IDEA League

Master of Science in Applied Geophysics

Research Thesis

Full-waveform inversion of seismic data using instantaneous-phase coherency

Jan Nicolaas Schaly

August 5, 2022

Full-waveform inversion of seismic data using instantaneous-phase coherency

Master of Science Thesis

for the degrees of Master of Applied Earth Sciences and Master of Science in Applied Geophysics

by

Jan Nicolaas Schaly

August 5, 2022

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *August 5, 2022*

Supervisor(s):

Deyan Draganov

Myrna Staring

Committee Members:

Deyan Draganov

Myrna Staring

Florian Wellmann

Kees Weemstra

Abstract

With a rapid increase in computational resources two dimensional full-waveform inversion is evolving into a promising tool for near-surface geophysics. However, near-surface applications suffer from local minima due to amplitude errors associated with two dimensional full-waveform inversion. By clever definition of the misfit function, the influence of the amplitude errors can be mitigated. Here, I use a recently proposed misfit function based on the instantaneous-phase coherency. The instantaneous-phase coherency misfit function uses complex trace analysis to create an amplitude unbiased misfit function. First, I compare the new misfit function to a traditional least-squares misfit function by inverting synthetic models where noise is added, a field dataset containing Rayleigh waves and a field dataset containing Love waves. Next, I perform inversions on layer cake models to investigate the accuracy of the full-waveform inversion using the new misfit function and finally, I test the robustness of the inversion by using a complex subsurface model. The inversions performed on the synthetic models show that the instantaneous-phase coherency misfit is more robust when noise is introduced to the data compared to the least-squares misfit. Furthermore the two field datasets, demonstrate the ability of the instantaneous-phase to deliver accurate near-surface results when used on field data. The results from the layer cake inversions where inconclusive, however I did demonstrate that a better selection of the bandwidth did improve the result. Finally, the results from the complex subsurface model show that the instantaneous-phase coherency is able to resolve parts of the complex subsurface model.

Acknowledgements

First, I would like to express my gratitude to Jianhuan Liu for investing his spare time to contribute to this master thesis. I would like to thank him for helping me get acquainted with the inversion software. Also, for always being available to give feedback and have a discussion with me.

From Fugro, I want to thank my supervisors Eline Leentvaar and Myrna Staring. For all the weekly startup meetings and the advice throughout my internship. Their enthusiasm and can do attitude has inspired me to put my best effort into this master thesis project.

From TU Delft, I would like to thank my supervisor Deyan Draganov, for advising me throughout this master thesis. Additionally, I would like to thank my graduation committee.

Finally, I would like to thank my family, my girlfriend and friends for the support and endurance throughout all my years of study.

To perform my thesis research I made use of the DENISE modelling/inversion code (Köhn, 2011; Dokter et al., 2017). The P/SV wave code and SH/Love wave code are available at, respectively, `https://github.com/daniel-koehn/DENISE-Black-Edition` and `https://github.com/daniel-koehn/DENISE-SH`. The SH field data is an open source dataset made available by Köhn et al. (2019) and can be found at `https://www.dropbox.com/s/5jgr389tbllmefd/Fossa_Carolina_LBL_FWI_repro_pack.rar?dl=0`.

Technische Universiteit Delft
August 5, 2022

Jan Nicolaas Schaly

Table of Contents

List of Figures

Acronyms

BC boundary conditions

CFL Courant-Friedrichs-Lewy

C-PML convolutional perfectly matched layers

DENISE subwavelenght detail resolving nonlinear iterative seismic inversion

FATT first arrival traveltime tomography

FD finite difference

FWI full-waveform inversion

L-BFGS limited-memory Broyden, Fletcher, Goldfarb, and Shanno

IC initial conditions

MASW multichannel analysis of surface-waves

MPI message passing interface

PDEs partial differential equations

PE processing elements

SNR signal-to-noise ratio

SSG standard staggered grid

SWMs surface-wave methods

TSE Taylor series expansion

V_p P-wave velocity

V_s S-wave velocity

Chapter 1

Introduction

The use of surface-wave methods (SWMs) is growing in popularity as they present a non-invasive method for near-surface characterisation. Therefore, SWMs are used by various disciplines to perform near-surface characterisation. Conventional SWMs such as multichannel analysis of surface-waves (MASW) however rely on the extraction and inversion of dispersion curves to obtain a 1D S-wave velocity (V_s) profile of the near-surface (Park et al., 1999; Xia et al., 1999). The 1D V_s profile is then interpolated to deliver a 2D profile of the near-surface. This means that methods like MASW are not able deal with strong lateral heterogeneity (Pan et al., 2019).

With the rapid increase in computational power, full-waveform inversion (FWI) is evolving into a promising tool for near-surface geophysical applications. For example, Groos et al. (2017), invert Rayleigh waves to reconstruct small scale structures in the shallow subsurface. Whereas, Köhn et al. (2019) successfully invert Love waves to characterise a medieval canal structure. FWI is a data fitting procedure that aims to find the best possible set of model parameters by using the full-waveform contained in the seismic data (Virieux and Operto, 2009). By using the full-waveform, FWI is able to deal with lateral heterogeneities unlike conventional SWMs such as MASW. Furthermore, since the entire waveform is used in the inversion process, images of the near-surface can be obtained at sub-wavelength resolution (Köhn, 2011; Virieux and Operto, 2009).

However, near-surface application of FWI is still somewhat limited due to the sensitivity of the misfit functions to amplitude errors. If the recorded waveforms are contaminated with amplitude errors, cycle skipping could occur and the ill-posedness of the inverse problem is increased. These amplitude errors are associated with 2D FWI of field data and are caused by non-uniform source amplitudes, variable coupling of receiver positions and insufficient 3D to 2D correction of geometrical spreading (Maurer et al., 2012; Heider, 2014; Forbriger et al., 2014; Schäfer et al., 2014). To reduce the ill-posedness of the FWI, multiple researchers have proposed alternative methods to define the misfit function. For example, the decomposition of the wavefield by application of a sequential approach where the bandwidth is increased per inversion sequence, called multiscale-FWI (Bunks et al., 1995; Fichtner et al., 2013; Groos et al., 2017). Another possibility is to define a misfit function that separates the phase and

amplitude of the data (Bozdağ et al., 2011) or by defining a misfit function that is fully unbiased by the amplitude of the data (Liu et al., 2022).

Here, the misfit function based on the instantaneous-phase coherency recently proposed by Liu et al. (2022) is used. The instantaneous-phase coherency makes use of complex trace analysis to implicitly define the instantaneous-phase of the data (Liu et al., 2022). By incorporating the instantaneous-phase cohency misfit into FWI an amplitude unbiased measure of the misfit is created. The instantaneous-phase coherency FWI should therefore provide more robust results, when compared to a least-squares norm FWI in near-surface applications.

1-1 Thesis objective

The primary objective of this thesis is to test the performance of the instantaneous-phase coherency misfit function designed by Liu et al. (2022). To test the performance of the instantaneous-phase coherency FWI, multiple inversions where performed on synthetic test cases and field datasets.

First, the performance of instantaneous-phase coherency FWI is compared to a traditional least-squares norm FWI. To do so a synthetic test case was used, where the signal-to-noise ratio (SNR) was reduced by adding random noise to the synthetic seismograms resembling the previously described amplitude errors. Next, the two misfit functions where applied to a field dataset dominated by Rayleigh waves and finally a field dataset containing Love waves.

Afterwards, the instantaneous-phase coherency was tested on synthetic models based on results of a geo-technical and geophysical survey performed by Fugro at an undisclosed location in the Middle East. To test the accuracy of the instantaneous-phase coherency FWI, layer cake models where drawn from one of the PS-logs of which the depth and thickness of the layers were randomly varied slightly, while using the PS-log as initial model. Finally, the result of the geophysical survey was used to create a near-surface model to test how well the instantaneous-phase coherency FWI is able to resolve complex near-surface features.

Chapter 2

Theory

This chapter provides an overview of the theory used in this thesis. First, the forward-modelling approach is reviewed and second the theory of FWI is reviewed.

2-1 The forward-modelling approach

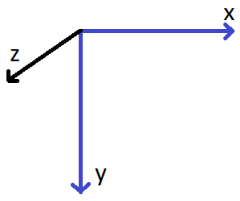


Figure 2-1: Coordinate system used in the DENISE modelling code.

Forward-modelling is a key part of the inversion process, since being able to extract information about the subsurface structure and composition from a seismic dataset requires the knowledge of wavefield propagation and how these wavefields are influenced by the subsurface. Unfortunately, an exact solution to the wave equation is generally not possible for most subsurface configurations. Hence, solutions of the wave equation are obtained through numerical methods, which in the case of seismic modelling are used to generate synthetic seismograms. Here the finite difference (FD) method, originally proposed by Virieux (1986), is used.

Throughout this thesis, it is assumed that all media are isotropic elastic, allowing 3D seismic wave propagation to be simplified to 2D seismic wave propagation, by assuming all motion takes place in the x-y plane. Therefore, 3D wave propagation is divided into two categories during the thesis: P/SV-wave propagation, where particle displacement only takes place in the x and y directions and SH/Love-wave propagation, where the particle displacement is along the z-axis in the x-y plane (Fig. 2-1).

In the following sections, first the formulation of the equation of motion used in the modelling code is given for both P/SV-waves and SH-waves. Next, the FD method is explained and

its implementation in the subwavelenght detail resolving nonlinear iterative seismic inversion (DENISE) modelling code (Köhn, 2011).

2-1-1 Wave propagation in an elastic medium

To describe the propagation of seismic waves in a general elastic medium the following system of linear partial differential equations (PDEs) is used

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i, \tag{2-1}$$

which is a formulation of the elastic equation of motion where the subscripts i,j denote either the x, y or z direction, ρ is the density, u_i the particle displacement, σ_{ij} the stress tensor and f_i being the body forces exerted on the medium. Eq. 2-1 describes the state of a general elastic medium. The influence of a force on the material properties of the medium can be described by the linear stress-strain relationship:

$$\sigma_{ij} = \lambda \theta \delta_{ij} + 2\mu \epsilon_{ij}, \tag{2-2}$$

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{2-3}$$

with strain ϵ_{ij}, the cubic dilatation of the medium $\theta = \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}$, and δ_{ij} the Kronecker's delta. The DENISE FD modelling code uses an alternate formulation of the elastic equation of motion, which is obtained by taking the time derivative of Eqs. 2-4 and 2-5. Assuming that the Lamé parameters λ and μ are independent of time and by taking into account that the time derivative of the particle displacement is equal to the particle velocity $\frac{\partial u_i}{\partial t} = v_i$, a new system of linear PDEs is obtained:

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i, \tag{2-4}$$

$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda \frac{\partial \theta}{\partial t} \delta_{ij} + 2\mu \frac{\partial \epsilon_{ij}}{\partial t}, \tag{2-5}$$

$$\frac{\partial \epsilon_{ij}}{\partial t} = \frac{1}{2}\left(\frac{\partial \partial u_i}{\partial t \partial x_j} + \frac{\partial \partial u_j}{\partial t \partial x_i}\right) = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right). \tag{2-6}$$

P/SV-wave propagation in an isotropic elastic medium

As mentioned above, during P/SV-wave propagation particle motion takes place in the x-y plane and will therefore be the only plane with non-zero particle displacement. For P/SV-wave propagation the stress-velocity formulation is used, resulting in the following system of equations (Ben-Menahem and Singh, 2012; Köhn, 2011):

$$\rho \frac{\partial v_x}{\partial t} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + f_x, \tag{2-7}$$

$$\rho \frac{\partial v_y}{\partial t} = \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + f_y, \tag{2-8}$$

where the particle velocities v_x and v_y together compose the particle velocity vector v_i and the body forces f_x and f_y compose the body force vector f_i. Taking the time derivative of the stress tensor results in the following equations

$$\frac{\partial \sigma_{xx}}{\partial t} = (\lambda + 2\mu) \frac{\partial v_x}{\partial x} + \lambda \frac{\partial v_y}{\partial y} + \frac{\partial \sigma_{xx0}}{\partial t}, \tag{2-9}$$

$$\frac{\partial \sigma_{yy}}{\partial t} = \lambda \frac{\partial v_x}{\partial x} + (\lambda + 2\mu) \frac{\partial v_y}{\partial y} + \frac{\partial \sigma_{yy0}}{\partial t}, \tag{2-10}$$

$$\frac{\partial \sigma_{xy}}{\partial t} = \mu \left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x} \right) + \frac{\partial \sigma_{xy0}}{\partial t}, \tag{2-11}$$

with σ_{xx0}, σ_{yy0} and σ_{xy0} being the components of the surface force. Since the medium is isotropic it is sufficient to define only the trace elements of the stress tensor σ_{xx}, σ_{yy} and one of the non-diagonal elements σ_{xy}.

SH/Love propagation

During SH/Love-wave propagation particle displacement takes place parallel to the z-direction and inside the x-y plane. Therefore, Eq. 2-1 only requires the xz and yz components resulting in the following system of PDEs:

$$\rho \frac{\partial v_z}{\partial t} = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial y} + f_z, \tag{2-12}$$

with particle velocity v_z and body force f_z. When defining the stress tensor for the SH/Love-wave case the dilatation will be zero, resulting in the following stress-strain relation:

$$\frac{\partial \sigma_{xz}}{\partial t} = \mu \left(\frac{\partial v_z}{\partial x} \right), \tag{2-13}$$

$$\frac{\partial \sigma_{zy}}{\partial t} = \mu \left(\frac{\partial v_z}{\partial y} \right). \tag{2-14}$$

2-1-2 Finite differences

For simple cases, the system of PDEs describing P/SV-wave propagation (Eqs. 2-7, 2-8 and 2-9) and the system of PDEs for SH/Love-wave propagation (Eq. 2-12) can be solved analytically. However, in most cases a numerical solution is required as the medium will be too complex to solve analytically. Here, the FD approach described by Virieux (1986) is used.

To obtain a numerical solution of the elastic equation of motion, Eqs. 2-7, 2-8, 2-9, 2-10 and 2-11 must be discretized in both time and space on a grid. Hence, the particle velocity v_i, the stress tensor σ_{ij} and the Lamé parameters λ and μ must be calculated at discrete grid points. The DENISE modelling code uses a Cartesian standard staggered grid (SSG) (Virieux, 1986) with coordinates x and y respectively being the discrete grid points i and j with grid spacing dh, with $i \in N|\,[1, NX]$ and $j \in N|\,[1, NY]$. Furthermore, the time t is defined as n with time step dt with $n \in N|\,[1, NT]$. The symbols NX and NY denote the spatial grid points in the x and y directions, respectively, and NT the total amount of time steps. The distribution of these parameters on the standard staggered grid used in DENISE is shown in Fig. 2-2.

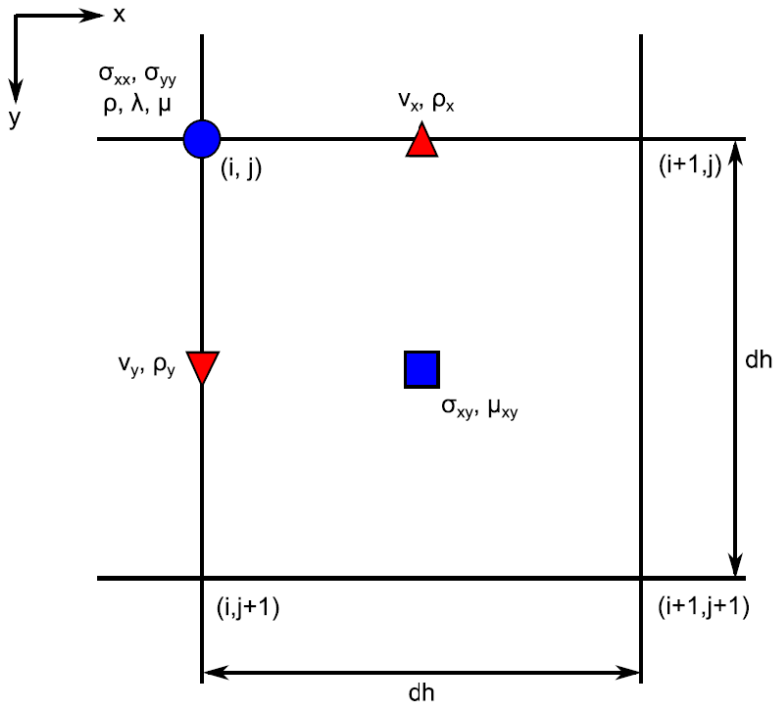


Figure 2-2: The grid geometry used by the DENISE modelling code (Köhn, 2011) based on the standard staggered grid originally proposed by Virieux (1986).

Moreover, the partial derivatives must be replaced with FD operators. The partial derivatives are obtained through means of a Taylor series expansion, which can be varying in length depending on the required accuracy. Here, a forward shifting operator D_k^+ is defined

$$D_k^+ [f_j] = \left[\frac{\partial f}{\partial x}\right]_{j+\frac{1}{2}} = \sum_{n=1}^{N} \beta_n \frac{f_{j+n} - f_{j-n+1}}{\Delta x}, \tag{2-15}$$

and a backward shifting operator D_x^-

$$D_k^+ [f_j] = \left[\frac{\partial f}{\partial x}\right]_{j-\frac{1}{2}} = \sum_{n=1}^{N} \beta_n \frac{f_{j+n-1} - f_{j-n}}{\Delta x}, \tag{2-16}$$

defining the partial derivative at location j with β_n are the weighting coefficients of the SSG which on a uniform grid are computed through means of a Taylor series expansion (TSE). The subscript k denotes the Cartesian coordinate direction on the grid. In this thesis a 4^{th} order operator is used for determining the partial derivatives. The order of the operator is determined by $2N$, which is the length of the TSE. For a 4^{th} order operator this means $N = 2$ and thus two equations are added to the operator (Falk et al., 1996).

Using the operators defined in Eqs. 2-15 and 2-16 the velocity field and stress tensor are updated for each grid location. The DENISE modelling code updates the stresses as follows

$$\sigma_{xx}^{n+\frac{1}{2}} [i] [j] = \sigma_{xx}^{n-\frac{1}{2}} [i] [j] + \Delta t \left((\lambda [i] [j] + 2\mu [i] [j]) D_x^- \left[v_x^n [i + \frac{1}{2}] [j] \right] + \lambda [i] [j] D_y^- \left[v_y^n [i] [j + \frac{1}{2}] \right] \right), \tag{2-17}$$

$$\sigma_{yy}^{n+\frac{1}{2}} [i] [j] = \sigma_{xx}^{n-\frac{1}{2}} [i] [j] + \Delta t \left(\lambda [i] [j] D_y^+ \left[v_x^n [i + \frac{1}{2}] [j] \right] + (\lambda [i, j] + 2\mu [i] [j]) D_y^+ \left[v_y^n [i] [j + \frac{1}{2}] \right] \right), \tag{2-18}$$

$$\sigma_{xy}^{n+\frac{1}{2}} = \sigma_{xy}^{n-\frac{1}{2}} [i + \frac{1}{2}][j + \frac{1}{2}] + \Delta t \mu \left[i + \frac{1}{2} \right] \left[j + \frac{1}{2} \right] \left(D_y^+ \left[v_x^n [i + \frac{1}{2}] [j] \right] + D_x^+ \left[v_y^n [i] [j + \frac{1}{2}] \right] \right), \tag{2-19}$$

and the velocity field

$$v_x^{n+1} [i + \frac{1}{2}] [j] = v_x^n [i + \frac{1}{2}] [j] + \frac{\Delta t}{\rho [i + \frac{1}{2}] [j]} \left(D_x^- \left[\sigma_{xx}^{n+\frac{1}{2}} [i] [j] \right] + D_y^- \left[\sigma_{xy}^{n+\frac{1}{2}} [i + \frac{1}{2}][j + \frac{1}{2}] \right] + f_x^{n+\frac{1}{2}} \right), \tag{2-20}$$

$$v_y^{n+1} [i] \left[j + \frac{1}{2} \right] = v_y^n [i] \left[j + \frac{1}{2} \right] + \frac{\Delta t}{\rho [i] [j + \frac{1}{2}]} \left(D_x^- \left[\sigma_x^{n+\frac{1}{2}} [i + \frac{1}{2}][j + \frac{1}{2}] \right] + D_y^+ \left[\sigma^{n+\frac{1}{2}_{yy}} [i] [j] \right] + f_y^{n+\frac{1}{2}} \right). \tag{2-21}$$

Furthermore, the model parameters μ and ρ have to be determined. For stability on the SSG, the shear modulus μ has to be harmonically averaged and the density ρ arithmetically averaged (Bohlen and Saenger, 2006)

$$\langle \mu \rangle [i + \frac{1}{2}][j + \frac{1}{2}] = 4 \left(\mu^{-1} [i] [j] + \mu^{-1} [i + 1] [j] + \mu^{-1} [i] [j + 1] + \mu^{-1} [i + 1] [j + 1] \right)^{-1}, \tag{2-22}$$

$$\rho_x [i + \frac{1}{2}] [j] = \frac{1}{2} \left(\rho [i] [j] + \rho [i + 1] [j] \right), \tag{2-23}$$

$$\rho_y [i] [j + \frac{1}{2}] = \frac{1}{2} \left(\rho [i] [j] + \rho [i] [j + 1] \right), \tag{2-24}$$

with $\langle \mu \rangle$ being the harmonically averaged shear modulus.

2-1-3 Initial and boundary conditions

In general, numerical simulations of seismic wave propagation in any unbound media varying in scale from local to global settings require a set of initial conditions (IC) and boundary conditions (BC) that ensure the medium is in rest before starting the simulation and cap of the computational domain, respectively.

Initial conditions

The set of IC used here, enforce the medium to be in rest before the simulation starts. The DENISE code sets the particle velocity and the time derivative of the particle velocity to zero

$$v_i\left(\mathbf{x}, t=0\right)=0,\ x \in \mathbf{V}, \tag{2-25}$$

$$\frac{\partial v_i\left(\mathbf{x}, t=0\right)}{\partial t}=0,\ x \in \mathbf{V}, \tag{2-26}$$

at time $t=0$ with $\mathbf{V}$ being the volume on which the IC are imposed.

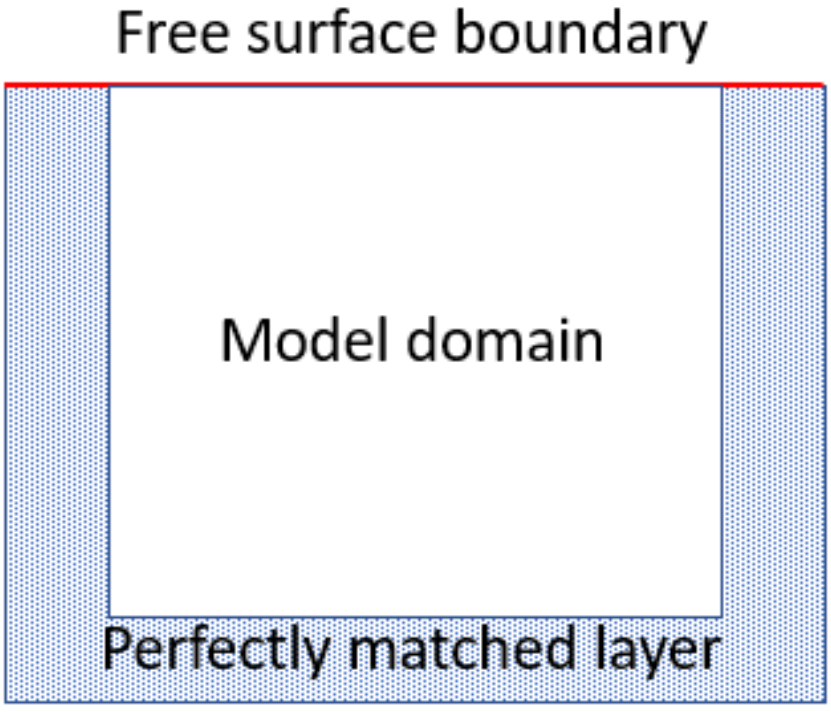


Figure 2-3: General concept of the implementations of the boundary conditions.

Free-surface boundary

Free-surface boundaries are used to mimic the interface between air and the earth's surface and are of great importance for near-surface seismic implementations as the surface of the earth has a strong impact on the outcome of the simulation (Robertsson et al., 1995). Free-surface

boundaries can be defined in an implicit or explicit manner. The DENISE modelling code uses an explicit implementation of the free-surface boundary. The explicit implementation of a free-surface boundary is achieved through a mirroring technique designed by Levander (1988), which is a stable and accurate manner in the case of a plane interface (Robertsson et al., 1995; Levander, 1988). The mirroring technique assumes that a planar free-surface is located at $i = h$ and that the stresses at this free-surface are set to zero; furthermore, the stresses directly below the free-surface are mirrored through means of an inverse sign resulting in the following definitions of the stress tensor (Levander, 1988; Robertsson et al., 1995; Köhn, 2011)

$$\sigma_{yy}(h, j) = 0, \tag{2-27}$$

$$\sigma_{yy}(h-1, j) = -\sigma_{yy}(h+1, j), \tag{2-28}$$

$$\sigma_{xy}\left(h-\frac{1}{2}, i+\frac{1}{2}\right) = -\sigma_{xy}\left(h+\frac{1}{2}, j+\frac{1}{2}\right), \tag{2-29}$$

$$\sigma_{xy}\left(h-\frac{3}{2}, i+\frac{1}{2}\right) = -\sigma_{xy}\left(h+\frac{3}{2}, j+\frac{1}{2}\right), \tag{2-30}$$

$$\sigma_{xx}(h, j) = \frac{4\left(\lambda\mu+\mu^2\right)}{\lambda+2\mu} u_{xx}. \tag{2-31}$$

Perfectly matched layers

Numerical simulations of the wave equation in any unbound medium generally require an artificial boundary to limit the size of the computational domain to the scope of the study, which can vary from local to global scale. These artificial boundaries cause spurious energy inside the modelling domain from waves reflecting of these boundaries (Igel, 2017). Therefore, a boundary layer absorbing wave energy rather than reflecting it back in to the medium is created surrounding the modelling domain (Fig. 2-3). Here, convolutional perfectly matched layers (C-PML) are used, which multiply the wavefield with a dampening function

$$c = -V_{PML} * \frac{log(\alpha)}{L}, \tag{2-32}$$

where V_{PML} denotes the P-wave velocity inside the absorbing boundary frame, the constant $\alpha = 1 \times 10^{-4}$, and the thickness of the boundary L (Komatitsch and Martin, 2007; Köhn, 2011). In this thesis V_{PML} is defined by the velocity at the edge of the model domain.

2-1-4 Stability of the simulation

In setting up any FD simulation some key steps should be undertaken to ensure a stable simulation. First, one should determine the minimum wavelength that is possible to model. Next, the spatial grid spacing Δx should be defined, and finally the temporal time step Δt should be chosen. The minimum wavelength can easily be determined by

$$\lambda_{min} = \frac{V_{min}}{f_{max}}, \tag{2-33}$$

with λ_{min} being the minimum wavelength, V_{min} being the minimum propagation velocity in the velocity model, and f_{max} being the maximum frequency of the source signal.

Generally, the Δx and Δt are optimised to decrease the computing time and performance demands for the simulation. In doing so, a balance is sought that allows for as large as possible Δx and Δt without the simulation of the wavefield being unstable. Therefore, the DENISE modelling code tests if the simulation satisfies the temporal and spatial sampling criterions (Köhn, 2011).

Grid dispersion

In constructing an FD grid selecting the appropriate spatial sampling rate Δx is of great importance, as a too large sampling rate will lead to grid dispersion. Grid dispersion is best demonstrated by the example shown in Fig. 2-4 where snapshots of two FD simulations of the seismic wave propagation are displayed. As can be seen in Fig. 2-4 a too large Δx leads to an increasingly dispersive wavefield during the simulation resulting in an inaccurate sampling of the wavefield, hence the name grid dispersion. To ensure a sufficiently small spatial sampling Δx rate is used the DENISE code checks if the spatial sampling criterion is met. Here, a modified version of the Nyquist-Shannon sampling theorem is used

$$\Delta x \leq \frac{\lambda_{min}}{n} = \frac{V_{min}}{n f_{max}}. \tag{2-34}$$

Parameter n varies depending on the length of the FD operators defined in Eqs. 2-15 and 2-16, in the original Nyquist-Shannon sampling theorem $n = 2$; however, in this thesis an operator length of 4 was used resulting in $n = 8$ (Köhn, 2011).

Temporal instability

To ensure stability of FD simulations it is required to have a small enough time step Δt. The influence of Δt is illustrated by Fig. 2-5 where snapshots of two FD simulations of the seismic wave propagation are displayed. From Fig. 2-5 it can be observed that an inappropriate Δt leads to a disproportional increase in energy from the wavefield resulting in a unstable simulation. To ensure a stable simulation of the seismic wavefield, the DENISE code checks if the temporal sampling criterion is met. The temporal sampling criterion or Courant-Friedrichs-Lewy (CFL) criterion is defined by

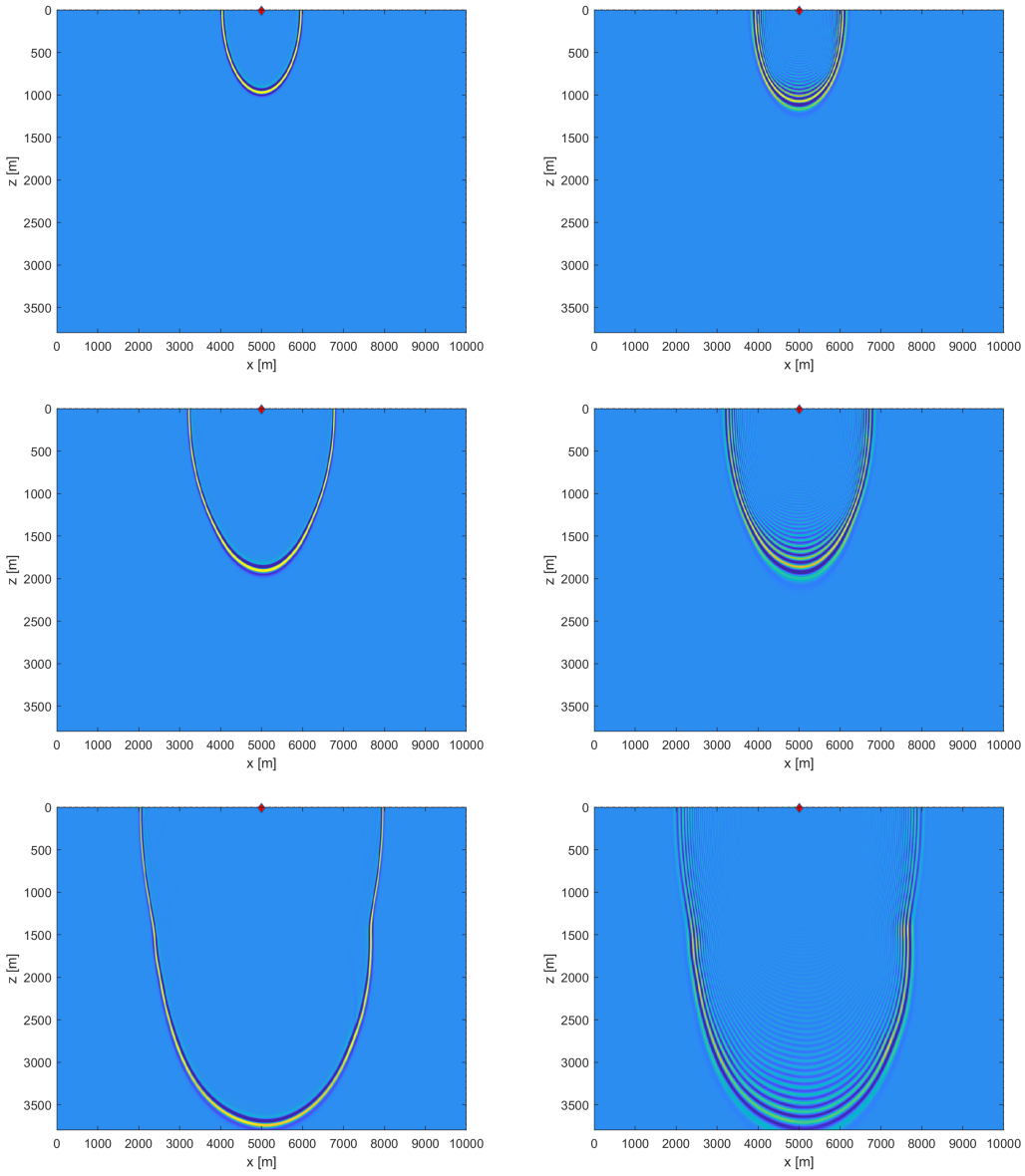


Figure 2-4: Snapshots of an FD simulation demonstrating a simulation without grid dispersion and a simulation where grid dispersion occurs. The left column displays snapshots of the simulation where Δx satisfies the sampling criterion. The right column displays snapshots when Δx does not satisfies the sampling criterion causing dispersion of the wavefield.

$$\Delta t \leq \frac{\Delta x}{h\sqrt{2}V_{max}}, \tag{2-35}$$

with the maximum velocity inside the velocity model V_{max} and parameter h. Parameter h is defined by summing over all the weighting coefficients β_n in the FD operators defined in Eqs. 2-15 and 2-16. As mentioned above, an operator length of 4 is used during the thesis, resulting in $h = \frac{7}{6}$ (Köhn, 2011).

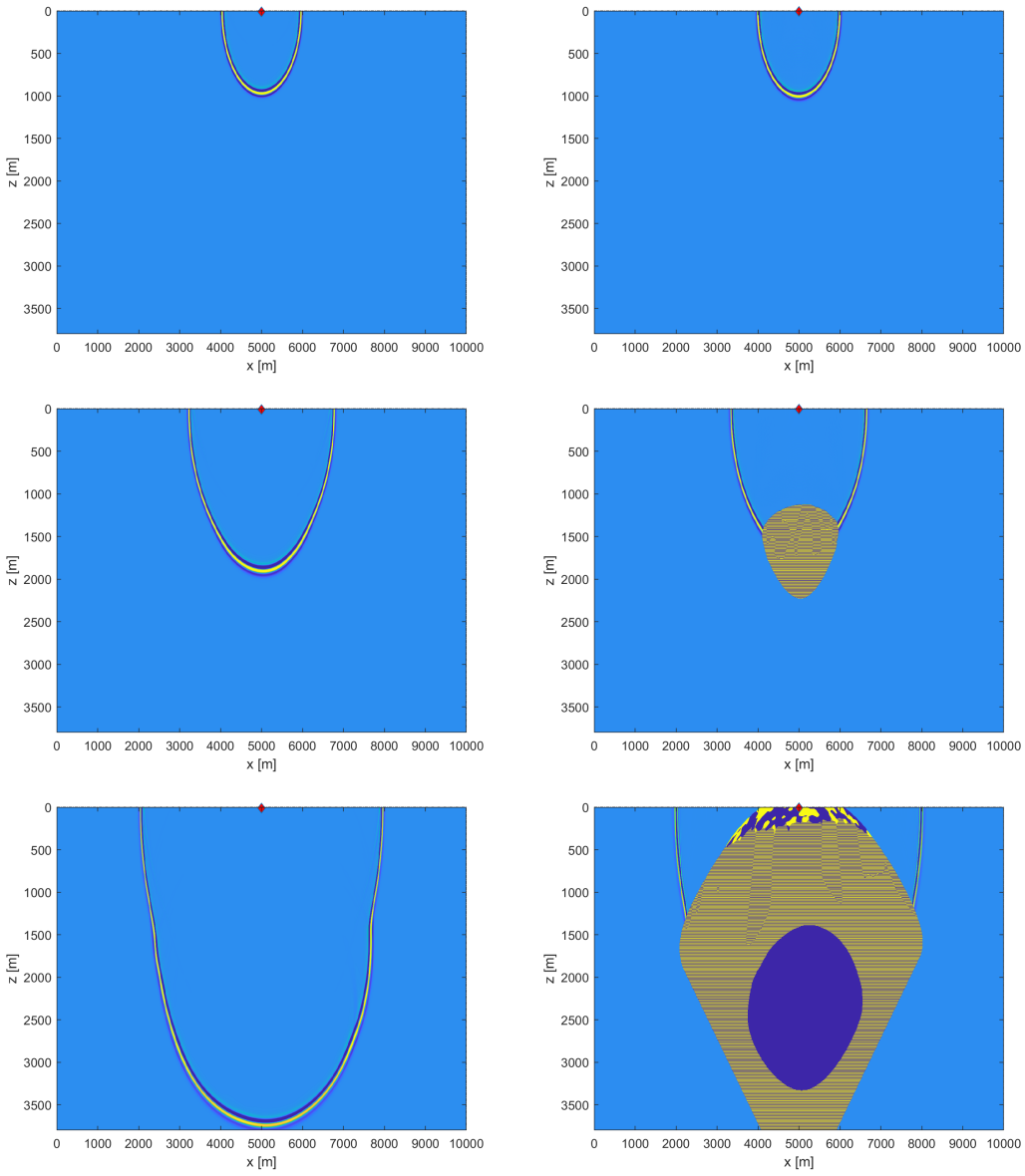


Figure 2-5: Snapshots of an FD simulation demonstrating a stable simulation and an unstable simulation. The left column presents samples where the CFL criterion is met, and the right column displays samples where the Δt does not meet the CFL criterion.

2-1-5 Parallelisation

When using FD for modelling wave propagation in complex media the demand for computational resources can be very high. Even with the steady increase in computational power, these resource demands often cannot be delivered by a single workstation. A strength of the DENISE modelling code, however, is its ability to use a computer cluster or multiple workstations to run the FD code in parallel by making use of a message passing interface (MPI). The ability to run the FD simulation in parallel vastly increases the capacity of the code to model complex media in detail. Parallelisation for FD simulations of wave propagation can be subdivided into two categories: shot parallelisation and domain decomposition. Shot parallelisation distributes each individual shot over all available CPU's, while domain decomposition partitions the domain into equally sized pieces and distributes those pieces over the available processing elements (PE). These PEs are generally a set of CPU nodes and have to interact with each other during the simulation. For relatively simple problems, shot parallelisation should result into better computing performance. However, in complex or large media the use of domain decomposition is superior (Bohlen, 2002). The domain decomposition strategy used in parallelisation of the DENISE code is designed by Bohlen (2002), where the PEs update the wavefield by computing Eqs. 2-17 - 2-21 individually, while passing along wavefield information at the edges of the domains with one another via the MPI. The padding layers are added specifically for communication among the PEs and always contain the most recent wavefield information from the neighbouring PEs. A schematic overview of the domain decomposition used can be seen in Fig. 2-6.

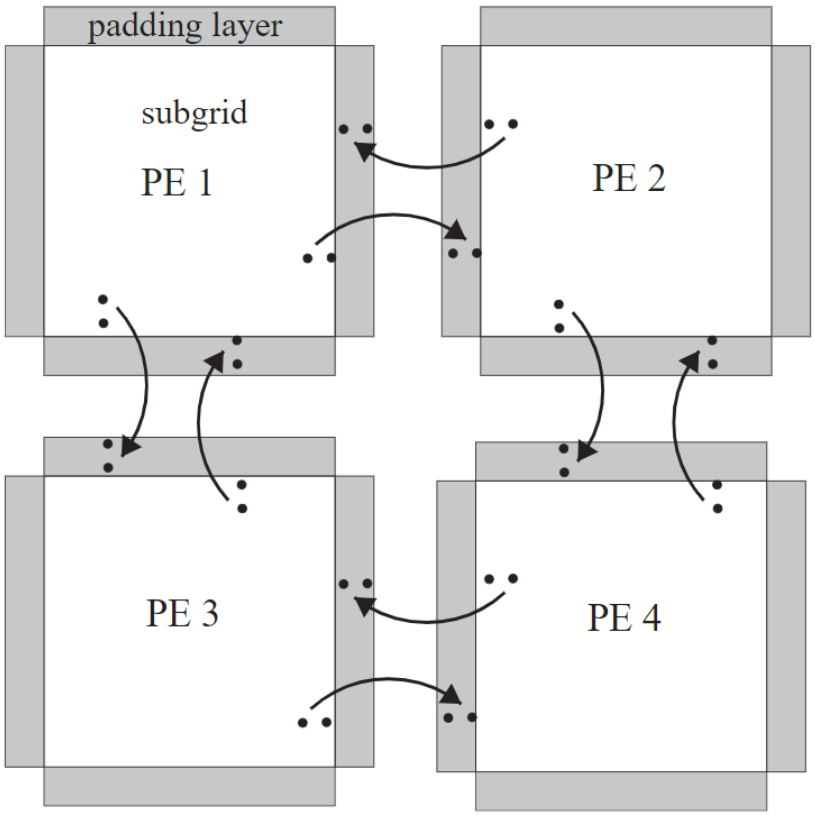


Figure 2-6: The domain decomposition strategy used by the DENISE modelling code, originally designed by Bohlen (2002).

2-2 Full-waveform inversion

The aim of FWI is to find the "best" possible model parameters from the full-waveform contained in the observed seismograms. To achieve this, the DENISE modelling code is used to simulate elastic wave propagation (Eqs. 2-17 - 2-21) through a proposed model to generate synthetic seismograms which are then compared to the observed seismograms by using a predefined misfit function. Model updates are proposed by minimising the misfit function, which is achieved by taking the gradient of the misfit function computed by the adjoint state method (Tarantola, 1984). Once the outcome of the misfit function is minimised, the model that best explains the observed data is found.

In the following subsections, definitions of the two misfit functions that are to be compared in this thesis are explained and finally the model-updating process is explained.

2-2-1 Misfit functions

The misfit function is used as a measure to describe the fit between observed and synthetic data. In this thesis, two misfit functions are compared: a classical approach to FWI by using a least-squares norm and instantaneous-phase coherency defined by Liu et al. (2022).

Least-squares norm

A classical misfit function used in FWI is the least-squares formulation of the misfit function (Tarantola, 1984). The misfit function is described by the least-squares norm of the difference between the measured data and the synthetic data summed over all receiver and source locations

$$J_1(\mathbf{m}) = \sum_{s,r} \int_0^T \left(\mathbf{u}^{obs}(x_s, x_r, t) - \mathbf{u}^{syn}(x_s, x_r, t, \mathbf{m}) \right)^2 dt, \tag{2-36}$$

where $\mathbf{u}^{obs}(x_s, x_r, t)$ is the observed data, $\mathbf{u}^{syn}(x_s, x_r, t, \mathbf{m})$ the synthetic data generated by forward-modelling of the wavefield and T is the total elapsed recording time (Tarantola, 1984; Liu et al., 2022). For brevity, from now on the dependency on the source and receiver location, respectively x_s and x_r, will be omitted from the expressions of the observed and synthetic data. Also the dependency on the model parameters $\mathbf{m}$ will be omitted from the expression of the synthetic data.

As mentioned above, to update the model, the misfit function Eq. 2-36 is minimised with the use of the gradient of the misfit function with respect to the model parameters $\mathbf{m}$ defined by

$$\frac{\partial J_1}{\partial \mathbf{m}} = \sum_{s,r} \int_0^T \left(\mathbf{u}^{syn}(t) - \mathbf{u}^{obs}(t) \right) \delta \mathbf{u}^{syn}(t)(t)\, dt = \sum_{s,r} \int_0^T \mathbf{r}(t)\, \delta \mathbf{u}^{syn}(t)\, dt, \tag{2-37}$$

with the residual of the wavefield $\mathbf{u}^{syn}(t) - \mathbf{u}^{obs}(t) = \mathbf{r}(t)$ and $\delta \mathbf{u}^{syn}(t)$ being a perturbation of the synthetic wavefield due to a change in the model space.

Instantaneous-phase coherency

The misfit function defined by Liu et al. (2022) is based on the principle of instantaneous phase coherency. The instantaneous phase coherency as a measure of model fit makes use of the exponential phase of the measured and synthetic datasets. The exponential phase of a seismic trace can be implicitly defined by

$$e^{i\phi(t)} = \frac{\mathbf{S}(t)}{\mathbf{A}(t)} = \frac{\mathbf{s}(t) + i\mathcal{H}\{\mathbf{s}(t)\}}{\sqrt{\mathbf{s}^2(t) + \mathcal{H}^2\{\mathbf{s}(t)\}}}, \tag{2-38}$$

with complex trace $\mathbf{S}(t)$, the amplitude of the complex trace $\mathbf{A}(t)$, real part of the seismic trace $\mathbf{s}(t)$ and imaginary part of the complex trace $\mathcal{H}\{\mathbf{s}(t)\}$ (Schimmel and Paulssen, 1997; Schimmel et al., 2011). By using the implicit definition of the exponential phase (Eq. 2-38) a misfit function using the instantaneous-phase coherency is defined

$$\begin{aligned} J_2(\mathbf{m}) &= \frac{1}{4}\sum_{s,r}\int_0^T \left\{ \left| e^{i\phi_{obs}(t)} - e^{i\phi_{syn}(t)} \right|^2 - \left| e^{i\phi_{obs}(t)} + e^{i\phi_{syn}(t)} \right|^2 \right\} dt \\ &= \frac{1}{4}\sum_{s,r}\int_0^T \left| \frac{\mathbf{u}^{obs}(t) + i\mathcal{H}\{\mathbf{u}^{obs}(t)\}}{\sqrt{\mathbf{u}^{obs2}(t) + \mathcal{H}^2\{\mathbf{u}^{obs}(t)\}}} - \frac{\mathbf{u}^{syn}(t) + i\mathcal{H}\{\mathbf{u}^{syn}(t)\}}{\sqrt{\mathbf{u}^{syn2}(t) + \mathcal{H}^2\{\mathbf{u}^{syn}(t)\}}} \right|^2 dt \\ &\quad - \frac{1}{4}\sum_{s,r}\int_0^T \left| \frac{\mathbf{u}^{obs}(t) + i\mathcal{H}\{\mathbf{u}^{obs}(t)\}}{\sqrt{\mathbf{u}^{obs2}(t) + \mathcal{H}^2\{\mathbf{u}^{obs}(t)\}}} + \frac{\mathbf{u}^{syn}(t) + i\mathcal{H}\{\mathbf{u}^{syn}(t)\}}{\sqrt{\mathbf{u}^{syn2}(t) + \mathcal{H}^2\{\mathbf{u}^{syn}(t)\}}} \right|^2 dt, \end{aligned} \tag{2-39}$$

with $e^{i\phi_{obs}(t)}$ the exponential phase of the observed data, $e^{i\phi_{syn}(t)}$ the exponential phase of the synthetic data generated by forward-modelling (Liu et al., 2022).

Like the least-squares norm, the model space is updated by minimising Eq. 2-39. To do so, the gradient of the misfit function with respect to the model parameters $\mathbf{m}$ is required, and is defined

$$\begin{aligned} \frac{\partial J_2}{\partial \mathbf{m}} &= \sum_{s,r}\int_0^T \left[\frac{\mathbf{u}^{syn}(t)\,\mathcal{H}\{\mathbf{u}^{obs}(t)\}\,\mathcal{H}\{\mathbf{u}^{syn}(t)\}}{\mathbf{A}_{obs}(t)\,\mathbf{A}_{syn}{}^3(t)} - \frac{\mathbf{u}^{obs}(t)\,\mathcal{H}^2\{\mathbf{u}^{syn}(t)\}}{\mathbf{A}_{obs}(t)\,\mathbf{A}_{syn}{}^3(t)} \right] \delta\mathbf{s}(t)\,dt \\ &\quad + \sum_{s,r}\int_0^T \left[\mathcal{H}\left\{ \frac{\mathbf{u}^{syn2}(t)\,\mathcal{H}\{\mathbf{u}^{obs}(t)\}}{\mathbf{A}_{obs}(t)\,\mathbf{A}_{syn}{}^3(t)} - \frac{\mathbf{u}^{obs}(t)\,\mathbf{u}^{syn}(t)\,\mathcal{H}\{\mathbf{u}^{syn}(t)\}}{\mathbf{A}_{obs}(t)\,\mathbf{A}_{syn}{}^3(t)} \right\} \right] \delta\mathbf{s}(t)\,dt \\ &= \frac{1}{4}\sum_{s,r}\int_0^T \tilde{\mathbf{r}}(t)\,\delta\mathbf{s}^{syn}(t)\,dt, \end{aligned} \tag{2-40}$$

with residual wavefield defined as $\tilde{\mathbf{r}}(t)$, and the instantaneous amplitude of the observed and synthetic waveforms respectively $\mathbf{A}_{obs}$ and $\mathbf{A}_{syn}$, respectively (Liu et al., 2022).

2-2-2 Finding the optimal model

FWI aims to infer the best fit of physical model parameters generally contained in a vector $\mathbf{m}$ on the basis of observed data $\mathbf{u}^{obs}$. For the problem of P/SV-wave propagation $\mathbf{m}$ consist of parameters λ, μ and ρ. To find the best fit of model parameters, the model space composed of all possible combinations of model parameters is searched by the FWI algorithm by minimising one of the previously defined misfit functions Eqs. 2-36 & 2-39.

Fig. 2-7, displays an arbitrary example of a model space composed of all possible combinations of two arbitrary model parameters m_x and m_y. Assuming that the model parameters influence the amount of residual-energy between observed and synthetic data, the optimal combination of model parameters can be found by searching for the lowest residual-energy, i.e., the minimum of the misfit function. Starting at the initial model $\mathbf{m}_1$ a search for the optimal model is performed in an iterative manner by taking a step in the model space

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \alpha_i \delta \mathbf{m}_i, \tag{2-41}$$

where $\delta \mathbf{m}_i$ is the search direction towards the minimum residual-energy and α_i the length of the step taken for the current iteration. This process is repeated until the lowest residual-energy value is attained.

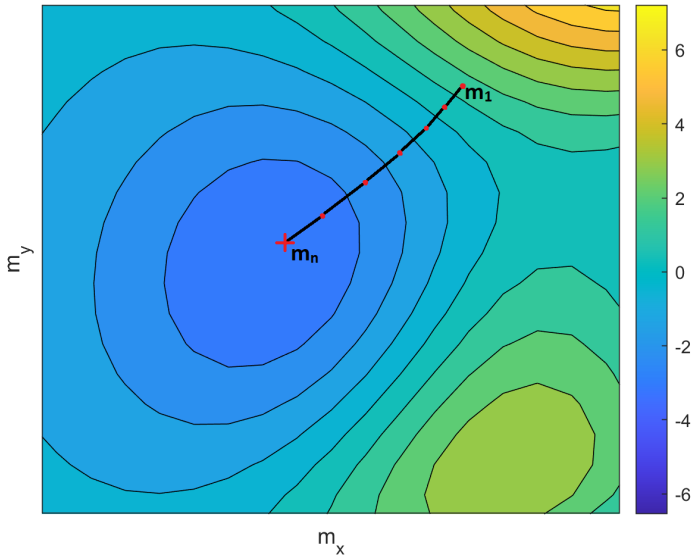


Figure 2-7: Schematic overview of how the model space is navigated from starting model $\mathbf{m}_1$ to the optimal model $\mathbf{m}_n$. Here, the model space is navigated by using the residual-energy as a function of two arbitrary model parameters m_x and m_y. The colours denote the residual-energy values as a function of the model parameters, with the red cross denoting the minimum residual-energy corresponding to the optimal combination of model parameters.

Determining the search direction

The search for the optimal model starts at an arbitrary point in the model space defined by a starting model $\mathbf{m}_1$. To find the search direction for the first iteration, one of the previously defined misfit functions is extended by means of a TSE near the initial model

$$J\left(\mathbf{m}_1 + \delta\mathbf{m}_1\right) \approx J\left(\mathbf{m}_1\right) + \delta\mathbf{m}_1 \left(\frac{\partial J}{\partial \mathbf{m}}\right)_1 + \frac{1}{2}\delta\mathbf{m}_1 \left(\frac{\partial^2 J}{\partial \mathbf{m}^2}\right)_1 \delta\mathbf{m}_1^{\mathbf{T}}, \tag{2-42}$$

with $\delta\mathbf{m}_1$ the search direction from the initial model $\mathbf{m}_1$ towards the new model (Köhn, 2011). Next, the derivative of Eq. 2-42 is taken with respect to the search direction and set to zero

$$\frac{\partial J\left(\mathbf{m}_1 + \delta\mathbf{m}_1\right)}{\partial \delta\mathbf{m}_1} = \left(\frac{\partial J}{\partial \mathbf{m}}\right)_1 + \delta\mathbf{m}_1 \left(\frac{\partial^2 J}{\partial \mathbf{m}^2}\right)_1 = \left(\frac{\partial J}{\partial \mathbf{m}}\right)_1 + \delta\mathbf{m}_1\mathbf{H}_1 = 0, \tag{2-43}$$

With $\mathbf{H}_1$ being the Hessian of the objective function. After reorganising Eq. 2-43, the expression for the optimal search direction is found

$$\delta\mathbf{m}_1 = -\mathbf{H}_1^{-1} \left(\frac{\partial J}{\partial \mathbf{m}}\right)_1, \tag{2-44}$$

with $\mathbf{H}_1^{-1}$ denoting the inverse Hessian of the misfit function. Furthermore, the gradient of the misfit function $\left(\frac{\partial J}{\partial \mathbf{m}}\right)_1$ denotes the direction of the steepest descent towards the minimum of the misfit function (Köhn, 2011). By generalising Eq. 2-44

$$\delta\mathbf{m}_i = -\mathbf{H}_i^{-1} \left(\frac{\partial J}{\partial \mathbf{m}}\right)_i, \tag{2-45}$$

a definition of the optimum search direction can be defined. However, Eq. 2-45 requires a solution to the inverse Hessian matrix, which is often computationally very expensive or cannot be solved at all (Nocedal and Wright, 2006; Virieux and Operto, 2009). The DENISE FWI algorithm avoids computation of the inverse Hessian by using a limited memory Quasi-Newton method called limited-memory Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS) (Köhn, 2011). The general concept of L-BFGS is to update an approximation of the inverse Hessian at each iteration using only a few gradients of the most recent iterations (Nocedal and Wright, 2006; Virieux and Operto, 2009). However, an initial guess of the Hessian has to be made, generally the inverse of the diagonal of the Hessian at step one can be used (Brossier, 2009). Replacing the inverse Hessian by its estimate a new definition of the optimal search direction is obtained

$$\delta\mathbf{m}_i = -\mathbf{B}_i^{-1} \left(\frac{\partial J}{\partial \mathbf{m}}\right)_i, \tag{2-46}$$

where $\mathbf{B}_i$ is the estimate of the inverse Hessian.

Computation of the gradient

As mentioned above, the gradient of the misfit function $\frac{\partial J}{\partial \mathbf{m}}$ with respect to the model parameters is used to determine the direction of steepest descent. To obtain a solution for Eq. 2-37 or 2-40, requires a solution to the Frechët derivatives which are the derivatives of the state variables with respect to the model parameters $\frac{\partial \mathbf{u}^{obs}}{\partial \mathbf{m}}$ and $\frac{\partial \mathbf{u}^{syn}}{\partial \mathbf{m}}$. Solving the Frechët derivatives can be expensive to compute; therefore, the DENISE code uses the adjoint method instead to solve the gradient of the misfit function (Plessix, 2006; Köhn, 2011).

To explain the adjoint method used for computing the gradient of the misfit functions, it is applied to the system of PDEs containing the formulation of the elastic equation of motion Eq. 2-1 and the stress-strain relationship Eqs. 2-2 & 2-3. However, now the stress tensor is defined by

$$\sigma_{ij} = T_{ij} - c_{ijkl}, \tag{2-47}$$

where T_{ij} denote the surface forces and c_{ijkl} the stiffness tensor of the medium (Köhn, 2011). Next, small perturbations are introduced into the equation of motion resulting in a new system of PDEs describing the perturbed equation of motion

$$\rho \frac{\partial^2 \delta u_i}{\partial t^2} = \frac{\partial \delta \sigma_{ij}}{\partial x_j} + \Delta f_i, \tag{2-48}$$

$$\delta \sigma_{ij} = \Delta T_{ij} - c_{ijkl} \delta \epsilon_{kl}, \tag{2-49}$$

$$\delta \epsilon_{ij} = \frac{1}{2} \left(\frac{\partial \delta u_i}{\partial x_j} + \frac{\partial \delta u_j}{\partial x_i} \right), \tag{2-50}$$

$$\Delta f_i = -\delta \rho \frac{\partial^2 u_i}{\partial t^2}, \tag{2-51}$$

$$\Delta T_{ij} = \delta c_{ijkl} \epsilon_{kl} = \delta_{jk} \delta_{lm} \delta \lambda + \left(\delta_{jl} \delta_{km} + \delta_{jm}, \delta_{kl} \right) \delta \mu, \tag{2-52}$$

where δ denoting the perturbed versions of the parameter or variable, Δ denoting the perturbed sources and the perturbed wavefield is propagating in the unperturbed medium (Köhn, 2011). Assuming we can obtain a solution to the wave equation by using Green's function $G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right)$ the following expression is obtained

$$\begin{aligned} \delta u_i(\mathbf{x}, t) = &- \int_V \left[\int_0^T G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right) \frac{\partial^2 u_j}{\partial t^2}\left(\mathbf{x}', t'\right) \, dt' \right] \delta \rho dV \\ &- \int_V \left[\int_0^T \frac{\partial G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right)}{\partial x_k'} \epsilon_{lm}\left(\mathbf{x}', t'\right) \delta_{jk} \delta_{lm} \, dt' \right] \delta \lambda \, dV \\ &- \int_V \left[\int_0^T \frac{\partial G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right)}{\partial x_k'} \epsilon_{lm}\left(\mathbf{x}', t'\right) \left(\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl} \right) \, dt' \right] \delta \mu \, dV, \end{aligned} \tag{2-53}$$

with $\delta u_i(\mathbf{x}, t)$ denoting the perturbed wavefield defined in terms of the Green's function and is a linearization of the forward problem known as the Born approximation (Köhn, 2011; Virieux and Operto, 2009). Note that, Eq. 2-53 can be separated into the Frechët derivatives used for the individual material parameters used in P/SV parameters:

$$\frac{\partial u_i}{\partial \rho} = -\int_0^T G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right) \frac{\partial^2 u_j}{\partial t^2}\left(\mathbf{x}', t'\right)\, dt', \tag{2-54}$$

$$\frac{\partial u_i}{\partial \lambda} = -\int_0^T \frac{\partial G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right)}{\partial x_k'} \epsilon_{lm}\left(\mathbf{x}', t'\right) \delta_{jk}\delta_{lm}\, dt', \tag{2-55}$$

$$\frac{\partial u_i}{\partial \mu} = -\int_0^T \frac{\partial G_{ij}\left(\mathbf{x}, t; \mathbf{x}', t'\right)}{\partial x_k'} \epsilon_{lm}\left(\mathbf{x}', t'\right) \left(\delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl}\right)\, dt'. \tag{2-56}$$

By making use of the adjoint method, the adjoint operator can be defined that maps a small change in the model space as a function of a perturbation of the wavefield in the data space

$$\delta \mathbf{m}'(\mathbf{x}) = \sum_s \int_0^T \sum_{r=1}^{N_{rec}} \delta u_i'\left(\mathbf{x}_r, t'\right) dt, \tag{2-57}$$

with $\mathbf{m}'$ denoting a small change in the model space caused by $\delta u_i'$ which is the perturbed wavefield, $\sum_s$ the summation over all source locations, $\sum_{r=1}^{N_{rec}}$ the summation over all receivers (Köhn, 2011; Virieux and Operto, 2009; Tarantola, 2005).

By making use of the adjoint operator Eq. 2-57 a wavefield is defined

$$\Psi(\mathbf{x}, t) = \sum_{r=1}^{N_{rec}} \int_0^T G_{ij}\left(\mathbf{x}_r, t'; \mathbf{x}, t\right) \delta u_i'\left(\mathbf{x}_r, t'\right) dt', \tag{2-58}$$

with Ψ denoting the wavefield that is generated as a function of the residual data δu_i that propagates from the receivers back in time (Tarantola, 1984; Virieux and Operto, 2009; Köhn, 2011). For the two misfit functions defined previously, wavefield Ψ denotes one of the residual wavefields defined in Eqs. 2-37 & 2-40.

Furthermore, with the help of Eq. 2-57 and the definition of the strain tensor Eq. 2-3 the gradients of the individual medium parameters can be defined for the P/SV case in isotropic elastic media

$$\delta \lambda' = -\sum_s \int \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y}\right) + \left(\frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y}\right) dt, \tag{2-59}$$

$$\delta \mu' = -\sum_s \int \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y}\right) + \left(\frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_y}{\partial y}\right) + 2\left(\frac{\partial u_x}{\partial x}\frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y}\frac{\partial \Psi_y}{\partial y}\right) dt, \tag{2-60}$$

$$\delta \rho' = -\sum_s \int \left(\frac{\partial u_x}{\partial x}\frac{\partial \Psi_x}{\partial x} + \frac{\partial u_y}{\partial y}\frac{\partial \Psi_y}{\partial y}\right) dt, \tag{2-61}$$

with Ψ_x and Ψ_y the components of the back propagated wavefield (either $\mathbf{r}$ or $\tilde{\mathbf{r}}$) and the derivatives of the wavefield with respect to the model parameters $\delta\lambda'$, $\delta\mu'$ and $\delta\rho'$ using the

displacement formulation of the equation of motion (Köhn, 2011). However, the DENISE FD modelling code used a formulation of the equation of motion based on particle velocity and the stress tensor rather than displacement (Eq. 2-4). This means that the displacements in Eqs. 2-59 - 2-61 have to be replaced with particle velocities $\mathbf{v}$ and stress tensor of the forward wavefield σ_{ij} and stress tensor of the back propagated wavefield Σ_{ij} resulting in

$$\delta\lambda' = -\sum_s \int \frac{(\sigma_{xx}+\sigma_{yy})(\Sigma_{xx}+\Sigma_{yy})}{4(\lambda+\mu)^2} dt, \tag{2-62}$$

$$\delta\mu' = -\sum_s \int \frac{\sigma_{xy}\Sigma_{xy}}{\mu^2} + \frac{1}{4}\left(\frac{(\sigma_{xx}+\sigma_{yy})(\Sigma_{xx}+\Sigma_{yy})}{(\lambda+\mu)^2} + \frac{(\sigma_{xx}-\sigma_{yy})(\Sigma_{xx}-\Sigma_{yy})}{\mu^2}\right) dt, \tag{2-63}$$

$$\delta\rho' = -\sum_s \int \frac{\partial v_x}{\partial t}\Psi_x + \frac{\partial v_y}{\partial y}\Psi_y dt, \tag{2-64}$$

where $\delta\lambda'$, $\delta\mu'$, and $\delta\rho'$ denote the derivatives of the particle velocities used for determining the gradient of the misfit functions (Groos, 2013; Köhn, 2011).

Step length

Finally, the updating process requires a definition of the step length α_i that is taken in each iteration when updating the model in search of the best model parameter fit (Eq. 2-41). To find the optimum step length, the DENISE inversion code attempts a parabolic line-search method (Nocedal and Wright, 2006; Köhn, 2011). This is achieved by testing a step length for each of the three model parameters separately

$$\alpha_\lambda = \epsilon \frac{max(\lambda_i)}{max(\delta\lambda_i)}, \tag{2-65}$$

$$\alpha_\mu = \epsilon \frac{max(\mu_i)}{max(\delta\mu_i)}, \tag{2-66}$$

$$\alpha_\rho = \epsilon \frac{max(\rho_i)}{max(\delta\rho_i)}, \tag{2-67}$$

with scaling parameter ϵ, the maximum of the model parameters $max(\lambda)$, $max(\mu)$, and $max(\rho)$, and the maximum of the gradients with respect to the model parameters $max(\delta\lambda_i)$, $max(\delta\lambda_i)$, $max(\delta\rho_i)$ (Groos, 2013; Köhn, 2011). After having defined three step lengths, estimates of their corresponding L_2 norms are made by fitting a parabola to these data points is defined

$$L_{2,i} = a\alpha_i^2 + b\alpha_i + c, \tag{2-68}$$

with unknowns a, b, and c where the shape of the parabola is used to determined the optimum step length in the DENISE FWI code (Köhn, 2011).

Chapter 3

Problem setup

As mentioned in chapter 1 the objective of the thesis is to investigate a new type of misfit function proposed by Liu et al. (2022). First, the performance of the instantaneous-phase coherency misfit (Eq. 2-39) is compared to a classical least-squares norm (Eq. 2-36). Next, the results of a geophysical and geo-technical survey performed by Fugro will be used. The results of the survey will be referred to as 3D dataset was used. Unfortunately the seismic data was not available at the time of this research, therefore the inversion result was used instead. The geo-technical data is used to create a test case for the accuracy of the instantaneous-phase coherency. Whereas the inversion result was used as a complex subsurface model to investigate the robustness of the instantaneous-phase when used for a complex subsurface. Since surface-wave propagation is dominated by shear-wave properties (Socco and Strobbia, 2004; Socco et al., 2010; Groos et al., 2017), only the V_s model is updated during the inversion, whereas the P-wave velocity (V_p) and ρ models are not updated. During all simulations a band-limited spike ($10 - 60$ Hz) was used as a source wavelet.

Below the models and experimental setup for each objective are presented.

3-1 Comparison of misfit functions

The instantaneous-phase coherency should be more suitable for near-surface purposes as it delivers a misfit measure that is independent of the amplitude (Liu et al., 2022). To verify the robustness of the method with respect to erroneous amplitude variations a synthetic test case is designed, where the observed seismograms were contaminated with random noise. Finally, two datasets from a near-surface geophysical survey are used to compare the performance of the instantaneous-phase coherency with the least-squares norm.

3-1-1 Synthetic test case

To test the robustness of Eq. 2-39 and Eq. 2-36 to amplitude variations, a series of inversions where performed on a 2D V_s model where noise was added to the data. The V_s model, shown

in Fig. 3-1a, has a background velocity of 300 m/s and contains two block shaped velocity anomalies of 270 m/s and 330 m/s. The initial model used for the inversion has a constant velocity of 300 m/s (Fig. 3-1b), the V_p is set to 1000 m/s and the density to 2000 kg/m^3. For simulation of the wavefield a receiver array is used with a 1 m spacing starting at $x = 5$ m and a total of 41 geophones are used. As source a band-limited spike is used with frequency band $10 - 60$ Hz, a total of 16 shots where performed and the source location is moved along the receiver array, starting at $x = 10$ m and a spacing between each shot location of 2 m.

After modelling the wavefield through the input model, a set of synthetic seismograms are obtained, which are regarded as the observed data. The observed data is then used to create four test cases with a varying SNR. The first test uses the uncontaminated observed data, while the three other cases are contaminated with noise. To generate the noise contaminated data a band-pass filtered $(10 - 60$ Hz) Gaussian noise is added to obtain datasets with a SNR of 5, 10 and 20. Fig. 3-2 displays four common shot gathers of the four test cases.

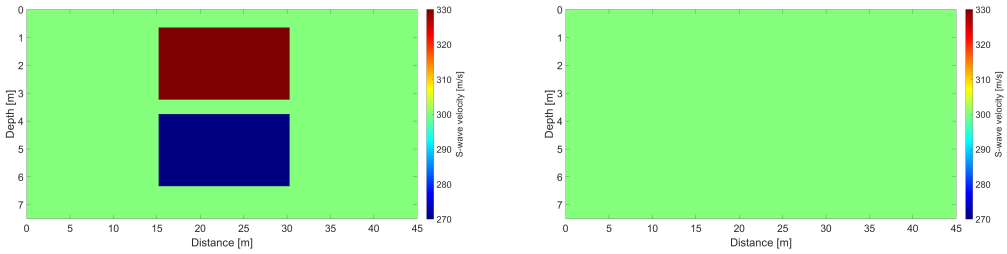


Figure 3-1: The synthetic V_s models used for testing the robustness to noise of the misfit functions. (a) the true model and (b) the initial model used in the inversion.

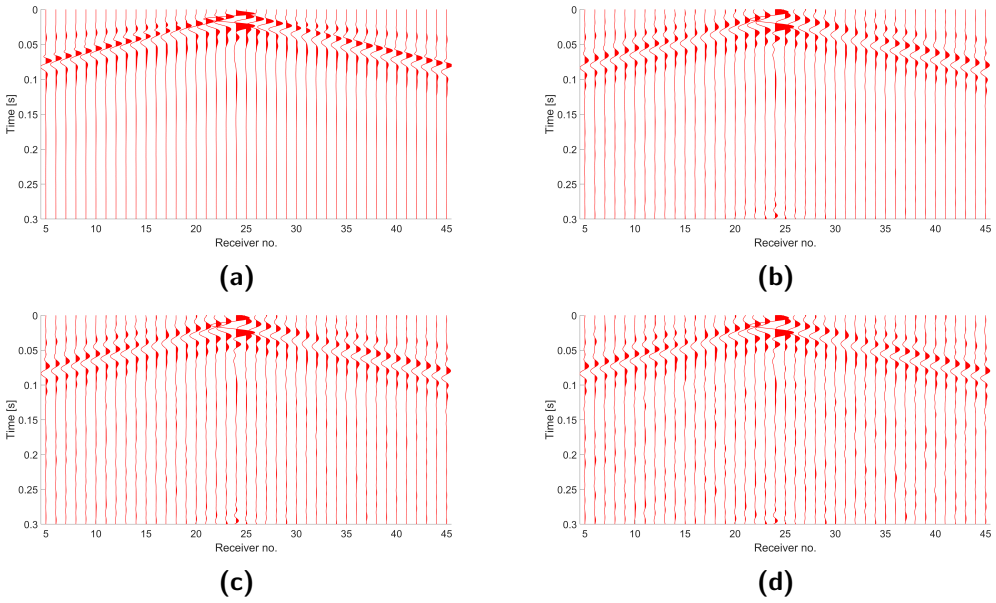


Figure 3-2: Common shot gathers obtained by forward simulation of the wavefield with increasing SNR of the 8^{th} shot. (a) without noise, (b) a SNR of 20, (c) a SNR of 10 and (d) a SNR of 5.

3-1-2 Field test 1

The first field dataset used was obtained during a seismic survey of an archaeological site near Ostia, Italy by Ghose et al. (2020). The survey line was partly positioned above the location of a known tumulus, which is buried at a depth of $0.5-2$ m, and located approximately at 20 m along the survey line (Ghose et al., 2020). The dataset used was pre-processed by Liu et al. (2022) and is dominated by Rayleigh waves, therefore the P/SV modelling code was used to simulate wave propagation. Fig. 3-3a displays the initial V_s model used for the inversion, with a gradual velocity increase from 140 m/s to 188.75 m/s. Similarly, the initial V_p model is set to increase from 261.79 m/s to 352.96 m/s, the density model is set to 2000 kg/m^3.

The source and receiver geometry used during forward-modelling of the wavefield was a roll-along type survey design. Here, two spreads where used to perform the roll-along. For both spreads a total of 120 geophones was used, with a receiver spacing of 0.25 m and a source interval of 1 m. The first spread was used for 9 shots, with the first shot taking place at 2 m and the receiver line started at 2 m. The second spread was used for the last 28 shots, with the location of the first shot at 11.5 m and the first receiver was located at 8 m. The inversion was performed sequentially, starting at low bandwidths and gradually increased: $0-10$ Hz, $0-20$ Hz, $0-30$ Hz, $0-40$ Hz, $0-50$ Hz and $0-60$ Hz.

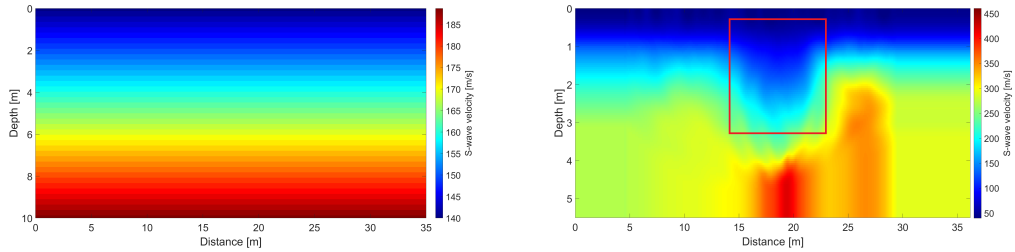
3-1-3 Field test 2

The second field data set is an open source dataset made available by Köhn et al. (2019) and was obtained during a seismological survey of the Fossa Carolina, Germany, which is a medieval canal structure. The survey was performed specifically to obtain SH/Love-wave data, meaning the SH modelling code was used to perform forward-modelling. The initial V_s model used for the inversion of the field data was obtained by performing first arrival traveltime tomography (FATT) (Fig 3-3b), which is a quick inversion method that uses the first arrival travel times to construct a velocity model of the subsurface, the density was set to 2000 km/m^3. The location of the channel structure is denoted by the red rectangle in Fig. 3-3b.

The forward simulation used a receiver array with 0.75 m spacing, with the first receiver being located at 0.8 m and a total of 48 receivers. The source was moved along the line, starting at 1.17 m with increments of 0.75 m for a total of 47 shots. The inversion was performed sequentially, starting at low bandwidths and gradually increased: $0-20$ Hz, $0-30$ Hz, $0-40$ Hz, $0-50$ Hz and $0-60$ Hz.

3-2 3D dataset

The 3D dataset displayed in Fig. 3-4 was the result of a large scale geo-technical and geophysical survey in the Middle East. The geo-technical data was used to construct a test case for testing the accuracy of the inversion code and the instantaneous-phase coherency. Whereas, the inversion result from the geophysical survey was used to generate a model to investigate the capability of the inversion when exposed to a complex subsurface. Both cases used the P/SV FD code as the survey design focused on the vertical components of the wavefield.



(a) **(b)**

Figure 3-3: The initial V_s models used during the inversion of the field datasets. (a) Initial model used for the inversion of the Ostia dataset and (b) the initial model used for the Fossa Carolina dataset (Köhn et al., 2019), the red square denotes the location of the river channel.

3-2-1 Accuracy of the instantaneous phase coherency

To investigate the accuracy of the inversion when using the instantaneous-phase coherency one of the PS-logs denoted by the coloured dots in Fig. 3-4 from the geo-technical survey was used. From Fig. 3-4, it can be observed that the result of the PS-log, often assumed as the ground truth, does not always coincide with the inversion result. Therefore the following question arose, can an initial model based on the ground truth from a nearby geo-technical survey be used to accurately resolve the ground truth at the current survey location?

To investigate the accuracy of the inversion, layer cake models were designed starting with the initial model. For the creation of the initial model, the low velocity layer from the inversion results was added for the first 2 m, next 7 data points from the PS-log were added and finally the model was smoothed (Fig. 3-5a). Next, a total of 20 models where created that were regarded as the ground truth at the location of the simulation. The true models used for the first 20 inversions where all generated by randomly varying the depth of the layers from the initial model within a range of $0-10$ m. The sequence in which the layers occurred did remain fixed. Afterwards, another series of 20 inversions were performed where the low velocity layer remained fixed between 0 m and 2 m. Finally, 10 models from the previous inversions where used again to experiment a different sequence of bandwidths for the inversion. Fig. 3-5b displays an representative example of one of the true models used during the first 20 inversions and Fig. 3-5c one of the models used for the last 30 inversions. The remaining models can be found in appendix A, B and C for respectively the first, second and third set of inversions.

For modelling the wavefield a receiver array is used with 1 m spacing, with the first receiver being located at 9 m and a total of 221 receivers. The source was moved along the line, starting at 27 m with increments of 5 m for a total of 37 shots. All 50 inversions where performed sequentially, starting at low bandwidths which where gradually increased. For the first 40 inversions the following bandwidths where used: $10-20$ Hz, $10-30$ Hz, $10-40$ Hz, $10-50$ Hz and $10-60$ Hz. The last 10 inversions used a sequence of bandwidths from: $0-10$ Hz, $0-20$ Hz, $0-30$ Hz, $0-40$ Hz, $0-50$ Hz and $0-60$ Hz.

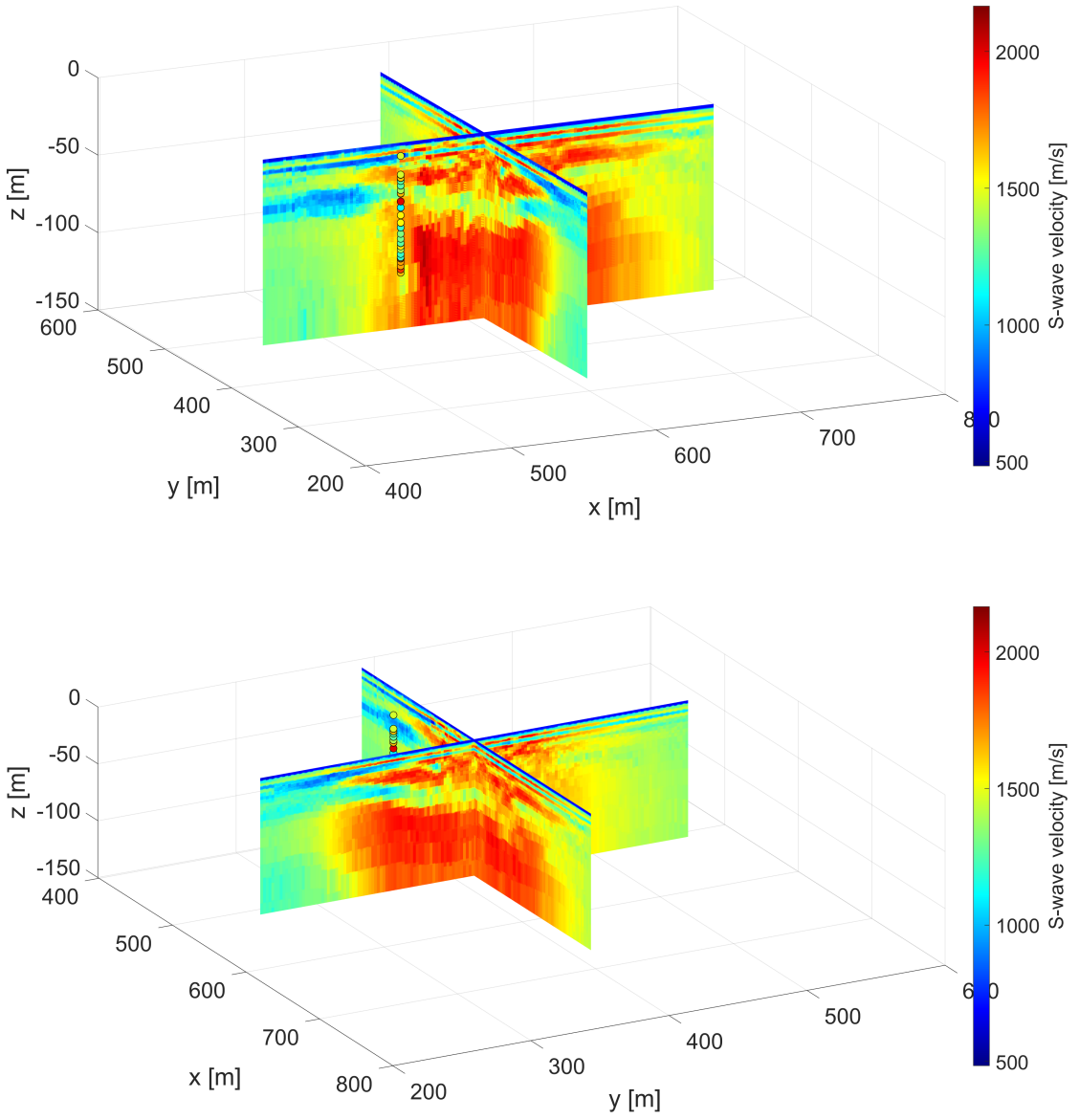


Figure 3-4: The 3D V_s model from which the complex subsurface model is build. The location of the PS-log used for constructing the accuracy tests is denoted by the coloured dots.

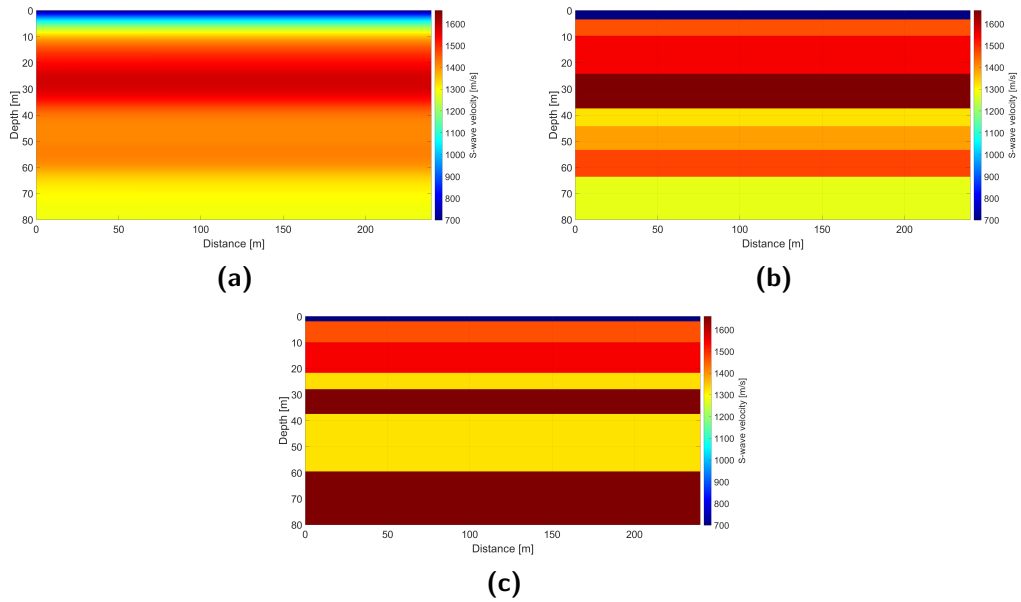


Figure 3-5: The models used during the accuracy experiments. (a) the initial model used for all test runs, (b) one of the true models used for the first 20 test runs and (c) one of the true models used for the remainder of the test inversions.

3-2-2 Complex subsurface model

To investigate whether the inversion using the instantaneous-phase was robust when the subsurface is complex, two inversions where performed. To imitate a complex subsurface a slice from the 3D dataset was taken to create the true model. The slice was adjusted slightly as the depth was halved and the edges where trimmed. The resulting model is shown in Fig. 3-6a. As some model features have a size of only a couple meters, this model creates an interesting test case for geophysical applications on a civil engineering scale.

Next, two initial models where designed. In general initial models used in FWI try to incorporate knowledge of the subsurface as local optimisation techniques depend heavily on a well defined starting model (Virieux and Operto, 2009). Therefore, two differently designed initial models where tested.

The first model was created as if there was PS-log information available in the middle of the survey line. Next, this information was used to create a smooth layer cake model which is shown in Fig. 3-6b. The second model was created as if there was already some estimate of the subsurface. Normally this is achieved by using another less cost heavy method first such as MASW or like the second field dataset FATT. However, here, a model was made by smoothing the true model so only some of the general subsurface features where maintained.

For modelling the wavefield, a receiver array is used with 2 m spacing, with the first receiver located at 10 m and a total of 75 receivers. The source was moved along the line, starting at 16 m with increments of 6 m for a total of 25 shots. The inversion was performed sequentially, gradually increasing the bandwidth starting at low frequencies: $10-20$ Hz, $10-30$ Hz, $10-40$ Hz, $10-50$ Hz and $10-60$ Hz.

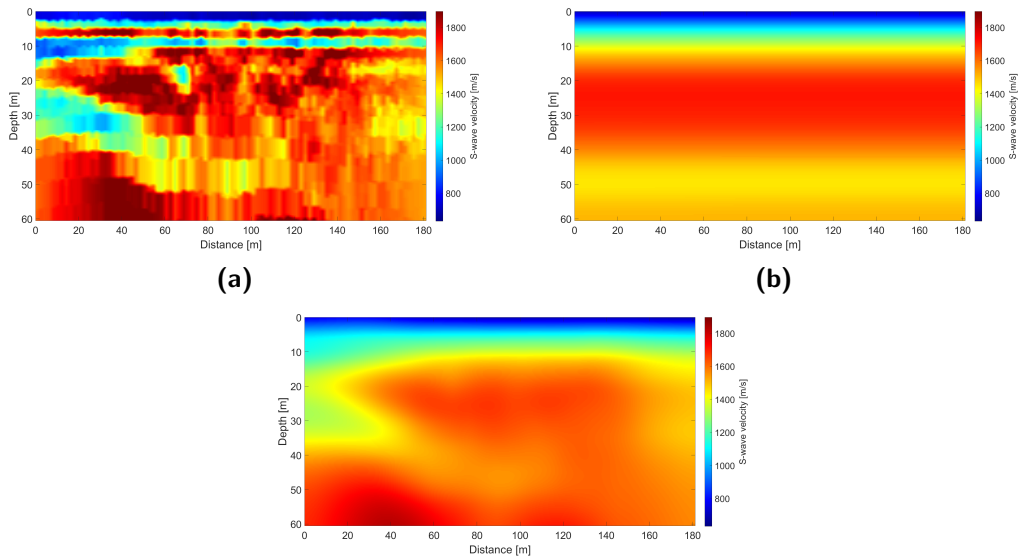


Figure 3-6: The V_s models used for the complex subsurface model inversions. (a) The slice of the 3D dataset used as true model, (b) the initial model as if a PS-log was taken and (c) the initial model based on a smoothed version of the true model.

Chapter 4

Results and discussion

This chapter will discuss the results of the inversions described in chapter 3. To properly evaluate the quality of the instantaneous-phase coherency, first the inversion results comparing it to a least-squares misfit will be discussed. Next, the accuracy of the inversion method is discussed and finally its robustness when used on a complex subsurface model.

4-1 Comparison of misfit functions

4-1-1 Synthetic test cases

To evaluate the influence of noise on the misfit functions, first the inversions where performed without introducing any noise to the data. Fig. 4-1, displays the results of the least-squares norm (Fig. 4-1a) and the instantaneous-phase coherency (Fig. 4-1b). From Fig. 4-1 it can be observed that both misfit functions perform well with the absence of noise and are capable of reconstructing the velocity anomalies contained in the true model (Fig. 3-1a).

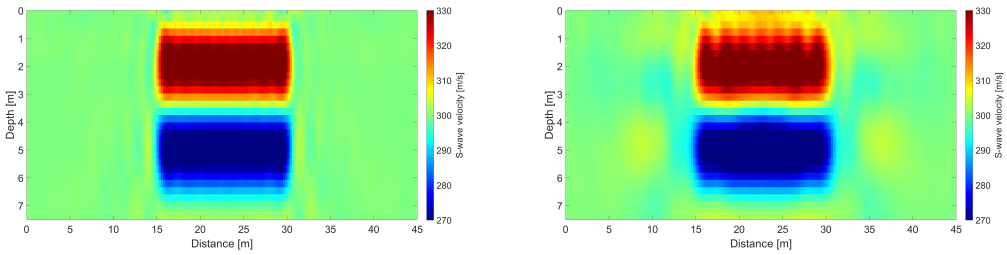


Figure 4-1: The inverted V_s models obtained from inversion of the noise free dataset. (a) The FWI result using the the least-squares norm and (b) the result when using the instantaneous-phase instead.

Next, the inversion results of the noise contaminated datasets are shown (Fig. 4-2). From Fig. 4-2, it can be observed that both inversions are able to reconstruct the velocity anomalies

contained in the true model. However, comparison of the V_s inversion results of the dataset with a SNR of 20 displays that the least-squares norm result (Fig. 4-2a) contains more artefacts than the instantaneous-phase coherency result (Fig. 4-2b). Demonstrating that the least-squares norm FWI is less robust to amplitude variations in the data content compared to the instantaneous-phase coherency FWI.

When inspecting the V_s inversion results with lower SNR, it is observed that for the least-squares norm FWI, artefacts become more prevalent when the SNR is decreased (Figs. 4-2c and 4-2e). Whereas, the results of the instantaneous-phase coherency FWI (Fig. 4-2d and 4-2f) demonstrate that the results of the inversion does not deteriorate when the SNR is decreased.

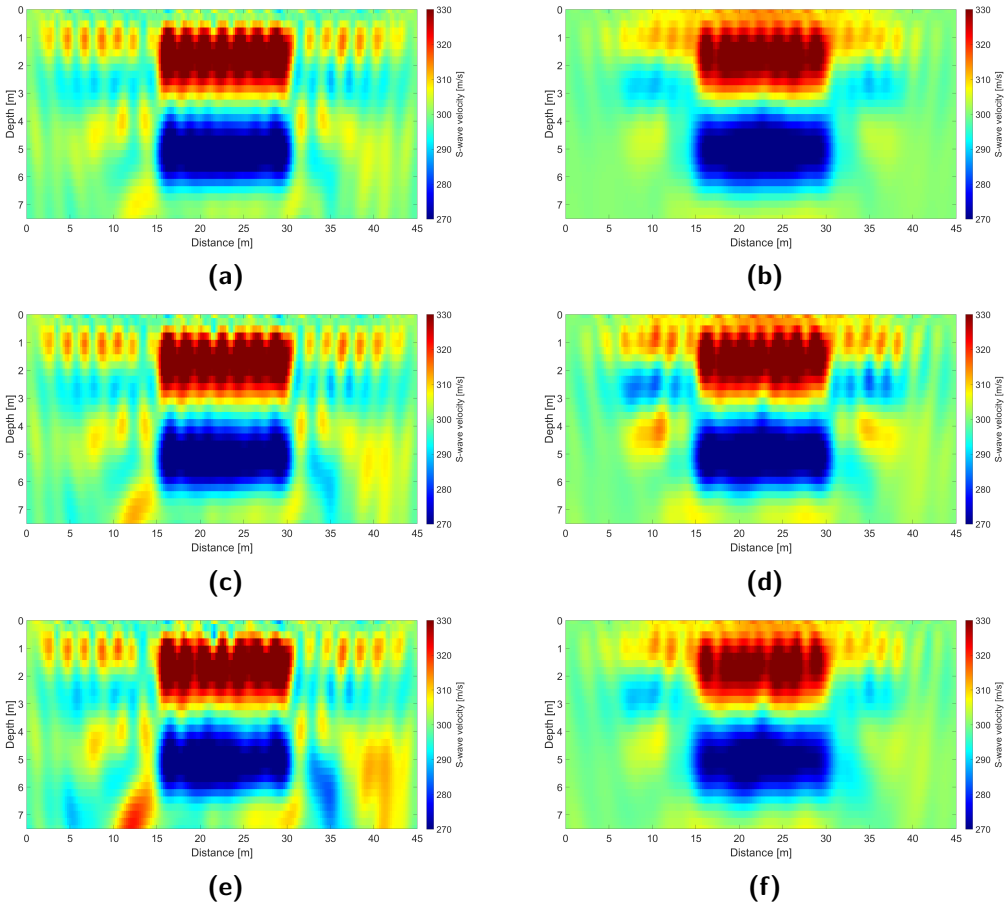


Figure 4-2: The inverted V_s models obtained from inverting the noise contaminated datasets, with an increasing SNR. (a),(c),(e) Denoting the models obtained from the data with respectively SNR of 20, 10 and 5 using least-squares norm. (b),(d),(f) Denoting the V_s models obtained form the data with respectively SNR of 20, 10 and 5 using the instantaneous-phase.

Although, the instantaneous-phase coherency misfit function should deliver an amplitude unbiased measure of the model fit (Schimmel et al., 2011; Liu et al., 2022), it can be observed from Figs. 4-2b - 4-2f, that the added random noise increases the uncertainty of the FWI solution. Further confirmed by comparison of the instantaneous-phase coherency FWI results

of the noise free data (Fig. 4-1b) and the noise contaminated data (Figs. 4-2b - 4-2f) demonstrating that the added noise does result in a more ill-posedness of the inversion.

However, from Figs. 4-2d and 4-2f it is observed that the result of the instantaneous-phase coherency FWI of the data with a SNR of 5 has less prominent artefacts than the result of the data with a SNR of 10. Indicating that the increased ill-posedness of the inversion is not directly related to an increase in noise content, rather the increased randomness of the data is a more likely cause.

4-1-2 P/SV-wave field test

The V_s inversion results of the Ostia field dataset are shown in Fig. 4-3. When comparing the outcome of the inversions, it is observed that both V_s models display similar features, however there are some major differences in first 2 m of the models. Further comparison remains difficult, as there is no knowledge of the ground truth. However, the approximate location of the tumulus is known and will therefore be used to judge the quality of the instantaneous-phase coherency compared to the least-squares norm.

When investigating the results obtained by instantaneous-phase coherency FWI shown in Fig. 4-3b, a low velocity anomaly is observed that coincides with the black rectangle used for denoting the location of the tumulus Ghose et al. (2020); Liu et al. (2022). Whereas, the results obtained from the least-squares norm FWI shown in Fig. 4-3a, display multiple velocity anomalies in the top 2 meters of the model. However, none of these coincide with the location of the tumulus.

A possible explanation for the least-squares norm FWI results not resolving the location of the tumulus could be due to amplitude errors in the shallow near-surface (Forbriger et al., 2014; Schäfer et al., 2014). These amplitude errors can lead to cycle-skipping, meaning the FWI algorithm has converged into a local minimum in the model space. It does however demonstrate the robustness of the instantaneous-phase coherency compared to the least-squares norm, as the FWI using the instantaneous-phase coherency was able to resolve the location of the tumulus accurately.

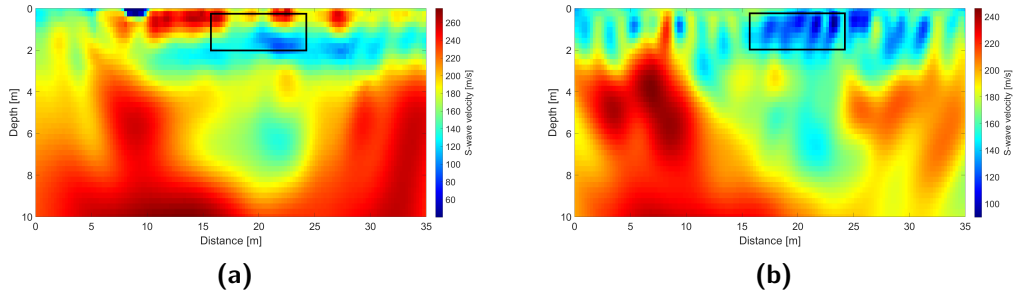


Figure 4-3: The V_s inversion results from the Ostia field dataset, with the black square denoting the location of the tumulus (Ghose et al., 2020; Liu et al., 2022). (a) The inversion result using the least-squares norm and (b) the inversion result using the instantaneous-phase coherency.

4-1-3 SH/Love-wave field test

Fig. 4-4 shows the V_s inversion results of the the Fossa Carolina field dataset. When inspecting the result of the least-squares norm FWI (Fig. 4-4a) it can be seen that the inversion result was unstable, demonstrating that the least-squares norm FWI is not able to deal consistently with near-surface applications. However, the V_s inversion result of the instantaneous-phase coherency (Fig. 4-4b) shows promising results. When comparing the result of the FATT (Fig. 3-3b) used as initial model for the FWI, it can be observed that the instantaneous-phase coherency FWI appears to add more detail to the canal structure denoted by the red oval in Fig. 4-4b. Which is consistent with the inversion results and archaeological sketch shown by Köhn et al. (2019).

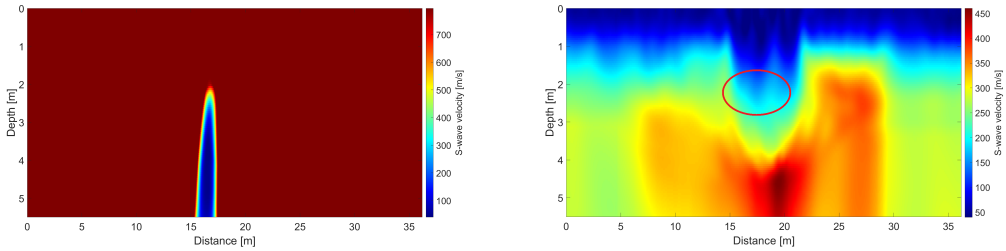


Figure 4-4: The V_s inversion results from the Fossa Carolina field dataset Köhn et al. (2019). (a) The inversion result using the least-squares norm and (b) the inversion result using the instantaneous-phase coherency.

4-2 3D dataset

Having shown the robustness of the instantaneous-phase coherency to amplitude variations when compared to the least-squares norm. Next, the results of the tests that used information from the 3D dataset are discussed.

4-2-1 Accuracy of the instantaneous phase coherency

To evaluate the accuracy of the instantaneous-phase coherency FWI, a total of 50 inversions where performed on layer cake like models. In the coming section a set of representative results from three test cases are presented. The other results of which the inversion was able to produce a result can be found in Appendix A, B and C for respectively the first, second and third set of inversions.

First the results of the layer cakes of which all layers could vary in depth and thickness are investigated. When investigating the results of the test runs Fig. 4-5 it can be observed that the stronger velocity contrasts in the true models (Figs. 4-5a and 4-5c) are better resolved compared to the more gradual velocity increases. Although there are some artificial layers introduced in the result between the two top most layers. These artificial layers are most likely the result of the initial model not being close enough to the true model, caused by the large contrast in V_s between the two top layers of 700 m/s and 1500 m/s. Since the initial

model has to be smoothed for the FWI algorithm to work, resulting in the strong velocity contrast to be smeared out over several meters, which in turn will cause the error.

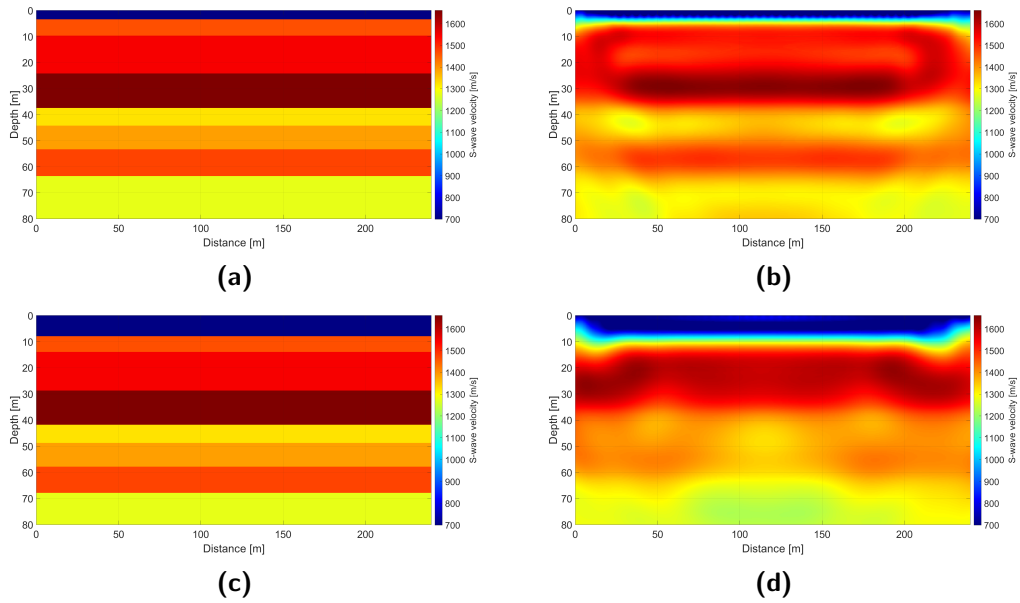


Figure 4-5: The V_s inversion results of the first layer cake test runs. (a) The true model of one of the better resolved models and (b) the inversion result. (c) The true model of one of the poorly resolved models and (d) its inversion result.

Since the typical soil profiles in the Middle East are generally only a few meters deep (Poulos, 2018), and the inversion results of the models that had a shallow top layer performed slightly better a new set of logs was created where all layers with exception of the top layer where randomly varying in depth and thickness. However, when investigating the result shown in Fig. 4-6b, it can be observed that the result does not resemble the true model (Fig. 4-6a) at all.

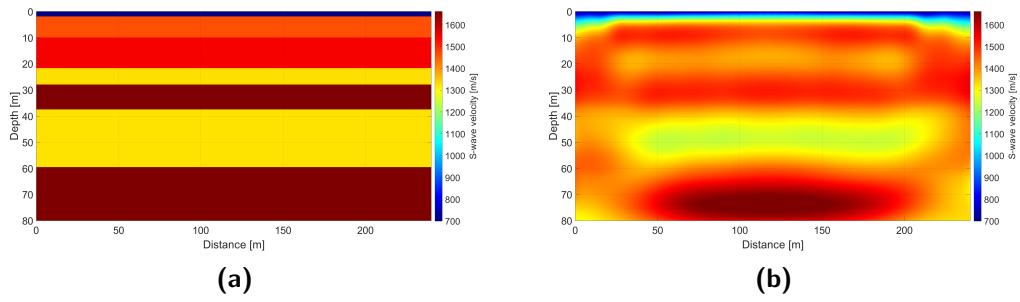


Figure 4-6: The V_s inversion result from one of the second layer cake test runs. (a) The true model of one of the better resolved models and (b) the inversion result.

Another possible explanation for the poorly resolved models could be due to the penetration depth. For the previous two tests a sequential strategy was used with the minimum frequency being set to 10 Hz. However, as a rule of thumb the effective depth that can be resolved is roughly $\lambda/3$. Therefore, a minimum frequency of 10 Hz and a minimum velocity of 700 m/s

result in a wavelength of 70 m and a resolution of approximately 23 m. To increase the vertical resolution of the inversion the first 10 models used for the second test, were reran with a source bandwidth of $0-60$ Hz and a sequential inversion approach with a minimum frequency of 0 Hz instead.

Fig. 4-7b shows the V_s inversion result from the same true model used to obtaining the result shown in Fig. 4-6b. When comparing the inversion results of the same true model (Figs. 4-6a and 4-7a) it can be observed that the long wavelength part of the model appears better resolved in the new inversion result (Fig. 4-7b). Furthermore, the high velocity layers are better resolved as well in the new inversion result. Demonstrating that the increase in bandwidth did improve the result somewhat. However, the overall quality of the inversions remains poor for the layer cake models.

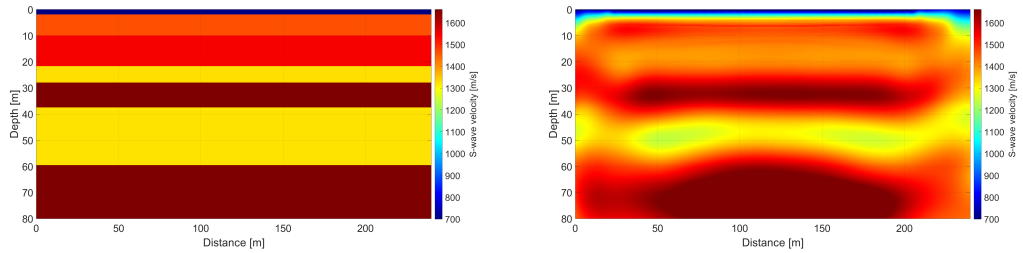


Figure 4-7: The V_s inversion result from one of the third layer cake test runs. (a) The true model, equal to the model used in Fig. 4-6 and (b) the inversion result.

4-2-2 Complex subsurface model

Finally, the V_s inversion results of the complex subsurface model are shown in Figs. 4-9 and 4-10. To evaluate the results, the true model (Fig. 4-8) is divided into separate zones annotated by the letters **A**, **B**, **C**, **D**, **E** and **F**.

When comparing the inversion results (Figs. 4-9 and 4-10) with the true model (Fig. 4-8) it can be observed that the high velocity body marked by **A**, is reasonably resolved up to a depth of $25-30$ m. Furthermore, the velocity body (**A**) is characterised by a lateral pinch-out to the left of the model (**C**), both inversions did resolve the shape of the pinch-out reasonably well. Furthermore, the low velocity zone marked by **G**, is reasonably well resolved for both inversions. As well as the high velocity body marked by **D**. The region characterised by the velocity changes (**F**) appears poorly resolved in both inversion results.

Another point of interest is the low velocity anomaly (**B**) located inside the high velocity body. As the low velocity anomaly is only a couple meters in width and height, it would be of potential interest to see which of the two approaches is able to resolve it more accurately. When comparing the two inversion results (Figs. 4-9 and 4-10) it becomes clear that the inversion result using the geo-technical initial model delivers a more accurate estimate of the low velocity body.

The thin layer marked by **E** is unresolved in both of the inversion results, with exception of the part near the marker. Furthermore, the thin layer appears to influence the solution

of the high velocity body (**A**) as well, as it appears to be seen as part of the body. This could be due to a inaccurate resolution caused by picking a bandwidth that is too small. A possible solution could be to raise the maximum frequency used during the modelling to 90 Hz. Similarly to the layer cake inversions a potential adjustment to the bandwidth could potentially improve upon this part of the solution.

With exception of the velocity anomaly marked by **B**, which is very accurately resembled in Fig. 4-9. The overall quality of the inversion result using Fig. 3-6c as initial model appears to be slightly better. This is a reasonable result as preferably taking a model close to the global minimum would increase the chance of obtaining the "best" possible model fit (Virieux and Operto, 2009).

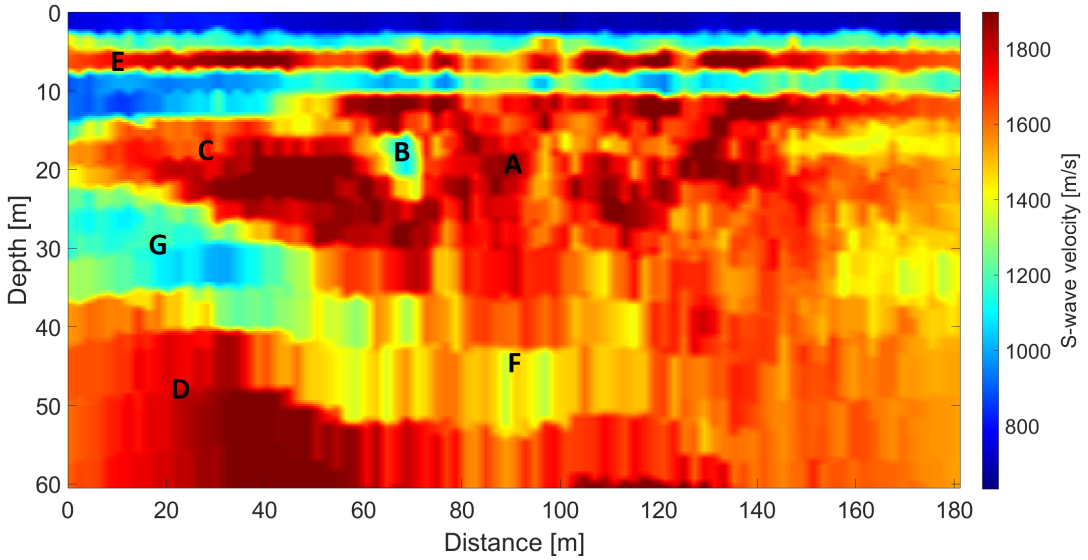


Figure 4-8: The true V_s models used to mimic the complex subsurface model inversions. **A** denotes the middle of a large body of high velocity, starting at approximately 10 m depth and pinches out towards the left of the model. **B** marks a low velocity anomaly. **C** denotes the pinched out part of the velocity body. **D** denotes the bottom right corner of the model, characterised by a V_s. **E** denotes a thin layer at a depth of approximately 8 m. **F** denotes a region below the large velocity body, characterised by some variations in the velocity. **G** denotes a low velocity body.

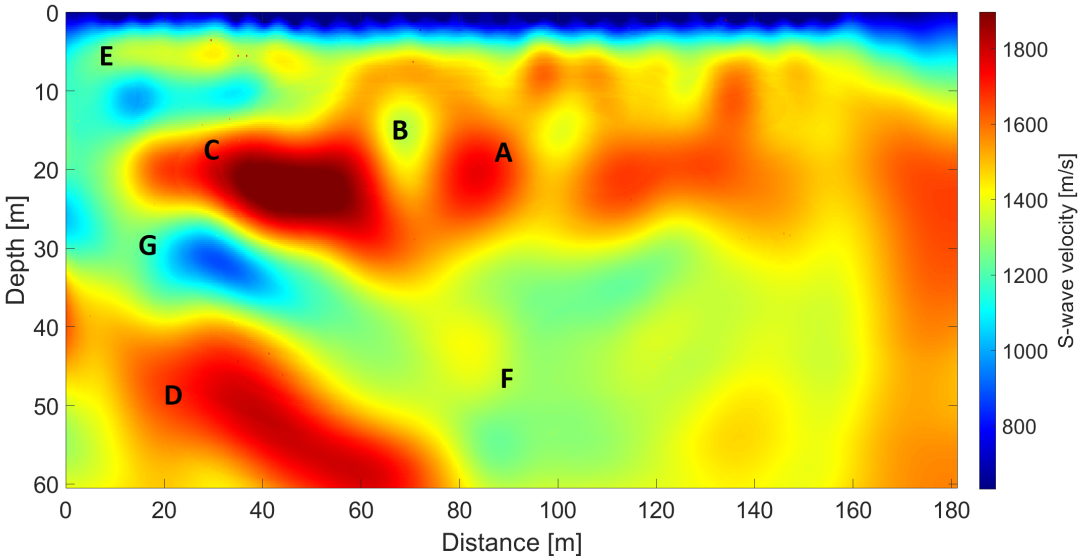


Figure 4-9: The V_s inversion result obtained with the geo-technical initial model. Letters **A**, **B**, **C**, **D**, **E** and **F** correspond to the same locations marked in Fig. 4-8.

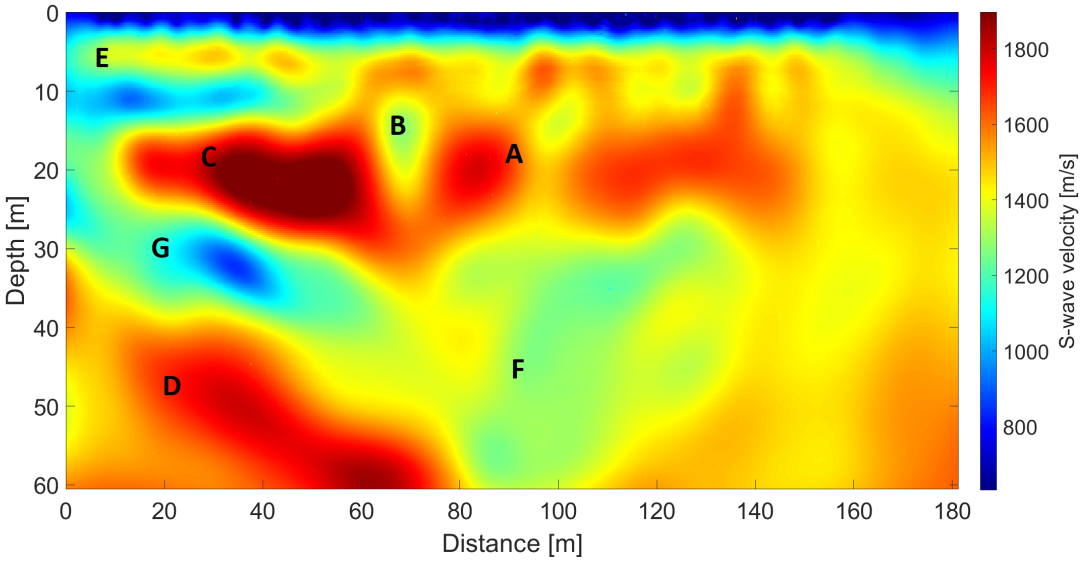


Figure 4-10: The V_s inversion result obtained with the smoothed true model as if results from another inversion method where used in building the initial model. Letters **A**, **B**, **C**, **D**, **E** and **F** correspond to the same locations marked in Fig. 4-8.

Chapter 5

Conclusion

This work investigated the instantaneous-phase coherency misfit function when used in FWI, and is proposed to have superior performance over a traditional least-squares norm when dealing with real datasets and noise contaminated synthetic data. The inversions results shown earlier clearly illustrate a promising performance for the instantaneous-phase coherency when applied for near-surface FWI.

From the inversion results using the synthetic test cases, it is shown that the instantaneous-phase coherency FWI delivers a more robust model estimate when noise is added. Furthermore, a decrease in SNR did not result in a lesser resolved model. Whereas, the least-squares norm FWI results decreased in quality with a decrease in SNR. Additionally, the results of the two field datasets, demonstrate the ability of the instantaneous-phase coherency to resolve complex near-surface models, whereas the least-squares norm is shown to be incapable of delivering accurate results. Therefore, it is concluded that the instantaneous-phase coherency delivers a robust and accurate result when inverting near-surface datasets.

Next, the instantaneous-phase coherency FWI was applied to layer cake models, in an attempt to determine the accuracy of the proposed method. The results however, are inconclusive, as some layers are resolved while others are not. Furthermore, the results shown artefacts as a results of inaccurate model estimation. A new selection of the inversion bandwidth however, did show improvements to the outcome of the inversion. Concluding that the instantaneous-phase coherency does require a initial model that is somewhat close to the true model, as well as appropriate selection of the bandwidth used for the source signal and the sequential inversion strategy will improve the outcome of the instantaneous-phase coherency FWI.

Finally, a complex subsurface model was created to investigate the ability of the instantaneous-phase coherency of resolving complex subsurface structures. The results demonstrate that the FWI code was partly able to reconstruct the complex subsurface, as it could accurately recover the small anomaly that was located inside the high velocity layer. However, the results most likely could have been improved by better design of the initial model and a better selection of the bandwidth used during the forward and inverse modelling.

Based on the conclusions presented in this thesis, a couple of recommendations are made for future research.

This thesis clearly demonstrates the importance of correct bandwidth selection for the forward and inverse modelling parts. In future studies, design of the source bandwidth as well as design of a potential band-pass filter used during the inversion should be done carefully. During the design of the source and band-pass filter bandwidth, the desired resolution should be leading.

Although, the instantaneous-phase coherency shows promising results, the layer cake models clearly demonstrate potential cycle skipping due to erroneous introduction of prior knowledge into the initial model. A possible solution could be to introduce prior knowledge into the misfit function through means of a prior model term. An example of such an approach is given by Asnaashari et al. (2013), where the initial model is used to model wave propagation and therefore used to determine the data misfit and the gradient of the misfit. While the prior model is only used during the optimisation step and as constraint on the inversion. The method shown by Asnaashari et al. (2013), shows an improved well-posedness of the problem and reduces the kinematic inaccuracy of the starting model.

Finally, if allowed by the computational resources available the method could be extended to 3D. This would greatly reduce the problems commonly associated with 2D FWI as these are partly due to correction of 3D to 2D geometrical spreading correction. However, the extension to 3D is very costly, as is shown by Butzer et al. (2013).

Bibliography

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., and Virieux, J. (2013). Regularized seismic full waveform inversion with prior model information. *Geophysics*, 78(2):R25–R36.

Ben-Menahem, A. and Singh, S. J. (2012). *Seismic waves and sources*. Springer Science & Business Media.

Bohlen, T. (2002). Parallel 3-d viscoelastic finite difference seismic modelling. *Computers & Geosciences*, 28(8):887–899.

Bohlen, T. and Saenger, E. H. (2006). Accuracy of heterogeneous staggered-grid finite-difference modeling of rayleigh waves. *Geophysics*, 71(4):T109–T115.

Bozdağ, E., Trampert, J., and Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.

Brossier, R. (2009). Seismic imaging of complex structures by 2d elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC63–WCC76.

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

Butzer, S., Kurzmann, A., and Bohlen, T. (2013). 3d elastic full-waveform inversion of small-scale heterogeneities in transmission geometry. *Geophysical Prospecting*, 61(6-Challenges of Seismic Imaging and Inversion Devoted to Goldin):1238–1251.

Dokter, E., Köhn, D., Wilken, D., De Nil, D., and Rabbel, W. (2017). Full waveform inversion of sh-and love-wave data in near-surface prospecting. *Geophysical Prospecting*, 65(S1):216–236.

Falk, J., Tessmer, E., and Gajewski, D. (1996). Tube wave modeling by the finite-difference method with varying grid spacing. *Pure and Applied Geophysics*, 148(1):77–93.

Fichtner, A., Trampert, J., Cupillard, P., Saygin, E., Taymaz, T., Capdeville, Y., and Villasenor, A. (2013). Multiscale full waveform inversion. *Geophysical Journal International*, 194(1):534–556.

Forbriger, T., Groos, L., and Schäfer, M. (2014). Line-source simulation for shallow-seismic data. part 1: theoretical background. *Geophysical Journal International*, 198(3):1387–1404.

Ghose, R., Liu, J., Draganov, D., Ngan-Tillard, D., Warnaar, M., Brackenhoff, J., van den Berg, J., and Stoger, H. (2020). Ultra-shallow shear-wave reflections locating near-surface buried structures in the unexcavated southern fringe of the ancient ostia, rome. In *Designating Place: Archaeological Perspectives of Built Environment in Ostia and Pompeii*, pages 51–62. Leiden University Press.

Groos, L. (2013). *2D full waveform inversion of shallow seismic Rayleigh waves*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT), Diss., 2013.

Groos, L., Schäfer, M., Forbriger, T., and Bohlen, T. (2017). Application of a complete workflow for 2d elastic full-waveform inversion to recorded shallow-seismic rayleigh waves. *Geophysics*, 82(2):R109–R117.

Heider, S. (2014). *2D Elastic Full-Waveform Tomography of Vibro-Seismic Data in Crystalline Host Rock at the GFZ-Underground-Lab, Freiberg*. PhD thesis, Universität Karlsruhe.

Igel, H. (2017). *Computational seismology: a practical introduction*. Oxford University Press.

Köhn, D. (2011). *Time domain 2D elastic full waveform tomography*. PhD thesis.

Köhn, D., Wilken, D., De Nil, D., Wunderlich, T., Rabbel, W., Werther, L., Schmidt, J., Zielhofer, C., and Linzen, S. (2019). Comparison of time-domain sh waveform inversion strategies based on sequential low and bandpass filtered data for improved resolution in near-surface prospecting. *Journal of Applied Geophysics*, 160:69–83.

Komatitsch, D. and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 72(5):SM155–SM167.

Levander, A. R. (1988). Fourth-order finite-difference p-sv seismograms. *Geophysics*, 53(11):1425–1436.

Liu, J., Ghose, R., and Draganov, D. (2022). Characterizing near-surface structures at the ostia archaeological site based on instantaneous-phase coherency inversion. *Geophysics*, 87(4):1–50.

Maurer, H., Greenhalgh, S. A., Manukyan, E., Marelli, S., and Green, A. G. (2012). Receiver-coupling effects in seismic waveform inversions. *Geophysics*, 77(1):R57–R63.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, New York, NY, USA, 2e edition.

Pan, Y., Gao, L., and Bohlen, T. (2019). High-resolution characterization of near-surface structures by surface-wave inversions: From dispersion curve to full waveform. *Surveys in Geophysics*, 40(2):167–195.

Park, C. B., Miller, R. D., and Xia, J. (1999). Multichannel analysis of surface waves. *Geophysics*, 64(3):800–808.

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(2):495–503.

Poulos, H. G. (2018). A review of geological and geotechnical features of some middle eastern countries. *Innovative Infrastructure Solutions*, 3(1).

Robertsson, J. O., Levander, A., Symes, W. W., and Holliger, K. (1995). A comparative study of free-surface boundary conditions for finite-difference simulation of elastic/viscoelastic wave propagation. pages 1277–1280. Society of Exploration Geophysicists.

Schäfer, M., Groos, L., Forbriger, T., and Bohlen, T. (2014). Line-source simulation for shallow-seismic data. part 2: full-waveform inversion—a synthetic 2-d case study. *Geophysical Journal International*, 198(3):1405–1418.

Schimmel, M. and Paulssen, H. (1997). Noise reduction and detection of weak, coherent signals through phase-weighted stacks. *Geophysical Journal International*, 130(2):497–505.

Schimmel, M., Stutzmann, E., and Gallart, J. (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale. *Geophysical Journal International*, 184(1):494–506.

Socco, L. and Strobbia, C. (2004). Surface-wave method for near-surface characterization: a tutorial. *Near surface geophysics*, 2(4):165–185.

Socco, L. V., Foti, S., and Boiero, D. (2010). Surface-wave analysis for building near-surface velocity models—established approaches and new perspectives. *Geophysics*, 75(5):75A83–75A102.

Tarantola, A. (1984). Linearized inversion of seismic reflection data. *Geophysical prospecting*, 32(6):998–1015.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.

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

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.

Xia, J., Miller, R. D., and Park, C. B. (1999). Estimation of near-surface shear-wave velocity by inversion of rayleigh waves. *Geophysics*, 64(3):691–700.

Appendix A

Appendix A

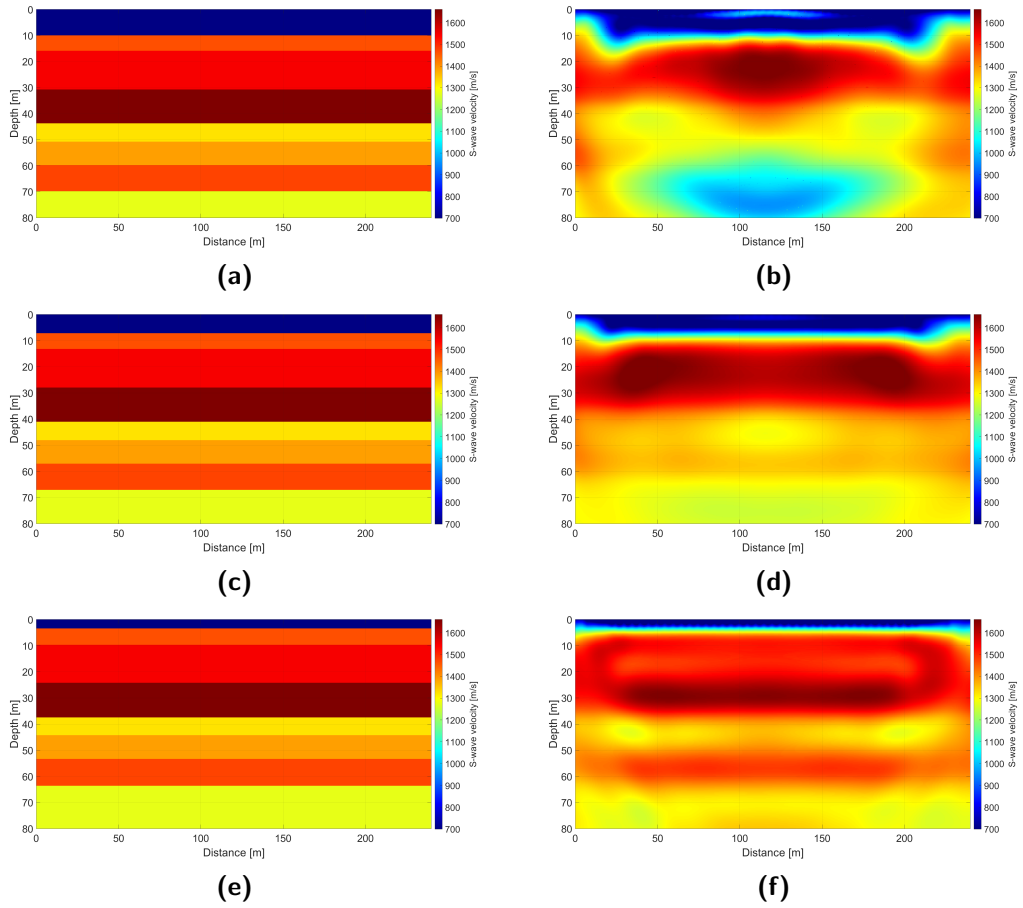


Figure A-1: V_s inversion results of the first set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

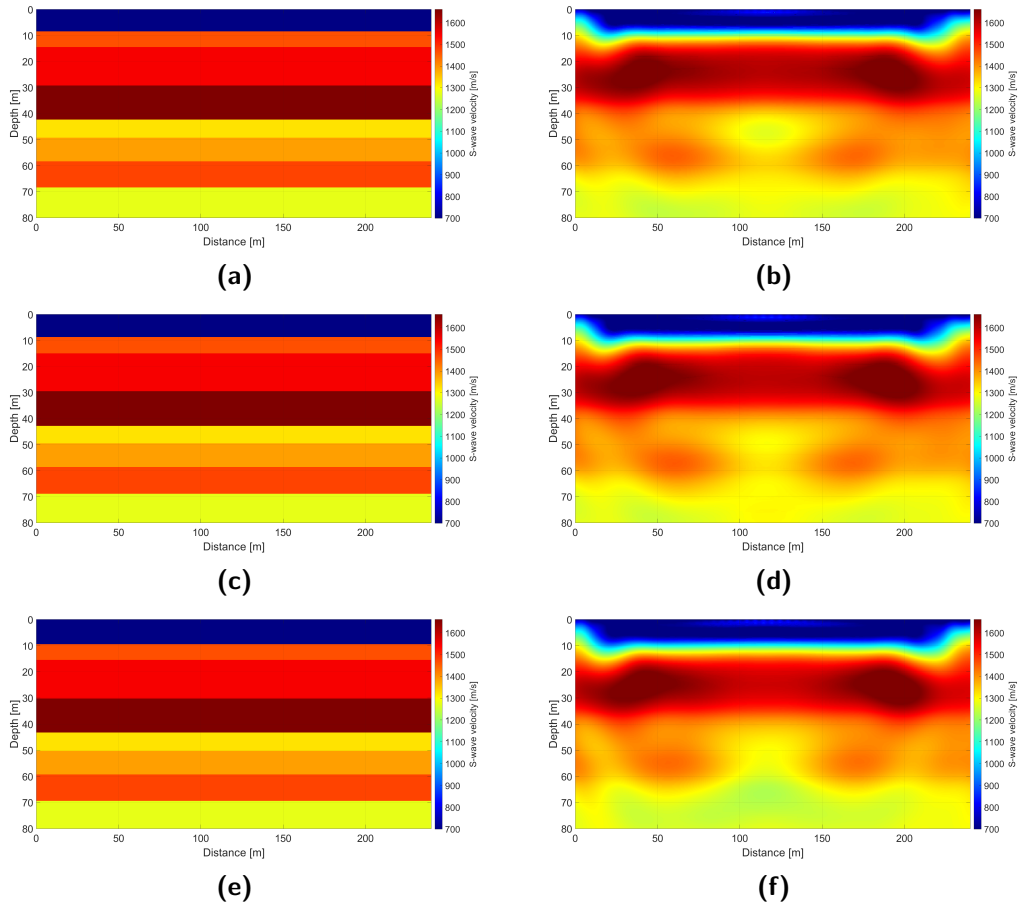


Figure A-2: V_s inversion results of the first set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

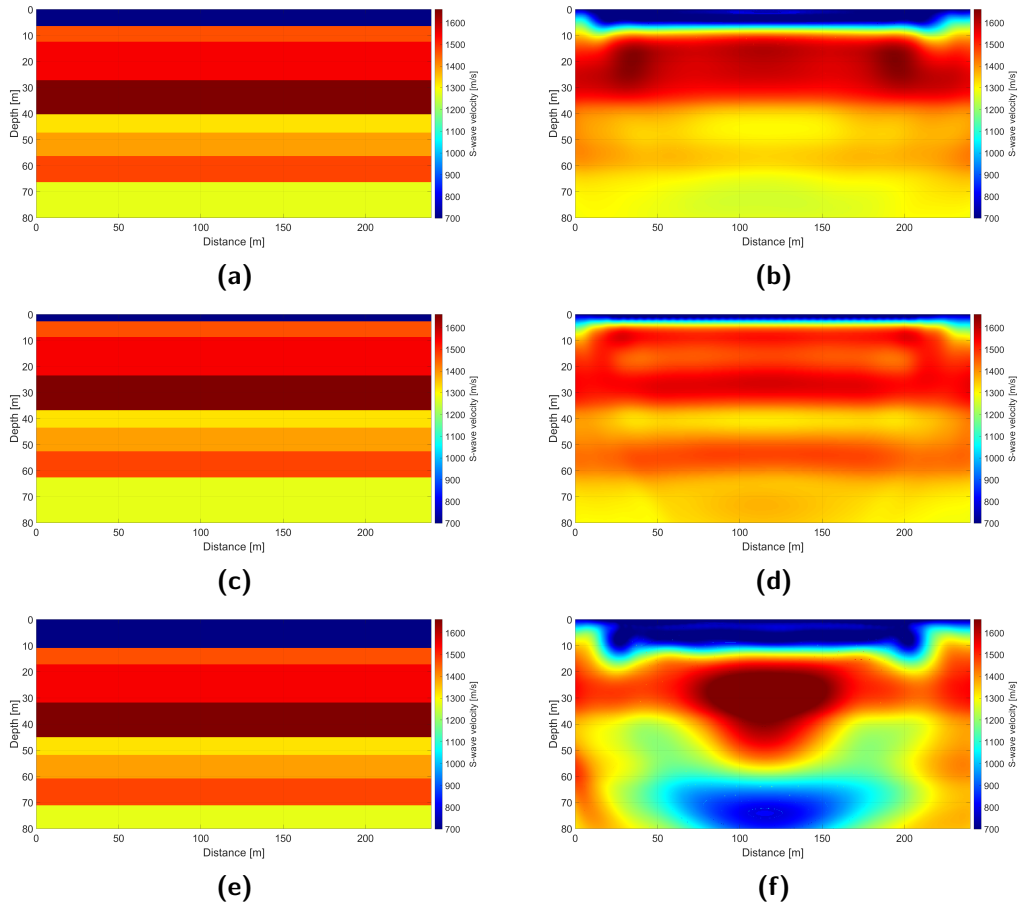


Figure A-3: V_s inversion results of the first set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

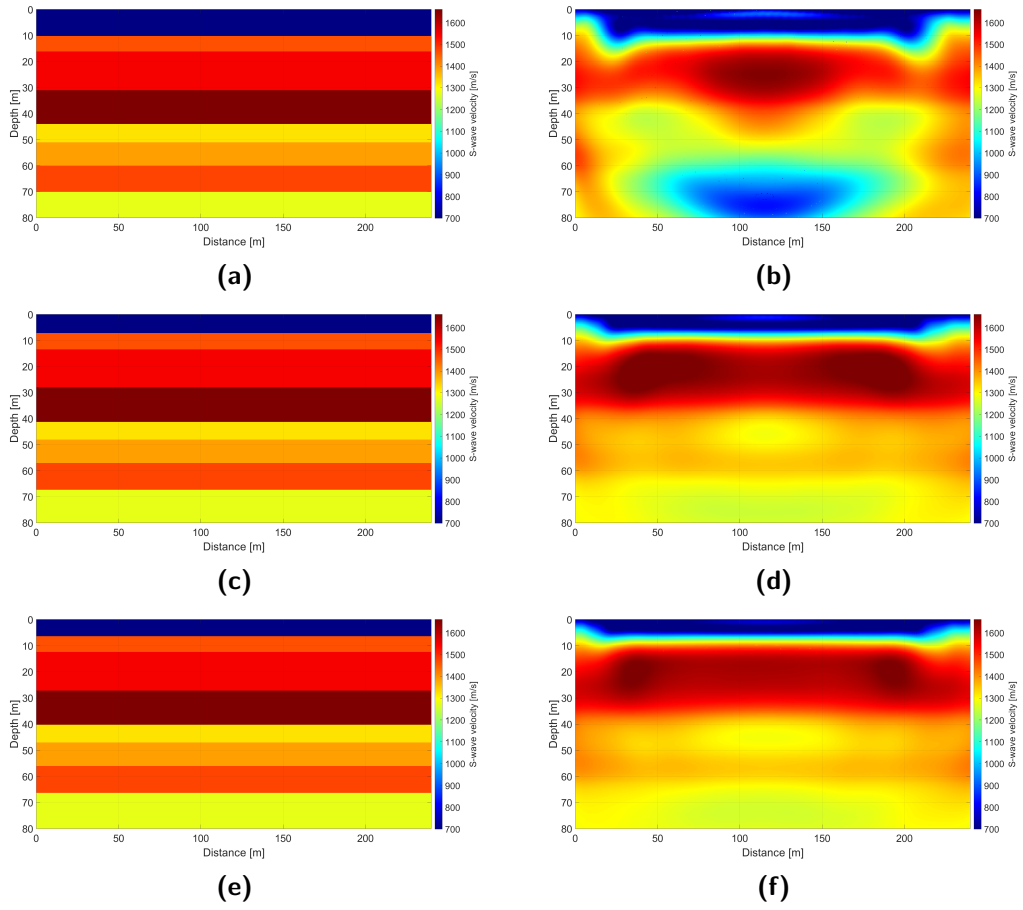


Figure A-4: V_s inversion results of the first set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

Appendix B

Appendix B

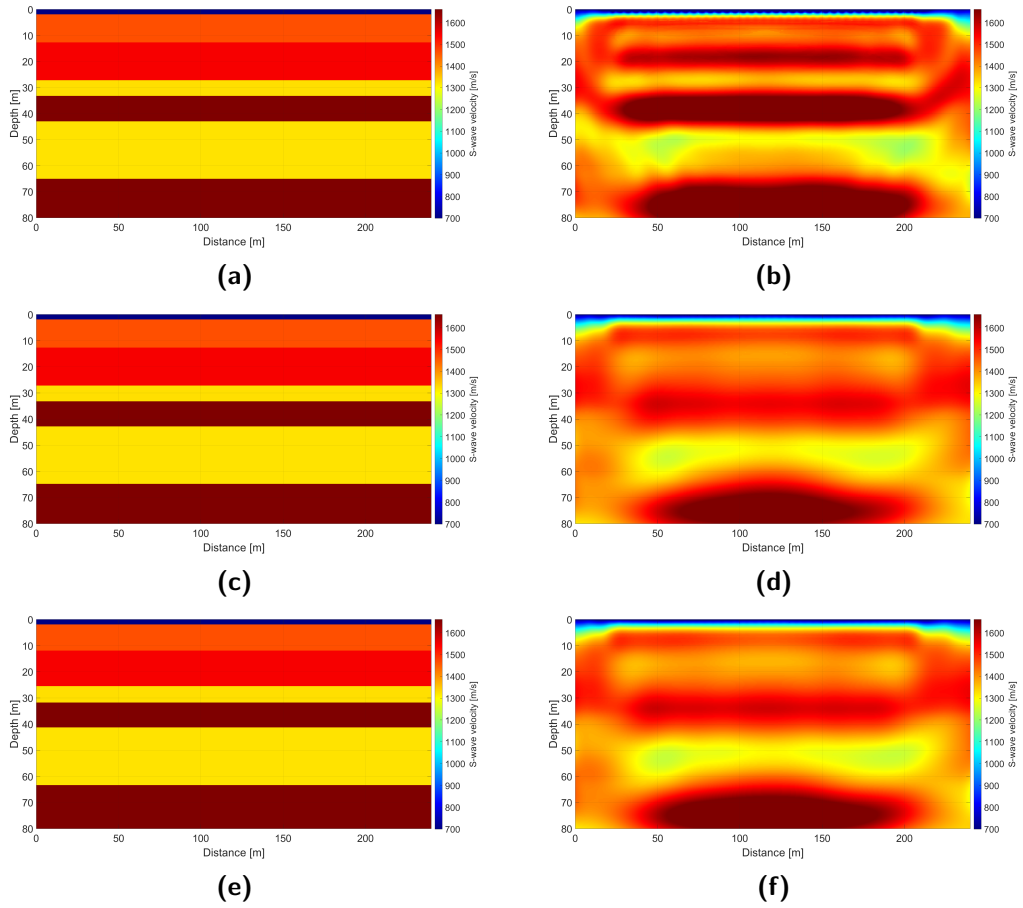


Figure B-1: V_s inversion results of the second set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

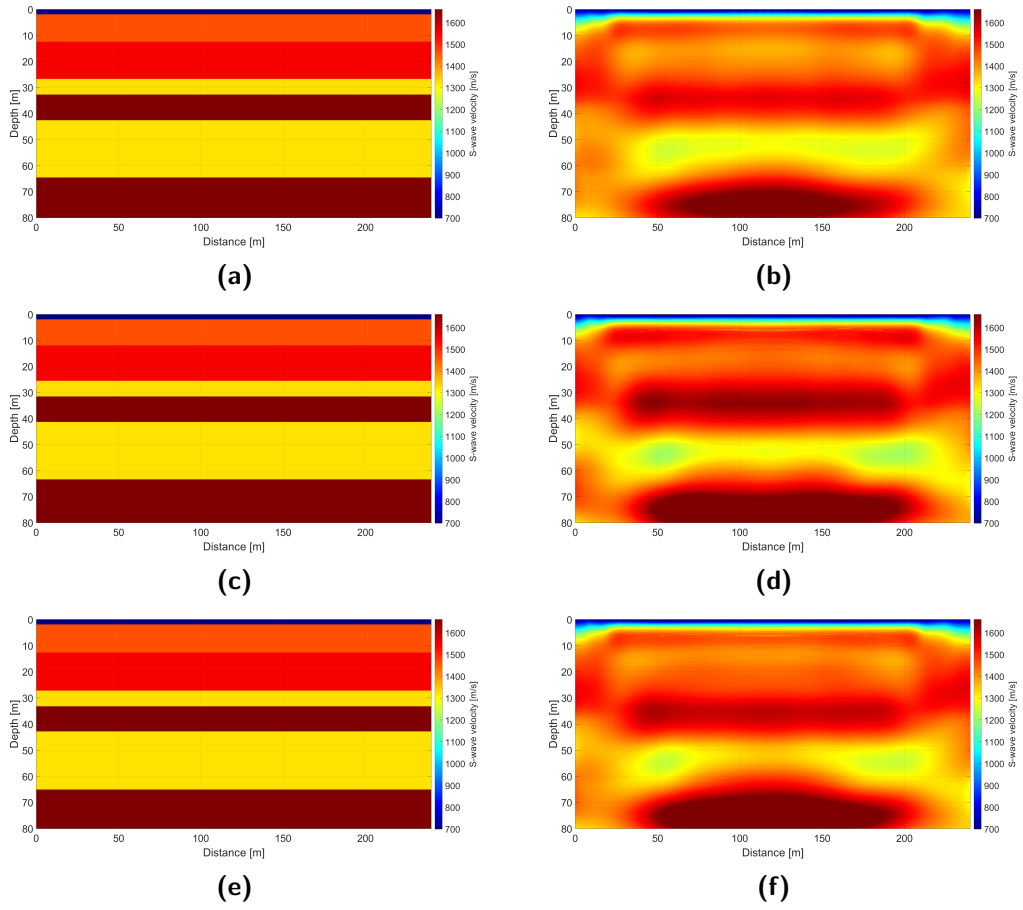


Figure B-2: V_s inversion results of the second set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the the corresponding inversion results.

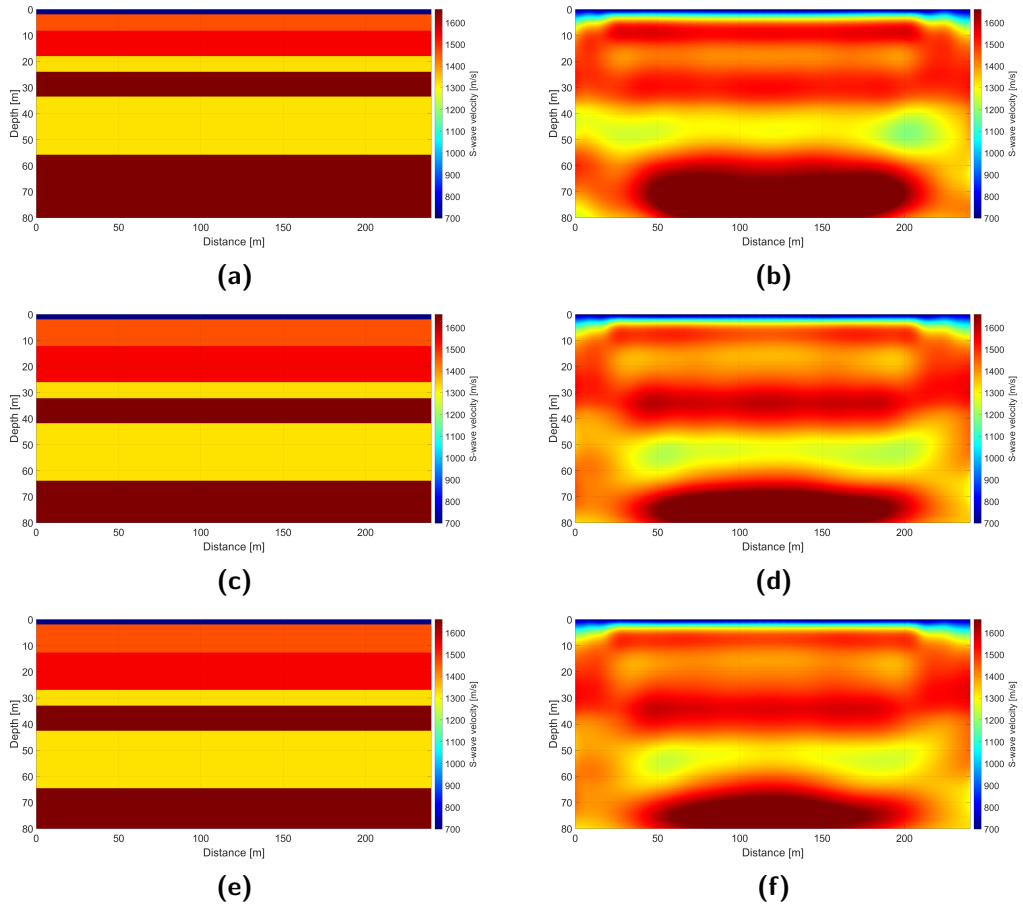


Figure B-3: V_s inversion results of the second set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

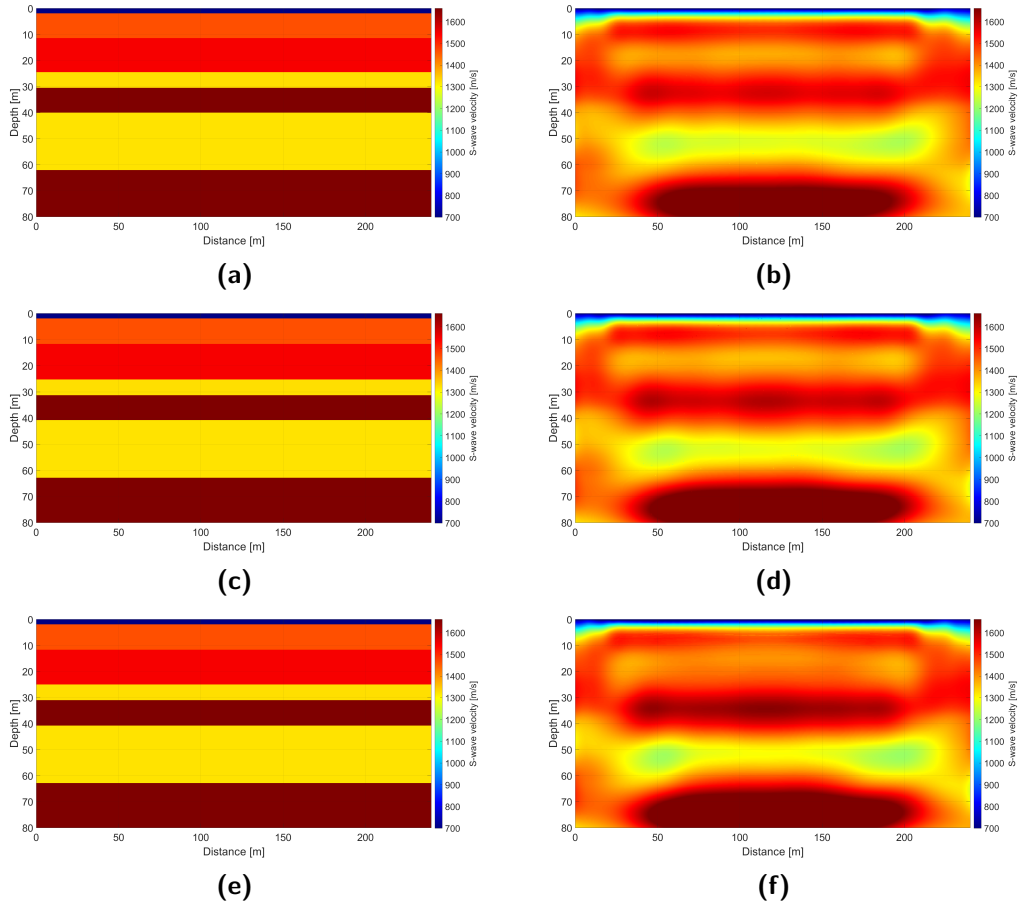


Figure B-4: V_s inversion results of the second set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

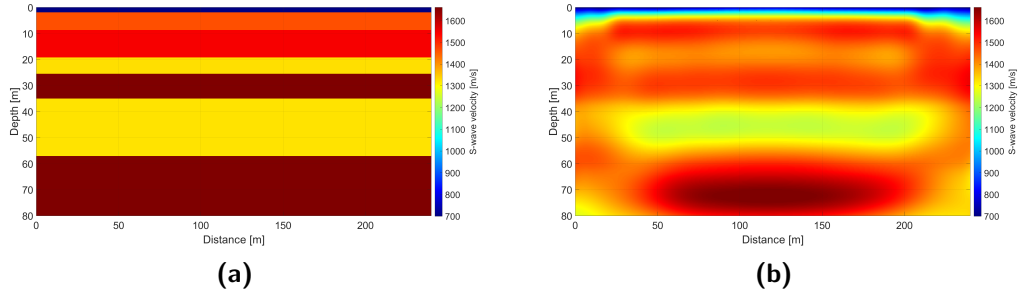


Figure B-5: V_s inversion result of the second set of layer cake test runs. (a) The true model used in the inversion with, (b) the inversion result.

Appendix C

Appendix C

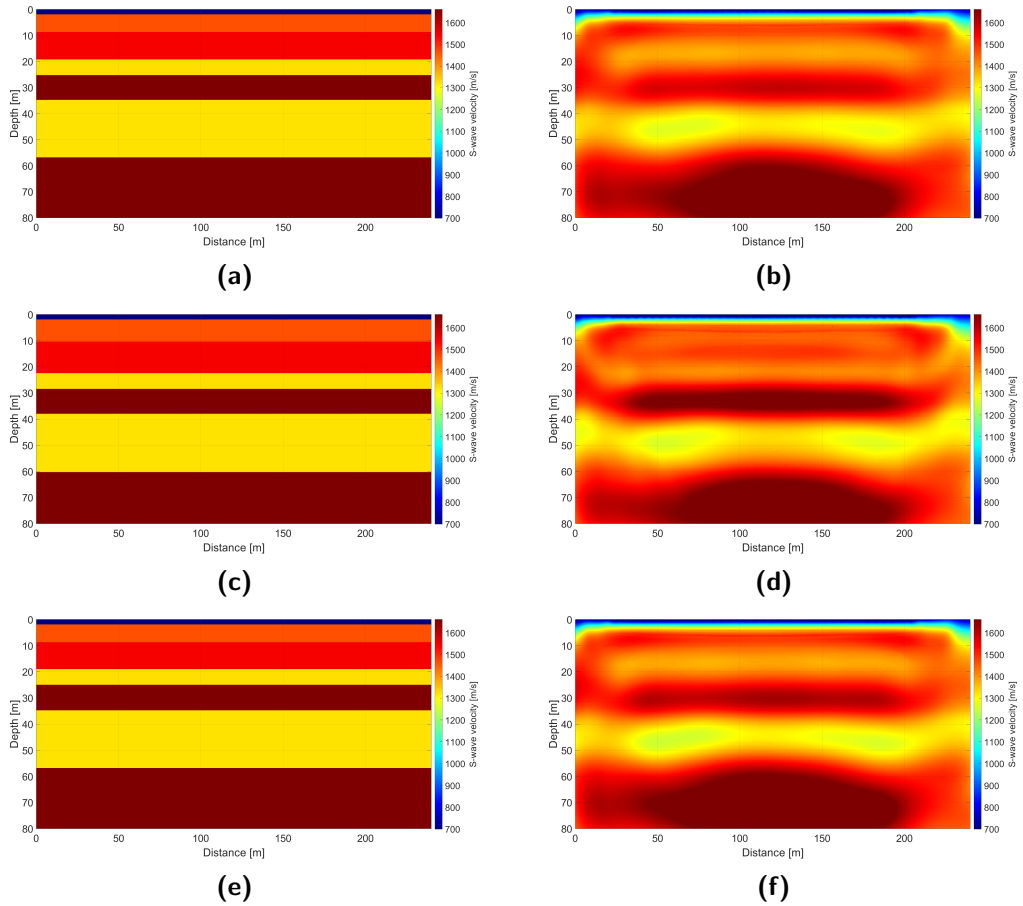


Figure C-1: V_s inversion results of the third set of layer cake test runs. (a), (c) and (e) The true models used in the inversion with, respectively, (b), (d) and (f) the corresponding inversion results.

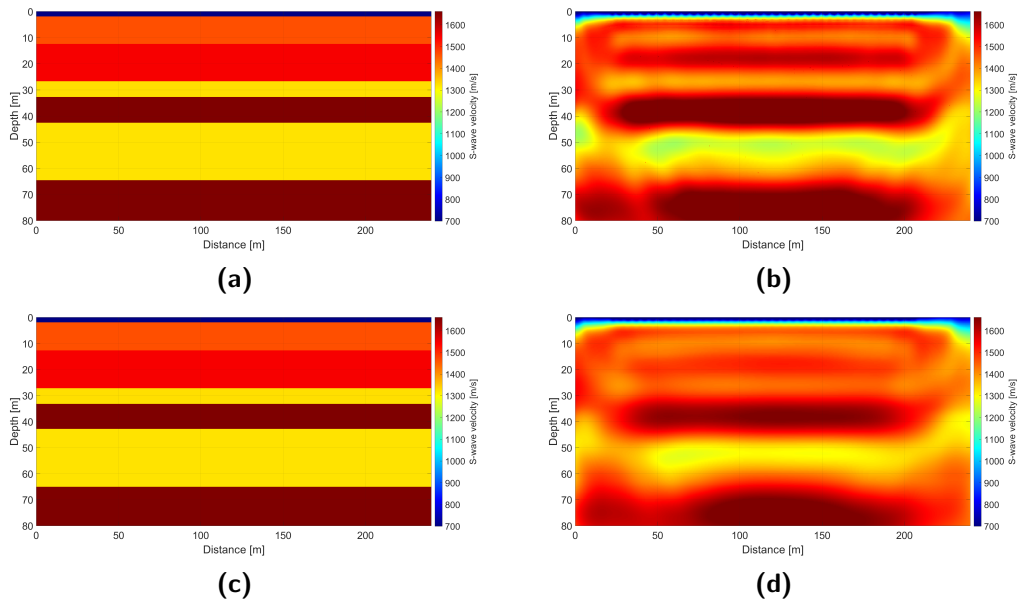


Figure C-2: V_s inversion results of the third set of layer cake test runs. (a) and (c) The true models used in the inversion with, respectively, (b) and (d) the corresponding inversion results.