INDIVIDUAL AND JOINT 2-D ELASTIC FULL-WAVEFORM INVERSION OF RAYLEIGH AND LOVE WAVES

F. Wittkamp and T. Bohlen

email: *florian.wittkamp@kit.edu*

keywords: *Near-surface investigation, Surface waves, full-waveform inversion*

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.

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 sledge-hammer 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. The main objectives of this work are (1) to compare the performance of the individual full-waveform inversion of Rayleigh and Love waves and (2) to explore the benefits of a simultaneous joint FWI of both types of surface waves.

THEORY

The basic concept of the 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 it minimizes the misfit between the synthetic data and the observed data. As its name suggests, 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, the FWI can achieve a resolution below the size of a wavelength.

Inverse problem

To formulate the inversion problem of the FWI, we have to parameterize the model space, $\mathbf{m} = (m_1, ..., m_N)^T$. 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, $\mathbf{d}_{syn}(\mathbf{m})$, based on a certain model, $\mathbf{m}$, we can use the nonlinear forward operator, f:

$$\mathbf{d}_{syn}(\mathbf{m}) = f(\mathbf{m}). \tag{1}$$

The data residuals, $\Delta\mathbf{d} = (\Delta d_1, ..., \Delta d_M)^T$, between the synthetic data, $\mathbf{d}_{syn}(\mathbf{m})$, and the observed data, $\mathbf{d}_{obs}$, are defined as:

$$\Delta\mathbf{d} = \mathbf{d}_{syn}(\mathbf{m}) - \mathbf{d}_{obs}. \tag{2}$$

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

$$E(\mathbf{m}) = \frac{1}{2} \cdot \Delta\mathbf{d}^T \cdot \Delta\mathbf{d}, \tag{3}$$

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 3 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:

$$\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{m}. \tag{4}$$

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

$$E(\mathbf{m}) = E(\mathbf{m}_0) + \Delta\mathbf{m}\left(\frac{\partial E(\mathbf{m}_0)}{\partial\mathbf{m}}\right)$$
$$+ \frac{1}{2}\Delta\mathbf{m}\left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial\mathbf{m}^2}\right)\Delta\mathbf{m}^T + O\{||\Delta\mathbf{m}||^3\}. \tag{5}$$

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

$$\frac{\partial E(\mathbf{m})}{\partial\mathbf{m}} = \frac{\partial E(\mathbf{m}_0)}{\partial\mathbf{m}} + \Delta\mathbf{m}\left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial\mathbf{m}^2}\right) \overset{!}{=} 0. \tag{6}$$

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

$$\Delta\mathbf{m} = -\left(\frac{\partial^2 E(\mathbf{m}_0)}{\partial\mathbf{m}^2}\right)^{-1}\frac{\partial E(\mathbf{m}_0)}{\partial\mathbf{m}} = -\mathbf{H}^{-1}\cdot\bigtriangledown_{\mathbf{m}}E(\mathbf{m}_0), \tag{7}$$

where $\bigtriangledown_{\mathbf{m}}E(\mathbf{m}_0)$ is the gradient of the objective function with respect to the N model parameters m_i. 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 4 and 7 we obtain the model update for iteration K by:

$$\mathbf{m}_{K+1} = \mathbf{m}_K + \Delta\mathbf{m}_K = \mathbf{m}_K - \mathbf{H}_K^{-1}\cdot\bigtriangledown_{\mathbf{m}}E(\mathbf{m}_K). \tag{8}$$

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 use to derive the gradients shown in following section. 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 complicated and computational too expensive for large-scale optimization problems like in the FWI, even on modern HPC systems. To overcome this limitation, we use a quasi-Newton limited-memory BFGS (L-BFGS) method, which avoids an explicit calculation of the Hessian. Instead, the L-BFGS method calculates an approximation to the inverse Hessian implicitly at every iteration by measuring the changes in gradients and models from the n most recent iterations (Nocedal and Wright, 2006). We thereby use a Wolfe conditions based step length search in order to ensure the stability of the L-BFGS algorithm. To perform a multi-parameter L-BFGS update we follow the approach of Brossier (2011), who proposes a dimensionless multi-parameter L-BFGS method, where normalized parameter classes are used within the L-BFGS algorithm. This approach allows to calculate updates for parameters with different units and magnitudes by a single L-BFGS algorithm. In this work, we use the arithmetic mean value for the normalization of the individual parameter classes.

Adjoint-state gradients

The calculation of the gradient of the objective function by an actual derivative of the objective function with respect to every single model parameter would need as much forward calculations as model parameters

(Virieux and Operto, 2009). 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.

Based on the isotropic and elastic SH and P-SV 2-D wave equations (Lay and Wallace, 1995; Virieux, 1986, 1984) and by using the adjoint-state approach according to Mora (1987), we derive the gradients for Love (SH) and Rayleigh (P-SV) waves for the density ρ and the Lamé-parameters μ and λ in stress-velocity formulation for one source. We hereby refer to the thesis of Wittkamp (2016) for a detailed derivation of these gradients. For simplicity we omit the temporal and spatial dependencies. The gradients for the SH waves are:

$$\frac{\partial E^{\text{SH}}}{\partial \rho} = -\int dt \cdot v_y^{\text{F}} \cdot v_y^{\text{B}}, \tag{9}$$

$$\frac{\partial E^{\text{SH}}}{\partial \mu} = -\int dt \cdot \frac{\left(\sigma_{xy}^{\text{F}} \cdot \sigma_{xy}^{\text{B}} + \sigma_{zy}^{\text{F}} \cdot \sigma_{zy}^{\text{B}}\right)}{\mu \cdot \mu}, \tag{10}$$

and for the P-SV waves:

$$\frac{\partial E^{\text{P-SV}}}{\partial \rho} = -\int dt \cdot \left[v_x^{\text{F}} \cdot v_x^{\text{B}} + v_z^{\text{F}} \cdot v_z^{\text{B}}\right], \tag{11}$$

$$\frac{\partial E^{\text{P-SV}}}{\partial \mu} = -\int dt \cdot \left[\frac{\left(\sigma_{xz}^{\text{F}} \cdot \sigma_{xz}^{\text{B}}\right)}{\mu \cdot \mu} + \frac{1}{4}\left(\frac{(\sigma_{xx}^{\text{F}} + \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} + \sigma_{zz}^{\text{B}})}{(\lambda + \mu)^2} + \frac{(\sigma_{xx}^{\text{F}} - \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} - \sigma_{zz}^{\text{B}})}{\mu \cdot \mu}\right)\right], \tag{12}$$

$$\frac{\partial E^{\text{P-SV}}}{\partial \lambda} = -\int dt \cdot \frac{(\sigma_{xx}^{\text{F}} + \sigma_{zz}^{\text{F}})(\sigma_{xx}^{\text{B}} + \sigma_{zz}^{\text{B}})}{4(\lambda + \mu)^2}, \tag{13}$$

where v describes the velocity and σ the stress. The upper indices correspond to the forward (F) propagated incident wave field or to the backward (B) propagated adjoint (residual) wave field, respectively. These gradients can be interpreted as a zero-lag cross-correlation between the incident and adjoint wave field.

For the derivation of the gradients we have chosen a model parameterization in terms of the Lamé-parameters and density, since this is proposed by Mora (1987) to be the simplest one for the derivation of the adjoint-state gradients. As a consequence, we also derived the gradients for these parameters. 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 $\mathbf{m} = (\boldsymbol{\rho}, \mathbf{v}_{\text{S}}, \mathbf{v}_{\text{P}})^T$. We derive the gradients for the seismic velocities and the density by using the chain rule of derivatives. The transformation of the parameterization reads (Köhn et al., 2012):

$$\frac{\partial E}{\partial v_{\text{P}}} = 2 \cdot \rho \cdot v_{\text{p}} \cdot \frac{\partial E}{\partial \lambda}, \tag{14}$$

$$\frac{\partial E}{\partial v_{\text{S}}} = -4 \cdot \rho \cdot v_{\text{S}} \cdot \frac{\partial E}{\partial \lambda} + 2 \cdot \rho \cdot v_{\text{S}} \cdot \frac{\partial E}{\partial \mu}, \tag{15}$$

$$\frac{\partial E}{\partial \rho'} = (v_{\text{P}}^2 - 2 \cdot v_{\text{S}}^2) \cdot \frac{\partial E}{\partial \lambda} + v_{\text{S}}^2 \cdot \frac{\partial E}{\partial \mu} + \frac{\partial E}{\partial \rho}. \tag{16}$$

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

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. 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.

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 3 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:

$$E_w(\mathbf{m}) = \frac{1}{2} \cdot \frac{\Delta \mathbf{d}^T \cdot \Delta \mathbf{d}}{\mathbf{d}_{\text{obs}}^T \cdot \mathbf{d}_{\text{obs}}}. \tag{17}$$

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.

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

$$E^{\text{JOINT}}(\mathbf{m}) = E_w^{\text{P-SV}}(\mathbf{m}) + E_w^{\text{SH}}(\mathbf{m}). \tag{18}$$

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.

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:

$$\delta \hat{\mathbf{g}}^{\text{JOINT}} = \left[\frac{\delta \hat{\mathbf{g}}^{\text{P-SV}}}{\max(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|)} + \frac{\delta \hat{\mathbf{g}}^{\text{SH}}}{\max(|\delta \hat{\mathbf{g}}^{\text{SH}}|)} \right] \cdot \left(\max\left(|\delta \hat{\mathbf{g}}^{\text{P-SV}}|\right) + \max\left(|\delta \hat{\mathbf{g}}^{\text{SH}}|\right) \right), \tag{19}$$

where $\delta \hat{\mathbf{g}} = \{ \frac{\partial E}{\partial \boldsymbol{\rho}}$ or $\frac{\partial E}{\partial \mathbf{v_s}} \}$. 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.

SYNTHETIC FWI EXPERIMENTS

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 the following, we introduce the subsurface model, the acquisition geometry and the FWI workflow. We choose a similar test setting as during the field measurements, in order to obtain relevant information which will be required during the field data inversion, where we will present the application of our methodology to a near-surface field dataset.

True and initial models

In preparation of the field measurements we choose the synthetic model to be close to the subsurface model at the desired test site of the field measurement. We obtain the subsurface model for this location from previous 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 synthetic 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 2 (left column), where we focus on an area around the trench. The model has a size of 60 m in the horizontal and of 16 m in the 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.

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. For v_S we use a gradient of 20.4 $\frac{\text{m/s}}{\text{m}}$ and for ρ we use a gradient of 154.2 $\frac{\text{Kg/m}^3}{\text{m}}$. 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.

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. In case of the P-SV simulations the receivers record the vertical velocity component, v_z, and in case of the SH simulations they record the horizontal crossline velocity component, v_y. We distribute the sources along the surface as well. 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:

$$s(t) = \sin(f_\text{d} \cdot \pi \cdot t)^3, \tag{20}$$

where f_d is the dominant frequency. We set the dominant frequency to 30 Hz, which generates frequencies between 0 and 60 Hz.

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 a time-domain finite-difference forward solver on a staggered-grid (Bohlen, 2002). 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 condition based on the mirroring technique of Levander (1988) and at the bottom as well as on the lateral boundaries we assume a C-PML boundary (Komatitsch and Martin, 2007). We set the temporal sampling to $2 \cdot 10^{-5}$ s and the total recording time to 0.8 s. We place the receivers and sources one grid point below the free surface to ensure an accurate amplitude scaling (Groos, 2013). The temporal derivations are calculated with second-order accuracy and the spatial derivations with a sixth-order FD-stencil. In figure 1 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. 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, 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.

FWI setup

For the synthetic reconstruction tests we further developed our existing 2-D P-SV FWI code (Köhn, 2011) by a SH FWI scheme as well as by the joint approach. For the gradient calculation we use the presented adjoint-state gradients (equations 9–13), which we transformed to a parameterization in terms of seismic velocities and density by equations 14–16. We obtain the actual model update by a normalized multi-parameter L-BFGS method, where the model and gradient differences of the last 20 iterations are used to estimate the inverse Hessian.

As objective function we use the weighted L_2-error (equation 17) between the normalized synthetic seismograms and the normalized pseudo-observed seismograms. We apply the normalization trace-wise, in oder to be consistent with the field data inversions, where we use this to be able to account for un-certainties and differences in the coupling of individual geophones to the ground. In the P-SV case, we calculate the objective function for the vertical displacement component and in the SH case for the hor-izontal 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 an approximation of the diagonal elements of the Hessian as a preconditioner for the gradient. This approximation is based on the sum of the amplitudes, u_i, of the forward modeled 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):

$$\mathbf{H}_a^{-1}(\mathbf{x}_s, \mathbf{x}) = \left[\epsilon + \int \mathrm{d}t |u_i(\mathbf{x}_s, \mathbf{x}, t)|^2 \left(\mathrm{asinh}\left(\frac{x_r^{\max} - x}{z} \right) - \mathrm{asinh}\left(\frac{x_r^{\min} - x}{z} \right) \right) \right]^{-1}, \quad (21)$$

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. We choose this water level empirically to $5 \cdot 10^{-3}$ for the SH waves and to $5 \cdot 10^{-2}$ for the P-SV waves. We calculate this preconditioner shot-wise and apply it normalized to the shot-wise 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
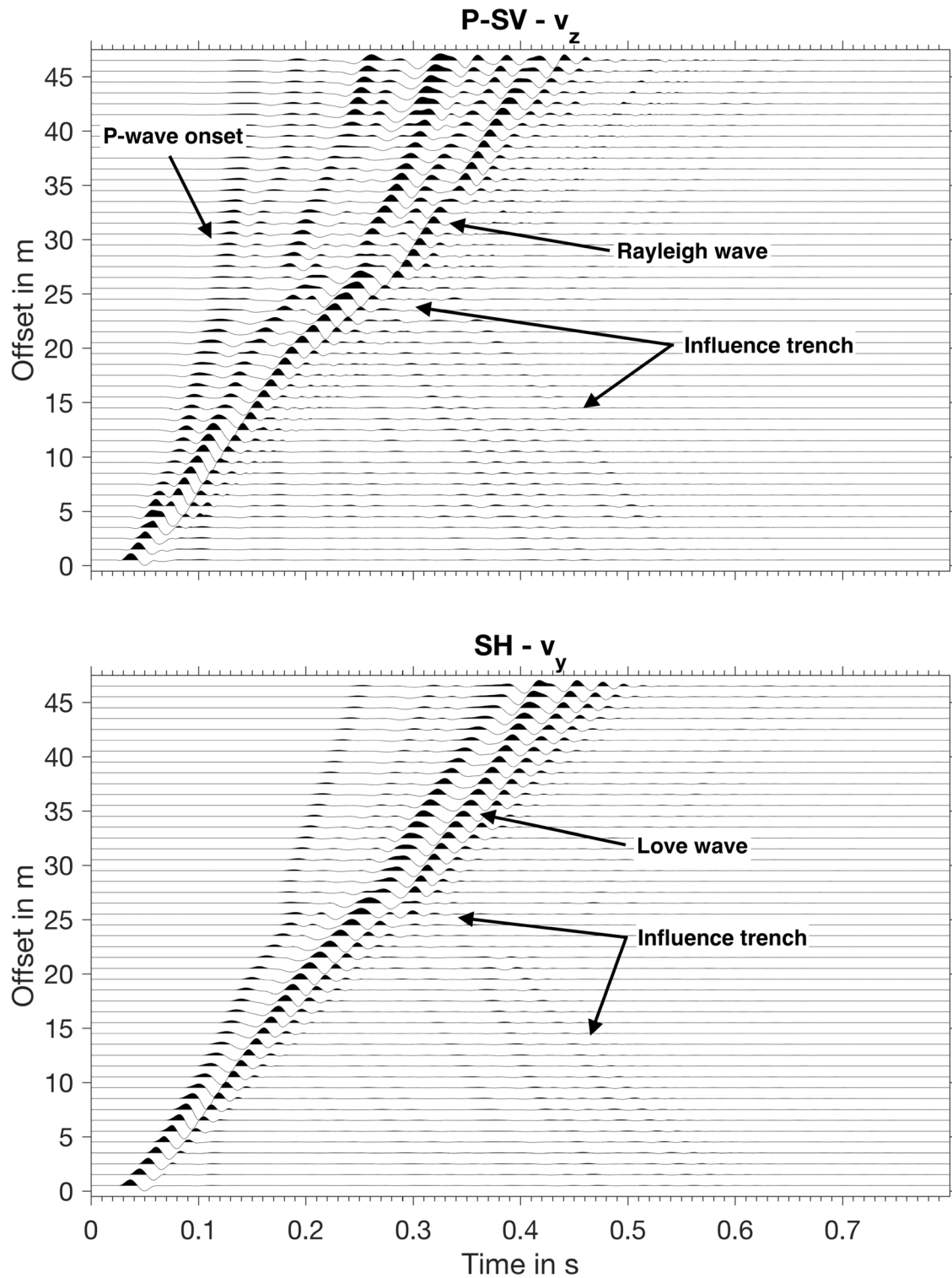
Figure 1: Trace-wise normalized shot-gather of the pseudo-observed seismic data. The corresponding source is located at the profile coordinate of 6.5 m. The seismograms are band-pass filtered between 4 and 60 Hz, representing the frequency range used in the FWI.

Table 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

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, 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 1. As abort criterion we use the reduction of the objective function from the second last to the current iteration:

$$AC_K(\text{in } \%) > \left(\frac{E_{K-2} - E_K}{E_{K-2}} \right). \tag{22}$$

Results

In the following, we discuss the result of each reconstruction experiment individually. In figure 2 we show a direct comparison of the final reconstructed models as well as the true model, where we focus on an area around the trench. Figure 3(a)-(c) shows a comparison of seismograms for the initial model, for the final reconstructed model and of the pseudo-observed data set for each inversion scheme, respectively.

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. 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(d). 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.

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. 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 later on. 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. 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. The inversion decreased the misfit smoothly in each frequency stage and reduced the objective function in total by four orders of magnitude, as shown in figure 3(d).

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(d). 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 joint objective function as a sum of both individual misfits, the total joint objective function shown in figure 3(d) is higher than both individual ones, since it is the sum of both.

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.

The inaccurate initial v_P model used in this case study 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
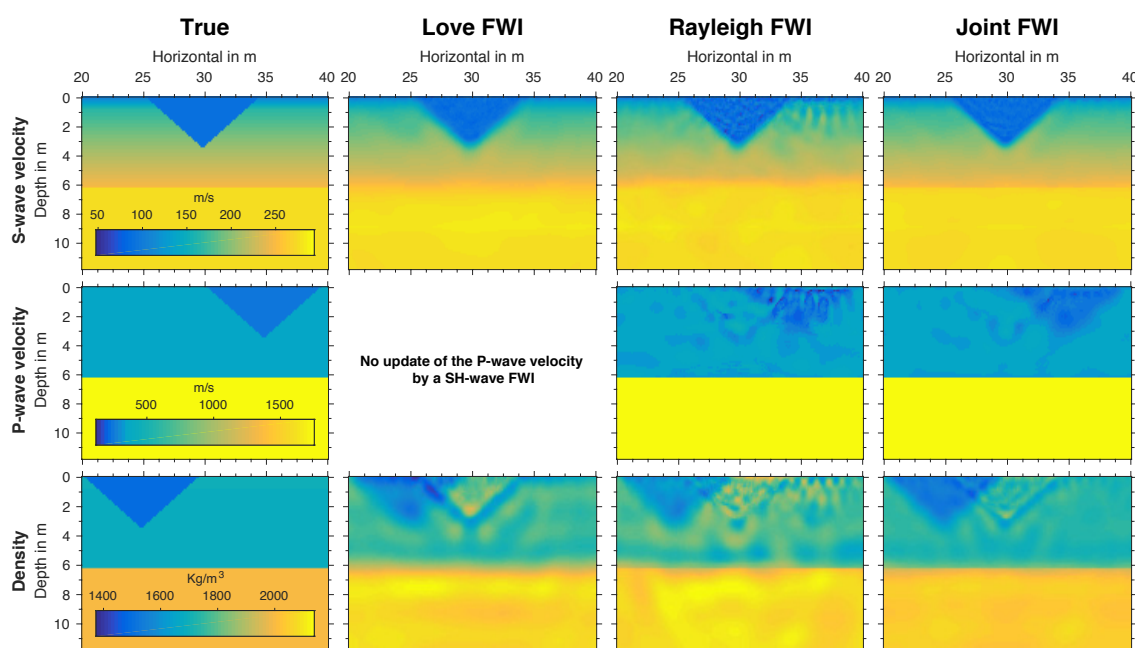
Figure 2: Inversions results of the synthetic reconstruction experiments. Comparison of the true subsurface model (first column) with the reconstructed model of individual Love wave FWI (second column), individual Rayleigh wave FWI (third column) and simultaneous joint FWI (fourth column). The elastic parameters v_S, v_P and ρ are shown row-wise from top to bottom, respectively.

of 9 m we used a constant gradient of 155 $\frac{m}{s \cdot m}$. 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 similar acquisition geometry in the field data application, this case study could help to interpret the reliability and potential artifacts in the field data FWI.

We present the result of this case study in figure 4, where we focus on an area around the trench positions. 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. 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.

We assume that these 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 receivers and 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

(a) Comparison of the seismograms for the individual Rayleigh wave FWI.

(b) Comparison of the seismograms for the individual Love wave FWI.

(c) Comparison of the seismograms for the simultaneous joint FWI.

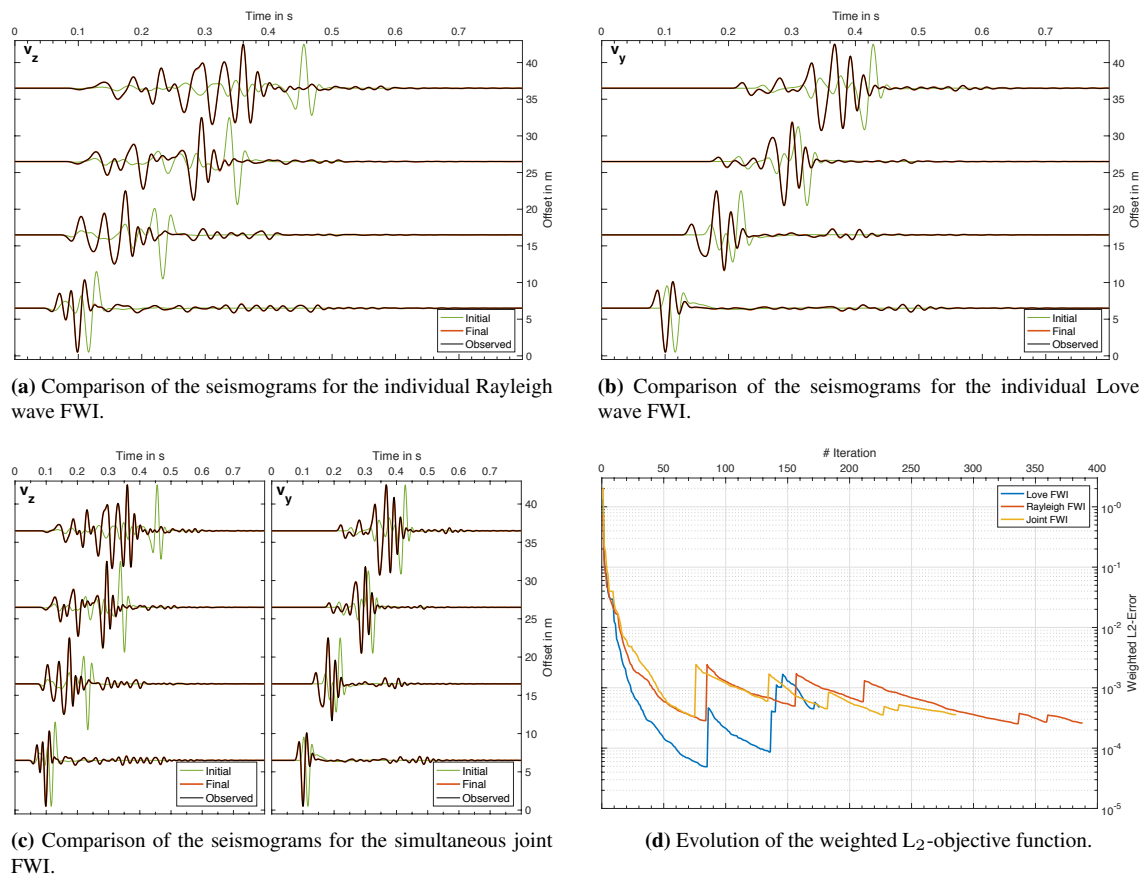**(d)** Evolution of the weighted L_2-objective function.

Figure 3: Data fit of the synthetic reconstruction experiments. Comparison of the seismograms for the initial model, for the final reconstructed model and of the pseudo-observed data set for all three inversions (a-c). Evolution of the objective function over the iterations (d).

Rayleigh FWI   Joint FWI
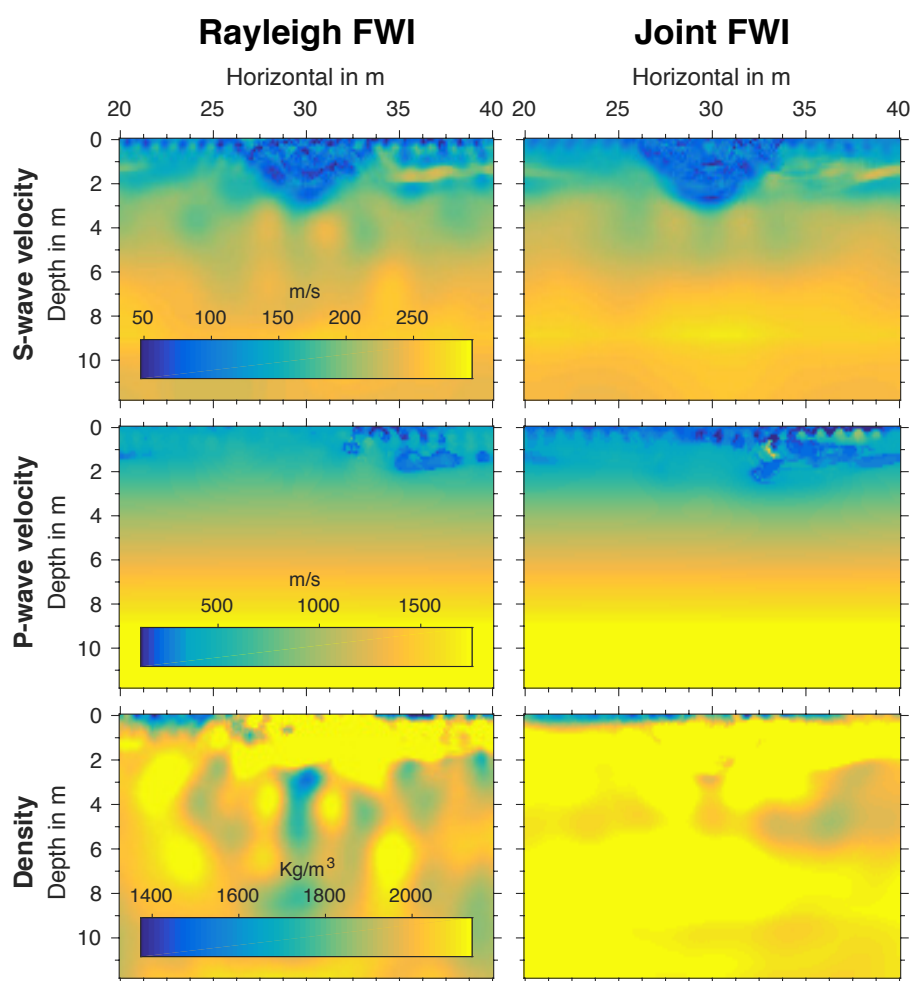
Figure 4: Inversions results of the synthetic reconstruction experiments for an inaccurate initial v_P model. Comparison of the reconstructed models of individual Rayleigh wave FWI (first column) and simultaneous joint FWI (second column). The elastic parameter v_S, v_P and ρ are shown row-wise from top to bottom, respectively.

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.

FIELD DATA APPLICATION

fig In this section, 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 recorded a field dataset on an airfield 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. We assumed such a 2-D subsurface model in the synthetic experiments, 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.

Test site

The location of the test site is on a glider airfield in Rheinstetten near Karlsruhe (Germany). Figure 5 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 fluviatile 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. 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.

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 5. 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. 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.

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, 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). An exemplarily shot gather of the preprocessed dataset is shown in figure 6. 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
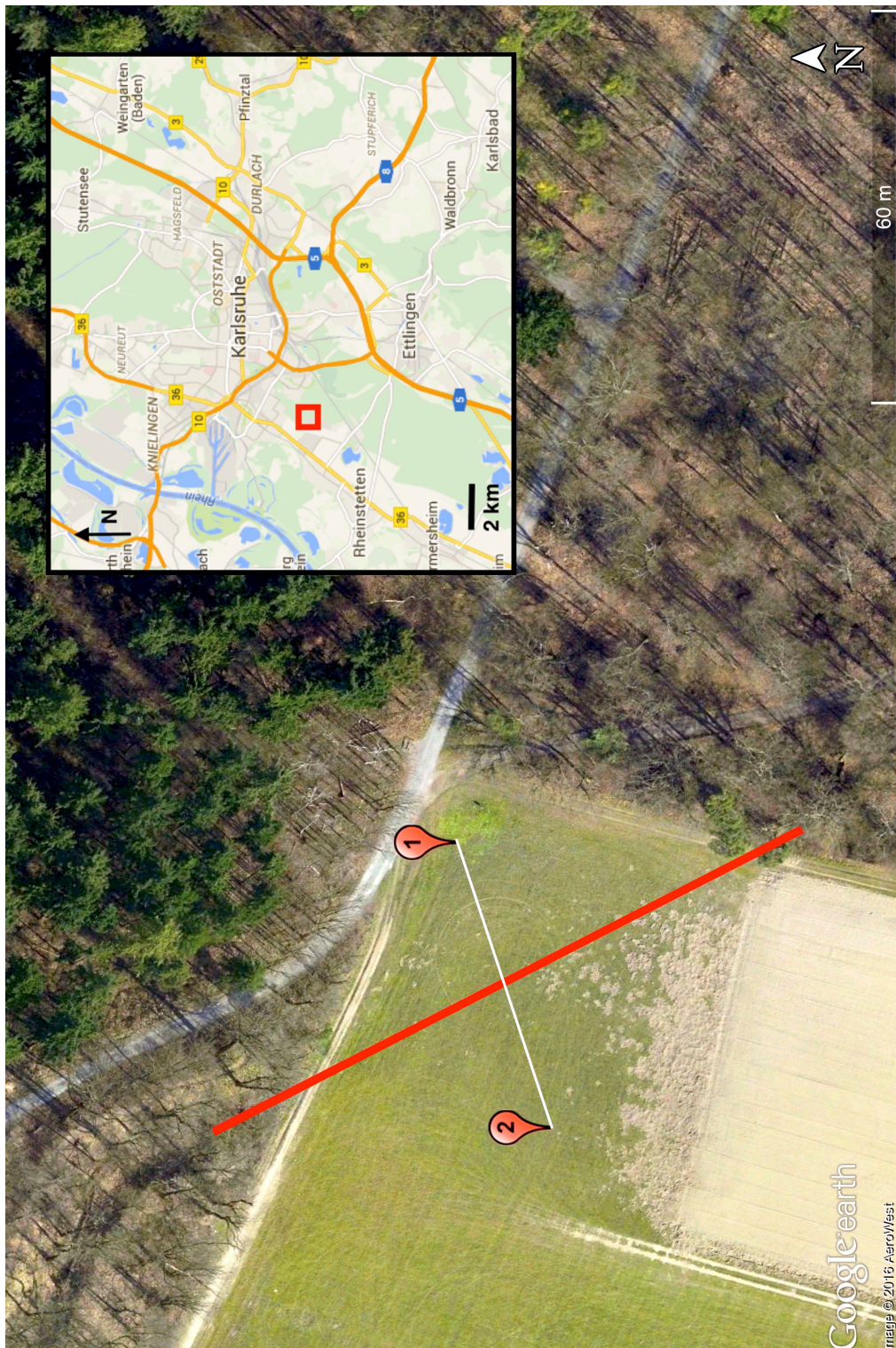
Figure 5: Overview map of the test site. The red line corresponds to the interpolated course of the "Ettlinger Linie" and the white line denotes the acquisition profile. Source: Google Earth (AeroWest, GeoBasis-DE/BKG).

we use them in the FWI. 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 seismogram, v_z, is 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 SH seismogram, v_y, is 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. After the preprocessing steps that we carried out the P-SV as well as the SH field dataset allow for the full-waveform inversion.

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 correspondents 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:

$$\rho = 0.31 \cdot v_P^{0.25}, \tag{23}$$

which assumes that v_P is given in units of m/s for a resulting ρ value in units of g/cm³ (Gardner et al., 1974). The obtained ρ model has a density of 1.325 g/cm³ in the upper layer and 2.142 g/cm³ in the lower half space. We obtain a first initial v_S model based on the minimum and maximum apparent phase velocities of the fundamental modes of the surface waves, which we derive from dispersion spectra. Based on this two values, we created a simple gradient model. 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 7 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 can partly compensate a to high Q-value. Taking this consideration, we set the Q-value to 15. We construct the attenuation model for the viscoelastic forward modelling by a Generalize Standard Linear Solid (GSLS) with three relaxation mechanisms. We follow Bohlen (2002) to calculate the relaxation frequencies to $f_1 = 0.2978$ Hz, $f_2 = 6.7325$ Hz and $f_3 = 84.6014$ Hz and the τ-value to 0.1576, where we use a reference frequency of 40 Hz. We show the final initial model for the three elastic parameters in figure 8.
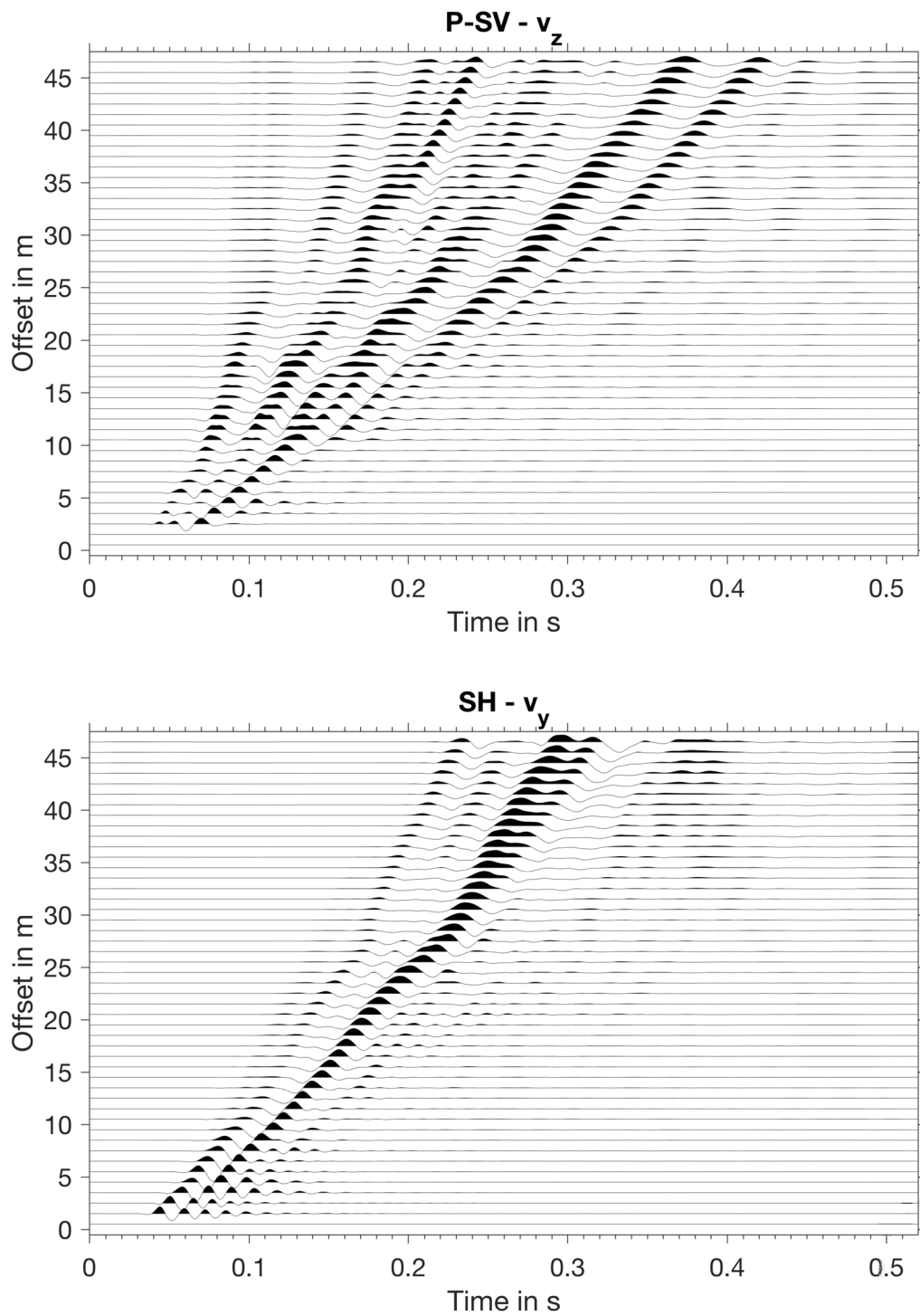
Figure 6: Trace-wise normalized shot-gathers of the preprocessed field dataset for the first source located at the profile coordinate of 10.5 m. The seismograms are band-pass filtered between 4 and 130 Hz, representing the frequency range used in the FWI. Traces near the source were overdriven and thus muted.
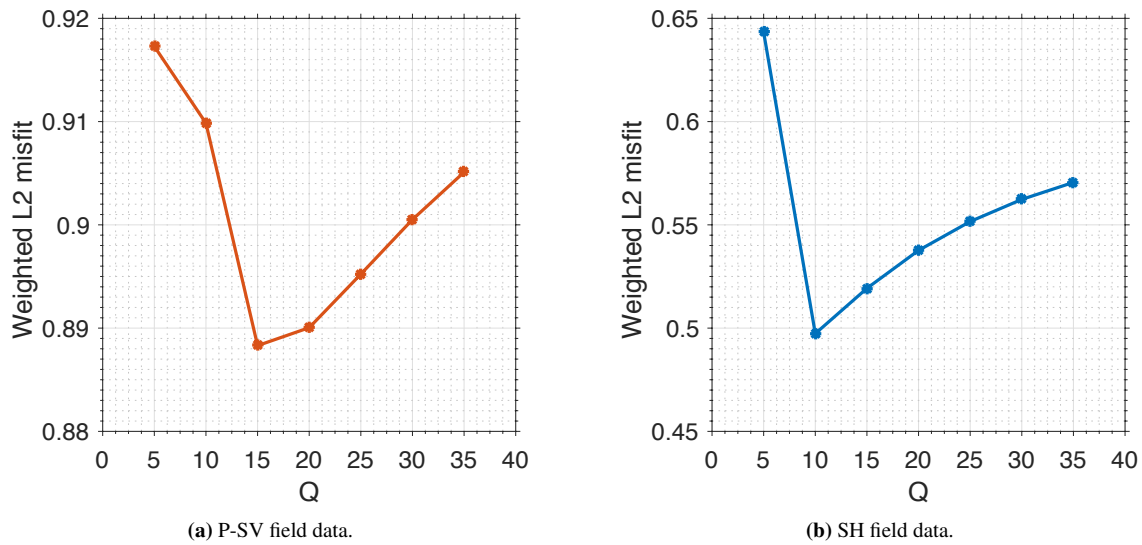
(a) P-SV field data.

(b) SH field data.

Figure 7: 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.

FWI setup

For the field data application we use a similar configuration as for the synthetic FWI experiments, thus, we stick with a brief discussion here.

Objective function   As objective function we again use the weighted L_2-error (equation 17) 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 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.

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 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 5) and of the last receiver at 57 m. The total propagation time of 0.52 s is identical to the time length of the observed seismograms.

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

Table 2: Workflow used in the field data FWI. Each stage is applied to the inversion until the automatic abort criterion AC (equation 22) 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 wether 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

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.

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 21) 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. One reason could be that the estimated initial size of the Hessian (scaled identity matrix as suggested by Nocedal and Wright (2006)) 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.

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 2.

Results

In the following, we discuss the result of each inversion individually. In figure 8 we show a direct comparison of the final inversion results as well as the initial model, where we focus on an area around the trench. Figure 10(a)-(c) shows a comparison of seismograms of the inversion result, of the observed data set as well as for the initial model,.

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 10(d). 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 9(b) 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.

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. 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 10(d). 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
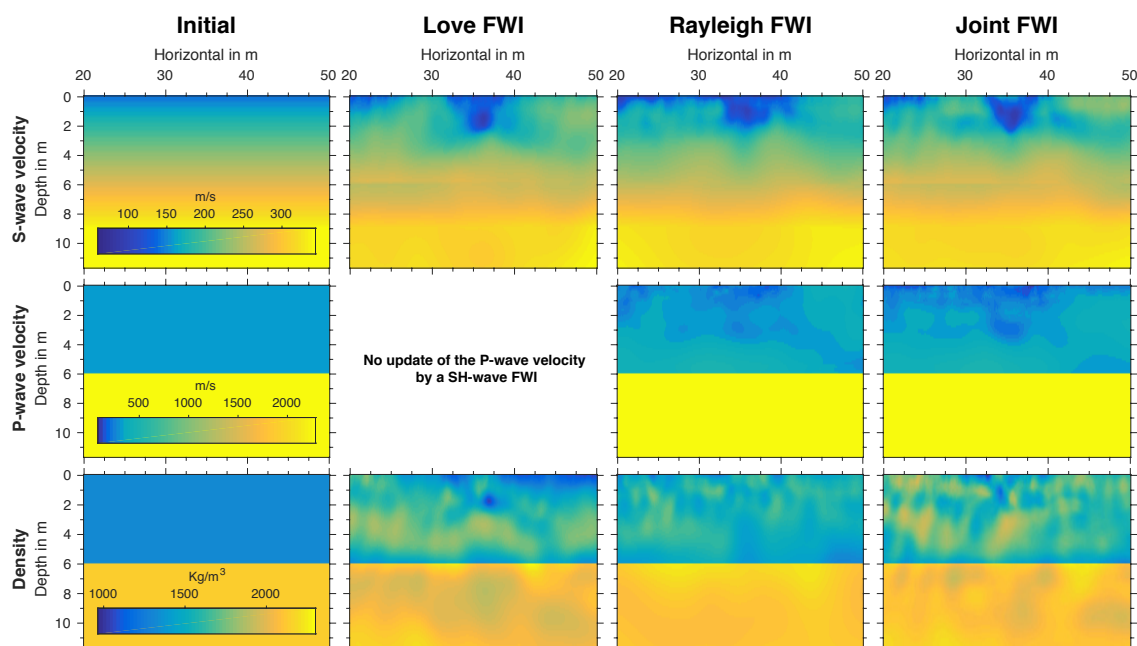
Figure 8: Inversions results of the field data application. Comparison of the initial model (first column) with the final model of the individual Love wave FWI (second column), the individual Rayleigh wave FWI (third column) and the simultaneous joint FWI (fourth column). The elastic parameters v_S, v_P and ρ are shown row-wise from top to bottom, respectively.

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 9(a).

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 10(d) 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 9(a,b) 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.
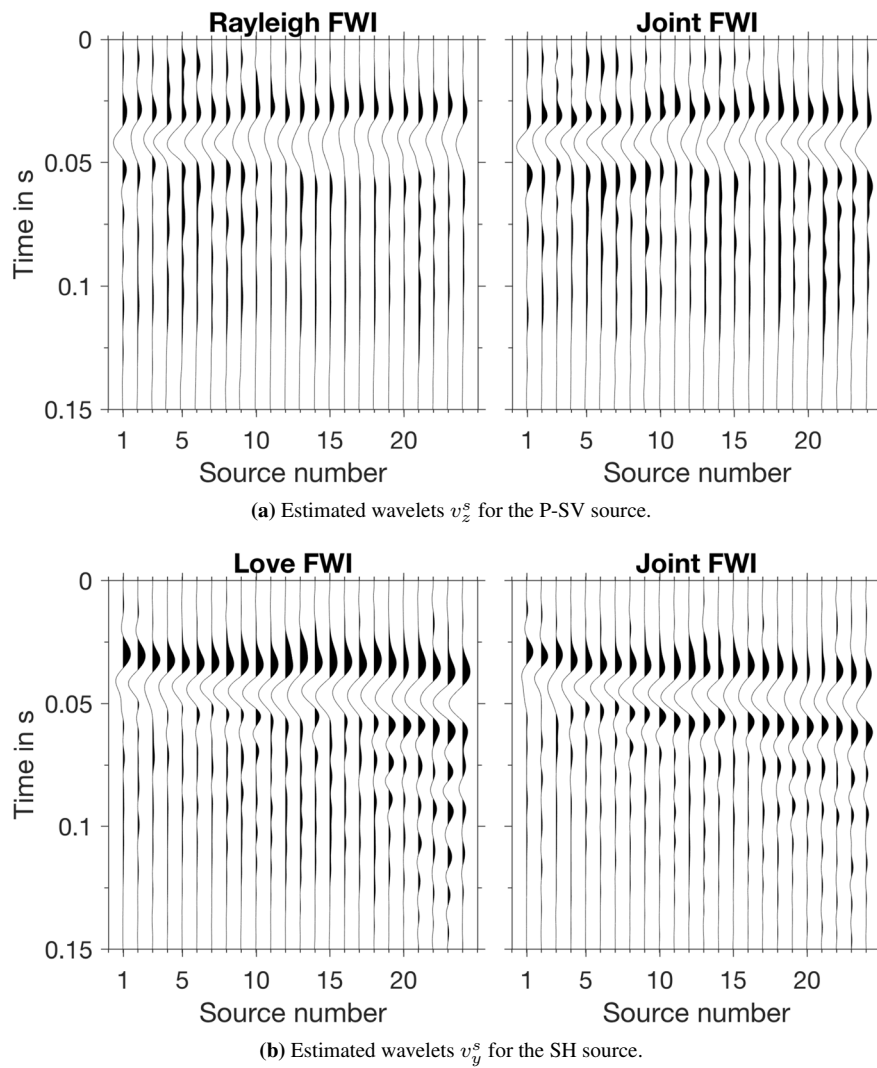
(a) Estimated wavelets v_z^s for the P-SV source.

(b) Estimated wavelets v_y^s for the SH source.

Figure 9: Estimated source time functions for the P-SV sources and the SH sources for the final model at 130 Hz of the individual wave type inversions (first column) and of the simultaneous joint FWI (second column). The P-SV source wavelets, v_z^s, are shown in the (a) and the SH source wavelets, v_y^s, in (b).
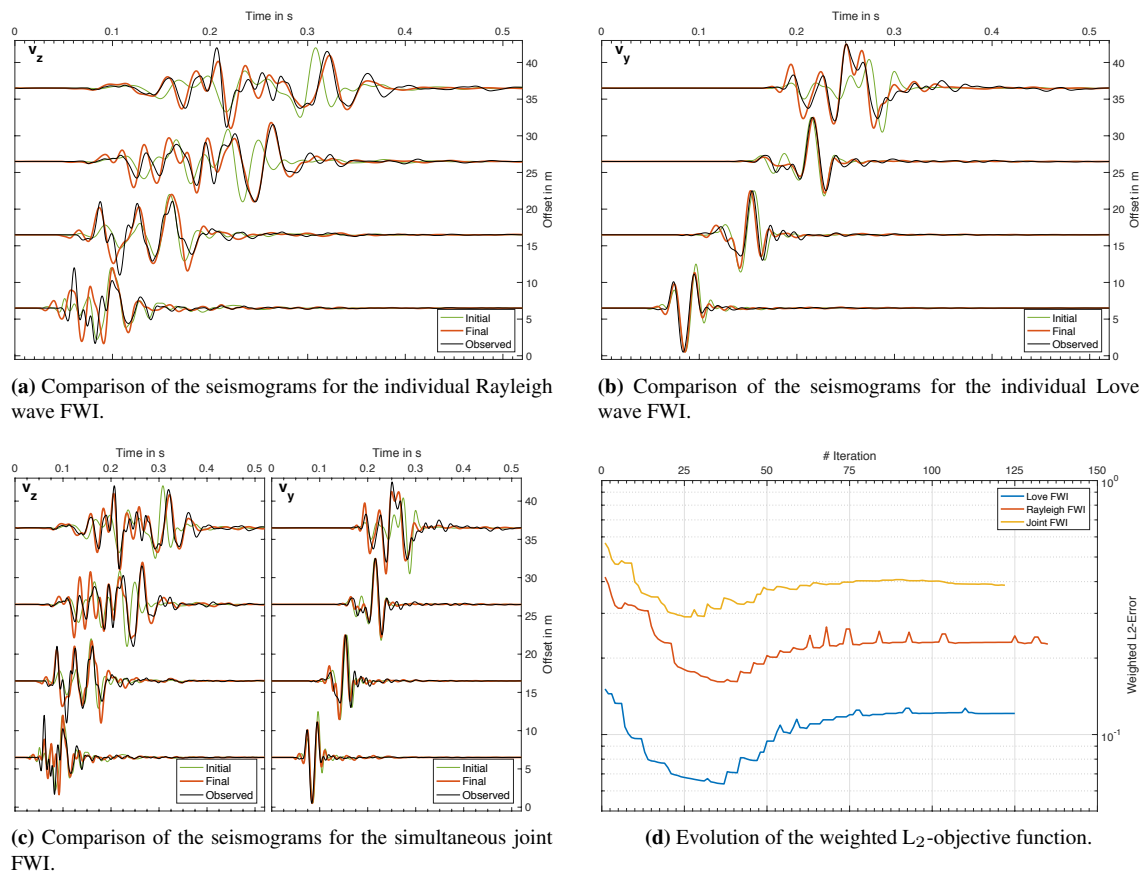
(a) Comparison of the seismograms for the individual Rayleigh wave FWI.

(b) Comparison of the seismograms for the individual Love wave FWI.

(c) Comparison of the seismograms for the simultaneous joint FWI.

(d) Evolution of the weighted L_2-objective function.

Figure 10: Data fit of the field data inversions. Comparison of the seismograms for the initial model, for the final model and of the observed data set for all three inversions (a-c). Evolution of the objective function over the iterations (d).

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. To overcome this limitation we compare the seismic FWI results obtained in this work with the result of a ground-penetrating radar (GPR) measurement.

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 5), however, since they had to relocated 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 11, 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.

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.

The qualitative comparison of the GPR image to the three FWI velocity models is shown in figure 11 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 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.

CONCLUSIONS

To conclude, 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.
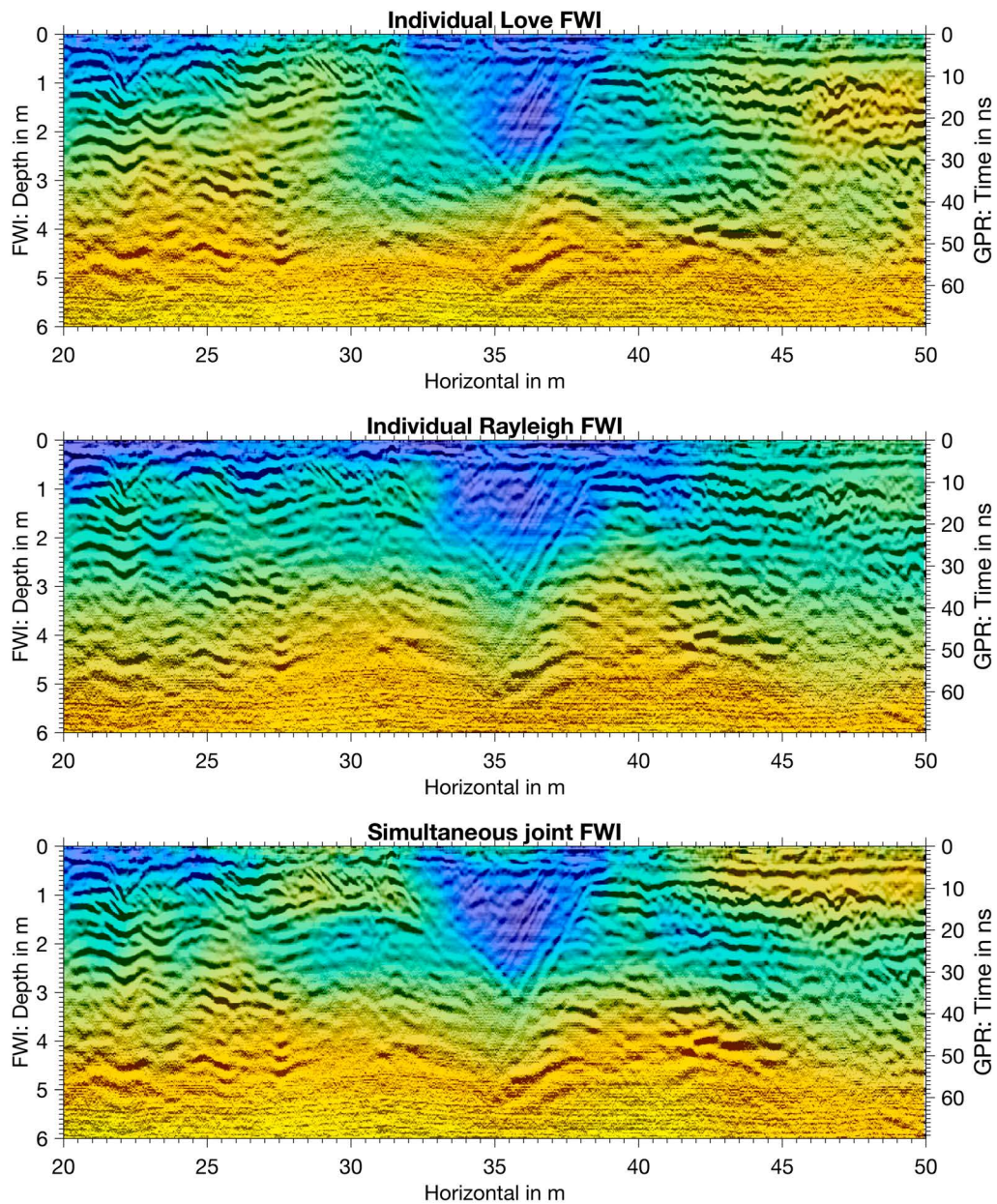
Figure 11: Qualitative comparison of the GPR result with the field data FWI results. Overlay of the final S-wave velocity models of individual Love wave FWI (top), individual Rayleigh wave FWI (middle) and simultaneous joint FWI (bottom) with the time-migrated image of the GPR measurement. A color-bar is not shown, because the transparency effect would falsify the color representation.

REFERENCES

Binnig, M. (2015). Full Waveform Inversion of shallow-seismic Rayleigh waves to characterize the "Ettlinger Linie". Master's thesis, Karlsruher Institut für Technologie (KIT).

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. (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.

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.

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*.

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.

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.

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.

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.

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.

Safani, 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).

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.

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. (2016). Individual and joint 2-d elastic full-waveform inversion of rayleigh and love waves. Master's thesis, Karlsruhe Institute of Technology (KIT), http://dx.doi.org/10.5445/IR/1000058955.

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.