Wiktor Waldemar Weibull

Seismic Processing and Velocity Analysis Using Reverse-Time Migration

Thesis for the degree of Philosophiae Doctor

Trondheim, November 2013

Norwegian University of Science and Technology
Faculty of Engineering Science and Technology
Department of Petroleum Engineering and
Applied Geophysics

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Engineering Science and Technology
Department of Petroleum Engineering and Applied Geophysics

ISBN 978-82-471-4868-6 (printed ver.)
ISBN 978-82-471-4869-3 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2014:3

Printed by NTNU-trykk

To my parents and family.

Acknowledgements

First, I would like to thank my adviser, Professor Børge Arntsen for his patience, support and guidance through these years. I feel very fortunate to have been his PhD student.

I thank Dr. Biondo Biondi, Dr. Leiv Gelius and Dr Alexey Stovas for serving on my thesis committee.

I acknowledge the financial support of Statoil ASA. I am especially grateful to Mark Thomson, Jon Andre Haugen, Marianne Houbiers, Joachim Mispel and Aleksander Kritski for support of this project and for valuable discussions on issues related to processing of seismic data.

All colleagues and staff at the Department of Petroleum Engineering and Applied Geophysics at NTNU are greatly acknowledged for their cooperation and for creating a good working environment. It was a great pleasure to meet persons like Lyubov, Olena, Pavel, Sthener, Mohammad Ashrafi, Amir Ghaderi, Hossein, Benjamin, Pamela, Torgeir, Mehran, Raheleh, Hadi, Tuhin, Anastasia, Sohrab, Sissel, Anders, Sandra, Lutz, Anastasia, Milan, Jesus, Mayembe, Mohammad Ghasemi, Amir, Torbjørn, Kate, and many others who shared their lunch and free times with me during these years.

Thanks to professors Lasse Amundsen, Martin Landrø and Ketil Hokstad for your lectures, and for being a great source of inspiration. Special thanks to Professors Alexey Stovas and Bjørn Ursin for providing helpful discussions on migration velocity analysis and anisotropic parameter estimation.

Thanks to the students of the NTNU Modeling and Imaging (MSIM) group, Espen Birger Raknes, Jon Marius Venstad, and Yi Liu for their help with issues related to mathematics and programming.

I acknowledge the NTNU HPC group, for their support during the implementation of the parallel computing applications, and for maintaining the cluster.

I would like to extend my biggest thanks to my mother Lillian, my brothers Jon and Oscar, my sister Vera, and my close friends for all support and encouragement and inspiration throughout my life.

Finally, I am also indebted to my wife Carla, and my children Lucas and Cecilia for understanding, patience and support during the work of this thesis.

Contents

List of Figures

Chapter 1

Introduction

The general goal of this thesis has been to improve seismic imaging over complex geological settings. Seismic imaging, or seismic migration as it is also called, is the problem of creating images of the upper part of the Earth's crust using elastic waves. The physics of wave-propagation is such that interfaces (or borders) between different geological units exhibiting changes in mechanical properties give rise to reflected waves which can be turned into images. The mathematical description of reflected seismic waves can be approximated by a set of linear equations, where the unknowns are the so-called reflection coefficients. The solution of the linear reflection coefficient problem, called prestack depth migration (PSDM), is a structural image of the subsurface. PSDM is a trivial routine practice in simple geological settings, and can be solved by relatively efficient algorithms based on ray approximations to the wave equation, such as Kirchhoff migration (Schneider, 1978). However, ray approximations have several shortcomings and often fail to be adequately accurate in complex geological structures with large and sharp velocity contrasts (Arntsen et al., 2009). In these areas, the problem is more complicated, and is better solved using more modern PSDM algorithms based on finite-frequency wavefield extrapolation, such as one-way wave equation migration (Berkhout, 1980), and reverse-time migration (Baysal et al., 1983). Independent of the approximations used, PSDM requires the knowledge of a "smooth" background approximation to the mechanical properties of the subsurface. These background models are usually expressed in terms of wave velocities and the problem of estimating them is known as velocity analysis. In this thesis, I combine state of the art PSDM by reverse-time migration (RTM) and velocity analysis in a non-linear iterative process called wave equation migration velocity analysis (WEMVA). The developed methods are explained in detail over two peer-reviewed papers included in chapters 2 and 3. In chapter 4 I include a paper where I derive methods to process seismic reflection data based on RTM with optimal application to WEMVA.

In this chapter, I present a brief introduction to the theory behind PSDM by RTM and WEMVA in a broad and general sense. At the end, I present a summary of the individual papers of the thesis.

1.1 The seismic experiment

In a seismic experiment, a wavefield is induced by a controlled source. Which can be a source of volume injection or a body force applied on the earth. The resultant wave propagates down into the earth, where part of the energy is backscattered due to reflection and refraction at interfaces that mark contrasts in the elastic properties of the earth. The wavefield is sampled in space and time by a limited number of pressure and/or displacement recorders. The source and/or the receivers are repositioned many times increasing the coverage area, as well as producing redundant data. There exist numerous different setups to acquire seismic data both on land, and over the ocean. In this introduction, as well as in this thesis, I will focus on marine seismic acquisition.

The most common acquisition method in marine seismic is the towed streamed setup. In this case, the source(s) and one or more cables equipped with hydrophones are towed at a shallow depth below the sea surface. The waves are induced by a source of volume injection, typically an airgun (Parkes and Hatton, 1986). Because this source approximates a pressure source and because water is an acoustic medium, the resultant waves are purely pressure waves (P-waves). Likewise the seismic data acquired at the water column will consist only of P-waves. Depending on the number of towed hydrophone cables the acquisition can be either of 2D or 3D data.

To acquire shear waves (S-waves) it is necessary to measure the seismic response at a solid. This can be done by deploying receivers at the sea bottom, or in a borehole. The first situation is generally called ocean bottom seismic acquisition (OBS), and typically uses four component receivers, one hydrophone plus a 3-component geophone. The second is called vertical seismic profiling (VSP), and uses three component geophones. In both cases, the resultant data is multicomponent. In OBS and VSP acquisition, the source is still deployed at the water column and therefore is a P-wave source. Which means that the S-waves recorded at the sea bottom are generated purely through P- to S-wave conversion (PS waves). Figure 1.1 shows snapshots of vertical particle displacement due to elastic wave propagation over a layered elastic model. The snapshots are overlaid on a model of P-wave velocity. The figure also illustrates the main types of marine seismic data acquisition. Note the conversion of P-waves into S-waves at the interfaces of the model.

1.2 Seismic imaging

The recorded seismic data can be processed into an image of the subsurface reflectors using PSDM. To achieve this, one need to have a model of wave propagation. The starting point for any method of seismic imaging are the equation of motion and the generalized Hooke's law. The equation of motion in a continuum are given by

$$\rho(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x},t)}{\partial t^2} = \frac{\partial \tau_{ij}(\mathbf{x},t)}{\partial x_j} + F_i(\mathbf{x},t), \tag{1.1}$$

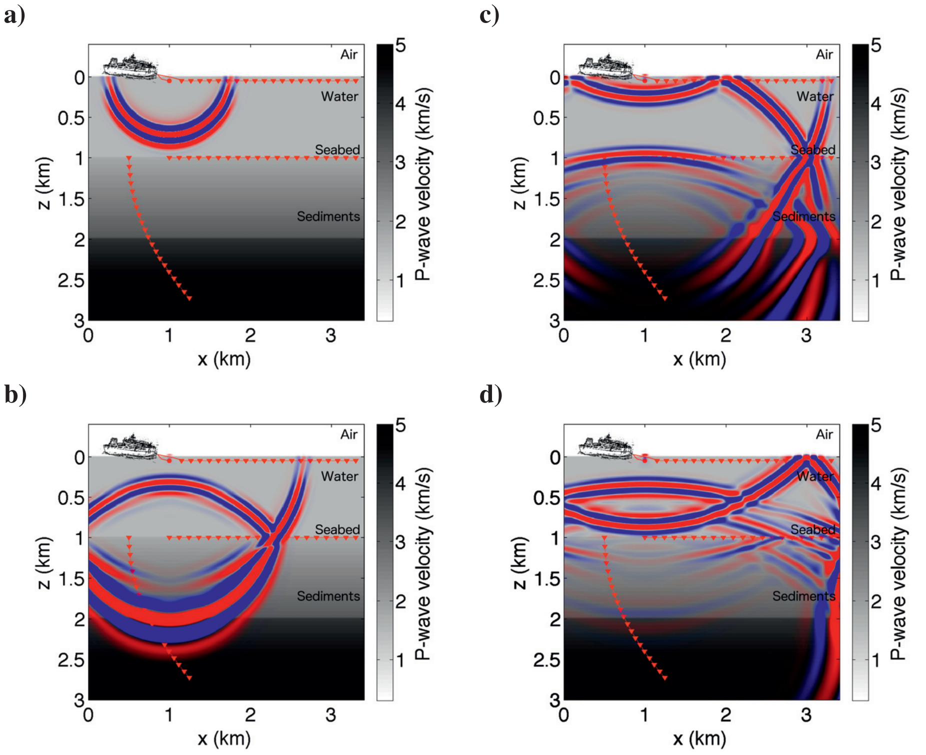


Figure 1.1: Snapshots of vertical particle displacement due to elastic wave propagation over a model of marine seismic acquisition. The snapshots are overlaid on a model of P-wave velocity. (a) Snapshot at 0.6 s; (b) Snapshot at 1.0 s; (c) Snapshot at 1.4 s; (d) Snapshot at 1.7 s. The red dot indicates the source position, while the inverted red triangles are receivers. There are three different acquisition geometries depicted. The shallowest receivers represent a towed streamer acquisition, while the line of receivers at 1 km depth represent an OBS acquisition line. Finally, the receivers in near vertical orientation represent a VSP acquisition setup.

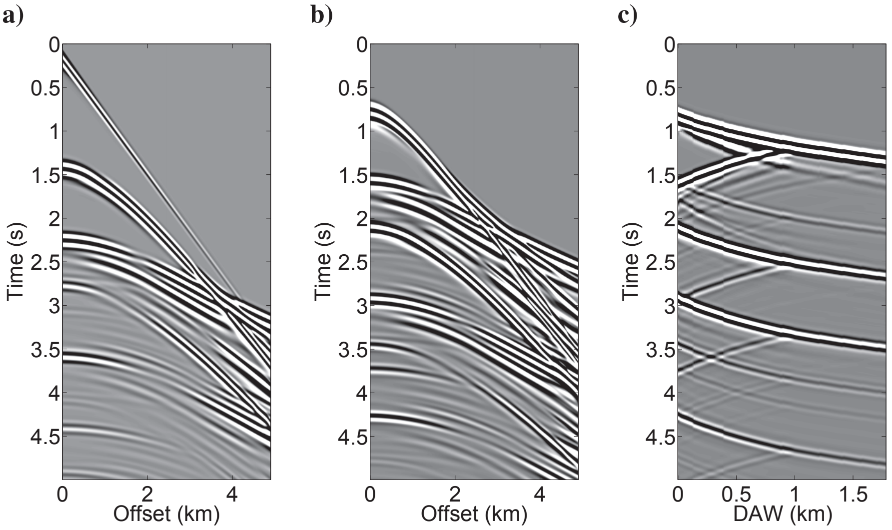


Figure 1.2: Seismic shot records obtained at the receivers depicted in Figure 1.1. (a) Towed streamer; (b) OBS acquisition; (c) VSP acquisition. DAW = distance along the well.

where $\mathbf{x} = (x, y, z)$ is the Cartesian basis; t is the time; $i, j = 1, 2, 3$, under Einstein's summation convention; ρ is the density; u_i are the particle displacements; τ_{ij} are the stresses; F_i is a body force vector acting as a seismic source.

The equation (1.1) is valid for any stress-strain relation including models of anelastic or anisotropic media. In this thesis, I deal with models of *elastic anisotropic* media. For these media the stress-strain relationship is approximated by the generalized Hooke's law

$$\tau_{ij}(\mathbf{x}, t) = c_{ijkl}(\mathbf{x}) e_{kl}(\mathbf{x}, t), \tag{1.2}$$

where the strains are related to the displacements by

$$e_{ij}(\mathbf{x}, t) = \frac{1}{2}\left(\frac{\partial u_i(\mathbf{x}, t)}{\partial x_j} + \frac{\partial u_j(\mathbf{x}, t)}{\partial x_i}\right), \tag{1.3}$$

and c_{ijkl} is a constant of proportionality known as the elasticity tensor.

The elasticity tensor contains information on how fast P- and S-waves travel as a function of direction at one particular point in the subsurface. In general, the elasticity tensor has 81 components, however, due to natural symmetries, the number of independent components is only 21, since $c_{ijkl} = c_{klij} = c_{jilk}$. Without loss of generality, the elasticity tensor can be written as the symmetric 6×6 matrix of elastic coefficients $\mathbf{C}$ using the Voigt notation,

$11 \rightarrow 1$, $22 \rightarrow 2$, $33 \rightarrow 3$, $23 \rightarrow 4$, $13 \rightarrow 5$, $12 \rightarrow 6$

$$[\mathbf{C}] = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix}. \tag{1.4}$$

It is important to state that to describe wave propagation over a fully heterogeneous and anisotropic medium (triclinic symmetry) we need 21 independent quantities for each point ($\mathbf{x}$) in the subsurface. Fortunately, the geology of most sedimentary basins is such that many simplifications (or symmetries) can be sought to reduce the number of parameters that are needed to describe wave propagation over the medium. The simplest case of all is that in which the medium is isotropic. In this case, $\mathbf{C}$ depends only on the Lamé parameters λ and μ

$$[\mathbf{C}_{ISO}] = \begin{pmatrix} \lambda+2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda+2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda+2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}. \tag{1.5}$$

The Lamé parameters are related to the P- and S-wave velocities through $v_p = \sqrt{\frac{\lambda+2\mu}{\rho}}$, and $v_s = \sqrt{\frac{\mu}{\rho}}$.

1.2.1 Acoustic wave equation

For an acoustic medium, the shear modulus μ is equal to zero, meaning that the equations (1.1), (1.2), and (1.3) can be simplified to

$$\left[\frac{1}{\rho(\mathbf{x})v_p^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\left(\frac{1}{\rho(\mathbf{x})}\nabla\right)\right] p(\mathbf{x},t) = -\nabla\cdot\left[\frac{1}{\rho(\mathbf{x})}\mathbf{F}(\mathbf{x},t)\right], \tag{1.6}$$

where p is the acoustic pressure, $\tau_{ij} = -p\delta_{ij}$, and δ_{ij} is the Kronecker delta.

For the particular case where the density is constant

$$\left[\frac{1}{v_p^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right] p(\mathbf{x},t) = -\nabla\cdot\mathbf{F}(\mathbf{x},t). \tag{1.7}$$

Equation (1.7) is the simplest wave equation that honors the kinematics of P-wave propagation in the subsurface. It is by large the most commonly used approximation in seismic imaging. This is mainly due to the fact that to solve this equation only the isotropic P-wave velocity need to be estimated for each point in the subsurface.

1.2.2 Acoustic reverse-time migration

In PSDM by reverse-time migration, wave propagation modeling is used to simulate the source wavefield forward in time, as well as to propagate the recorded wavefield backward in time. The acoustic wave equation (1.7) is solved by employing differential/spectral methods, either explicitly or implicitly. Whereas reflectivity at each point in the subsurface can be estimated by applying an imaging condition involving both reconstructed wavefields. The resultant reflectivity map of the subsurface is often referred to as the migrated image.

Here we will develop the theory for acoustic reverse-time migration in the time domain. The equations describing the forward- and reverse-time pressure modelings are given by

$$\left[\frac{1}{v_p^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right] w^s(\mathbf{x},t;s) = f(\mathbf{x},t;s), \tag{1.8}$$

and

$$\left[\frac{1}{v_p^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right] w^r(\mathbf{x},t;s) = d(\mathbf{x},T-t;s) \tag{1.9}$$

where, w^s and w^r are, respectively, the source and receiver acoustic wavefields; s is the source index; f is the source function for a particular seismic experiment s; and d is the recorded reflection data for that particular seismic experiment. Note that the recorded data is time-reversed, indicating that the modeling is done in reverse-time. Each seismic experiment, often referred to as a *shot* (firing of the source(s) and recording), is reconstructed using the above equations. This is commonly done using well known finite difference solutions to the wave equation (Virieux, 1986).

At each time step of the reconstruction, an image is constructed by employing an *imaging condition* relating the source and receiver wavefields at that time step (Claerbout, 1971; Rickett and Sava, 2002)

$$R(\mathbf{x};s) = \int_0^T w^s(\mathbf{x},t;s) w^r(\mathbf{x},T-t;s). \tag{1.10}$$

This is explained by the fact that at the point of reflection, the forward propagated source wavefield and the back propagated receiver wavefield will meet perfectly and produce a peek in the zero time lag of the cross-correlation. However, this is only true if the velocities along the wave paths from the source and receiver to the imaging point are correct, otherwise the peek will lie in some vicinity of the imaging point. Interestingly, in the context of non-linear inversion theory (Tarantola, 1984), this imaging condition appears as the sensitivity kernel of the least squares norm of the difference between the observed data and the data modeled over a smooth background medium.

A key component of WEMVA used in this thesis is the analysis of common image point gathers (CIGs). These gathers represent the reflectivity as a function of offset or angle. These can be used to build semblance or differential semblance error measures that capture the deviation of the maximum of cross-correlation from the actual imaging point. To

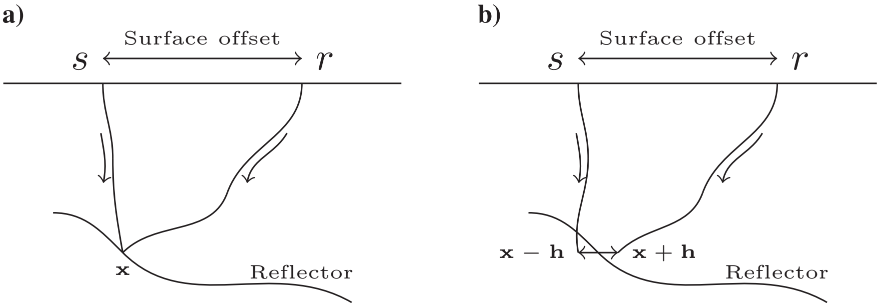


Figure 1.3: Wavefield reconstruction causes reflection energy to move towards the imaging points. The peeks in the cross-correlation of the source and receiver wavefields reveal the reflectors. (a) The zero spatial lag cross-correlation leads to an image of the reflectors; (b) Lagged cross-correlation produces CIGs as a function of space $\mathbf{x}$ and subsurface offset $\mathbf{h}$.

build CIGs, the imaging condition is extended to a lagged cross-correlation. The lags can be either spatial, temporal, or both (Sava and Vasconcelos, 2011). In this thesis, we use CIGs constructed using horizontal spatial lags

$$R(\mathbf{x}, \mathbf{h}; s) = \int_0^T w^s(\mathbf{x} - \mathbf{h}, t; s) w^r(\mathbf{x} + \mathbf{h}, T - t; s) \tag{1.11}$$

In the resultant image, the offset coordinates $\mathbf{h} = (h_1, h_2, 0)$ refer to the effective subsurface offset between the source and receiver wavefields (Figure 1.3). The resultant offset dependent images are commonly displayed as offset-depth panels. One important property of these CIGs is that if velocities are correctly estimated the energy from the sum of the images for all shots will be focused at zero lag, i.e. $\mathbf{h} = (0, 0, 0)$. That is, all peaks in cross-correlation will properly stack at the correct reflection points (Figure 1.4b-d). There are also imaging conditions that produce migrated images which are a function of angle of incidence instead of offset (deBruin, 1990), and methods that can be used to convert offset dependent images to angle dependent images (Rickett and Sava, 2002; Biondi and Symes, 2004). These images are displayed as angle-depth panels, and are referred to as angle domain CIGs (Figure 1.4c-e). In ADCIGs, the velocity accuracy is measured by the flatness of the migrated reflections (Shen and Symes, 2008).

1.2.3 Elastic anisotropic wave equation

The acoustic wave equation (1.7) has a long and productive history in seismic imaging (Gray et al., 2001; Etgen et al., 2009). However, theoretical studies, laboratory measurements and field studies have shown that many sedimentary rocks are anisotropic (Levin, 1979; Thomsen, 1986; Larner, 1993). Ignoring anisotropy in seismic imaging and velocity analysis often lead to significant differences between the spatial positions observed

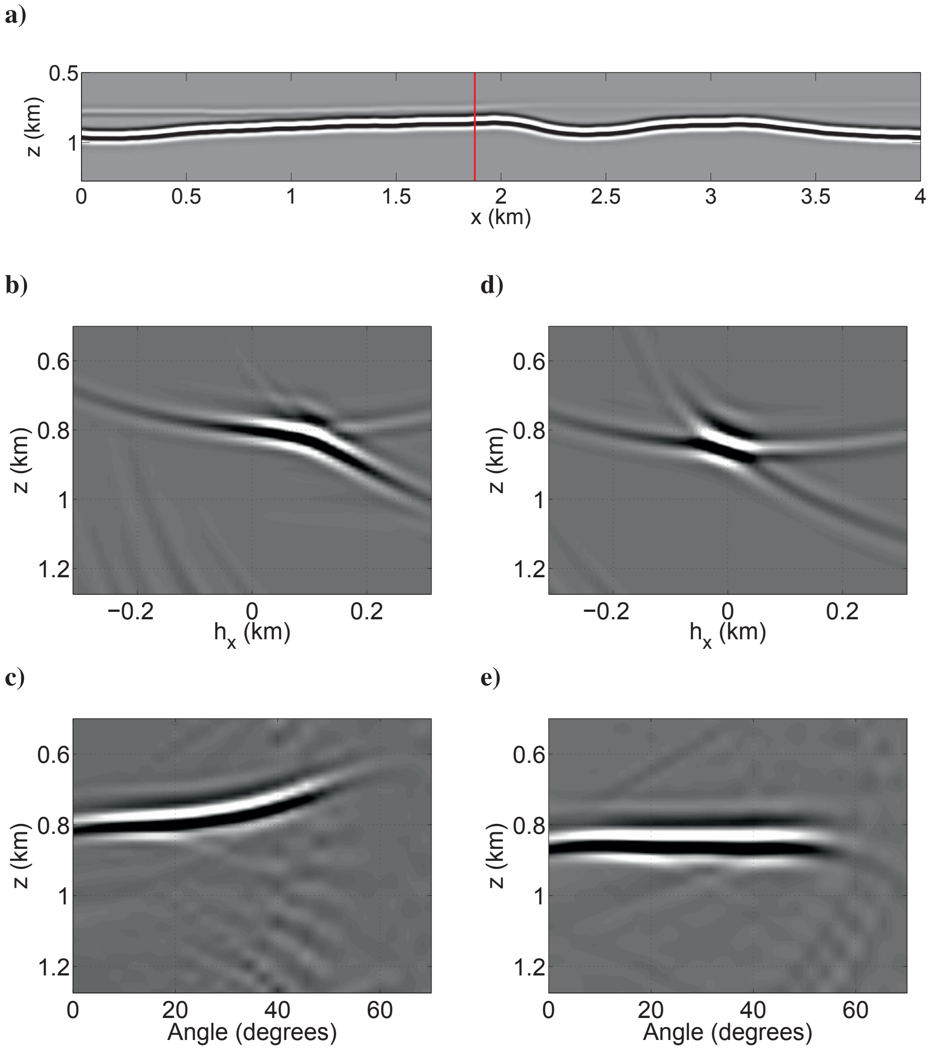


Figure 1.4: Images and CIGs constructed with RTM. (a) Zero lag cross-correlation image; (b) CIG at the position given by the vertical red line in (a), after migration with a wrong velocity model; (c) Angle domain CIG produced by slant stack of the CIG in (b); (d) CIG at the position given by the vertical red line in (a), after migration with the correct velocity model; (e) Angle domain CIG produced by slant stack of the CIG in (d).

in the seismic images and the actual geological positions observed in boreholes (Banik, 1984; Ball, 1995). The anisotropy of sedimentary rocks can be well approximated by a transverse isotropic medium (TI), with a symmetry axis perpendicular to the bedding plane (Byun, 1984; Audebert et al., 2006). In the case of TI media the matrix of elastic coefficients has five independent parameters plus two angles which control the direction of the symmetry axis. The simplest form of the matrix of elastic coefficients for TI medium corresponds to the case where the symmetry axis is vertical (VTI)

$$[\mathbf{C}_{VTI}] = \begin{pmatrix} C_{11} & C_{11} - 2C_{66} & C_{13} & 0 & 0 & 0 \\ C_{11} - 2C_{66} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{11} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}. \tag{1.12}$$

The tilted TI matrix C_{TTI} can be obtained by rotating the VTI matrix along x and y, or in index notation along x_1 and x_2. This can be most efficiently done using the Bond matrix

$$[\mathbf{M}] = \begin{pmatrix} r_{11}^2 & r_{12}^2 & r_{13}^2 & 2r_{12}r_{13} & 2r_{13}r_{11} & 2r_{11}r_{12} \\ r_{21}^2 & r_{22}^2 & r_{23}^2 & 2r_{22}r_{23} & 2r_{23}r_{21} & 2r_{21}r_{22} \\ r_{31}^2 & r_{32}^2 & r_{33}^2 & 2r_{32}r_{33} & 2r_{33}r_{31} & 2r_{31}r_{32} \\ r_{21}r_{31} & r_{22}r_{32} & r_{23}r_{33} & r_{22}r_{33}+r_{23}r_{32} & r_{21}r_{33}+r_{23}r_{31} & r_{22}r_{31}+r_{21}r_{32} \\ r_{31}r_{11} & r_{32}r_{12} & r_{33}r_{13} & r_{12}r_{33}+r_{13}r_{32} & r_{13}r_{31}+r_{11}r_{33} & r_{11}r_{32}+r_{12}r_{31} \\ r_{11}r_{21} & r_{12}r_{22} & r_{13}r_{23} & r_{12}r_{23}+r_{13}r_{22} & r_{13}r_{21}+r_{11}r_{23} & r_{11}r_{22}+r_{12}r_{21} \end{pmatrix}, \tag{1.13}$$

where $\mathbf{r} = (r_{ij})$ is the product of two elementary rotation matrices

$$[\mathbf{r}] = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ -\sin\phi\sin\theta & \cos\phi & \sin\phi\cos\theta \\ -\cos\phi\sin\theta & -\sin\phi & \cos\phi\cos\theta \end{pmatrix}, \tag{1.14}$$

and θ, and ϕ are the rotation angles about respectively the y-, and the x-axis. Using the Bond matrix the TTI matrix can be obtained from the VTI matrix according to

$$\mathbf{C}_{TTI} = \mathbf{M}\mathbf{C}_{VTI}\mathbf{M}^T. \tag{1.15}$$

To honor anisotropy in wave propagation it is necessary to use the elastic wave equation. Combining equations (1.1), (1.2), and (1.3) leads to the elastodynamic wave equation for the displacement field

$$\rho(\mathbf{x})\frac{\partial^2 u_i(\mathbf{x},t)}{\partial t^2} - \frac{\partial}{\partial x_j}\left[c_{ijkl}(\mathbf{x})\frac{\partial u_k(\mathbf{x},t)}{\partial x_l}\right] = F_i(\mathbf{x},t). \tag{1.16}$$

Most useful for seismic imaging and velocity analysis is the density normalized version of the equation (1.16)

$$\frac{\partial^2 u_i(\mathbf{x},t)}{\partial t^2} - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_k(\mathbf{x},t)}{\partial x_l}\right] = \frac{1}{\rho}F_i(\mathbf{x},t), \tag{1.17}$$

where $a_{ijkl} = c_{ijkl}/\rho$ is the density normalized elasticity tensor.

The elastic wave equation (1.17) accurately models the kinematics of both P- and S-waves in general heterogeneous anisotropic media. In 3D anisotropic media, the solution of the elastic equation leads to the propagation of three distinct, but in general coupled, wave modes: The quasi-P, quasi-SV, and quasi-SH waves. An alternative wave equation, called the pseudo-acoustic wave equation has been derived from the VTI elastic wave equation by (Alkhalifah, 2000), extended to TTI media by Zhang et al. (2003), and to orthorhombic media by Song and Alkhalifah (2012). The purpose of the pseudo-acoustic wave equation is to accurately, and efficiently, model the kinematic behavior of only the qP-waves, avoiding the qS-waves altogether. However, doing so, prescribes the possibility of using this equation in the emerging realm of multicomponent data. In this thesis, all anisotropic wave propagation is carried out using the elastic wave equation (1.17). Figure 1.5 show a comparison of snapshots of the solution of wave propagation using equation (1.17) over homogeneous isotropic, VTI and TTI models.

1.2.4 Elastic reverse-time migration

Similar to acoustic PSDM by RTM, elastic RTM is a two step process consisting of wavefield reconstruction and imaging condition. The main difference is that the wavefield reconstruction yields a displacement vector. The equations describing the wavefield reconstruction in an elastic anisotropic media are given by

$$\frac{\partial^2 u_i^s(\mathbf{x},t)}{\partial t^2} - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_k^s(\mathbf{x},t)}{\partial x_l}\right] = f_i(\mathbf{x},t), \tag{1.18}$$

$$\frac{\partial^2 u_i^r(\mathbf{x},t)}{\partial t^2} - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_k^r(\mathbf{x},t)}{\partial x_l}\right] = d_i(\mathbf{x},T-t), \tag{1.19}$$

Before applying the imaging condition, scalar wavefields can be computed from the displacement vector wavefields such as to separate P- and S-waves. This can be used to generate separate images for P- and S-waves. Over isotropic media the P- and S-waves can be separated according to

$$w_p(\mathbf{x},t) = \nabla \cdot u_i(\mathbf{x},t), \tag{1.20}$$

and

$$w_s(\mathbf{x},t) = \nabla \times u_i(\mathbf{x},t), \tag{1.21}$$

where w_p, and w_s represent, respectively, the P and S wavefields. In anisotropic media, the above equations do not perfectly separate the quasi-P, quasi-SV and quasi-SH wave modes. In this case, to achieve a better separation, more expensive methods must be used (Dellinger and Etgen, 1990; Yan and Sava, 2011). Once the scalar wavefields have been computed, an imaging condition relating both wave fields can be applied. For example, an

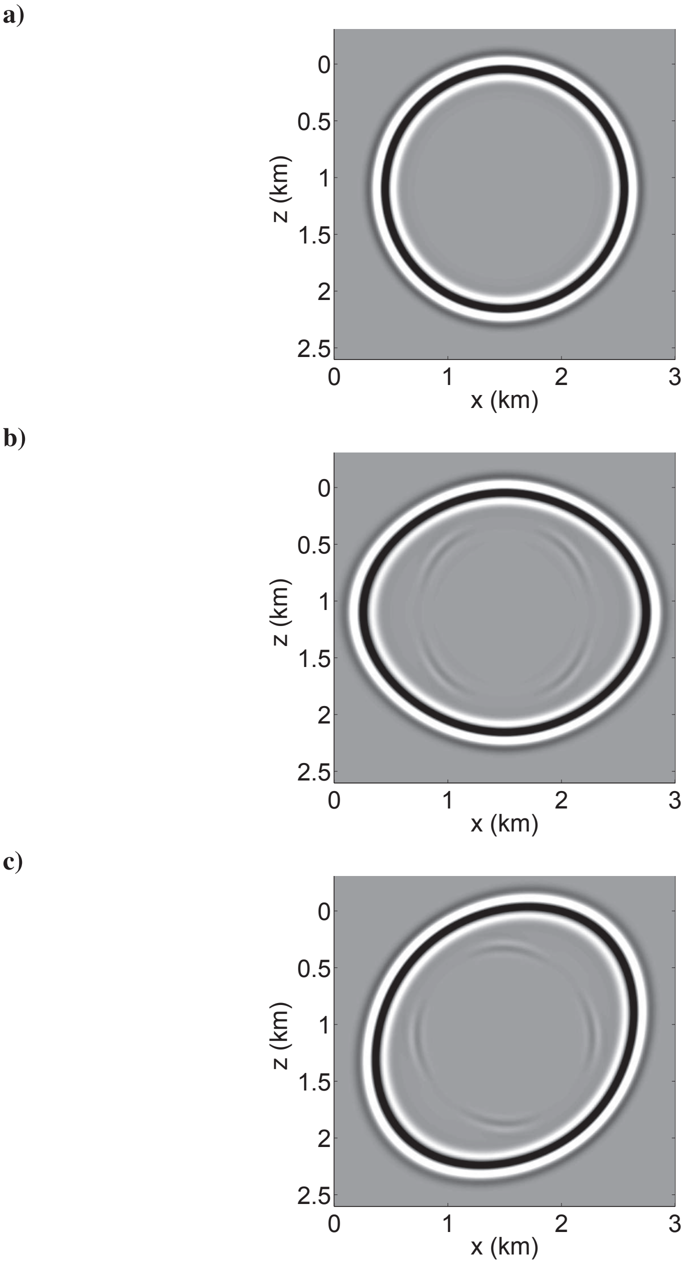


Figure 1.5: Comparison of 2D snapshots of elastic wave propagation over homogeneous (a) isotropic medium, (b) VTI medium, (c) TTI medium. The source is located at the center of the model.

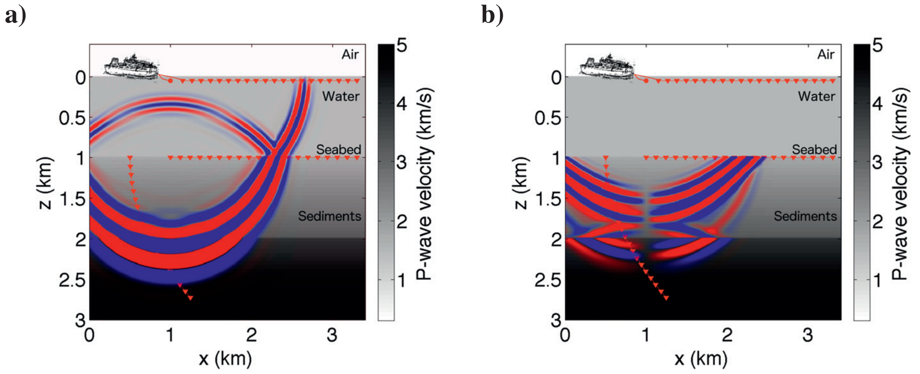


Figure 1.6: Example of P- and S-wave separation in wavefields using equations (1.20) and (1.21). (a) Scalar P-wave wavefield in snapshot of Figure 1.1b; (b) Scalar S-wave wavefield in snapshot of Figure 1.1b.

extended imaging condition producing subsurface horizontal offset CIGs of PS reflections is given by

$$R(\mathbf{x}, \mathbf{h}; s) = \int_0^T w_p^s(\mathbf{x} - \mathbf{h}, t; s) w_s^r(\mathbf{x} + \mathbf{h}, T - t; s) \tag{1.22}$$

1.3 Wave equation migration velocity analysis

When summing all seismic experiments redundant information is gathered, such as those derived from waves reflecting at the same points in the subsurface from different angles. This information can be used to estimate seismic velocities. However, velocity analysis is one the most challenging and time consuming steps in seismic processing. In practice, the problem requires large amount of computation time and human effort to converge to a solution. There is, therefore, a big interest in automatic methods that can solve the problem both accurately and efficiently.

1.3.1 The error function

WEMVA methods propose to automate the velocity analysis through non-linear minimization of an error function based on the seismic image domain (Chavent and Jacewitz, 1995; Shen et al., 2003; Mulder and ten Kroode, 2002; Sava and Biondi, 2004), as opposed to the data domain (Tarantola, 1984; Mora, 1987; Pratt, 1999). One the greatest challenges of non-linear minimization is exactly the non-linearity of the error function. In that respect, the image domain has the advantage that its error function is much more linear. One simple explanation for this fact is that in the image domain the error function explores the redundancy of the seismic data acquisition better than in the data domain. The disadvantage of image domain methods is that they are based purely on reflected data, and assume single scattering (Mulder and van Leeuwen, 2008). Which means that valuable information contained in refractions and multiple reflections is not used.

Implementation of WEMVA differ widely with respect to the way seismic imaging is carried out, and with respect to the way the error is quantified (Sava and Biondi, 2004). Under the isotropic assumption, WEMVA has been implemented based on ray-based Kirchhoff migration (Chauris and Noble, 2001; Mulder and ten Kroode, 2002), one-way wave equation migration (Sava and Vlad, 2008; Shen and Symes, 2008) or RTM (Chavent and Jacewitz, 1995; Mulder, 2008). Similarly, for anisotropic media, there are WEMVA implementations based on ray-based Kirchhoff migration (Brandsberg-Dahl et al., 2003; Behera and Tsvankin, 2009), one-way wave equation migration (Li and Biondi, 2011), or RTM (Li et al., 2012; Weibull et al., 2012). The misfit functions are typically a combination of semblance (a.k.a stack-power, or image-power), and differential semblance. Which, in turn, can be computed based on surface coordinates (Symes and Carazzone, 1991; Symes and Kern, 1994), or on subsurface coordinates (Shen et al., 2003; Shen and Symes, 2008; Mulder, 2008). In this thesis, I implement an error function for WEMVA using subsurface oriented differential semblance, semblance, and RTM. The error function can be written as (Shen and Symes, 2008; Weibull and Arntsen, 2011)

$$\mathcal{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{h}\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right\|^2 - \frac{\gamma}{2}\left\|\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},0)\right\|^2, \tag{1.23}$$

where $\mathbf{m}$ is a vector containing the velocity parameters, $\mathcal{R}(\mathbf{x},\mathbf{h}) = \sum_s R(\mathbf{x},\mathbf{h};s)$ are the stacked CIGs, $\|\cdot\|^2$ corresponds to the least squares norm, and γ is a constant that balances the contribution between differential semblance, given by the first term on the RHS of the equation, and semblance, given by the second term in the RHS of the equation. In principle, the error will be minimum when the energy in the CIGs is maximally focused at the zero subsurface offset. In chapters 2 and 3, I use this objective function to implement WEMVA over, respectively, acoustic and TTI media.

1.3.2 Optimization

The minimization problem can be solved using gradient based optimizers, such as steepest descent, non-linear conjugate gradient or quasi-Newton methods (Nocedal and Wright, 2000). These methods consist in iteratively moving in a descent direction until some convergence criteria is met. At the kth iteration the updated velocity parameters are calculated by

$$\mathbf{m}_{k+1}(\mathbf{x}) = \mathbf{m}_k(\mathbf{x}) - \alpha_i \mathbf{\Delta m}_k(\mathbf{x}), \tag{1.24}$$

where $k \in (1, 2, ..., N)$ is the iteration index, α_k is a positive step length, and $\mathbf{\Delta m}_k$ is a descent direction.

The descent direction $\mathbf{\Delta m}$ is computed from the gradient of the misfit function with respect to the velocity parameters $\nabla_{\mathbf{m}}\mathcal{J}$. The relationship between the gradient and the descent direction depends on the optimization method. The simplest relationship is found in the steepest descent method. In this case, the direction is given by

$$\mathbf{\Delta m}(\mathbf{x}) = -\nabla_{\mathbf{m}}\mathcal{J}(\mathbf{x}). \tag{1.25}$$

Once the direction is computed, the step length α can be computed using a line search (Nocedal and Wright, 2000).

1.3.3 Gradient computation

In problems with a large number of parameters, computing the gradient can be very expensive. However, the adjoint state method (Chavent and Lemonnier, 1974; Plessix, 2006) allows to compute the gradient at a similar cost as that of evaluating the objective function. In the adjoint state method, the optimization of the objective function is seen as a constrained optimization problem. The first step is to set up a Lagrangian function

$$\mathcal{L}(\mathbf{m}, u, u') = \mathcal{J}(\mathbf{m}, u) + \langle e(\mathbf{m}, u), u' \rangle \,. \tag{1.26}$$

Then the gradient is given by

$$\langle \nabla \mathcal{J}, \delta\mathbf{m} \rangle = \frac{\partial \mathcal{L}}{\partial \mathbf{m}}(\mathbf{m}, u, u')\delta\mathbf{m}, \tag{1.27}$$

where u is the solution of the state equation

$$e(\mathbf{m}, u) = 0, \tag{1.28}$$

and u' is the solution of the adjoint state equation

$$\frac{\partial \mathcal{L}}{\partial u}(\mathbf{m}, u, u')\delta u = 0. \tag{1.29}$$

In the above equations, u is called the state, and u' is called the adjoint state, and the state equation $e(\mathbf{m}, u)$ is the equation describing the forward map, for example, the acoustic, or the elastic wave equations. Further detail, and practical implementation of the adjoint state method for acoustic and elastic WEMVA can be found in Chapters 2 and 3 respectively.

1.3.4 Regularization

A problem common to most inverse problems is the inherent ambiguity in the solution. The effect of simplified physics, incomplete acquisition, lack of sensitivity, and the presence of noise, all limit the ability of the inversion to uniquely resolve the subsurface parameters. To mitigate the effects of the non-uniqueness, regularization is often applied (Tikhonov and Arsenin, 1977). The role of regularization is twofold: To avoid solutions that are far from the initial model, and to guide the solution towards the simplest model of the subsurface that explains the data. A typical regularization term is thus given by

$$\mathcal{J}_R = \frac{\epsilon}{2} \left\| \frac{\partial \mathbf{m}}{\partial x_i}(\mathbf{x}) - \frac{\partial \mathbf{m}^0}{\partial x_i}(\mathbf{x}) \right\|^2 , \tag{1.30}$$

where $\epsilon > 0$ is a weighting constant, and $\mathbf{m}^0$ represents the initial model of the velocity parameters.

1.3.5 Implementation

A typical organization of WEMVA can be seen in Figure 1.7. At the initial stage of the velocity analysis, the initial model $\mathbf{m}^{init}$ is provided to the optimizer. At each iteration or line search step the optimizer asks for the value of the error function and its gradient with respect to the model parameters. This process continues until a convergence criteria is met, at which point a final model $\mathbf{m}^{opt}$ is output.

WEMVA is a very computer intensive method, because it requires many iterations of PSDM. Fortunately, migration can be easily parallelized. Since each seismic experiment is independent of each other, parallelization scales with the number of shots. Also, the number of iterations required to converge to a solution can be reduced by using sophisticated optimizers such as the L-BFGS method (Nocedal and Wright, 2000).

1.4 Thesis Structure

This thesis consists of five chapters including this introduction. The core of the thesis are three independent peer-reviewed journal papers. These papers have their own introductions, conclusions and appendixes. Finally, the last chapter presents some concluding remarks and suggestions for possible future work.

1.4.1 Chapter 2 – Automatic velocity analysis with reverse-time migration

In this paper we implement a WEMVA method based on acoustic reverse-time migration. The error function measure the velocity inconsistency with help of differential semblance and semblance misfit functions. To improve the convergence properties of the error function, we introduce a vertical derivative operator which removes low frequency noise from the reverse-time migrated CIGs. The implemented method is tested in 2D synthetic and field data sets based on marine surface seismic acquisition. To improve the results of velocity analysis in the field data case, we introduced a reverse-time migration based demultiple procedure. This work has been presented at the ROSE Meeting in Trondheim in

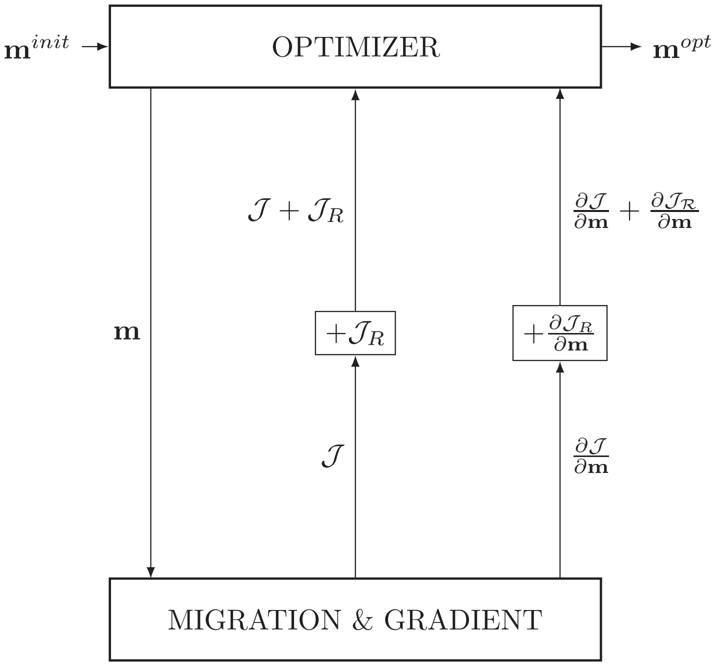


Figure 1.7: Simple flowchart showing the organization of the non-linear optimization in WEMVA.

2011 and at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC in June 2011. Paper accepted for publication in Geophysics 2013.

1.4.2 Chapter 3 – Anisotropic migration velocity analysis using reverse-time migration

In this paper we extend WEMVA to anisotropic media using elastic reverse-time migration. The method uses simple wavefield separation to create P-wave images that are input to an error function based on depth-oriented semblance and differential semblance misfit functions. The method is developed to estimate seismic velocities over 2D TTI media from marine surface seismic data. We show results of the 2D TTI WEMVA method on realistic 2D synthetic and 2D field data sets based on marine surface seismic acquisition. The introduction of anisotropy significantly complicates the ambiguity problem of the solution of the velocity analysis. This problem can not be solved with regularization. Instead the method must be provided with an additional source of information to achieve a unique solution. The method is a first step toward estimation of velocities from ocean bottom seismic data sets. This work has been presented at the ROSE Meeting in Trondheim 2012, at the 74th EAGE Conference and Exhibition incorporating SPE EUROPEC in June 2012, at the SEG/EAGE Summer Research Workshop in July 2012, and at the 81st SEG Annual International Meeting in September 2012. Paper submitted to Geophysics 2013.

1.4.3 Chapter 4 – Reverse-time demigration using the extended imaging condition

In this paper we present a method based on least squares inversion to reconstruct seismic reflection data from reverse-time migrated images. Using reverse-time migration makes the method applicable to complex geological media. Whereas migration using the extended imaging condition preserves the amplitude and phase information of both primaries and multiples, even in case of migration with an inaccurate velocity model. Applications of the method in seismic data reconstruction and image-based demultiple of seismic data are presented. The paper is scheduled for presentation at the 82nd SEG Annual International Meeting in September 2013. Paper submitted to Geophysics 2013.

Chapter 2

Automatic velocity analysis with reverse-time migration

Wiktor Weibull and Børge Arntsen

Norwegian University of Science and Technology, Trondheim, Norway

Abstract

In this paper we apply a method to automatically estimate the background velocities using reverse-time migration. The method uses a combination of differential semblance and similarity-index (a.k.a. semblance or stacking-power) to measure the focusing error in imaging and a non-linear optimization procedure to obtain the background velocities. A challenge in this procedure is that, for media consisting of complex and strongly refracting velocities, artifacts in the reverse-time migrated image (low-frequency noise) can cause the velocity analysis to diverge. We successfully overcome this issue by applying a simple vertical derivative filter to the image that is input to velocity analysis. The resultant velocity analysis method is tested in two 2D synthetic examples and one 2D field data example. Due to the assumptions inherent to prestack depth migration, the data that are input to velocity analysis must be singly scattered. In order to apply the method to multiple-rich data, we propose an image-based demultiple method. The method consists in muting events in the subsurface offset common image point gathers constructed with reverse-time migration, and remodeling the data using a kinematic demigration. A field data example shows how the image-based demultiple of the data helps to improve the velocity analysis in the presence of multiple scattering.

Presented at the ROSE Meeting in Trondheim in 2011 and at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC in June 2011; Paper accepted for publication in Geophysics 2013.

2.1 Introduction

An accurate estimate of the distribution of the subsurface seismic velocities is an indispensable component for obtaining an accurate image of Earth's reflectivity by prestack depth migration. Methods for velocity estimation in routine use today are often based on ray tomography and, although usually semi-automatic, requires time consuming picking of gathers and quality control. Due to the shortcomings of ray-theoretical depth migration approaches in areas with complex geology (Arntsen et al., 2009), one-way and more recently two-way wave equation imaging methods have become popular. These approaches are often combined with velocity estimation based on ray-theory. For consistency and improved resolution, velocity model building and seismic imaging should preferably be based on wave equation methods.

Wave equation migration velocity analysis (WEMVA) is based on focusing of seismic reflection data in the image domain and uses an automatic optimization procedure to estimate the velocity field, avoiding manual picking. The approach consists in formulating an objective function measuring to what extent subsurface offset- or angle-gathers are, respectively, focused or flattened, and then minimizing the objective function with respect to the velocity field.

Chavent and Jacewitz (1995) implemented WEMVA by using a similarity-index and reverse-time migration (RTM) to compute the velocity field. The procedure utilizes the complete wavefield and requires no picking. Biondi and Sava (1999) used one-way migration operators and image perturbations for computing corrections to the initial wavefield and Sava and Biondi (2004) extended this approach to a fully non-linear iterative scheme. The numerical implementation is described in detail by Sava and Vlad (2008).

Shen et al. (2003) used the Double Square Root approach to depth migration and an objective function based on differential semblance (Symes and Carazzone, 1991) to estimate the velocity field. The approach was extended to shot-profile migration based on one-way migration operators, through an objective function consisting of the difference between differential semblance and similarity-index (Shen and Symes, 2008). Mulder (2008) used depth migration based on the two-way wave equation in the frequency-domain and an objective function related to the differential semblance cost function divided by the similarity-index cost function to implement a non-linear scheme for computing the velocity field. Gao and Symes (2009) proposed to use a differential semblance cost function and RTM to solve the velocity estimation problem, and also gave an initial theoretical framework.

We implement an objective function for wave-equation migration velocity analysis using differential semblance, similarity-index, and RTM. In order to minimize problems related to amplitude-sensitivity caused by the two-way wave equation, we choose to modify the image by a spatial differentiation operator, similar to but simpler than the filter proposed by Mulder (2008). As we will show, this leads to an objective function with improved convergence properties. We also give complete expressions for computing the gradient of the cost function with respect to the velocity field, and show how the gradient can be used in a full non-linear optimization scheme illustrated with both synthetic and real data examples.

One challenge in applying WEMVA to field data is the presence of free-surface multiples (van Leeuwen and Mulder, 2008; Mulder and van Leeuwen, 2008). To overcome this problem, Mulder and van Leeuwen (2008) proposed to modify their objective function with an asymmetric weighting function. We extend this idea and develop a method of multiple attenuation based on reverse-time migration, muting and data reconstruction, and show how this method helps in improving the results of velocity analysis in the case of multiple-rich data.

Reverse-time migration and velocity analysis

2.1.1 Migration

In RTM, a Common Image Point gather (CIG), $\mathcal{R}$, can be produced by crosscorrelating a forward modeled source wavefield (D) with a reverse-time modeled scattered wavefield (U)

$$\mathcal{R}(\mathbf{x},\mathbf{h}) = \sum_s \int dt\, U_s(\mathbf{x}+\mathbf{h},t) D_s(\mathbf{x}-\mathbf{h},t), \tag{2.1}$$

where $\mathbf{x} = (x_1, x_2, x_3)$ is the subsurface mid-point coordinate, with x_3 being the vertical depth axis coordinate, $\mathbf{h} = (h_1, h_2, 0)$ is the subsurface horizontal half-offset, t is the time and s is the source index.

This imaging condition is an extension of Claerbout's principle (Claerbout, 1971; Rickett and Sava, 2002). According to this principle, given an accurate estimate of the material velocities, the crosscorrelation of the reconstructed source and receiver wavefields will have a maximum at zero lag in time and space. By parameterizing the image with an additional lag parameter we can capture the deviation of the maximum in crosscorrelation from zero lag, and use this to quantify the error in the estimates of the velocities. In this paper we only consider horizontal spatial lags in the crosscorrelation, however vertical spatial lags, and even temporal lags can be incorporated in the image (Biondi and Shan, 2002; Biondi and Symes, 2004; Sava and Fomel, 2006). The U_s and D_s wavefields are obtained through

$$D_s(\mathbf{x},t) = \int d\mathbf{x}'\, G(\mathbf{x},t;\mathbf{x}',0) * \sum_{sou=1}^{Nsou} \delta(\mathbf{x}'-\mathbf{x}_{sou}) S_s(\mathbf{x}',t), \tag{2.2}$$

and

$$U_s(\mathbf{x},t) = \int d\mathbf{x}'\, G(\mathbf{x},0;\mathbf{x}',t) * \sum_{rec=1}^{Nrec} \delta(\mathbf{x}'-\mathbf{x}_{rec}) P_s(\mathbf{x}',t), \tag{2.3}$$

where G is the Green's function for the constant-density two-way acoustic wave equation, and $*$ means time convolution. Here, P is the recorded reflection data and S is the source data, while $\mathbf{x}_{sou}$ are source datum coordinates for one shot, and $\mathbf{x}_{rec}$ are the receiver datum coordinates for one shot.

2.1.2 Velocity analysis

The velocity analysis is based on the optimization of the following objective function

$$\mathcal{J} = \mathcal{DS} - \mathcal{SI} + \mathcal{J}_{reg}. \tag{2.4}$$

The objective function is composed of three parts, the differential semblance misfit ($\mathcal{DS}$), the similarity-index ($\mathcal{SI}$) and a regularization term ($\mathcal{J}_{reg}$).

An important assumption in WEMVA is that when the velocity model is optimum, the CIGs are maximally focused at zero subsurface offset. Thus any deviation from perfect focus can be taken as an indication that the background velocity model must be improved. The differential semblance misfit represents a simple and direct way of quantifying the deviation from focus of CIGs

$$\mathcal{DS} = \frac{1}{2}\left\|\mathbf{h}\frac{\partial\mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right\|^2 = \frac{1}{2}\int d\mathbf{x}\int d\mathbf{h}\,\mathbf{h}^2\left[\frac{\partial\mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right]^2. \tag{2.5}$$

This differential semblance formulation is similar to that in Shen and Symes (2008). However, the fact that we use RTM to construct the image makes a fundamental difference. The solution of the one-way wave equation neglects any scattering from vertical contrasts in the velocities, while the same is not true for the solution of the two-way wave equation. The scattering during wavefield reconstruction produces undesired artifacts in the final RTM image. In the context of velocity analysis it is desirable to remove these artifacts, since they are coherent events and particularly sensitive to changes in the velocities, thus affecting the results of the optimization. Due to the low-wavenumber character and the predominantly vertical orientation of this effect, a simple vertical derivative filter acting over the image is sufficient to remove it (Guitton et al., 2007). A similar procedure, albeit involving multiplication in the wavenumber domain, has been applied by Mulder (2008) in his implementation of WEMVA using the two-way wave equation.

The differential semblance measure is based solely on kinematic considerations. To exploit the dynamic effect that the velocities can have through the improvement of the stack quality, the objective function can be augmented with the similarity-index (a.k.a. stacking-power or semblance) (Chavent and Jacewitz, 1995)

$$\mathcal{SI} = \frac{\gamma}{2}\left\|\frac{\partial\mathcal{R}}{\partial x_3}(\mathbf{x},0)\right\|^2 = \frac{\gamma}{2}\int d\mathbf{x}\left[\frac{\partial\mathcal{R}}{\partial x_3}(\mathbf{x},0)\right]^2. \tag{2.6}$$

where, γ is a constant that balances the weight of $\mathcal{SI}$ over $\mathcal{DS}$. Ideally, the weight should be chosen such that the similarity-index only acts as a regularization.

Finally, to further improve the wellposedness of the velocity analysis we also apply regularization (Tikhonov and Arsenin, 1977)

$$\mathcal{J}_{reg} = \frac{\alpha}{2}\left\|\frac{\partial v}{\partial \mathbf{x}}(\mathbf{x})\right\|^2 + \frac{\beta(\mathbf{x})}{2}\left\|v(\mathbf{x}) - v_{prior}(\mathbf{x})\right\|^2, \qquad (2.7)$$

where α is a scalar constant, β is a vector of constant weights, and v_{prior} is the vector containing a priori known values of velocity.

The optimization of equation 2.4 is performed using a gradient-based non-linear method (Byrd et al., 1995; Nocedal and Wright, 2000). The gradient of the misfit-function with respect to velocity is then required.

The gradient of equation 2.4 is given by

$$\nabla_v \mathcal{J}(\mathbf{x}) = \nabla_v \mathcal{DS}(\mathbf{x}) - \nabla_v \mathcal{SI}(\mathbf{x}) + \nabla_v \mathcal{J}_{reg}(\mathbf{x}). \qquad (2.8)$$

The adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Chavent, 2009) offers an exact and efficient way to compute the gradients of the differential semblance misfit and the similarity-index functions. In Appendix A, we show how to derive the gradient of equations 2.5 and 2.6 with respect to the acoustic velocities by the adjoint state method. In this case the gradient is given by

$$\begin{aligned}\nabla_v(\mathcal{DS} - \mathcal{SI})(\mathbf{x}) = &-\sum_s \int dt \frac{2}{v^3(\mathbf{x})}\frac{\partial^2 D_s}{\partial t^2}(\mathbf{x},t)D'_s(\mathbf{x},t)\\ &-\sum_s \int dt \frac{2}{v^3(\mathbf{x})}\frac{\partial^2 U_s}{\partial t^2}(\mathbf{x},t)U'_s(\mathbf{x},t). \qquad (2.9)\end{aligned}$$

The wavefields D'_s and U'_s are adjoint states associated with the constraints that the direct states (D_s and U_s), satisfy the constant density acoustic wave equation. These wavefields can be computed by the following adjoint modelings

$$D'_s(\mathbf{x},t) = \int \mathbf{dx}'\, G(\mathbf{x},0;\mathbf{x}',t) * \int d\mathbf{h}\, (\mathbf{h}^2 - \gamma\delta(\mathbf{h}))\frac{\partial^2 \mathcal{R}}{\partial x'^2_3}(\mathbf{x}'+\mathbf{h},\mathbf{h})U_s(\mathbf{x}'+2\mathbf{h},t) \qquad (2.10)$$

and

$$U'_s(\mathbf{x},t) = \int \mathbf{dx}'\, G(\mathbf{x},t;\mathbf{x}',0) * \int d\mathbf{h}\, (\mathbf{h}^2 - \gamma\delta(\mathbf{h}))\frac{\partial^2 \mathcal{R}}{\partial x'^2_3}(\mathbf{x}'-\mathbf{h},\mathbf{h})D_s(\mathbf{x}'-2\mathbf{h},t) \qquad (2.11)$$

where δ is the Kronecker delta.

The cost of computing the gradient in this way is approximately the same as that of evaluating the misfit function, however, it requires the state variables (D and U) to be stored for each shot, which can be expensive. In the discussion we suggest some measures to reduce this cost.

To complete the gradient of $\mathcal{J}$ we need to compute the gradient with respect to the regularization term

$$\nabla_v \mathcal{J}_{reg}(\mathbf{x}) = \beta(\mathbf{x})(v(\mathbf{x}) - v_{prior}(\mathbf{x})) - \alpha \frac{\partial}{\partial \mathbf{x}} \left[\frac{\partial v}{\partial \mathbf{x}}(\mathbf{x}) \right], \tag{2.12}$$

where $\frac{\partial v}{\partial \mathbf{x}}$ is taken to be zero at the boundaries.

2.1.3 Velocity preconditioning

The solution of differential semblance optimization is notoriously rough. Fei and Williamson (2010) indicated that the updated velocities can have artificial roughness features (vertical stripes). To ensure a smooth solution to the velocity analysis, we precondition the velocity model. In addition to speeding up the convergence, preconditioning also helps to make the velocity analysis well-posed, as it reduces the space of possible solutions and the number of parameters to be estimated. A popular choice of representing a velocity model is given by the cubic B-spline representation (Dierckx, 1993)

$$v(\mathbf{x}) = \sum_{\mathbf{i}} c_{\mathbf{i}} B_{\mathbf{i}}(\mathbf{x}), \tag{2.13}$$

where $B_{\mathbf{i}}$ are cubic splines defined at predetermined control points and $c_{\mathbf{i}}$ are coefficients to be determined by the velocity analysis. One advantage of using this representation is that it ensures continuous second derivatives, which is important for the derivative regularization implementation. At the same time, due to the local support of the cubic spline functions, this representation is also well suited to describe the spatial variations necessary for velocity analysis. When using the B-spline representation for optimization, the gradient must be transformed from the Cartesian to the spline basis

$$\nabla_c \mathcal{J}(\mathbf{i}) = \int d\mathbf{x}\, B_{\mathbf{i}}(\mathbf{x}) \nabla_v \mathcal{J}(\mathbf{x}). \tag{2.14}$$

Results

2.1.4 Synthetic data examples

We present the results of optimizing two 2D synthetic data sets. Both synthetic data sets were generated using 2D finite difference modeling, with synthetic density and acoustic velocity models. The geometry simulates a typical marine acquisition. We use a monopole point source and a Ricker wavelet with a peak frequency of 25Hz. The modeling was carried out with an algorithm that is 8th order accurate in space and 2nd order accurate in time (Virieux, 1986). The modeling aperture is taken as twice the cable length. To avoid reflections from the boundaries of the models, PML absorbing boundary conditions were implemented at all sides (Qin et al., 2009). This way, the resulting data sets are devoid of free-surface multiples. Although interbed multiples are still present in the data, they are significantly weaker than the primaries and are neglected in these examples. Preprocessing of the data sets consisted in muting the direct wave and the refracted waves at the receiver level.

Optimization is carried out with a L-BFGS method (Byrd et al., 1995). In both examples, the regularization parameter α was spatially invariable, and its value was chosen such that the initial derivative regularization error value was 1% of the initial $\mathcal{DS}$ value. The value of parameter β was chosen in a similar way as α, but it was set to zero outside of the topmost layer.

Shallow gas accumulation and leakage model

The first example consists of a 4-km long and 1-km deep model shown in Figure 2.1. This model simulates a dipping layered sediment succession with a small 'reef' like structure in the center. The density varies between 1800 and 2400 kg/m^3. The velocity model has the same structure as the density model, but with a localized low velocity lens (Gaussian with a peak of -500 m/s) under the reef structure.

This model simulates a scenario of shallow gas accumulation and leakage, a typical situation that, if not accurately predicted by the velocity model, can produce significant distortion on depth images.

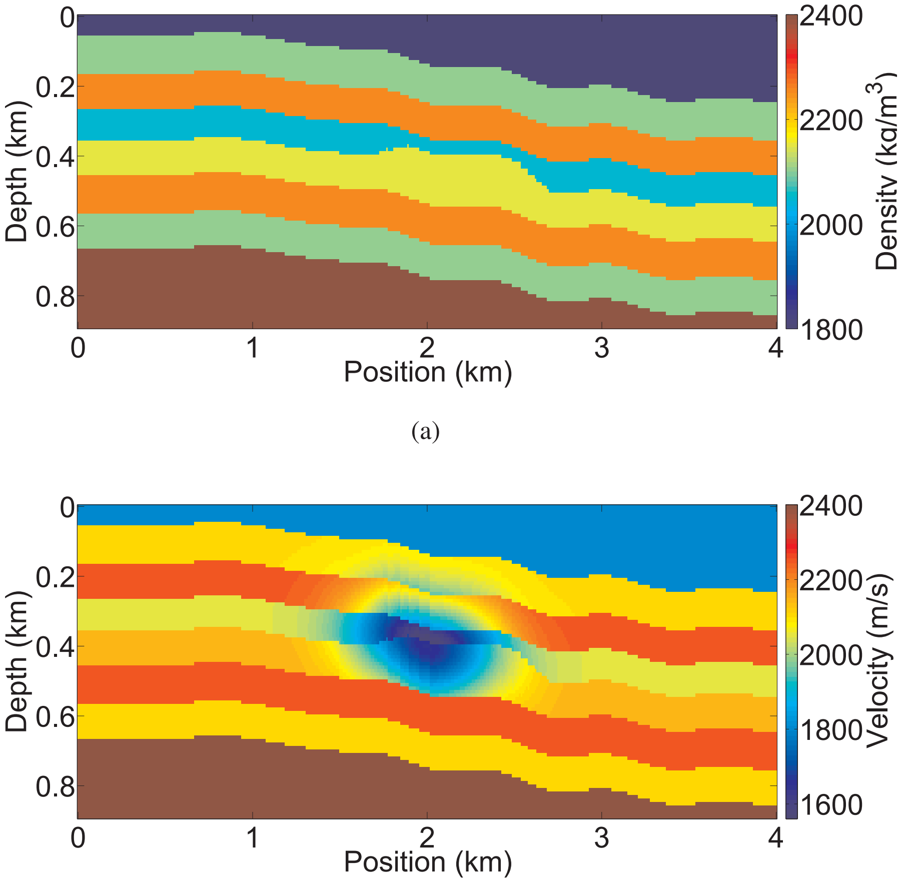
The acquisition geometry simulates a 2D marine acquisition with a 311 sources separated by 20 m. The cable length is 3.2 km with 10-m channel interval, with minimum offset of 500 m. Recording time length is 2 s. The initial model used for migration consists of a linear 1D velocity profile varying from 1.8 km/s to 2.4 km/s as shown in Figure 2.2a. For optimization, a bicubic B-spline spline with control points every 40 m in both x and z directions was then fitted to this initial model. The parameter γ was chosen so that the initial $\mathcal{SI}$ value was 20% of the initial $\mathcal{DS}$ value.

The CIGs are produced according to equation 2.1 with the half-offset axis varying between ± 400 m (81 offset samples).

The quality of our initial estimate of the velocity can be quantified by looking at the initial image shown in Figure 2.2b and the CIGs output by RTM displayed in Figure 2.2c. The image is obtained from the CIGs by taking only the zero subsurface offset component (i.e. zero lag crosscorrelation), and shows significant distortion, especially below the low velocity lens. At the same time, the energy in CIGs is significantly spread across the offset axis.

To justify the use of the spatial derivative in our modified differential semblance misfit function, we make two attempts at estimating velocities from this data. One without the spatial derivative, and other with our modified differential semblance objective.

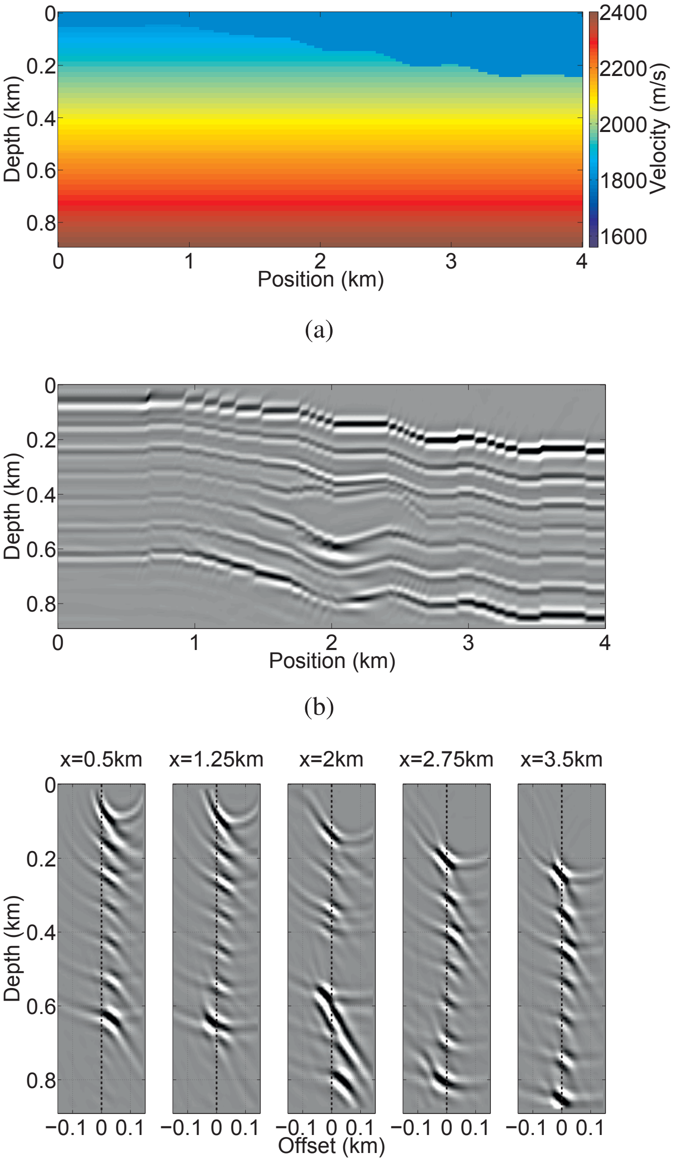
In the first attempt, we did not apply the spatial derivative. Optimization was stopped after 30 iterations because, despite the decrease in value of the objective function, the



(a)

(b)

Figure 2.1: Shallow gas accumulation and leakage synthetic model: (a) Density model (kg/m^3). (b) Acoustic velocity model (m/s).



(a)

(b)

(c)

Figure 2.2: (a) Initial velocity model (m/s). (b) Initial image. (c) Initial subsurface offset CIGs.

quality of the migrated image was getting worse with increasing number of iterations. This was taken as an indication that velocity analysis was converging to a non-optimal minimum. The updated velocities are shown in Figure 2.3a, the updated image is shown in Figure 2.3b and the updated subsurface CIGs are shown in Figure 2.3c. The optimization fails here because we allow too strong spatial variations in the updated velocities. These introduce artifacts in the reverse-time migrated image that are more sensitive to the velocities than the kinematic errors that we want to correct for.

In the second attempt we introduce the spatial derivative in the image. Now optimization converges after 17 iterations. The optimized velocity model is shown in Figure 2.4a, and it is clear that the low velocity lens and the layered structure of the velocity model are successfully detected.

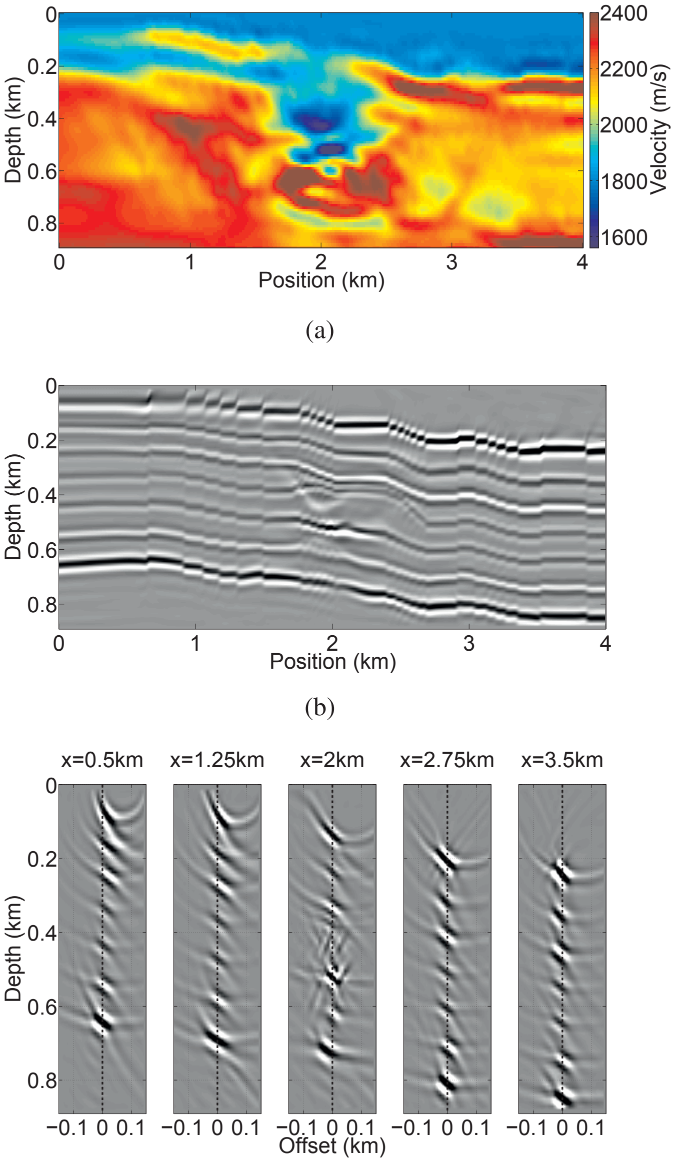
Figure 2.4b shows the results of the optimization on the image, while Figure 2.4c shows the optimized CIGs. The image migrated with the updated velocities is much better focused and the reflectors are now well positioned. At the same time, the CIGs are well focused at zero subsurface offset, indicating that the velocities are adequate to describe the kinematics of the data.

Figure 2.5a shows a comparison of the initial, updated and true traces of velocity at different spatial positions, while Figure 2.5b does the same for the reflectivity. The velocity traces show that, locally, the updated velocities can deviate significantly from the true velocities, though without compromising the reflectivity fit. This demonstrates the non-uniqueness inherent to the solution of this type of problem. From the reflectivity, we can see that the method is able to correct mispositioning errors that are, at times, larger than half-wavelength (see for example Figure 2.5b at $x = 2$ km).

In yet another experiment with this dataset, we compare the results of velocity analysis using the differential semblance misfit alone, the similarity-index alone and their combination. For this test, a bicubic B-spline with control points every 100 m in x direction and 50 m in the z direction was fitted to the initial model. In the combination result, the parameter γ Âăis chosen such that the initial $\mathcal{SI}$ value is equal to 50% of the initial $\mathcal{DS}$ Âăvalue.

The updated velocities obtained with the three different objective functions are shown in Figure 2.6, while the updated images are displayed in Figure 2.7. The results show that, in this example, all three objective functions converge to models that improve the quality of the initial migrated image. As can be seen in Figures 2.6a-b, the velocities obtained from velocity analysis based on $\mathcal{SI}$ and $\mathcal{DS}$ can be quite different. This reflects, in part, the fact that these objective functions have different sensitivities to the initial model. However, it is also a consequence of using different sources of information to constrain the velocity model.

Combining both objective functions seems to have the effect of averaging out artifacts and strengthening similarities between the different velocity models (Figure 2.6c). This helps to improve the quality of the final migrated image, as can be seen in Figure 2.7c.

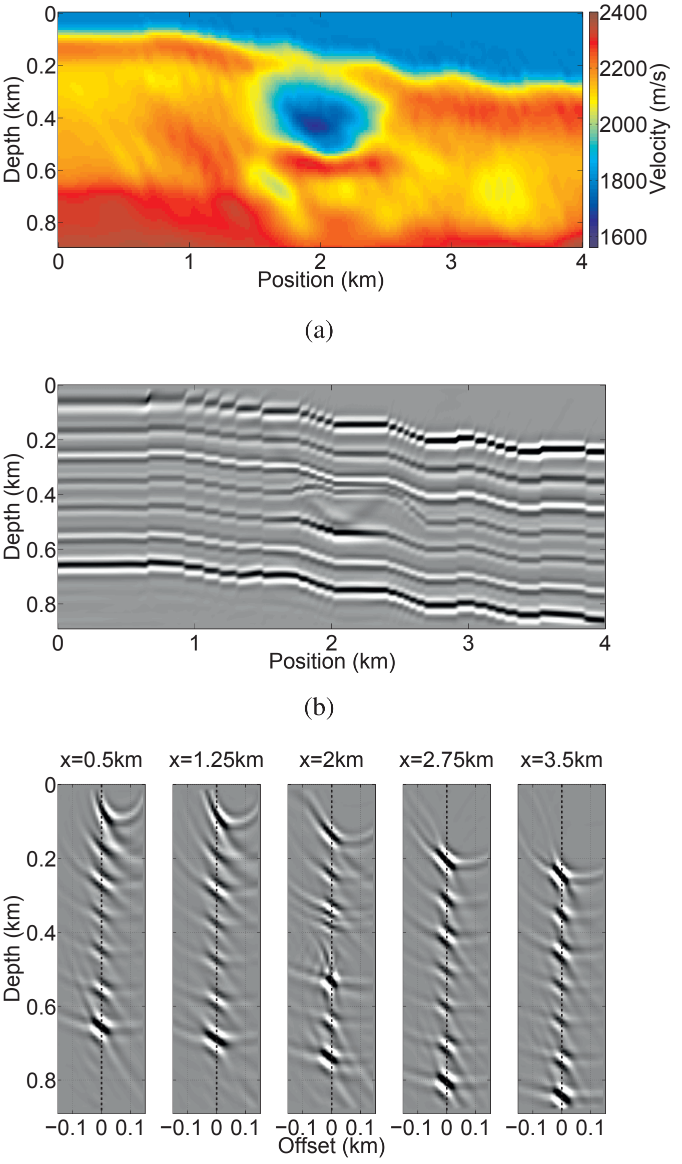


(a)

(b)

(c)

Figure 2.3: Updated velocity model (a), updated image (b), and updated subsurface offset CIGs (c) after 30 iterations of velocity analysis using the objective function without the spatial derivative.

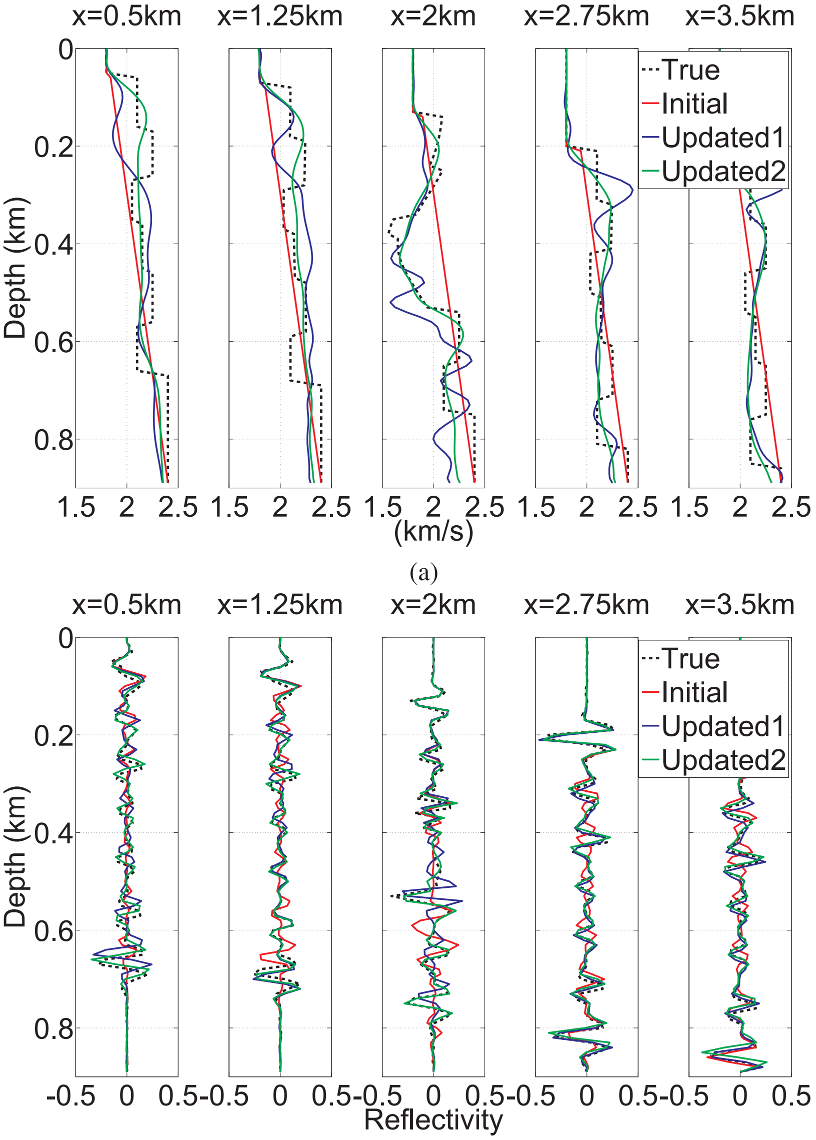


(a)

(b)

(c)

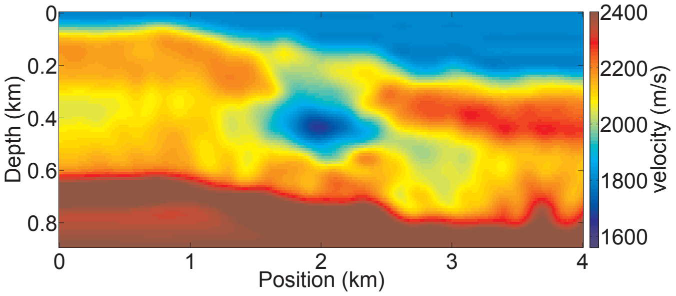
Figure 2.4: Updated velocity model (a), updated image (b), and updated subsurface CIGs (c) after 17 iterations of velocity analysis with the spatial derivative.



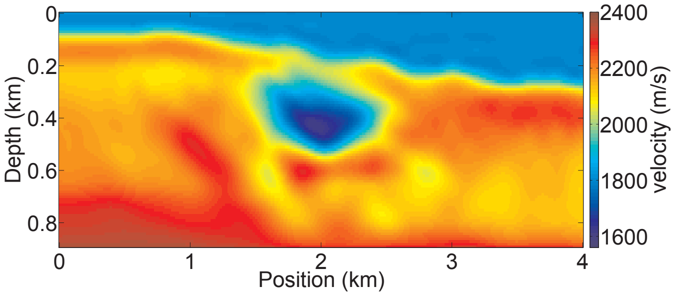
(a)

(b)

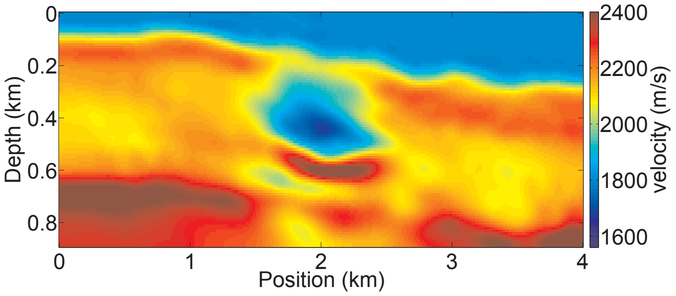
Figure 2.5: (a) Comparison of velocity traces. (b) Comparison of reflectivity traces. Updated 1 corresponds to the updated model and reflectivity from the optimization of the objective function without the vertical derivative, while updated 2 comes from the optimization of the objective function with the vertical derivative.



(a)

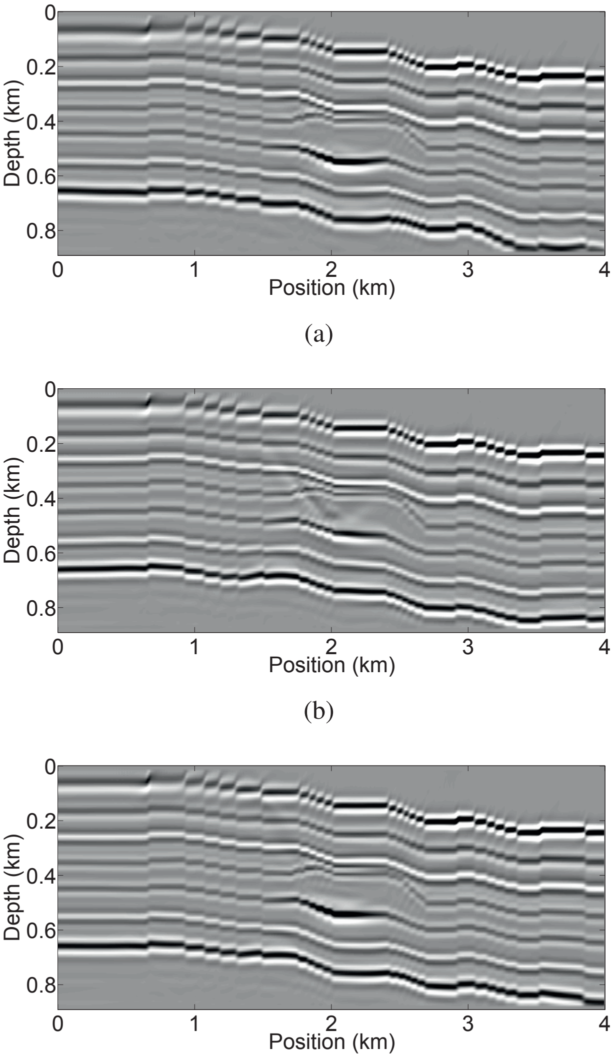


(b)



(c)

Figure 2.6: Updated velocity model using: (a) Similarity-index alone; (b) Differential semblance alone; (c) Differential semblance and similarity-index combined.



(a)

(b)

(c)

Figure 2.7: Updated image using: (a) Similarity-index alone; (b) Differential semblance alone; (c) Differential semblance and similarity-index combined.

Gullfaks model

The second data set is generated from a 2D synthetic model of the Gullfaks oil field, located in the Norwegian margin of the North Sea. The model is 3 km deep and 4 km across. At the top of the model there is a 200-m deep water layer. Beneath the water layer and down to about 1.8 km, the model consists largely of a layered sediment succession. The layered overburden unconformably overlies a set of rotated fault blocks. The density varies between 1200 and 2400 kg/m^3 (Figure 2.8a), while the velocity varies between 1.48 and 3.4 km/s (Figure 2.8b).

The geometry of the data in this example is a half-spread marine setup with with offsets ranging from 0 to 6 km, and spaced every 12.5 m. The recording time length is 4 s.

The initial model used for migration consists of a 1D velocity profile linearly varying from 1.48 to 3.4 km/s (Figure 2.9a). This model carries large deviations from the true model, and these stretch over several hundred meters. This results in significant traveltime errors (more than one wavelength at the considered frequencies). For optimization, a bicubic B-spline with control point spacings of 150 m in x direction and 50 m in z direction was fitted to this model. The parameter γ was chosen so that the initial $\mathcal{SI}$ value was 20% of the initial $\mathcal{DS}$ value. Migration and optimization are carried out in a similar way as in the first example.

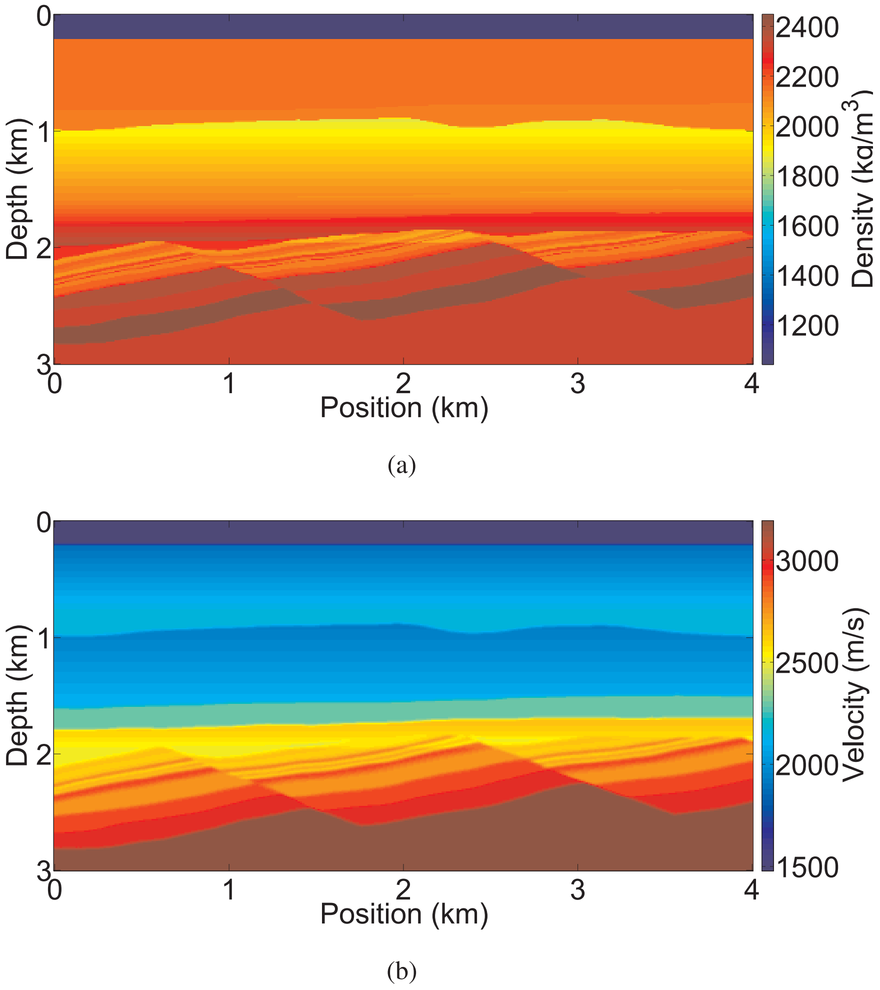
The results of migration with the initial model are shown in Figures 2.9b and 2.9c.

Despite the simplicity of the Gullfaks model, consisting mostly of flat layers with only mild variations in velocity, the convergence rate for this example is relatively slow. We interpret this as a consequence of the fact that there are not many reflectors in the upper 1.8 km of the model, which makes the inversion poorly constrained. Nevertheless, the optimization is stopped after 50 iterations, at which point the CIGs were deemed sufficiently focused.

The optimized velocity model is shown in Figure 2.10a. The results of optimization show that again, in this case the method is capable of improving the kinematics of the velocity model. This can be clearly seen if we compare the initial images, shown in Figures 2.9b and 2.9c, to the updated images, shown in Figures 2.10b and 2.10c. At the same time, Figures 2.11a and 2.11b show that the updated velocities and reflectivities are close to their true value (from the synthetic model).

2.1.5 Field data example

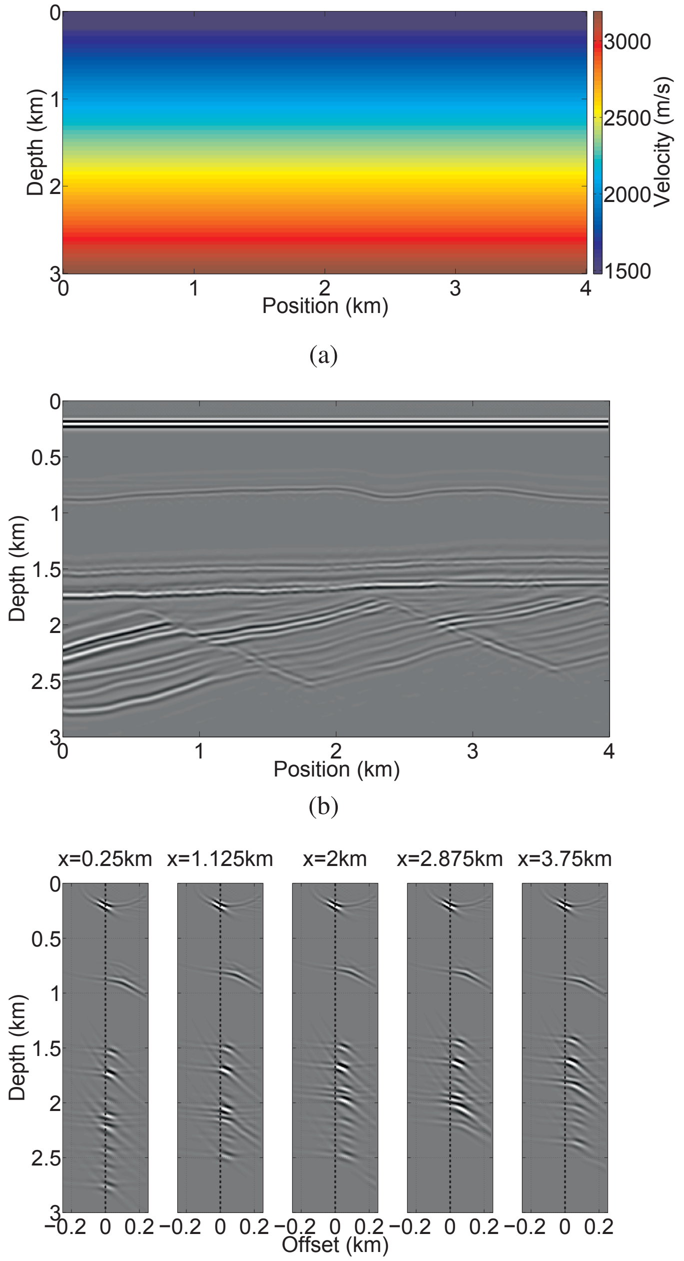
The method is tested on a field data set taken off the Norwegian North Sea. The data are originally a 3D data set, from which we extracted a 2D line. The geometry of the data consists of a line with minimum offset of 150 m and maximum offset of 5 km. The original receiver interval is 25 m and the original shot interval is 18.7 m. The data processing included multiple removal, and muting of direct wave, wide-angle reflections



(a)

(b)

Figure 2.8: Gullfaks synthetic model: (a) Density model (kg/m^3). (b) Acoustic velocity model (m/s).

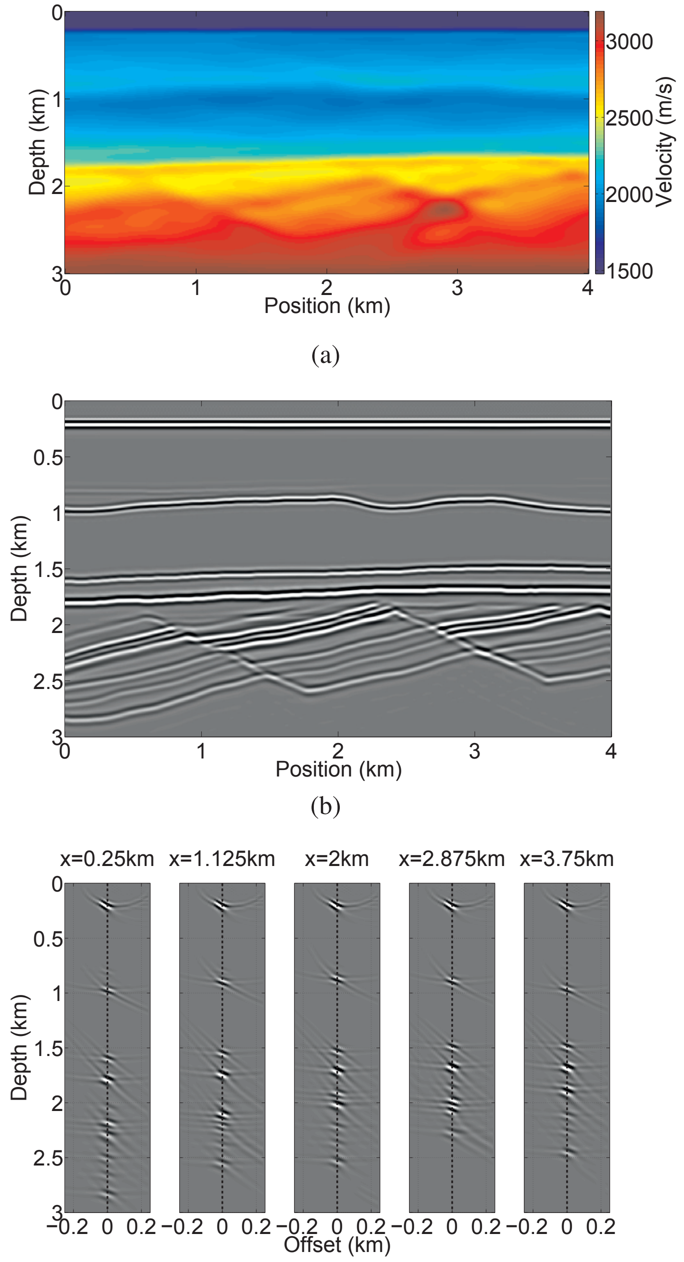


(a)

(b)

(c)

Figure 2.9: (a) Initial velocity model (m/s). (b) Initial image. (c) Initial subsurface offset CIGs.

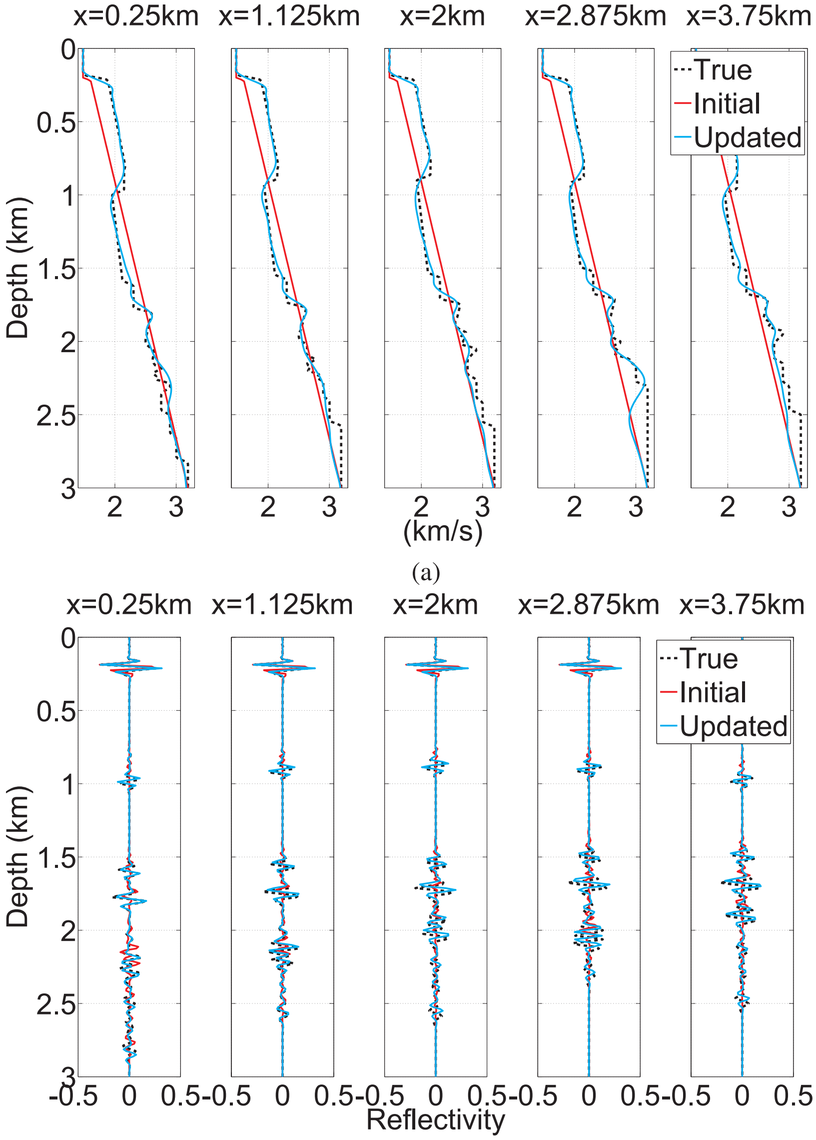


(a)

(b)

(c)

Figure 2.10: Updated velocity model (a), updated image (b), and updated subsurface CIGs (c) after 50 iterations of velocity analysis.



(a)

(b)

Figure 2.11: (a) Comparison of velocity traces. (b) Comparison of reflectivity traces.

and refractions. To increase the contribution from far offsets and deeper events, a power of 2 time gain (t^2) is applied to the data. The maximum frequency of the data was filtered down to 30 Hz, so that a grid of 20 by 20 m could be used for modeling and migration. However, for display purposes, the updated images are migrated on a finer grid of 10 by 10 m using frequencies up to 80 Hz. The choice of regularization parameters (α, β and γ) for optimization follow the same guidelines as in the synthetic examples.

The starting point for the velocity analysis is a 1D velocity model shown in Figure 2.12a. The model is constructed from a smoothed well log of P-wave velocities. For optimization, this initial model was fitted to a B-spline representation using a grid of control points spaced 600 m in the x direction and 100 m in the z direction.

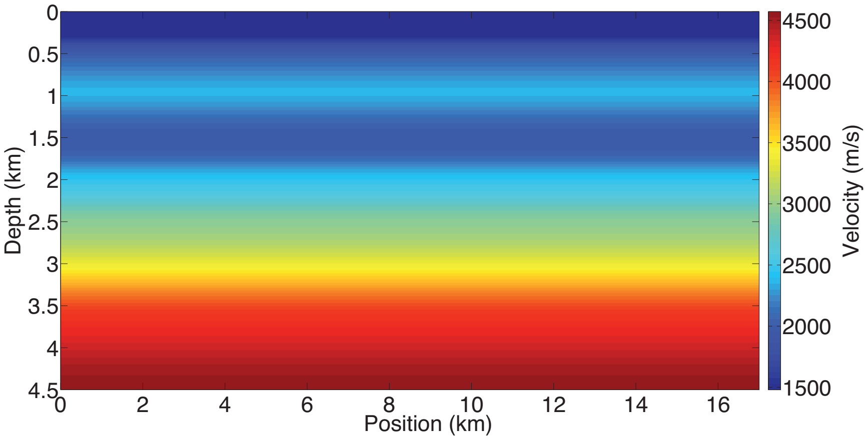
The initial image is shown in Figure 2.12b. Due to the approximately plane layered overburden, the initial image shows relatively well focused reflectors. Indeed the biggest challenge for velocity analysis in this data is the presence of resilient free-surface multiples that were not properly attenuated in preprocessing. To overcome this problem we developed a method based on identifying and muting the multiples directly on the subsurface offset CIGs. The muted CIGs are used in a demigration procedure that kinematically reconstructs the reflection data. The resultant demultipled data can then be used as input for the velocity analysis.

To identify the free-surface multiples in the subsurface offset CIGs, we first note that that these multiples focus at lower velocities than the primaries (Mulder and van Leeuwen, 2008). Thus, when migrating a data with free-surface multiples using an initial velocity model that is closer to the primaries, the multiples will, in general, appear as 'smiling events'. However, in the particular case of single-spread acquisition, events migrated with non-optimal velocities are also asymmetrically shifted relative to the zero subsurface offset. In the case of a source to the left of the streamer, the events that require lower velocity to focus will be shifted to the left, while those requiring higher velocity will be shifted to the right. Whereas Mulder and van Leeuwen (2008) introduce a bias towards higher velocities in the penalization of CIGs, we propose to mute the multiples and reconstruct the data as a preprocessing step. This is similar to the multiple attenuation procedure proposed by Sava and Guitton (2005), where multiples are attenuated by a dip filter in the subsurface angle gathers. The main differences here are that the multiple attenuation is carried out by a mute in the subsurface offset domain, and that, after the CIGs are muted, the data are reconstructed by 'demigrating' the muted subsurface offset CIGs. The data can be reconstructed from the muted CIGs according to

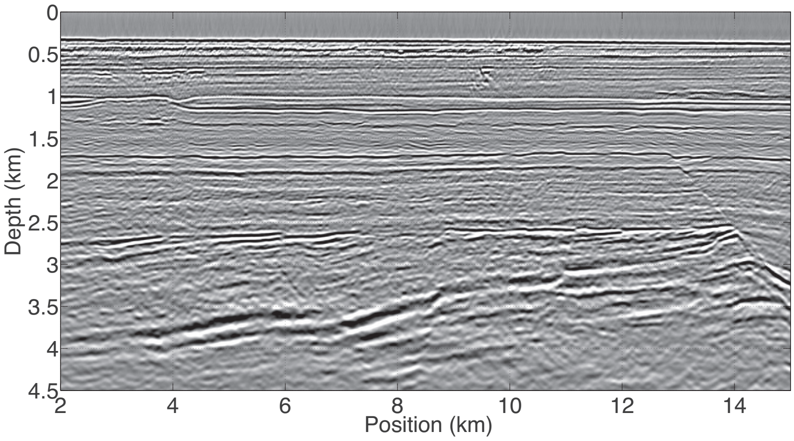
$$U_s(\mathbf{x_{rec}}, t) = \int \mathbf{dx}' \, G(\mathbf{x_{rec}}, t; \mathbf{x}', 0) * \int d\mathbf{h} \, \frac{\partial^2 \mathcal{R}_m}{\partial {x'}_3^2}(\mathbf{x}' - \mathbf{h}, \mathbf{h}) D_s(\mathbf{x}' - 2\mathbf{h}, t), \quad (2.15)$$

where $U_s(\mathbf{x_{rec}}, t)$ is the demigrated demultipled data at the receiver positions $\mathbf{x_{rec}}$, and $\mathcal{R}_m$ are the muted subsurface offset CIGs.

The data reconstruction avoids possible instabilities associated with the energy in the CIGs being moved in and out of the mute during the optimization procedure, which can



(a)



(b)

Figure 2.12: (a) Initial velocity model (m/s). (b) Image constructed using initial velocity model.

occur in the case of a 'static' weighting applied directly in the objective function (Mulder and ten Kroode, 2002; Mulder and van Leeuwen, 2008).

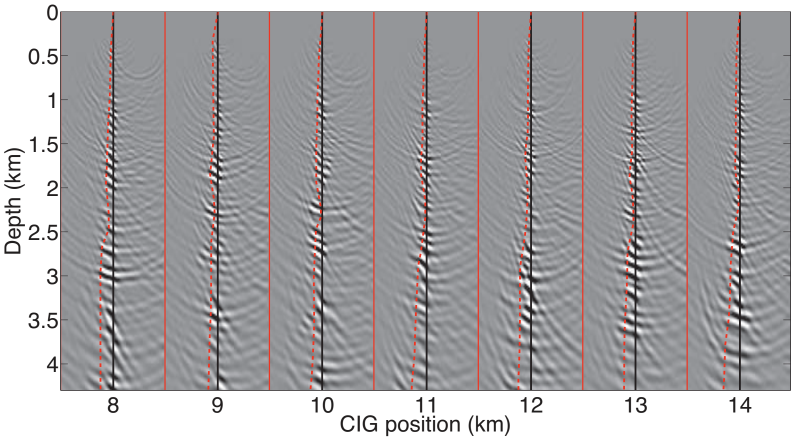
Note that the demultiple procedure is be done only once, as a preprocessing step, and requires no modification of the objective function. The cost of the demultiple is equivalent to the cost of one and a half reverse-time migrations, since the source wavefields, D_s, can be stored during the migration step (equation 2.1), and reaccessed later during the data reconstruction (equation 2.15).

The demultiple procedure requires picking a mute that separates multiple from primary events in the CIGs. One way to pick this mute is to perform an initial velocity analysis, where the goal is to converge to a velocity in between primaries and multiples. The primaries and multiples could then be separated in the resultant CIGs by a vertical mute at zero subsurface offset (Li and Symes, 2007). Here, we follow a more subjective approach, and pick a mute based on our own interpretation. As an example of the application of the demultiple procedure, Figure 2.13a shows a set of subsurface offset CIGs generated using the initial velocity shown in Figure 2.12a. The CIGs are separated by red vertical lines, while the black vertical lines show the position of the zero subsurface offset, and the dotted red lines show the position of picked mutes. The mutes are picked to the left of the strongest events in the CIGs, which are interpreted to be primaries. However, our experience with this data suggests that some strong events, in particular those between 2.5 and 3 km depth, are multiples and are, therefore, also included in the mute. Figure 2.13b shows the CIGs after the mute is applied. Whereas, Figures 2.14a-b show the data before and after the demultiple procedure. Note how low velocity events, indicated by steep moveouts, have been attenuated while the kinematics of primary events are preserved. This example shows that, although subjective, picking the mute based on interpretation allows great flexibility, and can be used to remove multiples and any other events that can be prejudicial to the velocity analysis.

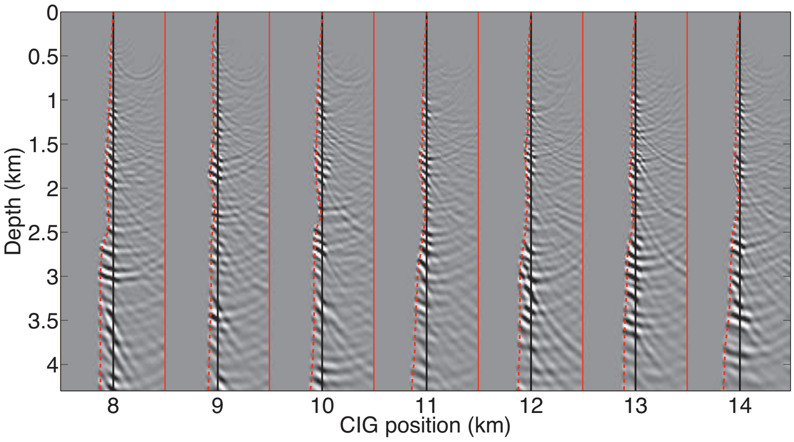
To demonstrate the efficiency of the demultiple method in reducing the sensitivity of WEMVA to free-surface multiples, we compare the results of velocity analysis using the original field data with those obtained using the demultipled data. The initial model used in both cases is the one shown in Figure 2.12a. For better comparison of the results, all updated images are constructed using the original data.

The updated velocities after 19 iterations of optimization using the original data are shown in Figure 2.15a, while the updated image is shown in Figure 2.15b. The results show that as expected the velocities updated by WEMVA represent a compromise, as the method attempts to simultaneously focus both primaries and multiples. As a consequence the final migrated image is locally distorted, and some reflectors that were well focused in the initial image (Figure 2.12b) are now clearly mispositioned.

In comparison, the velocities obtained after 26 iteration of velocity analysis using the demultipled data are shown in Figure 2.16a. The updated image (Figure 2.16b) now shows a slight improvement over the initial image (Figure 2.12b), as we would expect in this case.

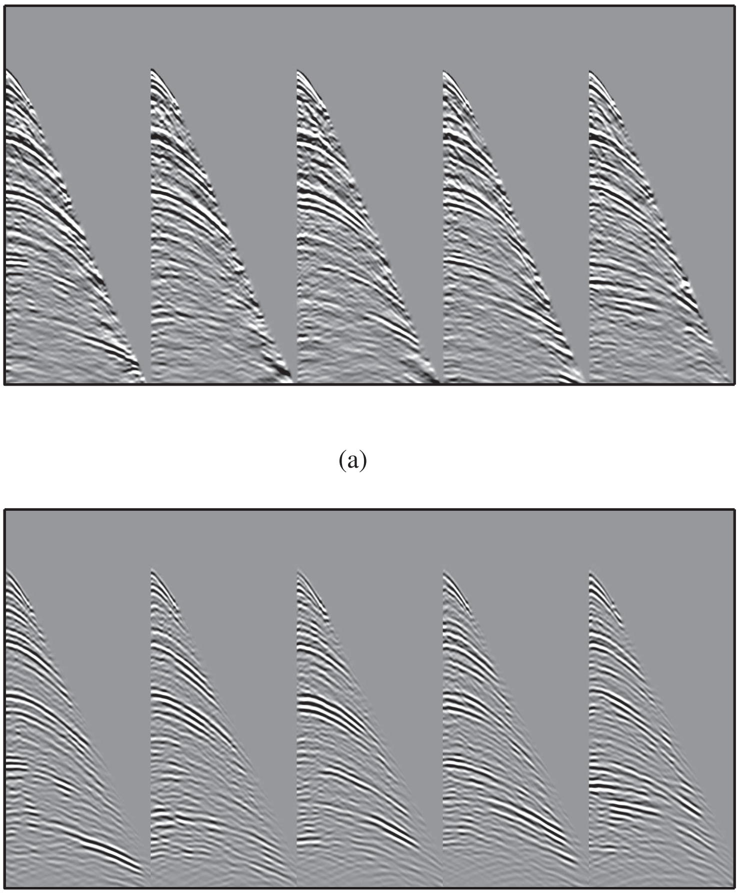


(a)



(b)

Figure 2.13: (a) Subsurface offset CIGs constructed using the initial velocity model in Figure 2.12a. The CIGs are separated by red vertical lines, while the dotted red lines mark the mute picks, and the black vertical lines mark the zero subsurface offset. (b) Subsurface offset CIGs after mute.



(a)

(b)

Figure 2.14: (a) 5 shots taken from the original data. (b) 5 shots taken from the data obtained from demigration of muted subsurface offset CIGs.

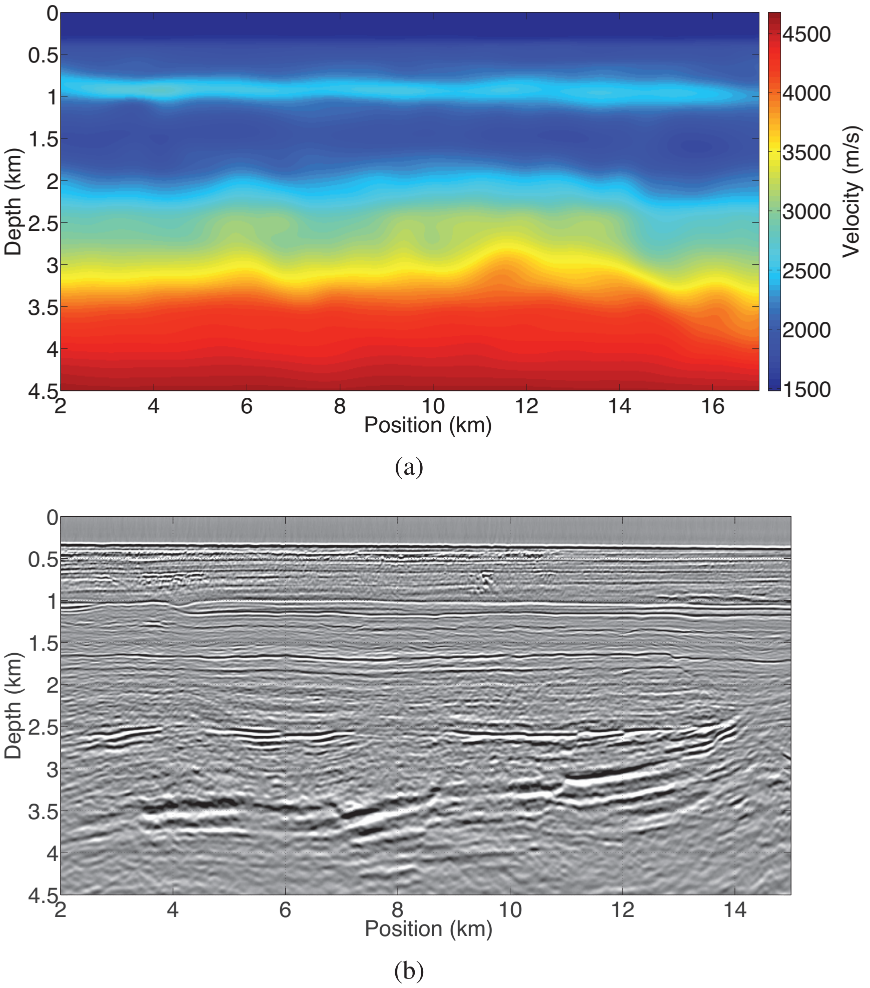
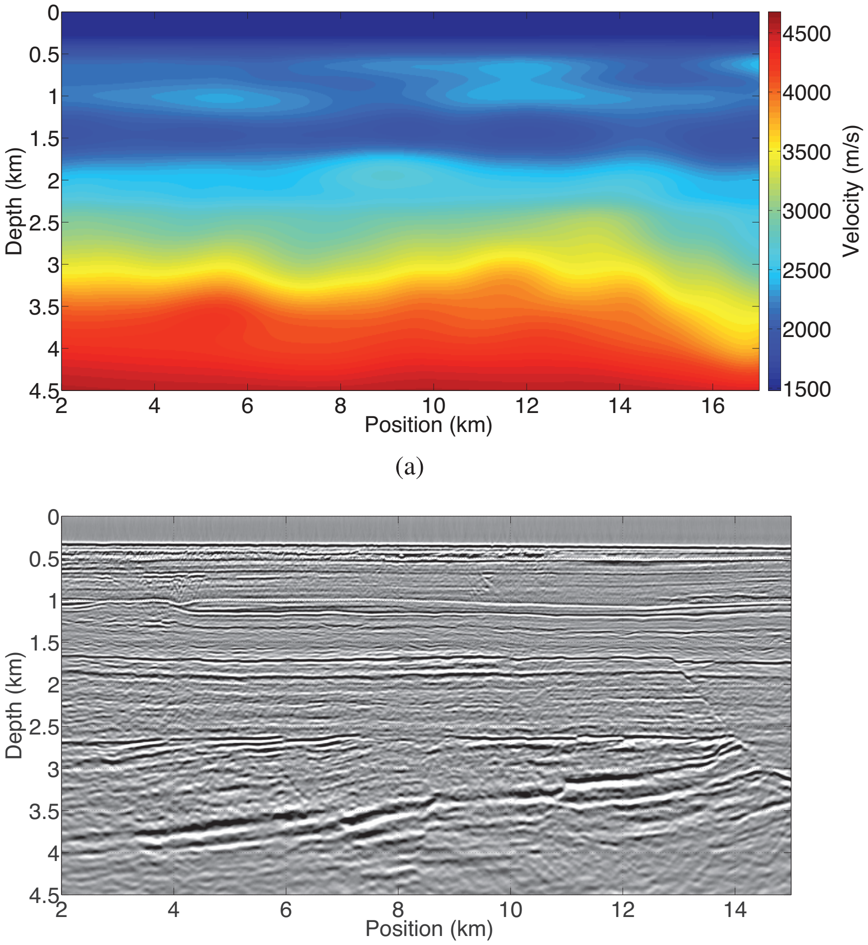


Figure 2.15: (a) Updated velocities from original data (m/s). (b) Image constructed using updated velocities in (a).



(a)

(b)

Figure 2.16: (a) Updated velocities from image-based demultipled data (m/s). (b) Image constructed using updated velocities in (a).

2.2 Discussion

RTM based WEMVA provides an automatic way of improving the quality of depth migrated images. However, artifacts in reverse-time migration which occur in presence of strong and sharp velocity contrasts can cause the method to diverge. In the first synthetic example we show how modifying the image with a vertical derivative operator improves the stability of the velocity analysis. In this example, the modification was necessary to ensure convergence and an adequate result. In general, we predict that this modification of the objective function is most significant when the velocity field has large velocity contrasts and/or is locally characterized by strong refracting mediums. In other cases, where one can accurately describe the kinematics of the acoustic medium with a more transparent velocity model (i.e. with only small acoustic impedance contrasts), such as in the second synthetic example, this modification may not significantly change the results.

The wellposedness of the velocity analysis is strongly dependent on the velocity preconditioning, and the regularization. To guarantee a stable solution, the number of parameters to be estimated needs to be adjusted to the quantity of information present in the image. In our numerical examples, we controlled the number of optimized parameters by choosing different grid spacings for the B-spline coefficients. In the first example, where we have well illuminated evenly distributed reflectors, the B-spline grid spacing could be made quite small (40×40 m). Whereas in the second example, with a scarcer number of reflectors to constrain the upper 2 km of the velocity field, the grid was purposely chosen more sparse (150×50 m). In the field data example, where there are large differences in the continuity and illumination of the reflectors a very sparse B-spline grid was adopted (600×100m). At the same time, the derivative regularization was used to prevent large artificial spatial velocity variations. These can appear during optimization due to the fact that the gradient of the differential semblance misfit and the similarity-index are scaled by the reflection coefficients of the image. However, it tends to slow down optimization. Therefore the regularization parameter should be chosen as small as possible. The choice of 1% of the initial differential semblance misfit value seemed to be satisfactory for all the examples.

Free-surface multiples have long been a problem for automatic velocity analysis methods (Mulder and ten Kroode, 2002; Li and Symes, 2007; van Leeuwen and Mulder, 2008; Mulder and van Leeuwen, 2008). Even if multiple attenuation is used as a part of preprocessing, some multiple energy might still remain and bias the velocity analysis (Li and Symes, 2007). Different methods have been proposed to reduce the sensitivity of WEMVA to the multiples. These methods are either based on including a filter in the objective function that bias the optimization towards higher velocities (Mulder and ten Kroode, 2002; Mulder and van Leeuwen, 2008), or iteratively modeling and subtracting the multiples as part of the velocity analysis (van Leeuwen and Mulder, 2008). We developed and tested a new preprocessing method where we mute the free-surface multiples in the subsurface offset CIGs. Instead of modifying the objective function, we use the muted CIGs to construct demultipled data, that is then used in the velocity analysis.

The distinction between multiples and primaries is done based on their focusing velocities, and might require interpretation. On the other hand, the procedure is very flexible and can also be used to remove other events from the data, such as steep dips and 'fast' interbed multiples, which are known to cause problems to WEMVA (Biondi and Shan, 2002).

In all examples, the initial models used as starting point for optimization, although 1D, were approximating well the true background trends of the actual velocities. These models could be found, in a first instance, by preconditioning the velocity model to the space of 1D velocity models or to the space of linear 1D models. This approach works here because, in the examples shown, the geology consisted approximately of plane layered overburdens. In more complicated geological environments, with strong and sharp contrast velocity variations, more refined initial models are likely to be required in order to avoid converging to non-optimal minima.

The high computational cost, both in terms of computation and storage, is currently limiting the method to 2D and low-frequency data sets. There are, however, several measures that can be used to reduce these costs. First in terms of computational cost, we suggest some strategies to improve the runtime of the method on large data sets.

The cost of RTM can be reduced significantly if shots can be combined and migrated together. If only two shots are combined the cost of migration is already halved. The speedup does not come for free though, since this approach introduces crosstalk in the result image. The crosstalk can be attenuated by using some sort of source encoding (Romero et al., 2000) . This approach has been widely experimented in RTM and FWI (Ben-Hadj-Ali et al., 2011). It is, however, not clear how the crosstalk artifacts could affect the results of WEMVA. Alternatively, the number of shots can be reduced through subsampling (Diaz and Guitton, 2011). In this case, only a subsample of the shots is taken and used for updating the velocity model at each iteration. By updating the subsampling along the iterative procedure, all shots are eventually used. A drawback of this procedure for WEMVA is that it may introduce aliasing in the subsurface offset CIGs, which may deteriorate the results.

In RTM a typical problem has been the need to model separately the incident wavefield and the scattered wavefield, which means that one of the two wavefields must be stored and reacessed later for the imaging step (checkpointing). In the case of differential semblance optimization, both fields need to be stored and accessed during gradient computation. Therefore the cost of storage is double of that of RTM. In this case, the cost of storing these fields can be reduced by optimal checkpointing, as suggested in Symes (2007).

2.3 Conclusion

We implemented a WEMVA method that can be used to estimate migration velocity fields from prestack seismic reflection data. The method estimates the velocities by minimizing an objective function based on differential semblance and similarity-index of subsurface offset CIGs constructed by reverse-time migration. Artifacts of reverse-time migration (low-frequency noise) can be very sensitive to changes in the velocities and cause WEMVA to diverge. We showed that by modifying the image with a simple spatial differentiation operator helps to stabilize the velocity analysis in the presence of strong velocity contrasts. We showed that multiples can be attenuated in a procedure based on reverse-time migration, muting and demigration, with optimal application for WEMVA.

2.4 Acknowledgments

We are grateful to Tamas Nemeth, Mauricio Sacchi, Jörg Schleicher, and five anonymous reviewers who greatly helped to improve the quality of this manuscript. The authors thank Statoil Petroleum AS, and the sponsors of the ROSE consortium for financial support of this work. We also acknowledge the partners in the Snorre license, Statoil Petroleum AS, Petoro AS, ExxonMobil Exploration and Production Norway AS, Idemitsu Petroleum Norge AS, RWE Dea Norge AS, and Core Energy AS for permission to publish the results.

2.5 Appendix A: Gradient computation

We now present the main steps required to derive the gradient of equations 2.5 and 2.6 by the adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Chavent, 2009).

The objective function function is given by

$$\mathcal{J}(\mathbf{U},\mathbf{D},\mathbf{v}) = \frac{1}{2}\int \mathbf{dh} \int \mathbf{dx}\, \hat{\mathbf{h}} \left[\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right]^2 , \tag{2.16}$$

where $\mathcal{R}(\mathbf{x},\mathbf{h}) = \sum_s \int dt\, U_s(\mathbf{x}+\mathbf{h},t) D_s(\mathbf{x}-\mathbf{h},t)$, and $\hat{\mathbf{h}} = \mathbf{h}^2 - \gamma\delta(\mathbf{h})$.

A Lagrangian function associated with the problem of minimizing equation 2.16 with respect to $\mathbf{v}$ can be written as

$$\begin{aligned}\mathcal{L}(\mathbf{U},\mathbf{D},\mathbf{U}',\mathbf{D}',\mathbf{v}) = {} & \mathcal{J}(\mathbf{U},\mathbf{D},\mathbf{v}) \\ & + \sum_s \left\langle \mathbf{U}'_\mathbf{s}, H^T \mathbf{U}_\mathbf{s} - \mathbf{P}_\mathbf{s} \right\rangle_{\mathbf{x},t} \\ & + \sum_s \left\langle \mathbf{D}'_\mathbf{s}, H \mathbf{D}_\mathbf{s} - \mathbf{S}_\mathbf{s} \right\rangle_{\mathbf{x},t} ,\end{aligned} \tag{2.17}$$

where $U'_s(\mathbf{x},t)$ and $D'_s(\mathbf{x},t)$ are Lagrange multipliers (adjoint states). The operator $H(\mathbf{v}) = \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)$ is the acoustic wave equation forward time marching operator. The transpose, $H^T(\mathbf{v})$, leads to a backward time marching scheme.

Here $\mathcal{L}$ is related to $\mathcal{J}$ by

$$\mathcal{J} = \mathcal{L}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}), \tag{2.18}$$

where $\mathbf{U}_\mathbf{v}$, $\mathbf{D}_\mathbf{v}$ denote one realization of the direct states for a particular shot with a velocity vector $\mathbf{v}$.

Implicit differentiation of the above equation with respect to $\mathbf{v}$ gives

$$\begin{aligned} \delta\mathcal{J} &= \frac{\partial\mathcal{L}}{\partial\mathbf{U}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{U} + \frac{\partial\mathcal{L}}{\partial\mathbf{D}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{D} \\ &\quad + \frac{\partial\mathcal{L}}{\partial\mathbf{v}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{v}. \end{aligned} \tag{2.19}$$

Now, if we assume $\mathbf{U}'$ and $\mathbf{D}'$ to satisfy

$$\frac{\partial\mathcal{L}}{\partial\mathbf{U}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{U} = 0, \tag{2.20}$$

and

$$\frac{\partial\mathcal{L}}{\partial\mathbf{D}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{D} = 0, \tag{2.21}$$

for all $\delta\mathbf{U}$ and $\delta\mathbf{D}$, then equation 2.19 reduces to

$$\delta\mathcal{J} = \frac{\partial\mathcal{L}}{\partial\mathbf{v}}(\mathbf{U}_\mathbf{v}, \mathbf{D}_\mathbf{v}, \mathbf{U}', \mathbf{D}', \mathbf{v}) \cdot \delta\mathbf{v}. \tag{2.22}$$

The problem now is to solve equations 2.20 and 2.21 for the adjoint states for each shot. Starting with equation 2.21. Noting that $\frac{\partial}{\partial x_3}^T = -\frac{\partial}{\partial x_3}$, and that

$$\int \mathbf{dh}\, \mathcal{R}(\mathbf{x}) U(\mathbf{x}+\mathbf{h}) \delta D(\mathbf{x}-\mathbf{h}) = \int \mathbf{dh}\, \mathcal{R}(\mathbf{x}+\mathbf{h}) U(\mathbf{x}+2\mathbf{h}) \delta D(\mathbf{x}), \tag{2.23}$$

the stationay points of the Lagragian $\frac{\partial\mathcal{L}}{\partial\mathbf{D}} = 0$ lead to the reverse-time problem

$$H^T(\mathbf{v}) D_s'(\mathbf{x}, t) = \int \mathbf{dh}\, \hat{\mathbf{h}} \frac{\partial^2\mathcal{R}}{\partial x_3^2}(\mathbf{x}+\mathbf{h}, \mathbf{h}) U_s(\mathbf{x}+2\mathbf{h}, t), \tag{2.24}$$

which can be be solved through a backward time marching scheme, starting from a final condition of rest, i.e. $D_s'(\mathbf{x}, T) = 0$.

While equation $\frac{\partial\mathcal{L}}{\partial\mathbf{U}} = 0$ leads to the forward-time problem

$$H(\mathbf{v}) U_s'(\mathbf{x}, t) = \int \mathbf{dh}\, \hat{\mathbf{h}} \frac{\partial^2\mathcal{R}}{\partial x_3^2}(\mathbf{x}-\mathbf{h}, \mathbf{h}) D_s(\mathbf{x}-2\mathbf{h}, t), \tag{2.25}$$

which can be solved through a forward time marching scheme, starting from an initial condition of rest, i.e. $U_s'(\mathbf{x}, 0) = 0$.

The solutions to equations 2.24 and 2.25 can be expressed in the form of Green's functions as

$$D'_s(\mathbf{x},t) = \int \mathbf{dx}'\, G(\mathbf{x},0;\mathbf{x}',t) * \int \mathbf{dh}\, \hat{\mathbf{h}} \frac{\partial^2 \mathcal{R}}{\partial {x'}_3^2}(\mathbf{x}'+\mathbf{h},\mathbf{h}) U_s(\mathbf{x}'+2\mathbf{h},t), \tag{2.26}$$

and

$$U'_s(\mathbf{x},t) = \int \mathbf{dx}'\, G(\mathbf{x},t;\mathbf{x}',0) * \int \mathbf{dh}\, \hat{\mathbf{h}} \frac{\partial^2 \mathcal{R}}{\partial {x'}_3^2}(\mathbf{x}'-\mathbf{h},\mathbf{h}) D_s(\mathbf{x}'-2\mathbf{h},t), \tag{2.27}$$

where $*$ means time convolution.

Finally we turn to the problem of finding the derivative of $\mathcal{J}$ with respect to velocity $(\mathbf{v})$. Differentiating equation 2.17 with respect to $\mathbf{v}$, and noting that $\left\langle \mathbf{D}', \frac{\partial H}{\partial \mathbf{v}} \mathbf{D} \right\rangle_t = \left\langle \mathbf{D}', \frac{\partial H}{\partial \mathbf{v}}^T \mathbf{D} \right\rangle_t = \left\langle \mathbf{D}', -\frac{2}{\mathbf{v}^3} \frac{\partial^2 \mathbf{D}}{\partial t^2} \right\rangle_t$, yields

$$\begin{aligned} \delta \mathcal{J} &= -\sum_s \left\langle \mathbf{D}'_\mathbf{s}, \frac{2\delta \mathbf{v}}{\mathbf{v}^3} \frac{\partial^2 \mathbf{D}_\mathbf{s}}{\partial t^2} \right\rangle_{x,t} \\ &\quad -\sum_s \left\langle \mathbf{U}'_\mathbf{s}, \frac{2\delta \mathbf{v}}{\mathbf{v}^3} \frac{\partial^2 \mathbf{U}_\mathbf{s}}{\partial t^2} \right\rangle_{x,t}, \end{aligned} \tag{2.28}$$

from which we obtain the gradient by picking the coefficients of $\delta \mathbf{v}$:

$$\begin{aligned} \nabla_\mathbf{v} \mathcal{J}(\mathbf{x}) &= -\sum_s \int dt\, \frac{2}{v^3(\mathbf{x})} \frac{\partial^2 D_s}{\partial t^2}(\mathbf{x},t) D'_s(\mathbf{x},t) \\ &\quad -\sum_s \int dt\, \frac{2}{v^3(\mathbf{x})} \frac{\partial^2 U_s}{\partial t^2}(\mathbf{x},t) U'_s(\mathbf{x},t). \end{aligned} \tag{2.29}$$

Chapter 3

Anisotropic migration velocity analysis using reverse-time migration

Wiktor Weibull and Børge Arntsen

Norwegian University of Science and Technology, Trondheim, Norway

Abstract

Seismic anisotropy, if not accounted for, can cause significant mispositioning of the reflectors in depth migrated images. Accounting for anisotropy in depth migration requires velocity analysis tools that can estimate the anisotropic background velocity field. We extend wave equation migration velocity analysis to deal with 2D tilted transverse isotropic media. The velocities are obtained automatically by non-linear optimization of the focusing and stack-power of common image point gathers constructed using an extended imaging condition. We use the elastic two-way wave equation to reconstruct the wavefields needed for both the image, and gradient computations. This leads to an anisotropic migration velocity analysis algorithm based on reverse-time migration. We illustrate the method with synthetic and field data examples based on marine surface seismic acquisition. The results show that the method significantly improves the quality of the depth migrated image. However, as it is common in the case of velocity analysis using surface seismic data, the estimation of anisotropic parameters seems to be strongly non-unique.

Presented at the ROSE Meeting in Trondheim 2012, at the 74th EAGE Conference and Exhibition incorporating SPE EUROPEC in June 2012, at the SEG/EAGE Summer Research Workshop in July 2012, and at the 81st SEG Annual International Meeting in September 2012; Paper submitted to Geophysics 2013.

3.1 Introduction

WEMVA can be described as a non-linear least squares inversion of prestack seismic reflection data in the image domain (Sava and Biondi, 2004). The procedure consists in setting up an objective function that can measure the misfit in the image due to a prestack depth migration with a non-optimal velocity model, and then minimize this function with respect to the velocity parameters. The objective function for WEMVA is typically based on the focusing of common image point gathers (Symes and Kern, 1994; Shen et al., 2003), stacking-power (Toldi, 1989; Chavent and Jacewitz, 1995), or a combination of both (Mulder, 2008; Shen and Symes, 2008). The procedure can be made fully automatic, and is, to a certain extent, robust against poor initial guesses of the velocity field (Shen and Symes, 2008). But due to the assumptions in the model of prestack depth migration, WEMVA is restricted to kinematic inversion of single scattering reflection data (Mulder and van Leeuwen, 2008).

Most implementations of WEMVA are based on the acoustic isotropic approximation (Sava and Vlad, 2008). However, in cases where the velocity field is anisotropic, velocity analysis under an isotropic assumption will ultimately lead to images that are well-focused, but mispositioned in space (Isaac and Lawton, 1999). In an attempt to overcome this problem, some research have been devoted to implement WEMVA under anisotropic assumptions. Li and Biondi (2011) proposed a method based on the depth-oriented extension of the differential semblance objective function (Shen et al., 2003) and one-way wave equation migration in a vertically transverse isotropic (VTI) medium. Li et al. (2012) also presented a method based on a similar objective function, but using the pseudo-acoustic two-way wave equation for a VTI medium (Alkhalifah, 1998). Weibull et al. (2012) presented a method where they use an objective function based on the depth-oriented extension of differential semblance combined with stacking-power maximization (Toldi, 1989), and elastic reverse-time migration to estimate anisotropic parameters over a VTI medium.

A VTI model can be a good approximation to some horizontally or nearly horizontally layered sequences (Levin, 1979; Banik, 1984; Sayers, 1994). But for tectonically deformed geological settings, such as fold thrust belts or at the flanks of salt diapirs, a tilted transverse isotropic (TTI) model is a better approximation (Isaac and Lawton, 1999). A general 2D TTI medium can be described by five spatially varying parameters, the P-wave velocity along the symmetry axis V_{P0}, the S-wave velocity along the symmetry axis V_{S0}, Thomsen (1986) parameters ε and δ, and the tilt θ of the symmetry axis with respect to the vertical. However, some assumptions can be used to reduce the number of parameters needed to describe the kinematics of the medium. In the context of P-wave velocity analysis, Tsvankin and Thomsen (1994) and Alkhalifah and Larner (1994) demonstrated that V_{S0} can be arbitrarily chosen. Also, a popular assumption, often referred to as structural transverse isotropy (STI), further reduces the number of parameters by assuming that θ is always perpendicular to the structure of the reflectors (Audebert et al., 2006).

In this paper, we extend WEMVA to deal with a 2D TTI model of the subsurface, and test it on synthetic and field surface seismic data. To account for anisotropy in the kinematics of wave propagation, we use a density normalized elastic wave equation. This is stable and accurately propagates waves at all angles, which is important for the estimation of anisotropic parameters. We use WEMVA to simultaneously estimate V_{P0}, ε, and δ. The parameter V_{S0} is assumed to have negligible influence on P-wave propagation and chosen arbitrarily, and θ is assumed to conform to the geology, and is estimated from the structure of the reflectors in the migrated image.

A major difficulty in the estimation of anisotropic velocities from the kinematics of surface reflection data is the inherent non-uniqueness related to the positioning of the reflectors in the subsurface and/or to the tradeoff between the different parameters (Vestrum et al., 1999; Grechka et al., 2002). The tradeoff between heterogeneity and anisotropy can, in principle, be reduced by considering the images in vertical time instead of depth, as suggested by Alkhalifah et al. (2001), or by considering a stretched depth axis as in Plessix and Rynja (2010). In this work, we use regularization to constrain the models to a physical set, and a coarse bi-cubic B-spline grid to confine the models to a sparse solution space. These measures help to obtain a convergent WEMVA algorithm, but are not sufficient to obtain a unique geological model of the subsurface. In practice, substantial additional information in the form of well logs and check shot surveys are required to narrow down the range of possible solutions to the problem (Yan et al., 2004; Bakulin et al., 2010).

This paper starts by explaining the method and showing what can be expected in the ideal condition. The method is then tested on 2D synthetic and field data sets. Next we discuss the main results and suggest potential ways forward. Finally, we present our conclusions.

3.2 2D TTI reverse-time migration

The basis for WEMVA is prestack depth migration. To build common image point gathers (CIGs) for velocity analysis, we use reverse-time migration with an extended imaging condition (Rickett and Sava, 2002)

$$\mathcal{R}(\mathbf{x},\mathbf{h}) = \int ds \int dt \, W^s(\mathbf{x}-\mathbf{h},t,s) W^r(\mathbf{x}+\mathbf{h},T-t,s), \tag{3.1}$$

where W^s are the forward modeled source wavefields, W^r are the reverse-time modeled receiver wavefields, $\mathbf{x} = (x,z)$ are the spatial coordinates, with z being the depth axis, $\mathbf{h} = (h_x, 0)$ is the subsurface horizontal half-offset, t is the time, and s is the source index.

The computation of the W^s and W^r wavefields depend on the choice of the wave equation. To take anisotropy into account, an anisotropic wave equation must be used in the reconstruction of the wavefields. We model the wave propagation in a 2D TTI medium, using a density normalized elastic wave equation (Ikelle and Amundsen, 2005)

$$\frac{\partial^2 u_i}{\partial t^2}(\mathbf{x},t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l}{\partial x_k}(\mathbf{x},t)\right] = F_i(\mathbf{x},t), \tag{3.2}$$

where u_i is the displacement field, a_{ijkl} is the density normalized elasticity tensor, F_i is a source term, and $i, j, k, l = x, z$ are indexes under the Einstein summation convention. The elastic wave equation and the density normalized elastic parameters a_{ijkl} are described in more detail in Appendix A.

To obtain the W^s and W^r wavefields using equation 3.2, we first model the displacement vector wavefields, u_i^s and u_i^r, according to the following equations

$$\frac{\partial^2 u_i^s}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l^s}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial S}{\partial x_i}(\mathbf{x_s}, t, s), \tag{3.3}$$

and,

$$\frac{\partial^2 u_i^r}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x})\frac{\partial u_l^r}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial P}{\partial x_i}(\mathbf{x_r}, T - t, s), \tag{3.4}$$

for respectively the source and receiver displacements. In equation 3.3, S is the source time function for source s, at location $\mathbf{x_s}$. Whereas in equation 3.4, P is the time-reversed recorded seismic reflection data for source s at receiver positions given by $\mathbf{x_r}$. Also note that equation 3.3 is to be solved forward in time, while equation 3.4 is to be solved in reverse-time.

We then extract a scalar wavefield from the source and receiver displacement wavefields by taking the divergence of the displacement vector scaled by the density normalized bulk modulus

$$W^s(\mathbf{x}, t, s) = V_{P0}^2(\mathbf{x})\frac{\partial u_i^s}{\partial x_i}(\mathbf{x}, t, s), \tag{3.5}$$

and,

$$W^r(\mathbf{x}, t, s) = V_{P0}^2(\mathbf{x})\frac{\partial u_i^r}{\partial x_i}(\mathbf{x}, t, s). \tag{3.6}$$

Note that, different from the pseudo-acoustic approximation of Alkhalifah (1998), equation 3.2 requires V_{S0} to be provided. However, if only the kinematics of P-wave propagation are considered, the S-wave velocities are of minor importance (Alkhalifah and Tsvankin, 1995). In this work, the S-wave velocities are heuristically chosen to be 0.9 km/s, and spatially invariant.

3.3 WEMVA

We quantify a misfit in the prestack depth migrated image using the same objective function as described by Shen and Symes (2008). The objective function consists of a combination of the depth oriented extension of differential semblance optimization (Shen et al., 2003) with stack-power maximization (Toldi, 1989; Chavent and Jacewitz, 1995; Zhou et al., 2009). The objective function can be written as

$$\mathcal{J} = \frac{1}{2}\left\|\mathbf{h}\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x}, \mathbf{h})\right\|^2 - \frac{\gamma}{2}\left\|\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x}, 0)\right\|^2 = \frac{1}{2}\int \mathbf{dx} \int \mathbf{dh}\, \hat{\mathbf{h}} \left[\frac{\partial \mathcal{R}}{\partial z}(\mathbf{x}, \mathbf{h})\right]^2, \tag{3.7}$$

where $\hat{\mathbf{h}} = \mathbf{h}^2 - \gamma\delta(\mathbf{h})$, with δ being the Kronecker delta and γ being a constant weight that balances the contribution of differential semblance and stack-power to the total value of the objective function. The spatial derivative operator ∂_z attenuates the low vertical wavenumber components in the RTM image (Guitton et al., 2007). This helps to improve the stability and convergence properties of WEMVA using the two-way wave equation (Mulder, 2008; Weibull and Arntsen, 2011). We use the vertical derivative operator for its simplicity and robustness. There are more sophisticated noise-reducing imaging conditions available, which also attenuate horizontally oriented noise, as well as better preserves vertical reflectors (Douma et al., 2010; Whitmore and Crawley, 2012).

In addition to quantifying the misfit in the CIGs, we use regularization to constrain the parameters to a feasible set, and also to prevent excessive roughness in the solution (Tikhonov and Arsenin, 1977). The regularization is implemented by adding the following term to the objective function

$$\begin{aligned}
\mathcal{J}_R \;=\; & \frac{\alpha_1}{2}\left\|\frac{\partial V_{P0}}{\partial x_i}(\mathbf{x}) - \frac{\partial V_{P0}^0}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_1(\mathbf{x})}{2}\left\|V_{P0}(\mathbf{x}) - V_{P0}^0(\mathbf{x})\right\|^2 \\
& + \frac{\alpha_2}{2}\left\|\frac{\partial \varepsilon}{\partial x_i}(\mathbf{x}) - \frac{\partial \varepsilon^0}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_2(\mathbf{x})}{2}\left\|\varepsilon(\mathbf{x}) - \varepsilon^0(\mathbf{x})\right\|^2 \\
& + \frac{\alpha_3}{2}\left\|\frac{\partial \delta}{\partial x_i}(\mathbf{x}) - \frac{\partial \delta^0}{\partial x_i}(\mathbf{x})\right\|^2 + \frac{\beta_3(\mathbf{x})}{2}\left\|\delta(\mathbf{x}) - \delta^0(\mathbf{x})\right\|^2, \qquad (3.8)
\end{aligned}$$

where V_{P0}^0, ε^0, δ^0 represent initial values of the target parameters; and α_n and β_n, with $n = 1, 2, 3$ are constant weights, one for each spatially varying parameter.

One limitation of the objective function given by equation 3.7 is that it is strictly valid under the single scattering assumption. Because multiples and primaries will focus at different velocities in the CIGs, the presence of multiples will introduce local minima in the objective function. One of the simplest solutions to this problem is to include multiple attenuation as a part of the preprocessing of the data used for velocity analysis, and use absorbing boundary conditions at all sides for the source and receiver wavefield reconstructions. In case multiple attenuation fails, there are other approaches that might be useful. Mulder and van Leeuwen (2008) propose to reduce the influence of the low velocity free-surface multiples by modifying the objective function with an asymmetric weighting function. Another proposed method consists in modeling and subtracting the multiples as a part of the velocity analysis (van Leeuwen and Mulder, 2008).

To check the sensitivity of the objective function to the parameters V_{P0}, ε and δ, we evaluated it in a simple 1D 3-layer model, as shown in Figure 3.1a. In a first stage, we used this acoustic model to simulate surface seismic data using a split-spread geometry with maximum offset of 1.4 km. We used a monopole point source, consisting of a Ricker pulse with dominant frequency of 15 Hz. No free surface was used in the modeling. The preprocessing of the data was limited to muting the direct wave and wide angle reflections. The resulting data are shown in Figures 3.1b-c. The data is virtually single scattering and therefore ideal within the assumptions of the method. Next, we perturbed the magnitude of the true parameters in the second layer for different values of V_{P0}, ε and δ, and migrated the data. In this procedure, only one parameter is perturbed at each time, leaving the other parameters set at their true value. The tests were conducted two times, one time with θ fixed to zero degrees, and a second time with θ fixed at 45 degrees. We then used the resulting images to compute the objective function values for three different combinations of the misfit functions: For stack-power alone, for differential semblance alone, and for the combination of stack power and differential semblance.

The results in Figure 3.2 show a comparison of the variation of the values of the different objective functions for each parameter. Figures 3.2a-c, show the results with θ fixed at zero degrees, while Figures 3.2d-f show the results with θ fixed at 45 degrees. What this idealized experiment shows is that the misfit functions are quasi-convex for a wide range of model perturbations, and may therefore be amenable for gradient based optimization.

Another fact shown is that, in this ideal case, if all but one parameter are known precisely, the unknown parameter can be uniquely determined. Within the Thomsens's parameters, the objective functions are more sensitive to errors in ε, than they are to errors in δ. The effect of having the wrong θ model is largest for δ, than it is for ε. And because the test model is isotropic, the sensitivity to the choice of θ is zero for the V_{P0} tests.

3.3.1 Gradient computation

To minimize the objective function, we use a L-BFGS method (Byrd et al., 1995; Nocedal and Wright, 2000). This method requires the evaluation of the objective function and its gradient with respect to the parameters at each iteration or line search step. We compute the gradient of equation 3.7 using the adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Plessix, 2006). This method gives the following equations for the gradients with respect to $V_{P0}(\mathbf{x})$, $\varepsilon(\mathbf{x})$, and $\delta(\mathbf{x})$

$$\begin{aligned}
\frac{\partial \mathcal{J}}{\partial V_{P0}}(\mathbf{x}) &= \int ds \int dt\, \frac{\partial a_{ijkl}}{\partial V_{P0}}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x},t,s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x},T-t,s) \\
&+ \int ds \int dt\, \frac{\partial a_{ijkl}}{\partial V_{P0}}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x},T-t,s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x},t,s) \\
+ \int ds \int dt\, 2V_{P0}(\mathbf{x}) & \frac{\partial u_i^s}{\partial x_i}(\mathbf{x},t,s) \int dh\, \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h},\mathbf{h}) W^r(\mathbf{x}+2\mathbf{h},T-t,s) \\
+ \int ds \int dt\, 2V_{P0}(\mathbf{x}) & \frac{\partial u_i^r}{\partial x_i}(\mathbf{x},T-t,s) \int dh\, \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}-\mathbf{h},\mathbf{h}) W^s(\mathbf{x}-2\mathbf{h},t,s), \qquad (3.9)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{J}}{\partial \varepsilon}(\mathbf{x}) &= \int ds \int dt\, \frac{\partial a_{ijkl}}{\partial \varepsilon}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x},t,s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x},T-t,s) \\
&+ \int ds \int dt\, \frac{\partial a_{ijkl}}{\partial \varepsilon}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x},T-t,s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x},t,s) \qquad (3.10)
\end{aligned}$$

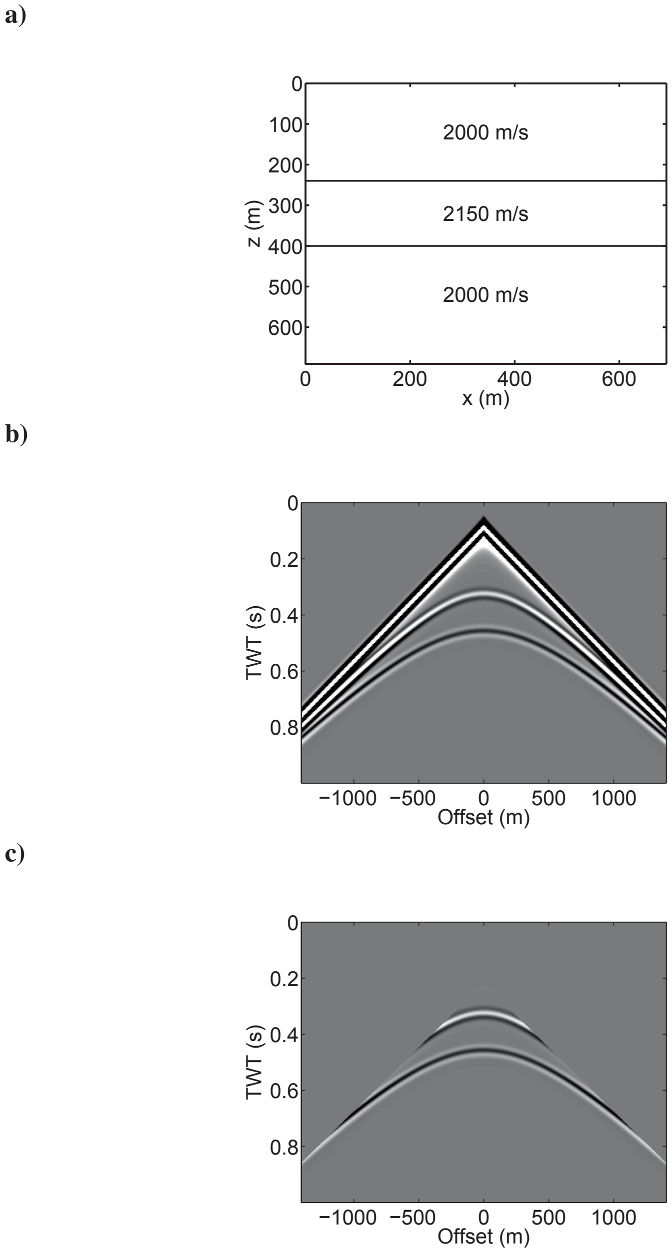


Figure 3.1: Model and data used to generate the sensitivity plots in Figure 3.2. (a) 1D Velocity model. (b) Synthetic shotgather modeled using the model in (a). (c) Shotgather in (b) after mute to remove direct wave and post-critical reflections.

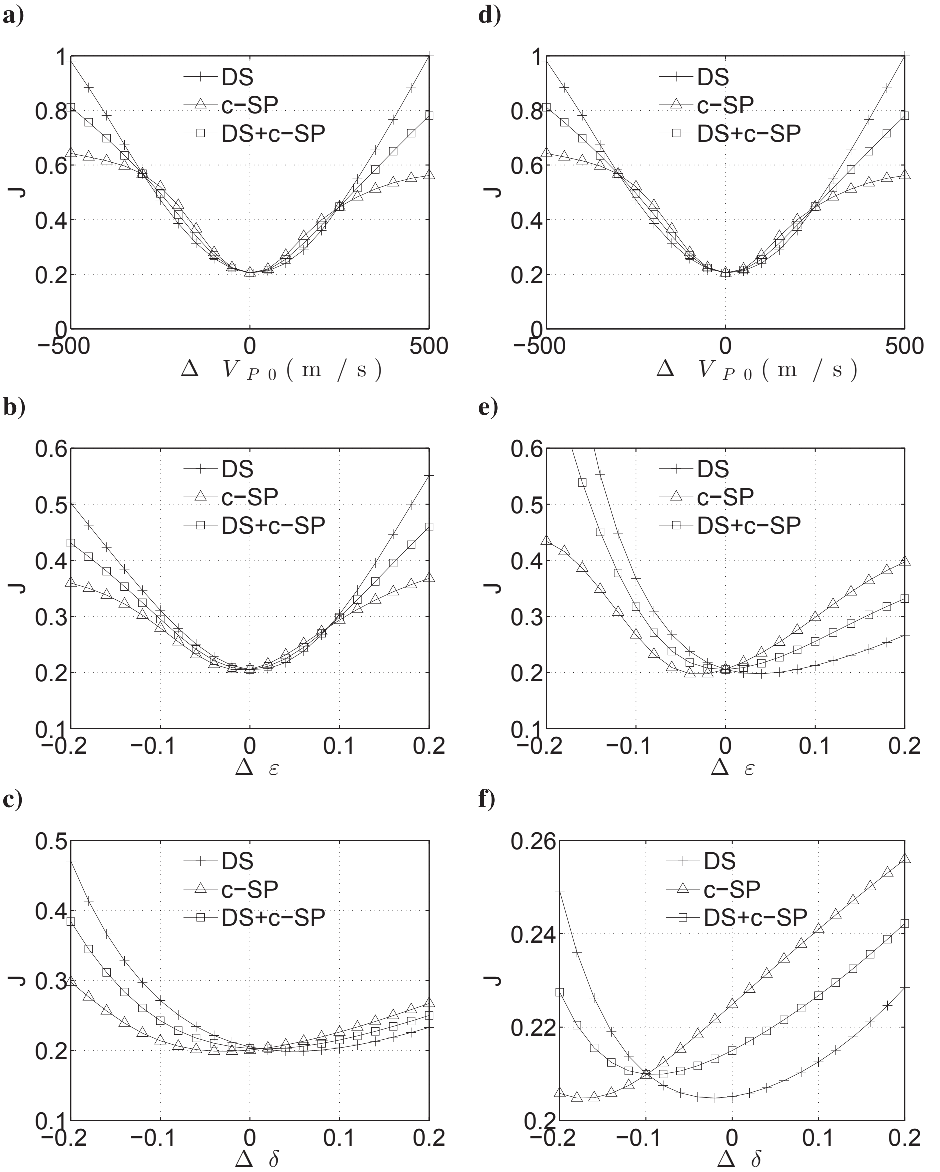


Figure 3.2: Objective function value computed from migrating the data in Figure 3.1c using the model in Figure 3.1a modified by perturbations in: (a) V_{P0}; (b) ε; (c) δ; (d), (e), and (f) are similar to (a), (b), and (c) but with θ set to 45 degrees. DS corresponds to the differential semblance misfit function, while SP corresponds to stack-power, with c being a constant added so that the minimum of $-SP$ is equal to the minimum of DS .

and,

$$\frac{\partial \mathcal{J}}{\partial \delta}(\mathbf{x}) = \int ds \int dt \, \frac{\partial a_{ijkl}}{\partial \delta}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x}, t, s) \frac{\partial \tilde{u}_i^s}{\partial x_j}(\mathbf{x}, T-t, s) + \int ds \int dt \, \frac{\partial a_{ijkl}}{\partial \delta}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x}, T-t, s) \frac{\partial \tilde{u}_i^r}{\partial x_j}(\mathbf{x}, t, s). \tag{3.11}$$

Note that because we scale the divergence of displacement vector by the density normalized bulk modulus V_{P0}^2, the gradient with respect to V_{P0} (equation 3.9) has two additional terms in the right hand side, when compared to the formula for the other gradients (equations 3.10 and 3.11).

The adjoint state wavefields $\tilde{u}_i^s$ and $\tilde{u}_i^r$ can be computed by the following adjoint modelings

$$\frac{\partial^2 \tilde{u}_i^s}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x}) \frac{\partial \tilde{u}_l^s}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial A^s}{\partial x_i}(\mathbf{x}, T-t, s) \tag{3.12}$$

and,

$$\frac{\partial^2 \tilde{u}_i^r}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[a_{ijkl}(\mathbf{x}) \frac{\partial \tilde{u}_l^r}{\partial x_k}(\mathbf{x}, t)\right] = \frac{\partial A^r}{\partial x_i}(\mathbf{x}, t, s), \tag{3.13}$$

where A^s and A^r are given by

$$A^s(\mathbf{x}, t, s) = V_{P0}^2(\mathbf{x}) \int dh \, \hat{\mathbf{h}} \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h}, \mathbf{h}) W^r(\mathbf{x}+2\mathbf{h}, T-t, s), \tag{3.14}$$

and,

$$A^r(\mathbf{x}, t, s) = V_{P0}^2(\mathbf{x}) \int dh \, \hat{\mathbf{h}} \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}-\mathbf{h}, \mathbf{h}) W^s(\mathbf{x}-2\mathbf{h}, t, s). \tag{3.15}$$

The adjoint sources $\frac{\partial A^s}{\partial x_i}$, and $\frac{\partial A^r}{\partial x_i}$ represent respectively the source and receiver side displacement residuals. These residuals are obtained by taking the kernel of the derivatives of the objective function with respect to the displacement wavefields u_i^s, and u_i^r. Because the displacement wavefields u_i^s, and u_i^r are originally shifted by respectively $+\mathbf{h}$ and $-\mathbf{h}$, a shift with opposite sign must be applied to the objective function prior to taking the respective derivatives. This explains why A^s have dependencies in $\mathbf{x}+\mathbf{h}$, and $\mathbf{x}+\mathbf{2h}$, and A^r have dependencies in $\mathbf{x}-\mathbf{h}$, and $\mathbf{x}-\mathbf{2h}$. For example, in the case of the source side we have

$$\int dx \int dh \, \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}, \mathbf{h}) W^s(\mathbf{x}-\mathbf{h}) W^r(\mathbf{x}+\mathbf{h}) = \int dx \int dh \, \frac{\partial^2 \mathcal{R}}{\partial z^2}(\mathbf{x}+\mathbf{h}, \mathbf{h}) W^s(\mathbf{x}) W^r(\mathbf{x}+\mathbf{2h}). \tag{3.16}$$

The same argument can be used for the receiver side. Finally, note that equation 3.12 is to be solved in reverse-time, while equation 3.13 is to be solved forward in time.

If regularization is applied, the gradients of equation 3.7 with respect to the velocity parameters must be augmented with the gradients of equation 3.8, which are given by

$$\frac{\partial \mathcal{J}_R}{\partial V_{P0}}(\mathbf{x}) = \beta_1(\mathbf{x}) \left(V_{P0}(\mathbf{x}) - V_{P0}^0(\mathbf{x})\right) - \alpha_1 \left(\frac{\partial^2 V_{P0}}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 V_{P0}^0}{\partial x_i^2}(\mathbf{x})\right), \tag{3.17}$$

$$\frac{\partial \mathcal{J}_R}{\partial \varepsilon}(\mathbf{x}) = \beta_2(\mathbf{x}) \left(\varepsilon(\mathbf{x}) - \varepsilon^0(\mathbf{x})\right) - \alpha_2 \left(\frac{\partial^2 \varepsilon}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 \varepsilon^0}{\partial x_i^2}(\mathbf{x}) \right), \tag{3.18}$$

$$\frac{\partial \mathcal{J}_R}{\partial \delta}(\mathbf{x}) = \beta_3(\mathbf{x}) \left(\delta(\mathbf{x}) - \delta^0(\mathbf{x})\right) - \alpha_3 \left(\frac{\partial^2 \delta}{\partial x_i^2}(\mathbf{x}) - \frac{\partial^2 \delta^0}{\partial x_i^2}(\mathbf{x}) \right). \tag{3.19}$$

3.3.2 Velocity preconditioning

To speed up convergence, and to restrict the space of possible solutions, we precondition the velocity parameters using bi-cubic B-splines (Dierckx, 1993)

$$V_{P0}(\mathbf{x}) = \sum_m b_{m1} B_{m1}(\mathbf{x}) + V_{P0}^0(\mathbf{x}), \tag{3.20}$$

$$\varepsilon(\mathbf{x}) = \sum_m b_{m2} B_{m2}(\mathbf{x}) + \varepsilon^0(\mathbf{x}), \tag{3.21}$$

$$\delta(\mathbf{x}) = \sum_m b_{m3} B_{m3}(\mathbf{x}) + \delta^0(\mathbf{x}), \tag{3.22}$$

where, B_{mn} are cubic splines defined at predetermined points $\mathbf{m} = (m_x, m_z)$ in a spline grid and b_{mn} are coefficients to be determined by the velocity analysis. The spacing of the spline grid controls the sparseness of the solution and can be chosen differently for V_{P0}, ε and δ, hence the $n = 1, 2, 3$ index. We only use B-splines to represent the velocity updates. This avoids having to fit the initial model to a B-spline basis, which would otherwise result in unnecessary smoothing.

The B-spline representation is attractive because it allows for local velocity variations, while, at the same time, ensures continuous second order spatial derivatives. These properties help to obtain a numerically stable and well-posed velocity analysis algorithm.

In practice, we compute the gradients in Cartesian coordinates and subsequently transform them to the spline basis

$$\frac{\partial \mathcal{J}_{m1}}{\partial b_{m1}} = \int d\mathbf{x}\; B_{m1}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial V_{P0}}(\mathbf{x}), \tag{3.23}$$

$$\frac{\partial \mathcal{J}_{m2}}{\partial b_{m2}} = \int d\mathbf{x}\; B_{m2}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial \varepsilon}(\mathbf{x}), \tag{3.24}$$

$$\frac{\partial \mathcal{J}_{m3}}{\partial b_{m3}} = \int d\mathbf{x}\; B_{m3}(\mathbf{x}) \frac{\partial \mathcal{J}}{\partial \delta}(\mathbf{x}). \tag{3.25}$$

3.3.3 Diagonal scaling

One problem associated with the estimation of more than one parameter at the same time is that the magnitudes of the gradients with respect to the different parameters can be very different. The different sensitivities are normally compensated for in a full-Newton optimization method, through the scaling given by the inverse Hessian matrix (Nocedal and Wright, 2000). However, for quasi-Newton methods this is not the case. The poor relative scaling causes the optimization to be dominated by the parameters with the largest gradient magnitudes. To mitigate this problem, we apply the so-called diagonal scaling (Nocedal and Wright, 2000), where the optimization variables are related to the bi-cubic spline coefficients by a linear coordinate transformation

$$b'_{mn} = \frac{b_{mn}}{k_n}. \tag{3.26}$$

By proper choice of the constants k_n for each spline coefficient array b_{mn}, we can rescale the gradients, since

$$\frac{\partial \mathcal{J}_{mn}}{\partial b'_{mn}} = k_n \frac{\partial \mathcal{J}_{mn}}{\partial b_{mn}}. \tag{3.27}$$

This helps to equalize the contribution of each parameter to the descent direction and hence simultaneous estimation of all parameters. One drawback of this approach, compared to a full-Newton method, is that the chosen scaling must be fixed at the beginning of the optimization, and can only be changed by restarting the optimization as a steepest descent.

3.3.4 Numerical optimization

The organization of the numerical optimization scheme in our 2D TTI WEMVA implementation is shown in Figure 3.3. In a first stage the L-BFGS algorithm is fed with an initial model (as shown in the left side of Figure 3.3), which corresponds to the initial diagonally scaled B-spline coefficient ${b'}_{mn}^{init}$. These are typically zero, since we are not fitting the initial models to B-Spline coefficients, and the initial updates are zero. At each iteration or line search step, the objective function and gradient need to be evaluated. The modeling variables V_{P0}, ε and δ necessary for the migration and gradient computation are obtained from the optimization variables b'_{mn}, through a two stage process, as shown in the left of Figure 3.3. First the diagonal scaling is removed by solving equation 3.26 for the B-spline coefficients b_{mn}. Then equations 3.20 through 3.22 are used to evaluate the B-splines. A step by step procedure for computing the objective function and gradient can be described as:

1. Construct $\mathcal{R}$ and evaluate the objective function, and in this procedure store the displacements u_i^s and u_i^r for each shot.

2. Perform separately the two adjoint modelings for each shot, according to equations 3.12 and 3.13 to compute, respectively, the adjoint states $\tilde{u}_i^s$ and $\tilde{u}_i^r$, and at each time step use equations 3.9, 3.10, and 3.11 to build the gradients.

3. Stack each gradient over all shots to obtain the full gradient.

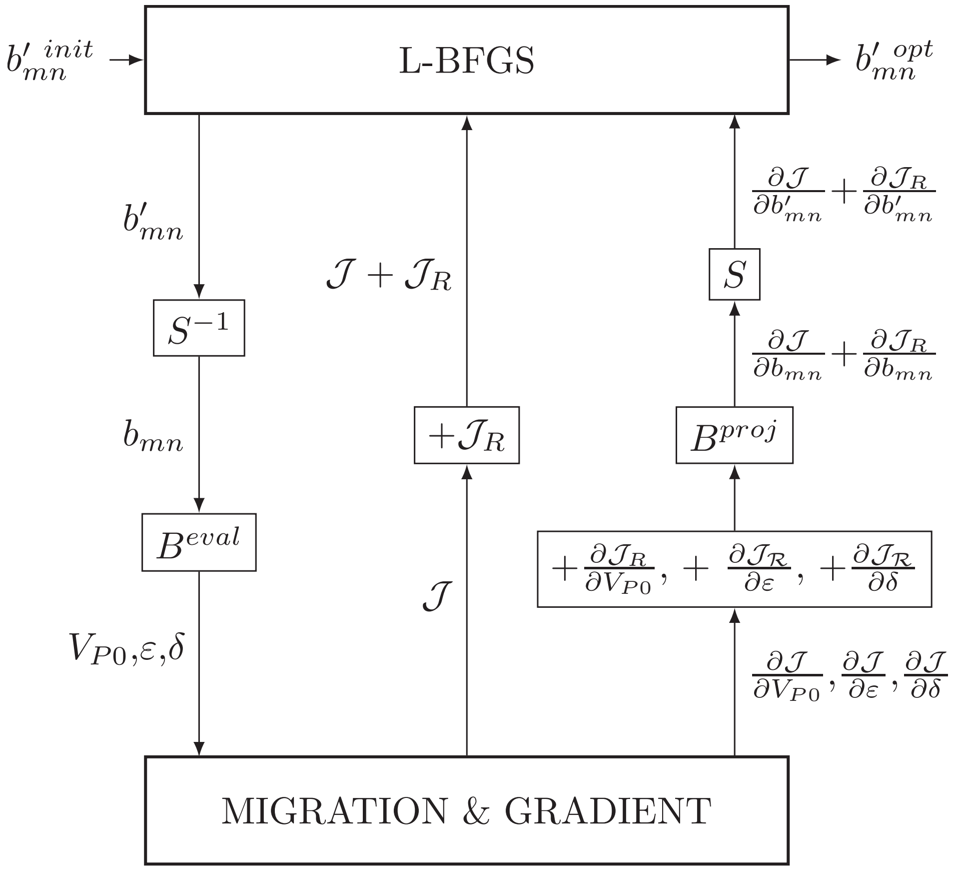


Figure 3.3: Organization of the 2D TTI WEMVA algorithm: B^{eval} and B^{proj} correspond to respectively B-Spline evaluation and projection; S and S^{-1} refer to respectively the forward and inverse diagonal scaling; $b'^{\,init}_{mn}$Âăare the initial diagonally scaled B-spline coefficients, while $b'^{\,opt}_{mn}$ are the optimized diagonally scaled B-spline coefficients.

The computed objective function $\mathcal{J}$ must be augmented by the regularization term $\mathcal{J}_R$, as shown in the center of Figure 3.3. While the gradients of $\mathcal{J}$ with respect to the parameters V_{P0}, ε and δ are augmented with the respective regularization gradients, as shown in the right side of Figure 3.3. Finally, these gradients are projected into a B-spline basis using equations 3.23 through 3.25, and diagonally scaled using equation 3.27.

The loop is repeated until some convergence criteria is met, or a predetermined number of iterations have been run. At which point, the optimized parameters $b'^{\,opt}_{mn}$ are outputted, as shown in the right of Figure 3.3

3.4 Synthetic 2D example

The first example of TTI WEMVA is based on the synthetic velocity model shown in Figure 3.4. The model is a 2D synthetic cross section of a North sea offshore reservoir. The anisotropic model simulates a 2D TTI medium. This model was used to generate synthetic seismic data using a finite difference solution to the elastic wave equation (Lisitsa and Vishnevskiy, 2010). The geometry of the data consists of a line with minimum offset of 0.15 km and maximum offset of 5 km. Absorbing boundary conditions were used to ensure that the data is free from surface related multiples. However, interbed multiples and converted waves are still present in the data.

In this example, as in the next example, we simultaneously estimate V_{P0}, ε and δ. All parameters are optimized over bi-cubic B-spline grids with 0.8 km spacing in the lateral direction and 0.2 km in the vertical direction. We assume an initial θ model and keep it constant over the course of the minimization. In this case, we use the true θ model, shown in Figure 3.4d.

Regularization consisted in constraining the anisotropic parameters ε and δ to be zero at the water layer and positive in the sediments. The constant α_1, controlling the derivative regularization of the V_{P0} model, was set so that the value of the regularization was 1% of the total initial objective function value, while α_2 and α_3 were set to values 10000 times larger than that of α_1. A taper is applied to mute the gradient in the water layer, simulating a situation where the velocity of the water is known. The maximum frequency of the data was filtered down to 30 Hz, so that a coarse grid of 0.02 by 0.02 km could be used for modeling and migration.

The starting point for the velocity analysis is an isotropic 1D velocity model, shown in Figure 3.5a. The model is constructed from a single smoothed trace of the true velocity model. The result of optimization on the parameters after 27 iterations are shown in Figure 3.5b-d. From this figure, we can see that the updated V_{P0} model is able to partially capture the main background features of the true V_{P0} model. On the other hand, the estimated anisotropic parameters show a strong imprint of the P-wave velocity, and a general lack of structure. This reveals a strong dependency between the different parameters, which in this case appears to be a major contributor to the non-uniqueness of the result. Of all

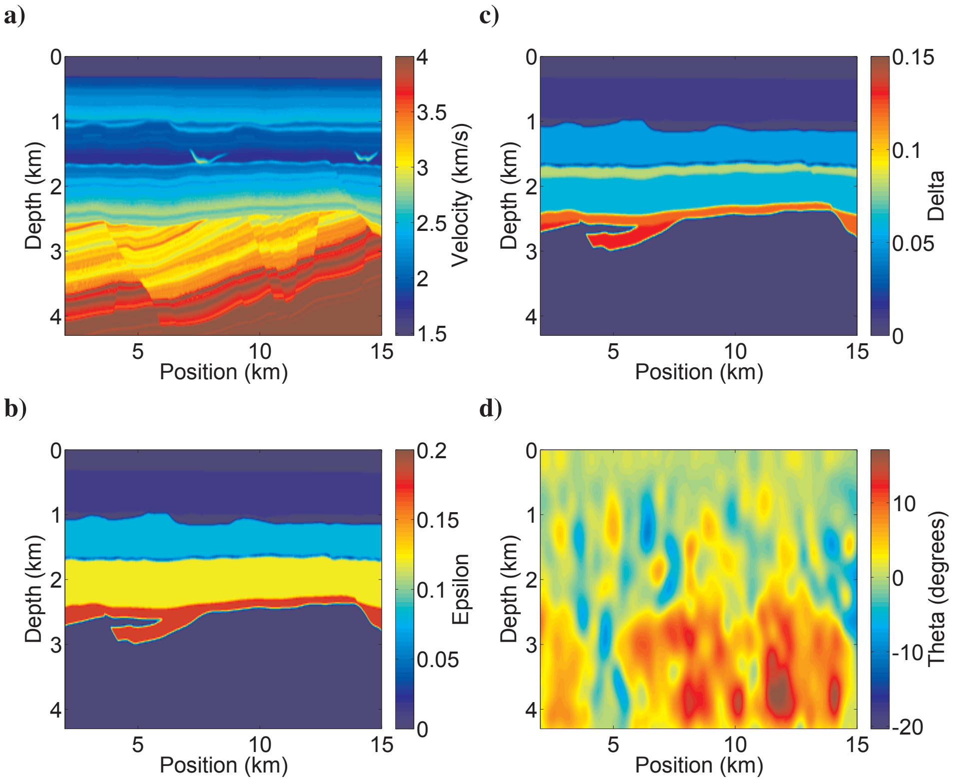


Figure 3.4: 2D TTI synthetic model of a North sea reservoir. (a) V_{P0}. (b) ε. (c) δ. (d) θ.

parameters estimated, δ seems to be the most poorly constrained. We note that there is a tendency for δ to be overestimated in the shallow parts of the model and underestimated in the deeper parts of the model. The reason for this behavior is twofold. The first reason is the poor separation between the effects of V_{P0} and δ on the kinematics of the image. The second reason is related to the poor scaling of the optimization, which causes the objective function to be dominated by the sensitivity to δ in the shallow parts of the image and by V_{P0} in the deeper parts. This means that in the shallow parts of the model, δ is mainly compensating for the kinematic errors introduced by an underestimated V_{P0}, and vice-versa for the deeper parts of the model.

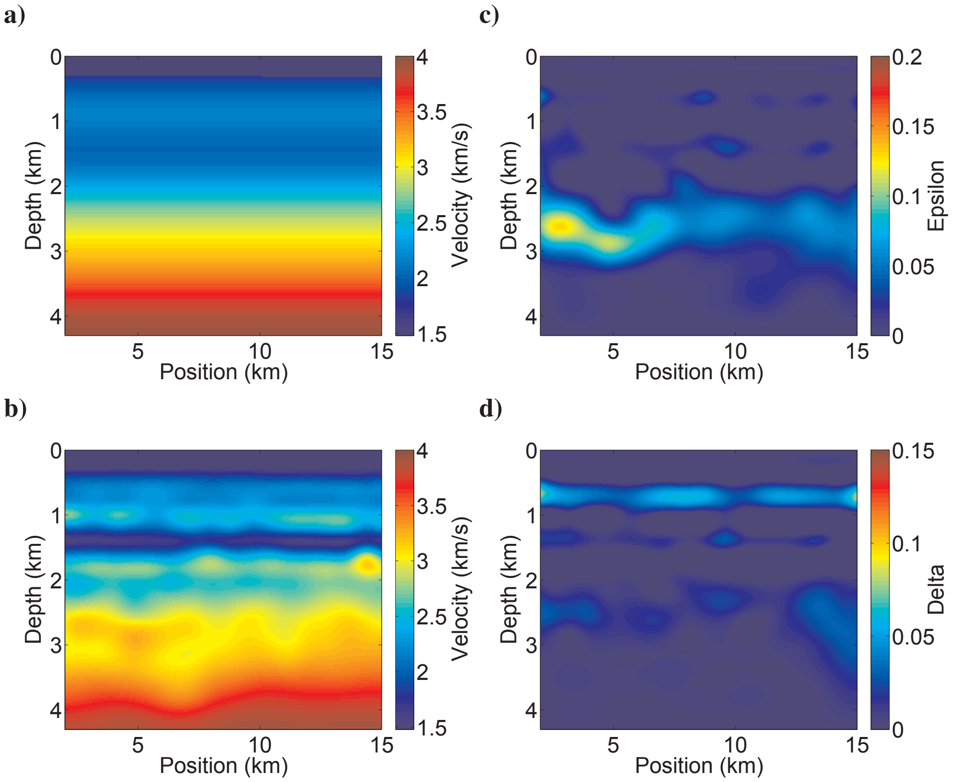


Figure 3.5: Initial and updated models used for the 2D synthetic data example. (a) Initial V_{P0} model. (b) Updated V_{P0} model after 27 iterations. (c) Updated ε model after 27 iterations. (c) Updated δ model after 27 iterations. The θ model used in this example is shown in Figure 3.4d

Figure 3.6 shows a comparison of the RTM images produced with the initial 1D, WEMVA, and true model parameters. The initial image has large mispositionings and is poorly focused due to the inaccurate initial background velocities. These issues are partially fixed in the optimized migrated image, which is better focused. But there are some mispositionings (up to more than 0.05 km) in the optimized image, in particular below 3 km depth. To help better compare the spatial positioning of the reflectors in the images, we have drawn three arrows in fixed positions in the images.

A selection of subsurface offset CIGs constructed using the initial model, the optimized model and the true model are shown in Figures 3.7a-c. The figure shows that the energy in the updated CIGs are now better focused at the zero lag, when compared to the initial CIGs. A comparison of the true model CIGs with the updated CIGs reveals that the updated CIGs are slightly mispositioned ($>$50 m) in depth. The fact that both the updated and true models can produce CIGs that are focused at zero lag, yet at different positions in space, reveals one source of non-uniqueness. This seems to repeat the generally known fact that, in presence of anisotropy, focusing of prestack depth migrated images does not guarantee a unique positioning of the subsurface reflectors.

Now we repeat the velocity analysis using the same parameters as before, but this time we use a fixed V_{P0} model obtained by smoothing the true model. The V_{P0} model is shown in Figure 3.8a, while the results of velocity analysis for ε, and δ are shown respectively in Figures 3.8a and 3.8b. The results show that by adding additional information, in this case V_{P0}, we can obtain better estimates of the anisotropic parameters.

3.5 Field 2D example

In the next example, we apply the method on a real data set taken off the North Sea, offshore Norway. The data is originally a 3D marine data set, from which we extracted a 2D line. The geometry of the data consists of a line with minimum offset of 0.15 km and maximum offset of 5 km. The data processing included multiple attenuation, and muting of direct wave, wide-angle reflections and refractions. The maximum frequency of the data was filtered down to 30 Hz, so that a coarse grid of 0.02 by 0.02 km could be used for modeling and migration.

The initial model for the optimization is shown in Figure 3.9a. It consists in an isotropic 1D model created by smoothing a well log of the vertical slowness. Here, we numerically estimate θ from the initial image, and keep it constant during optimization. We try to approximate an STI model of the subsurface, where the symmetry axis is perpendicular to the dip of the beddings. To estimate the tilt angle, we first estimate the smallest positive angle between the spatial gradient of the image and an unitary positive vertical vector. This can be done using the equation

$$\phi(\mathbf{x}) = \cos^{-1}\left[\frac{\left|\frac{\partial\mathcal{R}}{\partial z}(\mathbf{x},0)\right|}{\left(\frac{\partial\mathcal{R}}{\partial x}^2(\mathbf{x},0)+\frac{\partial\mathcal{R}}{\partial z}^2(\mathbf{x},0)\right)^{1/2}}\right]. \tag{3.28}$$

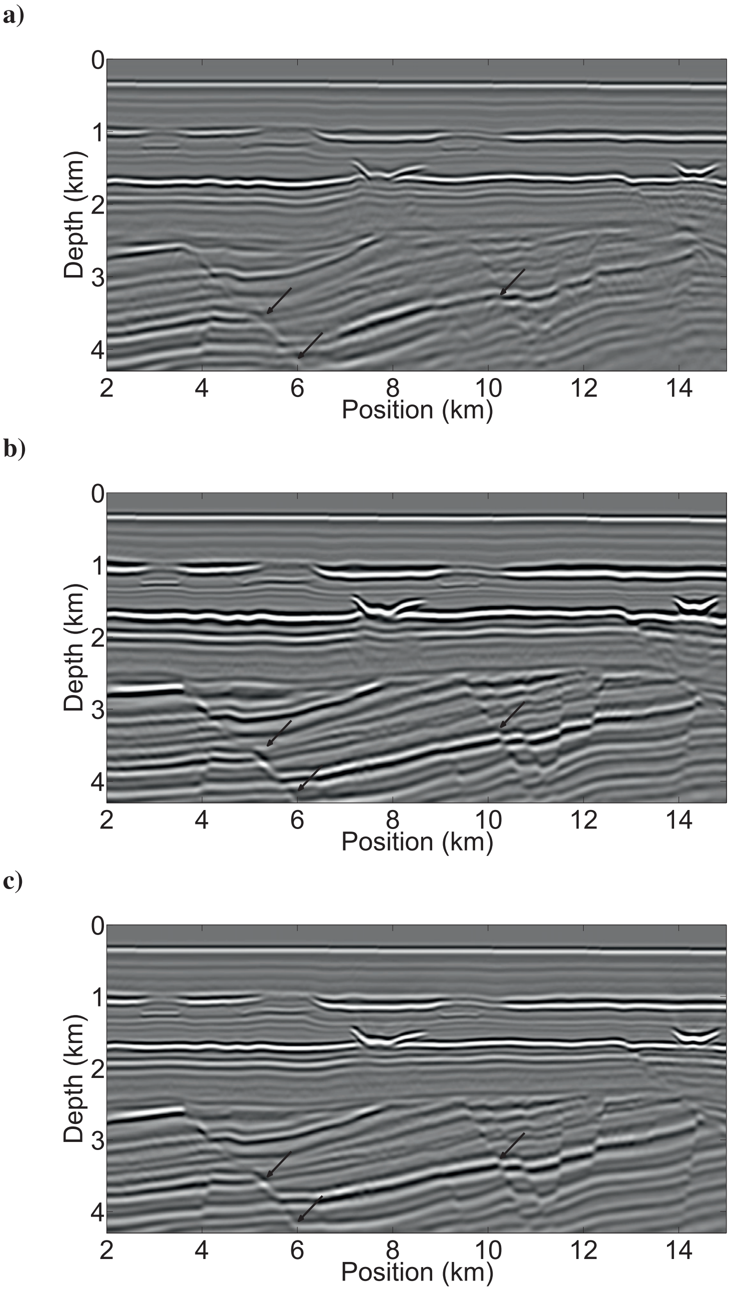


Figure 3.6: Images constructed using: (a) Initial model; (b) Optimized model; (c) True model. Arrows are drawn to help visualize the changes in the spatial positioning of the reflectors.

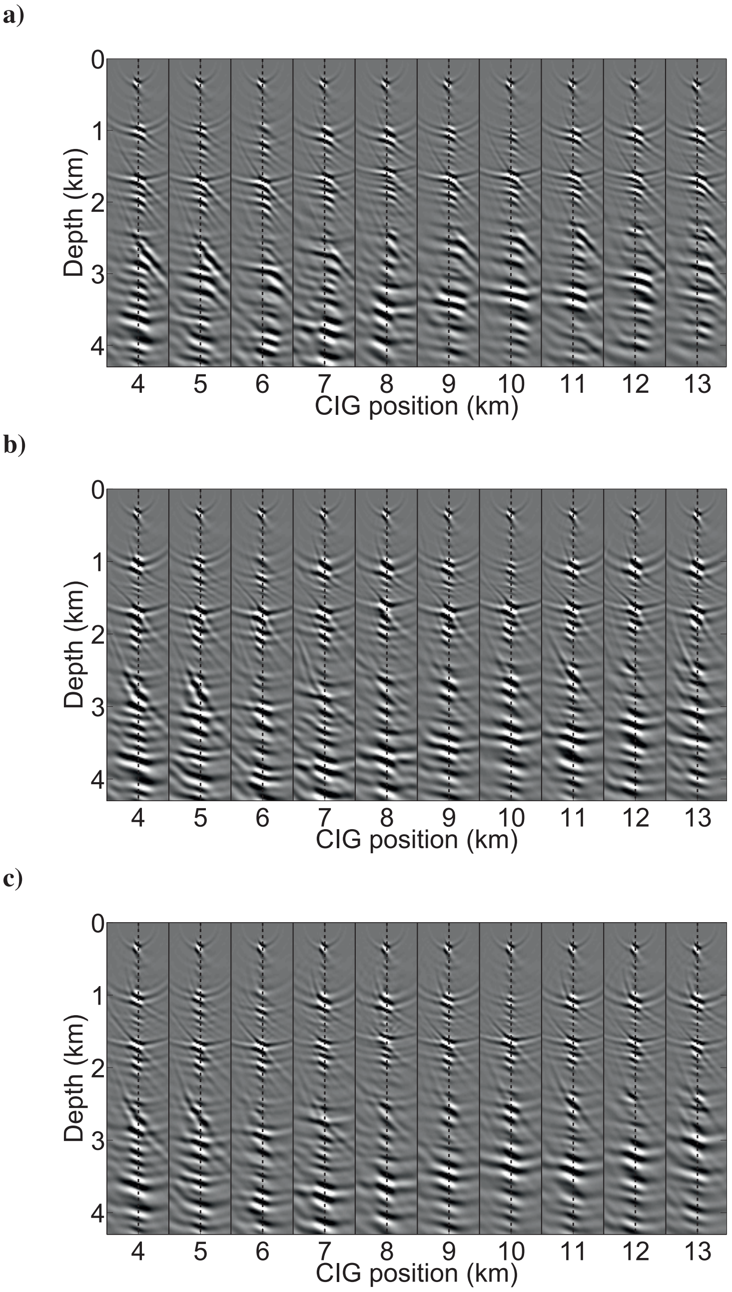


Figure 3.7: CIGs constructed using: (a) Initial model; (b) Optimized model; (c) True model. The black dotted lines mark the position of the zero subsurface offset. The offsets range between -0.5 and 0.5 km.

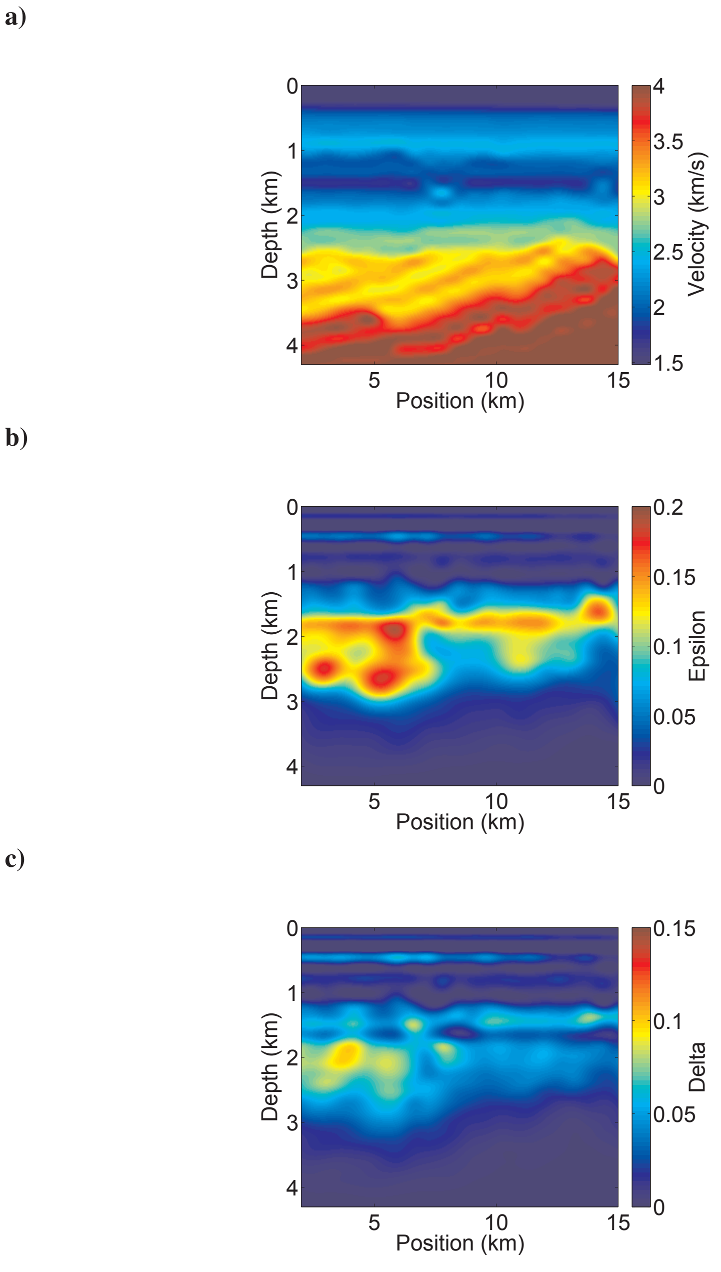


Figure 3.8: Results of 2D TTI WEMVA with a fixed V_{P0} model obtained by smoothing the true model. (a) Smoothed V_{P0}. (b) ε. (c) δ.

To find the sign of the angle, we use the following convention

$$\theta(\mathbf{x}) = \begin{cases} -\phi(\mathbf{x}) & \text{if } \frac{\partial \mathcal{R}}{\partial x}(\mathbf{x},0) \leq 0, \\ \phi(\mathbf{x}) & \text{if } \frac{\partial \mathcal{R}}{\partial x}(\mathbf{x},0) > 0. \end{cases} \tag{3.29}$$

To avoid excessive roughness in the estimate of θ, we low-pass filter the image gradient before computing ϕ, and subsequently low-pass filter θ. The estimated tilt angles are shown in Figure 3.9b. We can see that the estimated angles partially capture the general background trend of the structure in the image. However, some artifacts are apparent, which make the estimate look non-geological. These appear, in particular, at areas where there is a lack of reflectivity, or due to the smearing caused by the smoothing. Since these artifacts are of small angle magnitude ($|\theta|$ <5 degrees), we deem they should have a minor or negligible effect on wave propagation.

As in the previous example, the parameters are optimized over a bi-cubic B-spline grid with 0.8 km spacing in the lateral direction and 0.2 km in the vertical direction. Regularization followed the same guidelines as in the previous example.

The resultant estimated parameters after 47 iterations of 2D TTI WEMVA are shown in Figures 3.10. The estimated V_{P0} model reveal a, nearly 1D, plane-layered overburden down to 2 km. Below 2 km, as expected due to the normal faulted structures, the optimized V_{P0}, ε and δ models show more lateral variation.

The migrated images computed with the initial and updated models are shown in Figure 3.11, while Figure 3.12 shows a comparison on selected CIGs. We can see that optimization locally improves the focusing of the RTM image. These improvements are marked by arrows in Figure 3.11. Also, the energy in the updated CIGs is more focused at zero offset, as can be see in Figure 3.12. However, as in the synthetic data example, there is some uncertainty about the positioning of the reflectors in the final image.

3.6 Discussion

Prestack depth migration of P-wave surface seismic data in 2D TTI media using a density normalized elastic wave equation requires an estimate of the P-wave velocity along the symmetry axis V_{P0}, the S-wave velocity along the symmetry axis V_{S0}, Thomsen's parameters ε, and δ, and the tilt θ of the symmetry axis with respect to the vertical. To reduce the number of parameters we need to estimate, we assume that the kinematics of P-waves are independent of the choice of V_{S0}, and that θ is structurally conforming. Which reduces the velocity estimation problem to three parameters: V_{P0}, ε, and δ. Here we use 2D TTI WEMVA to simultaneously estimate these three parameters from surface seimic data. The results show that the method is able to improve the focusing of the depth migrated image. However, the synthetic data example clearly show that a unique set of anisotropic parameters is not constrained by the method. There are several contributors to this issue. One

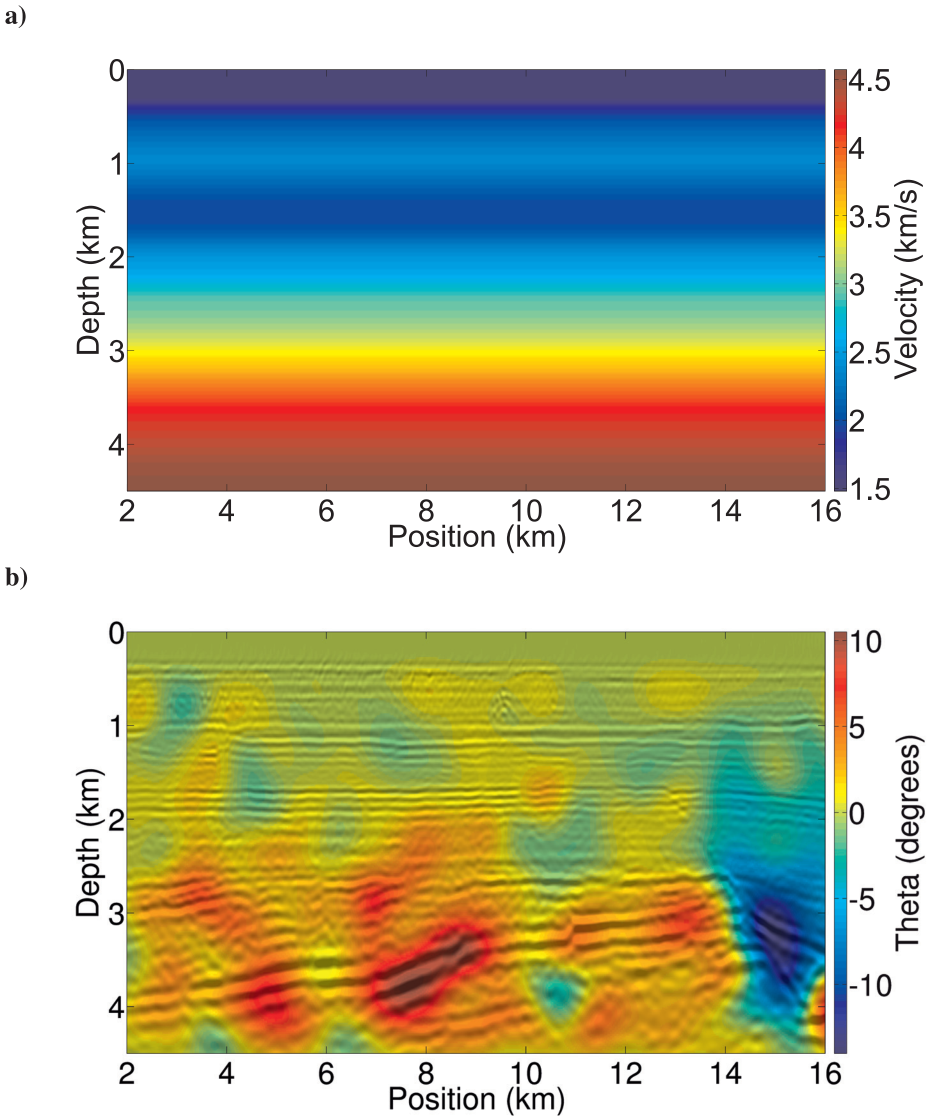


Figure 3.9: Initial models used for the 2D field data example. (a) Initial V_{P0} model. (b) θ model overlaid by the initial image. The tilt angles in (b) are estimated from the reflector dips of the initial image.

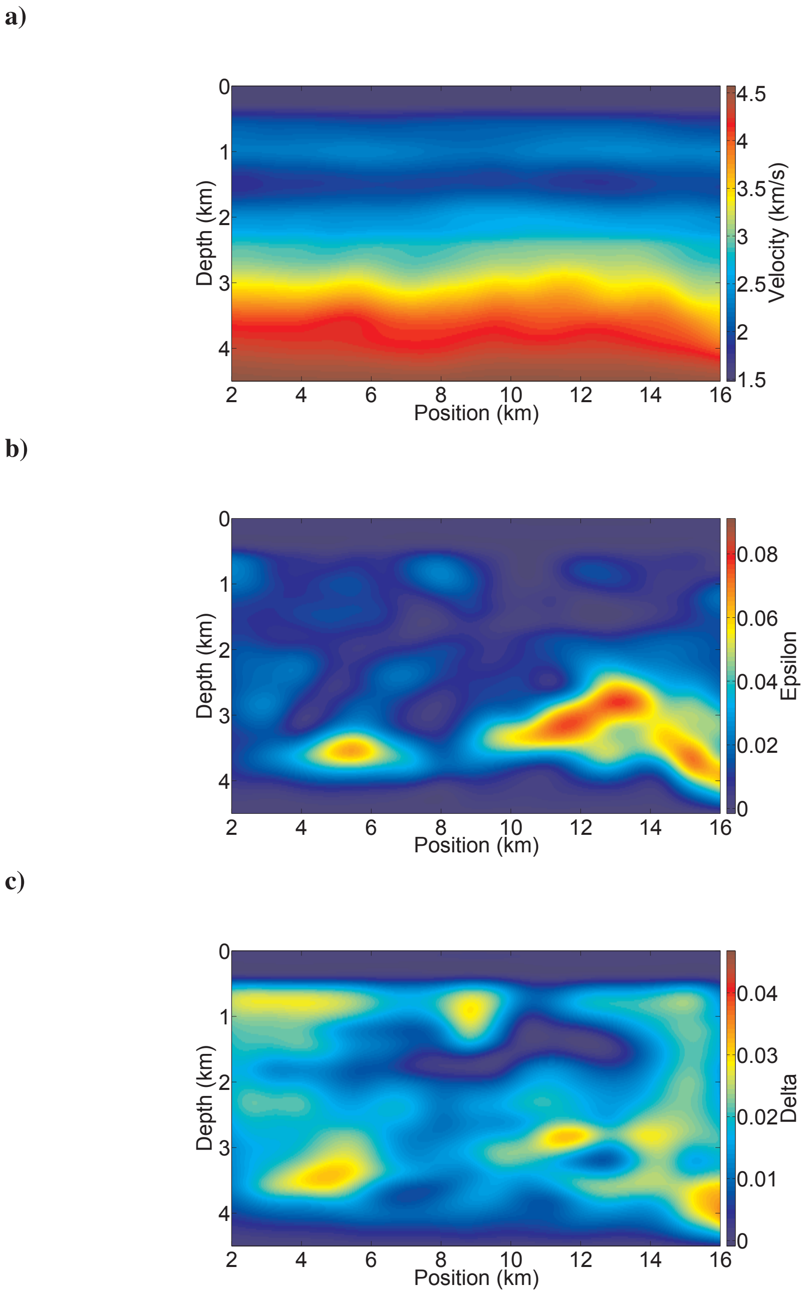
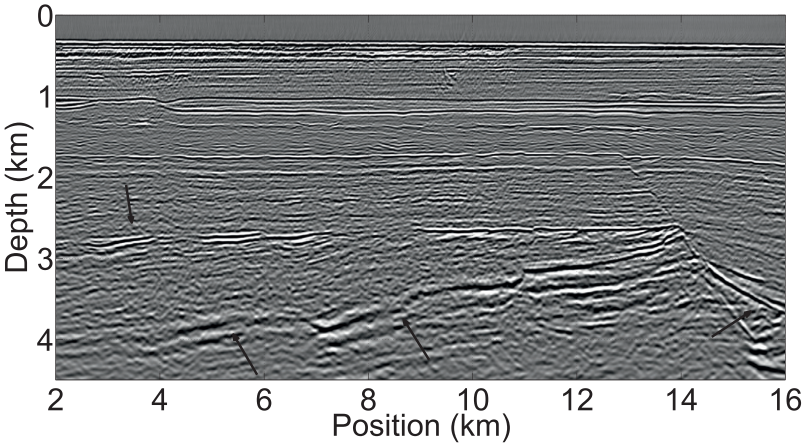


Figure 3.10: Optimized models after 47 iterations of 2D TTI WEMVA. (a) V_{P0}. (b) ε. (c) δ.

a)



b)

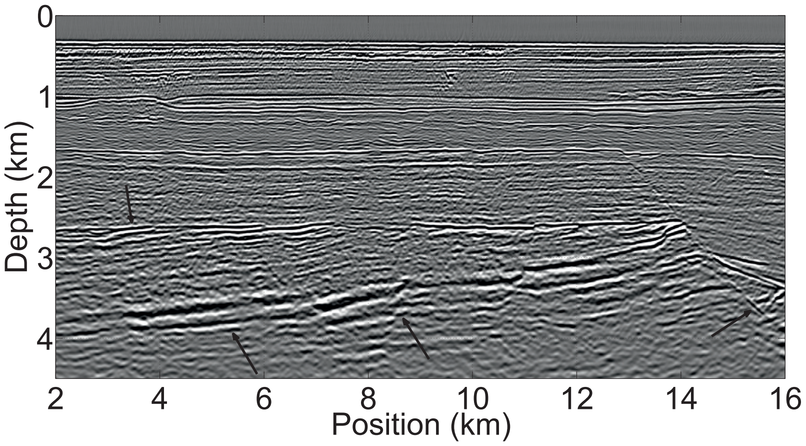


Figure 3.11: Images constructed using: (a) Initial model; (b) Optimized model. Arrows are drawn at fixed positions in the images to mark changes in the spatial positioning and focusing.

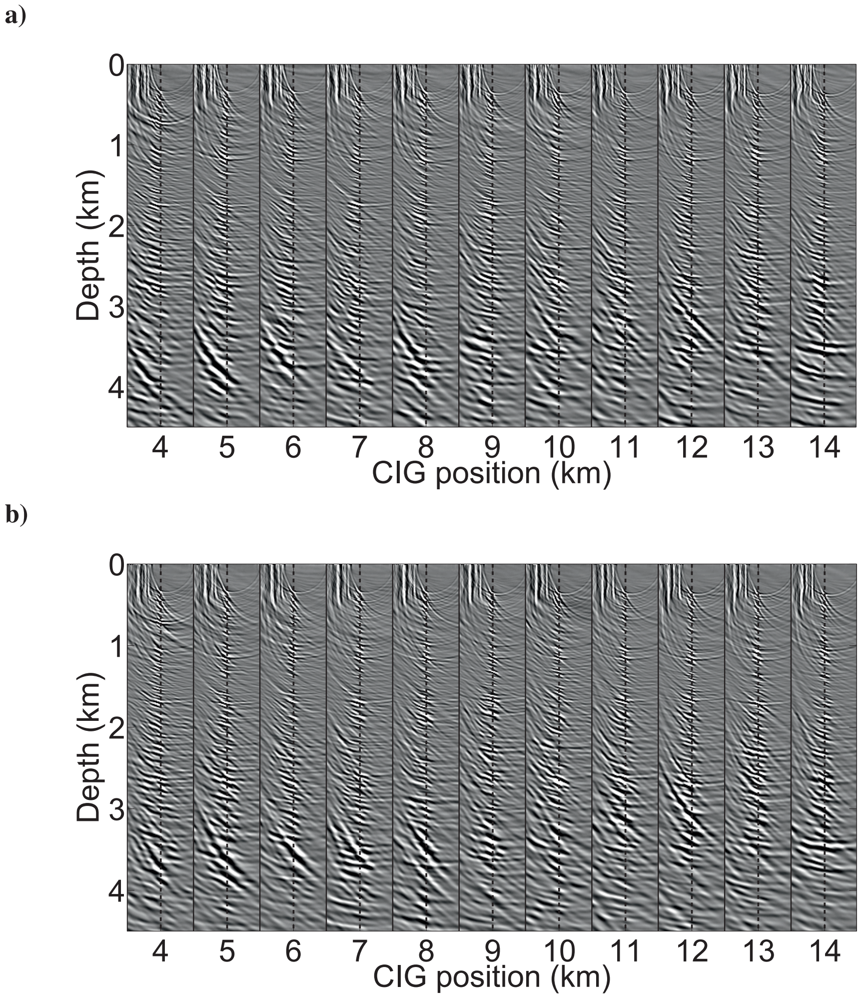


Figure 3.12: CIGs constructed using: (a) Initial model; (b) Optimized model. The black dotted lines mark the position of the zero subsurface offset. The offsets range between -0.5 and 0.5 km.

of them is that, over a transverse isotropic medium, focusing of the prestack depth images does not constrain a unique spatial positioning of the reflectors in a depth migrated image (Isaac and Lawton, 1999; Bakulin et al., 2010). Another reason is the existence of a strong interdependency between the parameters V_{P0}, ε, and δ (Alkhalifah and Tsvankin, 1995; Jones et al., 2003). In fact, Alkhalifah et al. (2001) shows that, under certain conditions, at most two parameters can be recovered from P-wave seismic reflection data. Therefore, if the goal is to obtain a geological model of the subsurface from surface seismic data, a combination of different methods, and additional information, such as well data, must be used to constrain the updated models to a narrower space of possible solutions (Alkhalifah and Tsvankin, 1995; Alkhalifah et al., 2001; Bakulin et al., 2010; Yan et al., 2004).

The choice of parameterization can be important for the results of velocity analysis, because it changes the scaling of the optimization problem (Nocedal and Wright, 2000). Here we use WEMVA to estimate V_{P0}, ε, and δ. In order to estimate more than one parameter simultaneously using a L-BFGS optimization algorithm, we apply diagonal scaling. The idea is to approximately equalize the sensitivity of the objective function to the different parameters. This procedure can be avoided if the parameters have the same units and vary approximately over the same ranges. It is thus possible that the parameterization given by V_{P0}, $V_h = V_{P0}\sqrt{1+2\varepsilon}$, and $V_n = V_{P0}\sqrt{1+2\delta}$, which was suggested by Alkhalifah and Tsvankin (1995), provides a better scaling for the problem. Although we develop the method using V_{P0}, ε and δ, it is trivial to change the parameterization to V_{P0}, V_h, and V_n, by substitution of the quantities on the density normalized stiffness tensor.

The main difference between isotropic and anisotropic WEMVA is in the reconstruction of the source and receiver wavefields. Here, we use a density normalized elastic wave equation to model the source and receiver displacement wavefields over 2D TTI media. And then extract quasi-P wavefields by taking the divergence of the displacements scaled by the density normalized bulk modulus. These wavefields are input to an extended crosscorrelation imaging condition to create PP CIGs of the subsurface. However, the elastic wave equation also naturally models S-waves. This opens for the possibility to apply WEMVA to multicomponent data. Instead of taking the divergence of the displacements, we could also extract scalar wavefields based on the curl of the displacement, which would allow us to create PS, and SS images of the subsurface. The 2D TTI WEMVA algorithm we present can be extended to use all these images, allowing to simultaneously estimate V_{P0}, and V_{S0}, as well as to improve the accuracy in the estimation of the anisotropic parameters ε and δ (Tsvankin and Thomsen, 1995; Grechka et al., 2002).

The WEMVA method we present is restricted to single scattering reflection data. This is a limitation shared by all methods based on the image domain, and prevents us from using valuable information contained in refractions, and multiple reflections. Data domain methods based on full-waveform inversion (FWI) can, in theory, use all information contained in the data (Tarantola, 1984; Virieux and Operto, 2009; Plessix and Rynja, 2010). On the other hand, for FWI to converge to the solution it requires either the presence of useful signal at very low frequencies (<4Hz), or an initial model that is kinematicaly close to the solution (Virieux and Operto, 2009). Combining FWI and WEMVA is non-trivial, since in practice the two methods have very different restrictions when it comes to the input data and initial models. However, one possible strategy is to simply use WEMVA to create initial models that are kinematically close enough to the solution, and then further refine them using FWI.

3.7 Conclusion

We present a WEMVA method based on anisotropic 2D TTI RTM. To reconstruct the wavefields needed for the imaging condition, we use a density normalized elastic wave equation. The WEMVA method can be used to simultaneously estimate the parameters V_{P0}, ε, and δ from surface seismic reflection data, using an objective function consisting of a combination of depth oriented differential semblance and stack-power maximization. In this procedure, the tilt of the symmetry axis with respect to the vertical is assumed to conform to the initial reflectivity geometry and not updated, and the S-wave velocities are heuristically chosen and also kept constant.

The method is tested on both synthetic and field data. The tests show that the method is well-posed and converges to a model which produces well focused images. On the other hand, the results also show that 2D TTI WEMVA of surface seismic data alone is not sufficient to constrain a unique anisotropic model of the subsurface. This means that to obtain a correct positioning of the reflectors in the subsurface, the method must be complemented with additional information.

Since we employ the elastic wave equation for wavefield reconstruction, the method can be easily generalized to any anisotropic medium. Furthermore, the quantities in the stiffness tensor can easily be transformed to honor other parameterizations than the one used in this paper.

3.8 Acknowledgments

Wiktor Weibull likes to thank Mariane Houbiers and Joachim Mispel for interesting discussions about the subject. This author also likes to thank Statoil Petroleum AS for financial support of this work. We acknowledge the partners in the Snorre license, Statoil Petroleum AS, Petoro AS, ExxonMobil Exploration and Production Norway AS, Idemitsu Petroleum Norge AS, RWE Dea Norge AS, and Core Energy AS for permission to publish the results.

3.9 Appendix A: 2D elastic TTI medium

We consider the following system describing 2D elastic wave propagation

$$\partial_t u_x = \partial_x \tau_{xx} + \partial_z \tau_{xz}, \tag{3.30}$$

$$\partial_t u_z = \partial_x \tau_{xz} + \partial_z \tau_{zz}, \tag{3.31}$$

$$\begin{pmatrix} \partial_t \tau_{xx} \\ \partial_t \tau_{zz} \\ \partial_t \tau_{xz} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} & a_{15} \\ a_{13} & a_{33} & a_{35} \\ a_{15} & a_{35} & a_{55} \end{pmatrix} \begin{pmatrix} \partial_x u_x \\ \partial_z u_z \\ \partial_x u_z + \partial_z u_x \end{pmatrix} + \begin{pmatrix} S_{xx} \\ S_{zz} \\ S_{xz} \end{pmatrix}, \tag{3.32}$$

where $\partial_t = \frac{\partial}{\partial t}$,$\partial_x = \frac{\partial}{\partial x}$, and $\partial_z \frac{\partial}{\partial z}$; u_i is the particle displacement vector, τ_{ij} Âăis a stress tensor, S_{ij} is an external source tensor; $A = (a_{ij})$ is the matrix of density normalized elastic coefficients. This matrix is obtained from the fourth order tensor a_{ijkl} using the Voight notation: $xx \rightarrow 1$, $zz \rightarrow 3$, $xz \rightarrow 5$ (Winterstein, 1990).

In a 2D TTI medium, A is given by the following elastic coefficients

$$\begin{aligned}
a_{11} &= a'_{11}\cos^4\theta + a'_{33}\sin^4\theta + 2(a'_{13} + 2a'_{55})\sin^2\theta\cos^2\theta, && (3.33)\\
a_{13} &= (a'_{11} + a'_{33} - 4a'_{55})\sin^2\theta\cos^2\theta + a'_{13}(\sin^4\theta + \cos^4\theta), && (3.34)\\
a_{15} &= (a'_{13} - a'_{11} + 2a'_{55})\sin\theta\cos^3\theta, && \\
&\quad +(a'_{33} - a'_{13} - 2a'_{55})\sin^3\theta\cos\theta, && (3.35)\\
a_{33} &= a'_{11}\sin^4\theta + a'_{33}\cos^4\theta + 2(a'_{13} + 2a'_{55})\sin^2\theta\cos^2\theta, && (3.36)\\
a_{35} &= (a'_{13} + 2a'_{55} - a'_{11})\sin^3\theta\cos\theta, && \\
&\quad +(a'_{33} - a'_{13} - 2a'_{55})\sin\theta\cos^3\theta, && (3.37)\\
a_{55} &= (a'_{11} + a'_{33} - 2a'_{13} - 2a'_{55})\sin^2\theta\cos^2\theta + a'_{55}(\sin^4\theta + \cos^4\theta), && (3.38)
\end{aligned}$$

where θ is the tilt angle of the symmetry axis with respect to the vertical, and $A' = (a'_{ij})$ are quantities given by

$$\begin{aligned}
a'_{11} &= V_{P0}^2(1 + 2\varepsilon), && (3.39)\\
a'_{13} &= \left[2\delta V_{P0}^2\left(V_{P0}^2 - V_{S0}^2\right) + \left(V_{P0}^2 - V_{S0}^2\right)^2\right]^{1/2} - V_{S0}^2, && (3.40)\\
a'_{33} &= V_{P0}^2, && (3.41)\\
a'_{55} &= V_{S0}^2, && (3.42)
\end{aligned}$$

where V_{P0} is the P-wave velocity along the symmetry axis, V_{S0} is the S-wave velocity along the symmetry axis, and ε and δ are Thomsen's anisotropic parameters.

Chapter 4

Reverse-time demigration using the extended imaging condition

Wiktor Weibull and Børge Arntsen

Norwegian University of Science and Technology, Trondheim, Norway

Abstract

The forward and inverse process of seismic migration and demigration or remodeling has many useful applications in seismic data processing. We present a method to re-obtain the seismic reflection data after migration, by inverting the common image point gathers produced by reverse-time migration with an extended imaging condition. This allows to convert the results of seismic data processing in the stacked image domain back to the prestack reflection data domain. To be able to reconstruct the data with high fidelity, we set up demigration as a least squares inverse problem, and solve it iteratively using a steepest descent method. Because we use an extended imaging condition, the method is not dependent on an accurate estimate of the migration velocity field, and is able to accurately reconstruct both primaries and multiples. At the same time, because the method is based on reverse-time migration it can accurately handle seismic reflection data acquired over complex geological media. Numerical results show the feasibility of the method, and highlight some of its applications on 2D synthetic and field data sets.

Scheduled for presentation at the 82nd SEG Annual International Meeting in September 2013. Paper submitted to Geophysics 2013.

4.1 Introduction

The motivation behind this work was to obtain a method to reconstruct seismic reflection data from common image point gathers (CIGs) constructed with reverse-time migration (RTM). The method should work without the need for an accurate velocity model, and the reconstructed data should have an acceptably small error in amplitude and phase. This would ultimately allow us to process data in the migrated domain, which can be an advantage in many situations, such as velocity analysis, illumination studies, data interpolation and multiple attenuation.

Demigration methods have a long history in seismic data processing. Loewenthal et al. (1976) introduced the concept of exploding reflector model, showing how to obtain zero-offset seismic data from a migrated stack using a background velocity model and wave theoretical methods. The Kirchhoff integral and the high frequency approximation have also been used for reconstruction of seismic data form migrated images (Jaramillo and Bleistein, 1999; Santos et al., 2000; Miranda, 2006). More recently, RTM has been used to recreate data from seismic images with the purpose of velocity analysis (Chauris and Benjemaa, 2010), and multiple attenuation (Zhang and Duan, 2012).

In their work, Chauris and Benjemaa (2010) uses the concept of the extended imaging condition (Sava and Vasconcelos, 2011) in a migration/demigration scheme. The advantage of the extended imaging condition over the classical cross-correlation imaging condititon (Claerbout, 1971), is that it preserves the phase and angle dependent amplitude information of the data in the migrated image, even in the case of migration with an inaccurate velocity model. Here we explore this fact and use extended images to set up demigration as an inverse problem. We try to reconstruct the prestack seismic reflection data from the migrated image by minimizing a least-squares function. And we solve the problem iteratively using a steepest descent method.

Results from 2D field and synthetic numerical examples show that phase information can be recovered after only one iteration, whereas the amplitude information can require many steepest descent iterations.

4.2 The method

The main purpose of the method is to be able to reconstruct seismic data from CIGs constructed using RTM with an extended imaging condition (Sava and Vasconcelos, 2011). In the extended imaging condition, instead to Claerbout (1971) classical cross-correlation of the source and receiver wavefields at the imaging point, CIGs are constructed by cross-correlating the source and receiver wavefields at symmetric lags around the imaging point. These cross-correlation lags can be either spatial (Rickett and Sava, 2002) or temporal lags (Sava and Fomel, 2006). The important point is that, different from the classical imaging condition, the extended imaging condition preserves the kinematic and angle dependent information of the data in the image, even in the case of migration with an inaccurate velocity model. We now show how we can use the extended imaging condition to set up a demigration method. We demonstrate the method using a time domain implementation of RTM with a space-lag cross-correlation imaging condition (Rickett and Sava, 2002)

$$\mathcal{R}(\mathbf{x},\mathbf{h}) = \int ds \int dt\, W_s(\mathbf{x}-\mathbf{h},t,s) \int d\mathbf{x}' \int dt'\, G(\mathbf{x}+\mathbf{h},t;\mathbf{x}',t') P(\mathbf{x}',t',s), \quad (4.1)$$

where $\mathcal{R}^0$ are CIGs (extended image), $\mathbf{x} = (x, y, z)$ are the spatial coordinates, $\mathbf{h} = (h_x, h_y, h_z)$ are spatial lags, t is the time, s is the source index, W_s are source wavefields, G is the acoustic Green's function, $P^0(\mathbf{x}', t', s)$ are common shot gathers.

The source wavefields are given by

$$W_s(\mathbf{x},t,s) = \int d\mathbf{x}' \int dt'\, G(\mathbf{x},t;\mathbf{x}',t') S(\mathbf{x}',t',s), \quad (4.2)$$

where S are source functions.

Assume now that we have the CIGs ($\mathcal{R}^0$), and we would like to obtain the data (P^0), that is, we are interested in the inverse procedure of equation 4.1. One approach is to apply the adjoint of migration, which, for the extended imaging condition, can be written as (Weibull and Arntsen, 2013)

$$P(\mathbf{x},t',s) = \int d\mathbf{x}' \int dt'\, G(\mathbf{x},t;\mathbf{x}',t') \int d\mathbf{h} \frac{\partial^2 \mathcal{R}^0}{\partial z^2}(\mathbf{x}'-\mathbf{h},\mathbf{h}) W_s(\mathbf{x}'-2\mathbf{h},t',s). \quad (4.3)$$

This equation has been successfully used by Weibull and Arntsen (2013) to reconstruct seismic data from muted CIGs. A similar equation has been used by Chauris and Benjemaa (2010) for velocity analysis. One problem with this modeling equation is that, even if it properly recreates the kinematics, it gives the wrong amplitudes for the data. Another approach, and the one that is proposed by this paper, is to cast the problem as a least squares inversion of the following objective function

$$\mathcal{J} = \frac{1}{2} \int d\mathbf{x} \int d\mathbf{h} \left[\frac{\partial \mathcal{R}^0}{\partial z}(\mathbf{x},\mathbf{h}) - \frac{\partial \mathcal{R}}{\partial z}(\mathbf{x},\mathbf{h}) \right]^2 . \quad (4.4)$$

Here $\mathcal{R}^0$ are CIGs to be demigrated, and $\mathcal{R}$ are forward mapped CIGs. The forward mapped CIGs are computed according to equation 4.1, but with unknown shot gathers $P(\mathbf{x}', t', s)$. Note that $\mathcal{R}$ is migrated using the same source wavefields W_s as $\mathcal{R}^0$. The vertical spatial derivatives in equation 4.4 are used to remove well known artifacts from RTM images (Guitton et al., 2007).

By minimizing this objective function, we seek to find the data that, when migrated, will approximate the image $\frac{\partial R^0}{\partial z}$ in a least squares sense. In principle, because of the linear relationship between the data and the receiver wavefields, the problem is linear and its solution can be sought explicitly. However, here we choose to solve the proposed least squares problem using a steepest descent method (Nocedal and Wright, 2000). This means that the demigrated shot gathers are updated iteratively according to

$$P_{i+1}(\mathbf{x},t,s) = P_i(\mathbf{x},t,s) - \alpha_i \frac{\partial \mathcal{J}}{\partial P_i}(\mathbf{x},t,s), \tag{4.5}$$

where $i \in (1, ..., N)$ is the iteration index, α_i is a positive step length, and $\frac{\partial \mathcal{J}}{\partial P_i}$ is given by

$$\frac{\partial \mathcal{J}}{\partial P_i}(\mathbf{x},t,s) = \int d\mathbf{x}' \int dt'\, G(\mathbf{x},t;\mathbf{x}',t') \int d\mathbf{h} \frac{\partial^2 \Delta\mathcal{R}_i}{\partial z^2}(\mathbf{x}'-\mathbf{h},\mathbf{h}) W_s(\mathbf{x}'-2\mathbf{h},t',s), \tag{4.6}$$

with the image residual $\Delta\mathcal{R}_i$ being given by

$$\Delta\mathcal{R}_i(\mathbf{x},\mathbf{h}) = \mathcal{R}^0(\mathbf{x},\mathbf{h}) - \int ds \int dt\, W_s(\mathbf{x}-\mathbf{h},t,s) \int d\mathbf{x}' \int dt'\, G(\mathbf{x}+\mathbf{h},t;\mathbf{x}',t') P_i(\mathbf{x}',t',s). \tag{4.7}$$

4.3 Numerical examples

We illustrate the method with some 2D seismic examples. The two first examples are based on the Marmousi model (Versteeg, 1993), which has become a benchmark model for complex geology. This model, shown in Figure 4.1, is used to generate seismic data using a finite difference modeling code (Virieux, 1986). We use a monopole point source and Ricker wavelet with dominant frequency of 20 Hz. The source spacing, and the receiver spacing are both 25 m. The shot gathers have a minimum offset of 0 km and a maximum offset of 5 km.

In the last example, we explore the application of demigration to field data. The data consist of a 2D marine seismic line acquired over the Norwegian North Sea. There are 460 shot gathers, each with a minimum offset of 75 m and a maximum offset of 1250 m. The source spacing, and the receiver spacing are both 12.5 meters. The dominant frequency of the data is $\approx$ 30 Hz, and the maximum frequency is $\approx$ 80 Hz.

4.3.1 Example 1

In the first example, we show the ability of the method to reconstruct the seismic prestack data from stacked migrated CIGs. In this procedure, the CIGs that are input for demigration, are the CIGs that are output from RTM with the original data, without any modification. The data that is output from demigration is then compared to the original data after 1, and after many iterations. The migration velocity model is shown in Figure 4.2. The migrated stacked image (zero lag) and a collection of CIGs are shown in Figure 4.3. The demigration is carried out without knowledge of the original data, except for its geometry. In other words, the initial shot gathers are a collection of zeroed traces.

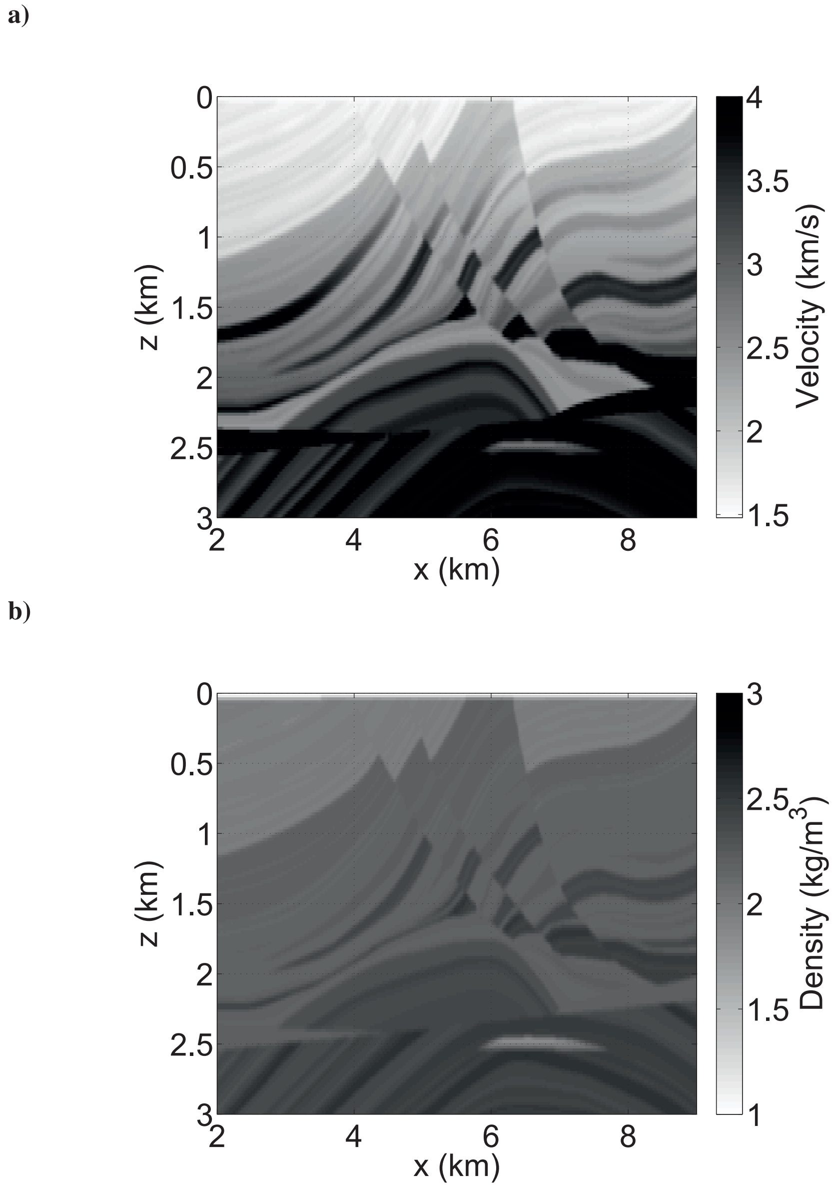


Figure 4.1: Marmousi acoustic model. a) Velocity model. b) Density model.

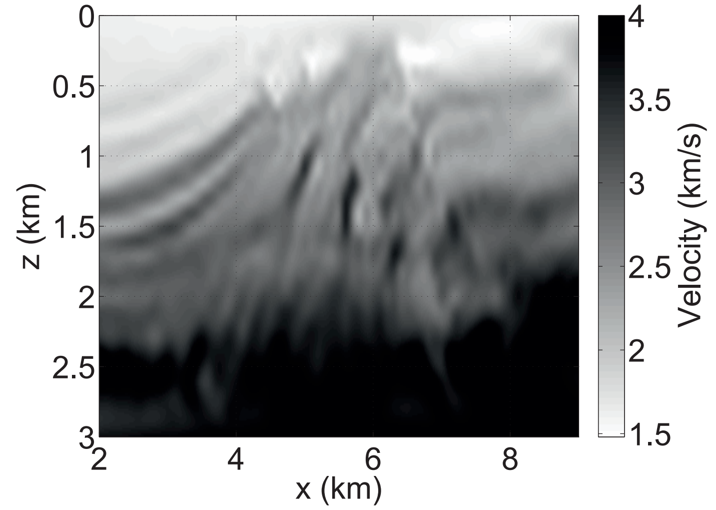


Figure 4.2: Migration velocity model for the Marmousi data set.

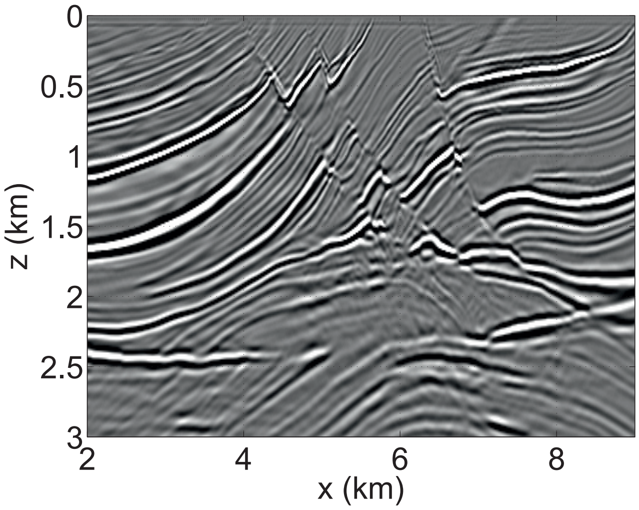
Figure 4.4 shows a comparison of a particular shot gather at source position 7.83 km of the original data with the result of demigration of the CIGs after one iteration, and after 5 iterations of demigration. In this comparison, the reconstructed shot gathers were scaled by an optimal scalar constant. This constant was found by minimizing the least squares difference of the amplitudes between the reconstructed shot gathers and the original shot gathers.

Figure 4.5 shows a comparison of traces from the original data, and from the demigrated data after 1 iteration, and after 5 iterations of demigration. The results show that the kinematics of the data are reconstructed already after one iteration. And after 5 iterations, the data amplitudes are getting closer to the ones in the original data, as can be seen in the comparison of the time traces (Figures 4.5a-c), as well as in the comparison of the amplitude spectra (Figure 4.5d).

4.3.2 Example 2

In the second example, we explore the application of demigration to interpolation of data. In this application, the Marmousi data of the previous example are decimated by only taking every 8th receiver. The resulting receiver interval of 200 m introduce severe dip aliasing in the seismic data recording. Migration using the decimated data results in CIGs such as the one shown in Figure 4.6a. The energy in the CIG that is outside the black dotted lines represents the migrated aliased events. This can clearly be concluded after comparing this CIG with the ones in Figure 4.3b, which were migrated with the original non-decimated geometry. Trying to reconstruct the data with the aliased energy in the

a)



b)

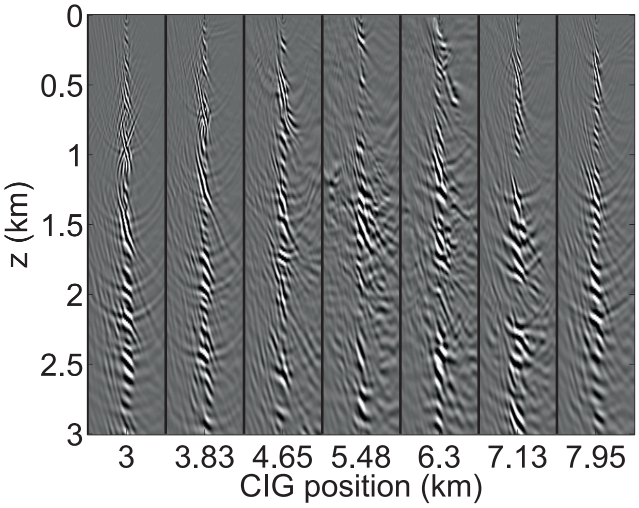


Figure 4.3: Marmousi migrated a) zero-lag image, and b) CIGs at several selected spatial positions.

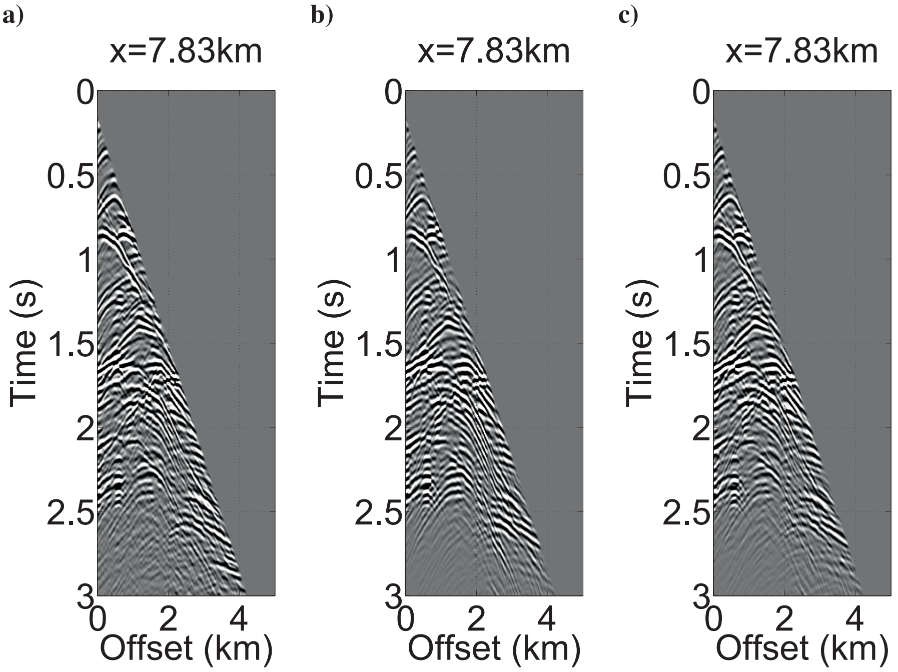


Figure 4.4: Marmousi shot gathers a) original, b) reconstructed after 1 iteration of demigration, c) reconstructed after 5 iterations of demigration.

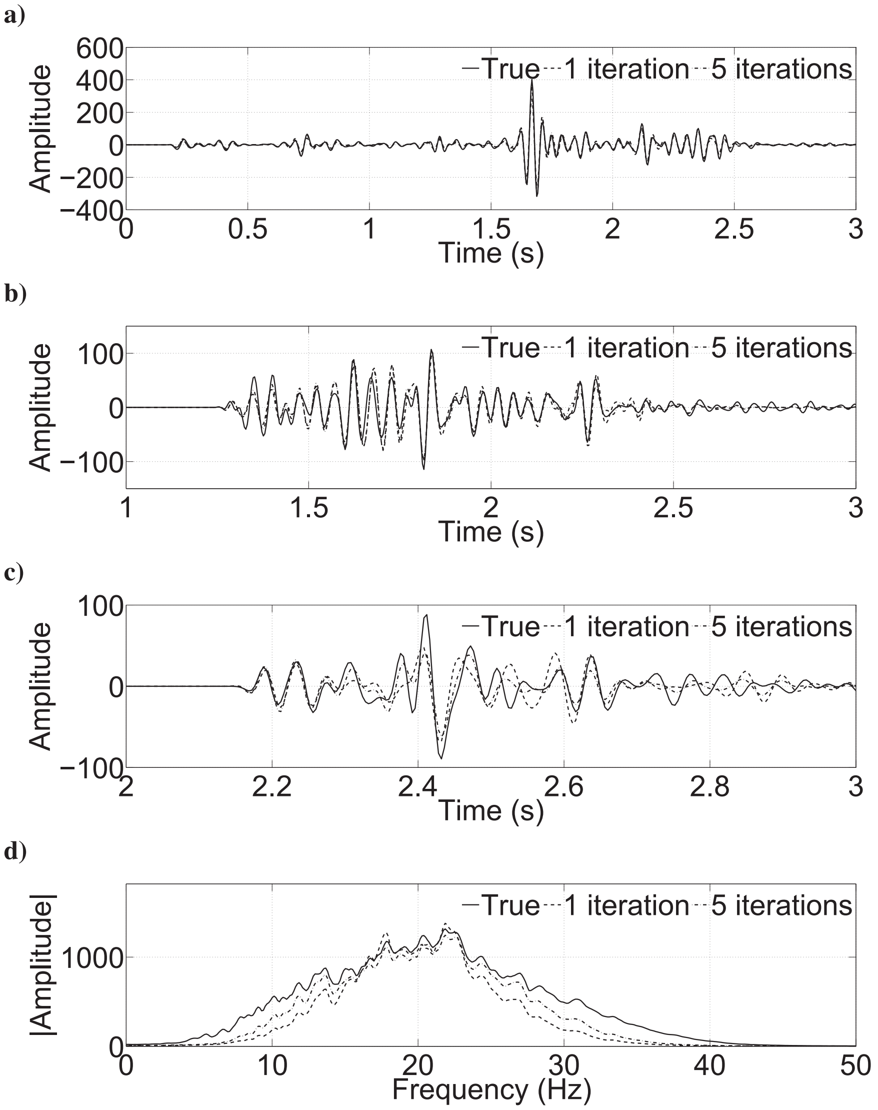


Figure 4.5: Comparison of traces of the Marmousi shot gather at source position 7.83 km a) at zero offset, b) at 1.65 km offset, c) at 3 km offset; d) Comparison of amplitude spectra averaged over all traces of the shot gather.

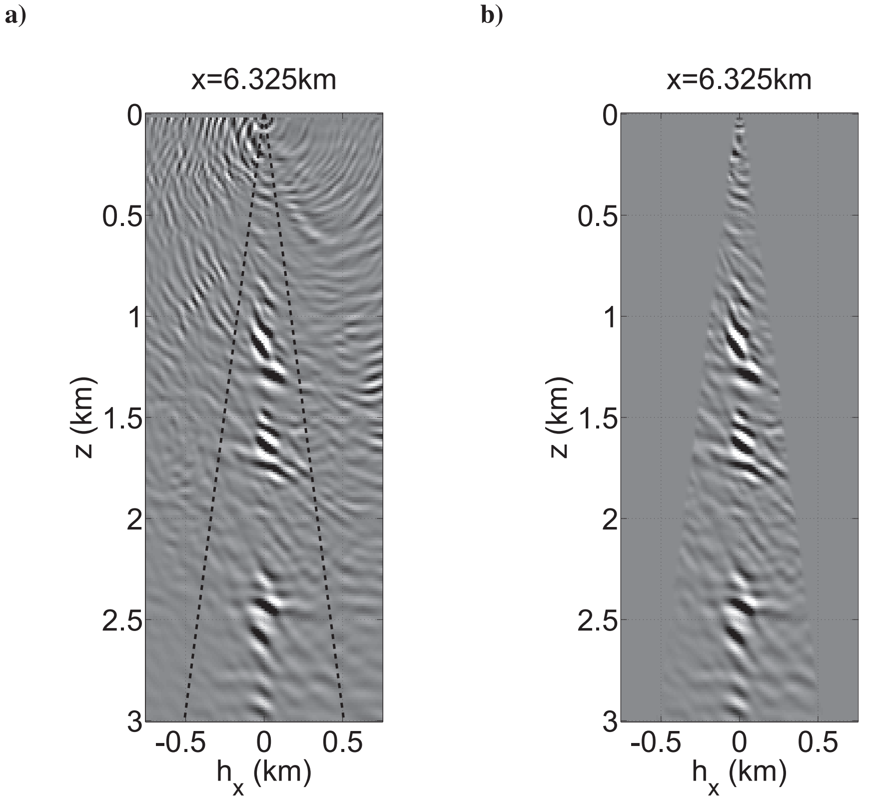


Figure 4.6: CIG at position 6.325 km a) before mute, b) after mute to remove aliased events. The black dotted lines in a) mark the position of the picked mute.

CIGs will, in principle, reconstruct the aliased events in the data. However, we want to reconstruct the data without aliasing. To achieve that we mute the aliased energy from the CIGs (Figure 4.6b).

After mute, we demigrate the muted CIGs, acquiring the data at the original receiver geometry. A comparison between the original shot gather, the decimated shot gather, and the reconstructed shot gather is shown in Figure 4.7.

Results show that demigration successfully reconstructs the kinematics of the original shot gather. Due to the smaller fold of the stack of the image constructed with the decimated data, the amplitudes of the reconstructed data will tend to converge to different values from those of the original data (about 8 times smaller in this case). Also, this procedure depends on having a migration velocity model that will focus the energy within the non-aliased part of the CIGs, that is, inside the black dotted lines indicated in Figure 4.6a. Otherwise, vital parts of the data may be muted together with the aliased energy. Despite these limitations, this type of reconstruction can find useful application in interpolating

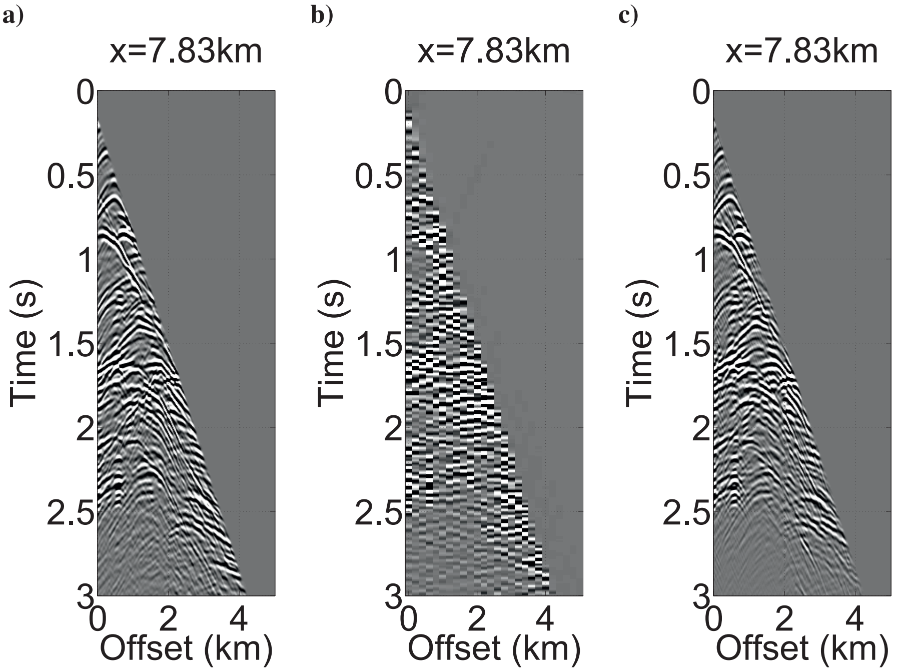


Figure 4.7: Shot gathers a) original, b) decimated, c) reconstructed shot gather after 5 iterations of demigration.

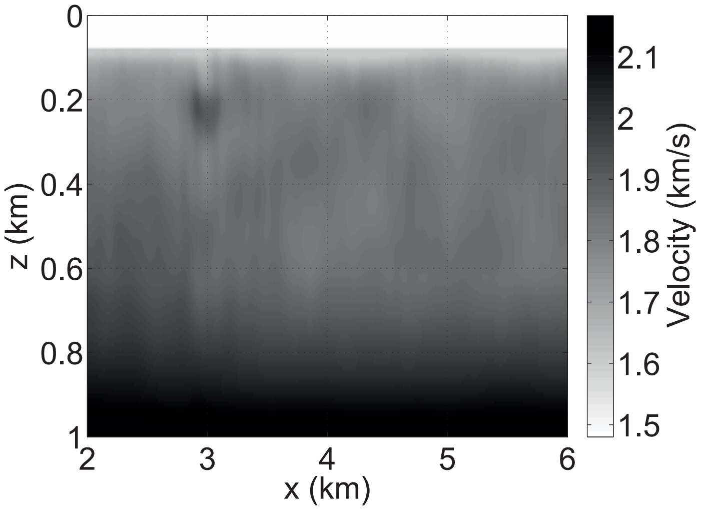


Figure 4.8: Migration velocity model for North Sea field data set.

irregularly sampled aliased data, where most classical methods fail (Zwartjes and Sacchi, 2007).

It is also possible to use this procedure to interpolate across shot gathers. This can be achieved in exactly the same way as above, but by using reciprocity (Ikelle and Amundsen, 2005), and demigrating common receiver gathers, instead of common shot gathers.

4.3.3 Example 3

We now present an application of the demigration method to free surface multiple attenuation.

The migration velocity model is shown in Figure 4.8. The migrated stacked image and CIGs are shown in Figure 4.9. The data used is the same as in the first example. To remove the multiples we explore a particular characteristic of the behavior of multiples in images migrated using the spatial lag extended cross-correlation imaging condition. In these CIGs, free surface multiples and primaries can be separated by noting that events requiring faster and slower velocities to focus are shifted in opposite directions relative to the zero lag. This behavior has been first pointed out by Mulder and van Leeuwen (2008), and later explored by Weibull and Arntsen (2013) to attenuate multiples before automatic velocity analysis. The demultiple procedure consists in muting the multiple events in the CIGs and reconstructing the data using the demigration method presented in this paper.

Figure 4.10 shows one particular CIG at position 4 km before and after muting the free surface multiple events. The black dotted line mark the position of the zero horizontal lag.

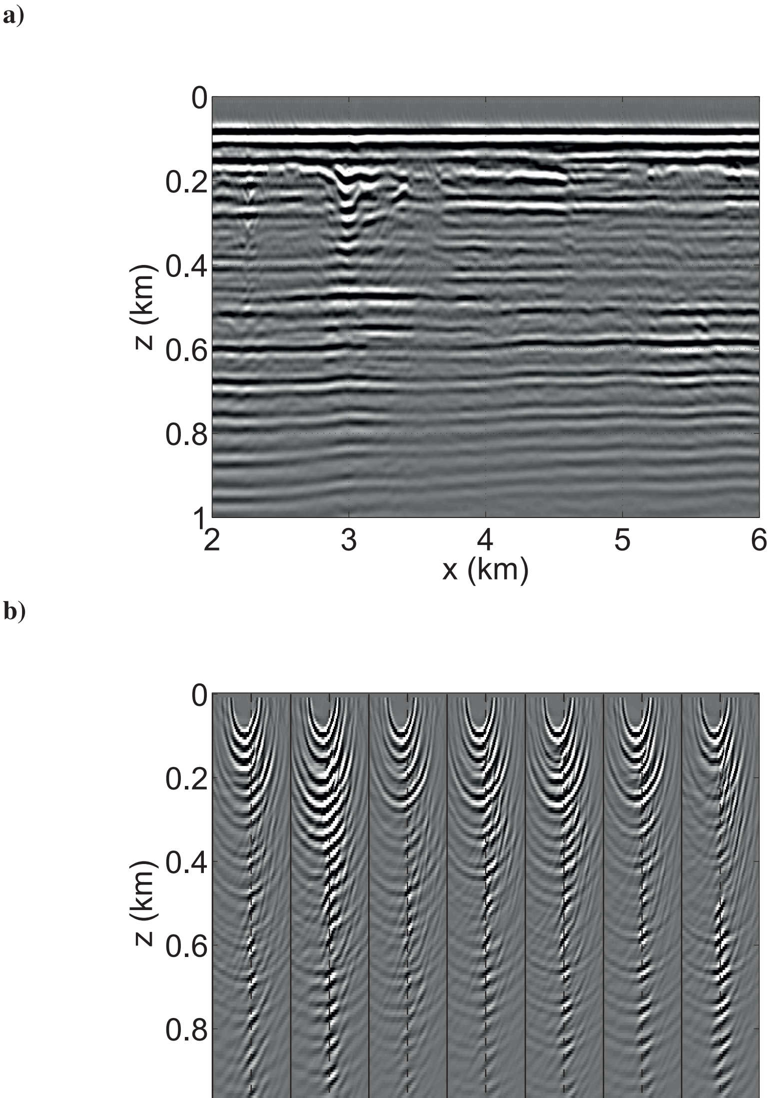


Figure 4.9: North Sea field data migrated a) zero lag image, and b) CIGs at selected spatial positions.

Note that the events to right of the black dotted lines represent events that require lower velocity to focus. These events are interpreted as multiples, and are therefore muted. In addition to muting the free surface multiple events, we also muted one particular event ocurring at a depth of about 800 m (Figure 4.10). The purpose is to show the ability of the method to remove events from the data by muting them in the extended image domain.

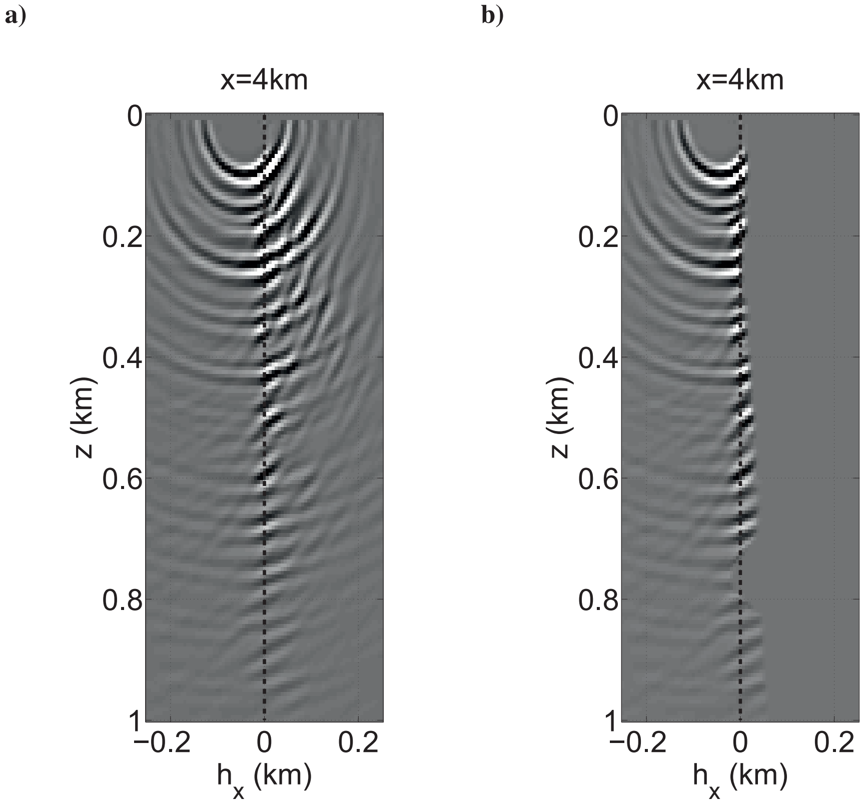


Figure 4.10: North Sea field data CIG at position 4 km a) original, b) after mute to remove free surface multiples.

We run the demigration for 50 iterations, using the original multiple-rich shot gathers as a starting point for the inversion. Figure 4.11a shows the original shot gather which was used to generate the CIGs, and also as a starting point for the inversion. Figure 4.11b shows the result of demigration on a particular shot gather at position 5.5 km. Finally, Figure 4.11c shows the difference between the original and demigrated shot gathers. As can be clearly seen, the events that have been muted in the CIGs, are attenuated in the demigrated shot gather. This includes the free surface multiple events, typically characterized by steep moveouts, and one particular primary event.

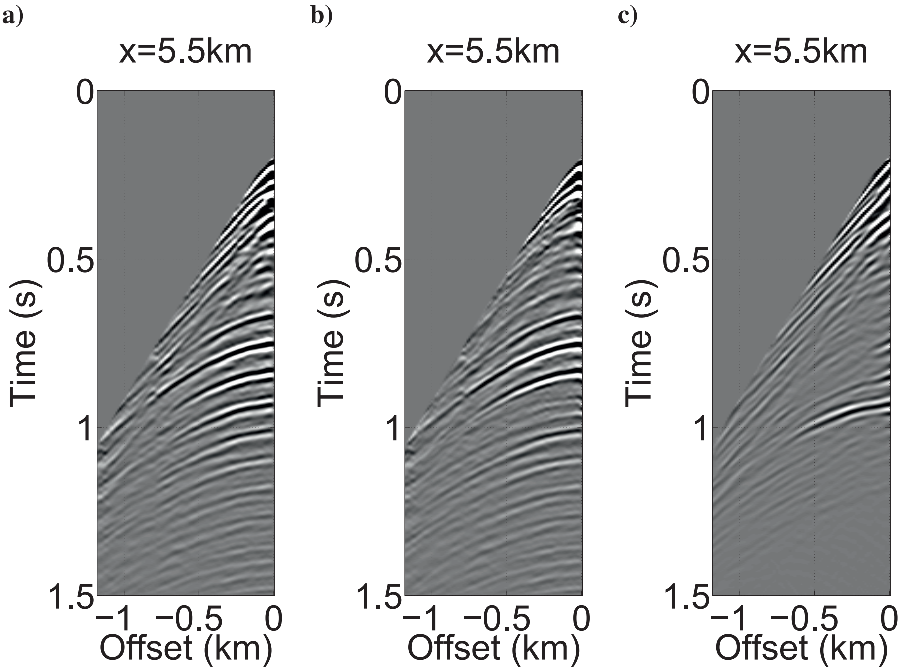


Figure 4.11: North Sea field data shot gathers at position 5.5 km a) original, b) after 50 iterations of demigration of the muted CIGs, c) difference between shot gathers a) and b).

4.4 Conclusion

We presented a method to reconstruct seismic reflection data from stacked CIGs constructed through RTM with an extended imaging condition. The method is based on least squares inversion, and is solved iteratively using a steepest descent approach. The numerical examples show that the extended imaging condition allows kinematic reconstruction of the prestack seismic reflection data after only one iteration, although many iterations are required to recover the correct amplitudes. The presented method has many interesting applications, such as image-based demultiple and data interpolation. And becuase the method is based on RTM, it can be applied to process data acquired over complex geological media.

Chapter 5

Concluding remarks

In this doctoral work I took a step in the direction of improving seismic imaging over complex geology. Seismic imaging over complex geological media is a challenging process. The strong and sharp contrasts, as well as the elastic nature of the sediments within the earth need to be honored by both the migration, and by the velocity estimation methods. With the purpose of imaging complex media, I developed and implemented three automatic parameter estimation methods based on non-linear least squares inversion in the image-domain. The main accomplishments and limitations of these methods are summarized below.

In Chapter 2, I extend WEMVA to deal with complex geological media by incorporating reverse-time migration and robust demultiple of surface seismic data. I introduce a vertical derivative operator to act over the reverse-time migrated image. This improves the convergence properties of the velocity analysis in areas of strong and sharp contrasts. The method is fully automatic and requires no picking. However, due to the very high computational cost, the implementation is restricted to 2D. 3D implementation is left as future work. The synthetic results show that the implemented WEMVA can recover accurate background subsurface models with large perturbations using both depth-oriented semblance and differential semblance as error measures. The method is limited to pre-critical single-scattering reflection data. This means that, prior to velocity analysis, post-critical reflections and refractions need to be muted, and the multiples must be attenuated. In the field data example, we introduce an image-based demultiple procedure and show how it helps to improve the results of WEMVA under the presence of free surface multiples. The method can be used to create complex initial models to be further refined by data domain methods such as full-waveform inversion.

The methodology of Chapter 2 was extended to 2D TTI elastic anisotropic media. To image anisotropic media, I use elastic reverse-time migration. The imaging can be described in two steps. First I use an elastic wave equation to reconstruct the source and receiver particle displacement wavefields. Then I use simple wave mode separation to obtain the scalar wavefields needed for the cross-correlation imaging condition. The method is developed for velocity analysis from PP CIGs constructed from surface seismic data.

Extension to deal with PS CIGs constructed from multicomponent data is left as future work. Numerical results on 2D synthetic and field data show that the method converges to models that improve the focusing of the depth migrated images. However, the assumption of anisotropy in the subsurface model strongly increases the ambiguity of the results. In this case, in order to converge to geologically meaningful models it is necessary to add additional information. The computational cost of method is very high, and future work is necessary to implement the method in 3D.

I also present a method to demigrate seismic reflection data using reverse-time migration with an extended imaging condition. The method, which is presented in Chapter 4, is related to the methods in Chapter 2 and 3, because the data is obtained by non-linear least squares inversion of an objective function defined in the image domain. The extended image condition produces CIGs that preserve amplitude and phase information of both primaries and multiple reflections. This means that they can be used to reconstruct the data with high fidelity, even in the presence of an inaccurate velocity model. The relationship between the data and the wavefields is linear, since the data is the source in the back propagation of the receiver wavefield. Thus the problem can be formulated as a linear least squares problem. However, due to the large size of the system, I cast the problem as a non-linear least squares problem and solve it using a steepest descent method. At each iteration the data is updated by adding to it the result of a Born modeling. The phase of the data is recovered after 1 iteration of the method, whereas the amplitudes require many iterations. Applications of the method in 2D synthetic models of complex geological media are presented. The results show the capability of the method to process complex seismic data. A 2D field data example is also included. Some applications, such as data reconstruction, might, to some extent, require an accurate estimate of the velocities. Again the cost of the method currently prohibits its use in 3D applications.

Bibliography

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

———, 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Alkhalifah, T., S. Fomel, and B. Biondi, 2001, The space-time domain: theory and modelling for anisotropic media.: Geophysical Journal International, **144**, 105–113.

Alkhalifah, T., and K. Larner, 1994, Migration error in transversely isotropic media: Geophysics, **59**, 1405–1418.

Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550–1566.

Arntsen, B., C. Gerea, and T. Rosten, 2009, Imaging salt bodies using explicit migration operators offshore norway: Geophysics, **74**, S25–S32.

Audebert, F., V. Dirks, and A. Pettenati, 2006, TTI anisotropic depth migration: What tilt estimate should we use?: 76th Annual International Meeting, SEG, Expanded Abstracts, 2382–2386.

Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2010, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, D27–D36.

Ball, G., 1995, Estimation of anisotropy and anisotropic 3-D prestack depth migration, offshore Zaire: Geophysics, **60**, 1495–1513.

Banik, N., 1984, Velocity anisotropy of shales and depth estimation in the North Sea basin: Geophysics, **49**, 1411–1419.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Behera, L., and I. Tsvankin, 2009, Migration velocity analysis for tilted transversely isotropic media: Geophysical Prospecting, **57**, 13–26.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: Geophysics, **76**, R109–R124.

Berkhout, A. J., 1980, Seismic migration : Imaging of acoustic energy by wavefield extrapolation: Elsevier.

Biondi, B., and P. Sava, 1999, Wave-equation migration velocity analysis: 69th Annual International Meeting, SEG, Expanded Abstracts, Society of Exploration Geophysicists, 1723–1726.

Biondi, B., and G. Shan, 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, SEG, 1284–1287.

Biondi, B., and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

Brandsberg-Dahl, S., B. Ursin, and M. De Hoop, 2003, Seismic velocity analysis in the scattering-angle/azimuth domain: Geophysical Prospecting, **51**, 295–314.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, A limited memory algorithm for bound contrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208.

Byun, B. S., 1984, Seismic parameters for transversely isotropic media: Geophysics, **49**, 1908–1914.

Chauris, H., and M. Benjemaa, 2010, Seismic wave-equation demigration/migration: Geophysics, **75**, S111–S119.

Chauris, H., and M. Noble, 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets: Geophysical Journal International, **144**, 14–26.

Chavent, G., 2009, Non-linear least squares for inverse problems: Theoretical foundations and step-by-step guide for applications: Springer.

Chavent, G., and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Chavent, G., and P. Lemonnier, 1974, Identification de la non-linéarité d'une équation parabolique quasilinéaire: Applied Mathematics & Optimization, **1**, 121–162.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

deBruin, C. G. M., 1990, Angle dependent reflectivity by means of prestack migration: Geophysics, **55**, 1223–1234.

Dellinger, J., and J. Etgen, 1990, Wave-field separation in two-dimensional anisotropic media: Geophysics, **55**, 914–919.

Diaz, E., and A. Guitton, 2011, Fast full waveform inversion with random shot decimation: 81st Annual International Meeting, SEG, Expanded Abstracts, SEG, 2804–2808.

Dierckx, P., 1993, Curve and surface fitting with splines: Oxford University Press.

Douma, H., D. Yingst, I. Vasconcelos, and J. Tromp, 2010, On the connection between artifact filtering in reverse-time migration and adjoint tomography: Geophysics, **75**, S219–S223.

Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, WCA5–WCA17.

Fei, W., and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: 80th Annual International Meeting, SEG, Expanded Abstracts, SEG, 4071–4076.

Gao, F., and W. Symes, 2009, Differential semblance velocity analysis by reverse time migration: Image gathers and theory: 79th Annual International Meeting, SEG, Expanded Abstracts, Society of Exploration Geophysicists, 2317–2321.

Gray, S. H., J. Etgen, J. Dellinger, and D. Whitmore, 2001, Seismic migration problems and solutions: Geophysics, **66**, 1640.

Grechka, V., A. Pech, and I. Tsvankin, 2002, P-wave stacking-velocity tomography for VTI media: Geophysical Prospecting, **50**, 151–168.

Guitton, A., B. Kaelin, and B. Biondi, 2007, Least-squares attenuation of reverse-time-migration artifacts: Geophysics, **72**, S19–S23.

Ikelle, L. T., and L. Amundsen, 2005, Introduction to petroleum seismology: Society of exploration geophysicists.

Isaac, J., and D. Lawton, 1999, Image mispositioning due to dipping TI media: A physical seismic modeling study: Geophysics, **64**, 1230–1238.

Jaramillo, H., and N. Bleistein, 1999, The link of kirchhoff migration and demigration to kirchhoff and born modeling: Geophysics, **64**, 1793–1805.

Jones, I. F., M. L. Bridson, and N. Bernitsas, 2003, Anisotropic ambiguities in TI media: First Break, **25**, 15–21.

Larner, K. L., 1993, Dip-moveout error in transversely isotropic media with linear velocity variation in depth: Geophysics, **58**, 1442–1453.

Levin, F., 1979, Seismic velocities in transversely isotropic media: Geophysics, **44**, 918–936.

Li, J., and W. W. Symes, 2007, Interval velocity estimation via NMO-based differential semblance: Geophysics, **72**, U75–U88.

Li, Y., and B. Biondi, 2011, Migration velocity analysis for anisotropic models: 81th Annual International Meeting, SEG, Expanded Abstracts, Society of Exploration Geophysicists, 201–206.

Li, Y. E., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: 82nd Annual International Meeting, SEG, Expanded Abstracts, Society of Exploration Geophysicists, 1–6.

Lions, J. L., and E. Magenes, 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag, Berlin.

Lisitsa, V., and D. Vishnevskiy, 2010, Lebedev scheme for the numerical simulation of wave propagation in 3D anisotropic elasticity: Geophysical Prospecting, **58**, 619–635.

Loewenthal, D., L. Lu, R. Roberson, and J. Sherwood, 1976, The wave equation applied to migration: Geophysical Prospecting, **24**, 380–399.

Miranda, F., 2006, Non-asymptotic seismic data reconstruction: 76th Annual International Meeting, SEG, Expanded Abstracts, 1–5.

Mora, P., 1987, Nonlinear twoâĂŘdimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

Mulder, W. A., 2008, Automatic velocity analysis with the two-way wave equation: Presented at the 70th Conference and Exhibition, EAGE, Expanded Abstracts, European Association of Geoscientists and Engineers.

Mulder, W. A., and A. P. E. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191.

Mulder, W. A., and T. van Leeuwen, 2008, Automatic migration velocity analysis and multiples: 78th Annual International Meeting, SEG, Expanded Abstracts, SEG, 3128–3132.

Nocedal, J., and S. J. Wright, 2000, Numerical optimization: Springer.

Parkes, G., and L. Hatton, 1986, The marine seismic source: Kluwer.

Plessix, R., and H. Rynja, 2010, VTI full waveform inversion: a parametrization study with a narrow azimuth streamer data example: 80th Annual International Meeting, SEG, Expanded Abstracts, SEG, 962–966.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Qin, Z., M. Lu, X. Zheng, Y. Yao, C. Zhang, and J. Song, 2009, The implementation of an improved NPML absorbing boundary condition in elastic wave modeling: Applied Geophysics, **6**, 113–121.

Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Santos, L., J. Schleicher, M. Tygel, and P. Hubral, 2000, Seismic modeling by demigration: Geophysics, **65**, 1281–1289.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Sava, P., and I. Vasconcelos, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

Sava, P., and L. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: Geophysics, **73**, 145–159.

Sava, P. C., and A. Guitton, 2005, Multiple attenuation in the image space: Geophysics, **70**, V10–V20.

Sayers, C. M., 1994, The elastic anisotrophy of shales: Journal of Geophysical Research: Solid Earth, **99**, 767–774.

Schneider, W., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76.

Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, 49–59.

Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2132–2135.

Song, X., and T. Alkhalifah, 2012, A linearized dispersion relation for orthorhombic pseudo-acoustic modeling: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1–6.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, SM213–221.

Symes, W. W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **5**, 654–663.

Symes, W. W., and M. Kern, 1994, Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples: Geophysical Prospecting, **42**, 1365–2478.

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

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.

Tikhonov, A. N., and V. Y. Arsenin, 1977, Solution of ill-posed problems: W. H. Winston and Sons.

Toldi, J. L., 1989, Velocity analysis without picking: Geophysics, **54**, 191–199.

Tsvankin, I., and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304.

———, 1995, Inversion of reflection traveltimes for transverse isotropy: Geophysics, **60**, 1095–1107.

van Leeuwen, T., and W. A. Mulder, 2008, Velocity analysis with multiples – NMO modeling for layered velocity structures: 78th Annual International Meeting, SEG, Expanded Abstracts, 1925–1929.

Versteeg, R. J., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, **58**, 873–882.

Vestrum, R., D. Lawton, and R. Schmid, 1999, Imaging structures below dipping TI media: Geophysics, **64**, 1239–1246.

Virieux, J., 1986, P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method: Geophysics, **51**, 889–901.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Weibull, W., B. Arntsen, M. Houbiers, and J. Mispel, 2012, Automatic anisotropic migration velocity analysis for reverse-time migration: 82nd Annual International Meeting, SEG, Expanded Abstracts, Society of Exploration Geophysicists, 1–5.

Weibull, W. W., and B. Arntsen, 2011, Reverse time migration velocity analsysis: Presented at the 73rd Conference and Exhibition, EAGE, Expanded Abstracts, European Association of Geoscientists and Engineers.

———, 2013, Automatic migration velocity analysis using reverse-time migration: Geophysics, **55**, 1070–1088. (accepted).

Whitmore, N. D., and S. Crawley, 2012, Applications of rtm inverse scattering imaging conditions: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1–6.

Winterstein, D. F., 1990, Velocity anisotropy terminology for geophysicists: Geophysics, **55**, 1070–1088.

Yan, J., and P. Sava, 2011, Improving the efficiency of elastic wave-mode separation for heterogeneous tilted transverse isotropic media: Geophysics, **76**, T65–T78.

Yan, L., L. Lines, and D. Lawton, 2004, Influence of seismic anisotropy on prestack depth migration: The Leading Edge, **23**, 30–36.

Zhang, L., J. W. Rector, and G. M. Hoversten, 2003, An acoustic wave equation for modeling in tilted TI media: 73rd Annual International Meeting, SEG, Expanded Abstracts, 153–156.

Zhang, Y., and L. Duan, 2012, Predicting multiples using a reverse time demigration: SEG Technical Program Expanded Abstracts 2012, 1–5.

Zhou, H., Q. Liao, and F. Ortigosa, 2009, Migration velocity inversion with semblance analysis (mVisa): Presented at the 71st Conference and Exhibition, EAGE, Expanded Abstracts, European Association of Geoscientists and Engineers.

Zwartjes, P., and M. Sacchi, 2007, Fourier reconstruction of nonuniformly sampled, aliased seismic data: Geophysics, **72**, V21–V32.