Dedicada a
papá y mamá, que
siempre me apoyaron y
me brindaron su infinito amor
y cariño. ¡Los Amo mucho!
Alex ☺

JOINT INVERSION OF REFLECTIVITY AND BACKGROUND SUBSURFACE COMPONENTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Alejandro Cabrales Vargas

August 2020

JOINT INVERSION OF REFLECTIVITY AND BACKGROUND SUBSURFACE COMPONENTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Alejandro Cabrales Vargas
August 2020

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Biondo Biondi, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Robert Clapp

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jerry Harris

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

Current state-of-the-art subsurface imaging techniques have based upon two-way wave propagation solutions. At the beginning of the 2000s, reverse-time migration constituted the "main character" of the imaging toolbox: the ultimate solution for geologically complex areas. Nowadays it represents the main "engine" working under the hood of the most acclaimed and researched imaging techniques, such as linearized waveform inversion (a.k.a. least-squares migration) for the estimation of the subsurface reflectivity, migration velocity analysis for the estimation of the subsurface background component, and full-waveform inversion for the estimation of both components as a single model parameter set. The latter is arguably the most ambitious attempt to obtain a complete and accurate picture of the subsurface. Unfortunately, it is vulnerable to cycle-skipping, which drives the optimization process to convergence into local minima. Strategies to attack this problem include better acquisition, data processing, and different implementations of the full-waveform inversion algorithm. In this thesis, I explore a different and certainly less ambitious approach, which consists of the joint inversion of the subsurface reflectivity and the subsurface background component as different parameter sets. During the early stages of my research, I posed the problem as a linear optimization problem. However, this approach does not have the expected properties. Maintaining the migration image nonlinear with respect to the background model corrected the problem. I test this new optimization problem scheme in synthetic 2D data and a 3D ocean-bottom node dataset from the Gulf of Mexico.

Preface

The electronic version of this report[1] makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR] are promises by the author about the reproducibility of each figure result. Reproducibility is a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

ER denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library[2]. We assume you have a UNIX workstation with Fortran, Fortran90, C, C++, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

[1] http://sepwww.stanford.edu/private/docs/sep179
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. The primary reasons for the CR designation is that the processing requires 20 minutes or more, MPI or CUDA based code, or commercial packages such as Matlab or Mathematica.

NR denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.6 (using the Intel compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. More information on reproducing SEP's electronic documents is available online[3].

[3]http://sepwww.stanford.edu/research/redoc/

Acknowledgments

This section is perhaps the most difficult one of this dissertation for me to write. Although whatever flaw found in my research is solely my responsibility, the opposite is the truth; whatever contribution written here constitutes a learning experience I embarked on, and I owe gratitude toward many people. The problem is having someone out by accident or not honoring someone properly. I apologize beforehand for these people. If that happens, it is unintentional and merely a by-product of my non-brilliant mind.

I am deeply grateful to Biondo Biondi, who accepted me as a member of the prestigious SEP group and guided me throughout my Ph. D. journey. I am not exactly the kind of smart person he was surely expecting (as virtually all the SEPers are and have been). For that reason, I am also immensely grateful for his patience in my slow learning. Similarly, I also thank Robert "Bob" Clapp for his invaluable help in almost everything. Is there something about seismic imaging that he ignores? Is there a piece of code that still resists his attacks to maximize the performance? I still have no answer to these questions.

I want to thank the other committee members, Jenny Suckale, Gary Mavko, and Jerry Harris. I did my second project with Jenny, and thereby I contributed on a tiny scale to rescue our planet (by understanding ice dynamics in Antarctica). Thanks to her merciless reviews, I improved my writing skills when it comes to redacting scientific papers. Gary was part of my committees until he retired, and then Jerry generously agreed to take the vacant place. I am very grateful to both of them.

Of course, I have special gratitude toward Jon Claerbout, not only because he is the reason SEP exists. My inspiration for getting into seismic imaging was reading Imaging the Earth's Interior. I reproduced his Ratfor codes in languages that I could use. After discarding Matlab, I tried C++ and then Fortran... just to come back to C++ when Bob declared the "death" of Fortran. I think of him as a forever young man who never gets tired of sparking new ideas.

I thank Stewart Levin for his assistance in understanding the 3-D data that I used in this work, and in general, his help to all of us, the students.

I am in debt with the SEP crew for having, directly and indirectly, supported me. Each of their contributions was so important that the order in which I mention the people is irrelevant. Something important that they gave me as a group was the sense of humbleness, something that I need to cherish and transmit to the highly arrogant oil & gas world in Mexico. Rather than monotonically thanking each SEPer, I implicitly thank them by describing my relationship and perception of each one of them. I do not wish to omit that my intrinsic personality, that of an introvert, always afraid of saying B.S., unfortunately, was the stumbling block that did not allow me to be closer to them (or to other people in the department for that matters).

Yi Shen was my office-mate, and still is my ever-smiling friend that was always there to make me feel good when the "pre-quals" stress manifested.

Huy Le is an introvert person, almost like myself. However, he is the anisotropy expert, and a dangerous player of the GPU computing game (and also the soccer game!). From him, I learned that, despite starting a project being utterly ignorant about something (in this case, seismic imaging), one could become an expert with enough hard work and dedication.

The Ettore-Guillaume duet has helped most students, and I am very grateful to them for the excellent solver that saved me one year of work. Ettore is always smart and enthusiastic at explaining and sharing his vast knowledge. As for Guillaume, I have admired his tenacity and persistence about thoroughly understanding whatever topic.

That is more amazing considering his advanced age (at least in appearance!)

Rahul Sarkar is undoubtedly the group's expert in inverse theory (and more recently in quantum computing). Thanks to Rahul, I understood why the method that I humbly propose in this dissertation had no hope to work within the linear regime. Unfortunately, I could not stay longer at Stanford to work with him. I have to say that I was fortunate enough to have the opportunity to know him better during the period that we worked together when he accepted the challenge of suffering my slow learning.

Our most recent post-doc, Ariel Lellouch, is also my friend and a person full of sense of humor. I admire his leadership (something I am still not good at) and his fast thinking, so he is prone to produce rapid and pertinent opinions about research works, even when he is an expert on a particular topic.

Stuart "The Young" Farris (to be distinguished from Stewart "The Wise" Levin), among the other young people of the group, taught me the valuable lesson that one should listen to and learn from everybody, notwithstanding that he or she is younger by more than one decade. Stuart constitutes an epitome of a person that can be smart, young, and mature at the same time.

Joseph Jennings, "mi amigo Joe", always helped me to practice my Spanish. Seriously, he helped me better understand the 3-D dataset that I use for Chapter 5 and to understand the adjoint method better. He is also an excellent friend that always was there to listen to my weird stories. I will miss him so much!

Taylor Dahlke and I started a friendship soon after we began interacting, almost by chance at the beginning. He happened to be working with the same 3D dataset that I used in Chapter 5 and helped me understand the OBC acquisition. Moreover, he performed most of the processing that made the data ready for imaging, which was invaluable for saving time and relaxing my tight deadline. I will also miss him!

Mademoiselle Fantine Hout is a polyglot lady that speaks many languages, including

Machine Learning, every day more ubiquitous in the Geophysics community (Some people say that even Guillaume would be learning it soon). Always prone to give away a warm and charming smile, among other things, she assisted me when I started working in the cluster with the now famous (and thoroughly undecipherable) Yang's code. She also played the role of my adopted daughter when we mischievously visited Disneyland during the SEG Annual Meeting in Anaheim.

Yinbin Ma has been quite an enigmatic character. He helped me understand a lot of useful stuff during the implementation of my codes, including the violation of causality that prevented my Born operator from passing the dot-product test. He also wrote the code that I (and some other people) use to compute the random boundary condition. However, I have the feeling that sometimes he does not appreciate his skills. Therefore, I want to publicly manifest that he is an expert in inverse theory and computation sciences, besides an excellent friend.

Rustam Akhmadiev and Siyuan Yuan spent quite a long time with me when they join the group. Siyuan was my office-mate (until I decided to work outdoors). Rustam has contagious enthusiasm and mastery in muscle-ups. He also taught me that you could do it when you want to learn something (for instance, Spanish!)

Milad Bader helped me during my TAship, in a very particular way. He was taking the courses. The other student in the class has a hyperactive and inquiring mind, and his questions quite often went beyond my knowledge. Milad came to assist me in answering such questions. I arguably learned more stuff than that that I helped to teach during such periods. Thanks to both of them!

Now it is the turn of my friend Julio Frigerio, who unfortunately joined the group too close to my departure. If I felt that my situation was desperate, having five years to achieve graduation in SEP, he came with the consign to finish in only three and a half! Nevertheless, he shows an untamed spirit and such a badass attitude that I cannot but admire. Moreover, I certainly cannot hide my awe at someone who understands Einstein's General Relativity.

Special thanks to Kurt Marfurt, my masters' advisor in Oklahoma. He encouraged me to work in the least-squares Kirchhoff time migration. It was the moment when I jumped from seismic interpretation to seismic imaging.

Speaking of the Mexican crew, I am very thankful to my mentor, Sergio Chávez, who, since our encounter in his class (which I was merely auditing at that time), has never run out of advice. I also thank him for having made me moved from water to oil & gas exploration. Also, thanks to the Pemex crew: My current boss, Humberto Salazar; my short-time boss, Carlos Caraveo; my colleagues, Alfredo Vázquez, Israel Ronzón, Rubén Ponce, Sergio Romahn, Karen Hernández, Alejandra Méndez, Silvino Domínguez, Jorge Chagolla, Madai Galero, and Juan Nájera. Special thanks to my former boss, Marco Vázquez, who firmly pushed me to apply for Stanford when I was reluctant about that. Thanks to my everlasting friends Javier "Oso" Sánchez, Óscar "Cyranus" Boni, and Vicente "Rayita" Hernández, with whom I took my first steps toward the steep and thorny road of geophysics. Last but not least, thanks to my colleagues Humberto Arévalo, Moisés Hernández, and their families, who faced the challenge of being a graduate student at Stanford before me.

Of course, I dedicate this work to the loves of my life: my wife Patricia, and my daughters, Zaira and Parisa. They endured with me a five-years life adventure at Stanford, far from our beloved ones. I also dedicate my work to my parents, Manolo and Lucha, my siblings, Manuel, Roberto, Lucy and Lupita, and all my relatives, who never lost faith in me!

"...when you have eliminated the impossible, whatever remains, however improbable, must be the truth." Sherlock Holmes to Dr. Watson in "The sign of the four", by Arthur Conan Doyle.

Contents

List of Tables

List of Figures

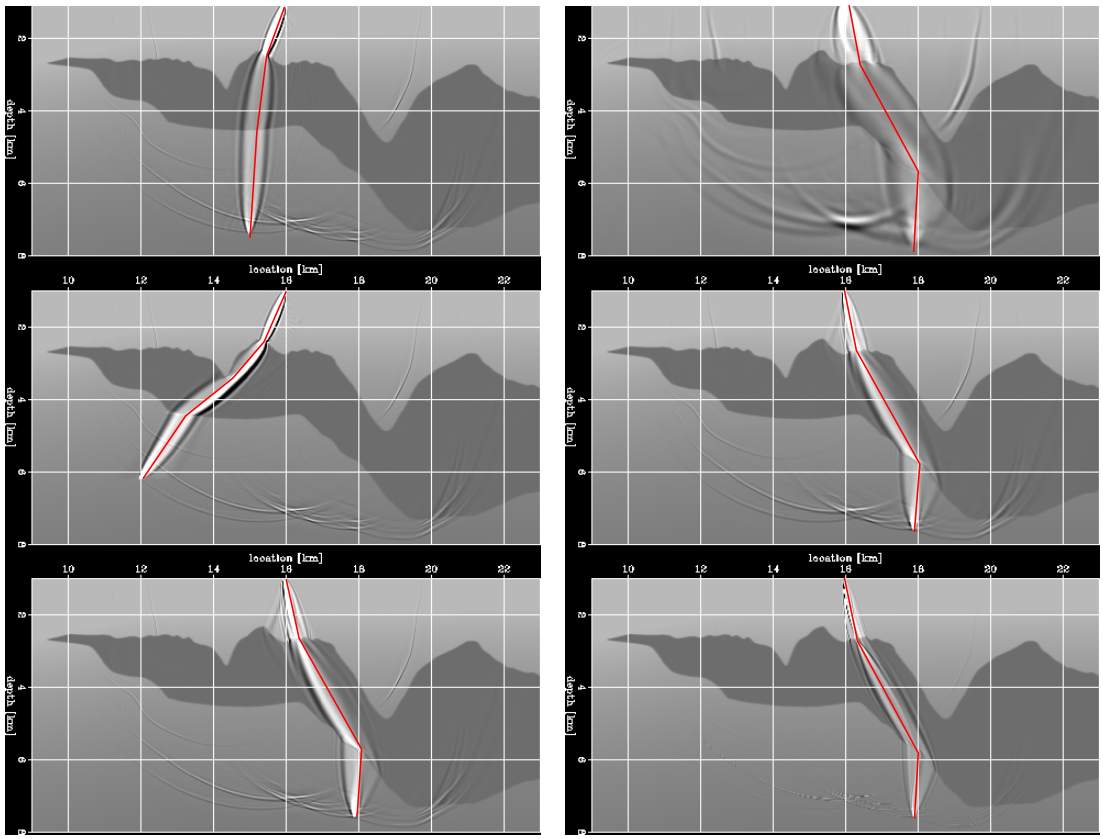
Chapter 1

Introduction

Until recently, worldwide crescent demand for hydrocarbons imposed oil & gas geoscientists the task of continuous improvements in exploration techniques. Nowadays, strategies have diversified and pointed toward more efficient recovery and less environmental impact by better drilling planning. Of particular importance is to have robust algorithms for accurately imaging the subsurface. We typically evaluate the robustness of imaging algorithms according to their ability to obtain 1) reliable subsurface velocity/density models, and 2) accurate estimation of the subsurface reflectivity.

We achieve the first objective when the estimated velocity and density models accurately position subsurface reflectivity contrasts during the imaging process. Early estimations of velocity make use of ray tracing tomography (still the workhorse of the industry), which resorts to the wave equation's high-frequency approximation. As such, wavepaths are represented as rays (Figure 1.1). However, such a representation frequently fails in the presence of sharp velocity contrasts, often leaving vast "shadows zones", i.e., areas of poor or null ray coverage. Better results can be obtained using wave-equation based velocity estimation methods, such as wave-equation migration velocity analysis (WEMVA) (Biondi and Sava, 1999; Biondi, 2006) and full-waveform inversion (FWI) (Tarantola, 1984a; Virieux and Operto, 2009), which perform the

optimization in the image domain and the data domain, respectively. These methods exploit the band-limited character of the wavepaths that makes them sensitive to broader areas. They constitute the well-known "banana-doughnut" sensitivity kernels (Tromp and Tape, 2005) that reduce the occurrence of shadow zones, even in the presence of complex geology, as illustrated in Figure 1.1.



(a)

Figure 1.1: Wavepaths vs. rays. Rays operate under the high-frequency approximation of the wave equation, whereas wavepaths are band-limited entities. Notice how the area of influence of the wavepaths (known as "banana-doughnut" sensitivity kernels, because of their shape in 3D) is significantly larger than that of the rays, which often leave shadow zones in complex areas such as underneath salt bodies. Modified from Biondi (2006). [**NR**]

Regarding the second objective, and assuming that we fulfilled the first one, we say that the reflectivity amplitude is accurate when it truly represents lithology and fluid

content heterogeneity in the subsurface (e.g. Ha and Marfurt, 2017). If we achieve this objective, we can rely on amplitude variation to perform qualitative and quantitative estimation of reservoir properties.

After estimating the velocity and density subsurface models (most of the time, we regard the latter as constant, for convenience), we can confidently image the subsurface seismic events by employing conventional imaging methods such as Kirchhoff migration and wave-equation migration. Seismic migration is *kinematically* correct. However, it does not generally recover amplitude information that serves the purpose of interpreting the rocks and fluids contrasts. In the case of Kirchhoff migration, seismic amplitudes can be corrected using weighting coefficients derived from approximate solutions to the transport equation (Schleicher et al., 2007). For wave-equation migration, such as reverse-time migration (RTM), we can balance the amplitudes to some extent by utilizing the deconvolution imaging condition (Valenciano and Biondi, 2003).

None of these conventional imaging methods addresses illumination problems. Likewise, they often have less than optimal seismic resolution because of the inherently blurring effect of seismic migration (Yu et al., 2006). These tasks can be achieved by means of least-squares migration (LSM) (Nemeth and Schuster, 1999), also known as linearized waveform inversion (LWI) (Tarantola, 1984b). LWI can be implemented in *data space* (e.g. Nemeth and Schuster, 1999; Duquet et al., 2000; Dai and Schuster, 2013; Wong et al., 2015; Fletcher et al., 2016), i.e., inverting for a reflectivity model, $\mathbf{r}$, from which data are synthesized by using the Born modeling operator, $\mathbf{L}$. The reflectivity model is iteratively updated until the misfit between the corresponding synthetic data, $\mathbf{d}(\mathbf{b}) = \mathbf{L}(\mathbf{b})\mathbf{r}$, and the recorded field data, $\mathbf{d}_{obs}$, becomes small in the least-squares sense.

We typically pose LWI in data domain as follows,

$$\Phi(\mathbf{r}) = \frac{1}{2}\|\mathbf{L}(\mathbf{b})\mathbf{r} - \mathbf{d}_{\text{obs}}\|_2^2, \tag{1.1}$$

where Φ is the misfit function, and $\mathbf{b}$ is the background component of the subsurface model, either velocity or slowness or slowness squared.

LWI can also be implemented in *model space* (e.g. Valenciano et al., 2006; Valenciano, 2008; Valenciano and Biondi, 2009; Fletcher et al., 2016). We invert for a reflectivity model that matches the migration image, $\mathbf{I}$, by means of the action of the Gauss-Newton Hessian of FWI, $\mathbf{H}$, as follows:

$$\Phi(\mathbf{r}) = \frac{1}{2}\|\mathbf{H}(\mathbf{b})\mathbf{r} - \mathbf{I}(\mathbf{b})\|_2^2, \tag{1.2}$$

where $\mathbf{H} = \mathbf{L}^{\mathrm{T}}\mathbf{L}$, and $\mathbf{L}^{\mathrm{T}}$ represents the *adjoint* of the Born modeling operator, a.k.a. the RTM operator. The advantage of implementing LWI in model space in comparison with the implementation in data space is that, once we estimated the Hessian, the inversion consists of matrix-like multiplications. However, the main drawback is that the Gauss-Newton Hessian needs to be pre-computed and stored. We often resort to affordable approximations such as point-spread functions (PSF) (Valenciano, 2008; Fletcher et al., 2016) for the Hessian estimation.

1.1 Setting the problem

Reflectivity estimation using LWI relies upon the availability of an accurate subsurface background model. Otherwise, the image will not be well focused and the inversion can exhibit slow convergence. There have been efforts to nullify this limitation, such as extending the domain in subsurface offset (e.g. Yang et al., 2018). However, remaining errors in the velocity model still make seismic reflections defocus. In this dissertation, I analyze the inaccuracy in the background model from a different point of view. In some cases, the background model can be accurate enough for seismic events to become reasonably well positioned after LWI. However, the accumulation of small inaccuracies in the background model can distort the reflectivity amplitude in some

areas. This situation can be critical when we need such a reflectivity to estimate reservoir properties.

Figure 1.2 portrays a hypothetical (though entirely plausible) situation in reservoir characterization works. In some geologic environments, it makes sense to interpret amplitude anomalies to map and delimit reservoir facies, under the assumption that their constituent rocks exhibit amplitude contrast with respect to rocks of non-reservoir facies. If seismic amplitudes become altered because of various small inaccuracies in the background model, the seismic interpreter may not detect prospect opportunities. Moreover, in some cases, he or she can end up proposing drilling targets based on false amplitude anomalies, with a dry hole as the ultimate consequence.

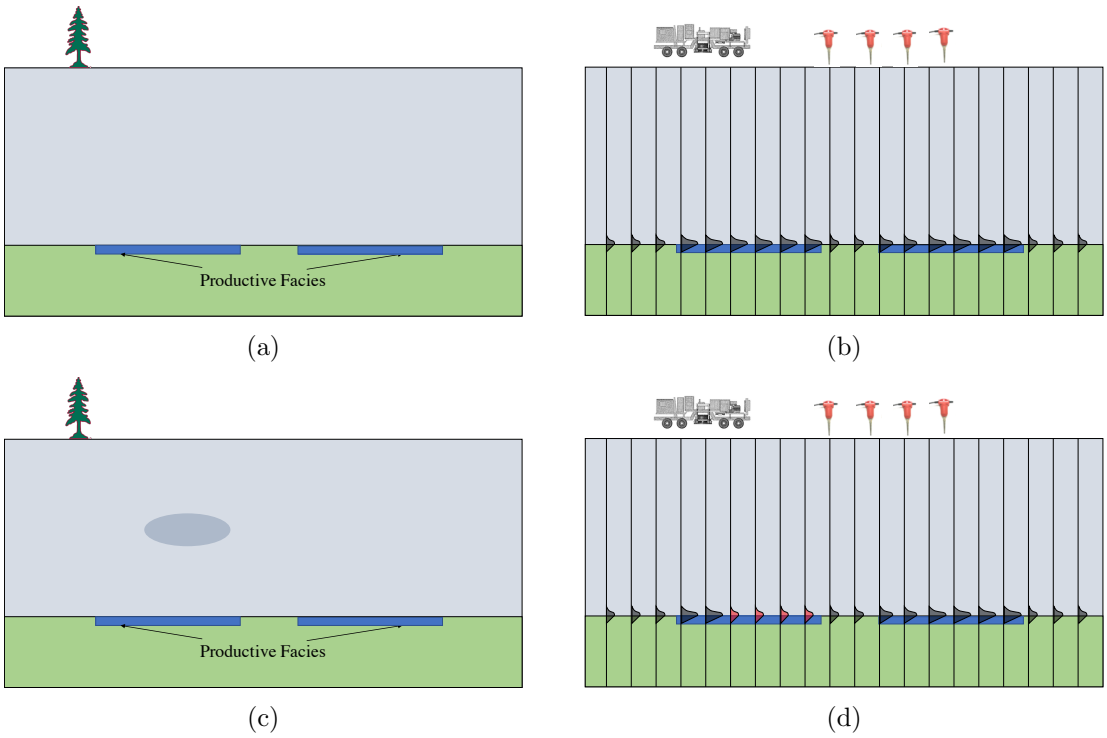


Figure 1.2: Impact of inaccuracies in the background subsurface model on the reflectivity estimation. a) Productive facies in the subsurface. b) After the seismic acquisition, processing, and reflectivity inversion, we can interpret the amplitude anomalies to map the reservoir rocks. c) Same as a), but including an unaccounted anomaly in the background. d) Reflectivity amplitude becomes altered (red traces), leading to incorrect mapping of the reservoir rocks. In more realistic scenarios, amplitudes would be affected by the combined effect of more than one anomaly. [**NR**]

1.2 Proposed solution

In this dissertation, I explore joint inversion of the subsurface reflectivity and background components, $\mathbf{r}$ and $\mathbf{b}$, respectively. Hereafter I will refer to such joint inversion as JIRB, standing for "Joint Inversion for Reflectivity and Background". The idea about incorporating $\mathbf{b}$ is to account for the inaccuracies of the background model during the reflectivity inversion. Hereafter I denominate the inaccurate background model as $\mathbf{b}_0$ so that the accurate background model is given as $\mathbf{b} = \mathbf{b}_0 + \mathbf{\Delta b}$, where $\mathbf{\Delta b}$ is the perturbation in the background. I incorporate the background model component into the inversion by using the WEMVA operator. The migration image, which depends on the background model, gets corrected as the inversion progresses. Simultaneously, the reflectivity is fitted to such an improved migration image using the action of the Gauss-Newton Hessian, similarly as LWI in model space operates. The reflectivity estimated with JIRB is expected to be more reliable than that obtained with conventional LWI using an inaccurate background model, $\mathbf{b}_0$.

1.3 Opportunity area for the new method

One drawback of the JIRB method proposed in this dissertation is the much longer computational time in comparison to conventional reflectivity inversion, mainly because of the addition of the WEMVA operator. In fact, why not performing velocity estimation separately and then going for conventional reflectivity inversion? My answer to this question is that jointly inverting for reflectivity and the background is not intended for large-scale exploration imaging, but detailed works in small areas for reservoir characterization purposes. In reconnaissance and exploration works, seismic interpreters "comb" the subsurface in the quest of oil and gas plays and opportunities. During this stage, drilling prospects are not examined in great detail because they are not yet proved to be productive. On the contrary, once hydrocarbon accumulations have been discovered, careful and detailed interpretation is performed at the

reservoir scale to propose a drilling plan for optimal exploitation. It is often required amplitude-preserving processing focused on the reservoir area. We may use LWI in model space to accomplish this task by pre-computing the Gauss-Newton Hessian and performing the inversion. However, if the background model is discovered not to be accurate enough, we may re-use the existing Gauss-Newton Hessian to run JIRB, rather than re-computing the Hessian with a previously re-estimated velocity.

There is another point of view regarding the implementation of the JIRB method. I hypothesize that the reflectivity estimation would improve if jointly performed with the background model. Conversely, the background model would benefit from the incorporation of the reflectivity in the inversion. For such a reason, in the numerical tests, I also compare the velocity estimation using WEMVA and JIRB.

1.4 Notation convention

Here I formally introduce the notation that I will use in this dissertation. Following Barnier and Almomin (2014), let $\mathbf{m}$ represent the subsurface model parameters, hereafter consisting of slowness squared. Based upon the "scale dichotomy" (Shen and Symes, 2008), let us assume that $\mathbf{m}$ can be split into the contribution of a background model (low-wavenumber component), $\mathbf{b}$, and a reflectivity model (high-wavenumber component), $\mathbf{r}$,

$$\mathbf{m} = \mathbf{b} + \mathbf{r}. \tag{1.3}$$

As aforementioned, we can further split the background component into the superposition of an inaccurate background, $\mathbf{b}_0$ and perturbation effect, $\mathbf{\Delta b}$,

$$\mathbf{b} = \mathbf{b}_0 + \mathbf{\Delta b}. \tag{1.4}$$

Therefore, substituting equation (1.4) into equation (1.3) we obtain

$$\mathbf{m} = \mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b} + \mathbf{r}. \tag{1.5}$$

The above assumption is fundamental in seismic imaging for techniques, including LWI and WEMVA. Hereafter I assume that the observed field data are well represented as Born linearized data, which I denote as $\mathbf{d}_{\text{obs}}$. Using the notation $\mathbf{L}(\mathbf{b})$ to denote the Born linearization operator evaluated at the background model, $\mathbf{b}$, and for the true background model and reflectivity model, $\mathbf{b}_{\text{true}}$ and $\mathbf{r}_{\text{true}}$, respectively, we obtain

$$\mathbf{d}_{\text{obs}} = \mathbf{L}(\mathbf{b}_{\text{true}})\mathbf{r}_{\text{true}}. \tag{1.6}$$

The migration image evaluated at the background model $\mathbf{b}$ will be denoted as $\mathbf{I}(\mathbf{b})$, and it is obtained as the adjoint of the Born linearization operator applied to the observed Born data $\mathbf{d}_{\text{obs}}$, constituting reverse-time migration,

$$\mathbf{I}(\mathbf{b}) = \mathbf{L}(\mathbf{b})^{\mathrm{T}}\mathbf{d}_{\text{obs}}. \tag{1.7}$$

The full-waveform inversion Gauss-Newton Hessian is similarly defined for a background model $\mathbf{b}$ as $\mathbf{H}(\mathbf{b}) = \mathbf{L}(\mathbf{b})^{\mathrm{T}}\mathbf{L}(\mathbf{b})$, while the wave equation migration velocity analysis operator, denoted as $\mathbf{W}(\mathbf{b})$, is defined as the derivative of the migration image with respect to the background model,

$$\mathbf{W}(\mathbf{b}) = \frac{\partial \mathbf{I}(\mathbf{b})}{\partial \mathbf{b}}. \tag{1.8}$$

1.5 Thesis overview

1.5.1 Chapter 2: Theory

In this chapter, I formally derive the JIRB method. I discuss the initial stage, where I had set the problem as a linear optimization process. Then I show that the linearization of the migration image with respect to the background model makes this approach fail. Finally, I show the current nonlinear implementation.

An additional term is needed to drive the background model to solutions that correct and focus the migration image. In the zero subsurface offset domain, we can enforce this requirement by maximizing the energy of the migration image (Tang, 2011a). This method is similar to semblance maximization (Soubaras and Gratacos, 2007) and stack power maximization (Zhang and Shan, 2013). There are two choices in the extended domain: maximization of the stacking power (for angle-domain common image gathers) or minimization of the image after applying differential semblance optimization (for offset-domain common image gathers).

1.5.2 Chapter 3: Random boundary condition

The random boundary condition (RBC) was originally proposed by Clapp (2009) for RTM, to avoid the storage of the whole source wavefield time history. It can also be used for WEMVA to avoid the storage of source and receiver wavefields time histories, enabling the application of the method under limited-memory computational settings, at the cost of additional wavefields propagations. In this chapter, I elaborate on the implementation of the RBC as part of the JIRB constituents, in particular for the WEMVA operator.

1.5.3 Chapter 4: Application to synthetic 2D data

In this chapter, I show the method's implementation in the sedimentary section of the Sigsbee model. I synthesized the data using the acoustic scalar Born modeling operator. I prepared five inversions. The first one is conventional LWI using as input the migration image with the wrong background. The second one is conventional LWI using the migration image with the correct background model. The third one is the JIRB test using the wrong background model. I evaluated the effect of the wrong background model in the estimated reflectivity, compared with the correct result, and to which extent JIRB is capable of estimating the correct background while correcting the reflectivity during the inversion. The fourth and fifth inversions consist of WEMVA and JIRB, respectively, whose background estimations are followed by conventional LWI for comparison purposes.

1.5.4 Chapter 5: 3D Application to 3D ocean-bottom node data from the Gulf of Mexico

In the last chapter, I apply the JIRB method to a 3D ocean-bottom node dataset provided by Shell. I perform the imaging experiments using the downgoing component as input field data, which has been previously obtained using the PZ-summation technique. I employ the well-known technique of mirror imaging. I initially planned on performing two inversions. The first one is conventional LWI using the migration image with the available velocity model. The second one is the JIRB method. The objective is to evaluate the ability to focus residual inaccuracies in the estimated reflectivity. However, the estimated reflectivity using JIRB significantly differs from that of LWI, rather than just constituting a better version of the latter. Therefore, I added numerical tests comparing the background models obtained with WEMVA and JIRB in RTM volumes produced with refined grids to better appreciate changes in the stratigraphy.

Chapter 2

Theory

In this chapter, I present formal derivations of different implementations the JIRB algorithm. I first show the linearization from the most general case and the reason for its failure. Next, I present the original derivation from the FWI objective function. Next, I show the final nonlinear JIRB objective function. Finally, I present some implementation ideas in the extended domain.

2.1 Derivation of JIRB: Most general cases

Let us recall equation (1.2) for conventional LWI in model space:

$$\Phi(\mathbf{r}) = \frac{1}{2} \left\| \mathbf{H}(\mathbf{b})\mathbf{r} - \mathbf{I}(\mathbf{b}) \right\|_2^2 . \tag{2.1}$$

The most general case of JIRB simply consist in the incorporation of $\mathbf{b}$ into the inversion:

$$\Phi(\mathbf{r},\mathbf{b}) = \frac{1}{2}\left\|\mathbf{H}(\mathbf{b})\mathbf{r} - \mathbf{I}(\mathbf{b})\right\|_2^2. \tag{2.2}$$

Note that both the Hessian and the migration image depend on the background model $\mathbf{b}$, which hereafter will correspond to the *accurate* background model that we do not know a priori.

In equation (2.2), the model space becomes expanded, and so it is the model null space (Aster et al., 2013). Therefore, it becomes necessary to incorporate another restriction to drive $\mathbf{b}$ away from spurious solutions. For zero subsurface offset imaging, we can maximize the energy of the migration image to constrain $\mathbf{b}$, obtaining

$$\Phi(\mathbf{r},\mathbf{b}) = \frac{1}{2}\left\|\mathbf{H}(\mathbf{b})\mathbf{r} - \mathbf{I}(\mathbf{b})\right\|_2^2 - \frac{\lambda}{2}\left\|\mathbf{I}(\mathbf{b})\right\|_2^2, \tag{2.3}$$

where λ is a trade-off parameter. The minus sign allows maximizing the second functional within a minimization scheme. The interpretation of equation (2.3) is that the reflectivity fits an updated migration image that gets improved as the background model is corrected. Now I substitute equation (1.4) into (2.3) to obtain

$$\Phi(\mathbf{r},\mathbf{b}_0+\boldsymbol{\Delta}\mathbf{b}) = \frac{1}{2}\left\|\mathbf{H}(\mathbf{b}_0+\boldsymbol{\Delta}\mathbf{b})\mathbf{r} - \mathbf{I}(\mathbf{b}_0+\boldsymbol{\Delta}\mathbf{b})\right\|_2^2 - \frac{\lambda}{2}\left\|\mathbf{I}(\mathbf{b}_0+\boldsymbol{\Delta}\mathbf{b})\right\|_2^2, \tag{2.4}$$

where hereafter $\mathbf{b}_0$ will represent the *inaccurate* background model, which we assume it is available. Now it is the perturbation $\boldsymbol{\Delta}\mathbf{b}$ what becomes a model parameter in the inversion.

2.2 Linearization attempt

Equation (2.4) is nonlinear with respect to $\mathbf{\Delta b}$ in both the Gauss-Newton Hessian and the migration image. We can make use of the following expansions:

$$\mathbf{H}(\mathbf{b}_0 + \mathbf{\Delta b})\mathbf{r} = \mathbf{H}(\mathbf{b}_0)\mathbf{r} + O(\mathbf{\Delta b}, \mathbf{r}), \tag{2.5a}$$

$$\mathbf{I}(\mathbf{b}_0 + \mathbf{\Delta b}) = \mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\mathbf{\Delta b} + O(\|\mathbf{\Delta b}\|^2). \tag{2.5b}$$

For a sufficiently small $\mathbf{\Delta b}$ we can drop the higher order terms, ($O(\mathbf{\Delta b}, \mathbf{r})$ and $O(\|\mathbf{\Delta b}\|)$). After doing that, equation (2.5a) means that the action of the Hessian upon the reflectivity when it is evaluated at the correct background, is not significantly different from the same action of the Hessian, but evaluated at the incorrect background. In this case the dropped higher order terms are second order in $\mathbf{r}$ and $\mathbf{\Delta b}$ simultaneously, which are both small with respect to $\mathbf{b}_0$. To support this claim I performed numerical experiments with the sedimentary section of the Sigsbee 2A model. Figure 2.1 show the true velocity, a perturbation in velocity, and the true reflectity. Figures 2.2, 2.3, and 2.4 compare the terms of expression (2.5a) for Gaussian anomalies similar to that of Figure 2.1(b), but with peak values of -200, -800, and -1600 ft/s for the Gaussian anomaly. Notice that the difference becomes evident only at 1600 ft/s. This value represents an error of approximately 20% with respect to the background average, which is larger than the residual anomalies intended for the either the WEMVA or the JIRB methods.

Expression (2.5b) constitutes the Taylor's series expansion of the migration image around the wrong background model, which we can truncate after the first-order term in $\mathbf{\Delta b}$. $\mathbf{W}$ represents the forward WEMVA operator, which is the derivative of the migration image with respect to the background model. This expression is valid for perturbation values such that $\|\mathbf{\Delta b}\| \ll \|\mathbf{b}_0\|$.

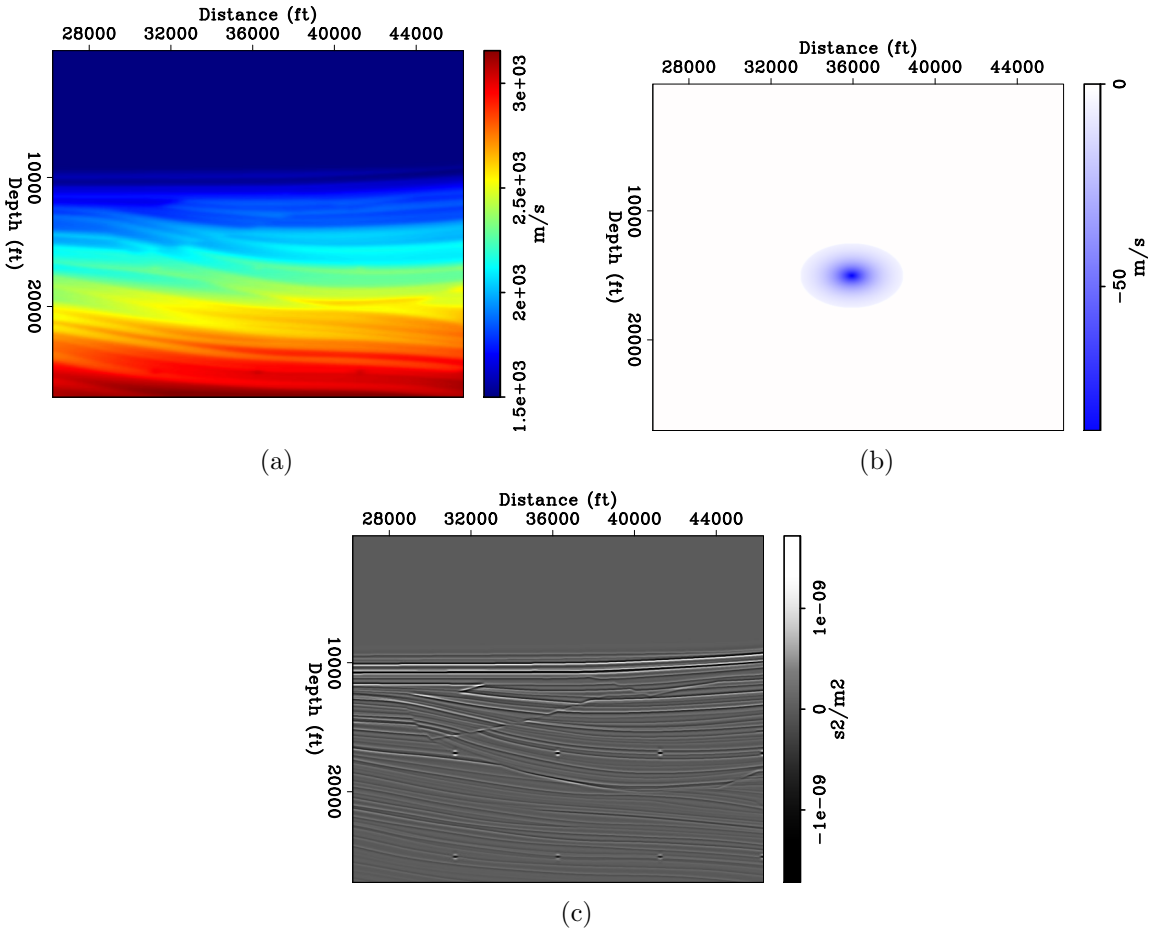


Figure 2.1: Sedimentary section of the Sigsbee model, showing: a) True velocity model. b) Gaussian anomaly. c) True reflectivity. [**CR**]

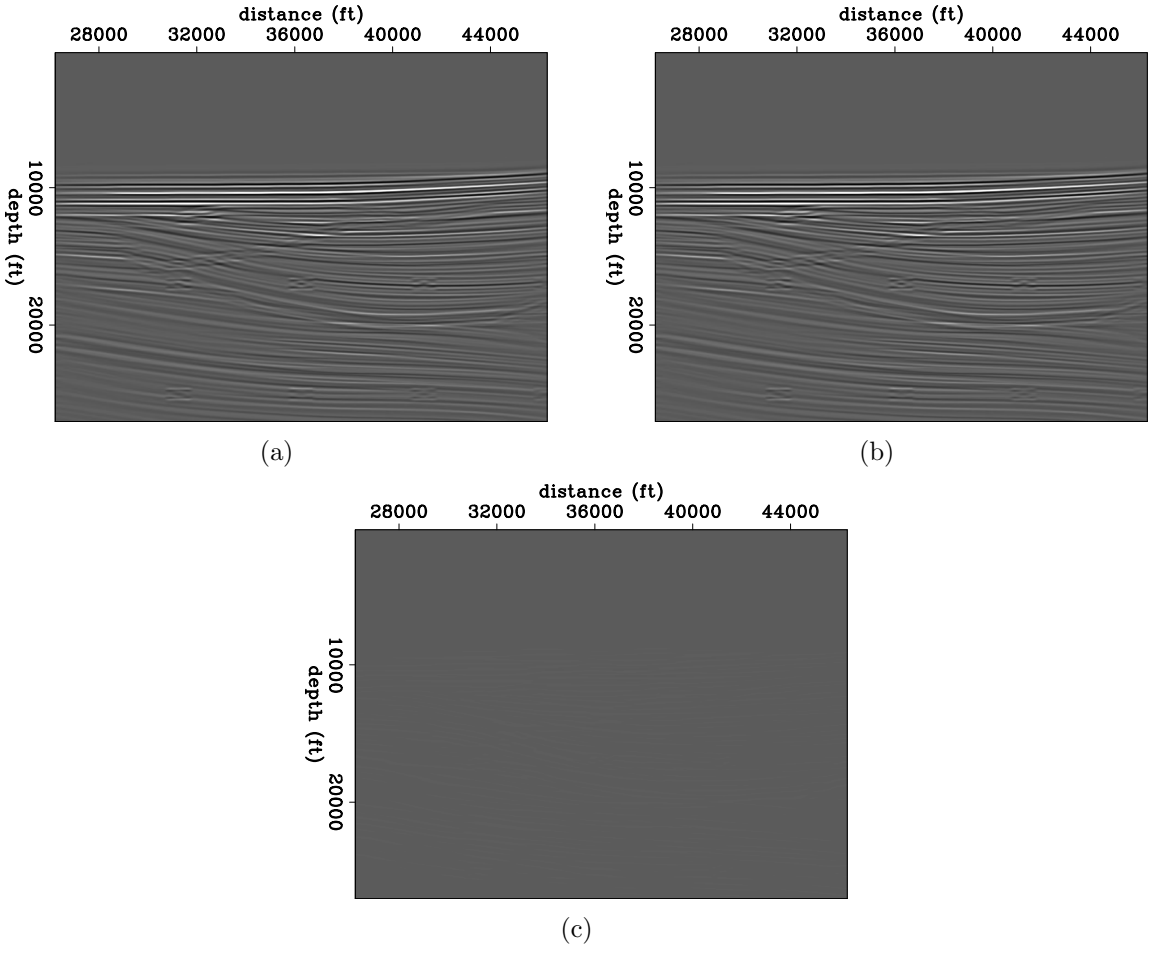


Figure 2.2: Test on the validity of expression (2.5a) for the Sigsbee model sedimentary section, using a Gaussian velocity perturbation of 200 ft/s: a) $\mathbf{H}(\mathbf{b})\mathbf{r}$; b) $\mathbf{H}(\mathbf{b}_0)\mathbf{r}$. c) Difference between images a) and b). [**CR**]

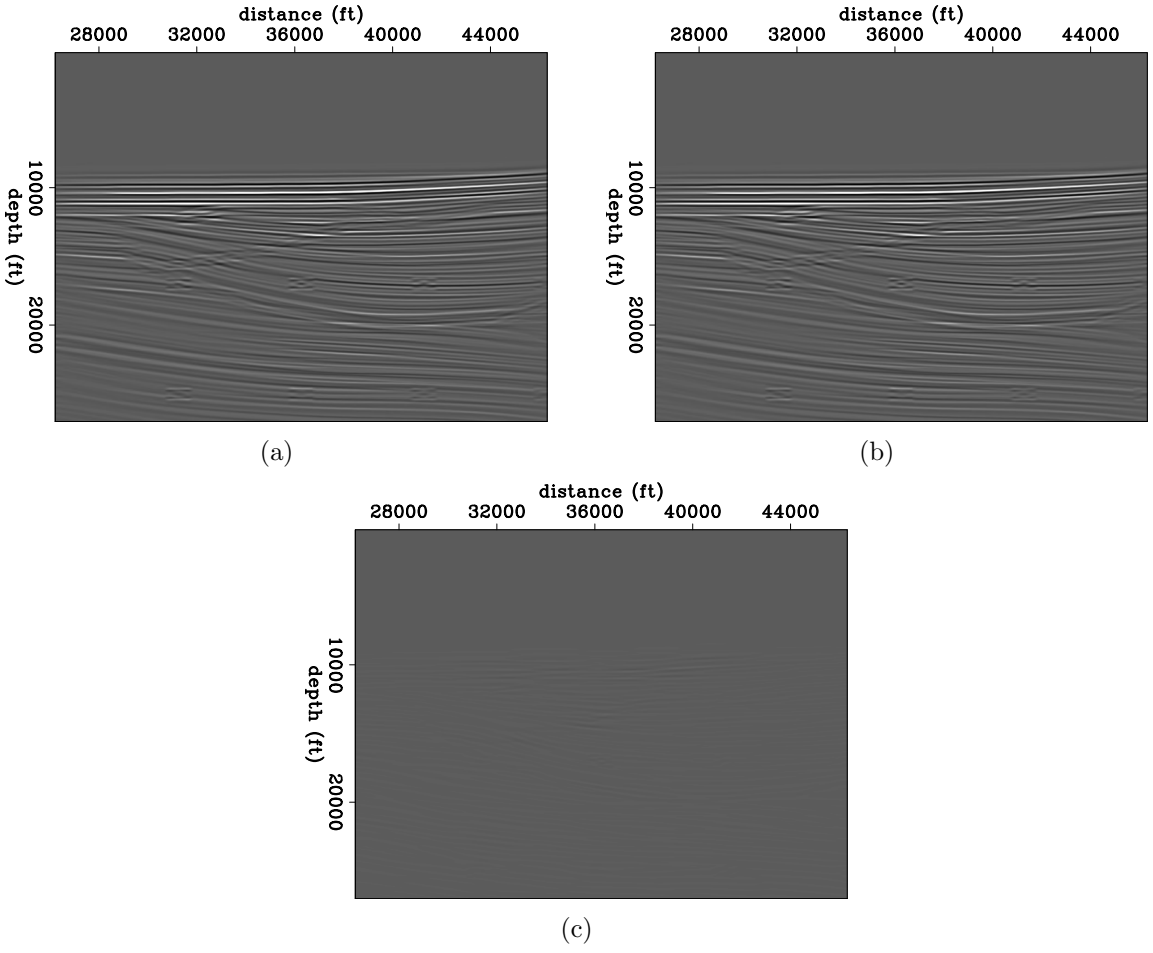


Figure 2.3: Test on the validity of expression (2.5a) for the Sigsbee model sedimentary section, using a Gaussian velocity perturbation of 800 ft/s: a) $\mathbf{H}(\mathbf{b})\mathbf{r}$; b) $\mathbf{H}(\mathbf{b}_0)\mathbf{r}$. c) Difference between images a) and b). [**CR**]

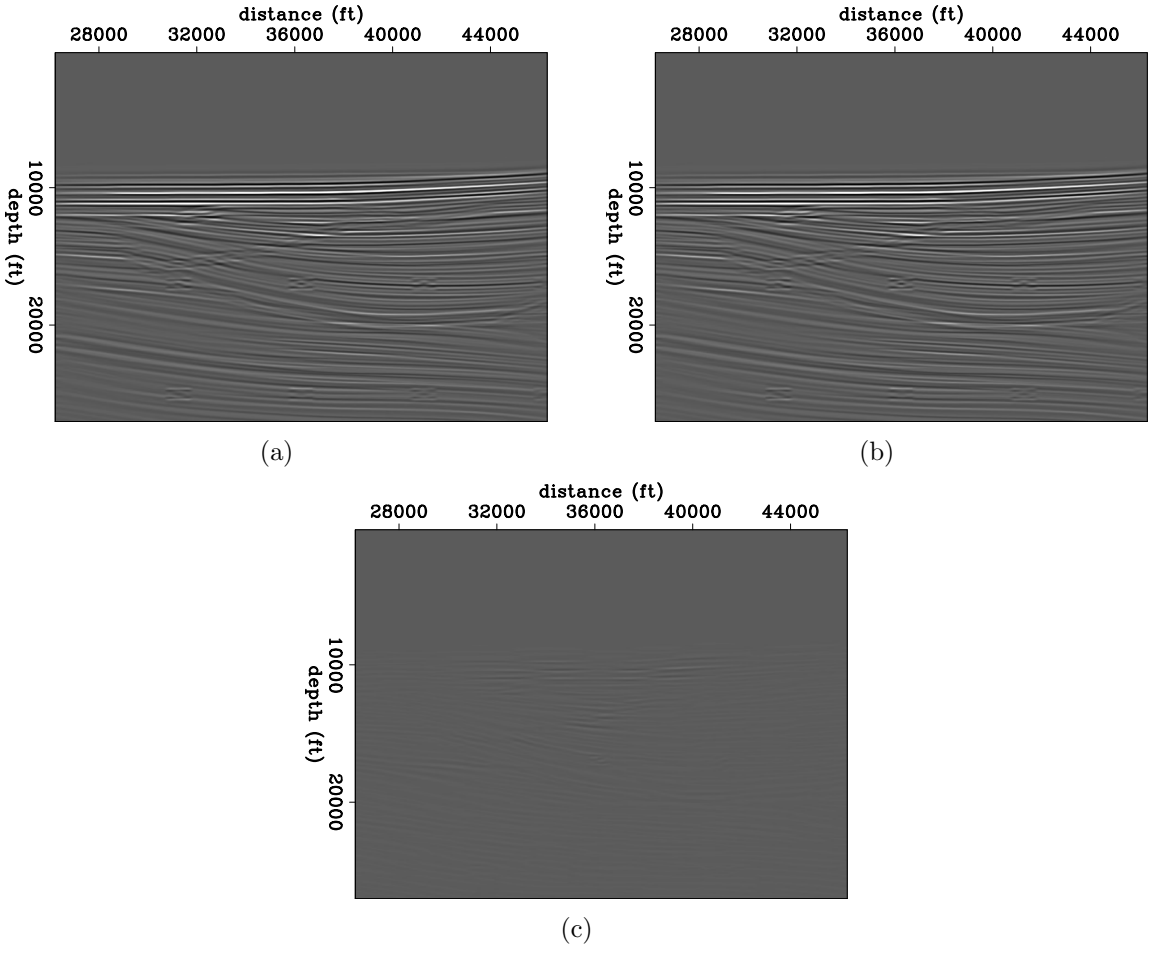


Figure 2.4: Test on the validity of expression (2.5a) for the Sigsbee model sedimentary section, using a Gaussian velocity perturbation of 1600 ft/s: a) $\mathbf{H}(\mathbf{b})\mathbf{r}$; b) $\mathbf{H}(\mathbf{b}_0)\mathbf{r}$. c) Difference between images a) and b). [**ER**]

Substituting equations (2.5a) and (2.5b) into equation (2.4) yields

$$\Phi(\mathbf{r}, \boldsymbol{\Delta}\mathbf{b}) = \frac{1}{2} \left\| \mathbf{H}(\mathbf{b}_0)\mathbf{r} - \mathbf{I}(\mathbf{b}_0) - \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} \right\|_2^2 - \frac{\lambda}{2} \left\| \mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} \right\|_2^2 . \quad (2.6)$$

The first functional in equation (2.6) has the expected behavior of minimizing the difference between the reflectivity under the action of the Hessian, and the migration image evaluated at the wrong background, corrected with the WEMVA term applied to the perturbation in the background. This term becomes a perturbation in the image that contains the corrections to the background model.

However, the second functional in equation (2.6) does not have the expected behavior. The reason is because, whereas the linearization in expression (2.5b) is valid for small $\boldsymbol{\Delta}\mathbf{b}$ values, substituting such a linearized term into the norm gives

$$\left\| \mathbf{I}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b}) \right\|_2^2 \approx \left\| \mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} \right\|_2^2 , \quad (2.7)$$

which is inaccurate. In fact, expression (2.7) preserves a second order term in $\boldsymbol{\Delta}\mathbf{b}$, but neglects another (see Appendix A for details). The consequence is that the linearized norm in the right-hand side of expression (2.7) behaves as a quadratic function of the type $Ax^2 + Bx + C$, with curvature term $A \geq 0$ because it is represented by a positive semidefinite matrix (Nocedal and Wright, 2006; Aster et al., 2013). In other words, the linearized norm is *convex*, thus there is no maximum at the true background, as it is expected from the original functional (left-hand side of expression (2.7)), which is *concave*.

To further illustrate the problem, I performed numerical tests on a flat-layer model and the sedimentary sector of the Sigsbee model (Figure 2.5). The flat-layer model includes a positive Gaussian anomaly in velocity. This model represents the true background model, $\mathbf{b}^{\text{true}} = \mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b}^{\text{true}}$. The incorrect background model $\mathbf{b}_0$ does

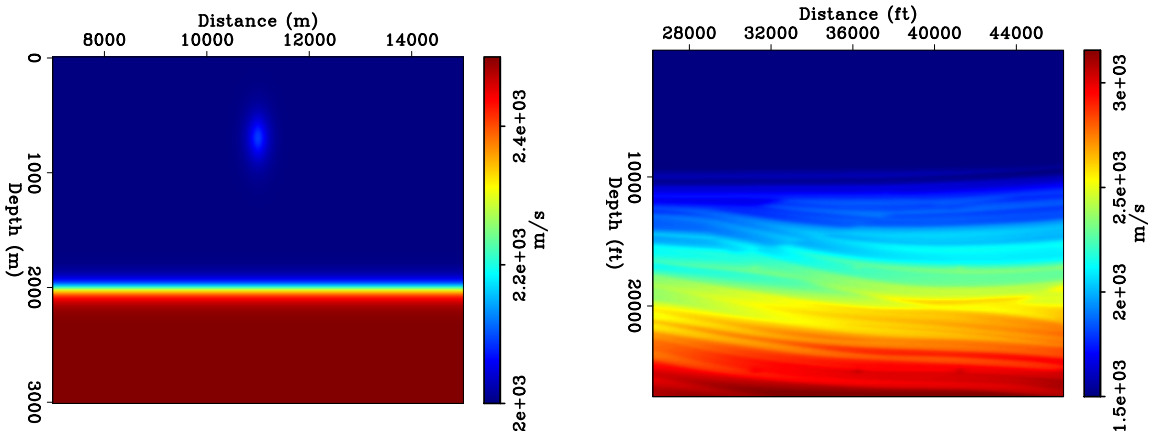
not contain the anomaly. Hence, this anomaly represents the true perturbation in the background, $\mathbf{\Delta b}^{\text{true}}$. On the other hand, the true Sigsbee background model does not contain a Gaussian anomaly. However, to construct an incorrect background model, I added a negative Gaussian anomaly in velocity to the true background model. The Gaussian anomalies for both models are shown in Figure 2.7.

I synthesized true data $\mathbf{d}^{\text{true}}$ in both cases by applying Born modeling to the corresponding reflectivity models (Figure 2.8) using the true background models. For the flat-layer model, the acquisition geometry consists of 101 shots regularly spaced every 140 m, and receivers every 20 m. For the Sigsbee model, the acquisition geometry consists of 54 shots spaced every 500 ft, and receivers every 75 ft.

The first test demonstrates that the image power maximization term in the linearized objective function (right-hand side of expression (2.5b)) does not have the desired property of concavity. To do this, I created trial perturbations in the background model, given as $\mathbf{\Delta b}^{\text{true}} + \alpha \mathbf{\Delta b}_{\text{rand}}$, where $\mathbf{\Delta b}^{\text{true}}$ is the true anomaly, $\mathbf{\Delta b}_{\text{rand}}$ represents randomly distributed perturbations whose amplitudes are confined within the interval $\big[-\max(\mathbf{\Delta b}^{\text{true}}), \max(\mathbf{\Delta b}^{\text{true}})\big]$, and α is a scalar parameter with values $-1, -0.9, ..., 0.9, 1$. Then I evaluated the functional, $\big\|\mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\big[\mathbf{\Delta b}^{\text{true}} + \alpha \mathbf{\Delta b}_{\text{rand}}\big]\big\|_2^2$, using different random realizations of $\mathbf{\Delta b}_{\text{rand}}$, and exploring along each direction with the α values, so $\alpha = 0$ corresponds to the evaluation of the objective function using the true background anomaly.

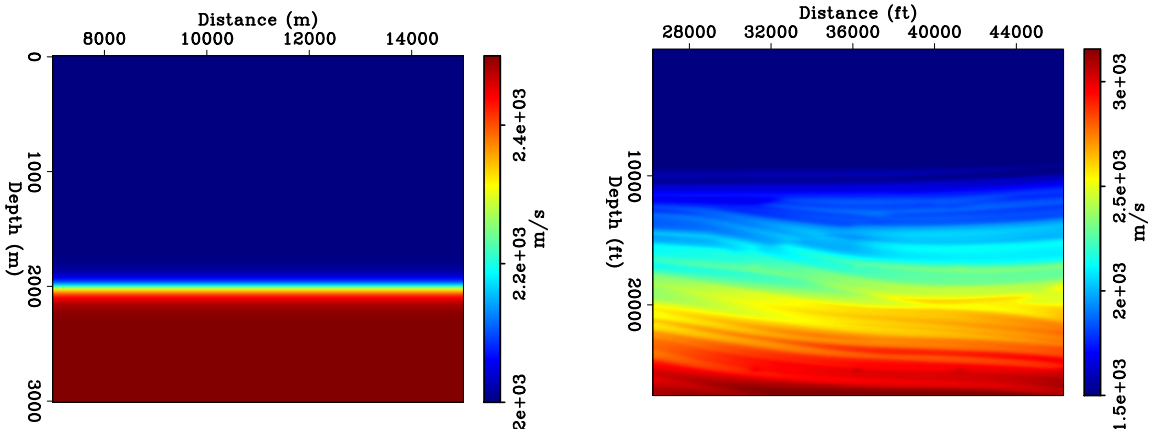
Figure 2.9 shows the result of 50 tests of the evaluation of the objective function using random directions, $\mathbf{\Delta b}_{\text{rand}}$, each one for the 21 values of α. Notice that there is no upper bound, and although the minima of the curves cluster around the vicinity of the true anomaly, none of them corresponds to it. These results prove that this objective function is convex, as I discussed before.

I performed similar tests for the fully nonlinear objective function of image power maximization (left-hand side of expression (2.5b)). Similar to the previous test, I evaluated this function at random directions. The results for both models are shown



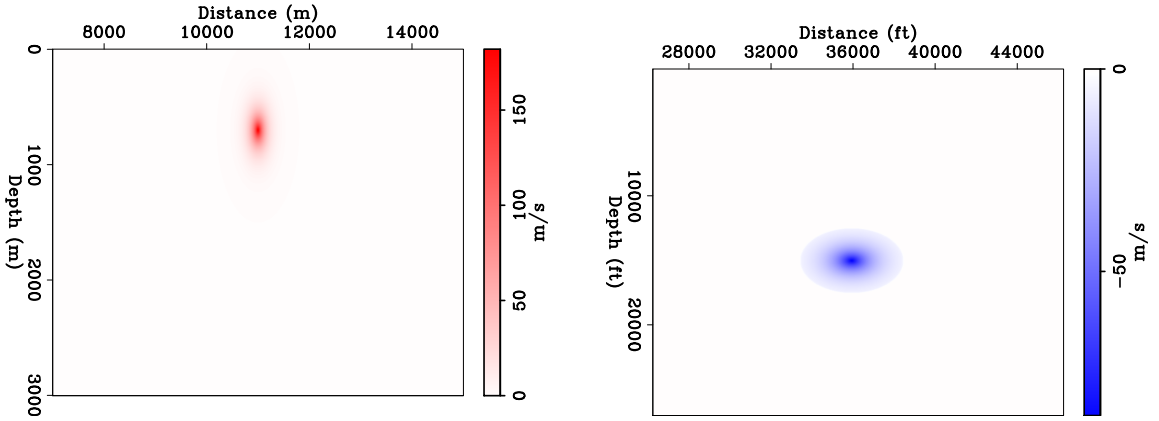
(a) (b)

Figure 2.5: True background models in velocity. a) Flat model. b) Sigsbee model. [**ER**]



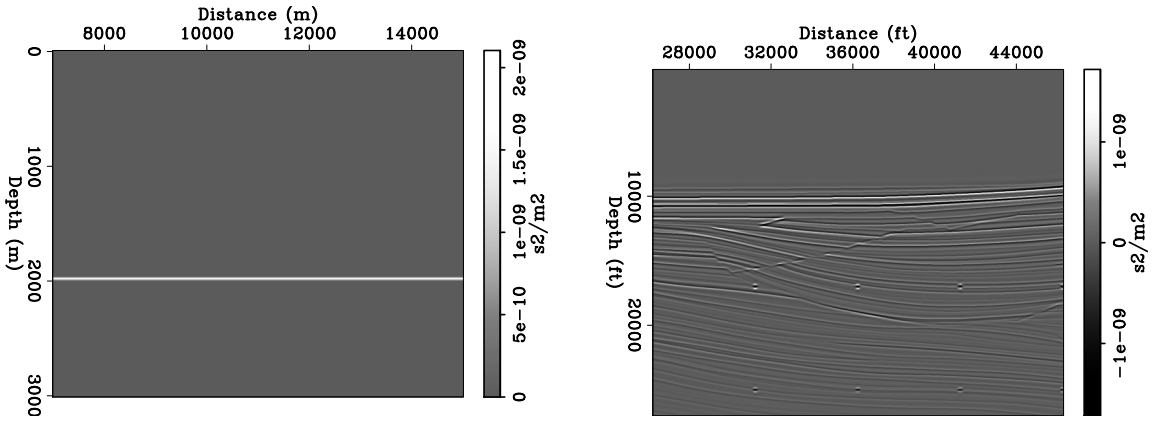
(a) (b)

Figure 2.6: Wrong background models in velocity. a) Flat model. b) Sigsbee model. [**ER**]



(a) (b)

Figure 2.7: Perturbations in the background models expresed in velocity. a) for the flat model. b) For the Sigsbee model. [**ER**]



(a) (b)

Figure 2.8: Reflectivity models in slowness squared. a) Flat model. b) Sigsbee model. [**ER**]

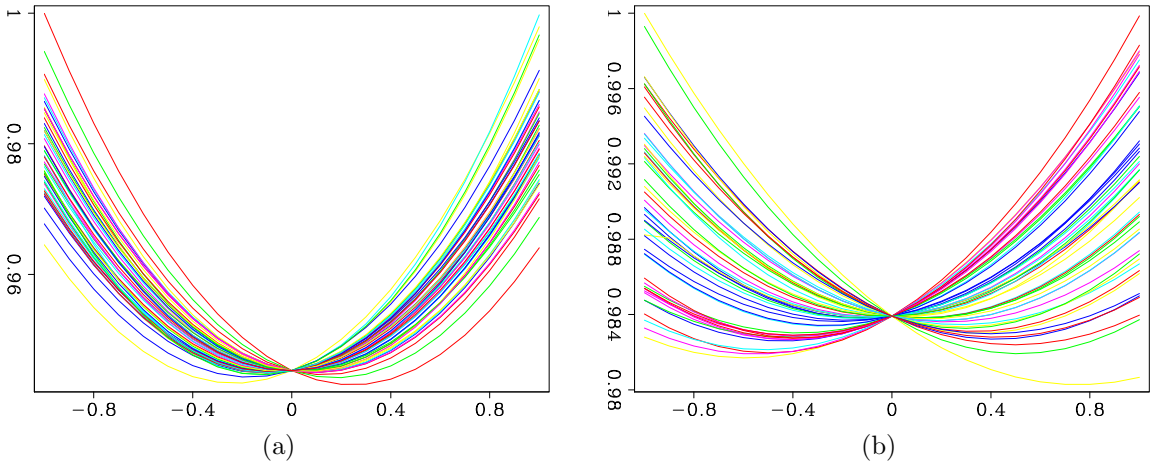


Figure 2.9: Experiment results of 50 random directions of the $\|\mathbf{I}(\mathbf{b}_0)+\mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}^{\text{true}}+\alpha\boldsymbol{\Delta}\mathbf{b}_{\text{rand}}\|_2^2$ functional in the a) flat model, and the b) Sigsbee model. Each random direction was explored with the α scalar. [**NR**]

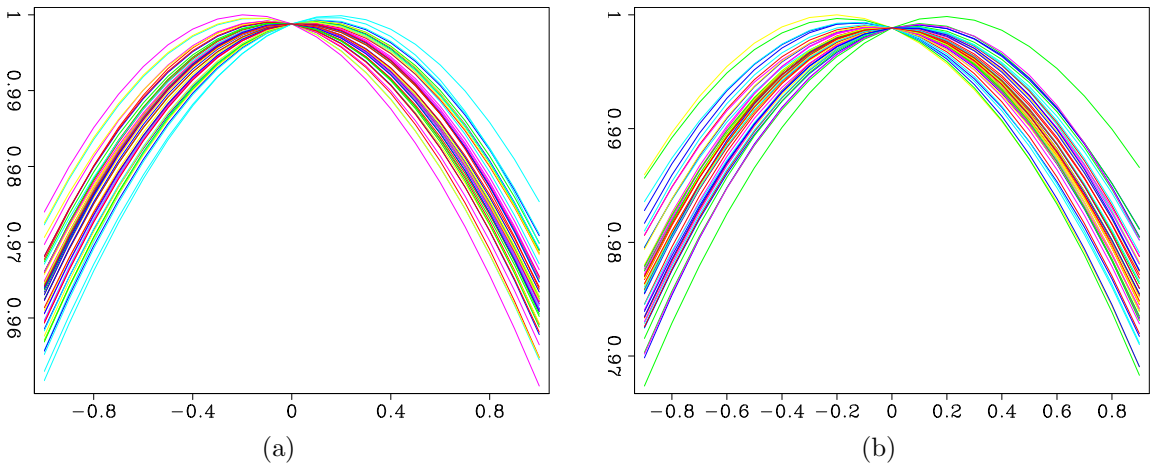


Figure 2.10: Experiment results of 50 random directions of the $\|\mathbf{I}(\mathbf{b}^{\text{true}}+\alpha\boldsymbol{\Delta}\mathbf{b}_{\text{rand}})\|_2^2$ functional in the a) flat model, and the b) Sigsbee model. Each random direction was explored with the α scalar. [**NR**]

in Figure 2.10. These results confirm that this objective function is concave along the directions explored. While these experiments do not tell us for sure if it is, in fact, concave, we can still conclude that this may be an approximate heuristic. However, even though the maximum values of the curves cluster around $\alpha = 0$, not all of them are exactly there.

2.3 Derivation from full-waveform inversion

The first attempt to envision a JIRB method was through the simplification of the Newton's equation associated to the FWI iterative solution (e.g. Cabrales-Vargas et al., 2016; Cabrales-Vargas, 2018).

Full-waveform inversion (Tarantola, 1984a; Pratt et al., 1998; Virieux and Operto, 2009) is a nonlinear optimization technique that, in principle, can estimate all the wavenumber components of the subsurface model parameters, $\mathbf{m}$. However, in practice, it suffers from cycle skipping if the difference between the initial model and the actual model is more than half a wavelength in the data space. Multiscaling approaches are often employed to circumvent the local minima problem (e.g. Virieux and Operto, 2009).

In a nutshell, FWI aims to minimize in the least-squares sense the misfit between the observed seismic data and synthetic data corresponding to model $\mathbf{m}$, which we can express as

$$\Phi(\mathbf{m}) = \frac{1}{2}\|\mathbf{d}(\mathbf{m}) - \mathbf{d}_{\text{obs}}\|_2^2, \tag{2.8}$$

where

$$\mathbf{d}(\mathbf{m}) = \mathbf{S}\big[\mathcal{L}(\mathbf{m})\big] \tag{2.9}$$

is the modeled data calculated by using the full-wave propagation operator, $\mathcal{L}$, collected at the receivers' position by the sampling operator, $\mathbf{S}$.

The solution of the FWI optimization problem in equation (2.8) using gradient descent optimization schemes or Hessian based methods (Nocedal and Wright, 2006) requires the computation of the gradient of the misfit function, which is given by

$$\nabla_{\mathbf{m}}\Phi(\mathbf{m}) = \big[\nabla_{\mathbf{m}}\mathcal{L}(\mathbf{m})\big]^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\big[\mathbf{S}\mathcal{L}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}\big] = \mathbf{L}(\mathbf{m})^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\big[\mathbf{S}\mathcal{L}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}\big], \quad (2.10)$$

where $\nabla_{\mathbf{m}}$ constitutes the gradient with respect to the subsurface model parameters, $\mathbf{S}^{\mathrm{T}}$ is the adjoint operator of sampling (hence it injects data at the receivers' positions), and $\mathbf{S}\mathcal{L}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}$ represents the data residuals that are projected into the model space by means of the RTM operator, $\mathbf{L}(\mathbf{m})^{\mathrm{T}} = \big[\nabla_{\mathbf{m}}\mathcal{L}(\mathbf{m})\big]^{\mathrm{T}}$.

The FWI solution can be obtained by local linearization where the current subsurface model, $\mathbf{m}_{\mathrm{i}}$, is updated by a direction vector $\delta\mathbf{m}$ such that $\|\delta\mathbf{m}\| << \|\mathbf{m}_{\mathrm{i}}\|$, thereby exhibiting a local quadratic behavior. Now we can expand the gradient in Taylor's series around $\mathbf{m}_{\mathrm{i}}$ and drop second- and higher-order terms, obtaining

$$\nabla_{\mathbf{m}}\Phi(\mathbf{m}_{\mathrm{i}} + \delta\mathbf{m}) \approx \nabla_{\mathbf{m}}\Phi(\mathbf{m}_{\mathrm{i}}) + \nabla_{\mathbf{m}}\nabla_{\mathbf{m}}\Phi(\mathbf{m}_{\mathrm{i}})\delta\mathbf{m} = \nabla_{\mathbf{m}}\Phi(\mathbf{m}_{\mathrm{i}}) + \mathbf{H}_{\mathrm{fwi}}(\mathbf{m}_{\mathrm{i}})\delta\mathbf{m}, \quad (2.11)$$

where $\mathbf{H}_{\mathrm{fwi}}(\mathbf{m}_{\mathrm{i}})$ is the FWI Hessian, not be confused with the FWI *Gauss-Newton* Hessian, $\mathbf{H} = \mathbf{L}^{\mathrm{T}}\mathbf{L}$, introduced in a previous section. As mentioned before, solving equation (2.11) represents the optimization of a quadratic problem in $\delta\mathbf{m}$. Therefore, the gradient evaluated at the updated model, $\mathbf{m}_{\mathrm{i}+1} = \mathbf{m}_{\mathrm{i}} + \delta\mathbf{m}$, becomes zero, and equation (2.11) reduces to

$$\mathbf{H}_{\text{fwi}}(\mathbf{m}_{\text{i}})\delta\mathbf{m} = -\nabla_{\mathbf{m}}\Phi(\mathbf{m}_{\text{i}}). \tag{2.12}$$

Equation (2.12) represents the well-known Newton's equation (Nocedal and Wright, 2006).

Now I will incorporate equation (1.5) into Newton's equation by defining the current model and the model step as follows:

$$\begin{aligned} \mathbf{m}_{\text{i}} &= \mathbf{b}_0, \\ \delta\mathbf{m} &= \boldsymbol{\Delta}\mathbf{b} + \mathbf{r}. \end{aligned} \tag{2.13}$$

On the other hand, according to Biondi et al. (2015) we can split the FWI Hessian into the FWI Gauss-Newton Hessian and the adjoint of the WEMVA operator:

$$\mathbf{H}_{\text{fwi}} = \mathbf{H} + \mathbf{W}^{\text{T}}. \tag{2.14}$$

Substituting equations (2.13) and (2.14) into the Newton's equation yields

$$\mathbf{H}_{\text{fwi}}(\mathbf{m}_{\text{i}})\delta\mathbf{m} = \big[\mathbf{H}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\big](\boldsymbol{\Delta}\mathbf{b} + \mathbf{r}). \tag{2.15}$$

As far as I am assuming the remaining inaccuracies in the background model to be small, it is valid to employ maximization of the image energy. Therefore, I dropped the transpose symbol in $\mathbf{W}$ because the WEMVA operator is self-adjoint in the zero-subsurface offset domain.

Re-arranging terms in the right-hand side of equation (2.15) we obtain

$$\big[\mathbf{H}(\mathbf{b}_0)\mathbf{r} + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}\big] + \big[\mathbf{H}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} + \mathbf{W}(\mathbf{b}_0)\mathbf{r})\big]. \tag{2.16}$$

The first part inside brackets in expression (2.16) contains the image term $\mathbf{H}(\mathbf{b}_0)\mathbf{r}$ employed in model-space LWI, and the term $\mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}$ that represents the perturbation in the image regarding the correction of the background model. I keep both terms because they have a predictable impact on primary events related to the subsurface reflectivity. The second part in brackets very likely contributes to second-order scattering events $\big[\mathbf{W}(\mathbf{b}_0)\mathbf{r}\big]$ and refinement of the low-wavenumber component $\big[\mathbf{H}\boldsymbol{\Delta}\mathbf{b}\big]$. I drop the first term because it is not related to the subsurface reflectivity's primary events, and I drop the second one because the method does not aim to retrieve a high-wavenumber background model.

Regarding the FWI gradient in the Newton's equation (right-hand side of equation (2.12)), it becomes $\nabla\Phi(\mathbf{b}_0)$. Substituting this gradient in equation (2.10) yields

$$\nabla\Phi(\mathbf{b}_0) = \big[\nabla\mathcal{L}(\mathbf{b}_0)\big]^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\big[\mathbf{S}\mathcal{L}(\mathbf{b}_0) - \mathbf{d}_{\mathrm{obs}}\big]. \tag{2.17}$$

Nonlinear wave propagation at the background model (free of reflections) will model only direct waves and diving waves. Therefore, the term $\mathbf{S}^{\mathrm{T}}\big[\mathbf{S}\mathcal{L}(\mathbf{b}_0) - \mathbf{d}_{\mathrm{obs}}\big]$ simply represents the negative of the perturbation in the data, $\boldsymbol{\Delta}\mathbf{d}_{\mathrm{obs}}$, (Barnier and Almomin, 2014), i.e., the data after removing the direct and diving waves. On the other hand, the adjoint of the gradient term represents Born adjoint modeling or RTM evaluated at $\mathbf{b}_0$. Thus, equation (2.17) becomes the negative of the RTM image,

$$\nabla\Phi(\mathbf{b}_0) = -\mathbf{L}(\mathbf{b}_0)^{\mathrm{T}}\boldsymbol{\Delta}\mathbf{d}_{\mathrm{obs}} = -\mathbf{I}(\mathbf{b}_0). \tag{2.18}$$

Substituting the first term in braces in expression (2.16) and equation (2.18) in the Newton's equation we obtain

$$\mathbf{H}(\mathbf{b}_0)\mathbf{r} + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} = \mathbf{I}(\mathbf{b}_0), \tag{2.19}$$

which I named Linearized Waveform Inversion with Velocity Updating (LWIVU), and is equivalent to minimizing the objective function

$$\Phi(\mathbf{r}, \boldsymbol{\Delta}\mathbf{b}) = \frac{1}{2} \left\| \mathbf{H}(\mathbf{b}_0)\mathbf{r} - \mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} \right\|_2^2 . \tag{2.20}$$

Note that this objective function represents the first functional in equation (2.6). The only difference is in the positive sign for $\mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}$, which is not a problem because the perturbation in the background can take positive or negative values.

2.4 Nonlinear scheme

Given the shortcoming of the objective function (2.6), Cabrales-Vargas and Sarkar (2019) proposed to keep the nonlinear objective function for the migration image, still maintaining the assumption $||\boldsymbol{\Delta}\mathbf{b}||_2 << ||\mathbf{b}_0||_2$, so that the approximation $\mathbf{H}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b}) \approx \mathbf{H}(\mathbf{b}_0)$ is accurate. Therefore, the JIRB objective function became

$$\Phi(\mathbf{r}, \boldsymbol{\Delta}\mathbf{b}) = \left\| \mathbf{H}(\mathbf{b}_0)\mathbf{r} - \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} - \mathbf{I}(\mathbf{b}_0) \right\|_2^2 - \lambda \left\| \mathbf{I}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b}) \right\|_2^2 . \tag{2.21}$$

This change makes the first term a linear least-squares problem over both $\boldsymbol{\Delta}\mathbf{b}$ and $\mathbf{r}$, while the second term is nonlinear and non-quadratic.

Although the objective function (2.21) behaves as desired, it is more straightforward and sensible to implement the non-linearity in both objective functions:

$$\Phi(\mathbf{r}, \mathbf{b}) = \left\| \mathbf{H}(\mathbf{b}_0)\mathbf{r} - \mathbf{I}(\mathbf{b}) \right\|_2^2 - \lambda \left\| \mathbf{I}(\mathbf{b}) \right\|_2^2 . \tag{2.22}$$

Objective function (2.22) can be solved (for local minima) using gradient-based methods such as steepest descent, nonlinear conjugate gradients, or quasi-Newton (Nocedal

and Wright, 2006). Notice that the Gauss-Newton Hessian is precomputed with the wrong background model, and does not require to be updated as it is the case for the migration image. Solving objective function (2.22) is the core of the JIRB method. Chapter 4 shows a variation where I merely add preconditioning and weighting factors.

2.5 Extended domain

The joint inversion of reflectivity and perturbation in the background can also be cast in the extended domain as

$$\Phi(\tilde{\mathbf{r}}, \mathbf{\Delta b}) = \frac{1}{2}\left\|\tilde{\mathbf{H}}(\mathbf{b}_0 + \mathbf{\Delta b})\tilde{\mathbf{r}} - \tilde{\mathbf{I}}(\mathbf{b}_0 + \mathbf{\Delta b})\right\|_2^2 + \frac{\lambda}{2}\left\|\mathbf{D}\tilde{\mathbf{I}}(\mathbf{b}_0 + \mathbf{\Delta b})\right\|_2^2, \tag{2.23}$$

where the tilde represents extension in subsurface offset and $\mathbf{D}$ represents the differential semblance optimization operator (DSO) (Symes and Carazzone, 1991). In this case, the second functional ought to be minimized to penalize non-zero subsurface offset energy originated from incorrect background model.

At first glance it appears that in the expanded case the second functional can be linearized by substituting $\mathbf{D}\tilde{\mathbf{I}}(\mathbf{b}_0+\mathbf{\Delta b}) \approx \mathbf{D}\left[\tilde{\mathbf{I}}(\mathbf{b}_0)+\tilde{\mathbf{W}}(\mathbf{b}_0)\mathbf{\Delta b}\right]$ in the norm, because both the nonlinear and the linearized functionals are convex, in contrast to the non-extended case where we maximize the image power, analyzed before. However, this approximation still lacks one second-order term that can lead to incorrect results (see Appendix A).

We can use a similar objective function as that of the zero subsurface offset domain:

$$\Phi(\tilde{\mathbf{r}}, \mathbf{b}) = \frac{1}{2}\left\|\tilde{\mathbf{H}}(\mathbf{b}_0)\tilde{\mathbf{r}} - \tilde{\mathbf{I}}(\mathbf{b})\right\|_2^2 + \frac{\lambda}{2}\left\|\mathbf{D}\tilde{\mathbf{I}}(\mathbf{b}_0 + \mathbf{\Delta b})\right\|_2^2. \tag{2.24}$$

The obvious caveat of equation (2.24) is that inversion in the extended domain is more costly than the inversion in the zero offset domain because the scattering operator and the imaging condition operator dominate the computations. This problem is particularly severe in 3D imaging. Some people have proposed strategies to alleviate the problem, such as the so-called "Nyquist approach"(Sun and Fu, 2013), where we reduce to the minimum the number of times that the imaging condition is applied during RTM, depending on the frequency range of the data. However, in order to achieve enough extension, the computational burden is still considerable. The DSO operator often requires special preconditioning to avoid high-wavenumber artifacts in the gradient (e.g. Shen and Symes, 2008; Tang, 2011b; Weibull and Arntsen, 2013). For such reasons, in this dissertation I implement the JIRB methodology exclusively in the zero subsurface offset domain.

Chapter 3

Random Boundary Condition

3.1 Introduction

The RTM implementation using the random boundary condition (RBC) was proposed by Clapp (2009) as an alternative to checkpointing (e.g. Symes, 2007), to avoid the storage of the source wavefield propagation history at the cost of an additional time-reverse propagation to re-generate it. Shen and Clapp (2011) further adapted the technique to cope with low-wavenumber limitations. This improvement is particularly beneficial to the proper computation of the FWI gradient (Shen and Clapp, 2015). More recently, Clapp and Alves (2016) adapted the method to elastic RTM.

Besides this "traditional" implementation of the RBC in RTM, I also adapted them for the WEMVA operator, which is one of the computationally expensive operators needed to implement JIRB. To my knowledge, RBC had not been used before in WEMVA. For this reason, this chapter is dedicated to explain how it works.

3.2 Brief review of the random boundary condition on RTM

When imaging using RTM, we perform the propagation of the wavefields in opposite time directions, i.e., the source wavefield's history is constructed by forward-time propagation of the injected source wavefield. In contrast, we construct the receiver wavefield by reverse-time propagation of the injected field data. For such reason, one of the wavefields (typically the source wavefield) remains either in RAM or physical memory during the propagation of the other wavefield, of which merely two history frames are kept in memory to perform the correlation imaging condition (Claerbout, 1971) on the fly. Even if storing only one of the wavefields, we still need to account for thousands of time frames that in 3D represent seismic volumes of considerable size, not to talk about the I/O latency that comes to the detriment of computational performance.

Different strategies can be implemented to reduce storage. Symes (2007) proposed the use of checkpoints. This strategy consists on saving a determined number of time frames for the re-computation of the wavefields from optimally chosen times. This method works well in practice, but it still requires a non-trivial selection of the optimum checkpoints (Symes, 2007).

Another strategy proposed by Clapp (2008) is saving the boundaries of the source wavefield in a rectangular halo beyond the imaging space, and re-injecting them to perform a reverse-time propagation of the source wavefield as performing the usual reverse-time receiver wavefield propagation, while multiplying corresponding time frames to build the image. This strategy allows a re-computation similar to checkpoints, with better applicability for large volumes (Clapp, 2008).

Both methods - checkpointing and saving the boundaries - require I/O work to some extent. For such reason, Clapp (2009) proposed the use of the RBC to recreate the source wavefield in reverse-time propagation. This operation only requires keeping

two time frames in RAM for both the source wavefield and the receiver wavefield, thereby constituting a tremendous saving in storage. Even though the RBC requires the re-computation of the source wavefield, it avoids excessive I/O operations.

Figure 3.1 schematically shows the RBC implementation in comparison to the conventional RTM computation.

The upper part of Figure 3.1 illustrates the stages of conventional RTM implementation were the source wavefield is forward-time propagated and its whole history of n_t samples (orange frames) is stored (a). Next, the receiver wavefield (blue frames) is reverse-time propagated (b-g) and cross-correlated with the receiver wavefield on the fly. Note that only two temporal frames of the receiver wavefield need be kept in memory. This strategy works well for small 2D datasets. However, for large 2D datasets or 3D datasets, the source wavefield needs to be stored in disk, and their corresponding time frames have to be accessed as needed during the cross-correlation.

The lower part of Figure 3.1 illustrates the stages of RTM using the RBC. The source wavefield is forward-time propagated as before, but only two temporal frames of its history are kept in memory (a). Then it trespasses the limit reaching sample $n_t + 1$. Next, the source wavefield is reverse-time propagated simultaneously with the receiver wavefield and cross-correlated on the fly (b-g). Note that only a total of four temporal frames are kept in memory, thereby wavefield storage is no longer needed.

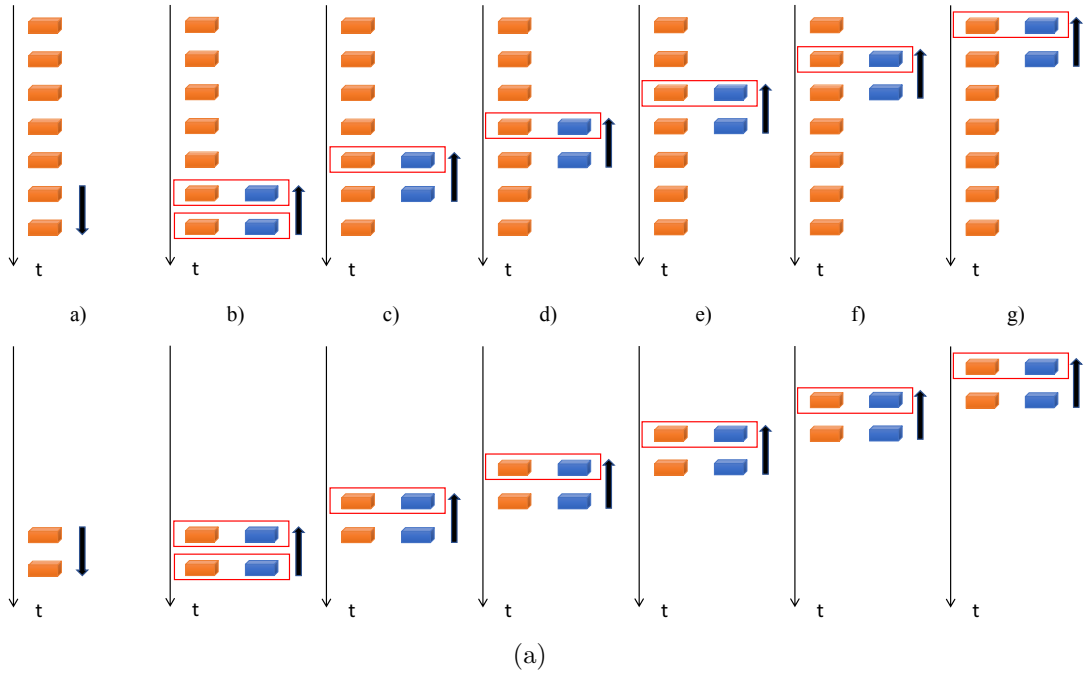


Figure 3.1: Comparison of conventional RTM implementation (top) and using the random boundary condition (bottom). [**NR**]

3.3 Application of the random boundary condition to WEMVA

The WEMVA operator has two components, namely the source-side component and the receiver-side component (Almomin and Tang, 2010; Tang, 2011a). Each component consists of two operations. For the source-side component, one propagates forward in time the source wavefield, then scatters this wavefield against the perturbation in the background model (forward operator) or the perturbation in the image (adjoint operator) and propagates the scattered wavefield. Finally, one propagates backward in time the receiver wavefield and cross-correlates it with the scattered source wavefield. We can use a similar argument for the computation of the receiver-side WEMVA component.

Figure 3.2 illustrates the different stages of the forward WEMVA operator. The WEMVA scattering stage is similar to Born scattering, whereas the WEMVA cross-correlation of the background wavefields with the scattered wavefields is similar to the RTM imaging condition. It is during the latter stage that we can apply the RBC.

We can implement the WEMVA operator as follows:

- Forward propagation of the source wavefield in the background slowness field.
- Backward propagation of the receiver wavefield in the background slowness field.
- Scattering of the source wavefield upon the perturbation in the image or the perturbation in the background model.
- Scattering of the receiver wavefield upon the perturbation in the image or the perturbation in the background model.
- Zero-lag time cross-correlation of the source wavefield and scattered receiver wavefield (receiver side).

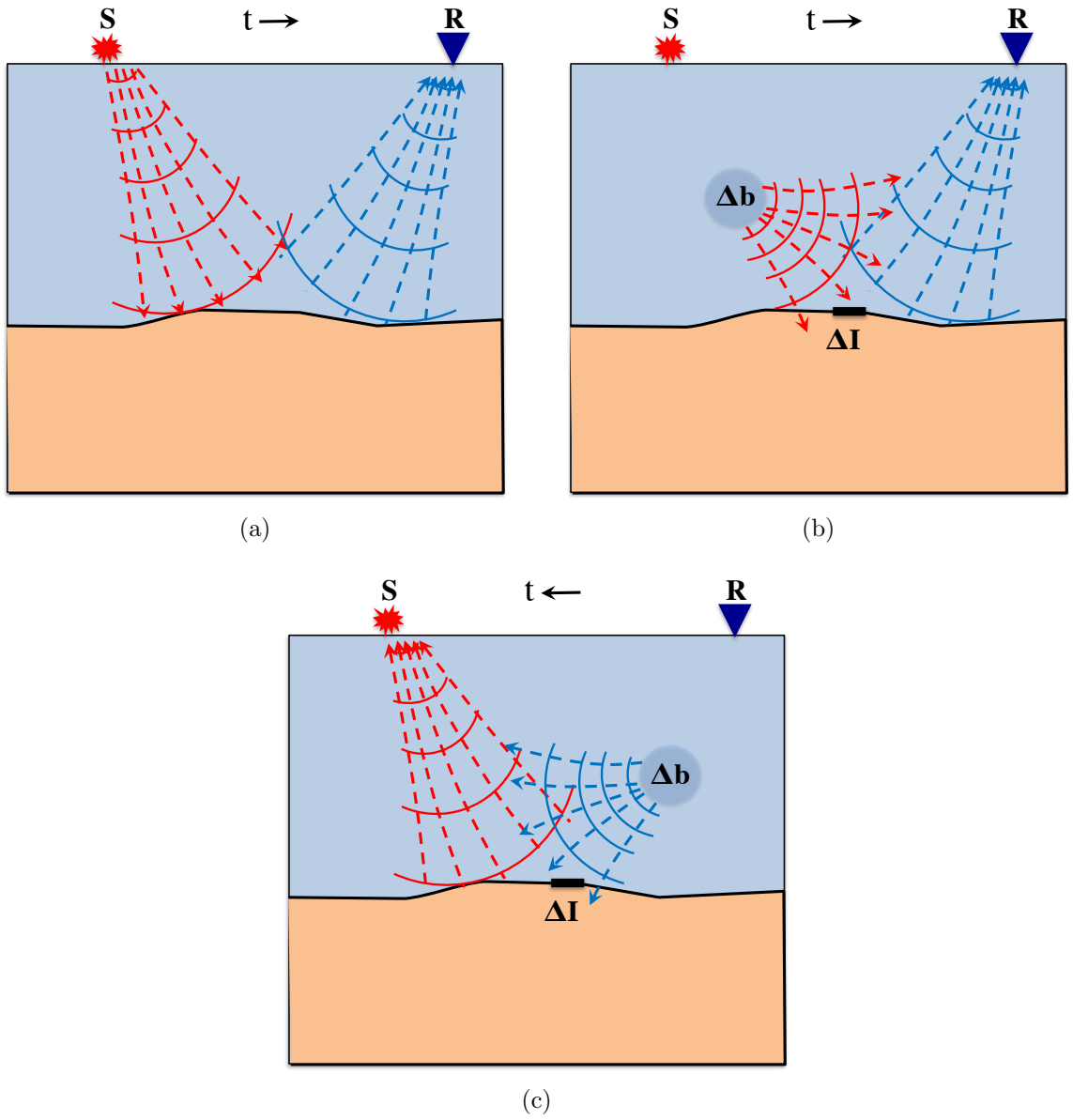


Figure 3.2: Stages of the forward WEMVA operator. a) Propagation of the background source and receiver wavefields. b) Propagation of scattered source wavefield and correlation with background receiver wavefield. c) Propagation of scattered receiver wavefield and correlation with the background source wavefield. The arrow in front of the symbol "t" represents time direction forwards (pointing right) or backward (pointing left). [**NR**]

- Zero-lag time cross-correlation of the receiver wavefield and scattered source wavefield (source side).
- Addition of source and receiver sides of WEMVA.

Next, I show the implementation of the above steps when we store the background source and receiver wavefields, then when we store only one background wavefield, and finally, when applying RBC that we do not have to store anything.

3.3.1 Implementation of the WEMVA operator storing the background wavefields

If we can afford to store the propagated wavefields, then we can execute the WEMVA process as described in Algorithm 1.

Algorithm 1 WEMVA implementation saving both source and receiver wavefields

- Forward *propagate* the source wavefield and store; then, scatter upon perturbation and forward *propagate* the scattered source wavefield.
- Backward *propagate* the receiver wavefield and store; then, scatter upon perturbation and backward *propagate* the scattered receiver wavefield.
- Perform cross-correlations during propagations of the scattered wavefields.

This procedure demands four propagations (indicated in italics): two propagations for the background wavefields and two propagations for the scattered wavefields. Note that only the source and the receiver wavefields need be stored, not the scattered wavefields.

3.3.2 Implementation of the WEMVA operator storing one background wavefield

Let us assume that we can afford to store only one background wavefield, for instance, the source wavefield. In this case, we can proceed as indicated in Algorithm 2.

Algorithm 2 WEMVA implementation storing one wavefield at a time

- Forward *propagate* the source wavefield and store it.
- Backward *propagate* the receiver wavefield and scatter upon perturbation; then, backward *propagate* the scattered receiver wavefield on the fly.
- Cross-correlate the scattered receiver wavefield with the stored source wavefield, whereas propagating the former.
- Delete the source wavefield.
- Backward *propagate* the receiver wavefield and store it.
- Forward *propagate* the source wavefield and scatter upon perturbation; then, forward *propagate* the scattered source wavefield on the fly.
- Cross-correlate the scattered source wavefield with the stored receiver wavefield, whereas propagating the former.

Notice that now we need to perform two additional propagations compared to the previous case. This is the price that we pay when we avoid storing one background wavefield.

3.3.3 Implementation of the WEMVA operator with RBC

One step further to prevent the storage of the background wavefields is using RBC to ensure the reversibility of propagations. We can proceed as indicated in Algorithm 3.

The cost of not saving wavefields at all is performing three further propagations with respect to the case of storing both wavefields. In total, we need to perform seven

Algorithm 3 WEMVA implementation storing none of the wavefields (using RBC)

- Forward *propagate* the source wavefield; then save the last two time frames.
- Backward *propagate* the receiver wavefield and scatter upon perturbation; then, backward *propagate* the scattered receiver wavefield "on the fly".
- At the same time, backward *re-propagate* the source wavefield and cross-correlate with the scattered receiver wavefield. Save the last two frames of the receiver wavefield.
- Forward *propagate* the source wavefield and scatter upon perturbation; then, forward *propagate* the scattered source wavefield "on the fly".
- At the same time, forward *re-propagate* the receiver wavefield and cross-correlate with the scattered source wavefield.

propagations. This scheme is the one that I employ in the computational codes that accompany this thesis.

Chapter 4

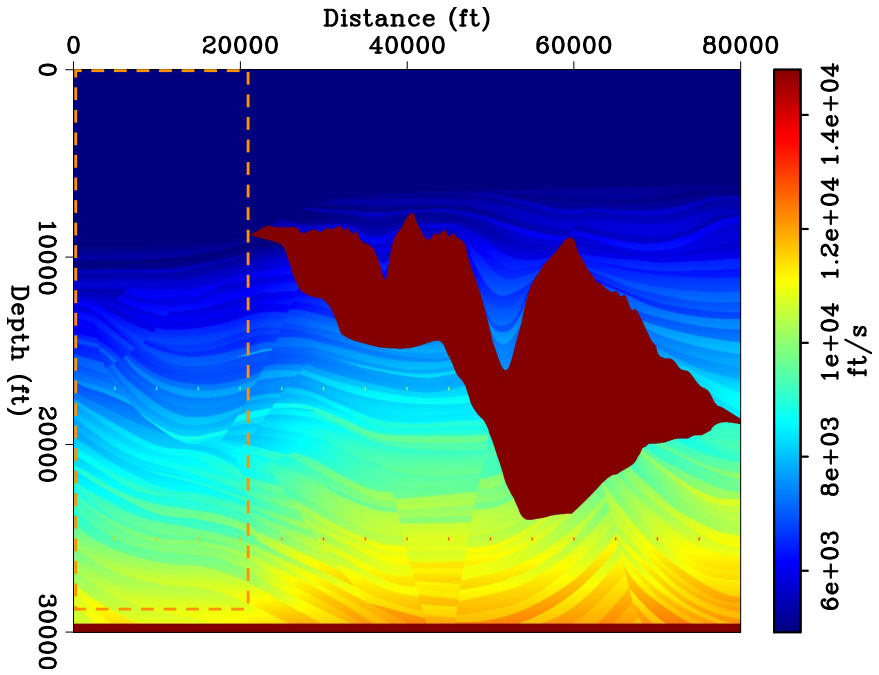
Application to synthetic 2D data

4.1 Experiment setup

In this section, I apply the JIRB method to the sedimentary section of the 2D Sigsbee 2A synthetic model (Figure 4.1). I display the background subsurface models in slowness squared, except when explicitly indicated. For the purpose of testing the validity of the method, I committed the inverse crime (see discussion in Schuster, 2017).

4.1.1 Synthetic data

I smoothed the original model and added a negative Gaussian anomaly in slowness squared (Figure 4.2(d)), corresponding to a positive velocity anomaly of approximately 850 ft/s (259 m/s) at its apex, obtaining the model shown in Figure 4.2(a). Such a model extends horizontally for 20000 ft (6096 m) spaced every 75 ft (22.86 m), and 27000 ft (8230 m) in depth spaced every 50 ft (15.24 m), and constitutes the true background model, $\mathbf{b}_{\text{true}}$, from which I synthesized Born modeled data, $\mathbf{d}(\mathbf{b}_{\text{true}})$, also using the reflectivity model shown in Figure 4.2(b). The acquisition geometry



(a)

Figure 4.1: Original Sigsbee A velocity model. The dashed rectangle indicates the sedimentary section employed for the numerical experiments. [**NR**]

consists of 54 split-spread shots spaced every 500 ft (152.4 m). Each shot gather contains 651 receivers spaced every 75 ft (22.86 m). The recording time is 11.92 s with 4 ms time step. Figure 4.3 shows a shot gather that stands at the center of the model.

For the inversion experiments, I prepared a model identical to $\mathbf{b}_{\text{true}}$, except for the Gaussian anomaly. This model without the anomaly became the wrong model, $\mathbf{b}_0$, shown in Figure 4.2(c). In the numerical implementation, the subsurface models include an extension of 26250 ft (8001 m) to the left and the right because the first and last shots stand at the limits of the image space. I obtain such an extension by replicating the last trace at each side. For the implementation of the random boundary conditions (RBC), I created velocity frames, initially within the spatial extension of each shot gather. Next, I included an extension of 1000 ft (304.8 m) to the left and to the right to account for the migration aperture. Next, I incorporated

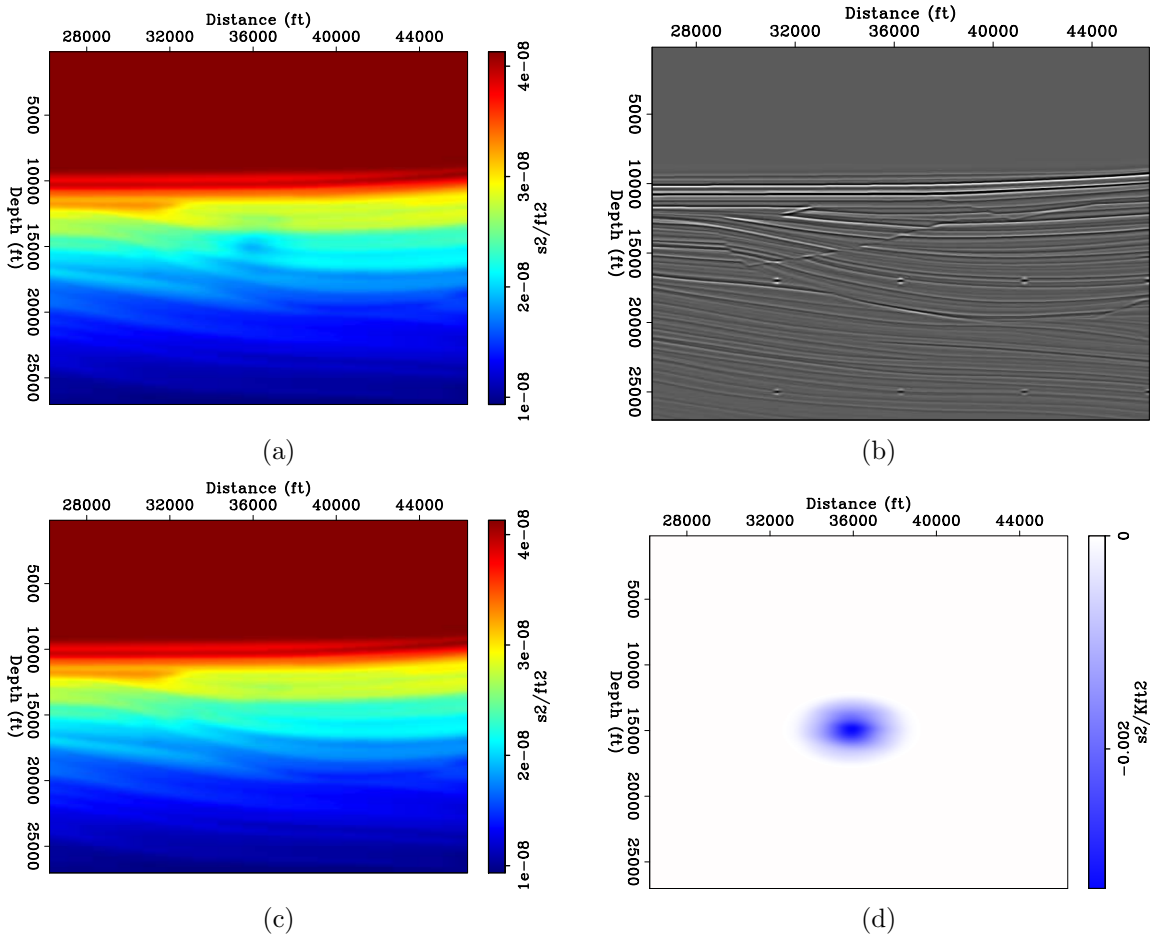
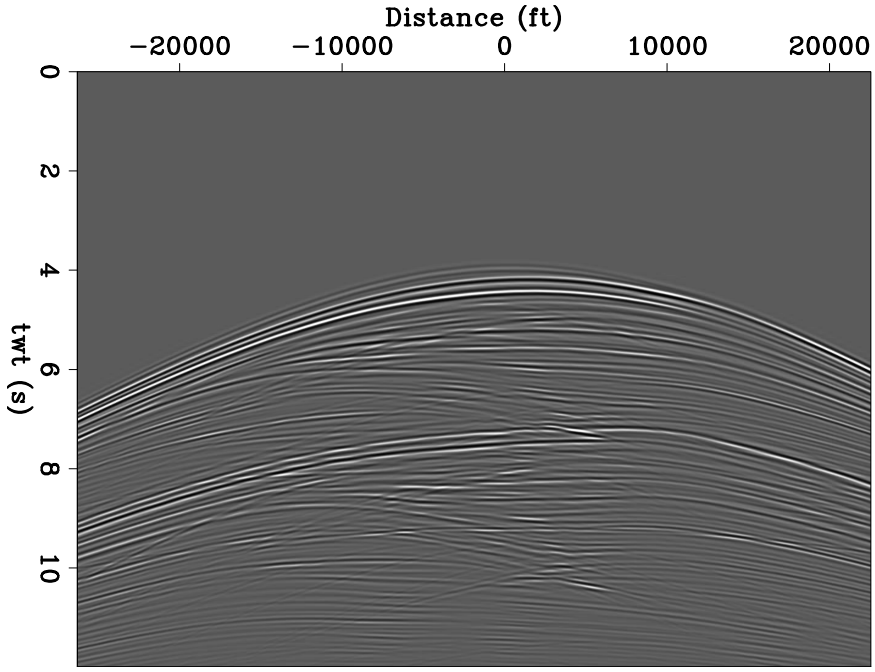


Figure 4.2: a) True background model ($\mathbf{b}_{\text{true}}$), b) true reflectivity model ($\mathbf{r}_{\text{true}}$), c) wrong background model ($\mathbf{b}_0$), and d) Gaussian anomaly representing the perturbation in the background ($\mathbf{\Delta b}$). [**ER**]

a halo of 1000 ft (304.8 m) that constitutes the support for the RBC. For the upper part of the model, I added extra space before the random halo to prevent artifacts that arise when the source stands at or close to the RBC. Figure 4.4 illustrates the velocity frame for a single shot close to the center of the model space.



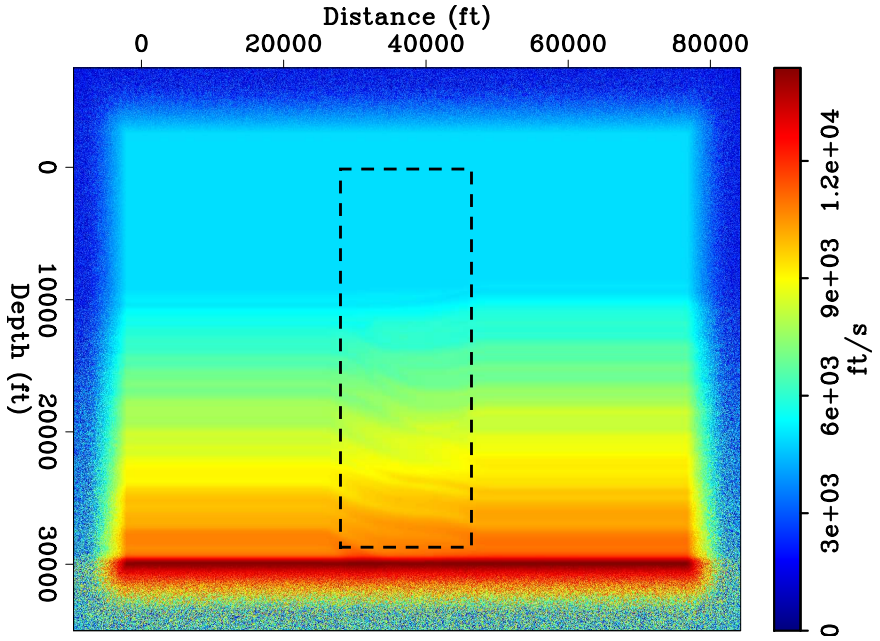
(a)

Figure 4.3: Single shot gather of the Sigsbee A model. [**CR**]

4.1.2 Inversion preliminaries

I pre-computed the Gauss-Newton Hessian of FWI, $\mathbf{H}(\mathbf{b}_0)$, using the point-spread functions (PSFs) shown in Figure 4.5. I produced this section by "seeding" spikes every 15 gridpoints and applying forward Born modeling followed by adjoint Born modeling. Gridpoints at the center of the PSFs correspond to zero-lag elements of the Hessian matrix, which are part of its main diagonal. Similarly, points picked at non-central locations of the PSFs correspond to off-diagonal elements of the Hessian matrix with the same non-zero lag.

For the Hessian application, I first pick samples corresponding to a determined lag. Next, I apply bilinear interpolation, followed by smoothing using a triangle filter, and then perform the Hessian multiplication for this lag "on the fly". I move on to the next lag and repeat the operation.



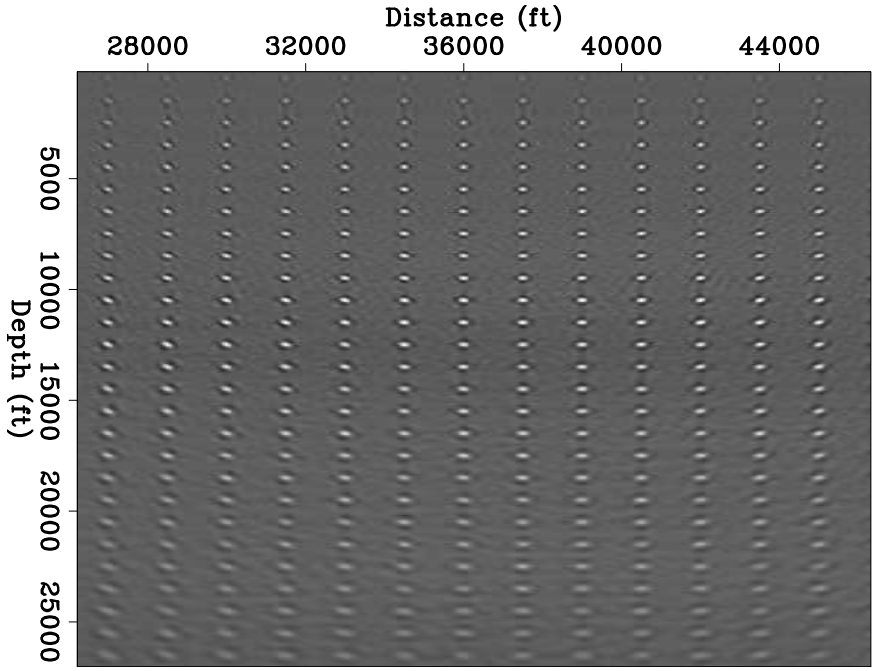
(a)

Figure 4.4: Velocity framework for the RBC for the central shot. The dashed square indicates the imaging area. [**NR**]

I had previously performed conventional WEMVA tests to invert for the background model component independently. During these tests, I incorporated a gaining weighting factor of the form $\mathbf{E} = \text{diag}(z_0^n, z_1^n, z_2^n, ...)$ (where z_i represents depth levels, and n is a positive integer), and applied it to the migration image to improve the inversion results (see Appendix B). I experimentally found $n = 2.5$ as the best value for the power coefficient. I incorporated this weighting factor to the WEMVA operator member of the JIRB implementation. Likewise, I also incorporated B-splines to precondition the gradient.

After including both weighting and preconditioning, the JIRB objective function becomes

$$\Phi(\mathbf{r}, \mathbf{p}) = \frac{1}{2} \left\| \mathbf{E}\big[\mathbf{H}(\mathbf{b}_0)\mathbf{r} - \mathbf{I}(\mathbf{B}\mathbf{p})\big] \right\|_2^2 - \frac{\lambda^2}{2} \left\| \mathbf{E}\mathbf{I}(\mathbf{B}\mathbf{p}) \right\|_2^2, \tag{4.1}$$



(a)

Figure 4.5: Point-spread functions computed on the sedimentary section of Sigsbee A. I gained and clipped the section for visualization purposes. [**CR**]

where the inverted background component model can be recovered using $\mathbf{b}_{\text{inv}} = \mathbf{B}\mathbf{p}_{\text{inv}}$. The JIRB experiments used the nonlinear steepest descent method. The inversion tests ran until the solver was unable to find an adequate step size to update the model.

I plotted the inverted reflectivity images displayed in the following section at the same color scale within the same figure set. For comparison purposes, I also computed conventional LWI (equation (1.2)) using the correct background model. Such inversion ran for 30 iterations.

For the background models, I compare perturbations in the background, $\mathbf{\Delta b} = \mathbf{b} - \mathbf{b}_0$, rather than the absolute background models, for better visualization. I plotted the results at the same color scale within each figure set, except when specified.

4.2 Numerical results

4.2.1 Comparing JIRB with LWI

Figure 4.6(a) shows the LWI reflectivity using the correct background model. Note how the amplitude values closely resemble those of the true reflectivity section in Figure 4.6(b). Figure 4.7 shows a zoom into the depth interval of $12000 - 22000$ ft for a better visualization. Note that seismic features at LWI have the correct kinematics. I hereafter use this zoom level for the rest of this section. I also regarded the LWI reflectivity model as the desired result.

Figure 4.8 shows the LWI reflectivity using the wrong background model compared with the true reflectivity. As expected, reflections at the center of the image get pulled upwards because of the absence of the Gaussian anomaly (note in particular the central spike relative to the grid lines in both cases). There is also unfocused energy in the reflectors of the lower part of the section.

Next, I display each JIRB result organized in sets of figures as follows: a) JIRB reflectivity vs. desired LWI reflectivity, both plotted with the same color scale and independent color scale, the latter to better appreciate reflectivity details otherwise obscured by the section whose amplitude dominates; b) JIRB perturbation in the background vs. true perturbation in the background; and c) trace comparisons of JIRB, LWI, and true reflectivities (I extracted the central trace for such comparisons).

Figure 4.9 shows the inverted reflectivity model obtained with JIRB using $\lambda = 12$. Figure 4.10 shows the comparison of the corresponding JIRB perturbation in the background with the true perturbation in the background. Here and hereafter I display the perturbations rather than the background models for the sake of better visualization. I chose this example as the baseline because I found that it approximately represents the minimum value of λ at which we can obtain the desired correction. Note how the reflectors moved back to their correct position, and the reflection events are almost

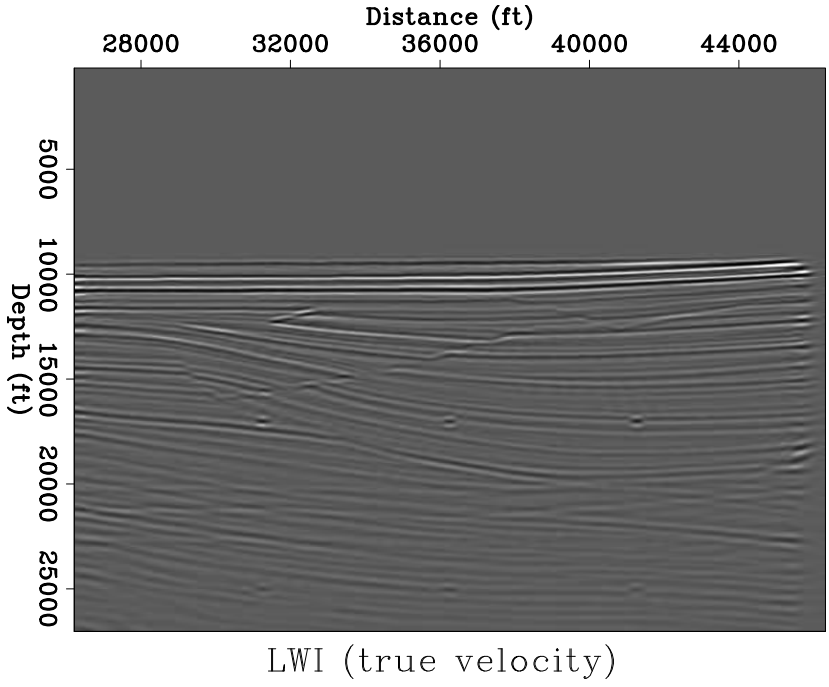
kinematically equivalent (Figures 4.9(c) and 4.9(d)). Compare with the LWI reflectivity. The correction in the reflectivity progresses at the same time as the background model gets corrected during the inversion. Nonetheless, the amplitudes are higher than the desired reflectivity (Figures 4.9(a) and 4.9(b)), as further corroborated by the trace comparison for JIRB, LWI, and the true reflectivity, shown in Figure 4.11. One solution to this problem can be the incorporation of the minimum norm regularization of the reflectivity into the JIRB objective function. The main drawback is the introduction of an additional trade-off parameter that we ought to estimate.

Elemental analysis of the objective function (4.1) promptly drives us to deduce that higher values of λ underweight the reflectivity-related gradient contribution. Conversely, smaller values of λ exacerbate the problem of higher amplitude, as Figures 4.15 and 4.17 exemplify for $\lambda = 10$, notwithstanding that the JIRB perturbation in the background appears to be unaffected (Figure 4.16). Going for smaller values, e.g., $\lambda = 5$, further underweights the background-related gradient to the point that it would not be able to correct the reflectivity inaccuracies. Figures 4.12 and 4.14 illustrate this problem. Note that the central spike does not reach the correct position. This issue is due to the comparatively poor recovery of the background component as observed in Figure 4.13, which the reader should compare to Figure 4.10(a)).

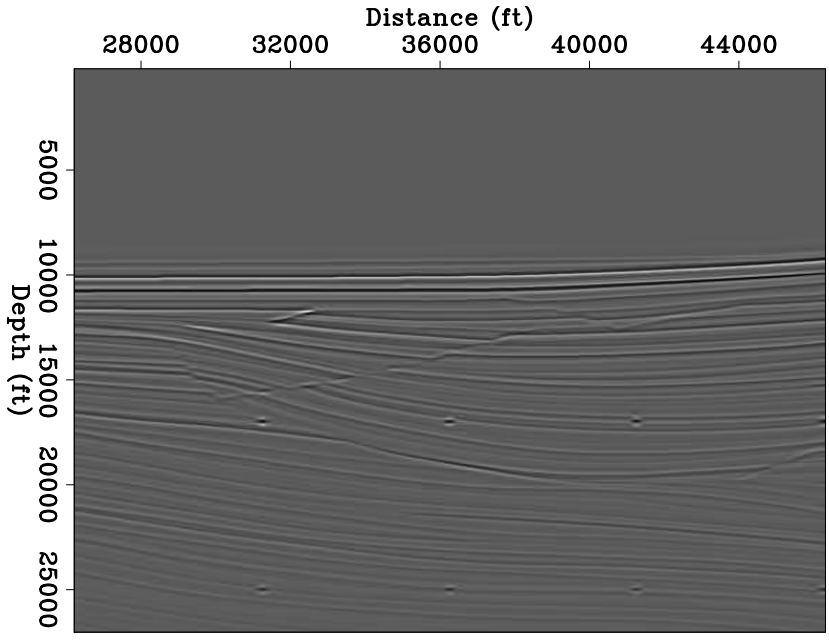
Going for higher values of λ, I experimented using $15, 20$, and 25. I also used $\lambda = 30$, but the sections are not very different from $\lambda = 25$, hence they are not shown. Figures 4.18, 4.20, and 4.22 respectively show the corresponding reflectivity sections, while Figures 4.19, 4.21, and 4.23 show the corresponding trace comparisons. I omitted the perturbations in the background because they do not significantly differ from the case of $\lambda = 12$. We observe that the reflectivity amplitudes of the JIRB results get reduced as λ values increase, as anticipated. The best result is arguably obtained using $\lambda = 25$. Nonetheless, note in trace comparisons of Figure 4.23 that, at shallow depths, amplitudes get underestimated, while at deep depths amplitudes still get overestimated.

One problem derived from increasing the λ value is that the solver stops at earlier

iterations. As a consequence, the JIRB images lose resolution in comparison to LWI. Conversely, keeping higher resolution by reducing the λ value yields the higher-than-ideal amplitudes observed before. For a better appreciation of this issue, Figure 4.24 shows the comparison of the amplitude spectra of the JIRB and LWI traces plotted above. From these plots, we can arguably choose the $\lambda = 20$ or the $\lambda = 25$ reflectivity result as the best compromise between amplitude accuracy and seismic resolution.

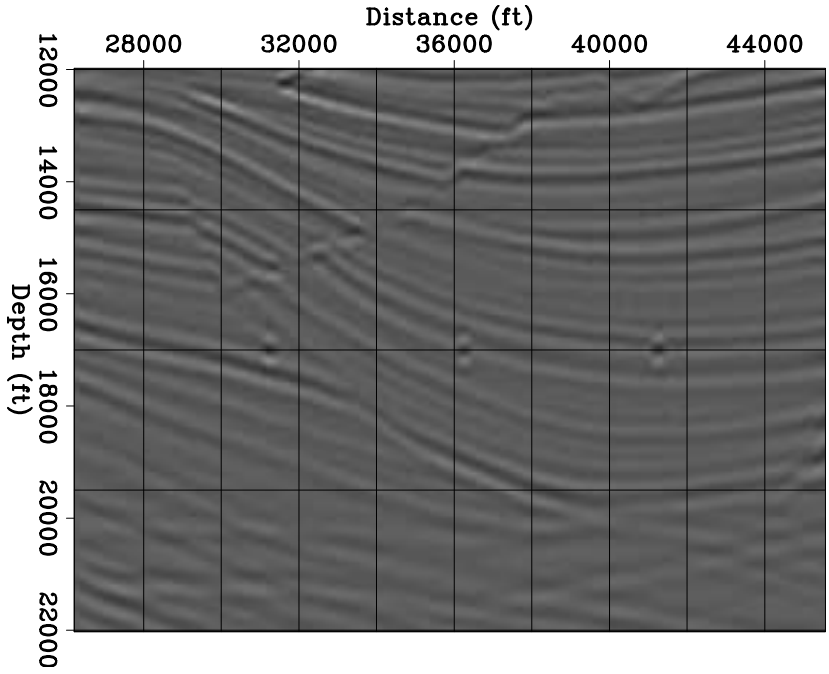


(a)

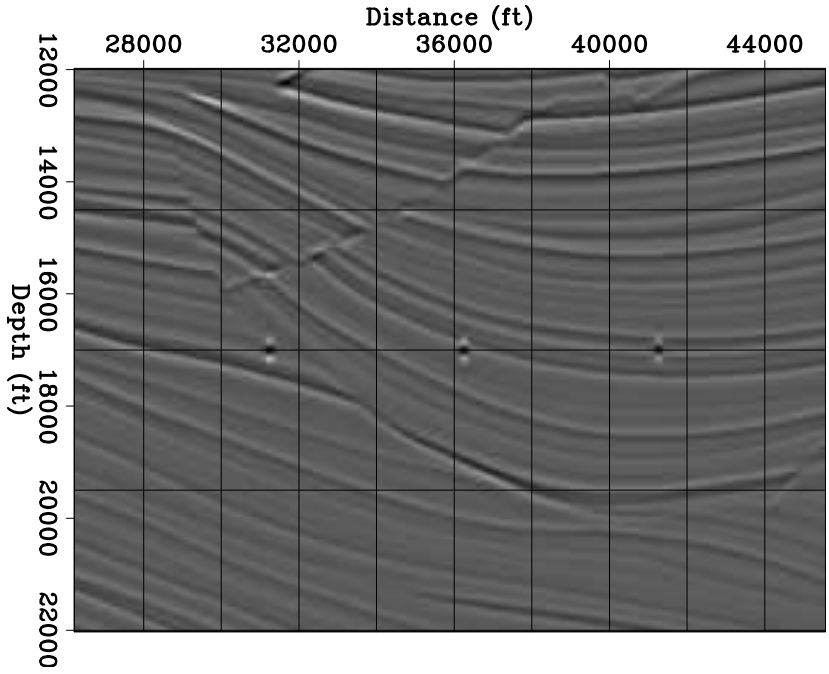


(b)

Figure 4.6: Comparison between a) conventional LWI with the true background model, and b) true reflectivity. [**CR**]

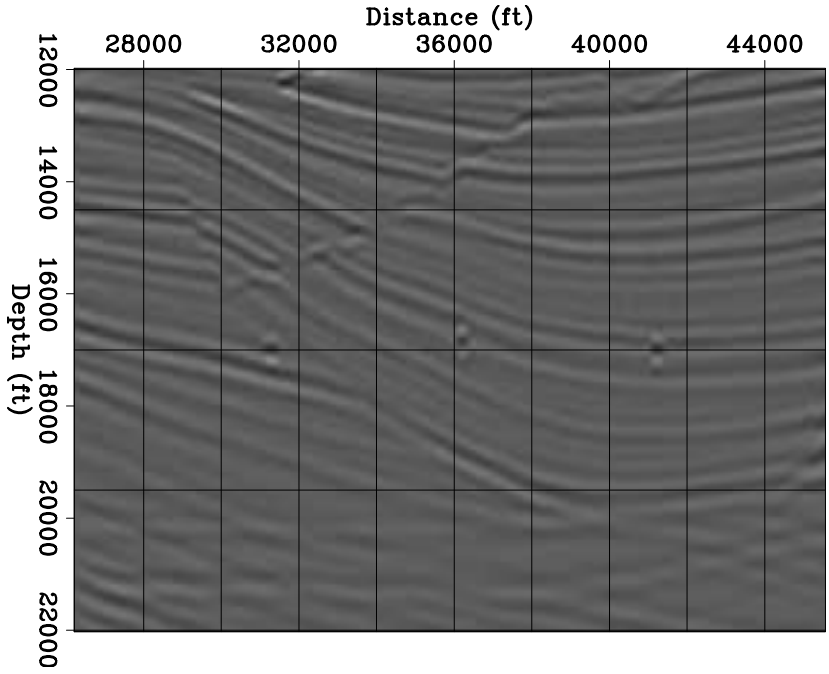


(a)

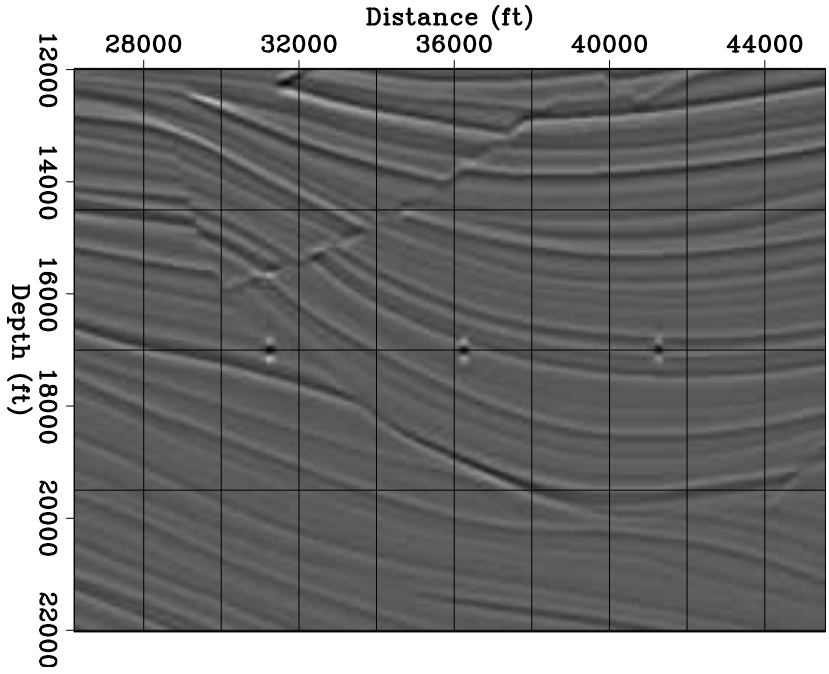


(b)

Figure 4.7: Zoom at sections in Figure 4.6: a) LWI with true background model; b) true reflectivity. [**CR**]



(a)



(b)

Figure 4.8: Zoom at a) LWI with wrong background model, and b) true reflectivity. [**CR**]

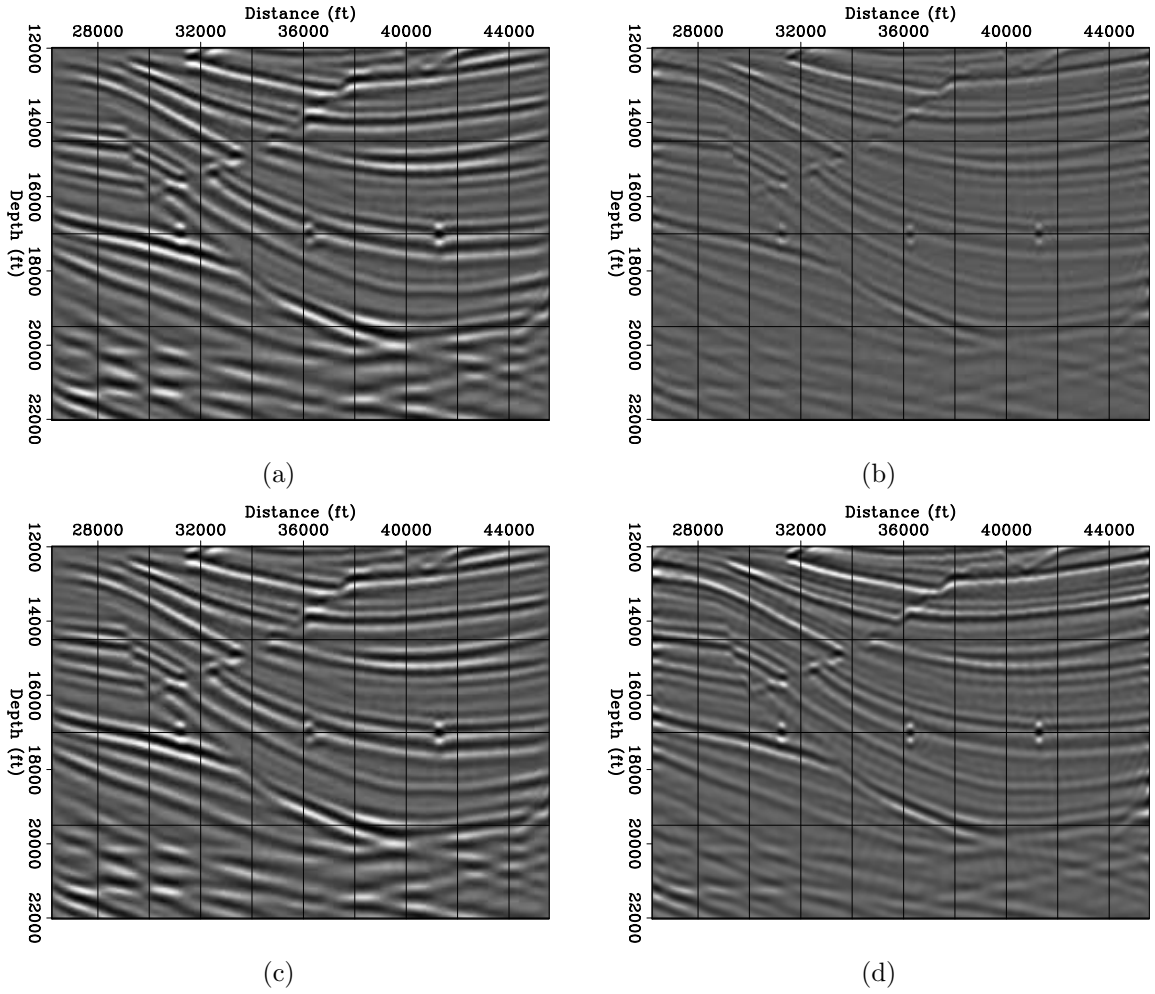
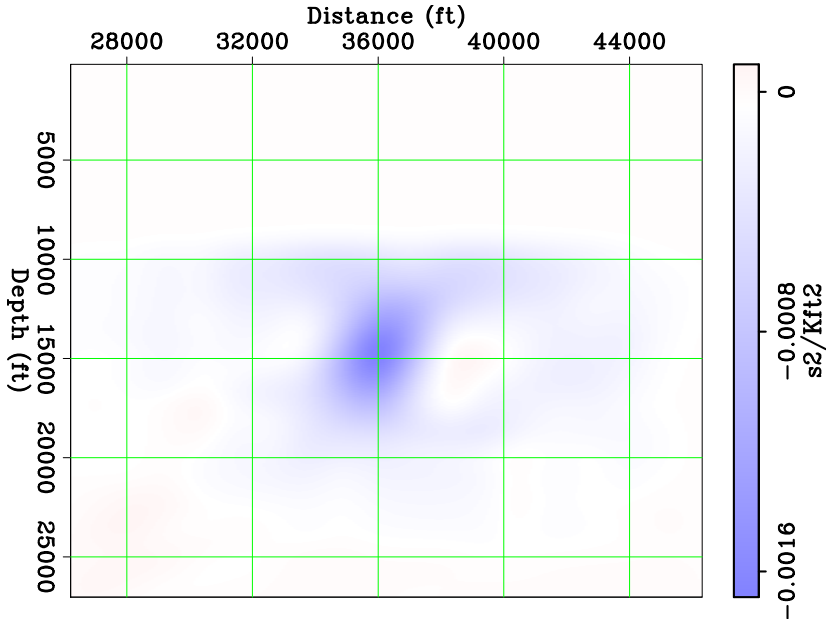
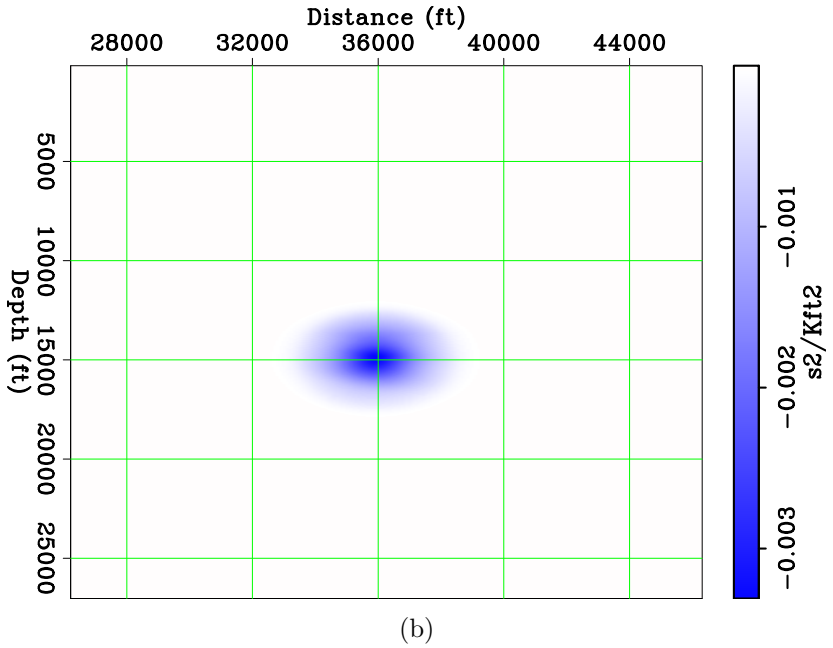


Figure 4.9: a) JIRB reflectivity using $\lambda = 12$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) but plotted with independent color scales. [**CR**]

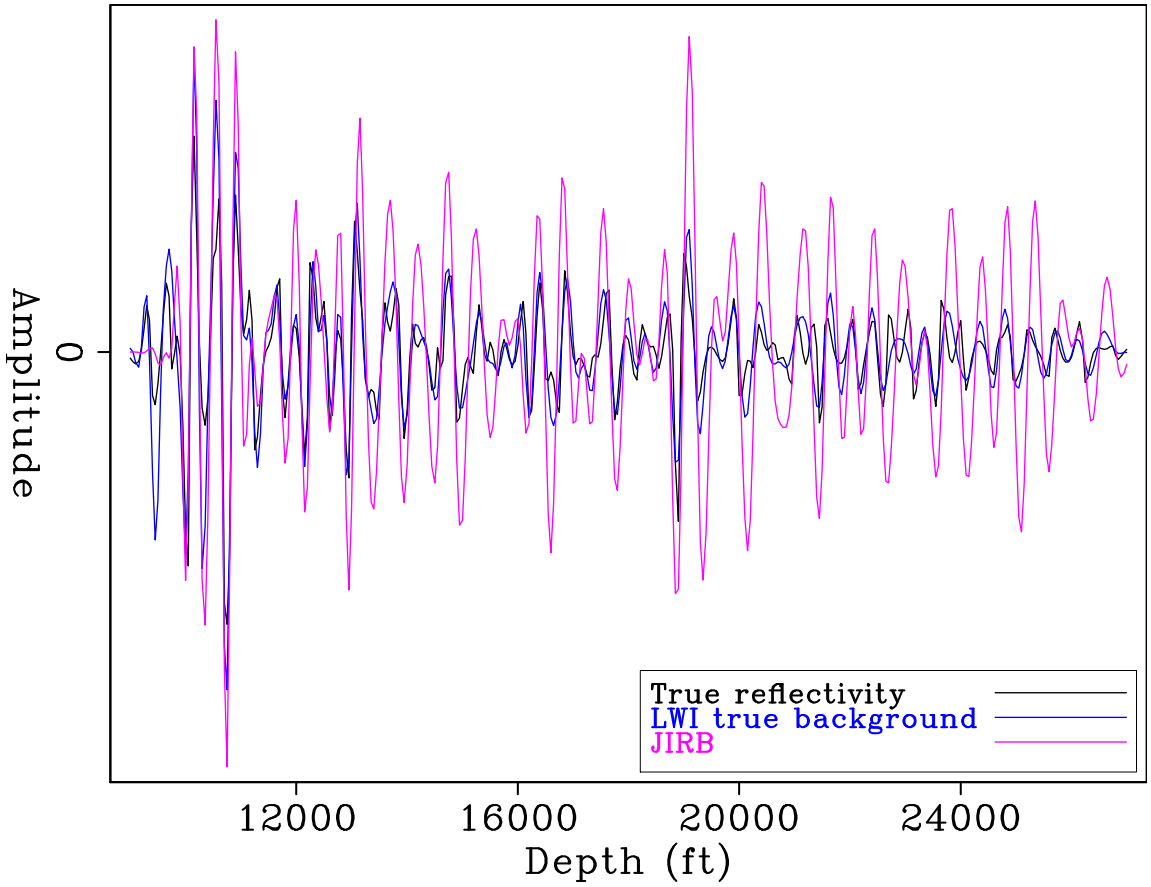


(a)



(b)

Figure 4.10: a) JIRB perturbation on the background using $\lambda = 12$. b) True perturbation in the background. [**CR**]



(a)

Figure 4.11: Trace comparison between JIRB ($\lambda = 12$), LWI, and true reflectivity. [**CR**]

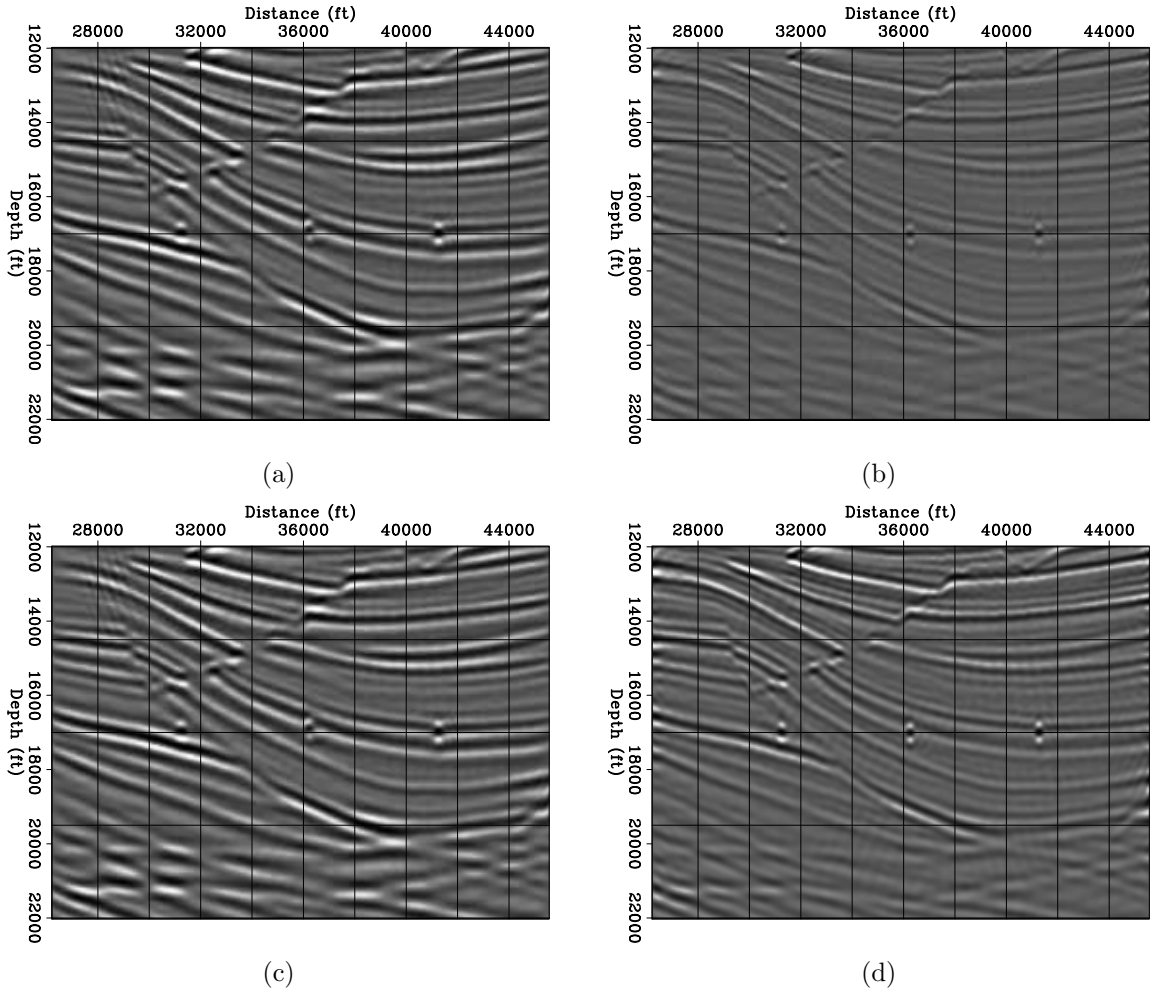


Figure 4.12: a) JIRB reflectivity using $\lambda = 5$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) but plotted with independent color scales. [**CR**]

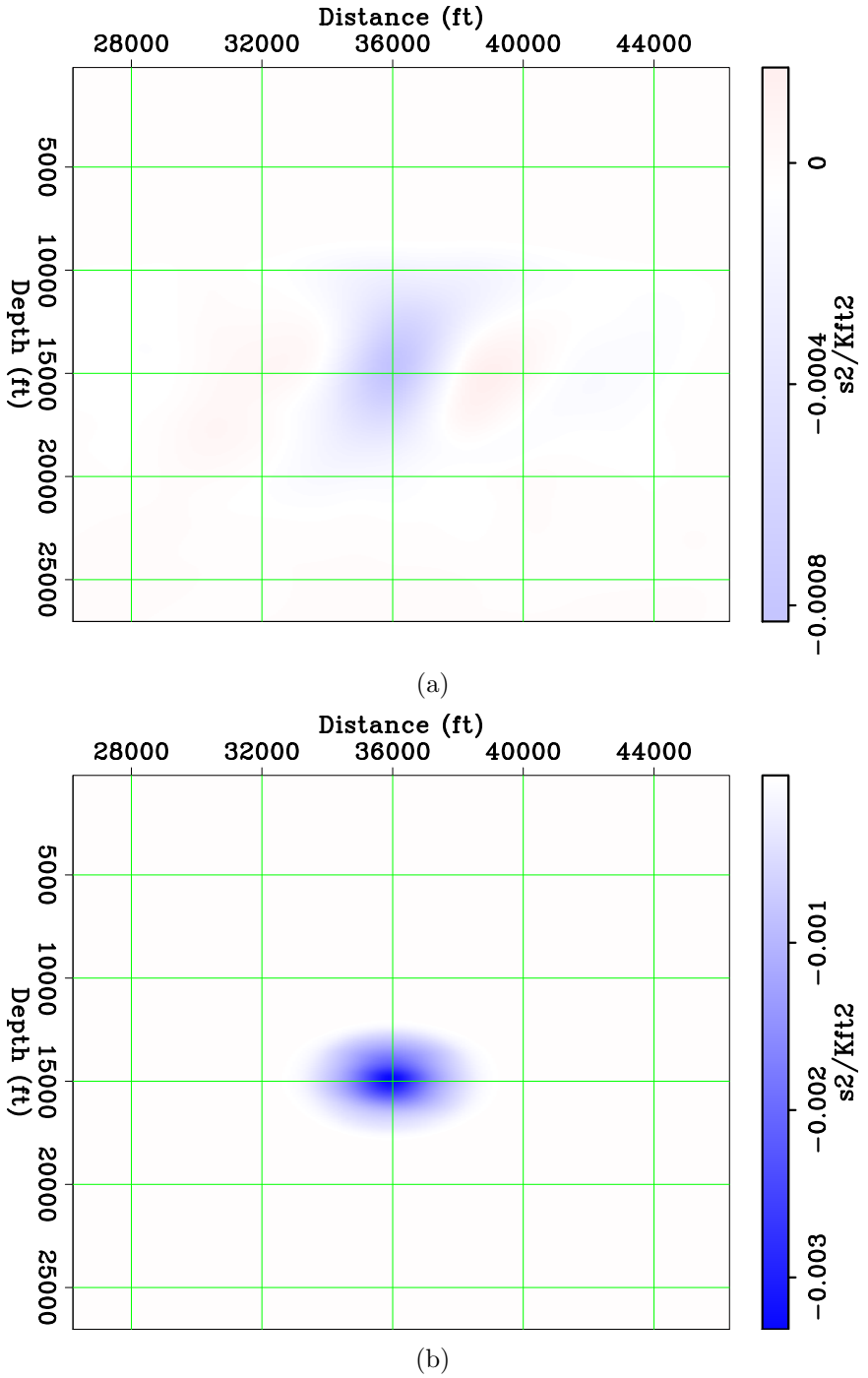
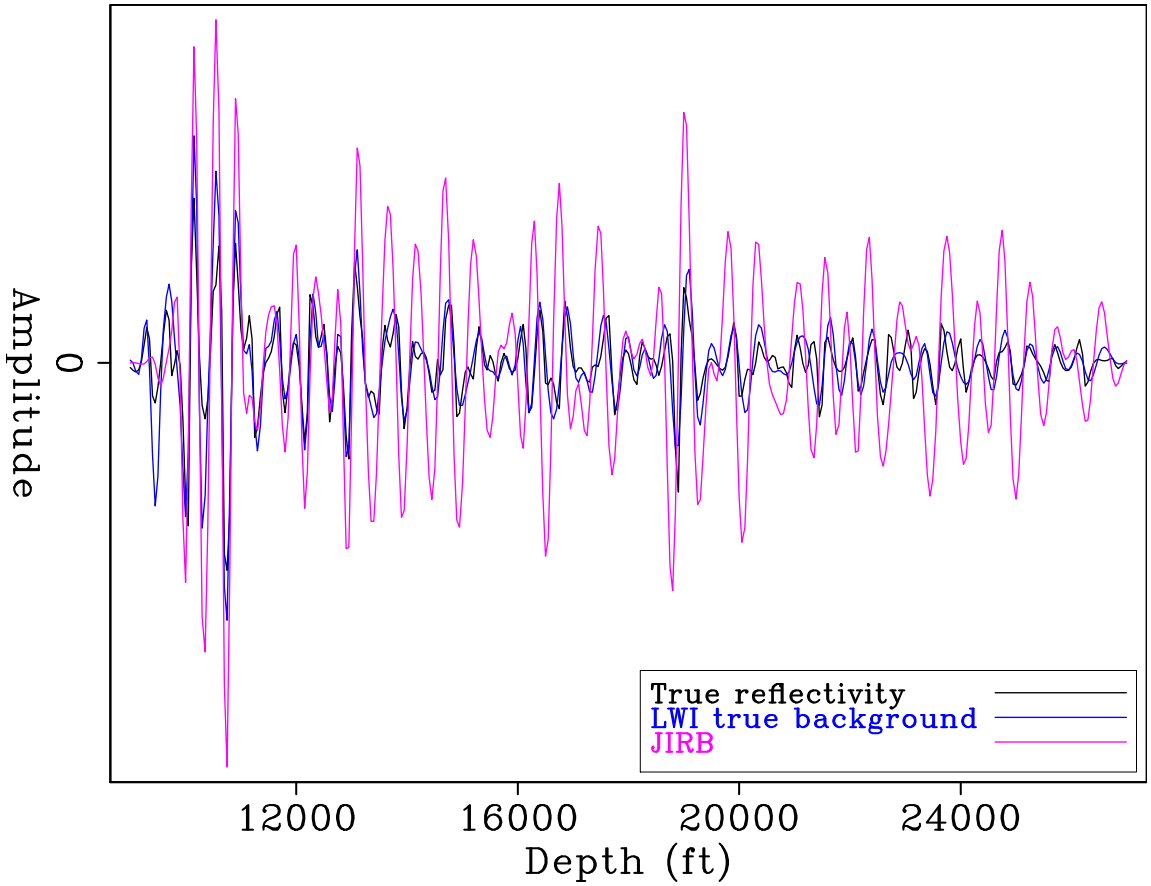


Figure 4.13: a) JIRB perturbation on the background using $\lambda = 5$. b) True perturbation in the background. [**CR**]



(a)

Figure 4.14: Trace comparison between JIRB ($\lambda = 5$), LWI, and true reflectivity. [**CR**]

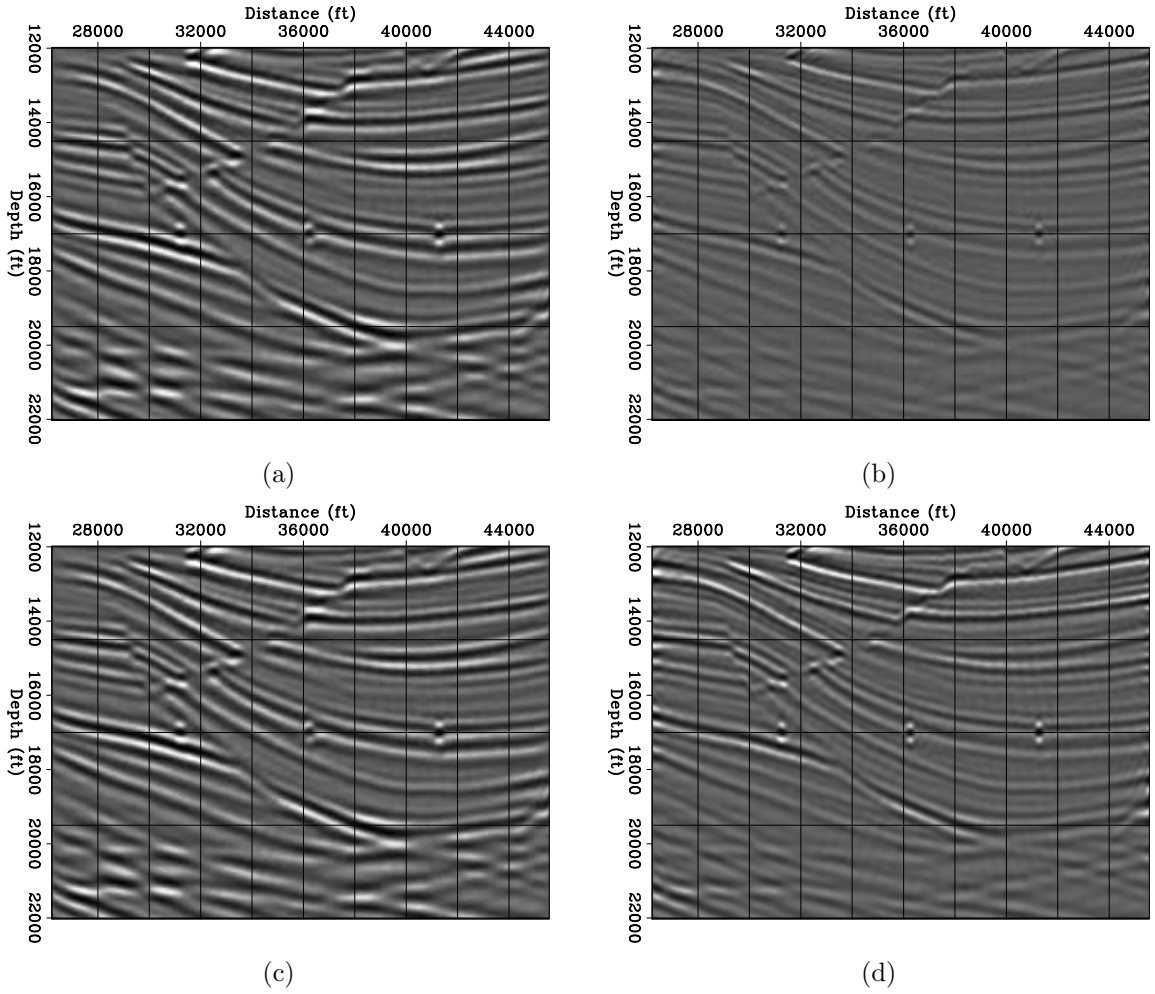


Figure 4.15: a) JIRB reflectivity using $\lambda = 10$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) but plotted with independent color scales. [**CR**]

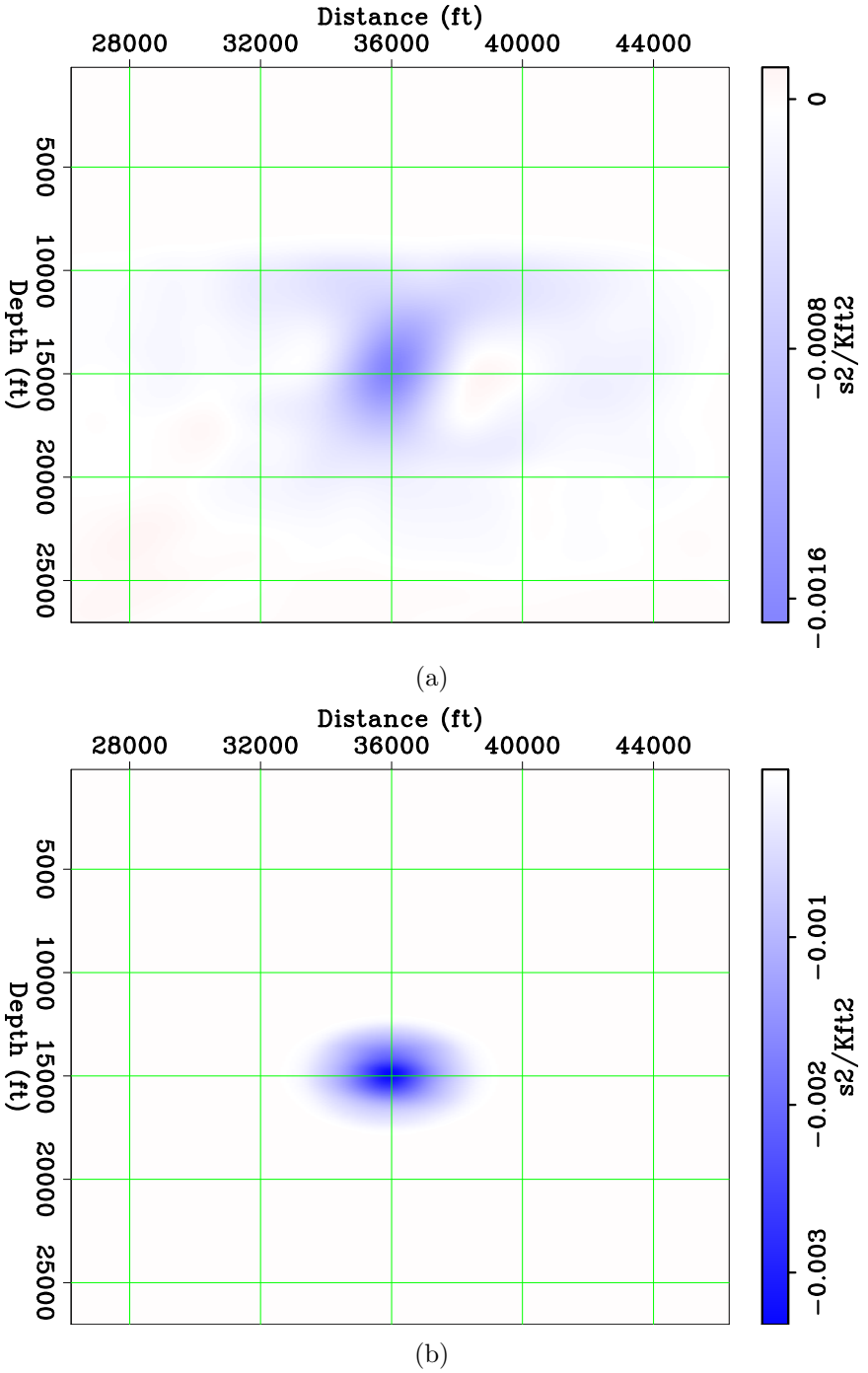
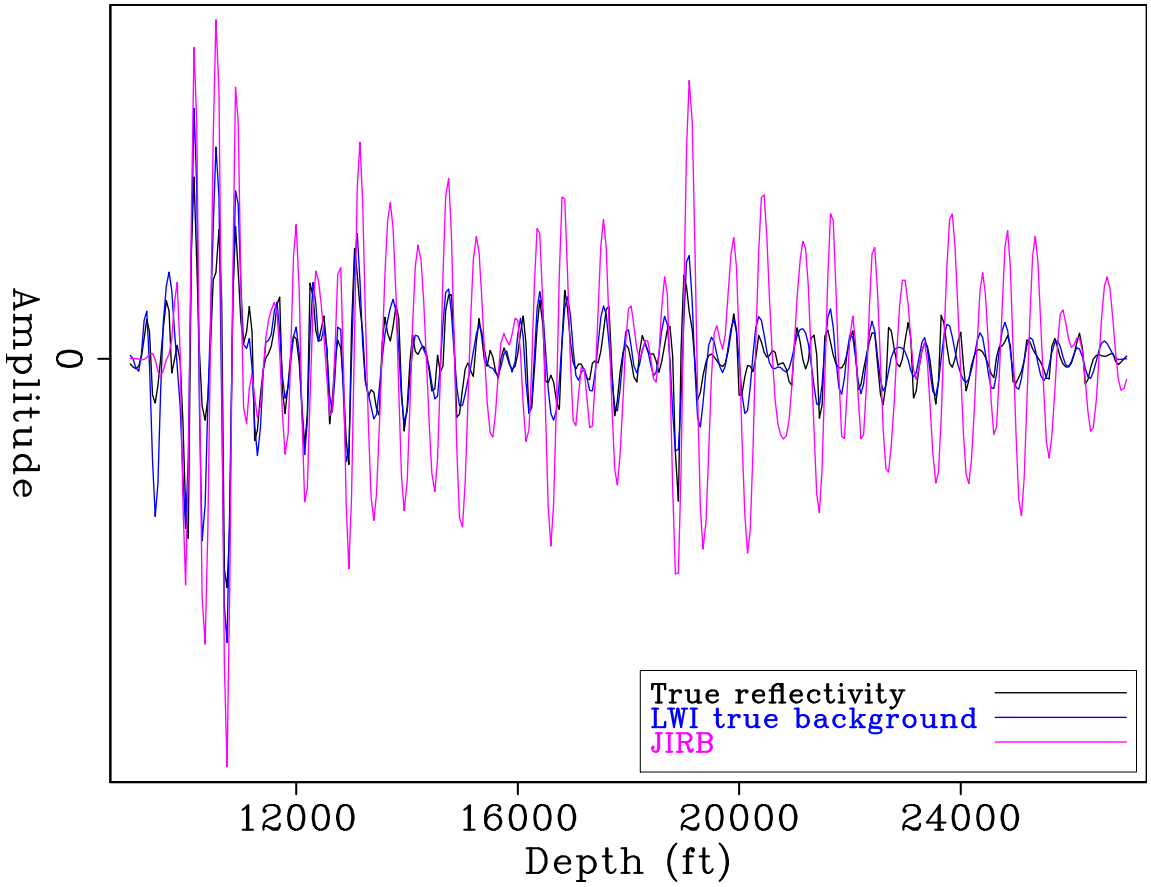


Figure 4.16: a) JIRB perturbation on the background using $\lambda = 10$. b) True perturbation in the background. [**CR**]



(a)

Figure 4.17: Trace comparison between JIRB ($\lambda = 10$), LWI, and true reflectivity. [**CR**]

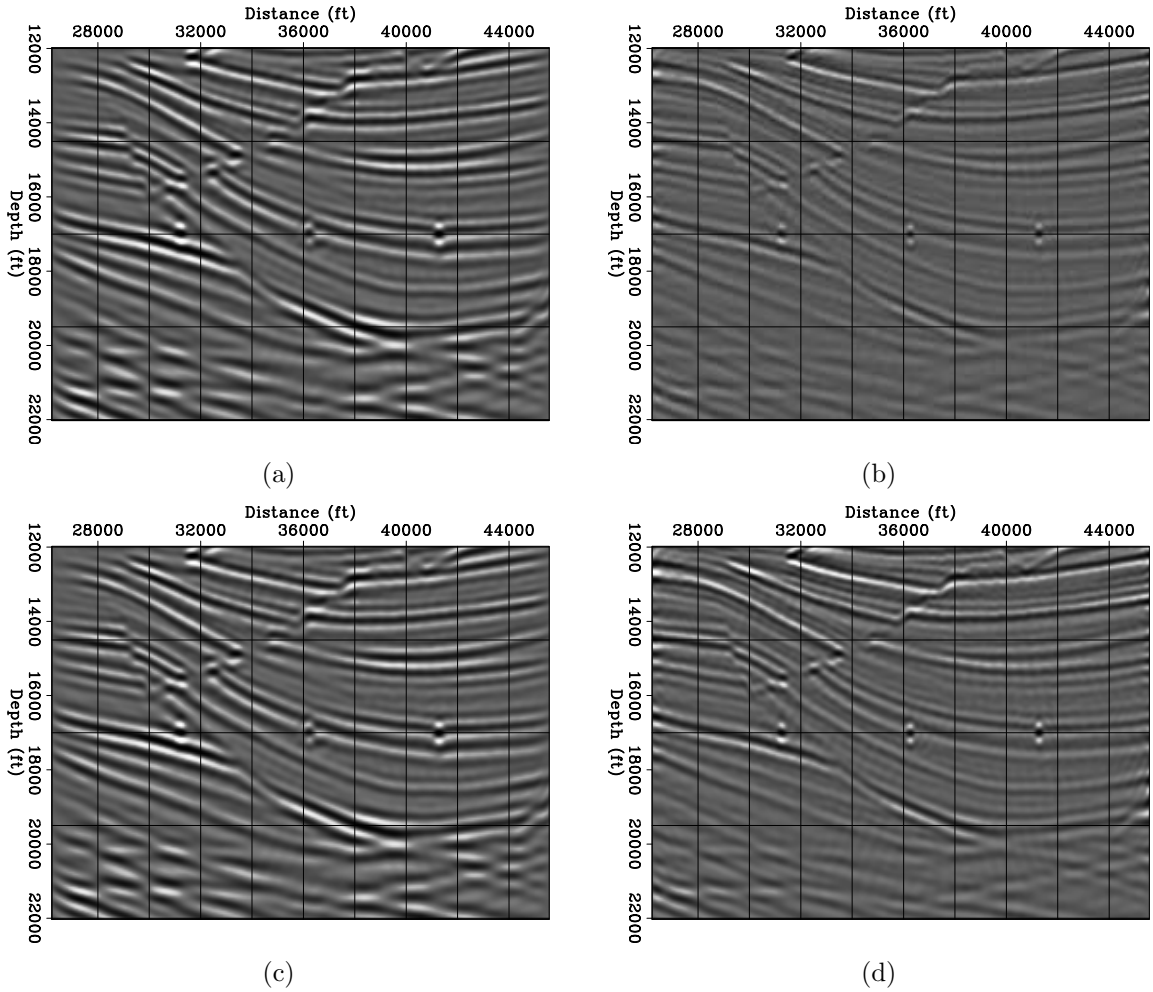
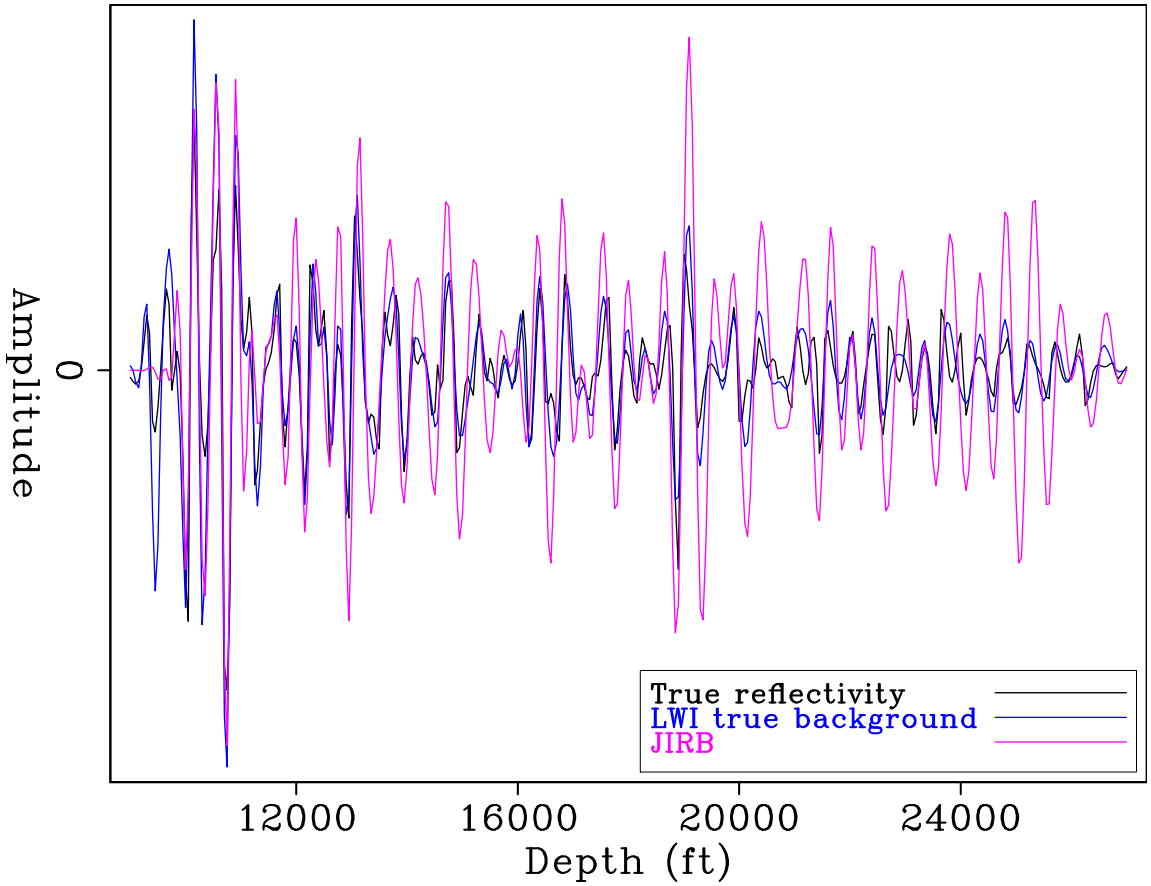


Figure 4.18: a) JIRB reflectivity using $\lambda = 15$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) with independent color scales. [**CR**]



(a)

Figure 4.19: Trace comparison between JIRB ($\lambda = 15$), LWI, and true reflectivity. [**CR**]

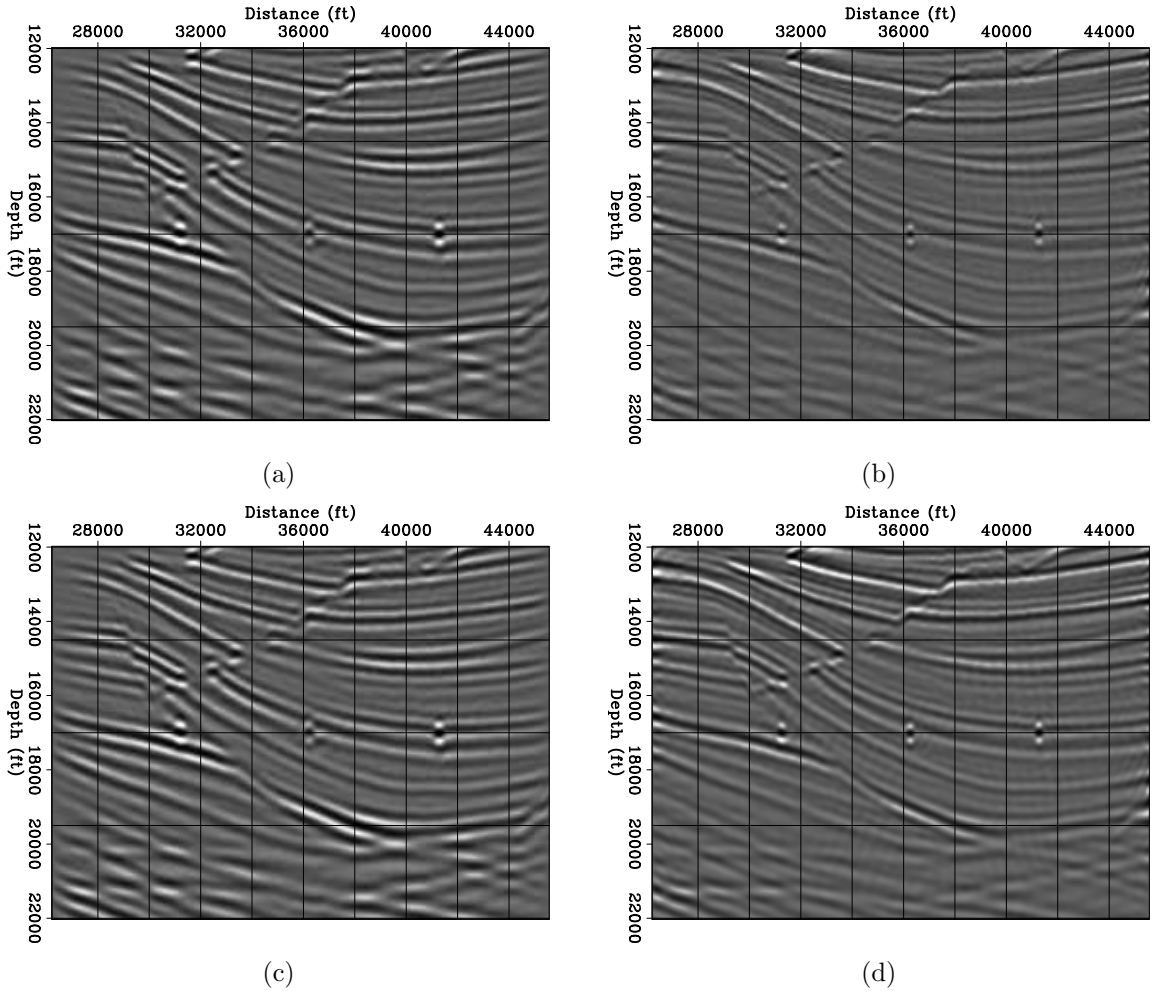
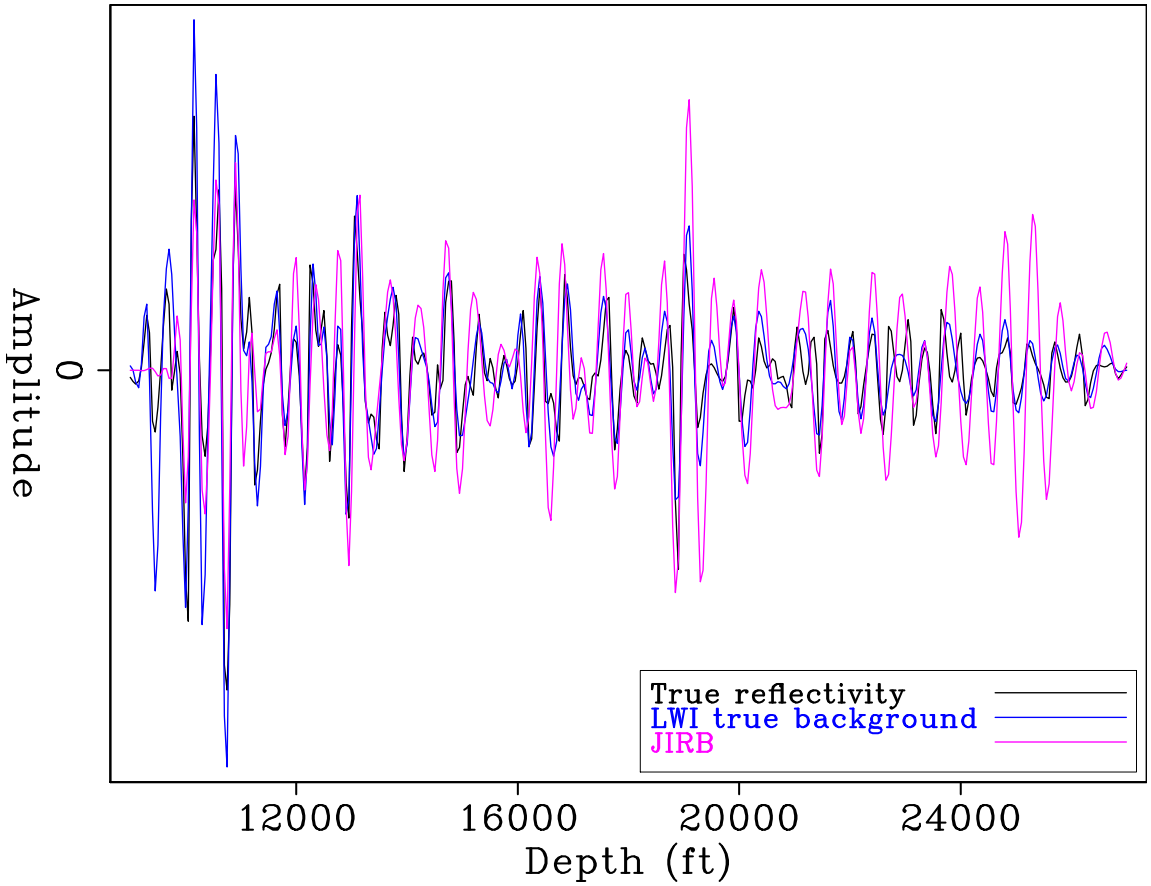


Figure 4.20: a) JIRB reflectivity using $\lambda = 20$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) with independent color scales. [**CR**]



(a)

Figure 4.21: Trace comparison between JIRB ($\lambda = 20$), LWI, and true reflectivity. [**CR**]

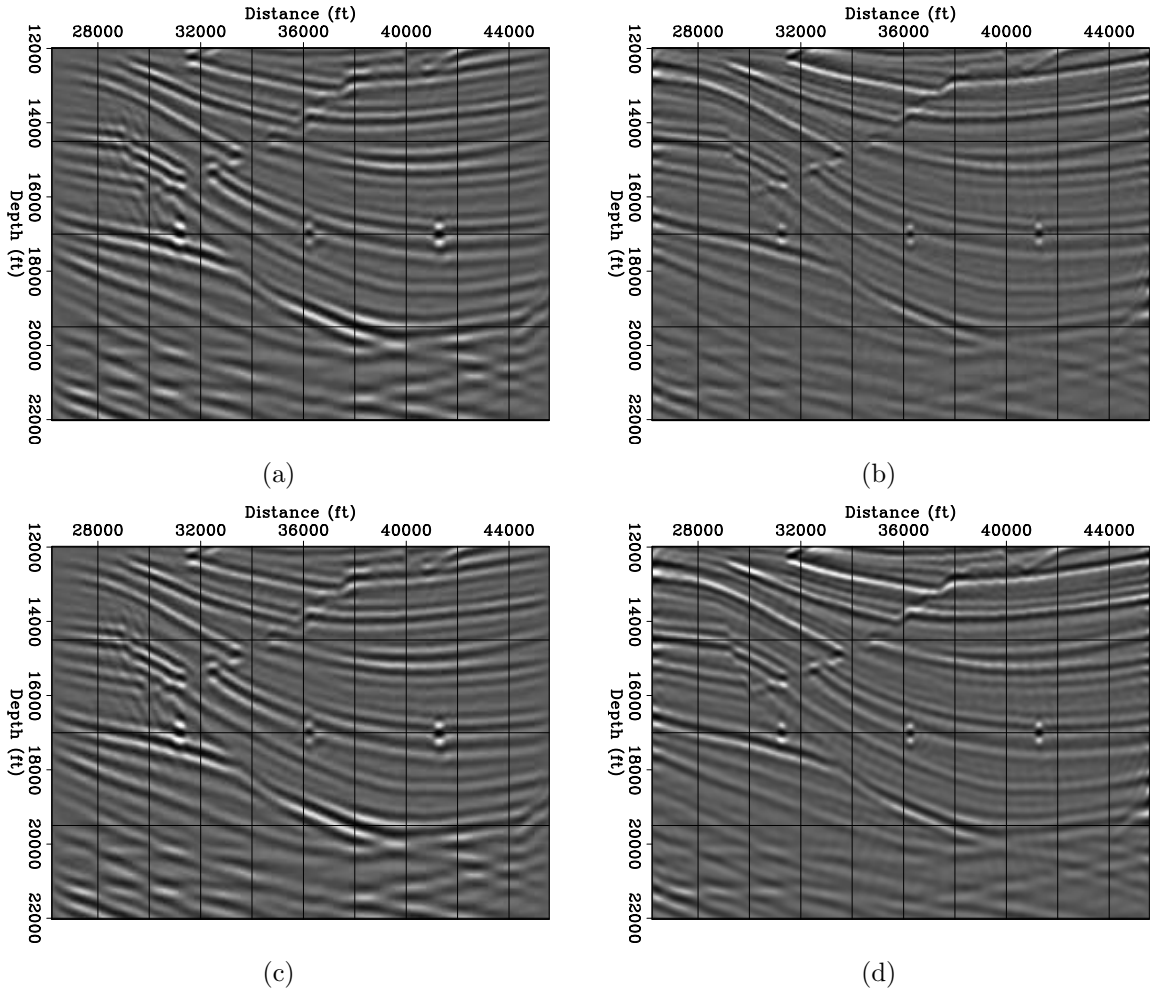
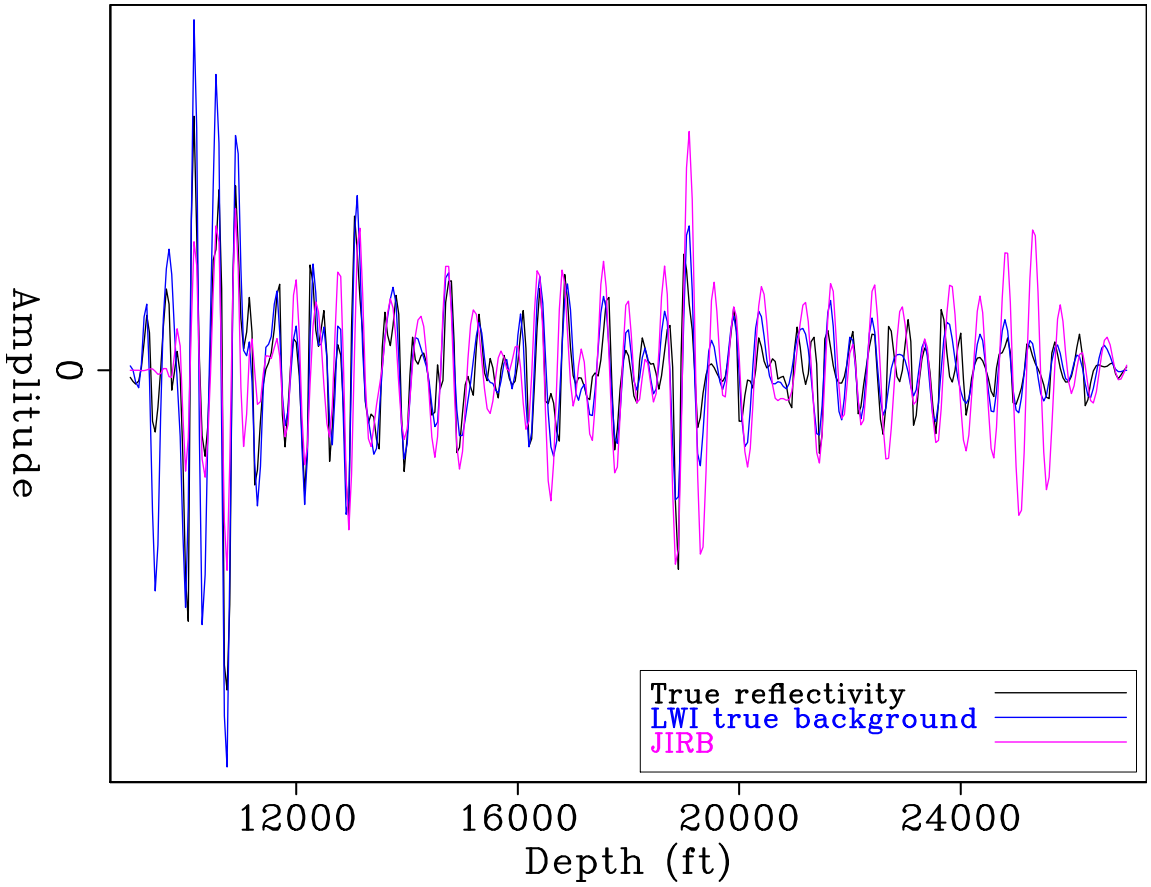


Figure 4.22: a) JIRB reflectivity using $\lambda = 25$. b) LWI reflectivity with true background model. c) and d) Same as a) and b) with independent color scales. [**CR**]



(a)

Figure 4.23: Trace comparison between JIRB ($\lambda = 25$), LWI, and true reflectivity. [**CR**]

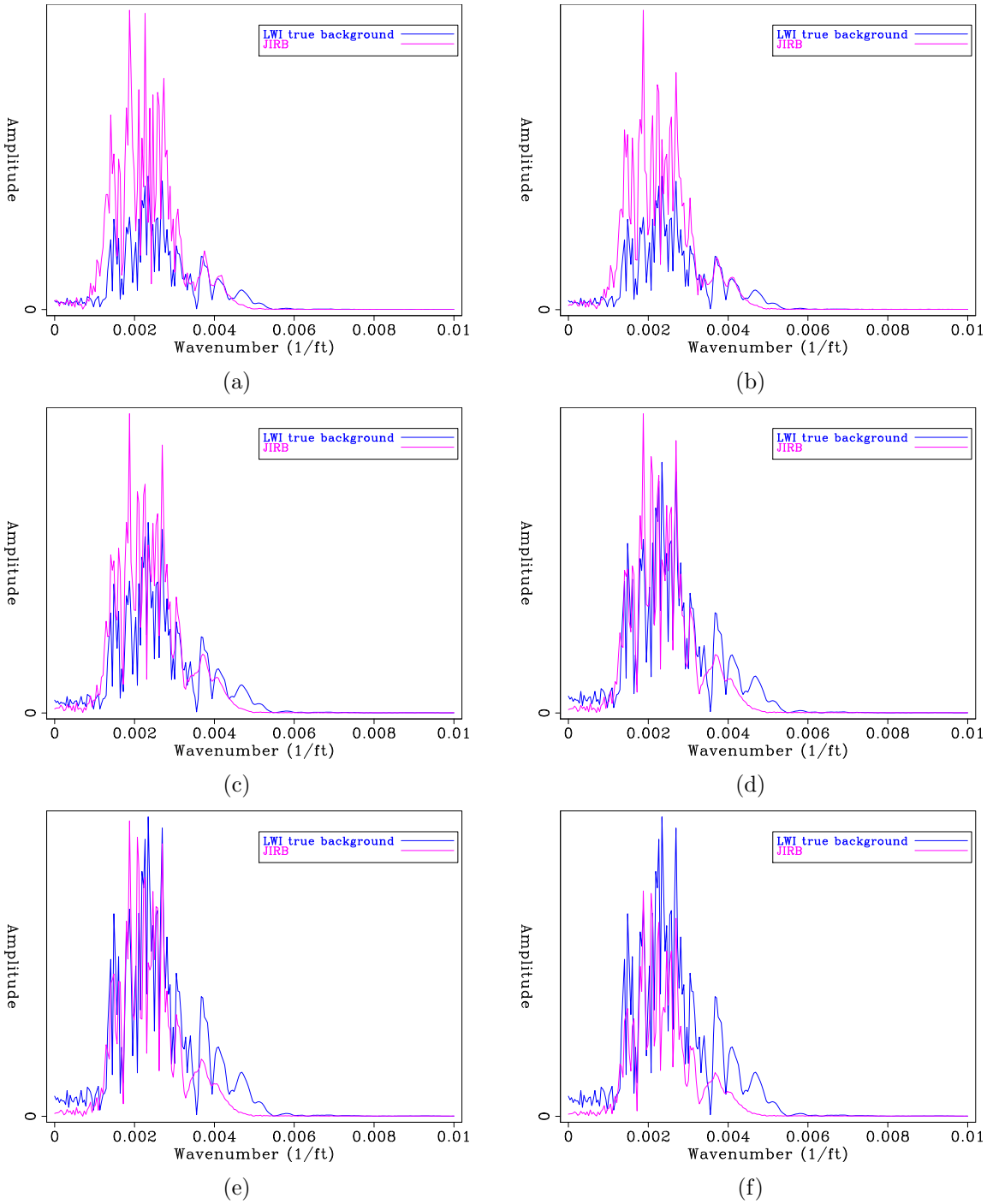
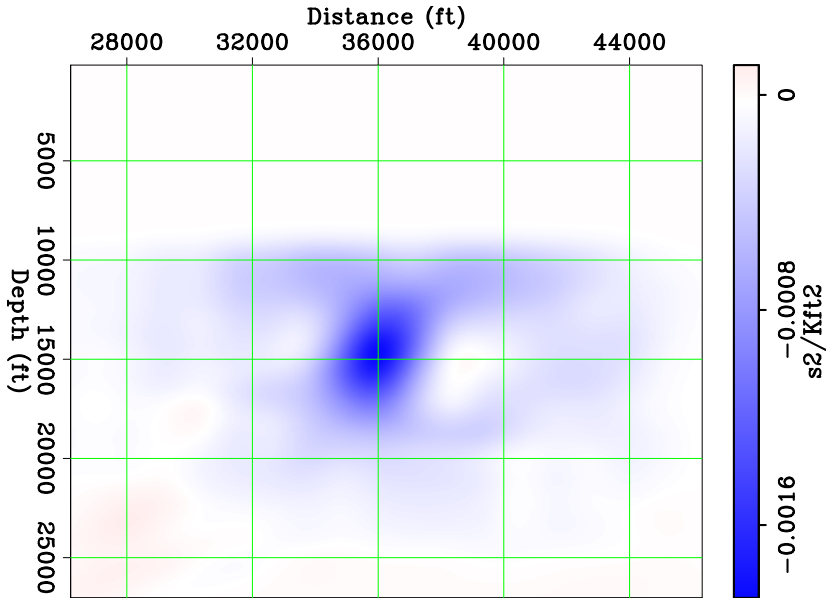


Figure 4.24: Amplitude spectra of the trace comparisons for a) $\lambda = 5$, b) $\lambda = 10$, c) $\lambda = 15$, d) $\lambda = 20$, e) $\lambda = 25$, and f) $\lambda = 30$. [**CR**]

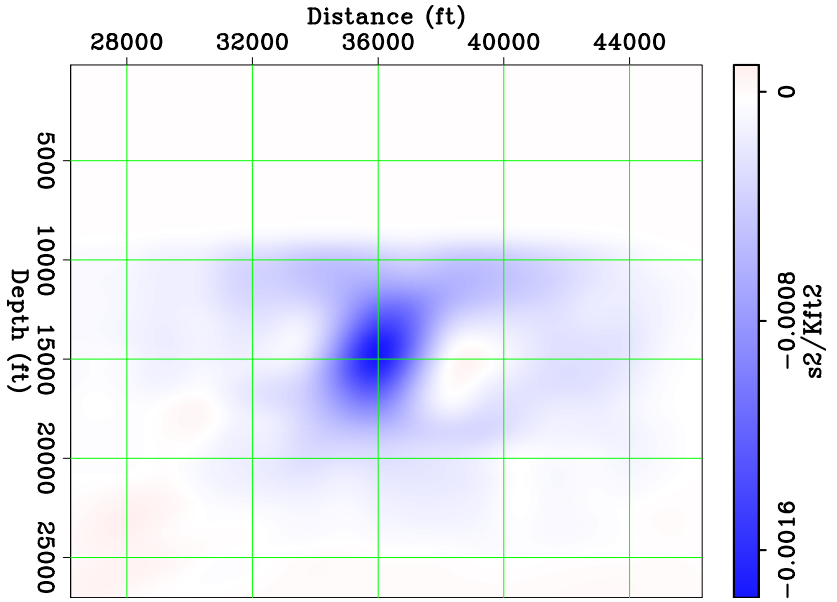
4.2.2 Using the JIRB background model for LWI

Similar to (but less sophisticated than) reflection FWI (e.g. Xu et al., 2012), JIRB incorporates the reflectivity into the inversion. Therefore, we can expect to recover a better estimation of the background model in comparison to WEMVA, where the background model is estimated alone. In this subsection, I performed the corresponding tests. The inverted background models using JIRB and WEMVA are employed to generate the migration images and the PSFs required by LWI in each case. I used the JIRB background model corresponding to $\lambda = 25$. Both JIRB- and WEMVA-based LWI tests ran for 30 iterations.

Figures 4.25(a) and 4.25(b) show the perturbations in the background obtained using JIRB and WEMVA, respectively. The difference between both appears to be negligible, although the comparison of the color scales shows that the JIRB perturbation has higher amplitude. Figures 4.26(a) and 4.26(b) show the reflectivity inversions for WEMVA and JIRB background models, respectively. I plotted the sections using the same color scale. Note how the JIRB result has more focused events, despite the small differences between the background models. The JIRB reflectivity also exhibits higher wavenumber content, as observed in the better resolved weak layers and confirmed by the traces comparison (Figure 4.27). However, noise arises in the form of high-dipping events crossing through the layers. In tests using the JIRB background model for other values of λ (not shown here), there was no significant difference in the estimated reflectivity.

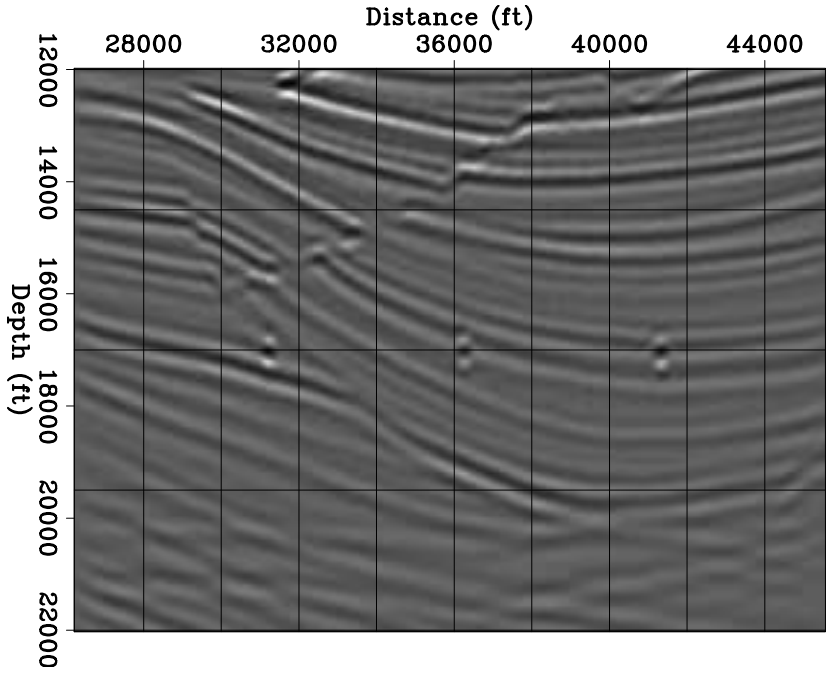


(a)

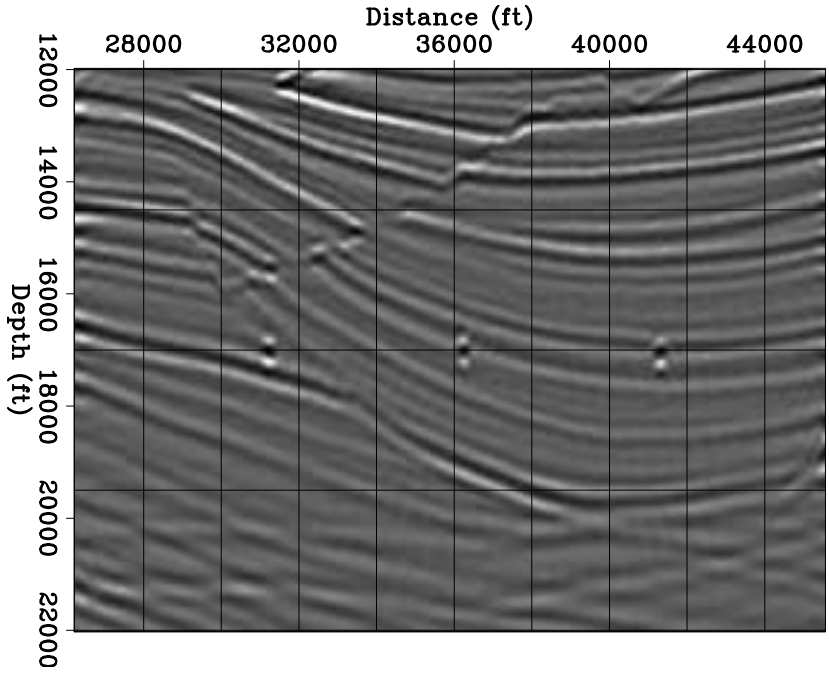


(b)

Figure 4.25: a) Perturbation in the background obtained using JIRB. b) Perturbation in the background obtained using WEMVA. [**CR**]

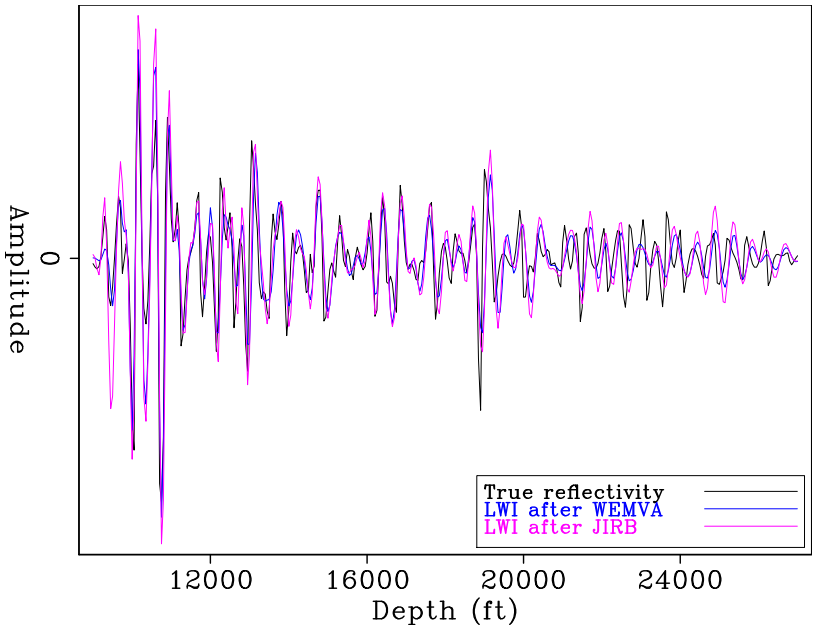


(a)

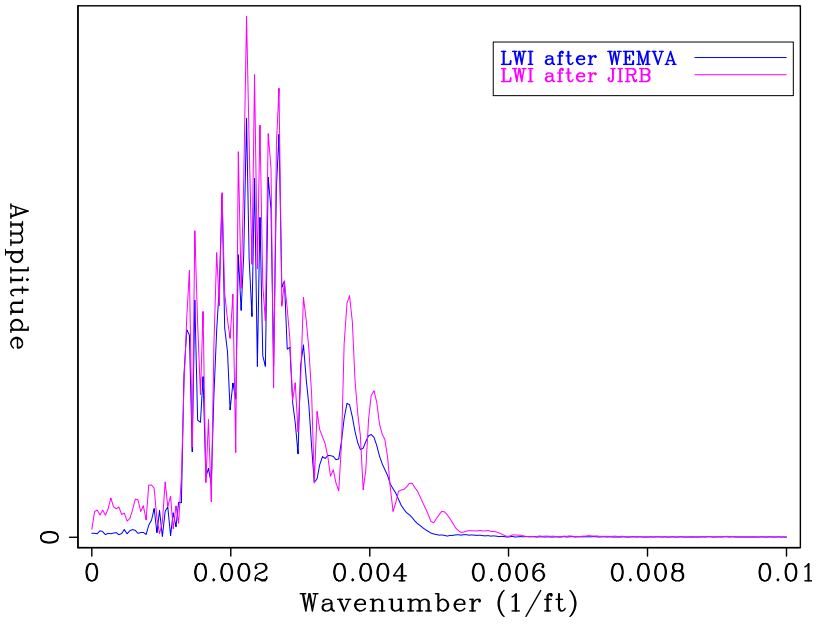


(b)

Figure 4.26: a) LWI reflectivity using the WEMVA background model. b) LWI reflectivity using the JIRB background model. [**CR**]



(a)



(b)

Figure 4.27: a) Trace comparison between LWI obtained with the WEMVA background model, LWI obtained with the JIRB ($\lambda = 25$) background model, and the true reflectivity. b) Spectra comparison of the LWI for WEMVA AND JIRB background models. [**CR**]

4.3 Conclusions

In this chapter, I showed the application of the JIRB method to a 2D synthetic dataset obtained from the sedimentary section of Sigsbee A.

The direct implementation of the joint inversion corrects the inaccuracies in the reflectivity at the same time that the background model is estimated. An adequate choice of the parameter λ is crucial to have the better trade-off between obtaining higher-than-optimal amplitudes and seismic resolution. The main limitation is the possibility of not running enough iterations to obtain the desired resolution, even for small values of λ.

I also tested conventional LWI with the JIRB background model in comparison to the WEMVA background model. There is a small, yet noticeable improvement in crispness and seismic resolution despite the small difference between the JIRB and the WEMVA background models.

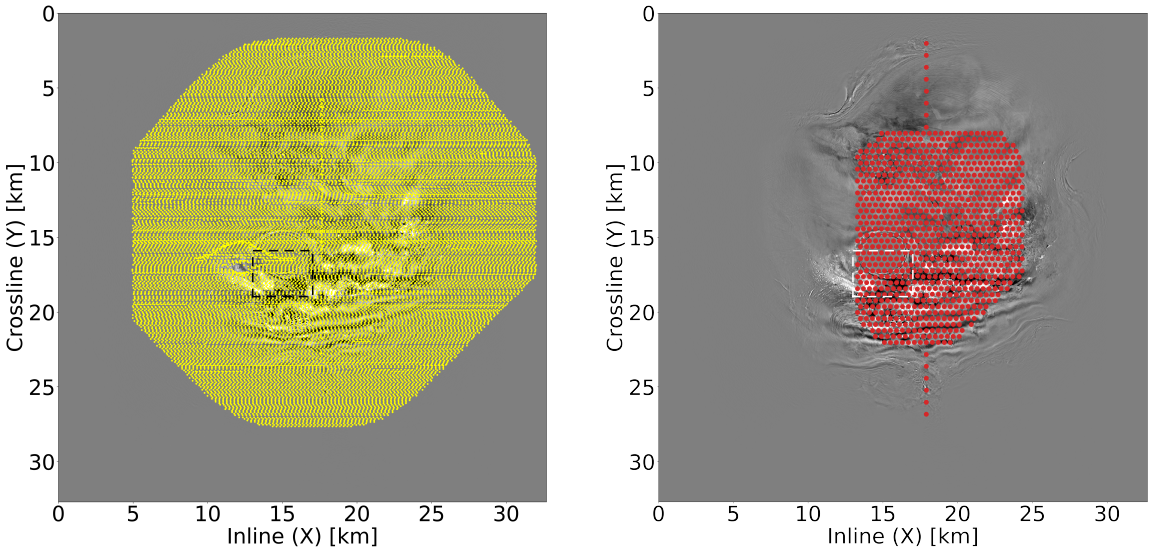
Chapter 5

Application to 3D marine data

In this chapter, I illustrate the JIRB method by presenting an application to an ocean bottom node (OBN) 3D dataset from the Gulf of Mexico, provided by Shell. I followed similar steps as in the synthetic tests of the previous chapter. With the tests shown in the current chapter, I evaluate the plausibility and performance of the JIRB method in real data, free of the inverse crime.

For comparison purposes, I prepared conventional LWI reflectivity. Similar to the previous chapter, I implemented JIRB and analyzed the corresponding background and reflectivity components. I also prepared an LWI test using the JIRB background model. Finally, I further evaluate the performance of JIRB versus WEMVA by utilizing the corresponding background models as input for fine-scale RTM.

I judged the results according to two criteria: 1) image focusing, and 2) enhancement of geological features. The seismic interpretation resorts to the limited available knowledge of the geology of the area. More accurate interpretations require more regional assessment of the geology, and incorporation of borehole data.



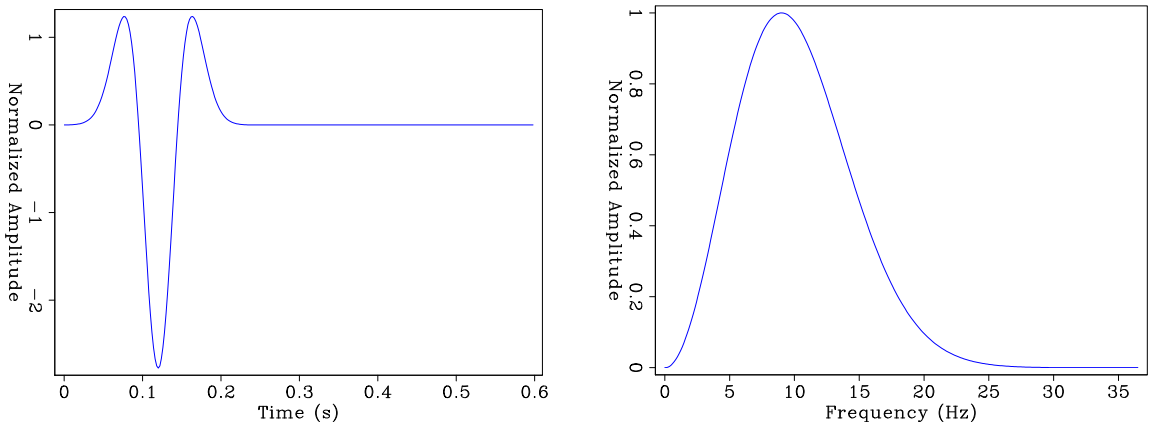
(a) (b)

Figure 5.1: Areal distribution of a) shots, b) receivers (nodes). The small squares indicate the image area for the numerical tests. [**CR**]

5.1 Preliminary steps

5.1.1 Data preparation

Figure 5.1 shows the areal distribution of the shots and the receivers (nodes) for the OBN acquisition campaign. The target consists of sedimentary reservoir rocks pinching out against the wall of a salt diapir, with apparent episodes of passive and active growths (Jackson and Hudec, 2017). Stratigraphic traps are also present, in particular related to deepwater deposits. The squares in both maps encompass the imaging area that I chose for the numerical tests. It contains part of the sedimentary volume and the diapir. There are 226 nodes contributing to the image within this area. The original recording time is 14 s sampled at 2 ms. For the numerical tests shown in this chapter, I only employed the first 8 s.



(a) (b)

Figure 5.2: a) Source wavelet and b) the corresponding amplitude spectrum. [**ER**]

Dahlke (2019) performed the PZ-summation to separate the upgoing and the downgoing wavefield components. This procedure consists of the adaptive addition and subtraction of the geophone vertical component and the hydrophone component of the data (Soubaras, 1996; Pacal et al., 2015). Dahlke (2019) also reshaped the wavelet and removed the bubble effect. I utilized these pre-processed data for the numerical tests shown in this chapter. Although the upgoing component produces sharper reflectivity contrasts, I chose the downgoing component because it has better coverage of the subsurface, and therefore, less tendency to produce aliasing artifacts in the images (Pacal et al., 2015). I employed the standard mirror technique (Grion et al., 2007; Pacal et al., 2015) for the imaging processes. I also resorted to the principle of reciprocity (Knopoff and Gangi, 1959), injecting the source wavelet at the nodes' positions, whereas I injected the data traces at the source positions in the corresponding common-receiver gathers (CRG).

I binned the downgoing component data (which I henceforth refer to simply as "the data") to obtain a regular grid of 25×25 m, which is suitable for imaging frequencies up to 30 Hz. For simplicity, I constructed the binned CRGs spanning the whole study area. I employed a 10.5 Hz dominant frequency Ricker wavelet as a source function, bandpassed to a maximum frequency of 30 Hz. Figure 5.2 shows the normalized wavelet and the corresponding amplitude spectrum.

5.1.2 Reverse-time migration

I prepared conventional RTM volumes to have the first view of the subsurface in the study area with my computational code [RTM images of the area are also available in Dahlke (2019)]. It allows the evaluation of my imaging toolbox on this OBN dataset. The RTM image also constitutes the input data for model-space LWI and JIRB.

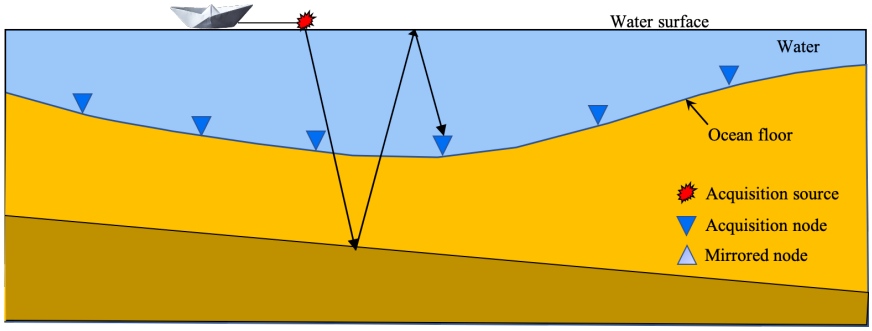
I prepared the velocity model to image using the downgoing component, as shown in Figure 5.3. First, I numerically expanded the water layer to enact the mirror imaging strategy. Then, I defined a datum corresponding to a numerical ocean floor mirrored about the water surface, and where the mirrored nodes will lie. Finally, as aforementioned, using reciprocity, I inject the source wavelet at the position of such mirrored nodes and the data at the source positions.

The migration aperture is 1250m (50 samples) in both x and y directions. For the random boundary condition (Clapp, 2009) I surrounded the volume with a halo of 50 samples. Like the synthetic data in Chapter 4, I included an extra layer of 50 samples on the top boundary to avoid placing the "sources" (which are the nodes acting as such, for reciprocity) upon the random layer's border, thereby preventing numerical artifacts. Figure 5.4 shows one velocity volume prepared for the migration of a single CRG. The lines indicate planes whereby the inline, the crossline, and the depth slice sections project to the lateral sides and top of the volume. Note that the randomness is mild at the volumes' interior, and increases toward the computational boundaries. I pre-computed 226 velocity volumes for different realizations of the engine code to generate different random halos. This strategy allows one to cancel out the incoherent random artifacts while reinforcing the subsurface signal in the image during stacking.

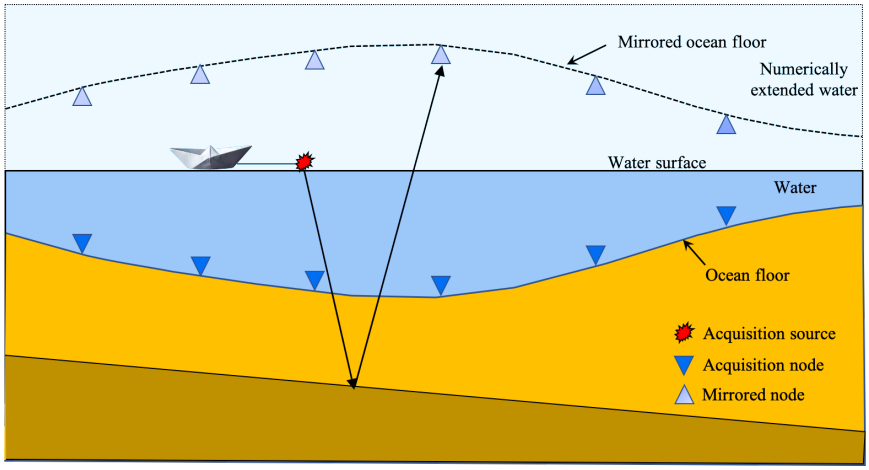
Figure 5.6(a) shows inline, crossline, and depth slice sections of the RTM volume. I applied a Laplacian filter to attenuate the low-wavenumber artifacts. The reflectors are almost flat except in the vicinity of the salt intrusion. The sedimentary units are tilted by the diapir growth, which is still deforming the ocean floor. I interpreted the presence of a possible channel complex at the indicated position. The presence of

this geological element can be expected in turbidite systems (e.g. Huang, 2018). See Figure 5.5. The diapir's walls are virtually vertical and are barely imaged with the chosen aperture. In other tests (not shown), I recovered such a diapir wall by using a larger migration aperture, but at the expense of introducing migration "smiles" that contaminated the reflectors. The purpose of the inversion tests is to properly image the sedimentary structures, rather than imaging the salt flanks. Hence, I keep the migration aperture small.

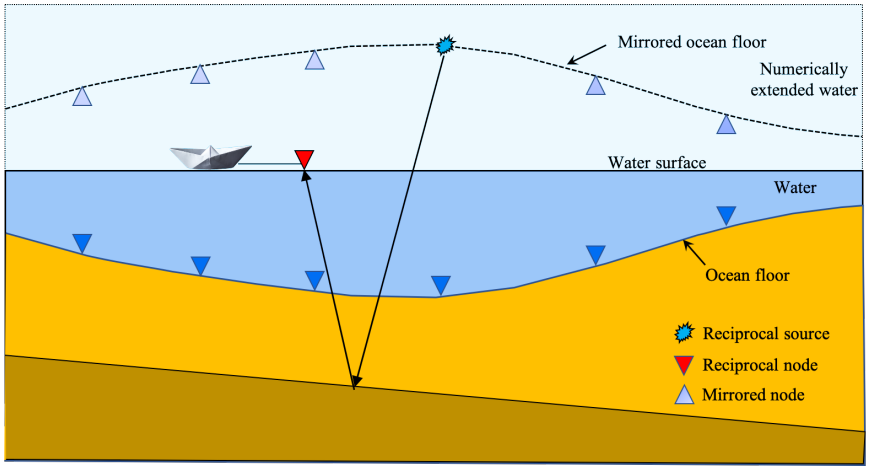
For the discussion on the image differences and improvements, I mainly focused on sedimentary features, such as those indicated by arrows in Figure 5.6(b), which probably correspond to a confined channel complex (e.g. Huang, 2018).



(a)

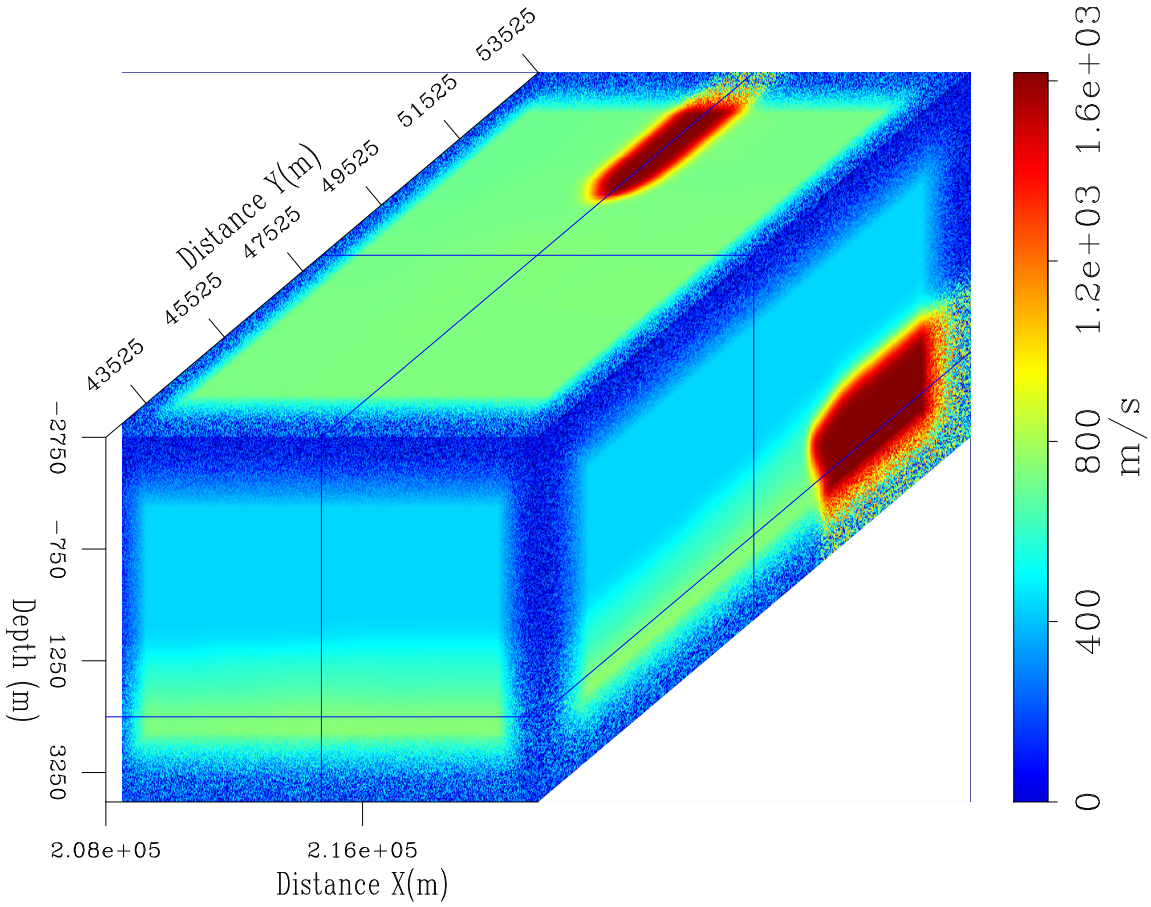


(b)



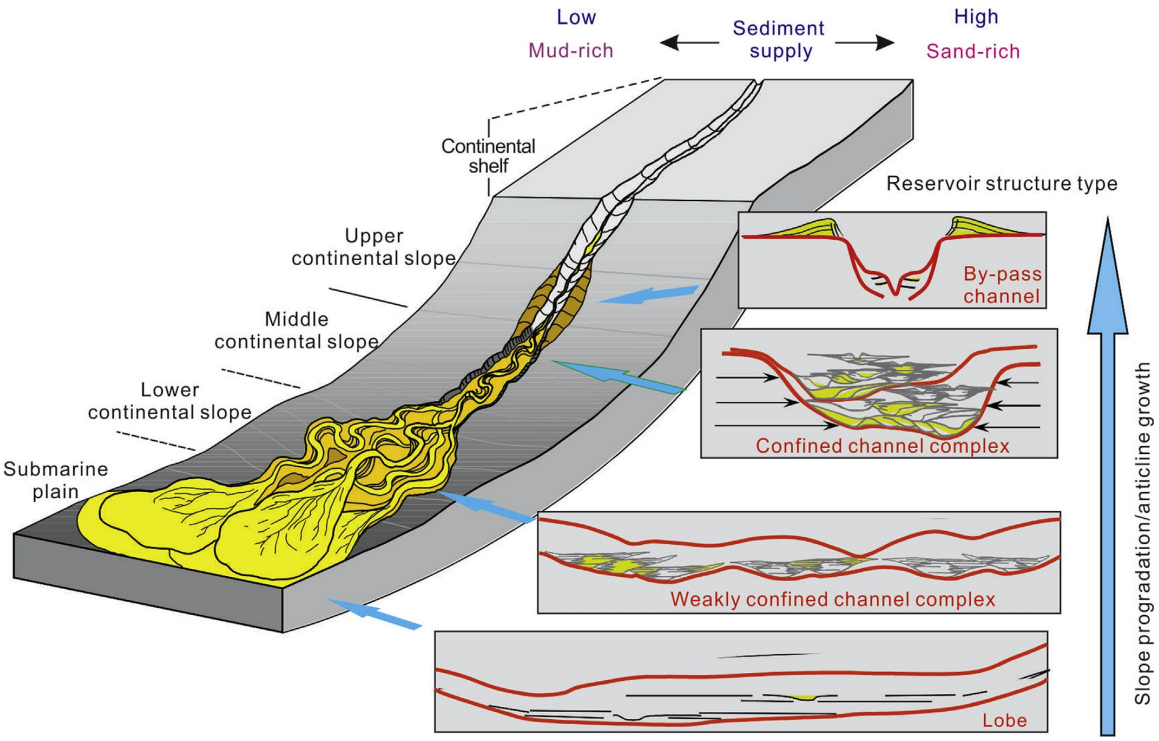
(c)

Figure 5.3: a) Ocean-bottom node acquisition cartoon, illustrating the water-surface multiple trajectory for a source-node pair. b) Numerical extension of the water velocity and re-datuming of the nodes, showing the equivalent source-node trajectory. c) After the application of the reciprocity principle to a source-node pair. [**NR**]



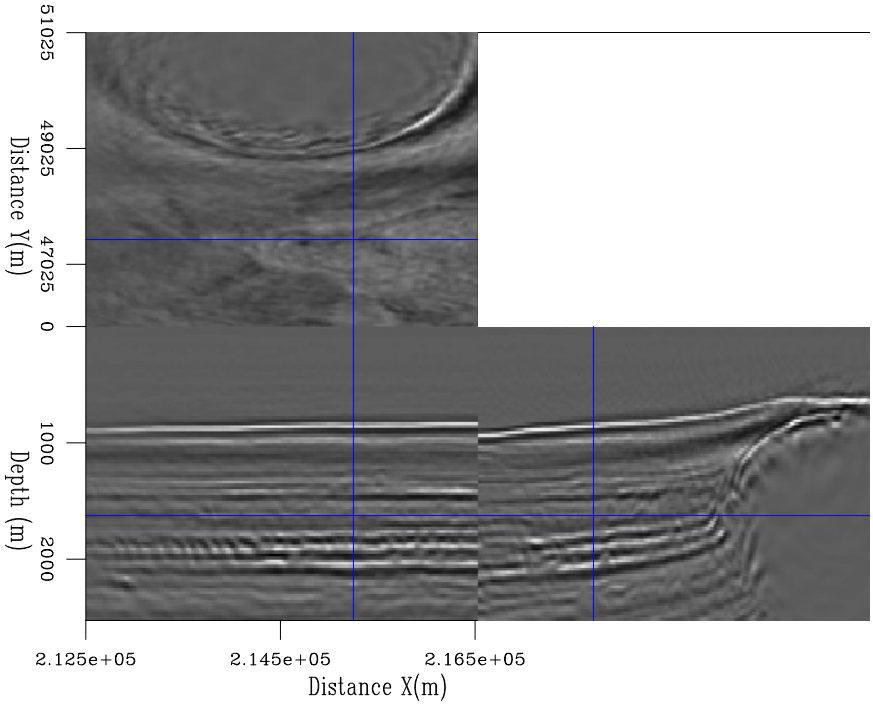
(a)

Figure 5.4: Random boundary frame for a single common-receiver gather. [**CR**]

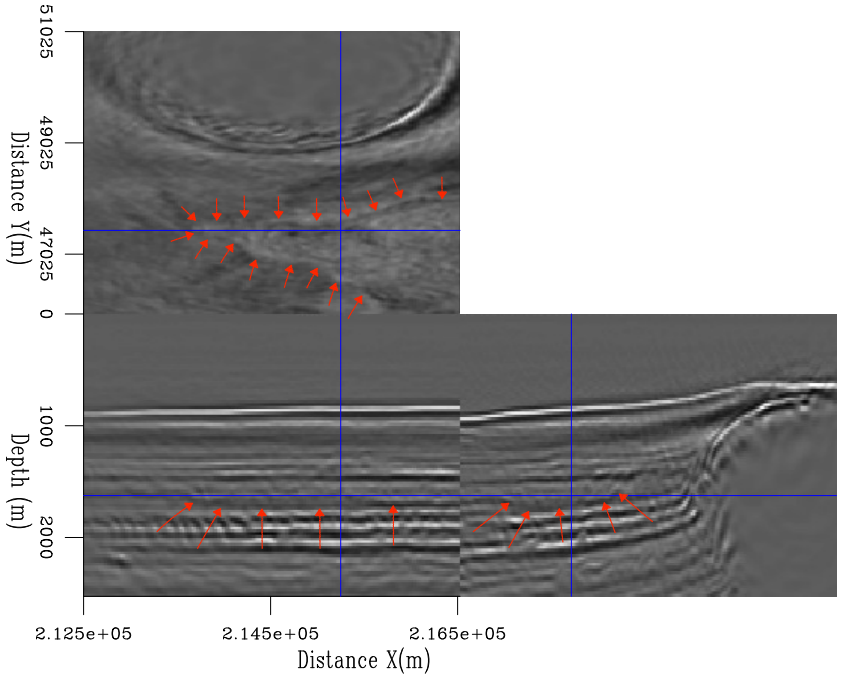


(a)

Figure 5.5: Illustration of turbidite system. Taken from Huang (2018). [**NR**]



(a)

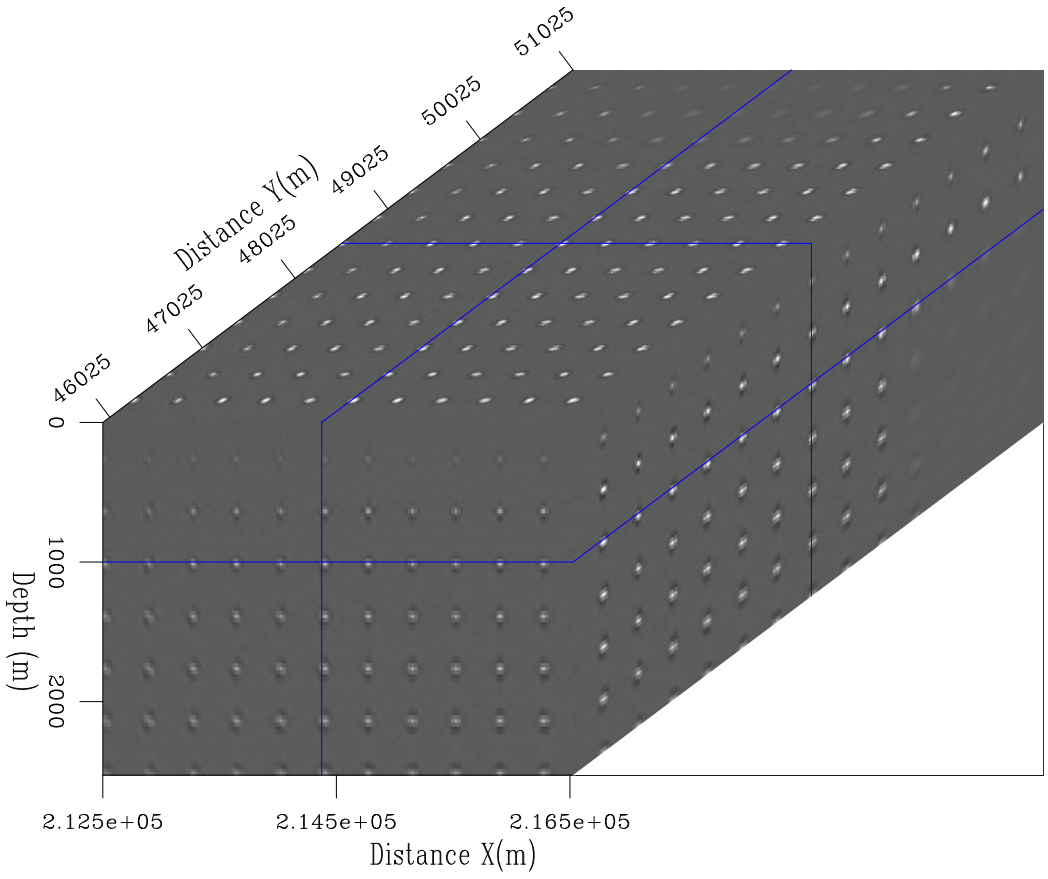


(b)

Figure 5.6: a) RTM image of the OBN data. b) Same as a), with arrows indicating the base and the lateral bounds of a possible channel complex. [**CR**]

5.1.3 Point-spread functions

Like the synthetic test of the previous chapter, I employed PSFs to estimate the Gauss-Newton Hessian for the 3D data tests. I seeded spikes every 15 gridpoint and cascaded Born modeling followed by RTM to compute such PSFs, shown in Figure 5.7. From this volume, the PSFs are read and interpolated "on the fly" to estimate Hessian terms. Notice that at the salt diapir's position, the energy of the PSFs becomes more smeared and faint than in the sedimentary part. Although Figure 5.7 shows the distribution of PSFs only in the model space, they span all across the computational space (i.e., including the boundaries halo).



(a)

Figure 5.7: PSF distribution in the image space. [**CR**]

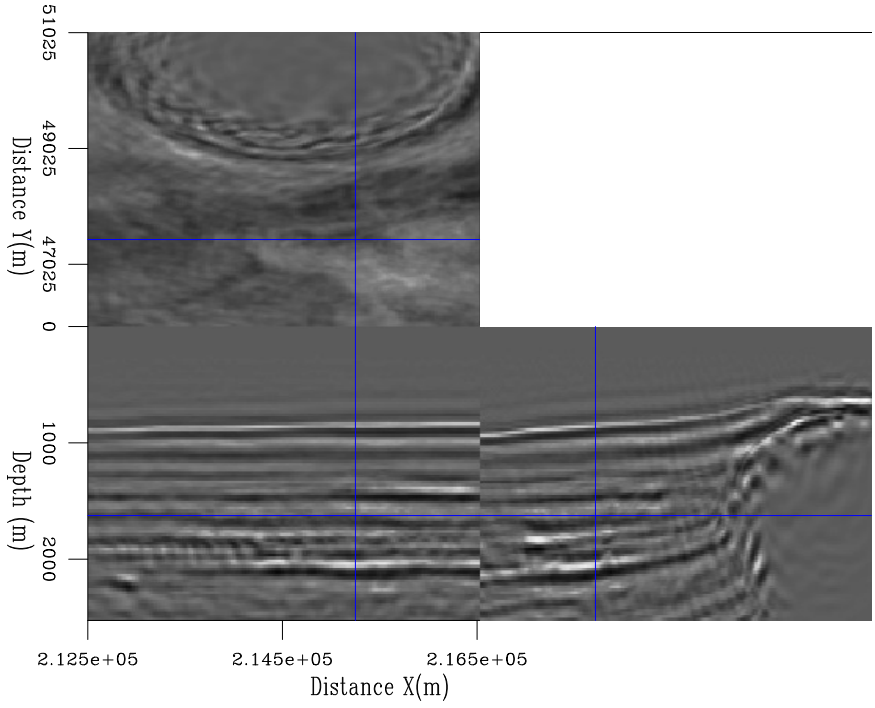
5.2 Inversion results

5.2.1 Reflectivity comparison

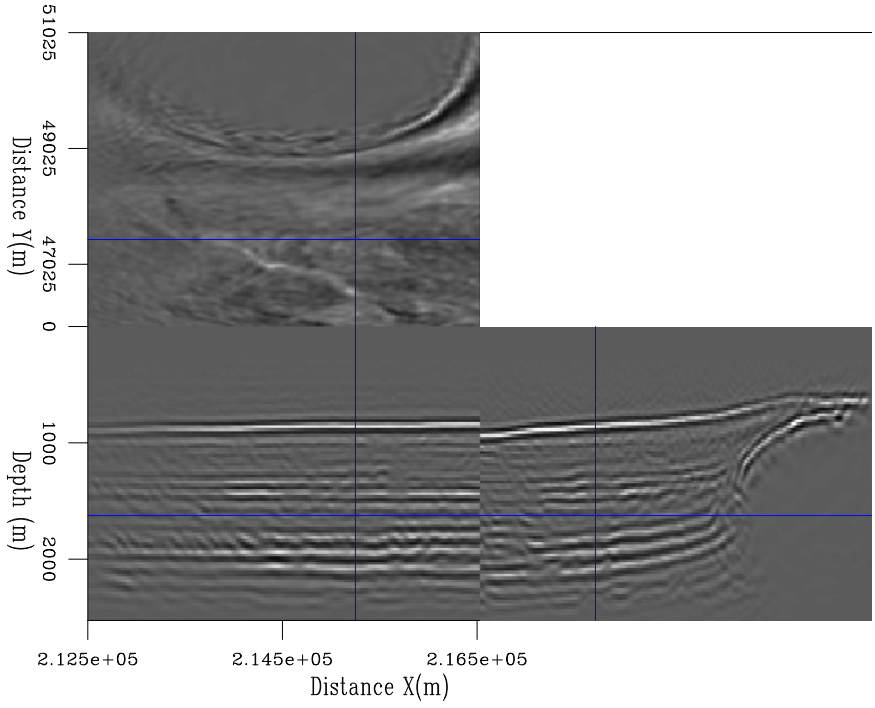
In this section, I compare the result of conventional LWI and JIRB for the reflectivity estimation. It is homologous to the numerical test in section 4.2.1.

Figure 5.8(a) shows the result of conventional LWI in model space, i.e., the result of solving equation (1.2). The background model is the original provided by Shell and is not updated during the LWI process. The Gauss-Newton Hessian was approximated by an "on the fly" interpolation of the PSFs shown in Figure 5.7. The inversion ran for 20 iterations, which was enough to flatten the objective function. The inversion recovered some high wavenumber components. It also introduced vertical artifacts in the estimated reflectivity image, even destroying the top of the salt diapir.

Figure 5.8(b) shows the result of JIRB reflectivity after 10 iterations, i.e., the result of solving equation (4.1). Notice that the events of the JIRB image are more focused. Such is a consequence of updating the background component at the same time as the reflectivity. However, this result is different from merely an improved version of the LWI result (Figure 5.8(a)). Logically, there is a more complex interaction between the model parameters (the reflectivity and background) than what I expected when I initially formulated the method as a linear problem.



(a)



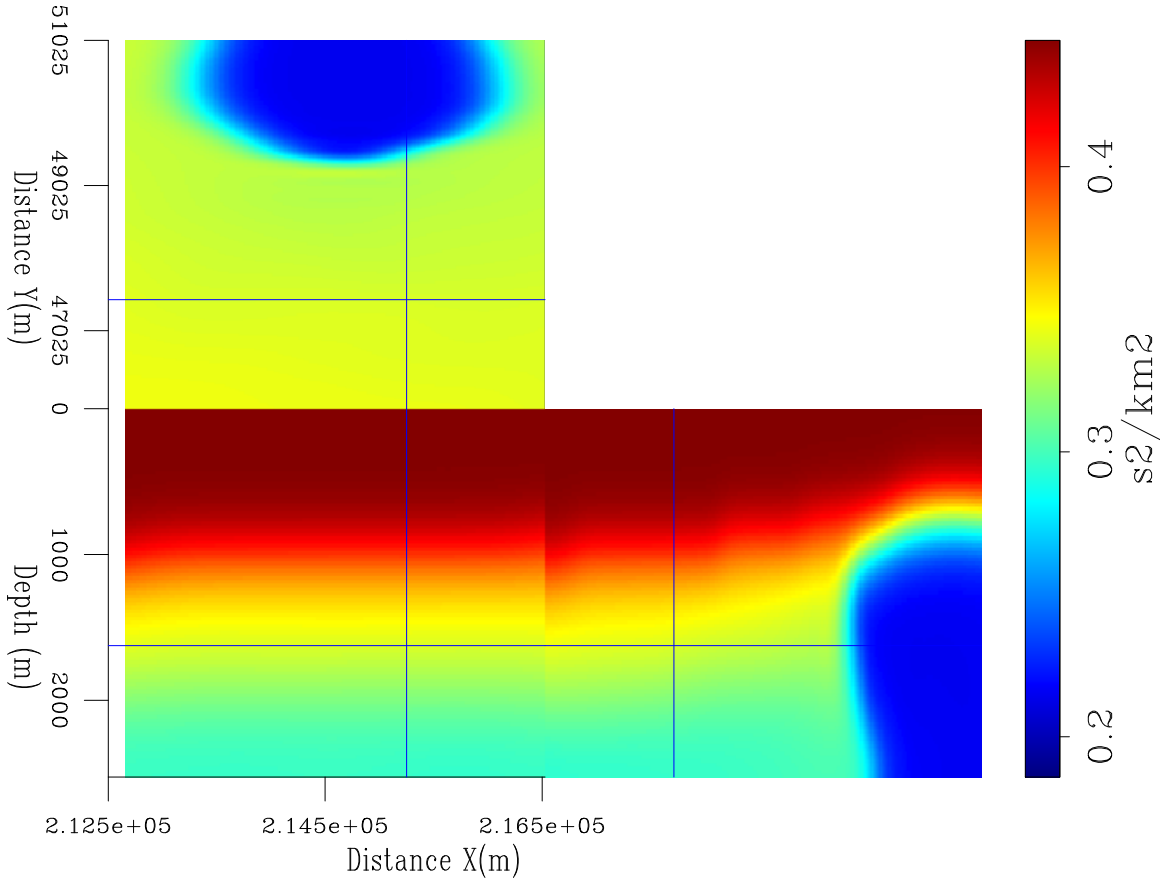
(b)

Figure 5.8: a) Reflectivity obtained with conventional LWI. b) Reflectivity component of JIRB. [**CR**]

5.2.2 Background component comparison

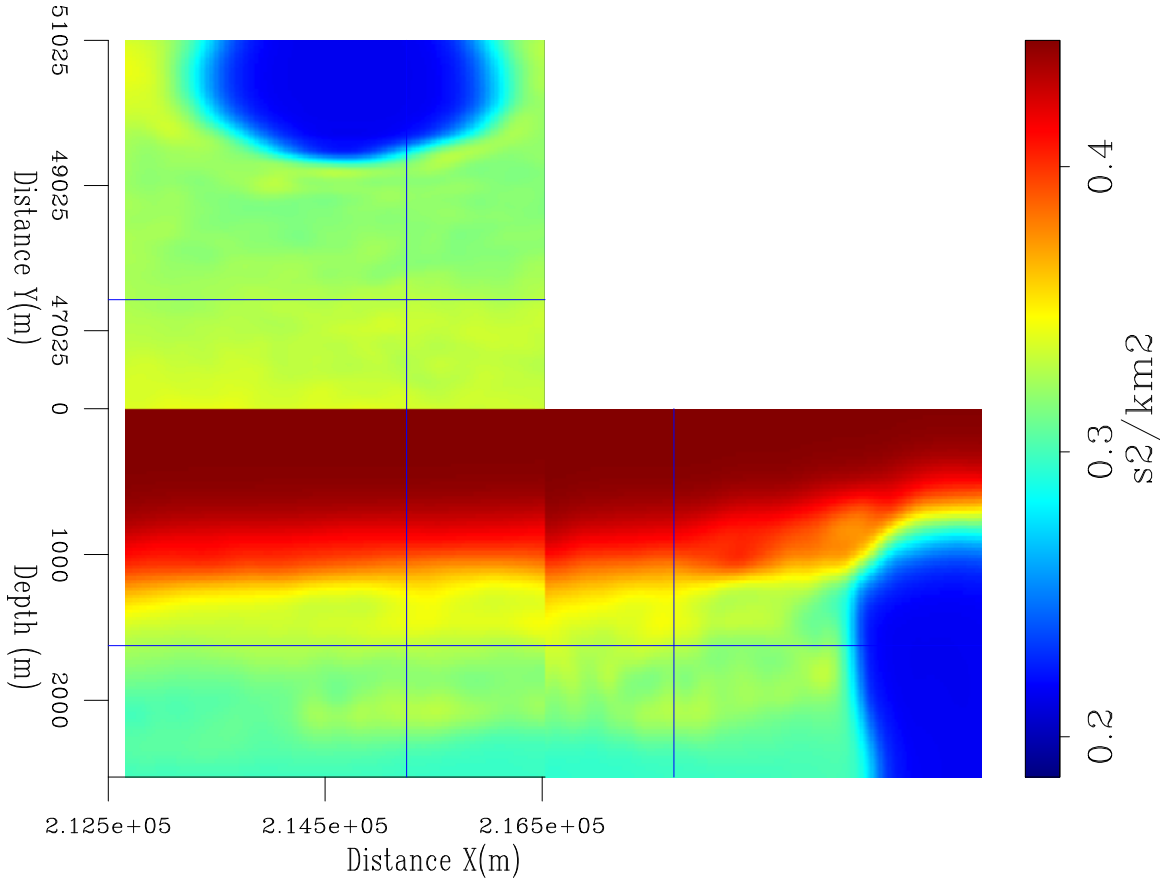
Background component inversions

Figures 5.9, 5.10, and 5.11 show the initial or original background model, the inverted background model using WEMVA, and the inverted background model using JIRB for $\lambda = 8$, respectively. During the inversions, I applied a mask to the gradient to prevent updating the squared slowness of the salt body and the water layer (see equation (B.3)). Notice that the initial model appears to be laterally smooth and relatively featureless in the sedimentary part, whereas the updated models exhibit vertical and lateral variations. This is better observed in the difference between the inverted and the initial background models in Figures 5.12 and 5.13. The additional detail obtained with the inversions should contribute to focus the image better. We observe that improvement in the estimated JIRB reflectivity image (Figure 5.8(b)). On the other hand, geologically speaking, it makes perfect sense to expect such heterogeneity in a deepwater depositional environment. In both inversions, JIRB and WEMVA, I preconditioned the background model using B-splines (see Appendix B), which helped prevent the occurrence of spurious, high wavenumber, artifacts.



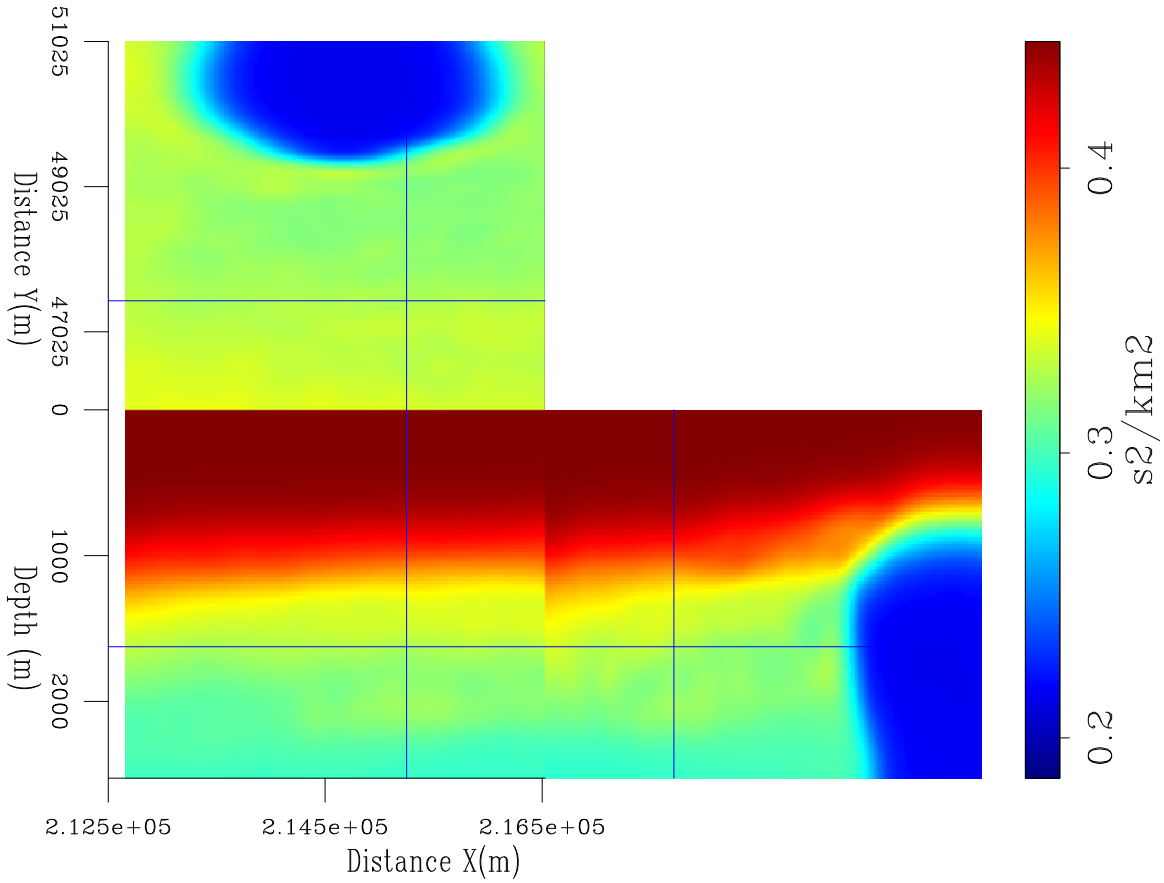
(a)

Figure 5.9: Original background model. ER



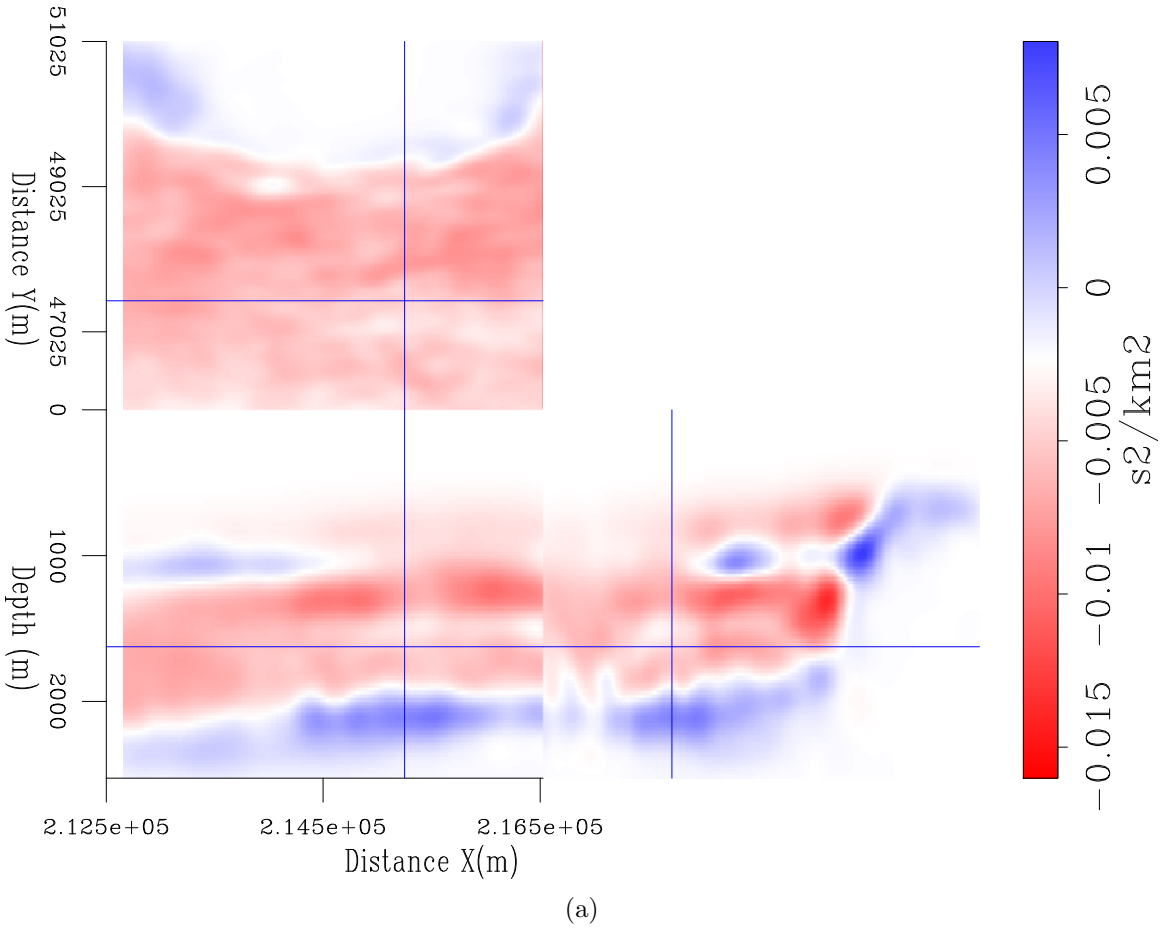
(a)

Figure 5.10: Background model obtained with WEMVA. ER



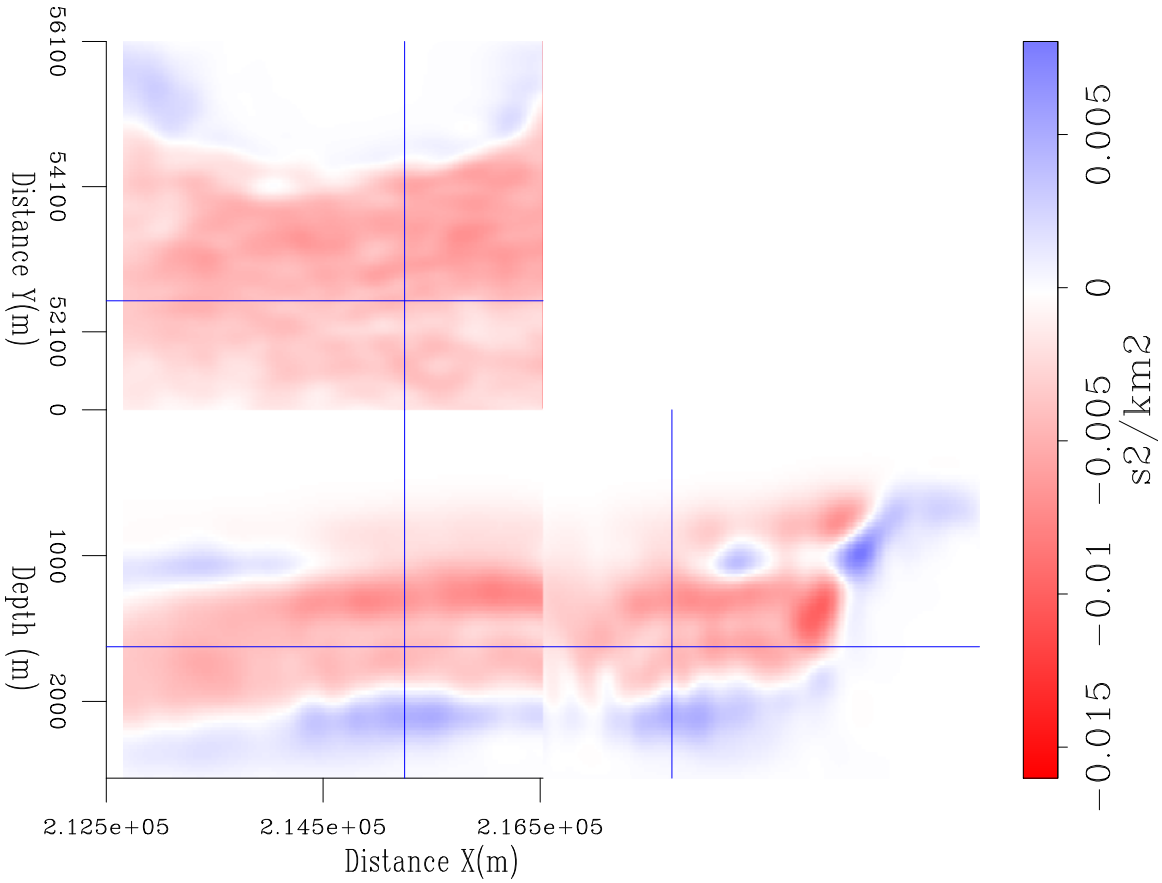
(a)

Figure 5.11: Background model obtained with JIRB. [**CR**]



(a)

Figure 5.12: Difference between WEMVA and initial background models. [**CR**]



(a)

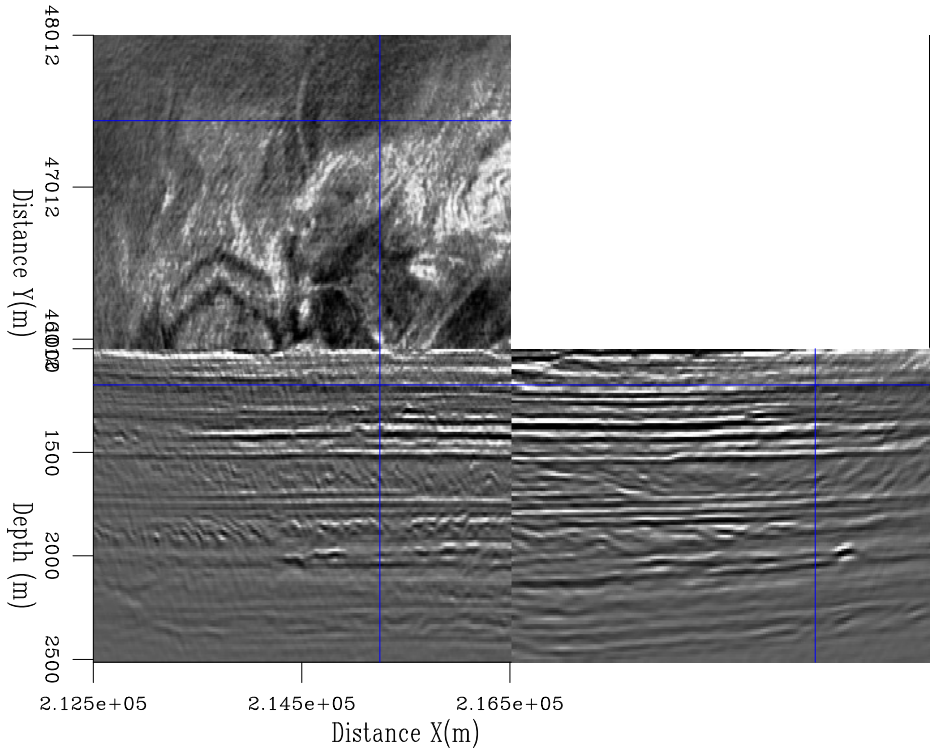
Figure 5.13: Difference between JIRB and initial background models. [**CR**]

Refined RTM images

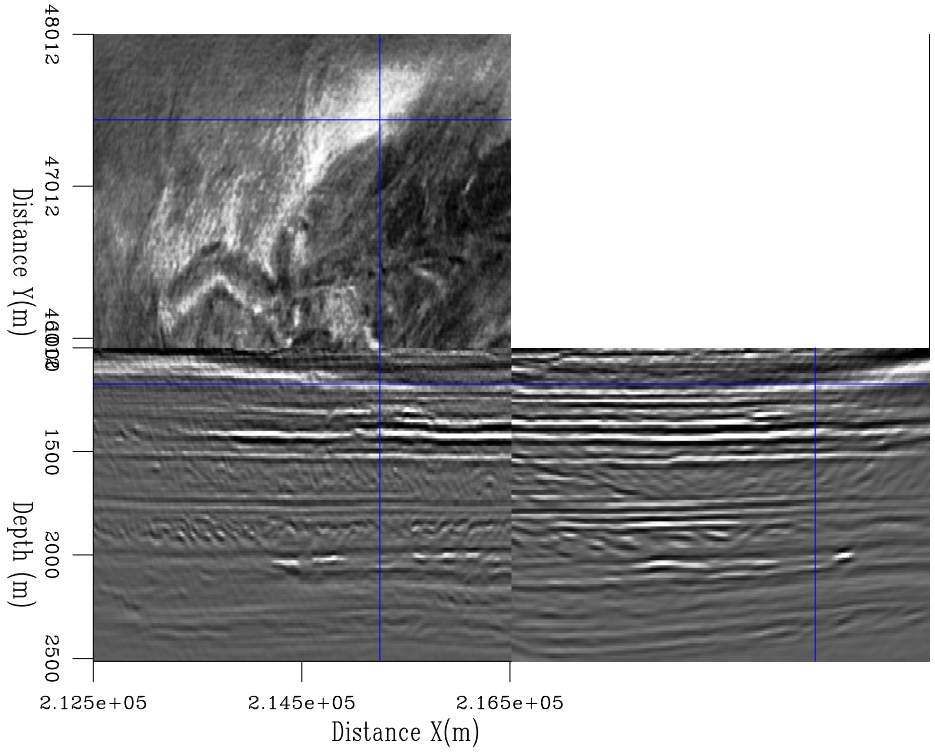
It remains in a quandary which one of the inversions achieved better result, WEMVA, or JIRB. It is very difficult to make a sensible judgement based solely on the results of Figures 5.9 to 5.13. One alternative is to proceed similarly as in Chapter 4: performing LWI using the corresponding inverted background models. However, for this 3D case, I performed refined RTM tests using the inverted models, reducing the grid size to $12.5 \times 12.5 \times 12.5$ m, i.e., half the original grid size and thereby duplicating the wavenumber of the migration image. This refinement allows for better imaging of the stratigraphic features that I want to enhance. The drawback of such refinement is the increase in computational time and storage by a factor of 8 or 16 in 3D.

Figure 5.14(a) shows the refined RTM image obtained with the original background model (Figure **??**). I zoomed into the sedimentary section to better appreciate the stratigraphy, excluding the water layer and the salt diapir. I show the same inline and crossline that I showed before in the previous results, but a shallower depth slice (approximately 1100 m). The latter exhibits channel signatures, which are very common in deep water environments. For comparison, Figure 5.14(b) shows the RTM refined image using the WEMVA inverted background model (Figure **??**). Notice how the stratigraphic features of the inline and crossline sections become more focused in the WEMVA volume, e.g., the channel features immediately below the reflector at 2000 m. However, the comparison is not valid for the depth slice because the velocity change introduced by the inversion significantly shifted the seismic events upward. Thus, while in the depth slice of Figure 5.14(a) we interpret channels, in the depth slice of 5.14(b) we can interpret other geological elements, a fan constituted by the white feature at the top of the depth slice, fed by a channel, and either crevasse splays or minor fan lobes originating from the same channel.

In the following, I compare the WEMVA- and the JIRB-based RTM images. Their differences are much more subtle than those with the RTM for the original background model. I highlighted some of such differences; others are difficult to observe without toggling the images.



(a)



(b)

Figure 5.14: Refined RTM volume for a) Original model, b) WEMVA model. [**CR**]

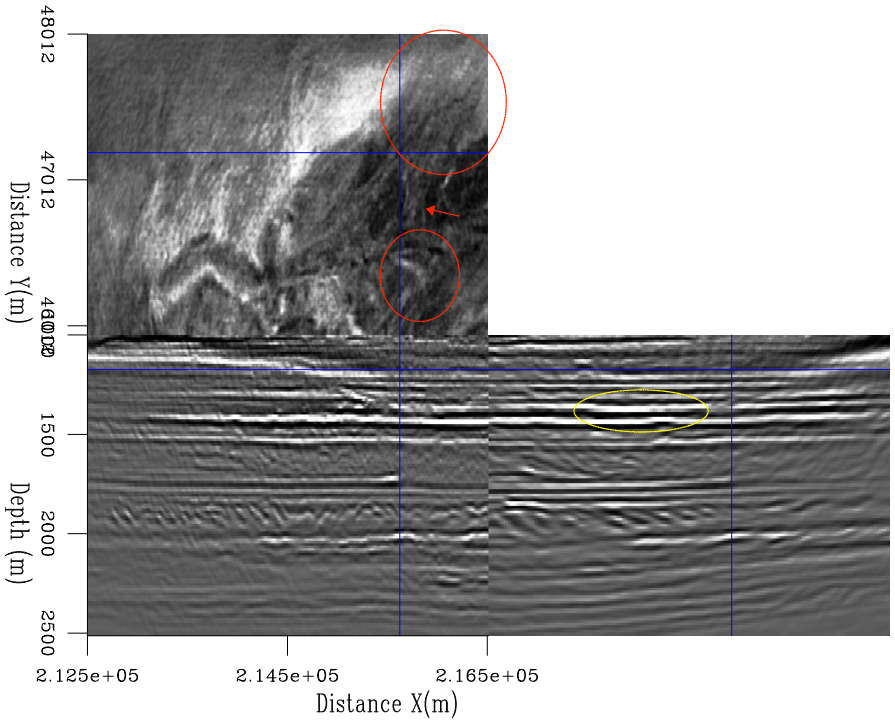
Figures 5.15(a) and 5.15(b) show the refined RTM image obtained with the WEMVA and the JIRB background models, respectively (Figures **??** and **??**). The depth slice shows improvement in a part of the fan and one of the channels, as indicated by the red ellipses and the red arrow. In the sections, the yellow ellipses indicate differences in the reflectors that suggest more focusing. There are also vertical adjustments of the seismic events due to the velocity discrepancy between the background models.

Figures 5.16(a) and 5.16(b) show the comparison for a deeper level (around 1300 m). The left half of the depth slice is featureless, possibly an abyssal plain. The right half shows a straight channel that bifurcates into two fans with a dendritic pattern. The rightmost feature appears to be another fan lobe, with internal channelization. The main improvements in the JIRB volume that I can identify are the better definition of the channels inside the first lobe (red oval) and the sharper border of the second one, both indicated by red ellipses. On the contrary, the straight channel, indicated by the red arrows, is better defined in the WEMVA volume. In the section, I highlighted some channel features that appear to be better focused when using JIRB.

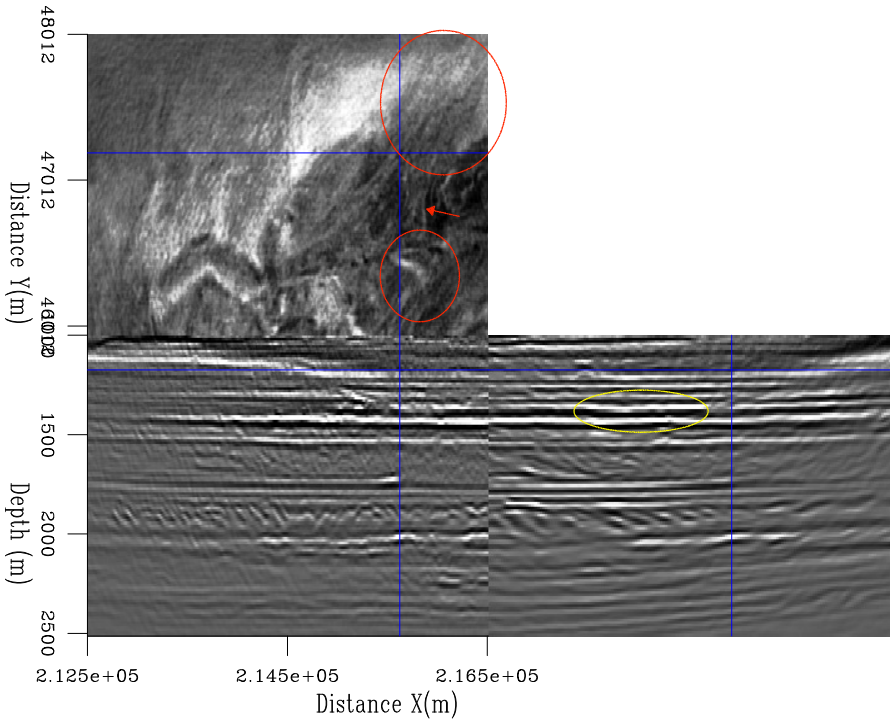
Figure 5.17 shows a deeper level, around 1800m, where some of the changes are indicated in red. It corresponds to the top of a stratigraphic unit between 1800 and 2000m that I interpret as a possible mass-transport complex that spans beyond the limits of the study area. In the vertical sections, we observe shingle geometries (short reflector units beginning in toplaps and ending in downlaps), and possible channels and fans deposited on top. Although there are features in the sections that noticeably become more focused in the JIRB volume (indicated by the yellow ellipses), the improvements are much more difficult to observe in the depth slice.

Figures 5.18(a) and 5.18(b) show a channelized level at 2000 m. The channels' signatures are distinguishable in the sections. In the depth slice, we observe that they constitute part of the channel complex in white upon the dark gray plain, although some other, grayish, channels, are visible at the lower part of the depth slice. The red ellipses indicate features that are better observed in the JIRB volume. Most of them are obscured in the WEMVA volume because the corresponding background model

was either inaccurately focus them, and/or because they are slightly shifted with respect to the depth slice. The correction in the JIRB volume brought them back to the stratigraphic level. Additional features that are slightly corrected by JIRB are indicated in the sections by the yellow ellipses and arrow.

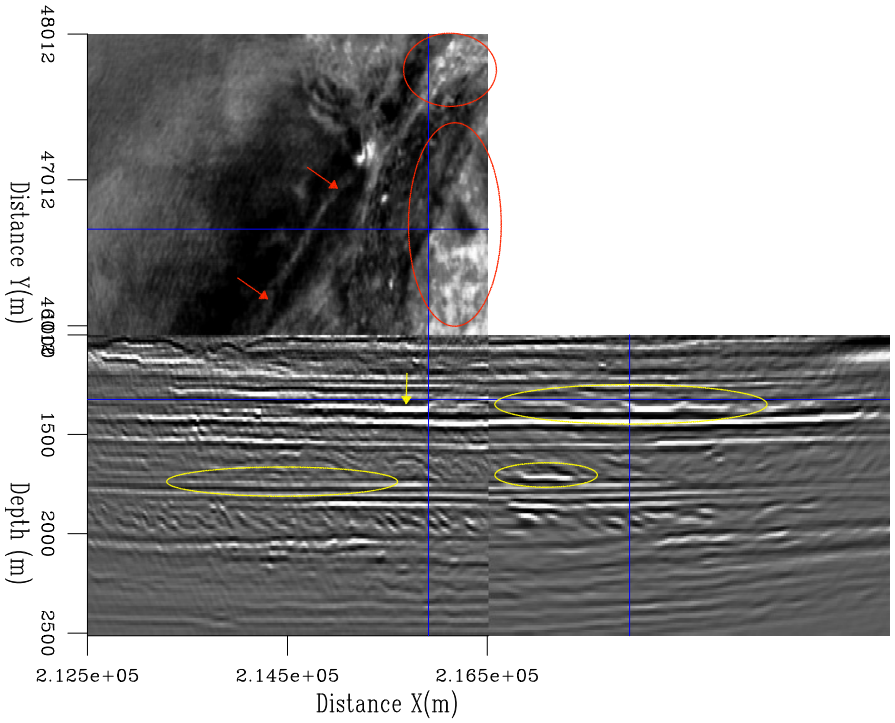


(a)

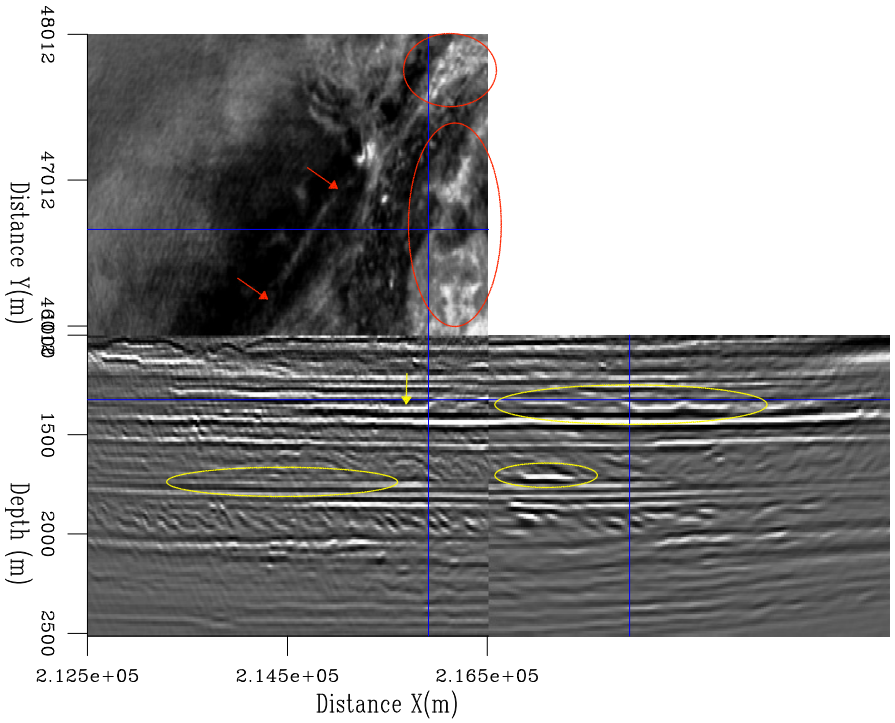


(b)

Figure 5.15: Refined RTM volume for a) WEMVA model, b) JIRB model. [**CR**]

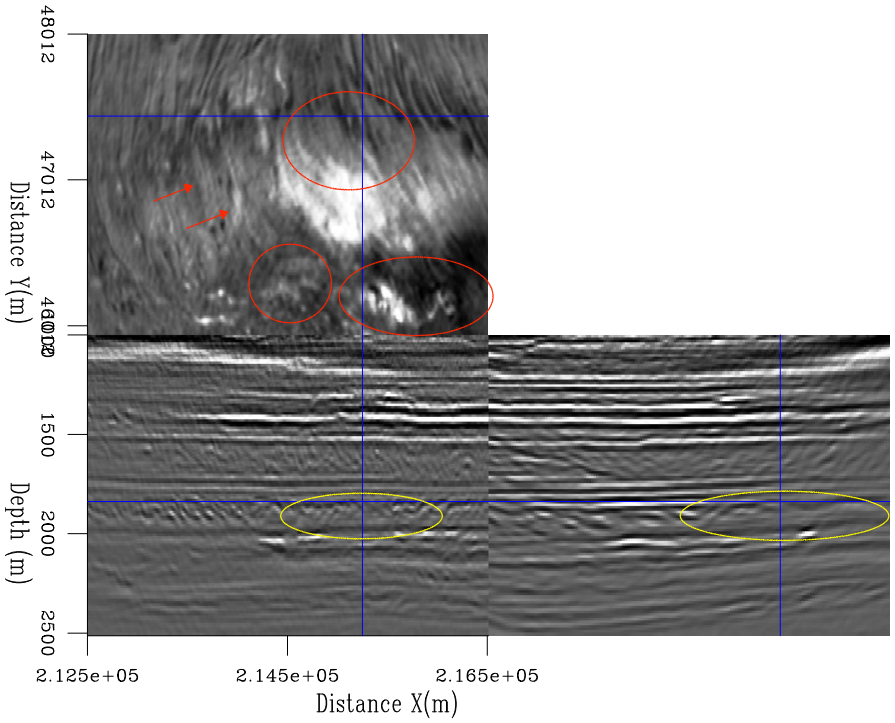


(a)

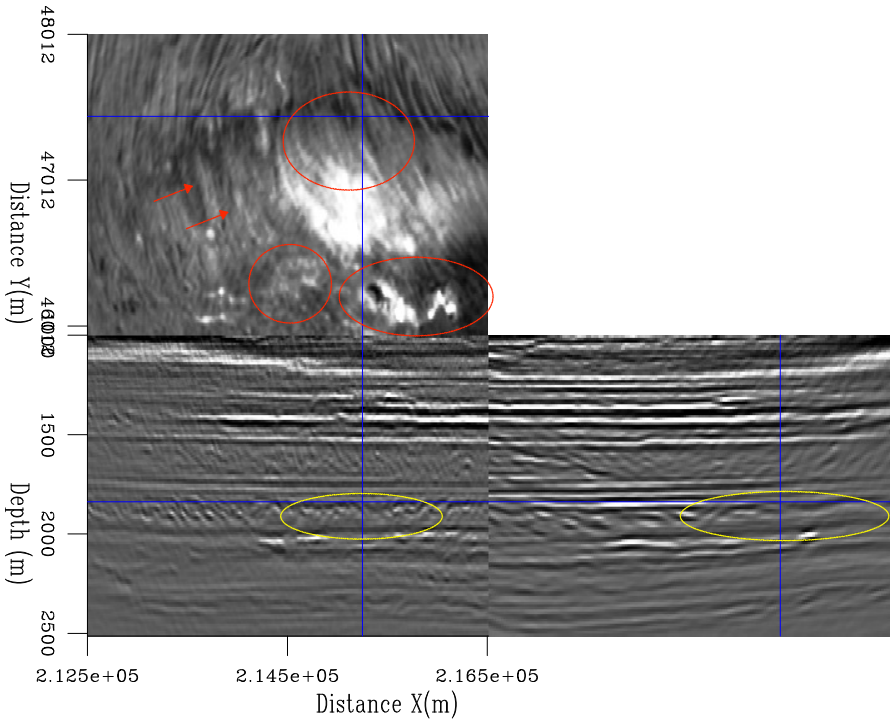


(b)

Figure 5.16: Refined RTM volume for a) WEMVA model, b) JIRB model. [**CR**]

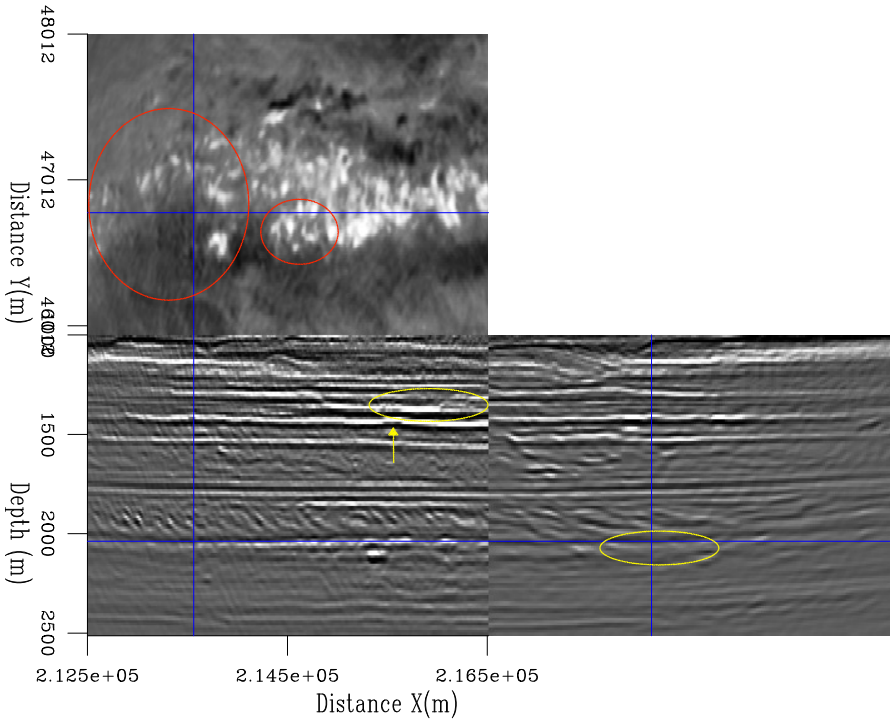


(a)

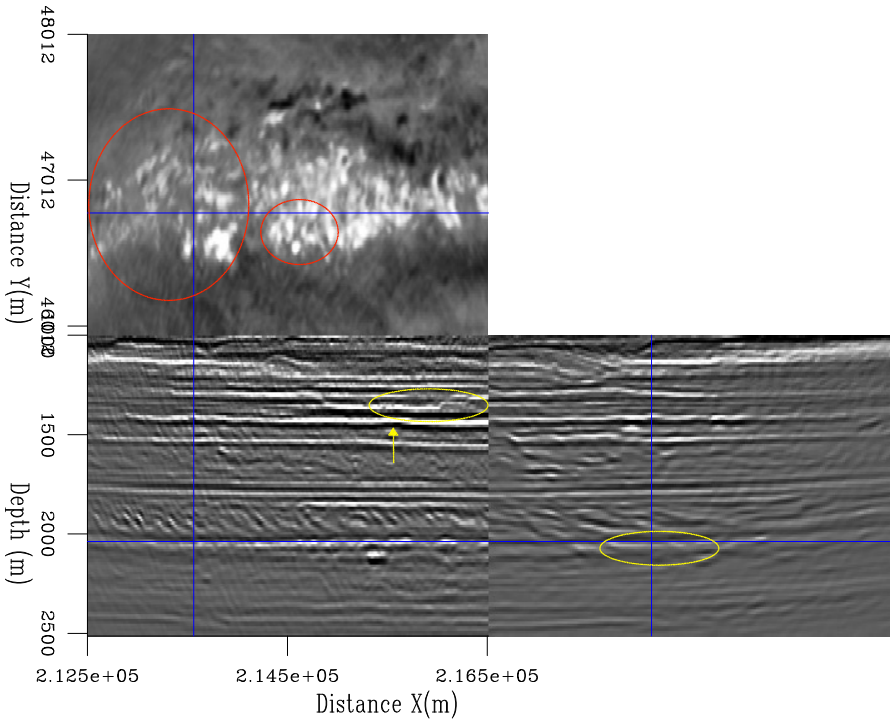


(b)

Figure 5.17: Refined RTM volume for a) WEMVA model, b) JIRB model. [**CR**]



(a)



(b)

Figure 5.18: Refined RTM volume for a) WEMVA model, b) JIRB model. [**CR**]

5.3 Conclusions

The JIRB method results on the OBN 3D dataset show some subtle improvements in comparison with the WEMVA results. It demonstrates that the joint inversion of the reflectivity and the background model by combining the least-squares RTM and the WEMVA operators can potentially correct residual inaccuracies in velocity, which in turn benefits the estimated reflectivity. The JIRB method yielded a better estimation of the reflectivity than LWI, but at a much higher cost. I had initially sought the improvement of the reflectivity by allowing updating the background model. However, as my project progressed, I also had to look at the method from a different perspective: to improve the background model by incorporating the reflectivity, instead of inverting for the background model alone. As so, I compared the JIRB vs. the WEMVA inversion, which are much more comparable in cost. There are small yet noticeable improvements with the JIRB method. Such improvements give a firm clue that incorporating the reflectivity into the background model inversion contributes to the correction of small inaccuracies in the latter. There is still plenty of room for future research on this topic. Of particular interest is obtaining an appropriate value of the λ parameter, which in this dissertation, I had to estimate by trial and error.

Appendix A

Problem in linearizing the image norm

The linear objective function (2.6) has a serious drawback in the second functional. The key point is that first performing the linearization of $\mathbf{I}(\mathbf{b})$ and plugging the expression into the norm $||\cdot||_2^2$, is not the same as the linearization of $||\mathbf{I}(\mathbf{b})||_2^2$. In fact, doing the former completely changes the properties of $||\mathbf{I}(\mathbf{b})||_2^2$.

Let us include another term in the expansion of the migration image

$$\mathbf{I}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b}) = \mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} + \mathbf{Z}(\mathbf{b}_0, \boldsymbol{\Delta}\mathbf{b}) + O(||\boldsymbol{\Delta}\mathbf{b}||^3), \tag{A.1}$$

where $\mathbf{Z}(\mathbf{b}_0, \boldsymbol{\Delta}\mathbf{b})$ incorporates all the second order terms in the expansion. Plugging (A.1) into the function $||\mathbf{I}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b})||_2^2$ gives

$$
\begin{aligned}
||\mathbf{I}(\mathbf{b}_0 + \boldsymbol{\Delta}\mathbf{b})||_2^2 &= \mathbf{I}(\mathbf{b}_0)^{\mathbf{T}}\mathbf{I}(\mathbf{b}_0) + 2\boldsymbol{\Delta}\mathbf{b}^{\mathbf{T}}\mathbf{W}(\mathbf{b}_0)^{\mathbf{T}}\mathbf{I}(\mathbf{b}_0) \\
&\quad + 2\mathbf{Z}(\mathbf{b}_0, \boldsymbol{\Delta}\mathbf{b})^{\mathbf{T}}\mathbf{I}(\mathbf{b}_0) + \boldsymbol{\Delta}\mathbf{b}^{\mathbf{T}}\mathbf{W}(\mathbf{b}_0)^{\mathbf{T}}\mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b} + O(||\boldsymbol{\Delta}\mathbf{b}||^3) \quad \text{(A.2)} \\
&= \|\mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}\|_2^2 + 2\mathbf{Z}(\mathbf{b}_0, \boldsymbol{\Delta}\mathbf{b})^{\mathbf{T}}\mathbf{I}(\mathbf{b}_0) + O(||\boldsymbol{\Delta}\mathbf{b}||^3).
\end{aligned}
$$

Notice the additional term $2\mathbf{Z}(\mathbf{b}_0, \boldsymbol{\Delta}\mathbf{b})^{\mathbf{T}}\mathbf{I}(\mathbf{b}_0)$, which is second order and is not accounted for by the second term in the objective function (2.6). Therefore, the approximation of such term of is not order accurate, which is problematic in this case as the curvature of the function is flipped near $\boldsymbol{\Delta}\mathbf{b}^{\text{true}}$, being no longer concave. In fact, the approximation $\|\mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}\|_2^2$ is actually a convex function with a constant Hessian that is at least positive semidefinite. As consequence, there is no upper bound or maximum close to the correct background model perturbation, as we observe in the numerical results in Chapter 2 (see Figure 2.9). This implies that one can move along any arbitrary direction with strictly positive curvature, and increase the objective function $\|\mathbf{I}(\mathbf{b}_0) + \mathbf{W}(\mathbf{b}_0)\boldsymbol{\Delta}\mathbf{b}\|_2^2$ indefinitely, which is clearly not what is desired.

Appendix B

Preconditioning in WEMVA

Inversion using WEMVA is intended to solely recover the low-wavenumber component of the subsurface. However, the high-wavenumber component, in other words, the reflectivity, often manifests its presence in the solution. The reason for that is because such a high-wavenumber component belongs to the model null space of the WEMVA operator. A similar problem arises when DSO is utilized in the extended domain and high-wavenumber artifacts are present in the gradient, thus affecting the final inverted model (Vyas and Tang, 2010). One popular way to combat such artifacts is preconditioning the gradient with B-splines (Weibull and Arntsen, 2013). On the other hand, the migration image in the WEMVA objective function might also required preconditioning to appropriately weight the amplitudes of the shallow part with respect to the deep part. In the following I show both preconditioning strategies (for the background model and for the image) that I applied in the numerical examples that illustrate this thesis.

B.1 Preconditioning the background model

Recall the WEMVA objective function for maximization of the image power,

$$\Phi(\mathbf{b}) = -\frac{1}{2}\left\|\mathbf{I}(\mathbf{b})\right\|_2^2 .$$

We can define $\mathbf{b} = \mathbf{Bp}$, where $\mathbf{B}$ represents the B-spline interpolation operator (Fomel, 2000; Tang, 2011b; Weibull and Arntsen, 2013), and $\mathbf{p}$ represents the projection of the background model onto a coarse grid. Operator $\mathbf{B}$ maps the subsurface model from a coarse grid onto a fine grid.

Substituting in the objective function gives

$$\Phi(\mathbf{p}) = -\frac{1}{2}\left\|\mathbf{I}(\mathbf{Bp})\right\|_2^2 . \tag{B.1}$$

This objective function is optimized with respect to the subsurface model projected onto the coarse grid, $\mathbf{p}$. Then, it is interpolated into the fine grid using operator $\mathbf{B}$. This procedure removes high-wavenumber components from the background model whilst retaining the expected low-wavenumber components. We can obtain the initial model $\mathbf{p}_0$ from $\mathbf{b}_0$ by iteratively solving the following normal equations:

$$\mathbf{B}^{\mathrm{T}}\mathbf{B}\mathbf{p}_0 = \mathbf{B}^{\mathrm{T}}\mathbf{b}_0$$

B.2 Preconditioning the image

It is often the case that the migration image exhibits imbalanced amplitudes, e.g., high amplitudes the shallow part that decrease as going deeper. We can include a weighting factor that compensates for such an effect, thereby obtaining

$$\Phi(\mathbf{p}) = -\frac{1}{2}\left\|\mathbf{EI}(\mathbf{Bp})\right\|_2^2 . \tag{B.2}$$

where

$$\mathbf{E} = \mathrm{diag}(\mathbf{z}^{\mathrm{n}}),$$

and $\mathbf{z}$ is the vector that contains the depth steps, and n is an integer positive value. In other words, the diagonal matrix operator $\mathbf{E}$ applies a depth-dependent amplitude compensation to the image.

B.3 Gradient of preconditioned objective function

We can calculate the gradient of the preconditioned objective function and obtain

$$\begin{aligned}\nabla\Phi(\mathbf{p}) &= -\left[\frac{\partial \mathbf{EI}(\mathbf{Bp})}{\partial \mathbf{p}}\right]^{\mathrm{T}} \mathbf{EI}(\mathbf{Bp}) = -\left[\frac{\partial \mathbf{I}(\mathbf{Bp})}{\partial \mathbf{b}}\frac{\partial \mathbf{b}}{\partial \mathbf{p}}\right]^{\mathrm{T}} \mathbf{E}^{\mathrm{T}}\mathbf{EI}(\mathbf{Bp}) \\ &= -\big[\mathbf{W}(\mathbf{Bp})\mathbf{B}\big]^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\mathbf{EI}(\mathbf{Bp}) = -\mathbf{B}^{\mathrm{T}}\mathbf{W}(\mathbf{Bp})^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\mathbf{EI}(\mathbf{Bp}).\end{aligned} \tag{B.3}$$

Note that the minus sign of the objective function ought to be preserved in the gradient.

Bibliography

Almomin, A. and Y. Tang, 2010, Migration velocity analysis based on linearization of the two-way wave equation: SEP-Report, **142**, 13–24.

Aster, R., B. Borchers, and C. Thurber, 2013, Parameter estimation and inverse problems: Academic Press.

Barnier, G. and A. Almomin, 2014, Tutorial on two-way wave equation operators for acoustic, isotropic, constant-density media: SEP-Report, **155**, 23–56.

Biondi, B., 2006, 3d Seismic Imaging: Society of Exploration Geophysicists.

Biondi, B., E. Biondi, M. Maharramov, and Y. Ma, 2015, Dissection of the full-waveform inversion hessian: SEP-Report, **160**, 19–38.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, 1723–1726.

Cabrales-Vargas, A., 2018, Improving reflectivity using linearized waveform inversion with velocity updating: SEP-Report, **172**, 193–208.

Cabrales-Vargas, A., B. Biondi, and R. Clapp, 2016, Linearized Waveform Inversion with Velocity Updating: Theory and first results: SEP-Report, **165**, 63–92.

Cabrales-Vargas, A. and R. Sarkar, 2019, A nonlinear scheme to perform linearized waveform inversion with velocity updating: SEP-Report, **176**, 85–100.

Claerbout, J., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, no. 3, 467–481.

Clapp, R., 2009, Reverse time migration with random boundaries: SEG Technical Program Expanded Abstracts, 2809–2813.

Clapp, R. G., 2008, Reverse time migration: Saving the boundaries: SEP-Report,

136, 137–144.

Clapp, R. G. and G. Alves, 2016, Random boundaries for elastic medium: SEP-Report, **163**, 41–48.

Dahlke, T., 2019, Velocity model building using shape optimization applied to level sets: PhD thesis, Stanford University.

Dai, W. and G. Schuster, 2013, Plane-wave least-squares reverse-time migration: Geophysics, **78**, no. 4, S165–S177.

Duquet, B., K. Marfurt, and J. Dellinger, 2000, Kirchhoff modeling, inversion for reflectivity and subsurface illumination: Geophysics, **65**, 1195–1209.

Fletcher, R., D. Nichols, R. Bloor, and R. Coates, 2016, Least-squares migation - data domain versus image domain using point spread functions: The Leading Edge, **35**, no. 2, 157–162.

Fomel, S., 2000, Inverse B-spline interpolation: SEP-Report, **105**, 79–108.

Grion, S., R. Exley, M. Manin, X. Miao, A. Pica, and G. P. R. S. Wang, Y., 2007, Mirror imaging of obs data: First Break, **25**, no. 25, 37–42.

Ha, T. and K. Marfurt, 2017, The value of constrained conjugate-gradient least-squares migration in seismic inversion: Application to a fractured-basement play, Texas panhandle: Interpretation, **5**, no. 3, SN13–SN27.

Huang, Y., 2018, Sedimentary characteristics of turbidite fan and its implication for hydrocarbon exploration in lower congo basin: Petroleum Research, **3**, 189–196.

Jackson, M. and M. Hudec, 2017, Salt tectonics: Cambridge University Press.

Knopoff, L. and A. Gangi, 1959, Seismic reciprocity: Geophysics, **24**, 681–691.

Nemeth, T. and G. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, no. 1, 208–221.

Nocedal, J. and S. Wright, 2006, Numerical optimization: Springer.

Pacal, E., R. Stewart, and a. Y. O. Baysal, E., 2015, Seismic Imaging with Ocean-Bottom nodes (OBN): Mirror Migration Technique: SEG Technical Program Expanded Abstracts, 4039–4043.

Pratt, R., C. Shin, and G. Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Schleicher, J., M. Tygel, and P. Hubral, 2007, Seismic true-amplitude imaging: Society of Exploration Geophysicists.

Schuster, G., 2017, Seismic inversion: Society of Exploration Geophysicists.

Shen, P. and W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, no. 5, VE49–VE59.

Shen, X. and R. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts, 2962–2965.

———, 2015, Random boundary condition for memory-efficient waveform inversion gradient computation: Geophysics, **80**, no. 6, R351–R359.

Soubaras, R., 1996, Ocean bottom hydrophone and geophone processing: SEG Technical Program Expanded Abstracts, 24–27.

Soubaras, R. and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, no. 5, U67–U73.

Sun, W. and L. Fu, 2013, Two effective approaches to reduce data storage in reverse time migration: Computer & Geosciences, **56**, 69–75.

Symes, W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, no. 5, SM214–SM221.

Symes, W. and J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

Tang, Y., 2011a, Fast automatic wave-equation migration velocity analysis using encoded simultaneous sources: SEP-Report, **145**, 27–46.

———, 2011b, Imaging and velocity analysis by target-oriented wavefield inversion: PhD thesis, Stanford University.

Tarantola, A., 1984a, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

———, 1984b, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015.

Tromp, J. and L. Q. Tape, C., 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, no. 1, 195–216.

Valenciano, A. and B. Biondi, 2003, 2d deconvolution imaging condition for shotprofile migration: SEG Technical Program Expanded Abstracts, 1059–1062.

———, 2009, Imaging by target-oriented wave-equation inversion: Geophysics, **74**, no. 6, WCA109–WCA120.

Valenciano, A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, no. 4, A35–A38.

Valenciano, A. A., 2008, Imaging by wave-equation inversion: PhD thesis, Stanford University.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion: Geophysics, **74**, no. 6, WCC1–WCC26.

Vyas, M. and Y. Tang, 2010, Gradients for the wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, 4077–4081.

Weibull, W. and B. Arntsen, 2013, Automatic velocity analysis with reverse-time migration: Geophysics, **78**, no. 4, S179–S192.

Wong, M., B. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: Geophysics, **80**, no. 6, S223–S235.

Xu, S., D. Wang, F. Chen, and G. Lambaré, 2012, Full waveform inversion for reflected seismic data: EAGE Extended Abstracts, W024.

Yang, W., Y. Li, A. Cheng, Y. Liu, and L. Dong, 2018, Least-squares reverse time migration with velocity errors: SEG Technical Program Expanded Abstracts, 4256–4260.

Yu, J., J. Hu, G. Schuster, and R. Estill, 2006, Prestack migration deconvolution: Geophysics, **71**, no. 2, 1942–2156.

Zhang, Y. and G. Shan, 2013, Wave-equation migration velocity analysis using partial stack-power maximization: SEG Technical Program Expanded Abstracts, 4847–4852.