IMAGE RECONSTRUCTION OF 2D VELOCITY MODELS IN SEISMIC IMAGING

A DISSERTATION

Submitted in partial fulfillment of the
requirement for the award of the degree

of

INTEGRATED MASTER OF TECHNOLOGY

in

GEOPHYSICAL TECHNOLOGY

By

SAGAR SINGH

DEPARTMENT OF EARTH SCIENCES
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
MAY, 2016

CANDIDATE'S DECLARACTION

I hereby declare that the work which is presented in this dissertation report titled ***"Image Reconstruction of 2D Velocity Models in Seismic Imaging***" in partial fulfillment of the requirements for the award of the degree of ***"INTEGRATED MASTER OF TECHNOLOGY"*** in ***GEOPHYSICAL TECHNOLOGY***, submitted to the Department of Earth Sciences, Indian Institute of Technology Roorkee is an authentic record of my own work carried out during the period from June 2015 to April 2016 under the supervision of Dr. S. Mukopadhyay, Professor, Department of Earth Sciences, Indian Institute of Technology Roorkee and Dr. Ali Ismet Kanli, Professor, Department of Geophysical Engineering, Istanbul University, Turkey.

I have not submitted the matter embodied in this dissertation report for the award of any other degree.

Date: **SAGAR SINGH** (11411024)

Place: IIT Roorkee Integrated M. Tech (Geophysical Technology)

CERTIFICATE

This is to certify that the above statement made by the candidate is correct to the best of my knowledge and belief.

I wish him a great success in life and bright future.

Date:

Place: IIT ROORKEE

Signature of co-guide	**Signature of guide**
Dr. Ali Ismet KANLI	Dr. Sagarika Mukopadhyay
Professor	Professor
Department of Geophysical Engineering	Department of Earth Sciences
Istanbul University, Turkey	IIT Roorkee, India

CERTIFICATE

I …..**SAGAR SINGH**…. hereby solemnly declare that the dissertation entitled …… **IMAGE RECONSTRUCTION OF 2D VELOCITY MODELS IN SEISMIC IMAGING**…… being submitted by me towards the partial fulfillment of the requirements for the award of ....**INTEGRATED M.TECH. GEOPHYSICAL TECHNOLOGY**…. degree is a record of my own work and that I have not copied the work of any other person(s) including published literature and material from any web site.

Where ever the work of other person(s) has been used, it has been duly acknowledged and quoted with proper reference to the original work.

I fully understand the implications of plagiarism and that if at any stage the above statement made by me is found to be incorrect; I shall be fully responsible for my act(s).

(Signature ……...)

Name …….SAGAR SINGH……...

Place……ROORKEE…….

Date…….

ACKNOWLEDGEMENT

I wish to express my deep sense of gratitude and sincere indebtedness to my supervisors **Dr. Sagarika Mukopadhyay**, Professor, Department of Earth Sciences, Indian Institute of Technology Roorkee; and co-supervisor **Dr. Ali Ismet Kanli**, Professor, Department of Geophysical Engineering, Istanbul University, Turkey, for their kind cooperation and encouragement that they have provided me for developing new ideas and tackle the various situations which came across while doing the dissertation work. Their undying determination to get the best out of their students served as inspiration for completion of this thesis.

I am grateful to the **Department of Earth Sciences, IIT Roorkee** and **Department of Geophysical Engineering, IU, Turkey** and the **administration** whose help and resources I have used from time to time, I thank the brains behind the scientific publications, earth science books and research papers whose guidance I have drawn on to make this thesis.

I convey my thanks to all my friends for their continuous support and help during different stage of my dissertation work.

Lastly, I will be always grateful to my parents whose blessings are always with me and acts as a beacon in finding the right path.

SAGAR SINGH

ABSTRACT

In this thesis, images of 2 D velocity models are reconstructed using Full Waveform Inversion which is based on the acoustic wave equation in both time and frequency domain. I have written MATLAB codes for Full Waveform Inversion (FWI) and then tested on few synthetic models. The main objective was to evaluate the capabilities of FWI to estimate velocities from the acquired synthetic data.

Standard seismic imaging tool such as Travel Time Tomography usually fail when applied to a complex geological structure. Indeed, the velocity contrast between the area of interest and the surroundings usually complicates the propagation of the seismic wave, thus making the imaging of the target body quite difficult. Full waveform Inversion (FWI) has shown the capabilities to image even in complicated subsurface in past years and emerged as a valuable seismic imaging tool which is now widely used in exploration industries.

In this thesis, firstly, the FWI code is tested on velocity models such as high velocity layer and low velocity layer models. These models represent a very familiar velocity structure found on near surface hence, dealing with these models is pretty reasonable. After successful testing, the FWI code was also tested on complicated Marmousi model in both time and frequency domain. One of the biggest contributions of the Marmousi model is that it demonstrated the limitations of first-arrival travel times in imaging complex media. Specifically, ***Geoltrain and Brac (1993)*** showed that multi-arrival travel times are needed in order to properly image the Marmousi model. In portions of the Marmousi model, or any other complex model, the first arrival is not necessarily the most energetic. Therefore, reflected energy from key horizons, such as the top of the reservoir, are not properly imaged by using only first arrival travel times (**Alkhalifah, 1998**).

The final results were then compared with the real velocity model to analyze the amount of recovered velocities. However, in order to initiate FWI an initial model is required which was derived from smoothing the true velocity model by using Gaussian smoother. This is an easy way to ensure an adequate starting model, as the FWI method is known to be sensitive on starting model.

The results in this master thesis demonstrate that image reconstruction done by FWI both in time and frequency domain matches very well with the true model used in this thesis. For the future, applying FWI to real data from more complex geological medium and developing a migration tool and test the effect of FWI on a migrated image, are interesting challenges.

CONTENTS

List of figures / List of Tables

CHAPTER 1

INTRODUCTION

Seismic imaging is a tool that bounces acoustic waves off subsurface rock structures to reveal possible crude oil and natural gas bearing formations. An ultrasensitive device called geophones is used to record those acoustic waves as they echo within the earth (Fig. 1.1). By studying the echoes, petroleum geologists seek to calculate the depth and structures of buried geologic formations. This analysis may help them identify hydrocarbon reservoirs hidden beneath the earth's surface.

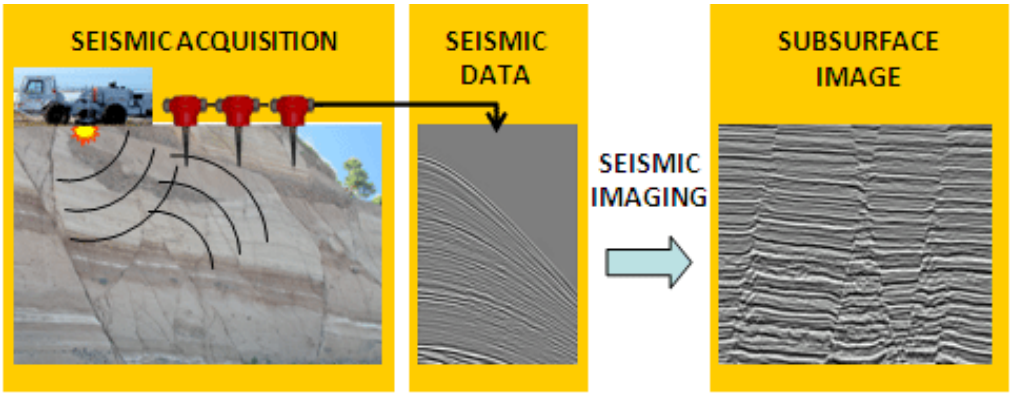


Figure 1.1: A Vibroseis source is used to image the subsurface using reflected events recorded by geophone. These records are called as seismogram. (**Source:** *http://www.eni.com/en_IT/innovation-technology/technological-focus/seismic-imaging/seismic-imaging.shtml*)

Recently, seismic depth imaging has become the preferred tool for seismic imaging of complex geological structures. This method starts with a smooth background velocity model before determining the short wavelength components of the structure by pre-stack depth

migration. The pre-stack depth migration method aims to locate the reflectors to their correct position in space. To be successful, this technique requires precise knowledge of the velocities of the ground. The estimation of a velocity model is a difficult process, especially when velocities changes rapidly in both lateral and vertical directions. Full-waveform inversion (FWI) is a seismic imaging tool for building such velocity models **(Thomassen, 2008)**.

In this thesis, I wrote FWI's MATLAB code to investigate the capabilities and the limits of this method from synthetic data. In addition, the theory and a general introduction about full-waveform inversion are also presented in this thesis.

1.1 Presentation of the synthetic 2D velocity models

Although, a few velocity models such as low velocity layer and high velocity layer models are used to test code, however, the main aim was to recover the velocities from the complicated Marmousi model.

The original Marmousi model was built to resemble an overall continental drift geological setting. Numerous large normal faults were created as a result of this drift. The geometry of the Marmousi is based somewhat on a profile of the North Quenguela through the Cuanza basin **(*Versteeg, 1993*)**.

The original model (Fig. 1.2) is nearly 9 Km long and 3 Km in depth. The target zone of reservoir located at a depth of about 2500 m. The model contains many reflectors, steep dips, and strong velocity variations in both the lateral and the vertical direction (with a minimum velocity of 1500 m/s and a maximum of 5500 m/s). The test model, shown in Figure 1.3, is however defined by 722 m long and 200 m in depth, re-sampled on 1 x 1m regular grid space (with a minimum velocity of 1.5 m/s and a maximum of 4.7 m/s). This modified model will be

referred as True Model. This necessary modification was done in order to have a reasonable size of model to handle. Both the Time Domain and Frequency domain FWI schemes used in this study will use Fig. 1.3 to estimate the capabilities of their algorithms.

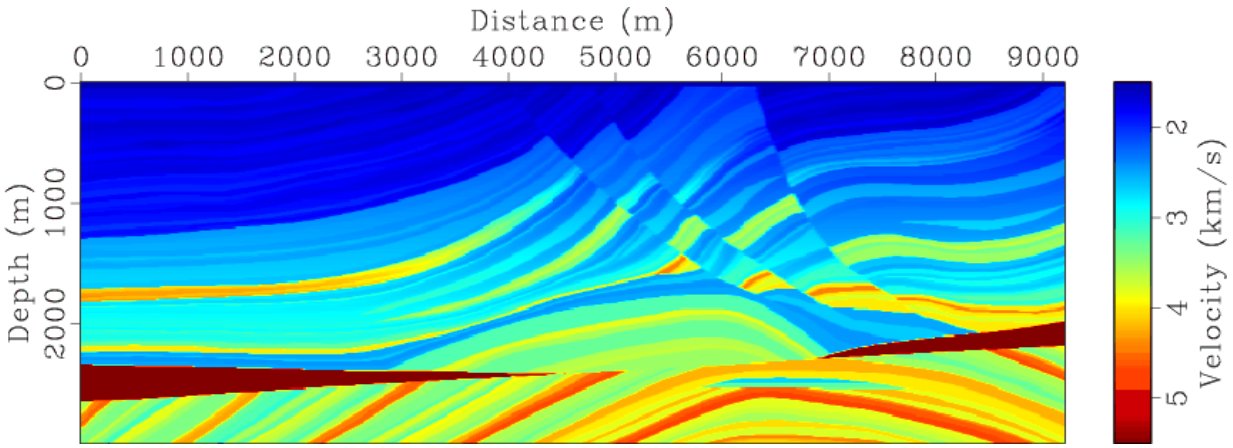


Figure 1.2: Original Marmousi Model (**Source:** Institut Français du Pétrole *http://www.ifp.fr/IFP/en/aa.htm*)

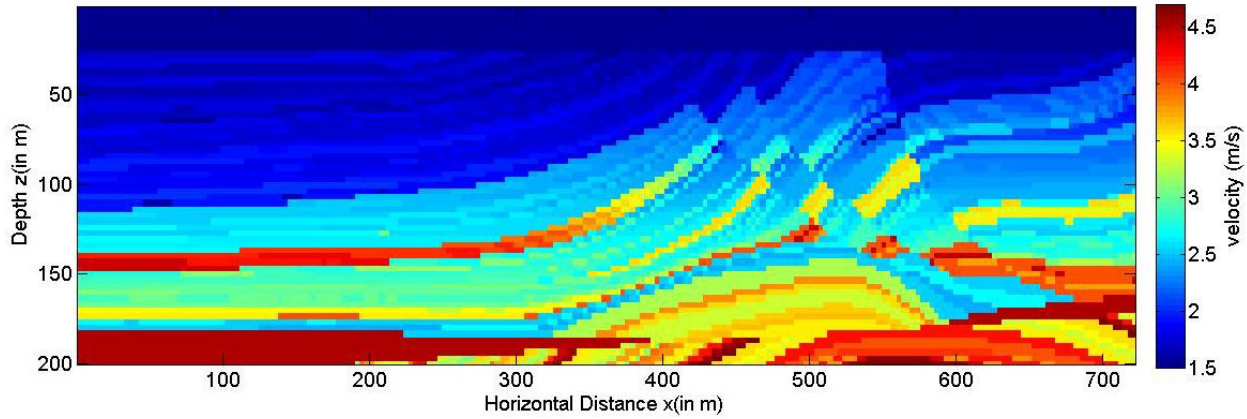


Figure 1.3: Mordified Marmousi Model

In Fig. 1.3 contains a low velocity water layer on top of the model from 1-25-m. Marmosui model is known as a benchmark model to test migration algorithm.

1.2 Contents of the Thesis

In Chapter 2 of this thesis the theory and general introduction to the inverse problems are discussed. The seismic inverse problem is presented, and why this problem is of any interest is discussed. Forward and Inverse modelling process of the seismic inverse problem are also discusses which includes some technique on how to solve it.

Chapter 3 presents the theory and methodology on how to perform Full Waveform Inversion in time domain. First, the reliability of the method was tasted by performing the algorithm on two simple models. Then numerical computation of this method was performed on Marmousi model, i.e benchmark model to test migration algorithm. The smoothed version of Marmousi model was taken as initial model. Inversion results are analyzed and discussed.

Chapter 4 discusses the frequency domain version of the Chapter 3. The benchmark model and initial model were kept same in this method as were in Chapter 3. Some key elements of the MATLAB code are also presented. Importance of frequencies and its sensitivity are included in this chapter. Inversion results are analyses and discussed.

Conclusions are presented in Chapter 6.

CHAPTER 2

GEOPHYSICAL INVERSE THEORY AND PROBLEMS

An inverse problem in Geophysics is the process of calculating from the set of observations the parameters or factors that produced them: for example, calculating the Earth's density from acquired data of its gravity field. This is the inverse of a forward problem, which starts with the causes and then calculates the results.

2.1 INTRODUCTION

Seismic inversion, in Geophysics, is the process of transforming seismic reflection data into a quantitative rock-property such as velocity, density etc. of a reservoir. The seismic inverse problem can be conceptually formulated as shown in Figure 2.1:

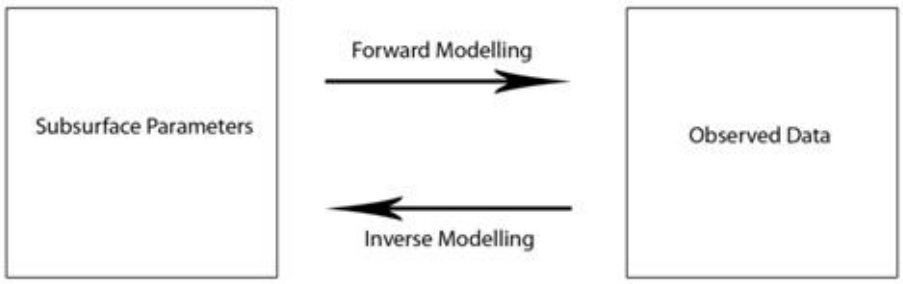


Figure 2.1: A simple sketch representing the relationship between forward and inverse problems.

The transformation from the observed data to subsurface parameters (or vice versa) is a result of the interaction of source wave-field with the medium that we wish to infer parameters

about. In other words, the transformation is the physics that relates the subsurface parameters to the observed data. In case of Seismic, our physical system is seismic waves (from active or passive source), the governing equation is wave equation, parameter is velocity (or density), and the observed data is pressure wave-field. The relationship between data and model can either be linear or non-linear as described by the following Equations **(Chadan et al., 1977)**

$$\boldsymbol{d} = G\boldsymbol{m} \qquad (2.1)$$

$$\boldsymbol{d} = g(\boldsymbol{m}) \qquad (2.2)$$

Where G (linear) and g (non-linear) are operators describing the explicit relationship between the observed data, $\boldsymbol{d}$, and model parameters $\boldsymbol{m}$, which is known.

Equation (2.1) and (2.2) represents the forward modeling process i.e. computing the data from a known model, while Equation (2.3) represents the backward or inverse modeling process if g is invertible.

$$\boldsymbol{m} = g^{-1}(\boldsymbol{d}) \qquad (2.3)$$

Inverse problems are typically ill posed in-contrast to the well posed problems which are more typical where the model parameters are known when modeling physical situation. Of the three conditions given by ***Hadamard (1902)*** (stability of the solutions, uniqueness, existence) the condition of stability is most usually violated.

Non uniqueness of inverse problems is caused by the non-linear relation between data and the model parameters of the inverse problem and Geophysical problems are almost always non-linear **(*Chabert, 2007*)**. The solution of an inverse problem is therefore not necessarily the true solution, but a solution that explains the data.

In Equation (2.3), if g is weekly non-linear, the problem can be linearised and a solution can be obtained by local methods (discussed in section 2.3.1). However, in case of strong non-

linear relationship between data and model parameters, global methods (discussed in section 2.3.2) are used to find a solution for inverse problem.

2.2 THE FORWARD MODELING

Assuming a model and then mathematically simulate seismic energy spreading through that model at a given time and position is forward modeling. In this thesis, I used acoustic wave equation for solving forward modeling. Equation (2.4) represents the non-homogeneous acoustic wave equation. u represents the pressure field, f represent the source term and m is the velocity model.

$$\frac{1}{\boldsymbol{m}(x,z)^2}\frac{\partial^2 \boldsymbol{u}(x,z,t)}{\partial t^2} - \nabla^2 \boldsymbol{u}(x,z,t) = \boldsymbol{f}(x,z,t) \qquad (2.4)$$

Many methods are known to solve this equation. In my thesis, the forward modeling is done in both time and frequency domain using Equation (2.4).

In time domain, the above equation can be solved by numerical methods. I used finite difference approximation method for discretization of Equation (2.4) where each node values contributed to the complete wave-field. Although, it is just simple calculations and suits for complex medium, however calculations can be time consuming. Computation for time domain forward modelling will be discussed in Chapter 3.

In frequency domain, Equation (2.4) must be transformed into frequency domain Helmholtz equation (Equation 2.5).

$$\frac{\omega^2}{\boldsymbol{m}(x,z)^2}\boldsymbol{u}(x,z,\omega) + \nabla^2 \boldsymbol{u}(x,z,\omega) = \boldsymbol{f}(x,z,\omega) \qquad (2.5)$$

where ω is the angular frequency. This system can also be represented as shown below

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \quad \text{where } \boldsymbol{A} = \left(\frac{\omega^2}{\boldsymbol{m}(x,z)^2} + \nabla^2\right) \qquad (2.6)$$

Now A can be inverted and wave-field u can be computed directly. Computation for frequency domain forward modelling will be discussed in Chapter 4.

2.3 INVERSE MODELLING

Inverse modelling consists of finding a set of model parameters ($\boldsymbol{m}$) that predicts the observed data ($\boldsymbol{d}_{obs}$) **(*Pratt et. al., 1998*).** Based on the current model parameters estimations, a forward modelling routine compute a data set ($\boldsymbol{d}_{cal}$). Our aim is to make the data residual ($\boldsymbol{r} = \boldsymbol{d}_{obs} - \boldsymbol{d}_{cal}$) small as possible and model parameters can be updated to reduce $\boldsymbol{r}$. Figure 2.2 illustrates the updation of model parameter as in an optimization process.

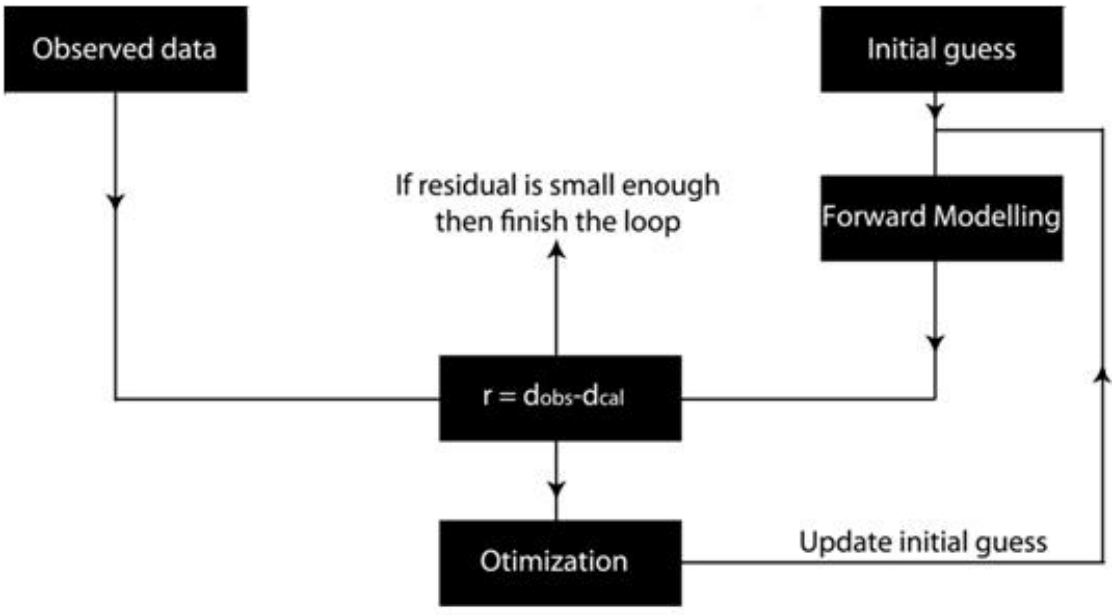


Figure 2.2: An illustration of the inverse modelling in an optimization process. Optimization process starts with an initial guess of the model parameters and calculated data is generated by forward modelling. If residual is small enough then the process will be terminated else the residual will undergo the specified optimization scheme to update the model parameters.

The residual, that we are trying to minimize is generally described by the l^2 (least square) norm of the error between is the observed and calculated data (Equation (2.7)).

$$\boldsymbol{O}(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{d}_{obs} - \boldsymbol{d}_{cal}\|^2 = \frac{1}{2}(\boldsymbol{r}'\boldsymbol{r}) \qquad (2.7)$$

Where $\boldsymbol{r}'$ is the transpose conjugate of $\boldsymbol{r}$. $\boldsymbol{O}(\boldsymbol{m})$ is known as objective function.

The model space consists of many numbers of possible models. If the number of distinct values of a parameter is k and total number of parameter is n then number of possible models in a model space is given by k^n. In case of the Earth, the model space consists of infinite number of models. Hence, we look for the model which can explain most parameters of the model. The inverse modelling process can solve by either linearising the problem and then solve the problem or using global optimization methods.

2.3.1 Linearised Methods

As discussed in section 2.1, if the relationship between the data and the model parameters is weekly non-linear then inverse problem might be solved by linearising the equation. The common way to linearise the problem is by using Born approximation.

In the framework of the Born approximation, we assume that updated model $\boldsymbol{m}$ can be written as the sum of the starting model $\boldsymbol{m}_0$ plus a perturbation model $\Delta\boldsymbol{m}$: $\boldsymbol{m} = \boldsymbol{m}_0 + \Delta\boldsymbol{m}$ ***(Virieux and Operto, 2009)***. A second order Taylor-Lagrange expression of objective function in the vicinity of $\boldsymbol{m}_0$ gives

$$\boldsymbol{F}(\boldsymbol{m}) = \boldsymbol{F}(\boldsymbol{m}_0 + \Delta\boldsymbol{m})$$

$$= \boldsymbol{F}(\boldsymbol{m_0}) + \sum_{j=1}^{M} \frac{\partial \boldsymbol{F}(\boldsymbol{m}_0)}{\partial \boldsymbol{m_j}} \Delta\boldsymbol{m_j} + \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M} \frac{\partial^2 \boldsymbol{F}(\boldsymbol{m}_0)}{\partial \boldsymbol{m}_j \partial \boldsymbol{m}_k} \Delta\boldsymbol{m}_j \Delta\boldsymbol{m}_k + O(\boldsymbol{m}^3) \qquad (2.8)$$

The higher order can be neglected, and equation (2.8) can be represented as

$$\boldsymbol{F}(\boldsymbol{m}_0 + \Delta\boldsymbol{m}) - \boldsymbol{F}(\boldsymbol{m_0}) = \sum_{j=1}^{M} \frac{\partial \boldsymbol{F}(\boldsymbol{m}_0)}{\partial \boldsymbol{m_j}} \Delta \boldsymbol{m_j} \qquad (2.9)$$

The above means that the relationship between the residual ($\boldsymbol{r} = \boldsymbol{F}(\boldsymbol{m}_0 + \Delta\boldsymbol{m}) - \boldsymbol{F}(\boldsymbol{m_0})$) and the model perturbation as be written as

$$\boldsymbol{r} = \boldsymbol{J}\Delta\boldsymbol{m} \qquad (2.10).$$

where $\boldsymbol{J}$ is the partial derivative of the objective function with respect to the model parameters. Equation (2.10) is now a linear equation, with $\boldsymbol{J}$ as the linear operator.

If it is somehow difficult to solve the linearised problem directly (as in Equation (2.10)) then it can be solved iteratively. The most common iterative local optimization methods are Gradient method, Newton method and Guess Newton method.

Gradient Method

Gradient method is simple iterative minimization technique. There are a variety of gradient methods available such as steepest decent, conjugate gradient and non-linear conjugate gradient. However, they all starts with a direction (negative of gradient) and cost function can always be reduced by going in that direction **(*Stekl, 1997*)**. If *k* is the current iteration, the gradient is described as

$$\nabla \boldsymbol{F}_k = \frac{\partial \boldsymbol{F}_k}{\partial \boldsymbol{m}_k} = Real(\boldsymbol{J}^t \boldsymbol{r}') \qquad (2.11)$$

where $\boldsymbol{J}^t$ is transposed Jacobian matrix (or sensitivity matrix), which is partial derivate of the objective function with respect to model parameters, as described in Equation (2.9).

In the gradient method, the model is updated according to the following equation (Equation 2.12)

$$\boldsymbol{m}_k = \boldsymbol{m}_{k-1} - \alpha_{k-1}\nabla \boldsymbol{F}_{k-1} \qquad (2.12)$$

where k is the current iteration number and α is the step length. Since, gradient provides no notion of scale, it need to be scaled by multiplying it with a proper step length (α), best found via line search algorithm **(*Pratt et. al. 1998*)**.

The gradient method is known to converse slowly and that it can be unstable **(*Ravaut et al., 2004*)**. To improve that, regularization, preconditioning can be used **(*Virieux and Operto, 2009*)**.

Newton's Optimization Method

Taking the derivative of Equation (2.8) with respect to the model parameter $\boldsymbol{m}_l$ results in

$$\frac{\partial F(\boldsymbol{m})}{\partial \boldsymbol{m}_l} = \frac{\partial F(\boldsymbol{m}_0)}{\partial \boldsymbol{m}_l} + \sum_{j=1}^{M} \frac{\partial^2 F(\boldsymbol{m}_0)}{\boldsymbol{m}_j \boldsymbol{m}_l} \Delta \boldsymbol{m}_j \qquad (2.13)$$

The minimum of the objective function in the vicinity of $\boldsymbol{m}_0$ is reached when the first derivative of the objective function vanishes. This gives:

$$\Delta \boldsymbol{m} = \boldsymbol{H}^{-1} \frac{\partial F(\boldsymbol{m}_0)}{\partial \boldsymbol{m}} \qquad (2.14)$$

where $\boldsymbol{H} = \left[\frac{\partial^2 \boldsymbol{F}(\boldsymbol{m}_0)}{\partial \boldsymbol{m}^2}\right]$ is called the Hessian matrix. In terms of Jacobian matrix Hessian can be represented as Equation (2.15).

$$\boldsymbol{H} = Real(\boldsymbol{J}^t \boldsymbol{J}') + Real\left(\left[\left(\frac{\partial \boldsymbol{J}^t}{\partial \boldsymbol{m}_1}\right)\boldsymbol{r}' \left(\frac{\partial \boldsymbol{J}^t}{\partial \boldsymbol{m}_2}\right)\boldsymbol{r}' \ldots\ldots \left(\frac{\partial \boldsymbol{J}^t}{\partial \boldsymbol{m}_n}\right)\boldsymbol{r}'\right]\right) = \boldsymbol{H}_a + \boldsymbol{R} \qquad (2.15)$$

Here, $\boldsymbol{H}_a$ is the approximated Hessian matrix. Since, $\boldsymbol{R}$ predicts the presence of first order multiple energy in the gradient, it works as a de-multiple operator **(*Pratt et al., 1998*)**.

Computation of the complete Hessian matrix is a difficult process. This process is simplified in Gauss-Newton Method.

Gauss-Newton Method

The computation for full Hessian involves time consuming inversions of very large matrices and a very difficult process. However, Gauss-Newton simplifies the process and uses only approximated Hessian matrix (Equation (2.16)) when updating to next model parameters (2.17). If the objective function is close to global minimum then $\boldsymbol{R}$ term contributes less to the overall solution **(*Pratt et al., 1998*)**.

$$\boldsymbol{H}_a = Real(\boldsymbol{J}^t \boldsymbol{J}') \qquad (2.16)$$

$$\boldsymbol{m}_k = \boldsymbol{m}_{k-1} - \boldsymbol{H}_a{}^{-1} \nabla \boldsymbol{F}_{k-1} \qquad (2.17)$$

In almost all cases, $\boldsymbol{H}_a$ is not invertible and a regularization term is added (Equation (2.18)).

$$\boldsymbol{m}_k = \boldsymbol{m}_{k-1} - (\boldsymbol{H}_a + \beta \boldsymbol{I})^{-1} \nabla \boldsymbol{F}_{k-1} \qquad (2.18)$$

Here, β is the weighing factor. For larger values of β, the above method approaches to gradient method and hence, can also include a step length, α, as for the gradient method **(*Sirgue, 2003*)**.

Apart from the above methods, Broyden-Fletcher-Goldfarb-Shanno (BFGS), Limited memory BFGS, Symmetric Rank 1 (SR1) have gained much attention and proved beneficial for faster convergence in limited time.

2.3.2 Global Methods

If the relationship between the data and the model parameters is strongly non-linear then global methods can serve the purpose to find a solution which can explain the model parameters. Few popular global optimization techniques are discussed below.

Monte Carlo Method

Monte Carlo method has different structures but they all tend to follow the similar pattern: it specify the model space, chooses a pool of models from the model space, runs forward modelling, and chooses the model that has the lowest objective function as the representative model of a system.

Simulated Annealing

In simulated annealing global optimization method, the model parameters are randomly perturbed. The change in the objective function caused by the perturbation is then computed to see if the perturbation is acceptable or not. If the objective function increases, the possibility for accepting the perturbation decreases. If this method converges, iteratively the model will become better **(*Collins and Kuperman, 1992*)**.

Genetic Algorithm

In genetic algorithm, a small number of models are chosen from the model space. The best models (parents) from the chosen model then form new models (children) by mutation, cross-over of the bit-string describing the model parameters. In turn, children replace the weakest models from the chosen model. Following this iterative strategy, the chosen models will gather towards a minimum of the objective function **(*Gerstoft, 1993*)**.

Apart from these, there are many global algorithms such as Artificial Neural Network, tunneling algorithm etc which are known to provide a solution efficiently that explains the model parameters. However, global methods are very dependent on fast forward modelling algorithm

and simple model parameterization because most time is spent on huge amount of forward modelling computation **(*Causse, 1998*)**. Now, computers are getting faster and then necessity of keeping simple model parameterization might decline with time. Therefore, global methods are probably the future in search for suitable solution for seismic inverse problem **(*Thomassen, 2008*)**.

CHAPTER 3

IMAGE RECONSTRUCTION IN TIME DOMAIN

In this chapter, the process of reconstructing the image using Full Waveform Inversion in time domain is discussed. The algorithms are written in MATLAB and are initially tested for few simple synthetic models i.e. low velocity layer model and high velocity layer model. After successful recovery of velocities from both models, the code was then run on complicated Marmousi model and results are discussed later in this chapter.

3.1 Introduction to Full Waveform Inversion

Full waveform inversion (FWI) is a non-linear data fitting process based on the seismic waveform data to estimate the model parameters, which usually appear as coefficients in a wave equation (**Liao and Cao, 2013**). FWI requires computation at three steps: solving the wave equation (also known as forward modelling) with the initial model, gradient computation of the objective function which is the half of the squared norm of the difference between the calculated and observed data, and lastly updating the model parameters with a valid optimization technique **(*Tarantola, 1984; Lailly, 1983; Virieux, 2009; Fichtner, 2011*)**. The numerical method (finite difference approximation) for the forward problem is discussed in section 3.2. Gradient based optimization techniques are utilized in most of the seismic FWI techniques to update the model parameters so for each iteration the gradient is needed. The efficiency and accuracy of the gradient computation determines the efficiency of the FWI method **(*Liao and Cao, 2013*)**. **Chavent (1974)**, introduced the adjoint state method in the theory of inverse problems, which soon became popular for providing an efficient and effective method for computing the gradient of the objective function in FWI **(*Plessix, 2006; Fichtner, 2011*)**.

As discussed in section 2.3, in the FWI workflow, there are several optimization techniques that can be used for updating the model parameters in order to minimize the objective function. In the Gauss-Newton method, where the approximated-Hessian matrix is needed to form is generally considered computationally costly. In fact, the full Newton method based on forming complete Hessian matrix (Eq. 2.15) is currently not being used in realistic FWI due to high computational cost **(*Virieux, 2009*)**. In practice, the simple alternative is to replace the inverse of the Hessian matrix with a carefully chosen step length. Such modification results in the so called gradient method, which are widely used in solving the gradient-based optimization problem **(*Liao and Cao, 2013*)**.

3.1.1 Forward Modelling in Time Domain

Forward modelling is a necessary step for solving any inverse problem. The governing equation for FWI in this thesis is 2D acoustic wave equation (see Equation (2.4)), where Laplacian operator is given by Equation (3.1).

$$\nabla^2 \boldsymbol{u}(x,z,t) = \frac{\partial^2 \boldsymbol{u}(x,z,t)}{\partial x^2} + \frac{\partial^2 \boldsymbol{u}(x,z,t)}{\partial z^2} \qquad (3.1)$$

The Laplacian operator can be approximated with the central difference operator. The approximation can either be five point or nine point approximation.

$$\nabla^2 \boldsymbol{u}_j^n \approx \frac{\boldsymbol{u}_j^{n+1} - 2\boldsymbol{u}_j^n + \boldsymbol{u}_j^{n-1}}{\Delta x^2} + \frac{\boldsymbol{u}_{j+1}^n - 2\boldsymbol{u}_j^n + \boldsymbol{u}_{j-1}^n}{\Delta z^2} \qquad five\ point\ approximation \qquad (3.2)$$

$$\nabla^2 \boldsymbol{u}_j^n \approx \frac{\left(-\frac{1}{12}\boldsymbol{u}_j^{n+2} + \frac{4}{3}\boldsymbol{u}_j^{n+1} - \frac{30}{12}\boldsymbol{u}_j^n - \frac{1}{12}\boldsymbol{u}_j^{n-2} + \frac{4}{3}\boldsymbol{u}_j^{n-1}\right)}{\Delta x^2} +$$

$$\frac{-\frac{1}{12}\boldsymbol{u}_{j+2}^n + \frac{4}{3}\boldsymbol{u}_{j+1}^n - \frac{30}{12}\boldsymbol{u}_j^n - \frac{1}{12}\boldsymbol{u}_{j-2}^n + \frac{4}{3}\boldsymbol{u}_{j-1}^n}{\Delta z} \qquad nine\ point\ approximation \qquad (3.3)$$

where n and j are x and y coordinate respectively as shown in Fig. 3.1.

In my code, in order to reduce computing time, the grid spacing was kept equal ($\Delta x = \Delta z$). As one would expect, the nine point approximation is more accurate, but is slower. Its main advantage is that it covers broader bandwidth. However, in my code I used five-point approximation which is much faster to compute.

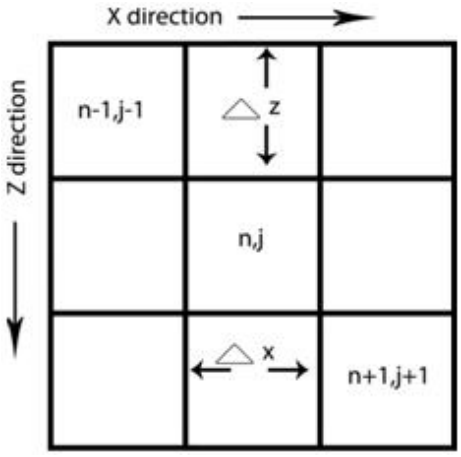


Figure 3.1: The computational grid of the approximations to the Laplacian oprator. (after: **Youzwishen and Margrave, 1999**)

Finally, each approximation scheme has a stability condition **(*Lines et al., 1998*)**. The stability condition for five and nine point approximations are respectively:

$$\Delta t \leq \sqrt{\frac{1}{2}} \frac{\Delta x}{\boldsymbol{m}_{max}} \qquad (3.4)$$

$$\Delta t \leq \sqrt{\frac{3}{8}} \frac{\Delta x}{\boldsymbol{m}_{max}} \qquad (3.5)$$

Here Δt is the temporal sampling rate, Δx is the grid spacing and $\boldsymbol{m}_{max}$ is maxing value of the model parameter (i.e. velocity in case of acoustic wave equation).

The time derivative is also calculated by a five point central difference approximation scheme.

$$\frac{\partial^2 \boldsymbol{u}(x,z,t)}{\partial t^2} \approx \frac{\boldsymbol{u}(t+\Delta t) - 2\boldsymbol{u}(t) + \boldsymbol{u}(t-\Delta t)}{\Delta t^2} \qquad (3.6)$$

By substituting Equations (3.6) and (3.2) in Equation (2.4), one can solve for wave-field at time $t+1$.

$$\boldsymbol{u}(t+\Delta t) = 2\boldsymbol{u}(t) - \boldsymbol{u}(t-\Delta t) + \frac{\Delta t^2}{\Delta \boldsymbol{m}_{n,m}{}^2}\left(\frac{\boldsymbol{u}_j^{n+1} - 2\boldsymbol{u}_j^n + \boldsymbol{u}_j^{n-1}}{\Delta x^2} + \frac{\boldsymbol{u}_{j+1}^n - 2\boldsymbol{u}_j^n + \boldsymbol{u}_{j-1}^n}{\Delta x^2}\right) + \boldsymbol{f} \qquad (3.7)$$

Equation (3.7) shows that the wave-field at time $t+\Delta t$ can be estimated by knowing the wave-field at time t and $t-\Delta t$. This process is known as time stepping.

In Equation (3.7) $\boldsymbol{f}$ is point source term (Equation (3.8)) and can be estimated by Ricker wavelet which is negative second derivative of a Gaussian (see Equation 3.9).

$$\boldsymbol{f}(x,z,t) = \boldsymbol{w}(t)\delta(x-x_s)\delta(z-z_s) \qquad (3.8)$$

$$\boldsymbol{w}(t) = (1 - 2\pi^2 v_0^2 (t-t_0)^2) exp(-\pi^2 v_0^2 (t-t_o)^2) \qquad (3.9)$$

Here, $\boldsymbol{w}$ is the time profile and δ indicates point source, v_0 is the peak frequency of Ricker wavelet. It is important that this function is causal ($\boldsymbol{w}(t)$ =0 for t<=0), so I introduced a time shift t_0. We can guarantee causality by setting $t_0 = 6/(\pi v_0\sqrt{2})$. Fig. 3.2 represents the Ricker wavelet at $v_0 = 20Hz$.

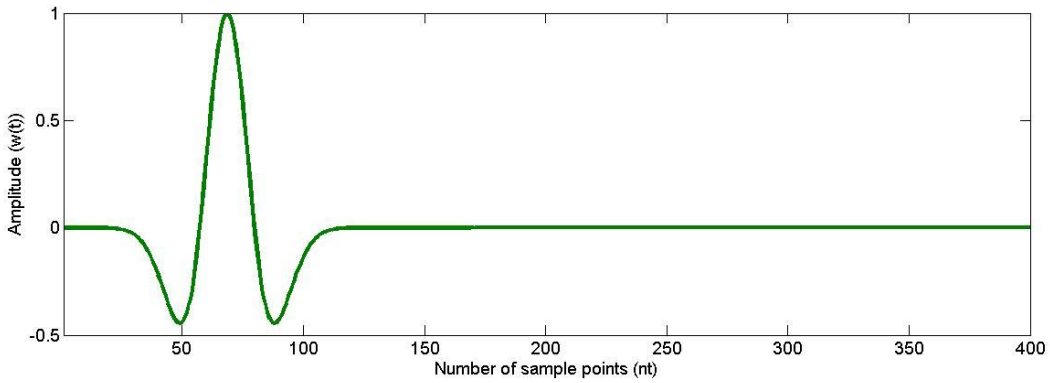


Figure 3.2: Ricker wavelet

Absorbing boundary conditions (ABC) are included in order to reduce reflection from the edges of the grid. The absorbing boundary conditions are constructed from paraxial approximations of the wave equation **(*Clayton and Enquist, 1977*)**. Figure 3.3 demonstrates the effect of ABC from grid edges.

Finally, the wave-fields are recorded at each time at each receiver positions for each source. These records are called as Synthetic Seismograms (for synthetic models) (Fig. 3.4). Multiple seismograms will be recorded from multiple source wave-fields.

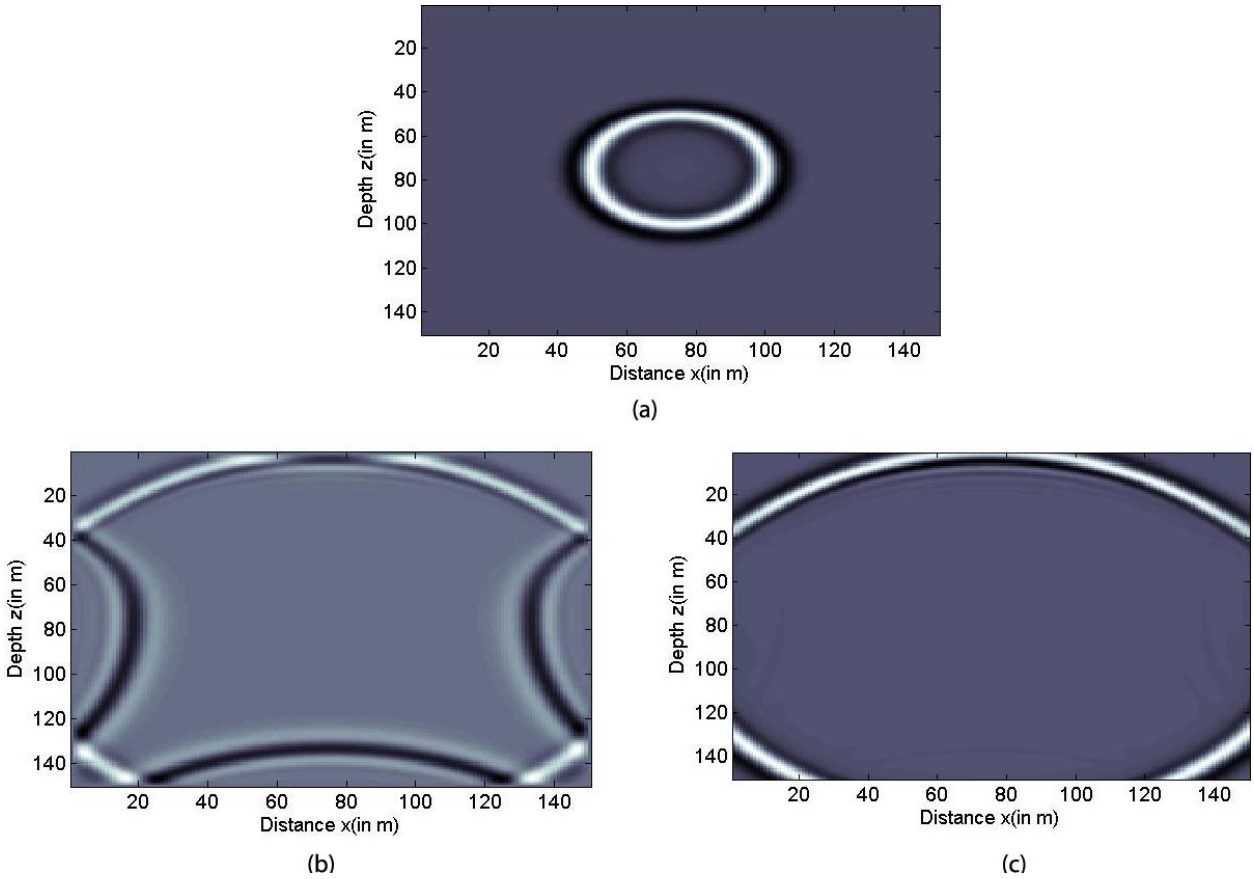


Figure 3.3: Effect of boundary conditions. (a) Wave-field at time 0.17 sec., (b) wave-field at time $t = 0.42$ sec. No boundary condition is used, hence intense reflections off the edges can be seen., (c) wave-field at time $t = 0.42$ sec. Absorbing boundary conditions is used, hence nearly no reflection off the edges.

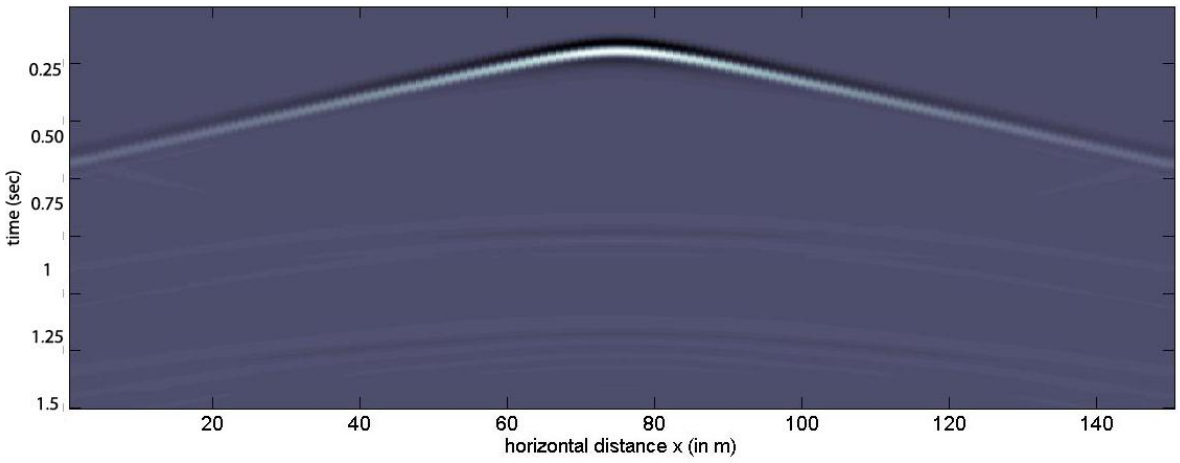


Figure 3.4: An example of synthetic seismogram. Receivers are places throughout the surface at 1 m spacing. Source is placed at 10 m depth in the middle of the medium. Here, reflection events indicating the presence of two reflectors.

3.1.2 Gradient Computation: Adjoint State Method

In section 3.2, I discussed how to implement forward modelling for model parameters $\boldsymbol{m}$**.** Simply, it can be represented as

$$\mathbb{F}(\boldsymbol{m}) = \boldsymbol{u} \qquad (3.10)$$

$$\boldsymbol{S}\mathbb{F}(\boldsymbol{m}) = d \qquad (3.11).$$

Here, $\mathbb{F}$ is the forward modelling map operator, $\boldsymbol{u}$ is the source wave-field, $\boldsymbol{S}$ is sampling operator of the forward map and d is our synthetic data.

In this section, I will implement F^* (adjoint operator) in order to compute the gradient of the Full Waveform Inversion non-linear least square objective function using imaging condition (Reverse Time Migration).

Let the incident wave-field $\boldsymbol{u}_0(x,z,t)$ is generated from the initial model $\boldsymbol{m}_0(x,z)$. The difference between the observed data $\boldsymbol{d}_{obs}$ and the calculated data d_{cal} (due to $\boldsymbol{u}_0(x,z,t)$) is called the residual $\boldsymbol{r}$.

$$\boldsymbol{r}(x,t) = \boldsymbol{d}_{obs}(x,t) - \boldsymbol{d}_{cal}(x,t) = \boldsymbol{d}(x,t) - \boldsymbol{S}\mathbb{F}[\boldsymbol{m}_0](x,t) \qquad (3.12)$$

From Equation (2.7), the objective function can also be represented as $\boldsymbol{O}(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{d}_{obs} - \boldsymbol{S}\mathbb{F}(\boldsymbol{m})\|$ then its gradient is given as $\frac{\partial \boldsymbol{O}}{\partial \boldsymbol{m}}[\boldsymbol{m}]$.

From waves and imaging notes **(Demanet, 2015)**, if $F = \frac{\partial \boldsymbol{S}\mathbb{F}}{\partial \boldsymbol{m}}[\boldsymbol{m}]$ the we have that the gradient of the objective function at a point $\boldsymbol{m}$ is

$$\frac{\partial \boldsymbol{O}}{\partial \boldsymbol{m}}[\boldsymbol{m}] = \nabla \boldsymbol{O}(\boldsymbol{m}) = -F^*(\boldsymbol{d} - \boldsymbol{S}\mathbb{F}[\boldsymbol{m}]) = -F^*\boldsymbol{r} \qquad (3.13)$$

Computation for full matrix $F = \frac{\partial \boldsymbol{S}\mathbb{F}}{\partial \boldsymbol{m}}[\boldsymbol{m}]$ and then transposing it is not considered as an empirical way to form F^*. More sophisticated way is to apply adjoint state method.

From Appendix (1):

$$F^*\boldsymbol{r} = -\int_0^T \boldsymbol{q}(x,z,t)\partial_{tt}\boldsymbol{u}_0(x,z,t)dt \qquad (3.14)$$

where $q(x,z,t)$ (back-propagated wave-field) is the solution of (including ABC)

$$\left(\frac{1}{\boldsymbol{m_0}(x,z)^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\boldsymbol{q}(x,z,t) = \boldsymbol{r}(x,t) \qquad (3.15)$$

and $\partial_{tt}\boldsymbol{u}_0(x,z,t)$ is the double time derivative of the incident wave-field.

A systematic, painless way for solving back-propagated wave-field (Equation (3.15)) is to follow the following steps in sequence **(Demanet, 2015)**:

1. Revert the residual $\boldsymbol{r}(x,t)$ at each receiver position in time.
2. Solve the Equation (3.15) in normal way with this new right hand side.
3. Revert the result at each receiver position in time.

Now, Equation (3.14) can be solved by taking the integral of the product of forward field ($\boldsymbol{u}_0$) (differentiated twice in time), and the back-propagated field $\boldsymbol{q}$, for each x and z independently in time.

Following the above process will produce an image which in seismology called as reverse time migration (**Demanet, 2015**).

For multiple sources, the gradient can be computed as,

$$\nabla \boldsymbol{O}(\boldsymbol{m}_0) = -\sum_s \int_0^T \boldsymbol{q}_s(x,z,t)\partial_{tt}\boldsymbol{u}_{0,s}(x,z,t)dt \qquad (3.16)$$

where the back-propagated field $\boldsymbol{q}_s$ is relative to the source s is:

$$\left(\frac{1}{\boldsymbol{m_0}(x,z)^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\boldsymbol{q}_s(x,z,t) = \boldsymbol{r}_s(x,t) \qquad (3.17)$$

The sum over s in Equation (3.16) is sometimes called a *stack* (Fig. 3.5). Stack makes use of the redundancy in the data to reveal more information about the data.

3.1.3 Gradient Based Optimization

The gradient based optimization method updates the velocity model according to the descent direction $\boldsymbol{P}_k$.

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k \qquad (3.18)$$

Here, k denotes the iteration number and α as step length.

By neglecting the terms higher than the 2nd order, the objective function can be expanded as

$$\boldsymbol{O}(\boldsymbol{m}_{k+1}) = \boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k) = \boldsymbol{O}(\boldsymbol{m}_k) + \alpha_k \langle \nabla \boldsymbol{O}(\boldsymbol{m}_k), \boldsymbol{P}_k \rangle + \frac{1}{2}\alpha_k^2 \boldsymbol{P}_k^{'} \boldsymbol{H}_k \boldsymbol{P}_k \qquad (3.19)$$

Here, H_k stands for Hessian matrix; $\langle .,. \rangle$ denotes inner product. Differentiation of the Equation (3.19) with respect to α_k gives

$$\alpha_k = -\frac{\langle \boldsymbol{P}_k, \nabla \boldsymbol{O}(\boldsymbol{m}_k)\rangle}{\boldsymbol{P}_k^{'}\boldsymbol{H}_k\boldsymbol{P}_k} = -\frac{\langle \boldsymbol{P}_k, \nabla \boldsymbol{O}(\boldsymbol{m}_k)\rangle}{\langle \boldsymbol{J}_k\boldsymbol{P}_k, \boldsymbol{J}_k\boldsymbol{P}_k\rangle} = \frac{\langle \boldsymbol{J}_k\boldsymbol{P}_k, \boldsymbol{d}_{obs} - \boldsymbol{d}_{cal}\rangle}{\langle \boldsymbol{J}_k\boldsymbol{P}_k, \boldsymbol{J}_k\boldsymbol{P}_k\rangle} \qquad (3.20)$$

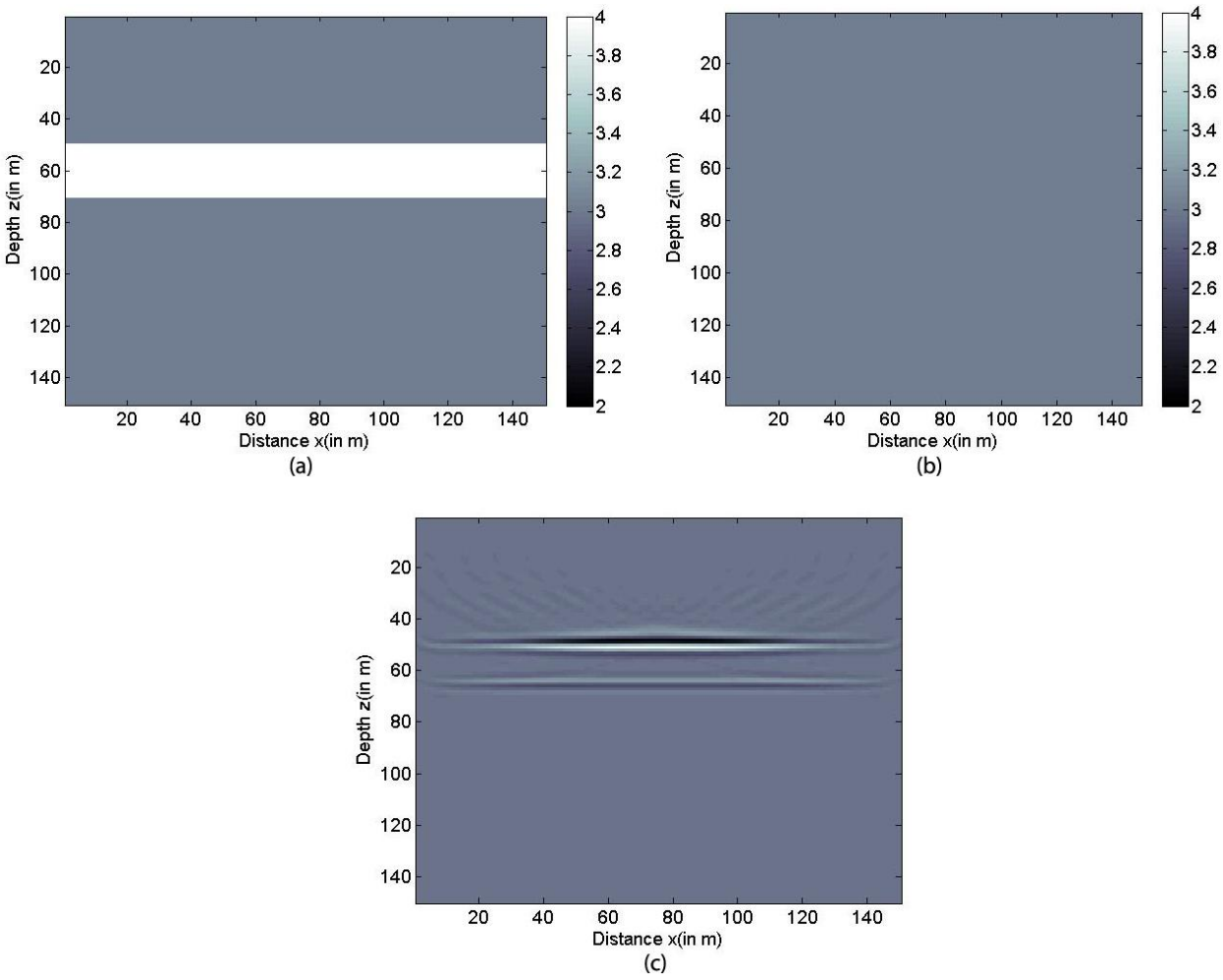


Figure 3.5: An illustration of adjoint state method using multiple source. (a) True velocity model, (b) initial velocity model, (c) stack image using (b) as initial model showing the reflectors at its exact location.

in which I used the approximated Hessian $\boldsymbol{H}_k \approx \boldsymbol{H}_a$.

To obtain a reasonable step size α_k in Equation (3.20), I estimated a small step length ϵ proposed by **Pica et al. (1990)**.

$$\max(\epsilon|\boldsymbol{P}_k|) \leq \frac{\max(|\boldsymbol{m}_k|)}{100} \qquad (3.21)$$

and the Taylor approximation

$$\boldsymbol{J}_k\boldsymbol{P}_k \approx \frac{\boldsymbol{S}\mathbb{F}(\boldsymbol{m}_k + \varepsilon\boldsymbol{P}_k) - \boldsymbol{S}\mathbb{F}(\boldsymbol{m}_k)}{\varepsilon} \qquad (3.22)$$

For computing direction P_k I used Non-linear Conjugate Gradient method, which decreases the objective function along the conjugate gradient direction.

$$\boldsymbol{P}_k = \begin{cases} -\nabla \boldsymbol{O}(\boldsymbol{m}_0), & k = 1 \\ -\nabla \boldsymbol{O}(\boldsymbol{m}_k) + \gamma_k \boldsymbol{P}_{k-1} & k > 1 \end{cases} \qquad (3.23)$$

There are number of ways to compute γ_k. I used a hybrid scheme combing Hestenes-Stiefel method and Dai-Yuan method **(Hagar and Zhang, 2006)** for better convergence.

$$\gamma_k = \max(0, min(\gamma_k^{HS}, \gamma_k^{DY})) \qquad (3.24)$$

In which

$$\begin{cases} \gamma_k^{HS} = \dfrac{\langle \nabla \boldsymbol{O}(\boldsymbol{m}_k), \nabla \boldsymbol{O}(\boldsymbol{m}_k) - \nabla \boldsymbol{O}(\boldsymbol{m}_{k-1}) \rangle}{\langle \boldsymbol{P}_{k-1}, \nabla \boldsymbol{O}(\boldsymbol{m}_k) - \nabla \boldsymbol{O}(\boldsymbol{m}_{k-1}) \rangle} \\ \gamma_k^{DY} = \dfrac{\langle \nabla \boldsymbol{O}(\boldsymbol{m}_k), \nabla \boldsymbol{O}(\boldsymbol{m}_k) \rangle}{\langle \boldsymbol{P}_{k-1}, \nabla \boldsymbol{O}(\boldsymbol{m}_k) - \nabla \boldsymbol{O}(\boldsymbol{m}_{k-1}) \rangle} \end{cases} \qquad (3.25)$$

This provide an automatic direction reset while over-correction of γ_k in conjugate gradient iterations. It reduces to steepest descent method when the subsequent search directions lose conjugacy. Figure 3.6 summarize the time domain image reconstruction algorithm of FWI.

3.1.4 Testing Time Domain Algorithm on simple Models

In order to test the codes, FWI algorithm was performed on two different velocity models separately. Velocity models chosen for testing were (1) low velocity step layer (2) high velocity step layer models.

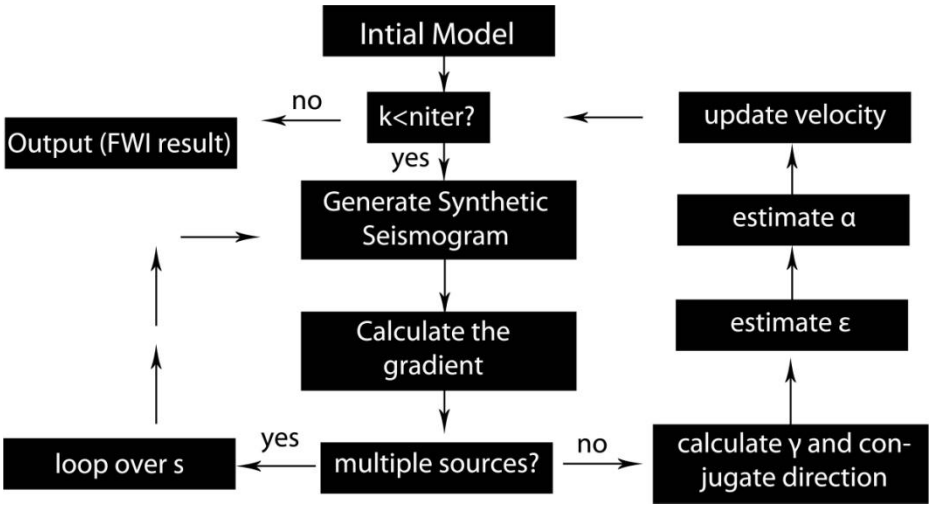


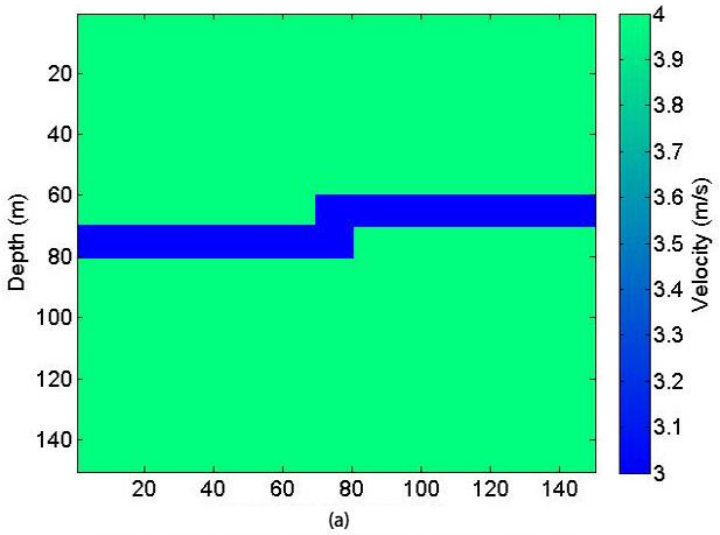
Figure 3.6: Full waveform Inversion flowchart

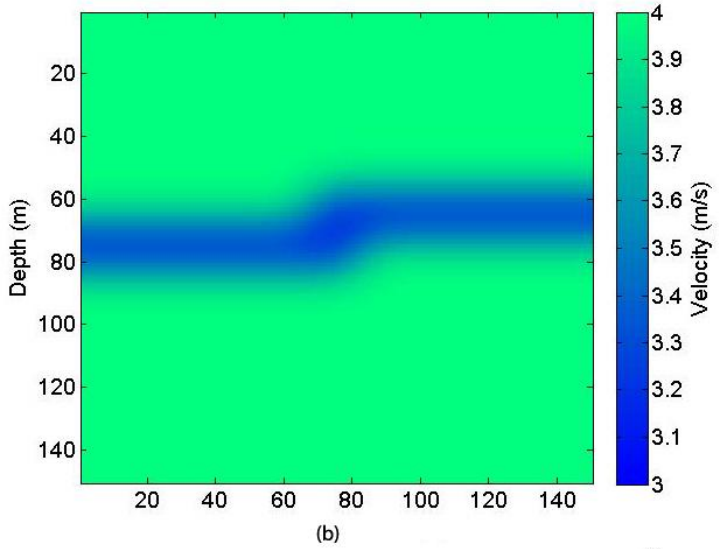
3.1.4.1 Low velocity layer model (LVL)

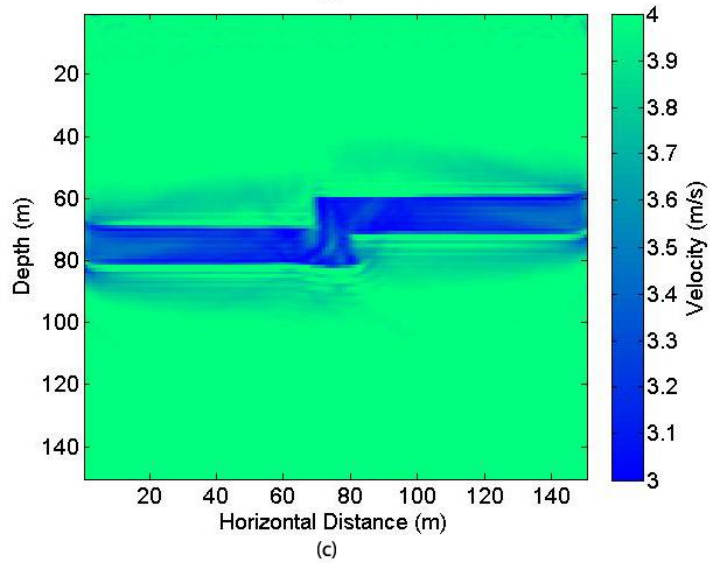
In this synthetic model (real model) (Fig. 3.7(a)) the field area is considered 150-m long and 150-m deep. The step like low velocity layer has a velocity of 3 m/s and a background velocity of 4 m/s. For Initial model (Fig. 3.7 (b)), I smoothen the true model using a Gaussian smoother function with a factor of 10. In order to avoid grid dispersion, Δx $(= \Delta z)$ must be chosen as

$$\Delta x < \frac{\max(\boldsymbol{m})}{4f} \qquad (3.26).$$

Here, f is the peak frequency. Sources were kept at a depth of 10-m from (10, 10) to (140, 10) with 10-m spacing between each source. Receiver were kept at a depth of 2-m from (1, 2) to (150, 2) with a spacing of 1-m between each receiver. ABCs were applied over all four boundaries. Number of iterations was kept 150. Final model is a result of FWI after such iterations and is shown in Fig. 3.7(c). The objection function's value (Fig. 3.7 (d)) is rapidly decreasing up to 25 iterations indicating that the objective function is decreasing in the negative conjugate direction, and then after in the steepest descent direction up to 200 iterations.


Depth (m)
20
40
60
80
100
120
140
20
40
60
80
100
120
140
Velocity (m/s)
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(a)


Depth (m)
20
40
60
80
100
120
140
20
40
60
80
100
120
140
Velocity (m/s)
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(b)


Depth (m)
20
40
60
80
100
120
140
20
40
60
80
100
120
140
Horizontal Distance (m)
Velocity (m/s)
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(c)

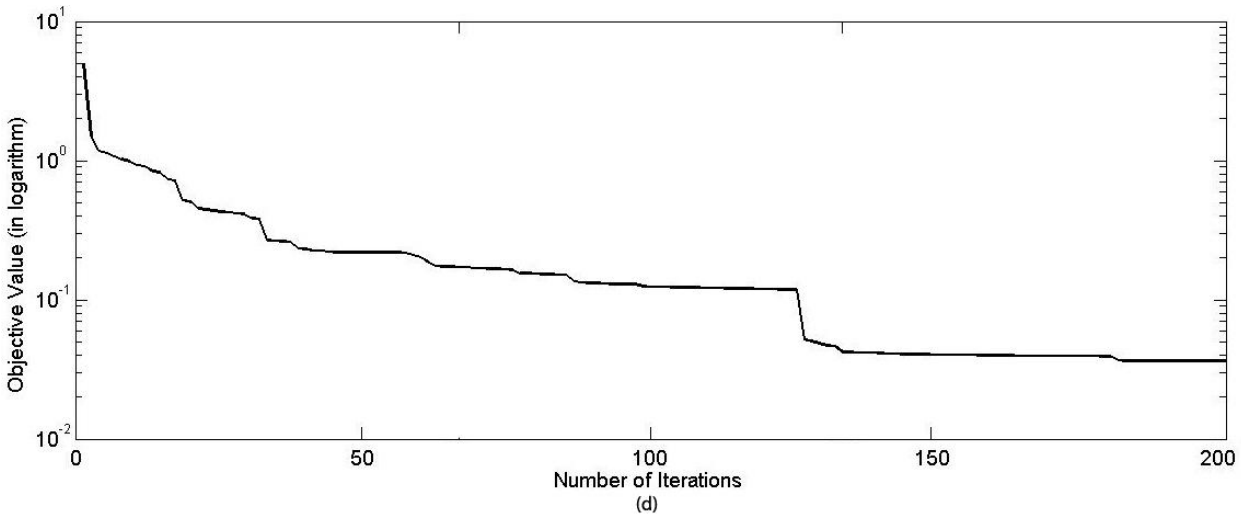
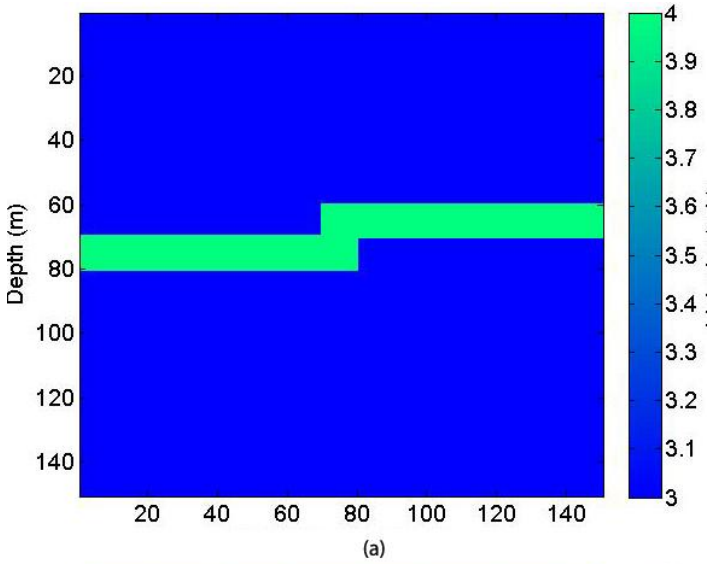


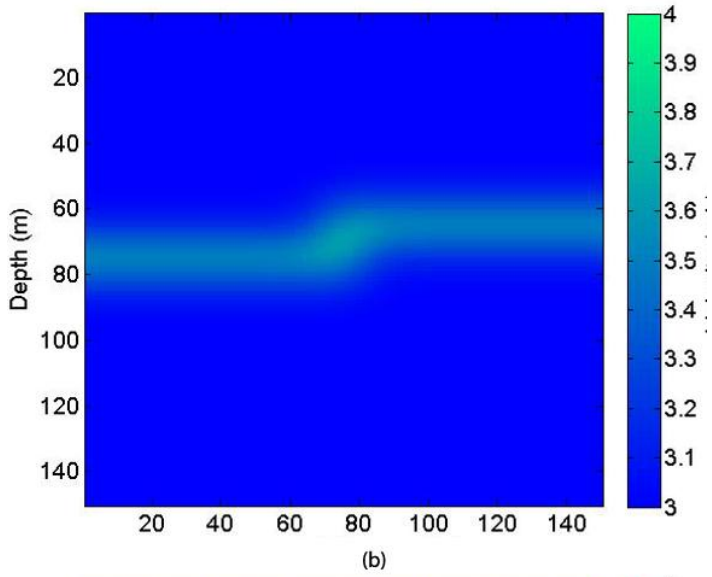
Figure 3.7: (a) Real synthetic model, (b) initial model, (c) final FWI result and (d) objective function

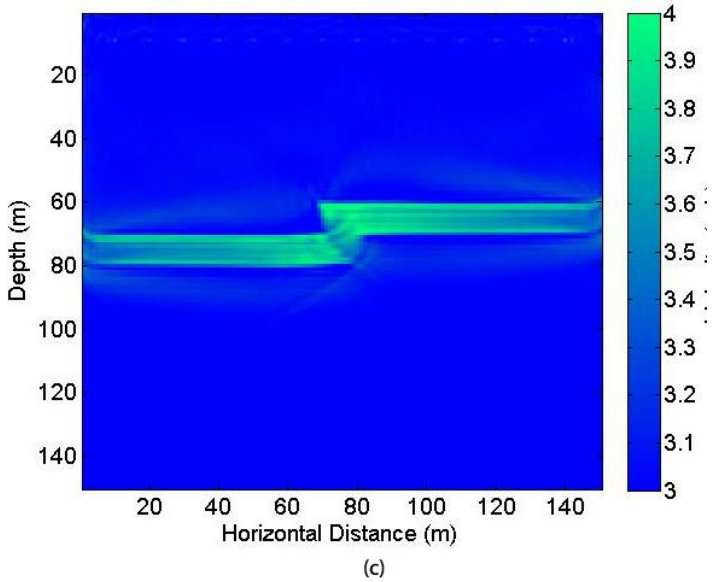
3.1.4.2. High velocity layer model (HVL)

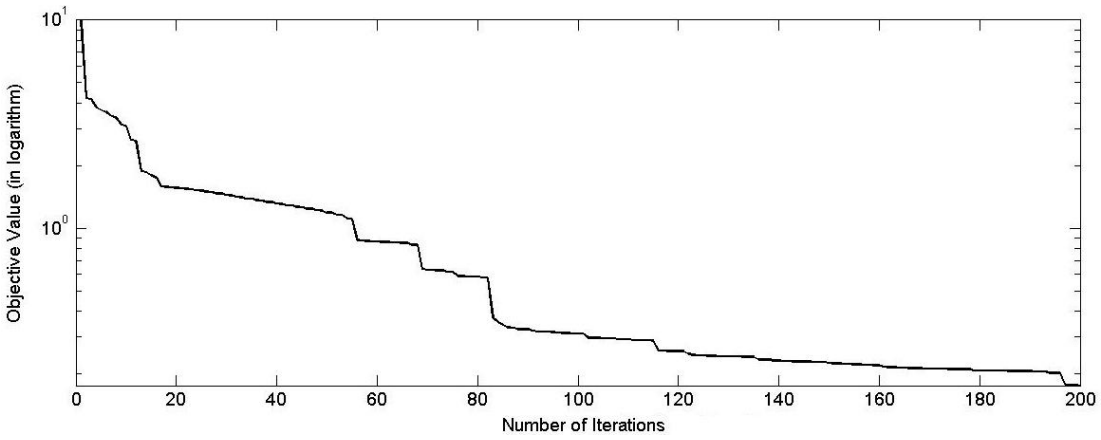
The synthetic model (true model) (Fig. 3.8(a)) is same as in Fig. 3.7(a) except the step like high velocity layer has a velocity of 4 m/s and a background velocity of 3 m/s. For Initial model (Fig. 3.8 (b)), the same concept is used as I did in LVL case.

Sources were kept at a depth of 10-m from (10, 10) to (140, 10) with 10-m spacing between each source. Receiver were kept at a depth of 2-m from (1, 2) to (150, 2) with a spacing of 1-m between each receiver. ABCs were applied over all four boundaries. Number of iterations was kept 228. Final model is a result of FWI after such iterations and is shown in Fig. 3.8(c). The objection function's value (Fig. 3.8 (d)) is rapidly decreasing up to 20 iterations indicating that the objective function is decreasing in the negative conjugate direction, and then after in the steepest descent direction up to 200 iterations.


Depth (m)
20
40
60
80
100
120
140
20 40 60 80 100 120 140
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(a)


Depth (m)
20
40
60
80
100
120
140
20 40 60 80 100 120 140
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(b)


Depth (m)
20
40
60
80
100
120
140
20 40 60 80 100 120 140
Horizontal Distance (m)
4
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
3
(c)



(d)

Figure 3.8: (a) Real synthetic model, (b) initial model, (c) final FWI result and (d) objective function.

3.2 Image Reconstruction for Marmousi Model

The main objective of this thesis is to evaluate Full Waveform Inversion applied to a very complex geological structure i.e. Marmousi model, and see if the method can be useful to recover velocity from it. The Marmousi dataset was used for the workshop on practical aspect of seismic data inversion at the 52nd EAGE meeting in 1990. This is a very challenging medium for seismic imaging.

The original Marmousi dataset is huge (Fig. 3.9). In order to handle the dataset, the original model was resampled into a much shorter dataset as shown in Fig. 3.10. This modification was done in order to have a reasonable size dataset which is easy to handle. Now, this modified model will be called as 'real model'.

In true model, field area is 722-m long and 200-m deep. The maximum velocity is 4.7 m/s. A low velocity water layer of thickness 25-m is places over the model and carries a constant velocity of 1.5 m/s.

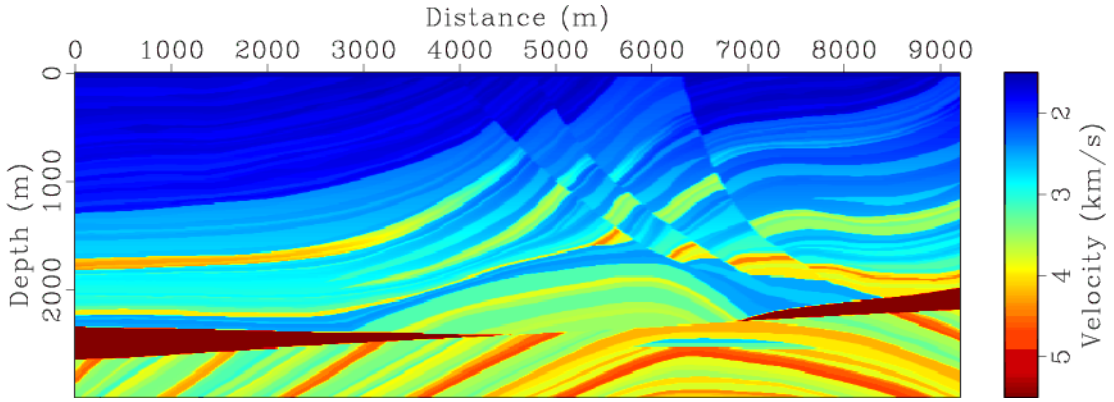


Figure 3.9: Complete Marmosui model

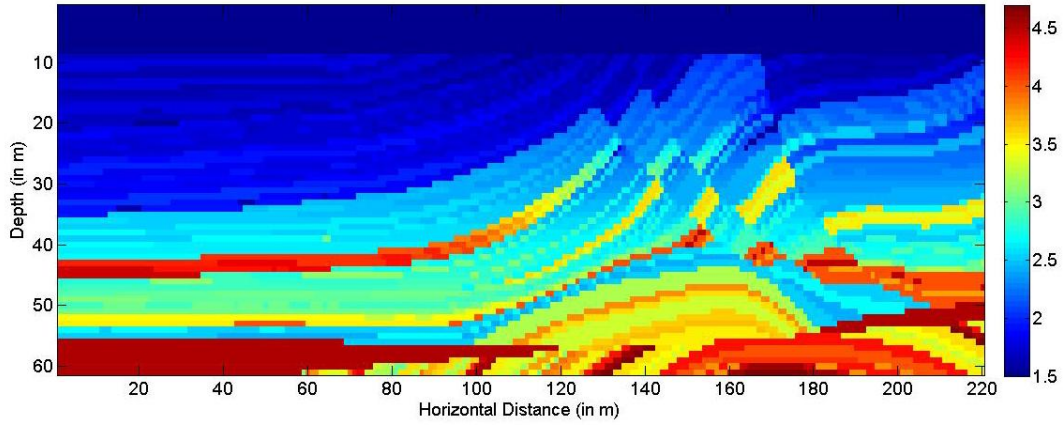


Figure 3.10: Resampled Marmousi model (True model for this thesis)

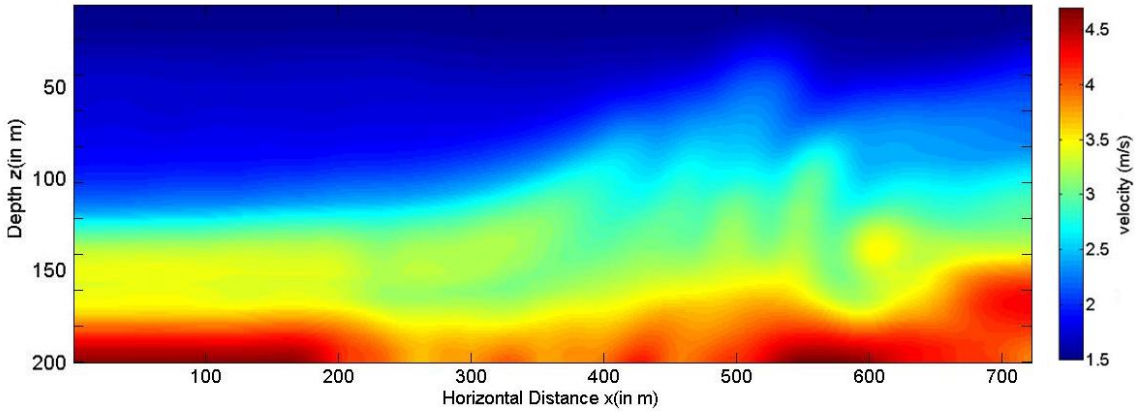


Figure 3.11: Initial model used to initiate FWI

Again, a Gaussian smoother of factor 10 is used to smooth the real model, which will be my initial model (Fig. 3.11) for FWI. The configuration of sources and receivers are shown in Fig. 3.11. There are 21 sources totally (from (10, 25) to (210, 25)) and 722 receiver for each source (at a depth of 6-m). The sources step is 10-m and the receiver step is 1-m. The sources line and receiver line is 10-m and 2-m below the surface respectively. The grid spacing was kept $\Delta x\ (= \Delta z)$ =0.01-m.

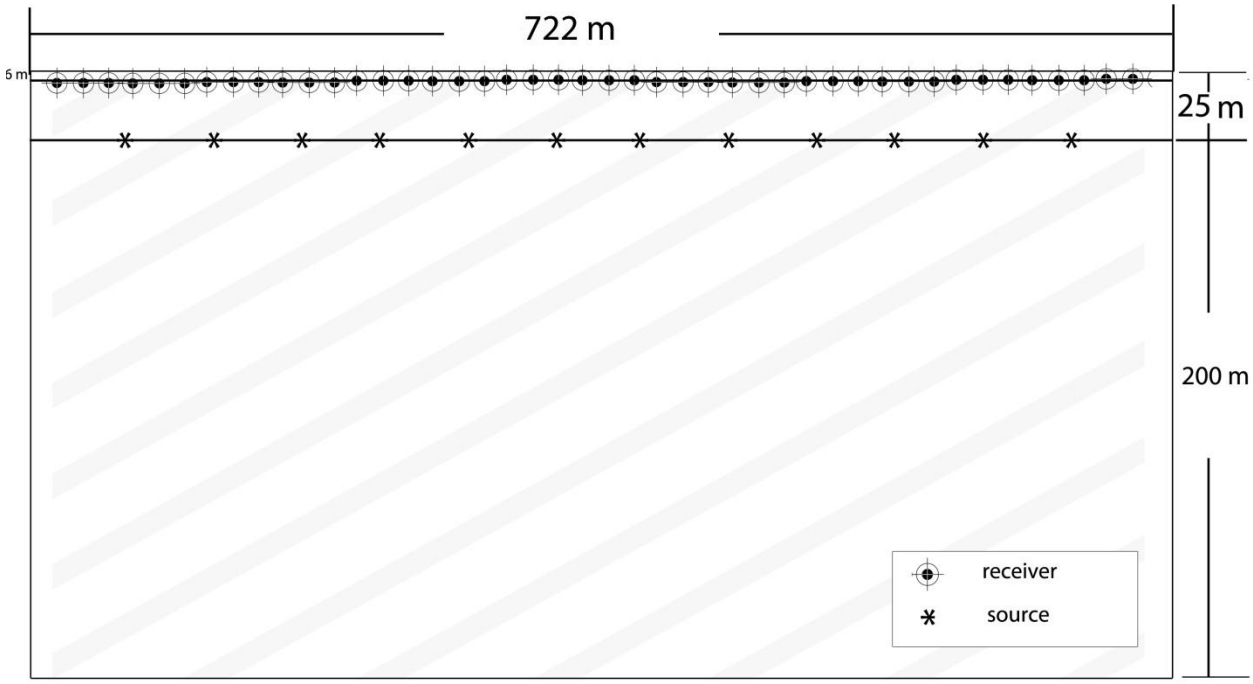


Figure 3.12: The configuration of sources and receiver for Full Waveform Inversion

In above configuration ABCs are applied on all four edges. The number of iterations set was 200. Figure 3.13 shows the Full Waveform Inversion result after 200th iteration with its continuously decreasing objective function (Fig. 3.14).

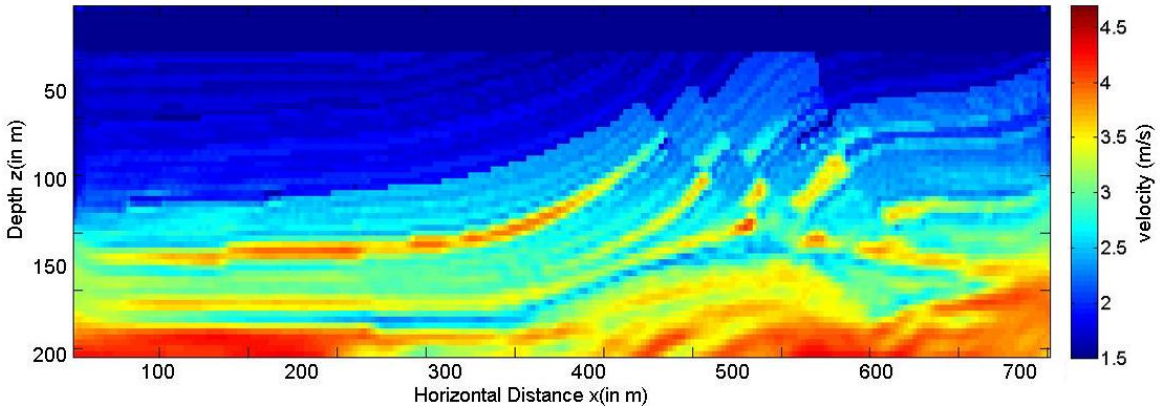


Figure 3.13: FWI's final model after 200th iteration

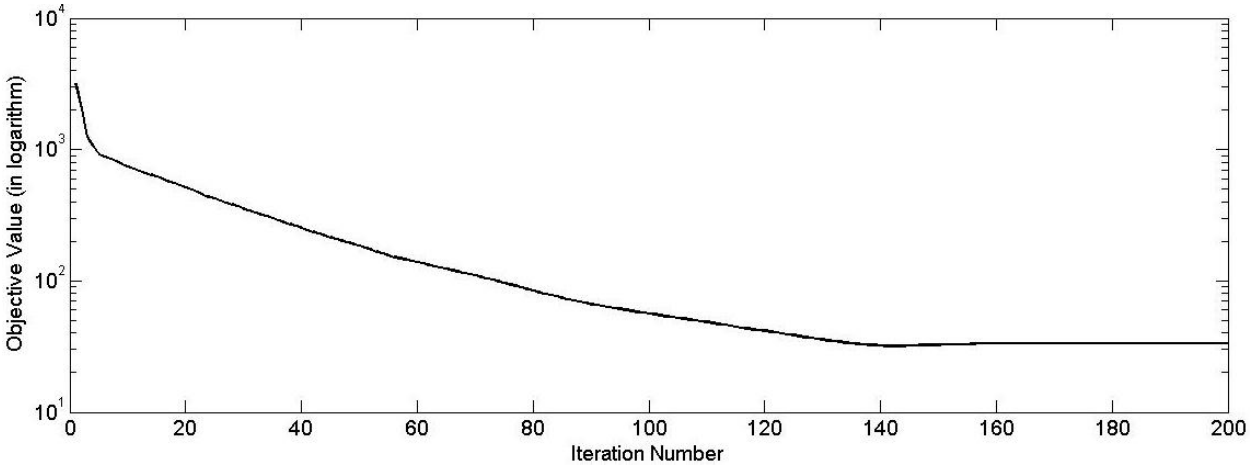


Figure 3.14: Objective function value at each iteration.

3.2.1 Discussion

A more qualified analysis of the velocities is easier when looking at the velocity profiles on Figure 3.15. Velocity logs at three different horizontal locations are extracted (at 200 m, 400 m and 600 m) from the true model, initial model and final model. In all three profiles (3.15(a), 3.15(b), 3.15(c)) the recovered velocities from FWI result matches the real velocities very well. The largest deviation from the real velocity is about 29% in a small portion of the 600 m profile (see Table. 3.1). Some portions of FWI are either slightly underestimated or overestimated due to

high to low (or low to high) velocity transition. If the average velocity for profile at x= 200m, 400m and 600m are 2.45 m/s, 2.48 m/s and 2.62 m/s, then average deviations from three different profiles will be -2.66%, -0.31%, -2.84 %. Hence, the Full-Waveform Inversion has managed to obtain precise results from the complex Marmousi model in time domain.

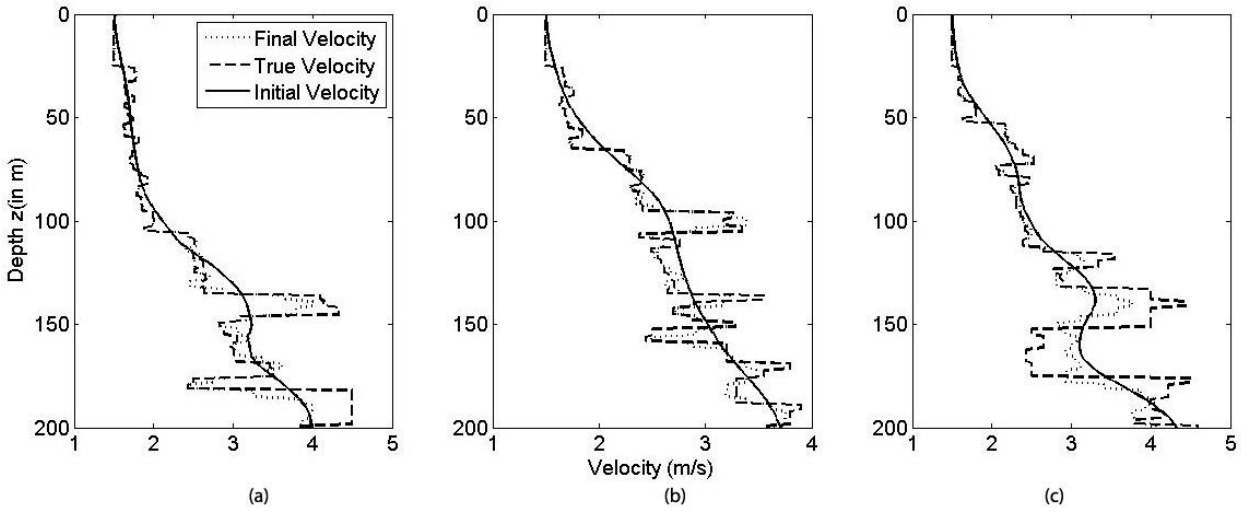


Figure 3.15: Velocity logs extracted at the horizontal distance x (a) 200 m (b) 400 m (c) 600 m.

Table 3.1: Table representing the true velocity, final recovered velocity and the % deviation along with the depth for each profile.

	Profile at x=200-m			*Profile at x=400-m*			*Profile at x=600-m*		
Depth (m)	**TRUE (m/s)**	**Final (m/s)**	**% Deviation**	**TRUE (m/s)**	**Final (m/s)**	**% Deviation**	**TRUE (m/s)**	**Final (m/s)**	**% Deviation**
1	1.5	1.52	1.57	1.5	1.50	0.01	1.5	1.50	0.13
11	1.5	1.50	0.09	1.5	1.49	-0.08	1.5	1.50	0.00
21	1.5	1.49	-0.08	1.5	1.50	0.01	1.5	1.49	-0.07
31	1.77	1.77	0.28	1.67	1.67	-0.10	1.58	1.58	0.13
41	1.76	1.75	-0.57	1.65	1.65	-0.01	1.66	1.65	-0.18
51	1.74	1.74	-0.21	1.69	1.70	0.07	1.63	1.63	-0.02
61	1.81	1.81	-0.30	1.72	1.72	-0.03	2.20	2.22	0.81
71	1.70	1.70	-0.01	2.29	2.29	0.21	2.53	2.51	-0.84
81	1.93	1.93	0.02	2.40	2.40	-0.28	2.47	2.45	-0.84
91	1.84	1.84	-0.02	2.39	2.35	-1.63	2.30	2.35	2.30
101	2.00	1.98	-0.84	3.16	3.39	7.01	2.39	2.40	0.45
111	2.51	2.54	1.06	2.75	2.60	-5.55	2.4	2.39	-0.01
121	2.63	2.51	-4.27	2.58	2.58	0.09	3.35	3.17	-5.33
131	2.61	2.46	-5.94	2.70	2.63	-2.83	2.89	2.82	-2.44
141	4.14	4.00	-3.49	2.89	3.17	9.81	4.45	3.76	-15.43
151	2.83	2.82	-0.48	3.28	3.23	-1.63	4	2.83	-29.16
161	2.97	3.03	1.92	3.2	2.89	-9.63	2.65	3.01	13.72
171	3.47	3.62	4.33	3.8	3.39	-10.72	2.5	3.12	24.93
181	2.44	2.46	0.84	3.3	3.21	-2.70	4.23	3.41	-19.20
191	4.5	4.00	-10.96	3.9	3.60	-7.54	4	3.94	-1.46

CHAPTER 4

IMAGE RECONSTRUCTION IN FREQUENCY DOMAIN

In this chapter, the process of reconstructing the image using Full Waveform Inversion in frequency domain will be discussed. Frequency domain Waveform Inversion, which includes solving the wave equation forward modelling and inversion, is similar to time domain Waveform Inversion. The solution of the wave equation (acoustic wave) in forward modelling is realized by numerical simulation **(*Wencai, 2002*)** of seismic wave propagation, and the inversion process can considered as solving a non-linear optimization problem. However, the efficiency of frequency domain Full Waveform Inversion depends on the number of frequencies used for inversion and is independent of the number of sources used (unlike time domain Full Waveform Inversion).

4.1 Introduction Domain

Since **Pratt (1990)** extended time domain waveform inversion into the frequency domain one, frequency domain Full Waveform Inversion has begun to develop rapidly **(*Meng et al., 2012*)**. Many theoretical studies have been made in the frequency domain waveform inversion field in recent years. **Sourbier et al (2009)** presented a preconditioned gradient method in frequency domain waveform inversion. **Zhang and Yuan (2009)** introduced two instances of frequency domain seismic waveform inversion for velocity imaging. **Long et al (2009)** provided the inversion of preconditioned gradient method to find the viscoelastic acoustic velocity structure. In this chapter, both the forward modelling and the inversion are performed in the frequency domain.

4.2 Forward Modelling in Frequency Domain

In general, forward modelling is the basis of the inversion process. The precision and efficiency of forward modelling have a substantial influence on the results of inversion. The advantage of frequency domain modelling compared with time domain modelling is that multi-experiment seismic data can be simulated economically by direct multiplication once the factors of the impedance matrix are calculated **(*Meng et al., 2012*)**. In addition, the forward wave equation in the frequency domain is more flexible as we can directly input the attenuation coefficient in the equation. Now, as described in section 2.2, the acoustic 2D wave equation is

$$\frac{1}{\boldsymbol{m}(x,z)^2}\frac{\partial^2 \boldsymbol{u}(x,z,t)}{\partial t^2} - \nabla^2 \boldsymbol{u}(x,z,t) = \boldsymbol{f}(x,z,t)$$

Taking Fourier transform on both sides of the two-dimensional time domain acoustic wave equation, the acoustic wave equation in frequency domain is as follows

$$\frac{\omega^2}{\boldsymbol{m}(x,z)^2}\boldsymbol{u}(x,z,\omega) + \nabla^2 \boldsymbol{u}(x,z,\omega) = \boldsymbol{f}(x,z,\omega) \qquad (4.1)$$

where ω is the frequency, $\boldsymbol{m}(x,z)$ is the acoustic propagation velocity in the medium, $\boldsymbol{u}(x,z,\omega)$ is the pressure field, $\boldsymbol{f}(x,z,\omega)$ is the source and ∇^2 is the Laplacian operator. From Equation (4.1) Helmholtz equation can be extracted and represented as

$$\frac{\omega^2}{\boldsymbol{m}(x,z)^2}\boldsymbol{u}(x,z,\omega) + \nabla^2 \boldsymbol{u}(x,z,\omega) = 0 \qquad (4.2)$$

At present, there are three common categories of frequency domain forward modelling methods: the finite difference, the finite element, and the finite volume method. The above Helmholtz equation (4.2) was discretised using finite difference scheme mentioned in **Erlangga (2008)**.The finite difference method is a numerical method based on the calculating difference principal. The boundary condition in the frequency domain is

$$\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}} - ik(x,y)\boldsymbol{u} = 0 \qquad (4.3)$$

where $\boldsymbol{n}$ is the outward normal of the boundary and i is the imaginary unit. Equation (4.3) is applied on all four edges of the field medium. By means of finite difference technology along with boundary condition, the wave equation (Equation (4.1)) can be recast in a matrix form as

$$\boldsymbol{Au} = \boldsymbol{f} \qquad (4.4)$$

In the above equation, A is a complex-valued, nx x nz (discritization along x and z direction) order impedance matrix. It depends on frequency, medium properties, finite difference format, and boundary conditions. $\boldsymbol{f}$ represents a vector consisting all point sources. The wave-field $\boldsymbol{u}$ (or pressure field) can be solved by

$$\boldsymbol{u} = \boldsymbol{A}^{-1}\boldsymbol{f} \qquad (4.5)$$

Figure (4.1), (4.2) and (4.3) represents the wave-field $\boldsymbol{u}$ acquired at different frequencies for Marmousi model (Fig. 1.4) keeping source in the middle.

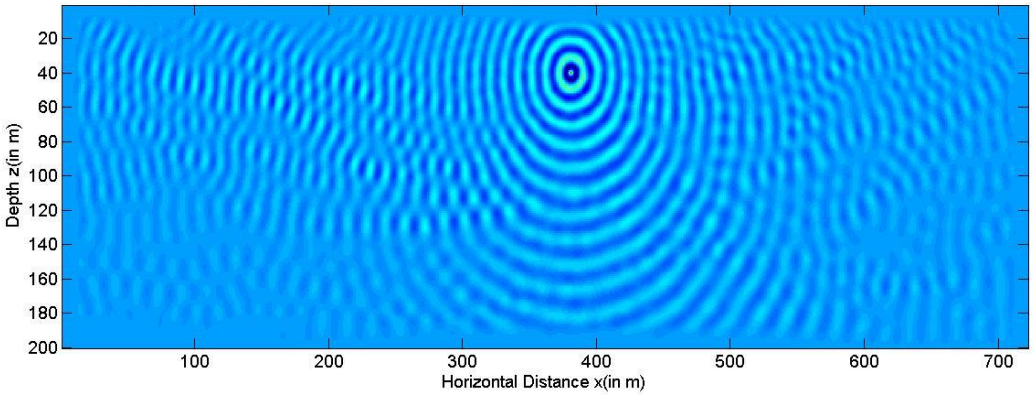


Figure 4.1: Wave-field at frequency 20 Hz.

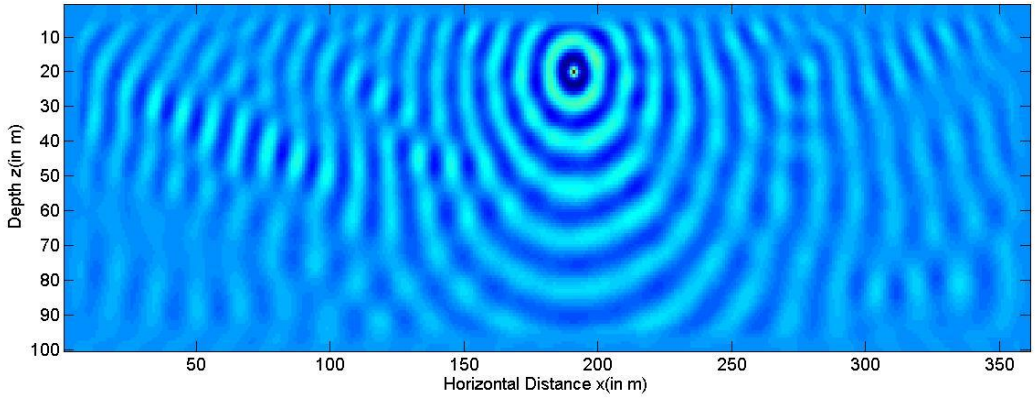


Figure 4.1: Wave-field at frequency 10Hz

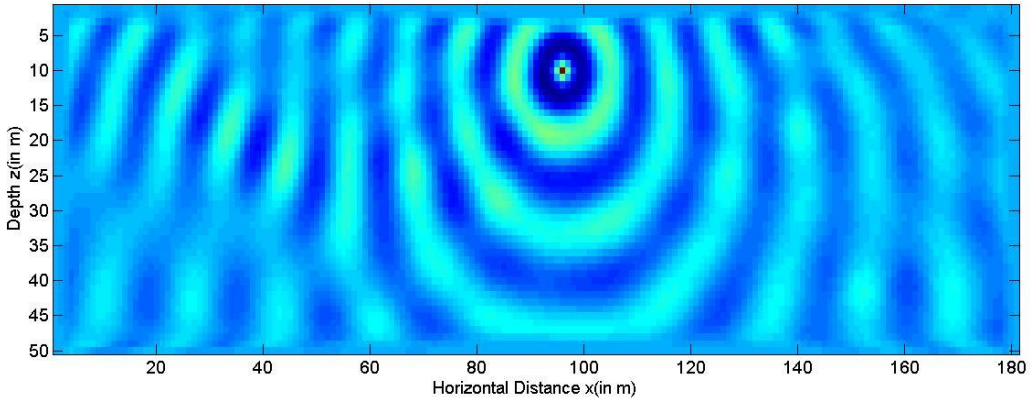


Figure 4.3: Wave-field at frequency 5Hz

If $\boldsymbol{P}$ is a vector consisting of point receivers then synthetic data d is be generated by

$$\boldsymbol{d} = \boldsymbol{P}' * \boldsymbol{u} \qquad (4.6)$$

4.2 Gradient Computation: Adjoint state method

A little modification in the description will suffice to express the time domain adjoint state equation into frequency domain. The idea in this section is that time reversal corresponds to complex conjugation in ω. For simplicity, I will assume single source.

The aim of adjoint state method is to find F^* such that $\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = \langle F^*\boldsymbol{d}, \boldsymbol{m} \rangle$ for all general $\boldsymbol{d}$ and $\boldsymbol{m}$. If $\boldsymbol{d}_r(t)$ is a function of receiver index r and time t then using Parseval formula, the data inner product $\langle \boldsymbol{d}, F\boldsymbol{m} \rangle$ can be represented in the frequency domain (Eq. 4.7) as

$$\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = 2\pi \sum_r \int_R \widehat{\boldsymbol{d}_r}(\omega)\overline{\widehat{(F\boldsymbol{m})}(x_r, z_r, \omega)}d\omega = \sum_r \int \boldsymbol{d}_r(t)(F\boldsymbol{m})(x_r, z_r, t)dt \qquad (4.7)$$

Since, the net quantity in the Equation (4.7) is real, it does not matter which of the two integrand's factors complex conjugate is placed on. From Appendix (2) the expression for the gradient is given by

$$F^*\boldsymbol{d}(x, z) = -2\pi \int_R \widehat{\boldsymbol{q}}(x, z, \omega)\omega^2\overline{\widehat{\boldsymbol{u}_0}(x, z, \omega)}d\omega \qquad (4.8)$$

If `V` $(= \widehat{\boldsymbol{q}_s}(x, z, \omega))$ is the back-propagated wave-field then in MATLAB $\widehat{\boldsymbol{q}_s}(x, z, \omega)$ is computed as

```
V = A'\(r*(Dobs-r'*(A\s)));
```

In synthetic modelling case, `Dobs` can be obtained from Equation (4.7) when applied on true model $(\boldsymbol{m})$, however in real subsurface modelling case, the available seismic data must be converted into frequency domain data set (Appendix (3)). The integral in ω in Equation (4.8) is over R and should be truncated to the number of frequencies used in the process.

4.2 Optimization: Limited Memory BFGS

Limited memory BFGS is an optimization algorithm in the family of quasi-Newton methods that approximates the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm using a limited amount of computer memory. According to **Wensheng Zhang (2013)**, the BFGS based method has higher inversion accuracy than the conjugate gradient based algorithm because the former uses the second order derivative information in computation.

Like the original BFGS, Limited memory-BFGS (L-BFGS) uses estimation to the inverse Hessian matrix to guide its search through variable space, but where BFGS stores a full approximation to the inverse Hessian, LBFGS stores only a few vectors that represent the approximation implicitly.

4.3.1 Computation for direction

LBFGS shares many attributes with other quasi-Newton algorithm, however it is different in the way the matrix multiplies for finding the search direction i.e., $\boldsymbol{P}_k = -\boldsymbol{H}_k \nabla \boldsymbol{O}_k$.

Here, I will take as given x_k (the model at iteration k), the position at the kth iteration, and $g_k = \nabla \boldsymbol{O}(\boldsymbol{m}_k)$ where O is the objective function, and all vectors are column vectors. The last m (this is not model $\boldsymbol{m}$) updates are stored in form $\boldsymbol{s}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$ and $\boldsymbol{y}_k = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$. I will define $\rho_k = \frac{1}{\boldsymbol{y}_k^T \boldsymbol{s}_k}$, and H_k^0 will be the initial approximation of the inverse Hessian that the estimation at iteration k will begin with. Then the direction can be computed as follows:

$$\boldsymbol{q} = \boldsymbol{g}_k$$

$$For\ i = k-1:-1:k-m$$

$$\alpha(i) = \rho(i) * s(i)^{'} * q$$

$$q = q - \alpha(i) * y(i) \quad \rangle \quad MATLAB\ format$$

$$end$$

$$\boldsymbol{H}_k^0 = \frac{\boldsymbol{y}_{k-1}^T \boldsymbol{s}_{k-1}}{\boldsymbol{y}_{k-1}^T \boldsymbol{y}_{k-1}} \quad \boldsymbol{z} = \boldsymbol{H}_k^0 \boldsymbol{q}$$

$$For\ i = k-m:1:k-1$$

$$\beta(i) = \rho(i) * y(i)^T * z$$

$$z = z + s(i) * \big(\alpha(i) - \beta(i)\big) \quad \rangle \quad MATLAB\ format$$

$$end$$

$$End\ with\ \boldsymbol{H}_k \boldsymbol{g}_k = \boldsymbol{z}$$

The above algorithm is valid either we are minimizing or maximizing. In case of minimizing, the search direction $\boldsymbol{z}$ $(= \boldsymbol{P}_k)$ would be the negative of z. The model update is now will be represented as

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k$$

Here $\boldsymbol{P}_k = z$ and α_k is the step length. α_k can be found by exact line search (see section 3.1.3) or by approximated line search methods. LBFGS does not require exact line search in order to converge. In my thesis, I used Wolfe's algorithm for computing α.

4.3.1 Computation for step length: Wolfe's Conditions (1969, corrected in 1971)

A popular inexact line search condition stipulates that α_k should first of all give sufficient decrease in the objective function ($\boldsymbol{O}(\boldsymbol{m}_k)$), as measured by the following inequality

$$\boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k) \leq \boldsymbol{O}(\boldsymbol{m}_k) + c_1 \alpha_k \nabla \boldsymbol{O}(\boldsymbol{m}_k)' \boldsymbol{P}_k \qquad where\ c_1 \in (0,1) \qquad (4.9)$$

Equation (4.9) requires that for the picked value of α_k the graph of $\boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k)$ lies below the line $\boldsymbol{O}(\boldsymbol{m}_k) + c_1 \alpha_k \nabla \boldsymbol{O}(\boldsymbol{m}_k)' \boldsymbol{P}_k$ (Fig. 4.4). By Taylor's theorem

$$\boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k) \leq \boldsymbol{O}(\boldsymbol{m}_k) + \alpha_k \nabla \boldsymbol{O}(\boldsymbol{m}_k)^T \boldsymbol{P}_k + \boldsymbol{R}(\alpha^2) \qquad (4.10)$$

Since, $\boldsymbol{P}_k$ is a descent direction, i.e., $\nabla \boldsymbol{O}(\boldsymbol{m}_k)' \boldsymbol{P}_k < 0$, such α_k exists.

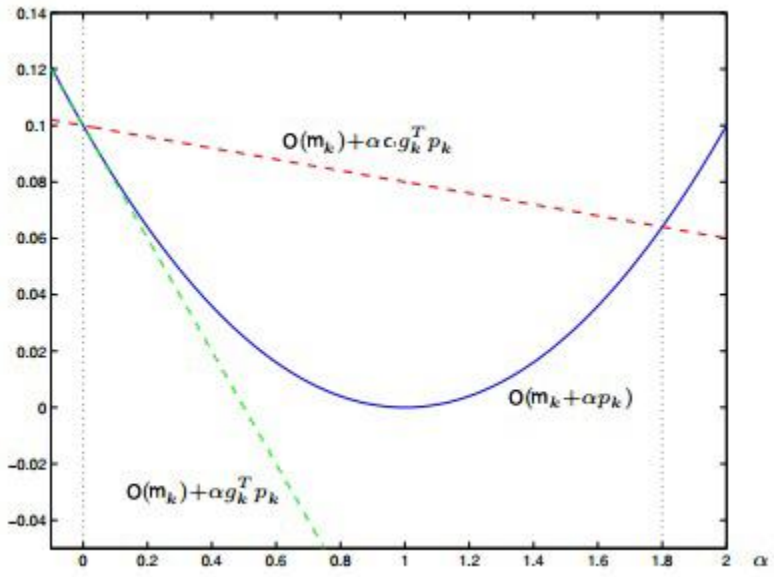


Figure 4.4: Graph illustrating the Equation (4.9). Here g_k represents the gradient ($\nabla O(\boldsymbol{m}_k)$).

(after: **Hauser, 2007**)

The sufficient decrease condition (in Equation (4.9)) is not enough to ensure convergence since this equation is satisfied for all small enough α. To rule out unacceptably small steps, the second requirement called a curvature condition is introduced.

$$\nabla \boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k)^T \boldsymbol{P}_k \geq c_2 \nabla \boldsymbol{O}(\boldsymbol{m}_k)^T \boldsymbol{P}_k, \qquad c_2 \in (c_1, 1) \qquad (4.11)$$

The curvature condition enforces to choose α large enough so that the slope of $\boldsymbol{O}(\boldsymbol{m}_k + \alpha k \boldsymbol{P} k$ is larger than $c2$ times the slope of $\boldsymbol{O}\boldsymbol{m}k$.

Equation (4.10) and (4.11) are Wolfe conditions. Sometimes the curvature condition can be amplified to out rule α's for which $\boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k)$ increases faster than$c_2 \nabla \boldsymbol{O}(\boldsymbol{m}_k)^T \boldsymbol{P}_k$. The resulting conditions are called strong Wolfe's conditions.

$$\boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k) \leq \boldsymbol{O}(\boldsymbol{m}_k) + c_1 \alpha_k \nabla \boldsymbol{O}(\boldsymbol{m}_k)' \boldsymbol{P}_k \qquad (4.12)$$

$$|\nabla \boldsymbol{O}(\boldsymbol{m}_k + \alpha_k \boldsymbol{P}_k)^T \boldsymbol{P}_k| \leq |c_2 \nabla \boldsymbol{O}(\boldsymbol{m}_k)^T \boldsymbol{P}_k| \qquad (4.13)$$

$$0 < c_1 < c_2 < 1 \qquad (4.14)$$

4.4 Importance of frequencies in Frequency domain FWI

When in the frequency domain, monochromatic wave-fields can be added separately in the inversion process. The low frequencies are more linearly related to the model perturbations than the higher frequencies **(*Sirgue, 2003*)**. Therefore the inversion process should start with low frequency components and progressively add higher frequencies. This helps in mitigating some of the non-linearity of the problem, hence increases the chance of getting global minimum.

4.4 Sensitivity to the starting model

The seismic inverse problem is a non linear problem, and the objective function generally contains several minima. Since the method used in this thesis is a local gradient approach, and not a global search in the model space, the gradient will descend to the nearest minimum on the objective function. It is therefore important that the initial model is near to the global minimum, so this will be the nearest minimum. An effort for constructing initial velocity model and its effect on final image reconstruction was made by **Kanlı (2009).** A sufficient initial model should be able to produce synthetic data with a misfit to the corresponding arrivals in the observed data of less than half a period of the starting frequency **(*Sun and McMechan, 1992*)**. Gaussian Smoother function can be used to generate initial model in synthetic case to ensure an adequate initial model. However, in case of real seismic data, first arrival travel time tomography can serve the purpose for initial model.

4.5 Resolution

Resolution is one of the most importance criteria for evaluating the reconstructed image in seismic imaging and acquisition geometries play a vital role in the resolution of the final reconstructed image (**Kanlı et al., 2008; Kanlı, 2008**).

One of the characteristics of the FWI is the high resolution of the results. The method can recover details with an accuracy of $\boldsymbol{\lambda/2}$ **(*Pratt et al., 1996*)** (where $\boldsymbol{\lambda}$ is the wavelength). Compared to travel-time tomography, which can express resolution of approximately $\boldsymbol{\sqrt{\lambda L}}$ (where $\boldsymbol{\lambda}$ is the wavelength and **L** is the offset) **(*Pratt et al., 1996*)**, this is a large improvement.

In seismic Imaging, it is desired to recover as much of the wave-number spectrum as possible. The classical approach in seismic imaging have been to first determine the low wave-

numbers of the velocity field by velocity analysis, followed by a separate reconstruction of the high wave-numbers. The low wave-numbers provides most of the large scale velocity values (tomography like contribution), while the higher wave-numbers improves the location and velocities form (migration-like contribution) **(*Mora, 1987*)**. Full Waveform Inversion has the possibility to recover the whole wave-number spectrum during the inversion process **(*Mora, 1987*)**, hence provide a velocity model where both the velocity values and shapes are accurately estimated.

4.6 Image Reconstruction for Marmousi Model

True model and the initial model (Fig. 4.6 and Fig. 4.7) were kept same as time domain-true and initial model for numerical computation of frequency domain Full Waveform Inversion. The configuration of the sources and receivers are represented in Fig. 4.5. There are 30 sources and 90 receivers equally spaced in the medium at a depth of 20-m. The absorbing boundary conditions were applied on all four edges of the medium. The inversion result after 25 iterations is shown in Figure 4.8.

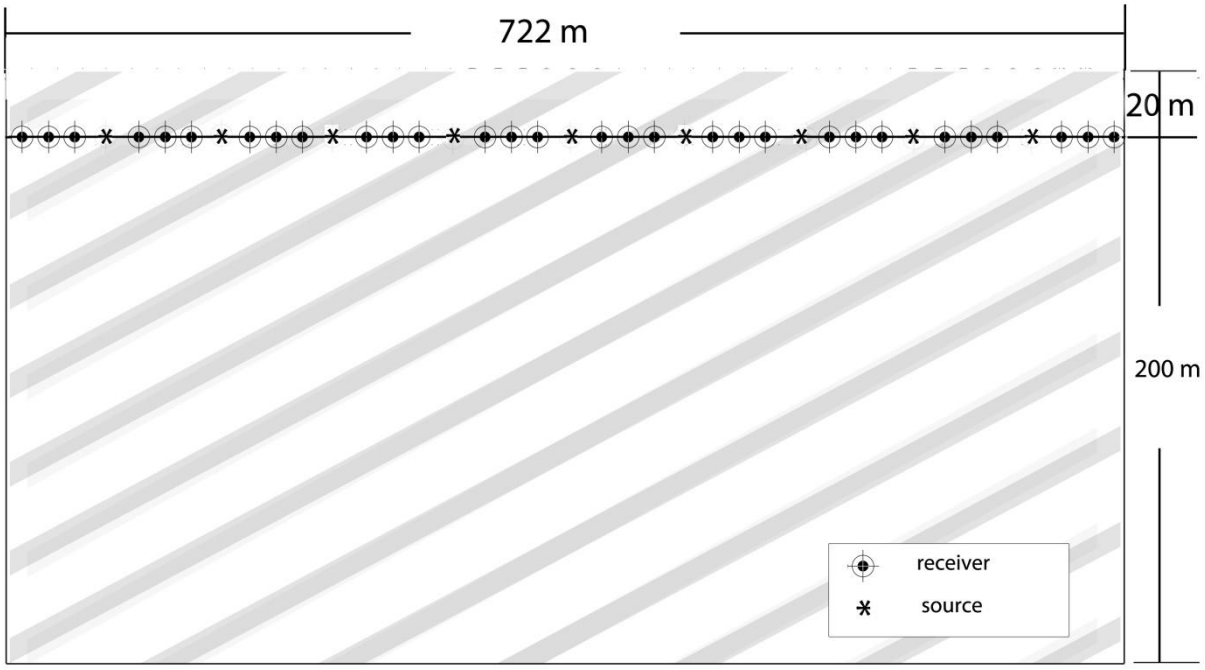


Figure 4.5: The configuration of sources and receiver for Full Waveform Inversion

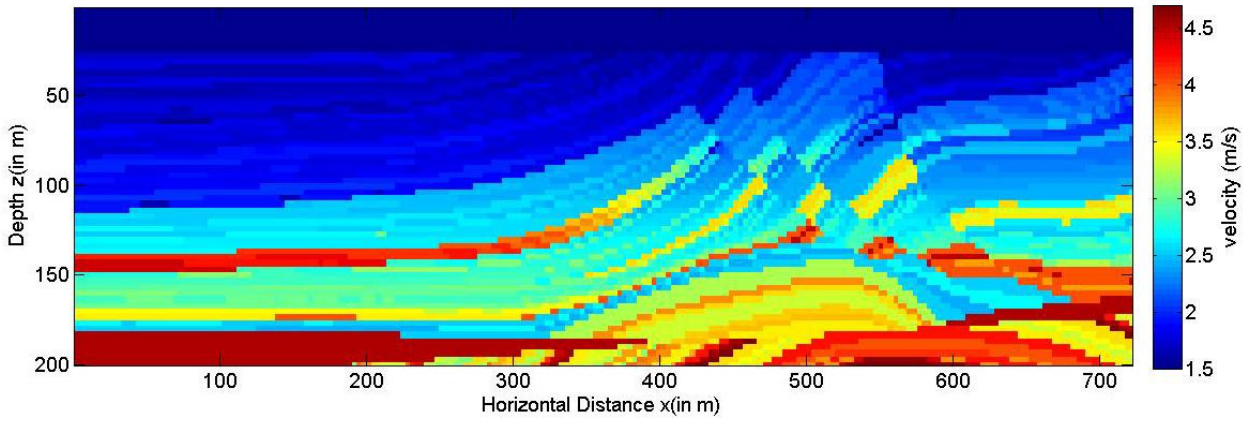


Figure 4.6: True Model for frequency domain FWI

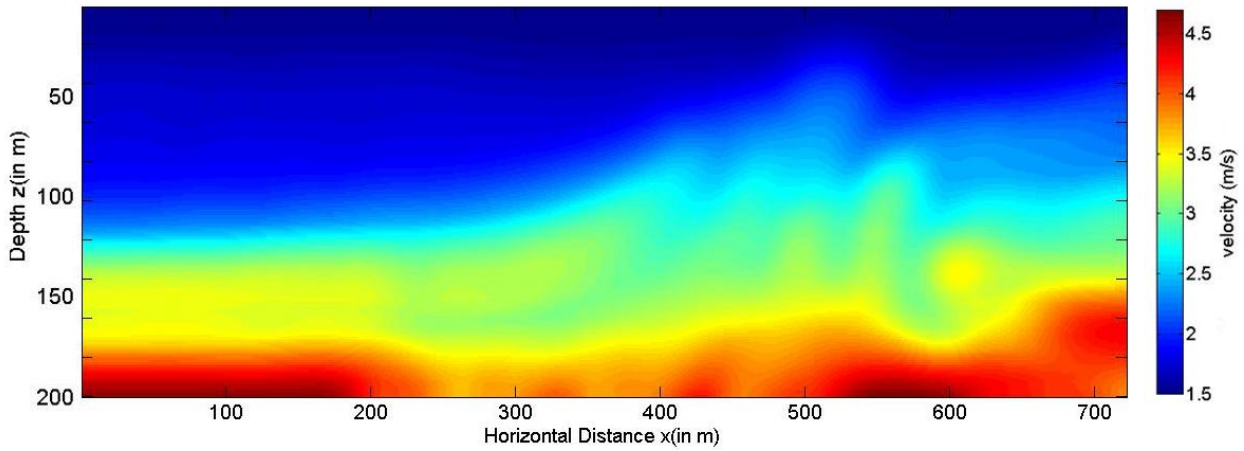


Figure 4.7: Initial Velocity Model

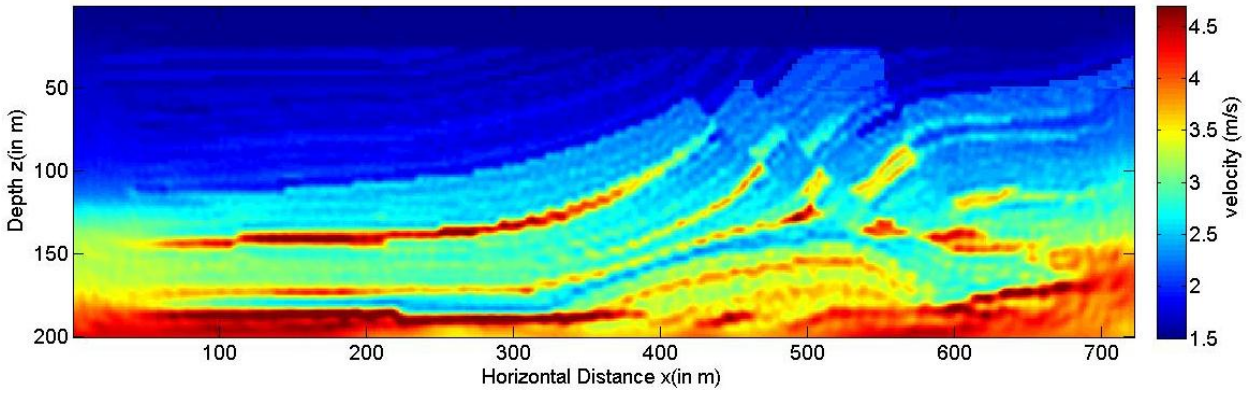


Figure 4.8: Inversion results after 25 iterations.

The objective function value received in frequency domain FWI was monotonically decreasing (Fig. 4.9).

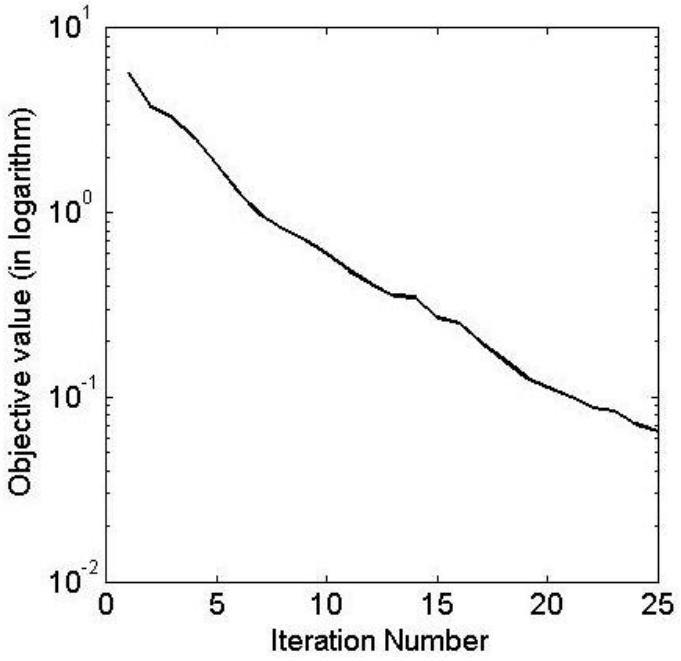


Figure 4.9: Objective function value at each iteration.

4.5.1 Discussion

In order to analysis the amount of velocities recovered in Figure 4.8, velocity logs at three different locations are extracted (at 200 m, 400 m and 600 m) from the true model, initial model and final model. In all three profile (Fig. 4.10(a), 4.10(b) and 4.10(c)) the recovered velocities from FWI result matches the true velocities very well. The largest deviation from the real velocity is about 21% in a small portion of the 600 m profile (see Table 4.1). Some portions of final recovered velocities are either slightly underestimated or overestimated due to transition between high to low (or low to high) velocities. If the average velocity for profile at x= 200m, 400m and 600m are 2.45 m/s, 2.48 m/s and 2.62 m/s, then average deviations from three different profiles will be -0.1%, -0.33%, 0.21 %. Hence, the Full-Waveform Inversion has managed to obtain precise results from the complex Marmousi model in frequency domain.

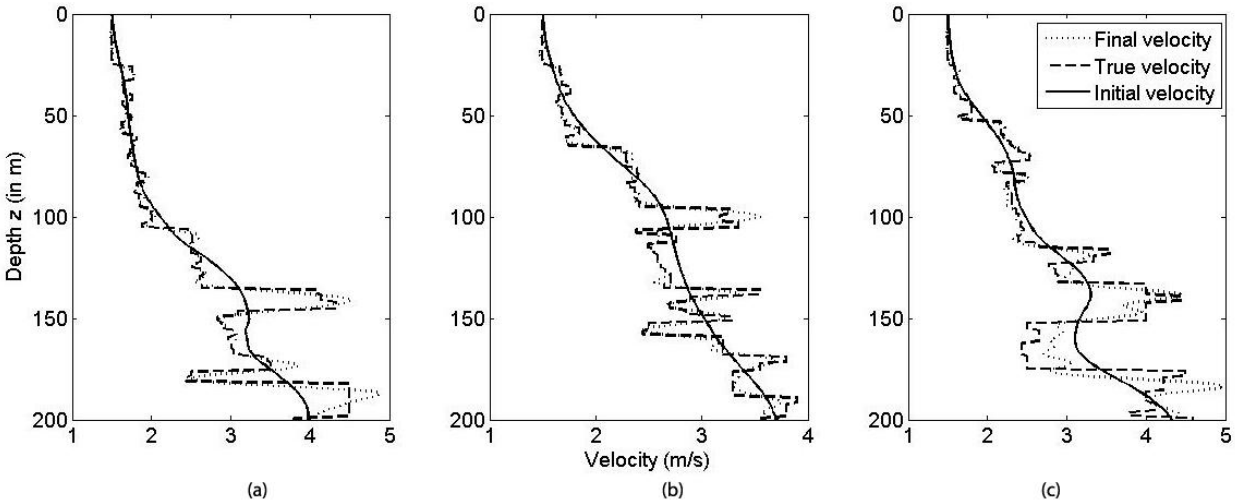


Figure 4.10: Velocity logs extracted at the horizontal distance x (a) 200 m (b) 400 m (c) 600 m.

Table 4.1: Table representing the true velocity, final recovered velocity and the % deviation along with the depth for each profile.

	Profile x=200 m			*Profile x=400 m*			*Profile x=600 m*		
Depth (m)	**True V (m/s)**	**Final V (m/s)**	**(% Deviation)**	**True V (m/s)**	**Final V (m/s)**	**(% Deviation)**	**True V (m/s)**	**Final V (m/s)**	**(% Deviation)**
1	1.5	1.49	-0.07	1.5	1.50	0.08	1.5	1.50	0.02
11	1.5	1.47	-1.41	1.5	1.48	-1.11	1.5	1.46	-2.29
21	1.5	1.52	1.38	1.5	1.48	-1.07	1.5	1.52	1.72
31	1.77	1.77	0	1.67	1.67	0	1.58	1.58	0
41	1.76	1.76	0	1.65	1.65	0	1.66	1.66	0
51	1.74	1.74	-0.02	1.69	1.67	-1.38	1.63	1.65	1.23
61	1.81	1.78	-1.85	1.72	1.72	-0.14	2.20	2.18	-0.69
71	1.70	1.73	1.75	2.29	2.35	2.63	2.53	2.51	-0.64
81	1.93	1.95	1.37	2.40	2.33	-3.07	2.47	2.52	1.91
91	1.84	1.91	3.56	2.39	2.35	-1.70	2.30	2.26	-1.80
101	2.00	1.98	-0.89	3.16	3.45	9.19	2.39	2.35	-1.47
111	2.51	2.57	2.42	2.75	2.60	-5.66	2.4	2.31	-3.55
121	2.63	2.63	0.06	2.58	2.58	0.18	3.35	3.17	-5.35
131	2.61	2.55	-2.23	2.70	2.56	-5.07	2.89	2.95	1.95
141	4.14	4.52	9.06	2.89	2.92	0.98	4.45	4.03	-9.25
151	2.83	2.82	-0.27	3.28	3.11	-5.12	4	3.14	-21.43
161	2.97	3.04	2.32	3.2	3.15	-1.54	2.65	2.79	5.57
171	3.47	3.55	2.27	3.8	3.69	-2.76	2.5	2.99	19.97
181	2.44	2.63	8.08	3.3	3.31	0.39	4.23	4.35	2.85
191	4.5	4.52	0.52	3.9	3.74	-4.06	4	4.00	0.03

CHAPTER 5

CONCLUSIONS

The Full Waveform Inversion method and its capabilities when applied to a complex geological model have been presented in this thesis. The FWI algorithm used in the study, performs an iterative search for a velocity model that minimises the residuals between the data computed in the velocity model and the observed data, i.e. the final result is a "best fit" model. The whole wave-field, including both waveform and phase, is being used as data. FWI computations are implemented in the time domain as well as frequency domain. However, time domain and frequency domain FWI have some disadvantages along with their advantages. Comparing with the frequency domain inversion methods, the time domain methods have the advantage of high efficiency in forward modelling as the wave modelling in frequency domain requires solving a large-scale system at each iteration which is time consuming for large scale problems. Time domain modelling provides the most flexible framework to apply time windowing of arbitrary geometry. In the frequency domain, the computational cost is only proportional to the number of frequencies used in the inversion, and not the number of sources. If the seismic data contains wide angle components, there will be a redundancy in the wave number spectrum present in the data, such a redundancy can be exploited in the frequency domain FWI by inverting fewer frequencies and hence save computational costs because it is possible to invert for single, discrete frequencies one at the time. Another great advantage with the frequency domain implementation is the ease and efficiency when adding multiple sources which is important in seismic methods.

The method is very sensitive to the non-linearity of the problem, and good input parameters, such as an accurate starting model and an initial frequency as low as possible, is important for obtaining good results. In this work it is shown that Full Waveform inversion has a great potential for estimating complex velocity models such as Marmousi Velocity Model when the acquisition parameters are as optimal as possible.

For the future, applying FWI to real data from more complex geological medium and developing a migration tool and test the effect of FWI on a migrated image, are interesting challenges.

APPENDICES

APPENDIX 1: Derivation for Gradient Expression in Time Domain

Considering one dimensional system, if $\boldsymbol{m}$ and $\boldsymbol{d}$ are two arbitrary functions (not necessarily relate to one another by the forward modelling) then by the definition of the adjoint of a linear operator for any $\boldsymbol{d}_r(t)$

$$\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = \langle F^* \boldsymbol{d}, \boldsymbol{m} \rangle \qquad (A1.1)$$

Here F is the linear operator. Left side is in data space and its inner product is given as

$$\langle \boldsymbol{d}, F(\boldsymbol{m}) \rangle = \sum_r \int_0^T \boldsymbol{d}_r(t) F\boldsymbol{m}(x_r, t)dt \qquad \text{here } \boldsymbol{u} = F\boldsymbol{m} \qquad (A1.2)$$

and the right side is in model space and its inner product is given as

$$\langle F^* \boldsymbol{d}, \boldsymbol{m} \rangle = \int_{R^n} (F^* \boldsymbol{d})(x)\boldsymbol{m}(x)dx \qquad (A1.3)$$

The relation $u = Fm$ is implicitly encoded by the two equation (for single source)

$$\left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta\right)\boldsymbol{u} = -\boldsymbol{m}\frac{d^2\boldsymbol{u}_0}{\partial t^2} \qquad (A1.4)$$

$$\left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta\right)\boldsymbol{u}_0 = \boldsymbol{f} \qquad (A1.5)$$

If d_{ext} is the projection of data onto the full wave-field then

$$\boldsymbol{d}_{ext}(x,t) = \sum_r \boldsymbol{d}_r(t)\delta(x - x_r). \qquad (A1.6)$$

Now, $\langle d, F(m) \rangle$ is represented as

$$\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = \int_{R^n} \int_0^T \boldsymbol{d}_{ext}(x,t)\, \boldsymbol{u}(x,t) dx dt \qquad (A1.7)$$

In order to use the wave equation for $\boldsymbol{u}$, the differential operator $\left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta\right)$ need to be materialized which can be done by considering an auxiliary wave-field $q(x,t)$that solves the same wave equation with extended data set $\boldsymbol{d}_{ext}$ on the right hand side (**Demanet, 2015**).

$$\left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta\right) \boldsymbol{q}(x,t) = \boldsymbol{d}_{ext}(x,t) \qquad (A1.8)$$

Substituting the Equation 8 in Equation 7 for and integrating by parts both in time and in space will give

$$\begin{aligned} \langle \boldsymbol{d}, F\boldsymbol{m} \rangle = & \int_V \int_0^T \boldsymbol{q}(x,t) \left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta\right) \boldsymbol{u}(x,t) dx dt \\ & + \int_V \boldsymbol{m}_0 \frac{\partial \boldsymbol{q}}{\partial t} \boldsymbol{u}|_0^T dx - \int_V \boldsymbol{m}_0 \boldsymbol{q} \frac{\partial \boldsymbol{u}}{\partial t}|_0^T dx \\ & + \int_{\partial V} \int_0^T \frac{\partial \boldsymbol{q}}{\partial n} \boldsymbol{u} dS_x dt - \int_{\partial V} \int_0^T \boldsymbol{q} \frac{\partial \boldsymbol{u}}{\partial n} dS_x dt \qquad (A1.9) \end{aligned}$$

In the above equation, V is a volume that covers the whole of R^n, and ∂V is the boundary of V- then in the limit $V = R^n$ the equality holds. In the limit of large V the boundary terms over ∂V disappears as they involve $\boldsymbol{u}$, a wave-field created by localized functions $\boldsymbol{f}, \boldsymbol{m}, \boldsymbol{u}_0$ and which does not have time to travel arbitrarily far within a time $[0,T]$ (**Demanet, 2015**). Because $\boldsymbol{u}(x,z,t=0) = 0$ and $\frac{\partial \boldsymbol{u}(x,t=0)}{\partial t} = 0$ the boundary terms vanishes at $t = 0$. In order to vanish the boundary terms at $t = T$, I need to apply

$$\boldsymbol{q}(x,t=T) = \frac{\partial \boldsymbol{q}(x,t=T)}{\partial t} = 0$$

Since I am only interested in the values of $\boldsymbol{q}(x,t)$ for $t \in [0,T]$, the above are final conditions rather than initial conditions, and the above equation is run backward in time

(**Demanet, 2015**). The boundary conditions need to be time reversed properly. A systematic way to compute $\boldsymbol{q}(x, t = T)$ was discussed in section 3.1.2.

Now, simplifying the left hand side will give

$$\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = \int_{R^n} \int_0^T \boldsymbol{q}(x,t) \left(\boldsymbol{m}_0 \frac{\partial^2}{\partial t^2} - \Delta \right) \boldsymbol{u}(x,t) dx dt$$

$$= \int_{R^n} \int_0^T \boldsymbol{q}(x,t) \boldsymbol{m}(x) \frac{\partial^2 \boldsymbol{u}_0}{\partial t^2} dx dt \qquad (A1.10)$$

As we discussed $\langle \boldsymbol{d}, F\boldsymbol{m} \rangle = \langle F^* \boldsymbol{d}, \boldsymbol{m} \rangle$ Hence, the gradient expression is given as

$$(F^* \boldsymbol{d})(x) = - \int_0^T \boldsymbol{q}(x,t) \frac{\partial^2 \boldsymbol{u}_0}{\partial t^2} dt \qquad (A1.11)$$

Equation 11 can be represented in two dimensional system as

$$(F^* \boldsymbol{d})(x,z) = - \int_0^T \boldsymbol{q}(x,z,t) \frac{\partial^2 \boldsymbol{u}_0(x,z,t)}{\partial t^2} dt \qquad (A1.12)$$

The above equation is known also as imaging condition.

APPENDIX 2: Derivation for Gradient Expression in Frequency Domain

Again considering one dimensional system, the extended data set in frequency domain is given as

$$\widehat{d_{ext}(x,\omega)} = \sum_r \widehat{d_r(\omega)}\delta(x - x_r)\delta(z - z_r) \qquad (A2.1)$$

From Equation (4.7), turning the sum over r into an integral over x then linearised scattered field is given as (**Demanet, 2015**)

$$\widehat{(Fm)}(x_r,\omega) = \int \hat{G}(x,y;\omega)m(y)\omega^2\widehat{u_0}(y,\omega)dy \qquad (A2.2)$$

In order to simplify Equation (4.7), I assume

$$\hat{q}(x,\omega) = \int \overline{\hat{G}(y,x;\omega)}m(y)\omega^2\widehat{u_0}(y,\omega)dy \qquad (A2.3)$$

where G is the Green's function. Using Equation (4.7) and Equation (A2.3), I will have

$$\langle d, Fm\rangle = \int m(y)[2\pi \int_R \hat{q}(y,\omega)\omega^2\overline{\widehat{u_0}(x,z,\omega)}\ d\omega]\, dy \qquad (A2.4)$$

Therefore,

$$F^*\boldsymbol{d}(y) = -2\pi \int_R \widehat{\boldsymbol{q}}(y,\omega)\omega^2\overline{\widehat{\boldsymbol{u}_0}(y,\omega)}d\omega \qquad (A2.5)$$

In two dimensional system Equation (A2.5) can be transformed to

$$F^*\boldsymbol{d}(x,z) = -2\pi \int_R \widehat{\boldsymbol{q}}(x,z,\omega)\omega^2\overline{\widehat{\boldsymbol{u}_0}(x,z,\omega)}d\omega \qquad (A2.6)$$

Equation (A2.6) is same as Equation (3.16) by Parseval's identity. Equation (A2.6) is the integral version of Equation (3.17) in the frequency domain. Since, in Fourier domain, the time

reversal of any real values function becomes the conjugation, the complex conjugate of $\hat{G}$ in Equation (A2.3) presents the fact that the equation is solved backward in time. $\widehat{\boldsymbol{q}} = \bar{\hat{\boldsymbol{G}}}\widehat{\boldsymbol{d}_r}$ can be interpret alternatively by applying an extra conjugate (Eq. A2.7),

$$\overline{\widehat{\boldsymbol{q}}} = \widehat{\boldsymbol{G}}\overline{\widehat{\boldsymbol{d}_r}} \qquad (A2.7)$$

The above equation be read as the sequence of operations given below:

1. Time reversal of $\boldsymbol{d}_r$,
2. Propagate it forward in time, and
3. Time reversal the results.

Following above sequence of operation provides an added advantage of not having to rethink the boundary condition for the backward equation (**Demanet, 2015**).

APPENDIX 3: Converting field seismic data into frequency domain data set.

Pre-processing of seismic data is necessary to convert it into frequency domain data set. In order to convert the time domain seismic data (Fig. 1) into frequency domain data set, each seismic trace must be Fourier transformed. One complex number can then be used to describe the frequency spectrum of a seismic trace at a given frequency component (**Thomassen, 2008**).

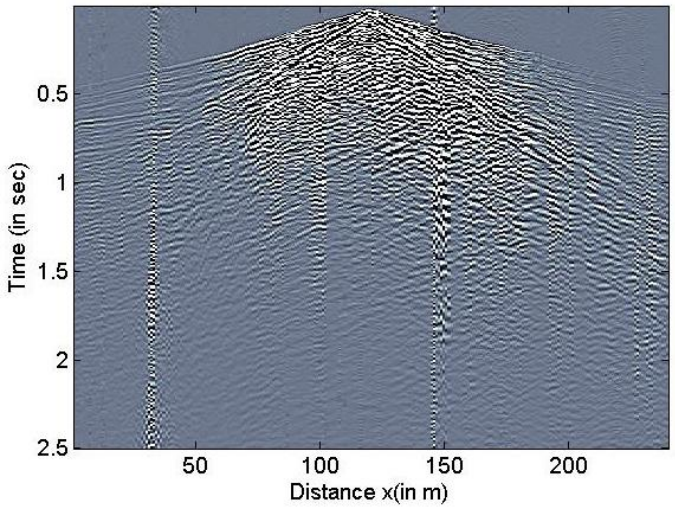


Figure A3.1: Seismic Data or Seismogram

Conversion from time domain data set to frequency domain data set is a simple process. For example, I extracted three traces (Fig. 2) from the seismogram. Each trace is converted into frequency domain using Fourier Transform. Single sided amplitude spectrum after Fourier Transform for each trace is shown in Figure 3. In order to make a data set of 9.8 Hz frequency component (any frequency component can be chosen), the corresponding amplitudes from their amplitude spectrum must be extracted. Figure 4 shows the frequency component data (for 9.8 Hz) where the red dot represents trace number 50, 100 and 150.

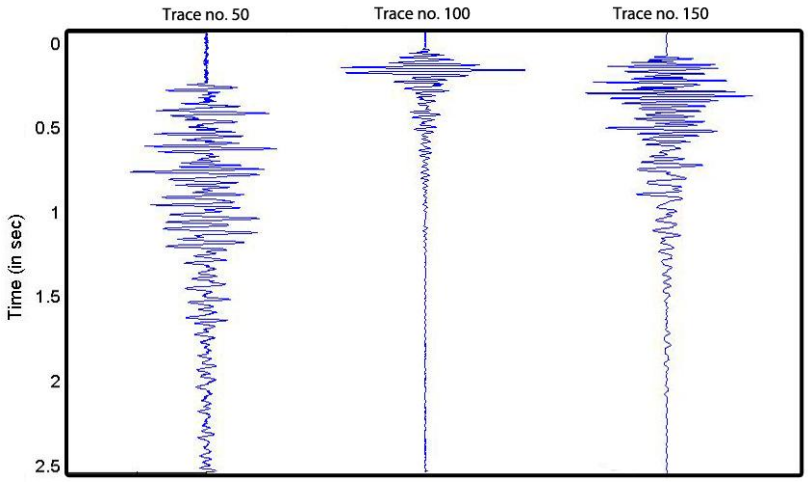


Figure A3.2: Traces extracted from the seismogram.

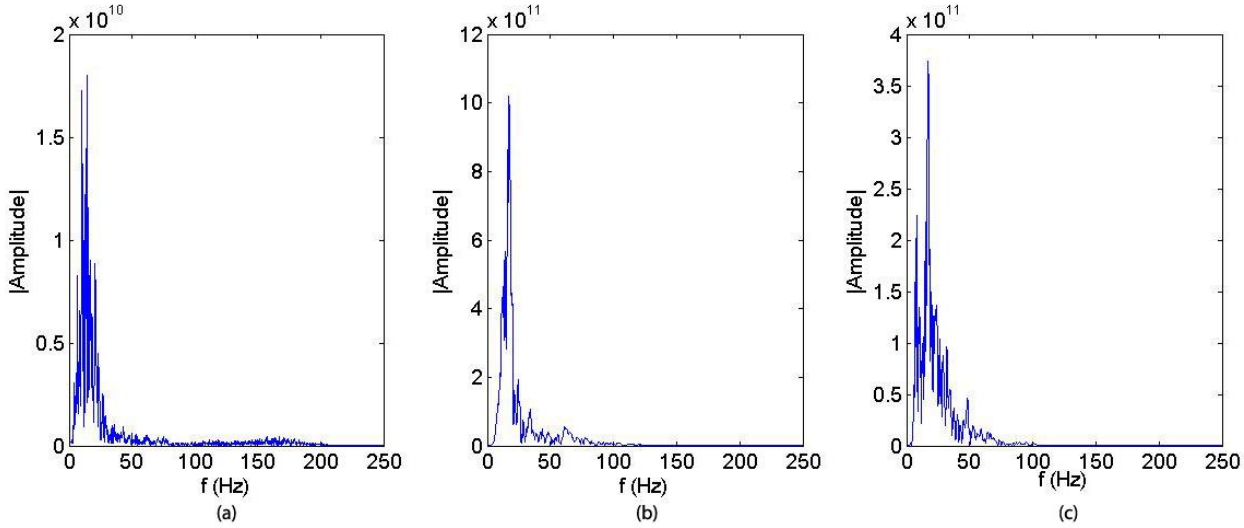


Figure A3.3: Single sided amplitude spectrum with trace no. (a) 50 (b) 100 and (c) 150

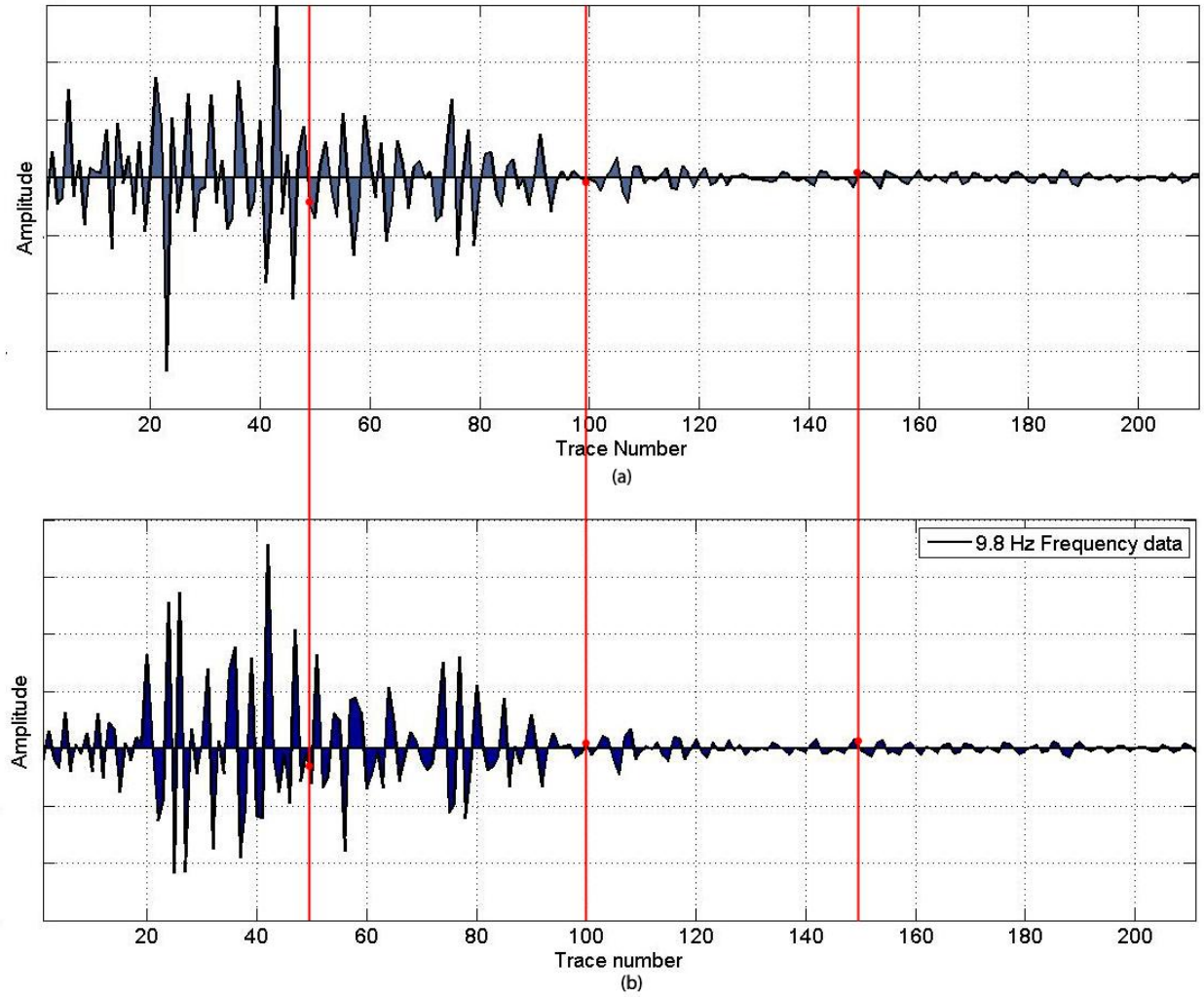


Figure A3.4: Frequency domain data set at 9.8 Hz; (a) real component (b) imaginary component

REFERENCES

Alkhalifah, T., 1998, Prestack Kirchhoff time migration for complex media: Stanford Exploration Project, Report 97, 45-60.

Causse, E., 1998, Seismic migration and inversion in viso-acoustic media: Thesis, Norwegian University of Science and Technology, Norway.

Chabert, A., 2007, Seismic imaging of the sedimentary and crustal structure of the hatton basin on the irish atlantic margin: Thesis, University College Dublin, Ireland.

Chadan, K. and Sabatier, P. C., 1977, Inverse Problems in Quantum Scattering Theory. Springer-Verlag.

Chavent, G., 1974, Identification of parameter distributed systems, in Goodson, R. and Polis, M. eds., Identification of function parameters in partial differential equations: American Society of Mechanical Engineers, 31-48.

Clayton, R. and Engquist, B., 1977, Absorbing Boundary Conditions for Acoustic and Elastic Wave Equations: Bulletin of the Seismological Society of America, 67, 1529-1540.

Collins, M. D., and Kuperman, W. A., 1992, Non-linear inversion for ocean bottom properties: The Journal of the Acoustical Society of America, 92, 2770–2782.

Demanet, L., 2015, Waves and Imaging. Class Notes: http://math.mit.edu/icg/resources/notes325.pdf

Erlangga, Y. A., 2008, Advances in Iterative Methods and Preconditioners for the Helmholtz Equation: Archives of Computational Methods in Engineering, 15, 37-66.

Fichtner, A., 2011, Full seismic waveform modeling and inversion: Advances in Geophysical and Environmental Mechanics and Mathematics, Springer Verlag.

Geoltrain, S., and Brac, J., 1993, Can we image complex structures with first-arrival traveltime?: Geophysics, 58, 4, 564-575.

Gerstoft, P., 1993, Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions: The Journal of the Acoustical Society of America, 95, 770-782.

Hadamard, J., 1902, Sur les problèmes aux dérivées partielles et leur signification physique: Princeton University Bulletin, 49-52.

Hager, W. W., and H. Zhang, 2006, A survey of nonlinear conjugate gradient methods: Pacific journal of Optimization, 2, 35-58.

Hauser, R., 2007, Line search Methods for Unconstrained Optimisation: Lecture 8, Numerical Linear Algebra and Optimisation. *https://people.maths.ox.ac.uk/hauser/hauser_lecture2.pdf*

Kanli, A. I., 2008, Image Reconstruction In Seismic And Medical Tomography: Journal of Environmental and Engineering Geophysics, 13, 85-97.

Kanlı, A. I., 2009, Initial Velocity Model Construction Of Seismic Tomography In Near-Surface Applications: Journal of Applied Geophysics, 67, 52-62.

Kanlı, A. I., Pronay, Z. and Miskolczi, R., 2008, The Importance Of The Spread System Geometry On The Image Reconstruction of Seismic Tomography: Journal of Geophysics and Engineering, 5, 77-85.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations, In: Bednar, J. B., Redner, R., Robinson, E., and Weglein, A., Eds., Conference on Inverse Scattering: Theory and Application, Society for industrial and Applied Mathematics.

Liao, W., and Cao, D., 2013, Full Waveform Inversion of crosswell seismic data using automatic differentiation: GeoConvention Integration: Geoscience Engineering Partnership, 6-12 May 2013, Calgary, AB, Canada.

Lines, L. R., Slawinski, R. and Bording, R. P., 1998, A recipe for stability analysis of finite-difference wave equation computations: 1998 Annual Research Report of the CREWES Project.

Long, G. H., Li, X. F. and Zhang, M. G., 2009, Visco-acoustic transmission waveform inversion for velocity structure in space frequency domain [J]: Acta Seismologica Sinica, 31, 32-41. .

Mora, P., 1987, Elastic wave-field inversion: Thesis, Stanford University, USA.

Pica, A., Diet, J. P. and Tarantola, A., 1990, Non-linear inversion of seismic reflection data in a laterally invariant medium: Geophysics, 55, 284-292.

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

Pratt, R., Song, Z.-M., Williamson, P., and Warner, M., 1996, Two dimentional velocity models from wide-angle seismic data by wave-field inversion: Geophysical Journal International, 124, 323-340.

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

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A., and Dell'Aversana, P., 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by

frequency-domain full-waveform tomography: application to a thrust belt: Geophysical Journal International, 159, 1032–1056.

Sirgue, L., 2003, Frequency domain waveform inversion of large offset seismic data: Thesis, l'Ecole Normale Superieure de Paris, France.

Sourbier, F., Operto, S., and Virieux, J., 2009, FWT2D: A massively parallel program for frequency-domain full-waveform tomography of wide-aperture seismic data-Part 1: Algorithm [J]: Computers & Geosciences, 35, 487-495.

Stekl, I., 1997, Frequency domain seismic forward modelling: A tool for waveform inversion: Thesis, Imperial College London, England.

Sun, R., and McMechan, G. A., 1992, 2-D full-wave-field inversion for wideaperture, elastic, seismic data: Geophysical Journal International, 111, 1–10.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, 51, 1893–1903.

Thomassen, E., 2008, Full Waveform Inversion Studies: Thesis, Norwegian University of Science and Technology, Norway.

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

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, 74, 1-26.

Wang, M., Zhang, D., Yao, Di., Qin, Q. and Xu, L., 2012, Application of Frequency-Domain Waveform Inversion Method in Marmousi Shots Dara: Wuhan University Journal of Natural Sciences, 17, 326-330.

Wencai, Y., 2002, Nonlinear geophysical inversion methods: Review and perspective [J]: Progress in Geophysics, 255-261.

Youzwishen, C. F. and Margrave, G. F., 1999, Finite difference modelling of acoustic waves in MATLAB: Annual Research Report, CREWES Project, 11.

Wolfe, P., 1969, Convergence Conditions for Ascent Methods", SIAM Rev., 11, 2, 226-235.

Wolfe, P., 1971, Convergence Conditions for Ascent Methods II: Some Corrections, SIAM Rev., 13, 185-188.

Zhang, R. and Yuan, C., 2009, Application of the seismic waveform inversion technique [J]: Natural Gas Industry, 29, 52-54.

Zhang., W. and Luo, J., 2013, Full-waveform Velocity Inversion Based on the Acoustic Wave Equation: American Journal of computational Mathematics, 2013, 13-20.