Time-lapse full waveform inversion in strongly scattering media

Aleksandr Tulapin

Thesis for the degree

Master of Science in Petroleum Geophysics

Department of Earth Science

University of Bergen

June 3, 2019

"There are few persons who have not, at some period of their lives, amused themselves in retracing the steps by which particular conclusions of their own minds have been attained. The occupation is often full of interest and he who attempts it for the first time is astonished by the apparently illimitable distance and incoherence between the starting-point and the goal".

The Murders in the Rue Morgue
— **Edgar Allan Poe**

In memory of my grandmother

Abstract

In petroleum seismology, full waveform inversion (FWI) is known as a powerful method taking into account all available in seismic data information with which one can determine a high-resolution subsurface image. In time-lapse seismic, the use of FWI is even more valuable since it makes it possible to monitor changes in model parameters. These changes are often associated with changes in physical conditions in a hydrocarbon reservoir. In high-contrast media, such as salt-affected regions for example, where petroleum traps are often located beneath large salt bodies, the application of FWI methods encounters two main problems: the absence of an accurate initial model to start an inversion workflow and, the absence of ultra-low frequencies in real seismic data. These issues often cause FWI algorithms to be ineffective. This thesis presents the Distorted Born Iterative T-matrix (DBIT) nonlinear inversion method integrated with smoothing and self-adaptive techniques to deal with these problems in the context of strongly scattering media models. The proposed approach was applied to synthetic data, which were generated using the T-matrix integral equation method.

The T-matrix integral equation method is a modelling technique used to model synthetic data. Through a systematic comparison of this method with finite-difference modelling, it was found that this technique can be considered to be exact. This makes it attractive to DBIT, i.e. T-matrix modelling based inversion.

A lack of ultra-low frequencies (1-4 Hz), when experimenting with the modified high-contrast EAGE/SEG salt models, caused the DBIT inversion algorithm to diverge. However, the application of smoothing (multi-scale regularization) and self-adaptive (a method for choosing the regularization parameter) techniques, solved this problem and the DBIT-based FWI method successfully recovered the analyzed models. The time-lapse application showed that for noisy data, only the high-contrast time-lapse velocity anomaly (velocity difference of 600 m/s) could be restored. For the low-contrast velocity difference (velocity difference of 200 m/s), the time-lapse anomaly on the recovered model failed to be resolved. A comparison of two different time-lapse strategies (parallel difference and sequential) demonstrated that the sequential approach worked faster and more accurate. The DBIT inversion method, in combination with the smoothing and self-adaptive techniques, can be considered as an effective tool for strongly scattering media model recoveries.

Acknowledgments

I would firstly like to express my gratitude to my supervisor, professor Morten Jakobsen for the useful comments, remarks and engagement through the learning process of this thesis. Spending hours on discussions, he opened for me the worlds of seismic modelling and inversion. I also want to thank my co-supervisor, associate professor Einar Iversen.

My research would have been impossible without MATLAB skills for which I am profoundly grateful to associate professor Henk Keers.

Special thanks go to a PhD candidate, Xingguo Huang for useful discussions.

I would like to thank all staff of the faculty of Mathematics and Natural Sciences of the University of Bergen and the Norwegian Government in general for giving me a chance to get a master's degree in petroleum geophysics in Norway.

I wish to acknowledge professor Shukhrat Radjabov from the National University of Uzbekistan and chief geophysicist Alexander Dlugach from JSC Arctic Marine Engineer Geological Expeditions for giving me feedbacks on this thesis.

I thank my friends Eldar Akhmetov and Alexey Baboshko for reading this thesis and giving their opinions about it.

Finally, I must express my very profound gratitude to my mother for providing me with unfailing support and continuous encouragement through the process of researching and writing this thesis.

Contents

List of Figures

List of Tables

List of abbreviations

1D One-dimensional

2D Two-dimensional

3D Three-dimensional

4D Four-dimensional

AVO Angle versus offset

DBA Distorted Born approximation

DBI Distorted Born inversion

DBIT Distorted Born iterative T-matrix

FDTD Finite-difference time domain

FWI Full waveform inversion

LS Lippmann-Schwinger

RMS Root mean square

SNR Signal to noise ratio

Chapter 1

Introduction

The only source of knowledge is experience.

—Albert Einstein

Currently, despite the increased interest in alternative energy sources, hydrocarbons still remain the most preferable because their extraction is relatively cheap (Ovcharenko et al., 2018). The development of new techniques and technologies in petroleum exploration and production has significantly enhanced oil and gas recoverability in conventional geological reservoirs and opened access to hydrocarbon deposits hidden in more complex geological settings (Leveille et al., 2011). High-contrast geological models presented by salt bodies is one of the examples of such complex environments (Farooqui et al., 2009; Jackson and Hudec, 2017). There are also plenty of cases where high-contrast structures consist of massive volcanic formations (Farooqui et al., 2009). In both geological situations, hydrocarbons reservoirs are typically located beneath these high-contrast objects (Farooqui et al., 2009; Jackson and Hudec, 2017).

Reservoir monitoring (as a part of enhanced oil recovery program) within hydrocarbon production sites, has become a trivial routine (Speight, 2015). The need for reservoir monitoring arises from two main reasons: economic benefit and environmental impact (Bjørlykke and Jahren, 2015). In the first case, it is economically feasible to invest in reservoir monitoring since the hydrocarbon production cost reduces through increased production rate and increased recovery (Johnston, 1997). In the second case, reduction of drilling holes and associated activities leads to reduced impact on the surrounding environment (Patin, 2001).

From a geophysical perspective, there are several methods to monitor reservoirs under production. These methods typically include seismic, electromagnetic and gravimetric surveys. Among them, seismic method is traditionally considered to be more accurate and reliable geophysical monitoring approach (Landrø, 2015). Seismic monitoring method is known as 4D or time-lapse seismics (Landrø, 2015).

In 4D seismics, a seismic survey is carried out over some given area or line repeatedly (Nguyen et al., 2015). In this case, the fourth dimension denotes calendar time (Yilmaz, 2001). In reservoir monitoring, calendar time is associated with different reservoir production stages (Landrø, 2015). The first survey (baseline survey) relates to reservoir imaging before production. The second and latest surveys (monitor survey) correspond to the seismic investigations conducted after production started. The difference between seismic wavefields obtained at the mentioned seismic monitoring surveys is known as time-lapse seismic data (Nguyen et al., 2015). Time-lapse seismic data analysis mainly concerns about wavefield changes in the reservoir zone. These changes are commonly linked to reservoir alterations (its physical conditions) induced by hydrocarbons depletion (Landrø, 2015). The analysis is typically accomplished in conjunction with available reservoir rock information obtained from well-log data and core samples (porosity, permeability) to model fluid movements in the reservoir (flow simulation) and to estimate reservoir production potential (Johnston, 1997). An example of time-lapse seismic data obtained at Sleipner CO_2 storage (a different kind of reservoir, but obeys the same seismic monitoring rules) is given in figure 1.1.

In order to be valuable for reservoir characterization, 4D seismic data must be appropriately acquired and processed (Landrø, 2015). Seismic differences corresponding to seismic data collected at different times are not only sensitive to changes in reservoir properties but also to differences in acquisition and processing technologies (Johnston, 1997). There exist several processing techniques that minimize inconsistencies between seismic data collected at different times (Yilmaz, 2001; Nguyen et al., 2015). However, this problem cannot still be avoided completely. For seismic data acquired within high-contrast media models, all the issues mentioned become even more important. It is common practice to consider subsurface imaging in the model (velocities) rather than seismic (amplitudes and travel times) domains. Esser et al. (2016), explaining a typical workflow to image subsalt structures (high-contrast models), put a particular emphasis on its complexity. The authors conclude that successful high-contrast model recovery relies heavily on the skills and experience of practitioners involved in the model restoration process. It is therefore important to have good and precise processing method. A good

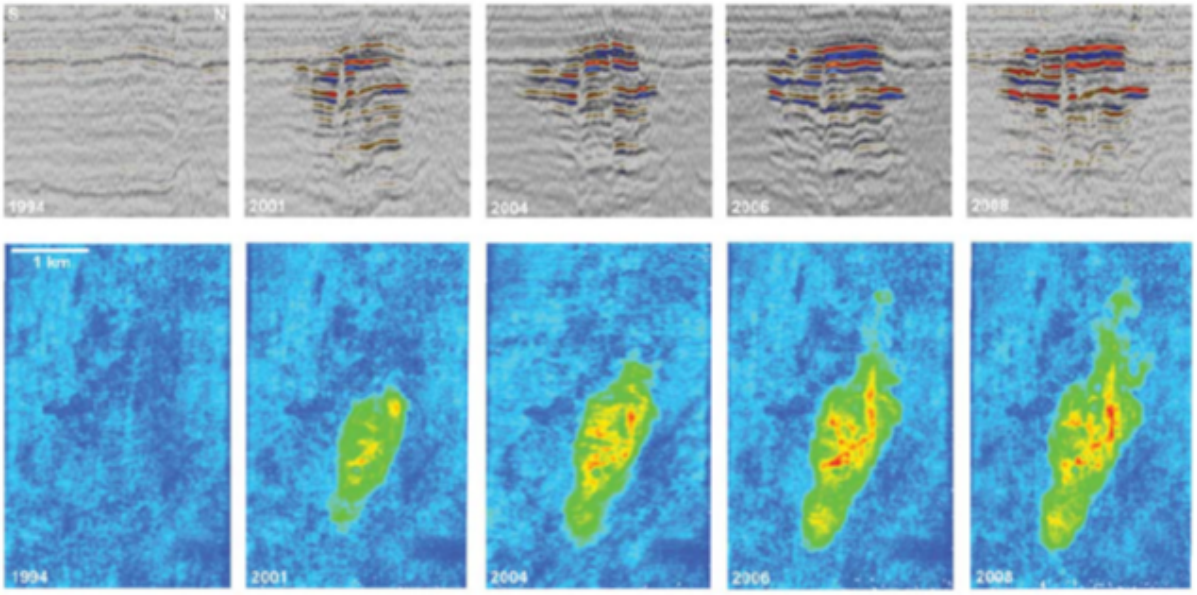

Figure 1.1: Time-lapse seismic images of the Sleipner CO_2 storage. Top panel- inline perspective on the 3D time-lapse slices; bottom panel- plan perspective on the 3D time-lapse slices (Chadwick et al., 2010).

candidate for this is Full Waveform Inversion (FWI).

FWI is a method that aims to reconstruct model parameters by considering all information in the recorded seismic wavefield (Queißer and Singh, 2013). An omportant feature that this method does not require data prepocessing procedures. Or, put another way, this model imaging approach excludes the need for manual interventions (Esser et al., 2016). However, application of FWI to high-contrast media cases such as salt trap models, where the main goal is to image the subsalt structure may be challenging. There are several reasons for this. Firstly, there is high contrast between the salt body and surrounded geological medium, which causes the subsalt structure to be poorly illuminated (Ovcharenko et al., 2018). Secondly, there is the absence of an accurate initial model to start the inversion workflow (Esser et al., 2016). Finally, real seismic data are typically lacking ultra-low frequencies. This can cause the inversion algorithm to get stuck at one of objective function local minima (Alkhalifah and Choi, 2012).

Thesis goal and objectives

The goal of this thesis is to study seismic time-lapse full waveform inversion in strongly scattering media by using the distorted Born iterative T-matrix inversion approach and different time-lapse inversion strategies.

The cornerstone of this research work is the T-matrix concept (Taylor, 1972) on which both the T-matrix integral equation modelling (Jakobsen, 2012) and distorted Born iterative T-matrix inversion (Jakobsen and Ursin, 2015) techniques are based on. Both techniques operate in the frequency domain. A choice of the T-matrix integral equation method as a modelling method in this thesis is explained by its modelling accuracy and ability to generate data for selected frequencies. For model recoveries based on the frequency domain inversion approaches only few frequencies are required (Sirgue and Pratt, 2004). This feature is very important if many sources in inversion experiments are used (Jakobsen and Ursin, 2015). On the other hand, the use of the same solver (T-matrix operator) when computing forward and inversion problems, may result in biased results (Colton and Rainer, 1998). Therefore, the first objective of this thesis is to systematically examine the T-matrix integral equation modelling approach and to show that this method provides results comparable with an exact solution.

The DBIT inversion method solves the nonlinear inverse problem by solving a series of related linear inverse problems (Jakobsen and Ursin, 2015). This linearization relies on the modelling principle underlying the Distorted Born Approximation (DBA) method (Jakobsen and Wu, 2018). To make sure that this approximation is reliable, it is necessary to demonstrate the modelling accuracy of the DBA modelling method. This is the second objective of this thesis.

The conventional DBIT-based inversion algorithm fails to recover models that include strongly scattering elements when data lack ultra-low frequencies and an accurate initial model. The third objective is to apply the DBIT inversion algorithm together with smoothing and self-adaptive techniques to deal with these issues.

Results obtained by different time-lapse model recovery techniques differ in computation cost and accuracy. The fourth objective is to show the effectiveness of two time-lapse inversion strategies when applied to strongly scattering media models.

For computational simplicity, in this thesis, both the modelling and the inversion problems are considered in the acoustic approximation. All inversion experiments are done using 2D synthetic data. Since the DBIT inversion is based on scattering theory, in this thesis, the term "high-contrast media" used in pair with the analyzed FWI method is considered to be a synonym to the term "strongly scattering media".

Thesis overview

This thesis contains 4 chapters. Each chapter is self-contained and focuses on a specific topic. Chapter 1 gives an overview of seismic reservoir monitoring and explains difficulties associated with the conventional processing when used for high-contrast media model reconstructions. It also introduces the concept of FWI and specifies some issues that could be encountered when one applies FWI to high-contrast models.

Chapter 2 focuses on the theory and practice of seismic modelling. First, it gives a brief overview of different seismic modelling methods. Then, based on scattering theory, it provides a derivation of the Lippmann-Schwinger integral equation and shows how it can be reduced to first order approximation (the Born approximation), which is applicable for modelling low-contrast media models. After discussing the discretization methodology of the Lippmann-Schwinger equation, it introduces the T-matrix concept and provides a derivation of the distorted Born approximation method. The chapter finishes with several numerical tests which demonstrate the consistency of the Born approximation, T-matrix modelling and distorted Born approximation approaches discussed in the theoretical part.

Chapter 3 discusses FWI theory and some applications of FWI. Firstly, some central aspects of inversion theory, including explanations of ill-posed and well-posed problems, regularization, linear and nonlinear inversion approaches and some elements of frequency domain inversion techniques are explained. After discussing Born inversion, some simple numerical experiments exemplifying the theory considered earlier are given. Further, it introduces the nonlinear Distorted Born Iterative T-matrix (DBIT) inversion method. It is shown that this method can be combined with smoothing and self-adaptive techniques to deal with strongly scattering inverse problems. Finally, a comparison of two different time-lapse strategies and their efficiency when applied to high-contrast models are demonstrated.

Chapter 4 provides conclusions of both the modelling and inversion results obtained previously. It also suggests some possible research directions in which the current study could be extended.

Appendix A presents a derivation of the finite-difference time domain modelling method and provides important points for its practical implementation.

Chapter 2

Seismic waveform modelling

It is the theory that decides what can be observed.

—Albert Einstein

2.1 Introduction

Seismic modelling is a technique that makes it possible to compute the propagation of seismic waves in the Earth's subsurface (Carcione et al., 2002). Historically, seismic modelling methods have been developed by studying 1D models. Later, with the computer power increase, 2D and 3D models have also been included in the consideration (Sayers and Chopra, 2009). In practice, 3D seismic modelling is a costly procedure, and even in the acoustic approximation (when the density is constant) the computations are enormous. At the same time, the computational cost can be significantly reduced if the modelling aim is clearly defined. In some cases, for example when the amplitude information is not crucial, the use of approximate seismic modelling methods (in order to estimate the optimal survey design, for instance) may be more desirable since they are faster and provide results comparable to exact modelling solutions (Carcione et al., 2002).

The theory behind seismic modelling methods can be found in various sources (Fichtner, 2011; Ikelle and Amundsen, 2005; Jakobsen, 2012). Carcione et al. (2002) categorize seismic modelling approaches as follows: direct methods, integral-equation methods, and ray-tracing methods.

Direct methods are traditionally considered to be exact methods (Carcione et al., 2002; Ikelle and Amundsen, 2005). In this case, seismic modelling involves performing two main steps: firstly, a physical model is approximated by its discretized version which contains a finite number of grid points; secondly, the seismic wavefield is computed using the wave equation (Ikelle and Amundsen, 2005). The main advantage of such methods is that there are no restrictions regarding the physical characteristics (composition contrast, for example) of the model employed and the modelling accuracy depends directly on the grid size. On the other hand, these methods are computationally expensive (Carcione et al., 2002). The most popular modelling techniques in this group are finite-difference, finite-element and pseudospectral methods (Schuster, 2017).

Integral-equation methods represent another group that is based on scattering theory (Lo and Inderwiesen, 1994). In this case, the total wavefield is represented by a superposition of the reference (background) and the perturbed (scattered) wavefields. The wavefield decomposition is very convenient since it make it possible to work with the included wavefields separately. On the other hand, model complexities in this type of methods (Carcione et al., 2002) are directly proportional to modelling errors, which have a tendency to accumulate from a scatterer to a scatterer. Therefore, these methods are more restrictive in application than the direct methods. However, they still can give accurate results when models are not very complex (Carcione et al., 2002). In recent years, the potential of the scattering methods has essentially increased. Jakobsen (2012), for example, presenting the T-matrix integral equation method, argues that this scattering theory-based modelling approach can provide modelling results comparable with the direct methods.

The final group represents ray-tracing (asymptotic) methods. These methods are based on the high-frequency (asymptotic) solution of the wave equation (Cerveny, 2001). On one hand, these methods are approximate, as they assume that the frequencies are high. On the other hand, they are the fastest methods which make them more useful, especially in 3D seismic modelling cases (Cerveny, 2001; Carcione et al., 2002).

In this thesis, four modelling approaches are considered. The first three are the scattering theory based methods, namely, the Born approximation, the T-matrix integral equation and the distorted Born approximation methods. The fourth approach is the finite-difference (exact) modelling technique.

2.2 Theory

This section introduces three seismic modelling techniques based on the Lippmann-Schwinger (LS) equation. They are the Born approximation, the T-matrix integral equation and the Distorted Born Approximation (DBA) methods.

2.2.1 General terms, equations and conditions

The wave equation, under the assumption that the density is constant, is defined as (Auld, 1973)

$$\left[\nabla^2 - \frac{1}{c^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}\right]\Psi(\mathbf{x},t) = -F_s(\mathbf{x},t), \tag{2.1}$$

where ∇^2 is the Laplace operator, $F_s(\mathbf{x},t)$ is the source function, and $\Psi(\mathbf{x},t)$ is the seismic wavefield at position $\mathbf{x}$ and some time t propagating in the medium with the velocity $c(\mathbf{x})$. The Fourier transform of equation (2.1) gives us the wave equation in the frequency domain (Helmholtz equation) written as (Jakobsen and Ursin, 2015)

$$L(\mathbf{x},\omega)\Psi(\mathbf{x},\omega) = -f_s(\mathbf{x},\omega), \tag{2.2}$$

with the wave operator $L(\mathbf{x},\omega)$ explicitly given by

$$L(\mathbf{x}) = \nabla^2 + k^2(\mathbf{x}), \tag{2.3}$$

and $\Psi(\mathbf{x},\omega)$ and $f_s(\mathbf{x},\omega)$ are the Fourier transforms of the wavefields $\Psi(\mathbf{x},t)$ and $F_s(\mathbf{x},t)$ respectively. The variable $k(\mathbf{x})$ depicts the wavenumber with the magnitude defined by

$$k(\mathbf{x}) = \frac{\omega}{c(\mathbf{x})}. \tag{2.4}$$

For the Helmholtz equation (2.2) to have a unique solution, one needs to consider the so-called Sommerfeld radiation conditions (Morse and Feshbach, 1953), which imply that the seismic energy propagates only in the outward direction. Mathematically, the Sommerfeld radiation conditions are given by

$$\frac{\partial\,\Psi}{\partial\,\mathbf{x}} \pm \frac{i\omega}{c(\mathbf{x})}, as\, x \to \pm\infty, \tag{2.5}$$

where wavefield Ψ is bounded for all $\mathbf{x}$ (Bleistein et al., 2001).

2.2.2 Greens functions and the Lippmann-Schwinger equation

A point-source solution to the Helmholtz equation

First, we consider the Helmholtz equation with a point-source. This means that the source function f_s on the right-hand side of equation (2.2) is a harmonic point pulse with the oscillating angular frequency ω. If we define δ as the Dirac-delta function and $\mathbf{x}'$ as the position of some arbitrary point within the perturbed medium, then

$$f_s(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}'), \tag{2.6}$$

In this case, there exists such a unique solution to equation (2.2), which is called the Green function $G(\mathbf{x}, \mathbf{x}')$ (Morse and Feshbach, 1953; Lo and Inderwiesen, 1994; Bleistein et al., 2001; Jakobsen and Ursin, 2015). The physical meaning of the Green function is that gives the wavefield in a point source. Substituting equation (2.6) into (2.2), we rewrite the Helmholtz equation in the following form:

$$L(\mathbf{x})G(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}'). \tag{2.7}$$

Following (Bleistein et al., 2001), we combine equations (2.2) and (2.7) and integrate the result over some volume Ω:

$$\int_\Omega \left[G(\mathbf{x}, \mathbf{x}')\nabla^2\Psi(\mathbf{x}) - \Psi(\mathbf{x})\nabla^2 G(\mathbf{x}, \mathbf{x}')\right] dV. \tag{2.8}$$

Using the divergence theorem, we rewrite the volume integral into a surface integral

$$\int_{\partial\Omega} \left[G(\hat{n}\cdot\nabla)\Psi - \Psi(\hat{n}\cdot\nabla)G\right] dS, \tag{2.9}$$

where $\partial\Omega$ is the bounding surface, $\hat{n}\cdot\nabla$ is the derivative normal to $\partial\Omega$, and n denotes the outward directed normal vector. Combining equations (2.7) and (2.9), we can solve for $\Psi(\mathbf{x}')$

$$\Psi(\mathbf{x}') = \int_\Omega G(\mathbf{x}, \mathbf{x}') f_s(\mathbf{x})\, dV - \int_{\partial\Omega} \left[G(\mathbf{x}, \mathbf{x}')\nabla\Psi(\mathbf{x}) - \Psi(\mathbf{x})\nabla G(\mathbf{x}, \mathbf{x}')\right] dS. \tag{2.10}$$

Since we consider only outgoing waves, it follows from the Sommerfeld conditions (equation (2.5)) that the surface integral in equation (2.10) is equal to zero. Applying the reciprocity principle (see Schuster (2017)), we end up with the point-source solution to the Helmholtz equation (Jakobsen and Ursin, 2015):

$$\Psi(\mathbf{x}) = \int G(\mathbf{x}, \mathbf{x}') f_s(\mathbf{x}')\, d\mathbf{x}'. \tag{2.11}$$

Green functions

The explicit forms of the Green functions for homogeneous unbounded media in one-, two, and, three- dimensions can be found analytically (see, for example Bleistein et al. (2001), Snieder (2004), Nowack (2010)). They are:

$$G^{1D}(\mathbf{x},\mathbf{x}') = \frac{i}{2k}e^{ik|\mathbf{x}-\mathbf{x}'|}, \tag{2.12}$$

$$G^{2D}(\mathbf{x},\mathbf{x}') = \frac{i}{4}H_0^{(1)}(k|\mathbf{x}-\mathbf{x}'|), \tag{2.13}$$

$$G^{3D}(\mathbf{x},\mathbf{x}') = \frac{1}{4\pi}\frac{e^{ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}. \tag{2.14}$$

Here, $H_0^{(1)}$ in equation (2.13) corresponds to the zero-order Hankel function of the first kind.

The Lippmann-Schwinger equation

The LS equation is a cornerstone in scattering theory (Ikelle and Amundsen, 2005). The derivations of the LS equation can be found in many literature sources (see, for example, Morse and Feshbach (1953), Lo and Inderwiesen (1994), Ikelle and Amundsen (2005), Bleistein et al. (2001), Jakobsen (2012), Jakobsen and Ursin (2015). In this thesis, a derivation of the LS equation is given mainly based on Lo and Inderwiesen (1994).

First, consider figure 2.1 which represents the scattering process schematically. Here, we can recognize two different wavefields associated with the scattering problem. The first one is the incident wavefield $\Psi^{(0)}(\mathbf{x})$ which was initiated by the source f_s located at some position $\mathbf{x}$ and propagates in the outward direction. The propagation velocity of this wavefront we denote as $c_0(\mathbf{x})$. Note, there is not any scattering wave initiated by the incident wave until it reaches the inhomogeneity. As soon as the incident field $\Psi^{(0)}(\mathbf{x})$ arrives at the inhomogeneity, Huygen's principle gives rise to the scattering wavefield $\Psi^{(1)}(\mathbf{x})$. According to the Huygen's principle, $\Psi^{(1)}(\mathbf{x})$ travels away from the point of disturbance in all directions. Generally, this process is not restricted to only one scattering interaction: scattering from other inhomogeneities also occurs. Mathematically, the superposition of the incident and scattered wavefields can be expressed in the following way

$$\Psi(\mathbf{x}) = \Psi^{(0)}(\mathbf{x}) + \Psi^{(1)}(\mathbf{x}). \tag{2.15}$$

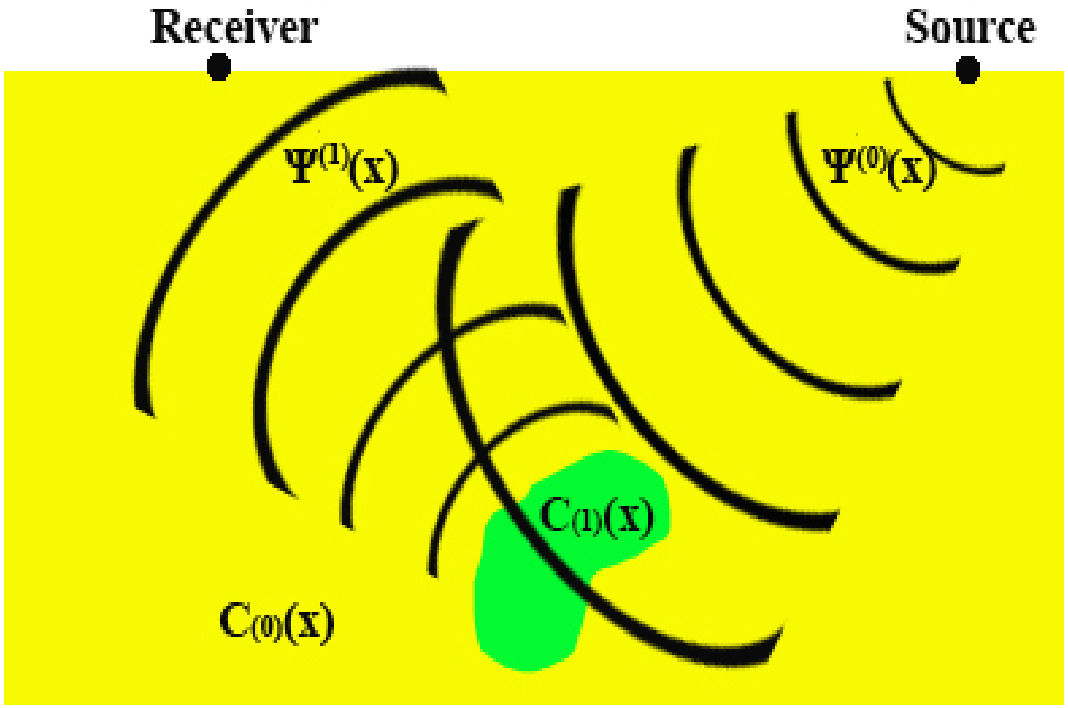

Figure 2.1: A scheme of the seismic scattering process.

Now, we reformulate the squared wavenumber $k^2(\mathbf{x})$ in equation (2.2) such that it represents both the background and perturbed media in the following form

$$k^2(\mathbf{x}) = k_0^2(\mathbf{x}) + \left[k^2(\mathbf{x}) - k_0^2(\mathbf{x})\right], \tag{2.16}$$

where the magnitude of $k_0(\mathbf{x})$ is given by

$$k_0(\mathbf{x}) = \frac{\omega}{c_0(\mathbf{x})}. \tag{2.17}$$

Pulling out $k_0^2(\mathbf{x})$ from the brackets in equation (2.16), we can write

$$k^2(\mathbf{x}) = k_0^2(\mathbf{x}) + k_0^2(\mathbf{x}) \left[\frac{k^2(\mathbf{x})}{k_0^2(\mathbf{x})} - 1\right], \tag{2.18}$$

or, in terms of velocities

$$k^2(\mathbf{x}) = k_0^2(\mathbf{x}) + k_0^2(\mathbf{x}) \left[\frac{c_0^2(\mathbf{x})}{c^2(\mathbf{x})} - 1\right]. \tag{2.19}$$

If we define the bracketed term in equation (2.19) as the model function $m(\mathbf{x})$ (Lo and Inderwiesen, 1994) (also known as the non-normalized contrast function (Jakobsen and Ursin, 2015)), then

$$m(\mathbf{x}) = \left[\frac{c_0^2(\mathbf{x})}{c^2(\mathbf{x})} - 1\right], \tag{2.20}$$

and so, equation (2.19) can be rewritten in more compact form

$$k^2(\mathbf{x}) = k_0^2(\mathbf{x}) + k_0^2 m(\mathbf{x}). \tag{2.21}$$

Substituting equations (2.15) and (2.21) in the Helmholtz equation (equation 2.2), we establish the relationship between the scattered field $\Psi^{(1)}(\mathbf{x})$ and the model function $m(\mathbf{x})$

$$\left[\nabla^2 + k_0^2(\mathbf{x}) + k_0^2(\mathbf{x})\, m(\mathbf{x})\right]\left[\Psi^{(0)}(\mathbf{x}') + \Psi^{(1)}(\mathbf{x}')\right] = -f_s(\mathbf{x}). \tag{2.22}$$

After some rearrangements in equation (2.22) and subtraction of the background field, we arrive at

$$\left[\nabla^2 + k_0^2(\mathbf{x})\right]\Psi^{(1)}(\mathbf{x}') = -k_0^2(\mathbf{x}) m(\mathbf{x})\left[\Psi^{(0)}(\mathbf{x}') + \Psi^{(1)}(\mathbf{x}')\right]. \tag{2.23}$$

Note that the right-hand side term in equation (2.23) can be treated as a virtual source (Jakobsen, 2012; Jakobsen and Ursin, 2015). Solving equation (2.23) for the perturbed wavefield $\Psi^{(1)}(\mathbf{x})$ we get

$$\Psi^{(1)}(\mathbf{x}) = k_0^2(\mathbf{x}) \int_\Omega G^{(0)}(\mathbf{x}, \mathbf{x}') m(\mathbf{x}')\left[\Psi^{(0)}(\mathbf{x}') + \Psi^{(1)}(\mathbf{x}')\right] d\mathbf{x}', \tag{2.24}$$

with Ω denoting the domain where the scattering potential is non-zero. The solution for the background wavefield can be easily obtained by referring to the Helmholtz equation (2.2) again

$$\left[\nabla^2 + k_0^2(\mathbf{x})\right]\Psi^{(0)}(\mathbf{x}) = -f_s(\mathbf{x}). \tag{2.25}$$

The Green function for the reference medium is assumed to be known. Therefore, after applying the source representation integral (equation 2.11), the solution for the background wavefield can be written as

$$\Psi^{(0)}(\mathbf{x}) = \int G^{(0)}(\mathbf{x}, \mathbf{x}') f_s(\mathbf{x}')\, d\mathbf{x}'. \tag{2.26}$$

Finally, substituting equations (2.26) and (2.24) for the background and perturbed fields in equation (2.15) for the total field, we end up with the well-known Lippmann-Schwinger equation

$$\Psi(\mathbf{x}) = \Psi^{(0)}(\mathbf{x}) + k_{(0)}^2(\mathbf{x}) \int_\Omega G^{(0)}(\mathbf{x}, \mathbf{x}')\, m(\mathbf{x}')\left[\Psi^{(0)}(\mathbf{x}') + \Psi^{(1)}(\mathbf{x}')\right] d\mathbf{x}'. \tag{2.27}$$

Note that the integral equation defining the perturbed field (2.24) is nonlinear. This nonlinearity arises from the product inside the integrand where the unknown perturbed field $\Psi^{(1)}(\mathbf{x})$ depends on the model function $m(\mathbf{x}')$. Being part of the LS equation, this nonlinearity causes computational difficulties when one tries to solve it directly. One of the ways to overcome these complications is to linearize the second term in equation

(2.27). Such linearization of the LS equation is known as the Born approximation (Ikelle and Amundsen, 2005).

In many modelling cases, linear approximations can be very useful. Saving computational time, they may provide results comparable to exact solutions. However, any linearization is often valid only under some specific conditions. This is also true for the Born approximation.

2.2.3 The Born series and the first Born approximation

Derivation

Motivated by Ikelle and Amundsen (2005), we first derive an expression for the so-called Born series by rearranging equation (2.27) in the following operator form:

$$\Psi = \Psi^{(0)} + G^{(0)}\delta L\Psi. \tag{2.28}$$

Here, δL corresponds to the modified model function that incorporates the squared wavenumber $k_0^2(\mathbf{x})$, which is explicitly defined as

$$\delta L = \delta L(\mathbf{x}) = k_0^2\, m(\mathbf{x}). \tag{2.29}$$

By expanding equation (2.28) as a Taylor series, we arrive at the Born (Neumann) scattering series (Ikelle and Amundsen, 2005)

$$\Psi = \Psi^{(0)} + G^{(0)}\delta L\Psi^{(0)} + G^{(0)}\delta LG^{(0)}\delta L\Psi^{(0)} + G^{(0)}\delta LG^{(0)}\delta LG^{(0)}G^{(0)}\delta L\Psi^{(0)} + \dots \tag{2.30}$$

Rewriting equation (2.30) in more compact form (Schuster, 2017), we have

$$\Psi = \sum_{n=0}^{\infty} = \left[G^{(0)}\delta L\right]^i \Psi^{(0)} = \Psi^{(0)} + \Psi^{(1)} + \Psi^{(2)} \dots \tag{2.31}$$

Here, the zeroth-order term $\Psi^{(0)}$ represents the direct wave propagating in the reference medium. The first-order term $\Psi^{(1)}$ describes the wave propagating from a source located in the reference medium to a scattering point, interacting with it, and then going to a receiver. The second order term $\Psi^{(2)}$ is responsible for the first order multiple response and so on (Schuster, 2017; Jabbari, 2016). Truncating the series in equation (2.31) to its first two terms, under the assumption that the contribution of the higher terms are negligibly small (Ikelle and Amundsen, 2005), we can write

$$\Psi^{(0)}(\mathbf{x}) + \Psi^{(1)}(\mathbf{x}) \approx \Psi^{(0)}(\mathbf{x}), \tag{2.32}$$

where the scattered wavefield $\Psi^{(1)}(\mathbf{x})$ is assumed to be much weaker than the reference wavefield $\Psi^{(0)}(\mathbf{x})$. Substituting equation (2.32) into (2.24), we can express the linearized form of the perturbed wavefield $\Psi^{(1)}(\mathbf{x})$ as

$$\Psi^{(1)}(\mathbf{x}) = k_0^2(\mathbf{x}) \int_\Omega G^{(0)}(\mathbf{x}, \mathbf{x}')\, m(\mathbf{x}')\Psi^{(0)}(\mathbf{x}')d\mathbf{x}'. \tag{2.33}$$

Finally, the mathematical representation of the first Born approximation (or just Born approximation), is now given by

$$\Psi(\mathbf{x}) = \Psi^{(0)}(\mathbf{x}) + k_0^2(\mathbf{x}) \int_\Omega G^{(0)}(\mathbf{x}, \mathbf{x}')\, m(\mathbf{x}')\Psi^{(0)}(\mathbf{x}')d\mathbf{x}'. \tag{2.34}$$

Validity of the Born approximation

The Born approximation, if the Green's function in the background medium is known, can be very attractive in seismic modelling for the following reasons (Lo and Inderwiesen, 1994; Ikelle and Amundsen, 2005):

- first, the use of only first two terms of the Born series makes it possible to compute primary waves only;
- secondly, the linear relationship between the data $\Psi(\mathbf{x})$ and model function $m(\mathbf{x})$ significantly simplifies the inverse problem.

Despite these advantages, for the Born approximation to be accurate, the velocity contrast between the background and perturbed media must be small, or, speaking mathematically, it must satisfy the following validity criterion (Bleistein et al., 2001; Schuster, 2017)

$$||k_0^2(\mathbf{x}) \int_\Omega G^{(0)}(\mathbf{x}, \mathbf{x}')\, m(\mathbf{x}')dx'|| \ll 1. \tag{2.35}$$

The Born approximation is also known as a weak-scattering approximation (Keller, 1969). Ignoring of equation (2.35) can lead to inaccuracies in both travel times and amplitudes (Ikelle and Amundsen, 2005). Ikelle and Amundsen (2005) tested the assumption behind the criterion in equation (2.35) by conducting a simple 1D modelling experiment. In those tests, the isotropic model included a homogeneous background medium with a single slab embedded (perturbation). Repeatedly changing the thickness and the velocity in the slab, the authors observed the accuracy of the Born modelling approach by comparing it with the exact modelling (analytical) solution. After performing a series of tests, the authors ended up with the following conclusions: when the contrast between the background

medium and embedded layer was significant, or if the propagation time through the slab was long enough, the arrival time from the bottom of the slab failed to display accurate results (since the Born approximation assumes that the velocity of the slab is equal to the velocity in the background medium). Also, if the contrast between the background medium and embedded layer was relatively high, the reflection amplitudes from the bottom of the slab displayed incorrectly.

2.2.4 Matrix representation of the Lippmann-Schwinger equation and T-matrix perspective

Discretization of any analytical equation is an important step toward practical solution. Using the discretization scheme developed by Jakobsen (2012), in the following, I will show how this can be implemented to solve the LS equation directly and with using the T-matrix approach.

Inspired by Jakobsen (2012), Jakobsen and Ursin (2015) and Jakobsen and Wu (2016), we start with rewriting the LS equation in the following form

$$\Psi(\mathbf{x}) = \Psi^{(0)}(\mathbf{x}) + \int_{\Omega} G^{(0)}(\mathbf{x}, \mathbf{x}')\, \delta L(\mathbf{x}') \Psi^{(0)}(\mathbf{x}')\, d\mathbf{x}'. \tag{2.36}$$

From the equation above and equations (2.11)and (2.26), one can deduce that the Green functions $G(\mathbf{x}, \mathbf{x}')$ and $G^{(0)}(\mathbf{x}, \mathbf{x}')$ for the actual and the reference medium are related via the following equation (Jakobsen and Ursin (2015))

$$G(\mathbf{x}, \mathbf{x}') = G^{(0)}(\mathbf{x}, \mathbf{x}') + \int_{\Omega} G^{(0)}(\mathbf{x}, \mathbf{x}'')\, \delta L(\mathbf{x}'')\, G(\mathbf{x}'', \mathbf{x}')\, d\mathbf{x}''. \tag{2.37}$$

Using equation (2.37), we rewrite equations (2.29) and (2.36) as

$$\Psi(\mathbf{x}) = \Psi(\mathbf{x}) + \int_{\Omega} \int_{\Omega} \bar{G}^{(0)}(\mathbf{x}, \mathbf{x}_1)\, \tilde{M}(\mathbf{x}_1, \mathbf{x}_2)\, \Psi(\mathbf{x}_2)\, d\mathbf{x_2} d\mathbf{x_1}, \tag{2.38}$$

and

$$\bar{G}^{(0)}(\mathbf{x}, \mathbf{x}') = \bar{G}^{(0)}(\mathbf{x}, \mathbf{x}') + \int_{\Omega} \int_{\Omega} \bar{G}^{(0)}(\mathbf{x}, \mathbf{x}_1)\, \tilde{M}(\mathbf{x}_1, \mathbf{x}_2)\, \bar{G}(\mathbf{x}_2, \mathbf{x}')\, d\mathbf{x}_2 d\mathbf{x}_1, \tag{2.39}$$

where

$$\tilde{M}(\mathbf{x}_1, \mathbf{x}_2) = m(\mathbf{x}_1)\, \delta(\mathbf{x}_1 - \mathbf{x}_2), \tag{2.40}$$

denotes the scattering potential of a non-normalized contrast function (equation (2.20)) compatible with the use of arbitrary heterogeneous media. Note, $\bar{G}^{(0)}(\mathbf{x}, \mathbf{x}')$ in equations (2.38) and (2.39) represent the modified Green's function

$$\bar{G}^{(0)}(\mathbf{x}, \mathbf{x}') = \omega^2\, G^{(0)}(\mathbf{x}, \mathbf{x}'). \tag{2.41}$$

Here, in equation (2.41), the angular frequency ω^2 absorbed into the modified Green function $\bar{G}^{(0)}(\mathbf{x}, \mathbf{x}')$ allows treating the remaining portion of the interaction (one now depends only on the spatial variation (Kouri and Vijay, 2003)) in a way suitable for inversion.

Discretization of the Lippmann-Schwinger equation

The discretization scheme developed by Jakobsen (2012) and Jakobsen and Ursin (2015) is as follows: first, we define a set of receivers located at positions x_r, where $r = 1, ..., N_r$. The volume Ω, where the scattering potential δL is non-zero, is split up into a number of equal N grid blocks with x_p and δV_p denoting the centers and volumes of these blocks respectively, where $p = 1, ..., N$, is a sequential number of the grid blocks. To avoid spatial aliasing, a size of the grid blocks should be chosen in conjunction with the spatial Nyquist criterion (see, for example, Yilmaz (2001)). If we assume that n corresponds to the index defining the field at a particular receiver position or inside a particular grid block, then the discretized versions of equations (2.38) and (2.39) are

$$\Psi_n = \Psi_n^{(0)} + \sum_{p=1}^{N} \sum_{q=1}^{N} \bar{G}_{np} M_{pq} \Psi_q \tag{2.42}$$

and

$$\bar{G}_{mn} = \bar{G}_{mn}^{(0)} + \bar{G}_{mp}^{(0)} M_{pq} \bar{G}_{qn}, \tag{2.43}$$

where

$$M_{pq} = m_p \delta v_p \delta_{pq}, \quad \text{(no sum over p)}. \tag{2.44}$$

Here, δ_{pq} is the Kronecker delta, with $\delta_{pq} = 1$ if $p = q$; $\delta_{pq} = 0$ if $p \neq q$.
The Green function, in turn, can be discretized, for example, according to Levinson and Markel (2016), implying

$$G_{pq}^{(0)} = G^{(0)}(\mathbf{x}_p, \mathbf{x}_q), \quad p \neq q, \tag{2.45}$$

and

$$\delta v_p G_{pp}^{(0)} = \int_{\Omega_p} d\mathbf{x} G^{(0)}(\mathbf{x}_p, \mathbf{x}). \tag{2.46}$$

Here, Ω_p in equation 2.46 indicates the domain corresponding to a single block centered at position x_p.

In the following, two groups of matrices corresponding to equations (2.42) and (2.43) will be distinguished: source-dependent and source-independent. There are two reasons for this division. Firstly, it will be convenient for the further computations, since the

only source-independent Green functions need to be multiplied by the angular frequency ω^2. Secondly, it is also consistent with the T-matrix modelling approach (which will be discussed later in this section).

For the source-independent group, we arrange the discretized fields (both reference and perturbed) at the receiver (R) and volume (V) positions into the corresponding vectors $\mathbf{\Psi}_R$ ($\mathbf{\Psi}_R^{(0)}$) and $\mathbf{\Psi}_V$ ($\mathbf{\Psi}_V^{(0)}$). From (2.42), we can write (Jakobsen and Ursin, 2015)

$$\mathbf{\Psi}_R = \mathbf{\Psi}_R^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)}\mathbf{M}\mathbf{\Psi}_R, \tag{2.47}$$

$$\mathbf{\Psi}_V = \mathbf{\Psi}_V^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)}\mathbf{M}\mathbf{\Psi}_V, \tag{2.48}$$

where, $\mathbf{M}$ is a diagonal $N \times N$ matrix containing the scattering potential on its diagonal. Similarly, from equation (2.43), we can deduce the relations for the Green functions $\mathbf{G}_{RV}$ ($\mathbf{G}_{RV}^{(0)}$) at the receiver-volume and $\mathbf{G}_{VV}$ ($\mathbf{G}_{VV}^{(0)}$) and volume-volume domains (Jakobsen and Ursin, 2015)

$$\bar{\mathbf{G}}_{RV} = \bar{\mathbf{G}}_{RV}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)}\mathbf{M}\bar{\mathbf{G}}_{VV}, \tag{2.49}$$

$$\bar{\mathbf{G}}_{VV} = \bar{\mathbf{G}}_{VV}^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)}\mathbf{M}\bar{\mathbf{G}}_{VV}. \tag{2.50}$$

The source-dependent Greens functions for both the reference and perturbed media at the source-receiver $\mathbf{G}_{RS}$ ($\mathbf{G}_{RS}^{(0)}$) and volume-source domains $\mathbf{G}_{VS}$ ($\mathbf{G}_{VS}^{(0)}$) are given by (Jakobsen and Ursin, 2015)

$$\mathbf{G}_{RS} = \mathbf{G}_{RS}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)}\mathbf{M}\mathbf{G}_{VS}, \tag{2.51}$$

$$\mathbf{G}_{VS} = \mathbf{G}_{VS}^{(0)} + \bar{\mathbf{G}}_{VS}^{(0)}\mathbf{M}\mathbf{G}_{VS}. \tag{2.52}$$

Here, the components of the $N_r \times N_s$ – dimensional matrix $\mathbf{G}_{RS}^{(0)}$ in equation (2.51) represents the source position $\mathbf{x}_s$ ($s = 1, \ldots, N_s$) and receiver position $\mathbf{x}_r$ ($r = 1, \ldots, N_r$). The elements of the $N \times N_s$ -dimensional matrix $\mathbf{G}_{VS}^{(0)}$ in equation (2.52) relate to source position $\mathbf{x}_s$ ($s = 1, \ldots, N_s$) and one scattering grid block position $\mathbf{x}_j$ ($j = 1, \ldots, N$). The solution of the LS equations can be obtained directly from equations (2.51) and (2.52) by solving for the Green functions. But, the forward modelling problem, as will be shown in the next subsection, can be solved more efficiently by applying the T-matrix approach, where the inversion of two matrices corresponding to the Green functions in equations (2.51) and (2.52) are replaced by inversion of that one corresponding to the T-matrix. (Jakobsen, 2012; Jakobsen and Ursin, 2015).

Transition operator and the T-matrix approach

The T-matrix, or a transition operator $\mathbf{T}$, is known from quantum mechanical scattering theory (Jakobsen and Ursin, 2015). Mathematically, it is defined by (Jakobsen, 2012;

Jakobsen and Ursin, 2015; Jakobsen and Wu, 2016, 2018)

$$\mathbf{M}\boldsymbol{\Psi}_V = \mathbf{T}\boldsymbol{\Psi}_V^{(0)}. \tag{2.53}$$

Using the above equation in combination with equation (2.48), we can write

$$\boldsymbol{\Psi}_V = \boldsymbol{\Psi}_V^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)}\mathbf{T}\boldsymbol{\Psi}_V^{(0)}. \tag{2.54}$$

Multiplying both sides of equation (2.54) by $\mathbf{M}$ and applying equation (2.53) again, we get

$$\mathbf{T}\boldsymbol{\Psi}_V^{(0)} = \mathbf{M}\boldsymbol{\Psi}_V^{(0)} + \mathbf{M}\bar{\mathbf{G}}_{VV}^{(0)}\mathbf{T}\boldsymbol{\Psi}_V^{(0)}. \tag{2.55}$$

Since $\boldsymbol{\Psi}_V^{(0)}$ is arbitrary, equation (2.55) can be reduced to the following form

$$\mathbf{T} = \mathbf{M} + \mathbf{M}\bar{\mathbf{G}}_{VV}^{(0)}\mathbf{T}. \tag{2.56}$$

Here, equation (2.56), being also known as the LS equation for the T-matrix, represents all nonlinear effects of multiple scattering. The explicit form for the T-matrix, can easily be obtained from equation (2.56) as

$$\mathbf{T} = (\mathbf{I} - \mathbf{M}\,\bar{\mathbf{G}}_{VV}^{(0)})^{-1}\mathbf{M}. \tag{2.57}$$

Note, the T-matrix itself, does not bring in any new information, however, being independent of the source-receiver configuration, it significantly simplifies a computation process.

Now, using the identity (2.53) as a fundamental definition for the T-matrix, equations (2.49-2.52) can be rewritten in the following forms:

$$\bar{\mathbf{G}}_{RV} = \bar{\mathbf{G}}_{RV}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)}\mathbf{T}\bar{\mathbf{G}}_{VV}^{(0)}, \tag{2.58}$$

$$\bar{\mathbf{G}}_{VV} = \bar{\mathbf{G}}_{VV}^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)}\mathbf{T}\bar{\mathbf{G}}_{VV}^{(0)}, \tag{2.59}$$

$$\mathbf{G}_{RS} = \mathbf{G}_{RS}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)}\mathbf{T}\mathbf{G}_{VS}^{(0)}, \tag{2.60}$$

and

$$\mathbf{G}_{VS} = \mathbf{G}_{VS}^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)}\mathbf{T}\mathbf{G}_{VS}^{(0)}. \tag{2.61}$$

Note, now, the use of equations (2.58-2.61) make it possible to solve for both the wavefields and Green functions in any heterogeneous background media.

Moving forward, we extend the modelling theory to be compatible with the use of the number of frequencies and sources. To do this, we start by considering equation (2.15) for the total field. The components of this equation are just products of the source vector

$\tilde{\mathbf{f}}_s$ containing all information about sources distribution, and the corresponding Green functions $\mathbf{G}_{RS}$ and $\mathbf{G}^{(0)}_{RS}$ at the receiver-source domains (Jakobsen, 2012)

$$\mathbf{\Psi}_{RS} = \mathbf{G}_{RS}\tilde{\mathbf{f}}_s, \tag{2.62}$$

$$\mathbf{\Psi}^{(0)}_{RS} = \mathbf{G}^{(0)}_{RS}\tilde{\mathbf{f}}_s. \tag{2.63}$$

If we define the perturbed wavefield as a vector $\mathbf{d}$, representing the difference between wavefields given in equations (2.62) and (2.63), then, we can write (Jakobsen and Wu, 2016; Eikrem et al., 2017; Jakobsen and Wu, 2018)

$$\mathbf{d} = [\mathbf{G}_{RS} - \mathbf{G}^{(0)}_{RS}]\tilde{\mathbf{f}}_s. \tag{2.64}$$

This equation in conjunction with equation (2.51), gives us

$$\mathbf{d} = \bar{\mathbf{G}}^{(0)}_{RV}\mathbf{M}\mathbf{G}_{VS}\tilde{\mathbf{f}}_s. \tag{2.65}$$

Assuming a set of discrete frequencies defined by ω_k ($k = 1, \ldots, N_\omega$), equation (2.65) can be rewritten exactly as (Jakobsen and Ursin, 2015)

$$\tilde{d}_{rn,sk} = \sum_{n=1}^{N} J_{rn,sk}\, m_n, \tag{2.66}$$

where

$$J_{rn,sk} = \bar{G}^{(0)}_{rn}(\omega_k)\delta M_n G_{ns}(\omega_k) f_s(\omega_k) \tag{2.67}$$

and

$$m_n = (c_n)^{-2} - (c^{(0)}_n)^{-2}, \tag{2.68}$$

with c_n and $c^{(0)}_n$ known wavespeeds in the background and perturbed media. From a practical perspective, it is more convenient to replace indices s, r and k by a single index α ($s, r, k \to \alpha$), where $\alpha = 1, 2, 3, ..., N_d$, with $N_d = N_s N_r N_w$. If we do this, then, equation (2.66) can be written as

$$d_\alpha = \sum_{n=1}^{N} J_{\alpha n}\, m_n, \tag{2.69}$$

or, in matrix form as

$$\mathbf{d} = \mathbf{J}\mathbf{m}. \tag{2.70}$$

Equation (2.70), being a system of linear equations, reflects a non-linear relation between the data-vector $\mathbf{d}$ and the contrast function $\mathbf{m}$ via the matrix $\mathbf{J}$. The physical meaning of the matrix $\mathbf{J}$ can be interpreted as the sensitivity of the scattered field with respect to a small perturbation of the wave-speed in the corresponding nth block of a volume for a given source, receiver and frequency. Note, the form of equation (2.70) is typical way to represent a modelling problem also suitable for inversion purposes (will be explained in chapter 3).

Discretization of the Born approximation

The integral equation for the perturbed wavefield associated with the Born approximation (2.33) can be easily discretized by using the same discretization scheme introduced in the previous subsection. Here, the similarities arise from the structure of equations (2.24) and (2.33). Replacing $\mathbf{G}_{VS}$ by $\mathbf{G}^{(0)}_{VS}$ in equation (2.65), we can define the data vector $\mathbf{d}$ for the perturbed wavefield as

$$\mathbf{d} = \bar{\mathbf{G}}^{(0)}_{RV}\mathbf{M}\mathbf{G}^{(0)}_{VS}\tilde{\mathbf{f}}_s \tag{2.71}$$

Again, assuming the set of frequencies and sources, and repeating all the corresponding steps described in the previous subsection, we end up with the matrix formulation of the perturbed wavefield (see, for exmaple, Eikrem et al. (2016)) for the Born approximation similar to equation (2.70). In this case, the sensitivity matrix $\mathbf{J}$ characterizes a linear relationship between the data-vector $\mathbf{d}$ and the contrast function $\mathbf{m}$.

The distorted Born approximation

The DBA is another seismic modelling method which provides almost exact linearized solution of the LS equation (Chew and Wang, 1990; Jakobsen and Ursin, 2015; Eikrem et al., 2017; Jakobsen and Wu, 2018). I start a derivation of the DBA method by decomposing the scattering potential as (Jakobsen and Ursin, 2015; Eikrem et al., 2017)

$$\delta\mathbf{M}^{(b)} = \mathbf{M} - \mathbf{M}^{(b)}. \tag{2.72}$$

Here, $\delta\mathbf{M}^{(b)}$ is a variation of the scattering potential $\mathbf{M}$ about the heterogeneous reference medium with the scattering potential $\mathbf{M}^{(b)}$. The last one leads to the variation in the Green function $\mathbf{G}^{(b)}_{RS}$, and given by

$$\delta\mathbf{G}^{(b)}_{RS} = \mathbf{G}_{RS} - \mathbf{G}^{(b)}_{RS}, \tag{2.73}$$

or, defined in terms of seismic fields, by

$$\delta\mathbf{\Psi}^{(b)}_{RS} = \mathbf{\Psi}_{RS} - \mathbf{\Psi}^{(b)}_{RS}. \tag{2.74}$$

If we consider $\delta\mathbf{\Psi}^{(b)}_{RS}$ and $\mathbf{\Psi}^{(b)}_{RS}$ as the perturbed fields with respect to the actual $\mathbf{\Psi}_{RS}$, we can write

$$\mathbf{\Psi}^{(b)}_{RS} = \mathbf{G}^{(b)}_{RV}\mathbf{M}^{(b)}\mathbf{G}^{(b)}_{VS}\tilde{\mathbf{f}}_s, \tag{2.75}$$

and

$$\delta\mathbf{\Psi}^{(b)}_{RS} = \mathbf{G}^{(b)}_{RV}\delta\mathbf{M}^{(b)}\mathbf{G}^{(b)}_{VS}\tilde{\mathbf{f}}_s. \tag{2.76}$$

Here, the unknown Green functions $\mathbf{G}^{(b)}_{VS}$ and $\mathbf{G}^{(b)}_{RV}$ (also known as the dynamic Greens functions) for the background medium correspond to the different parts of the scattering-path propagating from the source to the receiver via some volume. They are defined by the following equations Jakobsen and Wu (2018):

$$\mathbf{G}^{(b)}_{VS} = \mathbf{G}^{(0)}_{VS} + \mathbf{G}^{(0)}_{VV}\mathbf{T}^{(b)}\mathbf{G}^{(0)}_{VS}, \tag{2.77}$$

$$\mathbf{G}^{(b)}_{RV} = \mathbf{G}^{(0)}_{RV} + \mathbf{G}^{(0)}_{RV}\mathbf{T}^{(b)}\mathbf{G}^{(0)}_{VV}, \tag{2.78}$$

where, $\mathbf{T}^{(b)}$, is the T-matrix, defined by (Eikrem et al., 2017)

$$\mathbf{T}^{(b)} = (\mathbf{I} + \mathbf{M}^{(b)}\,\mathbf{G}^{(0)}_{VV})^{-1}\mathbf{M}^{(b)}. \tag{2.79}$$

Finally, the solution for the perturbed wavefield can be deduced from equation (2.74). To be implemented for multiple sources and the number of frequencies, one can use the implementation method discussed previously.

2.2.5 Seismic source

A formal definition of the seismic pulse in seismic modelling theory (Rabinovich et al., 2018) is that it represents "the solution of the corresponding nonlinear evolution equation describing the stress-strain states of the propagation medium". Due to the complexity of real media, it is often difficult to represent a point source corresponding to a seismic pulse mathematically. Despite this, there are still numerous of solutions offered (Rabinovich et al., 2018). The most known are the Ricker wavelet, the Berlage wavelet, the Gelfand wavelet (Rabinovich et al., 2018). In this thesis, for synthetic modelling, the Ricker wavelet (Ricker, 1953) is chosen. This choice is based on a close similarity of the Ricker wavelet to a real seismic pulse. Mathematically, the Ricker wavelet takes the following form:

$$R(t) = 1 - (t - \frac{1}{f_0})^2 f_0^2 \pi^2 e^{-(t-\frac{1}{f_0})^2\pi^2 f_0^2}, \tag{2.80}$$

where f_0 is the central frequency and t is recording time.
The Ricker wavelet, with two dominant frequencies 7.5 Hz and 15 Hz (and their amplitude spectra), is shown in figure 2.2.

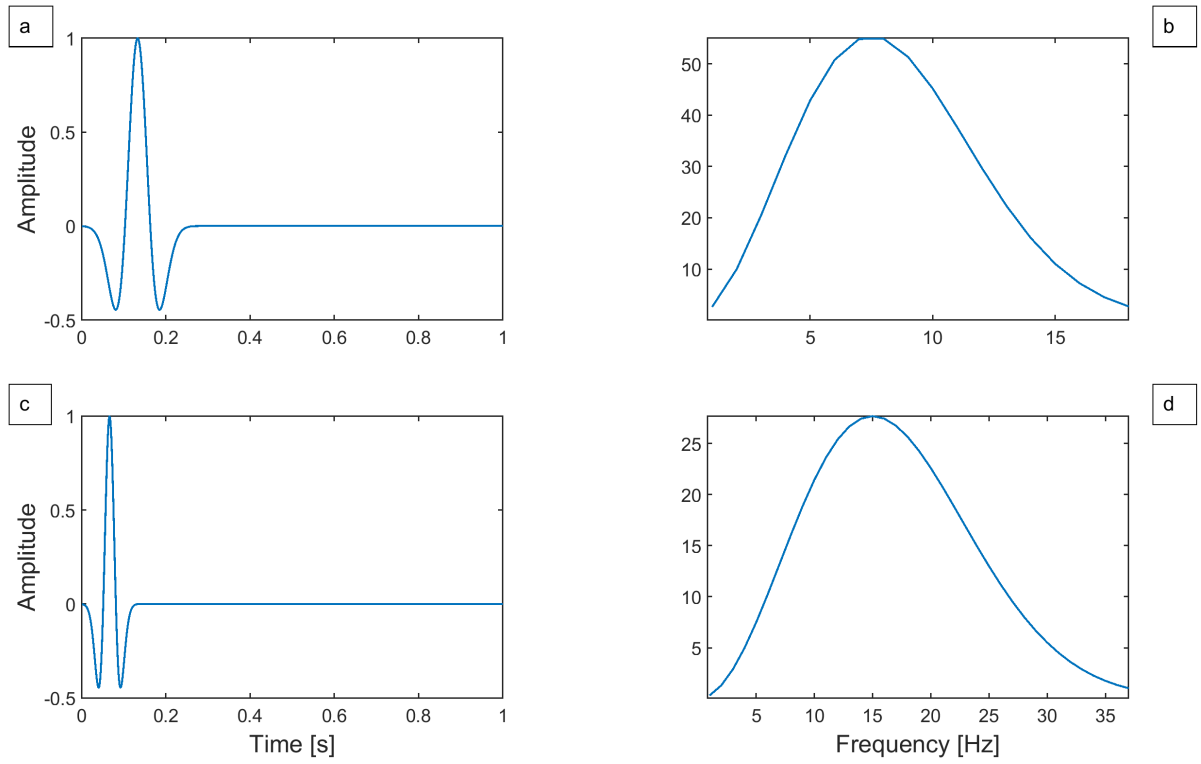

Figure 2.2: The Ricker wavelets with the dominant frequencies 7.5 Hz and 15 Hz ((a) and (c) respectively); and, their corresponded amplitude spectra ((b) and (d)).

2.2.6 Error estimations and random noise

Relative error and root mean square error

Quantitative data comparison may be obtained using different statistical methods. In this thesis, in order to compare data (in both modelling and inversion experiments), two statistical techniques are utilized: a method for the relative error estimation and a method for the root mean square error estimation.

The mathematical definition of the relative error is given by

$$\varepsilon_{rel} = \frac{x_{est} - x_{ex}}{x_{ex}} 100\%. \tag{2.81}$$

Here, x_{est} and x_{ex} denote estimated and exact measurements respectively.

The root mean square error can be found from

$$\varepsilon_{rms} = \sqrt{\frac{\sum_{t=0}^{n}(x_{1,t} - x_{2,t})^2}{n}}, \tag{2.82}$$

where $x_{1,t}$ and $x_{2,t}$ are two datasets for comparison and n is the number of samples in each dataset.

Mathematical definition of random noise

To simulate noisy data in the frequency domain, one can use a simple relation given by Jakobsen and Ursin (2015), in which some random Gaussian noise are added to each frequency component when modelling

$$\mathbf{d}_{noisy} = \mathbf{d} + \frac{|\mathbf{d}|}{\mathrm{SNR}} \frac{\boldsymbol{\mu}}{|\boldsymbol{\mu}|}. \tag{2.83}$$

Here $\mathbf{d}$ is the initial data-vector, $\boldsymbol{\mu}$ is the vector with random numbers added (which has the same length as the data-vector), and SNR is a value for the signal to noise ratio. The last one is defined by

$$\mathrm{SNR} = 10^{(\mathrm{SN}/20)}, \tag{2.84}$$

with SN denoting the value for a noise level given in dB.

2.3 Numerical results

The goal of this section is to demonstrate that the T-matrix integral equation method provides modelling results comparable to those of the exact modelling solution. All tests in the section are divided into two parts.

The first part intends to examine the accuracy of the Born approximation and T-matrix integral equation methods by comparing them with one of the exact modelling approaches. Starting from simple zero-offset experiments based on a single-slab model, I first investigate how consistent the modelling results of the Born and T-matrix modelling methods in response to variable thicknesses and contrasts of the embedded layer. After that, the study presents the examinations of multichannel cases, in order to show how accurate T-matrix modelling results at different offsets and times.

The second part focuses on time-lapse modelling experiments. Using simple geological models, which represent different stages of reservoir production, synthetic seismograms using the Born approximation, DBA and T-matrix techniques are computed. The seismograms are compared in order to determine their accuracy in time-lapse modelling.

The accuracy of the T-matrix integral equation method has already been studied before (Jakobsen and Ursin, 2015). Comparing the modelling results of the T-matrix and finite-difference methods (qualitatively) using the simple 2D model, the authors discovered that the difference between the generated wavefields was very small. This is useful, but as a next step, it is relevant to compare the modelling results quantitatively to see if the error is offset- and time-dependent. In this thesis, to perform such a quantitative comparison, the finite-difference seismic modelling as a reference solution for data comparison are used (similarly to Jakobsen and Ursin (2015)).

2.3.1 The finite-difference method and the methodology for data comparison

The finite-difference method is one of the most successful modelling techniques for solving the wave equation (Ikelle and Amundsen, 2005). In this method the fundamental differential equation (2.1) is discretized and implicitly solved. The technique is applicable in both time and frequency domains. In this thesis, for the purpose of comparing synthetics, the most popular and relatively simple time domain technique (Finite-Difference Time-Domain or FDTD) is used. Appendix A provides a mathematical derivation of the FDTD modelling method.

A direct comparison of the T-matrix (as well as Born and DBA) and FDTD modelling results meets some pitfalls. Scattering theory implies that the background medium of a model is infinite (Ikelle and Amundsen, 2005). This fact excludes any free-surface related reflections when using the scattering theory modelling approaches. In the case of the finite-difference method, on the contrary, these reflections take place. Therefore, for data to be compared, some adjustments in the modelling techniques are needed. There are two ways to adjust the analyzed modelling techniques. The first way involves applying absorbing boundary conditions to both modelling methods. When applied, the absorbing boundary conditions prevent the seismic energy reached the model boundaries from reflecting back to the model. The second way includes extending the free surface when generating finite-difference data such that the associated reflections do not appear within a given recording time. Even the first approach seems to be more attractive, the application of absorbing boundary conditions does not always work well and results often depend on the specifics of the boundary condition techniques employed (Ajo-Franklin, 2005). Therefore, in this thesis, the second approach is used when comparing the modelling data.

Finally, it should be mentioned that in the following data comparison, only the perturbed wavefields are analyzed. For the T-matrix method, this involves considering $\mathbf{\Psi}^{(1)}$ when modelling. For the FDTD method, however, the wavefield decomposition is impossible (the solution to equation (2.1) implies computing the total wavefield only). To obtain the perturbed wavefield by the FDTD modelling method, the following is done. First, the modelling problem is solved for the whole model. Then, the computations are repeated for the background model. Finally, the background wavefield is subtracted from the total wavefield.

2.3.2 Slab model

There two types of slab model are considered in the first set of tests. For the zero-offsets experiments, a simple slab model includes one embedded layer. For the multichannel experiments, additionally to the slab model, a "layer-cake" model simulating a series of horizontally aligned strata in the geological section is analyzed.

Zero-offset experiments

For the first zero-offset experiment, a single-slab model is 20 m in length and 800 m in depth. The thickness of the top, embedded and bottom layers are 80 m, 140-300 m and 580-420 m respectively (variable thicknesses of the embedded and bottom layers mean that they are subjects of change in the modelling tests). The velocities in the slab and background medium are 3000 m/s and 2000 m/s respectively. The size of the grid blocks is equal to 10 m in each direction. One source and one receiver, which are positioned exactly in the middle of the model (at the surface), are used. Total recording time and time sampling interval are 0.8 s and 0.004 s correspondingly. As a source function, the Ricker wavelet function with a dominant frequency of 7.5 Hz defined by equation (2.83) is used.

Discussion

Figure 2.3 shows two velocity models and the corresponding seismograms generated using the FDTD, Born approximation and T-matrix methods. Note, the signals reflected from the top of the slab are identical for all modelling methods. The reflections from the bottom of the slab are similar for the FDTD and the T-matrix modelling methods and quite different for the Born approximation.

In order to see how the amplitudes and travel times of the reflected signals change as a function of slab thicknesses, we gradually change the thickness in the interval from 140 m to 300 m. Choosing a thickness, which increases with the step of 20 m, we generate eight zero-offset seismograms for each modelling method. Some of the generated seismograms are shown in figure 2.4. To estimate the differences in amplitudes and travel times, we pick the first local minimum associated with the slab bottom in each modelled trace (crosses in figure 2.4), define the difference in amplitude and travel times between the first local minimum corresponding to the top of the slab and the picked values obtained before, interpolate the obtained differences to 100 points using spline interpolation (to make results more representative), and compute relative errors (assuming that the FDTD data are the reference data). The errors are shown in figure 2.5. Note the differences in error for the Born approximation (illustrations (a) and (b)) and the T-matrix modelling methods.

In the next zero-offset experiment, we fix the thicknesses of the top, embedded (slab) and bottom layers (80 m, 290 m and 430 m correspondingly) in the model. We vary the velocity in the slab from 2200 m/s to 3000 m/s with an interval of 100 m/s. Figure 2.6 demonstrates two selected models with corresponding seismograms generated for this test. It is worth noting that the T-matrix solution does not show any visual modelling differences in comparison with the FDTD modelling data, while the Born modelling results for the slab bottom are notably different.

Using the same strategy as in the previous test, we pick the first local minima (figure 2.7), estimate the difference in amplitudes and travel times between local minima corresponding to the top and bottom of the slab, interpolate (using spline interpolation) the picked results to 100 points and estimate relative errors. The computed modelling errors for the Born and T-matrix modelling methods are shown in figure 2.8. Again, the magnitudes of the estimated errors for the Born modelling method are much higher than those of the T-matrix method.

Multichannel experiments

For the multichannel experiments, two models are employed: the slab model (similar to the previous tests) and the "layer-cake" model (containing several horizontal layers). In the first case, velocities in the background and perturbed media are equal to 2400 m/s and 3000 m/s respectively (both models are shown in illustrations (a) (figures 2.9 and 2.10)). For the second case, velocities in the model vary from 2400 m/s to 3300

m/s. Both models have a horizontal dimension of 1500 m and a vertical dimension of 750 m. They consist of 1891 grid blocks which have a size of 25 m x 25 m. The other parameters required for the modelling are assumed to be the same for both models. A survey design involves 61 equidistantly spaced receivers located at the surface, and one source positioned exactly in the middle of the aperture (also at the surface). As a source function, the Ricker wavelet function with a central frequency equal to 7.5 Hz (equation 2.80) is used. Total recording time is 1.6 s. A time sampling interval is 0.004 s.

Discussion

Figures 2.9 and 2.10 represent the synthetic seismograms generated by the FDTD and T-matrix modelling methods (illustrations (b) and (c) respectively) and their differences (illustrations (d)). Note, the seismograms (illustrations (b) and (c)) in figure 2.10 include a strong reflection at times 0.55-0.65 s, which is absent in figure 2.9. This is an artificial reflection from the model bottom boundary. It did not appear in figure 2.9, because the velocity in the bottom layer is equal to the background velocity. Also, note that the differences (illustrations (c)) at far offsets in both figures are slightly different than compared to those at near offsets. To estimate modelling errors quantitatively, we do the following: the whole traces in the seismogram we divide into small trace groups (each trace group includes the equal number of traces) and, then, we compute RMS errors for each of these groups (using the FDTD data as reference data). Illustration (a) in figure 2.11 shows such RMS errors for the "layer-cake" model. Here, for the computations, the seismogram was divided into ten trace groups. We note that the RMS errors at far offsets are higher. However, as can be seen from the plot, the errors do not exceed 0.35.

Now, we investigate if the modelling differences are depth dependent. To do this, we choose three traces at different offsets (at the zero offset, at the half an offset and at the maximum offset), divide each of the chosen traces into equal time intervals and compute RMS errors for each of these intervals. Illustration (b) in figure 2.11 demonstrates such error estimations computed for the "layer-cake" model. Here, three traces corresponding to the offsets 0 m, 375 m and 750 m, and, ten time samples in each time group were used. Note, that at far offsets and higher depths, the modelling errors are more evident than for the near offsets.

2.3.3 Fault trap model

A fault trap model is a relatively simple geological model that includes two varieties. The first variety (the baseline model) corresponds to the first stage of the reservoir production (reservoir rocks saturated in hydrocarbons). The second variety (the monitor model) defines the second stage of the reservoir production (reservoir rocks depleted in hydrocarbons). The difference between these two models is the time-lapse model. These two models and their difference (time-lapse model) are shown in figure 2.12. The time-lapse anomaly is represented by a relatively small velocity variation (100 m/s).

A possible lithological interpretation for the trap model is provided in table 2.1. The size of the models (including model sampling), survey and recording parameters are chosen the same as in multichannel experiments in subsection (2.3.2).

Discussion

Figure 2.13 shows the noiseless and noisy synthetic seismograms generated using the baseline model (illustration (a) in figure 2.12) with the help of the T-matrix integral equation method. Note, how clean the wavefield is. All the geological elements in the wavefield can be identified. The seismograms in illustrations (b) and (c) look fuzzier because of the random noise. On the seismogram with a SNR equals to 5 dB (which corresponds to 56% noise), the reflections associated with the geological interfaces can hardly be detected.

Figure 2.14 represents three synthetic seismograms corresponding to the baseline, monitor and the time-lapse models in figure 2.12. All are generated using the FDTD method. Here, the time-lapse event (illustration (c)) is well defined (located approximately at 0.5-0.55 s). Figure 2.15 demonstrates a similar test, but for the Born approximation method. In this case, the time-lapse anomaly appears at times 0.6-0.55 s (approximately). This difference in arrival times arises from the fact that the Born approximation solution propagates in the model with the velocity of the background medium instead of the model velocities. Figure 2.16 shows the synthetic seismograms obtained by the DBA method. The time-lapse anomaly is at the same times as for FDTD modelling (illustration (c) in figure 2.14). For the T-matrix solution, as figure 2.17 shows, the seismic event corresponding to the time-lapse anomaly also comes in at the same time as in the FDTD modelling data. Figure 2.18 concludes the observations above showing the differences between the time-lapse seismograms obtained by the FDTD method, the Born

Table 2.1: Lithological interpretation of the fault trap model

Layers (from the top)	Lithological specification	Velocity $[m/s]$
Layer 1	Weekly consolidated deposits	2400
Layer 2	Sandy siltstones, mudstones	2600
Layer 3	Calcareous sandstones (reservoir rocks)	3000
Layer 4	Oil-saturated sandstones	2900
Layer 5	Limestones	3300

approximation (illustration (a)), DBA (illustration (b)) and T-matrix (illustration (c)) methods. For the Born approximation, the difference is quite large (illustration (a)). For the DBA and T-matrix modelling approaches, the residuals are very small (illustrations (b) and (c) correspondingly). Also, note that the DBA and T-matrix results are similar.

2.3.4 Concluding remarks

In this section, I briefly highlight some essential points of the results presented above. First, the zero-offset experiments (figures 2.3-2.8) showed that the modelling errors (relative errors) in amplitudes and travel times associated with the T-matrix integral equation method were minimal (around 1-2 %). Based on the multichannel tests (figures 2.9-2.11), it can be concluded that the modelling errors associated with the T-matrix technique are time- and depth depended. However, the errors (RMS errors) are small (do not exceed a value of 1). The analysis of the time-lapse seismograms generated by the Born approximation, DBA and the T-matrix modelling methods show that the DBA and T-matrix techniques are characterized by extremely small differences when compared to the FDTD results (figures 2.14, 2.16, 2.17 and 2.18 (illustration (b) and (c)), while the Born modelling approach, on the contrary, demonstrates large errors (illustration (a) in figure 2.18). It is also important to emphasize the close similarities between the DBA and T-matrix time-lapse modelling results (illustrations (b) and (c) in figure 2.18). These conclusions, in conjunction with the results provided by Jakobsen and Ursin (2015), I think, suggest that the T-matrix method can be considered to be an exact modelling method.

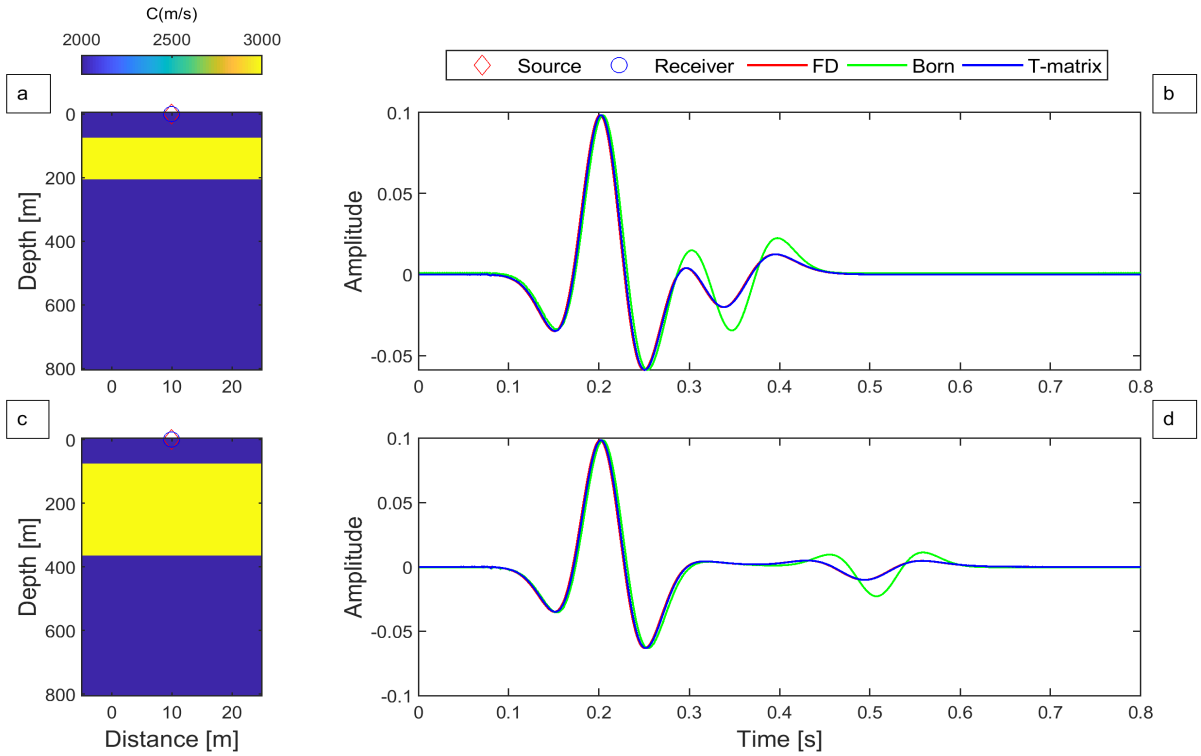

Figure 2.3: Zero-offset modelling experiment. (a) and (c) are the slab models with different thicknesses; (b) and (d) are the corresponding synthetic seismograms obtained using the FDTD, Born and T-matrix modelling methods.

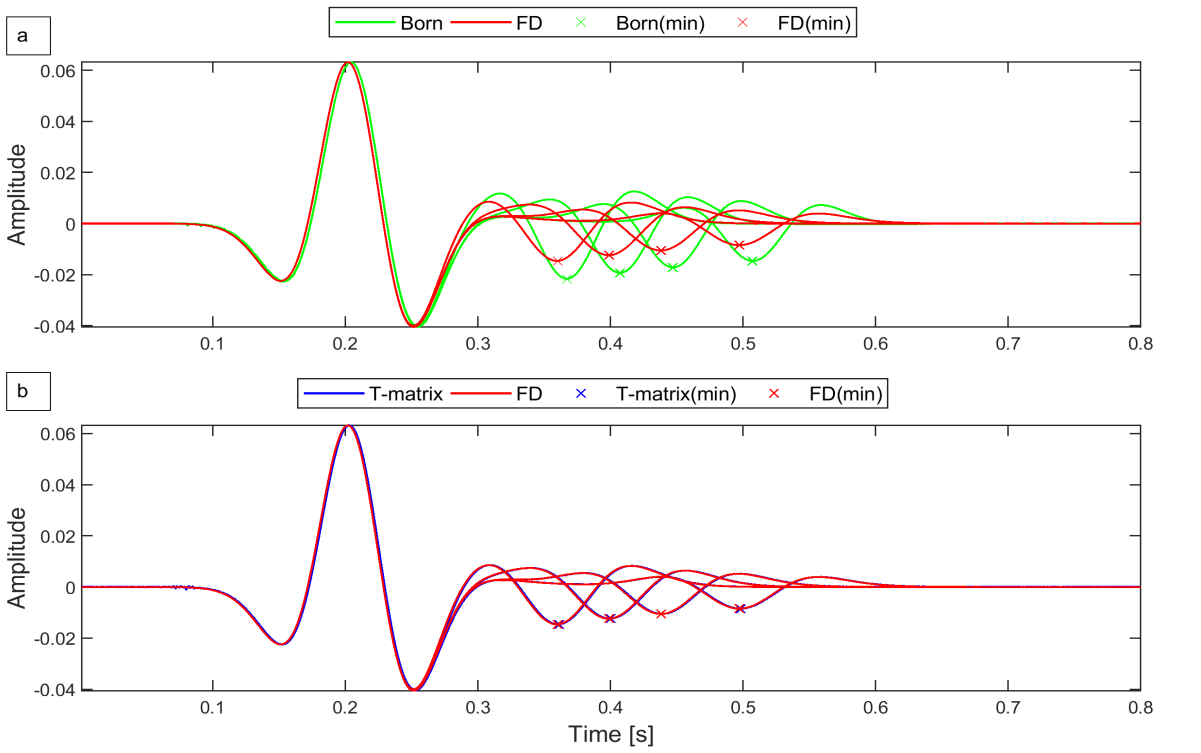

Figure 2.4: Zero-offset synthetic traces with the picked first minima (different slab thicknesses). The modelled traces obtained using the (a) Born approximation and the FDTD method; (b) T-matrix and FDTD methods.

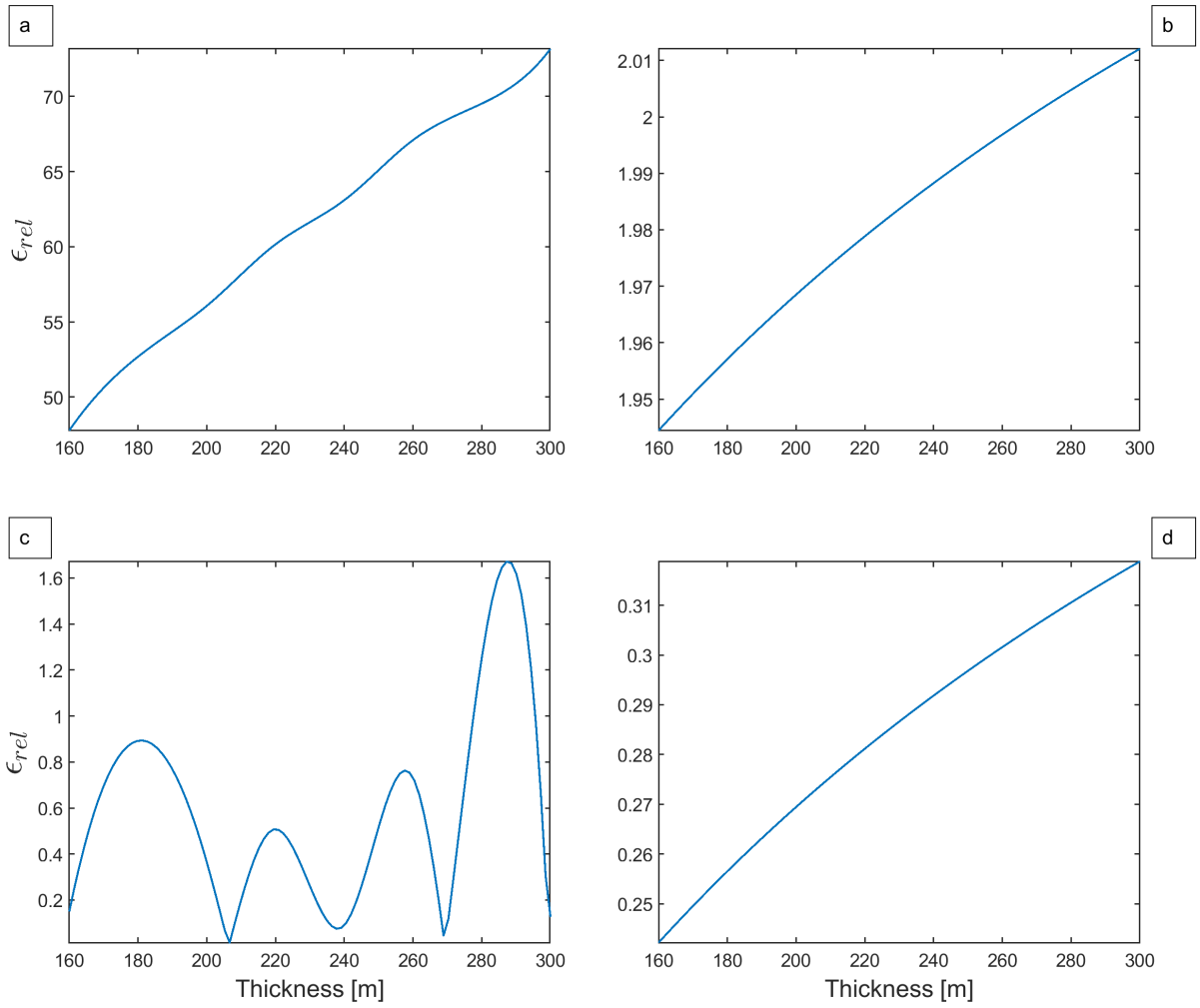

Figure 2.5: Relative errors for the Born approximation and T-matrix modelling methods in the slab models in figure 2.4. (a) and (c) correspond to the amplitude-related relative errors (the Born and T-matrix methods respectively); (b) and (d) correspond to the travel times-related relative errors (the Born and T-matrix methods respectively).

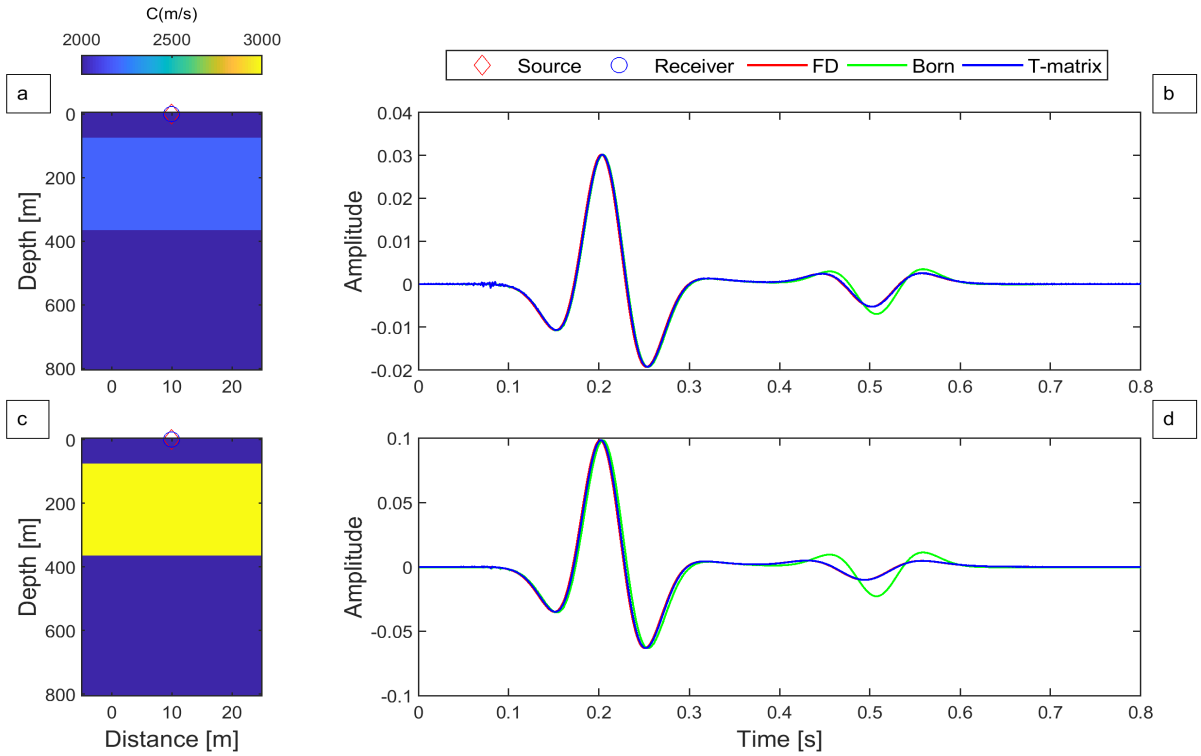

Figure 2.6: Zero-offset modelling experiment. (a) and (c) are slab models with different velocities; (b) and (d) are the corresponding synthetic seismograms obtained using the FDTD, Born and T-matrix modelling methods.

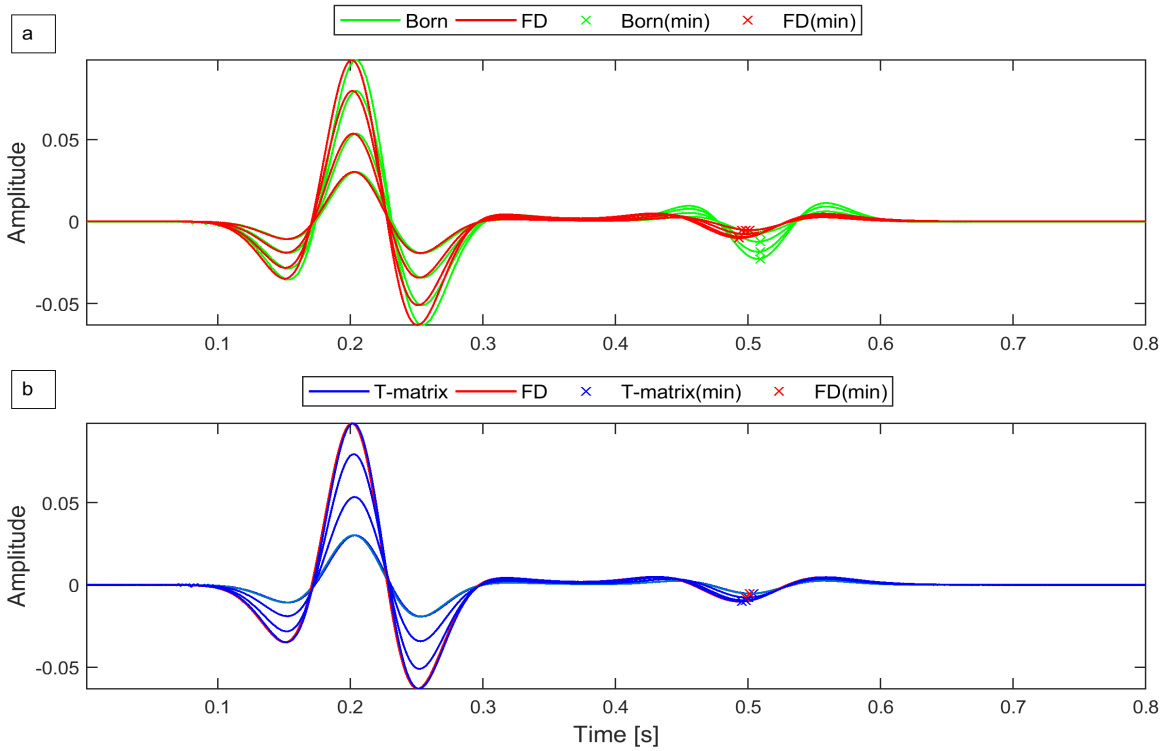

Figure 2.7: Zero-offset synthetic traces with the picked first minima (different slab velocities). The modelled traces obtained using the (a) Born approximation and FDTD method; (b) T-matrix and FDTD methods.

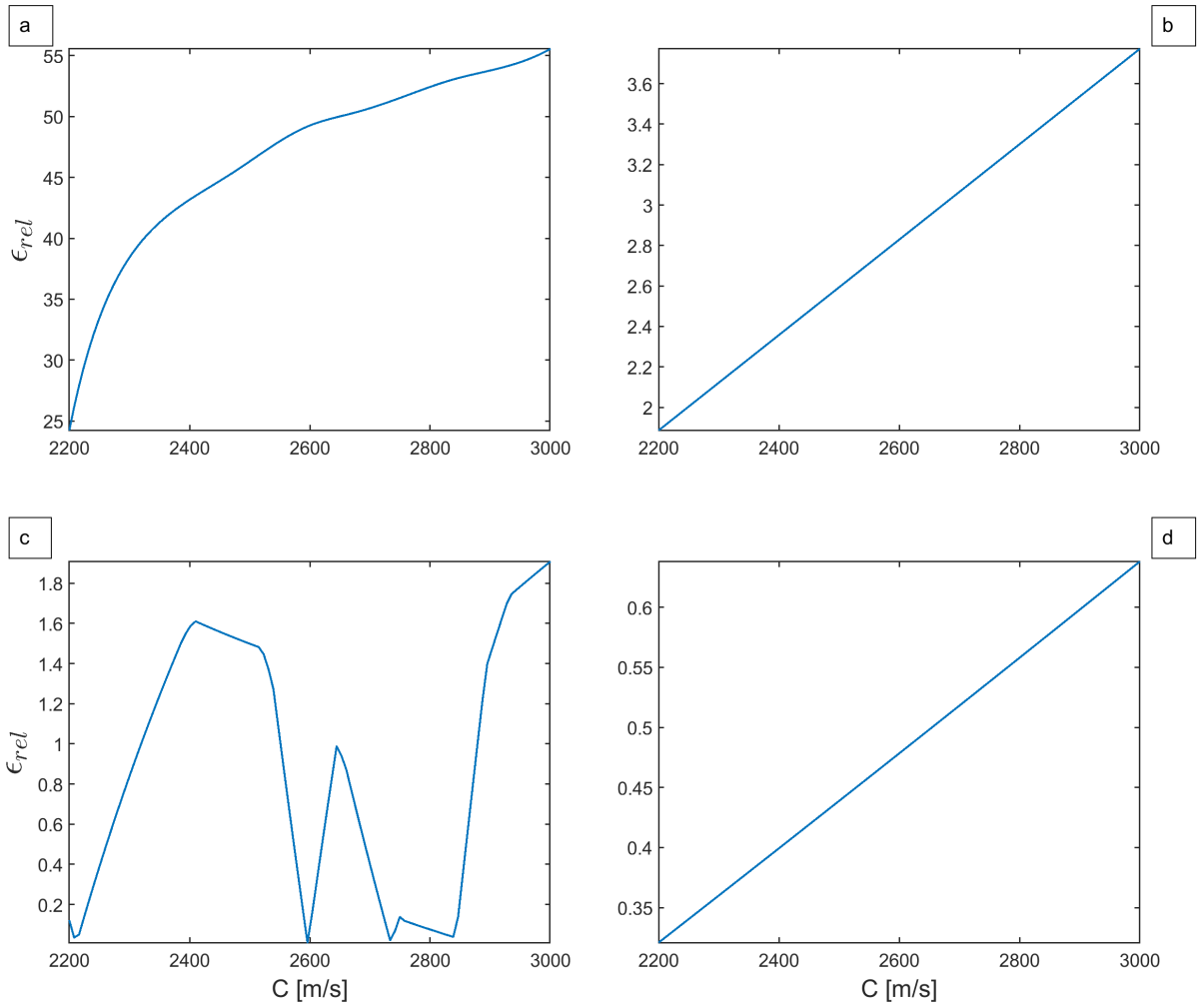

Figure 2.8: Relative errors for the Born approximation and T-matrix modelling methods in the slab models in figure 2.7. (a) and (c) correspond to the amplitude-related relative errors (the Born and T-matrix methods respectively); (b) and (d) correspond to the travel times-related relative errors (the Born and T-matrix methods respectively).

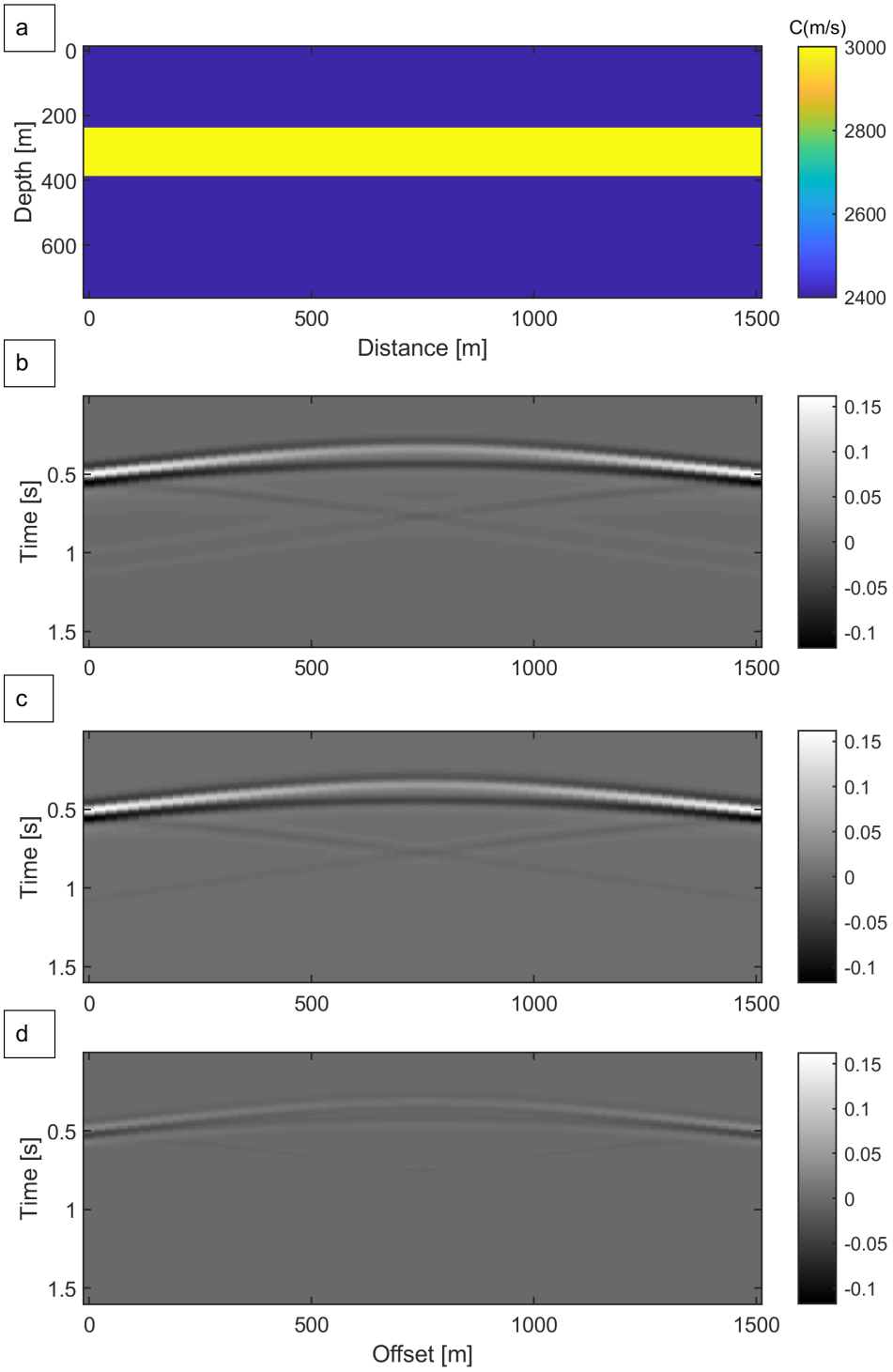

Figure 2.9: Multichannel modelling experiment. (a) the single-slab model; (b) the seismogram generated using the FDTD modelling method; (c) the seismogram generated using the T-matrix modelling method; (d) the difference between (b) and (c).

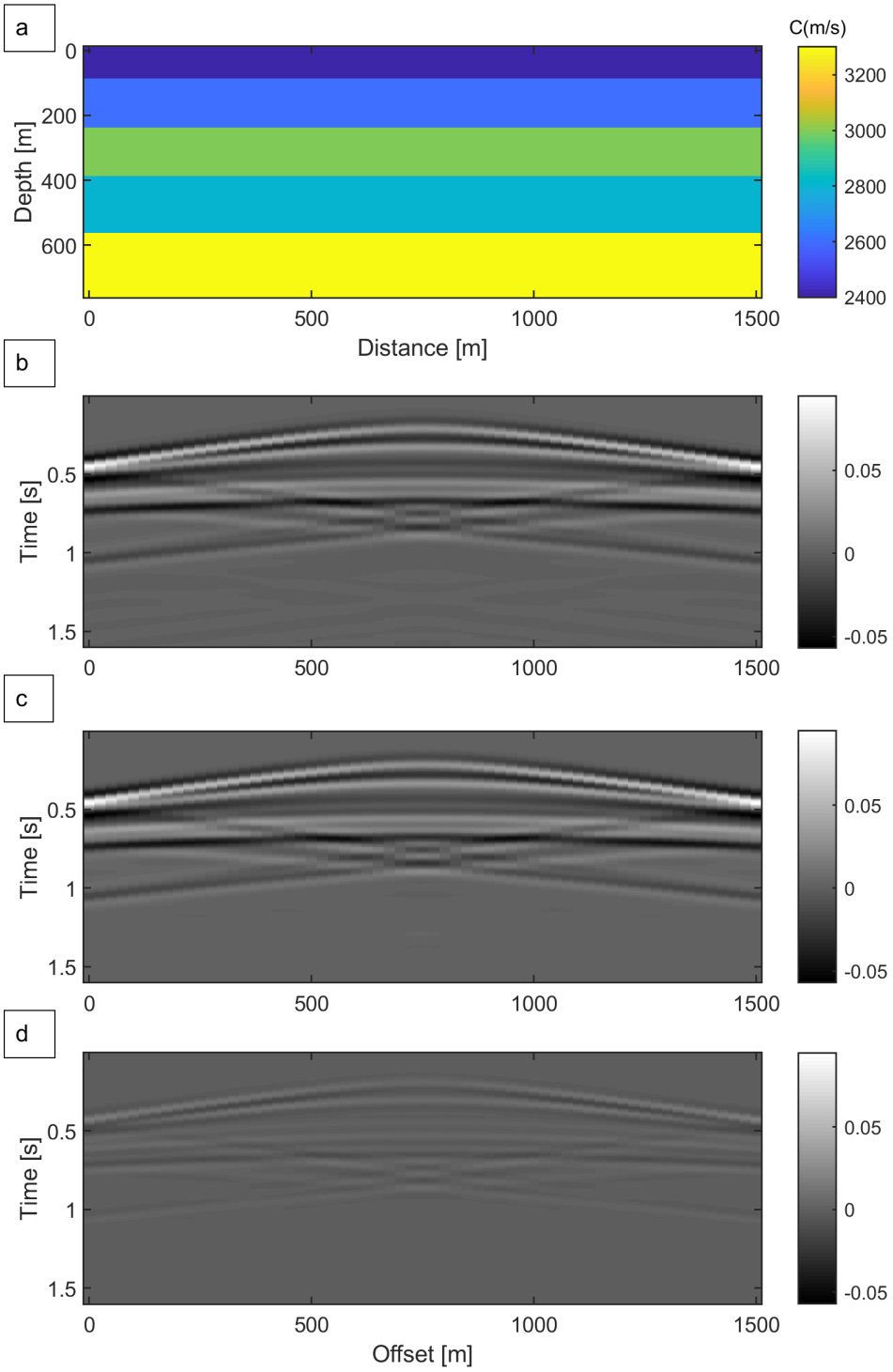

Figure 2.10: Multichannel modelling experiment. (a) the "layer-cake" model; (b) the seismogram generated using the FDTD modelling method; (c) the seismogram generated using the T-matrix modelling method; (d) the difference between (b) and (c).

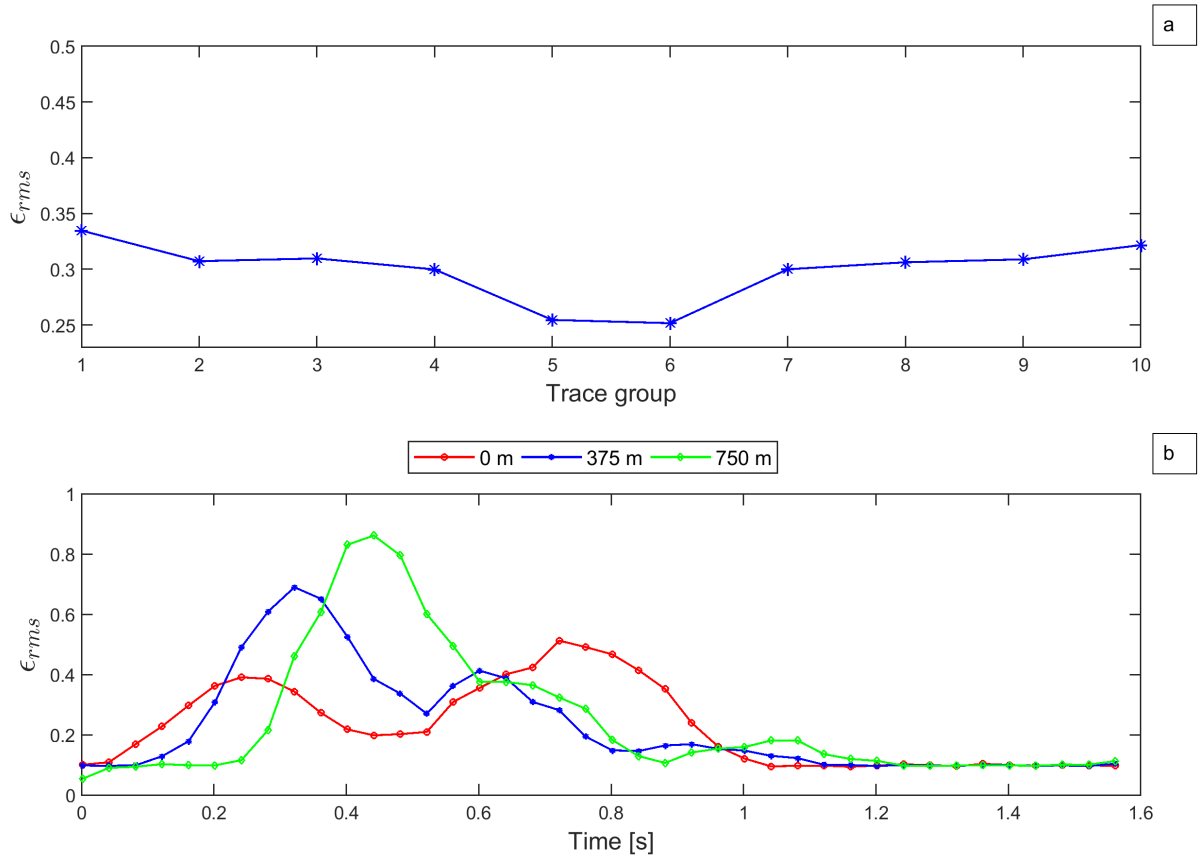

Figure 2.11: RMS errors associated with the "layer-cake" model (based on the T-matrix modelling method). (a) The offset-dependent RMS errors (each trace group includes six traces); (b) the time-dependent RMS errors at different offsets (each time group includes ten time samples).

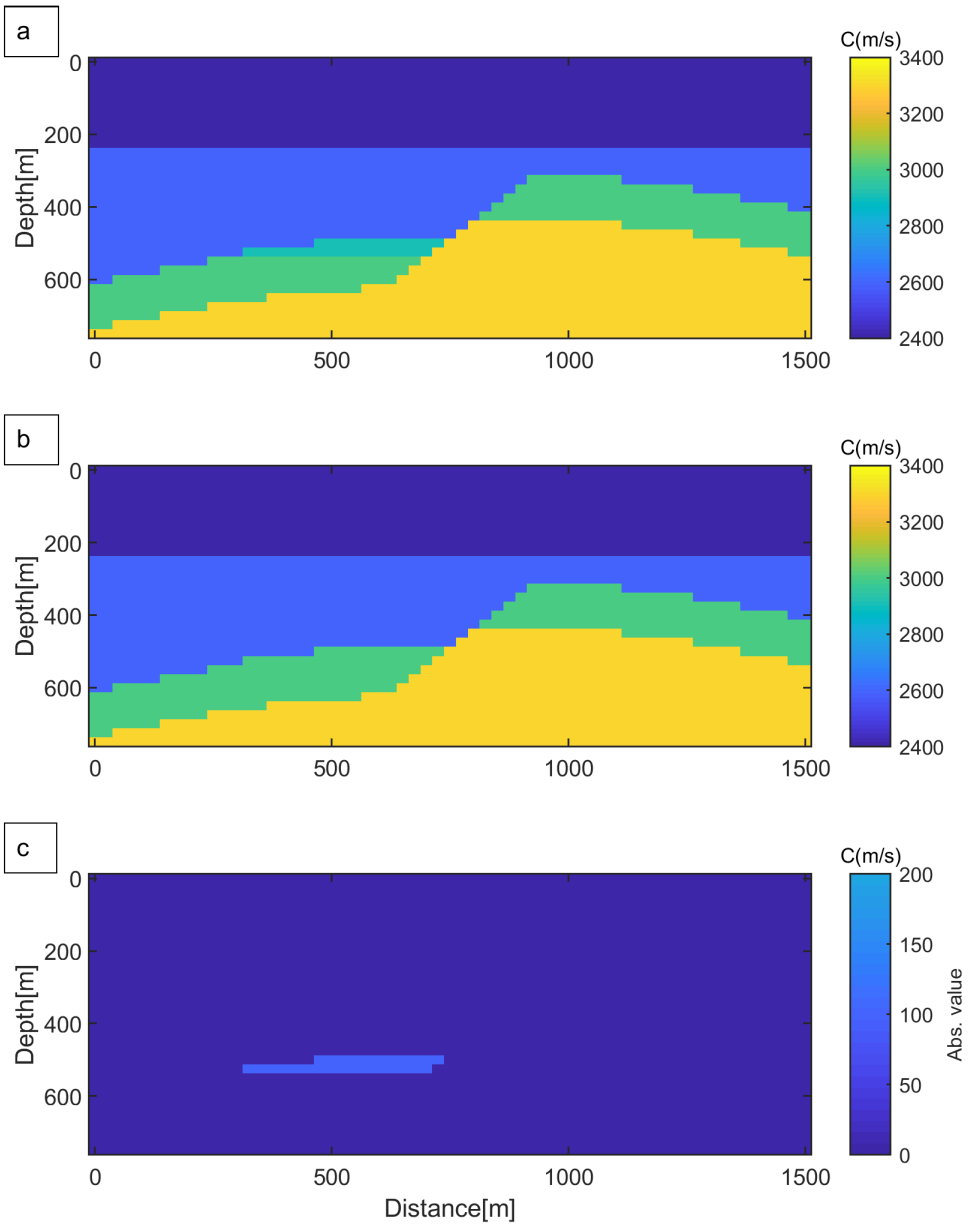

Figure 2.12: Fault trap model. (a) The baseline model; (b) the monitor model; (c) the time-lapse model (the difference between (a) and (b)).

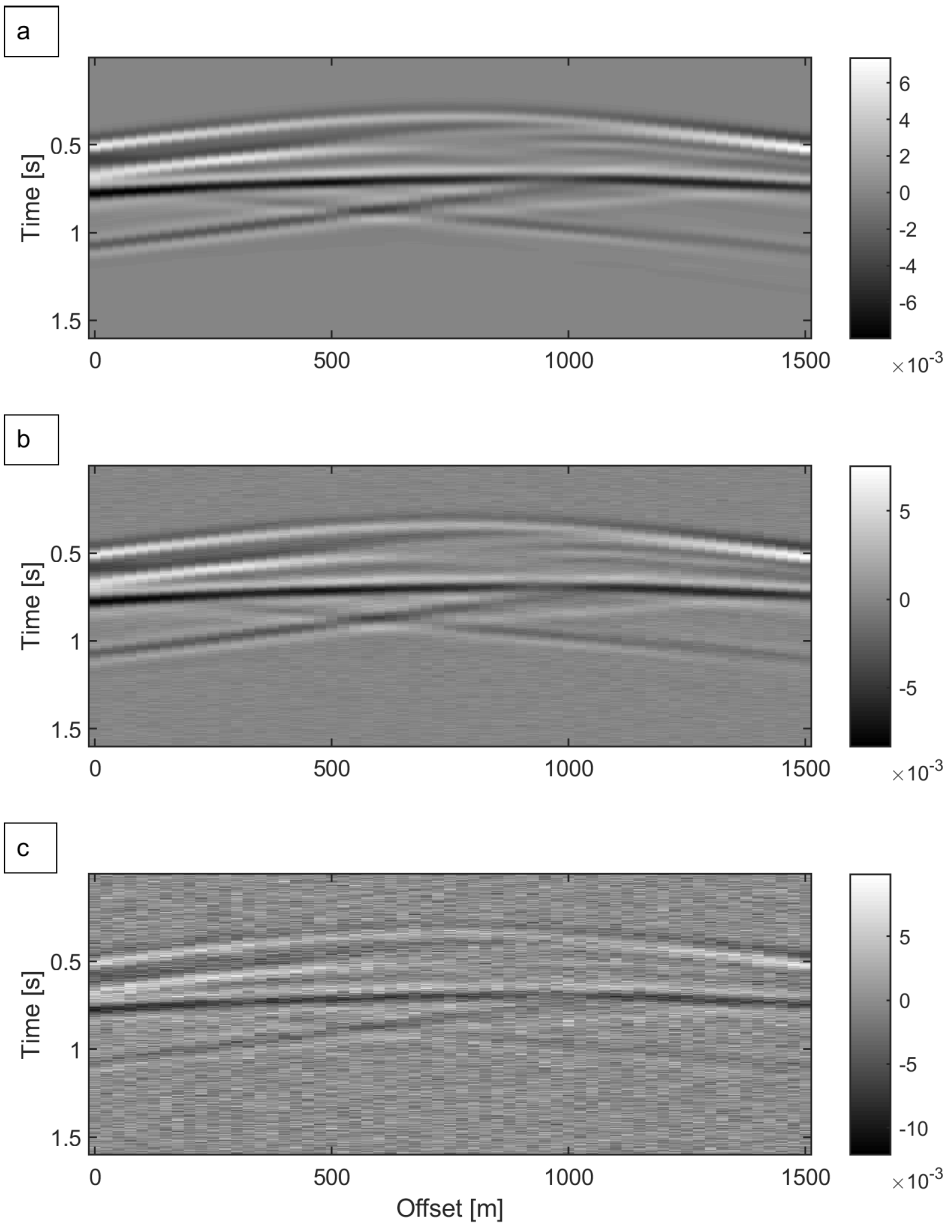

Figure 2.13: Synthetic seismograms generated from the baseline trap model using the T-matrix modelling method. (a) the clean seismogram (no noise added); (b) the noisy seismogram with the SNR equal to 20 dB (corresponds to 31% noise); (c) the noisy seismogram with the SNR equal to 5 dB (corresponds to 56 % noise).

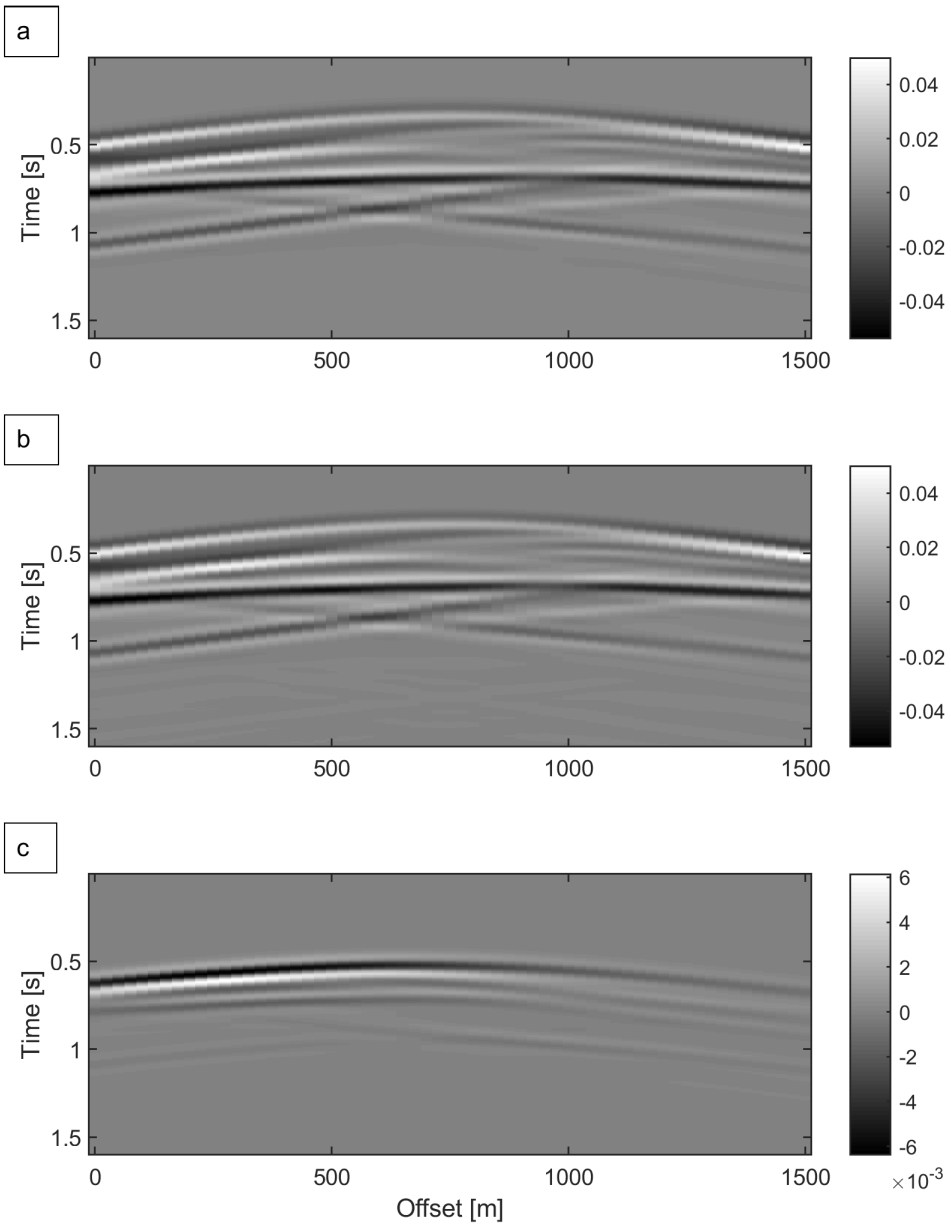

Figure 2.14: Synthetic seismograms generated from the models in figure 2.12 using the FDTD modelling method. (a) The baseline model seismogram; (b) the monitor model seismogram; (c) the seismogram difference (associated with the time-lapse model).

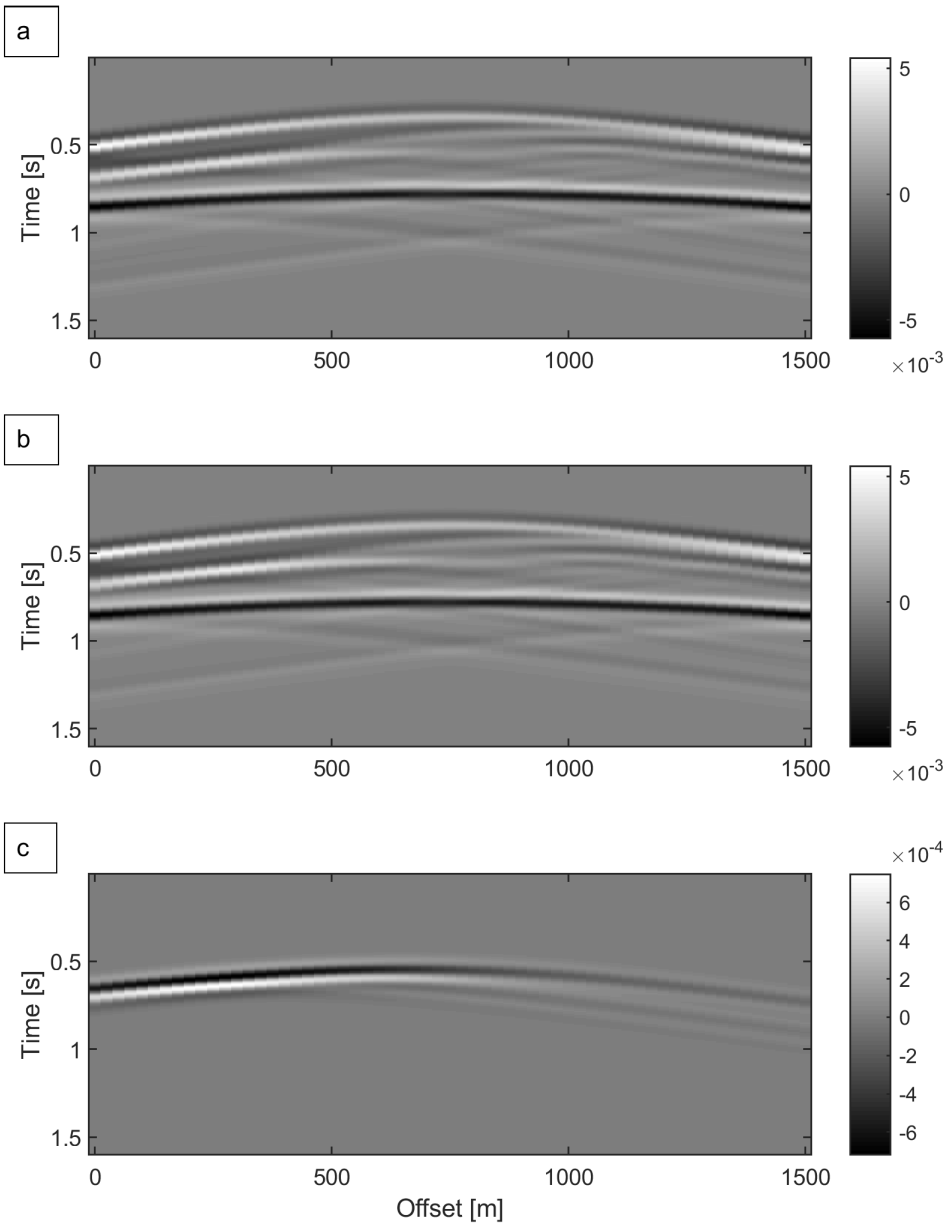

Figure 2.15: Synthetic seismograms generated from the models in figure 2.12 using the Born approximation. (a) The baseline model seismogram; (b) the monitor model seismogram; (c) the seismogram difference (assocated with the time-lapse model).

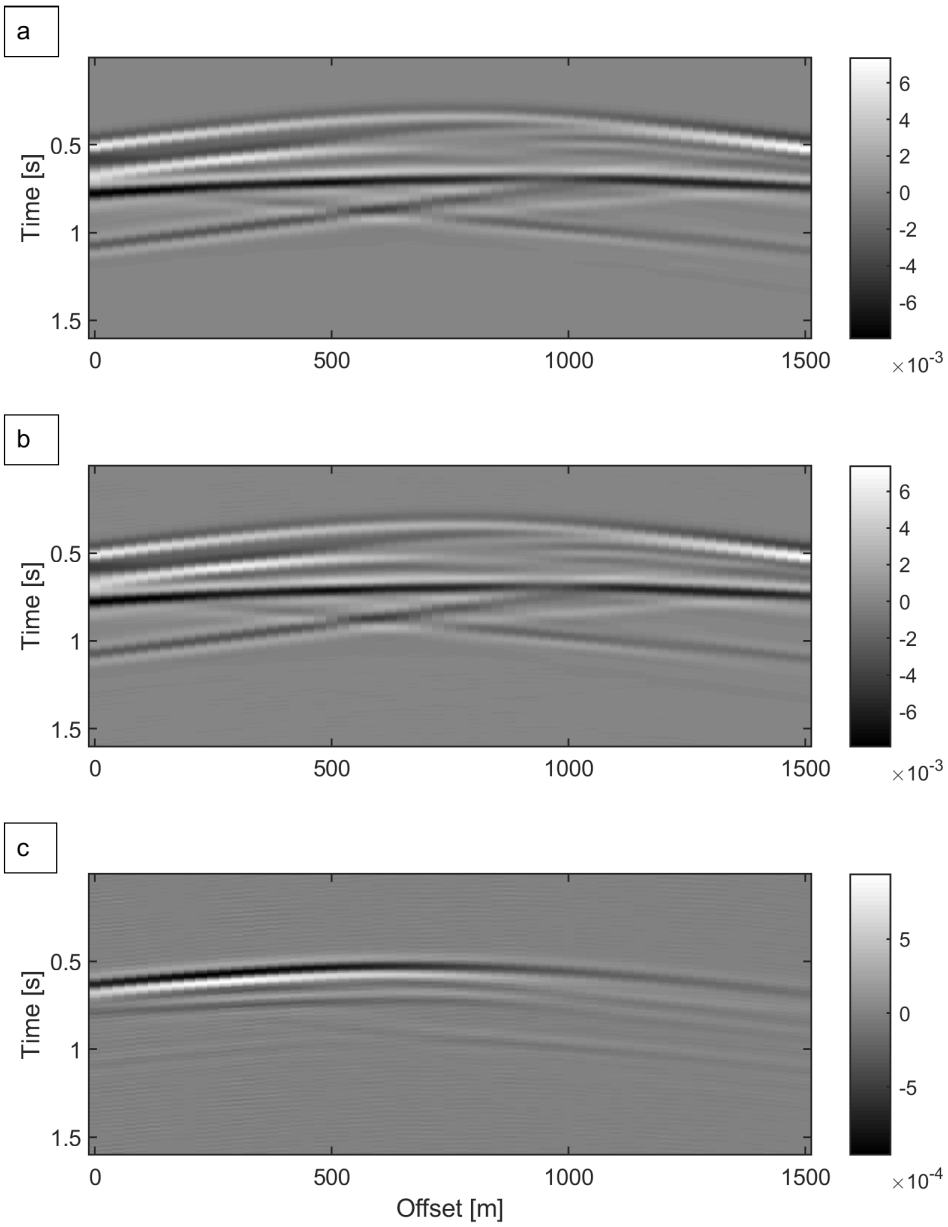

Figure 2.16: Synthetic seismograms generated from the models in figure 2.12 using the DBA modelling method. (a) The baseline model seismogram; (b) the monitor model seismogram; (c) the seismogram difference (associated with the time-lapse model).

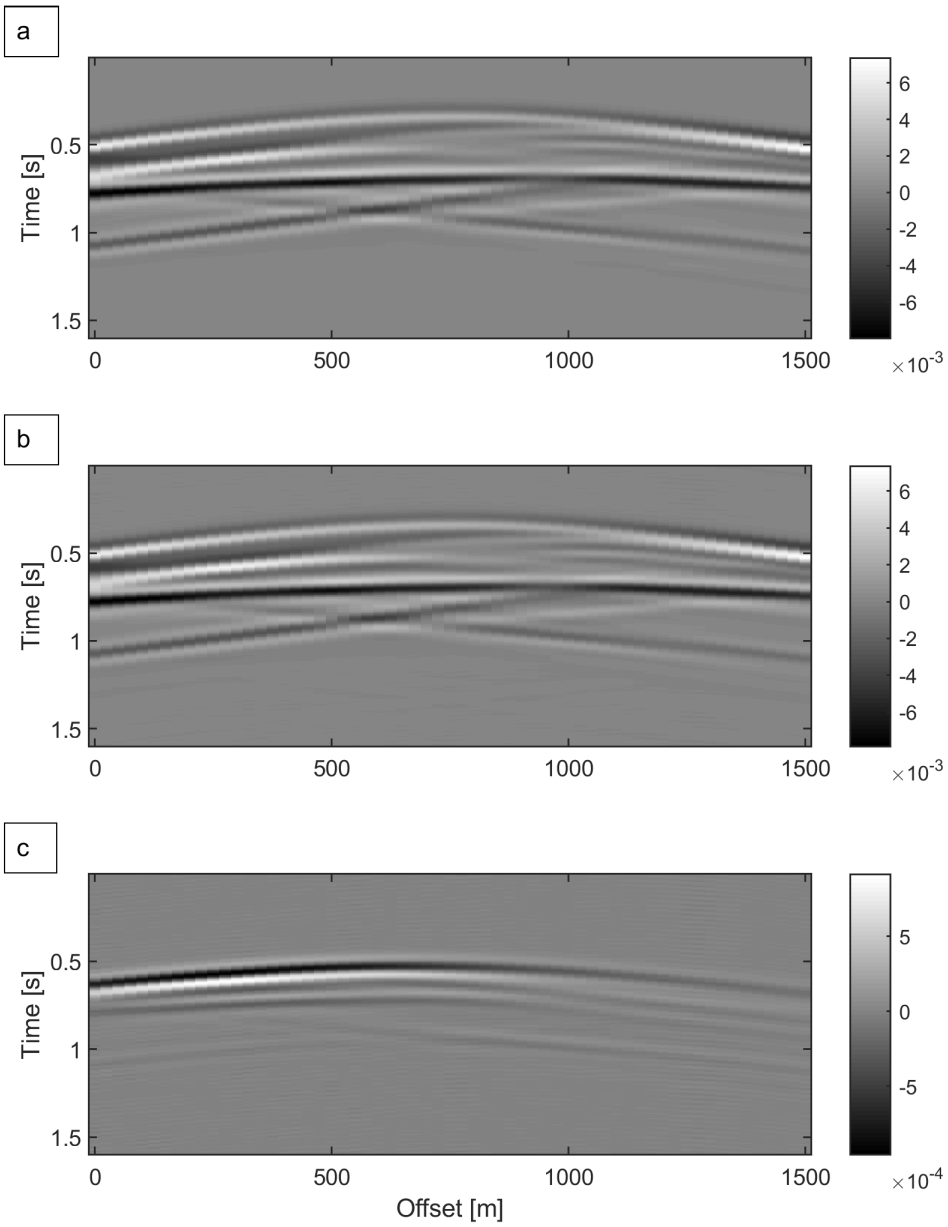

Figure 2.17: Synthetic seismograms generated from the models in figure 2.12 using the T-matrix modelling method. (a) The baseline model seismogram; (b) the monitor model seismogram; (c) the seismogram difference (associated with the time-lapse model).

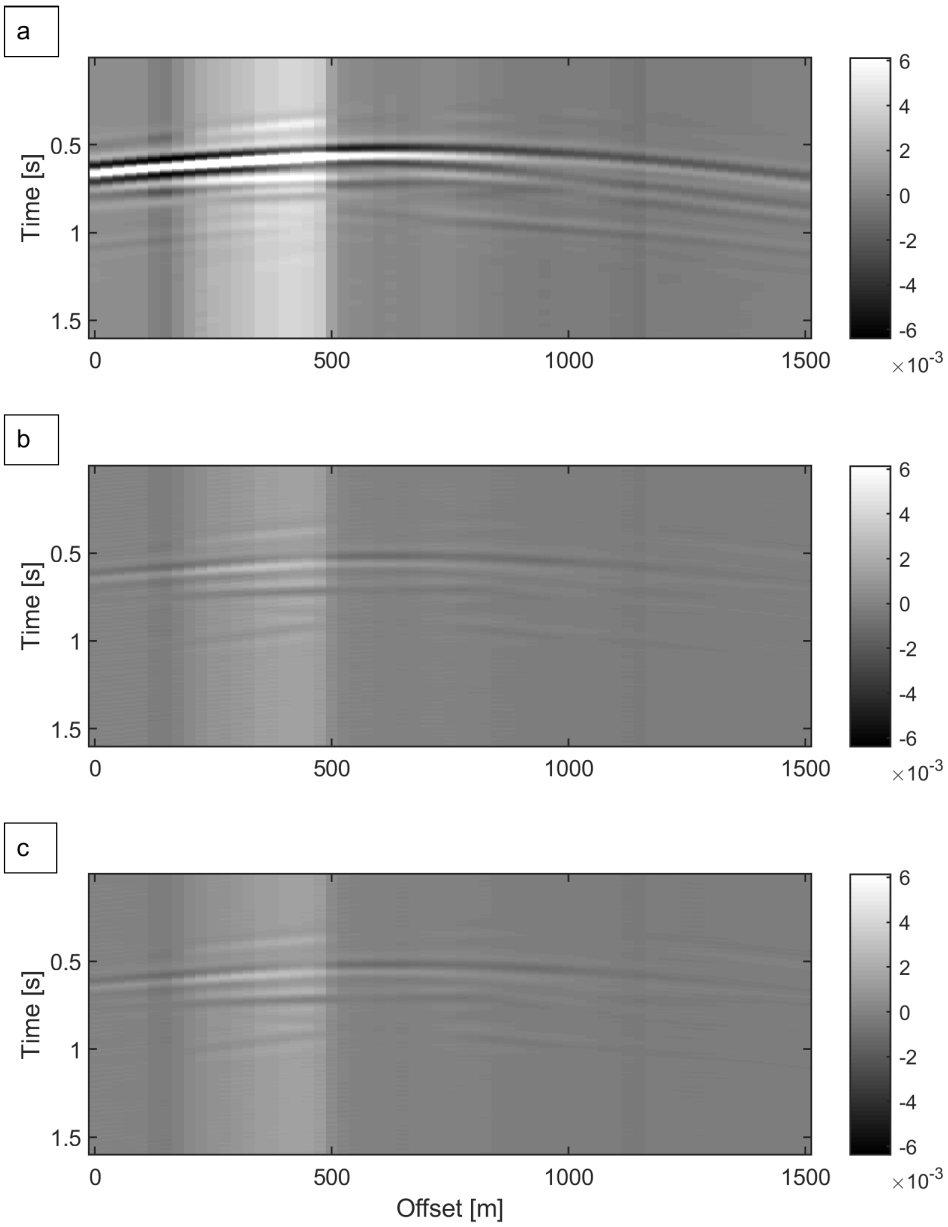

Figure 2.18: Differences between the time-lapse seismogram generated using the FDTD method and (a) the Born approximation, (b) the DBA method, (c) the T-matrix modelling method.

Chapter 3

Full waveform inversion

Everything should be made as simple as possible, but no simpler.

—Albert Einstein

3.1 Introduction

Seismic inversion is a process of converting seismic data into earth properties (Schuster, 2017). Several types of seismic inversion are recognized (Schuster, 2017). The phase inversion of travel time tomography intends to invert the picked travel times in seismic data for smoothly varying model velocities. This type of inversion is quite robust, but it is only valid for high-frequency approximations (Schuster, 2017). Another type is AVO (Amplitude Versus Offset) inversion. The combination of the AVO analysis (Castagna and Backus, 1993) and linear or nonlinear inversion theories (see, for example, Menke (2012), Aster et al. (2005)) makes it possible to detect fluid saturated zones and predict lithology (Hansen, 1992). In this case, for seismic data to be inverted, some preprocessing procedures are required (Asnaashari et al., 2015). One of the most well known types is FWI inversion, which is the main focus of this thesis.

In petroleum seismology, FWI is a model-based technique that aims to invert full seismic waveforms for the model parameters that produce these data (Ramírez and Lewis, 2010; Queißer and Singh, 2013). Model parameters estimation considers all the possible propagation effects present in seismic data: reflections, interbed multiples, converted waves, refractions, guided waves and others (Queißer and Singh, 2013).

Ikelle and Amundsen (2005) highlight the following issues that can be encountered when inverting for model parameters using FWI: (1) uniqueness problem (how one can be sure that the estimated model is the only one that fits the seismic data); (2) instability problem (in sense that small errors in data may result in considerable inconsistencies in the model estimate); (3) algorithm convergence problem (this issue relates to iterative inversion methods); (4) errors arising from inaccurate physical models; and, (5) the cost of the forward modelling step.

Seismic FWI, in principle, is a highly non-linear process. However, for models in which velocity contrast does not change significantly, the non-linear inversion problem can be replaced by a linearized version (Born inversion). Such approximation significantly simplifies the process of model parameters estimation.

In section 3.2, first, the main aspects of inversion theory are reviewed. Then, in section 3.3, Born inversion (the FWI method that is applicable for solving low-contrast perturbation problems) is introduced. Further, using a non-linear inversion approach, I present the Distorted Born Iterative T-matrix (DBIT) FWI method and demonstrate its application on time-lapse model examples (section 3.4). Finally, in section 3.5, the DBIT-based FWI inversion algorithm integrated with the smoothing and self-adaptive techniques as applied to high-contrast media models is discussed. As part of the study, I compare the different time-lapse inversion approaches (parallel difference and sequential) and analyze their effectiveness when applied to high-contrast media models.

3.2 Some aspects of inversion theory

3.2.1 Formulation of the inverse problem

The forward modelling problem in matrix (discretized) form is given as (Menke, 2012):

$$\mathbf{d} = \mathbf{J}(\mathbf{m}), \tag{3.1}$$

where $\mathbf{d}$ is the data-vector, $\mathbf{J}$ is the sensitivity matrix, and $\mathbf{m}$ is the model parameter. Both $\mathbf{d}$ and $\mathbf{m}$ vectors have lengths N and M, respectively. $\mathbf{J}$ is a N x M in size matrix (Menke, 2012). In the linear case, the forward modelling problem in equation (3.1) can be written as (Schuster, 2017)

$$\mathbf{d} = \mathbf{J}\mathbf{m}. \tag{3.2}$$

In a real situation, data always contain some noise. A "noisy" forward linear modelling problem can be expressed in the following way (Menke, 2012):

$$\mathbf{d} = \mathbf{Jm} + \boldsymbol{\mu}, \tag{3.3}$$

where $\boldsymbol{\mu}$ is an N-dimensional vector representing noise.

The solution to the inverse problem involves finding model parameters $\mathbf{m}$ such that equations (3.1), or (3.2), (3.3) is satisfied. Clearly, in these equations, an unique inverse solutions may not exist (if data contain less measurements than required to define all model parameters). In this situation, it is common to speak about the inverse solution that provides the best approximate model parameters (an estimate of model parameters) $\mathbf{m}_{est}$. One of the ways to obtain such approximate model parameters is by using the least squares method.

3.2.2 Least squares solution to the inverse problem

The simplest methods to solve a linear inverse problem involves finding such $\mathbf{m}_{est}$ that minimizes the misfit (residual) measure (in the least squares manner) from differences between the observed and predicted data, is defined by (Aster et al., 2005)

$$||\mathbf{d}_{obs} - \mathbf{Jm}_{est}||_2 = \sqrt{\sum_{i=1}^{N}(d_i - (\mathbf{Jm}_{est})_i)^2}. \tag{3.4}$$

The overall error E, given as the squared Euclidean norm (L_2 norm) of the residual vector (Menke, 2012) is then

$$E = (\mathbf{d}_{obs} - \mathbf{Jm}_{est})^{\mathrm{T}}(\mathbf{d}_{obs} - \mathbf{Jm}_{est}). \tag{3.5}$$

Finally, the model parameters $\mathbf{m}$ can be found from (see full derivation in Wang (2016), for example)

$$\mathbf{m}_{est} = [\mathbf{J}^{\mathrm{T}}\mathbf{J}]^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{d}_{obs}, \tag{3.6}$$

where, $\mathbf{J}^{\mathrm{T}}$ is the transpose of the sensitivity matrix $\mathbf{J}$.

3.2.3 Well-posed and ill-posed inverse problems

In applied mathematics, two types of inversion problems are recognized: well-posed and ill-posed (**?**). A well-posed inverse problem satisfies the following characteristics: it has a

solution; this solution is unique; and, finally, the solution "continuously" depends on data. If it does not, an inversion problem is known as ill-posed (Bleistein et al., 2001). From the above definition, all geophysical problems should be considered as ill-posed problems. Ill-posed inversion problems are also classified into underdetermined, overdetermined and mixed-determined subtypes (Menke, 2012). Mathematically, an undetermined inverse problem in equation (3.2) satisfies the inequality M > N (more unknowns than data). When there are more data than unknowns (M < N), the inverse problem is known as overdetermined. The mixed-determined situation occurs when, for some reason, some part of a model is overdetermined, and another is underdetermined (Menke, 2012). From this formulation, full waveform seismic inversion is a typical case of the mixed-determined problem: at shallow depths, the problem is usually overdetermined (there is more than enough information); at larger depths, on the contrary, the problem is undetermined (information deficiency).

3.2.4 Regularization

Prior information

In the case of purely undetermined inverse problem, there may exist several model estimates for a given data set for which the predicted error E equals to zero. Among all these solutions, only one corresponds to the true model. The estimation of this specific solution is a key problem in inversion theory. One of the options for estimating such an unique solution is to make the system of linear equations in (3.2) consistent, by adding some *a priori information* (Menke, 2012). *A priori information* is the information that is based on model-related logical expectations (Menke, 2012). At first sight, this formulation may seem a bit vague: for what is meant by logical expectations? To clear this up, consider a very simple example: suppose, after seismic data processing at some particular site the travel times (t) and the P-velocities (V) in the time window 0-2.5 s are related via the following cubic equation: $t = 0.033V^3 + 0.075V^2 + 0.0375V + 0.0063$. Assume, that one wants to determine velocities somewhere within the given time interval. After organizing travel times into the vector $\mathbf{t}$, and defining the sensitivity matrix $\mathbf{J}$, the inverse problem can easily be solved. Now, imagine, there is a need to estimate the P-velocities at the time interval 2.5-3 s. There are several guesses then could be made about the velocities in this interval based on some logical expectations. Firstly, velocities must be positive (this is obvious); secondly, they should probably increase with depths (this is true in many practical cases); finally, they may satisfy some simple linearly-increasing depth de-

pendence. These are the logical guesses about the probable velocity profile within the extended time interval. Given *a priori information*, then, the variety of possible solutions to the underdetermined inverse problem within the defined time interval (with the help of the guesses above) is significantly reduced. This minimizes the chance for the inversion algorithm to pick inappropriate $\mathbf{m}_{est}$.

Tikhonov regularization

The process of adding *a priori information* to the system of equations in (3.2) to make it consistent, is known as regularization (Wang, 2016). One of the widely used and simple techniques to deal with a discrete ill-posed problem is Tikhonov regularization (Aster et al., 2005). This technique considers the minimization of an objective function by solving the regularized least squares problem of the following form:

$$E(\mathbf{m}) = ||\mathbf{d}_{obs} - \mathbf{Jm}||_2^2 + \lambda^2 ||\mathbf{Lm}||_2^2. \tag{3.7}$$

Here, the first part of the equation is the residual norm discussed in section 3.2.2 (L2 norm). The second part consist of a product of the squared regularization parameter λ with the squared L_2 norm of the penalty term $\mathbf{Lm}$. The matrix $\mathbf{L}$ is a nth order derivative operator which is responsible for choosing the desired model smoothness. A choice of the zeroth order derivative operator (applied in this thesis) in equation (3.7) to minimize the penalty term leads to a normal equation (Zhdanov, 2015; Menke, 2012; Schuster, 2017; Wang, 2016) that is known as the regularized damped least squares solution

$$\mathbf{m}_{est} = [\mathbf{J}^{\mathrm{T}}\mathbf{J} + \lambda^2\mathbf{I}]^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{d}_{obs}, \tag{3.8}$$

with $\mathbf{I}$ denoting the identity matrix.

Selecting the regularization parameter

One of the methods for choosing a value for the regularization parameter is the Morozov discrepancy technique (Schuster, 2017). This technique involves assigning some value for the regularization parameter, inverting for the estimated model $\mathbf{m}_{est}$ using the modelled data $\mathbf{d}_{obs}$ and computing of the estimated data $\mathbf{d}_{est}$ (from the obtained previously estimated model). When we display λ^2 against $||\mathbf{d}_{est} - \mathbf{d}_{obs}||$ we can associate the regularization parameter λ to the value $||\mathbf{d}_{est} - \mathbf{d}_{obs}||$, which is equal to the estimated residual norm in the actual data (Schuster, 2017). The more sophisticated version of the discrepancy principle developed by Constable et al. (1987) makes it possible to solve linear and

nonlinear inversion problems based on iterative methods. Integrated with the cooling scheme introduced by Farquharson and Oldenburg (2004) or the self-adaptive technique (both will be discussed later in sections 3.4.2 and 3.5.5 respectively) presented by Ciric and Qin (1997), this principle works quite well for complex inverse problems.

Hansen (1992), suggested another approach for choosing the optimal λ, the so-called L-curve method. This involves constructing a log-log plot of the lengths of the model parameters versus the misfit vectors for different choices of the damping parameter. The maximum curvature point in such plot corresponds to an optimal value of the regularization parameter (Schuster, 2017).

3.2.5 Iterative methods

The need for iterative methods in the case of seismic FWI arises from the necessity to solve large linear equations. The extended explanations of some iterative methods used in FWI inversion algorithms can be found in Aster et al. (2005); Schuster (2017). The main principle behind such methods can be explained as follows. Consider a linear system $\mathbf{Ax} = \mathbf{b}$. The idea involves converting the linear system $\mathbf{Ax} = \mathbf{b}$ into the equivalent system in form $\mathbf{x} = \mathbf{Mx} + \mathbf{v}$ for some fixed matrix $\mathbf{M}$ and vector $\mathbf{v}$. After the initial vector $\mathbf{x}^0$ is being selected, a sequence of approximate solutions can be obtained by solving $\mathbf{x}^{k+1} = \mathbf{Mx}^k + \mathbf{v}$ for each $k = 0, 1, 2, ...$ (Vrahatis et al., 2003). Suppose some linear system comprised of hundreds or thousands of equations (as in case of FWI). In that situation, a choice for iterative methods is preferable for two main reasons: it takes less computation time and requires less computer memory. Additionally, if the accuracy of the computations is not an issue, the acceptable solution can be gained by reducing the number of iterations.

3.2.6 Nonlinear inversion

The discussion above has been mainly concerned about linear inverse problems. In FWI, however, all inversion problems are typically nonlinear due to the structural complexity of the Earth's subsurface. Therefore, it is also important to consider an inversion problem from a nonlinear perspective.

In practice, nonlinear inverse problem is often solved by approximating nonlinearity by

a number of linear inverse problems. Therefore, the computational principle remains the same. It is just sequentially applied to each linear problem involved. Iterative methods mentioned above are often the most convenient way to solve the associated systems of linear equations (Norton, 1988).

Any iterative scheme in the nonlinear inversion algorithm is designed to minimize the mean-square difference between the observed and synthetic data at each iteration. To start a new iteration, one requires information about the rate of change, or gradient, of the data with respect to the model (to know in which "direction" the inversion algorithm should adjust generated and observed data). In the scattering-based inversion methods, the gradient derivation follows from the assumption that the current Green function is computed on the basis of the previous model estimate with very small difference between them. At the next iteration, the previous model estimate becomes the background model, and then the Green function and the associated gradient are recalculated according to the new estimate. This iterative procedure is repeated until the mean-square error is minimized (Norton, 1988).

A practical problem of nonlinear inversion-based iterative schemes is the existence of local minima. If the appropriate initial model (*a priori information*) to launch the iterative process is not available, or data from which the model needs to be recovered are too noisy, the number of solutions satisfying the inversion problems increases. These false solutions become the objective function local minima, at which the inversion algorithm could stall causing to choose inappropriate model parameters (Mulder and Plessix, 2008). There are different approaches to deal with this problem (see, for example, Bunks et al. (1995), Chi et al. (2014)), and some of them are discussed later.

3.2.7 The inverse crime

Any new developed FWI algorithm requires thorough testing before being implemented in practice. Colton and Rainer (1998), when discussing the development of new inversion methods note that "for the inverse problem it is crucial that the synthetic data be obtained by a forward solver do not have any connection to the inverse solver". If a new inverse approach does not satisfy this simple requirement, the inverse solution is a subject of the "inverse crime". And, as Colton and Rainer (1998) point out, "not all of the numerical reconstructions which have appeared in the literature meet with this obvious requirement". Put another way, any inconsistencies in synthetic data caused by

modelling are compensated by the "imperfectness" of the inverse algorithm. This compensation is the inverse crime. One of the ways to minimize the inverse crime problem is to show that the forward modelling technique used to generate data provides an exact modelling solution.

3.2.8 Time vs frequency domain inversion

FWI inversion can be applied in both the time (Tarantola, 1986; Bunks et al., 1995; Shipp and Singh, 2002) and frequency (Pratt, 1990; Sirgue and Pratt, 2004) domains. When doing FWI in the frequency domain, it is important to note that not all frequency components in a signal contribute to model recovery (the most seismic energy is typically concentrated within the limited frequency interval located at the first tens of Hertz. Motivated by this fact, several frequency domain inversion studies that used limited frequencies have been performed (Pratt et al., 1996). The main question was the accuracy of such techniques (Bansal and Sen, 2010). Freudenreich et al. (2001) speculated that the frequency-limited approach required some specific strategy for selecting frequencies when inverting. Moreover, the authors, after some research, concluded that the frequency domain approach was not consistent with data acquired within limited offsets. Sirgue and Pratt (2004) argued this point and introduced the technique that made it possible with only a few frequencies (adequately selected) within limited offsets, to recover models successfully. When offsets are not limited, the frequency domain inversion problem can be reduced to the use of two main inversion techniques: simultaneous and sequential (Sirgue and Pratt, 2004). In the first case, all selected frequencies are inverted at one time. This approach gives more robust results when data are noisy. At the same time, this technique is costly, more sensitive to the choice of the initial model and suffers from the cycle-skipping problem (Jakobsen and Ursin, 2015). In the case of the sequential approach (also known as the frequency-hopping approach), one starts from the lowest frequency available in seismic data, then inverts for the next frequency and, so on, until all the selected frequencies are processed. To increase consistency of the results, inversion could be performed on frequency groups (Jakobsen and Ursin, 2015). Varying the number of frequencies in the group, it is possible to compromise between computation time and inversion stability.

3.3 Born inversion and its application

Born inversion is a linear FWI method. As discussed in subsection 2.2.4, the linearized modelling problem can be represented in the form of equation (3.2). Since a seismic FWI problem is typically mix-determined (subsection 3.2.3), a regularized solution based on Born inversion can be obtained from equation (3.8). Solving the Born inversion problem iteratively (subsection 3.2.5) for each selected frequency using the sequential frequency technique (subsection 3.2.8), we then obtain the required P-velocities by rearranging equation (2.20). All the computations involved are quite straightforward and follow directly from the inversion theory introduced above. Since Born inversion can be considered as a special case of nonlinear inversion, for now, I skip the Born implementation aspects (not to be repeated in derivation) and will refer to them later (section 3.4) explaining how they can be deduced from the DBIT inversion scheme.

From a practical point of view, the use of Born inversion is limited to weak perturbation models (Eikrem et al., 2016). Remember the criterion under which the Born approximation has been considered to be valid (subsection 2.2.3). For Born inversion to work, the data to which this linearized FWI method is applied, must be consistent with this validity criterion as well. If it holds, the expected results should be close to those obtained by nonlinear inversion methods, making Born inversion even more desirable due to reduced computation cost.

The application of Born inversion could be valuable, for example, in the reservoir monitoring. In this situation, the baseline velocity model is known. If so, the monitor model with the small velocity variation in the reservoir (consistent with the validity criterion mentioned) can be easily determined.

In what follows, a few examples illustrating the Born inversion method will be presented. To back up the inversion theory discussed above, firstly, the use of sequential frequency approach is demonstrated. Then, an example of the ill-posed inversion problem and ways to reduce ill-posedness are illustrated. Finally, several cases exemplifying model recoveries from data with different noises are provided (including recovery improvements when manipulating with the number of frequencies, receivers and shots).

All inversion examples, for simplicity reasons, are performed on a very simple model: a homogeneous background medium with a single-box embedded (in different examples, the heterogeneity is located at different depths). For this model, the velocity in the background medium is equal to 2000 m/s; the velocity in the heterogeneity is equal to

2050 m/s. The length of the model is 400 m; depth is 200 m. The size of the heterogeneity (box) is 40 x 40 m. In the first case, the depth of the box is defined at 20 m from the surface (illustration (a) in figure 3.1 (in the following will be named model 1)), and, in the second case is at 140 m from the surface (illustration (b) in figure 3.1 (in the following will be named model 2)). The model is discretized into 21 x 11 grid blocks which are 20 m by 20 m in size. A source function is the Ricker wavelet function (equation (2.80)) with a central frequency of 7.5 Hz. A time sampling interval is 0.004 s with the total record length is equal to 1 s. The numbers of shots, receivers and frequencies selected are specified in each particular case. To model the observed data, I use the Born approximation modelling method. To simulate noisy data, I refer to equation (2.83).

Figure 3.2 demonstrates the inversion results for both models in figure 3.1 obtained with different sets of frequencies. Consider the reconstructions of model 1. With the help of one frequency (2 Hz), only a general shape of this model was recovered (illustration (a)). The application of the frequency set [2, 6, 10 Hz] resulted in much better model reconstructions (illustration (b)). Finally, when frequencies [2, 6, 10, 14, 18, 22 Hz] were used, it was possible to obtain the original model. This is how the sequential approach based on frequency domain FWI works in practice. Now, compare the recovery evolutions for both models in figure 3.1. Note how different the final results are (illustrations (d) and (e)). The unsatisfactory model reconstructions for the heterogeneity positioned at the bottom of the model is explained by its indeterminacy (subsection 3.2.3). In the last case, information deficiency could be partially fixed if we increase the number of receivers, shots or frequencies (increase the number of equations in the corresponding linear system).

Figure 3.3 is the example illustrating changes in model recoveries with respect to the different number of receivers used (the number of shots and frequencies were kept constant). For model 1, the improvements are not very obvious (the heterogeneity is well determined), but they are notable for model 2 (compare illustrations (b), (d), (f) and (h)).

Figure 3.4 shows the enhancements in model restorations when the number of shots in the experiment was varied (the number of frequencies and receivers were kept constant). First, compare the reconstructions of model 1 in illustrations (a), (c), (e) and (g). Recovery improvements (even though they are not very distinct) can be identified. In the second case, a quality of model restorations increases commensurately with the number of receivers used (compare illustrations (b), (d), (f) and (h)).

Finally, figure 3.5 shows the model recoveries in relation to the different numbers of frequencies used (the number of shots and receivers were kept constant). Note how the recoveries of model 2 change (illustrations (b), (d), (f) and (h)).

In practice, observed data always contain noise. Figure 3.6 illustrates the inversion results corresponding to model 1 for the data with different amounts of random noise added. Note that the selected set of frequencies, the number of shots and receivers used to model data did not allow us to restore models for the data with SNR 2 dB (80% noise), 4 dB (63.3% noise), 6 dB (50% noise), 12 dB (25% noise) (illustrations (a), (b), (c), (d) and (e) respectively).

Figure 3.7 shows the case when variable combinations of modelling parameters (the number of frequencies, receivers and shots) applied to the noisy data (SNR=10 dB) corresponding to model 1 were used. In illustration (a) (the initial model), the heterogeneity cannot be identified. For this example, the data were generated with the help of 5 frequencies [2, 8, 12, 16, 22 Hz], 11 equidistantly spaced receivers, and 3 equidistantly spaced shots. Illustrations (b) and (c) corresponding to the frequency sets [2, 4, 8, 10, 12, 16, 22 Hz] and [2, 4, 6, 8, 12, 14, 16, 18, 22 Hz] (the number of shots and receivers were kept constant). Note the improvements. An increase in the number of receivers (21 and 31) led to even more enhanced recovery results (illustrations (d) and (e)). Finally, an increase in the number of shots (7 and 15) resulted in almost complete model recoveries (illustrations (f) and (h)).

Note, in the examples considered, there was no need to invert for all frequencies available in the frequency band in order to restore subsurface models. This observation is entirely consistent with the theory discussed in subsection 3.2.8. Now, suppose a seismic survey involves employing many shots. Clearly, in such situation, the frequency-limited approach demonstrated above has a significant advantage over the time domain FWI.

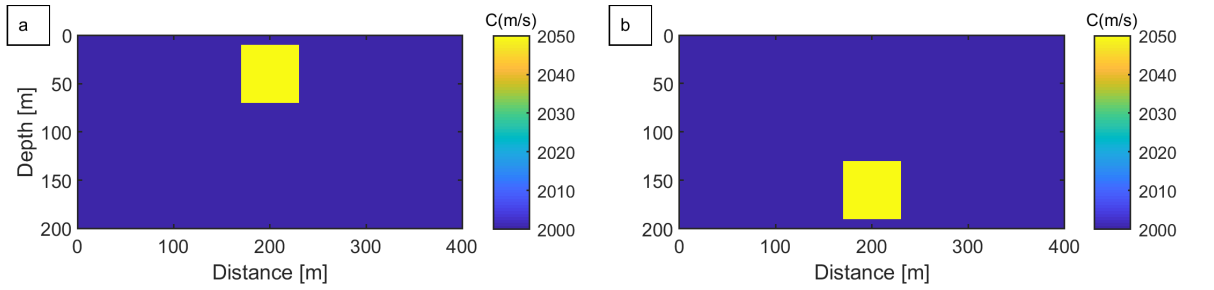

Figure 3.1: Box model. (a) Model 1 (the heterogeneity located at the top of the model); (b) Model 2 (the heterogeneity located at the bottom of the model).

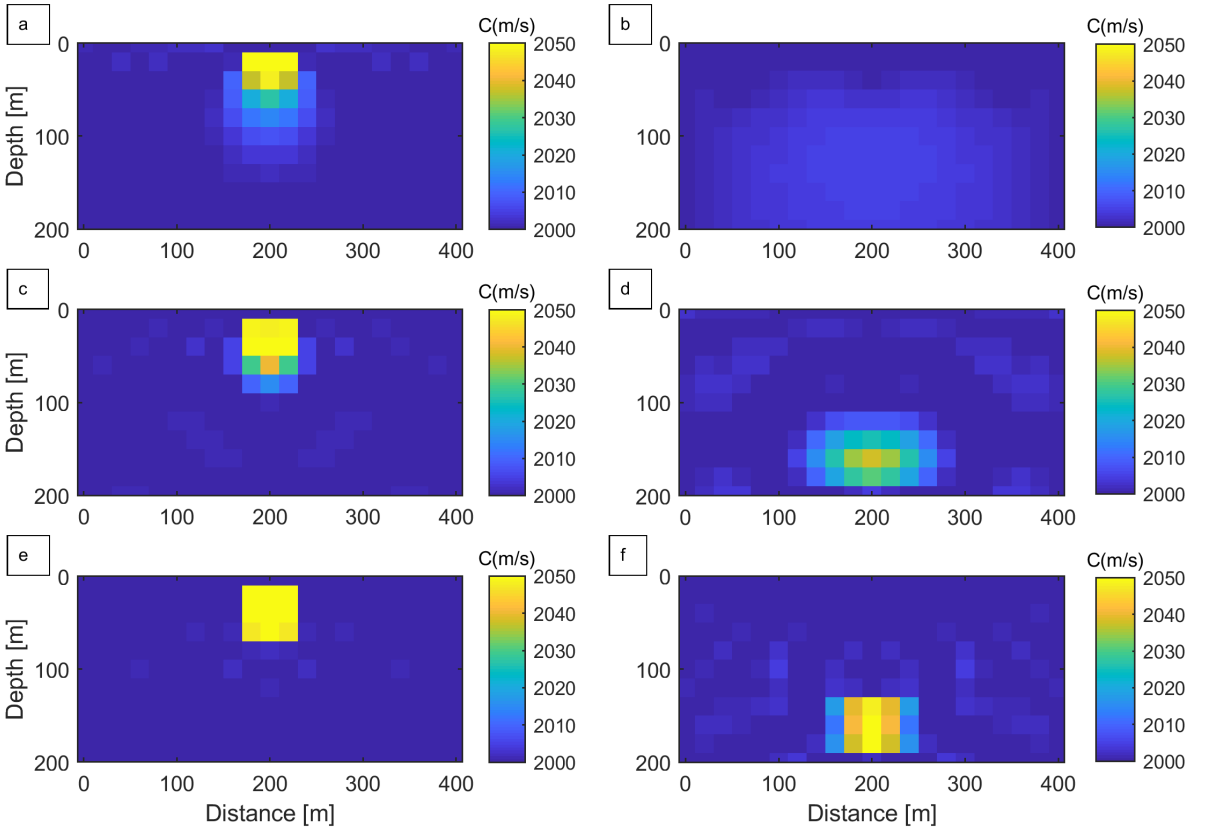

Figure 3.2: The sequential inversion approach. Models 1 (illustration (a) and 2 (illustration (b) recovered with a single frequency 2 Hz; (c) and (d) with a frequency set [2, 6, 10 Hz]; (e) and (f) with a frequency set [2, 6, 10, 14, 18, 22] Hz. For the recoveries 21 equidistantly spaced receivers and 1 shot (located at the centre of the acquisition system) were used.

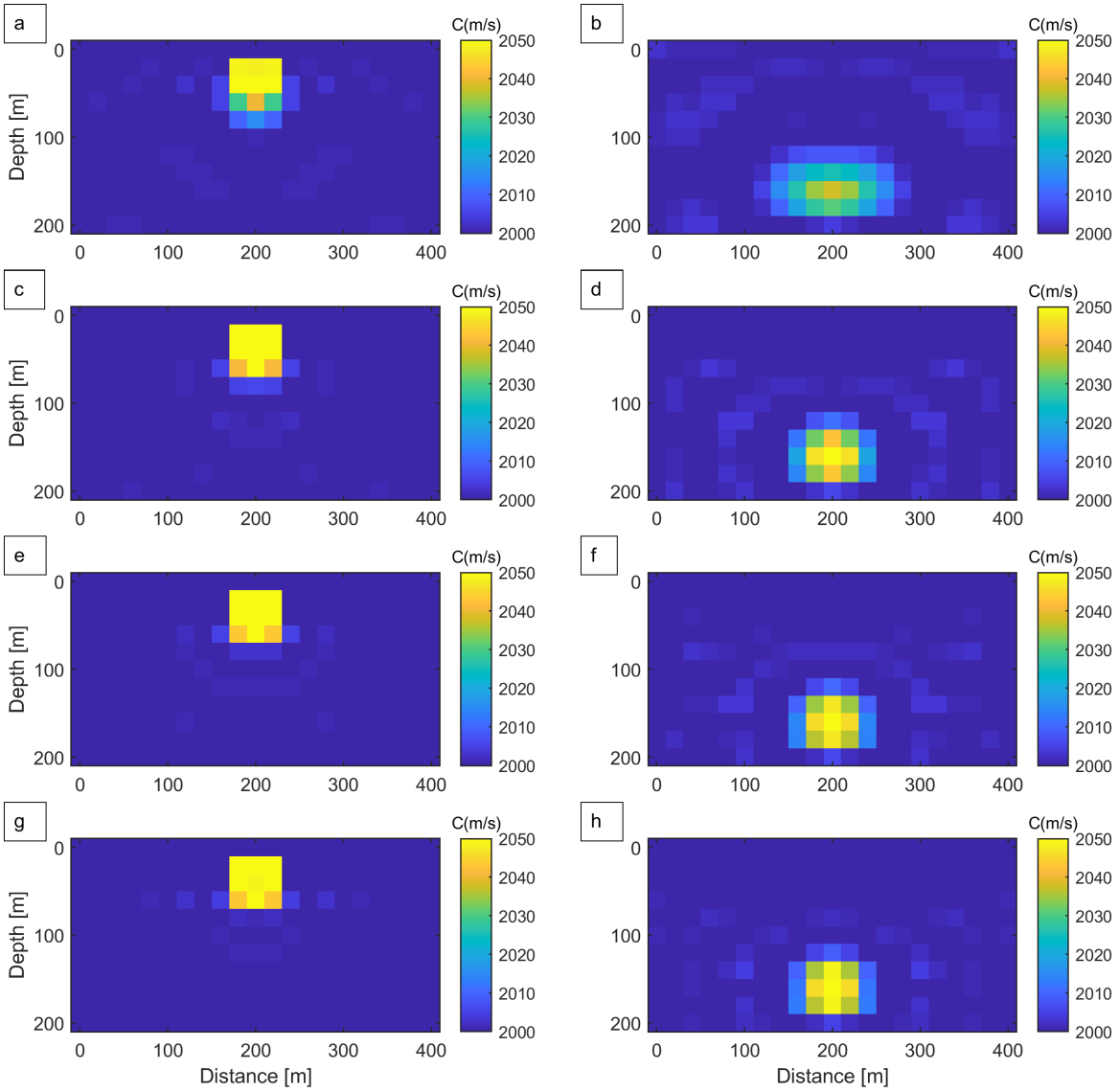

Figure 3.3: Reconstructions of model 1 and 2 with the help of different number of equidistantly spaced receivers. (a) and (b) 11 receivers; (c) and (d) 21 receivers; (e) and (f) 31 receivers; (g) and (h) 41 receivers. For the recoveries 3 equidistantly spaced shots and 6 frequencies [2, 6, 10, 14, 18, 22 Hz] were used.

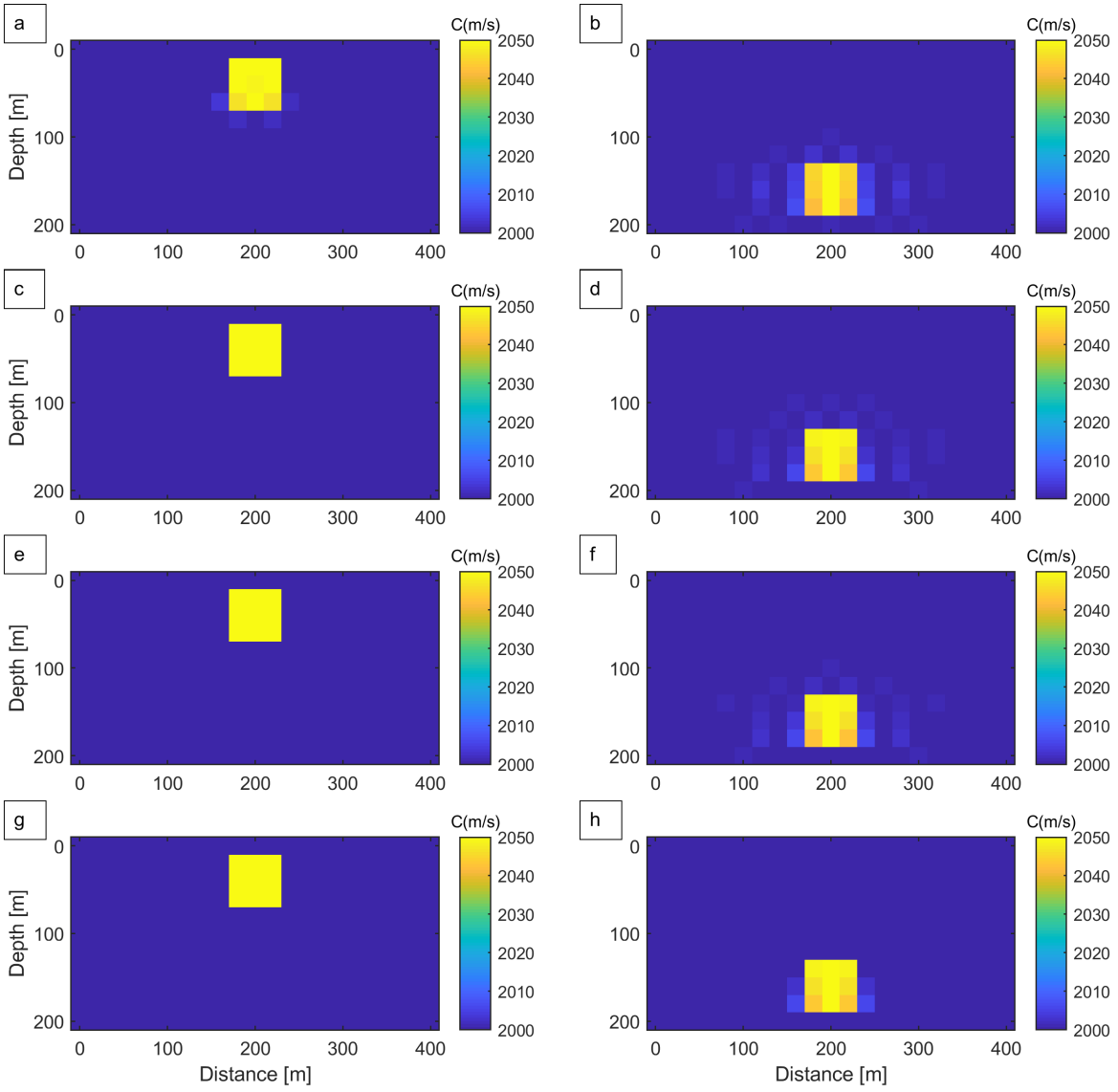

Figure 3.4: Reconstructions of model 1 and 2 with the help of different number of equidistantly spaced shots. (a) and (b) 1 shot (located at the centre of the acquisition system); (c) and (d) 3 shots; (e) and (f) 5 shots; (g) and (h) 7 shots. For the recoveries 21 equidistantly spaced receivers and 3 frequencies [2, 6, 10 Hz] were used.

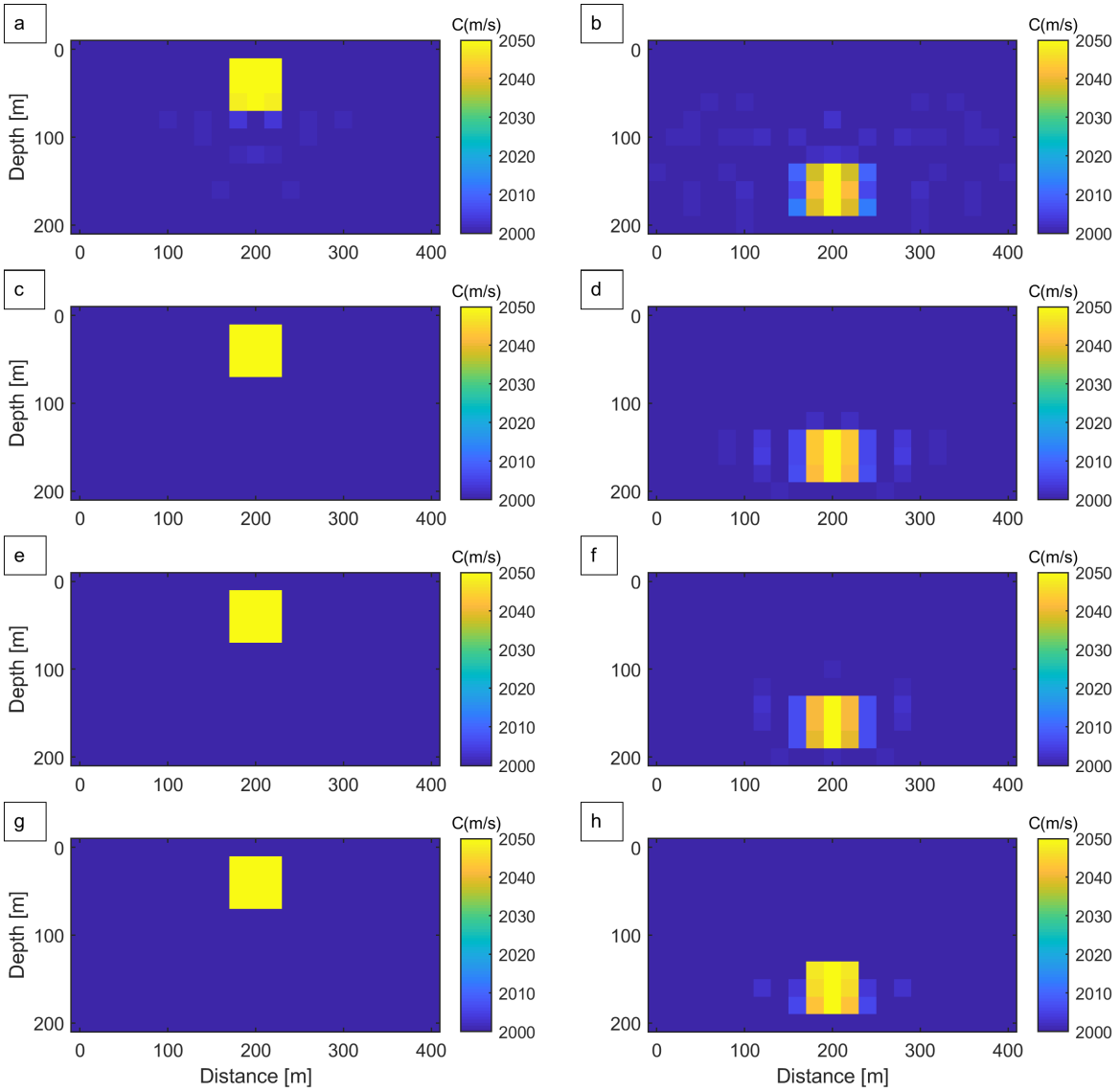

Figure 3.5: Reconstructions of model 1 and 2 with the help of different number of frequencies. (a) and (b) 3 frequencies [2,10,18 Hz]; (c) and (d) 6 frequencies [2, 6, 10, 14, 18, 20, 22 Hz]; (e) and (f) 11 frequencies (each even frequency in the interval 2-22 Hz); (g) and (h) 21 frequencies (each integer frequency in the interval 2-22 Hz). For the recoveries 3 equidistantly spaced shots and 21 equidistantly spaced receivers were used.

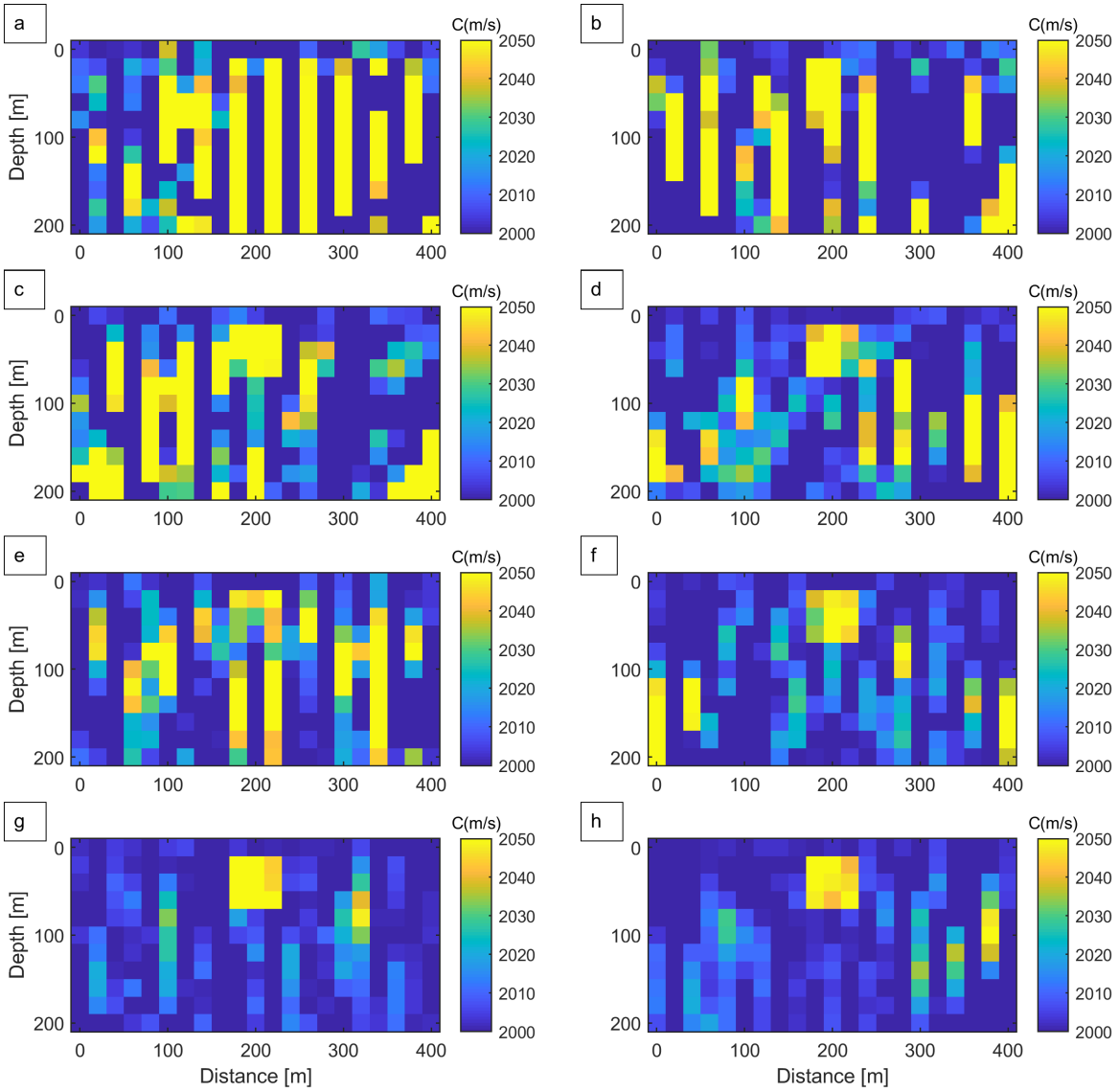

Figure 3.6: Reconstructions of model 1 from the data set with different noise levels (a) SNR=2 dB (correspond to 80% noise); (b) SNR=4 dB (corresponds to 63.3% noise); (c) SNR=6 dB (corresponds to 50% noise); (d) SNR=10 dB (corresponds to 32% noise); (e) SNR=12 dB (corresponds to 25% noise); (f) SNR=14 dB (corresponds to 20% noise); (g) SNR=16 dB (corresponds to 15.8% noise); (h) SNR=20 dB (corresponds to 10% noise). For the recoveries, 21 equidistantly located receivers; 6 frequencies [2, 6, 10, 14, 18, 22 Hz]; and, 3 equidistantly spaced shots were used.

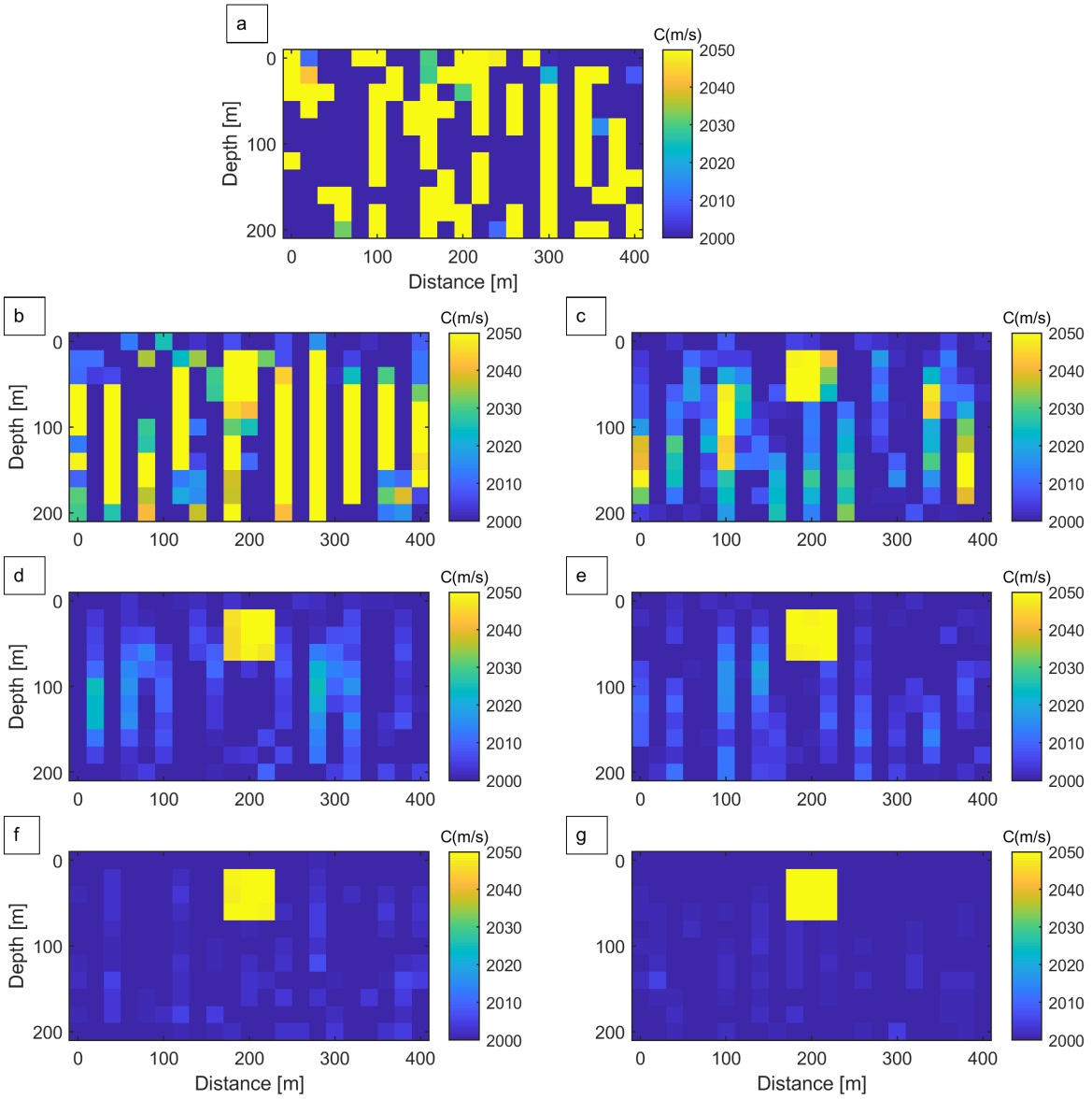

Figure 3.7: Reconstructions of model 1 from the data set with the SNR= 10 dB (corresponds to 32% noise). The reconstructions obtained with the help of (a) 5 frequencies ([2, 8, 12, 16, 22 Hz], 11 receivers, and 3 shots; (b) 7 frequencies [2, 4, 8, 10, 12, 16, 22 Hz], 11 receivers, and 3 shots; (c) 9 frequencies [2, 4, 6, 8, 12, 14, 16, 18, 22 Hz], 11 receivers, and 3 shots; (d) 9 frequencies (as previous), 21 receivers, and 3 shots; (e) the 9 frequencies (as previous), 31 receivers, and 3 shots; (f) 9 frequencies (as previous), 31 receivers, and 7 shots; (g) 9 frequencies (as previous), 31 receivers, and 15 shots.

3.4 Distorted Born iterative T-matrix inversion method

3.4.1 General principle

The DBIT method is a nonlinear FWI technique introduced by (Jakobsen and Ursin, 2015). The technique is based on the T-matrix principle integrated into the classic Distorted Born Iterative (DBI) method previously developed by (Chew and Wang, 1990). The inversion scheme behind the DBI method is similar to other scattering theory-based nonlinear approaches. Firstly, a nonlinear inverse problem is replaced by a series of the associated linear problems. Secondly, each of the linear problems is iteratively solved for both the perturbed and background media by updating the corresponding Green functions. The updating process is carried on until all the selected frequencies are processed (Chew and Wang, 1990). The main problem of this scheme is computation cost: the Green functions updates (at each new iteration) involve inverting large matrices. The T-matrix principle, as discussed before (subsection 2.2.4), changes this problem. Inversion of several matrices corresponding to Green functions is replaced by inversion of only one matrix corresponding to the transition operator $\mathbf{T}$.

In the following derivation of the DBIT inversion method, I assume that the reader is familiar with the the T-matrix and DBA modelling methods discussed in subsection 2.2.4. Not to be repeated, the details relevant to the modelling part will be omitted (they can be found in the subsection mentioned).

The mathematical representation of the DBIT inversion method we start from solving the forward problem for a variation in the scattering potential $\delta\mathbf{M}^{(i+1)} = \mathbf{M}^{(i+1)} - \mathbf{M}^{(i)}$ using the DBA method (Jakobsen and Ursin, 2015)

$$\delta\mathbf{d}^{(i)} \equiv (\mathbf{G}_{RS} - \mathbf{G}_{RS}^{(i)})\tilde{\mathbf{f}} = \bar{\mathbf{G}}_{RV}^{(i)}\delta\mathbf{M}^{(i+1)}\mathbf{G}_{VS}^{(i)}\tilde{\mathbf{f}}_s \tag{3.9}$$

Here, $\delta\mathbf{d}^{(i)}$ represents the linearized solution of the forward model around the inversion result gained after the ith iteration. Given an observation of the scattered field $\delta\mathbf{d}^{(i)}$. We can estimate the perturbation $\delta\mathbf{M}^{(i+1)}$ via a regularized least-squares inversion solution and replace the initial model by the inverted one (Jakobsen and Wu, 2018). After that, we update all Green functions which are responsible "for the various parts of the propagation path from the source to the receiver" (Jakobsen and Ursin, 2015)

$$\bar{\mathbf{G}}_{RV}^{(i+1)} = \bar{\mathbf{G}}_{RV}^{(i)} + \bar{\mathbf{G}}_{RV}^{(i)}\delta\mathbf{T}^{(i+1)}\bar{\mathbf{G}}_{VV}^{(i)}, \tag{3.10}$$

$$\bar{\mathbf{G}}_{VV}^{(i+1)} = \bar{\mathbf{G}}_{VV}^{(i)} + \bar{\mathbf{G}}_{VV}^{(i)} \delta\mathbf{T}^{(i+1)} \bar{\mathbf{G}}_{VV}^{(i)}, \tag{3.11}$$

and

$$\mathbf{G}_{RS}^{(i+1)} = \mathbf{G}_{RS}^{(i)} + \bar{\mathbf{G}}_{RV}^{(i)} \delta\mathbf{T}^{(i+1)} \mathbf{G}_{VS}^{(i)}, \tag{3.12}$$

$$\mathbf{G}_{VS}^{(i+1)} = \mathbf{G}_{VS}^{(i)} + \bar{\mathbf{G}}_{VV}^{(i)} \delta\mathbf{T}^{(i+1)} \mathbf{G}_{VS}^{(i)}, \tag{3.13}$$

where $\delta\mathbf{T}^{(i+1)}$ is the variation in the T-matrix between iteration i and $i+1$ explicitly defined by (Jakobsen and Ursin, 2015)

$$\delta\mathbf{T}^{(i+1)} = (\mathbf{I} - \delta\mathbf{M}^{(i+1)} \bar{\mathbf{G}}_{VV}^{(i)})^{-1} \delta\mathbf{M}^{(i+1)}. \tag{3.14}$$

Equations (3.10-3.14) can also be written in the following form

$$\bar{\mathbf{G}}_{RV}^{(i+1)} = \bar{\mathbf{G}}_{RV}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)} \mathbf{T}^{(i+1)} \bar{\mathbf{G}}_{VV}^{(0)}, \tag{3.15}$$

$$\bar{\mathbf{G}}_{VV}^{(i+1)} = \bar{\mathbf{G}}_{VV}^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)} \mathbf{T}^{(i+1)} \bar{\mathbf{G}}_{VV}^{(0)}, \tag{3.16}$$

$$\mathbf{G}_{RS}^{(i+1)} = \mathbf{G}_{RS}^{(0)} + \bar{\mathbf{G}}_{RV}^{(0)} \mathbf{T}^{(i+1)} \mathbf{G}_{VS}^{(0)}, \tag{3.17}$$

$$\mathbf{G}_{VS}^{(i+1)} = \mathbf{G}_{VS}^{(0)} + \bar{\mathbf{G}}_{VV}^{(0)} \mathbf{T}^{(i+1)} \mathbf{G}_{VS}^{(0)}, \tag{3.18}$$

with $\mathbf{T}^{(i+1)}$ given by (Jakobsen and Ursin, 2015)

$$\mathbf{T}^{(i+1)} = (\mathbf{I} - \mathbf{M}^{(i+1)} \bar{\mathbf{G}}_{VV}^{(0)})^{-1} \mathbf{M}^{(i+1)}. \tag{3.19}$$

3.4.2 Implementation

Assuming a number of sources $\tilde{\mathbf{f}}_s$ and a set of frequencies ω_k $(k = 1, \ldots, N_\omega)$ with the following application of the discretization scheme provided in subsection 2.2.4, equation (3.9) can be rewritten in component form exactly as (Jakobsen and Ursin, 2015)

$$\tilde{d}_{r,sk}^{(i)} = \sum_{n=1}^{N} J_{rn,sk}^{(i)} \delta m_n^{(i+1)}, \tag{3.20}$$

where

$$\tilde{d}_{r,sk}^{(i)} = \tilde{d}_{r,s}^{(i)}(\omega_k), \tag{3.21}$$

and

$$J_{rn,sk}^{(i)} = J_{rn,s}^{(i)}(\omega_k), \tag{3.22}$$

with

$$J_{rn,s}^{(i)}(\omega) = \left[\bar{G}_{rn}^{(i)}(\omega) \delta M_n G_{ns}^{(i)}(\omega)\right] f_s. \tag{3.23}$$

Replacing three indices s, r and k by a single index α, we rewrite equation (3.20) as

$$\delta d_{\alpha}^{(i)} = \sum_{n=1}^{N} J_{\alpha n}^{(i)} \delta m_{n}^{(i+1)}, \tag{3.24}$$

or, in matrix form as

$$\delta \mathbf{d}^{(i)} = \mathbf{J}^{(i)} \delta \mathbf{m}^{(i+1)}. \tag{3.25}$$

Equation (3.25) shows that a nonlinear problem has been reduced to a series of linear problems. Clearly, the solutions of each of the linear problem may be highly ill-posed, that may give unstable results. To manage this issue, we use the Tikhonov regularization method (subsection 3.2.4) and minimize the objective function at each inversion step as (Jakobsen and Ursin, 2015)

$$E(\delta \mathbf{m}^{(i+1)}) = ||\delta \mathbf{d}^{(i)} - \mathbf{J}^{(i)} \delta \mathbf{m}^{(i+1)}||^2 + (\lambda^{(i)})^2 ||\delta \mathbf{m}^{(i+1)}||^2. \tag{3.26}$$

A closed-form solution (Virieux and Operto, 2009) for the model parameters $\mathbf{m}^{i+1}$ of the DBIT inversion method then given by (Jakobsen and Ursin, 2015)

$$\mathbf{m}^{(i+1)} = \mathbf{m}^{(i)} + (\mathbf{H}^{(i)} + (\lambda^{(i)})^2 \mathbf{I})^{-1} \mathbf{V}^{(i)}, \tag{3.27}$$

where $\mathbf{H}^{(i)}$ denotes the Hessian matrix at the ith iteration, explicitly defined by

$$\mathbf{H}^{(i)} = \Re\big[(\mathbf{J}^{(i)})^{\dagger} (\mathbf{J}^{(i)})\big], \tag{3.28}$$

and $\mathbf{V}^{(i)}$ represents the gradient vector at the ith iteration, given by

$$\mathbf{V}^{(i)} = \Re\big[(\mathbf{J}^{(i)})^{\dagger} \delta \mathbf{d}^{(i)}\big]. \tag{3.29}$$

The data residuals vector $\delta \mathbf{d}^{(i)}$ at the ith iteration in equations (3.29) yields (Jakobsen and Ursin, 2015)

$$\delta \mathbf{d}^{(i)} = \delta \mathbf{d}_{obs}^{(i)} - \mathbf{J}^{(0)} \mathbf{m}^{(i)}, \tag{3.30}$$

where, $\delta \mathbf{d}_{obs}^{(i)}$ characterizes the scattered wavefield relative to the initial model (i=0). From equation (3.30), one can note that if the initial model is close to the real model, then the data residual vector $\delta \mathbf{d}^{(i)}$ converges toward 0, and consequently, the estimated model parameters in equation (3.27) converges toward the true model parameters (Jakobsen and Ursin, 2015).

The DBIT inversion scheme introduced above involves choosing a value for the regularization parameter at each iteration. This can be done by using different methods

mentioned earlier (see subsection 3.2.4). In this thesis, I focus only on the discrepancy principle developed by Constable et al. (1987). Two options for extending the discrepancy principle of Constable et al. (1987) will be considered. The first option assumes that the discrepancy principle is combined with the cooling scheme introduced by Farquharson and Oldenburg (2004). Another option suggests integrating the discrepancy principle with the self-adaptive technique presented by Ciric and Qin (1997). Both the cooling scheme and the self-adaptive technique can be used in combination with the DBIT method discussed above. However, experimenting with DBIT-based FWI in strongly scattering media, the author found that the self-adaptive technique was remarkably efficient in those cases. Thus, even both techniques perform the same function, I decided to give a description of the cooling scheme in this section (as a part of the general DBIT method presented above) and to discuss the use of the self-adaptive technique in subsection 3.5.5, when considering nonlinear inversion based on high-contrast media models.

The cooling technique used in this thesis is provided by the simplified version of the original scheme introduced by Farquharson and Oldenburg (2004). In this technique, a key point is to properly select a value for the initial regularization parameter (varies in diapason between 0 and 1) which is gradually decreased at each successive iteration ensuring algorithm convergence. Farquharson and Oldenburg (2004) explaining the importance of an initial choice note that if the starting regularization parameter is too small at early iterations, then the estimated model can accumulate artefacts to compensate for model errors associated with our lack of information about the model structure. To get rid of these artefacts later, one may require to perform lots of additional iterations. As a consequence, this leads to the higher cost. Therefore, as the authors advise, it is more efficient "if the starting value of the regularization parameter is fairly large and restrictions are placed on its greatest allowed decrease, thus enforcing a slow but steady introduction of structure into the model". There are not any specific rules to define this starting value. In each case, a proper choice mainly depends on the size of the model and data noise level. For this reason, it is useful to perform several tests in order to find an acceptable value, exploring the provided interval between 0 and 1.

Mathematically, the simplified cooling scheme, if we assume that $\lambda^{(i)}$ is the value of the regularization parameter at ith iteration, can be written as

$$\lambda^{(i)} = \lambda^{(0)} a^{(i-1)}, \tag{3.31}$$

where a varies between 0.1 and 0.9. Performing the whole routine in repeated manner for all iterations and frequencies, we end up with the final DBIT solution. The DBIT algorithm is shown in figure 3.8.

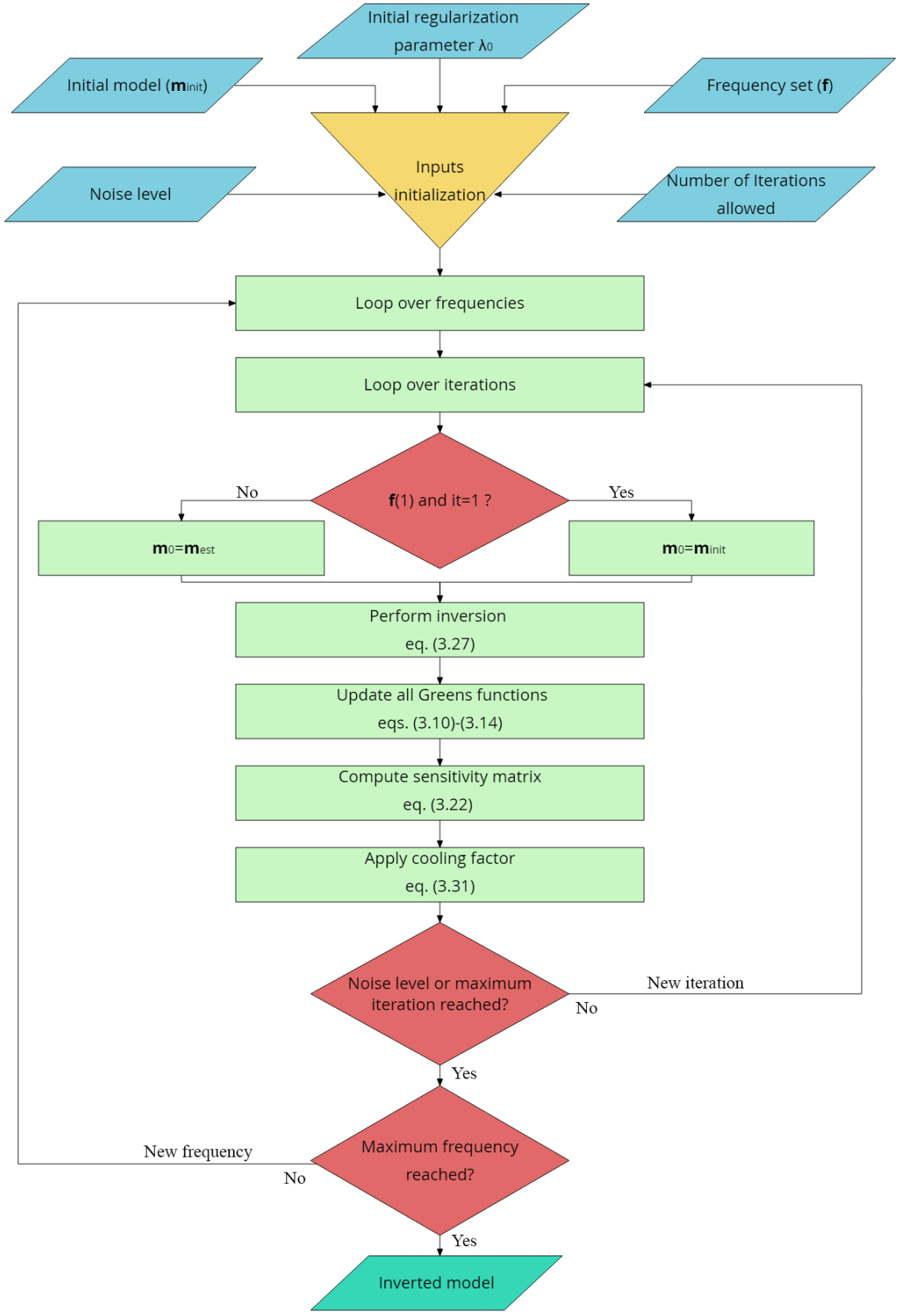

Figure 3.8: DBIT inversion algorithm.

The presented above DBIT inversion algorithm can be easily modified into the Born inversion routine used in section 3.3. The modifications imply only removing Green functions updates from the DBIT inversion method discussed (equations 3.10-3.19).

3.5 Time-lapse distorted Born iterative T-matrix full waveform inversion in strongly scattering media

3.5.1 Time-lapse inversion strategies

Time-lapse FWI is a high-resolution technique that allows quantitatively imaging variations in physical parameters within reservoir zones from seismic data acquired at different calendar times (Asnaashari et al., 2011).

The concept of time-lapse FWI involves two main steps (**?**). First, the recovery of the baseline model corresponding to the reservoir preproduction stage, second, the reconstruction of the monitor model correlating with the reservoir production stage. The time-lapse reconstruction is the differences between these two restored models. Since time-lapse changes in the reservoir are mainly related to the variations in the target area, the second step does not always require inverting for the whole model, reducing the inversion process to the reservoir zone only (Huang et al., 2018).

Generally, three main strategies for the time-lapse model reconstructions are recognized (Asnaashari et al., 2015). The first one is the parallel difference strategy that considers independent inversion of data sets corresponding to the baseline and monitor seismic surveys. Another one is the sequential strategy that implies using the baseline recovery as the starting model to invert for monitor model. The last one is the differential (double-difference) strategy that involves inverting only the difference between the baseline and monitor data (Asnaashari et al., 2015). In this thesis, only two time-lapse approaches are considered: the parallel difference strategy and the sequential strategy.

3.5.2 Hydrocarbon deposits in strongly scattering media

Unconventional hydrocarbon traps originated in high-contrast media are known in different geological settings (Dolson, 2016). One of the most known settings is usually associ-

ated with salt domes. The evolution of the salt dome-related hydrocarbon traps usually includes the following stages: at the start, salt masses, being mechanical weak, under applied stress (lithostatic pressure or some tectonic activity, for example) are pushed up through the overlapped hydrocarbon-saturated strata to the earth surface and form dome-like in shape geological structures. When such salt masses cross reservoir rocks, they block hydrocarbons migration pathways capturing them beneath the salt bodies (Jackson and Hudec, 2017). Another geological setting relates to massive, "mushroom-like" in shape, subvolcanic bodies known as laccoliths. Formations of the associated hydrocarbon traps occur in a way very similar to the previous type: liquid masses of volcanic rocks, due to internal magmatic pressure, breaking through the overlapped hydrocarbon-saturated rocks. When the balance between magmatic and lithostatic pressures is reached (often somewhere near the subsurface), a hot volcanic substance starts solidifying and forms laccoliths. Being cut off by these intrusions, hydrocarbons are become captured beneath them (Farooqui et al., 2009).

In both geological cases described above, the velocity contrast between the intruding and the surrounding rocks is incredibly high (Jackson and Hudec, 2017; Schutter, 2003). This fact leads to the following issue: most of the seismic energy is reflected back from the top of such bodies resulting in insufficient illumination beneath them (Ovcharenko et al., 2018).

In the following, when discussing high-contrast media, I will mainly address the salt-dome type models (since this one is more popular in petroleum seismology practice) but the reader should keep in mind that all the numerical experiments are also relevant for the subvolcanic type.

3.5.3 Seismic inversion in strongly scattering media

Subsalt imaging is traditionally considered as the most challenging problem in petroleum seismology. The illumination problem mentioned previously, requires applying advanced seismic techniques at all stages: from data collecting to seismic interpretation. Data collecting implies using the appropriate multi-azimuth long-offset acquisition system with a broadband signal to illuminate the salt-affected target properly. Processing often includes high-quality noise suppression (in order to get rid of all except primary P-wave reflections) and several depth migrations (to define the proper position of subsalt geological structures). Interpretation is usually performed manually and, therefore, highly reliant on the

skills and expertise of practitioners involved (**?**). At both, the processing and interpretation stages, some difficulties may occur. FWI, with its ability to obtain high-resolution velocity models, can significantly reduce the cost of the subsalt imaging by replacing the conventional workflow with the automatized model recovery (**?**). However, the use of FWI on high-contrast models may be challenging. The possible challenges are a lack of an accurate starting model and the absence of ultra-low frequencies (ULF) in real seismic data (Alkhalifah and Choi, 2012). Note, in seismic the ULF interval does not have any defined boundaries: this one varies depending on the model size. Off course, the FWI challenges mentioned, are applied to all kind of velocity models, but for high-contrast cases, these issues are much more significant.

In the first case (the absence of an accurate initial model), if the difference between estimated and observed data such that mismatches in travel times are more than half of period, then the so-called cycle-skipping problem takes place (Alkhalifah and Choi, 2012; Ovcharenko et al., 2018). In the model domain, when performing inversion, the presence of this problem is characterized by repeated contrast velocity anomalies (Virieux and Operto, 2009; Ovcharenko et al., 2018).

Several researchers have considered the importance of ULF in FWI. Bunks et al. (1995) applying the multiscale approach, have illustrated the difference in FWI results reconstructed with the help of different scales. Baeten et al. (2013) have demonstrated the significance of ULF comparing the model restorations obtained from real seismic data where the lowest frequencies in a frequency band were 1.5 Hz and 2 Hz. Even the difference between the lowest frequencies was very small (0.5 Hz), in the second case, some structural elements in the recovered model were lost (Baeten et al., 2013).

In a more general sense, both the issues discussed are interrelated. The absence of ULF in data, for example, can be compensated by the accurate initial model. The opposite is also true. In most cases, the developed techniques try to deal with the first issue. This is explained by the fact that the accurate initial model (especially in the case of high-contrast media models) is often not available (Chi et al., 2014; Ovcharenko et al., 2018). Ovcharenko et al. (2018) have divided all these techniques into two main approaches. The first approach considers modifying the data misfit functional (objective function smoothing techniques). The second approach includes methods employing "image or gradient manipulations with single iteration updates of FWI (processed gradients can lead to shorter paths toward the global minimum)".

In the light of the problems discussed above, for the DBIT inversion technique to

be applied to strongly scattering media models, some adjustments are required. In this thesis, I consider two techniques that make it possible to adapt the DBIT inversion algorithm to strongly scattering models. The first one is the smoothing (or spatial smoothing) technique developed by Jakobsen and Wu (2018). This technique has been designed to obtain an equivalent of the low-wavelength model estimate by repeatedly applying the moving average in order to smooth a data-vector and the corresponding rows of sensitivity matrix Jakobsen and Wu (2018). The second technique is the self-adaptive technique for selecting the regularization parameter. The self-adaptive technique cannot be referred to any approaches mentioned earlier since it has nothing to do with objective functions smoothing or gradient manipulations. However, due to the ability to select the regularization parameter at each successive iteration in a "wise way" (Ciric and Qin, 1997), it provides more stable results when applied to complex FWI problems (including strongly scattering cases).

3.5.4 Smoothing technique

The principle of the smoothing technique is quite simple. Consider data sets where ULF are absent. Application of the moving average to both the data residual and the sensitivity matrix for all sources available is "equivalent to the extraction of low-wavelength data component that can replace the missing ULF" (Jakobsen and Wu, 2018).

The mathematical foundation of the smoothing technique is explained as follows. Consider equation (3.25). It is a system of linear equations. Any mathematical operations applied to both sides of this system of equations do not change the balance if these operations are identical. Say, we multiply the residual data vector $\delta\mathbf{d}^{(i)}$ and each row $\mathbf{j}^{(i)}$ of a sensitivity matrix $\mathbf{J}^{(i)}$ by the same constant. By doing this, not any change in the model vector $\delta\mathbf{m}^{(i+1)}$ is expected. Now, if we assume that $\mathbf{d}_p^{(i)}$ is the P-point residual data-vector at ith iteration, with $p = 0, 1...(P-1)$, and N_d is the number of data points in the moving average operating on $\mathbf{d}_p^{(i)}$, then the smoothed data are given by

$$\tilde{\mathbf{d}}_p^{(i)} = \frac{1}{N_d} \sum_{j=p}^{p+N_d-1} \mathbf{d}_p^{(i)}. \tag{3.32}$$

Similarly, for each smoothed row $\mathbf{j}_p^{(i)}$ of the sensitivity matrix $\mathbf{J}_p^{(i)}$, we can write

$$\tilde{\mathbf{j}}_p^{(i)} = \frac{1}{N_d} \sum_{j=p}^{p+N_d-1} \mathbf{j}_p^{(i)}. \tag{3.33}$$

Now, the modified DBIT inversion algorithm (combined with the smoothing technique) that handles the absence of ULF in seismic data can be obtained by rewriting equations (3.26) and (3.27) in the following form:

$$E(\delta\mathbf{m}^{(i+1)}) = ||\delta\tilde{\mathbf{d}}^{(i)} - \tilde{\mathbf{J}}^{(i)}\delta\mathbf{m}^{(i+1)}||^2 + (\lambda^{(i)})^2||\delta\mathbf{m}^{(i+1)}||^2, \tag{3.34}$$

and

$$\mathbf{m}^{(i+1)} = \mathbf{m}^{(i)} + (\tilde{\mathbf{H}}^{(i)} + (\lambda^{(i)})^2\mathbf{I})^{-1}\tilde{\mathbf{V}}^{(i)}, \tag{3.35}$$

where

$$\tilde{\mathbf{H}}^{(i)} = \Re\big[(\tilde{\mathbf{J}}^{(i)})^\dagger(\tilde{\mathbf{J}}^{(i)})\big], \tag{3.36}$$

and

$$\tilde{\mathbf{V}}^{(i)} = \Re\big[(\tilde{\mathbf{J}}^{(i)})^\dagger\delta\tilde{\mathbf{d}}^{(i)}\big]. \tag{3.37}$$

Here, all the components have the same physical meanings as in equations (3.26-3.29)

3.5.5 Self-adaptive technique

The self-adaptive technique is a method for automatically selecting a value for the regularization parameter (Ciric and Qin, 1997). The selection is based on the error behaviour computed between real and estimated data at each successive iteration. The idea is very straightforward: apply a multiplier smaller than 1 to the regularization parameter from previous iteration if the error decreases, and large than 1 if that one increases.

The discrepancy principle with the cooling scheme integrated into the DBIT algorithm has shown very good results. In that case, a starting value λ_0 for the regularization parameter is selected arbitrarily in the defined interval (subsection 3.4.2). Being chosen, this one is multiplied by a factor, usually ranges from 0.1 to 0.9 at each successive iteration. By doing this, we assume that the error between the real and estimated data after each iteration decreases. What, if for some reason, the error unexpectedly starts increasing? In a worse case, the FWI algorithm diverges. Another option, the divergence process is stopped by the iteration threshold. In this case, at the next frequency, the algorithm probably tries to compensate for artefacts accumulated previously. However, even if this happens, the model reconstruction may still not be acceptable (since some structural elements corresponding to lower frequencies have been irretrievably lost).

The self-adaptive technique suggests choosing a value for the regularization parameter in a "wise way" (Ciric and Qin, 1997). An initial value for the regularization parameter in this technique is based on the statistical approach developed by Franklin (1970), which,

mathematically, can be formulated as (see for the full derivation Franklin (1970) or Ciric and Qin (1997), for example):

$$\lambda_0 = tr(\mathbf{H}^{(0)}/N), \tag{3.38}$$

where $\mathbf{H}^{(0)}$ is the Hessian matrix in equation (3.28) at iteration 0, N is the number of diagonal elements of the Hessian matrix. After computing the first model estimate (equation (3.38)), we multiply the initial regularization parameter by a factor of 0.5 and calculate the error between the real and estimated data using $||\mathbf{d}^{(i)} - \mathbf{J}\mathbf{m}^{(i)}||/||\mathbf{d}^{(i)}||$. Then, we repeat the same procedure for the second iteration and compare the corresponding errors. If the error difference decreases, the model estimate is accepted, and the regularization parameter is multiplied by a factor of 0.5 again. Otherwise, the model estimate is kept equal to that computed previously, and the regularization parameter is multiplied by a factor of 1.5. In the second case, multiplication by a factor of 1.5 will be repeated until the decrease in the error difference is achieved (Ciric and Qin, 1997).

3.6 Numerical results

The inversion tests are performed on two models. The first model is a relatively simple low-contrast fault model intending to show the application and general features of the time-lapse Born and DBIT-based FWI algorithms using the parallel difference approach, exemplify the inverse crime problem (subsection 3.2.7), and, demonstrate time-lapse inversion results for data with different noises added.

For the second series of experiments, the resampled SEG/EAGE salt (high-contrast) model is used. On these experiments, I show how consistent DBIT inversion results (integrated with smoothing and self-adaptive techniques) when applied to data lacking ULF and demonstrate which of the time-lapse strategies (section 3.5.1) is more efficient for time-lapse model reconstructions.

3.6.1 Fault trap model

The fault trap model has already been considered before (subsection 2.3.3). To generate data, the same modelling parameters as in subsection 2.3.3 are used. But now, modelling are performed using only selected frequencies defined for inversion.

To perform Born and DBIT inversion experiments, the frequency-hopping approach is

applied (see subsection 3.2.8) which, in this case, involves using nine frequencies starting from 2 to 20 Hz with a sample rate of 2 Hz. The initial regularization parameter λ_0 is defined to be 10^{-4} that, according to the cooling scheme-based discrepancy principle, is sequentially reduced by a factor of 0.5 at each successive iteration. The noise level and iteration threshold are defined to be 10^{-3} and 30 respectively. The initial velocity model (in order to start the iterative inversion process) is chosen to be a constant model with a speed equal to 2400 m/s.

In all tests, the time-lapse models are computed based on the parallel difference approach (subsection 3.5.1) which considers using the same initial model to restore the baseline and monitor models. The time-lapse anomaly reconstruction is computed as the difference between the baseline and monitor models.

Discussion

Figure 3.10 shows the inversion results for the baseline, monitor and time-lapse models in the case where data were generated by the Born approximation method and inverted with the help of the Born inversion technique. Note, how well the shapes of these models were restored (compare with illustrations (a) and (b) in figure 3.10). Figure 3.11 exemplifies the case where data modelled by the T-matrix method, were inverted using Born inversion. The models failed to be reconstructed. This is exactly what should be expected in such cases. Due to relatively high-contrast (in a sense that it does not correspond to the validity criterion given in equation (2.35)), the tectonic model cannot be restored by Born inversion. The successful restoration in figure 3.10 is a visual example of the inverse crime problem (subsection 3.2.7): the error generated at the modelling stage were compensated by the inversion algorithm because both the forward and inverse solutions rely on the same solver. Figure 3.12 displays another example where the model restorations were performed using the DBIT inversion method. In this case, the seismic data were computed using the T-matrix integral equation method. Now, all the models were recovered. In a more general sense, these reconstructions are also subjects of the inverse crime problem: both the T-matrix modelling method and DBIT-based FWI use the same T-matrix operator when computing. However, as demonstrated in subsection 2.3.4, the modelling errors associated with the T-matrix integral equation method are very small. Therefore, the inverse crime problem can be neglected.

Figures 3.13 and 3.14 display the DBIT inversion results obtained from the noisy data with the SNR equal to 20 dB (corresponds with 31 % noise) and 5 dB (corresponds

with noise 56 %). Note, in both cases, the main shapes of the tectonic trap models were satisfactorily restored (illustrations (a), (b) in figures 3.13 and 3.14). For the less noisy data (the SNR=20 dB) in figure 3.13 (illustration (c)), the time-lapse model was recovered with good quality and the anomaly can be easily spotted. For the noisier case (SNR=5 dB) in figure 3.14 (illustration (c)), the time-lapse anomaly cannot be clearly resolved.

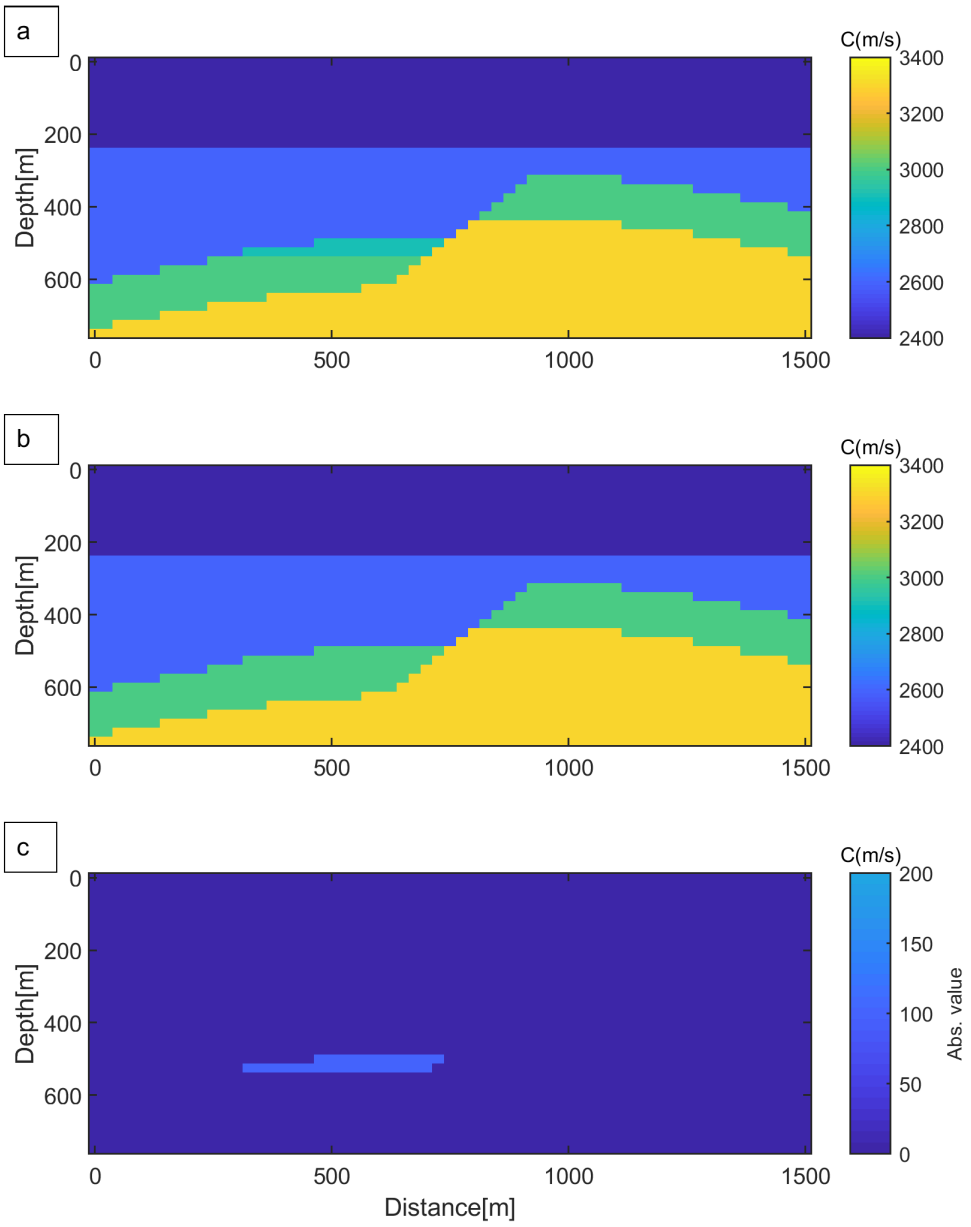

Figure 3.9: Fault trap model. (a) The baseline model; (b) the monitor model; (c) the time-lapse model

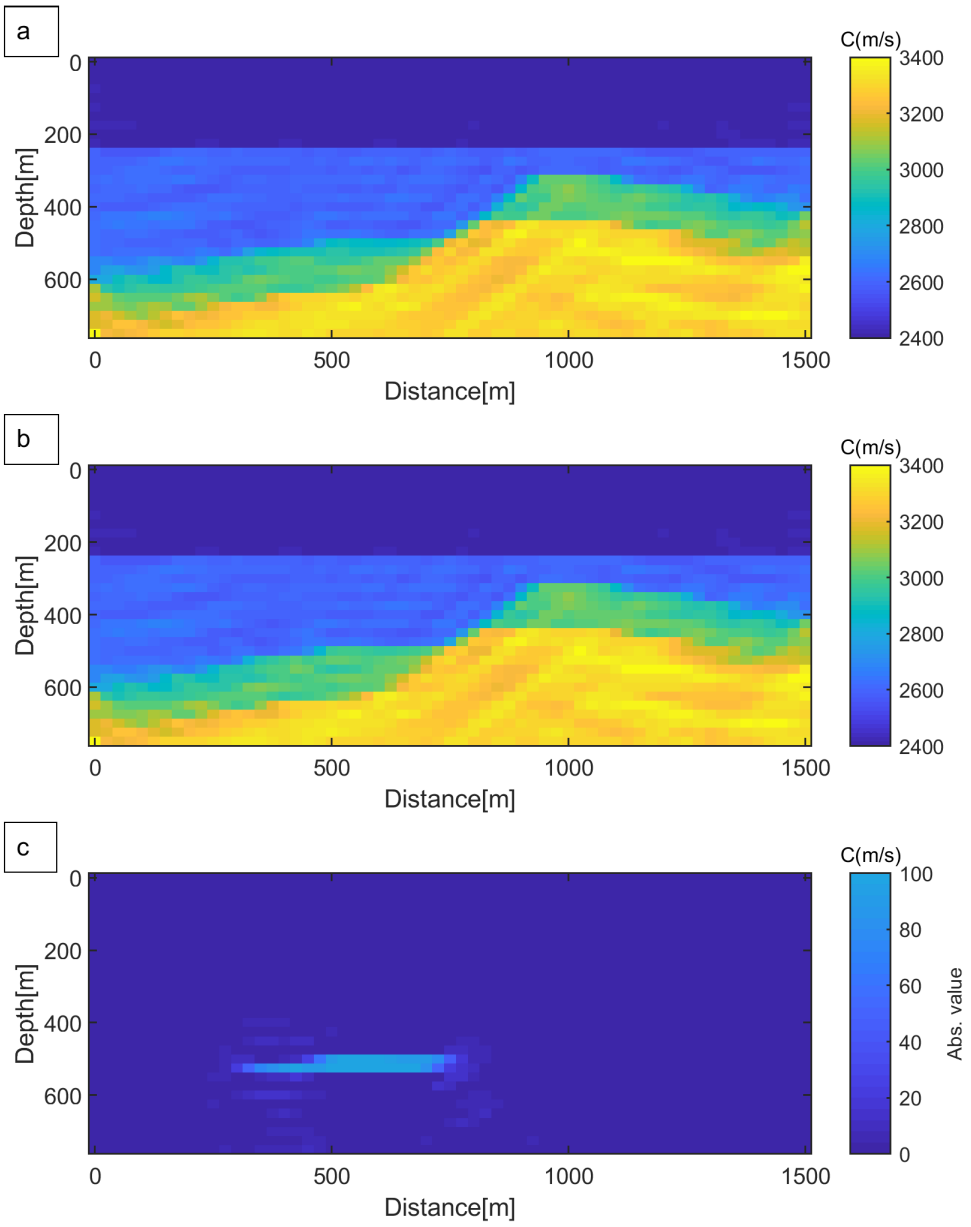

Figure 3.10: Recovered fault trap model (data: the Born approximation; reconstructions: Born inversion). (a) The baseline model; (b) the monitor model; (c) the time-lapse model. The frequency set used [2, 4, 6, 8, 10, 12, 14, 16, 18, 20 Hz]. The time-lapse model obtained using the parallel difference strategy

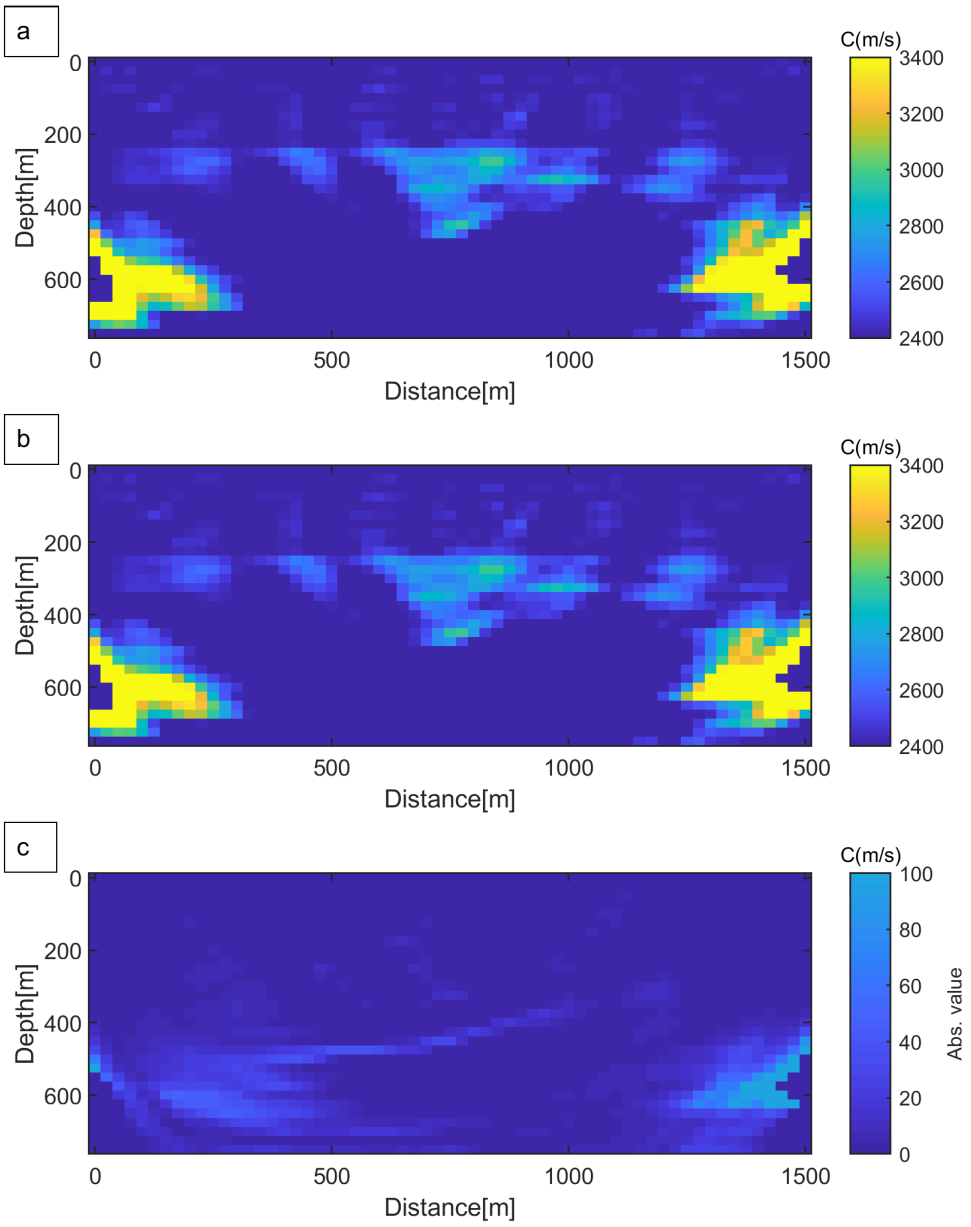

Figure 3.11: Recovered fault trap model (data: the T-matrix modelling method; reconstructions: Born inversion). (a) The baseline model; (b) the monitor model; (c) the time-lapse model. The rest parameters are the same as in figure 3.10.

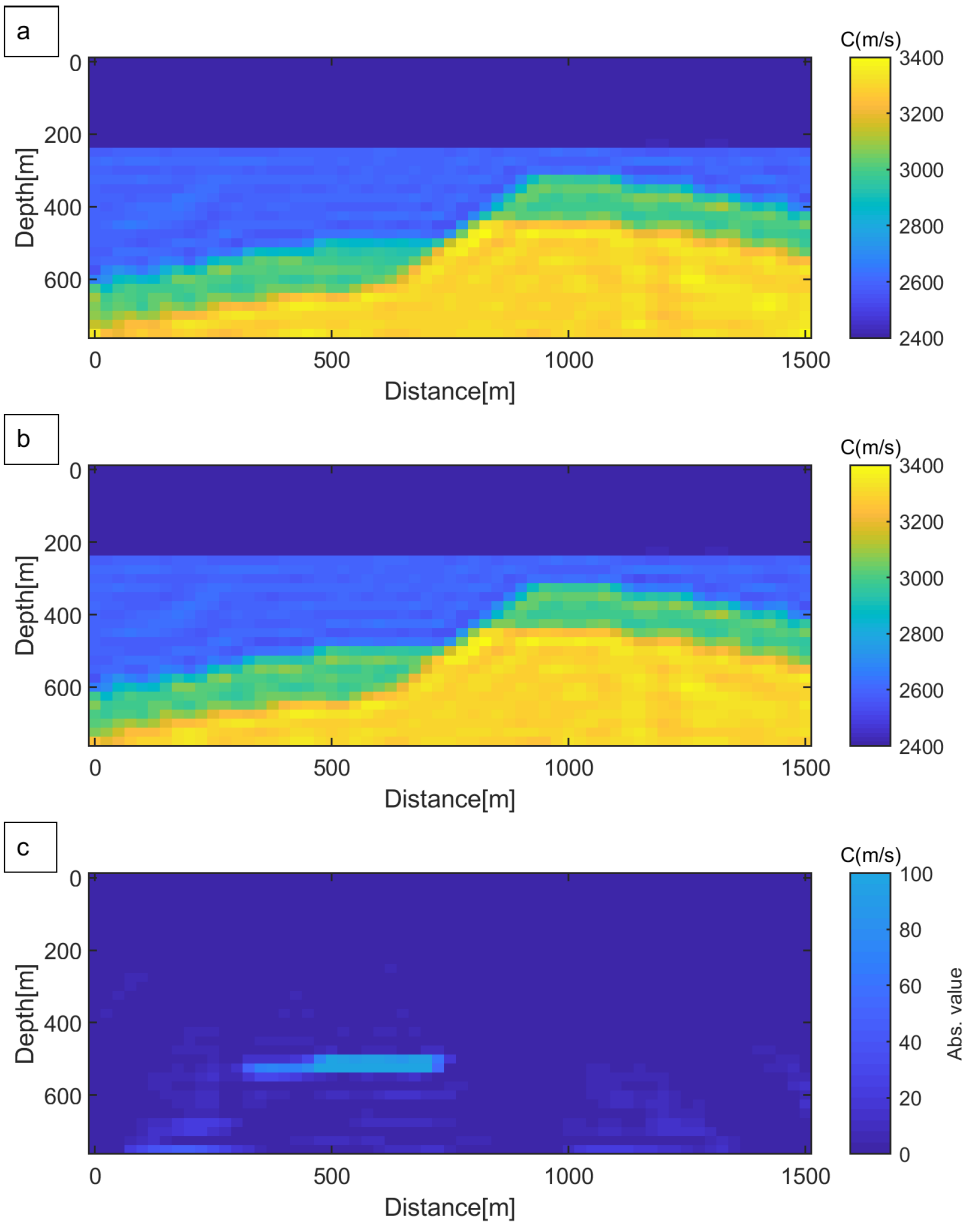

Figure 3.12: Recovered fault trap models (data: the T-matrix modelling method; reconstructions: DBIT inversion). (a) The baseline model; (b) the monitor model; (c) the time-lapse model. The rest parameters are the same as in figure 3.10.

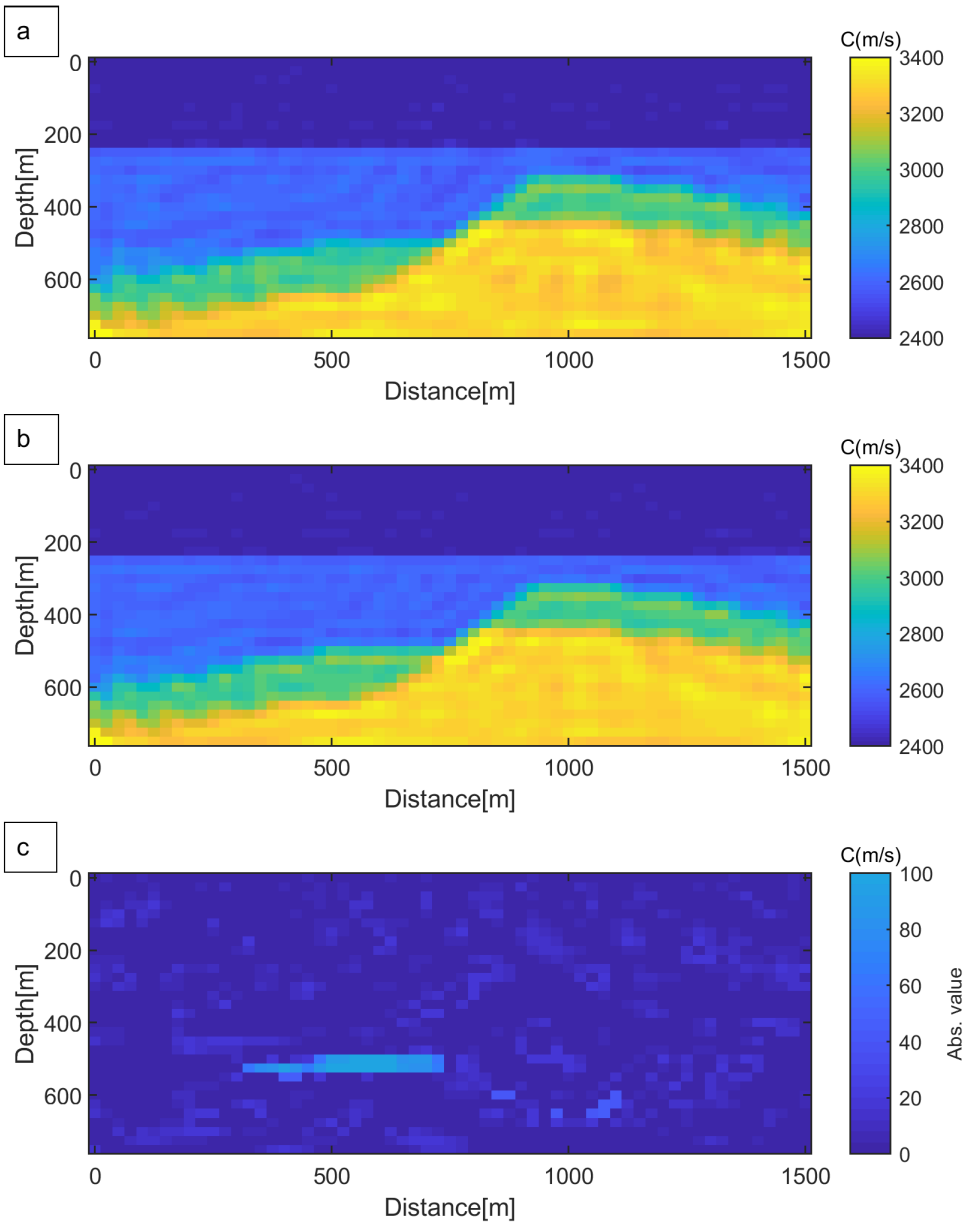

Figure 3.13: Recovered fault trap models (data: the T-matrix modelling method; reconstructions: DBIT inversion). (a) The baseline model; (b) the monitor model; (c) the time-lapse model. SNR=20 dB. The rest parameters are the same as in figure 3.10.

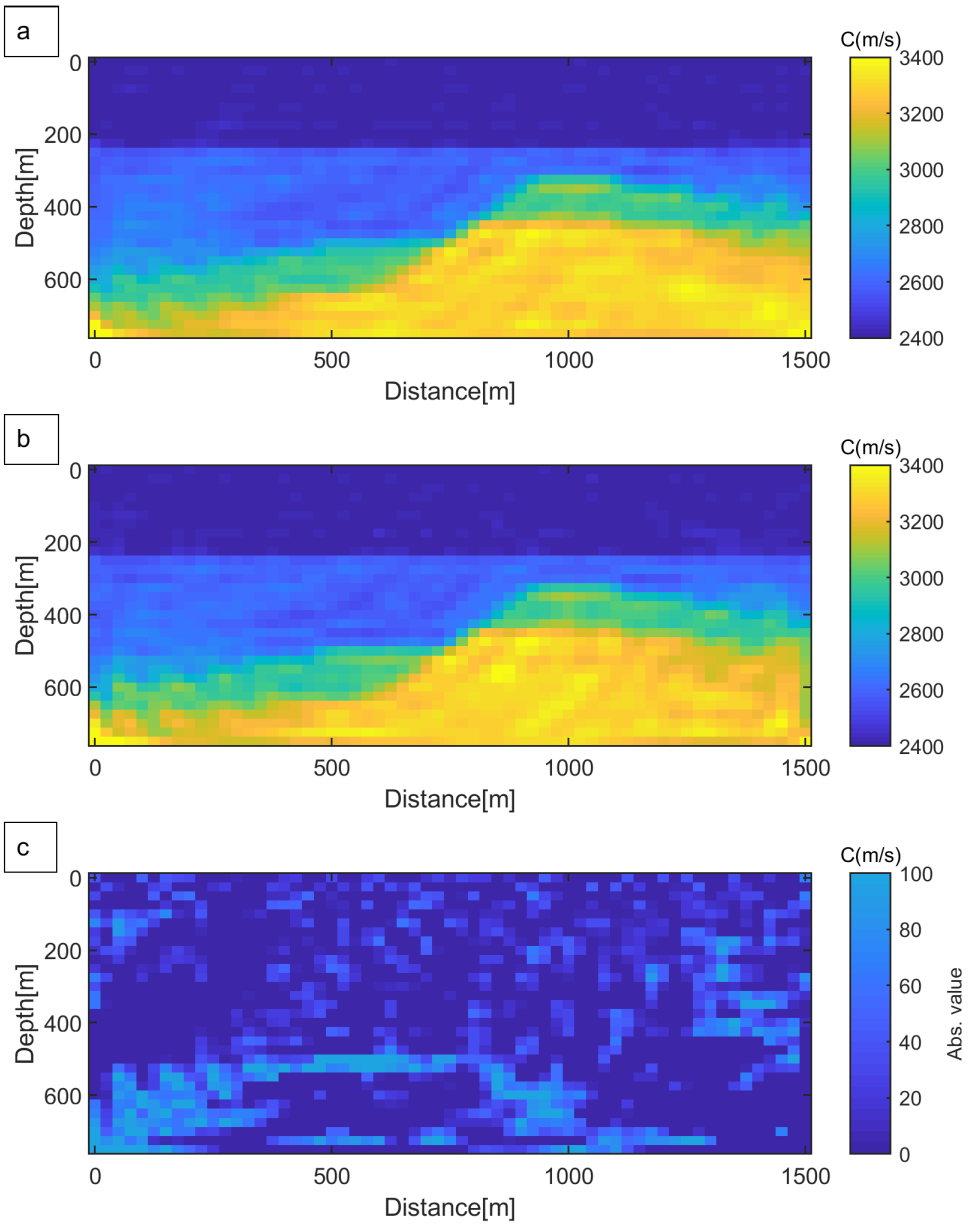

Figure 3.14: Recovered fault trap models (data: the T-matrix modelling method; reconstructions: DBIT inversion). (a) The baseline model; (b) the monitor model; (c) the time-lapse model. SNR=5 dB. The rest parameters are the same as in figure 3.10.

3.6.2 EAGE/SEG salt trap model

High-contrast models are the most challenging for any nonlinear inversion techniques. To test how reliable the DBIT inversion method for time-lapse imaging in strongly scattering media, the well-known EAGE/SEG salt model is considered. In the current study, to speed up the computational process, the original EAGE/SEG salt model is resampled such that the modified version consists of 116 x 24 grid blocks (2784 blocks in total). The size of each grid block is defined to be 24 m in each direction. The total length and depth of the modified EAGE/SEG salt model are 2760 m and 552 m (respectively) with the velocity variation from 1530 m/s to 4500 m/s.

Simulating the imaginary hydrocarbon reservoir, I slightly change the resampled EAGE/SEG salt model following the assumption that the hydrocarbon trap is located just beneath the salt body. In the following, I will assume two different cases of reservoir saturation (different baseline models). In the first case, reservoir rocks are mainly saturated with oil with the P-wave velocity equals to 2400 m/s. In the second situation, I propose that reservoir rocks are filled with a mixture of oil and gas with the P-velocity equals to 1600 m/s. The P-wave velocities in the monitor model (for both cases) I define to be equal 2200 m/s. Therefore, we have two different time-lapse models for the following experiments. The first one (will be named the low-contrast time-lapse velocity model) is the model with the P-wave velocity difference equals to 200 m/s. The second one (will be named the high-contrast time-lapse model) is the model with the P-wave velocity difference of 600 m/s. The low-contrast variation between the baseline and monitor models (first situation) is considered to be an extra challenge for the DBIT inversion method. All the models mentioned are provided in figures 3.15 (illustrations (a-e)). ULF, in the following experiments, denotes the frequency interval 1-4 Hz.

To synthesize observed data corresponded to the baseline and monitor models in figure 3.15, the T-matrix integral equation method (subsection 2.2.4) is used. The survey design in all experiments implies using 116 and 48 equidistantly spaced receivers and sources (respectively) located at the surface. As a source function, to initiate the seismic field, the Ricker wavelet function (equation (2.80)) with a dominant frequency of 7.5 Hz is employed. Total recording time and the time sampling rate are chosen to be 1 s and 0.004 s correspondingly. To experiment with inversion on noisy data, I refer to equation (2.81).

A set of frequencies required for the frequency-hopping approach to perform the DBIT inversion are defined in each specific case. In the experiments where the self-adaptive

technique is not used, the initial regularization parameter is equal to 10^{-3}. Otherwise, this one is defined according to equation (3.31). For the smoothing technique, the smoothing vector is chosen as follows: $\mathbf{N}_d$= [32 16 8 4 1 1 1 1] (each of the vector components corresponds to the number of points used in the moving average for the specific frequency). For all tests, the number of iterations allowed is equal to 30. For the initial model (to start inversion), I choose a linearly-increasing model in which velocity varies from 2000 m/s to 3500 m/s.

Reconstructions of the time-lapse models are performed using two different strategies: parallel difference and sequential (subsection 3.5.1). The baseline model reconstructions (illustrations (a) in figures 3.20-3.31) for all time-lapse experiments are computed based on the linearly-increasing initial model (illustrations (f) in figure 3.15). The monitor models, in turn, are obtained in the following ways: for the parallel difference strategy, in a way similar to the baseline model (starting from the linearly-increasing velocity model); for the sequential strategy, using the baseline reconstructions computed earlier. Note, the proposed self-adaptive and smoothing techniques (and their combination) are used only for the parallel difference time-lapse approach. For the sequential strategy, I use the traditional DBIT algorithm (based on the discrepancy principle with the integrated cooling scheme). Time-lapse models (for both time-lapse techniques) are computed as the difference between the baseline (illustrations (a) in figures 3.20-3.31) and corresponded monitor (illustrations (b) and (c) in figures 3.20-3.31) models.

Discussion

Firstly, before starting any experiments regarding time-lapse model reconstructions, it is important to see how ULF contribute to the inversion process in general. Secondly, when considering the implementation of the smoothing and self-adaptive techniques with the DBIT inversion method (data lacking ULF), it may be useful to observe how accurate and fast such the combined algorithms work.

Figure 3.16 demonstrates the reconstructions of the baseline model with the selected set of frequencies [1, 3, 5, 7, 15, 17 Hz]. Note that ULF are mainly responsible for recovering the general shapes of the model while the higher frequencies are building up smaller structural elements.

Figure 3.17 (illustration (a) and (b) exemplifies two cases: the baseline model recovery for the data with ULF [1, 3, 5, 7, 15, 17 Hz] and without ULF. Note, in the second case

the model failed to be recovered (illustration (b)). The plots (illustrations (c), (d), (e) and (f)) illustrate the performance of the DBIT algorithm for each of these cases. The absence of ULF, as seen from illustration (c), is characterized by the larger initial model error. In this case, the algorithm could not even finish the inversion process at frequency 5 Hz.

Figure 3.18 presents another example where the models were restored using the smoothing and self-adaptive techniques (illustrations (a) and (b) respectively) from the data lacking ULF. The DBIT algorithm for both methods has worked slightly different (illustrations (c)-(f)), but the results are comparable. An estimation of overall error often tends to average the discrepancies corresponded to small structural details. Therefore, it makes sense to compare the inversion results obtained with the help of these techniques at some specific locations within the reconstructed model (model estimate profiles). Figure 3.19 provides such a comparison (based on the baseline model) obtained at three different locations: 240 m, 1392 m and 2520 m. Visually, the difference in reconstruction can be hardly noted (compare illustrations (b)-(d) and (e)-(g)). The corresponded RMS errors (estimated for each profile) are 0.0185, 0.0929, and 0.0145 for the smoothing technique; and, 0.0221, 0.0998, and 0.0185 for the self-adaptive technique. Note, both techniques are characterized by less accurate reconstructions at the location beneath the salt body (illustrations (c) and (f)).

Now, we analyze different time-lapse strategies mentioned. Figure 3.20 shows the time-lapse model recoveries obtained from the data sets containing each odd frequency in the interval from 1 Hz to 17 Hz. The time-lapse model corresponding to the parallel difference strategy (illustration (e)) contains more artefacts comparing to that obtained by the sequential strategy (illustration (d)). Such the recovery difference arises from the initial models used to obtain the monitor models (illustrations (b) and (c)): for the sequential approach, the initial model only differs in the area corresponding to the time-lapse anomaly (therefore the monitor model is computed faster and with minimum artefacts); for the parallel difference approach, inversion requires repeating the whole iterative process from the linearly-increasing model (which is more costly).

Figure 3.21 shows the examples of the model recoveries when the observed data sets are lacking ULF. In this case, the absence of low-wavelength data component forced the DBIT algorithm to diverge when inverting for the baseline and monitor (the parallel difference approach) models (illustrations (a) and (b)). With the absence of the appropriate starting model for the sequential strategy (illustration (a)), the corresponding monitor (illustration (c) and time-lapse models (illustration (e)) failed to be restored as well.

Figure 3.22 shows the model inversion results for the data sets where the absence of ULF was compensated by applying the smoothing technique. All the related models were restored quite well. Compare the time-lapse reconstructions (illustrations (d) and (e)) computed by the different time-lapse approaches. The recoveries corresponding to the sequential approach are obviously better. The application of the self-adaptive technique integrated into the DBIT inversion algorithm (figure 3.23) provides results comparable to figure 3.22.

Figures 3.24 and 3.25 show the experiments for the smoothing and the self-adaptive techniques but applied to the data in which the frequency interval 1-6 Hz was absent. Note, the integrated techniques did not succeed in model recoveries. The next step is to combine both techniques together to see if it gives any recovery improvements.

Figure 3.26 shows the inversion results of the combined scheme applied to the data from the previous example. Note, the application of both techniques together allowed us to restore the baseline and monitor models (illustration (a) and (b)). See, how different the time-lapse reconstructions obtained by the different strategies (illustration (d) and (e)): in the first case, the time-lapse anomaly is surrounded by some artefacts and can be hardly identified if the exact location of the anomaly is not known; the sequential strategy, on the contrary, demonstrates the reconstruction allowing to spot the anomaly without any doubts. Even this experiment is far from practice, since the absence of frequencies in the interval 1- 6 Hz is not common, the fact that the combined scheme made possible to reconstruct the models leads to a conclusion that this combination provides more stable inversion results. Now, we go back and experiment on the data lacking ULF again, but now, using the combined approach. Here, we expect to see some recovery improvements compared to the situations where the techniques were used separately. Indeed, if we compare the model restorations in figure 3.27 with those in figure 3.22 and 3.23, we can note some positive changes. Inspired by these improvements, we finally investigate how this technique combination deals with noisy data.

Figures 3.28 and 3.29 exemplify the cases where the modified DBIT inversion algorithm (with the smoothing and self-adaptive techniques integrated) were applied to the data sets (lacking the ULF) with SNR=20 dB and SNR=5 dB correspondingly. Note, even in both cases the baseline and monitor models were restored (illustrations (a), (b) and (c)), the time-lapse models were not resolved (illustrations (d) and (e)). One of the possible explanations for such time-lapse recovery failure could be a small velocity contrast of the time-lapse anomaly (200 m/s). Therefore, in the following, to test this assumption, we will repeat the DBIT inversion experiments, but now, noisy data will be

generated from the models corresponding to the high-contrast time-lapse anomaly (600 m/s) given in figure 3.15 (illustration (b), (c) and (e)).

Figure 3.30 shows the inversion results for the time-lapse anomaly recovered from data with the SNR equals to 20 dB. Note, in this case, the time-lapse anomaly reconstructions corresponding to both the parallel difference and sequential strategies can be easily identified (illustrations (d) and (e)). The experiment with the high-contrast time-lapse anomaly, but on noisier data (the SNR equals to 5 dB) also provided acceptable results (illustrations (d) and (e) in figure 3.31). Both time-lapse results in figures 3.30 and 3.31 can be additionally improved if we assume that the P-wave velocities in dry and saturated reservoir rocks are known (from laboratory or sonic log measurements). Figure 3.32 shows the time-lapse results from figures 3.30 and 3.31 where the velocity differences outside the interval 300-900 m/s were removed. Note, now, the time-lapse anomalies can be identified even better.

3.6.3 Concluding remarks

Resuming the analysis of the time-lapse DBIT FWI method, I highlight the most remarkable discussion points. For simple reservoirs, as a fault trap model (figure 3.9) for example, the DBIT technique works very efficiently and makes it possible to reconstruct low-contrast time-lapse anomalies even from very noisy data (figures 3.13 and 3.14). When applied to high-contrast models, the model restorations encounter some challenges. First, the absence of ULF in data sets does not allow solving the inversion problem in principle (illustration (b) in figure 3.17). The solution, however, can still be obtained with the help of the smoothing or self-adaptive techniques (figure 3.18). In general, the performance of both techniques are comparable (figures 3.18 and 3.19). In terms of accuracy, the use of the smoothing and self-adaptive techniques together provides even more attractive results (figures 3.27, 3.22 and 3.23). For noisy cases (SNR equal to 20 dB and 5 dB), the low-contrast time-lapse anomaly (the velocity difference about 200 m/s) may not be resolved (figures 3.28 and 3.29). Applied to the high-contrast time-lapse models (the velocity difference about 600 m/s), the anomalies can be resolved (illustrations (d) and (e) in figures 3.30-3.31). The application of the different time-lapse approaches showed that the sequential technique was less costly and recovered time-lapse models with better quality. The additional benefit of the sequential strategy (**?**) is that this approach does not require being extra accurate in survey design when performing the monitor survey. On the other hand, the sequential time-lapse strategy does not take into

account the probable effect of reservoir rocks compaction (Landrø, 2015): hydrocarbon depletion in reservoirs leads to a pore-pressure decrease and forces the reservoir rocks to compact. Even the effect of this geomechanical problem is usually very small, some types of reservoirs may be sensitive to this issue (Landrø, 2015).

The potential of the DBIT inversion applied to high-contrast media is very promising. Still, the analysis cannot be full without mentioning the artificial reflections produced from the model boundaries which contribute to the model recovery process when inverting. Artificial boundaries in the scattering theory-based methods are the boundaries that separate the model environment from a homogeneous background medium. At the modelling stage, when boundary conditions are not applied, the T-matrix modelling method considers these boundaries as real geological interfaces, generating artificial reflections. The DBIT-based FWI, in turn, when restoring a velocity model from such data, treats these artificial reflections as true ones, and use their help in model recovery. Now, consider the salt models in figures 3.15 (illustrations (a), (b) and (c)). Due to high contrast, most of the primary seismic energy reflected from the top of the salt body, which causes the subsalt illumination problem. Here, refractions and internal multiples, in most cases, are the only sources of information about the subsalt structure. And, the contribution of internal multiples is probably more significant. Malcolm et al. (2009), for example, proposed the algorithm that made it possible to illuminate the structures under high-contrast geological bodies using internal multiples only. Therefore, not to be biased in conclusions regarding the efficiency of the DBIT scheme proposed, one can treat the bottom model boundary as a real geological interface, for example.

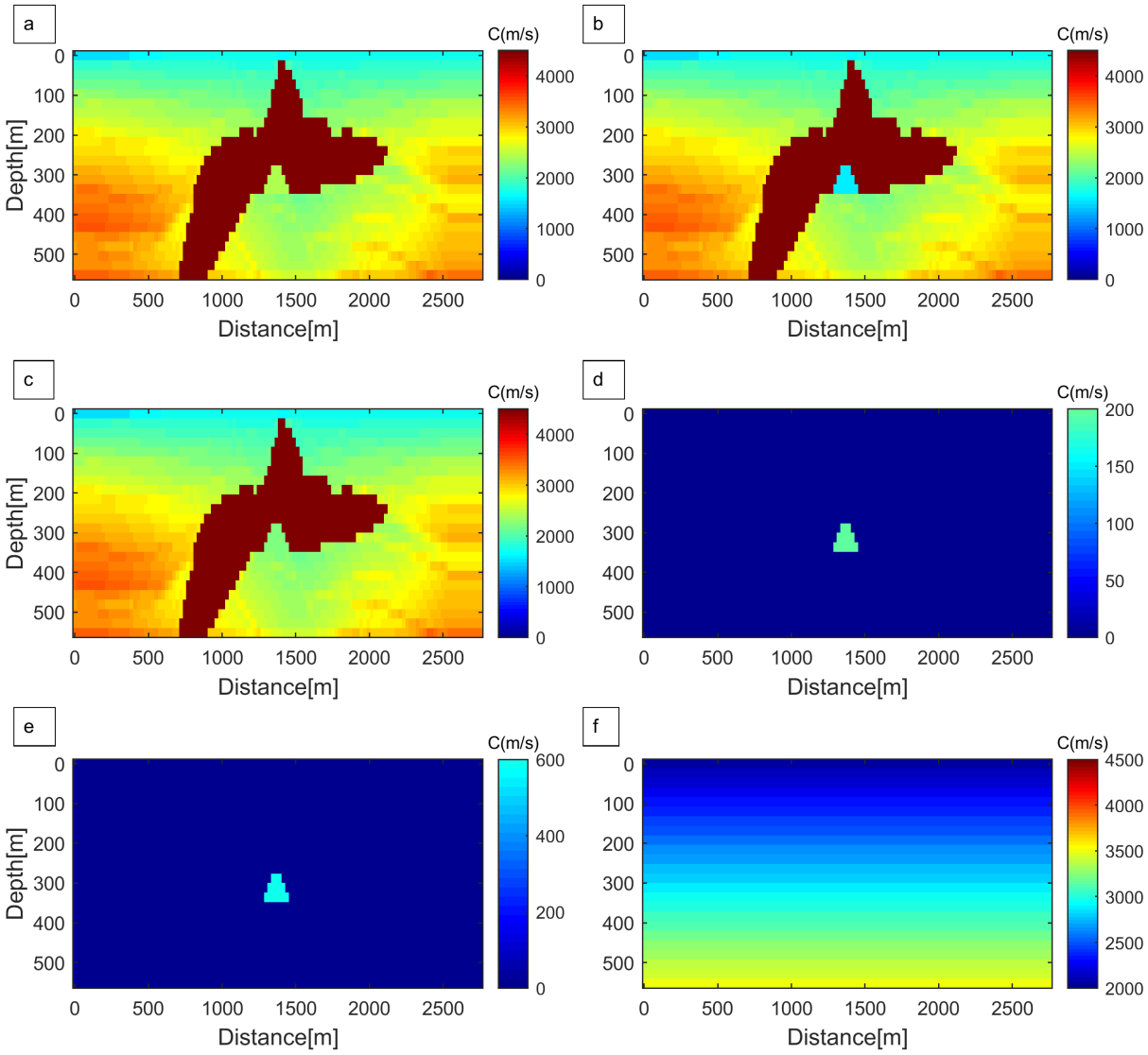

Figure 3.15: Salt trap model.(a) The baseline model (low-contrast anomaly); (b) the baseline model (high-contrast anomaly); (c) the monitor model; (d) the low-contrast time-lapse model (200 m/s); (e) the high-contrast time-lapse model; (f) (600 m/s); the linearly-increasing model.

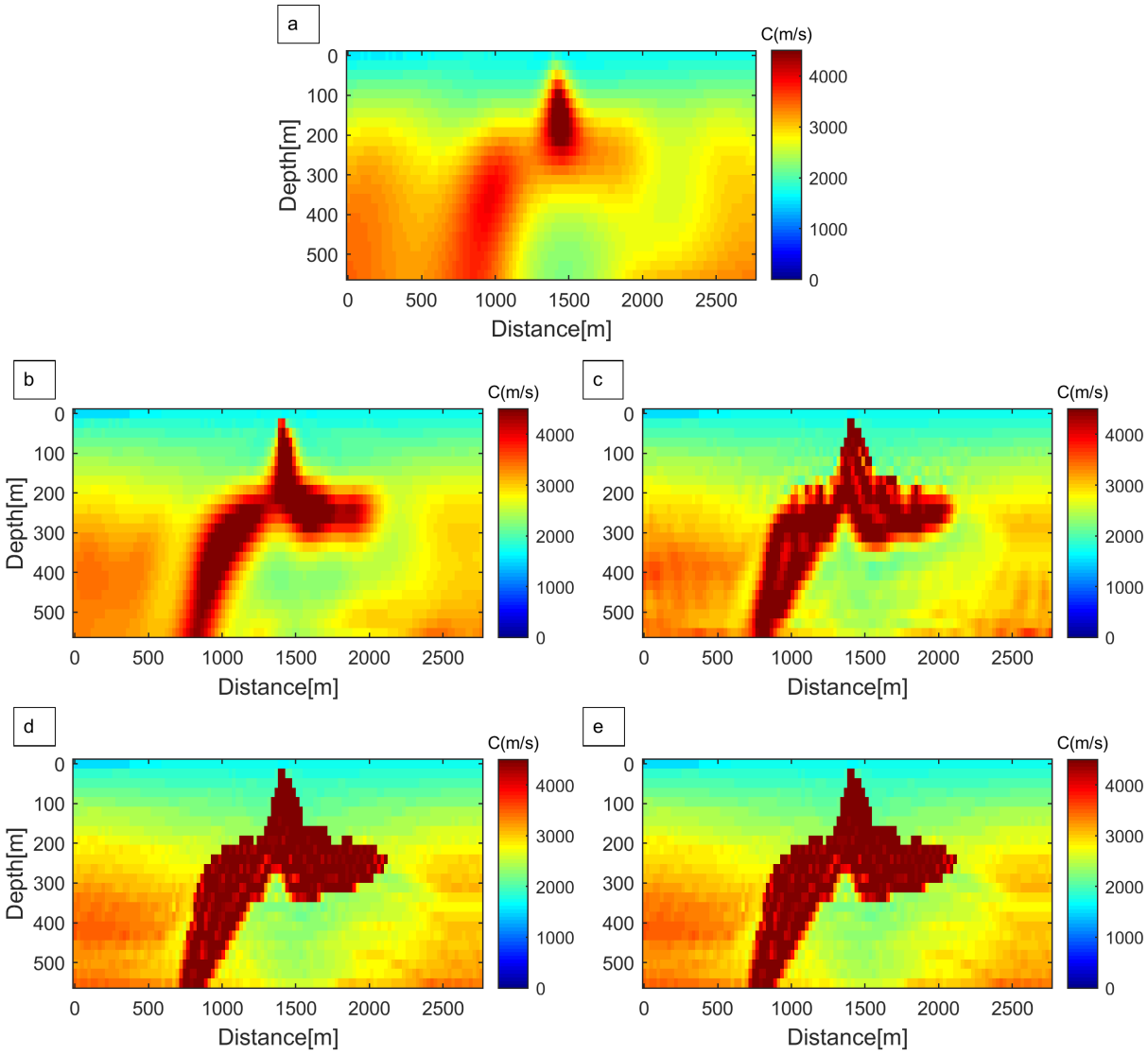

Figure 3.16: Reconstructions of the baseline model using the DBIT inversion method with the help of 5 selected frequencies. (a), (b), (c), (d) and (e) illustrations correspond to 1, 3, 5, 15 and 17 Hz, respectively.

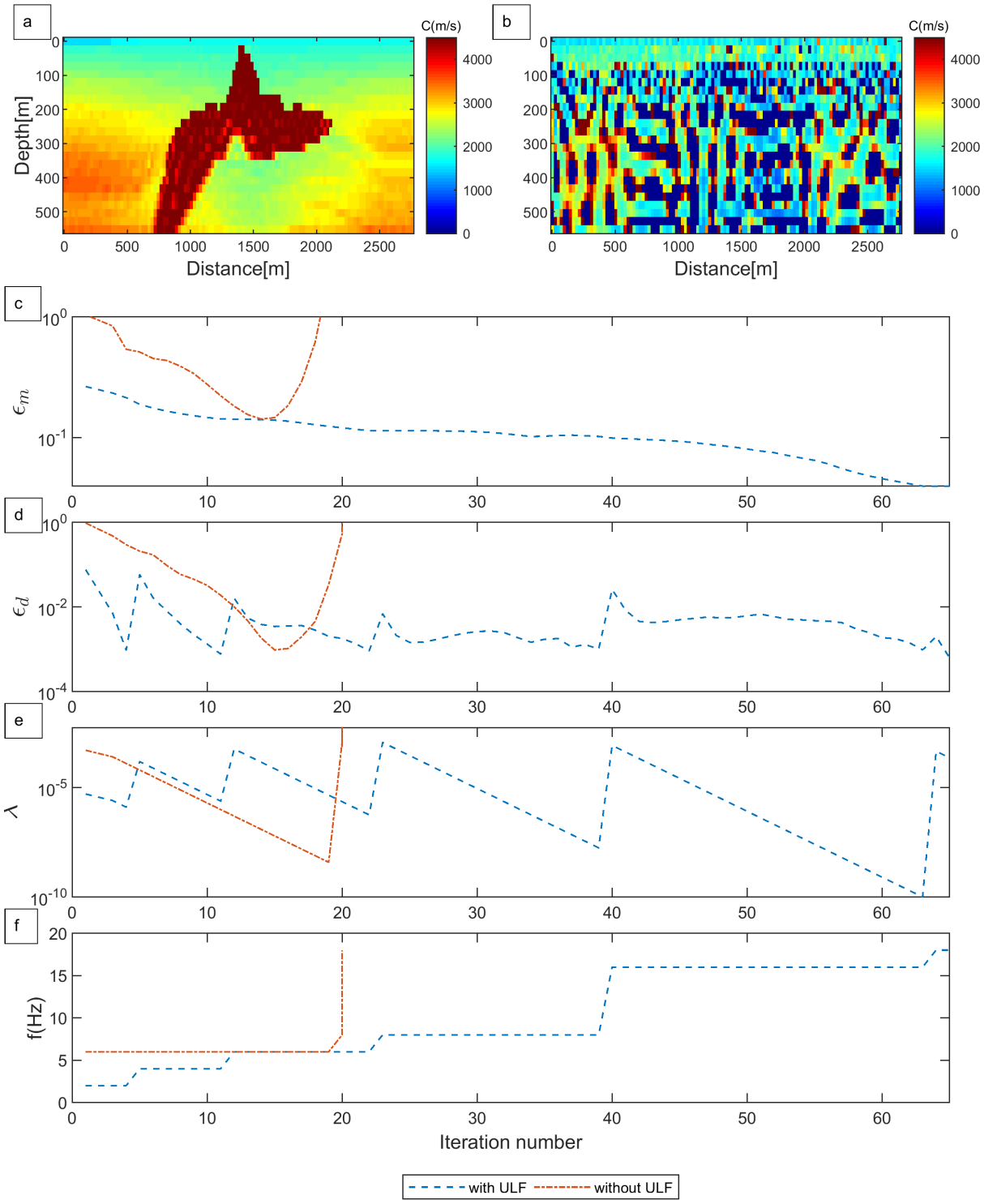

Figure 3.17: Performance of the DBIT inversion algorithm (based on the baseline model) for the data (a) with ULF; (b) without ULF; (c) the overall inversion error ϵ_m at ith iteration defined by $||\mathbf{m}_{(true)}-\mathbf{m}_{(i)}||/||\mathbf{m}_{(true)}||$; (d) the relative data residual error defined by $||\mathbf{d}^{(i)} - \mathbf{J}\mathbf{m}^{(i)}||/||\mathbf{d}^{(i)}||$; (e) regularization parameter λ at ith iteration defined in equation (3.31); (f) frequencies.

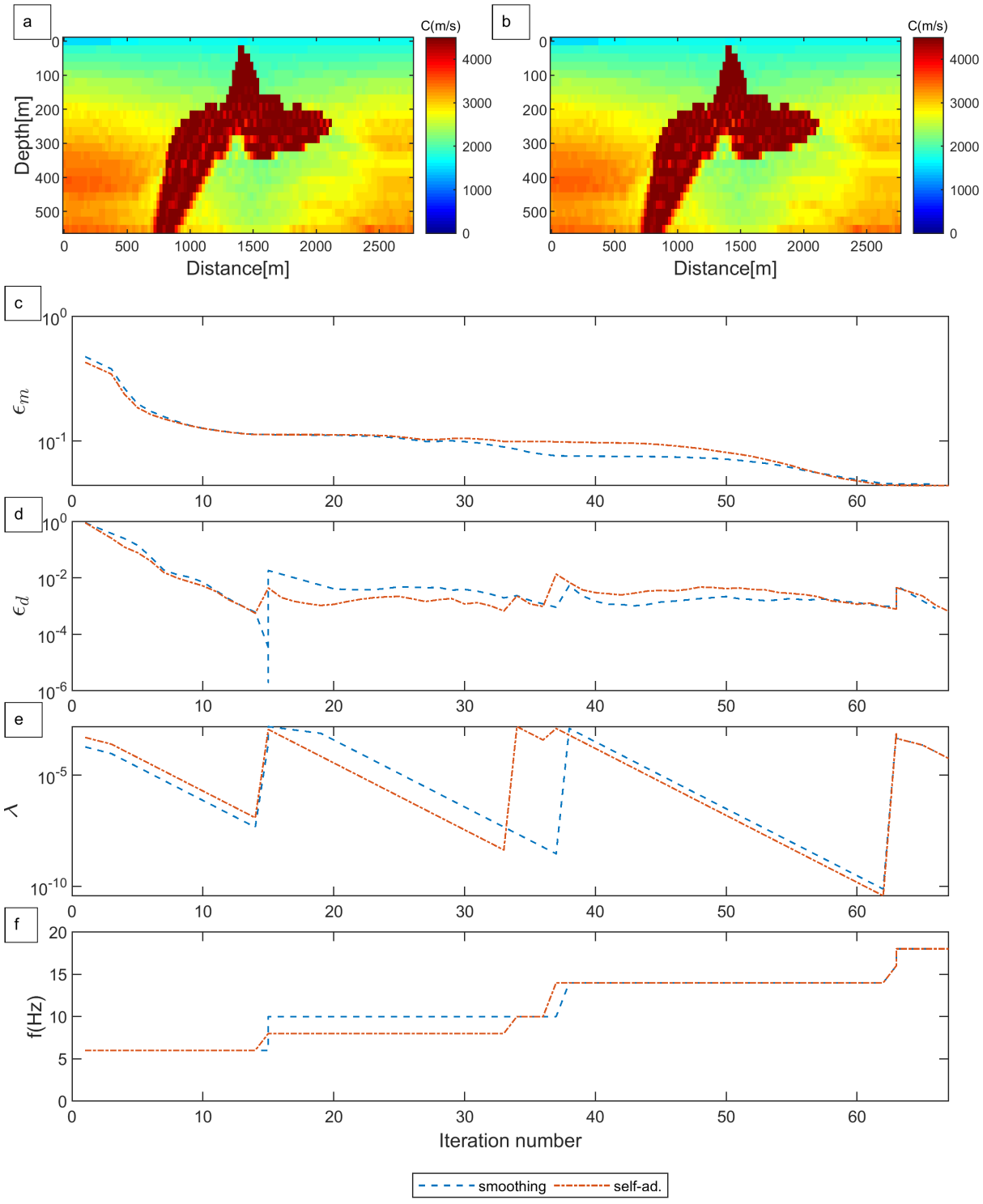

Figure 3.18: Similar to figure 3.17, but with the smoothing and self-adaptive techniques applied (data lacking ULF).

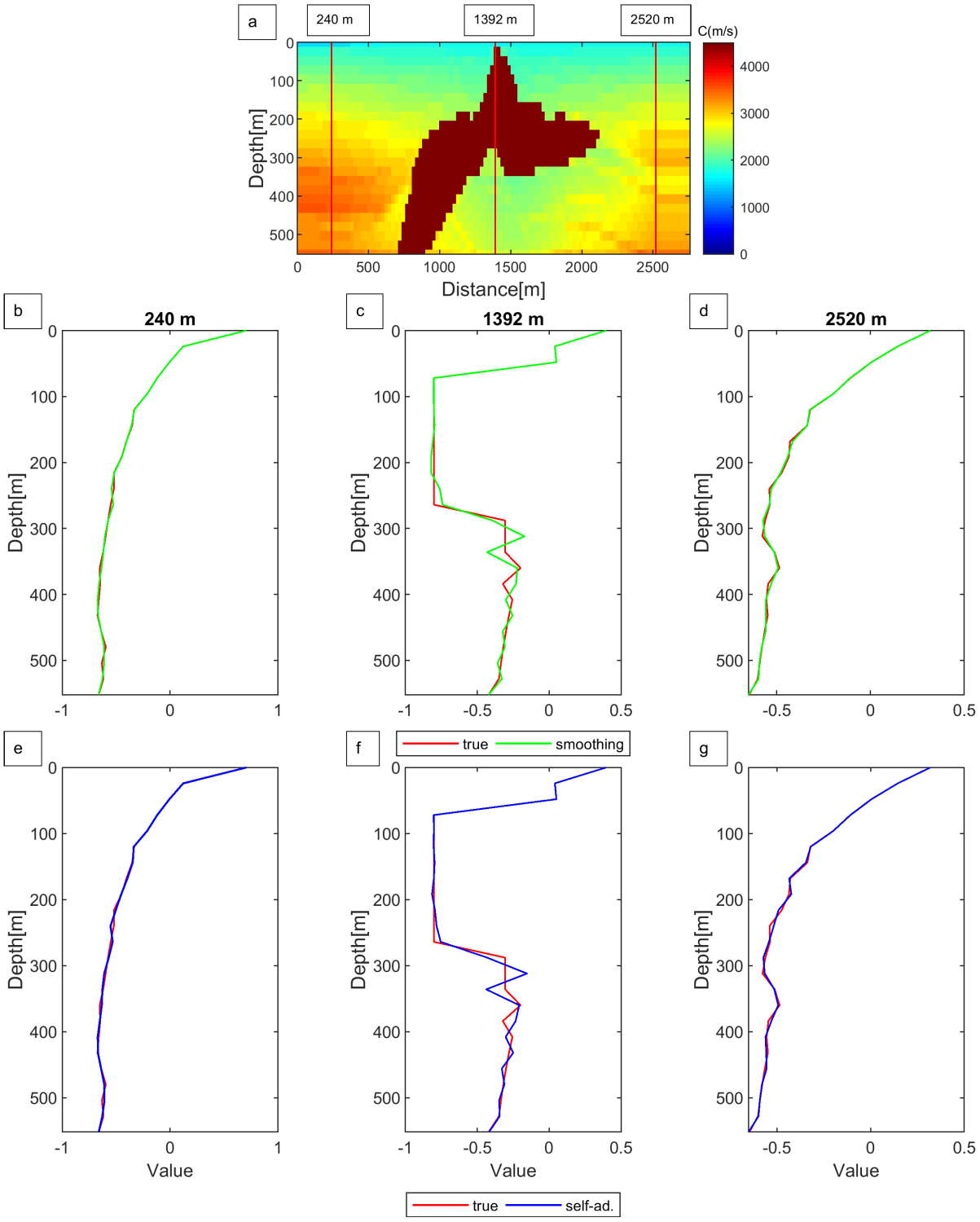

Figure 3.19: Model recoveries at different locations (profiles) within the baseline model obtained using the DBIT inversion algorithm combined with the smoothing (b, c and d) and self-adaptive (e, f and g) techniques.

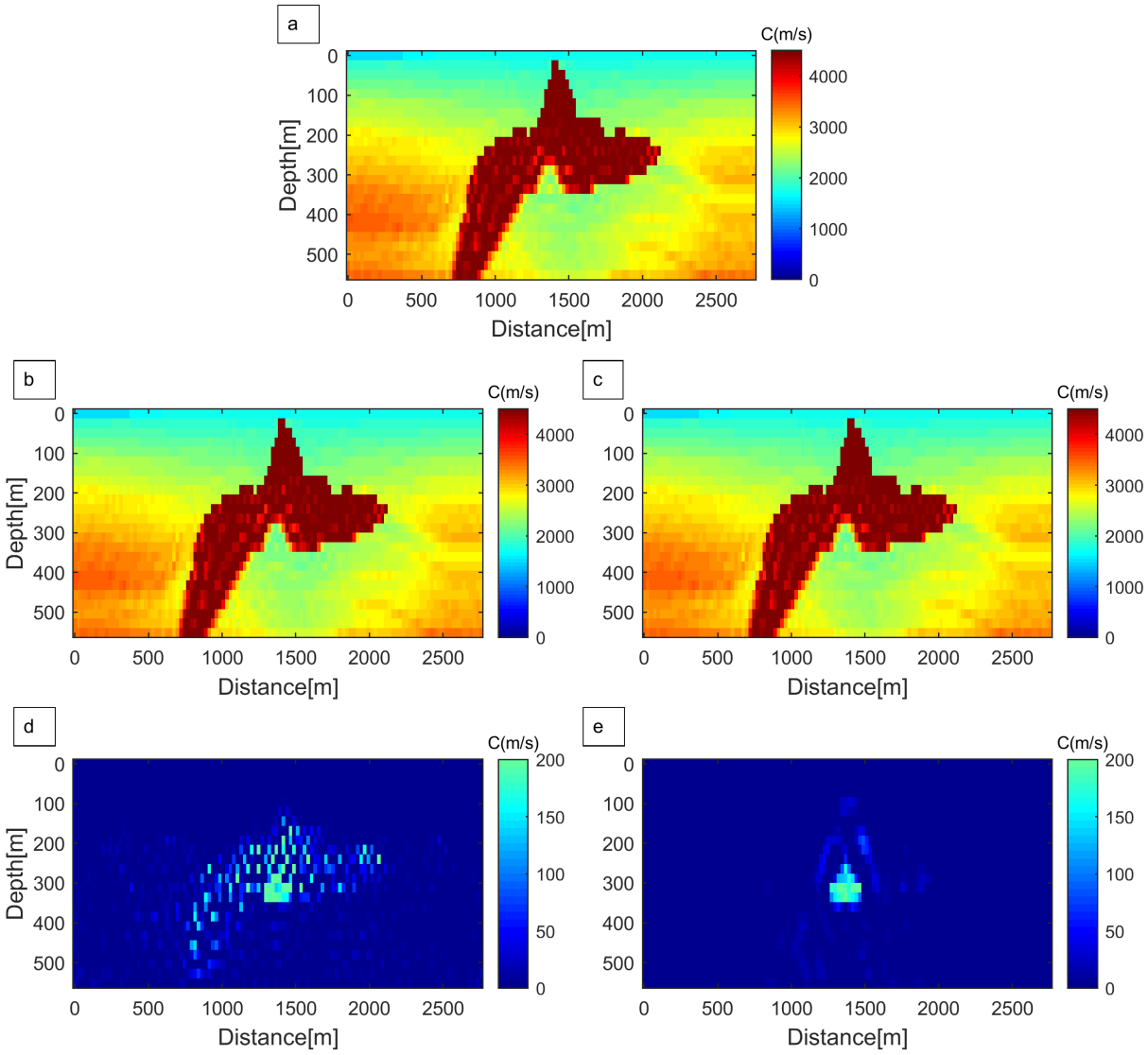

Figure 3.20: Model reconstructions obtained using the DBIT inversion algorithm. (a) The baseline model (based on the linearly increasing initial model); (b) the monitor model (based on the linearly increasing initial model); (c) the monitor model (based on the baseline model (illustration(a)); (d) the time-lapse model (obtained by the parallel difference strategy); (e) the time-lapse model (obtained by the sequential strategy). The frequencies used [1, 3, 5, 7, 9, 13, 15, 17 Hz]. The number of sources and receivers is 48 and 116 respectively. The time-lapse model corresponds to the low-contrast anomaly (200 m/s).

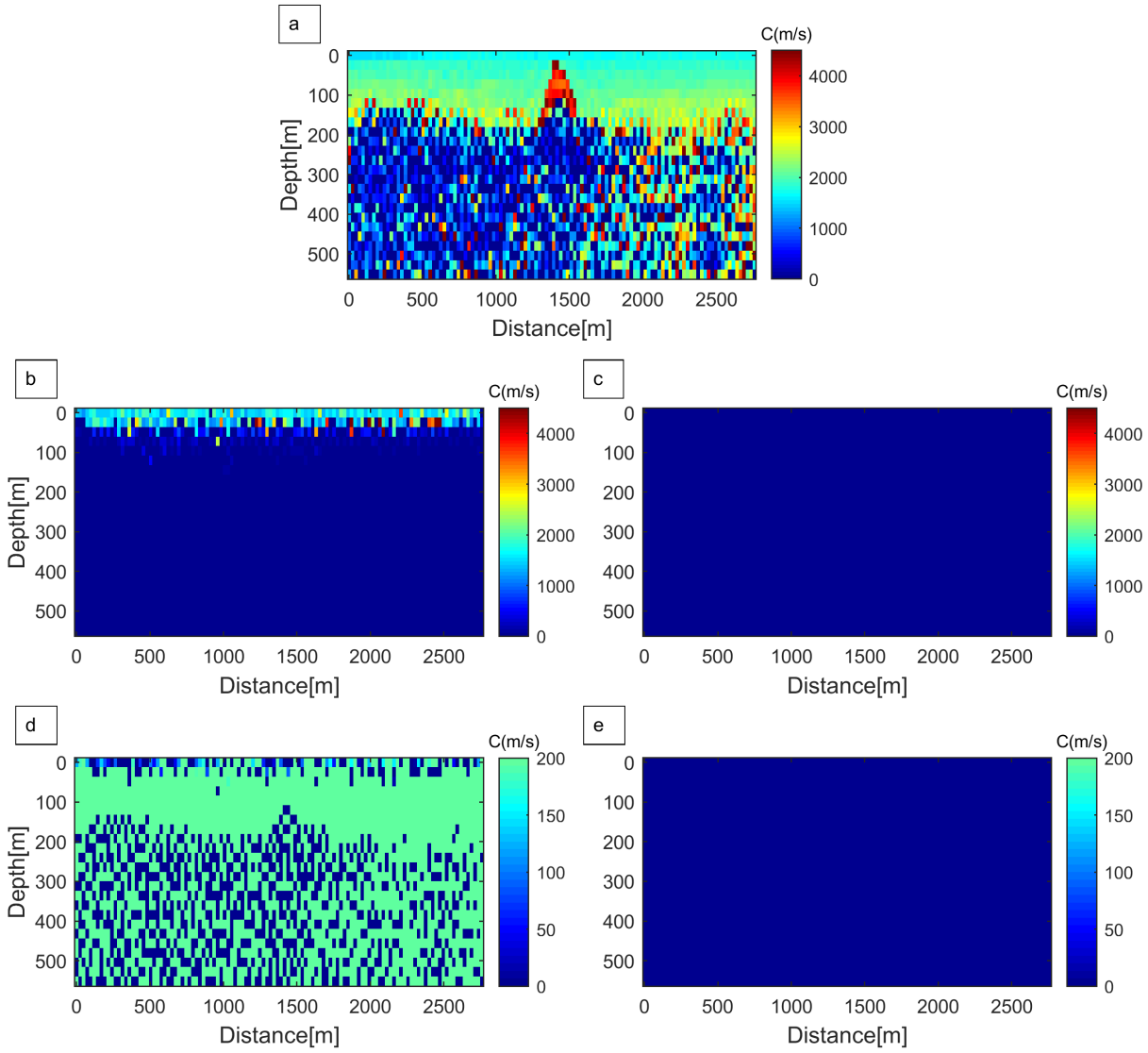

Figure 3.21: Similar to figure 3.20, but without ULF in the data.

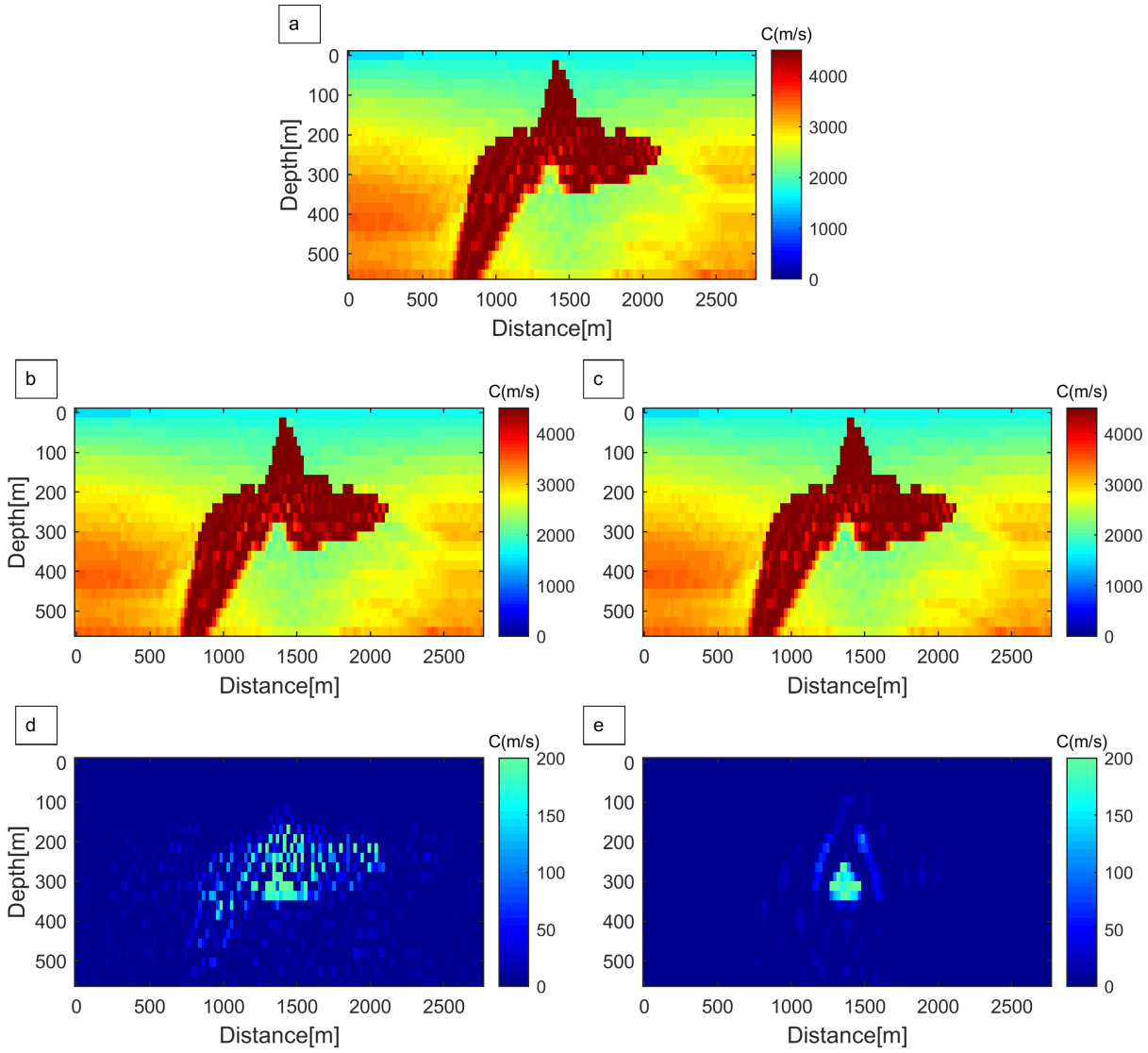

Figure 3.22: Similar to figure 3.20, but, without ULF in the data; and, with the smoothing technique applied.

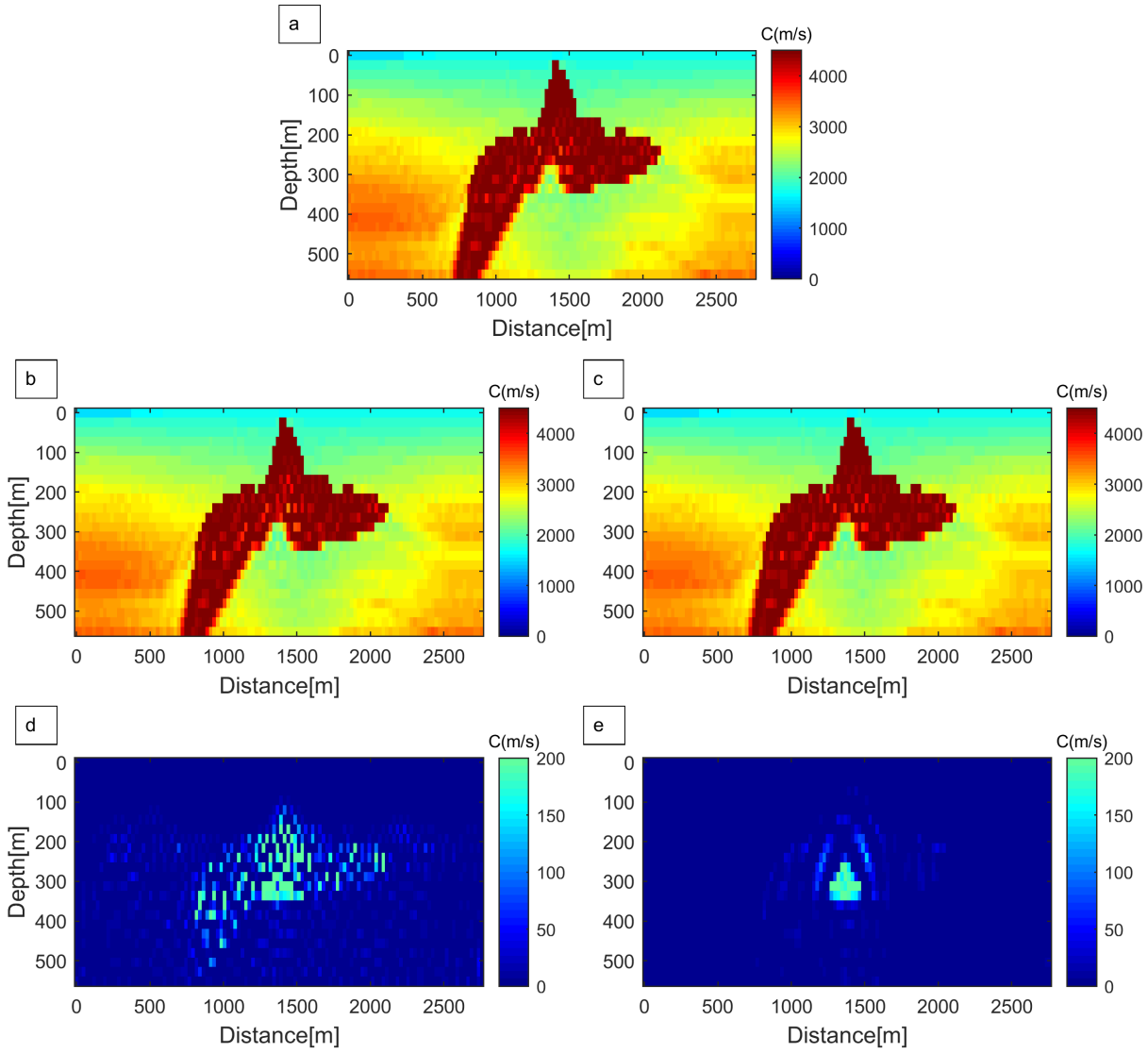

Figure 3.23: Similar to figure 3.20, but, without ULF in the data; and, with the self-adaptive technique applied.

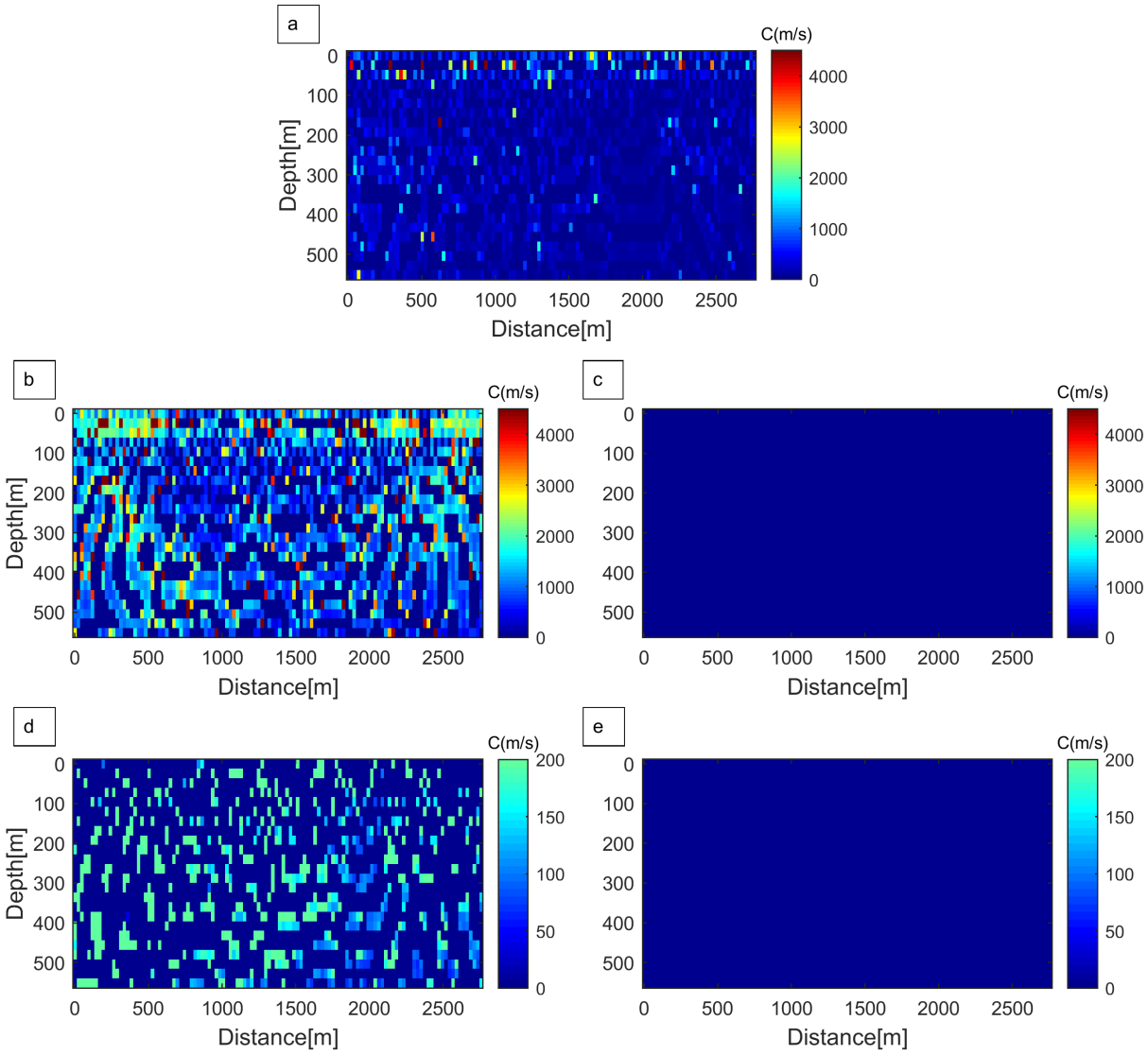

Figure 3.24: Similar to figure 3.20, but, without frequencies 1-6 Hz in the data; and, with the smoothing technique applied.

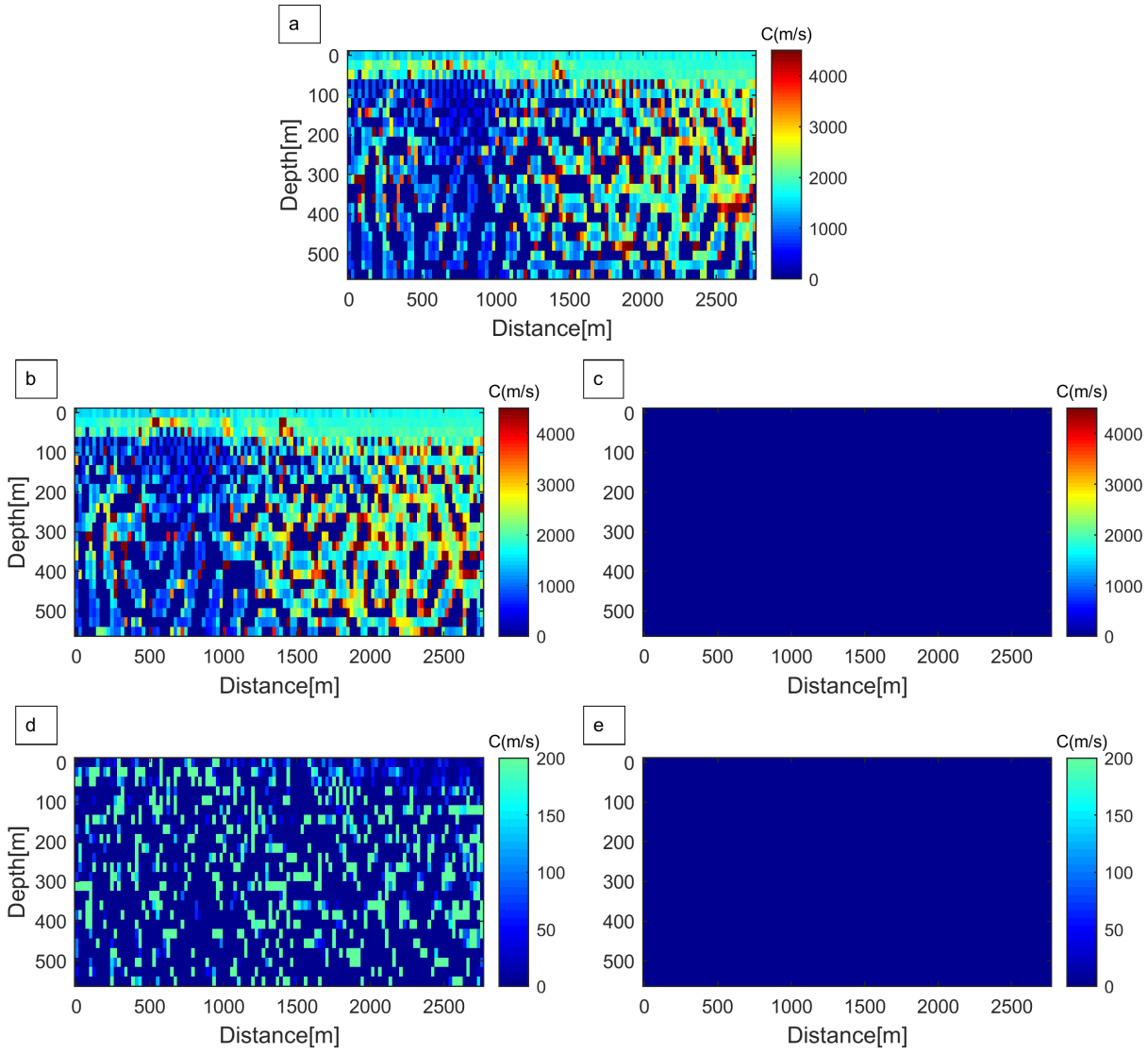

Figure 3.25: Similar to figure 3.20, but, without frequencies 1-6 Hz in the data; and, with the self-adaptive technique applied.

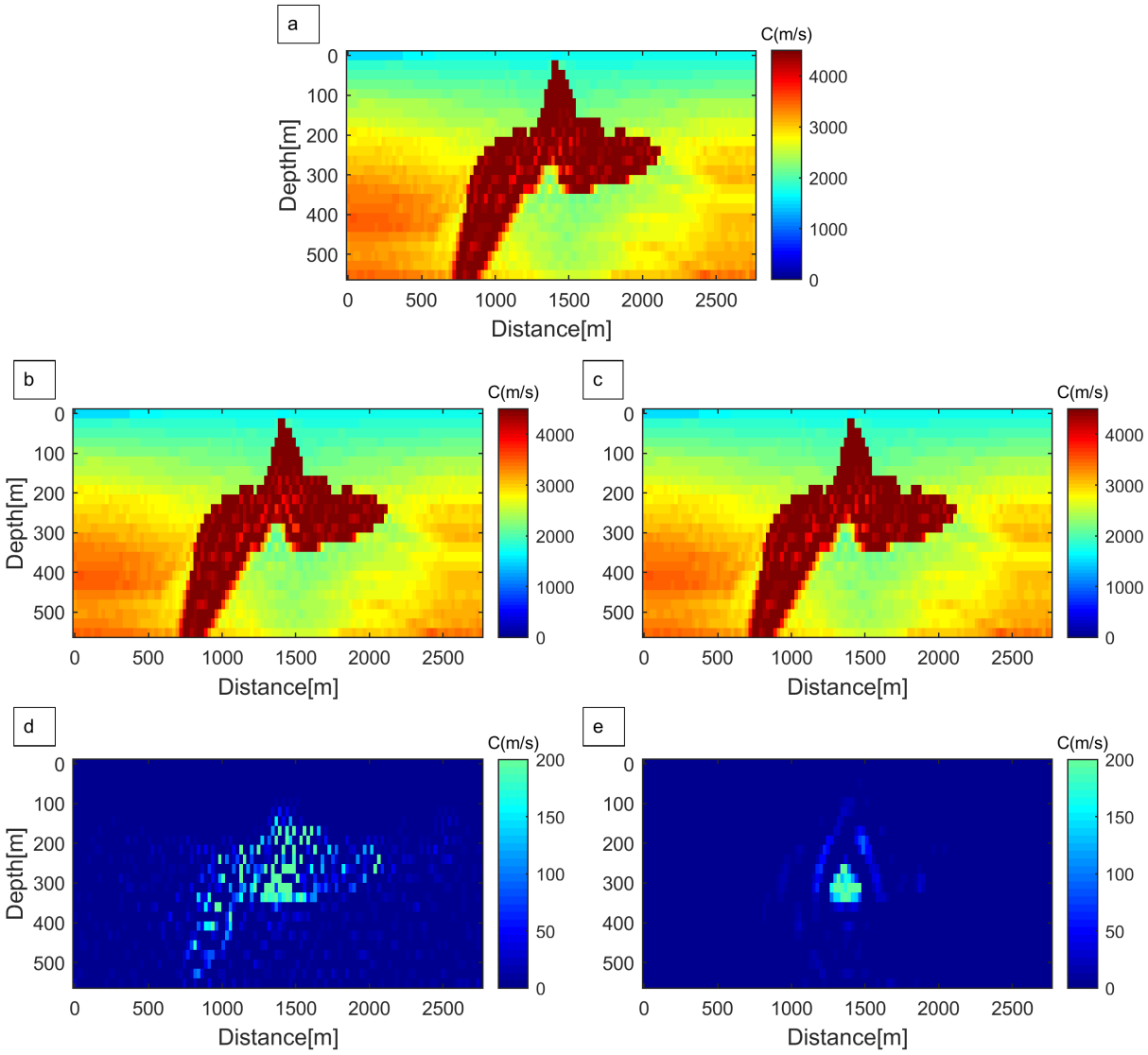

Figure 3.26: Similar to figure 3.20, but, without frequencies 1-6 Hz in the data; and with both the smoothing and self-adaptive techniques applied.

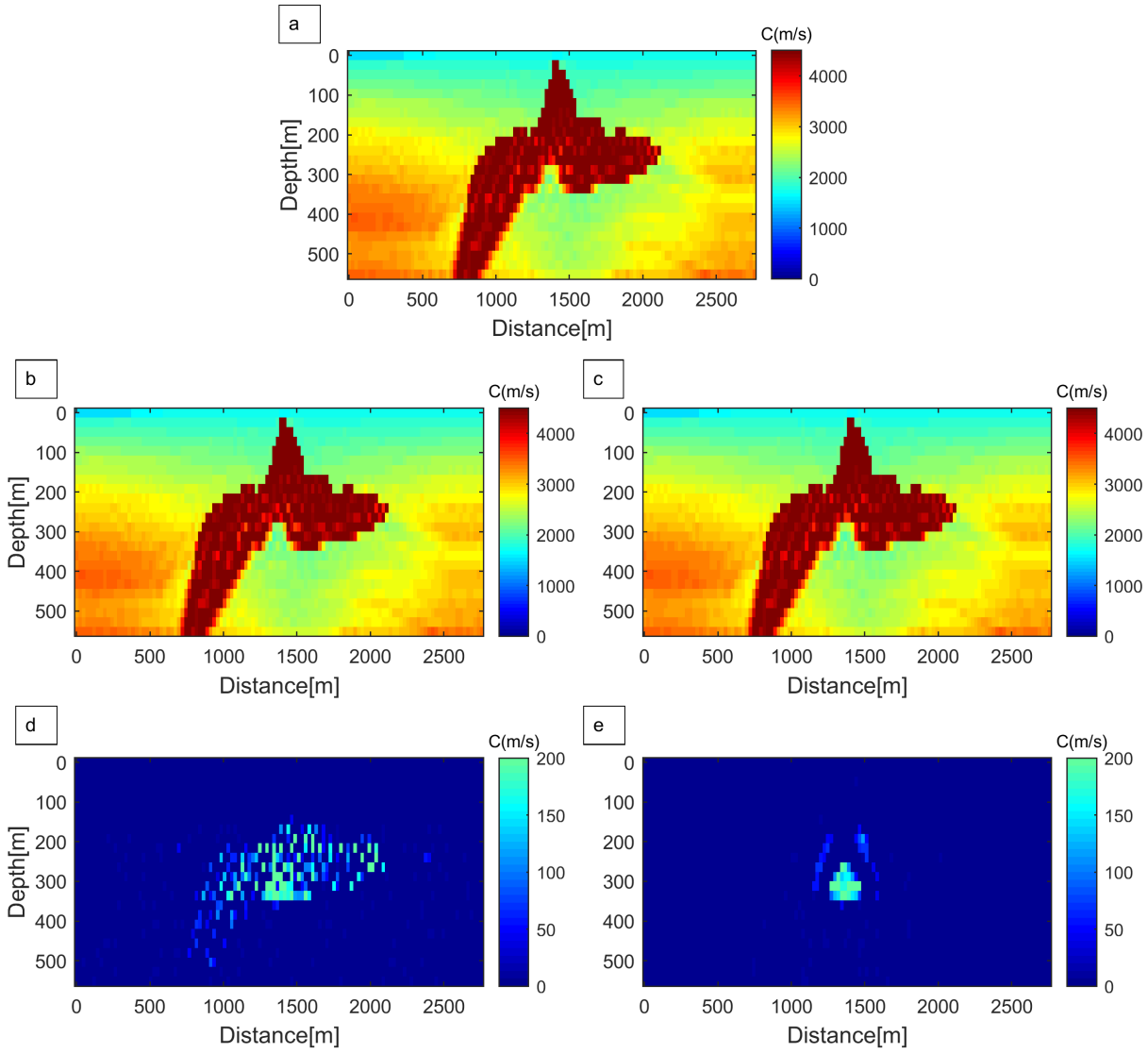

Figure 3.27: Similar to figure 3.20, but, without ULF in the data; and, with both the smoothing and self-adaptive techniques applied.

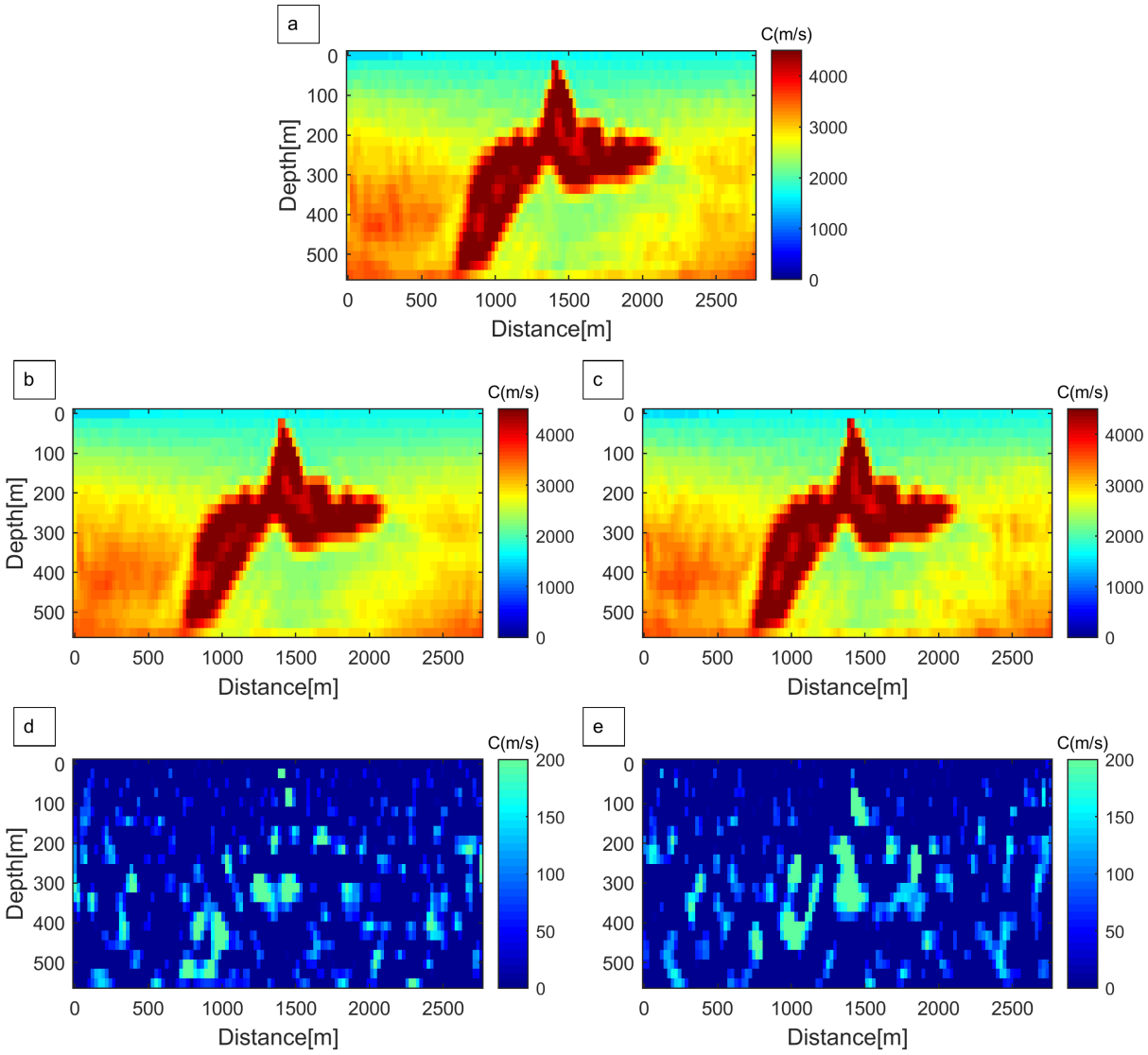

Figure 3.28: Similar to figure 3.20, but, without ULF in the data; with the smoothing and self-adaptive technique applied (the SNR equals to 20 dB).

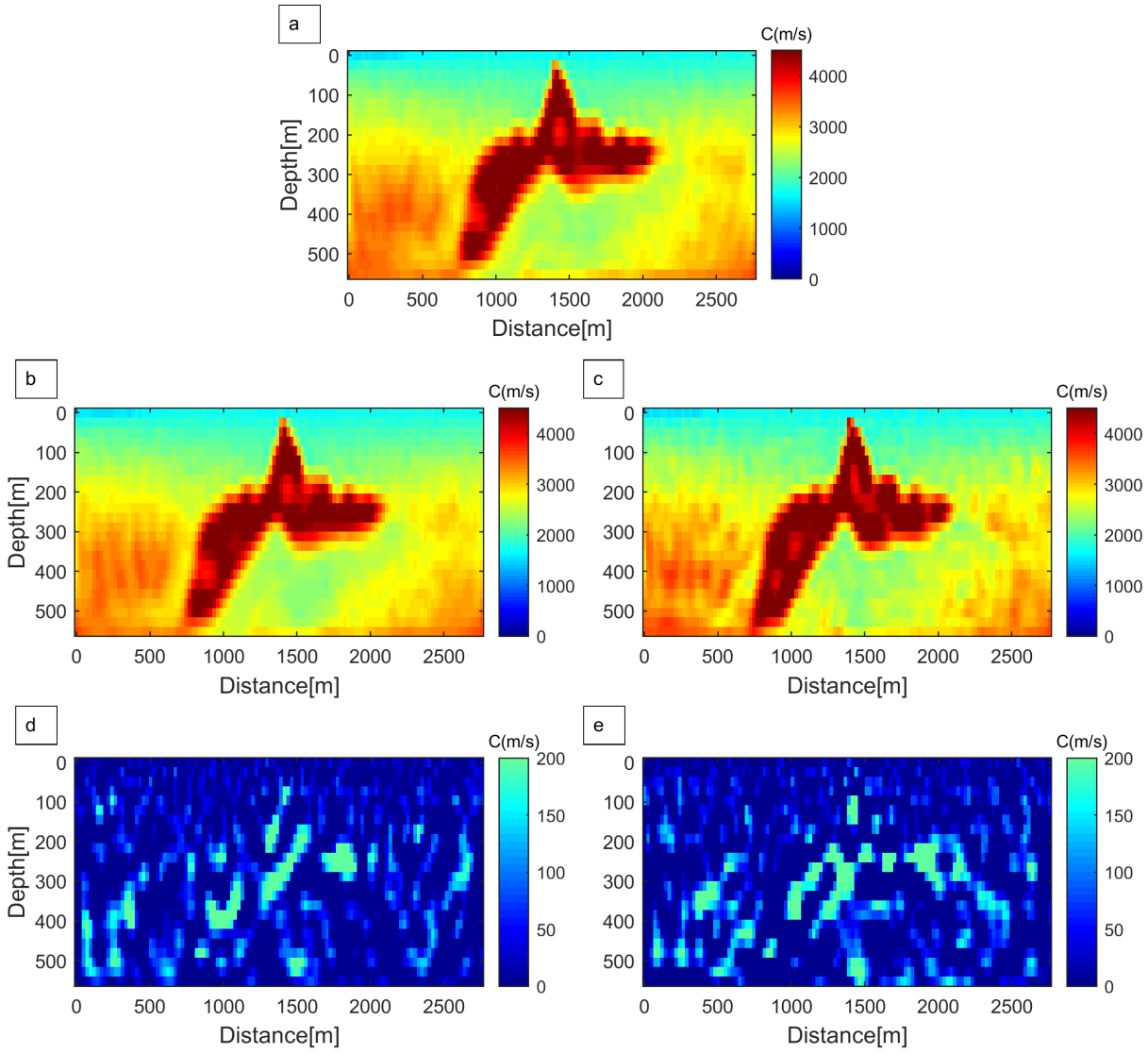

Figure 3.29: Similar to figure 3.20, but, without ULF in the data; with the smoothing and self-adaptive technique applied (the SNR equals to 5 dB).

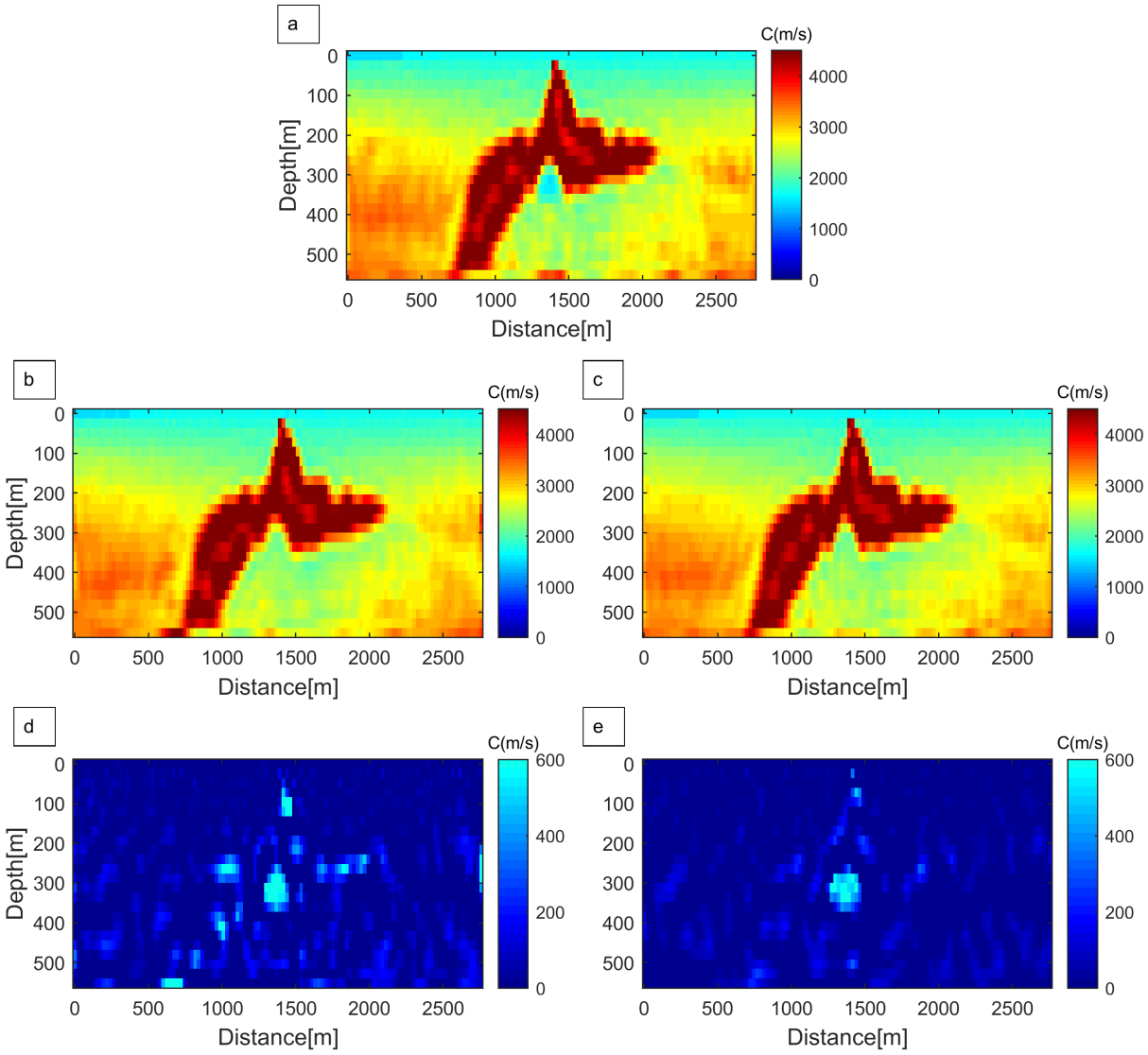

Figure 3.30: Similar to figure 3.20, but for high-contrast velocity anomaly (600 m/s) using the data with the SNR equals to 20 dB.

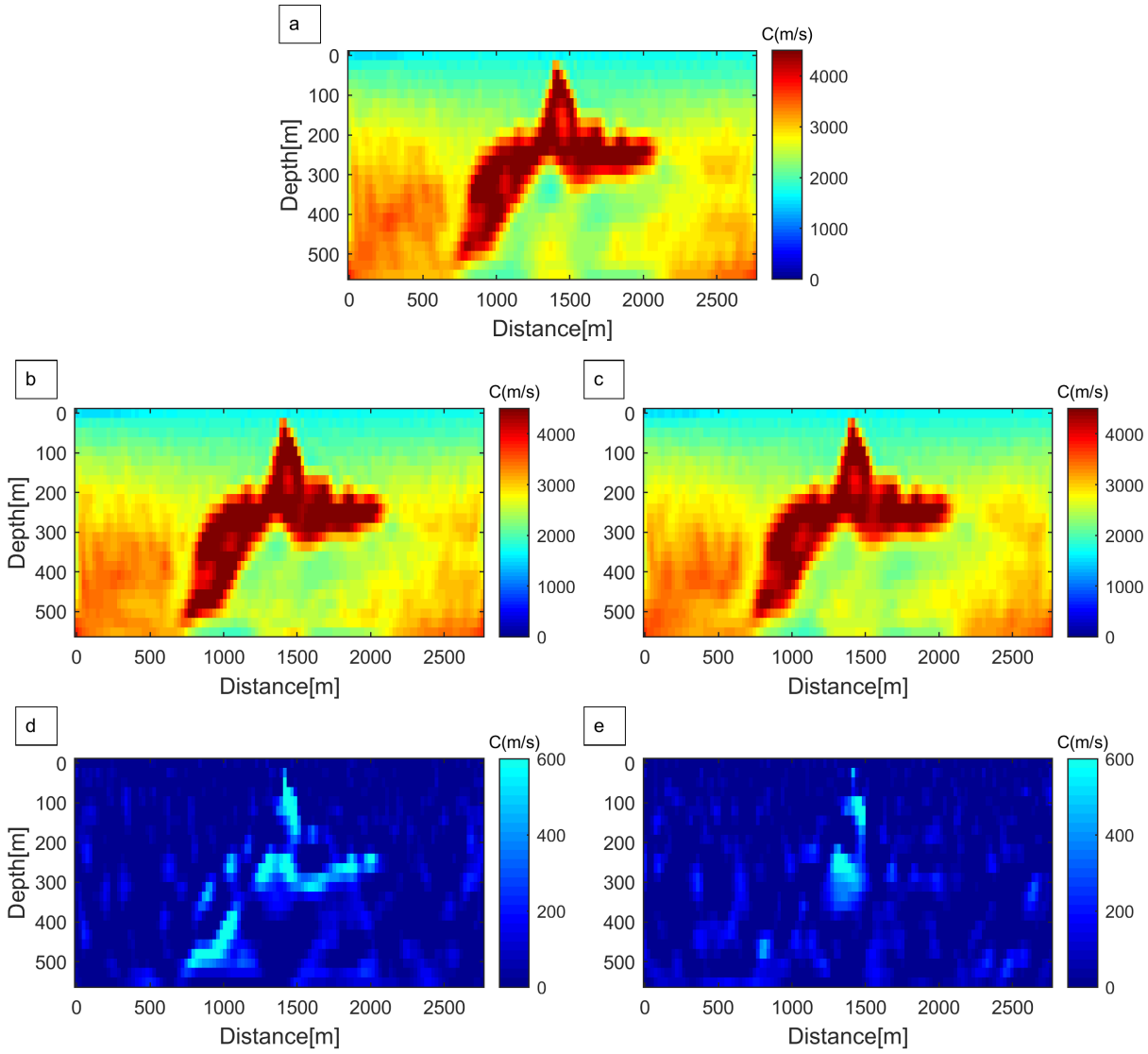

Figure 3.31: Similar to figure 3.20, but for high-contrast velocity anomaly (600 m/s) using the data with the SNR equals to 5 dB.

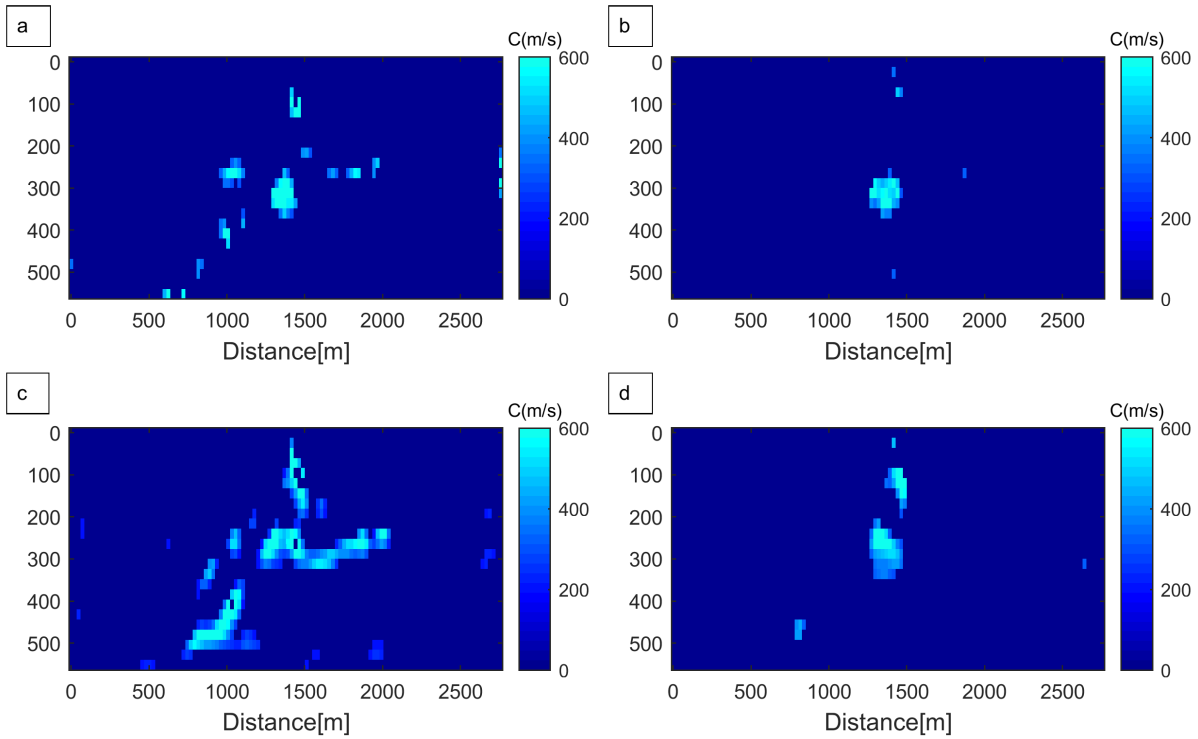

Figure 3.32: The time-lapse reconstructions from the noisy data in figures 3.30 and 3.31, but with the velocity differences outside the interval 300-900 m/s were removed. (a) and (c) obtained with parallel difference strategy (the SNR equal to 20 dB and 5 dB respectively); (b) and (d) obtained with parallel difference strategy (the SNR equal to 20 dB and 5 dB respectively).

Chapter 4

Conclusions and future work

Anyone who has never made a mistake has never tried anything new.

—Albert Einstein

4.1 Final summary

The main goal of this thesis was to investigate the application of time-lapse based DBIT FWI in strongly scattering media with different time-lapse strategeis. And this goal, as the author thinks, has been achieved.

In the modelling part of this research, using different 1D and 2D modelling examples, it was shown that the T-matrix integral equation method could be used as an exact modelling method. To check the accuracy of this modelling method, I compared data generated by the T-matrix approach with those obtained with the help of the finite-differnce time domain modelling technqie. The subsequent conclusion is consistent with the observations provided by Jakobsen and Ursin (2015).

A comparison between the time-lapse modelling data obtained by the T-matrix and those by DBA modelling method, showed the results were almost the same.

By experimenting with the conventional DBIT algorithm applied to strongly scattering media models (the resampled and modified SEG/EAGE salt model), it was demonstrated that this method did not work if the data were lacking ultra-low frequencies (1-4

Hz). The problem was solved when the traditional DBIT method was used together with smoothing and self-adaptive techniques. It was also shown that both techniques worked better in combination rather than separately.

Time-lapse model reconstructions corresponded with two different time-lapse scenarios (low-contrast and high-contrast) were analyzed and the following was discovered: when noiseless data were considered, the low-contrast model (velocity difference equals 200 m/s) was effectively recovered. In the case of noisy data, the low-contrast time-lapse anomaly from the reconstructed images was not identified. However, when noisy data were used to image high-contrast time-lapse model (velocity difference equals 600 m/s), the anomaly was successfully resolved. Therefore, to improve the low-contrast time-lapse model recovery, one could probably try to increase the number of receivers, shots or frequencies (or, even better, all of them together).

When I compared the parallel difference and sequential time-lapse inversion strategies, I found that the sequential approach worked faster and provided more accurate results.

The general conclusion is that, integrated with the smoothing and self-adaptive techniques, the DBIT inversion method can be considered as a powerful tool in solving time-lapse problems associated with strongly scattering media models.

4.2 Suggestions for future work

The DBIT-based time-lapse inversion method, when applied to the high-contrast model, shows very promising results. However, there is still plenty of room for further development.

One of the possible directions could be addressed to the model boundary problem. It would be interesting to see the inversion results obtained from data with applied boundary conditions when modelling. It would exclude any bias related to extra illumination of structures under high-contrast bodies with internal multiples. After implementing boundary conditions, one can try to reconstruct the time-lapse models with three time-lapse strategies: parallel difference, sequential and differential.

Another research direction of DBIT inversion could be, for example, the use of modelling data generated by the finite-difference frequency domain method (see, for example, Ajo-Franklin (2005). After applying appropriate boundary conditions, one could try to

generate data with only selected frequencies and, then, invert for model parameters. This would also solve the inverse crime problem providing additional test of the DBIT inversion method.

Additionally, the time-lapse DBIT-based FWI could be extended for the application to more realistic elastic cases with the aim to invert for both P- and S-velocity models. Finally, since originally geological media are three-dimensional, it would be interesting to adjust the DBIT inversion algorithm to 3D inversion cases. One could also try to experiment with real seismic data.

Finally, considering the scattering problem in general, one could try to solve the Lippmann-Schwinger equation more efficiently using iterative methods. One of the approaches in this direction is to use the modified Born series that guarantees the convergence independently of the model size and the strength of the velocity perturbation (Osnabrugge et al., 2016).

Appendix A

The finite-difference method

A.1 Derivation of the time domain finite-difference method for the acoustic wave equation

A derivation of the time domain finite-difference method is based on Schuster (2017). The following derivation considers a two-dimension case. The 2D acoustic wave equation can be written in the following form:

$$\frac{\partial^2(\mathbf{x}, t')}{\partial \mathbf{x}^2} - \frac{\partial^2 p\,(\mathbf{x}, t')}{\partial \mathbf{z}^2} - \frac{1}{c(\mathbf{x}, \mathbf{z})^2}\frac{\partial^2(\mathbf{x}, t')}{\partial\, t'^{\,2}} = f(\mathbf{x}, t'), \tag{A.1}$$

where $c(\mathbf{x}, \mathbf{z})$ denotes the velocity of the modelled medium, $p\,(\mathbf{x}, t')$ the pressure field, and $f(\mathbf{x}, t')$ the source function. Discretization of equation (A.1) in both time and space coordinates can be performed according to the following scheme:

$$(x, z, t) \leftrightarrow (i\,\Delta x, j\,\Delta z, t\,\Delta t). \tag{A.2}$$

Assuming i, j, and t in equation (A.2) are integers, the components of equation (A.1) can be expressed in discretized form as

$$p\,(x, z, t) \leftrightarrow p_{ij}^{t}, \tag{A.3}$$

$$f\,(x, z, t) \leftrightarrow f_{ij}^{t}, \tag{A.4}$$

and

$$c\,(x, z) \leftrightarrow c_{ij}. \tag{A.5}$$

For simplicity, we assume the sample interval in x and z directions are equal ($\Delta x = \Delta z$).

The second order central FD approximation scheme is given by

$$\frac{d^2 f(x)}{dx^2} \approx \frac{[f(x+\Delta x) - 2f(x) + f(x-\Delta x)]}{\Delta x^2}. \tag{A.6}$$

Applying (A.6) to equations (A.3), (A.4) and (A.5), we have

$$\frac{\partial^2 p}{\partial t^2} \approx \frac{[p_{ij}^{t+1} - 2p_{ij}^{t} + p_{ij}^{t-1}]}{\Delta t^2}, \tag{A.7}$$

$$\frac{\partial^2 p}{\partial x^2} \approx \frac{[p_{i+1j}^{t} - 2p_{ij}^{t} + p_{i-1j}^{t}]}{\Delta x^2} \tag{A.8}$$

and

$$\frac{\partial^2 p}{\partial z^2} \approx \frac{[p_{ij+1}^{t} - 2p_{ij}^{t} + p_{ij-1}^{t}]}{\Delta x^2}. \tag{A.9}$$

Assuming that initial conditions satisfy $p(x, t' = 0)$ and $\partial p(x, t' = 0)/\partial t'$ for all (x, z), such that the source wavefield is defined by the source function f_{ij}^{t}, the solution for the pressure field in dicretized form can be written as

$$\begin{aligned} p_{ij}^{t+1} = {} & 2p_{ij}^{t} + p_{i-1j}^{t} + \left(\frac{c_{ij}\,\Delta t}{\Delta x}\right)^2 [p_{i+1j}^{t} - 2p_{ij}^{t} + p_{i-1j}^{t}] \\ & + \left(\frac{c_{ij}\,\Delta t}{\Delta x}\right)^2 [p_{ij+1}^{t} - 2p_{ij}^{t} + p_{ij-1}^{t}] - (c_i j\,\Delta t \Delta x)^2 f_{ij}^{t}. \end{aligned} \tag{A.10}$$

Note, for the second-order finite-difference approximation to be accurate as a minimum of ten grid points per wavelength should be defined (Virieux, 1986). Additionally, the stable discretized solution requires for the time sampling interval to satisfy the following inequality (Ikelle and Amundsen, 2005):

$$\Delta t < 0.606\,\frac{\Delta x}{V_{max}}, \tag{A.11}$$

where Δt is the time interval, Δt is the space interval, and V_{max} is the maximum velocity in the model.

References

Ajo-Franklin, J. B. (2005). Frequency-domain modeling techniques for the scalar wave equation: An introduction.

Alkhalifah, T. and Y. Choi (2012). Taming waveform inversion non-linearity through phase unwrapping of the model and objective functions. *Geophysical Journal International 191*(3), 1171–1178.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux (2011). Sensitivity analysis of time-lapse images obtained by differential waveform inversion with respect to reference model. In *SEG Technical Program Expanded Abstracts 2011*, pp. 2482–2486.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux (2015). Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: which strategy? *Geophysical Prospecting 63*(1), 78–98.

Aster, R., B. Borchers, and C. Thurber (2005). *Parameter Estimation and Inverse Problems*. International geophysics series. Elsevier Science.

Auld, B. A. (1973). *Acoustic fields and waves in solids*.

Baeten, G., J. W. de Maag, R.-E. Plessix, R. Klaassen, T. Qureshi, M. Kleemeyer, F. ten Kroode, and Z. Rujie (2013). The use of low frequencies in a full-waveform inversion and impedance inversion land seismic case study. *Geophysical Prospecting 61*(4), 701–711.

Bansal, R. and M. K. Sen (2010). Ray-born inversion for fracture parameters. *Geophysical Journal International 180*(3), 1274–1288.

Bjørlykke, K. and J. Jahren (2015). *Sandstones and Sandstone Reservoirs*. Springer-Verlag Berlin Heidelberg.

Bleistein, N., J. K. Cohen, and J. W. Stockwell Jr (2001). *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*, Volume 13.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995). Multiscale seismic waveform inversion. *Geophysics 60*(5), 1457–1473.

Carcione, J. M., G. C. Herman, and A. P. E. ten Kroode (2002). Seismic modeling. *Geophysics 67*(4), 1304–1325.

Castagna, J. P. and M. M. Backus (1993). *Offset-dependent reflectivity : theory and practice of AVO analysis*.

Cerveny, V. (2001). *Seismic Ray Theory*. Cambridge University Press.

Chadwick, A., G. Williams, N. Delepine, V. Clochard, K. Labat, S. Sturton, M.-L. Buddensiek, M. Dillen, M. Nickel, A. L. Lima, R. Arts, F. Neele, and G. Rossi (2010). Quantitative analysis of time-lapse seismic monitoring data at the Sleipner CO2 storage operation. *29*(2), 170–177.

Chew, W. C. and Y.-M. Wang (1990). Reconstruction of two-dimensional permittivity distribution using the distorted born iterative method. *IEEE transactions on medical imaging 9*(2), 218–225.

Chi, B., L. Dong, and Y. Liu (2014). Full waveform inversion method using envelope objective function without low frequency data. *Journal of Applied Geophysics 109*, 36–46.

Ciric, I. R. and Y. Qin (1997). Self-adaptive selection of the regularization parameter for electromagnetic imaging. *IEEE Transactions on Magnetics 33*(2), 1556–1559.

Colton, D. and K. Rainer (1998). *Inverse Acoustic and Electromagnetic Scattering Theory* (2nd ed.).

Constable, S. C., R. L. Parker, and C. G. Constable (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics 52*(3), 289–300.

Dolson, J. (2016). *Understanding oil and gas shows and seals in the search for hydrocarbons*. Springer.

Eikrem, K. S., M. Jakobsen, and G. Nævdal (2017). Bayesian inversion of time-lapse seismic waveform data using an integral equation method. In *IOR 2017-19th European Symposium on Improved Oil Recovery*.

Eikrem, K. S., G. Nævdal, M. Jakobsen, and Y. Chen (2016). Bayesian estimation of reservoir properties—effects of uncertainty quantification of 4d seismic data. *Computational Geosciences 20*(6), 1211–1229.

Esser, E., L. Guasch, F. J. Herrmann, and M. Warner (2016). Constrained waveform inversion for automatic salt flooding. *The Leading Edge 35*(3), 235–239.

Farooqui, M., H. Hou, G. Li, N. Machin, T. Neville, A. Pal, C. Shrivastva, Y. Wang, F. Yang, C. Yin, et al. (2009). Evaluating volcanic reservoirs. *Oilfield Review 21*(1), 36–47.

Farquharson, C. G. and D. W. Oldenburg (2004). A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems. *Geophysical Journal International 156*(3), 411–425.

Fichtner, A. (2011). *Full Seismic Waveform Modelling and Inversion*.

Franklin, J. N. (1970). Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications 31*(3), 682–716.

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM review 34*(4), 561–580.

Huang, X., M. Jakobsen, K. S. Eikrem, and G. Nævdal (2018). A target-oriented scheme for efficient inversion of time-lapse seismic waveform data. In *SEG Technical Program Expanded Abstracts 2018*.

Ikelle, T. L. and L. Amundsen (2005). *Introduction to petroleum seismology*. Tulsa: Society of Exploration Geophysicists.

Jabbari, S. (2016). *A Theoretical Framework for Seismic Time-lapse Di erence AVO Analysis with Validation on Physical Modelling and Field Data*. Doctor of philosophy, Uiversity of Calgary.

Jackson, M. P. A. and M. R. Hudec (2017). *Influence of Salt on Petroleum Systems*, pp. 424–456. Cambridge University Press.

Jakobsen, M. (2012). T-matrix approach to seismic forward modelling in the acoustic approximation. *Studia Geophysica et Geodaetica* (56), 1–20.

Jakobsen, M. and B. Ursin (2015). Full waveform inversion in the frequency domain using direct iterative T-matrix methods. *Journal of Geophysics and Engineering* (12), 400–418.

Jakobsen, M. and R. S. Wu (2016). Renormalized scattering series for frequency-domain waveform modelling of strong velocity contrasts. *Geophysical Journal International 206*(2), 880–899.

Jakobsen, M. and R. S. Wu (2018). Accelerating the T-matrix approach to seismic full-waveform inversion by domain decomposition. *Geophysical Prospecting 66*(6), 1039–1059.

Johnston, D. H. (1997). A Tutorial on Time-Lapse Seismic Reservoir Monitoring. pp. 142–146.

Keller, J. B. (1969). Accuracy and Validity of the Born and Rytov Approximations. *J. Opt. Soc. Am. 59*(8), 1003–1004.

Kouri, D. and A. Vijay (2003). Inverse scattering theory: Renormalization of the Lippmann-Schwinger equation for acoustic scattering in one dimension. *Physical review. E, Statistical, nonlinear, and soft matter physics 67*, 46614.

Landrø, M. (2015). 4D Seismic. In K. Bjorlykke (Ed.), *Petroleum Geoscience:From Sedimentary Environments to Rock Physics* (2nd ed.)., pp. IX–508. Springer-Verlag Berlin Heidelberg.

Leveille, J. P., I. F. Jones, Z.-Z. Zhou, B. Wang, and F. Liu (2011). Subsalt imaging for exploration, production, and development: A review. *Geophysics 76*(5).

Levinson, H. W. and V. A. Markel (2016). Solution of the nonlinear inverse scattering problem by T-matrix completion. II. Simulations. *Phys. Rev. E 94*(4), 43318.

Lo, T.-w. and P. L. Inderwiesen (1994). *Fundamentals of Seismic Tomography*. Society of Exploration Geophysicists.

Malcolm, A. E., B. Ursin, and M. V. de Hoop (2009). Seismic imaging and illumination with internal multiples. *Geophysical Journal International 176*(3), 847–864.

Menke, W. (2012). *Geophysical data analysis: Discrete inverse theory*. Academic press.

Morse, P. M. and H. Feshbach (1953). *Methods of theoretical physics*. New York: McGraw-Hill.

Mulder, W. and R.-E. Plessix (2008). Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting 56*(6), 827–841.

Nguyen, P., M. J. Nam, and C. Park (2015). A review on time-lapse seismic data processing and interpretation. *Geosciences Journal 19*.

Norton, S. J. (1988). Iterative seismic inversion. *Geophysical Journal International 94*(3), 457–468.

Nowack, R. (2010). Seismic interferometry and estimation of the Green's function using Gaussian beams. *Earthquake Science 23*, 417–424.

Osnabrugge, G., S. Leedumrongwatthanakun, and I. M. Vellekoop (2016). A convergent born series for solving the inhomogeneous helmholtz equation in arbitrarily large media. *Journal of Computational Physics 322*, 113–124.

Ovcharenko, O., V. Kazei, D. Peter, and T. Alkhalifah (2018). Variance-based model interpolation for improved full-waveform inversion in the presence of salt bodies. *Geophysics 83*(5), R541–R551.

Patin, S. (2001, 03). Environmental impact of the offshore oil and gas industry. *Journal of Environmental Assessment Policy and Management 3*, 173–175.

Pratt, R., Z.-M. Song, P. Williamson, and M. Warner (1996). Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysical Journal International 124*(2), 323–340.

Pratt, R. G. (1990). Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging. *Geophysics 55*(5), 626–632.

Queißer, M. and S. C. Singh (2013). Full waveform inversion in the time lapse mode applied to CO2 storage at Sleipner. *Geophysical Prospecting 61*(3), 537–555.

Rabinovich, E. V., N. Y. Filipenko, and G. S. Shefel (2018, may). Generalized model of seismic pulse. *Journal of Physics: Conference Series 1015*, 052025.

Ramírez, A. C. and W. R. Lewis (2010). Regularization And Full-waveform Inversion: A Two-step Approach.

Ricker, N. (1953, jan). The Form and Laws of Propagation of Seismic Wavelets. *GEOPHYSICS 18*(1), 10–40.

Sayers, C. and S. Chopra (2009, may). Introduction to this special section: Seismic modeling. *The Leading Edge 28*(5), 528–529.

Schuster, G. T. (2017). *Seismic Inversion*. Society of Exploration Geophysicists.

Schutter, S. R. (2003). Hydrocarbon occurrence and exploration in and around igneous rocks. *Geological Society, London, Special Publications 214*(1), 7–33.

Shipp, R. M. and S. C. Singh (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International 151*(2), 325–344.

Sirgue, L. and R. G. Pratt (2004). Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics 69*(1), 231–248.

Snieder, R. (2004). *A Guided Tour of Mathematical Methods: For the Physical Sciences* (2 ed.). Cambridge University Press.

Speight, J. G. (2015). Chapter 6 - production. In J. G. Speight (Ed.), *Subsea and Deepwater Oil and Gas Science and Technology*, pp. 149 – 189. Boston: Gulf Professional Publishing.

Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics 51*(10), 1893–1903.

Taylor, J. R. (1972). *Scattering Theory: The quantum Theory on Nonrelativistic Collisions*. Wiley, New York.

Virieux, J. (1986). P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics 51*(4), 889–901.

Virieux, J. and S. Operto (2009, December). An overview of full-waveform inversion in exploration geophysics. *Geophysics 74*(6), WCC1 (online).

Vrahatis, M., G. Magoulas, and V. Plagianakos (2003). From linear to nonlinear iterative methods. *Applied Numerical Mathematics 45*(1), 59–77.

Wang, Y. (2016). *Seismic inversion: theory and applications*. John Wiley & Sons.

Yilmaz, Ö. (2001). *Seismic Data Analysis* (2 ed.). Tulsa: Investigations in Geophysics Society Of Exploration Geophysicists.

Zhdanov, M. S. (2015). *Inverse theory and applications in geophysics*, Volume 36. Elsevier.