

Challenges in near-surface seismic full-waveform inversion of field data

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

M.Sc. Nikolaos Athanasopoulos

Tag der mündlichen Prüfung:

06.11.2020

Hauptreferent:

Prof. Dr. Thomas Bohlen

Korreferent:

Prof. Dr. Hansruedi Maurer

Abstract

With the ever-increasing need for natural resources, the demand for detailed subsurface imaging of the Earth's interior has become more profound than ever. Among the plethora of geophysical methods, the study of seismic waves has proven itself as the primary way to extract detailed information of various subsurface scales. By deploying actively induced or passive sources and measuring the ground motion at various locations of the Earth's surface, scientists were able to delineate the relationship of wave propagation with regard to changes of the lithology inside the Earth's interior. Seismic theory constitutes also the majority of research related to the exploration of oil and gas. However, continuous depletion of oil resources and rise of renewable energies motivate the search for new applications for geophysical know-how originating from hydrocarbon exploration. A particular area where geophysical methods are required belongs to near-surface site investigations. Applications ranging from geotechnical soil characterization to targets of archaeological interest as well as water-rich environments have been the center of focus of such studies. However, in such targets seismics has been usually rejected due to the poor resolution and the lack of petrophysical relationships to important subsurface parameters that other methods provide at ease.

With the technological developments that came along during the 21st century, many techniques that were proven computationally prohibitive have found a new place in the highest ranks of advanced geophysical methods. Such is the case for full-waveform inversion, a technique that was introduced more than four decades ago, and was kept at hold until the technological improvements reached its demands. Full-waveform inversion (FWI), as the name suggests, uses the full information contained in seismic recordings to produce subsurface models of sub-wavelength resolution.

This thesis contributes to developments in modeling and inversion of shallow seismic data, through the development of sophisticated strategies and frameworks. The fundamental theory of scattered wavefields in the presence of strong surface waves is illustrated by a series of numerical experiments. The findings of this study give insights into the composition of shallow seismic recordings and provide a framework for their interpretation in future applications of FWI. In particular, it is shown that the scattering responses of attenuation and velocities for both P and S waves are very similar, which constitutes the main reason in the development of interparameter trade-off when preforming viscoelastic FWI. On the other hand, the scattering due to density perturbation has a unique signature in near surface applications. This implies that in such cases the retrieval of accurate density model is feasible, opposite to the experience acquired from the exploration industry.

After delineating the events that can arise from various subsurface structures, the application of seismic FWI from 3-component data is demonstrated here for the first time. Both synthetic and field studies are examined and reveal the benefits of multicomponent data for multiparameter FWI. Contrary to the main trends in shallow seismic research, both Rayleigh and Love waves are heavily employed to improve the subsurface characterization, which is revealed through synthetic studies. It is demonstrated that the

spatial connectivity between subsurface models is further enhanced through the enforcement of structural similarity constraints and joint inversion approaches with independent geophysical data.

Conventional multiparameter elastic FWI fails to reconstruct the P-wave velocity model due to its low sensitivity to Rayleigh waves which dominate the misfit. A key objective is thus to enable the accurate reconstruction of the P-wave velocity model. This is achieved through the development of a time-frequency windowing strategy, by gradually increasing the contribution of later arrivals in the course of the iterations. This approach shows improved reconstruction of all three model parameters, especially of the P-wave velocity and density model.

As a preparation for evaluating the techniques which are developed in this thesis, a preprocessing workflow is shown which can serve as a general guide for the application of FWI on field data and facilitate the extraction of reliable subsurface models. Two main target areas are evaluated from survey planning and sophisticated acquisition techniques to the retrieval of high-resolution subsurface models.

The application of 2D elastic full-waveform inversion of individual Rayleigh and Love waves is first demonstrated using a near-surface survey near Karlsruhe, Germany. The aim is to recover the subsurface information on an archaeological buried structure. The time-frequency windowing approach is applied to the field data set where it allows the reduction of the misfit function and retrieval of models of higher correlation compared to the conventional approaches. This is also the case when superimposing the velocity structure with the migrated results of independent ground-penetrating measurements that were performed on the site.

Secondly, the benefits of FWI are exploited as an effort to improve the aquifer characterization for better understanding of its properties and delineating its structure. By making use of a prismatic source, better coupling and coherency in the recordings allow for proper execution of joint inversion using data acquired from all three Cartesian coordinates in the test site near Krauthausen, Germany. For the first time, a high correlation of the density model from shallow seismic FWI with in-situ borehole information was obtained. The complex structure of the aquifer is revealed and can serve as a valuable source of information for further hydrogeophysical site characterization.

The developments presented here 1) establish a better understanding of surface wave scattering, 2) demonstrate advances towards multiparameter and multicomponent FWI, and 3) provide a complete workflow of FWI in near-surface field data applications.

Acknowledgments

*‘Σὰ βγεῖς στὸν πηγαμὸ γιὰ τὴν Ἰθάκη, νὰ εὔχεσαι νᾶναι μακρὺς ὁ δρόμος, γεμάτος
περιπέτειες, γεμάτος γνώσεις.....Κι ἂν πτωχικὴ τὴν βρῆς, ἡ Ἰθάκη δὲν σὲ γέλασε.
’Ἐτσι σοφὸς ποὺ ἔγινες, μὲ τόση πείρα, ἥδη θὰ τὸ κατάλαβες ἡ Ἰθάκης τί σημαίνουν’
— Κ.Π. Καβάφης, Ἰθάκη*

As my PhD journey reaches its end, I would like to take the time and thank everyone who contributed in one way or another on completing this dissertation, but also accompanying and supporting me throughout the whole time. This dissertation represents the end of this journey, but my main achievement is the personal development from the experiences and people that I met during this time, for without them I would not be able to reach my destination (Ἰθάκη).

First of all, I want to express my sincere gratitude to Thomas, for supervising and supporting my PhD project. His academic transparency and integrity, his enthusiasm on improving existing research and his broad knowledge and focus on detail inspired me to become a better scientist. Thank you for your continuous support through the last five years; it has been an honor working with you.

I am also extremely grateful to Hansruedi, for his supervision during my visit in ETH and for his involvement in my PhD project. It has been a great experience working with him and an inspiration for the last seven years since we first met. On this regard, I would also like to thank Edgar Manukyan, for his insights and suggestions during and after my visit in ETH, that greatly improved my thesis.

I am very grateful to Thomas Hertweck for his support, particularly during the writing phase of my thesis. His always timely input and suggestions helped me to keep on track and greatly improved the final outcome.

A special thanks to Claudia for her devotion throughout the whole period of my PhD, always taking care of any administration task but also for moral support during the course of this dissertation. I would like to express my gratitude to Thomas Bohlen, Thomas Hertweck, Laura, Michael Bachmann and Claudia for proof reading this work. Thanks to Petra, Thomas Hertweck and Michael Frietsch for the technical support.

Thanks to the Karlsruhe House of Young Scientists for financing my research visit to ETH Zürich, through the research travel grant. I would also like to thank Daniel Köhn for many conversations that we had throughout the EAGE conferences.

Thanks to all current and former members of the GPI group and in particular my fellow colleagues throughout these years, Tilmann, Florian, Leo, Martin, Mario, Renat, Tao, Mark, Laura, Niklas, Tingting, Adam, Mike, Michael, Andre, Lars - you made those years really enjoyable. Thanks to Mario for sharing the office with me and for our discussions during good and also stressful times. A very special thanks to Tilmann for being a great friend and support during the whole time. I will not forget our smoking

breaks although I hope to join you in quitting. Also thank you for our discussions on the derivations of the adjoint state method. I would like to thank Florian for his support in the early stage of my project and his patience, that greatly helped me catch up fast with the FWI software.

Thanks to the EEG group in ETH that made me feel welcome and included me in their activities. You made my stay one of the best experiences during my PhD. A special thanks to Nele, Patrick, Teo, Emily, Edu and Adrian for all the good times we had and will have in the future.

Thanks to my dear friends from the studies – Themis, Alex, Giannis, Manos, Vasilis, Panos, Nikos, Marianna, Kelly, Nicol, Vaggelio – for sticking around with me all these years and always supporting me.

Last but not least, I thank my family and Michael, for their continuous encouragement and support, emotional and mental, and for being always by my side when I needed them the most.

Contents

Abstract	i
Acknowledgment	iii
1 Introduction	1
1.1 General overview	1
1.2 Motivation and challenges	2
1.3 Overview	5
2 Theoretical background	7
2.1 Forward problem	8
2.1.1 Elastic wave equation	8
2.1.2 Viscoelasticity	12
2.1.3 Numerical representation of wave equation	14
2.2 Full-waveform inversion	17
2.2.1 Objective function of the inverse problem	17
2.2.2 Adjoint-state method	18
2.3 Numerical and physical aspects during FWI	23
2.3.1 Optimization method	23
2.3.2 Preconditioning	23
2.3.3 Multi-test step length estimation	24
2.3.4 Misfit definitions	24
2.3.5 Multi-scale approach	25
2.4 FWI workflow	26
3 Numerical study of seismic scattering	29
3.1 Scattering principles in shallow seismics	30
3.1.1 Isotropic body wave scattering	30
3.1.2 Scattering in heterogeneous media: Born approximation versus FD modelling	32
3.1.3 Viscoelastic wave equation of scattered wavefield	32
3.1.4 Instantaneous energy density of P- and S-waves	33
3.1.5 Numerical setup	35
3.2 Scattering effects caused by perturbations in the elastic parameters	35
3.2.1 Model description	35
3.2.2 Results	36
3.2.3 Conclusions	38
3.3 Dependency of scattering effects on inclusion characteristics	38

3.3.1	Influence of the depth of the inclusion	40
3.3.2	Influence of the size of the inclusion	40
3.3.3	Influence of the shape of the inclusion	40
3.3.4	Influence of the composition	42
3.3.5	Conclusions	42
3.4	Summary	42
4	Synthetic studies	43
4.1	Multiparameter inversion	44
4.1.1	Simultaneous joint inversion	44
4.1.2	Joint objective function of P-SV and SH waves	45
4.1.3	Joint gradients	45
4.1.4	Numerical setup and model description	46
4.1.5	Spatially correlated models	47
4.1.6	Spatially uncorrelated models	50
4.1.7	Influence of the initial P-wave velocity model	54
4.1.8	Conclusions	56
4.2	Time - frequency windowing	56
4.2.1	Theory and methodology	57
4.2.2	Effect of the taper function on frequency spectra	58
4.2.3	Effect of the taper function with offset	59
4.2.4	Synthetic reconstruction test	60
4.2.5	Results	60
4.2.6	Conclusions	64
4.3	Parameter sensitivity in far offset acquisition setup	64
4.3.1	Model description	64
4.3.2	Acquisition setup	66
4.3.3	Results	67
4.3.4	Conclusions	69
4.4	Structural similarity constraints	69
4.4.1	Theory of cross-gradients	70
4.4.2	Synthetic study and results	72
4.4.3	Conclusions	73
4.5	Summary	74
5	Preprocessing of field data: General overview	75
5.1	Data acquisition	75
5.2	Preprocessing workflow	76
5.3	Initial model building	78
5.3.1	Traveltime tomography	79
5.3.2	Multichannel analysis of surface wave (MASW)	79
5.3.3	Q estimation	80
5.3.4	Cycle skip evaluation	81
5.4	Source time function inversion	82

5.5	Summary	84
6	Field survey: archaeological investigation	85
6.1	Test site	86
6.2	Acquisition geometry and pre-processing	86
6.3	Initial models and FWI set-up	88
6.4	Results	89
6.4.1	Conventional versus time-frequency FWI	89
6.4.2	Inline P-SV inversion	93
6.4.3	SH inversion	95
6.5	Comparison with ground-penetrating radar	96
6.6	Summary	98
7	Field survey: aquifer characterization	99
7.1	Introduction	100
7.2	Study area	101
7.3	Prismatic source characterization	102
7.4	Acquisition set-up	105
7.5	Field data analysis and FWI parametrization	107
7.5.1	Preprocessing of field data	107
7.5.2	FWI set-up	107
7.6	Results	109
7.6.1	FWI results of profile 1 and 2	110
7.6.2	Evaluation of FWI density model with CPT measurements	116
7.6.3	Cross-section of P1 and P2	117
7.7	Summary	117
8	Summary and Conclusions	121
A	Individual scattering response for various inclusion cases	125
B	Cross-gradient implementation	129
C	Steel prism	131
D	Dispersion curves: profiles P1 and P2	133
E	Software and Hardware	135
	Bibliography	137

Chapter 1

Introduction

1.1	General overview	1
1.2	Motivation and challenges	2
1.3	Overview	5

1.1 General overview

Over the last century, the growing need for natural resources has increased the interest for subsurface imaging and characterization. Geophysical methods constitute the main tool for subsurface investigations, since they can provide information not only in regard to energy resources, i.e., oil and gas, but also for studying environmental, engineering or even archaeological targets. Additionally, geotechnical engineering demands detailed knowledge of the subsurface for creating proper foundations as part of various construction projects, which is also offered by geophysical studies. A key challenge for near-surface applications is the lithological sediment and aquifer characterization. These tasks require quantitative imaging of the topmost meters of the subsoil. With the groundwater constituting 98% of the Earth's unfrozen freshwater, the need for protection from anthropogenic pollution and ever-increasing water demands is of paramount importance (UNESCO, 2005).

With the continuous depletion of oil resources, it becomes apparent that there is a requirement for shifting the focus of interest and transferring the knowledge that was gained from hydrocarbon exploration into different fields. From the various geophysical methods, one of the most commonly used for subsurface imaging represents the study of seismic waves. Seismic surveying uses mechanical waves to map subsurface structures across a vast range of spatial scales. Conventional seismic methods are classified based on the specific wave types that are used to obtain subsurface models through inversion techniques. However, they are limited in the sense that they use only a fraction of the recorded wavefield.

About four decades ago, Lailly (1983) and Tarantola (1984) introduced the concept of full-waveform inversion (FWI). Full-waveform inversion has the potential to reveal subsurface structures without limitations regarding the subsurface heterogeneity, by making use of the whole information content included in seismic waves (i.e., amplitude and phase).

Improvements in computational resources have defined FWI as one of, if not the most, advanced techniques in seismic imaging, making its debut in various fields. Research areas range from medical imaging (Pratt

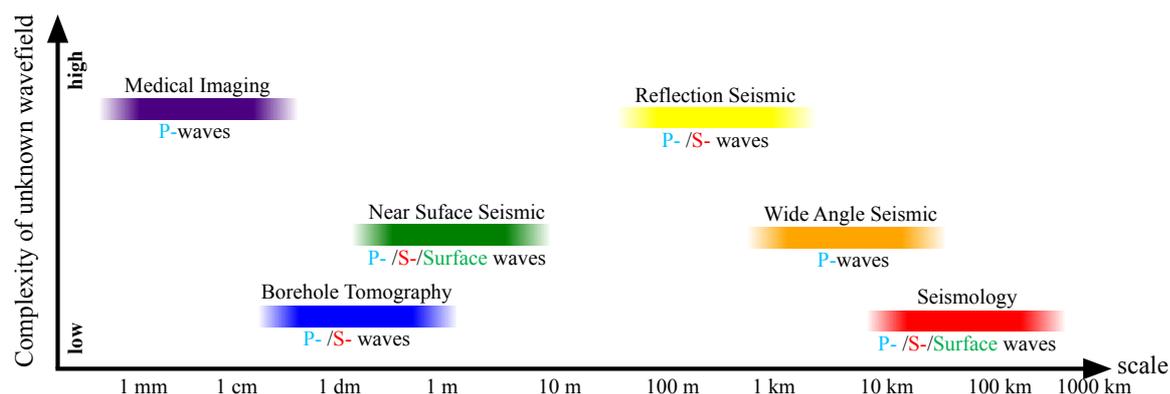


Figure 1.1: FWI applications over the last 20 years including a broad range of scales. The different applications are expressed as a function of complexity, spatial scale and wave types that are considered during the inversion. The color coding is shown in analogy to wavelengths of the optical electromagnetic spectrum.

et al., 2007), borehole tomography (Manukyan and Maurer, 2020), near-surface seismic (Gélis et al., 2007; Romdhane et al., 2011; Groos et al., 2017; Athanasopoulos et al., 2020), reflection seismic (Xu et al., 2012; Brossier et al., 2015), seismic exploration (Operto et al., 2004; Brossier et al., 2009; Gassner et al., 2019), wide-angle seismic (Morgan et al., 2013) and seismology (Bleibinhaus et al., 2007; Fichtner et al., 2009). Each field targets a different spatial scale and bears different complexity in terms of the recorded wavefield, as illustrated in Figure 1.1.

Near-surface seismic applications are highly attractive for various investigations, including geotechnical and hydrogeophysical characterization of the subsurface. For such surveys the recorded wavefield contains a vast variety of different wave types. Surface waves and compressional P-waves constitute the main wave types in such applications and have been applied extensively over the last decades due to their easy excitation and low acquisition cost. Surface waves which are generated in the presence of a free boundary (Earth's surface) consist of different types, which are classified based on the polarization of the ground motion during propagation.

The estimation of the S-wave velocity of the upper 30 m is entirely done by using surface waves and represents an important geotechnical parameter for site classification. Additionally, identification and proper processing of surface waves can be further used in exploration scales to either completely remove them (for seismic migration) or use them to improve the near-surface velocities.

Unlike marine studies, land seismics are particularly complex, due to the high sensitivity of the recorded signal to near-field effects, the presence of elastic waves (P, S and surface waves) and severe attenuation, which can lead to severe distortion of the wavefield (Foti et al., 2018). The same effect is also present in specific marine seismic studies, where ocean bottom surveys are employed. These effects are enhanced when dealing with short-offset acquisitions in the presence of surface waves.

1.2 Motivation and challenges

Common methods for extracting information from shallow seismic surface waves consist of the inversion of (local) dispersion curves or full slowness-frequency spectra (Dziewonski et al., 1969; Park et al., 1999;

Xia et al., 1999; Forbriger, 2003a; Socco et al., 2010). However, the main assumption of these techniques is that the subsurface properties vary mainly with depth. Therefore, they cannot properly account for situations where strong lateral variations exist. A solution to this problem is provided by FWI, since it makes use of all available information in the recorded wavefield possible. However, in shallow seismic applications FWI faces certain limitations due to the complexity of the recorded wavefield, caused by the interference of multiple surface wave modes and body waves (S and P waves).

Dominance of surface waves

The amplitudes of surface waves are primarily determined by the shallow S-wave velocity structure and only weakly by the P-wave velocity model. The amplitudes of P and S waves are primarily influenced by the corresponding velocity models. Mass density affects all wave types in different ways. The amplitudes of surface waves are generally much higher than the amplitudes of P waves. In such situations FWI fails to properly incorporate the contribution of P waves, as the misfit is dominated by the high-amplitude Rayleigh waves (Groos, 2013; Wittkamp et al., 2018). Most of the past research focused on the inversion of Rayleigh waves only (Tran et al., 2013; Groos, 2013; Schäfer, 2014; Athanasopoulos and Bohlen, 2017b) to reconstruct the S-wave velocity structure. However, the sensitivity of Rayleigh waves to P-wave velocity and density is relatively weak and therefore both P-wave velocity and density are not accurately retrieved.

The accurate estimation of a P-wave velocity model is crucial for determining the depth of geological interfaces, assessments of landslide environments, improving the S-wave velocity model (Groos, 2013) and interpreting aquifer systems with respect to saturation. Additionally, Wittkamp et al. (2018) applied joint FWI of Rayleigh and Love waves and have shown that unless an accurate P-wave velocity model is used, the joint inversion will be highly contaminated by inaccuracies of the Rayleigh wave inversion. In such cases the Love wave inversion would be more beneficial.

In the past few years, several approaches have been deployed to overcome the different contribution of surface and P waves in both the data and the model domain (Plessix and Mulder, 2004b; Nuber et al., 2015; Gibson et al., 1979; Brenders and Pratt, 2007). One of the main targets of this research lies on improving the reconstruction of the elastic parameters and especially addressing the problems in regard to the resolution and sensitivity of the P-wave velocity model. This will further allow to reduce the co-dependency of the P-wave velocity model on the S-wave velocity, which can lead to an overall improvement of FWI in shallow seismic applications.

Proper acquisition of multicomponent seismic data

The initial and most important step for any geophysical method is the acquisition of field data. A proper geological interpretation demands a suitable design of data acquisition, which will dictate the quality of subsurface information that can be retrieved (Maurer et al., 2010). Additionally, the quality of the recorded signal is crucial for the structural image of the subsurface. An advantage of performing shallow seismic inversion and processing in the presence of surface waves is that they can be easily excited and

recorded with standard field equipment such as a sledge-hammer. Their large amplitude provides a very high signal-to-noise ratio making their distinction and therefore processing more straightforward.

However, in cases where multicomponent data are required, proper coupling of the sources and receivers is of extreme importance. That is especially the case for techniques involving the simultaneous processing or inversion of multicomponent seismic data. This can be enabled by making use of more sophisticated seismic sources that can ensure the quality of the recordings while maintaining a uniform coupling. This topic will play a significant role in this thesis and the challenges but also successes that arise, through proper usage of multicomponent data, will be discussed in detail.

Reliability of inversion results and joint strategies

One advantage of near-surface seismic applications is the ease of acquisition of multiple independent data, often from various geophysical techniques. This is due to the relatively small size of the seismic surveys and the shallow depth of investigation, which makes the joint interpretation from various methods cost effective. Recent studies in multimethod subsurface imaging problems have shown that enforcement of structural similarity between different subsurface model parameters improves the tomographic reconstructions (Gallardo and Meju, 2003; Tryggvason and Linde, 2006; Doetsch et al., 2010). Manukyan et al. (2018) have shown that a structural similarity constraint is beneficial also for elastic seismic FWI. The methodology of the structural similarity constraint will be further discussed in the context of this work, along with its potential to enhance the resolution of the reconstructed models.

While typically most research regarding surface waves has been done using Rayleigh waves, the addition of Love waves has recently gained great interest. The main reason is that Love waves are propagating independently of the P-wave velocity, which as I mentioned represents a difficult parameter to obtain in sub-wavelength resolution. However, the acquisition of Love waves is in general more challenging due to the fact that a shear force is required and coupling can become less uniform. A main objective of this work will be to assess any potential benefits that can be obtained from FWI of both individual but also joint inversion of Rayleigh and Love waves.

The solution of the FWI problem is highly non-linear and demands sufficient low wave-numbers in the initial model or the observed data to avoid getting trapped in local minima. Since most techniques that are used consist of local optimization techniques, obtaining an uncertainty estimation is extremely difficult. Therefore, using independent data for comparison or joint interpretation is essential in order to obtain reliable results. For that reason, the results obtained in the field studies of this work, will be further compared with independent geophysical data and in-situ borehole information.

Unconventional targets

While seismic methods are often used in geotechnical site characterization, they can provide only a smoothed version of the actual subsurface structure. Due to this limitation many studies have excluded the application of seismic surveys for specific target areas, such as the detection of buried structures of archaeological interest and aquifer systems. In this work, the potential of the application of FWI to retrieve the structural architecture of an aquifer using seismics is investigated. In all previous studies the

main geophysical techniques included tracer experiments (Vereecken et al., 2000), cone penetration tests (Tillmann et al., 2008) and ground-penetrating radar measurements (Klotzsche et al., 2013; Gueting et al., 2015) to deduct information in such subsurface systems.

However, FWI facilitates the use of seismics in such target areas and can provide sub-wavelength resolution. The main objective of this research is the applicability of seismic FWI for such unconventional targets. Problems that arise will be discussed and comparison with other geophysical methods will provide insights for proper application of FWI in these specific cases.

1.3 Overview

Each chapter includes a specific introduction, a description of the methods used, and a discussion of the obtained results or observations. The thesis is organized as follows:

The theoretical basis and numerical implementation of forward and inverse problems, associated with elastic and viscoelastic wave propagation, are explained in detail in **Chapter 2**.

Chapter 3 deals with the fundamentals of scattering due to a subsurface inclusion in the presence of surface waves through a series of numerical experiments. It will provide the necessary knowledge to understand the different contributions of the parameter space to the actual recordings.

In **Chapter 4** a series of synthetic tests are shown. Different techniques will be employed in controlled synthetic studies and will be analyzed in terms of their theoretical background and applicability. Since the true subsurface model in these numerical tests is known, the results from FWI can be directly evaluated. The principles of joint inversion of seismic and non-seismic geophysical methods will be discussed. Additionally, a time-frequency windowing approach will be included in the FWI framework in order to improve the reconstruction of the P-wave velocity models.

Chapter 5 serves as a complete guide for preparing field seismic data for any seismic inversion scheme. It consists of essential steps for or prior to the inversion workflow which otherwise would most likely fail to obtain realistic results.

The application of FWI to field data is discussed in **Chapters 6** and **7**. The two chapters deal with the applicability of FWI for archaeological and hydrogeological site characterization. The techniques that are developed in the previous chapters are selectively chosen in each of the two field data cases, depending on specific requirements/limitations that will be analyzed in detail. Finally, since FWI is a non-linear problem and in order to assess the quality of the retrieved model parameters, a qualitative comparison with independent data is performed.

Conclusion and perspectives are provided in **Chapter 8** of this thesis.

On the Structure of this thesis

Chapters 3-7 have either been published or are being prepared for publication. However, they are self-contained only to a certain extent in order to avoid redundancies. This is intended to improve the reading experience of this thesis as a whole.

Chapter 2

Theoretical background

2.1	Forward problem	8
2.1.1	Elastic wave equation	8
2.1.2	Viscoelasticity	12
2.1.3	Numerical representation of wave equation	14
2.2	Full-waveform inversion	17
2.2.1	Objective function of the inverse problem	17
2.2.2	Adjoint-state method	18
2.3	Numerical and physical aspects during FWI	23
2.3.1	Optimization method	23
2.3.2	Preconditioning	23
2.3.3	Multi-test step length estimation	24
2.3.4	Misfit definitions	24
2.3.5	Multi-scale approach	25
2.4	FWI workflow	26

In order to characterize a medium, a series of physical measurements is performed. The measured quantity which is obtained can in general be anything from an electromagnetic wave, to a magnetic response or an elastic wave. The inversion approach which is inherited in this thesis and which was discussed in the previous chapter, is not exclusive for only mechanical waves, which are the focus of this work, but constitutes a general process for reconstructing physical parameters given a set of governing laws that predict their behavior in nature. This definition is the realm of what we call an inversion.

The theory that predicts the physical parameters from a given model corresponds to the forward problem. In contrast, retrieving the true model from a set of data measurements refers to the inverse problem. Certain general issues that arise by making use of these two definitions are related to the fact that the solution of the inverse problem is often connected to a solution of a highly non-linear problem. Therefore, iterative methods and linearized approximations have to be considered in order to obtain the set of model parameters that we are interested in. Additionally, we solve the inverse problem by defining a measure of the discrepancy between measured data and predicted data, leading to a minimization problem of a cost function. Thus, the general inversion scheme includes three essential elements: a physical theory, a numerical solution of the governing equations and an optimization algorithm.

For the purpose of this work the three above definitions correspond to the wave equation in viscoelastic media, a finite-difference time-domain scheme for its discretization and a local optimization algorithm for the minimization of the cost function.

This chapter addresses these three topics and shows the theoretical background required to perform each step. Additional challenges that arise from this process are highlighted and discussed with the aim to find an optimal inversion workflow.

2.1 Forward problem

I will initially introduce the elastodynamic equations which form the fundamental basics for wave propagation in elastic media. Afterwards, I will describe the shift from the continuous wave equation to its discretized counterpart in order to facilitate the numerical modelling of wave propagation.

2.1.1 Elastic wave equation

Initially, I consider wave propagation in homogeneous and isotropic elastic media. The parameters related to attenuation will be considered in the next section. There are two equations that constitute the basic principles of wave propagation in arbitrary elastic bodies: Newton's second law and Hooke's law. The next section will follow closely Lay and Wallace (1995) and Aki and Richards (2002).

Hooke's law: stress-strain theorem

General elasticity provides us with the concept of deformation which an elastic body will undergo when an external force is applied. The deformation which is described by the second-order symmetric tensor ϵ , is given by

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad (2.1)$$

where u is the displacement of a particle in direction of the i^{th} or j^{th} dimension of a Cartesian coordinate system, with location vector x . While the diagonal components of ϵ correspond to extension changes, the off-diagonal components describe the shear deformations. Given the current deformation acting on a body we can further determine the stress, describing traction that is acting on the surface of a body, which is also a second-order symmetric tensor. Once again the distinction between the diagonal and off-diagonal components of the stress tensor describe the normal and shear stresses, respectively.

The generalized Hooke's law states that for infinitesimally small deformations there can be assumed a linear relation between the stress tensor, σ , and the strain tensor, ϵ , acting on an elastic medium which is given by

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl}, \quad \text{where } i, j, k, l \in [1, 3], \quad (2.2)$$

using Einstein notation. While stress and strain are second-order tensors, the stiffness tensor C is of fourth order which means that it contains 81 components. Due to the symmetry between σ and ϵ , the components of C reduce to 36 and from energy conservation laws, it is then reduced further to 21 independent components. When considering isotropic media we can reduce C to two independent parameters, known as the Lamé parameters, λ and μ (shear modulus), leading to

$$\sigma_{ij} = \lambda\theta\delta_{ij} + 2\mu\epsilon_{ij} \quad (2.3)$$

$$\text{with } \theta = \epsilon_{11} + \epsilon_{22} + \epsilon_{33} \quad (2.4)$$

$$\text{and } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}, \quad (2.5)$$

where θ is the dilatation describing the relative variation of the volume and δ_{ij} is the Kronecker delta. The Lamé parameters in equation 2.3 are related to the seismic velocities of P and S waves, v_p and v_s , respectively, by

$$v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}} \quad \text{and} \quad v_s = \sqrt{\frac{\mu}{\rho}}. \quad (2.6)$$

In equations 2.3 and 2.6 both the Lamé parameters and the mass density, ρ , are considered time-invariant.

Newton's second law: equation of motion

Another basic concept underlying elastic wave propagation is described by Newton's second law, also known as the equation of motion. First we consider an arbitrary finite volume V with bounding surface S inside a body that can exhibit deformation, which gives us

$$\int_V \rho \frac{\partial^2 \vec{u}}{\partial t^2} dV - \int_V \vec{f} dV = \int_S \vec{T}(\vec{n}) dS. \quad (2.7)$$

The first term on the left-hand side of 2.7 represents the inertial forces of the volume V , while the second term the external body forces, \vec{f} , acting on a unit volume of V . The right-hand side, known as traction, \vec{T} , represents the surface forces acting on a unit area of S , and can be expressed by using the stress tensor as

$$T_i = \sigma_{ij} n_j, \quad (2.8)$$

where \vec{n} is of unit length and points outward normal to the surface S . What equation 2.7 tell us is that the sum of the volume force density and surface forces have to be in equilibrium.

From equations 2.7 and 2.8 and after applying the Gauss' divergence theorem to transform the surface to a volume integral, we get,

$$\int_V \left(\rho \frac{\partial^2 u_i}{\partial t^2} - f_i \right) dV = \int_V \frac{\partial \sigma_{ij}}{\partial x_j} dV. \quad (2.9)$$

Since this expression has to be independent of the volume V for the equilibrium condition to be fulfilled, the integrand has to be zero and after rewriting the terms we get the common equation of motion,

$$\rho \frac{\partial v_i}{\partial t} = f_i + \frac{\partial \sigma_{ij}}{\partial x_j}. \quad (2.10)$$

where \vec{v} is the particle velocity instead of the displacement in equation 2.9. We then take the time derivative of equations 2.3 and 2.1 leading to,

$$\frac{\partial \sigma_{ij}}{\partial t} = \lambda \frac{\partial \epsilon_{rr}}{\partial t} \delta_{ij} + 2\mu \frac{\partial \epsilon_{ij}}{\partial t}, \quad (2.11)$$

$$\frac{\partial \epsilon_{ij}}{\partial t} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right). \quad (2.12)$$

All three equations 2.10, 2.11 and 2.12 can be used to adequately describe wave propagation in an isotropic linear elastic medium.

Body and surface waves

Figure 2.1 illustrates how the different wave-types propagate through a medium, in terms of polarization with respect to the direction of propagation, but also particle motion (Müller and Weber, 2007). This sketch

helps us understand how we can proceed when recording the various wave types in field applications, but also how to decouple the elastic wave equation into individual wave equations.

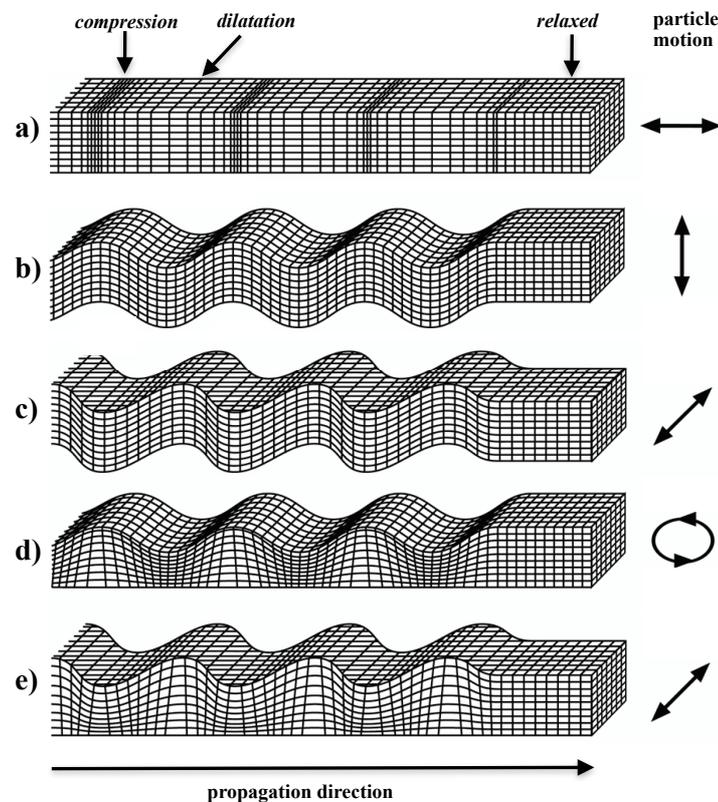


Figure 2.1: Types of seismic waves and particle motion as function of propagation time: a) compressional (P) wave traveling in a block of material, b) vertical polarized and c) horizontal polarized shear wave traveling in a block of material, d) Rayleigh wave traveling in a section of the Earth's surface, e) Love wave traveling along a section of the Earth's surface (modified after Bolt and Freeman, 1989).

Figure 2.1a shows that P waves involve both volume change and shearing, while the particle motion has a longitudinal polarization. On the other hand, the propagation of S waves does not involve volume change and the particle motion that occurs is always perpendicular (transverse) to the direction of propagation. Two cases that arise for the S-wave propagation, namely SV and SH waves, are due to their oscillation on the vertical or horizontal plane with respect to the direction of the propagation (Figures 2.1b and 2.1c, respectively).

In the presence of a free surface, as in the case of the Earth, other solutions of the seismic wave equation are also possible, which are known as surface waves. In the context of this work, I will focus only on Rayleigh and Love waves.

Figure 2.1d corresponds to the propagation of Rayleigh waves. Rayleigh waves originate from coupled P and SV waves and are therefore polarized in the vertical (SV) plane of propagation. The particle motion at the surface is elliptical and retrograde (counter clockwise), due to the phase shift between P and SV waves (Aki and Richards, 2002). On the other hand, Love waves (Figure 2.1e) consist mainly of SH waves and the particle motion is parallel to the surface (transverse motion). In both cases, the wave-amplitude decays exponentially over depth.

Decoupled wave equation

By omitting the temporal and spatial dependencies we can write the stress-velocity formulation of the elastic wave equation represented by a first-order hyperbolic system. We can split equation 2.10 into two separate partial differential equation systems, each independent from one another, for the P-SV and the SH waves.

These waves, as mentioned above, are decoupled in a 2D medium and therefore there is no conversion from one wave type into the other. Both waves types propagate in the same plane, but they differ in their polarization. In the following, the axis x , y and z of the Cartesian coordinate system are defined as x for the in-line, y for the cross-line and z for the vertical axis. The polarization direction of the P-SV waves in 2-D lies within the x - z plane, whereas the SH waves are polarized perpendicular to the x - z plane (y direction). The final system consists of the following equations:

P-SV waves	SH waves
$\dot{v}_x = \frac{1}{\rho} \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right)$ (2.13a)	$\dot{v}_y = \frac{1}{\rho} \left(\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial z} \right)$ (2.14a)
$\dot{v}_z = \frac{1}{\rho} \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right)$ (2.13b)	$\dot{\sigma}_{xy} = \mu \frac{\partial v_y}{\partial x}$ (2.14b)
$\dot{\sigma}_{xx} = \kappa \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) - 2\mu \frac{\partial v_z}{\partial z}$ (2.13c)	$\dot{\sigma}_{yz} = \mu \frac{\partial v_y}{\partial z}$ (2.14c)
$\dot{\sigma}_{zz} = \kappa \left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x} \right) - 2\mu \frac{\partial v_x}{\partial x}$ (2.13d)	
$\dot{\sigma}_{xz} = \mu \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right)$ (2.13e)	

where $\kappa = \lambda + 2\mu$ is the P-wave modulus and $\dot{\cdot}$ time differentiation.

2.1.2 Viscoelasticity

One of the basic principles underlying wave propagation in real elastic media is related to intrinsic damping of the waveform amplitudes, due to the dispersive properties of the media causing a part of the kinetic energy to be transformed into heat. This affects also the phases but to a lesser extend. By extend in order to account for the attenuating properties of the real media, we include the anelastic behavior in equation 2.2 by

$$\sigma_{ij} = (\dot{\Psi}_p - 2\dot{\Psi}_s) * \theta \delta_{ij} + 2\dot{\Psi}_s * \epsilon_{ij} , \quad (2.15)$$

where the multiplication is changed to convolution in time and the Lamé parameters changed to the relaxation functions for P- and S-waves, Ψ_p and Ψ_s , respectively. The relaxation functions Ψ_p and Ψ_s are computed by using a rheological model that can approximate the attenuation and dispersion properties of the medium, known as the generalized standard linear solid (GSLS) (Liu et al., 1976; Emmerich and Korn, 1987; Blanch et al., 1995). The GSLS consists of L Maxwell bodies (dashpot in series with a

spring) in parallel with a spring, where each represents the viscosity of the medium for a given relaxation frequency, f_l . These relaxation functions are

$$\Psi_p(t) = \kappa \left(1 + \sum_{i=1}^L \left(\frac{\tau_{cl}^p}{\tau_{\sigma l}} - 1 \right) e^{-t/\tau_{\sigma l}} \right) H(t), \quad (2.16)$$

$$\Psi_s(t) = \mu \left(1 + \sum_{i=1}^L \left(\frac{\tau_{cl}^s}{\tau_{\sigma l}} - 1 \right) e^{-t/\tau_{\sigma l}} \right) H(t), \quad (2.17)$$

where $H(t)$ is the Heaviside step function, τ_{cl}^p and τ_{cl}^s are the P- and S- wave strain retardation times, and $\tau_{\sigma l}$ is the stress relaxation time calculated by $\tau_{\sigma l} = 1/2\pi f_l$ for a given relaxation frequency f_l .

Additionally, in order to avoid the convolution in equation 2.15, Carcione et al. (1988) and Robertsson et al. (1994) suggested the use of memory variables, r , and Bohlen (2002) further corrected for the phase velocities. We can now modify equations 2.13 and 2.14 in order to construct the 2D viscoelastic wave equation¹ for P-SV and SH waves:

P-SV waves

$$\dot{\sigma}_{xx} = \kappa(1 + \tau^p) \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) - 2\mu(1 + \tau^s) \frac{\partial v_z}{\partial z} + \sum_{l=1}^L r_{xxl} \quad (2.18a)$$

$$\dot{\sigma}_{zz} = \kappa(1 + \tau^p) \left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x} \right) 2\mu(1 + \tau^s) \frac{\partial v_x}{\partial x} + \sum_{l=1}^L r_{zzl} \quad (2.18b)$$

$$\dot{\sigma}_{xz} = \mu(1 + \tau^s) \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) + \sum_{l=1}^L r_{xzl} \quad (2.18c)$$

$$\dot{r}_{xxl} = -\frac{1}{\tau_{\sigma l}} \left((\kappa\tau^p - 2\mu\tau^s) \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) + 2\mu\tau^s \frac{\partial v_x}{\partial x} + r_{xxl} \right) \quad (2.18d)$$

$$\dot{r}_{zzl} = -\frac{1}{\tau_{\sigma l}} \left((\kappa\tau^p - 2\mu\tau^s) \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) + 2\mu\tau^s \frac{\partial v_z}{\partial z} + r_{zzl} \right) \quad (2.18e)$$

$$\dot{r}_{xzl} = -\frac{1}{\tau_{\sigma l}} \left(\mu\tau^s \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) + r_{xzl} \right) \quad (2.18f)$$

SH waves

$$\dot{\sigma}_{xy} = \mu(1 + \tau^s) \frac{\partial v_y}{\partial x} + \sum_{l=1}^L r_{xy_l} \quad (2.19a)$$

$$\dot{\sigma}_{zy} = \mu(1 + \tau^s) \frac{\partial v_y}{\partial z} + \sum_{l=1}^L r_{zy_l} \quad (2.19b)$$

$$\dot{r}_{xy_l} = -\frac{1}{\tau_{\sigma l}} \left(\mu\tau^s \frac{\partial v_y}{\partial x} + r_{xy_l} \right) \quad (2.19c)$$

$$\dot{r}_{zy_l} = -\frac{1}{\tau_{\sigma l}} \left(\mu\tau^s \frac{\partial v_y}{\partial z} + r_{zy_l} \right) \quad (2.19d)$$

¹ Note that the formulation for the particle velocities (equations 2.13a,b and 2.14a) is not rewritten since it is changed explicitly by the spatial derivatives of the stress components.

In the above equations, r_{ijl} corresponds to the memory variable of the ij^{th} component for the l^{th} relaxation mechanism. The τ -method initially suggested by Blanch et al. (1995) allowed the reduction of the number of variables that are required to fully describe a GSLS. This further leads to an increase in computational savings and a reduction in memory requirements, due to the fact that the dimensionless variable τ is sufficient to describe the damping caused by a GSLS. The parameter τ allows us to remove the dependency of the P- and S-wave retardation times from the number of relaxation mechanisms that are used, and is given by

$$\tau = \frac{\tau_{el}}{\tau_{\sigma l}} - 1. \quad (2.20)$$

Finally, we can include the attenuation in wave propagation making use of the quality factor Q ($= 2\pi \frac{E}{\Delta E}$, with ΔE being the energy loss per cycle). We can approximate the Q -values for a given set of frequencies under a least-square approach (Bohlen, 2002) by

$$Q_{p/s}(\omega, \tau_{\sigma l}, \tau) = \frac{1 + \sum_{l=1}^L \frac{\omega^2 \tau_{\sigma l}^2}{1 + \omega^2 \tau_{\sigma l}^2} \tau}{\sum_{l=1}^L \frac{\omega \tau_{\sigma l}}{1 + \omega^2 \tau_{\sigma l}^2} \tau}, \quad (2.21)$$

where ω is the angular frequency. The relaxation times are calculated in a way that for the given frequency range the resulted Q -values are constant.

2.1.3 Numerical representation of wave equation

Solving equations 2.18 and 2.19 analytically is only possible for very simple models. In realistic cases, where complex structures are present, the wave propagation can only be estimated using numerical approaches. Two of the most popular methods to numerically calculate wave propagation are the finite-difference and the finite-element methods. In this work, I focus entirely on finite-difference methods (FD) due to mainly their numerical efficiency as well as their relatively simple implementation. One more reason is that I am interested in the near-surface scale, with short-offset acquisitions, where we are dealing with flat topography which also makes the FD approach straightforward.

2.1.3.1 Finite-difference scheme

The specific case that I am focusing is the finite-difference scheme in the time domain (FDTD). We approximate the spatial and temporal derivatives in the wave equations by finite-difference operators. This requires the discretization of the wave equation in both space and time using a Cartesian coordinate system of equidistant grid spacing. I therefore replace the continuous coordinates x and z by their discrete counterparts, such that $x = i\Delta h$ and $z = j\Delta h$, where i and j denotes the position of a specific grid point. For the time coordinate I similarly discretize the vector t such that $t = n\Delta t$ where n refers to a specific time step.

The exact placement of the viscoelastic material parameters in the chosen coordinate system is shown in Figure 2.2. I use the standard staggered-grid (SSG) approach following Virieux (1984) and Levander

(1988) to ensure stability and high accuracy of the forward modeling. In SSG the model parameters, ρ , λ and μ , and the diagonal stress components are defined on full grid points, while the off-diagonal stress and velocity components are distributed on half grid points.

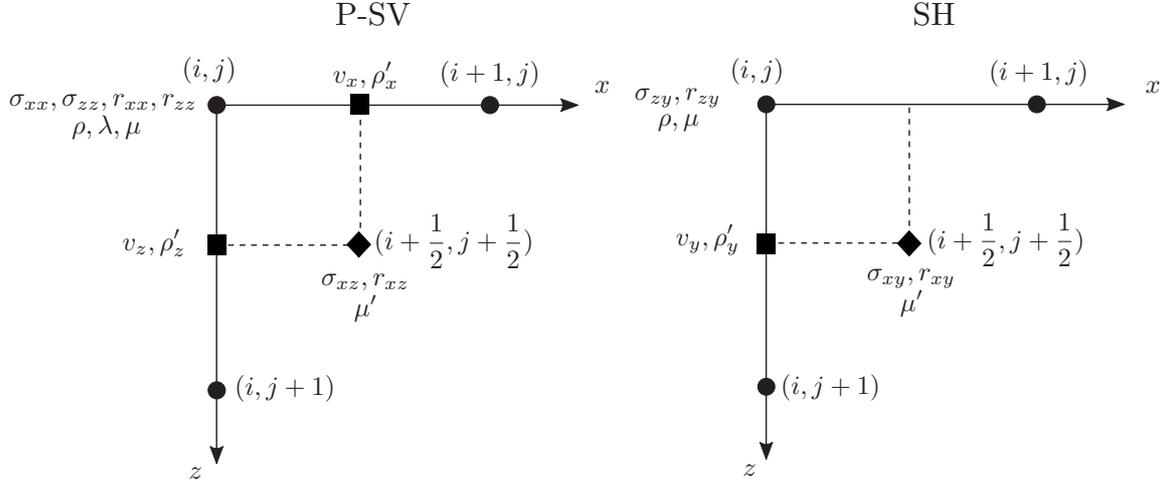


Figure 2.2: Distribution of viscoelastic parameters on a standard staggered-grid for P-SV waves (left) and SH waves (right).

After defining the placing of the material parameters in the SSG, we express the partial derivatives using second-order FD operators in both space and time for any continuous function f by

$$\frac{\partial f}{\partial x}(n, i) = \frac{f(n, i + \frac{1}{2}) - f(n, i - \frac{1}{2})}{\Delta h} + \mathcal{O}(\Delta h^2), \quad (2.22)$$

$$\frac{\partial f}{\partial t}(n, i) = \frac{f(n + \frac{1}{2}, i) - f(n - \frac{1}{2}, i)}{\Delta t} + \mathcal{O}(\Delta t^2), \quad (2.23)$$

for the i^{th} grid point and at the n^{th} time step. In this work, I use second-order approximations in time and sixth-order accuracy in space.

From Figure 2.2 we observe that certain parameters, namely μ' and $\rho'(x, y)$, require special treatment. In particular, the density values are obtained by arithmetically averaging the two adjacent grid points in each direction, while the μ values are calculated using the harmonic average of the four adjacent grid points.

2.1.3.2 Numerical accuracy

The discretization, although it makes the modeling of wave propagation feasible, comes with compromises in terms of numerical accuracy. More specific, the spatial sampling has to be chosen such that numerical dispersion is avoided. To ensure the proper spatial sampling the theory of signal processing demands that the following condition has to be met at all times:

$$\Delta h \leq \frac{\lambda_{\min}}{n} = \frac{v_{\min}}{nf_{\max}} \quad (2.24)$$

$$\text{where } n_{FD} = \{12, 8, 6, 5, 5, 4\} \quad \text{for } FD = 2, 4, 6, 8, 10, 12 \quad (2.25)$$

This means that the spatial sampling, Δh , should not surpass n points per minimum wavelength λ_{\min} . In the above equation, n depends on the order of the FD operator, while λ_{\min} depends on the maximum frequency of the source signal and the minimum seismic velocity of the medium.

Additionally, to ensure proper temporal sampling we need to satisfy the Courant-Friedrichs-Lewy criterion (Courant et al., 1928) which for a 2D case is given by

$$\Delta t \leq \frac{\Delta h}{\beta \sqrt{2} v_{\max}}, \quad (2.26)$$

meaning that the time required for a wave to propagate between two adjacent grid points has to be smaller than Δt . The parameter β contains the sum of the weighting coefficients of the FD operator.

2.1.3.3 Initial and boundary conditions

To simulate wave propagation in a given medium certain conditions have to be applied, in order to ensure both a stable but also a physically meaningful solution of the wave equation. These conditions may vary depending on the target that we wish to examine. In our case, we model wave propagation while we fulfill at all times the conditions shown below:

Initial conditions: The medium at time $t = 0$ is at rest since there are no forces applied to it and therefore it is at equilibrium. Without acting forces the particle velocities, the stresses and the memory variables (along with their time derivatives) vanish at every spatial location (Virieux, 1986, 1984), i.e.,

$$v_x|_{t=0} = v_y|_{t=0} = v_z|_{t=0} = 0, \quad (2.27)$$

$$\sigma_{xx}|_{t=0} = \sigma_{yy}|_{t=0} = \sigma_{zz}|_{t=0} = \sigma_{xz}|_{t=0} = \sigma_{xy}|_{t=0} = \sigma_{zy}|_{t=0} = 0, \quad (2.28)$$

$$r_{xx}|_{t=0} = r_{zz}|_{t=0} = r_{xz}|_{t=0} = r_{xy}|_{t=0} = r_{zy}|_{t=0} = 0. \quad (2.29)$$

Absorbing and free-surface boundaries: In order to avoid reflections at the locations of the numerical boundaries and allow the absorption of the wavefield, thus simulating the behavior of an infinite medium, we make use of the convolutional perfectly matched layers (CPMLs) method introduced by Komatitsch and Martin (2007). This boundary method exponentially decays any incoming wave, preventing it from reaching the numerical boundaries of the model, by stretching the coordinates in the frequency domain as the wavefield approaches the actual boundary. With the additional application of a damping factor, reflections that would otherwise be introduced (due to the discretization accuracy) can also be eliminated.

Finally, the top edge of the model corresponding to the Earth's surface demands a special treatment. For this boundary, the traction-free condition needs to be applied, since at the earth-to-air interface all stresses must vanish, which is done using the mirroring technique (Levander, 1988). All vertical stresses at the grid points where the free surface is located are set to zero, while the stresses above the free surface are calculated by mirroring the values below the surface, using the opposite sign.

2.2 Full-waveform inversion

The above formulations help us derive a solver to the forward problem, which essentially corresponds to the first step towards the solution of the inverse problem. As already discussed, the solution to the inverse problem in the framework of this thesis will be based on the principle known as full-waveform inversion. A local optimization technique, which through an iterative procedure that includes the minimization of a cost function, allows us to match simulated data with observations from real measured quantities. The solution of the inverse problem is highly ambiguous (compared to the forward problem), since the model retrieved by FWI is just one of many possible realizations that explain the given observations. In this section the basic principles of FWI will be discussed before moving on to certain techniques that will be employed to reduce this ambiguity/non-linearity of the solution of the inverse problem.

2.2.1 Objective function of the inverse problem

The general FWI methodology discussed in this section will follow Tarantola (1986). We initially parametrize the model space, $\mathbf{m} = (m_1, \dots, m_N)^T$, in terms of density and seismic velocities, which leads to $\mathbf{m} = (\rho, \mathbf{v}_S, \mathbf{v}_P)^T$. We define the non-linear forward operator f , which allows us to obtain synthetic data $\mathbf{d}_{\text{syn}}(\mathbf{m})$, given a model \mathbf{m} , through

$$\mathbf{d}_{\text{syn}}(\mathbf{m}) = f(\mathbf{m}) . \quad (2.30)$$

The goal of FWI is to find the model \mathbf{m} such that the difference between $\mathbf{d}_{\text{syn}}(\mathbf{m})$ produced by this model and \mathbf{d}_{obs} is minimum. This procedure takes place iteratively and the data residuals $\Delta \mathbf{d} = (\Delta d_1, \dots, \Delta d_M)^T$ between the synthetic data $\mathbf{d}_{\text{syn}}(\mathbf{m})$ and the observed data \mathbf{d}_{obs} are defined for every iteration as

$$\Delta \mathbf{d} = \mathbf{d}_{\text{syn}}(\mathbf{m}) - \mathbf{d}_{\text{obs}} . \quad (2.31)$$

The objective function, $E(\mathbf{m})$, can be defined arbitrarily. In my case I use the least-squares L_2 -norm between the synthetic and observed seismograms

$$E(\mathbf{m}) = \frac{1}{2} \sum_{r,s} \Delta \mathbf{d}^T \cdot \Delta \mathbf{d} , \quad (2.32)$$

where the contributions for every source and receiver are summed up. By assuming weak non-linearity of equation 2.32, we can make use of the perturbation theory under the Born approximation. Starting by adding a small perturbation $\Delta \mathbf{m}$ to our initial model \mathbf{m}_0 (as required when dealing with local optimization techniques) such that

$$\mathbf{m} = \mathbf{m}_0 + \Delta \mathbf{m} , \quad (2.33)$$

and expanding it through a 2nd-order Taylor series, we get

$$E(\mathbf{m}_0 + \Delta \mathbf{m}) = E(\mathbf{m}_0) + \Delta \mathbf{m} \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} + \frac{1}{2} \Delta \mathbf{m} \frac{\partial^2 E(\mathbf{m}_0)}{\partial \mathbf{m}^2} \Delta \mathbf{m}^T + \mathcal{O}(\Delta \mathbf{m}^3) . \quad (2.34)$$

Finding the minimum of equation 2.34 around \mathbf{m}_0 requires the derivative of $E(\mathbf{m}_0 + \Delta\mathbf{m})$ with respect to \mathbf{m} to vanish:

$$\frac{\partial E(m)}{\partial \mathbf{m}} = \frac{\partial E(m_0)}{\partial \mathbf{m}} + \Delta\mathbf{m} \frac{\partial^2 E(m_0)}{\partial \mathbf{m}^2} \stackrel{!}{=} 0 \quad (2.35)$$

By now solving equation 2.35 for $\Delta\mathbf{m}$, we obtain the perturbation in the model:

$$\Delta\mathbf{m} = - \left(\frac{\partial^2 E(m_0)}{\partial \mathbf{m}^2} \right)^{-1} \frac{\partial E(m_0)}{\partial \mathbf{m}} \quad (2.36)$$

By substituting equation 2.36 into 2.33 and the first and second-order derivative with respect to the model parameters with the gradient $\nabla_{\mathbf{m}}E(\mathbf{m}_n)$ and the Hessian matrix \mathbf{H} , respectively, we obtain the final model update for the n^{th} iteration by

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \Delta\mathbf{m}_n = \mathbf{m}_n - \mathbf{H}_n^{-1} \cdot \nabla_{\mathbf{m}}E(\mathbf{m}_n) . \quad (2.37)$$

The calculation of \mathbf{H} is computationally expensive for large-scale optimization problems, due to the fact that it is a square dense $M \times M$ matrix, where M is the size of the model parameters. Equation 2.37 is known as the Newton method. In my case, however, I will be using an approximation for the Hessian matrix and update the models using the approach of steepest-descent.

Under this modification, equation 2.37 is substituted by

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha \mathbf{P}_n \nabla_{\mathbf{m}}E(\mathbf{m}_n) , \quad (2.38)$$

where α is an adaptive step length and \mathbf{P} is a preconditioning matrix which depends on the optimization method (section 2.3.1). The gradient of the misfit function defines the direction of the update of the model parameters, while the step length defines the amount of change that is allowed in each iteration. It is therefore required to perform a parabolic search to estimate the best value for α (Kurzmann, 2012). This procedure is repeated for every iteration, where a total of three different step lengths and the misfit values are fit with a parabolic curve. The value on the vertex of this parabola is the one chosen. What is left now is to find a way to efficiently estimate the gradients $\nabla_{\mathbf{m}}E(\mathbf{m}_n)$. Calculating the gradient explicitly would require a tremendous amount of computational resources, since we would have to change each model parameter on each grid point by applying a small perturbation and then evaluating the misfit. Instead, I will make use of the adjoint state method (Tarantola, 1984; Mora, 1987), explained in the next section.

2.2.2 Adjoint-state method

As a first step for the derivation of the gradients required for the model updates (equation 2.38), I rewrite the wave equation in matrix-vector format for better clarity. For similar reason and to avoid unnecessary complexity, I will use the elastic wave equations for P-SV and SH waves without including the parameters related to attenuation. Note the change in vector notation ($\vec{\cdot}$) that is chosen in the section for better clarity and to avoid confusion with the matrix notation.

Matrix-vector formulation

Equations 2.13 and 2.14 can be formulated in a more compact form by expressing the individual terms as matrices and vectors, which leads to

$$\vec{F}(\vec{\chi}, m) = \mathbf{M}^{-1} \left(\frac{\partial \vec{\chi}}{\partial t} - \vec{b} \right) - \mathbf{Q}\vec{\chi} = 0. \quad (2.39)$$

The vectors $\vec{\chi}$ and \vec{b} represent the stress/velocity wavefields and the corresponding sources, respectively. Considering the 2D case, they can be expressed as:

$$\begin{array}{c} \text{P-SV waves} \\ \vec{\chi} = \begin{pmatrix} v_x \\ v_z \\ \sigma_{xx} \\ \sigma_{zz} \\ \sigma_{xz} \end{pmatrix} \quad \vec{b} = \begin{pmatrix} f_x/\rho \\ f_z/\rho \\ s_{\sigma_{xx}} \\ s_{\sigma_{zz}} \\ s_{\sigma_{xz}} \end{pmatrix} \end{array} \quad (2.40)$$

$$\begin{array}{c} \text{SH waves} \\ \vec{\chi} = \begin{pmatrix} v_y \\ \sigma_{xy} \\ \sigma_{zy} \end{pmatrix} \quad \vec{b} = \begin{pmatrix} f_y/\rho \\ s_{\sigma_{xy}} \\ s_{\sigma_{zy}} \end{pmatrix} \end{array} \quad (2.41)$$

The matrices \mathbf{M} and \mathbf{Q} in equation 2.39 include the elastic material parameters, m (λ , μ and ρ), and the stencil operator for the spatial derivatives, respectively:

$$\begin{array}{c} \text{P-SV waves} \\ \mathbf{M}^{-1} = \begin{pmatrix} \rho & 0 & 0 & 0 & 0 \\ 0 & \rho & 0 & 0 & 0 \\ 0 & 0 & A & B & 0 \\ 0 & 0 & B & A & 0 \\ 0 & 0 & 0 & 0 & \mu^{-1} \end{pmatrix} \end{array} \quad (2.42)$$

$$\begin{array}{c} \text{SH waves} \\ \mathbf{M}^{-1} = \begin{pmatrix} \rho & 0 & 0 \\ 0 & \mu^{-1} & 0 \\ 0 & 0 & \mu^{-1} \end{pmatrix} \end{array} \quad (2.43)$$

where $A = \frac{\lambda+2\mu}{4\mu(\lambda+\mu)}$ and $B = -\frac{\lambda}{4\mu(\lambda+\mu)}$

$$\begin{array}{c} \text{P-SV waves} \\ \mathbf{Q} = \begin{pmatrix} 0 & 0 & \partial_x & 0 & \partial_y \\ 0 & 0 & 0 & \partial_y & \partial_x \\ \partial_x & 0 & 0 & 0 & 0 \\ 0 & \partial_y & 0 & 0 & 0 \\ \partial_y & \partial_x & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (2.44)$$

$$\begin{array}{c} \text{SH waves} \\ \mathbf{Q} = \begin{pmatrix} 0 & \partial_x & \partial_z \\ \partial_x & 0 & 0 \\ \partial_z & 0 & 0 \end{pmatrix} \end{array} \quad (2.45)$$

Augmented functional

The derivation of the gradients in equation 2.38 is done using the method of augmented Lagrangian function, closely following the work of Plessix (2006). For this purpose, I first define the augmented functional, \mathcal{L} , as

$$\mathcal{L}(\vec{\chi}, \vec{\psi}, m) = E(\vec{\chi}, m) - \vec{\psi} \cdot \vec{F}(\vec{\chi}, m) \quad (2.46)$$

where the Lagrange multiplier $\vec{\psi}$ is independent of m . The function E is any objective function that we would like to minimize, given certain conditions. In my case it is the minimization of the residual energy between the observed and synthetic data. The augmented functional is therefore constructed by a multiplication of $\vec{\psi}$ with the constraints \vec{F} (the wave equation), subtracted from the function we would like to minimize. It is given by

$$\mathcal{L}(\vec{\chi}, \vec{\psi}, m) = \frac{1}{2} \int_{\mathbf{x}} \int_0^T (\vec{d}_{\text{syn}} - \vec{d}_{\text{obs}})^2 d\mathbf{x} - \int_0^T \int_{\mathbf{x}} \vec{\psi} \left(\mathbf{M}^{-1} \left(\frac{\partial \vec{\chi}}{\partial t} - \vec{b} \right) - \mathbf{Q} \vec{\chi} \right) dt d\mathbf{x} . \quad (2.47)$$

The parameter $\vec{\chi}$ is a physical realization and by definition we get

$$\vec{F}(\vec{\chi}, m) = 0, \quad (2.48)$$

and hence

$$\mathcal{L}(\vec{\chi}, \vec{\psi}, m) = E(m) . \quad (2.49)$$

In case that we do not use the vector-matrix format, we would require additional Lagrange multipliers in order to include both equations 2.10 and 2.12. I proceed by taking the derivative of E with respect to the model parameters m and apply the chain rule:

$$\frac{dE(\vec{\chi}, \vec{\psi}, m)}{dm} = \frac{\partial \mathcal{L}(\vec{\chi}, \vec{\psi}, m)}{\partial \vec{\chi}} \frac{\partial \vec{\chi}}{\partial m} + \frac{\partial \mathcal{L}(\vec{\chi}, \vec{\psi}, m)}{\partial m} \quad (2.50)$$

The solution to our problem requires to find $\vec{\psi}$ such that

$$\frac{\partial \mathcal{L}(\vec{\chi}, \vec{\psi}, m)}{\partial \vec{\chi}} = 0 \quad \text{and} \quad \frac{dE}{dm} = \frac{\partial \mathcal{L}(\vec{\chi}, \vec{\psi}, m)}{\partial m} \quad (2.51)$$

and therefore remove the derivatives over the stress/velocity wavefields, expressed through $\vec{\chi}$. To calculate the derivatives with respect to $\vec{\chi}$, I integrate equation 2.47 by parts in space and time, while neglecting the source terms \vec{b} and with boundary terms equal to zero (due to the absorbing boundaries and the initial condition, $t=0$):

$$\mathcal{L}(\vec{\chi}, \vec{\psi}, m) = \frac{1}{2} \int_{\mathbf{x}} \int_0^T \Delta \vec{d}^2 d\mathbf{x} + \int_0^T \int_{\mathbf{x}} \left(\mathbf{M}^{-1} \frac{\partial \vec{\psi}}{\partial t} \vec{\chi} - \mathbf{Q} \vec{\psi} \vec{\chi} \right) dt d\mathbf{x} \quad (2.52)$$

where $\Delta \vec{d} = \vec{d}_{\text{syn}} - \vec{d}_{\text{obs}}$. From the boundary terms that have been removed in the integration by parts, we obtain the boundary condition $\vec{\psi}(T) = 0$. We can now differentiate with respect to $\vec{\chi}$ and equate the result to zero:

$$\begin{aligned} \frac{\mathcal{L}(\vec{\chi}, \vec{\psi}, m)}{\partial \vec{\chi}} &= \int_{\mathbf{x}} \int_0^T \Delta \vec{d}(t) + \int_0^T \int_{\mathbf{x}} \left(\mathbf{M}^{-1} \frac{\partial \vec{\psi}(t)}{\partial t} \vec{\chi} - \mathbf{Q} \vec{\psi}(t) \right) dt d\mathbf{x} = 0 \\ \Rightarrow \int_0^T \int_{\mathbf{x}} \left(\mathbf{M}^{-1} \frac{\partial \vec{\psi}(t)}{\partial t} \vec{\chi} + \Delta \vec{d}(t) \right) dt d\mathbf{x} &= \int_0^T \int_{\mathbf{x}} \mathbf{Q} \vec{\psi}(t) dt d\mathbf{x} \end{aligned} \quad (2.53)$$

In order to fulfill the condition $\vec{\chi} = 0$ at $t = 0$ we need to write equation 2.53 reverse in time by replacing t with $T - t$, which leads to

$$\int_T^0 \int_{\mathbf{x}} \left(\mathbf{M}^{-1} \frac{\partial \vec{\psi}(t)}{\partial T-t} \vec{\chi} + \Delta \vec{d}(T-t) \right) dt d\mathbf{x} = \int_0^T \int_{\mathbf{x}} \mathbf{Q} \vec{\psi}(T-t) dt d\mathbf{x} \quad (2.54)$$

I additionally substitute $\tau = T - t$, $d\tau = -dt$ and $d\tau = -dt$, along with interchanging the boundaries T to zero and visa-versa. We then obtain:

$$\int_0^T \int_{\mathbf{x}} \left(\mathbf{M}^{-1} \frac{\partial \vec{\psi}(\tau)}{\partial \tau} \vec{\chi} - \Delta \vec{d}(\tau) \right) d\tau d\mathbf{x} = - \int_0^T \int_{\mathbf{x}} \mathbf{Q} \vec{\psi}(\tau) d\tau d\mathbf{x} \quad (2.55)$$

Finally, we obtain an equation similar to 2.39

$$\mathbf{M}^{-1} \frac{\partial \vec{\psi}(\tau)}{\partial \tau} - \Delta \vec{d}(\tau) = -\mathbf{Q} \vec{\psi}(\tau), \quad (2.56)$$

where the only differences are the signs and that the sources expressed by \vec{b} in equation 2.39 are now substituted by the data residuals $\Delta \vec{d}$. This equation is called the anti self-adjoint equation. The newly constructed $\vec{\psi}$ is the adjoint wavefield and corresponds to the solution of the forward problem (wave equation) with the difference that the source acting on the system is described by the data residuals and is called the adjoint source. Due to the fact that the initial conditions of the actual forward problem are equal to the terminal conditions at the modified forward problem, it means that we need to solve it backwards in time.

For the solution of equation 2.56 I will proceed in the same way as solving equation 2.39, while making the substitutions shown below for both P-SV and SH:

$$\vec{\psi} = \mathbf{\Lambda} \vec{\phi}, \quad \mathbf{\Lambda}_{\text{P-SV}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{\Lambda}_{\text{SH}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad (2.57)$$

where

$$\vec{\phi}_{\text{P-SV}} = (\phi_{v_x}, \phi_{v_z}, \phi_{\sigma_{xx}}, \phi_{\sigma_{zz}}, \phi_{\sigma_{xz}})^T \quad (2.58)$$

$$\vec{\phi}_{\text{SH}} = (\phi_{v_y}, \phi_{\sigma_{xy}}, \phi_{\sigma_{zy}})^T \quad (2.59)$$

Now that $\vec{\psi}$ is known, we can write the gradients at every x (therefore omitting the spatial integral) and upon substitution in 2.46 we get

$$\frac{\partial E}{\partial m} = - \int_0^T (\Lambda \vec{\Phi}) \left(\frac{\partial \mathbf{M}^{-1}}{\partial m} \frac{\partial \vec{\chi}}{\partial t} - \vec{b} \right) dt. \quad (2.60)$$

I finally expand the matrix-vector formulation and by substituting all quantities in equation 2.61 we get stress-velocity gradients for the P-SV and SH waves:

P-SV waves	SH waves
$\frac{\partial E}{\partial \rho} = - \int (v_x^F v_x^B + v_z^F v_z^B) dt \quad (2.61)$	$\frac{\partial E^{\text{SH}}}{\partial \rho} = - \int v_y^F v_y^B dt \quad (2.64)$
$\frac{\partial E}{\partial \mu} = - \int \left(\frac{(\sigma_{xz}^F \sigma_{xz}^B)}{\mu^2} + \frac{(\sigma_{xx}^F + \sigma_{zz}^F)(\sigma_{xx}^B + \sigma_{zz}^B)}{4(\lambda + \mu)^2} + \frac{(\sigma_{xx}^F - \sigma_{zz}^F)(\sigma_{xx}^B - \sigma_{zz}^B)}{4\mu^2} \right) dt \quad (2.62)$	$\frac{\partial E^{\text{SH}}}{\partial \mu} = - \int \frac{(\sigma_{xy}^F \sigma_{xy}^B + \sigma_{zy}^F \sigma_{zy}^B)}{\mu^2} dt \quad (2.65)$
$\frac{\partial E}{\partial \lambda} = - \int \frac{(\sigma_{xx}^F + \sigma_{zz}^F)(\sigma_{xx}^B + \sigma_{zz}^B)}{4(\lambda + \mu)^2} dt \quad (2.63)$	

where the indices correspond to the forward (F) propagated incident wavefield or to the backward (B) propagated adjoint (residual) wavefield, respectively. By making use of the above equations, we evaluate the gradients by crosscorrelating the forward and adjoint wavefields, while we additionally sum the gradients of all the individual shots in the case of multiple shots.

Finally, I proceed by using a different parameterization consisting of seismic velocities and density. This parameterization shows less cross-talk and less ambiguities compared to the one with density and the Lamé-parameters, due to similarities in the order of magnitude of seismic velocities and density, and is therefore preferred (Köhn et al., 2012). Another reason is the fact that their similar order of magnitude guarantees a more stable inversion. I make use of equation 2.6 and by applying the chain rule equations 2.61-2.65 become

$$\frac{\partial E}{\partial v_p} = 2\rho v_p \frac{\partial E}{\partial \lambda}, \quad (2.66)$$

$$\frac{\partial E}{\partial v_s} = -4\rho v_s \frac{\partial E}{\partial \lambda} + 2\rho v_s \frac{\partial E}{\partial \mu}, \quad (2.67)$$

$$\frac{\partial E}{\partial \rho} = (v_p^2 - 2v_s^2) \frac{\partial E}{\partial \lambda} + v_s^2 \frac{\partial E}{\partial \mu} + \frac{\partial E}{\partial \rho}, \quad (2.68)$$

where for the SH case any derivative of λ and the terms including v_p , vanish.

2.3 Numerical and physical aspects during FWI

The computational approaches which are discussed in this section constitute essential steps towards a successful application of FWI, for both synthetic and field data applications. Most of these techniques contribute features to reduce the non-linearity of the inverse problem and/or take advantage of physical properties of wave propagation to improve the convergence towards the true solution. Others deal with the reduction of the computational cost that the FWI algorithm introduces.

2.3.1 Optimization method

The steepest descent method introduced in equation 2.38 is a stable technique for solving local optimization problems, but its convergence speed is very slow (Nocedal and Wright, 2006). Instead, I use a preconditioned conjugate-gradient (PCG) scheme and update the model at iteration step n not over the gradient direction $\delta\mathbf{m}_n$, but along the conjugate direction $\delta\mathbf{c}_n$, namely

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \alpha\mathbf{P}_n\delta\mathbf{c}_n \quad \text{for } n > 1, \quad (2.69)$$

where

$$\delta\mathbf{c}_n = \delta\mathbf{m}_n - \beta_n\delta\mathbf{c}_{n-1}. \quad (2.70)$$

By considering information of the previous gradient we are able to abstain from abrupt changes of the gradient direction between successive iterations. The weighting factor β is estimated by Polak and Ribière (1969) and reads:

$$\beta_n = \frac{\delta\mathbf{m}_n^T(\delta\mathbf{m}_n - \delta\mathbf{m}_{n-1})}{\delta\mathbf{m}_{n-1}^T\delta\mathbf{m}_{n-1}}. \quad (2.71)$$

2.3.2 Preconditioning

Due to the non-uniqueness of FWI, physical characteristics of seismics should be used in order to constrain the number of possible solutions. Such constraints can be applied through a preconditioner in the calculation of the gradients required for the model updates, which is described by \mathbf{P}_n in equation 2.69.

Due to geometrical spreading, the absolute wave amplitude decays as body waves propagate away from the source, with a factor $1/\sqrt{r}$ (r = distance from excitation point). This leads to large updates in the shallow parts of the model, since at these locations the gradient has the largest amplitude. To compensate for this effect semi-circular tapers are applied to the gradients at the location of the sources. This taper sets the amplitude at the source position to zero and gradually increases the amplitude to one within a specified number of grid points. Since the taper is applied shot-wise, a model update is also obtained at the tapered source locations in the final gradient, due to the contribution of gradients from the sources with a different position.

Finally, a global preconditioner is used which refers to an approximation to the diagonal elements of the Hessian matrix. It is based on the sum of the amplitudes, u_i , of the forward modeled incident wavefield at each grid point. Additionally, the influence of the receivers is included by a geometrical estimation of the receiver's Green's functions. Following Plessix and Mulder (2004a) and Wehner et al. (2015) for a single shot we get

$$\mathbf{H}_a^{-1}(\mathbf{x}_s, \mathbf{x}) = \left[\epsilon + \int |u_i(\mathbf{x}_s, \mathbf{x}, t)|^2 \left(\operatorname{asinh} \left(\frac{x_r^{\max} - x}{z} \right) - \operatorname{asinh} \left(\frac{x_r^{\min} - x}{z} \right) \right) dt \right]^{-1}, \quad (2.72)$$

where x_r^{\max} and x_r^{\min} correspond to the maximum and minimum horizontal receiver position, respectively, for the source location \mathbf{x}_s . To stabilize the inversion of the expression above, a water level ϵ is added, which is empirically chosen for P-SV and SH waves. I calculate this preconditioner for each shot and apply it normalized to the gradients from each respective source.

2.3.3 Multi-test step length estimation

For the step length estimation, α (equation 2.69), although two additional forward calculations are sufficient to calculate a parabolic fit, we perform additional forward simulations to obtain a better fit, by using only selected shots though. The test step lengths for the individual material parameters are calculated by scaling the maximum of the gradient to the maximum of the actual models:

$$\alpha_{v_p} = \frac{\max(v_p)}{\lambda_i \max(\nabla_{v_p} E)} \quad (2.73)$$

$$\alpha_{v_s} = \frac{\max(v_s)}{\lambda_i \max(\nabla_{v_p} E)} \quad (2.74)$$

$$\alpha_{\rho} = \frac{\max(\rho)}{\lambda_i \max(\nabla_{\rho} E)}, \quad (2.75)$$

where λ corresponds to the maximum model changes allowed for i different step length trials.

2.3.4 Misfit definitions

The aim of the inversion process is to minimize the objective function iteratively, thereby finding a model of the subsurface that explains the observed data. The choice of the misfit function is of great importance, since both the gradient estimation and the adjoint sources will, as a consequence, be adjusted to the norm we use.

When it comes to field data applications, prior to the estimation of the residuals I trace-wise normalize both synthetic and observed data in order to account for differences in the recorded and simulated energy of the data. Additionally, this helps when we cannot precisely simulate the anelastic behavior underlying the field recordings and other unknown effects caused by, for instance, the coupling of individual geophones

to the ground and the energy of the force acting as a source. This normalization is done also in some synthetic cases that I will discuss later and is used in order to retain consistency with field data inversions.

While the objective function can be defined arbitrarily, throughout this thesis the misfit definition that I use is a weighted L_2 -norm between the synthetic and observed seismograms:

$$E_w(\mathbf{m}) = \frac{1}{2} \sum_{r,s} \frac{\Delta \mathbf{d}^T \cdot \Delta \mathbf{d}}{\mathbf{d}_{\text{obs}}^T \cdot \mathbf{d}_{\text{obs}}}, \quad (2.76)$$

where the contributions for every source and receiver are summed up. Compared to equation 2.32, the weighted objective function is defined as the ratio between the residual energy and the energy of the observed data, where a value of one would indicate that the residual energy is equivalent to the energy in the observed data.

Another misfit definition that I use in some of the synthetic cases is a least-squares norm of normalized wavefields, suggested by Choi and Alkhalifah (2012), and given by

$$E_n(\mathbf{m}) = \frac{1}{2} \sum_{r,s} \frac{\left(\frac{\mathbf{d}_{\text{syn}}}{\|\mathbf{d}_{\text{syn}}\|} - \frac{\mathbf{d}_{\text{obs}}}{\|\mathbf{d}_{\text{obs}}\|} \right)^2}{\left(\frac{\mathbf{d}_{\text{obs}}}{\|\mathbf{d}_{\text{obs}}\|} \right)^2}. \quad (2.77)$$

The advantage of this misfit function is that both near and far-offset traces contribute equally to the misfit and it is not sensitive to residuals in the amplitude decay with offset. However, it is sensitive to relative amplitude differences within one seismogram. Since in the FWI applications that are shown later the differences were negligible, throughout this thesis I will use the misfit definition given by Equation 2.76.

2.3.5 Multi-scale approach

As I have already discussed, FWI is the solution of a local optimization problem. In cases of gradient-based optimizations a convergence towards the global minimum is not guaranteed and the algorithm can converge in local minima. Especially in shallow applications and with the dominance of surface waves, FWI shows even higher non-linearity. While the choice of a strong a-prior information would increase the chances of converging towards the global minimum, in field data applications such information is usually not available.

Bunks et al. (1995) proposed a method to reduce the non-linearity included in the misfit function by means of a multiscale approach. They suggested to reduce the initial bandwidth of both observed and synthetic data through bandpass filtering, thus increasing the effective period of the data, resulting in a more linear polynomial function. Subsequently, the filter length is widened to allow more frequencies to be included in the inverse problem. Until today their method is one of the most common techniques for FWI due to the robustness of the approach.

An additional feature of the multiscale approach is that it allows the reduction of the trade-off between the involved parameters. Elastic parameters have a different sensitivity to different parts of the waveform. Therefore, a specifically designed filter can shift the focus towards the area of interest, by using only parts of the waveform that are sensitive to a specific parameter. In Figure 2.3, I show the application

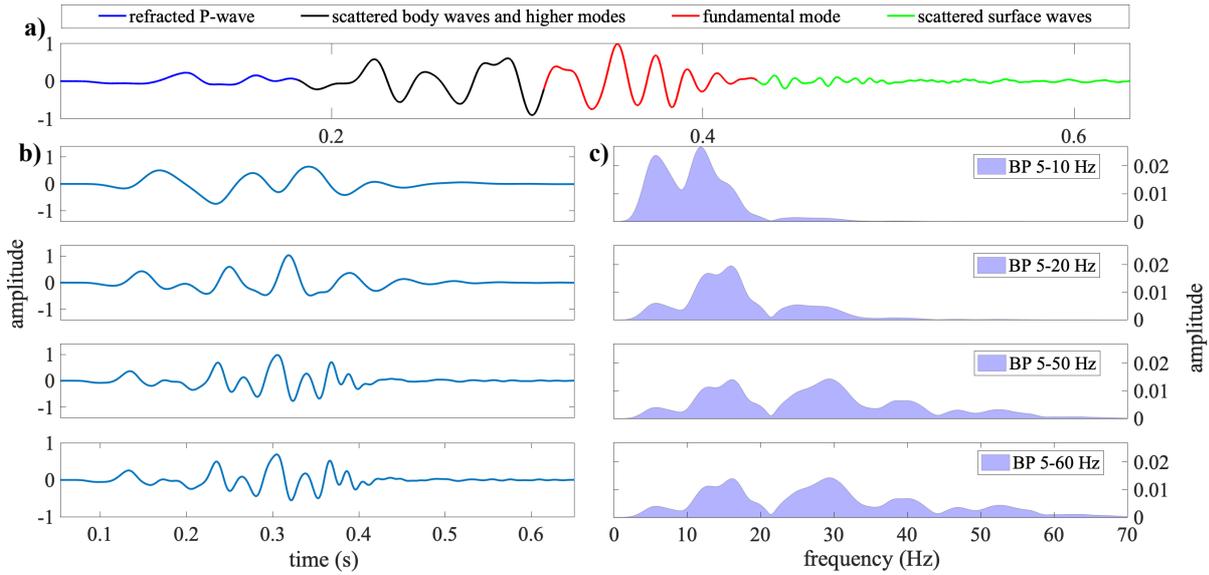


Figure 2.3: The effects of a multiscale approach on a synthetic seismogram in the time and frequency domain. a) Synthetic seismic reference trace. The arrivals of different wave types are highlighted. b) Seismogram after application of bandpass filters (BP). c) Corresponding amplitude spectra.

of a 4th order Butterworth filter on a representative seismic trace. The trace consists of various phases corresponding to different wavetypes and shows the complexity of the signal that will be used in the FWI algorithm. The application of the Butterworth filter shows that the waveform becomes less complex, which consequently leads to a simpler misfit function.

After the inversion reaches a certain abort criterion, we gradually increase the upper corner frequency of the Butterworth bandpass filter and thereby increase the bandwidth of the data. This procedure is repeated until the whole bandwidth of the data has been included. The criteria that I use to allow the algorithm to move to the next stage include the total number of iterations and a user-defined threshold value of relative change of the data misfit between the n^{th} iteration compared to the $n - 2^{\text{th}}$ iteration.

2.4 FWI workflow

To summarize the FWI workflow I illustrate a flowchart of all the individual parts, in Figure 2.4. The figure contains the starting model \mathbf{m}_n of every iteration, which generates the synthetic data \mathbf{d}_{syn} after performing the forward modeling, e.g., the solution of the 2D elastic wave equation (equation 2.39) given the elasticity of the medium and acquisition setup. At this point the recordings are inserted into the algorithm expressed as \mathbf{d}_{obs} .

By subtracting the data vectors \mathbf{d}_{obs} and \mathbf{d}_{syn} we obtain the data residuals $\Delta\mathbf{d}$ (equation 2.31). The residuals are then used in the misfit function $E(\mathbf{m}_n)$ in order to evaluate the difference between the data sets. They are further used as adjoint sources and backpropagated in order to compute the gradients of $E(\mathbf{m}_n)$ with respect to each model parameter. The gradients are calculated for each shot and then multiplied by the preconditioner \mathbf{P} and summed. The parabolic search calculates an optimal step length and with the inclusion of the gradient direction using the PCG optimization method, the final update

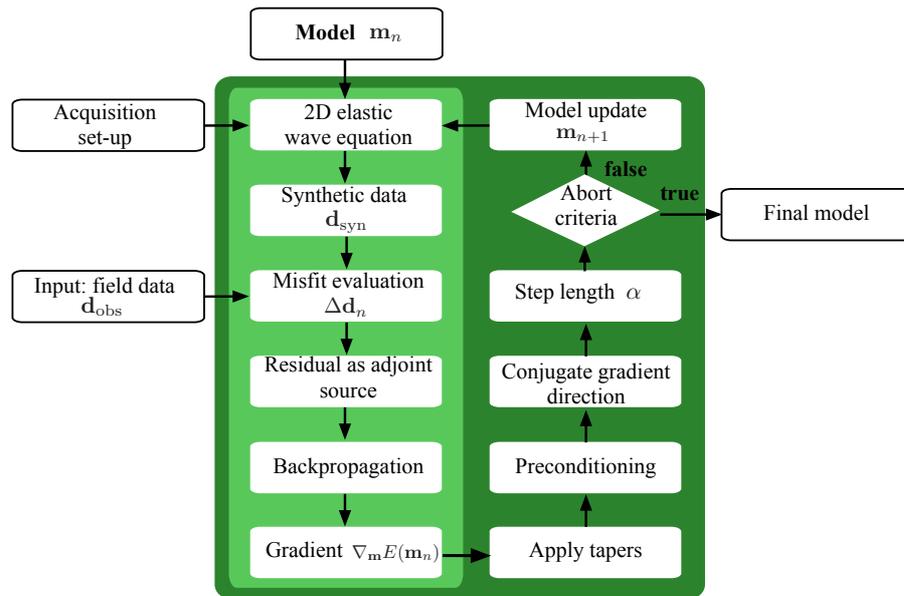


Figure 2.4: Schematic overview for one iteration of the FWI process. The dark green box corresponds to the loop over the number of iterations, while the light green box indicates the individual steps for each source.

is obtained (equation 2.38). In case the abort criteria were met the inversion finishes. Otherwise, it is repeated using as initial model, the updated one (\mathbf{m}_{n+1}).

Chapter 3

Numerical study of seismic scattering

3.1	Scattering principles in shallow seismics	30
3.1.1	Isotropic body wave scattering	30
3.1.2	Scattering in heterogeneous media: Born approximation versus FD modelling	32
3.1.3	Viscoelastic wave equation of scattered wavefield	32
3.1.4	Instantaneous energy density of P- and S-waves	33
3.1.5	Numerical setup	35
3.2	Scattering effects caused by perturbations in the elastic parameters	35
3.2.1	Model description	35
3.2.2	Results	36
3.2.3	Conclusions	38
3.3	Dependency of scattering effects on inclusion characteristics	38
3.3.1	Influence of the depth of the inclusion	40
3.3.2	Influence of the size of the inclusion	40
3.3.3	Influence of the shape of the inclusion	40
3.3.4	Influence of the composition	42
3.3.5	Conclusions	42
3.4	Summary	42

Part of this chapter was published as: Athanasopoulos, N., Bohlen, T. (2019), Shallow-seismic wavefield scattering and implications for viscoelastic FWI, Conference Proceedings in 25th European Meeting of Environmental and Engineering Geophysics, DOI: 10.3997/2214-4609.201902396.

As I discussed in Chapter 1, FWI suffers from so-called cross-talk between viscoelastic parameters: certain combinations of viscoelastic parameters may cause a similar seismic response and can therefore not be discriminated. In this chapter, I investigate the seismic response for a shallow structure by a series of numerical experiments using the 2D finite-difference method for solving the viscoelastic wave equation. As shallow seismic waves are dominated by surface waves I mainly analyze the scattering of surface-to-body waves and vice versa. This numerical study gives fundamental insight into the composition of shallow seismic recordings and provides a framework for their interpretation in future applications of FWI.

3.1 Scattering principles in shallow seismics

In typical applications of land seismics, we observe severe scattering of the seismic energy propagating into the shallow subsurface. This can occur due to various reasons, such as bedrock discontinuities, localized heterogeneities, dry river beds, karst environments and many others. An example of such case is illustrated in Figure 3.1. The scattering induced by such environments leads to strong surface-to-surface and body-to-surface wave combinations (Levander, 1990).

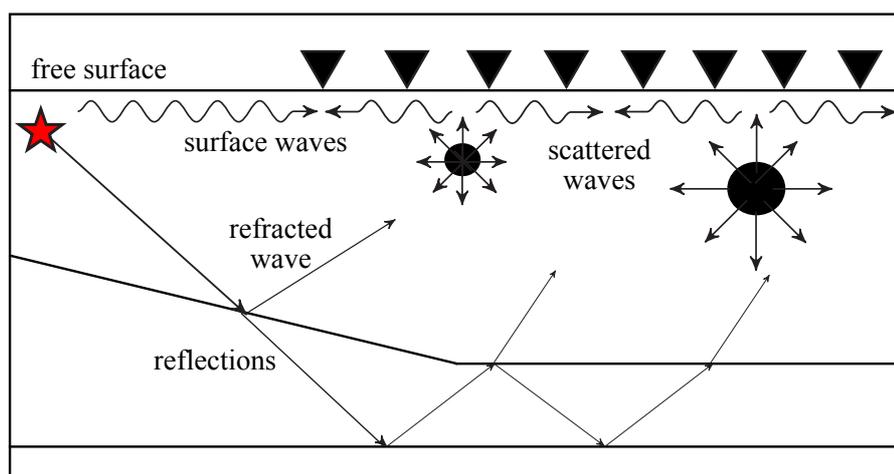


Figure 3.1: Sketch of seismic scattering due to shallow heterogeneities.

In the past decades, there have been many studies that focused on delineating the influence of arbitrarily shaped subsurface structures from seismic recordings. The approach they followed is based on understanding how seismic energy scatters in the Earth as it propagates through subsurface heterogeneities.

3.1.1 Isotropic body wave scattering

Following Wu and Aki (1985) and Aki and Richards (2002), I start off by examining the radiation patterns of body waves propagating in isotropic elastic media. The diffraction pattern or radiation pattern describes the distribution of amplitude of scattered energy as a function of scattering/diffraction angle. It depends on the directions of both the incident and scattered wavefields and describes the ratio between between

their amplitudes. It can be estimated by the derivative of the wave equation with respect to material matrix, \mathbf{M} (see equation 2.39).

Figure 3.2 shows the theoretical amplitude radiation patterns which are estimated using a ray and Born approximation (Wu and Aki, 1985) for isotropic homogeneous media. The diffraction pairs P-P, P-SV, SV-P, SV-SV and SH-SH are a result of perturbations in v_p , v_s and ρ for different planes.

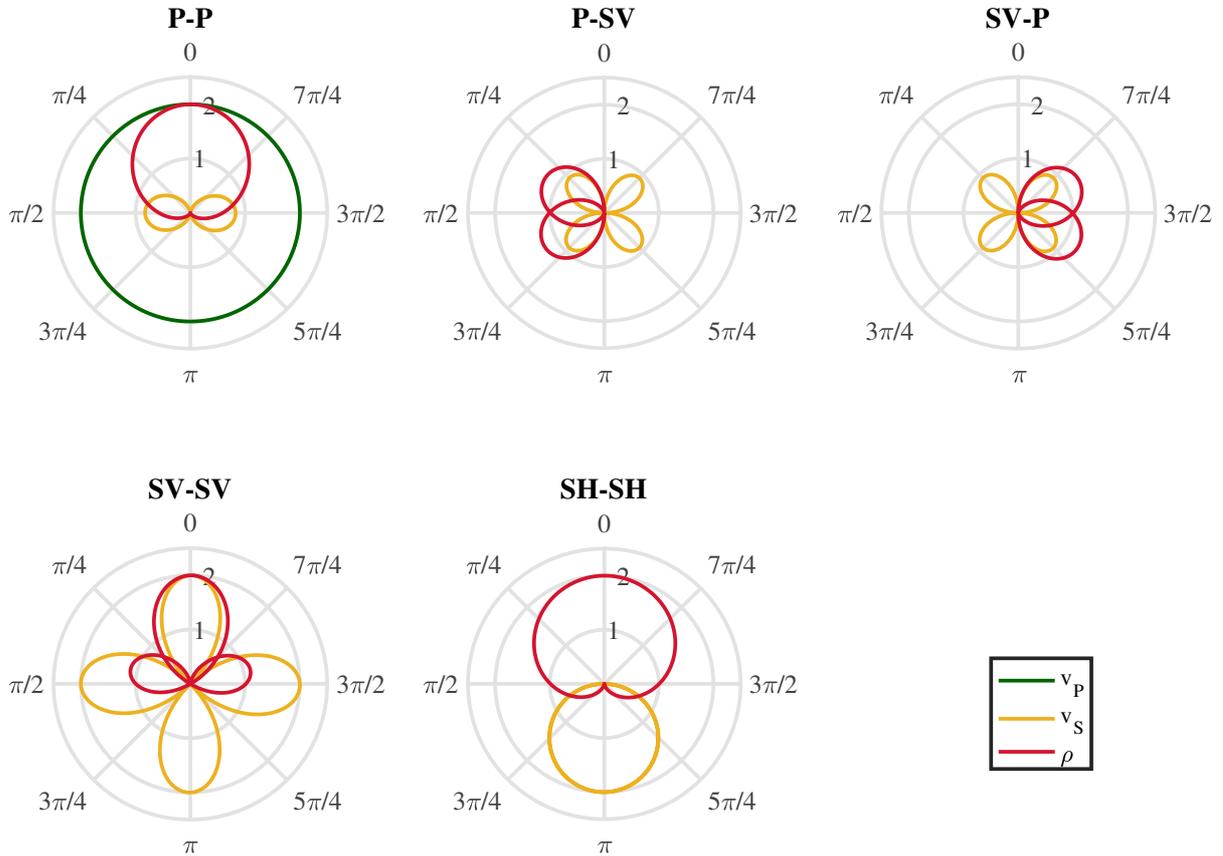


Figure 3.2: The radiation patterns of v_p , v_s and ρ perturbations for different modes.

Although the radiation patterns shown in Figure 3.2 strictly hold for homogeneous media only and the Earth is highly inhomogeneous, we can still see that at certain scattering angles there is an overlap between the different radiation patterns. This phenomenon which we meet also in heterogeneous media is responsible for parameter trade-offs during seismic processing and inversion techniques (Virieux and Operto, 2009).

In other words, the identical scattering pattern caused by certain parameters makes their distinction and separation in seismic recordings a very tedious task, if at all possible. Finally, it is also worth noting that the existence of shear rigidity increases greatly the complexity of the scattering pattern (Wu, 1989), i.e., when dealing with elastic versus acoustic wave propagation. In the following, I will try to identify radiation patterns which characterize not only body waves but also the distinct behavior of surface waves.

3.1.2 Scattering in heterogeneous media: Born approximation versus FD modelling

The radiation patterns I show in Figure 3.2 are not valid for inhomogeneous media. In the past decades, many studies have shown solutions of forward (Gubernatis et al., 1977; Wu and Aki, 1985; Wu, 1989) and inverse (Beylkin and Burridge, 1990; Blonk et al., 1995) scattering problems for imaging inhomogeneous media. Two of the main theories behind these studies are based on the perturbation method and the first-order Born approximation, while the third and relatively recent method is the use of FD numerical modelling.

Gubernatis et al. (1977) have shown that the scattering problem can be cast in the form of an integral equation similar to the scattering equation of quantum mechanics, by solving an iterative scheme up to the first iteration to calculate the scattering response (therefore Born approximation).

Wu and Aki (1985) further exploit this concept by expressing the scattering in terms of a decomposition into a reference and a perturbed medium. The scattering can then be implied by the wavefield induced at the position of the scatterer, acting as a secondary source. Therefore, while radiation patterns characterize scattering for point perturbations, they can be generalized to arbitrary perturbations that fit into the Born approximation (Wu and Aki, 1985), if the contrasts between the material parameters are not significant.

In the case of shallow seismics, the subsurface heterogeneities that we would like to examine and whose response we would like to identify, have a size comparable to the seismic wavelength. In such cases the equivalent source of scattering due to the existence of such a body/inclusion cannot be regarded as a point source, due to the fact that the phase differences of the incident field and of the scattered field from different parts of the inclusion are very dominant (Wu, 1989). Only in cases where the total scattered field is much weaker than the incident field (*Mie scattering*) and the size of the scatter is small (*Rayleigh scattering*) can the Born approximation be of use for calculating the scattered wavefield (Wu, 1989; Eaton, 1999).

The Finite-difference methods gained significant popularity with the improvements in computational power. They provide the ability to store particle velocities for any number of gridpoints and for all discrete timesteps (Robertsson et al., 1994; Bohlen, 2002). Robertsson and Chapman (2000) proposed a finite-difference-injection method where alterations of the seismic data can be incorporated to update seismic recordings. However, the efficiency of the method is limited to only first-order interactions with the scatterers. Bohlen et al. (2012) studied the influence of *composition* and *shape* of massive sulfide ore deposits hosted in crystalline rocks on the full scattered wavefield through a setup of numerical modelling experiments.

In this work, I will perform a similar study to Bohlen et al. (2012), where I will focus on the influence of seismic scattering caused by various types of inclusions, considering, however, the strong impact that the surface waves have on the recordings.

3.1.3 Viscoelastic wave equation of scattered wavefield

I formulate the theoretical background required in order to numerically examine the effects of elastic wave scattering caused by near-surface anomalies. For this purpose, I perform full wavefield simulations using the elastic 2D Finite-difference method which accounts for multiple scattering and mode conversions.

I consider 2D wave propagation in an isotropic viscoelastic medium described by the generalized standard linear solid (Carcione et al., 1988; Robertsson et al., 1994; Bohlen, 2002), expressed by equation 2.18. Two different models are used: a base background model that will remain unaltered throughout this study and a perturbed model as shown in Figure 3.3, which is the same as the background model with the addition of an inclusion/ body scatterer. The perturbed model will be modified in Section 3.3 for each case, dependent on the effects I wish to examine.

I perturb the P- and S-wave velocities (v_p and v_s , respectively), the density (ρ) and the attenuation parameters for P and S waves (q_p and q_s , respectively) as follows :

$$v_p' \rightarrow v_p + \delta v_p, \quad v_s' \rightarrow v_s + \delta v_s, \quad \rho' \rightarrow \rho + \delta \rho, \quad q_p' \rightarrow q_p + \delta q_p \quad \text{and} \quad q_s' \rightarrow q_s + \delta q_s \quad (3.1)$$

For the first study (section 3.2), each time I change only one of the viscoelastic parameters to study its individual contribution to the recorded data, while in section 3.3 I will perform a series of changes of the inclusion and observe the scattering response.

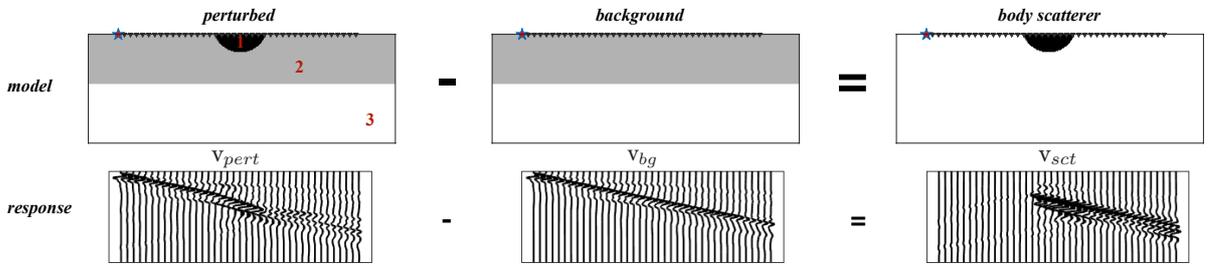


Figure 3.3: Calculation of the scattering response of a shallow local anomaly by subtracting the seismic response of the background model from the perturbed, i.e. total wavefield.

We are dealing with body-scatterers/inclusions and therefore I do not use the Born approximation as already discussed, but rather numerically investigate the effects of perturbations on the elastic parameters. By subtracting the perturbed (total) wavefield \mathbf{v}_{pert} from the background wavefield \mathbf{v}_{bg} I obtain the difference between these two, \mathbf{v}_{sct} , for each of the five cases shown in equation 3.1, given by

$$\mathbf{v}_{sct_i} = \mathbf{v}_{pert_i} - \mathbf{v}_{bg_i} \quad \text{where} \quad i = v_p, v_s, \rho, q_p \quad \text{and} \quad q_s. \quad (3.2)$$

A perturbation in a material parameter acts as a secondary source producing a wavefield that I refer to as *scattered* wavefield or scattering response (\mathbf{v}_{sct} , Figure 3.3).

3.1.4 Instantaneous energy density of P- and S-waves

In the case of isotropic heterogeneous media, we can decompose the response of an inclusion into a combination of two separate homogeneities (Athanasopoulos and Bohlen, 2019). This can be done since I examine a case of a discontinuous heterogeneity where the media are homogeneous outside and inside

of the inclusion, while sharp impedance contrasts occur at the boundaries. The equation of motion can therefore be written for a homogeneous space in vector format as:

$$\rho_i \ddot{\mathbf{u}} - (\lambda_i + 2\mu_i) \nabla \nabla \cdot \mathbf{u} + \mu_i \nabla \times \nabla \times \mathbf{u} = \mathbf{f} \quad \text{for } i = 0, 1, \quad (3.3)$$

where \mathbf{u} is the displacement vector, $\ddot{\mathbf{u}}$ its second time derivative and \mathbf{f} is the body force vector. ∇ , $\nabla \cdot$ and $\nabla \times$ are the gradient, divergence and curl operators, respectively. The space outside the inclusion is given by $i = 1$ and inside by $i = 0$.

In order to separate the responses of P and S waves, I calculate the curl and divergence of the particle velocity fields. Every vector field can be decomposed into a rotation-free and a divergence-free part. Since the divergence of a curl and the rotation of a divergence are zero, we get two independent equations which correspond to the viscoelastic wave equation for the propagation of compressional ($\nabla \cdot \mathbf{v}$) and shear ($\nabla \times \mathbf{v}$) waves (Morse and Feshbach, 1953). The vector product $\nabla \times \mathbf{v}$ corresponds to a rotation, describing a change of shape without any volume change, e.g., pure shearing, while the term $\nabla \cdot \mathbf{v}$ describes a volume change (compression and dilatation), which also contains some rotation-free shearing (except for hydrostatically compressed media). After applying the divergence and curl on 3.3 and neglecting the source terms, we get the decoupled (for homogenous media) wave equation for P and S waves:

$$\nabla \cdot \ddot{\mathbf{u}} - \frac{\lambda_i + 2\mu_i}{\rho_i} \nabla^2 (\nabla \cdot \mathbf{u}) = 0 \quad (3.4)$$

$$\nabla \times \ddot{\mathbf{u}} - \frac{\mu_i}{\rho_i} \nabla^2 (\nabla \times \mathbf{u}) = 0 \quad (3.5)$$

$$(3.6)$$

We can then define the instantaneous energy density (or intensities) for P- and S-wave particle velocities (E_p and E_s , respectively) in 2D following Morse and Feshbach (1953) by

$$E_p = \sqrt{\lambda + 2\mu} (\nabla \cdot \mathbf{u}) = \sqrt{\lambda + 2\mu} \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} \right) \quad \text{and} \quad (3.7)$$

$$E_s = \sqrt{\mu} (\nabla \times \mathbf{u})_z = \sqrt{\mu} \left(\frac{\partial u_x}{\partial y} - \frac{\partial u_y}{\partial x} \right). \quad (3.8)$$

In the above statement, we have preserved the divergence and curl sign information, while showing relative compressional and shear energy amplitudes (contribution of density is removed due to equality) similar to Dougherty and Stephen (1988). This walk-around is an essential step for studying the scattering response, because now we can assess the problem numerically. Morse and Feshbach (1953) have shown that exact solutions for scattering due to inclusions exist only for the case of a uniform sphere or a uniform circular cylinder in an infinite homogeneous medium. Therefore, for most cases of complex objects numerical methods represent the only viable option for studying the effects of scattering.

In the next section, I will use these principles to provide snapshots of the background and scattered wavefields based on the numerical solution of shallow seismic scattering from body and surface waves. Instead of displacement fields, I will show the particle velocities for consistency with the notations of the previous chapter.

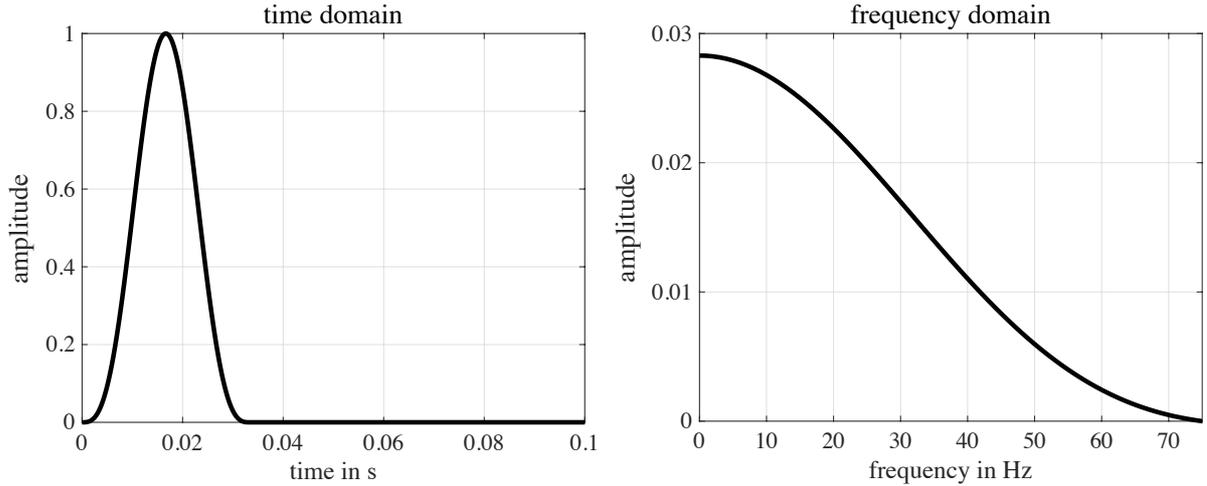


Figure 3.4: Left: A cubed sine wavelet defined by equation 3.9 with a center frequency of 30 Hz. Right: Its amplitude spectrum in the range from 0 Hz to 70 Hz.

3.1.5 Numerical setup

I examine the effects of near-surface heterogeneities on the recorded waveforms by simulating equations 2.18 using the background, perturbed and scattered models for each parameter as shown in Figure 3.3. The acquisition geometry of the study consists of a linear profile of 39 vertical component geophones with an equidistant spacing of 1 m. I use a vertical source with dominant frequency of 30 Hz. As source signal $s(t)$, I choose a cubed sine defined in the following way:

$$s(t) = \begin{cases} \sin^3(f_d \pi t), & \text{if } 0 < t < 1/f_d \\ 0 & \text{otherwise} \end{cases}, \quad (3.9)$$

where f_d is the dominant frequency. This wavelet and its amplitude spectrum are shown in Figure 3.4.

Receivers and sources are located at the surface, both to recreate a realistic scenario and to study surface-to-surface scattering which is of great interest for my case. The model is shown in Figure 3.1 and has a size of 50 m by 18 m. The interface between the top layer and the half space is at a depth of 8 m.

In order to get a better insight into the impact of various subsurface characteristics on the phase and amplitude of the scattered surface and body waves, I designed a series of numerical tests which are summarized in Table 3.1.

3.2 Scattering effects caused by perturbations in the elastic parameters

3.2.1 Model description

Figure 3.3 shows the models that I used in this section. The goal is to get a better understanding of how each individual parameter class affects the recordings and how the forward and back-scattered energy distribution caused by the inclusion look like. The body scatterer is located directly below the surface and it has the form of a semi-circular shape with dimensions 8.3 m (horizontal) by 2.9 m (vertical). The

Table 3.1: Examined cases and their corresponding figure number.

Studying effect	Figure #
individual parameter perturbation	3.5
depth of inclusion	3.6
size of inclusion	3.7
shape of inclusion	3.8
composition of inclusion	3.9

elastic parameters for each layer and the scatterer are given in Table 3.2. The values were chosen to be close to the subsurface structure that I investigate in the field data (Chapter 6), and better link the sensitivity results from my perturbation study to the actual recordings.

Table 3.2: Material properties of the models shown in figure 3.1. The percentage of reduction from the background model is shown as well.

	v_p (m/s)	v_s (m/s)	ρ (kg/m ³)	q_p/q_s (·)
1: body scatterer	355 (−23%)	92 (−45%)	1500 (−12%)	20 (−60%)
2: top layer	460	165	1700	50
3: bottom layer	1745	275	2000	200

The results of the simulation experiments are shown in Figure 3.5. The wavefield for the background medium is shown in Figure 3.5a.1-3. This wavefield is dominated by the Rayleigh wave which appears in the S-wave intensity (E_s , Figure 3.5a.1) and has weak P-wave contribution (E_p , Figure 3.5a.2 and 3.5a.3). By inserting a shallow anomaly with perturbation in all three elastic parameters (Table 3.2) we obtain the total wavefield (Figure 3.5b.1-3). In Figures 3.5c.-g., I show wavefields for single parameter perturbations in v_p , v_s , ρ , q_p and q_s , respectively, in order to further study the individual effects produced by each of the elastic parameter.

3.2.2 Results

Perturbation in v_p (−23%), Figure 3.5c: The scattered wavefield has similar intensity for both P- and S-waves. Mainly forward scattering into the direction of the Rayleigh wave occurs. A relatively high back-scattering intensity of P waves is also observed. The recorded seismic response at the surface is composed of forward scattered P and S waves and surface waves. This indicates that both the coda of the first arrival and the Rayleigh wave may contribute similarly to update the P-wave velocity model during FWI. Its amplitude, however, is small compared to the reference, revealing the difficulty of retrieving accurate P-wave velocities while using least-squares inversion schemes, in the context of near-surface

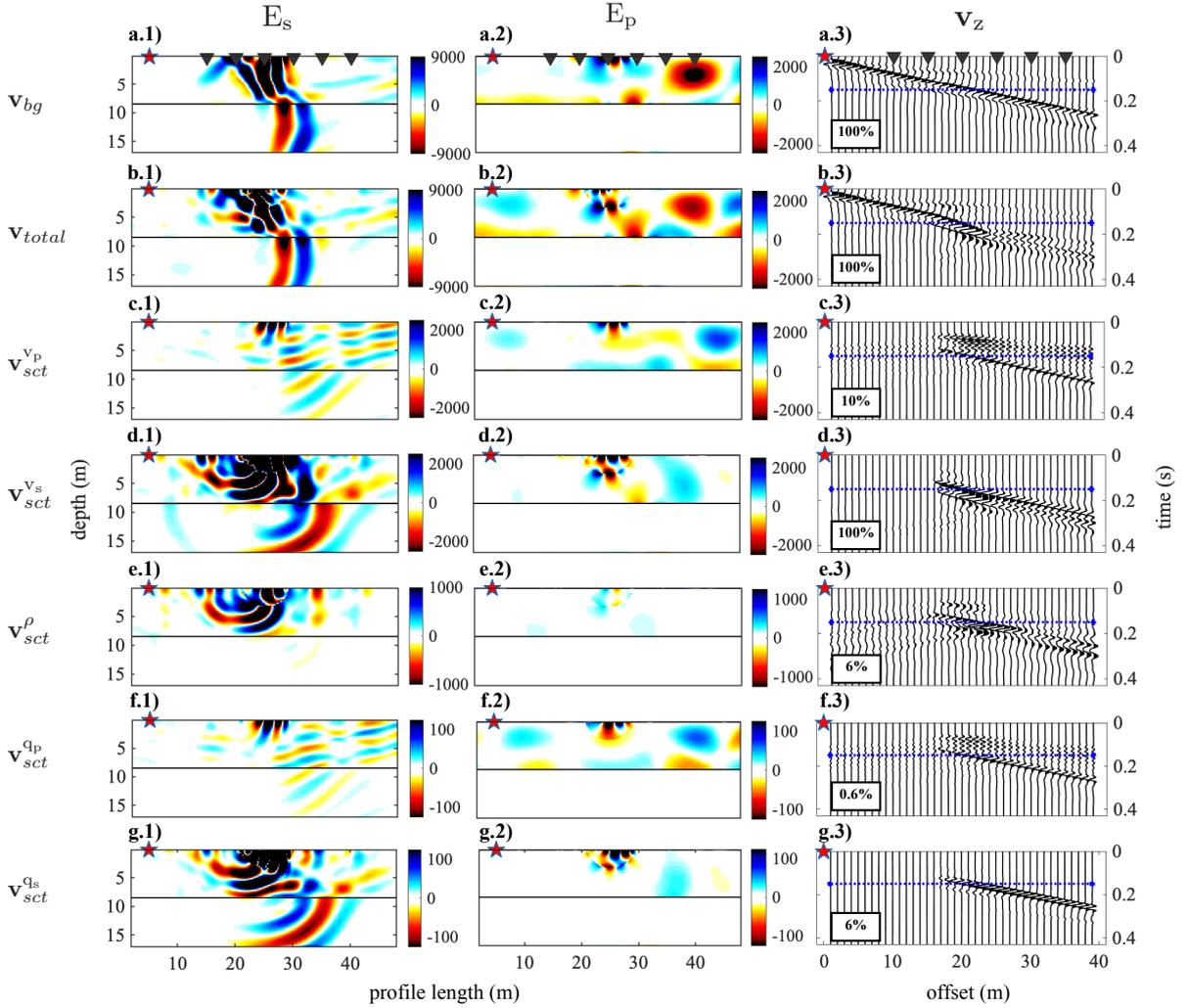


Figure 3.5: Snapshots (left two columns) of the S-wave (E_s) and P-wave (E_p) intensity of the background wavefield (top row) and scattered wavefields caused by perturbations of the viscoelastic parameters (Table 3.2). The corresponding seismograms of vertical particle velocity are shown in the right column. Triangles in the top-most figures mark the positions of some receivers and the stars indicate the shot location. For better visualization the amplitudes of the scattered wavefield are clipped to lower values as indicated by the colorbars. The percentages in the right column give the relative maximum amplitude of the scaled seismograms compared to the background (a.3).

seismics. Hierarchical approaches using time-windowing can therefore improve the reconstruction of P-wave velocity (Athanasopoulos and Bohlen, 2016).

Perturbation in v_s (-45%), Figure 3.5d: The scattering response due to an S-wave velocity perturbation is dominated by Rayleigh-wave scattering into the forward direction. The amplitudes of the forward scattered Rayleigh wave are comparable to the amplitude of the direct Rayleigh wave (Figure 3.5a). This indicates that FWI of Rayleigh waves is most sensitive to shallow S-wave perturbations, i.e., a reliable reconstruction of the S-wave velocity model can be achieved.

Perturbation in ρ (-12%), Figure 3.5e: The scattering response due to density perturbation differs fundamentally from perturbations in seismic velocities. From a physical point of view, density mainly influences the amplitude of waves, while the P- and S-wave velocity influence both traveltimes and amplitudes. We can observe a surprisingly strong back-scattering of P- and especially S-waves caused

by both the direct P- and Rayleigh wave. In particular the intensity of S-waves shows that the back-scattered energy is almost half the intensity of the forward scattering. The unique scattering pattern due to density perturbation, implies that in near-surface applications the retrieval of accurate density model is feasible. This is in contrast to the experience acquired from the hydrocarbon exploration industry and the use of body waves (Figure 3.2), where density radiation pattern overlaps with the ones of P- and S- wave velocities. Separating forward and back-scattered waves in FWI could potentially improve the reconstruction of density models. Additionally, one could employ an offset dependent taper in order to preserve the relative amplitude differences versus the offset, which should also assist FWI to retrieve more accurate the density model. However, we observe that the amplitude of a density perturbation constitutes a change of only 6 % relative to the reference recordings, making density especially challenging to invert for in FWI and therefore it is treated as a second-order parameter.

Perturbation in q_p and q_s (-60%), Figures 3.5f,g: The scattered wavefield due to perturbations in attenuation of P- and S-waves is similar to the intensities of the respective velocity perturbations (Figures 3.5c and 3.5d). This indicates that the velocity dispersion caused by attenuation controls the scattering behavior. Amplitude loss is negligible here due to the small size of the structure (body scatterer). Due to the similarity of the scattering responses and intensities of velocities and attenuation we can expect a significant cross-talk between v_p - q_p and v_s - q_s , making q_p and q_s the most challenging parameters to invert for. This effect may be reduced by accounting for frequency-dependent amplitude decay due to attenuation, which is not considered in Figure 3.5.

3.2.3 Conclusions

To summarize, we have seen that for the P-wave velocity changes the seismic response is composed mainly of forward scattered P-and S-waves and surface waves, while for changes in the S-wave velocity the seismic response is dominated by Rayleigh-wave scattering into the forward direction. A density perturbation produces strong back-scattering of P- and especially S-waves caused by both the direct P- and Rayleigh wave. Cross-talk between attenuation and velocity for both P- and S-waves is caused by the similarity of the corresponding scattering responses. This cross-talk may be reduced by incorporating the amplitude loss with offset caused by attenuation during FWI. In all cases of subsurface inclusions the scattering response of the shear wave is more dominant than the that of the P wave. Therefore, after a large distance of propagation and scattering, the scattered wave will be composed mainly of S-wave.

3.3 Dependency of scattering effects on inclusion characteristics

In this section, I modify the characteristics of the inclusion in terms of its location, size, shape and composition. I focus only on the scattered field and decompose it into its P- and S-wave intensities to better understand how the scattering develops in the forward and backward direction. Additionally, all the individual material contributions are given in Appendix A. In the following, I will show the P- and S-wave intensities from the previous chapter (column *a* of Figures 3.6, 3.7, 3.8 and 3.9) and compare them with variations of the inclusions.

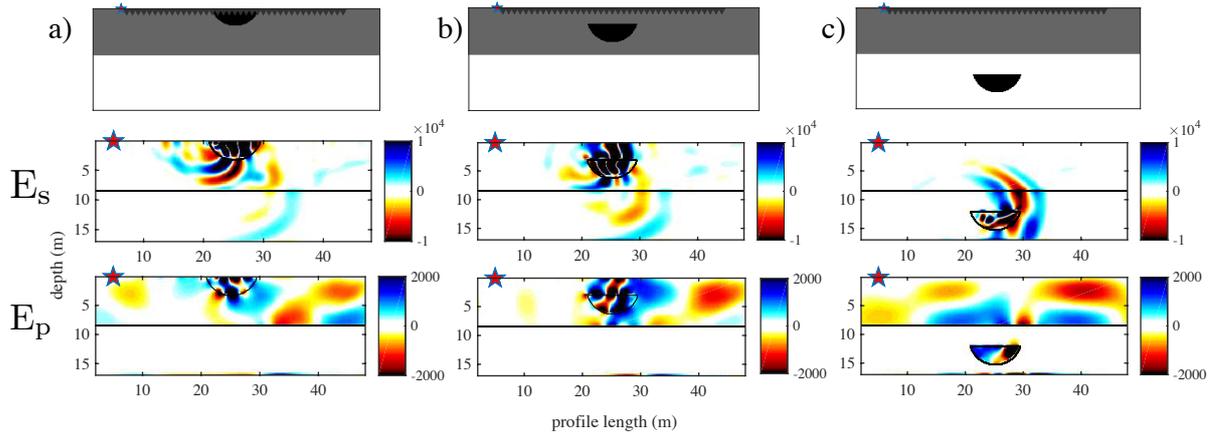


Figure 3.6: Scattered wavefields caused by perturbations of the viscoelastic parameters for various depths of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b.

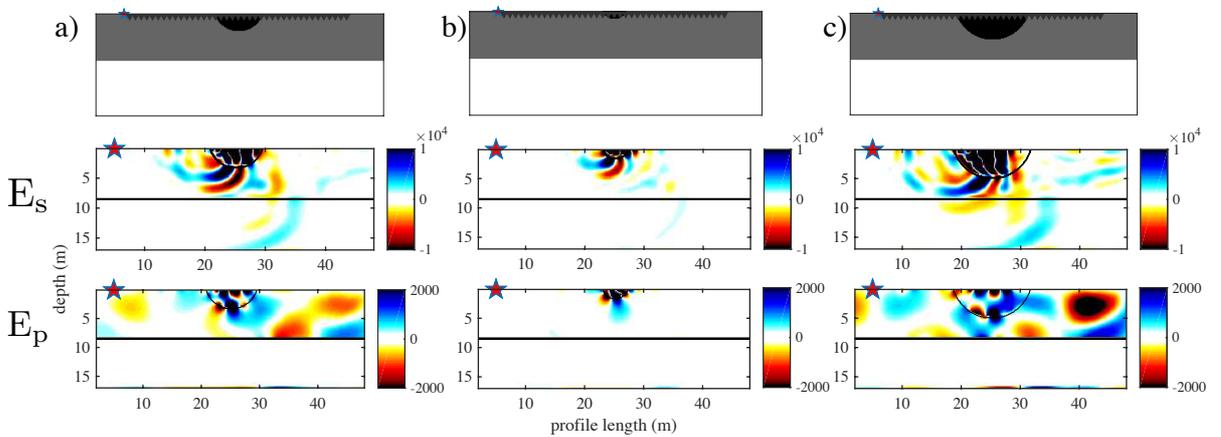


Figure 3.7: Scattered wavefields caused by perturbations of the viscoelastic parameters for various sizes of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b for the first column and then are halved for the second and doubled for the third column.

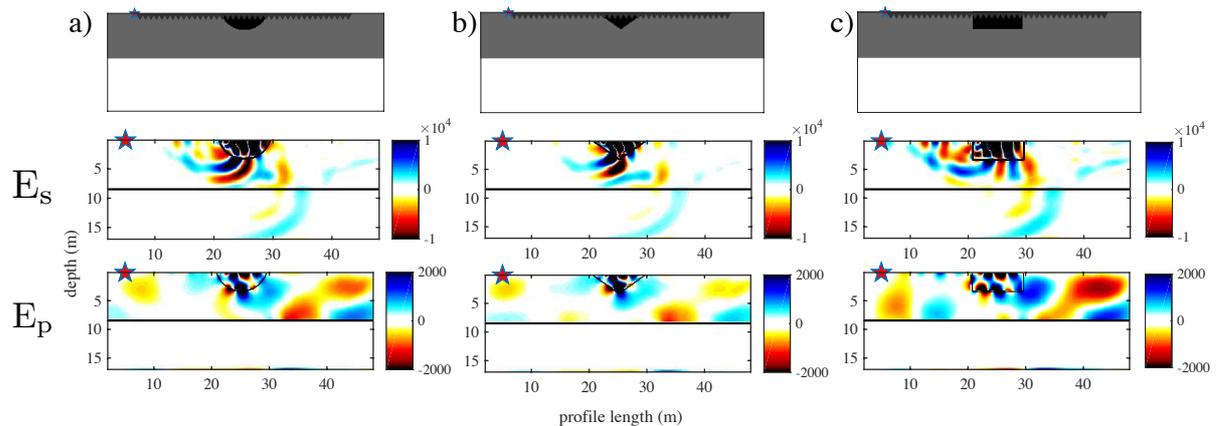


Figure 3.8: Scattered wavefields caused by perturbations of the viscoelastic parameters for various shapes of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b.

3.3.1 Influence of the depth of the inclusion

I keep the dimensions of the inclusion the same but I shift the initial midpoint of the center of the semicircle vertically from -1.2 m (i.e., 1.2 m above the subsurface) in Figure 3.6a to +1.2 and 10.8 m in Figures 3.6b and c, respectively. We observe that the excited surface wave energy is getting much weaker as the inclusion is buried deeper in the subsurface and completely vanishes for inclusions below a discontinuity. Since the inclusion has a semicircular geometry when the incident wave arrives, we observe significant changes in the kinematics of the scattered wavefield (Appendix A, Figure A.1), which are considered to be *shape* effects (no longer a smooth transition due to the sharp edge of the inclusion). The amplitude of the forward scattered P wave is higher when the inclusion is buried (at around $\lambda_p > 3$) below the discontinuity and exhibits similar back-scattered energy, compared to the other two cases. This could be due to the reduction of energy that would be used otherwise to convert body to surface wave. The amplitude of the S-wave remain similar for all three cases, which agrees with the the previous assumption.

3.3.2 Influence of the size of the inclusion

In this numerical test, I evaluate the scattering response induced by different sizes of the inclusion (Figure 3.7a, b and c), with maximal vertical length of 3, 1.2 and 4.4 m, respectively. As expected, the amplitude dependence of the scattered energy on the size of the inclusion is quite prominent and constitutes entirely a *shape* effect. In all three cases the scattered energy is higher in the forward direction. In Figure A.2 (Appendix A), we see a clear distinction between the three cases as far as the directivity of the scattered energy is concerned and how we get different body-to-surface wave conversions.

3.3.3 Influence of the shape of the inclusion

I investigate the effects of the *shape* of the scatterer by comparing three distinct geometrical shapes of the inclusion: a semi-circular, a triangular and a square shape. In general, when dealing with large complex shaped objects the scattering center will not lie necessarily at the center of the inclusion (Bohlen et al., 2012). A strong curvature in the shape of the inclusion will have a direct impact on the incident P wave, which will arrive first. This is prominent also in my results, since the intensity of the P wave shows a different scattering response for each of the three different cases (Figure 3.9, third row).

In each of the three examined cases, and for both P- and S-wave scattering, the shape of the inclusion has a dominant impact on the directivity of the scattered energy. There are no polarity reversals and the amplitudes of both intensities are similar for the three cases (Appendix A, Figure A.3). We generally observe the tendency of the forward-scattered energy to be dominant over backward scattering. However, this is not the case for the square-shape inclusion, at least as long as shear waves are concerned, making it obvious that this is entirely a *shape* effect. This was also shown by Bohlen et al. (2012), where for their experiments the Rayleigh scatterer (point-scatterer) caused the maximum P-P energy to propagate backwards, while for different shapes this was no longer the case. In my example, this can be observed to a lesser extent, since the size of the inclusions are of the same magnitude. We do, however, observe this effect as a relative increase or decrease of the backward and forward scattering, depending on the shape of the inclusion.

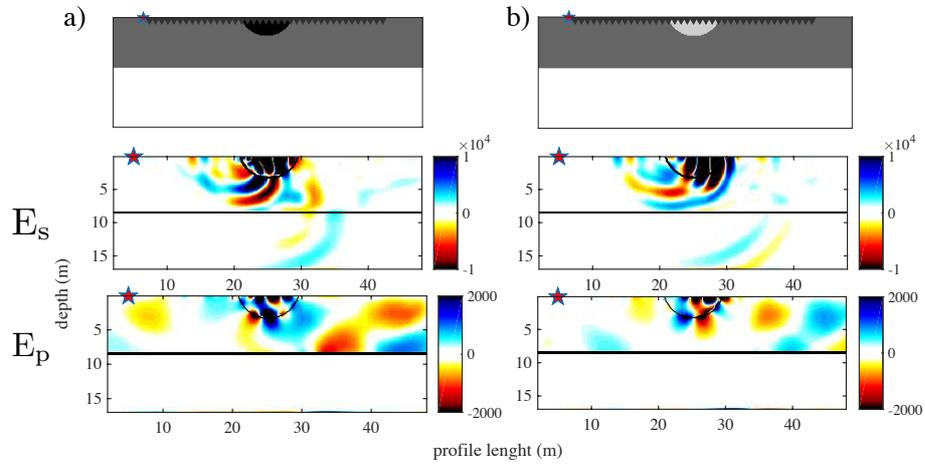


Figure 3.9: Scattered wavefields caused by perturbations of the viscoelastic parameters for different values of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b.

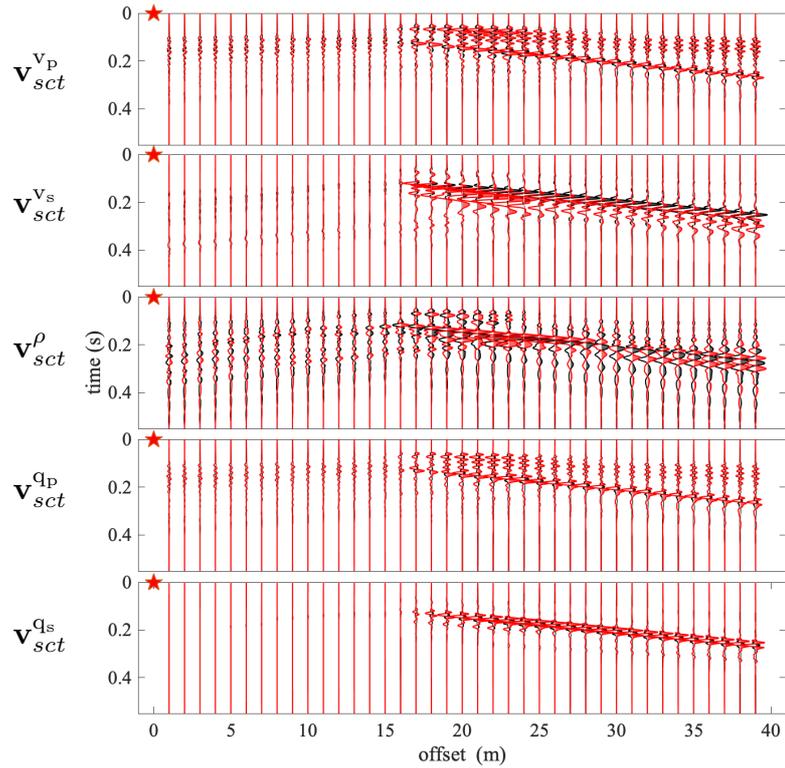


Figure 3.10: Seismograms showing the scattering response caused by perturbations of the viscoelastic parameters for different values of the inclusion. The two different compositions from the models in Figure A.4 are shown as red for negative and black for a positive change. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.

3.3.4 Influence of the composition

In this test, I swap the sign of the change of the inclusion relative to the background model (Table 3.1, first row). This refers to a change in the *composition* of the scatterer, since no *shape* changes occur. Figures 3.9 and 3.10 show that a variation in the composition of the scatterer leads to differences of the amplitude-versus-angle patterns of P- and mainly S-wave, as seen by the scattered velocity wavefields in the region of the refracted and surface waves. Both phase and amplitude of the recording are affected. The amplitude distribution observed in both cases have a tendency to be higher in the forward direction, signifying a that this constitutes *shape* effect.

Additional characteristic of this case is the phase reversals that are observed and become more obvious after overlapping the two different compositions for all individual perturbations (Figure 3.10). In my test, both forward and backward scattering are subject to phase reversals. Bohlen et al. (2012) also observed such behavior in their body-to-body wave scattering study, while they have further shown that the phase reversal scattering angle seems to be independent of the shape, even for large inclusions.

Figure 3.9 shows how the composition constrains significantly the development of the wavefield within the inclusion, where for both P and mainly S wave the scattered field has distinct behavior (in terms of the amplitude of the scattered energy, i.e., dynamics). The differences of the amplitudes of the arrival times for the far-offset traces (Figure 3.10) is also an indication of this characteristic.

3.3.5 Conclusions

This study presented insights on how the shallow scattering is affected by different characteristics of the inclusion. I thoroughly examined scattered waves due to body-to-body and body-to-surface wave conversions. We have seen that the directivity imposed by the *shape* is a first-order characteristic of the scattered wavefield, whereas composition effects are of secondary importance since mainly amplitude changes occur and phase-reversals. The above shows a clear distinction between *composition* and *shape* characteristics in the seismic scattering response. The *composition* effects usually affect the dynamic characteristics of the wavefield, whereas the kinematic behavior is mainly controlled by *shape* effects.

3.4 Summary

In this chapter, the focus was to better understand the footprint that a subsurface structure will cause in the viscoelastic parameters. A similar response, which was evaluated here in regards to similarity in the scattering pattern of each parameter, signifies the difficulty to accurately retrieve the individual structure of each parameter when using FWI or any other least-squares inversion scheme. By obtaining a more clear distinction in terms of their directivity (forward and backward), kinematic (wave conversions) and dynamic (amplitudes) behavior, one can better employ appropriate schemes to separate the contributions of each parameter and successfully reconstruct them. These findings will be the base of the methodologies that will be examined in the next chapter and applied to FWI.

Chapter 4

Synthetic studies

4.1	Multiparameter inversion	44
4.1.1	Simultaneous joint inversion	44
4.1.2	Joint objective function of P-SV and SH waves	45
4.1.3	Joint gradients	45
4.1.4	Numerical setup and model description	46
4.1.5	Spatially correlated models	47
4.1.6	Spatially uncorrelated models	50
4.1.7	Influence of the initial P-wave velocity model	54
4.1.8	Conclusions	56
4.2	Time - frequency windowing	56
4.2.1	Theory and methodology	57
4.2.2	Effect of the taper function on frequency spectra	58
4.2.3	Effect of the taper function with offset	59
4.2.4	Synthetic reconstruction test	60
4.2.5	Results	60
4.2.6	Conclusions	64
4.3	Parameter sensitivity in far offset acquisition setup	64
4.3.1	Model description	64
4.3.2	Acquisition setup	66
4.3.3	Results	67
4.3.4	Conclusions	69
4.4	Structural similarity constraints	69
4.4.1	Theory of cross-gradients	70
4.4.2	Synthetic study and results	72
4.4.3	Conclusions	73
4.5	Summary	74

Part of this chapter was published as:

Athanasopoulos, N., Manukyan, E., Bohlen, T., Maurer, H. (2018), Accurate reconstruction of shallow P-wave velocity model with time-windowed elastic full-waveform inversion, *Conference Proceedings in 80th EAGE Conference and Exhibition 2018*, DOI: 10.3997/2214-4609.201801305.

Athanasopoulos, N., Manukyan, E., Bohlen, T., Maurer, H. (2018), Application of structurally similarity constraints on time-domain elastic full-waveform inversion", *Conference Proceedings in 24th European Meeting of Environmental and Engineering Geophysics 2018*, DOI: 10.3997/2214-4609.201802630.

In this chapter a series of synthetic FWI tests is shown, in order to obtain a better understanding of the problems which can arise during FWI of shallow seismics and obtain ways to solve them. The main focus is to evaluate the reliability of the inversion results by examining the resolvability of the various elastic parameters in different cases. Effects of the inter-parameter trade-off are evaluated through the use of spatially uncorrelated models, while comparing the results to those obtained from spatially correlated models. Additionally, an emphasis is given on multiparameter FWI where both Rayleigh and Love waves are accounted during the inversion. The observations of the parameter sensitivity tests from Chapter 3 are used to define a new methodology based on a time-frequency windowing approach for improved model reconstruction. This approach is then evaluated in a more complex geological model of far-offset acquisition. Finally, an implementation of structural constraints in the FWI and its application shows the potential benefits of including information from independent geophysical methods in the seismic inverse problem.

4.1 Multiparameter inversion

This section deals with individual and joint 2D elastic full-waveform inversion of Rayleigh and Love waves for spatially correlated and uncorrelated models. First, I present the methodology used for the implementation of the joint inversion in the FWI scheme. After establishing the necessary steps for the joint inversion, I evaluate the performance of both individual and joint inversions and compare the results in terms of the reconstructed models and data fit.

4.1.1 Simultaneous joint inversion

As discussed in Chapter 2, the propagation of the P-SV and the SH waves is described by two independent wave equations, namely equations 2.13 and 2.14. Both the forward as well as the inverse problem of both wave types is decoupled. To invert both independent data sets the application of a joint inversion scheme is therefore required, during which both the P-SV and SH waves are inverted at the same iteration. By increasing the information required by the inversion and also taking into account different sensitivities of the recordings to the different material parameters we should theoretically improve the reconstruction of accurate subsurface models.

To perform the joint inversion, we first identify the material parameters that at any time both inversions are sensitive to. These two parameters for the elastic case correspond to the S-wave velocity and the

density model, respectively, since the SH waves are not sensitive to the P-wave velocity (lack of bulk modulus in equation 2.14). The objective function is adjusted to compensate for this dual minimization problem by merging the two individual data misfits produced by the P-SV and SH wave types. The same adjustments are required for the calculation of the model update, e.g., the gradient of v_s and density with respect to the total/joint objective function.

These adjustments to the FWI workflow are shown in more detail in the next sections, where I introduce the joint approach to combine the objective functions and the gradients of both individual wave type inversions.

4.1.2 Joint objective function of P-SV and SH waves

The objective function introduced in equation 2.76 represents a weighted objective function, where the ratio between the residual energy and the energy of the observed data set is estimated. A ratio of one would indicate that the residual energy is as big as the energy in the observed data set. This definition is helpful also for the joint inversion, since it is not certain that each data set has the same amount of energy, especially in the case of field data applications.

Since both wave types should be weighted equally, assuming that the medium is isotropic, I use a simple addition of both weighted objective functions to calculate the *joint objective function* (Wittkamp et al., 2018):

$$E_{\text{joint}}(\mathbf{m}) = E_{\text{P-SV}}(\mathbf{m}) + E_{\text{SH}}(\mathbf{m}) \quad (4.1)$$

4.1.3 Joint gradients

To combine the gradients of S-wave velocity and density from both the P-SV and SH inversion one can adopt different approaches. This combination is not as simple as the sum of the objective functions, due to the lack of an intuitive normalization between the two gradients. More specifically, the amplitude of the gradient depends on the slope of the objective function, since it is its derivative. Both objective functions $E_{\text{P-SV}}$ and E_{SH} should have a similar slope for the gradients to be of similar amplitude. A normalized addition of both gradients that is followed by a scaling with the sum of the applied normalization factors is being applied (Wittkamp et al., 2018). The maximum absolute gradient amplitude of each case is used as normalization factor. The joint gradient is then given by

$$\delta \hat{\mathbf{g}}^{\text{joint}} = \left(\frac{\delta \hat{\mathbf{g}}^{\text{P-SV}}}{\max(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|)} + \frac{\delta \hat{\mathbf{g}}^{\text{SH}}}{\max(|\delta \hat{\mathbf{g}}^{\text{SH}}|)} \right) \cdot \kappa, \quad (4.2)$$

where $\delta \hat{\mathbf{g}} = \left\{ \frac{\partial E}{\partial \rho} \text{ or } \frac{\partial E}{\partial v_s} \right\}$ and $\kappa = \max(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|) + \max(|\delta \hat{\mathbf{g}}^{\text{SH}}|)$.

Although I preserve the amplitude information by rescaling the joint gradient, it is not a necessary requirement. One could easily include a weighting factor that favors one over the other data sets, e.g., when one of the two data sets suffers more from poor signal-to-noise ratio (SNR).

4.1.4 Numerical setup and model description

In sections 4.1.5, 4.1.6 and 4.1.7, I evaluate the performance of the individual and joint inversions for Love and Rayleigh waves, while focusing on the understanding of the sensitivity of the material parameters in the different wave types. The inversion software I use and maintain is the code IFOS2D (Inversion of Full Observed Seismograms), which was initially developed by Köhn (2011). All the numerical and optimization aspects of the inversion code were explained in Chapter 2.

The acquisition geometry of all following studies is the same and consists of a linear profile of 39 three component geophones with an equidistant spacing of 1 m. I use the same source signal as in Chapter 3, given by equation 3.9, with the dominant frequency set to 30 Hz, which generates frequencies up to approximately 70 Hz (Figure 3.4). I use a vertically and horizontally polarized source for the P-SV and SH case, respectively. Receivers and sources are located at the surface, since I focus on non-invasive seismic measurements also in the field data cases that I will examine later on in Chapters 6 and 7. The model consists of 336 grid points in the horizontal direction and 120 grid points in the vertical direction with grid spacing of 0.15 m, resulting in a model of 50 m length in the horizontal and of 18 m depth in the vertical direction. At the top of the model I apply the free surface condition and at the bottom as well as on the lateral boundaries I use the CPML boundaries. The model contains a semicircular low-velocity/low-density inclusion at the surface similar to Chapter 3, as well as an interface at approximately 8 m depth. This interface represents the groundwater table, where all elastic parameters have a sharp discontinuity. The elastic parameters represent the geological situation found on a glider airfield in Rheinstetten in Western Germany which has been used in previous studies (Groos et al., 2017; Wittkamp et al., 2018; Athanasopoulos et al., 2020) and which I will further examine in Chapter 6.

In the P-SV case, I calculate the objective function for the vertical and horizontal inline displacement component and in the SH case for the horizontal crossline displacement component. The displacement seismograms are then obtained from the recorded velocity seismograms by numerical integration, however in all figures containing seismograms I show particle velocities. I choose a temporal sampling of $4.5 \cdot 10^{-5}$ s and the total recording time is 1 s.

Additional tapers are applied which are necessary for the stability of the inversion. Semicircular source tapers with a length of 3 grid points are used to precondition the shot-wise gradients. Basically, the source taper decays within 3 grid points from a value of one to zero at the location of the source. The approximation of the diagonal elements of the Hessian which is described by equation 2.72 is used, by additionally setting ϵ , empirically, to $5 \cdot 10^{-3}$ for the SH waves and to $5 \cdot 10^{-2}$ for the P-SV waves. It is necessary to further smooth the gradients using a 2D median filter, which has a size of 5 grid points in both spatial directions.

The inversion workflow showing the multiscale approach is presented in Table 4.1. The application of a Butterworth band-pass filter to the data is the last step, where the upper corner frequency of the filter (4th order) changes between stages and the lower corner frequency remains fixed at 5 Hz. I enforce the algorithm to perform a minimum of 4 iterations for each stage. However, in the case that the step length estimation fails (no step length that reduces the misfit function can be found given the number of trials that are evaluated), the algorithm directly switches to the next stage.

Table 4.1: Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter.

Stage	Update			LP in Hz
	v_P	v_S	ρ	
1	1	1	0	10
2	1	1	1	10
3	1	1	1	20
4	1	1	1	30
5	1	1	1	40
6	1	1	1	50
7	1	1	1	60

Note that the inversion at the very first stage ignores any update of the density model. I found that the gradient of the density and to some extent the updates of the density model become extremely unstable during early iterations, where the initial models for v_p and v_s are far from the true model.

In the three different studies shown in the next sections, both the interface and the values and locations of the inclusion will vary for different purposes, which I will refer to in more details in each of the following sections. The numerical setup is, however, the same and any deviations will be clearly mentioned. I assume the source wavelet to be known and that the medium is elastic, as otherwise this would add unnecessary complexity to the synthetic studies and it would not serve my purpose on assessing the sensitivity of SH and P-SV waves in the P- and S-wave velocity and density reconstruction.

4.1.5 Spatially correlated models

In this first study, I examine the inversion of the three elastic parameters using a very simple layout where the true models are all spatially correlated. The true models, shown in the first column of Figure 4.1, include an upper homogeneous layer (in this case for all three parameters) and the discontinuity as mentioned at a depth of 8 m. The size of the trench is 8.3 by 2.9 m in horizontal and vertical direction, respectively. The starting models consist of the true background model since, as mentioned above, I am interested in observing the stability of the inversion scheme and getting a first impression of the sensitivity of the data to each of the three models.

In the following, I will discuss the results of the individual P-SV and SH inversions along with the joint inversion for each of the model parameter separately. I will keep this format for the other two case studies as well.

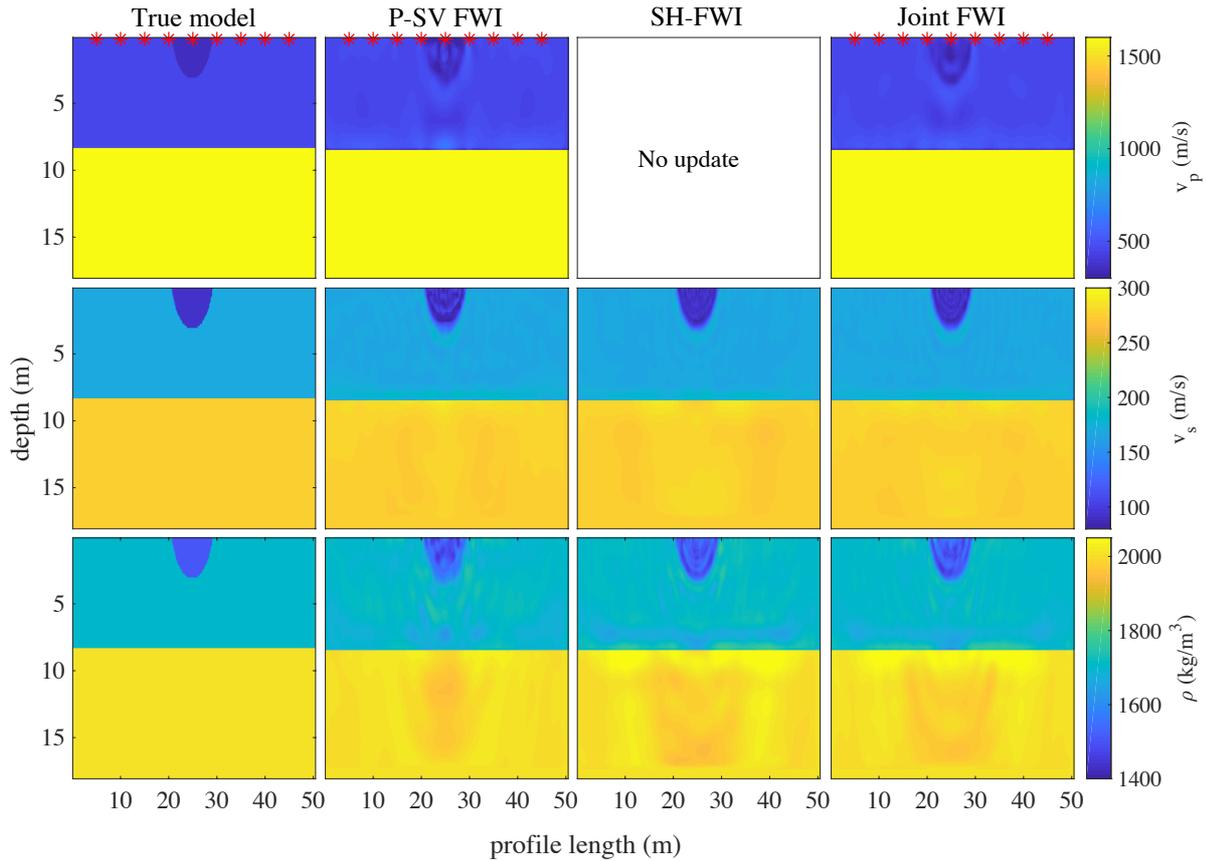


Figure 4.1: Models obtained by the inversion of synthetic data for a simple spatially correlated near-surface model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the final results of the Rayleigh wave (P-SV), Love wave (SH) and joint FWI. Red asterisks represent the source locations.

Results: S-wave velocity model

In near-surface seismic applications, as I already discussed, the surface waves which include most of the information in the recordings are most sensitive to the S-wave velocity. This was shown also in Chapter 3 where the highest amplitude of scattered waves originated from perturbations in the S-wave velocity model. Therefore, we already know since the data are dominated by surface waves, that we will be able to reconstruct the S-wave velocity accurately. In Figure 4.1 (second row) we observe that both the individual and the joint inversion reconstructed the trench quite accurately. Both the absolute values and the shape of the anomaly are close to the true model, especially in the case of the SH FWI and joint FWI (JFWI). In the P-SV case the shape deviates at the bottom of the trench and also the values are not as accurate. However, the overall reconstruction for such a simple case is very good.

Results: P-wave velocity model

For the P-wave velocity reconstruction I use the P-SV and JFWI (Figure 4.1, third row), as mentioned earlier. The anomaly in this simple case was reconstructed to a satisfying level; however, it is visible that although both techniques did not perfectly reconstruct the true model, the JFWI outperformed the P-SV case. Both the shape and values of the trench are closer to the true values. Additionally, in the case of P-SV FWI several artifacts arise inside the volume of the anomaly, something that we noticed also in

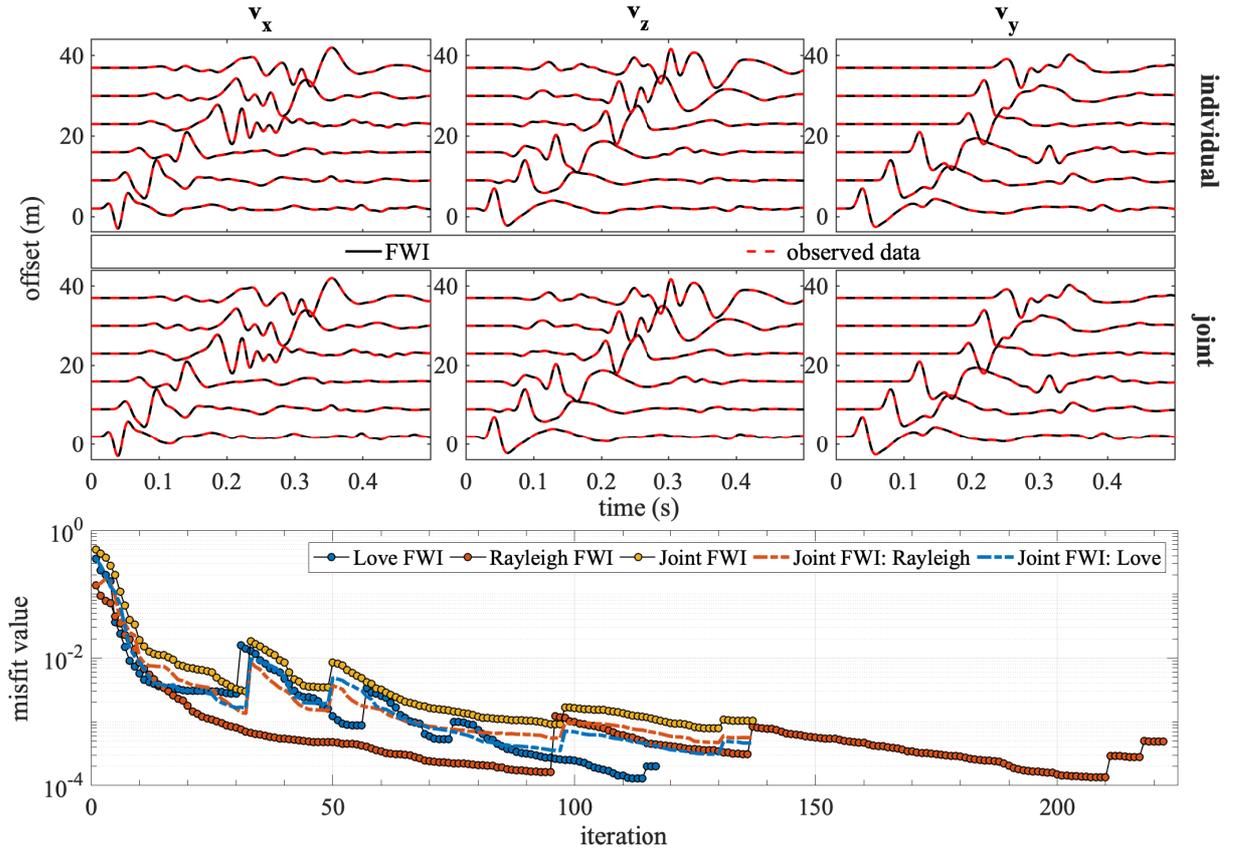


Figure 4.2: Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI), Love wave (SH FWI) and joint FWI shown in Figure 4.1. Top: Comparison of velocity seismograms for the horizontal inline (v_x), vertical (v_z) and crossline (v_y) components for 6 different offsets. Bottom: Misfit function for each of the inversions (filled circles). The dashed red and blue lines constitute the contribution of Rayleigh and Love waves, respectively, to the joint misfit function (yellow circles). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 4.1).

Wittkamp et al. (2018) and Athanasopoulos et al. (2020). Some more prominent artifacts appeared in the upper layer just below the anomaly at depths of around 6 m. The improvement of the v_p model in the case of the joint FWI is a result of the joint regularization and improvements of the v_s and the density model, since no further action is performed in the JFWI compared to the P-SV case. The inaccuracies in the v_p model have a direct impact on the reconstruction of the v_s model (Groos, 2013; Athanasopoulos and Bohlen, 2016).

Results: density model

The density, as I showed in the previous chapter, is a *weak* parameter in the process of FWI. Density perturbations mostly affect the amplitudes of the reflected waves rather than their phases as I show in Chapter 3. The reconstructed models have the lowest resolution among the three elastic parameters (Figure 4.1, third row). It is important to note that delaying the inversion of density is crucial, since otherwise the early updates of the density gradient become unstable. This arises from the inaccuracies of the velocity models at the early stages of the FWI workflow. A big improvement once again is obvious when comparing the P-SV case with the other two approaches. This shows that the Love waves, being

insensitive to the P-wave structure of the medium, can help reconstruct accurately the v_s and the density model, without suffering from cross-talk due to inaccuracies in the P-wave velocity model.

Results: data-fit comparison

In Figure 4.2, I show the seismograms for the different components of particle velocity. I show the observed data and the data inverted from FWI for six different offsets at the last frequency step (60 Hz). From the data fit we do not spot any significant residual energy in any of the components and wave types I used. To better quantify the results, I also plot the misfit functions for each of the inversions that I performed.

As we can see, the SH-FWI result has the lowest misfit, showing the accuracy of the inversion scheme, something which is in agreement with the reconstructed models for v_s and density. We have to consider that in this case the amount of data is less since for SH-waves we can only use the crossline components (v_y). The P-SV FWI resulted in the greatest number of iterations, while having in mind that the same abort criteria were used in all cases. The misfit was lower than the joint inversion from the Rayleigh waves but higher than the joint inversion from the Love waves.

For the joint inversion, the misfit function is the combination of the Rayleigh and Love wave FWI when the gradients are updated simultaneously, as mentioned at the beginning of this section. We observe an increase in the actual misfit values, which contradicts at first glance with the high resolution of the reconstructed models from the JFWI. However, this is actually not the case since a joint inversion has to compromise in terms of the individual components that it tries to minimize when more than one condition have to be fulfilled. The inversion converged in much less iterations compared to P-SV FWI and the error was of the same magnitude as the other two inversion schemes.

4.1.6 Spatially uncorrelated models

In this section, I examine the application of FWI in a more realistic subsurface scenario. Both the a-prior information regarding the initial model is not as strong compared to the previous section, and the models have a higher complexity. The true model is presented in Figures 4.3 and 4.4 (first column). In the upper layer v_s gradually increases with depth. I added a low-velocity horizontal layer in v_p in order to investigate the reconstruction of low-velocity layers by FWI, since this would be a challenging task for basic traveltimes tomography which is commonly used for refraction seismics. In order to reveal trade-off effects between the elastic parameters, I shift the horizontal location of the anomaly in each parameter. Stochastic fluctuations were added to make the model geologically more complex (Manukyan et al., 2012).

The initial models consist of a 1D structure with a smooth gradient of increasing velocity/density over depth up to the groundwater table which is shifted by 1 m compared to the true models (Figure 4.3, red solid lines). I choose to misplace the discontinuity of the groundwater table since, in the case of field data (Section 6), I do not have an exact depth of the bedrock and therefore this uncertainty can be included in the synthetic case. An accurate depth of such structures would greatly assist FWI to reconstruct the

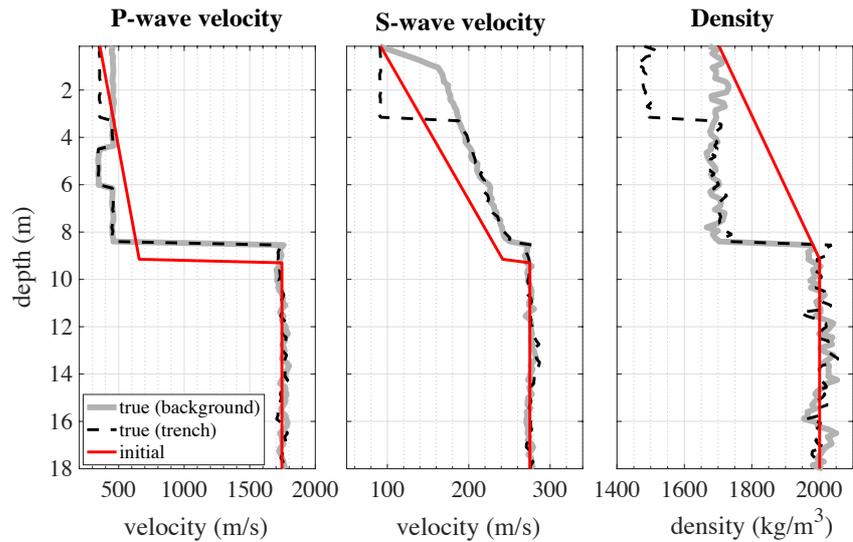


Figure 4.3: 1D depth profiles of the elastic parameters v_p , v_s and ρ for the true and the initial model.

elastic parameters more accurately (as shown in the previous section), which should in general be avoided when testing methodologies in a synthetic configuration.

Once again, I will discuss the results of the individual P-SV and SH FWI along with the joint inversion for each of the model parameter separately.

Results: S-wave velocity model

Figure 4.4 (second row) shows the reconstructed v_s model for the individual and joint inversions. Similar to the results of the spatially correlated study, the inversion retrieved the v_s of the medium very accurately. However, there is a visible improvement when using the SH and JFWI over the P-SV FWI. More specifically, in the P-SV FWI results, the v_s model is well resolved within the upper 8 m, with the exception of some artefacts around the discontinuity of the groundwater table. These can be related to some trade-off from inaccuracies in the v_p model which contaminate the reconstruction of the v_s model (Groos et al., 2014; Wittkamp et al., 2018). The interface separating the upper half-space and the groundwater table suffers from artefacts. The trench in all cases is accurately reconstructed in both shape and values (Figures 4.4 and 4.5), with some artifacts in the P-SV case inside the anomaly though. The Love wave FWI does not show any artifacts similar to the JFWI. Additionally, although the reconstruction of the groundwater discontinuity is not accurate, there is a clear increase in the v_s values directly above it, which is more consistent in these two cases over the P-SV FWI.

Results: P-wave velocity model

The reconstruction of the v_p model is not accurate, particularly in the P-SV FWI case (Figure 4.5, first column), for two main reasons: 1) the low sensitivity of the high-amplitude Rayleigh waves to v_p , and 2) the weak contribution of the low-amplitude P-waves in the misfit function. Although in the P-SV FWI there is a decrease of the velocity in the area where the low velocity horizontal layer is present, the shape and structure do not correlate with the true values. Additionally, the trench in the v_p model is not retrieved

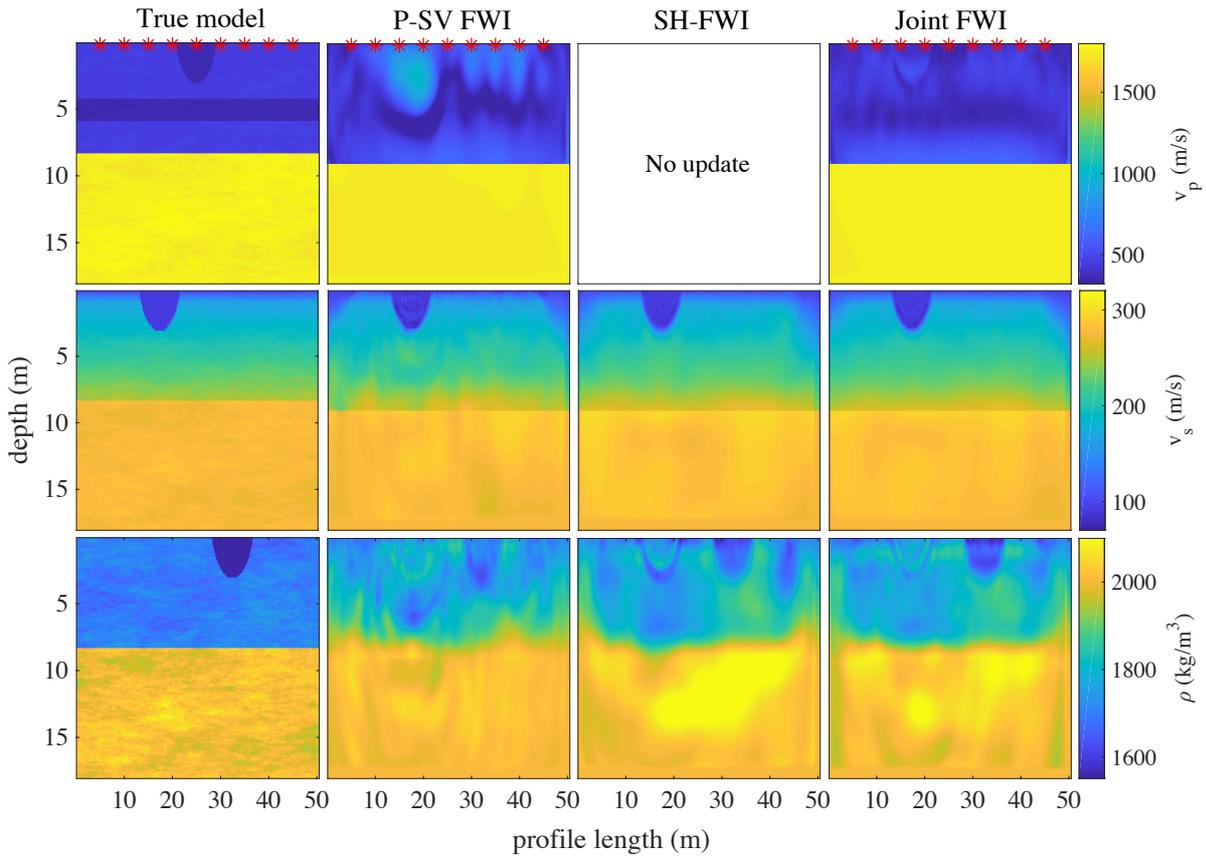


Figure 4.4: Models obtained by the inversion of synthetic data for a complex spatially uncorrelated near-surface model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the final results of the Rayleigh wave (P-SV), Love wave (SH) and joint FWI. Red asterisks represent the source locations.

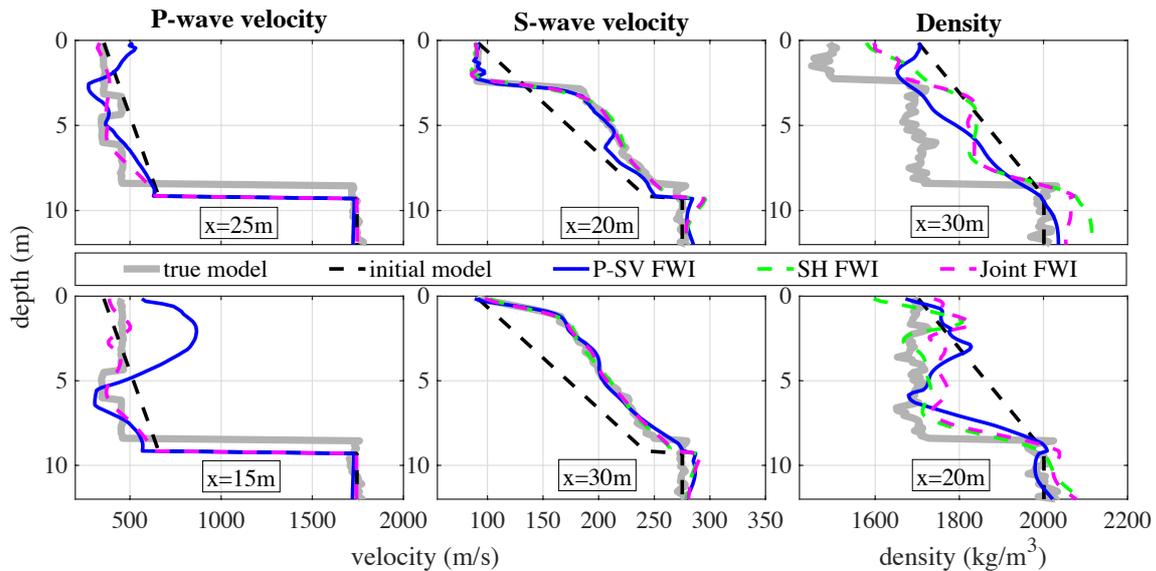


Figure 4.5: 1D models of the true, initial and reconstructed parameters shown in Figure 4.4. The first row shows vertical profiles of the true, initial and reconstructed models at the location of the shallow anomaly of each parameter. The second row shows vertical profiles through the background medium.

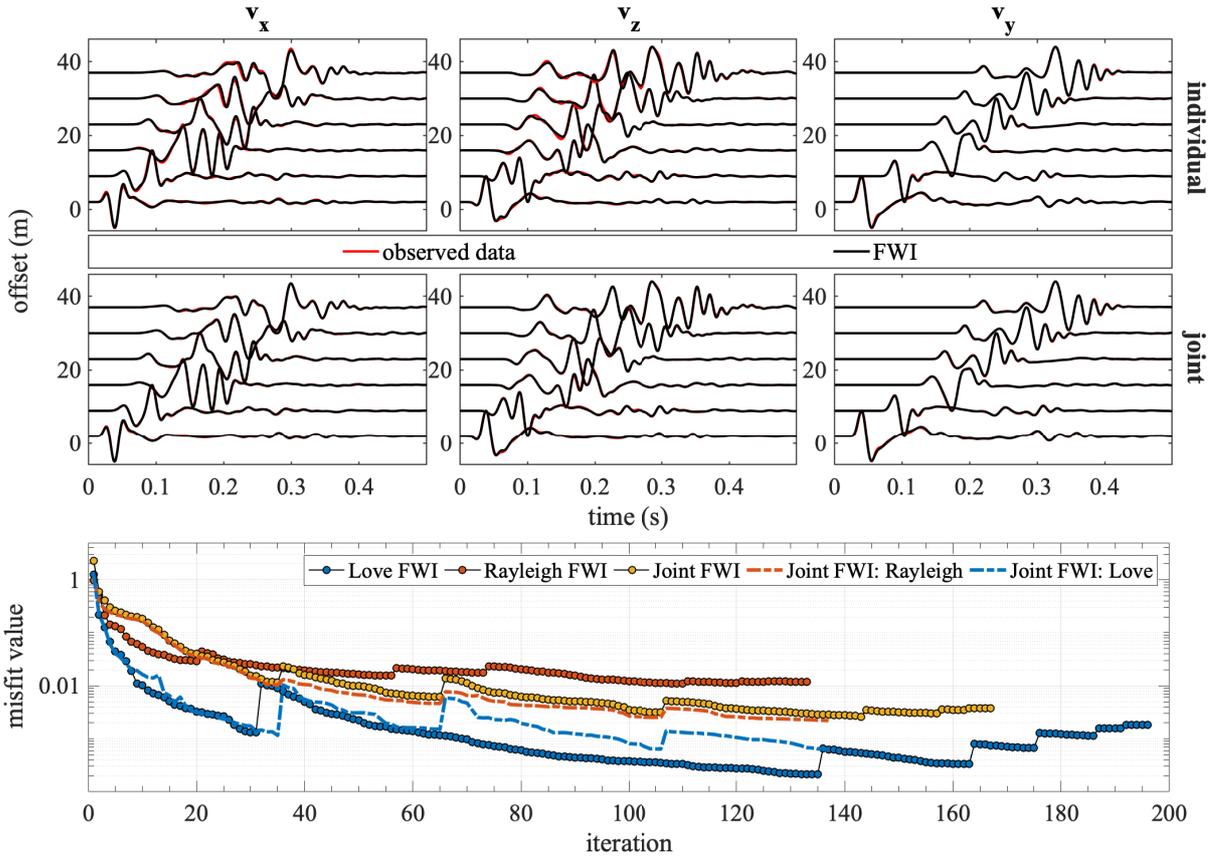


Figure 4.6: Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI), Love wave (SH FWI) and joint FWI shown in Figure 4.5. Top: Comparison of velocity seismograms for the horizontal in-line (v_x), vertical (v_z) and cross-line (v_y) components for 6 different offsets. Bottom: Misfit function for each of the inversions (filled circles). The dashed red and blue lines constitute the contribution of Rayleigh and Love waves, respectively, to the joint misfit function (yellow circles). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 4.1).

and there are artifacts at the locations of the sources. We observe an increase of v_p at the location of the trench in the v_s model due to cross-talk between v_s and v_p (Figure 4.5). The inaccuracies in the v_p model impacted also the reconstruction of the v_s model, which is clear from the footprint directly below the trench (Figure 4.4, v_s from P-SV FWI).

The joint FWI provided a better v_p model which once again appears to be the result of the joint regularization and improvements of the v_s and the density model. Both the low velocity layer and the trench are visible, however, they are a smooth version of the actual values and the interfaces are not precisely retrieved (Figure 4.5). The footprint from the trench in v_s is still present here but much less prominent. The overall v_p reconstruction shows that improvements are required in order to get a high resolution v_p model and this is the main focus of Section 4.2 and the application of FWI in the field data case of Chapter 6.

Results: Density model

The density model is poorly reconstructed and suffers from artifacts and cross-talk effects in all three cases (Figure 4.4, third row and Figure 4.5, third column). Although the location of the trench is accurate,

both its shape and absolute values are far from the true model. Contamination from the velocity models produces artifacts over the whole model space. Below the groundwater discontinuity there are high and low density values that do not correspond to the true values, and surprisingly this was most prominent when Love waves were considered. However, regarding the trench, there is a clear improvement in the resolution of the density model when Love waves are included in the inversion process, both in the case of SH and joint FWI.

An important difference compared to the spatially correlated study is that here I used a trace-wise normalization prior to the estimation of the objective function, for consistency reasons with field data applications. As mentioned before, the density mainly influences the absolute wave amplitude as a function of offset. Therefore, we will not be able to get the most out of FWI when using trace-wise normalization, but it is an important step to minimize the influence of a low signal-to-noise ratio which is always present in real recordings.

Results: data-fit comparison

In Figure 4.6, I compare the seismograms for six different offsets at the last frequency step. I increased the low-pass filter to 70 Hz compared to the previous section, due to the higher frequency content of the data, introduced by the modified model. Both the v_x and v_z (P-SV waves) components have a larger mismatch than the v_y (SH-waves) component. In the P-SV FWI, the mismatch is larger especially in the areas between the refracted waves (first arrivals) and the dominant mode of the Rayleigh wave, which is expected since v_p was not accurately retrieved. The amplitudes of the Rayleigh wave show also some differences to the true data. The v_y on the other hand shows that the SH-FWI managed to fit the observed and synthetic data much better. The joint FWI shows the smallest mismatch for all three components.

We can observe the improved fit of JFWI in Figure 4.6 where the misfit functions of each approach are shown. The misfit function of both the joint Rayleigh and Love wave FWI (blue and red dotted lines, respectively) are much lower compared to the respective individual inversions (blue and red circles, respectively). This shows that in case of more complex media and weak a-prior information the joint inversion outperforms the individual ones both with respect to the model resolution and data fit.

4.1.7 Influence of the initial P-wave velocity model

In this case study, I keep the model values of the spatially uncorrelated test of the previous section. However, the models are now spatially correlated and without the inclusion of the stochastic fluctuations. Furthermore, I removed the low velocity layer in v_p . The goal is to study the influence of the initial v_p in the FWI results.

I create two scenarios: in **case A** the initial v_p model is significantly different to the true model (Figure 4.7, 4th column), whereas in **case B** the initial v_p model corresponds to the true background P-wave velocity and accurate depth of the discontinuity (Figure 4.7, 2nd column). The same workflow is applied once again and I apply the trace-wise normalization of synthetic and observed data, as in the previous section.

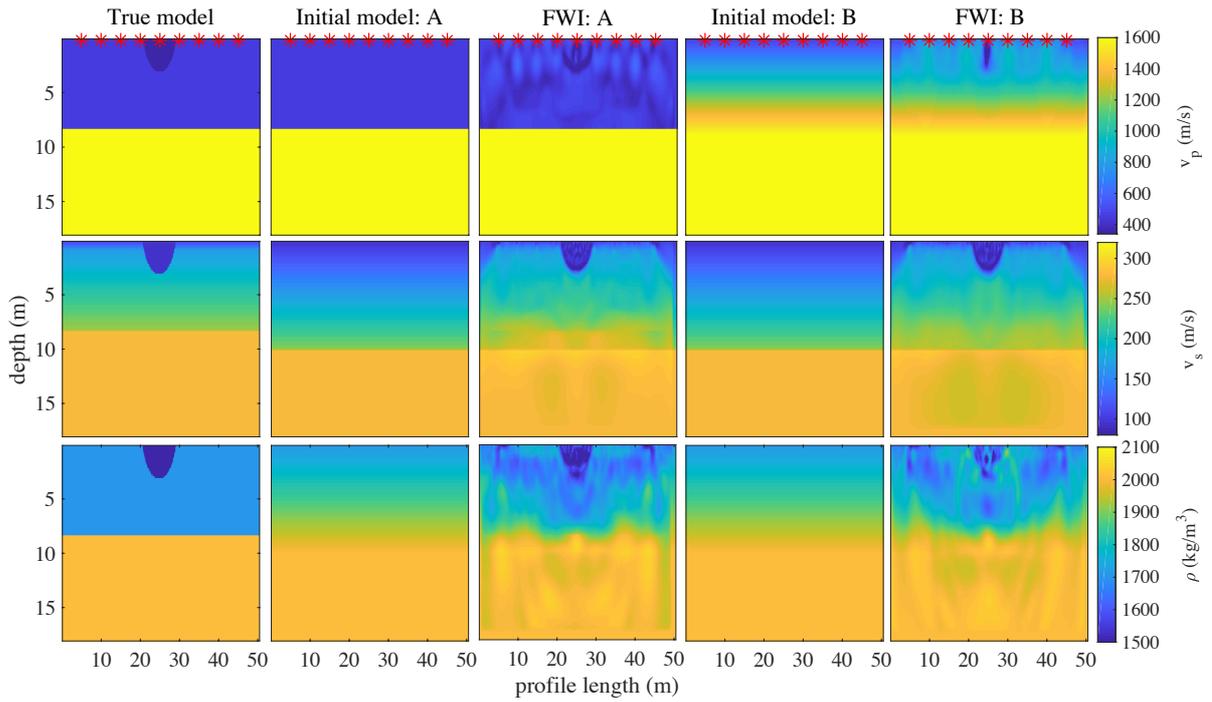


Figure 4.7: Models obtained by the inversion of synthetic data for a spatially correlated model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the initial and final results of the Rayleigh wave (P-SV) FWI using two different initial locations, namely A and B. Red asterisks represent the source locations.

Results

Case A: In this case, the elastic parameters are resolved to a much better extent compared to case B. The v_p model still suffers from some artifacts originating from the source locations, however the trench's shape and values are close to the true. The v_s model is still the parameter with the highest resolution, and the accurate background v_p assisted the inversion to correct for the true depth of the groundwater discontinuity, although the values still deviate to some extent from the true ones. The density, despite still being the weakest resolved parameter, was retrieved to a satisfying level. Especially the values and shape of the trench are quite accurately reconstructed, as well as the background values of the upper layer, which are closer to the true values compared to the case B. The high resolution in the model-space is in agreement with the high fit between the observed and inverted data, as shown by Figure 4.8, along with a one order of magnitude smaller misfit overall.

Case B: The issues which were discussed in previous sections become more prominent in the case of an inaccurate initial v_p model. There are barely any updates in v_p , with a small exception, however inaccurate, at the location of the trench. The FWI completely failed to account for such big differences and it was trapped in local minima. The poor reconstruction of the v_p model can be seen in Figure 4.8, as the P-wave onsets are completely mismatched to the actual observed data. The v_s model is quite accurate despite the inaccuracies of the v_p model, which in some cases cause also poor reconstruction of v_s (Athanasopoulos and Bohlen, 2016, 2017b; Wittkamp et al., 2018). However, the absolute values of the background model show deviations from the true model, both below and above the groundwater discontinuity. The correct depth of the discontinuity has not been retrieved as was the case in A. Finally, the density model was

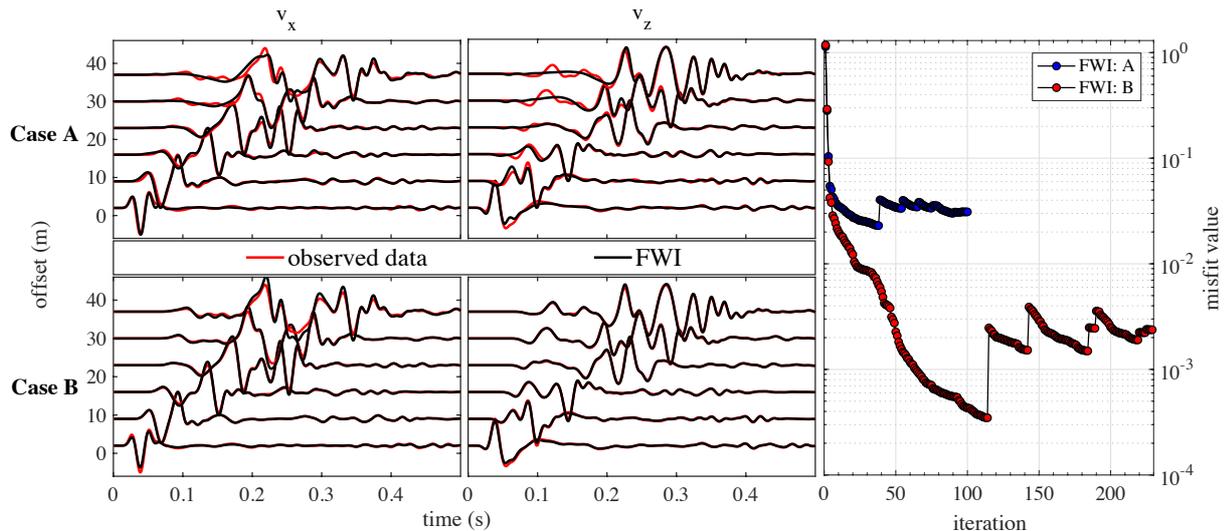


Figure 4.8: Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI) for the cases A and B shown in Figure 4.7. Comparison of velocity seismograms for the horizontal in-line (v_x) and vertical (v_z) components for 6 different offsets. Misfit function for the two different cases is shown in the third column.

also not accurately reconstructed, although there were some updates at the location of the trench, however inaccurate regarding the absolute values.

4.1.8 Conclusions

In a series of synthetic studies, I was able to determine the limitations of Rayleigh, Love wave and joint FWI. In all cases, the JFWI outperformed the P-SV FWI and in some cases also the SH-FWI. In the absence of an accurate initial model for the P-wave velocity, the Love wave FWI should be used to obtain the v_s model since (1) its convergence behavior is independent of the P-wave velocity, (2) it holds a smaller parameter space, which leads to less trade-off effects and (3) the SH wave equation is less complex than the P-SV wave equation, which allows a computationally efficient inversion. When however an accurate P-wave velocity can be retrieved, using the joint FWI has several advantages compared to both individual wave type inversions: (1) Decrease of the ambiguities of the inversion result, since more data are considered, (2) reduction of the parameter cross-talk and (3) further improvement on the resolution and accuracy of the inversion result.

4.2 Time - frequency windowing

In the past few years, several approaches have been deployed to overcome the different contribution of surface and P-waves in respect to both the data and the model domain. Plessix and Mulder (2004b) suggested an approximation to the diagonal elements of the Hessian by stacking the absolute values of the off-diagonal terms of a given line of the matrix with the diagonal term. Nuber et al. (2015) have demonstrated that by appropriate scaling of the approximate Hessian matrix they could enhance weak sensitivities relating to model updates that were disregarded. Gibson et al. (1979) used a layer stripping approach at which the model updates are obtained layer by layer from shallow to greater depths of the

model domain allowing a more accurate shallow reconstruction that would later benefit the inversion of the deeper parts of the subsurface. Brenders and Pratt (2007) made use of a finite-length offset-dependent time-damping function in the observed and synthetic data based on acoustic FWI in the frequency domain. Their goal was to improve the contributions of the earlier arrivals and relax the non-linearity of the inverse problem. In this section, I focus on the improvement of the reconstruction of the P-wave velocity model, by rebalancing the different wave types during FWI.

4.2.1 Theory and methodology

In order to separate the contribution of refracted P- and surface waves, I propose a time-frequency windowing approach (TF-FWI). I choose an adaptive time-window in combination with the multi-frequency (multi-scale) method so that I can better match the low-amplitude refracted waves. To achieve this I taper both the synthetic \mathbf{d}_{syn} and observed data \mathbf{d}_{obs} (Brossier et al., 2009; Heider, 2014; Choi and Alkhalifah, 2018; Athanasopoulos et al., 2020) with

$$\mathbf{d}_{damped} = e^{-\gamma t} \mathbf{d}, \quad (4.3)$$

where γ is the damping factor and t the recording time. The modified data residuals, $\Delta \mathbf{d}_{tf}$, between the synthetic data $\mathbf{d}_{syn}(\mathbf{m})$ and the observed data \mathbf{d}_{obs} then become

$$\Delta \mathbf{d}_{tf} = \mathbf{d}_{damped}^{syn}(\mathbf{m}) - \mathbf{d}_{damped}^{obs}. \quad (4.4)$$

The time-damping starts from zero time at all offsets in order to maintain consistency between the observed and synthetic data in terms of their kinematics, i.e., to have a higher flexibility of modeled arrival times which do not need to fall in a predefined time window. After time-windowing I apply the conventional multi-frequency approach (Bunks et al., 1995), i.e., I subdivide the inversion in stages of increasing frequency bands, and the time-windowing described by equation 4.3. I gradually increase the upper corner frequency of a Butterworth band-pass filter and thereby increase the bandwidth of the data. Additionally, I apply the time-damping as described by equation 4.3 but following a specific strategy for choosing the damping parameter γ to isolate the targeted wave events that we are interesting in.

The time-frequency approach is illustrated in Figure 4.9. Figure 4.9a shows a reference synthetic seismogram which corresponds to trace number 39 of Section 4.2.5 (red trace, Figure 4.11). The phases corresponding to specific wave types are highlighted. We can identify: a) the refracted wave as the earliest arrivals as phase 1, b) a superposition of S-wave, forward and scattered converted waves and higher Rayleigh waves modes as phase 2, c) the fundamental mode of Rayleigh waves as phase 3, and d) backscattered Rayleigh waves as phase 4. Each of these signals has a different signal amplitude, sensitivity towards subsurface structures.

Figure 4.9b shows the effect on the reference trace of the band-pass filtering (blue lines), along with the addition of the time-damping (red lines) for increasing bandwidth and decreasing damping parameter γ .

In the first phase, I apply a strong damping to taper any amplitudes after the first arrivals and invert only for the P-wave velocity model. The subsurface information is then gradually increased by relaxing the damping factor and jointly invert for P-wave velocity (v_p), S-wave velocity (v_s) and density (ρ) models,

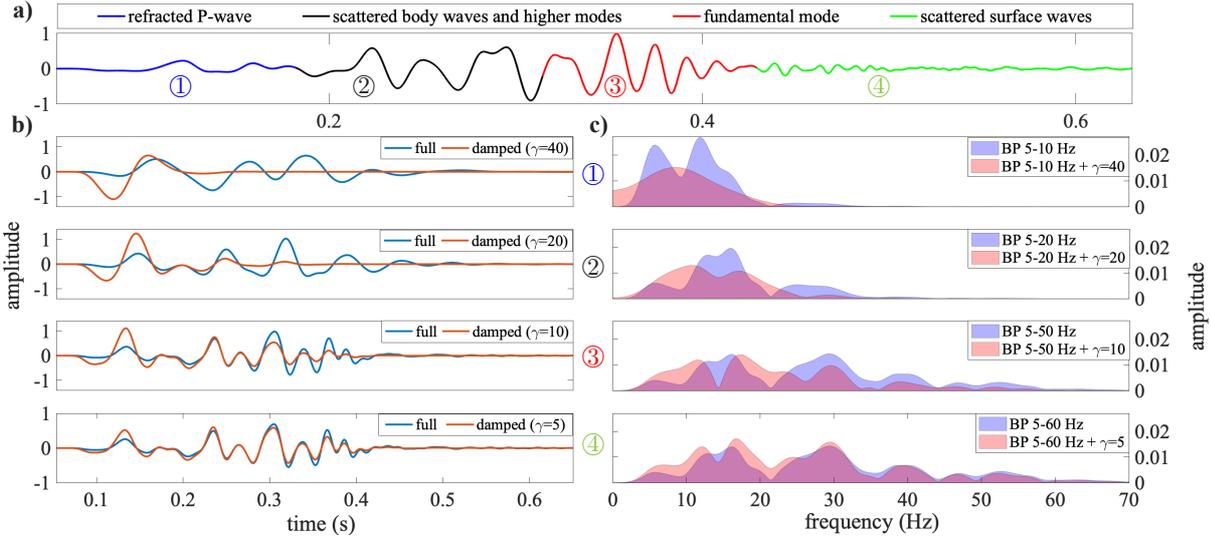


Figure 4.9: The effects of time-frequency windowing on a synthetic seismogram in the time and frequency domain. a) Synthetic reference seismic traces. The arrivals of different wave types are highlighted. b) Seismogram after application of pure low-pass filters (LP) and a combination of low-pass filter and time windowing (LP+TW). c) Corresponding amplitude spectra.

moving from phase 1 to 4 (every time the previous phases are included in the next one), while reducing the cycle-skipping problems by reducing the complexity of the waveform. This can also be seen in Figure 4.9c where the frequency content of each trace from Figure 4.9b is displayed. The application of the time window (equation 4.3) reduces the complexity of the signals and thereby the non-linearity of the inverse problem. In the frequency domain the exponential time windowing corresponds to smoothing the amplitude spectra. This can be observed in the amplitude spectra shown in Figure 4.10.

4.2.2 Effect of the taper function on frequency spectra

I evaluate the effect of the exponential damping factor in the seismic signal in terms of changes of the frequency content. The exponential function is defined as

$$g(t) = \begin{cases} e^{-\gamma t}, & t \geq 0 \\ 0 & t < 0 \end{cases}, \quad (4.5)$$

where γ is the damping factor.

During the TF-FWI approach, the function $g(t)$ is multiplied with each trace in the time domain in order to obtain the damped wavefield used for the calculation of the residuals and by extent of the adjoint source, for various values of γ . The multiplication in time leads to a convolution in the frequency domain. The frequency domain representation of equation 4.5 is given by:

$$G(\omega) = \int_{-\infty}^{+\infty} g(t)e^{-j\omega t} dt = \int_0^{+\infty} e^{-\gamma t} e^{-j\omega t} dt = \frac{1}{\gamma + j\omega} \quad (4.6)$$

Both functions are shown in Figure 4.10. $G(\omega)$ acts as a smoothing operator in the frequency domain. The multiplication of the seismic wavefield with the function $g(t)$ is similar to the application of a smoothing

operator in the frequency domain. The length of the smoothing operator (\mathcal{L}), i.e., the width of $|G(w)|$ at half maximum, has a linear relationship with γ given by

$$\mathcal{L} \approx 0.27\gamma, \quad (4.7)$$

We observe that the smaller the values of γ the more the frequency spectrum of the exponential function resembles a spike leading after convolution to barely any smoothing of the frequency spectrum of the resulting signal. On the other had, a strong damping factor leads to a wider spectrum which after the convolution results in a smooth version of the original frequency spectrum.

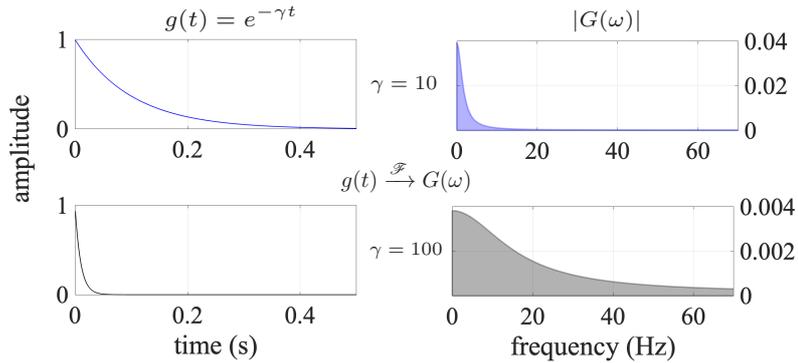


Figure 4.10: Taper function in the time and frequency domain for damping factor $\gamma = 10$ and 100 , respectively.

4.2.3 Effect of the taper function with offset

An additional step that is required in my approach contains the trace-wise normalization of both observed and synthetic shot gathers, since due to the exponential damping that I apply on the time series, the amplitude of the seismic traces decays with increasing offset. The compensation of the energy content of the far offsets due to normalization is illustrated in Figure 4.11, where we observe that in all stages a proper representation of the corresponding arrivals of the respective phase is taken into account. In field data applications, the first arrivals need to be picked so that we can neglect any energy in the data prior to the refracted waves, otherwise the exponential taper will also increase the noise level.

One main advantage of the time-windowed approach where the damping is applied on the whole data set (TF-FWI) is that I increase the contribution of P waves versus the Rayleigh waves but without completely neglecting the latter as was done in Athanasopoulos and Bohlen (2016, 2017b). This is essential due to the fact that the Rayleigh waves contribute also to the P-wave velocity model (as shown in Figure 3.5), although not as significantly as refracted waves. Another advantage is the reduction of the cycle-skipping, caused by a poor initial model, since the time window spans the whole simulation time, while in Athanasopoulos and Bohlen (2016, 2017b) the time window was very short (small time interval around the first arrivals) and therefore in the presence of severe kinematic mismatch the FWI would not converge towards the true solution. Finally, my approach does not contribute to an increase of the computational cost of FWI, since as shown by equation 4.4 the only modification is due to the multiplication with the exponential damping function. Therefore, the iteration number will be a direct indicative of the computational cost required for each method.

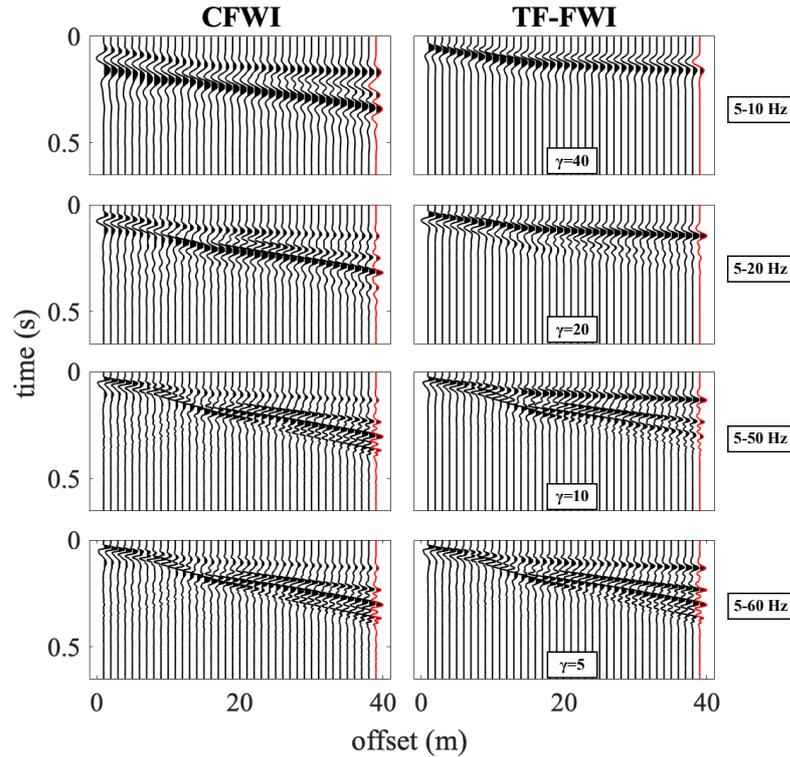


Figure 4.11: Data used in conventional FWI (CFWI, only multiscale approach) versus the data used in TF-FWI (time-frequency approach) for different inversion stages. The application of the trace-wise normalization allows to account for the offset-dependent energy-loss when applying a taper at time equal to 0.

4.2.4 Synthetic reconstruction test

In a synthetic study I compare the results of conventional FWI (CFWI) with TF-FWI to study the performance of the time-frequency approach. I use the spatially uncorrelated model from the previous section for a direct comparison. The effect of both the multiscale and the time-window approach on the first shot gather (observed data) is shown in Figure 4.11 for a few representative stages. The trace-wise normalization, as mentioned in Section 4.2, helps us to retain the amplitudes of the far-offset traces when starting the time-window at zero time. The general FWI setup in both cases is described in Section 2.2 and it is kept the same as before for consistency. Furthermore, I still use semicircular source tapers with a length of 3 grid points to precondition the shot-wise gradients and a smoothing using a 2D median filter, which has a size of 5 grid points in both spatial directions. I apply a Butterworth band-pass filter to the data, where the upper corner frequency of the filter (4th order) changes between stages and the lower corner frequency remains fixed at 5 Hz (Table 4.2).

4.2.5 Results

In Figures 4.12 and 4.13 I show a direct comparison of the final models for both CFWI and TF-FWI, along with the true and initial models. Figure 4.14 shows the comparison in data space where I compare the seismograms produced by the two methods and the true model.

Table 4.2: Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter. The column ' γ ' indicates the damping factor (equation 4.3). The first stage is only applied in TF-FWI.

Stage	Update			LP in Hz	γ
	v_P	v_S	ρ		
1	1	0	0	10	40
2	1	1	0	10	30
3	1	1	0	20	20
4	1	1	0	30	20
5	1	1	1	40	15
6	1	1	1	50	10
7	1	1	1	60	5
8	1	1	1	70	0

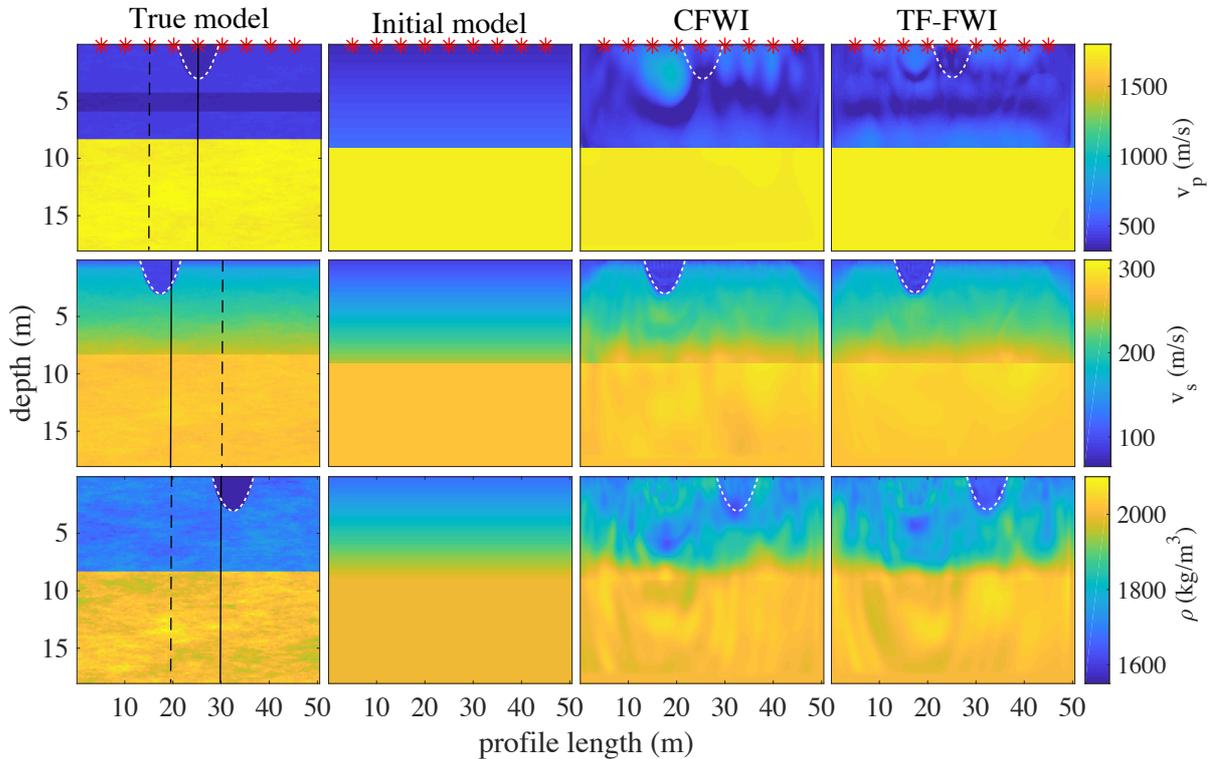


Figure 4.12: Models obtained by the inversion of synthetic data. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional FWI (CFWI) and time-frequency FWI (TF-FWI). Red asterisks represent the source locations. The true boundaries of the trench are highlighted. The black dashed and solid lines correspond to the 1D profiles in Figure 4.13.

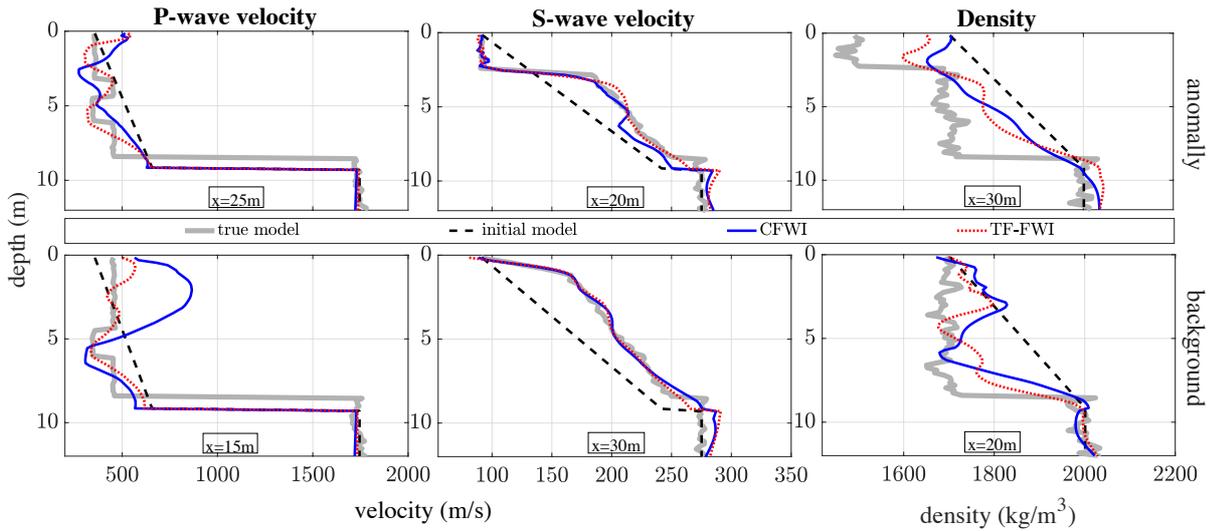


Figure 4.13: 1D models of the true, initial and reconstructed parameters shown in Figure 4.12. The first row shows vertical profiles of the true, initial and reconstructed models at the location of the shallow anomaly (solid vertical lines in Figure 4.12, first column). The second row shows vertical profiles through the background medium (dashed vertical lines in Figure 4.12, first column).

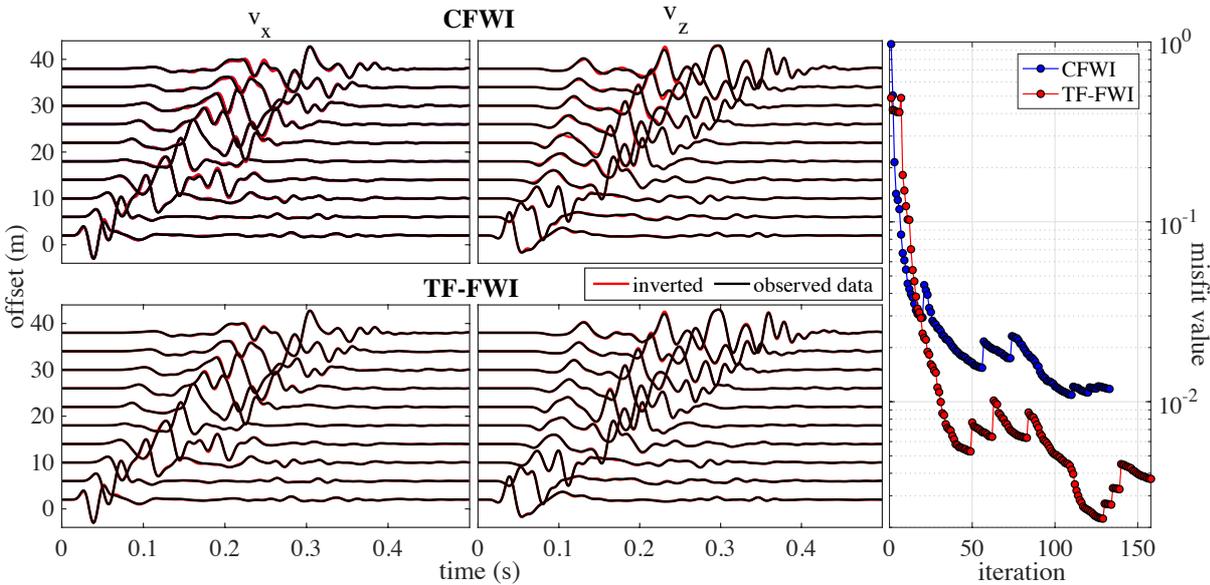


Figure 4.14: Synthetic reconstruction test. Comparison of the observed data with the data obtained using the final models of conventional FWI (CFWI) and time-frequency FWI (TF-FWI) shown in Figure 4.12. Comparison of velocity seismograms for the horizontal in-line (v_x , first column) and vertical (v_z , second column) components for 10 different offsets. b) Misfit function for the CFWI and TF-FWI (third column).

Conventional FWI: CFWI

With conventional FWI (Figure 4.12, column 3) I perform a multi-parameter inversion (according to Table 4.2) using the total wavefield. CFWI exhibits a very good reconstruction of the v_s model. This is a general strength of near-surface FWI since the wavefield is dominated by high-amplitude Rayleigh waves, which are mainly sensitive to the S-wave velocity structure. The v_s model is well resolved within the upper 8 m, with the exception of some artifacts around the discontinuity of the groundwater table. These can be related to some trade-off from inaccuracies in the v_p model which contaminate the reconstruction

of the v_s model (Groos et al., 2014). The interface separating the upper half-space and the groundwater table suffers from artifacts.

The reconstruction of the v_p model is not accurate for two main reasons: 1) The low sensitivity of the high amplitude Rayleigh waves to the v_p model, and 2) the weak contribution of the low-amplitude P-waves to the misfit function. Although there is a decrease of the velocity in the area where the low-velocity horizontal layer is present, the shape and structure do not correlate with the true values. Additionally, the trench in the v_p model is also not retrieved and there are artifacts at the locations of the sources. We observe an increase of v_p at the location of the trench in the v_s model due to cross-talk between v_s and v_p .

The density model is poorly reconstructed and suffers from artefacts and cross-talk effects. Although the location of the trench is accurate, both its shape and absolute value are far from the true model. Contaminations from the velocity models produce artifacts over the whole model area (Athanasopoulos and Bohlen, 2016; Athanasopoulos et al., 2018a).

Time-frequency FWI: TF-FWI

The final models from TF-FWI (Figure 4.12, column 4) exhibit a greater similarity to the true models compared to the conventional FWI results. The higher weighting of the refracted waves in the first stages and the gradual incorporation of the later arrivals allowed for an improved reconstruction of the v_p model. The trench is now visible in the v_p model and the absolute values match satisfactorily with the true velocity. However, the shape of the trench slightly differs from the true model at the boundaries. The artefacts in the vicinity of the sources are now reduced. Furthermore, the low-velocity horizontal layer is retrieved, in both shape and absolute value. The cross-talk from the v_s to the v_p model is better localized and has a smaller footprint compared to CFWI.

The v_s model is also reconstructed well. Compared to conventional FWI the improvements are small. We can identify small improvements immediately below the trench. This seems to be the result of the reduction of trade-off between v_p and v_s models. In Athanasopoulos et al. (2018a) I have shown that in case of an inaccurate initial v_p model, the quality of both the v_s model and the density model is greatly reduced, making the use of TF-FWI approach essential.

The reconstructed density model provides also a higher correlation to the true density. The artifacts around the groundwater table discontinuity are reduced, revealing a more coherent interface as given in the true model. Both the shape and density values of the trench are slightly improved. Also in this case, the cross-talk from the v_s model is better isolated.

1D vertical profiles

Vertical parameter profiles at the locations shown in Figure 4.12 (column 1) are plotted in Figure 4.13. The reconstruction of the absolute values of the v_p model is improved by TF-FWI, not only in the background but also at the location of the trench and the discontinuity (4.5-5.5 m depth). The v_s model from TF-FWI shows a better coherency to the true values, compared to the CFWI where we observe variations at a depth of 6 m below the trench. However, the overall quality of the v_s model is very good for both methods. The density as already discussed is the weakest parameter. Several artifacts are present in both methods, with

slight better the case of TF-FWI, especially at the location of the trench, where the absolute values are closer to the true ones.

Comparison in the data-space

In order to verify the quality of the reconstructed models I compare the seismograms for ten different offsets at the last frequency step (70 Hz) of the vertical and horizontal inline component for both approaches with the observed (true) data. Figure 4.14a shows the seismograms obtained from the observed (left) data and synthetic data from FWI (middle) and TF-FWI (right), respectively.

In both cases the fit of the data is very high for both horizontal and vertical components. However, mainly in the far offset traces, both the refracted and Rayleigh wave have a better fit in the case of TF-FWI in both components. The misfit as a function of iterations shows clearly that apart from the more accurate model reconstruction the residual energy can be further reduced by additional TF windowing.

4.2.6 Conclusions

Conventional multiparameter elastic FWI fails to reconstruct the P-wave velocity model due to its low sensitivity to Rayleigh waves which dominate the misfit. To overcome this problem, I suggested a time-frequency approach where we gradually increase the contribution of later arrivals in the course of the iterations. The TF-FWI approach showed improved reconstruction of all three model parameters, especially of v_p . The improved v_p allowed to reduce artifacts in v_s and obtain a more accurate density model. A comparison in the data-space revealed the improved fit of both the early and later arrivals.

4.3 Parameter sensitivity in far offset acquisition setup

As I have shown in the previous section, the reconstruction of accurate P-wave velocity models in short-offset acquisitions needs more sophisticated approaches compared to reconstructing the S-wave velocity. In this section, I will perform a similar numerical test but for a large-scale acquisition and examine the resolvability of the elastic parameters in such a scenario. I will then compare once again the TF-FWI with the conventional FWI.

4.3.1 Model description

The true model, shown in Figure 4.15 (contours) and 4.16 (first column), is a simplified version of the 2D synthetic model by Gray and Marfurt (1995), commonly known as the Canadian Foothills model. The original structure of the model is a linear velocity gradient in vertical direction overlaid with complex features of various sizes and shapes, with a mostly constant velocity. It corresponds to the geological features found beneath the foothills of the Canadian Rocky Mountains in the area of northeastern British Columbia.

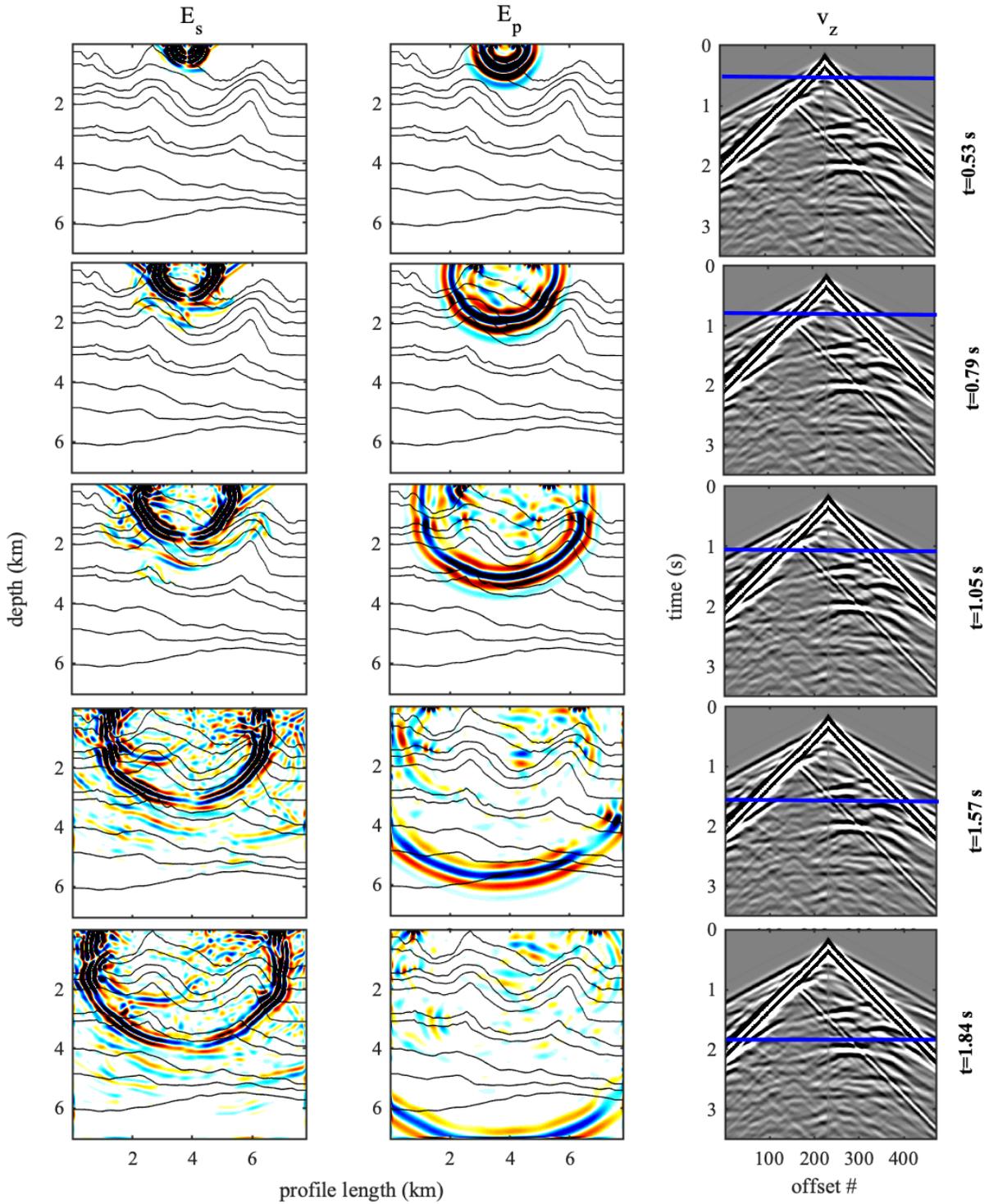


Figure 4.15: Snapshots (left two columns) of the S-wave (E_s) and P-wave (E_p) intensity with the contour of the underlying velocity model of the Canadian foothills. The corresponding seismograms of vertical particle velocity are shown in the right column. The source is located at the surface in the center of the model. The propagation time is increasing from top to bottom and the exact times of the snapshots are indicated by the blue lines in the seismograms.

The dimensions are 25 km in length and 10 km in depth with the sea level reaching depths of 2 km. It shows strong topographic variations with height differences of up to approximately 1.5 km. Originally, the P-wave velocity values (only acoustic modeling was considered) above the surface were set to 4 km/s, to avoid interaction of the wavefield with the interface. The modified model I use was initially created by Krieger (2019), when he evaluated FWI of land seismic data under strong topographic variations. In my case, since my focus is on showing the differences in the parameter reconstruction compared to the short-offset acquisitions (see previous sections), I set the values above the interface to 3.5 km/s.

I keep the original P-wave velocity, which varies from 3.5 to 5.9 km/s and estimate the S-wave velocity and density from empirical relationships. By choosing a Poisson's ratio of 0.25 (leading to $\frac{v_p}{v_s} = \sqrt{3}$), as the basis of the subsurface (from whose values the synthetic model was constructed) consists of crystalline rocks, I obtain the S-wave velocity. For the density estimation, I use Gardner's relation

$$\rho = 310 \frac{\text{kg s}^{\frac{1}{4}}}{\text{m}^{\frac{13}{4}}} v_P^{\frac{1}{4}}, \quad (4.8)$$

which holds for P-wave velocities above 1.5 km/s (Gardner et al., 1974). The final models are shown in Figure 4.16 (first column).

4.3.2 Acquisition setup

The acquisition geometry of the synthetic study consists of a linear profile of 471 vertical-component geophones, with an equidistant spacing of 15 m, and 48 vertical sources, with an equidistant spacing of 150 m. Both the receivers and sources are placed on the surface, as I am only interested in a large-scale model which includes complex features. The model consists of 1040 grid points in the horizontal direction and 940 grid points in the vertical direction with a grid spacing of 7.5 m, resulting in a profile of 7.8 km in the horizontal and of 7 km in the vertical direction. I use a Ricker wavelet as source signal with a dominant frequency of 5 Hz, which generates frequencies up to approximately 20 Hz. I choose a simulation time of 5 s with a time step interval of $dt = 7.5e - 04$ s. I apply a Butterworth band-pass filter to the data, where the upper corner frequency of the filter (4th order) changed between stages and the lower corner frequency remains fixed at 1 Hz (Table 4.3). Additional specifications include the use of semicircular source tapers with a length of 20 grid points to precondition the shot-wise gradients and a smoothing using a 2D median filter, which has a size of 15 grid points in both spatial directions. Table 4.3 shows the different stages of the inversion. I find that for such a model the optimal strategy is to modify the TF-FWI, such that the same frequency stage is repeated without the application of the time-damping as shown in Table 4.3, however, the differences to TF-FWI were not significant.

The initial models are shown in Figure 4.16 (second column) and they consist of linear gradients of increasing velocity/density over depth, where the upper and lower values correspond to the ones of the true model. In Figure 4.15 I show the P- and S-wave intensities along with the particle velocity recordings, similar to the studies done in Chapter 3, for increasing propagation time (from top to bottom). In this case the propagation of both P and S waves and the interactions on the different subsurface structures becomes more prominent, due to the frequency content of the data and the sufficiently long acquisition offsets.

Table 4.3: Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter. The column ' γ ' indicates the damping factor (equation 4.3).

Stage	Update			LP in Hz	γ
	v_P	v_S	ρ		
1	1	0	0	2	20
2	1	1	1	2	0
3	1	1	0	4	10
4	1	1	0	4	0
5	1	1	1	8	5
6	1	1	1	8	0
7	1	1	1	15	0
8	1	1	1	30	0

4.3.3 Results

In Figures 4.16 and 4.17 I show a direct comparison of the final models for both CFWI and TF-FWI along with the true and initial models. Compared to the parameter reconstruction from the previous examples, we observe that in the case of large-scale acquisitions v_p can be accurately retrieved, both from CFWI and TF-FWI (Figure 4.16 third and fourth column, respectively). This is mainly due to the fact that the refracted waves have a bigger presence in the recordings and therefore the least-squares scheme used in FWI can improve the fit of the refracted waves compared to the synthetic cases of the previous sections. To illustrate this phenomenon, I plot once more the P- and S-wave energy intensities as I did in Chapter 3 for different time steps along with the velocity seismograms. At the given frequencies and wavelengths the propagation of the body waves and their interactions with the various layers become more prominent compared to the small-scale example.

A good reconstruction in v_p helps both the v_s and density reconstruction. The v_s model is now not predominantly influenced by the surface waves, but also by body waves. As we have seen previously, an accurate density model can be retrieved only in the case where the other parameters which are involved in the inversion are accurately reconstructed. This allows the algorithm to focus on fitting the amplitudes of the waveform, since the phases are in a good agreement with the true model.

From Figure 4.16, we can also see the illumination of FWI which is directly dependent on the acquisition setup. At the edges of the model (around 800 m from the left and 1 km to the right), the reconstructed parameters in both cases are blurred with no accurate values or structures. However, this is entirely expected since the ray coverage is poor or non-existent in these regions. Additionally, the edges of the subsurface layers are smoother than in the true model, which is a result of insufficient high frequencies used during FWI. The illumination on the right side of the model is generally lower due to the subsurface structure (at 6 km there is an up-dip on the velocity and density structure), causing poor reconstruction of all the elastic parameters. Finally, it is visible that the resolution strongly degrades at depths larger than 4 km, which is another effect of the illumination related to the acquisition geometry.

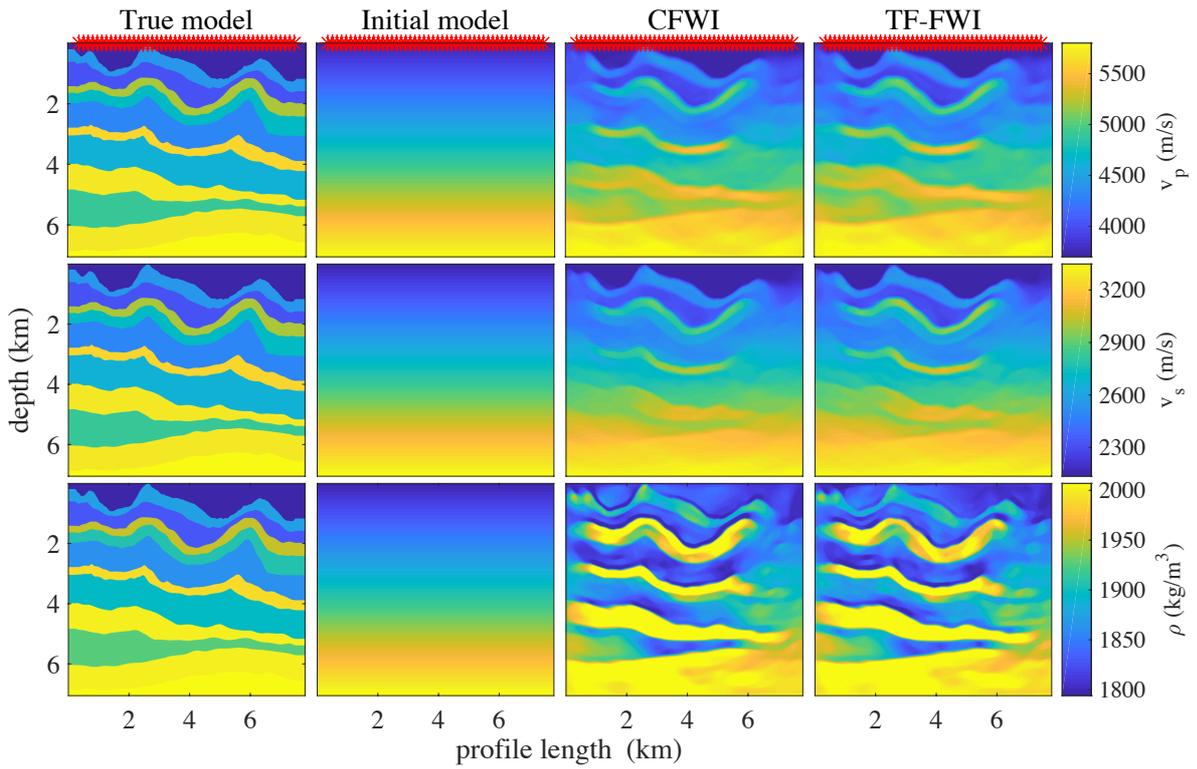


Figure 4.16: Models obtained by the inversion of synthetic data. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional (CFWI) and time-frequency (TF-FWI) inversion. Red asterisks represent the source locations.

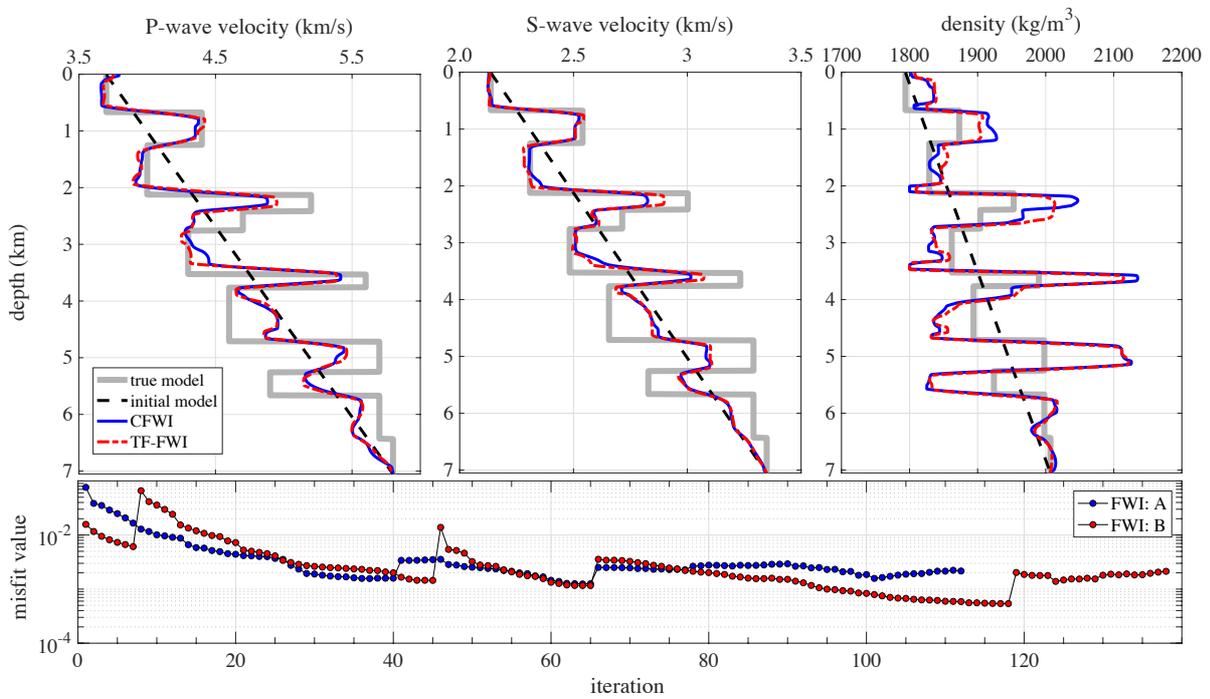


Figure 4.17: Top: 1D models of the true, initial and reconstructed parameters shown in Figure 4.16. Bottom: misfit functions for CFWI (A) and TF-FWI (B).

In Figure 4.17, we see 1D profiles from the center of the model (at 4 km). Although we could not spot significant differences between these two methods in Figure 4.16, we can see that despite the overall similar results the absolute values of the different layers is more accurately retrieved using TF-FWI. In general, we observe that the absolute values of each layer are not perfectly resolved in any of the three parameters, with the P-wave velocity model being the best and the density model the worst. The data misfit (Figure 4.17, bottom row) shows similar values for both cases, but in case of TF-FWI the values were consistently lower throughout the iterations.

4.3.4 Conclusions

In this section I inverted a large-scale seismic synthetic data set using CFWI and TF-FWI. In such cases the resolution of the reconstructed P-wave velocity model is greatly improved due to the better separation of the refracted waves from the high-amplitude surface waves. All three model parameters shown a good reconstruction, especially regarding the structural layout of the model. However, the absolute values of the individual layers were not perfectly retrieved, with the density model being the worst and the P-wave velocity model being the best. TF-FWI managed to improve the absolute values of the different layers to some extent, without increasing the overall computational cost (iteration number difference is negligible, see Figure 4.17).

4.4 Structural similarity constraints

Near-surface geophysical techniques are usually non-destructive low-cost methods. Often multiple geophysical methods are required in order to obtain reliable information, to reconstruct the reconstruction of the shallow subsurface. This is mainly due to the non-linear nature of these techniques, but also due to the various sensitivities of each method to specific structures of the subsurface. However, the fact that measurements are easily available gives a strong advantage in shallow applications, in the sense that several geophysical methods can be combined in order to obtain more geological information on the same target area.

Gallardo and Meju (2003) proposed the use of a structural similarity constraint in their inversion scheme in order to jointly invert resistivity and seismic data and demonstrated the benefits of the joint approach over the individual ones. Later Tryggvason and Linde (2006) showed that a cross-gradient constraint is beneficial also for inversion of P- and S-wave velocities in a local earthquake traveltime tomography study. Most recently Manukyan et al. (2018), employed a structurally constrained elastic FWI (CG-FWI) in a frequency-domain Gauss-Newton scheme to improve the quality of the reconstructed elastic parameters. Here I follow a similar approach but apply the cross-gradient in 2D elastic FWI as a constraint and not as a separate inversion problem. Additionally in my case, the inversion problem is solved by a preconditioned conjugate gradient approach using the adjoint-state method.

4.4.1 Theory of cross-gradients

By definition, the cross-gradient vector, \mathbf{cg} , can be expressed as the cross product of two individual gradients of models m_a and m_b , given by

$$\mathbf{cg}(m_a, m_b) = \nabla m_a(x, y, z) \times \nabla m_b(x, y, z). \quad (4.9)$$

To apply structural similarity between the two models equation 4.9 should become zero, meaning that the cross-product of two vectors is zero when the gradient vectors are either locally parallel, antiparallel or one or both are zero. For the 2D case, the x- and z-components of equation 4.9 vanish and the equation we would like to minimize becomes

$$cg_{a,b}(x, z) = \frac{\partial m_a}{\partial z} \frac{\partial m_b}{\partial x} - \frac{\partial m_a}{\partial x} \frac{\partial m_b}{\partial z} = E_{CG} = 0. \quad (4.10)$$

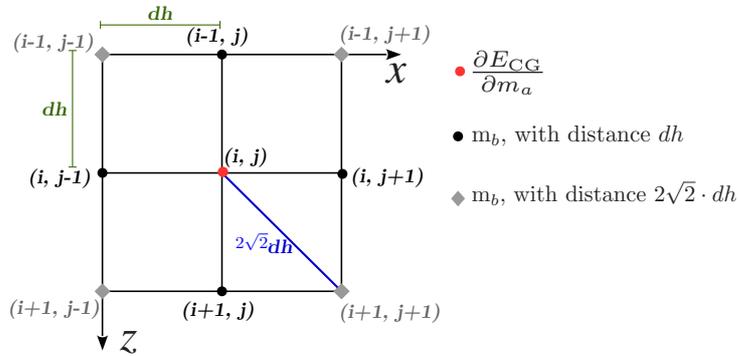


Figure 4.18: Eight grid-point stencil used to obtain the gradient of E_{CG} . Black circles correspond to the standard and grey rhombi to the rotated grid, both used to approximate the derivative at the central grid-point.

Joint gradients and objective function

As opposed to the previous studies, I suggest an alternative 8 grid-point stencil for the calculation of the gradient of the cross product. I calculate the cross-gradient function and its derivative over a model parameter m_a by using an eight-gridpoint scheme with second-order accuracy as shown in Figure 4.18. My scheme consists of a standard grid at which derivatives are approximated along the horizontal and vertical axis (Figure 4.18, black circles), and a rotated grid where the differential operators approximate derivatives along the diagonal (Figure 4.18, grey rhombi). This scheme can improve the calculation of the derivative for high-contrast discontinuities and it is therefore suitable for the case of the cross-gradient function (Athanasopoulos et al., 2018b).

The discretized formulation of the cross-gradient function after making use of the stencil shown in Figure 4.18 is given by

$$\begin{aligned} \hat{\mathbf{g}}^{\text{CG}} = & \frac{1}{dh^2} \left((m_a[i][j-1] - m_a[i][j+1])(m_b[i-1][j] - m_b[i+1][j]) \right. \\ & \left. - (m_a[i-1][j] - m_a[i+1][j])(m_b[i][j-1] - m_b[i][j+1]) \right) \\ & + \frac{1}{8dh^2} \left((m_a[i-1][j-1] - m_a[i+1][j+1])(m_b[i-1][j+1] - m_b[i+1][j-1]) \right. \\ & \left. - (m_a[i-1][j+1] - m_a[i+1][j-1])(m_b[i-1][j-1] - m_b[i+1][j+1]) \right). \end{aligned} \quad (4.11)$$

The discretized formulation of the gradient of the cross-gradient function using the same discretization is given by

$$\begin{aligned} \delta \hat{\mathbf{g}}^{\text{CG}} = & \frac{\partial E_{\text{CG}}}{\partial m_a}[i][j] = \left(\frac{1}{dh^2} \left((m_b[i][j-1] - m_b[i][j+1]) - (m_b[i-1][j] - m_b[i+1][j]) \right) + \right. \\ & \left. \frac{1}{8dh^2} \left((m_b[i-1][j-1] - m_b[i+1][j+1]) - (m_b[i-1][j+1] - m_b[i+1][j-1]) \right) \right) \times \text{sign}(\hat{\mathbf{g}}^{\text{CG}}), \end{aligned} \quad (4.12)$$

where the indices i and j refer to the discretized model blocks in the x- and z-directions, respectively, and dh is grid spacing. The joint inversion minimizes the combination of the least-squares misfit of the normalized wavefields and the cross-gradient term:

$$E_{\text{joint}}(\mathbf{m}) = E_{\text{seis}}(\mathbf{m}) + \gamma E_{\text{cg}}(\mathbf{m}), \quad (4.13)$$

where the parameter γ is a heuristically chosen parameter to weight the contribution of the cross-gradient term in the joint objective function. The last step required for FWI is to calculate the adjoint gradients but in this case accounting also for the joint objective function. I calculate gradients of both terms in equation 4.13 by using the adjoint-state method and equation 4.12. To stabilize the update and to balance the contribution of both the data misfit and of the structural constraint, I perform a normalized addition of both gradients, while I choose the maximum absolute gradient amplitude as normalization factor. The joint gradient reads

$$\delta \hat{\mathbf{g}}^{\text{JOINT}} = \left(\frac{(1 - \alpha) \delta \hat{\mathbf{g}}^{\text{seis}}}{\max(|\delta \hat{\mathbf{g}}^{\text{seis}}|)} + \frac{\alpha \delta \hat{\mathbf{g}}^{\text{CG}}}{\max(|\delta \hat{\mathbf{g}}^{\text{CG}}|)} \right), \quad (4.14)$$

where $\delta \hat{\mathbf{g}} = \{ \frac{\partial E}{\partial \rho}, \frac{\partial E}{\partial v_p}$ or $\frac{\partial E}{\partial v_s} \}$ and α is a parameter to weight the seismic gradient obtained by the adjoint-state method over the gradient of the cross-gradient objective function. This joint gradient approach weights both gradients as defined by α , but it does not preserve their amplitude information since the amplitude of the gradient of the cross-gradient function is much higher compared to the seismic one.

The variables underlying the cross-gradient implementation in the software IFOS2D are specified in Appendix B.

4.4.2 Synthetic study and results

In a synthetic study I compare the results of conventional FWI with CG-FWI. The acquisition setup is the same as in Sections 4.1 and 4.2 in terms of sources and receivers. The frequency content is once again the same ranging from 5 Hz to 70 Hz. The true model (Figure 4.19) is spatially correlated, but I am using the more complicated model of section 4.2 with the stochastic fluctuations included. The discontinuity of the two layers lies at approximately 6 m. Furthermore, in the upper layer v_s gradually increases with depth. The second column of Figure 4.19 depicts the initial models of v_p , v_s and ρ . The initial models of v_s and ρ consist of a smooth gradient of increasing velocity/density with depth.

In this case study, I use the true v_p model. The reason is that I wish to focus on the stability of the joint inversion that I propose and not account for other effects that arise from inaccuracies of the v_p inversion as shown in the previous sections. Therefore, I use the cross-gradient to pass structural information from v_p to v_s .

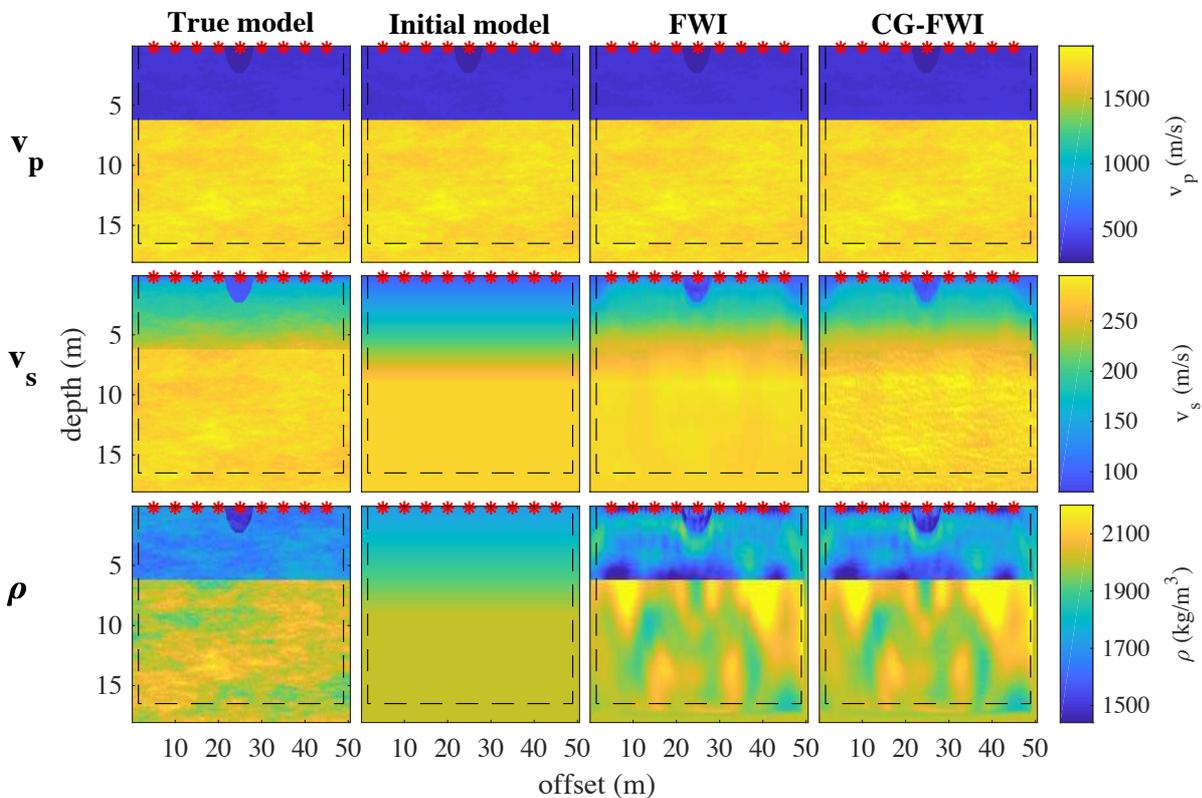


Figure 4.19: The rows represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional and structurally constrained FWI. Red stars represent the source locations. The true v_p model is used throughout the inversions and it is not updated. The PML boundaries are shown with dashed lines.

The conventional FWI (Figure 4.19, column 3) exhibits a very good reconstruction of v_s since the wavefield is dominated by high-amplitude Rayleigh waves, which are mainly sensitive to the S-wave velocity structure and also the true v_p is known. However, none of the small-scale structures were able to be resolved (due to the frequency content of the data) and the discontinuity is not perfectly retrieved. The overall density model is also quite accurate, due to its large dependency on the P-wave velocity. However,

the density values inside the trench are not well resolved. The final model from CG-FWI (Figure 4.19, column 4) exhibit a greater similarity to the true model than conventional FWI results. In v_s , the small-scale structures are resolved in high detail and also there is a clear interface of the discontinuity which highly matches the true model. However, since the choice of weighting for the cross-gradient plays a crucial role, the absolute values of the small-scale structures are not always precise. A more sophisticated normalization technique prior to the construction of the joint gradients would improve the structural information. Finally, the reconstructed density model is very similar to the conventional FWI with the exception of the trench which now matches better the true values.

The influence of the cross-gradient constraints on the FWI scheme can be easily observed by looking at the gradients which contribute to the actual updates of the models, as shown in Figure 4.20. The gradient of the cross-gradient term (Figure 4.20a) depicts the structural variations of the v_p model and is properly combined with the seismic gradient (Figure 4.20b) obtained by the adjoint state-method through the joint approach (Figure 4.20c), despite the large amplitude difference. This shows the potential of this approach to be used also in other combinations of models for which we wish to enforce structural similarity (e.g., v_p , density and v_s , density) and the direct applicability of the joint scheme with other geophysical methods. The α value that I used is 0.05 since I wanted to obtain most information from the seismic gradient.

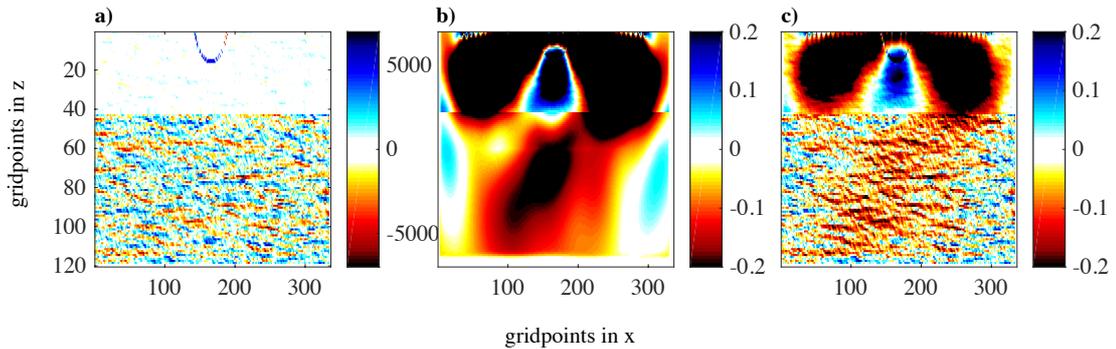


Figure 4.20: Gradients used to update the v_s model. a) Gradient of the cross-gradient term calculated by equation 4.12 from v_p . b) Seismic gradient of v_s obtained by the adjoint-state method. c) Combined gradient obtained by equation 4.13 with $\alpha = 0.05$.

4.4.3 Conclusions

I have developed a stable time-domain structurally constrained FWI scheme using an 8 grid-point stencil discretization scheme. By applying this scheme to a synthetic study, I demonstrated its ability to pass even small-scale structural information from v_p to v_s and improve the subsurface reconstruction compared to conventional FWI. On the downside, additional normalization of the cross-gradient updates is required to improve this technique and make it usable for field data applications. With the proposed joint approach, it becomes straightforward to combine other geophysical methods with seismic FWI, although the heuristic choice of the parameters γ and α require intense testing.

4.5 Summary

In this chapter a series of synthetic tests were shown. Different techniques were employed in order to construct strategies to better recover the subsurface parameters we wish to invert for. The principles of joint inversion of seismic and non-seismic geophysical methods were discussed, showing their benefits versus individual data inversions. Additionally, a time-frequency windowing approach was developed, which allowed the retrieval of improved P-wave velocity model, which is very challenging in near-surface short-offset acquisitions. The methodologies that were introduced and evaluated here will be applied in the field data cases of Chapters 6 and 7.

Chapter 5

Preprocessing of field data: General overview

5.1	Data acquisition	75
5.2	Preprocessing workflow	76
5.3	Initial model building	78
5.3.1	Traveltime tomography	79
5.3.2	Multichannel analysis of surface wave (MASW)	79
5.3.3	Q estimation	80
5.3.4	Cycle skip evaluation	81
5.4	Source time function inversion	82
5.5	Summary	84

In this chapter a complete guide for preparing field seismic data for any seismic inversion scheme will be presented. All the essential steps which are required prior to or during the inversion will be examined thoroughly, in a way that they can be further applied in any framework of seismic data inversion.

5.1 Data acquisition

As discussed in Chapter 2, seismic waves include body waves that travel three-dimensionally through the Earth's subsurface and surface waves that travel near its surface. From Figure 2.1, it became clear that in order to record P-SV waves we can use vertical polarized sources with either vertical or horizontal inline (along the direction of the sensor array) components. For SH recordings it is necessary to use a horizontally polarized shear force and record using crossline components.

In Figure 5.1 I show the three components of particle velocities from the field data study of Chapter 7. The detailed site description will be discussed later, however, I will use these data as an example of the preprocessing workflow. This workflow was applied to every real data case I examined. Because surface waves are constrained to propagate close to the surface, their amplitude decay as a function of source-distance is smaller compared to body waves, which propagate spherically in all three dimensions. Additionally, the propagation velocity of surface waves is frequency dependent. In Figure 5.1 both the high amplitudes of the surface waves and their dispersive nature are prominent.

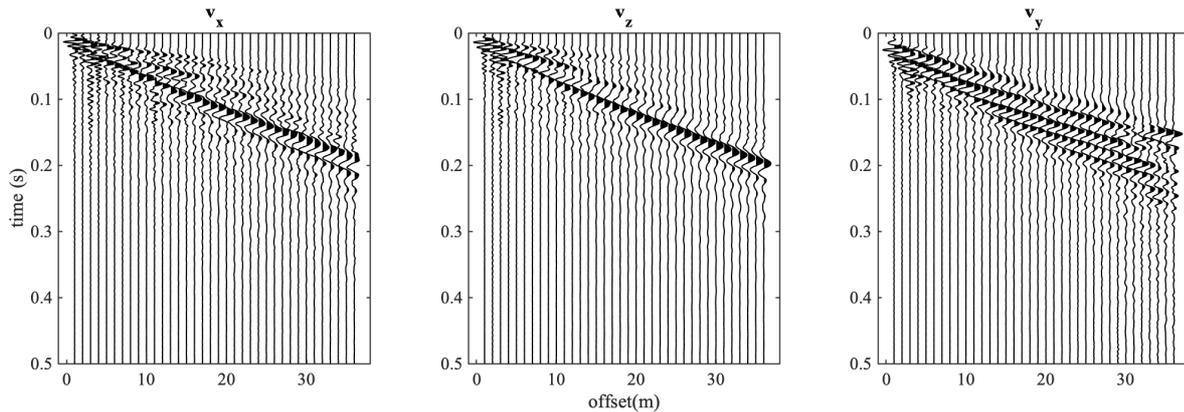


Figure 5.1: Trace-wise normalized shotgather of the seismic field data. The corresponding source is located at the profile coordinate of -3 m. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH waves. The gathers in both cases are dominated by the surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.

An advantage of performing shallow seismic inversion and processing in the presence of surface waves is that 1) they can be easily excited and recorded with standard field equipment such as a sledge-hammer, 2) their large amplitudes provide a very high signal-to-noise ratio making their distinction and therefore processing straightforward. From Figure 5.1 the Rayleigh wave can be identified by the wave group which is recorded at 0.18 s on the last trace of components v_x and v_z . Complementary in the v_y component, the Love wave is arriving slightly earlier and in this field case it is much more dispersive. It is obvious that the amplitudes of the refracted P-wave, recorded by v_x and v_z components, are highly clipped due to the high amplitudes contained in the surface waves. The coda between first arrivals at the onsets of surface waves corresponds to converted energy from various wave types and other scattering events, commonly identified as higher surface wave modes. An additional step often performed in the field is the stacking of multiple shots at the same location. This is done in order to account for the fact that the force generated from an impact between the sledge-hammer and the steel plate can vary. By stacking (usually around 5 times is sufficient, except in noisy environments) the signal-to-noise ratio also gets improved.

To summarize, the first step for real data seismic inversion consists of proper acquisition planning, in particular regarding the polarization of both sources and receivers. Then follows the identification of the various wave types and the evaluation as well of the signal-to-noise ratio.

5.2 Preprocessing workflow

In this section, I discuss all the steps required to make the raw recordings ready for use in an FWI algorithm (in my case the IFOS2D software). For purpose of illustration, I show the preprocessing of the vertical component shown in Figure 5.1. The results of each individual step are presented in Figure 5.2, starting from the raw data as acquired in the field and moving clockwise over the different steps.

During the acquisition for offsets of around 50 m, which is typical for resolving the upper 15-20 m in depth, we usually record between 1 to 2 s of signal with a sampling interval of 250 ms. From my experience this is sufficient for studies of this scale in order to properly include the surface waves, as is illustrated also from the synthetic studies. Additionally, the data should be stacked as mentioned in the

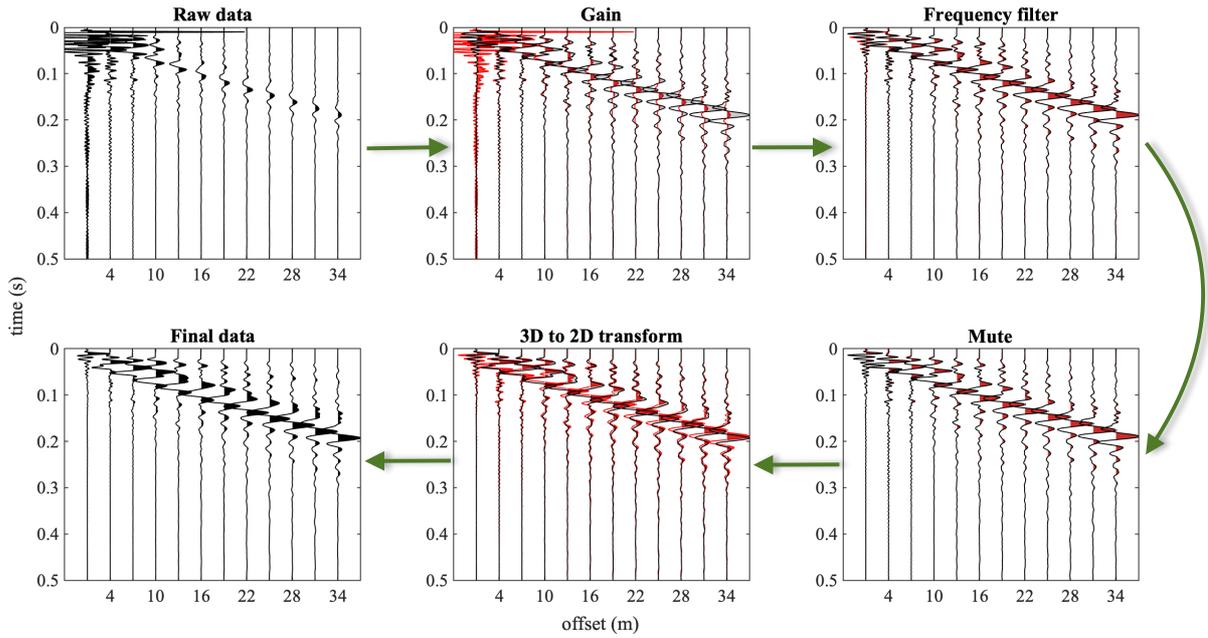


Figure 5.2: Complete preprocessing workflow for field seismic data. Starting from the raw data on top left image and moving clockwise through the various processing steps. With black color the result of the particular step is shown, while red indicates the previous step for comparison. Note that an transparency level is enforced to show both stages at all times. Every third trace is shown for better visualization of the various preprocessing steps.

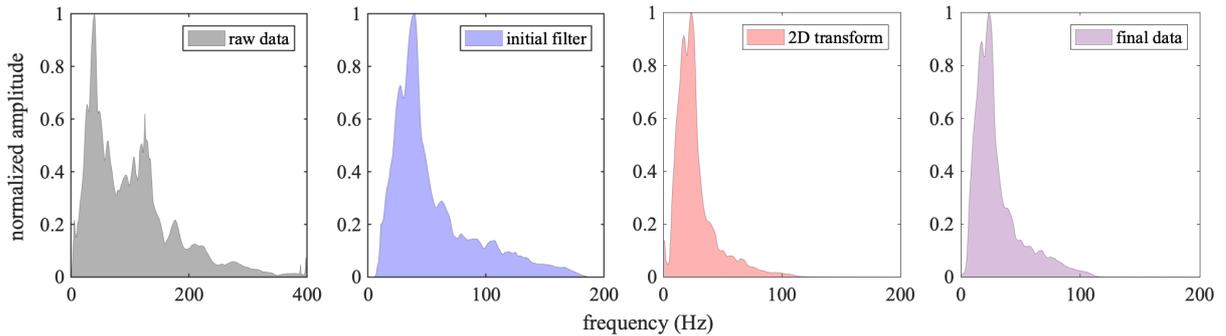


Figure 5.3: Frequency spectra of selective preprocessing workflow stages from Figure 5.2.

previous section to improve the signal-to-noise ratio and reduce the non-linearity of the induced source, e.g., the variability of the hammer impacts. Once the data are transferred to a format we can work on, in my example in SU-format (Stockwell, 1999) for compatibility reasons with IFOS2D, we can proceed by reducing the recorded time to only the time-window including the main events. This is done to reduce the computational time during FWI. As a next step, a gain is applied to the data which acts as a trace-wise normalization, in order to boost the arrival times of all near- and far-offset traces in a similar manner. This is important since as mentioned previously, the surface waves completely dominate the energy spectrum of the data set, especially compared to the first arrivals. Additionally, the gain helps to partly compensate for geometric spreading and dissipation due to wave propagation in real media.

I then filter out the low frequencies with respect to the eigenfrequency of the used sensors, which in the field data examples is 4.5 Hz. Additional filtering of the high-frequency content is important since it contains mostly noise. At this step it is also necessary to remove any amplitude prior to the onsets

of the first arrivals. Any event before these onsets corresponds to noise from the field or other passive interactions from the background environment.

After removing events which are most likely to be associated with noise, and enhancing the contribution of the various wave types, we now require to remove any out-of-plane events. More specifically, field data recordings include information from the full 3D volume of the subsurface. The recorded wavefield is excited by hammer blows which act as point-sources. Since in this work the inversion is performed on 2D sections of the subsurface, it is important to remove any events which are present in the data and correspond to subsurface structures outside the vertical plane of the acquisition array, since 2D FWI requires line-sources. There have been several studies to deal with the 3D to 2D conversion of seismic data. In my case, I use the method of *direct transformation* as suggested by Forbriger et al. (2014) and Schäfer et al. (2014), which includes a trace-wise convolution with $\sqrt{t^{(-1)}}$ and a multiplication with $r\sqrt{2}\sqrt{t^{(-1)}}$, where t is traveltime and r the offset. Since the convolution with $\sqrt{t^{(-1)}}$ corresponds to a half integration of the seismic traces (Figure 5.2), it is necessary to apply the high-pass filter as shown above in order to avoid artifacts due to low-frequency noise prior to the 3D to 2D transformation.

By having a more in-depth look over how the frequency content of the data is changed after each step, in Figure 5.3 I show the spectrum of the reference shot gather for selective processing stages. The raw data in this case contain energy up to approximately 400 Hz (Figure 5.3, black). The initial band-pass filter greatly narrows the amplitude spectrum that now spans from 5 up to 190 Hz, which is close to limitations of the acquisition's hardware (Figure 5.3, blue). The application of the 3D to 2D transform greatly reduces the frequency band, by mainly increasing the contribution of the lower frequencies (Figure 5.3, red). As noted above, some low-frequency noise arises due to the effect of the transformation, which needs to be removed. The final data consist of frequencies up to 100 Hz, with the main events associated with frequencies ranging from 8 to 60 Hz (Figure 5.3, magenta). For seismic applications of this scale, this is an expected/realistic frequency band as shown by many previous studies (Dokter et al., 2014; Groos et al., 2017; Wittkamp et al., 2018).

Final steps of the reprocessing workflow include 1) the application of a small time-delay of the whole data set (zero-padding) in order to avoid acausal effects in the inverted source time function (see Section 5.4) and 2) resampling in order to satisfy the discretization criteria imposed by the FD approach.

5.3 Initial model building

The choice of an initial model for any inversion scheme and in particular the FWI is of great importance. FWI is a highly non-linear problem and this is further enhanced in the near-surface region due to the dominance of the surface waves and their dispersive properties. Therefore, a proper investigation is required prior to FWI in order to obtain initial models that will include sufficient information of the actual subsurface and avoid cycle-skipping. These models have to predict the main wave phases and allow local convergence of the inversion. In the following, I will explain ways to obtain sufficiently good initial models for the viscoelastic parameters involved in the inverse problem.

5.3.1 Traveltime tomography

With the absence of far-offset traces in the scale I am examining in this thesis, obtaining an accurate initial P-wave velocity model is a rather difficult task. However, although the small-scale variations of the subsurface cannot be easily inferred, this does not hold for the depth of important discontinuities, such as the interface to a bedrock. By performing a simple 1D refraction tomography, e.g., identify the onsets corresponding to the direct and refracted wave, we can predict accurately the depth of such interfaces. Shot-gathers from opposite sides of the acquisition layout can be further used to detect dipping refractors. This is particularly important since FWI will not be able to correct for misplaced discontinuities when the contrasts in the values is rather high (see Section 4.1.7). Additionally, since the preference of the objective function lies on the minimization of data discrepancies associated with surface waves, it becomes clear that the true depth of such interfaces will not be recovered, unless accounted for by the initial model.

This initial P-wave velocity is used with Gardner's relation (equation 4.8) to provide first estimates of the initial density model. while further adjusting the density values using the Poisson's ratio after estimating the initial S-wave velocity model.

5.3.2 Multichannel analysis of surface wave (MASW)

During the last decades there have been many studies which used the surface waves to retrieve the S-wave velocity of shallow subsurface targets. Such studies included the inversion of dispersion curves through spectral analysis of surface waves, SASW (Dziewonski et al., 1969; Heisey J.S., 1982), multichannel analysis of surface waves, MASW (Park et al., 1999; Xia et al., 1999; Bohlen et al., 2004; Socco et al., 2010) and Fourier-Bessel expansion coefficients (Forbriger, 2003a,b).

The main assumption of these techniques is that the subsurface properties vary mainly with depth, therefore cannot properly account for situations where strong lateral variations exist. The 1D layered model and the plane wave assumptions which are considered in the calculation of surface-wave dispersion curves limits the dispersion-based imaging to subsurface without strong lateral heterogeneities. The identification of dispersion curves can also pose a problem, especially in the presence of higher modes (Zhang and Chan, 2003). However, since in my case I am interested in resolving an initial model which matches the main phases of the surface waves, the application of MASW is sufficient.

Initially, I perform a slant-stack to obtain the phase-slowness as a function of frequency (Figure 5.3) and compare it with the FK transformation (McMechan and Yedlin, 1981). A normalization at this point is necessary to boost the amplitudes of the full dominant mode of the surface wave. The dispersion curve image that is obtained carries information on the subsurface properties also in terms of higher modes, however, here I focus only on the fundamental mode. I proceed by picking the values of phase-slowness which correspond to the dominant mode (Figure 5.4, left).

For the inversion, I require a forward solution to the dispersion imaging, which in my case corresponds to Knopoff's technique (Knopoff, 1972), as it is quite reliable for the typical frequency range of shallow

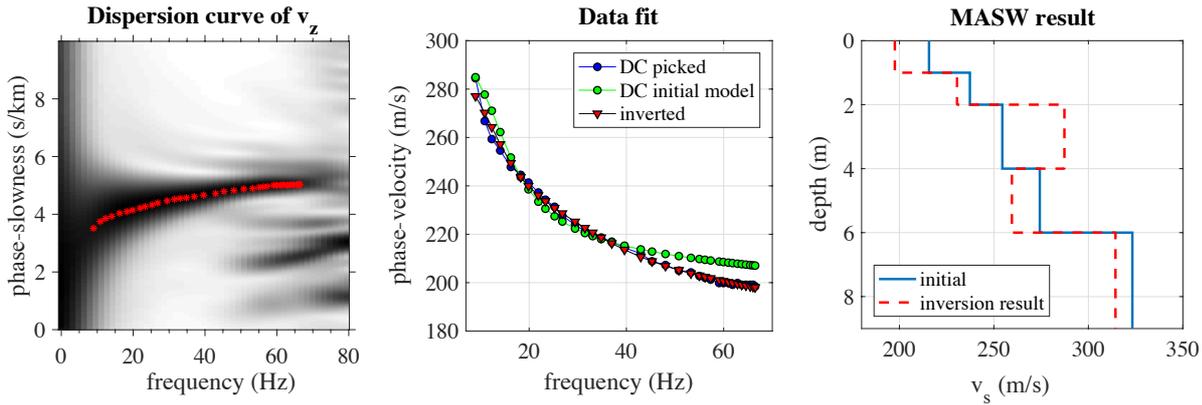


Figure 5.4: MASW inversion process. Right: slant-stack transformation and picking the fundamental mode of surface wave. Middle: observed, initial and inverted data. Left: inverted model from MASW versus the initial model.

seismics (Xia et al., 1999; Pan et al., 2016). Basically, the Rayleigh-wave phase velocity can be expressed as a function F of five variables and is given in its non-linear implicit form by

$$F(f_j, v_{R_j}, v_s, v_p, \rho, h) = 0, \quad (5.1)$$

where f is the frequency, v_{R_j} the phase velocity and h the thickness of each layer.

In terms of the acquisition setup which is required for proper dispersion imaging, I use a nearest source-receiver offset of twice the maximum investigation depth. I use an iterative solution to a weighted least-squares inversion problem of Rayleigh-wave phase velocities, as suggested by Xia et al. (1999), but any optimization technique should be able to provide sufficiently good results, since the non-linearity of the dispersion curve inversion is not as strong, as for example, for FWI. The inversion results of this reference shot-gather is shown in Figure 5.4 (right) along with the initial model which is estimated from empirical relations.

Figure 5.4 (middle) shows also how the inversion allowed to minimize the discrepancy between the observed and inverted data. After obtaining such a 1D representation of the subsurface, I move on by extrapolating the 1D values to a 2D plane, which I then use for the initial S-wave velocity model (further quality control is performed and will be discussed in Section 5.3.4).

5.3.3 Q estimation

In Chapter 2 I discussed the fundamental equations that describe the wave propagation in viscoelastic media. In case of shallow seismic investigations, the media consist of mainly unconsolidated sediments. In such cases, Schön (1996) have shown that the media are characterized by significant effects of anelastic damping, with Q varying from 5 to 80. I have shown how we can include the rheological model using a generalized standard linear solid (Blanch et al., 1995; Bohlen, 2002) in the elastic wave equations, by estimating the parameter Q through equation 2.21.

To estimate the quality factor I compare the field data with viscoelastically forward-modeled data which were obtained from the initial model. The attenuation parameters are therefore treated as prior information and will remain homogeneous and unchanged throughout the inversion process. Groos et al. (2014) have

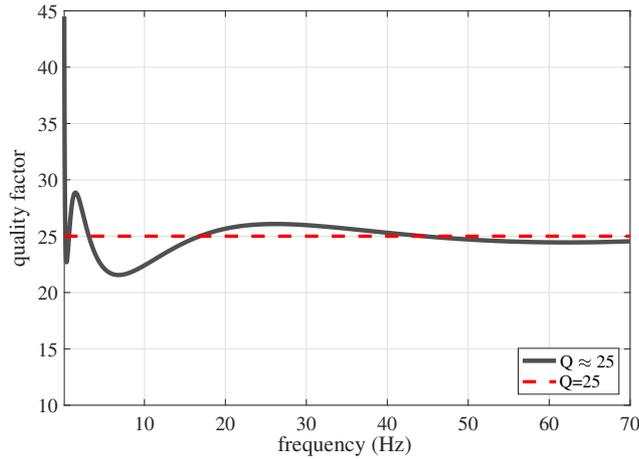


Figure 5.5: Exemplary rheological model. The quality factor of 25 is approximated by three relaxation mechanisms of a generalized standard linear solid with the relaxation frequencies for $\sigma=4.5912$, $\tau=0.0821$, $f_1=0.3579$, $f_2=6.7362$ and $f_3= 82.0317$ Hz.

shown that even a constant attenuation model is sufficient to significantly improve the inversion results. Furthermore, although there are already implementations of inverting for the attenuation parameters Q_p and Q_s (Fabien-Ouellet et al., 2017), the inter-parameter cross-talk due to the low impact of these two parameters on the waveform amplitudes significantly reduces the quality of the retrieved models (Chapter 3.2).

Before I compare the two data sets, the estimation of the source-time function (described in the next section) is required. I try different values of Q_p ($= Q_s$, I assume the same attenuation model for both P- and S-waves) for each shot and calculate the misfit for the sum of all shots and receivers. Using an L-curve criterion I am able to choose the Q value that minimizes the misfit between synthetic and observed data the most. An important feature of this methodology is that the choice of the appropriate Q value is done after the inversion of the source-time function. Groos et al. (2014) have shown that the source wavelet correction can account for differences between field data and synthetic data that are caused by too high quality factors used to generate the synthetic data. Therefore, I recommend a rather high quality factor, considering that the subsequently applied source-wavelet correction acts also as a low-pass filter.

The quality factor is approximated by three relaxation mechanisms in the field data cases that will follow. An example of such a model can be seen in Figure 5.5. I use three relaxation mechanisms due to the fact that I have frequencies of up to around 70 Hz in the field data. Using less relaxation mechanisms would cause a weak approximation for a homogeneous Q model, due to the fact that the frequency dependence of the quality factor were no longer be constant, as is approximately the case with three mechanisms.

5.3.4 Cycle skip evaluation

As a final step to assess the quality of the initial model, I perform a quality control check for every single shot and receiver between the observed and synthetic data. A well defined criterion to evaluate the existence of cycle-skipping between observed and synthetic data (from an initial model) is given by Virieux and Operto (2009). They state that the difference between the synthetic and observed data should be less than half the period, so that the Born approximation will be valid.

In my case, I designed two criteria to gauge the existence of cycle-skipping. I initially calculate the auto-correlation of the observed data and overlay the cross-correlation between observed and synthetic data (Figure 5.6). An important feature is that in general the accurate prediction of far-offset traces from an initial model is more challenging compared to the near to intermediate offsets (Figure 5.6), due to the longer propagation of the various wave types through the subsurface, which are not properly included in a typical initial model (usually a smooth version of a-prior information).

In the first criterion I estimate the difference of the maxima from the two series and restrict it to be less than half the period. In the second criterion, the two closest minima from the maximum value of the cross-correlation (left and right side) should include/surround the maximum of the autocorrelation (Algorithm 1).

Algorithm 1 Cycle-skipping evaluation - 2nd criterion

```

1:  $period = 1/(2 * f_{max\_iter\_stage}), nslimit = period/dt$ 
2: loop1:  $shot\_number = 1 : 1 : shots$ 
3: obs  $\leftarrow$  read observed data, syn  $\leftarrow$  read synthetic data
4: loop2:  $trace\_number = 1 : 1 : traces$ 
5: result(1) = index of max amplitude of auto-correlation (obs)
6: result(2,3) = index of two minima from cross-correlation (obs, syn)
7: difference(1,2) = result(1)-result(2,3)
8: if difference(1,2)(trace_number) > nslimit then
   cycle skip  $\leftarrow$  True
9: goto loop2
   goto loop1

```

The effect of the two different criteria is shown in Figure 5.6, where I compare the data from an initial model and the observed reference shot gather. The panels including the result of the two criteria have the cycle-skipped traces removed. As we can see, the first criterion is quite strict and leads to a higher trace rejection rate, while the second one seems more appropriate and closer to what Virieux and Operto (2009) recommend. In case a high number of traces is rejected after including all the shot gathers, then the initial model is rejected.

After thoroughly testing in all the inversion cases of this thesis and further additional studies, I highly recommend the use of the Algorithm 1 prior to any inversion. It can be easily included in any inversion framework or datafit evaluation.

5.4 Source time function inversion

An important step during an inversion of field data constitutes the characterization of the induced source, which corresponds to the impact of the hammer on a steel plate, as commonly used in shallow seismics. In order to obtain this information, I perform a stabilized deconvolution of the recorded wavefield with the synthetic one, following the work of Pratt (1999) and Forbriger et al. (2014). This procedure basically involves the solution of a damped, linear least-squares optimization problem. The linear relation between

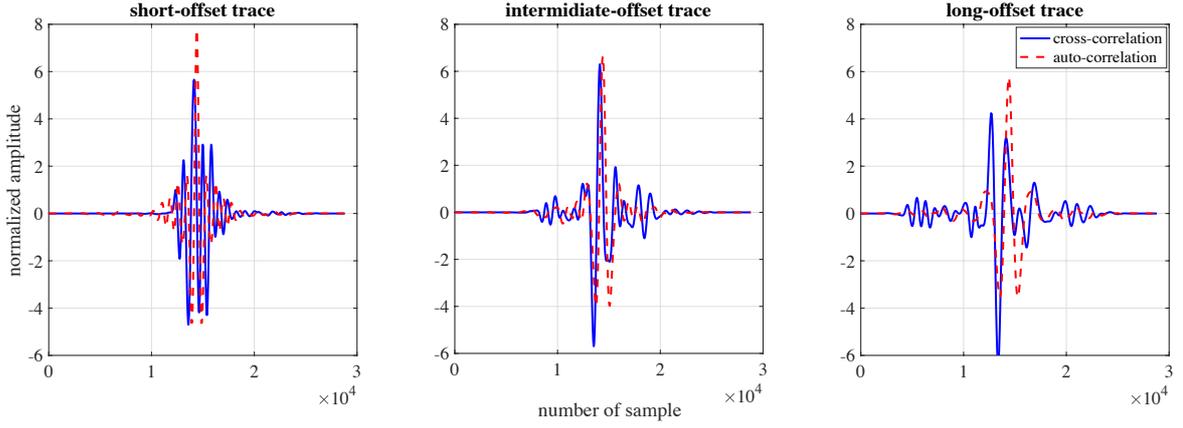


Figure 5.6: Overlay of the autocorrelation of the observed data and the crosscorrelation of them with the synthetic for different offsets. In the long-offset trace cycle-skipping is present.

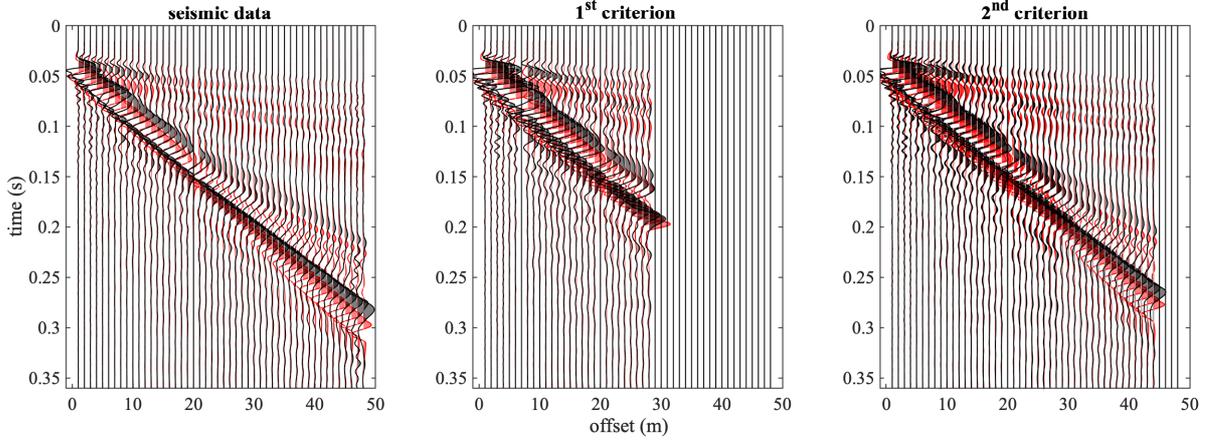


Figure 5.7: Overlay of field (black) and synthetic (red) data showing the effect of the two criteria for cycle-skipping evaluation. In the middle and right figure the traces classified as cycle-skipped are removed for illustration purposes.

the source wavelet and the seismograms can be used to obtain a wavelet correction filter. In the frequency domain this reads

$$\mathbf{S}\vec{u}(\omega_k) = s(\omega_k)\vec{f}(\omega_k), \quad (5.2)$$

where the tilde stands for the Fourier transform of the corresponding quantities and \mathbf{S} is a complex-valued impedance matrix which represents the model parameters. In the elastic case all parameters except \mathbf{S} depend on the discrete frequencies ω_k . It is necessary to provide a broader range of frequencies in the synthetic seismograms than in the case of the observed seismograms. Therefore, as initial source estimation the cubed sine wavelet (equation 3.9) which I used in the synthetic data is ideal, due to the inclusion of high-frequency content (Figure 3.4). Further delaying the energy, through zero-padding both observed and synthetic data, helps to mitigate potential acausal parts in the source wavelet.

We can then proceed to obtain this filter by minimizing the following misfit function:

$$s(\omega_k) = \frac{\vec{d}^T \vec{u}^*}{\vec{u}^T \vec{u}^* + \varepsilon}, \quad (5.3)$$

where the asterisks in equation 5.3 represent the complex conjugate. The goal is to find the complex scalar s such that its product with \vec{f} corresponds to the real source wavelet at ω_k . The parameter ε (commonly refer to as *white noise*) stabilizes the equation if the denominator is very small, which helps significantly in the presence of noise in the recordings (Forbriger et al., 2014).

The obtained source-wavelet correction filters remain unchanged within a frequency stage used during FWI and they are recalculated for each single shot when the frequency interval is increased. For the estimation of the source-wavelet correction filters only the near-offset traces per shot are used. However, the recordings in the vicinity of the source are ignored to avoid near-field effects, by applying the same offset windowing in all shot gathers. The source wavelet after application of the correction filter is the optimum source wavelet, since the residuals between synthetic and recorded data are minimized in a least-squares sense.

The source-wavelet correction filter estimation is required at the first iteration of every new frequency interval, due to the fact that both the model parameters and data content have been significantly changed and can therefore no longer be characterized by the initial source-time function. Additionally, the updated source-time function is required due to the presence of \mathbf{S} in equation 5.2, which shows that inaccuracies in the parameter space will reflect negatively on the estimated source wavelet.

5.5 Summary

The methodologies described in this chapter serve as a complete guide of preparing seismic field data for any inversion scheme and are not restricted to FWI applications. They constitute essential steps of the inversion workflow which otherwise would most likely fail to obtain realistic results. The various parameters that were described above have to be estimated heuristically, as they are entirely dependent on the acquired data and geological features of the subsurface. In the next two chapters, I will not repeat the methodologies described here. I will, however, give the parameter values/ranges that are necessary for recreating the same results.

Chapter 6

Field survey: archaeological investigation

6.1	Test site	86
6.2	Acquisition geometry and pre-processing	86
6.3	Initial models and FWI set-up	88
6.4	Results	89
6.4.1	Conventional versus time-frequency FWI	89
6.4.2	Inline P-SV inversion	93
6.4.3	SH inversion	95
6.5	Comparison with ground-penetrating radar	96
6.6	Summary	98

*Part of this chapter was published as: Athanasopoulos, N., Manukyan, E., Bohlen, T., Maurer, H. (2020), Time–frequency windowing in multiparameter elastic FWI of shallow seismic wavefield, *Geophysical Journal International*, **222**(2), 1164–1177, DOI: 10.1093/gji/ggaa242.*

In this chapter, I show the application of FWI in a field survey of archaeological interest. The survey is acquired on a glider airfield near the city of Karlsruhe (south-west Germany). The area has been extensively studied previously in order to identify the remnants of a trench-like structure which was initially constructed in the 18th century. In the beginning, I will introduce the test site (Section 6.1) and the acquisition set-up (Section 6.2), following with application of the methodologies introduced in Chapter 4. More specifically, the focus will be on the application of the TF-FWI approach to improve the resolution of the P-wave velocity (Section 6.4.1) using 2D FWI of Rayleigh waves. The knowledge of P-wave velocity in the test site is so far poor, since CFWI that has been previously performed, failed to retrieve an accurate P-wave velocity model. The SH FWI (Section 6.4.3) is further used to evaluate the results of the TF-FWI of P-SV waves. The results from the various inversions are discussed. Finally, I compare the findings with the results obtained from independent ground-penetrating radar measurements (Section 6.5).

6.1 Test site

The survey area is located on a glider airfield in Rheinstetten near Karlsruhe (Germany, Figure 6.1). The geology of the area was investigated by Hüttner et al. (1968), where they state that the subsurface consists of layered fluvial sediments of the late Pleistocene. Previous studies also performed in the area by Groos (2013), Groos et al. (2017) and Schäfer (2014) revealed a predominantly depth-dependent 1D subsurface. Also located in the area lies a shallow low-velocity structure (remnant of a man-made trench), which according to historic recordings served as a defensive line in the early 18th century and is known as the «Ettlinger Linie» (Lang, 1907). It was originally excavated to serve as a line of defence and was refilled several decades ago. In the wider area, the «Ettlinger Linie» is exposed without any deposition of soil. Figure 6.1 (blue line) shows the assumed location of the trench, by combining the north-west and south-east exposed traces of the trench. The low-velocity trench superimposes the lateral homogeneous subsurface locally, which validates the assumption of a 2D subsurface when the acquisition profile is set perpendicular to the trench (Figure 6.1, orange line).

As I discussed in Chapter 4 the results of the inversion of Rayleigh and Love waves showed that an inaccurate v_p model will lead to artefacts and inaccuracies in the reconstructed S-wave velocity model. The conclusion of this study was that Love wave inversion is more robust to accurately retrieve the S-wave velocity of the shallow subsurface, due to the fact that it is not sensitive to the P-wave velocity structure. However, since TF-FWI can improve the resolution of the P-wave velocity, I will focus on the applicability of the approach on this survey in order to get improved resolution of not only v_s but also v_p and density.

6.2 Acquisition geometry and pre-processing

The profile of the seismic survey crossed the path of the trench perpendicularly, as shown in Figure 6.1. The orientation of the profile is from north-east to south-west. The acquisition geometry consists of a linear profile of 48 3-component geophones (Geospace Technologies GSC-11D, eigenfrequency 4.5 Hz) with an equidistant spacing of 1 m. The data were acquired sequentially due to limited hardware, i.e., first only the vertical component, then only the horizontal inline and finally the horizontal crossline. The

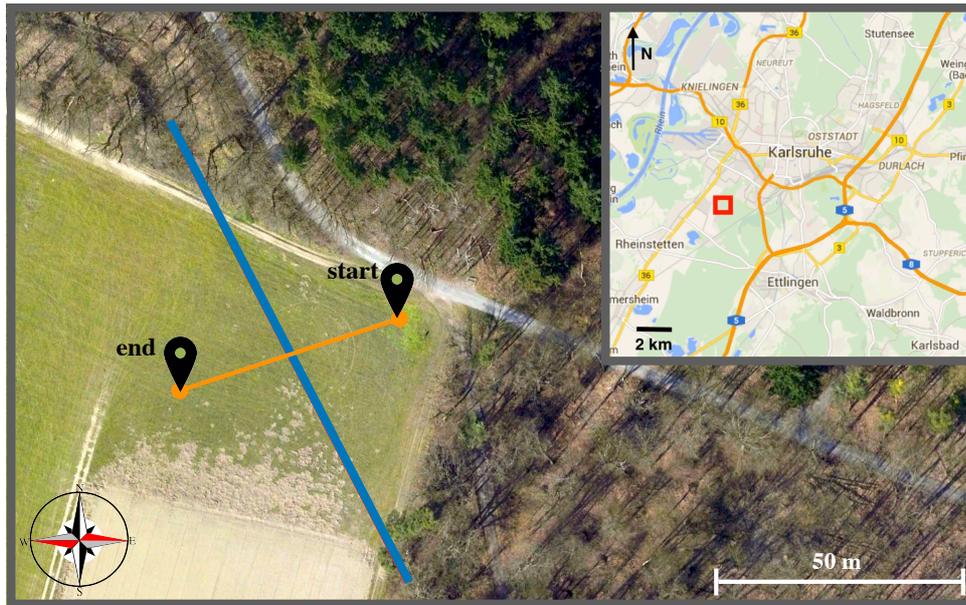


Figure 6.1: Map of the survey area on the glider airfield in Rheinstetten. The blue line corresponds to the interpolated path of the "Ettlinger Linie" and the orange line shows the acquisition profile. Start: N 48°58' 52.59", E 8°20' 46.73". End: N 48°58' 52.12", E 8°20' 44.57". Source: Google Earth (AeroWest, GeoBasis-DE/BKG).

sources are vertical hammer blows on a steel plate (24 in total) with an equidistant spacing of 2 m for the P-SV components, while a steel beam was used to generate shear force for the Love wave acquisition. The model space consists of 700 grid points in the horizontal direction and 200 grid points in the vertical direction, resulting in the actual dimensions of 70 m x 20 m (grid spacing is 0.1 m).

Since the actual source is unknown I estimate the source-time function by calculating a correction filter as discussed in Section 5.4. To avoid computation costs I update the source-time function only at every stage of FWI and not at every iteration. As mentioned previously, the estimated wavelet does not necessarily describe the actual force excited in the field measurement and can therefore be affected by residuals caused by an inaccurate parameter model. This provides us with an additional quality check for the inversion as coherency between the different source wavelets at different frequencies indicates an accurate reconstruction of the velocity model (Brenders, 2011). In the case of TF-FWI both synthetic and observed data are tapered prior to the calculation of the correction filter.

I recorded 1.5 s using a temporal sampling of $2.5 \cdot 10^{-4}$ s. The data were stacked with a fold of five to enhance the signal-to-noise ratio. I kept only up to 0.5 s, where the main signal was recorded and additionally upsampled the data to a sampling of $1.4 \cdot 10^{-5}$ s, in order to satisfy the stability criterion of the finite-difference forward solver. I picked the first arrival times and muted the energy prior to them. I zero-padded the signal adding 0.02 s in order to avoid non-causal effects in the inverted source-time function and then transformed the data from 3D to 2D. Finally, I normalized the seismograms trace-wise and applied a band-pass filter between 4 Hz and 120 Hz. The details of the procedures above were thoroughly explained in Section 5.2, along with any additional preprocessing steps. Figure 6.2 shows the data after preprocessing for each of the three components. Similar to Figure 5.1, the dominant energy of the recordings originates from the surface waves. The refracted waves are almost invisible in the presence of the high-amplitude surface waves and the converted waves in between. Such a scenario is ideal for the

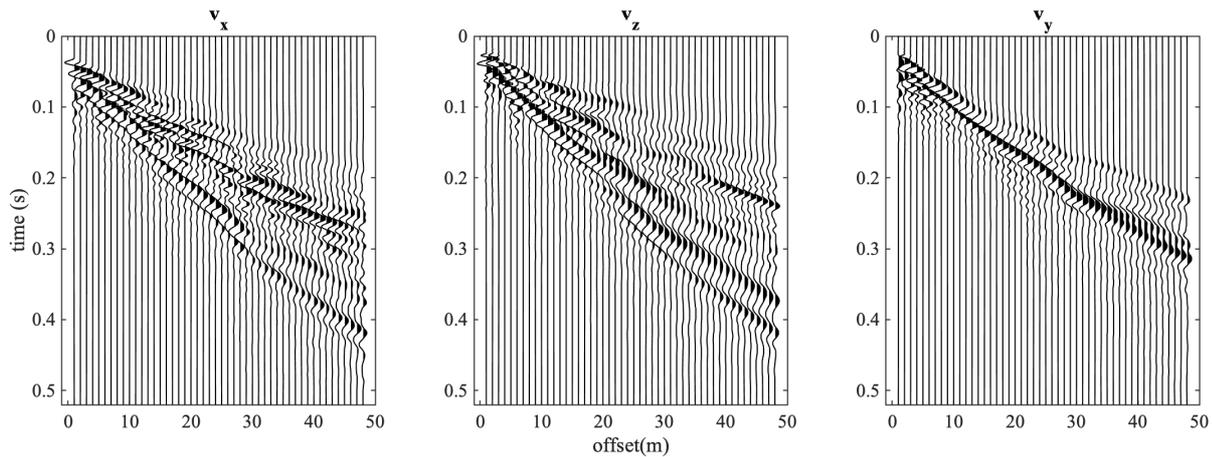


Figure 6.2: Trace-wise normalized shot-gather of the field seismic data. The corresponding source is located at the profile coordinate of 0.5 m. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH-waves. The gathers in both cases are dominated by the surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.

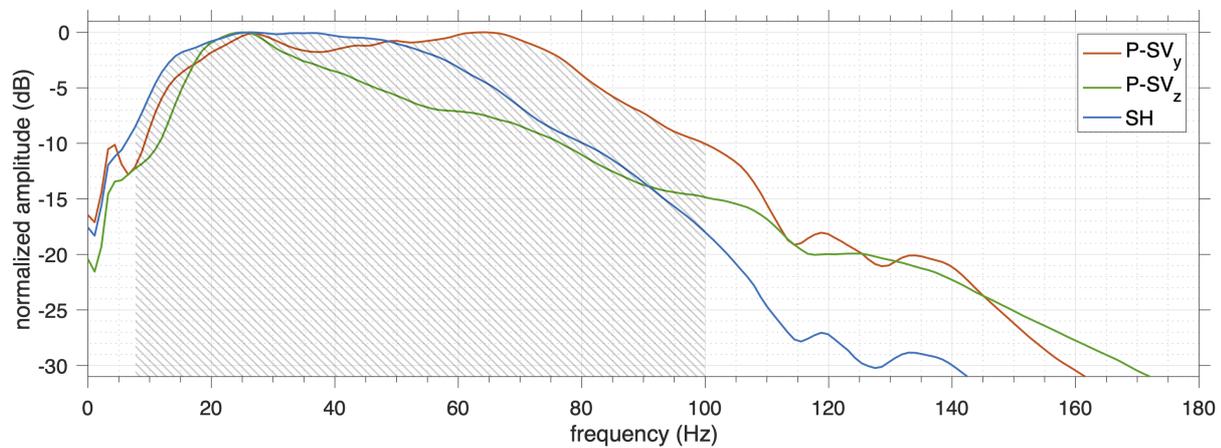


Figure 6.3: Amplitude-frequency spectra for each component of the field data set. The spectrum of the P-SV waves is shown in red and green for v_x and v_z components, respectively, and of the SH waves in blue. The shown spectra are the average of the spectra of normalized traces. The hatched area corresponds to the target frequency band which was used during FWI.

TF-FWI, since the exponential damping that is applied will enhance the refracted waves and allow them to be properly evaluated during the calculation of the misfit function.

In terms of frequency content, in Figure 6.3 I show the mean amplitude spectra as a function of frequency, measured by the three different components. There are no big differences between the three spectra, with an exception of the spectrum obtained from the horizontal inline component (red line) data. The main frequency content is located between 8 Hz and 100 Hz, which will be used during FWI.

6.3 Initial models and FWI set-up

I pick the first arrivals and perform basic P-wave traveltimes analysis (Section 5.3.1) to obtain an initial model for v_p (top layer 335 m/s, lower half-space 2284 m/s). Then I transform the data to FK domain and pick the dominant mode of the Rayleigh wave. I use these data points to perform a 1D MASW (Section 5.3.2) and obtain a 1D v_s model (140 m/s to 340 m/s up to 9 m). I extrapolate the values of the

1D model to obtain the final 2D models that will be used as initial models for FWI. Finally, I build an initial model for the density (top layer 1657 kg/m^3 , lower half-space 2143 kg/m^3) through a combination of previous knowledge of the geology in the area and empirical relationships for the shallow subsurface. Figure 6.4 (first column) shows the values of the initial models. In all three models the groundwater table discontinuity is estimated at 6.1 m, followed by a homogeneous layer.

As described in Section 5.3.3, I assume that the Q-values for P-waves and S-waves are identical and that a constant Q-value is sufficient for the whole model space (Groos et al., 2017). Then, I approximate a Q-value by comparing the misfit between synthetic and observed data for different Q-values. The resulting values of the local grid-search returns a $Q = 15$ and I then proceed by constructing the attenuation model by a GSLS with three relaxation mechanisms. The calculated relaxation frequencies are $f_1 = 0.29 \text{ Hz}$, $f_2 = 6.73 \text{ Hz}$ and $f_3 = 84.6 \text{ Hz}$ and the τ -value is 0.1576, where I use a reference frequency of 40 Hz.

After estimating the initial models I calculate the synthetic data and perform a quality control to verify that no cycle-skipping is present, as shown in Section 5.3.4, and correct accordingly (Köhn et al., 2019). The same preconditioning as for the synthetic reconstruction test (Section 4) was applied to the field data. This includes circular tapers around the source, the approximation to the diagonal elements of the Hessian and the application of a 2D median filter to the gradients, where the filter has a size of 0.8 m (8 grid points).

6.4 Results

In this section, I show the results from the different inversion types. Initially, I examine the vertical component recordings and compare CFWI with TF-FWI (Section 6.4.1). The reason I do not use both vertical and in-line recordings is the poor signal-to-noise ratio of the in-line recordings and the inability to couple the two data sets, potentially caused by the fact that they were not acquired simultaneously, as mentioned above. An inversion using both P-SV components was not successful after several trials. I investigate the reasons behind this issue in Section 6.4.2. In section 6.4.3, I show the results of SH-FWI and finally compare the different approaches with the results of the migrated image from ground penetrating radar measurements that have been performed on site.

6.4.1 Conventional versus time-frequency FWI

I investigate the results of both CFWI and TF-FWI and compare them both in model and data space (Figures 6.4 and 6.5, respectively). The workflow that I apply is shown in Table 6.1. In case of the CFWI γ is set to zero. In the first two stages of TF-FWI I update only the v_p model since I consider the refracted waves. In the first stage of CFWI, both the v_p and the v_s model is updated. The inversion of density in both cases is delayed to avoid cross-talk from the velocity updates at early stages (Section 4.1.5).

Conventional FWI

In the final v_s model of CFWI (Figure 6.4, column 2) I observe the trench centered at around 35 m. The S-wave velocity is around 120 m/s and the trench on the v_s model has mainly a triangular/circular shape.

Table 6.1: Workflow used for FWI of field data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass frequency filter. The parameter ' γ ' indicates the damping factor that is used during the corresponding stage (column 'Stage') in equation 4.3. The first two stages were only applied in TF-FWI for the v_p model.

Stage	Update			LP in Hz	γ
	v_P	v_S	ρ		
1	1	0	0	10	100
2	1	0	0	10	40
3	1	1	0	20	20
4	1	1	1	20	10
5	1	1	1	30	5
6–11	1	1	1	increment of 10	0

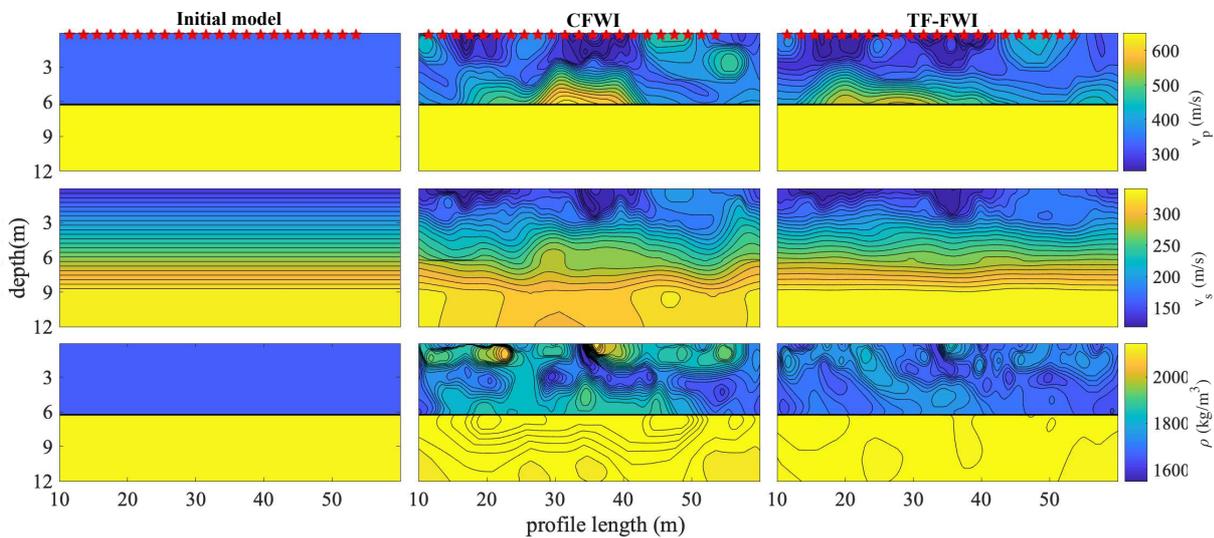


Figure 6.4: Models obtained by the inversion of field data. The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of CFWI and TF-FWI, respectively. Red stars represent the source locations.

It extends laterally 6 to 7 m and vertically approximately 2.6 m. At around 5.5 m depth and 28 to 42 m horizontal location the lateral coherency of the velocity is not consistent which is not expected since also it is located directly above the discontinuity of 6.1 m related to the groundwater table. Indeed, this structure is present in both v_p and density model and seems to be some cross-talk from inaccuracies of the model parameters. Another low-velocity anomaly is present close to the surface at around 10 to 20 m offsets which could potentially correlate to either looser sediments or high water saturation. At deeper parts of the profile the v_s model shows some variations below the groundwater table at 6 m depth which could signify inaccuracies of the velocity updates.

The v_p model shows some velocity variations which correlate only partly to the v_s model. In particular, although the velocity in the location of the trench is reduced, there is no good correlation to either the length or shape of the trench. Additionally, a higher increase of the v_p model below the trench causes

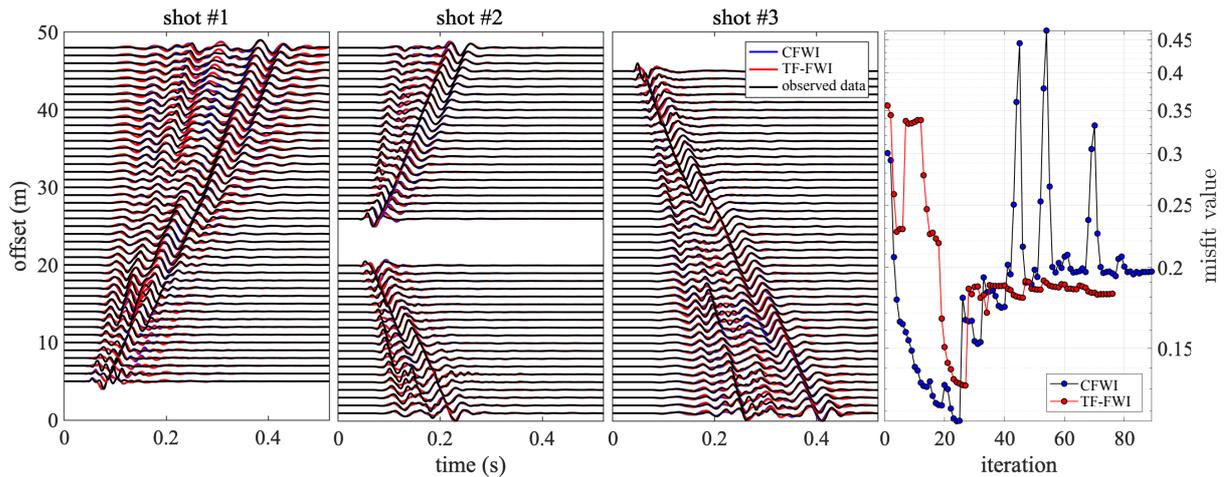


Figure 6.5: Comparison of the observed data with the data obtained using the final models of CFWI and TF-FWI shown in Figure 6.4. Comparison of velocity seismograms for three different shots at three locations. The misfit function for the CFWI and TF-FWI approach is shown at the fourth column.

artifacts to leak into both the v_s and the density model, similar to what I observed in the synthetic study. The low-velocity layer on the left of the trench is also present in v_p but with a different shape.

The density model is contaminated with many artifacts and its reliability is quite low due to the unrealistic absolute values. There is a one-sided structure at the location of the trench but this is interpreted as the result of cross-talk from v_s , which has been demonstrated in the synthetic study.

Time-frequency FWI

The v_s models of TF-FWI (Figure 6.4, column 2) and CFWI are very similar. The model shows less artifacts below the trench compared to the v_s from CFWI, maintaining a higher homogeneity overall. Our knowledge from previous studies indicates that there are no severe lateral heterogeneities, which comes in agreement with the results of TF-FWI.

The v_p model seems to be improved in the case of TF-FWI. It exhibits a high correlation with the v_s model structure. The trench now has a more triangular shape better matching the position of the shallow anomaly in v_s . The low-velocity structure on the left is also present. Similarly to the conventional, approach a slight increase of the v_p values on the right is observed, which could indicate higher water saturation.

The density model is smoother compared to the density retrieved from CFWI. It shows less local anomalies with weaker fluctuation in absolute values. We observe lower density values at the location of the trench. In the synthetic study I have shown that absolute density values can be slightly improved by TF-FWI. This can also be observed from the scattering signature of density (Section 3.2), which revealed to be quite distinct compared to both P and S waves, indicating as well that both models have to be accurate enough to produce a reliable density model. The above observations favor the density model retrieved from TF-FWI the most.

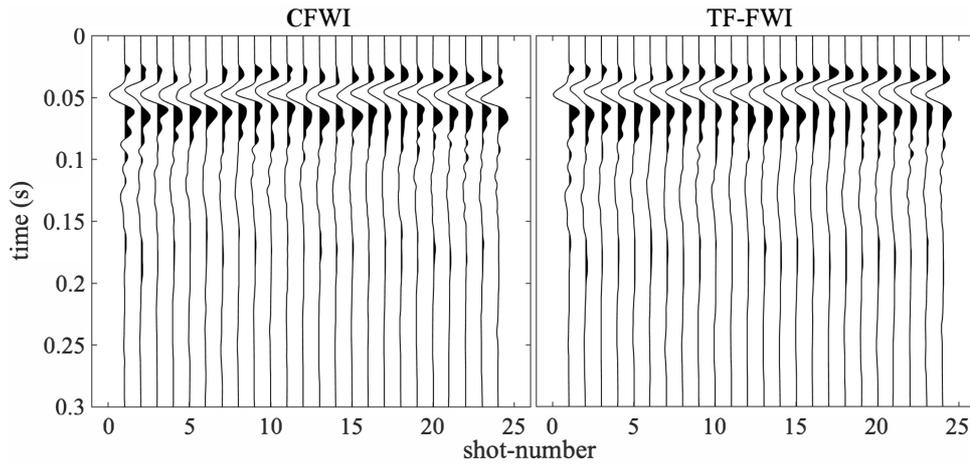


Figure 6.6: Estimated wavelets for the source-time function of CFWI and TF-FWI. The consistency between the inverted wavelets indicate that both inversions were able to provide reliable models, concerning the main subsurface features.

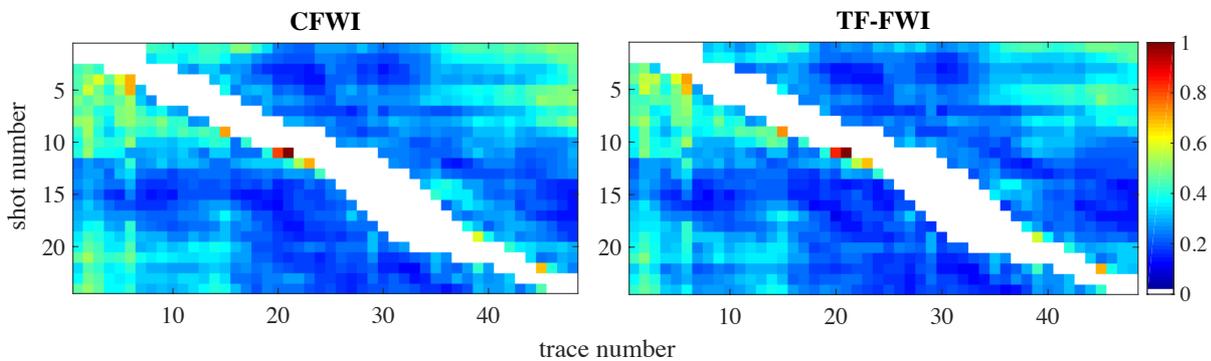


Figure 6.7: Data misfit of single shots and single receivers for CFWI and TF-FWI. The values are normalized by the maximum amplitude of the misfit in each case. The removal of the near-offset traces is indicated by the white area. These traces are not considered within the misfit calculation and in the inversion.

Comparison in the data-space

I compare the seismograms of three shot gathers at three different locations along the profile at the last frequency step (4-100 Hz) of the vertical component for both approaches in Figure 6.5. In both cases the fit of the data is satisfactory. Some differences can be observed at far offset traces where the refracted and Rayleigh waves have a slightly better fit in the case of TF-FWI. This was also the case in the synthetic study, but the differences here are much smaller. The misfit function allow us to better estimate the mismatch in this case. Figure 6.5b shows that the residual energy is further reduced using the TF-FWI. The misfit function of the CFWI exhibits stronger fluctuations at higher frequencies. In Figure 6.6 I compare the inverted source-time functions for every shot for the two methods. In both cases we observe similar wavelets for all source locations, indicating that both methods arrived at a consistent solution considering, however, that the source time functions are dominated by the Rayleigh waves and the v_s model. Finally, in Figure 6.7 I show the normalized misfit as a function of source and receivers, and we observe that there is a very similar pattern in terms of where the highest discrepancies between the synthetic and observed data lie. However, there is a slight reduction of the misfit for the far-offset traces, which aligns with the fact that the fit of the refracted waves has been improved with the TF-FWI.

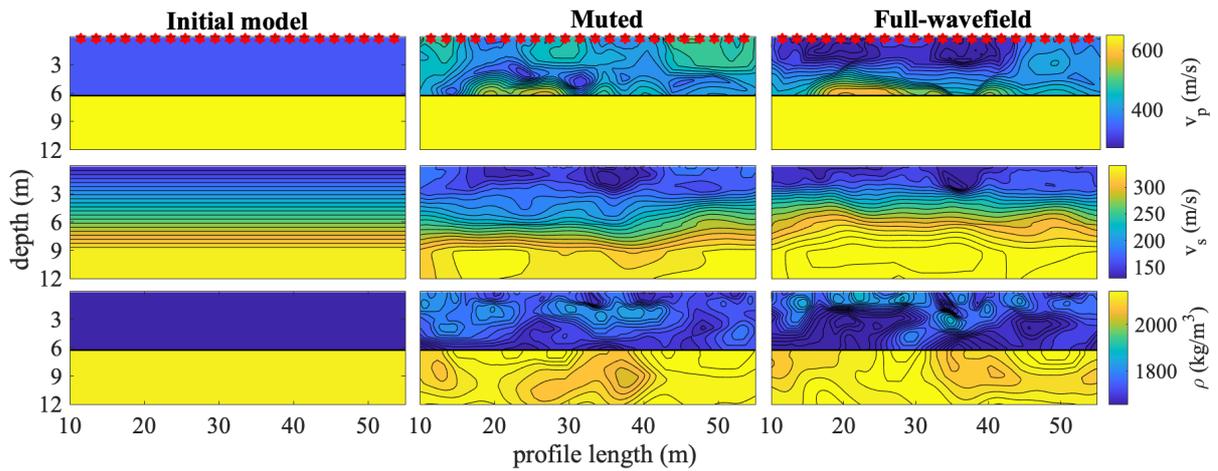


Figure 6.8: Models obtained by the inversion of field data. The rows of the Figure represent v_p , v_s and density models, respectively. The columns represent the initial model and the final results of P-SV FWI using only the v_x component, for the case that a strong mute is applied and the case that the full wavefield is considered for the inversion. Red stars represent the source locations.

6.4.2 Inline P-SV inversion

In this section, I show the results of P-SV FWI using only the horizontal in-line component. The poor signal-to-noise ratio in this case, along with the existence of noise around the area of the refracted waves makes the choice of inversion strategy very complicated. I proceed using the same setup as used for TF-FWI shown in Table 6.1. CFWI did not converge as the misfit function could not be reduced. At this point, I decided to create two cases: a) Application of strong muting around the first arrivals (along with the possibility of removing a significant part of their energy), and b) using the full wavefield without proper removal of the noise in the early arrivals (which creates inaccuracies at the early stages of the inversion, where a high γ also boosts noise).

The inverted models for both cases are shown in Figure 6.8, along with the initial models. The results using the full wavefield are in general more consistent with the previous findings (using the vertical component). The trench in v_s has a more circular shape but similar values and location. The low-velocity structure at the top left of the model is also present. The v_p model shows similar behavior, i.e., a smoothed version of the results from the vertical component, where the shapes are different but the location of the structures and their values are similar. The density shows few updates around the expected location of the trench.

The models resulting from the application of a strong mute show many deviations, especially in the v_p and density model. There is no indication of a trench in v_p , which shows the importance of the refracted waves for the accuracy of the retrieved P-wave velocity, as expected. In the v_s model, although the trench is visible, its dimensions deviate from the results of the vertical component (previous section). These results align with the synthetic findings, where I saw that inaccuracies in the v_p model will affect the reconstruction of the v_s model.

The data comparison is shown in Figure 6.9. However, the effect of muting is not visible here since I only show the results of the last iteration in this figure (where the value of γ is down to zero, see Table 6.1). In both cases FWI fitted the main phase of the fundamental surface wave mode. There are significant

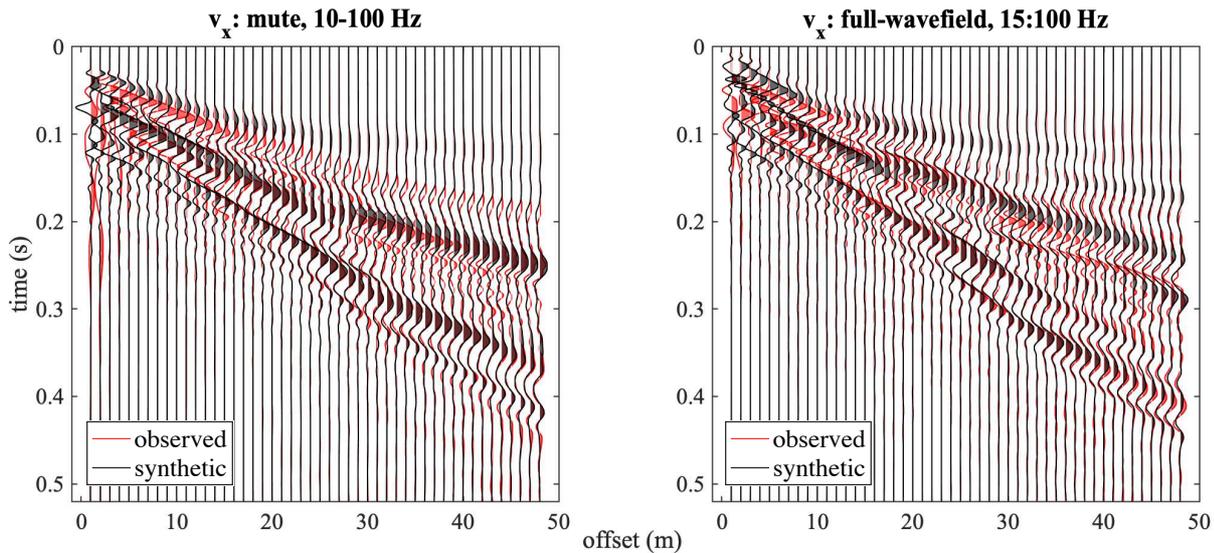


Figure 6.9: Comparison of the observed data with the data obtained using the final models of the two cases shown in Figure 6.8.

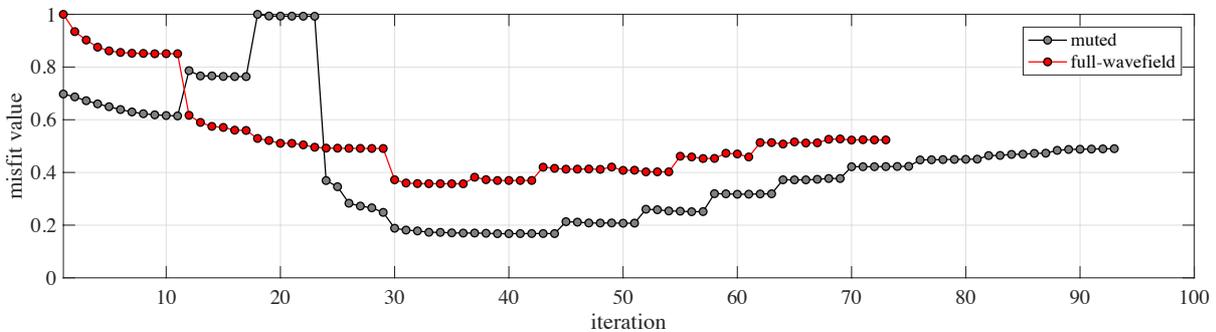


Figure 6.10: Comparison of the misfit function for the two cases of P-SV FWI using only the v_x component.

mismatches along the area between the coda of the first arrival and the fundamental Rayleigh wave mode. The misfit function is similar in both cases (Figure 6.10), however a direct comparison is not feasible due to the difference in the content of the data.

Finally, I show the inverted source wavelets for both cases (Figure 6.11) for the first and last inversion stages. It is clear that the source wavelets have many differences for the different locations in both cases. While in the first stage of each respective case the wavelets show some similarity (with some exceptions), it is obvious that there are inconsistencies. These can be related to either random field noise during acquisition, coupling issues of the sources (unlikely to be on the receiver side, since they were not relocated after the acquisition of the vertical components), subsurface variations causing out-of-plane energy (3D effects) recorded from the inline polarized receivers or a combination of all of the above. This final comparison shows that combining the v_x data with the v_z data is not feasible and therefore I focused only on the v_z data for the P-SV FWI, as opposed to what I did in the synthetic study (Chapter 4).

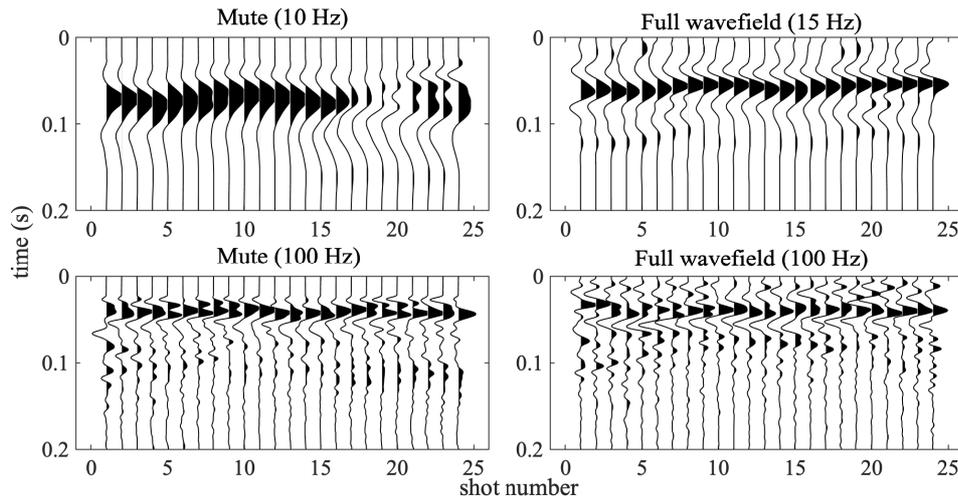


Figure 6.11: Estimated wavelets for the source-time function of v_x data for the case where a strong mute is applied and when the full wavefield is considered for the inversion, from the first (top) and last (bottom) inversion stages.

6.4.3 SH inversion

In this section, I use the recordings from the cross-line component to perform SH FWI. As mentioned before, the cross-line data set is dominated by the high-amplitude Love waves. Only two parameters are considered, namely the S-wave velocity and the density.

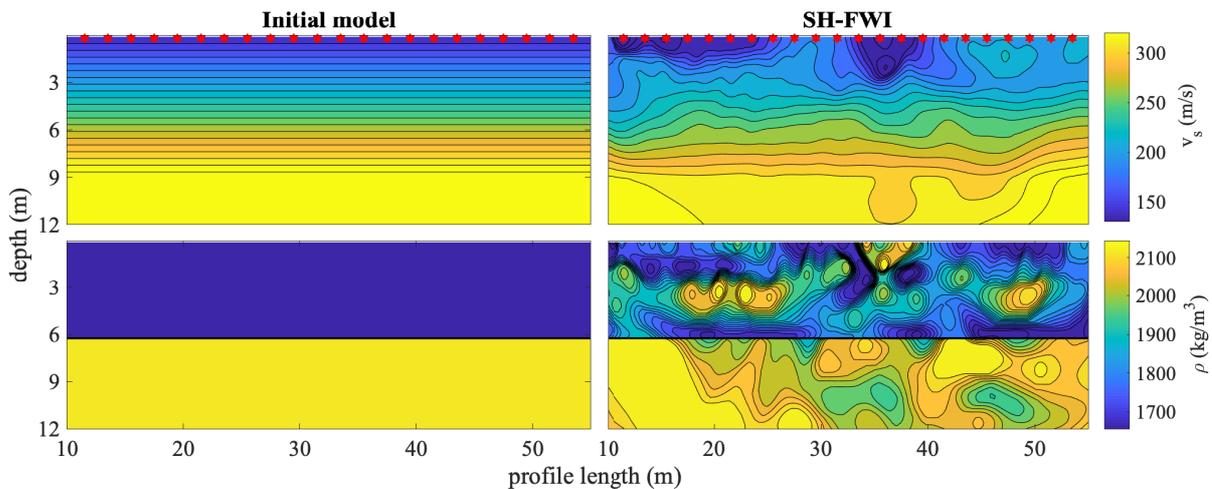


Figure 6.12: Models obtained by the inversion of field data. The rows of the Figure represent v_s and density models, respectively. The columns represent the initial model and the final results of SH FWI. Red stars represent the source locations.

The final models from the inversion are shown in Figure 6.12 along with the initial models. The v_s model is also here depth-dependent with small lateral variations across the profile. The trench is retrieved at the same location as in P-SV FWI and has similar dimensions, shape and values. It appears to have a more triangular shape and spans over the same length and depth as the previous inversion results. The low-velocity anomaly at the top left of the v_s model is also present. The density model shows a lot of small-scale variations, compared to the P-SV inversion results. The outline of the trench is present, but overall the model seems to suffer more from inversion artifacts. This can be related to the fact that there

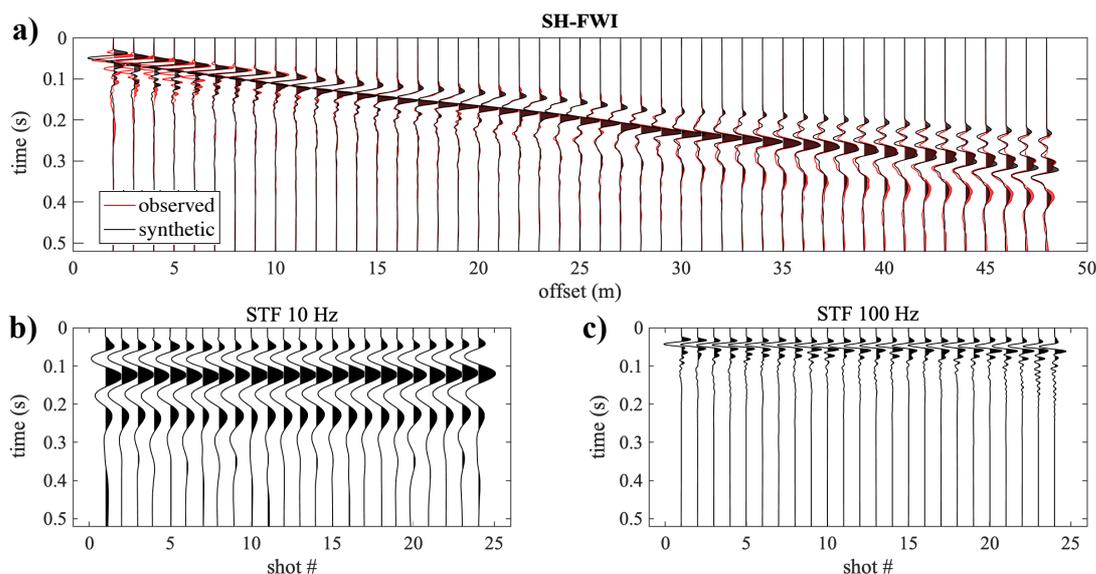


Figure 6.13: Results of SH FWI, corresponding to the models shown in Figure 6.12. a) Comparison of observed and synthetic velocity seismograms at the first shot location. b) Inverted source-time function at the first iteration. c) Inverted source-time functions at the last iteration.

is not significant information in the cross-line component to properly account for the density. Although these findings regarding the density do not align with the synthetic results, I have already expressed that obtaining reliable density model (especially in field data applications) from conventional FWI approaches, is a very difficult task. The combination of accurate P-wave and S-wave velocity, which was obtained through P-SV FWI, seems to provide more stable updates in the density model.

In Figure 6.13a, I show the comparison between the observed and the inverted data for the models shown in Figure 6.12. The datafit for both short and far offsets is exceptional (except for traces near the location of the source, which are not included in the inversion). All the phases of the various wave types have been fitted very well, signifying the successful application of SH FWI. Additionally, this can be observed by the consistency between the inverted source wavelets for all different source locations, both at the first inversion stage and at the last stage (Figure 6.13b and 6.13c, respectively).

The joint inversion of P-SV and SH FWI was part of the study presented in Wittkamp et al. (2018). Although, I repeated the joint inversion with my own setup, the results obtained are in high agreement with Wittkamp et al. (2018) for S-wave velocity reconstruction. The improvements shown here focus mainly on obtaining a more reliable P-wave velocity and by extent also density model by using the TF-FWI approach (since the S-wave can be more accurately retrieved by CFWI), which has not been investigated in that study.

6.5 Comparison with ground-penetrating radar

In order to evaluate the reliability of the FWI results with independent geophysical data, I compare them with the result of a ground-penetrating radar (GPR) measurement, similar to Wittkamp et al. (2018). For the zero-offset GPR measurements a 200 MHz antenna was used. The data were time-migrated using a constant-velocity Kirchhoff migration with a velocity of 0.1 m/ns. I applied a velocity factor of 0.086 m/ns

to scale the GPR traces from the time-domain to the depth-domain. In the migration image the trench is visible by boundary reflections which mainly consist of a triangular form. Inside the trench there are only a few reflections, suggesting a rather homogeneous filling of the trench. Several small-scale reflectors are present throughout the migration image.

I used the GPR image and superimpose it on top of the v_s and v_p models from CFWI and TF-FWI in Figure 6.14. In the GPR image the shape of the trench is underlined by the existence of strong reflections which form a triangular shape. The reflectivity inside the trench is weaker compared to the surrounding, indicating a rather homogeneous filling.

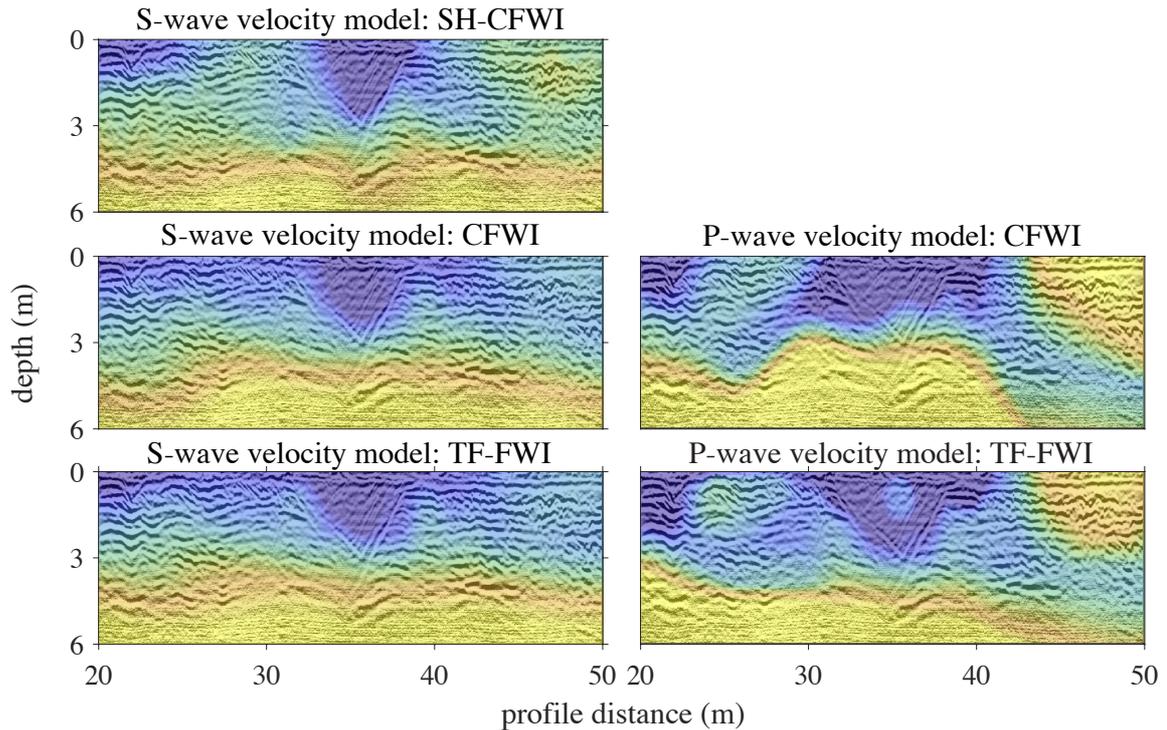


Figure 6.14: Qualitative comparison of the GPR result with the field data FWI results. Overlay of the final P- and S-wave velocity models of CFWI and TF-FWI of P-SV waves (middle and bottom row, respectively) and CFWI of SH waves (top row) with the time-migrated image of the GPR measurement. A color-bar is not shown, because the transparent overlay affects the color representation.

Both CFWI and TF-FWI produce low-velocity anomalies that fill the boundaries seen in the GPR image of the trench quite accurately. TF-FWI shows a more homogeneous filling with similar velocities in the shallow part at the left side of the trench. The horizontal extension of the trench is matched almost perfectly by both methods. The final v_p models exhibit clear differences between the two methods. The CFWI failed to properly match both horizontal and vertical extensions of the trench, and its shape. TF-FWI matches the trench reflectors quite well. Both horizontally but mainly vertically, the inclined reflectors fit the velocity contrast showing up in the v_p model. For further evaluation of CFWI and TF-FWI with other independent seismic information, I include in Figure 6.14 the final model of a CFWI using SH waves. The final v_s model from FWI of SH waves agrees well with the models derived by FWI of Rayleigh waves. Results from the joint inversion of P-SV and SH FWI were part of Wittkamp et al. (2018) and will not be presented here.

6.6 Summary

In this chapter I presented the application of 2D elastic full-waveform inversion of Rayleigh and Love waves to a near surface survey, with the aim to recover the subsurface information on an archaeological buried structure. The P-SV data set is dominated by a Rayleigh wave and the SH data set is dominated by a Love wave, both of which dominate the misfit function used in FWI. I applied the TF-FWI methodology to the field data set where I was able to reduce the misfit function and retrieve models of high correlation (compared to the CFWI). Compared to previous studies, the v_p for the first time in short-offset acquisitions could be reliably retrieved. The density model also was improved, where the TF-FWI approach allowed the reduction of artifacts which comes in agreement to the synthetic studies. This was also the case when superimposing the v_p and v_s with the migrated results of a GPR measurement that was performed on the site. As a last step to assess the reliability of the results, I performed the FWI of SH waves and obtained S-wave velocity structure of high similarity between the various methods, indicating (along with the GPR results) the successful implementation of FWI in the target area. I was not able to use the inline component during a joint P-SV FWI (including also the vertical component data simultaneously), since it contaminated the inversion results. However, I have shown that even in this case we can still retrieve valuable information for the elastic parameters. After evaluating the performance of the TF-FWI in both synthetic and field studies, I can highly recommend it in applications regarding near-surface seismic investigations. In the next chapter, I will further expand the time-frequency windowing approach by including it into a multicomponent multiparameter FWI framework.

Chapter 7

Field survey: aquifer characterization

7.1	Introduction	100
7.2	Study area	101
7.3	Prismatic source characterization	102
7.4	Acquisition set-up	105
7.5	Field data analysis and FWI parametrization	107
7.5.1	Preprocessing of field data	107
7.5.2	FWI set-up	107
7.6	Results	109
7.6.1	FWI results of profile 1 and 2	110
7.6.2	Evaluation of FWI density model with CPT measurements	116
7.6.3	Cross-section of P1 and P2	117
7.7	Summary	117

*Part of this chapter is in preparation for publication as: Athanasopoulos, N., Bohlen, T., Klotzsche, A. (2020), Aquifer characterization using multiparameter multicomponent viscoelastic full-waveform inversion, submission at *Geophysical journal international*.*

In this chapter the benefits of elastic full-waveform inversion (FWI) are exploited as an effort to improve aquifer characterization for better understanding of its properties and delineating its structure. Aquifer systems are complex near-surface targets, making their evaluation, especially when non-invasive geophysical methods are used, a challenging task. These sedimentary deposits are often composed of several distinct facies with strong stratigraphic boundaries that exhibit large variations in terms of subsurface properties, e.g., grain size, porosity or hydraulic conductivity. After an introduction, where I state the importance of studying aquifers (Section 7.1), I present the test site (Section 7.2). The aquifer is located in north-west Germany at the test site Krauthausen (Döring, 1997). With more than 30 years of various studies performed at this test site, the aim of this chapter is the addition of further information on the structural architecture of the aquifer using seismics for the first time. A specific acquisition setup is designed in order to optimize the acquisition time as well as facilitate the application of a multicomponent near-surface seismic study of such a demanding subsurface system (Section 7.3 and 7.4). The results of the joint inversion of Love and Rayleigh waves for this multicomponent study are then shown in Sections 7.5 and 7.6. Finally, the findings are compared with core penetrating tests from co-located measurements provided by colleagues of the research center in Jülich.

7.1 Introduction

A detailed characterization of aquifers strives to improve the prediction of groundwater flow and to evaluate possible contaminant hazards. Flow and transport processes in aquifers are mainly controlled by two parameters, the hydraulic conductivity and the porosity, where the first is the parameter exhibiting the highest variations (Dagan, 1989). Conventional methods (Döring, 1997) either have limited spatial sampling volume but high resolution (e.g., boreholes), or they record an average response over a large area (e.g., pumping tests). The inherently one-dimensional nature of these methods cannot provide information about the lateral heterogeneities of such subsurface environments. As a consequence, the simulation of groundwater flow and solute transport in aquifers becomes a very difficult task.

During the last decade, many studies were performed in order to facilitate the application of 2D tomographic approaches to aquifers (Ernst et al., 2007; Doetsch et al., 2010; Klotzsche et al., 2013; Gueting et al., 2015) in order to enhance the spatial coverage without suffering from loss of resolution. These studies mainly used ground penetrating radar (GPR) and electrical resistivity tomography (ERT) methods to obtain spatially highly resolved mappings of aquifer heterogeneities, taking advantage of the two-dimensionality of the tomographic images provided by these methods. Along with GPR and ERT, shallow seismic applications are also important for hydrogeological site characterization (Doetsch et al., 2010). While they cannot directly provide information about the dynamic characterization of an aquifer system (e.g., porosity, hydraulic conductivity, transport flow), they can, however, be a reliable source of information for the structural architecture (e.g., bulk density, velocity distribution).

Therefore, rather than relying on explicit petrophysical relationships, in this chapter I will focus on the assumption that geophysical variations found in the retrieved models can help to delineate the lithological zonation, which is in turn related to hydrogeological properties. In this context, the difficulty of defining one specific quantitative relationship required for the direct estimation of hydraulic parameters from seismic data can be ignored.

In Chapter 6, I have shown the applicability of FWI for characterizing shallow targets of archaeological interest, using individual data from all three Cartesian components. The focus of this chapter lies on the joint inversion of simultaneously using the vertical, horizontal inline and crossline components, where both Rayleigh and Love waves are included. As I have shown in Chapter 4, the application of the joint inversion can significantly improve the FWI results and provide models for P- and S-wave velocity and density of high resolution. In order to avoid previous issues when different components were simultaneously included (Section 6.4.2), a different kind of source is used here which I will discuss further in Section 7.3.

7.2 Study area

The location of the survey is the Krauthausen test site, situated in north-west Germany between the cities of Jülich and Düren (Figure 7.1a). The test site was set up in 1993 by the research center Jülich. A detailed description of the test site is given by Döring (1997) based on laboratory characterization of core samples. The subsurface on top of the aquifer can be divided into three layers (Figure 7.1b): (a) a bottom layer composed of sandy to gravelly grain size (from 6 to 11.5 m depth), (b) a well sorted sand layer (from 4 to 6 m depth), and (c) a poorly sorted gravel layer (from 1 to 4 m depth). The base of the aquifer is formed by thin layers of clay and sand at approximately 12 m depth. The groundwater level shows variations from 1 to 3 m depth depending on the annual season. The aquifer material is composed of alluvial terrace sediments, deposited by the river Rur. At the shallowest depth of the groundwater table (<1.5 m), the aquifer is semi-confined, below and up to the depth of the groundwater table it is semi-unconfined. The clay and silt content of the aquifer sediments vary between 0.5% and 7.5%, while the mean total porosity is around 26% (with a standard deviation of 7%).

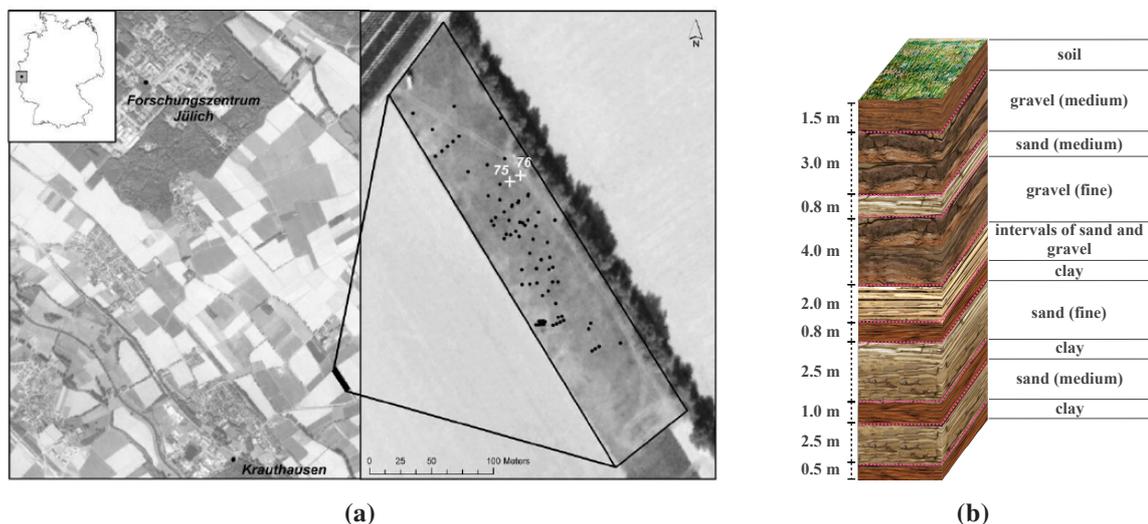


Figure 7.1: a) Overview of the Krauthausen test site in northwest Germany. Source: Döring (1997). b) A conceptual model of the geologic sequence for the aquifer in Krauthausen as studied by Döring (1997).

The description given above as well as the illustration depicted in Figure 7.1b represent only a simplified conceptual model of the aquifer as in reality the subsurface is much more complex, with high level of lateral heterogeneities.

Numerous studies were performed in the area with the goal to study the spatial distribution of aquifer parameters, including tracer experiments (Vereecken et al., 2000), cone penetration tests (Tillmann et al., 2008) and geophysical imaging methods (Klotzsche et al., 2013; Gueting et al., 2015). The true aquifer architecture exhibits lateral variations in layer thickness and properties and additional information from various geophysical methods is necessary for adequate characterization of these spatial distributions. Most recently Klotzsche et al. (2013) and Gueting et al. (2015, 2017) applied GPR full-waveform inversion and revealed the heterogeneous nature of the subsurface. The focus of my study lies on characterizing the aquifer in Krauthausen by applying seismic FWI and reconstruct the v_p , v_s and density models of the subsurface, as mentioned in the introduction.

7.3 Prismatic source characterization

In this section, I analyze the various impact sources which are used for shallow land seismic applications. In Chapter 5, I gave an overview of the different polarization directions of source-receiver pairs, which allow us to investigate specific wave types, depending on the polarization set-up. In the previous chapter, the field study included the use of two of the most common impact sources for shallow subsurface targets: A steel plate (Figure 7.2a) which allows us to retrieve P-SV waves by recording the vertical or horizontal inline component (on the receiver side), and a horizontal shear-beam (Figure 7.2b) which allows to retrieve SH waves by recording the horizontal crossline component.

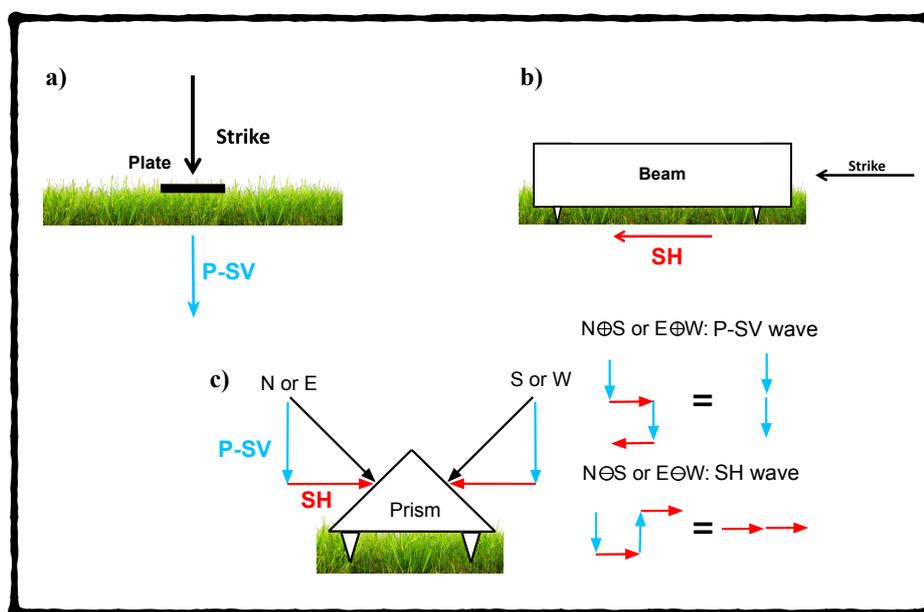


Figure 7.2: Sketch illustrating the various impact sources that were used in this study. a) Vertical strike on a steel plate used to record P-SV waves from vertical and inline-polarized receivers. b) Horizontal strike on a steel beam (which is coupled to the ground) used to record SH waves from subtraction of the signal from the two opposite sides from crossline-polarized receivers. c) Prismatic steel source that allows the separation into vertical and shear-polarized forces by either summation or subtraction, respectively, of the signal from the two opposite sides. The angle the prism source to the surface is 45 degrees. The main recorded wavetypes are highlighted for illustration purposes, however, more events can be produced depending on the source-receiver polarization that is applied.

Schmelzbach et al. (2016) suggested the use of a prismatic inclined source due to the advantage of simultaneously exciting of inline horizontally- and vertically-directed sources with uniform coupling. Additionally, by rotating the prism by 90 degrees it can also be used during the measurement of the horizontal crossline component. Thereby, we obtain the ground motion of all four azimuthal directions. The benefit of such an impact source is illustrated in Figure 7.2c. Summation of two opposite directed strikes leads to a cancellation of the horizontally-directed energy, which results in a vertically-directed source (x- and z-component). Subtraction results in enhancing¹ the horizontally-directed ground motion (y-component, given that the receivers are polarized horizontally in the crossline direction) and suppressing the vertically-directed energy.

An important advantage of this workflow using a prism source is the reduction of acquisition time. Acquiring all three components (x, y and z) is achieved by only one rotation of the source. Additionally, coupling issues are reduced for the same reason. I further modified the initial source design by increasing its overall volume for an easier deployment when used with a sledgehammer, while additionally increasing the coupling to the ground (size of the feet) in order to improve the horizontal recordings. The modified design is shown in Appendix C.

Next, I investigate the performance of the prism source compared to the conventional ones, similarly to Häusler et al. (2018). The comparison of the prismatic source with the conventional ones is done using reference data of profile 1 (for horizontal crossline recordings) and profile 2 (for vertical recordings) for the Krauthausen test site. The specific details of the acquisition set-up are discussed in the next section (Section 7.4). Here I focus on the direct comparison of the same polarized source-receiver pairs from different impact sources (Figure 7.2).

As a vector source, I used a sledgehammer in all cases. Hammer blows were struck normal to each of the two 45 degrees-inclined faces of the prism, and recorded separately. After rotation, another set of hammer blows on the two opposing faces of the prism were acquired. For the horizontal beamer, the hammer blows were performed on the two sides, leading to two separate recordings that were later subtracted, while for the steel plate a single perpendicular strike was necessary.

The resulting shot-gathers after separating the recordings into vertical and horizontal crossline components are shown in Figure 7.3a (v_z and v_y , respectively). For the evaluation of the recordings, 36 vertical component receivers and 48 horizontal crossline receivers were employed. All receivers were spaced at 1 m intervals. The general match of the the shot-gathers is great, with the vertical component being best. All wave types are present at same arrival times with no phase changes, but occasionally some amplitude differences. To better assess the similarity of the recordings, I plot the trace-by-trace normalized crosscorrelation for the two components (Figure 7.3b). The mean correlation for the vertical component is 0.9 (scale is 0 to 1, where one indicates perfect correlation) with a time difference of 1 ms, while for the horizontal crossline component the differences are higher at 0.73 mean correlation with 1.7 ms time difference.

I further compare the resulting amplitude spectra for the v_z and v_y recordings (Figure 7.4a and 7.4b, respectively). Overall, the spectra are similar. However, in the case of the vertical recordings, while

¹ By further adding or subtracting the signal, the component that remains can be further enhanced, due to the fact that it is a composite of two strikes polarized from both source and receiver side in the same way.

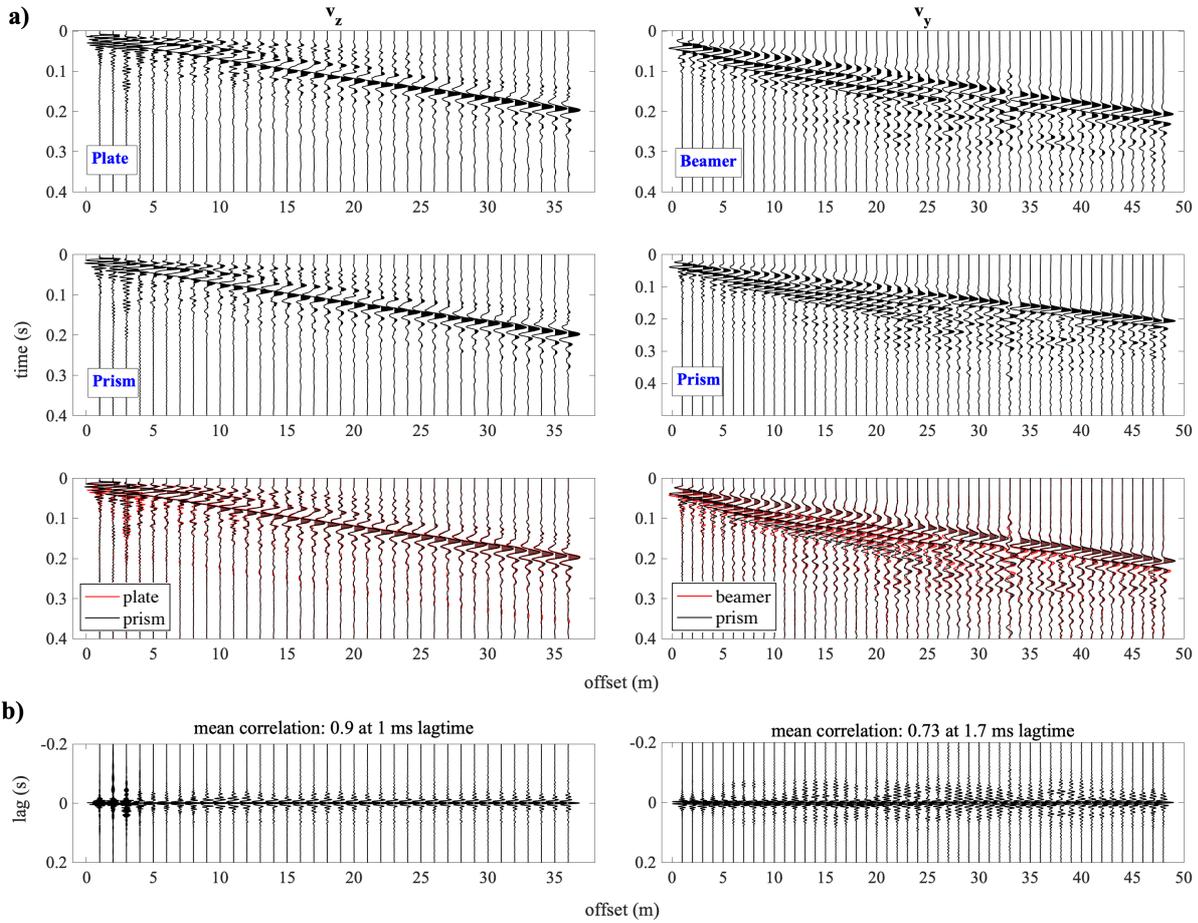


Figure 7.3: Reference shot-gathers induced by a steel plate (Figure 7.2a), a horizontal shear-beamer (Figure 7.2b) and the prismatic source (Figure 7.2c). a) The shot-gathers from the vertical and horizontal crossline recordings are shown (first and second column, respectively). b) Trace-by-trace crosscorrelation between the two source types of the same

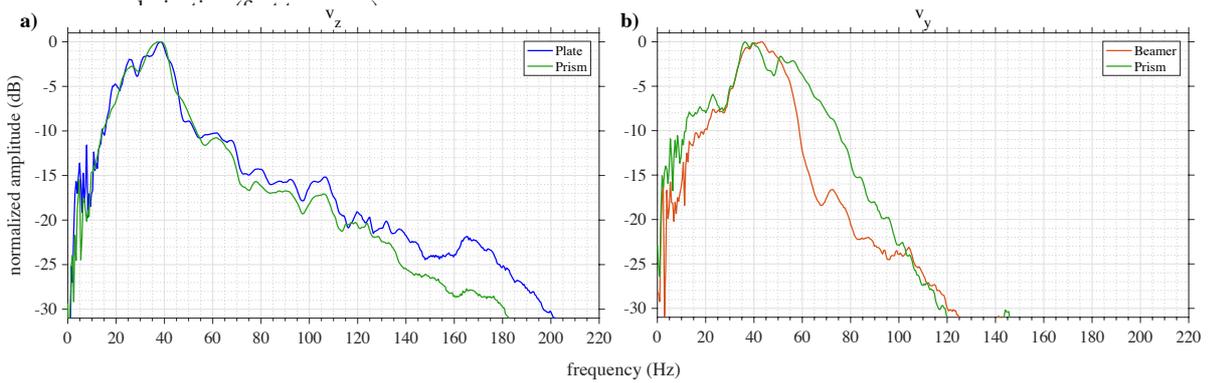


Figure 7.4: Amplitude-frequency spectra for each pair of sources with the same polarization. a) The spectrum of the P-SV recordings (v_z) for the vertical plate (blue line) and the vertical component of the prism (green line) obtained after summation of the two opposite impulses. b) The spectrum of the SH recordings (v_y) for the horizontal beamer (orange line) and the horizontal crossline component of the prism (green line) obtained after subtraction of the two opposite impulses (Figure 7.2c).

the relative changes over the frequency range are the same, the recordings using the plate as a source seem to have higher amplitude at frequencies above 70 Hz. On the other hand, in the horizontal crossline recordings, the prism induced recordings had consistently higher amplitudes compared to the ones induced from the shear-beamer. Additionally, the relative changes show some deviations. These results are in line with my own experience; a shear beam source provides less consistent results when compared to the prism source. Additionally, former tests with other shear sources confirmed these observations (Häusler et al., 2018). Moreover, shear sources produce lower amplitudes compared to the prism source.

The results of this study show that the overall performance of the prism source compared to the conventional sources is similar. However, since we can increase the consistency in terms of source coupling during acquisition, employ all three directional components just with one source and reduce the acquisition time, I strongly recommend the prism source for shallow seismic applications.

7.4 Acquisition set-up

The data were acquired on the 29th and 30th of September 2018 under relatively dry soil conditions. The layout of the seismic survey that I designed is shown in Figure 7.5. It consists of two profiles (P1 and P2) crossing each other perpendicularly, while passing through a series of boreholes where the GPR measurements took place. This was done for future investigations of coherences in the structures seen by the two geophysical approaches. To reduce any potential noise produced by the boreholes, a minimum distance of one meter was kept at all times, between the receiver array and the boreholes. The orientation of the two profiles is indicated in Figure 7.5 and North is indicated in the upper right. The geophones that were used are the same as described in Chapter 6 (3-component geophones: Geospace Technologies GSC-11D, eigenfrequency 4.5 Hz). For both profiles, the receiver spacing was 1 m and the source spacing 4 m.

For profile 1, 48 geophones were used to record the vertical and horizontal crossline component. For the vertical component, 14 hammer blows on a steel plate acted as a source (stacking 3 times). For the horizontal crossline component, 28 hammer blows on the prismatic source were employed (14 on each side, while stacking 4 times). Due to time and hardware constraints on profile 1, I was able to acquire only these two components. Therefore, the prismatic source was not used to record the signal on the vertical component. The data were acquired sequentially, i.e., first only the vertical component and then subsequently only the horizontal crossline. The profile crosses the boreholes 67, 31/62, 26/61, 22, 64 and 65, with the last one being the end of the receiver line.

For profile 2, 36 geophones were used to record all three components. The prismatic source was used while measuring with vertical and the horizontal crossline components (11 shots in each of the two sides, while stacking 4 times). The plate was used while measuring with horizontal inline and vertical component (11 shots, while stacking 3 times). Since the vertical component was acquired an additional time, I used this data for quality control by comparing the two signals. The response was very similar as was shown in Section 7.3. The profile 2 crosses the boreholes 48, 32, 38, 31, 62 and 30. An overview of the field conditions and the acquisition layout is shown in Figure 7.6.

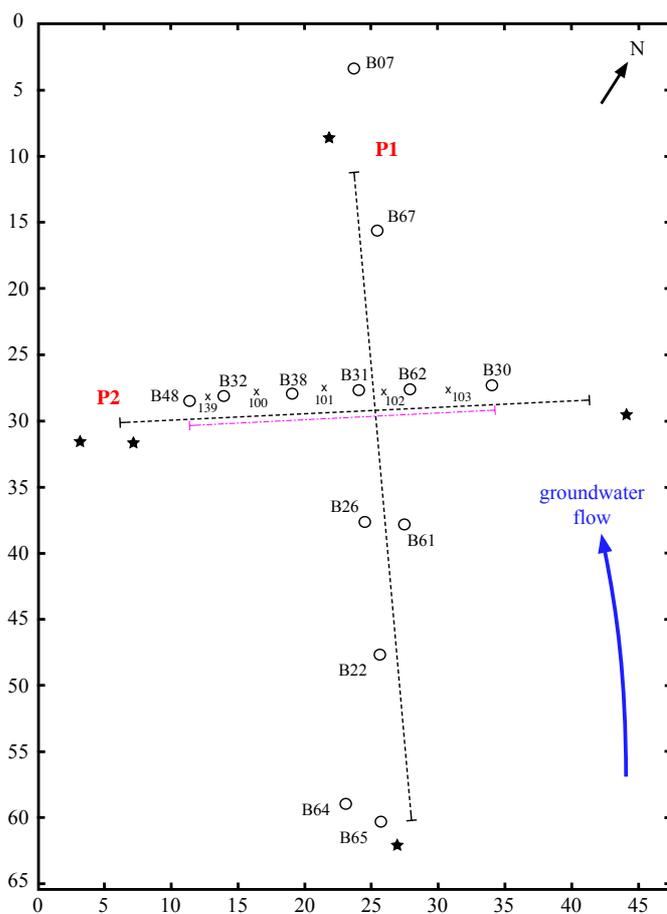


Figure 7.5: Map of the Krauthausen test site including the two seismic profiles and the boreholes adjacent to them (numbered circles). Black dashed lines show the receiver arrays and stars show the location of the first and last shot for each of the two profiles. Crosses indicate the location of cone penetrating measurements done by Tillmann et al. (2008). The purple dashed line shows the seismic profile that was used for a preliminary evaluation (Athanasopoulos and Bohlen, 2017a). The unit is m.



Figure 7.6: Overview of the field conditions and the acquisition layout, along with the prismatic source and its operation towards the inline component (direction of the receiver array).

7.5 Field data analysis and FWI parametrization

7.5.1 Preprocessing of field data

Seismic data were recorded for 2 s using a temporal sampling of $2.5 \cdot 10^{-4}$ s in all cases. However, during processing the total recording time was reduced to 0.6 s, as there were no events at later times. The preprocessing steps are very similar to the ones shown in Chapter 6 and included: 1) upsampling to $2.5 \cdot 10^{-6}$ s, 2) muting any events prior to the first arrivals, 3) zero-padding the signal with adding 0.02 s (required for the source-time function estimation), 4) transforming the data from 3D to 2D and 5) applying a band-pass filter between 4 Hz and 160 Hz.

Figure 7.7 shows the data after preprocessing for each of the three components and for both profiles. The dominant energy of the recordings originates from the surface waves, either Love waves for the crossline component or Rayleigh waves for the vertical and inline components. The refracted waves have a much lower amplitude relative to the surface waves, when compared to the data shown in Chapter 6.

One possible explanation might be that the aquifer, while being multi-layered, is seen as layers of low contrast in terms of seismic velocities. This makes the application of conventional techniques, such as refraction tomography, rather challenging. It is worth mentioning that there is no bedrock over the depth of investigation from the surveys I performed. In this case, although the TF-FWI provided lower overall misfit and a more stable inversion (especially in the early stages of FWI), it did not show significant updates in the P-wave velocity model. However, this does not necessarily mean that the inversion failed to update v_p accurately. Instead, I suggest that it did so relatively to the amount of information contained in the data. In terms of frequency content, in Figure 7.8, I show the mean amplitude spectra of the data measured by the three different components. There are no big deviations, with an exception of the horizontal crossline component (blue line) which exhibits a maximum at a slightly lower frequency compared to the P-SV components. The main frequency content of the data lies between 8 Hz and 120 Hz.

7.5.2 FWI set-up

The initial v_p , v_s and ρ models are calculated from the arrival times of the refracted waves, dispersion curve analysis, and through Gardner's relationship, respectively, as done in the previous chapter. In all three models the values of velocity and density increase with a smooth gradient across the whole model, due to the absence of a bedrock. Specifically, the initial models of v_p and v_s are given values from 480 to 1800 m/s and 190 to 450 m/s, respectively, while the density varies from 1200 to 2100 kg/m³.

The resulting values of the local grid-search returns $Q = 15$, which is expected for such unconsolidated environments. Then I proceeded by constructing the attenuation model by a GSLS with three relaxation mechanisms. The calculated relaxation frequencies are $f_1 = 0.68$ Hz, $f_2 = 8.77$ Hz and $f_3 = 81.64$ Hz and the τ -value is 0.05, where I use 35 Hz as a reference frequency. After estimating the initial models I calculate the synthetic data and perform a quality control to verify that no cycle-skipping is present, as shown in Section 5.3.4, and correct the initial model accordingly.

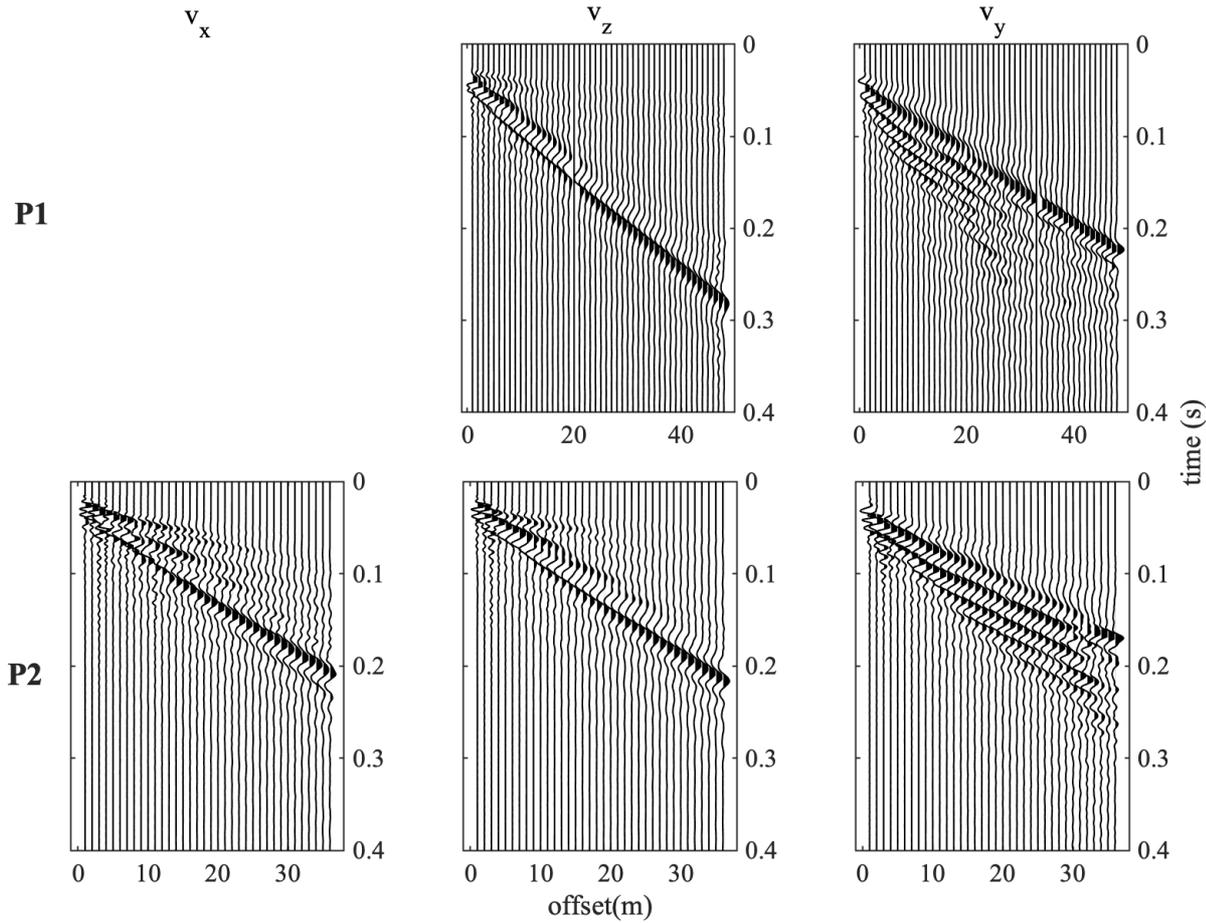


Figure 7.7: Trace-wise normalized shot-gathers of the field seismic data. The corresponding sources are located at profile coordinates of -3 m and -1.5 m, for profiles 1 and 2, respectively. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH-waves. The gathers in both cases are dominated by surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.

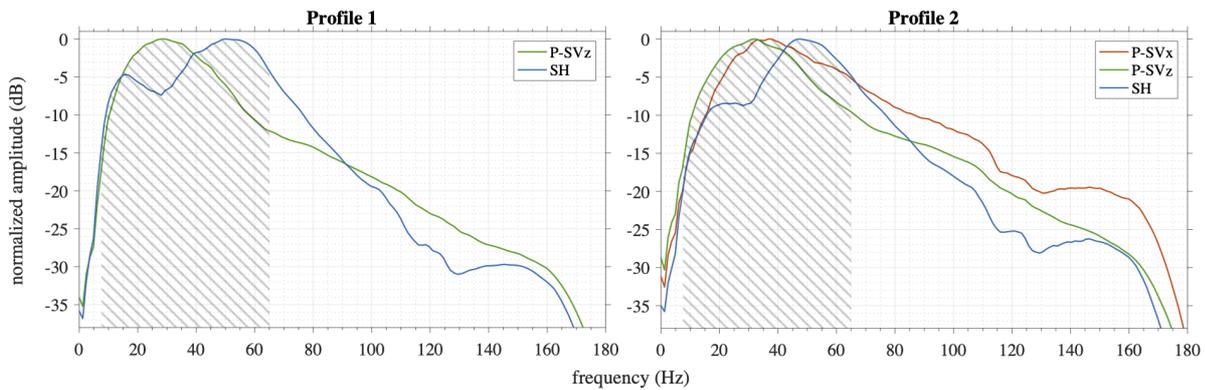


Figure 7.8: Amplitude-frequency spectra for each component of the two profiles, cumulative for every source and receiver. The spectrum of the P-SV waves is shown in red and green, for v_x and v_z components respectively and of the SH waves (v_y) in blue. The shown spectra are the average of the spectra of the normalized traces. The hatched area corresponds to the target frequency band which was used for FWI.

The model space consists of 540 (P1) and 440 (P2) grid points in the horizontal direction and 160 grid points in the vertical direction, resulting in a model space of approximately 70 m x 20 m for profile 1 and 55 m x 20 m for profile 2, with the grid spacing set to 0.125 m for both directions. The seismograms are normalized trace by trace. The source-time function is inverted as shown in Section 5.4 and updated

Table 7.1: Workflow used for FWI of field data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass frequency filter. The parameter ' γ ' indicates the damping factor that is used during the corresponding stage (column 'Stage') in equation 4.3.

Stage	Update			LP in Hz	γ
	v_P	v_S	ρ		
1	1	0	0	10	15
2	1	0	0	10	12
3	1	1	0	15	9
4	1	1	1	20	6
5	1	1	1	25	3
6	1	1	1	30	0
7	1	1	1	35	0
8	1	1	1	40	0
9–10	1	1	1	increment of 10	0

only at every new stage of the multiscale approach. The workflow parameters along with the respective tapering are the same for both inversions. The preconditioning applied to the synthetic reconstruction test (Section 4) is applied to the field data application. This includes circular tapers around the source (radius of 4 grid points), the approximation to the diagonal elements of the Hessian and the application of 2D median filter to the gradients, where the filter has a size of 1.5 m (12 grid points).

The inversion workflow is shown in Table 7.1. In general, 10 inversion set ups² were performed for each individual component and their possible pairs, both with CFWI and TF-FWI. In the synthetic studies, I have shown the improvements gained by the joint inversion and by the application of TF-FWI versus the conventional approach. Therefore, I focus here only on using all available data simultaneously and the TF-FWI approach.

7.6 Results

In this section I present the models obtained from multiparameter/multicomponent FWI. The section is subdivided into three parts. The first part shows the individual results from each profile and a detailed comparison of the data fit in both time and frequency representations. In the second part, the two profiles are combined in order to compare the various geological facies in a qualitative sense. Finally, in the subsequent section the results of seismic FWI are compared with CPT logs in the area obtained from the group at the research center in Jülich (Tillmann et al., 2008).

² For profile 1 the following inversion set-ups were tested: P_y^1 , P_z^1 and P_{yz}^1 . For profile 2 all three components and pairs of them were evaluated: P_x^2 , P_y^2 , P_z^2 , P_{xz}^2 , P_{xy}^2 , P_{zy}^2 and P_{xyz}^2 .

7.6.1 FWI results of profile 1 and 2

The models obtained through joint inversion of the P-SV and SH data, along with the initial models for both profile 1 and 2, are shown in Figures 7.9 and 7.10, respectively.

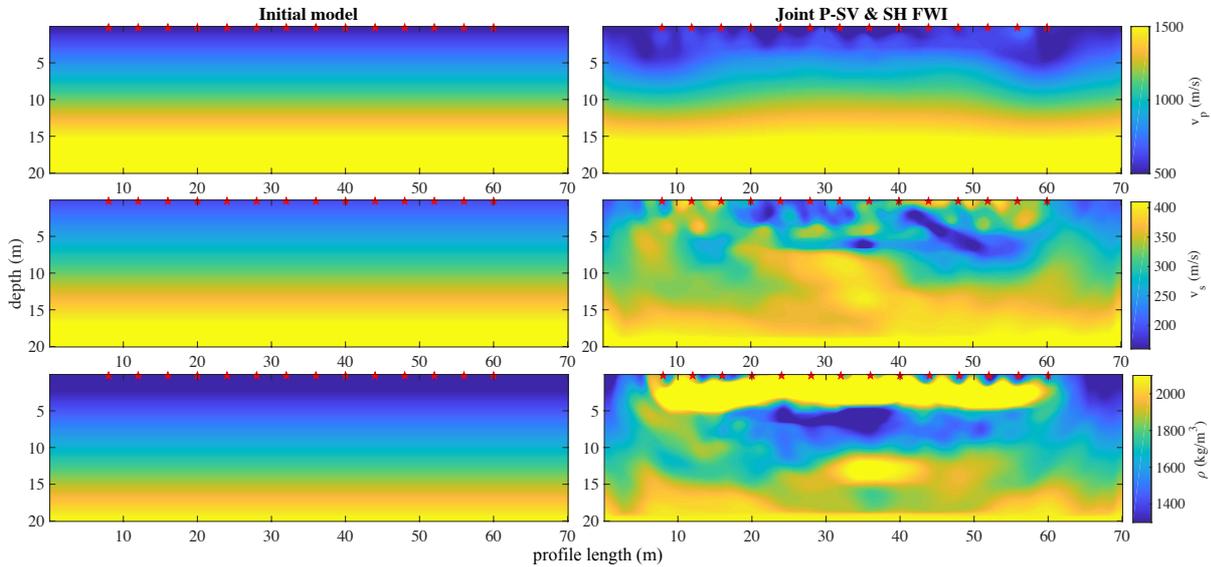


Figure 7.9: Models obtained by joint FWI of field data acquired at profile 1 (Figure 7.5). The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of joint FWI. Red stars represent the source locations.

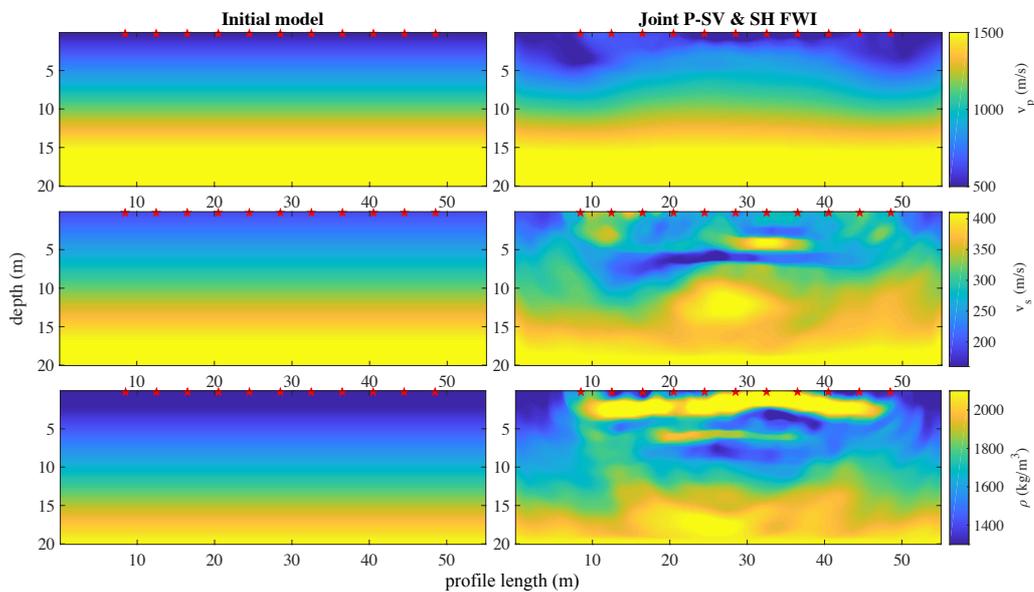


Figure 7.10: Models obtained by joint FWI of field data acquired at profile 2 (Figure 7.5). The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of joint FWI. Red stars represent the source locations.

P-wave velocity

As already mentioned, the v_p model did not get significant updates compared to the initial model, with the exception of the upper 3 to 5 m, for both profiles 1 and 2. Unconsolidated sediments with small contrasts in v_p would require much longer profiles in order to deduct information from the refracted waves or even applications of reflection seismics. This is also due to the lack of a bedrock over the depth of investigation. The location of the groundwater table at around 2 to 3 m, is probably the main reason that limited the v_p updates in the upper few meters of the subsurface only. The effect of water saturation on v_p is extremely high, leading to much higher v_p , while the v_s is barely affected.

The general fit of the refracted wave regarding its phase but also amplitudes (although less) is high (Figure 7.11a and 7.12a), for both profiles 1 and 2. This also holds when considering that this is a field data application where noise is also present in the recordings. The results indicate that the inversion was able to use the available information to update v_p .

S-wave velocity

As opposed to the v_p model, the retrieved v_s structure shows a high level of complexity. In particular, there appears to be distinct sedimentologic boundaries in the aquifer material at depths of 4, 5, 6.5 and 13 m, which correlate well for both profiles. In particular, v_s is low at the upper 3 m signifying the soil horizon with loose sediments. This layer also correlates with the depth the groundwater table and the upper part of the aquifer which is characterized by low saturation.

According to the geologic model of the area, we find a layer of poorly sorted gravel directly below the soil (Figure 7.1b). This structure correlates well with an increase in v_s (around 330 m/s) at depths of 4 to 5 m. An interesting feature that shows up in the inversion results is a thin layer of low v_s (around 250 m/s) at depths between 5 to 6.5 m, which according to information on the test site from previous studies (Gueting et al., 2015) is interpreted as the sand layer.

My interpretation, based on literature (Telford et al., 1990), is that a sandy soil should exhibit lower shear stress compared to gravelly soil for low compaction pressure, which is the case in unconsolidated sediments. The layer of sand in terms of the information from the test site has a thickness of around 1 m (with some spatial variations) and spans from depths of 4 to 6 m, at various locations. This correlates well with the obtained models from the joint FWI of both profiles. Especially, profile 2 shows a more distinct boundary between this layer and the two higher-velocity layers that are located above and beneath.

There is a small deviation with the prior knowledge of the test site in terms of the lateral width of this layer. However, the geological model which is derived from extraction of CTPs is limited to specific measuring points and therefore cannot be generalized for every location in the aquifer. I do believe that the margin of difference between the CTPs and the v_s inversion results allow for such an interpretation, especially since they correlate from depths of 5 to 6 m. Below this layer, v_s increases which is interpreted as the layer of fine gravel (higher degree of compaction) and comes in agreement with the borehole information, signifying the soil change from sand to gravel. The differences in the values compared to the shallower gravel layer do not come as a surprise, since shear wave velocity can vary within the same rock or soil

with weathering, compactness or even moisture content. All of these properties can vary significantly in such a subsurface system.

Finally, at depths of around 12-13 m, v_s reaches values of approximately 380 m/s, which is interpreted as the clay layer relative to its location from the boreholes. As clay content increases, the cohesion of clay minerals is expected to cause an increase in shear velocity.

Density model

The last row of Figures 7.9 and 7.10 shows the inverted density model. In both profiles 1 and 2, density shows very high values in the upper 3 to 4 m. This layer is very interesting, since there is no strong interface resolved in the v_p and v_s models that matches this layer, which could otherwise be interpreted as an effect of cross-talk. The density values drop for the next 1 m where presumably the gravel layer is located. There is a clear deviation between the two models at depth of 6 m, where we observe a layer of high density values at the location of the sand layer (from my interpretation of the v_s model) in profile 2. However, since profile 2 included all three components and based on the synthetic studies, this model has higher reliability. Below this layer, the density values decrease, signifying the layer of gravel. This interpretation is based on the matching values of the gravel layer with the one at 4 to 5 m depth.

At depths of around 12 to 13 m, the density of profile 2 rises at depths that match the clay layer. Profile 1 shows a further low density area at depths of 16 m. However, this is rather unreliable since the resolution at that depth is highly reduced compared to the shallower parts. Overall, the results of density correlate well with the structures of the v_s model, without necessarily sharing the same footprint (cross-talk effects). In Section 7.6.2 I will further investigate these results using independent data from logs that were extracted at the test site.

Data fit of FWI results

To evaluate the inversion results, I compare the velocity seismograms of the observed and synthetic data for all different components (Figure 7.11a and 7.12a, for profile 1 and 2, respectively), at the last frequency stage (60 Hz). For profile 1, the two components that were used match the observed data at a satisfactory degree. Both the fundamental modes of Rayleigh and Love waves are very well fitted. In particular, for the receiver 48 at the 1st shot location, the fundamental mode of Rayleigh waves arrives at approximately 0.21 s and the fundamental mode of the Love wave at 0.16 s (Figure 7.11a and 7.12a, top row). Similar results are obtained for all other shot-gathers. The refracted P-wave observed from the v_z component (only) also fits the field data, with some small deviations. The results on the v_y component show clearly higher misfit values, in comparison to the results on the v_z component. The highly dispersive behavior on this test site shows, in contrary to the previous field study (Chapter 6), that SH inversion can be more complicated than P-SV inversion despite the fact that it is a two-parameter versus three-parameter inversion. The difficulty of inverting for highly dispersive Love waves, in the context of FWI has also been observed by other studies (Köhn et al., 2019).

However, in profile 2 it seems that the addition of the v_x component in the P-SV case allowed the joint inversion to further improve the fit of the v_y component. In this case, all three components match the

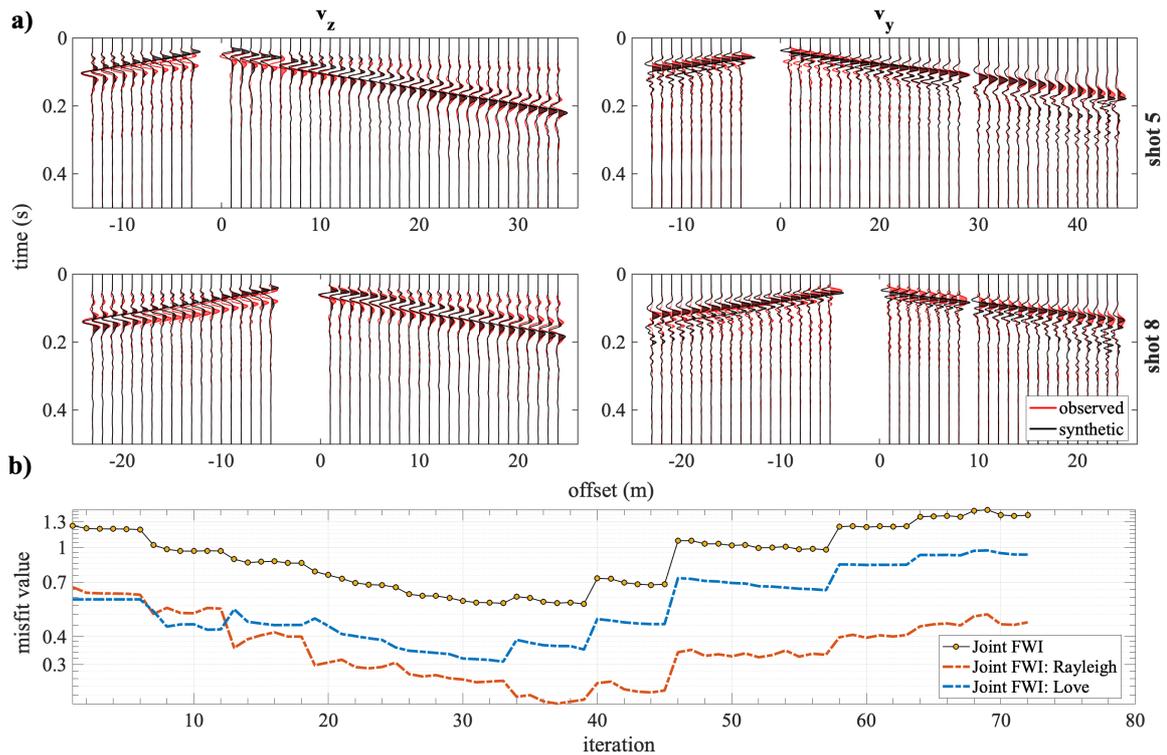


Figure 7.11: Inversion results of joint FWI for profile 1, corresponding to the models shown in Figure 7.10. a) Comparison of observed and synthetic velocity seismograms at the 1st and 6th shot location. b) Misfit as a function of the number of iteration. The dashed blue and red lines constitute the contribution of the Love and Rayleigh waves, respectively, to the joint misfit function (dashed yellow). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 7.1).

observed data to higher degree, which can be seen in Figures 7.11b and 7.12b. Specifically, despite the fact that in profile 2 the misfit is estimated from one additional component, the final misfit value is less compared to the misfit function of profile 1. Figure 7.12a shows the high correlation between observed data and the inversion results from different shot locations. The main events are fitted exceptionally well, while some mismatches are observed for far-offset traces, as is usually the case.

After evaluating the data fit in the time domain, I compare the frequency spectra of observed and synthetic data for the first shot location of each of the two profiles (Figure 7.13 and 7.14). The data in these figures agree with the seismogram fit, but additionally show that the highest mismatch in terms of frequency content occurs for the results of the v_y component of profile 1. Also the v_x component of profile 2 shows that the inversion result did not resolve the higher frequency content of the observed data. However, considering that this is a multicomponent inversion where all three data sets have to simultaneously be fitted by FWI, the results are more than satisfying.

A final step for evaluating the FWI results, as already mentioned, comes from the study of the inverted source-time functions. In Figures 7.15 and 7.16, the inverted source wavelets for each component and each shot are shown, for the first and last inversion stage. The high similarity in the first stage of FWI shows that the initial model was optimally chosen to fit the main events of both P-SV and SH waves. At the last stage, the coherency, especially for the P-SV waves of both profiles, is maintained. In the SH case, once again the results from profile 2 outperform the ones from profile 1. There is no clear indication to why this is the case. However, an interpretation could be that the influence from the boreholes located

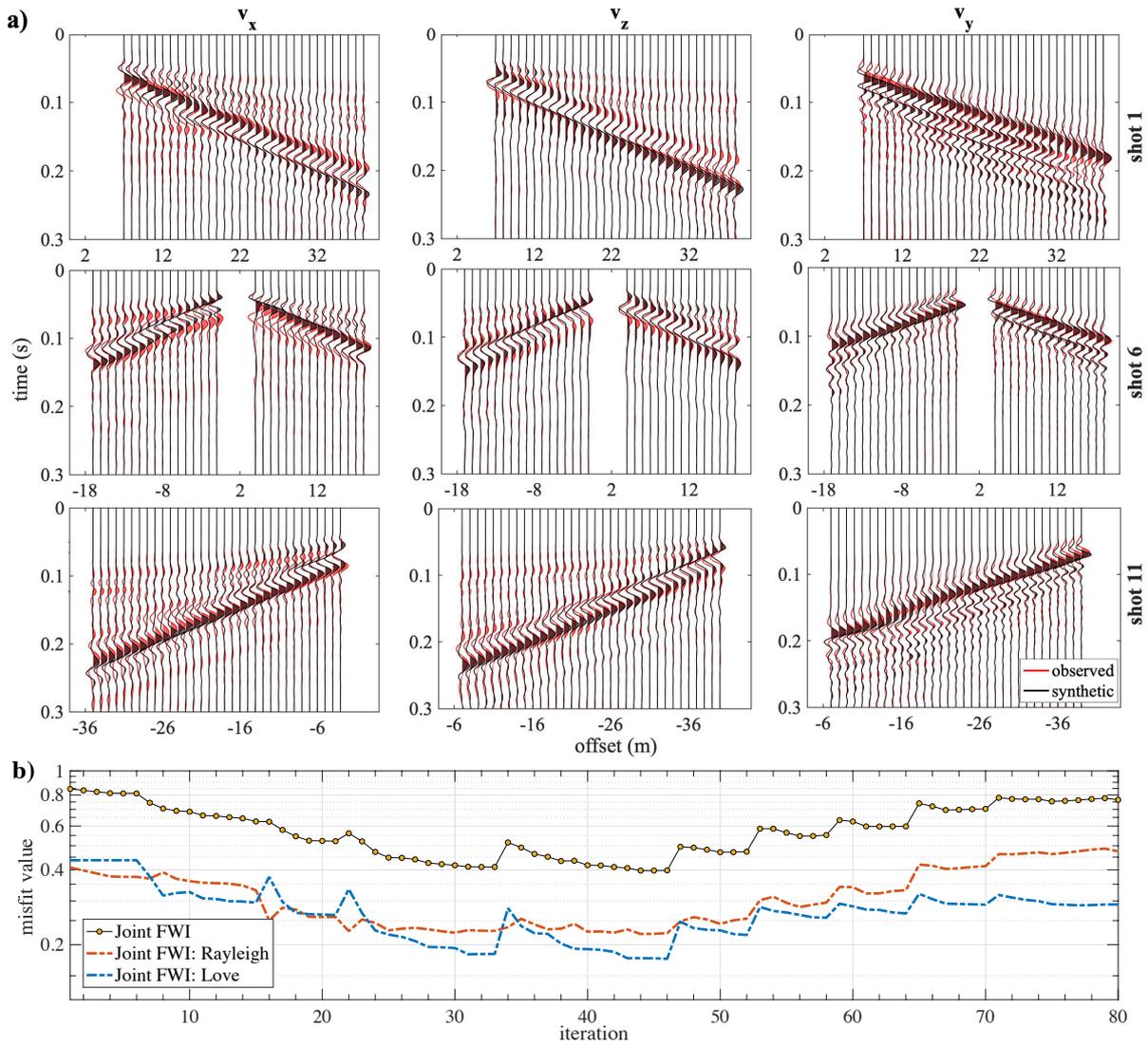


Figure 7.12: Inversion results of joint FWI for profile 2, corresponding to the models shown in Figure 7.10. a) Comparison of observed and synthetic velocity seismograms at the 1st, 6th and 11th shot location. b) Misfit as a function of the number of iteration. The dashed blue and red lines constitute the contribution of the Love and Rayleigh waves, respectively, to the joint misfit function (dashed yellow). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 7.1).

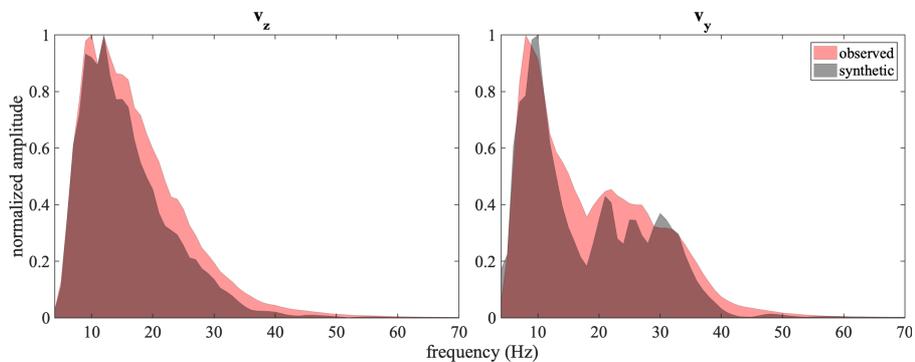


Figure 7.13: Frequency spectra of the observed and synthetic v_z and v_y seismic data at the first shot location (Figure 7.11.a.)

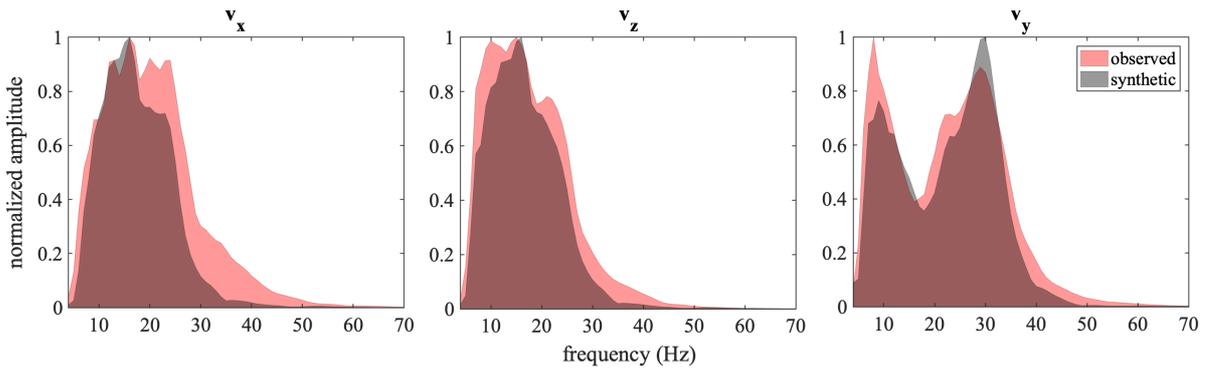


Figure 7.14: Frequency spectra of the observed and synthetic v_x , v_z and v_y seismic data at the first shot location (Figure 7.12.a).

across profile 1 was bigger compared to ones at profile 2, where a minimum of 1 m was kept at all times (Figure 7.5). Especially since this problem was identified on the crossline component, such an explanation could be relevant.

Additionally, I provide in Appendix D the dispersion curves extracted from the first shot location at each of the two profiles, where similar observations can be made. The data fit, in both the time and frequency

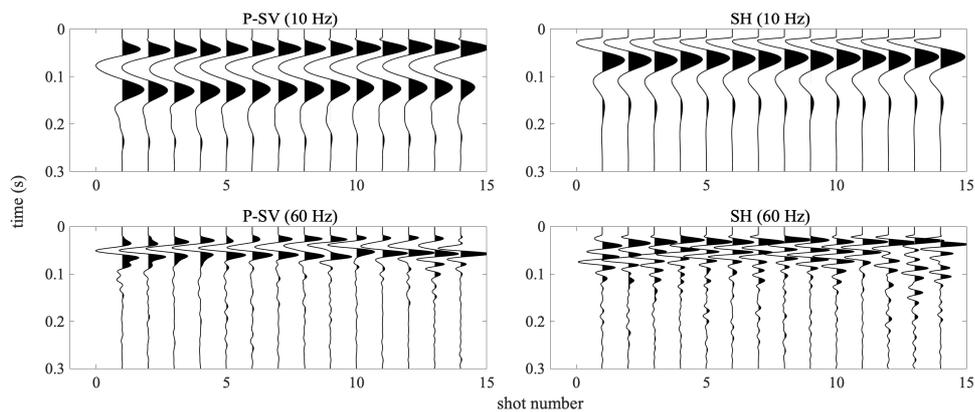


Figure 7.15: Estimated wavelets for the source-time function of P-SV (left) and SH (right) data acquired along profile 1, from the first (top) and last (bottom) inversion stages. The source wavelets are overall coherent for the various shot locations, for both first and final stages. The inverted wavelets from the crossline source are in general less coherent

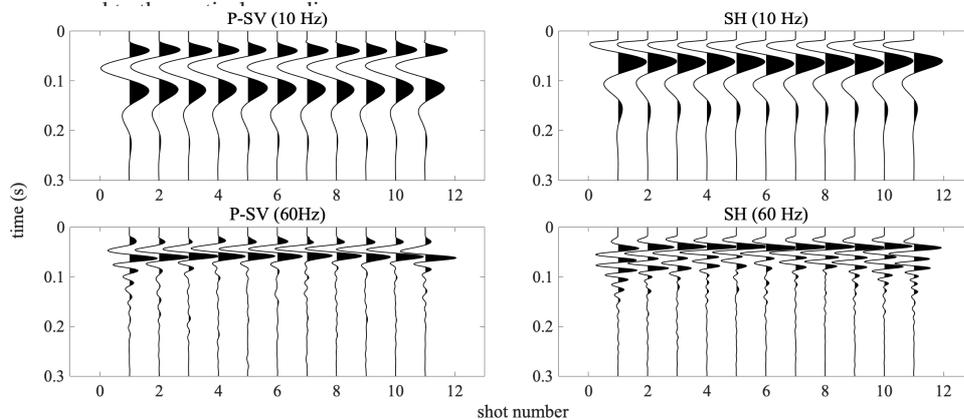


Figure 7.16: Estimated wavelets for the source-time function of P-SV (left) and SH (right) data acquired along profile 2, from the first (top) and last (bottom) inversion stages. The source wavelets are overall coherent for the various shot locations, for both first and final stages. The inverted wavelets from the crossline source are in general slightly less coherent compared to the vertical recordings.

domain, suggests that the application of the joint FWI in both profiles was overall successful. Once again, in field data applications obtaining the exact response of the medium under our specific numerical approximations is not feasible, since we additionally ignore the presence of noise in the data.

7.6.2 Evaluation of FWI density model with CPT measurements

As already mentioned, many models exist that produce the same data, making the decision of what actually constitutes a reliable result very challenging when it comes to field data. In the previous chapter, I used GPR migrated results and superimposed them onto the velocity FWI results in order to assess the reliability of the inversion. For the Krauthausen site I was given cone penetrating data that were extracted and studied by Tillmann et al. (2008) and Gueting et al. (2017). The locations of the CTPs are shown in Figure 7.5 and marked by their ID number (139, 100, 101, 102 and 103). At each of these locations, vertical profiles of mechanic cone resistance, natural gamma, bulk density and water content were measured up to a depth of approximately 13 m, with sampling intervals of 10 cm.

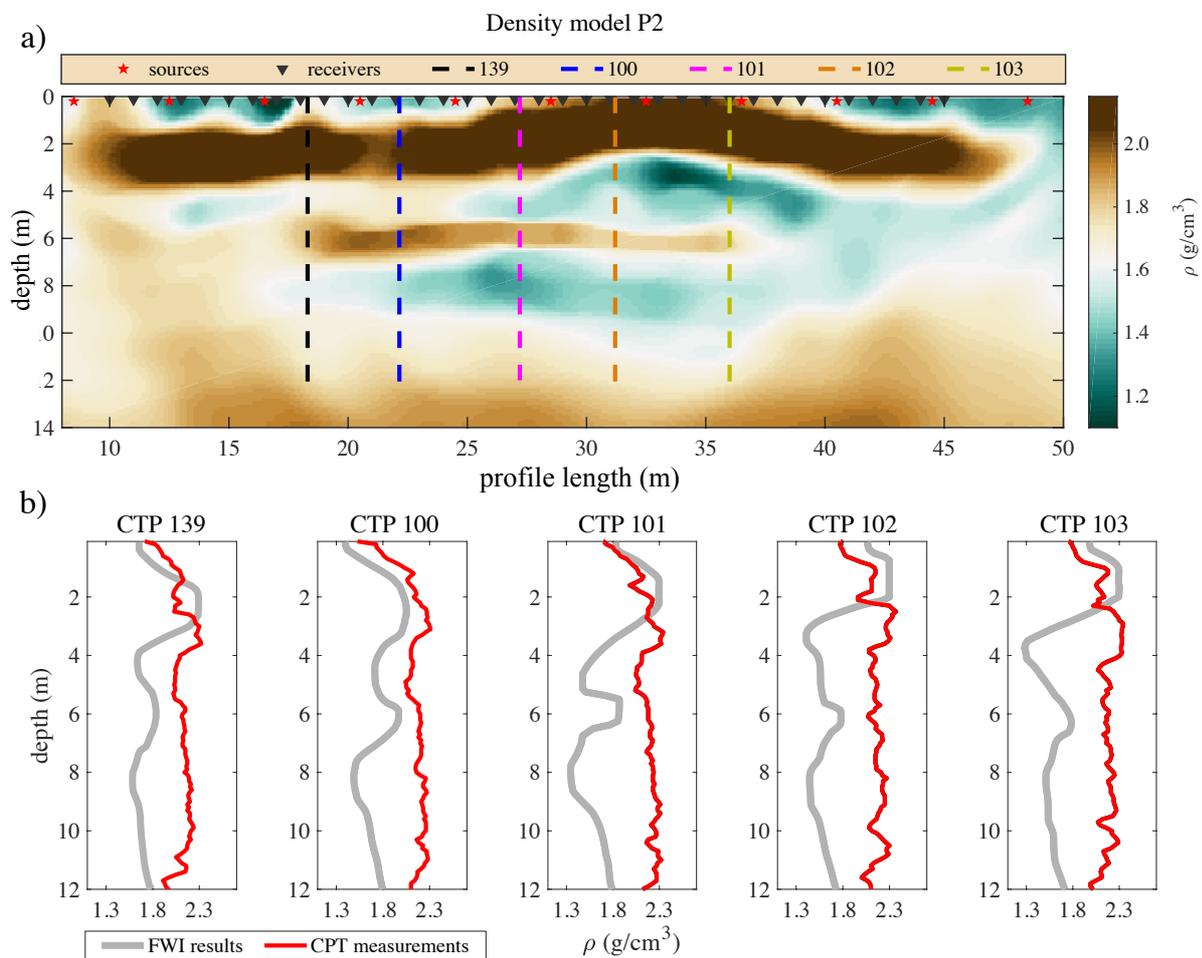


Figure 7.17: (a) Density obtained from FWI. (b) Comparison of vertical bulk density profiles derived from CPT with co-located 1D density models from FWI. The exact locations of the CPT profiles within the FWI model are indicated by the dashed lines.

From these data, the bulk density can be directly compared with the results of seismic FWI. The bulk density, ρ_{bulk} , according to Biot (1956), is given by

$$\rho_{bulk} = \rho_{fluid}\phi + \rho_{solid}(1 - \phi) \quad (7.1)$$

where ϕ is the porosity. In case I would solve the poroelastic wave equation, then a correction introduced by accounting for the porosity would be required (Liu and Bohlen, 2020). However, in the viscoelastic case that was used here the density from FWI is equivalent to the the bulk density of Equation 7.1.

Figure 7.17a shows the density model obtained from joint FWI at profile 2, along with the locations of the CTP measurements. In Figure 7.17b, I overlay the bulk density from the CTPs with 1D vertical profiles from the density model obtained by FWI. It is worth noting that the CTPs are located approximately 1 m west of the seismic survey. There is a very high correlation in the upper 2 m for all profiles. In particular, CTP₁₃₉ and CTP₁₀₀ matches the relative changes over depth up to 7 m, indicating that the inversion of the density was successful. There are higher deviations for the other profiles and for increasing depth. However, this can also be caused by the spatial distance between the CTP and seismic profiles. At depths from 7 to 12 m, all profiles show a relatively constant density, with fluctuations in terms of the absolute value. In Chapter 4, I have shown that in the case where the v_p and v_s models are well reconstructed, the density structure is also well resolved. Therefore, the high correlation of the CTPs with the FWI density model shows, indirectly, that the retrieved v_p and v_s models are also reliable. Considering that density in most studies is seen as a *weak* parameter, I have shown here that proper use of multicomponent data can greatly improve the inversion of density.

7.6.3 Cross-section of P1 and P2

As a final result, Figure 7.18 shows the results of all models from both profiles in a 3D layout. The two planes match well in the majority of the cases. For the v_p model, in both profiles the updates are contained mainly on the upper few meters, retaining the structure imposed by the initial model. The v_s model shows the highest similarity of the three in terms of spatial coherency between the two directional profiles, as was discussed in more detail above. The density model, after careful validation, revealed independent structures compared to the other two models, while also maintaining a high level of coherency between the two cross-sections. The general structure of the aquifer reveals a very complex architecture with multiple layers of small lateral volume. Therefore, it constitutes a challenging target to retrieve when dealing with non-invasive techniques, as in my case.

7.7 Summary

In this chapter, I presented results of a viscoelastic multiparameter multicomponent FWI on seismic data acquired at the Krauthausen test site, with the aim to structurally characterize the architecture of the aquifer. To perform FWI in such a complex medium, a specific source was employed that ensured a good coupling when recording multicomponent data and also reduced the total acquisition time. The comparison of this source with conventional ones revealed a high correlation, allowing me to suggest the use of the prismatic source for future applications of shallow land seismics. The acquisition set-up

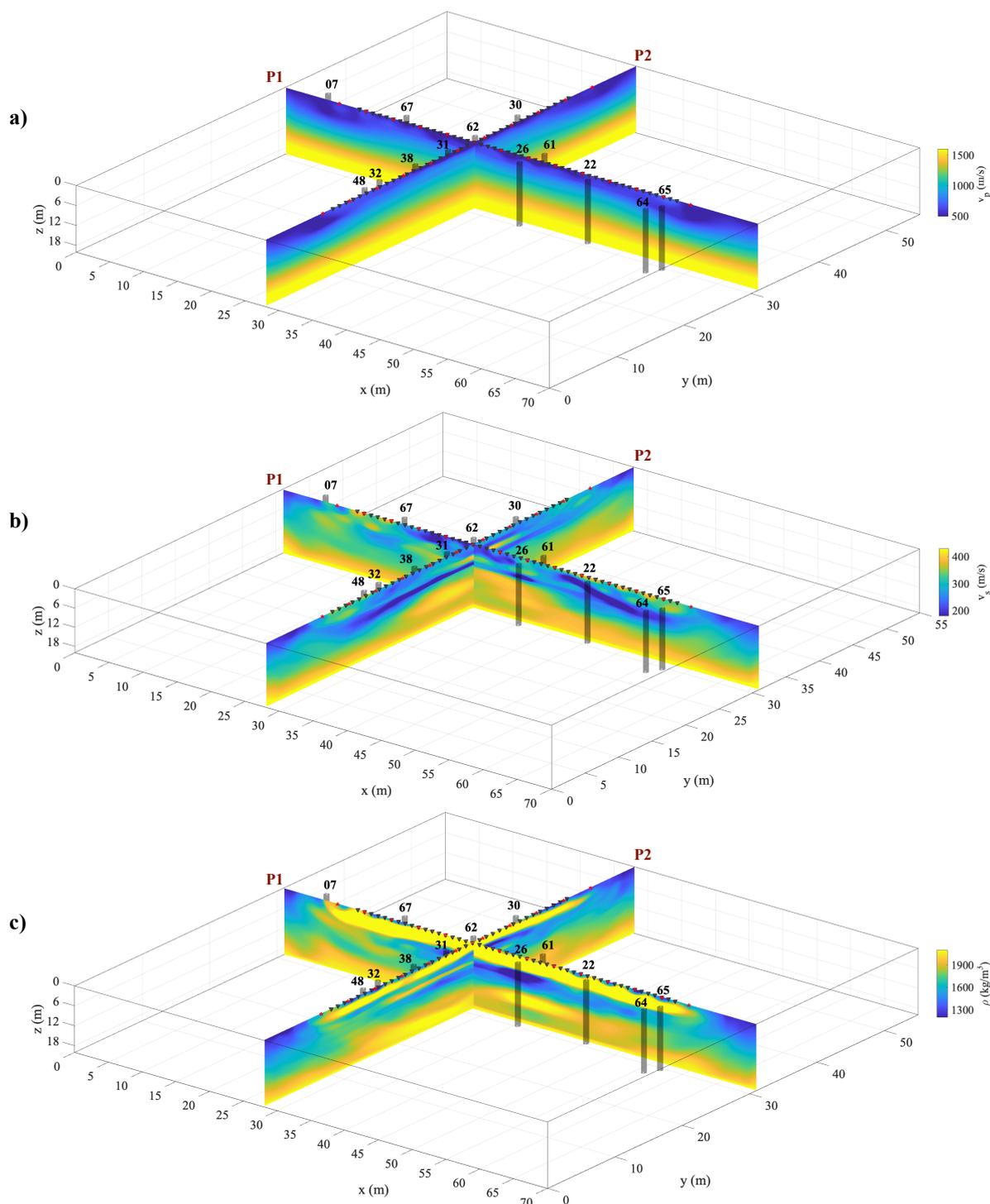


Figure 7.18: 3D plot of the retrieved v_p , v_s and density models from joint FWI of Rayleigh and Love waves of the two profiles. Black (opal) cylinders represent the boreholes in the area that were used to perform crosshole GPR FWI by Gueting et al. (2015, 2017).

was designed to match previous measurements of independent data that were recorded at the test site. The inversion was successful as both the residuals between observed and synthetic seismograms as well as the misfit function were minimized. Effects related to cycle-skipping were reduced through careful selection of the initial model and making use of TF-FWI, for which I already demonstrated its robustness in the previous chapter. The v_s model was most reliable due to the accurate fit of the Rayleigh wave

and its higher modes. The updates of the v_p model in this case were limited, due to the particularly low amplitudes of the refracted waves in the site and the shallow depth of the groundwater table. Contrary to previous studies, the performance of SH FWI does not always surpass P-SV FWI, despite having less parameters to invert for, and therefore less unknowns. Highly dispersive media place a challenge on the inversion of Love waves, but can be improved by making use of multicomponent strategies. Finally, in order to overcome the uncertainties of the reconstructed models, due to the inherent non-uniqueness of FWI, the FWI results were compared with cone penetrating data that were previously acquired. For the first time, a high correlation of the density model was obtained from several field data recordings. This also indirectly implies the reliability of the other two model parameters, as was illustrated by the synthetic studies (Chapter 4). The outcome of this study is that seismic methods, when used properly, can reveal significant information about the structural complexity of an aquifer system, and provide high spatial-resolution mapping of its architecture.

Chapter 8

Summary and Conclusions

The main outcome of this thesis is the development of strategies to ensure the successful application of FWI as a way to improve our knowledge of the subsurface. The studies that were performed focus on near-surface site investigations, with the aim to obtain high resolution models of geotechnical, archaeological and hydrogeophysical targets. The application of seismic FWI in these fields of research constitutes a new way to characterize the subsurface, by exceeding the resolution provided by conventional seismic techniques. The developments presented here: 1) establish a better understanding of surface wave scattering, 2) advances towards multiparameter and multicomponent FWI, and 3) provide a complete workflow of FWI in near-surface field data applications.

Surface wave scattering

In Chapter 3 it was shown that by using numerical methods the viscoelastic wave equation of a scattered wavefield can be obtained, as an effort to assess the various effects of an inclusion in the recorded wavefield. The scattering response from the individual model parameters was examined in terms of kinematics and directivity of the scattered wavefield, in order to delineate how the various parameters affect the recordings. It was demonstrated that for the P-wave velocity the seismic response is composed mainly of forward scattered P and S waves and surface waves, while for the S-wave velocity the seismic response is dominated by Rayleigh-wave scattering into the forward direction. Since the amplitude of Rayleigh waves also dominates the near-surface recordings, it becomes clear that the S-wave velocity can be easier retrieved through FWI, when compared to the P-wave velocity. Additionally, it was shown that similarity of the scattering responses between attenuation and velocity for both P and S waves, will result in the development of interparameter cross-talk when performing viscoelastic FWI. Finally, it was demonstrated that the scattering due to density perturbation bears a unique signature, causing a strong back-scattering of P and especially S waves caused by both the direct P and Rayleigh waves. Additional findings included the change of the scattering response from shape and composition variations of the inclusion. In particular, it was demonstrated that composition effects usually affect the dynamic characteristics of the wavefield (i.e. amplitude), whereas the kinematic behavior is mainly controlled by shape effects. This is in agreement to the behavior of scattered body waves as shown in previous studies.

Steps towards multiparameter and multicomponent FWI

Chapter 4 presented a series of synthetic studies which served as preparation for the application of FWI on the field data that followed. The benefits of multicomponent multiparameter FWI constitutes one of the main focus of this chapter. Both P-SV and SH recordings were used in FWI, either individually or jointly. In simple cases, with spatial coherent subsurface models, all three methods were able to provide similar results, and to perfectly reconstruct the model parameters. In cases where higher structural complexity and uncorrelated models were used the results revealed some interesting findings. The joint FWI outperformed the P-SV FWI and in some cases also the SH FWI. In the absence of an accurate initial model for the P-wave velocity, the P-SV FWI was not able to provide a reliable P-wave velocity model. Additionally, artifacts at the density and S-wave velocity model were introduced. Even with sufficiently good knowledge for the P-wave velocity, the P-SV FWI still struggled to provide reliable results for the P-wave velocity. In such cases, the SH FWI was able to reconstruct the subsurface models more accurately, due to the fact that its convergence behavior is independent of the P-wave velocity.

However, P-wave velocity is of great importance to improve site characterization, since when coupled with the S-wave velocity it can provide a spatial distribution of the Poisson's ratio and allow an accurate imaging of the area under examination. Opposite to long-offset seismic surveys, the reconstruction of P-wave velocity model was poor in every near-surface case that was evaluated, due to the preference of the FWI towards minimizing the high amplitude surface waves. To overcome this problem, a time-frequency windowing approach was developed, where the contribution of later arrivals was gradually increased in the course of iterations. The TF-FWI approach showed improved reconstruction of all three model parameters, especially of v_p . The improved v_p allowed to reduce artifacts in v_s and obtain a more accurate density model.

Additionally, it was revealed that an accurate reconstruction of the P- and S-wave velocity can also improve the density model. This is an important feature, since the so far experience from the hydrocarbon industry indicates that the density model is typically used only to compensate for physical effects that are not accounted for. In the case of near-surface applications, however, the inclusion of surface waves provide a valuable source of information, which leads to better reconstruction of the density model. This effect is further enhanced when multicomponent data are used during the inversion and played a key role on the successful retrieval of the density model in the field studies that followed.

As a final step, the development of a numerical scheme was shown, that allows for structurally constrained FWI. The scheme is not restricted to applications between only the elastic parameters, but can be coupled even with independent geophysical data.

Field data application of FWI on unconventional targets

Chapter 5 constitutes a general overview of proper field acquisition and advanced processing techniques, which serve as a preparation for field data FWI applications. Several techniques were presented to enable the construction of the initial model, by considering prior information that can be easily extracted from the data. An approach to gauge the quality of the initial model, in order to prevent cycle-skipping, is developed and recommended for every application of seismic FWI.

While FWI has been an established technique in hydrocarbon industry, only few studies so far have investigated its applicability in near-surface targets. Two near-surface targets were the focus of the multicomponent and multiparameter FWI and the techniques that were developed throughout this thesis.

First, the application of the 2D elastic full-waveform inversion of individual Rayleigh and Love waves is demonstrated to detect a buried structure of archaeological interest that is located near Karlsruhe, Germany. The time-frequency windowing approach was applied to the field data set where it allowed the reduction of the misfit function and retrieval of models of higher correlation, compared to the conventional FWI. For the first time, an accurate P-wave velocity model was obtained when using short-offset seismic surveys. This was further evaluated by superimposing the velocity structure with the migrated results of a ground-penetrating measurements that were performed on the site. The SH FWI provided exceptional results indicating once again its strength versus the P-SV FWI. Several limitations prevented the application of the three-component FWI, due to coupling issues between the vertical and horizontal inline components.

The second application was specifically designed in a way to avoid the coupling issues that were faced before. A prismatic source enabled proper coupling during acquisition of seismic data and enabled the joint inversion of all three-component data simultaneously. The target area was an aquifer system, where so far seismic techniques were not particularly employed. This was due to the poor resolution of the reconstructed models but also the lack of proper petrophysical relations. With the developments, however, that were shown in regards not only to the S-wave velocity but also the density and P-wave velocity model, a new window of opportunity had opened.

The survey included two profiles, crossing each other perpendicularly and aligned with several boreholes where other geophysical measurements have been employed in the past. The layout enabled additionally to study the continuity of the structures obtained by the 2D planes. The retrieval of the P-wave velocity was a challenging task, due to the low contrast values of the velocities that were present in the subsurface. Furthermore, the groundwater table being localized close to the surface and thereby influencing the P-wave velocity further complicated P-wave velocity retrieval. The inversion of the horizontal crossline component data proved to be the most challenging contrary to the results from the previous field study and the synthetic tests. The cause of this was the high dispersive nature of the Love wave in the examined test site, which shows that there is no global workflow one can adapt at every subsurface study. Flexibility and careful observation is a key to properly adapt the FWI workflow at different subsurface conditions. However, the simultaneous joint inversion of multicomponent data can help to provide more reliable results, since the amount of data versus the unknown elastic parameters is increased. For the first time, a high correlation of the density model from shallow seismic FWI with in-situ borehole information was obtained. Spatial continuity was achieved especially in regard to the S-wave velocity and density models. The complex structure of the aquifer was revealed and can now serve as a valuable source of information for further hydrogeophysical site characterization.

Outlook

In this section, I provide a framework for future developments that are required to further establish FWI as the preferred approach for delineating high resolution subsurface models in near-surface applications.

Due to the different signature of the scattering response caused by density perturbations, some implications arise that could improve the reconstruction of density from FWI. The trace-wise normalization that is required for a stable inversion, especially in field data applications, reduces the ability of FWI to retrieve the true absolute values of the density model. To avoid this issue, one could employ an offset dependent taper to retain the relative contributions of amplitudes versus offset which is present in the recorded data. This taper needs to be updated at the various stages of FWI and match the information in the data due to the increasing frequency band which is introduced.

In the absence of an accurate initial model for the P-wave velocity the Love wave FWI should be used to obtain the v_s model. The main reason is its in-dependency from the P-wave velocity, which also further decreases the number of model parameters which are unknown. The SH wave equation is less complex than the P-SV wave equation, which allows a computationally efficient inversion. However, in cases where the Love wave is highly dispersive, multicomponent FWI should be preferred instead.

The time-frequency windowing approach can allow the accurate reconstruction of the P-wave velocity model, however careful selection of the damping factors per stage is required. When properly employed, it will allow the reduction of the parameter cross-talk which is of paramount importance for retrieving the true subsurface model.

To summarize, the application of seismic FWI in near-surface surveys has been thoroughly examined. Different aspects of the inversion workflow were thoroughly presented; from the proper acquisition of multicomponent data to the sophisticated inversion strategies, that account for both the numerical and physical aspects of the inverse problem, to the evaluation and retrieval of high resolution subsurface models. The results establish the application of seismic FWI as the recommended method for obtaining the elastic parameters of the subsurface.

Appendix A

Individual scattering response for various inclusion cases

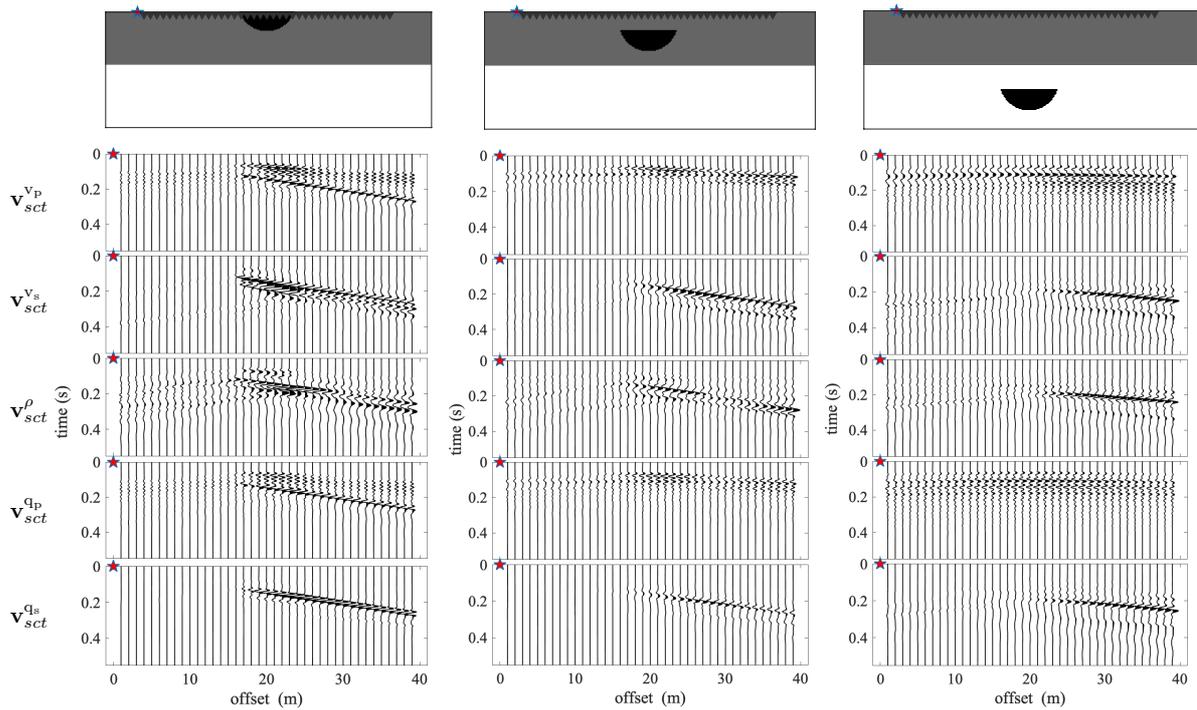


Figure A.1: Scattered wavefields caused by perturbations of the viscoelastic parameters for various depths of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.

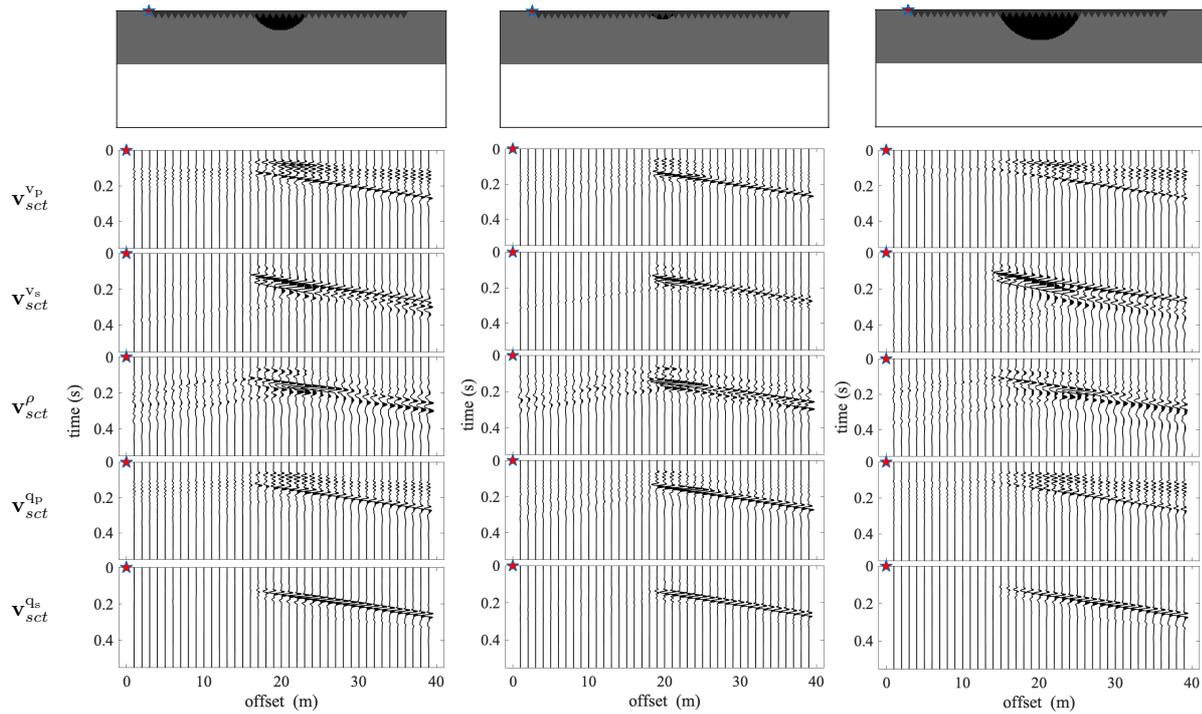


Figure A.2: Scattered wavefields caused by perturbations of the viscoelastic parameters for various sizes of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for the first column and then are halved for the second and doubled for the third column, for each parameter accordingly.

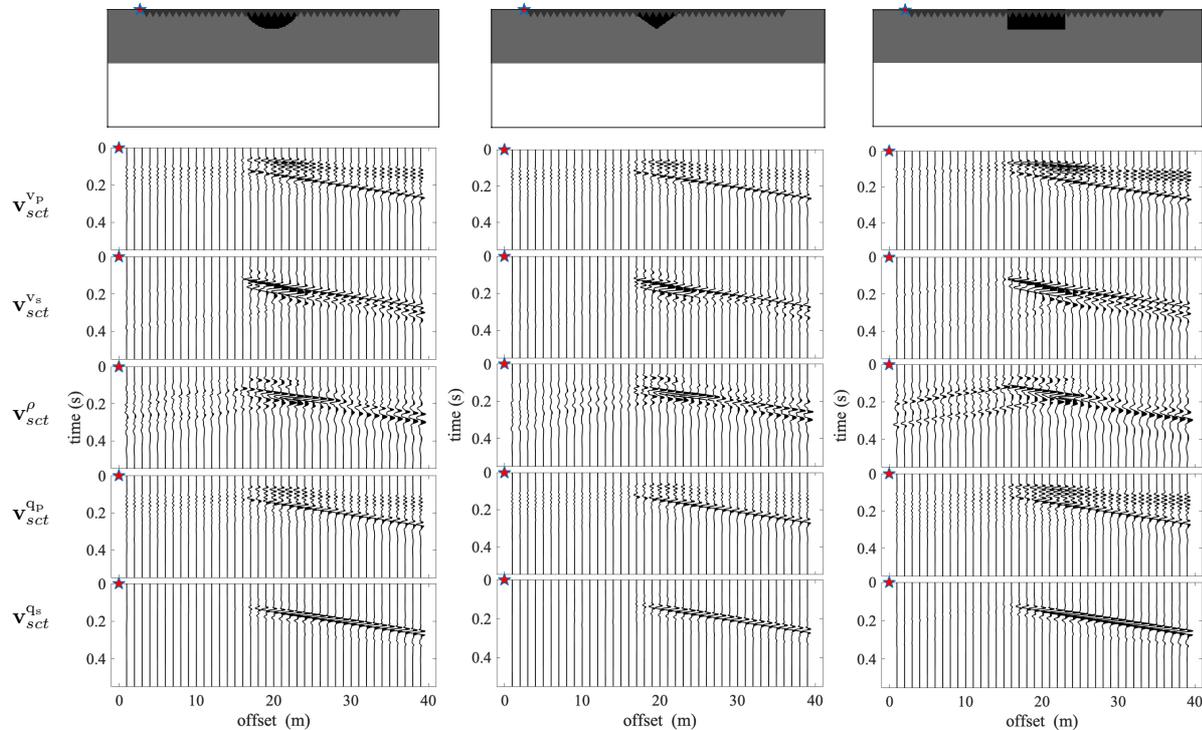


Figure A.3: Scattered wavefields caused by perturbations of the viscoelastic parameters for various shapes of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.

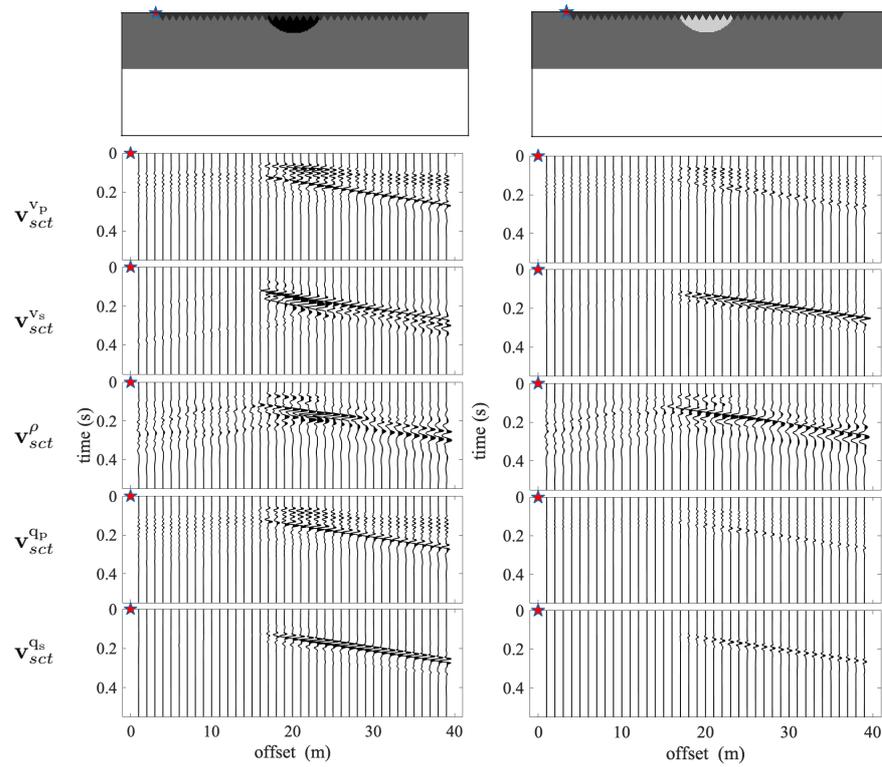


Figure A.4: Scattered wavefields caused by perturbations of the viscoelastic parameters for different values of the inclusion. The values for the first column are given in table 3.2, while for the second column the sign of the percentage added is shifted. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.

Appendix B

Cross-gradient implementation

At this section, I present a manual of the cross gradients into the IFOS2D code developed and maintained by our group (<https://git.scc.kit.edu/GPIAG-Software/IFOS2D>), for both usage and further future implementations. The parameter classes are given as:

```
"CROSS_GRAD" : "1",
"CGRAD_TYPE" : "1",
"CGRAD_NORM" : "1",
"NORMALIZE_CG_NORM" : "1",
"CGRAD_CALC" : "1",
"GAMMA_CGRAD" : "1",
"CGRAD_REG" : "1",
"JOINT_INVERSION_CG" : "1",
    "CGRAD_NORM" : "1",
"JOINT_INVERSION_CGRAD_ALPHA_VP" : "1",
"JOINT_INVERSION_CGRAD_ALPHA_VS" : "1",
"JOINT_INVERSION_CGRAD_ALPHA_RHO" : "1",
"CROSS_GRADIENT_MODEL" : "./model/cross_grad/cgmod",
```

Default value is:

CROSS_GRAD=0

By enabling the variable CROSS_GRAD, the cross gradient calculation will be performed during 2D elastic FWI. With the variable CGRAD_TYPE it is possible to define between which parameters the cross gradient constraints will be applied. There are currently 8 options available which are shown in Table B.1 (type 1, 4 and 8 are thoroughly tested). The variable CGRAD_NORM corresponds to the objective function which is used (currently L2norm). The parameter CGRAD_CALC represents a regularization term of the cross-gradient in the joint objective function (equation 4.13, parameter γ). With the variable NORMALIZE_CG_NORM the two norms are normalized and then γ represents the percentage of information passed in the objective function from the E_{cg} in equation 4.13.

The CGRAD_REG term allows to apply a 2D median filter or an external taper to the update of the cross-gradient. The parameter JOINT_INVERSION_CG needs to be set to 1 for estimating the joint gradients. If the variable CGRAD_NORM is set other than zero, three options are given for

Table B.1: All configurations for passing structural information from one model parameter to another.

CGRAD_TYPE	Structural constraint
1	v_s from v_p
2	ρ from v_p
3	ρ from v_s
4	v_s and ρ from v_p
5	ρ and v_s from v_p
6	ρ from v_s and v_p from v_s
7	all three parameters with one another
8	from external file to v_p , v_s and ρ

normalization of the seismic and cross-gradient gradients. `CGRAD_NORM==1` equation 4.14 is used, with `CGRAD_NORM==2` the normalization by the maximum gradient amplitudes is neglected and with `CGRAD_NORM==3` the correct amplitude is maintained for specific optimizations such is the case of LBFGS while the gradients are contributing equally (equation 4.14, variable α is ignored). The values `JOINT_INVERSION_CGRAD_ALPHA_VS`, `JOINT_INVERSION_CGRAD_ALPHA_VP` and `JOINT_INVERSION_CGRAD_ALPHA_RHO` correspond to the values of α in equation 4.14 for each of the three parameters v_s , v_p and ρ , respectively. With the variable `CROSS_GRADIENT_MODEL` an external model that we wish to structurally couple with the seismic parameters can be given.

Appendix C

Steel prism

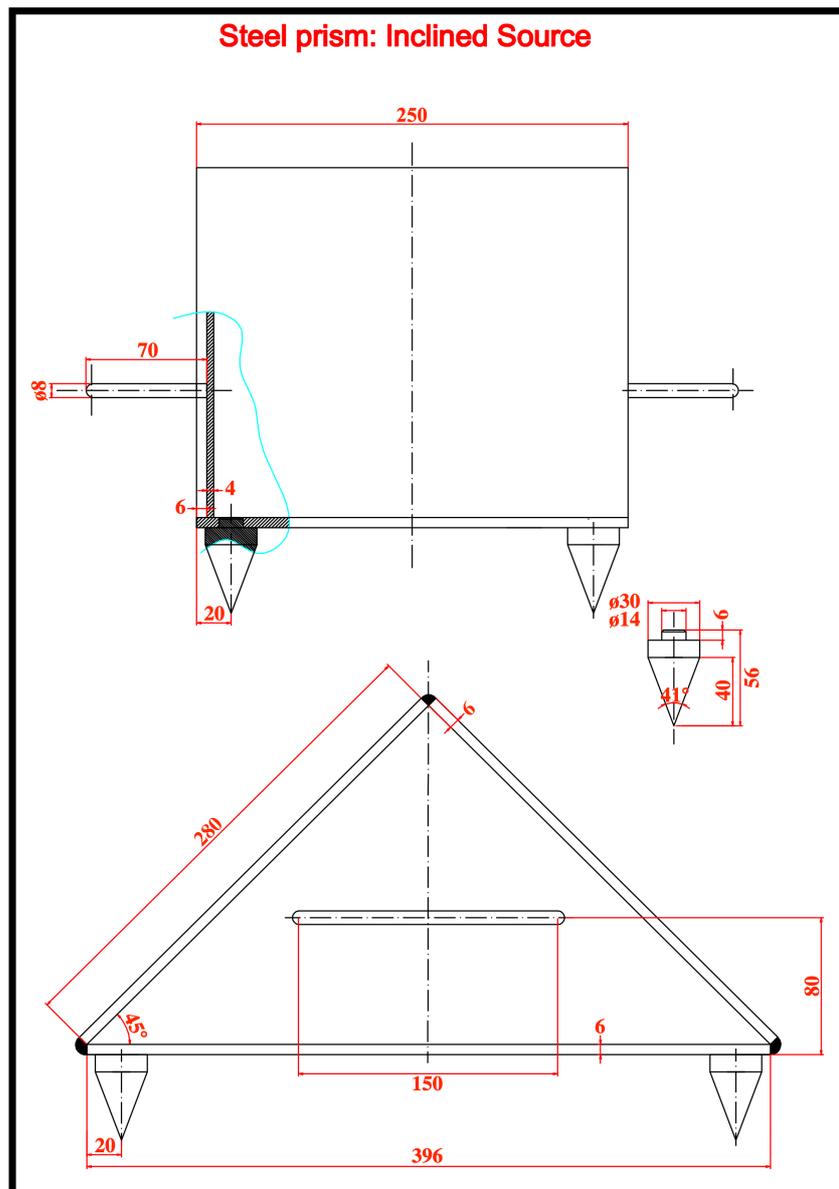


Figure C.1: Mechanical design of the steel prism. The dimensions are shown in mm.

Appendix D

Dispersion curves: profiles P1 and P2

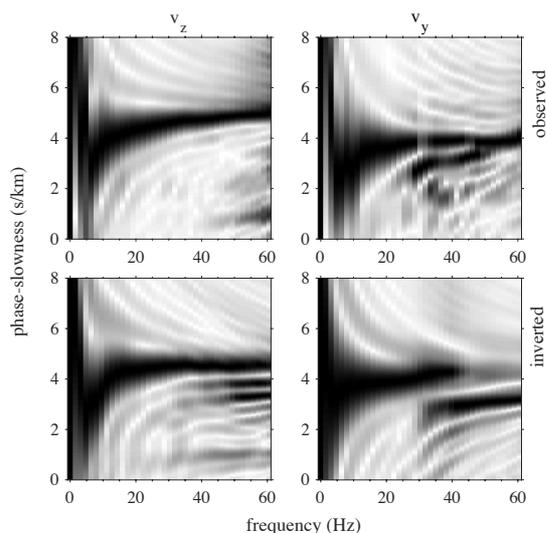


Figure D.1: Dispersion curves extracted from the data of profile 1. Phase-velocity over frequency calculated from the 1st shot of the observed data and inverted data (top and bottom rows, respectively) for the vertical (v_z) and horizontal crossline (v_y) components (left and right columns, respectively). The dispersion image of v_z reveals the fundamental mode of the Rayleigh wave, while the v_y component the fundamental mode of the Love wave, as well as the existence of a potential higher mode.

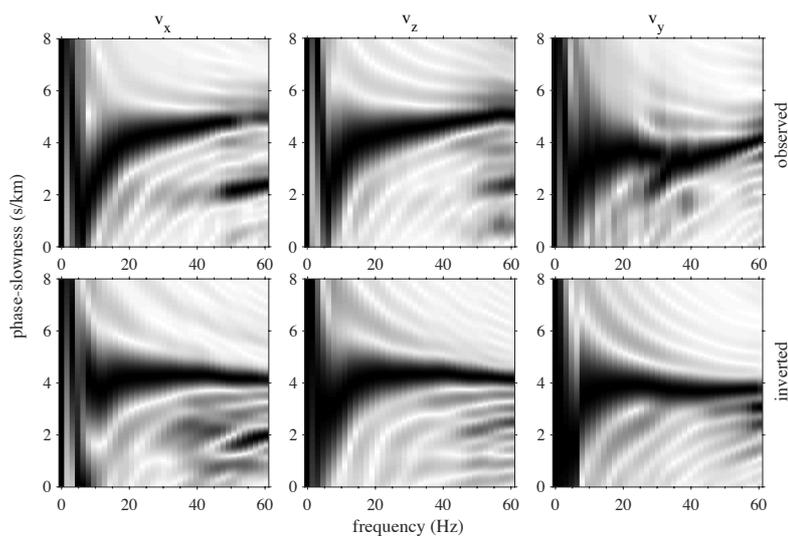


Figure D.2: Dispersion curves extracted from the data of profile 2. Phase-velocity over frequency calculated from the 1st shot of the observed data and inverted data (top and bottom rows, respectively) for the horizontal inline (v_x), vertical (v_z) and horizontal crossline (v_y) components (left and right columns, respectively). The dispersion images of v_x and v_z reveal the fundamental mode of the Rayleigh wave, while the v_y component the fundamental mode of the Love wave, as well as the existence of a potential higher mode.

Appendix E

Software and Hardware

Most of the software listed below was installed on a workstation with macOS operating system. This thesis was written with LATEX. The majority of the figures were produced by MatLab (technical figures) and Inkscape (sketches).

The software IFOS2D was used for the FWI of both synthetic and field data. It is currently maintained by me and the rest of the group of the Geophysical Institute of the Karlsruhe Institute of Technology (KIT) and is freely available under the GNU license at <https://git.scc.kit.edu/GPIAG-Software/IFOS2D>.

Most of the preprocessing of the field data, as well as the analysis of the inversion results were done with the MATLAB (MathWorks) and Seismic Un*x (Stockwell, 1999). The 3D to 2D transformation was performed with the software *lisousi* developed by Thomas Forbriger. For the inversion of the source-time function, the program *soutifu* was used and it is embedded in the IFOS2D code, provided also by Thomas Forbriger.

The inversion results were calculated on several high-performance-computing (HPC) systems, which I specifically mention in the section below.

Acknowledgments

I gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS (Jülich Supercomputing Centre, 2019) at Jülich Supercomputing Centre (JSC). Additional simulations were performed by the Euler supercomputing cluster at CSCS in Lugano.

Bibliography

- Aki, K. and Richards, P. G. (2002). *Quantitative seismology*, volume 1.
- Athanasopoulos, N. and Bohlen, T. (2016). Sequential full-waveform inversion of refracted and rayleigh waves. In *Near Surface Geoscience 2016-22nd European Meeting of Environmental and Engineering Geophysics*. Earthdoc.
- Athanasopoulos, N. and Bohlen, T. (2017a). Aquifer characterization using elastic full-waveform inversion. In *Near Surface Geoscience 2017-23rd European Meeting of Environmental and Engineering Geophysics*. Earthdoc.
- Athanasopoulos, N. and Bohlen, T. (2017b). Field data application of sequential full-waveform inversion of refracted and rayleigh waves. In *79th EAGE Conference and Exhibition 2017*. Earthdoc.
- Athanasopoulos, N. and Bohlen, T. (2019). Shallow-seismic wavefield scattering and implications for viscoelastic FWI. In *Near Surface Geoscience 2019-25th European Meeting of Environmental and Engineering Geophysics*. Earthdoc.
- Athanasopoulos, N., Manukyan, E., Bohlen, T., and Maurer, H. (2018a). Accurate reconstruction of shallow p-wave velocity model with time-windowed elastic full-waveform inversion. In *80th EAGE Conference and Exhibition 2018*. Earthdoc.
- Athanasopoulos, N., Manukyan, E., Bohlen, T., and Maurer, H. (2018b). Application of structurally similarity constraints on time-domain elastic full-waveform inversion. In *Near Surface Geoscience 2018-24th European Meeting of Environmental and Engineering Geophysics*. Earthdoc.
- Athanasopoulos, N., Manukyan, E., Bohlen, T., and Maurer, H. (2020). Time–frequency windowing in multiparameter elastic FWI of shallow seismic wavefield. *Geophysical Journal International*, 222(2):1164–1177.
- Beylkin, G. and Burridge, R. (1990). Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion*, 12(1):15–52.
- Biot, M. A. (1956). Theory of propagation of elastic waves in a fluid-saturated porous solid. i. low-frequency range. *The Journal of the Acoustical Society of America*, 28(2):168–178.
- Blanch, J. O., Robertsson, J. O., and Symes, W. W. (1995). Modeling of a constant Q: Methodology and algorithm for an efficient and optimally inexpensive viscoelastic technique. *Geophysics*, 60(1):176–184.
- Bleibinhaus, F., Hole, J. A., Ryberg, T., and Fuis, G. S. (2007). Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging. *Journal of Geophysical Research: Solid Earth*, 112(B6).

- Blonk, B., Herman, G. C., and Drijkoningen, G. G. (1995). An elastodynamic inverse scattering method for removing scattered surface waves from field data. *Geophysics*, 60(6):1897–1905.
- Bohlen, T. (2002). Parallel 3-D viscoelastic finite difference seismic modelling. *Computers & Geosciences*, 28(8):887–899.
- Bohlen, T., Kugler, S., Klein, G., and Theilen, F. (2004). 1.5 D inversion of lateral variation of Scholte-wave dispersion. *Geophysics*, 69(2):330–344.
- Bohlen, T., Müller, C., and Milkereit, B. (2012). 5. *Elastic Seismic-Wave Scattering from Massive Sulfide Orebodies: On the Role of Composition and Shape*, pages 70–89.
- Bolt, B. A. and Freeman, W. H. (1989). Earthquakes. *Earthquake Engineering & Structural Dynamics*, 18(3):451–452.
- Brenders, A. J. (2011). *Strategies for Waveform Tomography of Long-Offset, 2-D Exploration Seismic Data*. PhD thesis, The University of Western Ontario.
- Brenders, A. J. and Pratt, R. G. (2007). Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophysical Journal International*, 168(1):152–170.
- Brossier, R., Operto, S., and Virieux, J. (2009). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):105–118.
- Brossier, R., Operto, S., and Virieux, J. (2015). Velocity model building from seismic reflection data by full-waveform inversion. *Geophysical Prospecting*, 63(2):354–367.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Carcione, J. M., Kosloff, D., and Kosloff, R. (1988). Wave propagation simulation in a linear viscoelastic medium. *Geophysical Journal International*, 95(3):597–611.
- Choi, Y. and Alkhalifah, T. (2012). Application of multi-source waveform inversion to marine streamer data using the global correlation. *Geophysical Prospecting*, 60:748–758.
- Choi, Y. and Alkhalifah, T. (2018). Time-domain full-waveform inversion of exponentially damped wavefield using the deconvolution-based objective function. *GEOPHYSICS*, 83(2):R77–R88.
- Courant, R., Friedrichs, K., and Lewy, H. (1928). Über die partiellen Differenzgleichungen der mathematischen Physik. *Mathematische Annalen*, 100(1):32–74.
- Dagan, G. (1989). Flow and transport in porous formations. *Springer, New York*.
- Doetsch, J., Linde, N., Coscia, I., Greenhalgh, S. A., and Green, A. G. (2010). Zonation for 3d aquifer characterization based on joint inversions of multimethod crosshole geophysical data. *Geophysics*, 75(6):G53–G64.
- Dokter, E., Köhn, D., Wilken, D., and Rabbel, W. (2014). Application of Elastic 2D Waveform Inversion to a Near Surface SH-wave Dataset. In *76th EAGE Conference and Exhibition 2014*. Earthdoc.
- Döring, U. (1997). *Transport der reaktiven Stoffe Eosin, Uranin und Lithium in einem heterogenen Grundwasserleiter*. PhD thesis, Christian Albrecht Universität Kiel.

- Dougherty, M. E. and Stephen, R. A. (1988). Seismic energy partitioning and scattering in laterally heterogeneous ocean crust. *Pure and Applied Geophysics*, 128(1):195–229.
- Dziewonski, A., Bloch, S., and Landisman, M. (1969). A technique for the analysis of transient seismic signals. *Bulletin of the Seismological Society of America*, 59(1):427.
- Eaton, D. W. (1999). Weak elastic-wave scattering from massive sulfide orebodies. *Geophysics*, 64(1):289–299.
- Emmerich, H. and Korn, M. (1987). Incorporation of attenuation into time-domain computations of seismic wave fields. *Geophysics*, 52(9):1252–1264.
- Ernst, J. R., Green, A. G., Maurer, H., and Holliger, K. (2007). Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data. *Geophysics*, 72(5):53–64.
- Fabien-Ouellet, G., Gloaguen, E., and Giroux, B. (2017). Time domain viscoelastic full waveform inversion. *Geophysical Journal International*, 209(3):1718–1734.
- Fichtner, A., Kennett, B. L., Igel, H., and Bunge, H.-P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, 179(3):1703–1725.
- Forbriger, T. (2003a). Inversion of shallow-seismic wavefields: I. Wavefield transformation. *Geophysical Journal International*, 153(3):719–734.
- Forbriger, T. (2003b). Inversion of shallow-seismic wavefields: II. Inferring subsurface properties from wavefield transforms. *Geophysical Journal International*, 153(3):735–752.
- Forbriger, T., Groos, L., and Schäfer, M. (2014). Line-source simulation for shallow-seismic data. Part 1: Theoretical background. *Geophysical Journal International*, 198(3):1387–1404.
- Foti, S., Hollender, F., Garofalo, F., Albarello, D., Asten, M., Bard, P.-Y., Comina, C., Cornou, C., Cox, B., Di Giulio, G., Forbriger, T., Hayashi, K., Lunedei, E., Martin, A., Mercerat, D., Ohrnberger, M., Poggi, V., Renalier, F., Sicilia, D., and Socco, V. (2018). Guidelines for the good practice of surface wave analysis: a product of the interpacific project. *Bulletin of Earthquake Engineering*, 16(6):2367–2420.
- Gallardo, L. A. and Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2d inversion of dc resistivity and seismic data. *Geophysical Research Letters*, 30(13).
- Gardner, G., Gardner, L., and Gregory, A. (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.
- Gassner, L., Gerach, T., Hertweck, T., and Bohlen, T. (2019). Seismic characterization of submarine gas-hydrate deposits in the western black sea by acoustic full-waveform inversion of ocean-bottom seismic data. *Geophysics*, 84(5):311–324.
- Gélis, C., Virieux, J., and Grandjean, G. (2007). Two-dimensional elastic full waveform inversion using Born and Rytov formulations in the frequency domain. *Geophysical Journal International*, 168(2):605–633.
- Gibson, B. S., Odegard, M. E., and Sutton, G. H. (1979). Nonlinear least-squares inversion of traveltime data for a linear velocity-depth relationship. *Geophysics*, 44(2):185–194.

- Gray, S. H. and Marfurt, K. J. (1995). Migration from topography: Improving the near-surface image. *Canadian journal of exploration geophysics*, 31(1):18–24.
- Groos, L. (2013). *2D full waveform inversion of shallow-seismic Rayleigh waves*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Groos, L., Schäfer, M., Forbriger, T., and Bohlen, T. (2014). The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves. *Geophysics*, 79(6):R247–R261.
- Groos, L., Schäfer, M., Forbriger, T., and Thomas, B. (2017). Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves. *Geophysics*, 82(2):R109–R117.
- Gubernatis, J. E., Domany, E., Krumhansl, J. A., and Huberman, M. (1977). The born approximation in the theory of the scattering of elastic waves by flaws. *Journal of Applied Physics*, 48(7):2812–2819.
- Gueting, N., Klotzsche, A., van der Kruk, J., Vanderborght, J., Vereecken, H., and Englert, A. (2015). Imaging and characterization of facies heterogeneity in an alluvial aquifer using gpr full-waveform inversion and cone penetration tests. *Journal of Hydrology*, 524:680 – 695.
- Gueting, N., Vienken, T., Klotzsche, A., van der Kruk, J., Vanderborght, J., Caers, J., Vereecken, H., and Englert, A. (2017). High resolution aquifer characterization using crosshole gpr full-waveform tomography: Comparison with direct-push and tracer test data. *Water Resources Research*, 53(1):49–72.
- Heider, S. (2014). *2D Elastic Full-Waveform Tomography of Vibro-Seismic Data in Crystalline Host Rock at the GFZ-Underground-Lab, Freiberg*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Heisey J.S., Stokoe II K.H., M. A. (1982). Moduli of pavement systems from spectral analysis of surface waves. *Transp. Res. Rec.*, 852:22–31.
- Hüttner, R., Konrad, H.-J., and Zitzmann, A. (1968). Geologische Übersichtskarte 1:200000, Blatt CC7110 Mannheim. *Bundesanstalt für Geowissenschaften und Rohstoffe in Zusammenarbeit mit den Geologischen Landesämtern der Bundesrepublik Deutschland und benachbarter Staaten*.
- Häusler, M., Schmelzbach, C., and Sollberger, D. (2018). The galperin source: A novel efficient multicomponent seismic source. *Geophysics*, 83(6):P19–P27.
- Jülich Supercomputing Centre (2019). JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre. *Journal of large-scale research facilities*, 5(A135).
- Klotzsche, A., van der Kruk, J., Linde, N., Doetsch, J., and Vereecken, H. (2013). 3-d characterization of high-permeability zones in a gravel aquifer using 2-d crosshole gpr full-waveform inversion and waveguide detection. *Geophysical Journal International*, 195(2):932–944.
- Knopoff, L. (1972). Observation and inversion of surface-wave dispersion. 4:497–519.
- Köhn, D. (2011). *Time domain 2D elastic full waveform tomography*. PhD thesis, Christian-Albrechts-Universität zu Kiel.
- Köhn, D., De Nil, D., Kurzmann, A., Przebindowska, A., and Bohlen, T. (2012). On the influence of model parametrization in elastic full waveform tomography. *Geophysical Journal International*, 191(1):325–345.

- Köhn, D., Wilken, D., Nil, D. D., Wunderlich, T., Rabbel, W., Werther, L., Schmidt, J., Zielhofer, C., and Linzen, S. (2019). Comparison of time-domain sh waveform inversion strategies based on sequential low and bandpass filtered data for improved resolution in near-surface prospecting. *Journal of Applied Geophysics*, 160:69–83.
- Komatitsch, D. and Martin, R. (2007). An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, 72(5):SM155–SM167.
- Krieger, D. (2019). 2d elastic full-waveform inversion of land seismic data with topographic variations. Master's thesis, Karlsruhe Institute of Technology (KIT).
- Kurzmann, A. (2012). *Applications of 2D and 3D full waveform tomography in acoustic and viscoacoustic complex media*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In *Conference on Inverse scattering, Theory and application, Society for Industrial and Applied Mathematics*, 206–220.
- Lang, K. (1907). Die Ettlinger Linien und ihre Geschichte. *Veröffentlichungen des Karlsruher Altertumsvereins*.
- Lay, T. and Wallace, T. C. (1995). *Modern global seismology*, volume 58. Academic press.
- Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- Levander, A. R. (1990). Seismic scattering near the Earth's surface. *Pure and Applied Geophysics*, 132(1):21–47.
- Liu, H.-P., Anderson, D. L., and Kanamori, H. (1976). Velocity dispersion due to anelasticity; implications for seismology and mantle composition. *Geophysical Journal International*, 47(1):41–58.
- Liu, T. and Bohlen, T. (2020). An appetizer for the shallow seismic fwi in fluid-saturated poroelastic media: Sensitivity analysis. In *2020 SEG/SPWLA 7th Workshop on Porous Media: Structure, Flow, and Dynamics*.
- Manukyan, E., Latzel, S., Maurer, H., Marelli, S., and Greenhalgh, S. A. (2012). Exploitation of data-information content in elastic-waveform inversions. *Geophysics*, 77(2):R105–R115.
- Manukyan, E. and Maurer, H. (2020). Elastic vertically transversely isotropic full-waveform inversion using cross-gradient constraints — an application toward high-level radioactive waste repository monitoring. *Geophysics*, 85(4):R313–R323.
- Manukyan, E., Maurer, H., and Nuber, A. (2018). Improvements to elastic full-waveform inversion using cross-gradient constraints. *Geophysics*, 83(2):R105–R115.
- Maurer, H., Curtis, A., and Boerner, D. E. (2010). Recent advances in optimized geophysical survey design. *Geophysics*, 75(5):75A177–75A194.
- McMechan, G. A. and Yedlin, M. J. (1981). Analysis of dispersive waves by wave field transformation. *Geophysics*, 46(6):869–874.
- Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52(9):1211–1228.

- Morgan, J., Warner, M., Bell, R., Ashley, J., Barnes, D., Little, R., Roele, K., and Jones, C. (2013). Next-generation seismic experiments: wide-angle, multi-azimuth, three-dimensional, full-waveform inversion. *Geophysical Journal International*, 195(3):1657–1678.
- Morse, P. and Feshbach, H. (1953). *Methods of theoretical physics*. International series in pure and applied physics. McGraw-Hill.
- Müller, G. and Weber, M. (2007). Theory of elastic waves. <http://hdl.handle.net/11858/00-1735-0000-0001-32BE-0>.
- Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.
- Nuber, A., Manukyan, E., and Maurer, H. (2015). Enhancement of near-surface elastic full waveform inversion results in regions of low sensitivities. *Journal of Applied Geophysics*, 122:192 – 201.
- Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A., and Dell’Aversana, P. (2004). Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale travelttime and waveform inversions: a case study. *Geophysical Prospecting*, 52(6):625–651.
- Pan, Y., Xia, J., Xu, Y., Gao, L., and Xu, Z. (2016). Love-wave waveform inversion in time domain for shallow shear-wave velocity. *Geophysics*, 81(1):R1–R14.
- Park, C. B., Miller, R. D., and Xia, J. (1999). Multichannel analysis of surface waves. *Geophysics*, 64(3):800–808.
- Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- Plessix, R.-E. and Mulder, W. (2004a). Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157(3):975–987.
- Plessix, R.-E. and Mulder, W. A. (2004b). Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157(3):975–987.
- Polak, E. and Ribière, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue française d’informatique et de recherche opérationnelle, série rouge*, 3(1):35–43.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901.
- Pratt, R. G., Huang, L., Duric, N., and Littrup, P. (2007). Sound-speed and attenuation imaging of breast tissue using waveform tomography of transmission ultrasound data. In Hsieh, J. and Flynn, M. J., editors, *Medical Imaging 2007: Physics of Medical Imaging*, volume 6510, pages 1523–1534. International Society for Optics and Photonics, SPIE.
- Robertsson, J. O., Blanch, J. O., and Symes, W. W. (1994). Viscoelastic finite-difference modeling. *Geophysics*, 59(9):1444–1456.
- Robertsson, J. O. A. and Chapman, C. H. (2000). An efficient method for calculating finite-difference seismograms after model alterations. *GEOPHYSICS*, 65(3):907–918.
- Romdhane, A., Grandjean, G., Brossier, R., Rejiba, F., Operto, S., and Virieux, J. (2011). Shallow-structure characterization by 2D elastic full-waveform inversion. *Geophysics*, 76(3):R81–R93.

- Schäfer, M. (2014). *Application of full-waveform inversion to shallow-seismic Rayleigh waves on 2D structures*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Schäfer, M., Groos, L., Forbriger, T., and Bohlen, T. (2014). Line-source simulation for shallow-seismic data. Part 2: full-waveform inversion—a synthetic 2-D case study. *Geophysical Journal International*, 198(3):1405–1418.
- Schmelzbach, C., Sollberger, D., Greenhalgh, S., Horstmeyer, H., Maurer, H., and Robertsson, J. (2016). 9c seismic data acquisition for near-surface applications: recording, waveform reciprocity and 4c rotation. 2016(1):1–5.
- Schön, J. (1996). *Physical Properties of Rocks: Fundamentals and Principles of Petrophysics*. Pergamon.
- Socco, L. V., Foti, S., and Boiero, D. (2010). Surface-wave analysis for building near-surface velocity models—established approaches and new perspectives. *Geophysics*, 75(5):75A83–75A102.
- Stockwell, J. (1999). The cwp/su: Seismic un*x package,. *Computers & Geosciences*, 25:415–419.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (1986). A strategy for nonlinear elastic inversion of seismic reflection data. *Geophysics*, 51(10):1893–1903.
- Telford, W. M., Geldart, L. P., and Sheriff, R. E. (1990). *Applied Geophysics*. Cambridge University Press, 2 edition.
- Tillmann, A., Englert, A., Nyari, Z., Fejes, I., Vanderborght, J., and Vereecken, H. (2008). Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the krauthausen test site using cone penetration test. *Journal of Contaminant Hydrology*, 95(1):57 – 75.
- Tran, K. T., McVay, M., Faraone, M., and Horhota, D. (2013). Sinkhole detection using 2D full seismic waveform tomography. *Geophysics*, 78(5):R175–R183.
- Tryggvason, A. and Linde, N. (2006). Local earthquake (le) tomography with joint inversion for p- and s-wave velocities using structural constraints. *Geophysical Research Letters*, 33(7).
- UNESCO, International Hydrological Programme FR, U.-I. (2005). Recharge systems for protecting and enhancing groundwater resources. In *IHP-VI, Series on Groundwater No. 13, Proceedings of the 5th International Symposium on Management of Aquifer Recharge ISMAR5, Berlin, Germany*.
- Vereecken, H., Döring, U., Hardelauf, H., Jaekel, U., Hashagen, U., Neuendorf, O., Schwarze, H., and Seidemann, R. (2000). Analysis of solute transport in a heterogeneous aquifer: the krauthausen field experiment. *Journal of Contaminant Hydrology*, 45(3):329 – 358.
- Virieux, J. (1984). SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 49(11):1933–1942.
- Virieux, J. (1986). P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.
- Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.

- Wehner, D., Köhn, D., De Nil, D., Schmidt, S., al Hagrey, S., and Rabbel, W. (2015). A Combined Elastic Waveform and Gravity Inversion for Improved Density Model Resolution Applied to the Marmousi-II Model. In *77th EAGE Conference and Exhibition 2015*. Earthdoc.
- Wittkamp, F., Athanasopoulos, N., and Bohlen, T. (2018). Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves. *Geophysical Journal International*, 216(1):350–364.
- Wu, R. and Aki, K. (1985). Scattering characteristics of elastic waves by an elastic heterogeneity. *Geophysics*, 50(4):582–595.
- Wu, R.-S. (1989). The perturbation method in elastic wave scattering. *Pure and Applied Geophysics*, 131:605–637.
- Xia, J., Miller, R. D., and Park, C. B. (1999). Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, 64(3):691–700.
- Xu, S., Wang, D., Chen, F., Lambaré, G., and Zhang, Y. (2012). *Inversion on Reflected Seismic Wave*, pages 1–7.
- Zhang, S. X. and Chan, L. S. (2003). Possible effects of misidentified mode number on rayleigh wave inversion. *Journal of Applied Geophysics*, 53(1):17–29.

List of Figures

1.1	FWI applications over the last 20 years including a broad range of scales. The different applications are expressed as a function of complexity, spatial scale and wave types that are considered during the inversion. The color coding is shown in analogy to wavelengths of the optical electromagnetic spectrum.	2
2.1	Types of seismic waves and particle motion as function of propagation time: a) compressional (P) wave traveling in a block of material, b) vertical polarized and c) horizontal polarized shear wave traveling in a block of material, d) Rayleigh wave traveling in a section of the Earth's surface, e) Love wave traveling along a section of the Earth's surface (modified after Bolt and Freeman, 1989).	11
2.2	Distribution of viscoelastic parameters on a standard staggered-grid for P-SV waves (left) and SH waves (right).	15
2.3	The effects of a multiscale approach on a synthetic seismogram in the time and frequency domain. a) Synthetic seismic reference trace. The arrivals of different wave types are highlighted. b) Seismogram after application of bandpass filters (BP). c) Corresponding amplitude spectra.	26
2.4	Schematic overview for one iteration of the FWI process. The dark green box corresponds to the loop over the number of iterations, while the light green box indicates the individual steps for each source.	27
3.1	Sketch of seismic scattering due to shallow heterogeneities.	30
3.2	The radiation patterns of v_p , v_s and ρ perturbations for different modes.	31
3.3	Calculation of the scattering response of a shallow local anomaly by subtracting the seismic response of the background model from the perturbed, i.e. total wavefield.	33
3.4	Left: A cubed sine wavelet defined by equation 3.9 with a center frequency of 30 Hz. Right: Its amplitude spectrum in the range from 0 Hz to 70 Hz.	35
3.5	Snapshots (left two columns) of the S-wave (E_s) and P-wave (E_p) intensity of the background wavefield (top row) and scattered wavefields caused by perturbations of the viscoelastic parameters (Table 3.2). The corresponding seismograms of vertical particle velocity are shown in the right column. Triangles in the top-most figures mark the positions of some receivers and the stars indicate the shot location. For better visualization the amplitudes of the scattered wavefield are clipped to lower values as indicated by the colorbars. The percentages in the right column give the relative maximum amplitude of the scaled seismograms compared to the background (a.3).	37

3.6	Scattered wavefields caused by perturbations of the viscoelastic parameters for various depths of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b. . . .	39
3.7	Scattered wavefields caused by perturbations of the viscoelastic parameters for various sizes of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b for the first column and then are halved for the second and doubled for the third column.	39
3.8	Scattered wavefields caused by perturbations of the viscoelastic parameters for various shapes of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b. . . .	39
3.9	Scattered wavefields caused by perturbations of the viscoelastic parameters for different values of the inclusion. The values are given in Table 3.2, while a sketch of the model is shown on the top row. Triangles mark the positions of some receivers and the stars indicate the shot location. The P- and S-wave intensities are scaled with the same values as Figure 3.5b. . . .	41
3.10	Seismograms showing the scattering response caused by perturbations of the viscoelastic parameters for different values of the inclusion. The two different compositions from the models in Figure A.4 are shown as red for negative and black for a positive change. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly. .	41
4.1	Models obtained by the inversion of synthetic data for a simple spatially correlated near-surface model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the final results of the Rayleigh wave (P-SV), Love wave (SH) and joint FWI. Red asterisks represent the source locations.	48
4.2	Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI), Love wave (SHFWI) and joint FWI shown in Figure 4.1. Top: Comparison of velocity seismograms for the horizontal inline (v_x), vertical (v_z) and crossline (v_y) components for 6 different offsets. Bottom: Misfit function for each of the inversions (filled circles). The dashed red and blue lines constitute the contribution of Rayleigh and Love waves, respectively, to the joint misfit function (yellow circles). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 4.1).	49
4.3	1D depth profiles of the elastic parameters v_p , v_s and ρ for the true and the initial model. . .	51
4.4	Models obtained by the inversion of synthetic data for a complex spatially uncorrelated near-surface model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the final results of the Rayleigh wave (P-SV), Love wave (SH) and joint FWI. Red asterisks represent the source locations.	52
4.5	1D models of the true, initial and reconstructed parameters shown in Figure 4.4. The first row shows vertical profiles of the true, initial and reconstructed models at the location of the shallow anomaly of each parameter. The second row shows vertical profiles through the background medium.	52

4.6	Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI), Love wave (SH FWI) and joint FWI shown in Figure 4.5. Top: Comparison of velocity seismograms for the horizontal in-line (v_x), vertical (v_z) and cross-line (v_y) components for 6 different offsets. Bottom: Misfit function for each of the inversions (filled circles). The dashed red and blue lines constitute the contribution of Rayleigh and Love waves, respectively, to the joint misfit function (yellow circles). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 4.1).	53
4.7	Models obtained by the inversion of synthetic data for a spatially correlated model. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true models and the initial and final results of the Rayleigh wave (P-SV) FWI using two different initial models, namely A and B. Red asterisks represent the source locations.	55
4.8	Comparison of the observed data with the data obtained using the final models of the Rayleigh wave (P-SV FWI) for the cases A and B shown in Figure 4.7. Comparison of velocity seismograms for the horizontal in-line (v_x) and vertical (v_z) components for 6 different offsets. Misfit function for the two different cases is shown in the third column.	56
4.9	The effects of time-frequency windowing on a synthetic seismogram in the time and frequency domain. a) Synthetic reference seismic traces. The arrivals of different wave types are highlighted. b) Seismogram after application of pure low-pass filters (LP) and a combination of low-pass filter and time windowing (LP+TW). c) Corresponding amplitude spectra.	58
4.10	Taper function in the time and frequency domain for damping factor $\gamma = 10$ and 100 , respectively.	59
4.11	Data used in conventional FWI (CFWI, only multiscale approach) versus the data used in TF-FWI (time-frequency approach) for different inversion stages. The application of the trace-wise normalization allows to account for the offset-dependent energy-loss when applying a taper at time equal to 0.	60
4.12	Models obtained by the inversion of synthetic data. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional FWI (CFWI) and time-frequency FWI (TF-FWI). Red asterisks represent the source locations. The true boundaries of the trench are highlighted. The black dashed and solid lines correspond to the 1D profiles in Figure 4.13.	61
4.13	1D models of the true, initial and reconstructed parameters shown in Figure 4.12. The first row shows vertical profiles of the true, initial and reconstructed models at the location of the shallow anomaly (solid vertical lines in Figure 4.12, first column). The second row shows vertical profiles through the background medium (dashed vertical lines in Figure 4.12, first column).	62
4.14	Synthetic reconstruction test. Comparison of the observed data with the data obtained using the final models of conventional FWI (CFWI) and time-frequency FWI (TF-FWI) shown in Figure 4.12. Comparison of velocity seismograms for the horizontal in-line (v_x , first column) and vertical (v_z , second column) components for 10 different offsets. b) Misfit function for the CFWI and TF-FWI (third column).	62

4.15	Snapshots (left two columns) of the S-wave (E_s) and P-wave (E_p) intensity with the contour of the underlying velocity model of the Canadian foothills. The corresponding seismograms of vertical particle velocity are shown in the right column. The source is located at the surface in the center of the model. The propagation time is increasing from top to bottom and the exact times of the snapshots are indicated by the blue lines in the seismograms.	65
4.16	Models obtained by the inversion of synthetic data. The rows of the Figure represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional (CFWI) and time-frequency (TF-FWI) inversion. Red asterisks represent the source locations.	68
4.17	Top: 1D models of the true, initial and reconstructed parameters shown in Figure 4.16. Bottom: misfit functions for CFWI (A) and TF-FWI (B).	68
4.18	Eight grid-point stencil used to obtain the gradient of E_{CG} . Black circles correspond to the standard and grey rhombi to the rotated grid, both used to approximate the derivative at the central grid-point.	70
4.19	The rows represent v_p , v_s and density, respectively. The columns represent the true and initial models and the final results of the conventional and structurally constrained FWI. Red stars represent the source locations. The true v_p model is used throughout the inversions and it is not updated. The PML boundaries are shown with dashed lines.	72
4.20	Gradients used to update the v_s model. a) Gradient of the cross-gradient term calculated by equation 4.12 from v_p . b) Seismic gradient of v_s obtained by the adjoint-state method. c) Combined gradient obtained by equation 4.13 with $\alpha = 0.05$	73
5.1	Trace-wise normalized shotgather of the seismic field data. The corresponding source is located at the profile coordinate of -3 m. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH waves. The gathers in both cases are dominated by the surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.	76
5.2	Complete preprocessing workflow for field seismic data. Starting from the raw data on top left image and moving clockwise through the various processing steps. With black color the result of the particular step is shown, while red indicates the previous step for comparison. Note that an transparency level is enforced to show both stages at all times. Every third trace is shown for better visualization of the various preprocessing steps.	77
5.3	Frequency spectra of selective preprocessing workflow stages from Figure 5.2.	77
5.4	MASW inversion process. Right: slant-stack transformation and picking the fundamental mode of surface wave. Middle: observed, initial and inverted data. Left: inverted model from MASW versus the initial model.	80
5.5	Exemplary rheological model. The quality factor of 25 is approximated by three relaxation mechanisms of a generalized standard linear solid with the relaxation frequencies for $\sigma=4.5912$, $\tau=0.0821$, $f_1=0.3579$, $f_2=6.7362$ and $f_3= 82.0317$ Hz.	81
5.6	Overlay of the autocorrelation of the observed data and the crosscorrelation of them with the synthetic for different offsets. In the long-offset trace cycle-skipping is present.	83
5.7	Overlay of field (black) and synthetic (red) data showing the effect of the two criteria for cycle-skip evaluation. In the middle and right figure the traces classified as cycle-skipped are removed for illustration purposes.	83

6.1	Map of the survey area on the glider airfield in Rheinstetten. The blue line corresponds to the interpolated path of the "Ettliger Linie" and the orange line shows the acquisition profile. Start: N 48°58' 52.59", E 8°20' 46.73". End: N 48°58'52.12", E 8°20' 44.57". Source: Google Earth (AeroWest, GeoBasis-DE/BKG).	87
6.2	Trace-wise normalized shot-gather of the field seismic data. The corresponding source is located at the profile coordinate of 0.5 m. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH-waves. The gathers in both cases are dominated by the surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.	88
6.3	Amplitude-frequency spectra for each component of the field data set. The spectrum of the P-SV waves is shown in red and green for v_x and v_z components, respectively, and of the SH waves in blue. The shown spectra are the average of the spectra of normalized traces. The hatched area corresponds to the target frequency band which was used during FWI.	88
6.4	Models obtained by the inversion of field data. The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of CFWI and TF-FWI, respectively. Red stars represent the source locations.	90
6.5	Comparison of the observed data with the data obtained using the final models of CFWI and TF-FWI shown in Figure 6.4. Comparison of velocity seismograms for three different shots at three locations. The misfit function for the CFWI and TF-FWI approach is shown at the fourth column.	91
6.6	Estimated wavelets for the source-time function of CFWI and TF-FWI. The consistency between the inverted wavelets indicate that both inversions were able to provide reliable models, concerning the main subsurface features.	92
6.7	Data misfit of single shots and single receivers for CFWI and TF-FWI. The values are normalized by the maximum amplitude of the misfit in each case. The removal of the near-offset traces is indicated by the white area. These traces are not considered within the misfit calculation and in the inversion.	92
6.8	Models obtained by the inversion of field data. The rows of the Figure represent v_p , v_s and density models, respectively. The columns represent the initial model and the final results of P-SV FWI using only the v_x component, for the case that a strong mute is applied and the case that the full wavefield is considered for the inversion. Red stars represent the source locations.	93
6.9	Comparison of the observed data with the data obtained using the final models of the two cases shown in Figure 6.8.	94
6.10	Comparison of the misfit function for the two cases of P-SV FWI using only the v_x component.	94
6.11	Estimated wavelets for the source-time function of v_x data for the case where a strong mute is applied and when the full wavefield is considered for the inversion, from the first (top) and last (bottom) inversion stages.	95
6.12	Models obtained by the inversion of field data. The rows of the Figure represent v_s and density models, respectively. The columns represent the initial model and the final results of SH FWI. Red stars represent the source locations.	95
6.13	Results of SH FWI, corresponding to the models shown in Figure 6.12. a) Comparison of observed and synthetic velocity seismograms at the first shot location. b) Inverted source-time function at the first iteration. c) Inverted source-time functions at the last iteration.	96

6.14	Qualitative comparison of the GPR result with the field data FWI results. Overlay of the final P- and S-wave velocity models of CFWI and TF-FWI of P-SV waves (middle and bottom row, respectively) and CFWI of SH waves (top row) with the time-migrated image of the GPR measurement. A color-bar is not shown, because the transparent overlay affects the color representation.	97
7.1	a) Overview of the Krauthausen test site in northwest Germany. Source: Döring (1997). b) A conceptual model of the geologic sequence for the aquifer in Krauthausen as studied by Döring (1997).	101
7.2	Sketch illustrating the various impact sources that were used in this study. a) Vertical strike on a steel plate used to record P-SV waves from vertical and inline-polarized receivers. b) Horizontal strike on a steel beam (which is coupled to the ground) used to record SH waves from subtraction of the signal from the two opposite sides from crossline-polarized receivers. c) Prismatic steel source that allows the separation into vertical and shear-polarized forces by either summation or subtraction, respectively, of the signal from the two opposite sides. The angle the prism source to the surface is 45 degrees. The main recorded wavetypes are highlighted for illustration purposes, however, more events can be produced depending on the source-receiver polarization that is applied.	102
7.3	Reference shot-gathers induced by a steel plate (Figure 7.2a), a horizontal shear-beamer (Figure 7.2b) and the prismatic source (Figure 7.2c). a) The shot-gathers from the vertical and horizontal crossline recordings are shown (first and second column, respectively). b) Trace-by-trace crosscorrelation between the two source types of the same polarization (first two rows).	104
7.4	Amplitude-frequency spectra for each pair of sources with the same polarization. a) The spectrum of the P-SV recordings (v_z) for the vertical plate (blue line) and the vertical component of the prism (green line) obtained after summation of the two opposite impulses. b) The spectrum of the SH recordings (v_y) for the horizontal beamer (orange line) and the horizontal crossline component of the prism (green line) obtained after subtraction of the two opposite impulses (Figure 7.2c).	104
7.5	Map of the Krauthausen test site including the two seismic profiles and the boreholes adjacent to them (numbered circles). Black dashed lines show the receiver arrays and stars show the location of the first and last shot for each of the two profiles. Crosses indicate the location of cone penetrating measurements done by Tillmann et al. (2008). The purple dashed line shows the seismic profile that was used for a preliminary evaluation (Athanasopoulos and Bohlen, 2017a). The unit is m.	106
7.6	Overview of the field conditions and the acquisition layout, along with the prismatic source and its operation towards the inline component (direction of the receiver array).	106
7.7	Trace-wise normalized shot-gathers of the field seismic data. The corresponding sources are located at profile coordinates of -3 m and -1.5 m, for profiles 1 and 2, respectively. The v_x and v_z components show the P-SV wave types, while the v_y component shows the SH-waves. The gathers in both cases are dominated by surface waves, Rayleigh and Love waves for the P-SV and SH recordings, respectively.	108

7.8	Amplitude-frequency spectra for each component of the two profiles, cumulative for every source and receiver. The spectrum of the P-SV waves is shown in red and green, for v_x and v_z components respectively and of the SH waves (v_y) in blue. The shown spectra are the average of the spectra of the normalized traces. The hatched area corresponds to the target frequency band which was used for FWI.	108
7.9	Models obtained by joint FWI of field data acquired at profile 1 (Figure 7.5). The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of joint FWI. Red stars represent the source locations.	110
7.10	Models obtained by joint FWI of field data acquired at profile 2 (Figure 7.5). The rows of the Figure represent the v_p , the v_s and the density model, respectively. The columns represent the initial model and the final result of joint FWI. Red stars represent the source locations.	110
7.11	Inversion results of joint FWI for profile 1, corresponding to the models shown in Figure 7.10. a) Comparison of observed and synthetic velocity seismograms at the 1 st and 6 th shot location. b) Misfit as a function of the number of iteration. The dashed blue and red lines constitute the contribution of the Love and Rayleigh waves, respectively, to the joint misfit function (dashed yellow). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 7.1).	113
7.12	Inversion results of joint FWI for profile 2, corresponding to the models shown in Figure 7.10. a) Comparison of observed and synthetic velocity seismograms at the 1 st , 6 th and 11 th shot location. b) Misfit as a function of the number of iteration. The dashed blue and red lines constitute the contribution of the Love and Rayleigh waves, respectively, to the joint misfit function (dashed yellow). Jumps of increasing misfit value correspond to changes in the workflow stage (Table 7.1).	114
7.13	Frequency spectra of the observed and synthetic v_z and v_y seismic data at the first shot location (Figure 7.11.a.)	114
7.14	Frequency spectra of the observed and synthetic v_x , v_z and v_y seismic data at the first shot location (Figure 7.12.a).	115
7.15	Estimated wavelets for the source-time function of P-SV (left) and SH (right) data acquired along profile 1, from the first (top) and last (bottom) inversion stages. The source wavelets are overall coherent for the various shot locations, for both first and final stages. The inverted wavelets from the crossline source are in general less coherent compared to the vertical recordings.	115
7.16	Estimated wavelets for the source-time function of P-SV (left) and SH (right) data acquired along profile 2, from the first (top) and last (bottom) inversion stages. The source wavelets are overall coherent for the various shot locations, for both first and final stages. The inverted wavelets from the crossline source are in general slightly less coherent compared to the vertical recordings.	115
7.17	(a) Density obtained from FWI. b) Comparison of vertical bulk density profiles derived from CPT with co-located 1D density models from FWI. The exact locations of the CPT profiles within the FWI model are indicated by the dashed lines.	116
7.18	3D plot of the retrieved v_p , v_s and density models from joint FWI of Rayleigh and Love waves of the two profiles. Black (opal) cylinders represent the boreholes in the area that were used to perform crosshole GPR FWI by Gueting et al. (2015, 2017).	118

A.1	Scattered wavefields caused by perturbations of the viscoelastic parameters for various depths of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.	125
A.2	Scattered wavefields caused by perturbations of the viscoelastic parameters for various sizes of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for the first column and then are halved for the second and doubled for the third column, for each parameter accordingly.	126
A.3	Scattered wavefields caused by perturbations of the viscoelastic parameters for various shapes of the inclusion. The values are given in table 3.2. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.	126
A.4	Scattered wavefields caused by perturbations of the viscoelastic parameters for different values of the inclusion. The values for the first column are given in table 3.2, while for the second column the sign of the percentage added is shifted. Seismograms of vertical particle velocity of the scattered wavefield are shown, as well as the models used. Triangles mark the positions of some receivers and the stars indicate the shot location. The seismograms are scaled with the same values as Figure 3.5 for each parameter accordingly.	127
C.1	Mechanical design of the steel prism. The dimensions are shown in mm.	131
D.1	Dispersion curves extracted from the data of profile 1. Phase-velocity over frequency calculated from the 1 st shot of the observed data and inverted data (top and bottom rows, respectively) for the vertical (v_z) and horizontal crossline (v_y) components (left and right columns, respectively). The dispersion image of v_z reveals the fundamental mode of the Rayleigh wave, while the v_y component the fundamental mode of the Love wave, as well as the existence of a potential higher mode.	133
D.2	Dispersion curves extracted from the data of profile 2. Phase-velocity over frequency calculated from the 1 st shot of the observed data and inverted data (top and bottom rows, respectively) for the horizontal inline (v_x), vertical (v_z) and horizontal crossline (v_y) components (left and right columns, respectively). The dispersion images of v_x and v_z reveal the fundamental mode of the Rayleigh wave, while the v_y component the fundamental mode of the Love wave, as well as the existence of a potential higher mode.	133

List of Tables

3.1	Examined cases and their corresponding figure number.	36
3.2	Material properties of the models shown in figure 3.1. The percentage of reduction from the background model is shown as well.	36
4.1	Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter.	47
4.2	Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter. The column ' γ ' indicates the damping factor (equation 4.3). The first stage is only applied in TF-FWI.	61
4.3	Workflow used for the FWI of synthetic data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass Butterworth filter. The column ' γ ' indicates the damping factor (equation 4.3).	67
6.1	Workflow used for FWI of field data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass frequency filter. The parameter ' γ ' indicates the damping factor that is used during the corresponding stage (column 'Stage' ') in equation 4.3. The first two stages were only applied in TF-FWI for the v_p model.	90
7.1	Workflow used for FWI of field data. The column 'Update' indicates which of the specific elastic parameters is updated (yes=1, no=0). The parameter LP represents the upper corner frequency of the low-pass frequency filter. The parameter ' γ ' indicates the damping factor that is used during the corresponding stage (column 'Stage' ') in equation 4.3.	109
B.1	All configurations for passing structural information from one model parameter to another.	130