Full-waveform inversion of ground-penetrating radar data and its indirect joint petrophysical inversion with shallow-seismic data

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Tan Qin

Tag der mündlichen Prüfung: 04.11.2022

Referent: Prof. Dr. Thomas Bohlen
Korreferentin: Prof. Dr. Anja Klotzsche

Abstract

Both full-waveform inversion (FWI) of ground-penetrating radar (GPR) and shallow-seismic data have received special attention in the past decade because they allow the reconstruction of seismic and electromagnetic (EM) properties at high resolution. Research on the two FWIs includes: FWI of single geophysical data and joint FWI of multiple geophysical data. In this dissertation, I focus on GPR FWI in the former and joint petrophysical inversion (JPI) in the latter.

In GPR FWI, the high computational costs and frequency-independent assumption are two problems that limit its development. To reduce computational costs, I apply a subset FWI (SFWI) to multi-offset GPR data. SFWI uses the data simulated on a model subset to approximate the data simulated on an entire model. Thus it obtains theoretical speedup and memory saving factor equal to the size ratio of the model and its subset. SFWI has higher or lower than expected speedups when combined with the source parallelization and model domain parallelization, respectively. The model subset depends on the illumination of the acquisition geometry used, for which I give rules of thumb by quantifying its effect on the simulation and inversion. Both 2-D synthetic and field data validate that SFWI provides results comparable to FWI but requires much lower computational costs than FWI. This study makes FWI an affordable technique for general users and promotes its application in addressing near-surface problems.

The second problem means that dielectric permittivity and electrical conductivity are supposed to be frequency independent in most GPR FWI, which may lead to false estimates if they strongly depend on frequency. I develop frequency-dependent GPR FWI to solve this problem. Using the τ-method introduced from the seismic community, I define the permittivity attenuation parameter to quantify the attenuation resulting from the complex permittivity. The new parameter acts as a low-pass filter, distorting the waveform and decaying the amplitude of the EM waves. The use of permittivity attenuation reduces the number of reconstructed parameters in frequency-dependent GPR FWI. The 2-D synthetic examples illustrate that permittivity attenuation is necessary for reconstructing permittivity and conductivity models in frequency-dependent media. The 2-D field example shows that frequency-dependent GPR FWI provides a better fit to the observed data and a more robust conductivity reconstruction in a high permittivity attenuation environment than frequency-independent GPR FWI. This research greatly expands the application of GPR FWI in more complicated media.

Shallow-seismic and GPR FWI can provide complementary information for each other. Based on the sensitivity difference of the two data to petrophysical parameters, I propose indirect JPI, where seismic data are used for porosity reconstruction and GPR data are used for saturation reconstruction. Unlike conventional JPI, I first update the seismic and GPR parameters using non-petrophysical parametrizations and then transform the most reliable estimates to petrophysical parameters. The 2-D synthetic tests show

that indirect JPI can provide more accurate and consistent results than conventional JPI. In addition, due to the rational use of the sensitivity of geophysical data to parameters, indirect JPI is more robust when uncertainties exist in petrophysical *a priori* knowledge. More importantly, indirect JPI is flexible to integrate different types of seismic and EM waves and acquisition geometries depending on the target of interest, resulting in the best solution. Indirect JPI has been proven to be a promising approach for multiparameter reconstructions.

To validate if indirect JPI can solve the real problem, I apply it for the first time to Love-wave and multi-offset surface GPR field data. It provides consistent imaging of near-surface targets with good accuracy by estimating seismic, EM, and petrophysical models. The inversion results are validated by direct-push technology and borehole measurements. This application suggests that indirect JPI can avoid conflicting geological interpretations that may arise in individual inversions and shows higher efficiency of information exchange than joint structural inversion. Furthermore, this method is robust with different petrophysical initial models and coefficients, for instance, Archie's coefficients. The study verifies the feasibility of using indirect JPI to invert multiple geophysical field data and promotes the broader applicability of petrophysical methods.

In summary, this dissertation (1) reduces the computational costs of GPR FWI and extends GPR FWI to frequency-dependent media, and (2) proposes a new joint strategy to combine GPR and shallow-seismic data and validates its performance through 2-D synthetic and field examples.

Contents

1 Introduction

The main content of this dissertation consist of two published journal papers (Chapters 3 and 4) and two manuscripts prepared for submission (Chapters 2 and 5). Depending on the requirements of journals, I use different writing styles, such as British and American English. In the following, I briefly describe the motivations and main goals for these studies, since Section 'Introduction' in each chapter already includes the state of the art.

1.1 Motivation

Ground-penetrating radar (GPR) is a widely used geophysical method for exploring the Earth's shallow critical zone. Full-waveform inversion (FWI) of GPR data has received particular attention in the past decade because it can provide high-resolution subsurface models of dielectric permittivity and electrical conductivity. Up to now, GPR FWI still faces the following challenges:

- High computational costs: FWI is computationally expensive, mainly due to that several simulations of the full wavefield are required in each inversion iteration. It makes most of GPR FWI to be performed in a 2-D world and on high-performance computer. If the investigation area is large or many sources are needed, GPR FWI may become unaffordable for general users.
- High attenuation of electromagnetic (EM) waves: GPR emits and records EM waves propagating in the subsurface. The signal attenuates rapidly in conductive environments like clay. This limits the penetration depth of GPR signal and reduce the useful information that can be retrieved from GPR data.
- Frequency-independent assumption: In most GPR FWI, permittivity and conductivity are treated as constants over the main GPR bandwidth. In fact, many typical geological materials, especially in shallow weathered zones, are frequency-dependent. Frequency-independent assumption may not show the full potential of GPR FWI or may even lead to incorrect estimates if the frequency dependence is strong.
- Uncertainty of the inverse problem: FWI is a typical inverse problem that highly depends on the initial model and easily drops to the local minima. This uncertainty poses a significant challenge to data interpretation, which becomes worse when encountering the inherent limitations of GPR data.

To reduce uncertainty, one can use several geophysical methods to investigate the same targets because they can provide complementary information for each other. In this study, I focus on GPR and shallow-seismic method. The GPR signal is sensitive to water content, but the information it carries is limited by the subsurface environment and antenna frequency; shallow seismic data have a high signal-to-noise ratio and can detect the hardness of materials, but are difficult to determine changes in water. Independent inversions of these data may provide inconsistent interpretation due to the non-uniqueness of the inverse problem and the inherent limitations of these geophysical methods. Hence it is necessary to develop the joint inversion to make the best use of the information from different geophysical data.

The joint inversion can be mainly divided into two classes: joint structural inversion (JSI) and joint petrophysical inversion (JPI). They have the following challenges:

- Weak constraints of JSI: JSI exploits the structural similarity of different models. On the one hand, this is very advantageous for the wide application of JSI, on the other hand, it is a shortcoming because it uses only structural information and ignores other possible physical connections. For regions where the variation of one model is slight, it is difficult for JSI to improve the reconstruction of other coupled models.
- High dependence of JPI on *a priori* information: Compared to JSI, JPI has stronger constraint because it uses petrophysical relations to connect petrophysical properties and geophysical parameters. However, if the petrophysical relations or rock matrix parameters are inaccurate, the inversion results become poor. This makes JPI a site-dependent technique that is hard to apply to a wide range of areas.

The above challenges limit the development of GPR FWI and joint inversion and need to be solved urgently. This motivates the study of the dissertation.

1.2 Overview

The dissertation is organized as follows:

- Chapter 2: How to reduce the computational burden of GPR FWI? I use the model-dependent method to change the disadvantage of high EM wave attenuation to the advantage of speeding up the inversion and saving memory.
- Chapter 3: How to make GPR FWI more widely applicable and better feasible? I develop frequency-dependent GPR FWI to account for velocity dispersion and amplitude attenuation caused by complex permittivity.
- Chapter 4: How to make the best use of GPR and shallow-seismic data in joint inversion? I evaluate the sensitivities of these data to petrophysical parameters and propose an indirect JPI to make the joint inversion more robust when uncertainties appear in *a priori* information.

- Chapter 5: Is indirect JPI applicable to the real world? To validate this, I provide an example of using this method to solve a field problem and present its superior performance over individual inversions and JSI.
- Chapter 6: Conclusions of the full dissertation.

2 Application of subset full-waveform inversion to multi-offset ground-penetrating radar field data

Prepared as: Qin, T., Bohlen, T., Allroggen, N., & Klotzsche, A. Application of subset full-waveform inversion to multi-offset ground-penetrating radar field data. *GEOPHYSICS*.

Abstract

Ground-penetrating radar (GPR) is a widely used geophysical method for exploring the Earth's shallow critical zone. In practice, multi-offset GPR data are usually acquired with either surface or multi-borehole geometries. The recorded data can be analyzed by computationally expensive full-waveform inversion (FWI) algorithms. We apply a subset FWI (SFWI) to multi-offset GPR data to reduce computational costs by using the data simulated on a model subset to approximate the data simulated on an entire model. Due to the high attenuation characteristics of electromagnetic waves in typical soils, the size of the model subset can be much smaller than the model. Thus we obtain theoretical speedup and memory saving factor equal to the size ratio of the model and its subset. The properties of the model subset depend on the illumination of the chosen acquisition geometry. Our study reveals that, due to memory savings, SFWI has higher than expected speedups when using the source parallelization alone. When combined with the source parallelization and model domain parallelization, its performance degrades a little due to the proportional increase in communication time. Both 2D synthetic and field data validate that SFWI provides results comparable to FWI, but without the redundant computation present in FWI. In the field examples, a 45 m long survey line and a 21 m long multi-borehole plane, the SFWI obtain a speedup greater than eight times and three times, respectively. The former can be further accelerated by shortening the source-receiver array, but at a slight sacrifice in reconstruction precision. To sum up, this study makes FWI an affordable technique for general users and promotes its application to address real-world problems.

2.1 Introduction

Detailed characterization of the near-surface region is essential for engineering, hydrological, and archaeological applications. As a commonly used geophysical method, ground-penetrating radar (GPR) plays an important role in near-surface surveys (Jol, 2008). Up to now, most GPR data are still interpreted based on reflector picking and travel-time tomography, which may introduce artificial errors and the resolution may not meet practical needs (Paz et al., 2017). Another interpretation approach is full-waveform

inversion (FWI), which has been shown to have the potential for reconstructing high-resolution subsurface models of electromagnetic (EM) material properties (Klotzsche et al., 2019). However, GPR FWI is computationally expensive compared to other interpretation approaches. Multi-offset GPR data are usually collected at the surface or in multiple boreholes. FWI of these data may become computationally unaffordable for personal computers when the profile is long and many sources have to be modeled. Therefore, reducing computational costs is crucial in applying FWI to invert typical GPR data acquired in field surveys.

The cost of FWI is proportional to the number of sources and the size of the model. Hence, methods used to accelerate FWI can be divided into two main classes: source-dependent method and model-dependent method. The source-dependent method reduces the number of sources used in FWI. The encoded simultaneous-source FWI proposed by Krebs et al. (2009), one of the most popular source-dependent methods, merges n_s gathers into n_s^e supergathers, thus reducing the computational cost by a factor of $n_\mathrm{s}/n_\mathrm{s}^\mathrm{e}$, where n_s and n_s^e are the number of all sources and encoded sources, respectively. Nevertheless, this method requires the fixed-spread geometry for all sources, which is impractical for GPR data. Furthermore, crosstalk introduced by the encoding operation can interfere with the reconstruction results. To suppress the crosstalk and adapt the encoded simultaneous-source FWI to marine streamer seismic data, Choi and Alkhalifah (2012) developed the objective function using the global correlation norm. This objective function normalizes the trace; hence, it is not favorable to amplitude information, which is essential for reconstructing the conductivity from GPR data. To eliminate the crosstalk and acquisition limitation, Huang and Schuster (2012) proposed a frequency-division encoding method and Zhang et al. (2018) used the phase sensitive detection method. Both of them follow the same idea of assigning different frequency components to each source in one encoded source so that the wavefield of each source can be decoded from the wavefield of the encoded source by frequency filtering (Nihei and Li, 2007). However, the number of sources encoded into a supergather depends on the frequency band and the time window. Furthermore, each source contains much less information than a full-band signal. These limitations decrease the performance gain of the source-dependent methods (Bachmann and Tromp, 2020).

The model-dependent method reduces the model size by a factor of S, thus making the computation S times faster, where S is the size ratio of the model and the reduced model. One of the advantages of the model-dependent method compared to the source-dependent method is that it saves memory by S times. In the studies of Gueting et al. (2015), Gueting et al. (2017), and Zhou et al. (2021a), the survey plane is divided into $n_\mathrm{b} - 1$ crosshole planes by n_b boreholes where n_b is the number of boreholes. GPR FWI is implemented separately in each crosshole plane with the data acquired in two boreholes. The final GPR FWI results are obtained by splicing all $n_\mathrm{b} - 1$ crosshole planes. We refer to it as spliced GPR FWI. This method is computationally economical relative to the GPR FWI of an entire survey plane. However, spliced GPR FWI reconstructs each crosshole plane independently, which may lead to different convergences of data fitting and different degrees of model update in adjacent planes (Gueting et al., 2015; Zhou et al., 2021a). Moreover, when n_b is large, spliced GPR FWI becomes cumbersome due to the need to set up the acquisition for $n_\mathrm{b} - 1$ crosshole models. To apply FWI to land streamer seismic data more efficiently, for each source, Tran and Sperry (2018) performed forward modeling on a model subset (referred to as a model segment in their paper) rather than on the entire model. Unlike the spliced

GPR FWI, the method of Tran and Sperry (2018) is simple to operate as it requires only one acquisition setup. Hoang (2021) applied a similar approach to the land streamer seismic data acquired along a 1500 m profile and obtained a 15-fold acceleration with respect to the standard FWI. Unfortunately, neither Tran and Sperry (2018) nor Hoang (2021) offered an in-depth analysis of the principles for these methods and the selection of the model subset. On the other hand, in the 3D GPR forward modeling study of Koyan and Tronicke (2020), the authors examined the effects of the distance between the source-receiver array and the subset boundary on the simulated data, which is a first approach for quantifying the applicability and limitations of model-dependent methods to accelerate FWI.

In this paper, we adopt the subset FWI (SFWI) method, a model-dependent method introduced by Tran and Sperry (2018) in the shallow-seismic study, to process multi-offset GPR field data. Our contribution consists of two parts. First, we clarify the principles of SFWI and give rules of thumb for selecting the model subset by quantifying the performance difference between 2D SFWI and FWI. Second, we apply 2D SFWI for the first time to GPR field data acquired from different geometries. Up to now, GPR FWI has primarily been applied to crosshole radargrams and is rarely used to invert multi-offset surface GPR data (Klotzsche et al., 2019). The latter is more challenging because of reduced illumination, differences in sensitivity kernels and higher requirements for initial model building (Lavoué et al., 2014). In practice, however, surface data acquisition is faster to perform and does not require boreholes. It is therefore critical for computationally expensive tasks, such as 3D imaging. For these reasons, we evaluate SFWI's performance primarily using multi-offset surface GPR data and using crosshole GPR data only at the end of Section 2.4.

2.2 Methodology

2.2.1 Full-waveform inversion

The objective function Φ that we use in GPR FWI is a sum of the individual objective functions Φ_i

$$\Phi(\mathbf{m}) = \sum_{i=1}^{n_s} \Phi_i(\mathbf{m}), \quad \Phi_i(\mathbf{m}) = \frac{1}{2}||\mathbf{d}_i^{\text{syn}}(\mathbf{m}) - \mathbf{d}_i^{\text{obs}}||_2^2, \tag{2.1}$$

where $\mathbf{d}_i^{\text{syn}}$ and $\mathbf{d}_i^{\text{obs}}$ are the synthetic data and observed data of the ith source, respectively; $\mathbf{m}$ is the model parameters, i.e., dielectric permittivity ε and electric conductivity σ; n_s is the number of sources and $||.||_2$ represents the l_2 norm. We use the finite-difference time-domain (FDTD) method to simulate the propagation of the EM wavefields (Yee, 1966) and use the convolutional perfectly matched layer (CPML) to absorb the outgoing waves (Roden and Gedney, 2000). GPR FWI uses the adjoint-state method to calculate the derivative of $\Phi_i(\mathbf{m})$ with respect to $\mathbf{m}$ and sums it up to form an overall model gradient $\partial\Phi(\mathbf{m})/\partial\mathbf{m}$ as follows (Plessix, 2006; Meles et al., 2010):

$$\frac{\partial\Phi(\mathbf{m})}{\partial\mathbf{m}} = \sum_{i=1}^{n_s} \frac{\partial\Phi_i(\mathbf{m})}{\partial\mathbf{m}}. \tag{2.2}$$

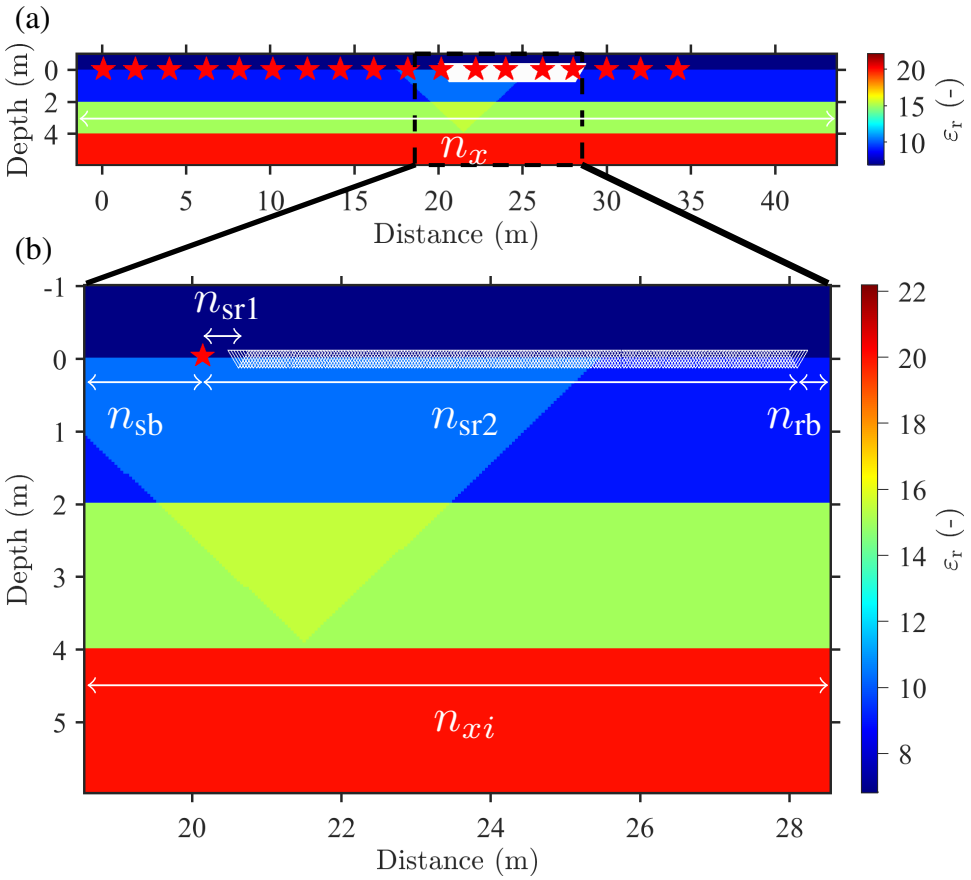

Figure 2.1: (a) The model $\mathbf{m}$ (1130×175 grids with grid spacing 0.04 m) and (b) the model subset $\mathbf{m}_i$ (250×175 grids with the same grid spacing). The black dashed rectangle in the model outlines the extraction area. The red stars are sources, and the white triangles are the receivers of the ith source. n_x and n_{xi} illustrate the widths of the model and its subset, respectively; n_{sb} and n_{rb} are the source-boundary and receiver-boundary distances, respectively; n_{sr1} and n_{sr2} are the minimum and maximum source-receiver offsets, respectively. The relative permittivity ε_{r} of the air is one (depth < 0 m).

The gradient is then applied in the preconditioned conjugate-gradient method to iteratively update the model parameters $\mathbf{m}$ (Qin et al., 2022b). Equation 2.2 demonstrates that $\partial\Phi(\mathbf{m})/\partial\mathbf{m}$ needs to be computed over the entire model space, which results in a high computational cost for GPR FWI if the model size is large. In the following, we show how to use a model-dependent method, i.e., SFWI, to solve this computational problem.

2.2.2 Subset full-waveform inversion

We build a model $\mathbf{m}$ in Figure 2.1a and use a 'walk-away' method to acquire multi-offset surface GPR data on one side of the source (Domenzain et al., 2021). Given the high attenuation of EM waves propagating in the subsurface, the reflected waves fall below the noise level when the transmitter-receiver offset exceeds a certain distance, depending on the signal-to-noise (S/N) ratio. Thus the source-receiver

array of the ith source illuminates only a fraction of the model (Meles et al., 2011), e.g., the model subset $\mathbf{m}_i$ in Figure 2.1b, which we can extract from the model by

$$\mathbf{m}_i = \mathbf{A}_i \mathbf{m}, \tag{2.3}$$

where $\mathbf{A}_i$ is an $n_{mi} \times n_m$ matrix for transformation. n_m and n_{mi} are the size (i.e., number of grids) of $\mathbf{m}$ and $\mathbf{m}_i$, respectively. Each row of $\mathbf{A}_i$ consists of a one corresponding to the coordinate index of one grid in $\mathbf{m}_i$ and $n_m - 1$ zeros elsewhere (Hoang, 2021). The matrix has only n_{mi} ones and is therefore highly sparse.

In SFWI, we simulate the EM wavefields on $\mathbf{m}_i$ and obtain the data $\mathbf{d}_i^{\text{syn}}(\mathbf{m}_i)$ with the same acquisition geometry as $\mathbf{d}_i^{\text{syn}}(\mathbf{m})$. The corresponding objective function is expressed as below:

$$\Phi(\mathbf{m}_i) = \sum_{i=1}^{n_s} \Phi_i(\mathbf{m}_i), \quad \Phi_i(\mathbf{m}_i) = \frac{1}{2} ||\mathbf{d}_i^{\text{syn}}(\mathbf{m}_i) - \mathbf{d}_i^{\text{obs}}||_2^2. \tag{2.4}$$

We use the adjoint-state method to compute $\partial\Phi_i(\mathbf{m}_i)/\partial\mathbf{m}_i$, the gradient of the model subset, and form a padded gradient $\partial\Phi_i(\mathbf{m}_i)/\partial\mathbf{m}$ by transforming the gradient subset back to its position in the entire model as follows:

$$\frac{\partial\Phi_i(\mathbf{m}_i)}{\partial\mathbf{m}} = \mathbf{A}_i^T \frac{\partial\Phi_i(\mathbf{m}_i)}{\partial\mathbf{m}_i}, \tag{2.5}$$

where the superscript ^{T} is the transpose operator. Equation 2.5 means that $\partial\Phi_i(\mathbf{m}_i)/\partial\mathbf{m}$ has the same size as $\mathbf{m}$ and is the result of padding $\partial\Phi_i(\mathbf{m}_i)/\partial\mathbf{m}_i$ to zero in areas outside the model subset (Tran and Sperry, 2018). If the padding areas have a negligible impact on the data of the ith source, we can approximate the data of the ith source simulated on $\mathbf{m}$ with that on $\mathbf{m}_i$, that is

$$\mathbf{d}_i^{\text{syn}}(\mathbf{m}) \approx \mathbf{d}_i^{\text{syn}}(\mathbf{m}_i). \tag{2.6}$$

Then we get $\Phi_i(\mathbf{m}) \approx \Phi_i(\mathbf{m}_i)$, $\Phi(\mathbf{m}) \approx \Phi(\mathbf{m}_i)$, and $\partial\Phi(\mathbf{m})/\partial\mathbf{m} \approx \partial\Phi(\mathbf{m}_i)/\partial\mathbf{m}$. In such a case, we approximate the gradient of the entire model $\partial\Phi(\mathbf{m})/\partial\mathbf{m}$ by summing the padded gradients of all model subsets as follows:

$$\frac{\partial\Phi(\mathbf{m})}{\partial\mathbf{m}} \approx \sum_{i=1}^{n_s} \frac{\partial\Phi_i(\mathbf{m}_i)}{\partial\mathbf{m}}. \tag{2.7}$$

We summarize the workflow of SFWI in Figure 2.2. The model $\mathbf{m}$ is split into n_s model subsets (equation 2.3), each of which should satisfy the data approximation (equation 2.6). For the convenience of parallelization, we define all model subsets of the same spatial size. The gradient corresponding to each source is calculated within the model subset $\mathbf{m}_i$ (equation 2.5) and then merged to form an approximated gradient (equation 2.7). We use the approximated gradient in the optimization for iteratively updating $\mathbf{m}$ (Qin et al., 2022b).

Our workflow differs from that of Hoang (2021) whose model subset is updated by its gradient $\partial\Phi_i(\mathbf{m}_i)/\partial\mathbf{m}_i$ directly. The workflow of Hoang (2021) violates the requirement of source parallelization, as different

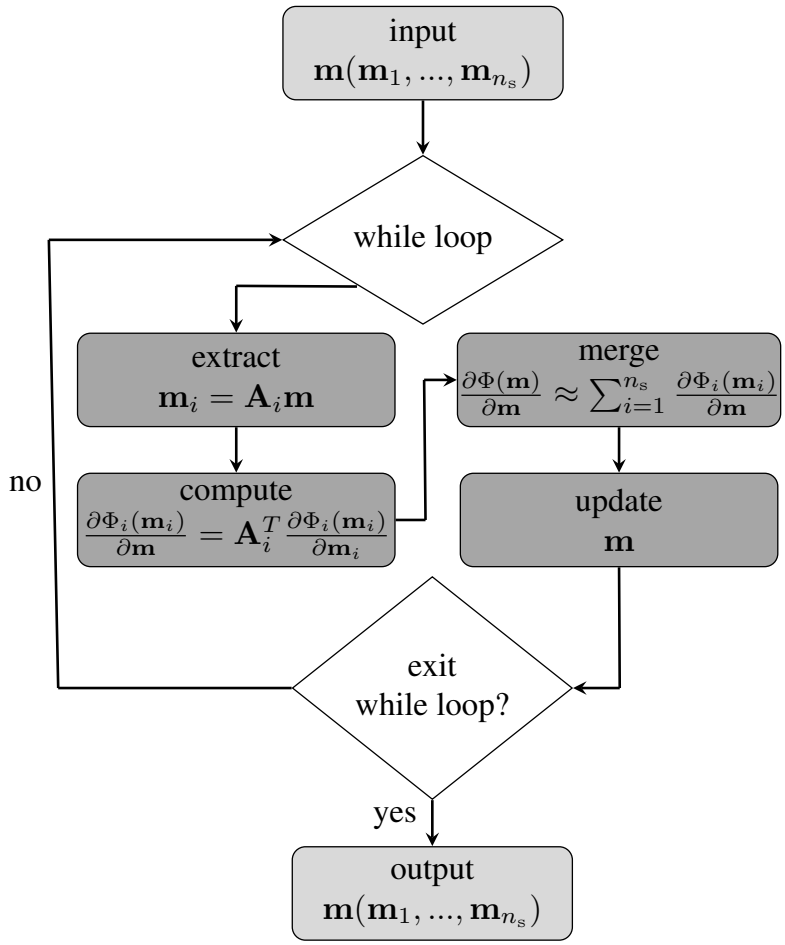

Figure 2.2: The subset full-waveform inversion (SFWI) workflow for multi-offset ground penetrating radar (GPR) data. The model $\mathbf{m}$ can be split into n_s model subsets, i.e., $\mathbf{m}_1$, ..., $\mathbf{m}_{n_s}$, where n_s is the number of sources.

model subsets may overlap each other and cannot be changed simultaneously in the inversion. We overcome this drawback by using the merged gradient of all sources (equation 2.7) to update the model. In this way, SFWI has the same compatibility as conventional FWI in terms of source parallelization, which is an important technique for high-performance computation.

As can be seen from the workflow (Figure 2.2), SFWI is very similar to FWI, except for the extra spatial transformation operations. Since the transformation matrix $\mathbf{A}_i$ is highly sparse, the extraction of model subset (equation 2.3) and the padding of gradient subset (equation 2.5) impose a negligible burden on SFWI. Therefore, the computational costs of gradient calculation for the ith source in SFWI depend on $\mathbf{m}_i$ not $\mathbf{m}$. The data approximation (equation 2.6) also applies for source-wavelet correction (Ernst et al., 2007a; Groos et al., 2014) and step length estimation (Meles et al., 2010). As a result, SFWI reduces all simulation-related tasks which are the most computational expensive part in FWI. Theoretically, in each iteration, SFWI obtains a speedup and memory saving factor equal to the size ratio of the model and its subset:

$$S = n_m / n_{mi}. \tag{2.8}$$

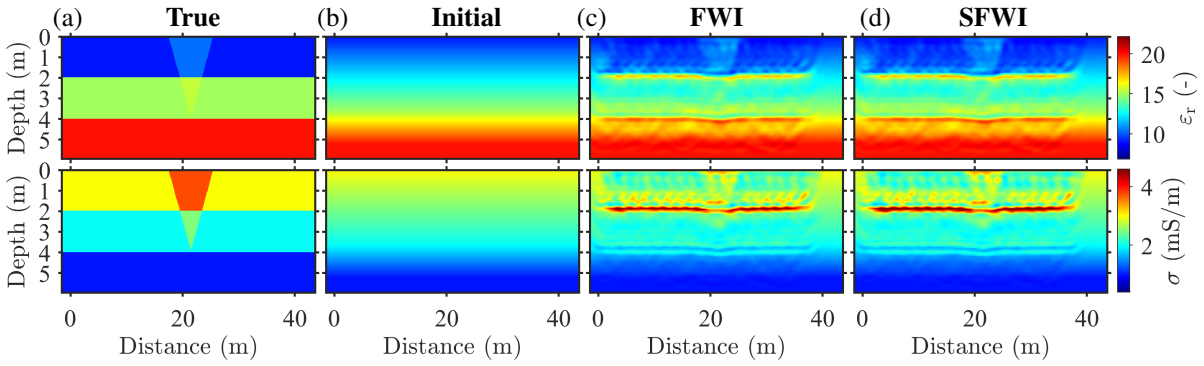

Figure 2.3: Models of the synthetic example, showing (a) the true models, (b) the initial models, (c) the reconstructed models of full-waveform inversion (FWI), and (d) the reconstructed models of SFWI.

In the 2D case (Figure 2.1), since that $\mathbf{m}$ and $\mathbf{m}_i$ have the same length in z direction (Cartesian coordinate system), we have the speedup factor $S^x_{2D} = n_x/n_{xi}$, where n_x and n_{xi} are the lengths of $\mathbf{m}$ and $\mathbf{m}_i$ in x direction.

2.3 Synthetic examples

The key point of our approach is that the data approximation (equation 2.6) must be satisfied, depending on the selection of the model subset, which in the 2D case is the selection of n_{xi}. As shown in Figure 2.1b, $n_{xi} = n_{sb} + n_{sr2} + n_{rb}$ where n_{sb}, n_{rb} and n_{sr2} are the source-boundary distance, receiver-boundary distance, and maximum source-receiver offset, respectively. The amplitude of data recorded at the last receiver position (offset = n_{sr2}) is much lower than that at the first receiver position (offset = n_{sr1} where n_{sr1} is the minimum source-receiver offset) due to geometry attenuation. Therefore, the influence of n_{rb} on data approximation is ignorable compared to that of n_{sb}. In the examples of this paper, we fix n_{sb} as one grid greater than n_{cpml} where n_{cpml} is the width of the CPML and is set to be 10 grids. By doing so, n_{rb} and n_{sr2} determine n_{xi}.

We take the mean structural similarity (MSSIM) index to measure the fidelity of the approximated data with respect to the data simulated on an entire model (Wang et al., 2004). Readers are recommended to refer to Boniger and Tronicke (2010) for an example of using MSSIM index in GPR data comparisons. Furthermore, we regard the reconstructed models of FWI as a reference and use MSSIM index to describe the alikeness of the reconstructed models of SFWI relative to the reference. We adopt the same settings as Boniger and Tronicke (2010) to compute MSSIM index. The closer the MSSIM index is to one, the more similar the two comparison objects are. In the following, we analyze the roles of n_{rb} and n_{sr2} in the simulation and inversion.

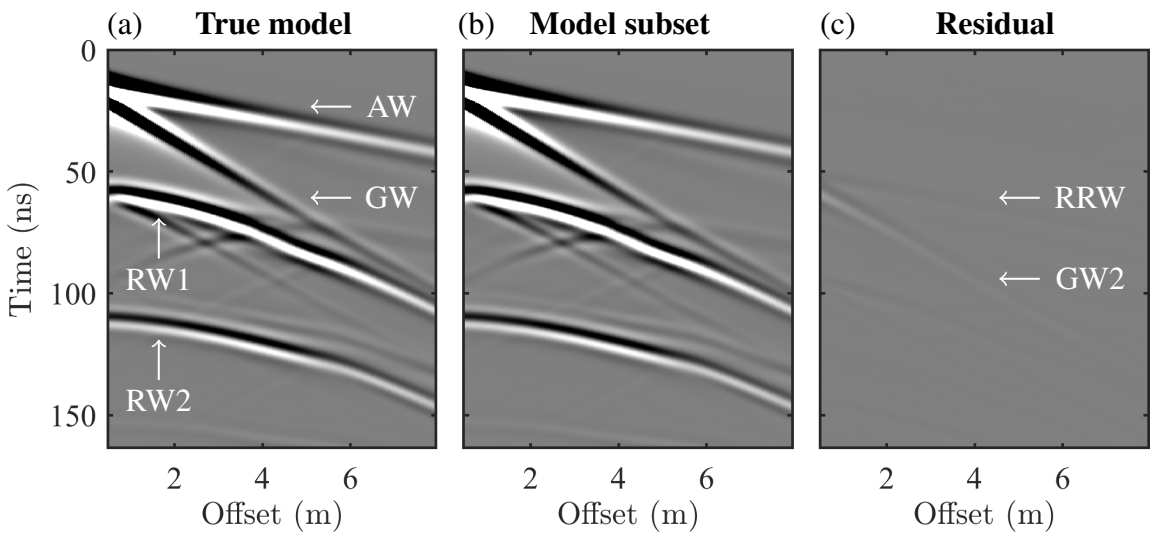

Figure 2.4: Radargrams of the 11th source simulated on (a) the true model shown in Figure 2.3a and (b) the model subset with n_{sb} = 1.6 m. The extraction area of the model subset is illustrated in Figure 2.1a. In (c) we show the residual between (a) and (b). The color range in (a–c) is clipped to 2 per cent of the highest amplitude of the true data in (a) for better visualization. AW = air wave, GW = ground wave, RW1 = the 1st reflected wave (2-m deep interface), RW2 = the 2nd reflected wave (4-m deep interface), RRW = reflected-refracted wave, GW2 = the 2nd ground wave.

2.3.1 Effect of the source-boundary distance

We build 2D synthetic models from the 2D field example in the next section. As shown in Figure 2.3a, the true models of the permittivity and conductivity consist of a three-layer structure and a trench anomaly embedded below the ground. The magnetic permeability is constant and equal to its value in vacuum. The acquisition geometry is the same as in Figure 2.1 where we place 18 sources, shifted Ricker wavelets (100 MHz), on the ground to generate the transversal magnetic waves in which there is only one electric field component perpendicular to the survey plane. The receivers record multi-offset electric field data on the right of the source.

Figure 2.4 shows the radargrams of the 11th source. To obtain the approximated data (Figure 2.4b), we implement forward modeling on the model subset with n_{sb} = 1.6 m. The data simulated on the true entire models mainly contain the air wave, ground wave, and the waves reflected from the subsurface. The approximated data are almost the same as the true data, except slight differences in the reflected-refracted wave (RRW) and the 2nd ground wave (GW2). The reason for these differences is that the distances from the 11th source to the 2-m deep interface and the left boundary of the trench are the same (see Figure 2.1b), resulting in the superposition of the 1st reflected wave (RW1) and the lateral reflected wave from the left boundary of the trench in the source location; the superimposed waves generate RRW and GW2; hence, RRW and GW2 change with the cut-off of the left boundary of the trench on the 11th model subset. However, the residuals have amplitude two orders of magnitude smaller than that of the true data and thus are considered to be ignorable.

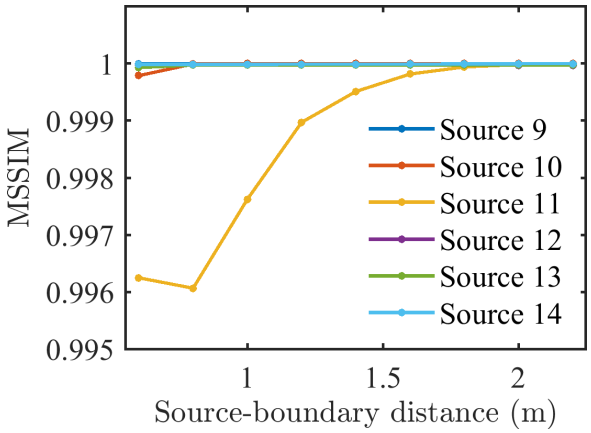

Figure 2.5: Mean structural similarity (MSSIM) index of the approximated data as a function of source-boundary distance. The six sources are generated on the model subsets extracted from the true model.

We simulate the wavefields of the 9th–14th sources on the model subset with n_{sb} varying from 0.6 m to 2.2 m, similar to Koyan and Tronicke (2020). The source-receiver spreading of the 9th–14th sources is across or close to the trench (see Figure 2.1). Figure 2.5 illustrates that the MSSIM index for all data is greater than 0.995, which means that the approximated data are highly consistent with the true data. The 11th source has a sharper decrease in the MSSIM index than other sources when n_{sb} decreases. This is due to the special position of the 11th source we mentioned above. We do not observe any effects of the cut-off of the model subset on the MSSIM index when n_{sb} = 2 m (the wavelength of the dominant frequency near the surface). Therefore, we take one wavelength as a rule of thumb for choosing the source-boundary distance. Note that we still use n_{sb} = 1.6 m in the following inversion in order to study the extent of its effect on the reconstruction result.

2.3.2 Effect of the maximum source-receiver offset

To investigate the influence of the maximum source-receiver offset n_{sr2} on the SFWI, we implement a 2D synthetic inversion example using the initial model shown in Figures 2.3b. For the observed data, we add 20 dB Gaussian noise to the data. We use a multiscale strategy in the inversion to avoid circle skipping (Bunks et al., 1995). We select five inversion stages to invert data with progressively expanding bandwidth and decreasing wavelength (Meles et al., 2012; Zhou et al., 2021b). From the first stage to the fifth stage, the frequency bands vary from 5 to 30, 40, 50, 70, and 100 MHz. We estimate a wavelet correction filter by a stabilized deconvolution and use it to correct the source time function at the beginning of each stage (Ernst et al., 2007a; Groos et al., 2014). The maximum number of iterations per stage is 15, and the abort criterion is that the relative data misfit improvement is less than 1 per cent. We smooth the gradient with a 1D Gaussian spatial filter in the horizontal direction to suppress the artifacts smaller than the FWI resolution. As the air layer is known, we multiply the gradient by a taper to zero out the gradient in the air layer.

Figures 2.3c and 2.3d demonstrate that the reconstructed results of FWI and SFWI are almost consistent. The permittivity and conductivity models, reconstructed mainly by small angle reflected waves, delin-

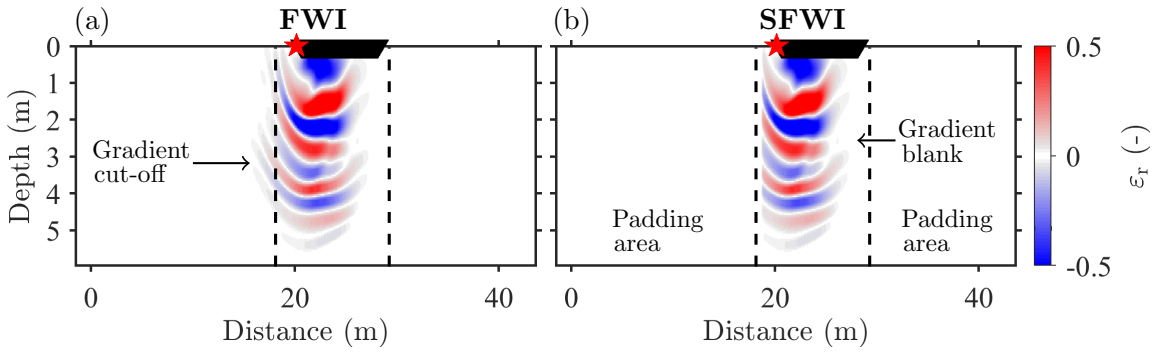

Figure 2.6: Permittivity gradient of the synthetic data example calculated by the 11th source in the first iteration of (a) FWI and (b) SFWI using n_{sb} = 1.6 m and n_{sr2} = 8 m. The red star presents the source, and the black triangles are the receivers. The black dashed lines outline the boundaries of the model subset used in SFWI. The color range in (a) and (b) is clipped to 50 per cent of the maximum value of the gradient in (a) for better visualization of the cut-off gradient.

eate the upper part of the triangular trench and the interfaces of the three horizontal layers. However, the values inside these geological bodies are not reliably reconstructed due to lack of long wavelength information in the surface multi-offset GPR data (Lavoué et al., 2014). Compared to the permittivity model, the conductivity model contains more artifacts, e.g., between the ground and the 2-m deep interface, which shows that conductivity estimation is more difficult and sensitive to noise.

In Figure 2.6 we show the permittivity gradient computed from the 11th source in the first iteration of FWI and SFWI. The illumination of one source in the FWI (Figure 2.6a) is rather narrow compared to the entire model space, caused by the high decay of the EM waves propagating in the subsurface and the focusing effect due to the source specific radiation pattern. This provides the possibility of using SFWI to save computational time and memory usage. In the gradient of SFWI (Figure 2.6b), the gradient outside the extraction area is missing because of the selection of n_{sb}. These cut-off errors are relatively small and do not contribute substantially to the reconstruction in Figure 2.3d.

The blank area on the right part of the model subset in the SFWI gradient (about 2 m width in Figure 2.6b) suggests that we can further reduce the maximum source-receiver offset n_{sr2} to obtain even larger computational savings. In Figure 2.7a we show the MSSIM index calculated by SFWI with different n_{sr2}. For consistency of data misfit in Figure 2.7b, we perform additional forward modeling on all reconstructed models with n_{sr2} = 8 m. FWI and SFWI result in the same permittivity reconstruction and data misfit when using n_{sr2} = 8 m. Due to the cut-off artifacts and sensitivity differences, the conductivity MSSIM index is lower than the permittivity MSSIM index. The model fidelity and data fitting become worse when n_{sr2} decreases. Nevertheless, the difference between n_{sr2} = 6 m and n_{sr2} = 8 m is relatively small ($<$ 0.01 MSSIM), implying that we can drop n_{sr2} to 6 m to save computational costs at the expense of a little reconstruction accuracy.

Thanks to the workflow in Figure 2.2, SFWI is compatible for parallelizations, such as source parallelization and model domain parallelization. In the case of the latter, one can use the model domain decomposition technique based on the message passing interface (Bohlen, 2002). We run the inversions

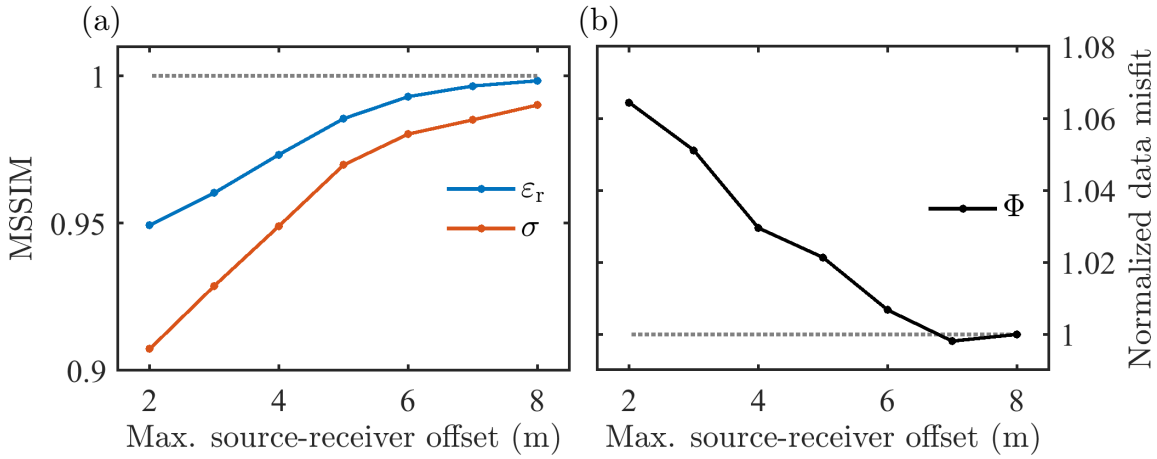

Figure 2.7: Model fidelity and data fitting comparison in the synthetic example. (a) MSSIM index of the model parameters reconstructed by SFWI as a function of n_{sr2}. (b) Final data misfit (equation 2.4) of SFWI as a function of n_{sr2}, normalized by that of FWI using n_{sr2} = 8 m. The gray dotted line marks the desired value.

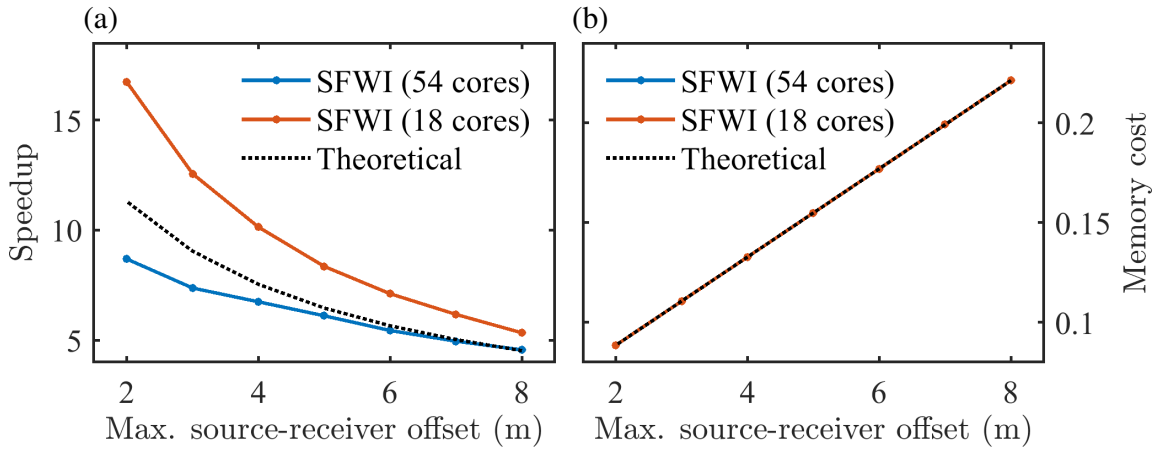

Figure 2.8: (a) Speedup and (b) memory cost of SFWI as a function of n_{sr2}, compared to FWI. The blue line indicates the real values observed in the SFWI using both source parallelization and model domain parallelization, and the red line illustrates those using only source parallelization. The dotted line is the theoretical values calculated by equation 2.8. The improvements in (a) and (b) are observed in the synthetic example and the field example at the Rheinstetten test site.

Table 2.1: Acquisition parameters of the multi-offset surface GPR data (200 MHz) at the Rheinstetten test site and those used within the FWI / SFWI.

Parameters	Raw	FWI
Number of sources	165	18
Traces per gather	56∼125	100∼175
Transmitter spacing	∼0.2 m	2 m
Receiver spacing	∼0.1 m	0.04 m
Minimum offset	0.2 m	0.3 m
Maximum offset	17 m	8 m
Sample rate	0.2 ns	0.08 ns
Recording window	200 ns	164 ns

of 18 sources on a central processing unit (CPU) server with an Intel(R) Core(TM) i5-9500 CPU @ 3.00 GHz and evaluate the performance of SFWI in two conditions: source parallelization alone and two parallelizations simultaneous. We allocate 18 cores with one core per source for the source parallelization alone. To perform both source parallelization and model domain parallelization, we allocate 54 cores with three cores per source, which means 18 source parallelizations and 3 model domain parallelizations.

Figure 2.8b shows that all parallelizations lead to almost the same theoretical and real memory savings for SFWI, as the memory usage is dominated by the wavefield storage for calculating the gradient and is determined by the actual model size used in numerical simulation. In Figure 2.8a we observe that the speedup of the source parallelization (the red line) is above the theoretical curve for all source-receiver offsets. It is due to that the memory usage of SFWI ($\leq$ 22 per cent) is much smaller than that of FWI and not considered in the theoretical speedup. However, when n_{sr2} is less than 6 m, the combination of source parallelization and model domain parallelization (the blue line) does not bring the expected acceleration. The reason is that we decompose the entire volume in x direction, similar to the extraction of the model subset in Figure 2.1a; FWI and SFWI take the same communication time for the decomposed subvolumes; the weight of the communication time in the running time increases as n_{sr2} decreases, which causes less improvement than the theoretical values when the weight exceeds a certain level (Bohlen, 2002). Note that these improvements are relative to the FWI using the same parallelizations. The real computational time of the SFWI using these two parallelizations simultaneously (54 cores with three cores per source) is shorter than the SFWI using source parallelization alone (18 cores with one core per source), but their ratio is greater than 1/3 for the reasons mentioned above.

Table 2.2: Multi-offset surface GPR data (200 MHz) preprocessing steps.

(1)	Data resampling in the frequency domain
(2)	Interpolation of clipped direct-arrival amplitudes
(3)	DC-shift removal and dewow
(4)	Bandpass filtering (5 – 400 MHz)
(5)	Bad traces removal and offset limitation
(6)	Data gridding in the time-offset domain
(7)	3D-to-2D transformation

2.4 Field examples

2.4.1 Surface GPR data

Surface data acquisition and preprocessing

We conducted the field measurement at the northeast corner of the gliding airfield Rheinstetten, Germany. Previous GPR migration imaging (Wegscheider, 2017) and shallow-seismic FWIs (Wittkamp et al., 2019; Gao et al., 2020; Pan et al., 2021; Irnaka et al., 2022) have validated that there is a V-shaped trench named Ettlinger Line, embedded in fine fluvial sediments of the Rhine river. The trench was excavated as a fortification in the early 17th century and has been leveled to the ground and is no longer visible from surface at the test site. Taking the 'walk-away' method (Domenzain et al., 2021), we acquired 165 multi-offset surface radargrams along a survey line perpendicular to the Ettlinger Line. The acquisition settings for the multi-offset surface GPR data are listed in Table 2.1. We used pulseEKKO Pro transmitter equipped with a pair of unshielded 200 MHz antennas and mounted a receiver on a sled for smooth movement. To track the receiver coordinates at the centimeter level accuracy, we adopted real-time kinematic (RTK) positioning with a self-tracking total station (Boniger and Tronicke, 2010).

We preprocess the field data in order to perform the inversion. Table 2.2 summarizes the main steps for data preprocessing. We first resample the data for forward modeling requirements (Irving and Knight, 2006) and interpolate the clipped amplitudes of high-energy direct arrivals that exceed the dynamic range of the acquisition unit (Benedetto et al., 2017). In steps (3) and (4), we eliminate noise at low and high frequencies. Then we delete the traces of waveform distortions and remove the traces with an offset larger than 8 m. Due to the uneven walking speed of the worker when moving the sled, the acquired data have irregular trace spacing. To ensure a balanced illumination in the measurement area, we apply the data gridding in step (6), i.e., 2D spline interpolation in the time-offset domain at regular trace spacing. Since our forward engine is 2D, we perform a 3D-to-2D transformation of reflected wave on the data in the last step (Forbriger et al., 2014). Table 2.1 shows the acquisition parameters used for FWI / SFWI after data preprocessing. To further save computational cost, we use only 18 sources in the inversions (see Figure 2.1a for their locations).

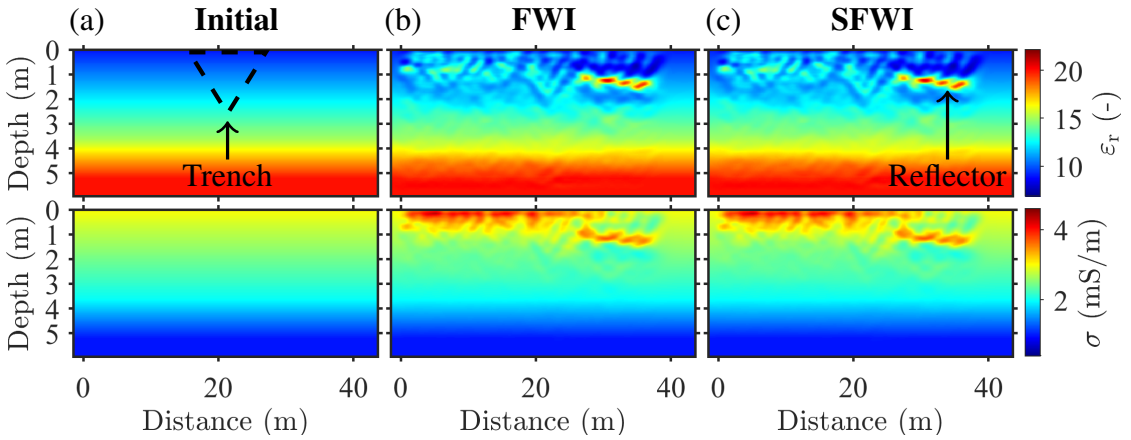

Figure 2.9: Models of the field example at the Rheinstetten test site, showing (a) the initial models, (b) the reconstructed models of FWI, and (c) the reconstructed models of SFWI. The dashed triangle in the initial model outlines the target trench, known as the Ettlinger Line.

Surface data inversion

We use the same inversion settings as for the synthetic examples presented in the previous section. Additionally, we normalize the gradients of each source before summing in order to eliminate energy differences in the data that may be due to variations in instrument performance, antenna coupling, or acquisition arrays. Figure 2.9a shows the initial models where the initial relative permittivity is 9 at the ground (the velocity of ground wave is 0.1 m ns^{-1}) and gradually increases to 20 at a depth of 6 m. On the other hand, the initial conductivity is 3 mS m^{-1} at the ground and decreases gradually to 1 mS m^{-1} at a depth of 6 m. We estimate the near-ground values of permittivity and conductivity from the ground wave in the field data (Annan, 2005).

As shown in Figures 2.9b and 2.9c, the reconstructed models of FWI and SFWI are almost identical, and both illustrate the presence of a triangular anomaly in the permittivity model, which is the expected location of the Ettlinger Line. SFWI can outline the trench structure at the interface between the refilled sand and underlying soil due to their dielectric differences. However, the permittivity values inside the trench are less reliable as the surface GPR data lack low wavenumber information, similar to our observation in the synthetic examples. On the right of the trench, we observe a strong reflector at 1.1 m depth that has a large contrast with the overlying materials in the permittivity and conductivity models. We cannot see the same reflector on the left side of the trench where the original ground surface was higher than on the right side and removed when leveling the trench (see schematics in Figure 3 in Irnaka et al. (2022)). It suggests that the reflector is likely to be interpreted as the original ground surface. On the left side of the trench, there is a near-surface conductive ($\sim$4 mS m^{-1}) layer that probably reveals a higher water content. Therefore penetration depth of the GPR signal is limited in this area.

Figure 2.10 displays the radargrams and the estimated source signals. As FWI and SFWI have the same reconstruction ability, we show only the results of SFWI (the red line corresponds to the model reconstruction in Figure 2.9c). We acquired the first 11 radargrams on the first day and the last 7 radargrams on the second day. The estimated source signals (Figure 2.10d) of the first 11 gathers differ from the last

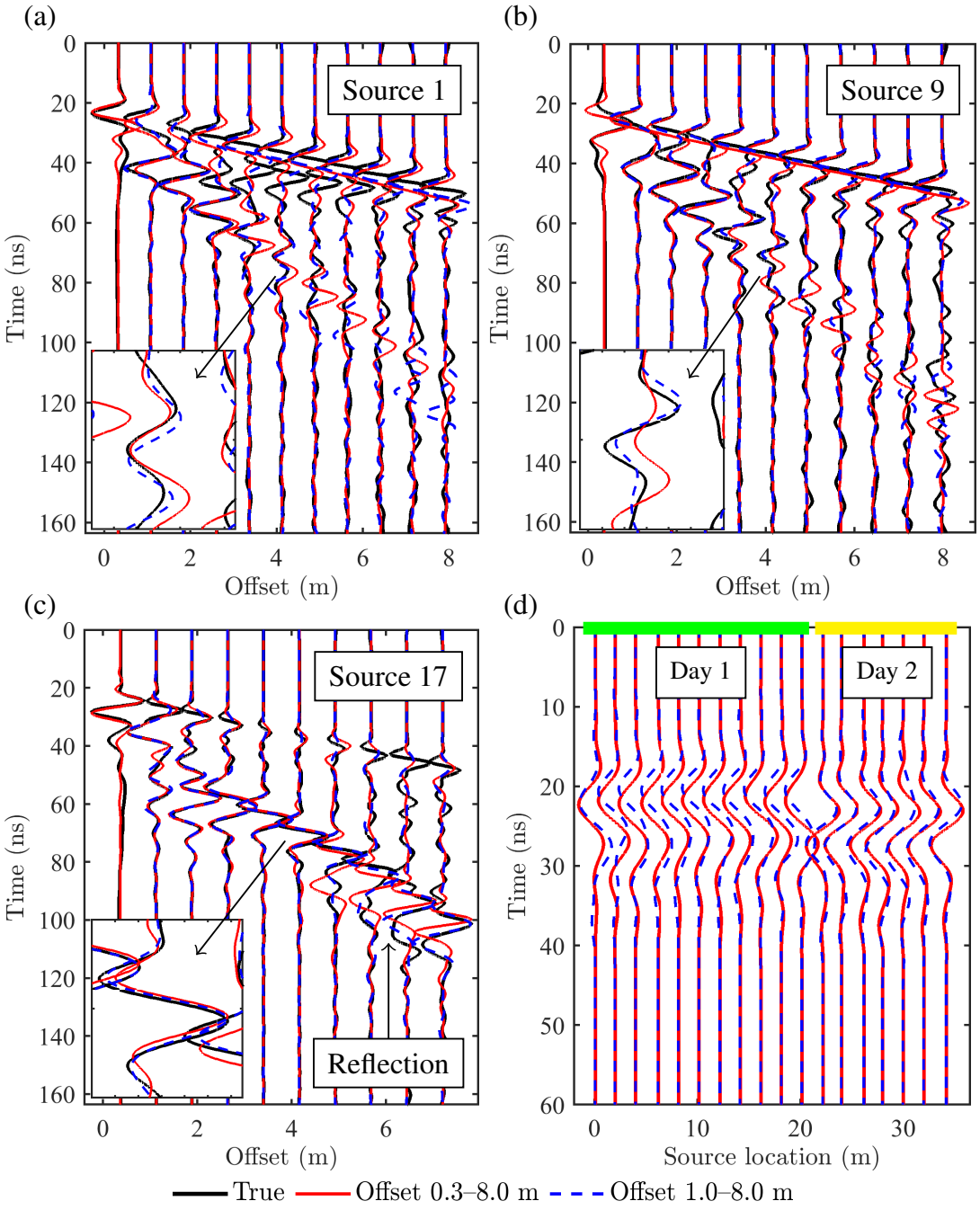

Figure 2.10: SFWI data fitting of (a) the 1st, (b) 9th, and (c) 17th sources of the field example at the Rheinstetten test site, normalized and shown once every 20 traces for better visualization. Note that the first trace (offset = 0.3 m) does not appear in the SFWI result using the data of offset 1.0–8.0 m. The rectangular windows highlight local waveforms. (d) Estimated source signals of 18 radargrams in SFWIs. The green and yellow lines mark the sources used on different days. A Butterworth bandpass filter (5–100 MHz in the fifth stage) is applied to the radargrams.

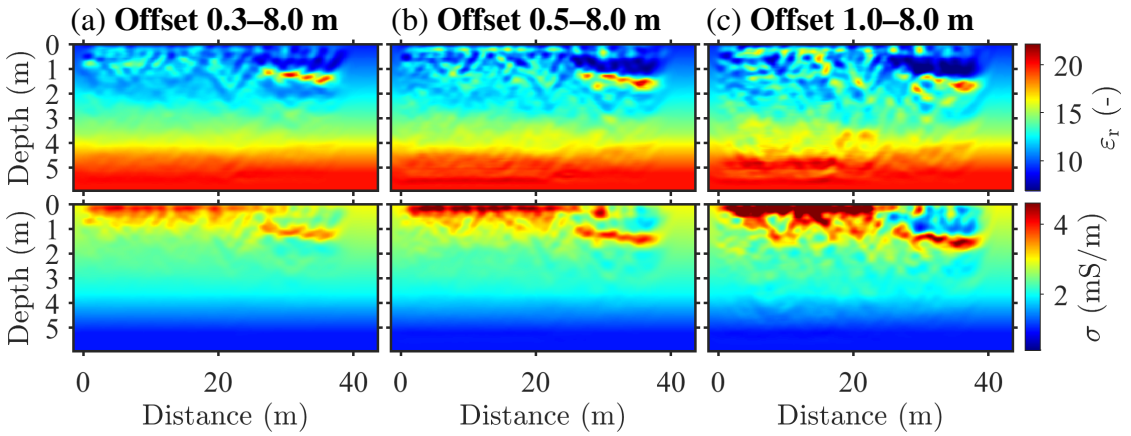

Figure 2.11: Reconstructed models of the field example at the Rheinstetten test site. We show SFWI results using data with offset (a) 0.3–8.0 m, (b) 0.5–8.0 m, and (c) 1.0–8.0 m.

7 gathers, probably due to antenna polarity changes or coupling differences caused by slight near-surface moisture variations resulting from drying or wetting at night. However, the source signals used on the same day present similar waveform shapes and travel times, which demonstrates the stability of the SFWI. Figures 2.10a and 2.10c indicate that the near-surface conductive layer (see Figure 2.9c) causes a higher attenuation of the ground and reflected waves in the 1st source than in the 17th source. In all data, the air and ground waves are hard to match because we use a 3D-to-2D transformation of reflected wave, and our 2D forward solver cannot account for the radiation patterns and appropriate antenna coupling in the real world. However, reflected waves are in good agreement with the field data for source-receiver offsets of less than 3 m. When the offset is over 3 m, the performance of SFWI in data fitting degrades.

Effect of the minimum and maximum source-receiver offsets

The mismatch of the first trace at 0.3 m offset in Figures 2.10a–c implies that some low-frequency components (e.g., near field effects) are not perfectly removed in the preprocessing. It raises the question of whether we can delete some near-offset traces to obtain a better fit to the data. To evaluate the effect of the near-offset data on SFWI's performance, we fix the maximum source-receiver offset n_{sr2} = 8.0 m and change the minimum source-receiver offset n_{sr1} to 0.5 and 1.0 m, respectively. Figure 2.10 presents the inverted data and estimated source signals when n_{sr1} = 1.0 m (the blue dashed line), and Figures 2.11b and 2.11c exhibit the reconstructed models. As expected, when increase n_{sr1} to 1.0 m, we do observe an improved data fitting in some traces, such as the zoomed waveforms in Figures 2.10a–c and the strong reflection event in Figure 2.10c. It should attribute to the shifted arrival times of some source signals (Figure 2.10d) and the deeper and stronger contrast of the right-hand reflector (1.4 m) in the permittivity and conductivity models (Figure 2.11c). Nevertheless, the trench boundaries become blurred in the permittivity model when n_{sr1} = 1.0 m, making it more difficult for us to identify the target. From this perspective, using the near-offset data with n_{sr1} = 0.3 m in SFWI is more favorable for geological interpretation.

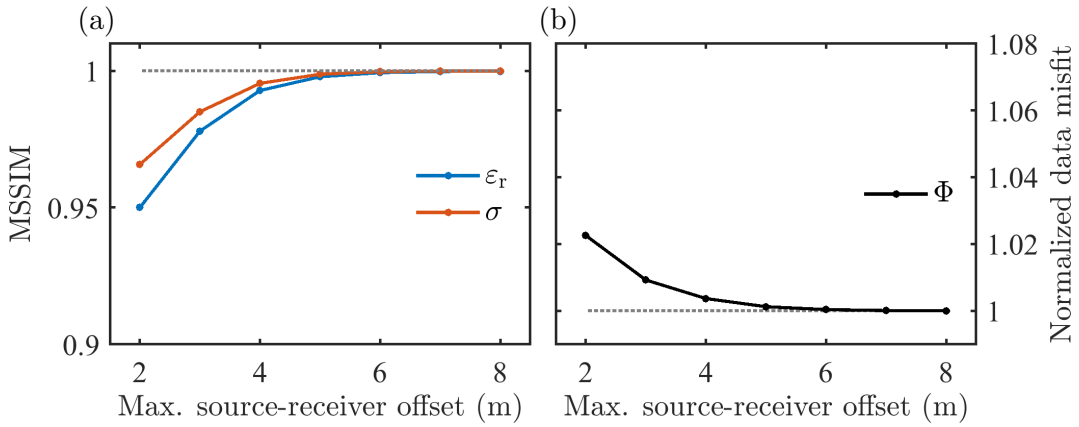

Figure 2.12: Model fidelity and data fitting comparison in the field example at the Rheinstetten test site. (a) MSSIM index of the model parameters reconstructed by SFWI as a function of the maximum source-receiver offset n_{sr2}. (b) Final data misfit (equation 2.4) of SFWI as a function of n_{sr2}, normalized by that of FWI using n_{sr2} = 8 m. The gray dotted line marks the desired value.

We keep n_{sr1} = 0.3 m and test the performance of SFWI with different n_{sr2}. The accelerations and memory savings for SFWI of the field data are the same as those of the synthetic data shown in Figure 2.8. Figure 2.12 displays the corresponding MSSIM index of the reconstructed models and the corresponding data misfits. Compared to the synthetic example in Figure 2.7, the field example suffers less from the reduction in n_{sr2} because the field data are produced by the source of higher frequency (200 MHz) and have lower S/N. Consequently, changing n_{sr2} from 8 m to 5 m in SFWI almost does not cause a change in the data misfit and model quality but can obtain a real speedup factor S^x_{2D} of 8.3472 when using source parallelization alone (see Figure 2.8a). Figure 2.12a shows that the conductivity fidelity is less affected by n_{sr2} than the permittivity fidelity, contrary to our observation in the synthetic example in Figure 2.7a. This may be due to the fact that in the field example, e.g., in Figure 2.11a, the area where the conductivity is rarely updated (at a depth of more than 1.5 m) is larger than the area where the permittivity is rarely updated (at a depth of more than 2.0 m), leading to persisting starting model and, thus, to a higher MSSIM index of conductivity.

Comparison with migration imaging

In Figure 2.13 we compare the SFWI result of multi-offset GPR data with the migration result of common-offset GPR data along the same profile. After performing a classical GPR processing including bandpass filter (50–400 MHz), zero-time correction, time depended scaling, we apply a Kirchhoff migration routine using a constant subsurface velocity of 0.1 m ns^{-1}. The resulting migrated GPR image (Figure 2.13a) shows the trench in the central location and the dipping high amplitude reflector at the right side of the trench. The migrated data on the left side of the trench appear more chaotic, with a lower penetration depth. These interpretations agree well with the FWI and SFWI results in Figure 2.9b and 2.9c. For a better comparison we also show the difference of the SFWI permittivity result to the initial model in Figure 2.13b. Within these differences, we observe that the area of the trench has been updated to higher permittivity values until a depth of ∼2.0 m. The same applies to the left side of the

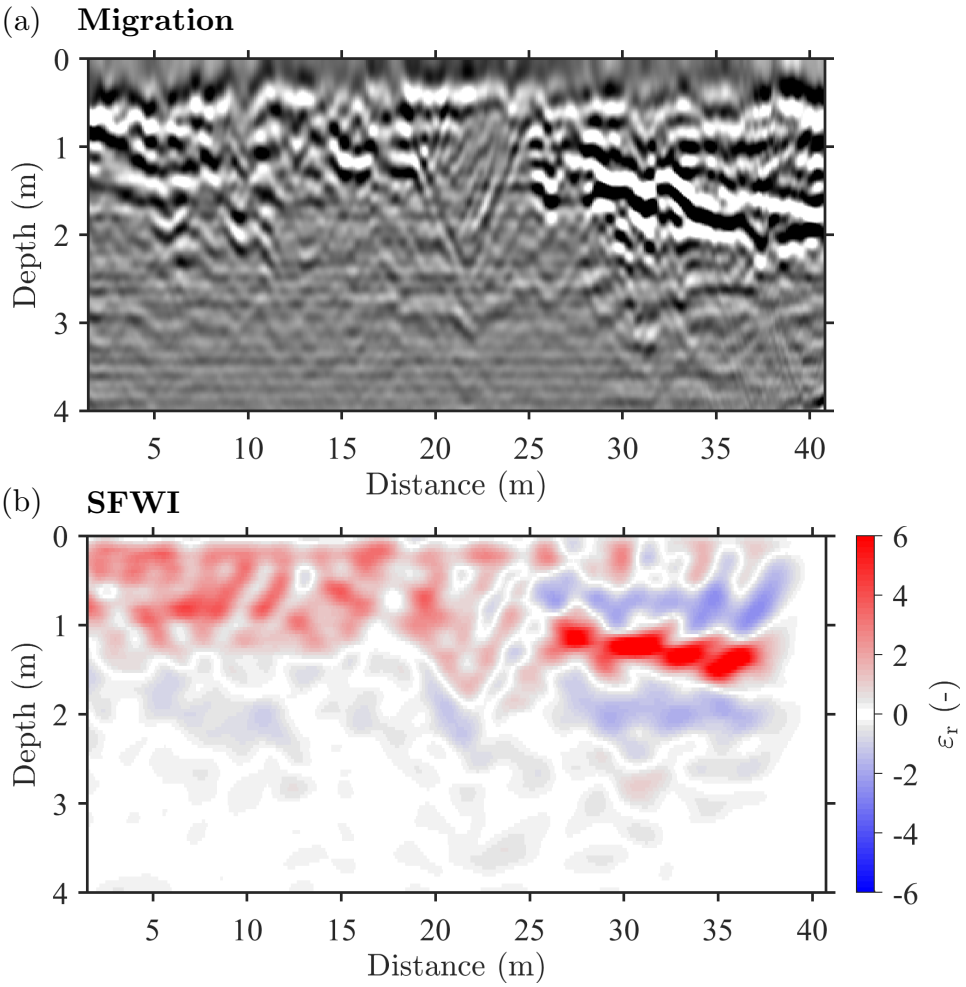

Figure 2.13: Comparison between (a) the migration result of common-offset data (offset = 0.5 m) and (b) the SFWI result of multi-offset data (offset = 0.3–8.0 m). In (b) we show the cumulative perturbations added to the initial model, i.e., the residuals between Figures 2.9c and 2.9a.

trench down to depth of ∼1.3 m. The strong reflector data on the right side is imaged by high and low permittivity update, following the shape of the reflector. These patterns indicate the structure of the reflector, but the absolute permittivity values (Figures 2.9c) are probably not correct due to the lack of low frequency information and the resulting insufficient inversion updates (as also observed in the synthetic example). Therefore, we conclude that GPR FWI and SFWI are capable to reconstruct the main structures in the subsurface.

2.4.2 Crosshole GPR data

The SFWI is also suitable for decreasing the calculation time when inverting crosshole GPR data. In this section, we show a well known field example of crosshole GPR data acquired at the Krauthausen test site. Readers are recommended to Gueting et al. (2015), Gueting et al. (2017), and Zhou et al. (2021a) for more details about the study site and acquisition settings. As shown in Figure 2.14a, the investigation plane is divided into four crosshole planes (i.e., B32–B38, B38–B31, B31–B62, and B62–B30) between

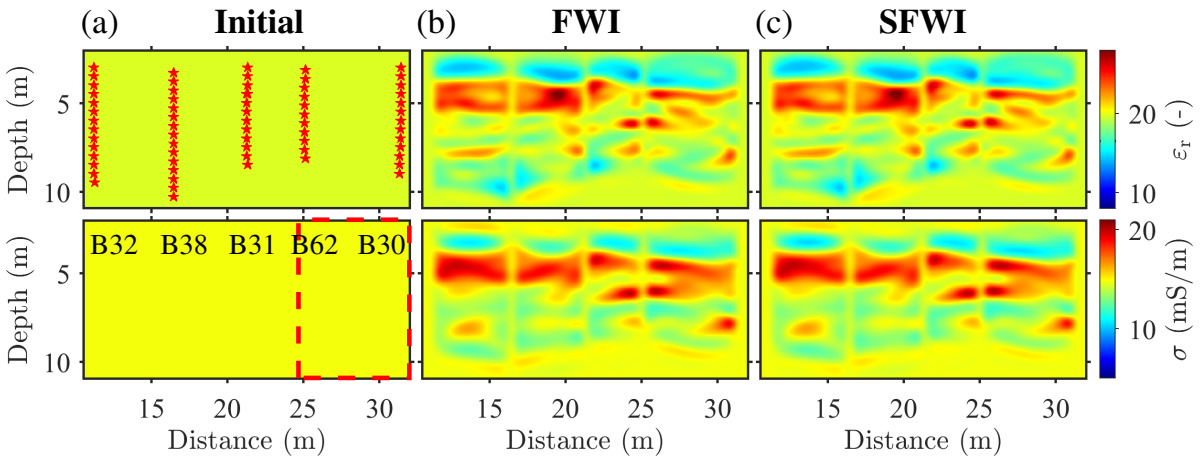

Figure 2.14: Models (430×180 grids with the grid spacing of 0.05 m) of the field example at the Krauthausen test site, showing (a) the initial models, (b) the reconstructed models of FWI, and (c) the reconstructed models of SFWI. In the initial models, the red stars represent the transmitters located at five boreholes (B32, B38, B31, B62, and B30), and the red dashed rectangle illustrates the extraction region of a model subset (150×180 grids).

five boreholes for GPR data acquisition. We use 102 radargrams (200 MHz) in the inversion, with 28, 27, 23, and 24 radargrams acquired from the left to the right crosshole planes, respectively. The preprocessing steps of the crosshole GPR data include data resampling in the frequency domain, DC-shift removal and dewow, bandpass filtering (5–400 MHz), and 3D-to-2D transformation of transmitted wave (Bleistein, 1986; Forbriger et al., 2014). After resampling, the data has a sample rate of 0.3 ns and a recording window of 307.2 ns.

The starting model is homogeneous ($\varepsilon_r = 20$, $\sigma = 15$ mS m^{-1}, see Figure 2.14a), and the inversion settings are analogous to those used in the field example of the surface data. For crosshole GPR FWI / SFWI, we adjust the forward solver, frequency bands (from 5 to 30, 50, 70, 100, and 130 MHz), gradient filter, and gradient taper. We use numerical simulations of transversal electric waves where there is only one magnetic field component perpendicular to the survey plane. To suppress potential artifacts, we apply a 2D Gaussian spatial filter to smooth the gradient since each crosshole plane is covered by more than 20 sources at different angles, and thus the crosshole illumination is bidirectional and more balanced than the surface illumination. We multiply the gradient by a taper to reject areas outside the source-receiver plane. For instance, for the sources located in the B32–B38 plane (see Figure 2.14a), the gradient at distances greater than B38 or less than B32 will be muted. In SFWI, the size of the model subset is selected to cover the B62–B30 plane that has the maximum crosshole distance (6.16 m, see the red dashed rectangle in Figure 2.14a).

The results shown in Figures 2.14b and 2.14c indicate that FWI and SFWI have the same performance in reconstructing the models. The MSSIM index of ε and σ is very close to one. Both FWI and SFWI reveal a poorly sorted gravel layer (depth shallower than 4 m), a well sorted sand layer (depth between 4 and 6 m), and a sandy gravel layer (depth deeper than 6 m) (Tillmann et al., 2008). Our inversion results are comparable to previous studies of Gueting et al. (2015), Gueting et al. (2017), and Zhou et al. (2021a) (note the corrections in Gueting et al. (2020)), although we use a homogeneous initial model.

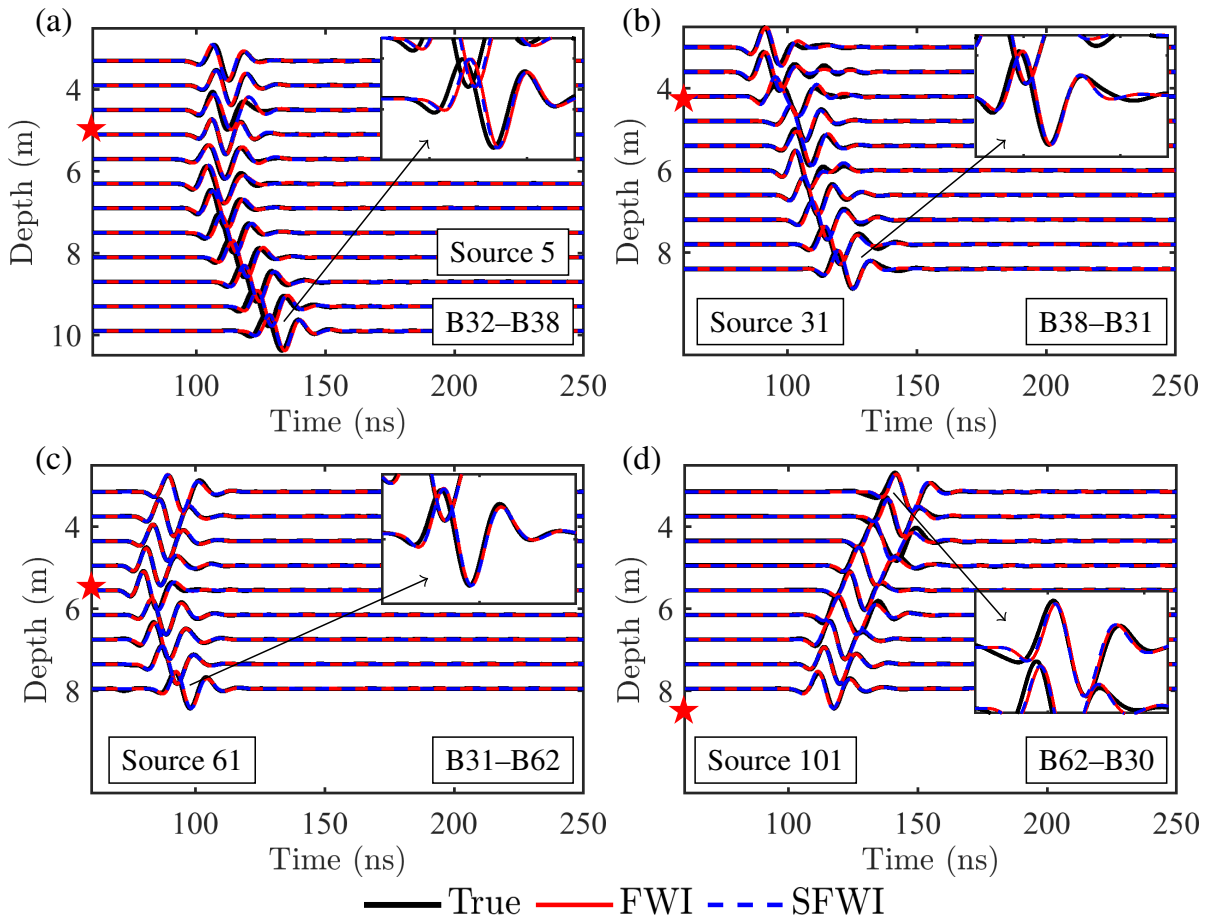

Figure 2.15: Data fitting of (a) the 5th, (b) 31st, (c) 61st, and (d) 101st sources of the field example at the Krauthausen test site, normalized and shown once every six traces for better visualization. The rectangular windows highlight local waveforms and the red stars mark the depth of the sources. A Butterworth bandpass filter (5–130 MHz in the fifth stage) is applied to the radargrams.

Figure 2.15 shows four radargrams from the four crosshole planes. The synthetic data of FWI and SFWI are very similar and match the observed data well (the data misfits of the two inversions are identical). Compared to the other sources, the 5th source has larger data fit errors at 8–10 m depths (see the enlarged window in Figure 2.15a), indicating that in the B32–B38 plane, the reconstructed permittivity at these depths could be slightly higher than the actual values. This can be improved by using the ray-based initial models (Gueting et al., 2015) or constraining the inversion with a priori information (Zhou et al., 2021a).

As shown in Figure 2.16, source signals estimated by FWI and SFWI have almost the same waveforms and amplitudes. In Figure 2.16c, the amplitude differences of signals in different crosshole planes may be due to the borehole spacing (n_{sr1} in Figure 2.16a) and horizontal variation of the media (Figure 2.14b and 2.14c), while those in the same crosshole plane may be caused by vertical variation of the media (Figure 2.14b and 2.14c). These suggest the need for gradient normalization when inverting field data. In the B32–B38 plane (Figure 2.16a and 2.16b), we find that sources located in the same borehole (source 1–14 in B32 borehole and source 15–28 in B38 borehole) have relatively large phase shifts, which leads

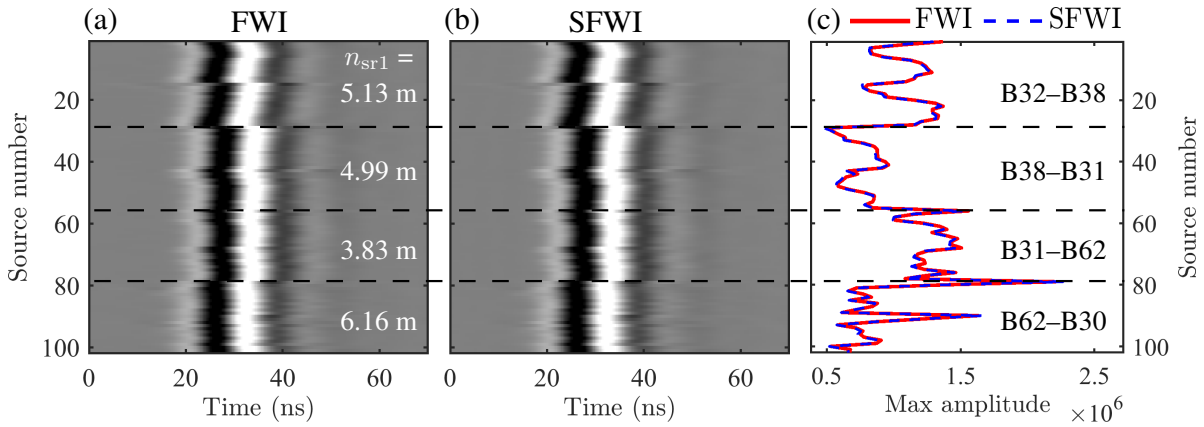

Figure 2.16: Estimated source signals of 102 radargrams in (a) FWI and (b) SFWI at the Krauthausen test site. Signals are normalized and their real amplitudes are plotted in (c). The black dashed lines are used to distinguish the source signals on different crosshole planes.

Table 2.3: Computational costs of FWI and SFWI in the field example at the Krauthausen test site. The ratio is obtained by dividing FWI by SFWI, and the running time of a inversion of 102 sources is measured on a CPU server using 34 cores (one core per source, looping three times in one simulation).

	FWI	SFWI	Ratio
Model size (grid)	430×180	150×180	2.8667
Running time (s)	2760.386	846.2059	3.2621
Memory cost (MB)	922.385	323.907	2.8477

to larger data mismatch in the radargrams (e.g., Figure 2.15a). It implies that the reconstruction of this crosshole plane is likely to have more uncertainty than the other three crosshole planes.

In Table 2.3 we show the computational costs of FWI and SFWI. Unlike the surface SFWI, where the maximum source-receiver offset n_{sr2} can adjust for higher acceleration and more memory savings (see Figure 2.8), the crosshole SFWI has a limited improvement. It is because n_{sr2} is fixed in the crosshole geometry, owning to the constraints of the borehole location. However, we still observe a speedup (3.2621) higher than the theoretical value (2.8667) when using only source parallelization. The memory reduction of SFWI approaches the theoretical value for the crosshole data (Table 2.3) and the surface data (Figure 2.8b). These examples prove that SFWI is applicable to multiple acquisition geometries.

2.5 Discussion

In this study, we use a 'walk-away' method to record multi-offset surface GPR data. Our field data acquisition is slower than Lavoué (2014), which extracted hundreds of multi-offset gathers from 15 common-offset gathers with an offset interval of 0.5 m. However, each trace in the multi-offset GPR data used in Lavoué (2014) could be produced by a different transmitter-ground coupling. In contrast, our measurement can be expected to have the same transmitter-ground coupling and denser trace intervals ($\sim$0.1 m) in a single gather.

In the field example of multi-offset surface GPR data (200 MHz), we find that near-offset data play an important role in delineating the subsurface structure. This observation may be site dependent and differs from other field applications of surface GPR FWI. For instance, Lavoué (2014) used 250 MHz antennas with a minimum source-receiver offset of 1.24 m to detect a dipping limestone reservoir starting from the surface to 8 m depth; Domenzain et al. (2021) used 50 MHz antennas with a minimum source-receiver offset of 7.5 m to explore coarse fluvial deposits up to 6 m depth. Limestone reservoir and coarse fluvial deposits allow for deeper penetration of GPR signals and more reflection information, hence the greater use of far-offset data in FWI. Our test site consists of fine gravel, sand, and clay and lacks the physical property differences that produce reflections below 2 m depth (see Figure 2.13a). Thus FWI / SFWI uses mainly the near-offset data to reconstruct the trench and areas to its left. However, the far-offset data become more important when imaging areas to the right of the trench where strong reflections occur (see Figure 2.10c).

We repeat the surface GPR SFWI of the first field example but with an initial permittivity model of homogeneous half-space ($\varepsilon_r = 9$). To our surprise, the data fitting and cumulative perturbations are very close to those in Figures 2.10a–c and Figure 2.13b. When we change the initial permittivity model to an otherwise homogeneous half-space with values far from those derived from the ground wave velocity, SFWI fails to match the observed data and reconstruct our target. We observe a similar but weaker change when testing the initial conductivity model in the same way, probably due to the weaker sensitivity of GPR data to conductivity. These tests demonstrate that, in the initial models, the accuracy of the near-surface part is more significant for FWI / SFWI than the deeper part. It is good news for surface GPR FWI / SFWI because the near-surface a priori information is easily available from the ground wave.

Figures 2.3 and 2.13 reveal a limitation of surface GPR FWI /SFWI, i.e., when strong reflection appears in the observed data, the inversion mainly focuses on outlining the interface not estimating the real values above and below the interface. This phenomenon rarely occurs in crosshole GPR FWI using transmitted wave because it comes from the migration kernel of the reflected wave (Xu et al., 2012). There are two ways to address this problem. One is to implement reflection-waveform inversion involving a tomographic kernel to reconstruct the background model (Yao et al., 2020). Another is to use the complementary information of other geophysical data by joint inversion (Domenzain et al., 2020; Qin et al., 2022b). Since previous shallow-seismic FWIs at the same test site have shown the potential to provide long-wavelength seismic models (Pan et al., 2021; Irnaka et al., 2022), we will investigate the joint inversion of shallow-seismic and multi-offset surface GPR data in the future.

In some cases where the long edge of the entire model is not n_x but n_z, SFWI still works via exchanging the roles of n_{xi} and n_{zi} in this paper, where n_z and n_{zi} are the lengths of the model $\mathbf{m}$ and its subset $\mathbf{m}_i$ in z direction. For example, for the crosshole GPR data acquired in two 50 m deep boreholes (Pinard et al., 2016), SFWI has S^z_{2D} times acceleration and memory savings over FWI, where $S^z_{2D} = n_z/n_{zi}$. On the other hand, we can extend SFWI to the 3D case by selecting a model subset ($n_{mi} = n_{xi} \times n_{yi} \times n_z$) from an entire model ($n_m = n_x \times n_y \times n_z$), where n_y and n_{yi} are the lengths of $\mathbf{m}$ and $\mathbf{m}_i$ in y direction. In this case, we obtain the computational improvement $S^{xy}_{3D} = n_x n_y/(n_{xi} n_{yi})$ from equation 2.8. It can be seen that $S^{xy}_{3D} \geq S^x_{2D}$ because $n_y \geq n_{yi}$, which suggests that SFWI is more economical in 3D imaging

than in 2D imaging. This advantage makes 3D SFWI affordable even for personal computers that cannot perform 3D FWI when the model size is too large.

It is worth to mention that SFWI is compatible with the source-dependent method. Encoded simultaneous-source FWI (Huang and Schuster, 2012; Zhang et al., 2018) theoretically speeds up n_s/n_s^e times by encoding n_s gathers into n_s^e supergathers (ignoring the time of encoding and decoding process). The maximum speedup of this method is n_s when $n_s^e = 1$. If we combine SFWI with this method, each encoded source consists of n_s/n_s^e adjacent sources and is generated on the model subset, the theoretical speedup becomes Sn_s/n_s^e where S is calculated by equation 2.8. The combined speedup can be greater than n_s if $S/n_s^e > 1$. Such a combination overcomes the acceleration limits of source-dependent and model-dependent methods and is very promising for the application of 3D GPR imaging.

SFWI is also applicable to land streamer seismic data (Tran and Sperry, 2018; Hoang, 2021) and marine streamer seismic data (Choi and Alkhalifah, 2012; Huang and Schuster, 2012). Note in the subsurface and water, seismic waves (1∼100 Hz) attenuate less than EM waves (10∼1000 MHz) for the same propagation distance. To ensure a good data approximation, the source-boundary and receiver-boundary distances of the model subset should be wider than those used in GPR SFWI. Therefore, the performance increase for SFWI can be expected to be larger for GPR data than for seismic data, when considering an equal profile length. However, due to absence of surface waves and partially limited reflectors, the reconstruction of GPR parameters might not reach the same quality and has to be interpreted carefully.

2.6 Conclusion

We presented SFWI, a model-dependent method, to reduce the computational costs of inverting multi-offset GPR data. SFWI makes use of the high attenuation property of EM waves, which in most cases is regarded as a drawback of GPR, to approximately simulate data on a model subset rather than on an entire model. The decrease in model space not only saves memory usage for gradient calculation, but also speeds up the inversion process. These computational improvements can be quantified in theory by the size ratio of the model and its subset. Nevertheless, the actual improvements are related to the parallelization approaches used. SFWI performs faster than the theoretical value when combined with the source parallelization alone, as the lighter memory burden is favorable for more efficient computation. When both source parallelization and model domain parallelization are used, the performance of SFWI is slightly poorer than expected if the model subvolume is smaller than a certain size, i.e., when the communication time between subvolumes takes a greater percentage of the total runtime.

SFWI's performance depends on the selection of the model subset, which in 2D case is determined by the source-boundary distance, receiver-boundary distance, and maximum source-receiver offset. To ensure a good data approximation, we provide rules of thumb for the first two: the source-boundary distance should be close to one wavelength of the dominant frequency in the subsurface; the receiver-boundary distance needs to be only one grid width larger than the absorbing boundary. However, choosing the optimal maximum source-receiver offset relies on the subsurface environment and GPR antenna parameters

and requires individual testing. A fundamental requirement for the maximum source-receiver offset is to cover the effective illumination area of the acquisition geometry.

The 2D field examples showed that SFWI has the equivalent ability to FWI in reconstructing the permittivity and conductivity models and matching the observed data. In the field example of a 45 m long multi-offset surface GPR profile, SFWI obtained an acceleration greater than eight times when using a maximum source-receiver offset of 5 m. This application also revealed the importance of using the near-offset traces to image the near-surface targets. In the field example of a 21 m long multi-borehole plane, SFWI got more than three times speedup with a maximum source-receiver offset of 6.16 m. This approach has proven to be more economical than conventional FWI for 2D multi-offset GPR data in a wide range of acquisition geometries, and has greater computational advantages for both seismic and GPR data in 3D applications. To further improve its performance, our approach can be extended by other approaches, e.g., the encoded simultaneous source technique. However, the actual performance of SFWI in these cases requires further studies.

Acknowledgments

(This section will be hidden for submission.) This work is financially supported by the China Scholarship Council (No. 201806260258). Tan Qin would like to thank Tilman Steinweg and Mark Wienöbst for their help in developing the WAVE-Toolbox and thank Duy Hoang Hoang for his preliminary work on the code. The authors sincerely thank Lars Houpt, Felix Bögelspacher, Leon Merkel, Michael Mayer, Hagen Steger, Roland Helfer, Philipp Koyan, and the master students at Geophysical Institute, Karlsruhe Institute of Technology for their help in field data acquisition at the Rheinstetten test site.

Data and materials availability

An open-source software (GPL) package containing the source code, models, and data used in this paper is provided in the WAVE-Toolbox (https://github.com/WAVE-Toolbox).

3 Full-waveform inversion of ground-penetrating radar data in frequency-dependent media involving permittivity attenuation

Published as: Qin, T., Bohlen, T., & Allroggen, N. (2022). Full-waveform inversion of ground-penetrating radar data in frequency-dependent media involving permittivity attenuation. *Geophysical Journal International*, 232(1), 504-522.

Abstract

Full-waveform inversion (FWI) of ground-penetrating radar (GPR) data has received particular attention in the past decade because it can provide high-resolution subsurface models of dielectric permittivity and electrical conductivity. In most GPR FWIs, these two parameters are regarded as frequency independent, which may lead to false estimates if they strongly depend on frequency, such as in shallow weathered zones. In this study, we develop frequency-dependent GPR FWI to solve this problem. Using the τ-method introduced in the research of viscoelastic waves, we define the permittivity attenuation parameter to quantify the attenuation resulting from the complex permittivity and to modify time-domain Maxwell's equations. The new equations are self-adjoint so that we can use the same forward engine to back-propagate the adjoint sources and easily derive model gradients in GPR FWI. Frequency dependence analysis shows that permittivity attenuation acts as a low-pass filter, distorting the waveform and decaying the amplitude of the electromagnetic waves. The 2-D synthetic examples illustrate that permittivity attenuation has low sensitivity to the surface multioffset GPR data but is necessary for a good reconstruction of permittivity and conductivity models in frequency-dependent GPR FWI. As a comparison, frequency-independent GPR FWI produces more model artefacts and hardly reconstructs conductivity models dominated by permittivity attenuation. The 2-D field example shows that both FWIs reveal a triangle permittivity anomaly which proves to be a refilled trench. However, frequency-dependent GPR FWI provides a better fit to the observed data and a more robust conductivity reconstruction in a high permittivity attenuation environment. Our GPR FWI results are consistent with previous GPR and shallow-seismic measurements. This research greatly expands the application of GPR FWI in more complicated media.

3.1 Introduction

Ground-penetrating radar (GPR) plays an increasingly important role in near-surface surveys (Jol, 2008). Full-waveform inversion (FWI) was first proposed by Tarantola (1984) for seismic reflection data, and

then introduced to crosshole GPR data by Ernst et al. (2007b) and Kuroda et al. (2007). GPR FWI has been successfully applied to crosshole data (Ernst et al., 2007a; Meles et al., 2010; Oberröhrmann et al., 2013; Gueting et al., 2015) as well as to surface recordings (El Bouajaji et al., 2011; Busch et al., 2012; Lavoué et al., 2014; Liu et al., 2018). In the past decade, it has been shown that GPR FWI has great potential for reconstructing high-resolution subsurface models of electromagnetic (EM) material properties, namely dielectric permittivity and electrical conductivity (Klotzsche et al., 2019). In most GPR FWIs, these two parameters are assumed to be frequency independent, i.e. they are constant values over the main GPR bandwidth. However, many typical geological materials in the shallow subsurface are frequency-dependent to EM waves (Turner and Siggins, 1994). In this case, GPR FWI may not reveal its full potential if the forward modelling cannot account for velocity dispersion and permittivity attenuation.

The time-domain modelling of EM waves requires solving Maxwell's equations involving convolution calculations that explain the frequency-dependent electrical properties. For replacing the convolution with the more efficient multiplication, Carcione (1996) used the Debye model and Kelvin-Voigt model to approximate the relaxation functions of the dielectric permittivity and electrical conductivity, respectively. Bergmann et al. (1998) further developed Carcione's approach in the time domain and performed the numerical modelling by the finite-difference time-domain (FDTD) method. Their Maxwell's equations are analogous to the viscoelastic (or viscoacoustic) equations with standard linear solid (SLS) mechanisms, as shown, for example in Blanch et al. (1995). These equations use similar physical models to characterize the frequency dependence of electrical parameters and seismic moduli (Carcione and Cavallini, 1995). It implies that GPR FWI can refer to viscoelastic FWI already developed in the seismic community, such as Fabien-Ouellet et al. (2017) and Jiang (2019). However, Maxwell's equations differ from the viscoelastic equations because the complex permittivity and conductivity always occur in a combined form and both cause attenuation (Turner and Siggins, 1994; Bradford, 2007). In viscoelastic case, the magnitude of dispersion and attenuation of a single wave (P or S waves) can be described by a single parameter (τ_P or τ_S), which corresponds to the case of considering only permittivity attenuation or conductivity attenuation in EM waves (Blanch et al., 1995).

In the last decade, most GPR FWIs still used frequency-independent sensitivity kernels when updating model parameters (Klotzsche et al., 2019). There are two reasons for this. One is that simultaneous reconstruction of permittivity and conductivity is already a challenging task in frequency-independent GPR FWI, while studies of frequency-dependent GPR FWI that require estimation of more parameters take a back seat. Another reason is that frequency-independent GPR FWI, when combined with source signal estimation, can handle weak permittivity attenuation environments where quality factor (Q) ≥ 20 (Belina et al., 2012). In the study of Belina et al. (2012), since their model gradients were practically the same as those of the frequency-independent GPR FWI and the static conductivity was ignored, researchers did a permittivity-only inversion with a priori Q model. To better image the realistic materials with $Q < 20$ and investigate crosstalk between multiple parameters, frequency-dependent GPR FWI must be developed.

FWI is also known as an inverse scattering problem in the microwave community. Winters et al. (2006) proposed the time-domain inverse scattering technique to estimate the frequency-dependent average di-

electric properties, using a short relaxation time approximation in Debye scatterers. Based on their work, Papadopoulos and Rekanos (2011) introduced an auxiliary differential equation (ADE) with the polarization current density, which extended the feasibility of microwave imaging. Deng et al. (2021) reported the high-performance computation of EM FWI in the single-pole Debye model, using the theoretical basis of the inverse scattering technique. These studies suggested that we can similarly develop frequency-dependent GPR FWI. Nevertheless, they did not give an explicit measure of permittivity attenuation in EM wave propagation, which makes the inversion algorithm less intuitive than viscoelastic FWI involving seismic attenuation.

In this paper, we use the τ-method, with reference to the viscoelastic waves (Blanch et al., 1995), to quantify the attenuation of EM waves caused by the complex permittivity. Based on this, we modify the time-domain Maxwell's equations proposed by Bergmann et al. (1998) and implement the frequency-dependent GPR FWI. With a frequency dependence analysis, we show how each model parameter affects EM wave propagation, including amplitude attenuation and velocity dispersion. Finally, we apply 2-D GPR FWI in the synthetic examples and field examples and compare the performance of the FWI considering and not considering permittivity attenuation. Since surface-based GPR plays an important role for characterizing near-surface targets, we perform a surface-based GPR FWI in this study to investigate the applicability and limitations of using reflected waves.

3.2 Methodology

3.2.1 Forward problem

The propagation of EM waves in heterogeneous media can be expressed by Maxwell's equations as follows:

$$\begin{aligned} -\partial_t \mathbf{B} - \nabla \times \mathbf{E} &= 0, \\ -\nabla \times \mathbf{H} + \partial_t \mathbf{D} + \mathbf{J}_c &= -\mathbf{J}_e, \end{aligned} \tag{3.1}$$

with three generalized constitutive relations:

$$\begin{aligned} \mathbf{B} &= \mu \mathbf{H}, \\ \mathbf{D} &= \varepsilon * \mathbf{E}, \\ \mathbf{J}_c &= \sigma * \mathbf{E}, \end{aligned} \tag{3.2}$$

where $\mathbf{B}$ is the magnetic flux, $\mathbf{H}$ is the magnetic field, $\mathbf{D}$ is the electric displacement and $\mathbf{E}$ is the electric field. $\mathbf{J}_c$ is the conduction current density and $\mathbf{J}_e$ is the electric current sources. ∇ is the Laplace operator, $\times$ is the curl operator and $*$ is the time convolution. The magnetic permeability μ is essentially frequency independent, while the dielectric permittivity ε and electrical conductivity σ are described as complex frequency-dependent quantities (Carcione, 1996; Bergmann et al., 1998). The attenuation and dispersion processes of EM waves depend on the ε and σ of the media.

Substituting eq. (3.2) into eq. (3.1), Maxwell's equations become

$$\begin{aligned} -\mu\partial_t\mathbf{H} - \nabla\times\mathbf{E} &= 0, \\ -\nabla\times\mathbf{H} + \varepsilon * \partial_t\mathbf{E} + \sigma * \mathbf{E} &= -\mathbf{J}_e. \end{aligned} \tag{3.3}$$

In order to solve eq. (3.3) explicitly in the time domain, we must calculate the convolutions $\varepsilon * \partial_t\mathbf{E}$ and $\sigma * \mathbf{E}$. We use the Debye model (eq. 3.4) to approximate the relaxation function of the permittivity in the time domain (Carcione, 1996), i.e. the frequency-dependent function in the frequency domain.

$$\varepsilon(t) = \partial_t\Psi_\varepsilon(t) \quad \text{with} \quad \Psi_\varepsilon(t) = \varepsilon_s\left[1 - \frac{1}{L}\sum_{l=1}^{L}\left(1 - \frac{\tau_{El}}{\tau_{Dl}}\right)e^{-t/\tau_{Dl}}\right]H(t), \tag{3.4}$$

where τ_{El} and τ_{Dl} are the relaxation times of the electric field and displacement of the lth Debye model, respectively. $\tau_{El} < \tau_{Dl}$ for the lth mechanism. $\tau_{Dl} = 1/(2\pi f_l)$ where f_l is the relaxation frequency of the lth Debye peak and L is the number of relaxation mechanisms. $H(t)$ is the Heaviside function. ε_s is the static dielectric permittivity, corresponding to the permittivity in the frequency-independent media. Following the τ-method proposed by Blanch et al. (1995), we define the permittivity attenuation $\tau_\varepsilon = 1 - \tau_{El}/\tau_{Dl}$ as the attenuation level of the lth Debye model for EM waves. By doing so, Ψ_ε can be rewritten as

$$\Psi_\varepsilon(t) = \varepsilon_s\left(1 - \tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}e^{-t/\tau_{Dl}}\right)H(t). \tag{3.5}$$

In eq. (3.5), one can see that τ_ε is independent of the lth Debye model, which means the same τ_ε can be used for different Debye models. The newly defined variable is dimensionless ($0 \le \tau_\varepsilon < 1$) and more straightforward for identifying attenuation and dispersion caused by the complex permittivity. Using the Debye model, the convolution $\varepsilon * \partial_t\mathbf{E}$ becomes

$$\varepsilon(t) * \partial_t\mathbf{E} = \partial_t\Psi_\varepsilon(t) * \partial_t\mathbf{E} = \varepsilon_s(1-\tau_\varepsilon)\,\partial_t\mathbf{E} + \varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_{Dl}}\mathbf{E} + \sum_{l=1}^{L}\mathbf{r}_l. \tag{3.6}$$

Due to space limitations in the main content, we omit the computational process of convolution (readers can find details of eqs 3.6, 3.7, 3.10, and 3.11 in Appendix A). We introduce an ADE with a memory variable $\mathbf{r}_l$ of the lth mechanism corresponding to $\mathbf{E}$:

$$\partial_t\mathbf{r}_l = -\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^2}\mathbf{E} - \frac{1}{\tau_{Dl}}\mathbf{r}_l. \tag{3.7}$$

Note that the definition of $\mathbf{r}_l$ is slightly different from the memory variable in Carcione (1996) and Bergmann et al. (1998). Similar to viscoelastic wave equations described in Bohlen (2002), we incorporate ε_s into $\mathbf{r}_l$. By doing so, we can easily develop self-adjoint wave equations for GPR FWI (shown

in Section 3.2.2). If only one Debye model is used, the memory variable $\mathbf{r}_l$ is related to the polarization current density $\mathbf{J}$ used in the microwave imaging of Papadopoulos and Rekanos (2011) by

$$\mathbf{J} = \mathbf{r}_l + \frac{\Delta\varepsilon}{\tau_{Dl}}\mathbf{E} \quad \text{with} \quad \Delta\varepsilon = \varepsilon_s\tau_\varepsilon \text{ and } l = 1. \tag{3.8}$$

On the other hand, the relaxation function of the conductivity can be described by a Kelvin-Voigt type model (Pipkin, 2012) as

$$\sigma(t) = \partial_t\Psi_\sigma(t) \quad \text{with} \quad \Psi_\sigma(t) = \sigma_s\left[H(t) + \tau_\sigma\delta(t)\right], \tag{3.9}$$

where τ_σ is the relaxation time including the out-of-phase component of conductivity, σ_s is the static conductivity, and $\delta(t)$ is the Dirac function. Then, the convolution $\sigma * \mathbf{E}$ becomes

$$\sigma(t) * \mathbf{E} = \partial_t\Psi_\sigma(t) * \mathbf{E} = \sigma_s\left(\mathbf{E} + \tau_\sigma\partial_t\mathbf{E}\right). \tag{3.10}$$

Hence, we replace two convolutions with multiplications

$$\varepsilon(t) * \partial_t\mathbf{E} + \sigma(t) * \mathbf{E} = \varepsilon_\infty^e\partial_t\mathbf{E} + \sigma_\infty^e\mathbf{E} + \sum_{l=1}^{L}\mathbf{r}_l, \tag{3.11}$$

with the effective optical permittivity ε_∞^e and effective optical conductivity σ_∞^e as below:

$$\varepsilon_\infty^e = \varepsilon_s(1-\tau_\varepsilon) + \sigma_s\tau_\sigma, \quad \sigma_\infty^e = \sigma_s + \varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_{Dl}}. \tag{3.12}$$

The real effective permittivity and real effective conductivity with respect to the angular frequency ω are given by (Bergmann et al., 1998)

$$\begin{aligned}
\varepsilon^e(\omega) &= \varepsilon_s(1-\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{\omega^2\tau_{Dl}^2}{1+\omega^2\tau_{Dl}^2}) + \sigma_s\tau_\sigma, \\
\sigma^e(\omega) &= \sigma_s + \varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{\omega^2\tau_{Dl}}{1+\omega^2\tau_{Dl}^2}.
\end{aligned} \tag{3.13}$$

In fact, the effective optical parameters are the real effective parameters at infinite frequency. The attenuation factor α, phase velocity v and quality factor Q are described as (Turner and Siggins, 1994)

$$\begin{aligned}
\alpha &= \omega\left[\frac{\mu\varepsilon^e(\omega)}{2}\left(\sqrt{1+\tan^2\delta}-1\right)\right]^{1/2}, \\
v &= \left[\frac{\mu\varepsilon^e(\omega)}{2}\left(\sqrt{1+\tan^2\delta}+1\right)\right]^{-1/2}, \\
Q &= \tan^{-1}\delta,
\end{aligned} \tag{3.14}$$

where the loss tangent $\tan\delta = \sigma^e(\omega)/[\omega\varepsilon^e(\omega)]$. With $\tan\delta = 1$, we obtain the transition angular frequency $\omega_t = 2\pi f_t = \sigma^e(\omega_t)/\varepsilon^e(\omega_t)$ where f_t is the transition frequency. In frequency-independent media where $\tau_\varepsilon = 0$, σ^e_∞ can be simplified to σ_s. Therefore the attenuation is determined by σ_s alone, which is called conductivity attenuation. In frequency-dependent media, however, both σ_s and τ_ε contribute to attenuation. Based on eqs (3.7) and (3.11), we rewrite eq. (3.3) as

$$\begin{aligned} &-\mu\partial_t\mathbf{H} - \nabla\times\mathbf{E} = 0, \\ &-\nabla\times\mathbf{H} + \varepsilon^e_\infty\partial_t\mathbf{E} + \sigma^e_\infty\mathbf{E} + \sum_{l=1}^{L}\mathbf{r}_l = -\mathbf{J}_e, \\ &\mathbf{E} + \frac{L\tau_{Dl}^2}{\varepsilon_s\tau_\varepsilon}\partial_t\mathbf{r}_l + \frac{L\tau_{Dl}}{\varepsilon_s\tau_\varepsilon}\mathbf{r}_l = 0, \quad l = 1,...,L. \end{aligned} \tag{3.15}$$

We solve eq. (3.15) using the FDTD method of second order in time and fourth order in space. In the standard staggered grid of Yee (1966), the memory variable $\mathbf{r}_l$ locates at the same position of the electric field $\mathbf{E}$. The electrical parameters ε^e_∞, ε_s, σ^e_∞ and τ_ε are averaged to the location of $\mathbf{E}$, while the magnetic parameter μ is averaged to the location of the magnetic field $\mathbf{H}$. The convolutional perfectly matched layer is included at the model boundary to absorb the outgoing waves (Roden and Gedney, 2000).

Unlike eqs (6)–(8) in Papadopoulos and Rekanos (2011), eq. (3.15) uses the effective optical conductivity, resulting from different definitions of the polarization current density and memory variable (see eq. 3.8). Our modification in eq. (3.15) takes three advantages of the τ-method (Blanch et al., 1995). First, we use one τ_ε for all Debye models and set τ_{Dl} as a priori, which means less memory usage and computations in numerical modelling of EM waves than using τ_{Dl} and τ_{El} for L Debye models. Second, if $\sigma_s \ll \sigma^e_\infty$, we can simply approximate a constant Q by $Q \approx 2/\tau_\varepsilon$. Thus, the attenuation is similar to a linear function of frequency and the simulator can account for the waveform distortion of EM waves (see Section 3.3 for an example). Third, we use only one τ_ε in eqs (3.13) and (3.14) to quantify the magnitude of attenuation and dispersion introduced by complex permittivity, which is intuitive and reduces the number of parameters reconstructed by frequency-dependent GPR FWI in Section 3.2.2.

For convenience, we express eq. (3.15) in a matrix-vector formalism

$$\mathbf{M}_1\partial_t\mathbf{u} + \mathbf{M}_2\mathbf{u} - \mathbf{A}\mathbf{u} = \mathbf{s}, \tag{3.16}$$

with

$$
\begin{aligned}
&\mathbf{u} = (H_x, H_y, H_z, E_x, E_y, E_z, r_{x1}, r_{y1}, r_{z1}, ..., r_{xL}, r_{yL}, r_{zL})^T, \\
&\mathbf{s} = (0,0,0,-J_{ex},-J_{ey},-J_{ez},0,0,0,...,0,0,0)^T, \\
&\mathrm{diag}(\mathbf{M}_1) = \left(-\mu,-\mu,-\mu,\varepsilon^e_\infty,\varepsilon^e_\infty,\varepsilon^e_\infty,\frac{L\tau^2_{D1}}{\varepsilon_s\tau_\varepsilon},\frac{L\tau^2_{D1}}{\varepsilon_s\tau_\varepsilon},\frac{L\tau^2_{D1}}{\varepsilon_s\tau_\varepsilon},...,\frac{L\tau^2_{DL}}{\varepsilon_s\tau_\varepsilon},\frac{L\tau^2_{DL}}{\varepsilon_s\tau_\varepsilon},\frac{L\tau^2_{DL}}{\varepsilon_s\tau_\varepsilon}\right), \\
&\mathbf{M}_2 = \begin{pmatrix} 0_3 & 0_3 & 0_3 & \dots & 0_3 \\ 0_3 & \sigma^e_\infty I_3 & I_3 & \dots & I_3 \\ 0_3 & I_3 & \frac{L\tau_{D1}}{\varepsilon_s\tau_\varepsilon} I_3 & & \\ \dots & \dots & & \ddots & \\ 0_3 & I_3 & & & \frac{L\tau_{DL}}{\varepsilon_s\tau_\varepsilon} I_3 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 0_3 & D & 0_3 & \dots & 0_3 \\ D & 0_3 & 0_3 & \dots & 0_3 \\ 0_3 & 0_3 & 0_3 & \dots & 0_3 \\ \dots & \dots & & \ddots & \\ 0_3 & 0_3 & & & 0_3 \end{pmatrix}, \\
&D = \begin{pmatrix} 0 & -\partial_z & \partial_y \\ \partial_z & 0 & -\partial_x \\ -\partial_y & \partial_x & 0 \end{pmatrix} = D_1\partial_x + D_2\partial_y + D_3\partial_z,\ D_i^* = D_i^T = -D_i, \\
&D_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, D_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, D_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, D^T = D, \\
&D^* = (D_i\partial_i)^* = -D_i^*\partial_i = D_i\partial_i = D \quad \Rightarrow \quad \mathbf{A}^* = \mathbf{A},
\end{aligned}
\tag{3.17}
$$

where 0_3 is the 3×3 zero matrix and I_3 the 3×3 unit matrix. The superscript * is the transpose conjugate operator which is equivalent to transpose operator ^{T} for real variables. When deriving the transpose conjugate of $\mathbf{A}$, we use the zero-valued boundary condition, i.e. eqs (54) and (55) in Yang et al. (2016).

3.2.2 Inverse problem

The objective function Φ that we use in GPR FWI is

$$\Phi(\mathbf{m}) = \Phi(\mathbf{u}) = \frac{1}{2}||R\mathbf{u}(\mathbf{m}) - \mathbf{d}^{\mathrm{obs}}||_2^2 = \frac{1}{2}||\Delta\mathbf{d}||_2^2, \tag{3.18}$$

where the synthetic data are extracted from the forward wavefield $\mathbf{u}$ at the receiver position by the restriction operator R; $\Delta\mathbf{d} = R\mathbf{u}(\mathbf{m}) - \mathbf{d}^{\mathrm{obs}}$ is the residual between the synthetic data and observed data $\mathbf{d}^{\mathrm{obs}}$. The FWI seeks to iteratively reconstruct the model parameters $\mathbf{m}$ by minimizing Φ as follows:

$$\mathbf{m}_{k+1} = \mathbf{m}_k + \lambda\Delta\mathbf{m}_{k+1}, \tag{3.19}$$

$$\Delta\mathbf{m}_{k+1} = -\mathbf{P}\frac{\partial\Phi}{\partial\mathbf{m}} + \gamma\Delta\mathbf{m}_k, \tag{3.20}$$

where the step length λ is calculated by line search (Pica et al., 1990). Considering the different sensitivities of model parameters to the objective function, we use an individual step length for each parameter (Ernst et al., 2007b). The model update direction for the $(k+1)$th iteration $\Delta\mathbf{m}_{k+1}$ is computed by the conjugate-gradient method, where γ is the scale factor (Polak and Ribiere, 1969) and $\mathbf{P}$ is the preconditioner (Plessix and Mulder, 2004). We use a multiscale strategy to avoid cycle skipping in FWI (Bunks et al., 1995). In the following, we show how to calculate the model gradient $\partial\Phi/\partial\mathbf{m}$.

Due to that $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{A}$ are self-adjoint, we obtain the adjoint-state equations as follows (Plessix, 2006):

$$\begin{aligned} &-\mathbf{M}_1\partial_t\hat{\mathbf{u}}+\mathbf{M}_2\hat{\mathbf{u}}-\mathbf{A}\hat{\mathbf{u}}=-R^*\Delta\mathbf{d},\\ &\hat{\mathbf{u}}=\left(\hat{H}_x,\hat{H}_y,\hat{H}_z,\hat{E}_x,\hat{E}_y,\hat{E}_z,\hat{r}_{x1},\hat{r}_{y1},\hat{r}_{z1},...,\hat{r}_{xL},\hat{r}_{yL},\hat{r}_{zL}\right)^T, \end{aligned} \tag{3.21}$$

where the waveform residual $R^*\Delta\mathbf{d}$ is used as the adjoint sources for backpropagation and $\hat{\mathbf{u}}$ is the adjoint wavefield. In contrast to the forward problem, the computation of eq. (3.21) is done backward from the recording time T to 0. Hence we reverse the time by substituting $t^{'}=T-t$, $\hat{\mathbf{u}}^{'}(t^{'})=\hat{\mathbf{u}}(T-t)$ and $\Delta\mathbf{d}^{'}(t^{'})=\Delta\mathbf{d}(T-t)$ in the above equation and get the self-adjoint-state equations as below:

$$\begin{aligned} &\mathbf{M}_1\partial_{t'}\hat{\mathbf{u}}^{'}+\mathbf{M}_2\hat{\mathbf{u}}^{'}-\mathbf{A}\hat{\mathbf{u}}^{'}=-R^*\Delta\mathbf{d}^{'},\\ &\hat{\mathbf{u}}^{'}=\left(\hat{H}_x^{'},\hat{H}_y^{'},\hat{H}_z^{'},\hat{E}_x^{'},\hat{E}_y^{'},\hat{E}_z^{'},\hat{r}_{x1}^{'},\hat{r}_{y1}^{'},\hat{r}_{z1}^{'},...,\hat{r}_{xL}^{'},\hat{r}_{yL}^{'},\hat{r}_{zL}^{'}\right)^T. \end{aligned} \tag{3.22}$$

Then we cross-correlate the forward wavefield $\mathbf{u}$ and the back-propagating wavefield $\hat{\mathbf{u}}$ to compute the model gradient

$$\begin{aligned} \frac{\partial\Phi}{\partial\mathbf{m}}&=\int_0^T\hat{\mathbf{u}}(t)^*\left(\frac{\partial\mathbf{M}_1}{\partial\mathbf{m}}\partial_t\mathbf{u}(t)+\frac{\partial\mathbf{M}_2}{\partial\mathbf{m}}\mathbf{u}(t)\right)dt\\ &=\int_0^T\hat{\mathbf{u}}^{'}(T-t)^*\left(\frac{\partial\mathbf{M}_1}{\partial\mathbf{m}}\partial_t\mathbf{u}(t)+\frac{\partial\mathbf{M}_2}{\partial\mathbf{m}}\mathbf{u}(t)\right)dt. \end{aligned} \tag{3.23}$$

Equations (3.22) and (3.23) indicate two advantages of our modification in eq. (3.15). One advantage is that we can use the same forward solver for the forward and back-propagating wavefield simulations without any additional programming works. Another advantage is that all model parameters are distributed at the diagonal positions of $\mathbf{M}_1$ and $\mathbf{M}_2$, which means that the form of model gradients is very simple (see Appendix B for details).

Theoretically, in eq. (3.23), $\mathbf{m}$ can be any model parameter of eq. (3.15). However, we are only interested in the electrical parameters $\mathbf{m}=(\varepsilon_\infty^e,\sigma_\infty^e,\varepsilon_s,\tau_\varepsilon)$ in GPR FWI. According to eq. (3.12) and the chain rule, we convert the gradient of the objective function from the parameter class $\mathbf{m}=(\varepsilon_\infty^e,\sigma_\infty^e,\varepsilon_s,\tau_\varepsilon)$ to $\mathbf{m}^{'}=(\varepsilon_s^{'},\sigma_s,\tau_\varepsilon^{'},\tau_\sigma)$ where $\varepsilon_s^{'}=\varepsilon_s$ and $\tau_\varepsilon^{'}=\tau_\varepsilon$ (see Appendix B for details):

$$\frac{\partial\Phi}{\partial\mathbf{m}^{'}}=\frac{\partial\Phi}{\partial\mathbf{m}}\frac{\partial\mathbf{m}}{\partial\mathbf{m}^{'}}. \tag{3.24}$$

In this study, we use the real effective permittivity ε^e and real effective conductivity σ^e at the reference angular frequency ω_0 as the input and output parameters of GPR FWI, where ω_0 is set as the peak frequency of the source wavelet. Thus we ensure the physical consistency of the results obtained by

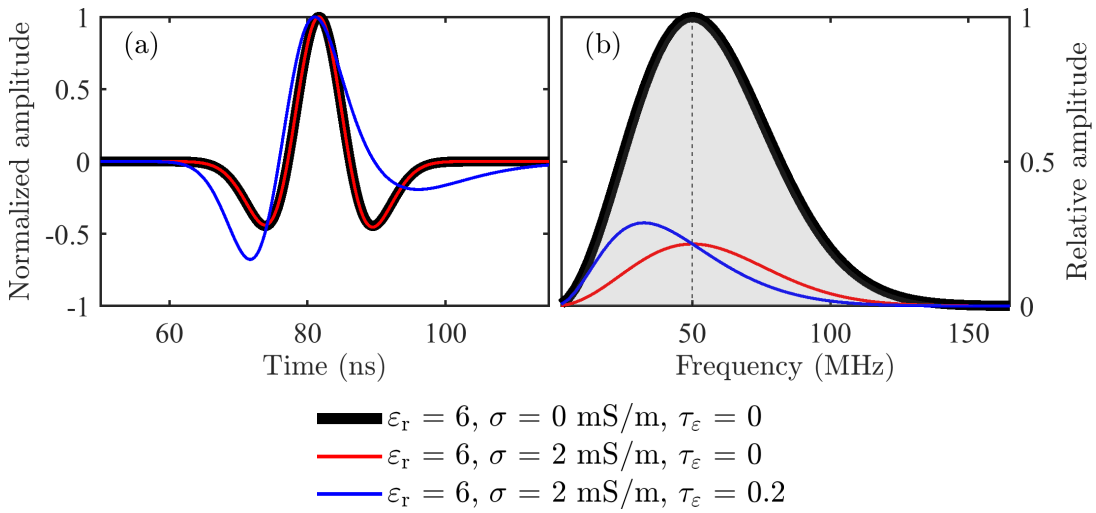

Figure 3.1: (a) Radargrams generated by a shifted 50 MHz Ricker source in 1-D homogeneous media and obtained by the receiver 10 m offset away from the source. The traces are normalized by their maximum amplitudes. (b) Corresponding amplitude spectra. The spectra are divided by the maximum amplitude of the thick black line. The translucent grey area represents the amplitude spectrum of the source wavelet, and the thin dashed line marks the peak frequency. One relaxation mechanism is used when $\tau_\varepsilon = 0.2$ (the blue line).

frequency-dependent GPR FWI and frequency-independent GPR FWI, similar to the correction for the phase velocity in viscoacoustic FWI (Kurzmann et al., 2013). In the following sections, $\varepsilon^e(\omega_0)$ and $\sigma^e(\omega_0)$ will be referred to as the permittivity ε (or relative permittivity ε_r) and the conductivity σ for convenience if not explicitly stated. In frequency-dependent GPR FWI, one can update the static parameters (ε_s and σ_s) using the gradients calculated by eq. (3.24), or update the effective parameters (ε and σ) by applying the chain rule again based on eq. (3.13) (see eq. B.8 for details).

3.3 Frequency dependence analysis

In order to analyse the effect of different model parameters on EM wave propagation, we make a comparison experiment using the 1-D analytical solution of EM wave in homogeneous media (Blanch et al., 1995). As shown in Fig. 3.1, we discuss three kinds of typical media in this section. All media have the same relative permittivity ($\varepsilon_r = 6$) so that the signals at the reference frequency propagate with the same phase velocity. The latter two attenuating media have the same conductivity ($\sigma = 2$ mS m^{-1}) in order to generate the same attenuation level at the reference frequency. We take the wave propagation in a non-attenuating medium as a reference (the thick black line). One relaxation mechanism is employed in frequency-dependent medium where $\tau_\varepsilon = 0.2$ and $\sigma_s = 0.1$ mS m^{-1}. Thus, according to eq. (3.13), the permittivity attenuation τ_ε dominates the conductivity (95 per cent) and contributes to the permittivity slightly (10 per cent) at the reference frequency. We set the relaxation time of the conductivity $\tau_\sigma = 0$ s in this paper so that we can focus more on the effect of τ_ε on the data. The value of τ_ε is chosen to approximate a constant Q (= 10) which has been observed in some geological materials (Turner and Siggins, 1994).

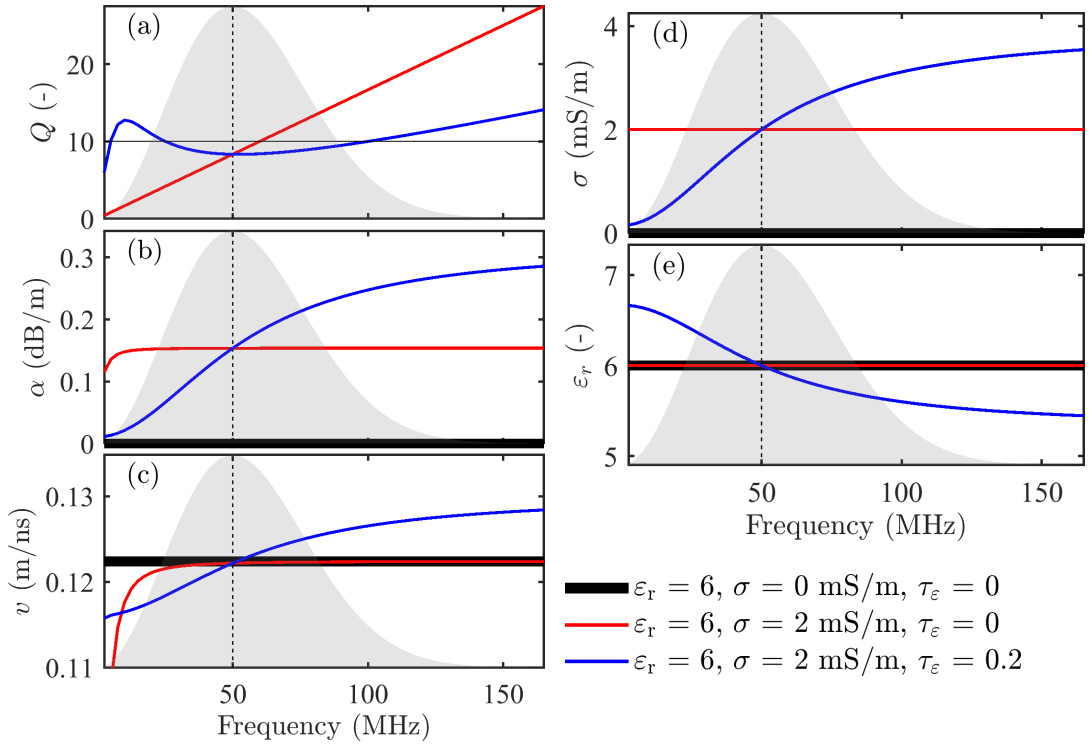

Figure 3.2: Frequency-dependent media characteristics. (a) Quality factor, (b) attenuation factor, (c) phase velocity, (d) real effective conductivity and (e) real effective permittivity. The translucent grey area represents the amplitude spectrum of the source wavelet, and the thin dashed line marks the peak frequency. The thin black line in (a) shows the desired Q value ($Q = 10$). Note that Q is infinite for the non-attenuating medium (the thick black line) and therefore not displayed in (a).

Compared to the reference medium, the attenuating media lead to a significant decrease in amplitude in Fig. 3.1(b). We observe different waveform and amplitude decreases in the frequency-dependent medium (the blue line) where τ_ε works as a low-pass filter, with more attenuation at frequencies above the peak frequency and less attenuation in the other frequency ranges. This is presented as a shift of the amplitude spectrum towards lower frequencies in Fig. 3.1(b), corresponding to the waveform deformation in Fig. 3.1(a). In contrast, the conductivity attenuation medium scales the amplitude of different frequencies equally, without changing the waveform shape.

To find the reasons of these phenomena, we show the phase velocity v, attenuation factor α and quality factor Q with respect to frequency in Figs 3.2(a)–(c). Those three characteristics are computed from the real effective parameters shown in Figs 3.2(d) and (e) through eq. (3.14). The reference medium has infinite Q value, no attenuation, and a constant velocity. Along with the dominant frequency range of the source wavelet (25–80 MHz), the conductivity attenuation medium has a linear increase of the quality factor, almost the constant attenuation level, and similar velocity with the reference medium due to that the transition frequency ($\approx$ 6 MHz) is much smaller than the peak frequency of the source wavelet (50 MHz). That results in almost identical travel time and amplitude changes of different frequency components in Fig. 3.1. Unlike the frequency-independent media, the frequency-dependent medium causes both phase velocity and attenuation factor to increase with frequency in the dominant frequency

Table 3.1: Background model parameters.

Layer	Thickness (m)	ε_r (-)	σ (mS m^{-1})	τ_ε (-)
Air	2	1	0	0
Soil	6	6	2	0.1
Rock	2	9	3	0.1

region, which is the reason for distorting the waveform and amplitude spectrum in Fig. 3.1. Although we use only one relaxation mechanism, the Q approximated is close to the desired value. With more relaxation mechanisms, the Q approximation can be further improved. As shown in Figs 3.2(d) and (e), dielectric permittivity and electrical conductivity are frequency-independent in the reference medium and conductivity attenuation medium and become frequency-dependent in the permittivity attenuation medium. When the frequency is greater than 25 MHz, the conductivity is proportional to the attenuation factor, and the permittivity is inversely proportional to the square of phase velocity.

3.4 Inversion of synthetic data

We use several synthetic examples of 2-D models ($x-z$ plane) to analyse the performance of frequency-dependent GPR FWI. Transverse magnetic (TM) mode waves are used in the simulation and inversion of GPR surface recordings. In frequency-independent media, there are wavefield components H_x, H_z and E_y and reconstructed model parameters ε and σ. In frequency-dependent media, additional memory variable r_{yl} and reconstructed model parameters τ_σ and τ_ε are added with $L = 1$. The magnetic permeability μ is constant and equal to its value of vacuum. We set $\tau_\sigma = 0$ s based on the analysis in the previous section.

The EM model space shown in Fig. 3.3 is 10 m $\times$ 36 m, and the grid spacing is 0.1 m. The three-layer background parameters are set according to Table 3.1. As the 2-m-thick air layer above the ground is not updated during the inversion, we do not show it in the figures of this paper. The true model for the synthetic examples consists of the background model and triangular perturbations, where the permittivity perturbations are relatively weak, so that we can better study the reconstruction of the conductivity and permittivity attenuation models in the inversion. Receivers are placed on the ground at 0.2 m intervals to record the radargrams of the electric field component with a time sampling of 0.15 ns and a time window of 300 ns. Eighteen transmitters generate the electric field at 2 m intervals on the ground (see red stars in Fig. 3.3). The source is a shifted Ricker wavelet with a centre frequency of 50 MHz. For each source, the receivers record data with an offset of 1–20 m. The observed data are simulated in the true model and are the same for the frequency-dependent GPR FWI and frequency-independent GPR FWI ($\tau_\varepsilon = 0$). The minimum wavelength of the EM waves observed in the soil layer is about 1 m.

We use the background model (Table 3.1) as the initial model to perform FWI. The model parameters are updated simultaneously in multiparameter inversion examples. Apart from the pre-conditioner shown in eq. (3.20), we also multiply the gradient by a user-defined taper to mitigate source and receiver artefacts. We use a multiscale strategy with five inversion stages and up to 15 iterations per stage. From the first stage to the fifth stage, we sequentially implement frequency band variations from 5 to 15, 25, 35, 50,

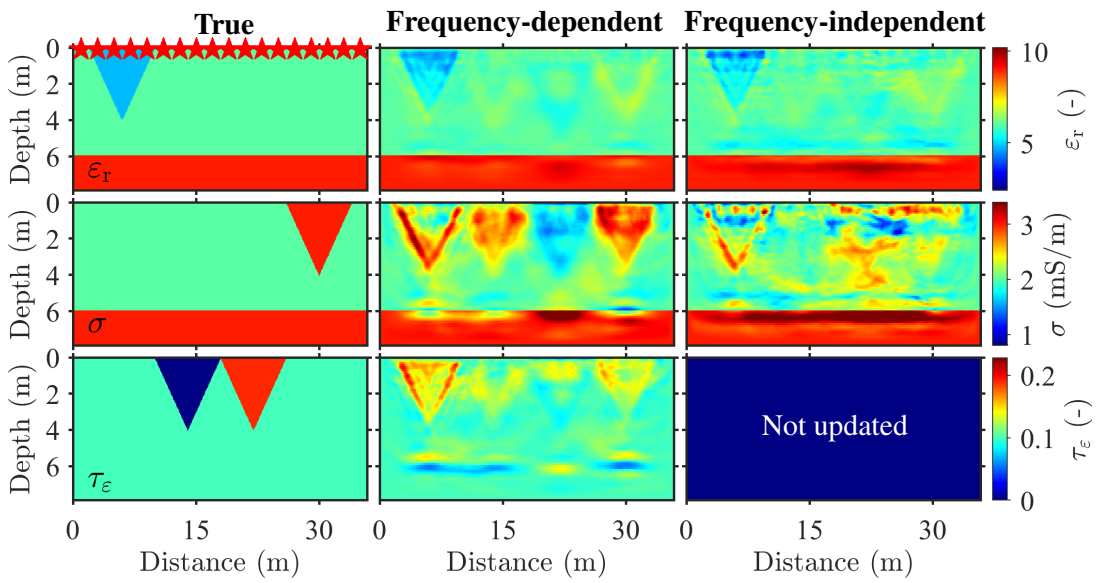

Figure 3.3: Models of the three-parameter synthetic example where all anomalies are spatially uncorrelated. The three columns are the true models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI results (ε_r and σ), respectively. The red stars on the true model are the transmitters.

and 80 MHz in a Butterworth bandpass filter. At the beginning of each stage, we apply the source-time function (STF) inversion (Groos et al., 2014). If the relative misfit change between the current iteration and the previous second iteration is less than 1 per cent, the inversion will switch to the next stage or stop if it is the last stage.

3.4.1 Uncorrelated model

Noise-free data

We build a true model consisting of the background model and several trench anomalies (triangles) spatially uncorrelated on each model. In Fig. 3.3, variations in the static parameters lead to anomalies in permittivity and conductivity models. To insulate the permittivity and conductivity models from permittivity attenuation anomalies, we adjust these static parameters at locations where τ_ε anomalies exist. The strong crosstalks between ε_r, σ, and τ_ε occur in the frequency-dependent GPR FWI results. In the conductivity model, we observe original conductivity anomaly and severe footprints generated by the anomalies of permittivity and τ_ε. Frequency-independent GPR FWI suffers even more from the crosstalks, which makes the original conductivity anomaly indistinguishable. In frequency-independent GPR FWI, more artefacts appear in the non-anomalous area of the permittivity and conductivity models because of the absence of permittivity attenuation. Although permittivity attenuation is weakly sensitive to the data, it is helpful for the correct reconstruction of the conductivity anomaly (the last trench) in the frequency-dependent GPR FWI.

To further investigate the performance of two FWIs, we compare the observed data and inverted data in Figs 3.4(a) and (b). In the observed data, the ground wave travels faster in the first source than in the last

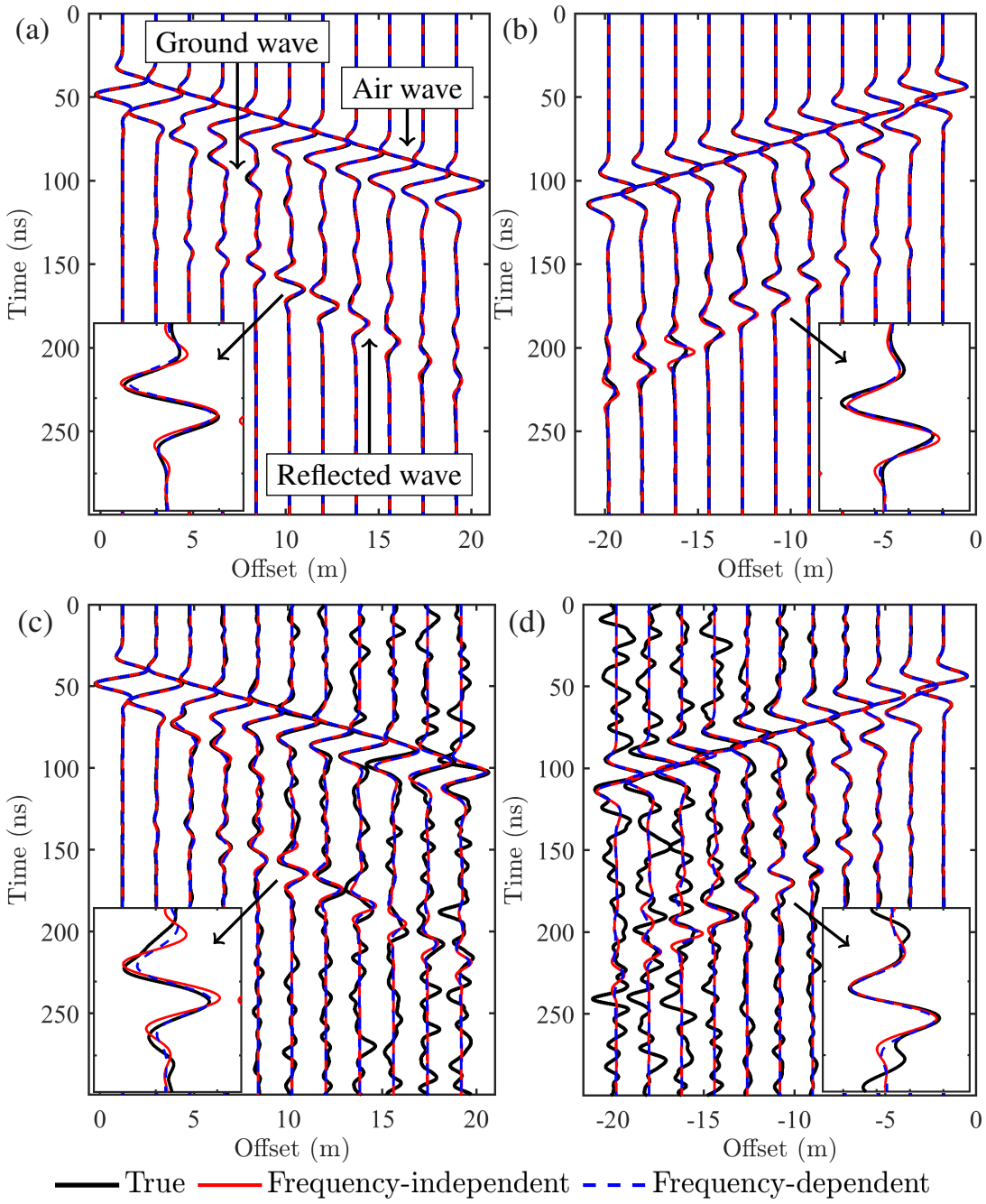

Figure 3.4: Data fitting of the three-parameter synthetic example where all anomalies are spatially uncorrelated. Gaussian noise (SNR = 20 dB) is added to the observed data in (c) and (d). (a and c) and (b and d) are the radargrams of the first and last sources, shown once every eight traces. Each trace is divided by the maximum amplitude of each trace of the observed data. The rectangular windows show the zoomed waveforms. A Butterworth bandpass filter (5–80 MHz in the fifth stage) is applied to the radargrams.

source due to the permittivity anomaly on the left side. The reflected wave of the first source propagates along the low permittivity attenuation anomaly ($\tau_\varepsilon = 0$) and shows a lower amplitude than that of the last source at an offset of 10–20 m, similar to the observation in Fig. 3.1(b). In GPR FWI, the STF inversion can provide an 'effective' source wavelet that is low-pass filtered (Belina et al., 2012). Thus it can account for the frequency-dependent effects of EM waves in weak permittivity attenuation environments ($Q \geq 20$ and $\tau_\varepsilon \leq 0.1$). Nevertheless, after travelling through the high permittivity attenuation regions, such an 'effective' source may generate waveforms different from the observed data, e.g., the reflected wave with offsets greater than 10 m in Fig. 3.4(b). As a consequence, frequency-independent GPR FWI attributes the waveform differences to perturbations throughout the model space, resulting in some unwanted artefacts in Fig. 3.3. On the contrary, frequency-dependent GPR FWI can reconstruct trench anomalies in the permittivity and conductivity models and reduce artefacts due to its ability to describe permittivity attenuation. Frequency-dependent GPR FWI, therefore, agrees highly with the observed data. This example illustrates that the combination of frequency-independent GPR FWI and STF estimates can only partially explain the waveform distortions caused by permittivity decay. It is necessary to consider frequency-dependent GPR FWI when reconstructing the conductivity model.

Noise-contaminated data

In the synthetic example above, the observed data are free of noise. It is unrealistic and may lead us to overestimate the performance of the GPR FWI approaches. Therefore, starting from this section, we add Gaussian noise with a signal-to-noise ratio (SNR) of 20 dB, with respect to the strongest amplitude, to the observed data. As a result, the reflected waves in the observed data shown in Figs 3.4(c) and (d) are heavily disturbed and below the noise level after a 15 m offset. We repeat the synthetic example using noisy data and present the inversion results in Fig. 3.5.

Compared to Fig. 3.3, we see fewer model updates of the two FWIs in Fig. 3.5. It is reasonable as noise slows down the convergence of data misfits. On the one hand, the reconstructions of the four trenches more or less deteriorate. On the other hand, the crosstalks and artefacts are also suppressed. Consequentially, frequency-independent GPR FWI shows clearer results of the conductivity model where the original conductivity anomaly becomes distinguishable. Fig. 3.4 also indicates that the existence of noise increases the difficulty of data fitting. In the last radargram, the two FWIs using noisy data have fewer differences in the waveform than those using noise-free data. To some extent, noise stabilizes frequency-independent GPR FWI and allows it to converge to better results than the noise-free case. Overall, frequency-dependent GPR FWI still possesses fewer data misfits and model artefacts than frequency-independent GPR FWI when the observed data are contaminated by noise.

3.4.2 Correlated model

Three-parameter model

In the true model of this example, the perturbations of the permittivity and conductivity models are located at the same position and have the same values in all trenches. However, the static permittivity and

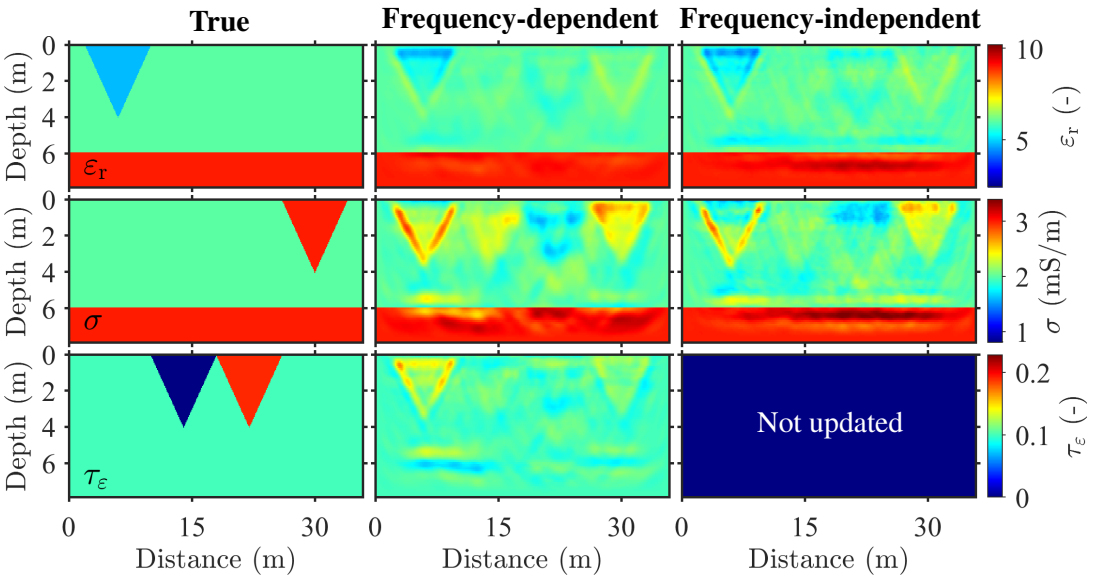

Figure 3.5: Models of the three-parameter synthetic example where all anomalies are spatially uncorrelated. Gaussian noise (SNR = 20 dB) is added to the observed data. The three columns are the true models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI results (ε_r and σ), respectively.

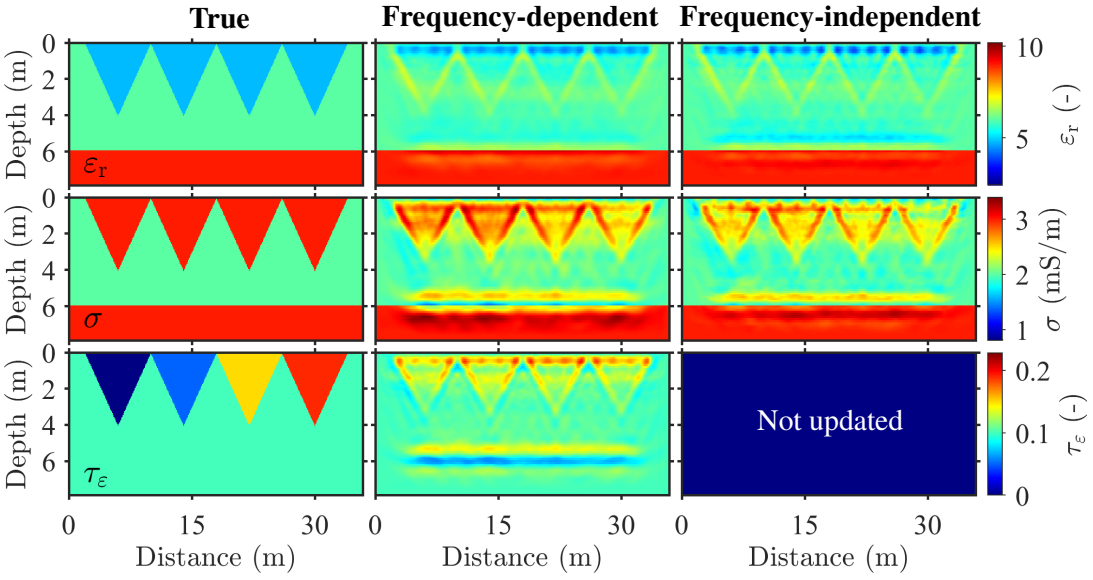

Figure 3.6: Models of the three-parameter synthetic example where all anomalies are spatially correlated. Gaussian noise (SNR = 20 dB) is added to the observed data. The three columns are the true models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI results (ε_r and σ), respectively.

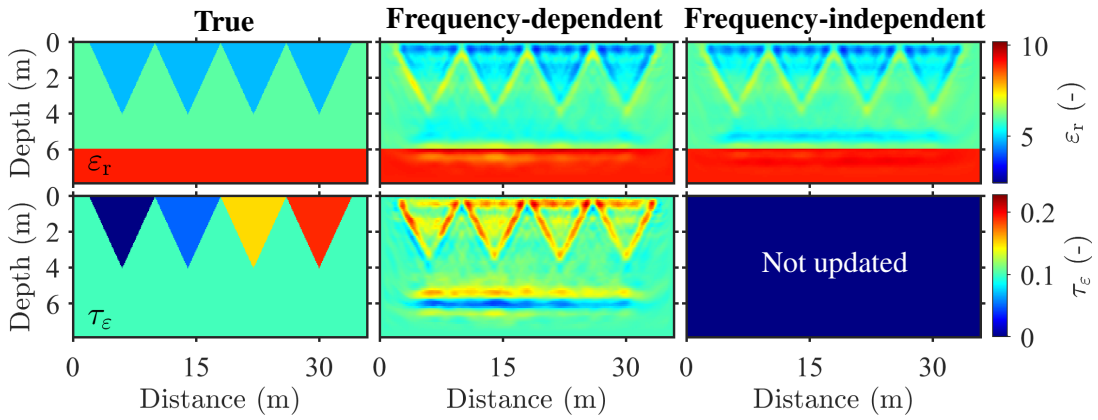

Figure 3.7: Models of the two-parameter synthetic example where all anomalies are caused by the static permittivity ε_s and permittivity attenuation τ_ε. Gaussian noise (SNR = 20 dB) is added to the observed data. The three columns represent the true models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI result (ε_r), respectively.

static conductivity in each trench are different due to the variations of the permittivity attenuation model. In Fig. 3.6, we see that frequency-dependent GPR FWI and frequency-independent GPR FWI have similar performance in terms of reconstructing the permittivity model. The estimation of the conductivity perturbation deteriorates from left to right as permittivity attenuation increases. Besides, frequency-dependent GPR FWI shows more in-trench reconstructions than frequency-independent GPR FWI. In the frequency-independent GPR FWI results, the upper boundary of the last conductivity trench becomes discontinuous and only the two sides can be distinguished. On the contrary, frequency-dependent GPR FWI still reconstructs the four conductivity trenches with good resolution, although the first two trenches in the permittivity attenuation model are incorrectly estimated.

Two-parameter model

We perform another two tests to investigate the crosstalk of different parameters in the GPR FWI. Fig. 3.7 shows the same true models of ε and τ_ε as Fig. 3.6. The static conductivity model is adjusted so that the conductivity model is unchanged. The increasing values of τ_ε from the left to the right trench means that the percentage of τ_ε in permittivity increases from 0 to 10 per cent. We observe heavy crosstalks between τ_ε and ε in the results of frequency-dependent GPR FWI. The four permittivity attenuation trenches are described as high-value anomalies, and the four permittivity trenches are recovered to different degrees. For example, the last trench is reconstructed better than the first one, even though they have the same value in the true model. This should attribute to the more realistic dispersion and attenuation in the last trench. These differences also occur in the permittivity result of frequency-independent GPR FWI. Since the small effect of τ_ε on the phase velocity (see Fig. 3.2c), the permittivity results obtained from frequency-dependent GPR FWI has only a slight improvement compared to that from frequency-independent GPR FWI (Fig. 3.7).

When it comes to the synthetic example of σ and τ_ε shown in Fig. 3.8, we adapt the static permittivity to make permittivity the same as the background values. Although four conductivity anomalies look the

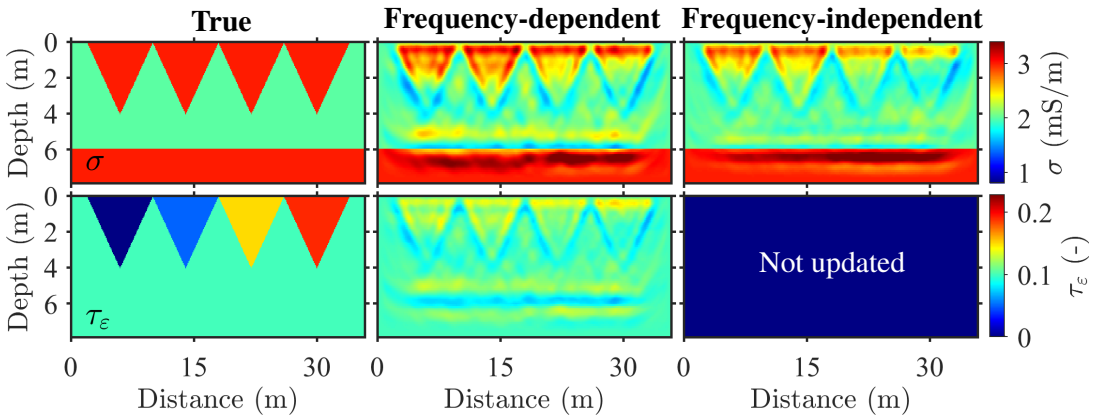

Figure 3.8: Models of the two-parameter synthetic example where all anomalies are caused by the static conductivity σ_s and permittivity attenuation τ_ε. Gaussian noise (SNR = 20 dB) is added to the observed data. The three columns represent the true models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI result (σ), respectively.

same in the true model, they are combinations of different static conductivity and permittivity attenuation. The percentage of τ_ε in these conductivity anomalies increase from 0 in the first one to 75 per cent in the last one. We observe severe interference of conductivity on permittivity attenuation in the results of frequency-dependent GPR FWI, where all τ_ε anomalies are reconstructed as high values. Similar to that shown in Fig. 3.6, it is difficult for frequency-independent GPR FWI to recover the conductivity anomalies dominated by permittivity attenuation. In the permittivity attenuation model reconstructed by frequency-dependent GPR FWI, the conductivity model introduces crosstalks that become weaker with depth (Fig. 3.8), while the permittivity model causes crosstalks around the trench boundaries and soil-rock interface (Fig. 3.7). It is due to different sensitivity kernels of permittivity and conductivity in FWI (Meles et al., 2011). Therefore, in Fig. 3.6, we observe a superposition of these effects when reconstructing three parameters simultaneously.

3.5 Inversion of field data

3.5.1 Data acquisition and pre-processing

We conducted field measurements at the northeast corner of the gliding airfield in Rheinstetten, Germany. This test site is well known from previous GPR and seismic studies (Schaneng, 2017; Wegscheider, 2017; Pan et al., 2018, 2021; Wittkamp et al., 2019; Gao et al., 2020; Irnaka et al., 2022). It is covered by sediments consisting of gravel and sand from the Rhine river. The ground layer is composed of partially saturated soil. At the test site, a defensive "V" shaped trench named Ettlinger Line (dashed triangle in Fig.3.9) was built in the early 17th century and has been refilled and is now completely flattened to the surface. The existence and shape of the Ettlinger Line have been delineated in detail via 3-D GPR migration imaging (Wegscheider, 2017), 3-D Rayleigh-wave dispersion inversion (Schaneng, 2017; Pan et al., 2018), 2-D joint elastic FWI of Rayleigh and Love waves (Wittkamp et al., 2019), 2-D viscoelastic

Table 3.2: Acquisition parameters of the surface GPR data (200 MHz) in the Rheinstetten test site and those used within the FWI.

Parameters	Raw	FWI
Number of sources	165	18
Traces per gather	56–125	100–175
Transmitter spacing	$\sim$0.2 m	2 m
Receiver spacing	$\sim$0.1 m	0.04 m
Minimum offset	0.2 m	0.3 m
Maximum offset	17 m	8 m
Sample rate	0.2 ns	0.08 ns
Recording window	200 ns	164 ns

Table 3.3: Surface GPR data (200 MHz) pre-processing steps.

(1)	Data resampling in the frequency domain
(2)	Interpolation of clipped direct-arrival amplitudes
(3)	DC-shift removal and dewow
(4)	Bandpass filtering (5–400 MHz)
(5)	Bad traces removal and offset limitation
(6)	Data gridding in the time-offset domain
(7)	3-D / 2-D transformation

FWI of Rayleigh waves (Gao et al., 2020), and 3-D viscoelastic FWI of the surface waves (Pan et al., 2021; Irnaka et al., 2022).

The surface GPR profile is positioned perpendicular to the Ettlinger Line. The two ends of the profile (from southwest to northeast) are in the same location as 'C' and 'B' in fig. 1(a) in Pan et al. (2018). The acquisition settings for the surface GPR data are listed in Table 3.2. Our GPR data were recorded using a single-channel pulseEKKO Pro GPR system equipped with a pulseEKKO Ultra receiver. The Ultra receiver stacked the records 256 times to obtain a higher SNR by reducing the random noise. The nominal centre frequency of the transmitter is 200 MHz. We deployed the transmitter-receiver orientation in HH mode to acquire TM wave data. The receiver was mounted on a sledge for smooth movement and tracked at the centimetre level by employing a real-time kinematic (RTK) positioning using a self-tracking total station as presented by Boniger and Tronicke (2010). To obtain multioffset GPR data for one source, we fixed the transmitter and moved the receiver towards the transmitter. Then we changed the transmitter location and moved the receiver away from the transmitter to produce the next gather. It took us two minutes to record one radargram and six hours for all 165 radargrams. Our measurements were slower than those of Lavoué (2014) who extracted hundreds of multioffset gathers from 15 common-offset gathers with an offset interval of 0.5 m. However, each trace in the multioffset GPR data used in Lavoué (2014) might be generated by a different transmitter-ground coupling. In contrast, our measurement had the same transmitter-ground coupling in one gather and had denser receiver intervals ($\sim$0.1 m).

In order to apply the FWI, we have to pre-process the GPR data first. The steps for data pre-processing are listed in Table 3.3, similar to those used in Domenzain et al. (2021). Due to the uneven walking

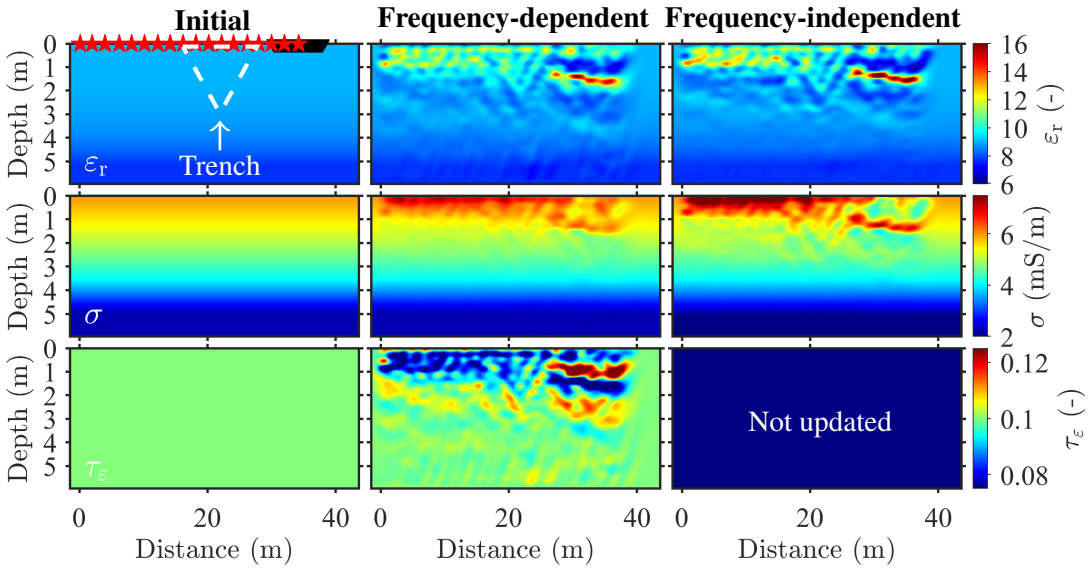

Figure 3.9: Models of the field data example in the Rheinstetten test site. The three columns are the initial models, the frequency-dependent GPR FWI results and the frequency-independent GPR FWI results, respectively. The white dashed triangle in the initial model outlines the target trench, known as the Ettlinger Line. The red stars are the transmitters, and the dense black triangles are the receivers of the 16th transmitter.

speed of the worker when moving the sledge, the data we acquired has irregular trace spacing. To ensure a balanced illumination in the measurement area, we apply the data gridding, that is 2-D spline interpolation in the time-offset domain at regular trace spacing. Finally, we transform the data acquired in the 3-D world into 2-D line-source data because we use a 2-D forward solver in GPR FWI. We use the transformation of the reflected waves to correct for phase and amplitude differences between the 3-D and 2-D data (Forbriger et al., 2014).

3.5.2 Data inversion

We build an initial model (7 m $\times$ 45.2 m with a grid spacing of 0.04 m) in Fig. 3.9, which does not show the air layer of 1 m thickness. The topographical variations along the survey line are minor and therefore ignored. The initial relative permittivity is 9 at the ground (the velocity of ground wave is 0.1 m ns^{-1}) and decreases slightly to 8 at a depth of 6 m. On the other hand, the initial conductivity is 6 mS m^{-1} at the ground and decreases gradually to 2 mS m^{-1} at a depth of 6 m. The initial permittivity attenuation is 0.1. We use similar inversion settings as the synthetic examples presented in the previous section, except that the frequency bands vary from 5 to 30, 40, 50, 70, and 100 MHz. These bands are chosen to avoid cycle skipping in GPR FWI since our initial model is relatively simple (Bunks et al., 1995). In frequency-dependent GPR FWI, the relaxation frequency of the Debye model f_l is 50 MHz and the reference angular frequency $\omega_0 = 2\pi f_l$, ensuring that the conclusions we drew in Sections 3.3 and 3.4 are still applicable to the field data example. To save computational time, we select 18 gathers for FWI with a source interval of 2 m (see Table 3.2).

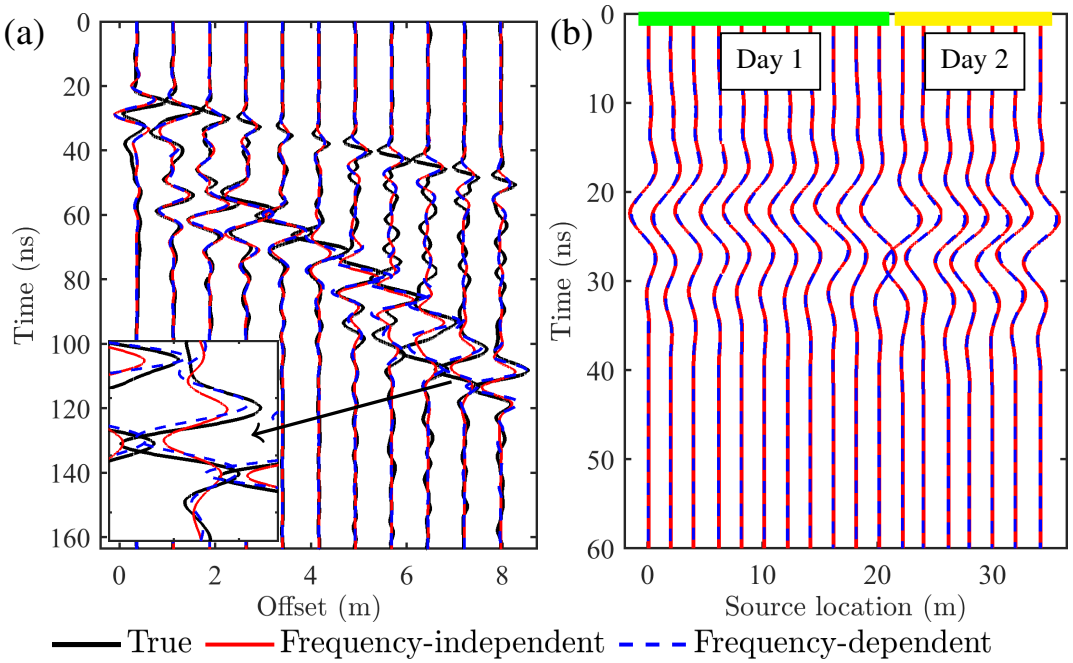

Figure 3.10: (a) Data fitting of the 16th radargram of the field data example in the Rheinstetten test site, shown once every 20 traces. Each trace is divided by the maximum amplitude of each trace of the observed data. The rectangular window shows the zoomed waveforms. (b) Estimated source signals of two FWIs. The green and yellow lines mark the sources used on different days. A Butterworth bandpass filter (5–100 MHz in the fifth stage) is applied to the radargrams.

The reconstructed permittivity models of the two FWIs in Fig. 3.9 illustrate the presence of a triangular anomaly, the Ettlinger Line, which is in high agreement with the 3-D GPR migration results of Wegscheider (2017). On the right of the trench, we observe a strong permittivity contrast (a slightly right-tilted reflector) at 1.1–1.5 m depth, described as a low *S*-wave anomaly in the 3-D shallow-seismic FWI results of Pan et al. (2021) and Irnaka et al. (2022). Unlike the shallow-seismic FWI with a resolution of 1 m, GPR FWI has a much higher resolution ($\sim$0.25 m). Consequently, the reflector appears to be connected to the trench in the GPR FWI results, whereas their connection is more ambiguous in the shallow-seismic FWI results. One cannot observe the same reflector on the left side of the trench because the ground surface on this side was higher than on the right side and has been levelled [see schematics in fig. 3 in Irnaka et al. (2022)]. This implies that the reflector might be the original ground surface. Besides, both FWIs reveal a conductive layer near the ground and ranging from a thickness of 1.0 m on the left to 1.5 m on the right. More importantly, frequency-dependent GPR FWI reconstructs a permittivity attenuation model consistent with the permittivity model, where the Ettlinger Line is also visible.

The main differences between the two FWIs come from the right part of the highly conductive layer. This part is discontinuous and more similar to the permittivity model in frequency-independent GPR FWI. To understand these differences, we show the radargrams of the 16th source and the estimated source signals in Fig. 3.10. For both FWIs, the air and ground waves are difficult to fit because we use a 3-D / 2-D transformation of the reflected wave, and the line source cannot describe the radiation patterns and antenna-ground coupling in the real world. In the 16th radargram, the reflected waves

become dominant at offsets greater than 2 m due to the strong permittivity contrast on the right side of the trench. Frequency-independent GPR FWI matches well to the reflection events with offsets shorter than 4 m, beyond 4 m its performance degrades. Frequency-dependent GPR FWI better fits the amplitude of reflection events at offsets greater than 4 m. It results from the high permittivity attenuation layer on the right of the trench, similar to the case in Fig. 3.4. The two FWIs have similar final data misfits (difference less than 1 per cent) due to strong decay characteristics of the surface recordings along the offset direction. Nevertheless, from the perspective of reflection fitting, frequency-dependent GPR FWI performs better and thus probably produces a more reliable conductivity model. Note that we acquired the first 11 radargrams on the first day and the last 7 radargrams on the second day. The estimated source signals of the first 11 sources are different from the last 7 sources in Fig. 3.10(b), probably due to coupling differences caused by slight near-surface moisture variations resulting from drying or wetting at night. However, the source signals used on the same day show similar waveform shapes and travel times, which indicates the stability of the two FWIs and source wavelet estimation.

3.6 Discussion

In the examples of frequency dependence analysis and inversion, we ignore the relaxation time of conductivity ($\tau_\sigma = 0$ s) and use only one relaxation mechanism ($L = 1$) for costing purposes. The importance of considering $\tau_\sigma > 0$ s and $L > 1$ deserves further study. In the field example, the reconstructed permittivity attenuation model delineates the subsurface targets and provides a meaningful geological interpretation. Its high heterogeneous distribution also demonstrates the necessity of including it in the inversion. Based on our observation, we regard it still to be challenging to accurately estimate the permittivity attenuation model due to its weak sensitivity to the data and the crosstalk between multiple parameters. To better image the permittivity attenuation, two strategies are available. One strategy is to decouple the radiation patterns of the reconstructed parameters. It requires choosing a suitable parametrization so that different gradients have opposite behaviour with respect to azimuth (Yao et al., 2018). Another strategy is to use the Hessian operator computed by Newton's method in optimization (Métivier et al., 2013; Operto et al., 2013; Gao et al., 2021). Truncated Newton method has been employed in frequency domain GPR FWI by Pinard et al. (2015). When it is introduced to the time domain GPR FWI, we have to find a balance between the elimination of crosstalk and the increase in computational cost, which needs further investigation.

Previous studies have shown the potential of joint inversion because other data can provide some complementary information to the GPR inversion (Linde et al., 2008; Domenzain et al., 2020; Qin et al., 2022b). For example, the electrical resistivity (ER) data are sensitive to electrical conductivity. Therefore, the joint inversion of GPR and ER data may help to improve the reconstruction of conductivity (Domenzain et al., 2021). Note that the estimated conductivity is frequency-dependent in our GPR FWI, but frequency-independent in the ER inversion. Transformation of the two can be achieved by eq. (3.13), but depends on the reliability of the permittivity attenuation reconstruction, which requires further improvement in the future.

3.7 Conclusion

In this paper, we quantified permittivity attenuation for EM wave propagation simulator using the τ-method introduced from the seismic community, which not only saves memory and reduces computations in forward modelling but also simplifies the constant Q approximation and frequency-dependent GPR FWI. By defining the parameter τ_ε, we proposed a new form of time-domain Maxwell's equations to describe the propagation of EM waves in frequency-dependent media. These equations have two advantages in GPR FWI. The first one is their self-adjoint property, which allows performing frequency-dependent GPR FWI without changing the backpropagation engine. The second advantage is that the model gradients derived from these equations are fairly simple, ensuring as few modifications as possible to the existing frequency-independent GPR FWI code. Frequency dependence analysis revealed that, in the GPR spectrum range, the attenuation caused by the static conductivity is frequency-independent, while that caused by τ_ε is frequency-dependent. The permittivity attenuation acts as a low-pass filter and leads to waveform deformation and less amplitude decay of EM waves.

The 2-D synthetic examples confirmed the effectiveness and limitations of frequency-dependent GPR FWI. The spatially uncorrelated examples indicated that surface multioffset GPR data are weakly sensitive to permittivity attenuation compared to permittivity and conductivity. They also demonstrated that both frequency-dependent GPR FWI and frequency-independent GPR FWI suffer heavy crosstalk of different parameters. The spatially correlated examples showed that frequency-dependent GPR FWI works well in estimating the permittivity and conductivity models coupled with different permittivity attenuation levels, and frequency-independent GPR FWI produces comparable permittivity results as the dispersion effect is weak. Although combined with STF estimation, frequency-independent GPR FWI cannot handle the reconstruction of conductivity anomalies dominated by permittivity attenuation. In this case, frequency-dependent GPR FWI can provide more reliable conductivity results even if the permittivity attenuation models fail to be estimated and the observed data contain a degree of noise. These synthetic examples suggested the need for frequency-dependent GPR FWI in the presence of dispersion effects and variable attenuation.

When applying to field data acquired at the Rheinstetten test site, both frequency-dependent GPR FWI and frequency-independent GPR FWI successfully outlined the Ettlinger Line as a triangle permittivity anomaly. On the right of the trench, we found a slightly right-tilted layer which exhibits strong permittivity attenuation. Frequency-dependent GPR FWI showed better data fitting and more continuous conductivity structure in this region, and was likely to be more robust than frequency-independent GPR FWI. Previous GPR migration imaging and shallow-seismic FWIs have verified our GPR FWI results. Crosstalk mitigation and joint inversion with other geophysical approaches need to be studied in the future. The use of other models describing relaxation phenomena of permittivity and conductivity also deserves further investigation.

Acknowledgments

This work is financially supported by the China Scholarship Council (No. 201806260258). Tan Qin would like to thank Tilman Steinweg and Mark Wienöbst for their help in developing the WAVE-Toolbox. The authors sincerely thank Lars Houpt, Felix Bögelspacher, Leon Merkel, Michael Mayer, Hagen Steger, Roland Helfer, Philipp Koyan and the master students at Geophysical Institute, Karlsruhe Institute of Technology for their help in field data acquisition, and the editor Rene-Edouard Plessix, the reviewer Hansruedi Maurer and another anonymous reviewer for their constructive comments.

Data availablity

An open-source software (GPL) package containing the source code, models and data used in this paper is provided in the WAVE-Toolbox (https://github.com/WAVE-Toolbox).

A Convolution calculation

The convolution of the dielectric permittivity ε and the electric field $\mathbf{E}$ is

$$
\begin{aligned}
\varepsilon(t) * \partial_t \mathbf{E} =& \partial_t \Psi_\varepsilon(t) * \partial_t \mathbf{E} = \left\{ \partial_t \left[\varepsilon_s \left(1 - \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} e^{-t/\tau_{Dl}} \right) H(t) \right] \right\} * \partial_t \mathbf{E} \\
=& \left\{ \varepsilon_s \left[\left(1 - \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} e^{-t/\tau_{Dl}} \right) \delta(t) + \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}} e^{-t/\tau_{Dl}} H(t) \right] \right\} * \partial_t \mathbf{E} \\
=& \varepsilon_s (1 - \tau_\varepsilon) \partial_t \mathbf{E} + \left\{ \varepsilon_s \tau_\varepsilon \frac{1}{L} \left[\sum_{l=1}^{L} \frac{1}{\tau_{Dl}} e^{-t/\tau_{Dl}} H(t) \right] \right\} * \partial_t \mathbf{E} \\
=& \varepsilon_s (1 - \tau_\varepsilon) \partial_t \mathbf{E} + \left\{ \varepsilon_s \tau_\varepsilon \frac{1}{L} \partial_t \left[\sum_{l=1}^{L} \frac{1}{\tau_{Dl}} e^{-t/\tau_{Dl}} H(t) \right] \right\} * \mathbf{E} \\
=& \varepsilon_s (1 - \tau_\varepsilon) \partial_t \mathbf{E} + \left\{ \varepsilon_s \tau_\varepsilon \frac{1}{L} \left[\sum_{l=1}^{L} \frac{1}{\tau_{Dl}} e^{-t/\tau_{Dl}} \delta(t) - \sum_{l=1}^{L} \frac{1}{\tau_{Dl}^2} e^{-t/\tau_{Dl}} H(t) \right] \right\} * \mathbf{E} \\
=& \varepsilon_s (1 - \tau_\varepsilon) \partial_t \mathbf{E} + \varepsilon_s \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}} \mathbf{E} - \left[\varepsilon_s \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}^2} e^{-t/\tau_{Dl}} H(t) \right] * \mathbf{E} \\
=& \varepsilon_s (1 - \tau_\varepsilon) \partial_t \mathbf{E} + \varepsilon_s \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}} \mathbf{E} + \sum_{l=1}^{L} \mathbf{r}_l .
\end{aligned}
\tag{A.1}
$$

Therefore we define the memory variable $\mathbf{r}_l$ of the lth mechanism corresponding to $\mathbf{E}$ as

$$\begin{aligned}
\mathbf{r}_l =& -\left[\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^2}e^{-t/\tau_{Dl}}H(t)\right]*\mathbf{E},\\
\partial_t\mathbf{r}_l =& -\left[\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^2}e^{-t/\tau_{Dl}}\delta(t)-\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^3}e^{-t/\tau_{Dl}}H(t)\right]*\mathbf{E}\\
=& -\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^2}\mathbf{E}+\left[\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^3}e^{-t/\tau_{Dl}}H(t)\right]*\mathbf{E}\\
=& -\frac{\varepsilon_s\tau_\varepsilon}{L\tau_{Dl}^2}\mathbf{E}-\frac{1}{\tau_{Dl}}\mathbf{r}_l.
\end{aligned} \tag{A.2}$$

The convolution of the conductivity σ and $\mathbf{E}$ is

$$\begin{aligned}
\sigma(t)*\mathbf{E} =& \partial_t\Psi_\sigma(t)*\mathbf{E}\\
=& \partial_t\left\{\sigma_s\left[H(t)+\tau_\sigma\delta(t)\right]\right\}*\mathbf{E}\\
=& \sigma_s\left[\delta(t)*\mathbf{E}+\tau_\sigma\delta(t)*\partial_t\mathbf{E}\right]\\
=& \sigma_s\left(\mathbf{E}+\tau_\sigma\partial_t\mathbf{E}\right).
\end{aligned} \tag{A.3}$$

We sum the two convolutions and get

$$\begin{aligned}
\varepsilon(t)*\partial_t\mathbf{E}+\sigma(t)*\mathbf{E} =& \varepsilon_s\left(1-\tau_\varepsilon\right)\partial_t\mathbf{E}+\varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_{Dl}}\mathbf{E}+\sum_{l=1}^{L}\mathbf{r}_l\\
&+\sigma_s\left(\mathbf{E}+\tau_\sigma\partial_t\mathbf{E}\right)\\
=& \left[\varepsilon_s\left(1-\tau_\varepsilon\right)+\sigma_s\tau_\sigma\right]\partial_t\mathbf{E}\\
&+\left[\sigma_s+\varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_{Dl}}\right]\mathbf{E}+\sum_{l=1}^{L}\mathbf{r}_l\\
=& \varepsilon_\infty^e\partial_t\mathbf{E}+\sigma_\infty^e\mathbf{E}+\sum_{l=1}^{L}\mathbf{r}_l,
\end{aligned} \tag{A.4}$$

where

$$\varepsilon_\infty^e=\varepsilon_s\left(1-\tau_\varepsilon\right)+\sigma_s\tau_\sigma,\quad \sigma_\infty^e=\sigma_s+\varepsilon_s\tau_\varepsilon\frac{1}{L}\sum_{l=1}^{L}\frac{1}{\tau_{Dl}}. \tag{A.5}$$

B Gradient calculation

For consistency with Papadopoulos and Rekanos (2011), we present the gradients as the zero-lag cross-correlation of forward wavefield $\mathbf{u}$ and adjoint wavefield $\hat{\mathbf{u}}$ in the following. Note that we actually replace $\hat{\mathbf{u}}(t)$ in eq. (3.23) by $\hat{\mathbf{u}}'(T-t)$ which is given by the same forward solver. By doing so, the cross-correlation of $\hat{\mathbf{u}}(t)$ and $\mathbf{u}(t)$ becomes the convolution of $\hat{\mathbf{u}}'(t)$ and $\mathbf{u}(t)$.

Thus, we obtain the gradients of the electrical parameters as follows:

$$\begin{aligned}
\frac{\partial \Phi}{\partial \varepsilon_\infty^e} &= \int_0^T \left(\hat{E}_x \partial_t E_x + \hat{E}_y \partial_t E_y + \hat{E}_z \partial_t E_z\right) dt, \\
\frac{\partial \Phi}{\partial \sigma_\infty^e} &= \int_0^T \left(\hat{E}_x E_x + \hat{E}_y E_y + \hat{E}_z E_z\right) dt, \\
\frac{\partial \Phi}{\partial \varepsilon_s} &= -\frac{L r_{\varepsilon\sigma}}{{\varepsilon_s}^2 \tau_\varepsilon}, \\
\frac{\partial \Phi}{\partial \tau_\varepsilon} &= -\frac{L r_{\varepsilon\sigma}}{\varepsilon_s \tau_\varepsilon^2},
\end{aligned} \tag{B.1}$$

where

$$\begin{aligned}
r_{\varepsilon\sigma} &= \int_0^T \sum_{l=1}^{L} \tau_{Dl} \left(\tau_{Dl} r_{\varepsilon l} + r_{\sigma l}\right) dt, \\
r_{\varepsilon l} &= \left(\hat{r}_{xl} \partial_t r_{xl} + \hat{r}_{yl} \partial_t r_{yl} + \hat{r}_{zl} \partial_t r_{zl}\right), \\
r_{\sigma l} &= \left(\hat{r}_{xl} r_{xl} + \hat{r}_{yl} r_{yl} + \hat{r}_{zl} r_{zl}\right).
\end{aligned} \tag{B.2}$$

Eq. (B.2) can be simplified by the following equation, based on eq. (A.2) where $\tau_{Dl}(\tau_{Dl}\partial_t \mathbf{r}_l + \mathbf{r}_l) = -\frac{\varepsilon_s \tau_\varepsilon}{L}\mathbf{E}$:

$$r_{\varepsilon\sigma} = -\frac{\varepsilon_s \tau_\varepsilon}{L} \int_0^T \sum_{l=1}^{L} \left(\hat{r}_{xl} E_x + \hat{r}_{yl} E_y + \hat{r}_{zl} E_z\right) dt. \tag{B.3}$$

Then the gradient of ε_s and τ_ε are rewritten as:

$$\begin{aligned}
\frac{\partial \Phi}{\partial \varepsilon_s} &= \frac{1}{\varepsilon_s} \int_0^T \sum_{l=1}^{L} \left(\hat{r}_{xl} E_x + \hat{r}_{yl} E_y + \hat{r}_{zl} E_z\right) dt, \\
\frac{\partial \Phi}{\partial \tau_\varepsilon} &= \frac{1}{\tau_\varepsilon} \int_0^T \sum_{l=1}^{L} \left(\hat{r}_{xl} E_x + \hat{r}_{yl} E_y + \hat{r}_{zl} E_z\right) dt,
\end{aligned} \tag{B.4}$$

According to eq. (A.5) and the chain rule, we convert the gradients of the objective function from the parameter class $\mathbf{m} = (\varepsilon_\infty^e, \sigma_\infty^e, \varepsilon_s, \tau_\varepsilon)$ to $\mathbf{m}' = (\varepsilon_s', \sigma_s, \tau_\varepsilon', \tau_\sigma)$ where $\varepsilon_s' = \varepsilon_s$ and $\tau_\varepsilon' = \tau_\varepsilon$.

$$\begin{aligned}
\frac{\partial \Phi}{\partial \varepsilon_s'} &= (1 - \tau_\varepsilon) \frac{\partial \Phi}{\partial \varepsilon_\infty^e} + \tau_\varepsilon \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}} \frac{\partial \Phi}{\partial \sigma_\infty^e} + \frac{\partial \Phi}{\partial \varepsilon_s}, \\
\frac{\partial \Phi}{\partial \sigma_s} &= \tau_\sigma \frac{\partial \Phi}{\partial \varepsilon_\infty^e} + \frac{\partial \Phi}{\partial \sigma_\infty^e}, \\
\frac{\partial \Phi}{\partial \tau_\varepsilon'} &= -\varepsilon_s \frac{\partial \Phi}{\partial \varepsilon_\infty^e} + \varepsilon_s \frac{1}{L} \sum_{l=1}^{L} \frac{1}{\tau_{Dl}} \frac{\partial \Phi}{\partial \sigma_\infty^e} + \frac{\partial \Phi}{\partial \tau_\varepsilon}, \\
\frac{\partial \Phi}{\partial \tau_\sigma} &= \sigma_s \frac{\partial \Phi}{\partial \varepsilon_\infty^e}.
\end{aligned} \tag{B.5}$$

For the reference angular frequency ω_0, we have

$$\begin{cases} \varepsilon^e(\omega_0) = a\varepsilon_s + \sigma_s\tau_\sigma \\ \sigma^e(\omega_0) = \sigma_s + b\varepsilon_s \end{cases} \Leftrightarrow \begin{cases} \varepsilon_s = [\varepsilon^e(\omega_0) - \sigma^e(\omega_0)\tau_\sigma]/(a - b\tau_\sigma) \\ \sigma_s = \sigma^e(\omega_0) - b\varepsilon_s \end{cases}, \tag{B.6}$$

with

$$\begin{aligned} a &= 1 - \tau_\varepsilon \frac{1}{L}\sum_{l=1}^{L} \frac{\omega_0^2\tau_{Dl}^2}{1+\omega_0^2\tau_{Dl}^2} = 1 - \bar{a}\tau_\varepsilon, \\ b &= \tau_\varepsilon \frac{1}{L}\sum_{l=1}^{L} \frac{\omega_0^2\tau_{Dl}}{1+\omega_0^2\tau_{Dl}^2} = \bar{b}\tau_\varepsilon. \end{aligned} \tag{B.7}$$

Alternatively, we can convert the gradients of the objective function from the parameter class $\mathbf{m}' = (\varepsilon_s', \sigma_s, \tau_\varepsilon', \tau_\sigma)$ to $\mathbf{m}'' = (\varepsilon^e, \sigma^e, \tau_\varepsilon'', \tau_\sigma')$ where $\tau_\varepsilon'' = \tau_\varepsilon'$ and $\tau_\sigma' = \tau_\sigma$.

$$\begin{aligned} \frac{\partial\Phi}{\partial\varepsilon^e} &= \frac{\partial\varepsilon_s'}{\partial\varepsilon^e}\frac{\partial\Phi}{\partial\varepsilon_s'} + \frac{\partial\sigma_s}{\partial\varepsilon^e}\frac{\partial\Phi}{\partial\sigma_s}, \\ \frac{\partial\Phi}{\partial\sigma^e} &= \frac{\partial\varepsilon_s'}{\partial\sigma^e}\frac{\partial\Phi}{\partial\varepsilon_s'} + \frac{\partial\sigma_s}{\partial\sigma^e}\frac{\partial\Phi}{\partial\sigma_s}, \\ \frac{\partial\Phi}{\partial\tau_\varepsilon''} &= \frac{\partial\varepsilon_s'}{\partial\tau_\varepsilon''}\frac{\partial\Phi}{\partial\varepsilon_s'} + \frac{\partial\sigma_s}{\partial\tau_\varepsilon''}\frac{\partial\Phi}{\partial\sigma_s} + \frac{\partial\Phi}{\partial\tau_\varepsilon'}, \\ \frac{\partial\Phi}{\partial\tau_\sigma'} &= \frac{\partial\varepsilon_s'}{\partial\tau_\sigma'}\frac{\partial\Phi}{\partial\varepsilon_s'} + \frac{\partial\sigma_s}{\partial\tau_\sigma'}\frac{\partial\Phi}{\partial\sigma_s} + \frac{\partial\Phi}{\partial\tau_\sigma}, \end{aligned} \tag{B.8}$$

where

$$\begin{aligned} &\frac{\partial\varepsilon_s'}{\partial\varepsilon^e} = \frac{1}{a - b\tau_\sigma}, \quad \frac{\partial\sigma_s}{\partial\varepsilon^e} = -b\frac{\partial\varepsilon_s'}{\partial\varepsilon^e}, \\ &\frac{\partial\varepsilon_s'}{\partial\sigma^e} = -\frac{\tau_\sigma}{a - b\tau_\sigma}, \quad \frac{\partial\sigma_s}{\partial\sigma^e} = 1 - b\frac{\partial\varepsilon_s'}{\partial\sigma^e}, \\ &\frac{\partial\varepsilon_s'}{\partial\tau_\varepsilon''} = \frac{\varepsilon^e - \sigma^e\tau_\sigma}{(a - b\tau_\sigma)^2}(\bar{a} + \bar{b}\tau_\sigma), \quad \frac{\partial\sigma_s}{\partial\tau_\varepsilon''} = -b\frac{\partial\varepsilon_s'}{\partial\tau_\varepsilon''}, \\ &\frac{\partial\varepsilon_s'}{\partial\tau_\sigma'} = \frac{b\varepsilon^e - a\sigma^e}{(a - b\tau_\sigma)^2}, \quad \frac{\partial\sigma_s}{\partial\tau_\sigma'} = -b\frac{\partial\varepsilon_s'}{\partial\tau_\sigma'}. \end{aligned} \tag{B.9}$$

4 Indirect joint petrophysical inversion of synthetic shallow-seismic and multi-offset ground-penetrating radar data

Published as: Qin, T., Bohlen, T., & Pan, Y. (2022). Indirect joint petrophysical inversion of synthetic shallow-seismic and multi-offset ground-penetrating radar data. *Geophysical Journal International*, 229(3), 1770-1784.

Abstract

Both full-waveform inversion (FWI) of shallow-seismic and ground-penetrating radar (GPR) surface recordings have received particular attention in the past decade since they can reconstruct seismic and electromagnetic properties at high resolution. For consistent near-surface imaging, seismic and GPR wavefields can be combined by joint petrophysical inversion (JPI) using classical time-domain FWI. In conventional JPI of wavefields, both shallow-seismic and multi-offset GPR data contribute to reconstructing the same petrophysical parameters. In this paper, we show that seismic and GPR wavefields have different sensitivities to these parameters assuming the widely established petrophysical model combining the Gassmann fluid substitution model, the complex refractive index model and Archie's equation. Based on this observation, we propose a new coupling strategy of petrophysical parameters which we call indirect JPI. In indirect JPI, seismic data are primarily used for porosity reconstruction, while GPR data are used only for saturation reconstruction. Unlike conventional JPI, we first update the seismic and GPR parameters using non-petrophysical parametrizations and then transform the most reliable estimates to petrophysical parameters. 2-D synthetic tests show that indirect JPI can provide more accurate and consistent results than conventional JPI. In addition, indirect JPI is more robust when uncertainties exist in petrophysical *a priori* knowledge. More importantly, indirect JPI has the flexibility to integrate different types of seismic and electromagnetic waves and acquisition geometries depending on the target of interest, resulting in the best solution. Indirect JPI has been proven to be a promising approach for multiparameter reconstructions (seven parameters in this study).

4.1 Introduction

Shallow-seismic surface wave and ground-penetrating radar (GPR) methods play increasingly important roles in near-surface investigation. Full-waveform inversion (FWI) was first proposed by Tarantola (1984) for seismic reflection data, and then introduced to shallow-seismic data (Romdhane et al., 2011;

Tran and McVay, 2012) and crosshole GPR data (Ernst et al., 2007b; Kuroda et al., 2007). In the past decade, it has been shown that FWI, specifically shallow-seismic FWI and GPR FWI, has great potential to reconstruct high-resolution multiparameter subsurface models of seismic and electromagnetic (EM) material properties (Pan et al., 2019; Klotzsche et al., 2019), respectively.

In the shallow subsurface, surface waves (Love wave and Rayleigh wave) usually dominate the seismic wavefield, which requires elastic rather than acoustic FWI (Romdhane et al., 2011). Love-wave FWI (Pan et al., 2016; Dokter et al., 2017), Rayleigh-wave FWI (Groos et al., 2017) and the joint FWI of Rayleigh and Love waves (Wittkamp et al., 2019) were successfully applied to field data. On the other hand, GPR FWI has so far been applied primarily to crosshole data (Ernst et al., 2007b; Meles et al., 2010; Oberröhrmann et al., 2013; Gueting et al., 2015). GPR FWI using surface recordings has typically been used to determine the shallow guided layer (Busch et al., 2012) or isolated targets (El Bouajaji et al., 2011; Liu et al., 2018). Lavoué et al. (2014) proved that on-ground multi-offset GPR data possesses higher vertical resolution than crosshole GPR data and deserves more attention. Independent FWI of shallow-seismic and GPR data, however, may provide inconsistent information because of the non-uniqueness of the inverse problem and the inherent limitations of these geophysical methods. In general, the combination of several data may help to reduce uncertainties and avoid conflicting interpretations (Linde et al., 2008; Rumpf and Tronicke, 2014; Wagner et al., 2019; Domenzain et al., 2020). Hence, it is important to investigate the feasibility of integrating shallow-seismic and multi-offset GPR data using joint inversion.

There are two common joint inversion approaches: joint structural inversion (JSI) and joint petrophysical inversion (JPI). In JSI, it is assumed that the models share consistent interfaces, where the structural similarity is often achieved by applying a structural similarity measure using, for instance, a cross-gradient function as an additional constraint (Gallardo and Meju, 2003). JPI is based on empirical or experimental petrophysical relations that connect different physical parameters through the common petrophysical properties (porosity and saturation in this study). In JSI and JPI, researchers address the hypothesis that the joint inversion of seismic and EM wavefields leads to consistent multiparameter models of seismic and EM material properties, which have higher accuracy than the models obtained by individual FWI. Feng et al. (2017) investigated the JSI of shallow-seismic reflection and multi-offset GPR data. Gao et al. (2012) performed the JPI of controlled-source electromagnetic measurements (CSEM) and acoustic seismic measurements to monitor a reservoir and pointed out that petrophysical parameters cannot be inferred uniquely by seismic or EM measurements alone. Abubakar et al. (2012) performed a comparison between JSI and JPI using CSEM and elastic seismic data and demonstrated that JPI produces better results than JSI if the adopted petrophysical relations are reliable.

In order to implement JPI of shallow-seismic and multi-offset GPR data in a near-surface context (depth < 10 m), two prerequisites have to be satisfied. The first one is that the results obtained by the two inversions have the comparable spatial resolution, which can be achieved by selecting the central frequency of the seismic source and GPR transmitter so that both methods produce similar wavelength and depth penetration. The second one is that reliable petrophysical relations must be known. The relations between, on the one hand, petrophysical parameters and seismic parameters (Gassmann, 1951; Biot, 1955) and, on the other hand, petrophysical parameters and GPR parameters (Annan, 2005; Altdorff et al., 2019; Yu

et al., 2020) have been well investigated and applied in many cases. Thus, the existing JPI scheme, for example, developed by Gao et al. (2012) or Abubakar et al. (2012) can be adopted to combine shallow-seismic and GPR data. We call it conventional JPI. Nevertheless, the conventional JPI incorporates the seismic and GPR data to reconstruct the same petrophysical parameters without fully considering the sensitivity and reliability of the information provided by each method. That may lead to unsatisfactory results as a petrophysical parameter cannot benefit much from the data set if its sensitivity to the data set is weak. Besides, the poor reconstructed seismic or GPR parameters may interfere with the estimation of petrophysical parameters.

In this paper, we conduct a synthetic feasibility study of JPI of shallow-seismic and multi-offset GPR data by considering the sensitivity of different data to petrophysical parameters. We first introduce the petrophysical relations used in JPI and analyse the sensitivity of seismic and GPR parameters to petrophysical parameters. We then propose an indirect petrophysical parametrization to implement JPI, which is referred to as indirect JPI. A series of 2-D synthetic inversion examples are used to compare the performance of indirect JPI and conventional JPI. Finally, we investigate the robustness of indirect JPI in the presence of inaccurate petrophysical relations and rock matrix parameters and the applicability of indirect JPI to different wave types and data acquisition geometries.

4.2 Methodology

The goal of JPI is to estimate petrophysical parameters $\mathbf{m}_{\mathrm{P}}$, that is porosity ϕ and saturation S_{w} in our case, based on the relations between petrophysical and seismic parameters on the one hand, and petrophysical and GPR parameters on the other hand. Therefore, reliable petrophysical relations are assumed to be known beforehand. We use the conventional least-squares objective function to quantify the misfit in both data domains:

$$\begin{aligned}
&\Phi_{\mathrm{S}}(\mathbf{m}_{\mathrm{S}}) = \frac{1}{2}||\mathbf{d}_{\mathrm{S}}^{\mathrm{syn}}(\mathbf{m}_{\mathrm{S}}) - \mathbf{d}_{\mathrm{S}}^{\mathrm{obs}}||^2, \quad \mathbf{m}_{\mathrm{S}} = [V_P, V_S, \rho]^T \,, \\
&\Phi_{\mathrm{EM}}(\mathbf{m}_{\mathrm{EM}}) = \frac{1}{2}||\mathbf{d}_{\mathrm{EM}}^{\mathrm{syn}}(\mathbf{m}_{\mathrm{EM}}) - \mathbf{d}_{\mathrm{EM}}^{\mathrm{obs}}||^2, \quad \mathbf{m}_{\mathrm{EM}} = [\varepsilon, \sigma]^T \,, \\
&s.t. \quad f(\mathbf{m}_{\mathrm{S}}, \mathbf{m}_{\mathrm{EM}}, \mathbf{m}_{\mathrm{P}}) = 0, \quad \mathbf{m}_{\mathrm{P}} = [\phi, S_{\mathrm{w}}]^T \,,
\end{aligned} \tag{4.1}$$

where Φ_{S} and Φ_{EM} are the objective functions of seismic and GPR inversions, respectively; $\mathbf{d}_{\mathrm{S}}^{\mathrm{syn}}$ is the synthetic seismograms and $\mathbf{d}_{\mathrm{S}}^{\mathrm{obs}}$ the observed seismograms; $\mathbf{d}_{\mathrm{EM}}^{\mathrm{syn}}$ is the synthetic radargrams and $\mathbf{d}_{\mathrm{EM}}^{\mathrm{obs}}$ the observed radargrams; $\mathbf{m}_{\mathrm{S}}$ represents seismic parameters including the P-wave velocity V_P, S-wave velocity V_S, and density ρ; $\mathbf{m}_{\mathrm{EM}}$ represents EM parameters including electrical conductivity σ and dielectric permittivity ε. The function f describes the petrophysical relations. We use separate objective functions in this study, rather than a combined objective function such as the additive cost functions in Wagner et al. (2019) and the multiplicative cost functions in Abubakar et al. (2012). By doing so, we avoid calculating the data weighting matrix and the scaling factor for balancing the contribution of

seismic and GPR data in JPI (Heincke et al., 2017). JPI seeks to iteratively update the model parameters $\mathbf{m}$ ($\mathbf{m}_\mathrm{S}$, $\mathbf{m}_\mathrm{EM}$ and $\mathbf{m}_\mathrm{P}$) by minimizing Φ (Φ_S and Φ_EM) as follows:

$$\mathbf{m}^{k+1} = \mathbf{m}^k + \lambda \Delta \mathbf{m}^{k+1} \,, \tag{4.2}$$

$$\Delta \mathbf{m}^{k+1} = -\mathbf{P} \frac{\partial \Phi}{\partial \mathbf{m}} + \gamma \Delta \mathbf{m}^k \,, \tag{4.3}$$

where the step length λ is calculated by line search and parabolic fitting; The model update direction at the $(k+1)$th iteration $\Delta \mathbf{m}^{k+1}$ is computed by the preconditioned conjugate-gradient method (eq. 4.3), where the scale factor γ follows Polak and Ribiere (1969); The pre-conditioner $\mathbf{P}$ is given by eq. (23) in Plessix and Mulder (2004) to suppress source and receiver artefacts.

4.2.1 Petrophysical relations

In order to link the seismic parameters $\mathbf{m}_\mathrm{S}$ to petrophysical parameters $\mathbf{m}_\mathrm{P}$, we employ the Gassmann fluid substitution model (Gassmann, 1951). In fluid-saturated rocks, the *P*-wave velocity V_P, *S*-wave velocity V_S, and density ρ are expressed as follows:

$$\begin{aligned} &V_P = \sqrt{\frac{K + 4/3\,\mu}{\rho}}, \qquad V_S = \sqrt{\frac{\mu}{\rho}} \,, \\ &\rho = (1-\phi)\,\rho_\mathrm{ma} + \phi\,[S_\mathrm{w}\rho_\mathrm{w} + (1-S_\mathrm{w})\,\rho_\mathrm{a}] \,, \end{aligned} \tag{4.4}$$

with

$$\begin{aligned} &K = (1-\beta)\,K_\mathrm{ma} + \beta^2 M, \qquad \mu = (1-\beta)\,\mu_\mathrm{ma} \,, \\ &\frac{1}{M} = \frac{\beta-\phi}{K_\mathrm{ma}} + \frac{\phi}{K_\mathrm{f}}, \quad \frac{1}{K_\mathrm{f}} = \frac{S_\mathrm{w}}{K_\mathrm{w}} + \frac{1-S_\mathrm{w}}{K_\mathrm{a}} \,, \\ &\beta = \phi/\phi_\mathrm{c} \quad \text{with} \quad 0 \le \phi < \phi_\mathrm{c} \,, \end{aligned} \tag{4.5}$$

where K is the bulk modulus and μ the shear modulus of fluid-saturated rock; $K_\mathrm{ma} = \rho_\mathrm{ma} V_{P\mathrm{ma}}^2 - 4/3\,\mu_\mathrm{ma}$ is the bulk modulus, $\mu_\mathrm{ma} = \rho_\mathrm{ma} V_{S\mathrm{ma}}^2$ the shear modulus, and ρ_ma the density of the rock matrix where $V_{P\mathrm{ma}}$, $V_{S\mathrm{ma}}$, and ρ_ma represent the *P*-wave velocity, *S*-wave velocity, and density of the rock matrix, respectively; K_f is the bulk modulus of pore fluid and M represents the resulting average modulus; K_w and K_a are the bulk moduli of water and air, respectively; ρ_w is the density of water and ρ_a the density of air; β is the Biot's coefficient which is a function of porosity. In this study, we choose the critical porosity ϕ_c = 0.4 above which the solid becomes a suspension (Nur, 1992).

The dielectric permittivity ε is expressed as a function of ϕ and S_w by the complex refractive index model (CRIM) (Birchak et al., 1974; Garambois et al., 2002), that is

$$\varepsilon = \varepsilon_0 \left[(1-\phi)\sqrt{\varepsilon_\mathrm{rma}} + \phi\left(S_\mathrm{w}\sqrt{\varepsilon_\mathrm{rw}} + (1-S_\mathrm{w})\sqrt{\varepsilon_\mathrm{ra}}\right)\right]^2 \,, \tag{4.6}$$

where ε_rw, ε_ra and ε_rma are the relative permittivities of water, air and the rock matrix, respectively. For convenience, we use the relative permittivity $\varepsilon_\mathrm{r} = \varepsilon/\varepsilon_0$ where ε_0 is the free space permittivity.

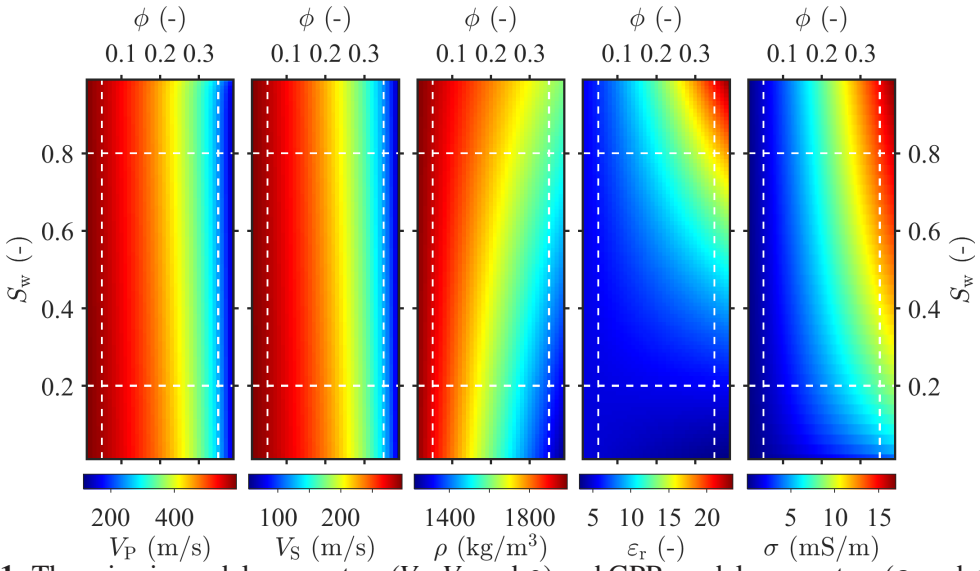

Figure 4.1: The seismic model parameters (V_P, V_S and ρ) and GPR model parameters (ε_r and σ) as functions of porosity $\phi \subset [0.01, 0.39]$ and saturation $S_w \subset [0.01, 0.99]$ using $V_{Pma} = 600$ m s^{-1}, $V_{Sma} = 300$ m s^{-1} and $\rho_{ma} = 2000$ kg m^{-3} in Gassmann's equations, $\sigma_w = 50$ mS m^{-1} in Archie's equation and $\varepsilon_{rma} = 5$ in CRIM. The white dashed lines indicate the slices shown in Fig. 4.2.

The electrical conductivity σ, ϕ and S_w follow Archie's equation (Archie et al., 1942):

$$\sigma = \frac{1}{a}\sigma_w \phi^m S_w^n \,, \tag{4.7}$$

where σ_w is the conductivity of groundwater, a is the tortuosity factor, m is the cementation exponent, and n is the saturation exponent.

In the above relations, the seismic petrophysical parameters of the rock matrix (V_{Pma}, V_{Sma} and ρ_{ma}), the EM petrophysical parameters of the rock matrix (ε_{rma} and σ_w) and Archie's coefficients (a, m and n) are obtained from well logging, core drilling or laboratory measurements. We use $a = 2$, $m = 0.4$ and $n = 1.13$ in this study, where the low values of m and n are chosen for loose clay, a typical near-surface sediment. These rock matrix parameters and Archie's coefficients are assumed to be known in the synthetic tests.

4.2.2 Sensitivity analysis

To analyse the sensitivity of the petrophysical parameters to the seismic model parameters (V_P, V_S and ρ) and the GPR model parameters (ε and σ), we plot their variation as a function of ϕ and S_w within the range $0 < \phi < \phi_c$ and $0 < S_w < 1$ in Fig. 4.1. It can be seen that the seismic parameters are mainly affected by porosity ϕ. Saturation S_w has weak influence on seismic velocities and a minor effect on density. The GPR parameters have a relatively strong sensitivity to both porosity and saturation. Consequently, for the empirical relations considered, seismic FWI can be used primarily to constrain porosity, whereas GPR FWI can constrain two petrophysical parameters simultaneously.

In order to emphasize the sensitivity of the petrophysical parameters, we plot the relative changes of seismic and GPR model parameters in Fig. 4.2. We choose $S_w = 0.2$ and $S_w = 0.8$ to represent low and

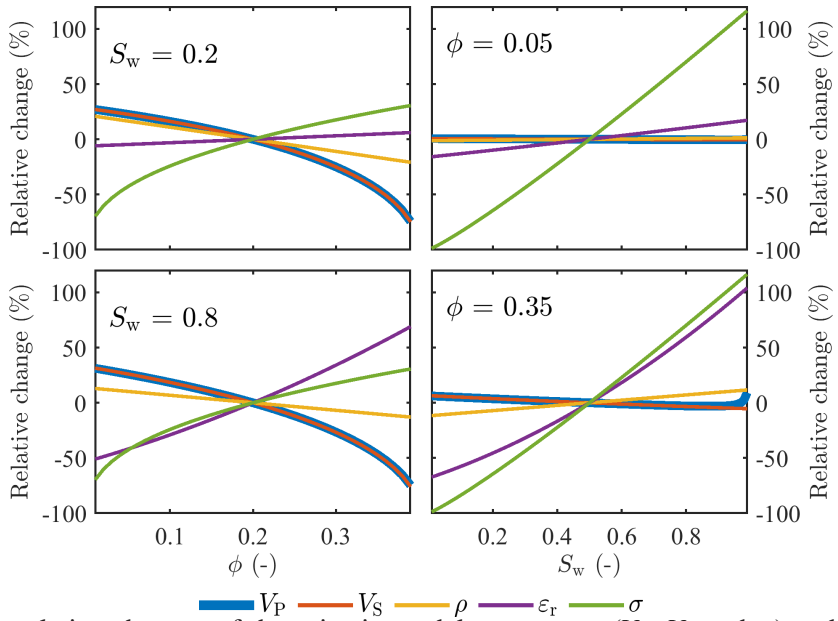

Figure 4.2: The relative changes of the seismic model parameters (V_P, V_S and ρ) and the GPR model parameters (ε_r and σ) as functions of porosity ϕ and saturation S_w around the reference porosity $\phi = 0.2$ (the left-hand column) and the reference saturation $S_w = 0.5$ (the right-hand column). The rock matrix parameters and Archie's coefficients are the same as in Fig. 4.1. The relative changes of V_P and V_S are very similar.

high saturation, and $\phi = 0.05$ and $\phi = 0.35$ to represent low and high porosity, respectively. On the left-hand column of Fig. 4.2, we observe that seismic velocities and conductivity have opposite trends with increasing porosity for both saturation cases. The dielectric permittivity ε is only sensitive to porosity variation in the high saturation case. As can be seen in the right-hand column of Fig. 4.2, the influence of saturation on the seismic model parameters is relatively weak. Permittivity is more sensitive to saturation in the high porosity environment. Although conductivity is greatly affected by porosity and saturation, this might not be a favourable factor for JPI because it is relatively difficult to estimate conductivity in GPR FWI.

4.2.3 Joint petrophysical inversion

Based on the sensitivity analysis in Section 4.2.2, we observe that, for the assumed petrophysical relations, the seismic parameters are mainly sensitive to porosity, and the EM parameters are affected by porosity and saturation. In conventional JPI shown in Fig. 4.3(a), the gradients of the seismic or GPR objective function with respect to porosity and saturation are given by the chain rule, and then directly used to update porosity and saturation through eqs (4.2) and (4.3). However, one may not expect shallow-seismic data to help the reconstruction of saturation because of the low sensitivity of seismic parameters to saturation (Figs 4.1 and 4.2). Besides, the conductivity gradient with low reliability can be projected onto the petrophysical gradients, resulting in a poorer reconstruction of petrophysical parameters than that produced by the individual petrophysical inversion. We have observed these phenomena in many numerical examples, which are not shown in this paper. To avoid these drawbacks, we can simplify conventional JPI in each iteration in four steps:

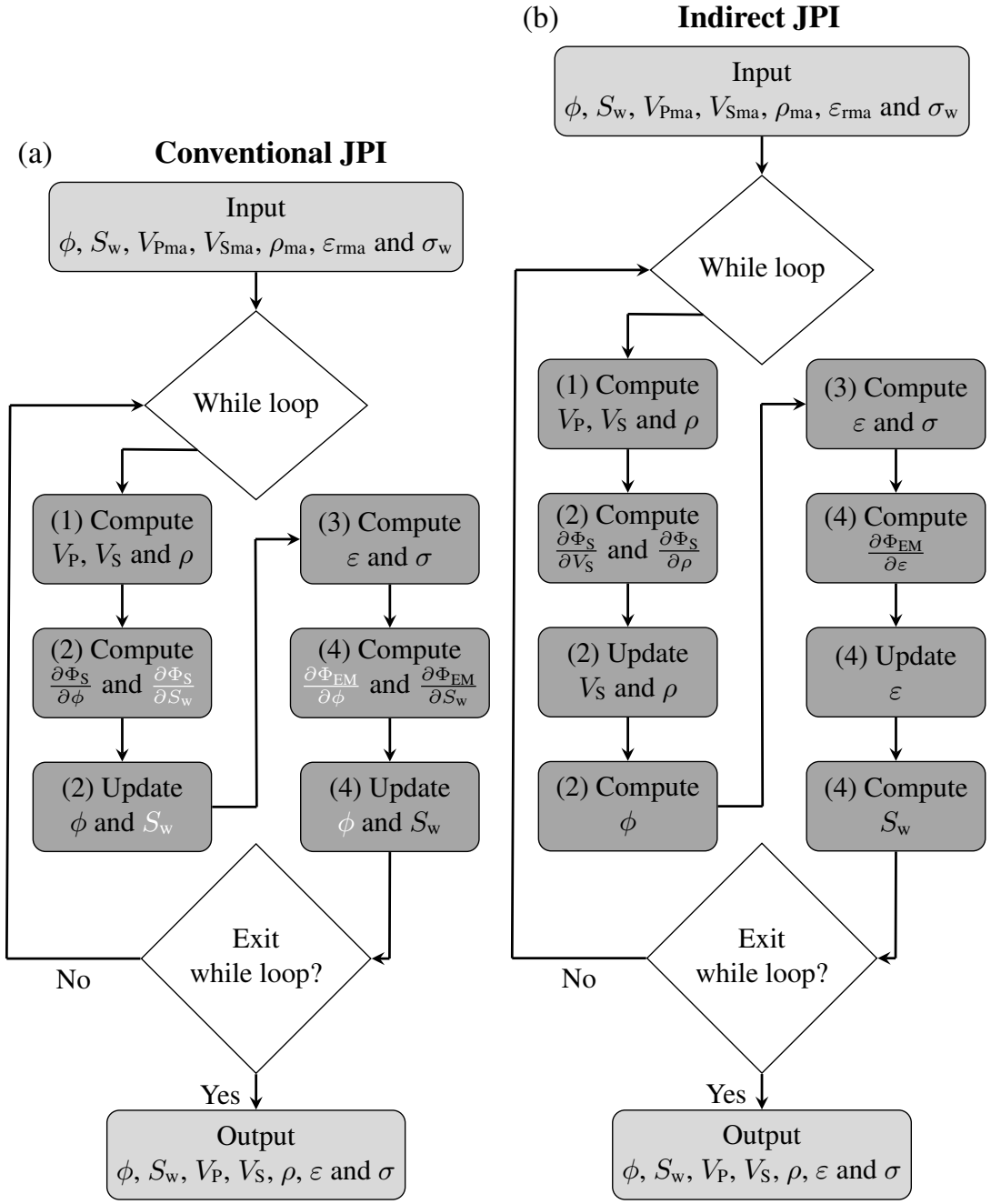

Figure 4.3: The workflows of joint petrophysical inversion (JPI). (a) Conventional JPI. (b) Indirect JPI. The gradients $\partial\Phi_S/\partial\phi$, $\partial\Phi_S/\partial S_w$, $\partial\Phi_{EM}/\partial\phi$ and $\partial\Phi_{EM}/\partial S_w$ in (a) are given by the petrophysical parametrization, similar to eqs (29) and (30) in Gao et al. (2012). The models and gradients in white will not be used in the simplified conventional JPI. The gradients $\partial\Phi_S/\partial V_S$ and $\partial\Phi_S/\partial\rho$ in (b) are yielded by the seismic velocity parametrization, similar to eq. (12) in Köhn et al. (2012). The gradient $\partial\Phi_{EM}/\partial\varepsilon$ in (b) is provided by the logarithmic parametrization, eq. (38) in Meles et al. (2010). The implementation steps of JPI are indicated by numbers one to four.

(1). We calculate the seismic parameters (V_P, V_S, and ρ) from the ϕ and S_w models through Gassmann's equations (eqs 4.4 and 4.5).

(2). In shallow-seismic FWI, we use the petrophysical parametrization to calculate the ϕ gradient and update the ϕ model only (eqs 4.2 and 4.3).

(3). The ϕ and S_w models are used to calculate the EM parameters (ε and σ) via CRIM (eq. 4.6) and Archie's equation (eq. 4.7).

(4). In GPR FWI, we employ the petrophysical parametrization to compute the S_w gradient and update the S_w model only (eqs 4.2 and 4.3).

The simplified conventional JPI is also shown in Fig. 4.3(a) (note the difference between the white and black variables). Unlike conventional JPI, the simplified conventional JPI does not use seismic data to calculate the gradient and update the model of saturation in step (2), nor does it use GPR data to calculate the gradient and update the model of porosity in step (4). This simplification allows one data set to contribute to only one of the two petrophysical parameters. The simplified conventional JPI ensures that the results are at least not worse than the individual petrophysical inversion results, which is a fundamental requirement that conventional JPI must fulfil. However, such a simplification may not be sufficient because the EM Jacobian matrix of saturation still includes the conductivity and the seismic Jacobian matrix of porosity consists of the bulk modulus (the *P*-wave velocity related), similar to eqs (29) and (30) in Gao et al. (2012). This is a bottleneck of conventional JPI when we estimate multiple parameters simultaneously.

In shallow-seismic FWI, where the wavefield is dominated by surface waves, the *S*-wave velocity is the parameter that can be reconstructed most reliably (Pan et al., 2019). Similarly, the permittivity, which quantifies the phase velocity of EM waves, can be inverted by GPR FWI with good accuracy (Klotzsche et al., 2019). We, therefore, propose an indirect petrophysical parametrization that establishes relations between porosity and *S*-wave velocity, as well as saturation and permittivity. Indirect petrophysical parametrization consists of seismic velocity parametrization and logarithmic parametrization which have proven to be efficient for shallow-seismic and GPR FWIs, respectively. For details of those two parametrizations, the reader is referred to, for instance, Köhn et al. (2012) and Meles et al. (2010). In the following, we first introduce two equations necessary for indirect petrophysical parametrization and then describe the steps to implement this parametrization in JPI.

In indirect petrophysical parametrization, the well recovered seismic model parameters (V_S and ρ) are transformed to ϕ by using Gassmann's equations as follows:

$$\phi = \phi_c \left(1 - \frac{\mu}{\mu_{ma}}\right), \quad \mu = \rho V_S^2, \quad \text{with} \quad 0 \leqslant \phi < \phi_c \,. \tag{4.8}$$

On the other hand, we estimate S_{w} by the well reconstructed GPR model parameter (ε) through eq. (4.9), which is another form of CRIM:

$$S_{\mathrm{w}} = \frac{\left[\sqrt{\varepsilon/\varepsilon_0} - (1-\phi)\sqrt{\varepsilon_{\mathrm{rma}}}\right]\phi^{-1} - \sqrt{\varepsilon_{\mathrm{ra}}}}{\sqrt{\varepsilon_{\mathrm{rw}}} - \sqrt{\varepsilon_{\mathrm{ra}}}}, \quad \text{with} \quad 0 \leqslant S_{\mathrm{w}} < 1\,. \tag{4.9}$$

We implement indirect JPI in each iteration through the following four steps:

(1). We calculate the seismic parameters (V_P, V_S, and ρ) from the ϕ and S_{w} models through Gassmann's equations (eqs 4.4 and 4.5).

(2). In shallow-seismic FWI, we use the velocity parametrization to compute the gradients and update the models of V_S and ρ (eqs 4.2 and 4.3). The ϕ model is then calculated from the recovered V_S and ρ by eq. (4.8).

(3). The ϕ and S_{w} models are used to calculate the EM parameters (ε and σ) via CRIM (eq. 4.6) and Archie's equation (eq. 4.7).

(4). We employ the logarithmic parametrization to compute the ε gradient and update the ε model in GPR FWI (eqs 4.2 and 4.3). Eq. (4.9) is then used to estimate S_{w} model from the reconstructed ε model.

As shown in Fig. 4.3, the main differences between conventional JPI and indirect JPI are steps (2) and (4). Unlike conventional JPI, which uses the petrophysical parametrization directly, indirect JPI applies non-petrophysical parametrizations (i.e. the seismic velocity parametrization and logarithmic parametrization) to compute the gradients and update the models of seismic and GPR parameters (Fig. 4.3b). It then transforms the best recovered parameters (the *S*-wave velocity and density or permittivity) to porosity or saturation. Thus, we improve the reliability of the petrophysical inversion results by indirectly calculating these less reliable parameters (the *P*-wave velocity and conductivity) with reliable information. Similar to the simplified conventional JPI, indirect JPI also takes into account the sensitivity of petrophysical parameters to seismic and EM parameters. Furthermore, indirect JPI uses the sensitivity of shallow-seismic and GPR data to seismic and EM parameters, which stabilizes the multiparameter reconstruction. Note that the simplified conventional JPI belongs to the conventional JPI family as it applies the petrophysical parametrization in the same way. Therefore, we still refer to this approach as conventional JPI throughout this paper.

4.3 Synthetic examples

In this section, we explore the characteristics of indirect JPI of shallow-seismic and multi-offset GPR data using a controlled environment with synthetic data. We first introduce a 2-D reference model and the surface data acquisition setup of Rayleigh waves and TM mode waves (i.e. the mode whose magnetic field components are always in the measurement plane). Then we apply indirect JPI on the forward-modelled synthetic data to simultaneously reconstruct all seven parameters (ϕ, S_{w}, V_P, V_S, ρ, ε and σ). Furthermore, we analyse the robustness of indirect JPI to inaccurate petrophysical relations and rock

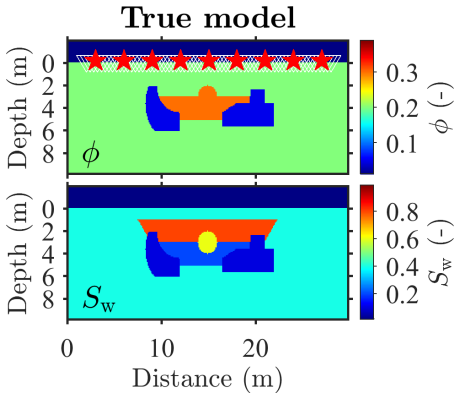

Figure 4.4: The reference petrophysical models used in the synthetic inversion tests. The red stars and white triangles represent the transmitters and receivers of GPR (0.1 m above the surface). The seismic data acquisition uses the same number and interval of sources and geophones as transmitters and receivers used in GPR data acquisition, but the sources and geophones are placed along the surface. The air layer (depth < 0 m) is known and will not be displayed in the inversion results.

matrix parameters. The results of (simplified) conventional JPI are compared. In the end of this section, we investigate the possibility of applying indirect JPI that combines other wave types and acquisition geometries.

Seismic waves and EM waves are modelled by the 2-D finite-difference time-domain (FDTD) method (Virieux, 1984, 1986; Yee, 1966). The convolutional perfectly matched layer (CPML) is included at the model boundaries, except for the free surface of seismic models where an imaging method is applied (Levander, 1988). We use a multi-scale strategy to avoid cycle skipping in FWI (Bunks et al., 1995). In the joint inversion approach, inversions can switch to the next stage if the abort criteria of both individual inversions are satisfied or if the maximum iteration number is reached. The seismic and electric source time functions are known in the inversion. A depth-dependent weighting scheme is applied to seismic inversion to enlarge the contribution of the deeper region (Plessix and Mulder, 2008).

4.3.1 Synthetic model and data acquisition

The true petrophysical models, modified from the synthetic study of Pan et al. (2020), exhibit a high structural similarity between porosity ϕ and saturation S_w (see Fig. 4.4). We use the same Archie's coefficients as we presented in Section 4.2.1. The rock matrix parameters are 1-D so that the values of background seismic parameters (V_P, V_S and ρ) and GPR parameters (ε and σ) increase with depth from 0 to 8 m and they are kept constant in the part deeper than 8 m. The values of the seismic rock matrix parameters ($V_{P\mathrm{ma}}$, $V_{S\mathrm{ma}}$ and ρ_{ma}) linearly increase with depth, while the values of the EM rock matrix parameters ($\varepsilon_{\mathrm{rma}}$ and σ_w) quadratically increase with depth to approximate a linearly decreasing EM wave velocity ($v \approx 1/\sqrt{\mu_m \varepsilon}$ where μ_m is magnetic permeability). We set $V_{P\mathrm{ma}} = 342.1531\ \mathrm{m\,s^{-1}}$, $V_{S\mathrm{ma}} = 207.6540\ \mathrm{m\,s^{-1}}$, $\rho_{\mathrm{ma}} = 1899.8\ \mathrm{kg\,m^{-3}}$, $\varepsilon_{\mathrm{rma}} = 2.4025$ and $\sigma_w = 11.9\ \mathrm{mS\,m^{-1}}$ at the surface, and $V_{P\mathrm{ma}} = 721.1296\ \mathrm{m\,s^{-1}}$, $V_{S\mathrm{ma}} = 437.1345\ \mathrm{m\,s^{-1}}$, $\rho_{\mathrm{ma}} = 3024.8\ \mathrm{kg\,m^{-3}}$, $\varepsilon_{\mathrm{rma}} = 20.0256$ and $\sigma_w = 53.5\ \mathrm{mS\,m^{-1}}$ at 8 m depth. Correspondingly, the true seismic and GPR models shown in Fig. 4.5 are built by those 1-D rock matrix models combined with the true petrophysical models through the petrophysical relations

mentioned in Section 4.2.1. Please note that although the air layer ($\varepsilon_r = 1$ and $\sigma = 0$ mS m^{-1}) is not displayed in the GPR models, it has to be included in the forward modelling.

Nine vertical particle velocity sources (40 Hz shifted Ricker wavelet) are generated with a source spacing of 3 m along the surface (almost coincident with the red stars in Fig. 4.4). Sixty geophones are also placed along the surface from 2 m to 28 m with an interval of 0.4 m to acquire vertical-component seismograms (almost coincident with the white triangles in Fig. 4.4). The GPR data acquisition uses the same number and interval of transmitters and receivers as sources and geophones used in the shallow-seismic data acquisition. The position of GPR antennas is 0.1 m above the surface (the red stars and white triangles in Fig. 4.4). The transmitters excite a 40 MHz shifted Ricker wavelet in the electric field perpendicular to the observation plane, and the receivers acquire the same component radargrams. To simulate a more realistic example, we add Gaussian noise with a signal-to-noise ratio (SNR) of 10 dB, with respect to the strongest amplitude, to the seismic data. In the radargrams, Gaussian noise of SNR = 25 dB is added. The different noise levels applied to seismograms and radargrams are determined by the features of shallow-seismic data and multi-offset GPR data (see Appendix C for details). The dominant wavelength of the observed surface waves is around 4 m, which is similar to the dominant wavelength of the EM waves that eliminate the direct waves.

4.3.2 Joint petrophysical inversion of noisy data

The homogeneous background parameters with $\phi = 0.2$ and $S_w = 0.4$, except for the air layer where $\phi = S_w = 0$, are used as initial petrophysical models. The initial seismic and GPR models are calculated from those initial petrophysical models and the 1-D rock matrix parameters mentioned earlier. In the multiscale strategy, the maximum number of iterations per stage is 15, and the abort criterion is that the relative data misfit change is less than 1 per cent. We select five inversion stages to sequentially use data with ever decreasing wavelength. From the first stage to the fifth stage, frequency-band variations from 5 to 20, 35, 45, 60, and 80 Hz are implemented for seismic inversion. Due to the high-frequency attenuation of EM waves, we apply frequency-band variations from 5 to 15, 25, 35, 50, and 80 MHz for GPR inversion. At each stage, the dominant wavelengths of the observed seismic waves and EM waves (eliminating the direct waves) are comparable.

We show the reconstructed seismic and GPR models in Fig. 4.5. Both conventional JPI and indirect JPI generate consistent inversion results of seismic velocities and density. This can be attributed to the internal constraint provided by petrophysical relations on seismic parameters. However, indirect JPI performs better than conventional JPI since the *P*-wave velocity model is indirectly calculated from the *S*-wave velocity and density models which are reconstructed more reliably in seismic FWI. On the other hand, conventional JPI provides disappointing results of the permittivity and conductivity models. The main reason is that conductivity, which is related to amplitude attenuation of radargrams, is relatively hard to estimate in GPR FWI, while conventional JPI uses this untrustworthy conductivity information in the gradient calculation of petrophysical parameters. Consequently, even the permittivity model, which should be well recovered in GPR FWI, is compromised. In this case, the petrophysical link projects the errors in conductivity onto permittivity. In indirect JPI, this negative effect is eliminated by transforming

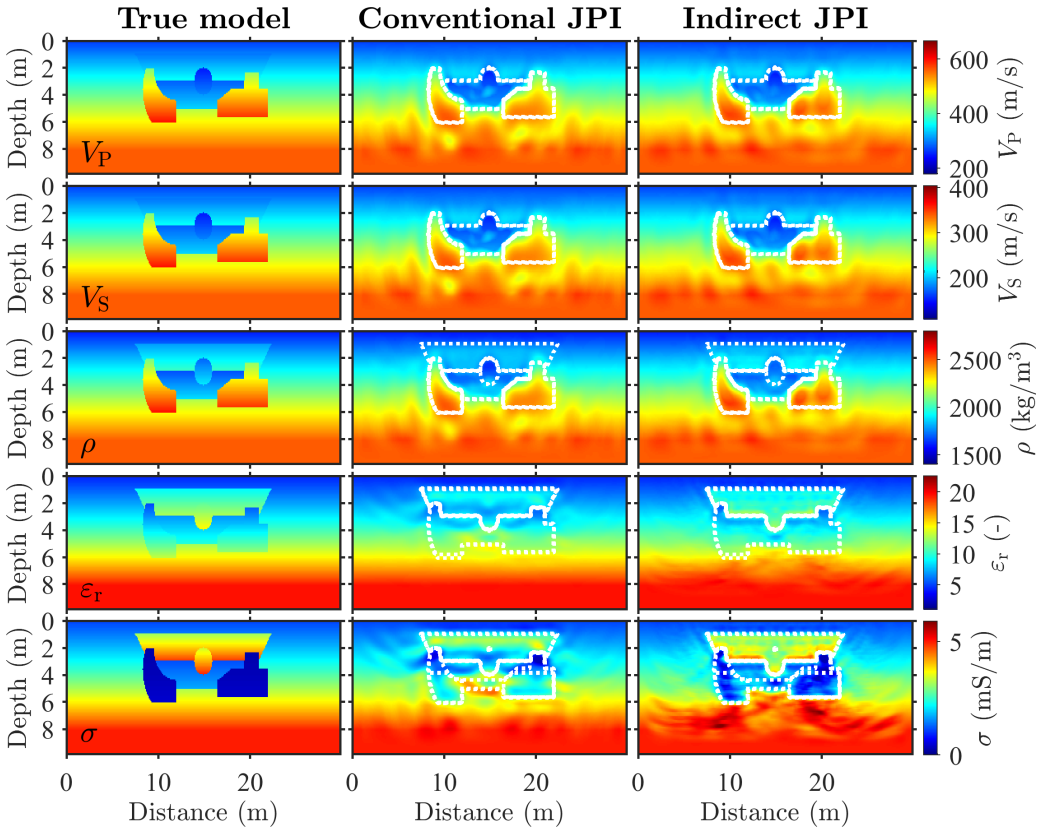

Figure 4.5: The seismic models (the upper three rows) and GPR models (the lower two rows) reconstructed by JPI using noisy data simulated in the reference model. The white dashed lines in JPI results indicate the outline of the true models.

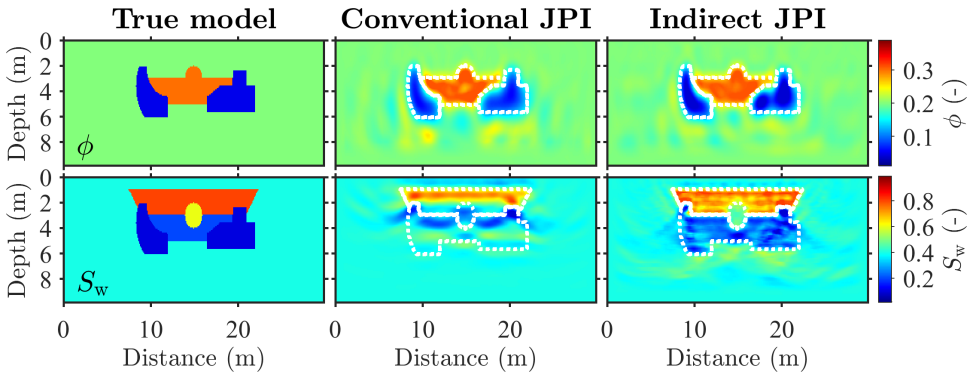

Figure 4.6: The petrophysical models reconstructed by JPI using noisy data simulated in the reference model.

a well-recovered permittivity model to a saturation model which is then fed back into the conductivity model. The conductivity model is therefore indirectly computed from the permittivity model which carries more accurate information in GPR FWI.

The petrophysical models obtained by JPI are shown in Fig. 4.6. One can find that conventional JPI is able to recover porosity, but fails to recover saturation. As we stated in Section 4.2.3, the less reliable conductivity gradient in GPR FWI misleads the saturation gradient to a wrong direction through the chain rule. To avoid such interference, indirect JPI mutes the contribution of conductivity in reconstructing

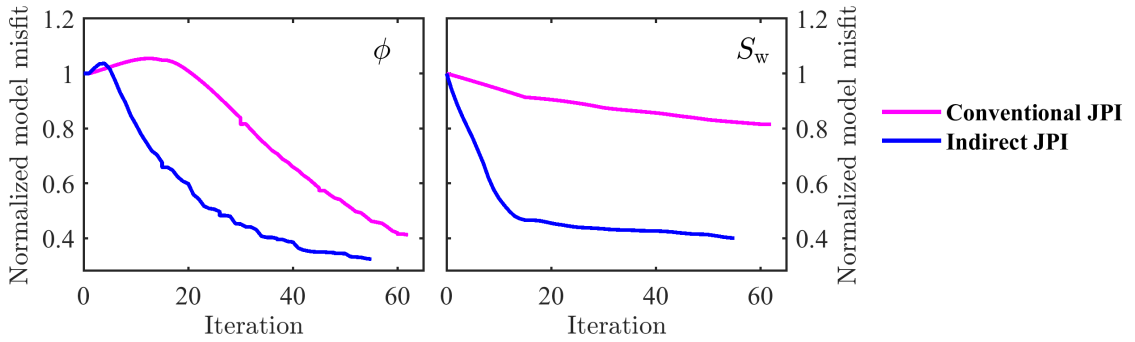

Figure 4.7: The petrophysical model misfits of JPI using noisy data simulated in the reference model.

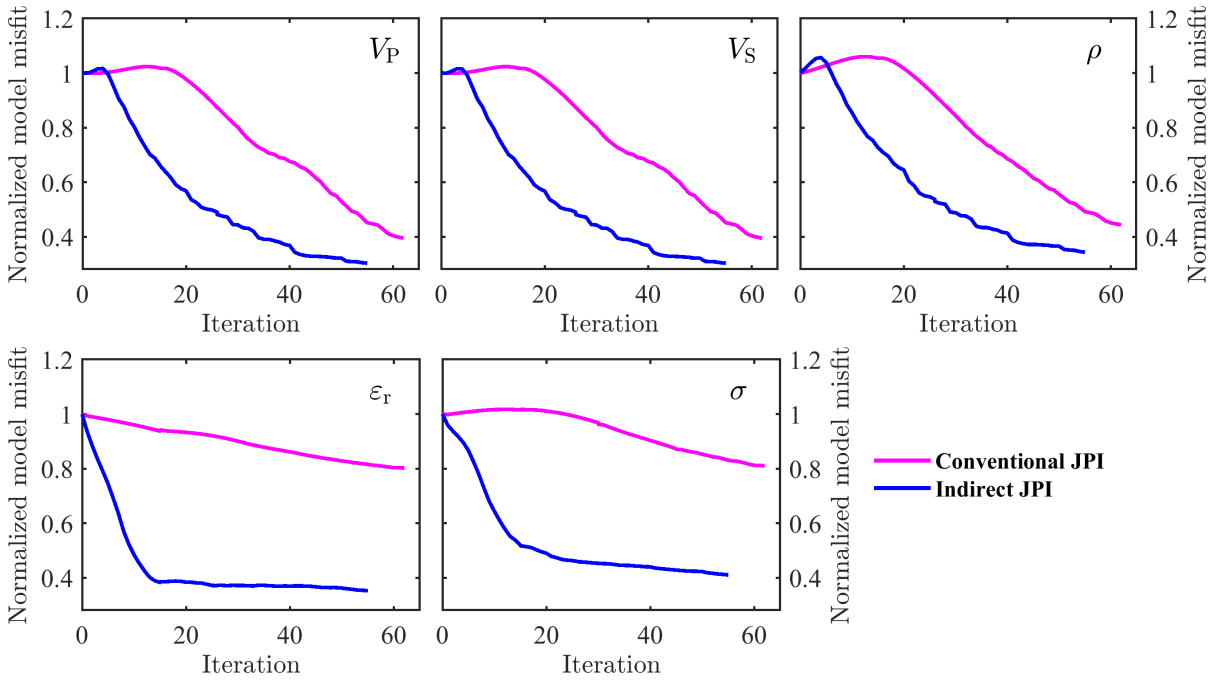

Figure 4.8: The seismic and GPR model misfits of JPI using noisy data simulated in the reference model.

saturation. Fig. 4.6 demonstrates that indirect JPI can recover porosity and saturation with high accuracy. In particular, a high-saturation layer and low-saturation anomalies are outlined clearly. Furthermore, the values in the low-porosity regions are much closer to the true values as the saturation model has been correctly reconstructed. That explains why we can identify the top layer with higher density and the deep low-conductive anomalies from the indirect JPI results shown in Fig. 4.5.

In order to evaluate the performance of the two JPIs, we calculate the L_2 norm model misfit, a deviation of the reconstructed model from the true synthetic model, as a function of iteration number. As shown in Fig. 4.7, conventional JPI reduces the final misfit of porosity and saturation to about 40 and 80 per cent of the initial model misfits, while indirect JPI can reach lower values, especially for saturation (40 per cent). In the first few iterations, we observe an increase in porosity misfit which corresponds to the reconstruction of the high-porosity anomaly in the centre. The subsequent decrease in porosity misfit implies that the low-porosity anomalies on two sides begin to reconstruct correctly. Compared to conventional JPI, indirect JPI starts to update the low-porosity region earlier, ultimately yielding a better porosity estimate. More importantly, indirect JPI performs much better in converging to lower

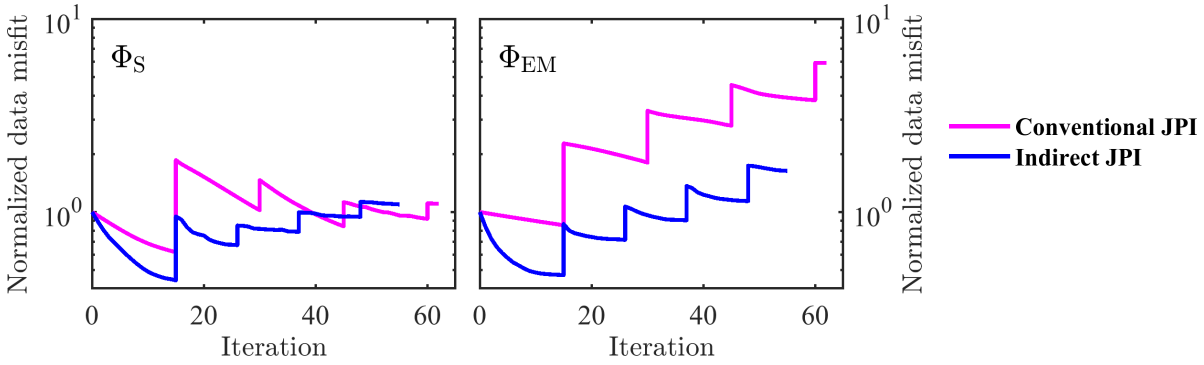

Figure 4.9: The seismic and GPR data misfits of JPI using noisy data simulated in the reference model.

saturation misfit. Comparing Fig. 4.7 and 4.8, we find similar variations of the model misfit in the seismic parameters (V_P, V_S and ρ) and porosity, in both conventional JPI and indirect JPI. It illustrates that the seismic parameters are strongly sensitive to porosity. On the other hand, we observe that the conductivity misfit in conventional JPI increases first and then decreases, resulting from the superposition of porosity and saturation. Permittivity misfit in conventional JPI seems to be more impacted by saturation than porosity. In indirect JPI, since the misfits of the GPR parameters and saturation converge much better, the misfits of the seismic parameters and porosity can further decrease.

Finally, we show the convergence curves of the data misfit in Fig. 4.9. The 'jumps' of data misfit happen, when the inversions change to a broader frequency band in the next stage. In the seismic data misfits, the two convergence curves end up with similar values because the *S*-wave velocity, which mainly affects the objective function of shallow-seismic FWI, is recovered to similar levels (30–40 per cent of the initial model misfit, see Fig. 4.8). The GPR data matching is improved significantly in indirect JPI compared to conventional JPI. In this synthetic test, indirect JPI is capable of reconstructing more accurate models and reaching lower levels of model misfit and data misfit at the same time.

4.3.3 Robustness tests

In the above research, the accurate Archie's coefficients and rock matrix models are exactly known. However, this is not realistic and might lead us to overestimate the performance of the JPI approach. Therefore, in this section, we test the robustness of JPI in the presence of errors in *a priori* information. The observed data used for JPI are the same noisy data simulated in the reference model. All the final model misfits are normalized by those obtained by the initial model with accurate Archie's coefficients and rock matrix parameters.

We change the value of Archie's coefficients m and n used in the inversion and compare the corresponding results of JPI with those using correct Archie's coefficients (m = 0.4 and n = 1.13). Table 4.1 reveals that the range of the final porosity misfit is small (41.22–42.85 per cent in conventional JPI and 31.69–33.81 per cent in indirect JPI) and slightly affected by the variation of m and n. This observation is reasonable because porosity is reconstructed by seismic data alone in conventional JPI and indirect JPI (note the simplification of conventional JPI in Section 4.2.3). The changes of m and n in conventional JPI cannot cause a worse impact on the final saturation misfit that is already quite large (around 80 per

Table 4.1: The final petrophysical model misfits with respect to different error levels of Archie's coefficients (m and n). The accurate Archie's coefficients and the corresponding misfits are shown in bold.

		Conventional JPI		Indirect JPI	
m	n	ϕ (per cent)	S_w (per cent)	ϕ (per cent)	S_w (per cent)
0.4	**1.13**	**41.22**	**81.52**	**32.38**	**40.01**
1.0	1.13	42.65	78.97	33.81	37.92
2.0	1.13	42.85	77.52	32.59	36.97
0.4	1.50	42.59	80.38	33.59	49.67
0.4	2.00	41.48	78.97	32.41	62.57
0.4	3.00	42.46	81.83	31.69	78.12

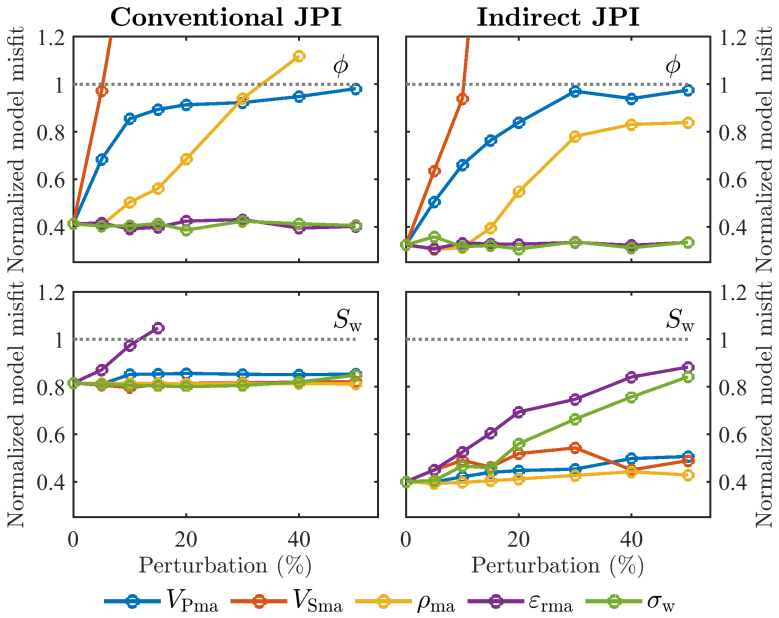

Figure 4.10: The final petrophysical model misfits as functions of the perturbation of rock matrix parameters. The grey dash line marks the initial model misfit. The colourful dots indicate the actual values of the perturbation and model misfit, and the colourful lines across those dots represent the trend of model misfit along with the perturbation range (≤ 50 per cent). Once the final model misfit of one parameter is higher than the initial model misfit in certain perturbations, we no longer display the results of higher perturbation.

cent). In indirect JPI, the saturation misfit does not suffer from an inexact cementation exponent m if the saturation exponent n is fixed. However, with n increasing, the saturation misfit increases because the reconstruction of permittivity is influenced by the poorly recovered conductivity and can damage the saturation. When n increases to 3.00, the recovered saturation deteriorates but remains similar to that yielded by conventional JPI using accurate Archie's coefficients.

We apply a constant homogeneous perturbation to the seismic and EM rock matrix parameters. When one rock matrix parameter is perturbed, the other rock matrix parameters are unaffected. The perturbation is quantified by the positive relative percentage of the average value of each rock matrix parameter. We use the perturbed model as an initial model of JPI. As shown in Fig. 4.10, V_{Pma}, V_{Sma} and ρ_{ma} have a great influence on the reconstruction of porosity in both conventional JPI and indirect JPI. In

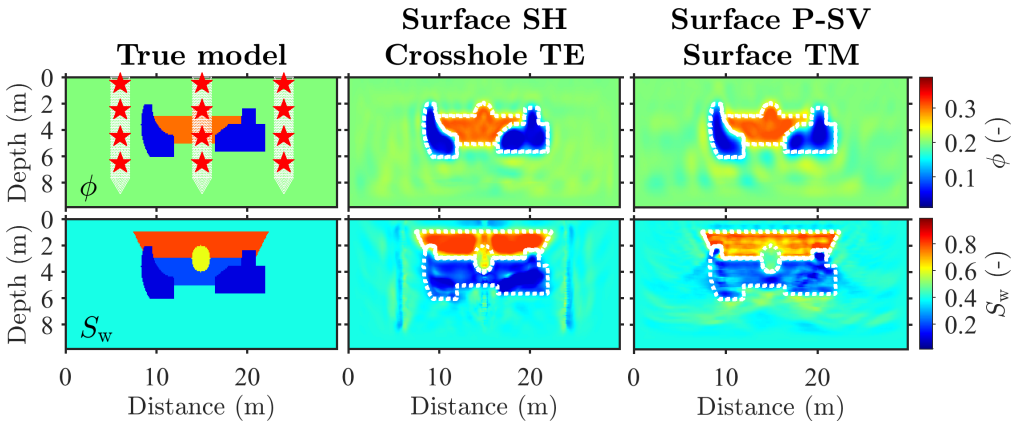

Figure 4.11: The results of indirect JPI using different data simulated by different acquisition geometries in the reference model (the air layer is not displayed). The red stars and white triangles represent the sources (transmitters) and geophones (receivers) used to obtain crosshole data. The results of indirect JPI using surface *P*-SV-wave data and surface TM-wave data are the same as those of the indirect JPI shown in Fig. 4.6.

particular, the final porosity misfit is higher than the initial model misfit when the perturbation of $V_{S\mathrm{ma}}$ > 5 per cent or $\rho_{\mathrm{ma}} > 30$ per cent in conventional JPI and the perturbation of $V_{S\mathrm{ma}} > 10$ per cent in indirect JPI, which means the inversion fails to reduce the model misfit. The perturbations of $\varepsilon_{\mathrm{rma}}$ and σ_{w} in a range of 50 per cent do not seem to affect a good estimation of porosity. When it comes to the final saturation misfit, the seismic rock matrix parameters perturbed in the 50 per cent range have a slight effect on the reconstruction of saturation in conventional JPI and indirect JPI. On the other hand, increasing perturbation of $\varepsilon_{\mathrm{rma}}$ causes an increase in saturation misfit. In conventional JPI, the saturation cannot get any improvement when the perturbation of $\varepsilon_{\mathrm{rma}} > 10$ per cent, while in indirect JPI, even though $\varepsilon_{\mathrm{rma}}$ perturbs to 30 per cent, the saturation result is still better than that of conventional JPI using accurate rock matrix parameters. The change of σ_{w} in conventional JPI seems not to affect saturation misfit because saturation reconstruction has been very poor. Although the saturation misfit increase with σ_{w} perturbation in indirect JPI, the highest saturation misfit (the perturbation of $\sigma_{\mathrm{w}} = 50$ per cent) is comparable with that produced by the conventional JPI without any errors in the rock matrix parameters. These examples prove that indirect JPI can tolerate more errors in the assumed petrophysical models compared with the conventional JPI.

4.3.4 Inconsistent acquisition geometries

It is possible to apply indirect JPI on other types of seismic and GPR data simulated by different acquisition geometries. We show an inversion example of indirect JPI using surface SH-wave data and crosshole TE-wave data (i.e. the mode whose electric field components are always in the measurement plane) in Fig. 4.11 and list all of the twelve combinations in 2-D case in Table 4.2. The surface data acquisition uses the same number and interval of sources (transmitters) and geophones (receivers) as presented in Section 4.3.1. In the simulated *P*-SV (Rayleigh-wave) and SH (Love-wave) surface recordings, Gaussian noise of SNR = 10 dB is added. Similarly, Gaussian noise of SNR = 25 dB is added to the simulated TE and TM surface recordings. In crosshole data acquisition (Fig. 4.11), twelve sources (transmitters) are

Table 4.2: The final petrophysical model misfits of indirect JPI with respect to the combination of different wave types and acquisition geometries. The misfits corresponding to the synthetic example in Section 4.3.2 are shown in bold. It is impossible to obtain a crosshole TM mode wave in practice due to observation limitations.

Wave types / Geometries	P-SV and TE		P-SV and TM		SH and TE		SH and TM	
	ϕ (per cent)	S_{w} (per cent)	ϕ (per cent)	S_{w} (per cent)	ϕ (per cent)	S_{w} (per cent)	ϕ (per cent)	S_{w} (per cent)
Surface and Surface	37.30	45.35	**32.38**	**40.01**	24.66	45.26	23.42	40.03
Surface and Crosshole	33.13	25.90	—	—	24.74	25.06	—	—
Crosshole and Surface	28.25	40.04	28.63	38.01	24.45	40.69	24.54	38.90
Crosshole and Crosshole	29.72	25.19	—	—	27.45	25.49	—	—

deployed in three vertical boreholes located at 6, 15 and 24 m distance; Each borehole has four sources (transmitters) placed from 0.5 to 6.5 m depth; Geophones (receivers) are deployed in the adjacent boreholes from 0.4 to 8 m depth with an interval of 0.2 m. For all simulated crosshole recordings, we add Gaussian noise of SNR = 20 dB. The same inversion configuration as in Section 4.3.2 is used for indirect JPI.

Indirect JPI using SH-wave data can reconstruct porosity more accurately than indirect JPI using *P*-SV-wave data, whether in the surface or crosshole geometries (see Table 4.2 and Fig. 4.11), because SH-wave FWI has fewer model parameters. On the other hand, compared with that obtained by surface TE-wave or TM-wave data, the saturation estimated by crosshole TE-wave data is greatly improved due to that crosshole GPR geometry has better illumination in the area deeper than 3 m (Fig. 4.11). As a result, indirect JPI of SH surface recordings and TE crosshole data provides the best reconstruction of porosity and saturation among all of the twelve combinations in Table 4.2. These tests show that indirect JPI is applicable to various types of waves and acquisition geometries. Therefore, one can make the best use of indirect JPI by selecting the combination wisely for the target of interest.

4.4 Discussion

We propose to apply indirect petrophysical parametrization iteratively in indirect JPI. We have also tried to use this parametrization at different points during the inversion process. One can find the implementation steps and synthetic inversion tests for these attempts in Appendix D. Their results are slightly worse than those obtained by applying this parametrization iteratively. However, they still show a significant improvement over the results of conventional JPI.

We have tested indirect JPI in other synthetic models, such as a checkerboard model (Oberröhrmann et al., 2013), a trench model (Gao et al., 2020) and the models with higher conductivity in the near-surface zone, and observed similar good performance. The success of indirect JPI demonstrates the importance of taking into account the sensitivity of data to petrophysical parameters, which can be divided into two parts: The first is the sensitivity of data to seismic and GPR parameters. For example, the high sensitivity of shallow-seismic data to the *S*-wave velocity, or the sensitivity of multi-offset GPR data to the dielectric permittivity. The second is the sensitivity of seismic and GPR parameters to petrophysical parameters

as we have discussed in Section 4.2.2. Those two parts are handled by indirect JPI through indirect petrophysical parametrization.

4.5 Conclusion

In this paper, we proposed indirect JPI of shallow-seismic and multi-offset GPR data for consistent and accurate imaging. The indirect JPI uses an indirect petrophysical parametrization which is developed based on the sensitivity of shallow-seismic and GPR data to petrophysical parameters. In indirect JPI, seismic data primarily contribute to recovering porosity, while GPR data dominate the reconstruction of saturation. Indirect JPI employs the seismic velocity parametrization and logarithmic parametrization to update seismic and GPR model parameters, respectively. The best reconstructed parameters are then transformed to petrophysical parameters. By doing so, we avoid the risk of damaging the inversion results by some less reliable information.

The 2-D synthetic inversion tests showed that indirect JPI outperforms conventional JPI in reconstructing more accurate petrophysical, seismic and GPR parameters. In particular, indirect JPI has significant improvements in estimating saturation and GPR parameters. The robustness tests indicated that, in JPI, the estimation of porosity rarely depends on Archie's coefficients and the EM rock matrix parameters, and the estimation of saturation is slightly affected by the seismic rock matrix parameters. On the other hand, incorrect seismic rock matrix parameters mainly cause poor porosity reconstruction, and the errors of Archie saturation component and EM rock matrix parameters only affect the reconstruction of saturation. We also found that indirect JPI suffers less from inaccurate *a priori* information by linking only the most reliably estimated seismic and GPR parameters to porosity and saturation, respectively. In addition, indirect JPI is suitable for different wave types (surface waves, transmitted waves, reflected waves, etc.) and acquisition geometries (surface observation, crosshole observation, etc.), and the extension of indirect JPI to include multiple geophysical data is straightforward.

The application and verification of JPI on field data is the next step and needs to be studied in the future. The use of other petrophysical relations and non-petrophysical parametrizations in JPI also needs to be further investigated.

Acknowledgments

This work is financially supported by the China Scholarship Council (No. 201806260258). Tan Qin would like to thank Tilman Steinweg and Mark Wienöbst for their help in developing the WAVE-Toolbox. The authors sincerely thank Thomas Hertweck for polishing the original manuscript and thank the editor Michal Malinowski, Florian Wagner and an anonymous reviewer for their constructive comments.

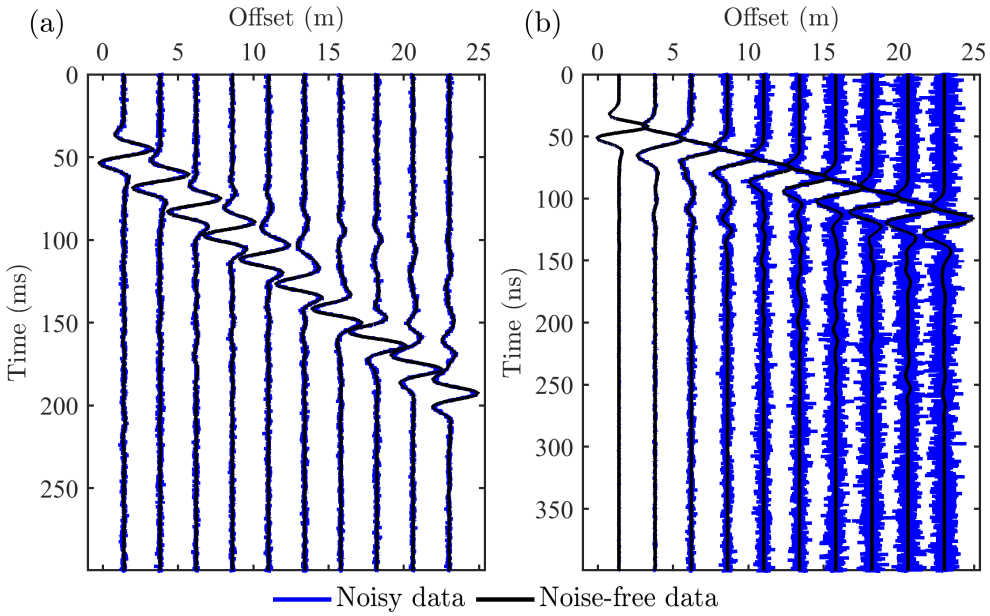

Figure C.1: The noise-free data versus the noise-contaminated data. (a) The vertical particle velocity component seismograms, and (b) the electric field component radargrams of the first shot. Both are trace-by-trace normalized and only every sixth trace is displayed for better visualization. Gaussian noise with a signal-to-noise ratio (SNR) of 10 dB is added to the seismic data. In the radargrams, Gaussian noise of SNR = 25 dB is added similarly.

Data availability

An open-source software (GPL) package containing the source code, synthetic models and data used in this paper is provided in the WAVE-Toolbox (https://github.com/WAVE-Toolbox).

C Data comparison

Fig. C.1 shows the impact of Gaussian noise on the seismograms and radargrams of the first shot. It is observed that the near-offset traces (offset $<$ 10 m) and far-offset traces (offset $>$ 10 m) in the seismic section are influenced slightly by the noise. However, radargrams are significantly disturbed by the noise in the far-offset range. For a more straightforward comparison, we plot the SNR as a function of offset in Fig. C.2. The SNR of seismograms is about 9 dB for the entire section because we eliminate the traces of offset $<$ 1 m and use the maximum offset of 25 m within which the amplitude decay of Rayleigh waves is weak. Although we add Gaussian noise with a higher SNR to the GPR data, the SNR of the radargrams decreases rapidly with offset and is smaller than 0 dB after 10 m offset which means the effective signal is below noise level in those areas. This strong attenuation results from the geometric propagation of EM waves and the presence of conductivity in underground space. Unlike crosshole GPR data, in which transmitted waves are dominant and contribute to the inversion, surface GPR data are dominated by the direct waves and provides less useful signals (reflected waves) for the inversion.

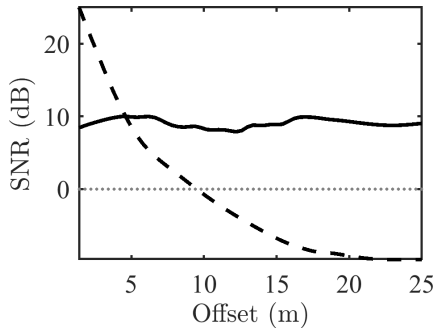

Figure C.2: The SNR as a function of offset, which is calculated from the corresponding seismograms and radargrams shown in Fig. C.1. The solid line is the SNR of the seismograms and the dashed line the SNR of the radargrams. The dotted line marks the reference SNR = 0 dB, that is noise and signal magnitudes are equal.

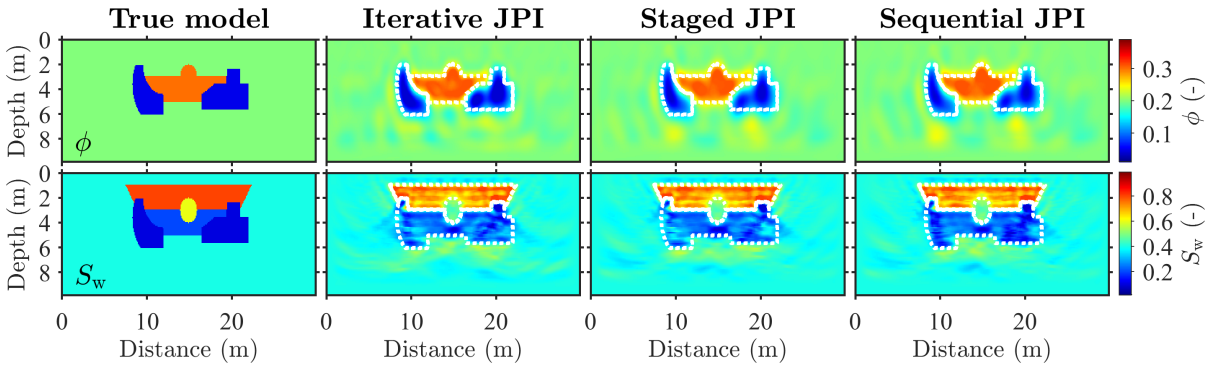

Figure D.1: The results of JPI using indirect petrophysical parametrization at different points during the inversion. The results of iterative JPI are the same as those of the indirect JPI shown in Fig. 4.6.

D Indirect petrophysical parametrization comparison

There are three options to implement indirect petrophysical parametrization in indirect JPI at different points during the inversion. The first option, as shown by steps (1–4) in Fig. 4.3(b), iteratively uses this parametrization. The second option is to execute this parametrization at the end of each inversion stage in the multi-scale strategy. The third option is to apply seismic inversion and GPR inversion sequentially and only use steps (2) and (3) once after the seismic inversion. For convenience, we refer to the indirect JPI using those three options as iterative JPI, staged JPI and sequential JPI. Compared with iterative JPI, staged JPI and sequential JPI possess several gaps where no petrophysical parameter exchange happens. We set those gaps to give the inversion a certain degree of freedom and see whether the reconstruction benefits from it.

We implement staged JPI and sequential JPI with the same inversion configuration as the iterative JPI shown in Section 4.3.2 and simultaneously reconstruct all seven parameters (ϕ, S_w, V_P, V_S, ρ, ε and σ). Their results are shown in Fig. D.1. The use times of indirect petrophysical parametrization in iterative JPI, staged JPI and sequential JPI are 55, 5 and 1, respectively. In the porosity model, staged JPI and sequential JPI generate more artefacts in the region deeper than 6 m. Iterative JPI is able to produce

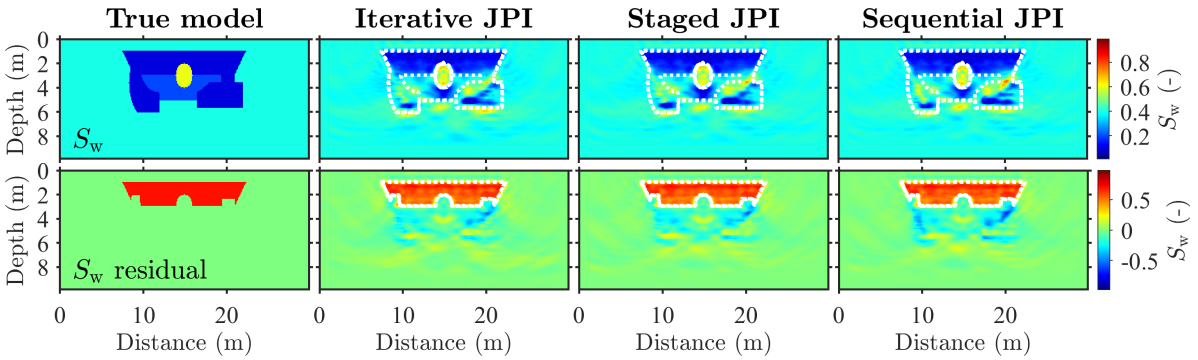

Figure D.2: The saturation inversion results of the time-lapse model and the saturation residuals caused by water outflow.

better porosity and saturation models due to the more frequent exchange of information between shallow-seismic and GPR FWI.

Dynamically monitoring the groundwater is one of the potential applications of JPI. To further evaluate the three options of indirect petrophysical parametrization, we introduce a time-lapse model simulating a water outflow process. The difference between the reference and time-lapse model is saturation: We reduce the saturation of the first trapezoid layer from 0.8 in the reference model (Fig. D.1) to 0.1 in time-lapse model (Fig. D.2). We conduct three JPIs again and display the inversion results of saturation in Fig. D.2. Iterative JPI, staged JPI and sequential JPI are able to recover a reliable saturation layer at 1–3 m depth. However, sequential JPI produces more saturation artefacts at 3–5 m depth. Most of them appear on the interfaces with distinct contrast in porosity. They are caused by the worse reconstructed porosity, as the seismic FWI is not aided by the GPR FWI in sequential JPI. Those synthetic tests demonstrate that the timely mutual constraint of the two data sets would help reduce the inversion artefacts. Hence we adopt an iterative strategy for indirect JPI in the main body of this paper.

5 Consistent Imaging of Near-Surface Targets Using Indirect Joint Petrophysical Inversion of Shallow-Seismic and Multi-Offset Ground-Penetrating Radar Field Data

Prepared as: Qin, T., Bohlen, T., Allroggen, N., & Pan Y. Consistent Imaging of Near-Surface Targets Using Indirect Joint Petrophysical Inversion of Shallow-Seismic and Multi-Offset Ground-Penetrating Radar Field Data. *Geophysical Research Letters*.

Abstract

In near-surface surveys, different geophysical methods may yield conflicting geological interpretations for the same targets. To solve this issue, we have developed an indirect joint petrophysical inversion (JPI) integrating shallow-seismic and multi-offset ground-penetrating radar (GPR) data. These data are used to reconstruct different petrophysical parameters depending on their sensitivities to parameters. Our application to Love-wave and surface GPR field data indicates that indirect JPI accurately reveals the degree of consolidation and water content in the subsurface. It suggests that indirect JPI can provide consistent images that are difficult to achieve with individual inversions and has a higher efficiency of information exchange than joint structural inversion. More importantly, this method is robust when there are uncertainties in petrophysical a priori information, significantly promoting its applicability in a broader range of situations.

Plain Language Summary

Shallow seismic and ground-penetrating radar (GPR) data have different sensitivities to near-surface materials. Separate inversions of these data may produce inconsistent results, causing difficulties in interpretation. For this reason, we have developed an indirect joint petrophysical inversion (JPI) to combine these data by considering their sensitivity differences. We apply this method to field data and find that it works nicely in simultaneously imaging multiple physical properties. It reveals that indirect JPI can offer a consistent geological interpretation that takes full advantage of each data. Compared to another joint inversion approach, our approach has greater strengths in information exchange. As a major contribution of this study, indirect JPI reduces the reliance on site information and empirical relations, which is a strong evidence that it can be widely used.

5.1 Introduction

The near-surface area, a few tens of meters underground, are closely related to social development and life safety. A detailed characterization of this area is essential for urban construction, engineering exploration, environmental assessment, archaeological investigation, hydrological monitoring, polar research (Everett, 2013). In near-surface surveys, shallow-seismic and ground-penetrating radar (GPR) methods are widely used geophysical techniques. Shallow-seismic data are sensitive to the mechanical parameters in the subsurface but cannot identify the moisture distribution (Gassmann, 1951). GPR data are highly sensitive to the water content but the penetration depth is limited by the conductivity (Annan, 2005). Separate inversions of these data may provide inconsistent interpretations and not fully use their advantages. Combining the two data via joint inversion can provide supplementary information for each inversion, decreasing uncertainty and avoiding ambiguity (Linde et al., 2008).

There are numerous joint inversion methods proposed to integrate multiple geophysical data (Linde et al., 2008; Wagner et al., 2019). They can be mainly divided into two classes: joint structural inversion (JSI) and joint petrophysical inversion (JPI). JSI assumes that different geophysical parameters have similar distributions. The cross-gradient function is one of the most common methods to quantify structural similarity (Gallardo and Meju, 2003). On the other hand, JPI supposes that different geophysical parameters are a function of the petrophysical properties. In practice, JSI has a broader application than JPI since it does not strictly require that geophysical parameters obey a certain petrophysical relation. However, the structural constraint is weaker than the petrophysical constraint and has a risk of generating parameter relations that deviate from the natural geological environment.

JPI is a valuable technique that combines all geophysical parameters through petrophysical relations. As a double-edged sword, the applicability of JPI is limited by a priori information, including the assumed petrophysical relations and the rock matrix properties. To address this problem, Qin et al. (2022b) analyzed the sensitivity of seismic and GPR data to petrophysical parameters and proposed an indirect JPI method. Their method indirectly uses petrophysical parametrization so that seismic and GPR data contribute to porosity and saturation updates, respectively. Compared to the conventional JPI (Abubakar et al., 2012), indirect JPI increases the robustness of the inversion under inaccurate a priori information and thus improves the applicability of the algorithm.

Indirect JPI was developed from full waveform inversion (FWI) of shallow-seismic and GPR data, a successful imaging technique that provides high-resolution multi-parameter subsurface models with geological interests (Tarantola, 1984). Shallow-seismic wavefield is dominated by Love or Rayleigh waves, which are the result of the interference of SH or *P*-SV waves. Compared to Rayleigh-wave FWI, Love-wave FWI has fewer model parameters, lower computational costs, and less trade-offs between multiple parameters (Dokter et al., 2017). Multi-offset surface GPR data are easy to obtain and have higher spatial resolution than crosshole GPR data of the same frequency (Lavoué et al., 2014). Here, we apply indirect JPI for the first time to Love-wave and multi-offset surface GPR field data acquired in Rheinstetten, Germany. We compare indirect JPI with individual inversions and JSI and verify our results with borehole samples and direct-push technology (DPT) measurements. The main goal of this paper is to assess the applicability of indirect JPI to solve practical problems.

5.2 Joint Petrophysical Inversion Method

The theoretical parts of indirect JPI, including detailed sensitivity analysis and comparison with conventional JPI, have been described in Qin et al. (2022b). Therefore, we only briefly introduce the main idea of this method in this paper. In indirect JPI of Love-wave and multi-offset GPR data, we use the least-squares objective functions to quantify the data misfit:

$$\Phi_{\mathrm{S}}\left(\mathbf{m}_{\mathrm{S}}\right)=\tfrac{1}{2}||\mathbf{d}_{\mathrm{S}}^{\mathrm{syn}}\left(\mathbf{m}_{\mathrm{S}}\right)-\mathbf{d}_{\mathrm{S}}^{\mathrm{obs}}||^{2}, \tag{5.1}$$

$$\Phi_{\mathrm{EM}}\left(\mathbf{m}_{\mathrm{EM}}\right)=\tfrac{1}{2}||\mathbf{d}_{\mathrm{EM}}^{\mathrm{syn}}\left(\mathbf{m}_{\mathrm{EM}}\right)-\mathbf{d}_{\mathrm{EM}}^{\mathrm{obs}}||^{2}, \tag{5.2}$$

subject to

$$f_{\mathrm{S}}(\mathbf{m}_{\mathrm{S}},\mathbf{m}_{\mathrm{P}})=0\,, \tag{5.3}$$

$$f_{\mathrm{EM}}(\mathbf{m}_{\mathrm{EM}},\mathbf{m}_{\mathrm{P}})=0\,, \tag{5.4}$$

where Φ_{S} and Φ_{EM} are the objective functions of seismic and electromagnetic (EM) data inversions, respectively; $\mathbf{d}_{\mathrm{S}}^{\mathrm{syn}}$ is the synthetic seismograms and $\mathbf{d}_{\mathrm{EM}}^{\mathrm{syn}}$ the synthetic radargrams; $\mathbf{d}_{\mathrm{S}}^{\mathrm{obs}}$ is the observed seismograms and $\mathbf{d}_{\mathrm{EM}}^{\mathrm{obs}}$ the observed radargrams; $\mathbf{m}_{\mathrm{S}}$ represents seismic parameters including the *S*-wave velocity V_S and density ρ; $\mathbf{m}_{\mathrm{EM}}$ represents EM parameters including dielectric permittivity ε and electric conductivity σ. $\mathbf{m}_{\mathrm{P}}$ is petrophysical parameters, i.e., porosity ϕ and saturation S_{w} in our study. The function f_{S} and f_{EM} are the petrophysical relations used to connect $\mathbf{m}_{\mathrm{P}}$ with $\mathbf{m}_{\mathrm{S}}$ and $\mathbf{m}_{\mathrm{EM}}$, respectively. The seismic and EM wave equations are solved respectively by the 2D finite-difference methods (Virieux, 1984; Yee, 1966). We iteratively update the model parameters $\mathbf{m}$ ($\mathbf{m}_{\mathrm{S}}$, $\mathbf{m}_{\mathrm{EM}}$ and $\mathbf{m}_{\mathrm{P}}$) by using the conjugate-gradient method to minimize Φ (Φ_{S} and Φ_{EM}) (Gao et al., 2020). A shown in equations (5.1) and (5.2), we use separate objective functions that already ensure a reasonable contribution of seismic and GPR data, thus avoiding the need to calculate the data weighting matrix and scale factors in the traditional joint inversion (Moorkamp et al., 2013).

We implement indirect JPI in each iteration through the following four steps:

(1). We calculate the seismic parameters (V_S and ρ) from the ϕ and S_{w} models through f_{S} (equation 5.3).

(2). In Love-wave FWI (equation 5.1), we use the velocity parametrization to compute the gradients and update the models of V_S and ρ. The ϕ model is then calculated from the recovered V_S and ρ by f_{S} (equation 5.3).

(3). The ϕ and S_{w} models are used to calculate the EM parameters (ε and σ) via f_{EM} (equation 5.4).

(4). We employ the logarithmic parametrization to compute the ε gradient and update the ε model in GPR FWI (equation 5.2). We then estimate S_{w} model from the reconstructed ε model by f_{EM} (equation 5.4).

In this paper, we use f_{S} the Gassmann fluid substitution model without bulk modulus (Gassmann, 1951), and f_{EM} the complex refractive index model (CRIM) (Birchak et al., 1974) and Archie's equation (Archie

et al., 1942). f_S used in step (2) and f_{EM} used in step (4) are exactly the same as those used in steps (1) and (3), but with different expressions for their respective purposes. Readers are recommended to refer to equations (4–9) in Qin et al. (2022b) for comparison. Steps (2) and (4) include seismic velocity parametrization (Köhn et al., 2012) and logarithmic parametrization (Meles et al., 2010) which have proven to be efficient for shallow-seismic and GPR FWIs, respectively.

In contrast to conventional JPI that directly uses the petrophysical parametrization (Abubakar et al., 2012), indirect JPI applies non-petrophysical parametrizations to compute the gradients and update the models of seismic and GPR parameters. Step (2) transforms the *S*-wave velocity to porosity through shear modulus since shallow-seismic FWI can reconstruct the *S*-wave velocity in high-quality (Pan et al., 2019). Likewise, step (4) transforms permittivity to saturation because permittivity is the parameter that can be most effectively estimated by GPR FWI (Klotzsche et al., 2019). We then derive these hard-to-recover parameters (density and conductivity) in steps (1) and (3). Thus, we use only the reliable information of each geophysical data, reducing the trade-offs of multi-parameter inversions. In the above implementation steps, we consider the sensitivity of seismic and EM parameters to petrophysical parameters and the sensitivity of shallow-seismic and GPR data to seismic and EM parameters. Their combination ensures that the information exchange between the two inversions is efficient and makes the joint inversion robust.

5.3 Application to Near-Surface Imaging

5.3.1 Inversion Preparation

We acquired the data in Rheinstetten, Germany, where a V-shaped trench, Ettlinger Line, was excavated on a sedimentary plain covered by gravel and sand from the Rhine river. This trench was refilled with sand a few decades ago and became invisible from surface at the test site. The current ground layer is composed of partially saturated soil, and the groundwater table is below 6 m depth (Wittkamp et al., 2019). We show field data examples in Figure 5.1 and the initial seismic and GPR models in Figure 5.2 where the trench is located between 17 and 27 m in horizontal direction. We conducted a 2D investigation with seismic and GPR profiles perpendicular to the Ettlinger Line. The profile crosses the previous 3D shallow-seismic survey area from the southwest corner to the northeast corner (Pan et al., 2021).

To record the Love waves, we laid out 48 geophones (horizontal crossline component) from -3.5 m to 43.5 m in the horizontal direction and used a hammer to blow on a steel beam source in the crossline direction. We acquired 12 seismograms with a shot spacing of 4 m and a fixed geophone spreading (see Figure 5.1a). Our GPR data were recorded using a single-channel GPR system with a transmitter of 200 MHz nominal center frequency. We deployed the transmitter-receiver orientation in HH mode and acquired 18 radargrams with a source spacing of 2 m. Unlike seismic data acquisition, GPR data acquisition used a ‘walk-away’ method where we fixed the transmitter and moved the receiver on the right side of the transmitter with offsets ranging from 0.2 m to 8 m (see Figure 5.1b). The steps for Love-wave and GPR data preprocessing are similar to those used in Wittkamp et al. (2019) and Qin et al. (2022a).

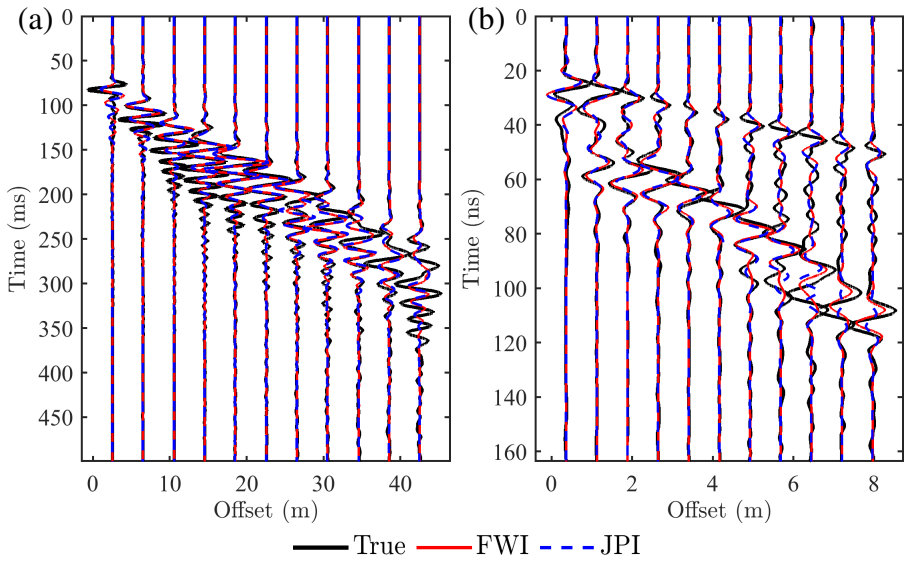

Figure 5.1: Comparison of the observed (true) data with the synthetic data corresponding to the results of individual full-waveform inversion (FWI) and indirect joint petrophysical inversion (JPI). (a) The horizontal velocity seismogram of the 1st shot, shown once every four traces. (b) The horizontal electric field radargram of the 16th ground-penetrating radar (GPR) source, shown once every twenty traces. For comparison, each trace is divided by the maximum amplitude of each trace of the observed data.

We use similar model space but different grid sizes for seismic models (0.12 m×0.12 m) and GPR models (0.04 m×0.04 m), as shown in Figure 5.2. The petrophysical models shown in Figure 5.3 have the same grid size as seismic and GPR models in each inversion. In GPR models, the air layer of 1 m thickness remains constant during the inversion and thus is not displayed. The seismic and GPR initial models change gradually with depth, where the near-ground values of the *S*-wave velocity, permittivity and conductivity are estimated from the data (Xia et al., 2012; Annan, 2005), and the density model is set empirically. We assume that the porosity initial model decreases with depth and saturation is a homogeneous half-space (Figure 5.3). We calculate the rock matrix parameters (including the groundwater conductivity) from the initial models based on the petrophysical relations using the critical porosity ϕ_c = 0.4 in Gassmann's equations, the tortuosity factor a = 2, the cementation exponent m = 1.4, and the saturation exponent n = 1.13 in Archie's equation. To account for the attenuation of *S*-wave velocity, we use a viscoelastic solver with one relaxation mechanism of 40 Hz relaxation frequency and set the attenuation level $Q_S \approx 13.3$ (Gao et al., 2020).

We adopt the inversion setups similar to those of Qin et al. (2022b), except that the frequency bands of GPR inversion vary from 5 to 30, 40, 50, 70, and 100 MHz, allowing for stable convergence. In addition, at the beginning of each inversion stage, we estimate the source wavelets by a correction filter (Groos et al., 2014). We apply a 1D Gaussian filter in the horizontal direction to the gradient to suppress the components shorter than the dominant wavelength.

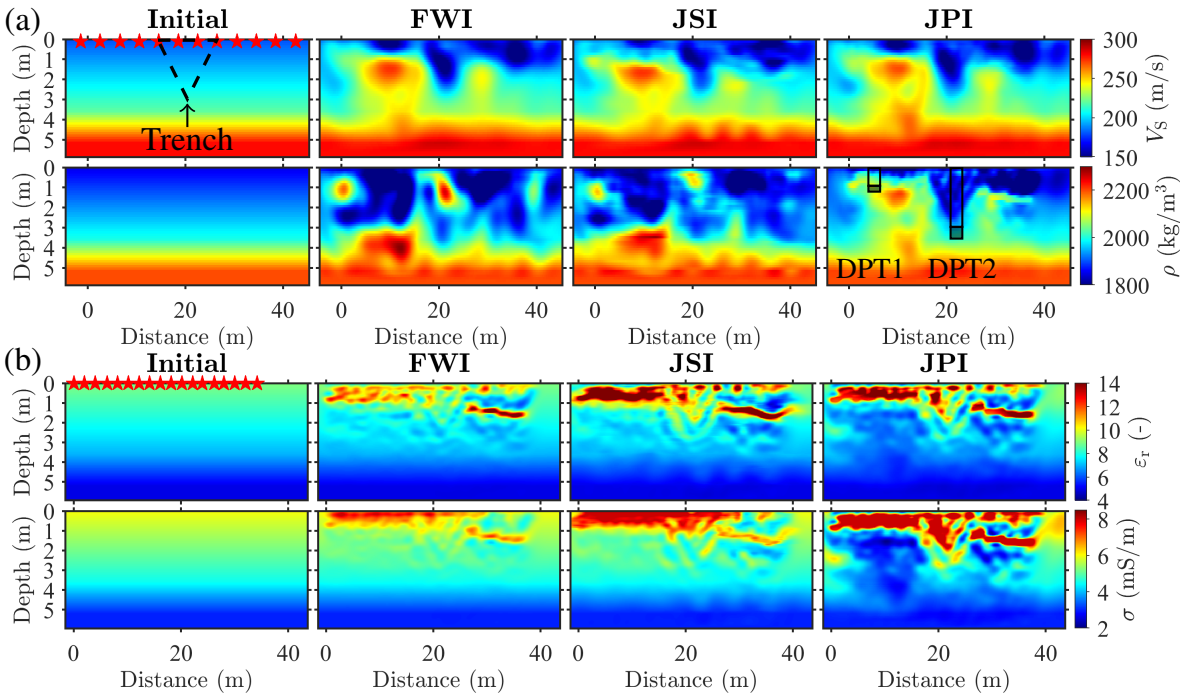

Figure 5.2: (a) Seismic models (S-wave velocity V_S and density ρ) and (b) GPR models (relative permittivity ε_r and conductivity σ). The four columns are the initial models, the reconstructed models of individual FWI, the reconstructed models of joint structural inversion (JSI), and the reconstructed models of indirect JPI, respectively. In the initial models, the red stars are the sources and the dashed triangle outlines the cross-section of the Ettlinger Line. The borehole histograms overlaid on the density model are the direct-push technology (DPT) results (Figures E.1a and E.1b), where the transparent and translucent areas present the unconsolidated and consolidated soil, respectively.

5.3.2 Inversion Results

As shown in Figures 5.2 and 5.3, the seismic, GPR, and petrophysical models reconstructed by indirect JPI successfully illustrate the presence of the Ettlinger Line. On the on hand, the S-wave velocity result is comparable to that of the 3D shallow-seismic FWIs of Pan et al. (2021). On the other hand, the permittivity result is in high agreement with the GPR migration image of Wittkamp et al. (2019). Due to the constraint of petrophysical relations, the density model also reveals the exact shape of the trench, which is difficult to see from past investigations of Wittkamp et al. (2019) and Gao et al. (2020). For the same reason, the conductivity model has a similar structure to the permittivity model. Note that in the GPR models, the boundaries of the trench become less visible compared to the high permittivity and high conductivity anomalies in the interior. It is due to two reasons. One is that the high conductivity environment near the surface degrades the penetration depth of the GPR signal. The other is that the trench boundaries are the transition zone of porosity (see Figure 5.3), where the low porosity outside makes saturation less sensitive to GPR data (Qin et al., 2022b).

For comparison, we show individual FWIs and JSI results in Figure 5.2. Individual FWIs use the same objective functions (equations 5.1 and 5.2) but are not subject to petrophysical relations. JSI is the same as individual FWIs and is additionally constrained by three cross-gradients, i.e., $t(V_S, \varepsilon)$, $t(V_S, \rho)$

and $t(\varepsilon, \sigma)$ where t is the cross product of the spatial gradients (Gallardo and Meju, 2003). The *S*-wave velocity models estimated by three inversions are of good quality, while the density models reconstructed by SH FWI and JSI show a high-density anomaly inside the trench and a low-density anomaly to the left. These anomalies may be artifacts caused by the crosstalk from *S*-wave velocity and the low sensitivity of Love-wave data to density. JSI improves the density model by correcting the trench bottom and reducing these artifacts. Nevertheless, these improvements are much fewer than those provided by indirect JPI.

Individual GPR FWI outlines the shape of the trench in the permittivity model but not in the conductivity model due to the weak sensitivity of GPR data to conductivity. JSI provides a valuable conductivity model by transforming the structure information from the more reliable permittivity model. Since surface GPR data is dominated by the short-wavelength information (Lavoué et al., 2014), the long-wavelength background models are difficult to update by GPR FWI alone. To overcome this drawback, joint inversions attempt to use the complementary information from Love-wave data. Unfortunately, JSI fails to realize this goal because the long-wavelength seismic background, e.g., the high *S*-wave velocity anomaly on the left side of the trench, has low cross-gradient values with the GPR model. Indirect JPI successfully offers the needed information for GPR FWI through the porosity model, thus allowing reconstruction of the low permittivity and low conductivity background.

The seismogram fitting of SH FWI and indirect JPI are comparable (Φ_S is 71.9759 and 73.3305 $m^2\,s^{-2}$ for each), while the radargram fitting of GPR FWI (Φ_{EM} = 2.7767E-4 $V^2\,m^{-2}$) is slightly better than that of indirect JPI (Φ_{EM} = 2.8824E-4 $V^2\,m^{-2}$). In Figure 5.1b, the reflected waves dominate in the radargram at offsets greater than 2 m due to the strong permittivity contrast on the right side of the trench. Indirect JPI matches well to the reflection events within the first 5 m offsets and becomes worse than GPR FWI when the offset is over 5 m. The air waves are hard to fit in both inversions as we use the 3D / 2D transformation of the reflected wave (Forbriger et al., 2014) and our 2D solver cannot well simulate the radiation patterns and antenna-ground coupling. Overall, the data misfit of indirect JPI is slightly higher than that of GPR FWI (within 4 per cent) because of the interaction of seismic and GPR data and the additional petrophysical constraint.

We compare indirect JPI with individual petrophysical inversions (IPIs) in Figure 5.3. SH IPI follows equations (5.1) and (5.3), and GPR IPI follows equations (5.2) and (5.4). Unlike indirect JPI, IPIs use petrophysical parametrization to directly update the petrophysical models, which means that weak sensitivity information, e.g., conductivity gradient, is also included. Consequently, SH IPI and GPR IPI generate conflicting results at 0.5–1.5 m depth on the right side of the trench. SH IPI describes this region as having high porosity and high saturation anomalies, while GPR IPI interprets it as a low porosity and low saturation anomalies. Neither of them can minimize the two data misfits simultaneously. However, indirect JPI describes this area as a water-poor layer consisting of high porosity and low saturation anomalies, ensuring that both data are well matched.

5.3.3 Comparison of Inversion Results with Borehole Information

To validate our interpretations of the density model, we compare two independent DPT measurements (DPT1 and DPT2) with the result obtained from indirect JPI in Figure 5.2a and Figures E.1a and E.1b

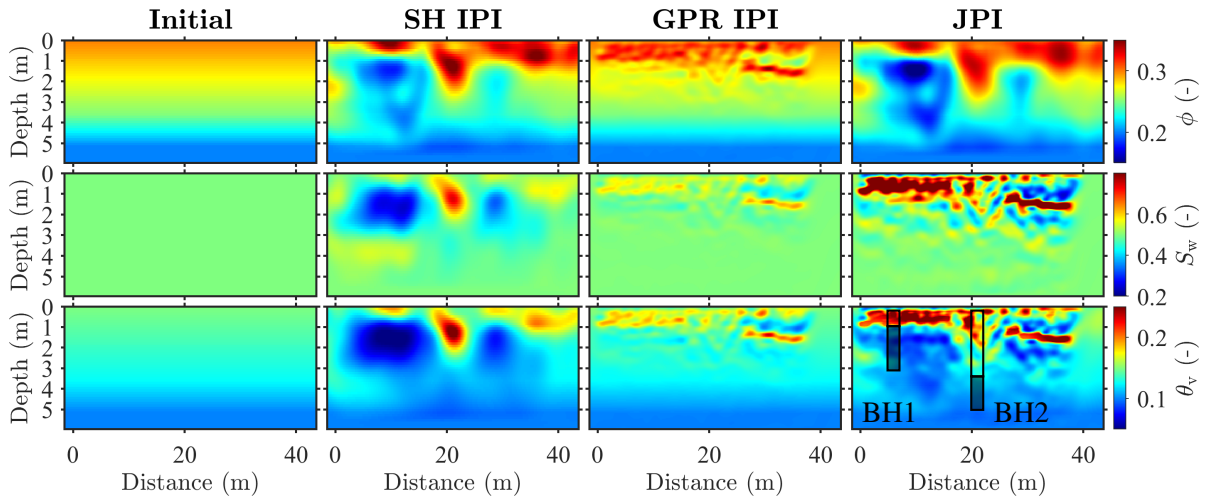

Figure 5.3: Petrophysical models (porosity ϕ and saturation S_w) and volumetric water content model θ_v. The four columns are the initial models, the reconstructed models of SH individual petrophysical inversion (IPI), the reconstructed models of GPR IPI, and the reconstructed models of indirect JPI, respectively. The borehole histograms overlaid on the water content model are the gravimetric water content θ_g given by borehole soil samples (Figures E.1c and E.1d), where the transparent and translucent areas exhibit high and low water content, respectively.

in Supporting Information. Our DPT measurements reveal a change from loose topsoil to compacted subsoil at 1 m depth at the DPT1 position and an interface between refilled sand and underlying soil at 3 m depth at the DPT2 location, which corresponds to the bottom of the trench (see Figure E.1b). These findings agree with the density model of indirect JPI in Figure 5.2a and prove that indirect JPI provides a more realistic density model than SH FWI and JSI.

Petrophysical inversions give the volumetric water content $\theta_v = \phi S_w$. We drilled two boreholes (BH1 and BH2) close to the DPT locations and measure the gravimetric water content θ_g. Figure 5.3 and Figures E.1c and E.1d suggest that indirect JPI fits well with the BH1 measurement in both high and low water content area. For BH2, SH IPI overestimates the water content in the trench and GPR IPI underestimates it. Indirect JPI matches the BH2 measurement with minimum error, although the performance at the trench bottom is not ideal due to the two limitations of the GPR technique mentioned above.

5.3.4 Robustness Tests

It is well known that the performance of petrophysical inversions tends to depend on a priori information, i.e., the rock matrix parameters and Archie's coefficients. Instead of giving the rock matrix parameters explicitly, we compute them from the initial models in this study. If we fix seismic and GPR initial models, the rock matrix parameters are calculated from petrophysical initial models $\mathbf{m}_P^{\text{initial}}$. To evaluate the inversion robustness, we take the mean structural similarity (MSSIM) index to measure the fidelity of the reconstructed models of indirect JPI using different $\mathbf{m}_P^{\text{initial}}$ relative to the reference. We adopt the same settings as Boniger and Tronicke (2010) to compute the MSSIM index. The closer the MSSIM index is to one, the more similar the two comparison objects are (Wang et al., 2004).

Table 5.1: Model MSSIM Index Obtained by Indirect JPI with Different Petrophysical Initial Models.

$\mathbf{m}_{\mathrm{P}}^{\mathrm{initial}}$		Seismic MSSIM		GPR MSSIM	
ϕ	S_{w}	V_S	ρ	ε	σ
0.2	0.5	0.9915	0.9877	0.9757	0.9000
0.2	0.7	0.9922	0.9851	0.9344	0.8601
0.3	0.5	0.9998	0.9986	0.9992	0.9929
0.3	0.7	0.9990	0.9926	0.9653	0.9276

Note. MSSIM is the mean structural similarity. $\mathbf{m}_{\mathrm{P}}^{\mathrm{initial}}$ is the initial model of ϕ and S_{w}.

Table 5.1 shows the MSSIM index where indirect JPI using $\mathbf{m}_{\mathrm{P}}^{\mathrm{initial}}$ shown in Figure 5.3 is the reference and other four indirect JPIs use homogeneous $\mathbf{m}_{\mathrm{P}}^{\mathrm{initial}}$ in half-space. Note that we adjust the initial permittivity model to a homogeneous half-space ($\varepsilon_{\mathrm{r}} = 9$) in this test to fulfill the requirement of CRIM. Compared to indirect JPI results shown in Figures 5.1–5.3, it leads to a slight change in data fitting (difference within 1 per cent) and petrophysical model reconstruction (MSSIM > 0.99) in the reference JPI but does not impact the conclusions we have made (see Figure E.2). Table 5.1 reveals that when using initial models with high porosity and medium saturation ($\phi = 0.3$ and $S_{\mathrm{w}} = 0.5$), indirect JPI obtains reconstruction results almost the same (MSSIM > 0.99) as the reference. This is due to the same initial saturation model and the higher sensitivity of the seismic and GPR parameters to the petrophysical parameters in this case and, therefore, the more efficient information exchange in indirect petrophysical parametrization (Qin et al., 2022b). When we change $\mathbf{m}_{\mathrm{P}}^{\mathrm{initial}}$ to different values, the seismic model estimates are stable (MSSIM > 0.985) compared to the GPR model estimates since the former are mainly related to porosity but the latter is governed by porosity and saturation. In the latter, we observe larger variations in conductivity than permittivity because conductivity is indirectly derived from permittivity and the seismic models (see implementation steps in Section 5.2) and is susceptible to cumulative errors. However, even in the inversion with the maximum conductivity difference (MSSIM = 0.8601 when $\phi = 0.2$ and $S_{\mathrm{w}} = 0.7$) from the reference, the reconstructed seismic and GPR models can delineate subsurface structures similar to the indirect JPI results shown in Figure 5.2 and provide a meaningful geological interpretation (see Figure E.2). It indicates that the separate use of seismic and GPR data promotes the robustness of the petrophysical inversions.

We also test indirect JPI with the same initial conductivity model (Figure 5.2) but different Archie's coefficients, corresponding to different groundwater conductivities (see Figure E.3). The range of m between 0.4 and 2.4 and n between 1.13 and 3.00 gives almost identical results (MSSIM > 0.99). It proves that Archie's coefficients in the test range do not pose any problems for indirect JPI, which should be attributed to the abandonment of weakly sensitive conductivity in indirect petrophysical parametrization (Qin et al., 2022b). Overall, these examples demonstrate that indirect JPI does not need to know accurate a priori information. With suitable seismic and GPR initial models, any petrophysical initial models and Archie's coefficients in a reasonable range can produce similar results. This strength makes indirect JPI a promising technique that can be easily applied to many field environments with as few assumptions about petrophysical relations as possible.

5.4 Conclusion

In this paper, we applied indirect JPI to Love-wave and multi-offset surface GPR field data for consistent imaging of the near-surface targets. Indirect JPI takes into account the sensitivity of shallow-seismic and GPR data to petrophysical parameters. The application at the Rheinstetten test site showed that this approach not only outperforms individual FWIs and JSI in reconstructing seismic and GPR parameters, but also provides more reliable petrophysical models than IPIs. Indirect JPI presented significant improvements in estimating saturation, density, and conductivity, therefore reducing the ambiguity and uncertainty of single geophysical techniques and facilitating the final geological interpretations, such as determining groundwater distribution and facies stratification. This study also suggested that this approach has a more efficient information communication due to the solid link of petrophysical relations. Thanks to the separate contributions of seismic and GPR data, indirect JPI reduced the reliance on a priori information and exhibited great potential for real-world applications. Our observations are in agreement with the DPT and borehole measurements.

Acknowledgments

The data used in this work can be obtained from repository KITopen (to be continue). An open-source software (GPL) package containing the source code used in this paper is provided in the WAVE-Toolbox (https://github.com/WAVE-Toolbox) and archived at repository Zenodo (https://doi.org/10.5281/zenodo.6875824). This work is financially supported by the China Scholarship Council (No. 201806260258). Tan Qin would like to thank Tilman Steinweg and Mark Wienöbst for their help in developing the WAVE-Toolbox. The authors sincerely thank Lars Houpt, Felix Bögelspacher, Michael Mayer, Hagen Steger, Roland Helfer, Philipp Koyan, and the master students at Geophysical Institute, Karlsruhe Institute of Technology for their help in field data acquisition.

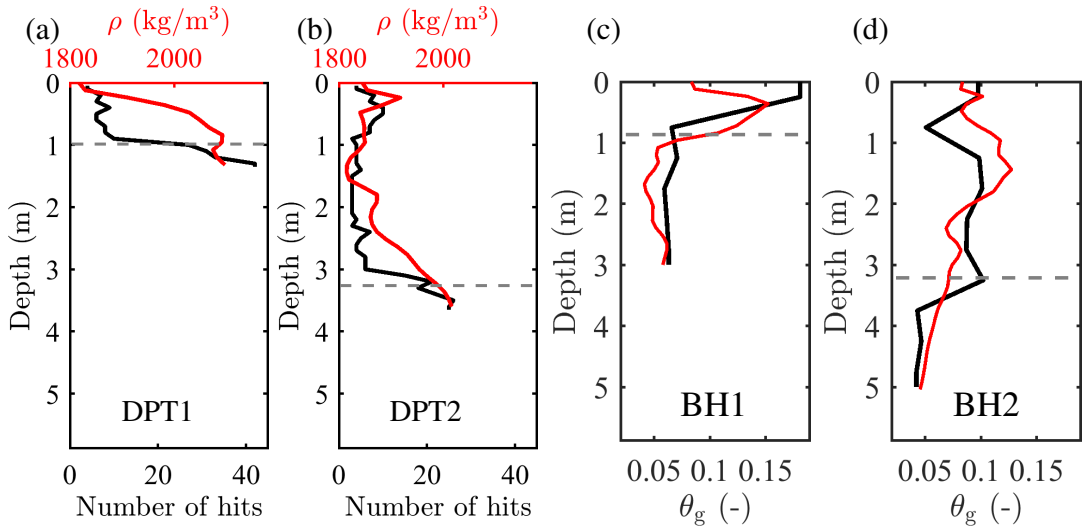

Figure E.1: Comparison of indirect joint petrophysical inversion (JPI), direct-push technology (DPT), and borehole measurements. The black lines in (a) and (b) are the number of hits measured by DPT, and the black lines in (c) and (d) are the gravimetric water content of borehole soil samples. The red lines are the results of indirect JPI, i.e., density in (a) and (b) and gravimetric water content in (c) and (d). The gray dashed lines mark the interfaces between transparent and translucent areas of the borehole histograms shown in Figures 5.2 and 5.3.

E Supporting Information

Figure E.1 is the results of direct-push technology (DPT) and borehole measurements. In DPT measurement, a metal pile was hit by a free-falling slide hammer (10 kg) and thus pushed into the ground. By recording the number of hits per 0.1 m depth pushed in, we measured the consolidation degree in the subsurface. The larger the number, the more compact the soil. Figures E.1a and E.1b show the results of DPT1 and DPT2, respectively. In borehole measurements, we collected soil samples every 0.5 m depth and then measured the gravimetric water content of the soil samples by drying operation in the laboratory. Petrophysical inversions give the gravimetric water content $\theta_g = \theta/(1-\theta)$ where $\theta = \theta_v/\rho$ (θ_v is the volumetric water content and ρ is density). Figures E.1c and E.1d show the results of BH1 and BH2, respectively. Note the uncertainty present in the density of converting θ_v to θ_g is another factor impacting the fitting of two curves.

Figure E.2 is the indirect joint petrophysical inversion (JPI) results with different petrophysical initial models. The models shown in Figures E.2a and E.2b are used to calculate the mean structural similarity (MSSIM) index in Table 5.1.

Figure E.3 illustrates the effect of different Archie's coefficients on the reconstruction of the petrophysical parameters.

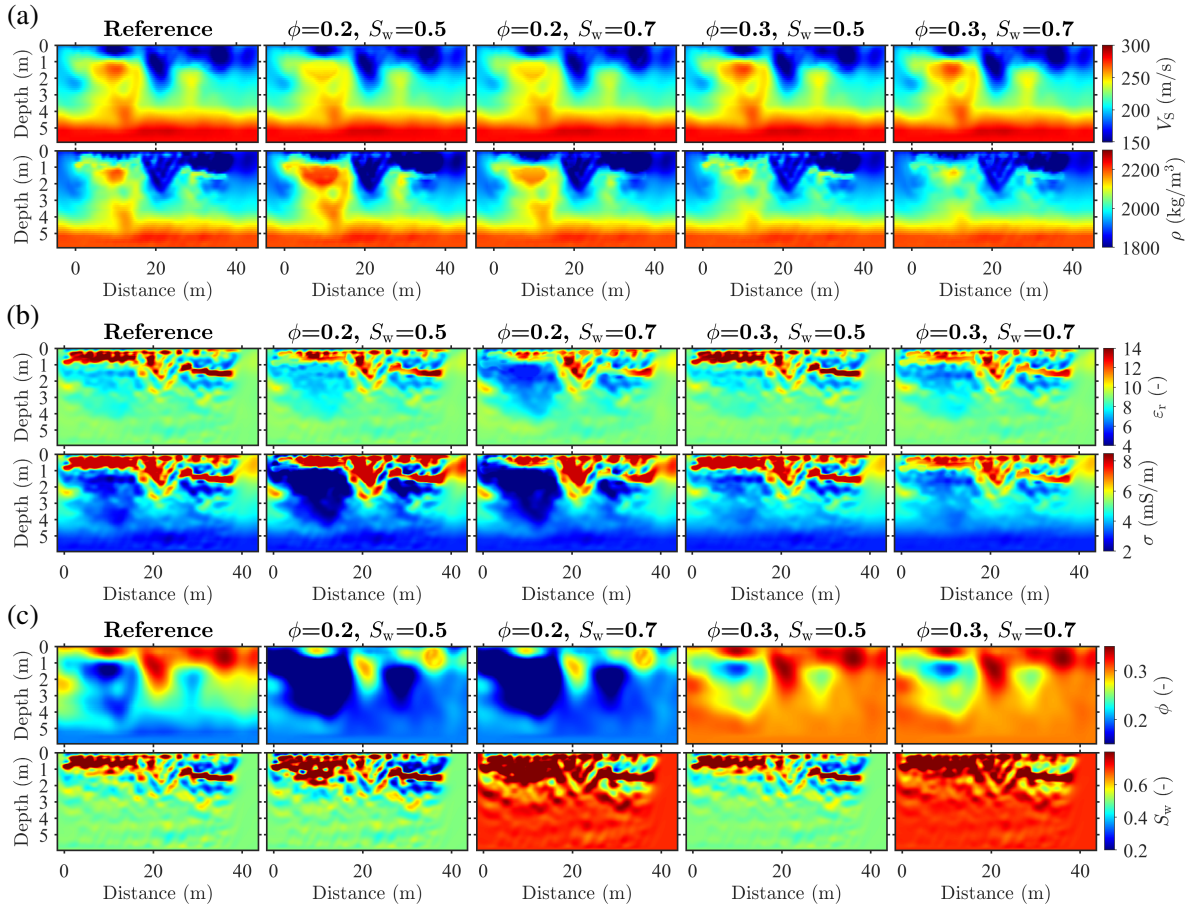

Figure E.2: (a) Seismic models (S-wave velocity V_S and density ρ), (b) ground-penetrating radar (GPR) models (relative permittivity ε_r and conductivity σ), and (c) petrophysical models (porosity ϕ and saturation S_w). The five columns are the indirect JPI results with different petrophysical initial models (see the title of each row). The five JPIs have the same seismic and GPR initial models where the seismic initial models and initial conductivity models are the same as in Figure 5.2, and the initial permittivity model is a homogeneous half-space ($\varepsilon_r = 9$). The referenced petrophysical initial models are the same as in Figure 5.3.

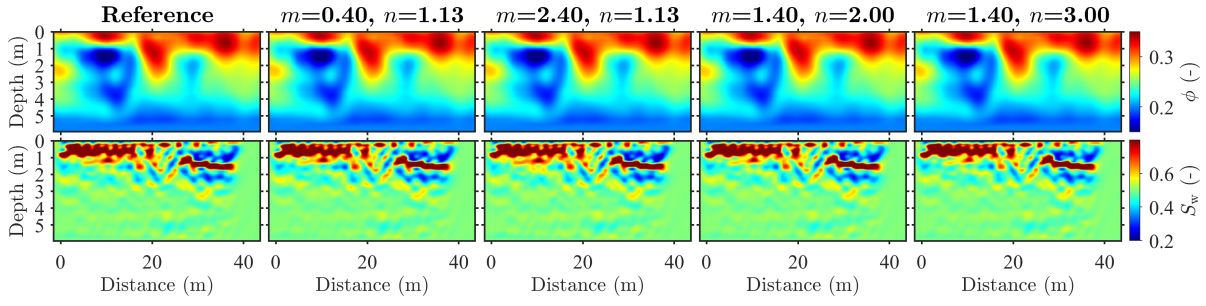

Figure E.3: Petrophysical models (ϕ and S_w). The five columns are the indirect JPI results with different Archie's coefficients (cementation exponent m and saturation exponent n). The initial models of five JPIs are the same as referenced initial models used in Figure E.2. m = 1.4 and n = 1.33 in the reference JPI.

6 Summary

In this dissertation, I presented the results of my research on FWI of GPR data and its joint inversion with shallow-seismic data. As I have drawn conclusions in Section 'Conclusion' and given outlooks in Section 'Discussion' of each chapter (except Chapter 5 where discussion is ignored), here I only summary the main contributions of these studies.

- Chapter 2: I presented SFWI to reduce the computational costs of inverting multi-offset GPR data. As a model-dependent method, SFWI can get a theoretical speedup and memory saving factor equal to the size ratio of the model and its subset. The practical improvements of SFWI can be higher or lower than this factor, depending on the combined parallelizations. Different from the previous study, I quantified the influence of model subset on simulation and inversion and offer rules of thumb and some suggestions for selecting model subset.

- Chapter 3: Using the τ-method introduced in the seismic community, I defined the permittivity attenuation parameter to quantify the attenuation resulting from the complex permittivity and to modify time-domain Maxwell's equations. It showed that permittivity attenuation acts as a low-pass filter, accounting for waveform distortion and amplitude attenuation. The frequency-dependent GPR FWI developed is not only simple to implement, but also ensures a better fit to the observed data and a more realistic reconstruction when the material is strongly frequency-dependent.

- Chapter 4: Based on the sensitivities of GPR and shallow-seismic data to petrophysical parameters, I proposed an indirect JPI. This technique is robust as it links only the most reliably estimated seismic and GPR parameters to porosity and saturation, respectively. As a result, indirect JPI can simultaneously estimate petrophysical, seismic and GPR parameters in high quality and can be applied in a wide range of environment in the presence of uncertainties in *a priori* information.

- Chapter 5: I provided a real world example to validate the feasibility of indirect JPI. It demonstrated that this method has outstanding performance over individual inversions and JSI. On the one hand, indirect JPI reduces the ambiguity of inverting single geophysical data and promotes consistent imaging of the near-surface targets. On the other hand, it takes advantage of the strong constraint from petrophysical relations to robustly reconstruct seismic and GPR parameters with different petrophysical initial models and the coefficients used in petrophysical relations.

In conclusion, I reduced the computational costs of GPR FWI in Chapter 2 and extended GPR FWI to frequency-dependent media in Chapter 3. I then developed an indirect JPI to combine GPR and shallow-seismic data and verify its performance through synthetic examples in Chapter 4 and field examples in

Chapter 5. Indirect JPI can be performed using the acceleration study in Chapter 2. By replacing the permittivity and conductivity with the real effective parameters at the reference frequency, indirect JPI can also be compatible with GPR data obtained in strongly frequency-dependent media (Chapter 3), but this needs further study.

Bibliography

Abubakar, A., Gao, G., Habashy, T. M., and Liu, J. (2012). Joint inversion approaches for geophysical electromagnetic and elastic full-waveform data. *Inverse Problems*, 28(5):055016.

Altdorff, D., Botschek, J., Honds, M., van der Kruk, J., and Vereecken, H. (2019). In situ detection of tree root systems under heterogeneous anthropogenic soil conditions using ground penetrating radar. *Journal of Infrastructure Systems*, 25(3):05019008.

Annan, A. P. (2005). GPR methods for hydrogeological studies. In *Hydrogeophysics*, pages 185–213. Springer.

Archie, G. E. et al. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME*, 146(01):54–62.

Bachmann, E. and Tromp, J. (2020). Source encoding for viscoacoustic ultrasound computed tomography. *The Journal of the Acoustical Society of America*, 147(5):3221–3235.

Belina, F., Irving, J., Ernst, J., and Holliger, K. (2012). Evaluation of the reconstruction limits of a frequency-independent crosshole georadar waveform inversion scheme in the presence of dispersion. *Journal of Applied Geophysics*, 78:9–19.

Benedetto, A., Tosti, F., Ciampoli, L. B., and D'amico, F. (2017). An overview of ground-penetrating radar signal processing techniques for road inspections. *Signal processing*, 132:201–209.

Bergmann, T., Robertsson, J. O., and Holliger, K. (1998). Finite-difference modeling of electromagnetic wave propagation in dispersive and attenuating media. *Geophysics*, 63(3):856–867.

Biot, M. A. (1955). Theory of elasticity and consolidation for a porous anisotropic solid. *Journal of applied physics*, 26(2):182–185.

Birchak, J. R., Gardner, C. G., Hipp, J. E., and Victor, J. M. (1974). High dielectric constant microwave probes for sensing soil moisture. *Proceedings of the IEEE*, 62(1):93–98.

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.

Bleistein, N. (1986). Two-and-one-half dimensional in-plane wave propagation. *Geophysical Prospecting*, 34(5):686–703.

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

Boniger, U. and Tronicke, J. (2010). On the potential of kinematic GPR surveying using a self-tracking total station: Evaluating system crosstalk and latency. *IEEE Transactions on Geoscience and Remote Sensing*, 48(10):3792–3798.

Bradford, J. H. (2007). Frequency-dependent attenuation analysis of ground-penetrating radar data. *Geophysics*, 72(3):J7–J16.

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

Busch, S., van der Kruk, J., Bikowski, J., and Vereecken, H. (2012). Quantitative conductivity and permittivity estimation using full-waveform inversion of on-ground GPR data. *Geophysics*, 77(6):H79–H91.

Carcione, J. and Cavallini, F. (1995). On the acoustic-electromagnetic analogy. *Wave motion*, 21(2):149–162.

Carcione, J. M. (1996). Ground-penetrating radar: Wave theory and numerical simulation in lossy anisotropic media. *Geophysics*, 61(6):1664–1677.

Choi, Y. and Alkhalifah, T. (2012). Application of multi-source waveform inversion to marine streamer data using the global correlation norm. *Geophysical Prospecting*, 60(4-Simultaneous Source Methods for Seismic Data):748–758.

Deng, J., Rogez, Y., Zhu, P., Herique, A., Jiang, J., and Kofman, W. (2021). 3D time-domain electromagnetic full waveform inversion in Debye dispersive medium accelerated by multi-GPU paralleling. *Computer Physics Communications*, 265:108002.

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

Domenzain, D., Bradford, J., and Mead, J. (2020). Joint inversion of full-waveform ground-penetrating radar and electrical resistivity data: Part 1. *Geophysics*, 85(6):H97–H113.

Domenzain, D., Bradford, J., and Mead, J. (2021). Joint full-waveform GPR and ER inversion applied to field data acquired on the surface. *Geophysics*, 87(1):1–77.

El Bouajaji, M., Lanteri, S., and Yedlin, M. (2011). Discontinuous Galerkin frequency domain forward modelling for the inversion of electric permittivity in the 2D case. *Geophysical Prospecting*, 59(Modelling Methods for Geophysical Imaging: Trends and Perspectives):920–933.

Ernst, J. R., Green, A. G., Maurer, H., and Holliger, K. (2007a). Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data. *Geophysics*, 72(5):J53–J64.

Ernst, J. R., Maurer, H., Green, A. G., and Holliger, K. (2007b). Full-waveform inversion of crosshole radar data based on 2-D finite-difference time-domain solutions of Maxwell's equations. *IEEE transactions on Geoscience and Remote Sensing*, 45(9):2807–2828.

Everett, M. E. (2013). *Near-surface applied geophysics*. Cambridge University Press.

Fabien-Ouellet, G., Gloaguen, E., and Giroux, B. (2017). Time domain viscoelastic full waveform inversion. *Geophysical Journal International*, 209(3):1718–1734.

Feng, X., Ren, Q., and Liu, C. (2017). Quantitative imaging for civil engineering by joint full waveform inversion of surface-based GPR and shallow seismic reflection data. *Construction and Building Materials*, 154:1173–1182.

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.

Gallardo, L. A. and Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2D inversion of dc resistivity and seismic data. *Geophysical Research Letters*, 30(13).

Gao, G., Abubakar, A., and Habashy, T. M. (2012). Joint petrophysical inversion of electromagnetic and full-waveform seismic data. *Geophysics*, 77(3):WA3–WA18.

Gao, L., Pan, Y., and Bohlen, T. (2020). 2-D multiparameter viscoelastic shallow-seismic full-waveform inversion: reconstruction tests and first field-data application. *Geophysical Journal International*, 222(1):560–571.

Gao, L., Pan, Y., Rieder, A., and Bohlen, T. (2021). Multiparameter viscoelastic full-waveform inversion of shallow-seismic surface waves with a pre-conditioned truncated Newton method. *Geophysical Journal International*, 227(3):2044–2057.

Garambois, S., Sénéchal, P., and Perroud, H. (2002). On the use of combined geophysical methods to assess water content and water conductivity of near-surface formations. *Journal of Hydrology*, 259(1-4):32–48.

Gassmann, F. (1951). Elasticity of porous media. *Vierteljahrsschrder Naturforschenden Gesselschaft*, 96:1–23.

Groos, L., Schäfer, M., Forbriger, T., and Bohlen, T. (2014). The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves. *Geophysics*, 79(6):R247–R261.

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

Gueting, N., Klotzsche, A., van der Kruk, J., Vanderborght, J., Vereecken, H., and Englert, A. (2015). Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests. *Journal of hydrology*, 524:680–695.

Gueting, N., Klotzsche, A., van der Kruk, J., Vanderborght, J., Vereecken, H., and Englert, A. (2020). Corrigendum to "imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests"[j. hydrol.(2015) 680-695]. *Journal of Hydrology*, 590:125483.

Gueting, N., Vienken, T., Klotzsche, A., Van Der Kruk, J., Vanderborght, J., Caers, J., Vereecken, H., and Englert, A. (2017). High resolution aquifer characterization using crosshole GPR full-waveform tomography: Comparison with direct-push and tracer test data. *Water Resources Research*, 53(1):49–72.

Heincke, B., Jegen, M., Moorkamp, M., Hobbs, R. W., and Chen, J. (2017). An adaptive coupling strategy for joint inversions that use petrophysical information as constraints. *Journal of Applied Geophysics*, 136:279–297.

Hoang, D. H. (2021). Random-shot full-waveform inversion of shallow seismic surface waves. Master's thesis, Karlsruhe Institute of Technology.

Huang, Y. and Schuster, G. T. (2012). Multisource least-squares migration of marine streamer and land data with frequency-division encoding. *Geophysical Prospecting*, 60(4-Simultaneous Source Methods for Seismic Data):663–680.

Irnaka, T., Brossier, R., Métivier, L., Bohlen, T., and Pan, Y. (2022). 3-D multicomponent full waveform inversion for shallow-seismic target: Ettlingen Line case study. *Geophysical Journal International*, 229(2):1017–1040.

Irving, J. and Knight, R. (2006). Numerical modeling of ground-penetrating radar in 2-D using MATLAB. *Computers & Geosciences*, 32(9):1247–1258.

Jiang, H. (2019). *Seismic imaging: strategies for visco-acoustic full waveform inversion*. PhD thesis, Université Paris sciences et lettres.

Jol, H. M. (2008). *Ground penetrating radar theory and applications*. elsevier.

Klotzsche, A., Vereecken, H., and van der Kruk, J. (2019). Review of crosshole ground-penetrating radar full-waveform inversion of experimental data: Recent developments, challenges, and pitfalls. *Geophysics*, 84(6):H13–H28.

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.

Koyan, P. and Tronicke, J. (2020). 3D modeling of ground-penetrating radar data across a realistic sedimentary model. *Computers & Geosciences*, 137:104422.

Krebs, J. R., Anderson, J. E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A., and Lacasse, M.-D. (2009). Fast full-wavefield seismic inversion using encoded sources. *Geophysics*, 74(6):WCC177–WCC188.

Kuroda, S., Takeuchi, M., and Kim, H. J. (2007). Full-waveform inversion algorithm for interpreting crosshole radar data: A theoretical approach. *Geosciences Journal*, 11(3):211–217.

Kurzmann, A., Przebindowska, A., Köhn, D., and Bohlen, T. (2013). Acoustic full waveform tomography in the presence of attenuation: A sensitivity analysis. *Geophysical Journal International*, 195(2):985–1000.

Lavoué, F. (2014). *2D full waveform inversion of ground penetrating radar data: towards multiparameter imaging from surface data*. PhD thesis, Université de Grenoble.

Lavoué, F., Brossier, R., Métivier, L., Garambois, S., and Virieux, J. (2014). Two-dimensional permittivity and conductivity imaging by full waveform inversion of multioffset GPR data: A frequency-domain quasi-Newton approach. *Geophysical Journal International*, 197(1):248–268.

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

Linde, N., Tryggvason, A., Peterson, J. E., and Hubbard, S. S. (2008). Joint inversion of crosshole radar and seismic traveltimes acquired at the South Oyster Bacterial Transport Site. *Geophysics*, 73(4):G29–G37.

Liu, T., Klotzsche, A., Pondkule, M., Vereecken, H., Su, Y., and van der Kruk, J. (2018). Radius estimation of subsurface cylindrical objects from ground-penetrating-radar data using full-waveform inversion. *Geophysics*, 83(6):H43–H54.

Meles, G., Greenhalgh, S., Van der Kruk, J., Green, A., and Maurer, H. (2012). Taming the non-linearity problem in GPR full-waveform inversion for high contrast media. *Journal of Applied Geophysics*, 78:31–43.

Meles, G. A., Greenhalgh, S. A., Green, A. G., Maurer, H., and Van der Kruk, J. (2011). GPR full-waveform sensitivity and resolution analysis using an FDTD adjoint method. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5):1881–1896.

Meles, G. A., Van der Kruk, J., Greenhalgh, S. A., Ernst, J. R., Maurer, H., and Green, A. G. (2010). A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9):3391–3407.

Métivier, L., Brossier, R., Virieux, J., and Operto, S. (2013). Full waveform inversion and the truncated Newton method. *SIAM Journal on Scientific Computing*, 35(2):B401–B437.

Moorkamp, M., Roberts, A., Jegen, M., Heincke, B., and Hobbs, R. (2013). Verification of velocity-resistivity relationships derived from structural joint inversion with borehole data. *Geophysical Research Letters*, 40(14):3596–3601.

Nihei, K. T. and Li, X. (2007). Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD—PSD). *Geophysical Journal International*, 169(3):1069–1078.

Nur, A. (1992). Critical porosity and the seismic velocities in rocks. *EOS Transactions AGU*, 73:66.

Oberröhrmann, M., Klotzsche, A., Vereecken, H., and van der Krak, J. (2013). Optimization of acquisition setup for cross-hole: GPR full-waveform inversion using checkerboard analysis. *Near Surface Geophysics*, 11(2):197–209.

Operto, S., Gholami, Y., Prieux, V., Ribodetti, A., Brossier, R., Metivier, L., and Virieux, J. (2013). A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice. *The leading edge*, 32(9):1040–1054.

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

Pan, Y., Gao, L., and Bohlen, T. (2021). Random-objective waveform inversion of 3D-9C shallow-seismic field data. *Journal of Geophysical Research: Solid Earth*, 126(9):e2021JB022036.

Pan, Y., Gao, L., and Shigapov, R. (2020). Multi-objective waveform inversion of shallow seismic wavefields. *Geophysical Journal International*, 220(3):1619–1631.

Pan, Y., Schaneng, S., Steinweg, T., and Bohlen, T. (2018). Estimating S-wave velocities from 3D 9-component shallow seismic data using local Rayleigh-wave dispersion curves–A field study. *Journal of Applied Geophysics*, 159:532–539.

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.

Papadopoulos, T. G. and Rekanos, I. T. (2011). Time-domain microwave imaging of inhomogeneous Debye dispersive scatterers. *IEEE Transactions on Antennas and Propagation*, 60(2):1197–1202.

Paz, C., Alcalá, F. J., Carvalho, J. M., and Ribeiro, L. (2017). Current uses of ground penetrating radar in groundwater-dependent ecosystems research. *Science of the Total Environment*, 595:868–885.

Pica, A., Diet, J., and Tarantola, A. (1990). Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, 55(3):284–292.

Pinard, H., Garambois, S., Métivier, L., Dietrich, M., Sénéchal, G., and Rousset, D. (2016). Full-waveform inversion of GPR data acquired between boreholes in Rustrel carbonates. In *E3S Web of Conferences*, volume 12, page 01002. EDP Sciences.

Pinard, H., Garambois, S., Métivier, L., Dietrich, M., and Virieux, J. (2015). 2D frequency-domain full-waveform inversion of GPR data: Permittivity and conductivity imaging. In *2015 8th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)*, pages 1–4. IEEE.

Pipkin, A. C. (2012). *Lectures on viscoelasticity theory*, volume 7. Springer Science & Business Media.

Plessix, R.-E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.

Plessix, R.-E. and Mulder, W. (2004). Frequency-domain finite-difference amplitude-preserving migration. *Geophysical Journal International*, 157(3):975–987.

Plessix, R.-E. and Mulder, W. (2008). Resistivity imaging with controlled-source electromagnetic data: depth and data weighting. *Inverse problems*, 24(3):034012.

Polak, E. and Ribiere, G. (1969). Note on convergence of conjugate direction methods. *Revue Francaise D Informatique De Recherche Operationnelle*, 3(16):35–43.

Qin, T., Bohlen, T., and Allroggen, N. (2022a). Full-waveform inversion of ground-penetrating radar data in frequency-dependent media involving permittivity attenuation. *Geophysical Journal International*, 232(1):504–522.

Qin, T., Bohlen, T., and Pan, Y. (2022b). Indirect joint petrophysical inversion of synthetic shallow-seismic and multi-offset ground-penetrating radar data. *Geophysical Journal International*, 229(3):1770–1784.

Roden, J. A. and Gedney, S. D. (2000). Convolution PML (CPML): An efficient FDTD implementation of the CFS–PML for arbitrary media. *Microwave and optical technology letters*, 27(5):334–339.

Romdhane, A., Grandjean, G., Brossier, R., Réjiba, F., Operto, S., and Virieux, J. (2011). Shallow-structure characterization by 2D elastic full-waveform inversion. *Geophysics*, 76(3):R81–R93.

Rumpf, M. and Tronicke, J. (2014). Predicting 2D geotechnical parameter fields in near-surface sedimentary environments. *Journal of Applied Geophysics*, 101:95–107.

Schaneng, S. P. (2017). Erstellung eines 3D modells der scherwellengeschwindigkeit im bereich der Ettlinger Linie (Rheinstetten) aus der 1D inversion der lokalen dispersion von Rayleigh-wellen [1D inversion of local Rayleigh wave dispersion curves to construct a local 3D S-wave velocity model of the Ettlinger Linie in Rheinstetten]. Master's thesis, Karlsruhe Institute of Technology.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.

Tillmann, A., Englert, A., Nyari, Z., Fejes, I., Vanderborght, J., and Vereecken, H. (2008). Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the Krauthausen test site using cone penetration test. *Journal of contaminant hydrology*, 95(1-2):57–75.

Tran, K. T. and McVay, M. (2012). Site characterization using Gauss–Newton inversion of 2-D full seismic waveform in the time domain. *Soil Dynamics and Earthquake Engineering*, 43:16–24.

Tran, K. T. and Sperry, J. (2018). Application of 2D full-waveform tomography on land-streamer data for assessment of roadway subsidence. *Geophysics*, 83(3):EN1–EN11.

Turner, G. and Siggins, A. F. (1994). Constant Q attenuation of subsurface radar pulses. *Geophysics*, 59(8):1192–1200.

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.

Wagner, F., Mollaret, C., Günther, T., Kemna, A., and Hauck, C. (2019). Quantitative imaging of water, ice and air in permafrost systems through petrophysical joint inversion of seismic refraction and electrical resistivity data. *Geophysical Journal International*, 219(3):1866–1875.

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612.

Wegscheider, S. (2017). Abbildung der Ettlinger Linie auf dem segelflugplatz Rheinstetten mittels georadar [Illustration of the Ettlinger line on the gliding airfield Rheinstetten by means of GPR]. Master's thesis, Karlsruhe Institute of Technology.

Winters, D. W., Bond, E. J., Van Veen, B. D., and Hagness, S. C. (2006). Estimation of the frequency-dependent average dielectric properties of breast tissue using a time-domain inverse scattering technique. *IEEE Transactions on Antennas and Propagation*, 54(11):3517–3528.

Wittkamp, F., Athanasopoulos, N., and Bohlen, T. (2019). Individual and joint 2-D elastic full-waveform inversion of Rayleigh and Love waves. *Geophysical Journal International*, 216(1):350–364.

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.

Xu, S., Chen, F., Lambaré, G., Zhang, Y., and Wang, D. (2012). Inversion on reflected seismic wave. In *2012 SEG Annual Meeting*. OnePetro.

Yang, P., Brossier, R., Métivier, L., and Virieux, J. (2016). A review on the systematic formulation of 3-D multiparameter full waveform inversion in viscoelastic medium. *Geophysical Journal International*, 207(1):129–149.

Yao, G., Da Silva, N. V., and Wu, D. (2018). Sensitivity analyses of acoustic impedance inversion with full-waveform inversion. *Journal of Geophysics and Engineering*, 15(2):461–477.

Yao, G., Wu, D., and Wang, S.-X. (2020). A review on reflection-waveform inversion. *Petroleum Science*, 17(2):334–351.

Yee, K. (1966). Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307.

Yu, Y., Klotzsche, A., Weihermüller, L., Huisman, J. A., Vanderborght, J., Vereecken, H., and van der Kruk, J. (2020). Measuring vertical soil water content profiles by combining horizontal borehole and dispersive surface ground penetrating radar data. *Near Surface Geophysics*, 18(3):275–294.

Zhang, Q., Mao, W., Zhou, H., Zhang, H., and Chen, Y. (2018). Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise. *Geophysical Journal International*, 215(3):1659–1681.

Zhou, Z., Klotzsche, A., Schmäck, J., Vereecken, H., and van der Kruk, J. (2021a). Improvement of ground-penetrating radar full-waveform inversion images using cone penetration test data. *Geophysics*, 86(3):H13–H25.

Zhou, Z., Klotzsche, A., and Vereecken, H. (2021b). Improving crosshole ground-penetrating radar full-waveform inversion results by using progressively expanded bandwidths of the data. *Near Surface Geophysics*, 19(4):465–487.

List of Figures

List of Tables

Acknowledgment

There are many people that I would like to appreciate during my doctoral studies. Without their help and company, I could not have completed this dissertation.

First of all, I would like to thank my supervisor, Mr. Thomas Bohlen, for his guidance and support of my research. He showed me how to be an independent researcher in solving scientific problems and presenting research results effectively. I enjoyed his patience and benefited from his profound knowledge and rigorous attitude towards science.

Then I would like to thank my colleague Yudi Pan, whose innovative ideas have inspired me a lot, my colleague Niklas Allroggen from Potsdam for fruitful discussions on GPR data acquisition and inversion, and my colleague Anja Klotzsche from Jülich for the help in GPR FWI.

I am also grateful to Thomas Hertweck for his help with scientific writing and much technical support, and to Tilman Steinweg and Mark Wienöbst for their help in developing the WAVE toolbox and programming with Linux systems. Special thanks to Claudia Payne for her continuous help with many administrative tasks that made my stay in Germany easier and more memorable.

I thank Lars Houpt, Felix Bögelspacher, Leon Merkel, Michael Mayer, Hagen Steger, Anna Albers, Roland Helfer, Philipp Koyan, Amelie Cathrine Nüsse, Tuan Dung Vu, Rafael Acosta and other master students at Geophysical Institute, KIT, for their help in field data acquisition. I would like to thank Jan Igel and Stephan Schennen for the nice laboratory measurements provided during my visit to Hanover.

I thank Shufan Hu, Xiagang Long, Linan Xu and Nikolaos Athanasopoulos for their help during the conference in The Hague, The Netherlands. I thank my master supervisor Yonghui Zhao for his constant support, and my friends Danhua Xin, Sichuan Yuan, Kai Xiong for their help in living and studying abroad.

Thanks to all current and former members of the GPI group for working together over the last four years, especially my colleagues: Laura Gaßner, Renat Shigapov, Mario Fernandez, Lingli Gao, Tingting Liu, Sarah Beraus, Duy Hoang Hoang, Priscilla Ntengue Chatue, Jonas Müller, Daniel Krieger, Marina Silva Borja, Sonia Adelina Sortan, Abdullah Jarah, Mehrdad Soleimani, Carina Becker, Maria Urquizar Carmona, Niloofar Alaei, Clara Fraile and Bessam Alubeyid.

Finally, I would like to sincerely thank my family for their love. I thank my wife for her company during the past five years. She has made me not alone and has filled my life with happiness. I thank my mother and father for giving me life and raising me, my grandmother for her care, my brother and sister-in-law for taking care of this family, and my mother-in-law and father-in-law for their support.