

IVGPR: A NEW PROGRAM FOR ADVANCED END-TO-END GPR PROCESSING

A Thesis

by

IVAN VAZQUEZ

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Chair of Committee, Mark E. Everett
Committee Members, David W. Sparks
Krzysztof A. Michalski
Head of Department, Michael C. Pope

December 2017

Major Subject: Geophysics

Copyright 2017 Ivan Vazquez

ABSTRACT

Ground penetrating radar (GPR) processing workflows commonly rely on techniques developed particularly for seismic reflection imaging. Although this practice has produced an abundance of reliable results, it is limited to basic applications. As the popularity of GPR continues to surge, a greater number of complex studies demand the use of routines that take into account the unique properties of GPR signals. Such is the case of surveys that examine the material properties of subsurface scatterers. The nature of these complicated tasks have created a demand for GPR-specific processing packages flexible enough to tackle new applications. Unlike seismic processing programs, however, GPR counterparts often afford only a limited amount of functionalities. This work produced a new GPR-specific processing package, dubbed IVGPR, that offers over 60 fully customizable procedures. This program was built using the modern Fortran programming language in combination with serial and parallel optimization practices that allow it to achieve high levels of performance. Within its many functions, IVGPR provides the rare opportunity to apply a three-dimensional single-component vector migration routine. This could be of great value for advanced workflows designed to develop and test new true-amplitude and inversion algorithms. Numerous examples given through this work demonstrate the effectiveness of key routines in IVGPR. Additionally, three case studies show end-to-end applications of this program to field records that produced satisfactory result well-suited interpretation.

DEDICATION

In dedication to Ivan and Dagmaris who instilled in me the will to persist until I
persevered.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Mark E. Everett for his patience, encouragement, and guidance throughout this pursuit. Not only was Dr. Everett a mentor to me as an undergraduate student, but he has been a key role model and major influence in my academic growth. In addition, I would like to thank Maria Pesek and her family for all of the help and love they have provided in the last couple of years. Akhil Amara, Patrick Wagner, and Daniel Elizondo have offer me great advice and a sincere friendship from the beginning of my graduate studies. Therefore, I also extend my gratitude to them. None of this would have been possible without the support and love of my father Ivan Vazquez. He is my greatest source of motivation and I would like to acknowledge him for all of his sacrifices.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Professor Mark E. Everett (advisor) and professor David W. Sparks of the Department of Geology and Geophysics along with Professor Krzysztof A. Michalski of the Department of Electrical & Computer Engineering.

The datasets analyzed in § 6.2 and § 6.3 were provided by Timothy S. de Smet and Akhil Amara, respectively. In addition, the data in §4.3.1 was obtained from an open-source repository in the USGS website and forms part of a study by Zaremba et al. that was published in (2016).

All other work conducted for the thesis was completed by the student independently.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

NOMENCLATURE

GPR	Ground penetrating radar
EM	Electromagnetic
A-scan	The display of a single GPR trace is commonly referred to as an A-scan
B-scan	The display of a multitude of adjacent traces.
IVGPR	Processing package for GPR data developed as part of this study
iGPRread	Graphical user interface used read and write encrypted GPR files.
iViz3D	Graphical user interface used displaying three-dimensional GPR datasets.
AGC	Automatic gain control
DR	Differential resolution
ATS	Average trace subtraction
MTS	Mean trace subtraction
ATM	Alpha-trimmed mean
TVSW	Time variant spectral whitening
FDSW	Frequency domain spectral whitening
GLCM	Gray-level co-occurrence matrix
BNS	Background noise suppression
FFT	Fast Fourier Transform
MSA	Multitaper spectral analysis

BMS	Background matrix subtraction
ER	Energy ratio
PSD	Power spectral density
CTA	Complex trace analysis

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
CONTRIBUTORS AND FUNDING SOURCES	v
NOMENCLATURE	vi
TABLE OF CONTENTS	viii
LIST OF FIGURES	xi
LIST OF TABLES	xix
1. INTRODUCTION	1
1.1 Motivation of research	2
1.2 Objectives	4
1.3 Research design	5
1.4 Thesis structure	6
2. THE IVGPR PROCESSING PACKAGE	8
2.1 Overview	8
2.2 Structure and data flow	8
2.3 Data visualization	10
2.3.1 MATLAB graphical interfaces	12
3. BASIC PROCESSING TECHNIQUES	14
3.1 Overview	14
3.2 Dewow	15
3.3 Filtering	19
3.3.1 The windowed-sinc filter	19
3.3.2 Basic smoothing filters	21
3.4 Time-zero correction	23

3.5	Time gains	26
3.5.1	Automatic gain control	26
3.5.2	Spherical and exponential compensation	30
3.5.3	Programmed gains	32
3.6	Desaturation	33
3.7	Background noise suppression	38
4.	ADVANCED PROCESSING TECHNIQUES	41
4.1	Overview	41
4.2	Image enhancement	41
4.2.1	Interpolating GPR data	42
4.2.2	Advanced filtering	43
4.3	Attribute analysis	48
4.3.1	Textural attributes	48
4.3.1.1	Uniformity, contrast, homogeneity, and entropy	50
4.3.2	Rapid assessment of structural and spectral content	51
4.3.3	Complex trace analysis	55
4.3.3.1	Envelope, instantaneous phase, and instantaneous frequency	57
4.4	Alternative techniques	60
4.4.1	Spectral Balancing	60
4.4.1.1	Time-variant spectral whitening	61
4.4.1.2	Frequency domain spectral whitening	64
4.4.1.3	Differential resolution	68
4.4.2	Background matrix subtraction	72
4.4.3	Attribute-based time-zero correction	76
4.4.3.1	Edge-preserving smoothing	79
5.	MIGRATION OF GPR DATA	82
5.1	Overview	82
5.2	Phase-shift migration	83
5.2.1	Downward continuation	84
5.2.2	Phase shift migration algorithm	89
5.3	Vector-migration	92
5.3.1	Algorithm and pseudocode	93
5.4	Performance analysis of migration algorithms	98
6.	CASE STUDIES	105
6.1	Overview	105
6.2	Case I: Geophysical archeology over the recreational yard of the main Alcatraz prison	105

6.3 Case II: Field record over exposed fault near Mason, Texas	110
7. CONCLUDING REMARKS AND FUTURE WORK.....	116
REFERENCES	117

LIST OF FIGURES

FIGURE	Page
1.1 Systematic approach used to design and test new modules.	5
2.1 iViz3D GUI. (a) Loads a set of GPR lines, (b) implements changes in the settings to the data, (c) sets the strength of the smoothing and the direction along which it is applied, (d) pads all lines to longest or trims all lines to shortest, (e) displays inlines, crosslines, and time slices, and (f) shows the current location within GPR cube, and (g) allows the user to navigate through the cube. Specific locations within the cube can be set in (f).	11
2.2 iGPRRead GUI. (a) Allows for the modification of headers, (b) displays the saved data, (c) flips the order of traces, (d) removes temporal gains applied during acquisition, (e) clips data edges, (f) specifies the input and output formats.	13
3.1 Fundamental processing steps for a bistatic GPR surveys. The orange box encloses a set of basic processing steps commonly applied at the site. These are usually enough when detection of a target is the primary goal. The gray box encloses offsite steps, which are typically applied in descending order.....	15
3.2 Example of wow noise in GPR trace and effect of dewow filter.	16
3.3 GPR field record showing broad horizontal bands caused by low frequency wow noise. The banding became prominent after applying a temporal gain.(b) Demonstrates the improvement obtained by applying a dewow filter to (a) prior to implementing the temporal gain.....	17
3.4 Average amplitude spectrum of GPR record. The region shaded red shows the portion of the spectrum removed by the filter. The filter used is also shown.	18
3.5 Illustrates a typical frequency filtering flow (similar to the one used in IVGPR).	20

3.6	Amplitude response of three bandpass windowed-sinc filter with identical cutoff frequencies (i.e. $f_{hp} = 300$ MHz and $f_{lp} = 1000$ MHz) but varying kernel sizes. This filters were created using the windowed-sinc routine in IVGPR.	22
3.7	Shows the product (b) obtained by smoothing a raw radargram (a) using a two-dimensional mean filter with a 5×5 window. (b) Shows that this kernel achieved a fair trade-off between resolution and smoothing.....	23
3.8	Four conventional locations for picking the time-zero point along the first reflected wavelet. Point A marks the first negative peak. Point B indicates the S-crossover. Point C indicates the first maximum. Point D corresponds to the Z-crossover.....	24
3.9	Shows a radargram before (a) and after (b) time-zero correction. The preferred location was set to be the first significant peak.	25
3.10	Shows the convolution of a finite duration pulse (left) with a reflectivity series (red). The resulting time series was attenuated using a linear decay down to 10% of the original amplitude. Without a gain, an interpreter could mistakenly deduce that the contrast in material properties varies with depth.	27
3.11	Scheme used to compute the gaining coefficients for the instantaneous and rms AGC. The original time series is depicted as a series of points sectioned into three parts (green, blue, and purple). The rms AGC algorithm computes the coefficients at the center of each time gate (box) without overlap. The instantaneous AGC computes the gain at the center of overlapping windows, which are shown in a stair case pattern for clarity. The yellow section indicates the region where values are extrapolated in both techniques.....	29
3.12	(a) Shows the input B-scan in raw form. Reflectors are clearly discernible only within 5 ns from the ground surface. (b) shows the output from the rms AGC routine. This greatly enhanced the number of visible features. A similar effect is obtained after instantaneous AGC (c) and robust AGC (d). The main differences between the tree techniques is the effect on relative lateral amplitudes and structural resolution.	31
3.13	Illustrates how the amplitude decay in a raw scan (a) can be ameliorated by applying the SEC algorithm. (b) The B-scan after SEC. A red arrow indicates where a deep feature of interest became apparent after the gain was applied.	32

3.14	(a) Workflow for programmed gains. (b) Raw B-scan affected by attenuation. (c) The B-scan after being processed by a programmed gain without smoothing.	34
3.15	Shows the manifestation of signal saturation along a GPR trace. The dashed red lines indicate the limits of the dynamic range for the receiver system. Clipped section appear as horizontal plateaus collinear with the limits of the dynamic range.	35
3.16	Interpolation results using a fourth-, fifth-, and sixth- order polynomial interpolation along with a cubic spline to estimate the amplitudes along a clipped section. Each interpolation was compared to the actual amplitudes using a percent error analysis. The number of neighboring points used were 5 (a and b), 9 (c and d), and 13 (e and f).	37
3.17	(a) Radargram showing extensive clipping. (b) Output from IVGPR's de-clipping routine with a fourth-order polynomial interpolation.	38
3.18	Shows a field record before (a) and after (b) it was treated by the ATS algorithm. The background noise matrix was estimated using a windows of 101 elements. A red square encloses the inferred space-time confinement of a hole filled moments before acquiring the data. The blue square surrounds events that show enhanced continuity applying the ATS algorithm.	40
4.1	Shows a bilinear interpolation of a raw GPR slice (a) originally composed of 29 lines. The interpolated slice (b) contains 290 lines. No obvious evidence of damaging artifacts can be observed for this time slice.	43
4.2	Extending the definition of cutoff frequency k_c as the perimeter of an ellipse.	45
4.3	Two-dimensional Butterworth transfer function (a) and a two-dimensional Tukey window (b). The same cutoff frequency was chosen along each axial direction resulting in the isotropic response shown.	46
4.4	Raw radagram (a) treated by the Wiener filter denoising procedure in IVGPR using an 11×11 window. Two boxes are use to highlight how this filter can improve signal-to-noise without deteriorating the resolution of the image. A smaller window size may be used to retain more detail.	47

4.5	Approach used to compute textural attributes. A moving window of size M by M is centered at each pixel of an image and the resulting sub-images are quantized to a prescribed gray level. The figures on the right show a sub-image quantized to 4 gray levels. Next, a GLCM is computed for every direction. The images on the left show how the 0° GLCM is derived. The resulting matrices are summed and the sum is normalized to produce C° . Textural attributes can then be easily computed from C°	49
4.6	(a) Uniformity. (b) Contrast. (c) Homogeneity. (d) Entropy. One can easily discern regions of potential interest (with high and low values) highlighted by these attributes. There are a number of locations where these attributes reinforce one another.	52
4.7	Shows a 2D multitaper spectral estimate for the entire radagram (a) and a 1D implementation for a synthetic trace created by combining three sinusoids with frequencies of 7.8, 15.6, and 125 MHz.	54
4.8	(a) Shows a raw radagram along with the extracted gradient amplitude (b) and direction (c) attributes.	55
4.9	Shows the complex trace analysis attributes included in IVGPR. These consists of envelope (b), instantaneous phase (c) and instantaneous frequency. The attributes were extracted from the raw B-scan shown in (a). The amplitudes for each plot was normalize by the absolute maximum. Additionally, the instantaneous frequency was treated using a Plank window to reduce the effect of high amplitudes at the edges of the data.	59
4.10	Time-variant spectral balancing algorithm. The input data is Fourier transformed (red arrows) and filtered. This produces N sub-records with different frequency content. After an inverse Fourier transformation (blue arrows), each sub-record is gained using a programmed gain. The gained sub-records are combined to produced the TVSW data.	62
4.11	Three adjacent bandpass filters (dashed curves) generated by setting $f_1 = 50$ MHz, $f_2 = 650$ MHz, and $N = 3$. Their superstition (solid purple curve) is equivalent to a single bandpass filter with cutoff frequencies f_{hp} and f_{lp}	63
4.12	(a) Raw radagram with three squares indicating regions of different depth and characteristics. (b) Shows the result after treating the raw radagram with the TVSW. The results within the three regions were used to qualitatively asses the performance of this routine.	65

4.13	Amplitude spectrum of a GPR trace before (blue) and after (red) FDSW. The spectral response appears broader and its peak frequency has shifted to a higher value.	66
4.14	Identical sections of a profile before (left) and after (right) FDSW. Three depth regions are highlighted to evaluate the effectiveness of the routine. Clear enhancements to the continuity and resolution of reflectors can be observed after FDSW (right).....	68
4.15	Plot of the terms in equation 4.28. The sporadic behavior of high order derivative is likely due to noise in the original trace.	70
4.16	(a) Shows the output trace (blue) generated after treating the original trace (orange) with DR. The power spectrum of both traces is shown in (b). Figures (c) and (d) show a GPR section before and after DR, respectively. Three boxes enclose different depth ranges used to evaluate the technique. A red arrow indicates improvements in resolution.....	71
4.17	Shows three background matrices corresponding to: (a) mean trace subtraction (MTS), (b) alpha-trimmed mean trace subtraction (ATM), and (c) background matrix subtraction (BMS). A dashed square shows an instance where MTS appears to overestimate the background noise. An arrow shows one of the sharp transitions peculiar to the BMS technique.	75
4.18	Automatic first-arrival picking workflow. Processes colored in red indicate steps that differ from Coppen’s method.	76
4.19	Shows the Time-zero correction of a raw B-scan (A) after applying the peak-finding based method (B), the energy-ratio method (C), and the entropy method (D).	79
4.20	Average trace (black curve) showing the time-zero picks that would result from the entropy method (A), energy-ratio method (B), and the peak-finding method (C–F). The red and blue dashed line shows the average absolute value for the first derivative of entropy and energy-ratio attributes, respectively.	79
4.21	(a) Original step function. (b) Step function after random Gaussian noise was added. (c) Noisy data treated by conventional moving average filter using a window size of 17 elements. (d) Noisy data treated by EPS using a window size of 17 elements.	81

5.1	Illustrates the effect of data collection and migration. In particular, a point scatterer in object (or physical) space becomes mapped as a hyperbolic event in data space.....	83
5.2	Illustrates how a hyperbolic event becomes more compact as the array of receivers gets closer to its source. The white circles represent wavefronts at the time they first reach the receivers. The dotted lines indicate the distance a wave front needs to travel to reach the edge receivers after it arrives at the middle receiver. On the right, the resulting hyperbolic event within the acquisition time (indicated by a gray band) are shown.	85
5.3	Up-going plane wave propagating at some angle θ with respect to the vertical axis. The signal reaching two stations (red and blue dots) is shown as stem plots on the right.....	87
5.4	Contour lines for the equiphase planes corresponding to the maximum value of a monochromatic plane wave. The propagation direction of the wave is indicated with a white line such that it makes an angle θ with respect to the vertical axis.....	88
5.5	Shows artifacts caused by the intrinsic periodic boundary conditions of the phase shift migration algorithm.....	91
5.6	Shows the vector migration workflow proposed by Streich et al. (2007). The steps shaded in gray indicate the FFT-based technique for finding \hat{G} . The outputs from a 100 MHz monostatic antenna configuration at a depth of $z = 5$ m are shown for different steps. Oversampling factors of $p_x = p_y = 4$ were used on a spatial (x, y) grid composed of 512 by 512 elements separated by a distance of 0.1 m in both directions.....	94
5.7	Method for computing \tilde{F} (right-hand side) from the Hadamard, or element-wise, product of the Tukey window and \tilde{D}^{-1} , which is found using equation 5.16. The Tuckey window shown was found using a transition width of $\mu = 1$, see §4.2.2.	97
5.8	Subsurface model of a kink-style syncline. A total of eight geometrical shapes and four unique materials were used to construct the model. These create a sand layer overlaying a folded layer composed of three clay section varying only in permittivity.	100
5.9	Synthetic B-scan of kink-style syncline after undergoing an instantaneous AGC gain. The expected bow-tie pattern can be seen at the center. Two green arrows point to a diffraction and a multiple within the B-scan.	101

5.10	Shows the raw synthetic radargram (a) next to the one migrated using the phase-shift algorithm (b). A velocity of 0.15 m/ns was used for the migration. A separable two-dimensional Plank-tapering window was applied to artifacts produced at the edges of the data. The same migration without tapering is shown in (c), which resulted in the manifestation of strong artifacts.....	102
5.11	GPR line treated by the vector migration algorithm using $\epsilon_r = 4$, $p_x = 8$, and $p_y = 16$. Each original depth slice contained 160×64 elements. A total of 1733 depth slices were migrated using 38 frequency components from 10 MHz to 400 MHz.	103
5.12	Parallel performance analysis for the single-component 3D exact radiation pattern vector migration routine in IVGPR. The run time (a) shows a quadratic decrease with number of threads activated while the speedup (b) was linear with a slope of 0.77.	103
5.13	Vector migration of field data. The migrated line (b) shows artifacts likely due to aliasing. The corresponding unmigrated line is shown in(a).	104
6.1	Workflow proposed by de Smet (2016).	106
6.2	Outputs from a typical workflow implemented in IVGPR. (a) shows the raw line. (b) shows the migrated data after AGC. (c) and (d) show the energy and instantaneous attributes respectively.	108
6.3	New workflow proposed to produce improved results on this dataset.	109
6.4	(a) Gradient direction attribute used to reveal the underlying geometry and to assess the level of noise in the image. The dashed line separates a the regions of high and low signal-to-noise ratios. (b) Shows the same B-scan after applying the 2D adaptive denoising routine in IVGPR. The coherency of events for this image appears to be much improved, which corresponds to a lower number of randomly oriented events.....	110
6.5	Line 32(a) and line 58(b) from the study conducted over the recreational yard of Alcatraz prison island. Two red arrows indicate the inferred location of underground tunnels. A red square encloses a region of anomalously low amplitudes.	111

6.6	(a) Map showing the recreational yard of the main prison in Alcatraz island. The green and red lines indicate the position of lines 32 and 58, repetitively. A rendering of the believed architecture of the military fortification that predates the prison can be seen overlaying the map. Line 32 traverses a tunnel near its mid-section while line 58 traverses a tunnel nearly 10 meters inline. (b) Shows the locations of the lines with respect to a depth slice from the original publication by de Smet (2016). Modified for this work with the consent of Mark E. Everett and Timothy S. de Smet..	112
6.7	Shows the location of a high-angle fault traversed by the GPR line A to A'. The line was acquired in the northward direction while the fault dips to the south-east.	113
6.8	Line A after a low-cut dewow and a programmed gain were applied in order to highlight strong reflectors masked by the direct arrival. A dashed box encloses a cluster of hyperbolic events initially believed to be caused by a local fault.	114
6.9	(a) Shows the B-scan produced using IVGPR after implementing the workflow proposed on the previous section. A number of visible fractures were mapped. Those marked with orange curves correspond to picks matching the ones made by Amara (2016). Newly interpreted picks are indicated by green curves and the inferred location of the large normal fault is shown by the yellow curve. Two red arrows indicate regions where the data processed by Amara (2016) (b) and that processed by the author (a) show a pair of peculiar features in good agreement.	115

LIST OF TABLES

TABLE	Page
2.1 Modules with most of their available routines and techniques. Those marked with a * contain one or more parallel routines. Any routine containing the symbol ‡ can be applied to one- or two-dimensional data. The symbol † indicates advanced techniques.	9

1. INTRODUCTION

The use of ground penetrating radar (GPR) systems for the detection and characterization of subsurface targets has become a highly sought-after service in a wide range of fields. This flexible yet powerful non-invasive profiling technique offers a way, even for novice users, to obtain high resolution images. Prior to the proliferation of GPR-specific data processing routines, the abundance of seismic equivalents offered an attractive alternative (Annan, 1999; Everett, 2013). Today, it remains common for practitioners to rely on processing and imaging techniques intended specifically for seismic data (Cassidy and Jol, 2009; van der Kruk, 2015). Although seismic processing tools are robust and can generate interpretable results, they are not designed to address GPR-specific noise. Hence, the resulting images may contain artifacts that are indistinguishable from *bona fide* geological responses. Furthermore, the seismic techniques are inadequate for studies aimed at gaining insight into the electromagnetic properties of the subsurface (van der Kruk, 2015). Then, the question to ask is: why has the reliance on seismic processing packages persisted despite potential inaccuracies? While exploring the answers to that question by querying a range of practitioners is outside the scope of this research, the following section highlights several shortcomings of modern GPR software that might provide a partial explanation. This is followed by a description of how IVGPR, an open-source GPR processing program that I developed as part of this thesis research, can overcome the key limitations found in the seismic counterparts. Primarily, IVGPR addresses the scarcity of freely available vector migration algorithms. The final three sections report the specific objectives, methodology, and structure of this thesis respectively.

1.1 Motivation of research

Readily available commercial GPR software tends to be highly priced and often works only with data formatted by acquisition tools from the same manufacturer. Recent quotes requested for RADAN, REFLEXW, GPR-SLICE, and EKKO_Project revealed prices ranging from 1000 to 11000 US dollars, excluding the cost of additional modules. All four software packages, while offering impressive visualization capabilities, contain a nearly identical set of basic processing routines. In particular, the complexity of the migration algorithms is limited to scalar or asymptotic methods that use far-field approximations. These algorithms do not account for the polarization of electric field and the radiation pattern of antennas in the near- and mid-field. As a result, the amplitudes of the images produced cannot be quantitatively related to the electromagnetic properties of targets in a straightforward manner (Streich and van der Kruk, 2007; van der Kruk, 2015). This limitation hinders the performance of advanced post-migration analyses that depend on accurate amplitude information (Sassen and Everett, 2009).

Some GPR applications require specific, or even novel, processing and imaging techniques that are not available in popular processing software packages. In these cases, researchers opting to develop their own codes might encounter a few obstacles. Firstly, since algorithm development tools within existing processing packages are rarely available, the data must be converted into a practical format. However, this conversion requires facing the laborious and sometimes unfeasible task of deciphering the binary configurations unique to each manufacturer's acquisition tools. An extensive web search failed to reveal any freely-available programs for this purpose. The work of Tzanis (2010), which produced MATgpr, provides a good platform for reading, visualizing, and processing GPR data at no additional cost to anyone with a MATLAB license. MATgpr is capable of reading data formatted as DZT-, DT1-, and SEG-Y-files. Yet, data can only be exported as one

of the binary structures that is readable by the program. In my study, the apparent scarcity of programs capable of reading and writing GPR data was addressed by developing the MATLAB interphase iGPRRead. This interface, like MATgpr, is capable of reading and displaying data in the most prominent GPR data formats. In contrast to MATgpr, iGPRRead can export data in such versatile formats as CSV and ASCII-files.

Secondly, researchers might find it difficult to analyze the combined effect of their own specialized routines when used along with the standard processing steps that are available in vendor software packages. This situation arises when source codes are not available, algorithms are vaguely explained, and/or vendors do not offer a way for users to incorporate new routines into their programs. Researchers could circumvent this obstacle by building external routines that involve repeatedly transferring data in and out of programs. Such options, however, are generally costly and impractical. By contrast, the program created in this work consists of a transparent and self-contained set of independent modules. The modular architecture allows users to easily modify, add, and maintain specialized processing routines. By also offering a collection of plotting functions, IVGPR eliminates the need to outsource data for visualization, thereby allowing researchers to focus on their primary scientific objectives.

Although the original aim of this study was to develop and explore the effectiveness of vector migration algorithms, the limitations described above imposed many practical obstacles to this aim, which motivated the creation of IVGPR. This program, however, was not created in order to rival any of the existing commercial processing packages. Rather, it simply provides a free and flexible platform for researchers to readily implement and test new specialized GPR processing routines.

1.2 Objectives

The central goal of this work is to develop an open-source GPR-specific processing package capable of implementing a vector migration algorithm. The advantages of this class of migration techniques have been well documented (Annan, 1999; Leuschen and Plumb, 2000; Slob, 2003; Sassen and Everett, 2009; Streich and van der Kruk, 2007; van der Kruk, 2015), but they are generally absent in the available processing software packages (Cassidy and Jol, 2009). As a result, analyses of their performance with field data remains sparse. An aim of this thesis is to provide new evidence of the practicality and applicability of these algorithms by evaluating their effectiveness treating synthetic and field data. Further, researchers using IVGPR in the future can generate additional contributions to the available body of work on these techniques.

A self-contained GPR processing package, IVGPR provides a library of widely used basic routines as well as implementations of modern approaches that have been shown to be effective throughout the literature. Thus, users have the rare opportunity to execute full processing workflows containing vector migration algorithms. This thesis investigates the effects on final GPR images of different workflows. This is done by varying important parameters and the order of processing routines within IVGPR. Sample outputs are presented alongside the theoretical development of the key routines. Practical examples are shown in a collection of case studies. Also, qualitative comparisons are made between images obtained using analogous workflows with IVGPR and vendor software packages.

A further goal of the thesis is to facilitate the future enhancement and expansion of IVGPR by adopting a modular design, providing the theoretical foundation of key processing routines, and striving for optimal performance. For the latter, any major processing routine with an estimated potential speedup ranging from two to four, based on Gustafson's law (Gustafson, 1988), was parallelized using OpenMP. A considerable effort

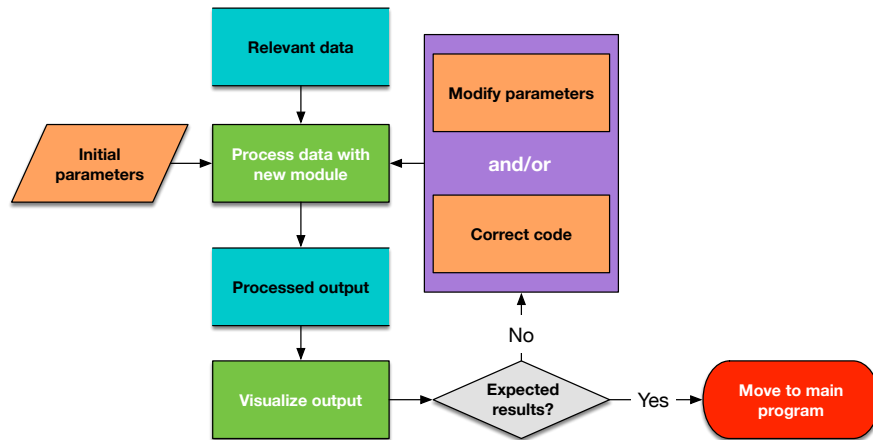


Figure 1.1: Systematic approach used to design and test new modules.

went into maintaining a consistent formatting and documentation style to the source code. These are just two of the several measures that I made to ensure efficient programming practice.

1.3 Research design

The first stage of the development of the IVGPR software package focused on including basic GPR processing routines (i.e. *data editing*, *dewow*, *time-zero correction*, *gain correction*, *background subtraction*, and *filtering*). Fortran was chosen as the programming language because of its advantages in performance and the availability of compilers at no charge. Data visualization within IVGPR was made possible through gnuplot in combination with a Fortran-based user interface I developed. To achieve additional clarity, two graphical interfaces for visualizing GPR data within MATLAB were also constructed. These, dubbed iGPRread and iViz3D, share no relation to the gnuplot interface. All three plotting interfaces will receive their proper description in a later section of this thesis. Figure 1.1 shows the systematic “minimal-working-example” workflow that I used to build and test each module independently.

The next stage of development was aimed towards the implementation and testing of two different migration techniques as well as a series of other advanced processing steps. The migration techniques included in IVGPR are the Gazdag phase-shift migration (Claerbout and Green, 2008) and a three-dimensional single-component vector migration (Streich and van der Kruk, 2007). The former is a scalar techniques commonly used in seismology, while the latter is a GPR-specific migration routine. Implementation of these two contrasting techniques makes it possible to draw meaningful conclusions on the benefits of accounting for the antenna radiation pattern and polarization associated with EM vector-wave propagation. Two controlled data sets, one from a numerical simulation and one from a field experiment, are used herein to quantitatively estimate the accuracy of the migration routines.

During the final stage, IVGPR was tested using GPR field data from surveys with pre-existing complementary interpretations made by other GPR practitioners. The IVGPR images were compared to those obtained by the colleagues using their own preferred software packages. In terms of performance, the primary goals were to ensure that results were stable and accurate irrespective of parametrization and the order and combination of processing steps. However, the available time frame of this project, in combination with its ambitious scope, allowed only a modest amount of testing. The author acknowledges that IVGPR, like any other software under active development, might contain glitches that were not identified during the limited testing. Future use of this program is likely to reveal new issues, which capable users can address.

1.4 Thesis structure

This thesis consists of five parts: a synopsis of IVGPR and related programs; basic processing routines; advance processing routines; migration algorithms; and case studies. The first three parts are especially beneficial to any future user of IVGPR wishing to fix,

expand, and/or modify the program. These sections provide a clear description of the general structure and theory of the main processing steps. The fourth part is devoted, primarily, to the development and testing of vector a migration technique. Lastly, a number of case studies will be presented to evidence how the program may serve as an effective tool for processing and interpreting GPR field records. The thesis concludes with a brief discussion and recommendations for further work.

2. THE IVGPR PROCESSING PACKAGE

2.1 Overview

IVGPR is a powerful end-to-end processing package intended for academic research involving GPR surveys. The number and variety of routines in the program surpasses that of most freely available counterparts. In addition, the programming language (i.e. Fortran 95) and components of IVGPR are nonproprietary, which expands its use beyond non-profit applications. Many architectural and implementation aspects were influenced by Clerman and Spector (2011). Perhaps most notable is the modular structure, which simplifies customization and maintenance. As a result, users wanting to expand the package can do so with only a basic understanding of the overall structure. This section will explore the architecture and core components of IVGPR.

2.2 Structure and data flow

The structure of the program consists of a main file linked to 22 modules and five open-source libraries. These libraries enable processes such as Fast Fourier Transforms, the conjugate gradient method, matrix inversion, spectral decomposition, and multitaper spectral analysis. Table 2.1 lists the modules and routines designed to aid in the editing, processing, and interpretation of GPR data. The ensemble of routines have dozens of lower level dependencies contained in various utility modules. Most of the routines in IVGPR underwent some degree of optimization in order to improve computational performance. Common optimization practices included the removal of redundant computations within loops, making effective use of Fortran's column-major order, and basic implementations of OpenMP parallelization. This work primarily focuses on the theory and algorithms related to the processing routines. Proper descriptions of lower level routines were given in the form of systematic documentation throughout the code.

Table 2.1: Modules with most of their available routines and techniques. Those marked with a * contain one or more parallel routines. Any routine containing the symbol ‡ can be applied to one- or two-dimensional data. The symbol † indicates advanced techniques.

Category	Modules	Routines and Techniques
Editing	Data management	Smooth muting
		Temporal cropping
		Trace removal
Processing	Dewow*	Median and weighted mean filter
		Low-pass filter
	Background subtraction*	Moving average trace subtraction
		Alpha-trimmed mean†
		Background matrix subtraction†
	Filtering*	Windowed-sinc
		Gaussian‡, mean‡, and median
		Edge preserving smoothing†
		2D elliptical Butterworth filter†
		Elliptical Tukey window†
	Time Zero	Peak finding
		Modified Coppen's method†
		Entropy method†
	Trace Balancing	RMS, median, and maximum amplitude
	Declipping*	Least squares polynomial fitting
		Cubic spline interpolation
	Gains	Spherical and exponential correction
		RMS, instantaneous, and robust AGC
		Programmed gain
	Spectral Balancing*	Time-variant spectral whitening
		Frequency domain spectral whitening
Differential resolution		
Migration*	Phase-shift migration	
	3D single-component exact-field vector migration†	
Interpretation	Attributes	Uniformity, homogeneity, contrast, and entropy‡
		Energy ratio
		Envelope‡
		Gradient direction
		Absolute gradient magnitude
		Instantaneous frequency‡
		Instantaneous phase‡
	Image enhancement	Box blur and Gaussian blur
		Sharpen mask and unsharp mask
		2D adaptive noise-removal

A comprehensive data flow design minimizes the effort needed to adapt new modules and functionalities. In particular, every input/output operation is controlled by a set of routines inside the data management module. All user-defined parameters and instructions are specified in a formatted text file, which utilizes Fortran's NAMELIST input statements. The parameters, along with the input GPR data, are stored within, and shared by, the data management module. Any external routine linked to the module gains access to this information. The output generated by each processing step can be saved in one of two formats, a CSV or DAT-file. The latter is especially formatted for IVGPR while the former is readable by most plotting softwares. It is worth noting that, while IVGPR currently lacks a graphical user interface, anyone wishing to append an interactive layer can exploit the fact that all communication with the program involves a single text file. Thus, an interface would only need to generate said file and execute the program.

2.3 Data visualization

IVGPR also provides a plotting module that interfaces with the open-source package gnuplot to generate figures without the need of proprietary plotting programs. Gnuplot is flexible, portable, and can generate publication-ready figures in a variety of formats (Janert, 2010). The module contains four plotting functions that can handle one- and two-dimensional data. The user can customize a number of features such as labels, font size, color maps, number of tics, and the overall dimension of the plot. Two of the functions can generate temporary on-the-fly plots useful for testing the quality of results during the development of new routines. Nearly every radargram in this work was produced using the plotting functions in IVGPR. One key limitation of this module is the absence of three-dimensional visualization capabilities. The next iteration of this program will be wrapped in python, and it will make use of the highly capable matplotlib package. This, in turn, will facilitate the adaptation of an interactive UI for visualization in three-dimensions.

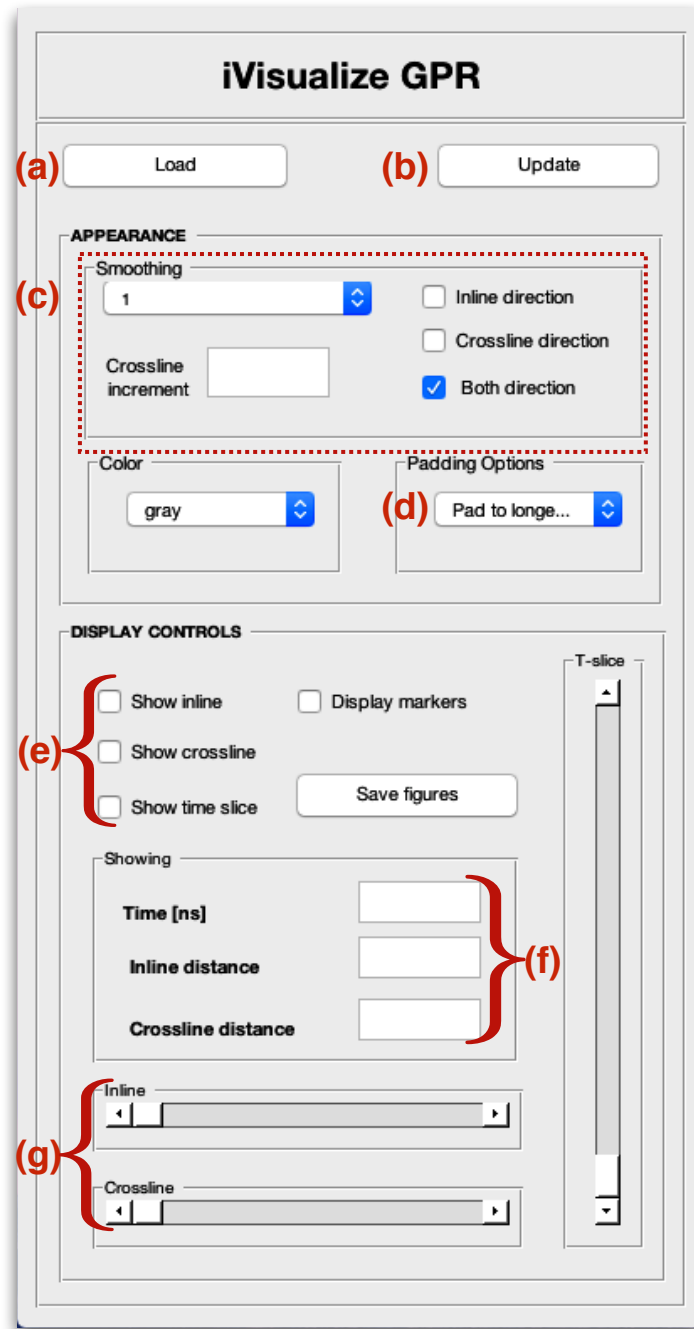


Figure 2.1: iViz3D GUI. (a) Loads a set of GPR lines, (b) implements changes in the settings to the data, (c) sets the strength of the smoothing and the direction along which it is applied, (d) pads all lines to longest or trims all lines to shortest, (e) displays inlines, crosslines, and time slices, and (f) shows the current location within GPR cube, and (g) allows the user to navigate through the cube. Specific locations within the cube can be set in (f).

2.3.1 MATLAB graphical interfaces

Although the program can read and process three-dimensional datasets, it lacks an intuitive way to display data in more than two dimensions. Furthermore, it cannot directly read the binary files generated by GPR systems. Two MATLAB interfaces, iViz3D and iGPRRead, were designed to address these limitations. The former affords a means to display and navigate through the inlines, crosslines, and depth slices of three-dimensional GPR records. Users can also enhance the smoothness of the displayed data by applying a linear interpolation. This requires setting a smoothing factor (ranging from 1 to 10) and the direction along which the interpolation should be implemented. Since this program was designed strictly for displaying purposes, a means of storing the records shown was not implemented. Figure 2.1 shows the iViz3D graphical user interface (GUI).

The iGPRRead interface was designed to convert data from the binary formats of major GPR vendors (e.g. DZT, DT1, and SEG Y) into those readable by IVGPR and MATLAB. This GUI also allows users to implement some data editing. The editing operations available can trim edge traces, remove temporal gains, and modify the header information. Figure 2.2 shows the iGPRRead interface. Future releases of the program will replace these interfaces with python based analogs. This, in turn, will eliminate the dependence on the proprietary scripting language MATLAB.

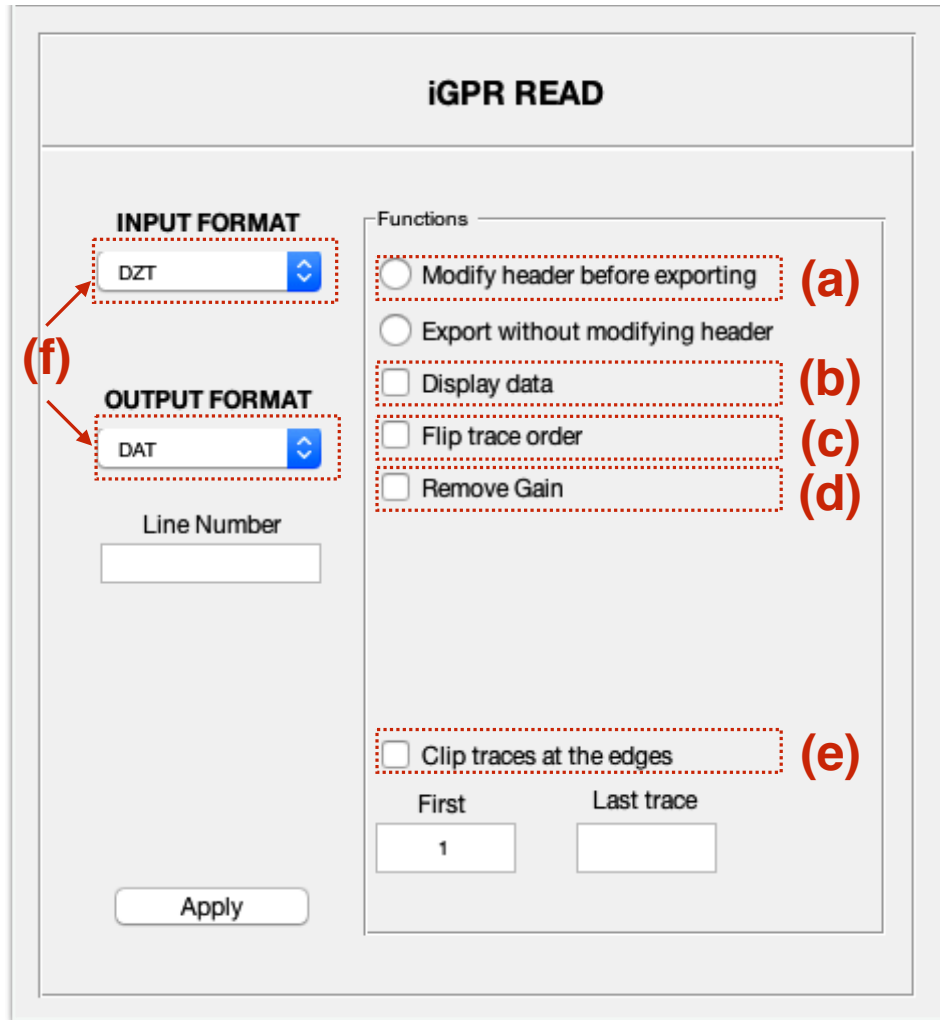


Figure 2.2: iGPRRead GUI. (a) Allows for the modification of headers, (b) displays the saved data, (c) flips the order of traces, (d) removes temporal gains applied during acquisition, (e) clips data edges, (f) specifies the input and output formats.

3. BASIC PROCESSING TECHNIQUES

3.1 Overview

In raw form, a radargram (or B-scan) represents a distorted image of the subsurface typically contaminated by unwanted signals. Processing offers a way to improve the quality and interpretability of the raw data. When properly implemented, this increases the signal-to-noise ratio, returns the data in a form that accurately represents the subsurface, or a combination of both. The need for data processing is heavily influence by the goals of the job, the nature of the raw data, and limitations in cost (Cassidy and Jol, 2009). Figure 3.1 shows a typical processing flow for bistatic GPR surveys. The orange region encloses a set of processes applicable at the acquisition site. In some cases this can be sufficient to yield a sound (and even accurate) interpretation. The gray region includes off-site processing adequate when a higher levels or data refinement is needed. The techniques colored green and blue are recommended for the majority of projects, where the blue techniques are linked to the specific survey types for which they apply.

Once the data is recorded, the next steps usually include removing bad traces, updating headers and background information, sorting files, and other steps that facilitate subsequent processing and maintenance of the survey (Annan, 1999; Cassidy and Jol, 2009). These fall under the category of data editing and tend to consume the most time in a processing flow. Data editing is followed by the application of basic processing routines. In most cases the list of basic processing routines includes dewow filtering, time-zero correction, time gains, as well as temporal and spatial filtering. Furthermore, they require little to no knowledge of a site and can be applied to most data sets without introducing much bias. Cassidy and Jol (2009) estimated that only about ten percent of GPR surveys need more than basic processing to make accurate interpretations. This section describes the theory,

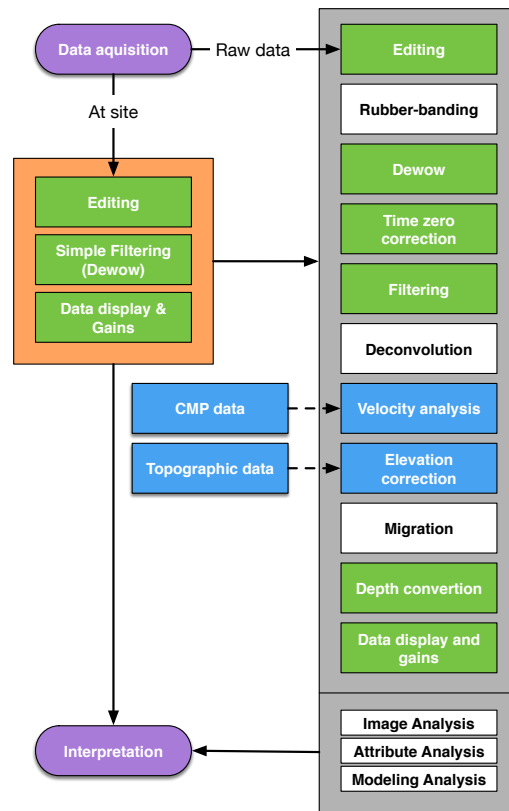


Figure 3.1: Fundamental processing steps for a bistatic GPR surveys. The orange box encloses a set of basic processing steps commonly applied at the site. These are usually enough when detection of a target is the primary goal. The gray box encloses offsite steps, which are typically applied in descending order.

and demonstrates successful applications, of the basic routines included in IVGPR.

3.2 Dewow

Wow is a nonlinear, non-stationary, low-frequency noise caused by inductive coupling and/or signal saturation from shallow scatterers (Cassidy and Jol, 2009). The affected traces can have a non-zero baseline amplitude (or DC bias) that slowly bows up or down with time, as shown in figure 3.2. However, dewow typically refers to the removal of low frequency noise, which may or may not conform with the characteristics defined by

Cassidy and Jol (2009). Classical dewow routines implement mean filters, median filters, and low-cut frequency filters. Gerlitz et al. (1993) compared these three approaches and found that residual median filters (RMF) generated superior results. Contrary to this, Schmelzbach et al. (2012) observed that low-cut frequency filters outperformed RMF for their datasets. One must remember that the performance of dewow techniques is subject to implementation bias and the characteristics of the data at hand. Therefore, an optimal dewow tool may not exist. The dewow bundle in IVGPR, and the subject of this section, contains the *mean*, *median*, and *low-cut filtering* techniques.

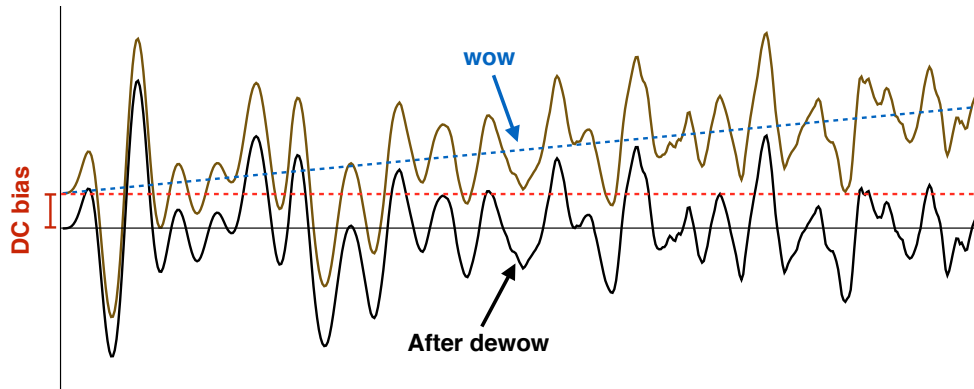
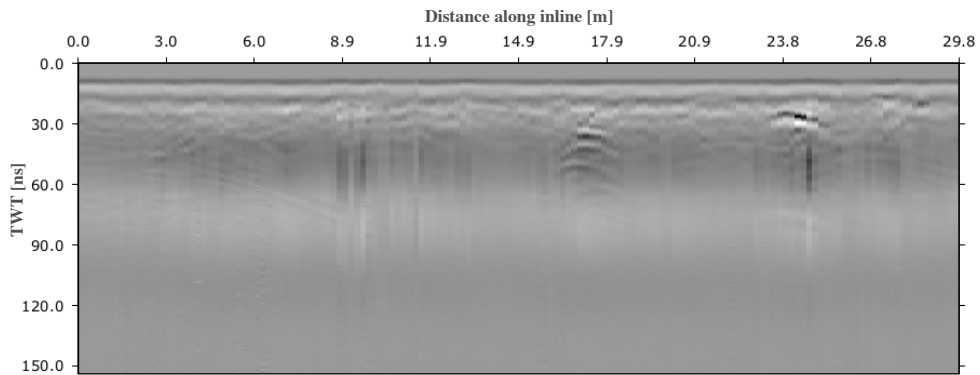


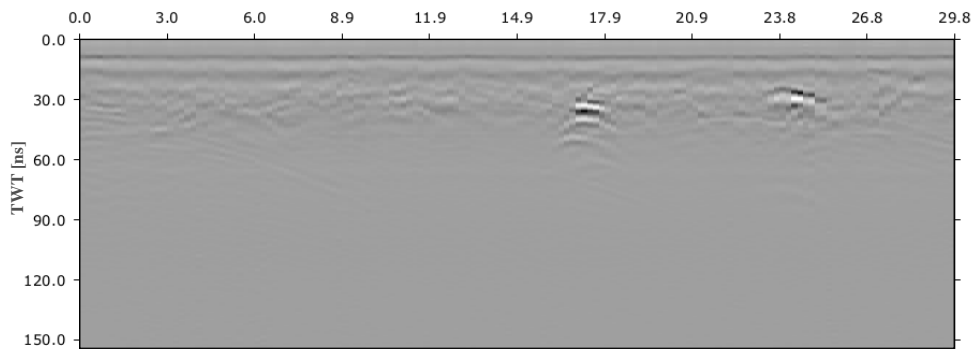
Figure 3.2: Example of wow noise in GPR trace and effect of dewow filter.

An example of wow noise in a field record is shown in figure 3.3. This dataset was first treated with a programmed gain to reveal structures masked by attenuation and limitations in the instrument’s dynamic range. Unfortunately, a low frequency noise was inadvertently amplified and a set of broad horizontal bands appeared throughout the radar-gram. Figure 3.9b shows the output from IVGPR obtained by removing the wow noise with a low-cut filter prior to implementing a programmed gain. The output image shows no sign of the unwanted noise and, in addition, the resolution and interpretability of the

image were effectively enhanced. Furthermore, a low-cut filtering dewow is straightforward to implement in IVGPR, requiring only a user defined filter kernel size and a cut-off frequency.



(a)



(b)

Figure 3.3: GPR field record showing broad horizontal bands caused by low frequency wow noise. The banding became prominent after applying a temporal gain.(b) Demonstrates the improvement obtained by applying a dewow filter to (a) prior to implementing the temporal gain.

After extensive testing, attempts to remove the same noise using the RMF and mean filter dewow proved futile. The amplitude spectrum computed by averaging the amplitude spectra of every trace is shown in figure 3.4. The low- frequency horizontal banding was

initially thought to be the result of relatively high amplitudes near 30 MHz. This was later confirmed by applying a set of low-cut filters and noting that a cut-off frequency of 30 MHz failed to remove the banding while 35 MHz generated the desired result. The filtered amplitude spectrum (black solid line) and the filter used are shown in figure 3.4. This noise corresponded to frequencies higher than what RMF and mean filter dewow are intended to effectively treat. Additionally, its frequency content was highly localized and required a filter optimized for frequency separation. Median and mean filters are generally inadequate for frequency filtering. This example, however, should not discourage the use of these two routines. Not only are they simple to implement, with the only required parameter being a temporal window size, but they remain a popular way to perform dewow filtering.

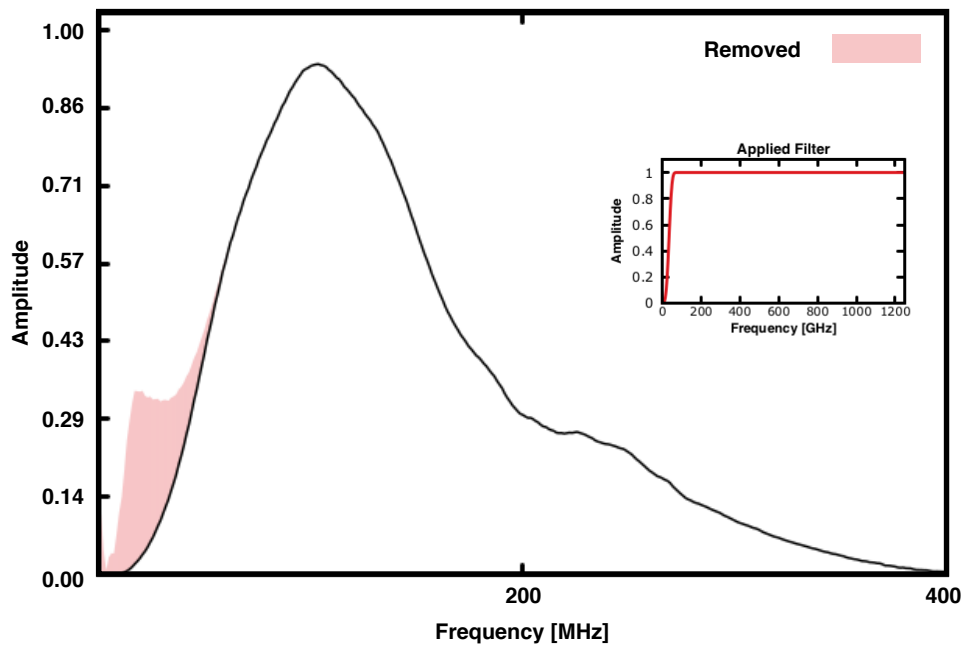


Figure 3.4: Average amplitude spectrum of GPR record. The region shaded red shows the portion of the spectrum removed by the filter. The filter used is also shown.

3.3 Filtering

An unprocessed GPR record generally constitutes a mixture of the subsurface response with other undesirable signals, e.g. ambient signals from electronic devices (Olhoeft, 2000) and inherent system noise (Rashed and Harbi, 2014). Filters offer ways to separate these signals and enhance (or suppress) features that would otherwise hinder the interpretability of records. While there exists a virtually unlimited variety of filters with varying levels of sophistication, one often finds that basic filtering routines are sufficient for most tasks (Smith et al., 1997). There are many excellent literary works on discrete-signal processing and digital filters where the interested reader can gain a robust understanding of filter design and implementation. The goal of this section is not to dwell on the details and intricacies of discrete signal processing. Rather, we will focus on how filtering is performed in IVGPR and observe some useful examples.

3.3.1 The windowed-sinc filter

One-dimensional frequency filtering in IVGPR utilizes the *digital windowed-sinc filter* technique described by Smith et al. (1997). This procedure can implement *low-pass*, *high-pass*, *band-pass*, and *notch* filters by convolving the corresponding temporal windowed-sinc kernels with the signal. Filters applied in this manner belong to the class of Finite Impulse Response (FIR) filters (Oppenheim et al., 1989; Smith et al., 1997). Figure 3.5 depicts a typical FIR filtering flow like the one in IVGPR. The first step is to create a filter kernel in time (space) domain. To this end, the user specifies the size, type, and cut-off frequencies. The i -th element of a temporal (spatial) windowed-sinc kernel (h) for a low-pass filter with cutoff frequency f_c is found using

$$h[i] = \frac{\sin(2\pi f_c(i - M/2))}{i - M/2} \left[0.42 - 0.5 \cos\left(\frac{2\pi i}{M}\right) + 0.08 \cos\left(\frac{4\pi i}{M}\right) \right], \quad (3.1)$$

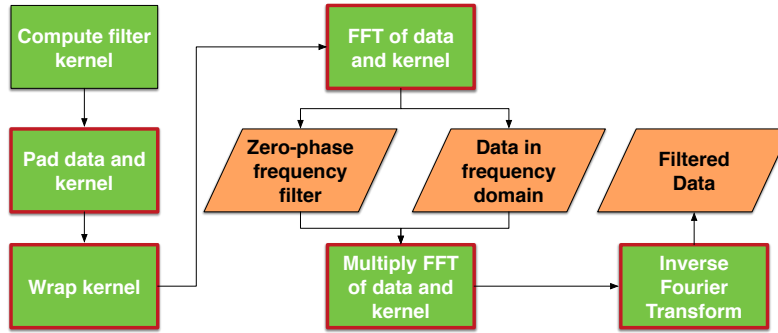


Figure 3.5: Illustrates a typical frequency filtering flow (similar to the one used in IVGPR).

where M is some odd integer describing the size of the filter kernel. High-pass, band-pass and notch filters can be created from low-pass filters using spectral inversion and spectral reversal (Smith et al., 1997).

The remaining steps describe a linear convolution between the filter kernel and the signal. Note that this is carried out as a multiplication in Fourier-domain to take advantage of the high computational performance of fast Fourier transforms (FFT). Since FFT based convolution is inherently circular, the filter kernel and data must be padded in order to avoid a wraparound of nonzero amplitudes. Namely, if a time series is of size N and the filter kernels is of size M , then the two should be padded to a size equal or bigger than $N + M - 1$. If the filter is zero-phase, the desired filtered output will be contained within the first N elements of the convolved response. For the phase of the filter to be zero, one must ensure that the elements $\{1, 2, \dots, \frac{M-1}{2} - 1\}$ and $\{\frac{2N+M+1}{2}, \frac{2N+M+1}{2} + 1, \dots, N + M - 1\}$ in the padded kernel correspond to the elements $\{\frac{M-1}{2} + 1, \frac{M-1}{2} + 2, \dots, M\}$ and $\{1, 2, \dots, \frac{M-1}{2}\}$ in the original kernel respectively.

Users should have a comfortable level of familiarity with the properties of windowed-sinc filters prior to implementing them. In particularly, they should understand how the

width of the transition band relates to the size of the kernel, how the cutoff frequency is defined, and the symmetrical aspects of the frequency response. The transition band is the region of the amplitude spectrum that separates the pass-band (or region of maximum gain) from the stop-band (or region of zero gain). The sharpness of the transition band, i.e. its width, is inversely proportional to the size of the filter kernel. Figure 3.6 shows the amplitude spectrum of three bandpass filters with identical cutoff frequencies but different kernel sizes. Note that, as the number of elements increases, the transition from a value of one to zero (and vice versa) occurs within a narrower frequency range. Also shown in figure 3.6 are two vertical lines placed at the cutoff frequencies, i.e. 300 MHz and 1000 MHz. These intersect the amplitude response of the filter at a value of 0.5, which differs from the conventional definition of 0.707 (or -3 dB). Finally, noticed that the transition band appears to have odd symmetry about the half maximum point. This is in fact the case and it explains why we can use a low-pass filter to derive the other filter types.

3.3.2 Basic smoothing filters

In some cases we may want to remove features in the data that, when analyzed in the Fourier domain, have an unconfined frequency response. A filter optimized to separate bands of frequencies is generally ineffective for such tasks (Smith et al., 1997). Consider, in particular, the treatment of spikes in a record. These features are highly localized in physical space and, in turn, have energies that spans a broad spectrum and commonly overlap with those of an otherwise noise-free signal. Therefore, it is complicated, if not unfeasible, to effectively implement a band-cut filter to treat this type of noise. Rather, we may find that a median filter effectively eliminates spikes without adversely influencing the character of the data. Another example is that of random noise, which is optimally treated by basic mean and gaussian filters. The median, mean, and gaussian filters constitute the set of basic smoothing filters in IVGPR. For the first two the user only

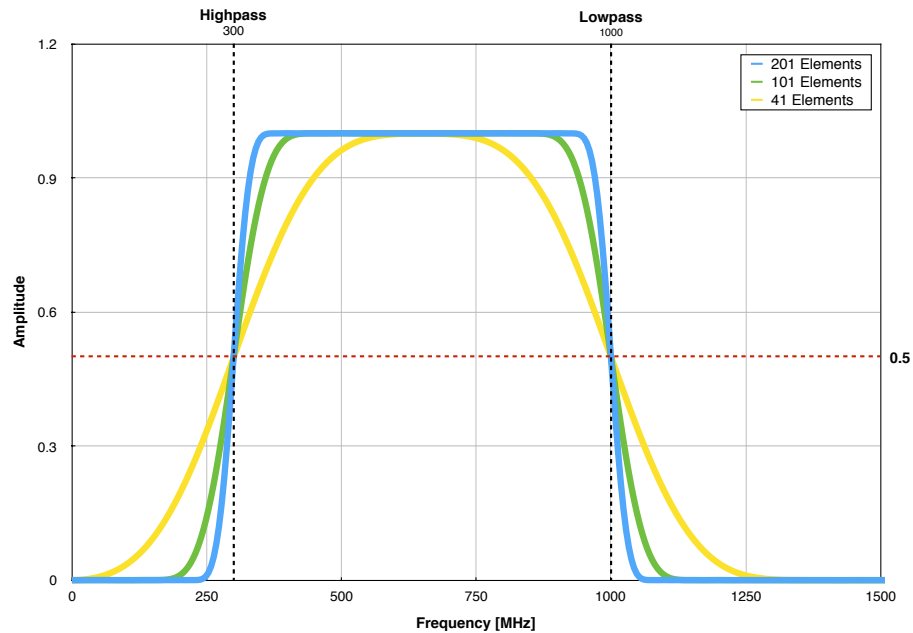


Figure 3.6: Amplitude response of three bandpass windowed-sinc filter with identical cut-off frequencies (i.e. $f_{hp} = 300$ MHz and $f_{lp} = 1000$ MHz) but varying kernel sizes. This filters were created using the windowed-sinc routine in IVGPR.

needs to define the size of a window (or kernel), while for gaussian filters one must also define the desired standard deviation. These filters can also be applied along two dimensions. The median and windowed-sinc filters may be applied using a separable scheme, i.e. two one-dimensional filters oriented along different dimension, or by a two-dimensional moving window. Figure 3.7 shows the results of a mean filter applied using a moving window. Since the means and gaussian smoothing filters have well-defined equations in two-dimensional Cartesian coordinates, they can be applied using a two-dimensional FFT-based convolution. The convolution of these filters with an image is commonly referred to as box and gaussian blur in image processing.

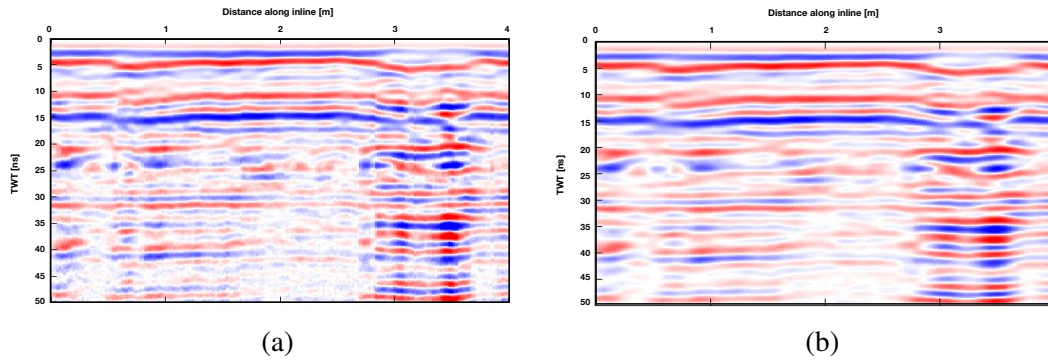


Figure 3.7: Shows the product (b) obtained by smoothing a raw radargram (a) using a two-dimensional mean filter with a 5×5 window. (b) Shows that this kernel achieved a fair trade-off between resolution and smoothing.

3.4 Time-zero correction

The time value representing the arrival of the earliest reflected wavelet is known as the time-zero point. This position can change because of variations in antenna elevations, cable length differences, electronic instabilities, and thermal drift (Cassidy and Jol, 2009). If left untreated, time-zero shifts can deteriorate lateral coherency and cause inaccurate depth estimations. Thus, it is necessary to adjust the zero time of all traces to a consistent and stable location along the earliest reflected wavelet. As noted by Yelf (2004), a global consensus for the exact location of the time-zero point on a wavelet does not exist. The choice usually depends on the options available in the algorithms of GPR manufacturers and the preference of the users. Figure 3.8 shows four typical time-zero pick locations. Yelf (2004) studied the stability of these picks and reported that the first negative peak (A) and the first crossover point (B) provide good stability for geological applications that use antennas with frequencies ranging from 25 to 500 MHz. In general, however, the stability of a time-zero algorithm largely depends on the quality of the data and the robustness of the method used. In IVGPR workers can apply two contrasting approaches to automatically adjust the zero time. The main difference between these techniques lies

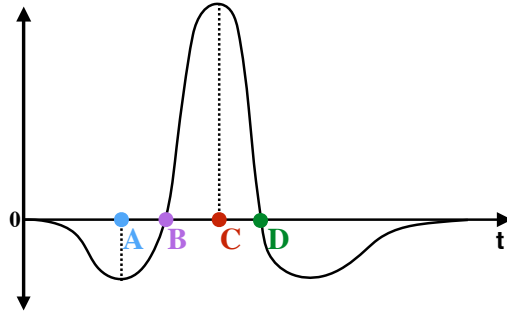


Figure 3.8: Four conventional locations for picking the time-zero point along the first reflected wavelet. Point A marks the first negative peak. Point B indicates the S-crossover. Point C indicates the first maximum. Point D corresponds to the Z-crossover.

in their respective levels of sophistication. This section will focus on the simple option, which relies on a *peak-finding* scheme to predict the time-zero point along each trace.

The time-zero correction based on peak-finding begins by searching for the first significant peak along a trace. It does this by first locating all peaks and determining the maximum peak amplitude. Recall that the i -th element along a trace is a peak if its amplitude, a_i , satisfies the relation $a_{i-1} < a_i > a_{i+1}$. This condition cannot be satisfied at the boundaries, which may have amplitudes surpassing the highest peak. Next, the algorithm determines the minimum significant amplitude from a user-defined percentage, p , such that all acceptable peak amplitudes exist in the range $[a_{\max} - a_{\max} \cdot p, a_{\max}]$. The following steps involve searching for the first occurrence of a significant amplitude, trimming the top of the record to reflect that the peak found is the zero time value, and padding (with a smooth transition) the end of the trace in order to have a consistent temporal duration for each trace. Troughs can be detected effortlessly after multiplying the input by -1 . Figure 3.9 shows a B-scan successfully treated by the basic time-zero correction algorithm in IVGPR.

Zero-crossover points can be found in a similar manner. In fact, a relevant zero-crossover point must lie adjacent to a significant peak (or trough). The neighboring ex-

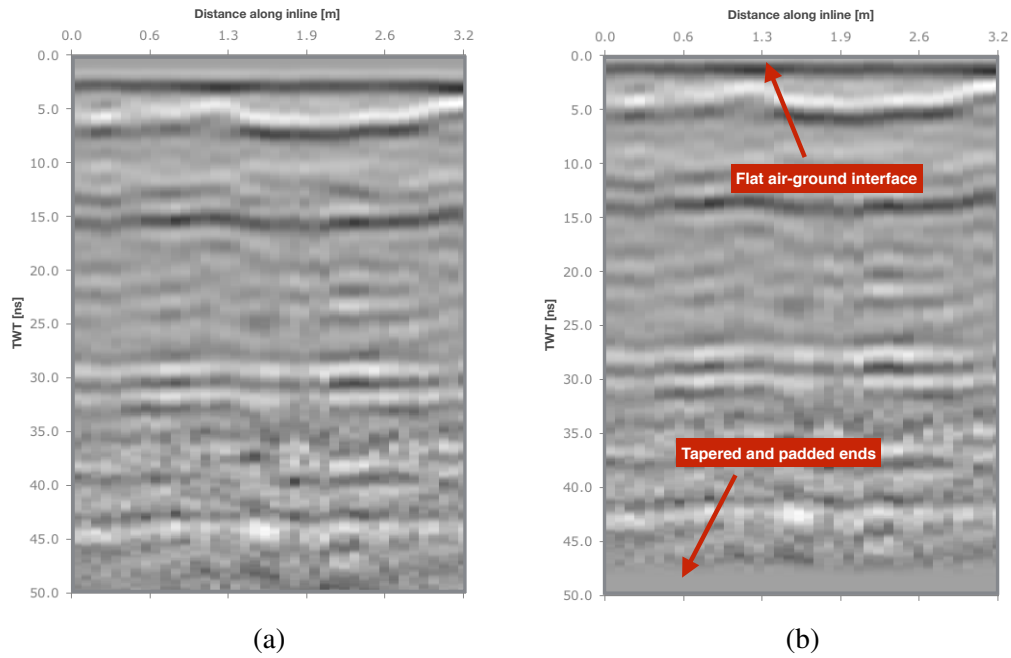


Figure 3.9: Shows a radargram before (a) and after (b) time-zero correction. The preferred location was set to be the first significant peak.

trema depends on the type of crossover point, i.e. S-crossover and Z-crossover. For example, consider the case where the user selects the S-crossover option for the time-zero points. The algorithm would then look for the first significant trough, a_i . Next, it would search for the first amplitude, a_j , that satisfies the relations $a_{j-1} < 0$ and $a_j \geq 0$ (where $j > i$). The crossover point would then be identified as the j -th element along the time series. Notice that the second condition for a S-crossover point admits values greater than zero. That is because a data point of zero amplitude prior to crossing into positive values might not exist.

The algorithms described here is highly sensitive to the character and quality of the data. For example, high levels of saturation in a trace can make it difficult to clearly define the amplitude relation and location of extrema. Furthermore, zero-crossover points might be inconsistent when there are sharp high amplitude fluctuations in the early parts of the

data. Both complications described can be ameliorated with additional processing steps. Therefore, workers should be familiar with the properties of their records and implement any necessary processing prior to implementing time-zero correction.

3.5 Time gains

Amplitudes in a raw A-scan generally show a decrease with time due to geometrical spreading and attenuation. As a result, late arrivals become difficult to discern and the weak amplitudes can be mistakenly linked to subtle changes in the electromagnetic properties of the medium. The latter can give the illusion that the deeper materials are vastly different from those in shallow sections. Additionally, the dynamic range of the raw data, which is influenced by amplitude variations, can exceed that of the displaying method in use. Gain functions attempt to compensate for time-dependent amplitude decays. Figure 3.10 depicts time gains as the correction to amplitude attenuation along the trace. Note that the attenuated trace obscures the fact that absolute reflection strength remains constant at all interfaces. The most common time gains in GPR applications are automatic gain control (AGC), spherical and exponential compensation (SEC), constant, and user defined gains (Cassidy and Jol, 2009). In practice, time gains work by generating a set of scaling factors that, when multiplied with the signal, equalize the average absolute amplitude at all time levels. There is a plethora of literary works focusing on gaining techniques, even within the context of GPR. Cassidy and Jol (2009) concisely describes the benefits and shortcomings of popular gaining techniques for GPR applications. This section focuses on the gaining algorithms included in IVGPR and gives examples to demonstrate their effectiveness.

3.5.1 Automatic gain control

An automatic gain control (AGC) equalizes the strength of signals across a trace for display purposes (Jol and Bristow, 2003). Unlike the exponential and spherical compensa-

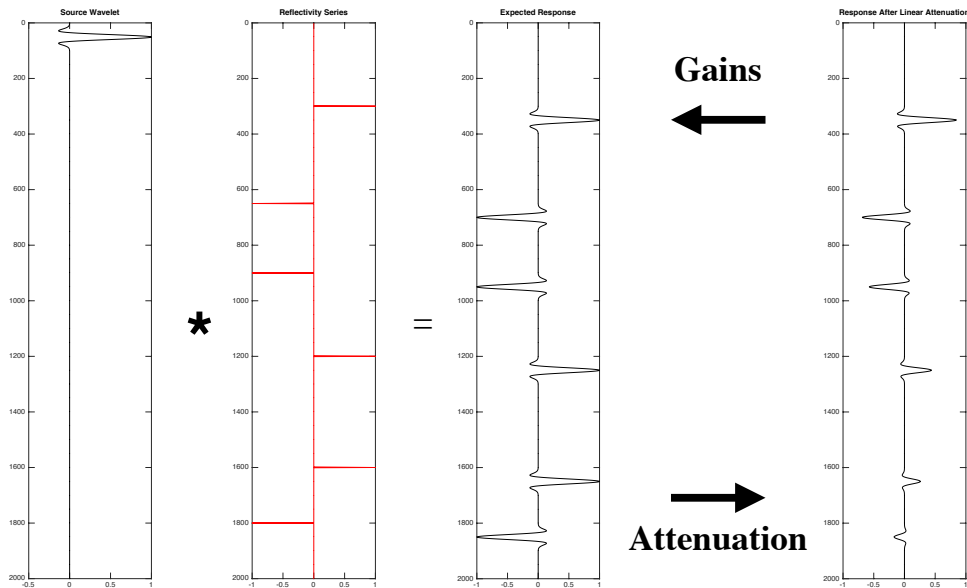


Figure 3.10: Shows the convolution of a finite duration pulse (left) with a reflectivity series (red). The resulting time series was attenuated using a linear decay down to 10% of the original amplitude. Without a gain, an interpreter could mistakenly deduce that the contrast in material properties varies with depth.

tion gains, AGC-type gains lack a direct connection to the physical phenomena that cause attenuation. These algorithms simply reduce the dynamic range of the data such that it can match, or be contained in, that of the displaying method. Their outputs show degraded or absent contrast in signal strength. This can inhibit the ability to learn about the properties of the medium, which correlate to the relative amplitude information (van der Kruk, 2015). Consequently, true amplitude processing workflows tend to avoid the use of AGC. The AGC-family performs exceptionally well in applications where the main objective is to assess the lateral continuity of an event, such as delineating geologic horizons (Annan, 1999). Workers with limited familiarity to this family of gains should only use them for visualization at the end of a workflow.

The AGC-type gains in IVGPR consist of an *instantaneous*, *root mean square (rms)*, and a *robust* AGC. The first two methods follow the popular development presented in

Yilmaz (2001). Some attractive features of the rms and instantaneous amplitude AGC include the need for only three basic parameters and the straight forward encoding. The gaining coefficient at the center of a time gate for all three techniques is of the form

$$G = \frac{R}{A}, \quad (3.2)$$

where R is a reference rms level and A is an averaging method for the values inside the gate. In the case of the instantaneous AGC, the averaging is performed using the absolute value of the amplitudes contained within a gate such that

$$A_{\text{inst}} = \frac{1}{N} \sum_{i=1}^N |x_i|, \quad (3.3)$$

where x_i represents the i -th amplitude within the time gate. The rms AGC algorithm uses the quadratic (or rms) average, which result in

$$A_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}. \quad (3.4)$$

Another important distinction between the rms and instantaneous AGC, as implemented in IVGPR, relates to how the gaining coefficients are computed for all time samples in a trace. In the case of the rms AGC, an incomplete set of coefficients is computed only at the center of non-overlapping time gates. The remaining coefficients are found after interpolating among the computed values (and extrapolating the edges). By contrast, the instantaneous AGC uses overlapping time gates to define the gaining coefficient at all time samples, edge samples excepted. Figure 3.11 illustrates how the rms and instantaneous AGC algorithms determining the gain coefficients for all samples in a trace. For the time gate shown, the rms AGC computes only three coefficients from the data while

deriving the rest through interpolation and extrapolation. The instantaneous AGC extracts five time more coefficients directly from the data. While the latter might appear more attractive, adjusting each value individually can boost the deterioration of relative amplitude information.

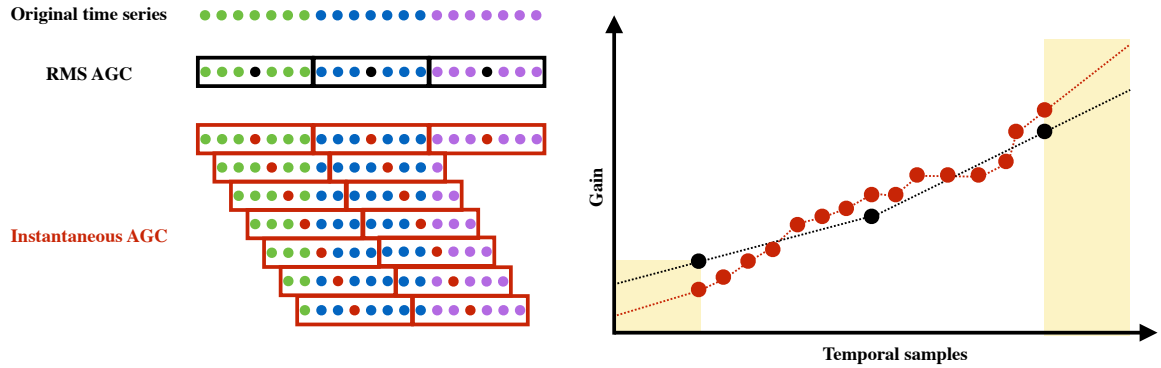


Figure 3.11: Scheme used to compute the gaining coefficients for the instantaneous and rms AGC. The original time series is depicted as a series of points sectioned into three parts (green, blue, and purple). The rms AGC algorithm computes the coefficients at the center of each time gate (box) without overlap. The instantaneous AGC computes the gain at the center of overlapping windows, which are shown in a stair case pattern for clarity. The yellow section indicates the region where values are extrapolated in both techniques.

Robust AGC is a modification to the instantaneous AGC that replaces the A_{inst} with the alpha-trimmed mean of the absolute values within the gate. Thus, we have that

$$A_{\text{rob}} = \frac{1}{N - 2[\alpha N]} \sum_{i=1+[\alpha N]}^{N-[\alpha N]} |x_i|, \quad (3.5)$$

where the square brackets denote the greater integer part and α determines what percentage of the samples are trimmed. The value of α should satisfy $0 \leq \alpha < 0.5$ (Bednar and Watt, 1984), where choosing $\alpha = 0$ results in A_{inst} . This algorithm allows the preservation of amplitude extremes, which produces gained traces that retain more of the true character

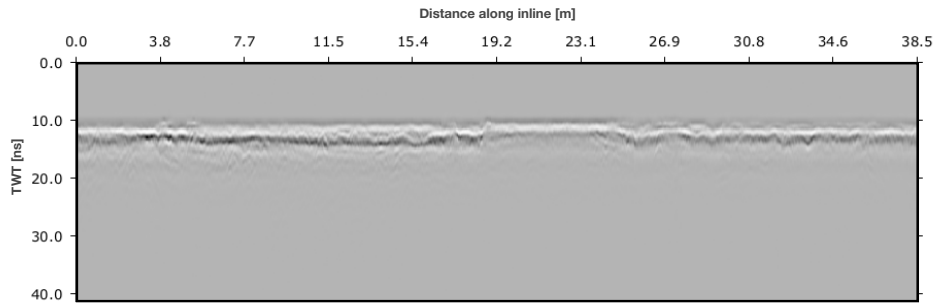
of the input. Thus, the robust AGC computes a data-derived gain function that better preserved relative amplitude information. Figure 3.12 shows sample outputs from all three automatic gain control techniques included in IVGPR identical time gates. This record used was acquired over the recreation yard in Alcatraz’s main prison. We see that, while all outputs are very similar, instantaneous AGC and robust AGS shows the most structural features. As a word of caution, a small time gate in all three algorithms can greatly amplify small amplitudes related to unwanted signal and obscure events of importance.

3.5.2 Spherical and exponential compensation

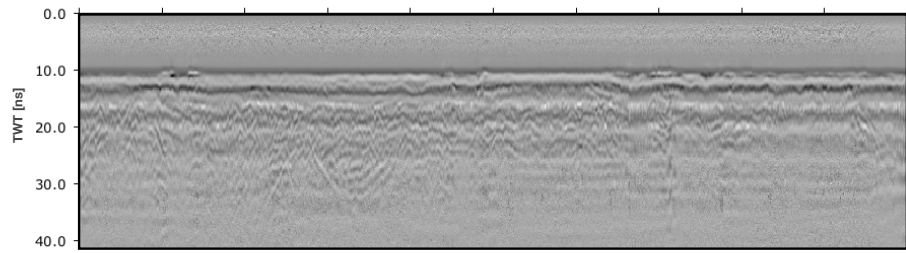
In order for a time gain to achieve an ideal restoration of the “true” amplitudes it must account for all physical processes distorting the desired amplitude information. Spherical and exponential compensation (SEC) gains follow this principal by reversing the effect of two theoretical models of attenuation, i.e. spherical divergence and exponential decay. The first effect relates to the amplitude decay of a spherically propagating wavefront, which is proportional to $1/r$ (Yilmaz, 2001). In addition, the wavefront can experience an exponential decay in amplitude due to material losses. The SEC time gain in IVGPR uses the equation proposed by Sandmeier (2002):

$$G(t) = \left(1 + \frac{\alpha}{W}t\right) \exp(\beta vt), \quad (3.6)$$

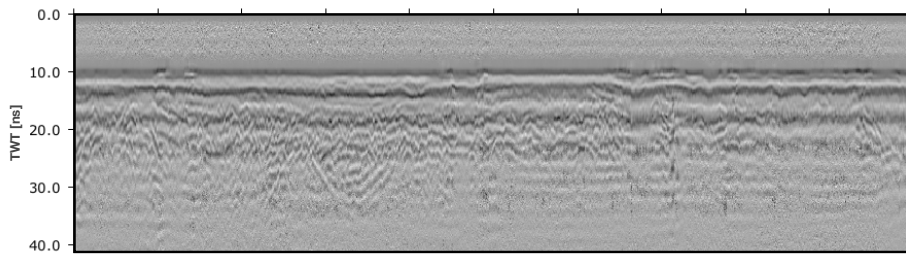
where β is the exponential damping, v is the velocity model of the propagation medium, W is the pulse width, and α represents the linear gain. Equation 3.6 is most effective when the velocity distribution and lossy properties of the subsurface can be accurately inferred, which is rarely the case. To make matters worst, electromagnetic wavefronts do not generally propagate spherically (isotropically) in geologic materials. Still, SEC gains offer an alternative to AGC that abides by a sensible physical model of attenuation and better preserve relative amplitude information. Figure 3.13 shows an example of the SEC rou-



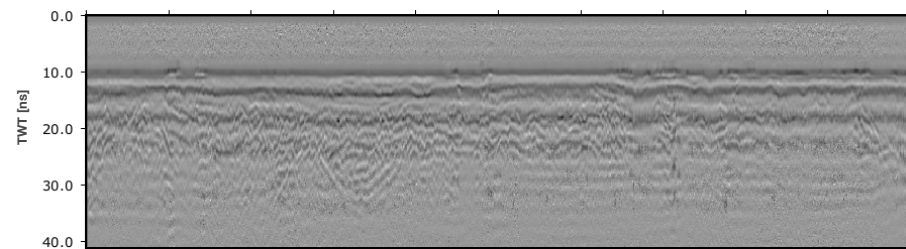
(a) B-scan before applying AGC.



(b) B-scan after applying rms AGC.



(c) B-scan after applying instantaneous AGC.



(d) B-scan after applying robust AGC.

Figure 3.12: (a) Shows the input B-scan in raw form. Reflectors are clearly discernible only within 5 ns from the ground surface. (b) shows the output from the rms AGC routine. This greatly enhanced the number of visible features. A similar effect is obtained after instantaneous AGC (c) and robust AGC (d). The main differences between the tree techniques is the effect on relative lateral amplitudes and structural resolution.

tine in IVGPR applied to a field record. The relative lateral resolution, judging by the top reflector, appears to be mostly unaffected. A red arrow points some features that emerged after the correction. The radargram appears to have gained as much as 10 ns of resolution, but many features remain obscured. Even achieving this modest amount of improvement proved to be a laborious task.

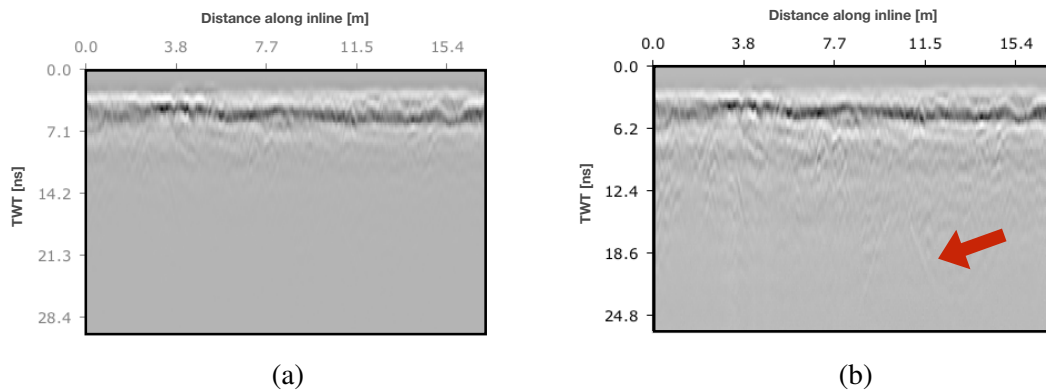


Figure 3.13: Illustrates how the amplitude decay in a raw scan (a) can be ameliorated by applying the SEC algorithm. (b) The B-scan after SEC. A red arrow indicates where a deep feature of interest became apparent after the gain was applied.

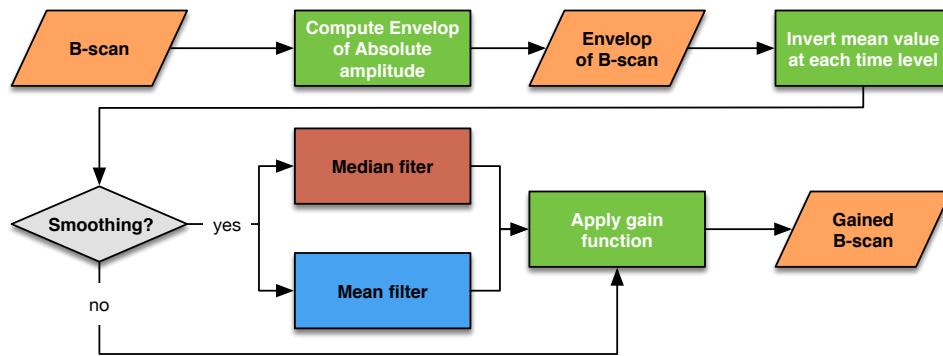
3.5.3 Programmed gains

The third type of gain in IVGPR belongs to the family of programmed gains, where a data driven technique is used to automatically estimate a global gain curve. Programmed gains (or programmed gained control) are notoriously simple to implement, require few user defined parameters, and do not generally suffer from a severe deterioration in relative amplitudes—as it is the case with AGC gains (Yilmaz, 2001). The implementation of these gains can be broken down into three basic steps: (1) estimate the amplitude decay; (2) invert, scale, and threshold the amplitude decay obtained from the first step in order to obtain a global gain function; and (3) apply the resulting gain function to all of

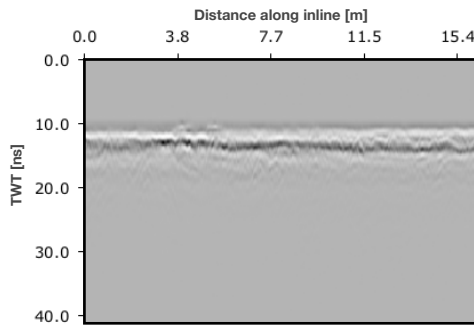
the traces using a point-wise multiplication. A workflow for this algorithm is shown in figure 3.17a. A worker applying a programmed gain only needs to specify parameters related to the method of estimating the amplitude decay curve. For example, in IVGPR an amplitude decay curve is first estimated for each trace using the envelop of the absolute amplitudes. This procedure is free of user-defined parameters, but the maximum gain level allowed can be threshold. Next, the algorithm combines the decay curves using a simple arithmetic average at each time level. The final step involves smoothing the decay curve, which requires another user-defined parameter. Workers can choose to apply a mean filter, median filter, or to use the original decay curve. The first two options require the specification of a window size, which adds an extra parameter to the implementation. However, every permissible choice of window size results in an acceptable gained section. Note that the lowest allowable size (i.e. one sample) outputs a decay curve identical to the input, while at the upper limit (i.e. all of the samples in the trace) the decay curve is simply scaled down by a constant without losing its character. Thus, program gains reduce the amount of user bias in time gaining by necessitating only three basic parameters (at most) that do not require any *a priori* knowledge of the subsurface materials. The preservation of signal character and ease of use make programmed gains the preferred choice in IVGPR. Figure 3.14c shows a satisfactory treatment of a raw trace (figure 3.14b) by this routine. Note that the amount of visible features is comparable to those seen after AGC gains but lateral amplitude information was better preserved.

3.6 Desaturation

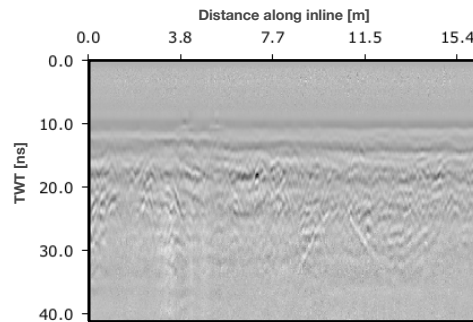
The amplitude of incoming signals sometimes exceeds the dynamic range of a receiver system, a phenomenon known as signal saturation (or clipping). Figure 3.15 illustrates how clipping manifests along a GPR trace. Highlighted sections show where the dynamic range “clipped” the true amplitudes. This also affects the smoothness and localizes the



(a) B-scan after applying robust AGC.



(b)



(c)

Figure 3.14: (a) Workflow for programmed gains. (b) Raw B-scan affected by attenuation. (c) The B-scan after being processed by a programmed gain without smoothing.

amplitude response in time. Signal saturation is common in GPR surveys and, in severe cases, can render a record unusable for the intended goals. It commonly results from the improper use of gains during acquisition. Perhaps the most serious consequences of saturation relates to the difficulties it creates during processing. Not only are temporal amplitude variations of affected traces distorted, but their frequency content becomes artificially enriched with high frequency harmonics. The latter can hamper the performance of frequency-domain routines like *spectral balancing* and *FT-based migrations*. Complications also arise in time-domain routines such as *time-zero correction* and *background matrix subtraction*. To illustrate an example, consider the application of *trace normal-*

ization. In some situations it is useful to normalize each trace to its absolute maximum amplitude in order to correct for variations in antenna-ground coupling (Cassidy and Jol, 2009). However, if consecutive traces are clipped, they will have identical normalization coefficients (i.e. the limit of the dynamic range) that do not reflect their true absolute maximum amplitude.

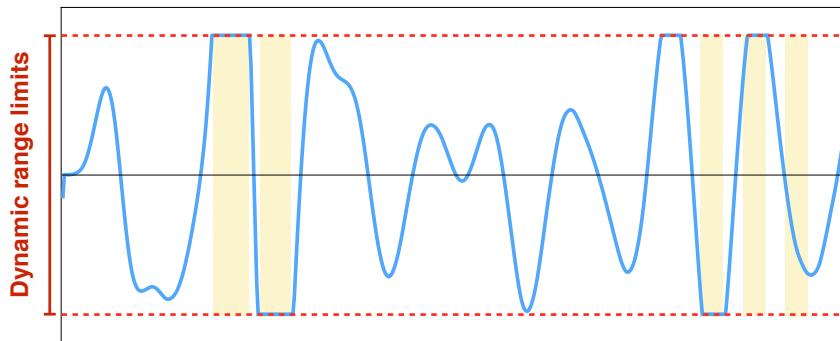
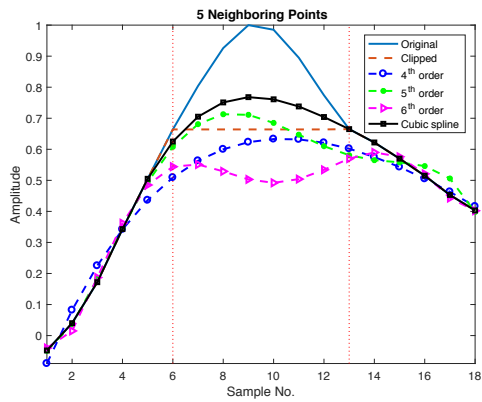


Figure 3.15: Shows the manifestation of signal saturation along a GPR trace. The dashed red lines indicate the limits of the dynamic range for the receiver system. Clipped sections appear as horizontal plateaus collinear with the limits of the dynamic range.

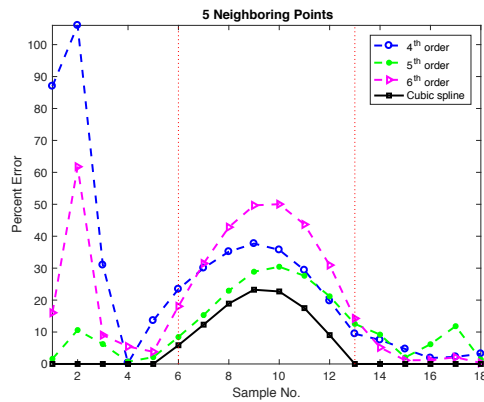
Clipping is not particular to GPR records, and many ways of ameliorating its effect have been proposed across multiple industries. Nonetheless, an exact reconstruction of affected signals is unfeasible. The two techniques included in IVGPR use a cubic spline or a least-square polynomial interpolation, respectively, to interpolate across clipped regions. The former is more common in commercial softwares (Cassidy and Jol, 2009) possibly due to its unconditional smoothness and stability. By contrast, standard polynomial interpolation is subject to instabilities and can be heavily influenced by the quality and abundance of points surrounding the damaged sections. In particular, poor data quality can lead to incoherent reconstructions while the number of available points limits the order and robustness of the polynomial fitting.

The performance of declipping in IVGPR was assessed with the help of a manually clipped trace. A section with no obvious symmetry was judiciously picked to generalize the results. Figure 3.16 shows the outcome of this experiment. A set of vertical dotted lines enclose the clipped amplitudes. The edge samples of clipped sections generally do not correspond to real values, since their locations are determined by the sampling rate and their amplitudes by the dynamic range. As a result, the sets of “neighboring points” used contain amplitudes leading to but not including those at the edges. As expected, the performance of cubic spline was independent of the number of datapoints used and returned the same correction each time. The same behavior was not observed for a fourth-, fifth-, and sixth-order polynomial interpolation. While the lowest percent error along the clipped amplitudes was achieved by the sixth-order polynomial using nine neighboring points, increasing the number of points to 13 reduced the accuracy of this fit. In fact, the accuracy of the polynomial fits seemed to converge as the number of points used increased.

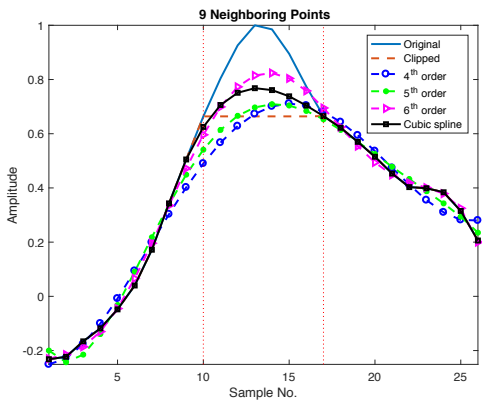
The steps following interpolation are identical for both routines. They include stitching the cap to the nearest neighbors and smoothing the connections. This is repeated until every trace is free of clipping. Both methods require the user to specify the minimum tolerable size of clipped regions, the number of neighboring points to use, the size of the smoothing window, and the number of points around the stitching to smooth. Additionally, least-square polynomial interpolation requires the specification of an order. Figure 3.17 shows a satisfactory application of declipping using a fourth-order polynomial interpolation. The stitching was smoothed using a three point window applied to the three surrounding elements. Notice that a reasonable enhancement lateral amplitude variation was achieved by this technique. Better results might be observed if declipping is followed by a proper smoothing scheme such as the structure-oriented smoothing of Hale (2009).



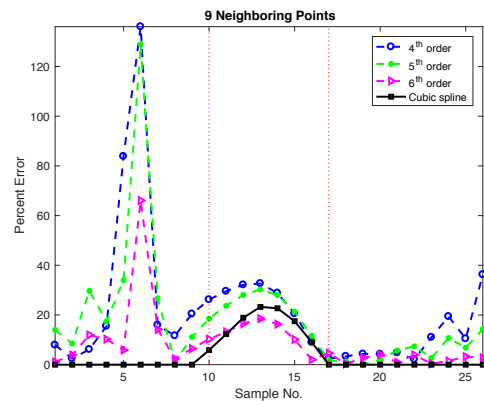
(a)



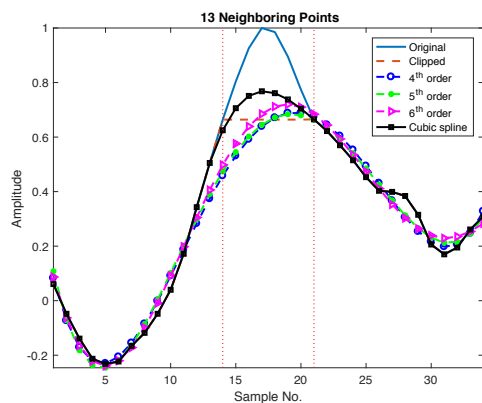
(b)



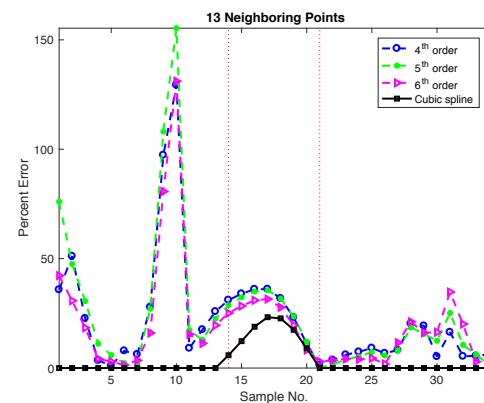
(c)



(d)



(e)



(f)

Figure 3.16: Interpolation results using a fourth-, fifth-, and sixth- order polynomial interpolation along with a cubic spline to estimate the amplitudes along a clipped section. Each interpolation was compared to the actual amplitudes using a percent error analysis. The number of neighboring points used were 5 (a and b), 9 (c and d), and 13 (e and f).

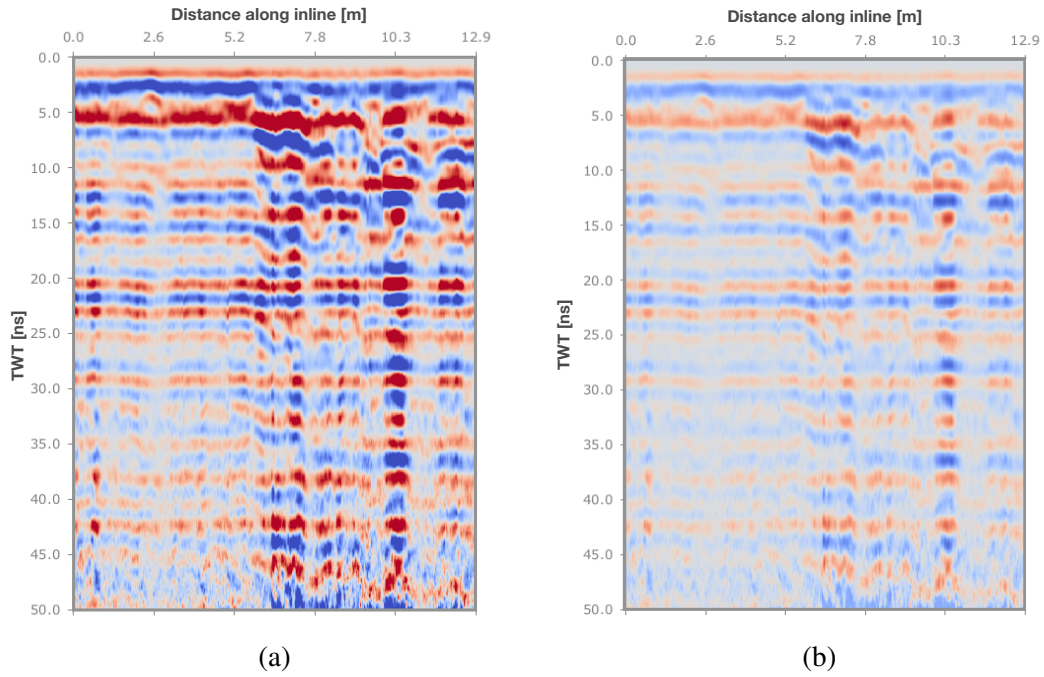


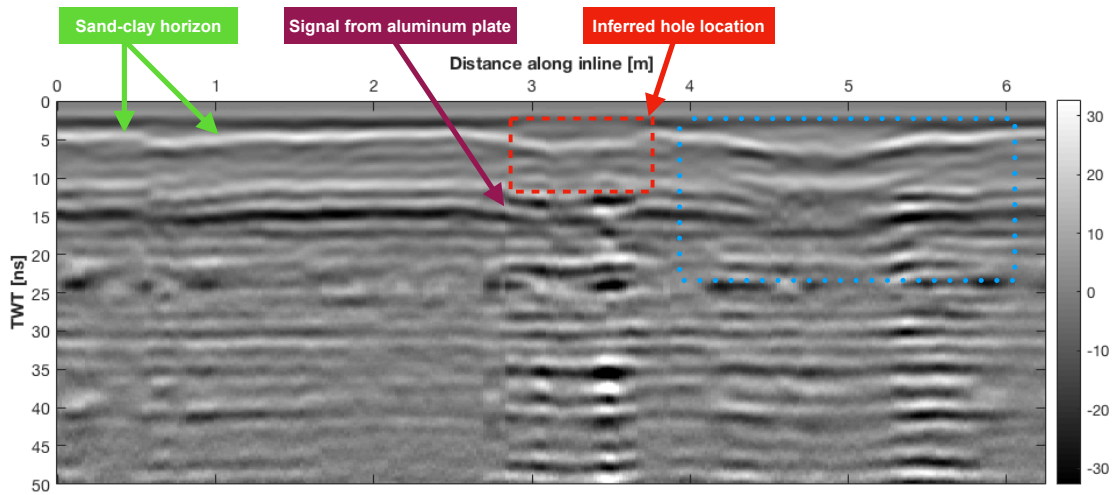
Figure 3.17: (a) Radargram showing extensive clipping. (b) Output from IVGPR's declipping routine with a fourth-order polynomial interpolation.

3.7 Background noise suppression

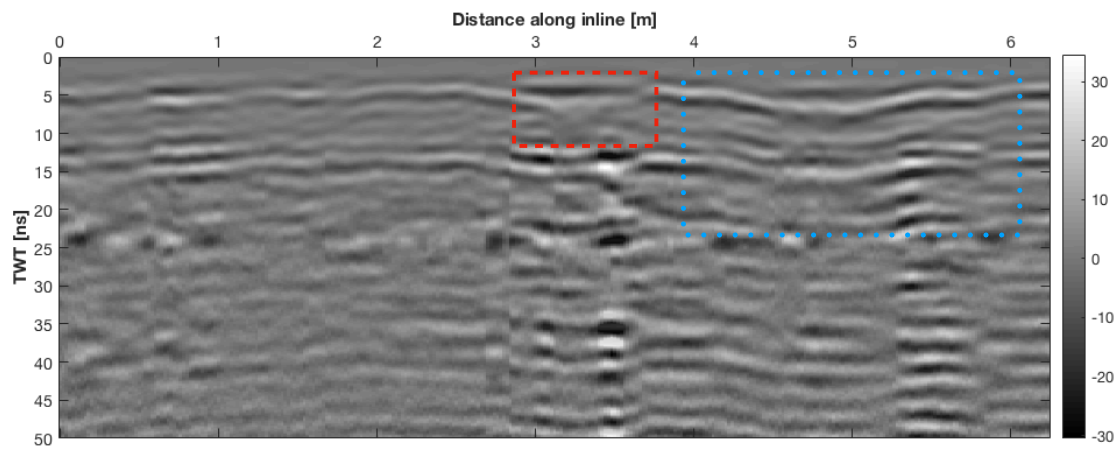
Background noise is an unwanted coherent signal generally introduced by properties inherent to the GPR equipment. For instance, when a GPR system sends the intended pulse, leftover currents continue to reverberate along the antenna's length sending secondary pulses (Rashed and Harbi, 2014). Also, the mismatch between the impedance of the antenna and that of the ground causes a backscattering of energy that can result in antenna ringing and, thus, horizontal striping across the B-scans. Common methods for background noise suppression (BNS) rely on *spatial averaging filters* and *low-cut wavenumber filters* (Annan, 1999). This work provides three instances of the former, which vary only in the techniques used to estimate the background noise matrix. The most elemental option, and the topic of this section, removes background noise using a spatial rectangular filter.

This approach is sometimes called mean trace subtraction (MTS), and it is the preferred option for most GPR surveyors (Rashed and Harbi, 2014). While appearing primitive, MTS typically produces satisfactory results at a minimal cost.

Since the MTS routine implements an unweighted averaging to estimate the background noise matrix, it does not discriminate between signal and noise. Nonetheless, MTS often yields satisfactory results while requiring only the specification of a one-dimensional window length. Figure 3.18 serves as evidence to the effectiveness of this routine. The survey line cut across a freshly filled hole (with dimensions of $0.9\text{ m} \times 0.9\text{ m} \times 0.6\text{ m}$) dug at about three meters from the left edge. The hole was filled with a random mixture of the excavated sediments placed over an aluminum plate, which measured $0.73\text{ m} \times 0.73\text{ m} \times 0.013\text{ m}$. A red dashed-square delineates the inferred location of the hole in the B-scan. The horizon separating a dry clean-sand (about 0.20 m thick) from a clay layer of unknown thickness is attributed to a strong event near the five nanosecond mark, see figure 3.18a. Note that this event appears to only suffer from a small deterioration in continuity near the hole. Additionally, seemingly coherent layering can be observed occurring within the inferred location of the hole in the data, i.e. the events from the 3 ns to the bottom 12 ns mark. As a result, an average surveyor could, understandably, misinterpret the untreated data. The sand-clay boundary in figure 3.18b exhibits clear discontinuities at the proper locations. One can also appreciate two flat events, corresponding to the top and base of the hole, capping a section that shows no obvious layering. This is consistent with the response one would predict. The blue dotted-square encloses several horizons that show an enhancement in continuity after MTS. In summary, this technique can effectively remove background noise and improve the interpretability of the data at little cost to the user. The author suggest the use of MTS prior to any alternative techniques.



(a)



(b)

Figure 3.18: Shows a field record before (a) and after (b) it was treated by the ATS algorithm. The background noise matrix was estimated using a windows of 101 elements. A red square encloses the inferred space-time confinement of a hole filled moments before acquiring the data. The blue square surrounds events that show enhanced continuity applying the ATS algorithm.

4. ADVANCED PROCESSING TECHNIQUES

4.1 Overview

Advanced processing routines can prove useful in complex georadar studies and when the state of a dataset does not permit a standard visual interpretation after basic processing. However, many routines in this class are notorious for their sensitivity to operator bias and often generate outputs that differ greatly from the input data (Annan, 1999). These are some reasons why numerous surveyors consider them to be unnecessary or even destructive. In my experience, these techniques have proven effective in enhancing the quality of data and extracting otherwise “hidden” information. The algorithms of this type included in IVGPR can be subdivided into image enhancement, attribute analysis, and alternative methods. We will explore these categories in more detail below. Since the particular techniques span several fields and vary greatly in difficulty, it would prove daunting to include detailed theoretical accounts of each. Instead, only those techniques expected to be of most value in a wider range of GPR applications were given extensive descriptions.

4.2 Image enhancement

Image enhancement in IVGPR focuses on suppressing noise and amending aliasing through the use of smoothing and interpolation techniques, respectively. Members of this module lack an explicit connection with the physical and practical mechanisms that produce the unwanted characteristics in GPR data that they aim to enhance. As a result, they might not be suited for every application. Additionally, these techniques are not meant to serve as ubiquitous replacements to careful data collection and survey planning. In the general case where the accuracy and reliability of an interpretation is central to a GPR study, surveyors should resort to image enhancement techniques only after exhausting the option of a more meticulous reacquisition of the data. We start by examining the adaptation of a

basic interpolation routine intended for cases where a refinement of the spatial sampling is necessary. This is followed by a comprehensive description of advanced multidimensional filtering techniques in IVGPR that are useful for smoothing and noise suppression.

4.2.1 Interpolating GPR data

Spatial aliasing in GPR data may result from unavoidable limitations in the time and cost of an acquisition. Interpolation offers a viable solution in such cases. This problem is common in seismic surveying and, in response, many sophisticated seismic interpolation algorithms have been proposed throughout the years (Liu and Fomel, 2011). The similarities between georadar and seismic reflection imaging allow many of these techniques to be effectively implemented in GPR studies. The interpolating options in IVGPR are limited to three rudimentary techniques developed after Press et al. (1992), and may not be well-suited for GPR data. These are a *bilinear interpolation*, a *bicubic interpolation*, and a *separable application of two one-dimensional polynomial interpolations*. No convincing evidence in support of these methods for treating GPR data was found throughout the literature and, therefore, users should weight their influence lightly during interpretation. Figure 4.1 shows the result of a bilinear interpolation along a depth slice at the maximum aggressiveness (or refinement) currently allowed in IVGPR. Notice how the character of the raw slice, 4.1a, is preserved, i.e. no detrimental artifacts or distortions are readily obvious. Refinement of this kind might be necessary when trying to apply novel algorithms with stringent sampling requirements to old datasets that fail to satisfy them. An example of such a routine is *vector migration*, which yields unaliased results for high frequency antennas only when a sufficiently fine sampling is used. Better suited interpolation algorithms (Feng et al., 2007; Martins et al., 2015) are currently being developed for this package and should be available in the near future.

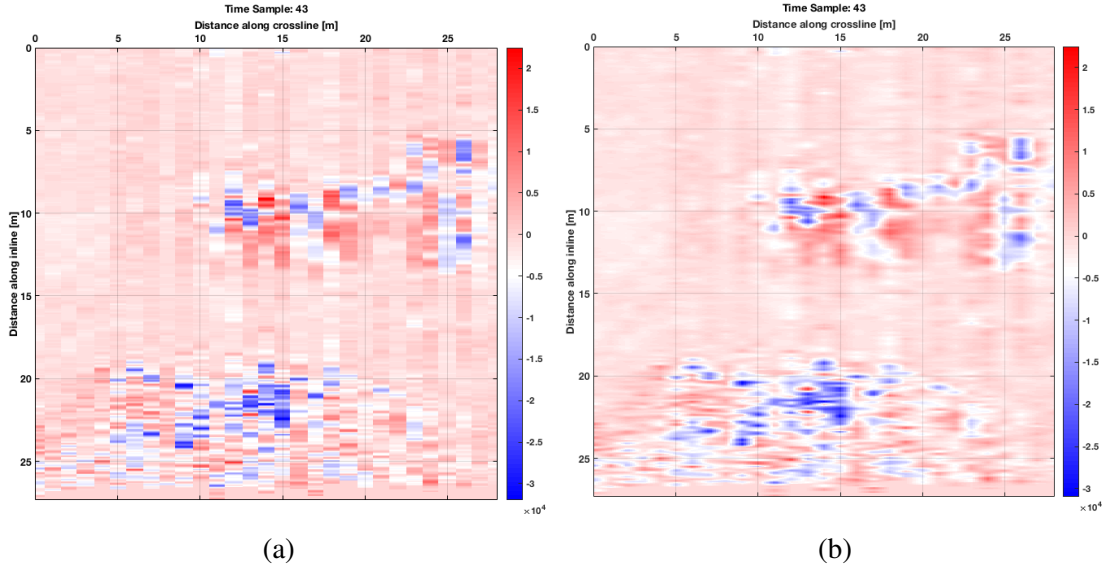


Figure 4.1: Shows a bilinear interpolation of a raw GPR slice (a) originally composed of 29 lines. The interpolated slice (b) contains 290 lines. No obvious evidence of damaging artifacts can be observed for this time slice.

4.2.2 Advanced filtering

Advanced filtering, as used in this dissertation, refers to any form of filtering deviating from the conventional approaches described in section 3.3. In particular, IVGPR offers an *elliptical Butterworth filter*, an *elliptical Tukey window* function, and a *Wiener denoising procedure*. The first two can be derived from their one-dimensional analogues by carrying out a trivial coordinate transformation. To verify this, let us consider the one-dimensional Tukey window in wavenumber domain as given by:

$$W(k) = \begin{cases} 1 & |k| < k_c \\ \cos^2\left(\frac{\pi[|k| - k_c]}{2\mu}\right) & k_c \leq |k| < k_c + \mu, \\ 0 & k_c + \mu \leq |k| \end{cases} \quad (4.1)$$

where k_c is the spatial cutoff frequency and μ controls the width of the transition band. In order to create a non-separable anisotropic point spread function from equation 4.1, we may start by considering the wavenumber k to be the polar radius $k(k_x, k_y)$, that is

$$k = k(k_x, k_y) = \sqrt{k_x^2 + k_y^2}. \quad (4.2)$$

Next, rather than representing the cutoff frequency with a fix value, i.e. a circle in polar coordinates, we may define the cutoff as the trace of an ellipse centered at the origin. Recall that the distance from the origin to the perimeter of an ellipse is, in general, dependent on direction. This means that $k_c \rightarrow k_c(\theta)$ such that

$$\begin{aligned} k_c = k_c(\theta) &= \frac{k_a k_b}{\sqrt{k_a^2 \sin^2 \theta + k_b^2 \cos^2 \theta}} \\ &= \sqrt{k_{c_x}^2 + k_{c_y}^2}, \end{aligned} \quad (4.3)$$

where k_a and k_b denote the horizontal and vertical semi-axes respectively. The rectangular components of $k_c(\theta)$ are indicated by k_{c_x} and k_{c_y} . Figure 4.2 depicts the configuration of the cutoff frequency described by equations 4.3. This, in turn, means that the components of the cutoff frequency satisfy the equation of an ellipse in \mathbf{k} -space, i.e.

$$\frac{k_{c_x}^2}{k_a^2} + \frac{k_{c_y}^2}{k_b^2} = 1. \quad (4.4)$$

After some algebra, we can arrive at the expression

$$k_c = k_a \sqrt{\frac{k_x^2 + k_y^2}{k_x^2 + k_y^2 / \epsilon}} \quad (4.5)$$

where $\epsilon = k_a/k_b$ characterizes the anisotropy of the filter (Bernstein et al., 2004). We then arrive at the desired equation, not shown here, by substituting 4.2 and 4.5 into 4.1.

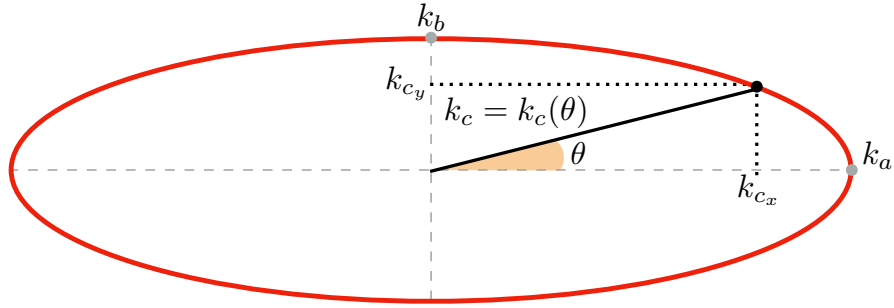


Figure 4.2: Extending the definition of cutoff frequency k_c as the perimeter of an ellipse.

A similar set of steps would allow us to derive the elliptical Butterworth transform function from its one-dimensional form, which would lead us to

$$H(k_x, k_y) = \frac{G_0}{1 + \left[\left(\frac{k_x}{k_{c_x}} \right)^2 + \left(\frac{k_y}{k_{c_y}} \right)^2 \right]^n}. \quad (4.6)$$

Here $H(k_x, k_y)$ is the amplitude response, n is the number of poles, and G_0 is the DC gain of the filter. A plot of the elliptical Butterworth transport function and elliptical Tukey window function is shown in figure 4.3. In this case the cutoff frequency was chosen to be the same in both direction, i.e. $\epsilon = 1$, producing radial non-separable filters. These filters may be designed to be highly directional, which could prove useful in treating coherent noise having some preferred orientation.

The third, and final, advanced filtering procedure implements an adaptive Wiener filter to reduce noise (Jin et al., 2003). For a brief account of the pertinent mathematical formalism, let us first denote the signal encoded in a GPR image at location (i, j) as:

$$R_{ij} = r_{ij} + n_{ij}, \quad (4.7)$$

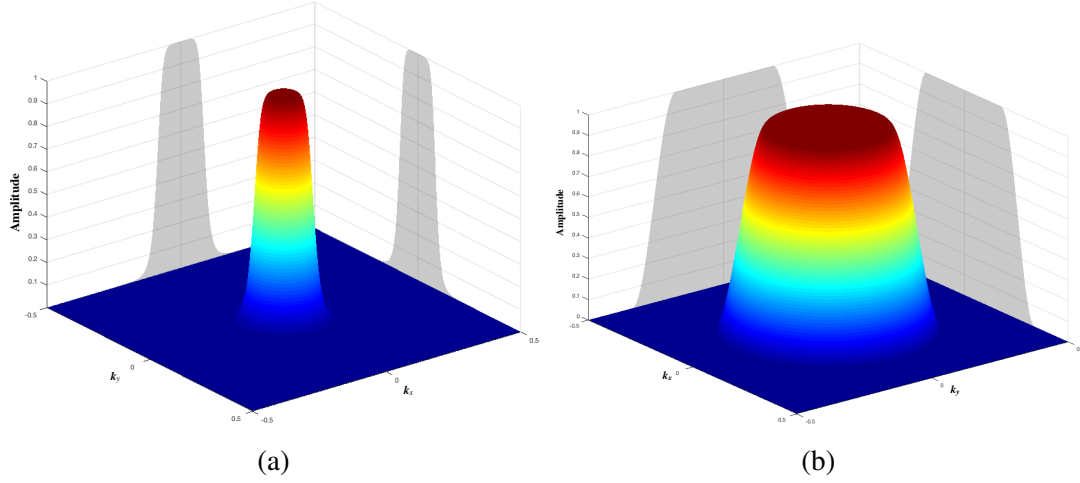


Figure 4.3: Two-dimensional Butterworth transfer function (a) and a two-dimensional Tukey window (b). The same cutoff frequency was chosen along each axial direction resulting in the isotropic response shown.

where r_{ij} represents the noise free component and n_{ij} is the noise. The estimate of r_{ij} obtained by the Wiener filter is

$$\hat{r}_{ij} = \mu_{ij} + \frac{\sigma_{ij}^2}{\sigma_{ij}^2 + \gamma_{ij}^2} (R_{ij} - \mu_{ij}), \quad (4.8)$$

for which

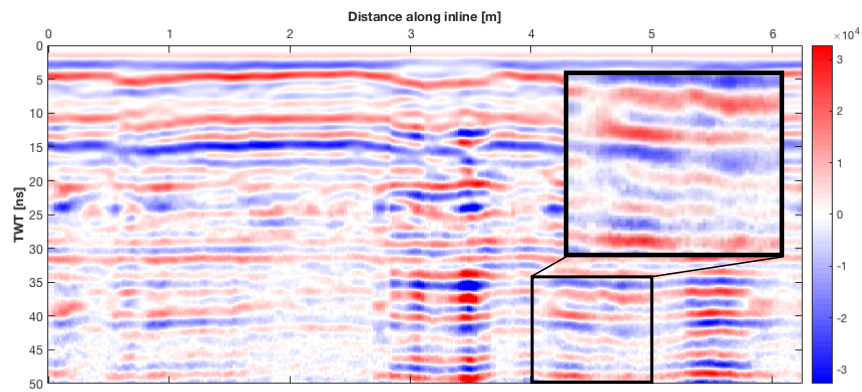
$$\mu_{ij} = \frac{1}{NM} \sum_{k,l \in \eta} R_{kl} \quad (4.9)$$

is the mean and

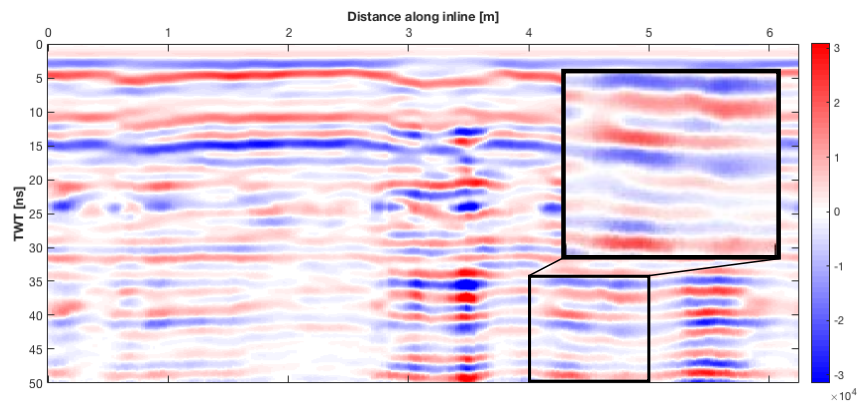
$$\sigma_{ij}^2 = \frac{1}{NM} \sum_{k,l \in \eta} R_{kl}^2 - \mu_{ij}^2 \quad (4.10)$$

is the variance computed in the $N \times M$ neighborhood of (i, j) denoted by η . If the noise variance at location (i, j) , i.e. γ_{ij}^2 , is not entered manually, then the algorithm estimates a homogenous noise variance by averaging the value of σ_{ij}^2 obtained at each location. In my experience, this routine has produced satisfactory reductions in noise without adversely

influencing other aspects of the image. Note that, unlike the frequency domain filters described above, the Wiener filter operates in object (or physical) space. It also produces an image processed by a two-dimensional moving average and a variance attribute as byproducts. Figure 4.4 shows a result from this procedure using an 11×11 window without a manual estimation of the noise variance γ_{ij}^2 . One can easily see that the filtered image shows an improved signal-to-noise ratio without serious losses in spatial resolution.



(a)



(b)

Figure 4.4: Raw radagram (a) treated by the Wiener filter denoising procedure in IVGPR using an 11×11 window. Two boxes are use to highlight how this filter can improve signal-to-noise without deteriorating the resolution of the image. A smaller window size may be used to retain more detail.

4.3 Attribute analysis

Attribute analysis refers to the extraction and examination of data-derived quantities ideally related to the structural and compositional characteristics of the subsurface. The aim is to assist the interpreter by bringing awareness to supplementary information encoded in the data that may remain undetected otherwise (Taner et al., 1979). Attributes can be categorized based on their dimensionality (i.e. one-, two-, or three-dimensions) as well as the domain in which they operate, e.g. physical space, complex space, and Fourier domain. The package currently offers twelve attributes; five are single-trace (or one-dimensional) and six are two-dimensional. Some are used in support of processing routines and have proper introductions elsewhere in this work. Here we begin by considering a set of textural attributes that, even though are seldom used in GPR studies, have recently shown promising results (Zhao et al., 2016). Next we will look at some techniques that may be useful for structural assessment and quality control in the early stages of a project. The final part of this section will focus on complex trace analysis, which has formed part of seismic prospecting for many decades.

4.3.1 Textural attributes

Textural evaluation has gained increasing approval in seismic stratigraphical studies as a viable tool for identifying structural fabrics (Gao, 2011). In spite of this, adaptations for georadar studies remain mostly unexplored throughout the literature (Zhao et al., 2016). This work offers further evidence on the effectiveness of textural attributes for GPR applications. The scheme I constructed quantifies four textures (i.e. *uniformity*, *contrast*, *homogeneity*, and *entropy*) utilizing the gray level co-occurrence matrix (GLCM). As shown in figure 4.5, the algorithm begins by creating a sub-image of size $M \times M$ centered at pixel (i, j) , where M is some odd integer such that $M \geq 3$. The sub-image is subsequently quantized into n gray levels and, in turn, a normalized omnidirectional GLCM is found.

Figure 4.5 illustrates how the 0-degrees, P^{0° , GLCM is created. It involves identifying all of the pixels with an equally valued neighbor to the right. Note also that the resulting P^{0° matrix is of size $n \times n$, which is a property of all gray level co-occurrence matrices. Furthermore, it is easy to see that $P^{180^\circ} = (P^{0^\circ})^T$. The normalized omnidirectional GLCM, C^0 , is found by summing the matrices corresponding to all relevant directions (i.e. $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ,$ and 315°) and normalizing the sum. Textural attributes are then computed from C^0 . Even though this process only requires two user-defined parameters (i.e. M and n), the overall computational cost is highly dependent on their values. A good rule of thumb is to start the search for the optimal parametrization with small, and inexpensive, values.

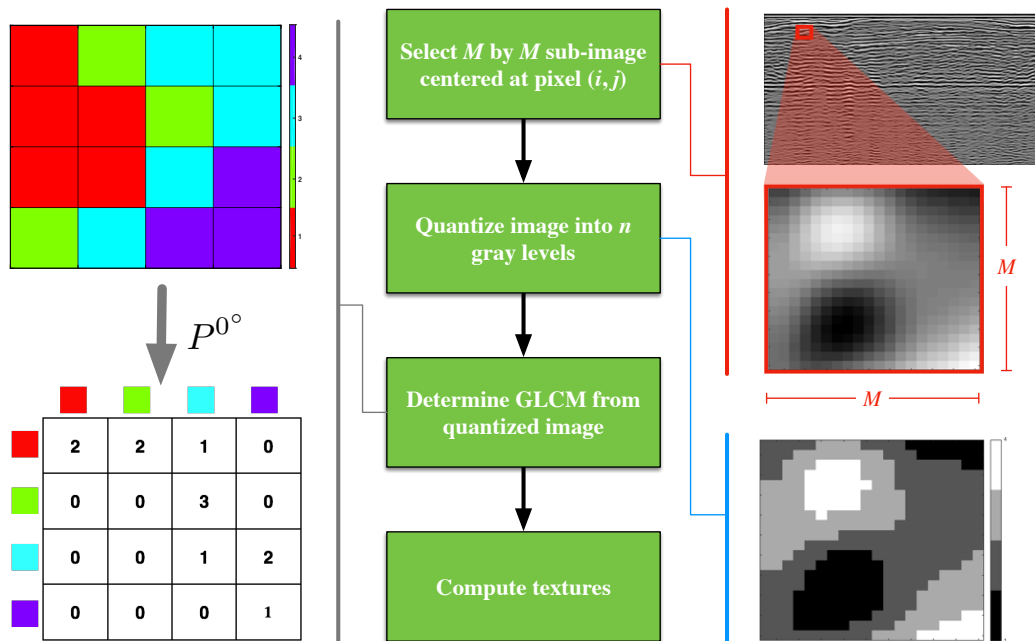


Figure 4.5: Approach used to compute textural attributes. A moving window of size M by M is centered at each pixel of an image and the resulting sub-images are quantized to a prescribed gray level. The figures on the right show a sub-image quantized to 4 gray levels. Next, a GLCM is computed for every direction. The images on the left show how the 0° GLCM is derived. The resulting matrices are summed and the sum is normalized to produce C^0 . Textural attributes can then be easily computed from C^0 .

4.3.1.1 Uniformity, contrast, homogeneity, and entropy

As implied by the name, uniformity (or energy) measures the extent to which an image is uniform. It reaches its maximum value when the image is constant, which corresponds to a C^o matrix with only one nonzero value. Therefore uniformity is also reflected by the sparsity of the GLCM. Areas of high geological complexity or low signal-to-noise may be detected by this attribute. The expression for uniformity at the (i, j) -pixel is

$$\sum_i \sum_j (C_{ij}^o)^2, \quad (4.11)$$

which produces values in the range $[0, 1]$.

Contrast characterizes the disparities in intensity between a pixel and its neighbors at each location along an image. The contrast attribute, contrary to uniformity, shows a minimum value of zero for unchanging amplitudes. Results from uniformity and contrast should be complementary in certain regions of an image and, therefore, may be used in combination to reinforce an observation. This attribute is found using

$$\sum_i \sum_j (i - j)^2 C_{ij}^o. \quad (4.12)$$

The spatial compactness of the distribution represented in C^o is calculated by homogeneity. A null value indicates a uniform distribution while unity symbolizes a diagonal C^o . One may compute homogeneity from the expression

$$\sum_i \sum_j \frac{C_{ij}^o}{1 + |i - j|}. \quad (4.13)$$

Entropy provides a statistical estimate of randomness in a sub-image. At its maximum, one expects the co-occurrence matrix to contain elements that are all equal. This corre-

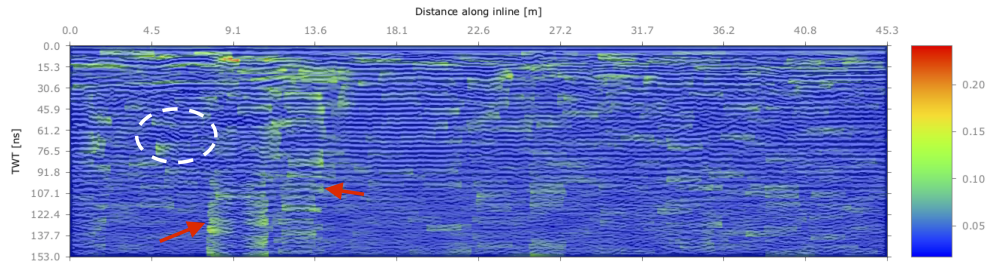
sponds to an equal probability of finding all possible amplitude combinations along the relevant directions. Entropy is found using the classical formula:

$$-\sum_i \sum_j C_{ij}^o \log C_{ij}^o. \quad (4.14)$$

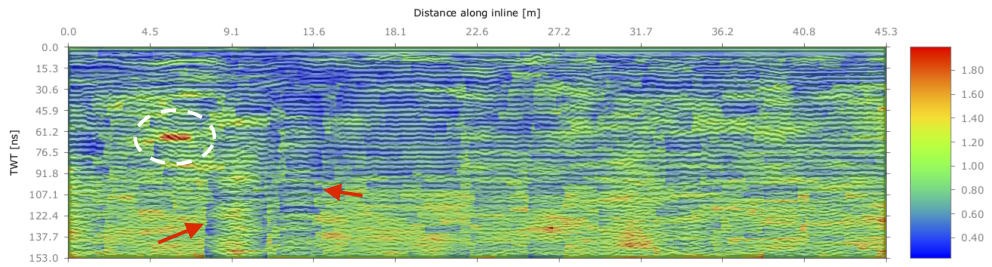
Figure 4.6 shows the four textural attributes overlaying the georadar used to extract them. The computation was carried out using a window size of 51 elements and 16 gray levels. The data is part of a study conducted by Zaremba et al. (2016) and can be freely obtained from the USGS website. Some features of interest include several vertical strips at late arrivals between the 9.1 and 13.6 meter marks (indicated by two red arrows) as well as some areas of high geological complexity highlighted by the measurements. Notice that all four attributes seem to indicate a set of characteristics for the vertical strips that agree with one another. For instance, these sections appear to have low complexity (or high uniformity) relative to adjacent regions. A feature of high contrast, encircled by dashed white lines, also shows consistent properties across all of the textural attributes. In summary, these attributes can help facilitate the identification and characterization of subsurface structures represented by the spatial and amplitude variations encrypted in the data.

4.3.2 Rapid assessment of structural and spectral content

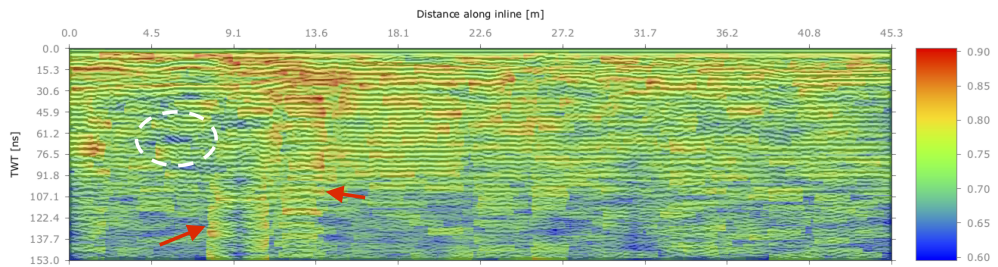
The *multitaper spectral analysis* (MSA), *gradient direction*, and *absolute gradient magnitude* attributes provide simple, yet effective ways to assess the spectral and structural content of GPR datasets. While IVGPR allows for similar evaluations using other techniques, experience has shown these to be the most effective. The first of these attributes estimates the power spectral density (PSD) along each trace. It relies on the proven adequacy of Prieto et al. (2009) mtspec library. The PSD is estimated using Thomson's



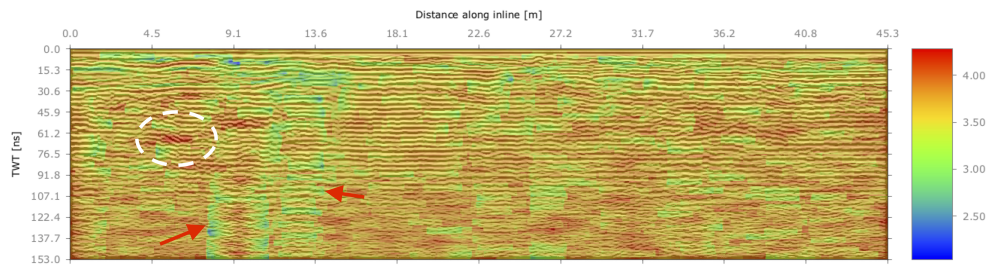
(a)



(b)



(c)



(d)

Figure 4.6: (a) Uniformity. (b) Contrast. (c) Homogeneity. (d) Entropy. One can easily discern regions of potential interest (with high and low values) highlighted by these attributes. There are a number of locations where these attributes reinforce one another.

multitaper method. Details on the theory related to PSD estimation, and those particular to mtspec, are beyond the scope of this work. The MSA attribute is simple to implement. The user needs only to specify a value for the time-bandwidth product and the number of tapers. One may also choose to operate along a single trace or the entire radargram. The latter paints a picture of the overall frequency content, which may highlight unwanted signals specific only to some traces. Figure 4.7a shows the power spectrum found for every trace of a radargram. This record was acquired using a 400 MHz antenna and, as a result, we see only an insignificant amount of energy at far higher frequencies. The effectiveness of this routine is further demonstrated in figure 4.7b. Here the PSD of a synthetic trace resulting from a linear combination of three sinusoids was analyzed using a TBP of 1.5 and 2 tapers. The frequencies of the sinusoids were 7.8, 15.6, and 125 MHz. As shown, the three sinusoids were successfully detected with an outstanding frequency resolution.

The gradient direction and absolute gradient magnitude, as presented here, stem from a widely popular image processing technique known as Sobel edge detection. The methodology is straightforward and does not require any parametrization. In general the gradient magnitude and gradient direction of a radargram, $R(x, t)$, are given by

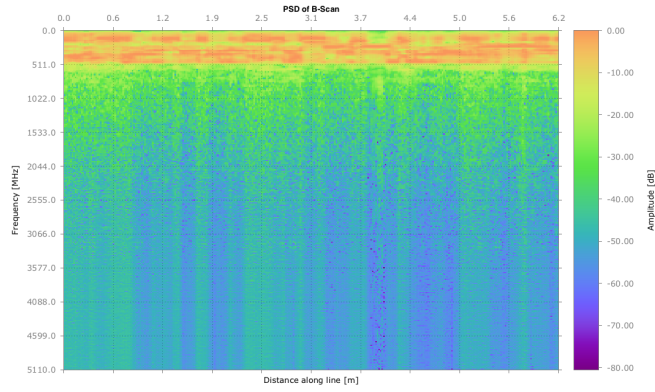
$$|\nabla R| = \sqrt{\left(\frac{\partial R}{\partial x}\right)^2 + \left(\frac{\partial R}{\partial t}\right)^2} \quad \text{and} \quad \theta_{\nabla} = \arctan\left(\frac{\partial R/\partial t}{\partial R/\partial x}\right), \quad (4.15)$$

respectively. The derivatives in equations 4.15 may be approximated as a two-dimensional convolution between $R(x, t)$ and the Sobel operators. Namely,

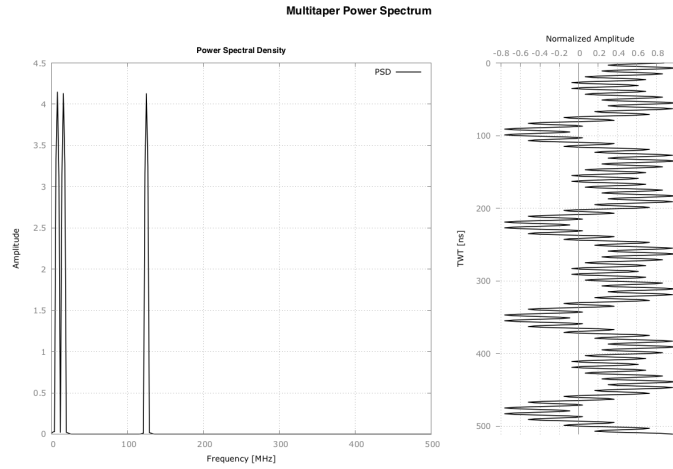
$$\frac{\partial R}{\partial x} \approx R * S_x$$

$$\frac{\partial R}{\partial t} \approx R * S_t$$

(4.16)



(a)



(b)

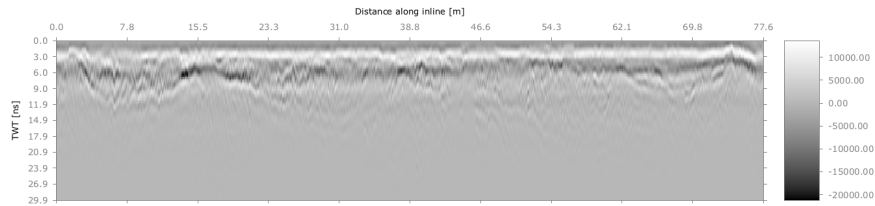
Figure 4.7: Shows a 2D multitaper spectral estimate for the entire radagram (a) and a 1D implementation for a synthetic trace created by combining three sinusoids with frequencies of 7.8, 15.6, and 125 MHz.

where we have that

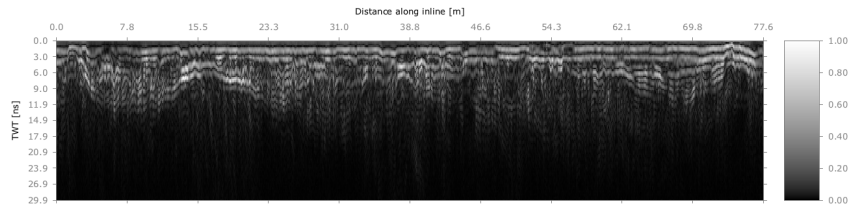
$$S_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \quad (4.17)$$

and $S_t = (S_x)^T$ (Jähne, 2005). Figure 4.8 shows how the two attributes can effectively emphasize structures not easy to distinguish in the raw data. As expected from equation 4.15,

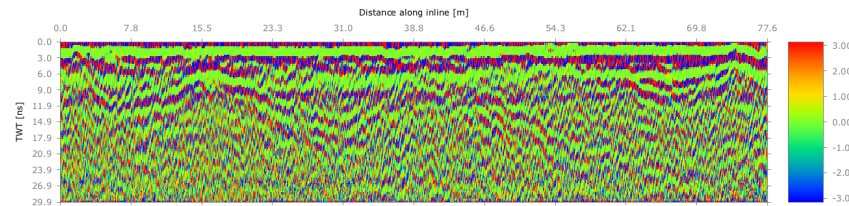
the values for gradient direction ranges from $-\pi$ to π and appear to be insusceptible to attenuation. The same is not true for gradient amplitude. Thus, a quick structural assessment without any additional processing is best obtained from the former.



(a)



(b)



(c)

Figure 4.8: (a) Shows a raw radargram along with the extracted gradient amplitude (b) and direction (c) attributes.

4.3.3 Complex trace analysis

Complex trace analysis (CTA) centers around the idea that transforming a trace into a discrete-time analytic signal permits the extraction of valuable attributes with local tempo-

ral significance (Taner et al., 1979). The use of CTA has prevailed in exploration seismology since first advocated by Taner et al. (1979). It also received early adoption in popular GPR processing packages (Annan, 1999). The set of CTA attributes in IVGPR is limited to instantaneous frequency, instantaneous phase, and envelope (or reflection strength). Other attributes, e.g. apparent polarity and weighted average frequency, may be derived from these three and will form part of future releases.

The particular transformation technique in IVGPR was designed after Marple (1999). There are at least two reasons for this. First, this procedure primarily relies on the FFT and, therefore, is computationally efficient. Secondly, and unlike some of the alternatives, each resulting complex-valued signal, $\{z(t)|t = n\Delta t \text{ and } n = 0, \dots, N - 1\}$, is orthogonal within its support, $T(= (N - 1)\Delta t)$, and has a real component equal to the input trace. We may write the transformation of a trace, $r(t)$, after Boashash (1992a) as

$$\begin{aligned} z(t) &= r(t) + i\mathcal{H}[r(t)] \\ &= A(t)e^{i\phi(t)}, \end{aligned} \tag{4.18}$$

where $A(n)$ is the reflection strength, $\phi(n)$ is the instantaneous phase, and \mathcal{H} is the discrete Hilbert transform. Here $z(n)$ satisfies the conditions

$$\text{Re}\{z(t)\} = r(t) \tag{4.19}$$

and

$$\sum_{t=0}^T r(t)\mathcal{H}[r(t)] = 0. \tag{4.20}$$

In practice the transformation is performed in three steps. For the first step, the trace is transformed into frequency domain by applying a FFT such that $r(t) \xrightarrow{\text{FFT}} \hat{r}(\omega)$, where $\omega = m\Delta f$ and $m = 0, \dots, N - 1$. Next we reshape $\hat{r}(\omega)$ into the FFT of $z(t)$, i.e. $\hat{z}(\omega)$,

by letting

$$\hat{z}(\omega) = \begin{cases} \hat{r}(0), & \text{for } m = 0 \\ 2\hat{r}(m\Delta f), & \text{for } 1 \leq m \leq \frac{N}{2} - 1 \\ \hat{r}\left(\frac{N}{2}\Delta f\right), & \text{for } m = \frac{N}{2} \\ 0, & \text{for } \frac{N}{2} + 1 \leq m \leq N - 1 \end{cases}. \quad (4.21)$$

Finally, the desired discrete time analytic signal is obtained by implementing an inverse FFT, that is $\hat{z}(\omega) \xrightarrow{\text{IFFT}} z(t)$.

4.3.3.1 Envelope, instantaneous phase, and instantaneous frequency

The envelope attribute is formally defined as the norm of the complex trace $z(t)$, i.e.

$$\begin{aligned} A(t) = |z(t)| &= \sqrt{\text{Re}[z(t)]^2 + \text{Im}[z(t)]^2} \\ &= \sqrt{r(t)^2 - (\mathcal{H}[r(t)])^2}. \end{aligned} \quad (4.22)$$

The extrema in $A(t)$ may not coincide with those of $r(t)$ and, therefore, could fail to represent the true reflection strength of events along a trace (Taner et al., 1979). High envelope amplitudes may identify major changes in lithology. Variations in lateral amplitude could be linked to fluctuations in bed thickness or the presence of faulting and fracturing. Envelope is also useful to assess the resolution of the data, and it is commonly used to estimate an amplitude decay curve.

Instantaneous phase represents the time dependent polar angle of the $z(t)$ phasor in the complex plane and is defined as

$$\begin{aligned} \phi(t) &= \tan^{-1} \left[\frac{\text{Re}[z(t)]}{\text{Im}[z(t)]} \right] \\ &= \tan^{-1} \left[\frac{\mathcal{H}[r(t)]}{r(t)} \right]. \end{aligned} \quad (4.23)$$

Since the instantaneous phase is independent of reflection strength, it can be used to analyze the continuity of coherent events. Therefore, it is adequate for showing faults, pinch-outs, and other types of discontinuities in the record.

The instantaneous frequency is most conveniently defined as the temporal derivative of the instantaneous phase, that is

$$f_i(t) = \frac{d}{dt}\phi(t). \quad (4.24)$$

This attribute is useful for identifying subsurface features that respond selectively to the frequency content of the incident pulse (Annan, 1999; Jol, 2008). For example, thin beds may have tuning frequencies that overlap with the spectrum of the signal and would, therefore, produce enhanced responses. The derivative in equation 4.24 is commonly approximated via numerical differentiation, which is inherently sensitive to random noise. In particular, IVGPR uses a central difference scheme to estimate $f_i(t)$. An extensive, yet instructive analysis of theoretical and practical details pertaining to the notion of instantaneous frequency is given in Boashash (1992a) and Boashash (1992b).

Figure 4.9 shows the CTA attributes in IVGPR computed from a shallow B-scan acquired over an urban area predominated by horizontally lying reflectors. The extraction does not require any parametrization. The result for instantaneous frequency shown in figure 4.9d, however, was treated with a Plank taper to overcome the masking of internal features caused by sporadic high amplitude events near the edges. The envelope (figure 4.9b) and the instantaneous phase (figure 4.9c) appear to adequately emphasize abrupt changes in reflection strength and the continuity of events, respectively. The instantaneous frequency shows sharp amplitude variations along horizons that could indicate a significant change in the nature of these boundaries. Interpreters are advised to consider these attributes collectively in order to obtain better results. One such case is given by Orlando

(2002), who used all three attributes to characterize the distribution of fluids around a water table.

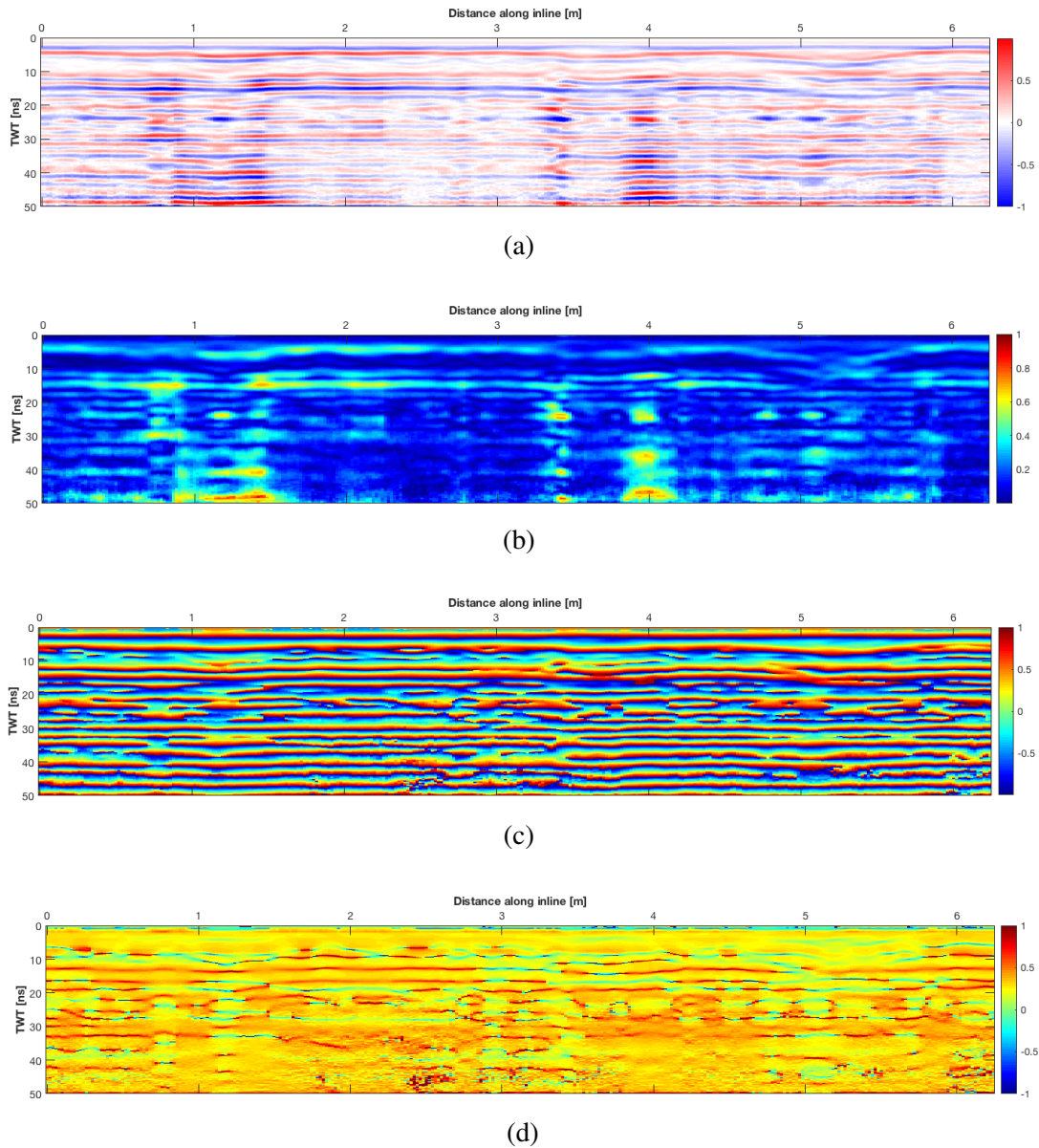


Figure 4.9: Shows the complex trace analysis attributes included in IVGPR. These consists of envelope (b), instantaneous phase (c) and instantaneous frequency. The attributes were extracted from the raw B-scan shown in (a). The amplitudes for each plot was normalize by the absolute maximum. Additionally, the instantaneous frequency was treated using a Plank window to reduce the effect of high amplitudes at the edges of the data.

4.4 Alternative techniques

The final set of advanced routines involves methods that offer an alternative way to implement otherwise basic procedures. Aside from improved performance, several of these routines promise a reduction in operator bias. Their peculiar theoretical formalities differ from that of conventional techniques and, therefore, might discourage their adoption. We begin this section exploring the topic of *spectral balancing*, which provides a substitute to a traditional *deconvolution*. Next, we study the advantages of replacing a moving average background correction scheme with a method proposed by Rashed and Harbi (2014). Lastly, a technique developed after Sabbione and Velis (2010) to automatically detect first arrivals is examined.

4.4.1 Spectral Balancing

To electromagnetic pulses with high dominant frequencies, e.g. the 10 to 1000 MHz range of geological radar sounding, the earth's preferential attenuation of high frequency content is intensified, since the corresponding wavelengths fall closer to the scale of inhomogeneities (Cassidy and Jol, 2009). Consequently, the spectral bandwidth of a source wavelet becomes narrower in proportion to the traveled distance. In the time domain a diminished spectral content unfolds as a deterioration of the temporal (vertical) resolution. The most common processing techniques to ameliorate this degradation in resolution are *deconvolution* and *spectral balancing*. The former, however, is notoriously ineffective in GPR applications, which is credited to the discrepancies between the character of GPR signals and the underlying assumptions of seismic deconvolution (Neto and de Medeiros, 2006; Belina et al., 2009; Tronicke et al., 2015). By contrast, Belina et al. (2009) showed that spectral balancing techniques (i.e. *spectral whitening* and *spectral blueing*) can adequately enhance the spectral bandwidth of GPR data without the high computational cost and laborious parametrization required by deconvolution. A more recent development by

Sajid and Ghosh (2014) suggests an alternative technique based on cascaded dipole filters. This method, known as *differential resolution*, compares to spectral whitening in simplicity and performance while requiring even fewer user-defined parameters. Tronicke et al. (2015) tested the effectiveness of spectral whitening, spectral blueing, and DR on synthetic and field data. Even though the authors' analyses are primarily qualitative, their results demonstrate that spectral balancing can enhance the frequency content and, consequently, the vertical resolution of GPR data. It is important to acknowledge that spectral balancing techniques lack rigorous mathematical developments to support and describe any perceived effectiveness (Yilmaz, 2001; Belina et al., 2009; Tronicke et al., 2015). As a result, spectral balancing is treated with suspicion by the GPR community and rarely appears in processing softwares (Tronicke et al., 2015). This section aims to provide a better understanding on how these algorithms can be implemented and further demonstrate their effectiveness in treating synthetic and field GPR data.

4.4.1.1 *Time-variant spectral whitening*

The implementation of time-variant spectral whitening (TVSW) in IVGPR was molded after the technique described by Yilmaz (2001). An overview of the TVSW workflow is given in figure 4.10. In the first stage the user defined the frequency band of interest, $\Delta f (= f_{lp} - f_{hp})$ and the number of distinct bandpass filters, N , to fit within this region. Each filter is applied to the original record, independently, resulting in N sub-records. Frequency dependent properties in the original data (e.g. attenuation) will differ between the sub-records, since their frequency content corresponds to those of the filters used to generate them. The next stage is to design and implement N temporal gain functions. Ideally, these gain functions should account for the different levels of attenuation in each sub-record and minimize the amount of user bias. The implementation in IVGPR achieves both of these goals using the programmed gain routine, refer to section 3.5.3. In the final

step, the sub-records are combined to yield the corrected data.

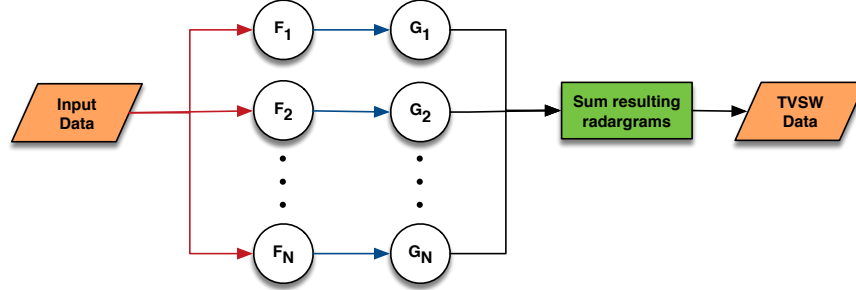


Figure 4.10: Time-variant spectral balancing algorithm. The input data is Fourier transformed (red arrows) and filtered. This produces N sub-records with different frequency content. After an inverse Fourier transformation (blue arrows), each sub-record is gained using a programmed gain. The gained sub-records are combined to produced the TVSW data.

In IVGPR, the bandpass filters are designed using the *windowed-sinc* temporal kernel, see section 3.3.1. The symmetry in the transition-band of windowed-sinc filters allows for an exact decomposition of a single bandpass filter into a set of bandpass filters overlapping at common cutoff frequencies. Thus, if the desired frequency range extends from f_{hp} to f_{lp} , the decomposition of a single bandpass filter into N filters is given by

$$\tilde{B}_{[f_{hp}, f_{lp}]} = \sum_{i=1}^N \tilde{B}_{[f_{hp}+(i-1)\alpha, f_{hp}+i\alpha]}, \quad (4.25)$$

where $\alpha(= \Delta f/N)$ is the bandwidth of the filters and $\tilde{B}_{[f_{hp}, f_{lp}]}$ represents the amplitude response of a bandpass filter with cutoff frequencies f_{hp} and f_{lp} . Figure 4.11 shows the amplitude responses of three adjacent bandpass filters and their sum. Note that the sum satisfies equation 4.25, i.e. it is equivalent to a bandpass filter with cutoff frequencies f_{hp} and f_{lp} . Another notable, and perhaps alarming, feature in figure 4.11 is the substan-

tial amount of overlapping between adjacent transition zones. Recall, however, that the width of the transition band (and the amount of overlapping) is inversely proportional to the width of the windowed-sinc kernel used to generate each filter. Thus, any adverse effect could be brought down to a negligible level by making the temporal filter kernel sufficiently large. A relationship between the amount of overlapping and the quality of the technique is not explored in this work. However, a qualitative and quantitative analysis could be achieved from judicial implementation of the TVSW algorithm and the mathematical formalism described in this document.

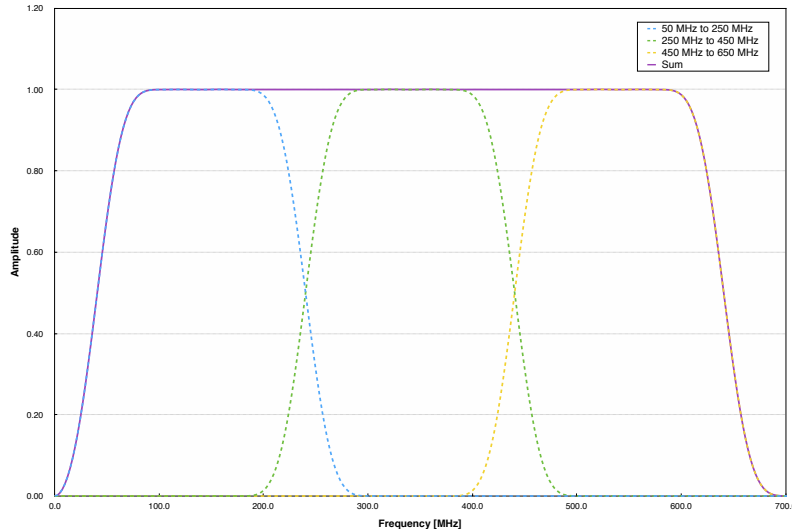


Figure 4.11: Three adjacent bandpass filters (dashed curves) generated by setting $f_1 = 50$ MHz, $f_2 = 650$ MHz, and $N = 3$. Their superposition (solid purple curve) is equivalent to a single bandpass filter with cutoff frequencies f_{hp} and f_{lp} .

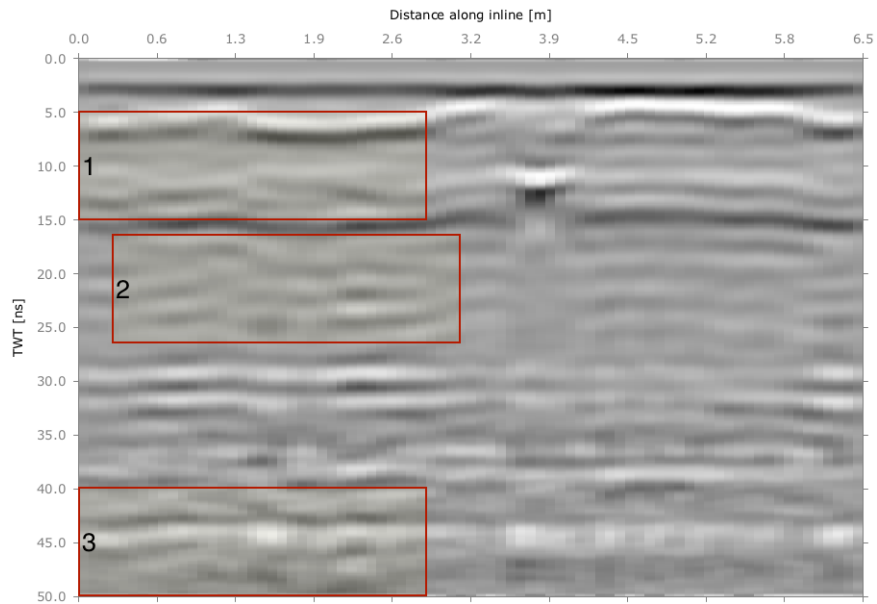
An effective application of TVSW should enhance the vertical resolution of the data without introducing artifacts. Additionally, the amplitude spectrum of the record should become broader and more homogeneous. Other indications of a successful implementation include the emergence of deeper features (where the attenuation is greatest) and increased

continuity of events. This algorithm is highly sensitive to parametrization. For example, improper selection of the frequency range can mute important features and/or amplify noise. A good rule of thumbs is to pick a frequency range that goes above and below the nominal frequency by 200–300 MHz. Users should expect the best results to come only after carefully selecting near optimal parameters tailored to their data.

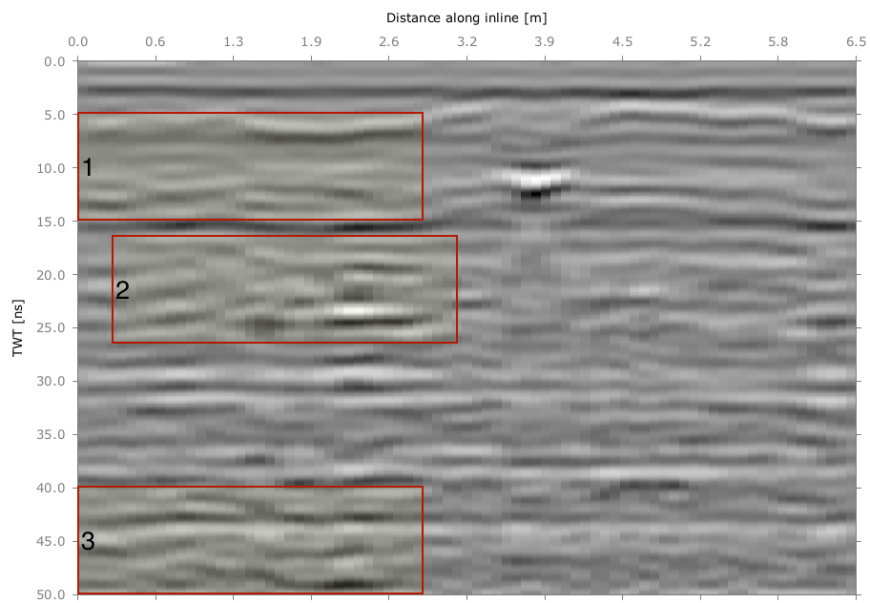
A satisfactory implementation of the algorithm is shown in figure 4.12. This profile was acquired using a co-polarized 500 MHz GSSI antenna–receiver pair. The near-surface geology of the surveyed area consisted primarily of a thin (10–20 cm) layer of clean sand overlaying a thick clay. An aluminum plate was buried near the 3.9 m mark, which is indicated by a bright reflector at about 10 ns. Figures 4.12a and 4.12b show the raw and balanced profiles respectively. The TVSW algorithm was applied using five filter-gain pairs to treat the frequency content ranging from 200 MHz to 800 MHz. Three boxes, label 1–3, highlight events at different depth ranges where the resolution, continuity, and emergence of events can be appreciated. In the shallowest box, the reflectors became noticeably thinner and more continuous after the correction. The second depth range shows some new features resulting from enhancements of the vertical and lateral resolution of the data. Lastly, the deepest box shows a significant increase in the number of visible structural features. As shown by this example, this algorithm can enhance the interpretability of the data while extending its usefulness to a deeper range.

4.4.1.2 Frequency domain spectral whitening

Unlike the time-variant spectral whitening (TVSW) algorithm, frequency domain spectral whitening (FDSW) operates directly on the amplitude spectrum of the data. A radar-gram successfully treated by FDSW should display an improved signal-to-noise ratio as well as enhancements to the vertical and structural resolution (Lee, 1986). In addition, the amplitude spectrum, as with TVSW, should become broader and more homogenous. An



(a) Radargram of unprocessed data.



(b) Radargram treated only by TVSW.

Figure 4.12: (a) Raw radargram with three squares indicating regions of different depth and characteristics. (b) Shows the result after treating the raw radargram with the TVSW. The results within the three regions were used to qualitatively assess the performance of this routine.

amplitude spectrum corrected by FDSW is shown in figure 4.13. Aside from an obvious broadening, we also see that the peak frequency shifted to a higher value that better agrees with the nominal frequency of the antenna (i.e. 500 MHz). It is important to note that the FDSW technique does not correct the phase of the data. In fact, the validity of the technique relies heavily on the assumption that the amount of distortion in the recorded phase spectrum is negligible. Lee (1986) thoroughly describes the mathematical formalism for FDSW and draws a number of quantitative comparisons to seismic deconvolution. Since the adaptation of FDSW in IVGPR follows the work in Lee (1986), a similar derivation is included in this work.

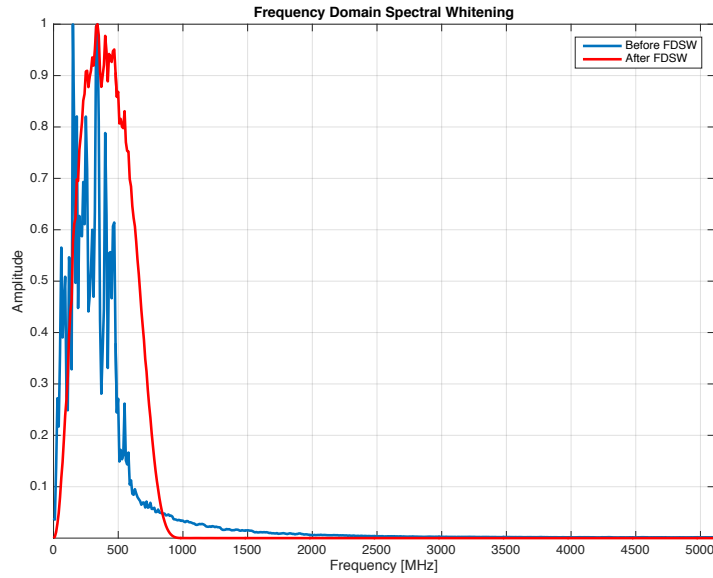


Figure 4.13: Amplitude spectrum of a GPR trace before (blue) and after (red) FDSW. The spectral response appears broader and its peak frequency has shifted to a higher value.

Consider a trace T such that its Fourier transform is given by $\mathcal{F}(T) \equiv \tilde{T} = \tilde{A}e^{i\phi}$, where \tilde{A} is the amplitude spectrum of the trace and ϕ is the corresponding phase spectrum.

Spectral whitening in the frequency domain can then be defined as

$$\tilde{T}_{\text{sw}} = \tilde{A}^\alpha e^{i\phi}, \quad (4.26)$$

where \tilde{T}_{sw} is Fourier transform of the corrected GPR trace, and α is an arbitrary constant ranging from zero to one. We can rewrite equation 4.26 in the form

$$\begin{aligned} \tilde{T}_{\text{sw}} &= \tilde{A}\tilde{A}^{\alpha-1}e^{i\phi} \\ &= \tilde{F}\tilde{A}e^{i\phi} = \tilde{F}\tilde{T}, \end{aligned} \quad (4.27)$$

and view $\tilde{F}(= \tilde{A}^{\alpha-1})$ as a filter applied to \tilde{T} .

It is instructive to note that, in the limit as $\alpha \rightarrow 0$ we have $\tilde{F} \rightarrow \tilde{A}^{-1}$. If the subsurface reflectivity is assumed to be random, the amplitude spectrum of the source wavelet can be approximated by that of the trace. Thus, the filter derived by FDSW resembles that of a spiking deconvolution in the limit as $\alpha \rightarrow 0$. A key parameter controlling the performance of FDSW is α , which thresholds the amount of whitening. A bandpass filter is commonly applied to ensure that the effect of the technique is contained within a reasonable frequency range, which could be determined from the properties antenna-receiver pair used and the geology of the medium. As with TVSW, an optimized parametrization is, in general, only evident after judiciously testing a range of possible values.

To see the effect of FDSW on physical data, we consider its application to the profile treated by TVSW in § 4.4.1.1. Two identical portion of the radargram, before and after applying the FDSW routine, are shown in figure 4.14. Boxes, labeled 1–3, enclosed three distinct depth regions. We first observe that the reflectors in all three boxes appear thinner and more continuous. While the shallowest and deepest boxes (i.e. 1 and 3) retained all of the major events after the correction, some events in box 2 became difficult to discern. Perhaps most striking, the density of visible structural features in the deepest box greatly

exceeds that of the other two depth ranges. We see a number of distinguishable thin layers that cannot be detected in the original record. Since a description of the geology beyond 10 ns was not acquired, the validity of this particular implementation cannot be confirm. However, this technique has been shown to reveal details in agreement with exposed geological features (Belina et al., 2009). Also, its simplicity alone make it a attractive way to seek improvements in the resolution of a record.

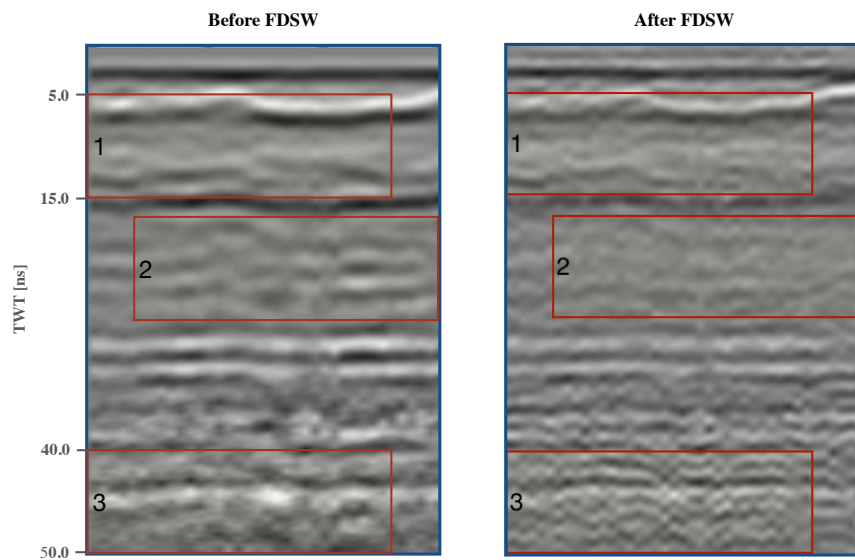


Figure 4.14: Identical sections of a profile before (left) and after (right) FDSW. Three depth regions are highlighted to evaluate the effectiveness of the routine. Clear enhancements to the continuity and resolution of reflectors can be observed after FDSW (right).

4.4.1.3 Differential resolution

Differential resolution (DR) offers another parametrization-lean approach to implement spectral balancing. This technique was a recent development by Sajid and Ghosh (2014) originally intended for seismic data, but Tronicke et al. (2015) successfully used the DR algorithm to treat synthetic and field GPR records. In their work, Tronicke et al.

(2015) found that DR produced an artifact free output when treating a basic synthetic profile. By contrast, visible artifacts appeared in the record treated by FDSW. One major drawback of DR, with no mention in previous works, is its high sensitivity to noise. This results from the reliance on high order derivatives (e.g. 4th and 6th) to enhance high frequencies in conjunction with the susceptibility to noise of numerical differentiation. In fact, all tolerable applications of DR to real data using IVGPR have only resulted in minor improvements to the vertical resolution.

In order to clearly understand the steps of DR, we begin by considering the equation describing a corrected trace as defined by Sajid and Ghosh (2014). Namely, we have that

$$R = Y + Y_s - Y^{II} + Y^{IV} - Y^{VI}, \quad (4.28)$$

where R corresponds to the normalized corrected trace, Y represents the normalized input trace, the subscript s indicates the smoothing of Y , and the roman numerals correspond to the orders of numerical differentiation applied to Y . The use of capital letters implies a normalization by the median of the absolute value, e.g.

$$Y = \frac{y}{\text{median}(|y|)}, \quad (4.29)$$

where y would be the input trace. Each derivative in equation 4.28 is obtained using alternating successive applications of forward and backward differences. The smooth trace is added to enhance low frequencies. While smoothing in IVGPR can be done in a number of ways, Sajid and Ghosh (2014) suggest a 10 pass convolution of Y with the kernel $[1, 2, 1]$.

The terms in equation 4.28 for a physical trace are shown in figure 4.15. The trace was not treated prior to applying DR and, as a result, the high order derivatives show sporadic

oscillations in the later sections of the trace where the signal-to-noise ratio depreciates. Therefore, it is suggested to implement a suitable denoising routine prior to applying this spectral balancing technique. After extensive testing, the adaptation of DR suggested by

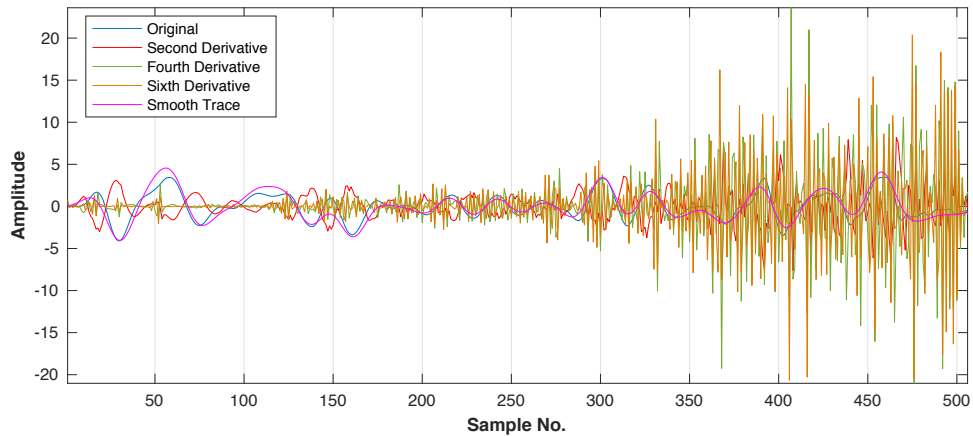
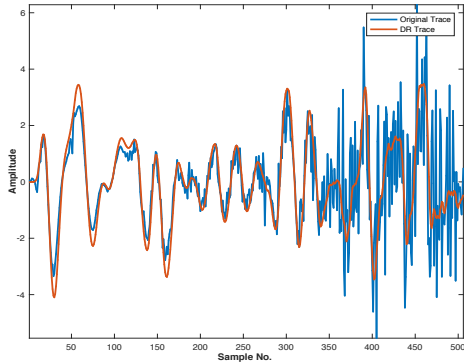


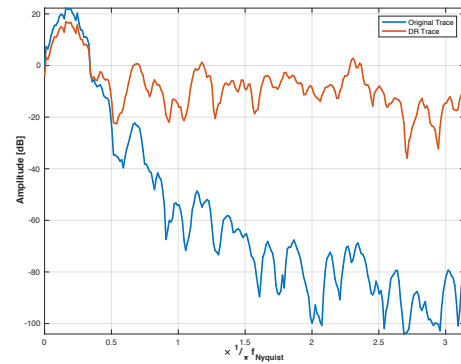
Figure 4.15: Plot of the terms in equation 4.28. The sporadic behavior of high order derivative is likely due to noise in the original trace.

Sajid and Ghosh (2014) failed to adequately treat most GPR field records. As evidenced by figure 4.16b, the DR algorithm can generate a sensible correction to the power spectrum of a trace. Namely, the lower frequency content remained virtually unchanged while the decay in higher frequencies was corrected. However, figure 4.16a shows that this boost in high frequency content gave the trace a noisy appearance and rendered much of the later section uninterpretable. In order to achieve better results, a bandpass filter was added as an additional step following the original implementation. With this modification, the DR algorithm was applied to the GPR section shown in figure 4.16c. This profile was chosen to allow a direct comparison between the DR and the two contrasting spectral balancing techniques in IVGPR, i.e. time-variant spectral whitening and frequency domain spectral whitening. While the result (shown in figure 4.16d) shows some clear improvements to

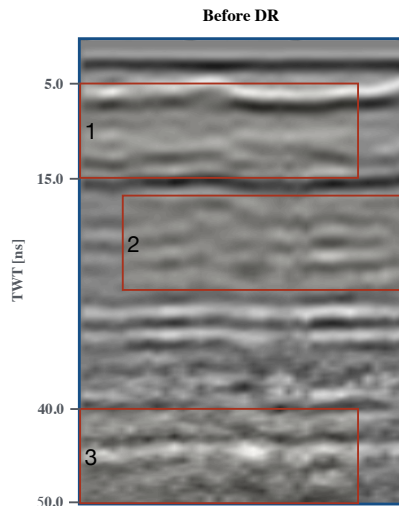
the vertical and structural resolution, the technique appears less effective than its spectral balancing counterparts. Nonetheless, this could be due to less artifacts being generated, which could be validated with a controlled experiment.



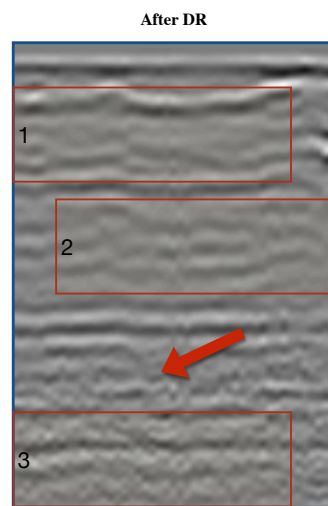
(a) Trace before and after DR.



(b) Power spectrum before and after DR.



(c) Raw B-scan.



(d) B-scan treated only with DR.

Figure 4.16: (a) Shows the output trace (blue) generated after treating the original trace (orange) with DR. The power spectrum of both traces is shown in (b). Figures (c) and (d) show a GPR section before and after DR, respectively. Three boxes enclose different depth ranges used to evaluate the technique. A red arrow indicates improvements in resolution.

4.4.2 Background matrix subtraction

Recently, Rashed and Harbi (2014) introduce a background noise suppression (BNS) scheme that, unlike conventional counterparts, performs a highly discriminatory averaging in order to estimate the background noise. In particular, the so-called background matrix subtraction (BMS) method replaces the single step unweighted average of mean trace subtraction (MTS), cf. § 3.7, with a multi-step adaptive weighted average. The weights should, ideally, scale each sample in a way that reflects the likelihood of these to constitute background noise. After extensive testing, however, this technique failed to produced significant gains while it added two user-defined parameters that proved difficult to adjust in order to achieve a satisfactory performance. Nonetheless, the results presented by Rashed and Harbi (2014) in combination with the principal of adaptive weighting to better estimate background noise motivated the inclusion of this procedure in IVGPR.

To compute the weights, the algorithm first determines the alpha-trimmed mean (Bednar and Watt, 1984) of the elements inside a horizontal window of length N centered at each data point (i, j) , where N is an odd integer. The alpha-trimmed mean of a set $W_{ij}(= \{x_k | k = 1, \dots, N\})$ is found by first sorting the x_k 's in ascending order and, in turn, using the formula

$$m_{ij}(\alpha) = \frac{1}{N - 2[\alpha N]} \sum_{\ell=[\alpha N]+1}^{N-[\alpha N]} x_{\ell}. \quad (4.30)$$

Here the set $\{x_{\ell} | \ell = 1, \dots, N\}$ represents the sorted W_{ij} , the symbol $[.]$ indicates the greater integer function, and $0 \leq \alpha < 0.5$. The number of edge samples discriminated in equation 4.30 is controlled by α . When $\alpha = 0$, equation 4.30 is equivalent to the arithmetic

mean used by MTS. The next step involves computing the weights from:

$$w_\ell = \begin{cases} |\underline{x}_\ell - m_{ij}(\alpha)|^{-s} & \text{if } \text{sign}(x_\ell) = \text{sign}(m_{ij}(\alpha)) \\ 0 & \text{otherwise} \end{cases}, \quad (4.31)$$

where s is an input parameter use to control the aggressiveness of the weights. Equation 4.31 prevents any element with a polarity opposite to that of $m_{ij}(\alpha)$ from contributing to the background noise matrix. In some cases, specially when the data is severely saturated, the difference $\underline{x}_\ell - M_{ij}(\alpha)$ can equal zero and lead to an undefined w_ℓ . IVGPR deals with this problem by issuing a warning and replacing the zero difference with a small value. The algorithm continues by normalizing the weights as

$$\hat{w}_\ell = w_\ell / \sum_{j=1}^N w_j, \quad (4.32)$$

Finally, the estimated background noise at location (i, j) is given by

$$b_{ij} = \sum_{\ell=1}^N \hat{w}_\ell x_\ell. \quad (4.33)$$

A number of redundant and incorrect steps in the original work are amended here. In particular, the equivalent of equations 4.32 and 4.33 in Rashed and Harbi (2014) yields an incorrectly scaled background noise matrix. Also, the authors do not provide a way to handle instances when $m_{ij}(\alpha)$ equals zero. Recall that zero is unsigned and, therefore, does not satisfy ether criterion in equation 4.30. To this end, equation 4.30 was modified

into

$$w_\ell = \begin{cases} |\underline{x}_\ell - m_{ij}(\alpha)|^{-s} & \text{if } \text{sign}(x_\ell) = \text{sign}(m_{ij}(\alpha)) \text{ or } m_{ij}(\alpha) = 0 \\ 0 & \text{otherwise} \end{cases}. \quad (4.34)$$

Since the window W_{ij} can go out of bound near the data edges, some form of edge handling must be applied prior to carrying out this technique. Rashed and Harbi (2014) suggest an edge mirroring scheme for this purpose, which is the default option in IVGPR. Other choices, e.g. extending and wrapping, can be activated by modifying the source code.

The BNS subroutine allows for three methods of estimating the background noise matrix. The first two are found directly from equation 4.30 and may be expressed as $B = [m_{ij}(0)]$ and $B = [m_{ij}(\alpha)|\alpha \neq 0]$ for the MTS and ATM options, respectively. The third option is to use the routine describe in this section, that is $B = [b_{ij}]$. Each method can be applied iteratively, where each successive background noise matrix is estimated from the previous one. A thorough comparison between MTS and BMS, which shows many advantages for the latter, is given by Rashed and Harbi (2014) and will not be included in this work. Currently no example of similar success with BMS has been found by the author.

Three background noise matrices are shown in figure 4.17. All three were found using a window size of 53 elements after performing 10 iterations. A value of $\alpha = 0.25$ was used for ATM and BMS (figures 4.17b and 4.17c respectively), and $s = 1.2$ was used for the latter. Even though there are noticeable differences between the background noise estimated by MTS and those of ATM and BMS, there last two achieve nearly identical results. A dashed square highlights one of the sections where MTS may have overestimated the background noise. This was inferred by comparing the extent to which a number of strong events stretches away from the enclosed area. An arrow indicates one of the coarse

transitions calculated by the BMS routine. This is peculiar to BMS and may have adverse effects in subsequent processing. Extensive testing revealed that MTS and ATM generally achieve better results without the added complications of BMS.

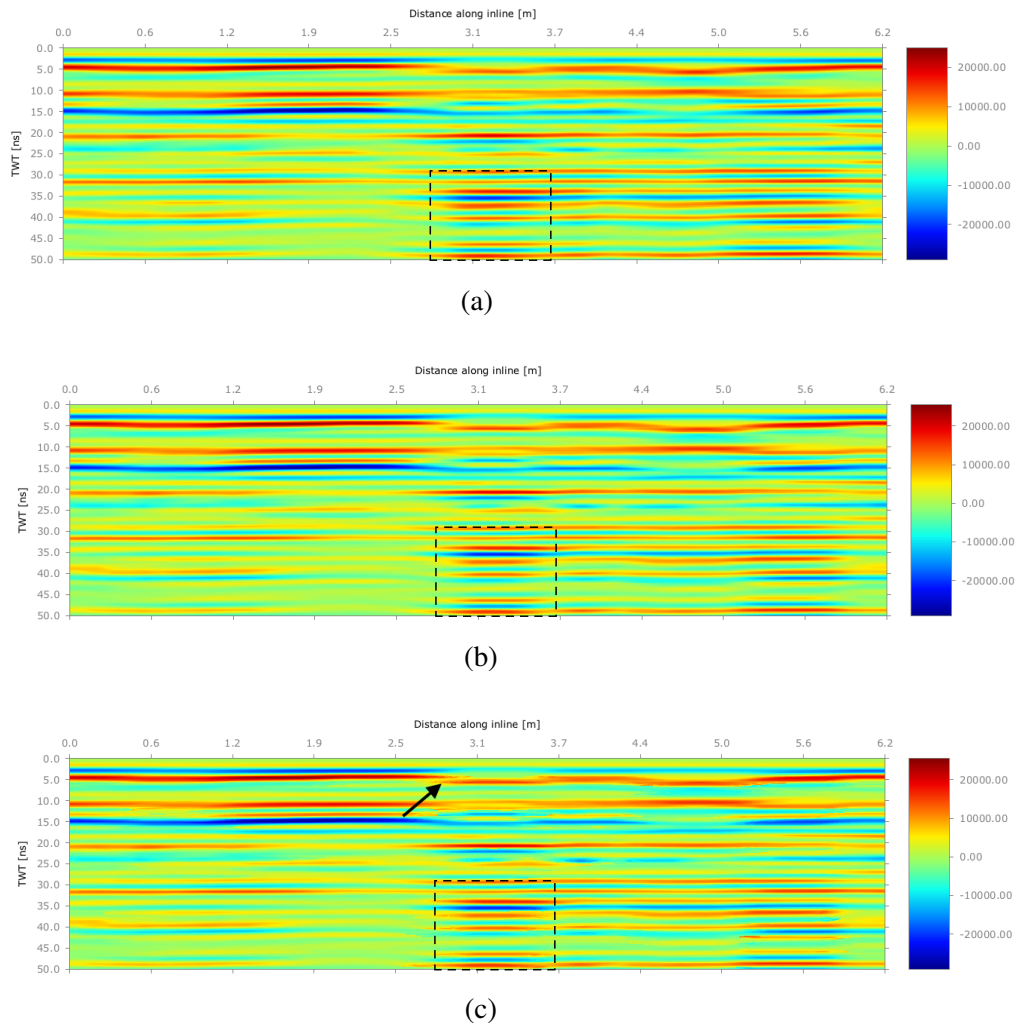


Figure 4.17: Shows three background matrices corresponding to: (a) mean trace subtraction (MTS), (b) alpha-trimmed mean trace subtraction (ATM), and (c) background matrix subtraction (BMS). A dashed square shows an instance where MTS appears to overestimate the background noise. An arrow shows one of the sharp transitions peculiar to the BMS technique.

4.4.3 Attribute-based time-zero correction

Sabbione and Velis (2010) proposed a way to improve the accuracy and stability of automatic first-arrival picking by use of one-dimensional attributes in conjunction with edge-preserving smoothing (EPS). Their procedure resembles the popular Coppens' algorithm (Coppens, 1985) with two major difference. First, it is not limited to the energy-ratio attribute. Instead, the authors suggest energy-ratio, entropy, and fractal dimension as suitable candidates for the task. Of these, however, only the first two are presently available in IVGPR. Secondly, an EPS filter (Liu et al., 2003) is applied to the attributes in order to accentuate the advent of significant shifts in amplitude level. The location of the first-arrival is found by differentiating the filtered attribute(s) and, in turn, finding the maximum absolute amplitude. Figure 4.18 shows a workflow describing this process. Four user defined parameters are needed by this routine and will be discussed below. For the last steps, IVGPR approximates the differentiation with a central difference scheme and finds the absolute maximum via an intrinsic index-searching function.



Figure 4.18: Automatic first-arrival picking workflow. Processes colored in red indicate steps that differ from Coppen's method.

Typical definitions for the energy along a curve often involves a direct proportionality to amplitude (or its value squared). This makes it a fitting attribute for detecting a transition between a “quiet” region and the arrival of the earliest pulse on a trace, since the two often differ by prominent variations in amplitude. In the case of energy-ratio (ER), the measurement at each point of a trace is estimated from the energy in its vicinity scaled

by the energy accumulated up to that point. This, consequently, favors early variations in energy. The equation to calculate ER, as given in Sabbione and Velis (2010), is:

$$\mathcal{E}(t) = \frac{E_1(t)}{E_2(t) + \beta}, \quad (4.35)$$

where $\beta[\geq 0]$ stabilizes the solution to help reduce the chance of erroneous picks. Here $E_1(t)$ and $E_2(t)$ represent the energies accumulated within the time ranges $[t - n_\ell + 1, t]$ and $[1, t]$, respectively. These quantities are computed using

$$E_1(t) = \sum_{i=t-n_\ell+1}^t s_i^2 \quad (4.36a)$$

$$E_2(t) = \sum_{i=1}^t s_i^2. \quad (4.36b)$$

Since $[t - n_\ell + 1, t] \in [1, t]$, we have that $E_2(t) \geq E_1(t)$ and $\mathcal{E}(t) \geq 1$. We see that $\mathcal{E}(t)$ requires two user defined parameters, i.e. β and n_ℓ . A third parameter, i.e. filter kernel size M , is needed by the EPS filter applied to the ER attribute. The particular algorithm does not involve any form of edge-handling and, thus, cannot estimate the energy of the first $n_\ell - 1$ terms.

The one-dimensional entropy attribute can be used to expose temporal variations in the statistical properties of a curve (Denis and Crémoux, 2002; Sabbione and Velis, 2010). For this reason, Sabbione and Velis (2010) propose its use to determine when the signal-to-noise ratio along the early section of a trace increases in response to the first arrival. Entropy is found from the expression:

$$H(t) = \log \left(\frac{1}{n_h} \sum_{i=t-n_h+1}^{t-1} |s_{i+1} - s_i| \right), \quad (4.37)$$

where n_h is the size of a moving window. Note that, as with ER, the computed value is

assigned to the leading sample in the window. As remarked by Sabbione and Velis (2010), too small of a value for n_h can result in large fluctuations in entropy while large windows may lead to unacceptably low resolutions. Both, energy-ratio and one-dimensional entropy, are helpful beyond time-zero correction. For example, they may be applied in combination with other attributes to formulate an integrated analysis.

Three time-zero corrections obtained from the energy-ratio method (C), entropy method (D), and the peak-finding method (B) are shown in figure 4.19. The first event in the original record (A) is recognized as the air-ground interface and, apart from a couple of exceptions near the 4.5 and 16 meter marks, it displays a continuous horizontal trend. The parameter values used to implement the ER and entropy methods were ($n_\ell = 7, \beta = 0.2, M = 11$) and ($n_h = 18, M = 15$), respectively. Picks made by the ER and peak-finding techniques appear to be of comparable quality and superior to that of the entropy-based technique. However, ER proved to be far more laborious as a result of its greater number of parameters. It should be pointed out that the entropy-based time-zero correction (D) was chosen from a large number of tests, all of which appeared to be less stable from a qualitative perspective. For this reason, the author suggest using the peak-finding algorithm prior to attempting the more involved approaches of this section.

Figure 4.20 indicates the possible time-zero picks for all three algorithms available in IVGPR. The picks by the entropy and ER methods (A and B respectively) correlate with the locations of the maximum absolute value for the derivative of the corresponding filtered attributes (shown by the red and blue curves). All the possible locations for the peak-finding algorithm (C–F) are marked, with C being closest to the picks by the other two techniques. One may argue that the methods from this section resulted in a time-zero that better represents the true advent of the first arrival, which is most accurately defined by the location marked as A. This observation has not being explored beyond what is presented. Also, Sabbione and Velis (2010) proposed an additional routine to improve the accuracy

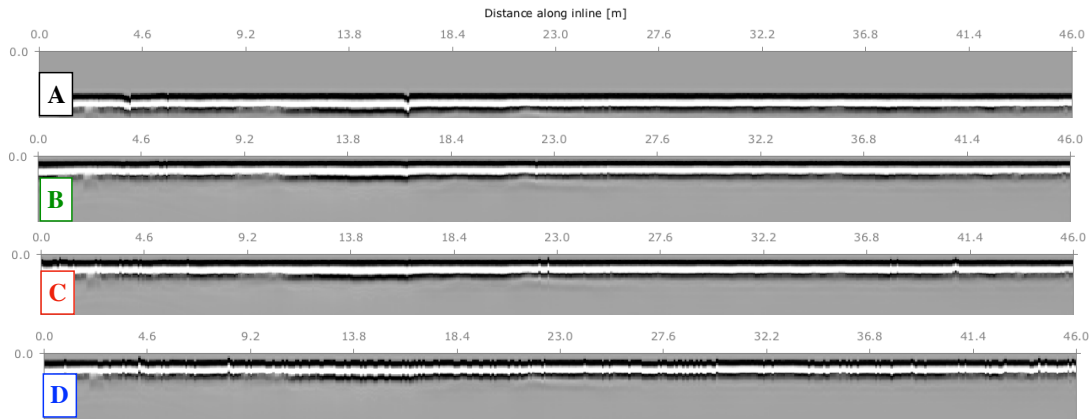


Figure 4.19: Shows the Time-zero correction of a raw B-scan (A) after applying the peak-finding based method (B), the energy-ratio method (C), and the entropy method (D).

and consistency of the picks made by their techniques, which has yet to be implemented in IVGPR. In summary, all the techniques available appear to yield acceptable results with the peak-finding method offering the best trade-off between stability and effort.

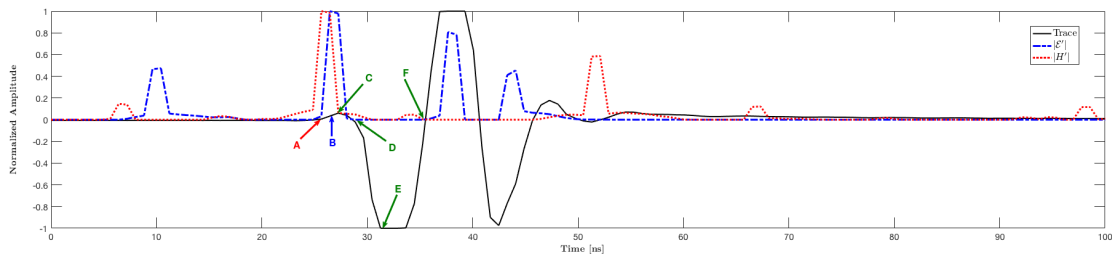


Figure 4.20: Average trace (black curve) showing the time-zero picks that would result from the entropy method (A), energy-ratio method (B), and the peak-finding method (C–F). The red and blue dashed line shows the average absolute value for the first derivative of entropy and energy-ratio attributes, respectively.

4.4.3.1 Edge-preserving smoothing

Edge-preserving smoothing filters are designed to provide a reduction in noise comparable to that of moving average filters while achieving a much sharper step response (Liu

et al., 2003; Sabbione and Velis, 2010). To explain how EPS operates, let us consider an input data series $A(= \{a_i | i = 1, 2, \dots, N\})$ treated using a kernel of size M , where M is some positive odd integer. Algorithm 1 describes a basic EPS routine, which applies to this scenario. It is worth remarking that EPS, as with many moving window filters, must be preceded by a proper handling of data edges. In particular, this gives us an extended data series $A^*(= \{a_j^* | j = 1 - \Delta, 2 - \Delta, \dots, N + \Delta\})$, with $\Delta(= (M - 1)/2)$ being the number of appended elements at either end. The i -th smoothed element, E_i , is computed by comparing the standard deviation inside M windows around the point $a_{i+\Delta}^*$, identifying which window location minimized the standard deviation, and, finally, assigning the arithmetic average inside that window to E_i . These three steps comprise lines four to seven in algorithm 1.

Algorithm 1 Edge-preserving smoothing, after Liu et al. (2003).

- 1: Expand edges so that $A \rightarrow A^*$
 - 2: $\Delta = (M - 1)/2$
 - 3: **for** $i = 1, 2, \dots, N$ **do** $\triangleright N := \text{total number of samples}$
 - 4: **for** $k = 0, 1, \dots, M - 1$ **do** $\triangleright M := \text{size of moving window}$
 - 5: $S_{k+1} = \sqrt{\frac{\sum_{j=i+k-\Delta}^{i+k+\Delta} (a_j^* - \mu)^2}{M - 1}}$ $\triangleright \mu$ is local average
 - 6: Locate index of minimum value in S_k , e.g. $\ell = \text{minloc}(S_k)$.
 - 7: $E_i = \frac{1}{M} \sum_{j=i+\ell-1}^{i+M+\ell-1} a_j$ \triangleright Smoothed i -th element
-

The excellent step response of EPS is clearly demonstrated in figure 4.21. For this case a step function, composed of 600 elements and centered at sample 300, was combined with random Gaussian noise to form the response shown in figure 4.21b. The outputs from a general moving average filter, figure 4.21c, was calculated using a 17 element sliding

window. The edge, originally located along the dashed vertical line shown, is markedly less pronounced in this figure. By contrast, figure 4.21d shows how an EPS filter of equal size preserved a transition of sharpness comparable to that of figure 4.21a.

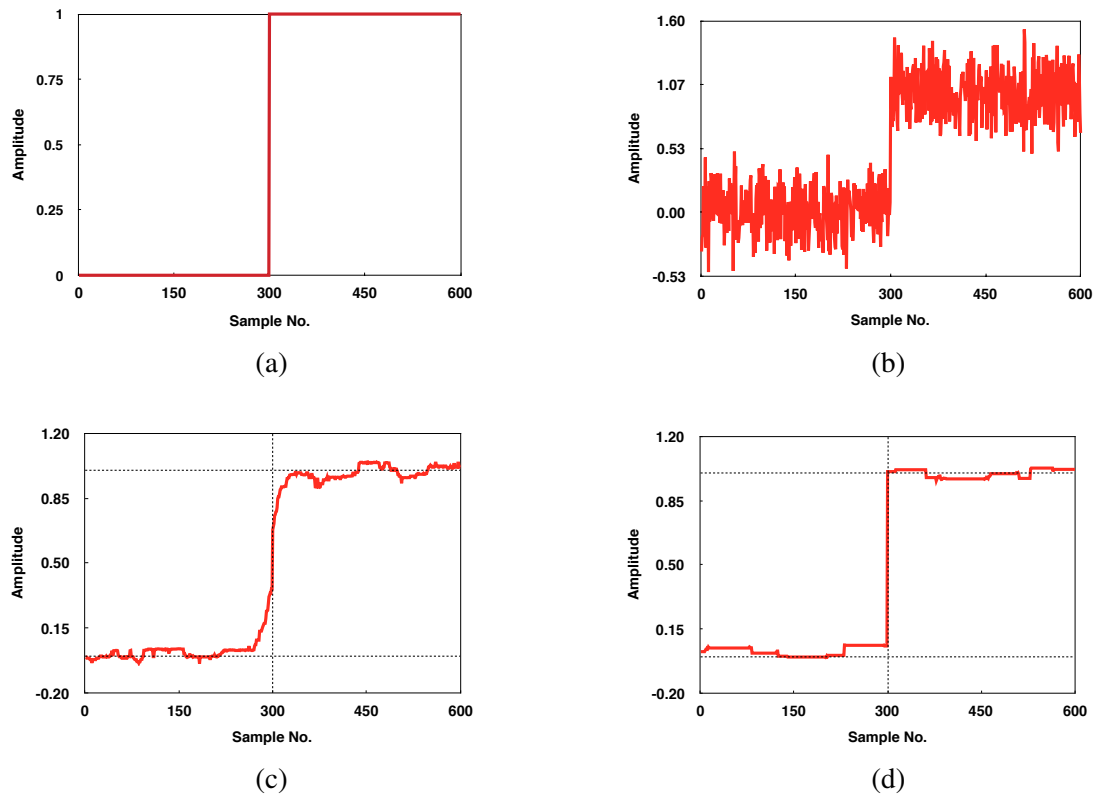


Figure 4.21: (a) Original step function. (b) Step function after random Gaussian noise was added. (c) Noisy data treated by conventional moving average filter using a window size of 17 elements. (d) Noisy data treated by EPS using a window size of 17 elements.

5. MIGRATION OF GPR DATA

5.1 Overview

Migration may be viewed as an attempt to map an image from data space into object (or physical) space, where the geometry and properties of subsurface targets can be accurately estimated (Leuschen and Plumb, 2001). As illustrated by figure 5.1, images in data space contain distortions inflicted by the acquisition process. One such example is the deterioration of spatial resolution, which results when energy reflected by an object gets mapped along a stretch of space far greater than its physical dimensions. In response, migration should reposition this energy within the object's "true" spatial confinement (Jones, 2014). Typical GPR migration techniques include: hyperbolic summation migration, Kirchhoff migration, frequency-wavenumber migration, and finite difference migration (Leuschen and Plumb, 2000; Cassidy and Jol, 2009; Qin et al., 2014; Ozdemir, Caner and Demirci, Sevket and Yigit, Enes and Kavak, Adnan, 2007). The performance of migration algorithm is generally hindered by assumptions and simplifications that do not adhere to the complex nature of field records (Slob, 2003). For instance, many algorithms assume a laterally invariant velocity model, which is generally inaccurate. This assertion is specially crippling considering the principal role that an accurate estimation of the subsurface velocity structure plays in most migrations algorithms (van der Kruk, 2001; Slob, 2003; Jones, 2014). The two techniques available in IVGPR are examples of wave-field extrapolation techniques. In particular, the user can decide between the widely popular phase-shift migration (Gazdag, 1978; Claerbout and Green, 2008) or an exact field vector-migration routine (Streich et al., 2007) specially designed for GPR data. This section presents a synoptic description of the theory underlying these algorithms while providing satisfactory examples of their implementation.

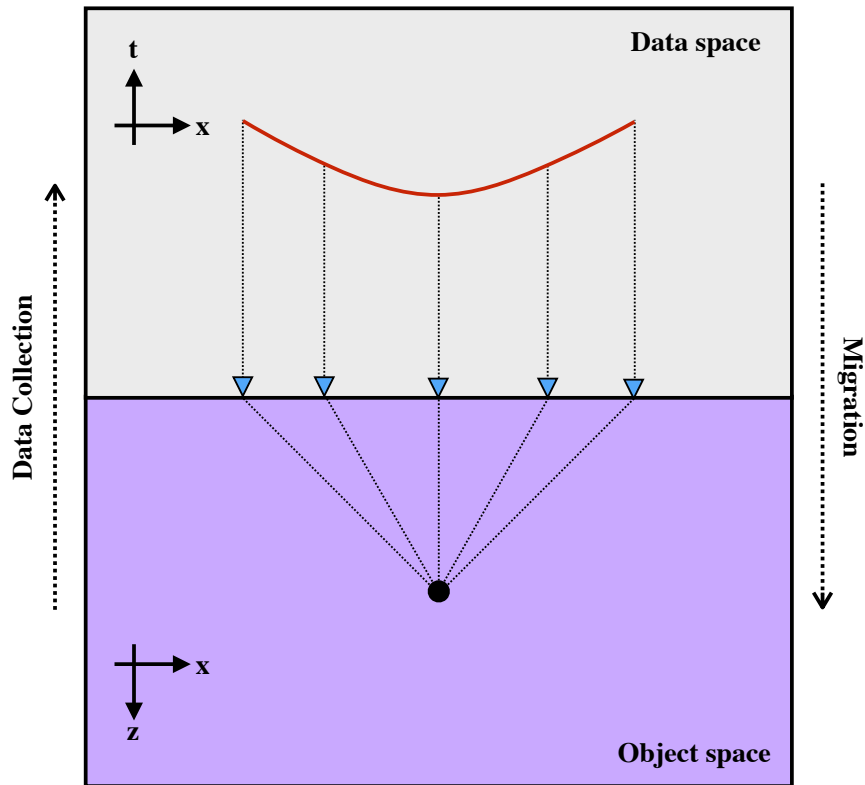


Figure 5.1: Illustrates the effect of data collection and migration. In particular, a point scatterer in object (or physical) space becomes mapped as a hyperbolic event in data space.

5.2 Phase-shift migration

Phase-shift migration is a frequency-domain technique introduced by Gazdag (1978) to treat zero-offset seismic records. As expected of a frequency-domain migration algorithms, it features a computational efficiency that far surpasses that of time-domain imaging techniques (Chun and Jacewitz, 1981). Most mainstream GPR processing softwares offer the option to perform a frequency-domain migration through the Stolt or Phase-shift algorithm, with the former being more abundant. One clear advantage of phase-shift migration over Stolt migration is the ability to easily account for vertical variations in velocity (Gazdag, 1978). This section begins by presenting a fictitious experiment to illustrates the

energy focusing property of downward-continuation and migration. Next, it gives a derivation of the downward-continuation formulas in Fourier domain, which forms the core of the phase-shift migration algorithm. A principal focus is to show how the phase-shift migration algorithm stems from the classic idea of *downward-continuation* brought forth by Claerbout (1971). One may regard this section as a concise but complete summary of past works related to the phase-shift migration method, primarily that of Claerbout and Green (2008). Additional details on the general theory of phase-shift migration (and similar frequency-domain techniques) can be found in Gazdag (1978), Chun and Jacewitz (1981), Scales (1995), Claerbout and Green (2008), and Stolt and Weglein (2012). For GPR specific applications, the reader can visit the works of Gu et al. (2004) as well as Zhang et al. (2013).

5.2.1 Downward continuation

Consider an impulse acoustic source located at some point (x_s, y_s, z_s) within a homogeneous half-space of known properties. The emitted pulses propagate isotropically (with radial symmetry) away from the source point at some velocity v . A linear array of receivers, measuring 2δ , is initially positioned on the $z = 0$ plane (parallel to the x -axis) and extends from $x_s - \delta$ to $x_s + \delta$. The receivers record the wavefield that arrives within a time-window T . After each recording, the array of receivers is lowered a distance Δz and the process is repeated. Figure 5.2 depicts this scenario. We see that as the array of receivers approaches the source, the hyperbolic event observed in the recorded data shrinks, i.e. the event within the gray section corresponding to the time range T . This is a desirable result during migration and a key aspect of downward-continuation. Ultimately we seek to focus the recorded energy down to a point, but that result is not obvious from this example. In fact, the hyperbola will appear as a point at $z = z_s$ only if the receivers are separated by a distance $\Delta x > vT$ along the array.

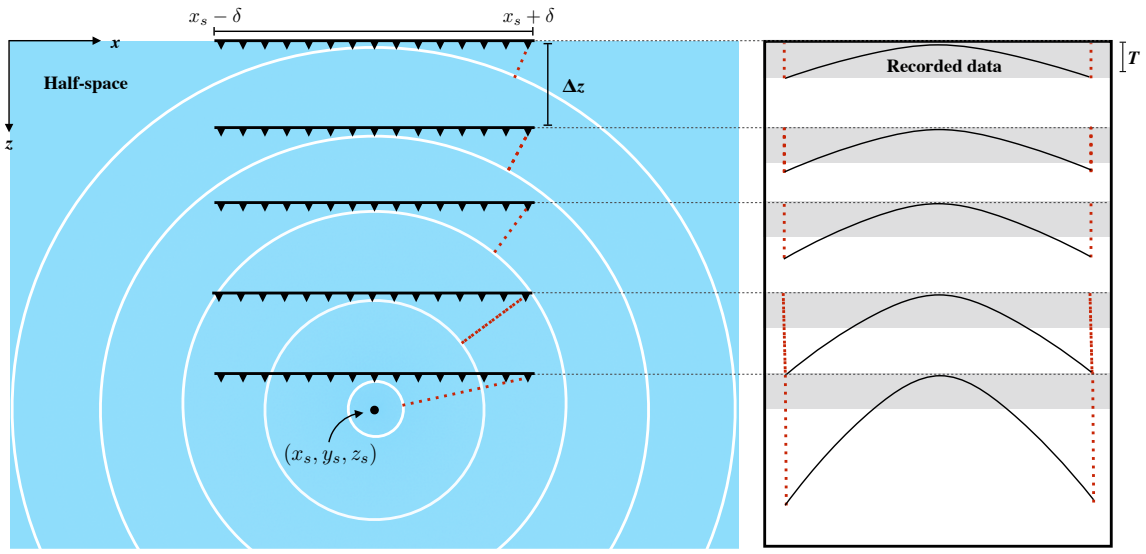


Figure 5.2: Illustrates how a hyperbolic event becomes more compact as the array of receivers gets closer to its source. The white circles represent wavefronts at the time they first reach the receivers. The dotted lines indicate the distance a wave front needs to travel to reach the edge receivers after it arrives at the middle receiver. On the right, the resulting hyperbolic event within the acquisition time (indicated by a gray band) are shown.

In practice, downward-continuation does not require moving the receiver array or having a source embedded in the medium. Instead, the wavefield recorded at the surface is decomposed into plane waves that are extrapolated along the depth-axis by means of time- or phase-shifting. To illustrate this consider a plane wave traveling parallel the z -axis of the half-space shown in figure 5.2. Hence, the equiphase planes and $z = \text{const.}$ are parallel. Assume that at time t we record the energy of the plane wave along the interface $z = 0$, such that the recorded wavefield can be expressed using $u(t, z = 0)$. Note that there is no dependence in the x - or y -coordinates since the measured response is constant along every xy -plane for a homogeneous and harmonic plane wave traveling along the z -axis. The response at some deeper section, $z = \Delta z$, can be obtained by applying a

time-shift to the measurement at $z = 0$ using

$$u(t, z = \Delta z) = u(t, z = 0) * \delta(t + \Delta t), \quad (5.1)$$

where $\Delta t [= \Delta z/v]$ is the temporal sampling rate and $\delta(t)$ is the *dirac delta function*. The convolution theorem allows us to write equation 5.1 in the frequency-domain as

$$\begin{aligned} U(\omega, z_0 + \Delta z) &= U(\omega, z = z_0) e^{-i\omega \Delta t} \\ &= U(\omega, z = z_0) e^{-i\omega \frac{\Delta z}{v}}. \end{aligned} \quad (5.2)$$

We can derive a similar equation for an arbitrarily oriented plane wave using the sketch shown in figure 5.3. Assume that a plane wave propagates through a homogeneous medium at some angle θ with respect to the vertical (depth) axis. Two recording stations, marked by a red and blue dot, are positioned along a vertical line. The stations are separated by a vertical distance Δz . Assume that at some time t_1 after we start recording the wave-front (shown in yellow) reaches the lower station. At a later time t_2 the same wave-front reaches the second station. The records shown in figure 5.3 (right) reveals that the signal at the deeper point can be determined by time-shifting the shallower response by $\Delta t = \Delta z \cos \theta / v$, where v is the velocity of propagation for the host medium. Thus, the downward-continuation equation becomes

$$\begin{aligned} U(\omega, \theta, z + \Delta z) &= U(\omega, \theta, z) e^{-i\omega \Delta t} \\ &= U(\omega, \theta, z) e^{-i\omega \frac{\Delta z \cos \theta}{v}}. \end{aligned} \quad (5.3)$$

Equation 5.3 can be generalized to work with depth-dependent velocity distributions, $v = v(z)$, but it does not account for amplitude variations introduced by reflections and transmission along interfaces. Moreover, our downward-continuation equations only ap-

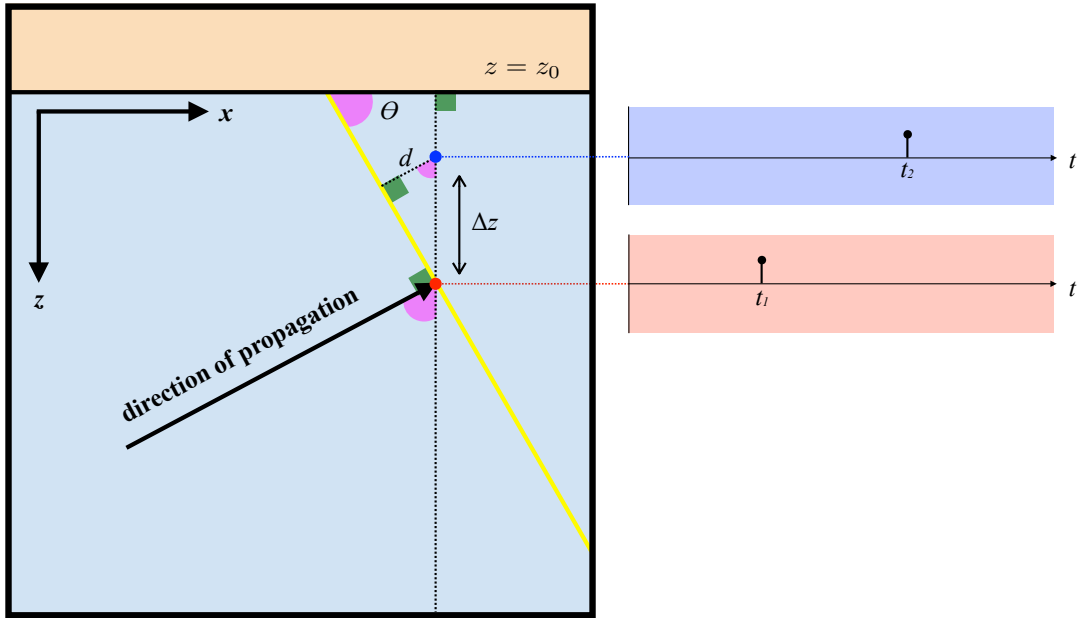


Figure 5.3: Up-going plane wave propagating at some angle θ with respect to the vertical axis. The signal reaching two stations (red and blue dots) is shown as stem plots on the right.

ply to plane waves and cannot handle most field records directly. Fortunately, field records can be decomposed into sets of plane waves. The Fourier transform lets us express a field record as the superposition of monochromatic plane waves represented by points in the wavenumber-frequency domain. We now need to modify equation 5.3 such that it can operate on Fourier transformed data. This requires finding a way to relate the propagation angle θ to the wavenumbers k_x and k_z .

Let us now consider a moving cosine wave propagating through a homogeneous medium at a velocity v and with angular frequency ω . The wave travels at an angle θ with respect to the z -axis and can be expressed using

$$u(x, z) = \cos \left[\omega \left(\frac{x}{v} \sin \theta + \frac{z}{v} \cos \theta - \varphi \right) \right], \quad (5.4)$$

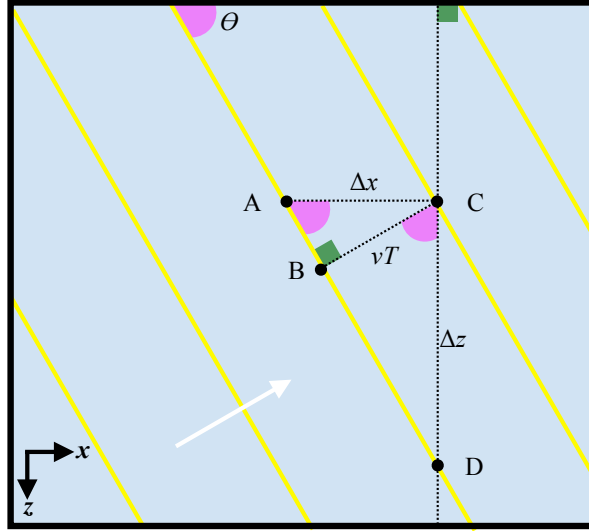


Figure 5.4: Contour lines for the equiphase planes corresponding to the maximum value of a monochromatic plane wave. The propagation direction of the wave is indicated with a white line such that it makes an angle θ with respect to the vertical axis.

where φ is some arbitrary phase-shift added for completeness. Figure 5.3 illustrates this scenario, where the yellow lines indicate the equiphase planes corresponding to the maximum amplitude. Lets define an apparent wavelength in the i -th direction as λ_i such that it describes the physical distance for two equal parts of a wave to repeat. We see from figure 5.3 that the x - and z -directed wavelength are equal to Δx (the segment \overline{AC}) and Δz (the segment \overline{CD}) respectively. The wavelength in the direction of propagation, λ_θ , is related to the velocity v and period T of the wave by $\lambda_\theta = vT$, which is equal in length to the segment \overline{BC} . We can relate wavelengths λ_x and λ_z to λ_θ using the propagation direction θ as

$$\lambda_x = \frac{\lambda_\theta}{\sin \theta} = \frac{v}{f \sin \theta}, \quad (5.5a)$$

$$\lambda_z = \frac{\lambda_\theta}{\cos \theta} = \frac{v}{f \cos \theta}. \quad (5.5b)$$

where $f [= 1/T]$ is the frequency of the moving cosine wave. When both sides of equation 5.5 are divided by 2π and inverted, the desired relationships between wavenumbers (k_x and k_z) and the angle θ are revealed as

$$\frac{\sin \theta}{v} = \frac{k_x}{\omega}, \quad (5.6a)$$

$$\frac{\cos \theta}{v} = \frac{k_z}{\omega} = \frac{1}{v} \sqrt{1 - \frac{k_x^2 v^2}{\omega^2}}. \quad (5.6b)$$

The downward propagator in equation 5.3, namely $\exp(-i\omega\Delta z [\cos \theta/v])$, becomes operational inside the Fourier domain after substituting equation 5.6b. Observe how the downward-continuation process, which originated as a time-shift, has become the phase-shift $\exp(-ik_z\Delta z)$ in Fourier domain. To recapitulate, an arbitrary data record, $u(t, x, z = z_0)$, is decomposed into a set of monochromatic plane waves, $U(\omega, k_x, z = z_0)$ after a Fourier transformation. Each element in U at z_0 can be downward-continued to a new depth $z = z_0 + \Delta z$ using the equation

$$\begin{aligned} U(\omega, k_x, z_0 + \Delta z) &= U(\omega, k_x, z_0) \exp(-ik_z\Delta z) \\ &= U(\omega, k_x, z_0) \exp\left(-\frac{i\omega\Delta z}{v} \sqrt{1 - \frac{v^2 k_x^2}{\omega^2}}\right). \end{aligned} \quad (5.7)$$

Alternative derivations of equation 5.7 based directly on the Helmholtz equation are given by Gu et al. (2004) and Zhang et al. (2013). While more direct, these approaches generally neglect the physical significance of the mathematics involved. Claerbout and Green (2008) achieve the same results with less mathematical rigor and greater clarity.

5.2.2 Phase shift migration algorithm

The imaging stage of the phase shift migration algorithm is generally broken down into three main steps: (1) decomposing the wavefield into plane waves by a Fourier transform in space and time; (2) extrapolate the field down into the earth using a phase-shift operator;

(3) estimate the reflectivity by evaluating the results from the previous step at $t = 0$ and applying an inverse Fourier transform. It is worth emphasizing that these steps do not comprise the entirety of the phase shift algorithm. Some important preliminary steps are not included, e.g. the discretization of the medium into homogeneous layers of thickness Δz . We start by addressing each of the steps in more detail while noting any pertinent assumption and limitation.

The imaging stage begins with a two-dimensional Fourier transformation to the GPR dataset. While migration algorithms based on the Fourier transform approach provide high computational efficiency, they suffer from two drawbacks. First, they are limited to homogeneous media (Scales, 1995). Secondly, space becomes periodic making the edges of the dataset act as periodic boundaries (Claerbout and Green, 2008), which can result in artifacts (Leuschen and Plumb, 2000). Figure 5.5 illustrates how an event reaching a periodic boundary reappears on the opposite side. Phase shift algorithms commonly treat the latter by allowing the user to pad the edges of the problems with zeros. This tactic was adopted in IVGPR.

Next, the measured wavefield is extrapolated downwards along the depth axis iteratively. As it stands, equation 5.7 only accounts for up-going waves, i.e. one-way propagation. If we assume a zero-offset (monostatic) configuration, then the up- and down-going waves follow identical paths. Consequentially, we can use a velocity profile twice as slow as the original in order to achieve the appropriate two-way travel times. The downward-continuation operator becomes

$$C = \exp \left(-i\omega\Delta\tau\sqrt{1 - \frac{v^2k_x^2}{4\omega^2}} \right), \quad (5.8)$$

where τ is the two-way vertical travel time such that $\Delta\tau = 2\Delta z/v$.

During the final imaging step, the extrapolated data is used to estimate the reflectivity,

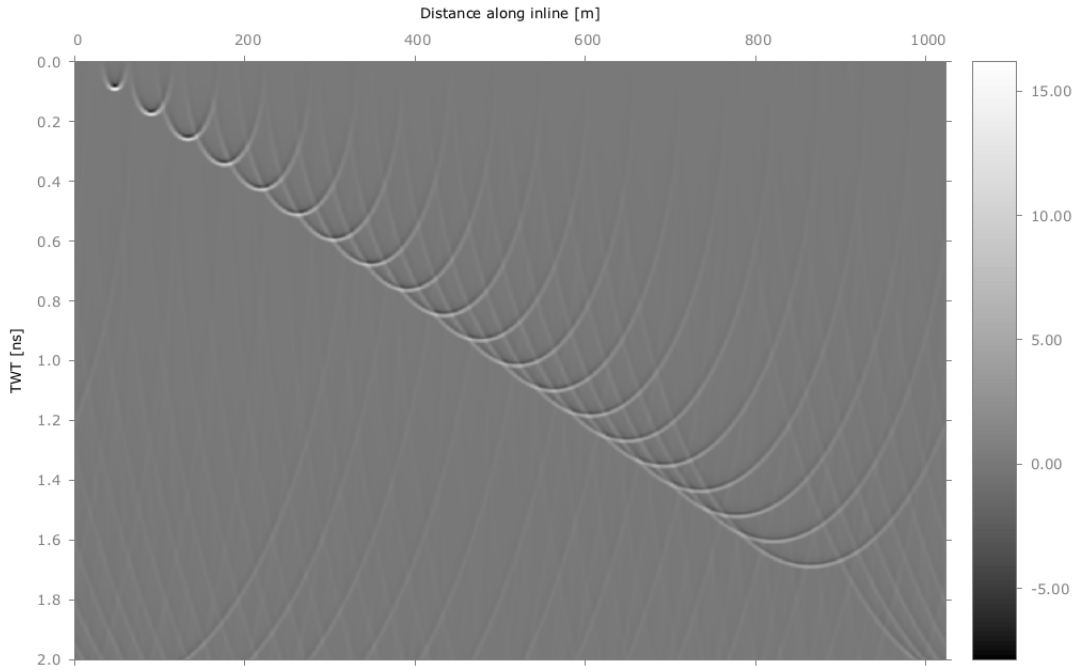


Figure 5.5: Shows artifacts caused by the intrinsic periodic boundary conditions of the phase shift migration algorithm.

Image(x, z), for every x at each new value of z . This is achieved by evaluating the extrapolated wavefield at $t = 0$, which is analogous to summing over all frequency components ω (Gazdag, 1978; Scales, 1995; Claerbout and Green, 2008). The migrated image for some arbitrary depth layer of thickness Δz is obtained after a one-dimensional inverse Fourier transform from k_x to x .

The pseudo-code detailing the implementation method adopted in IVGPR is shown in algorithm 2. This approach is commonly referred to as Claerbout’s phase-shift migration algorithm (Scales, 1995; Claerbout and Green, 2008). We see that line 1 performs the first step of the imaging stage. Line 2 is where different depth sections are picked, and it marks the beginning of the iterative steps. The second step corresponds to lines 3 through 6 while the third step is perform from lines 7 to 11.

Algorithm 2 Phase shift migration from Claerbout and Green (2008)

```
1:  $u(t, x, \tau = 0) \xrightarrow{FFT} U(\omega, k_x, \tau = 0)$   $\triangleright$  2D-Fourier transform
2: for  $\tau = \Delta\tau, 2\Delta\tau, \dots$ , end of time axis on B-scan do
3:   for all  $k_x$  do
4:     for all  $\omega$  do
5:        $C = \exp\left(-i\omega\Delta\tau\sqrt{1 - v^2k_x^2/4\omega^2}\right)$ 
6:        $U(\omega, k_x, \tau) = U(\omega, k_x, \tau - \Delta\tau) * C$ 
7:     for all  $k_x$  do
8:        $\text{Image}(k_x, \tau) = 0$ 
9:     for all  $\omega$  do
10:       $\text{Image}(k_x, \tau) = \text{Image}(k_x, \tau) + U(\omega, k_x, \tau)$ 
11:    $\text{Image}(k_x, \tau) \xrightarrow{IFFT} \text{Image}(x, \tau)$   $\triangleright$  Fourier transform along  $x$ 
```

5.3 Vector-migration

Proper GPR imaging requires migration techniques that address the vectorial nature of EM signals and the radiation pattern of antennas (van der Kruk et al., 2000; Slob, 2003; Streich and van der Kruk, 2007; van der Kruk, 2015). Most studies, however, use scalar methods similar to the one described in § 5.2, which are unfit for true amplitude processing and inversion. A new class of migration algorithms specifically designed for GPR systems (dubbed vector migration) has been developed in the past two decades. Their objective is to generate images with quantitative characteristics directly related to the physical properties of targets (van der Kruk, 2015). Most algorithms in this class implement far-field approximation to Green’s functions that are not proper for the typical depth ranges of GPR studies (about 1 to 15 meters or 1 to 5 wavelengths). By contrast, Streich et al. (2007) proposed an inexpensive way to obtain highly accurate solutions to Green’s functions as part of a three-dimensional single-component copolarized vector migration algorithm. The routine offered in IVGPR stems from the works of van der Kruk (2001) and Streich et al. (2007). The focus of this section is to provide a synaptic but adequate description of this algorithm. For an in-depth description of this technique the motivated reader should consult Streich

et al. (2007) and the complementary material mentioned in that publication.

5.3.1 Algorithm and pseudocode

Vector migration and phase-shift migration, see § 5.2, share a number of similarities that fall in line with the properties of wavefield extrapolation (WE) methods. Images in both methods are formed by applying a set of depth extrapolating operators, $\{\tilde{F}(\omega_i; \mathbf{k}, z_j)\}$, that are computed for each relevant depth level z_i and angular frequency ω_i . In practice, these operators function as two-dimensional filters acting along the \mathbf{k} -planes corresponding to the frequency components of the recorded wavefield and the desired depth levels. The migrated image along a depth level is calculated from the sum of all filtered frequency components for that depth. These steps form the basis of *the imaging condition* for wavefield extrapolation migration (Claerbout and Green, 2008; Jones, 2014) and should be consistent with the strategy outlined in § 5.2.2. Next we want to consider how this idea was transformed into the vector migration routine proposed by Streich et al. (2007) and implemented in IVGPR.

Figure 5.6 provides an illustrative description of the vector migration strategy proposed by Streich et al. (2007). This scheme is an amendment to the earlier work of van der Kruk (2001). The main difference between the two lies in the strategy used to estimate the f - x domain Green's functions necessary to find $\tilde{F}(\omega_i; \mathbf{k}, z_j)$. In particular, far-field analytical expressions (Engheta et al., 1982) are replaced by a four step scheme (shaded in gray) to obtain accurate approximations valid in the near- and mid-field ranges. The first step involves computing f - \mathbf{k} domain Green's functions, $\tilde{G}_{\alpha\beta}$, where $\alpha \in \{x, y, z\}$ represents the orientation of the electric field and $\beta \in \{x, y\}$ represents the inline-crossline directions. The symbols $\hat{\cdot}$ and $\tilde{\cdot}$ indicate whether a variable is in x - f or \mathbf{k} - f domain, respectively. For a horizontal x - or y -oriented co-polarized antenna pair, f - \mathbf{k} domain

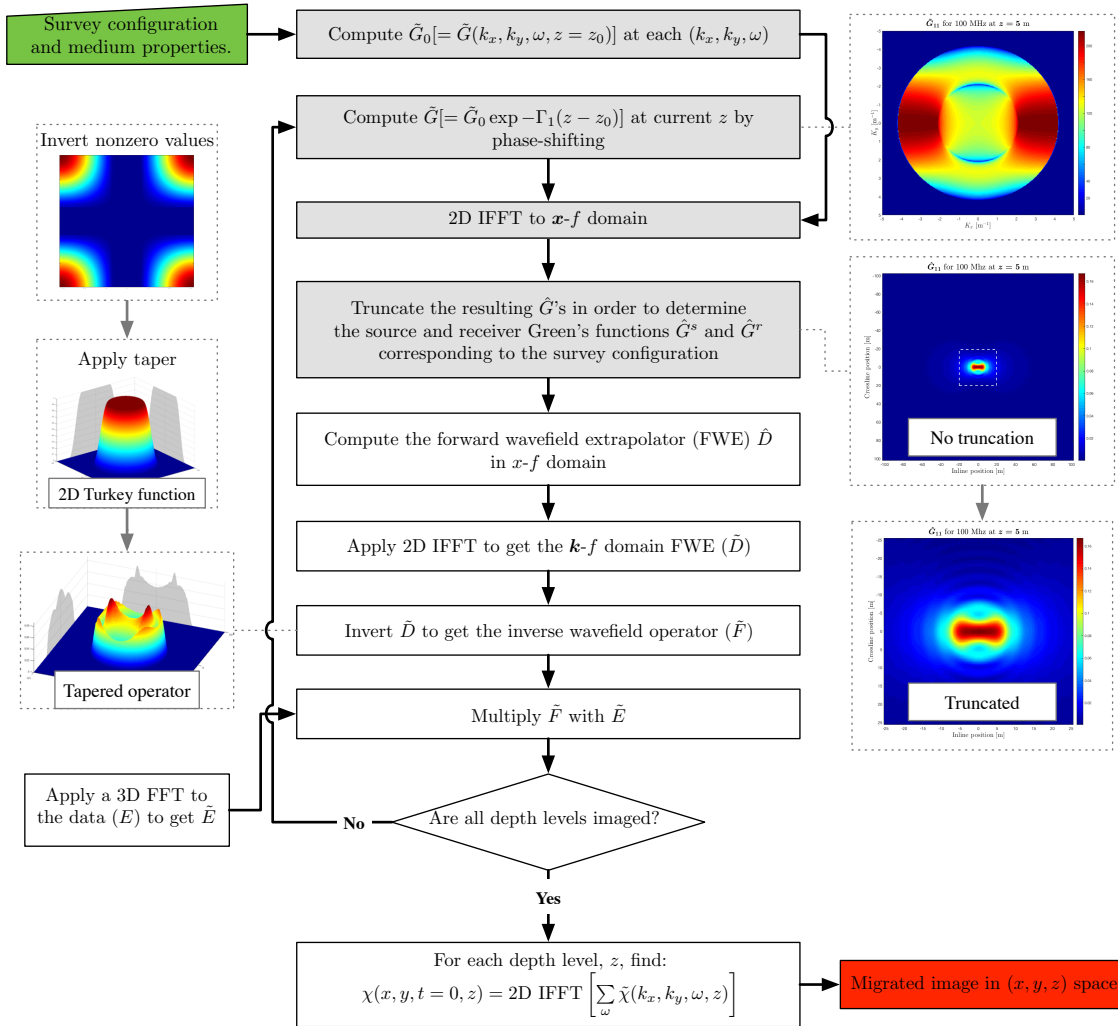


Figure 5.6: Shows the vector migration workflow proposed by Streich et al. (2007). The steps shaded in gray indicate the FFT-based technique for finding \hat{G} . The outputs from a 100 MHz monostatic antenna configuration at a depth of $z = 5$ m are shown for different steps. Oversampling factors of $p_x = p_y = 4$ were used on a spatial (x, y) grid composed of 512 by 512 elements separated by a distance of 0.1 m in both directions.

Green's functions at the initial depth z_0 are found, after Streich et al. (2007), using

$$\begin{pmatrix} \tilde{G}_{xx}(k_x, k_y, \omega, z_0) \\ \tilde{G}_{yx}(k_x, k_y, \omega, z_0) \\ \tilde{G}_{zx}(k_x, k_y, \omega, z_0) \end{pmatrix} = -\zeta \begin{pmatrix} k_x^2 \tilde{V} + \tilde{U} \\ k_x k_y \tilde{V} \\ -j k_x \Gamma_0 \tilde{V} \end{pmatrix} \quad (5.9)$$

and

$$\begin{pmatrix} \tilde{G}_{xy}(k_x, k_y, \omega, z_0) \\ \tilde{G}_{yy}(k_x, k_y, \omega, z_0) \\ \tilde{G}_{zy}(k_x, k_y, \omega, z_0) \end{pmatrix} = -\zeta \begin{pmatrix} k_x k_y \tilde{V} \\ k_y^2 \tilde{V} + \tilde{U} \\ -j k_y \Gamma_0 \tilde{V} \end{pmatrix}, \quad (5.10)$$

where we have that:

$$\begin{aligned} \tilde{U} &= \frac{\exp\{-\Gamma_1 z_0\}}{\Gamma_0 + \Gamma_1}, & \tilde{V} &= \frac{\exp\{-\Gamma_1 z_0\}}{\gamma_1^2 \Gamma_0 + \gamma_0^2 \Gamma_1}, \\ \Gamma_i &= \sqrt{\gamma_i^2 + k_x^2 + k_y^2}, & \gamma_i^2 &= \eta_i \zeta, \\ \eta_i &= \sigma_i + j\omega\epsilon_i, & \zeta &= j\omega\mu_0. \end{aligned} \quad (5.11)$$

Here γ_i describes the complex propagation constant for air ($i = 0$) and the subsurface ($i = 1$), η_i is the electric material parameters for conductivities σ_i and permittivities ϵ_i , and ζ is the magnetic material parameter with permeability μ_0 . Green's functions for some other depth interval z can be found using a vertical phase-shift on equation 5.9, such that

$$\tilde{G}_{m\alpha}(k_x, k_y, \omega, z) = \tilde{G}_{m\alpha}(k_x, k_y, \omega, z_0) \exp\{-\Gamma_1 (z - z_0)\}. \quad (5.12)$$

In practice, equation 5.9 is analyzed at each node of a three-dimensional rectangular mesh in (k_x, k_y, f) space. An interval refinement (or oversampling) is implemented along the k_x and k_y axes. This is equivalent to expanding the size of each wavefield extrapolating filter kernel, which generally enhances the results obtained by this class of migration algorithm (Hale, 1991). Particularly, the increase in kernel size manifests as a more accurate approximation to the “exact” Green's functions in f - x domain (Streich and van der Kruk, 2007). The refined wavenumber intervals are found using

$$\Delta k_\beta^{(o)} = \frac{\Delta k_\beta^{\text{data}}}{p}, \quad (5.13)$$

where $\Delta k_\beta^{\text{data}} [= 2\pi/N_\beta \Delta\beta]$ represents the original wavenumber sampling intervals (determined by the survey configurations) and p is the *oversampling* (or refining) factor. It is worth mentioning that large values of p could result in unfeasible computational and memory requirements for an average personal computer.

Once the f - \mathbf{k} domain Green's functions are computed, the next step is to obtain their f - \mathbf{x} counterparts (\hat{G}) by performing a 2-D IFFT in \mathbf{k} domain. The oversampling from the previous step results in \hat{G} 's that encompass a region of object space greater than the surveyed area. Thus, the edges of each \hat{G} must be truncated to ensure a dimensional agreement with the unmigrated data. When the transmitter-receiver offset is not zero, the Green's functions for the receiver (\hat{G}^r) are obtained by shifting the truncation boundaries by an amount equivalent to the offset between the antennas. This strategy is most accurate when the spacing between the transmitter and the receiver is some integer factor of the trace spacing, which is assumed to be uniform.

The forward wavefield extrapolator \hat{D} is found from the inner product of the source and receiver Green's functions. Namely, we have that

$$\hat{D}_\beta(x, y, \omega, z) = \sum_\alpha \hat{G}_{\alpha\beta}^s(x, y, \omega, z) \hat{G}_{\alpha\beta}^r(x, y, \omega, z) \quad (5.14)$$

for an antenna pair pointing in the β direction. Next we compute the inverse operator \tilde{F} using the expression

$$\tilde{F}(k_x, k_y, \omega, z) = \begin{cases} \left(\mathcal{F} \left\{ \mathcal{F} \left[\hat{D}(x, y, \omega, z) \right] \right\} \right)^{-1} & \text{for } k_x^2 + k_y^2 < -\gamma_1^2, \\ 0 & \text{otherwise} \end{cases}, \quad (5.15)$$

where \mathcal{F} is the forward Fourier transform. Note that equation 5.15 is nonzero only within the propagation regime of the \mathbf{k} -planes corresponding to each admissible frequency com-

ponent. In practice, however, a sharp transition like the one defined by this equation should be avoided. Recall the relationship between the sharpness of the transition band and the size of the kernel discussed in § 3.3.1. An “ideal” transition, like the one defined in equation 5.15, would correspond to a virtually infinite kernel prohibiting a working approximation to a linear convolution between the data and the inverse operator. This problem is circumvented by using an elliptical Tukey window function (Bernstein et al., 2004) to smoothly transition between the propagation and the evanescent regimes. In the IVGPR implementation, the ideal stop-band attenuation of Tukey windows was exploited by allowing the condition $k_x^2 + k_y^2 < -\gamma_1^2$ in equation 5.15 to be replaced with $|\tilde{D}| > 0$. In other words, we first find \tilde{D}^{-1} using

$$\tilde{D}^{-1} = \begin{cases} 1/\tilde{D} & \text{for } |\tilde{D}| > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.16)$$

and, subsequently, multiply our result with a Tukey window to get \tilde{F} . This approach, shown in figure 5.7 resulted in noticeable improvements in performance.

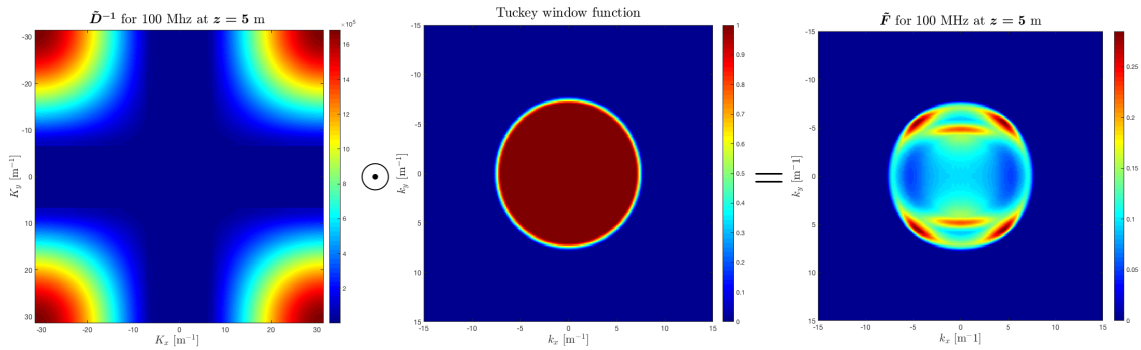


Figure 5.7: Method for computing \tilde{F} (right-hand side) from the Hadamard, or element-wise, product of the Tukey window and \tilde{D}^{-1} , which is found using equation 5.16. The Tukey window shown was found using a transition width of $\mu = 1$, see §4.2.2.

The final stage of each iteration in the migration algorithm involves multiplying \tilde{F} with the \mathbf{k} - f representation of the data, i.e. \tilde{E} . This results in a single-frequency component, $\tilde{\chi}(k_x, k_y, \omega, z)$, for the current depth level z . Once all of the frequency components have been computed, the migrated image at z is found using

$$\chi(x, y, z) = \mathcal{F}^{-1} \left\{ \mathcal{F}^{-1} \left(\sum_{\omega} \tilde{\chi}(k_x, k_y, \omega, z) \right) \right\}. \quad (5.17)$$

This procedure is repeated for each depth level until the final migrated cube is found. Algorithm 3 gives a programmatic representation of the steps described above. The backbones of the implementation in IVGPR follow algorithm 3 with the exception that it allows user to limit the number of frequency values, ω , to consider. For details on the physical significance and mathematical formalism of the wavefield extrapolator \tilde{D} , the reader can visit Slob (2003) as well as van der Kruk (2001). The ensuing section examines the performance of both, phase-shift and vector migration, techniques through primarily qualitative means.

5.4 Performance analysis of migration algorithms

Synthetic and field data were used to evaluate the performance of both, vector and phase-shift, migration algorithms. In particular, the accuracy of each technique was quantified by comparing the percent error between the dimensions of structures within migrated sections and those of the models used to generate the raw data. Additionally, the parallel efficiency of the vector migration algorithm was examined using the runtime differences from a serial to a quad-thread run using single thread increments while keeping the dimension and configuration of each run constant. Even though only a fundamental, and at times purely qualitative, analysis is given in this section, it captures how the two migration routines in IVGPR can generate satisfactory results upon judicious implementation. Details of practical importance, e.g. the effect of tapers and padding, are also included.

Algorithm 3 Single-component vector migration based on Streich et al. (2007)

- 1: $E(t; x, y, z_0) \xrightarrow{3D\ FFT} \tilde{E}(\omega; k_x, k_y, z_0)$ \triangleright Input GPR record
 - 2: $(\Delta k_x^{(o)}, \Delta k_y^{(o)}) = (\Delta k_y p_x^{-1}, \Delta k_x p_y^{-1})$
 - 3: **Compute** $\tilde{G}_{m\alpha}(\omega; k_x, k_y, z_0)$ **and image at** z_0 \triangleright Same steps as shown below
 - 4: but without a phase-shift
 - 5: **for** $z = \Delta z + z_0, 2\Delta z + z_0, \dots, N\Delta z + z_0$ **do**
 - 6: **for all** ω **do**
 - 7: **for all** (k_x, k_y) **do**
 - 8: $\tilde{G}_{\alpha\beta}(\omega; k_x, k_y, z) = \tilde{G}_{m\alpha}(\omega; k_x, k_y, z_0)e^{-\Gamma_1(z-z_0)}$ \triangleright Phase shifting step
 - 9: $\tilde{G}_{\alpha\beta}(\omega; k_x, k_y, z) \xrightarrow{2D\ IFFT} \hat{G}(\omega; x, y, z)$
 - 10: **Truncate and shift** $\hat{G}_{\alpha\beta}$ **to get** $\hat{G}_{\alpha\beta}^s$ **and** $\hat{G}_{\alpha\beta}^r$
 - 11: $\hat{D}_\beta = \sum_\alpha \hat{G}_{\alpha\beta}^s \hat{G}_{\alpha\beta}^r$
 - 12: $\hat{D}_\beta(\omega; x, y, z) \xrightarrow{2D\ FFT} \tilde{D}_\beta(\omega; k_x, k_y, z)$
 - 13: **for all** $\tilde{D}_\beta \neq 0 + i0$ **do**
 - 14: $\tilde{F}_\beta = \tilde{D}_\beta^{-1}$
 - 15: **Implement taper or filter** \tilde{H} **to get** $\tilde{H}\tilde{F}_\beta$
 - 16: $\tilde{\chi} = \tilde{E}\tilde{H}\tilde{F}_\beta$
 - 17: $\chi(t = 0; x, y, z) = 2D\ IFFT \left\{ \sum_\omega \tilde{\chi} \right\}$ \triangleright Migrated depth slice
-

The highly regarded modeling tool gprMax (Warren et al., 2015) was used to simulate the response of a kink style syncline. The corresponding subsurface model, shown in figure 5.8, was created using four unique materials and eight geometrical shapes. These were configured to represent a layer of clean sand (with $\epsilon_r = 4$, $\mu_r = 1$, and $\sigma = 0\text{mS/m}$) overlaying a clay layer with laterally varying properties. The clay layer is divided into three sections with different dielectric constants, i.e. $\epsilon_r \in \{10, 20, 40\}$, but the same value for conductivity and relative permeability ($\mu_r = 1$ and $\sigma = 100\text{mS/m}$). A Hertzian dipole emitting a Ricker wavelet with a nominal frequency of 100 MHz served as the source. The source and receiver were oriented perpendicular to the inline direction in a monostatic configuration. The synthetic B-scan comprised 160 traces spaced 0.1 m apart and each containing 2333 samples representing a total time window of 110 ns. Figure 5.9 shows the

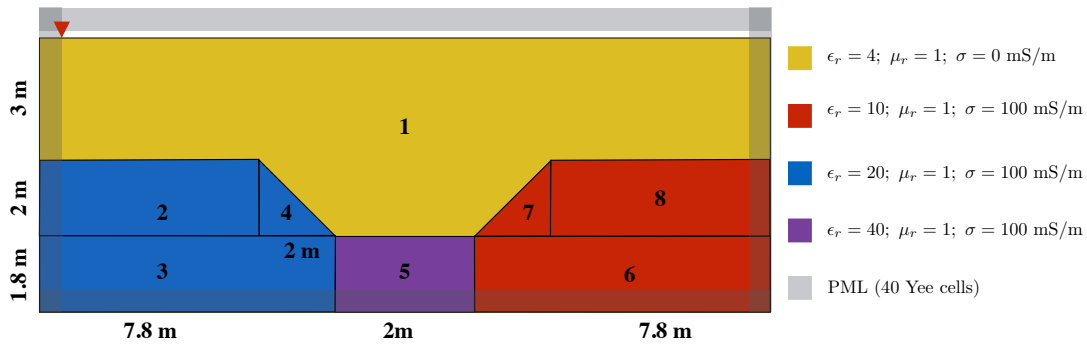


Figure 5.8: Subsurface model of a kink-style syncline. A total of eight geometrical shapes and four unique materials were used to construct the model. These create a sand layer overlaying a folded layer composed of three clay section varying in permittivity.

resulting B-scan after an instantaneous AGC gain was applied. This, in turn, revealed the bowtie pattern characteristic of synclines as well as multiples and diffractions within the simulated response. Since the vector migration algorithm requires three-dimensional data, a GPR cube composed of 64 lines at a hypothetical spacing of 0.1 meters was created by cloning the B-scan and assuming isotropy in the cross-line direction.

Figure 5.10 shows a juxtaposition of the raw radargram (a) and the result after the phase-shift routine (b). The migration was carried out using a velocity of 0.15 m/ns, which corresponds to $\epsilon_r = 4$. A comparison between the dimensions of the migrated syncline and the model revealed an average percent error of 6.1%. The horizontal edges of the migrated image were tapered during the migration and, therefore, present regions where data was lost. This, in combination with padding, can help reduce the number of artifacts that appear in the migrated radargram. By contrast, figure 5.10c shows the result of carrying out an analogous migration without applying a taper. Note that, while the reconstruction of the syncline is virtually identical, we can appreciate the emergence of two strong artifacts.

Figure 11 shows the central line in our GPR cube, i.e. line 32, after carrying out a

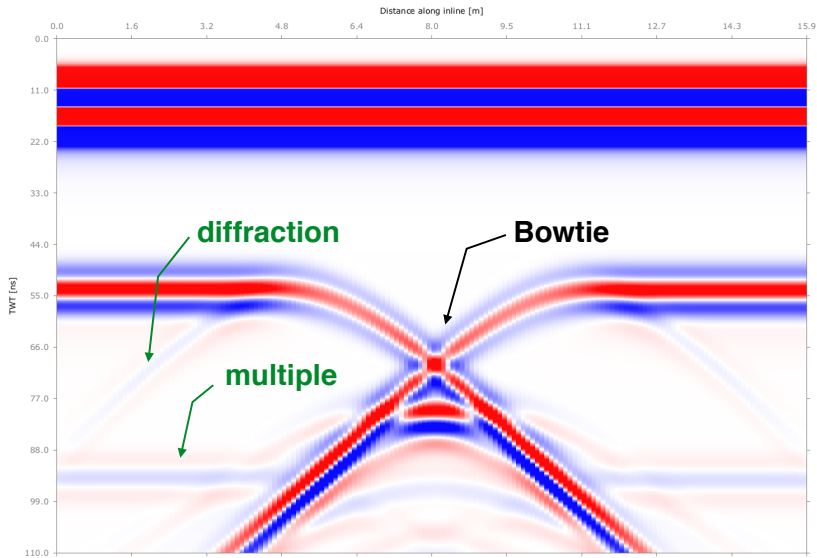


Figure 5.9: Synthetic B-scan of kink-style syncline after undergoing an instantaneous AGC gain. The expected bow-tie pattern can be seen at the center. Two green arrows point to a diffraction and a multiple within the B-scan.

vector migration. This run was configured to use a dielectric constant (ϵ_r) of 4, an inline oversampling factor (p_x) of 8, and a cross-line oversampling factor (p_y) of 16. Frequencies ranging from 10 MHz to 400 MHz were used, which amounted to 38 frequency elements. Note that the ground-to-surface reflection was removed as necessitated by this algorithm (Streich et al., 2007). This, in turn, resulted in traces with 1733 samples each, or a 26% reduction in size. Unlike the admissible percent error achieved by phase-shift migration, the dimensions of the radiograms migrated by this routine deviated, on average, by 25% from those of the model. Even though the exact cause for this discrepancy was not uncovered, three possibilities are: the required trimming of the direct wave, a non-optimal choice of parametrization, and the nature of the synthetic data. Another important distinction between the migrated data in figure 5.10 and figure 5.11 pertains to the amplitude variations along the sand-to-clay contact. In particular, the line treated by the vector migration shows a clearer difference between the reflection strength at the hinge and that of the the limbs

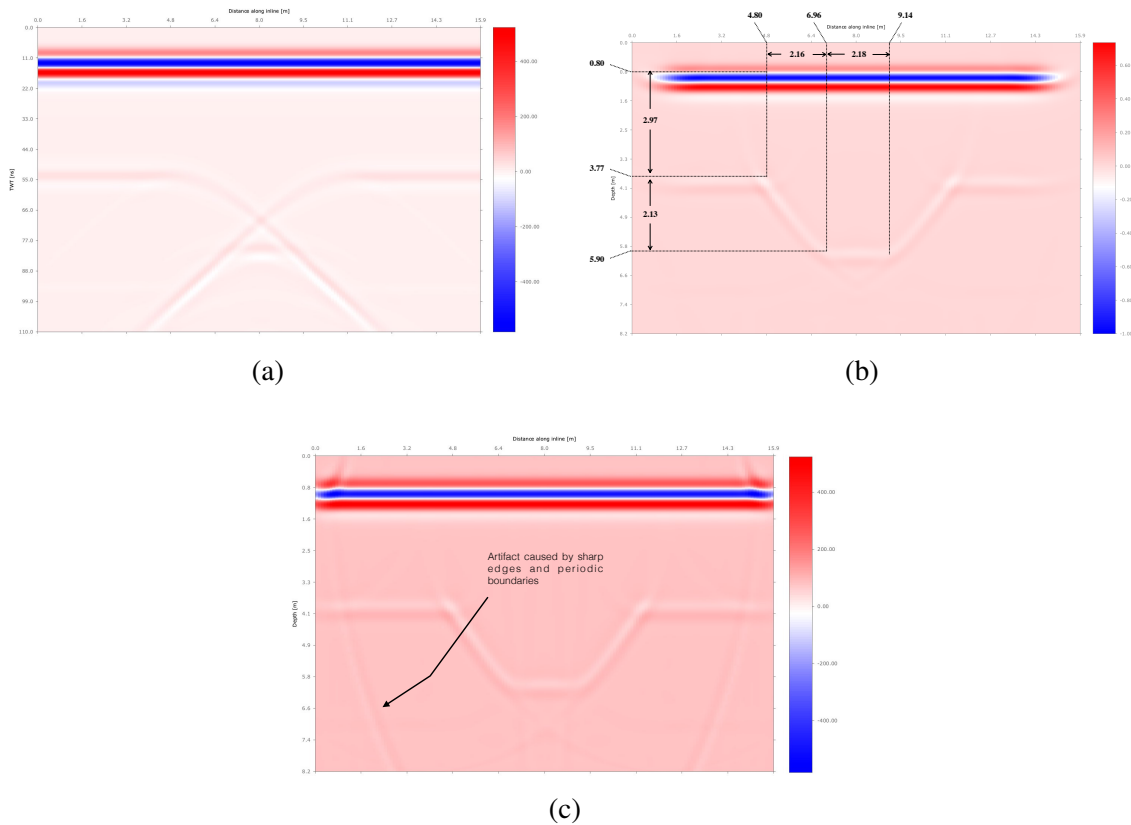


Figure 5.10: Shows the raw synthetic radargram (a) next to the one migrated using the phase-shift algorithm (b). A velocity of 0.15 m/ns was used for the migration. A separable two-dimensional Plank-tapering window was applied to artifacts produced at the edges of the data. The same migration without tapering is shown in (c), which resulted in the manifestation of strong artifacts.

and unfolded lateral sections. From these one could correctly infer that a greater contrast in properties exists in that region.

One hindering characteristic of vector migration is an uneconomical computational demand. Streich et al. (2007) reported running times of 8.4 hours for a modest survey using a high end processor. The implementation in IVGPR was designed using a parallel framework that allowed for a maximum speedup of 3.26 folds after engaging four threads. A parallel performance analysis for the vector migration routine is shown in figure 5.12.

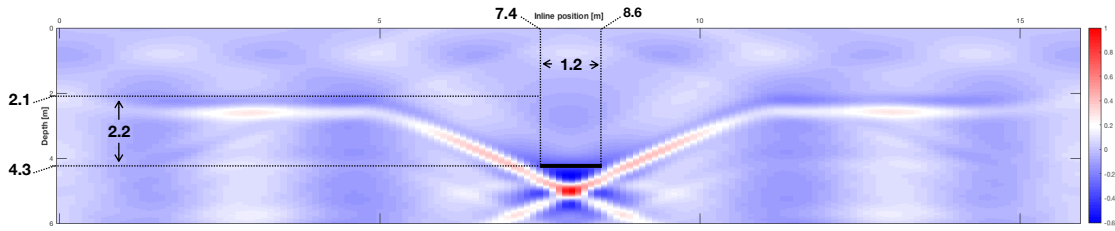


Figure 5.11: GPR line treated by the vector migration algorithm using $\epsilon_r = 4$, $p_x = 8$, and $p_y = 16$. Each original depth slice contained 160×64 elements. A total of 1733 depth slices were migrated using 38 frequency components from 10 MHz to 400 MHz.

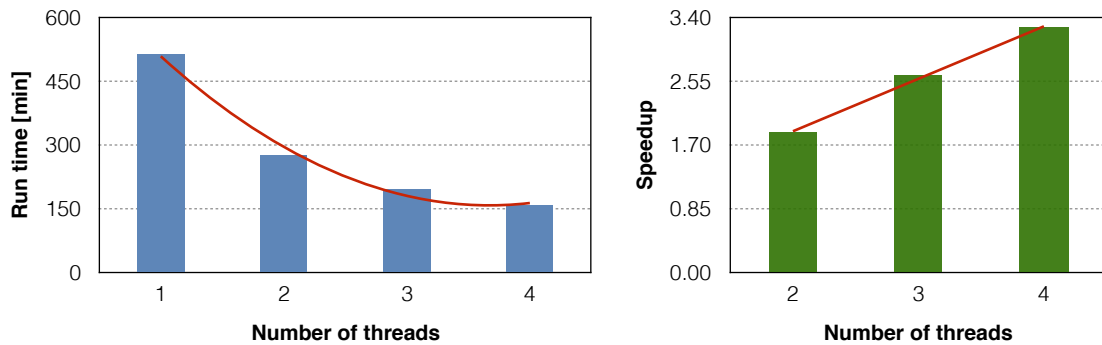


Figure 5.12: Parallel performance analysis for the single-component 3D exact radiation pattern vector migration routine in IVGPR. The run time (a) shows a quadratic decrease with number of threads activated while the speedup (b) was linear with a slope of 0.77.

This results were obtained using 2333 depth slices with 1024 by 64 elements each and a total of 11 frequency elements, which corresponds to more than 1.68×10^9 iterations. A quadratic decay in run time and a linear increase (of about 0.77 folds per processor) were found. The synthetic dataset used for figure 5.11 was comparable in size to the survey used by Streich et al. (2007), and it was migrated using a similar configuration. Yet, it took an average 127 minutes to complete each successful run using four processors with a clocking speed of 3.6 GHz each.

The performance of this vector migration algorithm was qualitatively assessed using a small field survey composed of 18 lines with 213 traces and 512 time samples. The

survey was conducted without any regards to the stringent requirements of this routine, which impaired the possibility of an effective migration. A slide corresponding to the line at the center of the migrated survey is shown in figure 5.13b. Even though we can appreciate a number of artifacts near the deeper sections, the overall reconstruction does not appear inadequate when compared to the unmigrated central line, figure 5.13a. A more informative quantitative analysis of this result was not pursued in this work.

In summary, it was demonstrated that the phase-shift and vector migration algorithms can adequately migrate GPR data if properly implemented. It appeared as though the phase shift migration algorithm offers an inexpensive operation cost and a more accurate geometrical reconstructions. On the other hand, the vector migration produced a more accurate rendering of reflection strengths but only a marginally acceptable geometrical reconstruction and intensive computational requirements.

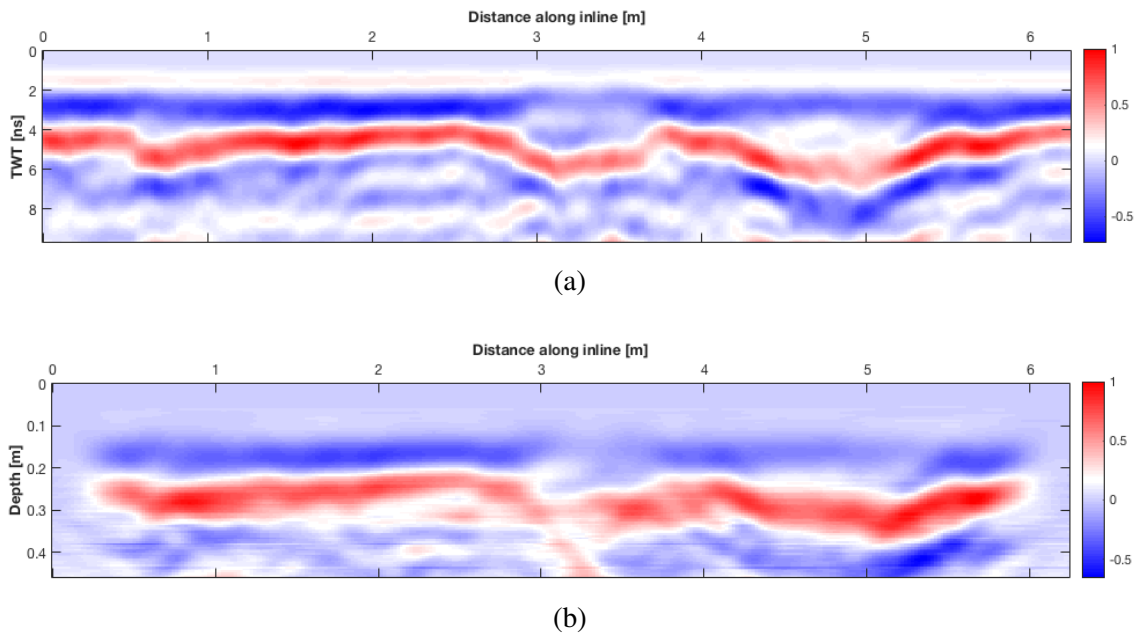


Figure 5.13: Vector migration of field data. The migrated line (b) shows artifacts likely due to aliasing. The corresponding unmigrated line is shown in(a).

6. CASE STUDIES

6.1 Overview

This section demonstrates how the numerous routines described throughout this work can be combined in order to create effective workflows capable of enhancing the interpretability and usefulness of GPR records. To that end, a total of two case studies where IVGPR served as the primary processing tool are presented. The datasets used were made available, in raw and processed form, by fellow GPR practitioners. The first survey took place at Alcatraz National Historic Landmark, while the second survey was performed to study a large exposed fault near Mason, Texas. Both surveys are introduced with a concise description of pertinent historical and geological elements. The results and interpretations included in this section are intended to demonstrate the robustness of the program and should not be used to rival or support past conclusions made using these datasets.

6.2 Case I: Geophysical archeology over the recreational yard of the main Alcatraz prison

The Alcatraz survey was carried out by Dr. Timothy S. de Smet and Dr. Mark E. Everett over the main prison's recreational yard. The goal of this study was to identify buried infrastructures matching the architectural records of a military fortification dating back to 1859. A manuscript by de Smet, currently under review, suggests that the recreational yard has a concrete foundation overlaying soil used as fill during the construction of the prison. The expected targets (e.g. man-made tunnels and the remains of buildings) should have coherent signatures distinguishable from that of surrounding filling materials. The GPR data was acquired with a PulseEkko system using a pair of 500 MHz antennas in a common-offset bistatic configuration. The survey consisted of 63 lines that varied in lengths but were separated from each other using a consistent line-spacing of 0.5 m. Each

line was recorded with a trace density of 40 stations (traces) per meter and a total of 500 time samples, which corresponded to an acquisition window of 100 ns. The processing workflows implemented by de Smet (2016), shown in figure 6.1, involved a dewow filter (applied as a bandpass filter with $f_{\text{high-pass}} = 300$ MHz and $f_{\text{low-pass}} = 700$ MHz), an ATS background noise filter, a 2D FFT Stolt migration, and an AGC. In addition, de Smet (2016) used the energy and instantaneous amplitude (or envelope) attributes to aid in the interpretation of the processed data. The following explores the product of this workflow using IVGPR as well as a modified version proposed by the author.

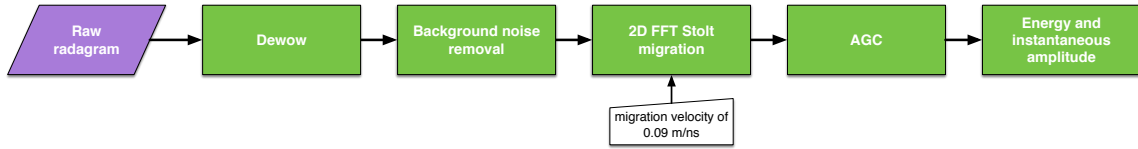


Figure 6.1: Workflow proposed by de Smet (2016).

The raw radagram for line 35 of this survey is shown in figure 6.2a. As evidenced by this image, raw records generally reveal only a limited amount of information. Figure 6.2b shows the result obtained after applying a Gazdag migration and, subsequently, an instantaneous AGC. The migration was carried out using a constant subsurface velocity of 0.09 m/ns (corresponding to $\varepsilon_r = 11.1$) and, therefore, reduces to the Stolt migration used by de Smet (2016). This migration was preceded by a low cut dewow filter and background noise removal, in agreement with figure 6.1. The instantaneous AGC was computed using a time gate of 2.2 ns. Figures 6.2c and 6.2d show the energy and instantaneous amplitude attributes, respectively. While the latter does not require any parameters, the former was obtained using 8 gray levels and a window size of 15 by 15 elements. It is worth mentioning that energy here was computed using the textural attribute routine of § 4.3.1, which

might differ from the method used by de Smet (2016). The shallow section of figure 6.2b (above 0.25 m) shows a dense set of energetic diffractions likely due to a noisy region that precedes the advent of the first arrivals. Without a sufficiently large number of padded elements, this energy would wrap around introducing numerous artifacts that could render the data uninterpretable. This problem can be avoided by properly implementing a time-zero correction prior to the migration step. Note that, unlike figure 6.2a, the migrated image and the attributes display an abundance of details about the subsurface and, therefore, are better suited for interpretation. This workflow is common among GPR practitioners and, as demonstrated in here, can be successfully implemented using the IVGPR package.

Figure 6.4 represents a modification to the workflow introduced above. The first step refers to the combination of three routines suited to derive a comprehensive assessment of the data's quality. This, in turn, helps to optimize ensuing steps in the processing flow. In particular, these routines are the gradient magnitude, multi-taper power spectrum estimator (MTPSE), and programmed gain. The first of these procedures can reveal the presence of random noise throughout the data and, thus, justify the use of a temporal trimming and/or denoising. It also provides images of the subsurface structure without the need of parametrization. The MTPSE generates a highly accurate rendering of the spectral content for the entire radagram and, therefore, may reveal the presence of unwanted signals that can be treated using frequency filtering. The programmed gain, like the gradient direction, can help unveil important features in signal depleted regions in a manner familiar to most users. Figures 6.4a and 6.4b shows an example in which the gradient direction attribute was used to uncover the presence of random noise and later confirm the effectiveness of the denoising procedure. A dashed line corresponds to the temporal level below which the signal-to-noise ratio exceeds the desired tolerance. Since the performance of these three quality assessment routines is virtually independent of parameters, users are advised to implement them at the early stages of a georadar investigation.

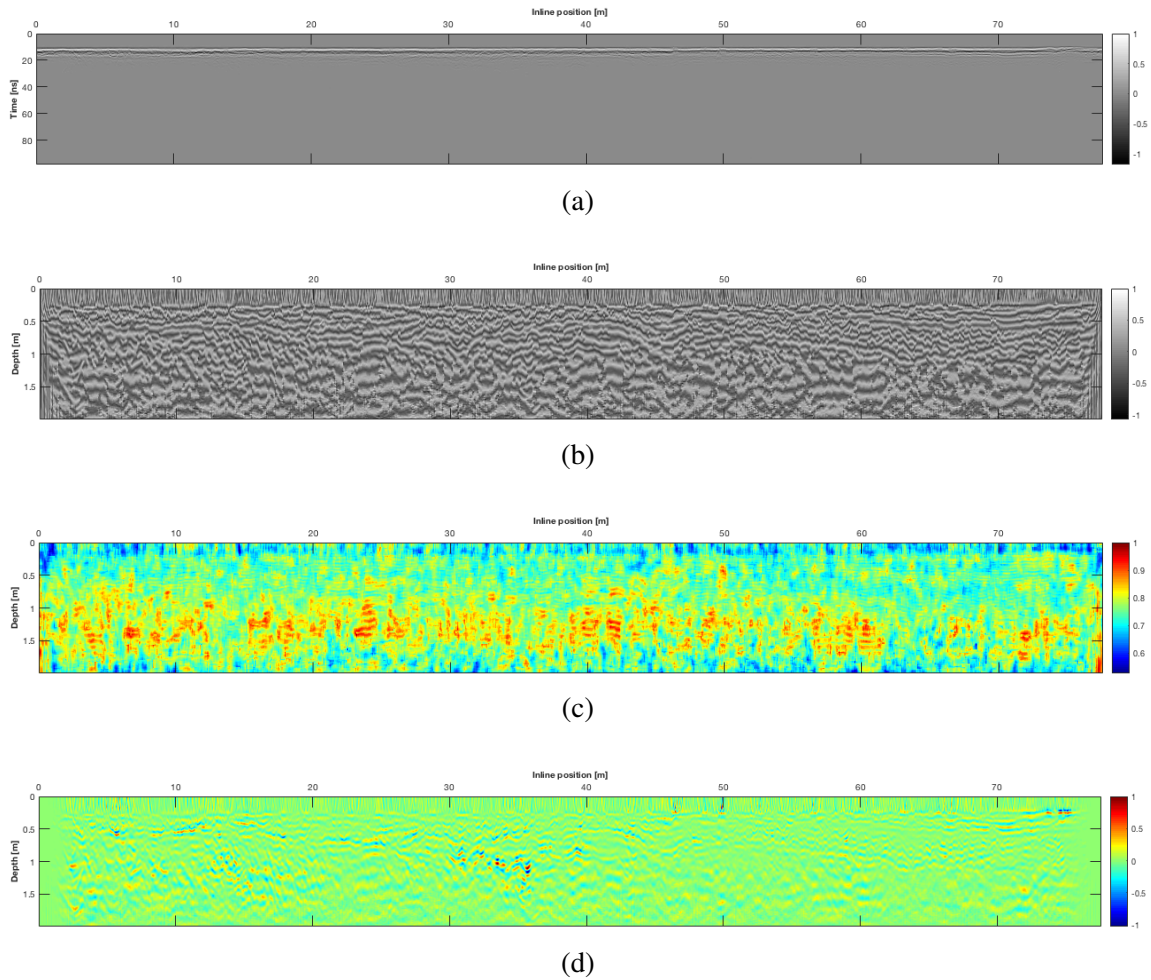


Figure 6.2: Outputs from a typical workflow implemented in IVGPR. (a) shows the raw line. (b) shows the migrated data after AGC. (c) and (d) show the energy and instantaneous attributes respectively.

The next two steps in the workflow involve an optional temporal trimming and a time-zero correction. The former should be implemented based on the results obtained from the quality analysis while the latter is applied following the methods described in §3.4 and § 4.4.3. Steps three through six are analogous to the ones in de Smet’s workflow, with the exception of a time-zero correction. Step seven involves the implementation of TVSW which, as described in § 4.4.1.1, acts similarly to a temporal gain but can enhance the

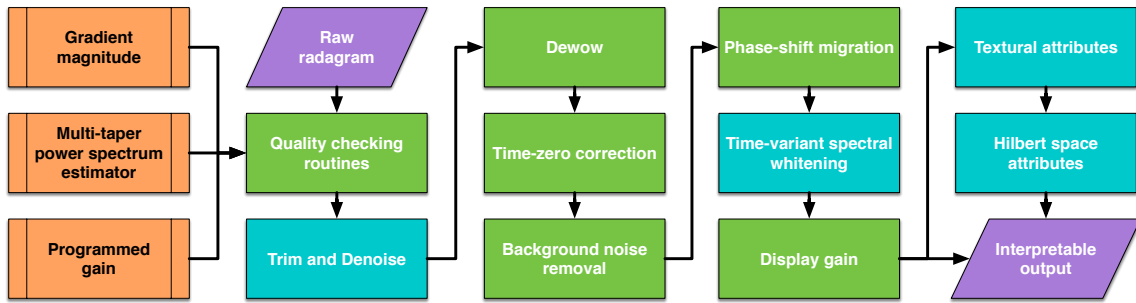
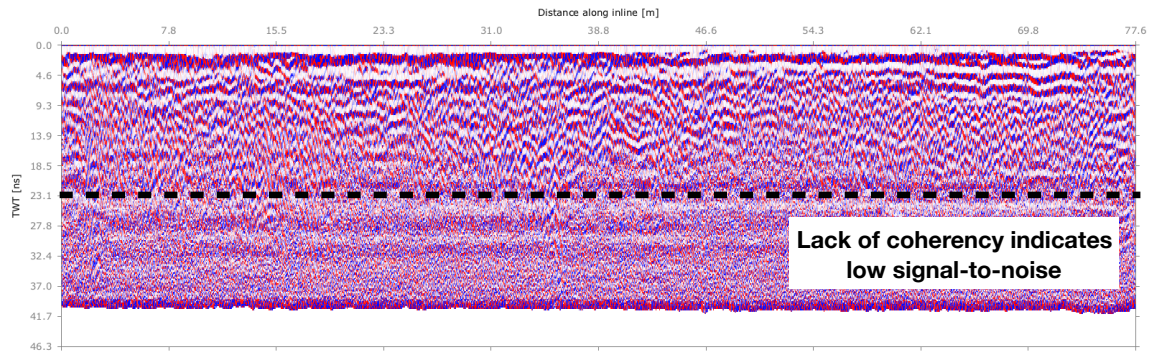


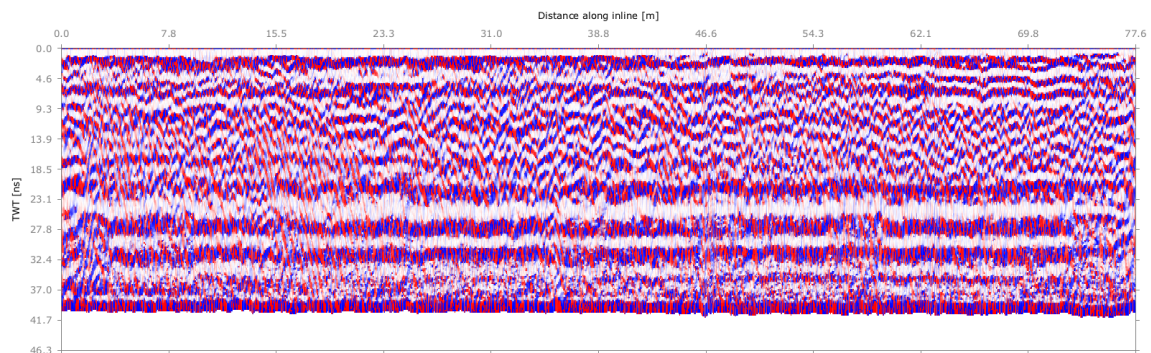
Figure 6.3: New workflow proposed to produce improved results on this dataset.

vertical resolution of the data. Even though spectral balancing routines are not typical in GPR studies, the author has found this technique pleasantly effective. The choice of display gain for step eight should align with the goals of the study. For instance, if a textural attribute analysis is to be performed, an AGC gain can prove detrimental by producing unphysical variations in amplitude.

Figure 6.5 shows line 32 and line 58 after applying steps one through seven of the proposed workflow. The location of these lines within the recreational yard are shown in figure ??a. Figure ??b shows a depth slice from the original survey and also indicates the position of the two lines. As shown in figure 6.6a, line 58 traverses an underground tunnel located about 10 meters inline. The migrated radiogram of figure 6.5b shows a trapezoidal feature (indicated by the red arrow) at a similar inline location and at a depth of 0.6 meters, which presumed to be related to this archeological structure. A region of anomalously low amplitude is enclosed by a dashed box. The source of this peculiar feature was not identified. Line 58 traverses a large tunnel believed to lie underneath the center of the recreational yard, as indicated by figure 6.6a. A strong reflector with the appropriate geometrical characteristics one would expect of this feature is evident in figure 6.5a. This case study demonstrated that IVGPR can be used to for practical GPR studies. In addition, a workflow useful for basic to mildly complex georadar surveys was



(a)



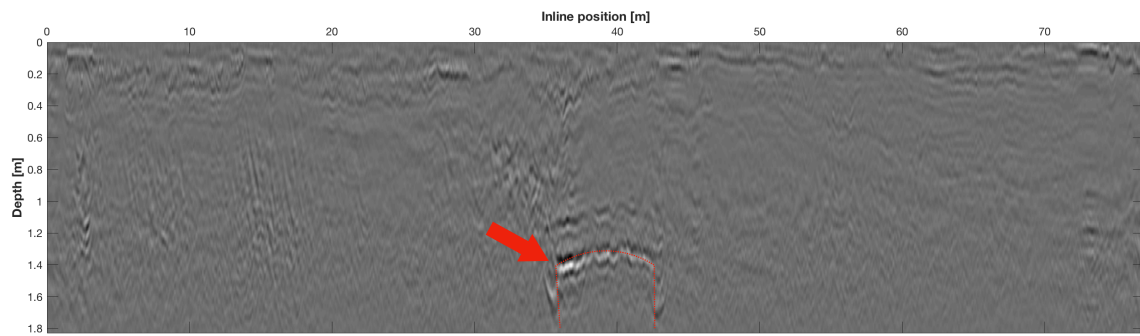
(b)

Figure 6.4: (a) Gradient direction attribute used to reveal the underlying geometry and to assess the level of noise in the image. The dashed line separates the regions of high and low signal-to-noise ratios. (b) Shows the same B-scan after applying the 2D adaptive denoising routine in IVGPR. The coherency of events for this image appears to be much improved, which corresponds to a lower number of randomly oriented events.

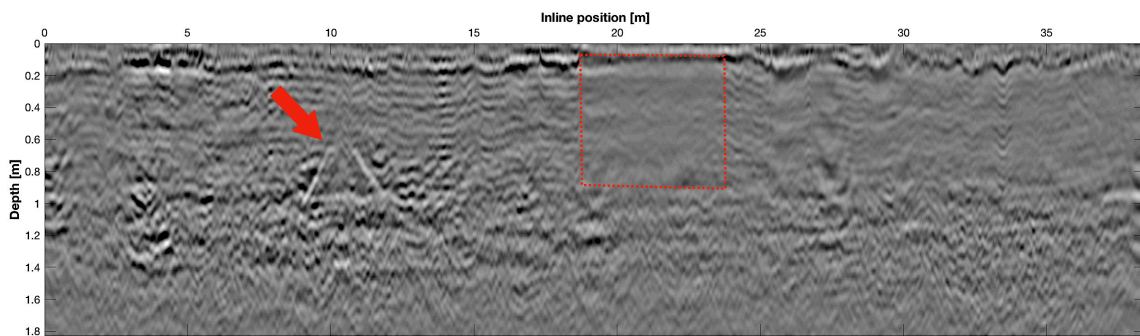
proposed. In the subsequent case study, this workflow will be applied to a field line in order to characterize a fracture network.

6.3 Case II: Field record over exposed fault near Mason, Texas

The record processed in this section was acquired over a densely fractured area located in the western portion of the Llano Uplift near Mason, Texas. It formed part of a multi-azimuth polarimetric GPR survey conducted by Amara (2016). In his work, Amara



(a)



(b)

Figure 6.5: Line 32(a) and line 58(b) from the study conducted over the recreational yard of Alcatraz prison island. Two red arrows indicate the inferred location of underground tunnels. A red square encloses a region of anomalously low amplitudes.

(2016) sought to obtain a detailed characterization of the underlying near-surface fracture network. The original survey constituted a total of four lines (labeled A through D) each acquired using three distinct antenna-receiver configurations. Line A in PP-mode was analyze for this case study, where PP indicates that the antenna and receiver where both parallel to the inline direction. This line traversed a high-angle normal fault, shown in figure 6.7, that trends north-east to south-west and dips to the south-east. Figure 6.7 depicts the position of the GPR line with respect to the fault. Note that the line was acquired in the northward direction. For a detailed account of the geology pertinent to this study, the

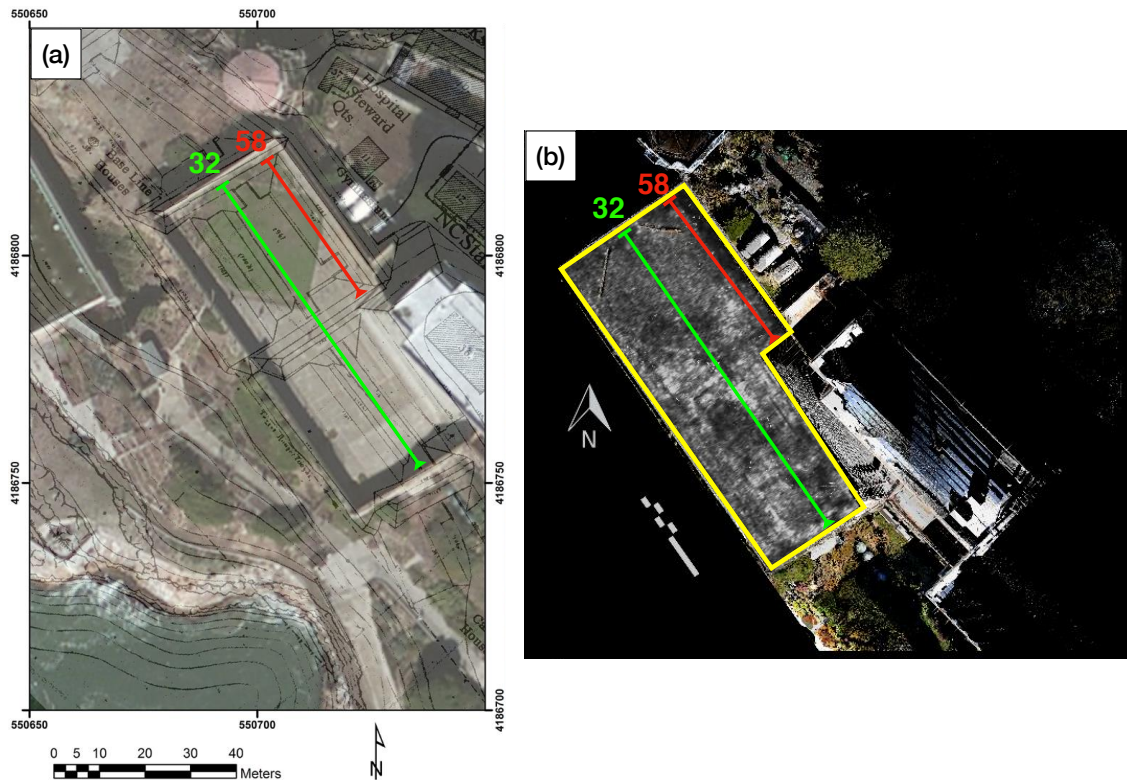


Figure 6.6: (a) Map showing the recreational yard of the main prison in Alcatraz island. The green and red lines indicate the position of lines 32 and 58, repetitively. A rendering of the believed architecture of the military fortification that predates the prison can be seen overlaying the map. Line 32 traverses a tunnel near its mid-section while line 58 traverses a tunnel nearly 10 meters inline. (b) Shows the locations of the lines with respect to a depth slice from the original publication by de Smet (2016). Modified for this work with the consent of Mark E. Everett and Timothy S. de Smet.

reader should visit Amara (2016). The examples in this section focus on demonstrating how IVGPR can generate accurate images that offer improvements from those published in the original work, which involved the use of a commercial processing package.

This survey was carried out using a PulseEKKO system with a 200 MHz antenna-receiver pair. The B-scan used here is composed of 150 traces uniformly spaced 0.2 meters apart. Each trace contains 750 samples that amounts to a total time window of 300 ns. In his work, Amara (2016) processed this line by applying a dewow filter, background



Figure 6.7: Shows the location of a high-angle fault traversed by the GPR line A to A'. The line was acquired in the northward direction while the fault dips to the south-east.

noise filter, Stolt migration, and an AGC. Note that, aside from the attribute analysis step, this workflow is identical to the one implemented by de Smet (2016). In IVGPR, line A was processed using steps one through eight of the modified workflow introduced in the previous section. This was followed by a manual mapping of numerous fractures in the processed record. The general location of the fault was first inferred during the initial quality assessment of the data, which corresponds to step one of the modified workflow. The author noticed that, after a programmed gain was applied, a cluster of hyperbolas appeared about 17 meters inline. A dewow high-pass filter with a cut-off frequency of 35 MHz was applied to generate figure 6.8, since applying a gain alone strengthened a high amplitude and low frequency signal that made it difficult to conduct an effective quality assessment.

Figure 6.9a shows the radagram processed in IVGPR as well as a network of fractures manually picked. This same line, processed and interpreted by Amara (2016), is shown in figure 6.9b. The fractures indicated with orange curves show instances where analogous picks were made in both images. On the other hand, fractures mapped with green curves

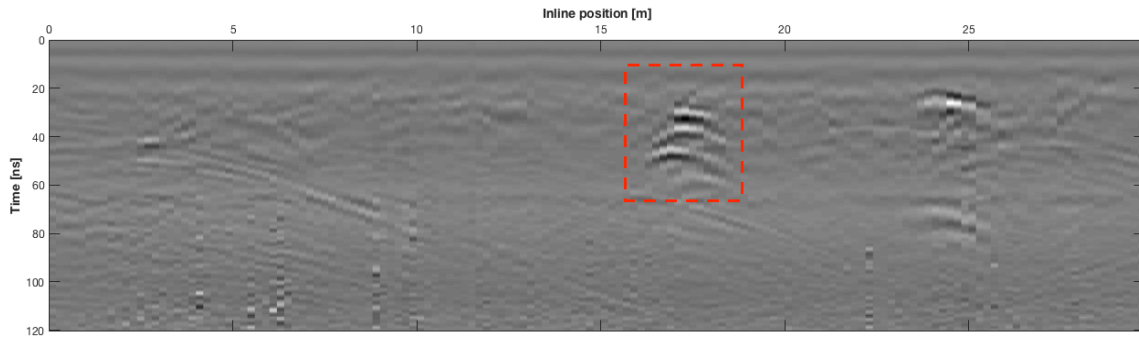
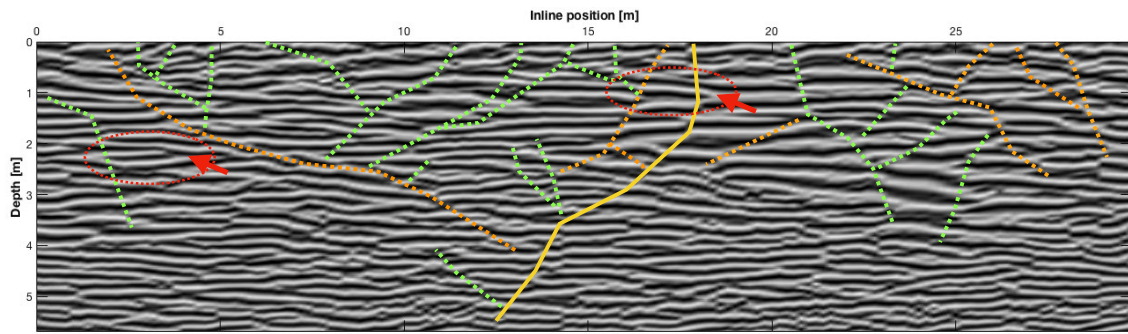


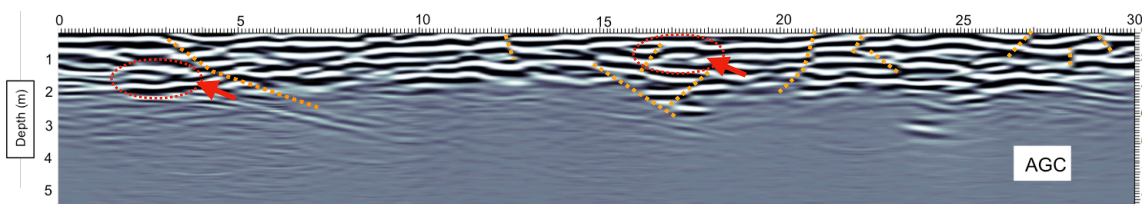
Figure 6.8: Line A after a low-cut dewow and a programmed gain were applied in order to highlight strong reflectors masked by the direct arrival. A dashed box encloses a cluster of hyperbolic events initially believed to be caused by a local fault.

belong to those newly found by the author. Two arrows indicate where a pair of peculiar features appear in both images, which shows that both processing flows generated comparable results. A yellow curve identifies the fault within the dataset. The most significant differences between these two images relate to the vertical resolution and depth for which effective interpretations can be drawn in each. In particular, the image processing using IVGPR shows a higher vertical and lateral resolution, manifested by visibly thinner beds and clearer discontinuities. Additionally, while both records show similar depth ranges, the image from Amara (2016) reveals little detail below the 2 m mark.

To process this data, the author first used a 2D adaptive denoising routine with a window size of 5 by 5 elements. Next, a high-pass dewow filter was used to remove amplitudes below 35 MHz and a time-zero correction set to s-crossover aligned the first arrivals at a common datum. This step was followed by an ATM background subtraction with $\alpha = 0.25$ and a window size of 51 elements. A phase-shift migration was carried out using a constant velocity of 0.09 m/ns in agreement with that used by Amara (2016). The final two steps involved a TVSW followed by an instantaneous AGC. The TVSW was applied using four filters within the frequencies of 50 to 350 MHz, while a time gate of 5.2 ns was used for the instantaneous AGC. It is worth noting some of the results of different time-zero



(a)



(b)

Figure 6.9: (a) Shows the B-scan produced using IVGPR after implementing the workflow proposed on the previous section. A number of visible fractures were mapped. Those marked with orange curves correspond to picks matching the ones made by Amara (2016). Newly interpreted picks are indicated by green curves and the inferred location of the large normal fault is shown by the yellow curve. Two red arrows indicate regions where the data processed by Amara (2016) (b) and that processed by the author (a) show a pair of peculiar features in good agreement.

correction methodologies. Amara (2016) used temporal clipping rather than an automatic or manual selection of the first arrival. The quality of the image produced by IVPGR allowed for a comprehensive fracture characterization in a greater amount of detail than that by an experienced user using a commercial software. This should further evidence the ability of this program to produce satisfactory results when treating field records.

7. CONCLUDING REMARKS AND FUTURE WORK

This work presented the major features of a new GPR processing package intended to be powerful, free, and flexible. The author hopes to have provided enough justification, through effective examples, in support of the adoption of IVGPR for academic georadar studies. An extensive search failed to reveal a free GPR processing package program that compares to IVGPR in the number and sophistication of the available functionalities. Furthermore, this program offers the rare opportunity to implement 3D vector migrations as part of a processing workflow. The unavailability of these techniques has been known to hinder the performance of advanced georadar procedures such as inversion and true amplitude processing.

Currently this program lacks a user interface that is easy to use for those without a comfortable level of expertise in Fortran programming. Future iterations of this program will be wrapped in python and will provide a dedicated GUI in order to remove the need to interact with the source code. In addition, the author will develop python-based methods to read encrypted GPR files and visualize 3D GPR records. In its current state, the author sees the usefulness of IVGPR limited to a small niche of advanced users in the academic sectors. By releasing a version better suited for the masses, this program can possibly replace many of the feature-deprived options currently available at no cost.

REFERENCES

- Amara, A., 2016, Using multi-azimuth and multi-polarization ground penetrating radar to characterize a fractured fault zone in Mason County, Texas [Master's thesis]: Texas A&M University.
- Annan, A., 1999, Practical processing of GPR data: Sensors and Software, Ontario.
- Bednar, J., and Watt, T., 1984, Alpha-trimmed means and their relationship to median filters.: IEEE Transactions on Acoustics, Speech & Signal Processing, v. **32**, no. 1, p. 145.
- Belina, F., Dafflon, B., Tronicke, J., and Holliger, K., 2009, Enhancing the vertical resolution of surface georadar data: Journal of Applied Geophysics, v. **68**, no. 1, p. 26–35.
- Bernstein, M.A., King, K.F., and Zhou, X.J., 2004, Handbook of MRI pulse sequences: Elsevier.
- Boashash, B., 1992a, Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals: Proceedings of the IEEE, v. **80**, no. 4, p. 520–538.
- Boashash, B., 1992b, Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications: Proceedings of the IEEE, v. **80**, no. 4, p. 540–568.
- Cassidy, N.J., and Jol, H., 2009, Ground penetrating radar data processing, modelling and analysis: Ground Penetrating Radar Theory and Applications, p. 141–176.
- Chun, J.H., and Jacewitz, C.A., 1981, Fundamentals of frequency domain migration: Geophysics, v. **46**, no. 5, p. 717–733.
- Claerbout, J.F., 1971, Toward a unified theory of reflector mapping: Geophysics, v. **36**, no. 3, p. 467–481.
- Claerbout, J.F., and Green, I., 2008, Basic earth imaging: Stanford University.
- Clerman, N.S., and Spector, W., 2011, Modern Fortran: Style and usage: Cambridge

University Press.

- Coppens, F., 1985, First arrival picking on common-offset trace collections for automatic estimation of static corrections: *Geophysical Prospecting*, v. **33**, no. 8, p. 1212–1231.
- de Smet, T.S., 2016, *Advances in archeological geophysics: Case studies from historical archeology* [Ph.D. thesis]: Texas A&M University.
- Denis, A., and Crémoux, F., 2002, Using the entropy of curves to segment a time or spatial series: *Mathematical Geology*, v. **34**, no. 8, p. 899–914.
- Engheta, N., Papas, C.H., and Elachi, C., 1982, Radiation patterns of interfacial dipole antennas.: *Radio Science*, v. **17**, p. 1557 – 1566.
- Everett, M.E., 2013, *Near-surface applied geophysics*: Cambridge University Press.
- Feng, X., Liu, C., Sato, M., et al., 2007, Interpolation using the Kirchhoff migration integral for visualization of an irregular GPR data, in 2007 SEG Annual Meeting: Society of Exploration Geophysicists.
- Gao, D., 2011, Latest developments in seismic texture analysis for subsurface structure, facies, and reservoir characterization: A review: *Geophysics*.
- Gazdag, J., 1978, Wave equation migration with the phase-shift method: *Geophysics*, v. **43**, no. 7, p. 1342–1351.
- Gerlitz, K., Knoll, M.D., Cross, G.M., Luzitano, R.D., and Knight, R., 1993, Processing ground penetrating radar data to improve resolution of near-surface targets, in 6th EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems: p. 561–574.
- Gu, K., Wang, G., and Li, J., 2004, Migration based SAR imaging for ground penetrating radar systems: *IEEE Proceedings-Radar, Sonar and Navigation*, v. **151**, no. 5, p. 317–325.
- Gustafson, J.L., 1988, Reevaluating Amdahl's law: *Communications of the ACM*, v. **31**, no. 5, p. 532–533.

- Hale, D., 1991, Stable explicit depth extrapolation of seismic wavefields: *Geophysics*, v. **56**, no. 11, p. 1770–1777.
- Hale, D., 2009, Structure-oriented smoothing and semblance: CWP.
- Jähne, B., 2005, Digital image processing: The Netherlands: Springer-Verlag.
- Jin, F., Fieguth, P., Winger, L., and Jernigan, E., 2003, Adaptive Wiener filtering of noisy images and image sequences, in Proceedings 2003 International Conference on Image Processing, v. 3: IEEE, p. III–349–52 vol.2.
- Jol, H.M., 2008, Ground penetrating radar theory and applications: Elsevier, 544 p.
- Jol, H.M., and Bristow, C.S., 2003, GPR in sediments: advice on data collection, basic processing and interpretation, a good practice guide: Geological Society, London, Special Publications, v. **211**, no. 1, p. 9–27.
- Jones, I., 2014, Tutorial: migration imaging conditions: *First Break*, v. **32**, no. 12, p. 45–55.
- Lee, M.W., 1986, Spectral whitening in the frequency domain, No. 86-108, US Geological Survey.
- Leuschen, C., and Plumb, R., 2000, A matched-filter approach to wave migration: *Journal of Applied Geophysics*, v. **43**, no. 2, p. 271–280.
- Leuschen, C.J., and Plumb, R.G., 2001, A matched-filter-based reverse-time migration algorithm for ground-penetrating radar data: *Geoscience and Remote Sensing, IEEE Transactions on*, v. **39**, no. 5, p. 929–936.
- Liu, Y., and Fomel, S., 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: *Geophysics*, v. **76**, no. 5, p. V69–V77.
- Liu, Y., Luo, Y., and Deng, Z., 2003, Edge preserving smoothing for oblique images: *Geophysical research letters*, v. **30**, no. 13.
- Marple, L., 1999, Computing the discrete-time "analytic" signal via FFT: *IEEE Transactions on signal processing*, v. **47**, no. 9, p. 2600–2603.

- Martins, S.S., Sacchi, M., Travassos, J.M., et al., 2015, Interpolating GPR data using Singular Spectrum Analysis (SSA) and Anti-alias SSA, in 2015 SEG Annual Meeting: Society of Exploration Geophysicists.
- Neto, P.X., and de Medeiros, W.E., 2006, A practical approach to correct attenuation effects in GPR data: *Journal of Applied Geophysics*, v. **59**, no. 2, p. 140–151.
- Olhoeft, G.R., 2000, Maximizing the information return from ground penetrating radar: *Journal of Applied Geophysics*, v. **43**, no. 2, p. 175–187.
- Oppenheim, A.V., Schafer, R.W., Buck, J.R., et al., 1989, *Discrete-time signal processing*, v. 2: Prentice-hall Englewood Cliffs.
- Orlando, L., 2002, Detection and analysis of LNAPL using the instantaneous amplitude and frequency of ground-penetrating radar data: *Geophysical Prospecting*, v. **50**, no. 1, p. 27–41.
- Ozdemir, Caner and Demirci, Sevket and Yigit, Enes and Kavak, Adnan, 2007, A hyperbolic summation method to focus B-scan ground penetrating radar images: An experimental study with a stepped frequency system: *Microwave and Optical Technology Letters*, v. **49**, no. 3, p. 671–676.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., 1992, *Numerical recipes in FORTRAN*: Cambridge university press.
- Prieto, G., Parker, R., and Vernon III, F., 2009, A Fortran 90 library for multitaper spectrum analysis: *Computers & Geosciences*, v. **35**, no. 8, p. 1701–1710.
- Qin, Y., Wang, Q.F., Qiao, L.H., and Ren, X.Z., 2014, Research on ground penetrating radar migration imaging technology: *Sensors & Transducers*, v. **180**, no. 10, p. 51.
- Rashed, M., and Harbi, H., 2014, Background matrix subtraction (BMS): A novel background removal algorithm for GPR data: *Journal of Applied Geophysics*, v. **106**, no. 0, p. 154 – 163.
- Sabbione, J.I., and Velis, D., 2010, Automatic first-breaks picking: New strategies and

- algorithms: *Geophysics*, v. **75**, no. 4, p. V67–V76.
- Sajid, M., and Ghosh, D., 2014, A fast and simple method of spectral enhancement: *Geophysics*, v. **79**, no. 3, p. V75–V80.
- Sandmeier, K.J., 2002, Manual of REFLEXW Version 2.5, Program for processing and Interpretation of seismic, acoustic or electromagnetic reflection, refraction and transmission Data.
- Sassen, D.S., and Everett, M.E., 2009, 3d polarimetric gpr coherency attributes and full-waveform inversion of transmission data for characterizing fractured rock: *Geophysics*, v. **74**, no. 3, p. J23–J34.
- Scales, J.A., 1995, *Theory of seismic imaging: Lecture notes in earth sciences: 55*, Berlin, Heidelberg : Springer Berlin Heidelberg, 1995.
- Schmelzbach, C., Tronicke, J., and Dietrich, P., 2012, High-resolution water content estimation from surface-based ground-penetrating radar reflection data by impedance inversion: *Water Resources Research*, v. **48**, no. 8.
- Slob, E., 2003, Toward true amplitude processing of GPR data, in *Proceedings of the 2nd International Workshop on Advanced Ground Penetrating Radar, 2003.* : IEEE, p. 16–23.
- Smith, S.W., et al., 1997, *The scientist and engineer's guide to digital signal processing: California Technical Pub. San Diego.*
- Stolt, R.H., and Weglein, A.B., 2012, *Seismic imaging and inversion : application of linear inverse theory.:* Cambridge ; New York: Cambridge University Press, 2012.
- Streich, R., and van der Kruk, J., 2007, Accurate imaging of multicomponent gpr data based on exact radiation patterns: *IEEE Transactions on Geoscience and Remote Sensing*, v. **45**, no. 1, p. 93–103.
- Streich, R., van der Kruk, J., and Green, A.G., 2007, Vector-migration of standard copolarized 3D GPR data: *Geophysics*, v. **72**, no. 5, p. J65–J75.

- Taner, M.T., Koehler, F., and Sheriff, R., 1979, Complex seismic trace analysis: *Geophysics*, v. **44**, no. 6, p. 1041–1063.
- Tronicke, J., Schennen, S., and Allroggen, N., 2015, Spectral enhancement of GPR data: Some practical considerations, in 8th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 2015: IEEE, p. 1–4.
- Tzanis, A., 2010, matGPR release 2: A freeware MATLAB® package for the analysis & interpretation of common and single offset GPR data: *FastTimes*, v. **15**, no. 1, p. 17–43.
- van der Kruk, J., 2001, Three-dimensional imaging of multi-component ground penetrating radar [Ph.D. thesis]: Delft University of Technology.
- van der Kruk, J., 2015, 11.07 Tools and Techniques: Ground-Penetrating Radar, in Schubert, G., ed., *Treatise on geophysics: 2nd Edition*, Oxford, Elsevier, p. 209–232.
- van der Kruk, J., Wapenaar, C., and Fokkema, J.T., 2000, Multicomponent 3D imaging of ground penetrating radar data using matrix inversion in the spatial Fourier domain, in 8th International Conference on Ground Penetrating Radar: International Society for Optics and Photonics, p. 508–513.
- Warren, C., Giannopoulos, A., and Giannakis, I., 2015, An advanced GPR modelling framework: The next generation of gprMax, in 8th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 2015: p. 1–4, doi:10.1109/IWAGPR.2015.7292621.
- Yelf, R., 2004, Where is true time zero?, in *Proceedings of the Tenth International Conference on Ground Penetrating Radar, 2004. GPR 2004.*, v. 1: IEEE, p. 279–282.
- Yilmaz, Ö., 2001, *Seismic data analysis*, v. 1: Tulsa, OK, Society of Exploration Geophysicists.
- Zaremba, N.J., Smith, K.E., Bishop, J.M., and Smith, C.G., 2016, Ground-penetrating radar and differential global positioning system data collected from Long Beach Island, New Jersey, April 2015, , US Geological Survey.

- Zhang, H., Benedix, W.S., Plettemeier, D., and Ciarletti, V., 2013, Radar subsurface imaging by phase shift migration algorithm, in Radar Conference (EuRAD), 2013 European: p. 531–534.
- Zhao, W., Forte, E., and Pipan, M., 2016, Texture attribute analysis of GPR data for archaeological prospection: *Pure and Applied Geophysics*, p. 1–15.