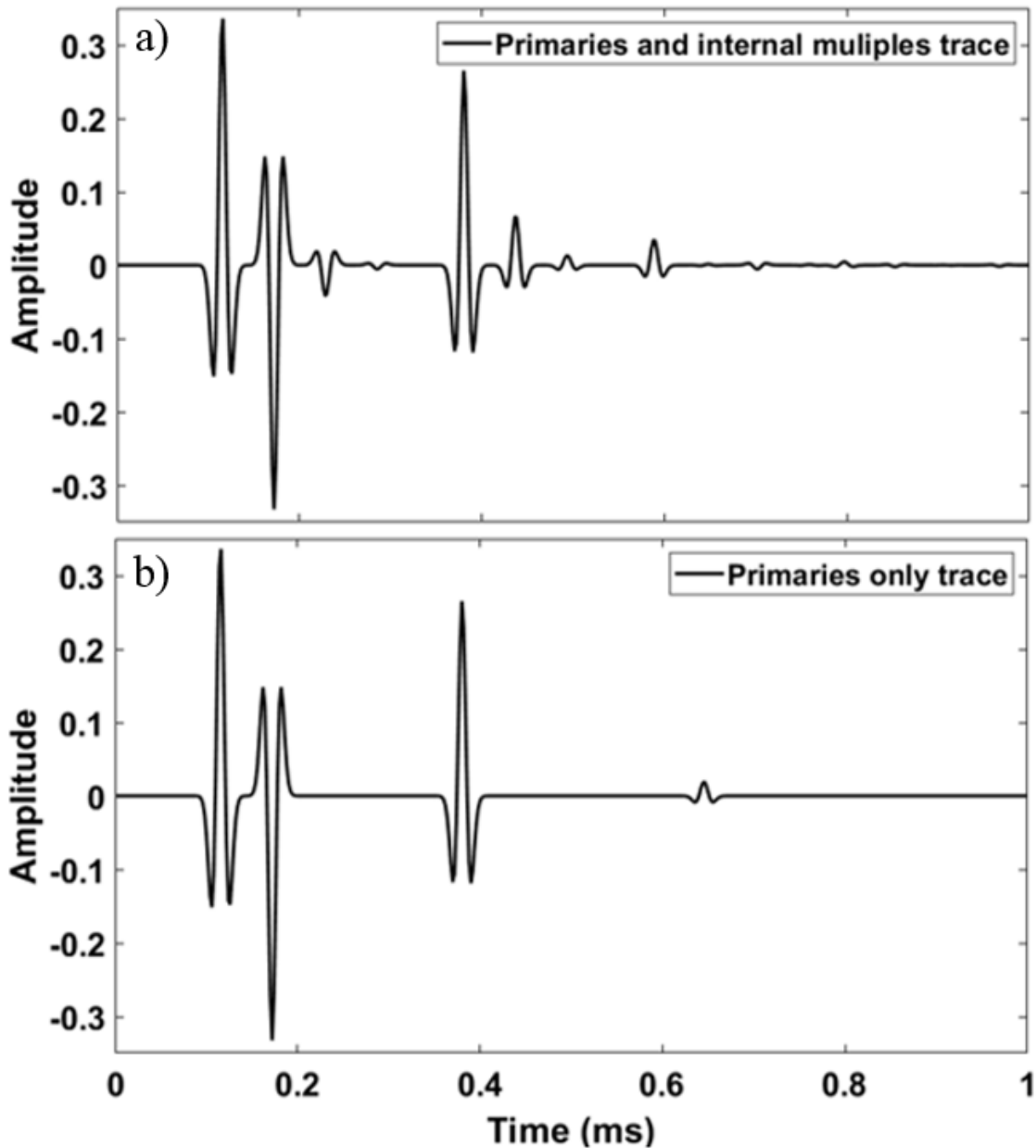


Figure 3.7 a) Primaries and multiples trace b) Primaries only trace highlighting location of missing primary due to internal multiples

For this example, the prediction is completed using the standard approach using the first term and the 1D adaptive subtraction and 2D adaptive subtraction in the downward generator space with the higher order terms. Both results are displayed (Figure 3.8).

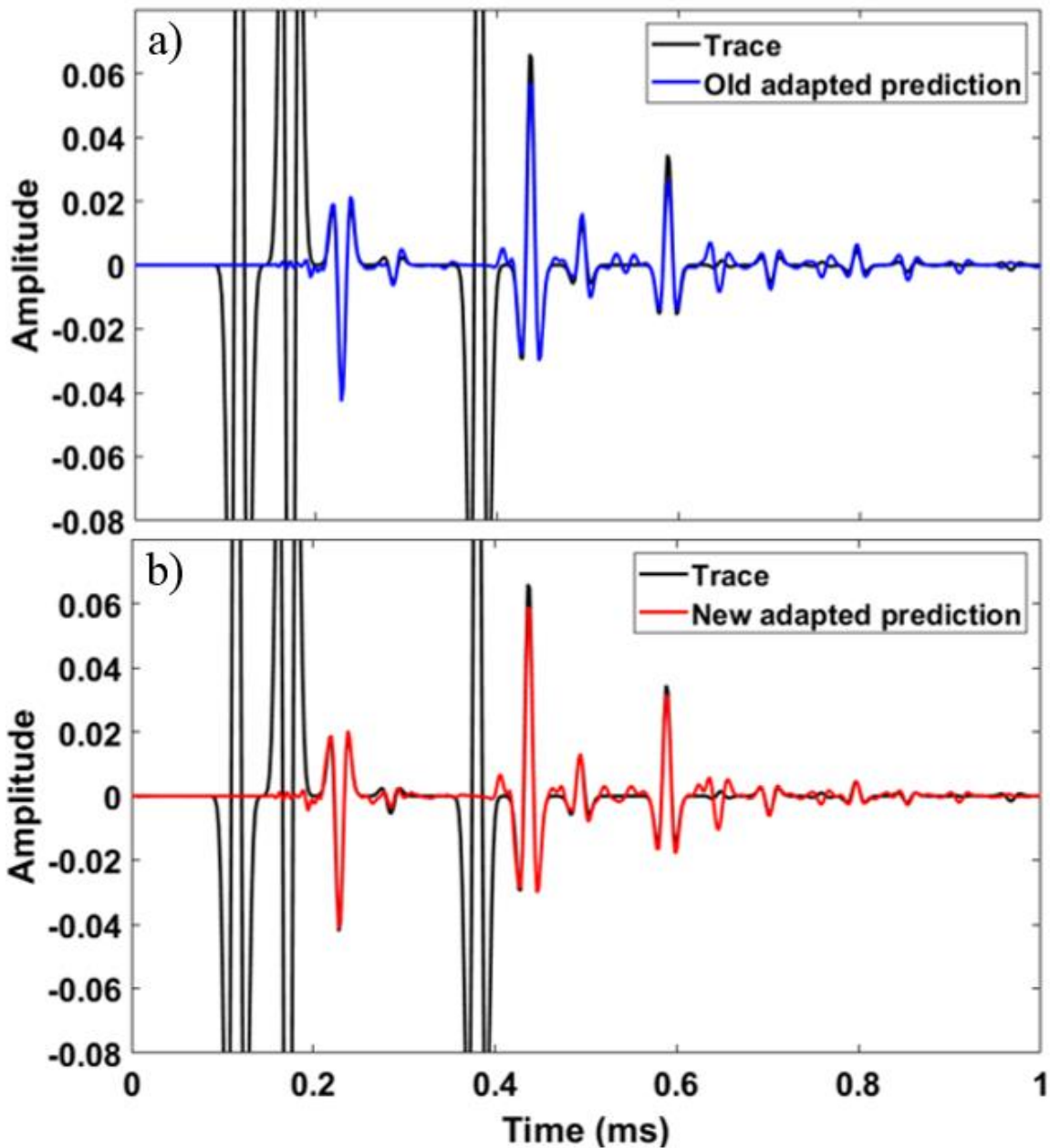


Figure 3.8 a) Internal multiple prediction with 1D adaptive subtraction b) Internal multiple prediction with 2D adaptive subtraction and higher order terms

As shown previously both can predict the internal multiples without predicting primaries. The amplitudes for both appear reasonable with a slightly improved match with the additional terms and 2D adaptive subtraction. The multiple attenuated trace is created for the first test with only the first term and 1D adaptive subtraction (Figure 3.9).

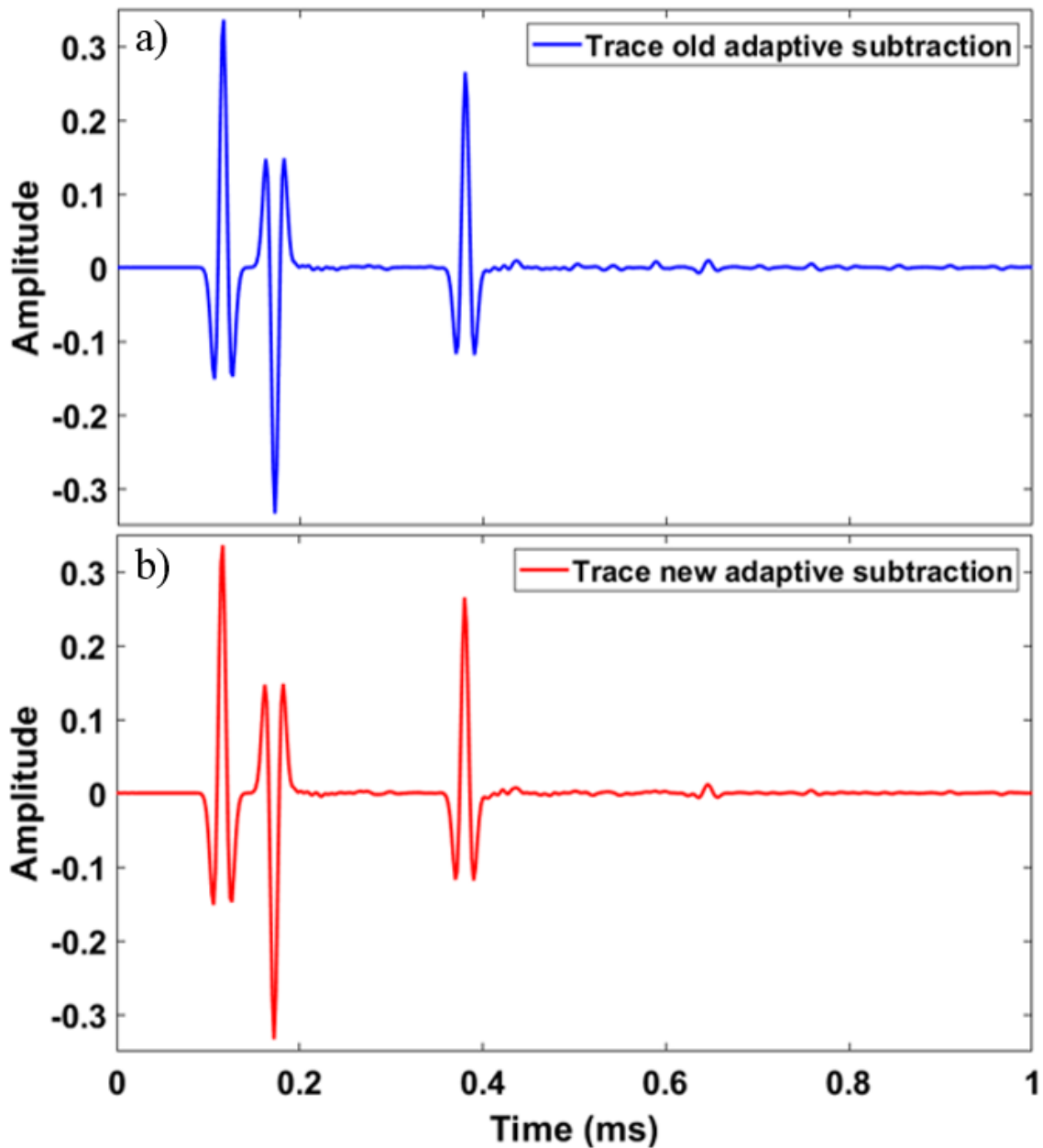


Figure 3.9 Trace with internal multiple attenuation using a) the 1D adaptive subtraction Trace with internal multiple attenuation and b) 2D adaptive subtraction and higher order terms

The original method result shows the three prominent primaries and several small amplitude events which make it difficult to distinguish if any of these are primaries. The multiple attenuated trace created with higher order terms and 2D adaptive subtraction has the

prominent three large primaries with the additional of the small fourth primary event now visible. This display the importance of correcting for these amplitude issues. Also giving increased confidence in the prediction to reduce potential harm to primary events.

3.6 Conclusions

It is displayed for a zero offset noise free synthetic how including the higher order terms and adaptive subtraction can accurately predict the internal multiples and their amplitudes. In practice with real data the higher order terms may not be as large an issue depending on the noise level of the recorded data. It is shown how the variation in predicted amplitudes between first order and higher order internal multiples can be detrimental to the predicted trace. Using higher order terms in the series the higher order multiples and artifacts can be isolated from the first order multiples. A critical step is how to scale the amplitudes prior to using the higher order terms. By including higher order terms in the series and 2D adaptive subtraction an accurate internal multiple prediction can be made which as displayed can minimize damage to primary amplitudes. A key question to be addressed is how practical this extra dimension will become with real data. Once various amplitude recovery techniques have been used, or if there are other factors such as attenuation or other losses across the downward generator will this extra dimension be able to assist.

Chapter Four: Offset-Time Domain Algorithm

4.1 Non-stationary seismic data

When a seismic wave propagates through a medium, reflections and transmissions result at an interface with varying properties. Another common outcome of a wave propagating through a medium are changes to the wavelet shape. Various mechanisms can create a non-stationary wavelet. One mechanism which will have this effect is attenuation which can be either intrinsic or extrinsic (Lv & Innanen, 2016). Attenuation will generally cause a loss of frequency on the high end of the amplitude spectrum and the wavelet becomes wider due to the emphasis of the low end of the amplitude spectrum. If intrinsic, this is an inherent property of the medium or if extrinsic the attenuation is caused by short path internal multiples (Lv & Innanen, 2016). If there is extrinsic attenuation due to short path multiples this can not be resolved by the ISS method. If a spatial dimension is used such as a shot gather the recording of a dipping wavefront can also cause an apparent broadening of the wavelet. There are cases where a stationary epsilon has been shown to be insufficient (Innanen & Pan, 2014; Innanen, 2017). When using the inverse scattering series internal multiple attenuation method, the epsilon term in the algorithm is used to account for the bandwidth of the data. If the bandwidth of the input data changes in space or time, then the algorithm must also be flexible to vary with the input data.

4.2 Non-stationary epsilon

The original algorithm uses a single stationary epsilon value for all space dimensions and time. There has been recent work to explore various calculation domains (Sun & Innanen, 2016). In some 1.5D domains epsilon can vary in the transformed spatial dimension such as

wavenumber (Innanen & Pan, 2014). These 1.5D transform domains allow for a variation of epsilon in the spatial direction (k_g) due to the method utilizing the 1D version of the algorithm over every spatial step. Thus, for every spatial step epsilon can vary as there is no communication between wavenumbers. This variable epsilon has been limited to the spatial dimensions because the time variable in the input data is not shared with output which is frequency. To allow for a variable epsilon the inverse scattering series equation for predicting internal multiples has been rewritten in the offset-time domain (Innanen, 2015b). Improvements in the 1D time prediction have been shown due to the nonstationary epsilon (Innanen, 2017).

4.3 Implementing the time domain algorithm

It was outlined in chapter one how the offset-time and time versions of the algorithms are derived from the original algorithm. Implementation of the algorithm is shown including differences between the time domain and transformed domain calculation. Beginning with the 1D time (Equation 1.29) and applying the integration limits directly to the data with the use of two Heaviside step functions (Innanen, 2015b).

$$B_3(t) = \int_{-\infty}^{\infty} dt' s_1(t' - t) \int_{-\infty}^{\infty} dt'' [O(t, t', t'') s_1(t' - t'')] s_1(t''), \quad (4.1)$$

Where the mask $O(t, t', t'')$, effectively applies the lower-higher-lower criteria directly to the data.

$$O(t, t', t'') = H[t'' - (t' - (t - \varepsilon))] H[t - \varepsilon - t''], \quad (4.2)$$

To calculate numerical convolution there are multiple ways to carry out the operation. It can be performed by careful tracking of indices for the set of multiplications and summations.

Convolution can also be carried out through the construction of a convolution matrix where the operation is then calculated through matrix multiplication.

4.3.1 Masking operators and pseudocode

Displayed in Figure 4.1 is the matrix multiplication approach of Equation 4.1 applying the convolution for the 1D time version of the algorithm (Innanen, 2015b).

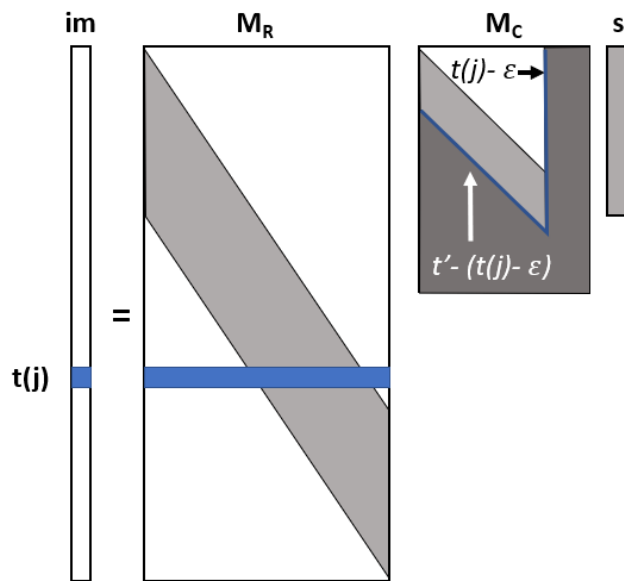


Figure 4.1, adapted from (Innanen, 2015), visual display of the calculation of convolutions and correlations for a given time $t(j)$ through matrix multiplication, with the mask matrix applied to the convolution matrix M_C .

The correlation matrix (M_R), convolution matrix (M_C) and the input data (s) after matrix multiplication gives the prediction (im). Note the mask matrix applied to the convolution matrix by the shaded region of the being zeroed from Equation 4.2. Figure 4.2 shows how the mask is applied to the convolution matrix.

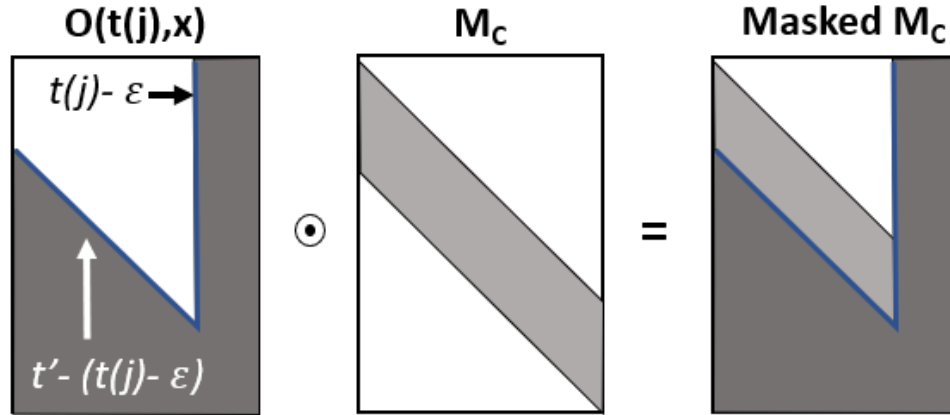


Figure 4.2, adapted from (Innanen, 2015), displays the mask matrix applied to the convolution matrix where the shaded region is set to zero and the bounds are determined such that the lower-higher-lower criteria is met for the given $t(j)$

The mask is a binary matrix applied to the convolution matrix where location of the change from zero to one ensure that the lower-higher-lower criteria is met. The mask also includes the epsilon term controlling the minimum event travel distance. Note that the internal multiple prediction is completed for each time step and the mask changes for each time step ensuring the lower-higher-lower criteria is obeyed for the given prediction time.

For the offset-time case, with the addition of the spatial dimension, the process of applying the matrix multiplication is like in 1D time. Now the equivalent of a 2D convolution in both offset and time is performed. The correlation and convolution operators become block matrices which compute both the spatial and temporal convolutions (Innanen, 2015b) (Figure 4.3). The block matrix is a larger 2D matrix which is built with smaller 2D matrixes within to calculate the convolutions and correlations for all offsets. The input trace ($s_1(x, t)$) is altered and stacked into a 1D vector for all offsets. This is also done for the prediction output vector (im). The masking matrix is applied in a similar manner to the 1D case but for every subset

block in the convolution block matrix. For the multidimensional case the prediction is calculated for a single time slice and across all offsets (Innanen, 2015b).

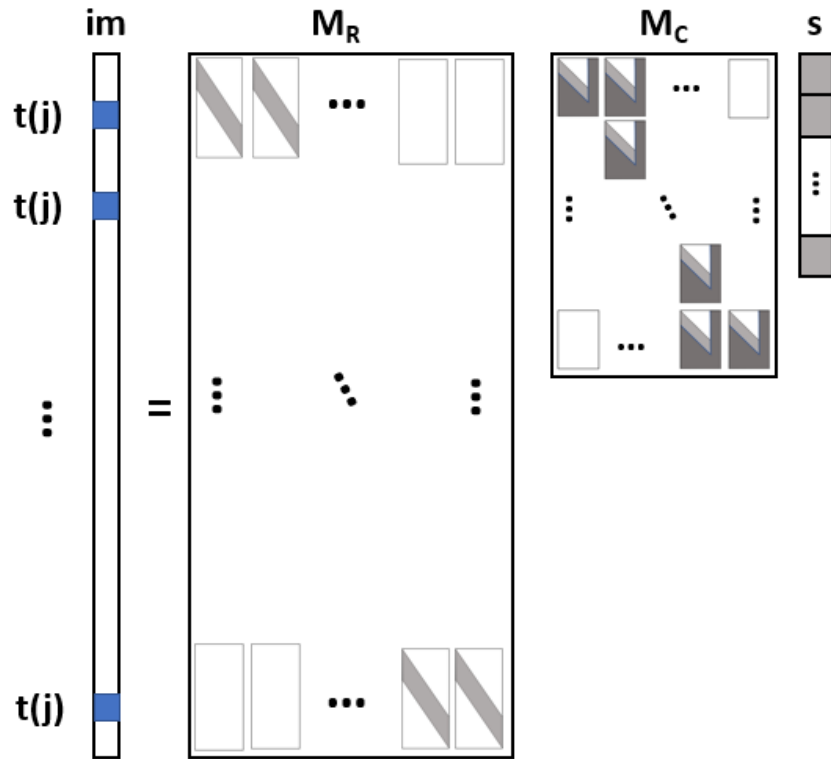


Figure 4.3 Display of the prediction algorithm for the (x, t) case for a given time $t(j)$ calculated for all offset with the convolution and correlation matrices, where the mask matrix is applied to each convolution matrix in the block matrix M_C . this is applied to the (x, t) trace (s) which has been stacked into a single column vector. Adapted from (Innanen, 2015).

With the addition of the spatial dimension relative to the 1D case these matrices become large, to the point where it becomes difficult to fully implement due to computer memory limitations. In practice these ideas of masking and convolution matrices can be equivalently completed with the use of a 2D convolution function (e.g. conv2 in MATLAB). This is likely completed using careful indexing to carry out the calculations. The Pseudo code of how this is implemented is displayed (Figure 4.4).

```

for it = tstart:tend {
    tindex = it-epsilon
    mask= 1
    mask(tindex:tend,x) = 0

    ms1= s1.*mask % apply mask to data b1
    cms1 = conv2(ms1,ms1) % 2D convolution with masked data
    b3(it,x) = conv2(cms1,s1R) % 2D convolution with cmb1 and time reversed data (s1R)
}

```

Figure 4.4 Pseudo code displaying the implementation of the (x, t) case with the use of 2D convolution functions for a stationary epsilon. The mask is applied in a similar manner where all values are set to zero given that they are either below the calculation time (it) and epsilon number of samples above (it).

Since the mask matrix is calculated for every time slice and the input and output domains are both offset and time, epsilon can be nonstationary in both dimensions. If epsilon is varied strictly in the time dimension, this is completed by having 1D epsilon schedule as a function of time for all offsets. If epsilon is to vary in both time and space, then a 2D epsilon schedule is built and the mask, which was previously displayed, will now vary with respect to offset. With the epsilon model matching the size of the input data. This is applied directly to the input data for each time step shown (Figure 4.5).

```

for it = tstart:tend {
  mask= 0
  for ix = xstart:xend {
    for iit = tstart:it-epsilon(ix,it) {
      mask(iit,ix) = 1
    }
  }
}

ms1= s1.*mask % apply mask to data b1
cms1 = conv2(ms1,ms1) % 2D convolution with masked data
b3(it,x) = conv2(cms1,s1R) % 2D convolution with cmb1 and time reversed data (s1R)
}

```

Figure 4.5 Pseudo code displaying the implementation of the (x, t) case with the use of 2D convolution functions for a nonstationary epsilon in both time and space dimensions

The epsilon schedules either varying in time or time and offset are displayed (Figure 4.6).

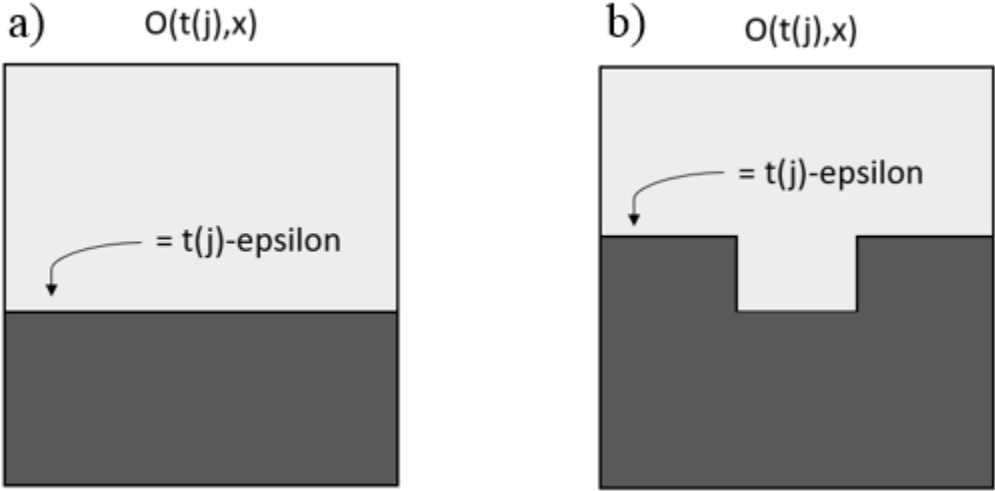


Figure 4.6 Visual display for the mask matrix applied to the data (s) prior to the 2D convolution function, where the bounds are determined such that the lower-higher-lower criteria is met for the given $t(j)$ for a) epsilon varying in time and constant with offset b) An example of epsilon varying in time and with offset with two epsilon values with a harsh cutoff

4.3.2 Differences between time and frequency domain

The underlying algorithm for the method in any of the domains is the same. By recasting the algorithm in different domains there have been variations in the prediction results. A result

of the varying domains to note at this stage is how the calculation is carried out numerically. In the previous transformed domains, the calculation was done one pseudo-depth at a time starting at zero and effectively predicting all internal multiples that would occur due to the given pseudo-depth and below with the result in the frequency domain. The time domain version is calculated differently as it is solved for a given output time. Effectively this means that, for a given time, it is searching for events which occur above this time that could have an internal multiple occurring at the output time. This results in either a top down approach for pseudo-depth or a bottom up approach for time. The way the lower-higher-lower criteria is implemented also varies. In the transformed domains it is done through varying the operating bounds of the algorithm. In the time domain the events which do not obey the criteria are effectively zeroed out. Another result of this, as I am solving for each time independently, is that if there is a given time interval of interest then this time could be isolated and solved for, without having to solve for all output times. If this was done for a frequency domain version of the algorithm, then all pseudo-depth above the desired one would have to be solved for.

4.4 Numerical example with 1.5D synthetic data

To evaluate the offset-time domain version of the algorithm a simple geologic model is used. This is the same geologic model and synthetic that was used previously (Figure 2.10). I create a shot record using finite difference modeling (`afd_shotrec` from the CREWES MATLAB toolbox) (Figure 2.11).

4.4.1 Internal multiple prediction

This geologic model has created a shot record which has a strong first order internal multiple. For the initial test of the time-offset version of the algorithm an epsilon value of 30 and 70 are tested (Figure 4.7).

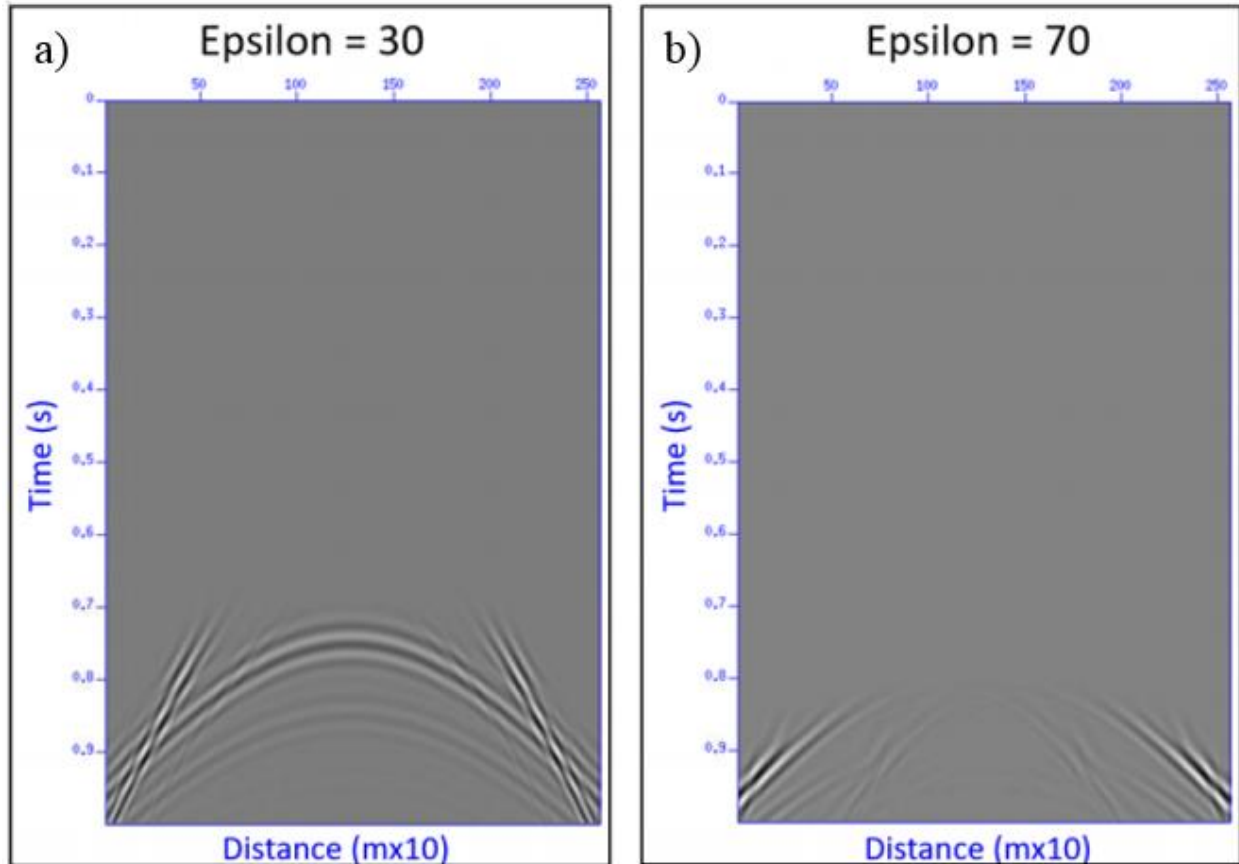


Figure 4.7 a) offset-time multiple prediction with epsilon=30, b) offset-time multiple prediction with epsilon=70

The epsilon of 30 has successfully predicted both the first order multiple and the higher order multiples. There is also a steeply dipping artifact which has been predicted. This issue has been shown previously in the (k_g, z) domain version of the algorithm (Innanen and Pan, 2014). The artifact was shown to be caused by the steeply dipping reflection from the input data becoming broad in the transformed spatial domain at larger k_g . Using an epsilon value of 70 has

diminished this artifact but epsilon has also become sufficiently large that it is not fully predicting the internal multiples, most notably at the near offsets. It is shown in the (k_g, z) domain how a k_g varying epsilon removed the artifact. This spatially varying epsilon is tested in the time-offset domain by combining the two epsilon values of 30 and 70 (Figure 4.8).

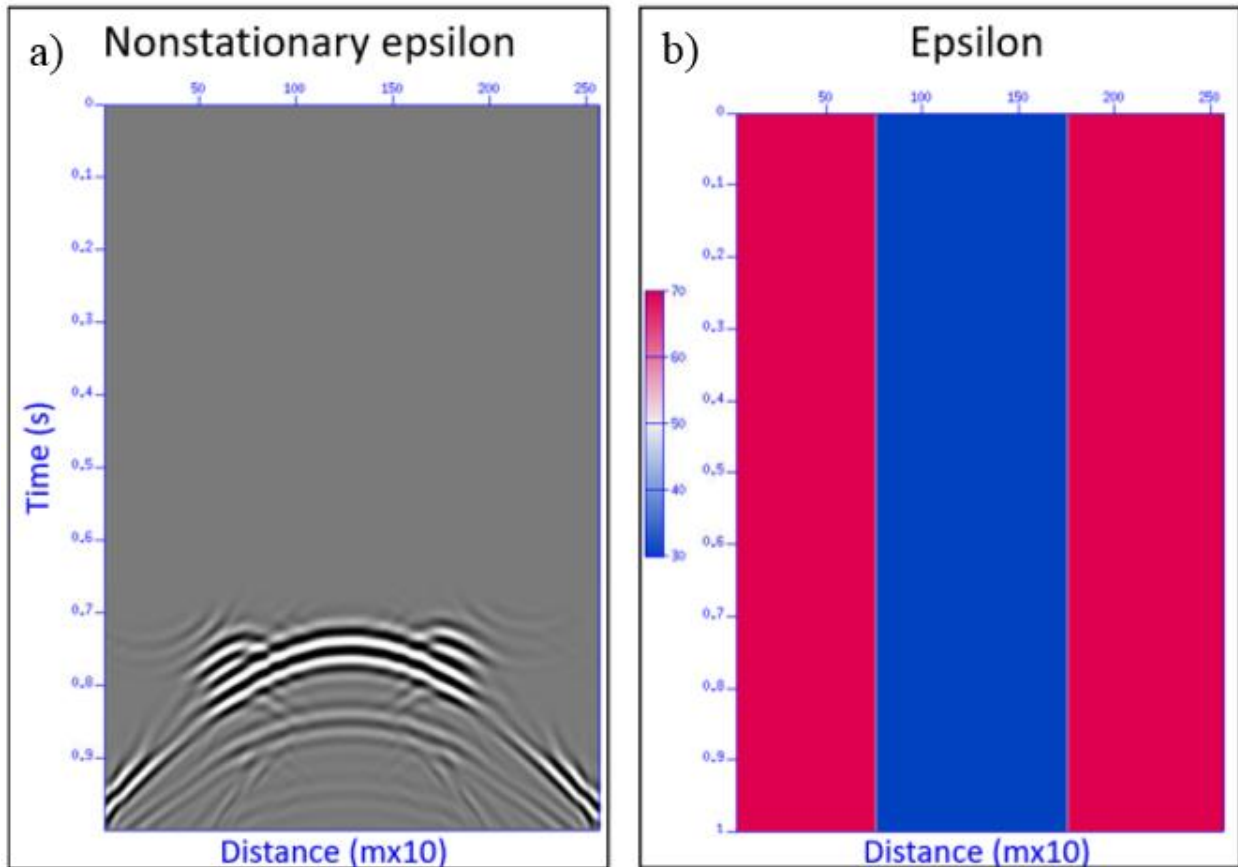


Figure 4.8 a) offset-time multiple prediction with spatially varying epsilon, b) epsilon schedule used for prediction with harsh cutoff

The internal multiple prediction has improved, but a new prediction artifact has been introduced. This 2D epsilon schedule was created with a sharp transition from the value of 30 to 70. This has assisted in removing the steeply dipping artifact but has appears to have introduced a diffraction off the edge of epsilon variation predicted on both sides of the first order internal

multiple. To mitigate this an epsilon schedule with a smooth linear trend from the 30 to 70 is implemented (Figure 4.9).

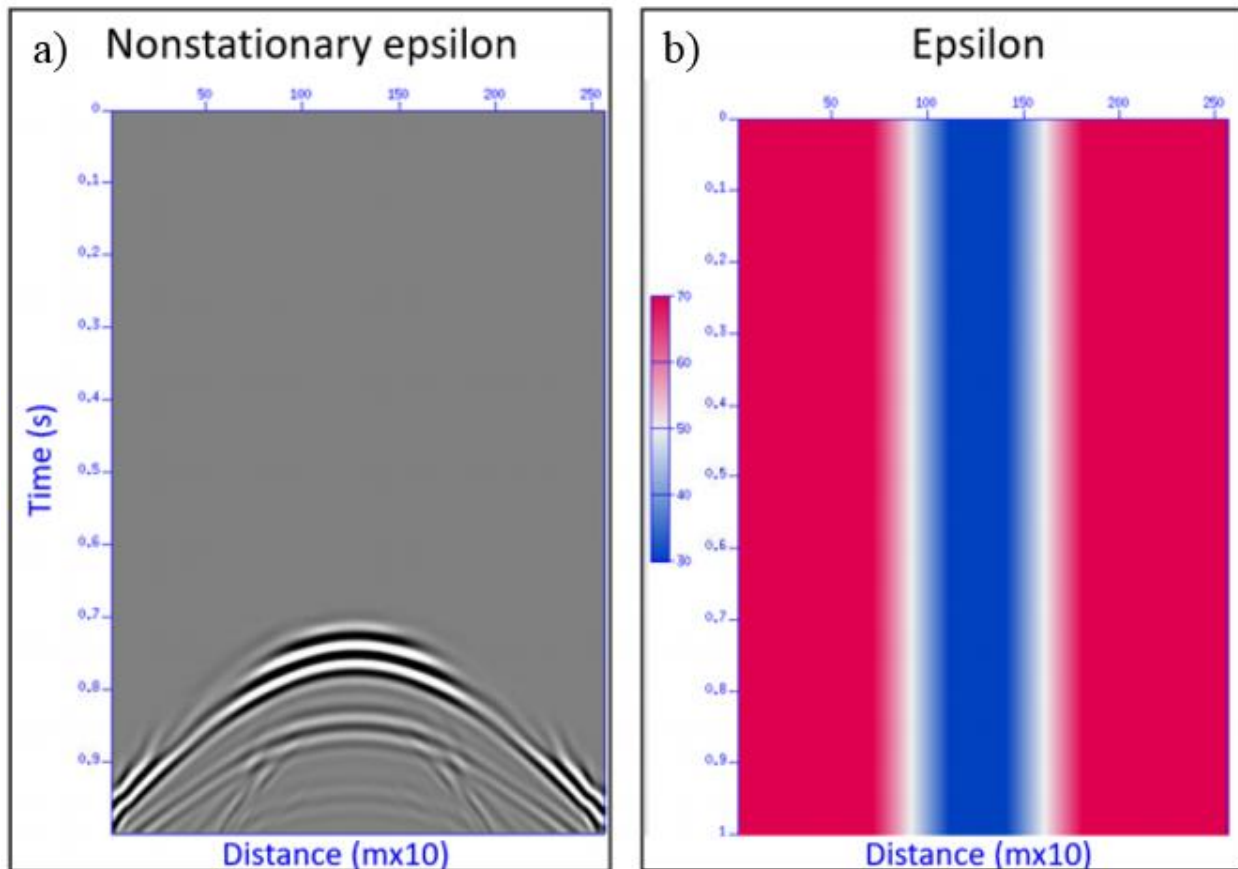


Figure 4.9 a) offset-time multiple prediction with spatially varying epsilon, b) epsilon schedule used for prediction with linear taper

This has improved the prediction of the internal multiples relative to the previously abrupt transition. The strong first order multiple along with the higher order multiples are predicted but with some artifacts remaining. Next an epsilon schedule which varies in both space and time is tested. This was designed to overlap the steeply dipping event in the input data and taper to minimize any diffraction artifacts (Figure 4.10).

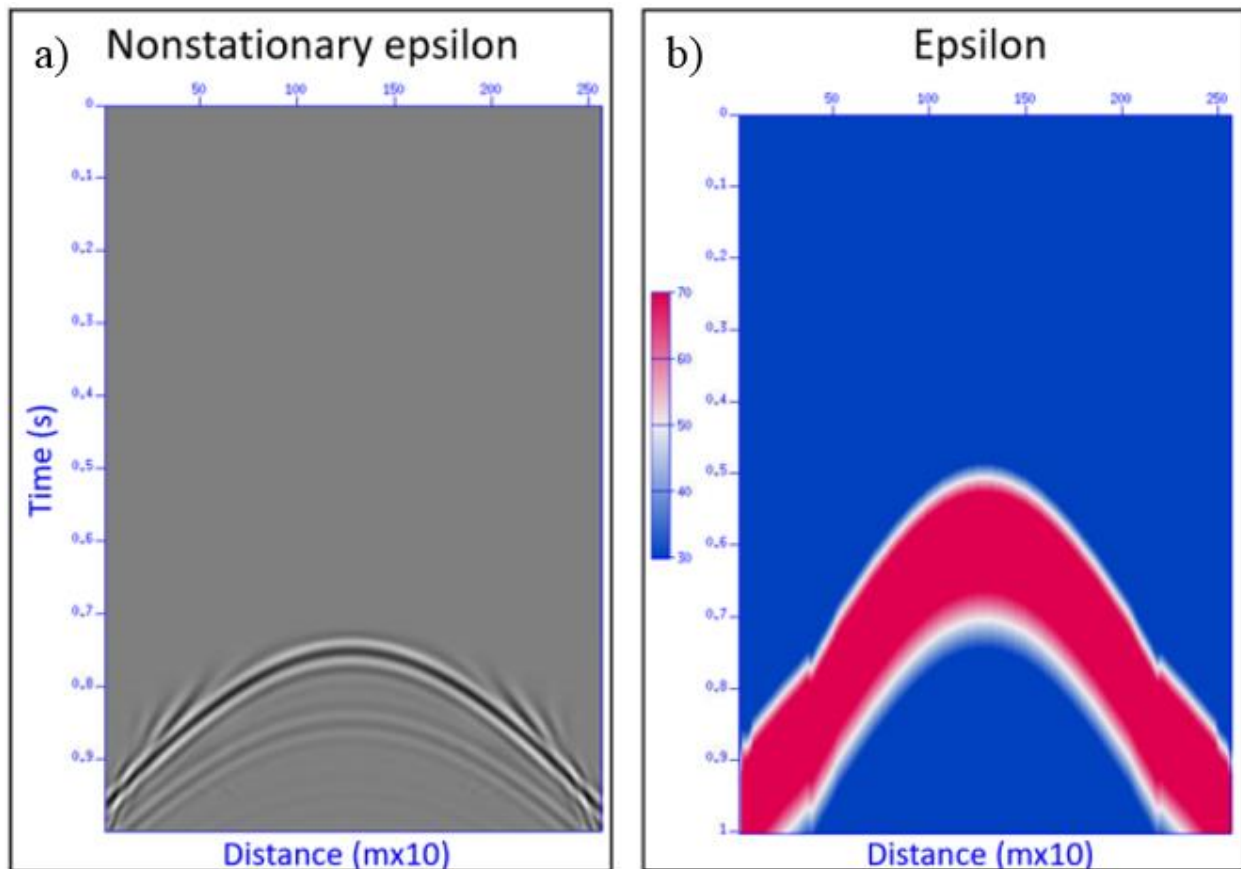


Figure 4.10 (Left) offset-time multiple prediction with spatially varying epsilon (Right) epsilon schedule used for prediction varying in both offset and time

This final schedule which varies in both offset and time is capable of both predicting the multiples in the data and minimizing the artifacts. Though there is still some residual artifact present from the steeply dipping event.

4.5 Numerical example with 1.5D synthetic data and irregular spatial sampling

All previous examples shown have had regular sampling in the spatial dimension. In practice, when recording 2D or 3D seismic data, the resulting spatial sampling will be irregular with offset. The transformed domains will handle irregular sampling differently than the time-offset version. Transforms such as tau-p can be setup to handle and transform data with irregular

sampling. Since the time offset version is all completed in the same domain, the irregular sampling will have to be handled in a different way, and how that sampling impacts the prediction is shown.

4.5.1 Random sampling numerical test

The first model to be tested is a two layer plus half-space (Figure 4.11). The two primaries and first order internal multiple were calculated using ray tracing and convolved with a 25 Hz minimum phase wavelet with 8ms sample rate. For this model the amplitudes for the two primaries and internal multiple are initially set to one for all offsets followed by the application of a spatial cosine taper.

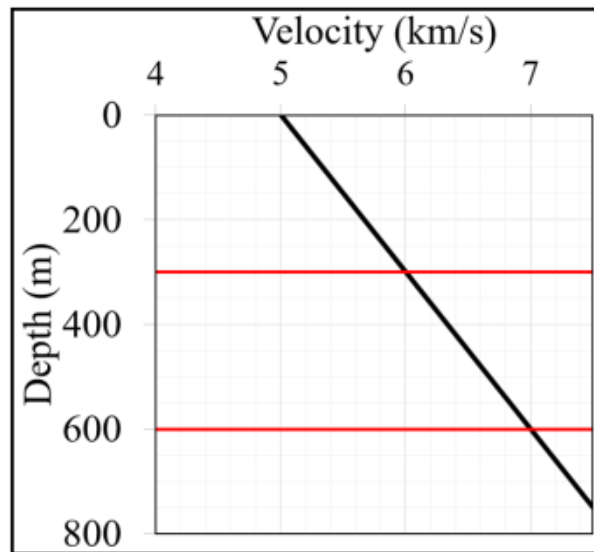


Figure 4.11 Velocity model with interface depths indicated by red lines

The first test is with the original dataset that has 256 offset traces and 64 temporal samples. This is displayed for both synthetic and prediction with the time-offset algorithm with two random decimation tests (Figure 4.12).

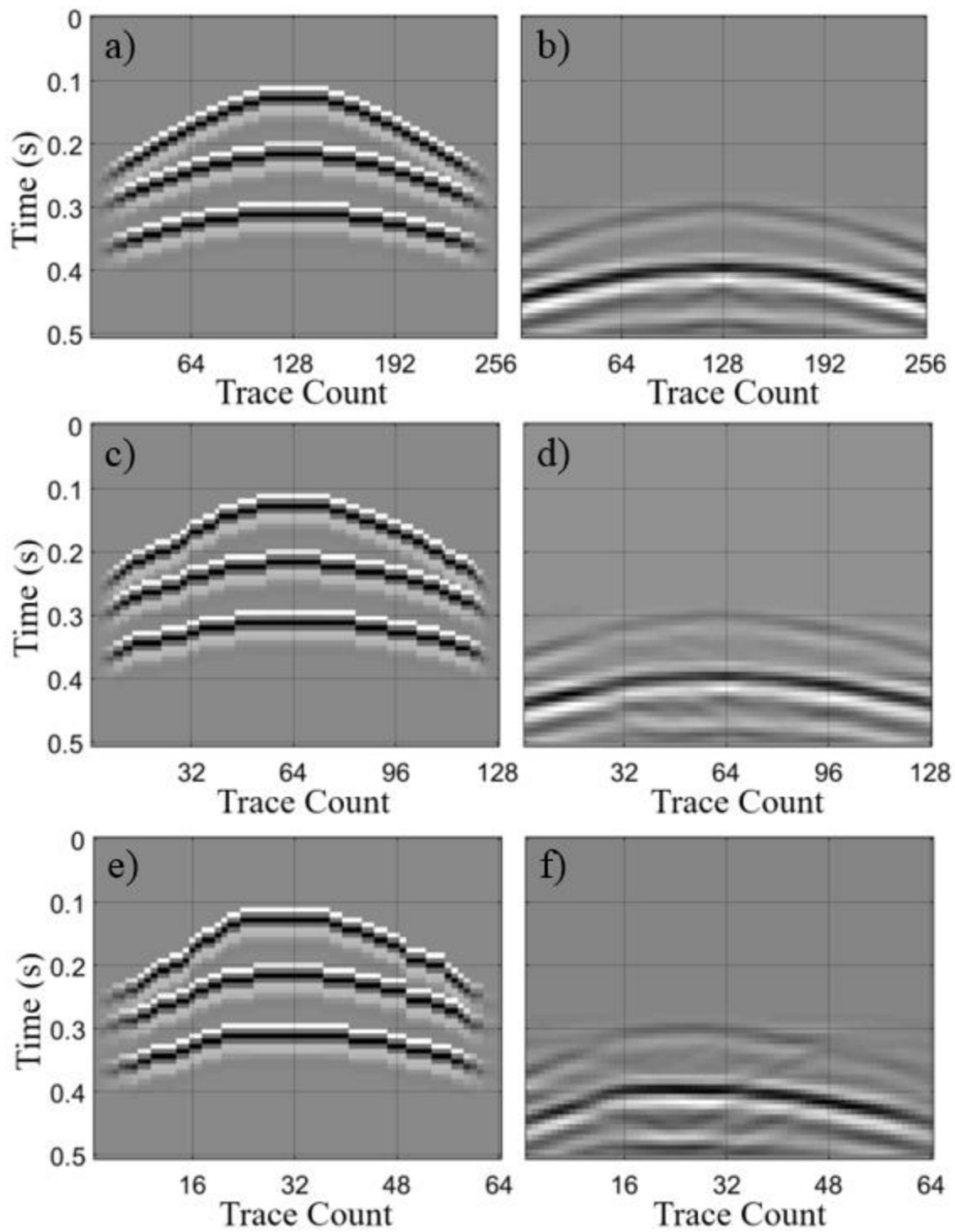


Figure 4.12 Synthetic data with total trace count of a) 256, c) 128 and e) 64. Internal multiple prediction for the given input data with trace count b) 256, d) 128 and f) 64

The algorithm has predicted the first order internal multiple without artifacts. There is an additional deeper event in the prediction that was not in the input data. This extra predicted event is the second order multiple from the two interfaces, which was not initially modeled with the raytracing. Note the amplitude of this higher order multiple is larger due to the previously noted feature of the algorithm overpredicting higher order events. In this first decimation test the method is observed to produce a reasonable internal multiple prediction. With the variations in input data the internal multiple can approximate the character of the decimated multiple. With the more significant decimation and irregular spatial sampling the method is struggling to produce the prediction. The result is poor relative to the final decimation test with 64 traces. The next test on this model is predicting all of the positive offsets from the original data with regular offset sampling (Figure 4.13).

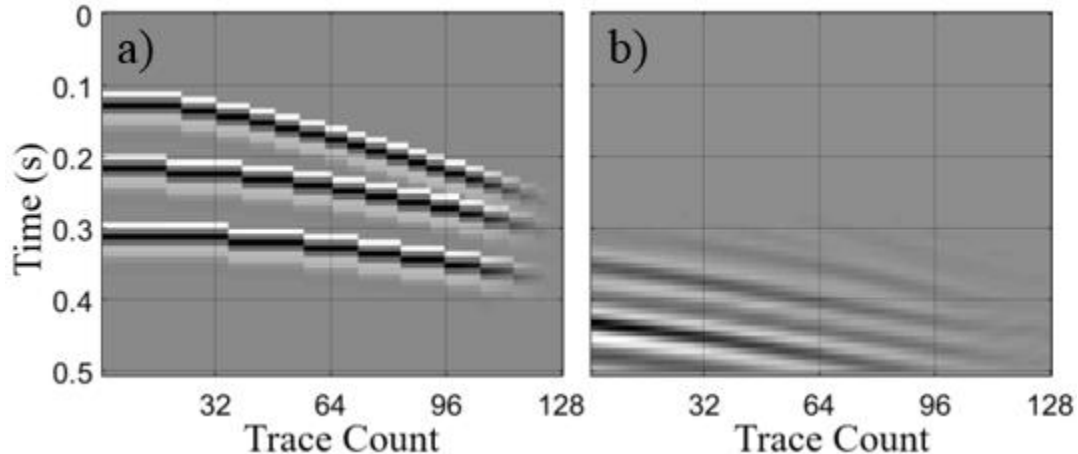


Figure 4.13 a) Positive offset synthetic, b) Internal multiple prediction

With only the positive offsets, the algorithm in the time-offset domain is unable to predict the internal multiples. The intercept time is incorrect, and the slope appears linear. The 1.5D time-offset version appears to require a split spread input and struggles with only using half of a shot record.

4.5.2 Orthogonal 3D survey geometry numerical test

When recording 3D seismic surveys, the offsets will not be regular, but also not perfectly random. A common recording will contain a lack of near and far offset data and an abundance of mid offset traces. To test how the algorithm functions on a model which is recorded with a 3D survey geometry synthetic model is created. The geologic model parameters were chosen to give significant time separation between primaries and internal multiples. This is done to avoid any steeply dipping reflections and allow for the use of a single epsilon. A spatial cosine taper has also been applied to the input data. This model was designed to allow for the testing and understanding of the single variable of spatial sampling without confusing the results between sampling and time variant epsilon. The model, synthetic and time offset prediction are displayed (Figure 4.14).

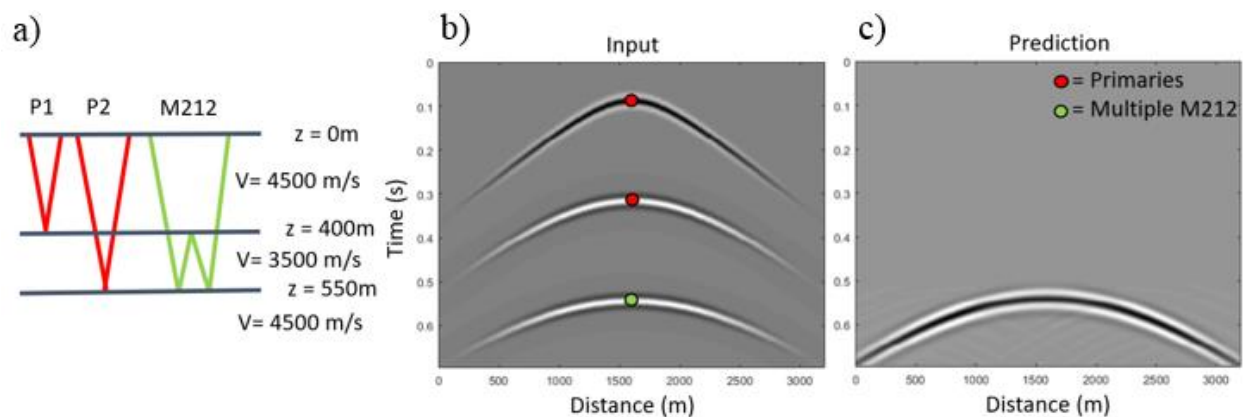


Figure 4.14 a) Velocity model, b) synthetic, c) time offset prediction

As intended by the model design, this single internal multiple is predicted without artifacts and is significantly separated from the primary events. The 3D geometry has been designed as orthogonal geometry, with 100m source and receiver line spacing and 20m source

and receiver spacing. This 3D geometry modeling was completed with PSDesign (Lawton., D. 2014). A single common mid point (cmp) location was chosen within this 3D geometry to be tested. The offset azimuth and reflection rays are displayed (Figure 4.15).

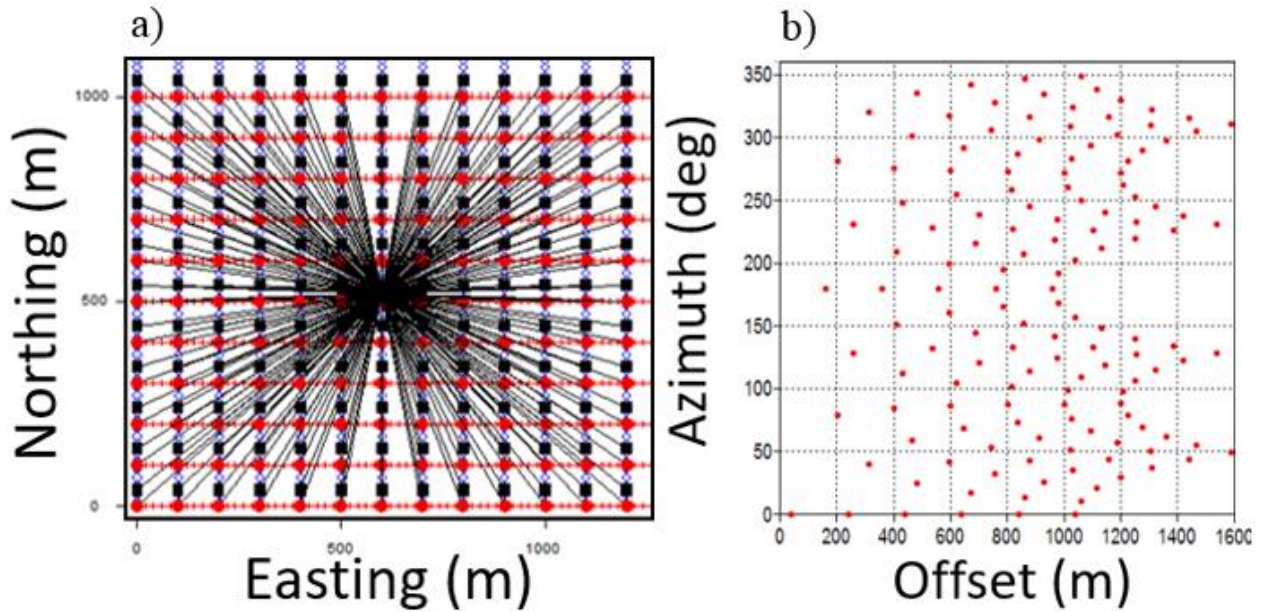


Figure 4.15 a) Source (red) and receiver (blue) geometry with rays for a single cdp, b) Offset and azimuth distribution for the displayed cdp

To create the cmp gather from the original trace a subset of offsets from the full regularly sampled trace were taken. Since the algorithm assumes a $v(z)$ medium, all azimuths were assumed to be equal when creating the trace. The results of taking this 3D trace and applying the time-offset prediction is displayed (Figure 4.16).

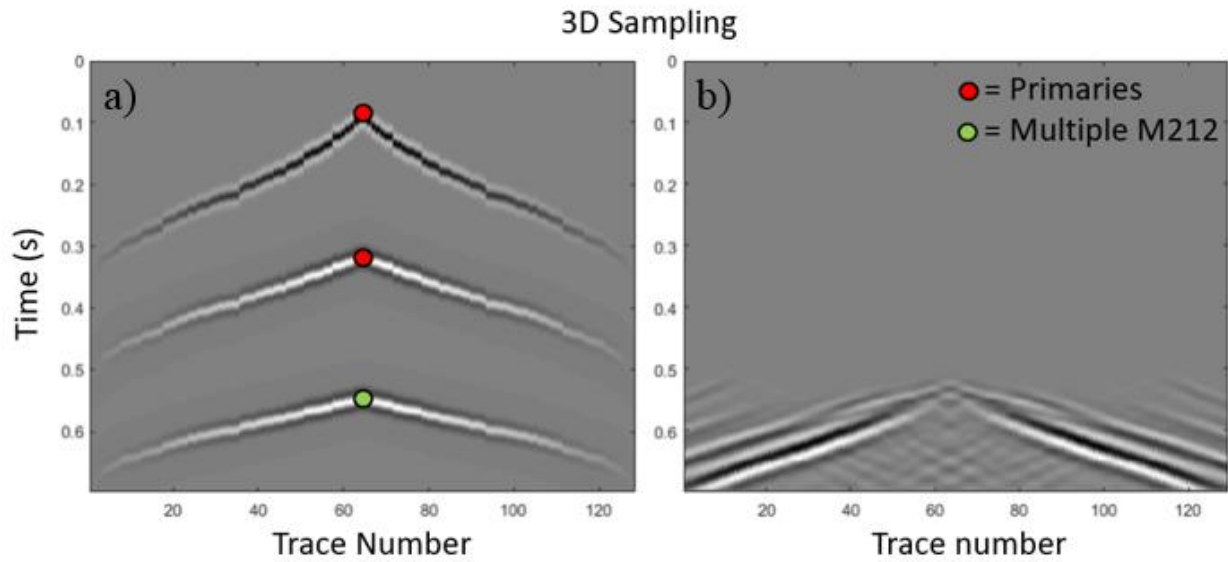


Figure 4.16 a) Input seismic trace from 3D seismic geometry, b) Internal multiple prediction

The resulting prediction has been unsuccessful at predicting the internal multiple. This appears to be due to the reflections from the irregular geometry being unrealistic. To create a slightly more realistic gather blank traces are inserted to give an apparent regular sampling (Figure 4.17).

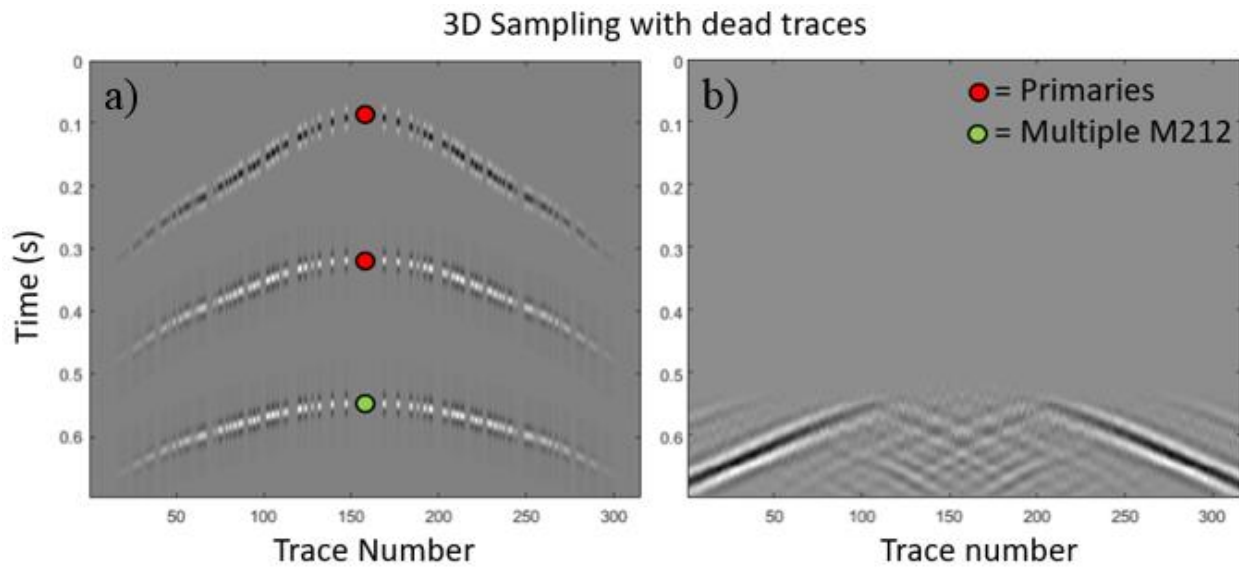


Figure 4.17 a) Input seismic trace from 3D seismic geometry with blank traces, b) Internal multiple prediction

By creating a more regular dataset this has significantly improved the prediction result. Though the prediction is struggling with the near offsets, the mid and far offsets appear to have been reasonably predicted. Note the method has predicted internal multiples at offsets where there was no data initially. The predominant artifact appears to be numerous diffractions at the locations of the missing traces. Like the previous test, as this is a wave equation-based method with minimal assumptions, it treats the input data as if this is what has been recorded on the surface, leading to the issues observed. The next test will take the geometry with the blanked traces but use a spline interpolation to fill in the blank traces. The result will be an approximately regular 2D line using a combination of the recorded traces and interpolated traces. The result and prediction are displayed (Figure 4.18).

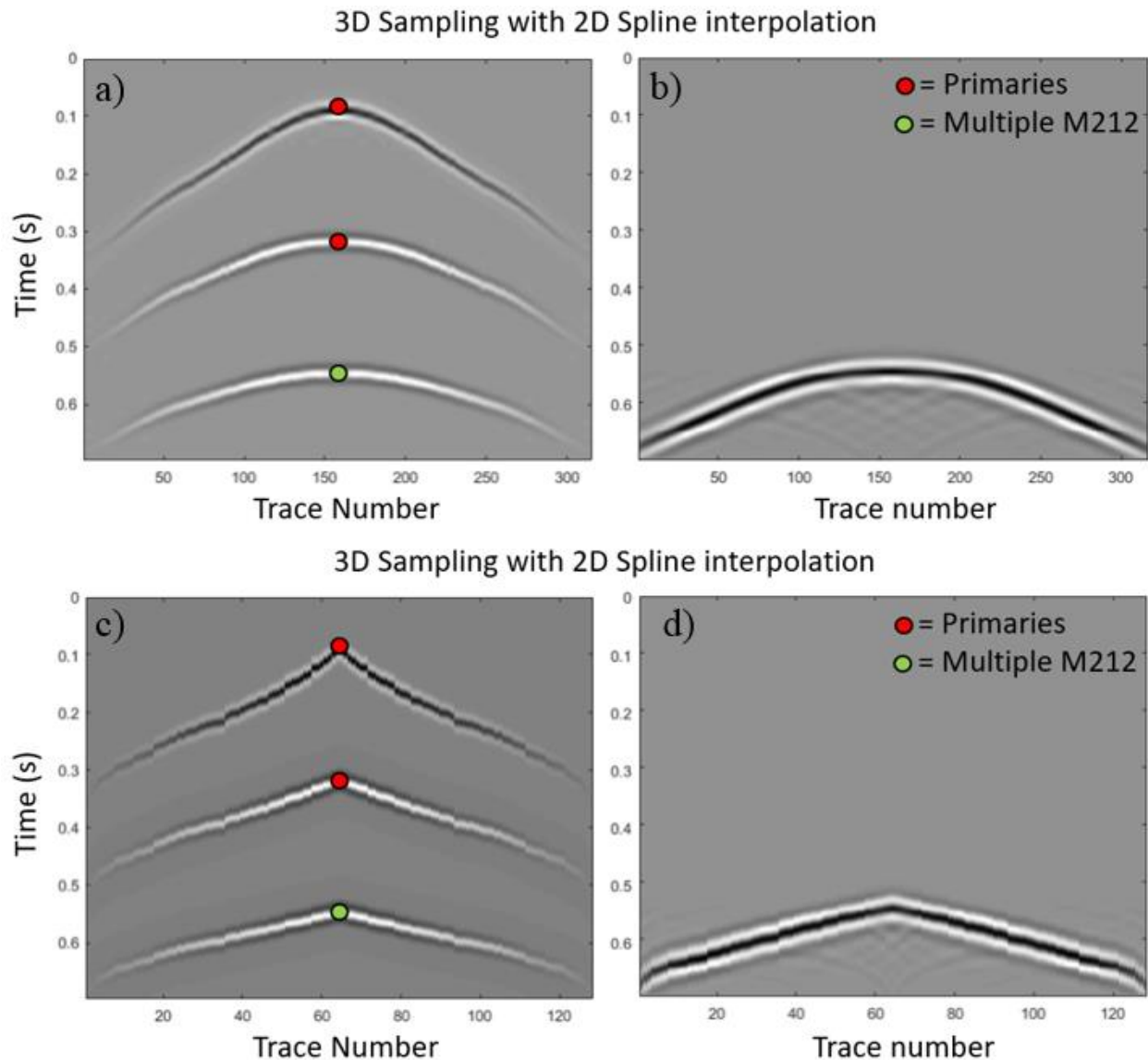


Figure 4.18 a) Input seismic trace from 3D seismic geometry with 2D Spline interpolation, b) Internal multiple prediction, c) Input seismic trace from 3D seismic geometry, d) Internal multiple prediction with the original traces displayed

The interpolation has significantly helped the prediction and again a reasonable internal multiple prediction is found. The only negative aspect to this approach is the increased computational cost with the higher trace count due to interpolation. The data decimated back to include only the original traces has given a reasonable prediction with minimal artifacts.

4.6 Conclusions

The method using the inverse scattering series for internal multiple prediction has been adapted to compute in offset-time domain (Innanen, 2015b). With this change in domain comes the ability to utilize a non-stationary epsilon. Displayed was an example of how varying epsilon can improve the prediction. This is completed on a model where a stationary epsilon will be insufficient. Also displayed is the how a sharp epsilon boundary in the epsilon schedule can cause new artifacts to be present in the prediction. The time-offset internal multiple prediction algorithm was also tested on two synthetic models with irregular sampling. The two tests showed that the 1.5D time-offset internal multiple prediction algorithms can accommodate some amount of irregular sampling. Once the variation between samples becomes significant and the wavefront is not smoothly varying, issues with the prediction begin to arise. Uplift was found both by regularizing spatial sampling either by inserting blank traces or by interpolation. The downside of this approach is the increase in computational cost.

Chapter Five: Devon Synthetic Data

5.1 Donated dataset

A dataset which includes pre and post stack 3D seismic, a vertical seismic profile (VSP) and a well with various well logs (gamma ray, porosity, sonic, density, ...) was donated to CREWES. Numerous research goals have been identified for the dataset, which is being assessed with tools such as FWI (Cova, et al., 2018). The primary purpose for the dataset donation was due to the presence of significant internal multiples in the seismic data. The substantial work on internal multiples that has occurred at CREWES led to the donation of data to test the algorithms. The first step to evaluate the inverse scattering series internal multiple attenuation effectiveness will be on synthetic data created from the well logs. A subset of the well logs is displayed (Figure 5.1).

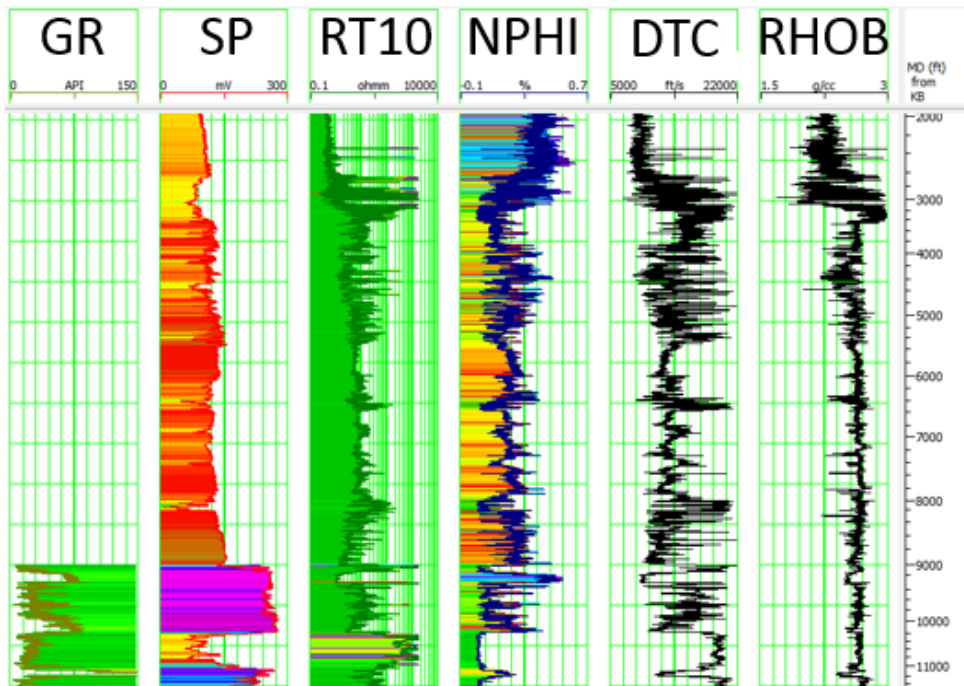


Figure 5.1 Well logs including gamma ray, spontaneous potential, resistivity, neutron porosity, sonic and density

The various well logs display rapid variability with depth. Both the sonic and density logs show variations in properties which cause significant impedance contrasts. These impedance contrasts could cause the noted issues with internal multiples. To this end, synthetic seismic is created using the well logs to both confirm the noted issue of internal multiples in the seismic data sets and to assess the applicability of the inverse scattering series internal multiple prediction algorithm to the dataset. Various domains, methods, issues and solutions have been outlined for the inverse scattering series internal multiple prediction. Here these methods are stressed beyond the simple models that were analyzed in the preceding chapters.

5.2 Vertical Seismic Profile

A vertical seismic profile (VSP) is acquired by placing geophones in a wellbore and initiating a seismic source on the surface. A VSP can have various source geometries including zero offset, walk away to obtain offset information or walk around to obtain azimuthal information. A critical advantage of VSP surveys is the ability to separate the acquired data into a primaries only and primaries plus multiples dataset. The setup for a zero offset VSP is displayed with commonly recorded wave types (Figure 5.2).

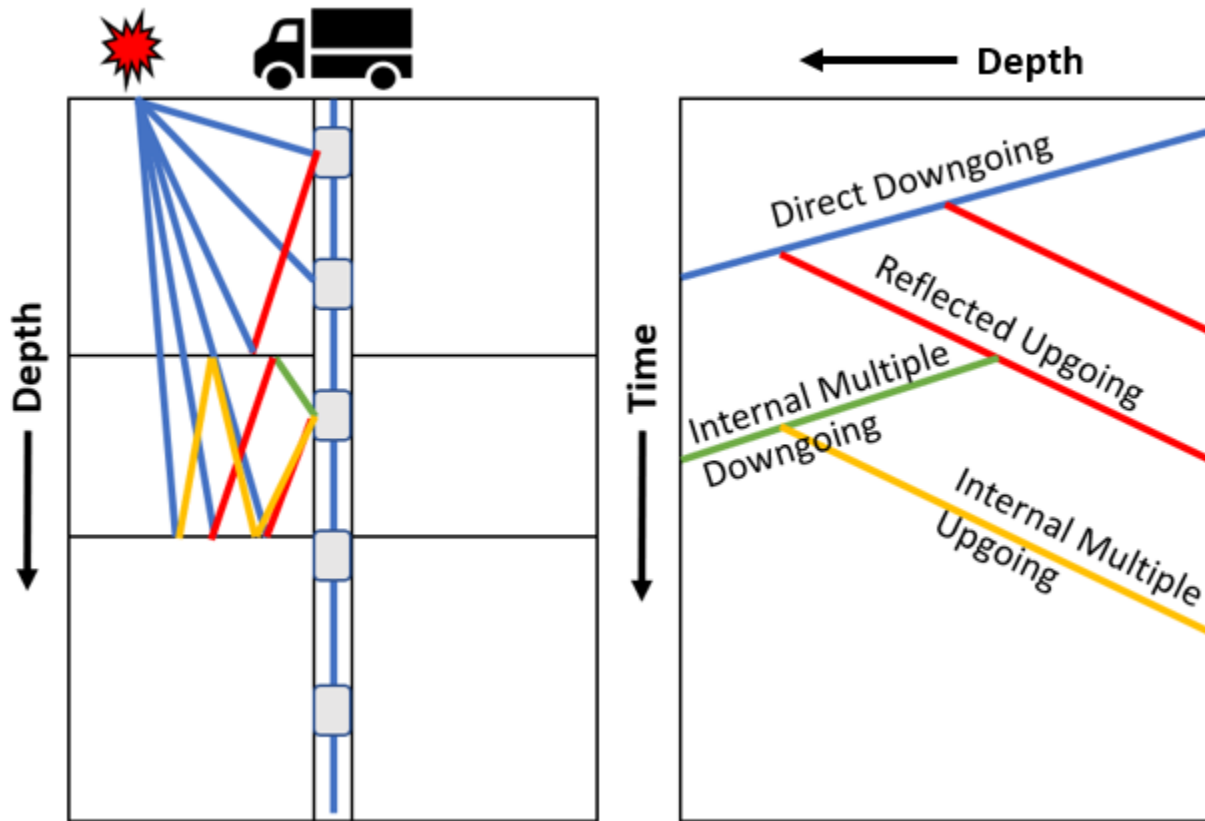


Figure 5.2 VSP with upgoing and down going energy and internal multiples

When the seismic source is initiated the direct downgoing wave is the first event recorded giving the first break. When an interface is crossed a reflection occurs and is recorded as an upgoing wave shortly after the direct downgoing wave. As the wavefronts continue, both primaries and multiples occur, and both the upgoing and downgoing wavefields are recorded. The result is a dataset which contains all the mentioned waveforms. Through processing steps to be discussed the upgoing and downgoing wavefields can be separated to give a dataset which contains only upgoing or downgoing waves. By stacking restricted portions of the wavefield separated VSP, primaries and multiples can also be separated. This is the benefit of a VSP and why it is an essential tool to evaluate seismic multiples.

5.3 Synthetic data modeling

The modelling utilizes the sonic and density logs sampled at 0.5ft intervals with an approximate length of 11000ft. The logs were not recorded to the surface, so a single replacement value is used for both the sonic and density to fill in the data from the surface to the top of the logged interval. Two sets of synthetic seismic data are generated to model the seismic from the VSP and the 3D seismic data. The two sets are required due to frequency content variations between the VSP and 3D seismic. The central frequency of the VSP is approximately 30 Hz, while for the PSTM data is 37 Hz. The workflow for synthetic VSP generation and multiple prediction analysis is shown for the 30Hz case, where only the results for the higher frequency synthetic data approximating the 3D seismic is shown.

5.3.1 Modeling algorithm

Layer propagator matrices, using the CREWES MATLAB toolbox (function `vspq`) are used to create the synthetic seismic VSP (Margrave & Daley, 2014). The method allows for fast calculations of a trace that includes all orders of internal multiples by creating propagator matrices as a function of the number of layers in the input geologic model. The propagator matrices method has the flexibility to create primaries only, primaries and internal multiples, surface multiples and various other combinations (Margrave & Daley, 2014). The VSP synthetic generated for this work only includes primaries and internal multiples, as the internal multiple prediction algorithm assumes these are the only events in the input trace. The function can also output the upgoing and downgoing wavefields directly so for the synthetic data the processing step of separating the wavefields is not required.

5.3.2 Synthetic VSP

The synthetic VSP was created using a zero phase 30 Hz Ricker wavelet, sampled at 0.002ms with a record length of 6 seconds and outputting displacement. The well data which was originally sampled at 0.5ft intervals was resampled to 5ft intervals and a geophone is placed at every layer. Though the well depth is only 11000ft with a total two-way travel time of approximately 2 seconds, the longer recording time is required to record the additional multiples. The resulting synthetic VSP with upgoing and downgoing waves is displayed and compared to the real VSP (Figure 5.3).

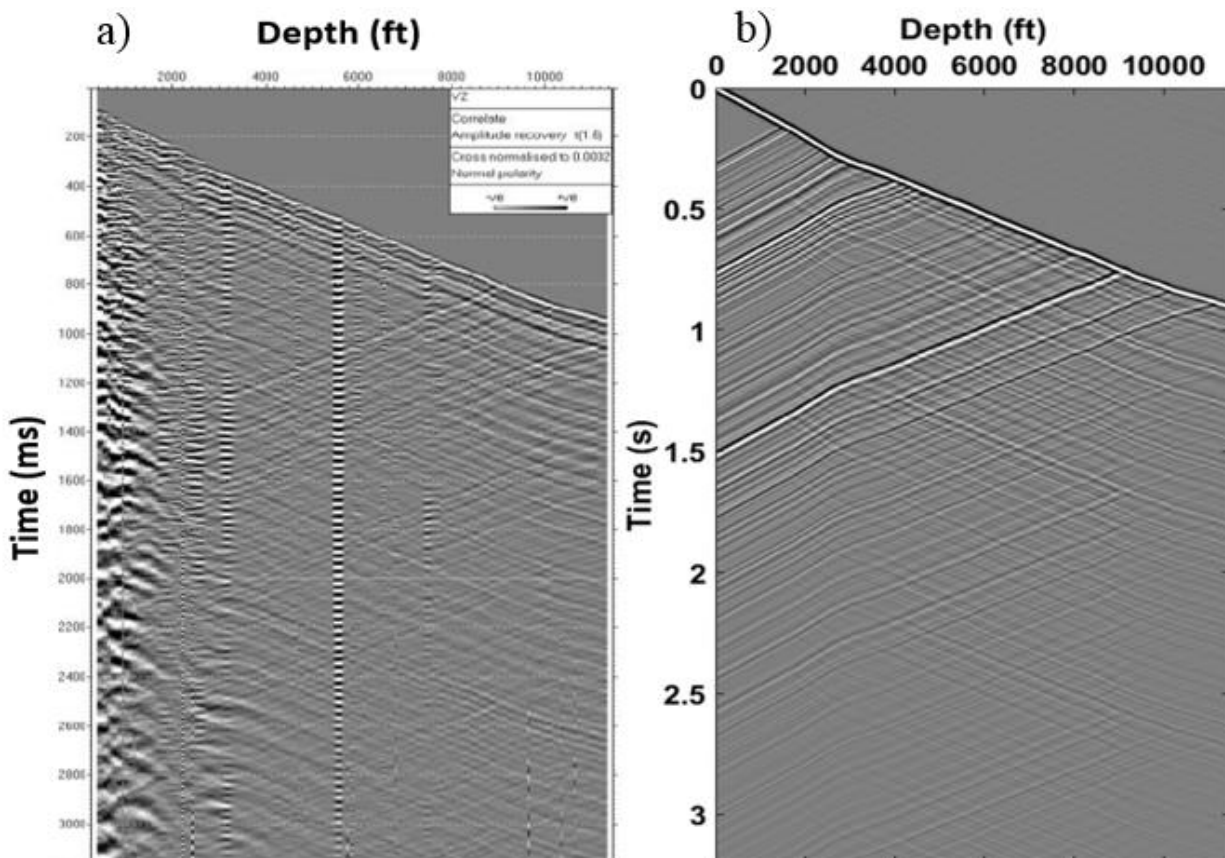


Figure 5.3 a) Recorded VSP with up and downgoing waves, b) Synthetic VSP with up and downgoing waves

The synthetic VSP shares similarities with the recorded VSP such as the slope of the first break and location of many of the reflections and upgoing waves. This method of modeling the data, input wavelet and blocking length chosen appears reasonable to give meaningful conclusions through the analysis of the data to later be applied to this real data. For the synthetic VSP the only processing step required is to flatten the data to create the corridor stacks. Since this is synthetic data and the velocities are known, a traveltime relationship can be calculated to flatten the data. When internal multiples are included in the synthetic modeling, the first break traveltimes no longer match with the calculated value. This is due to the short path internal multiples causing dispersion in the wavefront altering the arrival time of the direct wave (Figure 5.4).

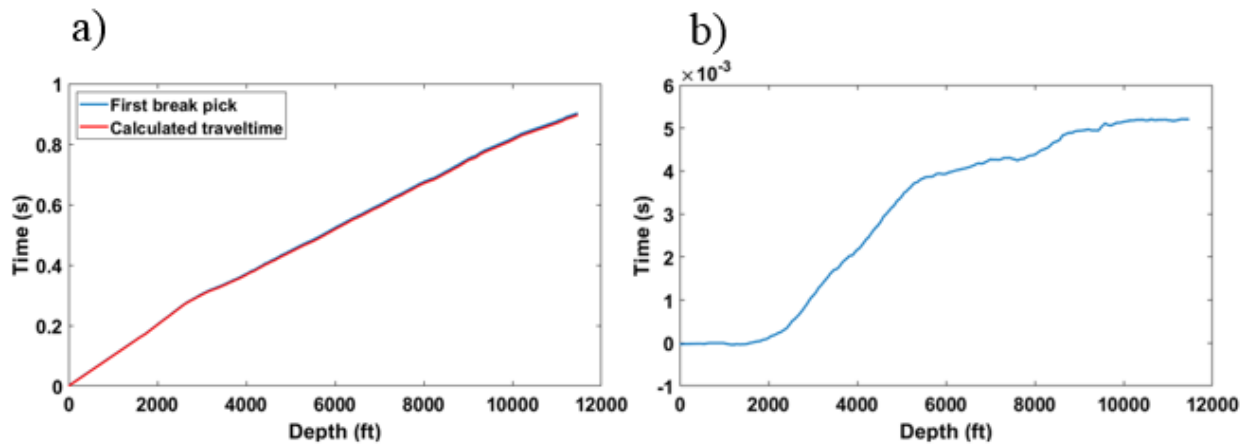


Figure 5.4 a) calculated one way travel time and picked one way travel time, b) Smoothed difference between the two one way travel times

To flatten the data the calculated traveltime from the sonic log is used and the smoothed difference between the calculated and picked travel time is added to include the effects of dispersion. This combination of calculated traveltimes and first break picks gives an accurate time shift which is applied to the upgoing wavefield synthetic VSP to flatten the data (Figure 5.5). The following analysis was done on the upgoing wavefield which can be obtained directly

when creating synthetic data, this removes the requirement of separating the wavefields with processing. This was done to get an exact upgoing wavefield with no concerns of the effectiveness of the wavefield separation processing step. This upgoing wavefield is also comparable to what would be recorded for a land surface seismic recording.

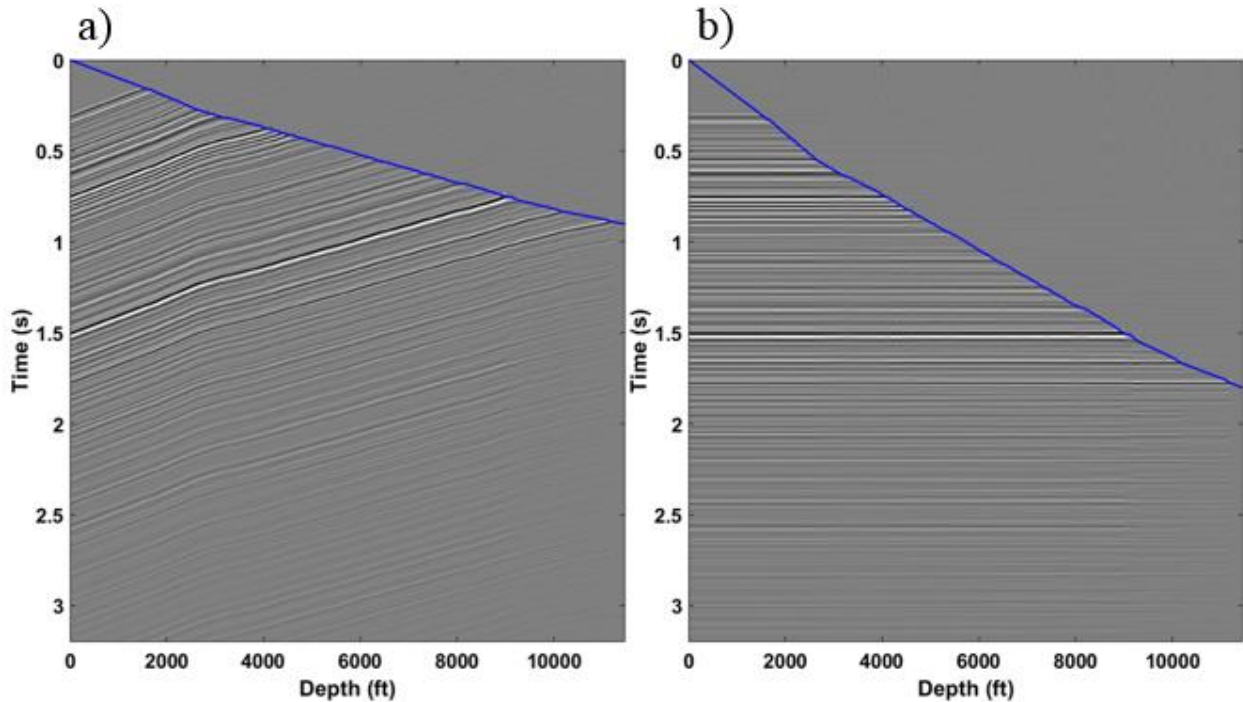


Figure 5.5 a) Synthetic VSP with first break pick displayed in blue with upgoing events, b) flattened synthetic VSP with first break pick displayed in blue with upgoing events

Creating this flat VSP is critical because any deviations from this will cause errors in the calculation of the corridor stacks. There are three traces that will be created from the flattened VSP and used in the internal multiple prediction analysis, the outside corridor stack, zero depth trace and full stack.

The outside corridor stack is created by stacking the outside traces (25 traces near blue line). These outside traces are comprised of the upgoing reflected seismic energy recorded

shortly after downgoing direct wavefront from the source. The resultant outside corridor stack approximates a primaries only stack. Since the stack is limited soon after the direct arrival there is insufficient time for long path multiples to occur. It is noted that short path internal multiples may be included.

The zero depth trace is the recording of the waves on the surface and is identical to a standard reflection survey. This trace will contain both primaries and internal multiples as is recorded on the surface.

The full corridor stack is created by stacking all the data to give another estimate of the primaries and internal multiples trace. In practice, when there is noise present in the data, the full stack can assist in removing noise through the stacking process resulting in an approximation of the primaries and multiples trace. For a noise free synthetic the full stack is not necessary as the zero depth trace serves the same purpose.

The full stack and zero depth trace are compared (Figure 5.6). This is a useful plot to see how accurate of an assumption the full stack is as this will be used to approximate the zero depth trace. It shows that some internal multiples are not properly estimated with the full stack.

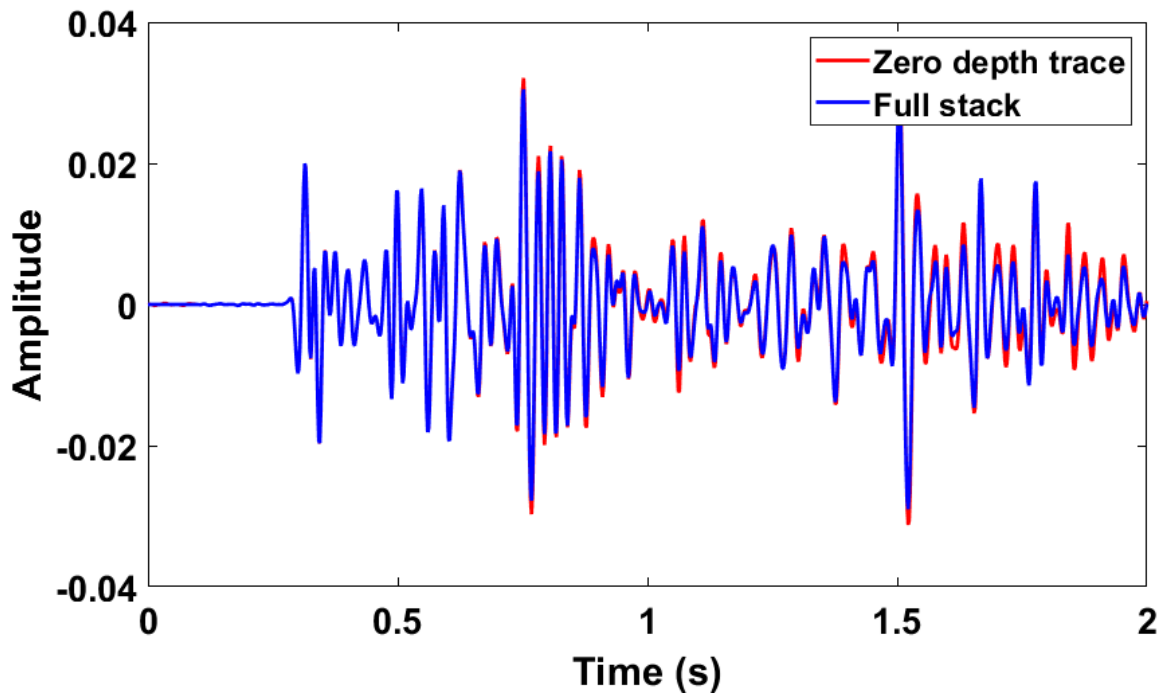


Figure 5.6 Full stack (primaries and multiples) and zero depth trace (primaries and multiples)

This is due to approximating the multiple energy from several geophone depths. But this is a reasonable estimate to use if the zero depth trace is not feasible. The next comparison is between the zero depth trace and outside corridor stack to assess how the internal multiples have degraded the seismic response (Figure 5.7). The outside corridor stack ends at approximately 1.9 seconds which corresponds to where the well log data ends. After this there are no longer any primaries and the subsequent data is all internal multiples because the model was only created to the depth where the well logs end.

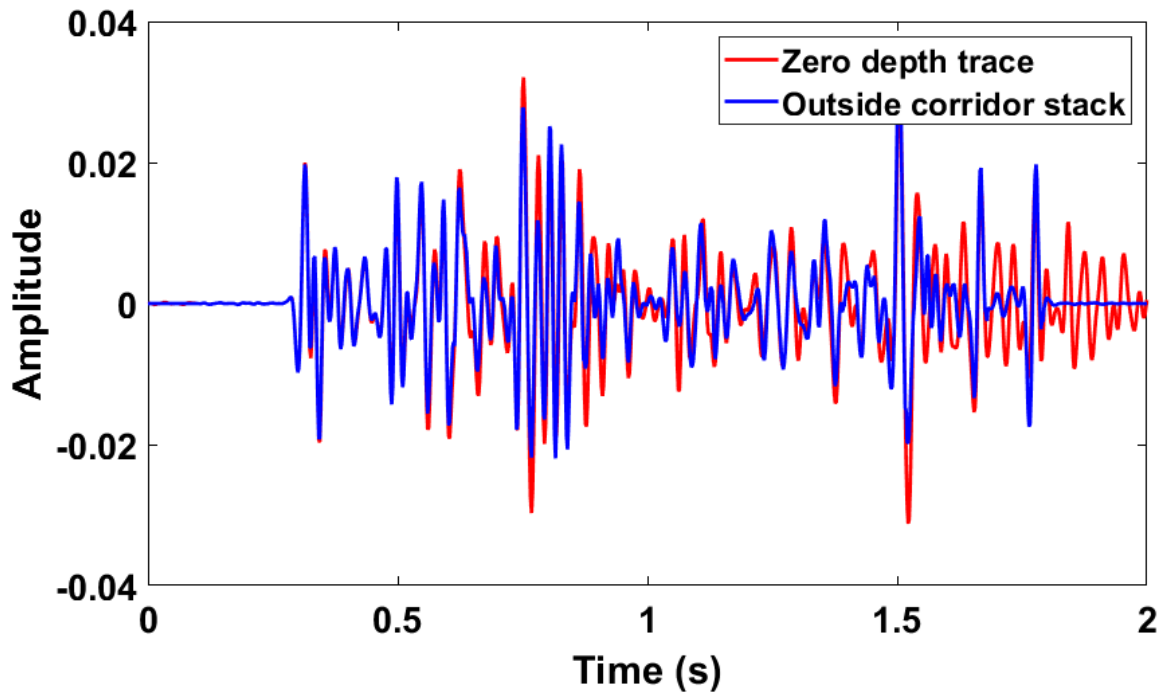


Figure 5.7 Outside corridor stack (primaries) and zero depth trace (primaries and multiples).

The resulting zero depth trace has similar character to the outside corridor stack but with drastically different amplitudes due to the inclusion of internal multiples. The locations with significant separation between the two traces such as between 1.5 and 1.9 seconds displays significant damage to the primaries. The higher amplitude portions of the trace could be incorrectly interpreted to be higher impedance contrasts in the subsurface. Within the time where there are primaries there are no locations of isolated internal multiples. The overlapping primary and multiple energy can make the application of adaptive subtraction challenging. By subtracting these two traces an estimate of the internal multiple trace as recorded on the surface is created.

5.4 Internal multiple prediction

The internal multiple prediction is carried out on the zerodepth trace using the 1D frequency domain version of the algorithm (Equation 1.42) with an epsilon value of 15. The epsilon value was chosen based on the sampling rate and frequency content to prevent artifacts. The prediction results are displayed in the downward generator space (Figure 5.8). Also displayed is the blue line along the outlining how the predicted internal multiples only reside in the upper right half of the space.

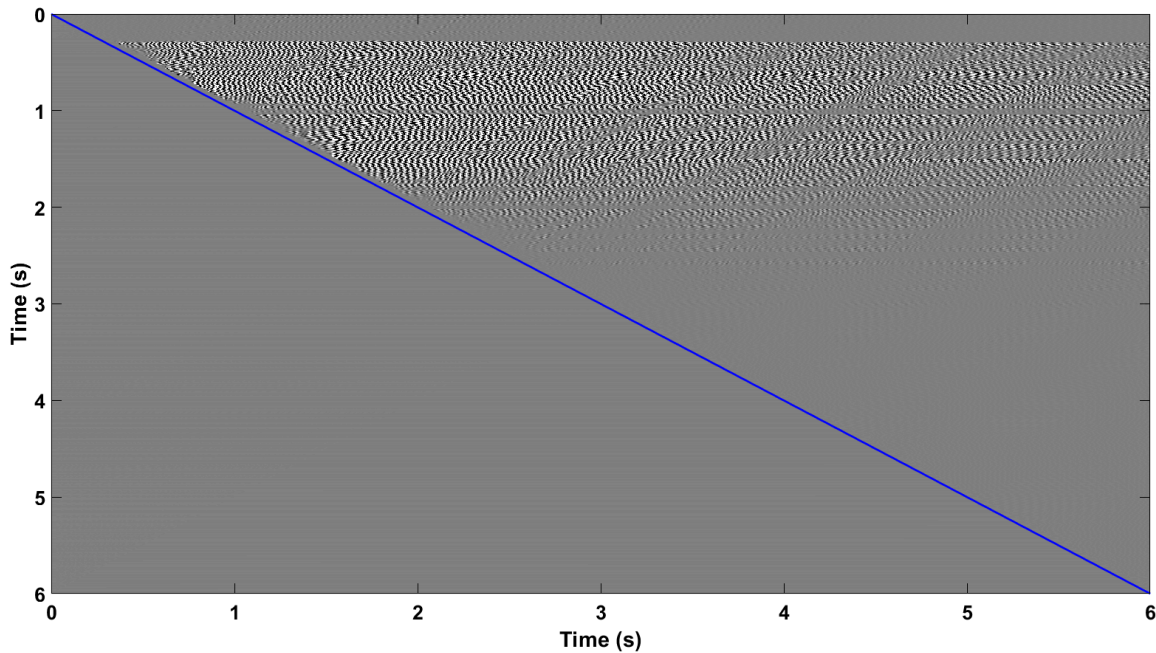


Figure 5.8 The 2D Downward generator space for the zero depth trace prediction

The downward generator space displays the complexity of the prediction when the input synthetic is also complex. For a given prediction time (vertical slice) there is a significant number of internal multiples which all contribute to the final multiple trace once stacked over all pseudodepths. This space does not contain the isolated internal multiples seen on previous synthetic examples. First the standard 1D prediction approach is done by summing over the all pseudo-depths. To determine the success of the prediction the multiples trace is created by

subtracting the zero depth trace and the outside corridor stack and compared to the prediction (Figure 5.9).

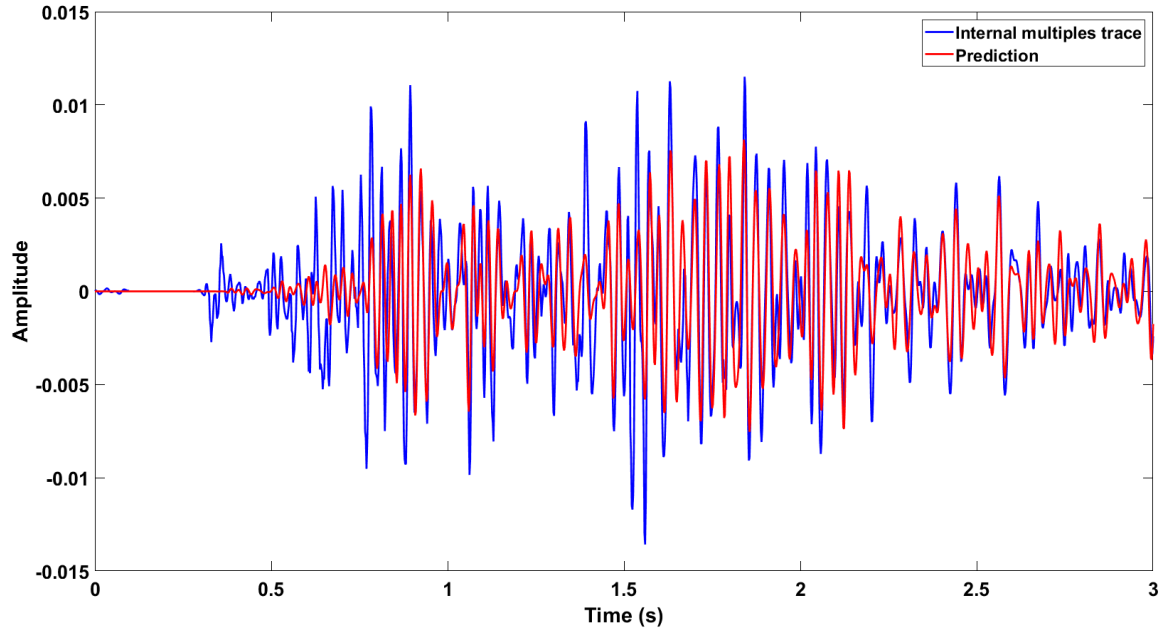


Figure 5.9 Internal multiples trace in red and internal multiple prediction in blue.

Overall the method has been able to predict the internal multiples of this example with an extremely complex multiple train. The largest errors occur early in the trace until approximately 0.6 seconds. After this the predicted trace shares similar character to the multiple trace with some issues in the amplitudes of the predictions.

5.4.1 Error quantification

To quantify the accuracy of the prediction a moving average of the absolute value of the trace is calculated. This is carried out on both the internal multiples trace and the subtraction of the internal multiples trace and the prediction. If the multiple is overpredicted or incorrectly predicted in terms of polarity then the plot will be higher than the baseline multiple trace. The plot of the difference will only reduce to zero if the multiple is predicted exactly. This plot also

displays the locations of significant prediction error with crossover between the two averaged traces. If the multiple is slightly under or overpredicted this will display as a reduction in the value. This again raises the question of what successful multiple attenuation is. Ideally the internal multiples would be predicted exactly but that will rarely be the case. Often there is the concern of damaging primary energy when removing multiples. If a multiple is underpredicted than it is thought to have been partially attenuated. If the multiple is overpredicted than there are concerns that you are damaging primaries. Is more detrimental to have partially attenuated multiples which obscure the primaries then overpredicted multiples and attenuate some of the primaries. If there are concerns about attenuating primaries, then the multiple prediction can simply be scaled back to the point that all multiples are underpredicted. This rolling average is displayed for the internal multiples trace and internal multiples trace minus the prediction (Figure 5.10).

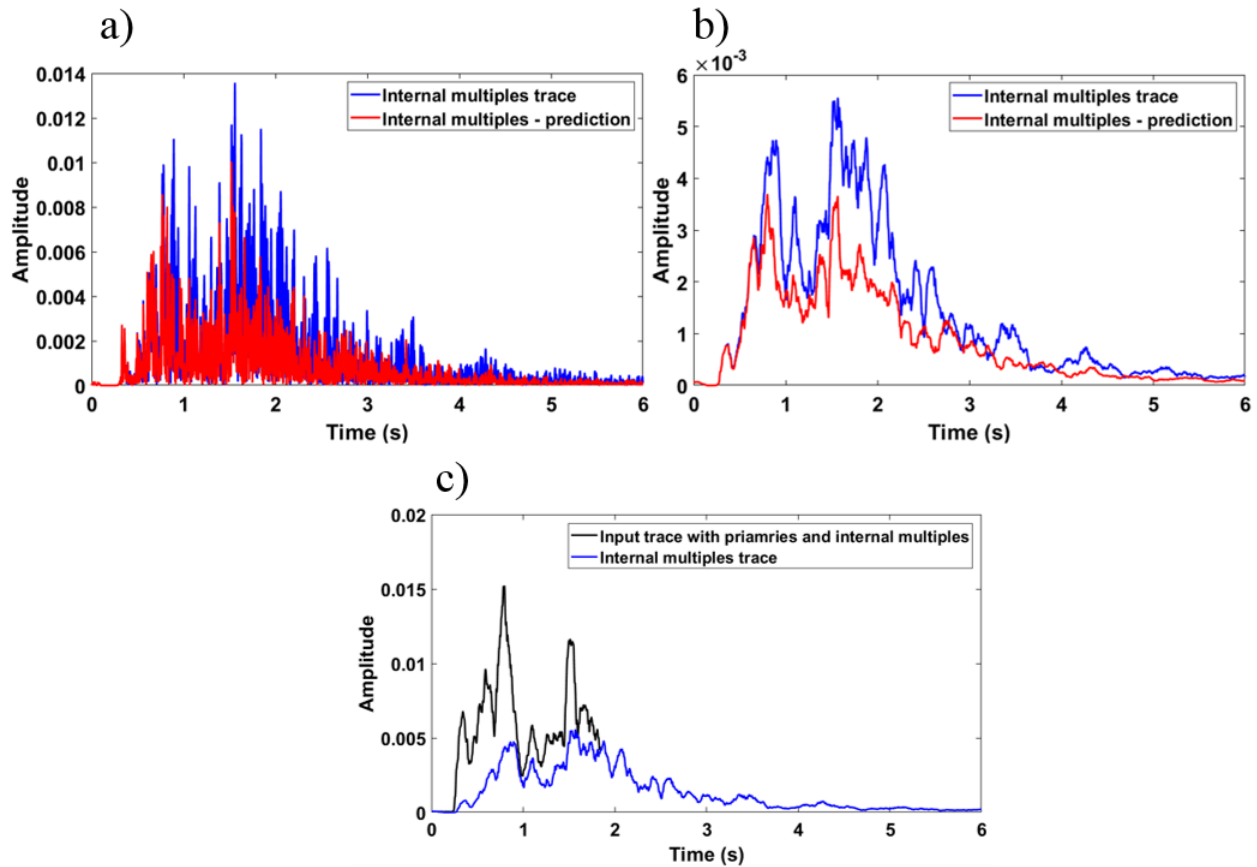


Figure 5.10 a) Multiples trace in red and internal multiple prediction in blue taken as absolute value of the multiples trace and difference with the multiples trace, b) a 50 point moving average window is used to smooth the traces c) a 50 point moving average window is applied to the input data trace for reference to the internal multiple trace

The prediction algorithm appears capable of successfully attenuating some of the internal multiples in the data. The amount of multiple energy is shown to have decreased due to multiple attenuation except for a few locations which display crossover. The amount of multiple energy relative to the input trace is also displayed. Early in the trace the majority of the events are primaries, but at later times beginning around 1 second the amount of internal multiples energy relative to the total input trace amplitude increases significantly. Next the 2D adaptive subtraction is applied to the downward generator space. The number of traces in the

pseudodepth dimension is reduced through stacking prior to 2D adaptive subtraction (Figure 5.11).

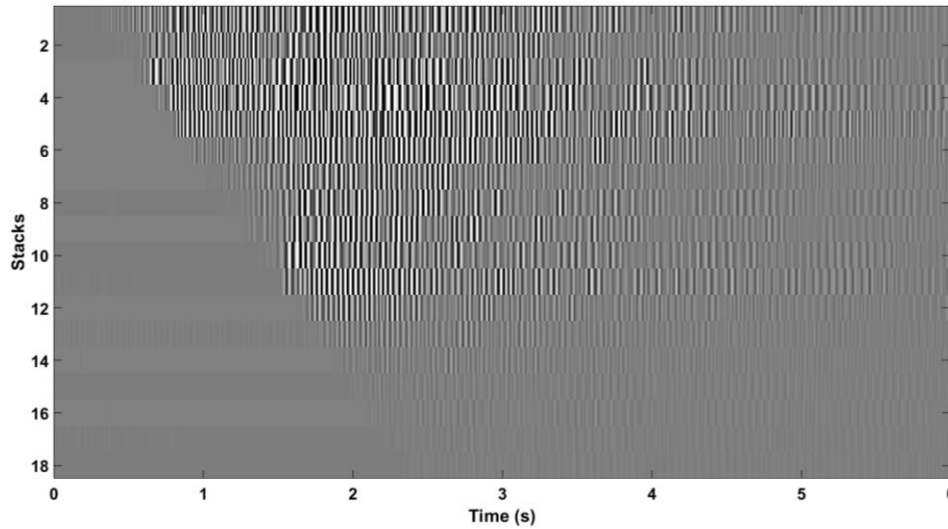


Figure 5.11 Displaying the 2D downward generator space for the synthetic VSP after stacking

The result of applying 2D adaptive subtraction is compared to the internal multiple trace (Figure 5.12).

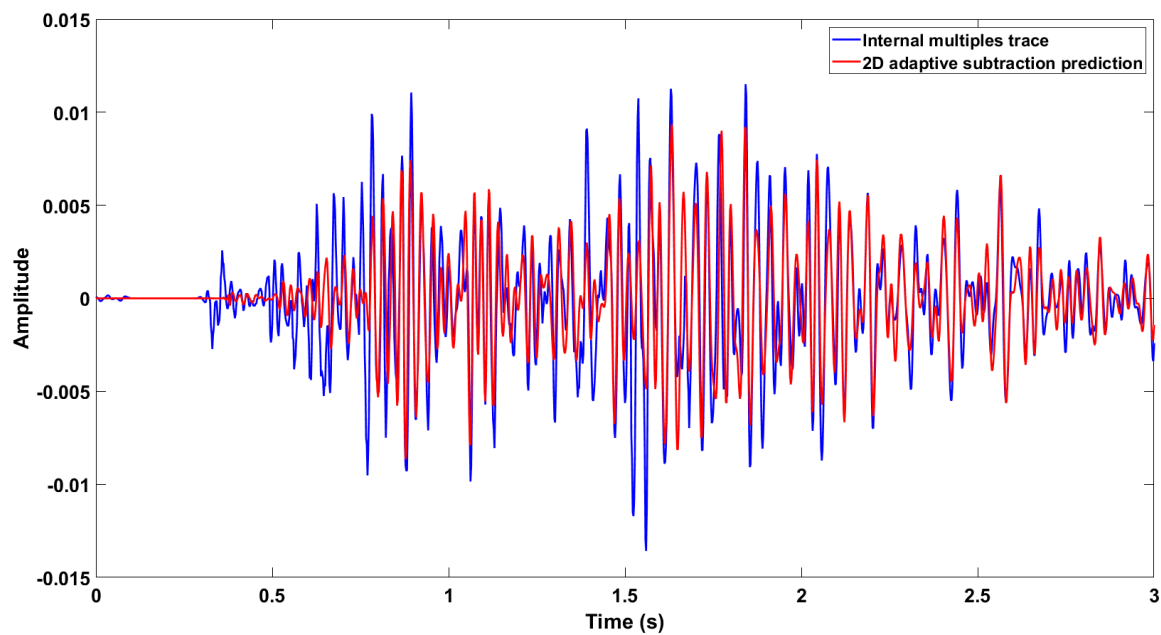


Figure 5.12 Multiples trace and internal multiple prediction from 2D adaptive subtraction.

There is a reasonable match to the internal multiple trace with the bulk of the mismatch occurring early in the trace. For this plot it is difficult to see the uplift due to the 2D adaptive subtraction as there are still amplitude mismatches between the predicted result and the multiples. The same moving average numerical plots are created to compare the prediction and prediction with 2D adaptive subtraction (Figure 5.13).

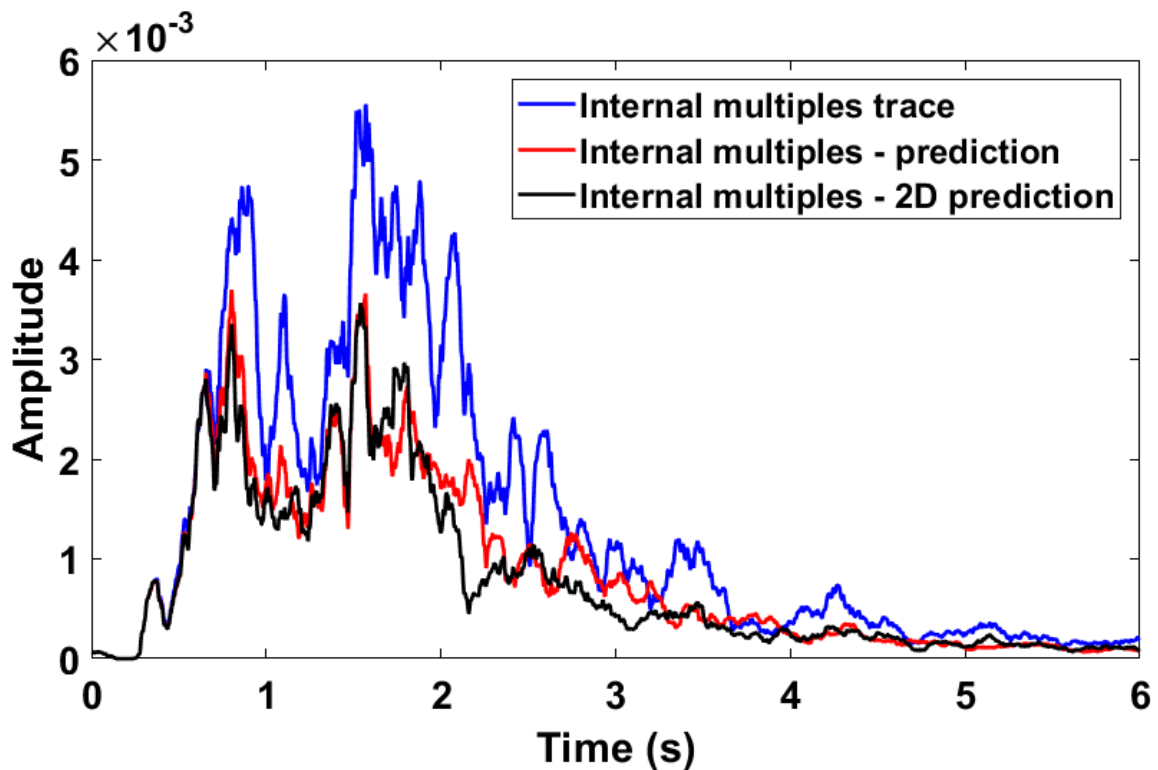


Figure 5.13 Absolute value of multiple trace, and difference with the two predictions with rolling average window of 50.

In both cases the multiple energy has been diminished. The 2D adaptive subtraction has improved the prediction at mostly later times in the trace with some locations throughout where it has been slightly detrimental. Note that there are no primaries after approximately 1.9 seconds. This may have assisted the adaptive subtraction as there are some locations with

isolated internal multiples. This may not be possible for recorded data cases as there will continually be primary energy.

All traces are compared qualitatively by viewing in conventional wiggle plot display (Figure 5.14). Included is the outside corridor stack which is the primaries only trace and goal to return to after internal multiple attenuation. The zero depth trace which is the input into the algorithm and includes primaries and internal multiples, and finally the result of attenuating internal multiples applying the prediction to the zero depth trace.

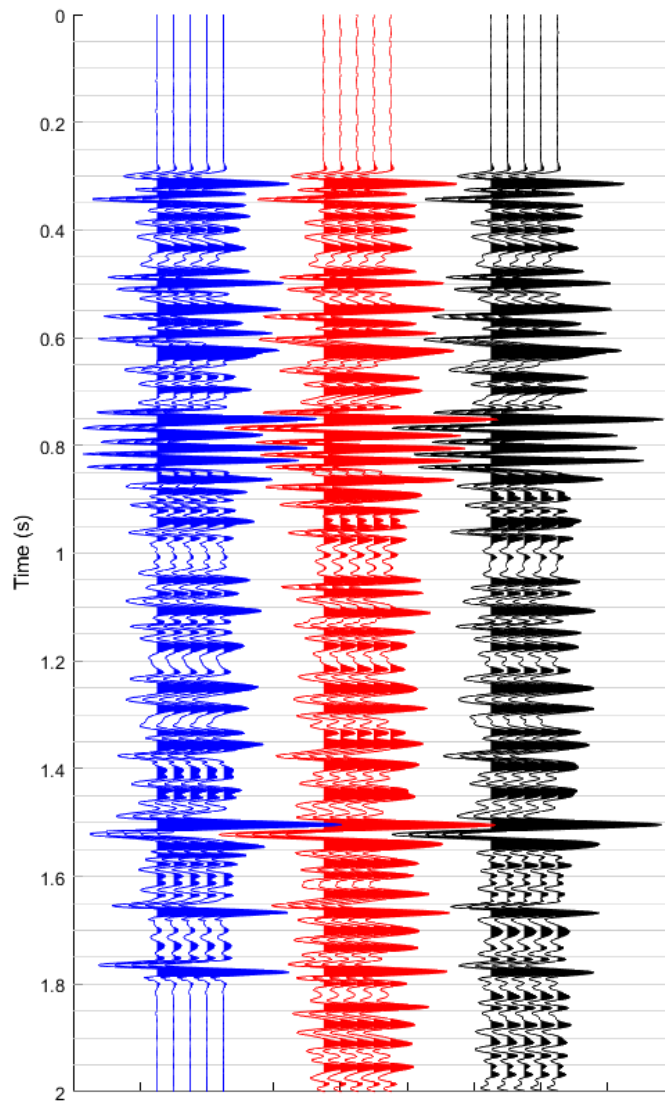


Figure 5.14 outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black.

The result is analyzed near the base of the trace where the initial internal multiple modeling shows significant internal multiples (Figure 5.15). For all images in the figure the blue and red traces remain constant, only the black trace changes as noted.

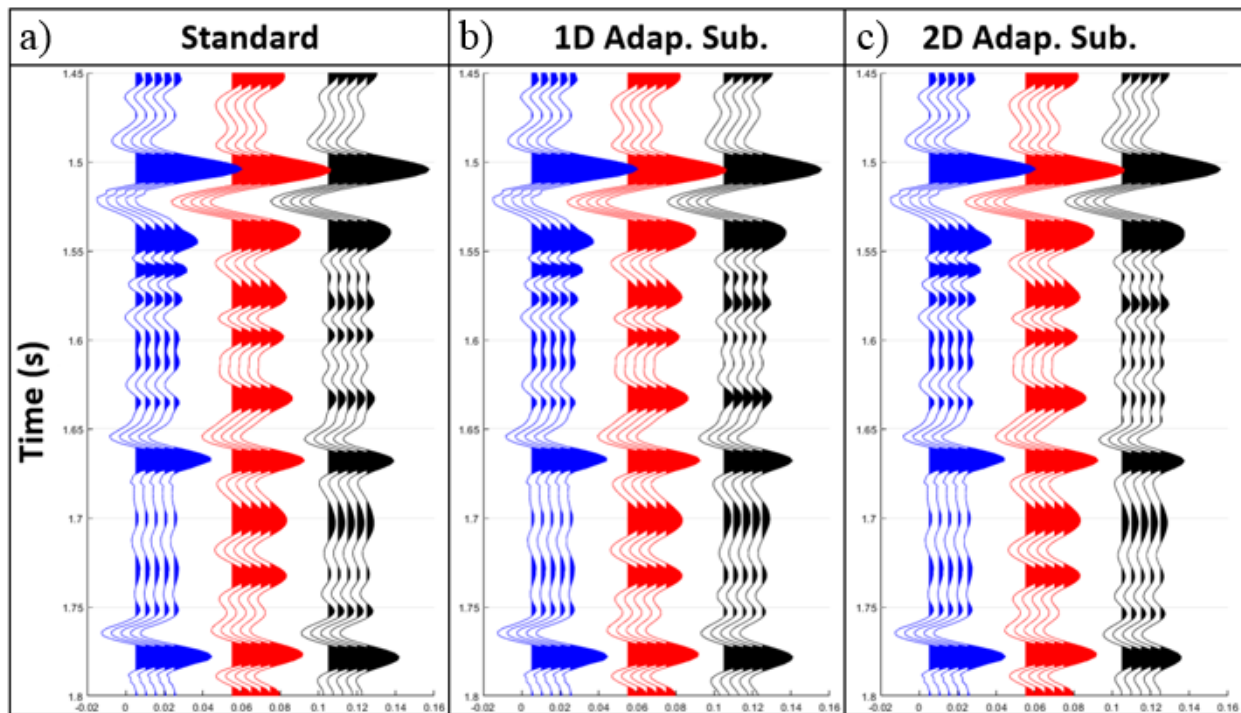


Figure 5.15 Outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black. a) b3 subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction

All versions of the subtraction have improved the dataset by taking the input trace which was significantly contaminated with multiples at this level and return the traces to a state where it resembles the primaries only trace. There are several locations with polarity reversals and events previously obscured events can now be imaged. Below 1.6 seconds there are two peaks that are only visible with the use of the 2D adaptive subtraction. Displayed is the result of following all the previously outlined steps but with a 37Hz Ricker wavelet (Figure 5.16).

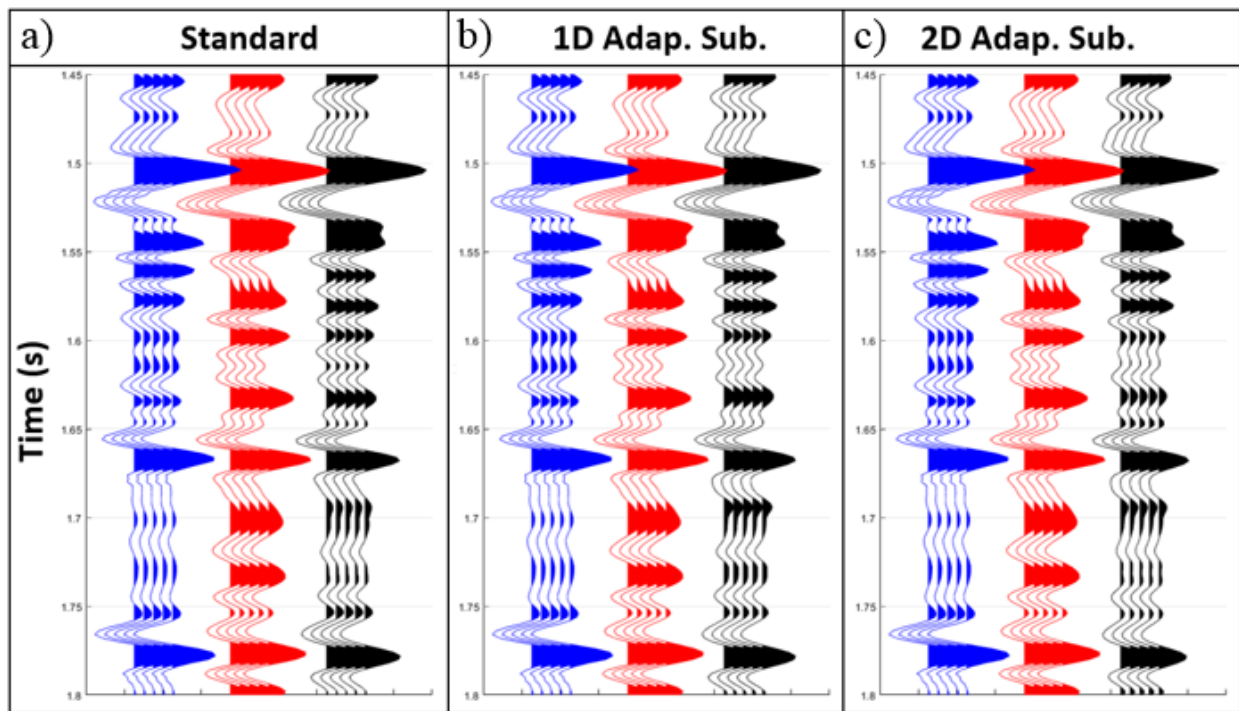


Figure 5.16 Using 37 Hz Ricker wavelet displaying outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black for a) b3 subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction

A significant improvement is seen on the trace with an enhanced similarity between the primaries only trace and the result of attenuating the multiple contaminated trace. For this higher frequency case the difference between the standard, 1D and 2D adaptive subtraction methods is more subtle. Though all options have improved the input dataset through internal multiple attenuation.

5.5 Conclusions

There were two main objectives for the synthetic test. To see how the algorithm performed on a significantly complex synthetic and to determine if the method is applicable for real data cases. The results showed it can succeed in this complex scenario. The direct subtraction and

the adaptive subtraction versions all displayed improvements due to multiple attenuation. This success gives confidence moving forward to apply the inverse scattering series internal multiple prediction method to real data.

Chapter Six: Devon Real Data

6.1 Application to recorded data

A continuing goal of CREWES is the application of the inverse scattering series to land seismic data. To this end the tools and workflows that have been developed are applied to several real datasets. The synthetic modeling from the well log data displayed the potential for significant internal multiples and it was shown how the ISS method could attenuate and improve the dataset. This improvement occurred deeper in the section where the internal multiples were more significant. This improvement was displayed both qualitatively and quantitatively. Given the success of the synthetic tests the algorithm is applied to the real datasets. This includes the VSP, the pre and post stack 3D seismic data.

6.2 Recorded VSP

The VSP was recorded both at zero offset and walkaway for a single azimuth, with 43 three component geophones and a geophone spacing of 49.34ft. The geophone array was moved six times to give the final dimensions of the survey with a top depth of 55ft and a bottom depth of 11304.52 ft from KB. The geometry of the source and receiver locations in the wellbore is displayed (Figure 6.1).

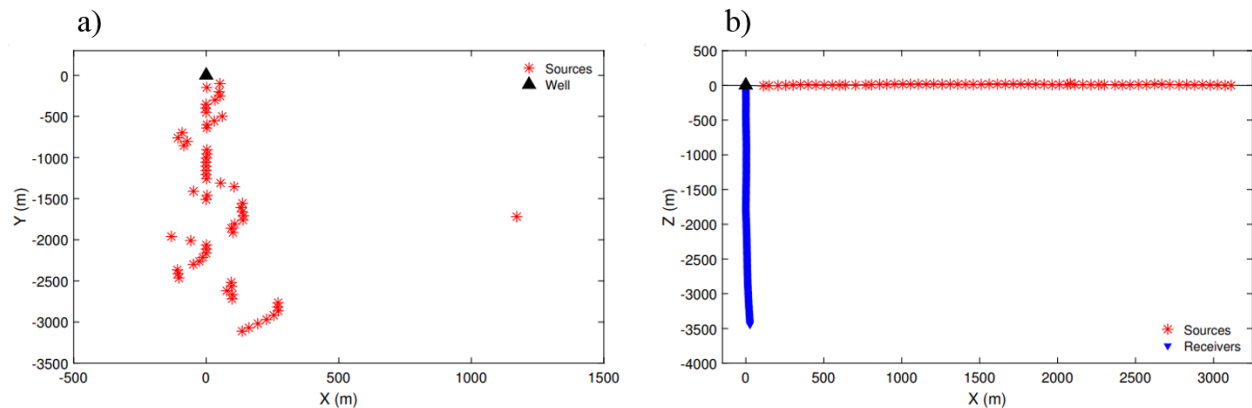


Figure 6.1 Figure from (Cova, et al., 2018) a) source locations in Easting and Northing, b) Receiver location in depth and in line with the sources

For the internal multiple analysis only the zero offset data is analyzed. This VSP used a Vibroseis source with a linear 16 second sweep from 2-140 Hz with 0.5 sec cosine tapers and two sweeps per vibe point had a 6 second recording length at 1ms sample rate.

6.2.1 Processing

The VSP was processed by Raul Cova at CREWES using the vertical component of the three-component geophone (Cova, et al., 2018). Three key steps to prepare the data for analysis are; amplitude recovery, wavefield separation and flattening. A gain of was applied to recover the amplitudes due to geometric spreading, transmission and other losses. The downgoing direct arrival was picked to use the arrival times to flatten the data. Next the upgoing and downgoing wavefields need to be separated as the demultiple analysis only requires the upgoing waves. This is done by flattening the data and using a trace median filter to remove the upgoing energy. Since the upgoing and downgoing waves have different kinematics, when the data is flattened on the downgoing and smoothed, any dipping energy gets attenuated. This results in a dataset with only the downgoing wavefield. Then the downgoing wavefield can be subtracted from the

original dataset to give the upgoing wavefield. No deconvolution is applied to the VSP to avoid any impacts the deconvolution could have on the multiples, as certain types of deconvolution act as a multiple attenuator. The result after processing is displayed (Figure 6.2).

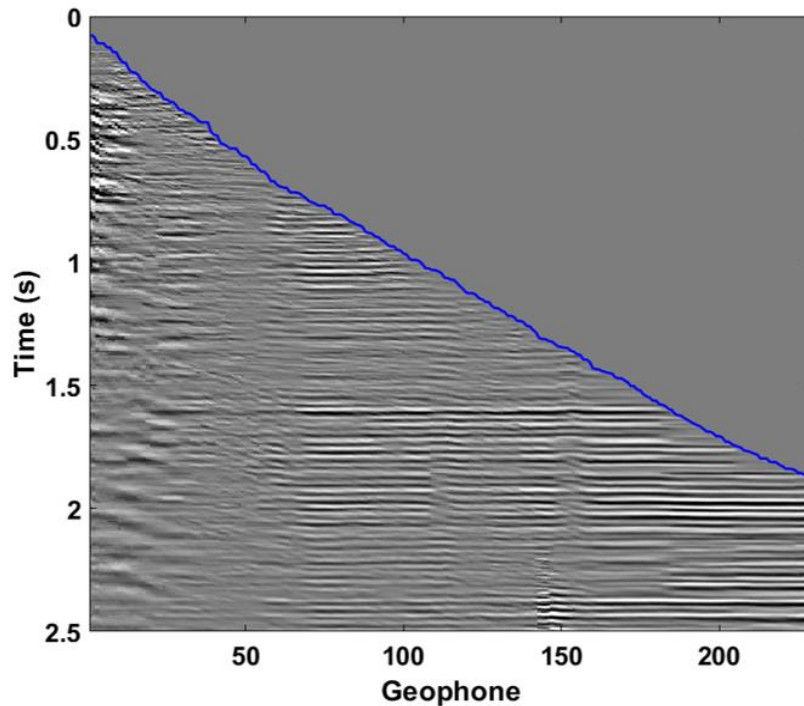


Figure 6.2 Recorded VSP after processing with upgoing wave

The data quality is acceptable with distinct upgoing and downgoing wavefields. At shallow wellbore depths, the data has significantly higher noise levels and the upgoing and downgoing waves can no longer be found. This is thought to be caused by poor geophone coupling to the wellbore. There are also a few traces with high amplitude noise through the VSP. The best data quality in terms of signal to noise occurs deeper in the wellbore. With the data processed the outside corridor stack (primaries only) and full stack (primaries and multiples) can be created. It would be preferred to use the zero depth trace or a smaller stack from the shallower wellbore depth. Due to the poor geophone coupling and high noise levels the results of the trace would not be reliable. The synthetic modeling and tests shown in Figure 5.6

indicates that, for this dataset, the assumption of the full stack being used in place of the zero depth trace is reasonable. The outside corridor stack was created using a window of 25 traces from the direct arrival. Displayed is the frequency spectrum of the full and outside corridor stack from the VSP where peak frequency is approximately 30 Hz (Figure 6.3).

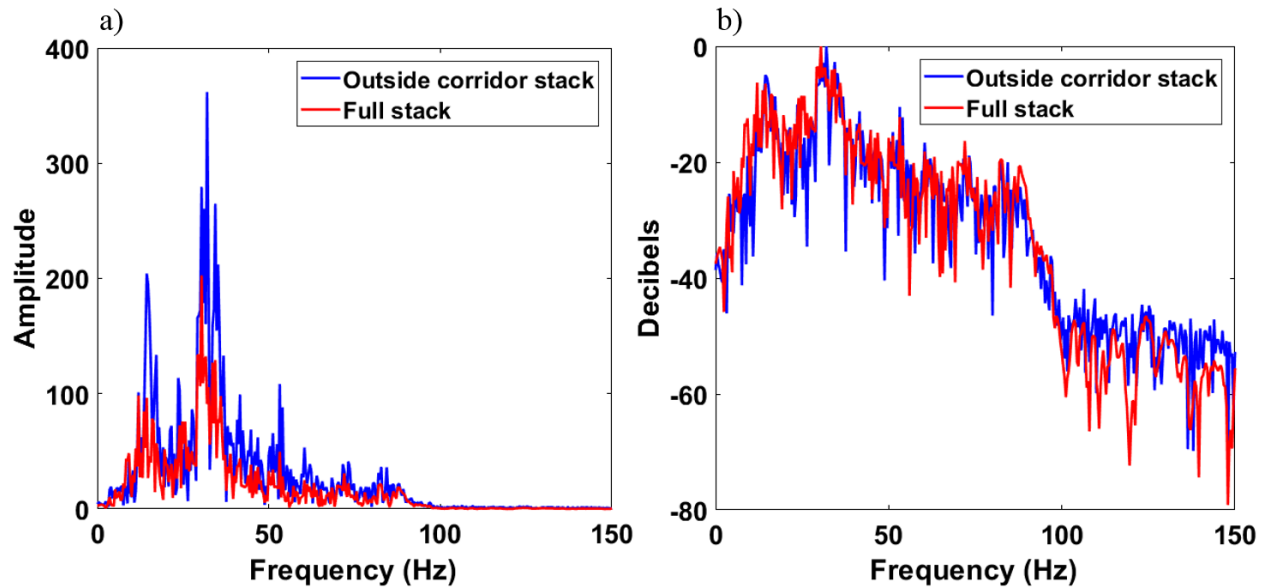


Figure 6.3 a) Frequency content differences between full stack (Primaries and multiples) to outside corridor stack (primaries only) and b) using a decibel scale

The amplitude spectrum has significant notching and variation in amplitude through the bandwidth.

6.2.2 Corridor stacks

The outside corridor stack and the full stack are compared to assess the level of multiple contamination present in the recorded data (Figure 6.4).

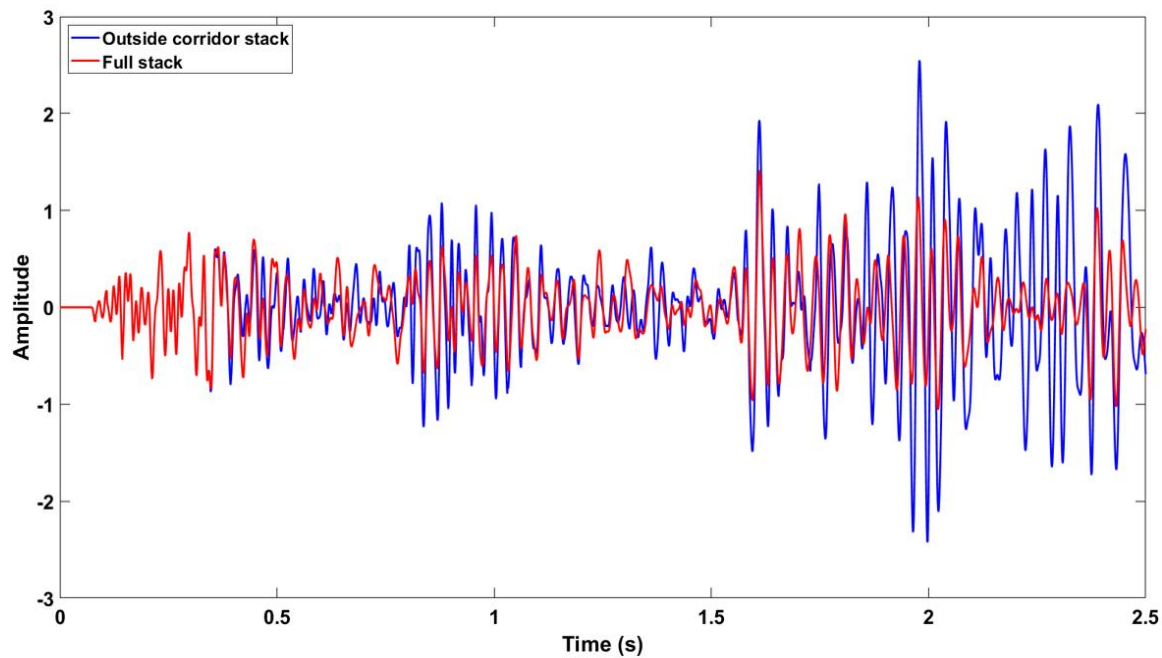


Figure 6.4 Full stack (Primaries and multiples) and outside corridor stack (primaries only)

There are significant variations between the two stacks. Like the synthetic test most of the differences appear to be the amplitudes of the traces at a given time. Some of this variation is due internal multiples present in the data. With increasing time, the amplitudes of the two traces diverge. The only scaling applied to these traces was done during the VSP processing. At this stage it is unclear to what extent the difference between the two is due to multiples and understanding this is key, as the multiple trace is created by subtracting the two. Note that it is assumed that all the difference between the two traces is due to multiples. In practice, there could be other noise which is contributing to the difference between the two traces. It is proposed that some of this may be due to additional amplitude gain being required for the full stack.

6.2.3 Internal multiple prediction

The full stack is used as the input into the internal multiple prediction algorithm. This is compared to the multiples trace created by subtracting the outside corridor stack and the zero depth trace (Figure 6.5).

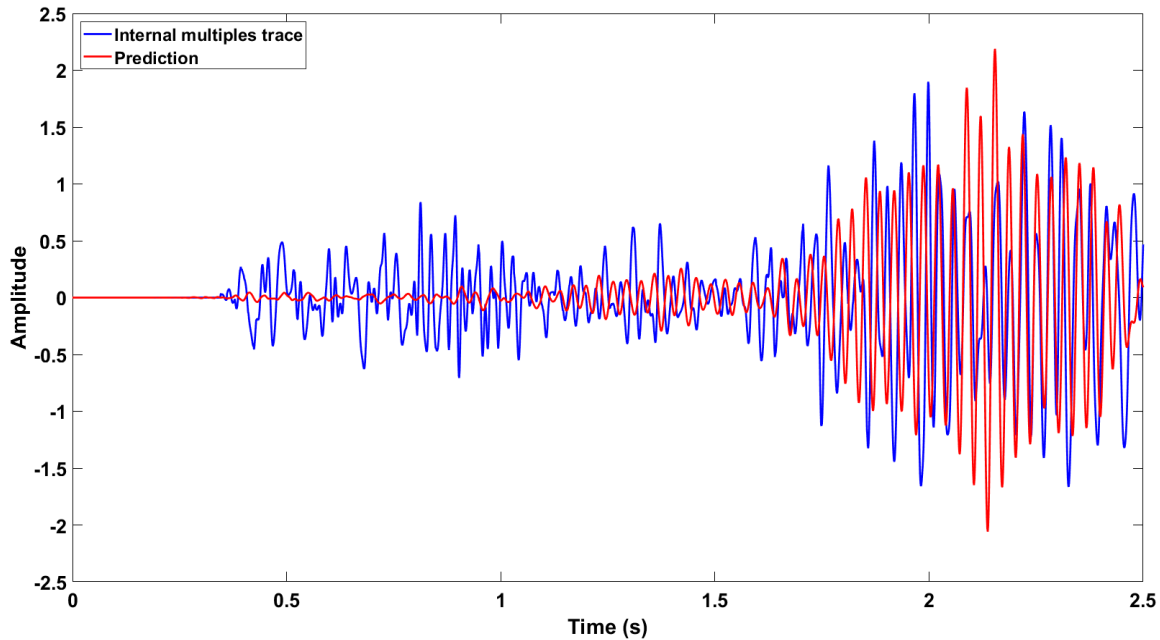


Figure 6.5 Internal multiple prediction and internal multiples trace

Although the prediction does not appear to match the internal multiple trace, the envelope between the two is approximately matching. At a few locations the predicted amplitudes match well but there are other portions of the trace with a poor match to the multiples estimate, with either the amplitudes being incorrect or the polarity of the events not matching. This variation from a reasonable to poor match occurs in quick succession. Due to the concerns with the amplitudes and how quickly the mismatch can vary in time, uplift may be found with the 2D adaptive subtraction. The dataset was stacked to give 8 traces as seen in Figure 6.6.

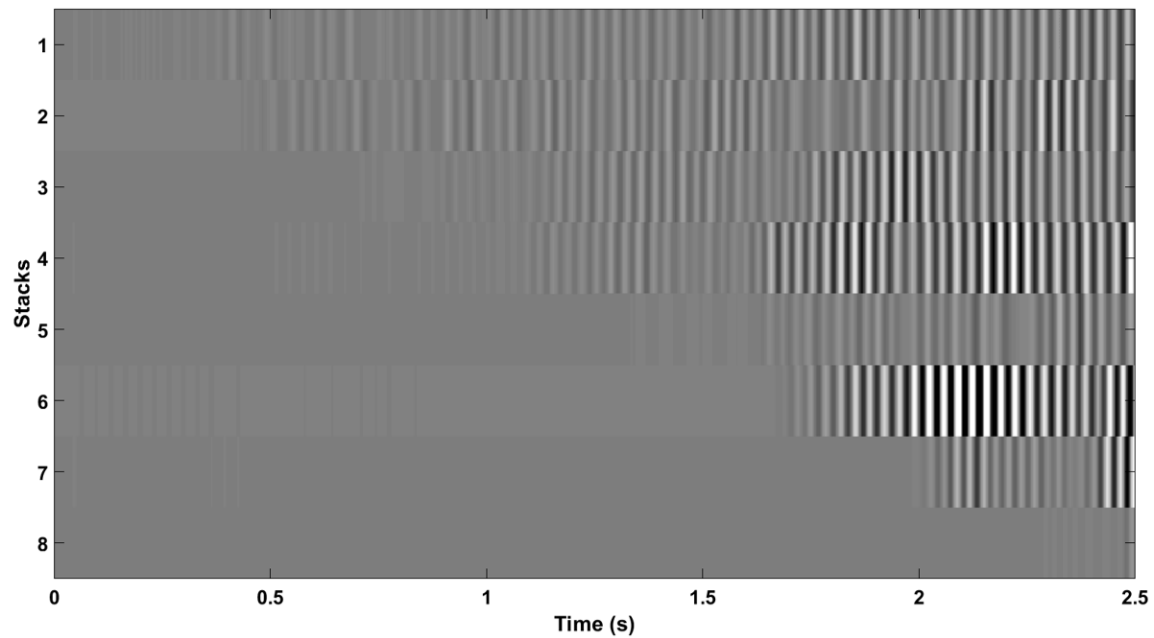


Figure 6.6 Downward generator space after stacking for 2D adaptive subtraction

There appears to be significant variability between the different pseudo-depth. As previously postulated, if there are any issues with the amplitudes or any scaling which was incorrectly done, it is possible that the 2D space of the downward generator may assist. The 2D adaptive subtraction result is compared to the internal multiples trace (Figure 6.7).

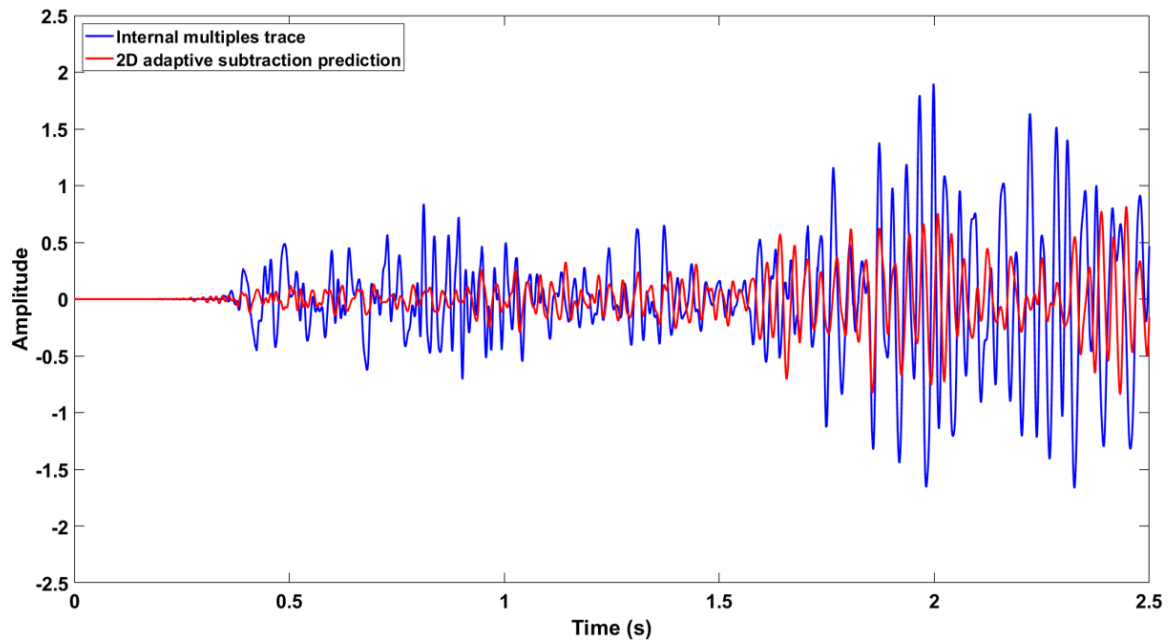


Figure 6.7 Internal multiple prediction with 2D adaptive subtraction and internal multiples trace

There appears to be an improved match to the internal multiples trace. Compared to the standard approach, the trace appears to match in terms of polarity at significantly higher rate. The amplitude of the events has generally decreased. Again, even the prediction envelope appears improved. The outside corridor stack is compared to both the full stack and the multiple attenuated full stack (Figure 6.8).

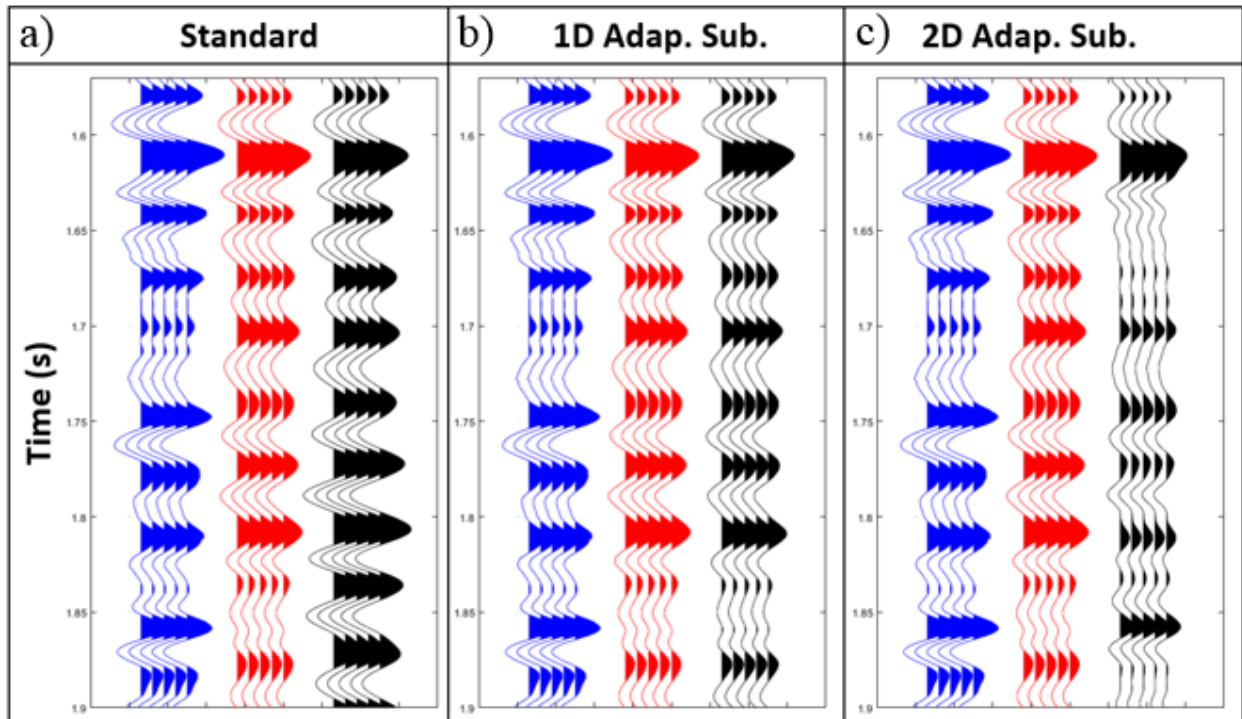


Figure 6.8 Outside corridor stack (primaries) in blue, zero depth trace (primaries and multiples) in red and zero depth trace after internal multiple attenuation in black for a) b3 subtraction, b) 1D adaptive subtraction and c) 2D adaptive subtraction

For the real VSP the result is more variable compared to the synthetic case. For the standard approach with no adaptive subtraction there is little uplift with multiple attenuation process. The 1D adaptive subtraction has greatly reduced the internal multiple energy and there is a minimal difference between the input zero depth trace and the final internal multiple attenuated trace. Lastly for the 2D adaptive subtraction, there are locations where it appears to have helped. The primary below 1.8s appears to have been recovered from this multiple attenuation but other locations such as above 1.6 seconds appear to have been detrimental. In the recorded data there are locations with improvements, but this is subjective.

6.3 3D PSTM stacked seismic data

The post stack PSTM 3D seismic data is analyzed. This has been processed through to anisotropic prestack time migration with steps including denoising, amplitude scaling, deconvolution and 5D interpolation (Figure 6.9).

```
C20 SOURCE: VIB/DYN SWEEP LENGTH: 18 SEC.  FREQ.:2-108 HZ  LAYOUT: 3
C21 SRC LINE/INT.: 990(N-S)/165 FT RCVR LINE/INT.: 1155(E-W)/165 FT
C22 GEOPHONES: I/O SENSOR  ARRAY: 3/6 FT CIRCLE  INSTRUMENT: GSR
C23 FOLD: 300  PATCH: 36 LN X 288 CHAN / 10368  NEAR/FAR OFFSET: 6 / 31575 FT
C24 FIELD RL: 6000 MS  FIELD SI: 2 MS  FIELD FILTER: 1 - 207 HZ  FORMAT: SEGD
C25 PROCESSING SEQUENCE (DOMAIN:TIME):
C26 Reformat; Powerline Noise Removal; Min. Phase; Geom. Assign; Trace Edits;
C27 Geometric Spreading; SC Scal.;Pre-Decon Noise Atten.;SC Scal.;SC Decon
C28 (600-2600MS@0FT,2100-4200MS@18000FT.,100MS OPER,0.01% PW);Unbias SC Scal.;
C29 Phase&Static Matching;Refrac. Stat.(Datum=1900FT,Vw=2500FT/S,Vr=8500FT/S);
C30 Vels.:1mi X 1mi; 1st Pass Res(500-2000MS,+/-16MS);Vels.:0.5mi X 0.5mi;
C31 2nd Pass Res(500-2000MS,+/-8MS);Noise Atten; 2nd Pass SC Decon;Noise Atten;
C32 Unbias SC Scal; 5D Interp.; NMO Removal;PSTM Vels:0.5mi X 0.5mi;Anisotropic
C33 Kirchhoff PSTM; RNMO; Stack; Polarity(Norm); SEG Y Out
```

Figure 6.9 EBCDIC header displaying acquisition parameters and processing

Specifics on the algorithms and parameters for the processing is unknown. Applying a data driven multiple attenuation as the final step in the processing may not be an ideal stage in the sequence. As there have been numerous processing steps leading up to this with several denoise and amplitude adjustment steps. The multiple attenuation is tested at this processing stage both to attempt to improve this individual dataset and determine the applicability of the method to be applied to fully processed data. If successful on this fully processed data, the internal multiple attenuation method can be easily applied to any data volume without requiring reprocessing. For the 3D seismic data tests, the trace to be analyzed is the one which is coincident with the well and VSP. Displayed is the frequency spectrum of the single PSTM stacked trace (Figure 6.10).

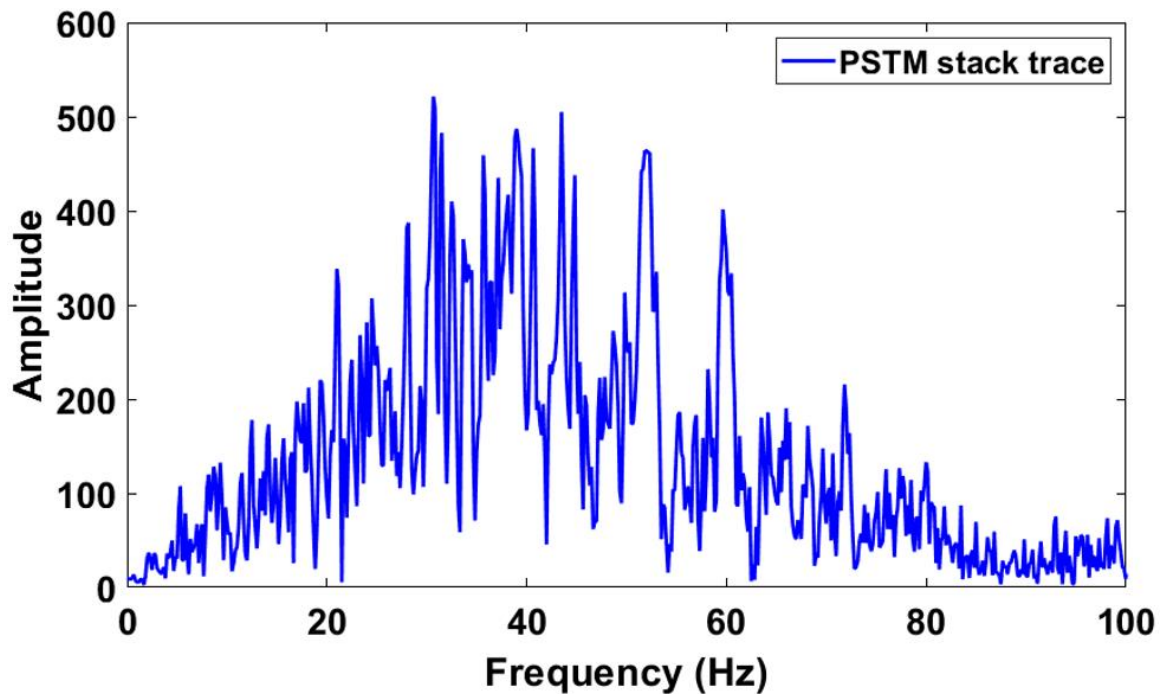


Figure 6.10 Frequency spectrum of PSTM Stack

This shows central frequency around 37 Hz and a broader range of frequencies than the VSP. The spectrum is not very white though there is less notching than the VSP amplitude spectrum.

6.3.1 Well tie

The well is tied to the seismic data to compare with the synthetic modeling and testing and is completed using Hampson-Russell software. To tie the well the checkshot was used from the VSP which applied a slight stretch and a bulk shift and a 37Hz Ricker wavelet is used. Also shown is the cross correlation of the synthetic seismic to the data (Figure 6.11). This synthetic was created with primaries only in Hampson Russell.

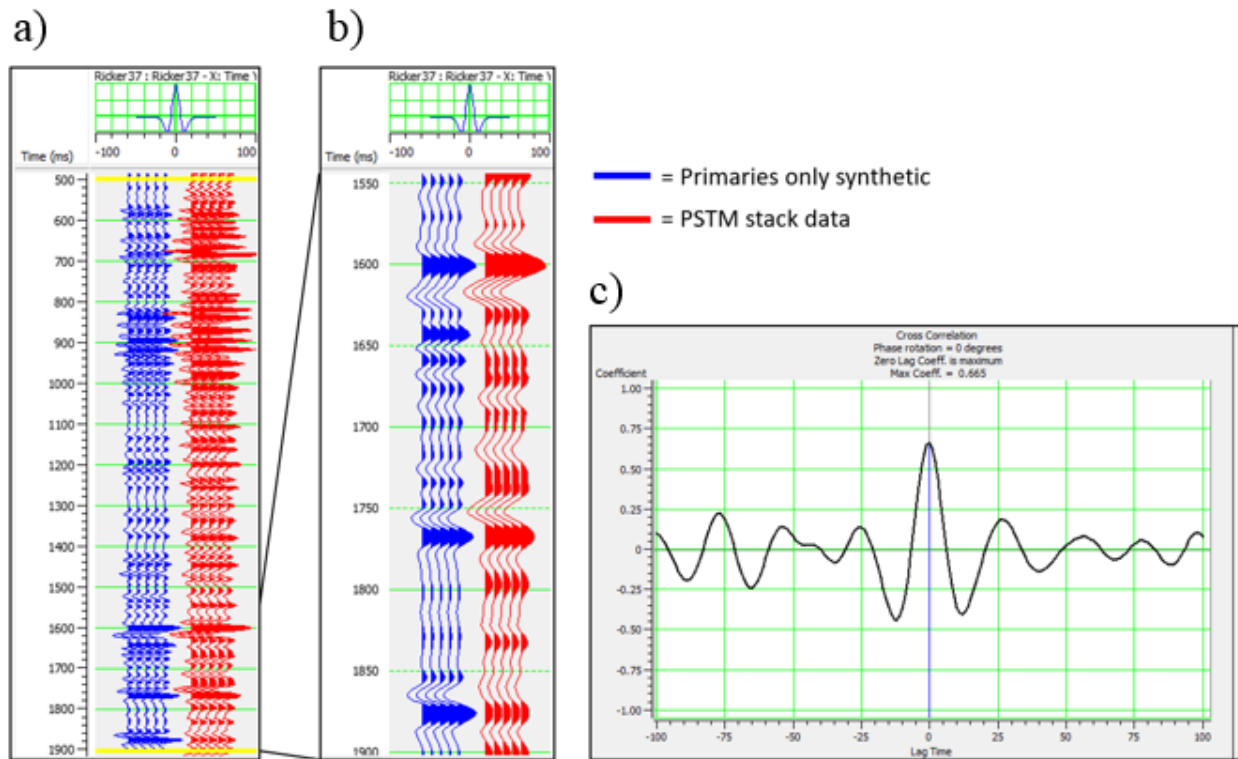


Figure 6.11 Well tie to 3D PSTM stacked data with a) synthetic in blue and stacked data in red, b) well tie over the time window of 1550–1900ms and c) cross correlation window with a max coefficient of 0.665

There appears to be a reasonable tie to the zero offset synthetic. The three large reflectors at 1.6 1.75 and 1.85 all seem to tie the stacked data well. Some of the dissimilarity between the two could be due to multiples as this synthetic does not include these events. The 3D stacked seismic data is also compared to the synthetic VSP. This is displayed both over the entire well log interval and the deeper section where the significant internal multiples were displayed and successfully attenuated. This is also compared by overlaying the traces directly (Figure 6.12).

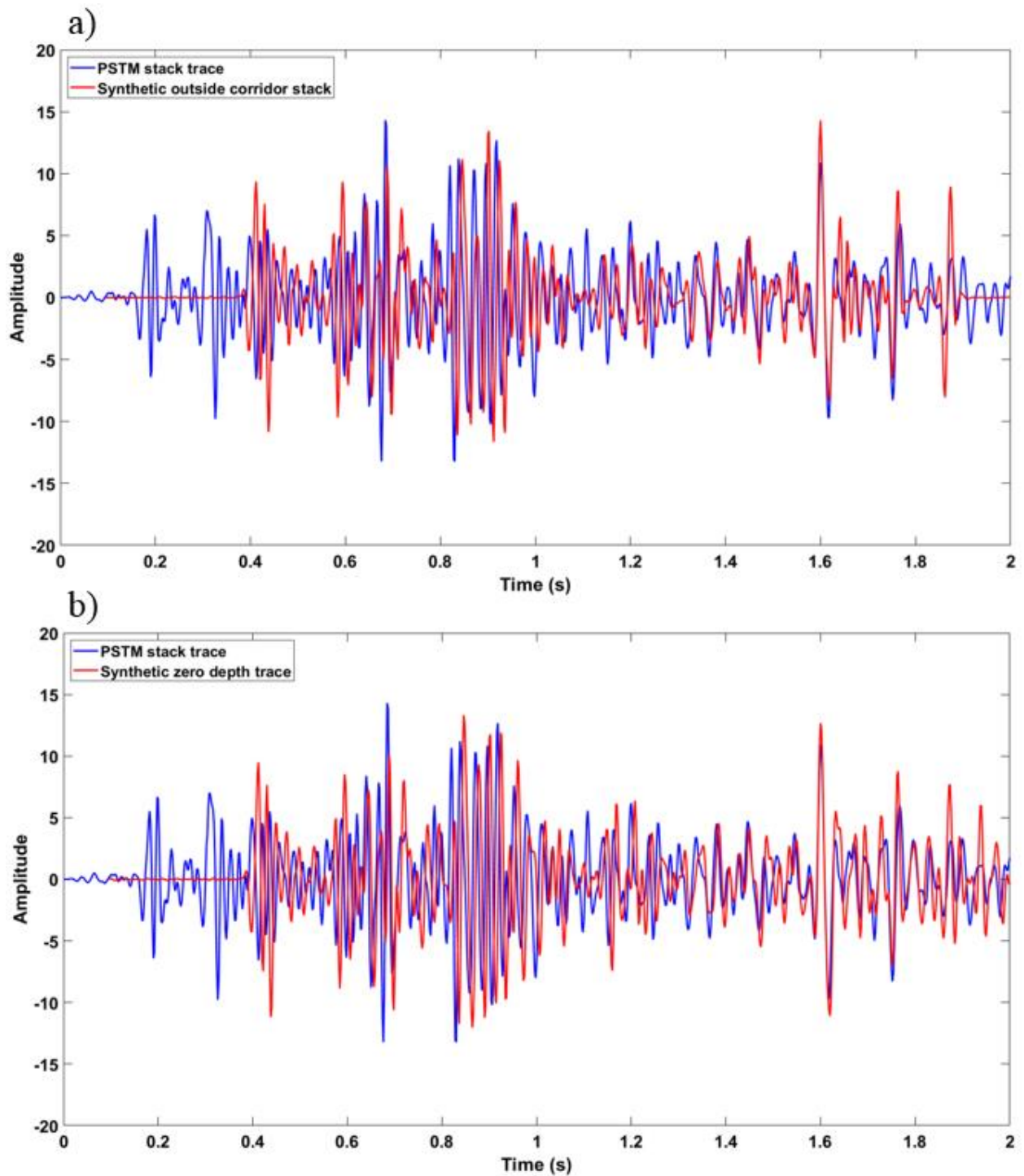


Figure 6.12 a) Synthetic outside corridor stack (Primaries) and trace from 3D PSTM Stack, b) Synthetic zero depth trace (Primaries and multiples) and trace from 3D PSTM Stack

Due to the high levels of noise in the VSP at shallow depths the synthetic VSP data was used to tie to the stack data. In a gross sense both the outside corridor stack and the zero depth trace match the stacked reflection data. This suggests the issue with the real data is due to internal multiples as that is all that was modeled in the synthetic case. Due to the similarity between the synthetic with multiples and the stacked trace this gives confidence that inverse scatter series method may be able to assist. There appears to be some stretch required to match the synthetic exactly and this was also seen in the checkshot survey. To match the real data to the synthetic the focus was deeper in the section. This stretch or temporal drift is possibly due to the frequency content differences between the tool recording the sonic log data and the vibroseis seismic data. Deeper in the section at approximately 1.8 seconds there is an improved amplitude match to the zero depth trace which contains both primaries and internal multiples, relative to the outside corridor stack.

6.3.2 Internal multiple prediction

Applying the multiple attenuation method to the stacked data becomes more difficult to judge the success of the algorithm. With the both the real and synthetic VSP or any other synthetic tests there are simple methods to isolate the trace which contains only internal multiples. With the stacked data this internal multiple trace does not exist. The inverse scattering series internal multiple prediction algorithm is used with an epsilon value of 15. The prediction is displayed and overlain with the input stacked data trace (Figure 6.13).

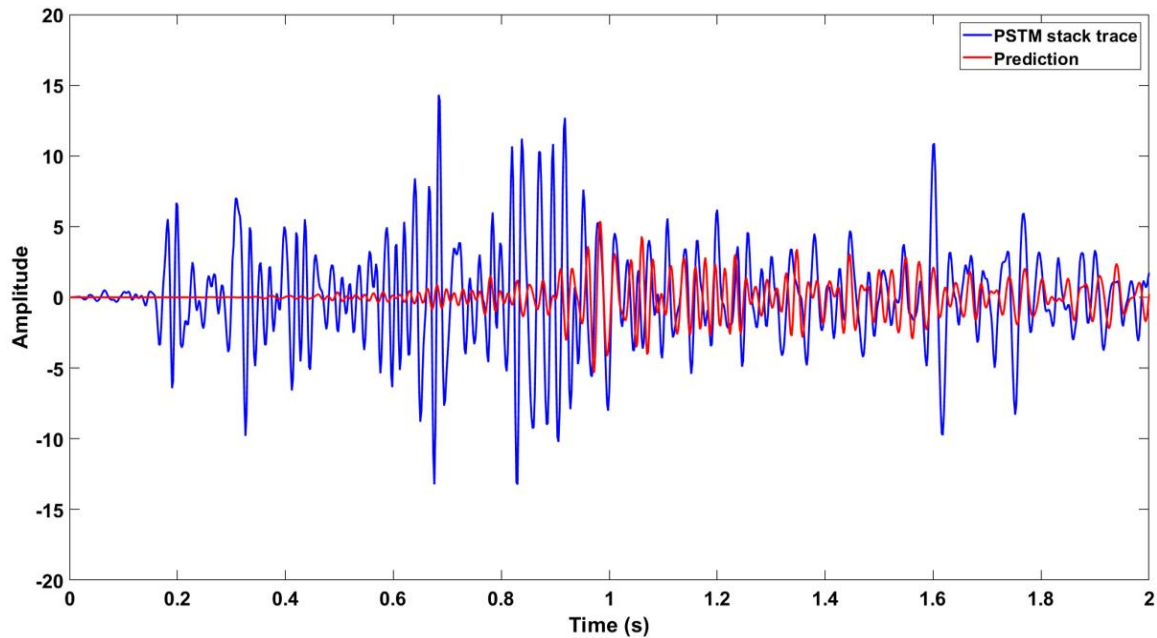


Figure 6.13 Internal multiple prediction and trace from 3D PSTM Stack

The amplitudes are scaled to approximately match those of the input data. The trace appears to largely overlay the input data with similar character. As seen previously on the synthetic this may be encouraging as the impact the multiples have on the data is largely amplitude adjustments. Displayed is the 2D Downward generator space for the prediction after stacking. Again, even after stacking as the 2D generator space is viewed the resulting prediction in this space is complex (Figure 6.14).

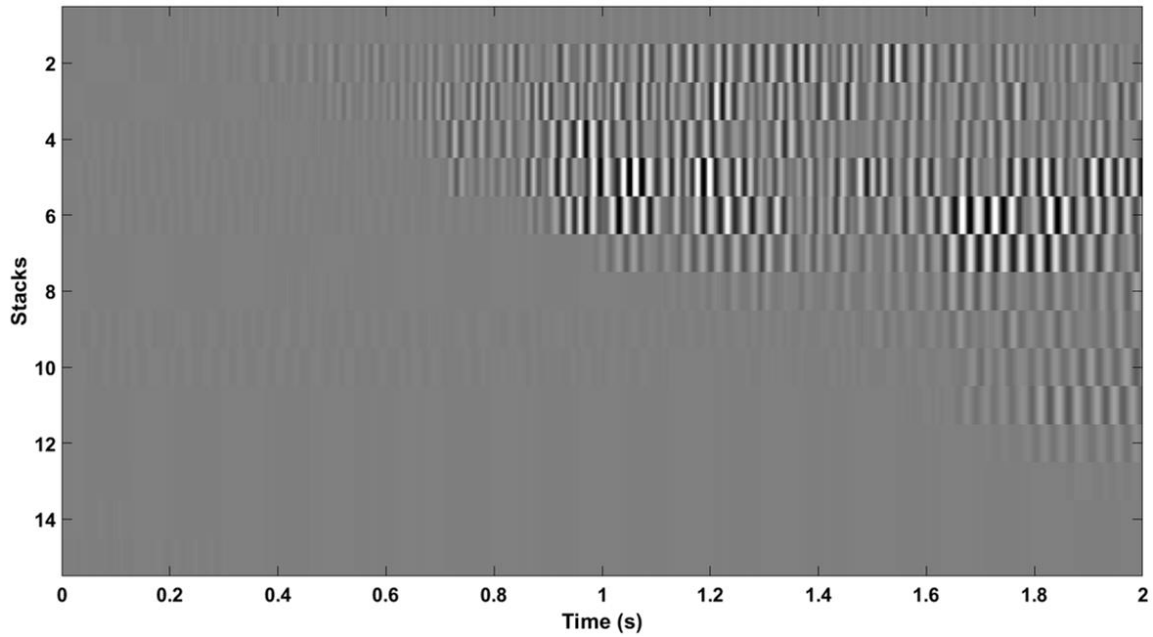


Figure 6.14 Downward generator space after stacking for 2D adaptive subtraction

The outside corridor stack from the synthetic VSP is used to judge the success of the algorithm. This is compared to the primaries only synthetic to see if the attenuated trace is better tied to the synthetic with no multiples (Figure 6.15).

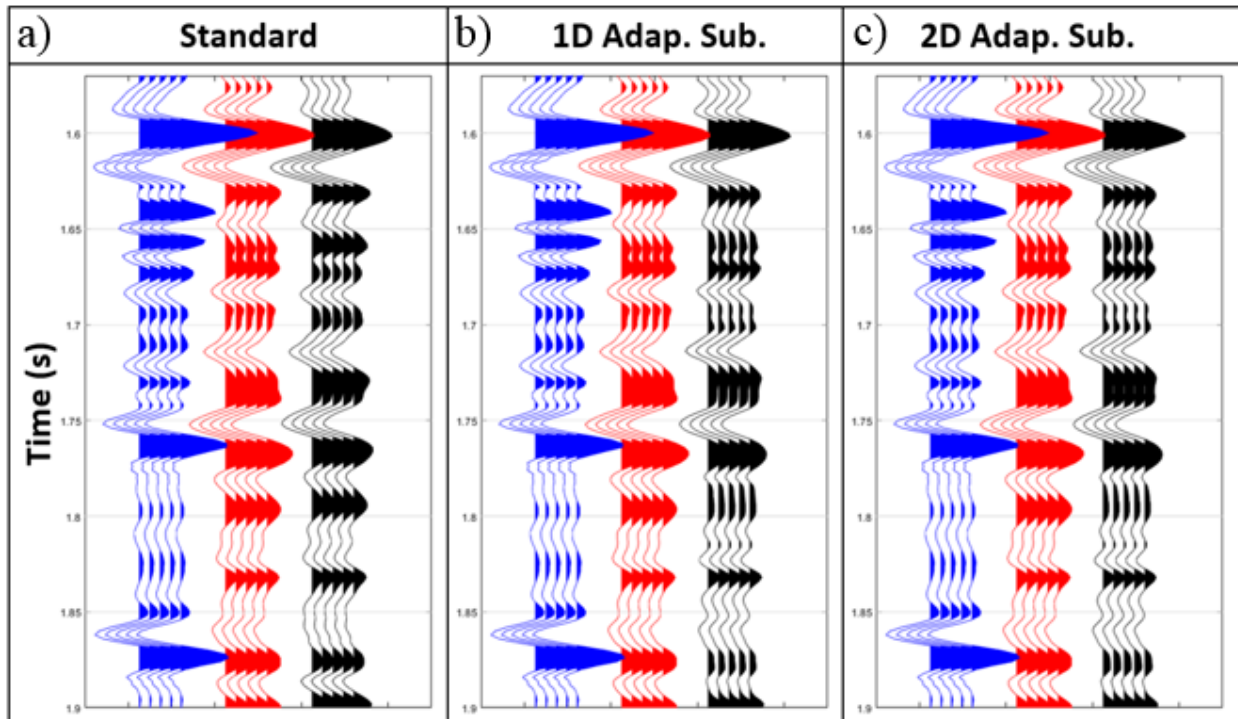


Figure 6.15 Outside corridor stack (primaries) in blue, PSTM trace (primaries and multiples) in red and PSTM trace after internal multiple attenuation in black for a) Standard subtraction, b) 1D adaptive subtraction, c) 2D adaptive subtraction

The results from the stack are also difficult to determine the success of the method in attenuating multiples. The results are also highly sensitive to the stack size used in the 2D adaptive subtraction. It was shown that there is significant similarity between the synthetic zero depth trace which contained primaries and multiple and the tie to the PSTM stack trace. There were encouraging results from this initial synthetic test as it appeared to display the ability to attenuate multiples with the inverse scattering series method for this dataset. With the decreased success on the real data this may display just how critical some of the amplitude differences between the synthetic input trace and the PSTM trace to the algorithm.

6.3.3 Crossline internal multiple prediction

Though the success at the control point is mixed, the internal multiple prediction method was then applied to a crossline from the data to compare the result on seismic section (Figure 6.16). The input data is relatively flat which can make it difficult to distinguish primaries from multiples as they will have similar structural patterns. If the structure varied greatly with time then this could aid in distinguishing primaries from multiples. In this case with the given flat data the resulting multiples will also be relatively flat. This is where the amplitudes of the events and how the multiples are attenuated from the data becomes critical.

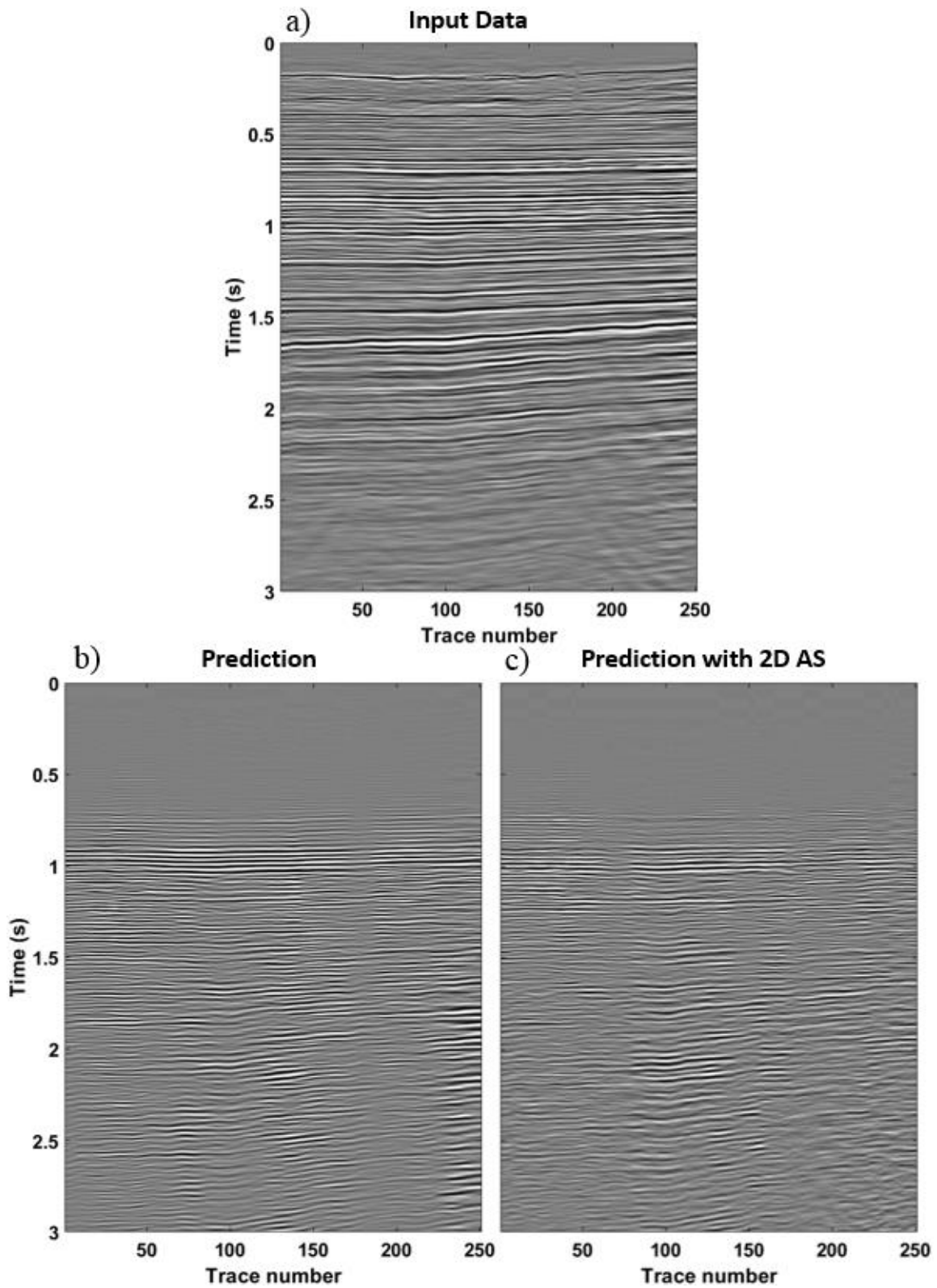


Figure 6.16 Crossline displaying 3D PSTM stack data a) input data, b) internal multiple prediction and c) internal multiple prediction with 2D adaptive subtraction

Displayed is the result of attenuating the internal multiples with the 2D adaptive subtraction in the downward generator space (Figure 6.17).

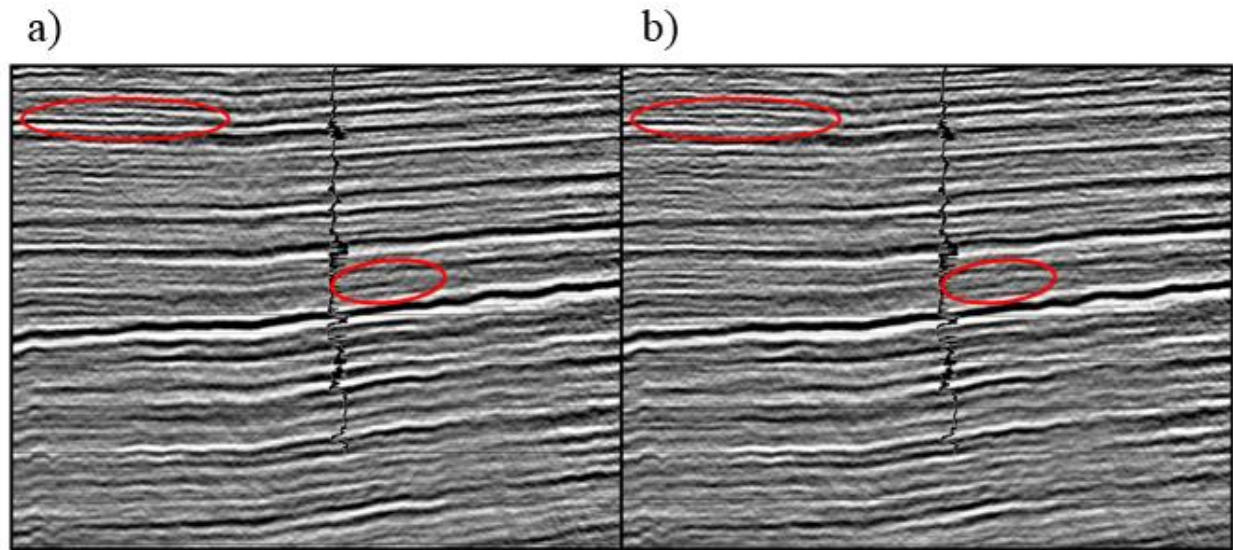


Figure 6.17 a) Crossline through PSTM stack. b) Crossline through PSTM stack after internal multiple attenuation, with red ovals highlighting significant areas of change due to internal multiple attenuation.

The results display minimal variation between the two but there are some small changes in coherency through the crossline. The well tie is compared after the internal multiple attenuation (Figure 6.18).

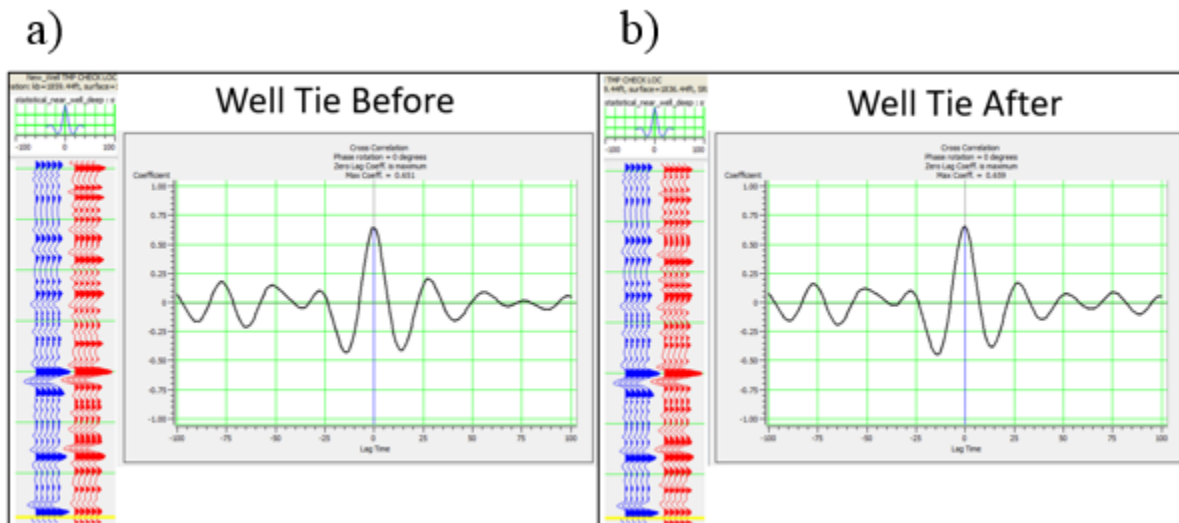


Figure 6.18 Well tie cross correlation a) before and b) after internal multiple attenuation

Comparing the cross correlation pre and post multiple attenuation the results show a negligible change in the tie. Due to the minimal quantitative differences a qualitative comparison is needed to see if there are improvements in space.

6.4 3D PSTM prestack seismic data

Three sets of prestack gathers were made available to test the method. The pre-interpolation, post interpolation and the migrated data. Displayed is the pre-interpolation data (Figure 6.19). This has irregular sampling which can impact which depending on the domain used for internal multiple prediction will impact the results.

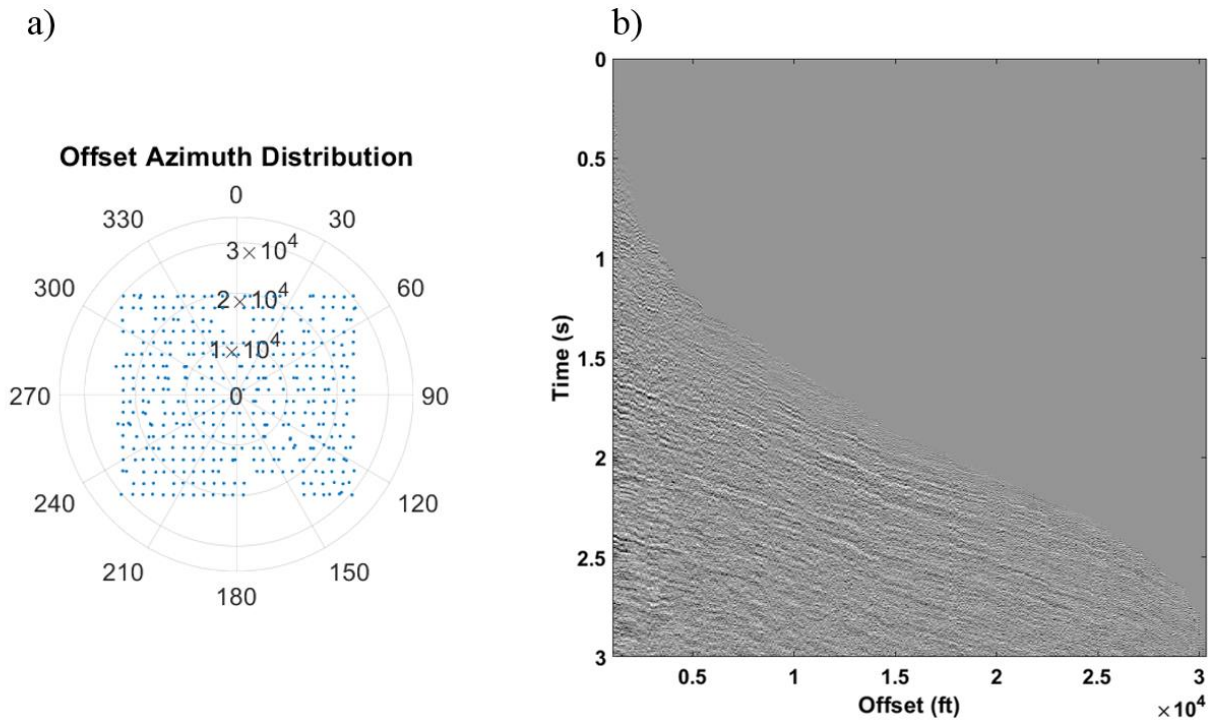


Figure 6.19 a) offset (ft) and azimuth acquisition distribution for the given CMP, b) seismic gather with the displayed acquisition before 5D interpolation

After 5D interpolation offsets are regularized and there is a significant improvement to the signal to noise ratio in the data (Figure 6.20). This will aid in the algorithm using these events and the increased signal to noise to produce and improved result. The regular sampling also assists the algorithm as the method expects smoothly varying wavefronts.

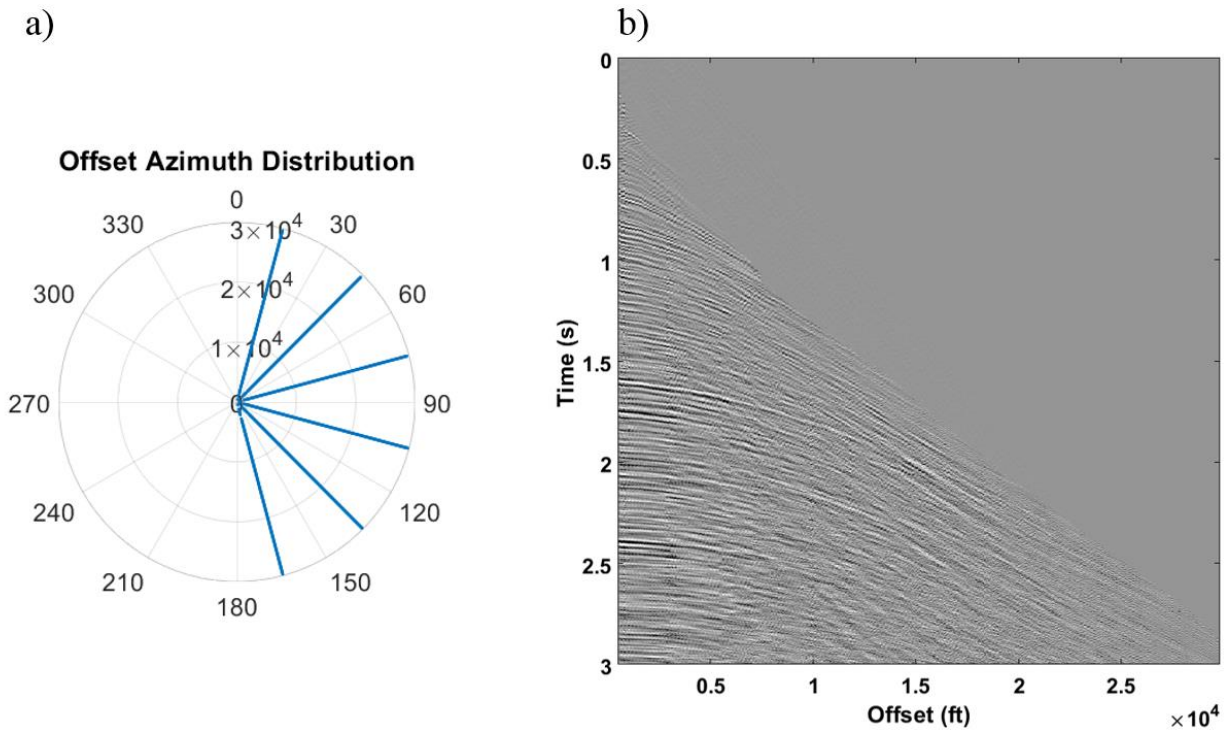


Figure 6.20 a) offset (ft) and azimuth acquisition distribution for the given CMP after 5D interpolation, b) seismic gather with the displayed acquisition after 5D interpolation

6.4.1 Time offset internal multiple prediction

The time offset version of the internal multiple prediction algorithm is used on the prestack data. As displayed in the Chapter 4.5.1 the algorithm requires a split spread input. If the input data only contains positive offsets, then the result will contain artifacts. The data is copied about the zero offset to give both positive and negative offsets prior to input to the internal multiple prediction algorithm. Displayed is the input and prediction of the internal multiples using the time offset version of the algorithm (Figure 6.21).

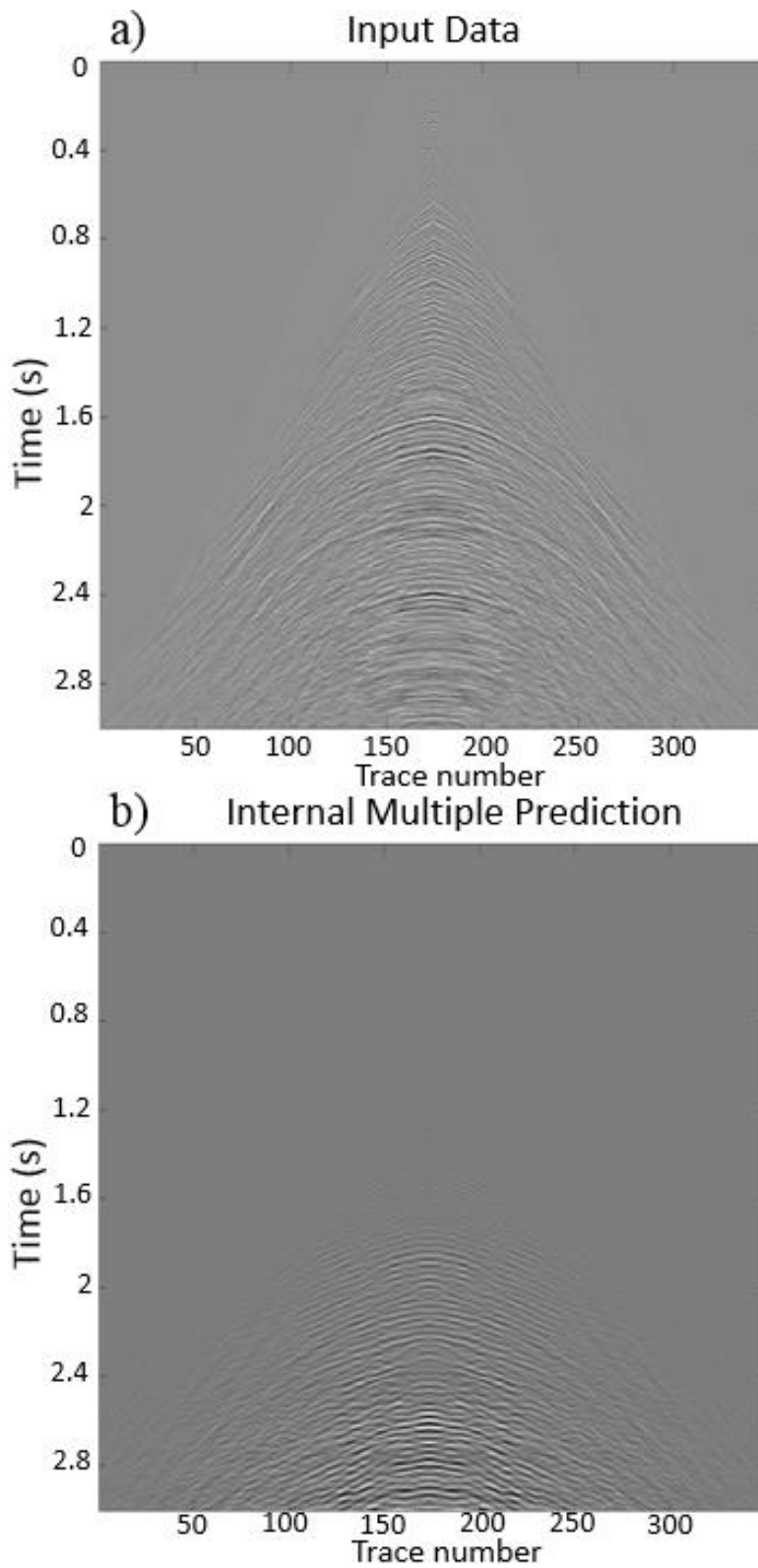


Figure 6.21 a) seismic gather after 5D interpolation, b) Time offset internal multiple prediction of seismic gather after 5D interpolation

There are numerous distinct events that can be seen in the prediction. There also appears to be linear events that have been predicted which can be seen in the input data. The prestack data may require further processing prior to the application of the internal multiple prediction algorithm.

6.5 Conclusions

These real data tests of the method have been more difficult to determine their success. Most of the real data cases displayed some locations which improved due to the attenuation of multiples with other locations where it had possibly been detrimental. The internal multiple attenuation results from the synthetic testing were encouraging. The 3D PSTM stack tied the zero depth trace synthetic reasonably well, but subtle differences between the two remained. The question that is raised is the possibility of additional processing that may assist in the prediction of internal multiples. The method assumes that the only events that remain in the data are primaries and internal multiples. It requires that these are both part of the input data as they will be used to predict events and the multiples must also be present for the subtraction. Has the migration or other processing steps altered the multiples in the data. If there is noise or other events in the data, these events will be used by the algorithm to predict the relevant internal multiples. This potentially demonstrates how critical preserving amplitudes through processing can be to assist these data driven methods. The next steps are to determine how to achieve this success with real data.

Chapter Seven: Conclusions

7.1 Conclusions

The inverse scattering series for internal multiple attenuation has been implemented with few assumptions and required inputs. It has shown promise as a wave equation based multiple attenuation method. The method is implemented in two parts, first part is the prediction and second part is the adaptive subtraction step to remove the multiples. The issues addressed in this work are to improve the initial prediction and reduce the strain on the adaptive subtraction to correct for errors and artifacts and achieve a more reliable result. This was accomplished through the development of the following tools and workflows.

The main tools used and developed: first, the downward generating space, which can assist in accounting for any losses at the downward generator level, and corrections needed to apply it. This also allows for an improved understanding of the generated multiples, as the uncollapsed multiples from each generator can be examined. Second, the use of higher order terms, which has been previously shown to assist but what was revisited in this work as an outline of how to use the higher order terms to improve the prediction. This was done in combination with the downward generator space and 2D adaptive subtraction. Third, a relatively new domain was tested, which show how a nonstationary epsilon can improve the result of the prediction and displayed its abilities to function in the case of irregular sampling. The combined impact of these results is an improved prediction capability and new Q.C. tools to evaluate the prediction algorithm.

These tools were applied to a variety of synthetic and real data examples and showed significant improvement on the simple synthetic cases. They displayed how correcting for these small amplitude mismatches can lead to increased seismic resolution in seismic data contaminated with internal multiples. Even on the more complex synthetics, it was shown how the 2D adaptive subtraction can improve the result in some instances. The real data tests were less conclusive.

7.2 Recommendations for future work

As is often the case, this work has potentially raised more questions than it has solved. The recommendations for future work are broken into two groups.

First, is to better understand the nature of internal multiples. It was shown how when multiples and primaries are sufficiently separated there can be a very accurate result and a high level of success. Once there are more short path multiples, this starts to cause errors in the prediction. I believe a more detailed method of how to determine the success of the method is required. There may be cases where it is believed that the method is unsuccessful, but it is possible that we have set the method up to fail. Some tests may be asking the method to resolve multiples that the inverse scattering series is not designed for. Future work should include how to successfully merge solutions which attack both short period internal multiples possibly through deconvolution like methods and long path internal multiples.

Second, to determine what must be done to the data to prepare it for this type of multiple removal. Though some of the issues were touched on including sampling and noise, a more

rigorous study on the application of the method and how the processing steps impact the result is warranted.

References

- Berkhout, A. J. and Verschuur, D. J., 1997. Estimation of multiple scattering by iterative inversion; Part 1, Theoretical considerations. *Geophysics*, 62(5), pp. 1586-1595.
- Coates, R. T., and Weglein, A. B., 1996. Internal multiple attenuation using inverse scattering: results from prestack 1D & 2D elastic synthetics. *SEG Technical Program Expanded Abstracts 1996*, pp 1522-1525.
- Cova, R., Innanen, K. and Rauch-Davies, M., 2018,.Walkaway VSP data conditioning for FWI. *CREWES Research Reprot*, Volume 30.
- Eaid, M., Sun, J., Keating, S., and Innanen, K. A., 2016. 1D and 1.5D internal multiple prediction in MatLab. *CREWES Annual Report*, 28.
- Groenestijn, G. J. A., and Verschuur, D. J., 2009. Estimating primaries by sparse inversion and application to near-offset data reconstruction. *Geophysics*, 74(3), A23-A26.
- Guitton, A., and Verschuur, D., 2004. Adaptive subtraction of multiples using the l1-norm. *Geophysical Prospecting*, 52.
- Hilterman, F., Nicholson, F. and Chen, Q., 2018. Seismic interpretation when short-period internal multiples are present. *The Leading Edge*, 37(1), pp. 19-26.
- Innanen, K. A., and Pan, P., 2014. Large dip artifacts in 1.5D internal multiple prediction and their mitigation. *CREWES Annual Report*, 26.
- Innanen, K. A., 2015a. A nonstationary search parameter for internal multiple prediction. *CREWES Annual Report*, 27.
- Innanen, K. A., 2015b. Time domain internal multiple prediction. *CREWES Annual Report*, 27.
- Innanen, K. A., 2017. Time- and offset-domain internal multiple prediction with nonstationary parameters. *Geophysics*, 82(2), V105-V116.

- Iverson, A., 2014. The impact of interbed multiples on the inversion and interpretation of prestack data. *CSEG Recorder*, p. 39(1).
- Iverson, A. and Innanen, K. A., 2017. Inverse scattering series internal multiple prediction with depth dependent scalars. *CREWES Annual Report*, p. 29.
- Iverson, A., Innanen, K. A., and Trad D. O., 2017. Internal multiple prediction in the time and offset domains. *CREWES Annual Report*, p. 29.
- Iverson, A., Innanen, K. A., and Trad, D., 2018a. Internal multiple prediction and subtraction: well log synthetic. *CREWES Annual Report*, p. 30.
- Iverson, A., Innanen, K. and Trad, D., 2018b. Internal multiple prediction in the offset-time domain: Response to irregular spatial sampling. *SEG Technical Program Expanded Abstracts*, pp. 4538-4542.
- Iverson, A., Innanen, K., and Trad D.O., 2018c. Internal multiple prediction in the time and offset domains. *Geoconvention 2018 Expanded Abstracts*.
- Iverson, A., Innanen, K. A., Trad, D. and Rauch-Davies, M., 2018d. Internal multiple prediction and subtraction: VSP, pre-and post-stack seismic data examples. *CREWES Annual Report*, p. 30.
- Iverson, A., Keating, S., Innanen, K. A. and Trad, D., 2018e. Internal multiple prediction with higher order terms and a new subtraction domain. *CREWES Annual Report*, p. 30.
- Iverson, A., Keating, S., Innanen, K. and D.O., T., 2019. Internal multiple prediction and the downward generator space. *Geoconvention 2019 Expanded Abstracts*.
- Kabir, N. M. M. and Marfurt, J. J., 1999. Toward true amplitude multiple removal. *The Leading Edge*, 18(1), pp. 66-73.
- Keating, S., Sun, J., Pan, P., and Innanen, K. A., 2015. Nonstationary 11 adaptive subtraction with application to internal multiple attenuation. *CREWES Annual Report*, 27.

- Liang, H., Ma, C., and Weglein, A. B., 2013. General theory for accommodating primaries and multiples in internal multiple algorithm: Analysis and numerical tests. *SEG Technical Program Expanded Abstracts 2013*, pp 4178-4183.
- Lv, S. and Innanen, K., 2016. Towards characterization of intrinsic and stratigraphic Q in VSP data with information measures. *CREWES Research Report*, Volume 28.
- Ma, C.*, and Weglein, A. B., 2015. A new Inverse Scattering Series (ISS) internal-multiple-attenuation algorithm that predicts the accurate time and approximate amplitude of the first-order internal multiples and addresses spurious events: Analysis and Tests in 2D. *SEG Technical Program Expanded Abstracts 2015*, pp 4402-4407.
- Margrave, G. F., and Daley, P. F., 2014. VSP modelling in 1D with Q and buried source. *CREWES Annual Report*, 26.
- Pan, P. and Innanen, K. A., 2013. A review of internal multiple prediction. *CREWES Annual Report*, Volume 25.
- Peacock, K.L., and Treitel, S., 1969. Predictive Deconvolution: Theory and Practice. *Geophysics*, 34(2), pp. 155-169.
- Perez, M. A., and Henley, D. C., 2000. Multiple attenuation via predictive deconvolution in the radial domain. *CREWES Annual Report*, 12.
- Pica, A., and Delmas, L., 2008. Wave equation based internal multiple modeling in 3D. *SEG Technical Program Expanded Abstracts 2008*, pp 2476-2480.
- Ramirez, A. C., and Weglein, A. B., 2008. Inverse scattering internal multiple elimination: Leading-order and higher-order closed forms. *SEG Technical Program Expanded Abstracts 2008*, pp 2471-2475.
- Ramirez, A. C., 2013. Analysis of data-driven internal multiple prediction. *Journal of Seismic Exploration*, Volume 22, pp. 105-128.

SEG wiki, 2019. Dictionary.

Sun, J., and Innanen, K. A., 2014. 1.5D internal multiple prediction in the plane wave domain. *CREWES Annual Report*, 26.

Sun, J., and Innanen, K. A., 2016. Literature review and discussions of inverse scattering series on internal multiple prediction. *CREWES Annual Report*, 28.

Sun, J., Innanen, K., Trad, D., and Geng, Y., 2017a. Multicomponent inverse scattering series internal multiple prediction Part I: Analytical analysis of input preparation. *CREWES Annual Report*, 28.

Sun, J., Innanen, K., Trad, D., and Geng, Y., 2017b. Multicomponent inverse scattering series internal multiple prediction Part II: Synthetic application. *CREWES Annual Report*, 29.

Verschuur, D. J., Berkhout, A. J. and Wapenaar, C. P. A., 1992. Adaptive surface related multiple elimination. *Geophysics*, 57(9), pp. 1166-1177.

Weglein, A. B., Gasparotto, F. A., Carvalho, P. M., and Stolt, R. H., 1997. An inverse-scattering series method for attenuating multiples in seismic reflection data. *Geophysics*, 62(6), 1975–1989.

Widess, M. B., 1973. How thin is a thin bed. *Geophysics*, 38(6), pp. 1176-1180.

Wu, Z. J., and Dragoset, B., 2011. Robust internal multiple prediction algorithm. *SEG Technical Program Expanded Abstracts 2011*, pp. 3541-3545.

Xiao, C., Bancroft, J. C., Brown, J. R., and Cao, Z., 2003. Multiple suppression: A literature review. *CREWES Annual Report*, 15.

Zhang, H. and Shaw, S., 2010. 1D analytical analysis of higher order internal multiples predicted via the inverse scattering series based algorithm. *SEG Technical Program Expanded Abstracts 2010*, pp. 3493-3498.

Zhang, H., Skorinski, D., Evans, P., Chebihat, A. M. B., Stannard, K., Fothergill, P., and Gerami, H., 2018. Effective Multiple Attenuation and Depth Imaging with Thousands Well Ties for Montney Reservoir Characterization: A Western Canada Wembley Valhalla Land Case Study. *Geoconvention Exapnded Abstracts 2018*.

Zou, Y., and Weglein, A. B., 2015. An internal-multiple elimination algorithm for all first-order internal multiples for a 1D earth. *SEG Technical Program Expanded Abstracts 2015*, pp 4408-4412.

Zou, Y., Ma, C., and Weglein, A. B., 2016. The first inverse-scattering-series internal multiple elimination method for a multidimensional subsurface. *SEG Technical Program Expanded Abstracts 2016*, pp 4550-4554.