HAL
open science

Two-dimensional near-surface seismic imaging with surface waves : alternative methodology for waveform inversion

Carlos Andrés Pérez Solano

To cite this version:

Carlos Andrés Pérez Solano. Two-dimensional near-surface seismic imaging with surface waves : alternative methodology for waveform inversion. Earth Sciences. Ecole Nationale Supérieure des Mines de Paris, 2013. English. NNT : 2013ENMP0023 . pastel-00932790

HAL Id: pastel-00932790

https://pastel.hal.science/pastel-00932790v1

Submitted on 17 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

Ecole doctorale n° 398 : Géosciences et Ressources Naturelles

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

l'École nationale supérieure des mines de Paris

Spécialité "Dynamique et Ressources des Bassins Sédimentaires"

présentée et soutenue publiquement par

Carlos Andrés PÉREZ SOLANO

Le 9 décembre 2013

Two-dimensional near-surface seismic imaging with surface waves: alternative methodology for waveform inversion

Directeur de thèse : Hervé CHAURIS

Maître de thèse : Daniela DONNO

Jury

M. Michel DIETRICH, Directeur de Recherche CNRS, Université Joseph Fourier	Rapporteur
M. Gilles GRANDJEAN, Ingénieur-Chercheur, BRGM	Rapporteur
M. Olivier KAUFMANN, Professeur, Université de Mons - Président du jury	Examinateur
M. Claudio STROBBIA, Géophysicien Sénior, Total	Examinateur
M. Hervé CHAURIS, Professeur, MINES ParisTech	Examinateur
Mme Daniela DONNO, Chargée de Recherche, MINES ParisTech	Examinatrice
M. Ludovic BODET, Maître de Conférences, UPMC	Invité
M. René-Édouard PLESSIX, Chercheur Sénior, Shell	Invité

T H E S E

MINES ParisTech
Centre de Géosciences
35 rue Saint-Honoré, 77305 Fontainebleau cedex

Acknowledgement

I would like to thank my advisers Daniela Donno and Hervé Chauris. I thank them for their availability, support and encouragement given along my stay at MINES ParisTech. The discussions with them have always been plenty of learning for me. They also contributed in making my stay in France very pleasant; they were there for me even when I needed support in administrative matters. The monitoring provided by Daniela let me advance in the research while reminding me the importance of the scientific fundamentals. Thanks to Hervé, I have been able to understand novel approaches as he has stimulated me to analyse all the new ideas. Together, we have discussed, developed equations and in addition we have shared breakfasts, lunches and dinners. I thank them also for having trusted in my ability to develop this research even if they did not know me before I started my PhD thesis.

I would like to thank the members of my jury. Thanks to Gilles Grandjean and Michel Dietrich for having accepted to review my thesis in detail. Thanks to Claudio Strobbia and Olivier Kaufmann for their questions and reviews as examiners of my thesis. I thank Ludovic Bodet and René-Édouard Plessix for their advices and suggestions given as invited members of my jury.

I especially acknowledge Gian Piero Deidda from the University of Cagliari for having provided the real data set used for the real application in this thesis. He also provided the tomography model used as initial model for inversion.

Many thanks to my colleagues in the Geophysics Team of the Geosciences Center at MINES ParisTech. Thanks to Alexandrine Gesret and Mark Noble for their availability, discussions and advice. I especially thank Véronique Lachasse for all the support and advice provided at Fontainebleau. Since my first day at the office and up to now, she has always been there ready to offer her advice regarding any preoccupation I may have. I also thank Dominique Vassiliadis from the Geosciences Center who helped me solving administrative matters.

I would like to thank my first officemates Nidhal Belayouni, Sébastien Penz and Jiangbo Liu. Thanks Nidhal and Sébastien for offering me valuable advice in both scientific and personal matters related to my stay in Fontainebleau. Special thanks to Elise Vi Nhu Ba who always encouraged me to speak French and who also offered me her friendship. The last officemates Fang Wang, Charles-Antoine Lameloise, Emmanuel Cocher and Sven Schilke offered me also their friendship and support. I have enjoyed a lot the billard games with Nidhal, Charles-Antoine, Sven and Emmanuel.

I participated in a three month internship at Shell. I would like to thank Alexandre Stopin, René-Édouard Plessix and Fons ten Kroode for the scientific and professional discussions.

Trough the three years spent at Fontainebleau, I had the opportunity to meet some friends. I would like to mention the students of the specialised formations CESAM, CESECO, CESPROMIN and CFSG. I thank Julian, Yecid and Edgar for their company and support during my first year. In the second year, I spent several good times with

Ivan, Miguel, Emma, Sary, Santiago, Ana Vanessa and Juan Carlos.

I especially thank Thomas Gillot, my colleague from the Geology Team, who offered me his friendship during my entire stay at Fontainebleau. He supported me and was present in important moments in my life. I especially thank Sabrina Prince for her friendship. We have had a lot of happy moments in Fontainebleau when making plans with them. Thanks Tom and Sab.

Finally, I would like to thank my family. My research would not have been possible without the support of my parents Francisco and Luz Mar as well as the support of my brother Fernando. Gracias por apoyar mis decisiones y por estimular el desarrollo de mis ideas, siempre contando con sus valiosos consejos. Un saludo de agradecimiento para mi cuñada Gloria y mi sobrina Ana María. Gracias a mis suegros Orlando y Genys y a mi cuñado Oscar por el apoyo que me han brindado.

Before continuing with the text of this research, I would like to especially thank my wife. Gracias Laura Meliza por compartir conmigo estos años. Gracias por tu apoyo, consejos y por escucharme siempre, primero desde la distancia, y ahora desde tu cercana compañia.

Abstract

In the context of hydrocarbon exploration, seismic imaging is widely used to characterise the first few kilometres of the Earth's interior. Information is carried on by elastic waves after their propagation through the Earth. Elastic waves can be divided in two groups: body and surface waves. The latter are the most energetic. However, body waves are more commonly used for seismic imaging while surface waves are usually considered as coherent noise.

Surface waves can be used to reconstruct the properties of the near surface (engineering, archaeological and environmental studies) and the Earth's deep interior (global seismology). Current surface wave inversion techniques derive local 1D property profiles, mainly by analysing wave properties from picked dispersion curves. Lateral variations are difficult to be handled with this approach. 2D surface wave imaging is thus needed.

A more general scheme to process seismic data is Full Waveform Inversion (FWI). In theory, FWI can be used to explain the complete wave propagation recorded in seismograms. In the presence of energetic surface waves, FWI needs a very accurate initial velocity model for preventing local minima problems.

This PhD thesis proposes an alternative methodology for waveform inversion applied to surface waves. It has been named windowed-Amplitude Waveform Inversion (w-AWI). It is similar to FWI, but formulated in the $f-k$ domain. The functional measures the misfit between the absolute value of $f-k$ spectra for computed and recorded data. With the absolute value, the phase information is partially destroyed. This is compensated for by considering local spatial windowing before computing $f-k$ spectra. As in FWI, waveform modelling is used. Local windowing also exists in classical analysis of dispersion curves, but here dispersion curve picking is not needed.

We have implemented the novel methodology in 2D elastic isotropic media. For numerical modelling, we use a finite-difference approach discretised in curved grids such that smoothly curved topography and interfaces can be handled. We have paid a particular attention to the free-surface conditions. As in classical FWI, local optimisation techniques can be used. The gradient of the objective function is computed using the adjoint-state method and then used in a quasi-Newton approach.

We have analysed w-AWI through synthetic tests comparing results to FWI, with a particular attention to low frequency data content. These tests show that w-AWI is more robust than FWI with respect to the choice of the initial velocity model. We have found that the length of the spatial windowing has a direct influence on the shape of the misfit functional: this additional parameter is an important element and can vary to help for converging towards the global minimum. We show how to estimate the source wavelet. Furthermore, we have applied the classical 1D surface wave inversion, FWI and the modified technique to a real data set acquired in an ultra-shallow high-velocity contrast environment. The w-AWI approach reveals some interesting features, even if further work would be needed to improve the results.

Keywords: surface waves, inverse problem, seismic modelling, waveform inversion.

Résumé

Dans le cadre de l'exploration pétrolière, l'imagerie sismique est largement utilisée pour caractériser les premiers kilomètres de l'intérieur de la Terre. Les propriétés de la sub-surface (vitesse, densité, ...) modifient la propagation des ondes élastiques. Ces ondes peuvent être séparées en deux groupes : les ondes de volume et les ondes de surface. Les dernières sont les plus énergétiques ; cependant les ondes de volume sont traditionnellement utilisées pour l'imagerie sismique pétrolière, tandis que les ondes de surface sont généralement considérées comme du bruit cohérent.

Les ondes de surface sont utilisées dans d'autres contextes pour reconstruire des propriétés de la proche sub-surface (pour des études d'ingénierie, archéologie, environnement) et de l'intérieur profond de la Terre (en sismologie globale). Les techniques communes d'analyse des ondes de surface se basent sur l'inversion des courbes de dispersion et estiment des profils des propriétés localement 1D. Les variations latérales sont difficiles à traiter avec cette approche. Une technique d'imagerie 2D avec des ondes de surface est donc nécessaire.

L'inversion des formes d'onde complète (Full Waveform Inversion, FWI) est un schéma général pour l'imagerie. En principe, cette approche peut être utilisée pour expliquer la propagation de toutes les ondes sismiques enregistrées par les sismogrammes. Mais en présence d'ondes de surface, la FWI a besoin d'un modèle de vitesse initial très précis pour s'affranchir des problèmes de minimaux locaux de la fonction objective.

Nous proposons une méthodologie alternative pour l'inversion des ondes de surface. Elle a été appellée ici comme "windowed-Amplitude Waveform Inversion" (w-AWI). Cette méthodologie est très proche de la FWI, mais elle est formulée dans le domaine de Fourier $f-k$. La fonction objective mesure le résidu au sens des moindres carrés entre les valeurs absolues des transformées $f-k$ des données calculées et des données enregistrées. Avec la valeur absolue, l'information de phase est partiellement détruite. Cela est compensé par le fenêtrage spatial appliqué avant de calculer les transformées $f-k$. Comme dans la FWI, la modélisation complète de la forme d'onde est utilisée. Le fenêtrage local existe aussi dans l'analyse classique des courbes de dispersion, mais dans la w-AWI les courbes de dispersion ne sont pas pointées.

Nous considérons des milieux 2D élastiques isotropes. Pour la modélisation numérique, nous utilisons une approche de différences finies discrétisée sur des grilles curvilignes (topographie et interfaces courbes peuvent être considérées). Une attention particulière a porté sur les conditions de la surface libre. Comme dans la FWI classique, des techniques d'optimisation locales sont utilisées. Le gradient de la fonction objective est calculé selon la méthode de l'état adjoint et est ensuite introduit dans une approche quasi-Newton.

Nous analysons la w-AWI avec des tests synthétiques en comparant les résultats à ceux de la FWI. Nous proposons une analyse du rôle du contenu basse fréquence des données. Ces tests montrent que la w-AWI est plus robuste que la FWI en ce qui concerne le choix du modèle de vitesse initial. Nous avons constaté que la longueur du fenêtrage spatial a une influence directe sur la forme de la fonction objective : ce paramètre supplémentaire

est un élément important et peut être modifié pour aider à converger vers le minimum global. Nous montrons aussi comment estimer l'ondelette source. Enfin, nous comparons trois approches sur un jeu de données réelles : l'inversion classique 1D des ondes de surface, la FWI et la technique modifiée. Le jeu de données réelles est acquis dans contexte d'imagerie des premiers mètres, avec de forts contrastes de vitesse. L'approche w-AWI révèle des caractéristiques intéressantes, même si des travaux supplémentaires seraient nécessaires pour affiner les résultats.

Mots clés : ondes de surface, problème inverse, modélisation sismique, inversion des formes d'onde.

Contents

1 Introduction **13**

1.1 Surface-wave properties 16

1.2 1D inversion of surface waves 18

1.3 Full Waveform Inversion 19

1.4 Research objectives 21

1.5 Thesis outline 22

2 Elastic wave modelling in curved 2D isotropic media **25**

2.1 Introduction 27

2.2 Wave propagation in rectangular domains 29

2.2.1 Velocity-stress formulation 29

2.2.2 Free-surface conditions 31

2.2.3 Non-reflective boundary conditions 33

2.2.4 Source term 36

2.2.5 Numerical results 36

2.3 Wave propagation in curved domains 40

2.3.1 Chain-rule approach 40

2.3.2 Tensorial approach 42

2.3.3 Computational cost 44

2.3.4 Free-surface conditions 46

2.3.5 Non-reflective boundary conditions 47

2.3.6 Source term 47

2.3.7 Numerical results 48

2.4 Conclusions 55

3 Inversion of surface waves **59**

3.1 Introduction 61

3.2 Inverse problem theory 62

3.2.1 Inverse problem solution 63

3.2.2 Adjoint-state method 65

3.2.3 Strategy for updating the parameters 65

3.3 Surface wave analysis (SWA) 67

3.3.1 Processing workflow for dispersion curve extraction 68

3.3.2 Inversion of dispersion curves 71

3.3.3 Conclusions . . . 85
3.4 Full waveform inversion (FWI) . . . 86
3.4.1 Gradient computation . . . 86
3.4.2 Source wavelet estimation . . . 88
3.4.3 Resolution and surface waves . . . 90
3.4.4 Alternative objective functions . . . 92
3.4.5 Conclusions . . . 93
3.5 Discussion . . . 94

4 An alternative approach for inversion of surface waves 95
4.1 Introduction . . . 97
4.2 The windowed amplitude waveform inversion (w-AWI) . . . 99
4.2.1 Model update and objective function . . . 99
4.2.2 Gradient computation . . . 103
4.2.3 Source wavelet estimation . . . 104
4.2.4 Gradient regularisation and preconditioning . . . 105
4.3 Analysis of w-AWI . . . 106
4.3.1 Objective function shape and local minima . . . 107
4.3.2 Cycle skipping . . . 110
4.3.3 Width of spatial windows . . . 117
4.3.4 Adjoint source and gradient . . . 123
4.3.5 Low frequencies . . . 124
4.4 Inversion of synthetic data . . . 129
4.4.1 No a priori information and known source wavelet . . . 129
4.4.2 Inversion with unknown source wavelet . . . 141
4.4.3 Multi-parameter inversion . . . 147
4.4.4 Inversion with low frequencies . . . 152
4.5 Conclusions . . . 165

5 Real data application 167
5.1 Introduction . . . 169
5.2 Seismic dataset . . . 171
5.2.1 Preprocessing . . . 172
5.2.2 Initial velocity model . . . 174
5.3 Seismic inversion . . . 174
5.3.1 SWA . . . 175
5.3.2 FWI and w-AWI . . . 177
5.4 Discussion . . . 189
5.5 Conclusion . . . 194

6 Conclusions and Perspectives **195**
6.1 Conclusions 195
6.1.1 Waveform modelling 195
6.1.2 Analysis of inversion techniques 196
6.1.3 A novel approach 197
6.2 Perspectives 198
6.2.1 Multi-scale strategies 198
6.2.2 Include more of the Earth's physics in modelling 199
6.2.3 Extension to 3D 199

A Residual source for the adjoint-state variable in w-AWI **201**

List of Figures **203**

Bibliography **215**

Chapter 1

Introduction

Contents

1.1 Surface-wave properties 16
1.2 1D inversion of surface waves 18
1.3 Full Waveform Inversion 19
1.4 Research objectives 21
1.5 Thesis outline 22

The characterisation of the subsurface using seismic techniques consists of recording a propagated wavefield, determining the properties of the propagation, and using the inferred properties for an imaging or inversion procedure. This process leads to a quantitative estimation of the geometrical characteristics and the dynamic behaviour of the parameters that have influenced the wave propagation. The quantitative estimation can be achieved by analysing the propagation of different types of waves (body and surface waves) recorded with different techniques and acquisition geometries and their conversion phenomena (reflection, refraction, diffraction, dispersion).

In the context of hydrocarbon exploration, the traditional seismic imaging techniques are based on body-wave analysis (in particular P-wave reflections) because these spread energy into the Earth's interior and carry interesting information about the deep subsurface. On the contrary, surface waves (often referred to as ground-roll) are usually considered as coherent noise that must be removed or attenuated, since they blur target reflections from deeper depths. Figure 1.1 shows an example of common shot gathers with highly-energetic surface waves masking reflections at short offsets.

Conventional methods for suppressing surface waves exploit differences between body and surface waves like their apparent velocity (by $f-k$ filtering (Yilmaz, 2001) or stacking-based analysis (Ulrych et al., 1999)), their polarisation (de Franco and Musacchio, 2001), or both (Donno et al., 2008). Stacking-based analysis consists of separating signal from noise by selecting an appropriate mask to Radon transformed data. However, when data is poorly spatially sampled, spatial aliasing might prevent an efficient $f-k$ filtering. In such a case, singular value decomposition (SVD) can be used used to estimate and

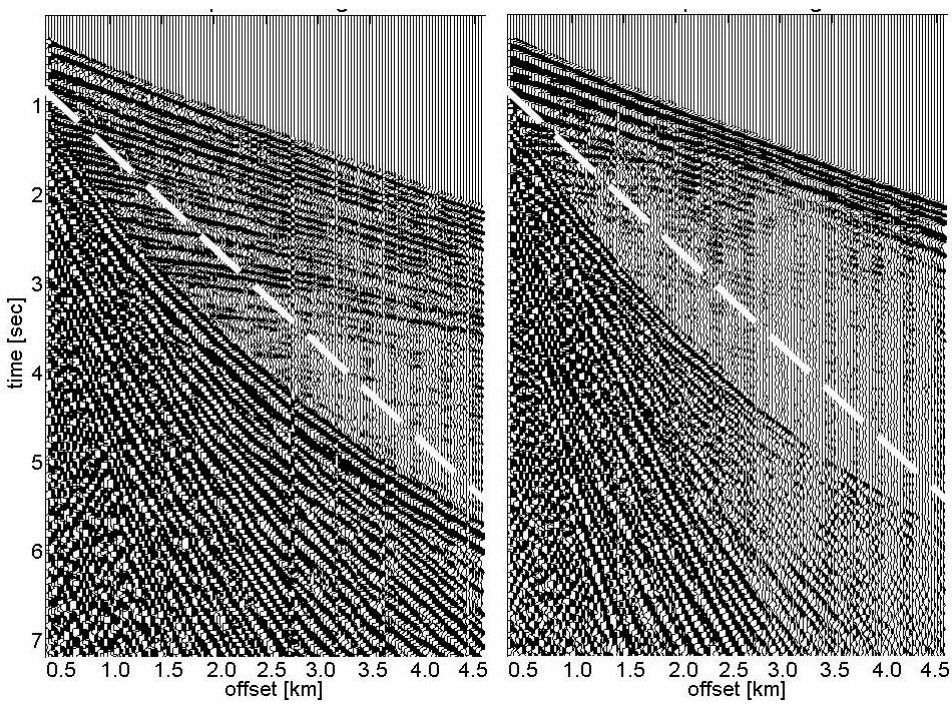


Figure 1.1: Land seismic shot gathers: radial (left) and vertical (right) components. Surface waves are below the white dashed line.

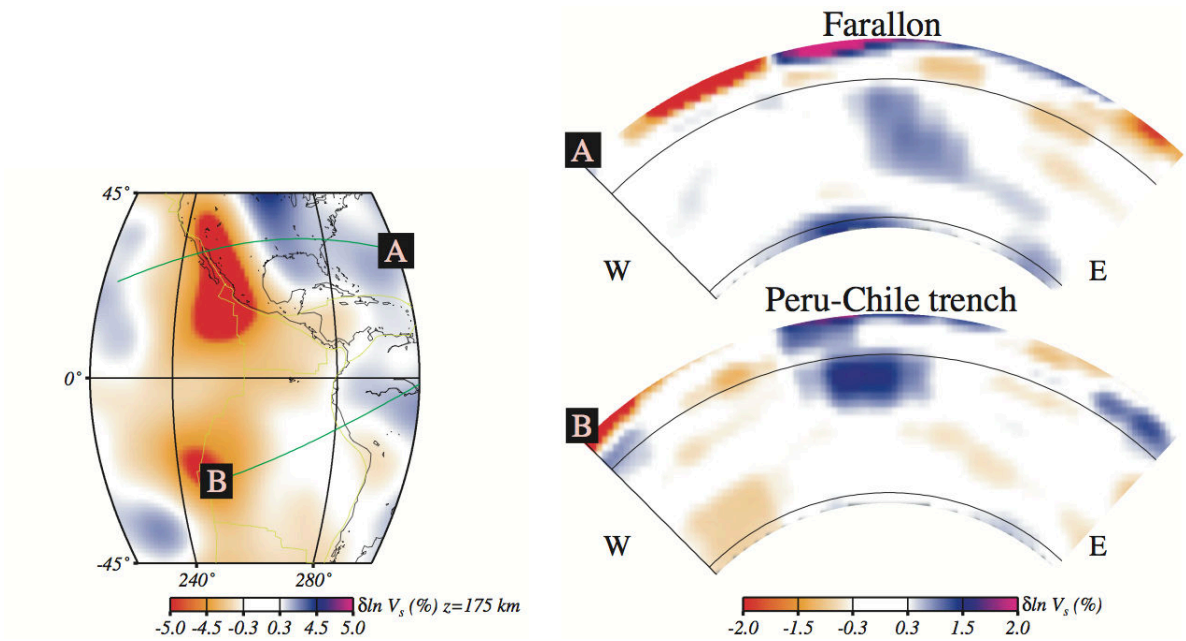


Figure 1.2: Characterisation of the subduction zone under America proposed by Mégnin and Romanowicz (2000). Body, surface and higher-mode waveforms have been analysed to obtain this mantle image.

separate interfering polarised wavefields (de Franco and Musacchio, 2001).

Surface waves play a curious role in seismic methods. While oil exploration handles them as noise, solid Earth seismology have been employing them for characterising the crust and the upper mantle structure during the last 60 years (Aki, 1960b,a; Dorman et al., 1960; Kanamori, 1970). These waves can travel around the Earth for a long time, even passing through the same location several times after large earthquakes. Lower frequencies ($\ll$ 1 Hz) propagate deeper and travel faster while higher frequencies usually provide coverage of the upper mantle. Romanowicz and Durek (2000) have reviewed the studies and applications of surface wave dispersion carried out during the second half of the 20th century in the context of solid Earth seismology. As an example, recordings of long period surface waves (also called mantle waves) were used by Woodhouse and Dziewonski (1984) to obtain the first 3D Vs characterisation of the upper mantle. Mégnin and Romanowicz (2000) derived a 3D Vs model of the whole mantle from the information analysis of body, surface and higher-mode waveforms (Figure 1.2). Moreover, in Godinho Ferreira (2005), lateral variations have been considered by surface wave modelling with ray theory.

There are other applications of surface waves. At a very small scale, ultrasonic Rayleigh waves have been used in non-destructive testing (NDT) for material characterisation. In this context, mechanical and structural properties are estimated to discover, for example, surface defects or the presence of cracking (Thompson and Chimenti, 1997). At an intermediate scale, surface waves are employed in geotechnical engineering to identify mechanical properties of the shallow subsurface (Stokoe II et al., 1988).

Only recently, the oil exploration industry has increased its interest in the large

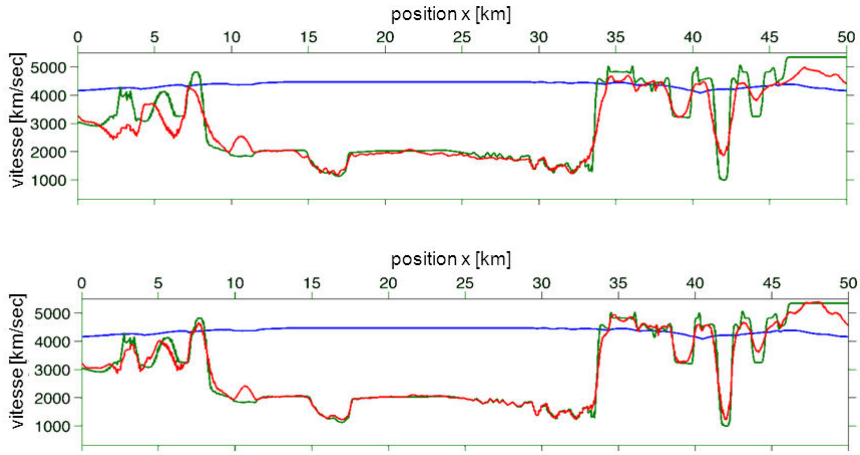


Figure 1.3: Example of tomography result (here a depth slice) extracted from Taillandier et al. (2009). The resulting velocity model after inversion (in red) better fits with the real model (in green) when the near surface is known (bottom plot) versus the case of unknown near surface (top plot). The initial velocity before inversion is in blue.

amount of information contained in surface waves. About 67% of the energy generated by a vertical source at the air-ground interface of a homogeneous half-space propagates in the form of surface waves (Lamb, 1904; Richart et al., 1970). This is a very strong reason for exploiting them.

For hydrocarbon exploration, the near surface is often a major source of wavefield distortion in land acquisition. The quality of the image at the exploration target level is reduced because of near surface effects. The types and properties of surface waves depend on the elastic parameters of the shallow subsurface. Lateral variations and heterogeneities strongly influence surface wave propagation. Therefore, surface wave properties can be inverted to provide a near-surface model. In seismic exploration, the proper knowledge of the near surface would allow improving static corrections (Mari, 1984) and model building in depth. Figure 1.3 shows an example of the improvements in depth of tomography when the first 100 m near the surface are known. Moreover, after using the information carried by surface waves, they can be removed before further processing by using tailored (with the inferred properties) filtering methods.

The objective of this PhD thesis is to investigate an innovative technique for near surface characterisation using surface waves. In this Chapter, we first analyse surface wave properties and then we review currently used techniques for near-surface characterisation using surface waves as well as the limitations of such methods. Finally, we present the outline of this thesis.

1.1 Surface-wave properties

The main properties of surface waves have to be reviewed to better understand surface-wave inversion methods. The properties of waves created along a free surface were first

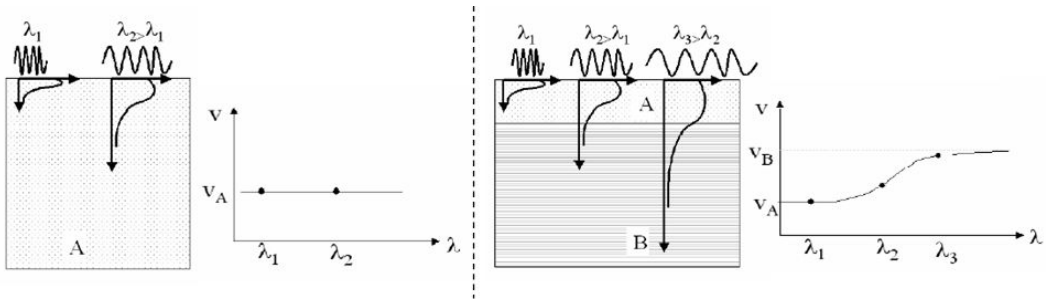


Figure 1.4: In homogeneous media (left), there is no geometric dispersion, while in heterogeneous media (right), each wavelength propagates with a different velocity (extract adapted from Strobbia (2002)).

investigated by Rayleigh (1885) in elastic isotropic media considering P-SV wave propagation, as well as by Love (1911) considering SH wave propagation. This particular group of waves can be described in terms of their propagation velocity, energy and particle motion.

Surface waves propagate parallel to the Earth's surface. Their amplitude decreases exponentially with the depth while at the surface their amplitude decreases with the square root of the distance from the source r (as for body waves in 2D). Thus, most of their energy propagates in a shallow zone with depth (roughly) above the shear wavelength λ_s for a considered frequency-velocity pair (Ewing et al., 1957). They propagate at slow velocities compared to body waves. For these reasons, surface waves are energetic and they mask body waves at their arrival at the geophones. A property of Rayleigh waves is its polarisation: their particle motion observed at the surface describes a retrograde ellipse in the vertical plane that contains the direction of propagation.

The propagation of surface waves in vertically heterogeneous media shows dispersive behaviour (Thomson, 1950; Haskell, 1953). Dispersion is associated with propagation of different frequencies at different phase velocities. In particular, surface waves are characterised by the geometric dispersion which depends on the survey geometry and the properties of the evaluated model, in opposition to the intrinsic dispersion of materials. Since the penetration depth of surface waves depends on the wavelength, the high frequencies (short wavelengths) usually propagate in shallow layers and their velocity depends on the shallow properties, while the low frequencies (long wavelengths) propagate in thicker layers and their velocity is influenced by deeper zone properties. In a homogeneous medium (Figure 1.4, left), the different wavelengths "sample" different depths of the subsurface, but being always the same medium, all the wavelengths have the same velocity. Such velocity, in the case of Rayleigh waves, ranges from 0.87 to 0.96 times the S-wave velocity (the maximum being reached in media with Poisson ratio equal to 0.5), while Love waves only can exist in the presence of heterogeneities (Love, 1911). If the medium is vertically heterogeneous (Figure 1.4, right), each wavelength propagates with different phase velocities. Hence, the velocity of propagation is strongly frequency-dependent and is function of the geometric distribution of subsurface properties.

The relation between the surface wave phase velocities and the frequency is represented by dispersion curves (Figure 1.5). Dispersion curves describe fundamental and higher modes of surface wave propagation (Strobbia, 2002). The fundamental mode com-

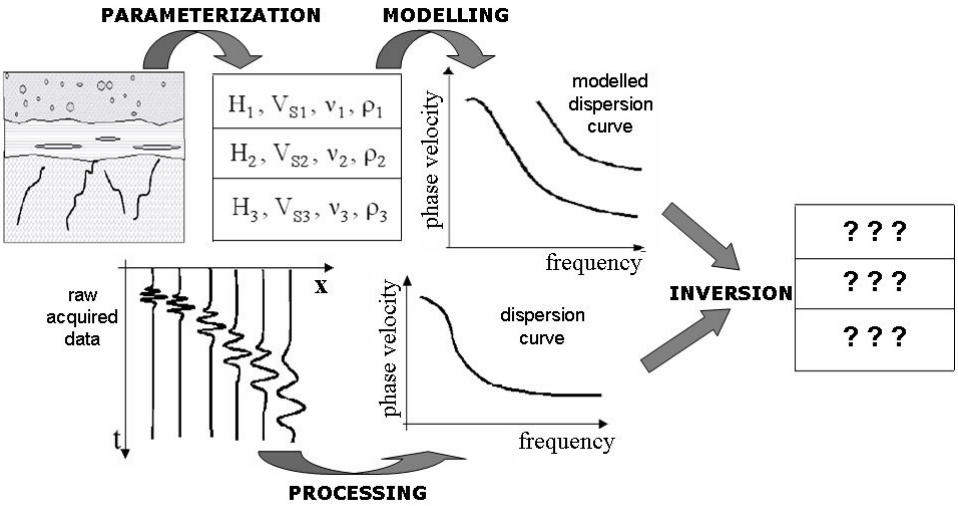


Figure 1.5: Scheme of surface-wave methods (extract adapted from Strobbia (2002)).

monly (but not always) corresponds to the most energetic one and identifies the slowest phase velocity for each frequency. Higher modes usually propagate faster than the fundamental mode but without exceeding the maximum shear-wave velocity in the model. Higher modes exist only above a certain frequency value (cut-off frequency).

1.2 1D inversion of surface waves

The geometric dispersion of surface waves is the principal property behind classical surface-wave analysis. Surface-wave methods consist of three phases (Figure 1.5, adapted from Strobbia (2002)): modelling of the surface-waves, processing of the real seismic data to estimate the dispersive characteristics at a site, and then inversion of these data to estimate the subsoil properties.

Surface-wave modelling commonly consists of the computation of dispersion curves. The subsurface is represented as a 1D profile with homogeneous elastic layers characterised by four parameters: thickness h, mass density ρ, and two mechanical parameters like, for example shear-wave Vs and compressional-wave velocity Vp. The 1D model is the most widely used in seismological and characterisation applications.

The dispersion curve can be experimentally measured from raw acquired data. Many methods have been proposed for the processing and extraction of dispersion curves. The most popular method consists of picking energy maxima in the frequency-wavenumber $(f - k)$ spectrum of data (Nolet and Panza, 1976; Tselentis and Delis, 1998), and then computing the phase velocities as $v = f/k$. The picked dispersion curves can be used to numerically estimate the subsoil properties by solving an inverse problem. After inversion, modelled dispersion curve should fit the curves picked from experimental data. Surface-wave inverse problem is non-linear and often mix-determined, because the information about the shallow layers is carried by all the frequencies, while that concerning deeper layers is carried by the low frequencies only.

Socco et al. (2010) presents a comprehensive review of the surface-wave inversion methods, mainly in the context of exploration geophysics and geo-technical engineering. After the introduction of the spectral analysis of surface waves (SASW) by Nazarian and Stokoe II (1984), higher modes were rapidly incorporated to such analysis (Gabriels et al., 1987). Furthermore, as larger acquisition capabilities were achieved, Park et al. (1999) and Xia et al. (1999) proposed the multichannel analysis of surface waves (MASW). In Strobbia (2002) and Socco and Strobbia (2004), dispersion curves plotted in the frequency-velocity ($f - v$) domain are used to retrieve plane-layered Vs and h profiles. Higher modes are incorporated in inversion to improve resolution and coverage (like in seismology) because they usually penetrate deeper with smaller wavelengths (Xia et al., 2003; Luo et al., 2007). Attenuation properties can also be incorporated in inversion (Lai, 1998). In this context, new misfit functionals and new ways of handling lateral variations have been recently proposed (Socco et al., 2009; Maraschini et al., 2010; Maraschini and Foti, 2010; Bergamo et al., 2012; Samyn et al., 2013).

The main limitation of these methods is that they only allow retrieving locally 1D models. In practice, it is also efficient because synthetic dispersion curves can be obtained by solving a well-known eigenproblem (Thomson, 1950; Haskell, 1953; Hisada, 1995; Lai, 1998). However, if the 1D model hypothesis is not verified at the scale of observation (which is often the case), then lateral variations can affect the results.

Dispersion curve picking can be a cumbersome task particularly difficult to be achieved in large seismic datasets. The fundamental mode is not always the most energetic one and sometimes only apparent dispersion curves can be obtained (Socco and Strobbia, 2004). Moreover, the computation of dispersion curves in 2D media is not clear. Several investigations have resulted in improvements regarding the high frequency stability of the original numerical solutions presented by Thomson (1950) and Haskell (1953). However, at least 2D modelling is necessary to handle lateral variations especially present at certain imaging scales.

1.3 Full Waveform Inversion

To overcome the limitations of 1D dispersion curve techniques, inversion of waveforms could be considered. Full waveform inversion (FWI) is a technique potentially able to explain even dispersive surface waves present in seismic gathers. In this technique, the objective is to explain the entire seismogram by simulating the full wavefield propagation in a 3D elastic model. As a result, such elastic model characterises the real medium where seismic data has been acquired. Figure 1.6 shows differences between a snapshot of acoustic (top) and elastic (bottom) wave propagation. Surface waves are generated at the elastic free surface, and they strongly complicate the interpretation of the seismic gathers.

Originally investigated by Lailly (1983), Tarantola (1984) and Mora (1987), FWI is a data-fitting procedure, often solved with a local optimisation method. The usual fitting criterium consists of the least-squares misfit between recorded and modelled data. Depending on the objectives, the forward problem has been formulated within the acoustic (Tarantola, 1984) or the elastic (Tarantola, 1986; Mora, 1987) wave equations, either in

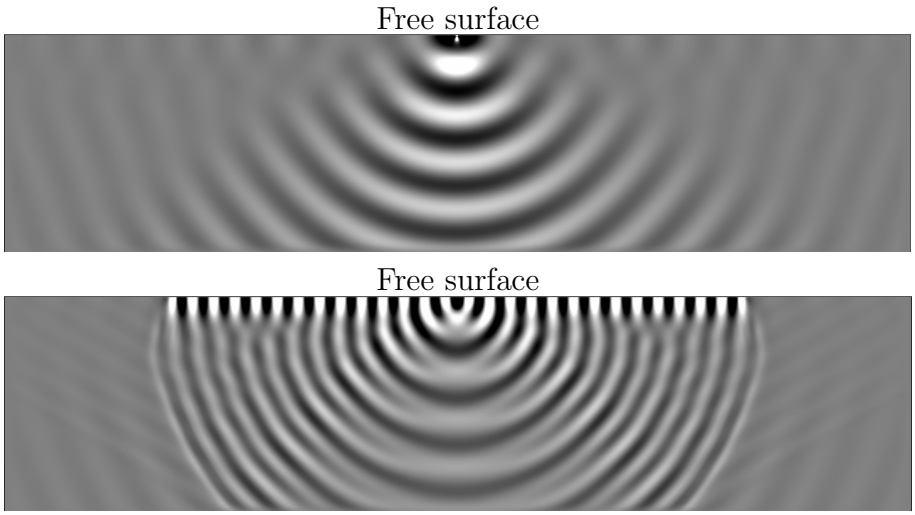


Figure 1.6: Snapshot of the wave propagation in acoustic (top) and elastic (bottom) homogeneous media with a free surface. Rayleigh waves are originated at the free-surface contact in the P-SV elastic case. A monochromatic wave source has been employed.

the time domain (Virieux, 1986; Gauthier et al., 1986) or in the frequency domain (Pratt and Chapman, 1990; Pratt et al., 1996). As convergence strategy, gradient methods as well as Newton type methods can be implemented (Pratt et al., 1998).

This inverse problem is ill-posed in the sense of Hadamard (1902). Besides, wave amplitudes are non-linear with respect to the Earth properties. Many studies has been carried out searching for mitigation of non-linear effects. Bunks et al. (1995) observed that at low frequencies, the objective function shape is smoother, therefore the initial model has more chances of belonging to the global minimum basin. Relying on frequency domain modelling, Sirgue and Pratt (2004) proposed a strategy for choosing the correct frequency to ensure the continuity of wavenumber coverage. Multiscale strategies were also adopted either in the case of 2D acoustic (Sirgue, 2003) and 2D elastic modelling (Brossier et al., 2009).

Local minima effects in FWI are enhanced by the energetic surface waves. The choice of initial velocity model is restrictively constrained in the presence of high frequency surface waves (Gélis et al., 2007). In order to avoid falling into a secondary minimum, the initial model should provide data that is in general accurate within less than half the wavelength. The usual techniques to select the initial velocity model consists of searching a (smooth) model that explains the kinematics of the data, for example using ray-based inversion methods and/or tomography. Furthermore, time-damping strategies have been proposed either to reduce strong non-linearities introduced by surface waves in the first iterations (Brossier et al., 2008; Romdhane et al., 2011), or to focus inversion onto the early arrivals (Sheng et al., 2006; Sears et al., 2008). However, there are not yet clear strategies in the context of FWI with surface waves.

Figure 1.7 show an example of application of FWI to a real land dataset with surface waves. In this case, acoustic FWI was used to find the model explaining diving waves propagated in a sedimentary context. Moreover, a recent hybrid approach has shown

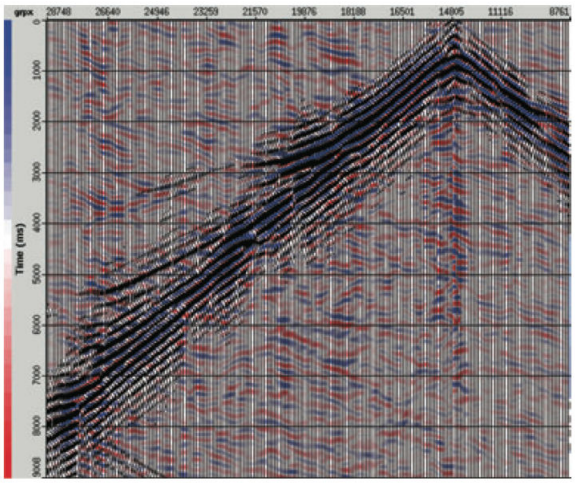


Figure 1.7: A real data set explained by FWI (Plessix et al., 2012). Acoustic FWI was used to find the model explaining diving waves propagated in a sedimentary context.

the potential of including near surface data from dispersion curve inversion and FWI for velocity model building (Droujinine et al., 2012).

We can now raise the following question: is the combination of time-damping and multiscale approaches, besides the choice of a smooth initial model, enough to handle highly-energetic surface waves in FWI?

1.4 Research objectives

In this thesis, the main objective consists of proposing an alternative formulation for inversion of surface waves in 2D media. Based on the analysis of two approaches potentially able to deal with surface waves at seismic exploration scale, we aim at combining their advantages in an intermediary approach that overcomes some of the identified drawbacks.

In this work, we have focused on 2D isotropic media considering that the passage to 3D anisotropic media is less difficult. Multiparameter inversion has not been deeply studied here and most of our efforts have been concentrated on the inversion of Lamé parameters (and density in a couple of examples).

Three specific tasks are addressed in order to accomplish our research objective. The first one is programming a wave propagation code able to handle free surface, topography and curved interfaces. The modelling tool should be simple with low computational cost, but should be capable of handling waveform propagation in complex models. Attenuating boundary conditions must also be implemented to make this code suited to be used in waveform modelling and iterative seismic inversion.

The second task is to better understand the theoretical and practical aspects of SWA and FWI. This leads to the identification of the desired properties that can be used in a novel inversion approach.

Finally, we propose and analyse the novel inversion approach. We compare the charac-

teristics and the results provided by FWI to the novel approach. Synthetic data inversion is used to evaluate the role of the initial velocity model and the final resolution. An application to real data is also considered.

1.5 Thesis outline

This thesis is organised into five chapters and one appendix. Chapter 2 deals with the wave propagation modelling. Chapter 3 is a review of the theory of 1D surface-wave inversion methods and FWI. The proposed approach for inversion of surface waves is described in Chapter 4, where synthetic examples are shown. Finally, Chapter 5 presents an application to real data. In the next few paragraphs, we give more details on the methods and results presented in each chapter.

In Chapter 2, we present the formulation of the wave propagation modelling. We have implemented the classical velocity-stress formulation for 2D elastic isotropic media presented by Virieux (1986), but discretised in a different staggered grid and implemented in a curved domain. Two implementations of free-surface conditions (Levander, 1988; Xu et al., 2007) have been compared to an analytical solution in order to choose the most accurate one. We have also analysed the theoretical computational cost of two approaches able to handle curved interfaces and topography: the chain rule (Zhang and Chen, 2006) and the tensorial approach (Komatitsch et al., 1996). We have found that the chain rule approach is better suited if less than sixth-order finite-difference discretisation is employed. Our code is second order in time and space and is programmed in Fortran 90. The codes are parallelised using a combination of MPI and OpenMP. We have employed one of the collocation grids described by Saenger et al. (2000). The attenuating boundary conditions proposed by Komatitsch and Martin (2007) were implemented.

In Chapter 3, we study the theory of seismic inversion in the context of SWA and FWI. We present original dispersion curve inversion examples as well as others considered in literature for benchmarking (Luo et al., 2007). We have programmed the inversion code in Matlab following the Marquardt-Levenberg approach presented by Xia et al. (1999). To compute dispersion curves in plane-layered media, we have employed the forward modelling code programmed by Lai (1998). Dispersion curves are extracted from shot gathers under our own processing workflow. On the side of FWI, we have studied gradient and Newton convergence strategies (Pratt et al., 1998). The gradient computation is investigated within the adjoint-state method (Plessix, 2006). We give explicit formula for computing the gradient with respect to Lamé parameters and density within the 2D isotropic velocity-stress formulation employed in Chapter 2. We have also analysed the wavelet estimation either as solution of a linear inverse problem (Pratt, 1999) and as part of the adjoint-state method (Tarantola, 1986).

In Chapter 4, we start by presenting the theory of the proposed inversion approach, which we call "windowed-Amplitude Waveform Inversion" (w-AWI). This approach is based on a modification of the classical FWI objective function. We develop the formulation for the residual source computation (using the adjoint-state method) and the wavelet estimation. We use a preconditioning approach proposed in the context of resistivity imaging (Plessix and Mulder, 2008). As convergence strategy, we employ a quasi-

Newton approach (Nocedal, 1980). We have implemented FWI and w-AWI in Fortran 90 including MPI and OpenMP. Several numerical tests are presented with the objective of analysing the characteristics of w-AWI in comparison to FWI. We highlight the fact that w-AWI has the same global minimum of FWI, but it is less constrained with respect to the choice of initial velocity model. We demonstrate through examples that the global minimum basin is wider in w-AWI. Finally, we support the numerical analysis by presenting synthetic inversion results. We use two exact velocity models: one inspired by a real model (Deidda and Balia, 2001) and also a layered model with curved interfaces.

In Chapter 5, we show an application to ultra-shallow real seismic data. The data were acquired on the ground model presented by Deidda and Balia (2001). Preprocessing applied to the data set includes amplitude correction, frequency filtering and noise attenuation. On the side of SWA, we invert a pseudo 2D Vs model from picked dispersion curves. Then, waveform inversion (FWI and w-AWI) is applied to the dataset. Different frequency bands are evaluated: low frequencies, high frequencies and full spectrum. The wavelet source is numerically estimated both with FWI and w-AWI. The obtained results are encouraging so that these techniques should be considered for further analysis.

Chapter 2

Elastic wave modelling in curved 2D isotropic media

Contents

2.1 Introduction **27**

2.2 Wave propagation in rectangular domains **29**

2.2.1 Velocity-stress formulation 29

2.2.2 Free-surface conditions 31

2.2.3 Non-reflective boundary conditions 33

2.2.4 Source term 36

2.2.5 Numerical results 36

2.3 Wave propagation in curved domains **40**

2.3.1 Chain-rule approach 40

2.3.2 Tensorial approach 42

2.3.3 Computational cost 44

2.3.4 Free-surface conditions 46

2.3.5 Non-reflective boundary conditions 47

2.3.6 Source term 47

2.3.7 Numerical results 48

2.4 Conclusions **55**

Résumé du chapitre
Dans ce chapitre, nous étudions des méthodes de résolution numérique pour la propagation des ondes dans des milieux élastiques. Nous nous sommes intéressés à prendre en compte des topographies variées mais toujours lisses. Plus précisément, nous analysons des approches par différences finies sur des grilles carrées ou courbées. Les conditions aux bords à la surface libre ainsi que les conditions absorbantes sont regardées avec attention. Enfin, nous présentons des tests numériques pour comparer les solutions obtenues à des solutions analytiques. Un exemple de modélisation dans un milieu complexe est discuté à la fin du chapitre.

Dans la partie 2.2, nous introduisons d'abord les équations de propagation des ondes dans les milieux élastiques (Aki and Richards, 1980). Nous avons choisi de résoudre ces équations pour les milieux 2D isotropes en utilisant la formulation vitesse-contrainte présentée par Virieux (1986). Une attention particulière a été portée aux conditions de surface libre. Deux approches ont été évaluées, par comparaison avec une solution analytique, aux grands offsets (distance source-récepteur) : *Stress image method* (Levander, 1988) et *acoustic-elastic approach* (Xu et al., 2007). Grâce aux bonnes performances cinématiques, nous préférons la première approche. Pour les conditions absorbantes, la méthode *Convolutional Perfect Match Layer* présentée par Komatitsch and Martin (2007) a été choisie. Des tests numériques ont été réalisés pour le problème de Lamb (Lamb, 1904), de Garvin (Garvin, 1956) et pour une interface entre deux milieux élastiques (Berg et al., 1994). Des comparaisons avec la solution analytique trouvée par la méthode de Cagniard-de Hoop (de Hoop, 1960) ont montré les très bonnes performances de notre outil de modélisation sur des grilles carrées (corrélation plus grande que 99% et erreur RMS inférieure à 1%).

Dans la partie 2.3, nous analysons deux approches pour la modélisation des ondes sur des grilles courbées. Ces approches sont la *tensorial approach* (Komatitsch et al., 1996) et la *chain rule* (Zhang and Chen, 2006). La méthode *chain rule* sera choisie après évaluation théorique des coûts de calcul numérique. La grille en quinconce doit être modifiée car la grille classique (Virieux, 1986) n'est pas adaptée pour la méthode choisie. Nous utilisons donc la grille en quinconce analysée dans Saenger et al. (2000). Les conditions à la surface libre ainsi qu'aux bords absorbants sont analysées pour les mettre en œuvre correctement sur des grilles courbées. Les mêmes tests réalisés pour des grilles carrées (section 2.2) sont aussi réalisés pour les grilles courbées, mais ici les interfaces sont toutes inclinées. Les résultats montrent que cette approche est également très performante. Cependant, le coût de calcul et la consommation de mémoire sont deux fois plus élevés que pour les grilles carrées. Néanmoins, des topographies faiblement courbées peuvent être prises en compte avec cette approche de modélisation de propagation des ondes élastiques dans des milieux isotropes 2D.

2.1 Introduction

Elastic wave propagation is commonly modelled by numerically solving the differential equations established in the velocity-stress formulation (Virieux, 1986). Several numerical methods have been proposed offering fast and reliable results when computing the propagation of shear and compressional waves in different (structured and unstructured) grids (Saenger et al., 2000; Kaser and Igel, 2001). We mainly consider here finite-difference approaches. Surface waves can be modelled by implementing free-surface conditions on these grids using one of the multiple proposed approaches (Levander, 1988; Robertsson, 1996; Mittet, 2002; Bohlen and Saenger, 2006; Xu et al., 2007). However, elastic wave propagation in realistic media with non-rectilinear (lateral and vertical) interfaces cannot be accurately modelled using classical approaches dedicated to rectangular grids. In this chapter, we revisit the way how to model the elastic wave propagation in 2D isotropic media with curved free surface and curved internal interfaces using finite-difference methods.

Four principal alternatives are described in the literature for accomplishing this goal. They are described as follows.

- The shape of layers (including the free surface) can be taken into account by implementing a rectangular grid and refining the grid in the vicinity of the interface contact (Robertsson, 1996; Moczo et al., 2007). Curved interfaces are approximated by small straight lines. This method is stable and easy to implement. However, some spurious diffractions can be generated from the squared ending of the grid, mostly when curved interfaces are discretised.

- The curved free surface can be implemented in rectangular grids by building fictitious values of the solution near the boundaries as presented in Lombard et al. (2007). In this approach, the curved shape of the surface is implicitly taken into account in the formulation. Internal curved interfaces can be modelled using a similar modelling approach (Lombard and Piraux, 2004). This method accounts for stability; the accuracy does not depend on the relative position of the boundary with respect to the grid. This approach requires defining all the unknown values at the same grid points, and thus only single-grid schemes can be used.

- Unstructured grids can be used for modelling the wave propagation in curved media. Differential operators can be implemented by using finite volume, finite element or spectral element methods (Kaser and Igel, 2001; Kaser and Dumbser, 2006). By implementing these methods on unstructured grids, the wave propagation can be modelled in complex curved media because the mesh can be adapted to a wide range of shapes. However, the accuracy may depend on the quality of the mesh built for each particular medium. Also, these methods are computationally more expensive than finite-difference methods.

- The velocity-stress formulation can be solved in curved grids built to fit the shapes described by the free surface and the internal interfaces. The formulation must be adapted to be solved when the discretised variables are distributed along curvilinear grids. The method we propose here is based on a finite-difference staggered-grid

alternative implemented on curved grids. In the following, we concentrate on this approach and analyse it more carefully.

Some formulations have been proposed with a topography-dependent modelling approach using the chain rule for horizontal derivatives only (Hestholm and Ruud, 1994, 1998, 2001; Tarrass et al., 2011). The vertical grid lines must always be straight. This restriction leads to difficulties in the implementation of more complex media.

A more general solution consists of considering the propagating variables as tensorial fields (McConnell, 1957; Malvern, 1969). The numerical tests conducted by (Komatitsch et al., 1996) show that wave propagation can be computed in general for any curved coordinate system. Nevertheless, we show in this chapter that the splitting of the covariant derivative, necessary for the divergence to be invariant under coordinate transformations, increases the computational cost when the transformed system is not rectilinear.

Differently from the methods mentioned above, we propose here to *apply the chain rule in a completely curved system* (i.e. all axes are curved). This solution is both general and computationally affordable, and allows us to address the problem of a wave propagating in complex media with curved interfaces. This approach is potentially suited for implementations in the spectral domain (Tessmer et al., 1992; Carcione and Wang, 1993; Carcione, 1994), as well as for finite-difference approaches (Zhang, 1997; Zhang and Chen, 2006).

For the numerical implementation of the chain rule, we use a *modified staggered grid* improving both computational cost and accuracy aspects. Free-surface conditions are set in the curved system by defining the normal vector to the surface. We have adapted the theory of the Convolutional Perfect Matching Layer (CPML) to curved staggered grids for avoiding reflective boundaries (Komatitsch and Martin, 2007). For the modelling of explosive and directional sources, we have found a constant which allows us (when numerical dispersion is negligible) to avoid the application of any additional amplitude correction, when the numerical results are compared to the analytical solution. The precision of the approach is tested by comparing the numerical result with the analytical one obtained by the Cagniard-de Hoop solution (de Hoop, 1960) in the horizontal and tilted Lamb's problem (Lamb, 1904), Garvin's problem (Garvin, 1956) and elastic/elastic interface problem (Berg et al., 1994). The results of these tests demonstrate the high accuracy of the implementation. Optimal results are obtained when the angle between the axes at the free surface tends to perpendicularity. As shown through an example, this approach allows us to model the wave propagation in more complex media, revealing its potential as a tool for further applications in models with topography or lateral varying velocities, among others. In very complex media though, curved grids are not well suited because they are rather adapted to smooth interface curvatures.

This chapter is organised as follows. Section 2.2 deals with the wave propagation in rectangular domains by describing the approach for the finite-difference formulation, the modelling of the free-surface, the non-reflective boundary conditions, and the implementation of the source. New aspects are developed in section 2.3 but dealing with the wave propagation in curved domains. The free-surface condition is very important for proper surface wave modelling. Numerical results are presented at the end of each section.

2.2 Wave propagation in rectangular domains

This section describes the so-called velocity-stress finite-difference formulation solved in Cartesian coordinates for the modelling of elastic waves in isotropic media (section 2.2.1). In section 2.2.2, we present two approaches for free-surface modelling. A performance analysis allows us to choose the stress image method (SIM) for our implementation. The CPML theory is described in section 2.2.3. For the implementation of explosive and directional sources (in section 2.2.4), we have found the constant that allows us to directly compare the numerical results to the analytical ones. Finally, we test the performance of the implementation by comparing our numerical results with the analytical solutions available in literature (section 2.2.5).

2.2.1 Velocity-stress formulation

The propagation of waves in the Earth is represented by a 3D elastic anisotropic equation (Aki and Richards, 1980). Such equation can be solved either in the frequency-space domain or in the time-space domain. As surface waves are dispersive waves (Chapter 1), we need to study the full frequency range of the wave propagation. Thus, we need an equation defined in the time domain (this reduces computational costs of full frequency computation). In general, this equation is represented in time t and space $\mathbf{x}$ by

$$\mathcal{C}(\mathbf{m}|\mathbf{x},t)\mathbf{u}(\mathbf{x},t) - \mathbf{s}(\mathbf{x},t) = \mathbf{0}, \tag{2.1}$$

where $\mathcal{C}(\mathbf{m}|\mathbf{x},t)$ is the matrix with partial differentiators and elastic model parameters, $\mathbf{u}(\mathbf{x},t)$ represents wave field variables and $\mathbf{s}(\mathbf{x},t)$ represents the wave source components.

In this thesis, we implement a 2D elastic wave propagation modelling because it is computationally more affordable than full 3D modelling. The wave propagation system is formulated as a first-order hyperbolic system (Virieux, 1986) and it is discretised in staggered grids making its implementation to be straightforward. The elastodynamic equations of the velocity-stress formulation (Virieux, 1986) in 2D isotropic media are:

$$\begin{aligned}
\dot{v}_x &= \frac{1}{\rho}(\partial_x\sigma_{xx} + \partial_z\sigma_{xz}) + \varphi_{v_x}, \\
\dot{v}_z &= \frac{1}{\rho}(\partial_x\sigma_{xz} + \partial_z\sigma_{zz}) + \varphi_{v_z}, \\
\dot{\sigma}_{xx} &= (\lambda + 2\mu)\partial_x v_x + \lambda\partial_z v_z + \varphi_{\sigma_{xx}}, \\
\dot{\sigma}_{zz} &= (\lambda + 2\mu)\partial_z v_z + \lambda\partial_x v_x + \varphi_{\sigma_{zz}}, \\
\dot{\sigma}_{xz} &= \mu(\partial_z v_x + \partial_x v_z) + \varphi_{\sigma_{xz}},
\end{aligned} \tag{2.2}$$

where x and z are the Cartesian coordinates; λ, μ and ρ are elastic model parameters (Lamé parameters and density) (part of $\mathcal{C}(\mathbf{m}|\mathbf{x},t)$ in equation 2.1), (v_x, v_z) and $(\sigma_{xx}, \sigma_{zz}, \sigma_{xz})$ are velocity and stress components ($\mathbf{u}(\mathbf{x},t)$ in equation 2.1)) and φ represents the source term ($\mathbf{s}(\mathbf{x},t)$ in equation 2.1)). A dot over a variable denotes time differentiation. The symbols $\partial_x = \partial/\partial x$ and $\partial_z = \partial/\partial z$ represent the partial derivatives for each Cartesian direction. Velocities of different kind of waves are related to the Lamé parameters and density such that, in this kind of media, compressional-wave velocity is expressed as $V_p = \sqrt{(\lambda + 2\mu)/\rho}$, and shear-wave velocity is $V_s = \sqrt{\mu/\rho}$.

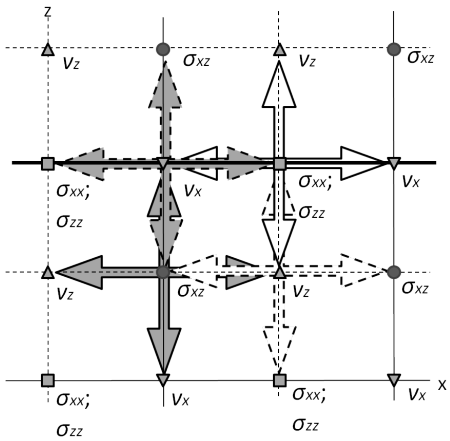


Figure 2.1: Staggered grid with the bold solid line specifying the free-surface boundary. Solid lines identify the principal grid and dashed lines represent the auxiliary staggered grid.

The source term considered in equation 2.2 is equivalent to an infinite line with cylindrical radiation in 3D media. A 2D to 3D amplitude correction can be applied by considering the principles of geometrical spreading: (1) body waves are attenuated as $1/r$ in 3D media and as $1/\sqrt{r}$ in 2D media and (2) surface waves are attenuated as $1/\sqrt{r}$ in 3D media and are not attenuated in 2D media. Hence, the 2D data can be partially corrected by using the so-called $1/\sqrt{t}$ filter (Bleistein, 1986), if the medium is such that it can be considered homogeneous in the y direction (2.5 D). Anisotropy can be modelled (Thomsen, 1986; Tsvankin, 2005), but with the purpose of analysing surface waves, in this study, we have decided to use equation 2.2 which is focused on heterogeneous isotropic media.

For the implementation of equation 2.2 in rectangular grids, we use the staggered grid proposed by Virieux (1986). In such a grid, the discretisation of equations 2.2 can conveniently (in computational terms) be implemented as presented in Figure 2.1.

We use a centred second-order discretisation in time and space. The time and space derivatives of a generic function $f(x_i)$ are discretised as

$$\begin{aligned} \dot{f}(x_i) &\approxeq \frac{1}{\Delta t}(f^{t+\frac{1}{2}}(x_i) - f^{t-\frac{1}{2}}(x_i)), \\ \partial_{x_j} f(x_i) &\approxeq \frac{1}{\Delta x_j}(f(x_1,.,x_{j+\frac{1}{2}},.,x_n) - f(x_1,.,x_{j-\frac{1}{2}},.,x_n)), \end{aligned} \tag{2.3}$$

where Δt and Δx_j represent the time and space differentiation steps. As it can be seen from Figure 2.1, the calculation of a variable is straightforward because its value depends on the centred numerical derivatives related to its four nearest neighbouring values.

The proper spatial sampling of the grid Δx_j is found according to the maximum dispersion condition accepted in each specific simulation, such that

$$\Delta x_j \leq \frac{\lambda_{min}}{n_\lambda} = \frac{V_{min}}{f_{max}} \frac{1}{n_\lambda}, \tag{2.4}$$

where λ_{min} corresponds to the minimum wavelength in the medium, V_{min} to the minimum

velocity in the medium, f_{max} is the maximum frequency of the source; the value n_λ is defined as the number of grid points per wavelength. It is commonly set above 16 points per minimum wavelength in order to perform a correct modelling of waves. Increasing n_λ reduces numerical dispersion (Virieux, 1986; Levander, 1988).

The time step Δt is found by satisfying the Courant condition for numerical stability

$$\Delta t V_{max} \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta z)^2}} \leq 1, \tag{2.5}$$

where V_{max} is the maximum velocity in the medium and $\Delta x = \Delta z = \Delta x_j$ for $j = 1, 2$.

2.2.2 Free-surface conditions

The free surface can be modelled according to the theoretical condition of zero stress for the component normal to the surface (Aki and Richards, 1980; Chapman, 2004). In rectangular grids, this condition can be achieved either via the so-called stress image method (SIM) (Levander, 1988; Robertsson, 1996), or by changing the elastic properties of the medium at the free surface (Mittet, 2002; Xu et al., 2007). The comparison of these methods has quantitatively shown their high accuracy when compared to the analytical solution for distances between the source and receiver below 10 times the shear wavelength (Bohlen and Saenger, 2006; Xu et al., 2007). In this paragraph, we present the formulation for the two approaches, and a comparison between the two methods for long offsets (up to 100 times the shear wavelength). We conclude that the SIM method presents a better performance.

Stress image method

Let $\widehat{n}_i^{fs} = \widehat{z}$ be the vector normal to the free surface. By imposing the zero-stress condition for the component normal to the surface $\sigma_{ji}^{fs}\widehat{n}_i^{fs} = 0$ and by substituting in the set of equations 2.2, we get:

$$\sigma_{ji}^{fs}\widehat{n}_i^{fs} = 0 \Rightarrow \begin{bmatrix} \sigma_{xx} & \sigma_{xz} \\ \sigma_{xz} & \sigma_{zz} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{ll} \sigma_{zz} = 0 \rightarrow & \partial_z v_z = -\frac{\lambda}{\lambda+2\mu}\partial_x v_x, \\ \sigma_{xz} = 0 \rightarrow & \partial_z v_x = -\partial_x v_z. \end{array} \tag{2.6}$$

This definition can be implemented in the finite-difference staggered-grid scheme by declaring the free-surface boundary at the position indicated by the bold line in Figure 2.1. The variables σ_{zz}, σ_{xx} and v_x are the variables at the boundary. Taking equation 2.6 and defining indices (i, j, k) as the discretised x-axis, z-axis and time, if we put the free-surface boundary in $j = 1$, as indicated in Figure 2.1, then the finite-difference scheme for the variables present at this boundary is formulated as

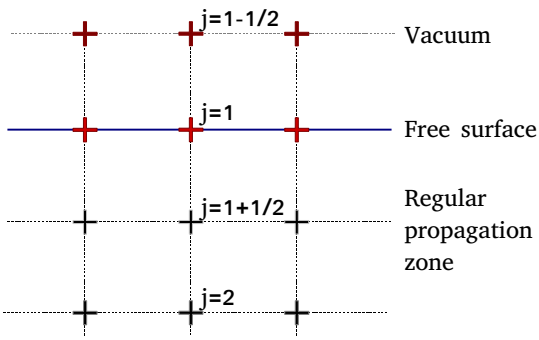


Figure 2.2: Discretisation scheme for the free-surface implementation.

$$
\begin{aligned}
\sigma_{zz}^{k}(i+\frac{1}{2},1) &= 0, \\
\sigma_{xx}^{k}(i+\frac{1}{2},1) &= \sigma_{xx}^{k-1}(i+\frac{1}{2},1) \\
&\quad + \Delta t \frac{4\mu(\lambda+\mu)}{\lambda+2\mu} \frac{v_x^{k-1/2}(i+1,1) - v_x^{k-1/2}(i,1)}{\Delta x}, \\
v_x^{k+1/2}(i,1) &= v_x^{k-1/2}(i,1) + \frac{\Delta t}{\rho} \left\{ \frac{\sigma_{xx}^{k}(i+\frac{1}{2},1) - \sigma_{xx}^{k}(i-\frac{1}{2},1)}{\Delta x} \right. \\
&\quad \left. + \frac{\sigma_{xz}^{k}(i,1+\frac{1}{2}) + \sigma_{xz}^{k}(i,1+\frac{1}{2})}{\Delta z} \right\},
\end{aligned}
\tag{2.7}
$$

where the required value for the vacuum, $\sigma_{xz}^{k}(i,1-\frac{1}{2})$, is found after supposing an odd parity of the stress with respect to the free-surface boundary (i.e. $\sigma_{xz}^{k}(i,1-\frac{1}{2}) = -\sigma_{xz}^{k}(i,1+\frac{1}{2})$). This approach is known as the Stress Image Method (SIM) (Levander, 1988; Robertsson, 1996).

Acoustic-elastic approach

An alternative method for modelling the free surface proposes a change in the elastic properties of the medium at the vacuum-solid contact (Mittet, 2002). An improved version of this method, called acoustic-elastic approach (AEA), consists of changing the Lamé parameters at the free-surface contact under the following rules: $\rho^{fs} = 0.5\rho$, $\lambda^{fs} = 0$ and $\mu^{fs} = \mu$ (Xu et al., 2007).

Accuracy test with respect to offset

The performances of these two methods have been compared with the analytical solution for a homogeneous medium (Berg et al., 1994). We have analysed the numerical dispersion of Rayleigh waves at far offsets, for distances between receiver and source up to 100 times the shear wavelength. In homogeneous media, the Rayleigh-wave velocity is constant and frequency independent (Dunkin, 1965). We use this property to compare the Rayleigh

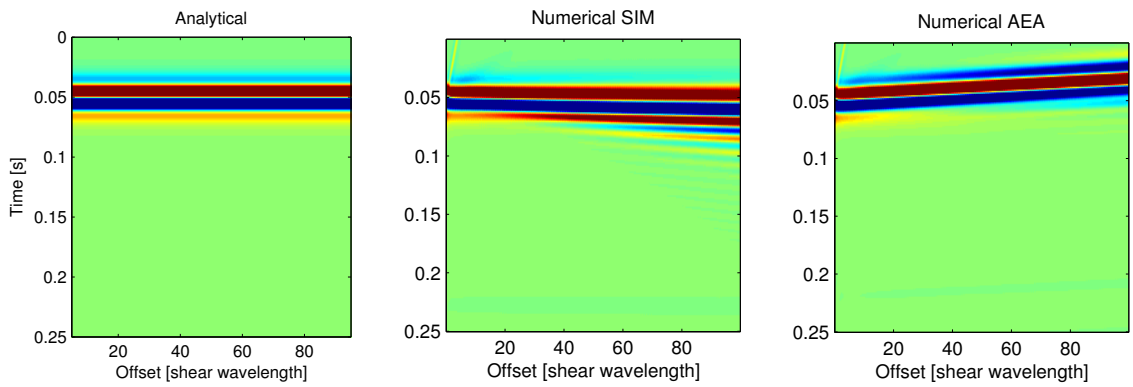


Figure 2.3: Numerical dispersion observed in Rayleigh-wave propagation for offsets up to 100 times the shear wavelength. The SIM (centre) and AEA (right) methods were implemented with $n_\lambda = 30$ points per shear wavelength. They are compared to the analytical solution (left). The same move-out correction has been applied to all the three shot gathers.

wave arrival times for each offset, after move-out correction. In Figure 2.3, we compare the analytical, SIM and AEA solutions. The results show an increasing phase shift for far offsets mainly in the AEA solution. The solution provided by SIM becomes numerically dispersive.

The derivative of the phase with respect to frequency can be used to know the delay of the signal with respect to frequency and offset (Oppenheim et al., 1998). In Figures 2.4 and 2.5, we present the phase in the frequency-offset domain (top left), the derivative of the phase with respect to frequency (top right) and a comparison of analytical and numerical Rayleigh waves for offsets equal to 5 (bottom left) and 33 (bottom right) times the shear wavelength. These results show that the SIM approach presents higher fidelity with the analytical solution for a wider range of offsets than the AEA approach. The SIM method shows a phase shift due to the numerical dispersion, which can be corrected by using more samples per wavelength for the discretisation (Bohlen and Saenger, 2006). The AEA method presents an increasing negative phase delay for far offsets, which cannot be explained by numerical dispersion. As a consequence, we have chosen in this work the direct implementation of the theoretical free-surface conditions as in the SIM method because of its higher accuracy for a variety of offsets.

2.2.3 Non-reflective boundary conditions

We chose to implement the CPML approach for its proven attenuating and non-reflective properties (Komatitsch and Martin, 2007). The formulation proposed by Komatitsch and Martin (2007) improves the behaviour of classical PML (Bérenger, 1994; Collino and Tsogka, 2001) at grazing incidence and is based on an unsplit convolutional approach. In literature, it is recommended that the length of the attenuation zone is at least 0.5 to 1 times the considered wavelength. Therefore, the number of needed grid points in the attenuation zone depends on the number of grid points per wavelength.

The CPML approach consists of substituting the derivatives in the absorbing layer

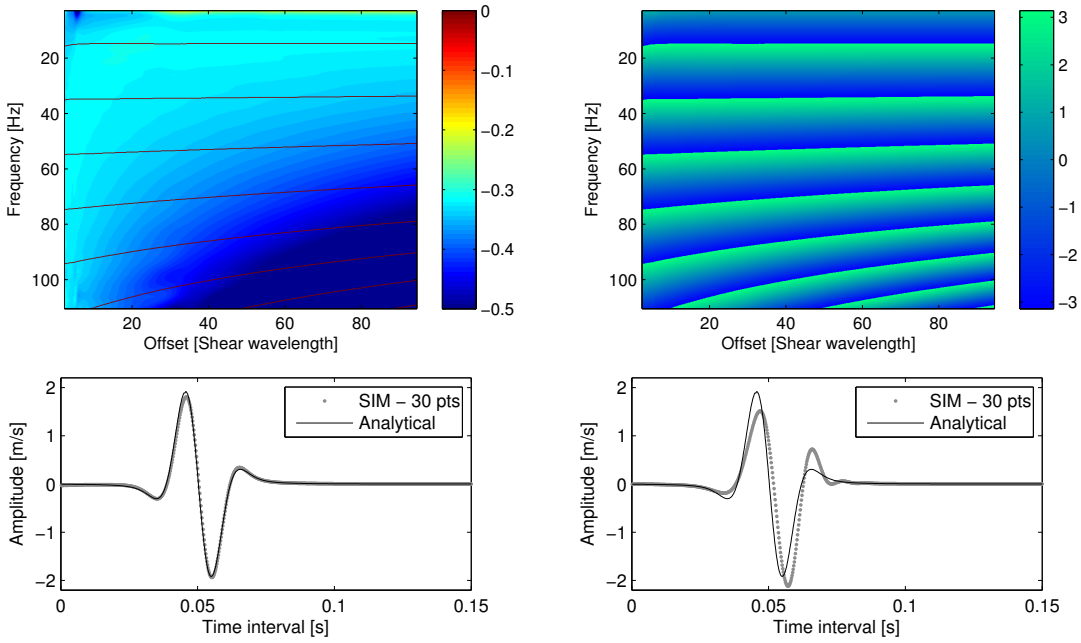


Figure 2.4: Rayleigh-wave properties for SIM method. Top left: phase angle in the frequency-offset domain. Top right: phase angle derivative with respect to frequency. Bottom: comparison between numerical and analytical signals using 30 points per shear wavelength for *offsets equal to 5 (left) and to 33 (right) times the shear wavelength.*

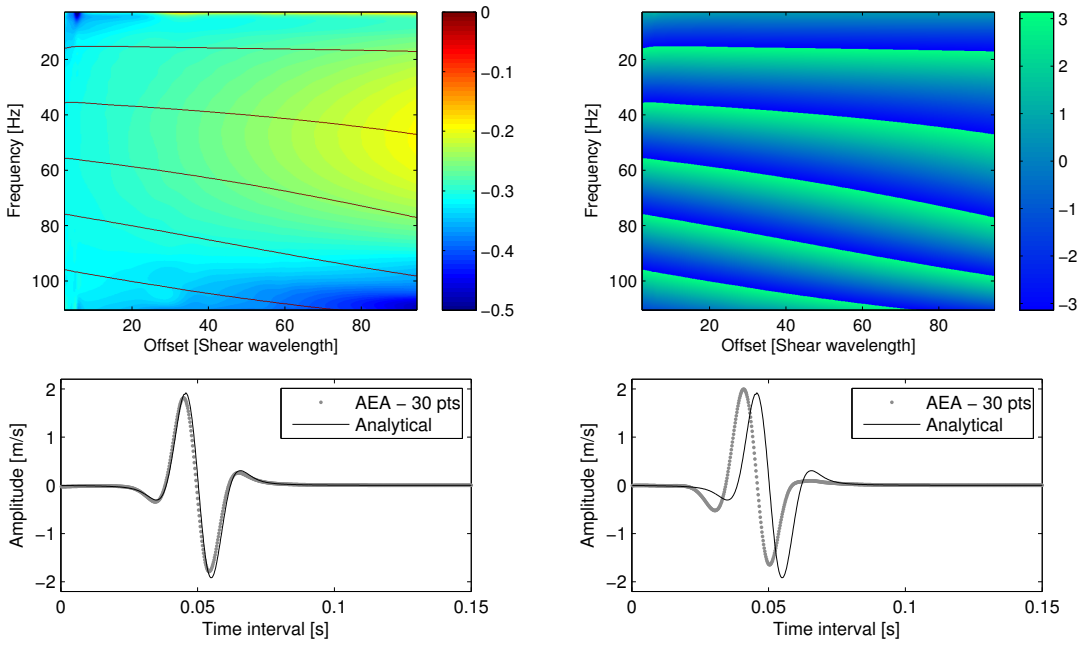


Figure 2.5: Rayleigh-wave properties for AEA method. Top left: phase angle in the frequency-offset domain. Top right: phase angle derivative with respect to frequency. Bottom: comparison between numerical and analytical signals using 30 points per shear wavelength for offsets equal to *5* (left) and to *33* (right) times the shear wavelength.

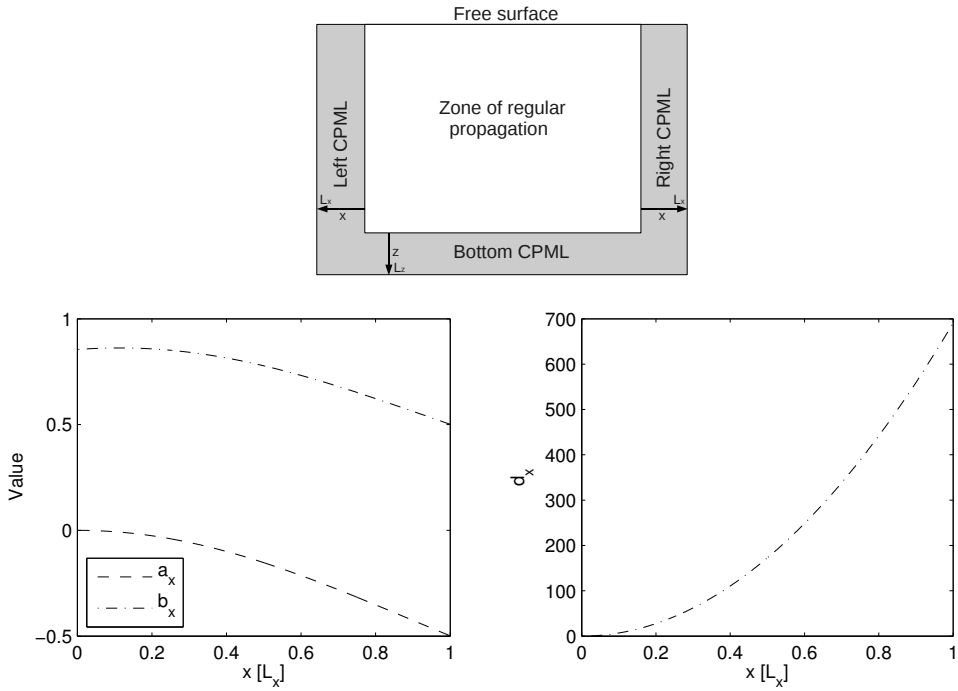


Figure 2.6: Top: General scheme for the boundary conditions. Bottom left: attenuation functions a_x and b_x for a specific case. Bottom right: CPML parameter d_x for a specific case (Δt=1 ms; $f = 50$ Hz and $V = 2000$ m/s).

zone with:

$$\partial_{\tilde{x}} = \partial_x + \psi_x, \tag{2.8}$$

where ψ_x is a memory variable defined, for a discretised time n, as

$$\psi_x^n = b_x \psi_x^{n-1} + a_x (\partial_x)^{n-1/2}, \tag{2.9}$$

where

$$\begin{aligned} b_x &= e^{-(d_x+\alpha_x)\Delta t}, \\ a_x &= \frac{d_x}{d_x+\alpha_x}(b_x - 1), \\ \alpha_x &= \pi f \frac{L_x - x}{L_x}, \\ d_x &= d_0 V \left(\frac{x}{L_x}\right)^2, \\ d_0 &= -\frac{3}{2L_x}\log(R), \end{aligned} \tag{2.10}$$

the variable x represents the Cartesian direction, the constant R is defined as $R = 0.001$, the parameter V is a chosen velocity in the model (commonly the maximum velocity V_p), f is a chosen frequency (commonly the source central frequency) and L_x is the length of the attenuation zone for the trajectory x. Figure 2.6 shows the a_x, b_x (bottom left) and d_x (bottom right) parameters of boundary conditions with respect to the length L_x for a specific case (Δt=1 ms; $f = 50$ Hz and $V = 2000$ m/s).

2.2.4 Source term

Two different sources of energy can be implemented in the modelling: explosive and directional source. The explosive source is modelled as

$$\begin{aligned} \varphi_{\sigma_{xx}} = \varphi_{\sigma_{zz}} = C_a(\lambda + 2\mu)A_{src}, \\ \varphi_{v_x} = \varphi_{v_z} = \varphi_{\sigma_{xz}} = 0, \end{aligned} \tag{2.11}$$

while the directional source is modelled as

$$\begin{aligned} \varphi_{v_x} &= C_a \frac{1}{\rho} A_{src} \sin\theta, \\ \varphi_{v_z} &= C_a \frac{1}{\rho} A_{src} \cos\theta, \\ \varphi_{\sigma_{xx}} &= \varphi_{\sigma_{zz}} = \varphi_{\sigma_{xz}} = 0, \end{aligned} \tag{2.12}$$

where A_{src} defines the amplitude and wave shape required in the model (usually derivatives of the Gaussian function) and θ is the inclination angle of the source with respect to the horizontal surface. The constant C_a is a necessary amplitude correction to match analytical and numerical responses. We used a heuristic approach with different velocity models to find the value of this constant. The value is

$$C_a = \frac{2.5^2}{(\Delta x \Delta z)} \tag{2.13}$$

for 2D models.

2.2.5 Numerical results

Three different experiments were used in order to evaluate the performances of the presented finite-difference method: Lamb's problem (Lamb, 1904), Garvin's problem (Garvin, 1956) and elastic/elastic problem (Berg et al., 1994). The grid used for the test as well as source and receiver positions are shown in Figure 2.7. A second layer at the bottom was used only in the case of the elastic/elastic problem; otherwise the entire grid describes a homogeneous medium with $V_p = 2500$ m/s, $V_s = 1200$ m/s, $\rho = 1000$ kg/m^3. The analytical response was obtained through the Cagniard-de Hoop semi-analytical solution (de Hoop, 1960). For the finite-differences calculation, the second order discretisation in time and space has been used.

Lamb's problem

The Lamb's problem consists of considering the propagation of Rayleigh, P, and S waves in an infinite homogeneous medium with a free surface and a directional (vertical) source (Lamb, 1904). The grid has 430×150 points with the same spatial steps in the two Cartesian directions equal to 0.5 m (Figure 2.7). A first-order Gaussian wavelet with central frequency $f_c = 75$ Hz and time delay $t_0 = 50$ ms is applied in the vertical component of velocity (i.e. $\theta = -90^o$). As the shear wavelength is 16 m, the number of grid points per shear wave is 32 (equation 2.4). The source and receiver R1 are located

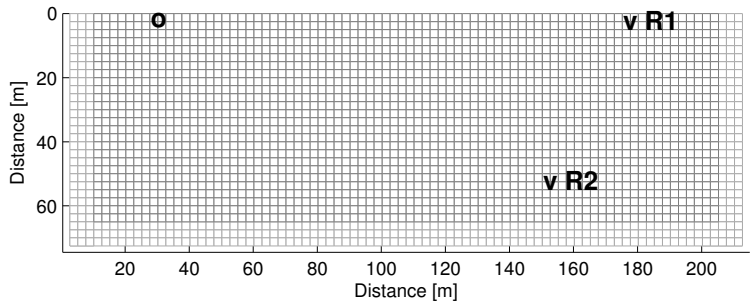


Figure 2.7: Grid used to validate the finite-difference approach on rectangular domains. Lines are plotted every 5 grid points. The lighter parts of the grid identify the CPML attenuation zone.

0.25 m (one grid point) below the surface. The receiver R2 is located 50.75 m below the surface. The distance between receivers and source is 146 m for R1 and 130.2 m for R2. These distances correspond to 9.1 and 8.1 times the shear wavelength. The simulated time is 200 ms with time step $\Delta t = 0.1$ ms.

In this experiment, we expect to have high-amplitude Rayleigh wave and low-amplitude coupled body waves (P and S waves) in R1. In the case of R2, body waves (P and S) are separated and show higher amplitudes, while the Rayleigh wave have lower amplitude. The seismograms are shown in Figure 2.8, each one with its respective mean squared error (MSE) and correlation (Corr) coefficients with respect to the analytical response, without any additional amplitude correction. Low values of MSE and high values of Corr lead to conclude that the method properly simulates the propagation of the seismic wave in the presence of free surface with high agreement in both Rayleigh and body waves, when a directional source is applied.

Garvin's problem

The Garvin's problem consists of considering the propagation of Rayleigh, P, and S waves in an infinite homogeneous medium with a free surface and an explosive source (P-wave source) (Garvin, 1956). Equation 2.11 is used to model this source. We have used a first-order Gaussian wavelet with frequency $f = 75$ Hz and time delay $t_0 = 50$ ms. The distances between source and receivers are the same as in the Lamb's Problem. The simulated time is 200 ms with a time step $\Delta t = 0.1$ ms.

The results are shown in Figure 2.9. In quantitative terms, the MSE between the analytical and the numerical result is low, while Corr is high. With this result, the accuracy of this finite-difference implementation for modelling the explosive source with free-surface conditions is demonstrated.

Elastic/elastic problem

This problem consists of modelling the direct and reflected P and S waves when an interface between two infinite homogeneous media is considered. In order to run this

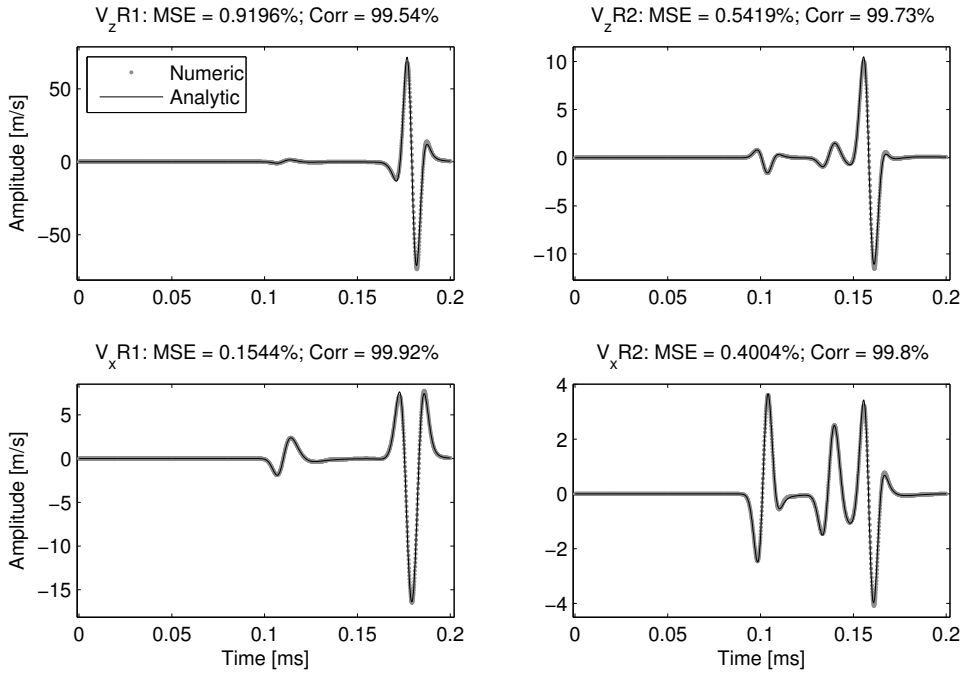


Figure 2.8: Analytical and numerical seismograms for the Lamb's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right).

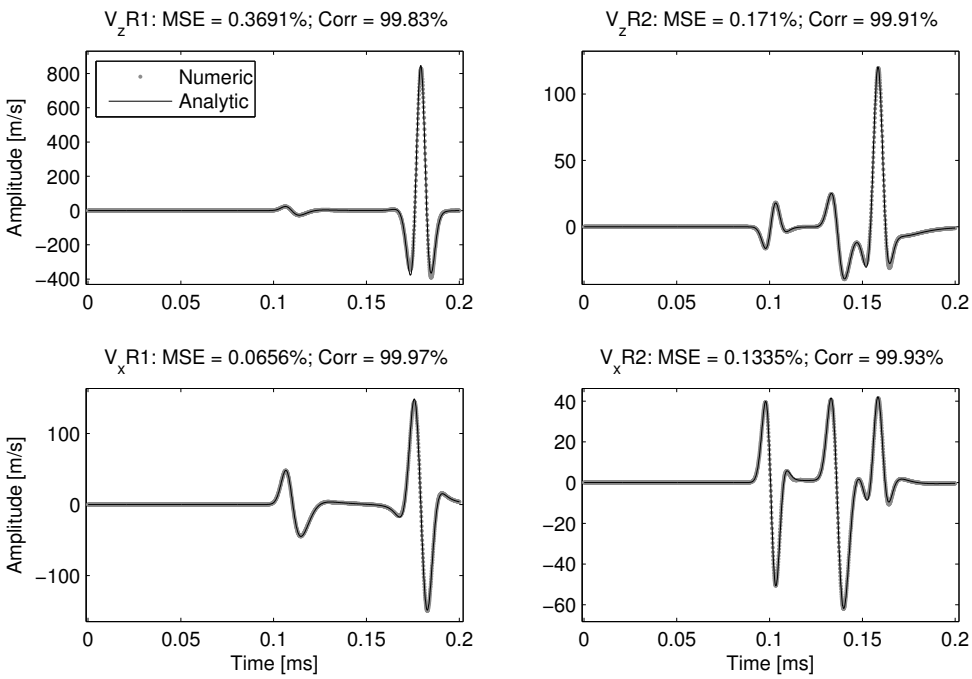


Figure 2.9: Analytical and numerical seismograms for the Garvin's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right).

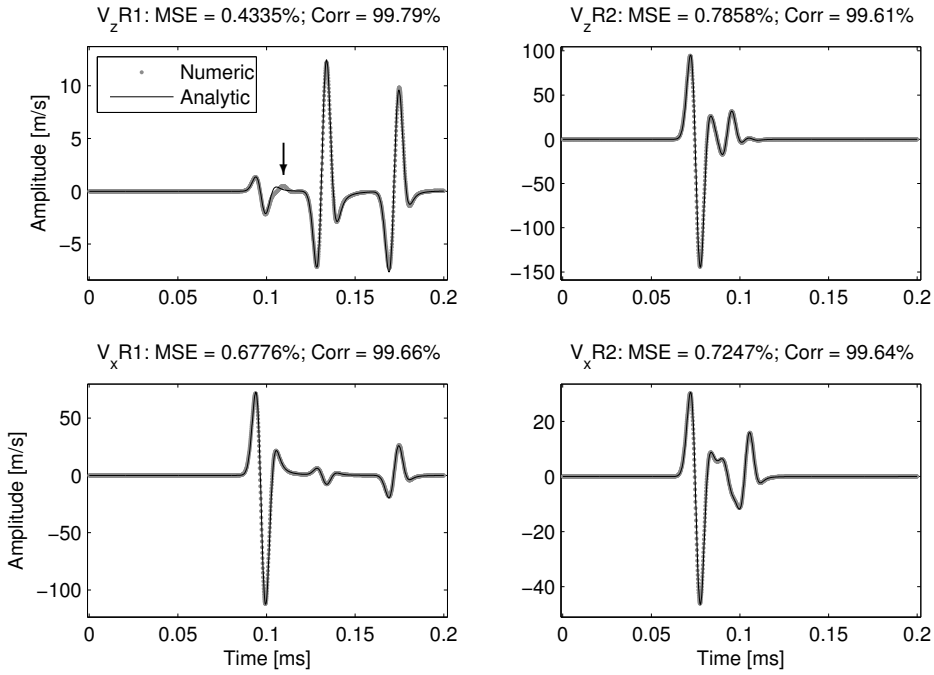


Figure 2.10: Analytical and numerical seismograms for the elastic/elastic problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right).

experiment, a second layer at the bottom part of the model shown in Figure 2.7 is added. The properties of the second layer are $V_p = 3000$ m/s, $V_s = 1600$ m/s, $\rho = 1500$ kg/m^3. The source is explosive and free-surface conditions do not apply to this problem. In fact, as the objective of this experiment is the verification of results for direct and reflected waves, the free-surface conditions have been replaced by CPML conditions, so that Rayleigh waves are not created. In the 380×380 point grid with $\Delta x = \Delta z = 0.5$ m, the explosive source is located 120 m above the top of the second layer. This source is described with the same properties as the one used in the Lamb's problem. The receivers R1 and R2 are located in the first medium at a horizontal distance of 146 m from the source and at vertical distances from the top of the second layer of 119.25 m and 19.25 m, respectively. The normalised distance between source and receiver is 9.1 shear wavelengths for R1 and 11.1 shear wavelengths for R2.

Due to the contrast in velocity and density between the two media and the use of a single P-wave source (explosive source), the expected results should show a low-amplitude direct P-wave, high-amplitude PP-wave (reflected) and PS-wave (reflected) in R1. For R2, the result should display a high-amplitude P-wave, low-amplitude PP-wave and PS-wave (a low-energy direct P-wave is transmitted to the second medium due to the specific contrast properties). Seismograms of the vertical and horizontal components of velocity are shown in Figure 2.10. The good agreement between analytical and numerical results is confirmed by the low MSE and high correlation values. The arrow in the top left plot indicates a non-physical reflection coming from the CPML zone, which is too close to the R1 position. Finally, the results confirm that the approach presented in this section is an effective modelling method with high accuracy when directional and explosive sources

are used in free-surface and contrasted layer media.

2.3 Wave propagation in curved domains

For curved domains, some aspects need special consideration, in particular, the free-surface implementation, attenuating boundaries and computational cost. We present formulation for the chain-rule approach and the tensorial approach (sections 2.3.1 and 2.3.2). Both approaches are general and solve the elastic wave propagation in curvilinear domains. The chain-rule approach is finally chosen based on a computational-cost criterion (section 2.3.3). Free-surface conditions are presented in section 2.3.4 considering the normal vector to the curved vacuum-solid interface. The CPML conditions are presented in section 2.3.5, taking into account the change in the direction of the derivative for curvilinear domains. The implementation of the source terms and their constants is presented in section 2.3.6. Numerical results are shown in section 2.3.7. We evaluate the implemented approach for the same reference cases: the tilted Lamb's problem, the tilted Garvin's problem and the tilted elastic/elastic interface. We analyse the case of non-perpendicular angles between the axes in the staggered grid at the free surface, and show an example of wave propagating in a laterally varying medium.

2.3.1 Chain-rule approach

The chain-rule approach consists in finding the contribution of each directional derivative of velocity and stress fields, when velocity and stress are discretised over a curved distribution of grid points in space with respect to the Cartesian components. In this case, the modelling grid obeys a curvilinear distribution defined by the functions $\xi(x, z)$ and $\eta(x, z)$, with x and z being the components in Cartesian coordinates and ξ and η being the so-called distribution functions. Using this approach, the partial derivative of a generic function $f(x, z)$ with respect to $\xi(x, z)$ is given by

$$\partial_\xi f = (\partial_\xi x)(\partial_x f) + (\partial_\xi z)(\partial_z f), \tag{2.14}$$

while the derivative of $f(x, z)$ with respect to $\eta(x, z)$ is calculated by

$$\partial_\eta f = (\partial_\eta x)(\partial_x f) + (\partial_\eta z)(\partial_z f). \tag{2.15}$$

The derivatives of x and z with respect to $\xi(x, z)$ and $\eta(x, z)$ can be calculated starting from the system of equations

$$\begin{aligned}
\partial_x x &= 1 = (\partial_x \xi)(\partial_\xi x) + (\partial_x \eta)(\partial_\eta x),\\
\partial_x z &= 0 = (\partial_x \xi)(\partial_\xi z) + (\partial_x \eta)(\partial_\eta z),\\
\partial_z x &= 0 = (\partial_z \xi)(\partial_\xi x) + (\partial_z \eta)(\partial_\eta x),\\
\partial_z z &= 1 = (\partial_z \xi)(\partial_\xi z) + (\partial_z \eta)(\partial_\eta z),
\end{aligned} \tag{2.16}$$

so that

$$\begin{aligned}
\partial_\xi x &= \mathcal{J}\partial_z \eta, & \partial_\eta x &= -\mathcal{J}\partial_z \xi,\\
\partial_\xi z &= -\mathcal{J}\partial_x \eta, & \partial_\eta z &= \mathcal{J}\partial_x \xi,\\
\mathcal{J} &= (\partial_z \eta \partial_x \xi - \partial_x \eta \partial_z \xi)^{-1},
\end{aligned} \tag{2.17}$$

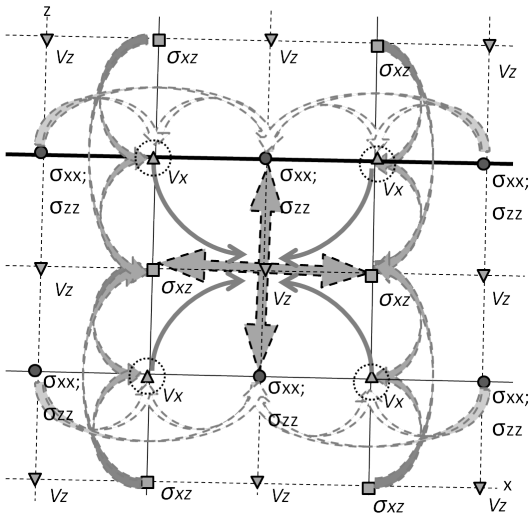


Figure 2.11: Staggered grid for rectangular domains. The arrows indicate how an averaging process would be performed in the case of curved domains.

where $\mathcal{J}$ is the Jacobian of the transformation. Note that $\mathcal{J}$ should be determined and non-zero at all points of the surface.

Applying the transformation described by equations 2.14 and 2.15 into the system of equations 2.2, we obtain the set of equations modelling the wave propagation in the curved system:

$$\begin{aligned}
\dot{v}_x &= \frac{1}{\rho}(\partial_\xi x \partial_x \sigma_{xx} + \partial_\xi z \partial_z \sigma_{xx} + \partial_\eta x \partial_x \sigma_{xz} + \partial_\eta z \partial_z \sigma_{xz}) + \varphi_{v_x}, \\
\dot{v}_z &= \frac{1}{\rho}(\partial_\xi x \partial_x \sigma_{xz} + \partial_\xi z \partial_z \sigma_{xz} + \partial_\eta x \partial_x \sigma_{zz} + \partial_\eta z \partial_z \sigma_{zz}) + \varphi_{v_z}, \\
\dot{\sigma}_{xx} &= (\lambda + 2\mu)(\partial_\xi x \partial_x v_x + \partial_\xi z \partial_z v_x) + \lambda(\partial_\eta x \partial_x v_z + \partial_\eta z \partial_z v_z) + \varphi_{\sigma_{xx}}, \\
\dot{\sigma}_{zz} &= (\lambda + 2\mu)(\partial_\eta x \partial_x v_z + \partial_\eta z \partial_z v_z) + \lambda(\partial_\xi x \partial_x v_x + \partial_\xi z \partial_z v_x) + \varphi_{\sigma_{zz}}, \\
\dot{\sigma}_{xz} &= \mu(\partial_\eta x \partial_x v_x + \partial_\eta z \partial_z v_x + \partial_\xi x \partial_x v_z + \partial_\xi z \partial_z v_z) + \varphi_{\sigma_{xz}},
\end{aligned} \tag{2.18}$$

where the notations for Lamé parameters, density, velocity, stress and source term are kept unchanged with respect to section 2.2.

Implementation in the staggered grid

The derivatives in the numerical curved domain can be computed using the finite-difference approximation described in equation 2.3. Nonetheless, if we use the same staggered grid as for rectangular domains (Virieux, 1986) then some of the partial derivatives of the system of equations 2.18 could not be directly computed in the needed grid points. For instance, in Figure 2.11 an example of the calculation for v_z in such a grid is showed. In this case, the derivatives $\partial_z \sigma_{xz}$ and $\partial_x \sigma_{zz}$ cannot be directly found. These derivatives would need to be computed at the surrounding points and then a linear averaging process would provide the needed values (see Figure 2.11). As a result, the computational cost

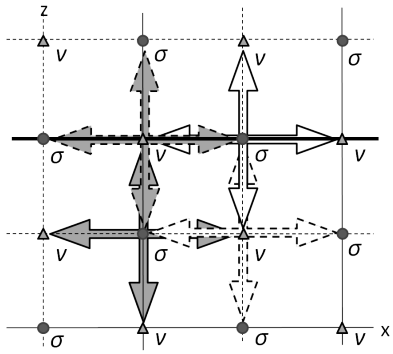


Figure 2.12: Modified staggered grid for curved domains. Velocity components are placed at the points identified with a triangle. The stress components are placed at the points identified with a circle. The arrows indicate how each centred derivative is computed. The bold line identifies interface with the free surface.

would be increased and, more importantly, numerical imprecisions would be introduced as a result of the averaging process.

We propose to adopt a different staggered grid to overcome the described computation difficulties. In our approach, two discretisation grids separated by half the spatial step are used: one for velocity components and the other for stress components (see Figure 2.12). The new staggered grid is different from the classical one (Virieux, 1986) because an additional grid point is introduced in the middle of every 4 points for each velocity and stress components (and also for the elastic model parameters). This approach requires to save in memory twice the grid point number of the classical approach, but no numerical imprecisions are introduced to the chain rule approach. The advantages of using this grid, where the components are placed in the needed position for straightforward computation, have been explained in details by Saenger et al. (2000) (Figure 1c in the article). The collocation method for this grid is well adapted for discretising the chain rule approach.

In the 2D curved system, the numerical stability is ensured when the Courant condition holds

$$\Delta t V_{max} \sqrt{\frac{1}{(\Delta\xi_{max})^2} + \frac{1}{(\Delta\eta_{max})^2}} \leq 1, \tag{2.19}$$

where $\Delta\xi_{max}$ and $\Delta\eta_{max}$ are the maximum spatial steps for each component in the curved distribution. We now present an alternative approach.

2.3.2 Tensorial approach

A set of universal equations for the wave propagation in solid media can be written by considering that the mechanics of continuous media must be invariant in any system of coordinates. In the tensorial approach, we preserve notation used in section 2.2, except for Einstein's notation which is used to denote summation over a repeated index.

The equation for the conservation of linear momentum is written in its invariant expression, for a general coordinate system defined by ξ^j, as (McConnell, 1957; Malvern,

1969; Komatitsch et al., 1996)

$$\rho \dot{v}_i = \nabla_{\xi^j} \sigma_i^j + \varphi_i, \tag{2.20}$$

where the invariant divergence term of the stress tensor is

$$\nabla_{\xi^j} \sigma_i^j = \frac{\partial \sigma_i^j}{\partial \xi^j} - \Gamma_{bi}^a \sigma_a^j + \Gamma_{ba}^j \sigma_i^a, \tag{2.21}$$

with indices i, j, a and b representing the spatial dimension (i.e. in 2D their only possible values are 1 and 2); the superscripts indicating the contra-variant component and subscripts the covariant component; and the symbol Γ_{ab}^c being the needed function (not tensor) to make the differentiation invariant under coordinate transformations (commonly called Christoffel symbols of the second kind). The definition of the Christoffel symbols of the second kind is presented at the end of this section.

The strain-velocity invariant relation is defined as

$$\dot{\varepsilon}_{ij} = \frac{1}{2}(\nabla_{\xi^i} v_j + \nabla_{\xi^j} v_i), \tag{2.22}$$

where ε represents the finite-strain tensor, and the covariant derivative of velocity is defined as

$$\nabla_{\xi^j} v_i = \frac{\partial v_i}{\partial \xi^j} - \Gamma_{bi}^a v_a. \tag{2.23}$$

The Hooke's law is the physical relationship, between stress and strain, that completes the continuum mechanics theory solved with the tensorial approach

$$\dot{\sigma}_i^j = \lambda \delta_i^j \dot{\varepsilon}_k^k + 2\mu \dot{\varepsilon}_i^j, \tag{2.24}$$

where δ_i^j is the Kronecker delta, defined as 1 when subscript and superscript have the same value and 0 for any other combination; index k represents the spatial dimension.

The expression relating the covariant tensor with the mixed tensor for the finite-strain tensor is given by:

$$\varepsilon_i^j = g^{mj} \varepsilon_{im}, \tag{2.25}$$

where index m represents the spatial dimension. The metric tensor g^{mj} carries on the complete information of the coordinate transformation. In this formulation, the metric tensor is defined by the products of the first order partial derivatives of the transformed coordinate system ξ^j with respect to the Cartesian coordinates (Komatitsch et al., 1996). For this reason, g^{mj} is constant only in the case in which the transformed coordinate system is rectilinear (McConnell, 1957). *This property has a direct influence in the computational cost of the tensorial approach implementation.*

As seen before, the Christoffel symbols of the second kind are the necessary components to preserve invariance for the covariant derivatives under coordinate system transformations (McConnell, 1957). These functions depend only on the metric tensor, such that

$$\Gamma_{ab}^c = \frac{1}{2} g^{cn} \left(\frac{\partial g_{bn}}{\partial \xi^a} + \frac{\partial g_{an}}{\partial \xi^b} - \frac{\partial g_{ab}}{\partial \xi^n}\right). \tag{2.26}$$

where indices c and n represent the spatial dimension. From equation 2.26, we can note that the Christoffel symbols are equal to 0 only in the case of a rectilinear transformation

for the coordinate system (i.e. when the metric tensor is constant). In such a case, the wave propagation equations are very similar to those we have when considering Cartesian coordinates (Cartesian coordinates are a specific case of rectangular coordinates). The only difference would be a constant in equation 2.25, which in its turn implies a proportional (constant) modification of the strain-velocity relation (equation 2.22) with respect to the Hooke's law (equation 2.24).

2.3.3 Computational cost

The two approaches presented above describe the wave propagation in curved domains. We have chosen to implement the chain-rule approach in our work on the basis of a computational-cost analysis. In this analysis, we calculate the number of operations required to model the wave propagation in a general curved domain. Even if, depending on the chosen programming strategy, the computational cost may slightly vary, our result leads to conclude that the chain-rule approach is less costly for low orders of finite-difference discretisation.

Chain-rule approach

Solving equations 2.18 at particular point in time and space implies the computation of 6 partial derivatives related to the stress and 4 partial derivatives related to the velocity. Let l be the number of operations required to calculate a partial derivative. The number of additional operations needed, without taking into account the source term, is: 6 additions and 10 multiplications for both components of velocity; and 7 additions and 13 multiplications (considering $\lambda + 2\mu$ as a fixed matrix during computation) for all the stress components. Overall, the order of the computation is

$$\mathcal{O}_{cr}(n \times (10l + 36)), \tag{2.27}$$

where n is the total number of spatial and time nodes (points) in the simulation scheme.

For comparison, 8 partial derivatives per node must be computed in the classical rectangular domain (equation 2.2), where 4 are needed for stress and 4 for velocity. Moreover, 2 additions and 2 multiplications are required for velocity component computation, and 3 additions and 5 multiplications are required for stress component computation. Hence, the order of the calculation is $\mathcal{O}_{rd}(n \times (8l + 12))$. For the second-order finite difference scheme (two operations per partial derivative), the computational cost of the chain-rule approach is twice the computational cost of the classical approach ($\mathcal{O}^2_{cr}(n \times 56)$ and $\mathcal{O}^2_{rd}(n \times 28)$).

Tensorial approach

Komatitsch et al. (1996) demonstrated that the computational cost for the tensorial approach is less than for the chain-rule approach when performing rectilinear transformations. However, 2D heterogeneous media impose using curvilinear coordinates instead of rectilinear coordinates. Therefore, the Christoffel symbols of the second kind are not zero and they must be taken into account in the computation of each variable.

The computational cost to compute the velocity components using equation 2.20 is calculated as follows. The number of operations required for the computation of the stress divergence in curvilinear 1D media is found from equation 2.21, where 2 multiplications, 2 additions and 1 partial derivative are needed. In the 2D case, we consider that the spatial indices (i, j, a, b) can be equal to 1 and 2, and that the computation of one stress-tensor divergence component needs the same number of operations as for 1D. The stress tensor is defined by 4 different components in 2D, so that the computational cost for the two velocity components in equation 2.20 is $4(l + 4) + 4$. The additional 4 operations are 2 divisions by ρ and 2 additions of the corresponding stress-tensor divergence component for each velocity component.

An equivalent procedure can be used to find the strain tensor at a particular point. From equation 2.23, the number of operations needed to calculate one component of the covariant derivative of velocity are: 1 partial derivative, 2 multiplications and 2 additions (for instance, if $i = j = 1$, the free index a can have two possible values). The number of operations required for the computation of the covariant strain tensor in equation 2.22 when subindex i and j are equal (considering the fact that, in this case, a mathematical reduction of operations can be done), is 1 partial derivative, 2 additions and 2 multiplications. If it is not the case $(i \neq j)$, we need to compute the two components of the covariant derivative of velocity, thus requiring 2 partial derivatives, 2 multiplications and 2 additions (in this case the Christoffel symbols are the same for both components of covariant velocity derivatives). Moreover we need two additional operations: 1 to compute the addition of the 2 components of the velocity derivative and 1 to compute the division by 2. Taking into account the fact that the strain tensor is symmetric (i.e. $\varepsilon_{12} = \varepsilon_{21}$), the number of operations to calculate the covariant strain tensor is: $2(l + 4) + (2l + 4 + 2)$.

The computation of the mixed strain tensor requires 1 addition and 2 multiplications for each of the 4 components (all of them are different even if the strain tensor is defined symmetrical in the covariant case). The total cost to compute the mixed strain tensor is: $2(l + 4) + (2l + 4 + 2) + 4 \times 3$.

The mixed stress tensor, found from the Hooke's law (equation 2.24), requires 8 additional operations for the 4 components: 2 additions and 6 multiplications. Summarising, the computational cost of this method is:

$$\mathcal{O}_{ts}(n \times (8l + 54)). \tag{2.28}$$

Comparing the computational cost of the two approaches for curved domains, we can note that, even if the number of partial derivatives in the tensorial approach is lower than in the chain rule case, the number of additional operations is higher, meaning the existence of a threshold dependent on the number of calculations per partial derivative (Figure 2.13). As the two curves intersect at around 10 operations per partial derivative, we conclude that the computational cost for the chain-rule approach is lower than for the tensorial approach, when using a centred finite-difference scheme up to the 6th order of discretisation.

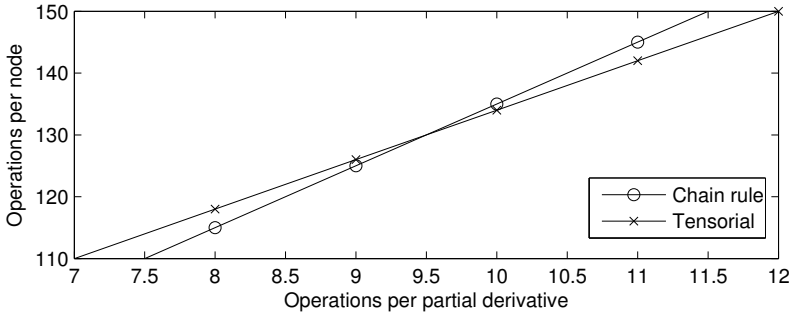


Figure 2.13: Theoretical computational cost of tensorial and chain rule solutions per node vs. number of operations per partial derivative

2.3.4 Free-surface conditions

The theoretical definition of the free-surface boundary condition states that the component of stress normal to the surface is equal to zero (Aki and Richards, 1980; Chapman, 2004). The vector normal to the surface needs to be defined. In our implementation, the free surface coincides with the top of the grid defined by the distribution functions $\xi(x, z)$ and $\eta(x, z)$. For this reason, the normal vector to the free surface can be easily found using the geometry defined by these distribution functions in the Cartesian coordinates (this result is also shown in Zhang and Chen (2006)). The normal vector to the free surface is

$$n^{fs} = \hat{x}\partial_\xi z + \hat{z}\partial_\xi x, \tag{2.29}$$

where superscript 'fs' denotes variables at the free surface.

The theoretical definition of the free-surface conditions is then

$$\sigma_{ji}^{fs}\widehat{n}_i^{fs} = 0 \rightarrow \begin{bmatrix} \sigma_{xx}^{fs} & \sigma_{xz}^{fs} \\ \sigma_{xz}^{fs} & \sigma_{zz}^{fs} \end{bmatrix} \begin{bmatrix} \partial_\xi z \\ \partial_\xi x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.30}$$

In our implementation, we need to express these vertical derivatives which cannot be calculated at the free surface: $\partial_z v_x$ and $\partial_z v_z$. The set of equations in 2.18 are used to replace the stress values in equation 2.30 (considering that equation 2.30 is invariant in time), so that the vertical derivatives of velocity can be found with respect to their horizontal derivatives as

$$\begin{bmatrix} \partial_z v_x \\ \partial_z v_z \end{bmatrix} = -X^{-1}Y \begin{bmatrix} \partial_x v_x \\ \partial_x v_z \end{bmatrix}, \tag{2.31}$$

where X and Y are defined as

$$\begin{aligned} X &= \begin{bmatrix} \chi\partial_\xi z\partial_\xi z + \mu\partial_\xi x\partial_\eta z & \lambda\partial_\xi z\partial_\eta z + \mu\partial_\xi x\partial_\xi z \\ \lambda\partial_\xi x\partial_\xi z + \mu\partial_\xi z\partial_\eta z & \chi\partial_\xi x\partial_\eta z + \mu\partial_\xi z\partial_\xi z \end{bmatrix}, \\ Y &= \begin{bmatrix} \chi\partial_\xi z\partial_\xi x + \mu\partial_\xi x\partial_\eta x & \lambda\partial_\xi z\partial_\eta x + \mu\partial_\xi x\partial_\xi x \\ \lambda\partial_\xi x\partial_\xi x + \mu\partial_\xi z\partial_\eta x & \chi\partial_\xi x\partial_\eta x + \mu\partial_\xi z\partial_\xi x \end{bmatrix}, \end{aligned} \tag{2.32}$$

with $\chi = \lambda + 2\mu$.

In the case of stress, if the vertical grid line is perpendicular to the free surface then the same SIM method used for rectangular domains can be implemented. If angles are

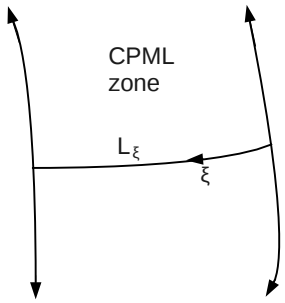


Figure 2.14: Example of the differentiation direction defined by ξ in the CPML attenuation zone of a curved domain.

non-perpendicular, the equivalence of equation 2.30 can be employed. For instance, if the value of $\partial_z \sigma_{xx}$ is needed to find a component of velocity (equation 2.18), the expression (derived from equation 2.30)

$$\sigma_{xx}^{fs} = -\frac{\partial_\xi x}{\partial_\xi z} \sigma_{xz}^{fs}, \tag{2.33}$$

is used together with a modification of the discretisation order. This modification is a reduction from centred finite difference to one-sided finite difference. In section 2.3.7, numerical results for different angles between the axes at the free surface are presented.

2.3.5 Non-reflective boundary conditions

The absorbing boundaries presented for rectangular domains can also be applied in curved domains. The CPML proposed by Komatitsch and Martin (2007) is based on a modification of the derivative in the attenuation zone. As a result, we consider that the directional derivatives used in the chain-rule approach can be approximated using equation 2.8 in the attenuation zone. In this process, the direction of the variable x, in equation 2.8, is the one that defines the variables required by equations 2.9 and 2.10. In curved domains, the direction x must describe the curved lines. In Figure 2.14, we show an example for a curved differentiation trajectory in the case of a generic variable ξ (curved version of x for the CPML formulation). In such a case, the derivative with respect to ξ can be implemented by replacing the partial derivative with the expression

$$\partial_{\tilde{\xi}} = \partial_\xi + \psi_\xi, \tag{2.34}$$

where ψ_ξ is found using equation 2.9.

2.3.6 Source term

The source can be modelled using the set of equations 2.11 and 2.12. Furthermore, considering that the use of distribution functions $\xi(x, z)$ and $\eta(x, z)$ can make the spatial steps to be variable, the constant for the amplitude correction must be modified accordingly.

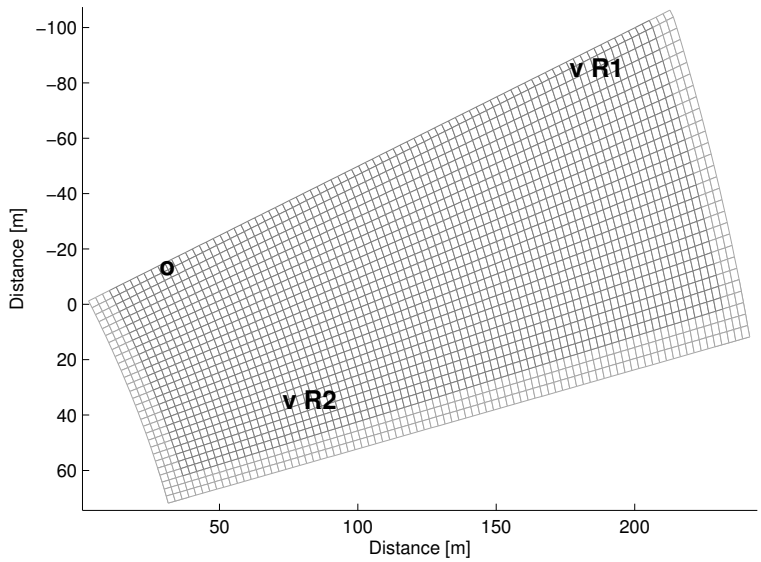


Figure 2.15: Curved grid used for numerical validation. The circle represents the source and the triangles represent the receivers. The lighter parts of the grid identify the CPML attenuation zone. Only one every 5 grid lines is plotted.

The modified constant is

$$C_a = \frac{2.5^2}{\Delta x \Delta z}(\partial_\xi x(x_s, z_s)\partial_\eta z(x_s, z_s) + \partial_\eta x(x_s, z_s)\partial_\xi z(x_s, z_s)), \tag{2.35}$$

where the coordinates (x_s, z_s) define the position of the source in the 2D media. This definition generalises the Cartesian grid constant to curved grids.

2.3.7 Numerical results

To analyse the chain-rule approach for curved domains, we propose comparing analytical and numerical solutions for three numerical experiments and a tilted free surface. We use the tilted Lamb's problem, the tilted Garvin's problem and the tilted elastic/elastic problem. Analytical solutions exist because the tilted-interface problems are rotated versions of the classical problem. For finite-difference calculation, a second-order discretisation scheme has been used. We have preserved orthogonality between axes at the free surface for the three mentioned experiments but we also present an analysis when this is not true. Finally, we compute seismograms for a complex model.

Tilted Lamb's problem

In order to validate the method, we perform a comparison between the analytical and numerical solutions when the free surface has a slope of 30°. The homogeneous model is identified with the properties $V_p = 2500$ m/s, $V_s = 1200$ m/s, $\rho = 1000$ kg/m^3. The grid is shown in Figure 2.15. For this experiment, no special angles between axes are used,

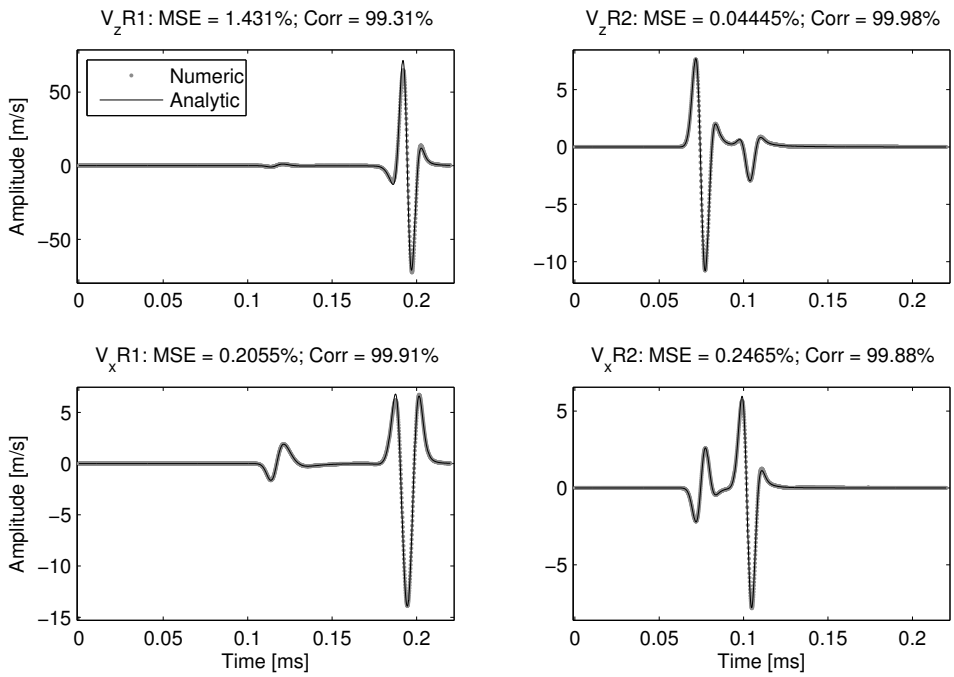


Figure 2.16: Analytical and numerical seismograms for the tilted Lamb's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.15.

except for the free surface grid points, where a 90° angle has been preserved. The source is placed 1.3975 m below the surface. One of the receivers is located at a horizontal distance of 162.1149 m from the source and 1.3975 m below the surface. The location of the second receiver with respect to the first one is -146.4859 m in the horizontal direction and 61.0350 m in the vertical direction. In all cases, the horizontal distance is supposed to be parallel to the surface and the vertical distance is perpendicular to the surface. The normalised distance between source and receiver is 10.1 shear wavelengths for R1 and 3.9 shear wavelengths for R2.

The grid used in this experiment is made of 430×150 points. The range of spatial steps defined by the distribution functions are $\Delta\xi = [0.5 - 0.75]$ m and $\Delta\eta = [0.5 - 0.75]$ m. The source has the same properties as the one used in the classical Lamb's problem. After calculating the Courant stability condition, the proper time step was found to be $\Delta t = 0.1$ ms. The total simulation time is 220 ms.

Since the classical Lamb's problem is defined with a force source applied on the surface along a direction perpendicular to the surface, a rotation of the directional source, with $\theta = 30^o$, is applied. Moreover, a rotation of the numerical results must be applied using the transformation described in the following equation

$$\begin{pmatrix} v_x^R \\ v_z^R \end{pmatrix} = \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix} \begin{pmatrix} v_x \\ v_z \end{pmatrix}. \tag{2.36}$$

Using the rotated numerical velocities v_x^R and v_z^R, the comparison with the analytical

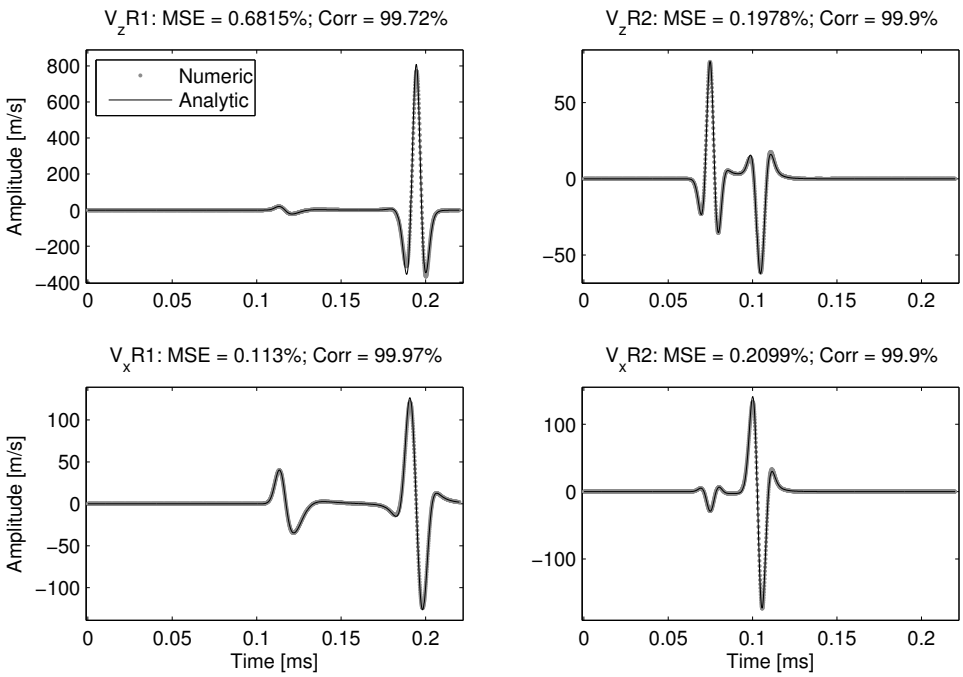


Figure 2.17: Analytical and numerical seismograms for the tilted Garvin's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.15.

solution can be performed. The MSE criteria gives a quantitative measure of misfit between the results. In the case of the vertical component, the misfit for R1 and R2 are 1.43% and 0.04% respectively. For the horizontal component they are 0.20% and 0.24% confirming the good agreement also shown by the high values ($> 99\%$) of Corr (Figure 2.16).

Tilted Garvin's problem

The compressional source modelled for this experiment is the one described in equation 2.11. It produces a spherical P-wave. The medium considered in this experiment corresponds to the one used in the tilted Lamb's problem. The numerical results must be rotated as described by equation 2.36. The positions of source and receivers are the same as for the tilted Lamb's problem, except for the vertical location of the source which, for this experiment, is 1.04 m below the surface.

Referring to the results shown in Figure 2.17, the highest MSE percentage is 0.6815 %, while the lowest Corr percentage is 99.72%, both for the vertical velocity component in R1. These results allow us to conclude that the propagation due to an explosive source in a homogeneous medium is modelled with high accuracy when compared to the analytical solution for surface and body waves, at least for smoothly curved grids.

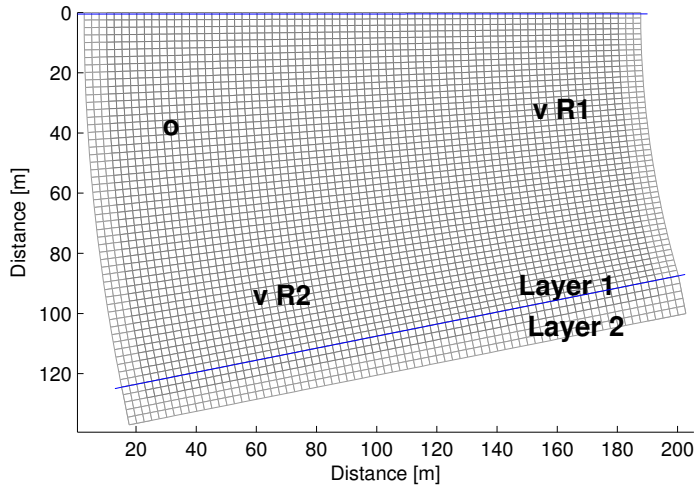


Figure 2.18: Grid used to model the two-layer medium for the tilted elastic/elastic problem.

Tilted elastic/elastic problem

A dipping interface between two homogeneous media is used in this experiment (see Figure 2.18). The properties of the top layer are $V_p = 2500$ m/s, $V_s = 1200$ m/s, $\rho = 1000$ kg/m^3. The second layer is identified by $V_p = 3000$ m/s, $V_s = 1600$ m/s, $\rho = 1500$ kg/m^3. An explosive source (identified by a circle in Figure 2.18) is placed 95.13 m above the interface. The P-wave source has the same properties as the one used in the Garvin's problem. One of the receivers is placed at 86.64 m above the interface and at 119.74 m horizontally from the source. The second receiver is at 34.35 m above the surface and is 17.16 m horizontally from the source. These locations are given accordingly with the 10% dipping slope of the interface. Horizontal distances are parallel to the interface and vertical distances are perpendicular to it. The normalised distance between source and receiver is 6 shear wavelengths for R1 and 4 shear wavelengths for R2.

As the coordinate system is rotated, the results are compared after applying the transformation described by equation 2.36. As shown in Figure 2.19, the MSE misfits are 0.44% and 0.80% for R1 and R2 for the vertical component of velocity, and 0.66% and 0.74% for the horizontal component. The results displayed in Figure 2.19 show the expected match in the direct P-wave, reflected P-wave and reflected S-wave for the two velocity components at both receiver locations.

The relative error between analytical and numerical solutions expressed by MSE and Corr values can be used to compare rectangular and curved domain implementations. These quantities are summarised in Table 2.1. In this comparative table, the relative distance (in terms of the shear wavelength) between source and receivers is included. In the tests with a free surface, the near surface receiver is one wavelength closer to the source in the rectangular case. The results are slightly better in the rectangular case both in terms of MSE and Corr values. For the buried receiver and free surface, the results are better in the curved case. However, the distance between source and receiver

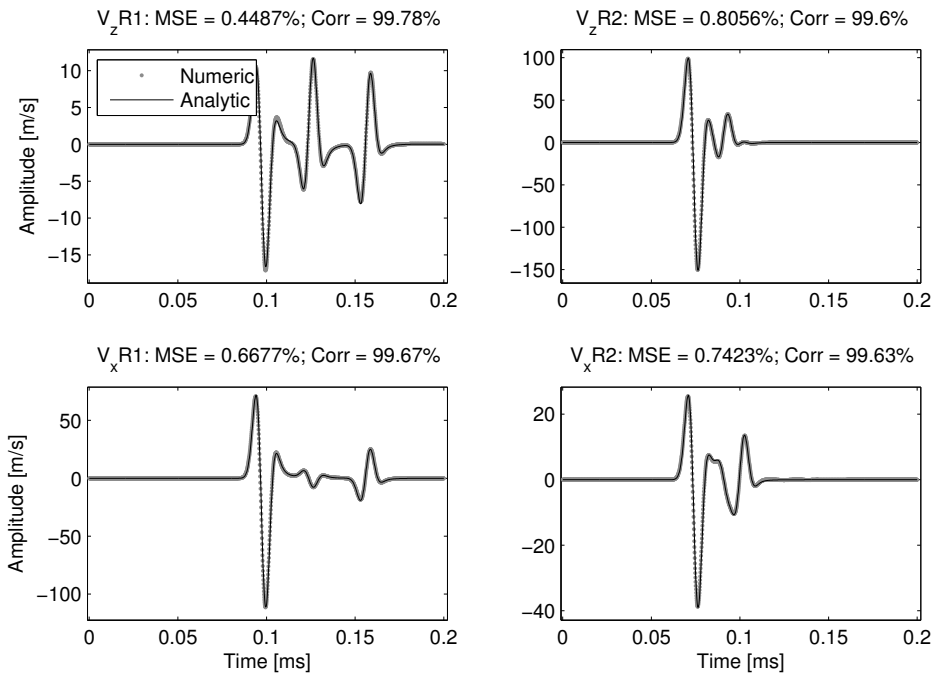


Figure 2.19: Analytical and numerical seismograms for the tilted elastic/elastic problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.18.

Receiver → Test ↓	d R1 (λ_s)	V_x R1 (MSE; Corr)	V_z R1 (MSE; Corr)	d R2 (λ_s)	V_x R2 (MSE; Corr)	V_z R2 (MSE; Corr)
Lamb	9.1	0.15; 99.92	0.92; 99.54	8.1	0.40; 99.90	0.54; 99.73
Tilted Lamb	10.1	0.21; 99.91	1.43; 99.31	3.9	0.25; 99.88	0.04; 99.98
Garvin	9.1	0.07; 99.97	0.37; 99.83	8.1	0.13; 99.93	0.17; 99.91
Tilted Garvin	10.1	0.11; 99.97	0.68; 99.72	3.9	0.20; 99.90	0.20; 99.90
Elastic/elastic	9.1	0.68; 99.66	0.43; 99.79	11.1	0.72; 99.64	0.79; 99.61
Tilted e./e.	6	0.67; 99.67	0.45; 99.78	4	0.74; 99.63	0.81; 99.60

Table 2.1: Results using rectangular and curved domain implementations (Figures 2.8 to 2.10, 2.16, 2.17 and 2.19). The 'd' stands for distance between source and receiver and λ_s for shear wavelength.

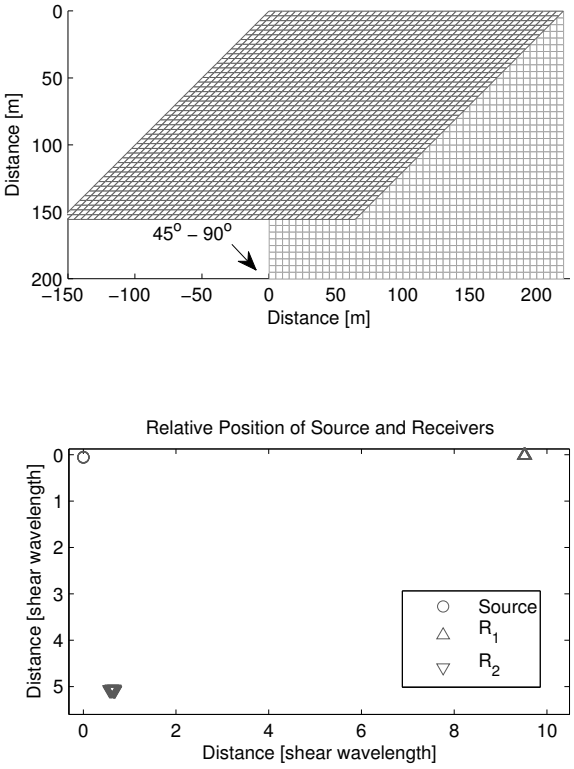


Figure 2.20: Top: different grids with 45° to 90° angles between axes have been used. Only the first and the last grid are plotted. Bottom: relative position of source and receivers for the complete set of experiments.

is half the distance of the rectangular case. For the elastic/elastic interface tests, the MSE and Corr values are fairly similar for both implementations although the distance is shorter in the curved one (66 % for R1 and 36 % for R2). The results given by the rectangular domain implementation are relatively better than the curved domain results. Nevertheless, the MSE is always smaller than 1.5% and the Corr is always higher than 99% for source-receiver distances up to 11.1 times the shear wavelength, which is an evidence of accuracy.

Analysis of variable angles at the free surface

The free-surface experiments presented so far have been run using orthogonal axes at the free surface. In this paragraph, we analyse results for angles at the free surface ranging from 45° to 90° (Figure 2.20). MSE and Corr are used to evaluate precision with respect to the angle between axes.

The medium is homogeneous with $V_p = 2500$ m/s, $V_s = 1200$ m/s, $\rho = 1000$ kg/m^3 and is discretised in a 450×450 grid. An explosive source (75 Hz Ricker wavelet) has been used. 32 points per shear wavelength and a Courant number of 0.48 in all the experiments ensures stability and comparability.

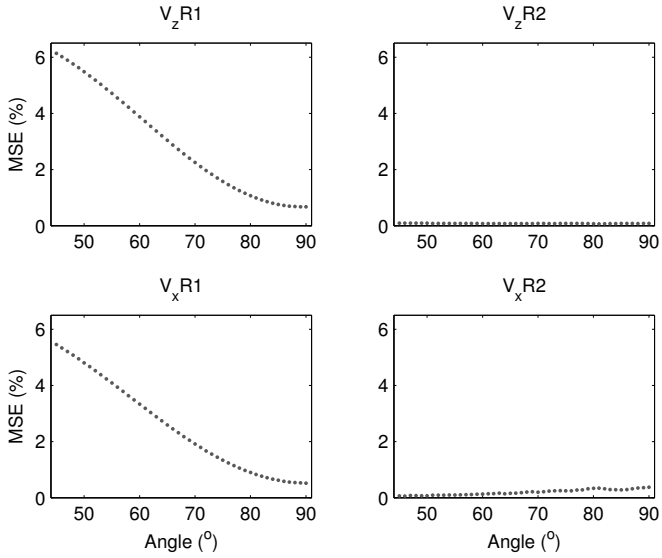


Figure 2.21: MSE for the range of angles from 45° to 90° measured at the two receivers described in Figure 2.20.

In Figure 2.20, the relative position of source and receivers is shown. The distance between the receiver located near the free surface and the source is 9.5 times the shear wavelength. For the buried one, the distance to the source is 5 times the shear wavelength. Note that the relative position of R2 is not exactly constant (Figure 2.20). Nevertheless, such position varies within 4% of the shear wavelength.

The angle has a stronger impact on the quality of the signal at the free surface (R1) than deeper in the model (R2) (Figures 2.21 and 2.22). The maximum MSE percentage is 6% and the minimum Corr percentage is 96.7%, both of them, for an angle of 45° and for receiver R1. In the case of R2, the vertical component presents almost no variations with angle (top left of Figures 2.21 and 2.22). Smaller variations in the horizontal component (R2) are principally related to the slight variation of its relative position.

We explain these results by considering the reduction in the order of the finite-difference discretisation at the free surface. In fact, in our approach we reduce the second-order discretisation to first order at the free surface. Moreover, as well as in the case of rectangular grids (see for example Xia et al. (2007)), the match between analytical and numerical solutions can be increased by considering more n_λ points. Further research could be done to find a more accurate implementation in the case of non-perpendicular grid angles at the free surface.

Wave propagation modelling in a complex medium

As an example of wave propagation modelling in a complex medium, we have created a velocity model with curved layers and a normal fault (Figure 2.23 top). The structure of this model has been implemented using a curved grid by letting the interfaces between layers coincide with the distribution functions $\xi(x, z)$ and $\eta(x, z)$. The implementation of such a grid avoids interpolations in the modelling of the structure. The profiles for V_p

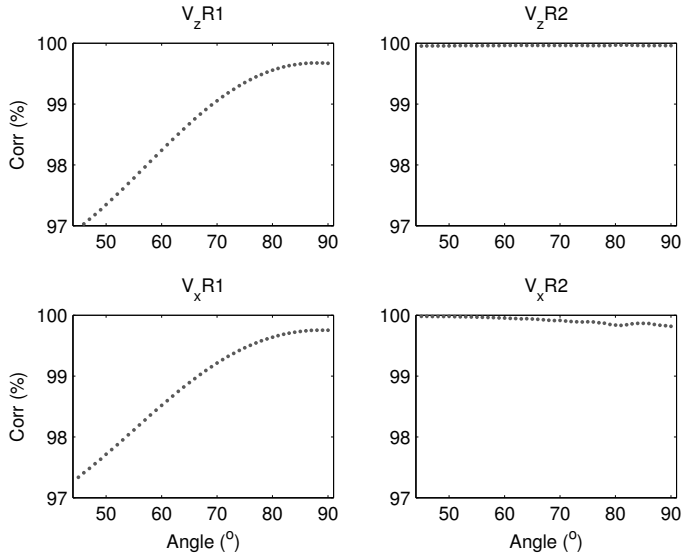


Figure 2.22: Correlation for the range of angles from 45° to 90° measured at the two receivers described in Figure 2.20.

and V_s, as well as the density ρ and the Poisson ratio are shown in Figure 2.23 (bottom), for the two locations indicated by arrows.

The spatial steps for modelling were defined in the ranges $\Delta\xi = [0.25 - 0.45]$ m and $\Delta\eta = [0.25 - 0.45]$ m. The curved grid is implemented with 860×300 points. In order to satisfy the stability condition, the time step is $\Delta t = 0.037.5$ ms. An explosive source (75 Hz Ricker wavelet delayed by 50 ms) was implemented.

The grid and a snapshot at t=0.125 ms are shown in Figure 2.24. Elastic waves have been identified in the snapshot (Figure 2.24): body (direct, transmitted, reflected) and surface (Rayleigh (1885)) waves. Rayleigh waves (P-SV media) are commonly known as ground roll in exploration seismology. Figure 2.25 shows vertical and horizontal velocities for a line of receivers located at the surface. We use this numerical result to show that elastic wave propagation in 2D isotropic media with curved interfaces and topography in addition to lateral variations can be handled with our implementation. Nonetheless, derivatives of the distribution functions must exist and the Jacobian of the transformation must be different than zero at all the grid points (i.e. smoothly curved grids).

2.4 Conclusions

We have presented an approach for elastic wave modelling in rectangular and curved grid domains. We have applied the chain rule to the velocity-stress formulation. As discretisation strategy, we have implemented second-order finite-differences in a modified staggered grid. Two distribution functions are used for the design of the grid, so that the partial derivatives of the stress and velocity are evaluated with respect to the distribution functions. Free-surface conditions have been modelled by implementing the theoretical definition (given in continuous mechanics) in curved grids and defining the free surface at

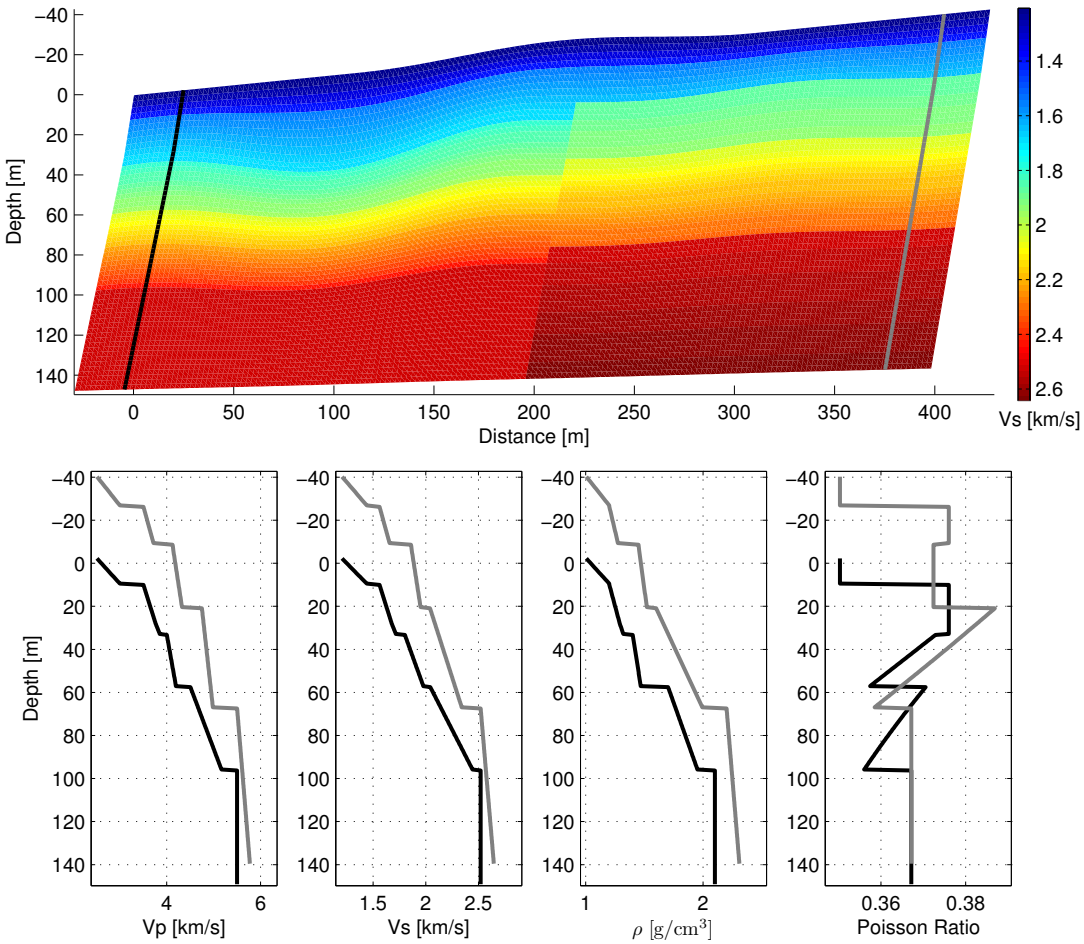


Figure 2.23: Complex medium. Top: 2D S-wave velocity model. Bottom: 1D elastic parameter profiles for the positions indicated in the 2D model by the black and the grey lines.

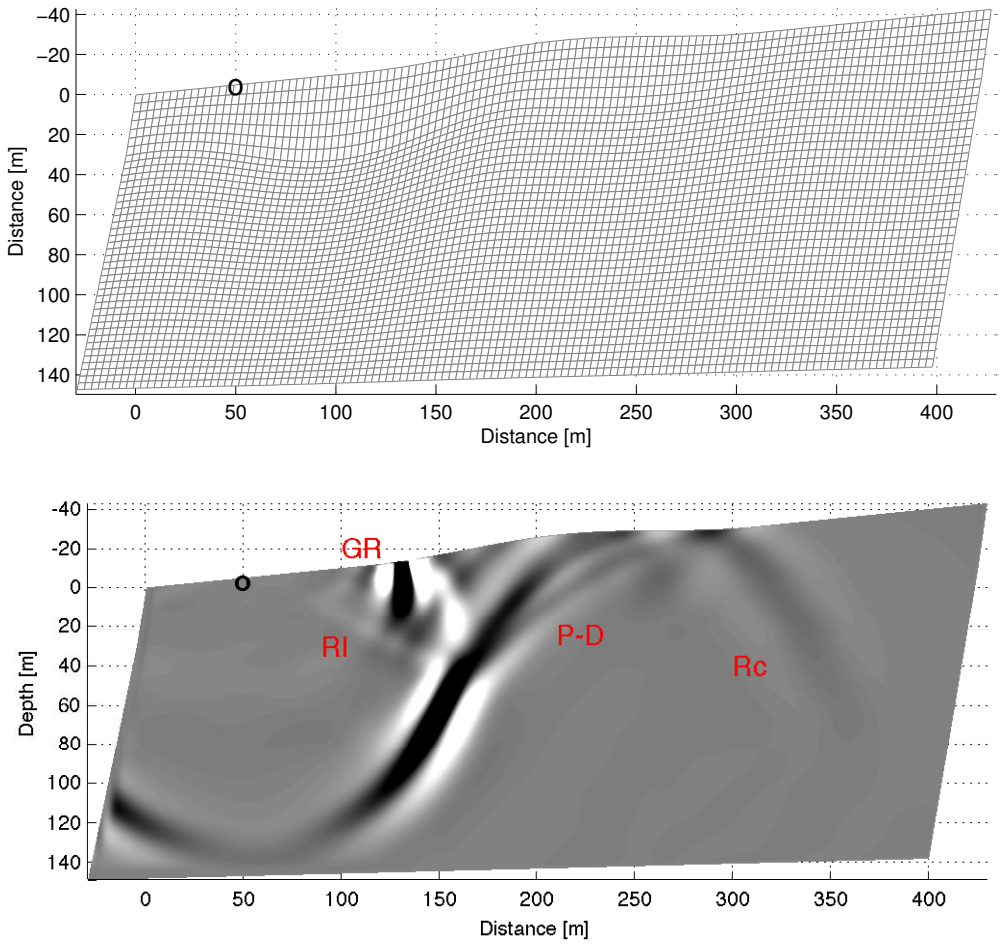


Figure 2.24: Discretisation grid (top) and snapshot of wave propagation at t=0.125 s (bottom) for V_z component. Only one every 7 grid lines is plotted. The circles indicate the source position. Identified waves are examples of: P direct (P-D), refracted (Rc), reflected (Rl) and surface waves (GR).

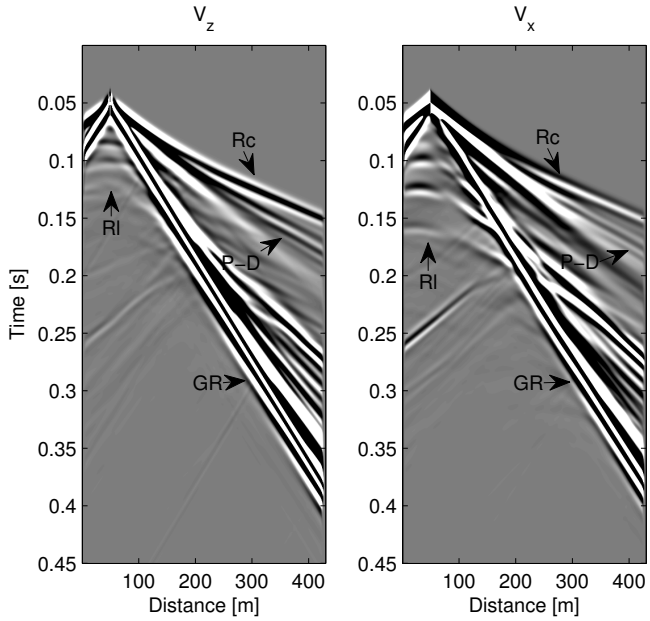


Figure 2.25: Shot gathers recorded at the surface (Figures 2.23 and 2.24). Seismic waves indicated by arrows are examples of waves identified in the snapshot (Figure 2.24). Clipping is stronger than in the snapshot.

the top of the grid. In such implementation, the required vector (normal to the surface) is found from the geometry defined by the distribution functions. We have implemented CPML attenuation conditions considering that the trajectory of differentiation depends on the distribution functions.

We have tested this approach by comparing analytical and numerical solutions for 2D homogeneous isotropic elastic wave propagation, both with horizontal and tilted interfaces, and using explosive and directional sources (Lamb's problem, Garvin's problem and elastic/elastic problem). The high agreement between analytical and numerical solutions demonstrates the reliability of this modelling approach. When varying the angle between the axes at the free surface, better results are obtained if the axes are orthogonal. Finally, we have computed the elastic wave propagation in a complex medium, using a curved grid, to show that our approach can be used to obtain the complete propagated field (surface and body waves) in the presence of curved interfaces, topography and lateral variations. This chain-rule based approach for modelling the elastic wave propagation in curved domains is an accurate tested modelling tool that will be employed in the following chapters for waveform inversion.

On the side of limitations, the chain-rule approach works only if the derivatives of the distribution functions exist and the Jacobian of the transformation is not equal to zero at all grid points. Therefore, our tool is limited to smoothly varying grids. In the presence of strongly distorted grids, numerical errors are expected. If strong heterogeneity / interface roughness is present, then staircase representation of the interfaces has to be used for the modelling.

Chapter 3

Inversion of surface waves

Contents

3.1 Introduction **61**

3.2 Inverse problem theory **62**

3.2.1 Inverse problem solution 63

3.2.2 Adjoint-state method 65

3.2.3 Strategy for updating the parameters 65

3.3 Surface wave analysis (SWA) **67**

3.3.1 Processing workflow for dispersion curve extraction 68

3.3.2 Inversion of dispersion curves 71

3.3.3 Conclusions 85

3.4 Full waveform inversion (FWI) **86**

3.4.1 Gradient computation 86

3.4.2 Source wavelet estimation 88

3.4.3 Resolution and surface waves 90

3.4.4 Alternative objective functions 92

3.4.5 Conclusions 93

3.5 Discussion **94**

Résumé du chapitre

Dans ce chapitre, nous introduisons la théorie du problème inverse. Cette théorie est ensuite appliquée à l'analyse des ondes de surface (Surface Wave Analysis, SWA) et puis à l'inversion des formes d'onde (Full Waveform Inversion, FWI). Une discussion est proposée à la fin du chapitre. Dans cette discussion, l'objectif est de donner au lecteur un aperçu des avantages et des limitations à prendre en compte pour le choix d'une nouvelle formulation (explicitée dans le chapitre 4).

Dans la section 3.2, la théorie du problème inverse est présentée de manière générale (Tarantola, 2005). Cette généralité est privilégié car cela nous permet de l'appliquer aux deux méthodes spécifiques d'inversion développées ici. La fonction objective classique mesure la différences au sens des moindres carrés entre les données observées et les données modélisées (calculées avec l'outil numérique developpé dans le chapitre précédent). Les paramètres du modèle sont actualisés dans la direction dictée par le gradient de la fonction objective. L'aspect essentiel consiste à calculer de manière efficace ce gradient ; c'est ce qui est fait en utilisant la méthode de l'état adjoint (Plessix, 2006). Enfin, les différentes stratégies d'actualisation du modèle sont abordées : Newton, Gauss-Newton, quasi-Newton, régularisations de la fonction objective.

Dans la section 3.3, les approches classiques de la SWA utilisées actuellement pour des applications dans la proche sub-surface sont présentées en détails. De plus, des développements originaux sont proposés : une méthodologie pour extraire les courbes de dispersion des enregistrements en point tir, et des règles heuristiques pour aider à la convergence de l'inversion. Pour les exemples d'inversion, nous considérons soit le mode fondamental des ondes de surface représenté par une seule courbe de dispersion (Xia et al., 1999), soit les multiples modes représentés par plusieurs courbes de dispersion (Socco and Strobbia, 2004; Luo et al., 2007). L'utilisation des plusieurs courbes de dispersion améliore les résultats. Cependant, les données sont parfois ambiguës car elles peuvent être expliquées par plusieurs modèles élastiques. En inversion des ondes de surface, tout le spectre de données doit être pris en compte puisque ces ondes sont sensibles aux différents paramètres en profondeur en fonction de la fréquence.

Dans la section 3.4, les aspects théoriques de la FWI sont introduits (Lailly, 1983; Tarantola, 1986; Virieux and Operto, 2009). Les équations de l'état adjoint pour le calcul du gradient de la fonction objective sont développées pour la formulation vitesse-contrainte utilisée dans le chapitre 2. L'estimation de l'ondelette source pour la propagation est en soi un problème inverse. Nous analysons deux approches pour cette estimation : une qui tire profit de la linéarité de ce problème inverse (par rapport à la source) (Pratt, 1999) et une autre qui utilise le gradient de la fonction objective et l'état adjoint (Tarantola, 1984). Enfin, nous présentons une revue bibliographique des fonctions objectives alternatives, car dans le chapitre 4 nous allons développer une approche intermédiaire basée sur une nouvelle fonction objective intermédiaire entre celle de la SWA et celle de la FWI.

3.1 Introduction

Interface waves (Rayleigh, Stoneley) are created at the interface between two media with contrasted P and S wave velocities. Seismic surface waves usually appear at the interface between the Earth and the air (Rayleigh, 1885). These waves can easily be observed because they are often the most energetic seismic events recorded in land datasets. In simple configurations, their amplitude decays as $1/\sqrt{t}$, where t is the traveltime, whereas the decay is as $1/t$ for body waves. Surface waves have been mostly studied in global seismology. For oil industry applications, the number of research works studying surface waves has increased through the last 20 years (Chapter 1). These waves either present challenging issues for imaging methods, or (as it has recently been recognised) they are a great source of information for near surface characterisation (less than a few hundred of metres). Socco et al. (2010) presented a review of the current approaches for the characterisation of the near surface using surface waves.

In this chapter, we present the solution of the inverse problem and its specific application in surface wave analysis (SWA), as well as in Full Waveform Inversion (FWI). These two methods are of interest because they can be used to retrieve the velocity model from surface waves. The first inversion method is used to invert for locally 1D layered velocity models by analysing dispersion curves. The second inversion method is used to retrieve 2D or 3D high-resolution velocity models by inversion of the complete recorded seismic data. We have implemented codes to analyse the 1D method in details. We then use this analysis as a guide to propose an intermediary approach in which we try to take advantage of both the 1D and the FWI methods. The new formulation is presented in Chapter 4 along with applications to synthetic data.

SWA

The penetration depth of surface waves depends on the wavelength. In heterogeneous media, surface waves are dispersive (each frequency propagate at different phase velocity) (Strobbia et al., 2010). In the presence of heterogeneities, surface waves propagate in several modes referred to as fundamental and higher modes (Dunkin, 1965; Ewing et al., 1957). In contrast, the Rayleigh wave (coupled P-SV motion) velocity is unique and slightly smaller than the S-wave velocity in homogeneous media, while Love waves (SH motion only) can only be present in heterogeneous media.

Surface wave propagation is usually analysed through dispersion curves plotted in frequency-velocity ($f - v$) domain. The dispersion curves can be obtained from common shot gathers after application of established processing flows, as explained later in this chapter.

Surface wave velocity for tabular media depends on four layer parameters: P-wave velocity, S-wave velocity, density and layer thicknesses (Thomson, 1950; Haskell, 1953). Consequently, these parameters can potentially be inverted from recorded surface waves (Dziewonski et al., 1969; Hermann, 1973). Xia et al. (1999) presents an inversion approach, taking into account the fundamental mode only, to obtain the S-wave velocity profile. Subsequently, several approaches have been proposed including higher modes in the analysis to improve results considering that inversion is better constrained in this way (Beaty et al., 2002; Xia et al., 2003; Luo et al., 2007; Maraschini et al., 2010). The objective function is usually the least-squares misfit between the dispersion curves extracted

from recorded data and synthetic dispersion curves computed for some specific frequency values. In a 1D layered velocity model, synthetic dispersion curves can be computed (Lai, 1998). This is however less clear in 2D or 3D models. For this reason, SWA is usually restricted to locally 1D models.

As convergence approach, local search methods are more commonly used. However, global search methods have also been used, as for example, the Monte-Carlo approach proposed by Maraschini and Foti (2010). Some lateral variations (pseudo-2D inversion) can been handled using laterally constrained inversion or by applying spatial interpolation of 1D results (Luo et al., 2009; Socco et al., 2009; Vignoli et al., 2011).

FWI

FWI is not restricted to 1D velocity models. The misfit to be minimised is the least-squares distance between the recorded shot gathers and the synthetic shot gathers (Lailly, 1983; Tarantola, 1984; Virieux and Operto, 2009). As FWI is a general method, it should in principle also handle surface waves in addition to transmitted and reflected waves. The expected velocity model resolution is then higher than picking-based techniques such as ray-based tomography or SWA.

FWI is commonly formulated within the local optimisation theory because otherwise inversion would be excessively demanding of computational resources (Tarantola, 1986). Compared to SWA, where dispersion curves should be picked, FWI also considers both the phase and the amplitude of the signal. However, FWI objective function is also known to have a large number of local minima. The non-linearity of seismic inversion can be reduced in FWI by implementing multi-scale approaches and time windowing (Bunks et al., 1995; Sirgue and Pratt, 2004; Brossier et al., 2009), at least for transmitted or reflected waves.

Including surface waves in inversion is a challenging issue (Gélis et al., 2007). One solution consists in gradually including them by considering larger time windows (Romdhane et al., 2011). Working with real data seems to be more challenging as, for the moment, only the low frequency components have been successfully inverted (Shafer et al., 2004). Other ways of reducing non-linearity, and the inherent dependence on the choice of initial model, consist in considering alternative objective functions (Shin and Cha, 2008; Liu et al., 2011; Bozdag et al., 2011).

This chapter is organised as follows. First, we present the theory behind the seismic inverse problem solution (Section 3.2). Second, we describe the classical method for 1D inversion of surface waves illustrated with some examples (Section 3.3). Finally, we recap the theory of FWI in elastic isotropic media (Section 3.4). These elements are important for understanding the alternative method presented in Chapter 4.

3.2 Inverse problem theory

The seismic inverse problem consists of exploring the model space to find the model parameter values that minimise a misfit functional between recorded and modelled data (Tarantola, 2005). If the least-squares misfit is considered, then the objective function to

be minimised is

$$\Phi(\mathbf{m}) = \frac{1}{2}\int (\mathbf{d}(\mathbf{m}|\mathbf{x},t) - \mathbf{d}^{obs}(\mathbf{x},t))^T \mathbf{P}(\mathbf{x},t)(\mathbf{d}(\mathbf{m}|\mathbf{x},t) - \mathbf{d}^{obs}(\mathbf{x},t))\mathrm{d}t, \tag{3.1}$$

where $\Phi(\mathbf{m})$ represents the data misfit value for a model parameter $\mathbf{m}(\mathbf{x})$, $\mathbf{d}(\mathbf{m}|\mathbf{x},t)$ is the synthetic data modelled for $\mathbf{m}(\mathbf{x})$, $\mathbf{d}^{obs}(\mathbf{x},t)$ is the recorded data, $\mathbf{P}(\mathbf{x},t)$ is a weighting matrix that scales the relative contribution of each component, $\mathbf{x}$ are the spatial coordinates, t represents the time and T denotes matrix transpose. The weighting matrix is defined as $\mathbf{P} = \mathbf{L}^T\,\mathbf{L}$, where $\mathbf{L}$ is a diagonal matrix (Xia et al., 1999; Brossier et al., 2009). The data are obtained as a result of a physical experiment whose behaviour is usually non-linear. The whole set of equations representing the physics of the seismic experiment can be associated with the model parameters by using a non-linear operand as

$$\mathbf{d}(\mathbf{m}|\mathbf{x},t) = \mathbf{g}(\mathbf{m}(\mathbf{x}),t), \tag{3.2}$$

where $\mathbf{g}$ is the non-linear forward operand relating the data $\mathbf{d}(\mathbf{m}|\mathbf{x},t)$ and the model $\mathbf{m}(\mathbf{x})$. For example, the operand $\mathbf{g}$ can represent the elasto-dynamic equations that describe the wave propagation in 2D elastic isotropic media shown in Chapter 2 (equation 2.2 in Section 2.2.1).

To minimise the objective function in equation 3.1, the model space can be explored using global or local optimisation methods. In global optimisation, a pseudo-random exploration is used to find the best solution within the considered part of the model space (e.g. Monte Carlo method, Tarantola (2005)). In local optimisation, the model parameters are retrieved starting from an initial set of model parameters. Local optimisation methods are computationally better suited than global methods for seismic waveform inversion as the seismic model space is defined by several thousands of parameters, making computation too expensive for a global exploration of the model space. One of the consequences of this solution is the existence of secondary minima (mainly originated by the non-linear nature of the problem). Thus, the local search must ideally be started in the vicinity of the objective function global minimum.

In this section, we describe the solution of the non-linear inverse problem with an iterative local search approach. We have made an effort to preserve generality so that both SWA and FWI can be explained with the same theory. There are two approaches to update the model parameter (equation 3.8): (1) by employing the gradient $\mathbf{G}(\mathbf{m}^k)$ and the Hessian inverse $\left[\mathbf{H}(\mathbf{m}^k)\right]^{-1}$, or (2) by employing the Jacobian $\mathbf{J}_k$ and the Jacobian derivative $\partial\mathbf{J}_k^T/\partial\mathbf{m}^T$. The corresponding formulations are provided in the following paragraphs.

3.2.1 Inverse problem solution

The non-linear nature of the direct problem prevents us from estimating the model parameter in only one iteration. The inverse problem is then solved iteratively to minimise the value of $\Phi(\mathbf{m})$ (equation 3.1). To compute the model parameter update, a linear *approximation* is applied to the misfit functional through truncation of its associated Taylor series (Pratt et al., 1998). The related equations are obtained as follows.

The model parameter at iteration $k+1$ is expressed as the sum of a background model parameter plus a perturbation as

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \delta\mathbf{m}, \tag{3.3}$$

where $\mathbf{m}^k$ is the model parameter at iteration k and $\delta\mathbf{m}$ is the model parameter perturbation. The misfit functional is represented through the Taylor series expansion as

$$\begin{aligned}\Phi(\mathbf{m}^{k+1}) &= \Phi(\mathbf{m}^k + \delta\mathbf{m}) \\ &= \Phi(\mathbf{m}^k) + \delta\mathbf{m}^T\mathbf{G}(\mathbf{m}^k) + \frac{1}{2}\delta\mathbf{m}^T\mathbf{H}(\mathbf{m}^k)\delta\mathbf{m} + O(\delta\mathbf{m}^3),\end{aligned} \tag{3.4}$$

where

$$\mathbf{G}(\mathbf{m}^k) = \frac{\partial\Phi(\mathbf{m}^k)}{\partial\mathbf{m}} \tag{3.5}$$

and

$$\mathbf{H}(\mathbf{m}^k) = \frac{\partial\mathbf{G}(\mathbf{m}^k)}{\partial\mathbf{m}} = \frac{\partial^2\Phi(\mathbf{m}^k)}{\partial\mathbf{m}^2} \tag{3.6}$$

The term $\mathbf{G}(\mathbf{m})$ is referred to as (objective function) gradient and $\mathbf{H}(\mathbf{m})$ as Hessian matrix. The higher order terms $O(\delta\mathbf{m}^3)$ *are not* equal to zero except for quadratic functions. The objective function reaches its minimum when its first order derivative is equal to zero. This condition gives

$$\frac{\partial\Phi(\mathbf{m}^{k+1})}{\partial\mathbf{m}} = \mathbf{G}(\mathbf{m}^k) + \mathbf{H}(\mathbf{m}^k)\delta\mathbf{m} = \mathbf{0}, \tag{3.7}$$

where $\mathbf{0}$ is the zero vector. In equation 3.7, we can isolate $\delta\mathbf{m}$ and then replace it into equation 3.3 to obtain the *model parameter update*:

$$\mathbf{m}^{k+1} = \mathbf{m}^k - \left[\mathbf{H}(\mathbf{m}^k)\right]^{-1}\mathbf{G}(\mathbf{m}^k). \tag{3.8}$$

The objective function (equation 3.1) can also be differentiated with respect to the model parameter to obtain formula in terms of the Fréchet derivatives for the terms in equation 3.8. The first order derivative yields

$$\mathbf{G}(\mathbf{m}^k) = \int \mathbf{J}_k^T\mathbf{P}\left(\mathbf{d}(\mathbf{m}^k) - \mathbf{d}^{obs}\right)\mathrm{d}t, \tag{3.9}$$

where

$$\mathbf{J}_k = \frac{\partial\mathbf{d}(\mathbf{m}^k)}{\partial\mathbf{m}} \tag{3.10}$$

is the Jacobian matrix at iteration k. The second order derivative yields

$$\mathbf{H}(\mathbf{m}^k) = \int\left(\left[\frac{\partial\mathbf{J}_k^T}{\partial\mathbf{m}^T}\right]\mathbf{P}\left[\left(\mathbf{d}(\mathbf{m}^k) - \mathbf{d}^{obs}\right)\ldots\left(\mathbf{d}(\mathbf{m}^k) - \mathbf{d}^{obs}\right)\right] + \mathbf{J}_k^T\mathbf{P}\mathbf{J}_k\right)\mathrm{d}t, \tag{3.11}$$

where the special notation for the Jacobian derivative defined by Pratt et al. (1998) has been used.

3.2.2 Adjoint-state method

In seismic inversion, the adjoint-state method is used to compute the gradient of a misfit functional (Plessix, 2006). This method can be used to numerically evaluate the gradient (equation 3.5) without needing explicit expressions of the Fréchet derivatives.

In this paragraph, we will work with the general wave equation given in Section 2.2.1 (equation 2.1). The misfit is considered to be a functional of the wave field $\mathbf{u}(\mathbf{x},t)$ and is expressed as

$$\Phi(\mathbf{u}) = \frac{1}{2}\int (\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x},t) - \mathbf{d}^{obs}(\mathbf{x},t))^T \mathbf{P}(\mathbf{x},t)(\mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x},t) - \mathbf{d}^{obs}(\mathbf{x},t))\mathrm{d}t, \tag{3.12}$$

where $\mathbf{K}(\mathbf{x})$ is used to select the signals at receiver positions as $\mathbf{d}(\mathbf{x},t) = \mathbf{K}(\mathbf{x})\mathbf{u}(\mathbf{x},t)$.

We use the Lagrangian formulation presented by Plessix (2006). The associated augmented Lagrangian is

$$\mathcal{L}(\mathbf{u},\boldsymbol{v},\mathbf{m}) = \Phi(\mathbf{u}) - \langle \boldsymbol{v}(\mathbf{x},t), \mathcal{C}(\mathbf{m}|\mathbf{x},t)\mathbf{u}(\mathbf{x},t) - \mathbf{s}(\mathbf{x},t)\rangle\,, \tag{3.13}$$

where $\boldsymbol{v}$ is the adjoint-state variable. The differentiation of the augmented Lagrangian gives

$$\frac{\mathrm{d}\mathcal{L}(\mathbf{u},\boldsymbol{v},\mathbf{m})}{\mathrm{d}\mathbf{m}} = \frac{\partial\mathcal{L}(\mathbf{m})}{\partial\mathbf{m}} + \frac{\partial\mathbf{u}}{\partial\mathbf{m}}\frac{\partial\mathcal{L}(\mathbf{u})}{\partial\mathbf{u}} + \frac{\partial\boldsymbol{v}}{\partial\mathbf{m}}\frac{\partial\mathcal{L}(\boldsymbol{v})}{\partial\boldsymbol{v}}, \tag{3.14}$$

In order to avoid the computation of the Fréchet derivatives (e.g. $\partial\mathbf{u}/\partial\mathbf{m}$), we suppose $\partial\mathcal{L}(\mathbf{u})/\partial\mathbf{u} = \mathbf{0}$ and $\partial\mathcal{L}(\mathbf{u})/\partial\boldsymbol{v} = \mathbf{0}$. It implies that $\mathrm{d}\mathcal{L}/\mathrm{d}\mathbf{m} = \partial\mathcal{L}/\partial\mathbf{m}$. The condition $\partial\mathcal{L}(\mathbf{u})/\partial\boldsymbol{v} = \mathbf{0}$ means that it should satisfy the state equation (forward propagation). $\partial\mathcal{L}(\mathbf{u})/\partial\mathbf{u} = \mathbf{0}$ leads to

$$\mathcal{C}^*(\mathbf{m}|\mathbf{x},t)\boldsymbol{v}(\mathbf{x},t) - \frac{\partial\Phi}{\partial\mathbf{u}} = \mathbf{0}, \tag{3.15}$$

where $\mathcal{C}^*(\mathbf{m}|\mathbf{x},t)$ is the adjoint-state equation. It is straightforward to associate equation 3.15 with the propagation system in equation 2.1. $\boldsymbol{v}(\mathbf{x},t)$ can be computed with a wave propagation code only if $\mathcal{C}(\mathbf{m}|\mathbf{x},t)$ is auto-adjoint (with $\partial\Phi/\partial\mathbf{u}$ as wave source). Because both second and third terms in the right-hand side of equation 3.14 are equal to zero, the gradient is defined as

$$\mathbf{G}(\mathbf{m}) = \frac{\partial\mathcal{L}}{\partial\mathbf{m}} = -\left\langle \boldsymbol{v}(\mathbf{x},t), \frac{\partial\mathcal{C}(\mathbf{m}|\mathbf{x},t)}{\partial\mathbf{m}}\mathbf{u}(\mathbf{x},t)\right\rangle. \tag{3.16}$$

For auto-adjoint systems, the gradient computation needs solving two direct problems: one to obtain $\mathbf{u}(\mathbf{x},t)$ and the other to obtain $\boldsymbol{v}(\mathbf{x},t)$. Moreover, the gradient is given by the special correlation defined in equation 3.16, considering the operator $\partial\mathcal{C}(\mathbf{m}|\mathbf{x},t)/\partial\mathbf{m}$. Explicit formula for the gradient computation in the case of the velocity-stress formulation (Virieux, 1986) is given later in Section 3.3.

3.2.3 Strategy for updating the parameters

The inverse of the Hessian matrix can be interpreted as a preconditioning matrix applied to the gradient and used to improve convergence (Pratt et al., 1998). The Hessian matrix

in equation 3.11 is obtained by retaining the Taylor-series terms up to quadratic order (Tarantola, 2005). The convergence strategy known as the **Newton method** uses the inverse of the exact Hessian.

The inverse of the exact Hessian (equation 3.11) is difficult to compute especially regarding the second-order differential term (Jacobian derivative). However, the convergence can be improved, with respect to gradient-based approaches such as simple gradient and conjugate gradient, if preconditioning with (at least) an approximation of the inverse Hessian matrix is used. This argument is demonstrated by, for example, Pratt et al. (1998) and Métivier et al. (2012).

The approximate Hessian matrix, written as

$$\mathbf{H}_a(\mathbf{m}^k) = \int \mathbf{J}_k^T \mathbf{P} \mathbf{J}_k \, \mathrm{dt}, \tag{3.17}$$

can be used if the second order differential term is sufficiently small to be ignored, which happens when (Tarantola, 2005): "the residuals are small or the forward equation is quasi-linear". Indeed, if the forward equation is linear, then the Jacobian does not depend on the model parameter and the exact Hessian, in such case, corresponds to the definition of equation 3.17. The convergence strategy that implements the Hessian defined in equation 3.17 is known as the **Gauss-Newton method**.

Regularisation may be needed before inverting the Hessian. The damped Hessian is defined as

$$\mathbf{H}_d = \mathbf{H}_a + \alpha \mathbf{I}, \tag{3.18}$$

where α is the damping term. This approach can otherwise be obtained by considering a simple regularisation to form the objective function

$$\Phi_t = \Phi + \alpha \Phi_m, \tag{3.19}$$

where $\Phi = \Phi(\mathbf{m})$ remains as defined in equation 3.1, α represents the damping term and

$$\Phi_m = \frac{1}{2} \left(\mathbf{m} - \mathbf{m}_0\right)^T \mathbf{I} \left(\mathbf{m} - \mathbf{m}_0\right), \tag{3.20}$$

with $\mathbf{m}_0$ being an initial model parameter estimation. Equation 3.20 defines the simplest case of the general Tikhonov regularisation (Tikhonov and Arsenin, 1977), where a specialised weighting matrix is used instead of the identity matrix. The inversion method that implements the damped Hessian is commonly known as the **Levenberg-Marquardt method** (Levenberg, 1944; Marquardt, 1963).

Finally, the **quasi-Newton method** consists in interactively computing an approximated version of the Hessian inverse. In particular, the L-BFGS algorithm (named after Broyden-Fletcher-Goldfarb-Shanno and with the L standing for limited memory, Nocedal (1980)) computes a good approximation, as proven by Métivier et al. (2012), by saving the computed gradient in memory for a few iterations and then performing scalar products and additions between the vectors.

In general, the seismic inverse problem is non-linear. The model is usually updated through a local search either by explicitly computing the Jacobian (equation 3.9) or by employing the adjoint-state method (3.16). We have presented both formulations

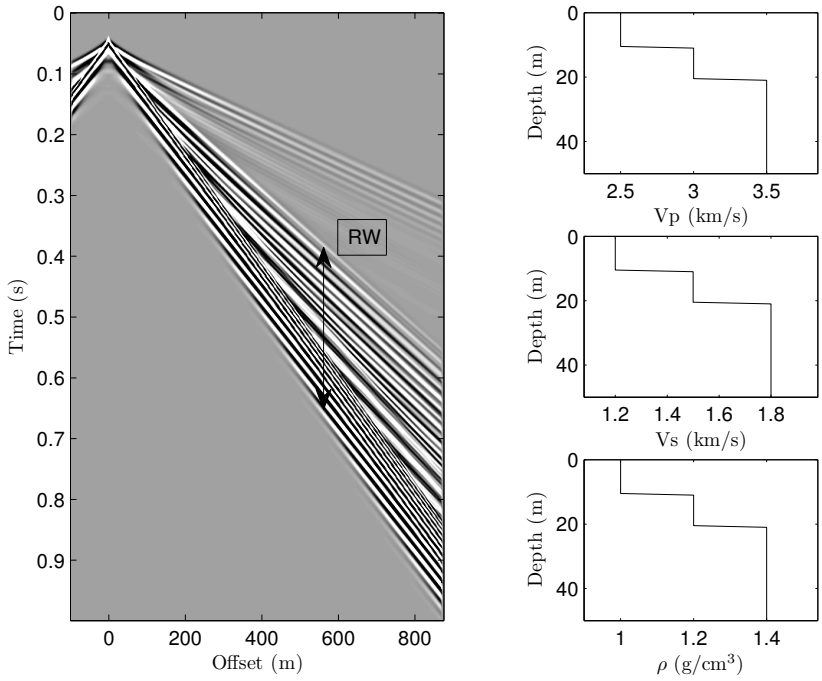


Figure 3.1: Synthetic shot-gather (left) computed for a 1D layered medium (right) using our finite-difference modelling code (Chapter 2). Rayleigh waves in the shot gather are dispersive and are indicated by the double arrow.

because SWA is often based on the first approach while FWI is based on the latter. As strategy for convergence, SWA will be formulated with a regularised method (Levenberg-Marquardt) and FWI with a quasi-Newton method. In the next two sections, we present the application of the theory presented here in SWA and FWI.

3.3 Surface wave analysis (SWA)

In surface wave methods, the data to be inverted are the dispersion curves (Lai, 1998; Xia et al., 1999; Socco and Strobbia, 2004). These curves are a representation of the surface wave velocity with respect to the frequency. This representation is valid for wave propagation in 1D layered velocity models. In this section, we first review the data processing to extract dispersion curves from common shot gathers. Second, we present the classical inversion algorithms together with some examples of fundamental mode inversion and multi-modal inversion. Some of these examples are adapted from the literature, others are proposed here to better illustrate our observations. The objective is to analyse in more details the methods and underline some drawbacks especially regarding ambiguity due to insufficiently constrained inversion.

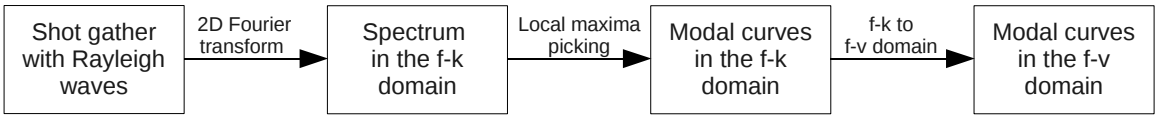


Figure 3.2: Surface wave processing workflow used to identify dispersion curves (modal curves) from shot gathers.

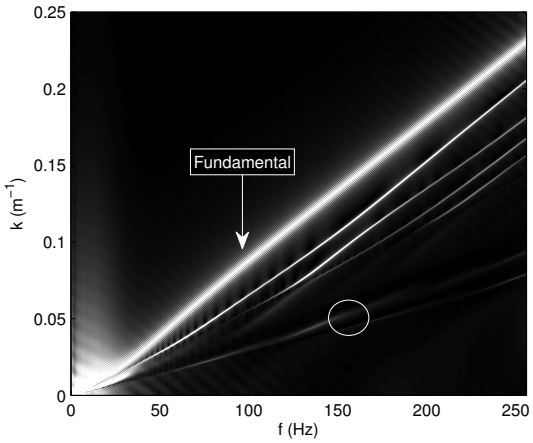


Figure 3.3: Spectrum amplitude for the shot gather in Figure 3.1. In this case, the fundamental mode is easily recognised because of its high energy. The energy indicated by the circle might not be a Rayleigh-wave mode (explanation given in the text).

3.3.1 Processing workflow for dispersion curve extraction

In this section, we show the steps used to build dispersion curves from shot gathers. As an example, we consider the shot gather shown in Figure 3.1, left, (computed for the elastic model shown in Figure 3.1, right) and identify the dispersion curves of Rayleigh waves. The dispersion curves obtained with the processing approach presented here are compared to a theoretical solution for plane-layered elastic media. We obtain this theoretical solution with the code implemented by Lai (1998), which is based on the work of Kennett and Kerry (1979); Luco and Apsel (1983); Chen (1983) and Hisada (1995). More details on surface wave modelling in plane-layered media are given later on when describing direct and inverse problems (section 3.3.2).

The processing steps to get the dispersion curves in the $f-v$ domain are described in the workflow of Figure 3.2. The $f-k$ spectrum is obtained by computing the 2D Fourier transform of the shot gather $d(x,t)$ as

$$D(k,f) = \iint d(x,t)e^{-i2\pi ft}e^{i2\pi kx}\mathrm{d}t\mathrm{d}x, \tag{3.21}$$

where k is the wavenumber, f is the frequency, x is the horizontal Cartesian coordinate.

The maximum offset of the shot gather (Figure 3.1) is $x_{max} = 850$ m and the maximum recorded time is $t_{max} = 1$ s. The sampling intervals are $\Delta x = 2$ m and $\Delta t = 2$ ms, hence,

the Nyquist wavenumber is $k_N = 1/(2\Delta x) = 0.25$ m^{-1} and the Nyquist frequency is $f_N = 1/(2\Delta t) = 250$ Hz. The fundamental mode of Rayleigh-wave propagation is easily recognised in the $f-k$ domain because of its high energy (Figure 3.1). Some other modes can be observed as well. The energy identified with a circle must be analysed in order to define whether it corresponds to surface waves or not. This analysis is done later in this section when presenting the relationship between $f-k$ and $f-v$ domains.

Referring to the workflow in Figure 3.2, the step called "local maxima picking" consists in tracking dispersion curves. In this example, we have used a simple semi-automatic semblance-based (Neidell and Taner, 1971) algorithm that identifies the "location" of each propagation mode in the $f-k$ image (Figure 3.3). The first step is to select each one of the propagation modes by manually picking one single pixel for each mode. Then, the algorithm "tracks" each selected mode by searching the pattern that identifies it within the image. This tracking step is done by a pattern recognition algorithm using the semblance criterium. Once the algorithm has identified the location of all the selected modes, a simple determination of the energy peaks gives the $f-k$ coordinates of the dispersion curves.

The described semi-automatic picking algorithm works properly if the dispersion modes are sufficiently separated in a particular domain. Some model parameter and acquisition geometry configurations yield overlapped propagation modes. In such cases, the extraction of dispersion curves is difficult and should be performed manually. In addition, high resolution of the $f-k$ spectrum might be useful to avoid errors in the semi-automatic or manual identification. In some configurations, only "apparent" dispersion curves can be identified and used in inversion (Strobbia, 2002).

The dispersion curves identified in the $f-k$ domain are transformed to the $f-v$ domain. The expression that relates frequency, wavenumber and velocity is

$$v = \frac{f}{k}. \tag{3.22}$$

This transformation step can be interchanged with the previous one such that the dispersion curve picking is done in the $f-v$ image instead. In that case, if the resolution is not modified then the semi-automatic picking should give similar results but unphysical velocity values could be rapidly identified. In the literature, the identification is done in the $f-k$ domain (Nolet and Panza, 1976; Tselentis and Delis, 1998), in the $f-v$ domain (Xia et al., 1999; Strobbia, 2002; Luo et al., 2007) and also in the $f-p$ (p stands for slowness) domain (McMechan and Yedlin, 1981; Luo et al., 2008, 2009). The latter is done by application of the Radon transform instead of the Fourier transform, which especially accounts for high resolution (Luo et al., 2009).

Equation 3.22 can also be used to analyse the modes identified by a circle in Figure 3.3. The Rayleigh-wave velocity is theoretically defined to be in the range [0.87-0.96]$\times v_s$ for homogeneous media (Strobbia, 2002). For the medium described in Figure 3.1, the Rayleigh-wave velocity related the fundamental mode cannot be faster than 0.96×1800 m/s (1728 m/s) and can not be slower than 0.87×1200 m/s (1044 m/s). For higher modes the maximum Rayleigh-wave velocity is equal to the maximum S-wave velocity 1800 m/s. In Figure 3.4, we plotted the $f-v$ region bounded by the maximum and minimum expected Rayleigh-wave velocities (left) and the corresponding $f-v$ domain (right) for the $f-k$ spectrum of Figure 3.3. Therefore, we can conclude that the

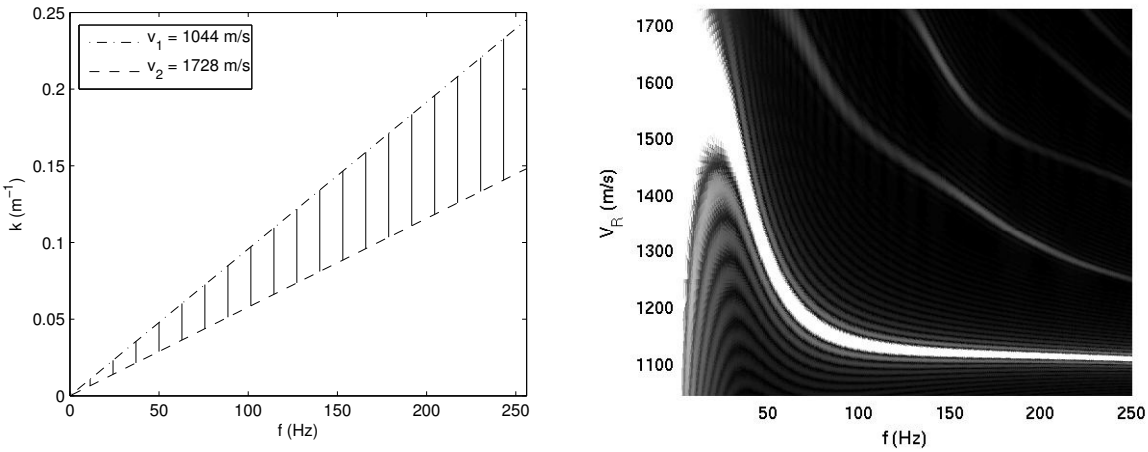


Figure 3.4: Left: the theoretical range of Rayleigh-wave velocities for the model presented in Figure 3.1 is indicated with the vertical lines. Right: the spectrum of Figure 3.3 is transformed to the $f-v$ domain. Note that the $f-v$ domain on the right-hand side image is a stretched and mirrored version of the theoretical range indicated on the left-hand side image.

modes identified with a circle in Figure 3.3 correspond to seismic events different from Rayleigh waves, probably P guided waves (explained by their higher velocity). When working with real data, the theoretical limit of Rayleigh-wave velocity cannot be known without considering *a priori* model information. In such case, high energy and small velocity values should be a natural way to distinguish surface waves from other seismic events in the $f-k$ domain. By contrast, a priori information could be used to enhance Rayleigh waves while attenuating other waves.

The output of the processing workflow described in Figure 3.2 is dispersion curves plotted in the $f-v$ domain. In Figure 3.5, the theoretical dispersion curves are shown for comparison. This solution has been computed with the implementation developed by Lai (1998) (details will be given in section 3.3.2). The extracted from processing and the theoretical curves have similar characters. The greatest mismatch is observed for the fundamental mode at frequencies below 20 Hz. This mismatch comes from a strong stretching applied to the low-frequency part of the spectrum when transforming the data from $f-k$ to the $f-v$ domain (see Figure 3.4). As a result, at lower frequencies the semi-automatic identification algorithm has reduced resolution. For instance, if a constant velocity $v = 1000$ m/s should be picked in the $f-k$ domain, then at $f = 100$ Hz a wavenumber overestimation $\Delta k = 0.01$ m^{-1} yields an error in velocity $\Delta v = -91$ m/s, while at $f = 10$ Hz the same wavenumber overestimation yields an error $\Delta v = -500$ m/s. Therefore, the dispersion curves picked in the $f-k$ domain and then transformed to the $f-v$ domain using equation 3.22 might contain erroneous information in the low frequencies. A velocity decay is also observed at frequencies above 200 Hz. This decay is not physical and is a consequence of grid dispersion in the finite-difference modelling code (Levander, 1988).

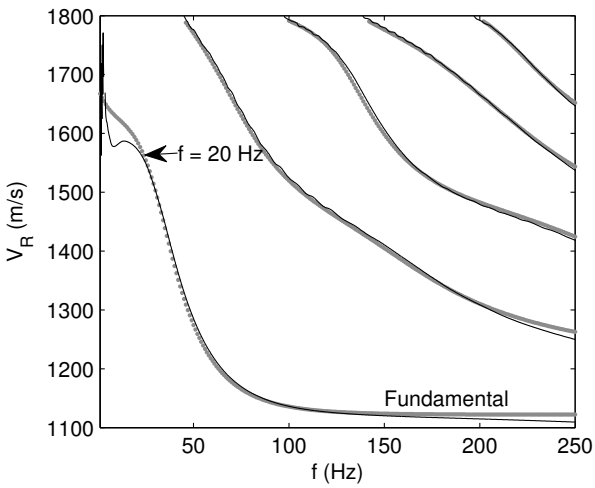


Figure 3.5: Dispersion curves corresponding to the model shown in Figure 3.1. The curves identified with processing (Figure 3.2) are plotted with solid lines. The theoretical curves are plotted with dashed lines.

3.3.2 Inversion of dispersion curves

Surface waves propagate dispersively in heterogeneous media. Dispersion is a property of surface waves whose propagation velocity depends on the frequency. In theoretical studies (Dunkin, 1965; Strobbia, 2002), it was demonstrated that surface waves present an exponential energy decay with depth. Most of the energy propagates within the zone with a maximum depth of $2\times\lambda_s$, λ_s being the shear wavelength. Therefore, surface waves are dispersive and propagate in different dispersion modes only if the medium presents heterogeneities (in P-wave velocity, S-wave velocity or density) within $2\times\lambda_s$ (Dunkin, 1965; Strobbia, 2002). In homogeneous media, Rayleigh waves propagate with constant velocity (Rayleigh, 1885) and Love waves can not exist (Love, 1911).

In this section, the SWA problem is described using the classical Levenberg-Mardquart formulation (Xia et al., 1999; Socco and Strobbia, 2004). The forward and the inverse problem representations are explained. We use some synthetic data examples to show the inversion of fundamental mode as well as the combined inversion of fundamental mode plus higher modes.

Description of the direct and the inverse problem

The forward problem in SWA is defined as computing dispersion curves for a known set of 1D model parameters (P-wave velocity $\mathbf{v}_p$, S-wave velocity $\mathbf{v}_s$, density $\boldsymbol{\rho}$ and layer thickness $\mathbf{h}$). The inverse problem consists in finding the model parameters when the only known information are the recorded surface waves usually represented by picked dispersion curves. (Figure 3.6). Mathematical definitions are given in the next paragraphs We begin by briefly reviewing the previously mentioned dispersion curve computation (Lai, 1998).

The Rayleigh and the Love wave dispersion theory is a solution of the Helmholtz-

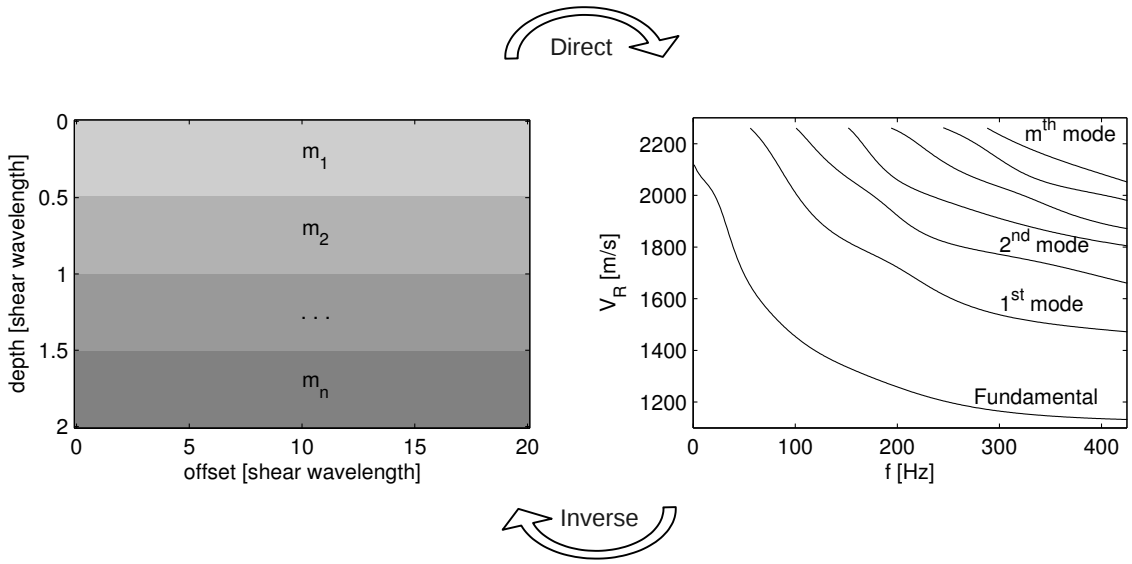


Figure 3.6: The surface-wave problem in a layered medium. Left: A layered model in the zone of Rayleigh wave propagation ($2 \times \lambda_s$). Each layer is identified with model parameters $\mathrm{m}_i = (\mathrm{v}_{pi}, \mathrm{v}_{si}, \rho_i, \mathrm{h}_i)$ for $i = 1, 2, ..., n$. Right: the dispersion curves for the Rayleigh waves propagating in this medium.

theorem based representation of the Navier equation of elasticity (Aki and Richards, 1980) (page 63). The velocity, displacement and energy of Rayleigh waves can be analytically obtained in homogeneous media (Strobbia, 2002). In such media, the Rayleigh wave velocity depends on the Poisson ratio and ranges over $v_R = [0.87; 0.96] \times v_s$, where the maximum value is reached for a Poisson ratio of 0.5.

In heterogeneous 1D layered media, the Rayleigh and the Love wave velocities are found by solving an eigenvalue problem developped as a secular equation. In general, the mathematical representation of the "dispersion equation" is

$$\mathbf{D}(c, k) = \det|\mathbf{U}^T\mathbf{T}\mathbf{V}| = \mathbf{0}, \tag{3.23}$$

where $\mathbf{U}$ and $\mathbf{V}$ represent the boundary conditions (free-surface and infinite half-space conditions) and $\mathbf{T}$ is the propagator matrix imposing conditions of stress and displacement continuity across the layers. The eigenvalues are then the only non-trivial solutions of the dispersion equation (secular function) and those (frequency and wavenumber) couple of values describe each dispersion mode.

A review of the development of numerical methods for dispersion curve computation was presented by Buchen and Ben-Hador (1996). Thomson (1950) and Haskell (1953) pioneered the theory of Love and Rayleigh wave dispersion in plane-layered models. Later on, in the 60's and 70's, several developments were proposed either for correcting the high-frequency instabilities of the original method (Pestel and Leckie, 1963; Dunkin, 1965; Thrower, 1965; Abo-Zena, 1979), or for introducing fundamental modifications in its implementation (Knopoff, 1964; Schwab and Knopoff, 1970). Such unstable effects were subsequently avoided by employing the so-called Reflection-Transmission matrix method (Kennett and Kerry, 1979; Luco and Apsel, 1983). Lai (1998) has implemented

the Reflection-Transmission method based on the revisitation made by Chen (1983) and Hisada (1995). This is the code, we have employed for dispersion curve computation in this section. The computational cost remains low mainly because there are only few model parameters in comparison to waveform modelling.

In our study, the (non-linear) forward problem is stated as

$$\mathbf{d}_c(f|\mathbf{m}) = \mathbf{g}(\mathbf{m}), \tag{3.24}$$

where $\mathbf{d}_c(f|\mathbf{m})$ represents the dispersion curves computed for a 1D layered model, $\mathbf{g}(\mathbf{m})$ is the non-linear operand for the direct problem and $\mathbf{m}$ represents the four model parameters. Equation 3.24 can be discretised for a finite number l of frequencies $f_j = (f_1, f_2, ..., f_l)$ and a finite number n of layers, such that, for instance, the S-wave velocities are $\mathrm{v}_{si} = (\mathrm{v}_{s1}, \mathrm{v}_{s2}, ..., \mathrm{v}_{sn})$.

In the inverse problem, the objective function to be minimised is

$$\Phi(\mathbf{m}) = \frac{1}{2}(\mathbf{d}_c(f_j|\mathbf{m}) - \mathbf{d}_c^{obs}(f_j))^T\mathbf{P}(\mathbf{d}_c(f_j|\mathbf{m}) - \mathbf{d}_c^{obs}(f_j)), \tag{3.25}$$

where $\mathbf{d}_c^{obs}(f_j)$ represents the (observed) dispersion curves picked from recorded data. To update the model parameter, the Gauss-Newton and the Levenberg-Marquardt methods are usually preferred (see Section 3.2.3 for details on these methods). We use here the Levenberg-Marquardt method as presented by Xia et al. (1999). The model parameter is updated as

$$\mathbf{m}^{k+1} = \mathbf{m}^k - \left(\mathbf{J}_k^T\mathbf{P}\mathbf{J}_k + \alpha\mathbf{I}\right)^{-1}\mathbf{J}_k^T\mathbf{P}\left(\mathbf{d}_c^k(f_j) - \mathbf{d}_c^{obs}(f_j)\right). \tag{3.26}$$

where $\mathbf{J}_k = \partial\mathbf{d}(f_j)/\partial\mathbf{m}^k$ is the Jacobian matrix computed for discretised dispersion curves in a layered model and $\mathbf{I}$ is the $n \times n$ identity matrix.

The model parameters reconstructed in inversion are usually the S-wave velocity $\mathbf{v}_s$ and/or thickness $\mathbf{h}$ (Socco and Strobbia, 2004; Socco et al., 2010). The other two parameters ($\mathbf{v}_p$ and $\boldsymbol{\rho}$) cannot be easily inverted because surface waves are mostly sensitive to $\mathbf{v}_s$ and $\mathbf{h}$. Conclusions regarding sensitivity have been presented by Song et al. (1989) and Xia et al. (1999) in the case of Rayleigh waves.

Besides local search methods, like the one presented here, global search methods have also been employed because of the low forward-modelling computational cost. For example, a Monte-Carlo approach is used by Maraschini and Foti (2010).

Examples

We illustrate the sensitivity matrix analysis, the inversion algorithm, and the inversion results of 9 different problems. Inversion was set to retrieve the S-wave velocity or the thickness. The input data are either the fundamental mode only or the fundamental mode plus higher modes (multi-modal inversion). We start from the simple three-layer profile of Figure 3.1 and end with a 7-layer model presented by Luo et al. (2007). We show that in complex models the initial model should be well constrained. Finally, a test demonstrates that the data can be completely explained by a velocity model different from the exact one.

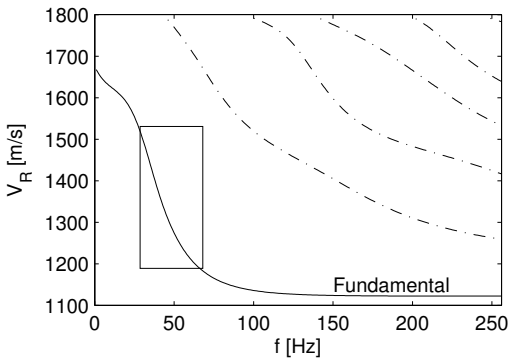


Figure 3.7: Dispersion curves for the model of Figure 3.1. The box indicates the region of chosen frequencies for a fundamental mode sensitivity analysis.

Sensitivity matrix: We analyse the sensitivity of the Rayleigh-wave fundamental mode for model in Figure 3.1. We compute the theoretical dispersion curves using the implementation of Lai (1998). We define a set of 35 frequencies with uniformly-sampled values in the range (25-60) Hz. One (fundamental mode) or two (fundamental and one higher modes) Rayleigh-wave dispersion modes are present for each frequency value (see Figure 3.7) but we only consider the fundamental mode in this sensitivity analysis.

The sensitivity matrix corresponds here to the Jacobian matrix representing the derivative of the fundamental mode (represented by $\mathbf{d}_c(f_j)$, where $j = 1, ..., 35$) with respect to the S-wave velocities (v_{si}, where $i = 1, 2, 3$ and v_{s1} is the shallowest layer velocity). The size of the Jacobian matrix is 35×3. We compute the partial derivatives using second-order finite differences as

$$\mathbf{J} = \frac{\partial \mathbf{d}_c(f_j)}{\partial \mathrm{v}_{si}} \approx \frac{\mathbf{d}_c^+(f_j) - \mathbf{d}_c^-(f_j)}{0.002 \times \mathrm{v}_{si}} \tag{3.27}$$

where the $\mathbf{d}_c^+(f)$ and $\mathbf{d}_c^-(f)$ represent the velocities of the fundamental mode computed for $(1 + 0.001)\mathrm{v}_{si}$ and $(1 - 0.001)\mathrm{v}_{si}$, respectively, and sampled at the chosen frequency values. The normalised Jacobian matrix (equation 3.27 divided by its norm) is plotted in Figure 3.8. The three lines in this Figure represent the sensitivity of the Rayleigh-wave fundamental-mode velocity with respect to the three S-wave velocities for the chosen range of frequencies. This sensitivity analysis shows some important properties of Rayleigh waves in layered media: as expected, the higher frequencies are more sensitive to v_{s1}, and thus, tend to propagate with low velocity. The lower frequencies are more sensitive to v_{s3}, and thus, tend to propagate faster.

Inversion algorithm and example: We present the inversion algorithm using the Marquardt-Levenberg method and we apply it to retrieve the S-wave velocities for the model depicted in Figure 3.1). The other three model parameters are considered known and equal to the exact profile. The observed data are a set of 15 Rayleigh-wave velocities sampled from the fundamental mode (represented by $\mathbf{d}_c(f)$) corresponding to 15 uniformly-spaced frequencies into the range (25-60) Hz (Figure 3.7). The inversion algorithm consists of the following steps:

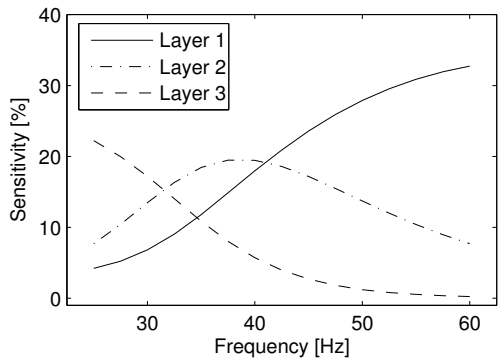


Figure 3.8: Sensitivity analysis for the Rayleigh-wave fundamental mode of Figure 3.7 with respect to the S-wave velocities. The velocity model is shown in Figure 3.1.

1. *Select a good initial model:* the initial profile velocity values v_{si}^0 differ from the exact profile values up to 50%. The set of initial values is calculated as $\mathrm{v}_{s1}^0 = 0.8 \times \mathrm{v}_{s1} = 960$ m/s, $\mathrm{v}_{s2}^0 = 1.2 \times \mathrm{v}_{s2} = 1800$ m/s, and $\mathrm{v}_{s3}^0 = 1.5 \times \mathrm{v}_{s3} = 2700$ m/s.

2. *Compute the dispersion curves:* the Rayleigh-wave fundamental mode is computed using the theoretical solution for the model parameters at iteration k. The corresponding Rayleigh-wave dispersion curve is $\mathbf{d}_c^k(f_j)$.

3. *Compute the Jacobian matrix:* the approximated Jacobian matrix for the current dispersion curve $\mathbf{d}_c^k(f)$ and model v_{si}^k is computed using equation 3.27. For the three-layer model, 6 forward modelling problems are computed because a second-order finite difference approximation was considered.

4. *Update the model parameter:* the S-wave velocity model is updated using equation 3.26 for $\mathbf{m}^k = \mathrm{v}_{si}^k$. The damping factor α is exponentially decreasing with iterations, such that, at the final iterations the variations in the model parameter are unpenalised and the misfit value may descend to zero (only in ideal cases).

The algorithm iterates from step 2 to 4 until the S-wave velocity model presents small variations with respect to the previous iteration. For the current example, 5 iterations are necessary to correctly converge. The chosen α values and the v_{si}^k through iterations are shown in Figure 3.9.

The computational cost of the inversion algorithm is proportional to the number of forward problem computations performed in step 3. Here, the cost of the forward problem depends on the number of model layers. As we were interested in the analysis for a small number of parameters, we have not defined more efficient approaches (e.g. adjoint-state method).

Estimation of 5 and 7 layer S-wave velocity using the fundamental mode only: We have carried out S-wave velocity inversions for two different models: (1) a 5-layer model (Table 3.1) and (2) a 7-layer model (Table 3.2). $\mathbf{v}_p$, $\boldsymbol{\rho}$ and $\mathbf{h}$ are considered known and equal to the exact ones. We consider the 7-layer model as more complex than the 5-layer one. The purpose behind these two examples is to show that convergence is usually achieved only if the choice of initial velocity model is well constrained. Furthermore, a damping factor dynamically changing with iterations can improve convergence.

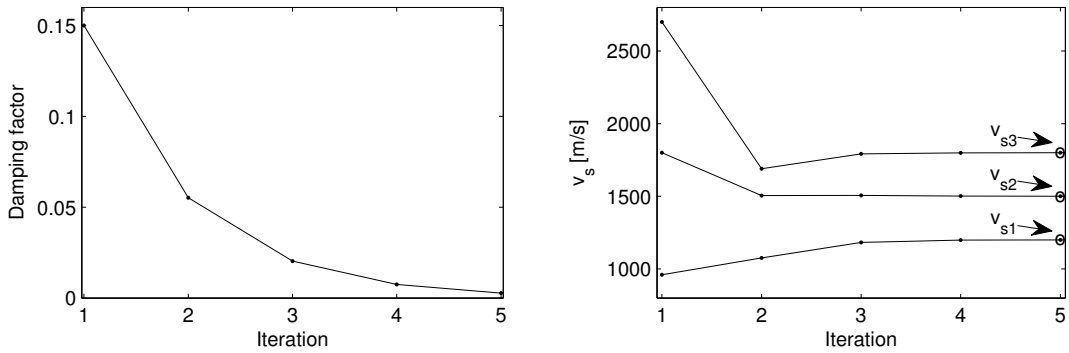


Figure 3.9: Damping factor (left) and S-wave velocities (right) for each iteration using the Marquardt-Levenberg method for surface-wave inversion applied to the model of Figure 3.1.

Layer number	$\mathbf{v}_s$ (m/s)	$\mathbf{v}_s$ (m/s)	$\boldsymbol{\rho}$ (kg/m^3)	$\mathbf{h}$ (m)
1	1200	2500	1000	6
2	1500	3000	1200	9
3	1800	3500	1400	13
4	2040	4125	1500	15
5	2160	4250	1600	∞

Table 3.1: Elastic parameters for a 1D 5-layer medium.

Layer number	$\mathbf{v}_s$ (m/s)	$\mathbf{v}_p$ (m/s)	$\boldsymbol{\rho}$ (kg/m^3)	$\mathbf{h}$ (m)
1	800	1500	750	5
2	1000	2300	1000	9
3	1440	3000	1200	11
4	1680	3500	1400	12
5	2040	4125	1500	15
6	2160	4250	1600	18
7	2400	4500	1750	∞

Table 3.2: Elastic parameters for a 1D 7-layer medium.

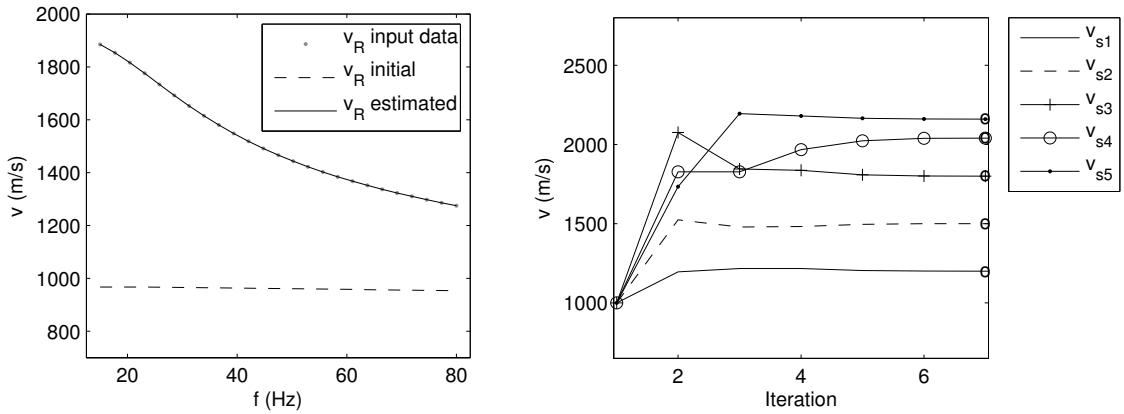


Figure 3.10: 5-layer model (Table 3.1) inversion with constant damping factor and no weighting. Left: fundamental mode of Rayleigh-wave velocity. Right: S-wave velocities for each iteration using the Marquardt-Levenberg inversion method.

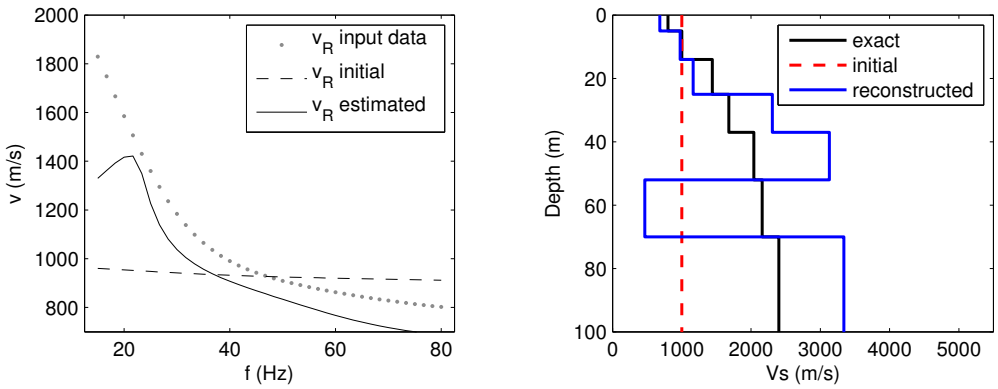


Figure 3.11: 7-layer model (Table 3.2) inversion *without weighting or regularisation*. Left: the data are 25 samples of the fundamental mode of Rayleigh-wave velocity. Right: S-wave profiles.

For the 5-layer medium, the observed dispersion curve contains 25 values of the fundamental mode corresponding to 25 uniformly-sampled frequencies within the range 15-80 Hz (Figure 3.10, left). The initial S-wave velocity model is homogeneous with $\mathrm{v}_s^0(l_i) = 1000$ m/s, for $i = 1, 2, ..., 5$. The inversion algorithm correctly converged to the exact model velocities after 7 iterations (Figure 3.10, right). The initial velocity model is easy to be chosen because the exact model is simple (5 layers) in addition to known $\mathbf{v}_p$ and $\boldsymbol{\rho}$.

For the 7-layer medium, the observed data are the fundamental mode sampled with the same frequency range used in the 5-layer case (Figure 3.11, left). The initial S-velocity model is homogeneous ($\mathrm{v}_s^0(l_i) = 1000$ m/s). In this case, inversion diverged when regularisation and weighting were not included (Figure 3.11). We have computed the objective function for several models whose velocities range from the initial S-velocity model to the exact one (Figure 3.12). As there are no secondary minima (at least in the considered velocity range), the non-regularised inversion probably overestimated the

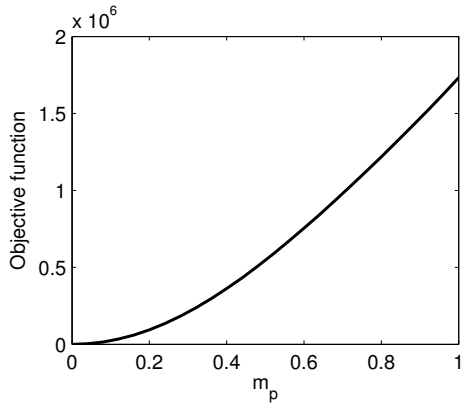


Figure 3.12: Objective function when the initial Vs model is controlled by one single model parameter. For m_p=0, the initial model is the exact one. For m_p=1, the initial model is the constant model shown in 3.11. No local minima is observed in this range.

model update based on the strong misfit at low frequencies which leads to instability. Furthermore, convergence is achieved when considering dynamic weighting and dynamic regularisation (Figure 3.13).

The strategy consists in including heuristic criteria to chose $\mathbf{L}$ and α in inversion. The diagonal values of $\mathbf{L}$ ($\mathbf{L}_j$) are related to the misfit for each frequency value. The objective is to decrease excessively high misfit contributions like, for example, the misfit of low frequencies in Figure 3.11. At each iteration the weighting values are computed as

$$\mathbf{L}_j = \text{minval}\left\{0.1; \frac{r_{max}}{\mathbf{d}_c^k(f_j) - \mathbf{d}_c^{obs}(f_j)}\right\}, \tag{3.28}$$

where 'minval' means minimum value and $r_{max} = \text{maxval}\{\mathbf{d}_c^k(f_j) - \mathbf{d}_c^{obs}(f_j)\}$ is the maximum misfit value. A minimum weight value (0.1) is included because too small misfit contributions might lead to neglecting the related residuals whereas we want to fit the complete data set.

For the damping factor, we have considered that it should be smaller when approaching convergence, i.e., objective function global minimum and smaller gradient projection. Each iteration starts with $\alpha_0^k = |\bar{\mathbf{G}}_k|\Phi(\mathbf{m}^k)$, where $\mathbf{G}_k = \mathbf{J}_k^T\mathbf{L}(\mathbf{d}_c^k(f_j) - \mathbf{d}_c^{obs}(f_j))$. This starting factor is simply the average of the absolute value of the gradient multiplied by the objective function. The damping factor is then linearly decreased ($\alpha^k = 10^{-p}\alpha_0^k$) such that the objective function is minimised for one p ranging from 0 to 20.

We analyse an additional result with the 7-layer model. We have switched the order of layers 3 and 4, and 5 and 6, to test the inversion algorithm in the presence of velocity inversions. Referring to Table 3.2, the order of layers for this test is: 1, 2, 4, 3, 6, 5, 7. The inversion converged when using the dynamic damping and weighting (Figure 3.14). In this particular case, it has been proved that this inversion method can retrieve the correct velocity model even in the presence of velocity inversions in the model.

The Marquardt-Levenberg method with dynamic damping and weighting and inverting only the fundamental mode of Rayleigh waves yields excellent results when the 1D inverse surface-wave problem is well constrained: initial velocity model sufficiently close

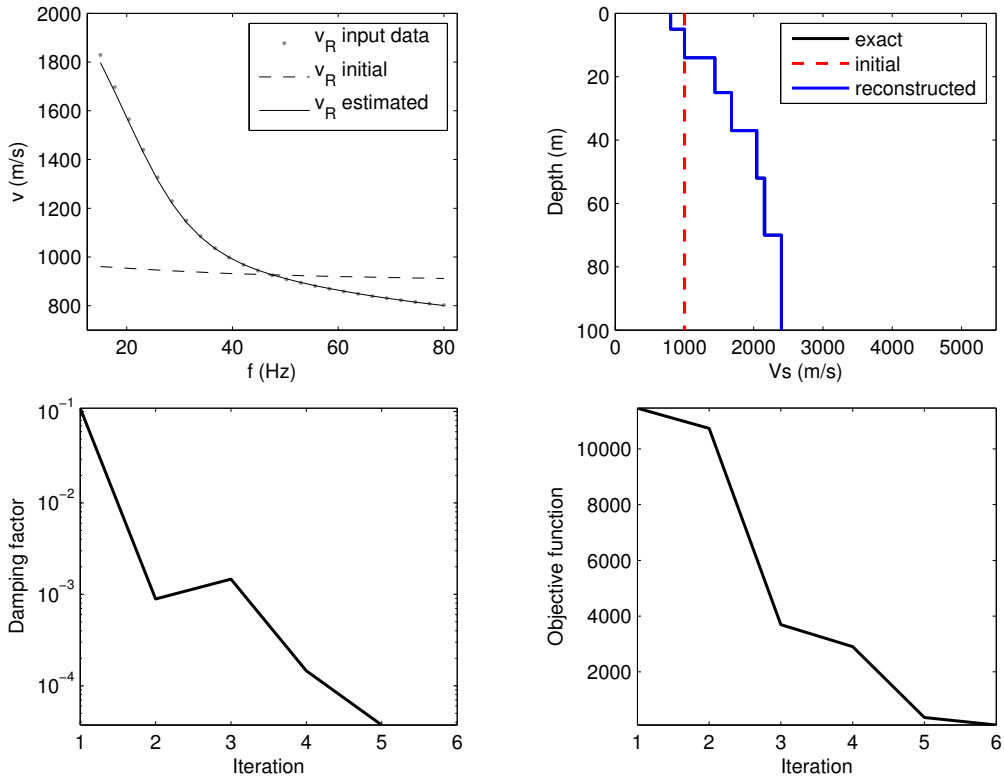


Figure 3.13: 7-layer model (Table 3.2) inversion *including dynamic weighting and dynamic regularisation*. Top: data (left) and S-wave profiles (right). Bottom: dynamic damping factor (left) and objective function (right) with iterations.

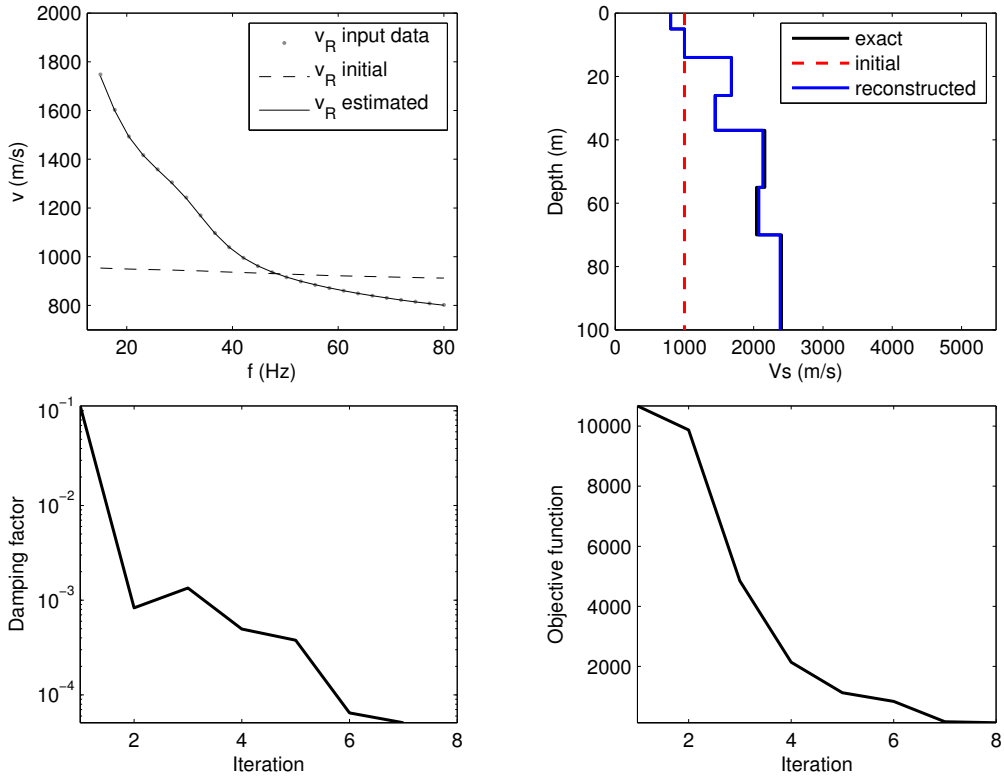


Figure 3.14: 7-layer model (Table 3.2) inversion when the order of layers is changed to 1, 2, 4, 3, 6, 5, 7. *Dynamic weighting and dynamic regularisation are included.* Top: data (left) and S-wave profiles (right). Bottom: dynamic damping factor (left) and objective function (right) with iterations.

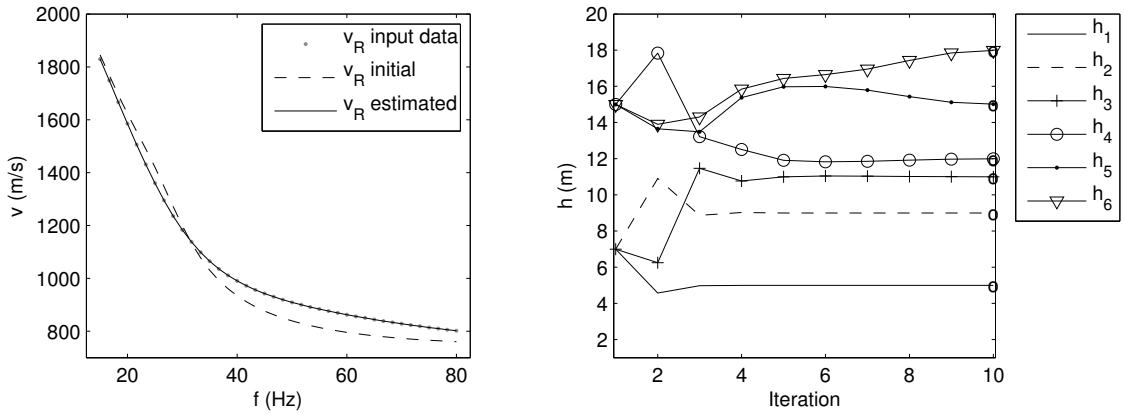


Figure 3.15: 7-layer model (Table 3.2) *thickness* inversion. Left: the data are 25 samples of the fundamental mode of Rayleigh-wave velocity. Right: thickness values through iterations.

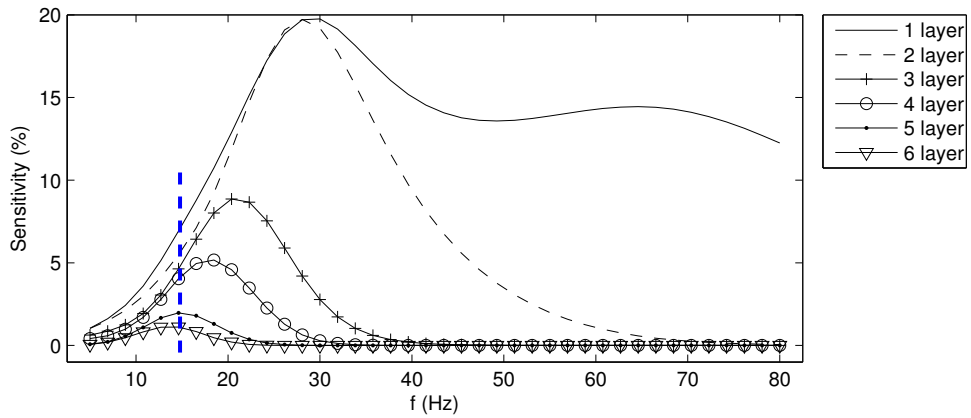


Figure 3.16: Sensitivity analysis of the Rayleigh-wave velocity with respect to variations in thickness for the model of Table 3.2. The dashed blue line indicates the beginning of the chosen frequencies for the inversion process (15-80 Hz).

to the exact one, a non-overlapped fundamental mode and good a priori information (P-wave velocity, density and thickness were supposed to be known in these examples). When the other model parameters are unknown higher-order Rayleigh-wave modes should be included in the inversion to correctly converge (Strobbia, 2002; Luo et al., 2007).

The set of frequencies should be chosen in a range such that the fundamental mode is sensitive to all the layer's velocity. For instance, if a small S-wave velocity variation in one of the layers does not reflect any considerable change in the Jacobian matrix, then the S-wave velocity of this layer cannot be retrieved by inversion. Finally, the convergence rate appears to be controlled by the damping factor and the weighting matrix as shown here using a heuristic definition.

Thickness estimation using the fundamental mode only: We used inversion to estimate layer thickness for the 7-layer model (Table 3.2). The initial values of thickness

Layer number	$\mathbf{v}_s$ (m/s)	$\mathbf{v}_p$ (m/s)	$\boldsymbol{\rho}$ (kg/m^3)	$\mathbf{h}$ (m)
1	275	650	1820	1.2
2	325	650	1820	1.2
3	375	750	1860	1.2
4	425	750	1860	1.2
5	475	1400	1910	1.2
6	525	1400	1910	1.2
7	575	1800	1960	1.2
8	625	1800	1960	1.2
9	675	2150	2020	1.2
10	725	2150	2020	1.2
11	775	2150	2020	1.2
Half space	825	2800	2090	∞

Table 3.3: Parameters of a model whose layer thicknesses *are* proportional to the exact model ones (Table 3.2).

are 7 m for the first three layers and 15 m for the other four layers. $\mathbf{v}_p$, $\mathbf{v}_s$ and $\boldsymbol{\rho}$ are considered known and equal to the exact ones. The layer thicknesses are correctly retrieved after 10 iterations (Figure 3.15). In this example, the thicknesses of the three shallowest layers are retrieved at the 5th iteration (Figure 3.15, right), while the thicknesses of the deepest layers are retrieved at the 10th iteration. Such a behaviour can be explained by a sensitivity analysis with respect to thickness. The rows of the Jacobian matrix are plotted in Figure 3.16. This Figure shows that Rayleigh waves are less sensitive to variations in the thickness of the deepest layers for the chosen range of frequencies. In this case, the convergence rate can be improved if the range of frequencies starts from 5 Hz instead of 15 Hz. However, resolution issues of the dispersion curves at lower frequencies could prevent us from including such low frequencies. An alternative solution that can improve the sensitivity to deepest layers consists in including higher-order modes in inversion (following example).

Inversion of one, two or three dispersion curves: In this example, the *exact* model is the 7-layer one presented in Luo et al. (2007). For the first three tests, the *initial* model parameters are taken from Luo et al. (2007) (Table 3.3). Although the initial values of thickness are supposed to be unknown, they are considered here as being proportional to the thicknesses in the exact model. We inverted either the fundamental mode (test 1), fundamental plus one higher mode (test 2) or fundamental plus two higher modes (test 3). The results of inversion are shown in Figure 3.17. For test 1, the inversion algorithm succeeded at matching observed and modelled data. However, the retrieved velocity model does not correspond to the exact velocity model (Figure 3.17, top). For test 2, the two considered dispersion curves are correctly explained and the retrieved velocity model is closer to the exact one than for test 1 (Figure 3.17, middle). For test 3, the three considered dispersion curves are correctly reconstructed and the retrieved velocity model indeed corresponds to the exact velocity model (Figure 3.17, bottom). We can therefore conclude that the lack of initial information about the thickness generated an

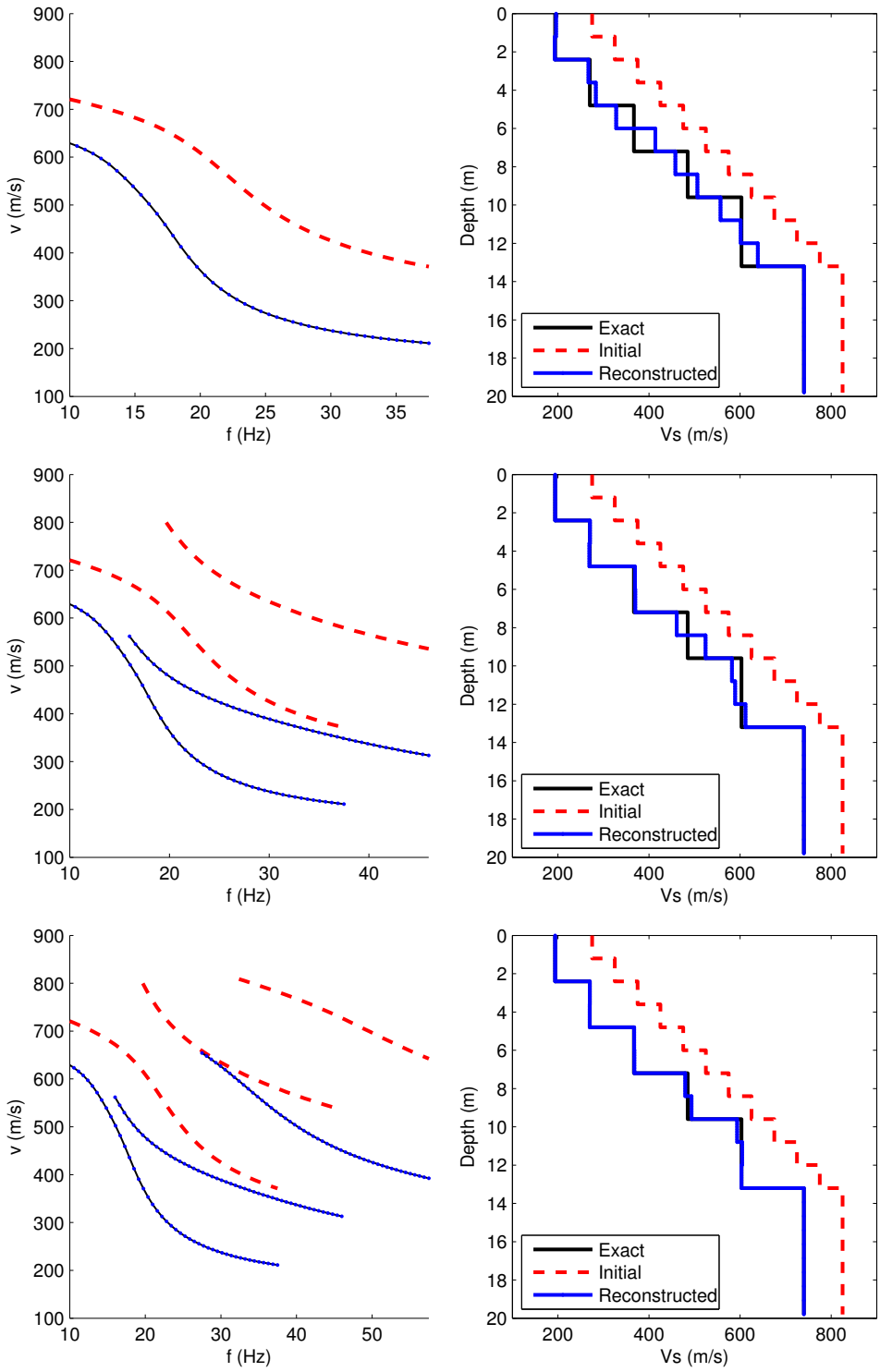


Figure 3.17: Inverted data (left) and S-wave velocity profile (right) for exact, initial and reconstructed models (legends are the same at right and left). Exact model thicknesses *are* multiples of the initial model ones. Top: inversion of only the fundamental mode. Middle: inversion of two modes. Bottom: inversion of three modes.

Layer number	$\mathbf{v}_s$ (m/s)	$\mathbf{v}_p$ (m/s)	$\boldsymbol{\rho}$ (kg/m^3)	$\mathbf{h}$ (m)
1	275	650	1820	1
2	275	650	1820	1
3	325	700	1840	1
4	325	750	1860	1
5	375	750	1860	1
6	425	1400	1910	1
7	475	1400	1910	1
8	525	1800	1960	1
9	575	1800	1960	1
10	575	1950	1990	1
11	625	2150	2020	1
12	675	2150	2020	1
13	725	2150	2020	1
14	775	2800	2090	1
15	825	2800	2090	1
Half space	825	2800	2090	∞

Table 3.4: Parameters of a model whose layer thicknesses *are not* proportional to the exact model ones (Table 3.2).

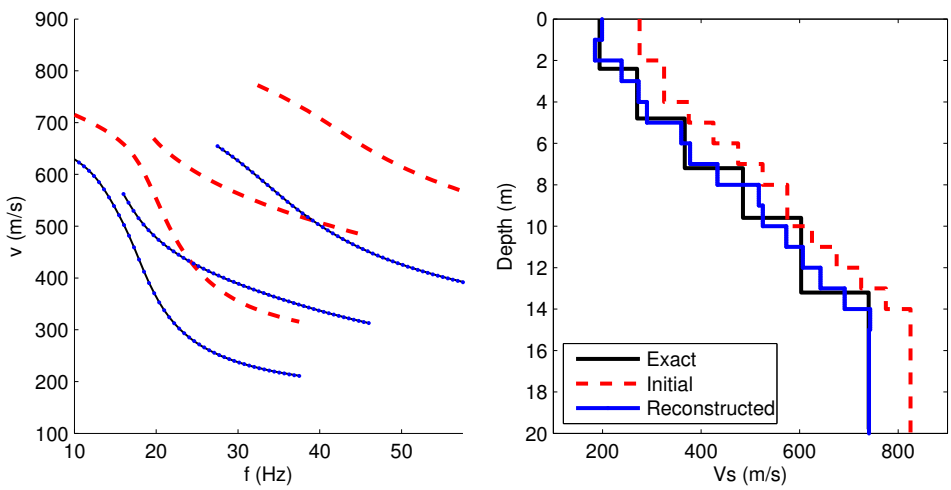


Figure 3.18: Same as Figure 3.17 but exact model thicknesses *are not* multiples of the initial model ones.

erroneous velocity model after inversion of only the fundamental mode. The solution to this problem is to consider more dispersion curves. The best result was obtained by inversion of three dispersion curves.

In a final test, we invert three dispersion curves as for test 3 but with a different initial model. We choose different values of thickness such that the exact model thicknesses are not multiples of the initial model ones (see the elastic parameters in Table 3.4). The results of this test are shown in Figure 3.18. The data are correctly reconstructed but the velocity model does not coincide with the exact velocity model. In this case, the observed data are ambiguous as two different velocity models can completely explain the observed data. The objective behind this test was to show the data ambiguity in SWA.

3.3.3 Conclusions

The estimation of model parameters from surface waves recorded in the field is usually done in three steps: data acquisition, processing and inversion (Socco and Strobbia, 2004). We have described the last two steps presenting (1) the classical processing approach used to extract the dispersion curves from shot gathers and (2) a least-squares fitting method to solve the inverse problem.

The classical processing approach consists of computing the $f-k$ spectrum, via the 2D Fourier transform, identifying the dispersion curves, and transforming them to the $f-v$ domain. The extraction of dispersion curves is successful only if certain conditions are honoured. For example, the $f-k$ resolution can affect the result especially at low frequencies due to the $f-k$ to $f-v$ transformation. Moreover, the mode identification in the $f-k$ image can be difficult if the propagation modes are not sufficiently separated and distinguishable. The first drawback could be avoided if the mode identification is done directly in the $f-v$ domain and some a priori information is available. The second drawback is difficult to be handled and in some cases only apparent dispersion curves can be extracted.

The goal of inversion is to find a set of 1D model parameters that minimise the dispersion curve mismatch. The dispersion curves can be computed for 1D velocity models but their computation is not clear for 2D and 3D models. Inversion can be applied to one (fundamental mode) or several (fundamental and higher modes) dispersion curves. The inverted model parameter is usually the distribution of S-wave velocity as a function of depth because surface waves are more sensitive to this parameter. We have presented examples using the least-squares inversion method with dynamic damping and dynamic weighting. The results are satisfactory in well constrained inversion problems. However, the dispersion curves can be explained by more than one velocity model, hence the data are sometimes ambiguous in SWA. The inversion convergence depends on several factors: initial velocity model sufficiently close to the exact one, good *a priori* information on the other three model parameters, non-overlapped dispersion curves in the $f-k$ image, and the choice of a frequency range broad enough such that the surface-wave velocity is sensitive to all the relevant S-wave velocities.

3.4 Full waveform inversion (FWI)

The objective of FWI is to minimise the misfit between recorded shot gathers and synthetic ones, from now on referred to as recorded data and modelled data. The synthetic shot gathers are computed with wave propagation modelling. We work with the velocity-stress formulation presented by Virieux (1986) (details in Chapter 2). In this section, we present explicit formula for the gradient computation with the adjoint-state method in 2D elastic isotropic media. A description of the two approaches used for the wavelet source estimation is then given. Finally, we describe some alternatives to FWI that include modified objective functions. Numerical experiments are not presented here but in Chapter 4 (different than for section 3.2).

3.4.1 Gradient computation

The misfit functional $\Phi(\mathbf{m})$ is defined in equation 3.12, with $\mathbf{u} = [v_x, v_z, \sigma_{xx}, \sigma_{xz}, \sigma_{zz}]$. The model parameter $\mathbf{m}(\mathbf{x})$ for isotropic elastic media corresponds, for example, to the Lamé parameters and density ($\boldsymbol{\lambda}(\mathbf{x})$, $\boldsymbol{\mu}(\mathbf{x})$, $\boldsymbol{\rho}(\mathbf{x})$), velocities and density ($\mathbf{v}_p(\mathbf{x})$, $\mathbf{v}_s(\mathbf{x})$, $\boldsymbol{\rho}(\mathbf{x})$), or impedances and density ($\mathbf{I}_p(\mathbf{x})$, $\mathbf{I}_s(\mathbf{x})$, $\boldsymbol{\rho}(\mathbf{x})$) (Tarantola, 1986). We only work with the first set of model parameters understanding that any other combination can be chosen depending on the objectives of inversion (Forgues and Lambaré, 1997; Prieux et al., 2013a,b). Diffraction pattern analysis useful to define the best set of model parameters for inversion of surface waves have not been done in this work.

Equations for elastic isotropic media

We would like to compute the gradient of the misfit function (equation 3.12) in the case of the velocity-stress formulation (equation 2.2). The augmented Lagrangian is defined in this case as

$$\begin{aligned}\mathcal{L}(\mathbf{u}, \boldsymbol{v}, \mathbf{m}) = \Phi(\mathbf{u}) - \\ \sum_{src} \int \Big(& \langle \beta_x^*, \rho \partial_t v_x - \partial_x \sigma_{xx} - \partial_z \sigma_{xz} - \varphi_{v_x} \rangle \\ & + \langle \beta_z^*, \rho \partial_t v_z - \partial_x \sigma_{xz} - \partial_z \sigma_{zz} - \varphi_{v_z} \rangle \\ & + \langle \sigma_{xx}^*, \partial_t \sigma_{xx} - (\lambda + 2\mu) \partial_x v_x - \lambda \partial_z v_z \rangle \\ & + \langle \sigma_{zz}^*, \partial_t \sigma_{zz} - (\lambda + 2\mu) \partial_z v_z - \lambda \partial_x v_x \rangle \\ & + \langle \sigma_{xz}^*, \partial_t \sigma_{xz} - \mu (\partial_x v_z - \partial_z v_x \rangle \Big) \mathrm{d}t,\end{aligned} \tag{3.29}$$

where $\boldsymbol{v} = [\beta_x^*, \beta_z^*, \sigma_{xx}^*, \sigma_{xz}^*, \sigma_{zz}^*]$ is the adjoint-state variable. The initial boundary conditions of state variables have been set to zero and are not shown in equation 3.29. These conditions are equal to the final boundary conditions of the adjoint-state variable.

For conveniency in the implementation, we would like to find an auto-adjoint system.

Thus, we have considered a change of variables for the adjoint stress components as

$$\begin{aligned} s^*_{xx} &= (\lambda + 2\mu)\sigma^*_{xx} + \lambda\sigma^*_{zz}, \\ s^*_{zz} &= (\lambda + 2\mu)\sigma^*_{zz} + \lambda\sigma^*_{xx}, \\ s^*_{xz} &= \mu\sigma^*_{xz}, \end{aligned} \tag{3.30}$$

with $[s^*_{xx}, s^*_{xz} s^*_{zz}]$ being the new adjoint stress variables. The adjoint-state equations are obtained by differentiation of the augmented Lagrangian with respect to the state variable, as shown in equations 3.14 and 3.15. The result is the same either if the change of variables is applied before or after differentiation.

The adjoint-state equations are obtained by differentiating equation 3.29 by applying the change of variables of equation 3.30. This operation yields

$$\begin{aligned} \partial_t \beta^*_x &= \frac{1}{\rho}(\partial_x s^*_{xx} + \partial_z s^*_{xz}) + \frac{\partial \Phi}{\partial v_x}, \\ \partial_t \beta^*_z &= \frac{1}{\rho}(\partial_x s^*_{xz} + \partial_z s^*_{zz}) + \frac{\partial \Phi}{\partial v_z}, \\ \partial_t s^*_{xx} &= (\lambda + 2\mu)\partial_x \beta^*_x + \lambda \partial_z \beta^*_z, \\ \partial_t s^*_{zz} &= (\lambda + 2\mu)\partial_z \beta^*_z + \lambda \partial_x \beta^*_x, \\ \partial_t s^*_{xz} &= \mu(\partial_z \beta^*_x + \partial_x \beta^*_z). \end{aligned} \tag{3.31}$$

These equations are similar to those of the direct functional but with different source terms $\partial\Phi/\partial v_x$ and $\partial\Phi/\partial v_z$, which are so-called residual sources. The equation system 3.31 has final boundary conditions equal to zero, which in the practice means that residual sources must be back-propagated in time. Residual sources are obtained by differentiating the objective function (equation 3.12) with respect to the components of velocity. Such procedure gives

$$\begin{aligned} \frac{\partial \Phi}{\partial v_x}(\mathbf{x}, t) &= P_x(\mathbf{x}, t) K_x(\mathbf{x})(K_x(\mathbf{x}) v_x(\mathbf{x}, t) - d^{obs}_x(\mathbf{x}, t)), \\ \frac{\partial \Phi}{\partial v_z}(\mathbf{x}, t) &= P_z(\mathbf{x}, t) K_z(\mathbf{x})(K_z(\mathbf{x}) v_z(\mathbf{x}, t) - d^{obs}_z(\mathbf{x}, t)), \end{aligned} \tag{3.32}$$

where $P_x(\mathbf{x}, t)$, $P_z(\mathbf{x}, t)$, $K_x(\mathbf{x})$ and $K_z(\mathbf{x})$ are the functional representations of the weighting matrix $\mathbf{P}(\mathbf{x}, t)$ and the receiver selecting matrix $\mathbf{K}(\mathbf{x})$ in equation 3.12.

The expression for the gradient is obtained by differentiating equation 3.29 with respect to the model parameters. The results of this operation considering the Lamé parameters and density are

$$\frac{\partial \Phi}{\partial \lambda}(\mathbf{x}) = \mathbf{G}(\lambda) = \sum_{src} \int \frac{(s^*_{xx} + s^*_{zz})\,(\partial_x v_x + \partial_z v_z)}{2(\lambda + \mu)} \mathrm{d}t, \tag{3.33}$$

$$\begin{aligned} \frac{\partial \Phi}{\partial \mu}(\mathbf{x}) = \mathbf{G}(\mu) = \sum_{src} \int \Bigg(& \frac{((\lambda + 2\mu) s^*_{xx} - \lambda s^*_{zz})\,\partial_x v_x}{2\mu(\lambda + \mu)} \\ & + \frac{((\lambda + 2\mu) s^*_{zz} - \lambda s^*_{xx})\,\partial_z v_z}{2\mu(\lambda + \mu)} \\ & + \frac{s^*_{xz}\,(\partial_x v_z + \partial_z v_x)}{\mu} \Bigg) \mathrm{d}t, \end{aligned} \tag{3.34}$$

and

$$\frac{\partial \Phi}{\partial \rho}(\mathbf{x}) = \mathbf{G}(\rho) = -\sum_{src} \int (\beta_x^* \partial_t v_x + \beta_z^* \partial_t v_z)\, \mathrm{d}t, \tag{3.35}$$

where $\mathbf{G}(\lambda)$, $\mathbf{G}(\mu)$ and $\mathbf{G}(\rho)$ are respectively the gradients with respect to λ, μ and ρ.

3.4.2 Source wavelet estimation

In real seismic inversion, the source wavelet is usually unknown. The estimation of the source wavelet is in itself an inverse problem that can be solved either using the adjoint-state method (Tarantola, 1984) or directly solving the linear inverse problem (Pratt, 1999).

Solution with the adjoint-state method

In the approach of (Tarantola, 1984), the wavelet source is stated as

$$\varphi_{v_z}(\mathbf{x}, t) = s(t)\delta(\mathbf{x} - \mathbf{x}_s), \tag{3.36}$$

where $s(t)$ is the source wavelet (or wavelet signature) and $\delta(\mathbf{x}-\mathbf{x}_s)$ is a Dirac distribution indicating the spatial position of the source at $\mathbf{x}_s$. In these equations, we refer only to source on the vertical component φ_{v_z}, but equations are the same for the horizontal component φ_{v_x}. The source wavelet estimate is iteratively updated according to

$$s^{k+1}(t) = s^k(t) - \mathbf{W}\frac{\partial \Phi}{\partial s(t)}, \tag{3.37}$$

where $\mathbf{W}$ represents the inverse of the Hessian matrix or its approximation. The augmented Lagrangian in equation 3.29 can be used to find an expression for the gradient as

$$\frac{\partial \Phi}{\partial s(t)} = \frac{\partial \mathcal{L}}{\partial s(t)} = \int \beta_z(\mathbf{x}, t)\delta(\mathbf{x} - \mathbf{x}_s) d\mathbf{x} = \beta_z(\mathbf{x}_s, t). \tag{3.38}$$

In this approach, the gradient of equation 3.38 needs modelling the wave propagation only once to obtain $\beta_z(\mathbf{x}_s, t)$ (back-propagation of residual sources).

Solution with linear inverse problem

In the approach of (Pratt, 1999), a component of the modelled data can be written as

$$d_z(\mathbf{x}, t) = s(t) * u_z(\mathbf{x}, t), \tag{3.39}$$

where $s(t)$ is the source wavelet, the symbol $*$ denotes convolution in the time domain and $u_z(\mathbf{x}, t)$ is the Green's function which can also be written as $u_z(\mathbf{x}, t) = S(\mathbf{x})p_z(\mathbf{x}, t)$, where $S(\mathbf{x})$ is the matrix that selects the signals at the position of receivers and $p_z(\mathbf{x}, t)$ is the complete velocity field for an impulsive source $\varphi_{v_z}(\mathbf{x}, t) = \delta(\mathbf{x}-\mathbf{x}_s)$. From equation 3.39, the objective function for a single source is

$$\Phi = \frac{1}{2}\int\int (s(t) * u_z(\mathbf{x}, t) - d_z^o(\mathbf{x}, t))^2\, d\mathbf{x}dt. \tag{3.40}$$

The perturbation theory can be used to find the source wavelet as

$$s(t) = s_0(t) - \left(\frac{\partial^2 \Phi}{\partial s_0^2(t)}\right)^{-1} \frac{\partial \Phi}{\partial s_0(t)}, \tag{3.41}$$

where $s_0(t)$ is a known background wavelet source,

$$\frac{\partial \Phi}{\partial s_0(t)} = \int u_z(\mathbf{x}, t) * (s_0(t) * u_z(\mathbf{x}, t) - d_z^o(\mathbf{x}, t))\, d\mathbf{x} \tag{3.42}$$

and

$$\frac{\partial^2 \Phi}{\partial s_0^2(t)} = \int u_z(\mathbf{x}, t) * u_z(\mathbf{x}, t) d\mathbf{x}. \tag{3.43}$$

Expressions 3.42 and 3.43 can be substituted in equation 3.41 to obtain

$$s(t) = s_0(t) - s_0(t) * \frac{\int u_z(\mathbf{x}, t) * u_z(\mathbf{x}, t) d\mathbf{x}}{\int u_z(\mathbf{x}, t) * u_z(\mathbf{x}, t) d\mathbf{x}} + \frac{\int u_z(\mathbf{x}, t) * d_z^o(\mathbf{x}, t) d\mathbf{x}}{\int u_z(\mathbf{x}, t) * u_z(\mathbf{x}, t) d\mathbf{x}} \tag{3.44}$$

and finally find the expression

$$s(t) = \frac{\int u_z(\mathbf{x}, t) * d_z^o(\mathbf{x}, t) d\mathbf{x}}{\int u_z(\mathbf{x}, t) * u_z(\mathbf{x}, t) d\mathbf{x}}. \tag{3.45}$$

As wave propagation modelling codes are usually used to compute $d_z(\mathbf{x}, t)$ for a wavelet source $s^c(t)$ (index c stands for current), equation 3.45 can be computationally implemented as

$$s(t) = s^c(t) * \frac{\int d_z(\mathbf{x}, t) * d_z^o(\mathbf{x}, t) d\mathbf{x}}{\int d_z(\mathbf{x}, t) * d_z(\mathbf{x}, t) d\mathbf{x}}. \tag{3.46}$$

The wavelet source estimation of equation 3.46 is often implemented in the frequency domain (see Pratt (1999)). In such implementation, convolutions are replaced by multiplications. It is also a common practice to include a weighting matrix to focus the source estimation on chosen seismic traces, for example on the near offset ones (Plessix and Cao, 2011).

In conclusion, the wavelet source can be numerically estimated with two different approaches. One of them is a gradient-based method (Tarantola, 1984). Back-propagation of residuals is needed to obtain such gradient. The other approach includes the Hessian and, because the inverse problem with respect to the source is linear in FWI, the best source wavelet for the current velocity model is obtained in one iteration (Pratt, 1999). We are interested in the first approach because in Chapter 4 we will work with an objective function which is non-linear with respect to the source wavelet.

In the next section, we briefly describe one of the usual ways of analysing the resolution in FWI. Then, we review some research works that investigate FWI with modified objective functions. This information is needed because we are presenting an alternative formulation in Chapter 4.

3.4.3 Resolution and surface waves

Acoustic media

The interpretation of the wavepath can give an insight into the resolution power of FWI (Virieux and Operto, 2009). We have implemented the equations presented in Woodward (1992) to obtain the wavepath for a source-receiver pair in a *3D acoustic model* (Figure 3.19). The model is homogeneous with Vp=2000 m/s. Free-surface conditions are not considered. The maximum extension of the first Fresnel zone is $F_1 = \sqrt{\lambda o_{sr}}$, where λ is the wavelength and o_{sr} is the source-receiver offset. λ depends on the considered frequency.

The wavepaths for two mono-frequency sources are shown in Figure 3.19 (F_1 is indicated in blue and a vertical section across the wavepath is plotted in red). The length of F_1 defines the resolution of transmitted/refracted waves (Williamson, 1991; Pratt et al., 1996). Outside the first Fresnel zone, isochrones can be observed (the first one is F_1). The distance between the peaks in the vertical section defines the resolution of reflected waves. Such distance decreases as the frequency of the source is increased which indicates the improvement of resolution offered by the high frequencies. FWI is a high resolution inversion method because in principle it can invert all the events belonging to the wavepath.

Elastic media

In the presence of elasticity and free surface, near surface interference is created by surface waves. This effect represents an additional difficulty in elastic FWI (Brossier et al., 2009; Gélis et al., 2007). To show this, we have computed two shot gathers (Figure 3.20) considering the 1D velocity profiles employed in one example of SWA (Figure 3.13). In this case, the initial data present strong phase shifs with respect to observed data due to dispersion effects (Figure 3.20). Besides, even if the Vp and ρ models are known, the lack of heterogeneities in Vs will prevent FWI from converging because of the initial phase mismatch. The 1D gradients of the misfit function computed for this example show high amplitudes concentrated in the first 10 m depth (Figure 3.21). Phase shifs of more than half the wavelength and gradients with excessively concentrated amplitude in the near surface are surface wave effects that need to be corrected.

Alternatives to reduce surface wave effects include, for example, the application of

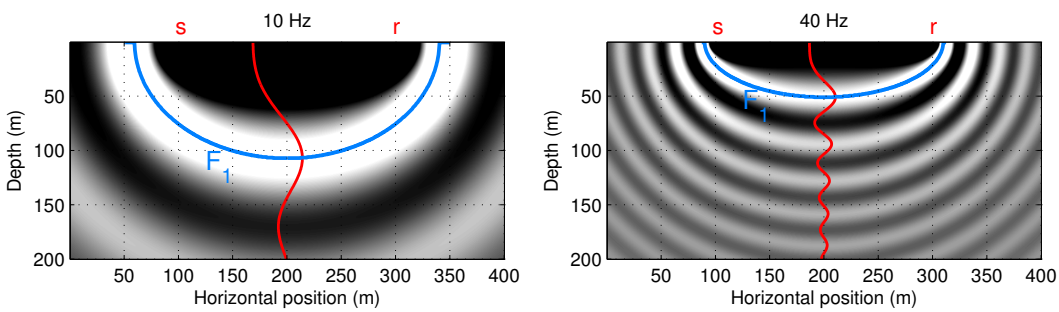


Figure 3.19: *Acoustic* wavepaths in a 3D medium (Woodward (1992)). Two monochromatic sources have been employed. The limit of the first Fresnel zone F_1 is indicated in blue. A vertical section across the wavepath is plotted in red.

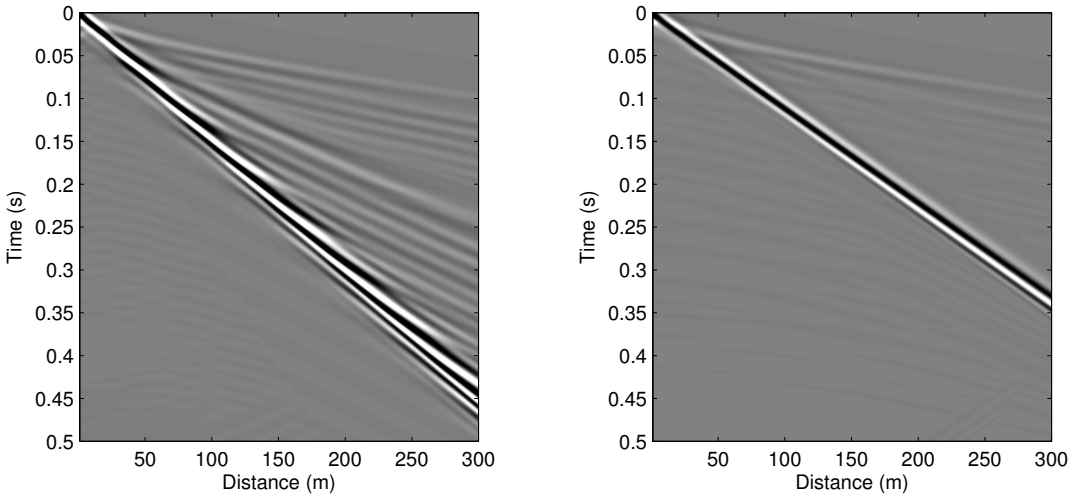


Figure 3.20: *Elastic* shot gathers (vertical component of velocity) at the surface for two different Vs models (Figure 3.13, top right). Left: the Vs model is layered. Right: the Vs model is homogeneous. The Vp and ρ models are identical for both shot gathers. Strong phase differences prevent FWI from converging.

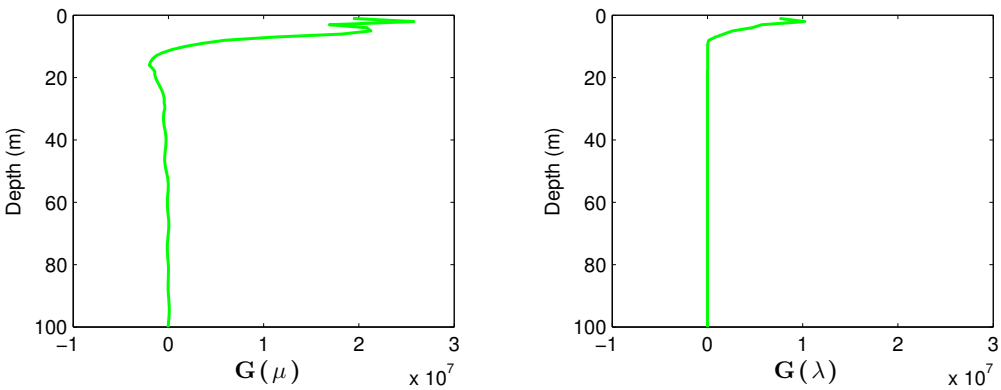


Figure 3.21: 1D FWI gradients (for μ on the left, for λ on the right) considering the initial profile of Figure 3.13. Higher values are concentrated in the first 10 m depth.

temporal windows which increase the first arrival energy and attenuate surface waves (Brossier et al., 2009; Romdhane et al., 2011). A different approach formulated in the $f-k$ domain will be presented and analysed in Chapter 4.

3.4.4 Alternative objective functions

FWI is a high resolution imaging technique that can be used to invert the complete recorded seismograms (Lailly, 1983; Tarantola, 1984; Virieux and Operto, 2009). Non-linearities lead to local minima in the classical least-squares objective function. The global minimum of the misfit function is related to the frequency content of the data: high frequencies reduce the basing of the global minimum, Bunks et al. (1995). In the presence of dispersive surface waves, high energy is expected in the high frequency content. Frequency and time masking strategies have widely been investigated (Pratt et al., 1998; Shin and Cha, 2008; Brossier et al., 2009; Bozdag et al., 2011; Liu et al., 2011; Romdhane et al., 2011; Métivier et al., 2012; Moghaddam and Mulder, 2012; Alkhalifah and Choi, 2013; Baek et al., 2013; Masoni et al., 2013).

In some research works, it is proposed to solve the difficulties in the application of FWI by implementing alternative objective functions. Those objective functions are usually designed to make inversion more robust by reducing local minima effects, complement the classical approach, or reduce the difficulty of including surface waves in inversion. In this section, we briefly comment on some of those research works. We study this because we are proposing an alternative objective function that hopefully will improve the performance of waveform inversion when the data contains dispersive surface waves (Chapter 4).

An alternative formulation that works in the Laplace domain is presented in Shin and Cha (2008). The modified objective function is

$$\Phi_a(\mathbf{m}) = \frac{1}{2}\sum_{src}\int_{-S}^{S}(\ln(\tilde{\mathbf{v}}(\mathbf{m})) - \ln(\tilde{\mathbf{d}}^{obs}))^2 \mathrm{d}s \tag{3.47}$$

where the variables with the tilde ($\tilde{}$) have been transformed to the Laplace domain and ln denotes the natural logarithm. The direct problem is formulated in the Laplace domain and the gradient is computed with similar formula as for classical FWI. There is however a difference with respect to FWI related to the residual source (more details can be found in Shin and Cha (2008)). In that work, it is shown that the implementation of an alternative objective function is useful to reconstruct a large wavelength (smooth) model with almost no a priori information.

Another couple of alternative objective functions is presented in Bozdag et al. (2011). In that work, the authors compare two misfit functions based on the envelope and on the instantaneous phase of signals. The two considered measurements are found by means of the analytic signal constructed from the seismograms and their Hilbert transform. The adjoint kernels of the considered misfit functions are presented through numerical examples in global seismology. Bozdag et al. (2011) conclude that the weighting characteristics yielded by the proposed misfit functions help mitigating problems related to high amplitudes. The authors suggest using these inversion procedures to complement FWI.

Another example of implementation of an alternative objective function is introduced in Liu et al. (2011). In this case, the objective function is written as

$$\Phi_d(\mathbf{m}) = \frac{1}{2}\sum_{src}\iint_0^{T_m}\left(\frac{\int_0^t (P(\mathbf{x},\tau))^2 \mathrm{d}\tau}{n(P(\mathbf{x},t))} - \frac{\int_0^t (P^{obs}(\mathbf{x},\tau))^2 \mathrm{d}\tau}{n(P^{obs}(\mathbf{x},t))}\right)^2 \mathrm{d}t\mathrm{d}\mathbf{x}, \tag{3.48}$$

where $P(\mathbf{x},t)) = P(\mathbf{x},t|\mathbf{m})$ is the pressure field for model $\mathbf{m}(\mathbf{x})$ and the denominators in the right-hand side term are functions that compute normalisation factors as $n(P(\mathbf{x},t)) = \int_0^{T_m}(P(\mathbf{x},\tau))^2\mathrm{d}\tau$ and $n(P^{obs}(\mathbf{x},t)) = \int_0^{T_m}(P^{obs}(\mathbf{x},\tau))^2\mathrm{d}\tau$, with T_m representing the total recorded time. By construction, the approach compares signals increasing with time and not oscillatory signals as in the classical formulation. Therefore local minima effects are reduced (Chauris et al., 2012; Donno et al., 2013). The residual source for the adjoint-state computation is found by differentiation of the objective function of 3.48 with respect to the pressure.

Some other strategies have been proposed to overcome the difficulty of defining an initial velocity model. Alkhalifah and Choi (2013) propose an objective function based on the instantaneous traveltime that can be used in inversion as a natural transition from travel time tomography to FWI. Baek et al. (2013) proposes a least-squares misfit based on registration as in image processing. The results with synthetic data show that this inversion approach can converge to the exact velocity model in the presence of some cycle-skipping problems.

Alternative objective functions are designed to deal with surface wave inversion in (Masoni et al., 2013). The implementation of the least-squares misfit in addition to weighted cross-correlation (van Leeuwen and Mulder, 2008) may result in a wider global minimum basin. The singular value decomposition (SVD), which has several applications in the context of surface waves could also be included in inversion (Moghaddam and Mulder, 2012).

We will focus here on the $f-k$ domain. We have seen that the main problem related to surface wave inversion using FWI is related to the mismatch of data phase (Figure 3.20). For this reason, special attention will be paid to the amplitude of surface waves in the $f-k$ domain while the phase will be partially neglected or ignored.

3.4.5 Conclusions

We have described the theoretical solution of the inverse seismic problem in the context of full waveform inversion (FWI). The model parameter is updated trough a local search whose direction is given by the gradient of the objective function. The adjoint-state method proves efficienct for the computation of the gradient. The Hessian matrix can be used in the context of Newton methods. The inverse of the exact Hessian matrix is however difficult to compute efficiently. However, the L-BFGS algorithm can be implemented to compute an approximation (Nocedal, 1980).

We have developed mathematical expressions for the computation of the misfit function gradient with the adjoint-state method for Lamé parameters and density in 2D elastic isotropic media (velocity-stress formulation) (Virieux, 1986). In our 2D elastic Lagrangian formulation, the wave-propagation system is auto-adjoint, and therefore, the

same implementation can be used for forward propagation and back-propagation modelling. The inverse problem to estimate the source wavelet is linear in FWI. We will focus on the gradient-based approach to solve this problem because we are proposing in Chapter 4 an alternative objective function which is non-linear with respect to the source wavelet (see Chapter 4). We did not considere other elastic parameters as radiation patterns are not yet investigated in the specific case of surface waves.

FWI is a high resolution imaging technique because all the recorded seismic waves can theoretically be fitted to synthetic data computed for a set of model parameters. Some potential problems may be encountered. For example, the global minimum vicinity is strongly related to the frequency content of the data. The choice of initial model is more difficult in the presence of surface waves. Some of the approaches proposed to overcome these difficulties include the formulation of FWI with alternative objective functions. We have reviewed some of these approaches because an alternative objective function is proposed in Chapter 4. Several numerical tests will be used to evaluate results of classical FWI and the novel technique (Chapter 4).

3.5 Discussion

We have reviewed two seismic inversion methods that can potentially deal with surface waves: SWA and FWI. Inversion of dispersion curves yields excellent results when the problem is well constrained in terms of a good initial model and good sensitivity with respect to the relevant model parameters. If the problem is not well constrained, then the retrieved velocity model is not equal to the exact one due to data ambiguities and non-uniqueness of the solution. The retrieved model in such cases satisfies the same boundary and continuity conditions across the layers. On the other hand, FWI can in principle invert the amplitude and phase of recorded seismic waves. The dispersion of surface waves creates and additional difficulty in the choice of the initial velocity model. We consider the advantages and limitations of these two inversion techniques in order to propose an intermediate approach for the inversion of surface waves.

In the next Chapter, we present an intermediate inversion approach together with an analysis of its characteristics. We compare results of the proposed approach to FWI to see if the constraint regarding the initial velocity model is reduced.

Chapter 4

An alternative approach for inversion of surface waves

Contents

4.1 Introduction 97

4.2 The windowed amplitude waveform inversion (w-AWI) . . 99

4.2.1 Model update and objective function 99

4.2.2 Gradient computation 103

4.2.3 Source wavelet estimation 104

4.2.4 Gradient regularisation and preconditioning 105

4.3 Analysis of w-AWI 106

4.3.1 Objective function shape and local minima 107

4.3.2 Cycle skipping 110

4.3.3 Width of spatial windows 117

4.3.4 Adjoint source and gradient 123

4.3.5 Low frequencies 124

4.4 Inversion of synthetic data 129

4.4.1 No a priori information and known source wavelet 129

4.4.2 Inversion with unknown source wavelet 141

4.4.3 Multi-parameter inversion 147

4.4.4 Inversion with low frequencies 152

4.5 Conclusions 165

Résumé du chapitre

Dans ce chapitre, nous présentons une nouvelle approche pour l'inversion des ondes de surface dans des milieux 2D que nous appelons *windowed-Amplitude Waveform Inversion* (w-AWI). Cette approche utilise une fonction objective définie dans le domaine $f-k$. Elle est intermédiaire entre celles utilisées en *Surface Wave Analysis* (Xia et al., 1999; Socco and Strobbia, 2004) et *Full Waveform Inversion* (Lailly, 1983; Tarantola, 1986; Pratt et al., 1998; Virieux and Operto, 2009). La modélisation de la propagation des ondes est faite avec le champ d'onde complet en utilisant la formulation vitesse-contrainte (Virieux, 1986) détaillée au chapitre 2. L'objectif de l'approche d'inversion intermédiaire consiste à rendre plus flexible le choix du modèle de départ par comparaison avec la FWI. En outre, les courbes de dispersion n'ont pas à être pointées comme pour la SWA. Une série des tests numériques sur des données synthétiques démontre les avantages de cette nouvelle approche.

La section 4.2 est consacrée aux développements théoriques de la w-AWI. La mise-à-jour du modèle est réalisée avec une approche basée sur gradient de la nouvelle fonction objective. Ce gradient est efficacement calculé selon la méthode de l'état adjoint (Plessix, 2006). L'ondelette source est estimée avec une approche de gradient car la nouvelle fonction objective n'est pas linéaire par rapport à cette ondelette. Nous présentons aussi la régularisation et le préconditionnement appliqué au gradient pour aider à la convergence.

La section 4.3 montre les résultats des tests numériques réalisés pour analyser et comparer les fonctions objectives des approches w-AWI et FWI. Nous étudions en particulier la forme de la fonction objective, le phénomène de saut de phase, l'influence de la longueur de la fenêtre en w-AWI et aussi le rôle des basses fréquences. D'après ces analyses, w-AWI est une approche dont sa fonction objective a un bassin d'attraction autour du minimum global plus large que celui pour la FWI. Les minimaux secondaires sont plus loin et moins présents que pour la FWI grâce à la définition de la fonction objective dans le domaine $f-k$. La longueur des fenêtres est un atout en w-AWI qui peut être exploité pour définir des stratégies d'inversion. Enfin, tout le spectre de fréquences des données peut être utilisé en w-AWI.

La partie 4.4 est dédiée aux tests d'inversion de données synthétiques avec w-AWI et FWI. Nous montrons des exemples où le modèle initial est éloigné du modèle exact. L'ondelette source est supposée inconnue dans certains tests. Les résultats montrent que w-AWI est plus flexible que FWI en ce qui concerne le choix du modèle initial. L'approche w-AWI peut retrouver un modèle de vitesse Vs qui est proche du modèle exact. La FWI pourrait être appliquée ensuite pour reconstruire la phase des données qui n'a pas été prise en compte par w-AWI. Le modèle élastique final sera ainsi amélioré.

Cette nouvelle formulation sera ensuite appliquée à un jeux de données réelles dont les résultats sont présentés dans le chapitre 5.

Abstract

We propose an alternative approach based on waveform inversion for reconstructing 2D model parameters from surface waves. We replace the classical full waveform inversion (FWI) objective function with a similar but alternative objective function formulated in a different domain. The classical FWI objective function suffers from severe local minima problems in the presence of reflected or dispersive waves. The presence of local minima requires an excellent initial model. In our approach, the data are first split using windows in the time-space $t - x$ domain. For each window, the amplitude of the frequency-wavenumber $f - k$ spectrum is computed. The objective function is the least-squares misfit between the observed and modelled 2D Fourier transformed datasets. We call this formulation the windowed amplitude waveform inversion (w-AWI). The w-AWI objective function reduces local minima problems as shown here through some numerical examples. The global minimum basin is wider in the w-AWI approach than in the FWI one. 2D inversion examples show that convergence towards the correct velocity model can be achieved even when cycle-skipping is observed in the $t - x$ domain. Besides, full advantage of surface waves can be taken as w-AWI performs a good weighting of low and high frequencies. This surface-wave inversion procedure can be used to retrieve near-surface model parameters in lateral-varying media.

4.1 Introduction

The characterisation of the near surface (the first hundreds of meters) is essential for improving seismic imaging of both shallow and deeper exploration targets. Conventional seismic characterisation is done by analysing body waves. For example, first-arrival traveltime tomography is used to reconstruct the long wavelength velocity model (Taillandier et al., 2009). Moreover, near-surface characterisation based on common-depth-point (CDP) reflection profiling requires ultra high-frequency seismic data acquisition (a few hundred Hz) (Knapp and Steeples, 1986). In these imaging methods, surface waves are considered to be coherent noise that should be eliminated from the seismograms to enhance body waves. However, surface waves commonly represent more than half the seismic energy recorded in shot gathers and carry useful information. Surface waves are dispersive in heterogeneous media (Thomson, 1950). Such property can be used to retrieve model parameters and characterise the near surface (Nazarian and Stokoe II, 1984; Park et al., 1999). Our objective is to use surface waves for reconstructing 2D high-resolution near-surface velocity models. We propose a surface-wave inversion approach based on a combination of the properties of two classical techniques: Surface Wave Analysis (SWA) and Full Waveform Inversion (FWI).

In SWA, the objective function consists of the least-squares misfit between extracted-from-shotgathers and modelled dispersion curves (Socco et al., 2010). The dispersion curves are a representation of the frequency-dependent surface-wave velocities mapped in the frequency-velocity ($f - v$) domain. We can often distinguish the fundamental and the higher dispersion modes. Xia et al. (1999) used a Marquardt-Levenberg (Marquardt, 1963) inversion approach to estimate the S-wave velocity profile from the fundamental mode only. Later on, multi-modal inversion was implemented to achieve reconstruction of

more complex models (Xia et al., 2003; Socco and Strobbia, 2004). Classical multimodal SWA requires separation and identification of different dispersion modes in the spectrum (Maraschini et al., 2010). Difficulties can arise when the fundamental mode can not easily be distinguished from higher modes.

The result of classical multi-modal SWA is a plane-layered depth-dependent model described in terms of S-wave velocity, P-wave velocity and density. Surface-wave propagation is more sensitive to S-wave velocity (Nazarian and Stokoe II, 1984; Xia et al., 1999). As a consequence, surface-wave inversion is commonly used to reconstruct S-wave velocity as a function of depth. Elastic waveform modelling is not required in this inversion approach. Only the dispersion curves must be modelled (Lai, 1998) requiring only modest computational efforts. Lateral velocity variations are usually handled by interpolating 1D results (Tian et al., 2003), or by laterally constrained inversion (Socco et al., 2009; Bergamo et al., 2012).

On the other side, FWI is an inversion procedure used to estimate a high resolution model of the subsurface. The classical objective function consists of the least-squares misfit between the complete recorded dataset and a synthetic dataset obtained by wave propagation modelling (Lailly, 1983; Tarantola, 1984; Gauthier et al., 1986; Mora, 1988; Pratt et al., 1998). In FWI, the inverse problem is currently solved with a local optimisation approach. The gradient of the objective function is computed at each iteration by cross-correlating forward and back-propagated wave fields (Tarantola, 1986). The synthetic data are computed in the frequency domain (Brossier et al., 2009), or in the time domain, by means of finite-element methods (Marfurt, 1984) or finite-difference methods (Virieux, 1986). FWI has a high computational cost but is in principle able to exploit the whole information carried by body and/or surface waves for the reconstruction of the velocity model and other elastic parameters. Virieux and Operto (2009) presented an overview of FWI including also the identification of the principal challenges to make FWI as popular as migration techniques. They enumerated the challenges as (1) building accurate starting models, (2) implementing minimisation criteria to increase robustness when performing multi-parameter inversion, and (3) improving the computational efficiency to make 3D elastic FWI feasible. In the 2D approximation, present-time computers are able to handle elastic wave propagation in reasonable time; however local minima effects due to inaccurate starting models are still an important challenge when inverting data that contains reflections and/or dispersive waves.

In this chapter, we present an inversion approach designed to reconstruct 2D elastic properties by inverting seismic data containing surface waves. We call it windowed-Amplitude Waveform Inversion (w-AWI). It is based on a modification of the classical FWI objective function proposed here for the purpose of reducing the presence of local minima. In our alternative objective function, some simple preprocessing of data are needed before computing the objective function value: (1) the data are split into subsets of consecutive receivers (spatial windowing), (2) each subset is transformed to the frequency-wavenumber ($f-k$) domain and *only the $f-k$ amplitude is considered*, and (3) the least-squares misfit between recorded and modelled subsets is computed. The general objective function value is obtained by summing the misfits of each individual spatial windowed subset. The data spatial windowing guarantees that the surface-wave properties contributing to the objective function in a particular area are associated to

the medium that is below the considered receivers. In the same area, this is a way of decreasing the dependence on the initial velocity model, and therefore, reducing local minima. Besides, since the misfit is defined in the $f-k$ domain, the w-AWI objective function minimises not only the distance between dispersion modes as in SWA, but also the wave energy as in FWI.

This chapter starts by presenting the theory underlying w-AWI (section 4.2). We show that the model update in w-AWI is computed in similar ways as in FWI with the adjoint-state method. We give details on the source wavelet estimation and the preconditioning approach implemented by us. In section 4.3, we analyse the effects of local minima, cycle skipping, width of spatial windows, residual source, gradient and low frequencies in w-AWI compared to FWI. In section 4.4, we show several examples of synthetic data inversion with or without *a priori* information. We also include examples regarding multi-parameter inversion and low-pass data filtering. We present these examples with the aim of explaining the advantages and limitations of the proposed surface-wave inversion technique.

4.2 The windowed amplitude waveform inversion (w-AWI)

We propose an inversion procedure that we call *windowed-Amplitude Waveform Inversion* (w-AWI). The w-AWI is an approach similar to FWI but it uses a different objective function. The goal of implementing such objective function is to increase the inversion robustness when the seismic data contains surface waves while retrieving a high resolution velocity model. The proposed objective function is a measure of the least-squares misfit between the modes of propagation of surface waves present in modelled and observed data mapped as amplitude information into the $f-k$ spectrum. In this section, we present the w-AWI objective function and compare the gradient computation to classical FWI. Moreover, we explain how to estimate the source wavelet. Finally, we discuss the regularisation and preconditioning used to improve the convergence rate and the illumination weighting.

4.2.1 Model update and objective function

This inversion approach consists of a local search of a model parameter $\mathbf{m}$ in the model space $\mathbf{M}$ that minimises a least-squares misfit functional $\Phi(\mathbf{m})$. The misfit functional is stated as

$$\Phi(\mathbf{m}) = \frac{1}{2}\sum_{src} \mathbf{R}^T(\mathbf{m})\mathbf{R}(\mathbf{m}) \tag{4.1}$$

where src stands for sources and $\mathbf{R}(\mathbf{m}) = \mathbf{R}(\mathbf{d}(\mathbf{m}), \mathbf{d}^{obs})$ is a residual vector (defined later in equation 4.3) between the synthetic data $\mathbf{d}(\mathbf{m})$ computed for a model parameter $\mathbf{m}$ and the recorded (or observed) data $\mathbf{d}^{obs}$. The data consist of a physical measure of the propagated wave, usually the particle velocity at the surface. The data are associated with this velocity information as $\mathbf{d}(\mathbf{m}) = \mathbf{K}\mathbf{v}(\mathbf{m})$, where $\mathbf{K}$ is the matrix used to select

velocity $\mathbf{v}(\mathbf{m})$ components at receiver positions. The computation of velocity $\mathbf{v}(\mathbf{m})$, which defines the forward problem, is done with equations 2.2 (Chapter 2).

The model parameter $\mathbf{m}$ vector is updated through iterations. This means that a model parameter at iteration $k+1$, represented by $\mathbf{m}^{k+1}$, is updated starting from an initial model parameter $\mathbf{m}^0$. In this approach, the misfit functional gradient is employed to update the model parameter using a quasi-Newton method (details in section 3.2). The model parameter is updated by

$$\mathbf{m}^{k+1} = \mathbf{m}^k - \mathbf{W}^{-1}(\mathbf{m}^k)\mathbf{G}(\mathbf{m}^k), \tag{4.2}$$

where $\mathbf{G}(\mathbf{m}^k)$ is the misfit functional gradient and $\mathbf{W}^{-1}(\mathbf{m}^k)$ is the approximate inverse of the Hessian. The gradient and Hessian correspond to the first and second order derivatives of the objective function with respect to the model, respectively.

As in FWI, the w-AWI model parameter $\mathbf{m}$ vector contains at least a few thousands of elements. Besides, wave propagation modelling is needed to explain the complete set of registered data. These two characteristics impose computational restrictions on the computation of gradient and inverse of the Hessian. Therefore, the gradient is computed using the adjoint-state method (Plessix, 2006), and the approximate inverse of the Hessian is computed using the limited memory-Broyden-Fletcher-Goldfarb-Shanno algorithm including boundaries to prevent unphysical model updates (L-BFGS-B) (Nocedal, 1980).

Objective function

In w-AWI, the misfit depends on the absolute value of the $f-k$ spectra of modelled and observed data such that both the energy and dispersion properties of surface waves are considered. Let us define the following *intermediary residual*

$$\mathbf{R}(\mathbf{m}) = (|\mathbf{D}(\mathbf{m})| - |\mathbf{D}^{obs}|), \tag{4.3}$$

where $|.|$ denotes the absolute value,

$$\mathbf{D}(\mathbf{m}) = \mathcal{F}_{2D}\{\mathbf{d}(\mathbf{m}|t,x)\} = \mathcal{F}_{2D}\{\mathbf{d}\}(\mathbf{m}|f,k) \tag{4.4}$$

and

$$\mathbf{D}^{obs} = \mathcal{F}_{2D}\{\mathbf{d}^{obs}(t,x)\} = \mathcal{F}_{2D}\{\mathbf{d}^{obs}\}(f,k), \tag{4.5}$$

with $\mathcal{F}_{2D}$ denoting the 2D Fourier transform expressed for a generic function $g(t,x)$ as

$$\mathcal{F}_{2D}\{g\}(f,k) = \iint g(t,x)e^{-i2\pi ft}e^{i2\pi kx}\mathrm{d}t\mathrm{d}x, \tag{4.6}$$

where t represents time, x the horizontal Cartesian coordinate, f the frequency and k the wavenumber. In classical FWI, the residual vector is defined as $\mathbf{R}(\mathbf{m}) = \mathbf{d}(\mathbf{m}) - \mathbf{d}^{obs}$. In this chapter, the weighting matrix $\mathbf{P}(\mathbf{x},t)$ (equation 3.1) is equal to one, thus it is not included in the objective function.

Please note that the residual defined in equation 4.3 is not equivalent to the classical FWI residual because the absolute value is applied to modelled and observed data

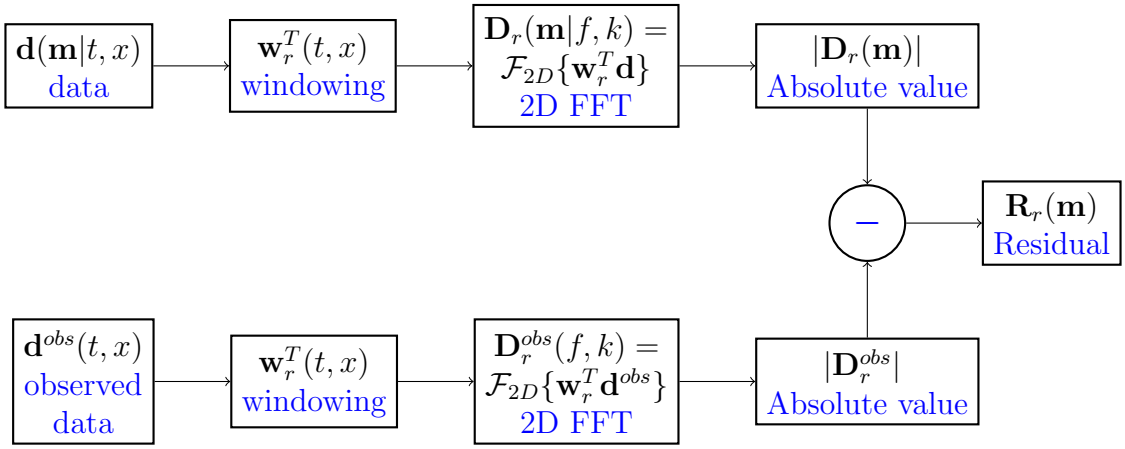


Figure 4.1: Workflow for w-AWI residual computation (equation 4.8).

separately. On the contrary, a residual like $\tilde{\mathbf{R}}(\mathbf{m}) = |\mathbf{D}(\mathbf{m}) - \mathbf{D}^{obs}|$ defines an objective function similar to the classical FWI objective function (this can be proven using Parseval's theorem, Oppenheim et al. (1999), p. 60).

The phase information is neglected in the intermediary residual (equation 4.3). As a result, lateral variations cannot be handled because velocity model perturbations could not be properly localised. In such a case, inversion would only be successful for 1D models. Otherwise, the local dependence of surface-wave dispersion can be used to relocalise velocity perturbations. For this reason, w-AWI employs spatial windows to separate the data into subsets of consecutive receivers. In this way, the residual principally depends on the dispersion effect that occurred within the zone below the considered receivers.

The spatial windows are represented by $\mathbf{w}_r^T(t,x)$ for $r = 1, ..., n_w$ with n_w being the number of windows. The *w-AWI objective function* is

$$\tilde{\Phi}(\mathbf{m}) = \frac{1}{2}\sum_{src}\sum_{r}\mathbf{R}_r^T(\mathbf{m})\mathbf{R}_r(\mathbf{m}), \tag{4.7}$$

where the residual is defined as

$$\mathbf{R}_r(\mathbf{m}) = (|\mathbf{D}_r(\mathbf{m})| - |\mathbf{D}_r^{obs}|), \tag{4.8}$$

and the data are computed as

$$\mathbf{D}_r(\mathbf{m}|f,k) = \mathcal{F}_{2D}\{\mathbf{w}_r^T(t,x)\mathbf{d}(\mathbf{m}|t,x)\} \tag{4.9}$$

and

$$\mathbf{D}_r^{obs}(f,k) = \mathcal{F}_{2D}\{\mathbf{w}_r^T(t,x)\mathbf{d}^{obs}(t,x)\}. \tag{4.10}$$

Equations 4.8 to 4.10 have been represented in the workflow shown in Figure 4.1

An example of the data considered in the w-AWI residual is shown in Figure 4.2. The synthetic shot gather $\mathbf{d}(\mathbf{m}|t,\mathbf{x})$ is shown on the top left (Figure 4.2). Besides, two different windows named $\mathbf{w}_a(t,x)$ and $\mathbf{w}_b(t,x)$ are shown on the top right (Figure 4.2). These windows will be used for computation of transformed data $\mathbf{D}_r(\mathbf{m}|f,k)$ (equation 4.9). The result of multiplying the shot gather by the windows is shown in the middle panels

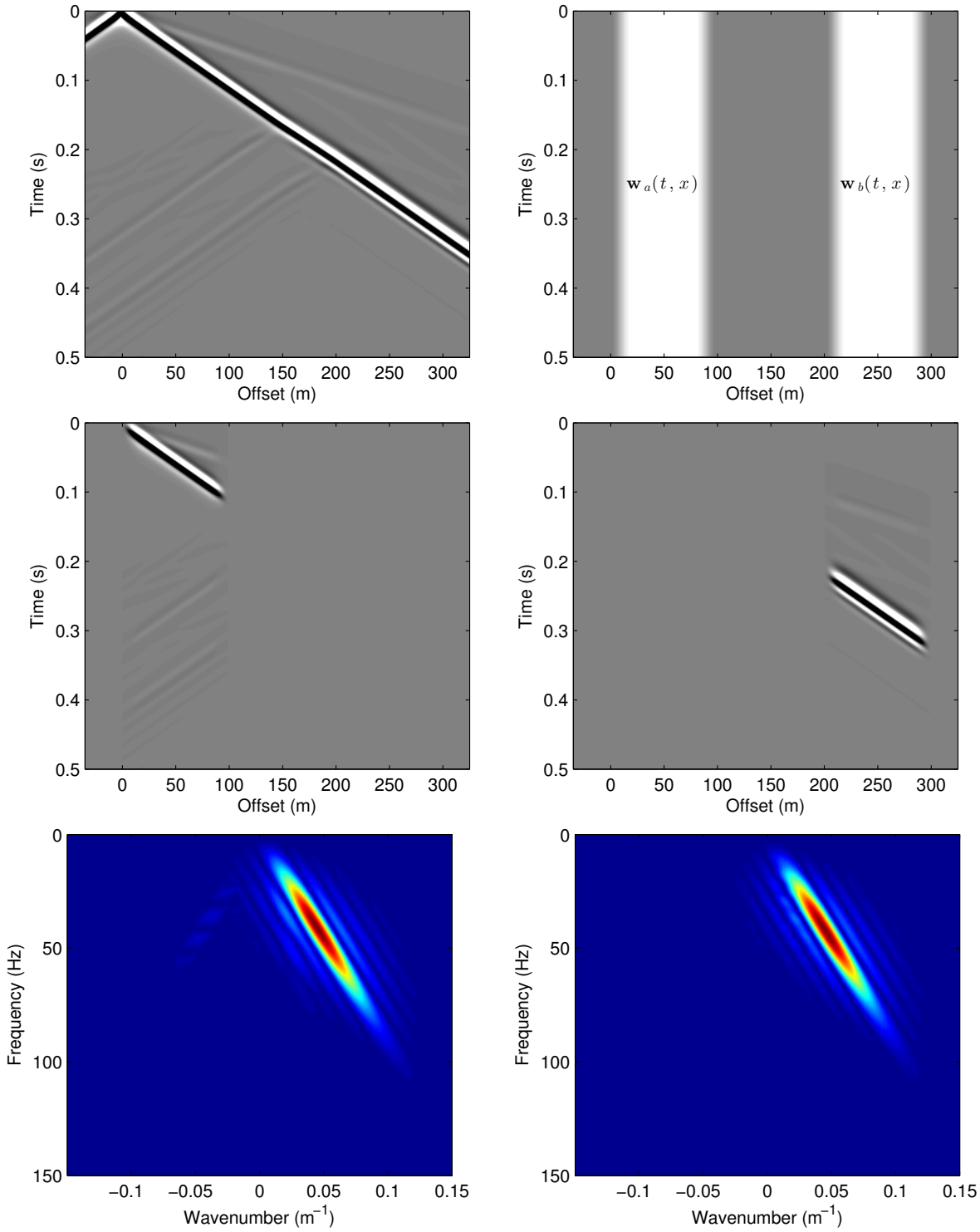


Figure 4.2: Example of data considered in the w-AWI residual. Top left: input shot gather. Top right: two windows used to select traces in the shot gather (equation 4.9). Middle: two subsets obtained by multiplying the shot gather by the windows ($\mathbf{w}_a(t,x)$ on the left; $\mathbf{w}_b(t,x)$ on the right). Bottom: absolute value of the 2D Fourier transforms of the two subsets shown in the middle.

(Figure 4.2). The two images at the bottom (Figure 4.2) correspond to the $f-k$ transformed data $|\mathbf{D}_r(\mathbf{m}|f,k)|$ to be used in the w-AWI residual (equation 4.8). Numerical examples will show that this data transformation allows reducing objective function local minima (created by phenomena like cycle skipping) while retrieving the same global minimum as in FWI. A detailed numerical comparison of FWI and w-AWI is presented in section 4.3.

4.2.2 Gradient computation

The misfit functional gradient is obtained using the adjoint-state method (Plessix, 2006). As described in Chapter 3, this method requires computing one forward modelling and one backward modelling. The gradient corresponds to the correlation between forward propagated and back-propagated data.

The forward wave propagation in 2D isotropic elastic media is modelled with a first-order hyperbolic system (Virieux, 1986). This system is

$$\begin{aligned}
\partial_t v_x &= \frac{1}{\rho}(\partial_x \sigma_{xx} + \partial_z \sigma_{xz}) + \varphi_{v_x}, \\
\partial_t v_z &= \frac{1}{\rho}(\partial_x \sigma_{xz} + \partial_z \sigma_{zz}) + \varphi_{v_z}, \\
\partial_t \sigma_{xx} &= (\lambda + 2\mu)\partial_x v_x + \lambda \partial_z v_z, \\
\partial_t \sigma_{zz} &= (\lambda + 2\mu)\partial_z v_z + \lambda \partial_x v_x, \\
\partial_t \sigma_{xz} &= \mu(\partial_z v_x + \partial_x v_z),
\end{aligned} \tag{4.11}$$

where x and z are the Cartesian coordinates, λ and μ are the Lamé parameters, ρ is the density, (v_x, v_z) are the velocity components, $(\sigma_{xx}, \sigma_{zz}, \sigma_{xz})$ are the stress components and $(\varphi_{v_x}, \varphi_{v_z})$ represent the source term.

The adjoint-state system (or back-propagation system) is found with equation 3.29 as

$$\begin{aligned}
\partial_t \beta_x^* &= \frac{1}{\rho}(\partial_x s_{xx}^* + \partial_z s_{xz}^*) + \frac{\partial \tilde{\Phi}}{\partial v_x}, \\
\partial_t \beta_z^* &= \frac{1}{\rho}(\partial_x s_{xz}^* + \partial_z s_{zz}^*) + \frac{\partial \tilde{\Phi}}{\partial v_z}, \\
\partial_t s_{xx}^* &= (\lambda + 2\mu)\partial_x \beta_x^* + \lambda \partial_z \beta_z^*, \\
\partial_t s_{zz}^* &= (\lambda + 2\mu)\partial_z \beta_z^* + \lambda \partial_x \beta_x^*, \\
\partial_t s_{xz}^* &= \mu(\partial_z \beta_x^* + \partial_x \beta_z^*),
\end{aligned} \tag{4.12}$$

where $\beta_x^*, \beta_z^*, s_{xx}^*, s_{xz}^*$ and s_{zz}^* are the adjoint-state variables (details in Section 3.4). Note that these equations are similar to those of the wave propagation system but with some modified sources $\partial\tilde{\Phi}/\partial v_x$ and $\partial\tilde{\Phi}/\partial v_z$. The adjoint-state system can be computed with the same forward wave propagation code. We refer the reader to Chapter 2 for more details on the forward modelling implemented here (second-order finite-difference on curved grids).

The adjoint-state system (equation 4.11) has the following final boundary conditions: at $t = T_{max}$, the adjoint-state variables and their time derivatives are equal to zero. These

conditions are a result of the initial boundary conditions of the state equations. The practical implementation of the final boundary conditions consists of back-propagating residual sources in time (Plessix, 2006). For more details, please refer to section 3.4.

The backward propagation system defined in equation 4.12 is almost exactly the same as in FWI for 2D elastic isotropic media (equation 3.31). The only difference between the gradient computation of FWI and w-AWI is the residual source ($\partial\tilde{\Phi}/\partial v_x$, $\partial\tilde{\Phi}/\partial v_z$). These new sources are retrieved by differentiating the misfit functional (equation 4.7) with respect to velocity. This operation yields

$$\frac{\partial\tilde{\Phi}(\mathbf{m})}{\partial\mathbf{v}(\mathbf{m})} = \mathbf{K}^T \sum_r \mathbf{w}_r \Re\left\{\mathcal{F}_{2D}^{-1}\left\{\frac{\mathbf{D}_r(\mathbf{m})}{|\mathbf{D}_r(\mathbf{m})|}\mathbf{R}_r(\mathbf{m})\right\}\right\}, \tag{4.13}$$

where $\Re$ stands for the real part and $\mathcal{F}_{2D}^{-1}$ represents the 2D inverse Fourier transform.

In general, equation 4.13 can be divided into three terms: (1) the multipliers $\mathbf{K}^T$ and $\mathbf{w}_r$, (2) the normalised term $\mathbf{D}_r(\mathbf{m})/|\mathbf{D}_r(\mathbf{m})|$ and (3) the residual $\mathbf{R}_r(\mathbf{m})$. The multipliers are the transposed of the ones used in the residual (equation 4.8), the normalised term is related to the differentiation of $|\mathbf{D}_r(\mathbf{m})|$, and the residual comes from considering the least-squares misfit functional (equation 4.7). Details on the mathematical development leading to equation 4.13 are shown in Appendix A. For comparison, the residual source in FWI corresponds to $\partial\Phi(\mathbf{m})/\partial\mathbf{v}(\mathbf{m}) = \mathbf{K}^T\mathbf{R}(\mathbf{m})$. Indeed, the expression of the w-AWI residual source is slightly more complicated than the FWI one. However, its computation represents a minor computational cost in comparison to the wave-propagation modelling. In the practice, if the term $|\mathbf{D}_r(\mathbf{m})|$ approaches to zero, then a small (10^{-3} of the maximum amplitude) term ϵ_0 is added to avoid numerical errors.

The equations for the correlation used to compute the w-AWI gradient are identical to those of FWI (see Chapter 3 for details). We have defined the Lamé parameters (λ and μ) and the density (ρ) as model parameters. The gradients for λ, μ and ρ are

$$\mathbf{G}(\lambda) = \frac{\partial\tilde{\Phi}}{\partial\lambda} = \sum_{src}\int \frac{(s_{xx}^* + s_{zz}^*)\,(\partial_x v_x + \partial_z v_z)}{2(\lambda+\mu)}dt, \tag{4.14}$$

$$\begin{aligned}\mathbf{G}(\mu) = \frac{\partial\tilde{\Phi}}{\partial\mu} = \sum_{src}\int & \frac{((\lambda+2\mu)s_{xx}^* - \lambda s_{zz}^*)\,(\partial_x v_x)}{2\mu(\lambda+\mu)} \\ & + \frac{((\lambda+2\mu)s_{zz}^* - \lambda s_{xx}^*)\,(\partial_z v_z)}{2\mu(\lambda+\mu)} \\ & + \frac{(s_{xz}^*)\,(\partial_x v_z + \partial_z v_x)}{\mu}dt,\end{aligned} \tag{4.15}$$

and

$$\mathbf{G}(\rho) = \frac{\partial\tilde{\Phi}}{\partial\rho} = -\sum_{src}\int \beta_x^*\partial_t v_x + \beta_z^*\partial_t v_z dt. \tag{4.16}$$

4.2.3 Source wavelet estimation

The source wavelet is an additional unknown in seismic data inversion. To estimate this unknown, Pratt (1999) takes into account the linearity of the source wavelet inverse

problem in FWI and retrieves formula to directly compute the optimal wavelet at each iteration. The expression for the source wavelet estimation in FWI is stated as

$$\mathbf{s}_P(f) = \frac{\mathbf{u}^T(f)\mathbf{d}^{obs}(f)}{\mathbf{u}^T(f)\mathbf{u}(f)}, \tag{4.17}$$

where $\mathbf{s}_P(f)$ is the source wavelet in frequency f, $\mathbf{u}(f)$ is the data modelled for an impulsive source and $\mathbf{d}^{obs}(f)$ represents the observed data in frequency (Pratt, 1999). In w-AWI, however, the source wavelet inverse problem is non-linear because of the definition of its residual (equation 4.8). Tarantola (1984) presented a gradient-based wavelet estimation approach using the adjoint-state variable. We propose estimating the source at each iteration using the quasi-Newton method.

The source wavelet must minimise the *same misfit functional used for the model parameter inversion* (equation 4.7). The unknown is the source wavelet and the model parameter is considered to be known at iteration k. The modelled data are defined as a convolution of the source wavelet and the Green's function. The data $\mathbf{d}(\mathbf{m}^k|t)$ of equation 4.9 is represented in the *frequency domain* at iteration k by

$$\mathbf{d}(\mathbf{m}^k|f) = \mathbf{s}_s(f)\mathbf{u}(\mathbf{m}^k|f), \tag{4.18}$$

where $\mathbf{s}_s(f)$ represents the source wavelet (or wavelet signature) defined in the frequency domain and $\mathbf{d}(\mathbf{m}^k|f)$ represents the Fourier transform in time of the modelled data.

The source wavelet is updated iteratively using

$$\mathbf{s}_s^{l+1} = \mathbf{s}_s^l - \mathbf{W}_s^{-1}\mathbf{G}_s, \tag{4.19}$$

where $\mathbf{G}_s = \partial\tilde{\Phi}(\mathbf{s}_s^l)/\partial\mathbf{s}_s^l$ represents the gradient of the misfit functional with respect to the source wavelet, and $\mathbf{W}_s = \partial^2\tilde{\Phi}(\mathbf{s}_s^l)/\partial(\mathbf{s}_s^l)^2$ is the inverse of the Hessian matrix (or an approximation). In this work, the approximate inverse of the Hessian is computed using the L-BFGS method (Nocedal, 1980). For the gradient, the first order derivative can be separated as

$$\frac{\partial\tilde{\Phi}(\mathbf{s}_s)}{\partial\mathbf{s}_s} = \frac{\partial\mathbf{d}(\mathbf{s}_s)}{\partial\mathbf{s}_s}\frac{\partial\tilde{\Phi}(\mathbf{s}_s)}{\partial\mathbf{d}(\mathbf{s}_s)}, \tag{4.20}$$

such that the needed gradient is written as

$$\mathbf{G}_s(\mathbf{s}_s) = \mathbf{u}^T(\mathbf{m}^k)\mathcal{F}\left\{\frac{\partial\tilde{\Phi}(\mathbf{s}_s)}{\partial\mathbf{d}(\mathbf{s}_s)}\right\}, \tag{4.21}$$

where $\mathcal{F}\{\partial\tilde{\Phi}(\mathbf{s}_s)/\partial\mathbf{d}(\mathbf{s}_s)\}$ corresponds to the Fourier transform in time of the residual source stated in equation 4.13 without including the term $\mathbf{K}^T$. For each source update, this approach needs computing the 2D Fourier transform. Back-propagation of residuals is *not* required.

4.2.4 Gradient regularisation and preconditioning

The convergence rate can be improved by gradient preconditioning and/or regularisation. We have implemented here depth-dependent preconditioning and gradient smoothing. The implemented approach is defined as follows.

Surface-wave amplitude decays exponentially with depth similarly to electromagnetic-wave amplitude. Therefore, an amplitude compensation used in resistivity imaging with electromagnetic data could also be applied in surface-wave inversion (Plessix and Mulder, 2008). We included a scaling of the model with depth such that the scaled model is

$$\hat{\mathbf{m}} = \frac{1}{z}\mathbf{m}, \tag{4.22}$$

which turns the objective function into

$$\hat{\Phi}(\hat{\mathbf{m}}) = \Phi(z\hat{\mathbf{m}}) = \Phi(\mathbf{m}), \tag{4.23}$$

and the gradient into

$$\partial_{\hat{\mathbf{m}}}\hat{\Phi}(\hat{\mathbf{m}}) = z\partial_{\mathbf{m}}\Phi(\mathbf{m}) \tag{4.24}$$

The scaled model is updated with

$$\hat{\mathbf{m}}^{k+1} = \hat{\mathbf{m}}^{k} - (\hat{\mathbf{W}}\partial_{\hat{\mathbf{m}}}\hat{\Phi}(\hat{\mathbf{m}}))^{k} \tag{4.25}$$

where $\hat{\mathbf{W}}$ is the approximate inverse of the Hessian for $\hat{\Phi}(\hat{\mathbf{m}})$ and $\hat{\mathbf{m}}$.

As a gradient regularisation approach, we smooth the gradient at each iteration by means of a 2D gaussian window. This window is computed as

$$g(x_0, z_0) = \frac{1}{\pi\alpha_x\alpha_z} e^{-\left(\frac{x-x_0}{\alpha_x}\right)^2} e^{-\left(\frac{z-z_0}{\alpha_z}\right)^2} \tag{4.26}$$

where (x_0, z_0) is the position of the sample to be smoothed in the model and $(\alpha_x,\ \alpha_z)$ define the shape and the dimensions of the 2D gaussian window.

4.3 Analysis of w-AWI

FWI is used to solve an ill-posed and non-linear inverse problem. The presence of local minima in the objective function is only avoidable if very low frequencies are considered (Bunks et al., 1995; Mulder and Plessix, 2008). Convergence toward a secondary minimum is commonly caused by the so-called cycle-skipping problems (phase difference of more than half the wavelength) (Bunks et al., 1995; Virieux and Operto, 2009). In FWI, the data may not suffer from cycle skipping in the $t - x$ domain. The w-AWI objective function is formulated in the $f - k$ domain and therefore an analysis of its behaviour with respect to cycle skipping and local minima in this domain must be performed.

The aim of this section is to analyse the properties and performance of the w-AWI objective function and to compare it to FWI using 2D synthetic data. Please note that we compute the gradients with respect to Lamé parameters and density (equations 4.14, 4.15 and 4.16), but inversion results are shown from now on in terms of Vp (P-wave velocity) and Vs (S-wave velocity) as these parameters are more commonly used in geophysics. Regarding the tests conducted in this section, we evaluate the objective function shape using a simple 2D velocity model. In our example, w-AWI shows a global minimum basin wider than FWI. We also evaluate the role of the window width in w-AWI approach. Finally, we show some examples where low-pass frequency filtering is necessary to avoid cycle-skipping effects in FWI but is not necessary in w-AWI.

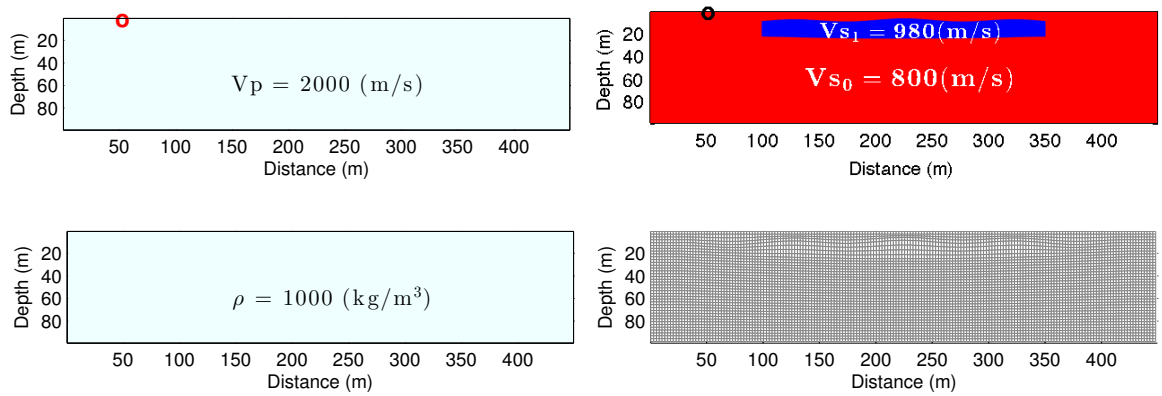


Figure 4.3: Elastic models used to compute synthetic data representing our observed data. Left: P-wave velocity (top) and density (bottom) homogeneous models. Right: S-wave velocity model (top) and the discretisation grid (bottom) conforming with the anomaly.

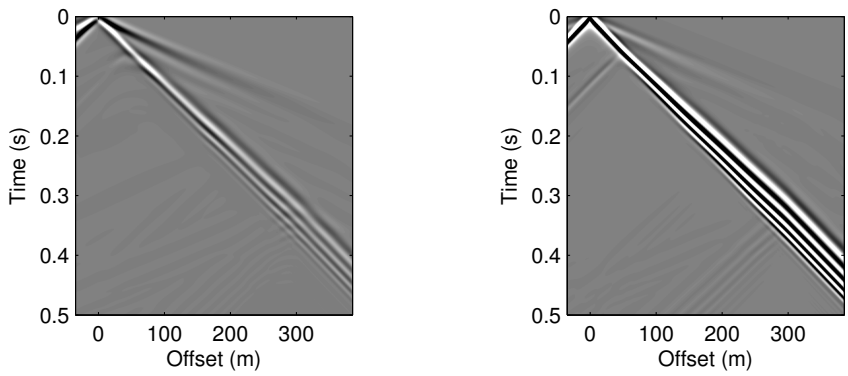


Figure 4.4: Horizontal (left) and vertical (right) velocity component of observed data. The data represent surface recordings and were computed for the model shown in Figure 4.3.

4.3.1 Objective function shape and local minima

We perform here a numerical evaluation of w-AWI and FWI objective functions using a simple model that depends only on two parameters. The 450×100 m model has constant P-wave and density (Vp = 2000 m/s and $\rho = 1000$ kg/m^3, see Figure 4.3). The S-wave velocity profile consists of a homogeneous background ($\mathrm{Vs}_0 = 800$ m/s) containing a non-rectangular velocity anomaly ($\mathrm{Vs}_1 = 980$ m/s) as shown in Figure 4.3 (top right). The grid conforms with the anomaly to prevent the appearance of non-physical diffractions (Figure 4.3 bottom right). Free-surface conditions are imposed on the top boundary of the model and CPML conditions are set on the lateral and bottom boundaries.

The observed data are the horizontal and vertical velocity components. The source considered is a 40 Hz Ricker wavelet applied on the vertical component of velocity and placed 0.5 m below the surface at the horizontal position of 50 m (see Figure 4.3). A total of 450 receivers placed at the source depth and equally spaced of 1 m apart were considered to compute the vertical and horizontal components of velocity. The observed

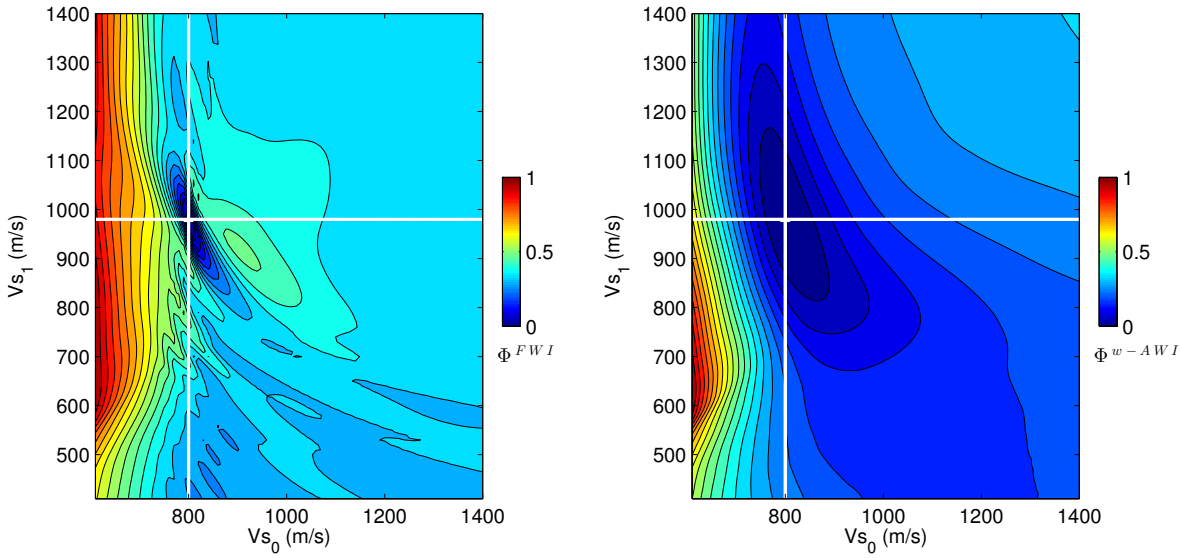


Figure 4.5: Normalised objective function value with respect to two parameters: background (Vs_0) and anomaly (Vs_1) velocities. Left: FWI. Rigth: w-AWI.

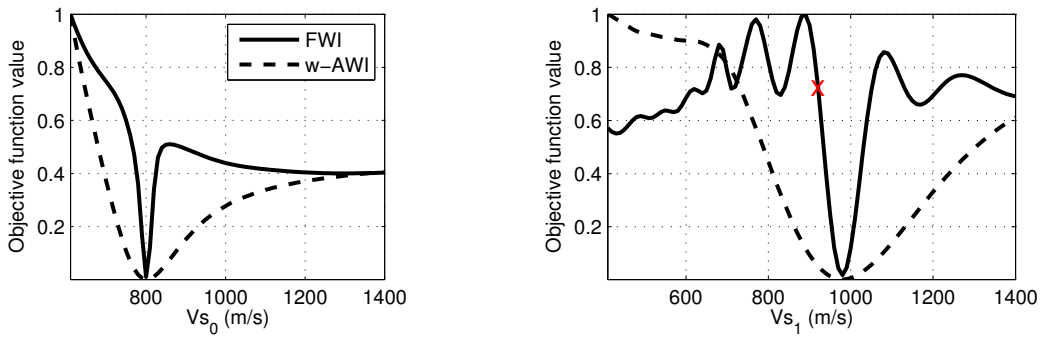


Figure 4.6: Normalised objective function curves taken from Figure 4.5 at Vs_1=980 m/s (left) and Vs_0 = 800 m/s (right). The red 'x' on the right-hand side plot indicates Vs_1=920 m/s.

shot gathers are displayed in Figure 4.4. The observed data contains dispersive surface waves that are created when the direct S wave impinges onto the velocity anomaly. A direct P wave and some other low-energy reflections from surface waves can be observed as well.

We compute the objective function values using several velocity models with different background and anomaly S-wave velocities. The P-wave velocity and density profiles remain unchanged and equal to the exact model. The S-wave background velocities range from 600 m/s to 1400 m/s whereas the anomaly velocities range from 400 m/s to 1400 m/s. In w-AWI computation, the overlapping windows used for data separation are spaced 4 m and are 128 m wide. The resulting objective function for FWI and w-AWI are shown in Figure 4.5. This example shows that w-AWI has a wider global minimum basin than FWI. As an interpretation of this result, one can say that convergence with w-AWI is guaranteed for the complete evaluated velocity range while FWI should converge

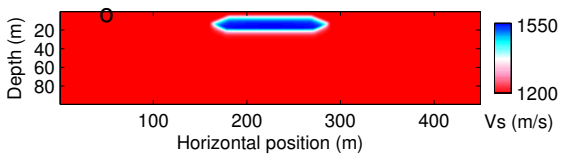


Figure 4.7: S-wave velocity model Vs_0 composed of a constant background and an anomaly. The background velocity is $Vs_0 = 1200$ m/s. In this image, the anomaly velocity is $Vs_1 = 1.15Vs_0$. The black circle indicates the source position.

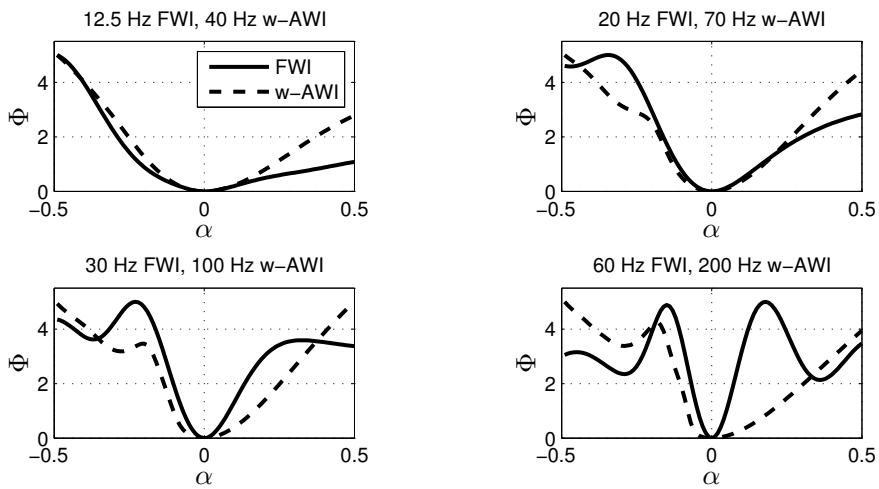


Figure 4.8: Objective function with respect to one model parameter α and computed with Ricker wavelets centred at different frequency values (as indicated in each plot title). The legend is the same for the four plots.

only for a reduced zone (i.e. the global minimum basin). Two slices of the objective functions of Figure 4.5 are shown in Figure 4.6. Those slice plots are taken (1) for the correct background velocity and varying anomaly velocity and (2) for the correct anomaly velocity and varying background velocity as indicated by the white lines in Figure 4.5. We can observe that for both tests the w-AWI basin is at least 3 times wider than the FWI basin. However, the shape of w-AWI basin does not seem to conform with a typical quadratic function.

As an additional test, we computed FWI and w-AWI objective functions using Ricker wavelets with different central frequencies. We aim at evaluating the frequency value that FWI needs to provide a global minimum basin having similar width than that of w-AWI. The P-wave and density models are homogeneous (Vp=2500 m/s and $\rho = 1000$ kg/m^3). The S-wave velocity model is homogeneous (Vs_0=1200 m/s) and contains an anomaly with velocity $Vs_1 = (1 + \alpha)Vs_0$, where α is the anomaly parameter. The exact S-wave velocity model is homogeneous ($\alpha = 0$) and, as initial models, we consider several values of the anomaly parameter ranging from $\alpha = -0.5$ to $\alpha = 0.5$. An example of initial velocity model, with $\alpha = 0.15$, is shown in Figure 4.7.

The results are shown in Figure 4.8. The FWI and w-AWI global minimum basins are similarly wide if a 12.5 Hz Ricker wavelet is used in FWI and a 40 Hz one is used in

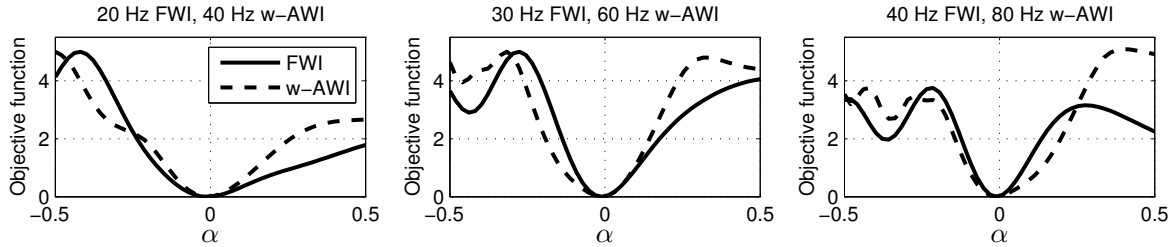


Figure 4.9: Same as Figure 4.8 but using monochromatic signal sources. The legend is the same for the three plots.

w-AWI (Figure 4.8, top left). Moreover, there is equivalence between FWI at 20 Hz and w-AWI at 70 Hz (Figure 4.8, top right), FWI at 30 Hz and w-AWI at 100 Hz (bottom left) and finally FWI at 60 Hz and w-AWI at 200 Hz (bottom right). In average, the width of the FWI global minimum basin is similar to the w-AWI one if the FWI Ricker wavelet central frequency is approximately 0.3 times the frequency of the w-AWI one.

The previous test can be complemented by considering monochromatic source time functions instead of the Ricker wavelet. The results (Figure 4.9) indicate that FWI and w-AWI global minimum basins are similar if the frequency used in FWI is approximately half the frequency used in w-AWI.

The w-AWI objective function has a wider global minimum basin when using both a broad-band source and a monochromatic source, which might be the result of a better frequency weighting in w-AWI. As a consequence, the minimum source frequency related to the choice of an initial velocity model is at least two times higher in w-AWI than in FWI.

4.3.2 Cycle skipping

We present an analysis of the cycle skipping in FWI and w-AWI. We consider the observed data obtained with the model of Figure 4.3, and the S-wave velocity initial models shown in Figure 4.10. Two S-wave initial models are homogeneous with Vs_0=600 m/s (model 1) and Vs_0=800 m/s (model 2). The third initial model has background velocity Vs_0=800 m/s and contains an anomaly with Vs_1=920 m/s (model 3). The modelled data ($\mathbf{d}$) for the three different models are shown in Figure 4.11. Dispersive surface waves are present only in the data computed for the model containing the anomaly. For the two other cases, the Rayleigh wave velocity is constant because the model is homogeneous. In Figures 4.12, 4.13 and 4.14, we compare seismic traces of observed and modelled data at three different offsets indicated by the red dashed lines in Figure 4.11. For model 1, the Rayleigh wave of modelled data has cycle skipping for the three offsets. For model 2, the Rayleigh waves propagate with similar velocities but the dispersion generated by the velocity anomaly create phase problems in traces 2 and 3 (Figure 4.13). Finally, none of the three traces for model 3 are cycle skipped because the initial model anomaly is only 6 % lower than the exact one. Indeed, model 3 is the only initial model falling into the basin of the objective function global minimum (the red cross in Figure 4.5, right).

We now consider the input data processed in the $f-k$ domain. First, we present

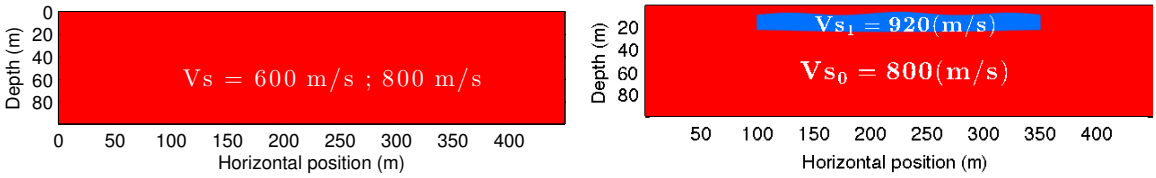


Figure 4.10: Initial S-wave velocity models. Synthetic data are computed for three different initial models: model 1 and 2 (left) are homogeneous with Vs=600 m/s and Vs=800 m/s, and model 3 has homogenous background with Vs_0 = 800 m/s and one anomaly with Vs_1 = 920 m/s (right).

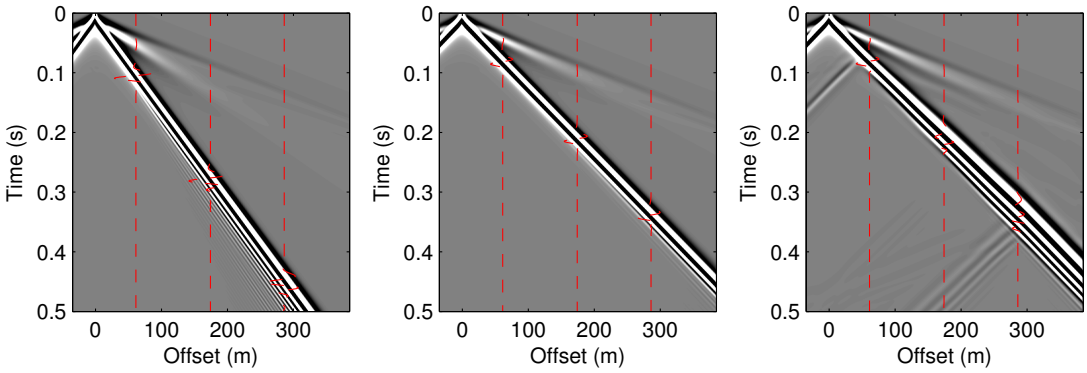


Figure 4.11: Initial data (vertical component of velocity) computed with the S-wave velocity models of Figure 4.10. Left: model 1. Centre: model 2. Right: model 3.

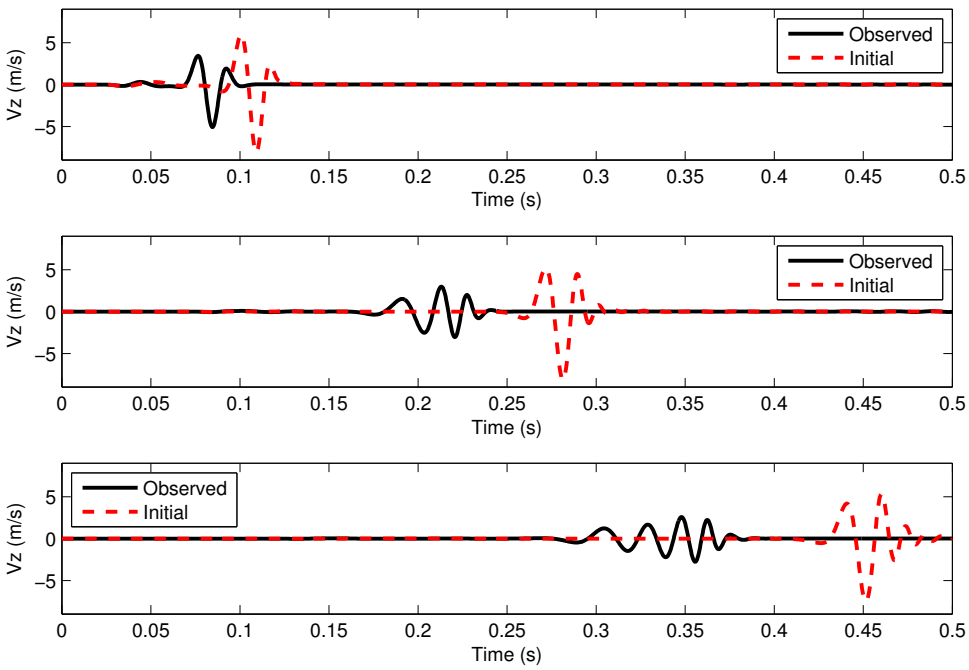


Figure 4.12: Observed and modelled data for *model 1* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom).

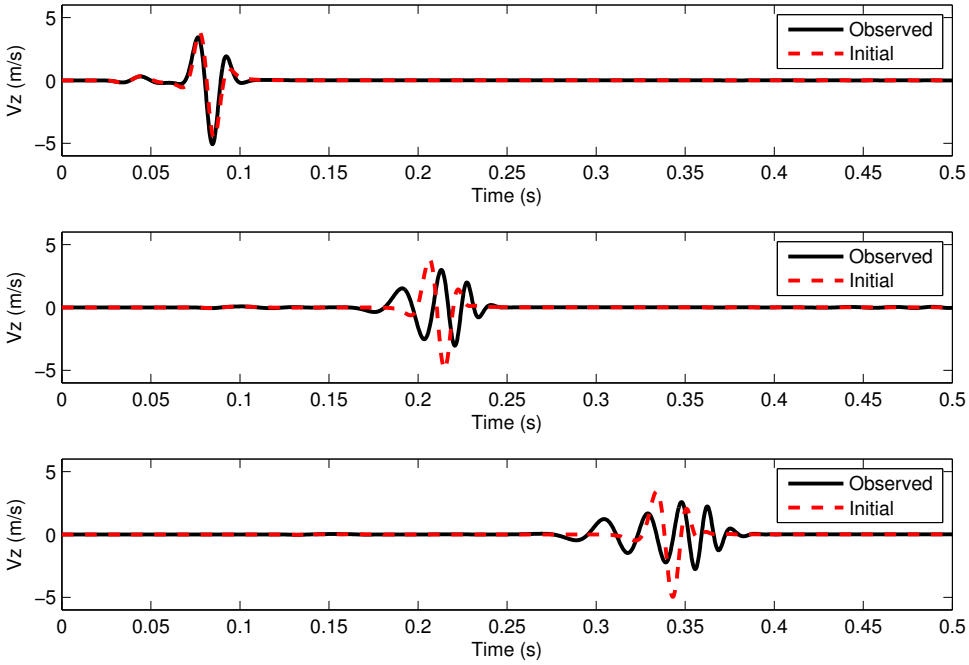


Figure 4.13: Observed and modelled data for *model 2* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom).

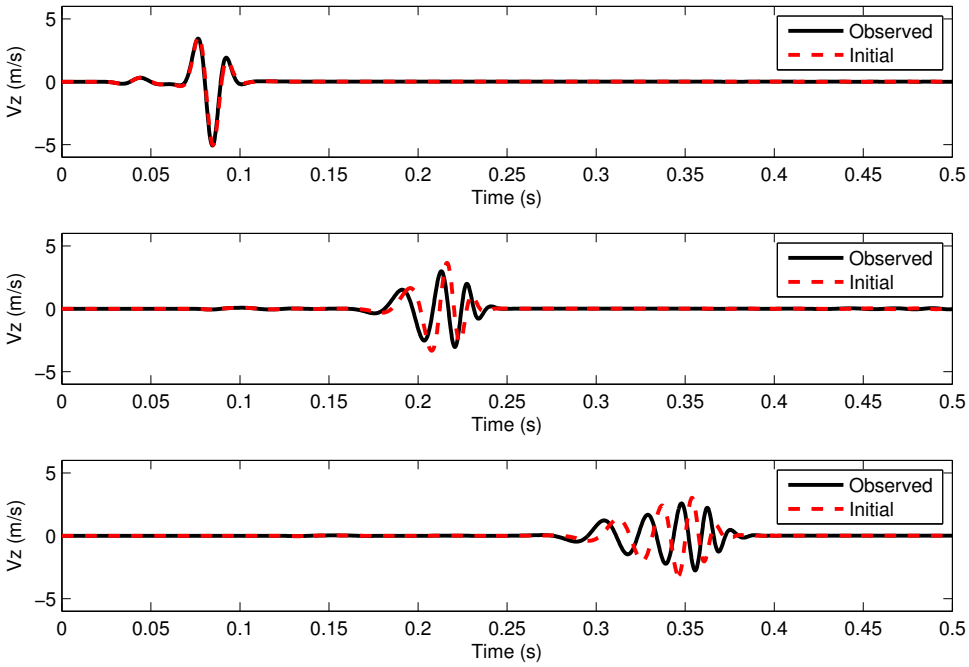


Figure 4.14: Observed and modelled data for *model 3* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom).

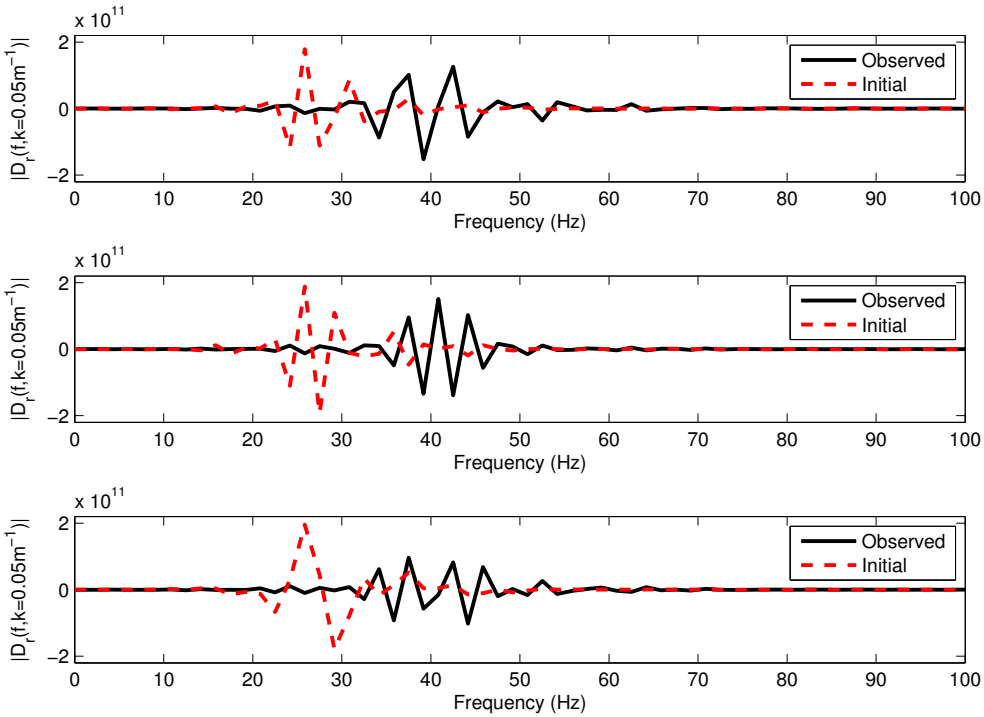


Figure 4.15: Real part of the 2D Fourier transform of observed and modelled data for *model 1* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

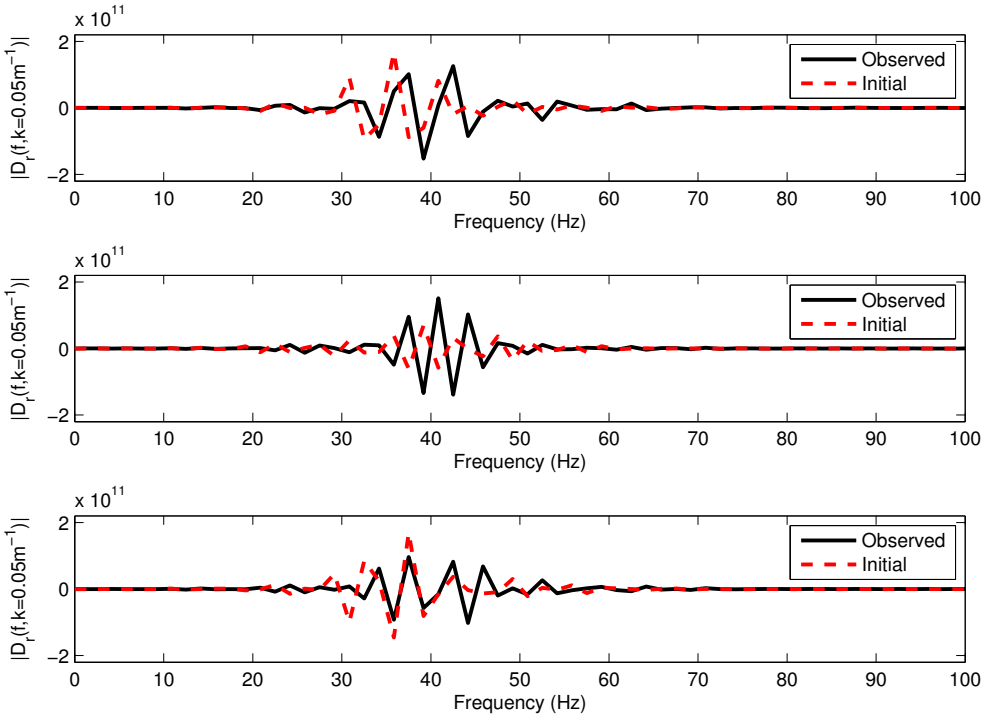


Figure 4.16: Real part of the 2D Fourier transform of observed and modelled data for *model 2* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

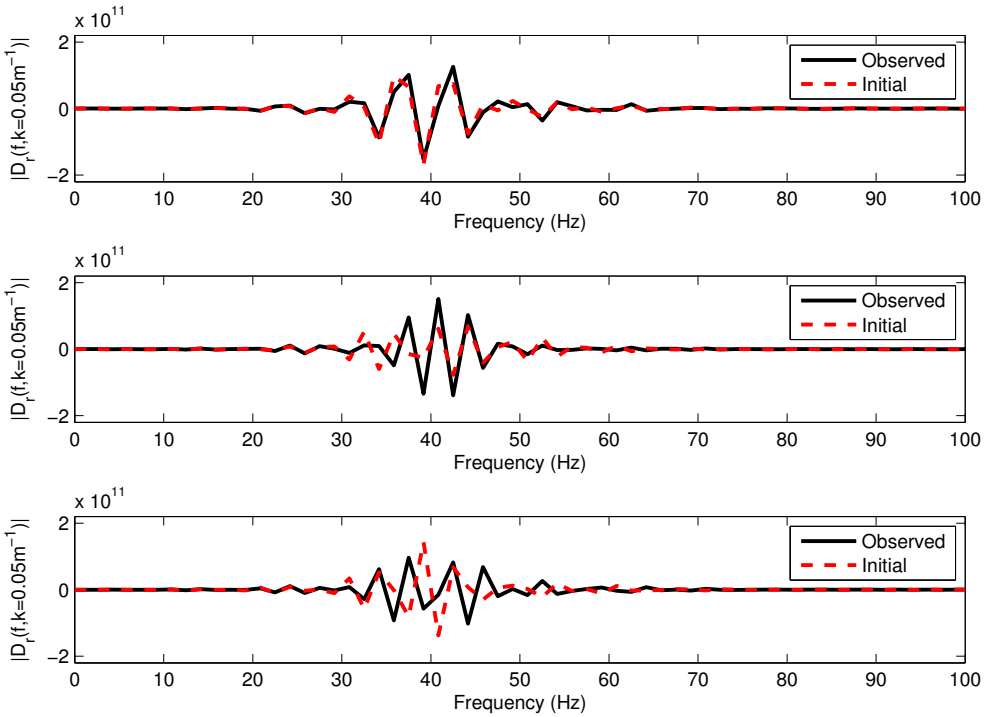


Figure 4.17: Real part of the 2D Fourier transform of observed and modelled data for *model 3* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

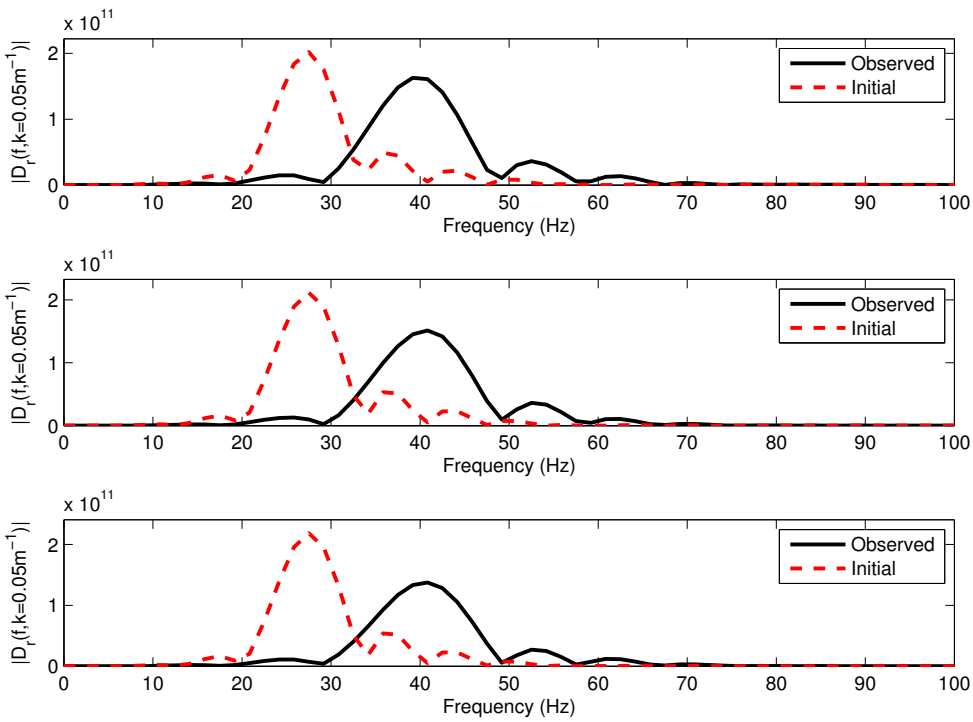


Figure 4.18: Absolute value of the 2D Fourier transform of observed and modelled data for *model 1* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

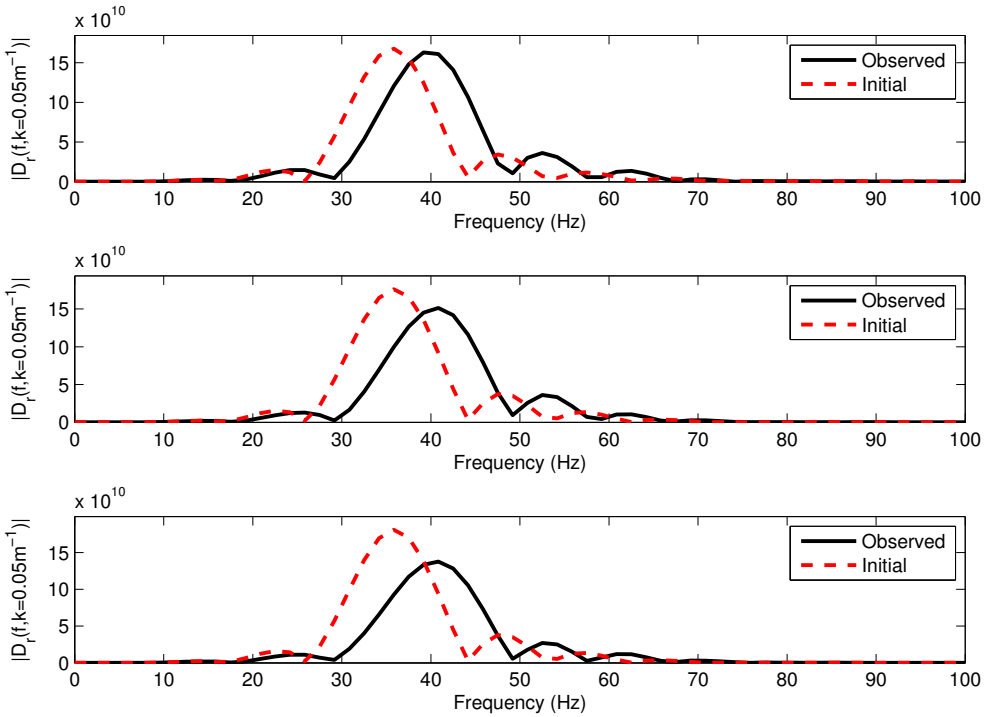


Figure 4.19: Absolute value of the 2D Fourier transform of observed and modelled data for *model 2* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

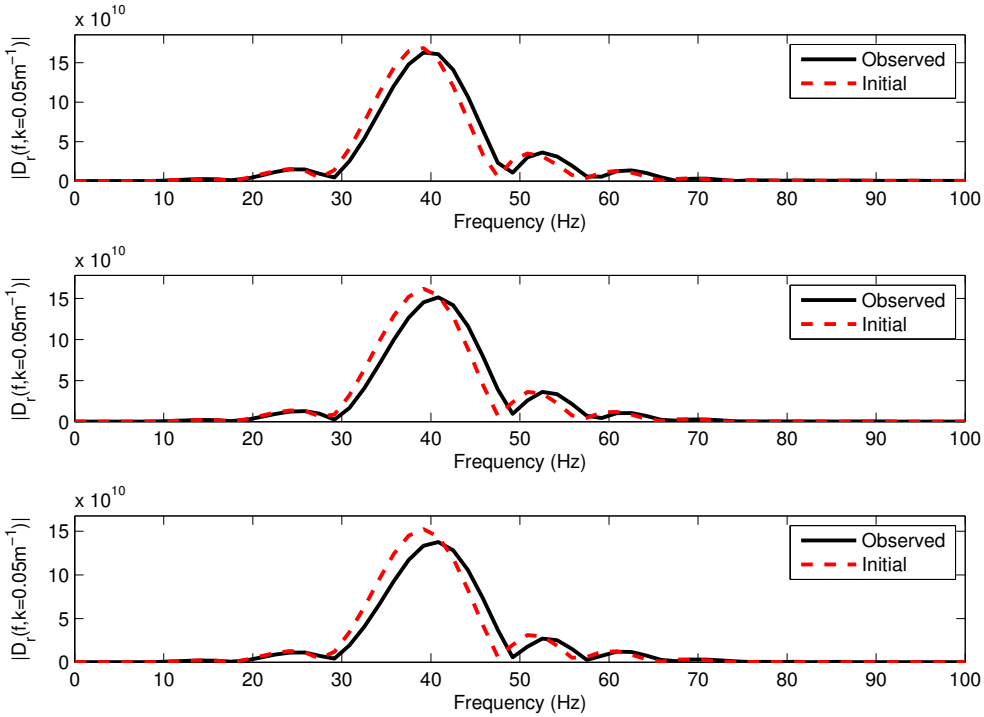


Figure 4.20: Absolute value of the 2D Fourier transform of observed and modelled data for *model 3* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom).

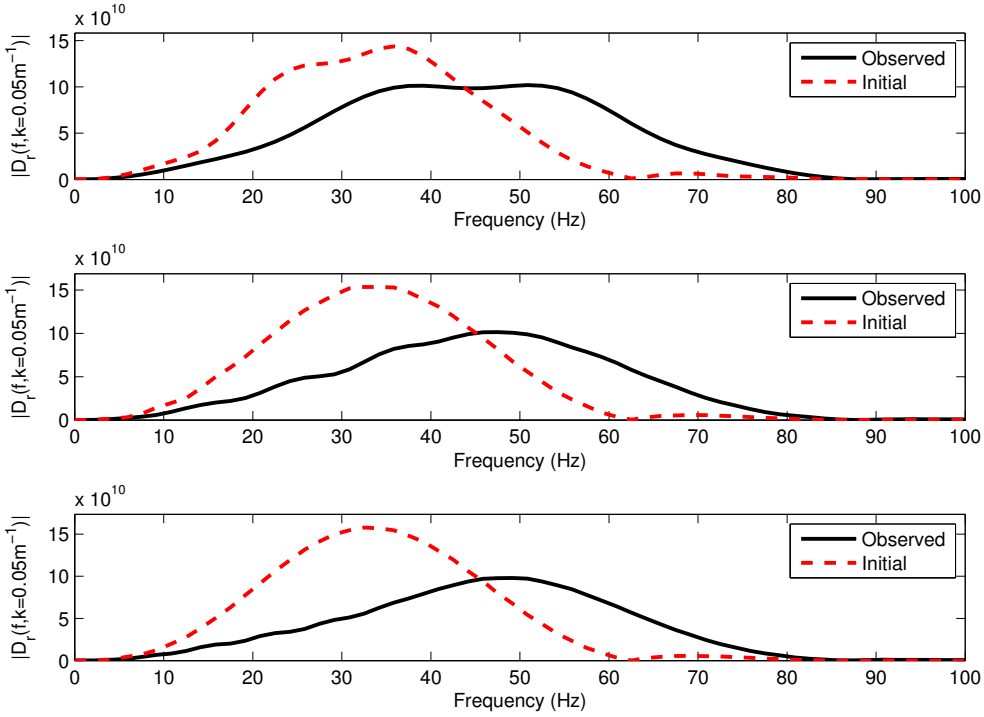


Figure 4.21: Same as Figure 4.18 (*model 1*) but using $l_w = 64$ m windows.

comparisons of sections taken at $k = 0.05$ m^{-1} in the real part of the $f - k$ spectra of the seismic responses of the initial and exact models. We computed the 2D Fourier transform of three subsets of the shot gathers using windows whose centres coincide with the selected three traces shown in Figures 4.12 to 4.14. The window width is $l_w = 128$ m. Results are shown in Figures 4.15, 4.16 and 4.17. For model 1 (Figure 4.15), the slices are not in correct cycle matching for any of the three positions similarly to the result of Figure 4.12. For model 2 (Figure 4.16), the slices have less cycle mismatch than for the seismic traces (Figure 4.13), however, it is not clear if they are in correct cycle matching. Finally, as for seismic traces (Figure 4.14), the three slices associated to model 3 are in correct cycle match (Figure 4.17) because the initial model is quite close to the exact model.

Contrary to the seismic traces and the real part of Fourier transformed data, the absolute value of Fourier transformed data can be in correct cycle matching even for the most different exact and initial models of these tests. We computed the 2D Fourier transform of observed and initial data using the same spatial selection windows as for the real part of the $f - k$ spectra (shown previously). In Figures 4.18, 4.19 and 4.20, we show the amplitude of spectra for the same wavenumber value ($k = 0.05$ m^{-1}) as for the real part. The comparisons show that observed and initial transformed data are in correct cycle match for models 2 and 3 (Figures 4.19 and 4.20). Note however that there is a cycle mismatch in model 1 (Figure 4.18) We have found that it can be avoided if a smaller spatial window is used. As an example, we show in Figure 4.21 the slices for model 1 when the windows width is reduced to $l_w = 64$ m. We observe that the energy of the fundamental mode is more extended in the frequency domain than for the result using the wider window. This result is interesting when designing strategies for

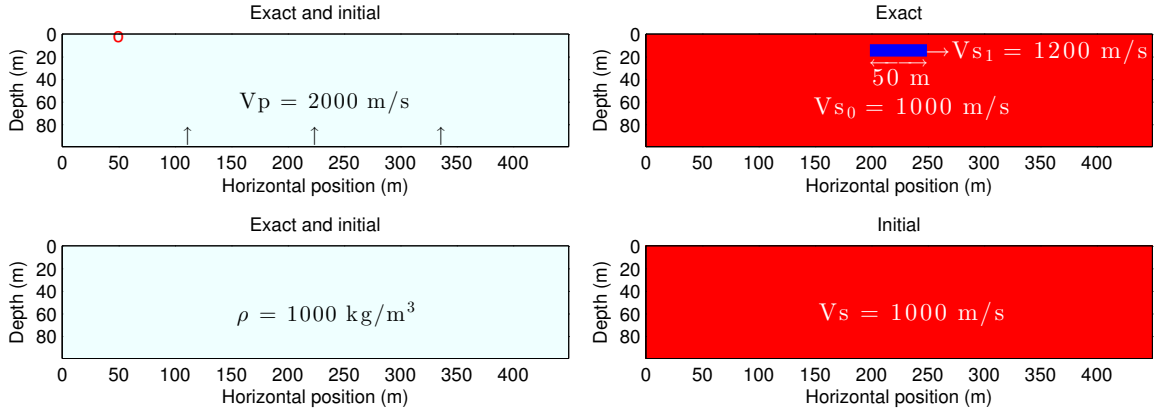


Figure 4.22: Exact and initial elastic models. The exact and initial P-wave velocity (top, left) and density (bottom, left) models are homogeneous. The exact (top, right) and initial (bottom, right) S-wave velocity models only differ for the velocity inclusion.

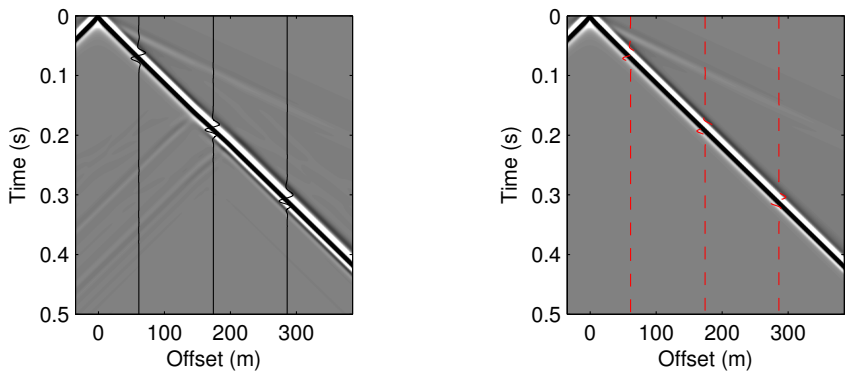


Figure 4.23: Shot gathers showing the vertical component of velocity. Observed data (left) and initial data (right).

convergence from initial models that are significantly different from the exact model. However, implementing shorter windows may generate losses in resolution when applying the 2D Fourier transform.

The images presented in this section support the results shown in section 4.3.1 as there are no visible secondary minima in w-AWI for the evaluated range of velocities (see Figure 4.5). However, this result does not mean that w-AWI does not have secondary minima. This example is a numerical test that shows how wide the global minimum basin can be with w-AWI with respect to FWI. We use this result as a motivation for inversion examples where the FWI reconstructed model does not converge because of cycle-skipping phenomena generating local minima effects, and we check if w-AWI would instead converge.

4.3.3 Width of spatial windows

In this section, we analyse the influence of the width of the spatial window in w-AWI (equations 4.9 and 4.10) on the smoothness of the objective function. We consider ob-

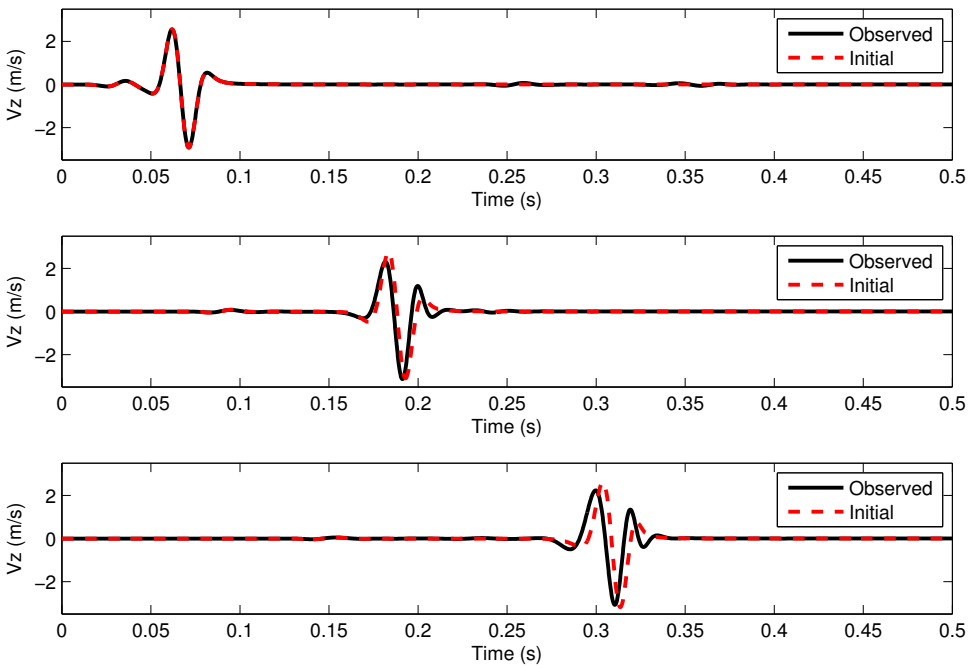


Figure 4.24: Observed and modelled seismic traces taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom).

served and initial data computed for two simple elastic models. The exact and initial P-wave and density models are homogeneous with Vp = 2000 m/s and $\rho = 1000$ kg/m^3 as shown in Figure 4.22 (left). The S-wave exact model consists of a homogeneous background (Vs = 1000 m/s) containing a velocity anomaly (Vs = 1200 m/s) as shown in Figure 4.22 (top, right). The initial S-wave velocity model is completely homogeneous (Vs = 1000 m/s, see Figure 4.22 bottom right). The observed and initial data are the vertical component of velocity recorded by 450 receivers equally spaced 1 m at a depth of 0.5 m. The source time function is Ricker wavelet with a dominant frequency of 40 Hz placed at the horizontal position of 50 m at 0.5 m depth, as indicated by the red circle in the P-wave model of Figure 4.22. The computed shot gathers are shown in Figure 4.23. The main difference is the presence of slightly dispersive waves and reflections in the observed data generated by the anomaly, whereas for the initial data the Rayleigh wave has constant velocity and no reflections are present. Three seismic traces, before/within/after the anomaly, as indicated in Figure 4.23, are compared in Figure 4.24. The traces before the anomaly (top of the figure) are almost identical (the difference are the very low-energy reflections at 0.25 s and 0.35 s). The traces within and after the anomaly are more different. Please note that the Rayleigh wave after the anomaly should have the same velocity as before the anomaly, however the seismic traces after the anomaly are different as the waveform is modified when it passes through the anomaly. This seems to be quite obvious but later on we will show that the input data of w-AWI behaves differently, depending on the length of the employed spatial windows.

In Figure 4.25, we show the 2D Fourier transform of the whole observed (on the left) and initial (on the right) shot gathers. The two sets of transformed data do not seem to be identical but it is not possible to localise the anomaly as we could do with the shot

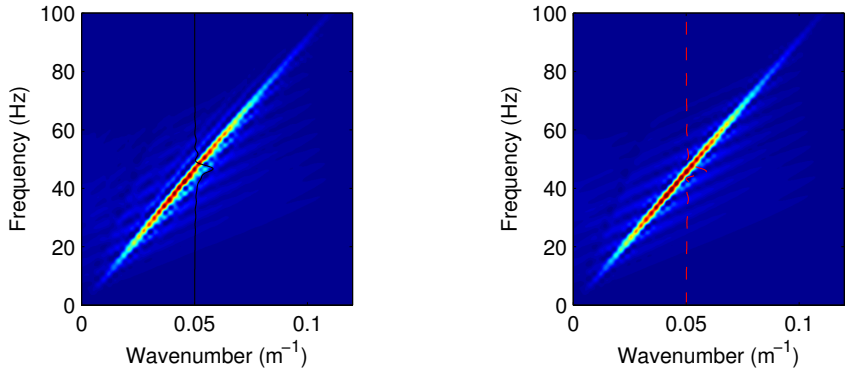


Figure 4.25: Absolute value of the 2D Fourier transform of observed (left) and initial (right) data. The window used here contains the whole set of traces ($l_w = 450$ m).

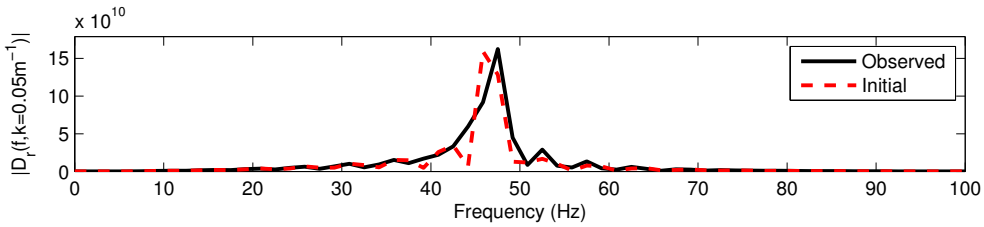


Figure 4.26: Slices of the data shown in Figure 4.25 taken at $k = 0.05$ m^{-1}.

gathers of Figure 4.23. Moreover, we can take a slice of the data for a given wavenumber value, as shown in Figure 4.26, to confirm that by taking the 2D Fourier transform of the whole shot gather it is not possible to know where the anomaly is localised. That is the consequence of ignoring the phase in the computation of the objective function in w-AWI. This is the reason for using a spatial windowing that reintroduces some information about the lateral localisation of model misfits (i.e. velocity anomalies).

We use several windows to gather subsets of receiver traces before applying the 2D Fourier transform. Each 2D window includes a $\sin^2(2\pi\alpha)$ taper at the four boundaries, with α being the length of the taper to avoid ringing when applying the Fourier transform. Three different window lengths are considered: l_w =256 m, 128 m and 64 m. In Figure 4.27, we show the amplitude of the 2D Fourier transform of the observed and initial data for three windows with l_w =256 m whose centres are indicated by the arrows of Figure 4.22 (top left). Observed and initial data spectra are different, but it is not possible to identify the anomaly position (Figure 4.27). The slice plot (Figure 4.28) comparing initial and observed data for a given wavenumber value displays a clear misfit for the three windows. But in this case, the parameter l_w defines a long window that includes the receivers above the anomaly in all three windows. The spectra for windows with $l_w = 128$ m are shown in Figure 4.29. Even if spectra seem similar, the slice comparison (Figure 4.30) reveals differences only in one of the three windows: the window centred above the velocity anomaly. In this case, the other two windows do not include any of the receivers located above the anomaly. For $l_w = 64$ m (Figure 4.31), the energy of the fundamental mode is extended and is not as concentrated in the frequency domain as for the previous cases. In the slice comparison (Figure 4.32), we observe that the only

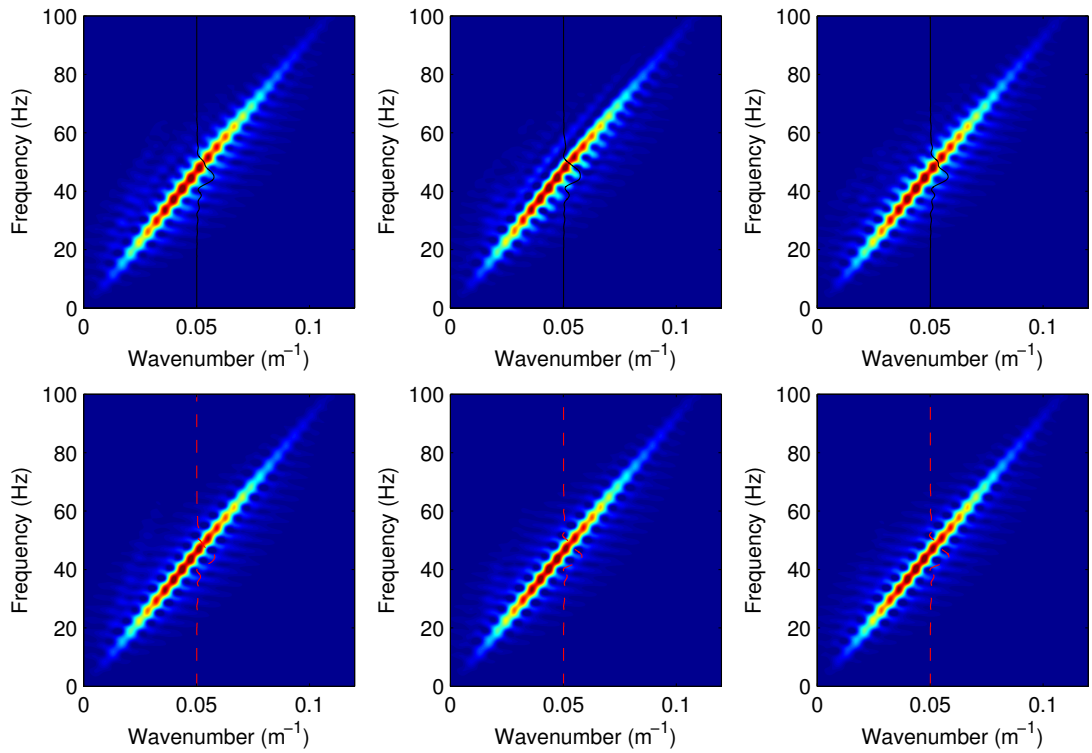


Figure 4.27: Absolute value of the 2D Fourier transform of observed (top) and initial (bottom) data. The $l_w = 256$ m windows are centred at 62 m (left), 174 m (middle), and 286 m (right).

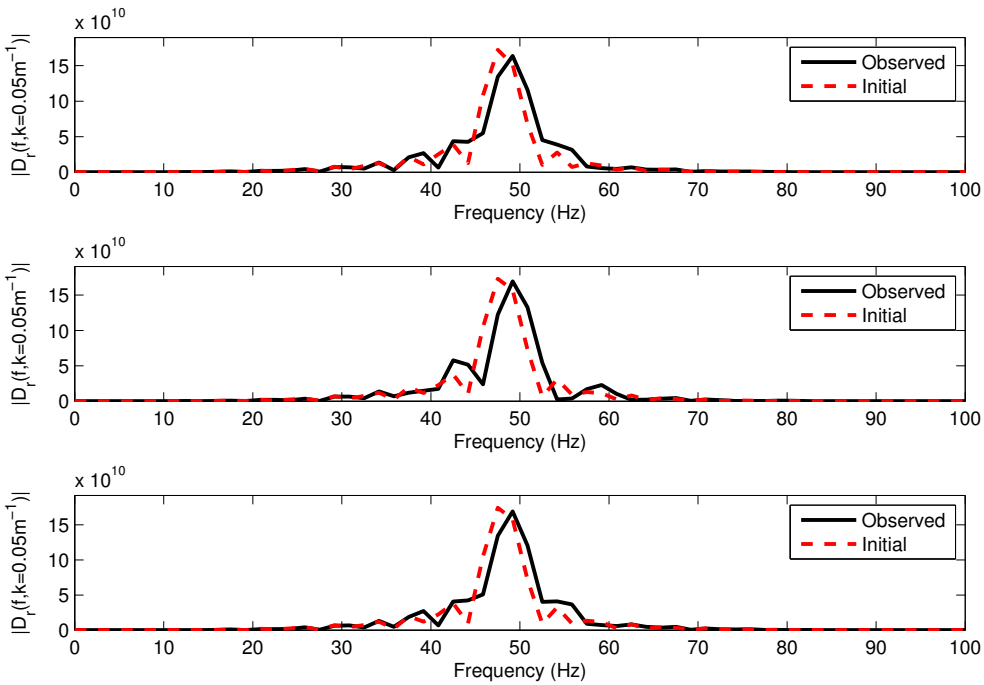


Figure 4.28: Slices of the data shown in Figure 4.27 taken at $k = 0.05$ m^{-1}. The $l_w = 256$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom).

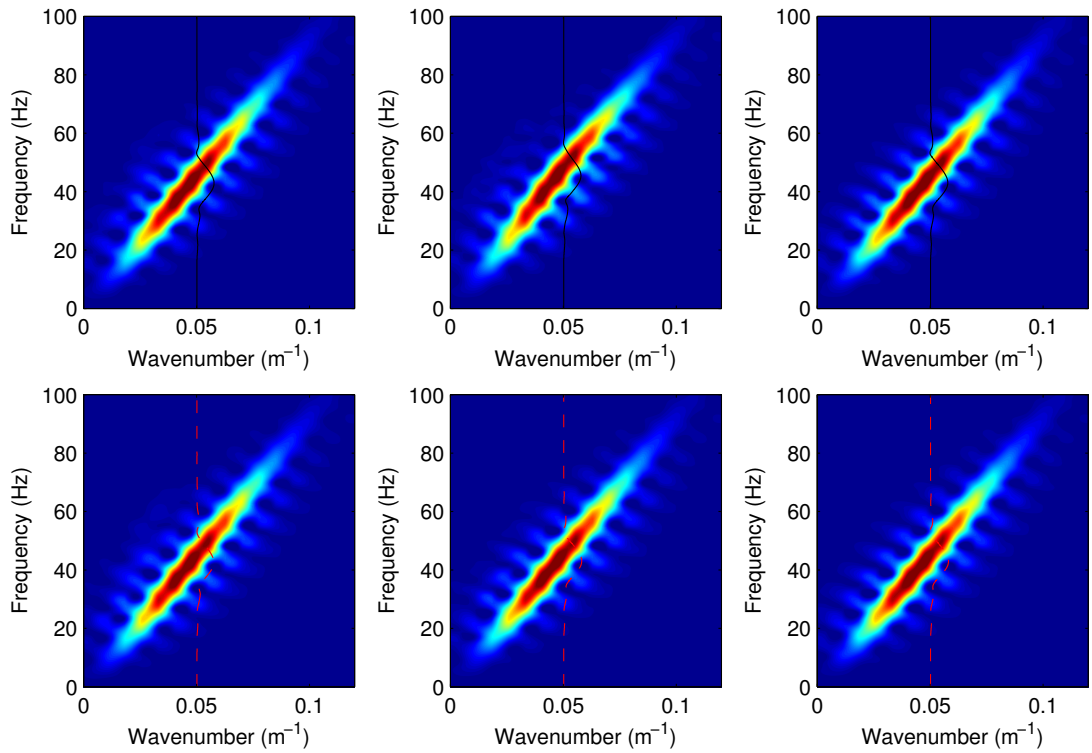


Figure 4.29: Same as Figure 4.27 but using $l_w = 128$ m windows.

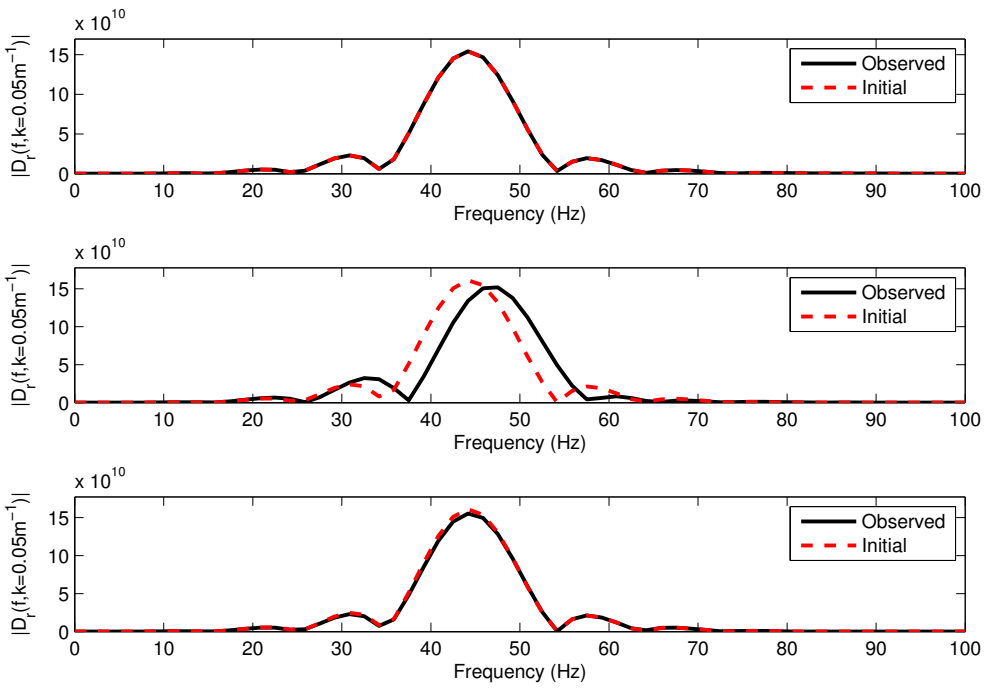


Figure 4.30: Slices of the data shown in Figure 4.29 taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom). There is almost no difference in the bottom panel because the corresponding windows mostly select receivers (222-350 m) that are not above the anomaly (200-250 m).

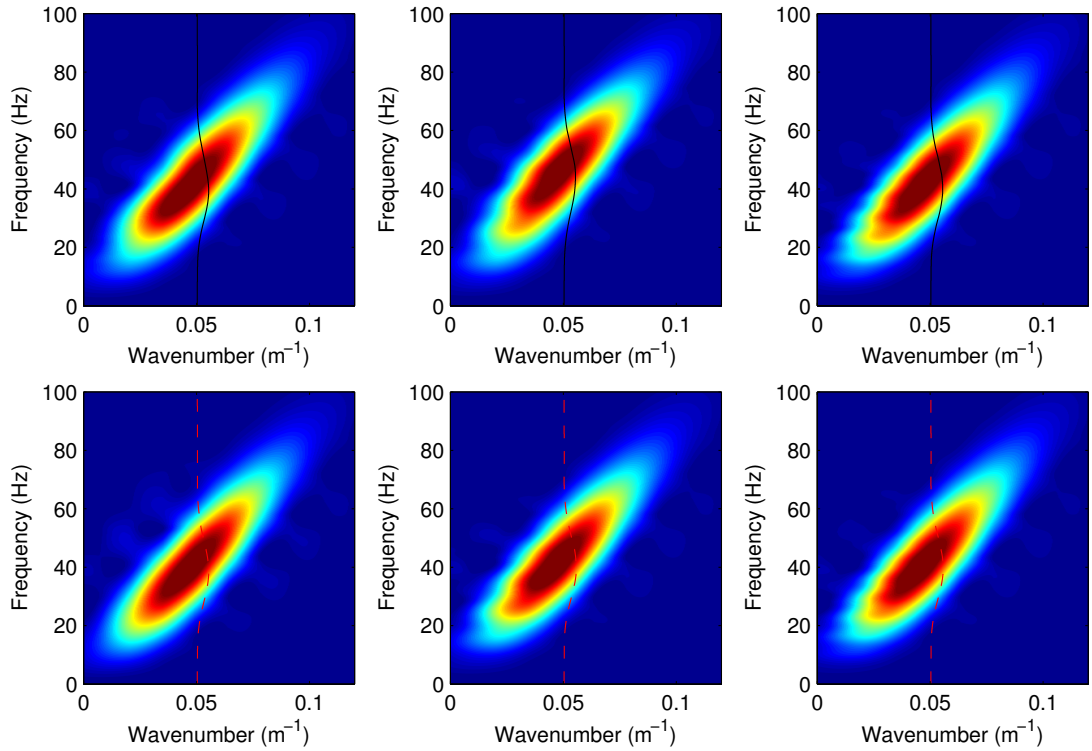


Figure 4.31: Same as Figure 4.27 but using $l_w = 64$ m windows.

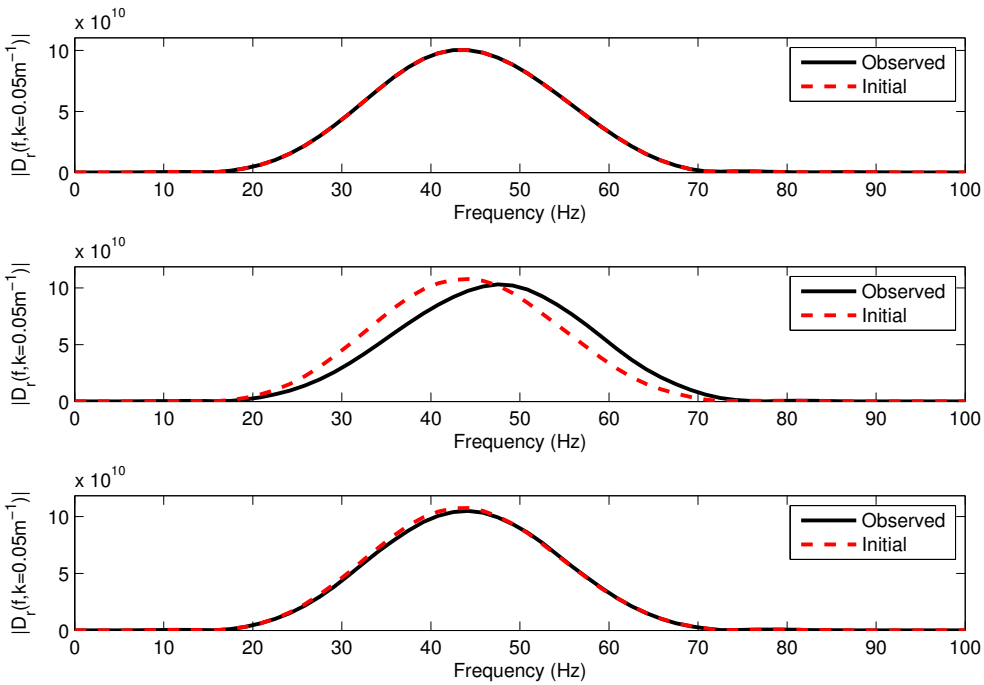


Figure 4.32: Slices of the data shown in Figure 4.31 taken at $k = 0.05$ m^{-1}. The $l_w = 64$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom).

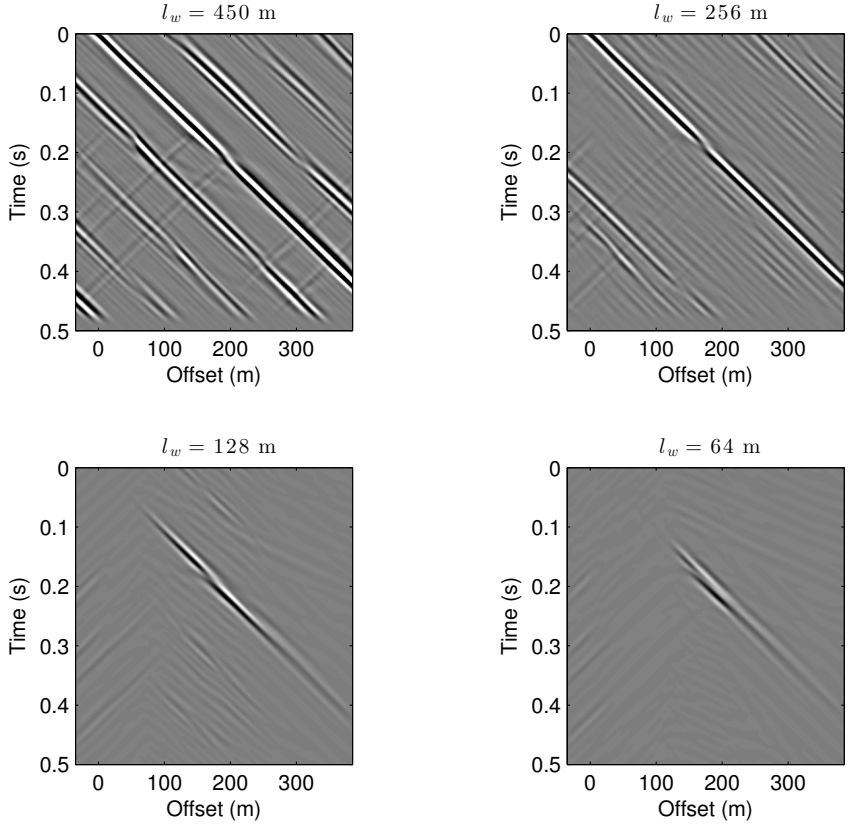


Figure 4.33: Adjoint sources computed with windows of different widths l_w. All the windows centres are separated by 4 m, at the exception of the one with $l_w = 450$ m.

window with a remarkable misfit is the one including receivers located above the anomaly (i.e. the central window).

4.3.4 Adjoint source and gradient

In this test, we evaluate the adjoint-state residual source and the gradient for w-AWI. The observed and initial data are the ones shown in Figure 4.23. We compute the residual source related to the vertical component of velocity. Different than for the test in section 4.3.3, we use 4 sliding windows with $l_w = 450, 256, 128$ and 64 m. For the $l_w = 450$ m window, the complete shot gather is transformed only once because the window width is equal to the shot gather width. For the other windows, several windows are used to separate the shot gather into subsets of receiver traces. Their centres are separated by 4 m. Please note that the computation of the residual source consists in separating the shot gather into windows of width l_w, then computing equation 4.13 which includes the sum of the 2D inverse Fourier transform over the windows.

The adjoint sources computed with the 4 different window widths are shown in Figure 4.33. For the 450 m window, the adjoint source contains several repetitions of the residual misfit. This effect is a consequence of the absolute value that neglects the phase information and thus the location of events in the $t-x$ domain. For the 3 other windows, the adjoint source contains more concentrated energy with less repetitions. The source

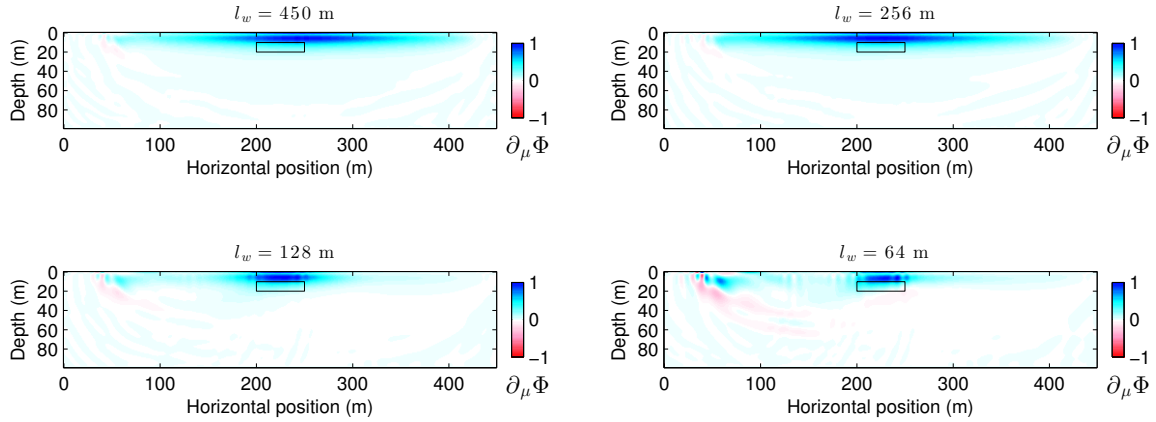


Figure 4.34: The w-AWI gradient computed with the window configurations of Figure 4.33. The position of the exact velocity anomaly of the model is contoured in black.

obtained with the 256 m wide window contains less repetitions as some information about the phase is reintroduced thanks to the implementation of the localised windows. However, the energy is excessively spread over the middle section which could prevent the localisation of the anomaly in the gradient. The 128 m window provides a good adjoint source as the anomaly location can be identified by the energy concentration towards the centre of the gather. The 64 m window offers the clearest adjoint source as most of the energy is concentrated between offsets 150 m and 200 m, which is the actual location of the anomaly.

The gradients obtained with the 4 sliding windows are shown in Figure 4.34. The use of the 450 m window leads to a gradient with maximum amplitude distribution non-centred with respect to the exact anomaly position because the phase is neglected. Besides, the amplitude is excessively extended outside of the anomaly limits (Figure 4.34). This gradient would update the model in an erroneous direction. The gradient computed with the 256 m window has an amplitude centred with respect to the anomaly position. However, the amplitude is still extended outside the anomaly position. For the last two window widths, the amplitude distribution is correctly centred with respect to the anomaly and coincides more with its location. However, the gradient obtained with the 64 m wide window exhibits some amplitude oscillations that might affect the model update. We conclude that even if the 128 m window is roughly 2.5 times longer than the anomaly (50 m as shown in Figure 4.22), the gradient shows a good trade-off among all the analysed results.

4.3.5 Low frequencies

Low frequencies in the data can help reducing the phase mismatch between observed and modelled data. Multi-scale approaches have been proposed to reduce the phase mismatch by starting the inversion with lowpass-filtered data and then gradually introducing higher frequency contents (Bunks et al., 1995; Brossier et al., 2009; Romdhane et al., 2011; Köhn et al., 2012). We have not yet established if a multi-scale approach would be suitable in waveform inversion of surface waves as it is the case for transmitted/diving wave

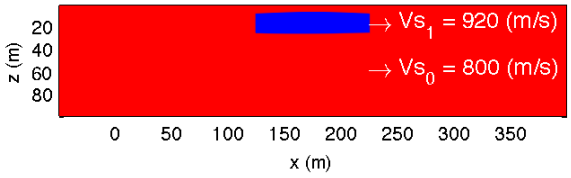


Figure 4.35: S-wave velocity profile consisting of a homogenous background and a velocity anomaly.

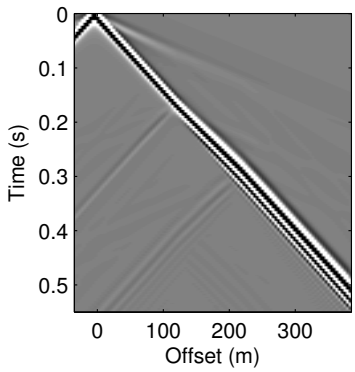


Figure 4.36: Seismic common shot gather showing the vertical component of velocity recorded near the surface. The source signature is a 40 Hz Ricker wavelet. Source excitation is along the vertical direction.

inversion (Mulder and Plessix, 2008; Plessix et al., 2012). In this paragraph, we analyse the advantage offered by inversion of low frequencies in shot gathers containing surface waves.

We consider a 450×100 m velocity model whose P-wave velocity and density models are homogeneous (Vp = 2000 m/s and ρ = 1000 kg/m^3). The S-wave velocity model (Figure 4.35) consists of a homogeneous background (Vs_0 = 800 m/s) with a non-perfectly rectangular velocity anomaly (Vs_1 = 920 m/s). Unphysical diffractions are avoided by employing a discretisation grid conforming with the anomaly shape.

The observed data are a synthetic shot gather (a single source). The source time function is a 40 Hz Ricker wavelet and is applied on the vertical component of velocity, placed at 0.5 m depth and x=0 m (Figure 4.35). A total of 112 receivers placed at the source depth and spaced 4 m apart have recorded the vertical component of velocity (Figure 4.36). The observed data contains dispersive surface waves that are created when the direct S wave impinges on the velocity anomaly. A direct P wave and some low-energy reflections of surface waves can be observed as well.

A seismic shot gather computed with the same geometry as the one employed for the observed data but with a second velocity model is used as initial data. The second model has the same velocity and density values as the exact model but does not contain the velocity anomaly. In Figure 4.37 (top), three groups of 16 seismic traces of the observed shot gather are compared to the corresponding traces of modelled data in the same offset intervals. From Figures 4.35 and 4.37, it can be established that the observed

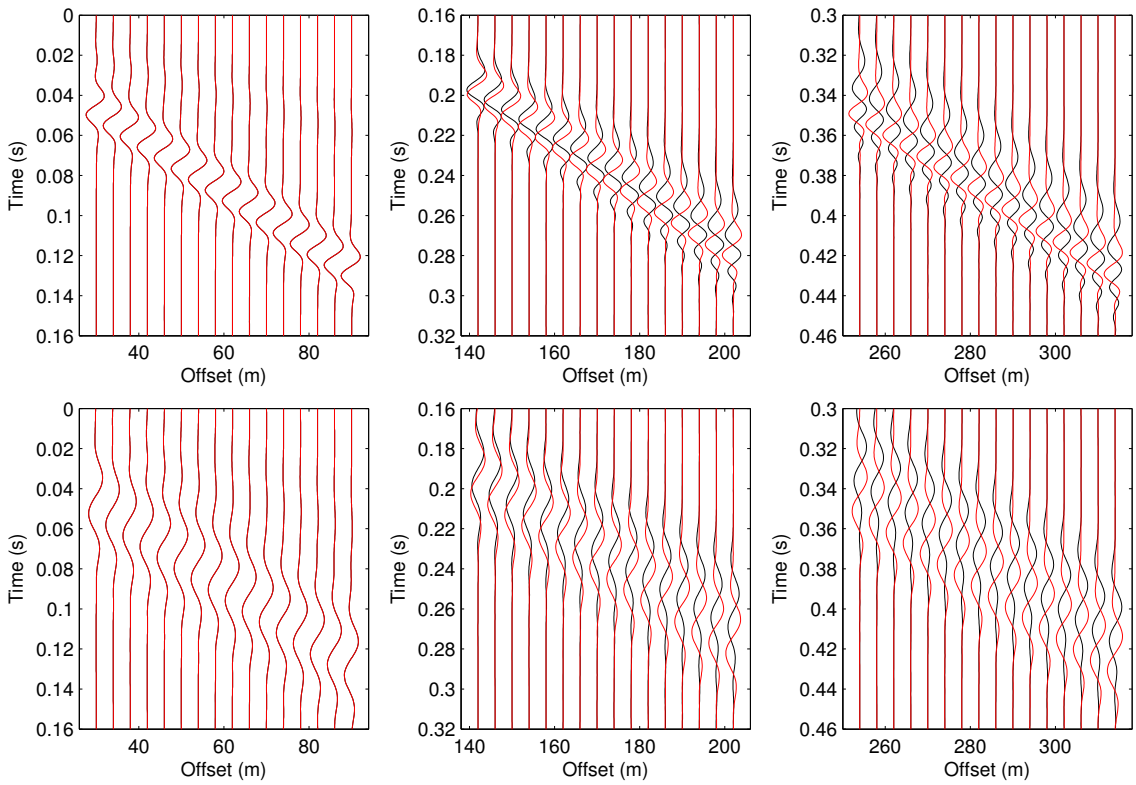


Figure 4.37: Observed (black) and modelled (red) seismic traces for three different source-receiver offset intervals indicated on the horizontal axes. Top: data computed with a 40 Hz Ricker wavelet. Bottom: band-passed filtered data (5-30 Hz).

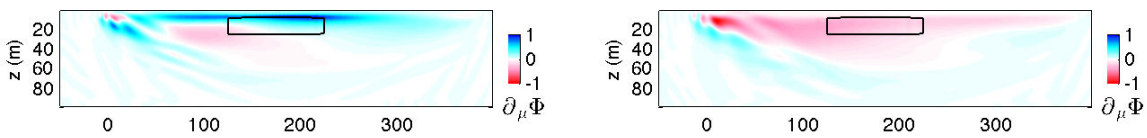


Figure 4.38: Gradients of FWI objective function with respect to μ. Left: data computed with a 40 Hz Ricker wavelet. Right: band-passed filtered data (5-30 Hz). The initial model is the same for both gradients.

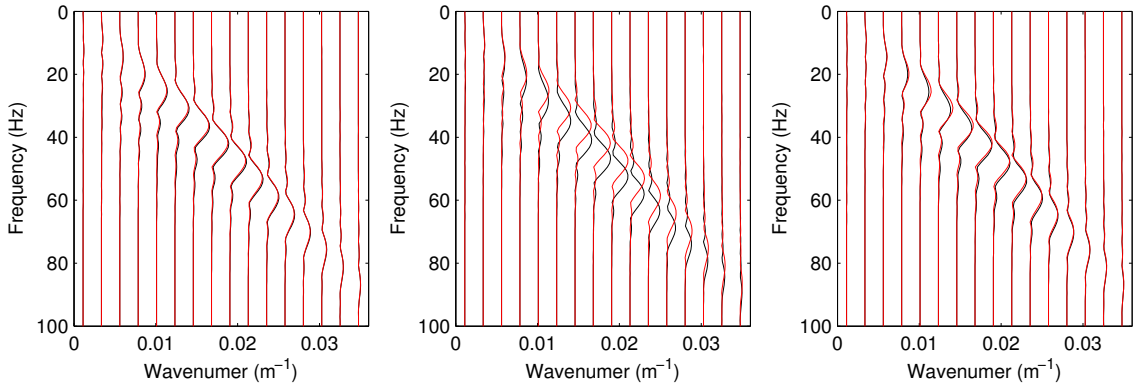


Figure 4.39: Absolute value of the 2D Fourier transform of observed (black) and modelled (red) data (Figure 4.37). Wiggle plotting is used for graphical purposes regarding amplitude match.

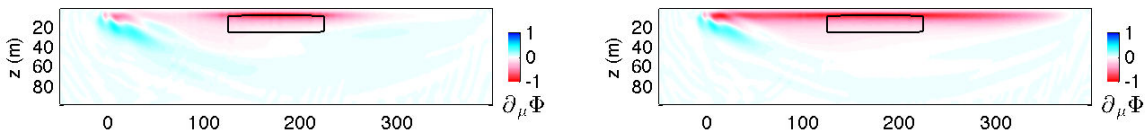


Figure 4.40: Gradients of w-AWI objective function with respect to μ using windows spaced 4 m apart. The initial model is the same for both gradients. Left: windows with l_w=128 m. Right: windows with l_w=256 m.

waveforms differ from the modelled waveforms for offsets greater than 125 m, location which propagated waves are physically modified by the velocity anomaly. Besides, most of the observed traces are not in phase with modelled traces as a consequence of the velocity anomaly. The computation of the classical FWI gradient is useful to show the effect of phase problems. In Figure 4.38 (left), the FWI gradient exhibits positive amplitude within the anomaly location. However, the value of μ is greater in the anomaly than in the homogenous model. The expression for the model update (equation 4.2) tells us that the gradient should have the opposite sign of the needed model update. Given that the necessary model update should be positive, the positive amplitude of the gradient suggest that the FWI gradient is in the wrong direction for this example.

We applied a band-pass filter (5-30 Hz) to observed and modelled data. In Figure 4.37 (bottom), the filtered data are plotted for the three offset ranges. Seismic traces in this frequency band are correctly matched because the wavelengths are larger than those of unfiltered data for similar propagation velocities. The FWI gradient computed with filtered data (shown in Figure 4.38 right) exhibits negative amplitude within the anomaly location which suggests that, contrary to the result of unfiltered data, the gradient is now in the good direction.

A different behaviour regarding cycle matching can be observed in the $f-k$ domain when w-AWI formulation is used. The $f-k$ transformation is applied separately to the seismic data in the three offset intervals shown in Figure 4.37. The absolute value of

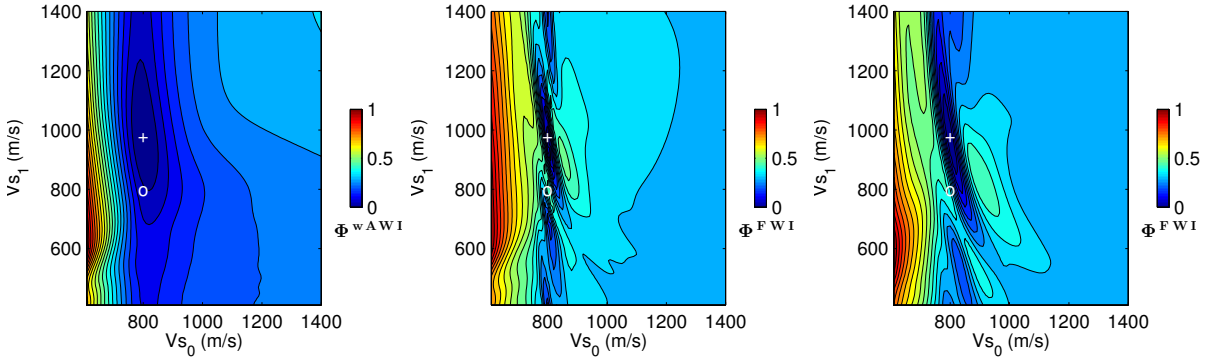


Figure 4.41: Objective function value for several initial models with different background and anomaly S-wave velocities. Left: w-AWI with $l_w = 128$ m. Middle: FWI. Right: FWI with band-pass (5-30 Hz) filtered data. The white cross indicates the exact model. The white circle indicates a model with homogeneous S-wave velocity of 800 m/s.

transformed data are plotted in Figure 4.39. The data are in correct cycle match for the three offset intervals (cycle is associated to amplitude lobes). The velocity anomaly generates a more visible difference between the cycles in the middle window (Figure 4.39, middle). This difference is expected because most of the receivers included in this window are located above the model velocity anomaly. Moreover, modelled data remains very close to observed data at the right of the anomaly (Figure 4.39, right) almost as they were at the left (Figure 4.39, left). The implicit phase removal in the misfit function explains the correct cycle match in the right-hand side window. In FWI, a low-pass frequency filtering is necessary to avoid cycle-skipping effects. In w-AWI, such a filtering is not necessary because there is no cycle skipping in the data. The w-AWI gradient using non-filtered data are in the good direction for convergence because negative amplitude is contained within the anomaly location as shown in Figure 4.40. However, the precision in the lateral localisation of velocity anomalies can be impacted by the window width in w-AWI. As an example, two w-AWI gradients are computed using windows spaced 4 m apart with widths $l_w = 128$ m and $l_w = 256$ m (Figure 4.40). The gradient computed with $l_w = 128$ m shows good lateral localisation of the anomaly (e.g. 100 m wide). But in the $l_w = 256$ m gradient the energy seems to be smeared out on both sides of the anomaly position. Nonetheless, the gradient seems to be in the good direction and properly centred in both w-AWI gradients.

The results of this test can be explained through an analysis of the shape of the objective functions of FWI and w-AWI. We compute the objective function values for synthetic data coming from models with different background and anomaly S-wave velocities. The considered background velocities ranges from 600 m/s to 1400 m/s whereas the anomaly velocity ranges from 400 m/s to 1400 m/s. In w-AWI computation, the windows used for data selection are spaced 4 m apart and are 128 m wide. In FWI, the objective function is computed for non-filtered data and band-pass filtered data ($5 - 30$ Hz). Results are shown in Figure 4.41.

For non-filtered data, this test shows that among the two considered objective functions w-AWI has the widest global minimum basin. When performing inversion with an initial model within the velocity values of Figure 4.41, the presence of secondary minima

in FWI may prevent convergence toward the global minimum, whereas in the case of w-AWI convergence may be achieved as only one minimum is observed. The objective functions computed for the data in the lower frequency band (5-30 Hz) show that the FWI global minimum basin can be broadened thus reducing constraints with respect to the initial model (Bunks et al., 1995). Local minima effects in low-frequency FWI are reduced as longer wavelengths are less sensitive to velocity variations in the model (Mulder and Plessix, 2008). The initial model for computation of FWI gradients is located at the position indicated by the white circle in the objective function image (Figure 4.41). The low-frequency FWI gradient (Figure 4.38, right) is in the good direction because the starting point is already in the FWI global minimum basin.

4.4 Inversion of synthetic data

In this section, we show inversion results using w-AWI and FWI for synthetic 2D velocity models. Please note that we compute the gradients with respect to the Lamé parameters and density and, to ease interpretation, we only show inversion results for Vp, Vs and ρ. In a first example, we test inversion considering that a-priori information is unknown. In this example, the initial velocity model consists of a linear velocity gradient. We use initial velocity models that are close to, or on the contrary, far from the exact velocity model to evaluate convergence results in FWI and w-AWI. The source wavelet is unknown in the second example. In a third example, we test multi-parameter inversion and show the result that best minimises the data misfit. Finally, we analyse results for frequency filtered data.

4.4.1 No a priori information and known source wavelet

We evaluate FWI and w-AWI for the reconstruction of a 2D elastic model (Figure 4.42). The exact model consists of a heterogenous elastic medium composed of several areas where the velocity is either constant or increasing with depth. This synthetic model is inspired on the ground model of Deidda and Balia (2001). The internal interfaces between velocity zones are parallel to the Cartesian axes or tilted. Both P-wave and S-wave velocities are, in general, increasing with depth with the exception of the layer located under the two high velocity zones. These high velocity zones have both parallel-to-Cartesian-axes and tilted interfaces. For the density, we consider two profiles: one 2D (third panel of Figure 4.42) and one homogeneous with $\rho = 1000$ kg/m^3 (bottom panel of Figure 4.42). The observed data consists of 20 common shot gathers modelled for sources that are evenly distributed between x=0 m and x=47 m as indicated by the red circles in Figure 4.42. The sources are all located 0.2 m below the surface. The source wavelet is a 40 Hz Ricker wavelet applied on the vertical component of velocity. The data are computed for 151 vertical velocity receivers spaced 0.4 m apart and located between x=-5 m and x=55 m at the source depth.

We perform three inversion tests (inversion 1, 2 and 3) using FWI and w-AWI. The source wavelet is considered known. The inversion reconstructs Lamé parameters λ and μ. Density is homogeneous and known in inversions 1 and 2 (and thus it is not updated). In inversion 3, the 2D density profile is included in the exact model (Figure 4.42) but it is

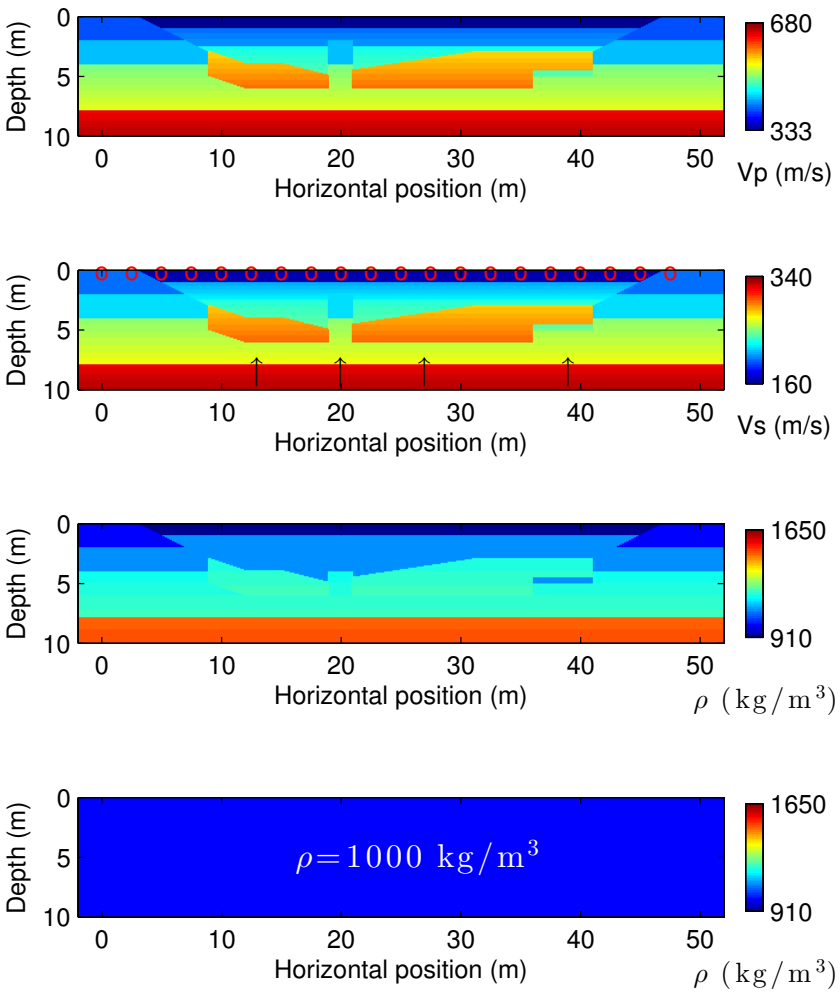


Figure 4.42: Exact velocity model for 2D inversion. Red circles indicate the position of 20 seismic sources used to generate the data. Density is either 2D (third panel) or homogenous (bottom panel) depending on the test.

assumed to be homogeneous in inversion (similarly to inversions 1 and 2). We employ the L-BFGS-B (Nocedal, 1980) and precondition the gradient as described in section 4.2.4. The boundaries of the L-BFGS-B algorithm are chosen to be 0.25 and 4 times the initial values. A small 2D gaussian window with 0.9 m radius ($0.9 = \alpha_x = \alpha_z$ in equation 4.26) is used to smooth the gradient at each iteration. As initial models, we consider linear increases of the properties with depth. Therefore, no a priori information is considered in the initial velocity model. The initial models are close to and far from the exact model in inversion 1 and 2, respectively. The initial model of inversion 3 is the same as for inversion 2 (far) but the density profile is considered unknown as mentioned previously. The objective behind inversions 1 and 2 is to evaluate the w-AWI result when the initial velocity model causes cycle skipping prevents convergence of FWI towards the global minimum. In inversion 3, we will mostly evaluate the accuracy of the w-AWI result when the density is variable and unknown.

The initial model and the result of inversion 1 with FWI and w-AWI are shown in Figure 4.43. The initial P-wave and S-wave velocity models are correctly updated by reconstruction of μ and λ both with FWI and w-AWI. The reconstructed Vp and Vs models contain the velocity features of the exact model and display a good representation of the tilted interfaces. 1D profiles of Vp and Vs show the high similarity of inversion results compared to the exact profiles (Figure 4.44), notably in the low-velocity zones under and between the two high-velocity zones. Observed and initial shot gathers for one source are shown in Figure 4.45. Although both shot gathers are similar in character regarding diving and Rayleigh waves, surface wave dispersion is stronger in observed data (Figure 4.45). The initial model for this inversion is close enough to the exact one and small phase problems do not prevent convergence towards the correct result. The shot gathers are completely reconstructed as shown in Figure 4.46. Moreover, the least-squares misfit is reduced to 0.24 % by FWI and to 0.11 % by w-AWI. The very good minimisation is confirmed by comparing seismic traces (Figure 4.46). FWI and w-AWI results match almost perfectly the observed data. This result is a reference for the following inversion tests.

Figure 4.47 shows the initial model together with the FWI and w-AWI results for inversion 2. The utilisation of an initial model far from the exact one results in non-convergence with FWI and convergence with w-AWI. The model retrieved by FWI is mostly updated at the very near surface but local minima yield inaccurate results in deeper parts. The divergent character of the FWI result can be observed in 1D profiles (Figure 4.48). Conversely, good profile similarity is obtained by w-AWI (Figure 4.48). These results are explained by the marked difference between initial and observed shot gathers as it can be observed in Figure 4.49. Indeed, observed and initial shot gathers differ both in diving and surface waves. This causes a strong cycle-skipping effect in FWI preventing convergence. A secondary minimum is the solution retrieved by the local search. By contrast, w-AWI converges because the global minimum basin of its objective function is wider as it depends on different input data ensuring correct convergence. In this case, in spite of the presence of cycle-skipping in the $t-x$ domain (section 4.3.1). The error is reduced to 23.98 % by FWI and to 0.11 % by w-AWI. As shown in Figure 4.50, the evolution of w-AWI objective function with iterations is similar to inversion 1 while less significant minimisation is observed in FWI. The comparison of traces allows us to observe that FWI converged to a secondary minimum (Figure 4.50) as the final data

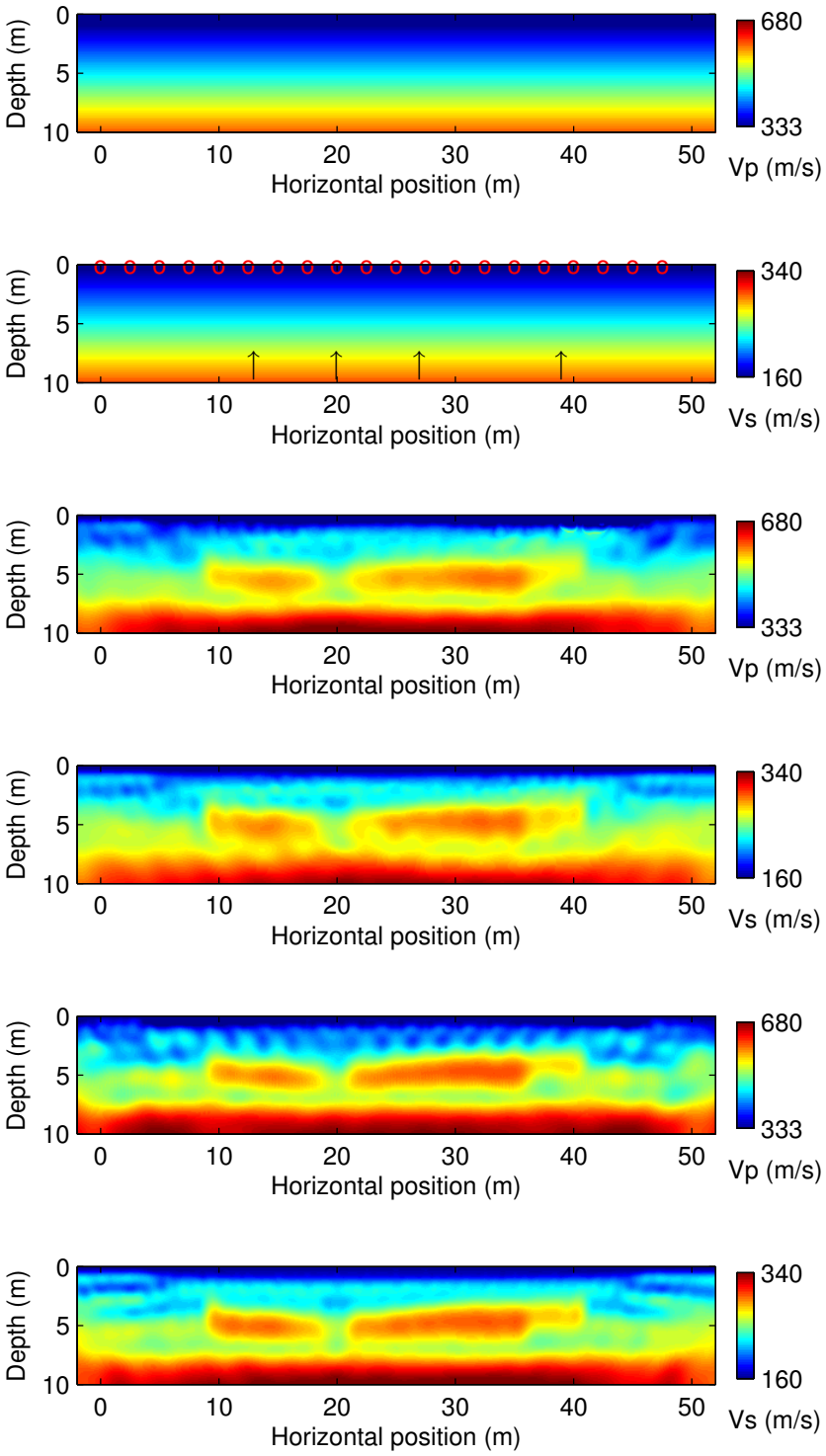


Figure 4.43: Inversion 1. Initial model (first and second panels), FWI result (third and fourth panels) and w-AWI result (fifth and last panels). Density is considered known (homogeneous) and is not updated during inversion.

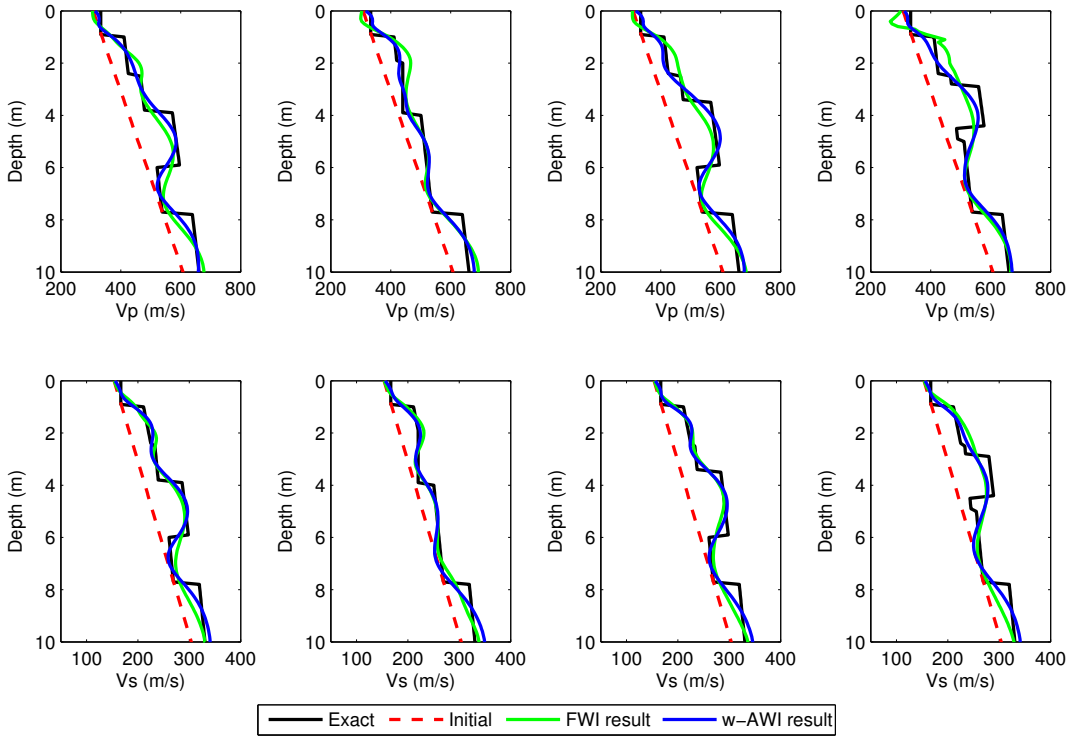


Figure 4.44: Inversion 1. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m.

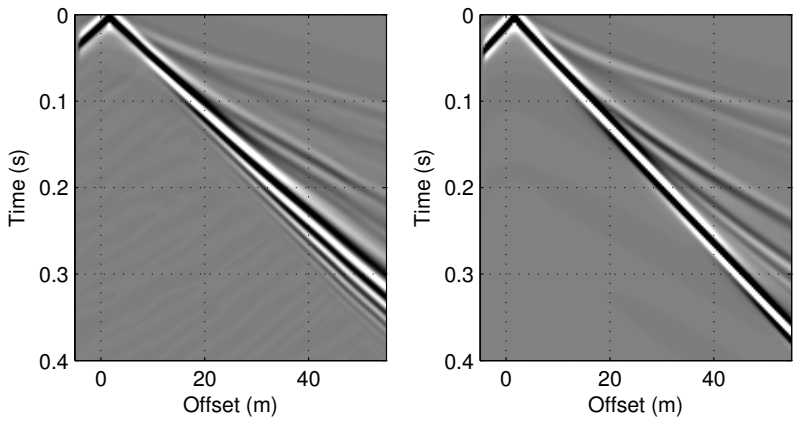


Figure 4.45: Example of observed (left) and initial (right) common shot gathers.

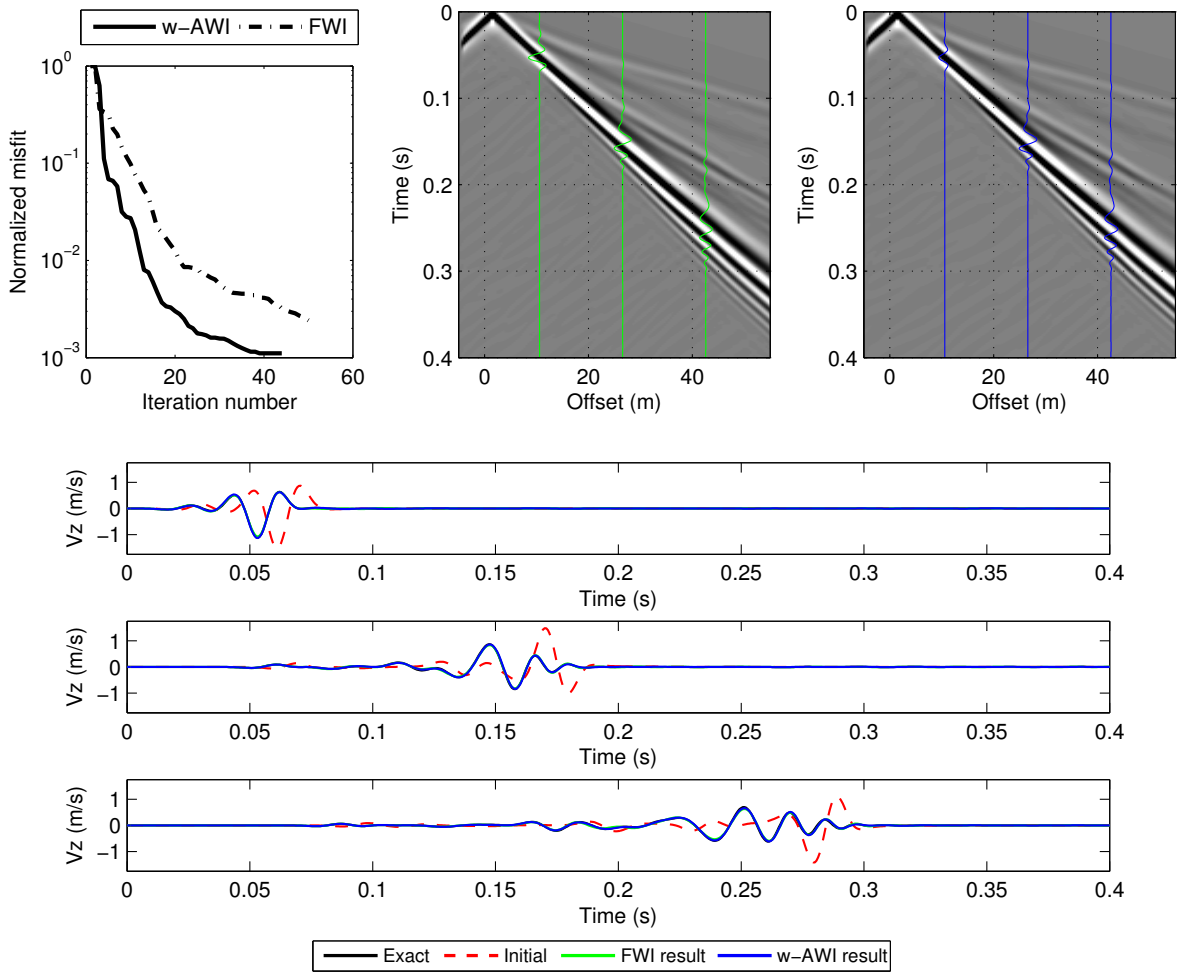


Figure 4.46: Inversion 1. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

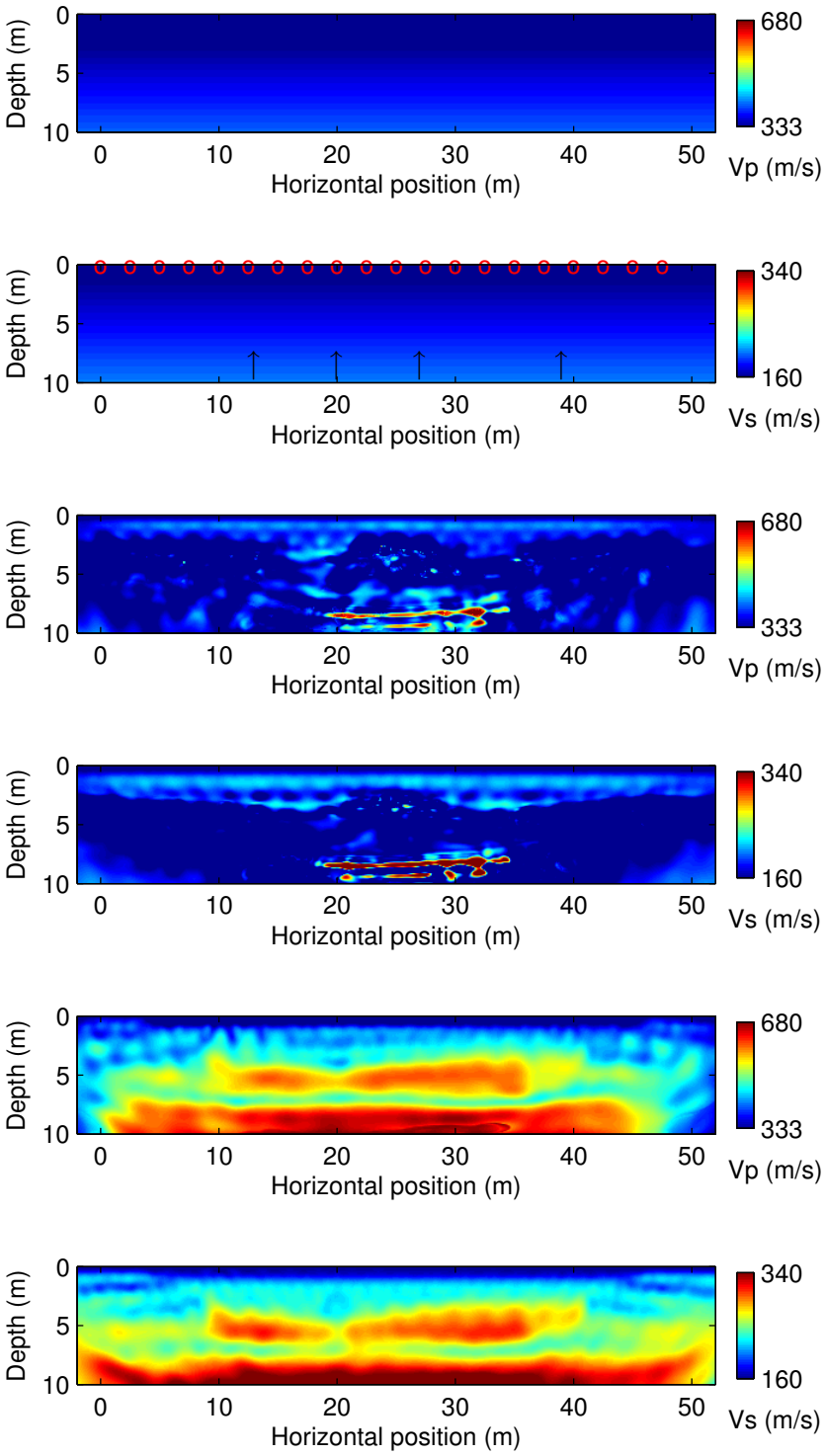


Figure 4.47: Inversion 2. Initial model (first and second panels), FWI result (third and fourth panels) and w-AWI result (fifth and last panels). Density is considered known (homogeneous) and is not updated during inversion.

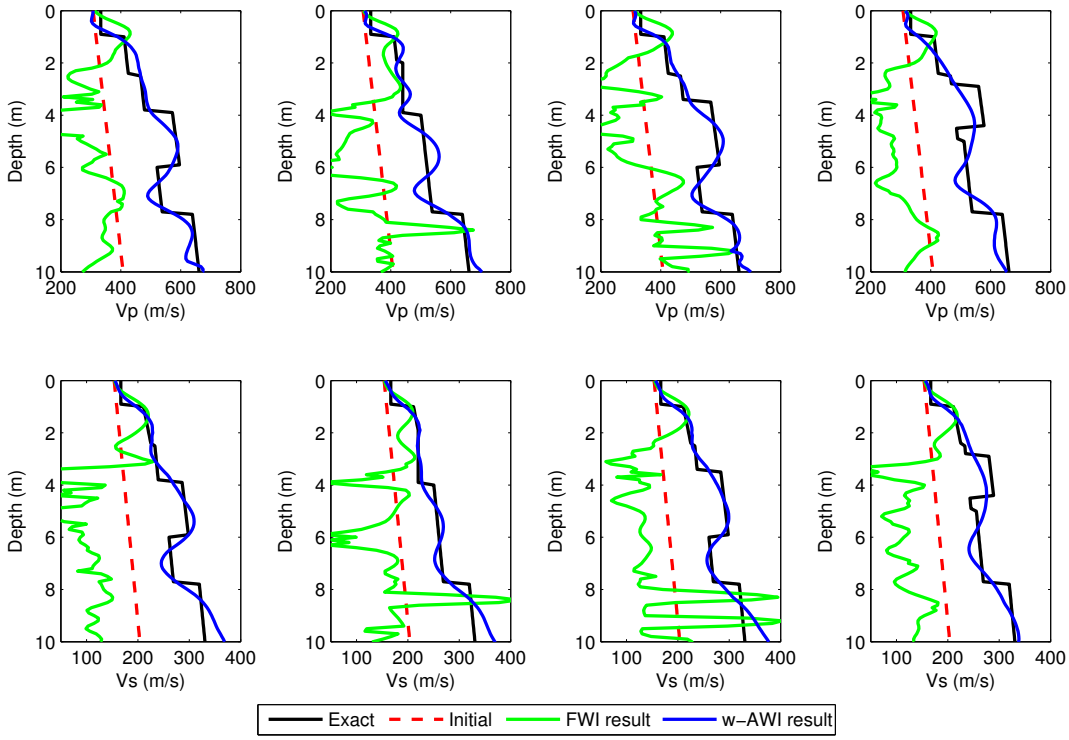


Figure 4.48: Inversion 2. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m.

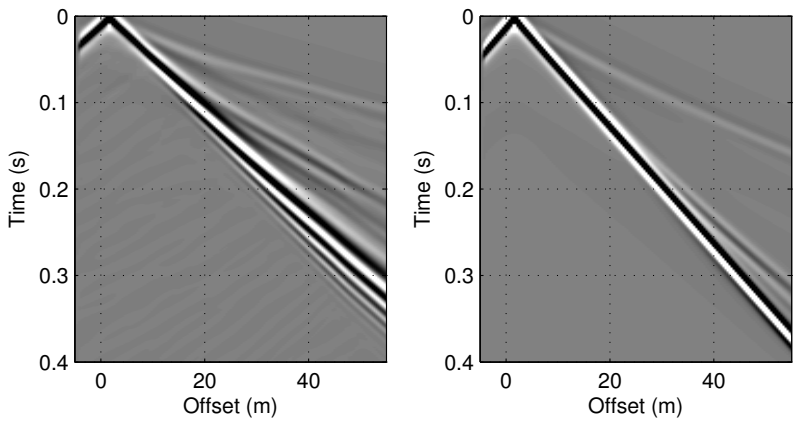


Figure 4.49: Inversion 2. Observed (left) and initial (right) common shot gathers. Only one shot gather is shown.

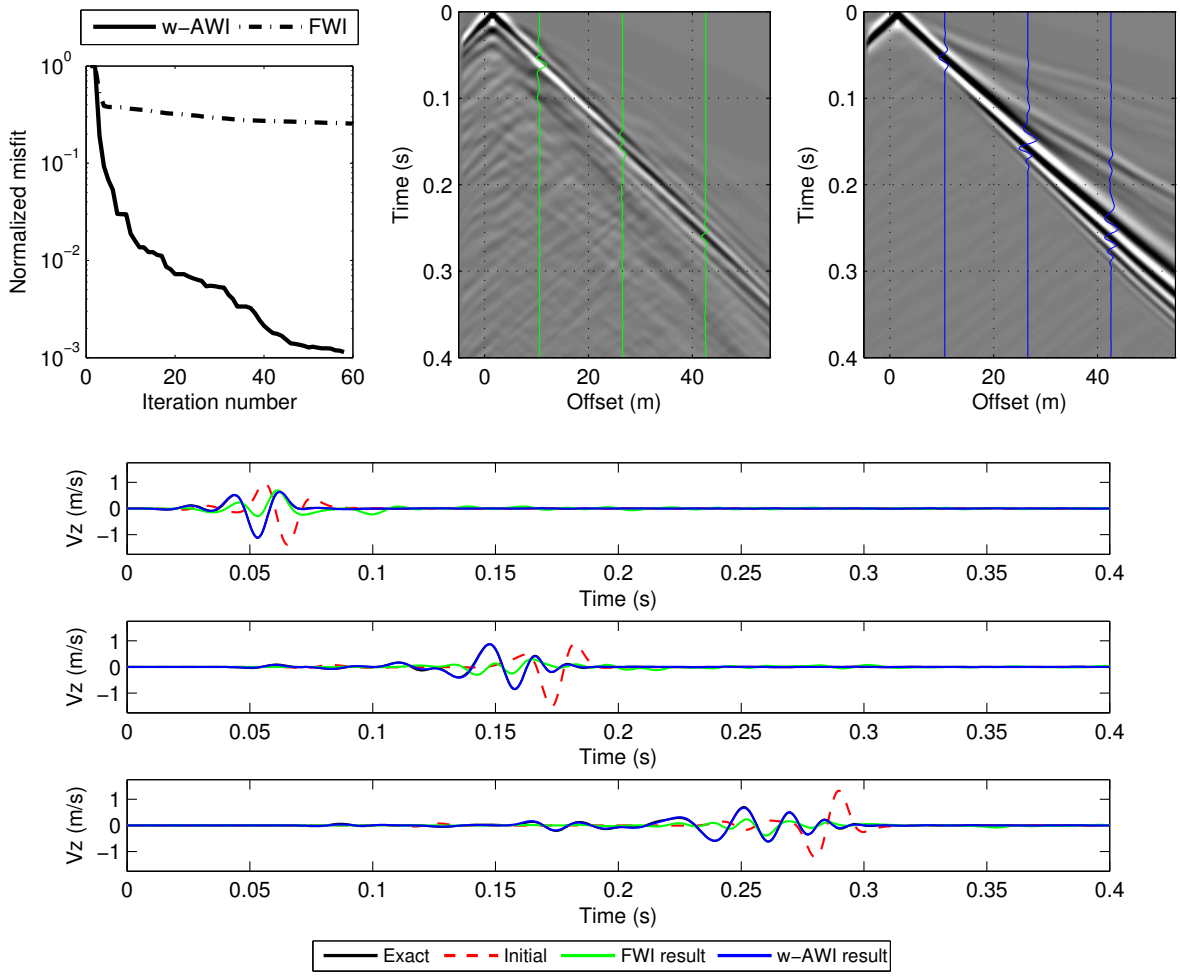


Figure 4.50: Inversion 2. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

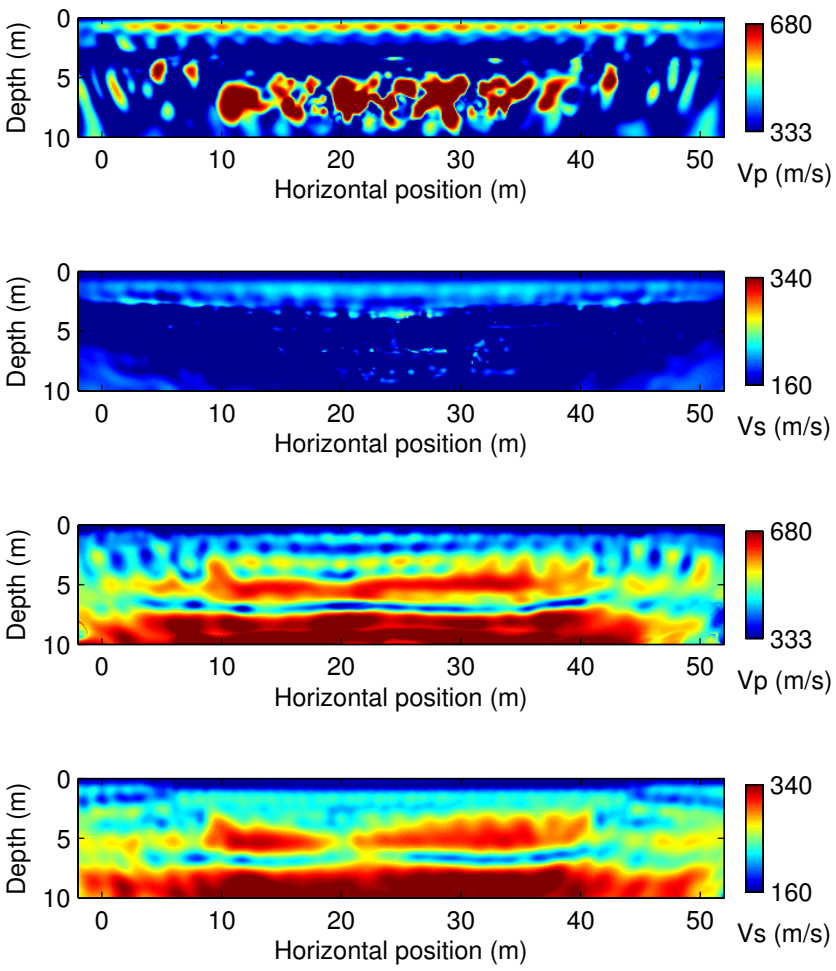


Figure 4.51: Inversion 3. FWI result (first and second panels) and w-AWI result (third and last panels). The exact density model is 2D (Figure 4.42). The initial density model is homogeneous and is not updated during inversion.

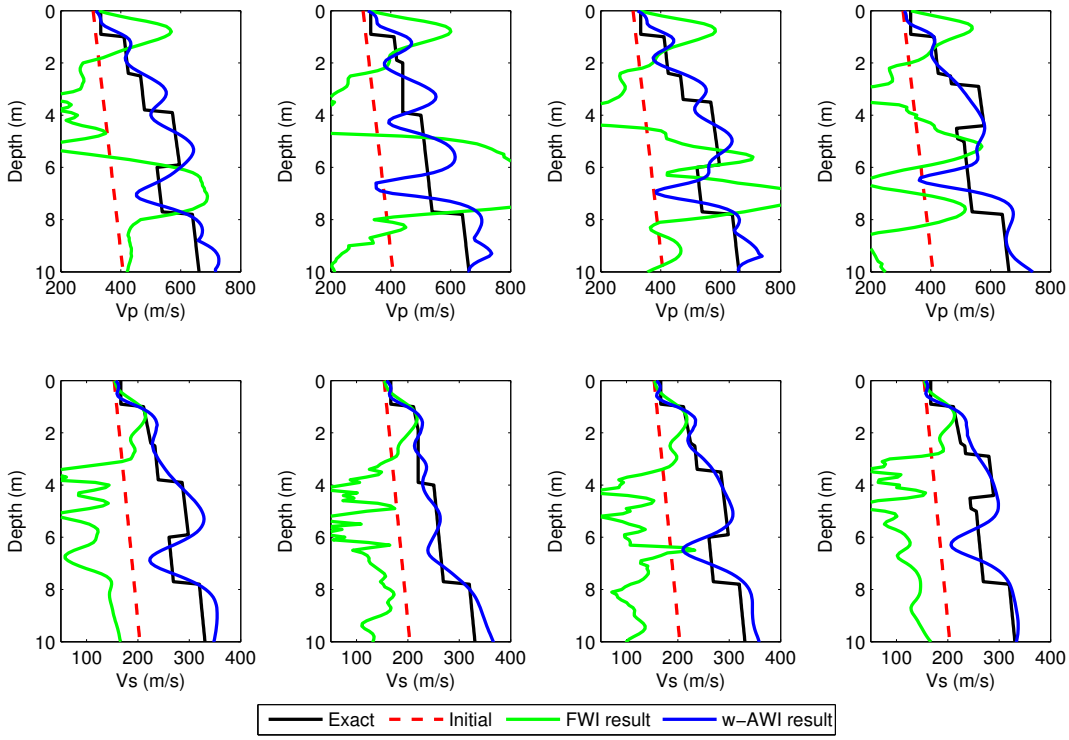


Figure 4.52: Inversion 3. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m.

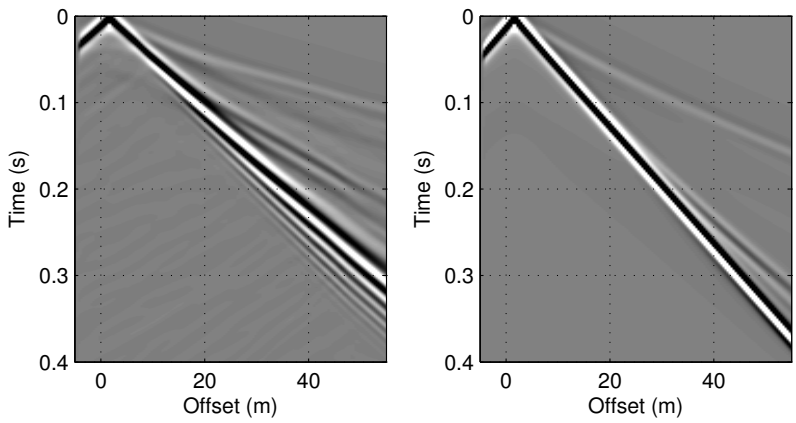


Figure 4.53: Inversion 3. Observed (left) and initial (right) common shot gathers. Only one shot gather is shown.

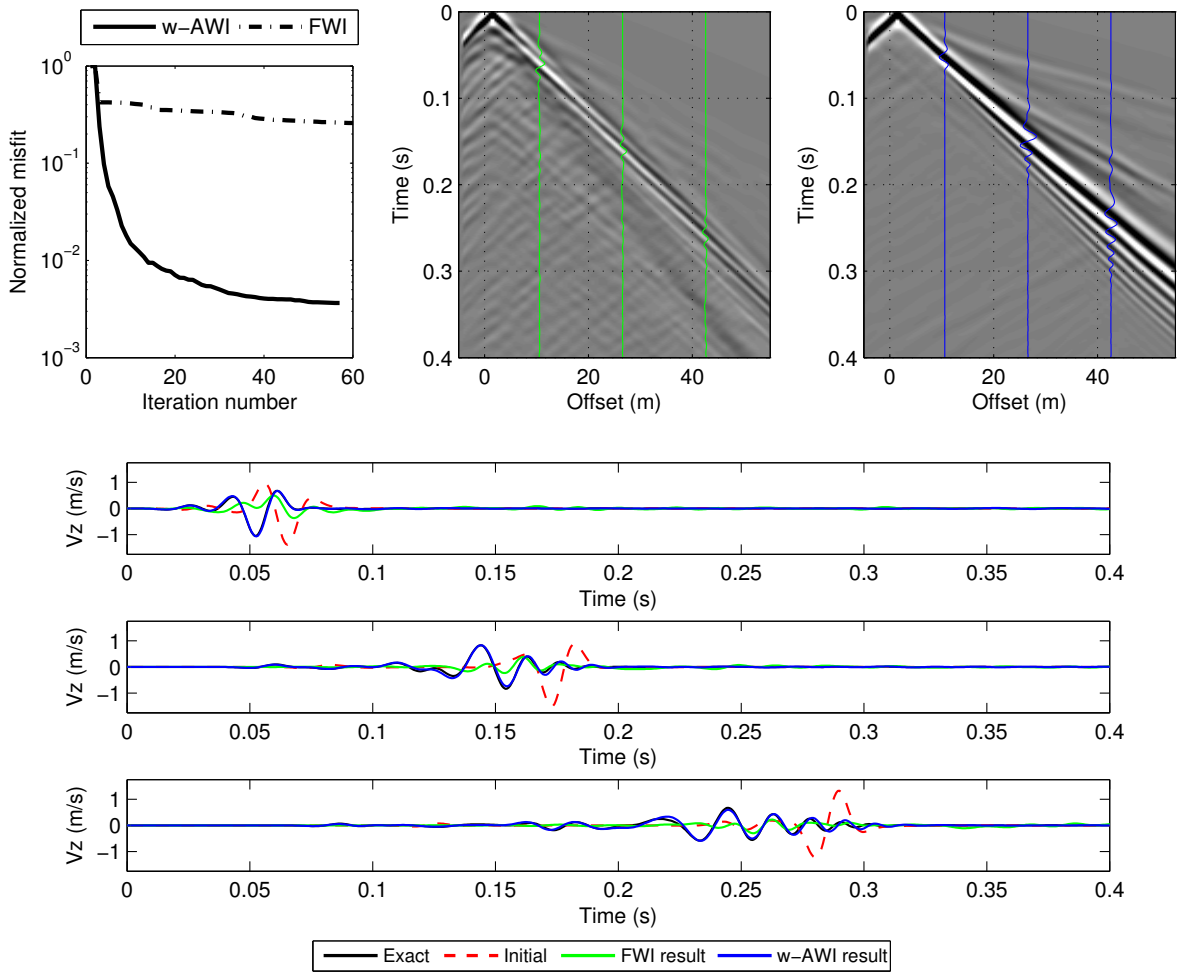


Figure 4.54: Inversion 3. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

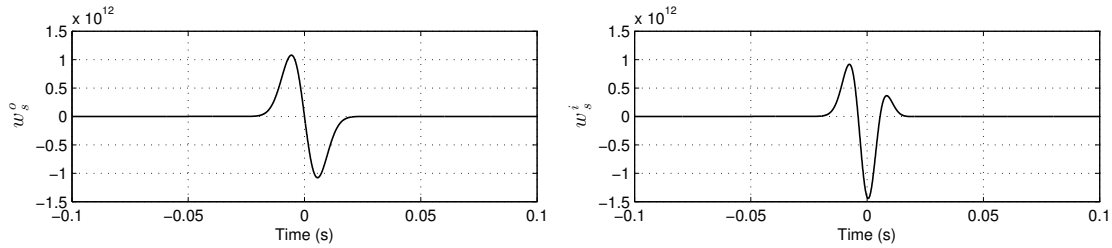


Figure 4.55: Exact (left) and initial (right) source wavelets.

seem to be more or less in phase with observed data but their amplitudes are smaller. FWI tried to explain the surface waves of observed data with a model whose near surface velocity is close to the exact one but deeper velocities are erroneous. On the contrary, with w-AWI, the data show almost no differences with observed data.

For inversion 3, the initial velocity models are the same as for inversion 2 (Figure 4.47, first and second panels), but in this case the exact density model is not homogeneous (Figure 4.42, third panel). The result of FWI diverges from the exact model because the initial model was already far in addition to the unknown density model (Figure 4.51, first and second panels). The result of w-AWI converges to the exact model especially in the case of S-wave velocity (Figure 4.51, third and last panels). The 1D profiles (Figure 4.52) show that FWI updates the very near surface and w-AWI reconstructs the whole model. However, the result of w-AWI is less accurate than in inversions 1 and 2 because the method has adjusted the velocity parameters to compensate for the strong density difference. Among the P-wave and S-wave velocities, the first one is the most affected by compensation effects as indicated by the stronger velocity oscillations (Figure 4.52, first row). The analysis presented by Tarantola (1986) and Forgues and Lambaré (1997) explains that density perturbations may cause updates in λ (and vice-versa) as their diffraction patterns are similar for P-P diffractions at short offsets. We have not analysed diffraction patterns in the case of surface waves. But our results seem to agree with body wave analysis. Finally, the error is minimised to 24.38 % and to 0.37 % of its initial value by FWI and w-AWI respectively. In data reconstruction, FWI retrieved data in phase with observed data but with incorrect amplitude, while w-AWI reconstructs the observed data although showing a slightly reduced accuracy than for inversion 1 and 2 due to the difference in density (Figure 4.54).

4.4.2 Inversion with unknown source wavelet

In this section, we evaluate FWI and w-AWI when the source wavelet is unknown. We propose three different tests (called test 1, 2 and 3). For the three tests, the observed data are computed with the model shown in Figure 4.42 including the homogeneous density model. The exact source wavelet is a 40 Hz Ricker wavelet (Figure 4.55, left). In test 1, we apply inversion considering the source wavelet as known to have a reference result. Test 2 is similar but the source wavelet is considered as unknown. The initial model for these two tests is a version of the exact model smoothed with an elliptical 2D gaussian window with vertical axis of 8 m (α_z in equation 4.26) and horizontal axis of 16 m (α_x in equation 4.26). In test 3, the source wavelet is considered as unknown and no a-priori

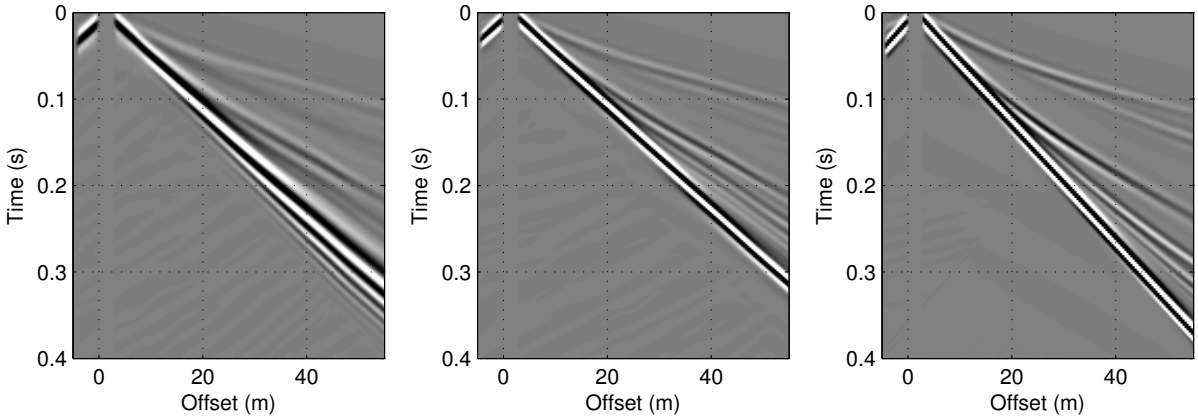


Figure 4.56: Observed (left) and initial common shot gathers (centre for test 1 and 2 and right for test 3). Only shot gathers due to one of the 20 considered sources are displayed.

information is considered in the initial velocity model. The initial velocity model in test 3 is the same as in inversion 1 of Section 4.4.1 (close velocity gradient). For tests 2 and 3 (unknown source wavelet), the initial data are computed with a source wavelet defined as

$$\varphi_z(x,t) = \delta(x - x_s)(0.3R^0(t, 40Hz) + 0.003R^1(t, 60Hz)), \tag{4.27}$$

where x_s defines the horizontal position of the source and $R^n(t, f)$ is the nth derivative of a Ricker wavelet with central frequency f. This initial wavelet is shown in Figure 4.55 (right). FWI and w-AWI are used to reconstruct the Lamé parameters only. Density is considered known and is not updated in inversion. The near offset traces are masked in inversion because in real data inversion the near offset geophones are commonly saturated in energy, which often make them non-usable. Observed data are represented by the shot gather in Figure 4.56 (for a source located at a horizontal position of 0 m). The corresponding initial shot gathers computed with the wavelet defined in equation 4.27 are shown as well (Figure 4.56, centre for the smooth initial model and right for simple-gradient initial model).

The inversion results for test 1 (known source wavelet) are shown in Figure 4.57. Four 1D profiles are shown in Figure 4.58. The overall results are excellent, as expected, because the initial velocity model contains good a priori information. The final misfit is smaller in w-AWI than in FWI (Figure 4.59). The data reconstructed by these two methods coincide almost perfectly with the observed data as shown in the traces comparison of Figure 4.59. The error between exact and final velocity model is slightly smaller in w-AWI (we present this result later on using a comparative table).

For test 2, the source wavelet is estimated at the begin of each L-BFGS-B iteration using the approach of (Pratt, 1999) in the case of FWI and the formulation presented in Section 4.2.3 in the case of w-AWI. The resulting source wavelets after inversion are shown in Figure 4.60 (left for FWI and right for w-AWI). Some oscillations that are not present in the exact wavelet are created in the FWI wavelet probably as a compensation for the lacking dispersion in the reconstructed data. The shape of the w-AWI source wavelet is not similar to the shape of the exact wavelet because the phase of the data are not used in the w-AWI inversion. The 2D velocity models retrieved with the two inversion

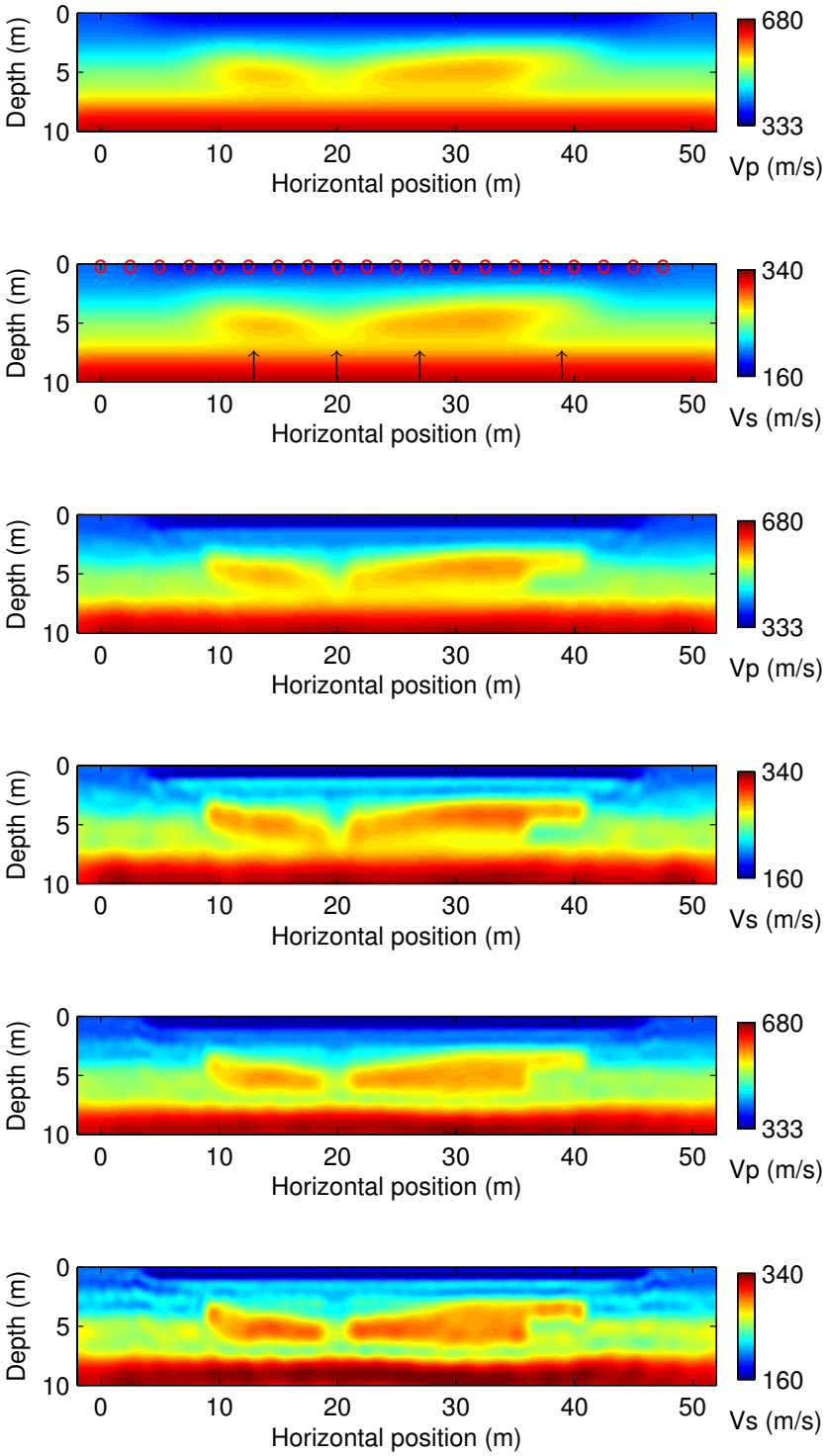


Figure 4.57: Test 1. Initial model (first and second panels), FWI (third and fourth panels) and w-AWI results (fifth and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is known and is not updated during inversion.

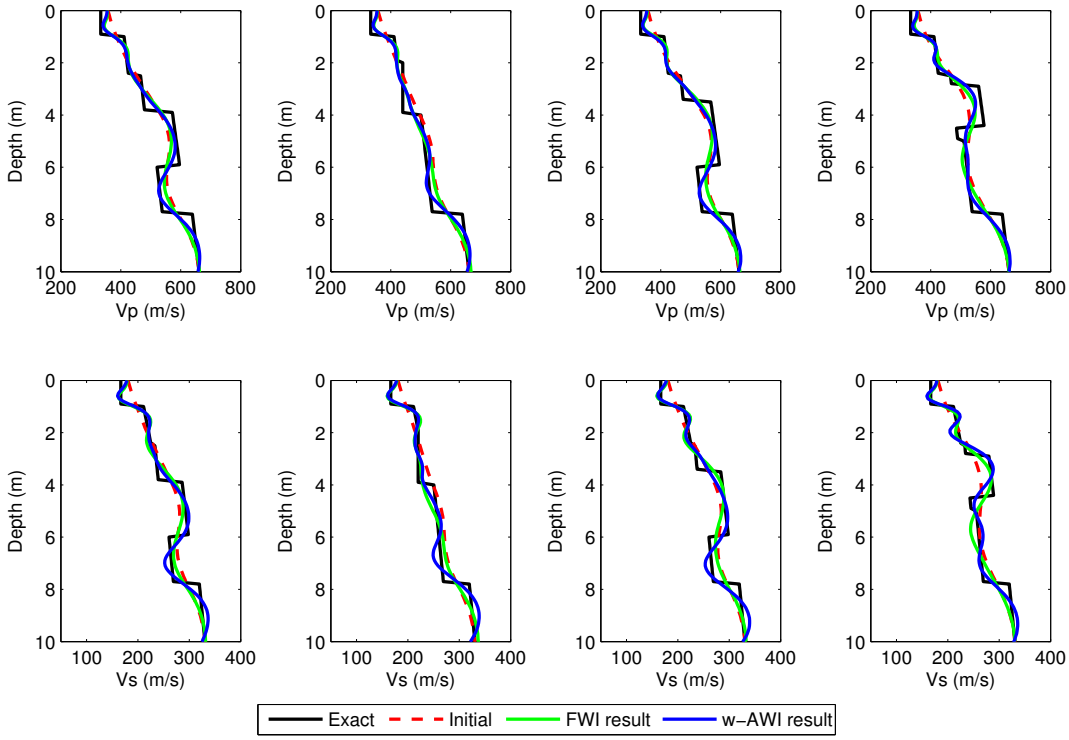


Figure 4.58: Test 1. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m.

procedures correctly correspond to the exact velocity models as shown in Figure 4.61. The 1D profiles shown in Figure 4.62 display an abnormal high velocity in the Vs result of FWI. This kind of fictitious layers are usually created in the presence of deeper velocity contrasts as is the case with the interface located at 4 m depth approximately. The data misfit is better minimised in w-AWI (on the order of 10^{-3}) than in FWI (on the order of 10^{-2}) as shown at Figure 4.63. However, the comparison of seismic traces shown on the bottom of Figure 4.63 reveals that in w-AWI the reconstructed data are not in phase with the exact data. This is a result of ignoring the phase of data in inversion. As a solution for shot-gathers reconstruction, a few iterations of FWI after w-AWI should suffice to give in-phase data because w-AWI has already converged to the exact velocity model. The same solution can be applied to reconstruct the phase of the source wavelet.

For test 3, we respectively obtain a convergent and a divergent result (from the exact model) with w-AWI and FWI. The final source wavelets are shown in Figure 4.64. The waveform of the FWI estimated source wavelet differs with the exact source wavelet (Figure 4.64, left). The w-AWI source wavelet is similar to the one estimated in the test 2 (Figure 4.64, right). The final 2D velocity profiles are shown in Figure 4.65. In these results, the FWI velocity features do not match with the exact model ones. The model reconstructed with w-AWI preserves similarities with the exact velocity model, however results are less good than for the first and second tests. 1D profiles are shown in Figure 4.66. FWI results show large velocity model updates but less divergent results than in Section 4.4.1. In the case of w-AWI, the velocity profiles show oscillating values

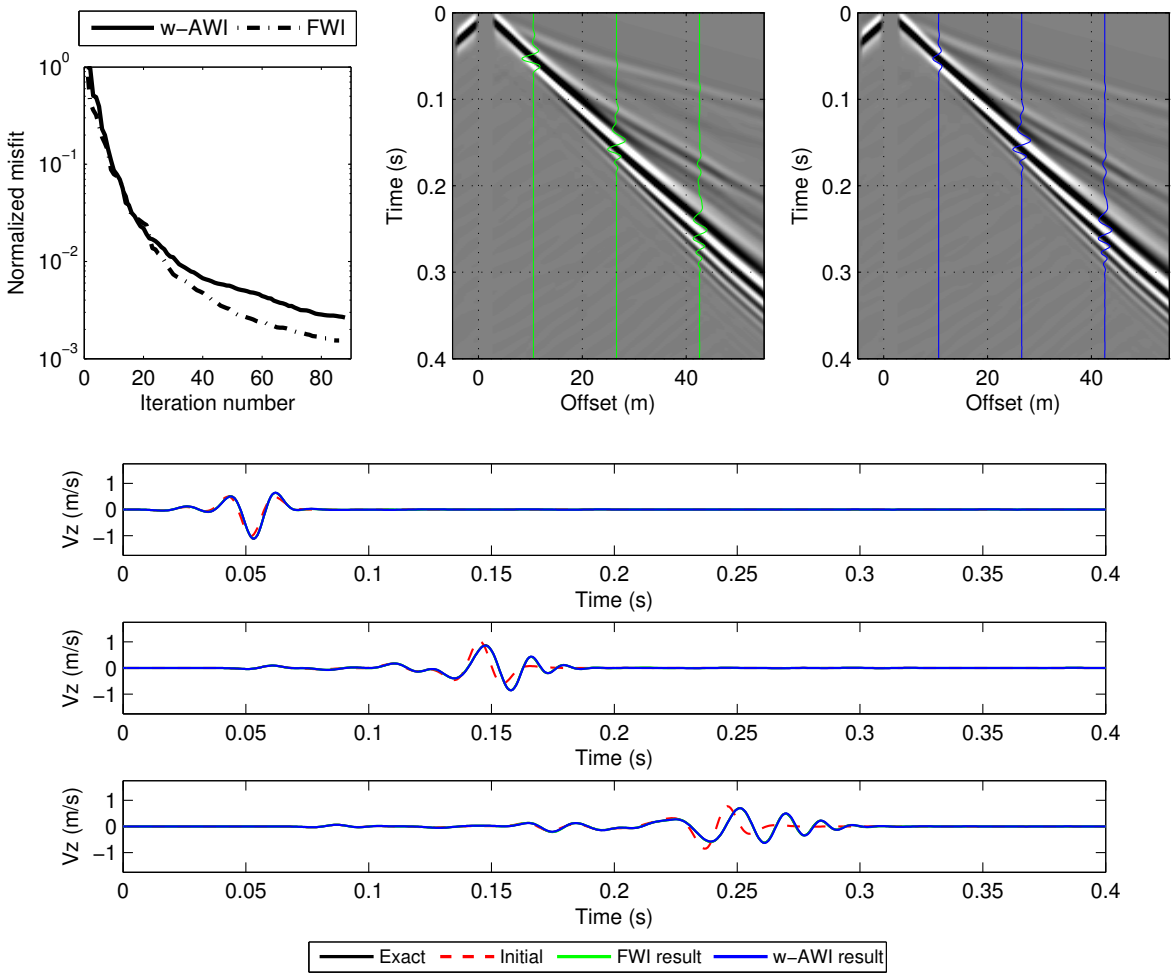


Figure 4.59: Test 1. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

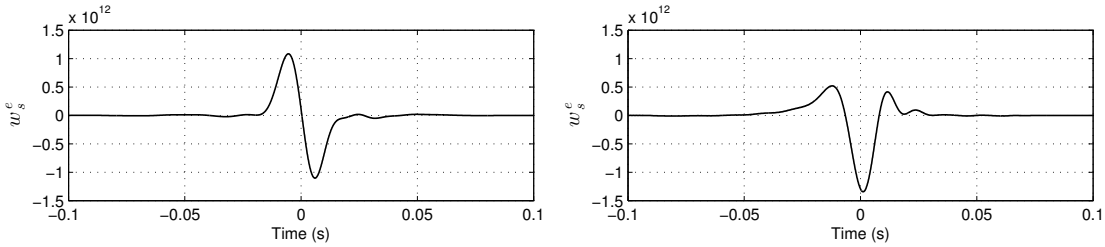


Figure 4.60: Test 2. Final source wavelets estimated with FWI (left) and w-AWI (right). The estimated wavelet is more similar to the exact one (Figure 4.55 left) in the case of FWI because, differently from w-AWI, the phase is taken into account.

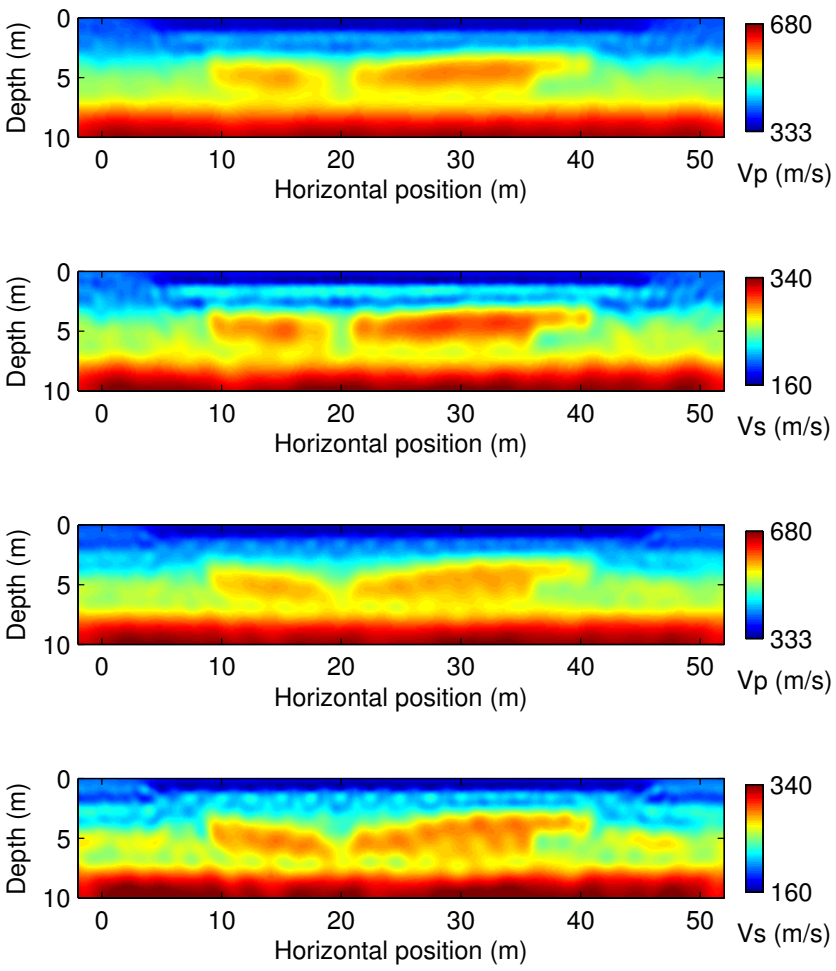


Figure 4.61: Test 2. FWI (first and second panels) and w-AWI results (third and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is unknown and updated during inversion.

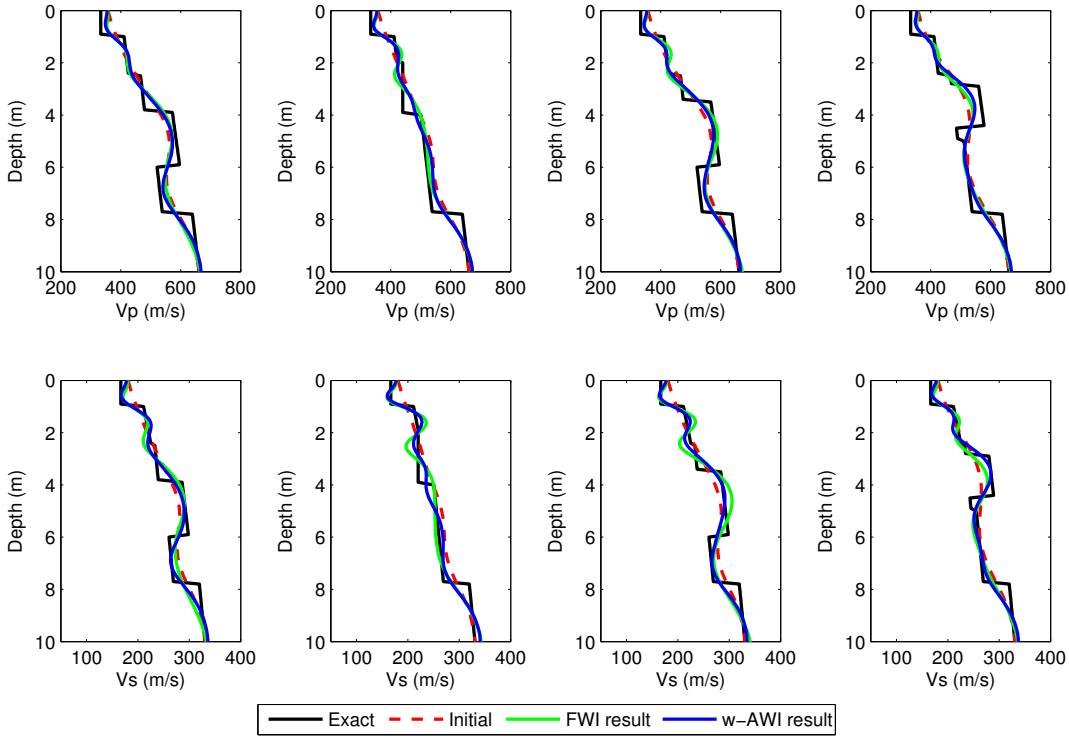


Figure 4.62: Test 2. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m.

that are not present in the exact profiles. The data misfit is minimised with w-AWI to a value in the order of 10^{-3} like in test 2 (Figure 4.67). The data reconstructed with w-AWI (Figure 4.67) shows again a mismatch in phase that can be fixed with a few FWI iterations.

We have computed the relative error between initial and exact velocity models as well as between reconstructed and exact velocity models. These values are shown in Table 4.1. Table 4.1 shows that the error is smaller in w-AWI than in FWI for all the tests. Results are better when the wavelet is known. When the wavelet is unknown and estimated through iterations, the result is better when using the smooth initial model, i.e., when considering a-priori information.

4.4.3 Multi-parameter inversion

In this section, we employ FWI and w-AWI to reconstruct multiple model parameters (λ, μ and ρ) considering no a-priori information (except for the source wavelet which is considered known). In our computations, we tested different inversion strategies divided in "steps". In each step, the inversion codes are used to reconstruct one, two or the three model parameters. Among the considered inversion strategies, we have only obtained one result in which the data are well explained. In this test, the final relative error between velocity models is smaller than the starting one. However, this relative error is not strictly minimised in all the steps Indeed, the error sometimes increases while the data misfit is

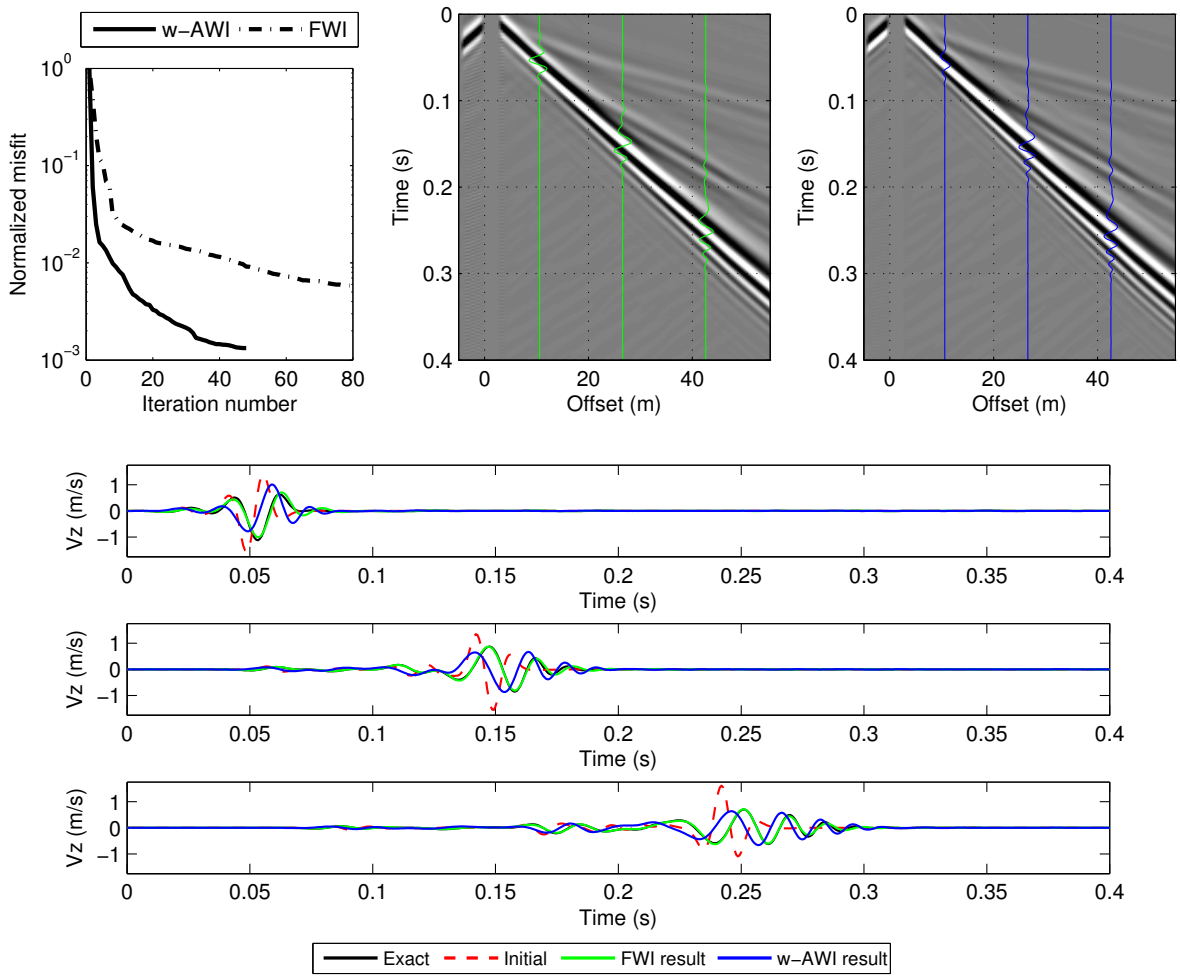


Figure 4.63: Test 2. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

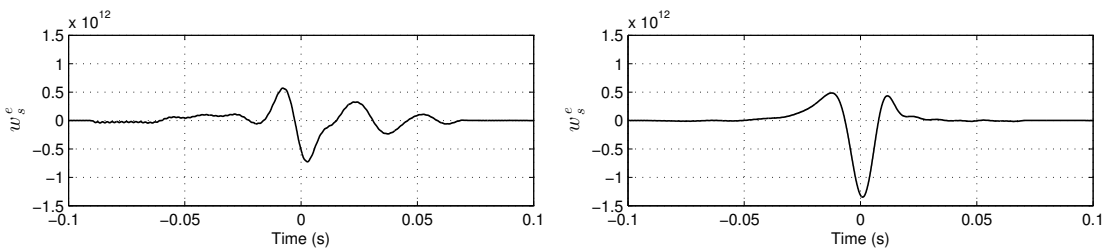


Figure 4.64: Test 3. Final source wavelets estimated with FWI (left) and w-AWI (right). The wavelet estimated with FWI shows convergence problems as it do not coincide to the exact one (Figure 4.55 left). The wavelet estimated with w-AWI allows convergence even if its shape is not similar to the exact one, which is a result of neglecting the phase.

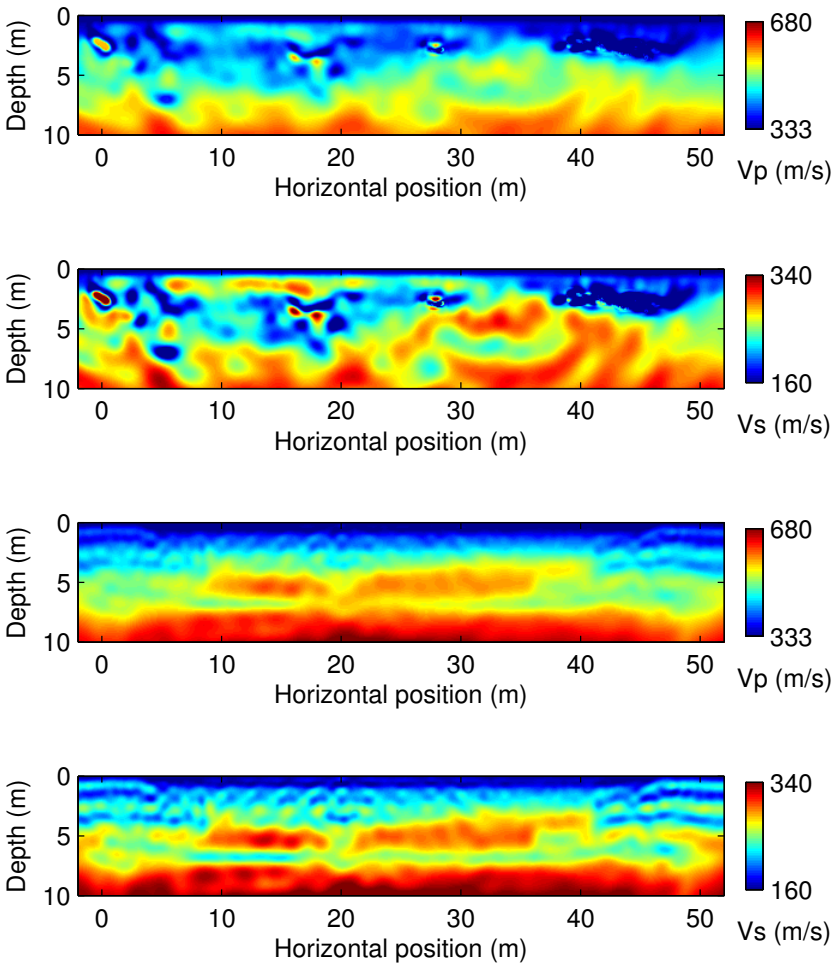


Figure 4.65: Test 3. FWI (first and second panels) and w-AWI results (third and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is unknown and updated during inversion.

Test Wavelet and velocity model →. Parameter ↓	1 Known wavelet, smooth model	2 Unknown wavelet, smooth model	3 Unknown wavelet, close velocity gradient
Vp (initial)	4.12	4.12	11.1
Vs (initial)	4.25	4.25	11.41
Vp (FWI)	3.24	3.97	9.49
Vs (FWI)	3.13	4.15	11.91
Vp (w-AWI)	2.67	3.24	4.2
Vs (w-AWI)	2.78	3.31	4.7

Table 4.1: Relative error in % of reconstructed models using exact or estimated wavelet.

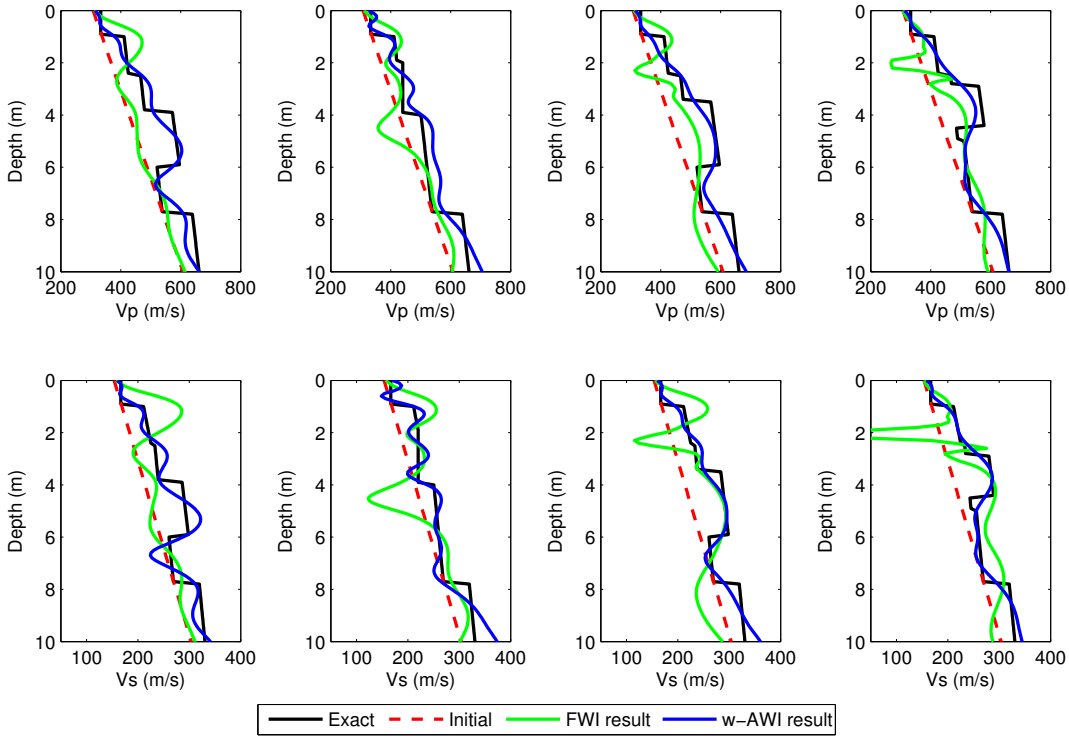


Figure 4.66: Test 3. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m.

minimised. We have obtained this result using w-AWI while FWI results diverged for all the considered strategies (probably due to the initial model). The inversion parameters and the best inversion strategy (among the considered ones) are described in the next paragraph.

The observed data are obtained with the model of Figure 4.42 including the 2D density profile. As initial velocity models, we have considered the same as for inversion 1 of section 4.4.1. The initial density model is a simple linear gradient (shown later in Figure 4.71). We apply inversion in three steps (1) simultaneous λ and μ reconstruction, (2) ρ inversion and (3) simultaneous λ and μ reconstruction. Step 1 is chosen based on the result of inversion 3 (section 4.4.1) where the compensation for density was small. In step 2, the density is updated supposing an already good velocity model. Finally, step 3 is necessary to tune the λ and μ distributions with an improved density model (from step 2). Step 1 results are shown in Figure 4.68 (2D profiles) and Figure 4.69 (1D profiles). As in inversion 1 (section 4.4.1), results are close to the exact model and the compensation for density is not strong because the initial model is close to the exact one. The misfit is reduced to 1.06 % after 22 L-BFGS-B iterations (Figure 4.70). The data are in general well reconstructed, except for a small phase mismatch (Figure 4.70). In step 2, only the density is updated. The result shows similarities to the exact model in the shallow part but some oscillations are created in the deeper part (Figures 4.71 and 4.71). However, the final data show a good match with almost imperceptible misfit with respect to the observed data (Figure 4.72. The misfit minimisation with respect to initial data (final

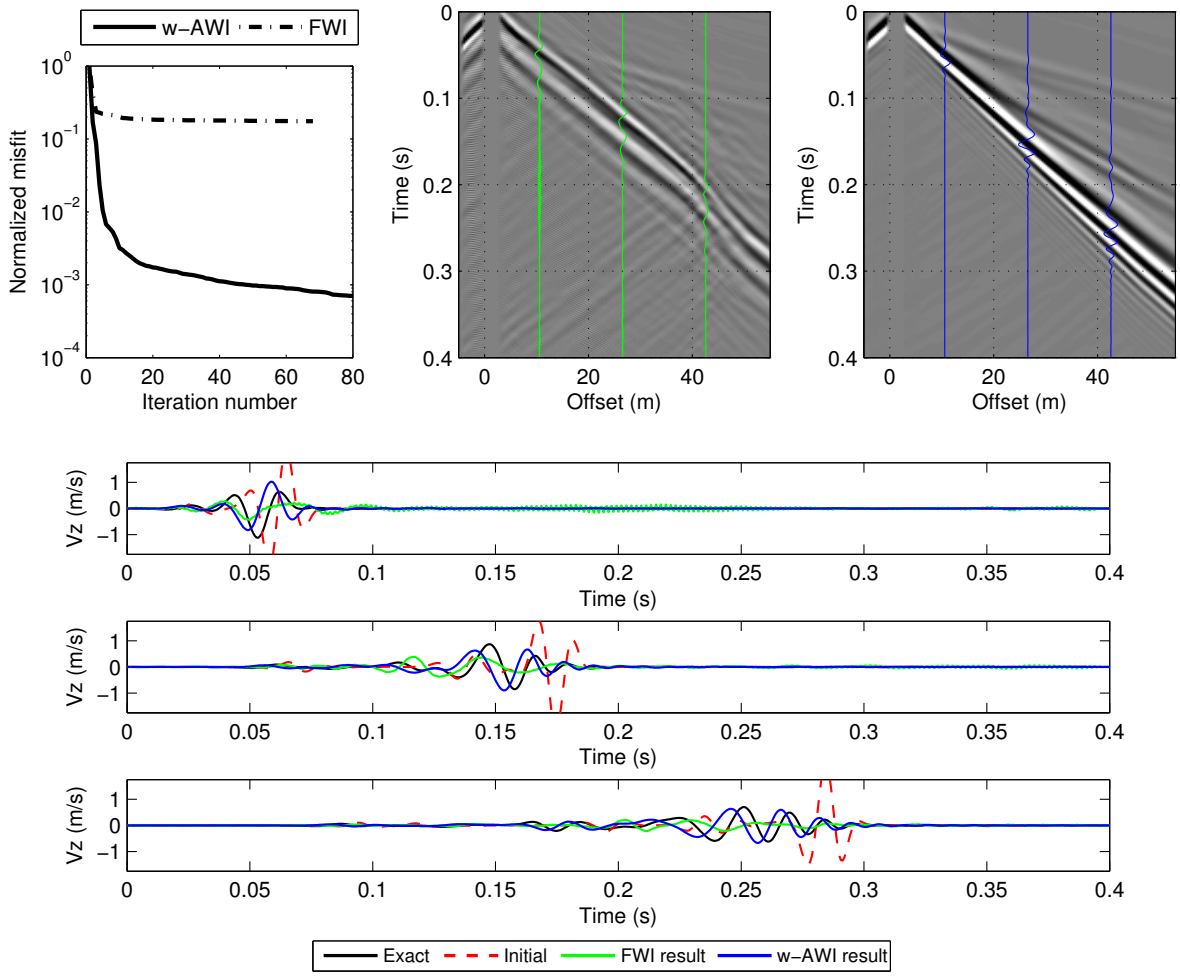


Figure 4.67: Test 3. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

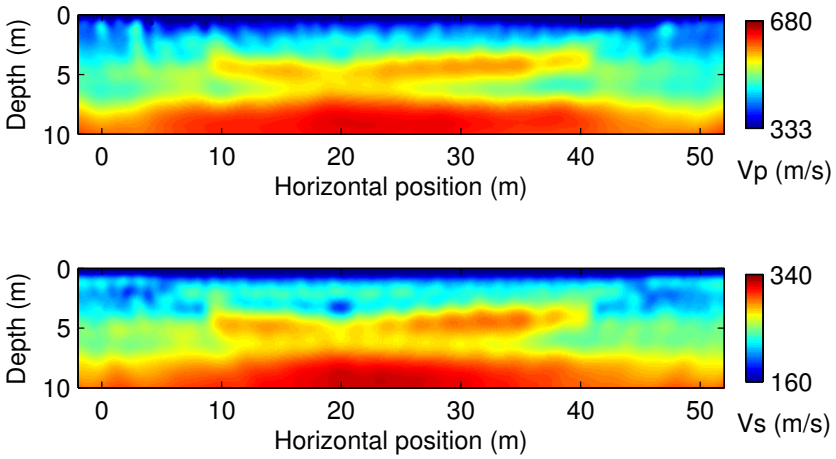


Figure 4.68: w-AWI multi-step inversion result. First step: only λ and μ are updated.

Step number and considered parameter →	initial error	1 λ and μ	2 ρ	3 λ and μ
Vp	11.1	5.7	6.4	5.9
Vs	11.4	5.5	6.1	5.6
ρ	7.5	7.5	8	8

Table 4.2: Relative error (in %) of models at the beginning (initial error) and at the end of each inversion step for a multi parameter inversion strategy using w-AWI.

data of step 1) is 52.88 %. Results of step 3 are shown in Figure 4.73 (2D) and Figure 4.74 (1D). The observed data are completely reconstructed as shown in Figure 4.75 with an additional misfit minimisation of 56.91 %. Overall, the misfit is minimised to 0.32 % after the three inversion steps. The final model error of this strategy, obtained after step 3, is smaller than the error at the beginning of inversion for λ and μ (see Table 4.2). The first step is the only one in which the errors are reduced but these quantities have been measured all over the model (including shallow and deeper parts). In the second and third steps, the error is increased but the data are better explained. The main problem in ρ came from the inversion in the deeper part of the model (Figure 4.71).

4.4.4 Inversion with low frequencies

In this test, we compare the result of FWI with w-AWI with low-frequency data.

The exact velocity model is layered with curved interfaces; velocity increases with depth and there is one lateral discontinuity at a horizontal position of 400 m (Figure 4.76). The grid conforms with the curved interfaces to avoid numerical diffractions. The observed data consists of 16 common shot gathers computed for sources applied on the vertical component of velocity and buried 0.5 m. The source signature is a 40 Hz Ricker wavelet. The initial data are computed with the same wavelet. The initial velocity models

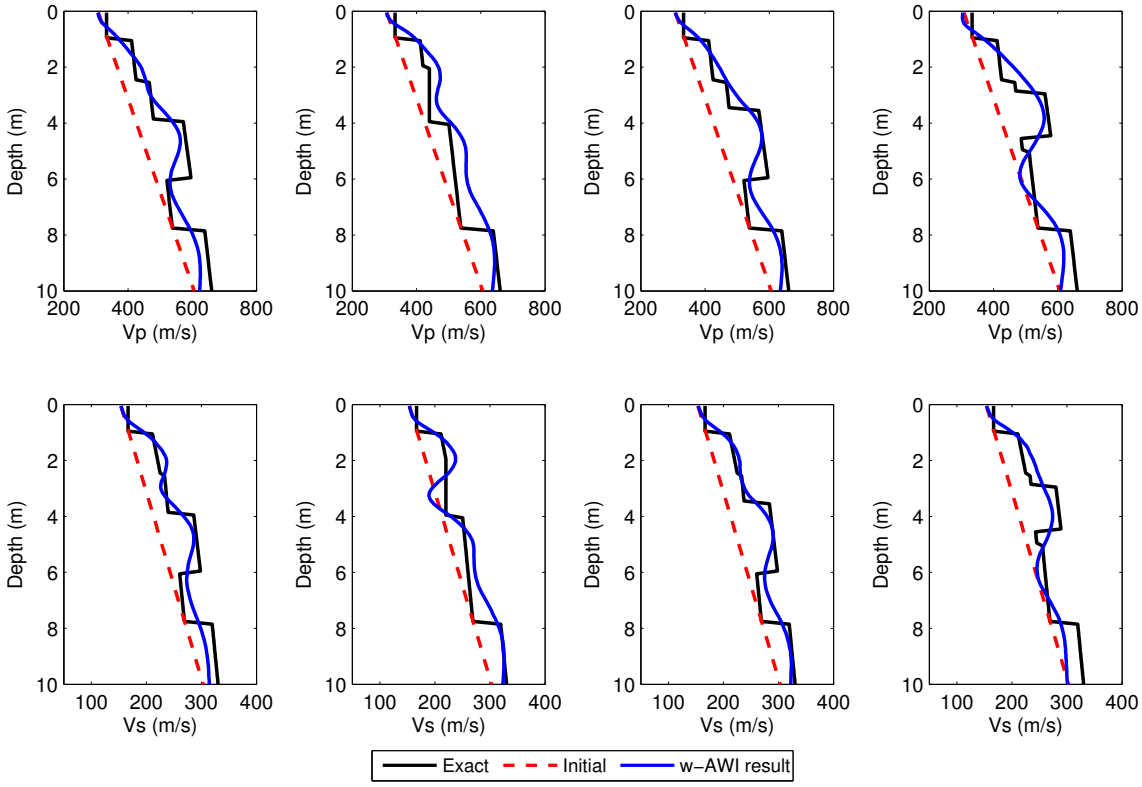


Figure 4.69: 1D plots of exact, initial and reconstructed models by w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m.

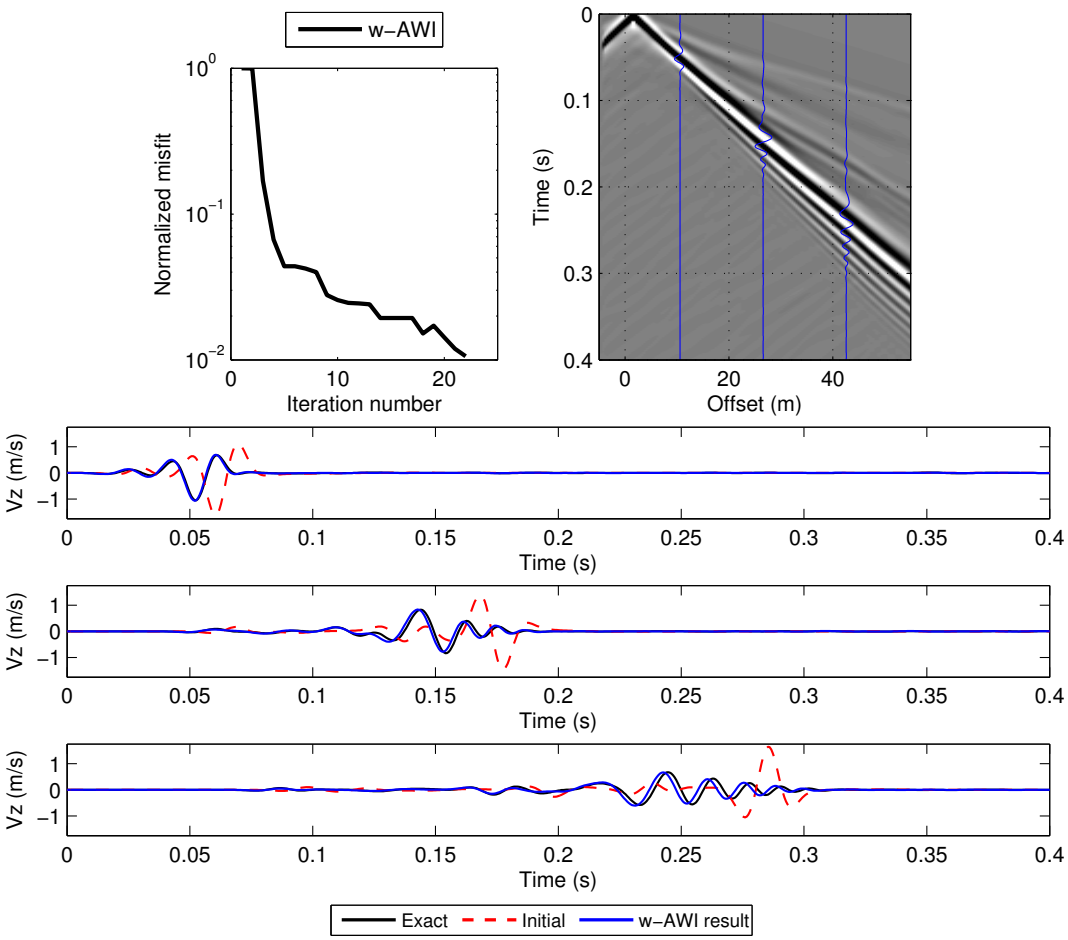


Figure 4.70: Misfit minimisation (top left) and final data obtained with w-AWI (top right). Bottom: seismic traces for 3 offset positions as indicated in the shot gathers.

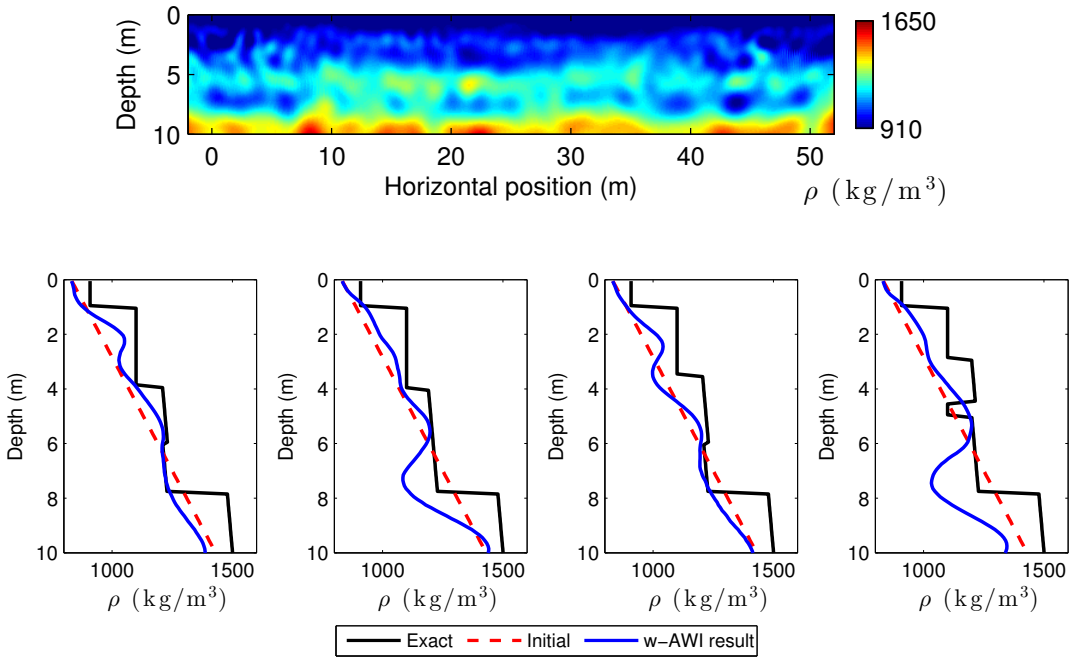


Figure 4.71: w-AWI multi-step inversion result. Second step: only density is updated. Top: final 2D density model. Bottom: 1D exact, initial and final density profiles. From left to right: x=13, 20, 27, 39 m.

are scaled versions of the exact models and they are defined as

$$\mathbf{m_i} = (1 - a_m)\mathbf{m_e} + a_m\mathbf{m_h}, \tag{4.28}$$

where $\mathbf{m_i}$ defines the initial model for Vp, Vs and ρ; $\mathbf{m_e}$ is the exact model and $\mathbf{m_h}$ is a homogeneous model defined for each parameter as Vp_h=2500 m/s, Vs_h=1500 m/s and ρ_h=1000 kg/m^3. In this example, a_m is equal to 0.15.

The purpose is to quantitatively evaluate inversion results. Thus, the relative errors of final models after FWI and w-AWI have been computed when considering non-filtered data (Figure 4.77) and band-pass filtered (2-9 Hz) data (Figure 4.78). The model retrieved from unfiltered data (Figure 4.77) shows smaller errors for w-AWI than for FWI. In w-AWI, velocities up to 60 m depth are reconstructed (dark blue in Figure 4.77). FWI recovers velocities up to 35 m depth including some stronger error zones near the surface (red zones in Figure 4.77). For filtered data, both methods show a similar behaviour; they are able to recover velocities up to the maximum considered depth (Figure 4.78). The final model relative errors (normalised by the initial error) are shown in Table 4.3. For low frequencies, FWI and w-AWI present similar results regarding these final errors (Table 4.3). The results of unfiltered data inversion show that the global minimum basin is wider in w-AWI since the first 60 m depth were reconstructed, even in the presence of energetic high frequencies.

The misfit is minimised down to 70.66 % by FWI and down to 6.96 % by w-AWI when considering unfiltered data. For low-frequency filtered data, the misfit is minimised down to 0.31 % by FWI and down to 1.89 % by w-AWI. FWI obtained the result in less iterations than w-AWI for unfiltered data while for filtered data the opposite behaviour

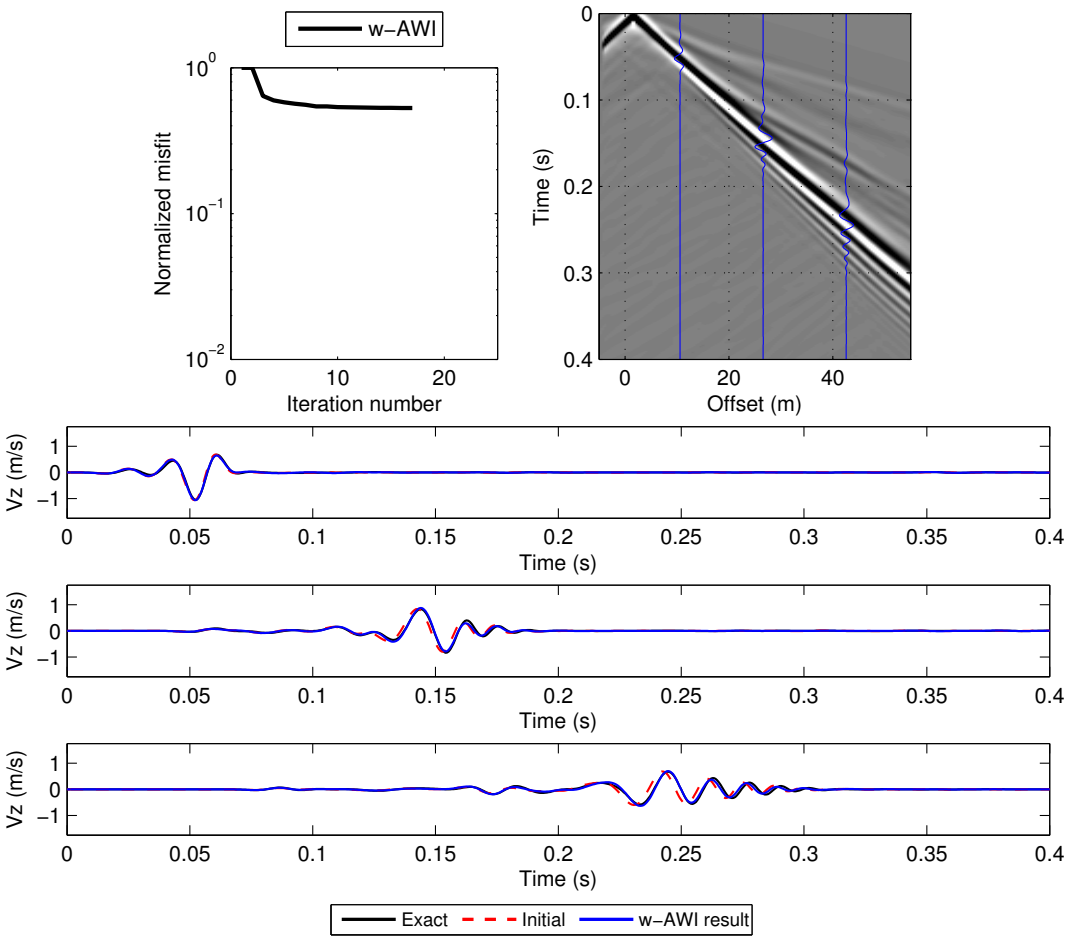


Figure 4.72: Misfit minimisation (top left) and final data obtained with w-AWI (top right) after density inversion. Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers.

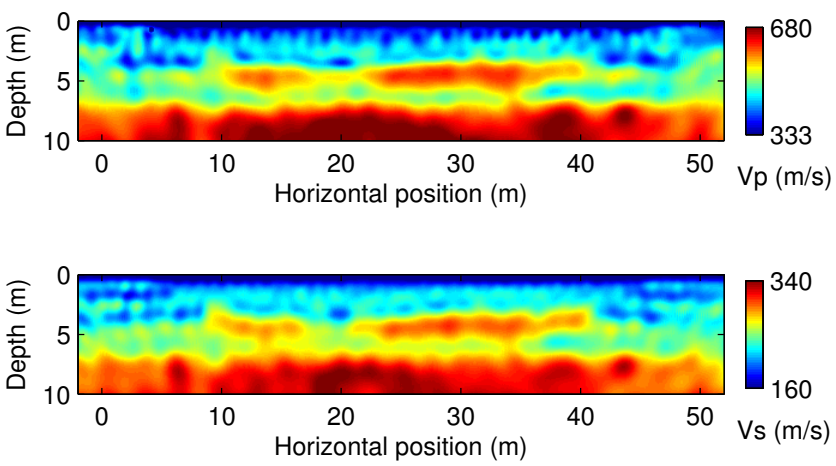


Figure 4.73: w-AWI multi-step inversion result. Third step: only λ and μ are updated.

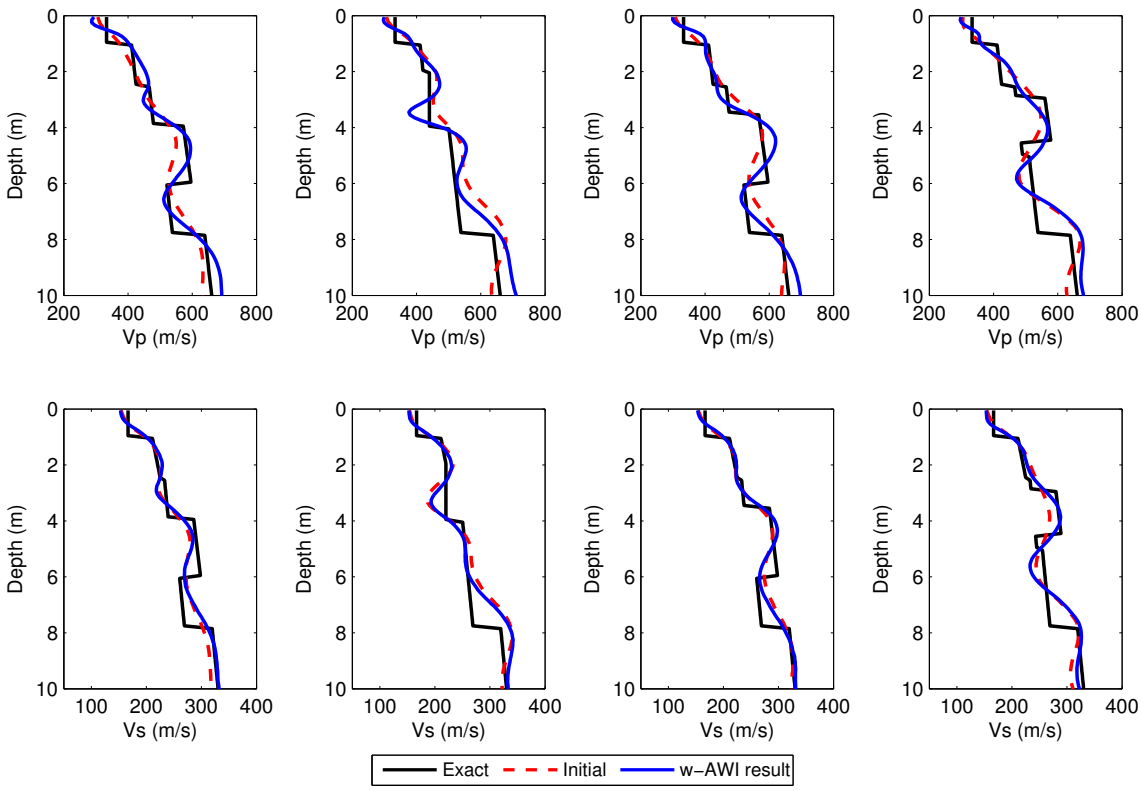


Figure 4.74: 1D plots of exact, initial and reconstructed models by w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m.

	FWI	FWI	w-AWI	w-AWI
	unfiltered	(2-9 Hz)	unfiltered	(2-9 Hz)
Vp	107.32	51.18	71.19	51.8
Vs	103.56	50.88	67.53	51.9

Table 4.3: Relative error (ξ) in % of reconstructed models with filtered and unfiltered data. These values have been normalised dividing by the initial error.

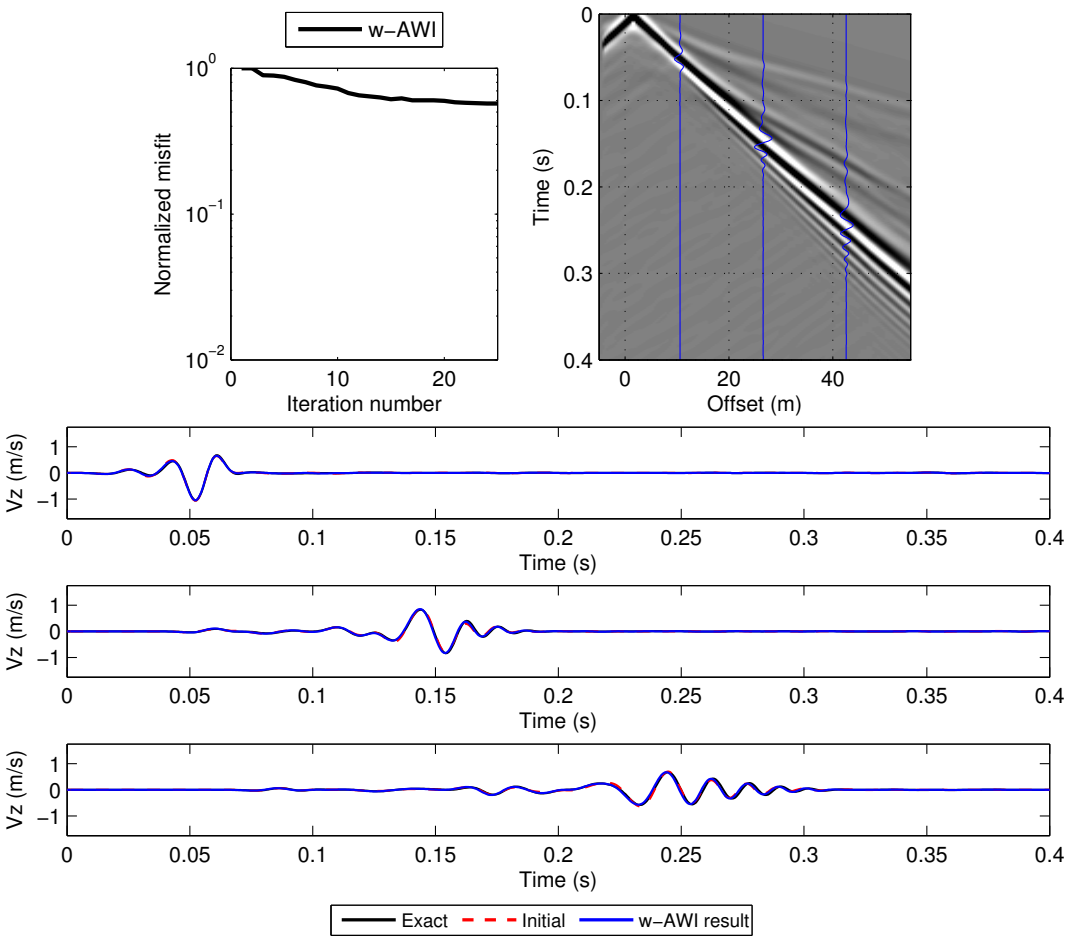


Figure 4.75: Misfit minimisation (top left) and final data obtained with w-AWI (top right). Bottom: seismic traces for 3 offset positions as indicated in the shot gathers.

is observed (Figure 4.79). A comparison between observed and initial data show the high energy of dispersive waves in the case of filtered data and how the frequency filtering can help avoiding cycle-skipping (Figure 4.80). The representation of seismic traces for observed, initial, and final data (with FWI and w-AWI) help observing the mismatch in the phase that resulted in FWI problems at the near surface (Figure 4.81 and 4.82) On the other side, the phase match accomplished with low-pass filtering resulted in the improvement of the FWI result. For w-AWI, the result is not affected by cycle-skipping effects in the tested inversions either with filtered or unfiltered data.

As stated in (Rix and Leipski, 1991), surface-wave imaging is expected to have an illumination depth equal to the wavelength λ^f of the considered frequency. For unfiltered data, considering a central frequency of 40 Hz this value is $\lambda^{40\text{Hz}}$=960(m/s)/40(Hz)=24 m. For 9 Hz, the maximum wavelength is $\lambda^{9\text{Hz}}$=960(m/s)/9(Hz)=106.67 m. Let us compare these values with the misfit function gradient at the first iteration (Figures 4.83 and 4.84). As expected, FWI and w-AWI gradients with unfiltered data show high amplitudes up to $\lambda^{40\text{Hz}}$. The divergence from the exact model obtained by FWI (Figure 4.77) could be related with an overestimation of the near-surface layer, especially at the left of the lateral discontinuity (horizontal position of 400 m). In the case of w-AWI, the first gradient exhibits energy up to $\lambda^{40\text{Hz}}$. Moreover, w-AWI has deepened the illumination up to at least $2\lambda^{40Hz}$ as shown in Figure 4.77, mostly because of its larger global minimum basin. For filtered data, the energy is extended over the whole gradient because $\lambda^{9\text{Hz}}$ is slightly larger than the maximum model depth. In this case, the low frequency illumination is

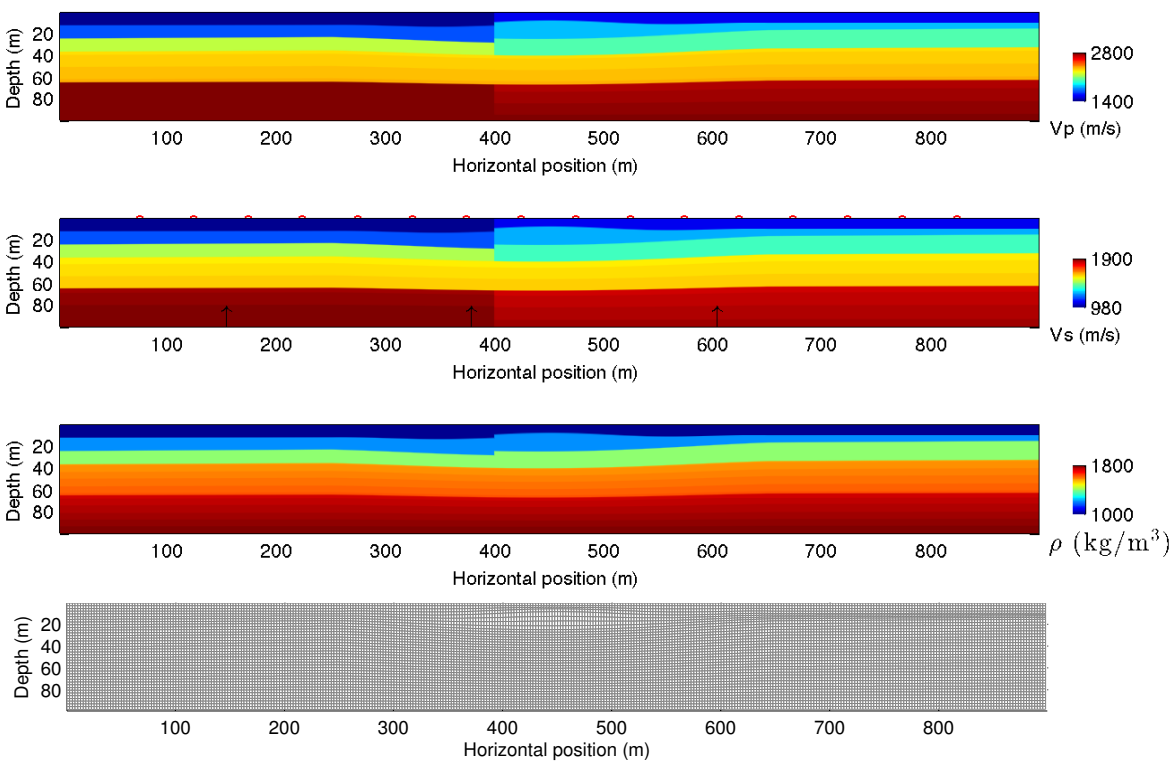


Figure 4.76: Exact model. From top to bottom: Vs, Vp, ρ and grid for modelling and inversion.

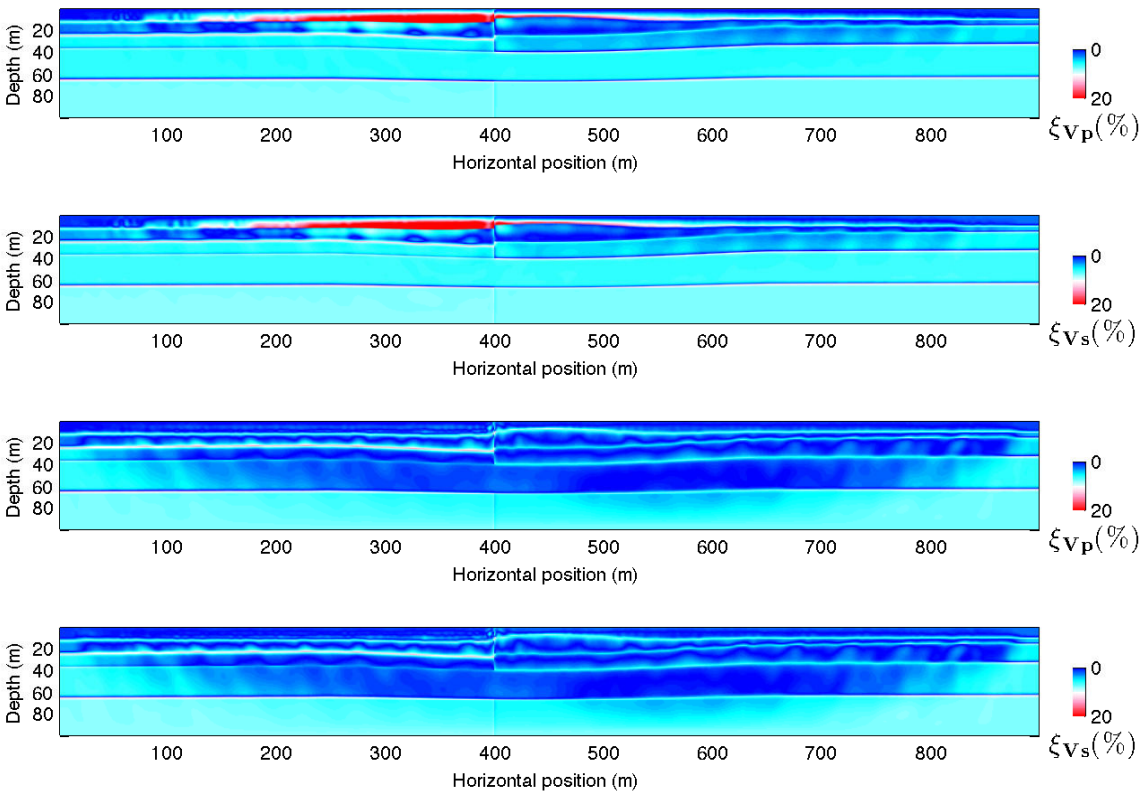


Figure 4.77: Unfiltered data inversion. Relative error between final and exact models. Top: Vp and Vs errors for FWI. Bottom: Vp and Vs errors for w-AWI.

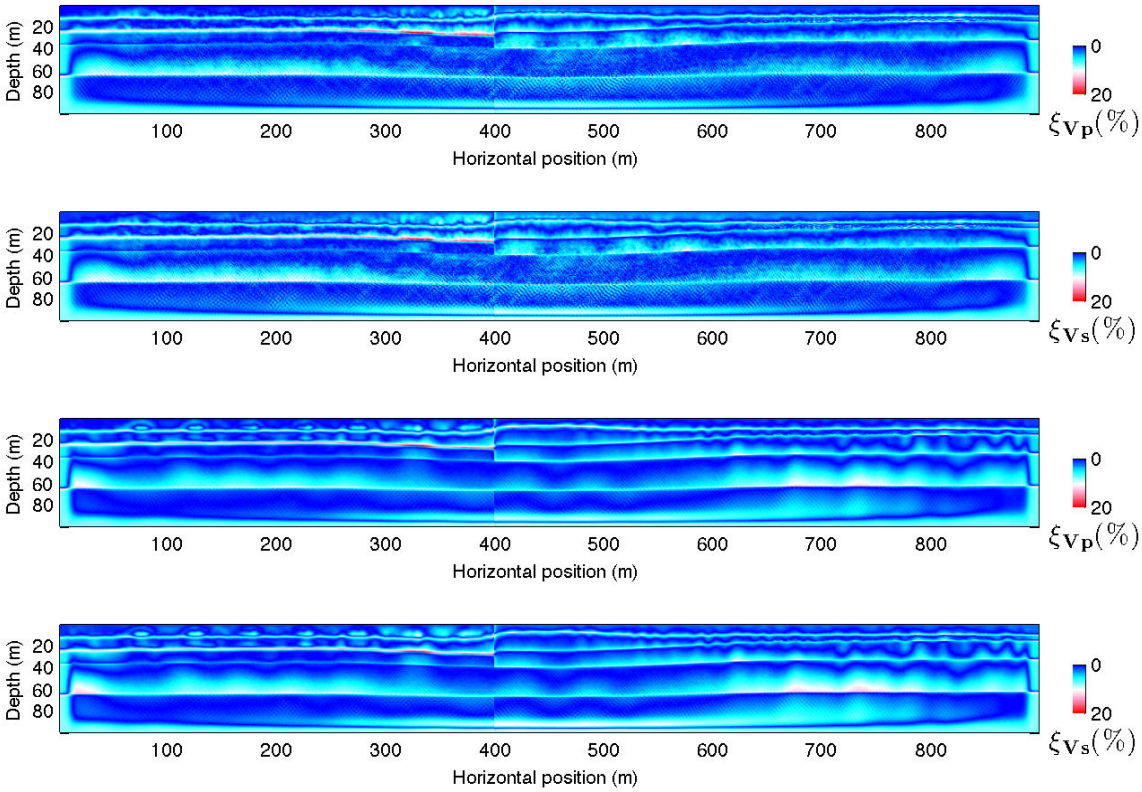


Figure 4.78: Bandpass filtered (2-9 Hz) data inversion. Relative error between final and exact models. Top: Vp and Vs errors for FWI. Bottom: Vp and Vs errors for w-AWI.

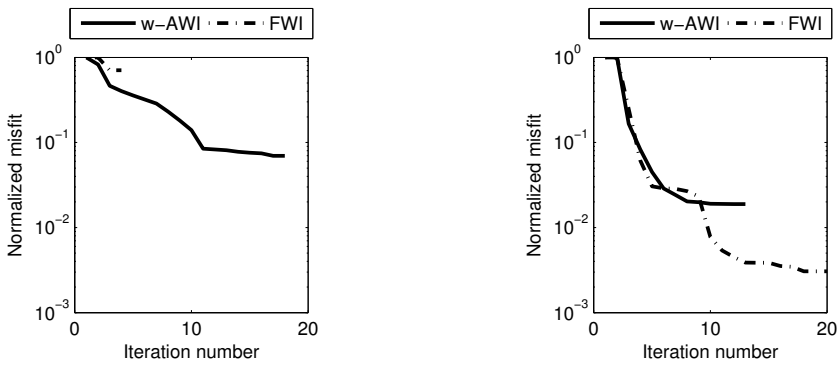


Figure 4.79: Objective function evolution with iterations. Left: unfiltered data. Right: filtered data.

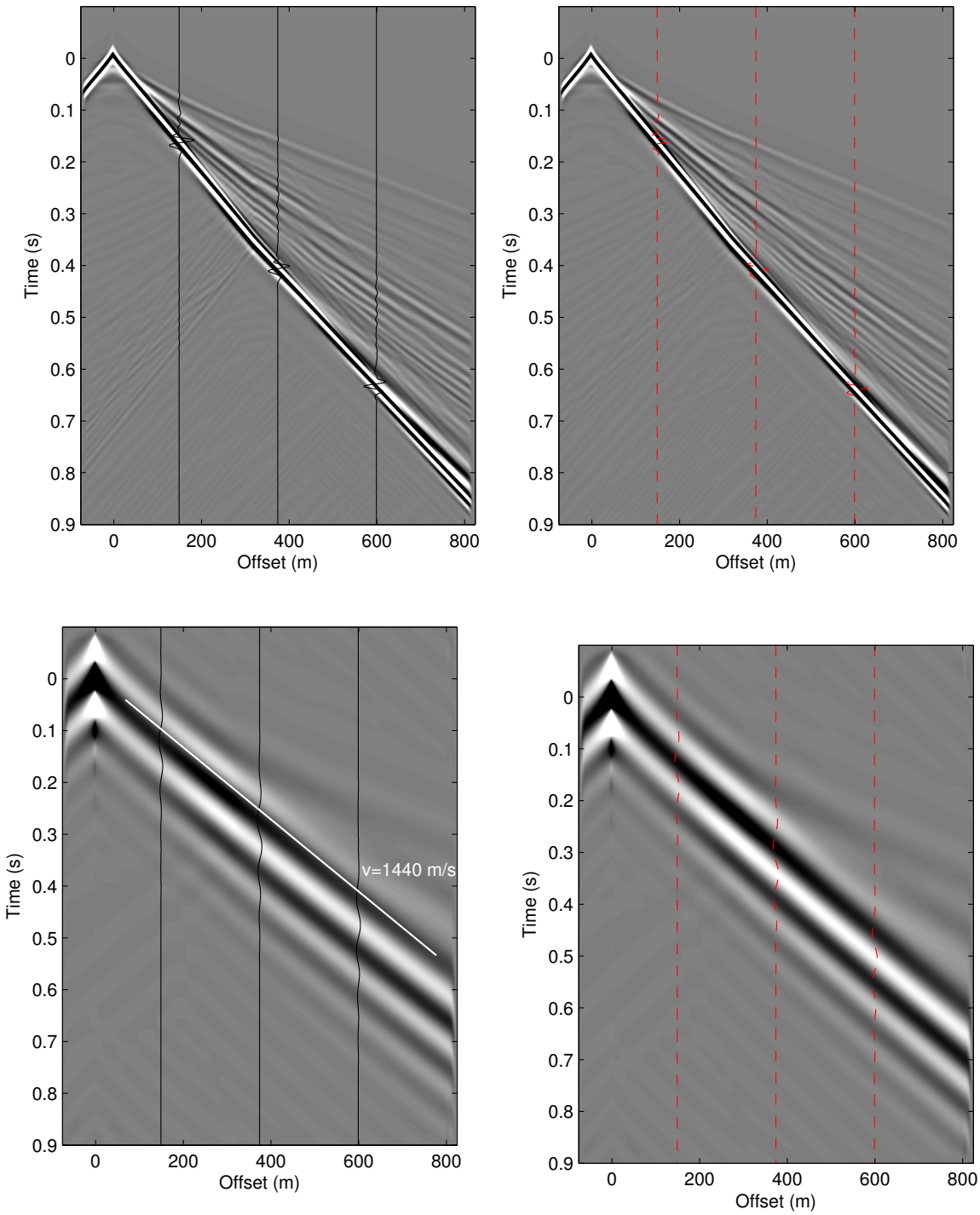


Figure 4.80: Shot gathers of observed (left) and initial (right) data without filtering (top) and with bandpass filtering (2-9 Hz) (bottom). Clipping level is equal for images in the same row.

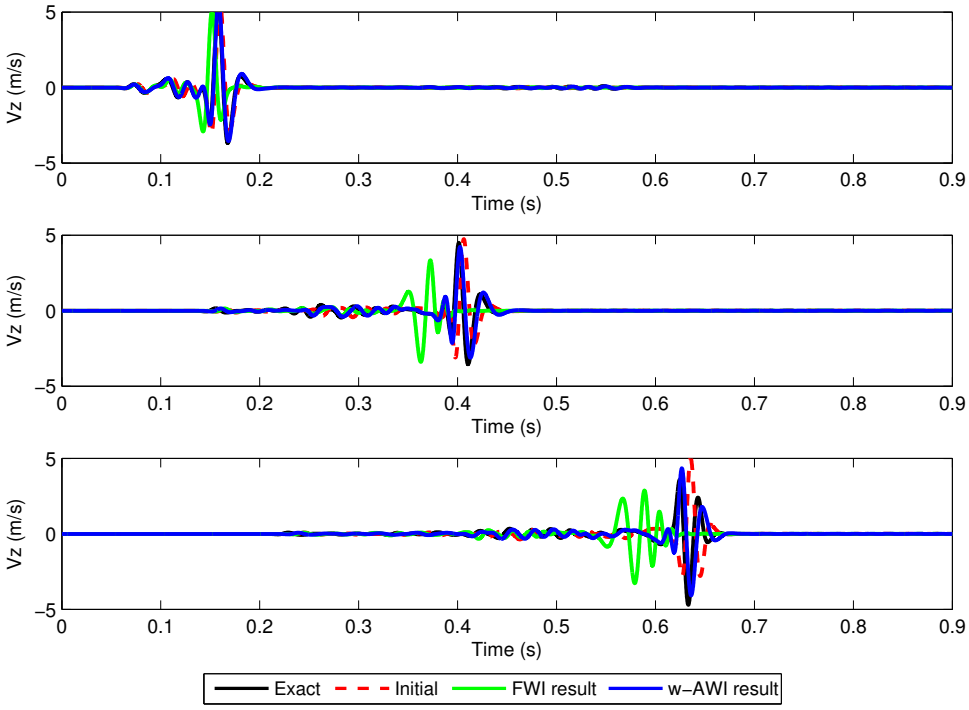


Figure 4.81: Comparison of seismic traces for unfiltered data at three offsets: 150 m, 375 m and 600 m.

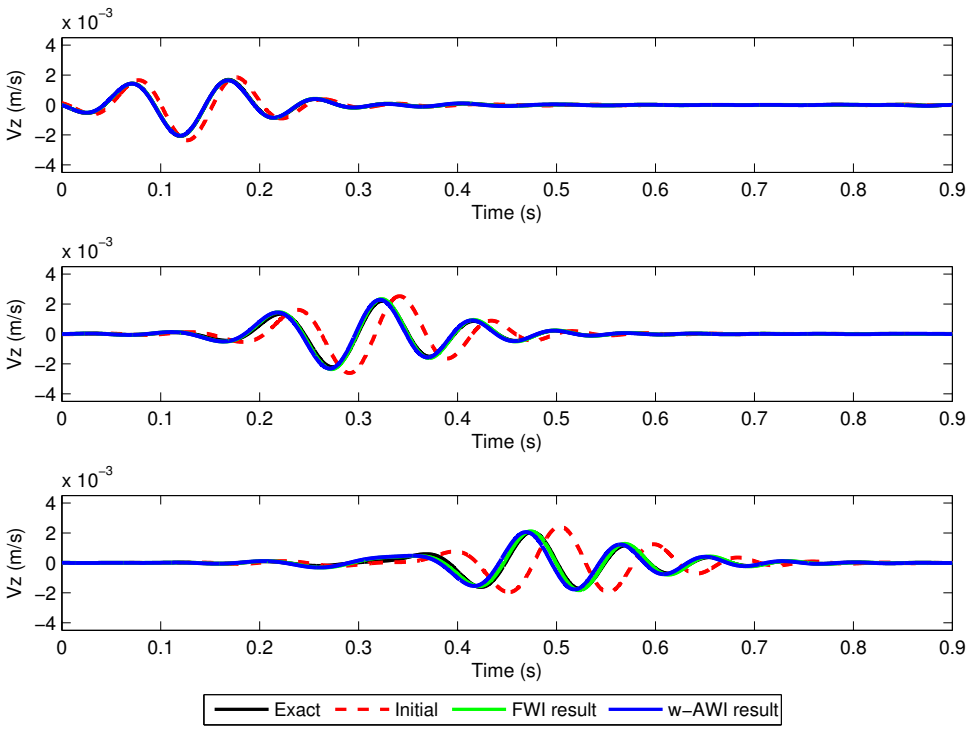


Figure 4.82: Comparison of seismic traces for filtered data (2-9 Hz) at three offsets: 150 m, 375 m and 600 m.

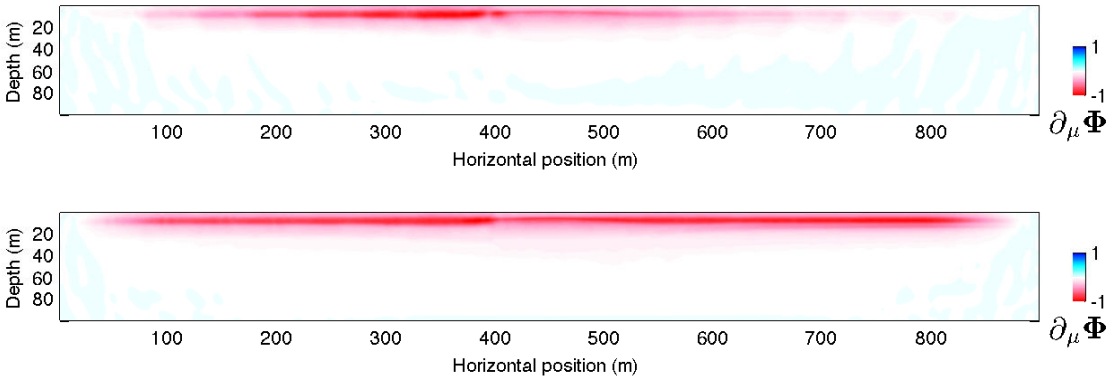


Figure 4.83: Gradient of the FWI (top) and w-AWI (bottom) objective function with unfiltered data.

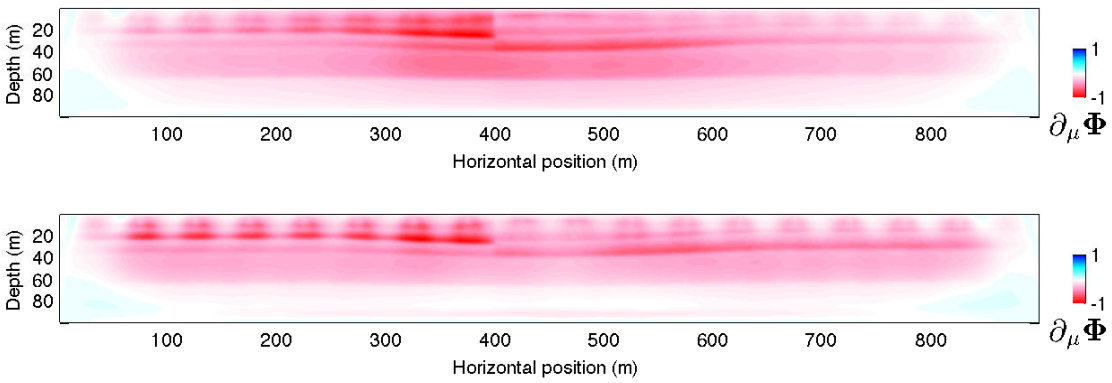


Figure 4.84: Gradient of the FWI (top) and w-AWI (bottom) objective function with filtered (2-9 Hz) data.

mainly accomplished by waves propagating at approximately v=1440 m/s (Figure 4.80). Such velocity value can either belong to the range of low-frequency surface-wave dispersion or correspond to S-wave diving events.

4.5 Conclusions

We have presented a methodology to invert surface waves in 2D media. The windowed amplitude waveform inversion (w-AWI) is an intermediary approach between Full Waveform Inversion (FWI) and classical Surface Wave Analysis (SWA). The principal difference with respect to FWI is the misfit functional. The new misfit functional is more similar to the one used in SWA. In order compute the misfit, the seismic gathers are first separated using spatial windows, and then the 2D Fourier transform is applied to the different (windowed) gathers. The misfit value is equal to the sum, over the windows, of the least-squares misfit between the absolute value of transformed gathers. *By taking only the absolute value of the transform, some local minima effects can be avoided.* Besides, the implementation of windows helps improving the localisation of velocity perturbations, provided a proper window width is chosen. The windows help reintroducing information about the phase.

Analysis of w-AWI

We have performed numerical tests comparing the properties of w-AWI and FWI. An analysis of local minima has shown that the basin of the global minimum of the objective function is wider in w-AWI. In the considered tests, the w-AWI basin is at least twice as wide as the FWI basin. We have explained this result by analysing the "cycle skipping" due to errors in the initial velocity model. In FWI, there is cycle skipping if the phase shift between observed and modelled data is larger than half the dominant wavelength. We have shown that in w-AWI, the cycle-skipping limit is more than half the wavelength, although rarely greater than a complete wavelength. *This result proves that w-AWI is approximately twice as flexible as FWI regarding the choice of an initial velocity model, which is a needed property when inverting dispersive surface waves.*

In other numerical tests, we have shown that the w-AWI gradient resolution obeys a spatial-windows width trade-off: small windows can improve the localisation of velocity perturbations but excessively small windows lead to resolution problems in the Fourier transform. It is known that in FWI, cycle skipping can sometimes be avoided by inverting low-frequent data content first. We have shown that w-AWI may achieve convergence in the case of unfiltered data. Furthermore, low-pass filtered data may increase surface-wave illumination depth in both techniques.

Inversion of synthetic data

We have also analysed the influence of the source wavelet in FWI and w-AWI. The two inversion techniques yield similarly good results if the inversion is well constrained and if the source wavelet is known. Moreover, if the initial model is far from the exact one (high errors in the model) then w-AWI may converge because it is less constrained with

respect to the choice of initial model. In our examples, energetic high frequency surface waves can explain divergence in FWI. If the source wavelet is unknown and inversion converges, then the data reconstructed by w-AWI usually contain the amplitudes of the data only (because of the absolute value consideration). The final data are then visually different from the observed data due to the missing phase in the definition of the objective function. However, the data phase could be recovered by applying FWI after w-AWI.

The inversion of density is difficult with all the considered techniques. If the density is not updated in inversion, then Vp and Vs models tend to explain the errors in density. However, if the density is updated then the density errors may increase even if the data misfit decreases. This result could be due to a stronger non-linearity with respect to the density that easily leads to ambiguous results. This observation could be validated by diffraction pattern analysis in the presence of a free surface in elastic media.

Finally, a test conducted with unfiltered data have revealed that w-AWI is able to converge towards the correct model even when the data contain all low and high frequencies, extending the normal illumination depth to at least $2\lambda_s^{f_c}$ (two times the S wavelength for the central frequency). For the same test, FWI did not converge. Moreover, FWI and w-AWI can be used to recover deeper parts of the model if low frequency data are used.

Chapter 5

Real data application

Contents

5.1 Introduction **169**

5.2 Seismic dataset **171**

5.2.1 Preprocessing 172

5.2.2 Initial velocity model 174

5.3 Seismic inversion **174**

5.3.1 SWA 175

5.3.2 FWI and w-AWI 177

5.4 Discussion **189**

5.5 Conclusion **194**

Résumé du chapitre

Dans ce chapitre applicatif, nous nous sommes intéressés à estimer un modèle élastique qui explique des données sismiques réelles. Ces données ont été acquises sur un milieu connu et présenté par Deidda and Balia (2001). L'objectif de l'inversion est d'imager une structure en béton qui se caractérise par une vitesse de propagation élevée. Nous appliquons l'approche *1D Surface Wave Inversion* (SWA), et aussi les approches en forme d'onde *Full Waveform Inversion* (FWI) et *windowed-Amplitude Waveform Inversion* (w-AWI). Les résultats obtenus sont encourageants, même si des analyses supplémentaires seront nécessaires pour améliorer le résultat d'inversion.

Dans la partie 5.2, nous présentons le jeu de données réelles et le modèle Vp préliminaire issu de la tomographie. La ligne sismique est composée de 61 point de tirs, chacun associés à 72 géophones. Même si la qualité des données est bonne, des étapes de pré-traitement doivent être appliquées pour l'inversion. Nous considérons une correction d'amplitude de 3D à 2D (Bleistein, 1986; Pratt, 1999). Des interpolations (Spitz, 1991; Naghizadeh and Sacchi, 2007) sont réalisées pour améliorer la cohérence des événements sismiques. Enfin, un masque relatif aux temps d'arrivée ainsi que des filtrages en fréquence sont considérés pour les tests d'inversion. Un modèle de vitesse des ondes P (Vp) obtenu par tomographie des temps d'arrivée nous a été fourni par Gian Piero Deidda (qui a aussi mené l'acquisition des données sismiques). Ce modèle sera le modèle de départ pour les inversions de forme d'onde.

Dans la section 5.3, les résultats des tests d'inversion sont montrés et analysés. Nous faisons séparément des inversions avec les approches SWA, FWI et w-AWI. Une inversion jointe qui consiste à utiliser la w-AWI dans un premier temps, puis à appliquer la FWI, est préféré car le résultat montre une structure à vitesse élevée placée à la position théorique de la structure en béton. Dans d'autres tests, l'importance de la fréquence minimale des données est démontrée. Cependant, la w-AWI a aussi besoin de fréquences élevées puisque la sensibilité des ondes de surface est répartie sur le spectre (partie 3.3).

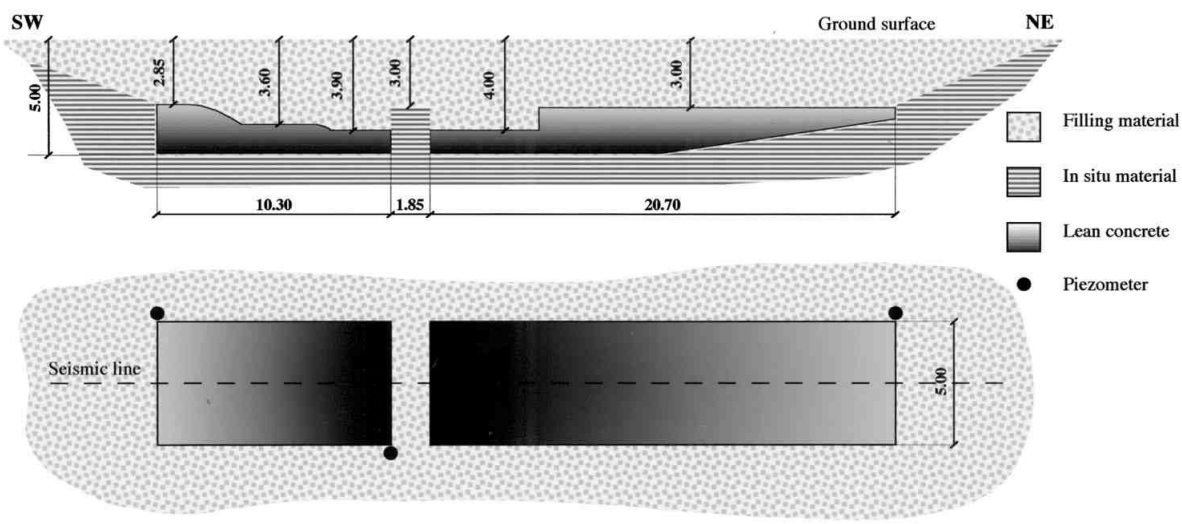
Enfin, une discussion est proposée avec pour but d'analyser les résultats obtenus (partie 5.4). Comme conclusion, nous obtenons que tout le spectre des données devrait être inclus dans l'inversion menée avec w-AWI car cette méthode travaille surtout avec des ondes de surface.

5.1 Introduction

We present an application on a ultra-shallow seismic dataset with a target at around 5 metre depth. The objective is to evaluate Surface Wave Analysis (SWA), Full Waveform Inversion (FWI) and windowed-Amplitude Waveform Inversion (w-AWI) implementations. Note that waveform-based techniques applied to engineering and environmental studies are not common approaches (Gao et al., 2007; Gélis, 2005).

The real seismic dataset under study was acquired in 2013 on the purpose-built ground model presented in Deidda and Balia (2001). The acquisition was carried out by the cited authors. The subsurface is described in simple terms as 3 laterally-varying layers which are from bottom to top: (1) unconsolidated sediment layer (silt and clay) (2) concrete layer and (3) compacted back fill material. Filling material refers to sediments extracted before concrete casting. The model (Figure 5.1) was originally built with the purpose of evaluating common-depth-point seismic reflection surveys in geological targets typically encountered in engineering investigations. The acquisition geometry and geophone description for our seismic dataset is given in detail in section 5.2.

The P-wave and S-wave velocities (Vp and Vs) shown in Figure 5.1 were measured in walk-away preliminary tests (Deidda and Balia, 2001). These velocity values are much lower than normally expected for the unconsolidated part. Compressional wave speeds are around 345 m/s in the air (1 bar and 20C) and 3500 m/s in quartz. An unconsolidated medium with porosity close to 40 % at the surface has a velocity even slower than in



	P-wave velocity	S-wave velocity	Density
Filling material	**135 m/s**	**72 m/s**	**1800 kg/m^3**
Lean concrete	**1800 m/s**	**900 m/s**	**2230 kg/m^3**

Figure 5.1: Purpose-built ground model, extracted from Deidda and Balia (2001). Velocity and density values obtained with walk-away tests are indicated. The displayed distances are in metres. Piezometers were placed to check water table levels.

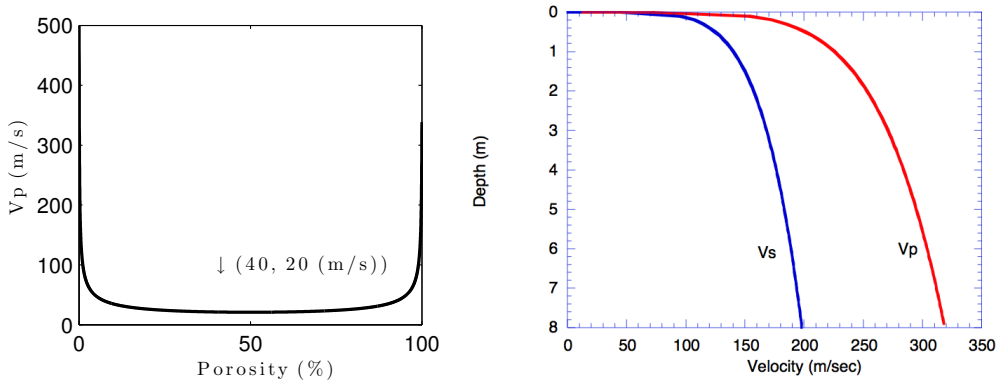


Figure 5.2: Theoretical wave speed values in the near surface. Left: waves can propagate as slow as 20 m/s (at critical porosity) (Bachrach et al., 1998). Right: velocity increases rapidly with depth, Bachrach and Nur (1998).

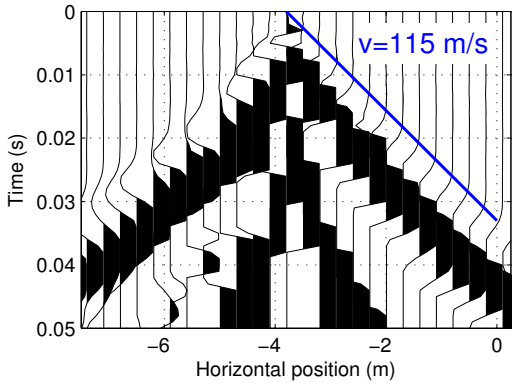


Figure 5.3: The apparent velocity within the first 3 m offset is around 115 m/s (shot-gather 20 of the real dataset).

the air. This has been noticed by some authors as Lester (1932); Domenico (1974); Bachrach et al. (1998). The key element defining near surface velocities is porosity. High porosity values represent almost suspension of grains in the air. When an acoustic wave propagates in such a medium, part of the energy is transformed to grain motion. Density of the sand grains is about 2000 times higher than density in the air. As a result, the wave propagates with velocity lower than the velocity in the air. Maybe surprisingly, Bachrach et al. (1998) presented simple formula for which compressional wave can be as slow as 13 m/s. Using the same equation, we show that for unconsolidated sand and for porosity values between 10 to 40 % the velocity is slower than 36 m/s (Figure 5.2). The lowest velocity measured by Bachrach et al. (1998) is 50 m/s. These low values are only expected at ultra-shallow depth (tens of centimetres). For depths larger than a few tens of cm, the velocity rapidly increases due to compaction effects and wave speed can be in the order of 230 m/s at 1 m depth Bachrach and Nur (1998). The lowest apparent velocity in our real data set is close to 115 m/s (Figure 5.3).

This chapter is organised as follows. The seismic data acquisition parameters and

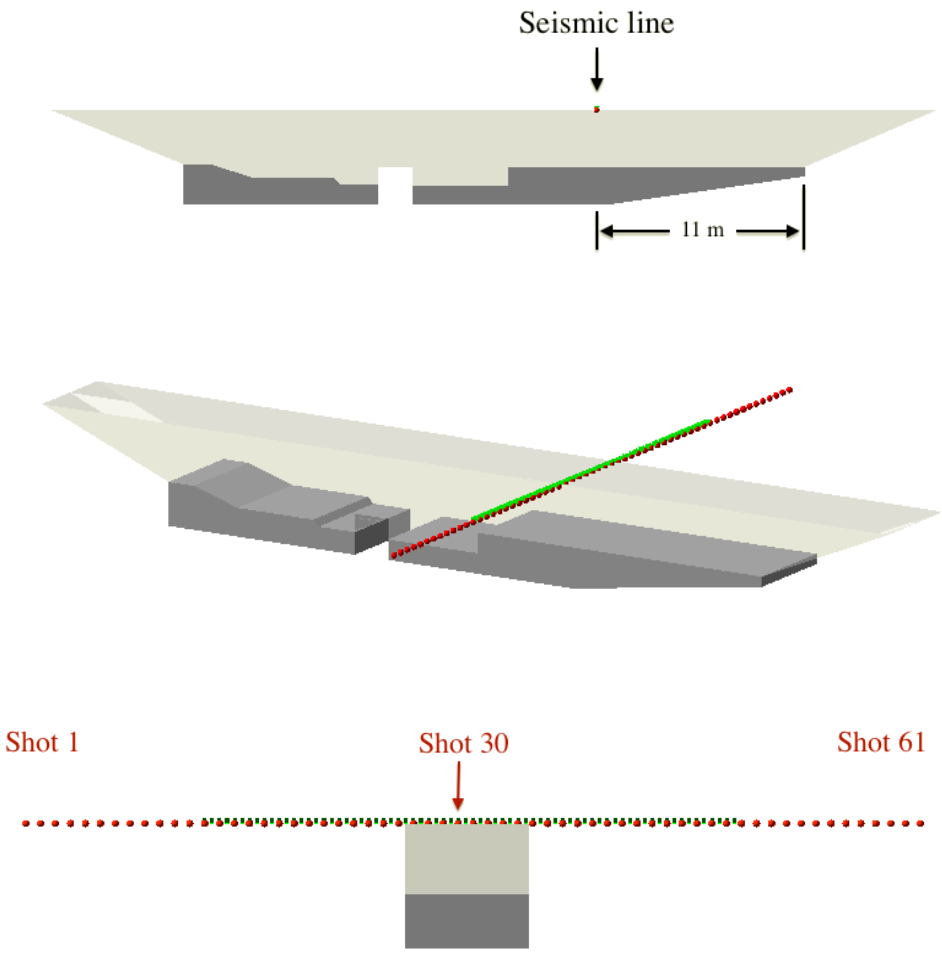


Figure 5.4: Sketch of the acquired seismic line and the ground model. Geophone and source positions are respectively indicated in green and red. Top: cross section. Middle: 3D view. Bottom: cross-line cross section.

data preprocessing are presented in section 5.2. We show the compressional-wave velocity model reconstructed with first-arrival traveltime tomography. This model was kindly provided by Gian Piero Deidda. Section 5.3 is devoted to seismic inversion results. Traveltime tomography Vp model is used as initial model in our tests. We have applied SWA based on the fundamental mode only (results are presented in Section 5.3.1). Shear-wave velocity was reconstructed from this analysis. Finally, we use FWI and w-AWI to invert filtered data using different frequency bands (results in Section 5.3.3).

5.2 Seismic dataset

The seismic line direction is perpendicular to the inline direction of the concrete body. The seismic dataset consists of 61 shot gathers with sources spaced 0.6 m apart. A total

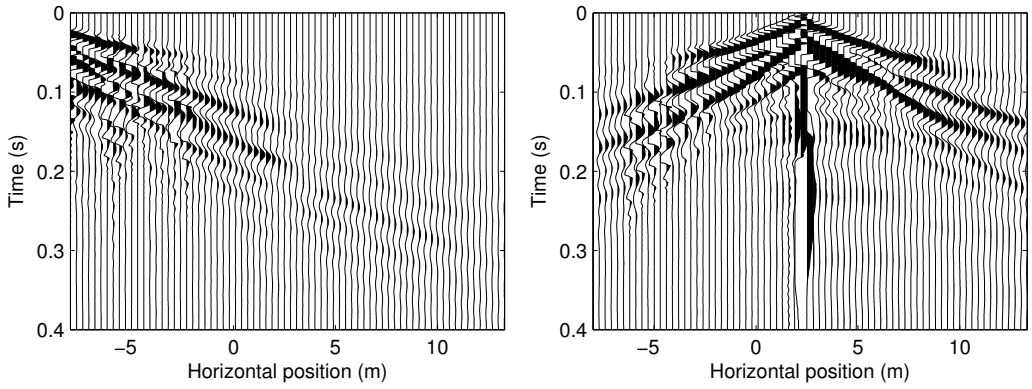


Figure 5.5: Representative raw shot gathers 10 and 30 before preprocessing.

of 72 vertical geophones were used at the surface according to the geometry shown in Figure 5.4 (sources in red, receivers in green). The distance between geophones is 0.3 m. The first geophone is positioned at 8.1 m outside of one of the lateral model boundaries, while the first source point is 7.2 m from the first geophone (15.3 m from the model boundary). The source line is laterally offset by 15 cm relative to the geophone line. The maximum source-receiver distance (offset) is 25.2 m.

Sledgehammer blows on a metallic base plate were used as a seismic source (one per source position). Geophone frequency response is flat between 4.5 and 160 Hz. Three different groups of geophones were employed (differing by date of fabrication). This is an additional difficulty for this dataset as geophone may possibly differ in amplitude and frequency.

Two raw seismic shot gathers are shown in Figure 5.5. The data quality is in general good. However, some traces on the left (negative horizontal position) seem to be noisy. The energy recorded by near offset geophones exceeds their dynamic range. As a result, the near-offset traces are clipped and do not contain clear seismic wave information.

5.2.1 Preprocessing

The goal is to preprocess the data for subsequent processing (SWA, FWI and w-AWI). The preprocessing stages that we have applied are

- 3D-to-2D geometrical spread correction. The first stage consists in multiplying the data by $\sqrt{t}$, where t is the recorded time. This correction works for surface and body waves in 2.5D media (Bleistein, 1986; Pratt, 1999).

- Seismic event coherency improvement. The coherency was improved by replacing noisy traces with interpolated ones. First, all the signals were checked to identify good and noisy traces. Then, one of two interpolation approaches was applied. F-X interpolation was applied, if both the traces on the left and on the right were good (Spitz, 1991). If such condition is not satisfied, then linear prediction filters are applied. Such filters are especially suited for irregularly missing data reconstruction

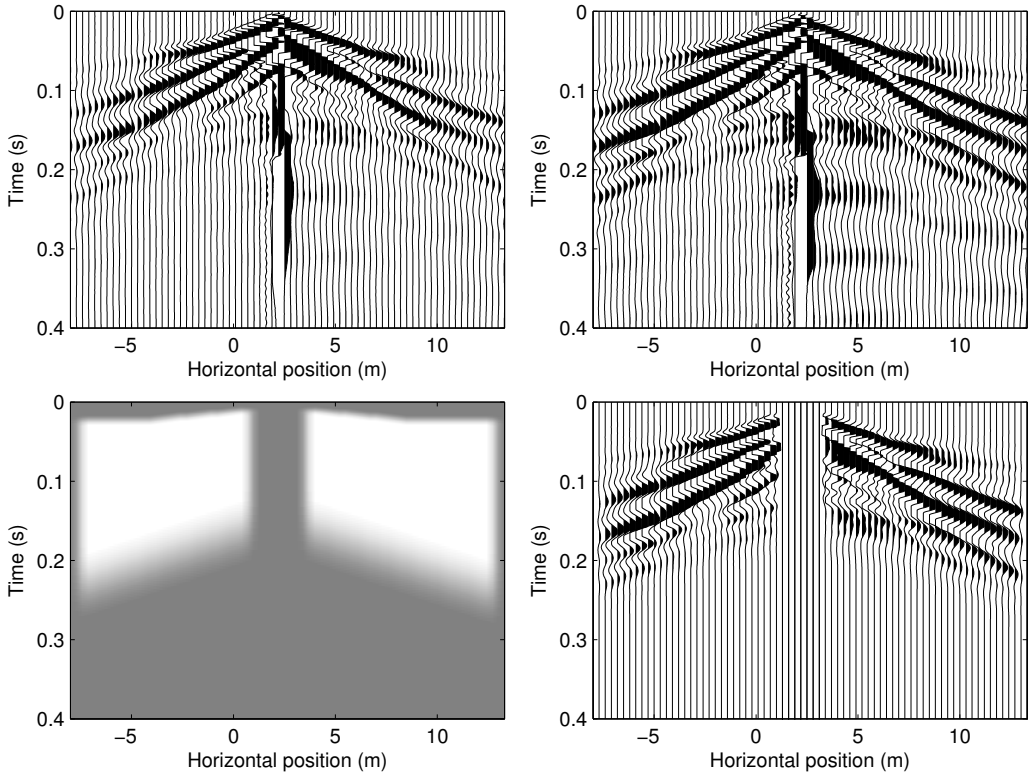


Figure 5.6: Shot gather 30 after different preprocessing stages. Top left: after interpolation to remove noisy traces. Top right: after interpolation and geometrical spreading correction. Bottom left: mask used to mute near offset traces and late arrivals. Bottom right: preprocessed data after multiplication with the muting mask. The mask is tapered to prevent inversion artifacts.

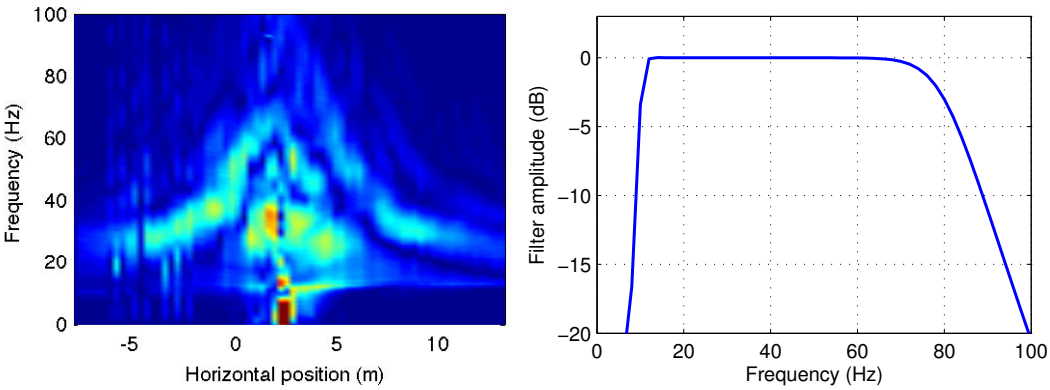


Figure 5.7: Left: frequency-space spectrum of raw shot gather 30 (Figure 5.5). Right: example of the 10th-order Butterworth band-pass filter considered for inversion. In this case, the band-passed is 10-80 Hz.

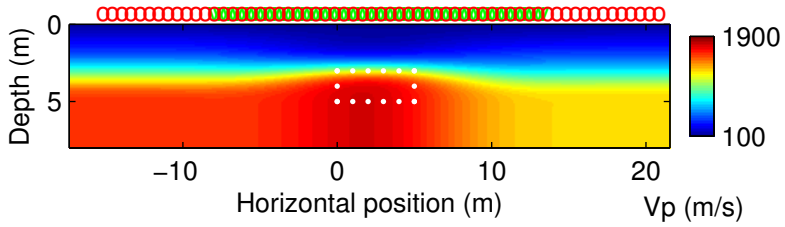


Figure 5.8: Vp model obtained with first-arrival traveltime tomography. The white points indicate the theoretical position of the concrete anomaly.

(Naghizadeh and Sacchi, 2007). Interpolation was applied in no more than 10 % of the seismic traces.

- Near offset and late arrivals muting. A mask was considered here (Figure 5.6) to remove unwanted events enhancing useful signals of surface and transmitted body waves.

The result of preprocessing is presented for shot gather 30 (Figure 5.6).

Further preprocessing includes frequency filtering (for waveform inversion). We employed 10th-order zero-phase digital Butterworth band-pass frequency filters (Oppenheim et al., 1998). We selected a 10 Hz low-cut frequency after spectra inspection. This frequency should commonly be dictated by the geophone low-cut frequency, although we have chosen it by inspection because different types of geophones were employed. In Figure 5.7, the absolute value of the spectrum for a representative shot gather is shown. The high-cut frequency is chosen between 18, 25, 36, 50 or 80 Hz depending on the waveform inversion objectives (results shown in Section 5.3). One of the considered frequency filters (10-80 Hz) is shown in Figure 5.7 (right).

5.2.2 Initial velocity model

A smooth Vp velocity model was obtained and kindly provided to us by Gian Piero Deidda using picked first-arrival traveltime tomography (Figure 5.8). The velocity is restricted to the first 5 m. Nevertheless, this result will be used as initial velocity model in waveform inversion because the first arrivals have been explained with a low RMS error (1.6 %). The Vs and density (ρ) models are less known and the only indication consists of the values indicated in Figure 5.1.

5.3 Seismic inversion

We have applied the inversion techniques presented in Chapters 3 and 4 (SWA, FWI and w-AWI) to the considered real dataset. The main objective of this application is to retrieve a model anomaly (Vp, Vs, ρ) in the central part of the model (Figures 5.4 and 5.8). In the concrete, higher velocity values are expected. In the next paragraphs, we describe the inversion tests and associated results.

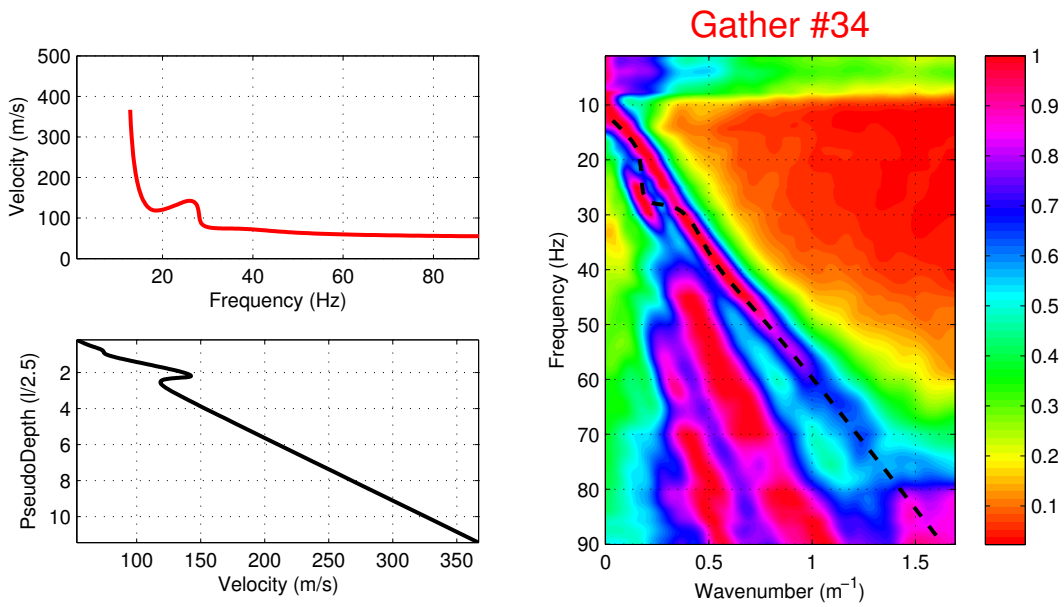


Figure 5.9: Left: fundamental mode plotted in the f-v domain (top) and with respect to pseudo depth (bottom). Right: surface-wave fundamental mode picked in gather # 34 (dashed black line).

5.3.1 SWA

The analysis here consists in inverting several 1D Vs models located at different horizontal positions which can then form a pseudo 2D model. The surface-wave fundamental mode was picked in 72 $f-k$ gathers. These gathers were obtained as follows

- Each shot gather was separated using 72 different spatial windows centred around each geophone position. The window only selects traces with offset smaller than 9.3 m for a maximum of 64 traces per window.
- The $f-k$ gathers were obtained by computing the 2D Fourier transform of each individual window.
- 72 stacked $f-k$ gathers were obtained by summing individual gathers corresponding to identical receiver positions: this means that different shots may contribute to the same $f-k$ gather.
- The fundamental mode was semi-automatically picked in the stacked gathers. A manual picking correction was applied when it was considered necessary (mostly at frequency below 30 Hz).

One of the results of the picking process is represented in Figure 5.9 (right). On the top left, the picked mode plotted in the frequency-velocity axis is shown. A pseudo-depth ($\lambda/2.5$) velocity model is shown on the bottom left (Figure 5.9). This profile is used as initial model for fundamental mode inversion.

The extracted dispersion curves were inverted using the approach described in Chapter 3. The Vp model used in inversion was obtained by averaging and layering the Vp tomography model where high velocity contrasts were observed (Figure 5.8). Surface waves

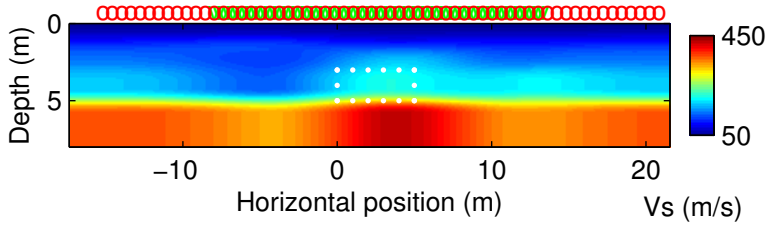


Figure 5.10: Pseudo 2D Vs model formed by 72 1D Vs profiles. The white points indicate the velocity anomaly. A higher velocity ($\approx$ 250 m/s) zone can be distinguished at the approximate anomaly position. Constant $\rho = 2000$ kg/m^3 was kept fixed during inversion.

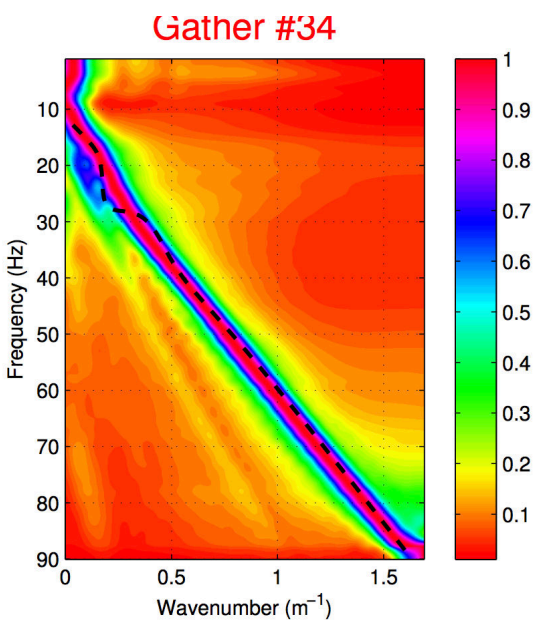


Figure 5.11: Waveform modelling for the final model retrieved with SWA (Figure 5.10). The same gathering process of real data are applied to modelled data. The fundamental mode picked from real data (dashed line) is near to the fundamental mode of modelled data but not centred at its energy maximum.

are less sensitive to density (Xia et al., 1999), and thus we have considered homogeneous profiles with $\rho = 2000$ kg/m^3 (around the values initially provided, Figure 5.1). The final Vs model is shown in Figure 5.10. Below 5 m depth, the retrieved velocities are vertically constant because we chose this limit for the half space in SWA and thickness was not inverted. A higher velocity zone was retrieved by inversion at the approximate anomaly position. The corresponding velocity value (inside the mentioned shape) is in average 250 m/s. This value is low in comparison to the expected value (Figure 5.1). Nonetheless, a velocity contrast around the theoretical anomaly position has been retrieved.

In SWA, dispersion curves are built locally, based on a 1D assumption. This is why contributions from different shots were stacked. To check the validity of the result, we compute the data in the final model and in a fully 2D sense. For that, we perform waveform modelling (Chapter 2). Then, we apply the same gathering process to the modelled data. Finally, the fundamental mode extracted from real data gathers is superposed to the $f - k$ gather of synthetic data. The real data fundamental mode approximately corresponds to the fundamental mode on synthetic data (Figure 5.11).

The Vs model has been determined employing SWA. Vp and ρ were not inverted because surface waves are less sensitive to these parameters. The theoretical anomaly position has been identified by velocity contrasts. Determined velocities are higher inside the anomaly position. However, they are low in comparison to the expected values (250 m/s compared to 900 m/s). This result could be due to a lack of surface wave interaction with the anomaly. However, in order to obtain a more meaningful result, higher modes should be included in inversion. It seem possible to pick them in the $f - k$ gathers (Figure 5.9). Fundamental plus higher modes can better constrain inversion as shown in Chapter 3.

5.3.2 FWI and w-AWI

In this section, we apply waveform inversion techniques (FWI and w-AWI) to the real data set. The goal is to identify a high velocity value at the anomaly position (Figure 5.4). Besides, different frequency bands of data are considered in inversion.

The initial Vp model is obtained by traveltime tomography (Figure 5.8). We performed a first waveform inversion test using the SWA Vs result as initial model, but the velocity at depths below 5 m was too high and surface waves could not penetrate. Besides, there was no clear strategy for selecting appropriate Vs values below 6 m depth. Because of this, we decide using a linear velocity gradient ranging from 50 m/s at the surface to 250 m/s at 12 m depth. In the literature, a linear gradient is considered to be appropriate for FWI when considering surface waves, at least in low-frequency inversion (Schäfer et al., 2013) . The ρ model is a linear gradient ranging from 2000 to 2800 kg/m^3. The previously described initial models are depicted in Figure 5.12.

For all the inversions, a preconditioning factor equal to the depth was applied to the gradient. Besides, the gradient was smoothed using a 1.6 m radius 2D Gaussian window (see Chapter 4 for details). The Lamé parameters (λ and μ) are the inversion unknowns, while ρ is not updated in inversion. The window length used in all w-AWI inversions has been fixed in $l_w = 9.3$ m (equivalent to 32 consecutive receivers).

Contrary to for the tests shown in Chapter 4, a single L-BFGS-B routine was used to

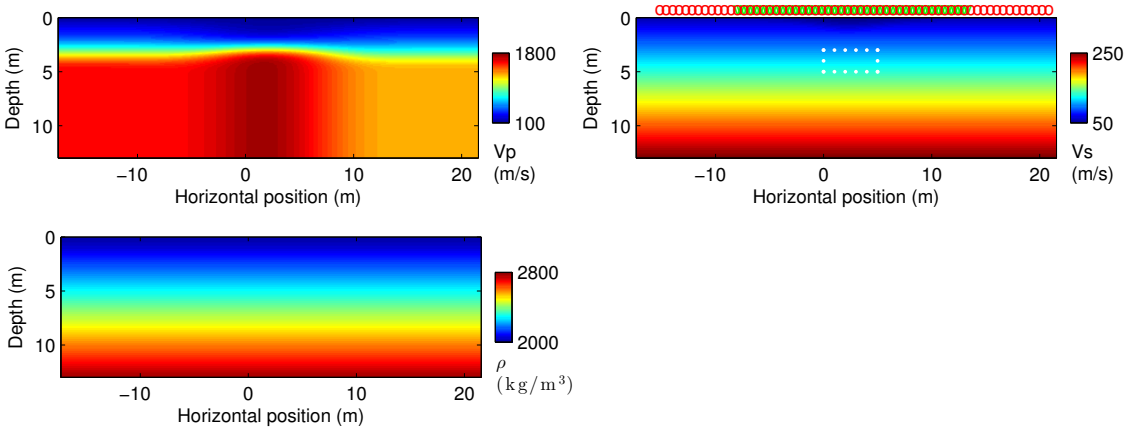


Figure 5.12: Initial Vp (top left), Vs (top right) and ρ (bottom) models for FWI and w-AWI.

compute only one approximate inverse of the Hessian for both λ and μ at each iteration. The objective was to take into account the possible coupling between λ and μ in inversion.

The unknown source wavelet is estimated using the approach presented by Pratt (1999) in the case of FWI, and using the approach described in Chapter 4 for w-AWI. The initial wavelet is a Ricker wavelet with dominant frequency of 25 Hz, except for low-frequency (10-18 Hz) inversion. In this case, a 14 Hz Ricker wavelet is employed.

Filtered data inversion

The results of inverting filtered data (with high-cut frequencies at 18, 25, 36, 50 and 70 Hz) are shown in Figures 5.13, 5.14, 5.15 and 5.16. The Vp final model does not change for depths greater than 5 m for any of the considered frequency bands (Figure 5.13). The Vp update has mostly acted in the shallow part of the model and is principally due to the updates of μ (Figure 5.14).

The Vs final models for the low frequency data (top of Figures 5.15 and 5.16) are smooth and mostly updated with positive values. For the other frequency ranges, the final Vs models contain a pair of high-low velocity layers updated between 0.5 and 3.5 m depth approximately (Figure 5.15). We have initially identified two main reasons for this (two other reasons are given below in the text). The first one is related to the high Vp values starting at 5 m depth. That feature partially prevents modelled waves from penetrating deeper and inversion must explain real waves by including a phantom reflector. The other possible reason can be associated with the very strong contrast between the low velocity unconsolidated top layer and the concrete structure. In this case, the initial velocity model is probably rather far from the objective function global minimum, which leads the inversion algorithms fall into a secondary minimum.

In our opinion, an interesting model update is the one obtained by w-AWI when considering the 10-70 Hz band-passed data. Besides the previously described pair of layers, there is a stronger velocity update at the approximate vertical position of the anomaly (bottom left of Figures 5.15 and 5.16). Horizontally, this update is smeared out from the anomaly position. However, we chose to analyse this result in more detail as it

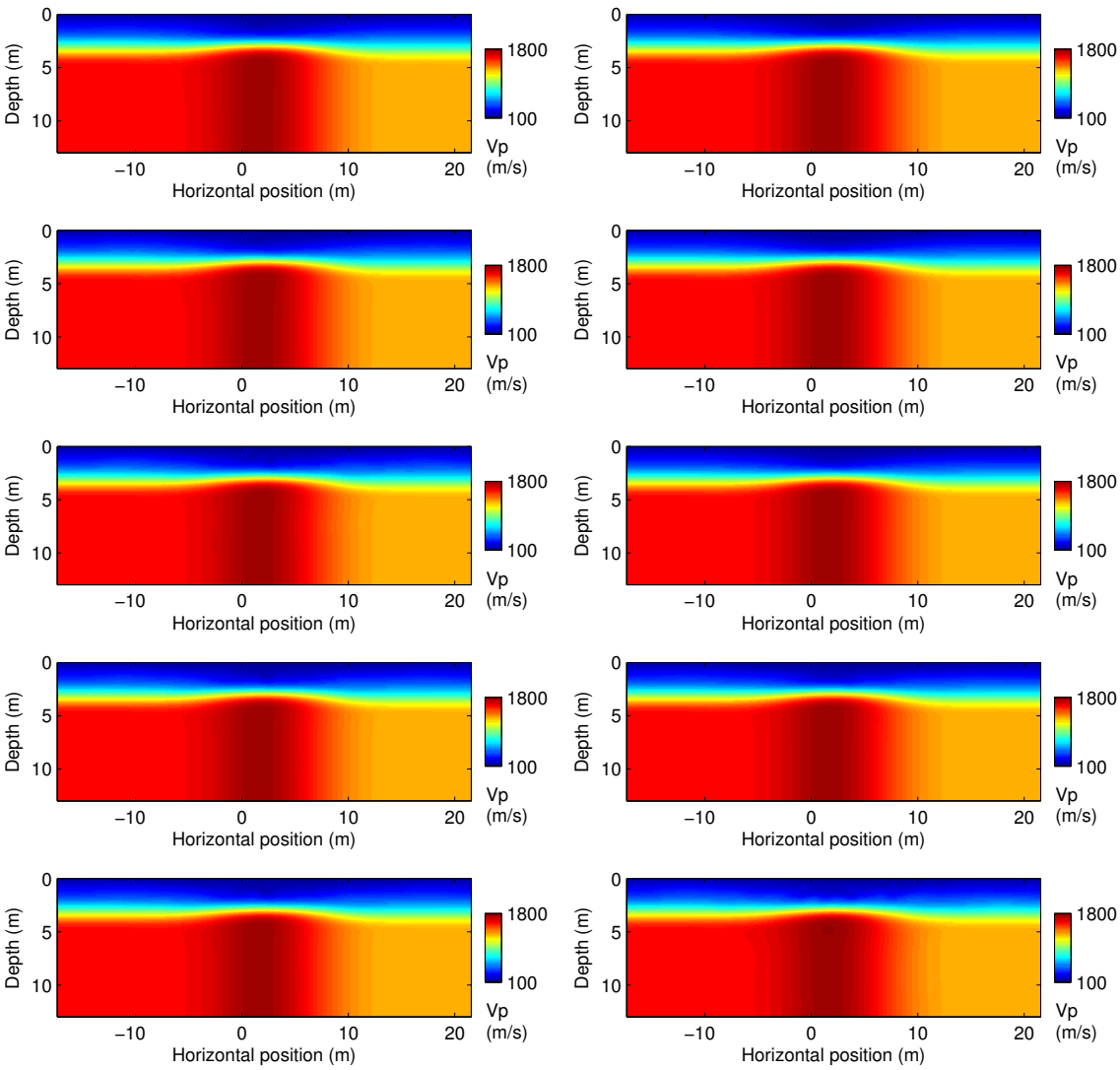


Figure 5.13: Vp *final models* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz.

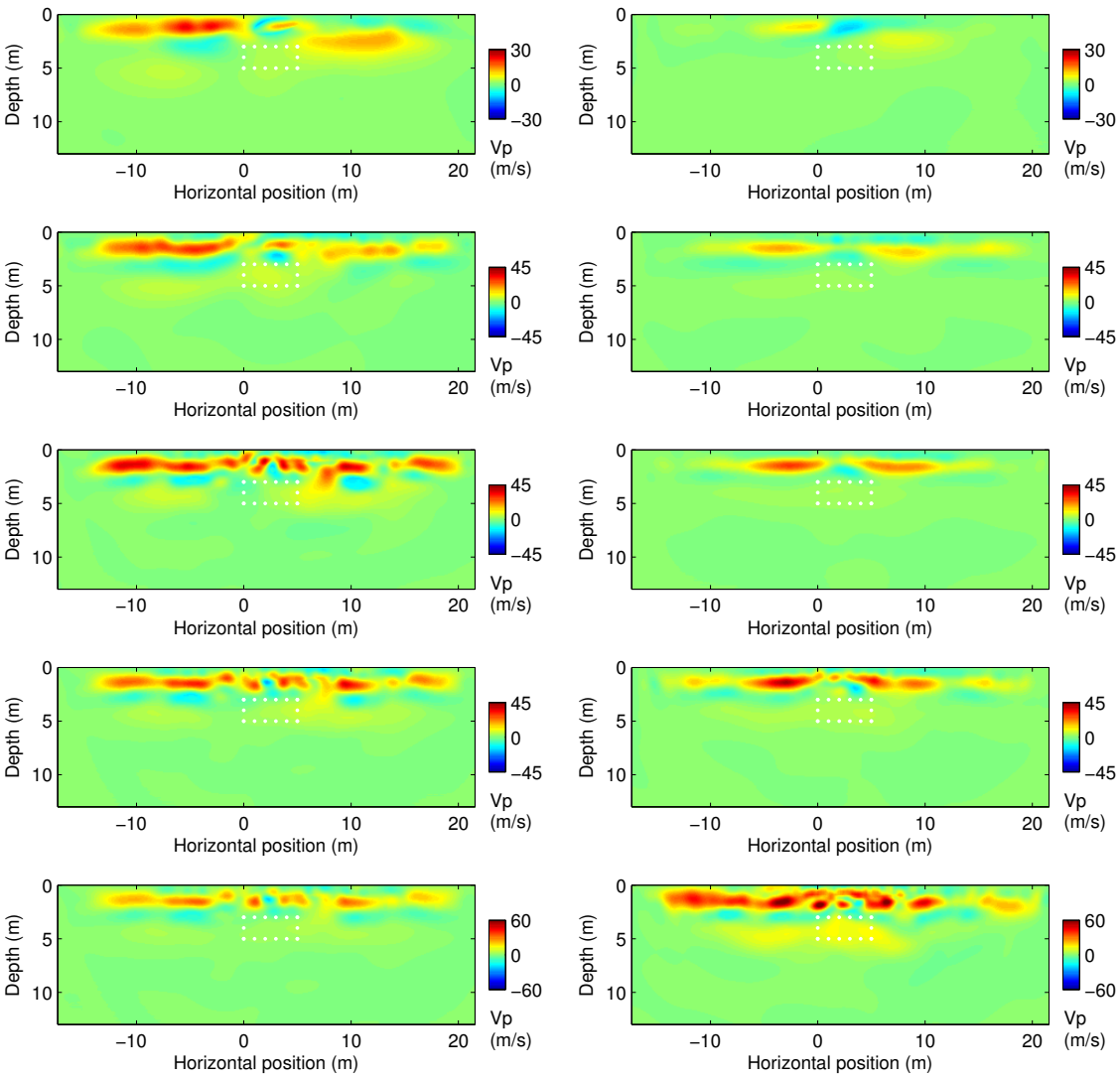


Figure 5.14: Vp *model updates* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz.

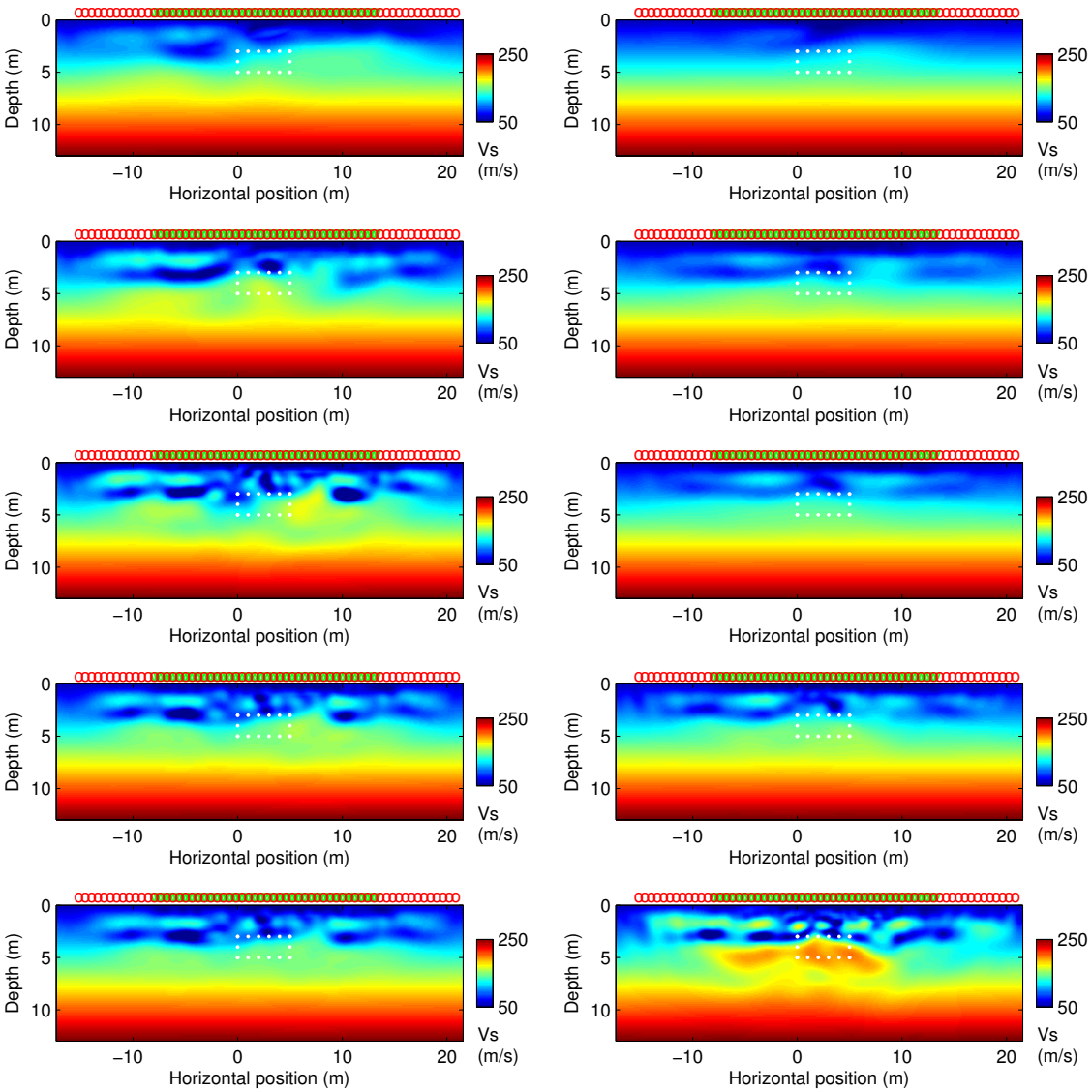


Figure 5.15: Vs *final models* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz.

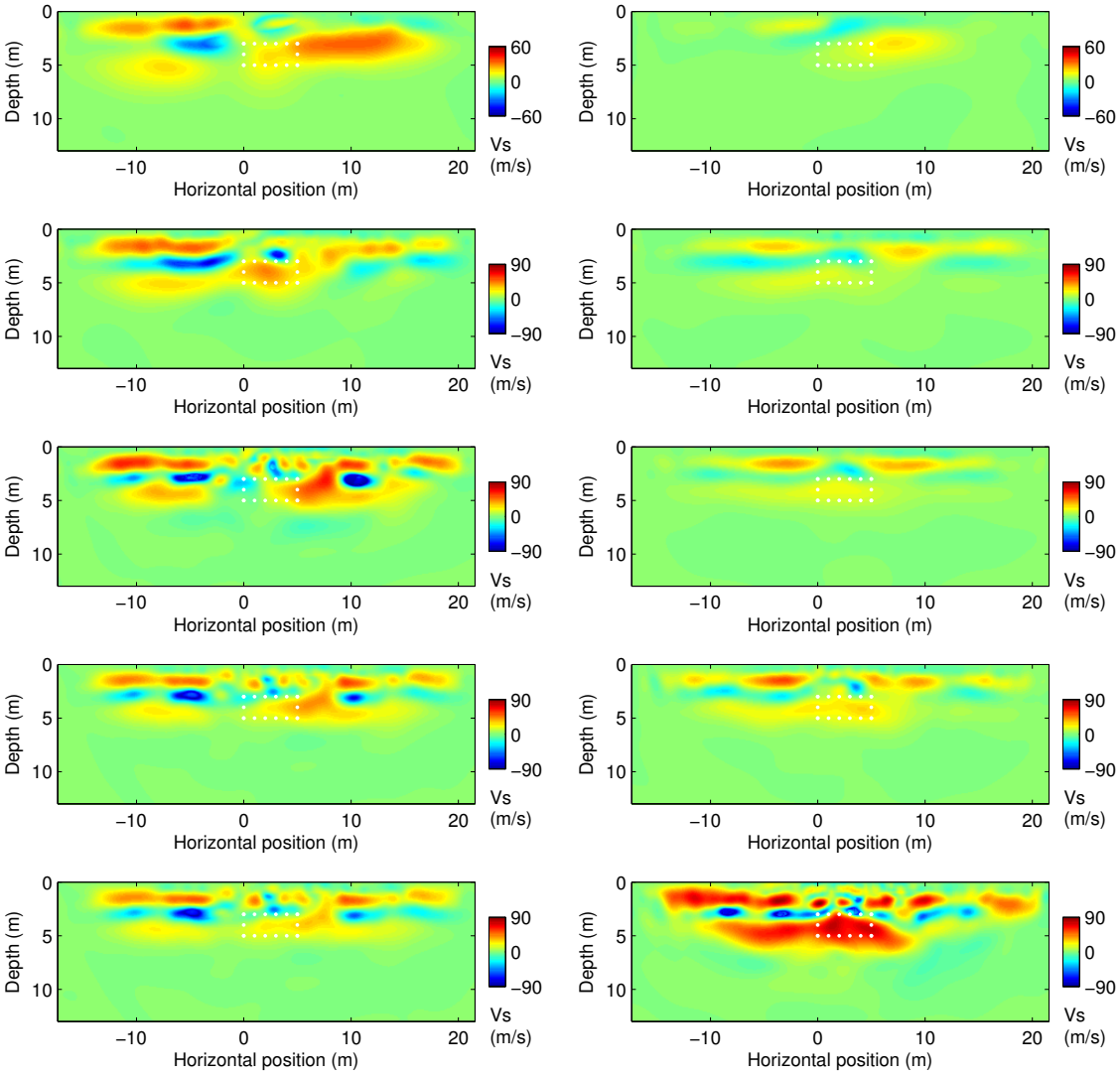


Figure 5.16: Vs *model updates* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz.

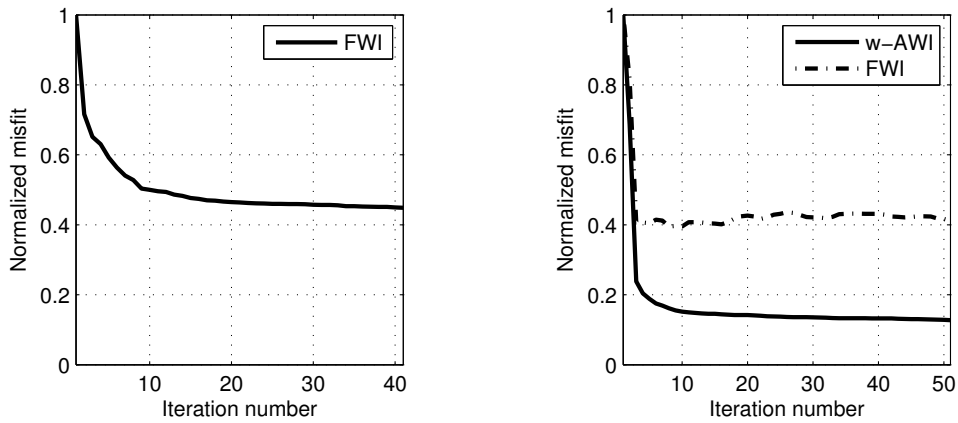


Figure 5.17: Objective function evolution with iterations for the 10-70 Hz bandpass data. Left: FWI results. Right: w-AWI was employed, but the FWI objective function computed through w-AWI iterations is also shown.

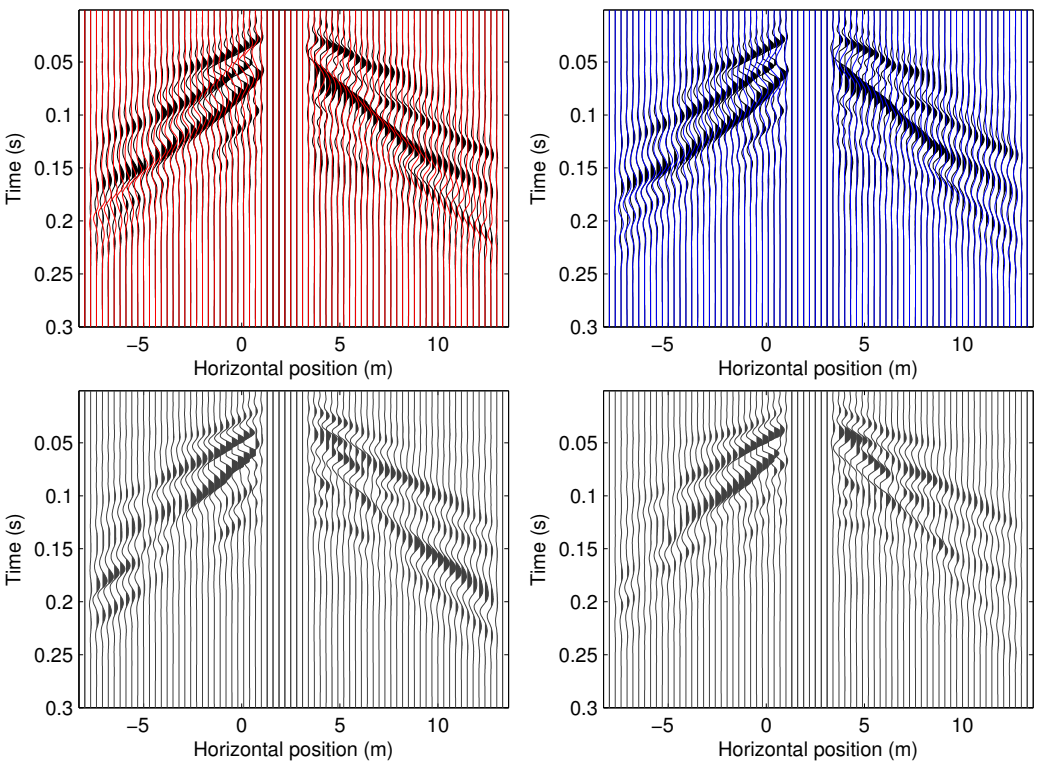


Figure 5.18: Representative result (shot gather 30) for FWI with 10-70 Hz band-passed data. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals.

is the one that best defines the anomaly position.

The misfit value minimisation is shown in Figure 5.17. On the right, the w-AWI and FWI misfit values are shown when using w-AWI for the model update. On the left (Figure 5.17), only the FWI misfit is shown when using FWI. The minimisation is apparently more effective with w-AWI. This is however not an indicator of better results in this case because the data phase is neglected in w-AWI. We observe that the data misfit is approximately the same (around 0.42) when employing FWI or w-AWI.

We study the data in order to better analyse this result. The FWI result for one of the shot gathers (30) is displayed in Figure 5.18. The quality of initial data (in red) is good with respect to observed data thanks to a pre-inversion wavelet estimation. At the end of iterations, the synthetic data (in blue) correctly mimics the observed data. This result principally explains surface waves. Regarding first arrivals, the result is not the same (almost no first arrivals are explained). The missing Vp update is one of the causes for this, besides the unknown and not updated density model. Nevertheless, the residuals confirm the surface wave reconstruction (bottom of Figure 5.18).

In the case of w-AWI, the data are shown in Figure 5.19 for a spatial window including negative offset receivers, and in Figure 5.20 for positive offsets. In these figures, the fundamental mode (most energetic one) appears to be correctly explained. However, higher modes are not completely explained. In the spectra slice comparison, similar features are displayed between observed and final data. The misfit minimisation (Figure 5.17) is confirmed by the comparison of residual data (Figures 5.19 and 5.20). Furthermore, the real data negative offset window (Figure 5.19) shows an uncommon fundamental mode. Fundamental and higher modes seem to be mixed at approximately 35 Hz and 0.1 m^{-1}. This can be due to the real medium configuration, or this could be a third reason for the appearance of the pair of layers previously described. Indeed, interpolation and the presence of noise could have changed the spectrum and masked or destroyed some of the useful information.

Common shot gathers obtained with w-AWI are shown in Figure 5.21. The initial data are different than in FWI (Figure 5.18) even if the initial model is the same because a source wavelet is estimated before starting inversion. The final data are not in correct phase with real data because w-AWI partially neglects the phase. Some information about the phase is reintroduced by using horizontal windows (see Chapter 4 for more details). However, the frequency amplitude is well explained as shown previously (Figures 5.19 and 5.20).

This result (w-AWI, 10-70 Hz band-passed data) can be considered as an initial model for a supplementary FWI run. The Vs and Vp models are smoothed before the supplementary FWI run in order to reduce some of the oscillatory effects caused by the shallow pair of layers (between 0.5 and 3.5 m). An elliptical (2.5 m vertically and 5 m horizontally) Gaussian window was employed for this purpose. The source wavelets found after w-AWI are the initial ones for FWI.

We first show the objective function minimisation (Figure 5.22). At the end of w-AWI, the measured misfit value (measured with FWI objective function) was approximately 42 % of the initial value; now by applying FWI in a second inversion run the misfit value was additionally reduced to 70 % of the w-AWI result. This roughly means that by applying this two stage (w-AWI then FWI) inversion the data misfit was minimised to

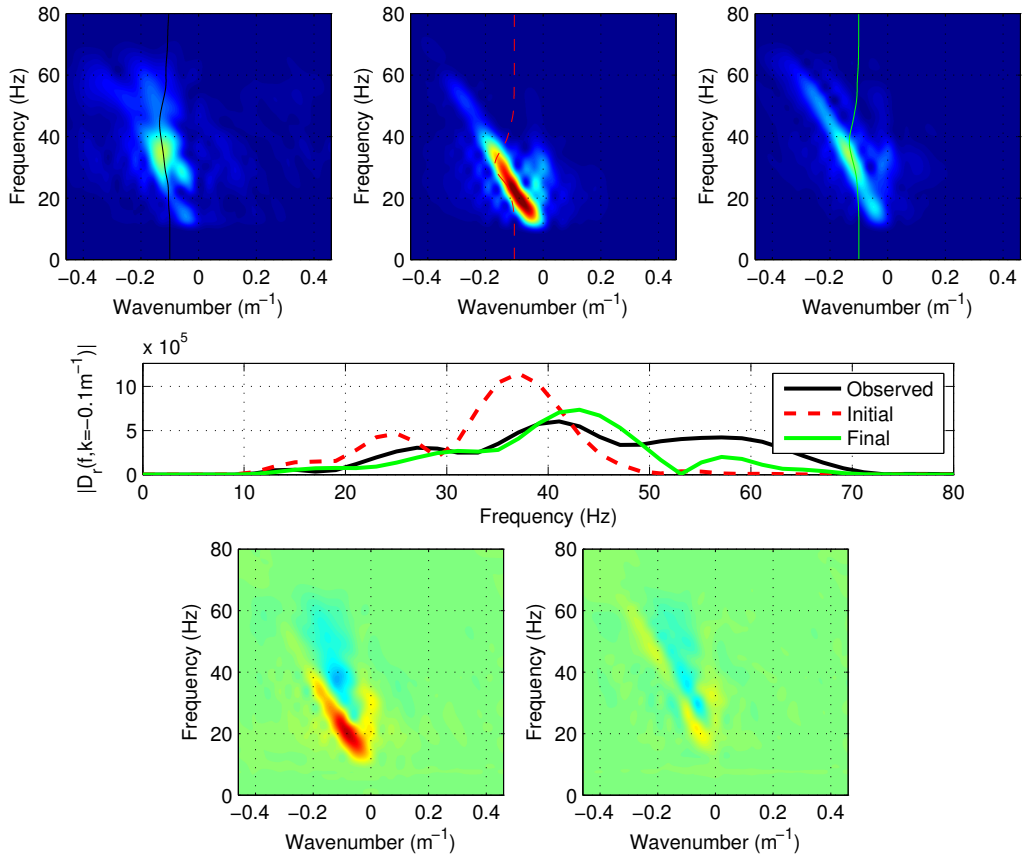


Figure 5.19: Representative result (shot gather 30, negative offset window) for w-AWI with 10-70 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Middle: spectra slice comparison. Bottom: initial (left) and final (right) residuals.

28 % of the initial value. This result can be observed in Figure 5.23, where initial (after w-AWI) and final data are shown. The final data seem indeed closer to real data than in the previous tests. Besides, the final data residual have also smaller amplitude. Finally, one of the effects of w-AWI can be observed in the source wavelets (Figure 5.24). After the first stage (w-AWI), the wavelets are in phase. After the second stage (FWI), the wavelets are shifted. This phase shift could be related to high velocities in the model which are compensated for by FWI with wavelet delays.

The velocity model obtained after the two stage inversion (Figure 5.25) seems to explain features of the real structure (Figure 5.4). However, the two abnormal shallow layers described in the first stage result (with w-AWI) are still present. Nevertheless, their shapes are smoother after the second stage (with FWI).

Local minima in w-AWI

The results of two tests applied to data filtered with higher low-cut frequencies are shown in this section. Instead of leaving the initial 10 Hz low-cut frequency (from preprocessing, Figure 5.6), we have employed low-cut frequencies of 25 and 36 Hz. We only applied w-AWI to this data to evaluate the effect of missing low-frequencies.

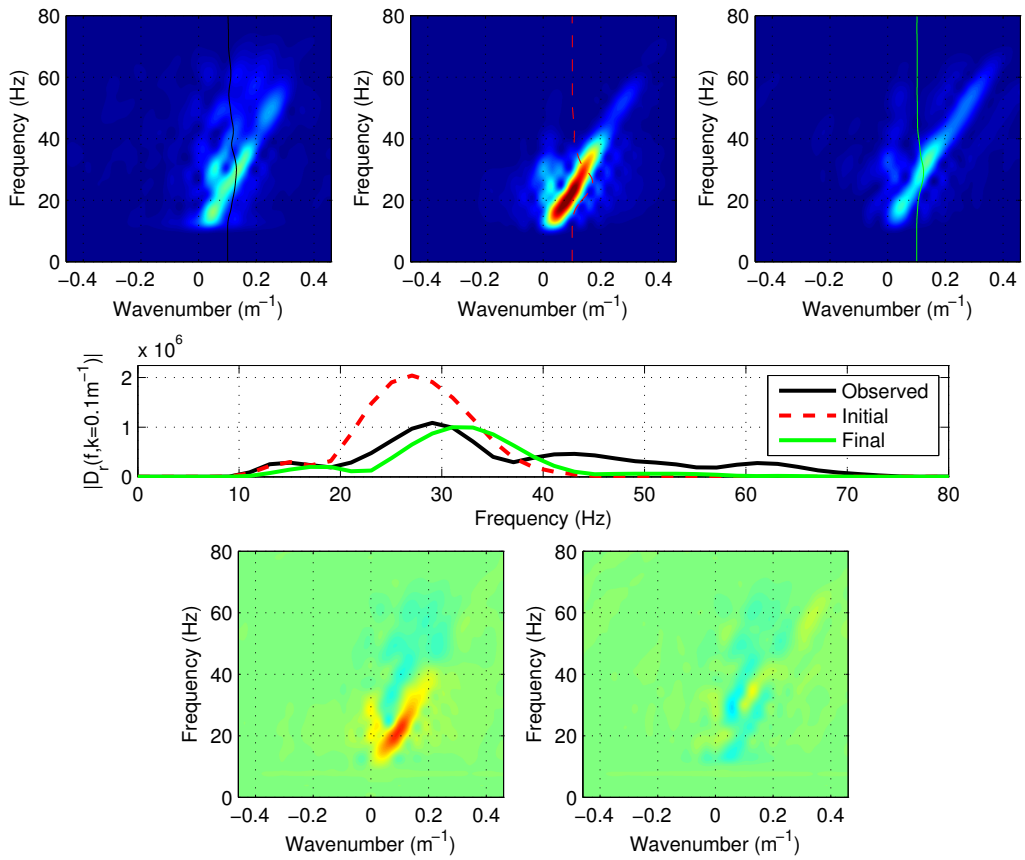


Figure 5.20: Representative result (shot gather 30, positive offset window) for w-AWI with 10-70 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Middle: spectra slice comparison. Bottom: initial (left) and final (right) residuals.

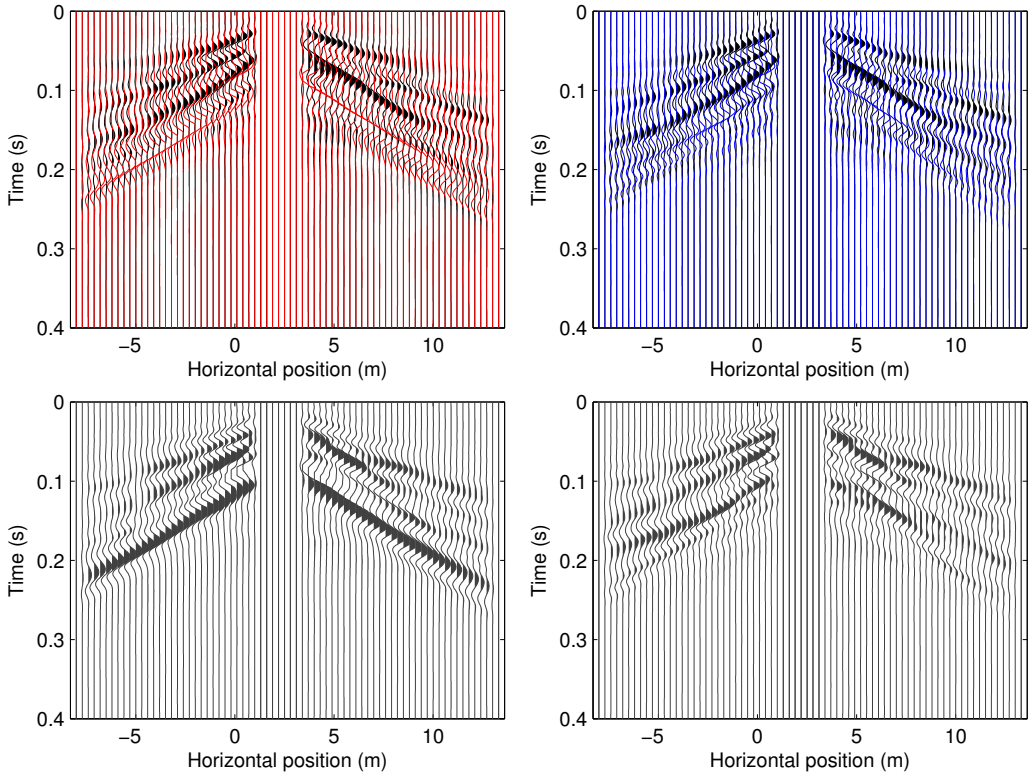


Figure 5.21: Representative result (shot gather 30) for w-AWI with 10-70 Hz band-passed data. The differences are mainly related to the phase. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals.

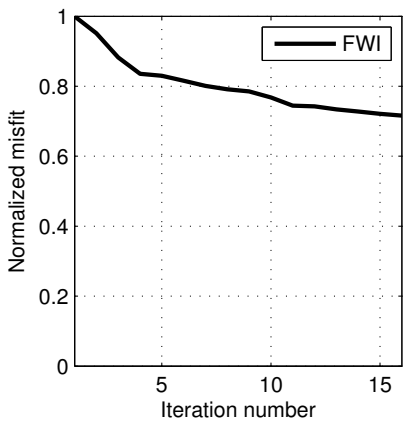


Figure 5.22: Objective function evolution with iterations. This is the second stage of inversion (FWI after w-AWI).

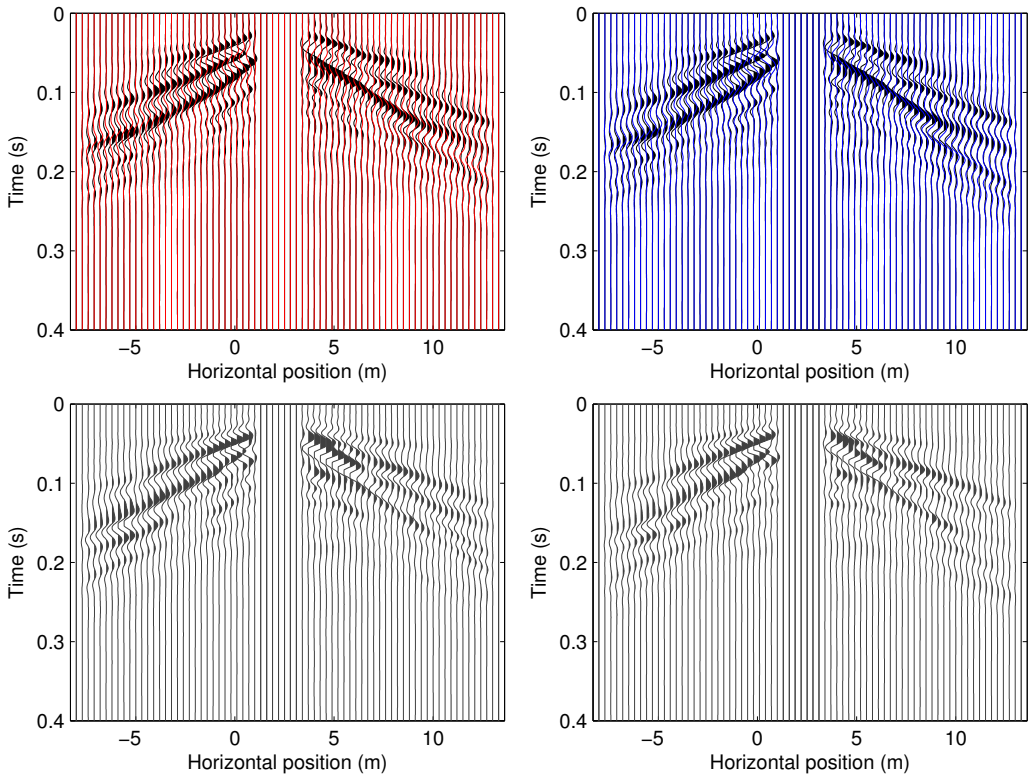


Figure 5.23: Representative result (shot gather 30) for FWI with 10-70 Hz band-passed data after w-AWI. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals.

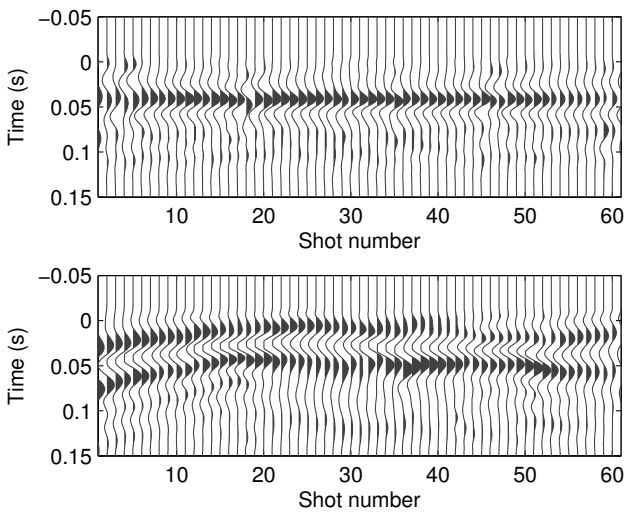


Figure 5.24: Source wavelets found after w-AWI (top) and w-AWI + FWI (bottom) inversion. Wavelets found by w-AWI are not shifted mainly because the phase is not inverted.

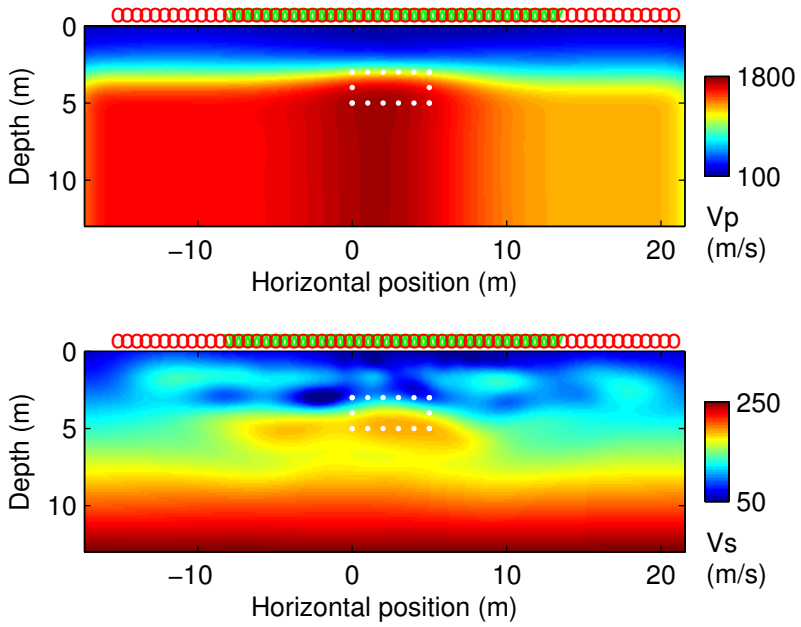


Figure 5.25: Second stage inversion result (FWI after w-AWI). Final result for Vp (top) and Vs (bottom).

In this case, a high velocity layer is created in the shallow part of the model after inversion (results in Figures 5.26 and 5.27). The usual low-velocity layer retrieved by inversion starting at 10 Hz (previously referred to as pair of layers) is not present in these results. Regarding the update at the anomaly position, the 25-80 Hz inversion has tried to do it correctly although at an erroneous vertical position. On the contrary, the 36-80 Hz inversion was not successful in updating the anomaly velocity.

The misfit was minimised to lower values in comparison to low frequency inversions (less than 0.1 compared to 0.15, Figure 5.28). However, the surface wave fundamental mode was only partially reconstructed (Figures 5.29 to 5.32). Besides, for the higher frequency band (36-80 Hz), the fundamental mode of synthetic data seems to be erroneously explaining one of the real data higher modes. Consequently, very shallow high velocity layers had to be created in order to match the synthetic fundamental mode to with one of the real data higher mode. This is a clear effect of local minima created by "cycle skipped" modes in $f - k$ gathers.

5.4 Discussion

We have applied SWA, FWI and w-AWI to a 2D real seismic profile at ultra-shallow depth with 61 shot gathers (containing 72 traces each). The data ware acquired above a purpose-built ground model, where a concrete anomaly was casted after digging a pit. The data quality is in general good, although some preprocessing stages had to be applied, especially for solving coherency problems between adjacent traces. Besides, the data were high-pass filtered by the *geophones* at around 10 Hz. This value is not constant for all the receivers because three different types of geophones were used during acquisition. The

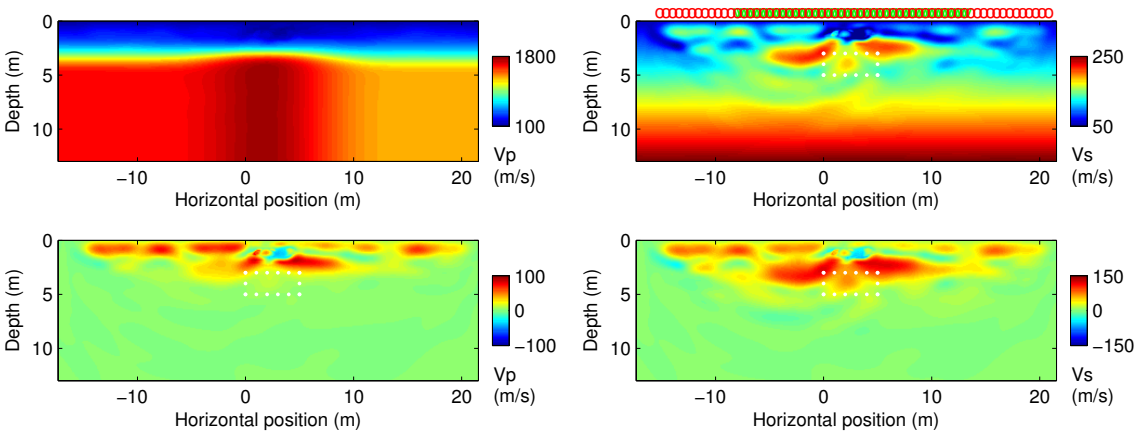


Figure 5.26: Result of w-AWI with 25-80 Hz band-passed data. Top: Vp (left) and Vs (right) final result. Bottom: final update for Vp (left) and Vs (left).

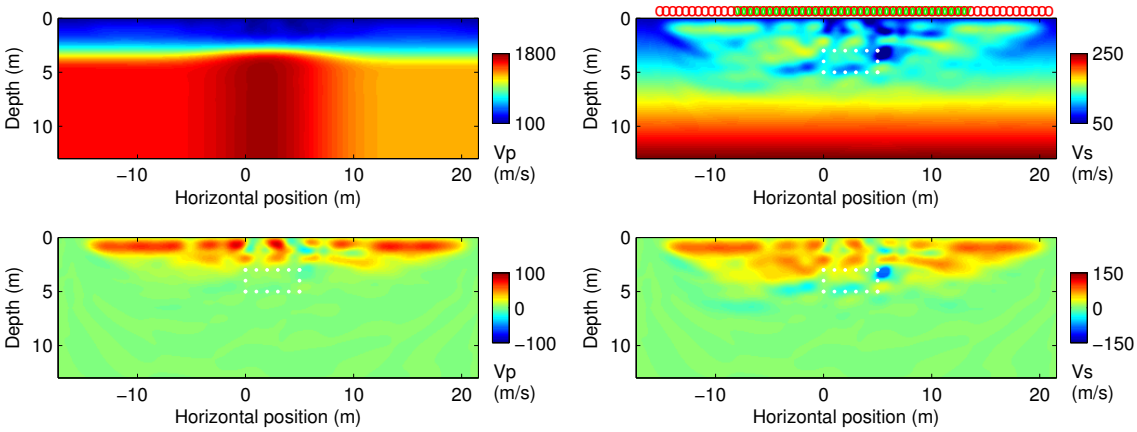


Figure 5.27: Result of w-AWI with 36-80 Hz band-passed data. Top: Vp (left) and Vs (right) final result. Bottom: final update for Vp (left) and Vs (left).

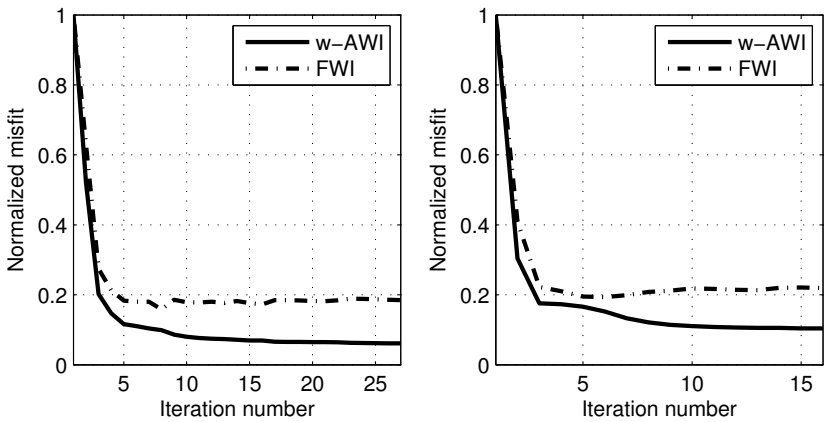


Figure 5.28: Objective function evolution with iterations. Left: 25-80 Hz. Right: 36-80 Hz. Only w-AWI was employed, but both w-AWI and FWI misfits are displayed.

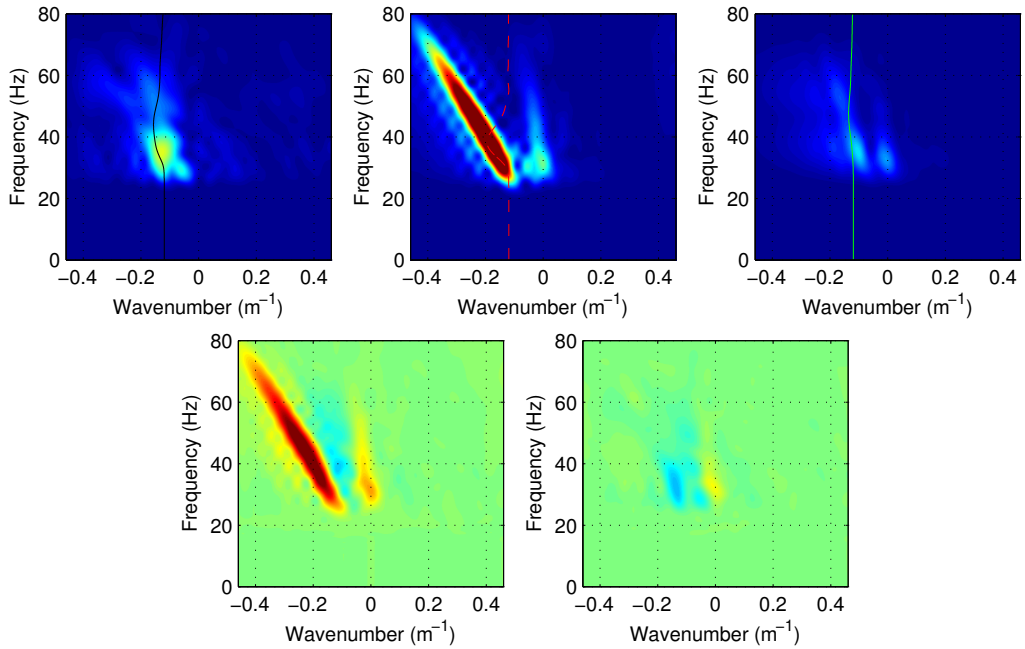


Figure 5.29: Representative result (shot gather 30, negative offset window) for w-AWI with 25-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals.

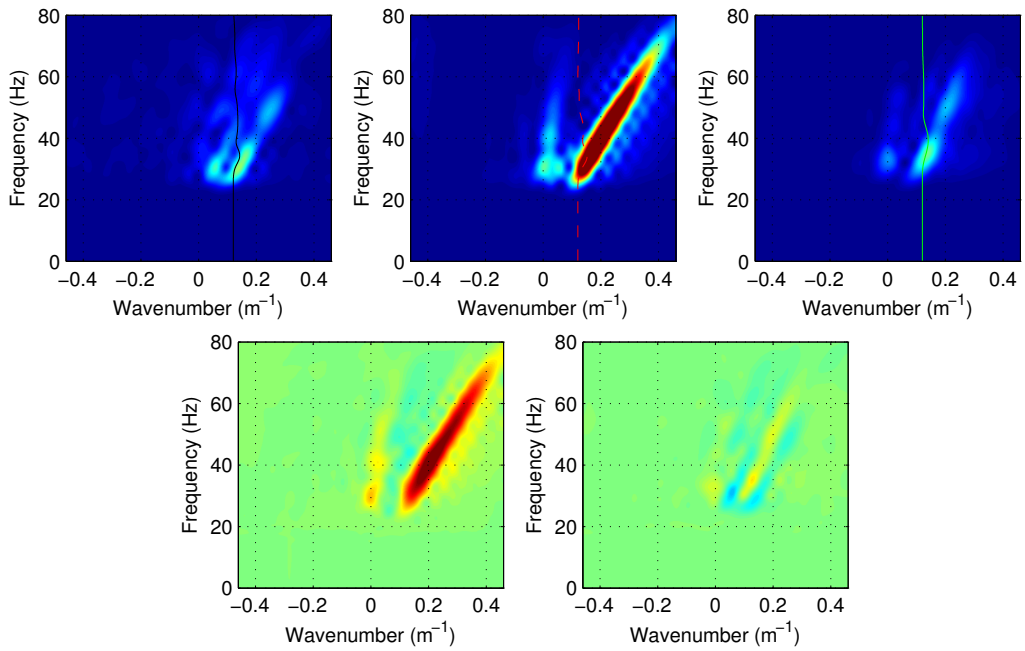


Figure 5.30: Representative result (shot gather 30, positive offset window) for w-AWI with 25-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals.

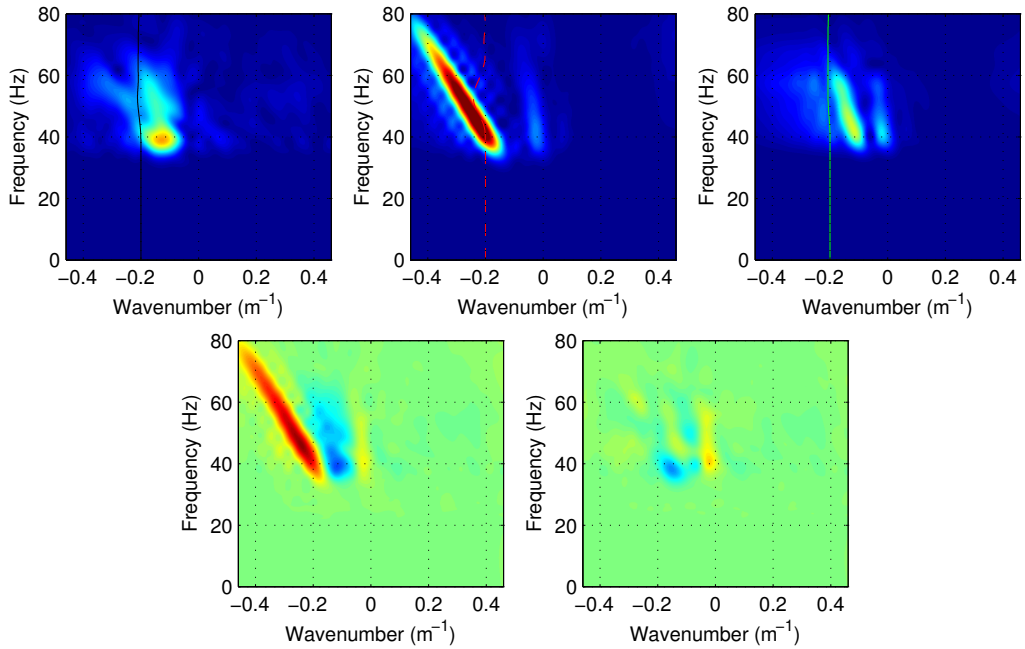


Figure 5.31: Representative result (shot gather 30, negative offset window) for w-AWI with 36-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals.

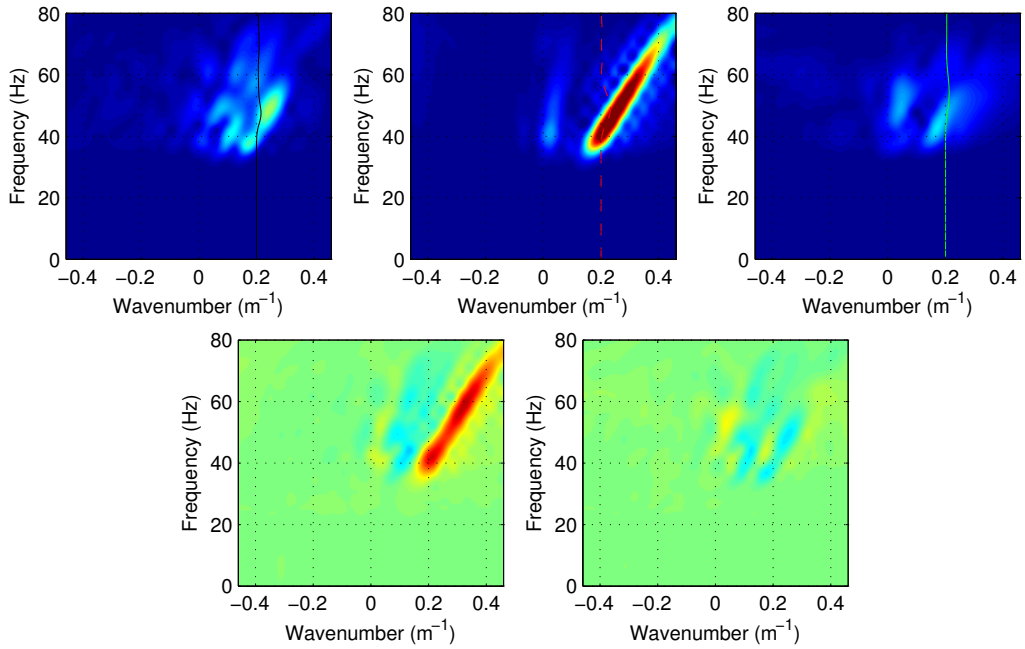


Figure 5.32: Representative result (shot gather 30, positive offset window) for w-AWI with 36-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals.

two nearest offset traces have been muted in preprocessing to avoid introducing spurious or distorted events.

As initial model, we considered a smooth Vp model inverted via first-arrival traveltime tomography. Some information about Vs and ρ is given by Deidda and Balia (2001). However, below 5 m, none of the elastic parameters are known. Besides, the concrete structure was placed around 13 years ago, and this, in addition to the sedimentary nature of the geological context, gives a reason to interrogate ourselves about the current anomaly position. The initial model could be improved by considering low-frequency FWI to wide aperture data (Gao et al., 2007; Plessix et al., 2012). However, the data set considered here mainly contains surface waves and early arrivals (transmitted waves) which are not sufficiently separated from other seismic waves.

SWA

We have applied a simple SWA inversion strategy using the picked fundamental mode. The dispersion curves were semi-automatically extracted from stacked $f-k$ gathers. This inversion was useful for us to better understand the data. The SWA analysis allowed us to identify a Vs = 250 m/s structure at the approximate anomaly position. We believe this result can be improved if higher modes are included in inversion; they can be distinguished and picked in the $f-k$ gathers. This new information could improve the anomaly resolution because higher modes penetrate deeper. For example, for a frequency of 20 Hz (contained in the data) the fundamental mode penetrates approximately up to 4 m (considering an average of Vs = 80 m/s in the first two meters). As the anomaly is placed between 3 and 5 m depth, we should be able to improve results in depth only by extending the fundamental mode penetration by a factor of 1.25.

FWI and w-AWI

Concerning waveform inversion, we have applied FWI and w-AWI to different frequency bands of data: 10-18 Hz, 10-25 Hz, 10-36, 10-50 and 10-70 Hz. With the lower frequency band, a smooth result was obtained with both techniques, but the anomaly position could not be defined. For all other intermediate frequency bands, a pair of high and low velocity layers was imaged by inversion in the near surface (0.5 to 3.5 m depth). We have identified four main reasons for this effect:

- High Vp values starting from 5 m depth. Our strategy for complementing the smooth 5 m thick Vp model (from traveltime tomography) consisted in extending the velocity to greater depths. This, however, might not correspond to reality because there should be a second velocity contrast below the anomaly, at around 5 m depth.

- High velocity contrast in the real medium. The interface between the sediments and the concrete is likely to create strong reflections that impose more constraints on the choice of initial velocity model.

- Noisy traces as well as data interpolation might have created the unclear fundamental mode observed in the negative abscissa part of the profile (Figure 5.19).

- Non-modelled physical aspects. As a final reason, we consider that anisotropy, attenuation and 3D effects are not modelled in our codes.

The integrated inversion process including a first w-AWI run followed by FWI has recovered a high velocity body ($\approx$ 200 m/s) near the approximate anomaly position. Besides, low velocities were imaged below this anomaly shape. The near-surface transitions appear smooth. Therefore, this result seems to explain some of the ground-model features.

5.5 Conclusion

Contrary to FWI, the best w-AWI strategy does not consist in starting inversion from the low-frequency content and then gradually including the higher frequencies. The useful information for w-AWI is the dispersion of surface waves. This technique needs to take into account the complete range of frequencies contained in the data. This is the reason why time domain modelling is necessary in w-AWI. In SWA, the dispersion curves must be sensitive to all the layers in the model. As w-AWI objective function is similar to SWA (in the sense of dispersion analysis), the complete data spectrum has to be included. This was shown in this chapter by inverting low to high frequency bands. The best result was obtained when considering the entire spectrum. Nonetheless, low-frequency content is also needed, as without it a high-velocity superior mode could erroneously be explained at high frequencies by the low-velocity fundamental mode.

Chapter 6

Conclusions and Perspectives

6.1 Conclusions

Surface Wave Analysis (SWA) can be used to determine elastic properties of the Earth at two different scales to image the near surface (less than 100 m) or the Earth's deep interior (crust and mantle). However, surface waves are rarely used in the intermediate scale, for example, to improve seismic imaging at exploration target depth (less than 10 km). Their propagation remain in the near surface and far from the target. They are thus often considered as coherent noise. Instead, body waves are preferred in seismic exploration techniques. In this PhD thesis, we have tried to see if the concept of SWA can be used to modify a technique commonly applied to body waves, Full Waveform Inversion (FWI), in order to characterise 2D media with surface waves. As a result, we have proposed an intermediary approach between SWA and FWI that we call windowed-Amplitude Waveform Inversion (w-AWI). This research focused on three main aspects. The conclusions of our study are presented as follows.

6.1.1 Waveform modelling

As a first matter of research, we implemented elastic wave propagation modelling with free-surface conditions in 2D elastic isotropic media (Chapter 2). The classical velocity-stress formulation was revisited to make it work in the presence of smoothly curved interfaces and topography. A computational cost analysis indicates that the chain rule approach is theoretically less costly than the tensorial approach if finite-difference stencils of less than the sixth order are employed. As one of our goals concerns the implementation of a modelling tool computationally efficient for FWI, we implemented the chain-rule approach with second-order finite differences. Our solution is implemented in the time-space domain because in this way surface wave dispersion can be modelled with less computational cost compared to frequency-space domain waveform modelling approaches.

We propose employing a modified staggered grid which makes the computation of partial derivatives to be straightforward within the chain-rule modelling framework. The accuracy of the proposed modelling scheme was evaluated by comparing results to an analytical solution (Cagniard-de Hoop method) showing high accuracy at least for Rayleigh and direct waves in homogenous media. The computational cost remains approximately

twice as large as the one corresponding to to the classical rectangular domain implementation. Nonetheless, wave propagation can be modelled without grid artefacts in smoothly curved 2D media.

6.1.2 Analysis of inversion techniques

The characteristics of SWA have been investigated. This research started with a bibliographic review followed by an analysis of synthetic data examples (Chapter 3). Dispersion curves representing surface wave propagation in plane-layered media is the essential element used in SWA. In dispersion curve inversion, even if the computational cost of the differential eigenvalue direct problem (secular function approach) remains low (allowing testing inversion results several times), searching the damping factor that guarantees both convergence through objective function minimisation can be a cumbersome task. Thus, we showed that the classical Levenberg-Marquardt method used in SWA can be improved in terms of convergence if the damping factor changes dynamically with iterations instead of being a fixed factor chosen on a trial-and-error basis. We have proposed a heuristic approach in which the damping ratio and residual weighting are chosen automatically according to the minimisation of the objective function at each iteration. In addition, a Tikhonov regularisation could also be considered (second-order Laplacian operator, weighting covariance matrix, etc).

The exact 1D Vs model can be found by inverting dispersion curves if accurate initial Vp and density models are provided, as well as a good approximation of the expected layer thicknesses. Moreover, inversion can find a model explaining the dispersion curves, *even completely*, but not corresponding to the exact model. In such a case, dispersion curves are ambiguous data as they can be explained by different models satisfying the same data. In some cases, even the use of higher mode incorporation fails to better constrain inversion. Therefore, more information should be incorporated to this surface wave inversion approach. This is one of the motivation for having analysed alternative approaches for surface wave inversion such as FWI (Chapter 3).

In principle, surface waves could be inverted by FWI directly from seismic shot gathers. This approach seems however difficult to apply successfully because high frequencies of surface waves increase the presence of local minima in the objective function. The global minimum basin becomes smaller which makes it difficult to choose an appropriate initial velocity model. When using FWI with body waves, starting inversion with low frequencies can retrieve a preliminary velocity model that explains the long to intermediate wavelengths. However, it is still not clear if such a strategy would guarantee a good result in the case of dispersive surface waves. Hence, the combination of SWA-type dispersion analysis and waveform modelling was investigated to find a waveform inversion approach especially suited for surface waves.

Separation of shot gathers into subsets of receivers was initially analysed with the purpose of developing a better understanding of surface wave dispersion in 2D media. It was observed that surface waves are dispersive only if there are model heterogeneities within a depth range smaller than two wavelengths. This result is already demonstrated in SWA. Furthermore, by separating the shot gathers into subsets and then analysing the absolute value of their 2D Fourier transform, we realised that the dispersion "observed"

in each subset of receivers is principally due to the medium properties localised below the position of each subset. This initial observation prompted our interest in modifying the classical FWI objective function in order to measure the misfit like in SWA but in 2D media and without picking dispersion curves.

Closed-form equations for gradient computation in the specific case of the velocity-stress formulation for 2D media using the adjoint-state method has been provided as well (given in the FWI theoretical review in Chapter 3). We applied a change of variables in order to make wave propagation auto-adjoint in the sense of the adjoint-state method which is practical because the same code can be used for propagation and back-propagation of residuals.

6.1.3 A novel approach

In Chapter 4, the theory of an alternative surface-wave inversion approach, w-AWI, is developed and then analysed through several numerical tests. The use of windows that select consecutive receivers has been introduced in this work. The classical FWI objective function has been consequently modified in order to include the previously mentioned observations (spatial windows, 2D Fourier transform, absolute value).

Implementing a different objective function evidently implies that its gradient is different than that of classical FWI. Nonetheless, we have used the adjoint-state method for gradient computation which only differs from classical FWI in the residual source for the adjoint-state variable.

The incorporation of windows and the absolute value lead to a wider global minimum basin compared to classical FWI. Synthetic tests using mono-frequency sources show that the w-AWI global minimum basin is as wide as the FWI one if the frequency considered in the latter is half the one considered in w-AWI. This feature of w-AWI is essential in surface wave inversion as it is automatically translated into a relaxation of constraints regarding the choice of initial velocity model. We proved this characteristic in the presence of surface waves but we did not consider the case of body waves (an analysis especially appealing in the case of reflections). Besides, the basin of the global minimum can potentially be tuned by including a smaller number of receivers (shorter windows). Lost of resolution in the $f-k$ domain is prevented by imposing a minimum number of receivers in each considered window. Conversely, too wide windows may result in unwanted property averaging.

The proposed technique w-AWI has been evaluated through synthetic data inversion tests and has been applied to a real seismic data set to image the shallow depths. We can summarise our results as follows.

- Provided that there is an initial velocity model that is close enough to the exact one (within the global minimum basin), both techniques can potentially find global minimums that are closely related. Resolution of w-AWI is principally controlled by the spatial window shape. We have qualitatively observed the change of resolution in the gradient computed with different window lengths. The number of receivers included in each window can be changed if needed in each inversion test.

- The capacity of w-AWI to find the correct model when the source wavelet is unknown has been proven. Formulation for source wavelet estimation has been pro-

vided in this work. We have tested it in the case the initial velocity model is either a smooth version of the exact one (an smooth initial model close to exact one) or a simple linear velocity gradient. The correct velocity model was retrieved in both cases.

- In some configurations, w-AWI correctly converges when FWI does not. This was observed on both synthetic and real data inversions (real data was acquired above a purpose-built ground model, thus with a good idea of the underground structure). This effect is mostly due to its wider global minimum basin. Such minimum basin is approximately twice as wide as for classical FWI. Local minima effects have been observed when high-pass filtering the real data. In w-AWI, a superior mode of observed data may erroneously be explained by the fundamental mode of modelled data if low-frequency data content is not included.

The proposed approach can be used to retrieve 2D properties of the near surface using surface waves. Wider frequency bands are needed in order to perform a complete dispersion analysis in w-AWI. By inverting wider frequency bands, localised features like fluid parameter variations could be estimated since higher frequencies can improve the imaging resolution especially needed in time-lapse reservoir monitoring (4D seismic) (Asnaashari, 2013). Synthetic tests have demonstrated this along with a real data inversion. An additional inversion strategy has been introduced by spatially windowing the data. Real data inversion has shown that w-AWI is likely situated somewhere between low-frequency transmitted-wave FWI and full spectra FWI. An accurate Vs model can be inverted from surface waves by exploiting their larger sensitivity to this elastic parameter. However, if a sufficiently accurate Vp (and also density) model is not provided, w-AWI has to explain surface wave propagation by creating lower or higher velocity layers in the Vs model. Besides, the presence of noisy seismic traces is still a matter of research. For this reason, multiple perspectives are envisaged.

6.2 Perspectives

Surface waves carry information about the near surface properties that have influenced their propagation. Some strategies can be explored in order to extend the results presented here.

6.2.1 Multi-scale strategies

The spatial windows defined in w-AWI can be used in different ways. It is possible to make the basin wider by including less receivers into each windowed subset. By doing this, resolution of the velocity model will be impacted. Thus, a smooth velocity background will result in this case. If this first step is successful, then the number of receivers per window can be increased. However, this number of receivers should be related to the target dimensions as averaging effects can decrease resolution in the end.

As a second strategy, multi-frequency and $f - k$ masking can be explored. Low to high frequency inversion strategies are usually incorporated in elastic and acoustic FWI

(Bunks et al., 1995; Sirgue, 2003; Sirgue and Pratt, 2004; Brossier et al., 2008; Romdhane et al., 2011). Surface waves are in general more sensitive to Vs (Xia et al., 1999; Song et al., 1989), however the sensitivity analysis could be revisited by considering low and high frequencies separately in addition to considering also the energy of dispersion modes. Moreover, masking can be applied to different zones in the $f-k$ gathers of w-AWI. This is a novel feature introduced by w-AWI. FWI usually includes masking in time and frequency but not in the $f-k$ domain. As an example, once the fundamental mode have been explained, $f-k$ masking could be used to focus inversion on higher modes. As well, dynamic weighting of specific zones in the $f-k$ domain could be applied similarly to SWA, where weighting is applied to the misfit between dispersion curves in the $f-v$ domain. In such a way, low weights could be assigned to unwanted high energy effects in the fundamental mode (effects due to noise or acquisition problems) in inversion.

As a third option, we suggest a workflow integrating w-AWI in this order: traveltime tomography, low-frequency FWI, w-AWI and full spectra FWI. First, tomography can be used to obtain a long-wavelength velocity model that will be the starting model for low-frequency FWI. FWI is then applied to long-offset wide-aperture seismic events (diving/transmitted waves) (Gao et al., 2007; Plessix et al., 2012). This second step leads to higher resolution in Vp and density. Then, w-AWI could be applied to full frequency data in order to find an accurate Vs model from surface waves. Finally, full band FWI can be applied leading to reconstruction of the missed data phase while improving resolution of the model. Results of this strategy could be validated in locally 1D media by comparing results to SWA and evaluating if surface wave dispersion modes have been correctly explained.

6.2.2 Include more of the Earth's physics in modelling

The effect of near surface anisotropy in the amplitude of surface waves needs further investigation. In global seismology, surface waves have been used to characterise anisotropic features in the upper mantle (Gaherty, 2004). Surface wave attenuation can also be used in inversion in order to determine medium elastic properties (Lai, 1998).

6.2.3 Extension to 3D

In this work, we have addressed the problem of 2D surface wave inversion. We consider the transfer of the developed methodology to 3D is feasible and without major difficulties. Besides, 3D wave propagation modelling leads to reliable amplitudes an this should improve the result of amplitude based techniques (w-AWI). Dispersion analysis in the extra dimension is not fundamentally different from what it is in 2D (dispersion gathers in the $f-k$ domain). Finally, spatial aliasing should be considered in 3D seismics. If it is present, then the missing information will be in the high-frequency large-wavenumber events which mostly correspond to the tail of the surface wave fundamental mode. This part of the fundamental mode is mainly sensitive to the slow velocities of the very shallow near surface. Thus, the effect of spatial aliasing may decrease the quality of the near surface seismic imaging.

Appendix A

Residual source for the adjoint-state variable in w-AWI

We define the w-AWI objective function shown in equation 4.7 as

$$\tilde{\Phi}(\mathbf{m}) = \frac{1}{2}\sum_{src}\sum_{r}\int_{-K}^{K}\int_{-F}^{F}(|\mathbf{D}_r(\mathbf{m}|f,k)| - |\mathbf{D}_r^{obs}(f,k)|)^2 \mathrm{d}f\mathrm{d}k, \tag{A.1}$$

where, instead of using the least-squares representation, we use the frequency and wavenumber integrals. The adjoint-state residual source $\partial\tilde{\Phi}(\mathbf{m})/\partial\mathbf{v}(\mathbf{m})$ in w-AWI is defined as

$$\frac{\partial\tilde{\Phi}(\mathbf{m})}{\partial\mathbf{v}(\mathbf{m})} = \sum_{src}\sum_{r}\int_{-K}^{K}\int_{-F}^{F}\frac{\partial|\mathbf{D}_r(\mathbf{m}|f,k)|}{\partial\mathbf{v}(\mathbf{m})}(|\mathbf{D}_r(\mathbf{m}|f,k)| - |\mathbf{D}_r^{obs}(f,k)|)\mathrm{d}f\mathrm{d}k \tag{A.2}$$

We must first find a solution for

$$\frac{\partial|\mathbf{D_r}(\mathbf{m}|f,k)|}{\partial\mathbf{v}(\mathbf{m})} = \frac{\partial|\mathcal{F}_{2D}\{\mathbf{S}^T(x,z)\mathbf{w}_r(x,t)\mathbf{v}(\mathbf{m}|x,z,t)\}|}{\partial\mathbf{v}(\mathbf{m})}. \tag{A.3}$$

To do this, we replace the absolute value to its expression in terms of the real and imaginary parts as

$$|\mathbf{D_r}| = \sqrt{\Re^2\{\mathbf{D_r}\} - \Im^2\{\mathbf{D_r}\}}, \tag{A.4}$$

where $\Re$ and $\Im$ are the operands that take the real and imaginary parts, respectively (we have neglected the dependence of $\mathbf{D_r}$ on the model parameter and the coordinates). By using equation A.4, we obtain

$$\begin{aligned}\frac{\partial|\mathbf{D_r}|}{\partial\mathbf{v}} &= \frac{1}{2\sqrt{\Re^2\{\mathbf{D_r}\} - \Im^2\{\mathbf{D_r}\}}}\frac{\partial\left(\Re^2\{\mathbf{D_r}\} - \Im^2\{\mathbf{D_r}\}\right)}{\partial\mathbf{v}} \\ &= \frac{1}{|\mathbf{D_r}(f,k)|}\left(\Re\{\mathbf{D_r}\}\frac{\partial\Re\{\mathbf{D_r}\}}{\partial\mathbf{v}} - \Im\{\mathbf{D_r}\}\frac{\partial\Im\{\mathbf{D_r}\}}{\partial\mathbf{v}}\right)\end{aligned} \tag{A.5}$$

To find the result of equation A.5, let us rewritten the 2D Fourier transform by using the identity $e^{i\theta}e^{i\psi} = \cos(\theta+\psi) + i\cdot\sin(\theta+\psi)$ as

$$\mathcal{F}_{2D}\{g(x,t))\} = \iint g(x,t)(\cos(2\pi(kx-ft)) + i\sin(2\pi(kx-ft)))\mathrm{d}t\mathrm{d}x, \tag{A.6}$$

such that the real and imaginary parts of the 2D Fourier transform are expressed as

$$\Re\{\mathcal{F}_{2D}\{g(x,t))\}\} = \iint g(x,t)(\cos(2\pi(kx-ft)))\mathrm{d}t\mathrm{d}x, \tag{A.7}$$

and

$$\Im\{\mathcal{F}_{2D}\{g(x,t))\}\} = \iint g(x,t)(\sin(2\pi(kx-ft)))\mathrm{d}t\mathrm{d}x. \tag{A.8}$$

The derivatives of real and imaginary parts of the 2D Fourier transform with respect to $g(x,t)$ are

$$\frac{\partial\Re\{\mathcal{F}_{2D}\{g(x,t))\}\}}{g(x,t)} = \cos(2\pi(kx-ft)), \tag{A.9}$$

and

$$\frac{\partial\Im\{\mathcal{F}_{2D}\{g(x,t))\}\}}{g(x,t)} = \sin(2\pi(kx-ft)), \tag{A.10}$$

We use equations A.9 and A.10 to develop equation A.5 to

$$\frac{\partial|\mathbf{D_r}|}{\partial\mathbf{v}} = \frac{\mathbf{S}^T\mathbf{w}_r}{|\mathbf{D_r}|}\left(\Re\{\mathbf{D_r}\}\cos(2\pi(kx-ft)) - \Im\{\mathbf{D_r}\}\sin(2\pi(kx-ft))\right). \tag{A.11}$$

We consider the identity

$$\Re\{\mathbf{D_r}\}\cos(2\pi(kx-ft)) - \Im\{\mathbf{D_r}\}\sin(2\pi(kx-ft)) = \Re\left\{e^{-i2\pi(kx-ft)}\mathbf{D_r}\right\}, \tag{A.12}$$

to replace it into the second term at the right-hand side of equation A.11 and obtain

$$\begin{aligned}\frac{\partial|\mathbf{D_r}|}{\partial\mathbf{v}} &= \frac{\mathbf{S}^T\mathbf{w}_r}{|\mathbf{D_r}|}\Re\left\{e^{-i2\pi(kx-ft)}\mathbf{D_r}\right\} \\ &= \mathbf{S}^T\mathbf{w}_r\Re\left\{e^{-i2\pi(kx-ft)}\frac{\mathbf{D_r}}{|\mathbf{D_r}|}\right\}\end{aligned} \tag{A.13}$$

With the result of equation A.12, we can find an expression for equation A.2 as

$$\begin{aligned}\frac{\partial\tilde{\Phi}}{\partial\mathbf{v}} &= \mathbf{S}^T\sum_r\mathbf{w}_r\int_{-K}^{K}\int_{-F}^{F}\Re\left\{e^{-i2\pi(kx-ft)}\frac{\mathbf{D_r}}{|\mathbf{D_r}|}\right\}(|\mathbf{D}_r| - |\mathbf{D}_r^{obs}|)\mathrm{d}f\mathrm{d}k \\ &= \mathbf{S}^T\sum_r\mathbf{w}_r\Re\left\{\int_{-K}^{K}\int_{-F}^{F}e^{-i2\pi(kx-ft)}\frac{\mathbf{D_r}}{|\mathbf{D_r}|}(|\mathbf{D}_r| - |\mathbf{D}_r^{obs}|)\mathrm{d}f\mathrm{d}k\right\}\end{aligned} \tag{A.14}$$

which can be expressed as

$$\frac{\partial\tilde{\Phi}}{\partial\mathbf{v}} = \mathbf{S}^T\sum_r\mathbf{w}_r\Re\left\{\mathcal{F}_{2D}^{-1}\left\{\frac{\mathbf{D_r}}{|\mathbf{D_r}|}(|\mathbf{D}_r| - |\mathbf{D}_r^{obs}|)\right\}\right\}, \tag{A.15}$$

where the operand $\mathcal{F}_{2D}^{-1}$ stands for the inverse 2D Fourier transform.

List of Figures

1.1 Land seismic shot gathers: radial (left) and vertical (right) components. Surface waves are below the white dashed line. 14

1.2 Characterisation of the subduction zone under America proposed by Mégnin and Romanowicz (2000). Body, surface and higher-mode waveforms have been analysed to obtain this mantle image. 15

1.3 Example of tomography result (here a depth slice) extracted from Taillandier et al. (2009). The resulting velocity model after inversion (in red) better fits with the real model (in green) when the near surface is known (bottom plot) versus the case of unknown near surface (top plot). The initial velocity before inversion is in blue. 16

1.4 In homogeneous media (left), there is no geometric dispersion, while in heterogeneous media (right), each wavelength propagates with a different velocity (extract adapted from Strobbia (2002)). 17

1.5 Scheme of surface-wave methods (extract adapted from Strobbia (2002)). 18

1.6 Snapshot of the wave propagation in acoustic (top) and elastic (bottom) homogeneous media with a free surface. Rayleigh waves are originated at the free-surface contact in the P-SV elastic case. A monochromatic wave source has been employed. 20

1.7 A real data set explained by FWI (Plessix et al., 2012). Acoustic FWI was used to find the model explaining diving waves propagated in a sedimentary context. 21

2.1 Staggered grid with the bold solid line specifying the free-surface boundary. Solid lines identify the principal grid and dashed lines represent the auxiliary staggered grid. 30

2.2 Discretisation scheme for the free-surface implementation. 32

2.3 Numerical dispersion observed in Rayleigh-wave propagation for offsets up to 100 times the shear wavelength. The SIM (centre) and AEA (right) methods were implemented with $n_\lambda = 30$ points per shear wavelength. They are compared to the analytical solution (left). The same move-out correction has been applied to all the three shot gathers. 33

2.4 Rayleigh-wave properties for SIM method. Top left: phase angle in the frequency-offset domain. Top right: phase angle derivative with respect to frequency. Bottom: comparison between numerical and analytical signals using 30 points per shear wavelength for *offsets equal to 5 (left) and to 33 (right) times the shear wavelength*. 34

2.5 Rayleigh-wave properties for AEA method. Top left: phase angle in the frequency-offset domain. Top right: phase angle derivative with respect to frequency. Bottom: comparison between numerical and analytical signals using 30 points per shear wavelength for offsets equal to *5* (left) and to *33* (right) times the shear wavelength. 34

2.6 Top: General scheme for the boundary conditions. Bottom left: attenuation functions a_x and b_x for a specific case. Bottom right: CPML parameter d_x for a specific case (Δt=1 ms; f = 50 Hz and V = 2000 m/s). 35

2.7 Grid used to validate the finite-difference approach on rectangular domains. Lines are plotted every 5 grid points. The lighter parts of the grid identify the CPML attenuation zone. 37

2.8 Analytical and numerical seismograms for the Lamb's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right). 38

2.9 Analytical and numerical seismograms for the Garvin's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right). 38

2.10 Analytical and numerical seismograms for the elastic/elastic problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right). 39

2.11 Staggered grid for rectangular domains. The arrows indicate how an averaging process would be performed in the case of curved domains. 41

2.12 Modified staggered grid for curved domains. Velocity components are placed at the points identified with a triangle. The stress components are placed at the points identified with a circle. The arrows indicate how each centred derivative is computed. The bold line identifies interface with the free surface. 42

2.13 Theoretical computational cost of tensorial and chain rule solutions per node vs. number of operations per partial derivative 46

2.14 Example of the differentiation direction defined by ξ in the CPML attenuation zone of a curved domain. 47

2.15 Curved grid used for numerical validation. The circle represents the source and the triangles represent the receivers. The lighter parts of the grid identify the CPML attenuation zone. Only one every 5 grid lines is plotted. 48

2.16 Analytical and numerical seismograms for the tilted Lamb's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.15. 49

2.17 Analytical and numerical seismograms for the tilted Garvin's problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.15. 50

2.18 Grid used to model the two-layer medium for the tilted elastic/elastic problem. 51

2.19 Analytical and numerical seismograms for the tilted elastic/elastic problem. Vertical (top) and horizontal (bottom) velocities are plotted for R1 (left) and R2 (right), accordingly to the scheme of Figure 2.18. 52
2.20 Top: different grids with 45° to 90° angles between axes have been used. Only the first and the last grid are plotted. Bottom: relative position of source and receivers for the complete set of experiments. 53
2.21 MSE for the range of angles from 45° to 90° measured at the two receivers described in Figure 2.20. 54
2.22 Correlation for the range of angles from 45° to 90° measured at the two receivers described in Figure 2.20. 55
2.23 Complex medium. Top: 2D S-wave velocity model. Bottom: 1D elastic parameter profiles for the positions indicated in the 2D model by the black and the grey lines. 56
2.24 Discretisation grid (top) and snapshot of wave propagation at t=0.125 s (bottom) for $\mathrm{V_z}$ component. Only one every 7 grid lines is plotted. The circles indicate the source position. Identified waves are examples of: P direct (P-D), refracted (Rc), reflected (Rl) and surface waves (GR). . . . 57
2.25 Shot gathers recorded at the surface (Figures 2.23 and 2.24). Seismic waves indicated by arrows are examples of waves identified in the snapshot (Figure 2.24). Clipping is stronger than in the snapshot. 58

3.1 Synthetic shot-gather (left) computed for a 1D layered medium (right) using our finite-difference modelling code (Chapter 2). Rayleigh waves in the shot gather are dispersive and are indicated by the double arrow. . . 67
3.2 Surface wave processing workflow used to identify dispersion curves (modal curves) from shot gathers. 68
3.3 Spectrum amplitude for the shot gather in Figure 3.1. In this case, the fundamental mode is easily recognised because of its high energy. The energy indicated by the circle might not be a Rayleigh-wave mode (explanation given in the text). 68
3.4 Left: the theoretical range of Rayleigh-wave velocities for the model presented in Figure 3.1 is indicated with the vertical lines. Right: the spectrum of Figure 3.3 is transformed to the $f-v$ domain. Note that the $f-v$ domain on the right-hand side image is a stretched and mirrored version of the theoretical range indicated on the left-hand side image. 70
3.5 Dispersion curves corresponding to the model shown in Figure 3.1. The curves identified with processing (Figure 3.2) are plotted with solid lines. The theoretical curves are plotted with dashed lines. 71
3.6 The surface-wave problem in a layered medium. Left: A layered model in the zone of Rayleigh wave propagation ($2 \times \lambda_s$). Each layer is identified with model parameters $\mathrm{m}_i = (\mathrm{v}_{pi}, \mathrm{v}_{si}, \rho_i, \mathrm{h}_i)$ for $i = 1, 2, ..., n$. Right: the dispersion curves for the Rayleigh waves propagating in this medium. . . 72
3.7 Dispersion curves for the model of Figure 3.1. The box indicates the region of chosen frequencies for a fundamental mode sensitivity analysis. 74

3.8 Sensitivity analysis for the Rayleigh-wave fundamental mode of Figure 3.7 with respect to the S-wave velocities. The velocity model is shown in Figure 3.1. 75
3.9 Damping factor (left) and S-wave velocities (right) for each iteration using the Marquardt-Levenberg method for surface-wave inversion applied to the model of Figure 3.1. 76
3.10 5-layer model (Table 3.1) inversion with constant damping factor and no weighting. Left: fundamental mode of Rayleigh-wave velocity. Right: S-wave velocities for each iteration using the Marquardt-Levenberg inversion method. 77
3.11 7-layer model (Table 3.2) inversion *without weighting or regularisation*. Left: the data are 25 samples of the fundamental mode of Rayleigh-wave velocity. Right: S-wave profiles. 77
3.12 Objective function when the initial Vs model is controlled by one single model parameter. For m_p=0, the initial model is the exact one. For m_p=1, the initial model is the constant model shown in 3.11. No local minima is observed in this range. 78
3.13 7-layer model (Table 3.2) inversion *including dynamic weighting and dynamic regularisation*. Top: data (left) and S-wave profiles (right). Bottom: dynamic damping factor (left) and objective function (right) with iterations. 79
3.14 7-layer model (Table 3.2) inversion when the order of layers is changed to 1, 2, 4, 3, 6, 5, 7. *Dynamic weighting and dynamic regularisation are included*. Top: data (left) and S-wave profiles (right). Bottom: dynamic damping factor (left) and objective function (right) with iterations. . . . 80
3.15 7-layer model (Table 3.2) *thickness* inversion. Left: the data are 25 samples of the fundamental mode of Rayleigh-wave velocity. Right: thickness values through iterations. 81
3.16 Sensitivity analysis of the Rayleigh-wave velocity with respect to variations in thickness for the model of Table 3.2. The dashed blue line indicates the beginning of the chosen frequencies for the inversion process (15-80 Hz). . 81
3.17 Inverted data (left) and S-wave velocity profile (right) for exact, initial and reconstructed models (legends are the same at right and left). Exact model thicknesses *are* multiples of the initial model ones. Top: inversion of only the fundamental mode. Middle: inversion of two modes. Bottom: inversion of three modes. 83
3.18 Same as Figure 3.17 but exact model thicknesses *are not* multiples of the initial model ones. 84
3.19 *Acoustic* wavepaths in a 3D medium (Woodward (1992)). Two monochromatic sources have been employed. The limit of the first Fresnel zone F_1 is indicated in blue. A vertical section across the wavepath is plotted in red. 90
3.20 *Elastic* shot gathers (vertical component of velocity) at the surface for two different Vs models (Figure 3.13, top right). Left: the Vs model is layered. Right: the Vs model is homogeneous. The Vp and ρ models are identical for both shot gathers. Strong phase differences prevent FWI from converging. 91

3.21 1D FWI gradients (for μ on the left, for λ on the right) considering the initial profile of Figure 3.13. Higher values are concentrated in the first 10 m depth. 91

4.1 Workflow for w-AWI residual computation (equation 4.8). 101

4.2 Example of data considered in the w-AWI residual. Top left: input shot gather. Top right: two windows used to select traces in the shot gather (equation 4.9). Middle: two subsets obtained by multiplying the shot gather by the windows ($\mathbf{w}_a(t,x)$ on the left; $\mathbf{w}_b(t,x)$ on the right). Bottom: absolute value of the 2D Fourier transforms of the two subsets shown in the middle. 102

4.3 Elastic models used to compute synthetic data representing our observed data. Left: P-wave velocity (top) and density (bottom) homogeneous models. Right: S-wave velocity model (top) and the discretisation grid (bottom) conforming with the anomaly. 107

4.4 Horizontal (left) and vertical (right) velocity component of observed data. The data represent surface recordings and were computed for the model shown in Figure 4.3. 107

4.5 Normalised objective function value with respect to two parameters: background (Vs_0) and anomaly (Vs_1) velocities. Left: FWI. Rigth: w-AWI. . 108

4.6 Normalised objective function curves taken from Figure 4.5 at Vs_1=980 m/s (left) and Vs_0 = 800 m/s (right). The red 'x' on the right-hand side plot indicates Vs_1=920 m/s. 108

4.7 S-wave velocity model Vs_0 composed of a constant background and an anomaly. The background velocity is Vs_0 = 1200 m/s. In this image, the anomaly velocity is $Vs_1 = 1.15Vs_0$. The black circle indicates the source position. 109

4.8 Objective function with respect to one model parameter α and computed with Ricker wavelets centred at different frequency values (as indicated in each plot title). The legend is the same for the four plots. 109

4.9 Same as Figure 4.8 but using monochromatic signal sources. The legend is the same for the three plots. 110

4.10 Initial S-wave velocity models. Synthetic data are computed for three different initial models: model 1 and 2 (left) are homogeneous with Vs=600 m/s and Vs=800 m/s, and model 3 has homogenous background with Vs_0 = 800 m/s and one anomaly with Vs_1 = 920 m/s (right). 111

4.11 Initial data (vertical component of velocity) computed with the S-wave velocity models of Figure 4.10. Left: model 1. Centre: model 2. Right: model 3. 111

4.12 Observed and modelled data for *model 1* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom). 111

4.13 Observed and modelled data for *model 2* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom). 112

4.14 Observed and modelled data for *model 3* (Figure 4.10) taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom). 112

4.15 Real part of the 2D Fourier transform of observed and modelled data for *model 1* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 113

4.16 Real part of the 2D Fourier transform of observed and modelled data for *model 2* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 113

4.17 Real part of the 2D Fourier transform of observed and modelled data for *model 3* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 114

4.18 Absolute value of the 2D Fourier transform of observed and modelled data for *model 1* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 114

4.19 Absolute value of the 2D Fourier transform of observed and modelled data for *model 2* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 115

4.20 Absolute value of the 2D Fourier transform of observed and modelled data for *model 3* (Figure 4.10) taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows used here are centred at 62 m (top), 164 m (middle), and 286 m (bottom). 115

4.21 Same as Figure 4.18 (*model 1*) but using $l_w = 64$ m windows. 116

4.22 Exact and initial elastic models. The exact and initial P-wave velocity (top, left) and density (bottom, left) models are homogeneous. The exact (top, right) and initial (bottom, right) S-wave velocity models only differ for the velocity inclusion. 117

4.23 Shot gathers showing the vertical component of velocity. Observed data (left) and initial data (right). 117

4.24 Observed and modelled seismic traces taken at three offset positions: 62 m (top), 164 m (middle) and 286 m (bottom). 118

4.25 Absolute value of the 2D Fourier transform of observed (left) and initial (right) data. The window used here contains the whole set of traces ($l_w =$ 450 m). 119

4.26 Slices of the data shown in Figure 4.25 taken at $k = 0.05$ m^{-1}. 119

4.27 Absolute value of the 2D Fourier transform of observed (top) and initial (bottom) data. The $l_w = 256$ m windows are centred at 62 m (left), 174 m (middle), and 286 m (right). 120

4.28 Slices of the data shown in Figure 4.27 taken at $k = 0.05$ m^{-1}. The $l_w = 256$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom). 120

4.29 Same as Figure 4.27 but using $l_w = 128$ m windows. 121

4.30 Slices of the data shown in Figure 4.29 taken at $k = 0.05$ m^{-1}. The $l_w = 128$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom). There is almost no difference in the bottom panel because the corresponding windows mostly select receivers (222-350 m) that are not above the anomaly (200-250 m). 121

4.31 Same as Figure 4.27 but using $l_w = 64$ m windows. 122

4.32 Slices of the data shown in Figure 4.31 taken at $k = 0.05$ m^{-1}. The $l_w = 64$ m windows are centred at 62 m (top), 174 m (middle), and 286 m (bottom). 122

4.33 Adjoint sources computed with windows of different widths l_w. All the windows centres are separated by 4 m, at the exception of the one with $l_w = 450$ m. 123

4.34 The w-AWI gradient computed with the window configurations of Figure 4.33. The position of the exact velocity anomaly of the model is contoured in black. 124

4.35 S-wave velocity profile consisting of a homogenous background and a velocity anomaly. 125

4.36 Seismic common shot gather showing the vertical component of velocity recorded near the surface. The source signature is a 40 Hz Ricker wavelet. Source excitation is along the vertical direction. 125

4.37 Observed (black) and modelled (red) seismic traces for three different source-receiver offset intervals indicated on the horizontal axes. Top: data computed with a 40 Hz Ricker wavelet. Bottom: band-passed filtered data (5-30 Hz). 126

4.38 Gradients of FWI objective function with respect to μ. Left: data computed with a 40 Hz Ricker wavelet. Right: band-passed filtered data (5-30 Hz). The initial model is the same for both gradients. 126

4.39 Absolute value of the 2D Fourier transform of observed (black) and modelled (red) data (Figure 4.37). Wiggle plotting is used for graphical purposes regarding amplitude match. 127

4.40 Gradients of w-AWI objective function with respect to μ using windows spaced 4 m apart. The initial model is the same for both gradients. Left: windows with l_w=128 m. Right: windows with l_w=256 m. 127

4.41 Objective function value for several initial models with different background and anomaly S-wave velocities. Left: w-AWI with $l_w = 128$ m. Middle: FWI. Right: FWI with band-pass (5-30 Hz) filtered data. The white cross indicates the exact model. The white circle indicates a model with homogeneous S-wave velocity of 800 m/s. 128

4.42 Exact velocity model for 2D inversion. Red circles indicate the position of 20 seismic sources used to generate the data. Density is either 2D (third panel) or homogenous (bottom panel) depending on the test. 130

4.43 Inversion 1. Initial model (first and second panels), FWI result (third and fourth panels) and w-AWI result (fifth and last panels). Density is considered known (homogeneous) and is not updated during inversion. . 132
4.44 Inversion 1. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m. 133
4.45 Example of observed (left) and initial (right) common shot gathers. . . . 133
4.46 Inversion 1. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 134
4.47 Inversion 2. Initial model (first and second panels), FWI result (third and fourth panels) and w-AWI result (fifth and last panels). Density is considered known (homogeneous) and is not updated during inversion. . 135
4.48 Inversion 2. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m. 136
4.49 Inversion 2. Observed (left) and initial (right) common shot gathers. Only one shot gather is shown. 136
4.50 Inversion 2. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 137
4.51 Inversion 3. FWI result (first and second panels) and w-AWI result (third and last panels). The exact density model is 2D (Figure 4.42). The initial density model is homogeneous and is not updated during inversion. . . . 138
4.52 Inversion 3. 1D profiles of exact, initial and reconstructed models by w-AWI and FWI. Top: P-wave velocity profiles. Bottom: S-wave velocity profiles. From left to right: x = 13, 20, 27, 39 m. 139
4.53 Inversion 3. Observed (left) and initial (right) common shot gathers. Only one shot gather is shown. 139
4.54 Inversion 3. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 140
4.55 Exact (left) and initial (right) source wavelets. 141
4.56 Observed (left) and initial common shot gathers (centre for test 1 and 2 and right for test 3). Only shot gathers due to one of the 20 considered sources are displayed. 142
4.57 Test 1. Initial model (first and second panels), FWI (third and fourth panels) and w-AWI results (fifth and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is known and is not updated during inversion. 143
4.58 Test 1. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m. 144

4.59 Test 1. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 145

4.60 Test 2. Final source wavelets estimated with FWI (left) and w-AWI (right). The estimated wavelet is more similar to the exact one (Figure 4.55 left) in the case of FWI because, differently from w-AWI, the phase is taken into account. 145

4.61 Test 2. FWI (first and second panels) and w-AWI results (third and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is unknown and updated during inversion. 146

4.62 Test 2. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m. 147

4.63 Test 2. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 148

4.64 Test 3. Final source wavelets estimated with FWI (left) and w-AWI (right). The wavelet estimated with FWI shows convergence problems as it do not coincide to the exact one (Figure 4.55 left). The wavelet estimated with w-AWI allows convergence even if its shape is not similar to the exact one, which is a result of neglecting the phase. 148

4.65 Test 3. FWI (first and second panels) and w-AWI results (third and last panels). Density is known (homogeneous) and is not updated during inversion. Source wavelet is unknown and updated during inversion. 149

4.66 Test 3. 1D plots of exact, initial and reconstructed models by FWI and w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m. 150

4.67 Test 3. Misfit minimisation (top left) and final data obtained with FWI (top centre) and w-AWI (top right). Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 151

4.68 w-AWI multi-step inversion result. First step: only λ and μ are updated. 152

4.69 1D plots of exact, initial and reconstructed models by w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m. 153

4.70 Misfit minimisation (top left) and final data obtained with w-AWI (top right). Bottom: seismic traces for 3 offset positions as indicated in the shot gathers. 154

4.71 w-AWI multi-step inversion result. Second step: only density is updated. Top: final 2D density model. Bottom: 1D exact, initial and final density profiles. From left to right: x=13, 20, 27, 39 m. 155

4.72 Misfit minimisation (top left) and final data obtained with w-AWI (top right) after density inversion. Bottom panels: seismic traces for 3 offset positions as indicated in the shot gathers. 156

4.73 w-AWI multi-step inversion result. Third step: only λ and μ are updated. 156

4.74 1D plots of exact, initial and reconstructed models by w-AWI. Top: P-wave velocity profile. Bottom: S-wave velocity profile. From left to right: x = 13, 20, 27, 39 m. 157

4.75 Misfit minimisation (top left) and final data obtained with w-AWI (top right). Bottom: seismic traces for 3 offset positions as indicated in the shot gathers. 158

4.76 Exact model. From top to bottom: Vs, Vp, ρ and grid for modelling and inversion. 159

4.77 Unfiltered data inversion. Relative error between final and exact models. Top: Vp and Vs errors for FWI. Bottom: Vp and Vs errors for w-AWI. . 160

4.78 Bandpass filtered (2-9 Hz) data inversion. Relative error between final and exact models. Top: Vp and Vs errors for FWI. Bottom: Vp and Vs errors for w-AWI. 161

4.79 Objective function evolution with iterations. Left: unfiltered data. Right: filtered data. 161

4.80 Shot gathers of observed (left) and initial (right) data without filtering (top) and with bandpass filtering (2-9 Hz) (bottom). Clipping level is equal for images in the same row. 162

4.81 Comparison of seismic traces for unfiltered data at three offsets: 150 m, 375 m and 600 m. 163

4.82 Comparison of seismic traces for filtered data (2-9 Hz) at three offsets: 150 m, 375 m and 600 m. 163

4.83 Gradient of the FWI (top) and w-AWI (bottom) objective function with unfiltered data. 164

4.84 Gradient of the FWI (top) and w-AWI (bottom) objective function with filtered (2-9 Hz) data. 164

5.1 Purpose-built ground model, extracted from Deidda and Balia (2001). Velocity and density values obtained with walk-away tests are indicated. The displayed distances are in metres. Piezometers were placed to check water table levels. 169

5.2 Theoretical wave speed values in the near surface. Left: waves can propagate as slow as 20 m/s (at critical porosity) (Bachrach et al., 1998). Right: velocity increases rapidly with depth, Bachrach and Nur (1998). 170

5.3 The apparent velocity within the first 3 m offset is around 115 m/s (shot-gather 20 of the real dataset). 170

5.4 Sketch of the acquired seismic line and the ground model. Geophone and source positions are respectively indicated in green and red. Top: cross section. Middle: 3D view. Bottom: cross-line cross section. 171

5.5 Representative raw shot gathers 10 and 30 before preprocessing. 172

5.6 Shot gather 30 after different preprocessing stages. Top left: after interpolation to remove noisy traces. Top right: after interpolation and geometrical spreading correction. Bottom left: mask used to mute near offset traces and late arrivals. Bottom right: preprocessed data after multiplication with the muting mask. The mask is tapered to prevent inversion artifacts. 173

5.7 Left: frequency-space spectrum of raw shot gather 30 (Figure 5.5). Right: example of the 10th-order Butterworth band-pass filter considered for inversion. In this case, the band-passed is 10-80 Hz. 173

5.8 Vp model obtained with first-arrival traveltime tomography. The white points indicate the theoretical position of the concrete anomaly. 174

5.9 Left: fundamental mode plotted in the f-v domain (top) and with respect to pseudo depth (bottom). Right: surface-wave fundamental mode picked in gather # 34 (dashed black line). 175

5.10 Pseudo 2D Vs model formed by 72 1D Vs profiles. The white points indicate the velocity anomaly. A higher velocity ($\approx$ 250 m/s) zone can be distinguished at the approximate anomaly position. Constant ρ = 2000 kg/m^3 was kept fixed during inversion. 176

5.11 Waveform modelling for the final model retrieved with SWA (Figure 5.10). The same gathering process of real data are applied to modelled data. The fundamental mode picked from real data (dashed line) is near to the fundamental mode of modelled data but not centred at its energy maximum. 176

5.12 Initial Vp (top left), Vs (top right) and ρ (bottom) models for FWI and w-AWI. 178

5.13 Vp *final models* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz. 179

5.14 Vp *model updates* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz. 180

5.15 Vs *final models* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz. 181

5.16 Vs *model updates* obtained with FWI (left) and w-AWI (right). The high-cut frequencies are considered from top to bottom as 18, 25, 36, 50 and 70 Hz. 182

5.17 Objective function evolution with iterations for the 10-70 Hz bandpass data. Left: FWI results. Right: w-AWI was employed, but the FWI objective function computed through w-AWI iterations is also shown. . . 183

5.18 Representative result (shot gather 30) for FWI with 10-70 Hz band-passed data. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals. 183

5.19 Representative result (shot gather 30, negative offset window) for w-AWI with 10-70 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Middle: spectra slice comparison. Bottom: initial (left) and final (right) residuals. 185
5.20 Representative result (shot gather 30, positive offset window) for w-AWI with 10-70 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Middle: spectra slice comparison. Bottom: initial (left) and final (right) residuals. 186
5.21 Representative result (shot gather 30) for w-AWI with 10-70 Hz band-passed data. The differences are mainly related to the phase. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals. 187
5.22 Objective function evolution with iterations. This is the second stage of inversion (FWI after w-AWI). 187
5.23 Representative result (shot gather 30) for FWI with 10-70 Hz band-passed data after w-AWI. Top: initial (red), reconstructed (blue) data superposed to observed (black) data. Bottom: initial (left) and final (right) residuals. 188
5.24 Source wavelets found after w-AWI (top) and w-AWI + FWI (bottom) inversion. Wavelets found by w-AWI are not shifted mainly because the phase is not inverted. 188
5.25 Second stage inversion result (FWI after w-AWI). Final result for Vp (top) and Vs (bottom). 189
5.26 Result of w-AWI with 25-80 Hz band-passed data. Top: Vp (left) and Vs (right) final result. Bottom: final update for Vp (left) and Vs (left). . . . 190
5.27 Result of w-AWI with 36-80 Hz band-passed data. Top: Vp (left) and Vs (right) final result. Bottom: final update for Vp (left) and Vs (left). . . . 190
5.28 Objective function evolution with iterations. Left: 25-80 Hz. Right: 36-80 Hz. Only w-AWI was employed, but both w-AWI and FWI misfits are displayed. 190
5.29 Representative result (shot gather 30, negative offset window) for w-AWI with 25-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals. 191
5.30 Representative result (shot gather 30, positive offset window) for w-AWI with 25-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals. 191
5.31 Representative result (shot gather 30, negative offset window) for w-AWI with 36-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals. 192
5.32 Representative result (shot gather 30, positive offset window) for w-AWI with 36-80 Hz band-passed data. Top: observed (left), initial (centre) and final (right) data. Bottom: initial (left) and final (right) residuals. 192

Bibliography

Abo-Zena, A., 1979, Dispersion function computations of unlimited frequency values: Geophysical Journal of the Royal Astronomical Society, **58**, 91–105.

Aki, K., 1960a, Further study of the mechanism of Circum-Pacific earthquakes from Rayleigh waves: Journal of Geophysical Research, **65**, 4165–4172.

———, 1960b, Study of earthquake mechanism by a method of phase equalization applied to Rayleigh and Love waves: Journal of Geophysical Research, **65**, 729–740.

Aki, K., and P. G. Richards, 1980, Quantitative seismology: W.H. Freeman & Co.

Alkhalifah, T., and Y. Choi, 2013, From tomography to FWI with a single objective function: Proceedings of the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC. (London, UK).

Asnaashari, A., 2013, Quantitative 4D seismic imaging in complex media using 2D full waveform inversion. PhD thesis: Université de Grenoble.

Bachrach, R., J. Dvorkin, and A. Nur, 1998, High-resolution shallow-seismic experiments in sand, Part II: Velocities in shallow unconsolidated sand: Geophysics, **63**, 1234–1240.

Bachrach, R., and A. Nur, 1998, Ultra shallow seismic reflection in unconsolidated sediments: Rock physics base for data acquisition: Society of Exploration Geophysicists Technical Program Expanded Abstracts, 866–869.

Baek, H., H. Calandra, and L. Demanet, 2013, Registration-based guided least-squares waveform inversion: Proceedings of the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC. (London, UK).

Beaty, K. S., D. R. Schmitt, and M. Sacchi, 2002, Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure: Geophysical Journal International, **151**, 622–631.

Bérenger, J. P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: Journal of Computational Physics, **114**, 185–200.

Berg, P., P. If, P. Nilsen, and O. Skovgaard, 1994, Analytical reference solutions: Advanced seismic modeling, in k. Helbig, ed.: Modeling the earth for oil exploration: Pergamon Press, 421–427.

Bergamo, P., D. Boiero, and L. V. Socco, 2012, Retrieving 2D structures from surface-wave data by means of spatial-varying spatial windowing: Geophysics, **77**, EN39–EN51.

Bleistein, N., 1986, Two and one half dimensional in plane wave propagation: Geophysical Prospecting, 686–703.

Bohlen, T., and E. H. Saenger, 2006, Accuracy of heterogeneous staggered-grid finite-difference modeling of Rayleigh waves: Geophysics, **71**, 109–115.

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

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, WCC105–WCC118.

Brossier, R., J. Virieux, and S. Operto, 2008, Parsimonious finite-volume fre- quency-

domain method for 2-D P-SV-wave modelling: Geophysical Journal International, **175**, 541–559.

Buchen, P. W., and R. Ben-Hador, 1996, Free-mode surface-wave computations: Geophysical Journal International, **124**, 869–887.

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

Carcione, J. M., 1994, The wave equation in generalized coordinates: Geophysics, **59**, 1911–1919.

Carcione, J. M., and J.-P. Wang, 1993, A chebyshev collocation method for the elastodynamic equation in generalized coordinates: Computational Fluid Dynamics Journal, **2**, 269–290.

Chapman, C. H., 2004, Fundamental of seismic waves propagation: Cambridge University Press.

Chauris, H., D. Donno, and H. Calandra, 2012, Velocity estimation with the normalized integration method: Proceedings of the 74th EAGE Conference & Exhibition incorporating SPE EUROPEC. (Copenhagen, Denmark).

Chen, X., 1983, A systematic and efficient method of computing normal modes for multilayered half space: Geophysical Journal International, **15**, 391–409.

Collino, F., and C. Tsogka, 2001, Application of the PML absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media: Geophysics, **66**, 294–307.

de Franco, R., and G. Musacchio, 2001, Polarization filter with singular value decomposition: Geophysics, **66**, 932–938.

de Hoop, A. T., 1960, A modification of Cagniard's method for solving seismic pulse problems: Applied Scientific Research, **B8**, 349–356.

Deidda, G. P., and R. Balia, 2001, An ultrashallow SH-wave seismic reflection experiment on a subsurface ground model: Geophysics, **66**, 1097–1104.

Domenico, S. N., 1974, Effect of water saturation on seismic reflectivity of sand reservoirs encased in shale: Geophysics, **39**, 759–769.

Donno, D., H. Chauris, and H. Calandra, 2013, Estimating the background velocity model with the normalized integration method: Proceedings of the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC. (London, UK).

Donno, D., A. Nehorai, and U. Spagnolini, 2008, Seismic velocity and polarization estimation for wavefield separation: IEEE Transactions on Signal Processing, **56**, 4794–4809.

Dorman, J., M. Ewing, and J. Olivier, 1960, Study of the shear-velocity distribution in the upper mantle by Rayleigh waves: Bulletin of the Seismological Society of America, **50**, 87–115.

Droujinine, A., R. Plessix, and F. Ernst, 2012, Integration of Dispersion Curve and Full Waveform Inversion Techniques for Onshore Velocity Model Building Inner Mongolia Study: Proceedings of the 74th EAGE Conference & Exhibition incorporating SPE EUROPEC. (Copenhagen, Denmark).

Dunkin, J. W., 1965, Computation of modal solutions in layered elastic media at high frequencies: Bulletin of the Seismological Society of America, **55**, 335–358.

Dziewonski, A., S. Bloch, and M. Landisman, 1969, A technique for the analysis of transient seismic signals: Bulletin of the Seismological Society of America, **59**, 427–444.

Ewing, W. M., W. S. Jardetzky, and F. Press, 1957, Elastic waves in layered media: McGraw Hill Book Co Inc.

Forgues, E., and G. Lambaré, 1997, Parametrization study for acoustic and elastic ray+Born inversion: Journal of Seismic Exploration, **6**, 253–277.

Gabriels, P., R. Snieder, and G. Nolet, 1987, In situ measurements of shear- wave velocity in sediments with higher-mode Rayleigh waves: Geophysical Prospecting, **35**, 187–196.

Gaherty, J. B., 2004, A surface wave analysis of seismic anisotropy beneath eastern North America: Geophysical Journal International, **158**, 187–196.

Gao, F., A. Levander, R. G. Pratt, C. A. Zelt, and G. Fradelizio, 2007, Waveform tomography at a groundwater contamination site: surface reflection data: Geophysics, **72**, G45–G55.

Garvin, W., 1956, Exact transient solution of buried line source problem: Proceedings of the Royal Society of London, **234**, 528–541.

Gauthier, O., A. Tarantola, and J. Virieux, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Geophysics, **51**, 1387–1403.

Gélis, C., 2005, Inversion des formes d'onde élastique dans le domain espace-fréquence en deux dimensions. Application á la caractérisation de la subsurface dans le cadre de la détection de cavités souterraines. PhD thesis: Université de Nice-Sophia Antipolis.

Gélis, C., J. Virieux, and G. Grandjean, 2007, Two-dimensional elastic full waveform inversion using Born and Rytov formulations in the frequency domain: Geophysical Journal International, **168**, 605–633.

Godinho Ferreira, A. M., 2005, Seismic surface waves in the laterally heterogenous Earth. PhD thesis: University of Oxford.

Hadamard, J., 1902, Sur les problémes aux derivées partielles et leur signification physique: Princeton University Bulletin, **46**, 49–52.

Haskell, N. A., 1953, The dispersion of surface waves in multilayered media: Bulletin of the Seismological Society of America, **43**, 17–34.

Hermann, R. B., 1973, Some aspects of band-pass filtering of surface waves: Bulletin of the Seismological Society of America, **63**, 663–671.

Hestholm, S., and B. Ruud, 1994, 2D finite-difference elastic wave modeling including surface topography: Geophysical Journal International, **118**, 643–670.

———, 1998, 3-D finite-difference elastic wave modeling including surface topography: Geophysics, **63**, 613–622.

———, 2001, 2D surface topography boundary conditions in seismic wave modelling: Geophysical Prospecting, **49**, 445–460.

Hisada, Y., 1995, An efficient method for computing Green's functions for a layered halspace with sources and receivers at close depths, Part 2: Bulletin of the Seismological Society of America, **85**, 1080–1093.

Kanamori, H., 1970, Synthesis of long-period surface waves and its application to earthquake source studies - Kurile islands earthquake of October 13, 1963: Journal of Geophysical Research, **75**, 5011–5027.

Kaser, M., and M. Dumbser, 2006, An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes - I. The two-dimensional isotropic case with external source terms: Geophysical Journal International, **166**, 855–877.

Kaser, M., and H. Igel, 2001, Numerical simulation of 2D wave propagation on unstructured grids using explicit differential operators: Geophysical Prospecting, **49**, 607–619.

Kennett, B. L. N., and N. J. Kerry, 1979, Seismic waves in a stratified half space: Geophysical Journal Royal Astronomical Society, **57**, 557–583.

Knapp, R. W., and D. W. Steeples, 1986, High-resolution common-depth-point seismic reflection profiling: field acquisition parameter design: Geophysics, **51**, 283–294.

Knopoff, L., 1964, A matrix method for elastic wave problems: Bulletin of the Seismological Society of America, **54**, 431–438.

Köhn, D., D. De Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen, 2012, On the influence of model parametrization in elastic full waveform tomography: Geophysical Journal International, **191**, 325–345.

Komatitsch, D., F. Coutel, and P. Mora, 1996, Tensorial formulation of the wave equation for interfaces: Geophysical Journal International, **127**, 156–168.

Komatitsch, D., and R. Martin, 2007, An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation: Geophysics, **72**, 155–167.

Lai, C. G., 1998, Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization. PhD dissertation: Georgia Institute of Technology.

Lailly, P., 1983, The seismic problem as a sequence of before-stack migrations: Inverse Scattering: Theory and Applications, ed. Bednar, J., SIAM, Philadelphia, PA, 206–220.

Lamb, H., 1904, Exact transient solution of buried line source problem: Philosophical Transactions of the Royal Society of London, **204**, 1–42.

Lester, O. C. J., 1932, Seismic weathered or aerated surface layer: American Association of Petroleum Geologists Bulletin, **16**, 1230–1234.

Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: Geophysics, **53**, 1425–1436.

Levenberg, K., 1944, A method for the solution of certain nonlinear problems in least squares: Quarterly of Applied Mathematics, **2**, 164–168.

Liu, J., H. Charuis, and H. Calandra, 2011, The Normalized Integration Method - An Alternative to Full Waveform Inversion?: Proceedings of the 17th Near Surface Conference. (Liecester, UK).

Lombard, B., and J. Piraux, 2004, Numerical treatment of two-dimensional interfaces for acoustic and elastic waves: Journal of Computational Physics, **195**, 90–116.

Lombard, B., J. Piraux, C. Gélis, and J. Virieux, 2007, Free and smooth boundaries in 2-D finite-difference schemes for transient elastic waves: Geophysical Journal International, **172**, 252–261.

Love, A. E. H., 1911, Some problems of geodynamics: Cambridge University Press.

Luco, J. E., and R. J. Apsel, 1983, On the Green's functions for a layered half space, Part 1: Bulletin of the Seismological Society of America, **73**, 909–929.

Luo, Y., J. Xia, J. Liu, Q. Liu, and S. Xu, 2007, Joint inversion of high-frequency surface waves with fundamental and higher modes: Journal of Applied Geophysics, **62**, 375–384. (Elsevier B. V.).

Luo, Y., J. Xia, R. D. Miller, Y. Xu, J. Liu, and Q. Liu, 2008, Rayleigh-wave dispersive energy imaging using a high-resolution linear radon transform: Pure and Applied Geophysics, **165**, 903–922.

———, 2009, Rayleigh-wave mode separation by high-resolution linear Radon transform: Geophysical Journal International, **179**, 254–264.

Malvern, L. E., 1969, Introduction to the mechanics of a continuous medium: Prentice-Hall, Series in Engineering of the physical sciences.

Maraschini, M., F. Ernst, S. Foti, and L. V. Socco, 2010, A new misfit function for multimodal inversion of surface waves: Geophysics, **75**, G31–G43.

Maraschini, M., and S. Foti, 2010, A Monte-Carlo multimodal inversion of surface waves: Geophysical Journal International, **182**, 1557–1566.

Marfurt, K., 1984, Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation: Geophysics, **49**, 553–549.

Mari, J. L., 1984, Estimation of static corrections for shear-wave profiling using the dispersion properties of Love waves: Geophysics, **49**, 1169–1179.

Marquardt, D. W., 1963, An algorithm for least squares estimation of nonlinear parameters: Journal of the Society of Industrial Applied Mathematics, **2**, 431–441.

Masoni, I., R. Brossier, J. Virieux, and J. L. Boelle, 2013, Alternative misfit functions for FWI applied to surface waves: Proceedings of the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC. (London, UK).

McConnell, A. J., 1957, Applications of tensor analysis: Dover Publications.

McMechan, G. A., and M. J. Yedlin, 1981, Analysis of dispersive wave by wave field transformation: Geophysics, **46**, 869–874.

Mégnin, C., and B. Romanowicz, 2000, The three-dimensional shear wave velocity of the mantle from inversion of body, surface and higher-mode waveforms: Geophysical Journal International, **143**, 709–728.

Métivier, L., R. Brossier, J. Virieux, and S. Operto, 2012, Toward Gauss-Newton and exact Newton optimization for full waveform inversion: Proceedings of the 74th EAGE Conference & Exhibition incorporating SPE EUROPEC. (Copenhagen, Denmark).

Mittet, R., 2002, Free-surface boundary conditions for elastic staggered-grid modeling schemes: Geophysics, **67**, 1616–1623.

Moczo, P., J. O. A. Robertsson, and L. Eisner, 2007, The finite-difference time-domain method for modeling of seismic wave propagation, in Advances in wave propagation in heterogeneous Earth, vol. 48: Academic press.

Moghaddam, P., and W. Mulder, 2012, The Diagonalator, an alternative cost functional for wave-equation inversion: Proceedings of the 74th EAGE Conference & Exhibition incorporating SPE EUROPEC. (Copenhagen, Denmark).

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

———, 1988, Elastic wavefield inverison of reflection and transmission data: Geophysics, **53**, 750–759.

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

Naghizadeh, M., and M. D. Sacchi, 2007, Multistep autoregressive reconstruction of seismic records: Geophysics, **72**, V111–V118.

Nazarian, S., and K. H. Stokoe II, 1984, In situ shear wave velocities from spectral analysis of surface waves: Proceedings on the 8th Conference on Earthquake Engineering, **3**, 31–38. (Prentice Hall, San Francisco, USA).

Neidell, N. S., and M. T. Taner, 1971, Semblance and other coherency measures for multichannel data: Geophysics, **36**, 482–497.

Nocedal, J., 1980, Updating quasi-Newton matrices with limited storage: Mathematics of Computation, **35**, 773–782.

Nolet, G., and G. F. Panza, 1976, Array analysis of seismic surface waves: Limits and

possibilities: Pure and Applied Geophysics, **114**, 775–790.

Oppenheim, A. V., R. W. Schafer, and J. R. Buck, 1998, Discrete-time signal processing: Prentice Hall.

———, 1999, Discrete-time signal processing. Second edition: Prentice Hall.

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

Pestel, E., and F. A. Leckie, 1963, Matrix methods in elasto-dynamics: MGraw-Hill.

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

Plessix, R.-É., G. Baeten, J. W. de Maag, F. ten Kroode, and Z. Rujie, 2012, Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set: Geophysical Prospecting, **60**, 733–747.

Plessix, R.-É., and Q. Cao, 2011, A parametrisation study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: Geophysical Journal International, **185**, 539–556.

Plessix, R.-É., and W. Mulder, 2008, Resistivity imaging with controlled-source electromagnetic data: depth and data weighting: Inverse problems, **24**, 034012, 1–22.

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

Pratt, R. G., and C. H. Chapman, 1990, Frequency domain elastic wave modeling by finite differences: a tool for cross-hole seismic imaging: Geophysics, **55**, 626–632.

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

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

Prieux, V., R. Brossier, S. Operto, and J. Virieux, 2013a, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: imaging compressional wave speed, density and attenuation: Geophysical Journal International, **194**, 1640–1664.

———, 2013b, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 2: imaging compressive-wave and shear-wave velocities: Geophysical Journal International, **194**, 1665–1681.

Rayleigh, J. W. S., 1885, On waves propagated along the plane surface of an elastic solid: Proceedings of the London Mathematical Society, **17**, 4–11.

Richart, F. E., R. D. Woods, and J. R. Hall, 1970, Vibration of soil and foundations: Prentice-Hall.

Rix, G. J., and E. A. Leipski, 1991, Accuracy and resolution of surface wave inversion: Recent advances in instrumentation, data acquisition and testing in soil dynamics: American Society of Civil Engineers, 17–32.

Robertsson, J., 1996, A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography: Geophysics, **61**, 1921–1934.

Romanowicz, B., and J. Durek, 2000, Seismological constrains on attenuation in the earth: a review: Geophysical Monograph, **117**, 161–180.

Romdhane, A., G. Grandjean, R. Brossier, F. Rejiba, S. Operto, and J. Virieux, 2011, Shallow-structure characterisation by 2D elastic full-waveform inversion: Geophysics, **76**, R81–R93.

Saenger, E. H., N. Gold, and S. A. Shapiro, 2000, Modeling the propagation of elastic waves using a modified finite-difference grid: Wave Motion, **31**, 77–92.

Samyn, K., A. Bitri, and G. Grandjean, 2013, Imaging a near-surface feature using cross-correlation analysis of multi-channel surface wave data: Near surface geophysics, **11**, 1–10.

Schäfer, M., L. Groos, T. Forbriger, and T. Bohlen, 2013, 2D Full waveform inversion of recorded shallow surface waves on a significantly 2D structure: Proceedings of the 19th Near Surface Conference. (Bochum, Germany).

Schwab, F., and L. Knopoff, 1970, Surface-wave dispersion computations: Bulletin of the Seismological Society of America, **60**, 321–344.

Sears, T., S. Singh, and P. Barton, 2008, Elastic full waveform inversion of multi-component OBC seismic data: Geophysical Prospecting, **56**, 843–862.

Shafer, M., L. Groos, T. Forbriger, and T. Bohlen, 2004, Surface wave methods for near-surface characterisation: A tutorial: Expanded Abstracts, 19th European Meeting of Environmental and Engineering Geophysics, Near Surface Geoscience. (Bochum, Germany).

Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Early arrival waveform tomography on near-surface refraction data: Geophysics, **71**, U47–U57.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: Geophysical Journal International, **173**, 922–931.

Sirgue, L., 2003, Inversion de la forme donde dans le domaine fréquentiel de donnes sismiques grands offsets. PhD thesis: Université Paris XI.

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

Socco, L. V., D. Boiero, S. Foti, and R. Wisén, 2009, Laterally constrained inversion of ground roll from seismic reflection records: Geophysics, **74**, G35–G45.

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

Socco, L. V., and C. Strobbia, 2004, Surface wave methods for near-surface characterisation: A tutorial: Near Surface Geophysics, **2**, 165–185.

Song, Y. Y., J. P. Castagna, R. A. Black, and R. W. Knapp, 1989, Sensitivity of near-surface shear-wave velocity determination from Rayleigh and Love waves: 50th Annual International Meeting of the Society of Exploration Geophysicists. Expanded Abstracts, **8**, 509–512.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **56**, 785–794.

Stokoe II, K. H., S. Nazarian, G. J. Rix, I. Sanchez-Salinero, J. Sheu, and Y. Mok, 1988, In situ seismic testing of hard-to-sample soils by surface wave method: Earthquake Engineering and Soil Dynamics II - Recent advances in ground-motion evaluation, 264–277.

Strobbia, C., 2002, Surface wave method: acquisition, processing and inversion. PhD dissertation: Politecnico di Torino.

Strobbia, C., P. Vermeer, A. Laake, A. Glushckenko, and S. Re, 2010, Surface waves:

processing, inversion and removal: First Break, **28**, 85–91.

Taillandier, C., M. Noble, H. Chauris, and H. Calandra, 2009, First arrival travel time tomography based on the adjoint-state method: Geophysics, **74**, WCB57–WCB66.

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

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

———, 2005, Inverse problem theory and methods for model parameter estimation: Society for Industrial and Applied Mathematics.

Tarrass, I., L. Giraud, and P. Thore, 2011, New curvilinear scheme for elastic wave propagation in presence of curved topography: Geophysical Prospecting, **59**, 889–906.

Tessmer, E., D. Kosloff, and A. Behle, 1992, Elastic wave propagation simulation in the presence of surface topography: Geophysical Journal International, **108**, 621–632.

Thompson, D. O., and D. E. Chimenti, 1997, Review of progress in quantitative nondestructive evaluation: Springer.

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

Thomson, W. T., 1950, Transmission of elastic waves through a stratified solid medium: Journal of Applied Physics, **21**, 89–93.

Thrower, E. N., 1965, The computation of the dispersion of elastic waves in layered media: Journal of Sound and Vibrations, **2**, 210–226.

Tian, G., D. W. Steeples, J. Xia, and K. T. Spikes, 2003, Useful resorting in surface-wave method with the Autojuggie: Geophysics, **68**, 1906–1908.

Tikhonov, A., and V. Arsenin, 1977, Solution of ill-posed problems: Winston.

Tselentis, G. A., and G. Delis, 1998, Rapid assessment of S-wave profiles from the inversion of multichannel surface wave dispersion data: Annali di Geofisica, **41**, 1–15.

Tsvankin, I., 2005, Seismic signatures and analysis of reflection data in anisotropic media, 2nd edition: Elsevier.

Ulrych, T. J., M. D. Sacchi, and J. M. Graul, 1999, Signal and noise separation: Art and science: Geophysics, **64**, 1648–1656.

van Leeuwen, T., and W. Mulder, 2008, Velocity analysis based on data correlation: Geophysical Prospecting, **56**, 791–803.

Vignoli, G., C. Strobbia, G. Cassiani, and P. Vermeer, 2011, Stastistical multioffset phase analysis for surface-wave processing in laterally varying media: Geophysics, **76**, U1–U11.

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

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

Williamson, P., 1991, A guide to the limits of resolution imposed by scattering in ray tomography: Geophysics, **56**, 202–207.

Woodhouse, J. H., and A. M. Dziewonski, 1984, Mapping the upper mantle: three dimensional structure of Earth's structure by inversion of seismic waveforms: Journal of Geophysical Research, **89**, 5953–5986.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

Xia, J., R. D. Miller, and C. B. Park, 1999, Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves: Geophysics, **64**, 697–700.

Xia, J., R. D. Miller, C. B. Park, and G. Tian, 2003, Inversion of high frequency surface waves with fundamental and higher modes: Journal of Applied Geophysics, **52**, 45–57.

Xia, J., Y. Xu, and R. D. Miller, 2007, Generating an image of dispersive energy by frequency decomposition and slant stacking: Pure and Applied Geophysics, **164**, 941–956.

Xu, Y., J. Xia, and R. D. Miller, 2007, Numerical investigation of implementation of air-earth boundary by acoustic-elastic boundary approach: Geophysics, **75**, SM147–SM153.

Yilmaz, O., 2001, Seismic data processing: Society of Exploration Geophysicists.

Zhang, J., 1997, Quadrangle-grid velocity-stress finite-difference method for elastic-wave-propagation simulation: Geophysical Journal International, **131**, 127–134.

Zhang, W., and X. Chen, 2006, Traction image method for irregular free surface boundaries in finite difference seismic wave simulation: Geophysical Journal International, **167**, 337–353.

Imagerie sismique de la proche sub-surface : modification de l'inversion des formes d'onde pour l'analyse des ondes de surface

RESUME

L'amélioration des images sismiques peut aider à mieux contraindre l'exploration des hydrocarbures. Les ondes élastiques qui se propagent dans la Terre peuvent être classifiées comme ondes de volume et ondes de surface. Si ces dernières sont les plus énergétiques, seules les ondes de volume sont couramment considérées comme des signaux utiles. Cependant, les ondes de surface sont utiles pour caractériser la proche sub-surface. Classiquement, les ondes de surface sont analysées dans des contextes de propriétés élastiques localement 1D.

Nous proposons une modification de l'inversion des formes d'onde classique pour reconstruire des profils de propriétés 2D (la windowed-Amplitude Waveform Inversion, w-AWI). La w-AWI est spécialement robuste en ce qui concerne le choix du modèle initial. Nous appliquons la w-AWI aux données synthétiques ainsi qu'aux données réelles, montrant que cette approche permet de récupérer des propriétés 2D.

Mots clés : ondes de surface, problème inverse, modélisation sismique, inversion des formes d'onde.

Two-dimensional near-surface seismic imaging with surface waves: alternative methodology for waveform inversion

ABSTRACT

High-resolution seismic imaging is essential to improve results of hydrocarbon exploration. Elastic waves propagate in the Earth as body and surface waves, the latter being the most energetic ones. Body waves are preferred for exploration seismic imaging while surface waves are usually considered to be noise. However, it has been recognised that the near surface can be characterised by analysing surface waves and that such result may improve the outcome of body-wave processing. Currently, surface waves analysis leads to retrieve local 1D property profiles.

We propose a waveform-based inversion procedure to derive 2D velocity models from surface waves. This method consists of a misfit functional modification of classical Full Waveform Inversion and we call it windowed-Amplitude Waveform Inversion (w-AWI). We show that w-AWI is robust regarding the choice of initial velocity model. We apply w-AWI to synthetic and real data obtaining encouraging near-surface imaging results.

Key words: surface waves, inverse problem, seismic modelling, waveform inversion.