

Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves

Individuelle und simultane elastische
2-D Wellenforminversion von Rayleigh und Love Wellen

Master's Thesis of

Florian Wittkamp

at the Department of Physics
Geophysical Institute (GPI)

Reviewer: Prof. Dr. Thomas Bohlen
Second reviewer: Prof. Dr. Friedemann Wenzel

1. September 2015 – 31. August 2016

Karlsruher Institut für Technologie
Fakultät für Physik
Postfach 6980
76128 Karlsruhe

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

Karlsruhe, August 12, 2016

.....

(Florian Wittkamp)

Abstract

We investigate the performance of the individual 2-D elastic full-waveform inversion (FWI) of Rayleigh and Love waves as well as the feasibility of a simultaneous joint FWI of both wave types. The FWI of surface waves can provide a valuable contribution to near-surface investigations, since they are mainly sensitive to the S-wave velocity and hold a high signal-to-noise ratio. In synthetic reconstruction tests we compare the performance of the individual wave type inversions and explore the benefits of a simultaneous joint inversion. In these tests both individual wave type inversions perform similarly well, given that the initial P-wave velocity model is accurate enough. In this case the joint FWI further improves the result. For an inaccurate initial P-wave velocity model, we observe artifacts in the results of the Rayleigh wave FWI and the joint FWI. Subsequently, we recorded a near-surface field dataset to verify the results by a realistic example. In the field data application the Love wave FWI is superior to the Rayleigh wave FWI, possibly due to the initial P-wave velocity model. Also in this case the joint FWI further improves the inversion result.

Zusammenfassung

In dieser Arbeit vergleichen wir die Eigenschaften der individuellen Wellenforminversion von Rayleigh und Love Wellen. Des Weiteren untersuchen wir die Realisierbarkeit einer simultanen Wellenforminversion beider Oberflächenwellen. Die Wellenforminversion von Oberflächenwellen leistet einen wichtigen Beitrag zur Erkundung des oberflächennahen Untergrundes. Insbesondere deshalb, da Oberflächenwellen äußerst sensitiv zur S-Wellengeschwindigkeit sind und ein hohes Signal-Rausch-Verhältnis aufweisen. In synthetischen Experimenten vergleichen wir die Fähigkeiten der beiden individuellen Wellenforminversionen den oberflächennahen Untergrund zu rekonstruieren und untersuchen, ob eine simultane Wellenforminversion beider Oberflächenwellen Vorteile bietet. In diesen Experimenten erzielen beide individuellen Wellenforminversionen vergleichbare Ergebnisse, vorausgesetzt, das Startmodell für die P-Wellengeschwindigkeit ist ausreichend genau bekannt. Ist dies der Fall, so führt eine simultane Wellenforminversion beider Oberflächenwellen zu einer Verbesserung des Inversionsergebnisses. Ist jedoch nur ein unzureichend genaues Modell für die P-Wellengeschwindigkeit bekannt, lassen sich Artefakte im Resultat der individuellen Rayleigh Wellenforminversion, als auch im Resultat der simultanen Wellenforminversion beider Wellentypen beobachten. Um diese Ergebnisse mit einem realitätsnahen Beispiel zu verifizieren, führten wir Feldmessungen durch. Die Anwendung der individuellen Wellenforminversion auf die aufgezeichneten Love Wellen führte zu einem besseren Resultat, als die individuelle Wellenforminversion der aufgezeichneten Rayleigh Wellen. Dieses Resultat hängt womöglich mit dem initialen Modell für die P-Wellengeschwindigkeit zusammen. Verglichen mit den individuellen Wellenforminversionen verbessert auch hier die gleichzeitige Inversion beider Oberflächenwellen das Ergebnis der Inversion.

Contents

Abstract	i
Zusammenfassung	iii
1. Introduction	1
1.1. Main objectives	2
1.2. Thesis structure	2
2. Methodology	5
2.1. Wave Equation	5
2.1.1. Elastic medium	6
2.1.2. Viscoelastic medium	8
2.2. Forward modelling	10
2.2.1. Discretization	10
2.2.2. Initial and boundary conditions	12
2.3. Full-Waveform Inversion	13
2.3.1. Inverse Problem	13
2.3.2. Gradient calculation: Adjoint State Method	15
2.3.3. Hessian calculation: L-BFGS	23
2.3.4. Multi-stage approach	26
2.4. Simultaneous Joint Inversion	29
2.4.1. Joint objective function	29
2.4.2. Joint gradients	30
2.5. Summary	31
3. Synthetic FWI experiments	33
3.1. Setting	33
3.1.1. True and initial models	33
3.1.2. Acquisition geometry	34
3.1.3. Pseudo-observed data	37
3.1.4. FWI setup	37
3.2. Results	40
3.2.1. Individual Love wave FWI	40
3.2.2. Individual Rayleigh wave FWI	40
3.2.3. Simultaneous joint FWI	41
3.2.4. Comparison	46
3.3. Case study: Influence of the initial P-wave velocity	49
3.4. Summary	52

4. Field data application	53
4.1. Setting	53
4.1.1. Test site	53
4.1.2. Acquisition geometry	55
4.1.3. Observed data	56
4.1.4. Initial model	59
4.1.5. FWI setup	60
4.2. Results	65
4.2.1. Individual Love wave FWI	65
4.2.2. Individual Rayleigh wave FWI	66
4.2.3. Simultaneous joint FWI	67
4.2.4. Comparison	74
4.3. Comparison with ground-penetrating radar	78
4.4. Summary	81
5. Conclusions and Outlook	83
5.1. Conclusions	83
5.2. Outlook	85
Bibliography	87
List of Figures	93
List of Tables	95
A. FD-modelling and inversion parameters	97

1. Introduction

The analysis of shallow-seismic surface waves provides a valuable contribution to near-surface investigations. Their propagation is mainly influenced by the shear-wave velocity, which is an important geotechnical parameter. The acquisition of shallow-seismic surface waves is simple and cost-efficient, since they can be easily excited, for instance, by sledgehammer blows on the surface and recorded by conventional single component geophones. Furthermore, they show a high signal-to-noise ratio in shallow-seismic field data recordings, which makes them even more attractive to a broad spectrum of near-surface studies.

The conventional methods for the analysis of shallow-seismic surface waves are the inversion of dispersion curves (e.g. [McMechan and Yedlin, 1981](#); [Park et al., 1999](#); [Xia et al., 1999](#)) or Fourier-Bessel expansion coefficients ([Forbriger, 2003a,b](#)). These approaches are based on a transformation of the recorded wave fields into the velocity/slowness–frequency/wavenumber domain, where 1-D inversion methods are applied to obtain synthetic 1-D subsurface models. However, these approaches are limited to lateral homogeneous or smooth heterogeneous subsurfaces, where in the latter case averaged material properties are obtained. [Socco et al. \(2010\)](#) present an overview of several techniques to overcome the limitation to 1-D subsurface models, like the analysis of data subsets along the profile, where local 1-D depth-dependent models are calculated and subsequently combined to a 2-D subsurface model ([Bohlen et al., 2004](#)). Nevertheless, all of those methods have a limited lateral resolution and are not applicable in highly heterogeneous media.

The full-waveform inversion (FWI) of seismic recordings, as proposed by [Lailly \(1983\)](#), [Tarantola \(1984\)](#) and [Mora \(1987\)](#), can reveal 2-D as well as 3-D subsurface structures without limitations regarding the subsurface heterogeneity. The FWI makes use of the whole information content included in seismic waves, such as the amplitude and the phase, which allows to achieve a resolution below the size of a wavelength. The main drawback of the FWI compared to the conventional methods is the requirement of large computational facilities, which are required for the numerical simulation of wave propagation. While this requirement prevented its application in the past decades, today's high-performance computing (HPC) systems provide enough computation power to make the FWI feasible. In recent time, the FWI has been successfully applied to a wide range of scales, such as in seismology (e.g. [Bleibinhaus et al., 2007](#); [Fichtner et al., 2009](#)), in seismic exploration (e.g. [Operto et al., 2004](#); [Brossier et al., 2009](#)) or in near-surface investigations (e.g. [Gélis et al., 2007](#); [Romdhane et al., 2011](#)). However, the application of the FWI to field data is still challenging, in particular the application to shallow-seismic surface waves is ambitious, since their propagation is highly nonlinear in complex earth media. So far, there are only a few publications which present successful 2-D FWI field data applications using shallow-seismic surface waves to reveal shear-wave velocity models. In most of the publications Rayleigh waves were used (e.g. [Tran et al., 2013](#); [Groos, 2013](#); [Schaefer, 2014](#)), whereas Love waves were only rarely used ([Dokter et al., 2014](#); [Pan et al., 2016](#)). Nevertheless,

Xia et al. (2012), who investigated the dispersion curve inversion of Rayleigh and Love waves, observed three main advantages of the Love wave compared to the Rayleigh wave inversion: (1) The inversion of Love wave data is more stable, since they are independent of the P-wave velocity, (2) Love wave dispersion curves are simpler than those of Rayleigh waves, for the same reason as (1), and (3) the dispersion curves of Love waves show a higher signal-to-noise ratio than those of Rayleigh waves. Safani et al. (2005) made similar observations and concluded that Love waves exhibit a higher sensitivity as well as inversion stability and show an improved signal-to-noise ratio in dispersion spectra compared to Rayleigh waves, respectively. However, to the best of our knowledge, there are no studies which compare the performance of the Love wave FWI and the Rayleigh wave FWI or which investigate a coupling of both inversions in order to perform a joint FWI.

1.1. Main objectives

The main objectives of this thesis are to compare the performance of the individual full-waveform inversion of Rayleigh and Love waves and to explore the benefits of a simultaneous joint FWI of both types of surface waves. For this purpose, we extend our existing 2-D P-SV FWI code by a SH FWI scheme. To perform a joint FWI of both independent FWI schemes we propose a joint approach, which we implement into our FWI code as well. In synthetic reconstruction experiments we investigate the performance of both individual wave type inversions as well as of the joint inversion. We then use the gained experience in these synthetic experiments in a field data application, where we verify the synthetic results by a realistic example.

1.2. Thesis structure

This thesis is divided into four main chapters:

In the **second chapter**, we describe the theory used in this thesis. We introduce the wave equations for an elastic and viscoelastic medium and present the finite-difference method, which we use to solve the wave equations numerically. We then discuss the inverse problem of the full-waveform inversion and derive the gradients for the P-SV and the SH waves. Furthermore, we describe the multi-parameter quasi-Newton L-BFGS method, which we use for the calculation of the model corrections. Finally, we propose our joint approach that is required to perform a simultaneous joint inversion of the P-SV and the SH waves.

To explore the performance of the individual and the joint 2-D elastic full-waveform inversion of Rayleigh and Love waves, we perform synthetic reconstruction tests which we present in the **third chapter**. For these reconstruction tests we assume a true subsurface model which we then use to generate pseudo-observed seismograms. The knowledge of the true model allows us to study the reconstruction ability of the FWI by comparing the true model to the reconstructed models. In preparation of the field data application, we choose a subsurface model which simulates the subsurface expected at the test site.

In the **fourth chapter**, we present the application of the individual and the joint 2-D

elastic full-waveform inversion of Rayleigh and Love waves to a near-surface field dataset. We show the recorded field dataset along with various preprocessing steps that we applied. The configuration of the inversion is introduced as well. After the discussion of the inversion results, we compare them to a ground-penetrating radar measurement to study the reliability and quality of the FWI results.

In the **fifth chapter**, we draw final conclusions based on our results.

2. Methodology

In this chapter, we describe the theory used in this thesis. First, we introduce the wave equations for an elastic as well as a viscoelastic medium. We then present a numerical solution to the wave equations by a time-domain finite-difference (FD) method. This method allows to obtain an explicit time-stepping FD-scheme which we use for the efficient simulation of wave propagation through the subsurface. Furthermore, we discuss the general inverse problem of the full-waveform inversion (FWI). We use the adjoint state method to calculate the gradient of the objective function computationally efficiently. We then describe the multi-parameter quasi-Newton L-BFGS method which is used to calculate the actual model correction. Afterwards, we introduce our approach to carry out a simultaneous joint inversion of the P-SV and the SH waves and finally give a summary of the full-waveform inversion process as a whole. In the following, we use the Einstein notation.

2.1. Wave Equation

The wave equation is the key to describe the propagation of seismic waves through an arbitrary medium. In the following, we will shortly derive the wave equation, as shown by [Lay and Wallace \(1995\)](#).

To start, we consider the equation of motion. The equation of motion is the most fundamental equation in seismology as it provides a relation of forces in a medium to measurable displacements ([Lay and Wallace, 1995](#)). The equation of motion reads:

$$\rho \cdot \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i, \quad (2.1)$$

where ρ denotes the density, u the displacement, t the time, σ the stress tensor, f the external body forces and x is the vector in space. The lower indices correspond to the direction in space $\{i, j\} = \{x, y, z\}$ (see figure 2.1).

To derive the 2-D equation of motion from equation 2.1, we assume a subsurface in which all properties vary only in the $x-z$ plane (see figure 2.1). In such a subsurface, all properties in the y direction are constant and derivations in the y direction vanish. By considering equation 2.1 and writing out all the summations explicitly the equation of motion splits up into two separate partial differential equation systems. One system describes the P-SV waves:

$$\rho \cdot \frac{\partial^2 u_x}{\partial t^2} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} + f_x, \quad (2.2a)$$

$$\rho \cdot \frac{\partial^2 u_z}{\partial t^2} = \frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{xz}}{\partial x} + f_z, \quad (2.2b)$$

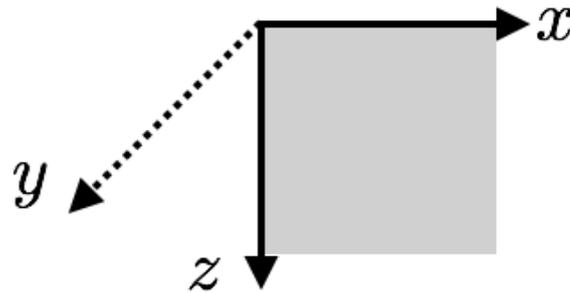


Figure 2.1.: Used cartesian coordinate system. The $x - z$ plane (grey area) represents the used 2-D plane. The z direction refers to depth. In the y direction all properties are assumed to be constant.

and the other system describes the SH waves:

$$\rho \cdot \frac{\partial^2 u_y}{\partial t^2} = \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial z} + f_y. \quad (2.3)$$

Both systems are independent of each other, which means the P-SV and the SH waves are decoupled in a 2-D medium, hence, there is no conversion from one wave type into the other. Both waves types propagate in the $x - z$ plane, but they differ in their polarization. The polarization direction of the P-SV waves lies within the $x - z$ plane, whereas the SH waves are polarized vertical to the $x - z$ plane, namely in the y direction.

Since both partial differential equation systems hold more unknowns than equations, there is no unique solution to them. Therefore, further information of how the medium will react to applied forces are necessary. There are two kinds of such reactions: (1) A purely elastic response and (2) a viscoelastic response.

2.1.1. Elastic medium

A purely elastic medium reacts spontaneous and linear to an applied force. The relation between stress, σ , and strain, ϵ , of such a medium follows the linear Hook's law (Lay and Wallace, 1995):

$$\sigma_{ij} = C_{ijkl} \cdot \epsilon_{kl}, \quad (2.4)$$

where the elastic moduli, C_{ijkl} , describes the properties of the medium. The strain tensor, ϵ_{kl} , yields:

$$\epsilon_{kl} = \frac{1}{2} \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right). \quad (2.5)$$

Considering an isotropic medium the elastic moduli, C_{ijkl} , simplifies to:

$$C_{ijkl} = \lambda \cdot \delta_{ij} \delta_{kl} + \mu \cdot (\delta_{il} \delta_{jk} + \delta_{ik} \delta_{jl}), \quad (2.6)$$

where the Kronecker delta, δ_{ij} ,

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \quad (2.7)$$

and the Lamé parameters μ and λ are used. The two Lamé parameters are sufficient to describe the material properties of an isotropic and elastic medium. For such a medium the stress-strain relationship simplifies to:

$$\sigma_{ij} = \lambda \cdot \epsilon_{kk} \cdot \delta_{ij} + 2 \cdot \mu \cdot \epsilon_{ij}. \quad (2.8)$$

The relation between the Lamé parameters and the seismic P-wave velocity, v_p , and S-wave velocity, v_s , is (Lay and Wallace, 1995):

$$v_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \quad v_s = \sqrt{\frac{\mu}{\rho}}. \quad (2.9)$$

The equation of motion and the stress-strain relation contain all required information to describe wave propagation in an isotropic elastic medium, which allows us to derive the wave equation. In stress-velocity formulation, where the particle velocity, $v = \frac{\partial u}{\partial t}$, is the wave field parameter, the wave equation is represented by a first-order hyperbolic system. This first-order system allows for an efficient numerical solution by a time-domain finite-difference method.

For simplicity, in the following we omit the temporal and spatial dependencies. For P-SV waves the elastic stress-velocity wave equation is (Virieux, 1986):

$$\frac{\partial v_x}{\partial t} = \frac{1}{\rho} \cdot \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right), \quad (2.10a)$$

$$\frac{\partial v_z}{\partial t} = \frac{1}{\rho} \cdot \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right), \quad (2.10b)$$

$$\frac{\partial \sigma_{xx}}{\partial t} = M \cdot \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) - 2 \cdot \mu \cdot \frac{\partial v_z}{\partial z}, \quad (2.10c)$$

$$\frac{\partial \sigma_{zz}}{\partial t} = M \cdot \left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x} \right) - 2 \cdot \mu \cdot \frac{\partial v_x}{\partial x}, \quad (2.10d)$$

$$\frac{\partial \sigma_{xz}}{\partial t} = \mu \cdot \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right), \quad (2.10e)$$

where $M = (\lambda + 2\mu)$ is the P-wave modulus. The elastic stress-velocity wave equation for the SH waves reads (Virieux, 1984):

$$\frac{\partial v_y}{\partial t} = \frac{1}{\rho} \cdot \left(\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial z} \right), \quad (2.11a)$$

$$\frac{\partial \sigma_{xy}}{\partial t} = \mu \cdot \frac{\partial v_y}{\partial x}, \quad (2.11b)$$

$$\frac{\partial \sigma_{zy}}{\partial t} = \mu \cdot \frac{\partial v_y}{\partial z}. \quad (2.11c)$$

2.1.2. Viscoelastic medium

In contrast to a purely elastic medium which does not attenuate the wave energy, wave propagation in real earth media is always dispersive and attenuated by the transformation of motion energy into heat. A viscoelastic rheological model can account for these effects and consequently allows for an accurate description of wave propagation.

The stress-strain relation for an isotropic viscoelastic medium reads:

$$\sigma_{ij} = \left(\dot{\Psi}_P - 2 \cdot \dot{\Psi}_S \right) * \epsilon_{kk} \cdot \delta_{ij} + 2 \cdot \dot{\Psi}_S * \epsilon_{ij}, \quad (2.12)$$

where Ψ_P and Ψ_S describe the relaxation functions for P- and S-waves, respectively (Bohlen, 1998; Christensen, 2012). The asterisk denotes convolution in time.

A rheological model that can approximate the attenuation and dispersion properties of a medium is the generalized standard linear solid (GSLS) (Emmerich and Korn, 1987; Blanch et al., 1995). A GSLS consists of several Maxwell bodies (dashpot in series with a spring) in parallel with a spring. Thereby a Maxwell body refers to a relaxation mechanism and by extent represents the viscosity of the medium. With a combination of several relaxation mechanisms a specific viscoelastic model with a certain attenuation versus frequency relation can be obtained (Robertsson et al., 1994; Liu et al., 1976).

The relaxation functions for a GSLS with L relaxation mechanism are:

$$\Psi_P(t) = M \cdot \left[1 + \sum_{l=1}^L \left(\frac{\tau_{\epsilon l}^P}{\tau_{\sigma l}} - 1 \right) e^{-t/\tau_{\sigma l}} \right] \cdot H(t), \quad (2.13)$$

$$\Psi_S(t) = \mu \cdot \left[1 + \sum_{l=1}^L \left(\frac{\tau_{\epsilon l}^S}{\tau_{\sigma l}} - 1 \right) e^{-t/\tau_{\sigma l}} \right] \cdot H(t), \quad (2.14)$$

where M is the relaxed P-wave modulus, μ is the relaxed S-wave modulus and $H(t)$ is the Heaviside step function (Bohlen, 1998; Liu et al., 1976; Carcione et al., 1988). The parameters $\tau_{\epsilon l}^P$ and $\tau_{\epsilon l}^S$ are the P- and S-wave strain retardation times for the l -th relaxation mechanism, respectively, and $\tau_{\sigma l}$ are the corresponding stress relaxation time.

Blanch et al. (1995) suggested the use of the τ -method to reduce the number of variables that describe the GSLS and thus, reduce the calculation time and the memory requirements. This method relies on the observation that one dimensionless variable, τ , can describe the level of damping caused by a GSLS. Moreover, with this method it is possible to approximate a frequency independent seismic quality factor, Q , in a limited frequency range. The parameter τ , which is independent of L , is defined as:

$$\tau = \frac{\tau_{\epsilon l}}{\tau_{\sigma l}} - 1 \quad (2.15)$$

and replaces the L -dependent retardation times.

The attenuation versus frequency relation of a GSLS which consists of L relaxation mechanisms can be described by the quality factor, Q :

$$Q(\omega, \tau_{\sigma l}, \tau) = \frac{1 + \sum_{l=1}^L \frac{\omega^2 \tau_{\sigma l}^2}{1 + \omega^2 \tau_{\sigma l}^2} \tau}{\sum_{l=1}^L \frac{\omega \tau_{\sigma l}}{1 + \omega^2 \tau_{\sigma l}^2} \tau}, \quad (2.16)$$

where ω denotes the angular frequency. In order to obtain a constant quality factor, Q , in a limited frequency range, the τ value and the L relaxation times, $\tau_{\sigma l}$, can be determined by a least-square inversion of equation 2.16 (Bohlen, 2002). While it is sufficient to use the same relaxation times, $\tau_{\sigma l}$, for P- and S-waves, the τ -values can be calculated separately to use different values of Q for both wave types (Bohlen, 2002). Instead of relaxation times it is quite often to use the relaxation frequencies, $f_l = \frac{1}{2\pi \cdot \tau_{\sigma l}}$.

The stress-strain relationship in equation 2.12 together with the equation of motion 2.1 contain the complete information of a viscoelastic medium and allows us to describe the wave propagation. However, the convolution in equation 2.12 is unsuitable for the discretization by a time-domain finite-difference method. To overcome this issue, Robertsson et al. (1994) and Carcione et al. (1988) proposed the usage of memory variables, r , in order to avoid the convolution.

The obtained first-order wave equation in stress-velocity formulation for the P-SV waves reads (Bohlen, 1998):

$$\frac{\partial v_x}{\partial t} = \frac{1}{\rho} \left(\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} \right), \quad (2.17a)$$

$$\frac{\partial v_z}{\partial t} = \frac{1}{\rho} \left(\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} \right), \quad (2.17b)$$

$$\frac{\partial \sigma_{xx}}{\partial t} = M \cdot (1 + \tau^p) \cdot \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) - 2 \cdot \mu \cdot (1 + \tau^s) \frac{\partial v_z}{\partial z} + \sum_{l=1}^L r_{xxl}, \quad (2.17c)$$

$$\frac{\partial \sigma_{zz}}{\partial t} = M \cdot (1 + \tau^p) \cdot \left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x} \right) - 2 \cdot \mu \cdot (1 + \tau^s) \frac{\partial v_x}{\partial x} + \sum_{l=1}^L r_{zzl}, \quad (2.17d)$$

$$\frac{\partial \sigma_{xz}}{\partial t} = \mu \cdot (1 + \tau^s) \cdot \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) + \sum_{l=1}^L r_{xzl}, \quad (2.17e)$$

$$\frac{\partial r_{xxl}}{\partial t} = -\frac{1}{\tau_{\sigma l}} \left((M \cdot \tau^p - 2\mu \cdot \tau^s) \cdot \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) + 2\mu\tau^s \frac{\partial v_x}{\partial x} + r_{xxl} \right), \quad (2.17f)$$

$$\frac{\partial r_{zzl}}{\partial t} = -\frac{1}{\tau_{\sigma l}} \left((M \cdot \tau^p - 2\mu \cdot \tau^s) \cdot \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) + 2\mu\tau^s \frac{\partial v_z}{\partial z} + r_{zzl} \right), \quad (2.17g)$$

$$\frac{\partial r_{xzl}}{\partial t} = -\frac{1}{\tau_{\sigma l}} \left(\mu \cdot \tau^s \cdot \left(\frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) + r_{xzl} \right). \quad (2.17h)$$

Thereby τ^p denotes the τ value for the P-waves and τ^s corresponds to the S-waves. The variables r_{ijl} describe the memory variable for the l -th relaxation mechanism.

The viscoelastic wave equation for the SH waves reads (Bohlen, 1998):

$$\frac{\partial v_y}{\partial t} = \frac{1}{\rho} \left(\frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{zy}}{\partial z} \right), \quad (2.18a)$$

$$\frac{\partial \sigma_{xy}}{\partial t} = \mu \cdot (1 + \tau^s) \cdot \frac{\partial v_y}{\partial x} + \sum_{l=1}^L r_{xy_l}, \quad (2.18b)$$

$$\frac{\partial \sigma_{zy}}{\partial t} = \mu \cdot (1 + \tau^s) \cdot \frac{\partial v_y}{\partial z} + \sum_{l=1}^L r_{zy_l}, \quad (2.18c)$$

$$\frac{\partial r_{xy_l}}{\partial t} = -\frac{1}{\tau_{\sigma l}} \left(\mu \cdot \tau^s \cdot \frac{\partial v_y}{\partial x} + r_{xy_l} \right), \quad (2.18d)$$

$$\frac{\partial r_{zy_l}}{\partial t} = -\frac{1}{\tau_{\sigma l}} \left(\mu \cdot \tau^s \cdot \frac{\partial v_y}{\partial z} + r_{zy_l} \right). \quad (2.18e)$$

The comparison of equation 2.18 to equation 2.17 shows that the wave equation for the SH waves is much simpler than for the P-SV waves. A more complex wave equation will lead to a higher complexity of the forward solver, which will result in an increased calculation time and extended memory requirements. This reveals a crucial advantage of the SH wave equation, as it can be solved faster and with less memory requirements than the P-SV wave equation. The complexity of the SH wave equation is comparable to the acoustic approximation (Kurzman, 2012), however, the SH wave equation allows the description of (visco-)elastic SH-wave propagation.

2.2. Forward modelling

The inversion of full seismic waveforms requires an accurate and computational efficient forward solver for the numerical simulation of seismic waves through complex earth models. A well-proven class of forward solvers for the elastic and viscoelastic wave equations are finite-difference (FD) methods. The FD-methods approximate the spatial and temporal derivatives in the wave equations by finite-difference operators. To calculate these discrete operators, the wave equation is discretized in both space and time domain.

2.2.1. Discretization

The simplest way to discretize in the space domain is by discrete grid points on a rectangular grid with equidistant spacing, Δh . For such a grid the relation between the continuous coordinates x and z and their discrete equivalents is:

$$x = i \cdot \Delta h \quad \text{and} \quad z = j \cdot \Delta h, \quad (2.19)$$

where a specific point (i, j) refers to a grid point. The discrete transformation of the time vector reads:

$$t = n \cdot \Delta t, \quad (2.20)$$

where n refers to a specific time step.

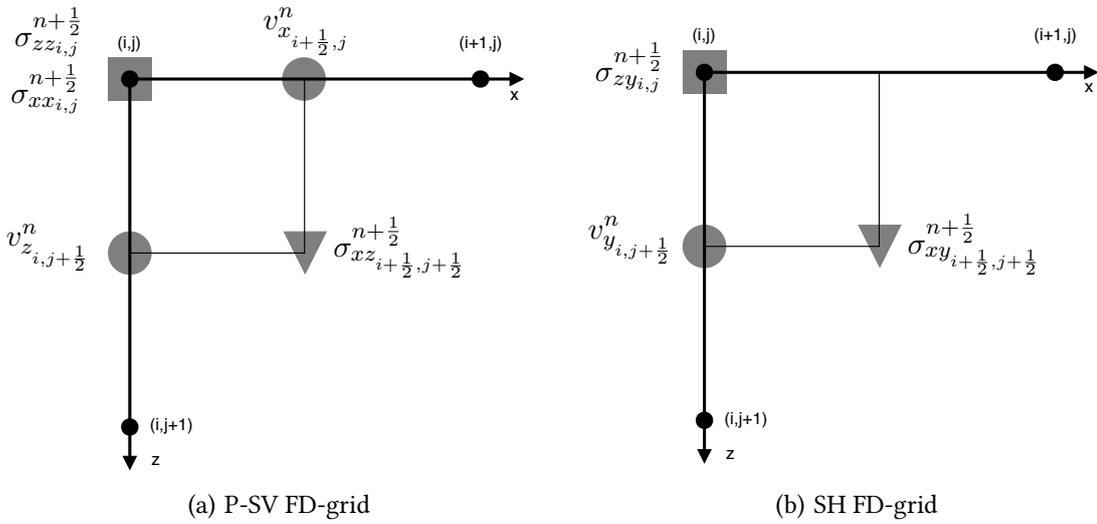


Figure 2.2.: Staggered 2-D finite-difference grid for the P-SV (left) and the SH (right) elastic wave equation. Black dots correspond to full grid points.

In this work, we follow the staggered-grid approach as proposed by [Virieux \(1984\)](#) and [Levander \(1988\)](#). This approach has several advantages compared to a conventional grid such as higher accuracy and a relaxed stability limit ([Liu et al., 1976](#)). On the staggered-grid the model parameters (μ , λ and ρ) and the diagonal stress components, σ_{kk} , are discretized on full grid points (i, j) , whereas off-diagonal stress components and particle velocities are discretized on half grid points. Moreover, the stress components are defined at half time steps $n + \frac{1}{2}$, whereas particle velocities are defined at full time steps n . Figure 2.2 shows the staggered-grid for the elastic P-SV and SH waves.

As already indicated, the basic concept of the FD-method is the replacement of the temporal and spatial derivatives in the wave equations by finite-difference operators. To illustrate such a staggered-grid approximation to the spatial and temporal derivation a second-order accurate operator is shown for both domains:

$$\left. \frac{\partial}{\partial x} f(x) \right|_i = \frac{f(i + \frac{1}{2}\Delta x) - f(i - \frac{1}{2}\Delta x)}{\Delta x} + \mathcal{O}(\Delta x^2), \quad (2.21)$$

$$\left. \frac{\partial}{\partial t} f(t) \right|_n = \frac{f(n + \frac{1}{2}\Delta t) - f(n - \frac{1}{2}\Delta t)}{\Delta t} + \mathcal{O}(\Delta t^2). \quad (2.22)$$

The FD-operators used in this work are second-order accurate in time and higher-order accurate in space. Higher-order spatial approximations are achieved by a Taylor expansion of equation 2.21. Higher-order temporal accuracy would be possible as well (eg. [Bohlen and Wittkamp \(2016\)](#)), but is not considered in this thesis. Instead, the accuracy of the temporal discretization is ensured by an appropriate temporal sampling.

The explicit time-stepping FD-scheme can be obtained by applying the staggered-grid approximation to the derivatives in the wave equation and considering the orientation of the staggered-grid as shown in figure 2.2. The work of [Bohlen \(1998\)](#) shows this procedure

exemplarily, however, in his case the orientation of the staggered-grid is shifted by half a grid point compared to the used grid in this work.

Due to the numerical discretization, the wave propagation suffers from numerical dispersion. Hence, the spatial sampling have to be selected small enough to held the numerical dispersion insignificant. Additionally, to ensure numerical stability of the FD-method a stability limit, the so-called Courant-Friedrich-Levy (CFL) condition (Courant et al., 1928), has to be satisfied. The used forward solver which is based on the software *SOFI2D* developed by Bohlen (2002) checks automatically for both stability and dispersion. A detailed description of numerical dispersion as well as numerical stability is given by Virieux (1986, 1984) and Levander (1988).

2.2.2. Initial and boundary conditions

Initial conditions

The initial conditions for all spatial locations, x , for the particle velocity are:

$$v_i(x, t = 0) = 0, \quad \dot{v}_i(x, t = 0) = 0, \quad \ddot{v}_i(x, t = 0) = 0, \quad (2.23)$$

and for the stress:

$$\sigma_{ij}(x, t = 0) = 0, \quad \dot{\sigma}_{ij}(x, t = 0) = 0, \quad \ddot{\sigma}_{ij}(x, t = 0) = 0. \quad (2.24)$$

For a viscoelastic simulation the initial conditions for the memory variables, r_{ij} , are:

$$r_{ij}(x, t = 0) = 0, \quad \dot{r}_{ij}(x, t = 0) = 0, \quad \ddot{r}_{ij}(x, t = 0) = 0. \quad (2.25)$$

Free surface boundary condition

To accurately simulate wave propagation at the earth's surface, a free surface condition has to be considered. At the interface from earth to air all stresses have to vanish. To fulfill this condition, we applied the mirroring technique from Levander (1988) on the top of the model. This approach is limited to a planar surface, however, this is sufficient for our research.

Perfectly matched layer boundary condition

To allow an efficient forward modelling, we truncated the used models as far as possible. Nevertheless, in small and finite models seismic waves will reach the model boundaries and will cause reflections at the locations of these artificial boundaries. Thus, it is important to use a boundary layer, which will suppress such artificial reflections. An effective method is the convolutional perfectly matched layers method (C-PMLs). The C-PMLs are a coordinate transformation in the complex numbers that will stretch the coordinates at the model boundary in such a way that the waves will never reach the actual boundary. The used C-PML implementation is based on Komatitsch and Martin (2007) and is applied at both lateral boundaries, as well as on the bottom of the model.

2.3. Full-Waveform Inversion

The basic concept of the full-waveform inversion (FWI) is to find a model of the subsurface that describes the observed seismic data in the most accurate way. A model is considered as the best, if the misfit between the synthetic data and the observed data is minimized. The FWI uses the full seismic waveforms, hence, the whole information content (e.g. wave amplitude, phase) is taken into account to find an appropriate model. For that reason, it can achieve a resolution below the size of a wavelength. In contrast, conventional seismic tomography methods evaluate only parts of the waveform, like the first-arrival times. However, wave propagation is highly nonlinear and hence the FWI is a nonlinear problem as well. Thus, there is no direct solution to the inverse problem. There are two possible options: (a) To use global grid-search methods or (b) to use local optimization methods. In grid-search methods the misfit is determined for a certain number of random models. The result of the inversion is the model, which holds the lowest misfit between synthetic and observed data. However, for a high number of model parameters this method is computational not affordable, even on current high-performance computing (HPC) systems. Local optimization methods are computational more efficiently. These methods depend on an initial model, which serves as a starting point for the iterative calculation of model updates. The drawback of this method is the requirement of a priori information about the subsurface and the limitation of the solution to the vicinity of the initial model. In the case the initial model is not accurate enough the inversion could end in a local minimum which is far away from the global minimum. However, due to the efficiency of the local optimization methods, these methods are mainly used in the FWI and will be used in this work as well.

2.3.1. Inverse Problem

To formulate the inversion problem of the FWI, we have to parameterize the model space, $\mathbf{m} = (m_1, \dots, m_N)^T$. In seismics two parameterizations are common: A parameterization in terms of density and seismic velocities, which yields $\mathbf{m} = (\boldsymbol{\rho}, \mathbf{v}_S, \mathbf{v}_P)^T$ or in terms of density and the Lamé parameters, which yields $\mathbf{m} = (\boldsymbol{\rho}, \boldsymbol{\mu}, \boldsymbol{\lambda})^T$.

To obtain synthetic data, $\mathbf{d}_{\text{syn}}(\mathbf{m})$, based on a certain model, \mathbf{m} , we can use the nonlinear forward operator, f :

$$\mathbf{d}_{\text{syn}}(\mathbf{m}) = f(\mathbf{m}). \quad (2.26)$$

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 as:

$$\Delta \mathbf{d} = \mathbf{d}_{\text{syn}}(\mathbf{m}) - \mathbf{d}_{\text{obs}}. \quad (2.27)$$

To measure the fit of the synthetic data to the observed data, we use the least-squares L_2 -norm of the data residuals:

$$E(\mathbf{m}) = \frac{1}{2} \cdot \Delta \mathbf{d}^T \cdot \Delta \mathbf{d}, \quad (2.28)$$

where $E(\mathbf{m})$ is called misfit or objective function. Thereby, the objective function refers to a summation of the data residuals over all time samples and all source-receiver pairs. Moreover, this definition has a special physical meaning, since it describes the residual energy which cannot be described by the current synthetic model. The aim of the inversion process is to minimize the objective function iteratively and therewith find a model of the subsurface that explains the observed data.

In the following, we assume only weak nonlinearity of equation 2.28 in order to use the Born approximation. In local optimization methods a local minimum of the objective function is searched in the vicinity of an initial model, \mathbf{m}_0 . Therefore, we add a model perturbation, $\Delta\mathbf{m}$, to the initial model to obtain an updated model:

$$\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{m}. \quad (2.29)$$

We now consider the objective function for this updated model which we expand in a Taylor series up to second-order accuracy:

$$E(\mathbf{m}_0 + \Delta\mathbf{m}) = E(\mathbf{m}_0) + \Delta\mathbf{m} \left(\frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} \right) + \frac{1}{2} \Delta\mathbf{m} \left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial \mathbf{m}^2} \right) \Delta\mathbf{m}^T + \mathcal{O}(\|\Delta\mathbf{m}\|^3). \quad (2.30)$$

To find a minimum of this objective function in the vicinity of the initial model, \mathbf{m}_0 , the derivative with respect to \mathbf{m} is required to vanish:

$$\frac{\partial E(\mathbf{m})}{\partial \mathbf{m}} = \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} + \Delta\mathbf{m} \left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial \mathbf{m}^2} \right) \stackrel{!}{=} 0. \quad (2.31)$$

Rearranging to the model correction, $\Delta\mathbf{m}$, gives the desired model update:

$$\Delta\mathbf{m} = - \left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial \mathbf{m}^2} \right)^{-1} \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} = -\mathbf{H}^{-1} \cdot \nabla_{\mathbf{m}} E(\mathbf{m}_0), \quad (2.32)$$

where $\nabla_{\mathbf{m}} E(\mathbf{m}_0)$ is the gradient of the objective function with respect to the N model parameters \mathbf{m}_0 . The second-order derivative with respect to the model parameters contains the curative information of the objective function and is called Hessian, \mathbf{H} .

With equation 2.29 and 2.32 we obtain the model update for iteration K by:

$$\mathbf{m}_{K+1} = \mathbf{m}_K + \Delta\mathbf{m}_K = \mathbf{m}_K - \mathbf{H}_K^{-1} \cdot \nabla_{\mathbf{m}} E(\mathbf{m}_K). \quad (2.33)$$

This second-order accurate model update is called Newton-method, since it considers the curative information of the objective function (Nocedal and Wright, 2006).

This means, in order to find a Newton-method model update the gradient and the Hessian of the objective function have to be calculated. The gradient can be calculated efficiently by the adjoint state method, which we will discuss in section 2.3.2. However, the Hessian, H , is a dense $N \times N$ matrix, where second-order derivatives of the objective function have to be calculated. The calculation of such second-order derivatives can be cumbersome and computational too expensive for large-scale optimization problems like in the FWI, even on large HPC systems. Therefore, we use a multi-parameter quasi-Newton L-BFGS method, which we will describe in section 2.3.3.

Besides the methods which we have chosen for the calculation of the gradient and Hessian, there are several other methods, like sensitivity kernels or full-Newton methods. A general overview and comparison of different methods within the FWI framework is given by Virieux and Operto (2009).

2.3.2. Gradient calculation: Adjoint State Method

To find a model update it is necessary to calculate the gradient of the objective function. The implementation of this calculation by an actual derivative of the objective function with respect to every single model parameter would need as much forward calculations as model parameters. To overcome this, [Tarantola \(1984\)](#) and [Mora \(1987\)](#) proposed the usage of the adjoint state method within the geophysical community. This method requires only two forward calculations in order to obtain the descent direction of the objective function.

In the following, we derive the gradient for the isotropic and elastic wave equation by the adjoint state method. Subsequently, we will derive the explicit gradients for the P-SV and the SH waves. The shown derivation follows [Mora \(1987\)](#) closely.

To start, we consider the influence of a small model perturbation, $\delta\mathbf{m}$, to synthetic data:

$$\mathbf{d}_{\text{syn}}(\mathbf{m}) = f(\mathbf{m}) = f(\mathbf{m}_0 + \delta\mathbf{m}), \quad (2.34)$$

where we again use the nonlinear forward operator, f . To simplify this expression we approximate the forward operator by a Taylor series, which gives:

$$f(\mathbf{m}_0 + \delta\mathbf{m}) = f(\mathbf{m}_0) + \frac{\partial f(\mathbf{m}_0)}{\partial \mathbf{m}} \delta\mathbf{m} + \mathcal{O}(\|\delta\mathbf{m}\|^2). \quad (2.35)$$

Calculating the difference between the synthetic data of the perturbed model and the synthetic data of the *unperturbed* background model, we get:

$$\delta\mathbf{d} = f(\mathbf{m}_0 + \delta\mathbf{m}) - f(\mathbf{m}_0), \quad (2.36)$$

$$= f(\mathbf{m}_0) + \frac{\partial f(\mathbf{m}_0)}{\partial \mathbf{m}} \delta\mathbf{m} - f(\mathbf{m}_0), \quad (2.37)$$

$$= \left[\frac{\partial f(\mathbf{m}_0)}{\partial \mathbf{m}} \right] \cdot \delta\mathbf{m}. \quad (2.38)$$

This is a linear relation between a perturbation, $\delta\mathbf{d}$, in the data space and a perturbation, $\delta\mathbf{m}$, in the model space. Hereby, the linear operator, $\frac{\partial f(\mathbf{m}_0)}{\partial \mathbf{m}}$, is called Frechét derivative. According to [Mora \(1987\)](#), the linearized forward problem shown in equation 2.38 can be written in continuous form as follows:

$$\delta\mathbf{d}(D) = \int_M dM \frac{\partial \mathbf{d}(D)}{\partial \mathbf{m}(M)} \delta\mathbf{m}(M), \quad (2.39)$$

where D denotes the data space and M the model space. This means, if the kernel, $\frac{\partial \mathbf{d}(D)}{\partial \mathbf{m}(M)}$, is known, all perturbations in the model space can be integrated to find the corresponding perturbation in the data space.

We now consider the local derivative of the objective function in equation 2.31 and reformulate it as follows:

$$\frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} = \left[\frac{\partial \mathbf{d}_{\text{syn}}(\mathbf{m}_0)}{\partial \mathbf{m}} \right]^T \cdot \delta\mathbf{d} = \left[\frac{\partial f(\mathbf{m}_0)}{\partial \mathbf{m}} \right]^T \cdot \delta\mathbf{d}. \quad (2.40)$$

This expression has a similar mathematical form as equation 2.38, even though the Frechét derivative is transposed. To be more general, the Frechét derivative is adjoint to the Frechét derivative in equation 2.38. However, since we have not considered any covariance or weighting, the adjoint becomes the transpose in this special case (Tarantola, 1984). For the sake of generality, in the following we will refer to the adjoint. For more details we recommend the work of Tarantola (2005). Nevertheless, equation 2.40 denotes a linear relation between a perturbation in the data space and a perturbation in the model space. Equation 2.40, which represents the gradient of the objective function, can be given in continuous form by an adjoint expression to equation 2.39 as follows (Mora, 1987):

$$\delta \hat{\mathbf{m}}(M) = \frac{\partial E(\mathbf{m}_0)}{\partial \mathbf{m}} = \int_D \mathbf{d}D \left[\frac{\partial \mathbf{d}(D)}{\partial \mathbf{m}(M)} \right]^* \delta \mathbf{d}(D), \quad (2.41)$$

where the asterisk denotes the adjoint. Note, the hat means that $\delta \hat{\mathbf{m}}$ and $\delta \mathbf{m}$ are not identical. In fact, $\delta \hat{\mathbf{m}}$ refers to the gradient of the objective function, whereas $\delta \mathbf{m}$ can be interpreted as the model correction $\Delta \mathbf{m}$ when $\delta \mathbf{d}$ is interpreted as the data residuals $\Delta \mathbf{d}$. Since the Frechét operator represents a linear operator, its kernel is identical to its adjoint kernel (Tarantola, 2005):

$$\left[\frac{\partial \mathbf{d}(D)}{\partial \mathbf{m}(M)} \right] = \left[\frac{\partial \mathbf{d}(D)}{\partial \mathbf{m}(M)} \right]^*. \quad (2.42)$$

Thus, if the Frechét operator of the linearized forward operator in equation 2.39 can be obtained, the corresponding adjoint equation 2.41 can be used to map from the data space to the model space. This mapping is the key to transfer a data residual to a model correction.

In the next step, we have to transform the shown equations into the seismic inverse problem. In seismics the data, \mathbf{d} , is the discrete particle displacement, u_i . The particle displacement is measured by a receiver at the time, t , at the location, \mathbf{x}_R . The seismic waves are excited by a source at the position, \mathbf{x}_S , at time zero ($t = 0$). We parameterize the model by $\mathbf{m} = (\boldsymbol{\rho}, \boldsymbol{\mu}, \boldsymbol{\lambda})^T$. According to Mora (1987), this parameterization is the simplest one to derive the gradients of the seismic objective function. Other parameterizations can be obtained by a transformation of the final gradients (see section 2.3.2.2). For simplicity, in the following we assume a single source. The total gradients can be obtained by a summation of the shot-wise gradients.

Adjusted to the seismic problem equation 2.39 yields:

$$\delta u_i(\mathbf{x}_r, t) = \int_V dV(\mathbf{x}) \frac{\partial u_i(\mathbf{x}_r, t)}{\partial \mathbf{m}(\mathbf{x})} \cdot \delta \mathbf{m}(\mathbf{x}), \quad (2.43)$$

and equation 2.41 yields:

$$\delta \hat{\mathbf{m}}(\mathbf{x}) = \int dt \sum_{\text{Receivers}} \frac{\partial u_i(\mathbf{x}_r, t)}{\partial \mathbf{m}(\mathbf{x})} \cdot \delta u_i(\mathbf{x}_r, t). \quad (2.44)$$

The aim is to find an expression which is equivalent to the forward problem in equation 2.43 in order to identify the kernel, $\frac{\partial u_i(\mathbf{x}_r, t)}{\partial \mathbf{m}(\mathbf{x})}$. Once we found an expression for this kernel, we can derive the gradients by the corresponding adjoint equation 2.44.

To find an explicit expression of the linearized forward problem in equation 2.43, we reconsider the equation of motion and the elastic stress-strain relationship:

$$\rho \cdot \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i, \quad (2.1)$$

$$\sigma_{ij} = C_{ijkl} \cdot \epsilon_{kl}. \quad (2.8)$$

We now introduce the following perturbations:

$$u_i \rightarrow u_i + \delta u_i, \quad (2.45)$$

$$\rho \rightarrow \rho + \delta \rho, \quad (2.46)$$

$$C_{ijkl} \rightarrow C_{ijkl} + \delta C_{ijkl}, \quad (2.47)$$

$$\sigma_{ij} \rightarrow \sigma_{ij} + \delta \sigma_{ij}, \quad (2.48)$$

$$\epsilon_{kl} \rightarrow \epsilon_{kl} + \delta \epsilon_{kl}. \quad (2.49)$$

Inserting these perturbations into equations 2.1 and 2.8 gives:

$$(\rho + \delta \rho) \cdot \frac{\partial^2}{\partial t^2} (u_i + \delta u_i) - \frac{\partial}{\partial x_j} (\sigma_{ij} + \delta \sigma_{ij}) = f_i, \quad (2.50)$$

$$(\sigma_{ij} + \delta \sigma_{ij}) - (C_{ijkl} + \delta C_{ijkl}) \cdot (\epsilon_{kl} + \delta \epsilon_{kl}) = 0. \quad (2.51)$$

The expressions above can be simplified by defining new body and surface source terms:

$$\Delta f_i = -\delta \rho \frac{\partial^2 u_i}{\partial t^2}, \quad (2.52)$$

$$\Delta T_{ij} = \delta C_{ijkl} \epsilon_{kl}. \quad (2.53)$$

With these source terms, we obtain a wave equation, which describes a displacement perturbation, δu_i , as a function of the introduced source terms:

$$\rho \frac{\partial^2 \delta u_i}{\partial t^2} - \frac{\partial \delta \sigma_{ij}}{\partial x_j} = \Delta f_i, \quad (2.54)$$

$$\delta \sigma_{ij} - C_{ijkl} \cdot \delta \epsilon_{kl} = \Delta T_{ij}. \quad (2.55)$$

The new sources Δf_i and ΔT_{ij} are called secondary or adjoint sources. The new wave equation defined by equations 2.50 and 2.51 has the same mathematical form as the unperturbed wave equation (equations 2.1 and 2.8), thus, we can obtain a solution in terms of Green's functions as follows (Aki and Richards, 2002):

$$\begin{aligned} \delta u_i(\mathbf{x}_r, t) = & + \int_V dV \cdot G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0) * \Delta f_j(\mathbf{x}, t) \\ & - \int_V dV \cdot \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial x_k} * \Delta T_{jk}(\mathbf{x}, t), \end{aligned} \quad (2.56)$$

where the asterisk denotes convolution in time.

This solution of the perturbed wave equation has the same shape as the linearized forward problem in equation 2.43. Consequently, we can use this equation to identify the desired kernel and thus, we can obtain the adjoint expression shown in equation 2.44.

To find a more explicit expression we perform further simplifications. By considering an isotropic medium and by substitution of the corresponding stress-strain relationship 2.8, we can write the adjoint source term, ΔT_{ij} , as:

$$\Delta T_{ij} = \delta_{ij} \cdot \epsilon_{kk} \cdot \delta\lambda + (\epsilon_{ij} + \epsilon_{ji}) \cdot \delta\mu = \delta_{ij} \cdot \frac{\partial u_k}{\partial x_k} \cdot \delta\lambda + \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \cdot \delta\mu. \quad (2.57)$$

Substitution of both adjoint sources into equation 2.56 results in:

$$\begin{aligned} \delta u_i(\mathbf{x}_r, t) = & - \int_V dV \cdot \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial t} * \frac{\partial u_j(\mathbf{x}, t)}{\partial t} \cdot \delta\rho \\ & - \int_V dV \cdot \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial x_j} * \frac{\partial u_m(\mathbf{x}, t)}{\partial x_m} \cdot \delta\lambda \\ & - \int_V dV \cdot \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial x_k} * \left(\frac{\partial u_j(\mathbf{x}, t)}{\partial x_k} + \frac{\partial u_k(\mathbf{x}, t)}{\partial x_j} \right) \cdot \delta\mu. \end{aligned} \quad (2.58)$$

This equation allows us to identify the desired kernels for each parameter class by comparison to the forward problem shown in equation 2.43. We obtain the following adjoint expressions by inserting the kernels into equation 2.44:

$$\delta \hat{\rho}(\mathbf{x}) = - \int dt \sum_R \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial t} * \frac{\partial u_j(\mathbf{x}, t)}{\partial t} \cdot \delta u_i(\mathbf{x}_R, t), \quad (2.59a)$$

$$\delta \hat{\mu}(\mathbf{x}) = - \int dt \sum_R \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial x_k} * \left(\frac{\partial u_j(\mathbf{x}, t)}{\partial x_k} + \frac{\partial u_k(\mathbf{x}, t)}{\partial x_j} \right) \cdot \delta u_i(\mathbf{x}_R, t), \quad (2.59b)$$

$$\delta \hat{\lambda}(\mathbf{x}) = - \int dt \sum_R \frac{\partial G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0)}{\partial x_j} * \frac{\partial u_m(\mathbf{x}, t)}{\partial x_m} \cdot \delta u_i(\mathbf{x}_R, t). \quad (2.59c)$$

We can further simplify these adjoint expressions by the use of the following convolution and integration theorems:

$$\int dt \cdot [f(t) * g(t)] \cdot h(t) = \int dt \cdot f(-t) \cdot [g(t) * h(-t)], \quad (2.60)$$

$$\int dt \cdot f(-t) \cdot g(-t) = \int dt \cdot f(t) \cdot g(t), \quad (2.61)$$

and by considering the reciprocity property of the Green's function:

$$G_{ij}(\mathbf{x}_R, t; \mathbf{x}, 0) = G_{ij}(\mathbf{x}, t; \mathbf{x}_R, 0). \quad (2.62)$$

The simplified adjoint expressions read:

$$\delta \hat{\rho}(\mathbf{x}) = - \int dt \sum_R \frac{\partial u_j(\mathbf{x}, t)}{\partial t} \cdot \frac{\partial}{\partial t} G_{ij}(\mathbf{x}, -t; \mathbf{x}_R, 0) * \delta u_i(\mathbf{x}_R, t), \quad (2.63a)$$

$$\delta \hat{\mu}(\mathbf{x}) = - \int dt \sum_R \left(\frac{\partial u_j(\mathbf{x}, t)}{\partial x_k} + \frac{\partial u_k(\mathbf{x}, t)}{\partial x_j} \right) \cdot \frac{\partial}{\partial x_k} G_{ij}(\mathbf{x}, -t; \mathbf{x}_R, 0) * \delta u_i(\mathbf{x}_R, t), \quad (2.63b)$$

$$\delta \hat{\lambda}(\mathbf{x}) = - \int dt \sum_R \frac{\partial u_m(\mathbf{x}, t)}{\partial x_m} \cdot \frac{\partial}{\partial x_j} G_{ij}(\mathbf{x}, -t; \mathbf{x}_R, 0) * \delta u_i(\mathbf{x}_R, t). \quad (2.63c)$$

We now introduce the backward propagated residual wave field:

$$\Psi_j(\mathbf{x}, t) = \sum_R G_{ij}(\mathbf{x}, -t; \mathbf{x}_R, 0) * \delta u_i(\mathbf{x}_R, t), \quad (2.64)$$

which has the same shape as the Green's function solution in equation 2.56, besides the negative sign in front of the time, t . This negative sign refers to a backward propagation in time. The source term of equation 2.64 corresponds to the data residuals at the receivers. This means that the residual wave field, Ψ , can be obtained by temporal backward propagation of the data residuals from the receiver locations.

Substitution of the residual wave field into the adjoint expressions gives the final gradients for the parameter classes ρ , μ and λ :

$$\delta \hat{\rho}(\mathbf{x}) = \frac{\partial E(\mathbf{m}_0)}{\partial \rho(\mathbf{x})} = - \int dt \cdot \frac{\partial u_j(\mathbf{x}, t)}{\partial t} \cdot \frac{\partial \Psi_j(\mathbf{x}, t)}{\partial t}, \quad (2.65a)$$

$$\delta \hat{\mu}(\mathbf{x}) = \frac{\partial E(\mathbf{m}_0)}{\partial \mu(\mathbf{x})} = - \int dt \cdot \left(\frac{\partial u_j(\mathbf{x}, t)}{\partial x_k} + \frac{\partial u_k(\mathbf{x}, t)}{\partial x_j} \right) \cdot \frac{\partial \Psi_j(\mathbf{x}, t)}{\partial x_k}, \quad (2.65b)$$

$$\delta \hat{\lambda}(\mathbf{x}) = \frac{\partial E(\mathbf{m}_0)}{\partial \lambda(\mathbf{x})} = - \int dt \cdot \frac{\partial u_m(\mathbf{x}, t)}{\partial x_m} \cdot \frac{\partial \Psi_j(\mathbf{x}, t)}{\partial x_j}. \quad (2.65c)$$

These adjoint expressions depend on the forward propagated incident wave field, u , and the backward propagated residual wave field, Ψ . Both can be obtained by the nonlinear forward solver, f , due to the reciprocity of the Green's function. Therefore, only two forward simulations are required to get the adjoint expressions and therewith the gradients in equation 2.65. Thereby, the gradients can be interpreted as a zero-lag cross-correlation between the incident and residual wave field.

We now recapitulate the properties of the assumed 2-D subsurface, to derive the gradients for the P-SV and the SH waves, respectively. In this subsurface all properties in the y direction are constant and derivations in the y direction vanish. Moreover, we reconsider the independency of the wave equations for the P-SV and the SH waves. For simplicity, in the following we omit the temporal and spatial dependencies. The gradients for the SH waves are:

$$\delta \hat{\rho}^{\text{SH}} = - \int dt \cdot \frac{\partial u_y}{\partial t} \cdot \frac{\partial \Psi_y}{\partial t}, \quad (2.66a)$$

$$\delta \hat{\mu}^{\text{SH}} = - \int dt \cdot \left(\frac{\partial u_y}{\partial x} \cdot \frac{\partial \Psi_y}{\partial x} + \frac{\partial u_y}{\partial z} \cdot \frac{\partial \Psi_y}{\partial z} \right), \quad (2.66b)$$

and for the P-SV waves:

$$\delta \hat{\rho}^{\text{P-SV}} = - \int dt \cdot \left(\frac{\partial u_x}{\partial t} \cdot \frac{\partial \Psi_x}{\partial t} + \frac{\partial u_z}{\partial t} \cdot \frac{\partial \Psi_z}{\partial t} \right), \quad (2.67a)$$

$$\delta \hat{\mu}^{\text{P-SV}} = - \int dt \cdot \left[\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x} \right) \cdot \left(\frac{\partial \Psi_x}{\partial z} + \frac{\partial \Psi_z}{\partial x} \right) + 2 \left(\frac{\partial u_x}{\partial x} \frac{\partial \Psi_x}{\partial x} + \frac{\partial u_z}{\partial z} \frac{\partial \Psi_z}{\partial z} \right) \right], \quad (2.67b)$$

$$\delta \hat{\lambda}^{\text{P-SV}} = - \int dt \cdot \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z} \right) \cdot \left(\frac{\partial \Psi_x}{\partial x} + \frac{\partial \Psi_z}{\partial z} \right). \quad (2.67c)$$

2.3.2.1. Stress-velocity formulation

Since we solve the forward problem by a stress-velocity finite-difference scheme, it is more practical to express the gradients in terms of velocity and stress instead of displacement. We therefore use the first-order stress-velocity wave equations to eliminate the displacement as suggested by [Shipp and Singh \(2002\)](#).

We obtain the following stress-velocity gradients for the SH waves:

$$\delta\hat{\rho}^{\text{SH}} = - \int dt \cdot v_y^{\text{F}} \cdot v_y^{\text{B}}, \quad (2.68a)$$

$$\delta\hat{\mu}^{\text{SH}} = - \int dt \cdot \frac{(\sigma_{xy}^{\text{F}} \cdot \sigma_{xy}^{\text{B}} + \sigma_{zy}^{\text{F}} \cdot \sigma_{zy}^{\text{B}})}{\mu \cdot \mu}, \quad (2.68b)$$

and for the P-SV waves:

$$\delta\hat{\rho}^{\text{P-SV}} = - \int dt \cdot [v_x^{\text{F}} \cdot v_x^{\text{B}} + v_z^{\text{F}} \cdot v_z^{\text{B}}], \quad (2.69a)$$

$$\delta\hat{\mu}^{\text{P-SV}} = - \int dt \cdot \left[\frac{(\sigma_{xz}^{\text{F}} \cdot \sigma_{xz}^{\text{B}})}{\mu \cdot \mu} + \frac{1}{4} \left(\frac{(\sigma_{xx}^{\text{F}} + \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} + \sigma_{zz}^{\text{B}})}{(\lambda + \mu)^2} + \frac{(\sigma_{xx}^{\text{F}} - \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} - \sigma_{zz}^{\text{B}})}{\mu \cdot \mu} \right) \right], \quad (2.69b)$$

$$\delta\hat{\lambda}^{\text{P-SV}} = - \int dt \cdot \frac{(\sigma_{xx}^{\text{F}} + \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} + \sigma_{zz}^{\text{B}})}{4(\lambda + \mu)^2}. \quad (2.69c)$$

Thereby the upper indices correspond to the forward (F) propagated incident wave field or to the backward (B) propagated residual wave field, respectively.

2.3.2.2. Parameterization

For the derivation of the gradients we have chosen a model parameterization in terms of the Lamé parameters and density, as proposed by [Mora \(1987\)](#). As a consequence, we also derived the gradients for the Lamé parameters and the density. However, other parameterizations can be resolved with less ambiguities between the individual parameter classes. Especially a parameterization by seismic velocities and density shows less ambiguities ([Köhn et al., 2012](#); [Tarantola, 1986](#)). Hence, we will parameterize the actual FWI experiments by seismic velocities and density $\mathbf{m} = (\boldsymbol{\rho}, \mathbf{v}_S, \mathbf{v}_P)^T$.

The relation between both parameterizations is ([Lay and Wallace, 1995](#)):

$$v_P = \sqrt{\frac{\lambda + 2 \cdot \mu}{\rho}}, \quad v_S = \sqrt{\frac{\mu}{\rho}}, \quad \rho' = \rho; \quad (2.70)$$

and vice versa:

$$\lambda = \rho' \cdot (v_P^2 - 2 \cdot v_S^2), \quad \mu = \rho' \cdot v_S^2, \quad \rho = \rho'. \quad (2.71)$$

We derive the gradients for the seismic velocities and the density by using the chain rule of derivatives (Butzer, 2015). We obtain for v_P :

$$\frac{\partial E}{\partial v_P} = \frac{\partial E}{\partial \lambda} \cdot \frac{\partial \lambda}{\partial v_P} + \frac{\partial E}{\partial \mu} \cdot \frac{\partial \mu}{\partial v_P} + \frac{\partial E}{\partial \rho} \cdot \frac{\partial \rho}{\partial v_P} \quad (2.72)$$

$$= 2 \cdot \rho \cdot v_P \cdot \frac{\partial E}{\partial \lambda}, \quad (2.73)$$

for v_S :

$$\frac{\partial E}{\partial v_S} = \frac{\partial E}{\partial \lambda} \cdot \frac{\partial \lambda}{\partial v_S} + \frac{\partial E}{\partial \mu} \cdot \frac{\partial \mu}{\partial v_S} + \frac{\partial E}{\partial \rho} \cdot \frac{\partial \rho}{\partial v_S} \quad (2.74)$$

$$= -4 \cdot \rho \cdot v_S \cdot \frac{\partial E}{\partial \lambda} + 2 \cdot \rho \cdot v_S \cdot \frac{\partial E}{\partial \mu}, \quad (2.75)$$

and finally for ρ' :

$$\frac{\partial E}{\partial \rho'} = \frac{\partial E}{\partial \lambda} \cdot \frac{\partial \lambda}{\partial \rho'} + \frac{\partial E}{\partial \mu} \cdot \frac{\partial \mu}{\partial \rho'} + \frac{\partial E}{\partial \rho} \cdot \frac{\partial \rho}{\partial \rho'} \quad (2.76)$$

$$= (v_P^2 - 2 \cdot v_S^2) \cdot \frac{\partial E}{\partial \lambda} + v_S^2 \cdot \frac{\partial E}{\partial \mu} + \frac{\partial E}{\partial \rho}. \quad (2.77)$$

Due to the change of parameterization, also the density gradient has changed.

2.3.2.3. Preconditioning

The calculation of the gradients is based on the forward and backward propagated wave fields where the gradients are obtained by a cross-correlation of both wave fields. However, the absolute wave amplitude decays with the distance R from the source by $1/\sqrt{R}$, due to geometrical spreading. Hence, the gradients would have high amplitudes in the vicinity of the sources and receivers and comparatively small amplitudes in the faraway area. Consequently, the model corrections would be concentrated in the area of high gradient amplitude. To compensate this effect and to allow a smooth convergence we apply a preconditioning to the gradients. In this thesis, we use two kinds of preconditioner: (1) A local preconditioner and (2) a global preconditioner.

The first is a local preconditioner which damps the high amplitudes in the vicinity of the sources. In order to do so, we apply a small-scale circular taper to the shot-wise gradients at the location of the source. This taper is designed to set the amplitude at the source position to zero and to gradually increase the amplitude to one within a few grid points. Since we calculate the final gradient as a sum over all shot-wise gradients, we obtain the gradient information at the tapered source position from the sources with shifted source position.

The second is a global preconditioner which accounts for the effect of geometrical spreading. For this purpose, we calculate an approximation to the diagonal elements of the Hessian. This approximation is based on the sum of the amplitudes, u_i , of the forward modelled incident wave field at each grid point. The influence of the receivers is included by a

geometrical estimation of the receivers Green's functions. The approximation for a single shot reads (Plessix and Mulder, 2004; Wehner et al., 2015):

$$\mathbf{H}_a^{-1}(\mathbf{x}_s, \mathbf{x}) = \left[\epsilon + \int dt |u_i(\mathbf{x}_s, \mathbf{x}, t)|^2 \left(\operatorname{asinh} \left(\frac{x_r^{\max} - x}{z} \right) - \operatorname{asinh} \left(\frac{x_r^{\min} - x}{z} \right) \right) \right]^{-1}, \quad (2.78)$$

where x_r^{\max} and x_r^{\min} correspond to the maximum and minimum horizontal receiver position for the source location \mathbf{x}_s , respectively. To stabilize the inversion of the expression above, a water level, ϵ , is added. This water level will be chosen experimentally and will damp the preconditioner in areas where less or no waves propagate. We calculate this preconditioner shot-wise and apply it normalized to the shot-wise gradients. The preconditioned gradient for all shots can be obtained by:

$$\delta \hat{\mathbf{m}}(\mathbf{x}) = \sum_{\text{shots}} \frac{\mathbf{H}_a^{-1}(\mathbf{x}_s, \mathbf{x})}{|\mathbf{H}_a^{-1}(\mathbf{x}_s, \mathbf{x})|} \cdot \delta \hat{\mathbf{m}}(\mathbf{x}_s, \mathbf{x}). \quad (2.79)$$

We introduce a normalization of the preconditioner to prevent a change of the absolute gradient amplitude. Our practical experience revealed, that a manipulation of the gradient amplitude can influence the stability of the quasi-Newton L-BFGS method negatively.

2.3.3. Hessian calculation: L-BFGS

In equation 2.33 we derived a second-order accurate Newton-method model update. To carry out such a model update the Hessian matrix, which contains the second-order derivatives of the objective function, have to be calculated. However, the memory requirements and especially the calculation of the second-order derivatives can be cumbersome and computational too expensive for large-scale optimization problems like the FWI. Thus, instead of a full-Newton method we use a quasi-Newton method. Quasi-Newton methods do not calculate the Hessian matrix explicit, instead they use an approximation of the Hessian, which will be updated at each iteration. This approximation is obtained by measuring the changes in gradients and models, which can be used to construct a model of the objective function. In some cases quasi-Newton methods can be even more efficient than full-Newton methods, since they do not require complex calculations of second-order derivatives (Nocedal and Wright, 2006).

In this work, we use a limited-memory version of the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, which is known as L-BFGS. The memory saving property of the L-BFGS method is achieved by deriving the curative informations from only the n most recent iterations. Thus, only for this n iterations the gradient differences:

$$\mathbf{y}_K = \nabla_{\mathbf{m}}E(\mathbf{m})_{K+1} - \nabla_{\mathbf{m}}E(\mathbf{m})_K, \quad (2.80)$$

and the model differences:

$$\mathbf{s}_K = \mathbf{m}_{K+1} - \mathbf{m}_K, \quad (2.81)$$

have to be kept in memory. Nocedal and Wright (2006), who give a detailed description of the underlying theory, suggest that relative small values of n (<20) are sufficient. Moreover, instead of calculating the Hessian explicitly, the L-BFGS method calculates the inverse Hessian implicit at every iteration. The introduction of the short cut:

$$\rho_K = \frac{1}{\mathbf{y}_K^T \cdot \mathbf{s}_K}, \quad (2.82)$$

allows to perform the L-BFGS method by a simple two-loop recursion, which is shown in algorithm 1. This L-BFGS algorithm returns directly the product of the inverse Hessian with the gradient, as follows:

$$\Delta\mathbf{m}_K = -\mathbf{H}_K^{-1} \cdot \nabla_{\mathbf{m}}E(\mathbf{m})_K. \quad (2.83)$$

To initiate the two-loop recursion the algorithm requires an initial guess to the Hessian, \mathbf{H}_{0K} . Nocedal and Wright (2006) proposed a scaled identity matrix, which estimates the size of the true Hessian:

$$\mathbf{H}_{0K} = \frac{\mathbf{s}_{K-1}^T \cdot \mathbf{y}_{K-1}}{\mathbf{y}_{K-1}^T \cdot \mathbf{y}_{K-1}} \cdot \mathbf{I}. \quad (2.84)$$

This approximation effects that the model update obtained by the L-BFGS method is scaled in physical units, therefore, it can be directly used to update the model. This is

Algorithm 1: L-BFGS two-loop recursion (Nocedal and Wright, 2006)

```

q  $\leftarrow \nabla_{\mathbf{m}}E(\mathbf{m})_K$ ;
for  $i = K - 1, K - 2, \dots, K - n$  do
     $\alpha_i \leftarrow \rho_i \cdot \mathbf{s}_i^T \cdot \mathbf{q}$ ;
     $\mathbf{q} \leftarrow \mathbf{q} - \alpha_i \cdot \mathbf{y}_i$ ;
end
 $\mathbf{r} \leftarrow \mathbf{H}_{0_K} \cdot \mathbf{q}$ ;
for  $i = K - n, K - n + 1, \dots, K - 1$  do
     $\beta \leftarrow \rho_i \cdot \mathbf{y}_i^T \cdot \mathbf{r}$ ;
     $\mathbf{r} \leftarrow \mathbf{r} + \mathbf{s}_i \cdot (\alpha_i - \beta)$ ;
end
stop with result  $\mathbf{r} = \mathbf{H}_K^{-1} \cdot \nabla_{\mathbf{m}}E(\mathbf{m})_K$ ;

```

a significant advantage of the L-BFGS method compared to classical conjugate gradient methods, which need a manual scaling of the model correction. Nonetheless, a more accurate approximation to the initial Hessian would further enhance the performance of the L-BFGS algorithm (Brossier, 2011).

To ensure that the obtained model correction points downhill the objective function, the so-called curvature condition has to be satisfied:

$$\mathbf{s}_K^T \cdot \mathbf{y}_K > 0, \quad (2.85)$$

which implies that the inverse Hessian is positive definite.

With the model correction, $\Delta \mathbf{m}_K$, we can calculate the model, \mathbf{m}_{K+1} :

$$\mathbf{m}_{K+1} = \mathbf{m}_K + \alpha_K \cdot \Delta \mathbf{m}_K, \quad (2.86)$$

where the step length, α_K , is added to the Newton-method update in equation 2.33. An appropriately scaled step length is necessary since the quasi-Newton L-BFGS method underlies the curvature stability condition. This condition can be satisfied by a step length that fulfills the Wolfe conditions, which are shown in section 2.3.3.2. We describe the line search which we use to choose such a step length in section 2.3.3.3.

Since the L-BFGS algorithm rebuilds the objective function from the gradient and model differences, it is required to reset the algorithm every time the objective function is exchanged. For instance, if the frequency content of the seismic data is changed.

2.3.3.1. Multi-parameter L-BFGS

In the FWI a multi-parameter inversion has to be performed, where parameter classes, like density and the shear-wave velocity, with different physical units and different orders of magnitude have to be updated. This would mean that each parameter class requires an own L-BFGS algorithm. As a consequence, it would not be possible to approximate non-diagonal elements of the Hessian, which could reduce trade-off effects between the parameter classes. To overcome this limitation, Brossier (2011) suggests to perform only one dimensionless L-BFGS algorithm by using normalized parameter classes. Thereby,

each parameter class, \mathbf{m}^j (where $j = \{\rho, v_p, v_s\}$), will be normalized by a representative value, m^{j0} . In this work, we use the arithmetic mean value for this normalization. In the following, we denote the gradients of the objective function as $\delta\hat{\mathbf{m}}$, whereas the actual model correction is denoted as $\Delta\mathbf{m}$. The normalized parameters are labeled by an over line. The normalized model reads:

$$\overline{\mathbf{m}}^j = \frac{\mathbf{m}^j}{m^{j0}}. \quad (2.87)$$

To normalize the gradient of each parameter class we use the chain rule of derivation:

$$\delta\overline{\mathbf{m}}^j = \frac{\partial E}{\partial \overline{\mathbf{m}}^j} = \frac{\partial E}{\partial \mathbf{m}^j} \cdot m^{j0}. \quad (2.88)$$

However, a normalized L-BFGS method will result in a normalized model correction. Therefore, the obtained model correction has to be denormalized:

$$\Delta\mathbf{m}^j = \Delta\overline{\mathbf{m}}^j \cdot m^{j0}. \quad (2.89)$$

This approach has the advantage that it allows to reduce the trade-offs between different parameter classes, because of the implicit estimation of non-diagonal elements of the Hessian. The complete algorithm of the multi-parameter L-BFGS method is presented in algorithm 2. This algorithm shows the procedure, which is required to calculate a model correction at iteration K with a L-BFGS method that uses the last n model and gradient differences. The shown algorithm works from the second iteration on, due to the fact that it requires an initial set of differences. Therefore, it is necessary to carry out a classical steepest descent update in the first iteration, where a parabolic line search can be used to estimate a step length (Nocedal and Wright, 2006).

2.3.3.2. Wolfe conditions

In order to ensure that the model correction obtained by the quasi-Newton L-BFGS method will point downhill the objective function, the inverse Hessian has to be positive definite. This can be ensured by a step length, α , which satisfies the Wolfe conditions (Nocedal and Wright, 2006). The first Wolfe condition is called sufficient decrease condition and reads:

$$E(\mathbf{m} + \alpha \cdot \Delta\mathbf{m}) \leq E(\mathbf{m}) + c_1 \cdot \alpha \cdot \nabla_{\mathbf{m}}E(\mathbf{m})^T \cdot \Delta\mathbf{m}, \quad (2.90)$$

where $c_1 \in (0, 1)$ is a constant. This condition requires that the new misfit lies below the straight line defined by the right hand side of the equation above. The second Wolfe condition is called curvature condition and reads:

$$\nabla_{\mathbf{m}}E(\mathbf{m} + \alpha \cdot \Delta\mathbf{m})^T \cdot \Delta\mathbf{m} \geq c_2 \cdot \nabla_{\mathbf{m}}E(\mathbf{m})^T \cdot \Delta\mathbf{m}, \quad (2.91)$$

where $0 < c_1 < c_2 < 1$. This condition rules out too small reductions of the objective function. In contrast, the first Wolfe condition accepts every reduction of the objective function. Nocedal and Wright (2006) suggest to choose $c_1 = 10^{-4}$ and $c_2 = 0.9$ for the quasi-Newton L-BFGS method.

Algorithm 2: Multi-parameter L-BFGS at iteration K (modified after [Brossier \(2011\)](#))

```

input gradients and models;
if  $K > n$  then
  | Discard the vector pair  $(\mathbf{y}_{K-n-1}, \mathbf{s}_{K-n-1})$ 
end
if  $K > 1$  then
  | for  $j = \{\rho, v_p, v_s\}$  do
  | | Store normalized model differences  $\mathbf{s}_{K-1} = (\mathbf{m}_K^j - \mathbf{m}_{K-1}^j) / m^{j0}$ ;
  | | Store normalized gradient differences  $\mathbf{y}_{K-1} = (\delta \hat{\mathbf{m}}_K^j - \delta \hat{\mathbf{m}}_{K-1}^j) / m^{j0}$ ;
  | end
end
Compute normalized model correction  $\Delta \bar{\mathbf{m}}$  with algorithm 1;
for  $j = \{\rho, v_p, v_s\}$  do
  | Denormalize model correction  $\Delta \mathbf{m} = \Delta \bar{\mathbf{m}} \cdot m^{j0}$ ;
end
Search appropriate step length  $\alpha_K$  with algorithm 3;
stop with final model correction  $\alpha_K \cdot \Delta \mathbf{m}$ ;

```

2.3.3.3. Line search

The step length search for the L-BFGS method is initiated with a step length of $\alpha = 1$, due to this step length fulfills the Wolfe conditions in most cases ([Nocedal and Wright, 2006](#)). In case that the step length of the previous iteration fulfills the Wolfe conditions, this step length will be retried. However, if it does not fulfill the Wolfe conditions, a line search algorithm is called. In algorithm 3 we show such a line search, which is based on the SEISCOPE OPTIMIZATION TOOLBOX. This algorithm is called until an appropriate step length is found. Nevertheless, in order to save calculation time, the L-BFGS updated will be skipped, if after a certain number of step length searches no step length could be found that satisfies the Wolfe conditions. Then a steepest descent update will be carried out, where the step length is estimated by a parabolic fit. In such a case the L-BFGS algorithm has to be reseted, because its stability is not ensured. It is important to mention, that every time the Wolfe conditions are checked the full gradient for the current test step length has to be calculated at the cost of two forward runs per source position.

2.3.4. Multi-stage approach

To make the full-waveform inversion more robust against local minima and to allow for a smooth convergence we use a multi-stage approach. The multi-stage approach divides the inversion process in different stages that are performed successively. A specific stage can refer, for instance, to a certain amount of data or to a specific parameter class that will be updated.

The amount of data that is fitted by the FWI can be controlled, for example, by a band-pass frequency filter or an offset based spatial filter. Thereby, the frequency filtering

Algorithm 3: Line search (SEISCOPE OPTIMIZATION TOOLBOX Manual)

```

Initialize  $\alpha_{\min}$  and  $\alpha_{\max}$  to 0 and  $\alpha$  to 1;
while  $\alpha$  do not satisfy the Wolfe conditions do
  if the first condition is not satisfied then
     $\alpha_{\max} = \alpha$ ;
     $\alpha = 0.5 \cdot (\alpha_{\min} + \alpha_{\max})$ ;
    continue;
  end
  if the second condition is not satisfied then
     $\alpha_{\min} = \alpha$ ;
    if  $\alpha_{\max} = 0$  then  $\alpha = 10 \cdot \alpha$ ;
    if  $\alpha_{\max} \neq 0$  then  $\alpha = 0.5 \cdot (\alpha_{\min} + \alpha_{\max})$ ;
    continue;
  end
end
stop with step length  $\alpha$  that satisfy the Wolfe conditions;

```

is of particular importance to allow for a smooth convergence. Since the wave lengths correlate to frequencies contained in the dataset, it is possible to use a frequency filter to control the wave lengths that are considered in the inversion. Moreover, the frequency content highly influences the shape of the objective function. In figure 2.3 we show an objective function for five different frequency ranges. From the top to the bottom the frequency content contained in the dataset is lowered. With decreasing frequency content the objective function becomes smoother and the number of local minima decreases significantly. Hence, it will become less likely that the FWI end in a local minimum far away from the global minimum.

To avoid trade-offs and cross-talk between different parameter classes, we also apply the multi-stage approach to the multi-parameter inversion itself. Cross-talk and trade-offs describe the effect that an anomaly which is present in only one parameter class, will be foot-printed into other parameter classes during the inversion. This is possible, due to the ambiguity of the inverse problem, where trade-offs between parameters can explain the dataset equally accurately. A well proven concept to reduce such effects is to successively start the update of each individual parameter class. To the beginning of the inversion only the parameter class is updated, which the waves have the highest sensitivity to, for instance, the S-wave velocity in the case of surface waves. After some progress in terms of misfit reduction, the second most sensitive parameter is updated as well and finally a full multi-parameter inversion can be used. This allows on the one hand to avoid cross-talk and on the other hand can speed-up the convergence. Köhn et al. (2012) give a detailed analysis of cross-talk effects and strategies to reduce them.

We realize the multi-stage approach within the FWI algorithm by an automatic workflow implementation. The workflow defines the different inversion stages which divide the inversion process. Each workflow stage is applied to the inversion until an automatic abort criterion (AC) for this individual stage is reached. As criterion we use the reduction

of the objective function from the second last to the current iteration. The abort criterion for iteration K can be formulated by an inequality as follows:

$$AC_K(\text{in } \%) > \left(\frac{E_{K-2} - E_K}{E_{K-2}} \right), \quad (2.92)$$

where the workflow stage will be switched in the case the inequality is fulfilled. By this criterion we can ensure that the inversion makes enough progress in terms of misfit reduction within a workflow stage, otherwise the algorithm switches to the next inversion stage. Thereby it can be more efficient to switch to a new workflow stage early, than to invest a lot of iterations to find the exact location of the local minimum within the current workflow stage. This is comparable to the reason that inaccurate line searches can be more efficient than accurate ones, since it is faster to go in a new direction than searching for the exact minimum along a given direction (Nocedal and Wright, 2006).

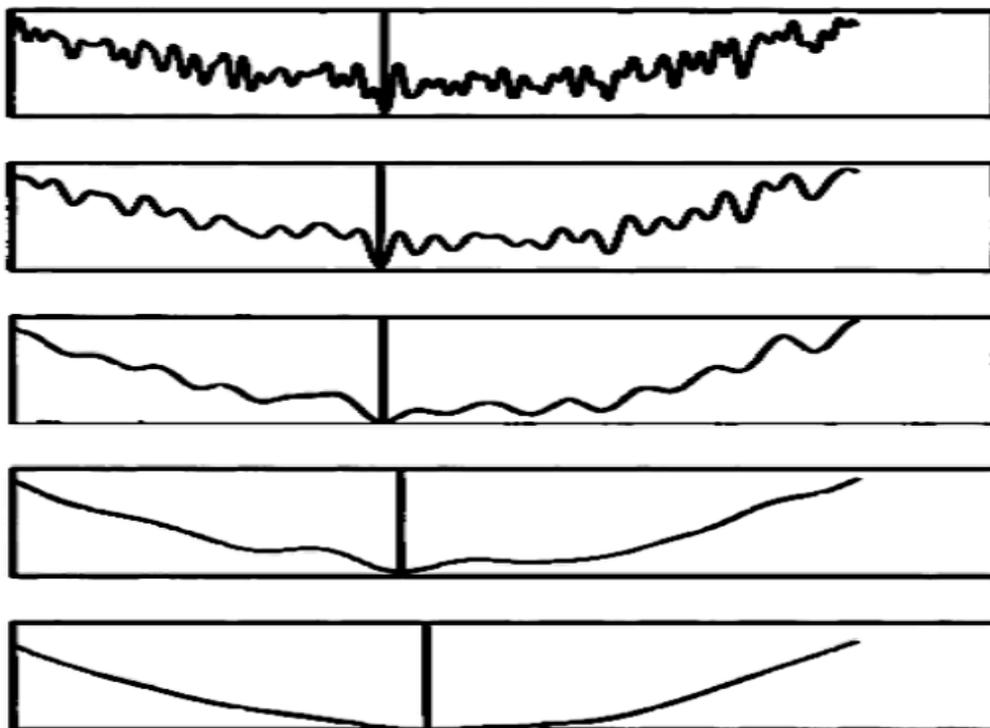


Figure 2.3.: Objective function for different scale lengths. The scale length increases from the top to the bottom. Source: [Bunks et al. \(1995\)](#).

2.4. Simultaneous Joint Inversion

In two dimensions the propagation of the P-SV and the SH waves is described by two independent wave equations, thus, the forward as well as the inverse problem of both wave types is decoupled. As a consequence, a manual coupling has to be applied to both individual inversions, which we refer to as joint approach, in order to carry out a joint inversion of both wave types. We call the joint inversion a simultaneous joint inversion, since we will invert the information content of both wave types at the same time in a single inversion. The aim of the joint inversion is to improve the final inversion result and to decrease the vulnerability to local minima and to ambiguities by making use of more information. The joint inversion allows to consider the full information content exploited in a 2-D seismic measurement, since the full 2-D three-component seismic dataset can be inverted simultaneously. This allows a significant increase of the wave coverage in the subsurface, hence, the joint dataset contains redundant information on the subsurface, which will improve the reliability of the inversion result. The first step to couple both individual wave type inversions is to bound both to one single parameter model, because the joint inversion should reveal a single parameter model that accounts for both datasets. Second, we have to merge both individual objective functions, in order to measure the total fit of the synthetic data to the observed data. Moreover, to combine the model update of both individual inversions, we have to apply a joint approach to the gradients of both inversions, since the P-SV as well as the SH inversion are sensitive to the S-wave velocity and to the density. The SH waves are not sensitive the P-wave velocity, we therefore do not apply a joint approach to the update of the P-wave velocity.

In the following, we introduce the joint approach to combine the objective functions and the gradients of both individual wave type inversions.

2.4.1. Joint objective function

To obtain a single measure of the fit between the synthetic and the observed data for the P-SV as well as for the SH waves, we have to combine both individual objective functions. The objective function introduced in equation 2.28 describes the residual energy that cannot be explained by the current synthetic model. However, this definition is not practical for the combination of both inversions, since both could contain a different amount of energy, which does not necessarily correspondent to the information quantity included or to the reliability of the specific dataset.

We therefore weight the residual energy with the energy of the observed data:

$$E_w(\mathbf{m}) = \frac{1}{2} \cdot \frac{\Delta \mathbf{d}^T \cdot \Delta \mathbf{d}}{\mathbf{d}_{\text{obs}}^T \cdot \mathbf{d}_{\text{obs}}}. \quad (2.93)$$

This weighted objective function is defined as ratio between the residual energy and the energy of the observed dataset. A ratio of one would indicate that the residual energy is as big as the energy in the observed dataset.

In general, our joint approach is not restricted to this specific kind of weighting. A normalization of each objective function with its initial value at the beginning of a new

inversion stage is also possible. In our research the energy based weighting is used, since it seems more intuitive to us.

Since both wave types should be weighted equally, we use a simple addition of both weighted objective functions to calculate the joint objective function:

$$E^{\text{JOINT}}(\mathbf{m}) = E_w^{\text{P-SV}}(\mathbf{m}) + E_w^{\text{SH}}(\mathbf{m}). \quad (2.94)$$

In the joint inversion this joint objective function is used for the steering through the parameter space. Furthermore, the quasi-Newton L-BFGS method will approximate the Hessian implicitly based on this objective function.

2.4.2. Joint gradients

Both individual wave type inversions are sensitive to the S-wave velocity as well as to the density. Consequently, both return gradients for these two parameter classes, which have to be combined for the sake of a joint inversion. This combination is not as simple as the coupling of the objective functions, due to the lack of an intuitive normalization. The amplitude of a gradient depends on the slope of the objective function, since a gradient is the derivative of an objective function. This means in our case that both gradients would only have a similar amplitude, if both objective functions hold a similar slope. However, this is not necessarily fulfilled since both inversions have their own objective function.

We therefore propose a normalized addition of both gradients that is followed by a scaling with the sum of the used normalization factors. We choose the maximum absolute gradient amplitude as normalization factor, respectively. The joint gradient reads:

$$\delta \hat{\mathbf{g}}^{\text{JOINT}} = \left[\frac{\delta \hat{\mathbf{g}}^{\text{P-SV}}}{\max(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|)} + \frac{\delta \hat{\mathbf{g}}^{\text{SH}}}{\max(|\delta \hat{\mathbf{g}}^{\text{SH}}|)} \right] \cdot \left(\max(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|) + \max(|\delta \hat{\mathbf{g}}^{\text{SH}}|) \right), \quad (2.95)$$

where $\delta \hat{\mathbf{g}} = \{\delta \hat{\rho} \text{ or } \delta \hat{\mathbf{v}}_s\}$.

Since the SH waves are not sensitive to the P-wave velocity, we do not calculate a joint gradient for this parameter. This joint gradient approach weights both gradients equally and preserves their amplitude information. The latter is important to provide the amplitude information to the L-BFGS algorithm, which relies on the evaluation of the gradient differences. In the case of a purely gradient-based inversion, like steepest descent, a normalized addition is sufficient since mostly the gradient amplitude is not further evaluated.

2.5. Summary

In this chapter, we presented the process of the full-waveform inversion as a whole. The inversion is iterative, this means, model updates are calculated with respect to the previous model. Therefore, it is sufficient to describe the procedure of one iteration, since the same procedure is repeated in all iterations. In figure 2.4 we illustrate the tasks that the algorithm carries out in the K -th iteration .

The first step in each iteration is to call the automatic workflow implementation. This implementation controls the configuration of the inversion by applying the workflow. The workflow divides the inversion into separate stages and allows for a multi-stage inversion. A stage can refer, for example, to a certain frequency filter or to an offset based trace filter. The workflow implementation uses an automatic abort criterion to switch to new workflow stages, each time this criterion is fulfilled. The criterion is based on the evolution of the objective function. The multi-stage strategy is essential to mitigate local minima of the objective function.

After the workflow is applied, the inversion algorithm enters the main part. The main part consists of two independent inversion engines (grey boxes). The first inversion engine is the P-SV wave engine, which can invert the vertical, d_z , and the horizontal inline displacement component, d_x . The other inversion engine is the SH wave engine, which can invert the horizontal crossline displacement component, d_y . To invert the full-waveforms each engine uses the classical adjoint state method. This method allows to calculate the gradients of the objective function by two forward simulations. First, a forward modelling is done based on the current synthetic model, where the calculated wave field is stored in order to be used in the gradient calculation. After the forward modelling, the data residuals and the data misfit are calculated. The data residuals are then back-propagated from the receiver positions. During this back propagation the gradients can be calculated as zero-lag cross-correlation between the stored forward propagated incident wave field and the back-propagated residual wave field. Each engine returns a data misfit and a gradient for each parameter class it is sensitive to.

After the gradients are obtained, the algorithm allows for preconditioning of the gradients in order to speed-up the convergence and to avoid local minima.

In the case of an individual wave type inversion, the algorithm executes only one of both engines and in the case of a simultaneous joint inversion it executes both engines in parallel. The joint inversion allows a full multi-component and multi-parameter inversion. However, in case of a joint inversion the gradients and the objective function of both individual inversions have to be combined. For this combination the algorithm calls the joint approach.

Afterwards, the quasi-Newton L-BFGS method is used to calculate the model correction. A line search determines a step length that fulfills the Wolfe conditions, in order to ensure the stability of the L-BFGS algorithm. Finally, the algorithm calculates the model at iteration $K + 1$ by the step length and the model correction.

This process is repeated iteratively until the inversion converges to a solution. As convergence criterion the reduction of the objective function is used. In the case the inversion algorithm is not able to reduce the objective function in two following iterations more than a given value, it is assumed that the inversion reached convergence to a local solution.

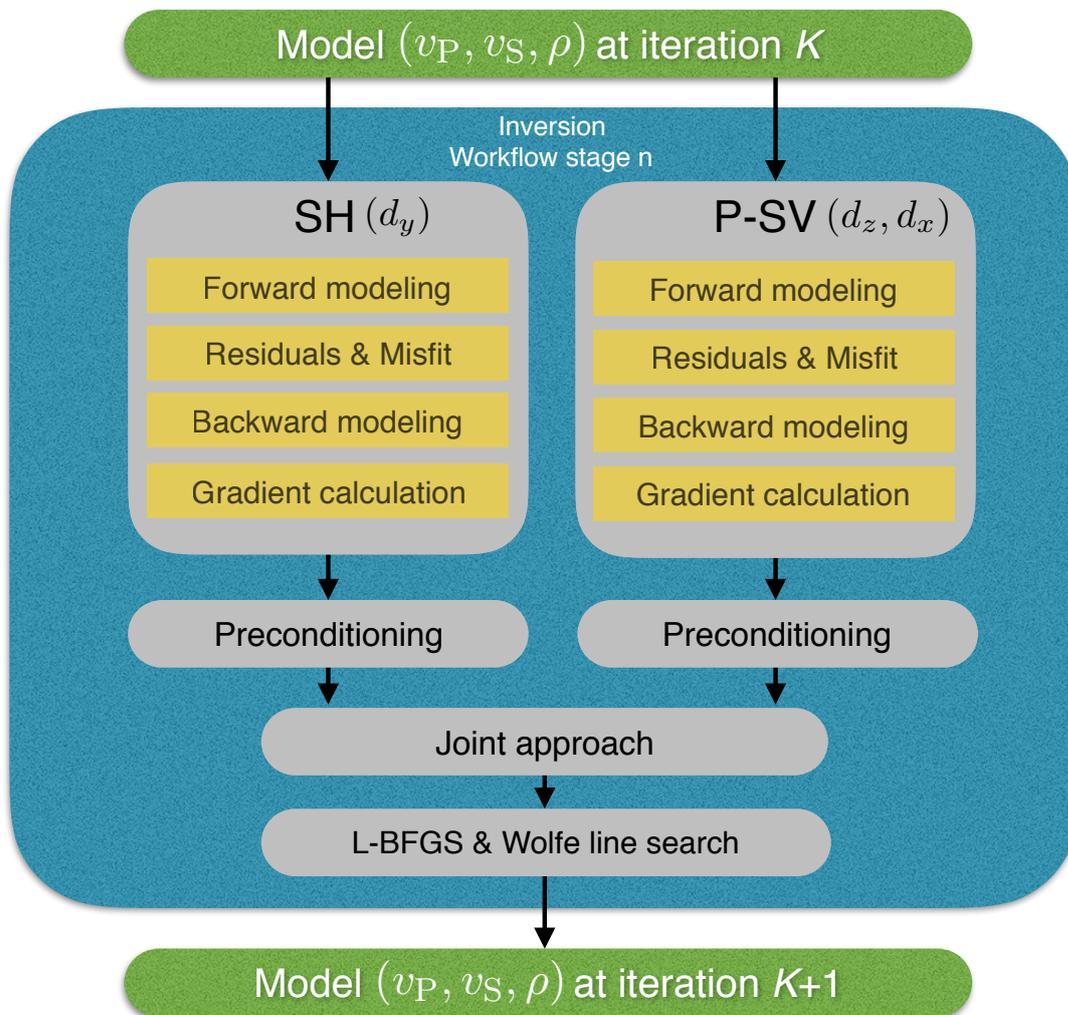


Figure 2.4.: Flow chart of the FWI algorithm to obtain a model correction for a single iteration. The blue box corresponds to the inversion part, which is performed for a certain workflow stage n .

3. Synthetic FWI experiments

In this chapter, we perform synthetic reconstruction tests to explore the properties of the individual and the joint 2-D elastic full-waveform inversion of Rayleigh and Love waves. For these reconstruction tests we assume a true subsurface model which we then use to generate pseudo-observed seismograms. The knowledge of this model allows us to directly study the reconstruction ability of the FWI by comparing the true model to the reconstructed models. In preparation of the field measurements (chapter 4), we use a subsurface model that simulates the expected model at the test site. So far several Rayleigh wave FWI studies have been carried out on the same test site which allow us to derive the subsurface model. In this synthetic study, we perform an individual Love wave FWI, an individual Rayleigh wave FWI and a simultaneous joint FWI of both wave types.

The structure of this chapter is as follows: First, we introduce the test setting, the assumed true subsurface model and the initial model for the reconstruction tests. Second, we present the inversion results of the three FWI studies: (a) Individual Love wave FWI, (b) individual Rayleigh wave FWI and (c) simultaneous joint FWI. We then perform a direct comparison of the inversion results. In a case study, we research the influence of the initial P-wave velocity model. Finally, we summarize the main results of the synthetic experiments.

3.1. Setting

In this section, we introduce the subsurface model, the acquisition geometry and the FWI workflow, which we will use in the synthetic reconstruction experiments. We choose the test setting to be close to the field measurements, in order to obtain relevant information which will be required in chapter 4, where we will present the application of our methodology to a near-surface field dataset.

3.1.1. True and initial models

To study the reconstruction properties of the three FWI schemes, we need to assume a true subsurface model. In preparation of the field measurements we choose the synthetic model to be close to the subsurface model at the desired test site. The location of the test site is on a glider airfield in Rheinstetten near Karlsruhe (Germany). We obtain the subsurface model for this location from previous Rayleigh wave FWI studies (Groos, 2013; Groos et al., 2014; Binnig, 2015). Their inversion results suggest a predominantly depth dependent 1-D background model, which is superimposed by a shallow small-scale low-velocity trench. For simplicity, we assume a purely elastic subsurface in our synthetic experiments.

In order to draw conclusions about the resolution in each individual parameter class as well as to explore trade-off and cross-talk effects between them, we shift the horizontal location

of the trench in each parameter class. In v_S , we place the trench in the middle of the model space. In v_P , we shift the trench by 5 m to the right and in ρ by 5 m to the left, relative to the v_S trench respectively. We hold the horizontal displacement between the three parameter class small to ensure a similar wave coverage for all three anomalies. We show the assumed true subsurface model in figure 3.1 (left) and present depth profiles across the three elastic parameters in figure 3.2. The model has a size of 60 m in the horizontal and of 16 m in vertical direction. In a depth of 6.3 m, we assume a water table where all elastic parameters contain a sharp contrast. Below the water table, all three elastic parameters are homogeneous. Above the water table, the background model consists of a gradient model for v_S and of a homogeneous layer for v_P and ρ . The gradient in v_S is steep in the uppermost part and becomes weaker in a depth of 1 m. The trench has a triangular shape and a length of 10 m at the surface. The lower edge of the trench lies in a depth of 3.5 m. The maximum anomaly of the trench with respect to the background model is 55% in v_S , 28% in v_P and 12% in ρ .

In figure 3.1 (right) we present the initial model for the inversion and in figure 3.2 the corresponding depth profiles. The initial models of v_S and ρ consist of linear gradient models up to a depth of 9 m. Below 9 m the initial models of this two parameters are identical to the true models. In contrast to the initial models of v_S and ρ , we use a high amount of a priori information for the initial v_P model, where we use the true background model as initial model. We thereby assume, that a simple two-layer v_P model like in this synthetic test could be obtained in a field measurement in a similar quality, for instance by a P-wave travel time analysis.

3.1.2. Acquisition geometry

For the seismic acquisition we use 48 multi-component receivers. We distribute them along the model surface with an equidistant spacing of 1 m. We place each receiver one grid point below the free surface to ensure an accurate amplitude scaling (Groos, 2013). In case of the P-SV simulations the receivers record the vertical, v_z , velocity component and in case of the SH simulations they record the horizontal crossline, v_y , velocity component. The orientation of the components can be identified in figure 2.1, where the profile lies on the x -axis. We distribute the sources along the surface as well, where we place them in the same depth as the receivers. We set the equidistant spacing of the sources to 2 m. For the P-SV simulations the source is a vertical force and for the SH simulations the source is a horizontal force in the crossline direction. These source types correspond to vertical and horizontal hammer blows in field measurements. As a source signal, $s(t)$, we choose a cubed sine:

$$s(t) = \sin(f_d \cdot \pi \cdot t)^3, \quad (3.1)$$

where f_d is the dominant frequency. We set the dominant frequency to 30 Hz, which generates frequencies between 0 and 60 Hz. We show the used source signal in the time-domain as well as in the frequency-domain in figure 3.3. In figure 3.1 we illustrate the positions of the 24 sources by yellow stars and of the receivers by red points. The configuration of the acquisition geometry of our synthetic experiments is identical to our field measurements.

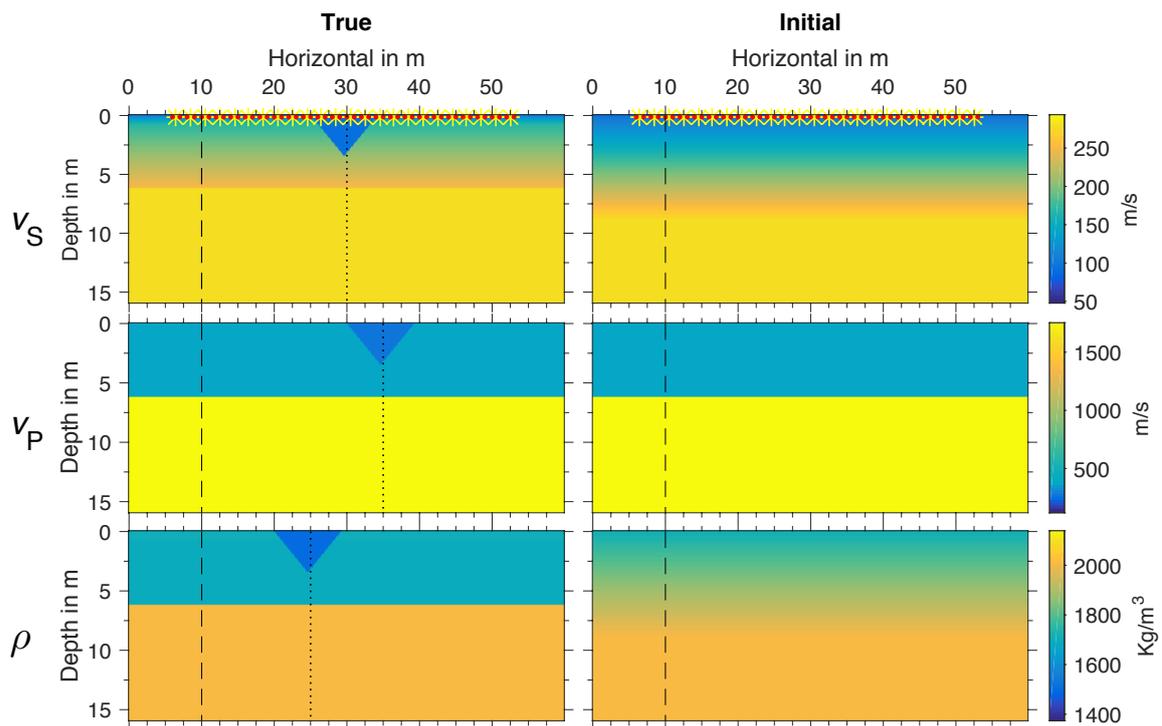


Figure 3.1.: True subsurface model for v_S , v_P and ρ used in the synthetic FWI experiments (left). Initial model for the inversion (right). The yellow stars indicate the source positions and the red points the receiver positions. Depth profiles for the dotted and dashed lines are shown in figure 3.2.

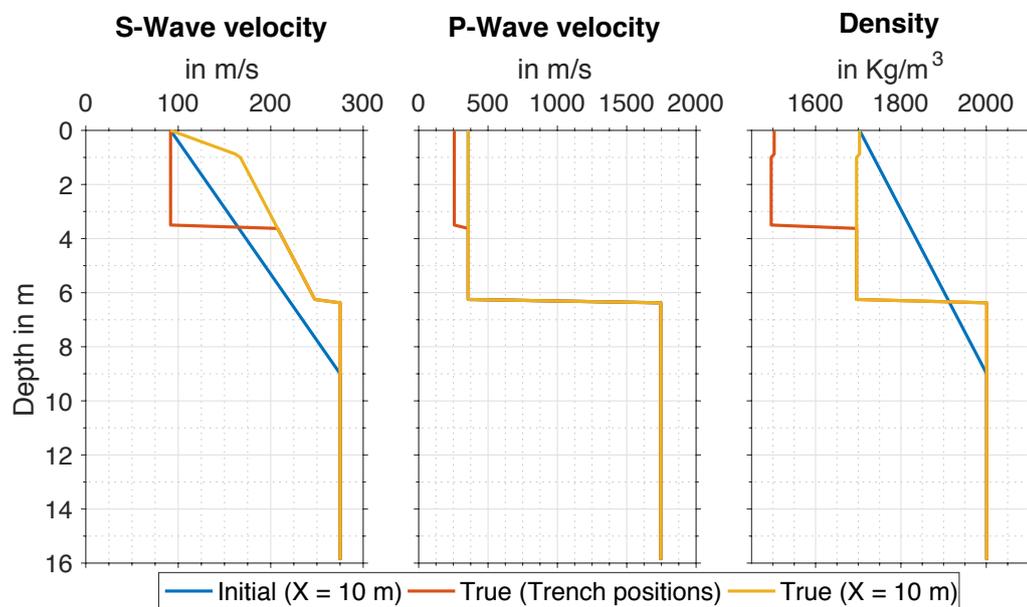


Figure 3.2.: Depth profiles of the elastic parameter v_S , v_P and ρ for the true model and for the initial model, as indicated in figure 3.1. At depths, where graphs are not visible, they are identical to the yellow graph.

3. Synthetic FWI experiments

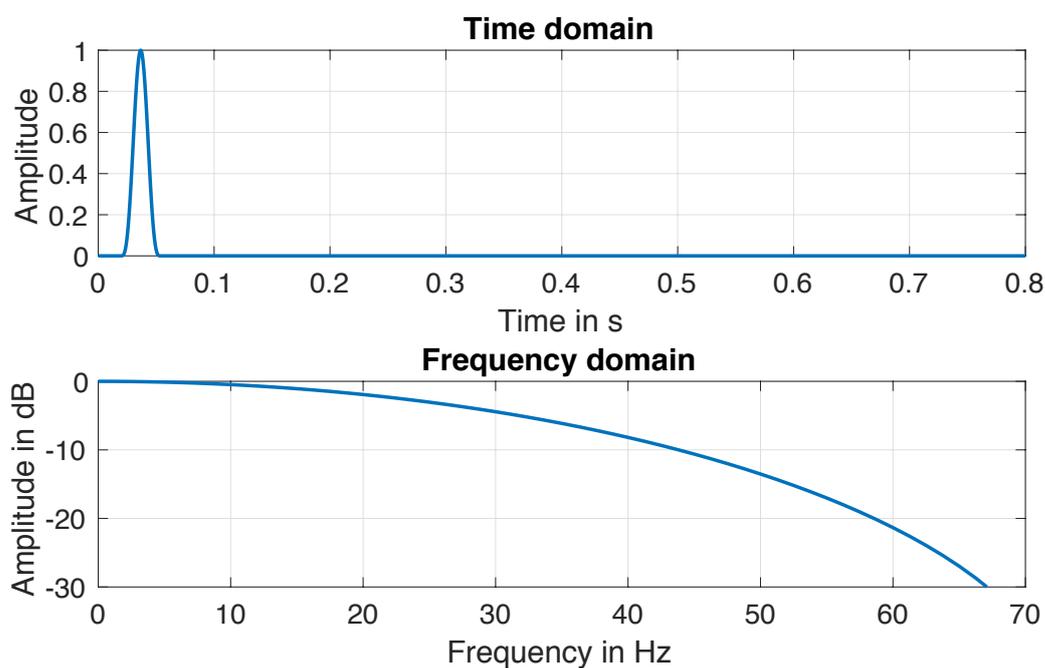


Figure 3.3.: Cubed sine with a main frequency of 30 Hz, which is used as source signal in the synthetic FWI experiments. The upper figure corresponds to the time-domain and the lower one to the frequency-domain.

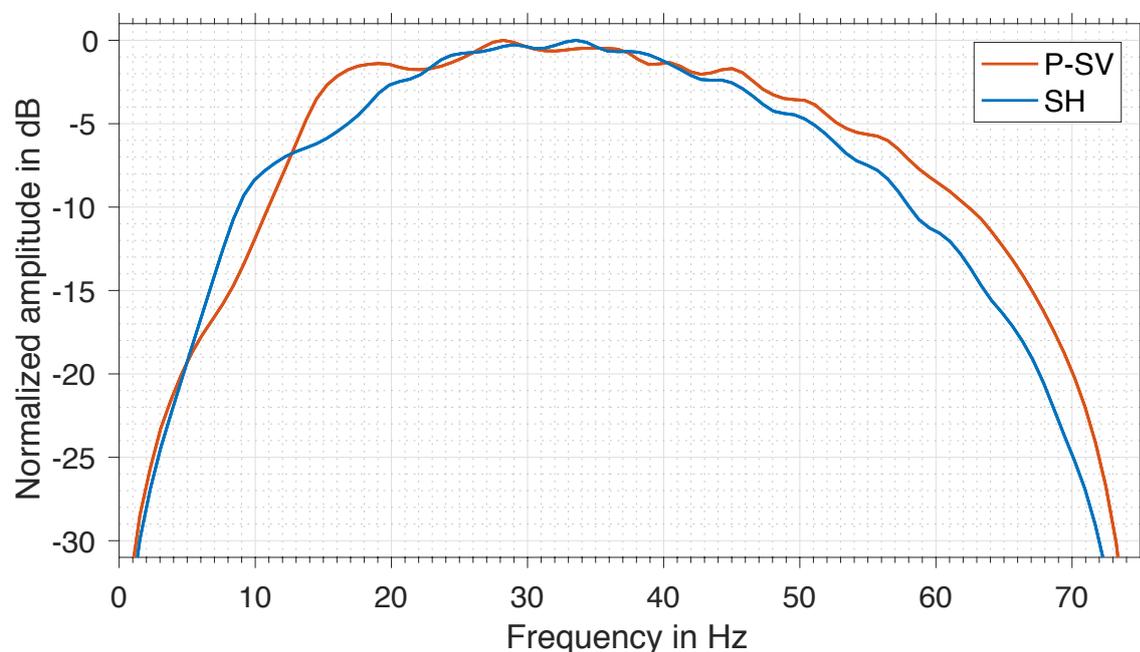


Figure 3.4.: Amplitude spectra of the pseudo-observed seismograms. The spectrum of the P-SV waves is shown in red and of the SH waves in blue. The shown spectra are the average of the spectra of the normalized traces.

3.1.3. Pseudo-observed data

We generate the pseudo-observed dataset based on the assumed true subsurface model and the shown acquisition geometry. For the simulation of these data we use the build-in finite-difference forward solver of the used FWI software code (section 3.1.4). For the finite-difference modelling, we discretize the model on an equidistant grid with a spacing of 0.125 m. On the top of the model space we apply a free surface and at the bottom as well as on the lateral boundaries we assume a C-PML boundary (section 2.2.2). We set the temporal sampling to $2 \cdot 10^{-5}$ s and the total recording time to 0.8 s. A detailed description of the parameters used for the forward-modelling can be found in appendix A in table A.1. In figure 3.5 we show an exemplary shot-gather for the P-SV velocity component v_z as well as for the SH component v_y . The corresponding source position for the shown shot-gather lies at the profile coordinate of 6.5 m (figure 3.1). We normalized the seismograms trace-wise and low-pass filtered them to 60 Hz, since no higher frequencies are used in the reconstruction experiments. In addition, we applied a 4 Hz high-pass filter, in order to simulate the response of geophones with an eigenfrequency of approximately 4.5 Hz. The P-SV seismogram, v_z , consists mainly of the Rayleigh surface wave. The Rayleigh wave carries most of the seismic energy and is by far the most dominant signal in this recording. P-wave onsets are visible, but they have smaller amplitude compared to the Rayleigh wave. The SH seismogram, v_y , shows no P-wave onset, which is obvious since only SH waves are propagated in SH simulations. Nevertheless, the v_y seismogram is dominated by surface waves as well, which we identify as Love waves. At an offset of about 25 m the influence of the low-velocity trench is visible in both recordings. At this offset, the wave fields are scattered and seismic energy is reflected backwards. In figure 3.4 we present the amplitude spectra of the frequency filtered seismograms. The spectra contain frequencies higher than 60 Hz, due to the cut-off frequency of the fourth-order Butterworth low-pass filter is not sharp. The frequency content of both recordings shows no significant difference.

3.1.4. FWI setup

For the synthetic reconstruction tests we use the FWI software code IFOS (*Inversion of Full Observed Seismograms*), which is based on the FWI code of Köhn (2011). IFOS is maintained by the Geophysical Institute of the Karlsruhe Institute of Technology and is available freely under the GNU license on opentoast.de. This software code provides a time-domain finite-difference method for the forward modelling during the inversion, which we configure identical as for the simulation of the pseudo-observed dataset (see table A.1). For the gradient calculation the FWI software uses the classical adjoint state method. To obtain the actual model update we use a normalized multi-parameter L-BFGS method. Thereby the model and gradient differences of the last 20 iterations are used to estimate the curative information of the objective function. We assume the source wavelet to be a priori known, hence, we use the same wavelet in the inversion as for the generation of the pseudo-observed seismograms.

As objective function we use the weighted L_2 -error (equation 2.93) between the normalized synthetic seismograms and the normalized pseudo-observed seismograms. We apply the normalization trace-wise, in order to be consistent with the field data inversions and also to

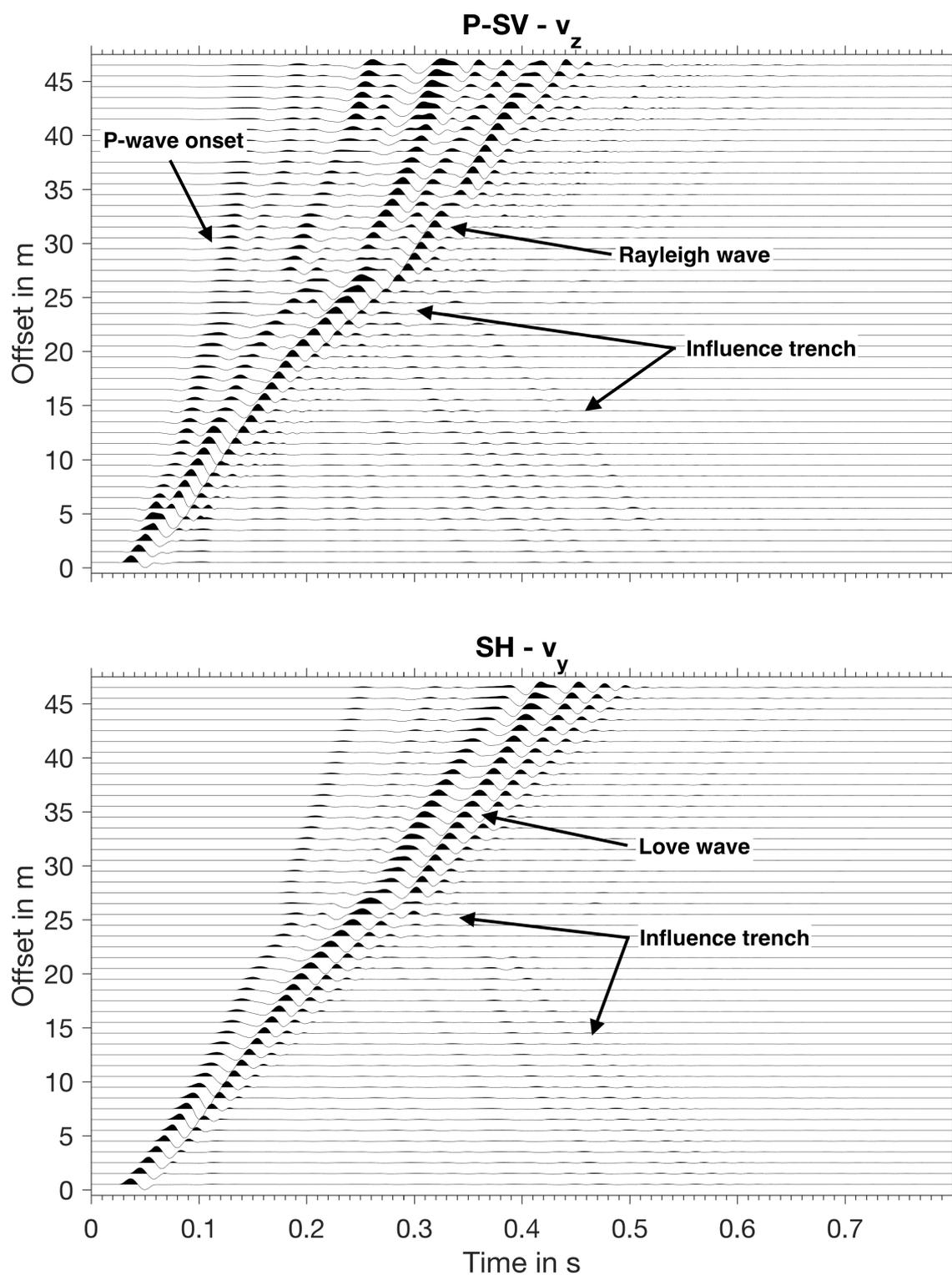


Figure 3.5.: Trace-wise normalized shot-gather of the pseudo-observed seismic data. The corresponding source is located at the profile coordinate of 6.5 m. The seismograms are band-pass filtered between 4 and 60 Hz, representing the frequency range used in the FWI.

be able to account for uncertainties and differences in the coupling of individual geophones to the ground (Maurer et al., 2012). In the P-SV case, we calculate the objective function for the vertical displacement component and in the SH case for the horizontal crossline displacement component. We obtain the displacement seismograms from the recorded velocity seismograms by numerical integration.

To precondition the shot-wise gradients we apply circular source tapers, which decay within 5 grid points from a value of one to zero at the actual source position, where zero refers to absolute damping. Moreover, we use the approximation of the diagonal elements of the Hessian (equation 2.78) as a preconditioner for the gradient. We set the water level, ϵ , empirically to $5 \cdot 10^{-3}$ for the SH waves and to $5 \cdot 10^{-2}$ for the P-SV waves. We apply this approximation empirically iteration-wise to the SH gradients and shot-wise to the P-SV gradients. In addition, the gradients become smoothed with a 2-D median filter, which has a size of 50 cm (4 grid points). To ensure stability of the forward solver as well as to obtain a physical meaningful elastic parameter model, we force a minimum v_P/v_S ratio of 1.2 during the inversion by increasing the P-wave velocity, if necessary.

We control the multi-stage inversion process by an automatic workflow implementation. We use this workflow implementation to increase the frequency content of the dataset gradually from 4 to 60 Hz, by increasing the corner frequency, LC , of a fourth-order Butterworth low-pass filter. Moreover, we apply a successive update strategy to the multi-parameter inversion by the workflow implementation. In the first iterations we only allow updates of v_S , until the automatic abort criterion, AC , (equation 2.92) is reached. Afterwards, in case of the individual P-SV as well as of the joint inversion we update v_S and v_P simultaneously, again until the automatic abort criterion is reached and finally we use a full multi-parameter inversion. We decide to use this successive update of the parameter classes to account for the different sensitivities of the surface waves to the individual elastic parameters. Their propagation is mainly influenced by the S-wave velocity. We divide the workflow in eight separate stages, which we describe in table 3.1. We force the algorithm to carry out at least three iterations at the beginning of each workflow stage. However, in the case that the step length estimation fails, the algorithm directly switches to the next stage.

Table 3.1.: Workflow used in the synthetic FWI experiments. Each stage is applied to the inversion until the automatic abort criterion AC is reached. The update columns indicate which of the specific elastic parameter is updated (yes=1) or not (no=0). LP represents the corner frequency of the low-pass frequency filter.

Stage	Update			LP in Hz	AC in %
	v_S	v_P	ρ		
1	1	0	0	10	20
2	1	1	0	10	10
3	1	1	1	10	1
4	1	1	1	20	1
5	1	1	1	30	1
6	1	1	1	40	1
7	1	1	1	50	1
8	1	1	1	60	0.5

3.2. Results

In the following, we discuss the result of each reconstruction experiment individually. Afterwards, we make a direct comparison. We present the result for the individual Love wave FWI in figure 3.7, for the individual Rayleigh wave FWI in figure 3.8 and for the simultaneous joint FWI in figure 3.9. We use the same form of presentation for all three figures. In figure 3.10 we show a direct comparison of the final reconstructed models, where we focus on an area around the trench. Finally, we present horizontal and vertical profiles across the true and final models in figure 3.11.

3.2.1. Individual Love wave FWI

The individual Love wave FWI reconstructed the v_S model successfully by recovering the low-velocity trench sharply and in full extension. At the depth of 6.3 m, the water table is smoothly visible. The SH waves are not sensitive to the P-wave velocity, thus, the v_P model is not updated. The reconstruction of the ρ model is surprisingly well, especially if we consider that we use trace-wise normalized seismograms to calculate the objective function and the fact that the impact of density is mainly to the absolute wave amplitude as a function of offset. Nonetheless, the Love wave FWI reconstructed the trench in the ρ model satisfactorily in both its size and value, but the contour of the trench is less sharp compared to the v_S model. The water table is not sharply restored in the ρ model. As a result of a cross-talk by v_S , the reconstructed ρ model shows a footprint of the v_S anomaly. Especially the outline of the v_S anomaly is clearly visible within the ρ model, since the density values are decreased there (see figure 3.11). The anomaly in the ρ model is not visible in the reconstructed v_S model, which indicates that the individual Love wave FWI could restore the v_S model with less ambiguity than the ρ model, due to a higher sensitivity of the Love wave to v_S than to ρ .

The fit of the final synthetic seismograms (red) to the the observed seismograms (black) is very well at all offsets and times. A remaining residual between the synthetic and observed seismograms is only hardly visible. The decrease in the objective function is smooth and in total the inversion reduced the misfit by four orders of magnitude, as shown in figure 3.6. At each new workflow stage the algorithm increases the frequency content and the objective function increases as well, due to the fact that more information is considered for the residual calculation.

To sum up, the individual Love wave FWI reconstructed the v_S and ρ model successfully. However, the ρ model suffers from a footprint of the trench located in the v_S model.

3.2.2. Individual Rayleigh wave FWI

The individual Rayleigh wave FWI revealed the trench in the v_S model accurately, where both its velocity value and shape are correct. In a depth of 6.3 m, the inversion imaged the water table sharply. Thereby it is likely that the Rayleigh wave FWI benefits from the sharp water table included in the initial v_P model (see figure 3.2). Nevertheless, the v_S model suffers from small-scale artifacts that are present inside the trench and especially to the right side of it. The reconstruction of the trench in the v_P model is satisfactorily.

The contour is not clearly visible and vertically orientated artifacts are observed inside the trench. These artifacts could be caused due to wrong P-wave velocities in combination with receiver related artifacts, since they occur directly underneath the receiver positions at areas with wrong P-wave velocities and not where the initial model contains the true velocity. We observed similar artifacts in the case study on the influence of an inaccurate initial v_P model, which we will present in section 3.3. The artifacts within the v_P trench correspond to the artifacts in the v_S model, where the v_S model suffers from a cross-talk by v_P . This cross-talk is especially visible in figure 3.11. The v_P model itself suffers from a slight footprint of the v_S trench. In general, we expect the resolution in the v_P model to be lower compared to the v_S model, due to the longer wavelengths of the P-waves than of the S-waves. The reconstructed ρ model matches the true model satisfactorily. The inversion recovered the shape of the trench in the ρ model sufficiently, but reproduced the density value slightly higher than its actual value. Moreover, small-scale artifacts are present in the ρ model at the position of the v_S and v_P trench, most likely caused by a cross-talk. The water table is visible as a sharp contrast, where the Rayleigh wave FWI could again benefit from the initial v_P model.

The fit of the synthetic seismograms to the pseudo-observed seismograms is very well, since the residuals are nearly vanished. In figure 3.6 we show the evolution of the objective function. The inversion decreased the misfit smoothly in each frequency stage and reduced the objective function in total by four orders of magnitude.

To sum up, the individual Rayleigh wave FWI reconstructed the trench in the v_S and the ρ model successfully. However, the inversion did not recover the trench in the v_P model accurately enough. The ρ model as well as the v_P model suffer from a cross-talk by v_S , while the v_S model shows a cross-talk by v_P .

3.2.3. Simultaneous joint FWI

The simultaneous joint FWI reconstructed the trench in the v_S model very well in terms of shape and velocity values. The inversion imaged the water level sharply, where it could benefit from the sharp contrast included in the initial v_P model. There are no cross-talk effects of the v_P or the ρ model visible in the v_S model. The reconstruction of the trench in the v_P model is acceptable, although the outline of the trench is not as sharp as in the case of the v_S model. The final v_P model as well as the ρ model suffer from a light footprint of the v_S trench. Besides this cross-talk effect, the inversion recovered the ρ model successfully in its shape and value. The reconstructed trench is sharp and the density values matches the true value. The water table is sharply visible in the ρ model, where a positive influence from the initial v_P model is most likely.

The joint FWI fitted the P-SV seismograms as well as the SH seismograms to the pseudo-observed seismograms without significant residuals. The joint objective function could be reduced by four orders of magnitude, as shown in figure 3.6. Compared to the individual wave type inversions the joint FWI was able to reduce the individual objective functions of the Love and the Rayleigh wave inversion even further. More precisely, in the last iteration the absolute misfit value of the P-SV and the SH waves is lower in the case of the joint FWI than in the case of the individual wave type inversions. However, since we calculated the

3. Synthetic FWI experiments

joint objective function as a sum of both individual misfits, the total joint objective function shown in figure 3.6 is higher than both individual ones, since it is the sum of both.

To summarize, the simultaneous joint FWI of Rayleigh and Love wave reconstructed successfully the trench in all three elastic parameters. The joint FWI reduced the individual objective functions even further than the individual wave type inversion. The v_p model as well as the ρ model suffer from a light footprint of the v_s trench.

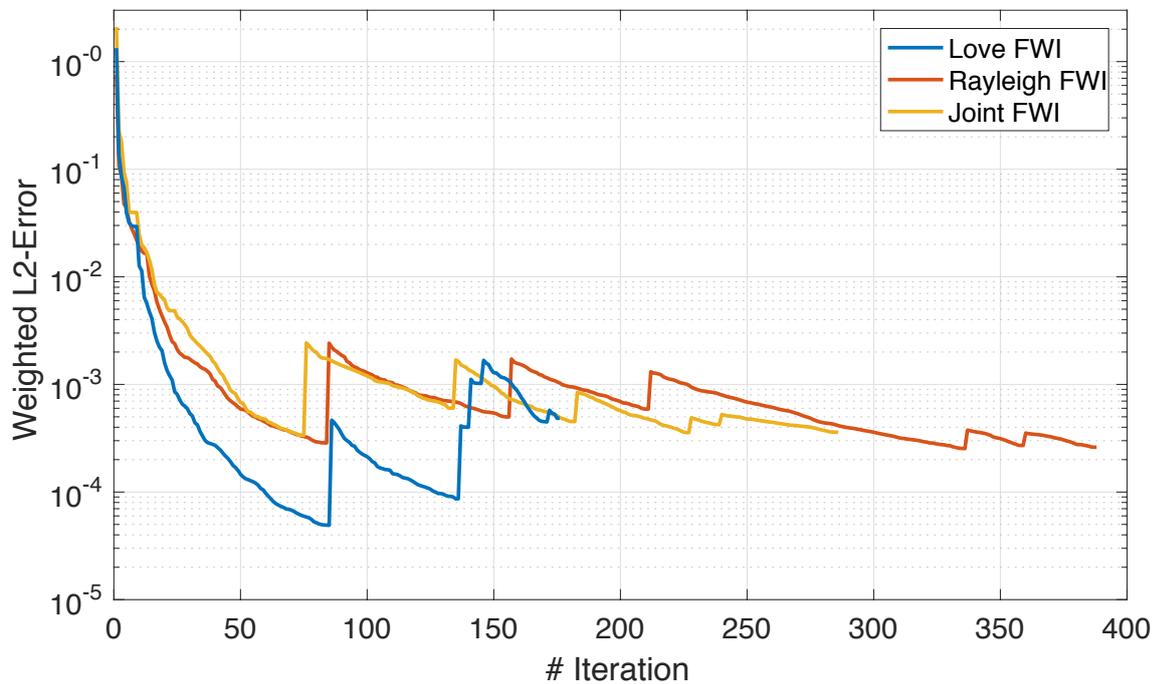


Figure 3.6.: Evolution of the L_2 -error during the synthetic Love wave FWI, Rayleigh wave FWI and simultaneous joint FWI over the iterations. The L_2 -error is weighted with the energy of the observed data. The sharp misfit jumps correspond to an increase of the low-pass corner frequency LP (see workflow table 3.1).

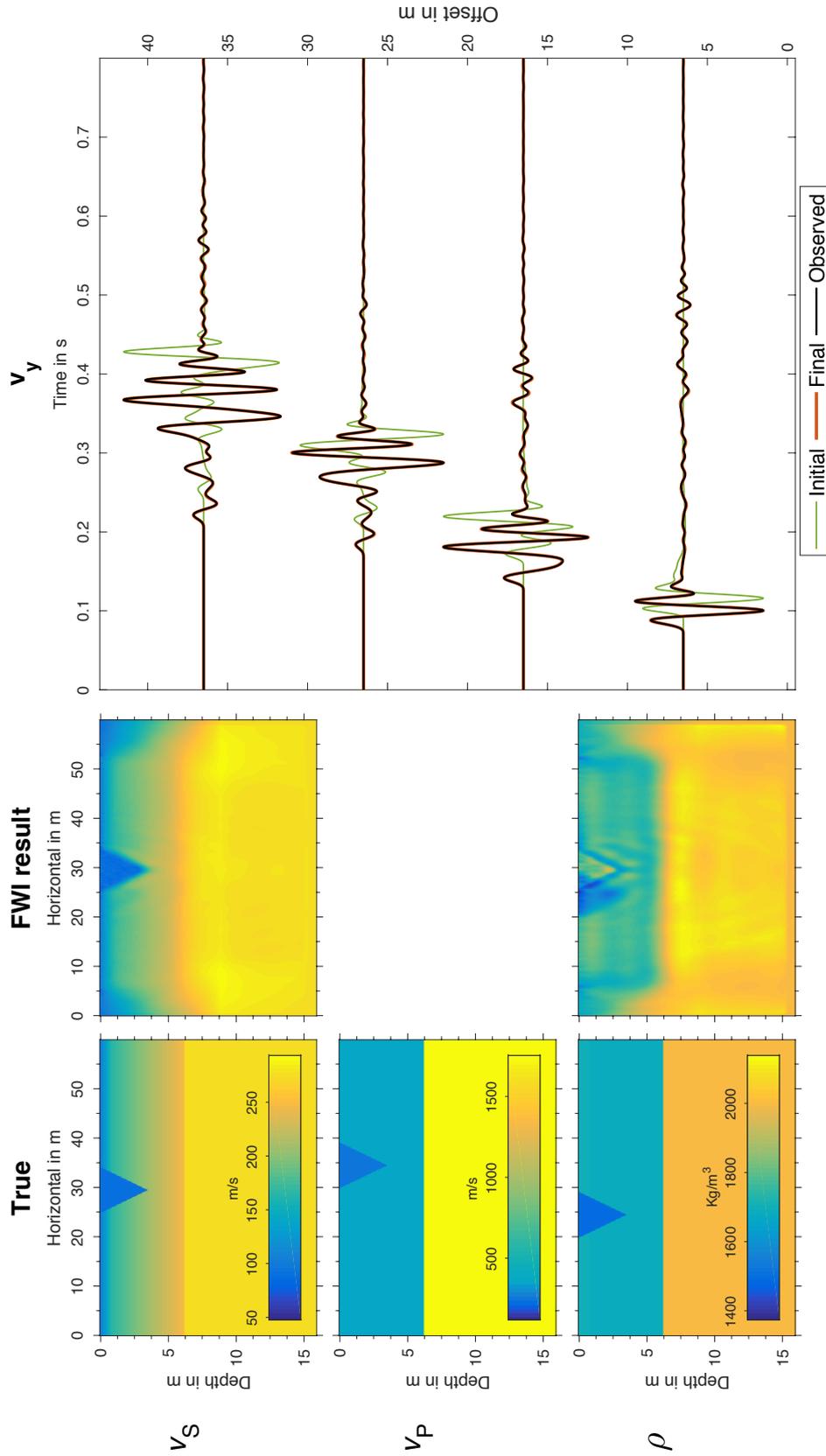


Figure 3.7.: Reconstruction result of **individual Love wave FWI**. True subsurface model besides final reconstructed models (left). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Seismograms at four different offsets (right). For each offset a seismogram for the true subsurface model (black), for the initial model (green) and for the final reconstructed model (red) is shown. The corresponding source is located at the profile coordinate of 6.5 m. The seismograms are band-pass filtered between 4 and 60 Hz, representing the frequency range used in the FWI.

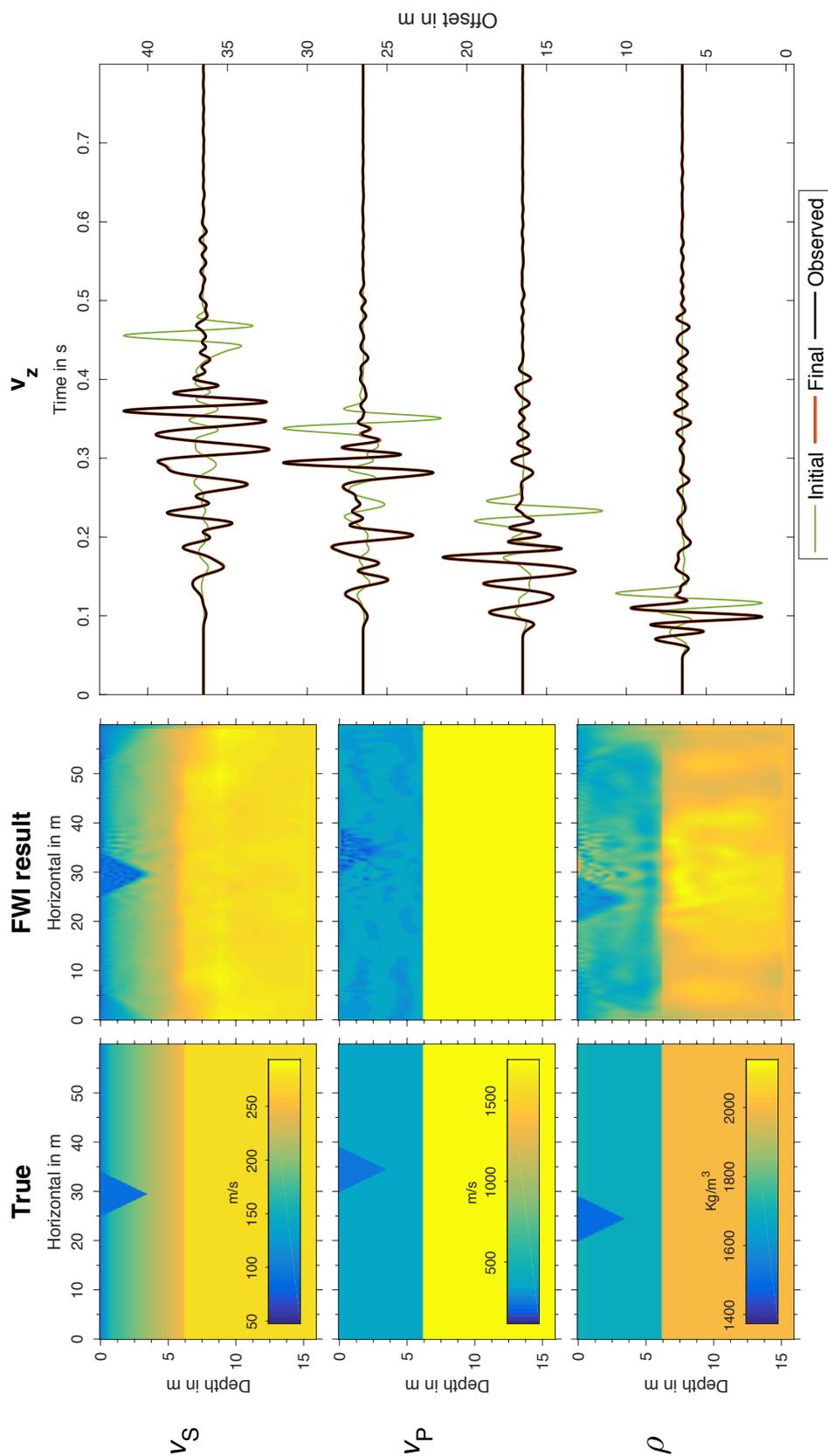


Figure 3.8.: Reconstruction result of **individual Rayleigh wave FWI**. True subsurface model besides final reconstructed models (left). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Seismograms at four different offsets (right). For each offset a seismogram for the true subsurface model (black), for the initial model (green) and for the final reconstructed model (red) is shown. The corresponding source is located at the profile coordinate of 6.5 m. The seismograms are band-pass filtered between 4 and 60 Hz, representing the frequency range used in the FWI.

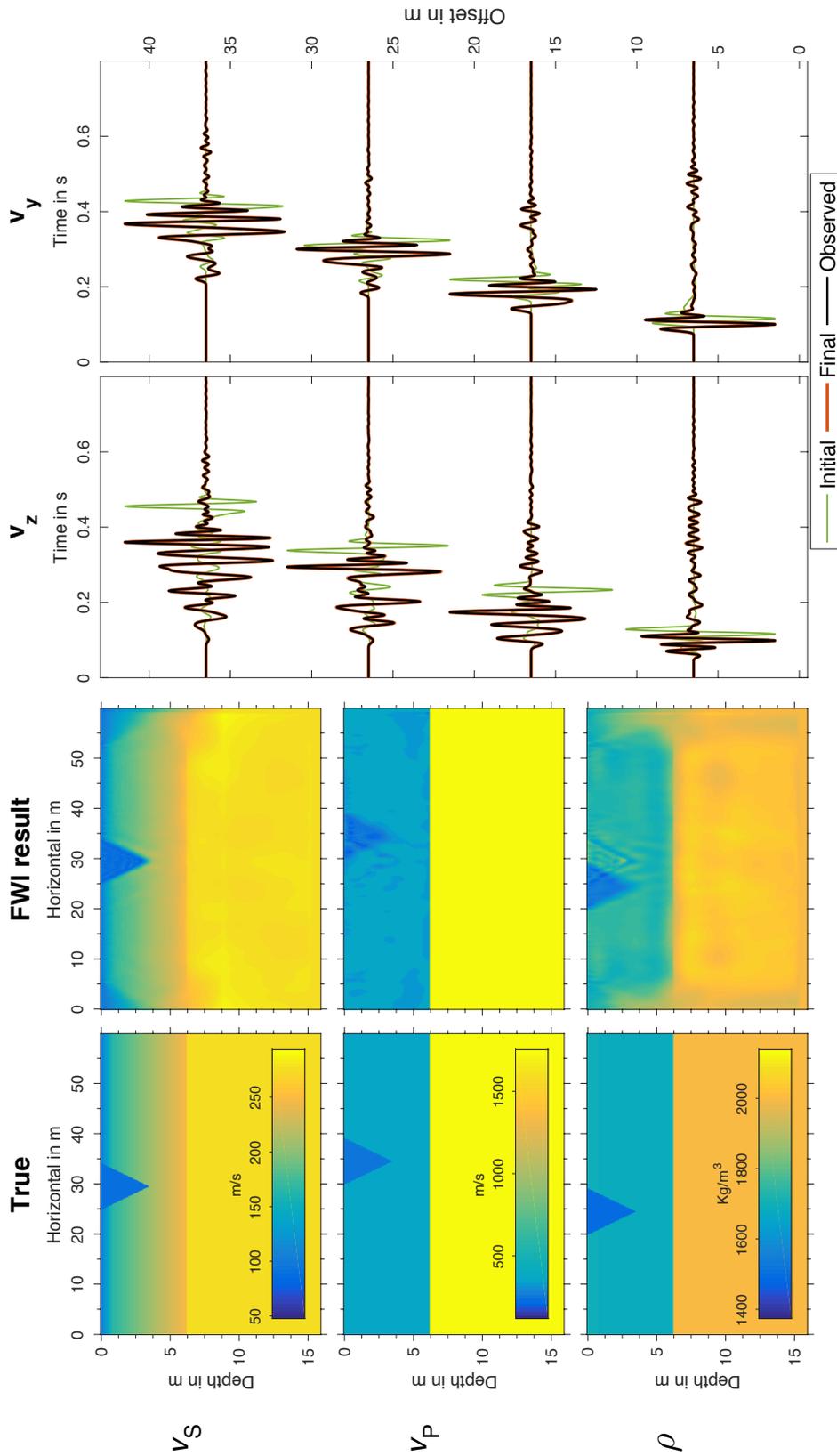


Figure 3.9.: Reconstruction result of **simultaneous joint FWI** of Love and Rayleigh waves. True subsurface model besides final reconstructed models (left). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Seismograms at four different offsets (right). For each offset a seismogram for the true subsurface model (black), for the initial model (green) and for the final reconstructed model (red) is shown. The corresponding source is located at the profile coordinate of 6.5 m. The seismograms are band-pass filtered between 4 and 60 Hz, representing the frequency range used in the FWI.

3.2.4. Comparison

After the individual interpretation of the inversion results we compare them now directly. Figure 3.10 shows the final models column-wise side by side. We thereby focus on a sub volume of the full model representing the area around the trench. Moreover, figure 3.11 shows the horizontal and vertical profiles, whose positions are indicated by figure 3.10.

S-wave velocity v_S :

Both the individual wave type inversions as well as the simultaneous joint inversion reconstructed the trench in the v_S model accurately in its shape and value. However, the individual Rayleigh wave FWI suffers from small-scale artifacts inside the trench and especially to the right of the trench, where a cross-talk by v_P caused the latter artifacts. This cross-talk is not present in the result of the joint FWI. The individual Love wave FWI does not show any artifacts in the v_S model as well. Moreover, all three inversions recovered the water table. The individual Rayleigh wave FWI and the joint FWI could benefit from the sharp water table that is included in the initial v_P model.

P-wave velocity v_P :

The individual Rayleigh wave FWI and the joint FWI reconstructed the trench in the v_P model with lower resolution compared to the other parameter classes, even though we used the true background model as the initial model. The individual Rayleigh wave FWI reconstructed the contour of the trench blurred and suffers from strong artifacts inside the trench. The joint FWI recovered the trench without artifacts in a smoothed kind. However, since we did not apply a joint approach to v_P , the improvement of the v_P model in the case of the joint FWI is most likely a result of the joint regularization of v_S and ρ . Nevertheless, both the individual Rayleigh and the joint FWI suffer from a footprint of the v_S trench. The individual Love wave FWI is not sensitive to P-wave velocities and therefore in this case we did not update the v_P model in this case.

Density ρ :

The reconstructed ρ model of all three inversions suffers from a footprint of the v_S trench. Nevertheless, all of them reconstructed the trench in an acceptable level. This is surprising, since we used trace-wise normalized seismograms for the calculation of the objective function and the fact that the influence of density is mainly to the absolute wave amplitude as a function of offset. The contour of the trench in the joint FWI result is sharper than for both individual wave type inversions. Additionally, the result of the joint FWI matches the true density value best and shows significant less cross-talk by v_S and v_P as the results of the other two inversions. The final ρ model of the individual Love wave FWI suffers from a strong footprint of the v_S trench and the individual Rayleigh wave FWI shows small-scale artifacts in the ρ model at the positions of the trench in v_S and v_P . All inversions recovered the water level successful. The individual Rayleigh wave FWI and the joint FWI could benefit thereby from the initial v_P model.

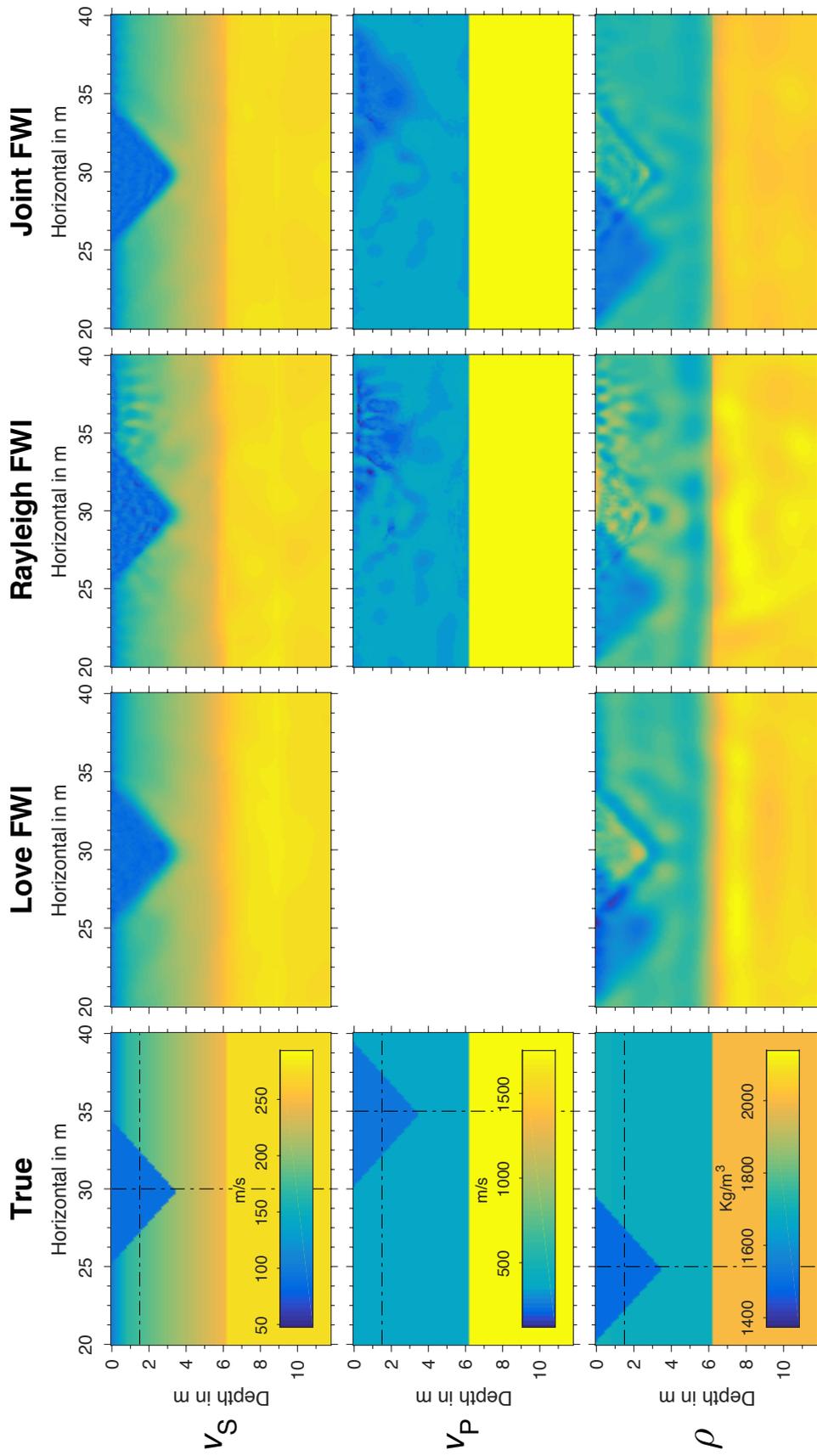


Figure 3.10.: Comparison of the true subsurface model (first column) with the reconstructed model of individual Love wave FWI (second column), individual Rayleigh wave FWI (third column) and simultaneous joint FWI (fourth column). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Horizontal and vertical profiles across all models are shown in figure 3.11.

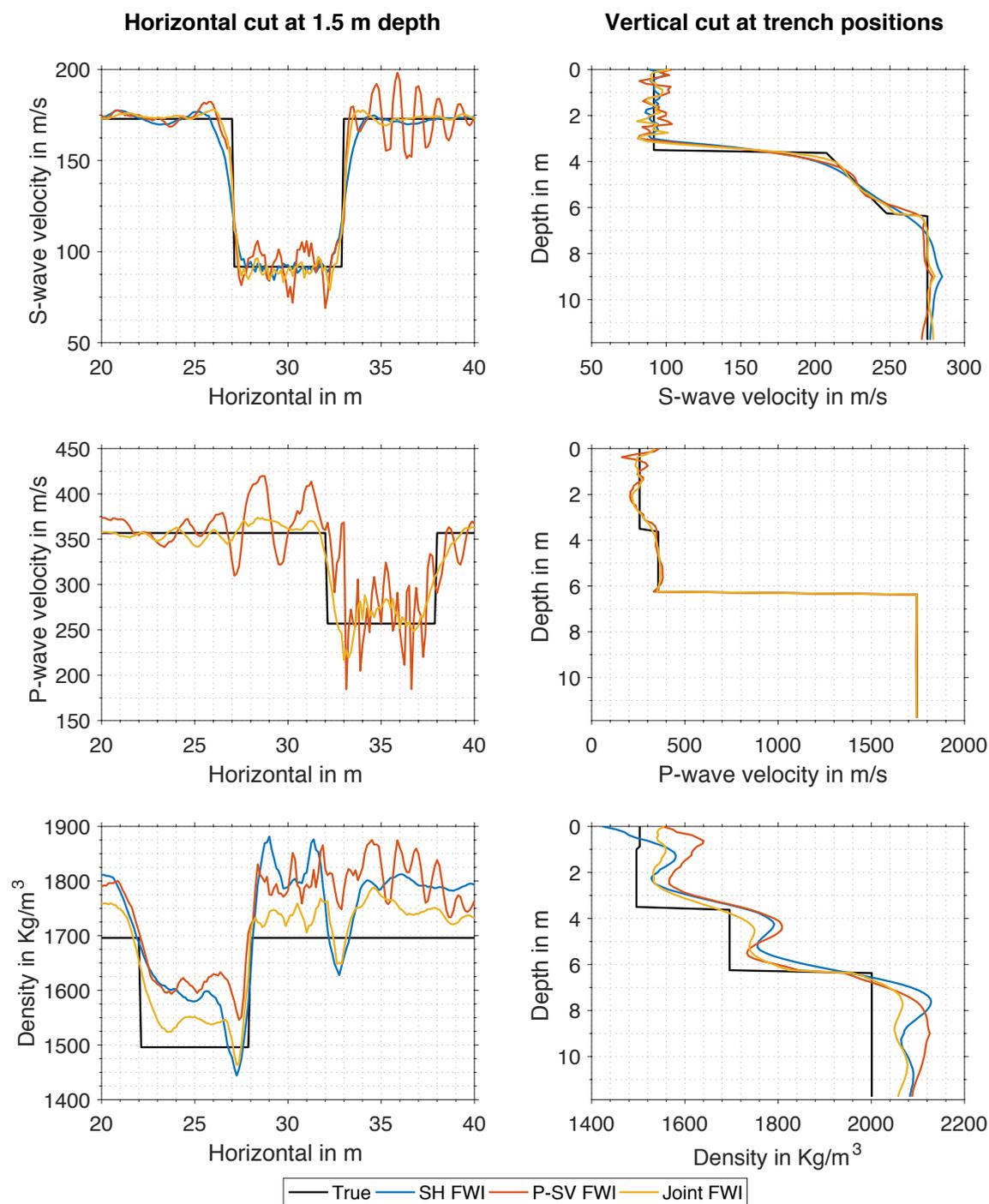


Figure 3.11.: Horizontal profiles at 1.5 m depth (left) and vertical profiles of the final models at the location of the trench positions (right), as indicated in figure 3.10. The elastic parameters v_S , v_P and ρ are shown row-wise from top to bottom, respectively.

3.3. Case study: Influence of the initial P-wave velocity

In this case study, we investigate the influence of the initial P-wave velocity model. In the synthetic experiments presented above we use the true v_p background model as initial model, because we assumed that a P-wave travel time analysis could provide an initial model of similar quality. In this test we do not follow this assumption and instead we use a simple gradient model as initial v_p model.

Problem and previous work:

In a previous study, we found that the initial v_p model highly influences the individual Rayleigh wave FWI and the simultaneous joint FWI (partly published in Wittkamp and Bohlen (2016)). In that study a similar configuration as in this work was used. However, only frequencies up to 20 Hz were considered and a different acquisition geometry was used. That study revealed that an unsuitable initial v_p model could lead to strong vertically orientated artifacts underneath the source positions in both the v_s model and the ρ model. Similar artifacts were also observed by Groos (2013). The individual Love wave FWI did not suffer from such artifacts, since it is not affected by v_p . The artifacts were most likely caused by the high amplitudes and the focused radiation pattern of the Rayleigh waves (Kähler and Meissner, 1983). Wrong P-wave velocities in the vicinity of the sources enhanced the artifacts in the gradients of v_s and ρ even more. The usage of the true v_p as initial model mitigated these artifacts significantly. We observed the influence of the initial v_p model to the overall convergence behavior to be quite severe. Moreover, we found that the v_p model highly influences the propagation of the Rayleigh waves, however, the ability to reconstruct the v_p model itself by Rayleigh waves is low. Nevertheless, as the full-waveforms contain the P-wave onsets and the fact the the P-wave is mainly influenced by the v_p model, the FWI theoretically could reconstruct the v_p model. However, the Rayleigh wave has a significant higher amplitude as the P-wave onsets, thus, the Rayleigh wave is dominating the objective function and the reconstruction of the v_p model depends mainly on the Rayleigh wave. To overcome this issue, we would have to consider the P-wave onsets separately which however is beyond the scope of this work.

Test setting:

In this case study we use a gradient model as initial v_p model. The acquisition geometry, the FWI setting, the true model and the initial model for the other two elastic parameters is identical to the synthetic experiments presented above. This allows us to access the influence of an inaccurate initial v_p model to the reconstruction ability of the individual Rayleigh wave FWI and the joint FWI based on the actual test setting. Since we use the same acquisition geometry in the field data application, this case study could help to interpret the reliability and potential artifacts in the field data FWI.

In figure 3.12 we show the used inaccurate initial v_p model. This initial model matches the true model at the surface and again at a depth of 9 m and from there on the true and initial model are identical. Between the surface and a depth of 9 m we used a constant gradient of $155 \frac{\text{m}}{\text{s}\cdot\text{m}}$. We present the result of this case study in figure 3.13, where we focus on an area around the trench positions.

Results:

The inaccurate initial v_p model does not influence the **individual Love wave FWI**, as the P-wave velocity does not affect SH waves at all.

In contrast, the result of the **individual Rayleigh wave FWI** exhibits a significant influence. The reconstructed v_s model suffers from circular low-velocity artifacts around the receiver positions. These artifacts are likely to be transferable to the source artifacts observed in our previous study, since the receivers will act as sources when the residuals are back propagated from the receiver positions. Taking into account this result as well as our previous work, we conclude that the presence of source or receiver artifacts depends on the acquisition geometry, e.g. the spacing between sources or receivers. Nevertheless, the shape of the trench is visible in the reconstructed v_s model, but the velocity inside the trench is rough. The v_p model does not fit the true model and shows strong artifacts at the position of the trench. These artifacts are also foot-printed to the v_s model. The updated ρ model contains systematically too high density values and therefore we cannot evaluate this model.

The result for the **simultaneous joint FWI** is similar to the individual Rayleigh wave FWI, especially for the v_p and the ρ model. However, the joint FWI recovered the v_s model slightly better, since the receiver artifacts are less prominent and the velocity values within the trench are smoother.

Summary:

In this synthetic sub-experiment we revealed a significant influence of an inaccurate initial v_p model to the reconstruction ability of the individual Rayleigh wave FWI. As in our previous study we found that the v_p model highly influences the propagation of Rayleigh waves, however, the Rayleigh wave FWI could not reconstruct the v_p model from an inaccurate starting point. With such an inaccurate initial v_p model only a rough reconstruction of the v_s model is possible, while the reconstructed ρ model does not allow for evaluation. We also observed artifacts in the v_s model caused by the incorrect v_p model. Furthermore, we found that the simultaneous joint FWI is not able to improve the inversion result significantly, despite its combination with the Love wave FWI, which is not influenced by v_p .

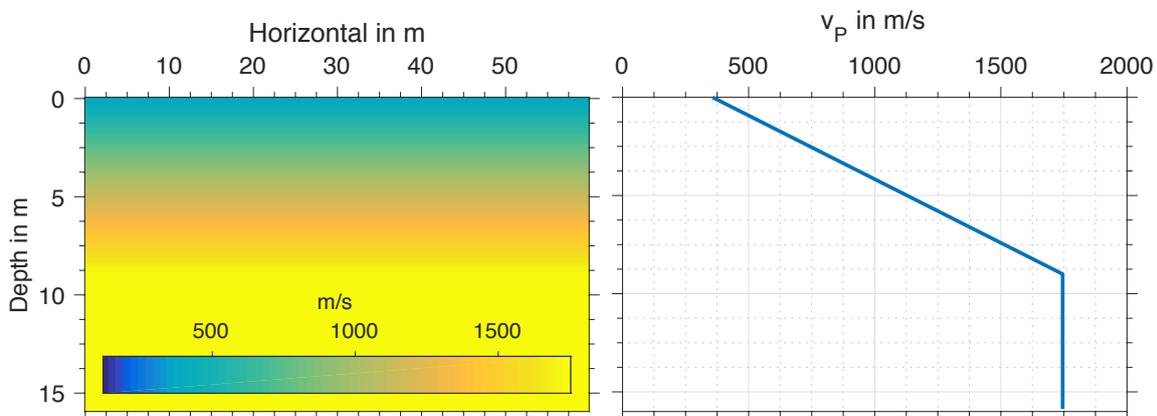


Figure 3.12.: Inaccurate initial v_p model used in the case study.

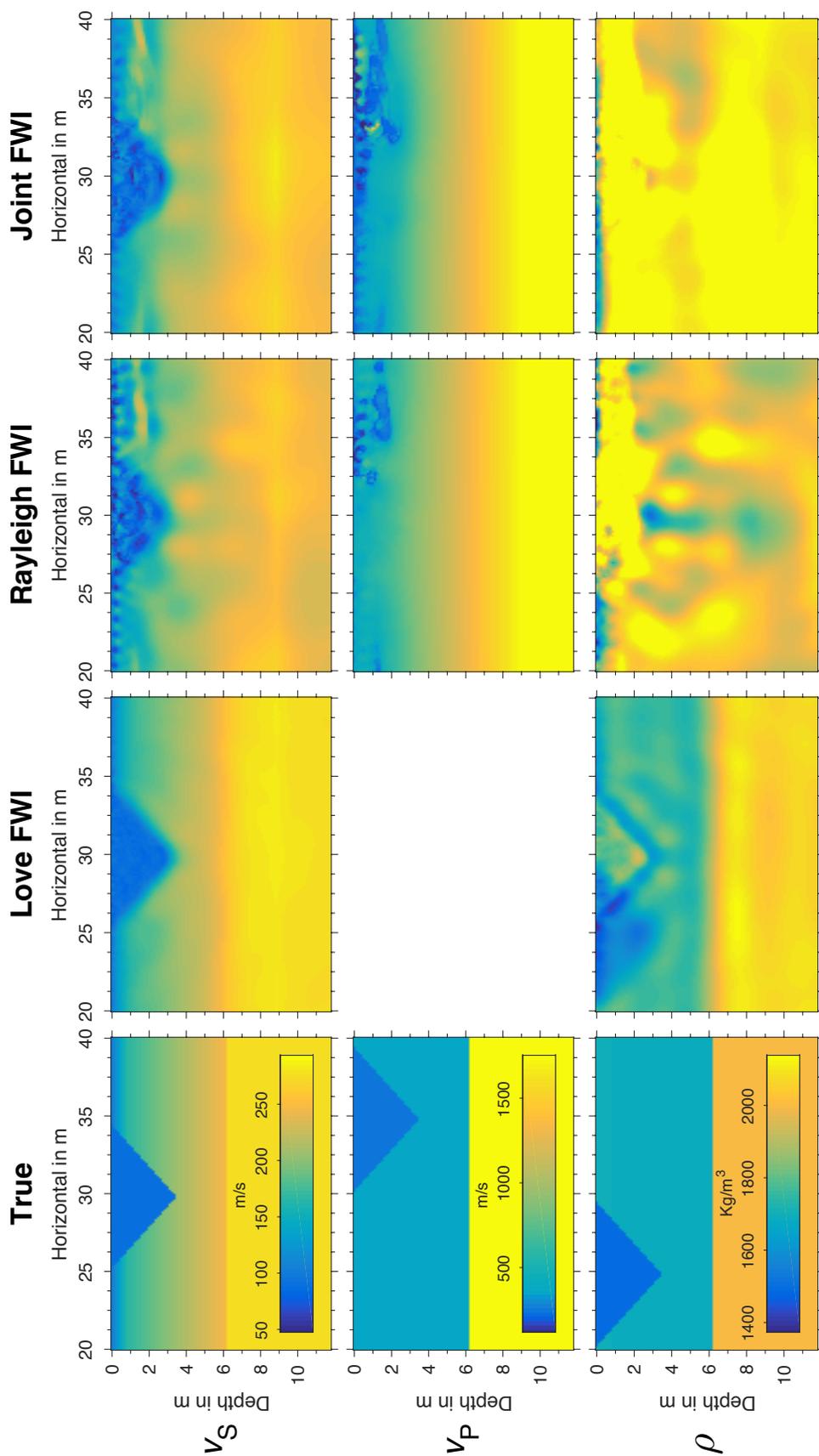


Figure 3.13.: Comparison of the true subsurface model (first column) with the reconstructed models of individual Love wave FWI (second column), individual Rayleigh wave FWI (third column) and simultaneous joint FWI (fourth column) for an inaccurate initial v_p model. The elastic parameter v_s , v_p and ρ are shown row-wise from top to bottom, respectively.

3.4. Summary

We performed synthetic reconstruction experiments to explore the properties of individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves. In these reconstruction tests we knew the true subsurface model, thus, we could study the reconstruction ability of the FWI by comparing the true model and the reconstructed models.

For these reconstruction tests, we assumed a predominantly depth dependent subsurface model that is superimposed by a shallow small-scale low-velocity trench. We thereby shifted the horizontal location of the trench in each parameter class to investigate cross-talk effects. For the initial model of v_S and ρ we used only a small a priori information, whereas we used the true background model as initial model for v_P .

In our synthetic experiment the individual Love wave FWI converged smoothly and reconstructed the subsurface models of v_S and ρ successfully, although the reconstructed ρ model suffered from a footprint of the v_S trench. The individual Rayleigh wave FWI recovered the v_S model and the ρ model successfully as well. The v_P model was rebuilt roughly and suffered from artifacts inside the trench. The ρ model as well as the v_P model suffered from a cross-talk by v_S . The v_S model showed a cross-talk by v_P . The simultaneous joint FWI reconstructed the trench in all three elastic parameters very well. The result of the joint FWI showed lesser trade-off effects compared to the results of the individual wave type inversions. The joint FWI decreased the objective functions of both wave types more than the individual wave type inversions. Altogether, the simultaneous joint FWI further improved the reconstruction of the subsurface.

To explore the influence of the initial v_P model, we carried out an additional case study. We found that the v_P model highly influences the propagation of the Rayleigh wave, however, the reconstruction ability of the Rayleigh wave FWI for the v_P model itself is low. Moreover, we observed that wrong P-wave velocities in the initial model lead to artifacts in the v_P model which are foot-printed to the v_S model. The simultaneous joint FWI did not provide a significant improvement in this case.

To sum up, we could successfully reconstruct the v_S model with all three inversions. In the case the initial v_P model is accurate, both individual wave type inversion performed similarly well and the joint FWI further improved the inversion result. In the absence of an accurate initial v_P model, the reconstructed v_S model of the individual Rayleigh wave FWI and the joint FWI is less accurate than in the case of an accurate initial v_P model. The Love wave FWI performed identical as in the case of an accurate initial v_P model, since Love waves are not influenced by the v_P model.

4. Field data application

In this chapter, we present the application of individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves to a near-surface field dataset. We recorded a field dataset on a glider airfield in Rheinstetten near Karlsruhe (Germany). Previous Rayleigh wave FWI studies took place on the same test site and proved the suitability of it for 2-D FWI. They propose a predominantly depth dependent subsurface that is superimposed by a shallow small-scale low-velocity trench. This trench is known as "Ettlinger Linie". It was originally excavated to serve as a line of defense and was refilled several decades ago. We assumed such a 2-D subsurface model in the synthetic experiments (chapter 3), where we verified the theoretical applicability of both the individual wave type inversions as well as the simultaneous joint FWI. In this experiment, we investigate the applicability of all three inversions to the recorded near-surface field dataset and evaluate their performance. This chapter is organized as follows: First, we introduce the test site, the test setting and the necessary preparatory work for FWI. Second, we present the inversion results of the three FWI studies: (a) Individual Love wave FWI, (b) individual Rayleigh wave FWI and (c) simultaneous joint FWI. We then perform a direct comparison of the inversion results. To draw conclusions on the quality of the obtained results, we compare them with the results from ground-penetrating radar measurements, which were acquired in the same area. Finally, we summarize the main results of the field data application.

4.1. Setting

In this section, we introduce the test site and the acquisition geometry. We then present the preparatory steps applied to the field dataset. Additionally, we describe the procedure to estimate an initial subsurface model and describe the chosen configuration of the inversion.

4.1.1. Test site

The location of the test site is on a glider airfield in Rheinstetten near Karlsruhe (Germany). Figure 4.1 shows an overall map of the area. This test site exhibits a planar surface and does not suffer from humanistic noise in the direct surrounding. The geological map by [Hüttner et al. \(1968\)](#) states that the subsurface consists of layered fluvial sediments of the late Pleistocene. Several shallow-seismic studies were carried out on this area, which provide further information on the subsurface. [Groos \(2013\)](#) and [Schaefer \(2014\)](#) performed a dispersion curve inversion and a FWI of Rayleigh waves on the north-west part of the airfield and propose a predominantly depth dependent 1-D subsurface. [Lüttschwager \(2014\)](#) investigated the north-east corner of the runway and discovered a shallow small-scale low-velocity anomaly (trench) that proceeds straight from the north-west to the south-east.



Figure 4.1.: Overview map of the test site on the glider airfield in Rheinstetten. The red line corresponds to the interpolated course of the "Ettlinger Linie" and the white line denotes the acquisition profile. Marker one (N 48°58'52.59", E 8°20'46.73") indicates the first receiver and marker two (N 48°58'52.12", E 8°20'44.57") the last receiver. The marker positions are accurate to several meters. Source: Google Earth (AeroWest, GeoBasis-DE/BKG).

Binnig (2015) confirmed this hypothesis by a 2-D Rayleigh wave FWI. His results suggest that this low-velocity trench locally superimposes the 1-D subsurface proposed by the previous studies. According to historic recordings this trench can be identified as the "Ettlinger Linie". It was originally excavated to serve as a line of defense and was refilled several decades ago (Lang, 1907). Outside the borders of the airfield the "Ettlinger Linie" is still uncovered and traceable, which allows to easily interpolate the course of the refilled trench within the airfield.

Such a subsurface structure suites well for 2-D FWI experiments. The low-velocity trench proceeds straight for about 70 m and superimposes the lateral homogeneous background subsurface locally. Hence, the assumption of a 2-D subsurface is valid in the case that the acquisition profile crosses the trench vertically, which is important since the 2-D FWI cannot account for signals, such as reflections, caused by anomalies located off the 2-D profile. Figure 4.1 shows an overview of the interpolated course of the trench (red) and the acquisition profile (white).

4.1.2. Acquisition geometry

To image the cross-section of the "Ettlinger Linie" with the 2-D FWI, we placed the acquisition profile to cross the interpolated course of the trench vertically, as illustrated in figure 4.1. We shifted the profile to contain the trench in its center, in order to obtain a high wave coverage within the low-velocity trench. The orientation of the profile is from north-east (marker one) to south-west (marker two).

For the seismic recording we used 48 three-component geophones with 4.5 Hz eigenfrequency of the type Geospace Technologies GSC-11D. We set the geophone spacing equidistantly to 1 m and adjusted the local orientation of the geophones to the profile, in order to ensure an accurate recording of the horizontal component. The total length of the receiver line was 47 m. In figure 4.1 marker one (N 48°58'52.59", E 8°20'46.73") indicates the first receiver position and marker two (N 48°58'52.12", E 8°20'44.57") indicates the last receiver position. For the P-SV dataset we recorded the vertical particle velocity and for the SH dataset we recorded the horizontal crossline component.

We set the spacing between the 24 sources to 2 m, where the first source was located between the first and the second receiver. All source positions were located within the receiver line. The source-receiver offset ranges from 0.5 m to 46.5 m. We used vertical hammer blows on a steel plate to excite the P-SV dataset and horizontal hammer blows in the crossline direction on a steel source rack to excite the SH dataset.

The distribution of the sources and receivers is illustrated in figure 4.2. Figure 4.7 (top left) shows the acquisition geometry as a whole. In the synthetic experiments we used an identical acquisition geometry.

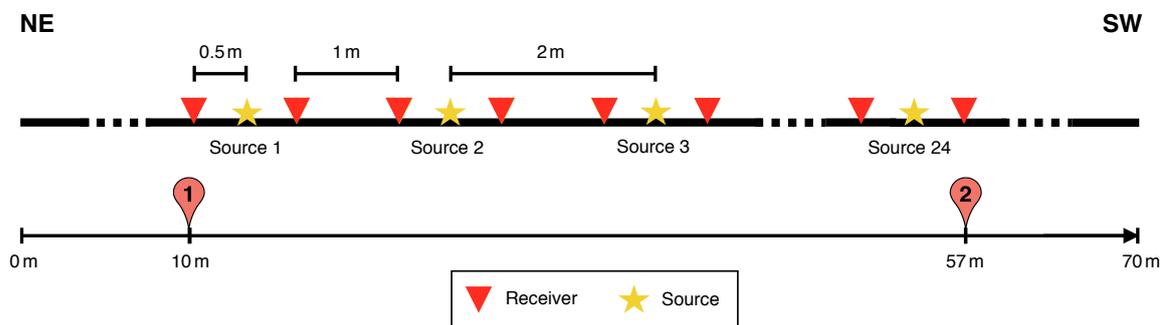


Figure 4.2.: Sketch of the acquisition geometry of the field data recording. The sketch is not to scale. Yellow stars indicate the source positions and red triangles indicate the receiver positions. The profile coordinates represent the whole model space. Marker one and two correspond to figure 4.1.

4.1.3. Observed data

The total recording time is 1.5 s and the temporal sampling is $2.5 \cdot 10^{-4}$ s. During the measurement we stacked the data with a fold of five to enhance the signal-to-noise ratio. However, the raw field dataset is not appropriate for a direct application of the 2-D FWI, thus, we perform a few preparatory steps. First of all, we shorten the data to 0.5 s, due to the absence of any significant energy at later recording times. We then up sample the data by a spline interpolation to a sampling of $1.4 \cdot 10^{-5}$ s, in order to satisfy the stability criterion of the finite-difference forward solver. To suppress signals before the actual wave onset, we apply a muting at the beginning of each trace. For the sake of avoiding non-causal effects in the inverted source time function (see section 4.1.5), we delay the whole dataset by 0.02 s, which results in a total time length of 0.52 s. Furthermore, we perform a 3-D to 2-D transformation, which is necessary, due to the fact that the recorded wave fields are excited by hammer blows that act like point-sources. However, the 2-D forward solver assumes line-sources. We transform the field data to an equivalent 2-D line-source by a trace-wise convolution with $\sqrt{t^{-1}}$ followed by a multiplication with $r \cdot \sqrt{2} \cdot \sqrt{t^{-1}}$, where t denotes the travel time and r the offset. This transformation is introduced as direct-wave transformation by [Forbriger et al. \(2014\)](#).

Exemplarily shot gathers of the preprocessed dataset are shown in figure 4.3 (first source) and 4.4 (last source). We normalized the seismograms trace-wise and applied a band-pass filter between 4 Hz and 130 Hz, which corresponds to the same way as we use them in the FWI. In figure 4.5 we show the mean amplitude-frequency spectra of the field data recordings. The main frequency content is located between 10 Hz and 100 Hz. There is no significant difference in the frequency content of both wave types.

The P-SV seismograms, v_z , are dominated by the Rayleigh wave. The Rayleigh wave is visible in the fundamental mode as well as in several higher modes. The direct and the refracted P-wave can be identified as well. The P-waves have much smaller amplitude compared to the Rayleigh wave. The change of the phase-velocities along the offset indicates the presence of a 2-D structure. In figure 4.12(a) we show the phase-slowness-frequency spectrum for the first shot, where we normalized the spectrum to the maximum amplitude of each frequency. The dispersion curve of the fundamental Rayleigh mode is dominant for frequencies up to approximately 40 Hz, where it reaches a phase-slowness of approximately 7 s/km. Above 25 Hz several higher modes coincide with the fundamental mode. The phase-slowness of the higher modes varies between approximately 1.8 to 5 km/s. The dispersion curves of the higher modes are not sharp enough to identify single modes.

The SH seismograms, v_y , are dominated by the Love wave, which is present in the fundamental mode. Compared to the Love wave the direct as well as the refracted S-wave have smaller amplitude and are only slightly visible. Figure 4.12(b) shows the phase-slowness-frequency spectrum for the first shot. The fundamental Love mode dominates the spectrum up to approximately 90 Hz, where it reaches a phase-slowness of about 7.5 s/km. A higher mode is present between approximately 30 to 45 Hz, whose phase-slowness vary between 3 to 4 s/km. The variation of the phase-slowness indicates again a 2-D structure within the survey area.

After the preprocessing steps that we carried out the P-SV as well as the SH field dataset allow for the full-waveform inversion.

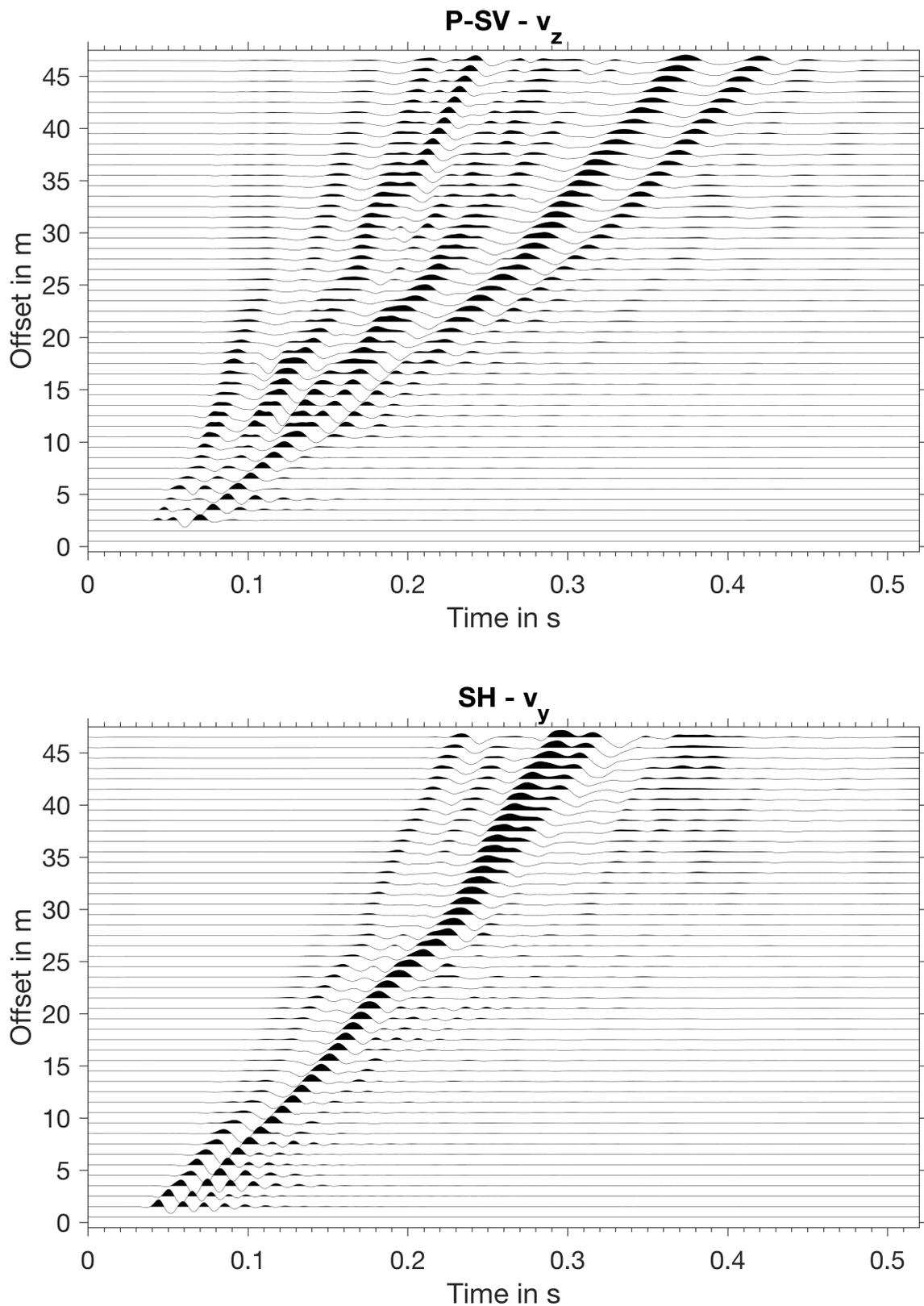


Figure 4.3.: Trace-wise normalized shot-gathers of the preprocessed field dataset for the first source located at the profile coordinate of 10.5 m. The seismograms are band-pass filtered between 4 and 130 Hz, representing the frequency range used in the FWI. Traces near the source were overdriven and thus muted.

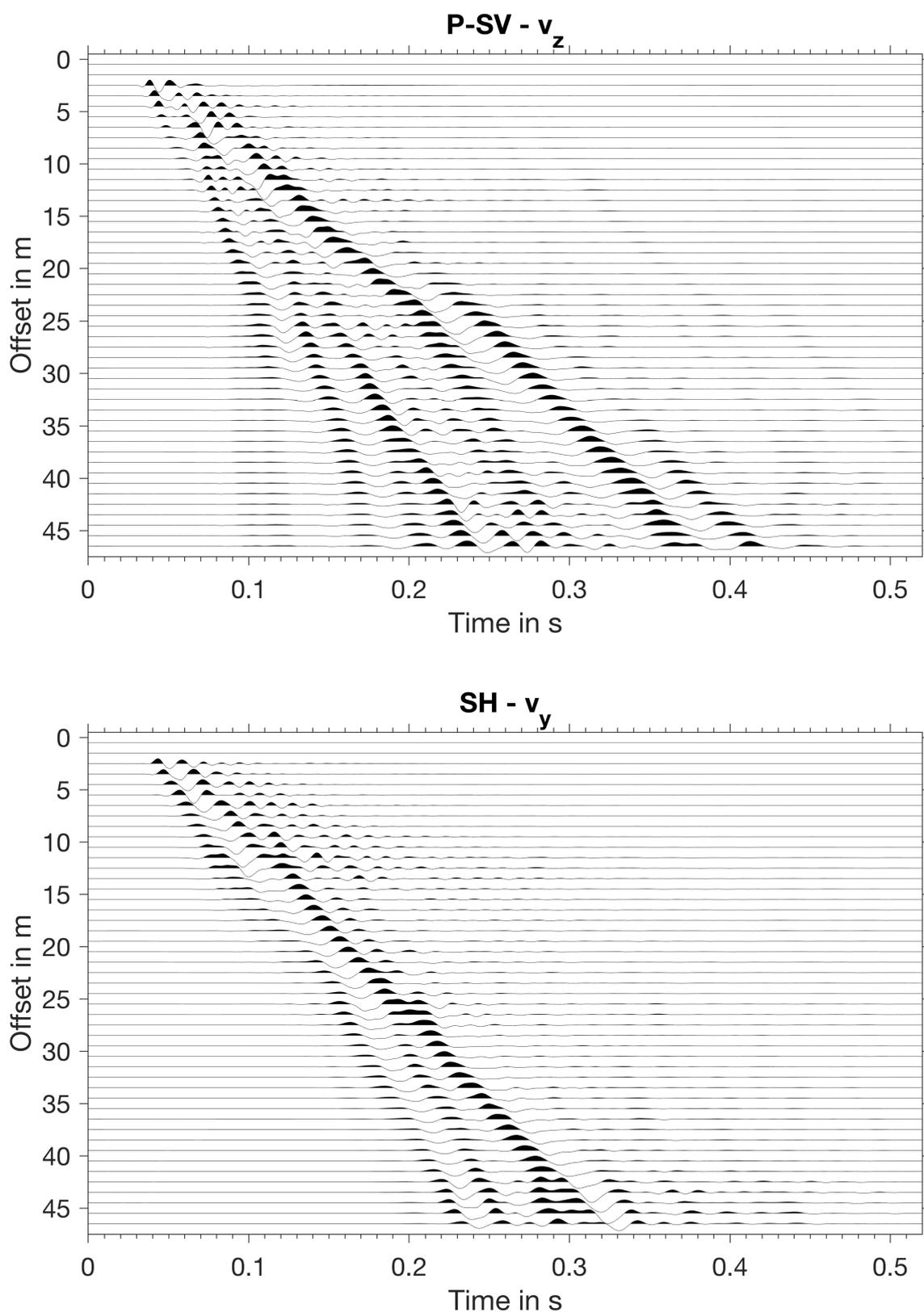


Figure 4.4.: Trace-wise normalized shot-gathers of the preprocessed field dataset for the last source located at the profile coordinate of 56.5 m. The seismograms are band-pass filtered between 4 and 130 Hz, representing the frequency range used in the FWI. Traces near the source were overdriven and thus muted.

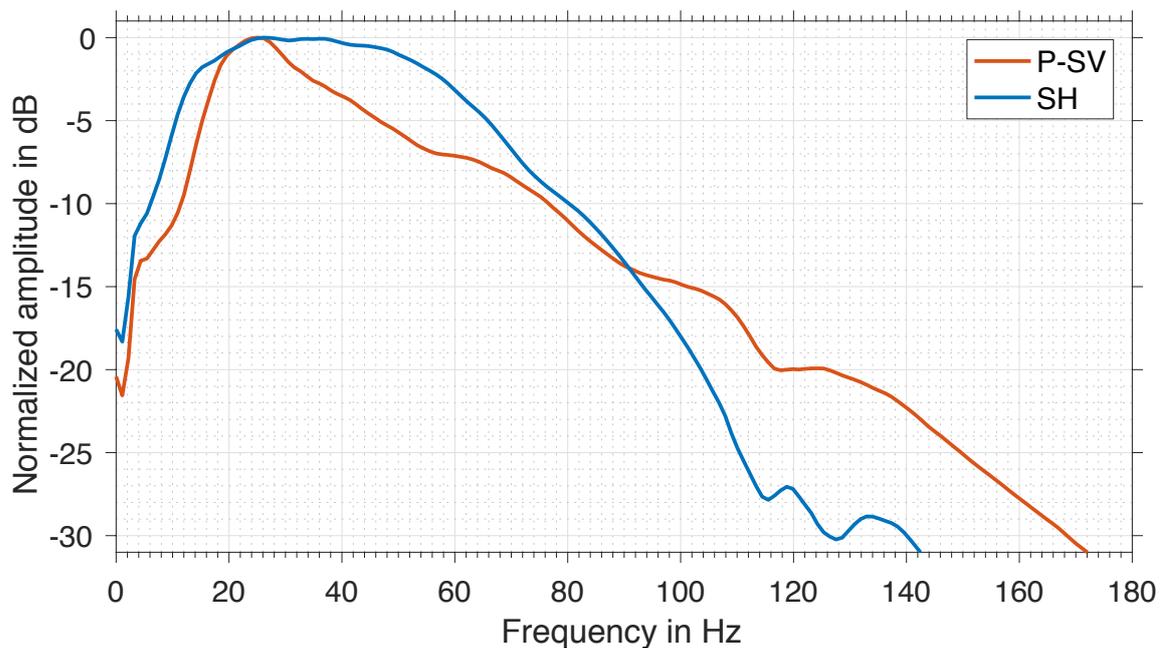


Figure 4.5.: Amplitude-frequency spectra of the field dataset. The spectrum of the P-SV waves is shown in red and of the SH waves in blue. The shown spectra are the average of the spectra of the normalized traces.

4.1.4. Initial model

Since we use a local optimization method for the full-waveform inversion, we have to assume an initial model for the parameters v_S , v_P and ρ . This model has to predict the main wave phases well enough to allow local convergence of the inversion. Moreover, we have to derive an initial model for the quality factors Q_S and Q_P , in order to describe the attenuation properties of the subsurface. The attenuation properties will not be updated during the inversion, instead we will use it as passive model parameters. In the following we assume the initial models to vary only with depth, since we expect the background model to be predominantly depth dependent.

First, we perform a P-wave travel time analysis to obtain an initial model for v_P . We therefore pick and evaluate the onsets of the direct and the refracted P-wave. The obtained model consists of two layers, where the interface lies in a depth of 6.1 m. For the upper layer we calculate a P-wave velocity of 335 m/s and for the lower half space we calculate a velocity of 2284 m/s. We assume that the sharp contrast in a depth of 6.1 m corresponds to the water table, as opposed to Groos (2013), who observed the water table in a depth of 6.8 m. This is in good agreement with our result, since it was raining the days before the field measurement took place, which could lead to a higher ground water table.

To obtain an initial model for ρ from the v_P model we use the empirical Gardner's relation:

$$\rho = 0.31 \cdot v_P^{0.25}, \quad (4.1)$$

which assumes that v_P is given in units of m/s for a resulting ρ value in units of g/cm^3 (Gardner et al., 1974). The obtained ρ model has a density of 1.325 g/cm^3 in the upper layer and 2.142 g/cm^3 in the lower half space.

We obtain a first initial v_S model from the work of Groos (2013), who used a gradient model which starts at 100 m/s at the surface and ends at 369 m/s in a depth of 9 m followed by a homogenous half space. However, neither the Love wave nor the Rayleigh wave FWI allow for convergence with this model. To overcome this issue we perform a local grid-search by vary the S-wave velocity at the surface and in a depth of 9 m. Finally, we were able to predict all main phases of the Rayleigh and Love waves for a S-wave velocity of 140 m/s at the surface and 340 m/s in a depth of 9 m.

Furthermore, we estimate an attenuation model. This is an important step, since inelastic damping has a significant influence to shallow-seismic recordings (Groos et al., 2014). We use two assumptions to obtain the Q -values from the field dataset: (1) We assume that the Q -values for P-waves and S-waves are identical and (2) that a constant Q -value is sufficient for the whole model space. Then, we can approximate a Q -value by calculating and comparing the misfit between synthetic and observed data for different Q -values. We calculate the misfit based on the presented initial model of v_S , v_P and ρ in the frequency range between 4 Hz and 100 Hz, which contains the main frequency content. The result of the local grid-search is shown in figure 4.6 for Q -values ranging from 5 to 35. For the P-SV dataset we obtain a minimum misfit for $Q = 15$, whereas the SH dataset reveals a minimum misfit for $Q = 10$. For the simultaneous joint FWI of both wave types we use the same parameter model, thus, we choose only one Q -value for the P-SV and the SH waves. Groos et al. (2014) suggested to use rather a higher Q -value than a smaller one, since the source time inversion (see section 4.1.5) can partly compensate a to high Q -value. Taking this consideration, we set the Q -value to 15. We construct the attenuation model by a GSLS with three relaxation mechanisms. The calculated relaxation frequencies are $f_1 = 0.2978$ Hz, $f_2 = 6.7325$ Hz and $f_3 = 84.6014$ Hz and the τ -value is 0.1576, where we use a reference frequency of 40 Hz.

The final initial model for the three elastic parameter classes v_S , v_P and ρ is presented in figure 4.7. This initial model allows to predict all main phases included in the P-SV and SH field dataset (left column of figures 4.9 and 4.10). Thus, the main requirement for the application of a local optimization method in the FWI is fulfilled. Based on this initial model we calculated phase-slowness-frequency spectra for the P-SV and SH waves, which is shown are figure 4.12 (second column). The fundamental modes of the surface waves dominate the spectra of both wave types, while higher modes are not present.

4.1.5. FWI setup

For the field data application we again use the FWI software IFOS. This software package uses a finite-difference method for the forward modelling and the adjoint state method for the calculation of the gradients. We use the quasi-Newton L-BFGS method and subsequently a conjugate gradient method to calculate the actual model update.

4.1.5.1. Objective function

As objective function we use the weighted L_2 -error (equation 2.93) between the normalized synthetic and normalized observed seismograms. We apply the normalization trace-wise to the seismograms, in order to mitigate uncertainties and differences in the ground coupling

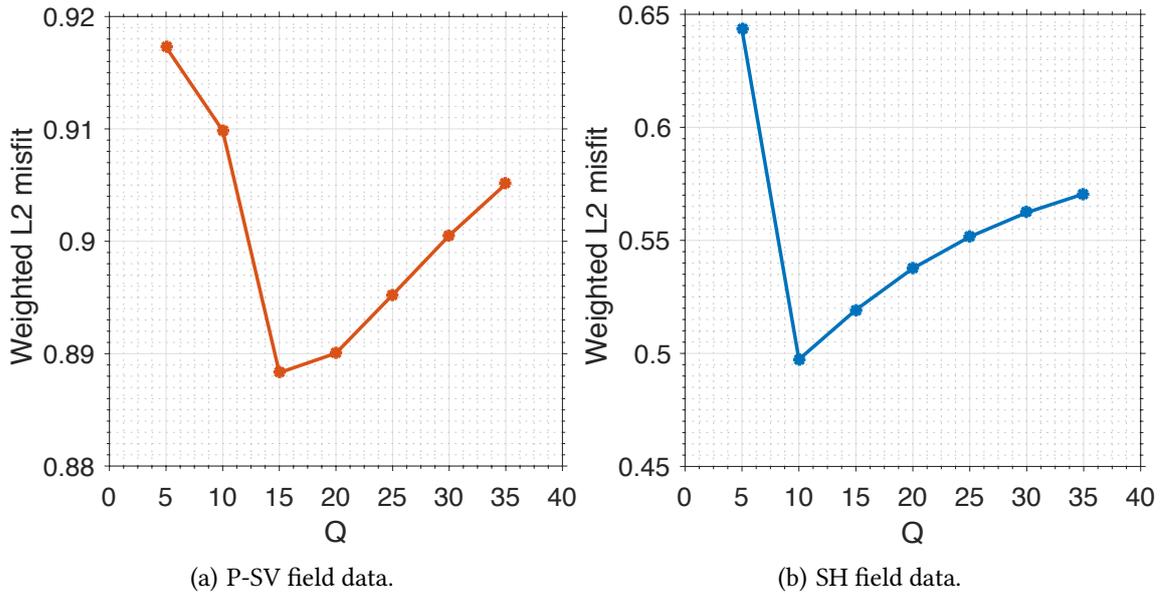


Figure 4.6.: Q -value estimation for the field dataset. The misfit between the synthetic data and the observed data is calculated in the frequency range of 4–100 Hz, which contains the main frequency content (see figure 4.5).

of individual geophones (Maurer et al., 2012). In the case of the P-SV FWI we calculate the objective function for the vertical displacement component and in the case of the SH FWI for the horizontal crossline displacement component. We obtain the displacement seismograms from the recorded velocity seismograms by numerical integration.

4.1.5.2. Forward modelling

For the synthetic wave propagation we use an explicit time-domain finite-difference scheme, where we set the time step interval to $\Delta t = 1.4 \cdot 10^{-5}$ s and the grid spacing to $\Delta h = 0.125$ m. The temporal derivations are calculated with second-order accuracy and the spatial derivations with a sixth-order FD-stencil. The model space has a size of 560 grid points in the horizontal direction and 160 grid points in the vertical direction, resulting in the actual dimensions of 70 m x 20 m. The location of the first receiver is at 10 m (figure 4.2) and of the last receiver at 57 m, which provides enough space to the lateral boundary during modelling. Receivers or sources located too close to the lateral boundary could cause strong artifacts in the gradients. We implement the P-SV source as a vertical force and the SH source as a horizontal crossline force. The boundary conditions are satisfied by a C-PML boundary with a size of 10 grid points at the bottom as well as at the lateral boundaries. Moreover, we apply a free-surface on the top of the model to accurately simulate surface waves. The total propagation time of 0.52 s is identical to the time length of the observed seismograms. A detailed description of the parameters that are used in the FD-simulations is given in appendix A, table A.2.

4. Field data application

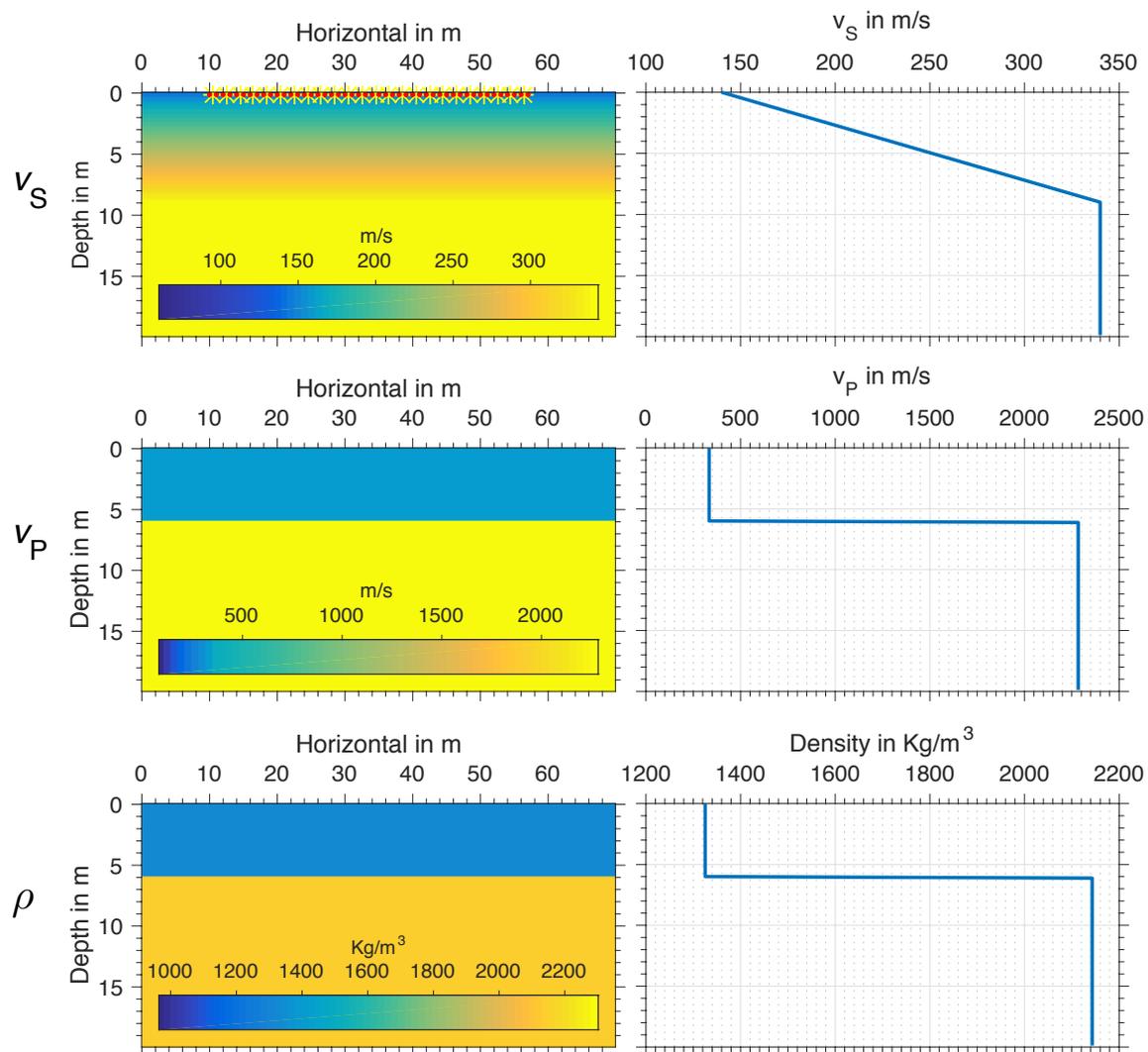


Figure 4.7.: Initial model for v_s , v_p and ρ used in the field data application (left). Depth profiles for the three elastic parameters (right). The yellow stars indicate the source positions and the red points the receiver positions.

4.1.5.3. Source wavelet estimation

In our field data measurements the actual excited source wavelets were not directly recorded, therefore they represent additional unknowns of the inverse problem (Pratt, 1999). To mitigate the effect of an unknown source wavelet, we perform a separate source time function inversion. For the estimation of a source wavelet the linear relation between the source wavelet and the seismograms can be used. This relation allows to calculate a wavelet correction filter by a stabilized deconvolution of the observed seismograms with the synthetic seismograms (Virieux and Operto, 2009; Groos et al., 2014). With such a correction filter an improved source wavelet can be obtained by a convolution of the synthetic wavelet with the estimated wavelet correction filter. For the stability of the deconvolution it is crucial to provide broader frequencies in the synthetic seismograms than in the case of the observed seismograms. However, the improved wavelet does not necessarily represent the actual source wavelet excited in the field measurement, instead it is the wavelet that minimizes the residuals between synthetic and observed seismograms. Therefore, the estimated wavelet might suffer from a trade-off as it could account for residuals caused by an inaccurate parameter model. This trade-off can be mitigated during the inversion, due to the iterative character of the inversion.

In the field data inversions we calculate the wavelet correction filter individually for each source and recalculate it in the case that a change in the inversion configuration occurs. As initial guess we use a cubed sine wavelet with a dominant frequency of 100 Hz. The frequency content of this initial wavelet covers the whole frequency band of the observed dataset and allows for a stable source wavelet estimation. For the source wavelet estimation we only consider traces with a source-receiver offset between 5 m and 10 m. The traces near the source could suffer from source artifacts like overdriven geophones, whereas traces far from the source could be influenced by heterogeneities in the subsurface. In figure 4.13 (first column) we show the corrected source wavelets for the initial parameter model at the first frequency stage (4 to 10 Hz). The estimated wavelets are homogeneous along the whole profile and do not show any artifacts.

4.1.5.4. Preconditioning

To precondition the shot-wise gradients we apply circular source tapers, which decay within 5 grid points around the source position. Additionally, we use the approximation to the diagonal elements of the Hessian (equation 2.78) as preconditioner for the gradients, where we use an identical configuration as in the synthetic experiments. In order to smooth the inversion result, we apply a 2-D median filter to the gradients, where the filter has a size of 1 m (8 grid points). To obtain a physical meaningful elastic parameter model and to ensure stability of the forward solver, we force a minimum v_p/v_s ratio of 1.2 by increasing the P-wave velocity, if necessary. For the calculation of the actual model update we use a normalized multi-parameter L-BFGS method, where the model and gradient differences of the last 20 iterations are evaluated. However, for frequencies above 50 Hz the L-BFGS method did not converge, since no step length could be found that fulfills the Wolfe conditions (section 2.3.3.2). One reason could be that the estimated initial size of the Hessian (scaled identity matrix, equation 2.84) is not accurate enough to allow further convergence. This problem might be solved by providing an external first guess of the

size of the Hessian to the L-BFGS algorithm, for example by the second-order adjoint method (Fichtner and Trampert, 2011). To overcome this issue in the field data inversions, we switch the optimization method to a conjugate gradient method for frequency stages above 50 Hz. In this case we use an inaccurate step length search in combination with a parabolic fitting to estimate an appropriate step length (Nocedal and Wright, 2006). This approach provides stable updates also for frequencies above 50 Hz.

4.1.5.5. Workflow configuration

We use the workflow implementation to increase the frequency content of the dataset gradually from 4 to 130 Hz by lifting up the corner frequency, LC , of a low-pass filter in steps of 5 Hz. In addition, we again apply a successive update strategy to the multi-parameter inversion. In the first iterations we only allow updates of v_S and v_P (for Love FWI only v_S), until the automatic abort criterion is reached and finally we use a full multi-parameter inversion. In total, the workflow is divided in 26 separate stages, which we present in table 4.1.

At the beginning of each new workflow stage we force the algorithm to perform at least three iterations, since both the L-BFGS method as well as the conjugate gradient method rely on the previous iterations. However, in the case of the L-BFGS method we allow the algorithm to switch directly to the next workflow stage in the case that within these three iterations no step length could be found that fulfills the Wolfe conditions. This is required to ensure stability of the quasi-Newton algorithm. In the case of the conjugate gradient method we perform an update with a small step length in the case that within these three iterations no step length could be found that reduces the objective function. We do not check the Wolfe conditions in the case of the conjugate gradient updates, since these are not a stability criterion for this method.

Table 4.1.: Workflow used in the field data FWI. Each stage is applied to the inversion until the automatic abort criterion AC (equation 2.92) is reached. The update column indicates which of the specific elastic parameter is updated (yes=1) or not (no=0). The parameter LP represents the corner frequency of the low-pass frequency filter. The method column indicates whether the L-BFGS or the conjugate gradient method is used for optimization.

Stage	Update			LP in Hz	AC in %	Method
	v_S	v_P	ρ			
1	1	1	0	10	10	L-BFGS
2	1	1	1	10	1	L-BFGS
3–9	1	1	1	Increment of 5	1	L-BFGS
10	1	1	1	50	1	L-BFGS
11	1	1	1	55	1	CG
12–25	1	1	1	Increment of 5	1	CG
26	1	1	1	130	0.5	CG

4.2. Results

In the following, we discuss the result of each field data inversion individually. Afterwards, we make a direct comparison. We present the result for the individual Love wave FWI in figure 4.9, for the individual Rayleigh wave FWI in figure 4.10 and for the simultaneous joint FWI in figure 4.11. We used the same form of presentation in all three figures. In figure 4.14 we show a direct comparison of the final models, where we focus on an area around the trench. Finally, we present horizontal and vertical profiles across the initial and final models in figure 4.15 and 4.16.

4.2.1. Individual Love wave FWI

The final v_S model is still predominantly depth dependent, but contains 2-D lateral variations. At the expected position of the refilled trench the inversion revealed a low-velocity anomaly. The anomaly of the trench has a smooth triangular shape and exhibits a length of 8 m at the surface and a depth of about 2.6 m. To the left of the trench the v_S model contains a second low-velocity anomaly, which is elongated and only present in the shallow part. This anomaly could be related to an increased saturation of the shallow soil, which coincides with the observed soil conditions in this area during the measurements. The SH waves are not sensitive to the P-wave velocity, thus, the v_P model is not updated. The 2-D variations observed in the v_S model are not visible in the ρ model. The inversion increased the density values in both layers and added strong small-scale variations.

The inversion improved the fit of the synthetic seismograms, however, a residual is still present. The fit of the near offset traces is better than of the far offset traces. The waves that arrive at the receivers in the far offset travelled farther and deeper, hence, more subsurface heterogeneities influence those recordings than the recordings at near offset receivers. The inversion decreased the misfit relative to the initial misfit up to the 30 Hz frequency stage, as shown in figure 4.8. For higher frequency stages the inversion could not again decrease the misfit below the misfit level of the previous frequency stage, most likely due to an increase of the noise level at higher frequencies. However, the inversion decreased the misfit within each frequency stage. In figure 4.13(b) (second column) we show the estimated source wavelets for the final model at 130 Hz. The wavelets are quite similar for all source positions, but the source wavelets for the higher source numbers show a ringing effect. This effect could be related to the saturated soil at the end of the profile, which allowed the steel source rack to oscillate slightly. In figure 4.12(b) (third column) the phase-slowness-frequency spectrum of the final synthetic seismograms is shown besides the spectrum of the field dataset. Although we did not perform a dispersion curve fitting, the visual comparison of the synthetic to the observed spectrum is improved. The FWI increased the phase-slowness of the fundamental mode and added the higher mode to the spectrum.

To sum up, we successfully applied the individual Love wave FWI to the near-surface field dataset. The updated v_S model contains a low-velocity anomaly at the expected position of the trench. This velocity anomaly allowed us to derive the size and shape of the refilled trench. The final ρ model provides no further information on the size of the trench.

4.2.2. Individual Rayleigh wave FWI

The individual Rayleigh wave FWI revealed several 2-D structures in the v_S model, which superimpose the mainly lateral homogeneous background model. In the middle of the profile the v_S model contains a square shaped low-velocity anomaly, which corresponds to the refilled trench. The lower edge of this anomaly lies in a depth of approximately 2.2 m. However, the contours of the anomaly are blurred, especially at the surface, where we cannot estimation the horizontal length of the anomaly. To the left of the trench the v_S model contains an elongated shallow second low-velocity anomaly, which could be related to the increased saturation of the shallow soil in this area. In general, the v_S model suffers from slight vertically orientated artifacts underneath some source positions, in particular at the positions of the low-velocity anomalies. We observed similar artifacts in the synthetic example at positions with inaccurate P-wave velocities (see section 3.3). However, we expect potential artifacts to be less dominant in the case of the field data application than in the case of the synthetic example, since we set the size of the median smoothing filter to 1 m for the field data FWI, which is twice as big as in the synthetic example. The overall variations in the final v_P model are small compared to the v_S model. At the position of the trench as well as at the position of the second anomaly in the v_S model we observe slightly reduced P-wave velocities. As demonstrated in the synthetic example, the v_P model could suffer from a cross-talk by v_S , hence, the light anomalies in the v_P model might be a result of cross-talk. Moreover, in the synthetic example we have shown that the resolution in the v_P model is lower compared to the resolution in the v_S model, due to the longer wavelengths of the P-waves than of the S-waves. The final ρ model does not contain any of the anomalies that are present in the v_S model. The inversion increased the density values in the first and second layer and added smooth small-scale vertically orientated artifacts.

The final seismogram fit is not as accurate as in the case of the individual Love wave FWI, however, the P-SV wave field is more complex than the SH wave field. The inversion fitted the fundamental Rayleigh mode quite well at all offsets. Generally, the fit of the phases is better than of the amplitudes. The evolution of the objective function is shown in figure 4.8. The inversion reduced the misfit relative to the initial misfit up to the frequency stage of 30 Hz. From 30 Hz on the inversion only reduced the misfit within each frequency stage, but could not reduced the misfit to the level of the previous frequency stage. The source time inversion revealed similar wavelets for all source positions, as indicated in figure 4.13(a) (second column). In figure 4.12(a) (third column) we present a phase-slowness-frequency spectrum for the final synthetic data as well as for the observed data. The synthetic model describes accurately the dispersion properties of the fundamental mode and the higher modes.

To sum up, we were able to improve the data fit of the P-SV dataset by an individual Rayleigh wave FWI. The final v_S model shows a low-velocity anomaly at the position of the refilled trench. No significant anomaly of the trench is present in the v_P model or the ρ model.

4.2.3. Simultaneous joint FWI

The final v_S model of the simultaneous joint FWI contains a low-velocity anomaly at the expected position of the refilled trench. The anomaly has an identical size as in the case of the individual Love wave FWI. The shape of the anomaly is triangular and the contour is sharp. The anomaly of the trench holds higher velocities in the shallow part than in the lower part. Moreover, the final v_S model contains an elongated shallow second low-velocity anomaly to the left of the trench, which could be related to an increased saturation of the shallow soil within this area. Altogether, the v_S model is still predominantly depth dependent. The variations of the v_P model are light compared to the initial model. At the positions of the v_S anomalies the v_P model shows slightly lower values, which could be a result of a cross-talk. We expect the resolution in the v_P model to be low, as explored in the synthetic experiments. The ρ model contains higher values than the initial model, especially in the upper layer. Neither of the two v_S anomalies can be observed in the density model. However, the ρ model suffers from vertically orientated small-scale lateral heterogeneities.

The fit of the synthetic seismograms to the observed seismograms is not as accurate as in the case of the individual wave type inversions, as opposed to the synthetic experiments, where the joint FWI decreased the misfit of both wave types even further. One reason could be slight anisotropic effects that would influence the propagation of both wave types differently, due to the contrasting polarization of the SH and the P-SV waves. Another reason could be a limited accuracy of the initial v_P model. Nonetheless, the fit of the SH data is better than the fit of the P-SV dataset. Figure 4.8 illustrates the evolution of the joint objective function. The joint FWI reached the lowest misfit for the 20 Hz frequency stage. From 20 Hz on the joint FWI decreased the misfit only within each frequency stage and could not decrease it below the misfit of the previous frequency stages. In figure 4.13 (top and bottom, third column) we present the improved source wavelets for the final frequency stage of 130 Hz. For both wave types the source wavelet estimation revealed homogeneous wavelets across the whole profile, despite a slight ringing effect for the sources with higher source numbers. The fit of the synthetic phase-slowness-frequency spectra to the field data spectra is similar well as in the individual wave type inversions, as shown in figure 4.12 (top and bottom, fourth column).

To sum up, we successfully applied a simultaneous joint FWI of Love and Rayleigh waves to a near-surface field dataset. We were able to reduce the objective functions of both wave types simultaneously and to obtain a single parameter model that accounts for both datasets. The joint inversion revealed a sharp anomaly of the trench in the v_S model. However, no clear anomaly of the trench is visible in the v_P model or the ρ model.

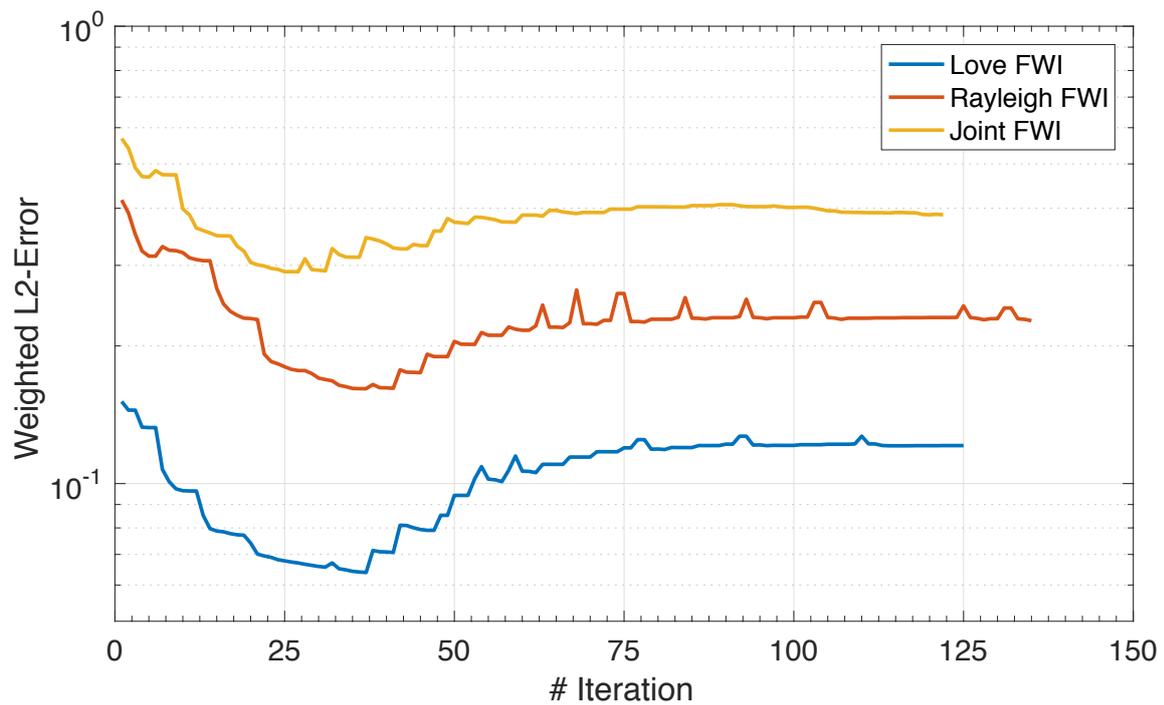


Figure 4.8.: Evolution of the L_2 -error during the field data Love wave FWI, Rayleigh wave FWI and joint FWI over the iterations. The L_2 -error is weighted with the energy of the observed data. The sharp misfit jumps correspond to an increase of the low-pass corner frequency LP (see workflow table 4.1).

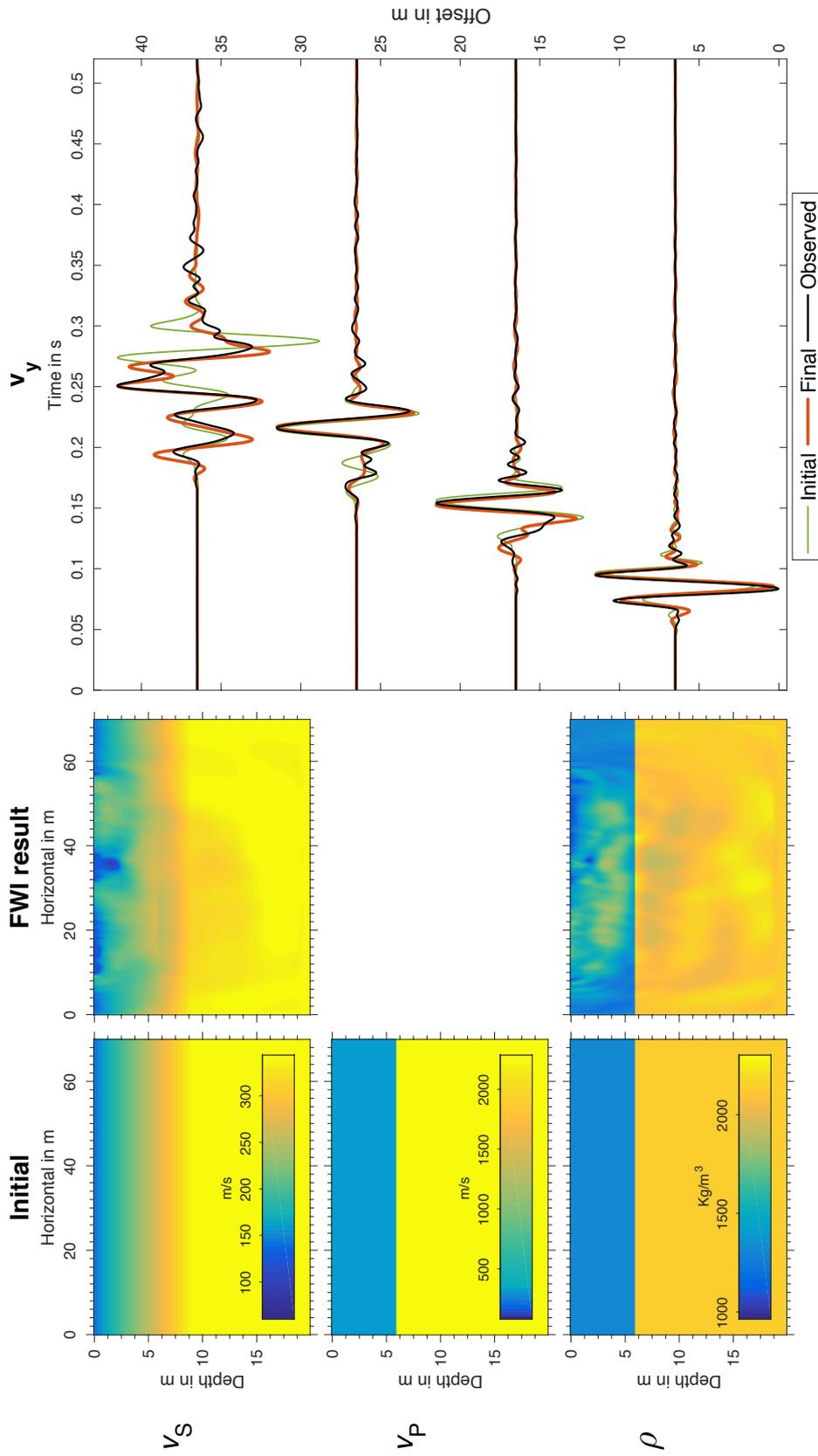


Figure 4.9.: Inversion result of **individual Love wave FWI**. Initial subsurface model besides the final model (left). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Seismograms at four different offsets (right). For each offset a seismogram for the observed dataset (black), for the initial model (green) and for the final model (red) is shown. The source is located at the profile coordinate of 10.5 m. The seismograms are band-pass filtered between 4 and 130 Hz.

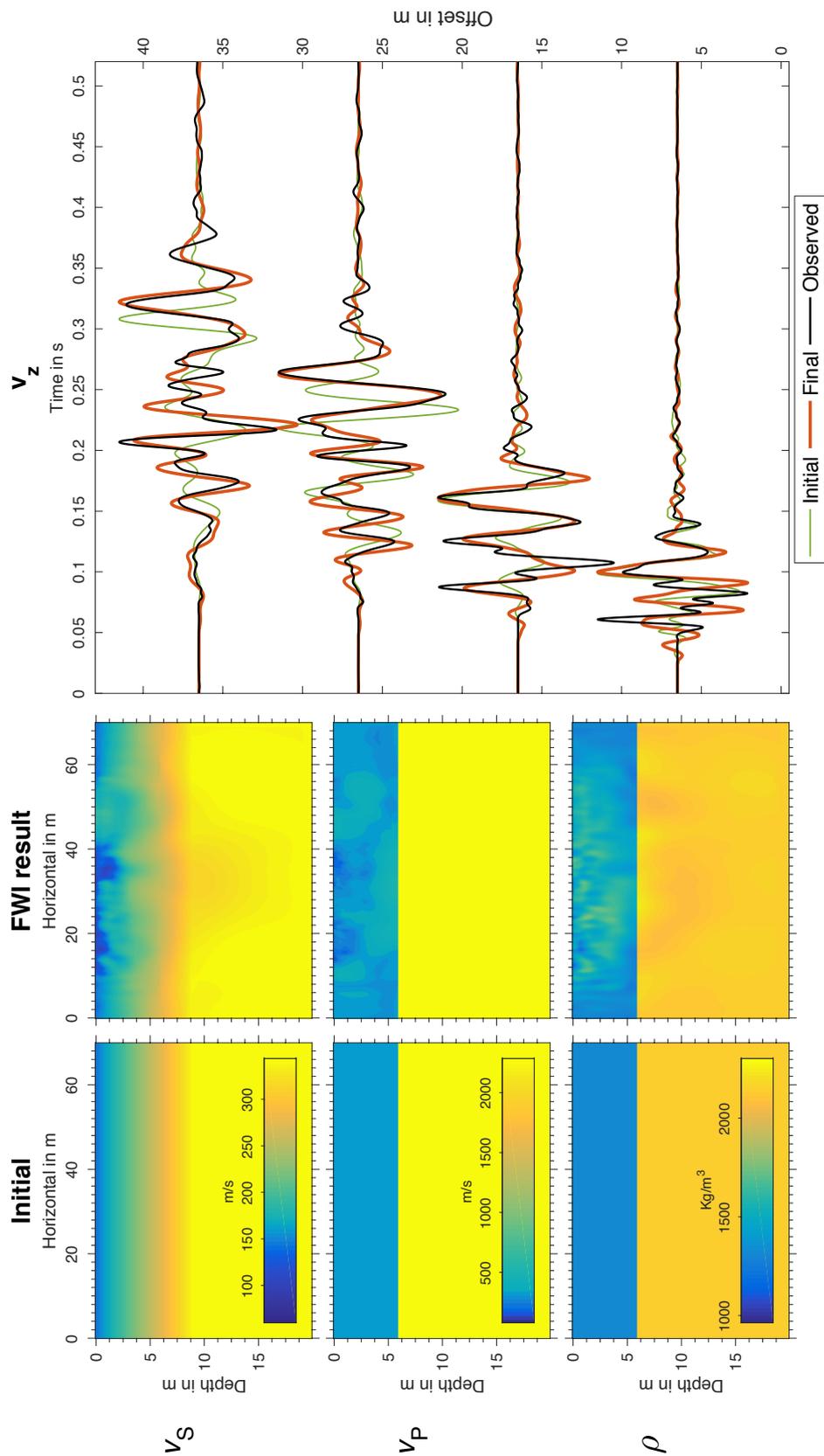


Figure 4.10.: Inversion result of **individual Rayleigh wave FWI**. Initial subsurface model besides the final model (left). The elastic parameters v_s , v_p and ρ are shown row-wise. Seismograms at four different offsets (right). For each offset a seismogram for the observed dataset (black), for the initial model (green) and for the final model (red) is shown. The source is located at the profile coordinate of 10.5 m. The seismograms are band-pass filtered between 4 and 130 Hz.

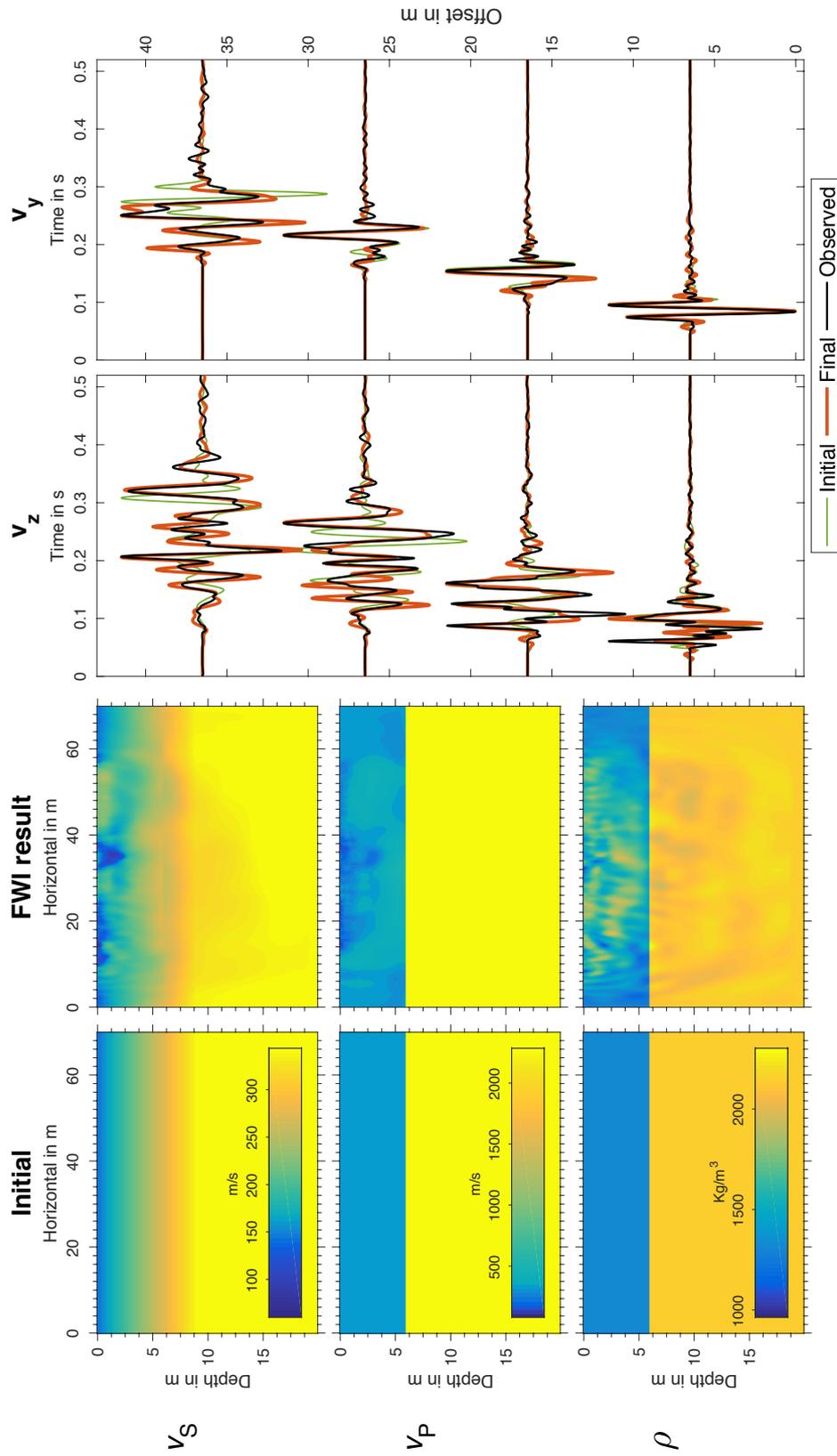


Figure 4.11.: Inversion result of **simultaneous joint FWI**. Initial subsurface model besides the final model (left). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Seismograms at four different offsets (right). For each offset a seismogram for the observed dataset (black), for the initial model (green) and for the final model (red) is shown. The source is located at the profile coordinate of 10.5 m. The seismograms are band-pass filtered between 4 Hz and 130 Hz.

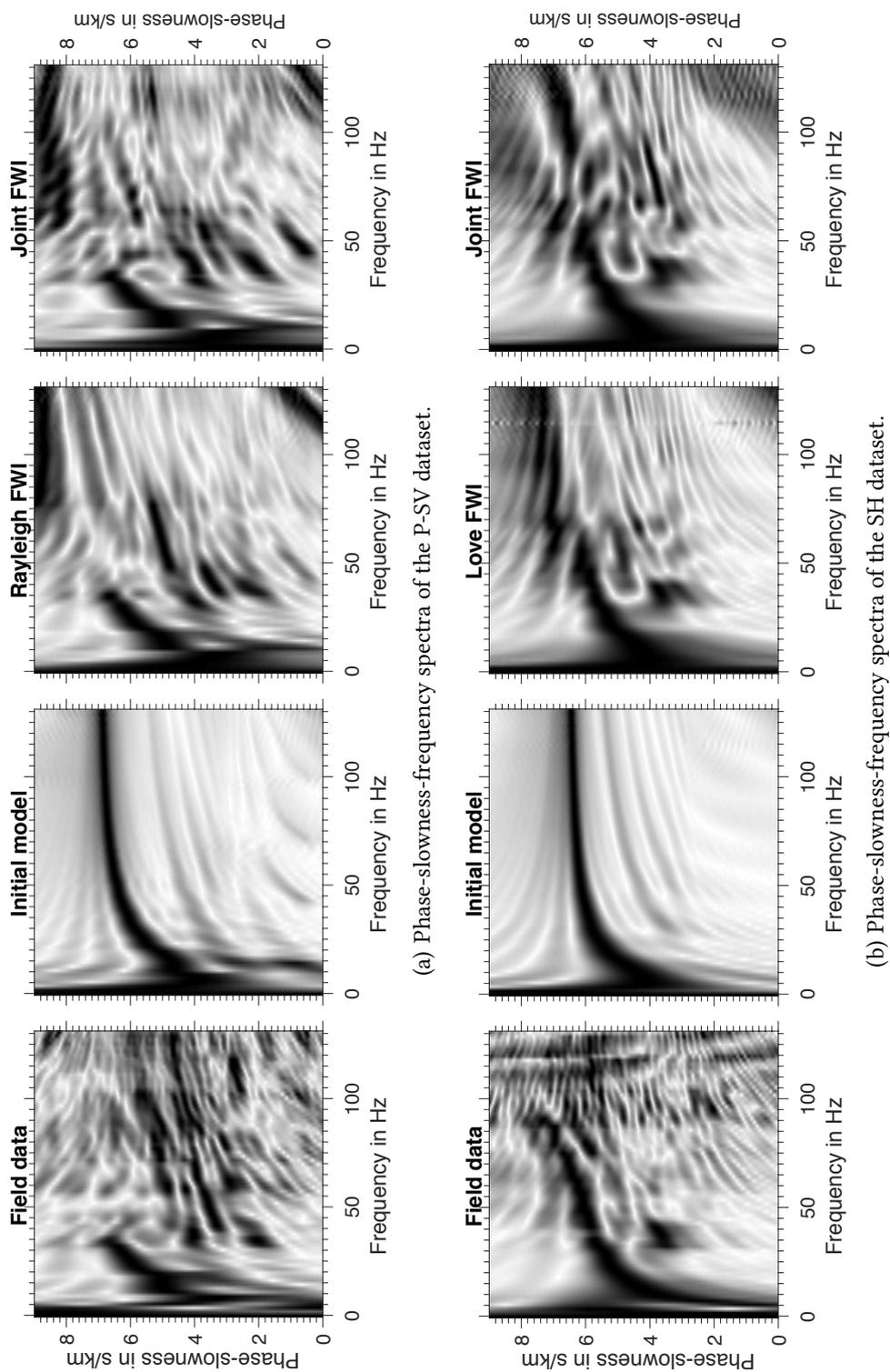
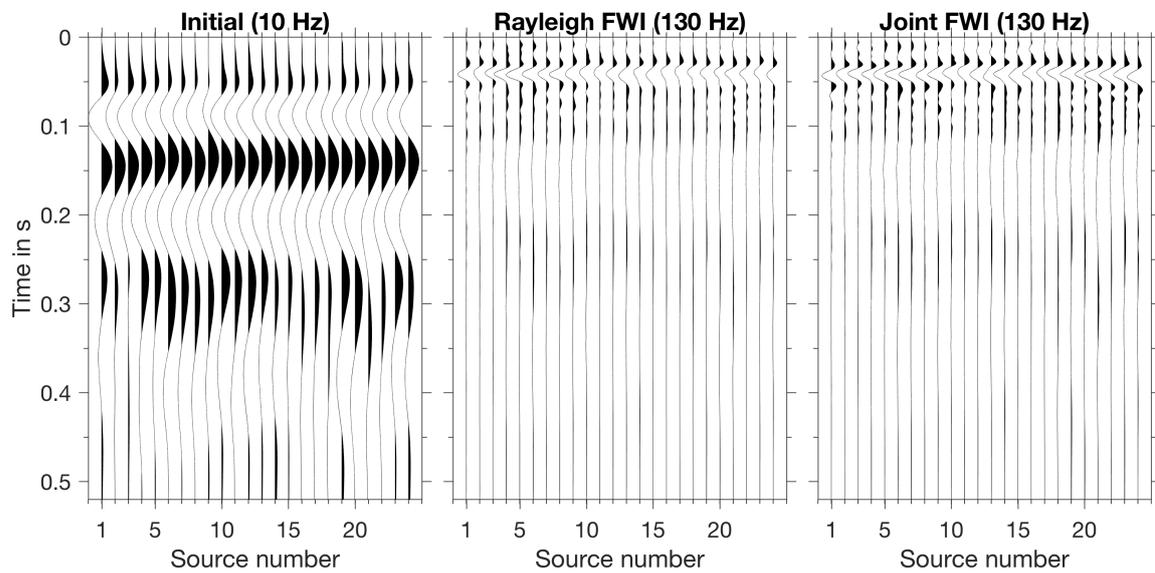
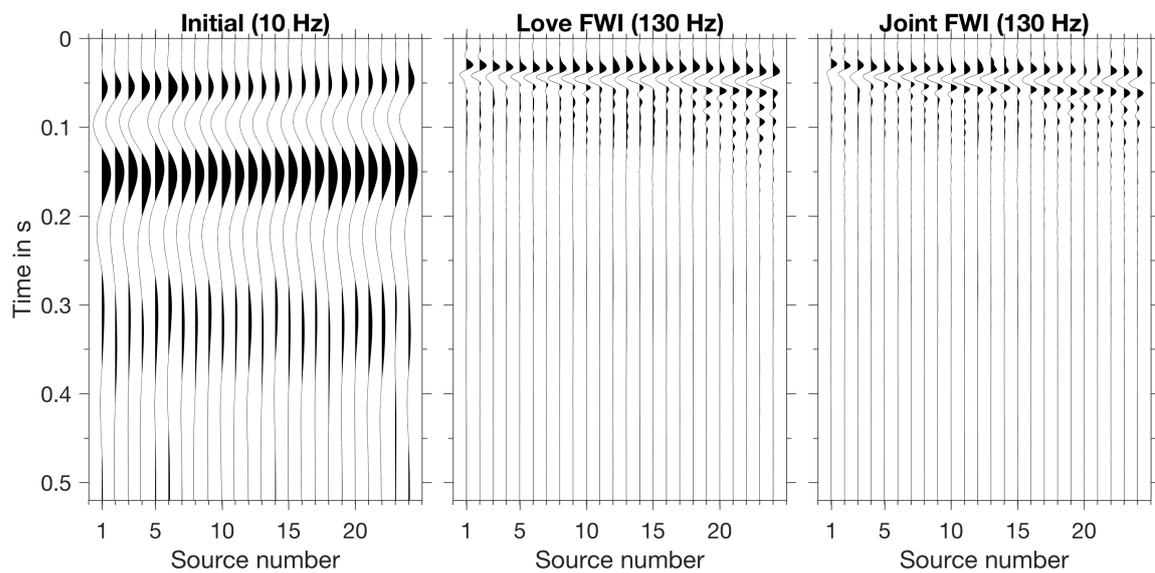


Figure 4.12.: Phase-slowness-frequency spectra for the first source, which is located at the profile coordinate of 10.5 m. Field data (first column). Synthetic data for initial model (second column). Synthetic data for the final model of the individual wave type FWI (third column). Synthetic data for the simultaneous joint FWI (fourth column).



(a) Estimated source time functions for the P-SV source.



(b) Estimated source time functions for the SH source.

Figure 4.13.: The estimated source time functions for the P-SV sources and the SH sources for the initial model at 10 Hz (first column), for the final model at 130 Hz in the case of the individual wave type inversions (second column) and for the final model at 130 Hz in the case of the simultaneous joint FWI (third column).

4.2.4. Comparison

After the individual interpretation of the inversion results we compare them now directly. Figure 4.14 shows the final models column-wise side by side. We thereby focus on a sub-volume of the full model representing the area around the trench. Moreover, figures 4.15 and 4.16 show the horizontal and vertical profiles, whose positions are indicated by figure 4.14.

S-wave velocity v_S :

All three updated v_S models are still mainly depth dependent. However, all inversions revealed a low-velocity anomaly at the expected position of the refilled trench (Ettlinger Linie). The individual Love wave FWI and the simultaneous joint FWI revealed an anomaly of similar shape and size. The anomaly has a triangular shape with a length of approximately 8 m at the surface and a depth of about 2.6 m. Nevertheless, the anomaly holds a sharper contour in the case of the joint FWI than in the case of the individual Love wave FWI. The anomaly revealed by the individual Rayleigh wave inversion holds a square shape, where the lower edge lies in depth of about 2.2 m. In this case the horizontal length of the trench is not clearly imaged. All results contain a second low-velocity anomaly to the left of the trench. This anomaly is horizontally elongated and only present in the shallow part. It is likely that this anomaly is caused by an increased water saturation of the shallow soil within this area, which coincides with the observed soil conditions during the field measurements.

P-wave velocity v_P :

The overall change of the v_P model is small for both the individual Rayleigh wave FWI as well as the joint FWI. Both models show slightly lower P-waves velocities at the positions of the anomalies in the v_S models. Neither of the two inversions changed the velocity of the lower layer. In general, both v_P models are quite similar. However, the slight low-velocity anomaly in the v_P models at the position of the trench could be a result of a cross-talk by v_S , as observed in the synthetic experiments. The SH waves are not sensitive to the P-wave velocity and we therefore do not apply a joint approach to the updates of v_P .

Density ρ :

All three updated ρ models hold higher values than the initial model. In addition, all of them show numerous small-scale variations, which are significantly stronger than in the other two parameter classes. In direct comparison of the three ρ models, the model of the individual Rayleigh wave FWI exhibits the smoothest appearance. The result of the Love wave FWI as well as of the joint FWI exhibit stronger small-scale variations and show several punctual low-density anomalies in the vicinity of the trench. As found in the synthetic experiments, the ρ model is highly sensitive to cross-talk from the v_S model and partly from the v_P model. Moreover, since we normalized the seismograms trace-wise during the calculation of the objective function, we expect the resolution in the ρ model to be generally low. In particular as the influence of density is mainly to the absolute wave amplitude as a function of offset. All in all, the updated ρ models do not provide further information on the shape of the trench.

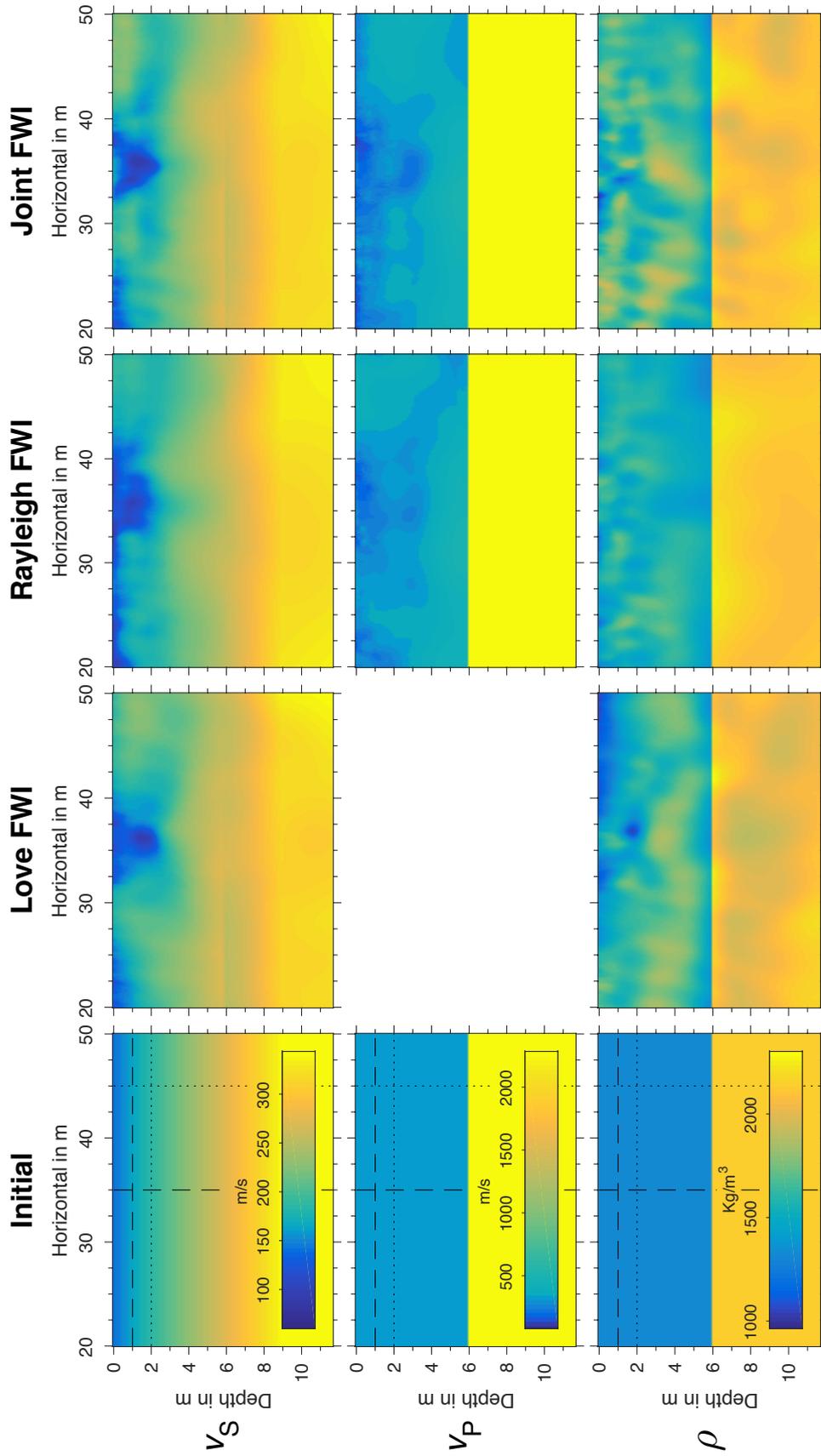


Figure 4.14.: Comparison of the initial model (first column) with the final model of the individual Love wave FWI (second column), the individual Rayleigh wave FWI (third column) and the simultaneous joint FWI (fourth column). The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively. Horizontal and vertical profiles across all models are shown in figure 4.15 for the dashed lines and in figure 4.16 for the dotted lines.

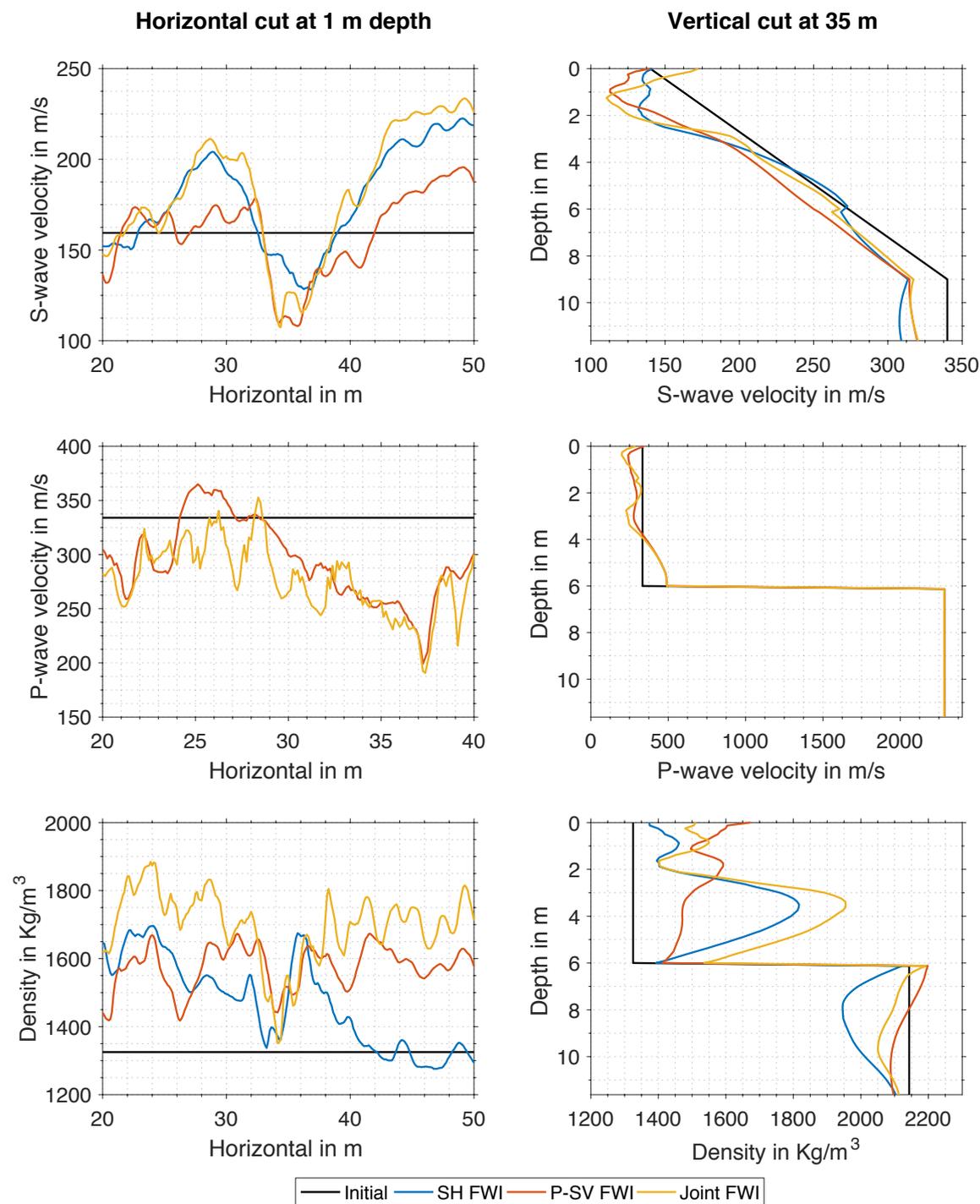


Figure 4.15.: Horizontal profiles at 1 m depth (left) and vertical profiles at the profile coordinate of 35 m (right) across the final models, as indicated by dashed lines in figure 4.14. The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively.

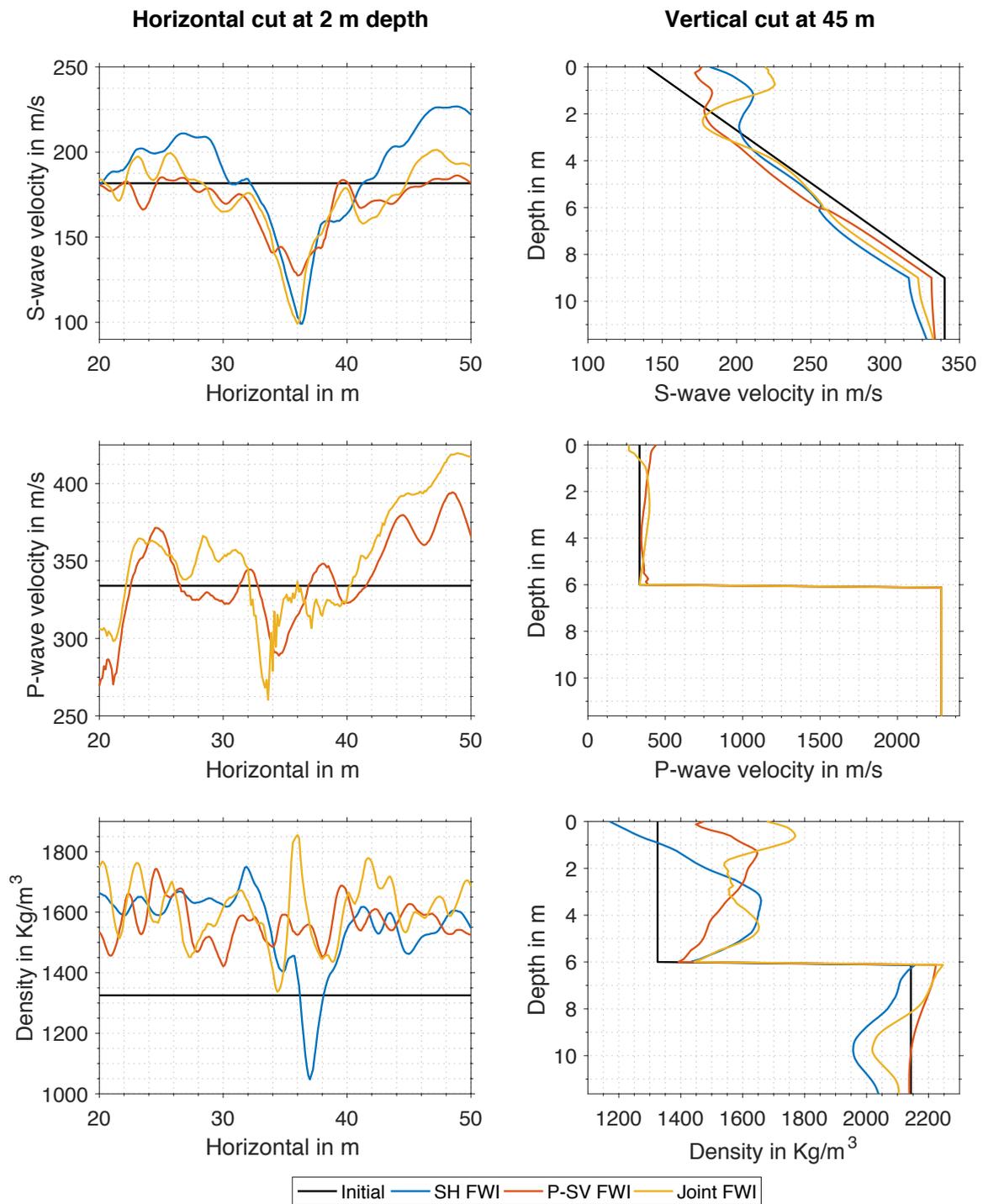


Figure 4.16.: Horizontal profiles at 2 m depth (left) and vertical profiles at the profile coordinate of 45 m (right) across the final models, as indicated by dotted lines in figure 4.14. The elastic parameters v_s , v_p and ρ are shown row-wise from top to bottom, respectively.

4.3. Comparison with ground-penetrating radar

Due to the fact that the true subsurface is unknown in the field data inversions, we cannot draw conclusions on the quality or the reliability of the inversion results. One way to overcome this limitation would be to do test drillings and perform laboratory analysis of soil samples. However, this is costly and mostly neither preferred nor allowed. Another way to overcome this limitation is to verify the inversion results by other measurement techniques. We therefore compare the seismic FWI results obtained in this work with the result of a ground-penetrating radar (GPR) measurement.

GPR measurement and result:

The GPR measurement took place at a later date than the seismic measurements and was carried out in the framework of a master's thesis by [Wegscheider \(2017\)](#). They used the same profile as we did (see map in figure 4.1), however, since they had to relocate the profile by GNSS coordinates the accuracy is limited to 3 – 4 m. For the zero-offset GPR measurement they used a radar of the manufacture IDS GeoRadar with a 200 MHz antenna. They time-migrated the dataset by a constant-velocity Kirchhoff migration, where they chose the migration velocity to 0.1 m/ns. We were provided with the final migrated image shown in figure 4.17, where the trench is visible by boundary reflections which reveal a triangular form. Inside the trench few reflections are visible, suggesting a homogeneous filling of the trench. Besides the location of the trench a lot of small-scale reflectors are present.

Comparability:

We used the final S-wave velocity models to compare the FWI results to the GPR image, since in near-surface applications the FWI could resolve this parameter with higher resolution and lesser ambiguities compared to the other elastic parameters. We call this comparison of both techniques qualitative, since we only link the location of velocity anomalies to reflections in the GPR image. As a consequence of the limited accuracy during the relocation of the seismic acquisition profile, the absolute horizontal position of both results is not comparable. We therefore corrected the lateral position of the GPR image manually. Due to the difficulties relating an accurate depth-migration of the GPR result, we adjusted the depth axis of the GPR result manually in order to fit the FWI results. We used a velocity factor of 0.086 m/ns to transfer the GPR result from the time-domain to the depth-domain of the FWI results.

Qualitative comparison of GPR and FWI results:

The qualitative comparison of the GPR image to the three FWI velocity models is shown in figure 4.18 as an overlay of both results.

The result of the **individual Love wave FWI** matches the GPR image quite well. The velocity model reveals contrasts at positions where strong reflections at the boundaries of the trench are visible. The extension of the low-velocity anomaly is mainly concentrated to the enclosed part of the these boundary reflections. Structures within the trench are not visible in the velocity model. The low-velocity anomaly to the left of the trench correlates with near-surface reflectors, which could represent impermeable layers that

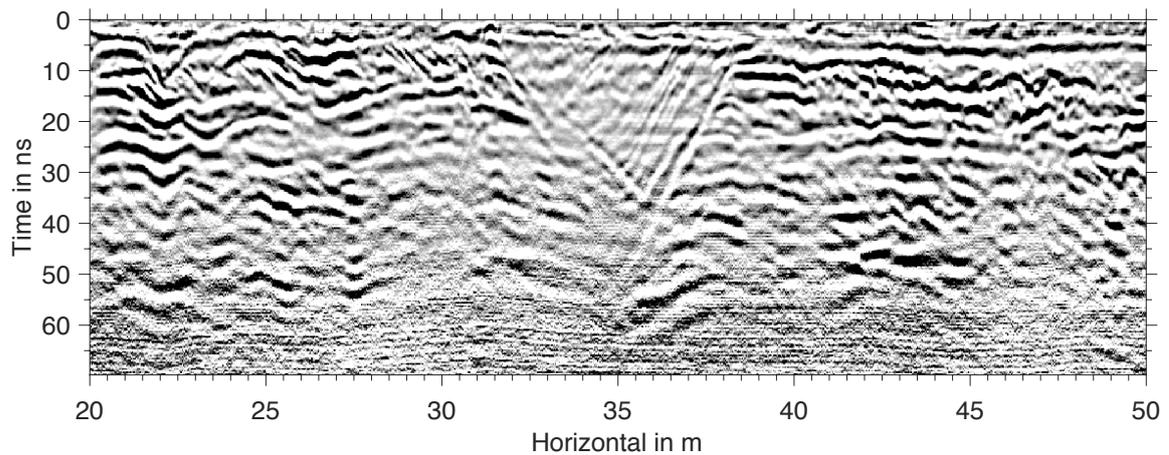


Figure 4.17.: Time-migrated image of the GPR measurement. A linear time-dependent weighting is applied to enhance the visibility of deep reflectors. The GPR result is provided by courtesy of [Wegscheider \(2017\)](#).

lead to accumulated water at the surface.

The velocity model of the **individual Rayleigh wave FWI** satisfactorily matches the migrated GPR image. The center part of the low-velocity anomaly lies within the enclosed area of the boundary reflections of the trench. However, we could not match the horizontal extension of the trench between both images. The lower part of the trench is not visible in the velocity model. The second low-velocity anomaly to the left of the trench correlates again with shallow reflections.

The **simultaneous joint FWI** revealed a low-velocity anomaly that fills the enclosed part of the boundary reflections of the trench accurately. These boundary reflections are visible as sharp contrast within the velocity model. The lower part of the trench matches between both images. The shallow low-velocity anomaly that is present on the left side correlates again with near-surface reflectors.

Summary:

This comparison of the FWI results with the GPR measurement shows that the results of both techniques are in good agreement. Both methods revealed the shape of the trench to be of a triangular form. The S-wave velocity model of the simultaneous joint FWI matches the time-migrated GPR image best. The boundary reflections of the trench can be directly mapped to sharp velocity contrasts. The result of the individual Love wave FWI matches the migrated image similar well. However, the contour of the trench is not as sharp and consistent to the migrated image as in the case of the joint FWI. The velocity model of the individual Rayleigh wave FWI fits the GPR image the least. The boundary of the trench cannot be matched between both results. The lower part of the trench that is indicated in the GPR image is not visible within the velocity model.

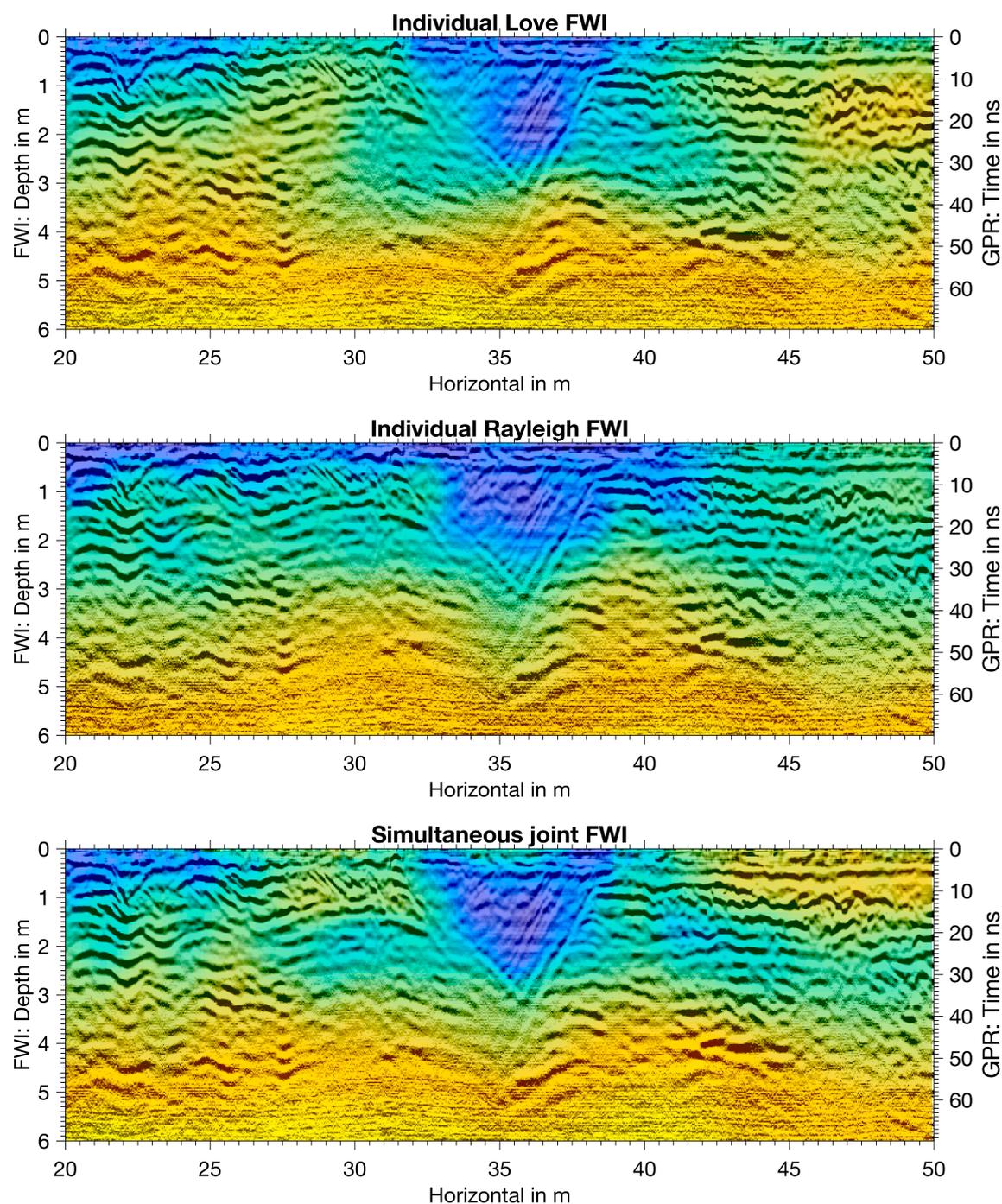


Figure 4.18.: Qualitative comparison of the GPR result with the field data FWI results. Overlay of the final S-wave velocity models of individual Love wave FWI (top), individual Rayleigh wave FWI (middle) and simultaneous joint FWI (bottom) with the time-migrated image of the GPR measurement. No color-bar is shown, because the transparency effect falsify the color representation.

4.4. Summary

In this chapter, we presented the application of the individual and the joint 2-D elastic full-waveform inversion of Rayleigh and Love waves to a near-surface field dataset.

The recording of the field dataset took place on a glider airfield in Rheinstetten near Karlsruhe (Germany). Previous studies on this test site propose a predominantly depth dependent 1-D subsurface that is locally superimposed by a low-velocity anomaly. This anomaly corresponds to a refilled trench that was originally excavated to serve as a line of defense known as "Ettlenger Linie". The recorded field dataset holds good quality for a full-waveform inversion after preprocessing. The P-SV dataset is dominated by a Rayleigh wave and the SH dataset is dominated by a Love wave, which are both surface waves. Based on the recorded dataset we derived initial models for the elastic parameters v_p , v_s and ρ as well as for the attenuation properties. The obtained initial model predicts all main phases of the field dataset and all three inversions allowed for convergence based on this model. The individual Love wave inversion revealed a low-velocity anomaly in the v_s model that allows to derive the size of the triangular shaped trench, although the shape of the trench is not visible in the ρ model. The final v_s model of the individual Rayleigh wave FWI contains a blurred square shaped low-velocity anomaly at the position of the trench, while no significant anomaly of the trench is present in the ρ model and the v_p model. The simultaneous joint inversion of both wave types was successful as well, since it decrease both objective functions simultaneously. At the position of the trench the joint inversion revealed a sharp triangular shaped low-velocity anomaly in the v_s model, where the trench is not visible in the ρ and v_p model. However, the joint inversion could not decrease the objective function of both wave types below the level of the individual wave type inversions.

We compared the S-wave velocity models to the result of a GPR measurement on the same profile, in order to draw conclusions on the quality of the FWI results. Both methods revealed a triangular shaped anomaly of the trench. The result of the simultaneous joint FWI matches the GPR image well, since the velocity anomaly of the trench accurately fills the enclosed part of the reflections at the boundary of the trench. These boundary reflections are represented by sharp contrasts in the velocity model. The result of the individual Love wave FWI fits the GPR result similar well as the result of the joint FWI. Nevertheless, the contours of the trench within the velocity model are not as sharp in the case of the Love wave FWI than in the case of the joint FWI. The individual Rayleigh wave FWI fits the GPR image the least, since we could not match the boundaries of the trench between both images.

To sum up, we successfully applied all three inversions to the near-surface field dataset. The expected trench was revealed as low-velocity anomaly in the v_s models, which allowed us to derive the shape and size of the trench. The individual Love wave FWI and the simultaneous joint FWI produced similar results. Nevertheless, the shape of the trench was sharper in the case of the joint FWI, in particular compared to the GPR image. The result of the individual Rayleigh wave FWI was blurred and allowed only a rough comparison to the GPR image.

5. Conclusions and Outlook

In this thesis, we compared the performance of the individual 2-D elastic full-waveform inversion (FWI) of Rayleigh and Love waves and explored the benefits of a simultaneous joint FWI of both types of surface waves. We therefore extended our existing 2-D elastic P-SV FWI code by a SH FWI scheme as well as by a joint approach, which is required to perform the simultaneous joint FWI. First, we explored the performance of the three inversions in synthetic reconstruction experiments and subsequently verified the results by a field data application.

5.1. Conclusions

Methodical aspects:

The FWI of shallow-seismic surface waves allows to accurately reconstruct the shallow subsurface by the evaluation of the whole information content included in seismic waveforms. The FWI can achieve a resolution below the size of a wavelength and, in contrast to the inversion of dispersion curves, is not limited in terms of subsurface heterogeneity. The acquisition of shallow-seismic surface waves is simple and cost-efficient, since conventional seismic field equipment is sufficient. For these reasons, the FWI provides a valuable contribution to near-surface investigations.

This thesis focused on the application of the FWI to Love and Rayleigh waves, which are both surface waves. The propagation of Love and Rayleigh waves is decoupled in a 2-D medium and thus, the inverse problems are decoupled as well. From this it follows that both wave types require a separate forward solver and that the gradients of both objective functions have to be calculated separately. While the Love waves are influenced by the S-wave velocity as well as by the density, the Rayleigh waves are additionally sensitive to the P-wave velocity. However, both wave types are mainly influenced by the S-wave velocity. The corresponding inversions can only reveal the material parameters which the respective wave type is sensitive to.

A simultaneous joint FWI of Love and Rayleigh waves inverts the waveforms of both wave types at the same time and results in a single subsurface model, which accounts for both datasets. It is based on a joint approach, which combines both objective functions as well as the gradients to joint equivalents. The joint FWI evaluates more data compared to the individual wave type inversions and thus, has the potential to decrease the ambiguities of the inversion result. We formulated the joint approach, without loss of generality, for a joint FWI of SH and P-SV waves recorded in a 2-D isotropic medium, hence, the approach is not limited to the special case of Love and Rayleigh waves, instead it can be used for a joint FWI of arbitrary SH and P-SV waves.

Synthetic reconstruction experiments:

In synthetic reconstruction experiments we investigated the performance of the individual FWI of Love and Rayleigh waves as well as the feasibility of a simultaneous joint FWI of both wave types. In these experiments we knew the true subsurface model, which allowed us to directly study the reconstruction ability of the three inversions by comparing the true model to the reconstructed models. We thereby chose the test setting to be close to the field data application, in order to make both comparable.

In our synthetic experiment both the individual Love wave FWI and the individual Rayleigh wave FWI performed similarly well in terms of accuracy and resolution, given that the initial P-wave velocity model is accurate. However, the individual Love wave FWI provided also an accurate reconstruction of the S-wave velocity model in the absence of an accurate initial P-wave velocity model, whereas the individual Rayleigh wave FWI allowed only for a rough reconstruction of the S-wave velocity model in this case. This reveals a crucial advantage of the individual Love wave FWI as it is independent of the P-wave velocity and therefore comprises a smaller parameter space, which is less likely to suffer from cross-talk effects. From this study, we also conclude that the propagation of Rayleigh waves is highly influenced by the P-wave velocity, however, the reconstruction ability of the Rayleigh wave FWI for the P-wave velocity itself is low.

Furthermore, we found that the simultaneous joint FWI improved the inversion result in the case the initial P-wave velocity model is accurate, whereas it did not improve the inversion result in the case the initial P-wave velocity model is inaccurate. We conclude that the joint FWI can further improve the inversion result provided that both individual inversions hold a similar convergence behavior. However, in the case that one of the individual inversions suffers from a convergence issue, the joint inversion might not be able to overcome this issue.

Therefore, an accurate initial P-wave velocity model is required to make the simultaneous joint FWI of Love and Rayleigh waves feasible and to allow a smooth convergence of the individual Rayleigh wave FWI. However, even in the case the initial P-wave velocity model is inaccurate, a rough reconstruction of the S-wave velocity model is possible by the individual Rayleigh wave FWI and the joint FWI, due to the fact that the surface waves are mainly sensitive to the S-wave velocity. The individual Love FWI is independent of the P-wave velocity and provides also in this case an accurate reconstruction of the S-wave velocity model.

Field data application:

In the field data application we investigated the performance and the applicability of all three inversions in a realistic scenario. For this purpose, we recorded a near-surface field dataset on a glider airfield in Rheinstetten near Karlsruhe (Germany).

In our field data application the individual Love wave FWI and the simultaneous joint FWI obtained similar results, whereas the result of the individual Rayleigh wave FWI deviated from the other two results. One reason could be that the initial P-wave velocity model was not accurate enough, which was shown in the synthetic study to have a negative influence on the convergence behavior of the Rayleigh wave FWI. Nevertheless, there we have also found that an inaccurate initial P-wave velocity model influences the joint FWI negatively. However, due to the fact that the inversion result of the joint FWI is

quite similar to the result of the individual Love wave FWI, which is not influenced by the P-wave velocity, we conclude that the field data joint FWI was not negatively influenced by an inaccurate P-wave velocity model.

To draw conclusions on the quality of the FWI results, we compared the final S-wave velocity models to a ground-penetrating radar (GPR) measurement, which took place on the same profile. The results of the individual Love wave FWI and the simultaneous joint FWI showed an accurate match to the GPR image. Nevertheless, compared to the individual Love wave FWI result, the result of the joint FWI matched the GPR image more precisely, in particular the velocity contrasts were sharper in the case of the joint FWI. The result of the individual Rayleigh wave FWI matched the GPR image only roughly. Based on this qualitative comparison we draw two conclusions. First, the individual Love wave FWI is superior to the individual Rayleigh wave FWI in this field data application, as indicated by the comparison to the GPR image. Second, the joint FWI further improved the inversion result, because its inversion result held sharper contrasts and matched the GPR image most precisely compared to both individual wave type inversions, respectively.

Moreover, this comparison shows the potential of the FWI, as it allows to assign elastic parameters to anomalies indicated in the migrated GPR image, which will help in the interpretation of subsurface anomalies. An approach, where in a first step the GPR is used to locate and map anomalies and subsequently the FWI is applied to obtain the elastic parameters of these anomalies, can be an efficient way to investigate the near-surface.

Recommendations for near-surface investigations by the FWI of surface waves:

To conclude this thesis, we propose two recommendations for near-surface investigations of the S-wave velocity distribution by the FWI of shallow-seismic surface waves.

In the absence of an accurate initial model for the P-wave velocity, we recommend the individual Love wave FWI for three main reasons: (1) Its convergence behavior is smooth and independent of the P-wave velocity, (2) it therefore holds a smaller parameter space, which leads to less cross-talk effects and (3) the SH wave equation is less complex than the P-SV wave equation, which allows a computationally efficient inversion.

In the case an accurate initial model for the P-wave velocity is available, we recommend the individual Love wave FWI against the individual Rayleigh wave FWI for the same reasons. However, in this case a simultaneous joint FWI of both wave types has several advantages compared to both individual wave type inversions: (1) It decreases the ambiguities of the inversion result, since more data is evaluated, (2) it reduces the cross-talk between the elastic parameter classes and (3) it further improves the resolution and accuracy of the inversion result.

5.2. Outlook

Today, the method of choice in shallow-seismic investigations of the S-wave velocity distribution is the inversion of dispersion curves. It can be performed on conventional personal computers, produces stable inversion results and requires only a few preparatory steps. Nevertheless, this method is not suitable for heterogeneous subsurfaces and is limited in terms of resolution. The FWI has the potential to overcome these limitations and

can produce high resolution multi-parameter images of the subsurface. However, the FWI requires large computational facilities and a high amount of preparatory work, such as the estimation of accurate initial models. At the present time there are only few field data applications of the FWI using shallow-seismic surface waves, compared to the dispersion curve inversion, which is applied on a daily basis in engineering offices. For this reason, there is significantly less experience in the application of the FWI than of the dispersion curve inversion. However, the mentioned restrictions of the FWI are only temporary, as computer power increases, experience is gained and the amount of preparatory work is decreased by further methodical developments.

No doubt, the FWI will become the method of choice in shallow-seismic investigations.

Bibliography

- Aki, K. and Richards, P. G. (2002). *Quantitative seismology*, volume 1.
- Binnig, M. (2015). Full Waveform Inversion of shallow-seismic Rayleigh waves to characterize the "Ettlinger Linie". Master's thesis, Karlsruher Institut für Technologie (KIT).
- 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).
- Bohlen, T. (1998). *Viskoelastische FD-modellierung seismischer Wellen zur Interpretation gemessener Seismogramme*. PhD thesis, Christian-Albrechts-Universität zu Kiel.
- 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. and Wittkamp, F. (2016). Three-dimensional viscoelastic time-domain finite-difference seismic modelling using the staggered Adams–Bashforth time integrator. *Geophysical Journal International*, 204(3):1781–1788.
- Brossier, R. (2011). Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, 37(4):444–455.
- 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):WCC105–WCC118.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Butzer, S. (2015). *3D elastic time-frequency full-waveform inversion*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Carcione, J. M., Kosloff, D., and Kosloff, R. (1988). Wave propagation simulation in a linear viscoelastic medium. *Geophysical Journal International*, 95(3):597–611.

- Christensen, R. (2012). *Theory of viscoelasticity: an introduction*. Elsevier.
- Courant, R., Friedrichs, K., and Lewy, H. (1928). Über die partiellen Differenzgleichungen der mathematischen Physik. *Mathematische Annalen*, 100(1):32–74.
- 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*.
- Emmerich, H. and Korn, M. (1987). Incorporation of attenuation into time-domain computations of seismic wave fields. *Geophysics*, 52(9):1252–1264.
- 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.
- Fichtner, A. and Trampert, J. (2011). Hessian kernels of seismic data functionals based upon adjoint techniques. *Geophysical Journal International*, 185:775–798.
- 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.
- Gardner, G., Gardner, L., and Gregory, A. (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics*, 39(6):770–780.
- 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.
- 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.
- 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*.
- Kähler, S. and Meissner, R. (1983). Radiation and receiver pattern of shear and compressional waves as a function of poisson’s ratio. *Geophysical Prospecting*, 31(3):421–435.

- 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.
- 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.
- 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.
- 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.
- Lüttschwager, G. (2014). Simulation and surveying of the near field radiation of seismic vibration sources. Master's thesis, Karlsruher Institut für Technologie (KIT).
- Maurer, H., Greenhalgh, S. A., Manukyan, E., Marelli, S., and Green, A. G. (2012). Receiver-coupling effects in seismic waveform inversions. *Geophysics*, 77(1):R57–R63.
- McMechan, G. A. and Yedlin, M. J. (1981). Analysis of dispersive waves by wave field transformation. *Geophysics*, 46(6):869–874.
- Métivier, L. and Brossier, R. (2016). The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication. *GEOPHYSICS*, 81(2):F1–F15.
- Mora, P. (1987). Nonlinear two-dimensional elastic inversion of multioffset seismic data. *Geophysics*, 52(9):1211–1228.
- Nocedal, J. and Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

- 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 traveltime 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. and Mulder, W. (2004). Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157(3):975–987.
- 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.
- Robertsson, J. O., Blanch, J. O., and Symes, W. W. (1994). Viscoelastic finite-difference modeling. *Geophysics*, 59(9):1444–1456.
- 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.
- Safari, J., O'Neill, A., Matsuoka, T., and Sanada, Y. (2005). Applications of Love wave dispersion for improved shear-wave velocity imaging. *Journal of Environmental & Engineering Geophysics*, 10(2):135–150.
- Schaefer, M. (2014). *Application of full-waveform inversion to shallow-seismic Rayleigh waves on 2D structures*. PhD thesis, Karlsruhe, Karlsruher Institut für Technologie (KIT).
- Shipp, R. M. and Singh, S. C. (2002). Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, 151(2):325–344.
- Socco, L. V., Foti, S., and Boiero, D. (2010). Surface-wave analysis for building near-surface velocity models—established approaches and new perspectives. *Geophysics*, 75(5):75A83–75A102.
- Tarantola, A. (1984). 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.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. siam.
- Tran, K. T., McVay, M., Faraone, M., and Horhota, D. (2013). Sinkhole detection using 2D full seismic waveform tomography. *Geophysics*, 78(5):R175–R183.

- 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.
- Wegscheider, S. (2017). In preparation: Geophysikalische Untersuchung der Ettlinger Linie auf dem Segelflugplatz in Rheinstetten. Master's thesis, Karlsruher Institut für Technologie (KIT).
- 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*.
- Wittkamp, F. and Bohlen, T. (2016). Individual and Joint 2-D Elastic Full Waveform Inversion of Rayleigh and Love Waves. In *78th EAGE Conference and Exhibition 2016*.
- 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.
- Xia, J., Xu, Y., Luo, Y., Miller, R. D., Cakir, R., and Zeng, C. (2012). Advantages of using multichannel analysis of Love waves (MALW) to estimate near-surface shear-wave velocity. *Surveys in Geophysics*, 33(5):841–860.

List of Figures

2.1.	Cartesian coordinate system in 2-D.	6
2.2.	2-D FD-grid for P-SV and SH elastic wave equations.	11
2.3.	Objective function for different scale lengths.	28
2.4.	Flow chart of the FWI algorithm.	32
3.1.	True and initial model for the synthetic FWI experiments.	35
3.2.	Depth profiles of the true and initial synthetic models.	35
3.3.	Cubed sine with a main frequency of 30 Hz, which is used as source signal.	36
3.4.	Amplitude spectrum of the synthetic data.	36
3.5.	Pseudo-observed seismograms in the synthetic reconstruction tests.	38
3.6.	Evolution of the L_2 -error for the synthetic experiments.	42
3.7.	Synthetic reconstruction result of individual Love wave FWI.	43
3.8.	Synthetic reconstruction result of individual Rayleigh wave FWI.	44
3.9.	Synthetic reconstruction result of simultaneous joint FWI.	45
3.10.	Comparison of the synthetic reconstructed models.	47
3.11.	Horizontal and vertical profiles across the final reconstructed models.	48
3.12.	Initial v_P model for the synthetic case study.	50
3.13.	Comparison of the reconstructed models for an inaccurate v_P initial model.	51
4.1.	Overview map of the test site.	54
4.2.	Source and receiver distribution of the field measurement.	55
4.3.	Observed seismograms for the first source position.	57
4.4.	Observed seismograms for the last source position.	58
4.5.	Amplitude spectrum of the field data.	59
4.6.	Q -value estimation for the field dataset.	61
4.7.	Initial model for the field data application.	62
4.8.	Evolution of the L_2 -error for the field data FWI.	68
4.9.	Field data inversion result of individual Love wave FWI.	69
4.10.	Field data inversion result of individual Rayleigh wave FWI.	70
4.11.	Field data inversion result of simultaneous joint FWI.	71
4.12.	Phase-slowness-frequency spectra of the field data.	72
4.13.	Estimated source time functions of the field dataset.	73
4.14.	Comparison of the final models of the field data FWI.	75
4.15.	Horizontal and vertical profiles across the final models of the field data FWI.	76
4.16.	Horizontal and vertical profiles across the final models of the field data FWI.	77
4.17.	Time-migrated image of the GPR measurement.	79
4.18.	Qualitative comparison of the GPR result with the field data FWI results.	80

List of Tables

3.1.	Workflow used in the synthetic reconstruction tests.	39
4.1.	Workflow used in the field data application.	64
A.1.	Configuration of the finite-difference forward modelling and of the inversion algorithm for the synthetic reconstruction tests.	97
A.2.	Configuration of the finite-difference forward modelling and of the inversion algorithm for the field data application.	98

A. Appendix

FD-modelling and inversion parameters

Table A.1.: Configuration of the finite-difference forward modelling and of the inversion algorithm for the synthetic reconstruction tests.

Modelling parameter:	
Model size ($NX \cdot NZ$)	460x128 Grid points
Grid spacing	0.125 m
Total time length	0.8 s
Temporal sampling	$2 \cdot 10^{-5}$ s
Spatial accuracy	6. Order
Temporal accuracy	2. Order
PML boundary size	10 Grid points
PML reference frequency	30 Hz
PML reference velocity	1500 m/s
Free surface	Imaging technique
Inversion parameter:	
Min. number of iterations per stage	3
Frequency filter	4. Order Butterworth
Frequency bandwidth	4 Hz – 60 Hz
Frequency steps	10 Hz
L-BFGS Historic steps	20
Circular source tapers	5 Grid points
Median gradient smoothing	4 Grid points
Hessian water level ϵ (P-SV)	$5 \cdot 10^{-2}$
Hessian water level ϵ (SH)	$5 \cdot 10^{-3}$

Table A.2.: Configuration of the finite-difference forward modelling and of the inversion algorithm for the field data application.

Modelling parameter:	
Model size ($NX \cdot NZ$)	560x160 Grid points
Grid spacing	0.125 m
Total time length	0.52 s
Temporal sampling	$1.4 \cdot 10^{-5}$ s
Spatial accuracy	6. Order
Temporal accuracy	2. Order
τ -value for Q-approximation	0.1576
Relaxation frequencies	0.2978 Hz, 6.7325 Hz, 84.6014 Hz
Reference frequency	40 Hz
PML boundary size	15 Grid points
PML reference frequency	50 Hz
PML reference velocity	1500 m/s
Free surface	Imaging technique
Inversion parameter:	
Min. number of iterations per stage	3
Frequency filter	4. Order Butterworth
Frequency bandwith	4 Hz – 130 Hz
Frequency steps	5 Hz
L-BFGS Historic steps	20
Circular source tapers	5 Grid points
Median gradient smoothing	8 Grid points
Hessian water level ϵ (P-SV)	$5 \cdot 10^{-2}$
Hessian water level ϵ (SH)	$5 \cdot 10^{-3}$

Danksagungen

Abschließend möchte ich mich bei allen Personen bedanken, welche diese Arbeit erst in diesem Umfang ermöglicht haben.

Bei Professor Thomas Bohlen möchte ich mich für die Betreuung der Arbeit und für die zahlreichen Ratschläge bedanken. Des weiteren bin ich dankbar für die Möglichkeit mich auch weiterhin wissenschaftlich zu verwirklichen. Professor Friedemann Wenzel möchte ich danken für die Übernahme des Korreferates.

Ein weiter Dank geht an die gesamte Arbeitsgruppe, welche für Diskussionen jeglicher Art zur Verfügung standen. Speziell bedanken möchte ich mich bei Niko, Laura und Martin für das Korrekturlesen dieser Arbeit. Bei Claudia möchte ich mich für die Hilfe bei Verwaltungsakten bedanken. Thomas Forbriger bin ich dankbar für die Unterstützung bei der Fehlersuche in IFOS und für die Bereitstellung zahlreicher Programme und Bibliotheken. André hat in mir das Interesse für die numerische Modellierung geweckt, vielen Dank dafür und für all die Ratschläge in der Vergangenheit. Besonders bedanken möchte ich mich bei Amelie, für das Korrekturlesen, für deine Unterstützung in den letzten Jahren und dafür, dass du mich auch mal auf andere Gedanken gebracht hast.

Bei meiner Familie möchte ich mich für die unbegrenzte und kontinuierliche Unterstützung während des ganzen Studiums bedanken. Ohne diese tiefe Unterstützung wäre mein Studium nicht möglich gewesen.

Acknowledgements

This work was performed on the high-performance computer InstitutsCluster II (IC2) at Steinbuch Centre for Computing at Karlsruhe Institute of Technology and on the computational resource ForHLR I funded by the Ministry of Science, Research and the Arts Baden-Württemberg and DFG ("Deutsche Forschungsgemeinschaft")