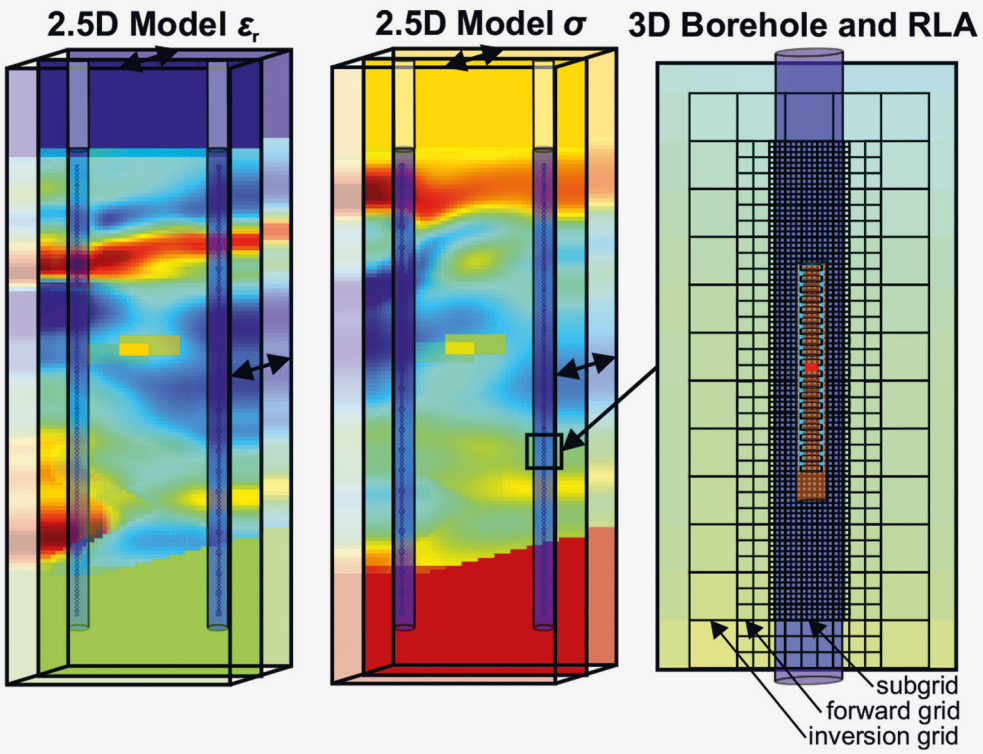

Multi-dimensional GPR full-waveform inversion for small-scale hydrogeophysical soil characterization

Dominik Hoven

Energie & Umwelt / Energy & Environment
Band / Volume 643
ISBN 978-3-95806-781-3

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 643

Forschungszentrum Jülich GmbH
Institut für Bio- und Geowissenschaften (IBG)
Agrosphäre (IBG-3)

Multi-dimensional GPR full-waveform inversion for small-scale hydrogeophysical soil characterization

Dominik Hoven

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 643

ISSN 1866-1793 ISBN 978-3-95806-781-3

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb
Umschlaggestaltung:	Grafische Medien, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH
Copyright:	Forschungszentrum Jülich 2024

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 643

D 82 (Diss. RWTH Aachen University, 2024)

ISSN 1866-1793
ISBN 978-3-95806-781-3

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.

Abstract

A detailed understanding of the processes within the critical zone, which covers the area from the earth's surface down to the aquifer, is essential for sustainable resource management and environmental protection. This zone exhibits complex flow and transport processes and supports critical ecosystem services such as water supply, agriculture, and climate regulation. However, imaging the complex critical zone accurately, especially at high resolutions required for a detailed analysis, presents significant challenges because of the variability of soil water content and complex subsurface structures. This thesis introduces a novel 2.5D ground penetrating radar (GPR) full-waveform inversion (FWI) method that enhances subsurface imaging by accurately incorporating 3D geometries, such as air and water filled boreholes, finite length antenna models, and lysimeter geometries, in the forward modeling of the GPR FWI. Furthermore, the 3D-to-2D data transformation with its assumptions, e.g. for the far-field, necessary for 2D GPR FWI, is not required with this method.

We show in synthetic studies with different inversion methods (2D FWI, 2.5D FWI, 2.5D FWI with borehole, and 2.5D FWI with borehole and antenna) an improved source wavelet reconstruction with the inclusion of realistic borehole and antenna geometries for the data. The inclusion of these geometries in the forward model of FWI approaches can significantly improve the accuracy of conductivity reconstructions, with a reduction in the mean relative absolute error of conductivity of more than 20% compared to simple 2D FWI and 2.5D FWI. The improvement is particularly noticeable in high-contrast zones. Although including antenna geometries significantly increases computational requirements by a factor of $\sim$10, the quality of reconstruction remains similar to the case with only borehole inclusion. In contrast to ray-based inversion (RBI), where artifacts arise when using high-angle data (72.35°), FWI still provides reliable results.

In a following analysis, we tested if a model that includes boreholes and finite length antenna models for experimental data measured with transmitter and receiver positioned in air and water filled boreholes can improve the effective source wavelet estimation. A synthetic test shows that using this approach, only one wavelet can be used for the reconstruction of both the unsaturated and saturated zone. However, we still observed challenges

with the current antenna model to account for the different coupling in air filled boreholes for measured data. Using the new 2.5D FWI with borehole and antenna models and a single source wavelet, the results of the saturated zone reconstruction were similar to those observed in previous studies where four effective source wavelets were considered. To obtain reliable results in the unsaturated zone, it is necessary to adapt the antenna model to resolve existing discrepancies.

Next to an improved reconstruction of small-scale structures in aquifers, small-scale processes in the soil-plant-atmosphere continuum are also of interest. In order to achieve a higher reconstruction resolution with the FWI for these processes, higher frequencies are necessary. In a first part, we indicate the constraints imposed by high-frequency GPR data, which require more precise starting models to fulfill the half-wavelength criterion of the GPR FWI. This cannot be met by the regular starting model approach of using RBI models. We show that a frequency-hopping approach can be used to generate starting models that meet these requirements. Furthermore, we investigated the influence of first-arrival and amplitude changes in the source wavelets on high-frequency GPR FWI. Utilizing an adapted heterogeneous model, we were able to show a more detailed reconstruction with higher frequency data compared to lower frequency data.

In a next step, we extended the model building process of the 2.5D GPR FWI and are now able to include more complex geometrical structures like lysimeters in the forward model. As we faced challenges to use the 2D GPR FWI on experimental high-frequency data acquired on lysimeters, we first investigated the different GPR waves in synthetic studies at lysimeters filled with homogeneous and heterogeneous soils. We show the complexity of the GPR data, that includes air, direct, and reflected waves. We created a synthetic 3D GPR lysimeter dataset with a center frequency of 450 MHz and applied the novel 2.5D GPR FWI to this dataset. It demonstrates an exceptional good reconstruction of the soil and fit of the dataset by the inversion results, effectively simulating air-, real soil-, and reflected waves as well as revealing intricate soil properties.

The newly developed 2.5D GPR FWI presented in this thesis enables the modeling and reconstruction of small-scale structures with high resolution. The application ranges from aquifer characterization to the now possible inversion of GPR data measured at lysimeter, providing a foundational framework for future research in high resolution subsurface imaging.

Zusammenfassung

Ein detailliertes Verständnis der Prozesse in der kritischen Zone, die den Bereich von der Erdoberfläche bis hinunter zum Grundwasserleiter umfasst, ist für eine nachhaltige Ressourcenverwaltung und den Umweltschutz unerlässlich. Diese Zone weist komplexe Strömungs- und Transportprozesse auf und unterstützt wichtige Ökosystemleistungen, wie Wasserversorgung, Landwirtschaft und Klimaregulierung. Eine hohe Auflösung der kritischen Zone für eine detaillierte Analyse stellt jedoch jedoch aufgrund der Variabilität des Bodenwassergehalts und der komplexen Strukturen im Untergrund eine große Herausforderung dar. In dieser Arbeit wird eine neue 2.5D Bodenradar (GPR) Vollewellenforminversion (FWI) vorgestellt, die die Modellierung des Untergrundes verbessert, indem 3D-Geometrien, wie luft- und wassergefüllte Bohrlöcher, Antennenmodelle und Lysimetergeometrien, in die Vorwärtsmodellierung der GPR FWI einbezogen werden. Darüber hinaus entfällt bei dieser Methode die 3D-zu-2D Datentransformation, welche für die 2D GPR FWI notwendig ist (und ihre Annahmen, z.B. für das Fernfeld).

Wir zeigen in synthetischen Studien mit verschiedenen Inversionsmethoden (2D FWI, 2.5D FWI, 2.5D FWI mit Bohrloch und 2.5D FWI mit Bohrloch und Antenne) eine verbesserte Rekonstruktion des Quellsignals durch die Einbeziehung von realistischen Bohrloch- und Antennengeometrien. Dadurch kann die Rekonstruktion der Leitfähigkeit erheblich verbessert werden, was sich in einer Verringerung des mittleren relativen absoluten Fehlers der Leitfähigkeit um mehr als 20% im Vergleich zur einfachen 2D FWI und 2.5D FWI zeigt. Die Verbesserung ist besonders in kontrastreichen Zonen erkennbar. Obwohl die Einbeziehung von Antennengeometrien den Rechenaufwand um einen Faktor $\sim$10 erhöht, bleibt die Qualität der Rekonstruktion ähnlich wie bei der bloßen Einbeziehung von Bohrlöchern. Im Gegensatz zur strahlenbasierten Inversion (RBI), bei der es bei der Verwendung von Daten mit hohem Winkel (72.35°) zu Artefakten kommt, liefert die FWI weiterhin zuverlässige Ergebnisse.

In einer folgenden Analyse haben wir getestet, ob ein Modell, das Bohrlöcher und Antennenmodelle für experimentelle Daten mit eischließt, wobei sich Sender und Empfänger in luft- und wassergefüllten Bohrlöchern befanden, die effektive Schätzung des Quellsignals verbessern kann. Ein synthetischer Test zeigt, dass mit diesem Ansatz nur ein Quellsignal für die Rekonstruktion sowohl der ungesättigten als auch der gesättigten Zone verwendet

werden muss, anstatt wie zuvor vier. Wir haben jedoch festgestellt, dass das aktuelle Antennenmodell die unterschiedliche Kopplung in luftgefüllten Bohrlöchern für gemessene Daten noch nicht berücksichtigen kann. Bei Verwendung der neuen 2.5D FWI mit Bohrloch- und Antennenmodellen und nur einem Quellsignal waren die Ergebnisse der Rekonstruktion der gesättigten Zone ähnlich wie in früheren Studien, die vier effektive Quellsignale berücksichtigten. Um verlässliche Ergebnisse in der ungesättigten Zone zu erhalten, ist es notwendig, das Antennenmodell anzupassen, um bestehende Unterschiede zu beseitigen.

Neben einer verbesserten Modellierung kleinskaliger Strukturen in Wasserleitern sind auch kleinskalige Prozesse im Boden-Pflanze-Atmosphäre-Kontinuum von Interesse. Eine höhere Auflösung erfodert die Nutzung höherer Frequenzen. Zuerst zeigen wir die Einschränkungen auf, die durch hochfrequente GPR-Daten entstehen und die präzisere Startmodelle erfordern, um das Halbwellenlängenkriterium der GPR FWI zu erfüllen. Dieses kann mit dem üblichen Startmodell-Ansatz, der Verwendung von RBI-Modellen, nicht erreicht werden. Wir zeigen, dass ein Frequenzsprungverfahren verwendet werden kann, um Startmodelle zu erzeugen, die diese Anforderungen erfüllen. Darüber hinaus untersuchen wir den Einfluss von Erstankunfts- und Amplitudenänderungen in den Quellsignalen auf die hochfrequente GPR FWI. Unter Verwendung eines angepassten heterogenen Modells können wir eine detailliertere Rekonstruktion mit hochfrequenten Daten im Vergleich zu niederfrequenten Daten zeigen.

Im nächsten Schritt haben wir den Modellaufbauprozess der 2.5D GPR FWI erweitert und sind nun in der Lage, komplexere geometrische Strukturen wie Lysimeter in das Vorwärtsmodell einzubeziehen. Da die Anwendung der 2D GPR FWI auf hochfrequente Daten, die an einem Lysimeter gemessen wurden, eine Herausforderung darstellt, analysieren wir zunächst die verschiedenen GPR-Wellen in synthetischen Studien an Lysimetern, die mit homogenen und heterogenen Böden gefüllt waren. Wir zeigen die Komplexität der GPR-Daten, die Luft, direkte und reflektierte Wellen umfassen. Wir haben einen synthetischen 3D-GPR-Lysimeter-Datensatz mit einer Mittenfrequenz von 450 MHz erstellt und die neue 2.5D GPR FWI auf diesen Datensatz angewendet. Die Ergebnisse zeigen eine außergewöhnlich gute Rekonstruktion des Bodens und eine gute Anpassung an den Datensatz durch die Inversionsergebnisse, wobei Luft-, reale Boden- und reflektierte Wellen effektiv simuliert werden und komplizierte Bodeneigenschaften sichtbar werden.

Die in dieser Arbeit vorgestellte neu entwickelte 2.5D GPR FWI ermöglicht die hochauflösende Modellierung und Rekonstruktion kleinskaliger Strukturen. Die Anwendung reicht von der Charakterisierung von Grundwasserleitern bis hin zur Inversion von GPR-Daten, die an Lysimetern gemessen wurden. Daher bietet sie eine Verbesserung für Problemstellungen im Bereich der hochauflösenden Untergrunddarstellung.

Contents

1. Introduction

Imaging the critical zone, which compromises both the saturated and unsaturated zones of variably saturated soil-aquifer systems (see Figure 1.1), is a challenging task which is essential to improve our understanding of the complex dynamics of flow and transport processes from the surface to the aquitard. The soil functions of the critical zone are vital for sustaining human life, for example through their role in buffering and filtering water. Therefore, the critical zone is extensively exploited for clean water supply, agriculture, and waste storage (Banwart et al., 2013; Banwart et al., 2019; Field et al., 2015). With this, mankind has undeniably altered the soil-plant-atmosphere continuum, as soils are essential for regulating climate and biogeochemical cycles (Lal et al., 2021). To understand these changes and the original natural processes in detail, a high-resolution characterization is required. The need for sustainable agricultural practices and the assessment and mitigation of environmental risks associated with human activities drive this interest, more than just for academic purposes (Binley et al., 2015; Guo and Lin, 2016; Thomsen et al., 2004). Small-scale structures are important to map and characterize because they can have a significant impact on zones of preferential flow, impermeable clay lenses or root water uptake (Vereecken et al., 2016). Furthermore, they have a significant impact on the processes that support clean water and nutrient availability.

In the last two decades, we have seen an tremendous increase in developments and applications of non-destructive geophysical methods to map and characterize the critical zone (Binley et al., 2015; Romero-Ruiz et al., 2018), such as electrical resistivity tomography (ERT) (e.g., Garré et al., 2011; Beff et al., 2013), electromagnetic induction (EMI) (e.g., Blanchy et al., 2020) and ground-penetrating radar (GPR) (e.g., Klotzsche et al., 2018). Although EMI and ERT can either map on a fast large scale or monitor parameters on fixed installed profile lines, these methods struggle to detect and monitor small-scale processes. Furthermore, both are in contrast to GPR not directly linked to the SWC. Especially GPR has shown in the last decade its high potential to identify small-scale structures. Due to the use of high-frequency electromagnetic waves, these non- or minimally invasive high

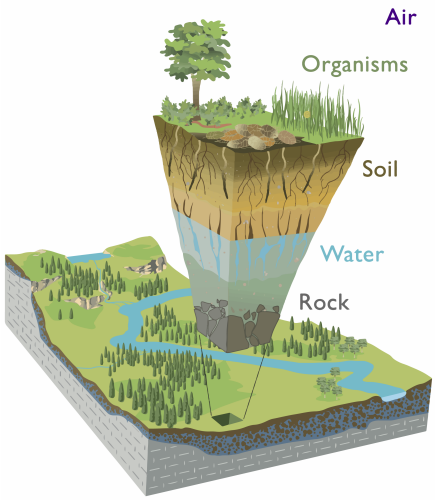

Figure 1.1.: Schematic view on the critical zone, which compromises the upper meter above the surface down to the bedrock (adapted from in Chorover et al. (2007); modified from artwork by R. Kindlimann; licensed under the Creative Commons AttributionShare Alike 4.0 International license).

resolution methods can then be used to constrain hydrological models, or help to derive predictions for plant growth for agricultural and the environmental purposes (Binley et al., 2015) based on petrophysical relationships that are controlled by small-scale subsurface structures. Particularly in the context of climate change, these technologies have facilitated the study of processes within various soil-plant-atmosphere continuums (Lärm et al., 2023). Since the critical zone is highly complex and processes can take place at different scales, different environments are needed depending of the target. To investigate flow and transport processes at the aquifer scale, controlled test sites are essential. Research sites, which are highly scientifically analyzed, like the Boise Hydrogeophysical Research Site (Barrash and Routh, 2006) or Krauthausen test site (Vanderborght et al., 2012) are particularly suitable to extend our knowledge of the vadose zone, which extends from the surface to the water table, and aquifers. While smaller processes such as the soil root interaction require controlled laboratory or plot-scale studies (e.g., Cai et al., 2016; Klotzsche et al., 2019a; Lärm et al., 2023; Lärm et al., 2024). Controlled laboratory studies, which can be scaled up to the field level through tools like lysimeters – soil filled cylinders with a diameter of around 1.0 m to 1.2 m (e.g., Garré et al., 2011) – offer a controlled environment to measure soil water content (SWC) and nutrient fluxes, providing a detailed view of soil properties

and processes. In order to image the properties and processes inside lysimeters, traditional geophysical methods such as ERT and time-domain reflectometry (TDR) are often considered to map soil structures and flow patterns (Garré et al., 2010; Garré et al., 2011). However, ERT has certain limitations, such as the smearing out of small-scale structures, caused by the inversion approaches, and the alternation of the signal depending on temperature, clay content, salinity, and SWC. One possible solution to overcome this issue is to combined ERT and GPR (Looms et al., 2008), or use complementary information such as soil sensors or coring.

In contrast to EMI and ERT, impulse GPR has the advantage of providing the highest possible resolution of geophysical field methods of the subsurface because of the use of high-frequency electromagnetic pulses between 25 MHz and 2.6 GHz (e.g., Klotzsche et al., 2018). GPR uses electromagnetic wave propagation and can provide information on the relative dielectric permittivity ϵ_r and electrical conductivity σ of the subsurface. These parameters can be linked to important soil properties such as porosity and SWC for permittivity, and soil texture and salinity for conductivity. Generally, GPR is most successfully applied in low-conductivity materials, where energy loss factors are much smaller than energy storage (Jol, 2009). Because of the large contrast in relative dielectric permittivity between air and water (see Table 1.1), the method is well suited to derive SWC using empirical or petrophysical models (e.g., Huisman et al., 2003; Klotzsche et al., 2018). To obtain subsurface images from GPR measurements, conventional ray-based inversion (RBI) approaches are typically applied (e.g., Dafflon et al., 2011). These methods only consider the first-arrival times and the first-cycle amplitudes of the measured waves in the inversion. Therefore, only a limited resolution can be achieved with RBI methods, and objects or even layers that are smaller than the dominant wavelength cannot be resolved. In contrast, full-waveform inversion (FWI) uses all the information contained in the measured data, including secondary events such as reflections and refractions. Hence, the theoretical resolution is in the order of half the dominant wavelength $\lambda/2$ of the signal (Virieux and Operto, 2009). FWI was first introduced by Tarantola (1984) for the inversion of seismic data. It is widely used in the seismic community and has continuously been developed further in the last 50 years (see overview provided by Virieux and Operto (2009)). As a result, a wide range of adaptations and applications for acoustic, elastic, and viscoelastic FWI approaches in the time- and frequency-domain have been proposed over the years (e.g., Pratt and Worthington, 1990; Pratt and Shipp, 1999; Marelli et al., 2012; Bohlen and Wittkamp, 2016; Liu, 2023). It should be noted that in the seismic community, multi-scale and multi-frequency approaches are used (Bunks et al., 1995; Sirgue and Pratt, 2004), which

Table 1.1.: Material properties of common materials in GPR studies: relative permittivity ε_r and electrical conductivity σ (modified after Blindow (2005)).

Material	$\varepsilon_r[-]$	σ [mS/m]
air	1	0
destilled water	80	0.01
fresh water	80	0.5
borehole water	80	50
sea water	80	4000
dry sand	3 - 5	0.01
saturated sand	20 - 30	0.1 - 1
silt	5 - 30	1 - 100
clay	5 - 40	2 - 1000
limestone	4 - 8	0.5 - 2
granite	6	0.01 - 1
ice	3.18	0.01

means that different frequency ranges are considered to create a suitable starting model for the inversion. Multi-scale approaches use a sequential inclusion of low to high-frequency components, resulting in a higher computational efficiency and more linear convergence. Multi-frequency approaches involve selecting different frequencies simultaneously to analyze and mitigate crosstalk artifacts. Furthermore, there are already some 3D FWI approaches developed and applied in seismics (e.g., Ben-Hadj-Ali et al., 2008; Warner et al., 2013; Butzer, 2015; Amestoy et al., 2016; Boddupalli et al., 2021).

Although FWI is widely used in the seismic community, it was rarely applied to GPR. The first adaptations to apply time-domain FWI to GPR data were carried out in the last two decades (Ernst et al., 2007b; Ernst et al., 2007a; Kuroda et al., 2007). The measurement types of GPR data used for GPR FWI can be roughly divided into crosshole, surface and off-ground. Similar to seismics, GPR FWI approaches can be divided into a time-domain and a frequency-domain approach. Frequency-domain approaches have been successfully investigated in several synthetic studies (e.g., Busch et al., 2012; Lavoué et al., 2014; Feng et al., 2021). However, the focus has rarely been extended to experimental data (e.g., Ellefsen et al., 2011; Pinard et al., 2016). So far, most of the applications for experimental data have been performed in the time-domain with more than 100 different datasets from various test sites (e.g., Klotzsche et al., 2019b; Güting et al., 2017; Zhou et al., 2020; Haruzi et al., 2022). All these studies investigated different aquifers and successfully demonstrated the potential to derive small-scale high resolution images of the permittivity and electrical

conductivity, showing variablities in the decimeter-scale. Surface FWI applications often address layered media models (e.g., Busch et al., 2012; Liu et al., 2018). Furthermore, there are also coupled strategies, e.g., FWI with GPR data with an indirect joint petrophysical inversion with shallow-seismic data by Qin (2022). Another approach is presented by Domenzain et al. (2020), where a joint inversion of GPR full-waveform and electrical resistivity data was performed for synthetic data to enhance the electrical conductivity models. Note that until now, combined time- and frequency-domain approaches are rarely applied in the GPR community. The idea of the seismic multi-scale FWI, is in the GPR community referred to as progressively expanded bandwidth approach or the frequency-hopping approach, which combines frequency- and time-domain benefits. Yang et al. (2014), Zhou et al. (2021), and others, showed the potential for GPR applications of combining time- and frequency-domain approaches to enhance model reconstruction by considering the benefits of both methods.

An important aspect for successfully performing FWI is the definition of adequate starting models. For time-domain FWI starting models that yield synthetic data within half the wavelength of the measured signals throughout the entire domain have to be used (Meles et al., 2011). This is to avoid cycle-skipping, where the wrong phase is fitted within the gradient approach. Ray-based approaches can provide starting models, that fulfill this half-wavelength criteria, which is required for a reliable application of time-domain FWI. However, in regions with high contrasts in permittivity and conductivity, it is often required to adapt these models to meet this criterion, for example with an amplitude analysis approach (e.g., Klotzsche et al., 2014; Zhou et al., 2020). Such updates can, for example, be performed by adding and testing different layers of permittivity in the starting model based on the result of an amplitude analysis approach and investigating the resulting data fit (e.g., Klotzsche et al., 2019b). It should also be noted that for experimental data, it is critically important to have an accurate starting model, as this is used to define the effective source wavelet that is used for the FWI. Therefore, errors in the starting model will directly propagate into the wavelet and affect the FWI results. Furthermore, note that while frequency-domain approaches are less dependent on this criterion, the choice of frequencies and the missing low-frequency components in experimental data limits the application of this approach.

One of the most common applications for time-domain GPR FWI is crosshole GPR. After the first pioneering work of Ernst et al. (2007b), Meles et al. (2010) improved the 2D crosshole GPR FWI with considering the field quantities as vectorial functions and

a simultaneous update of the permittivity and electrical conductivity distributions. An overview of the further developments, especially for the application to experimental data, is provided by Klotzsche et al. (2019b). Note that there are multiple pre-processing steps necessary for the experimental data, such as the time-zero correction. In addition to determining the starting models, estimating the effective source wavelet is another important step for the FWI. Especially for changing borehole filling effects, the FWI is challenging because of the different coupling phenomena that can not be described by only one source wavelet estimation (Klotzsche et al., 2019c; Mozaffari et al., 2022). Furthermore, the 2D nature of the forward model prevents the implementation of 3D structures, which have a great influence on the wave propagation, as well as secondary events like reflections. In particular, for the inversion of data measured both in unsaturated and saturated soils, the data has to be separated and each area inverted by its own. Klotzsche et al. (2019c) showed that FWI is possible, if a unique source wavelet is estimated for all combinations of transmitter and receiver in the unsaturated and saturated soil. An advantage of crosshole GPR FWI is that the distance between the antennas is known and the subsurface between the boreholes can be mapped with a large amount of data at different angles (Klotzsche et al., 2019b). The typical distance between boreholes for GPR FWI applications is between 2.5 m and 15 m, with a depth of up to 20 m. The distance is mostly limited by the attenuation of the electromagnetic wave, which is highly dependent on the local soil. Antenna frequencies for crosshole applications often range from 100 MHz to 250 MHz.

Although the 2D GPR FWI has shown to be a beneficial tool for high resolution aquifer characterization, the approach has also its limitations. Since the method considers a 2D forward model, the experimental data, which are measured in 3D, need to be transformed with the so called Bleistein filter to account for the different radiation pattern and amplitude decay in 3D and 2D. This filter causes errors especially at late arrival high amplitude data and is only valid in the far-field of the antenna, which complicates the extension towards surface data as everything there is in the near-field. Therefore, gradient-based 2D surface GPR FWI is not applied to experimental data for permittivity and conductivity inversion. To overcome the limitations of a 2D forward model, recent developments aim to move towards full 3D GPR FWI. The steady increase in computational capacity have made it possible to extend 2D GPR FWI to 2.5D by invariantly extending the 2D model in the perpendicular direction to create a 3D domain, as previously done in seismics (Song and Williamson, 1995; Tabarovsky et al., 1996). As shown in Mozaffari et al. (2020), a 2.5D GPR FWI is possible and reduces errors, especially in conductivity, since the 3D-to-2D correction with the Bleistein data transformation for the measured data is not necessary

anymore. Although this 2.5D GPR FWI has shown very promising results, developments are still ongoing, especially in terms of computational optimization to improve calculation times. Furthermore, the advancement of GPR modeling tools raises new possibilities, such as introducing detailed antenna geometries and extending FWI to 3D, since using graphic processing unit (GPU) architectures speeds up the simulations (Warren et al., 2019). Including antenna geometries in GPR modeling tools is important to better understand electromagnetic wave propagation and to improve the model towards reality. Simply creating a geometry is not sufficient to ensure that the signal matches reality as material properties often need to be adjusted as well. Several approaches have been developed using Taguchi's optimization method (e.g., Warren and Giannopoulos, 2011; Stadler and Igel, 2018), a hybrid linear/nonlinear FWI approach (e.g., Giannakis et al., 2019), or a particle swarm optimization algorithm (e.g., Stadler and Igel, 2022) have been proposed. Mozaffari et al. (2022) included boreholes and antennas in 3D forward simulations to show their influence and the influence of high angles on the signals of crosshole GPR data. In addition, a novel borehole-fluid effect correction was proposed that can be applied in the pre-processing steps to account for higher apparent velocities at higher angles in borehole data. It is important to note that Mozaffari et al. (2022) only included the boreholes and antennas in a regular forward model and performed travel time inversions but no FWIs. Including these geometries in the FWI forward model was not possible because of the high computational cost at that time.

Most of the time-domain 2D GPR FWI case studies conducted so far used frequencies up to 250 MHz and were able to derive decimeter-scale subsurface images. To image finer structures and properties related to the root zone, e.g., preferential flow paths in the vadose zone, higher frequencies are needed that can further enhance the resolution for materials that are commonly found in the critical zone. As a result of the stronger attenuation of higher frequencies in the subsurface, this also requires shorter offsets between the transmitter and receiver antennas. Therefore, the use of lysimeters, where antennas are placed around the soil column (similar to the procedure described by Schmalholz et al. (2004)), or boxes like Mangel et al. (2012), are ideal methods to study and enhance the understanding of centimeter-scale processes and the soil-root continuum using GPR FWI. Note that using higher frequencies imposes greater constraints on the FWI starting model because of the shorter wavelengths of the signals and the requirement to meet the half-wavelength criterion.

1.1. Thesis objective and outline

The primary objective of this thesis is to advance the current GPR FWI method towards a 3D GPR FWI for a high resolution subsurface characterization. To achieve this, we present a novel 2.5D GPR FWI approach that can include small-scale geometries into the forward modeling. In this thesis, we state our objectives as the following hypotheses:

1. Using a more complex forward model, including borehole fillings and finite-length antenna models, can provide source wavelets closer to reality.
2. Using an improved source wavelet based on a forward model, including borehole fillings and finite-length antenna models, can improve the accuracy of reconstruction.
3. A 2.5D GPR FWI with a realistic forward model can account for variable borehole fillings and antenna coupling effects with one estimated effective source wavelet.
4. Higher frequencies in 2D GPR FWI yield improved reconstruction resolution within the sub-decimeter-scale.
5. A 2.5D GPR FWI, including lysimeter geometries, can account for the different wavetypes and successfully invert synthetic lysimeter data in a sub-decimeter-scale.

The hypothesis 1 and 2 were both investigated in Chapter 3, while 3, 4, and 5 are tested in Chapter 4, Chapter 5, and Chapter 6. The thesis aims to enhance the geophysical community's comprehension of subsurface phenomena, particularly in the critical zone. It has the potential to significantly improve non-destructive subsurface investigation for environmental, agricultural, and resource management applications.

After the introduction, we propose a new 2.5D GPR FWI in Chapter 3, which is able to include additional geometries in the forward model. With this, we do not have to investigate different 3D specific effects such as borehole filling or antenna effects (e.g., Irving and Knight, 2005; Mozaffari et al., 2022) separately from the GPR FWI, but are able to account for them in the new 2.5D GPR FWI. Therefore, we investigate the influence of the inclusion of borehole fillings and finite-length antenna geometries in the forward model of a 2.5D GPR FWI. This ranges from the source wavelet estimation to the inversion results and also the influence of higher angles.

We then use the knowledge based on Chapter 3 in Chapter 4 to investigate the source wavelet estimation for unsaturated and saturated soils. An approach for 2D GPR FWI was proposed by Klotzsche et al. (2019c), which is based on four different estimated source

wavelets, as the coupling of the antenna to the borehole and soil is different if the antenna is in the unsaturated or saturated area. Including borehole and antenna models in the forward model of the FWI should make this obsolete, as the different couplings should now be included in the improved forward model. Therefore, we will investigate the possibility of performing a 2.5D GPR FWI on measured data without estimating different source wavelets for the unsaturated and saturated zones. We include air and water filled boreholes and a finite-length antenna model in the forward model for the source wavelet estimation and also in the 2.5D GPR FWI. If this forward model is accurate enough, it should account for the different coupling effects and yield one effective source wavelet for all possible combinations. After this we will show inversion results of a 2.5D GPR FWI with one source wavelet and including all data of the unsaturated zone as well as the saturated zone.

To investigate high resolution reconstruction in GPR we investigate the resulting challenges in Chapter 5 with the 2D FWI. High-frequency GPR data leads to stricter specifications of the starting models of the FWI. We introduce the frequency-hopping approach to meet these requirements. Until now, it was used at low frequencies to achieve better reconstructions (e.g., Meles et al., 2011; Zhou et al., 2021). Furthermore, we show that the use of high resolution reconstruction can improve the reconstruction of a more heterogeneous model.

In Chapter 6, we demonstrate why it is necessary to use 2.5D GPR FWI for GPR data measured on lysimeters. Based on a synthetic 3D dataset with a 450 MHz source wavelet, we investigate the reconstruction of the soil within a lysimeter with the 2.5D GPR FWI, which was not possible until now (Klotzsche, 2023).

The conclusion of this thesis, as well as an outlook for further improvements and applications, are presented in Chapter 7.

Appendix A is a preparatory study for Chapter 5. Different types of reflections inside a lysimeter are investigated in a synthetic case and measured data. A comparison between an unsaturated and saturated soil provides an overview of the influence of the medium on the reflections. Therefore, it can be considered a guideline for analyzing the different reflections.

2. Theory

This section provides an overview of the theory used in this thesis, from the basic introduction of fundamental electromagnetic theory to the numerical modeling of electromagnetic waves, and their application in GPR FWI.

2.1. Electromagnetic theory

GPR is based on measuring electromagnetic waves. Maxwell's equations describe how electric and magnetic fields interact and form their waves. Their description in this section is widely based on Taflove and Hagness (2005). If we consider a spatial region without electric or magnetic current sources, while taking into account the possible presence of materials capable of absorbing electric and magnetic field energy, we can express the time-dependent Maxwell's equations in their differential form as:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.1}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J} \tag{2.2}$$

$$\nabla \cdot \mathbf{D} = 0 \tag{2.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.4}$$

where $\mathbf{E}$ is the electric field [V/m], $\mathbf{D}$ the electric flux density [C/m^2], $\mathbf{H}$ the magnetic field [A/m], $\mathbf{B}$ the magnetic flux density [W/m^2], and $\mathbf{J}$ the electric current density [A/m^2].

These equations imply the following: Faraday's law in Equation (2.1) states that a time-varying magnetic field induces an electric field. Ampere's law with Maxwell's extension in Equation (2.2) indicates that magnetic fields can be generated by electrical currents (as described by Ampere's original law) and also by changing electric fields (Maxwell's addition). Gauss's law for electricity in Equation (2.3) implies that the field lines neither diverge nor converge. Gauss's law for magnetism in Equation (2.4) states that there are no

magnetic monopoles. Consequently, magnetic field lines are continuous and form closed loops or extend to infinity without ever beginning or ending at a point.

If we consider only linear, isotropic, and nondispersive materials, it is possible to relate **D** to **E** and **B** to **H**. Note that nondispersive materials are characterized by their field-independent, direction-independent, and frequency-independent electric and magnetic properties. The constitutive relationships are:

$$\mathbf{D} = \varepsilon\mathbf{E} = \varepsilon_r\varepsilon_0\mathbf{E}; \quad \mathbf{B} = \mu\mathbf{H} = \mu_r\mu_0\mathbf{H}, \tag{2.5}$$

where ε is the electrical permittivity [F/m], ε_r the relative permittivity [-], ε_0 the free-space permittivity (8.854×10^{-12} [F/m]), μ the magnetic permeability [H/m], μ_r the relative permeability [-], and μ_0 the free-space permeability $4\pi \times 10^{-7}$ [H/m]). Although these assumptions are commonly used, especially for the modeling of electromagnetic wave propagation, it is important to be aware that some materials on earth are non-linear, anisotropic, and dispersive.

The current density $\boldsymbol{J}$ can be a distinct source of electric field energy, referred to as $\boldsymbol{J}_{source}$. Materials that possess isotropic, nondispersive characteristics in terms of electric loss, lead to the attenuation of the electric fields through the conversion into heat energy, resulting in:

$$\mathbf{J} = \mathbf{J}_{source} + \sigma\mathbf{E}, \tag{2.6}$$

where σ is the electric conductivity [S/m]. Considering this, we derive two new equations for linear, isotropic, nondispersive, and lossy materials, which are also called Maxwell's curl equations:

$$\frac{\partial\mathbf{H}}{\partial t} = -\frac{1}{\mu}\nabla \times \mathbf{E} \tag{2.7}$$

$$\frac{\partial\mathbf{E}}{\partial t} = \frac{1}{\varepsilon}\nabla \times \mathbf{H} - \frac{1}{\varepsilon}\left(\mathbf{J}_{source} + \sigma\mathbf{E}\right). \tag{2.8}$$

By eliminating either the electric or magnetic field the wave character of the equations can be derived. In the case of eliminating the magnetic field it yields:

$$\nabla \times \nabla \times \mathbf{E} - \mu\varepsilon\frac{\partial^2\mathbf{E}}{\partial t^2} - \mu\sigma\frac{\partial\mathbf{E}}{\partial t} = 0. \tag{2.9}$$

Note that the electromagnetic waves are transverse, resulting in an oscillation of the electric field perpendicular to the direction of propagation of the wave.

The material properties that determine electromagnetic waves and fields are therefore electrical permittivity, magnetic permeability, and electrical conductivity. For the application of GPR the electrical permittivity and electrical conductivity are the primary parameters, as the relative permeability is often assumed to be constant and equal to 1 and therefore be $\mu = \mu_0$. The analysis of the material influence is based on Jol (2009).

The electrical permittivity quantifies the electric polarizability of the material. Together with the magnetic permeability we can calculate the wave velocity v of an electromagnetic wave travelling through a medium:

$$v = \frac{1}{\sqrt{\varepsilon\mu}}, \tag{2.10}$$

with the relative permeability being constant, this shows, that the electrical permittivity affects the speed of which the electromagnetic wave through a medium. In the case of vacuum where $\varepsilon_r = \mu_r = 1$ we get the speed of light: $\frac{1}{\sqrt{\varepsilon_0\mu_0}} \simeq 3 \times 10^8 \text{m/s} = c$. It is also possible to calculate the velocity based on the wavelength λ and the frequency f of the wave through:

$$v = \lambda f. \tag{2.11}$$

The attenuation of an electromagnetic wave in a medium is given by:

$$\alpha = \frac{\mu\sigma v}{2} = \frac{\sigma}{2\sqrt{\frac{\varepsilon}{\mu}}}. \tag{2.12}$$

Electromagnetic waves change if they encounter different media. Snell's law brings the angles of incidence θ_i and refraction θ_t together. In the case of low loss media we get:

$$\frac{\sin\theta_i}{v_1} = \frac{\sin\theta_t}{v_2}, \tag{2.13}$$

with the media velocities v_1 and v_2. It has to applied every time there is a change in medium interface. The angle of reflection θ_r is given by:

$$\theta_i = \theta_r. \tag{2.14}$$

Since parts of the incident wave are reflected (R) or transmitted (T) into the second medium, using Fresnel equations can provide the different contributions dependent on the polarization of the incident wave to the plane of incidence. In the case of only non-magnetic materials

(i.e., $\mu_1 = \mu_2 = \mu_0$) and if the polarization of the wave is perpendicular to the plane of incidence (s-polarized) the reflected and transmitted part can be described as:

$$R_s = \left| \frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t} \right|^2 = \left| \frac{n_1 \cos\theta_i - n_2 \sqrt{1 - \left(\frac{n_1}{n_2} \sin\theta_i\right)^2}}{n_1 \cos\theta_i + n_2 \sqrt{1 - \left(\frac{n_1}{n_2} \sin\theta_i\right)^2}} \right|^2, \tag{2.15}$$

$$T_s = 1 - R_s, \tag{2.16}$$

with the refractive indexes n_1 and n_2, which can be calculated as:

$$n = \frac{c}{v}. \tag{2.17}$$

If the polarization of the wave is parallel to the plane of incidence (p-polarized) the reflected and transmitted parts are:

$$R_p = \left| \frac{n_1 \cos\theta_t - n_2 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i} \right|^2 = \left| \frac{n_1 \sqrt{1 - \left(\frac{n_1}{n_2} \sin\theta_i\right)^2} - n_2 \cos\theta_i}{n_1 \sqrt{1 - \left(\frac{n_1}{n_2} \sin\theta_i\right)^2} + n_2 \cos\theta_i} \right|^2, \tag{2.18}$$

$$T_p = 1 - R_s. \tag{2.19}$$

Note that we do not include dispersive media in this work. For more information please refer to Zhang and Li (1998).

2.2. Numerical modeling of electromagnetic waves

In this section, we discuss the simulation of electromagnetic waves (EM) in more detail. Computational techniques are essential for simulating and analyzing electromagnetic interactions and various methods are possible. The most commonly used one is the finite-difference time-domain method. Other numerical modeling options are the time-domain finite element method (TDFEM) and the finite integration technique (FIT). For more information about the TDFEM please refer to Lee et al. (1997) and for the FIT to Weiland (1977) and Weiland (2003).

2.2.1. Finite-difference time-domain method

The finite-difference time-domain (FDTD) method was introduced by Yee (1966) and has become the most often used approach for time-domain GPR FWI. Major parts of this section are based on Taflove and Hagness (2005). This section provides only an overview of the method. For more detailed information, please refer to Taflove and Hagness (2005). The development of the FDTD method represented a significant advance in computational electromagnetics, because of its ability to solve the partial differential equations (PDEs) that describe electromagnetic phenomena. The solutions to these PDEs are generally unknown except for a few special cases, making the need for numerical methods such as FDTD critical to understand wave propagation and electromagnetic fields in complex environments. It was first introduced for two-dimensional (2D) problems, a significant development at the time, as it was considered highly computationally expensive at the time. Nowadays, as the long-run index of integer computations per second has increased from 1×10^5 to 1×10^8 in the 1960s to 1×10^{13} to 1×10^{16} for the 2020s (based on Coyle and Hampton (2024)) even 3D problems can be easily solved.

Yee's approach involves translating Maxwell's partial differential equations into a set of finite-difference equations. These finite-difference equations can be solved to approximate the electromagnetic fields in time and space. Within the FDTD the domain is subdivided by the so called Yee grid (Figure 2.1), which is a staggered grid that separates the electric and magnetic fields into two nested grids in space.

Thereby, one grid is dedicated to the electric field and another spatially, for half a cell, shifted grid for the magnetic field. This arrangement ensures that each electric field component is surrounded by four magnetic field components and vice versa, allowing direct numerical simulation of Maxwell's curl equations. In addition, the electric and magnetic field components are centered in time in a leapfrog scheme, where the electric field is updated in integer time steps and the magnetic field is updated in half-integer time steps. This leapfrog time-stepping scheme is fully explicit and achieves second-order accuracy through central-difference time derivatives. Additionally, this scheme is very suitable for a parallel computing implementation.

For a simulation using the FDTD method, a domain must be specified in which material properties are assigned to each cell. These material properties (relative permittivity, electrical conductivity, relative permeability, and magnetic loss) are already known from Section 2.1. In addition, sources with appropriate wavelets must be defined to initiate the

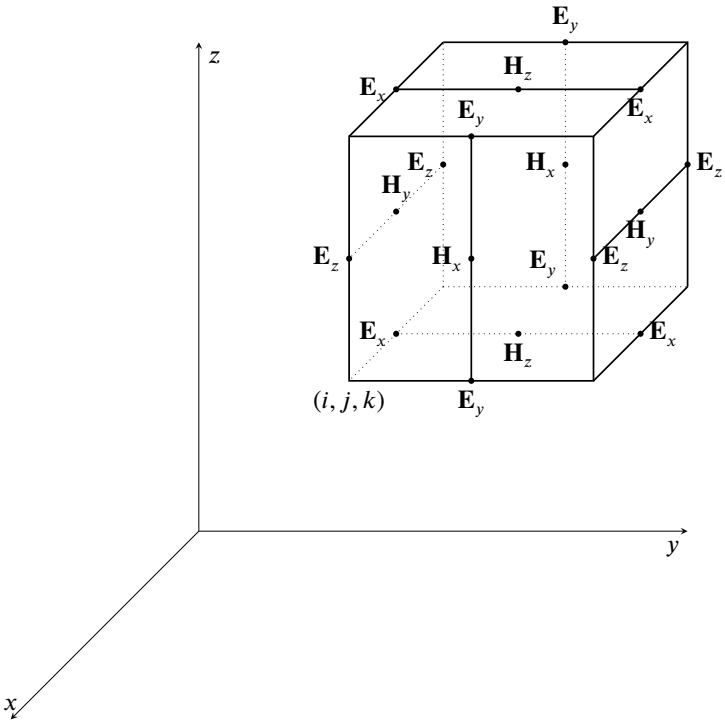

Figure 2.1.: 3D representation of a Yee grid cell including **E** and **H** field contributions (based on Yee (1966)).

electromagnetic wave propagation. The simulation proceeds by solving alternately for the electric field vector components at a given time step and then for the magnetic field vector components at the next half time step, over the same spatial volume. This process is repeated for all time steps.

To guarantee a stable simulation, the space and time steps must satisfy the Courant-Friedrichs-Lewy (CFL) stability condition (Courant et al., 1928), otherwise the leapfrog integration used to solve the partial differential equation is likely to become unstable:

$$\Delta t \leq \frac{1}{c\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2} + \frac{1}{\Delta z^2}}}, \tag{2.20}$$

where Δt is the time step, c is the speed of light in the vacuum or the wave propagation speed in the medium, Δx, Δy, and Δz are the spatial discretization steps in the x, y, and z

directions, respectively. For 2D, we consider the unused spatial discretization step Δu to be infinite.

Another issue is the numerical dispersion that can lead to erroneous simulation results. This effect occurs when the electromagnetic wave can not be properly be resolved with the given grid size. In this case, the physical phase velocity is not equal to the numerical phase velocity. Therefore, to account for this, we need to consider the minimum velocity v_{min}:

$$v_{\text{min}} = \frac{c}{\sqrt{\varepsilon_{\text{max}}}}, \tag{2.21}$$

the minimal wavelength in medium λ_{min}:

$$\lambda_{\text{min}} = \frac{v_{\text{min}}}{\text{f}_{\text{max}}}, \tag{2.22}$$

the maximum grid cell size Δ_{max}:

$$\Delta_{\text{max}} = max(\Delta x, \Delta y, \Delta z), \tag{2.23}$$

the Courant stability factor C_{stab}:

$$\text{C}_{\text{stab}} = \frac{c\Delta t}{\Delta_{\text{max}}}, \tag{2.24}$$

and the grid sampling density N:

$$\text{N} = \frac{\lambda_{\text{min}}}{\Delta_{\text{max}}}, \tag{2.25}$$

with the maximum permittivity ε_{max} and the maximum frequency f_{max} of the model. The f_{max} can be assumed to be 1.5 times of the center frequency of the signal. Using these equations, we can calculate the numerical phase velocity v_p and the resulting physical phase-velocity error $\Delta v_{\text{physical}}$:

$$v_p = \frac{\pi}{\text{N} \arcsin\left(\frac{1}{\text{C}_{\text{stab}}} \sin\left(\frac{\pi \text{C}_{\text{stab}}}{\text{N}}\right)\right)} \tag{2.26}$$

$$\Delta v_{\text{physical}} = \frac{v_p - c}{c} 100. \tag{2.27}$$

A common approach in discretization is to determine the maximum grid cell size Δ_{max} with (gprMax, 2024):

$$\Delta_{\text{max}} = \frac{\lambda_{\text{max}}}{10}. \tag{2.28}$$

Although the FDTD method does not explicitly require the source wavelet to be continuous or steady in the mathematical sense, there are practical considerations and requirements to ensure accurate simulation results. The grid must be fine enough to accurately resolve the highest frequency components in the source wavelet. This is constrained by the Nyquist criterion, which implies that the grid spacing (Δx, Δy, Δz) must be small enough to capture at least two points per wavelength of the highest frequency present in the source wavelet. A source wavelet that rises smoothly from zero helps to avoid introducing a wide range of frequencies and reduces the chance of exciting resonances unrelated to the physical problem being modeled. Sharp transitions or discontinuities in the wavelet can excite high-frequency components that the grid may not resolve well, introducing oscillations.

Most FDTD approaches use an equidistant grid. This is a valid approach for many applications. However, it may not efficiently resolve fine features or sharp discontinuities in the geometry, materials, or electromagnetic fields. In addition, the grid may not align with the physical boundaries of geometries or materials, leading to staircasing errors. Round geometries in particular suffer from this. The main solution for this would be to significantly reduce the grid size, which would drastically increase the computational requirements like computation time and required memory. To address this, implementing a subgridding approach can effectively reduce the grid size while mitigating the increase in computational demands. In a subgridding approach several blocks of grid cells, smaller than the main grid cells, can be placed at specific locations in the main grid without modifying it. A possible example is the Huygens subgridding approach (Bérenger, 2006), implemented by Hartley et al. (2018) in the open-source FDTD solver gprMax (Warren et al., 2016).

Absorbing boundary conditions (ABCs), like the perfectly matched layer (PML) (Bérenger, 1994), are necessary to replicate the effect of an unbounded domain within a finite simulation domain. The PML is an absorbing layer capable of absorbing electromagnetic waves with minimal reflection, regardless of the wave's angle of incidence, polarization, or frequency. The PML works by surrounding the simulation domain with an artificial layer, where the electromagnetic fields are exponentially attenuated without inducing spurious reflections back into the simulation domain.

To achieve the 2D FDTD method the 3D FDTD is split into transverse electric (TE) and transverse magnetic (TM) modes, depending of the invariant direction and of the chosen preference of electric or magnetic fields. Therefore, the computation time and memory is reduced drastically, compared to the full 3D FDTD. In 2D FDTD simulations, structures are assumed to be invariant along one axis. Hence, commonly line sources are used, to excite

the fields, emitting radially propagating cylindrical waves, having an amplitude scaling of $\frac{1}{\sqrt{r}}$. This type of simulation is very efficient for modeling planar structures. However, 2D FDTD is limited in its ability to account for the effects related to changes along the invariant axis, e.g., 3D structures.

In 3D FDTD simulations even complicated structures can be included. Note that these are restricted by the grid size as well. Furthermore, the computational requirements are significantly higher because of the added dimension. Two types of sources, soft and hard, are commonly used to excite the electromagnetic fields. Soft sources are characterized by their ability to specify a current density term at a field location, which is used to simulate a hertzian dipol. On the other hand, hard sources are idealized as point sources that inject fields directly into the computational domain. Both of these sources generate a spherical wave front, with its energy spread over the surface of a sphere, resulting in an amplitude scaling of $\frac{1}{r}$.

2.3. Ground-penetrating radar full-waveform inversion

Before the FWI can be applied to observed data, the data has to be pre-processed, a starting model to be generated, and a source wavelet to be estimated. An overview about the individual steps, as well as the FWI is given by the flowchart in Figure 2.2. In the following we will introduce the steps closer.

2.3.1. Pre-processing

In numerous cases, measured data is affected by noise caused by various technological factors such as radio stations, cell phones, power substations, power lines, or even the radar system itself and unwanted coupling effects. During the measurement, stacking can be used to reduce noise (Stephan et al., 2022). However, there are still many datasets where noise is nevertheless a dominant error factor. Commonly a noise filter is applied to reduce high- and low-frequency noise. Often simple filters, like band-pass filters which lets through a specific range of frequency components, prove to be quite effective (Annan, 1993). Due to the saturation of the signal by early arrivals a dewow filter is necessary (Huisman et al., 2003).

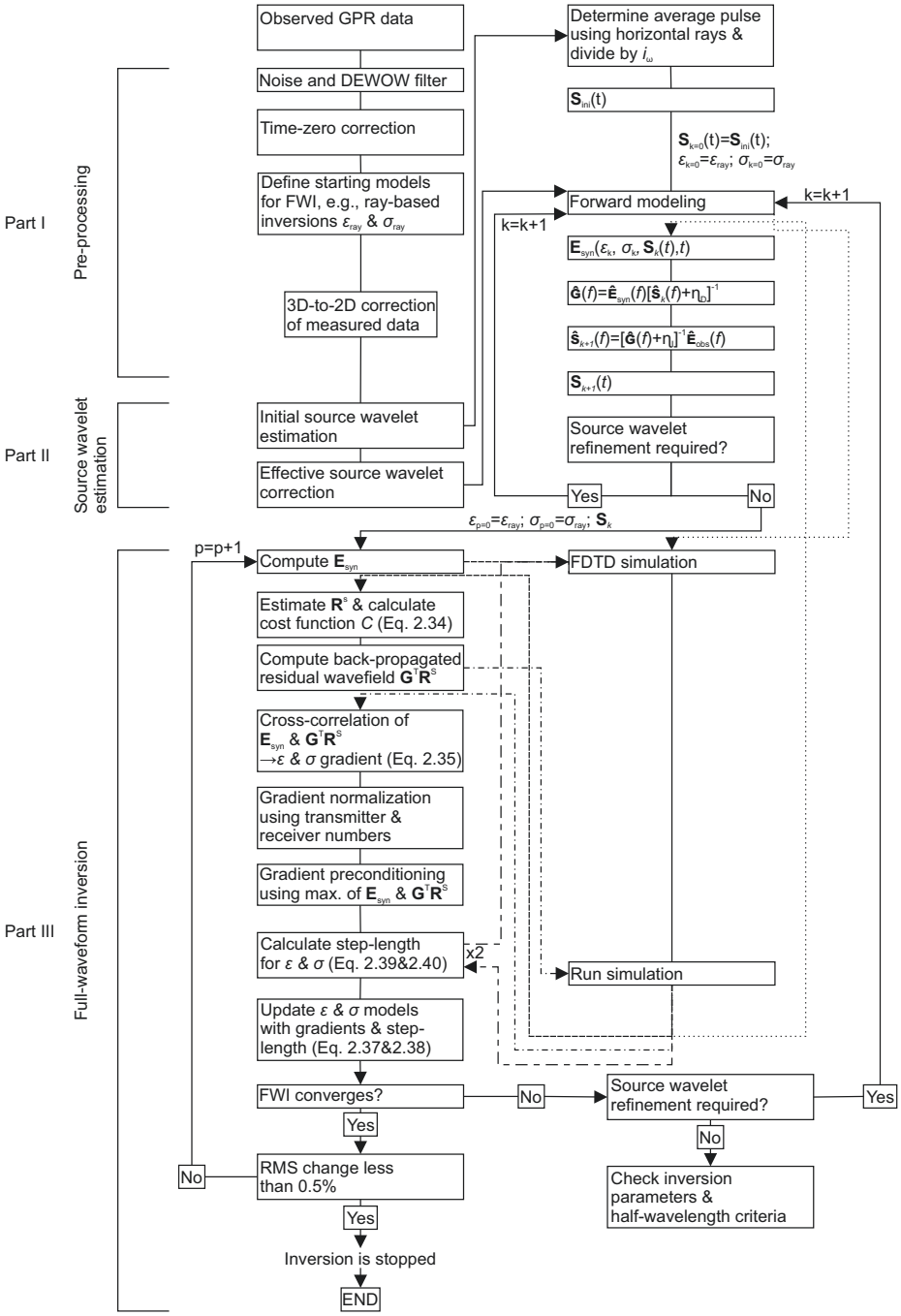

Figure 2.2.: Flowchart of the 2D GPR FWI inversion steps, adapted from Klotzsche et al. (2019b).

The accurate time-zero position of the radar signals is important for the FWI because inaccuracies in the inversion directly affect the permittivity results. Different data acquisition settings and equipment or instrumental delays can cause a misalignment. To align the signal's starting time, a time-zero correction is necessary. Typically, common midpoint (CMP) or wide-angle reflection and refraction (WARR) measurements in air are performed before and after the multiple offset gather (MOG) measurements (Klotzsche et al., 2019b). Interpolating the time-zero for each MOG provides the absolute time-zero. However, because of the potential variation in time-zero shifts during measurements and the sensitivity of the full waveform to even small time-zero errors, this interpolation may not be sufficiently accurate. Oberröhrmann et al. (2013) proposed a method utilizing equal transmitter (Tx) receiver (Rx) combinations. Zero-offset (ZOP) traces and horizontally traveled traces within the MOGs are cross-correlated to calculate the time-zero individually for the MOGs.

2.3.2. Starting model generation

To prevent cycle skipping during the FWI as well as to achieve more stable and reliable results, an overlap between the starting model data and the measured data of at least half a wavelength is required (Meles et al., 2011), which is referred to as the half-wavelength criterion. To obtain such an accurate starting model, the regular approach is to perform a ray-based inversion (RBI) on the measured data. Another approach for accurate starting models is the frequency-hopping approach (Zhou et al., 2021).

2.3.2.1. Ray-based inversion

The main idea of the RBI, is as the name states, the assumption that the wave travels as a ray between the Tx and Rx. The inversion of first-arrival times can give information about the velocity v and the inversion of the first-cycle amplitude about the attenuation α of the medium. The simplest approximation is to assume a straight travel path, indicating not changing velocities in the ray's path. A more precise method is to consider the varying velocities along the ray's path (based on Snells' law in Equation (2.13)), which can lead to curved travel paths. However, it also offers the possibility of more realistic velocity distributions Lanz et al. (1998). With the travel times for all transmitter-receiver combinations known, a finite-difference implementation of the Eikonal equation (Vidale, 1990) can be used to solve the forward problem. The eikonal solver for ray-based inversion is based on an equidistant grid where velocities are specified at each point. To minimize discretization

errors, as noted by Rabbel (2009), the spacing between grid points, both horizontally and vertically, should be narrower than the spacing between receivers. Note that this section is based on Klotzsche (2013).

First-arrival travel time inversions can be used to estimate a velocity model, which is inversely proportional to the ε_r, see (2.10). After defining a starting model, the inversion steps are:

1. Generate the travel times for the current model with forward modeling (e.g., an eikonal solver).
2. Estimate the error between the observed and simulated travel times.
3. Convert the error to an improved model.
4. Update the model.

The travel time misfit function C_{TT} in step 2 can be described as:

$$C_{\mathrm{TT}} = \sum_{s} \sum_{r} \frac{\left(t^{sr}_{\mathrm{obs}} - t^{sr}_{\mathrm{syn}}\right)^2}{n}, \tag{2.29}$$

where n, s and r are indicating the datapoints, the source and the receiver number, respectively, for the corresponding observed t^{sr}_{obs} and simulated travel times t^{sr}_{syn}. Step 3 can be expressed as solving a sparse matrix, for which multiple approaches can be used, e.g., like a least squares proposed by Lanz et al. (1998).

First-cycle amplitude inversions are used to estimate the attenuation α of the electromagnetic waves. As the name states, the main information for this inversion is the first-cycle amplitude, which is an exponential decay of the initial source amplitude and proportional to α (Giroux et al., 2007). The attenuation is depending on the permittivity (velocity) as well as the conductivity (see Equation (2.12)). A first-arrival travel time inversion and first-cycle amplitude inversion are commonly combined. In addition to the first-cycle amplitude, other details such as travel times and radiation patterns are necessary. For more information, please consult Holliger et al. (2001) and Maurer and Musil (2004).

The low computational demands of the RBI make it very attractive for a preliminary investigation. However, the inversion is only depending on two parameter sets, the first-arrival time and the amplitude of the first-cycle, neglecting a big amount of data stored in the measured signals. Therefore, the reconstruction resolution is limited. Since the conductivity is also dependent on the permittivity, most commonly the reconstruction of the

permittivity is better than the conductivity reconstruction. For standard ray-based crosshole GPR inversion approaches the reconstruction resolution scales approximately with the diameter of the first Fresnel zone, or $\sqrt{\lambda L}$ with λ and L indicating the dominant wavelength and path length, respectively (Williamson, 1991). Due to this resolution, issues often occur at high-contrast zones. Note that the ray-based models need to be updated in these high-contrast regions such as water table and waveguiding structures (Klotzsche et al., 2014) if they shall be used as starting models for the FWI.

2.3.2.2. Frequency hopping

The frequency hopping approach utilizes the wide frequency spectrum of the measured GPR data. It was used in frequency-domain for microwave imaging applications (Chew and Lin, 1995; Dubois et al., 2009) and for seismic data (Pratt et al., 1998; Maurer et al., 2009). First a low-frequency bandwidth inversion is performed and the results are used for the next inversion with a progressively higher frequency bandwidth, which is repeated till the centre frequency of the data. Meles et al. (2011) proposed a workflow for synthetic GPR data where the frequency of the source wavelet is iteratively increased, while the observed data retains the full bandwidth. Due to problems with real data, Zhou et al. (2021) proposed a workflow, which is the basis for the used frequency hopping approach in this work. A tapered bandpass filter is applied to the effective source wavelet (used to generate the synthetic data), as well as the observed data. Then, a specified number of FWI iterations are performed and the result is used for the next-largest frequency. This is continued until the entire frequency range is covered and subsequently the original effective source wavelet and data are considered. Usually a RBI model is used for the initial permittivity and conductivity of the frequency hopping. With this approach it is possible to overcome the problems of the RBI in the presence of high contrast layer, where the ray-based starting models for the FWI are inaccurate and need adaptions (e.g., Klotzsche et al., 2012; Klotzsche et al., 2013; Klotzsche et al., 2014).

2.3.3. 3D-to-2D correction

Measured GPR data can always be considered as 3D data. As mentioned in Section 2.2.1, there are differences in the modeling when considering 2D or 3D. For example, there are different radiation patterns, from spherical in 3D to cylindrical in 2D, which results in a different amplitude decay A. This decay is proportional to the travelled distance r and

time t, where $r = vt$. In 3D it is $A \propto \frac{1}{r} \propto \frac{1}{t}$ and in 2D $A \propto \frac{1}{\sqrt{r}} \propto \frac{1}{\sqrt{t}}$. This means, that if we want to use 3D data, we have to multiply the data with $\sqrt{t}$ if we use a 2D forward simulation in a FWI. Furthermore, there is a phase shift, as 2D sources are considered as infinitive long extended lines of point dipole. If we use a 2D forward model, we have to compensate for this, too. Most commonly a 3D-to-2D data transformation based on Bleistein (1986), as proposed by Ernst et al. (2007a) is used. Note that the correction here is based on Klotzsche (2013).

In our case we use a phase shift of $\frac{\pi}{4}$ with a scaling factor of $\frac{1}{\sqrt{\omega}}$ in the frequency domain, resulting in:

$$\hat{\mathbf{E}}_{2D}(x_s, x_r, \omega) = \hat{\mathbf{E}}_{obs}(x_s, x_r, \omega)\sqrt{\frac{2\pi t_{obs}(x_s, x_r)}{-j\omega\varepsilon^{mean}\mu}} \tag{2.30}$$

where $\hat{\mathbf{E}}_{2D}$ are the corrected and $\hat{\mathbf{E}}_{obs}$ the original measured data, while the $\hat{}$ indicates the frequency domain, x_s and x_r indicate the transmitter and receiver location, respectively, and $t_{obs}(x_s, x_r)$ are the travel times of the data that will be corrected.

Note that while Ernst et al. (2007a) attested a good agreement between 3D and corresponding 2D data in the far-field, Mozaffari et al. (2020) showed that the highest amplitudes do not always correlate to the first-arrival times. This is especially the case in high contrast layers. Such discrepancies can introduce errors, affecting the accuracy of the permittivity and conductivity reconstruction. In addition, out-of-plane arrivals in the data, caused by the presence of 3D geometries, can introduce further errors.

2.3.4. Source wavelet estimation

The last preparation for the FWI is the source wavelet estimation (see Figure 2.2 Part II). The source wavelet is unknown for measured data, as the antenna as well as the coupling effects change the emitted source wavelet. Since a source wavelet is necessary for the forward modeling with the FDTD method, it has to be estimated. The process can be divided in the initial source wavelet estimation (estimation of the shape of the wavelet) and the source wavelet correction (shape and amplitude of the wavelet). This is based on Klotzsche et al. (2019b).

The initial source wavelet is estimated by determining the average pulse by using the normalized and aligned approximately horizontally traces. Data with erroneous wave

shapes, e.g., from reflections, can be identified by cross-correlating the profiles and then be excluded. Since the electric field is proportional to the time derivative of the current density source, which corresponds to a multiplication by $j\omega$ in the frequency domain, it is necessary to divide the Fourier-transformed mean pulse by $j\omega$, to derive the initial source wavelet in the frequency spectrum. The wavelet is then normalized to 1 in the time-domain.

The source wavelet correction uses a deconvolution approach to correct the initial source wavelet $\mathbf{S}_{\text{ini}}$ in shape and amplitude. We first use the FDTD to generate a synthetic dataset $\mathbf{E}_{\text{syn}}$ for all transmitter and receiver combinations with $\mathbf{S}_{\text{ini}}$ and the starting model (ray-based). In the time-domain, a GPR dataset can be represented as a convolution of the source wavelet with the medium's impulse response, the Greens function $\mathbf{G}$. The effective source wavelet $\mathbf{S}_k$ can then be calculated by the deconvolution of the observed data $\mathbf{E}_{\text{obs}}$ with $\mathbf{G}$. To calculate $\mathbf{G}$, we can apply a spectral division on the synthetic field $\mathbf{E}_{\text{syn}}$ with the initial wavelet spectrum. A new estimated source wavelet $\mathbf{S}_{k+1}$ can be calculated by dividing $\mathbf{E}_{\text{obs}}$ with the newly calculated $\mathbf{G}$. To ensure numerical stability and prevent division by zero, prewhitening factors η_D and η_I are introduced. Through iterative deconvolution, which can be applied prior to or during FWI, the time-domain source wavelet $\mathbf{S}_{k+1}$ is progressively refined until its shape and amplitude consistently align with the observed data.

2.3.5. Full-waveform inversion[1]

The FWI used in this study is based on Meles et al. (2010) and the extensions illustrated in Klotzsche et al. (2019b). Therefore, in this section we will give an overview about the important steps that are essential (Figure 2.2 Part III). The full-waveform inversion of GPR data is based on solving Maxwells equations in a forward problem in the time-domain to estimate electromagnetic fields for certain model parameters. Note that this approach, as the majority of GPR FWI, assumes frequency-independent medium properties. Only few approaches investigate GPR FWI with frequency-dependent medium properties (Qin et al., 2022).

[1] adapted from D. Hoven, A. Mester, H. Vereecken, and A. Klotzsche (2023). "Evaluation of starting model approaches and effective source wavelet variations for high-frequency ground-penetrating radar full-waveform inversion". In: *GEOPHYSICS* 88.2, KS27–KS45. DOI: 10.1190/geo2021-0683.1

Maxwells equations (Equations (2.1), (2.2), (2.3), and (2.4)) can also be expressed in a compact form as (Meles et al., 2010):

$$\mathbf{M}(\varepsilon,\sigma)\begin{bmatrix}\mathbf{E}^s\\ \mathbf{H}^s\end{bmatrix}=\begin{bmatrix}\mathbf{J}^s\\ \mathbf{0}\end{bmatrix} \tag{2.31}$$

with the linear operator $\mathbf{M}(\varepsilon,\sigma)$, which can be rewritten as:

$$\begin{bmatrix}-\varepsilon(\mathbf{x})\frac{\partial}{\partial t}-\sigma(\mathbf{x}) & \nabla\times\\ -\mu_0\frac{\partial}{\partial t} & \nabla\times\end{bmatrix}\begin{bmatrix}\mathbf{E}^s(\mathbf{x},t)\\ \mathbf{H}^s(\mathbf{x},t)\end{bmatrix}=\begin{bmatrix}\mathbf{J}^s(\mathbf{x},t)\\ \mathbf{0}\end{bmatrix}. \tag{2.32}$$

The fields are vector quantities, exist for each individual transmitter, marked with the index s for source, and are locally defined at any point of space $\mathbf{x}$ and time t. They thus cover the entire space-time-domain of interest. The magnetic permeability is often simplified to its value in free space ($\mu_r = 1$); it seldom has an influence on GPR wave propagation, and leads for non-magnetic material to $\mu = \mu_0$. The medium properties are scalar functions of position, while the field quantities (in bold) are vectorial functions of both space and time. If we only investigate the electric field we can rewrite Equation (2.32) to:

$$\mathbf{E}^s = \hat{\mathbf{G}}\mathbf{J}^s, \tag{2.33}$$

where $\hat{\mathbf{G}}$ is the Green's operator of $\mathbf{M}$.

Time-domain FWI uses a conjugate gradient-type method that minimizes the misfit function C between measured data $\mathbf{E}^s_{\text{obs}}$ and modeled data $\mathbf{E}^s_{\text{syn}}$:

$$C(\varepsilon,\sigma)=\frac{1}{2}\sum_s\sum_r\sum_\tau\left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)-\mathbf{E}^s_{\text{obs}}\right]^T_{r,\tau}\cdot\delta(\mathbf{x}-\mathbf{x}_r,t-\tau)\left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)-\mathbf{E}^s_{\text{obs}}\right]_{r,\tau}, \tag{2.34}$$

where T represents the adjoint operator. To extract segments of interest from the electric fields, the Dirac delta δ function is used, which is necessary because the electric fields include the combined information from all sources s, receivers r, and observation times τ. To calculate $\mathbf{E}^s_{\text{syn}}$ and minimize the FWI cost function, an accurate forward model that solves Maxwells equations is required. To minimize the cost function and to optimize

the distributions for permittivity and conductivity, corresponding gradients are calculated (more details in Meles et al. (2010)):

$$\begin{bmatrix} \nabla C_\varepsilon(\mathbf{x}') \\ \nabla C_\sigma(\mathbf{x}') \end{bmatrix} = \sum_s \begin{bmatrix} \delta(\mathbf{x}-\mathbf{x}')\partial_t \mathbf{E}^s \\ \delta(\mathbf{x}-\mathbf{x}')\mathbf{E}^s \end{bmatrix}^T \hat{\mathbf{G}}^T \mathbf{R}^s \tag{2.35}$$

with

$$\mathbf{R}^s = \sum_r \sum_\tau [\Delta \mathbf{E}^s]_{d,\tau} = \sum_r \sum_\tau \delta(\mathbf{x}-\mathbf{x}_r, t-\tau) \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma) - \mathbf{E}^s_{\text{obs}}\right]_{r,\tau}. \tag{2.36}$$

A cross-correlation at all locations of the medium for all time steps and all transmitter-receiver combinations of the incident wavefield $\mathbf{E}^s_{\text{syn}}$ of each transmitter, with the residual wavefield $\mathbf{R}^s$ back-propagated by all receivers, yields the gradients (for further details Meles et al. (2010)). These gradients indicate the location and direction of the update of the permittivity and conductivity distributions to reduce the cost function C. To update the permittivity and conductivity, two step lengths ζ_ε and ζ_σ are needed, respectively. This enables the updating of both distributions in each iteration at the same time:

$$\varepsilon_{\text{upd}} = \varepsilon - \zeta_\varepsilon \nabla C_\varepsilon, \tag{2.37}$$

and

$$\sigma_{\text{upd}} = \sigma - \zeta_\sigma \nabla C_\sigma. \tag{2.38}$$

with the step lengths ζ_ε and ζ_σ:

$$\zeta_\varepsilon = \frac{\sum_s \sum_r \sum_\tau \left[\mathbf{E}^s_{\text{syn}}(\varepsilon+\kappa_\varepsilon \nabla C_\varepsilon,\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]^T_{r,\tau} \delta(\mathbf{x}-\mathbf{x}_r,t-\tau) \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma) - \mathbf{E}^s_{\text{obs}}\right]_{r,\tau}}{\sum_s \sum_r \sum_\tau \left[\mathbf{E}^s_{\text{syn}}(\varepsilon+\kappa_\varepsilon \nabla C_\varepsilon,\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]^T_{r,\tau} \delta(\mathbf{x}-\mathbf{x}_r,t-\tau) \left[\mathbf{E}^s_{\text{syn}}(\varepsilon+\kappa_\varepsilon \nabla C_\varepsilon,\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]_{r,\tau}} \tag{2.39}$$

and

$$\zeta_\sigma = \frac{\sum_s \sum_r \sum_\tau \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma+\kappa_\sigma \nabla C_\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]^T_{r,\tau} \delta(\mathbf{x}-\mathbf{x}_r,t-\tau) \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma) - \mathbf{E}^s_{\text{obs}}\right]_{r,\tau}}{\sum_s \sum_r \sum_\tau \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma+\kappa_\sigma \nabla C_\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]^T_{r,\tau} \delta(\mathbf{x}-\mathbf{x}_r,t-\tau) \left[\mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma+\kappa_\sigma \nabla C_\sigma) - \mathbf{E}^s_{\text{syn}}(\varepsilon,\sigma)\right]_{r,\tau}}. \tag{2.40}$$

The gradient calculations and step-length calculations depend on the number of transmitter and receiver locations used, and therefore for each new dataset $\mathbf{E}^s_{\text{obs}}$ with different antenna

numbers, new perturbation factors κ_ε and κ_σ have to be selected and adapted with great care. Yang et al. (2014) introduced an update for the gradient calculation, which allowed the use of similar perturbation factors for different acquisition setups. Because most of the current experimental data investigations used similar frequency ranges between 50 MHz and 120 MHz (center frequency of the data), no update of the perturbation factors was needed. To perform a parameter update in the vicinity of the antennas and to avoid a muting zone, a gradient precondition based on van der Kruk et al. (2015) is applied. For this gradient precondition, additional stabilization factors for permittivity and conductivity are required.

Overall, to judge the final FWI results, four criteria should be fulfilled, which were defined especially for experimental data applications (Klotzsche et al., 2019b). First, the root-mean-square error (RMS) must converge smoothly. As a convergence criterion, a change of the RMS of less than 0.5% between the single iterations is chosen. Second, the final RMS should be at least 50% lower than the starting model based on the ray-based inversion. Third, the gradient of the final iteration should be sufficiently attenuated. Finally, the correlation coefficient between the measured data and the data obtained with the models of the final iteration should be higher than 0.8. If the half-wavelength criterion is not satisfied, at least one of the previously mentioned criteria are often also not satisfied, which makes the FWI results unreliable.

2.3.6. Computational requirements

The computation of the forward model with the FDTD forms the main part of the computational requirements of the FWI, this short overview is based on Klotzsche (2013). During each iteration of the FWI, the forward problem has to be solved four times. The first solution is used to compute the residual wavefield, the second solution is necessary to calculate the model update directions (gradients), and two FDTD calculations are required to determine the step lengths. The calculations for each transmitter position are performed independently, allowing a parallelization with the message passing interface (MPI). A master is responsible for coordinating communication and calculating gradients and model updates, while every transmitter requires a slave. The overhead for distributing the calculations is approximately 10% when using the MPI system (Ernst et al., 2007b). The forward modeled

field $\mathbf{E}^s$ from the first forward modeling remains in the computer memory till the end of the second forward modeling. The computation time T_{comp} can be calculated with:

$$T_{\text{comp}} = 4 \cdot 1.1 \cdot T_{\text{forward}} \cdot N_{\text{iter}}, \tag{2.41}$$

where 4 represents the four calculations of the FDTD within one FWI iteration, T_{forward} the time for a single FDTD calculation including the build of the grid, and N_{iter} the number of iterations (Meles et al., 2010). Note that T_{forward} changes drastically between 2D and 3D, scaling with the additional cells in the extra dimension. Most of the full-waveform calculations are conducted on JURECA-DC, a modular supercomputer at Jülich Supercomputing Centre (Jülich Supercomputing Centre, 2018).

3. Next generation 2.5D GPR FWI including borehole and antenna models[1]

In previous studies, e.g., Mozaffari et al. (2022), it was only possible to perform a standalone 3D forward simulation to investigate borehole and antenna coupling effects. With new developments in the forward solver gprMax and increased computational resources, it is now possible to implement boreholes, as well as antennas in the forward modeling of the GPR FWI. This is possible by using subgrids. Note here, that with regular FDTD approaches the necessary computation time would be too high to use these geometries in the FWI, as for the forward and backward propagation and the step length calculation four forward models have to be simulated within one inversion step (Klotzsche et al., 2019b). Using the new 2.5D GPR FWI we investigate the benefits of this novel approach by applying it to a synthetic aquifer dataset with water filled boreholes, finite-length antenna models and high angles between transmitter and receiver. Thereby, especially the enhancements of the source wavelet estimation and the influence on the reconstruction resolution play an important role.

3.1. New 2.5D GPR FWI

Since the first developments of 2.5D GPR FWI a new version of gprMax, written in Python with computational intensive parts in Cython, as well as support for graphics processing units (GPUs), was released. To use this new gprMax as a forward model in our existing GPR FWI, we had to rewrite this code from C++ to Python. Note that with the 2.5D GPR FWI

[1] adapted from D. Hoven, C. Warren, J. van der Kruk, H. Vereecken, and A. Klotzsche (in preparation). "Including the effects of borehole fillings and finite-length antenna models in 2.5D crosshole ground penetrating radar full-waveform inversion". Planned for publication in GEOPHYSICS.

the 3D-to-2D correction of measured data is not necessary anymore (marked in blue in Figure 3.1). We extend the regular 2D models of dielectric relative permittivity and electrical conductivity for a customizable number of cells in the perpendicular direction and save the models utilizing HDF5 (The HDF Group et al., 2020) in a way, that the model can be directly loaded into gprMax. Note that in the following dielectric permittivity and electrical conductivity will be referred to as permittivity and conductivity, respectively. Furthermore, we allowed the definition of additional forward modeling mediums and structures, that will be added to the plain 2.5D model in every forward simulation. This can be, e.g., a borehole, or even a finite-length antenna model. For the latter, we also included an option for automatically placing antennas at every transmitter and receiver position. When using finite-length antenna models in the 2.5D case, it is important to note that it is not possible to calculate all receivers with one transmitter simultaneously because of the overlapping antenna geometries. To prevent this issue, a clustering technique was implemented where the receiver antennas are arranged in a way that the next antenna is outside the near-field of its nearest antenna. This means that only every N_{cluster}-th antenna is simulated, resulting in N_{cluster} times as many simulations.

The new version of gprMax, includes a Huygens subgridding approach (Hartley et al., 2018), which allows a local refinement of the model space. We added the subgridding to the flowchart in Figure 3.1. However, this means that it is necessary to check whether a transmitter, receiver, or geometry is in the main or in the subgrid and place it accordingly. Furthermore, the model of the permittivity and conductivity has to be added separately to the subgrid. Note that the subgrid areas are currently not included in the calculation of the gradients in the FWI, as they are placed tight around the transmitter and receiver, allowing a transmitter-receiver muting.

In the regular 2.5D case, the computation time is roughly scaling with the number of cells in the perpendicular direction. When using subgrids, the calculation increases according to the computation time of each subgrid, as they do not replace the main grid in that area, but are added to it. If clustering is applied, the regular computation time must be multiplied with N_{cluster}. Note that the effective computation time, especially when finite-length antennas are included and clustering is activated, depends heavily on the available computational resources.

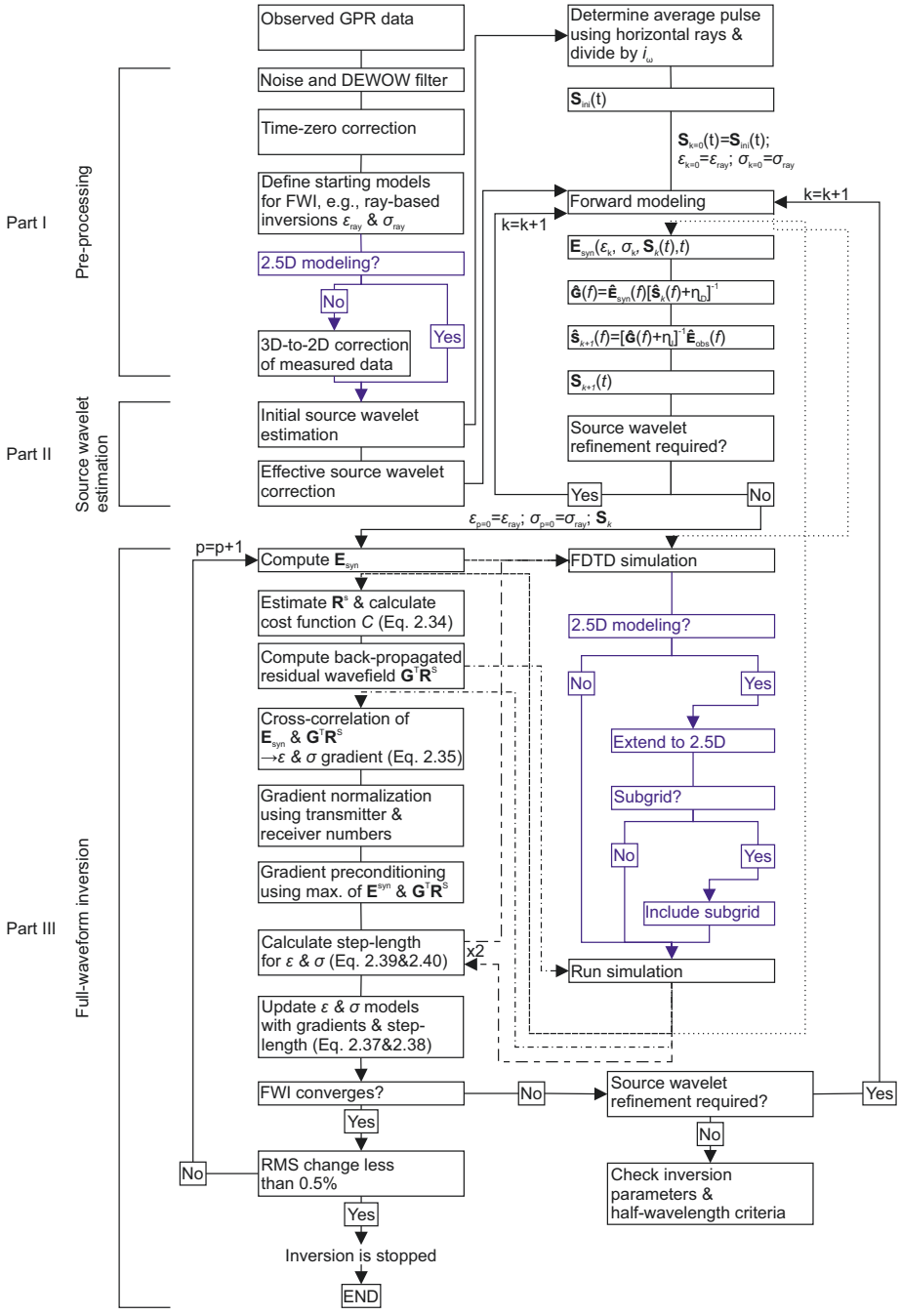

Figure 3.1.: New flowchart of the 2.5D GPR FWI inversion steps with subgridding. Based on the 2D GPR FWI flowchart by Klotzsche et al. (2019b). The changes are marked in blue.

3.2. Realistic synthetic model

To demonstrate the potential and improvement of the novel 2.5D FWI , we defined a realistic synthetic model based on 2D GPR FWI crosshole results from the well-defined Widen test site in Switzerland (Klotzsche et al., 2012). To investigate the impact of high angle data, the original model was stretched to a size of 4.95 m x 13.68 m, and, uses a cell size of 9 cm for the inversion and 3 cm for the standard forward modeling (Figure 3.2a). The upper 1.98 m of the model simulates the unsaturated zone of an aquifer with a homogeneous layer ($\varepsilon_r = 5$; $\sigma = 12$ mS/m). Below this, a smoothed layer with a width of 0.36 m exists to mimic the capillary fringe, followed by the stretched and compressed inversion results from Klotzsche et al. (2012) illustrating a saturated aquifer. Boreholes are positioned at a distance of 0.9 m and 4.05 m and extend to a depth of 12.6 m.

To increase the complexity of the model, we added three features to the models similar to those in Meles et al. (2010). The first one is a rectangular shaped structure in the middle of the domain with a dimension of 1.35 m x 0.45 m ($\varepsilon_r = 20$; $\sigma = 10$ mS/m). Between 9 m and 10.8 m depth, we added triangular structures extending from each borehole towards the inner medium. The third structure is a rising layer ($\varepsilon_r = 18$; $\sigma = 25$ mS/m) that begins at the left borehole at a depth of 11.34 m and ends in the right borehole at 10.8 m depth. Compared to their surrounding areas, these structures indicate high-high, high-low, and low-high combinations of relative permittivity and electrical conductivity. Note that we do this, to investigate if the FWI can reconstruct specific high contrast permittivity and conductivity combinations better than others.

To realistically model crosshole GPR data, we used in total 46 Tx and 222 Rx in a semi-reciprocal setup with each 23 Tx and 111 Rx per side, starting at a distance of 0.9 m or 4.05 m. The antenna locations start at a depth of 2.43 m and stop at a depth of 12.33 m, creating angles up to 72.35°. The spacing of Tx and Rx are 0.45 m and 0.09 m, respectively. With this setup we are able to investigate high angles up to 72.35° between the transmitters and receivers.

In a first step, for the transition of 2D to 2.5D, we extend the forward model in the perpendicular direction (y) for 100 cells (Figure 3.2b). Thereby, the Tx and Rx are located in the middle of the perpendicular direction. In a second step, we added at the locations of the Tx and Rx boreholes with a diameter of 12 cm in the 2.5D model to investigate the borehole filling effects. These are filled with air ($\varepsilon_r = 1$; $\sigma = 0$ mS/m) in the unsaturated

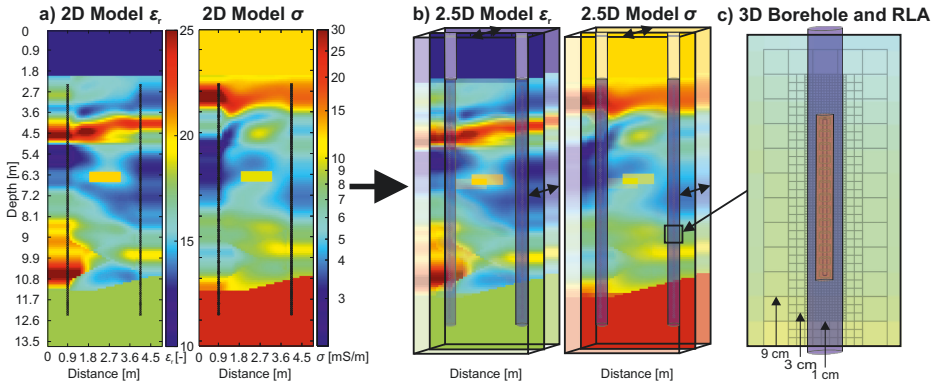

Figure 3.2.: Synthetic model setup for 2D and 2.5D FWI. a) ε_r and σ input model defined in a 2D plane, adapted from the Widen dataset (Klotzsche et al., 2012). b) Medium properties are extended in the perpendicular direction resulting in 2.5D models. Transmitter and receiver locations are indicated with circles and crosses, respectively. c) 3D model including circular 3D boreholes filled with air in the unsaturated zone and water in the saturated zone. 3D cylindrical resistive loaded finite-length antennas (Mozaffari et al., 2022) are present inside the water-filled borehole. The different modeling cell sizes are shown in c).

zone from 0 m to 1.98 m and with water ($\varepsilon_r = 80$ (Jol, 2009); $\sigma = 50$ mS/m (Coscia et al., 2011)) in the saturated zone below 1.98 m.

Additionally, subgrids with a grid size of 1 cm are used for the boreholes, reducing staircasing errors, and to add resistive-loaded finite-length antennas (RLA) for the Tx and Rx (illustrated in Figure 3.2c). We use a resistive-loaded finite-length antennas similar to Mozaffari et al. (2022), which is based on the Sensors and Software crosshole 200 MHz PulseEKKO system. The antenna model, which has a total length of 0.6 m, contains ten 1 cm long resistor elements on each side of the feeding point with a 1 cm gap between them (see Figure 3.3). The feeding point is located 25 cm from the left end of the antenna. Symmetrically extending from the feeding point, the cable (perfect electric conductor) has a length of 24 cm on both sides. Both resistor elements ($\varepsilon_r = 20$; $\sigma = 0.1$ mS/m) and the cable have a radius of 2 cm and are embedded in the insulator ($\varepsilon_r = 4$; $\sigma = 1e-7$ mS/m) with a radius of 3 cm. To allow a better comparison between the point source and the finite length antennas effects, we decreased the radius of the boreholes to 6 cm for the point sources such that the medium around the different sources is the same.

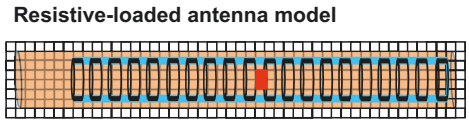

Figure 3.3.: Resistive-loaded finite length antenna model similar to Mozaffari et al. (2022), which is based on the Sensors and Software crosshole 200 MHz PulseEKKO system. The resistor is colored in orange, the cable in blue, and the feeding point in red. Note the size of one cell equals 1 cm x 1 cm.

3.3. GPR full-waveform inversion using different forward model approaches

To analyze the impact of increased complexity in the forward model on FWI reconstruction, we generated 'realistic' input GPR data using the full 2.5D forward model, which includes subgrids with the air and water filled boreholes and antenna models (see Figure 3.2c). Using the full 3D modeled crosshole GPR data, we applied four different inversion scenarios with increased complexity in each step:

- 2D FWI
- 2.5D FWI
- 2.5D FWI with borehole fillings in the subgrids (marked as 2.5D FWI BH)
- 2.5D FWI with borehole fillings and antennas in the subgrids (marked as 2.5D FWI BH A)

As Irving and Knight (2005) showed, we need to be careful using high ray-path angle data in the inversion. We can see a similar behavior of our realistically modelled 3D data as Irving and Knight (2005), where we detect an increased apparent velocity with higher ray-paths angle between the Tx and Rx (Figure 3.4). Therefore, we use in the different inversion data until a ray-path angle of 45° and the full dataset with ray-path angles until 72.35°. Note that we have 5106 traces for the full dataset, while we have only 2628 for the 45° dataset. The use of first break picks of ray-path angles higher as 45° is expected to cause numerical artefacts in the ray-based tomograms, while the effect on the FWI has not been investigated in much detail so far.

In a first pre-processing step, the 3D data are transformed to 2D using the Bleistein filter (Equation (2.30)) to be used in the 2D FWI. Note that the 3D and 2D datasets are normalized to its global maximum to avoid large numbers within the gradient calculation

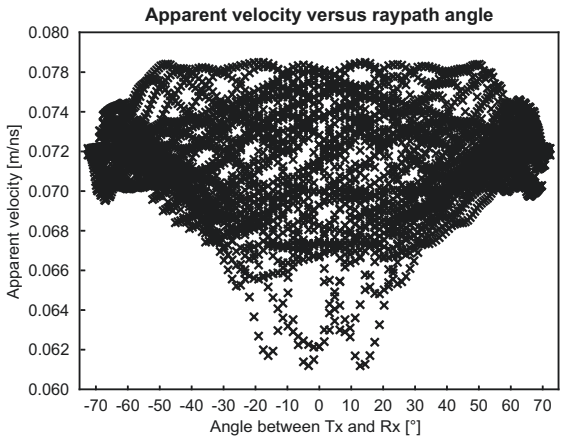

Figure 3.4.: Apparent velocity versus raypath angle for the synthetic data generated with the 3D synthetic aquifer model with water-filled boreholes and finite-length antenna models for transmitter (Tx) and receiver (Rx).

with the use of perturbation factors (necessary for the step length calculation in Equation (2.39) and (2.40)). Following the typical pre-processing steps for experimental data (see flowchart in Figure 2.2), we first define the starting models and then the effective source wavelet for each of this inversion.

3.3.1. Ray-based inversion

As stated in Chapter 1, good starting models for the FWI are mandatory and in most cases provided by the ray-based inversion results. Therefore, we performed the travel time and the first-cycle amplitude inversion (see Section 2.3.2.1) using data until a ray-path angle of 45° and the full dataset with ray-path angles until 72.35° (Figure 3.5).

The first-arrival times and first-cycle amplitudes of the realistic dataset are picked with an automatic routine. The starting model for the 45° angle dataset ray-based inversion was a homogeneous model with a velocity of 65 m/μs, while we used 70 m/μs for the full dataset. In both, a damping and smoothing factor of 1 was used.

The travel time inversion yields a clear layered reconstruction for the 45° dataset (Figure 3.5b). The high permittivity layer around 4 m depth also shows a small rise from the left to the right borehole. Additionally, we can see a slightly higher permittivity for the rectangular object, which is surrounded by low permittivity layers. The RMS decreased

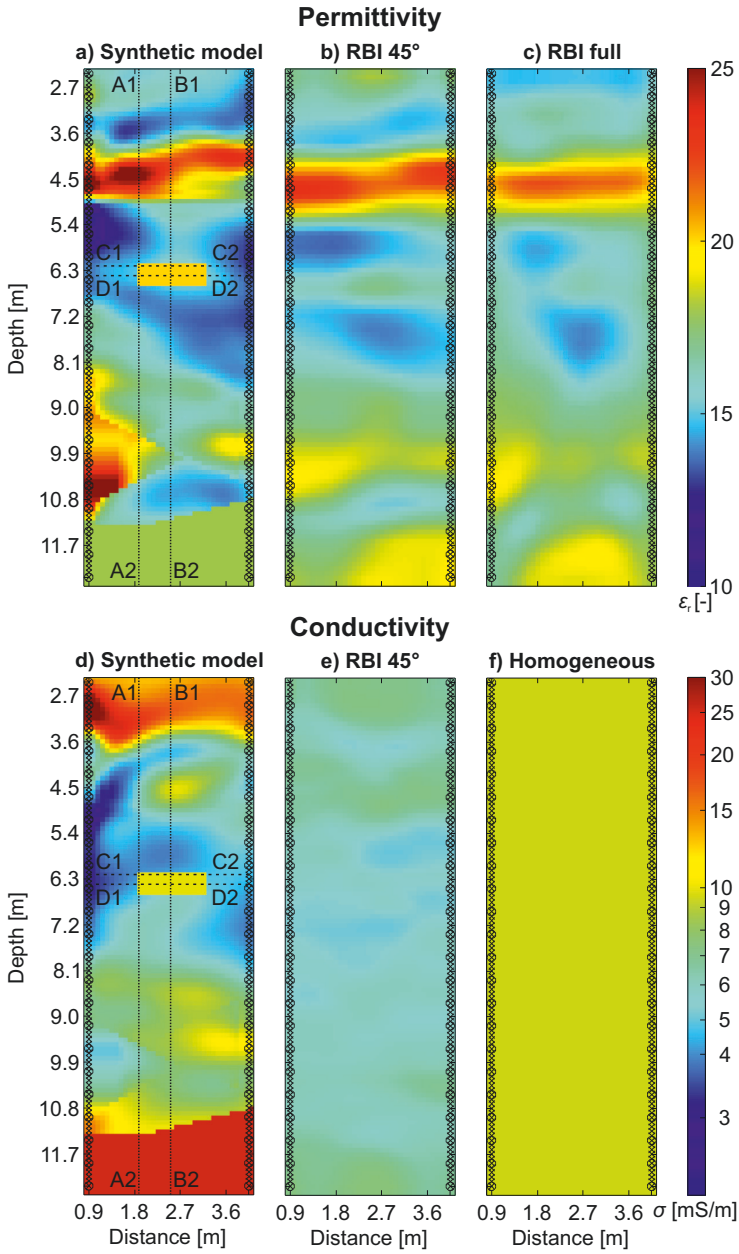

Figure 3.5.: Comparison of the RBI results (travel time and first-cycle amplitude inversion) for the 45° and full dataset with the synthetic model. Lines in the synthetic model indicate the vertical (A1-A2, B1-B2) and horizontal sections (C1-C2, D1-D2) that will be used to investigate the reconstruction through the rectangular object in more detail. Transmitter and receiver locations are indicated with circles and crosses, respectively.

from 6.42 ns to 0.36 ns. The results for the full angle dataset (Figure 3.5c), show only a straight high permittivity layer. In the low permittivity area between 4.8 m and 8.5 m only two smaller areas are visible. Furthermore, the low permittivity around 10.8 m is divided by a higher permittivity area into a left and right half. The RMS decreased from 4.19 ns to 0.73 ns. Overall, the structures are better resolved with the travel time inversion for the 45° dataset. Therefore, we used these as a starting model for permittivity for the FWIs. Note that we only have first-cycle amplitude inversion results for the 45° dataset (Figure 3.5e), as the inversion failed for the full dataset. Since the first-cycle amplitude inversions provide a relatively smooth image, we use a homogeneous conductivity model with a value of 9.5 mS/m (Figure 3.5f) for the different angle FWIs. Such an approach is similar to most of the experimental data applications (e.g., Klotzsche et al., 2012) and showed best robustness.

3.3.2. Effective source wavelets

The source wavelet is crucial for the forward modeling of electromagnetic wave propagation. While the source wavelet is known for most synthetic studies it is especially for experimental data unknown. Therefore, we estimate an effective source wavelet for each of our four different FWIs using the deconvolution approach (Section 2.3.4). We generate synthetic data using our ray-based results as the input model. Note that the forward modeling (Figure 2.2) uses the same setup as in the corresponding FWI. Therefore, in the forward model with the included finite length antenna model, the effective source wavelet for the antenna is estimated.

As the 3D and 2D dataset are normalized and the radiations in 3D and 2D are different, a clear difference in the amplitude of the 2D effective source wavelet and the 2.5D wavelets can already be seen (Figure 3.6a and b). Furthermore, some of the amplitude differences between 2D FWI and 2.5D FWI are based on different implementations in the FWI forward model. Additionally, we can observe an increase in amplitude for the 2.5D FWI wavelets when we include boreholes and the antennas. For a better visual comparison, we normalize the wavelets to their maximum (Figure 3.6c and d). We observe that the 2D FWI source wavelet starts later than the original source wavelet and also has stronger later amplitudes. Furthermore, its center frequency (80.9 MHz) is lower compared to the original source wavelet (86.1 MHz). The 2.5D FWI is similar in both the time and frequency domain (80.9 MHz) to the 2D FWI source wavelet. Although it looks as if the source wavelet for the 2.5D FWI BH is closer to the input source wavelet in the time-domain, a significantly lower

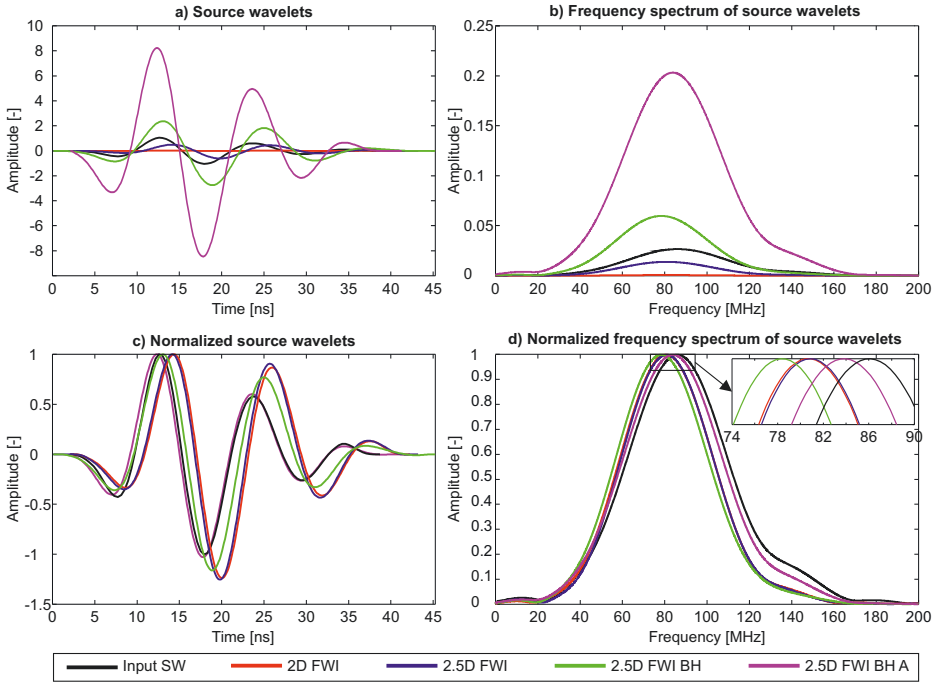

Figure 3.6.: Estimated effective source wavelets in a) time-domain and b) frequency-domain. Same wavelets normalized to 1 in c) and d) for better comparison, respectively. Note the detailed section in d) for a better visualization of the center frequencies. The input source wavelet is shown in black. The source wavelets for 2D FWI, 2.5D FWI, 2.5D FWI with borehole (BH), and 2.5D FWI with borehole and antenna (BH A) have been estimated using the deconvolution approach and are shown in red, blue, green, and purple, respectively.

center frequency (78.3 MHz) can be noticed. We assume, that the water filled borehole acts as a frequency filter. The estimated source wavelet for the 2.5D FWI BH A is the closest to the input source wavelet in shape and also in center frequency (84 MHz). This comparison demonstrated that with increasing complexity of the forward model, the effective source wavelet is closer to the real input wavelet of the antenna and only minor differences can be observed. These observations support the hypothesis that the more complex the forward model, the closer the effective source wavelet is to the real input wavelet of the antenna.

3.3.3. FWI of the different approaches

Using the RBI starting models and the corresponding source wavelet for each FWI approach, the different permittivity results for the 45° dataset and the full angle dataset are shown in Figure 3.7. In contrast to the RBI, the FWI approaches yield very similar results for both datasets, regardless the common assumption, that the full datasets may cause artefacts related to the increased apparent velocity. Note that all FWI results fulfilled the four criteria mentioned in Klotzsche et al. (2019b). We saw a proper reduction in the RMS and no remaining gradients in the final iteration of the FWI approaches. In contrast to the RBI, all the FWI results provide more details in the tomograms and structures in the decimeter-scale are resolved. We observed similar structures and an overall good fit to the input model for all FWI approaches. Having a closer look at the details in the models, we can observe that the FWI approaches utilizing more complex forward models, are able to enhance these reconstructions, e.g., between 3.6 m and 10.8 m depth, particularly at the rectangular object at 6.3 m depth and the triangular object between 9 m and 10.8 m depth. Please note here, that errors related to a low data coverage in the top and the bottom of the model increase with the more complex FWIs. In addition, we can see vertical artefacts in the area of the boreholes, probably caused as a result of the absence of gradient normalization (van der Kruk et al., 2015), which is not yet implemented for the 2.5D FWI.

While we see only minor improvements in the permittivity tomograms of the different FWI approaches, we can notice more effects on the electrical conductivity reconstructions in Figure 3.8. We see that the higher conductivity at a depth of 2.7 m and a distance of 0.9 m was not reconstructed with the 2D FWI, but only using the different 2.5D FWI approaches. The 2.5D FWI approach using boreholes (BH) most accurately determined the localization of the lens at a depth of 4.5 m and a distance of 2.7 m. The simple 2.5D FWI tends to overshoot between 5.4 m and 8.1 m depth, an effect that was reduced with the more complex 2.5D FWI. Furthermore, compared to the 2D FWI, the 2.5D FWI reconstructed the rectangular structure more clearly. The 2D FWI and 2.5D FWI yielded very similar results in the range between 8.1 m and 10.8 m depth, as did the 2.5D FWI BH and 2.5D BH A. A possible explanation for this is the influence of the borehole integration in the forward model. Note that contray to our expectations, the 2.5D FWI BH A shows more difficulties especially in the lower parts of the models compared to the 2.D FWI BH. The reconstruction results with the full dataset were similar to those of the reduced dataset, without any additional anomalies. As noted in the permittivity results, artefacts were present in the upper and lower areas of the reconstructions because of the limited data coverage.

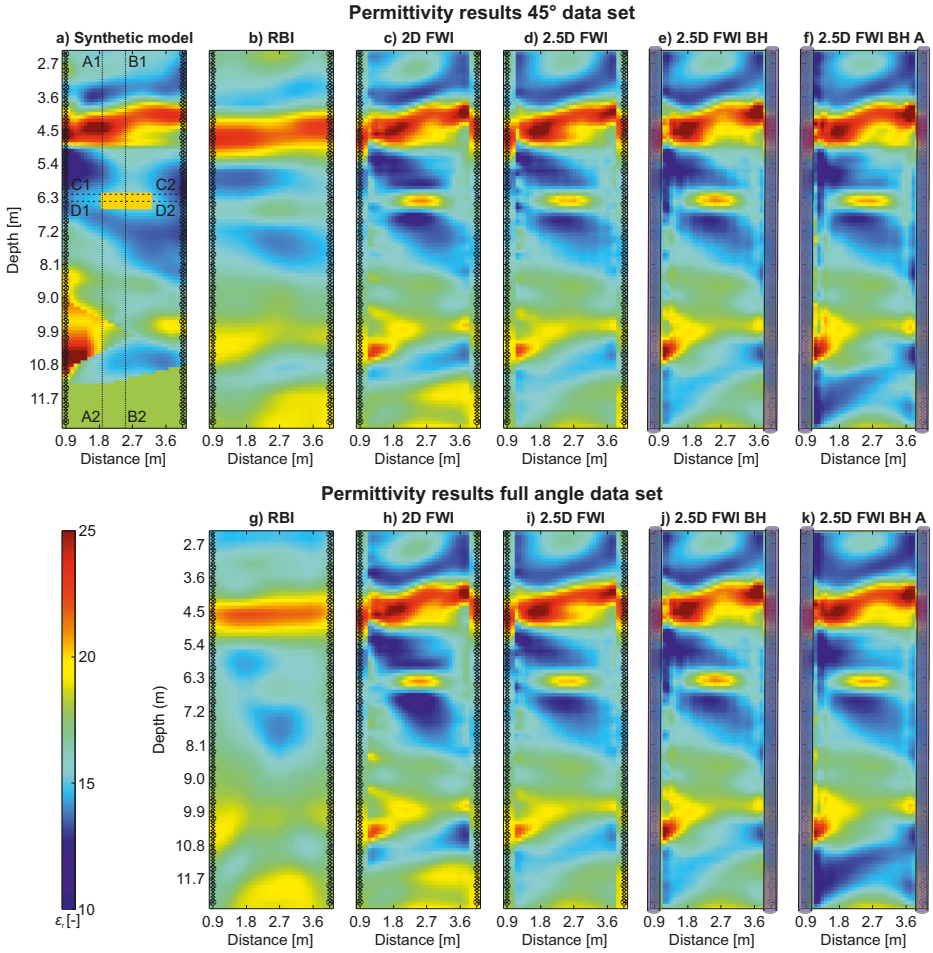

Figure 3.7.: Overview of the different FWI permittivity results: a) True model, b) RBI, c) 2D FWI, d) 2.5D FWI, e) 2.5D FWI with borehole (BH), and f) 2.5D FWI with borehole and antenna (BH A) for the 45° dataset, whereas g), h), i), j) and k) show the results for the full angle dataset, respectively. In all tomograms, the transmitter and receiver locations are indicated with circles and crosses, respectively, and boreholes are indicated by blue cylinders. Lines in the synthetic model indicate the vertical (A1-A2, B1-B2) and horizontal sections (C1-C2, D1-D2) that will be used to investigate the reconstruction through the rectangular object in more detail.

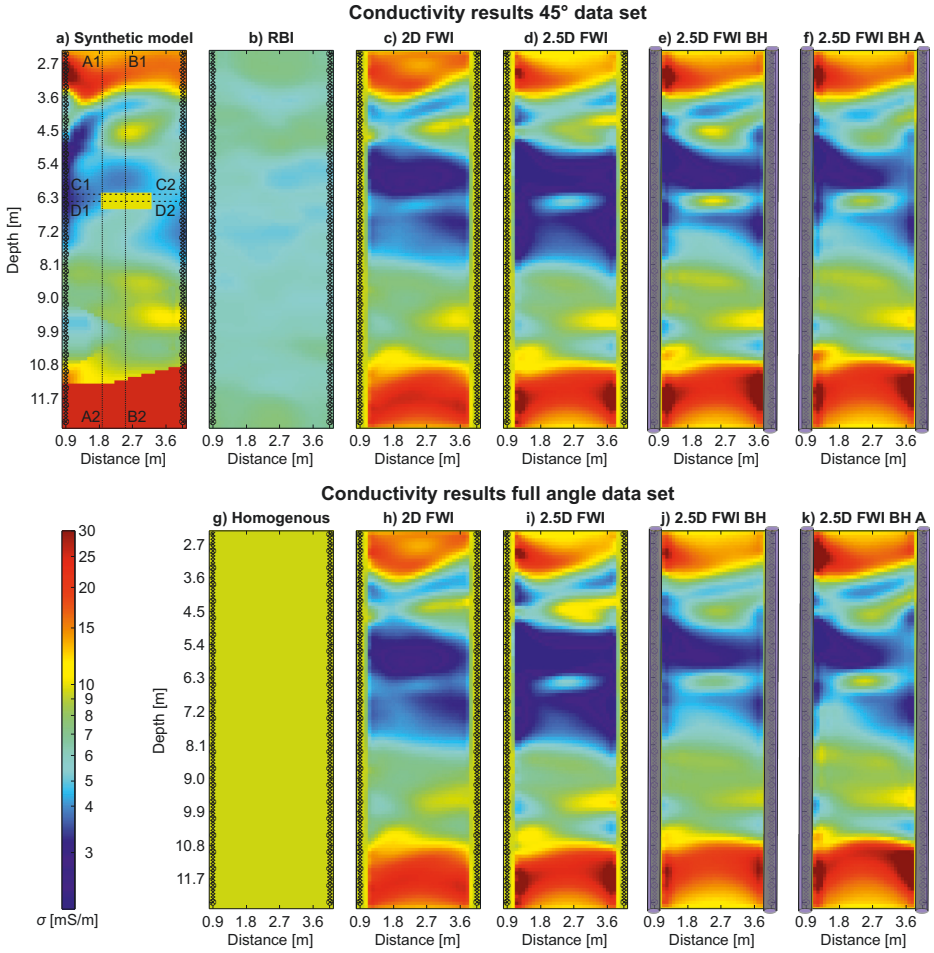

Figure 3.8.: Overview of the different FWI conductivity results: a) True model, b) RBI, c) 2D FWI, d) 2.5D FWI, e) 2.5D FWI with borehole (BH) and f) 2.5D FWI with borehole and antenna (BH A) for the 45° dataset, while h), i), j) and k) show the results for the full angle dataset, respectively. Note that the RBI failed for the full dataset, instead we show the homogeneous model in g) that was used for all FWI for both datasets. In all tomograms, the transmitter and receiver locations are indicated with circles and crosses, respectively, and boreholes are indicated by blue cylinders. Note the logarithmic scales for the conductivity. Lines in the synthetic model indicate the vertical (A1-A2, B1-B2) and horizontal sections (C1-C2, D1-D2) that will be used to investigate the reconstruction through the rectangular object in more detail.

To investigate the reconstruction of the parameters in more detail, we plotted vertical slices through the rectangular object, as illustrated in Figure 3.9. The positions are marked as dotted lines in Figure 3.5. Here, Figure 3.9a and b correspond to the left edge (profile A1-A2) and Figure 3.9c and d to the center (profile B1-B2). In the permittivity results for profile A1-A2, the RBI reconstruction is shifted by about 25 cm in depth. The sections of all FWI reconstructions are very similar. Differences mainly occur in the upper area down to a depth of 3.6 m and in the lower area starting from a depth of 10.8 m, which are the areas with low data coverage. The reconstructions with 2.5D FWI BH and BH A are better than those with 2D FWI and 2.5D FWI, especially in the middle area (see Figure 3.9b). The results in Figure 3.9c are similar to those in Figure 3.9a, with 2D FWI overshooting at the boundaries of the rectangular object. However, the conductivity reconstruction in 2.5D FWI for profile B1-B2 shows a stronger tendency to resolve higher values than 2D FWI. The 2.5D FWI BH A reconstruction shows the least overshoot in the border region of the rectangular object, but has some difficulties in the area around 4.5 m. In general, 2.5D FWI BH and BH A provide the best reconstructions.

The horizontal sections in Figure 3.10 offer a detailed view into the horizontal reconstruction of the rectangular object, with Figure 3.10a and b cutting through the object's upper boundary (profile C1-C2), and Figure 3.10c and d providing insights into the core region (D1-D2). Only a very small shift of the RBI can be seen in the area of the rectangular object. This is probably caused by the offset we observed in the vertical sections. In the permittivity results for profile C1-C2 it is visible that the 2D FWI has the lowest coverage in the width of the object and also provides higher values on the outer sides of the area. The 2.5D FWI and 2.5D FWI BH A give similar results, with the 2.5D FWI BH showing the best fit. While the rectangular object is not visible in the 2D FWI results in the conductivity reconstruction in profile C1-C2, the 2.5D FWI shows this structure even though the conductivities are clearly too low. The 2.5D FWI BH and BH A provide very similar results, with the 2.5D FWI BH giving the best results in Figure 3.10b. The results for the permittivity profile D1-D2 show only small differences between the FWI results, while for the conductivities in Figure 3.10d we can again see a similar behavior as for profile C1-C2, showing the superiority of the 2.5D FWI BH and BH A.

Transitioning from the spatial analysis to the quantitative evaluation of inversion accuracy, we investigate the RMS for the different inversion approaches as illustrated in Figure 3.11. The effective RMS is higher for the inversions with a full dataset because of the larger amount of data used in the calculation, but the RMS decrease is similar for

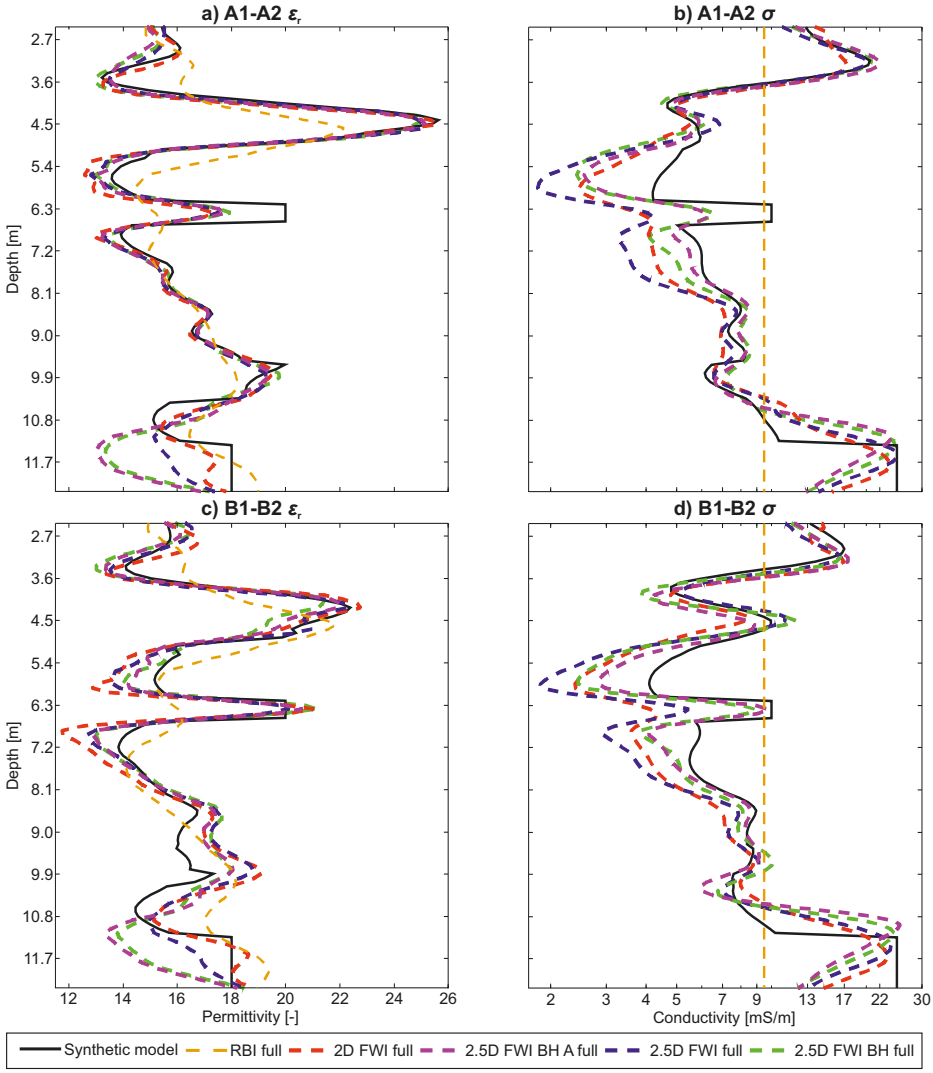

Figure 3.9.: Vertical sections through the rectangular object. a) and b) show a comparison of the permittivity and conductivity results, respectively for a horizontal section A1-A2 (see Figure 2), whereas c) and d) represent the section through B1-B2. The true model is shown in black. The results for RBI, 2D FWI, 2.5D FWI, 2.5D FWI with borehole (BH), and 2.5D FWI with borehole and antenna (BH A) are shown in orange, red, blue, green, and purple, respectively, and marked with a dotted line for the full dataset, respectively. Note the logarithmic scales for the conductivity.

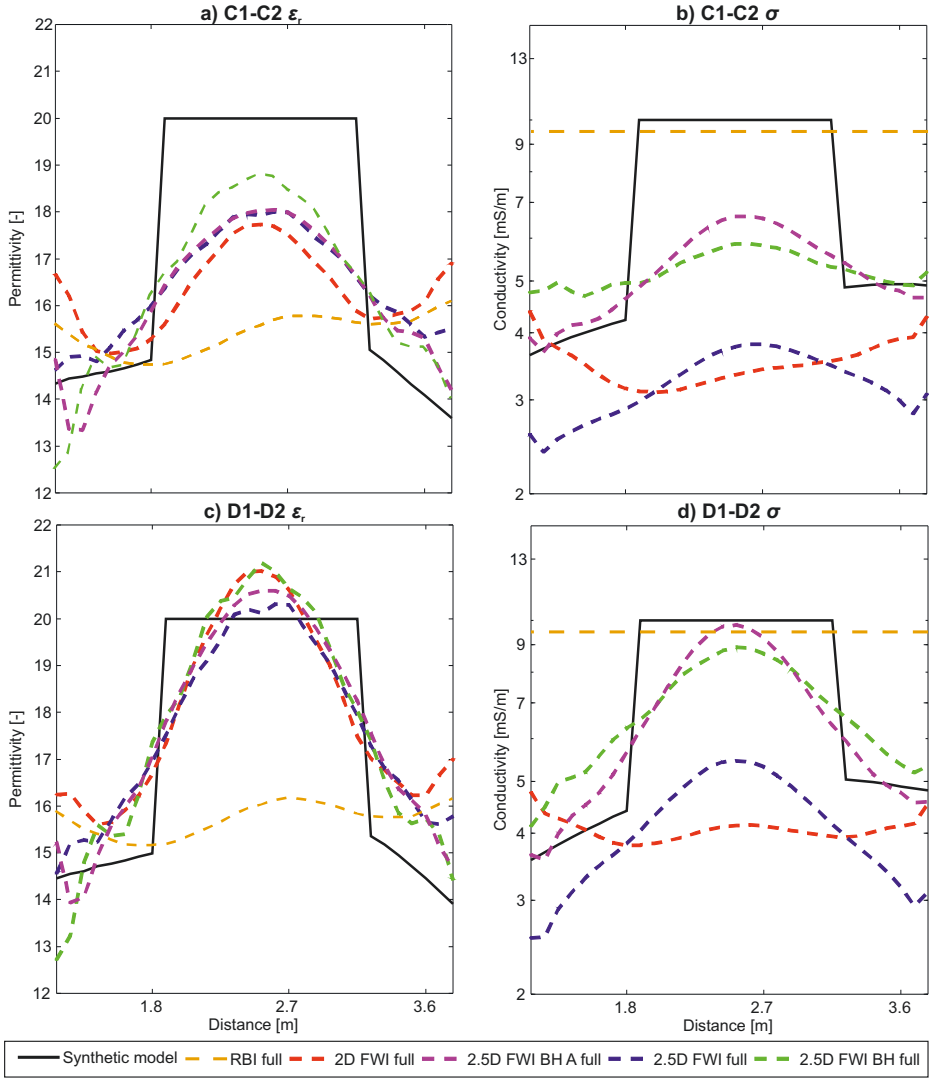

Figure 3.10.: Horizontal sections through the rectangular object. a) and b) show a comparison of the results for a horizontal section through C1-C2 (see Figure 2), whereas c) and d) represent the section through D1-D2. The true model is shown in black. The results for RBI, 2D FWI, 2.5D FWI, 2.5D FWI with borehole (BH), and 2.5D FWI with borehole and antenna (BH A) are shown in orange, red, blue, green, and purple, respectively, and marked with a dotted line for the full dataset, respectively. Note the logarithmic scales for the conductivity.

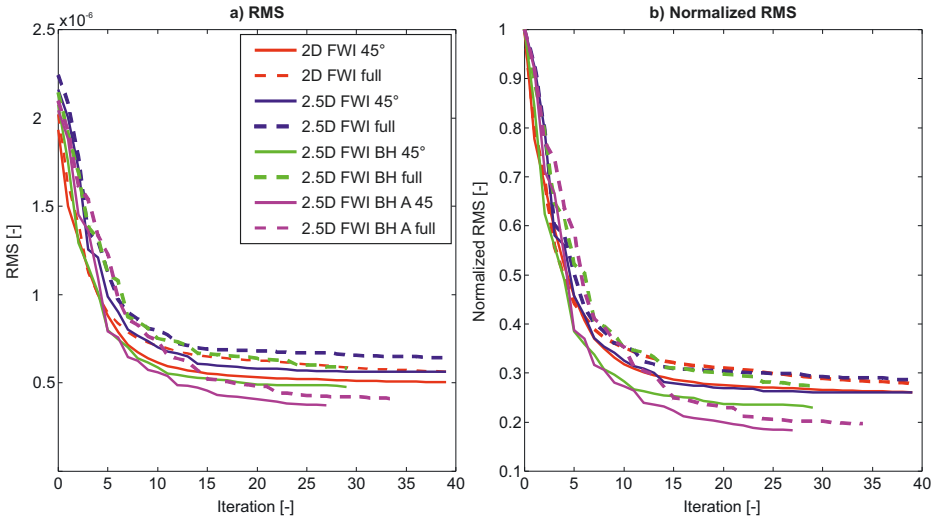

Figure 3.11.: Root Mean Square error (RMS) of the inversion iterations. a) shows the effective root mean square error of the different inversions for every iteration, whereas b) shows the normalized RMS, which is normalized to the RMS with the data generated by using the starting models and the corresponding estimated wavelet. The results for 2D FWI, 2.5D FWI, 2.5D FWI with borehole (BH), and 2.5D FWI with borehole and antenna (BH A) are shown in red, blue, green, and purple, respectively, and marked with a continuous and dotted line for the 45° and full dataset, respectively.

all methods. It appears that 2D FWI is better than 2.5D FWI, but the normalized RMS (Figure 3.11b) shows that the reduction compared to the initial RMS is very similar to 2.5D FWI. The effective RMS difference is partially attributed to the use of the Bleistein transformed dataset to compute the RMS for the 2D FWI. The 2.5D FWI BH A yields the lowest effective RMS. Additionally, the inclusion of borehole (and antenna) results in a significant improvement in RMS reduction. However, it is worth noting that even in the normalized cases, the reduction with the full dataset is not as good as in the dataset with reduced angles.

The final RMS in Table 3.1 shows that the final RMS for the 2D FWI is smaller compared to the 2.5D FWI for both the 45° and full angle dataset. Nevertheless, the RMS reduction is quite similar for both of them. 2.5D FWI BH and BH A show a decrease in the final RMS, compared to the FWI with the less complex forward model. In all of the cases the 2.5D FWI BH A yields the highest RMS reduction. Note here, that the initial

Table 3.1.: Initial versus final RMS with its reduction factor as well as the $\text{MRAE}_{\varepsilon_r}$ and MRAE_{σ} of the final inversion results.

Inversion Method	Initial RMS	Final RMS	RMS Reduction	MRAE ε_r	MRAE σ
2D FWI 45°	1.93e-6	5.25e-7	72.80%	5.03%	24.98%
2.5D FWI 45°	2.16e-6	5.93e-7	72.55%	5.35%	24.56%
2.5D FWI BH 45°	2.07e-6	4.90e-7	76.33%	5.52%	18.42%
2.5D FWI BH A 45°	2.04e-6	3.78e-7	81.47%	6.59%	19.28%
2D FWI full	2.02e-6	6.28e-7	68.91%	5.36%	21.84%
2.5D FWI full	2.24e-6	6.89e-7	69.24%	5.31%	26.20%
2.5D FWI BH full	2.14e-6	5.62e-7	73.74%	5.46%	21.37%
2.5D FWI BH A full	2.09e-6	4.23e-7	79.76%	6.76%	22.11%

RMS increases with the full angle dataset, but the more complex the model, the smaller the increase. Furthermore, the RMS reduction is decreaed with the full dataset, while the effect decreases with more complex approaches, as the 2D FWI reductions is decreased by 3.89% the 2.5D FWI BH A is only decreased by 1.71%. The mean relative absolute error (MRAE) ε_r in contrast is increasing, especially the 2.5D FWI BH A for both the 45° and full angle dataset. This can be attributed to the poorer reconstruction in the area below 10.8 m depth. The MRAE σ shows a significant decrease of nearly 6% if we compare the results of both plain 2D FWI and 2.5D FWI with the 2.5D FWI BH and BH A for the 45° dataset. The MRAE σ for the full dataset show no such trend, and the 2.5D FWI performs poorly independently of the others, which is because of the overshooting in the reconstruction of the area 4.5 m to 8.1 m depth.

The RMS progressions in Figure 3.11 provide us a metric about how good the approaches can reduce the error compared to each other. Nevertheless, it is important to also investigate the RMS for the different FWIs for each transmitter-receiver combination for the final iteration with the full dataset (Figure 3.12). Because of the reciprocal model setup we first show the RMS with the transmitter in the left borehole and then in the right borehole. Here we see that the RMS error distributions are very similar for both the 2D FWI and the 2.5D FWI. In particular, we can see that higher errors occur in the upper part of the model. These are already reduced by the 2.5D FWI BH, while the 2.5D FWI BH A shows negligible RMS errors in this range. This shows the significant improvement in optimization when boreholes and antennas are included in the FWI forward model. Note that for the 2.5D FWI BH A, the errors increase especially in the lower range of the transmitters from about 11 m. This correlates with the higher errors in the lower reconstruction area, which are attributed to problems in the inversion because of a poorer data coverage. The

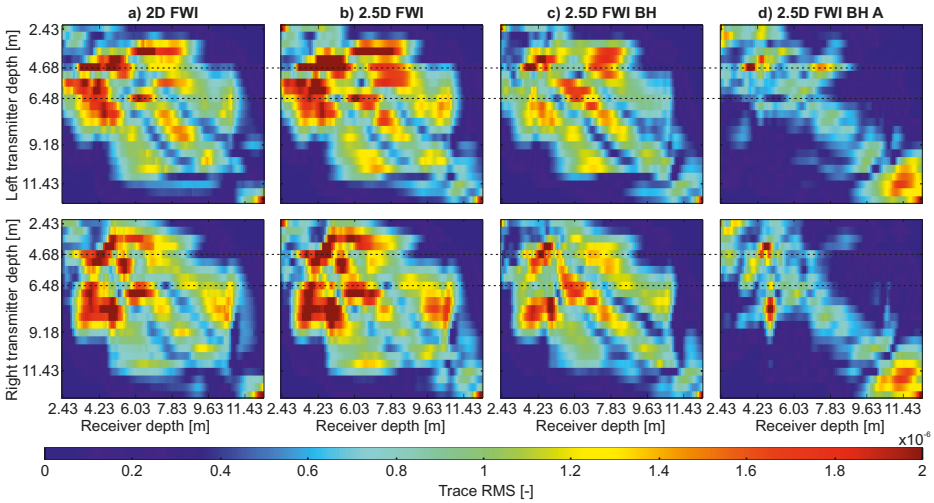

Figure 3.12.: Root mean square error (RMS) for every transmitter receiver combination, in the upper part the transmitter in the left borehole and below in the right, for the full angle data. For the final results of a) 2D FWI, b) 2.5D FWI, c) 2.5D FWI with borehole (BH), and d) 2.5D FWI with borehole and antenna (BH A). Note the dashed lines indicate the positions of transmitter 6 and 10.

described improvements and results can be seen regardless of whether the transmitter is in the left or right borehole, however the RMS is not strictly reciprocal.

It is important to consider the actual datafit (Figure 3.13). The transmitter is positioned in the left borehole and the receiver is in the right borehole. The differences between the synthetic and modeled data for the 2D FWI are smaller than for the 2.5D FWI, for both transmitter depths (4.68 m and 6.48 m). This is because of the Bleistein transformed dataset and the 2D modeling. Note that transmitter 6 is still in the waveguide (a high permittivity layer). Errors in the traces that are located in the waveguide and above are more noticeable in the 2D FWI and 2.5D FWI, but less so in the 2.5D BH and BH A. The differences in the 2.5D FWI are greater in the traces where the receiver is located below the waveguide compared to the 2D FWI. The differences below the waveguide obtained with the 2.5D FWI BH are similar to those obtained with the 2D FWI. The 2.5D FWI BH A has particularly low errors in the traces compared to the other differences, indicating exceptionally good optimization by the inversion. Transmitter 10 is positioned at the lower edge of the rectangular object. This is visible in all the data, particularly from about 65 ns with a receiver depth of about 6.5 m. The 2D FWI shows errors in the traces between 6 m and 8 m receiver

depth, since there is a possible discrepancy between the Bleistein transformed dataset and the 2D simulated data regarding reflections in the rectangular object. These issues are not present in any of the differences for the 2.5D FWI. From a depth of 6 m, the 2D FWI has smaller errors than the 2.5D FWI. Both 2.5D FWI and 2.5D FWI BH produce higher errors between 6 m and 8 m depth. At a depth of 8 m, the errors of the 2.5D FWI are higher compared to the 2D FWI, while the errors of the 2.5D FWI BH are similar to those of the 2D FWI. Similar to the previous results, the 2.5D FWI BH A shows a significant reduction in errors.

The computational demand increased as the FWI methods became more complex. The 2D FWI approach required 1.28 core-hours per iteration. Transitioning to 2.5D FWI increased computation time to 158.72 core-hours per iteration because of the three-dimensional forward modeling, with 100 cells in the y-direction. The 2.5D FWI BH method introduced additional complexity as it used subgrid modeling for borehole integration, requiring 560 core-hours per iteration. The 2.5D FWI BH A method required the highest computational resources, using 5099.52core-hours per iteration. This was caused by a special parallelization strategy that is required to integrate the antenna geometries into the inversion process, since it is not possible to simulate all receivers for a transmitter in one forward model, as the antenna geometries would overlap. It is important to note that subgrids were also utilized in this approach. It is important to point out that these are the total core-hours for an iteration, which are distributed across all nodes and their CPU cores. Note that the core-hour metric is dependent on the supercomputer used. In our case it is the system presented in Jülich Supercomputing Centre, 2018.

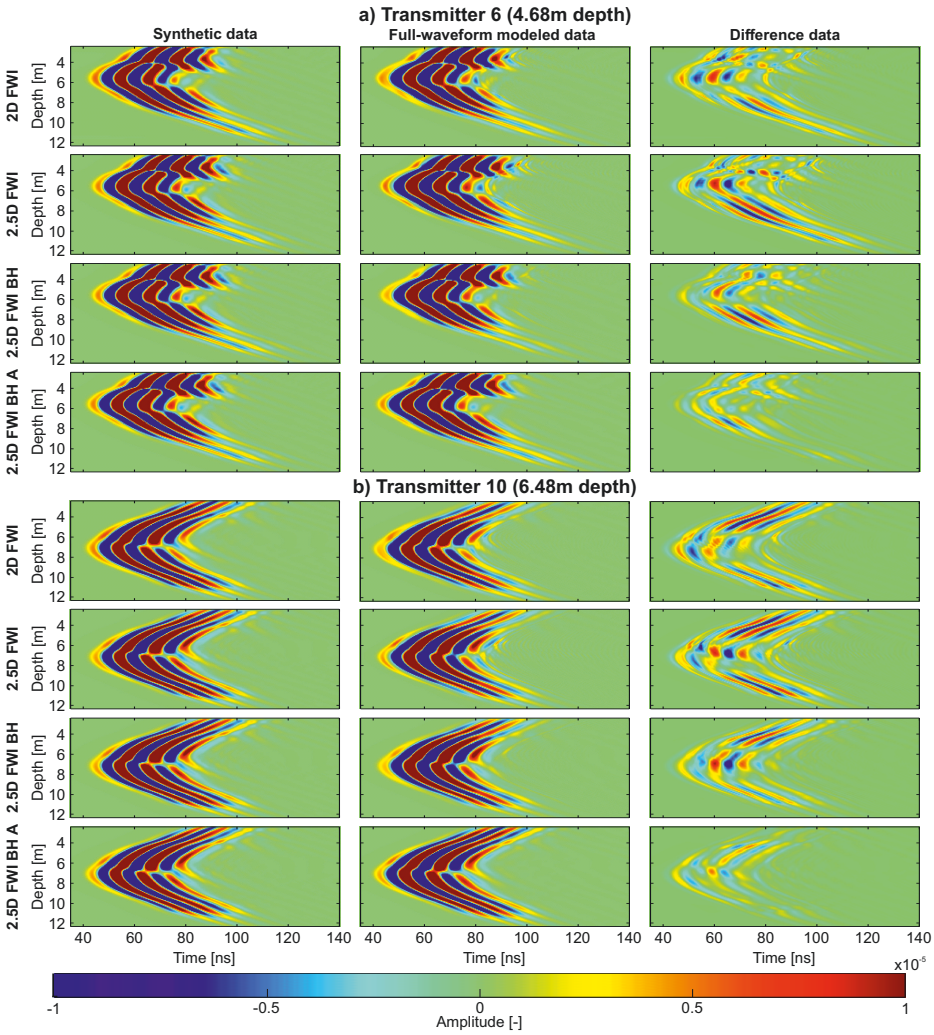

Figure 3.13.: Comparison of the synthetic and final full-waveform modeled data (full angle) with the corresponding differences for a) transmitter 6 and b) transmitter 10 at 4.68 m depth and 6.48 m depth, respectively. Results are shown for 2D FWI, 2.5D FWI, 2.5D FWI with borehole (BH), and 2.5D FWI with borehole and antenna (BH A).

3.4. Conclusions and outlook

In conclusion, our study utilized a synthetic model based on 2D GPR FWI crosshole results from the extensively surveyed Widen test site in Switzerland. We created a synthetic 3D dataset by extending 2D models in the perpendicular dimension and incorporating boreholes and resistive-loaded finite-length antennas to closely mimic realistic antenna radiation characteristics. The source wavelet estimations showed that a very good reconstruction of the wavelet can be achieved, especially with the inclusion of boreholes and antennas. The RBI results are inferior to the FWI results and suffer from artefacts with higher angle data, and we were not able to perform a first-cycle amplitude inversion for high angle data. Our FWI strategies, ranging from 2D to the more advanced 2.5D with boreholes and antennas, showed that incorporating detailed models for boreholes and antennas in the forward model significantly improved the conductivity inversion results. The advanced 2.5D FWI methods were superior, particularly in identifying a rectangular structure and other (high-contrast) structures within the domain, while the simple FWI methods failed to detect high-contrast conductivity features. In addition, the mean relative absolute error of conductivity was reduced by more than 20% with the advanced 2.5D FWI methods compared to simple 2D FWI and 2.5D FWI. The superiority of the complex 2.5D FWIs was further demonstrated through vertical and horizontal sections, indicating their higher spatial reconstruction. However, challenges remain in areas of low data coverage. Furthermore, additional investigation is required for vertical artefacts in the borehole region. We expect to solve this problem by implementing a gradient normalization. The RMS analysis highlights the improved performance of the more complex FWI methods, despite their higher computational requirements. Although the 2.5D FWI BH A yields the best RMS values, its reconstructions are not significantly superior to those of the 2.5D FWI BH. Therefore, we recommend using the 2.5D FWI BH for the time being if high-contrast zones are expected, particularly in conductivity. If the main interest is in permittivities and high-contrast zones are not expected, the 2D FWI may be sufficient as it already delivers significantly better results than the RBI.

In a next step we want to investigate the influence of borehole and antenna inclusion in the forward model on measured data. Especially a dataset with both air and water filled boreholes should benefit from the newly added geometries.

4. Benefits of the 2.5D GPR FWI for variably saturated soil-aquifer system

The application of the common 2D FWI to experimental data measured in variable aquifers is challenging because of the changing borehole filling. Klotzsche et al. (2019c) investigated how the 2D FWI can be applied to such crosshole GPR data acquired at the Widen test site in Switzerland. It was shown that four effective source wavelets have to be estimated to account for the variable borehole filling and antenna coupling effects caused by the different combinations transmitter-receiver combinations in the unsaturated and saturated zone to be able to invert such data. In the previous chapter, it was shown that the effective source wavelet becomes closer to the input source wavelet when borehole filling and finite-length antenna models are included in the forward modeling. Therefore, we will investigate the possibility of applying the new 2.5D FWI with borehole and antenna inclusion to data of a variably saturated soil-aquifer system. The more complex forward model should account for the coupling effects, and therefore only one source wavelet has to be estimated.

4.1. Simulation setup for unsaturated and saturated soil FWI

Based on the previous study by Klotzsche et al. (2019c), we created a schematic simulation setup for the FWI (see Figure 4.1a). According to this study, the transmitter (Tx) and receiver (Rx) were positioned in either the unsaturated zone (5 Tx and 26 Rx) or in the saturated zone (23 Tx and 114 Rx). Our model accounts for the transition between the two zones by moving from an air to a water filled borehole, with the water table located at a depth of 4.2 m, as in the previous study. Note that Klotzsche et al. (2019c) partly

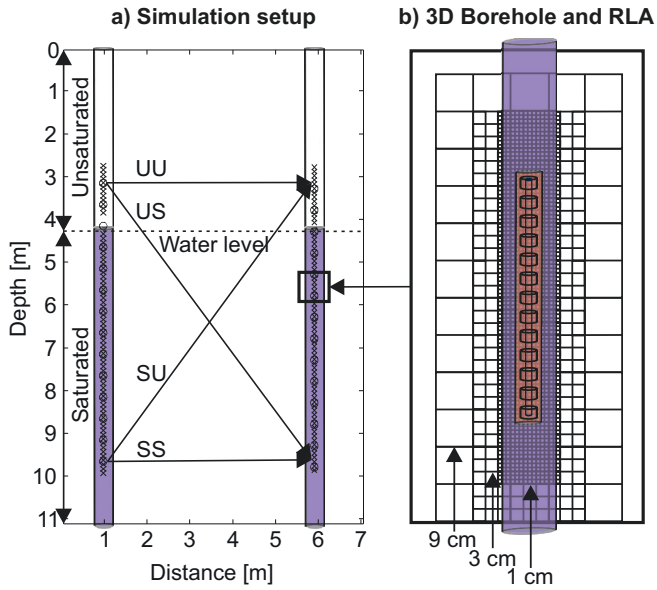

Figure 4.1.: a) Schematic view of the simulation setup based on Klotzsche et al. (2019c). b) 3D borehole and RLA. Tx and Rx locations are based on experimental data and indicated with circles and crosses, respectively. Tx and Rx can be placed in the unsaturated (UU) and saturated zone (SS) respectively. Furthermore, Tx can be placed in the unsaturated zone while Rx in the saturated zone (US) or vice versa (SU).

excluded Tx and Rx in the area of the water table, as the 2D FWI could not account for the half submerged antenna coupling. We plan to conduct a study in the future to investigate whether we can include these Tx and Rx with the method presented in this chapter.

Similar to the setup in Chapter 3, we use a 9 cm inversion grid size, 3 cm for the forward model main grid and a 1 cm forward model subgrid for the boreholes and antennas (Figure 4.1b). As the monitoring boreholes at the test site have a diameter of 11.4 cm, we approximated the boreholes with a diameter of 12 cm. Note that Klotzsche et al. (2019c) measured the data with the RAMAC Ground Vision system of Malå Geosience with 250 MHz antennas. As developing a new antenna model is time consuming, we tested in a first attempt the antenna model used in Chapter 3, which is a resistive-loaded finite-length antenna similar to a design used by Mozaffari et al. (2022), which is based on the Sensors and Software crosshole 200 MHz PulseEKKO system. We therefore can not expect perfectly fitted data, but can hope to achieve better results than with a point source.

In the study by Klotzsche et al. (2019c), the UU combination had the highest center frequency at 123 MHz, while the SS combination had the lowest at 85 MHz. The US and SU combinations had similar center frequencies of 94 MHz and 95 MHz, respectively. Note that in the previous 2D study, the data was processed similiar to the study by Klotzsche et al. (2010) with dewowing, pre-processing, a time-zero correction, and a 3D-to-2D correction, before the source wavelet was estimated.

4.2. Synthetic case: Study on the source wavelet estimation with synthetic data

First, we want to investigate the source wavelet estimation, based on a deconvolution approach (see Section 2.3.4) in a synthetic study to proof the concept. Therefore, we calculate a realistic synthetic dataset $\mathbf{E}_{\text{obs}}$ based on the final inversion results from Klotzsche et al. (2019c), shown in Figure 4.2c with their estimated effective source wavelet $\mathbf{S}_{SS}$. Furthermore, we added realistic air and water filled borehole geometries, according to the water table of 4.2 m, as well as the finite length antennas described in detail in Section 3.2. Note that as we include the finite length antenna model in the forward model, we estimate the effective source wavelet for the antenna. Since the geometries should account for different couplings between antennas, borehole, and soil, the same effective source wavelet should be estimated, for the four different effective source wavelets related to the Tx-Rx combinations in the unsaturated and saturated zone.

To verify the source wavelet estimation, we consider $\mathbf{E}_{\text{syn}} = \mathbf{E}_{\text{obs}}$ and calculated them with the same input source wavelet $\mathbf{S}_{SS}$. Then, the source wavelet estimation should again yield $\mathbf{S}_{SS}$ for all combinations. The results for the different Tx-Rx combinations in Figure 4.3a and b show that the estimated effective source wavelet for all combinations are overlapping with $\mathbf{S}_{SS}$, indicating that the source wavelet estimation works error free.

In a next step, we test the effect of the starting model on the source wavelet estimation. For this, we compare the estimated effective source wavelets of $\mathbf{E}_{\text{obs}}$ calculated with $\mathbf{S}_{SS}$ and a new $\mathbf{E}_{\text{syn}}$, based on the starting model of Klotzsche et al. (2019c) (see Figure 4.2a), including air and water filled boreholes and finite length antenna models in the 2.5D model, and using the same source wavelet $\mathbf{S}_{SS}$.

The estimated effective source wavelets have similar shapes and amplitudes, indicating that the estimation can capture the essential characteristics of the input source wavelet (see

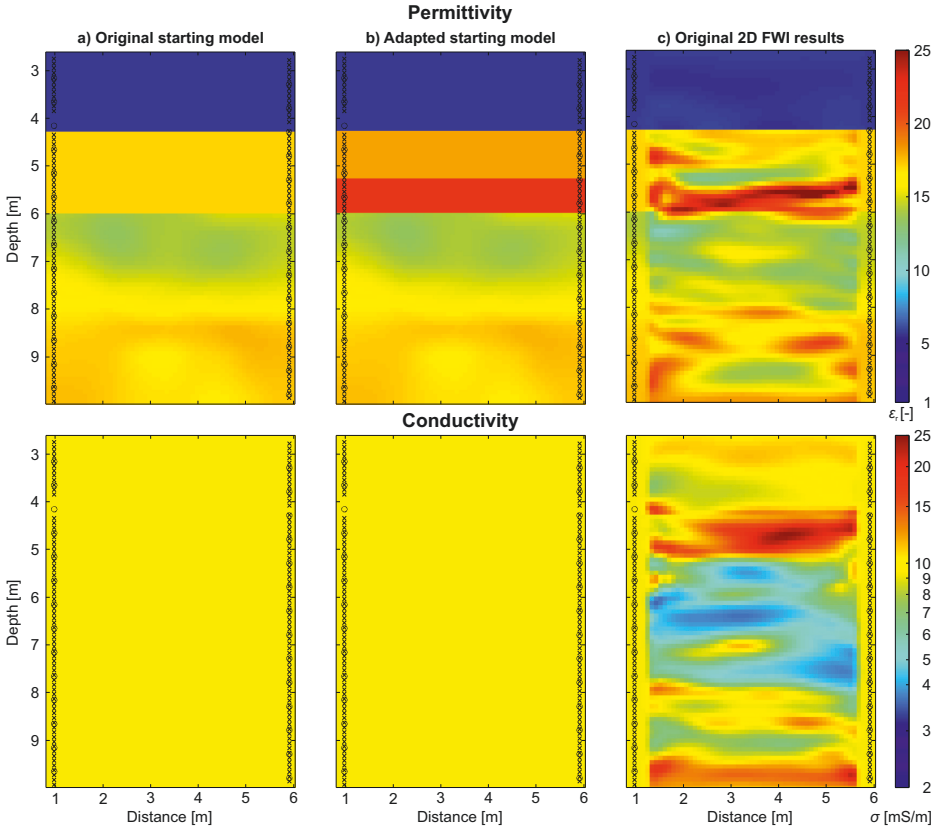

Figure 4.2.: a) Starting model as used in Klotzsche et al. (2019c), b) adapted starting model, similar to Mozaffari et al. (2020) to fulfill the half-wavelength criteria for the 3D data. c) final inversion results of Klotzsche et al. (2019c). Tx and Rx locations are indicated with circles and crosses, respectively.

Figure 4.4a). Differences in phase and amplitude are partly noticeable, which can be related to the errors introduced as the starting model is missing details from the original model. Note that the starting model is based on ray-based inversion analysis and synthetic tests for the 2D case. While UU and SS have similar first-arrival times, US and SU show a time shift (Figure 4.4b). Note that US and SU include larger amount of higher angle data and are more affected by the starting model error. Furthermore, the amplitude of SS decreases after the first maximum. This is especially visible with the second maximum at around 22 ns. US and UU are lower than the original at the first minimum, while at the second minimum, US, SU, and SS are higher than UU and the original (see Figure 4.4c). In the

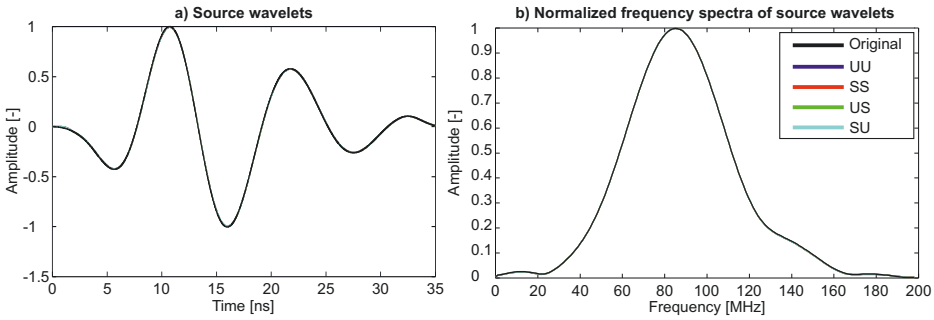

Figure 4.3.: Estimated source wavelets where $\mathbf{E}_{\text{syn}}$ equals $\mathbf{E}_{\text{obs}}$ and are calculated with the same source wavelet $\mathbf{S}_{SS}$. Estimated source wavelet in a) the time-domain and b) the normalized frequency spectrum. Original in black represents the input source wavelet $\mathbf{S}_{SS}$, while UU, SS, US, and SU represent the estimated source wavelets, corresponding to the Tx-Rx combination. Note that all four combinations are overlapping. Furthermore, the source wavelets in a) are normalized to the maximum of all for visualization purposes.

frequency domain, US has fewer frequencies after 130 MHz, but is otherwise similar to the original, as are UU and SU (Figure 4.4d). Of particular note is the lower center frequency of the SS source wavelet (78.5 MHz), which indicates a mismatch of the permittivities in the RBI as compared to the original synthetic model. The center frequencies of UU (85.6 MHz), US (84.4 MHz), and SU (84.4 MHz) are close to the original (85.6 MHz). A possible explanation could be the "stretched" phase of the $\mathbf{S}_{SS}$ after 20 ns, which indicates a lower frequency component. This indicates, that it is possible to estimate an effective source wavelet similar to the original source wavelet, if the forward model is close enough to the original model.

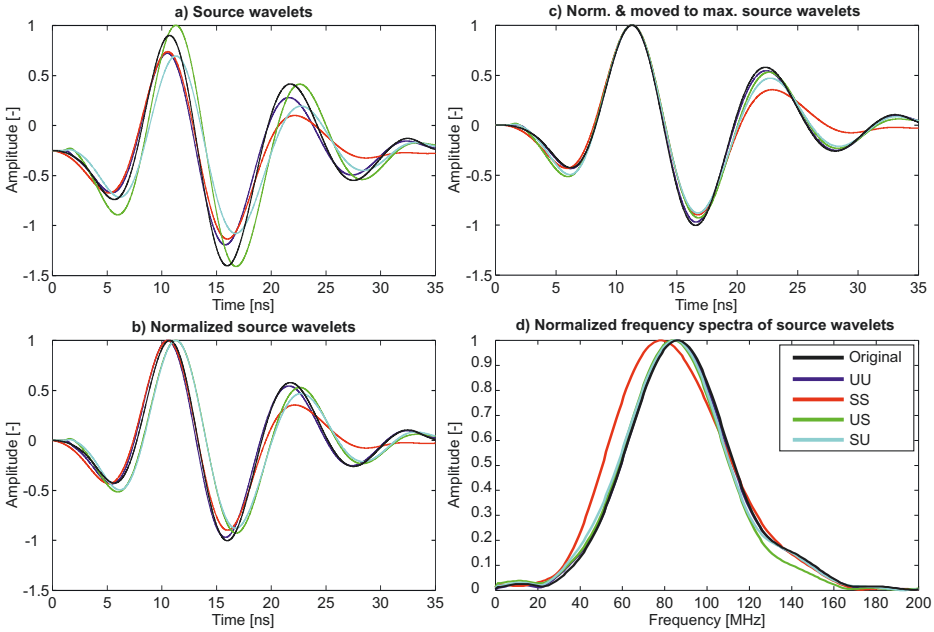

Figure 4.4.: Estimated source wavelets where $\mathbf{E}_{\text{obs}}$ is calculated with 4.2c and $\mathbf{E}_{\text{syn}}$ with 4.2a using $\mathbf{S}_{SS}$. Estimated source wavelets in a) the time-domain, b) normalized in the time-domain in, c) normalized and moved to the maximum in the time-domain, and d) the normalized frequency spectrum. Original in black represents the input source wavelet $\mathbf{S}_{SS}$, while UU, SS, US, and SU represent the estimated source wavelets, corresponding to the Tx-Rx combination. Note that the source wavelets in a) are normalized to the maximum of all for visualization purposes.

4.3. Experimental case: Source wavelet estimation with measured data

The synthetic test showed that the inclusion of borehole filling and the finite-length antenna model yielded similar source wavelets, thus opening the possibility of using a single source wavelet for the variable aquifer. Therefore, we want to investigate the source wavelet estimation with experimental data as used in Klotzsche et al. (2019c), that incorporates data measured in the unsaturated and the saturated zone. As we are using a 2.5D FWI, we do not apply the 3D-to-2D conversion to the data in the pre-processing step. It is important to note that the unconverted 3D data has a higher noise level compared to the converted data.

This is because the 3D-to-2D correction is an integration that amplifies low frequencies and not high frequencies, resulting in a reduction of the frequency content of the data and a decrease in high-frequency noise. However, since we do not conduct any further pre-processing of the data, it includes more noise, particularly in the unsaturated area. In the first test, we used the same water table of 4.2 m as in previous studies.

4.3.1. Source wavelet estimation using the starting model of the 2D case study

First, we test the same starting model as in the 2D case study. Note that already in this study, difficulties with the starting model in UU were indicated. Following the deconvolution approach (see Section 2.3.4), we used the measured dataset from Klotzsche et al. (2019c) for $\mathbf{E}_{\text{obs}}$ and calculated with their original starting model (Figure 4.2a) and estimated source wavelet of SS ($\mathbf{S}_{SS}$) the synthetic dataset $\mathbf{E}_{\text{syn}}$ to derive the effective source wavelets. First thing to note is, that the estimated effective source wavelets show a decrease in amplitude from UU to a similar amplitude for US and SU down to SS (Figure 4.5a), which can be related to the water in the boreholes for these Tx-Rx combinations. Similar to the synthetic case, we can observe a similar first-arrival for UU and SS, whereas there is a time shift for US and SU (Figure 4.5b). Note that US and SU match till the second maximum at around 20 ns, while SS and US show a different behaviour after 10 ns. Furthermore, we see that the second maximum in UU is higher than the first, leading to a mismatch if we move all to the maximum. Therefore, we normalized everything to the minimum und moved it there (see Figure 4.5c for visual comparison), which highlights the similar shape of the estimated effective source wavelets. Note that we can observe a more steep start of the estimated effective source wavelet for SS. As the signals in $\mathbf{E}_{\text{syn}}$ arrive too late compared to $\mathbf{E}_{\text{obs}}$, the deconvolution approach adjusts the source wavelet to an earlier arrival. As the time filter is applied after the frequency filter, low frequency components are introduced in the frequency spectrum of SS mainly visible in the time-domain after 20 ns (Figure 4.5d). Furthermore, a wide gap between the center frequencies from UU (132.7 MHz) to SS (94.6 MHz), US (104.4 MHz), and SU (100.7 MHz) is visible (see Figure 4.5d). US and SU are closer together and SS has a slightly lower center frequency.

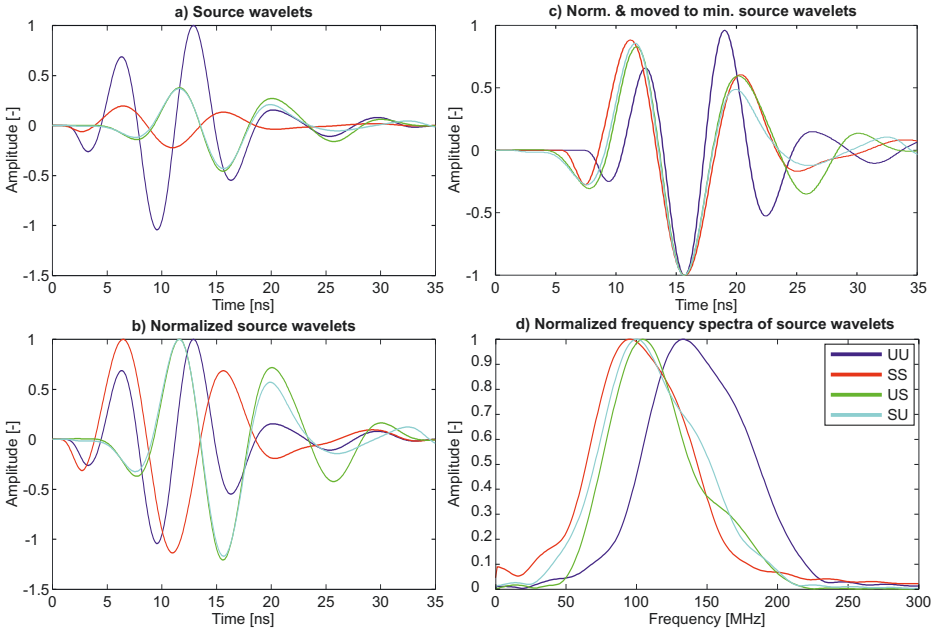

Figure 4.5.: Estimated source wavelets with $\mathbf{E}_{obs}$ being the measured dataset from Klotzsche et al. (2019c) and $\mathbf{E}_{syn}$ calculated with 4.2a. Estimated source wavelets in a) the time-domain, b) normalized in the time-domain, c) normalized and moved to the minimum in the time-domain, and d) the normalized frequency spectrum. UU, SS, US, and SU represent the estimated source wavelets, corresponding to the Tx-Rx combination. Note that the source wavelets in a) are normalized to the maximum of all for visualization purposes.

4.3.2. Source wavelet estimation using an updated starting model

In a second test, we used an adapted starting model, because we observed a misfit in the upper layers (in the range of half of a wavelength) when comparing the calculated data based on Figure 4.2a with the measured data. Note that Klotzsche et al. (2019c) also observed problems, with deriving the starting model, especially in the unsaturated zone. Compared to the RBI results, the layer was updated, as the ray-based picking uses a single shift of a maximum or minimum to the assumed arrival time for all traces the same, resulting in incorrect ray-based travel times. This is because the frequencies for the unsaturated and saturated zones are different. It is possible that this only occurs with the real 3D data, because of the filtering capacities of the 3D-to-2D correction also change the arrival of

the wave. To reduce this misfit, we adjusted the starting model by splitting the second layer into two layers with permittivity values of 18 and 22, respectively. This approach is similar to the one used by Mozaffari et al. (2020) to update the SS starting model for use in 2.5DFWI (Figure 4.2b).

Using this new starting model, we generated a new synthetic dataset $\mathbf{E}_{\text{syn}}$ that better aligns with the measured data $\mathbf{E}_{\text{obs}}$, particularly in the upper area. We then derived with the deconvolution approach new effective source wavelets for the different Tx-Rx combinations and compare them in detail with the results from Klotzsche et al. (2019c). All effective source wavelets that are estimated using $\mathbf{E}_{\text{syn}}$ with the new complex 2.5D forward model show an earlier first-arrival compared to the 2D estimations (Figures 4.6a and b). Note that there is a difference of approximately 3-4 ns between SS, US, and SU, while there is only a 1 ns difference for UU. The source wavelets for US and SU are further apart than with the previous starting model. This difference can be attributed to the antenna geometry, as well as the water filled boreholes. Additionally, in the 2D case, the estimated effective source wavelet for SU arrived before US, whereas it is reversed in the 2.5D case. A reduced amplitude difference between the estimated effective source wavelets SS, US, and SU can be observed when transitioning from 2D to 2.5D. As we do not change the conductivity of the model, this has to be related to the different refraction, caused by the changed medium properties. Compared to the 2.5D estimation for the original starting model, we are able to observe an improvement in the amplitudes with the adapted RBI model, as UU is now more similar to SS, US, and SU with the first maximum now higher than the second (Figure 4.6c). Overall, the center frequencies of SS (94.6 MHz), US (98.3 MHz), and SU (95.8 MHz) are closer together and UU (126.5 MHz) is also slightly lower with the new adapted RBI starting model for 2.5D compared to the 2D estimation (Figure 4.6d). Note that similar to the estimated effective source wavelet for SS with the original starting model, low frequency components are introduced for SS with the new starting model.

The results indicate that it is feasible to estimate a single source wavelet for the SS, US, and SU cases, which is a notable improvement compared to Klotzsche et al. (2019c). However, when both Tx and Rx are in the unsaturated zone, the center frequency is evidently too high compared to SS, US, and SU. This discrepancy may be attributed to a different coupling between the antenna, the air filled borehole, and the soil in reality. The reason for this may be that our antenna model (Sensors and Software crosshole 200 MHz PulseEKKO system) does not match the system used to measure the data (250 MHz RAMAC Ground

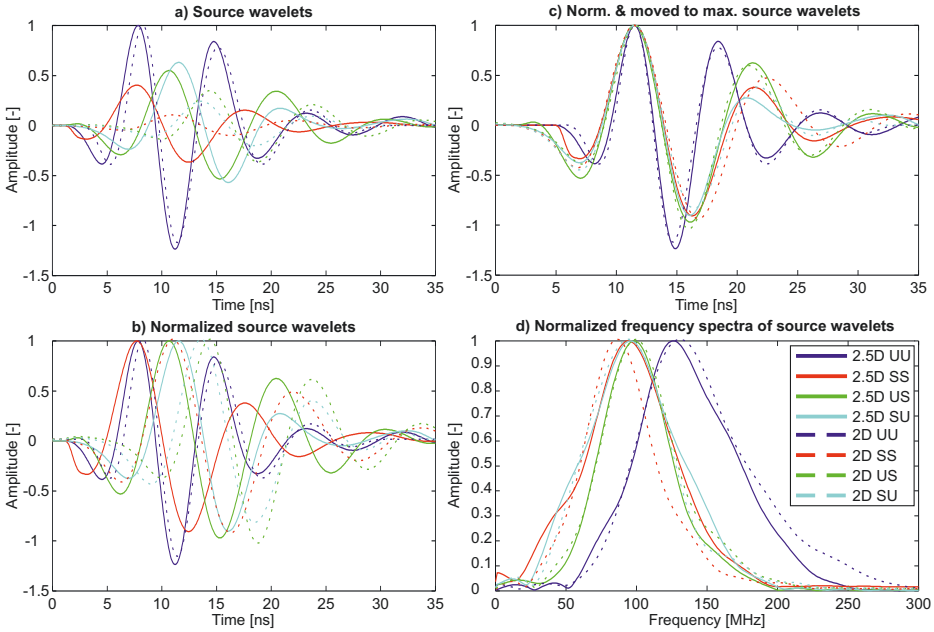

Figure 4.6.: Estimated source wavelets with $\mathbf{E}_{obs}$ being the measured dataset from Klotzsche et al. (2019c) and $\mathbf{E}_{syn}$ calculated with 4.2b. Estimated source wavelets in a) the time-domain, b) normalized in the time-domain, c) normalized and moved to the maximum in the time-domain, and d) the normalized frequency spectrum. UU, SS, US, and SU represent the estimated source wavelets, corresponding to the Tx-Rx combination. Note that the estimated source wavelets from Klotzsche et al. (2019c) are indicated with a dashed line. Note that the source wavelets in a) are normalized to the maximum of all for visualization purposes.

Vision system from MalåGeoscience). Therefore, a further adaption of the antenna model, especially for air filled boreholes, is needed.

Note that the depth of the water table measured by the GPR data depends on the thickness of the capillary fringe of the medium (Igel et al., 2013; Igel et al., 2016). In environments characterized by unsaturated and saturated gravel, the capillary transition zone is expected to be very thin (10 cm to 20 cm). However, the water table measured in the boreholes may vary. Therefore, we altered the water table (4.2 m based on Klotzsche et al. (2019c)) in the boreholes by 20 cm to investigate the effect of the capillary fringe. Furthermore, we also investigated the source wavelet estimation with point sources. Moreover we investigated different starting models for permittivity, where we: adjusted only the top

layer ($\varepsilon_r = 5$); the top layer ($\varepsilon_r = 7$) and divided the second layer into two ($\varepsilon_r = 22$ and $\varepsilon_r = 18$); divided the second layer in $\varepsilon_r = 22$ and have kept the other half as in Klotzsche et al. (2019c); changed the second layer to $\varepsilon_r = 18$. All these investigations showed no further improvements.

4.4. 2.5D GPR FWI of a variably saturated soil-aquifer system with one source wavelet

Although a perfect fitting source wavelet for all Tx-Rx combinations was not found, we want to investigate a FWI with the combined estimated source wavelet from Figure 4.6 with 4.2b as a starting model. We are aware, that this will introduce errors caused by the mismatch of the source wavelet for the unsaturated zone, but we expect to be able to reconstruct similar structures for the saturated zone as Klotzsche et al. (2019c) and improve the saturated zone. Furthermore, we compare our inversion results to the final results from Mozaffari et al. (2020), who investigated only the saturated part using the 2.5D FWI.

The permittivity distributions of the different results show a remarkable degree of similarity between the models, with slight variations in the different layers (Figure 4.7). The 2.5D FWI reconstructs lower permittivity values in the 4.8 m - 5.2 m region compared to both the 2D FWI and 2.5D FWI with borehole and antenna models (BH A). The 2.5D FWI BH A reconstruction shows low permittivity "wings" extending upward into the unsaturated zone from 4.2 m to 5.3 m depth, in contrast to the higher permittivity layer observed in the other models at similar depths. Across all models, a high permittivity layer exists at a consistent depth between 5.3 m and 6 m, although with varying intensity. Inside this layer is a high permittivity concentration, predominantly to the right for the 2D FWI and 2.5DFWI, and additionally to the left for the 2.5D FWI BH A. Notably, both the 2D FWI and 2.5D FWI BH A reveal a lens-like feature at a depth of 5.5 m to 6 m and a distance of 1.5 m, which is missing in the 2.5D FWI reconstruction. The reconstructions consistently identify a low permittivity layer spanning from 6 m to 8 m, with visible variations in the reconstruction of a central lens. The 2.5D FWI shows a gradual dip from the left to the right, while the 2D FWI and 2.5D FWI BH A show a more horizontally layered structure at about 7.5 m depth. This layer is followed by a higher permittivity layer in all models appear in similar configurations, although the 2D FWI structure differs slightly from that of the 2.5D FWI and 2.5D FWI BH A models. At the base of the domain, there is

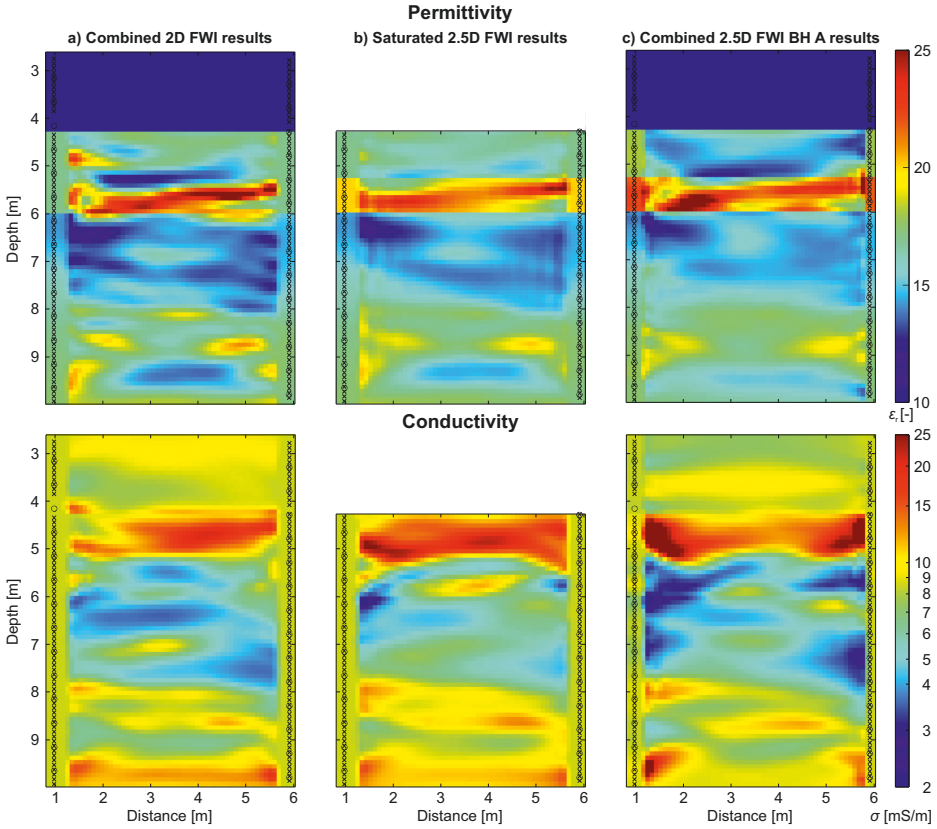

Figure 4.7.: Final FWI results for permittivity and conductivity for the experimental data from the Widen test site. The final results from using a) a combined 2D FWI for the unsaturated and saturated zone (Klotzsche et al., 2019c), b) a 2.5D FWI for the saturated zone (Mozaffari et al., 2020), and c) the new 2.5D FWI with borehole and antenna (BH A) combined for the unsaturated and saturated zone. Note that four source wavelets were used in a), while b) and c) used only one. Tx and Rx locations are indicated with circles and crosses, respectively.

an area with reduced permittivity. The 2D FWI shows lower values than the 2.5D FWI and 2.5D FWI BH A. Furthermore, a curvature in the layer, associated with the ray coverage is reduced in the 2.5D FWI BH A reconstruction.

The analysis of conductivity is limited till a depth of 4.2 m to 2D FWI and 2.5D FWI BH A because Mozaffari et al. (2020) only investigated the saturated zone. We can observe a high conductivity layer in the 2D FWI reconstruction, followed by a lower conductivity while we see a low conductivity for the 2.5D FWI BH A, then high conductivity and then

again low conductivity values. Each model reconstructs a high conductivity layer between 4 m and 5 m. The 2D FWI displays predominantly straight borders, while the 2.5D FWI and 2.5D FWI BH A display curved lower borders. The 2.5D FWI notably features a zone of increased conductivity on the left side, and existing or the 2.5D FWI BH A also on the right, which is not present in the 2D FWI reconstructions. The results show a lower conductivity layer extending to 7.8 m. A few higher conductivity lenses exhibit similarities in position and amplitude. However, the amplitude of the central lens at 5.5 m depth is higher in the 2.5D FWI than in the 2D FWI and 2.5D FWI BH A. Low conductivity intrusions are visible in the 2.5D FWI BH A. Starting from a depth of 8 m, there is a noticeable high conductivity zone that transitions into a layer of lower conductivity at around 9 m in all reconstructions. The lowest layer shows similar conductivity in both the 2D and 2.5D FWI. However, in the 2.5D FWI BH A, there is a significant concentration of higher conductivity near the boreholes.

The analysis of the RMS reduction shows a 49.36% reduction for the 2D FWI, while it is 30.67% for the 2.5D FWI BH A. It is important to note that the reduction for 2.5D FWI BH A was calculated using the updated starting model to fulfill the half-wavelength criterion. If we calculate the RMS reduction based on the initial RMS with the original starting model used for 2D FWI, we observe a reduction of 33.30%. The lower RMS reduction for the 2.5D FWI and 2.5D FWI BH A may be related to additional noise in the original 3D data, compared to the 2D data. This may be related to the Bleistein transformation, which lowers the frequency content of the data and thus reduces the high-frequency noise. Additionally, the 2D FWI has an advantage in the unsaturated area with an specifically for that area adapted higher frequency source wavelet. We achieved better results than the 2.5D FWI, which only considered the saturated area and showed an RMS reduction of 23.63%.

To further investigate the RMS, a visualization of the RMS distribution of the final inversion results for 2D FWI and 2.5D FWI BH A is visible in Figure 4.8. The RMS for each Tx-Rx combination was normalized to the mean RMS for all Tx-Rx combinations of the starting model. The 2.5D FWI with borehole antennas (BH A) in particular allows to distinguish between the unsaturated and saturated zone. This is indicated by the increased RMS in the unsaturated region because of the misfit created by using a lower frequency source wavelet. The 2D FWI utilizes an adapted source wavelet with a higher frequency for the unsaturated area, resulting in a better fit and lower RMS values in this area. Furthermore, we observe higher amplitudes for the unsaturated area, resulting in a higher RMS compared to the saturated zone. The 2D FWI shows higher RMS values for transmitters

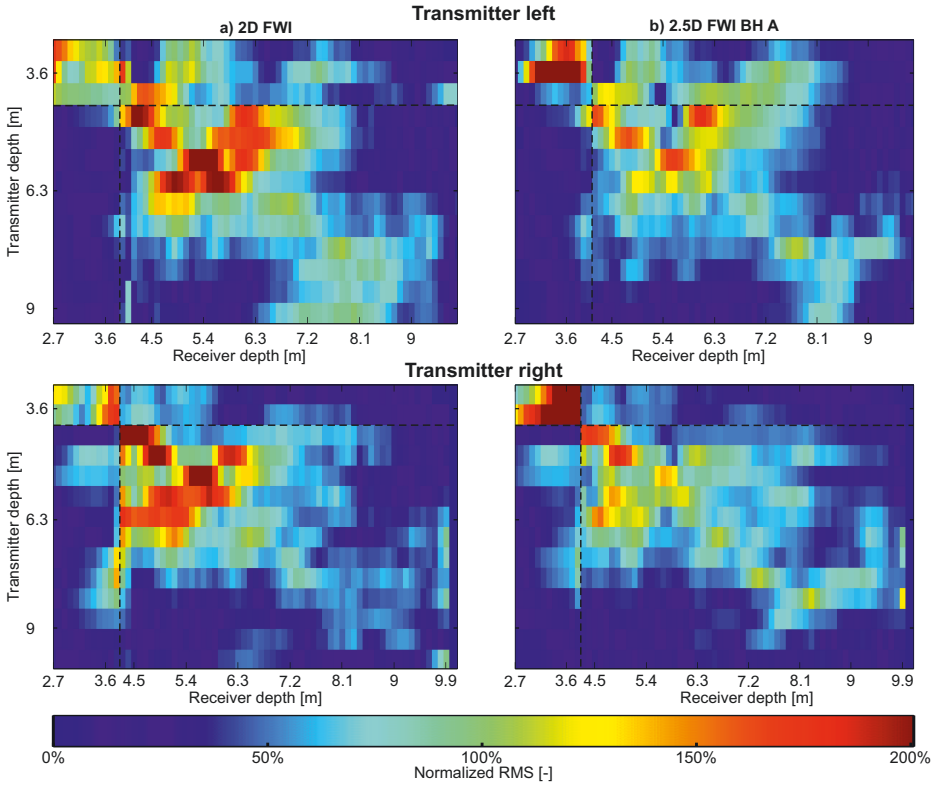

Figure 4.8.: Root mean square error (RMS) normalized to the mean RMS of the starting model for every transmitter receiver combination, in the upper part the transmitter in the left borehole and below in the right. For the final results of a) 2D FWI and b) 2.5D FWI with borehole and antenna (BH A). Note the dashed lines, indicating the water table and high values shown in red.

located between depths of 3.6 m and 7 m, particularly in the central part of the domain, when compared to the 2.5D FWI BH A. This area also includes the waveguide, indicating a better fit with the 2.5D FWI BH A. However, both inversion strategies exhibit comparable RMS profiles below 6.5 m and across the remaining domain. It is worth noting that when the transmitter is placed in the left borehole at a depth of 7 m, the 2.5D FWI BH A exhibits lower errors than the 2D FWI. This is reversed when the transmitter is positioned in the right borehole, with the 2.5D FWI BH A showing increased RMS errors compared to the 2D FWI for receivers beyond 7 m depth. This could be explained by a better fit in the left part of the model as then the coupling of the antenna emitting the wave into the ground is better.

Note that the coefficient of determination, r^2, which is the square of the correlation coefficient, for all traces decreases from 0.7686 to 0.6026, when we move from 2D FWI with four estimated effective source wavelets to 2.5D FWI BH A with one estimated effective source wavelet. This shows that further investigation of the antenna model is inevitable for a better data fit of the inversion.

4.5. Conclusions and outlook

In the first part of this chapter, we conducted a synthetic study on source wavelet estimation using a deconvolution approach to determine its functionality and how errors in a realistic model affect the source wavelet. Our observations revealed that errors primarily affect the time of rise in the source wavelet, with minor changes in amplitude. Additionally, we noted a lower center frequency for the source wavelet for SS. During our analysis of the source wavelet estimation using measured data, we observed changes in amplitude that correlated with the positioning of the transmitter and receiver. Specifically, we observed a decreasing amplitude from UU to US and SU down to SS. This can be related to the presence of water in the boreholes, which attenuates the electromagnetic wave. Our findings are consistent with those of Klotzsche et al. (2019c), a previous study that used 2D FWI and four effective source wavelets and also observed that UU yielded the highest frequency source wavelet and SS the lowest, with US and SU falling in between. An update to the starting model resulted in source wavelets with similar center frequencies for US, SU, and SS. Our hypothesis of a better fit for all source wavelets only worked for partly water filled boreholes. As the antenna model does not fully match the antennas used for the measurement, we must consider its different coupling behavior. Especially in air filled boreholes it proves to be a major issue.

Finally, we performed a 2.5D GPR FWI of unsaturated and saturated soil with one source wavelet. Comparing it to two previous studies on the same dataset, we were able to reconstruct very similar structures in the saturated zone. Since there is no ground truth, we can only evaluate the data misfit, which was particularly high for the unsaturated zone. Considering that the antenna model is based on the Sensors and Software crosshole 200 MHz PulseEKKO system and not on the RAMAC Ground Vision system of Malå Geosience with 250 MHz antennas that was used to measure the data, allows the conclusion that the results for the unsaturated area cannot be assessed as reliable. In order to use one source wavelet for the entire domain, the antenna model needs improvement.

The reconstruction of the unsaturated and saturated zone could be further improved by using a customized antenna model. Especially, by adapting the resistive-loaded finite-

length antenna based on the Sensors and Software crosshole 200 MHz PulseEKKO system to the RAMAC Ground Vision system of Malå Geosience with 250 MHz antennas, which was used for the measured data. For this, we plan to collaborate with experts in the field, such as Peter Annan, and explore an optimization technique for anntenna modeling. This could include Taguchi's optimization method (e.g., Warren and Giannopoulos, 2011; Stadler and Igel, 2018), a hybrid linear/nonlinear FWI approach (e.g., Giannakis et al., 2019), or a particle swarm optimization algorithm (e.g., Stadler and Igel, 2022). The implementation was beyond the scope of this study because of its time-consuming nature. However, we plan to implement it in the future to enable a more detailed investigation of other datasets measured with the same antenna.

In the current implementation of the 2.5D GPR FWI, we disregard the bending of the borehole, a common drilling problem. This can result in a displacement of the antenna from its central position in the borehole, causing not centered. As a result, the energy transferred to the soil is no longer symmetrical. In a future study, we want to quantify the influence of this displacement, since it introduces a significant implementation challenge, as the entire antenna geometry must be deviated within the equidistant grid of the FDTD grid. To accurately model the deviated antenna and avoid introducing staircasing errors, a much finer subgrid is necessary. However, this results in computational times that are currently not feasible. With the steady increase in computing capacity of supercomputers, it is reasonable to expect that it will be possible to implement these deviations in the near future.

In the 2D FWI case the data was transformed from 3D-to-2D with the Bleistein filter. As this lowers the frequency content of the data, high-frequency noise can be reduced in the 2D data. In a next step, we want to apply a noise filter to the measured 3D data and investigate its influence on the reconstruction.

The current implementation roughly covers the capillary fringe with a grid size of 9 cm. This causes the implementation and reconstruction to treat it as a hard contrast. However, the capillary fringe is actually not a distinct contrast between the unsaturated and saturated zones, but rather a transition. Therefore, errors are introduced as harder contrasts result in stronger reflections and data where the feeding point of the antenna was near the capillary fringe could not be used. If a subgrid is inserted in the area of the capillary fringe, it is possible to implement a smoothing of the hard contrast to a smooth transition from two cells by interpolating the medium over several cells, making this area more realistic. However, care must be taken to ensure a correct representation of the prevailing water column.

5. Evaluation of starting model approaches and effective source wavelet variations for high-frequency ground-penetrating radar full-waveform inversion[1]

A high resolution reconstruction is necessary for the investigation of small-scale structures. The theoretical reconstruction resolution of the FWI is in the order of half the dominant wavelength of the signal (Virieux and Operto, 2009). As GPR FWI can use high frequencies (Klotzsche et al., 2018) and the wavelength is smaller with higher frequencies (Equation (2.22)), we should be able to increase the reconstruction resolution with higher frequency pulses. In the following, we will investigate regular approaches for high-frequency GPR data, investigate their limitations and propose an alternative solution. Note that this study was conducted in the computationally efficient 2D domain.

5.1. Synthetic aquifer model

To test FWI for higher-frequency GPR data (up to 700 MHz), we defined a synthetic aquifer model based on FWI results obtained from the Widen site (Klotzsche et al., 2012). The geometry of the model (Figure 5.1) was chosen for its similarity to a typical soil column or lysimeter setup, which are often used in laboratory studies to investigate preferential flow using geophysical methods, for example (Garré et al., 2011; Schmalholz et al., 2004). This

[1] adapted from D. Hoven, A. Mester, H. Vereecken, and A. Klotzsche (2023). "Evaluation of starting model approaches and effective source wavelet variations for high-frequency ground-penetrating radar full-waveform inversion". In: *GEOPHYSICS* 88.2, KS27–KS45. DOI: 10.1190/geo2021-0683.1

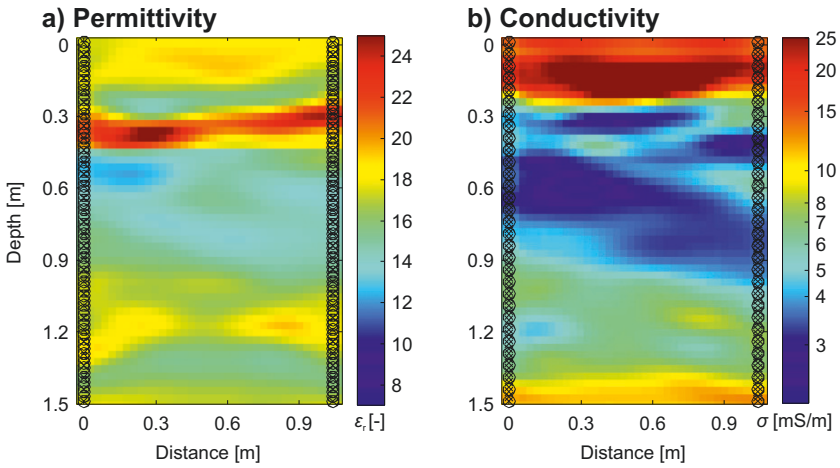

Figure 5.1.: a) ε_r and b) σ distributions for the synthetic aquifer model. Input models are based on the Widen data (Klotzsche et al., 2012) and downscaled to the size of a lysimeter. Note the logarithmic scale for the conductivity image. Tx and Rx locations are indicated with circles and crosses, respectively.

heterogeneous and challenging model was chosen, because it is based on a gravel aquifer and contains a waveguiding structure related to zones of preferential flow that causes late-arrival, high-amplitude events in the data. Such events are difficult to detect with RBI and can also cause problems when defining an optimal starting model for FWI.

The dimensions of the model are based on a real lysimeter with a height of 1.5 m and a diameter of 1.1 m. To avoid reflections at the model boundaries in the FDTD, which is used to calculate the electromagnetic fields, perfectly matched layers were defined around the forward modeling domain. They require a distance of approx. 0.5 m to the edge and were added to the model as additional cells. We tested these boundary conditions, and they satisfy the assumption of no reflection. Since the inversion is based on a cell size of 3 cm, the constructed models will have a height of 2.52 m and a width of 2.1 m. Note that the forward model has a finer discretization of 1 cm to decrease numerical errors in the FDTD. The permittivity and conductivity values of the gravel aquifer were interpolated to fit the dimensions of the lysimeter. The area next to the lysimeter towards the domain boundaries was filled with the last value inside the lysimeter. Similarly to crosshole studies, we positioned our transmitter Tx and receiver Rx antennas around the soil column. The area enclosed by Tx and Rx has a height of 1.5 m and a width of 1.08 m, and is located at the center of the model. Tx and Rx spacing are 0.05 m and 0.02 m for the multi-offset gathers,

respectively, and we applied a reciprocal setup, which means that we switched the sides of Tx and Rx and repeated the measurements. For the simulations, Ricker source wavelets with center frequencies of 200, 450 and 700 MHz were used. We chose 700 MHz as the highest frequency, because this is approximately the center frequency of GPR data provided by a commercial antenna (e.g., EKKO Pro 1000 MHz, Sensors & Software, ON, Canada) placed around a soil column. RBI requires first break times to be picked, which is difficult due to the slow increase of the Ricker wavelets. Therefore, we modified these wavelets by shifting the starting time towards 0 ns, and we also tapered the wavelet amplitudes after this time shift to guarantee a start at 0 amplitude at 0 ns. The maximum amplitudes of the wavelets were normalized to 1 by division with their maximum amplitude, and therefore the amplitudes are unitless.

As discussed before, one crucial point for the FWI is to obtain starting models within the half-wavelength criterion. Therefore, we investigated the allowed maximum difference for the permittivity for our model and considered thereby the planned frequencies in order to guarantee reliable and stable FWI results. To calculate this maximum difference $\Delta\varepsilon_r$ for the FWI relative permittivity starting model, a number of parameters, described below, need to be known. The first break of the GPR signals is the time required for a wave to travel the specified distance in a material with a certain permittivity, and is provided in ns. As a function, the first break fb is then defined as:

$$fb(d, \varepsilon_r) = \frac{d}{c}\sqrt{\varepsilon_r} \tag{5.1}$$

with the distance d between Tx and Rx, c the speed of light ($c \sim 0.3$ m/ns) and the relative permittivity of the material ε_r . It should be noted that for all further applications, we used the relative dielectric permittivity, which is unitless. For this case, half the time between the maximum and minimum of the wavelet was chosen as half the wavelength. The unitless maximum difference $\Delta\varepsilon_r$ of the relative permittivity is then defined as:

$$\Delta\varepsilon_r(fb, \lambda_h, d) = \left(\frac{fb \pm \lambda_h}{d}c\right)^2, \tag{5.2}$$

with the half-wavelength λ_h. Here, the half-wavelength λ_h is not the wavelength of the data but the parameter of the half-wavelength criterion in nanoseconds, and depends on the data. The first break, half-wavelength criterion, and maximum differences $\Delta\varepsilon_r$ were calculated in Table 5.1 for the minimum, mean, and maximum ε_r in the synthetic aquifer model (12.32, 16.54, and 25.87). For calculating the requirements of the half-wavelength

Table 5.1.: Parameters to estimate the half-wavelength parameter for the different frequencies: First break, half-wavelength criterion time parameter and maximum difference $\Delta\varepsilon_r$ for 200, 450, and 700 MHz. Calculations are performed for the minimum, mean, and maximum ε_r in the synthetic aquifer model (12.32, 16.54 and 25.87).

Center frequency	200 MHz			450 MHz			700 MHz		
ε_r	12.32	16.54	25.87	12.32	16.54	25.87	12.32	16.54	25.87
First break [ns]	12.88	14.92	18.66	12.88	14.92	18.66	12.88	14.92	18.66
λ_h [ns]	0.97			0.43			0.28		
Max. difference $\Delta\varepsilon_r$									
ε_{min}	10.53	14.46	23.25	11.51	15.60	24.69	11.79	15.93	25.10
ε_{max}	14.24	18.76	28.63	13.15	17.51	27.08	12.86	17.17	26.65

criterion, we consider these values for the entire domain instead of heterogeneous medium properties. We assumed a distance d of 1.1 m to obtain the first break and the maximum difference $\Delta\varepsilon_r$. The most severe constraints on the starting model are in areas with a high permittivity. Therefore, we will address this case ($\varepsilon_r = 25.87$) in particular. While for 200 MHz, the half-wavelength criterion time parameter λ_h allows a shift of almost 1 ns and a deviation of approximately ± 2 in permittivity, for 700 MHz this shift is reduced to 0.3 ns and the permittivity should not deviate by more than ± 0.5 for such high contrast zones. It is clear that for the higher frequency in particular, the criterion is harder to reach, and difficulties can be expected. Since we calculated the criterion for a homogeneous material, high restrictions are only the case for high permittivity layers. Nevertheless, if the starting model deviates by more than this in such areas, the FWI will probably not converge for this starting model and remaining gradients will be present (Klotzsche et al., 2010). This implies that for low-frequency data less accurate starting models can be used, while for higher frequency data a refined starting model approach is necessary. Inaccurate starting models are often indicated in the final inversion results by remaining gradient in certain domains or a bad data fit between the measured and modeled data. An often-used option for refinements of the starting models is to add structures in the ray-based inversion results, in the domain where the observed and modeled data not fulfill the half-wavelength criterion. Additionally, an amplitude analysis approach of the measured data can be used to identify areas in the domain with high contrast that can be added to the starting models (Zhou et al., 2020). Such approaches are well suited for lower frequencies but struggle with higher frequency data and often require some training on the applicant to identify such problematic zones. In contrast, frequency-hopping allows a refinement that covers the entire inversion domain, which makes it particularly attractive for high frequencies.

5.2. FWI for high-frequency data

Until now, FWI has only been applied to frequencies up to 250 MHz. To achieve an improved characterization and resolution of the subsurface and thus to monitor changes related to small-scale processes below the decimeter-scale, we first tested the effect of higher frequencies on FWI results and how the inversion parameters need to be updated. Therefore, we produced GPR data for our synthetic aquifer model (Figure 5.1) using the three modified Ricker wavelets, and performed individual FWI using different starting model strategies. In the following discussion, all FWI results fulfill the four criteria indicating reliable FWI results that were described by Klotzsche et al. (2019b), unless otherwise stated. An important aspect to consider for the inversion is the choice of perturbation and stabilization factors, which are also affected by the frequency of the measured data, choice of frequency steps in the frequency-hopping, and starting models. It should be noted that we optimized these factors in the beginning for each dataset with the different frequencies (e.g., 200, 450 and 700 MHz) and starting model approaches, to fulfill the four criteria indicating reliable FWI results by testing various options and set of parameters. The factors remain constant during the frequency-hopping.

First, we investigated the performance of the FWI using the 200 MHz data to show the ability of the FWI to correctly invert the data generated with the synthetic aquifer model with a lower frequency. For the 200 MHz FWI, we tested the three starting model approaches that are commonly considered for this frequency range (Figure 5.2): smoothed ε_r input model, ε_r ray-based inversion result and modified ray-based ε_r. The modified permittivity model was chosen in accordance with Klotzsche et al. (2012), who showed that the pure ray-based results cannot explain the results of the Widen site, and therefore a homogeneous layer was added to the FWI starting model to fit the half-wavelength criterion. For the FWI results presented here, a homogeneous starting model of the conductivity was considered in a similar way to Klotzsche et al. (2012), where different starting models for the conductivities were tested, and we therefore chose the same best conductivity starting model described in that study ($\sigma = 9.5$ mS/m). The perturbation factors for the step lengths (Equation 2.39 and 2.40) for 200 MHz need to be between 1e8 and 1e7 for ε_r and 1e5 for σ. We can see that the starting models (Figure 5.2a) differ mainly in the permittivity values between 0 m and 0.4 m depth as well as 1 m to 1.5 m. In these regions, we can also find the greatest difference between the corresponding FWI results (Figure 5.2b and d). Differences in the permittivity tomograms between 0 m and 0.2 m are visible in a direct comparison of the FWI results (Figure 5.2b) and by calculating the difference images (Figure 5.2c).

The permittivity results of the FWI with the RB starting model are showing larger values in the upper area between 0 m and 0.2 m in contrast to the input model, while the results with the RB-L starting model are slightly too low, which is consistent with the starting models. The influence of the starting models on permittivity in the area between 1.2 m and 1.5 m is only visible in the difference plot (Figure 5.2c). Furthermore, larger errors can be seen in the waveguide (0.3 m to 0.4 m) in addition to smaller errors throughout the model. All FWI conductivity results (Figure 5.2d) show a bending in the layer between 0 m and 0.2 m as well as between 1.3 m and 1.5 m. This is caused by insufficient ray path coverage in these areas, which is due to the positioning of Tx and Rx. Moreover, the fitting is noticeably worse in the areas peripheral to the antennas between 0.3 m and 0.9 m. The difference plot (Figure 5.2e) shows the previously mentioned problematic areas, as well as minor deviations in the inner model area.

Comparing the final RMS values for the three inversions in Table 5.2, the best values were obtained from the FWI using the smoothed starting model, although it should be taken into account that this option is not available for experimental data applications. This is also true when considering the reduction of the RBI RMS (1.58e4) to the final RMS of the FWI, stated here as a percentage below the RMS values. The RBI RMS was derived by performing an FDTD simulation based on the ray-based results and comparing the resulting traces to the real traces. For all FWI results, there is a 95% reduction in RMS compared to the RBI RMS. A similar behavior can be observed for the mean absolute error (MAE) of the standard deviation of the absolute error (STDAE) for the permittivity and conductivity tomograms. For the calculations of these performance values, the zone between the antenna positions was considered. RB provides the next-best results for the ε_r values. RB-L was noted to have the best values for σ MAE and σ STDAE despite having the poorest ε_r values. If FWI shows smaller errors in even one parameter, it must be kept in mind that in a multi-parameter inversion, all parameters need to be acceptable. By comparing the FWI tomograms and the reconstruction of the parameters, we can conclude that the modified ray-based results deliver the most optimal results, given that smoothed starting models cannot be considered for experimental data due to the unknown subsurface.

Similarly, to the procedure for the 200 MHz FWI, we generated 450 MHz data and performed FWI for different starting models (Figure 5.3a). In addition to the S, RB, and RB-L starting models for the 450 MHz FWI, we added two more starting models based on the RB and RB-L models from the 200 MHz FWI. These two new starting models are indicated by the suffix IT4. They are the permittivity distributions of the fourth iteration

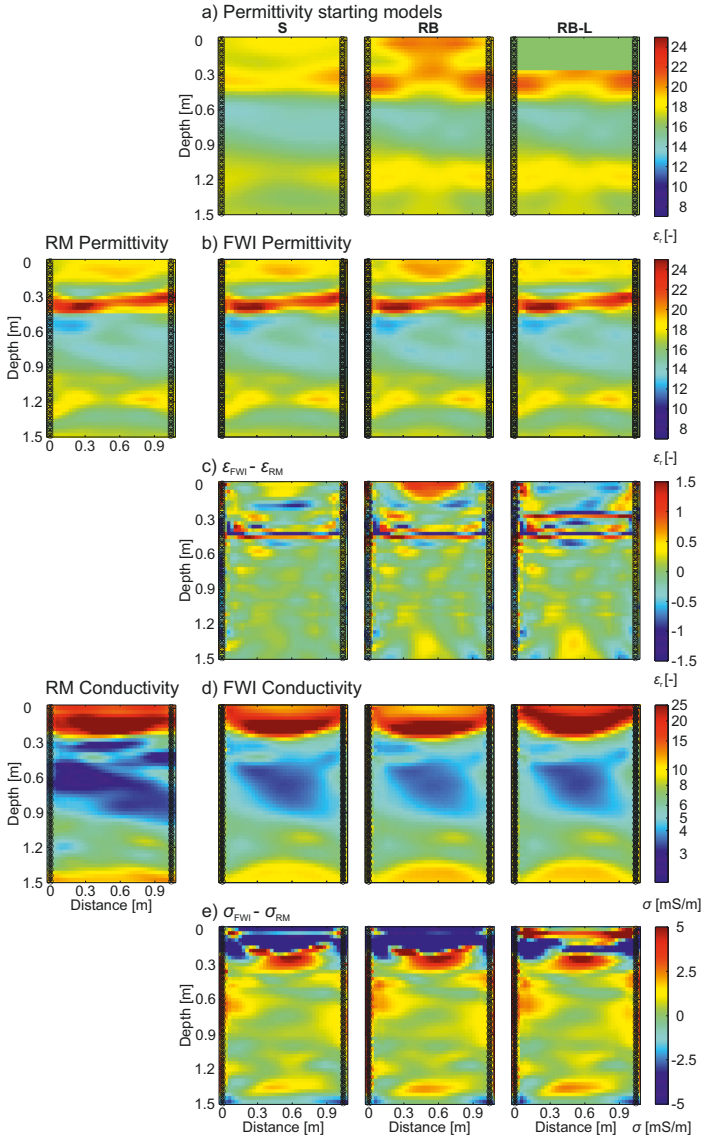

Figure 5.2.: Overview of the 200 MHz FWI results. The synthetic aquifer models (with RM signifying real model) are shown on the left. a) shows the different permittivity starting models (S: smoothed input model, RB: ray-based inversion result, RB-L: ray-based inversion result and added homogeneous layer), while the conductivity starting models are always homogeneous. The corresponding permittivity and conductivity FWI results are presented in b) and d), respectively. The errors between the synthetic aquifer model and the FWI results are shown for c) ε_r and e) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

Table 5.2.: Overview of the 200 MHz FWI results in terms of root-mean-square error (RMS), mean absolute error (MAE), and the standard deviation of the absolute error (STDAE) for permittivity and conductivity using three different starting models. The smoothed, ray-based, and modified ray-based models are indicated by S, RB, and RB-L.

	S	RB	RB-L
RMS	5.8e-6 3.59%	7.2e-6 4.57%	7.9e-6 5.03%
ε_r MAE	0.19	0.24	0.29
ε_r STDAE	0.24	0.28	0.31
σ MAE	1.48	1.80	1.40
σ STDAE	2.45	3.00	1.83

of the 200 MHz FWI, with the respective starting model. This approach is based on the frequency-hopping previously described. Comparing the five starting models, the difference made by the additional four iterations of the 200 MHz FWI can be seen in RB-IT4 and RB-L-IT4. The waveguide with its higher permittivity is more separated and has increased in amplitude. In the RB-IT4 starting model, we can already see the curved shape of the layer between a depth of 0 m and 0.2 m caused by the insufficient ray path coverage. In contrast, a homogeneous uniform layer with a lower permittivity value can be seen in the RB-L-IT4 starting model. The area with a lower permittivity at a depth of 0.6 m and between a distance of 0 m and 0.6 m is now visible, whereas the higher permittivity at a depth of 1.2 m and a distance between 0 m and 0.3 m increases. Compared to the 200 MHz FWI permittivity results, we can see negligible changes in the 450 MHz S starting model results (Figure 5.3b). For RB, the layer between 0 m and 0.2 m is curved, while the fit for RB-L decreases on the sides between a distance of 0 m to 0.3 m and 0.8 m to 1.1 m in this layer. The permittivity results for RB-IT4 are similar to those for RB, although the upward dipping layer between 0 m and 0.25 m depth is narrower in distance but increased in depth. The error plot (Figure 5.3c) shows the described changes in the upper layer, as well as errors in the waveguide as previously noted for 200 MHz. However, there is a noticeable diagonal line from a depth of 0.5 m and a distance of 0 m to a depth of 1.5 m and a distance of 0.8 m with higher permittivities. This diagonal effect, which is probably a numerical artefact, is not visible in the S starting model.

The conductivity results (Figure 5.3d) show a curve in the upper and lower layers, as they did at 200 MHz. For the S, RB, RB-IT4 and RB-L-IT4 models, the curve in the upper layer has a higher conductivity, while above it is a thin layer with a lower conductivity. For

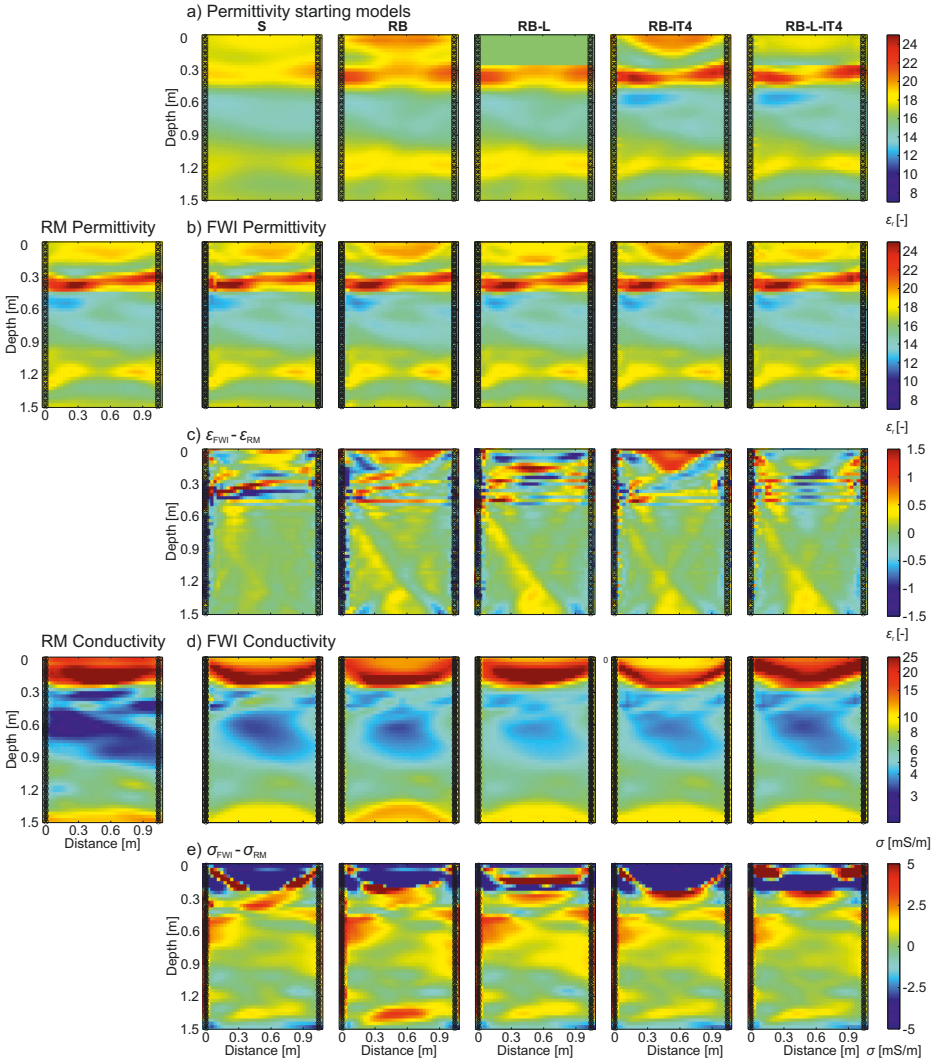

Figure 5.3.: Overview of the 450 MHz FWI results. The synthetic aquifer models (with RM signifying real model) are shown on the left. a) shows the different permittivity starting models (S: smoothed input model, RB: ray-based inversion result, RB-L: ray-based inversion result and added homogeneous layer, RB-IT4: result of the fourth iteration of the 200 MHz FWI based on the RB starting model, RB-L-IT4: result of the fourth iteration of the 200 MHz FWI based on the RB-L starting model), while the conductivity starting models are always homogeneous. The corresponding permittivity and conductivity FWI results are presented in b) and d), respectively. The errors between the synthetic aquifer model and the FWI results are shown for c) ε_r and e) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

RB-L, the conductivity in this layer is distributed in a similar way to the RM model. RB has increased conductivity in the lower layer compared to the other results. All of the results for 450 MHz show problems with the fit in the outer region of the middle, similar to the results previously obtained for 200 MHz. The fit of the conductivity in the range between 0.9 m and 1.3 m depth is similar for all models.

The RMS, ε_r MAE, and ε_r STDAE values of the results (Table 5.3) are lower for the two frequency-hopping starting model approaches than the corresponding values of the respective standard version, while for the σ STDAE value the opposite is the case. For the σ MAE, RB-IT4 is marginally better than RB, whereas RB-L-IT4 is significantly inferior to RB-L. When comparing the RMS and the reduction of the RBI RMS (6.42e4) to the final RMS of the FWI, the use of a frequency-hopping approach is recommended. The decision to use an approach based on an RB or RB-L starting model should be retested for each dataset. It should be noted that such layers only need to be introduced if high-contrast layers are present in the subsurface, which cannot be resolved by ray-based approaches. In such cases, additional constraints on the starting model need to be fulfilled from the beginning of the FWI.

Similarly to the 200 and 450 MHz tests, for the 700 MHz FWI, the three starting models used at 200 MHz were applied again and the starting models based on frequency-hopping were updated. To do so, we chose the permittivity distributions of the fourth iteration of the 450 MHz FWI for which the corresponding IT4 starting models were used. In these two starting models (Figure 5.4a), the existing structures increase in intensity. In the FWI results for permittivity (Figure 5.4b), we can see that a successful fit to the synthetic aquifer model is only possible for the RB-IT4 and RB-L-IT4 starting models. As before,

Table 5.3.: Overview of the 450 MHz FWI results using five different starting models: smoothed (S), ray-based (RB), adapted ray-based (RB-L), and updated starting modes after 4 iterations of RB and RB-L (RB-IT4 and RB-L-IT4). The RMS, MAE, and STDAE for permittivity and conductivity are presented.

	S	RB	RB-L	RB-IT4	RB-L-IT4
RMS	3.3e-5 5.09%	4.2e-5 6.54%	3.9e-5 6.03%	2.9e-5 4.47%	3.1e-5 4.86%
ε_r MAE	0.23	0.28	0.27	0.23	0.24
ε_r STDAE	0.35	0.30	0.30	0.25	0.28
σ MAE	1.97	2.14	1.62	2.11	2.60
σ STDAE	4.35	4.56	3.00	4.64	6.40

some errors occur for RB-IT4 in the upper layer and for both in the waveguide. Additionally, the diagonal line that already occurred at 450 MHz is also visible here. For S, RB, and RB-L, at least one of the four criteria for reliable FWI results is violated, regardless of the perturbation and stabilization factors. This is due to a violation of the half-wavelength criterion. As already shown in Table 5.1, the restrictions are tighter at higher frequencies and thus limit the possible range of values in the starting model. For S, this is only violated in the region between 0 m and 0.6 m depth, which leads to an incorrect fit in this region, while the fit in the rest of the model is reliable, with the exception of minor errors. In addition to the range between 0 m and 0.6 m depth, the criterion is also violated in the range between 1.1 m and 1.5 m for RB and RB-L, with the result that the fitting is incorrect there. Errors in permittivity have a direct influence on conductivities (Figure 5.4d). As FWI tries to minimize the error in permittivity by adjusting the conductivities, this leads to overshooting, which can be observed for S, RB, and RB-L. The conductivity results for RB-IT4 and RB-L-IT4 show major fitting problems in the range from 0 m to 0.3 m depth (Figure 5.4e). These errors occur because the permittivities are difficult to fit in this region because of the high medium contrasts and the starting model is probably not fitting the half-wavelength criterion in this region. As a result of the multi-parameter inversion, the conductivity gradient tries to minimize the errors in this domain and hence the conductivities are overshooting. Therefore, the inversion tries to compensate for the permittivity error with conductivity, and therefore becomes trapped in a local minimum. Slight errors are present between 0.3 m and 1.5 m depth.

The misfit in S, RB, and RB-L is also reflected in the FWI results (Table 5.4). Specifically, the RMS and the reduction to the RBI RMS (1.44e3), ε_r MAE, ε_r STDAE, and σ MAE are significantly higher than in RB-IT4 and RB-L-IT4. However, it was observed that σ STDAE is also relatively high in RB-IT4 and RB-L-IT4. RB-L-IT4 shows better values than RB-IT4.

For frequencies of approximately 700 MHz, we do not recommend the use of conventional starting models, since these no longer meet the increased requirements (see Table 5.1). Instead, we propose the use of either a frequency-hopping approach or our adapted frequency-hopping approach based on a modified ray-based starting model. Note that for higher frequency data the conductivities of the FWI show more troubles to fit the input data as for lower frequencies. This behavior is caused by the nature of the multi-parameter inversion, since small changes in permittivity can result in large changes in conductivity. This is something already known from past studies (e.g., Mozaffari et al., 2020). One way

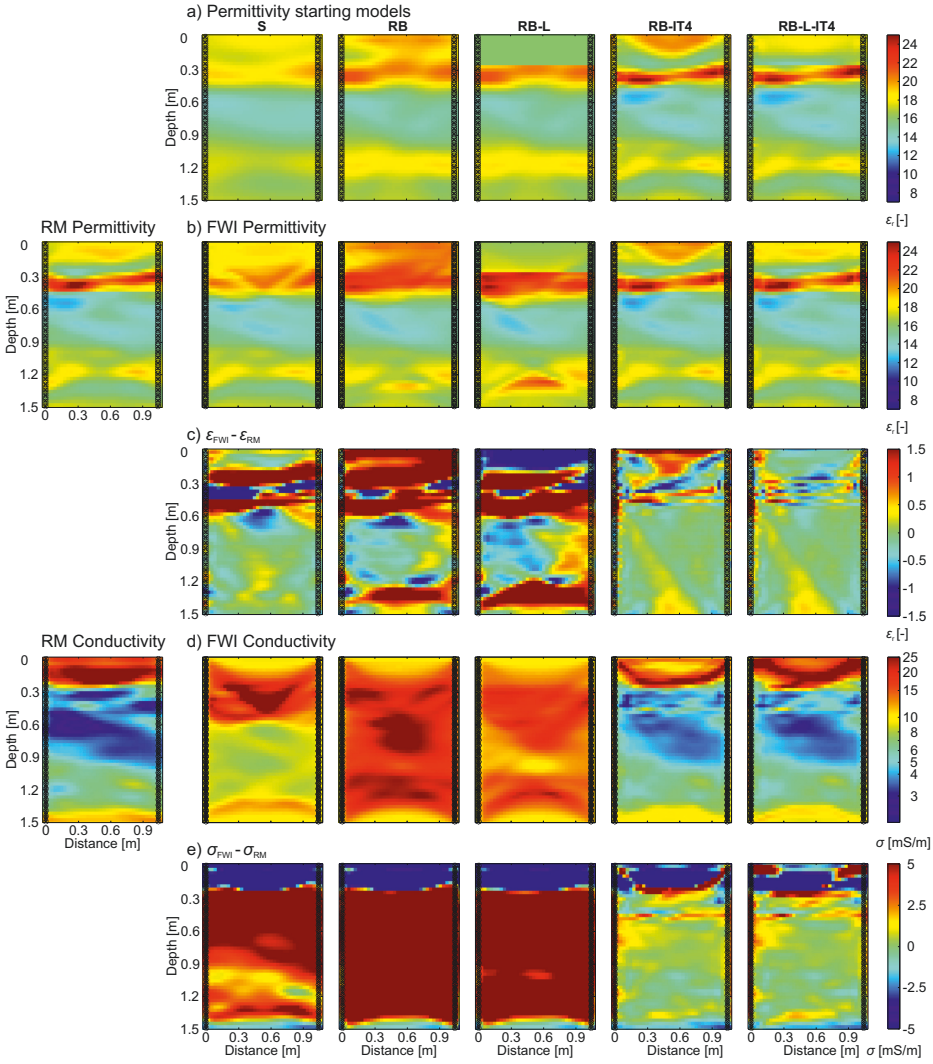

Figure 5.4.: Overview of the 700 MHz FWI results. The synthetic aquifer models (with RM signifying real model) are shown on the left. a) shows the different permittivity starting models (S: smoothed input model, RB: ray-based inversion result, RB-L: ray-based inversion result and added homogeneous layer, RB-IT4: result of the fourth iteration of the 450 MHz FWI based on the RB-IT4 starting model, RB-L-IT4: result of the fourth iteration of the 450 MHz FWI based on the RB-L-IT4 starting model), while the conductivity starting models are always homogeneous. The corresponding permittivity and conductivity FWI results are presented in b) and d), respectively. The errors between the synthetic aquifer model and the FWI results are shown for c) ε_r and e) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

Table 5.4.: Overview of the 700 MHz FWI results using five different starting models: smoothed (S), ray-based (RB), adapted ray-based (RB-L), and updated starting models after 4 iterations of RB and RB-L (RB-IT4 and RB-L-IT4). The RMS, MAE, and STDAE for permittivity and conductivity are presented.

	S	RB	RB-L	RB-IT4	RB-L-IT4
RMS	7.2e-4 49.79%	9.5e-4 65.93%	9.7e-4 67.45%	1.1e-4 7.89%	0.8e-4 5.73%
ε_r MAE	0.71	1.30	1.50	0.28	0.21
ε_r STDAE	1.10	1.50	1.63	0.35	0.24
σ MAE	7.82	14.02	10.34	2.54	2.34
σ STDAE	8.47	7.28	6.89	6.50	5.83

to improve this would be to use finer spatial sampling of the transmitter and receiver spacing, as shown by Oberröhrmann et al. (2013). Furthermore, attention should be paid to the perturbation factors for the final inversion with 700 MHz data, which range from 1e11 and 1e10 for ε_r to 5e8 and 1e7 for σ. These are approximately 2 to 4 orders of magnitude smaller than for 200 MHz data.

5.3. Effect of source wavelet variations on high-frequency FWI

An estimation of the unknown effective source wavelet is essential for FWI (for more details, see Klotzsche et al. (2019b)). Nevertheless, this can result in errors in the amplitude as well as in time shifts of the wavelet, because the estimation is based on the starting models used for the FWI. In the following section, this effect is considered first separately and then in combination with the 700 MHz effective source wavelet (Figure 5.5). Please note that for these comparisons, the RB-L-IT4 from Figure 5.4a was used as a starting model as this was shown to be the most adequate option for experimental data FWI.

For the time shift, the source wavelet was shifted $\pm 5\%$ and $\pm 1\%$ of its total length. The time shift mainly influences the permittivity. The full results are listed in Table 5.5, while Figure 5.6 shows an overview of the $\pm 5\%$ changes. The influence on the FWI permittivity results (Figure 5.6a) can be seen for the -5% time shift in the lower permittivity values. For $+5\%$, the differences from the results obtained with the original source wavelet are only visible in the difference plot (Figure 5.6b). The differences from the synthetic aquifer

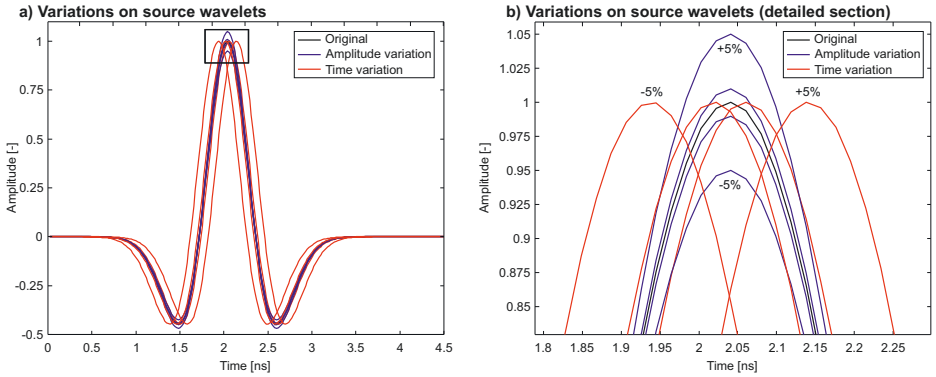

Figure 5.5.: Comparisons of the different source wavelets used to evaluate the influence of uncertainties in the effective source wavelet for high-frequency FWI in the time-domain. The source wavelets are shown in a) and b) is a detailed section, marked with a rectangle in a). The original source wavelet is based on a center frequency of 700 MHz, indicated in black, and its amplitude- and time variation are indicated in blue and red, respectively.

model increase between 0 m and 0.5 m depth and decrease between 0.5 m and 1.5 m. The conductivity clearly overshoots at -5%, whereas the results at $+5\%$ in the area between 0.3 m and 1.5 m depth are better than the original. This is due to the different results for permittivity.

The weak results for the σ MAE and σ STDAE at -5% stand out in the overview of the results (Table 5.5). These lead to the significantly increased RMS value, which is 1144.25% higher than without changes in the effective source wavelet. Also noticeable in the table are the good values for σ MAE and σ STDAE at T+5%. They decreased by 18.44% and 25.59% compared to the original source wavelet, especially since ε_r MAE is 36.06% higher than in the results obtained with the original source wavelet. This was not observed for a time shift of $+1\%$, where the results are consistently worse than those obtained with the original source wavelet.

For the amplitude variation, the amplitude of the source wavelet was varied by $\pm 5\%$ and $\pm 1\%$ of its maximum. Amplitude is strongly related to conductivity in FWI. This is visible in the FWI results obtained with the varied amplitude (Figure 5.7). The misfit in the permittivity (Figure 5.7b) shows only minor differences at the varied amplitudes. The results for the $\pm 1\%$ variation are only included in Table 5.6, since the differences are not visible. As expected, major changes in conductivity (Figure 5.7c) are visible for the $\pm 5\%$

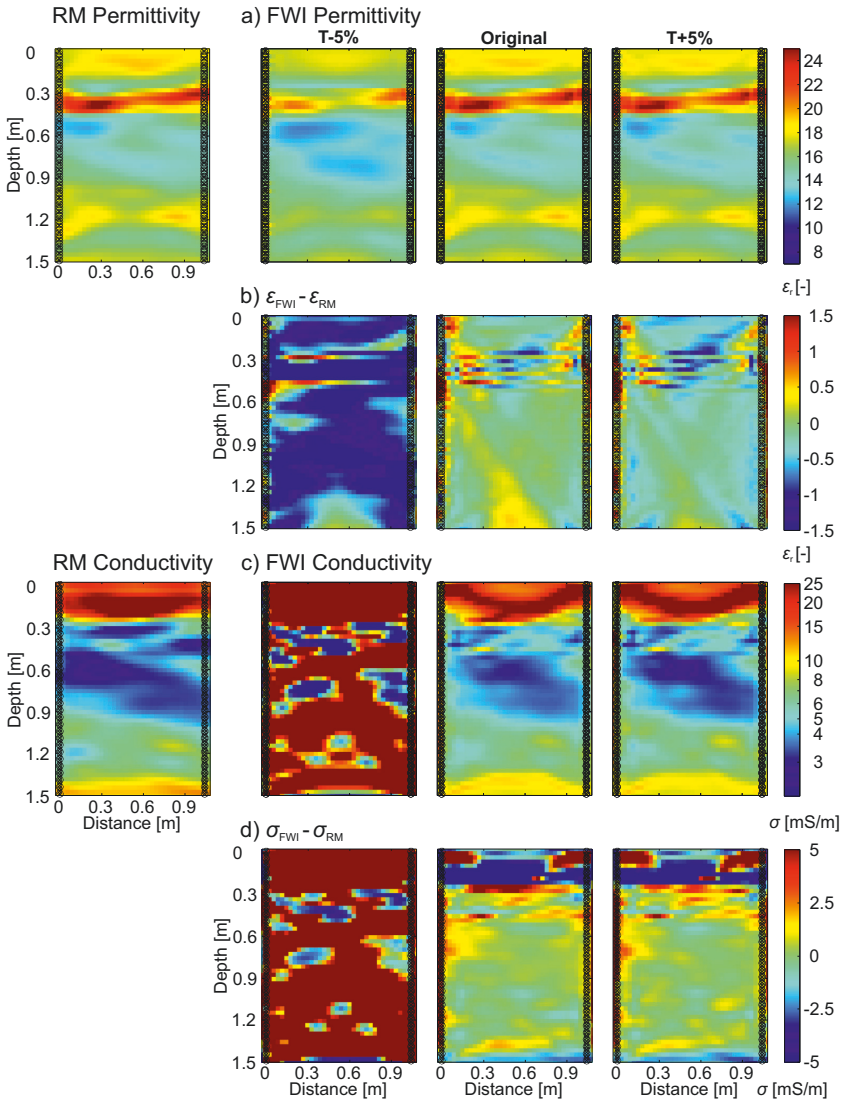

Figure 5.6.: Overview of the FWI results with time variation in the source wavelet. The synthetic aquifer models (with RM signifying real model) are shown on the left. The RB-L-IT4 model for 700 MHz was used as a starting model for the permittivity, while the conductivity starting models are always homogeneous. The permittivity and conductivity FWI results are presented in a) and c), respectively. For T-5% and T+5%, the source wavelet was shifted by 5% of its length to the left and right, while for the original, the source wavelet was not changed. The errors between the synthetic aquifer model and the FWI results are shown for b) ε_r and d) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

Table 5.5.: Overview of the time variation (T) FWI results. The percentage values indicate the difference from the values of the original wavelet varying between ±1% and ±5%. The RMS, MAE and STDAE for permittivity and conductivity for each FWI are presented.

	T-5%	T-1%	Original	T+1%	T+5%
RMS	102.9e-5 1144.25%	8.9e-5 7.13%	8.3e-5	9.1e-5 9.55%	8.5e-5 2.78%
ε_r MAE	1.21 469.53%	0.22 1.28%	0.22	0.22 4.14%	0.29 36.06%
ε_r STDAE	0.80 233.05%	0.24 1.92%	0.24	0.24 0.17%	0.22 -6.73%
σ MAE	282.44 11981.88%	2.49 6.49%	2.34	2.43 4.11%	1.91 -18.44%
σ STDAE	384.27 6494.48%	6.62 13.68%	5.83	5.99 2.74%	4.34 -25.59%

variation. For a lower amplitude, the conductivity decreases, while for a higher amplitude it increases. This occurs throughout the model but is most noticeable between 0.3 m and 1 m depth.

The comparison of the information in Table 5.6 shows that the results for the −1% amplitude are slightly better, and the RMS is 2.54% smaller, than the original for all parameters except σ STDAE. This is due to the fact that in the original, the fit in the area between 0.3 m and 1.5 m depth is slightly too weak, which results in too-high conductivities. With the reduced amplitude, the conductivity also decreases further in this range, resulting in a lower error. Since this is a multi-parameter inversion, better conductivity results can positively affect the permittivity results, which can be seen at A-1%. If the amplitude is reduced even further (−5%), the error increases. Due to the weaker fit with the original source wavelet and the shift, which leads to a better fit at −1%, the RMS at −5%with a deviation of 14.03% from the results with the original source waveletis smaller than at +5%, where the deviation is 19.7%.

The combination of time- and amplitude variation was carried out for the extrema of each case (±1% and±5%). In the following sections, only the percentages of variation are specified and in Figure 5.8, only the largest variation (±5%) is shown. The first value represents the time-, and the second the amplitude variation. The results for the permittivity (Figure 5.8a) are similar to those for the time. The misfit for the −5%−5% combination is clearly visible. For −5%+5%, the misfit is similar to a −5% variation of the time only,

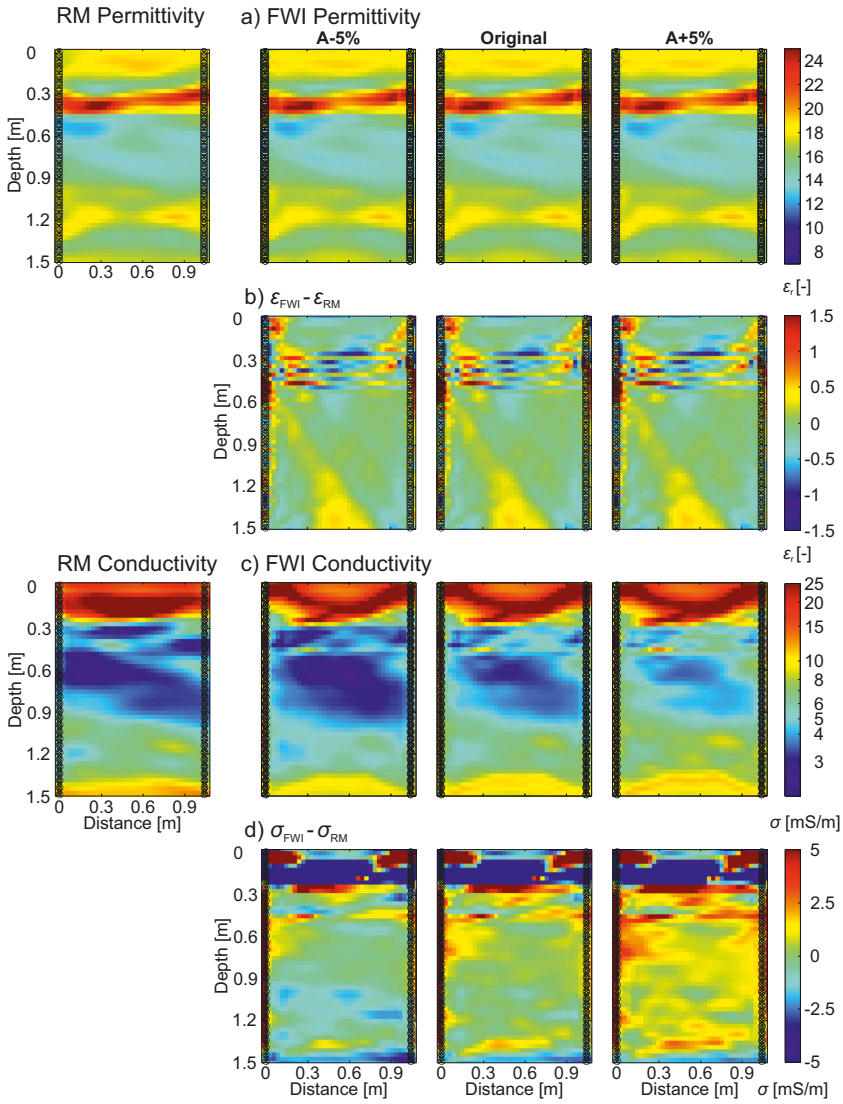

Figure 5.7.: Overview of the FWI results with amplitude variation in the source wavelet. The synthetic aquifer models (with RM signifying real model) are shown on the left. The RB-L-IT4 model for 700 MHz was used as a starting model for the permittivity, while the conductivity starting models are always homogeneous. The permittivity and conductivity FWI results are presented in a) and c), respectively. For A-5% and A+5%, the source wavelet amplitude was decreased or increased by 5% of the maximum amplitude, while for the original, the source wavelet was not changed. The errors between the synthetic aquifer model and the FWI results are shown for b) ε_r and d) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

Table 5.6.: Overview of the amplitude variation (A) FWI results. The percentage values indicate the difference from the values of the original wavelet varying between ±1% and ±5%. The RMS, MAE, and STDAE for permittivity and conductivity for each FWI are presented.

	A-5%	A-1%	Original	A+1%	A+5%
RMS	9.4e-5 14.03%	8.1e-5 -2.54%	8.3e-5	8.4e-5 1.24%	9.9e-5 19.70%
ε_r MAE	0.22 1.74%	0.21 -0.80%	0.21	0.21 0.71%	0.22 4.61%
ε_r STDAE	0.25 2.26%	0.24 -1.25%	0.24	0.24 0.50%	0.25 6.14%
σ MAE	2.54 8.46%	2.29 -0.20%	2.34	2.43 4.13%	2.97 26.85%
σ STDAE	5.90 1.17%	5.86 0.56%	5.83	6.08 4.39%	6.64 13.87%

but the waveguide is more consistent in the results for the combination. The differences for the combinations with a +5% time variation are best visible in the difference plot (Figure 5.8b). For both +5%−5% and +5%+5%, the error increases in the area between 0 m and 0.5 m depth and decreases between 0.5 m and 1.5 m depth. The results for the conductivity (Figure 5.8c) directly show the overshooting for the combination with a time shift of −5%. We can also see that the conductivity decreases for the +5%−5% combination in the area between 0.3 m and 1.5 m depth. The opposite is true for the +5%+5% combination.

The overview of the information values of the FWI results in Table 5.7 shows that the errors for the ±1% combination are similar to the basic variation. They all deviate by less than 10% compared to the original results, except for σ STDAE at −1%+1%. Note that the structures in permittivity and conductivity remain stable for the ±1% combination of both quantities. For a time shift of −5% the conductivity is overshooting and therefore a higher RMS is present. There, ε_r MAE, ε_r STDAE, σ MAE, and σ STDAE are also significantly increased. The combination of +5%−5% yields better results than using the original source wavelet for all parameters except ε_r MAE, while at +5%+5%, only σ STDAE is better than for the original source wavelet. The good results for +5%−5% are due to the fact that the −5% variation in amplitude reduces the conductivity, making it closer to the correct model, which in turn leads to a better reconstruction of the permittivity. Overall, we see an increasing mismatch with increasing phase and amplitude errors.

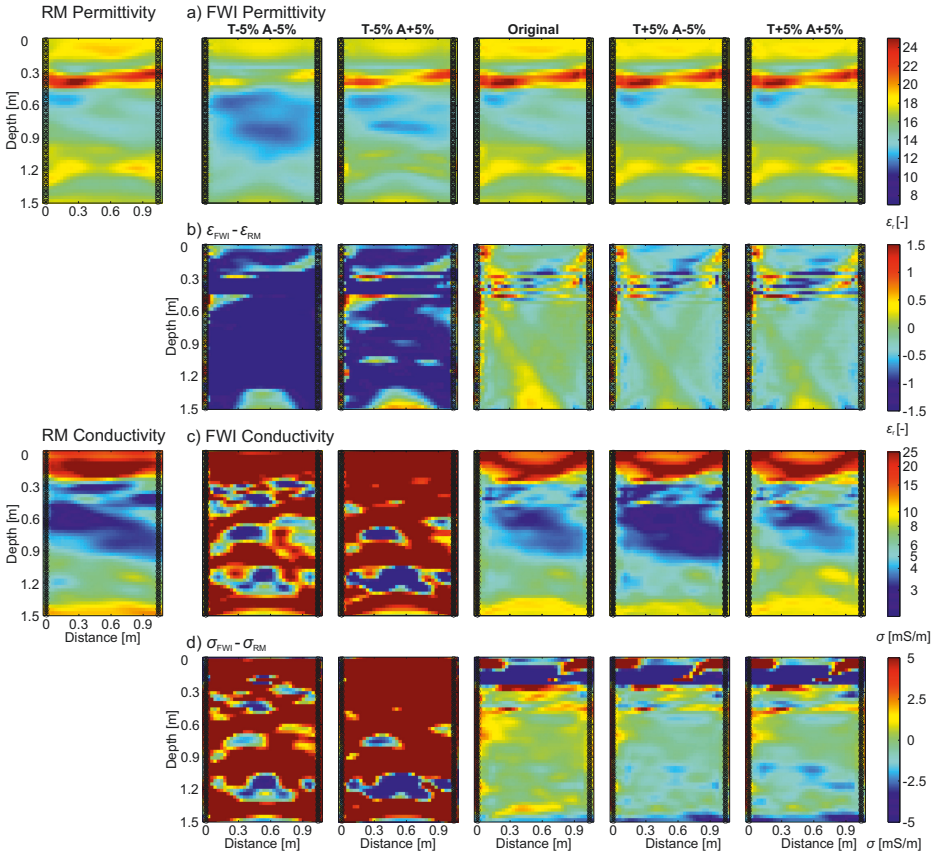

Figure 5.8.: Overview of the FWI results with a combination of time and amplitude variation in the source wavelet. The synthetic aquifer models (with RM signifying real model) are shown on the left. The RB-L-IT4 model for 700 MHz was used as a starting model for the permittivity, while the conductivity starting models are always homogeneous. The permittivity and conductivity FWI results are presented in a) and c), respectively. For T5% and T+5%, the source wavelet was shifted by 5% of its length to the left and right. For A-5% and A+5%, the source wavelet amplitude was decreased or increased by 5% of the maximum amplitude, while for the original, the source wavelet was not changed. The errors between the synthetic aquifer model and the FWI results are shown for b) ε_r and d) σ. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

Table 5.7.: Overview of the combined time and amplitude variation FWI results. The percentage values indicate the difference from the values of the original wavelet varying between ±1% and ±5%. The RMS, MAE and STDAE for permittivity and conductivity for each FWI are presented. The first value in the first line indicated the time and the second value the amplitude variations.

	-5% -5%	-5% +5%	-1% -1%	-1% +1%	Original	+1% -1%	+1% +1%	+5% -5%	+5% +5%
RMS	103.1e-5	103.1e-5	8.9e-5	8.8e-5	8.3e-5	8.9e-5	8.6e-5	8.3e-5	11.3e-5
	1146.80%	1146.68%	7.93%	6.78%		7.76%	4.19%	-0.26%	36.60%
ϵ_r MAE	2.24	1.13	0.21	0.22	0.21	0.22	0.22	0.28	0.30
	951.06%	433.00%	0.38%	1.13%		3.71%	3.76%	33.85%	40.90%
ϵ_r STDAE	1.4105	0.63	0.24	0.26	0.2393	0.24	0.24	0.21	0.25
	489.43%	165.06%	1.59%	2.47%		0.04%	-0.88%	-10.61%	3.51%
σ MAE	66.84	219.80	2.34	2.51	2.34	2.39	2.43	2.27	2.42
	2759.29%	9302.39%	-0.09%	7.36%		2.25%	4.11%	-2.93%	3.72%
σ STDAE	169.64	310.60	5.93	6.65	5.83	5.84	5.99	4.78	4.37
	2811.24%	5230.15%	1.68%	14.14%		0.25%	1.55%	-18.01%	-24.97%

5.4. Evaluation of starting models using frequency-hopping

In the first part of this study, we used an adapted and simplified version of the frequency-hopping approach as described above. This involved using two datasets with lower frequencies (200 MHz and 450 MHz) to generate a starting model for the 700 MHz data. These datasets were also created with the synthetic aquifer model. In reality, doing this would mean measuring three datasets with different frequencies for the lysimeter. Such an approach is often impossible due to the additional measurement time required and the limited frequency range of the antennas. Therefore, the results presented in this section were obtained using a frequency-hopping approach based on Zhou et al. (2021) applied to the 700 MHz GPR data. To generate data as well as effective source wavelets for lower frequencies, a low-pass filter was applied. We used a raised cosine filter with a roll-off factor of $\cos\left(\frac{\pi}{f+100\text{MHz}}\right)$, where f is the frequency to be filtered. These results had to be further smoothed with a time filter at the edges to avoid problems with the underlying forward model of the FWI, the finite-difference time-domain method. Finally, the results were normalized by dividing the maximum of the traces before filter application by the maximum after application, and then multiplying the traces after filter application with this value. Note that the applied bandpass filter reduces the amplitudes of the data, which we counterbalance with normalization of each frequency-hopping dataset. This step is necessary, to avoid that the conductivity FWI models for the filtered data will update with extremely high values to compensate for the differences in amplitude. The resulting source wavelets

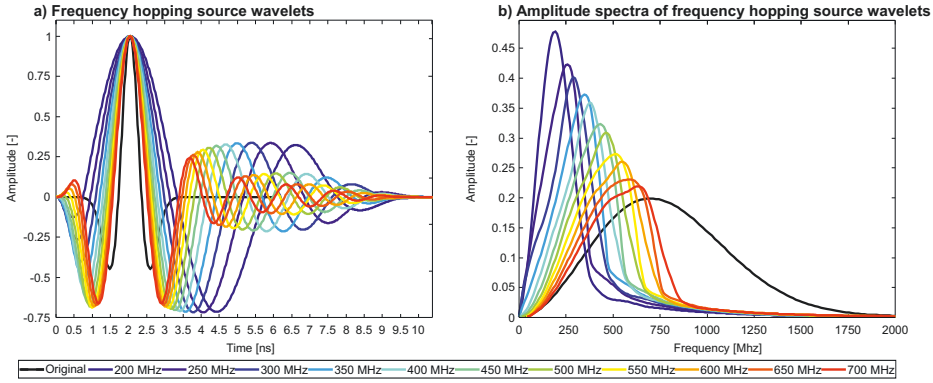

Figure 5.9.: Comparisons of the different effective source wavelets used for frequency-hopping in a) the time-domain and b) the frequency-domain. The filtered signals were normalized to the maximum of the original source wavelet in the time-domain, which results in higher amplitudes in the frequency domain.

for frequency-hopping are shown in Figure 5.9a. The normalization in the time-domain leads to a larger amplitude in the frequency-domain (Figure 5.9b), as the target frequency is smaller. The frequencies in Figure 5.9 indicate the corresponding beginning of the low-pass taper. The frequency ranges were chosen by investigating the frequency spectra of the original GPR data and the source wavelet. Note that this needs to be checked for every dataset individually since different GPR antennas, coupling conditions and the medium itself can influence the measured frequency spectra. For practical applications, we also need to consider that the frequency ranges of measured data in low loss media are reduced compared to the emitted pulse in air.

The RB and RB-L starting models were used as the starting model for the frequency-hopping. Two different frequency-hopping approaches were tested: first, constant perturbation factors with frequency steps of 50 MHz (suffix -50), and second, frequency steps of 250 MHz (suffix -250) were used. The last permittivity distributions before the final step with the complete 700 MHz data are shown in Figure 5.10a. When comparing these final starting models, it was observed that they are very similar below 0.3 m and differ mainly in the upper layer between 0.2 m and 0.3 depth. While the layer with increased permittivity is shaped like a V in the pure ray-based starting models (RB-50 and RB-250), the basic shape in the adapted starting models with an added layer (RB-L-50 and RB-L-250) is improved. In the case of RB-L-50, however, an area shaped like a V can also be seen in this layer. A similar behavior was observed in Klotzsche et al. (2012), indicating that the

starting model is trapped in a local minimum as a result of an inaccurate starting model. The waveguide is poorest resolved at RB-L-50, and the permittivity above this layer is too low. In addition, between 0.6 m and 0.9 m depth and between 0.8 m and 1 m depth, an area with a too-low permittivity can be observed, which is not visible with the other strategies. For both starting models based on 50 MHz frequency-hopping steps, it was observed that the permittivity in the range between 0.5 m and 0.6 m depth at a distance of 0.1 m to 0.4 m is lower than at the 250 MHz frequency-hopping steps. In the FWI results for permittivity, only RBL-L-250 converged towards the synthetic aquifer model in the range between 0 m and 0.2 m depth. For RB-L-50, the zone of low permittivity above the waveguide, which was already present in the starting model, is still present in the corresponding FWI results. The closest conductivity reconstruction (Figure 5.10c) to the synthetic aquifer model was achieved using RB-L-250, whereas a reconstruction with RB-L-50 led to the largest differences. Both RB starting model approaches show lower conductivity values between 0 m and 0.1 m depth. An acceptable reconstruction was not achieved for any of the results in the range between 0.2 m and 0.9 m depth, whereas a better reconstruction could be achieved between 0.9 m and 1.3 m depth.

Our investigation of the RMS, MAE, and STDAE for both permittivity and conductivity of the four results (Table 5.8) shows that all parameters are smallest for RB-L-250 and highest for RB-L-50. While RMS, permittivity, and σ STDAE values are better for RB-50, the σ MAE is lower for RB-250, which is related to the improved conductivity reconstruction in the range between 0.3 m and 0.9 m depth.

Table 5.8.: Overview of the frequency-hopping FWI results for 700 MHz data for the four different strategies using different starting models and frequency ranges. RB and RB-L indicate the ray-based and ray-based adapted models, respectively, and 50 and 250 illustrate the frequency steps used for the frequency-hopping. The RMS, MAE, and STDAE for permittivity and conductivity are presented.

	RB-50	RB-L-50	RB-250	RB-L-250
RMS	7.3e-5	14.9e-5	10.8e-5	6.8e-5
ε_r MAE	0.39	0.55	0.42	0.30
ε_r STDAE	0.35	0.58	0.40	0.27
σ MAE	2.57	4.19	2.47	1.89
σ STDAE	6.46	15.73	6.56	4.49

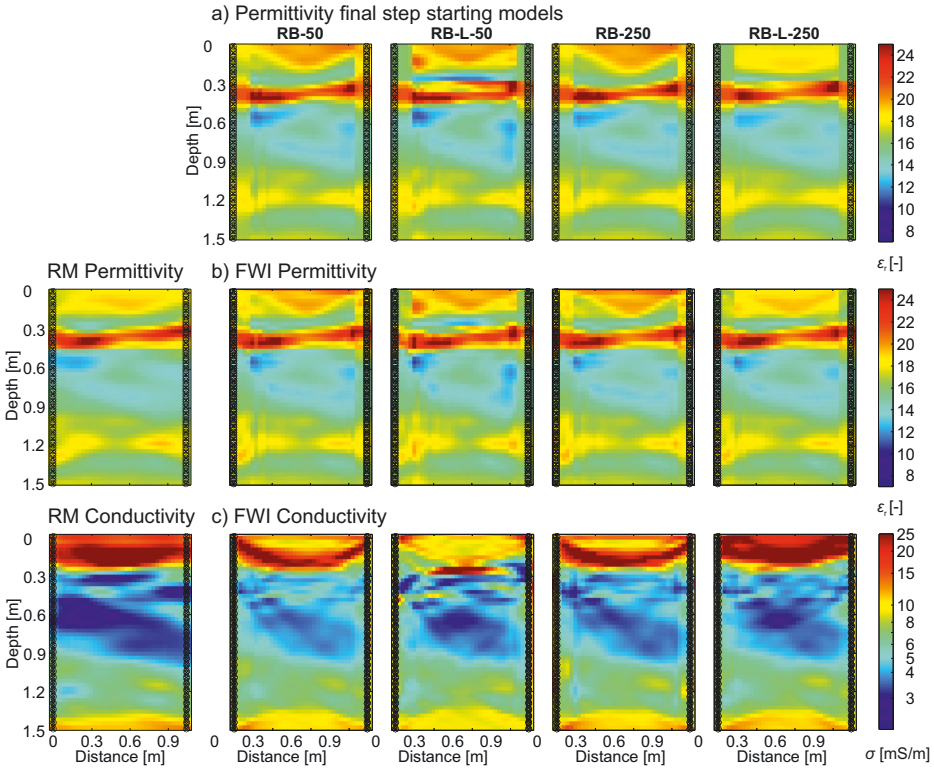

Figure 5.10.: Overview of the frequency-hopping FWI results. The synthetic aquifer models (with RM signifying real model) are shown on the left. a) shows the different permittivity starting models for the final frequency-hopping step (RB-C-50: ray-based inversion result as first starting model and constant perturbation factors for FWI steps of 50 MHz from 200 MHz to 700 MHz; RB-L-C-50: ray-based inversion result with an added homogeneous layer as first starting model and constant perturbation factors for FWI steps of 50 MHz from 200 MHz to 700 MHz; RB-A-50: ray-based inversion result as first starting model and adapted perturbation factors for each FWI step of 250 MHz from 200 MHz to 700 MHz; RB-L-A-50: ray-based inversion result with an added homogeneous layer as first starting model and adapted perturbation factors for each FWI step of 250 MHz from 200 MHz to 700 MHz), while the conductivity starting models are always homogeneous. The corresponding permittivity and conductivity FWI results are presented in b) and d), respectively. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scales for the conductivity image.

5.5. Noise in higher-frequency data

Several synthetic studies for crosshole GPR FWI investigated the effect of noise and the performance of the inversion using low-frequency data (e.g., Keskinen et al., 2021; Haruzi et al., 2022). It can be assumed that noise in the data will have a larger effect for high-frequency data than low-frequency data. Therefore, we investigate the effect of noise on the final FWI results using the best starting model approach for the 700 MHz data. Other synthetic studies either added 5% noise to the data (Meles et al., 2010) or applied different signal-to-noise ratio (S/N) levels (e.g., 25, 20 and 15 dB) to the data (Feng et al., 2021). In recent investigations for experimental higher-frequency GPR data under controlled conditions (e.g., laboratory or crosshole GPR), we did not observe a noise level higher than 25 dB. Therefore, we considered for our application a 25 dB S/N. Note that if in other scenarios a higher noise level would be observed in the GPR data, appropriate pre-processing steps are probably required to reach a similar low noise level to allow a reliable and stable FWI.

As a starting model approach, we use the previously described RB-L-250. The noise was generated using normally distributed random numbers, normalized to the range of -1 to 1, and scaled them accordingly to the S/N. To each of the 200, 450, and 700 MHz frequency filtered data and the full spectrum data, the same noise level was added. Due to the noise level, we adjusted the stopping criterion for the FWI from 0.5% to 1% change between subsequent iterations to avoid overfitting of the noise.

The FWI results in Figure 5.11 show negligible differences for the permittivity. This is because the added noise does not significantly change the phase of the data, which has an influence on the permittivity. In contrast, the results for conductivity with noise differ substantially from those without. Although the overall structures are still captured, the quantitative results are different, and, higher and lower values can be observed. Additionally, the shape of the layer between 0 m and 0.3 m depth has changed and artefacts between 0.3 m and 0.6 m depth are visible. The conductivity values between 0.4 m and 0.9 m depth are overall too low. Particularly noticeable is an area between 1 m and 1.3 m depth in the results with noise, as there is now an area with a significantly lower conductivity. The RMS of the final FWI inversion increased up to 5.00e-04 from before 0.68e-04. Since the noise mainly effects the amplitudes of the GPR data, most errors in the conductivity tomograms can be related to the noise level. Additional noise tests with lower SNR values (20 and 15 dB) showed a similar trend. Whereas the permittivity tomograms remained almost

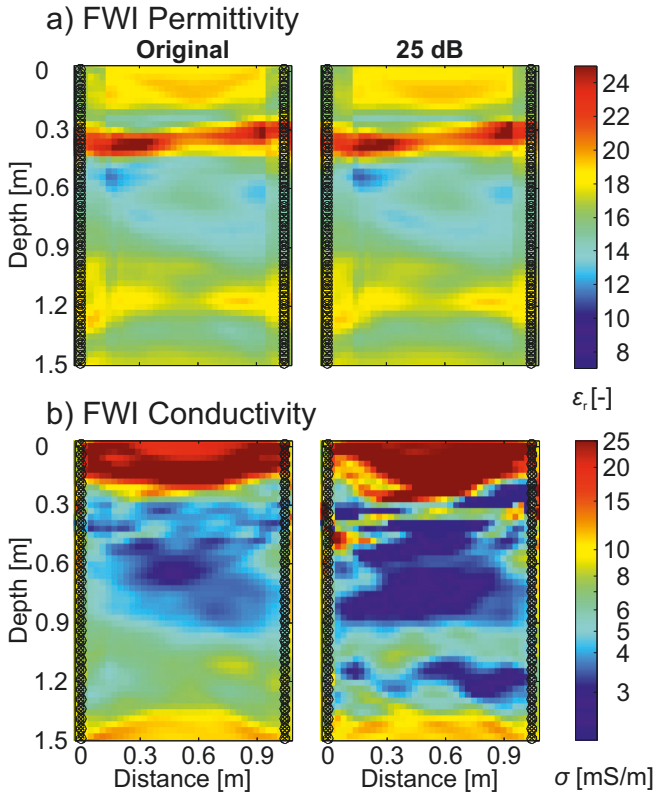

Figure 5.11.: Overview of the FWI results with the RB-L-250 approach and a signal-to-noise ratio of 25 dB. The results of the RB-L-250 approach without noise are shown on the left and on the right with noise. The permittivity and conductivity FWI results are presented in a) and b), respectively. In all tomograms, the transmitter and receiver location are indicated with circles and crosses, respectively. Note the logarithmic scale for the conductivity image.

stable, the quantitative conductivity values were less good while the structures were still visible (not shown). Note that we did not further pre-process the noisy GPR data. Therefore, we recommend for experimental data, if a higher noise level is present in the data, to pre-process the data to increase the SNR to avoid errors in the conductivity models. A possible method for the pre-processing for the frequency-hopping approach can be, for example, the use of an automatic band-pass filtering procedure as proposed by Economou, 2016, which suggests a method for band pass filtering without altering the amplitudes.

5.6. Stochastic aquifer model

One of the goals of using higher-frequency GPR data for FWI is to achieve higher-resolution images and an improved reconstruction of the subsurface properties. Therefore, we added a perturbation of the permittivity and conductivity real models with a variogram. We used a horizontal and vertical correlation length of 0.5 m and a Hurst exponent of 0.01 resulting in small-scale variations in the permittivity and conductivity images. We scaled it to a magnitude in permittivity and conductivity of 20% of the maximum permittivity and conductivity values of our original synthetic model (Figure 5.1). Afterwards, we added the scaled variogram to the original synthetic model, in order to generate an updated synthetic model with a higher spatial variability in permittivity and conductivity (Figure 5.12). Based on this, we calculated new 200 MHz, 450 MHz and 700 MHz GPR data. The new input models marked as RM (signifying real model) in Figure 5.12 show more details and variation in the models. Similar to the previous test, we performed FWI using the modified ray-based starting model with an added layer for the 200 MHz data and the simplified frequency-hopping approach with the same starting model for 700 MHz (cf. RB-L-IT4 in Figure 4). The permittivity results at 200 MHz resemble a smooth version of the RM. In contrast, the results for the 700 MHz data, show finer structures, especially between 0.4 m and 1.3 m depth. While the conductivity results for 200 MHz in the range between 0 m and 0.3 m depth are more accurate than for 700 MHz, the results between 0.3 m and 1.4 m depth are better for the 700 MHz FWI. With the 700 MHz FWI, structures with a cell size of 9 cm can apparently be resolved, whereas at 200 MHz the resolution is limited to the known decimeter-scale. To show the advantages and an improved reconstruction of the 700 MHz FWI results in more detail, we plotted a vertical cross-section through the model at a distance of 0.6 m (indicated with the dashed lines in RM in Figure 5.12). Comparing the permittivity and conductivity values of the stochastic aquifer model with the 200 MHz FWI results, relatively smooth curves can be observed. In contrast, the results with 700 MHz show a finer reconstruction of the medium properties, for example between 0.5 m and 0.9 m depth, where also small-scale changes in permittivity are resolved. Additionally, the conductivity values are closer to the values of the stochastic model than for the 200 MHz FWI results. In fact, a much closer fit to the original data can be seen in both permittivity and conductivity. Nevertheless, it should be noted that the conductivity results still provide a poorer fit to the original input values than the permittivity results. However, we advise using higher frequencies, if possible, to resolve finer structures.

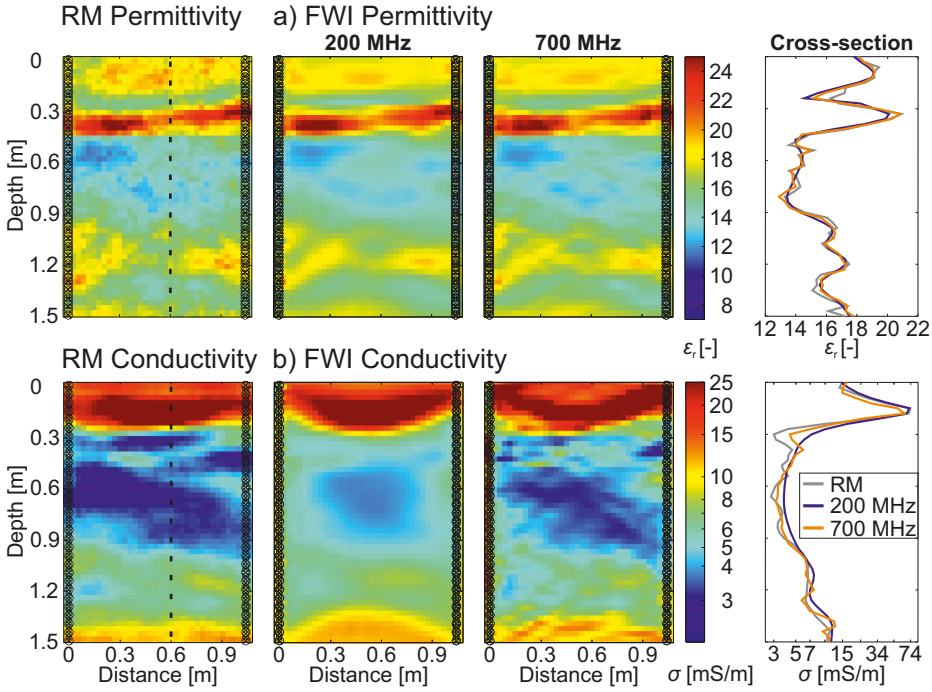

Figure 5.12.: Overview of the FWI results with a stochastically changed version of the synthetic aquifer model. The stochastic aquifer models (with RM signifying real model) are shown on the left. The ray-based inversion result for 200 MHz for the synthetic aquifer model was used for 200 MHz as starting model. The fourth iteration of this 200 MHz FWI was used as a starting model for 450 MHz, the fourth iteration of which was then used as a starting model for the 700 MHz FWI. The permittivity and conductivity FWI results for 200 MHz and 700 MHz data are presented in a) and b), respectively. Transmitter and receiver location are indicated with circles and crosses, respectively. The dashed line in the RM shows the location of the plotted cross-section of the permittivity and conductivity at a distance of 0.6 m. The gray, blue and orange lines represent the cross-sections through the three corresponding models. Note the logarithmic scale for the conductivity images.

5.7. Conclusions and outlook

In this chapter, we highlight the possible challenges and pitfalls that can occur when applying GPR FWI to higher-frequency data. Thereby, special focus was on an optimal starting model definition and the importance of an accurate effective source wavelet estimation. For our study, we considered a challenging aquifer structure including small-scale high contrast layers to prove the effectiveness of our approaches. Thereby, we demonstrated that a starting model selection based on ray-based inversion results no longer yields reliable FWI results at frequencies higher than 450 MHz since difficulties arise to fit the half-wavelength criterion. This implies that at low frequencies less accurate starting models can be used, while for higher frequencies the frequency-hopping approach should be utilized. Therefore, we adapted an existing frequency-hopping approach to create a simplified version for higher frequencies and showed its potential to produce reliable FWI results, with some limitations. When using the frequency-hopping approach in the presence of high-contrast layers, particular care still needs to be taken by adapting the ray-based starting models to reduce problems with the half-wavelength criterion. In addition, based on our test results, we recommend updating the perturbation factors for the step-length calculations for the simplified frequency-hopping approach for higher frequencies. As the frequencies increase, the perturbation factors decrease. If the chosen perturbation factors are too small, the FWI resolves fewer structures and truncation errors can emerge, while perturbation factors that are too big cause overshooting of the FWI results. In the worst case, the inversion may then be trapped in a local minimum. Note that the choice of the inversion parameters needs further investigation for experimental data and need to be adapted for different frequency ranges.

For the electrical conductivity, the medium reconstruction is more challenging since it is also more sensitive to errors with increasing frequency. Future work will investigate how the frequency-hopping approach can be applied to improve conductivity starting models. Additionally, when applying the frequency-hopping approach to other datasets with different frequency spectra, the frequency steps and the starting frequency need to be adjusted to guarantee sufficient frequency data coverage at the beginning of the FWI. Overall, we can conclude that to successfully perform FWI for higher-frequency GPR data, frequency- and time-domain methods need to be combined to enhance the starting model based on lower-frequency data.

One other critical criterion for the FWI is the estimation of the unknown effective source wavelet. Since in our synthetic study the wavelet was known, we could analyze the

effect of errors in the effective wavelet that may be caused by errors in the starting model definition resulting in shift in time and/or frequency. The analysis of the influence of the variation of the effective source wavelet at high frequencies showed that time shifts of 5% are fatal for the FWI results, while the results for +5% time shifts are still acceptable. Amplitude variations cause a lower electrical conductivity result for a lower amplitude and higher values for a higher amplitude, but still allow reliable conclusions to be drawn about the model. The multi-parameter FWI problem cannot be solved uniquely. Thus, changes in the source wavelet or even artefacts can quickly lead to problems in the inversions. Since the effective source wavelet estimation is based on the starting models and the data processing methods, for experimental data, time and amplitude errors can be present which will then affect the final FWI tomograms. Note that errors in the time-zero correction are not accounted for and cannot be validated with the FWI tomograms. Therefore, the knowledge about the errors that we have gained here can be directly linked to experimental data applications to add a further quality check for results. We would also like to point out that the validation of FWI results not only requires a consideration of individual components – such as the data fit, the final RMS, the permittivity or conductivity images, or even the statistical values – but also a careful consideration of all of these criteria together. It should be noted that we ignored signal dispersion in our study, as we did not expect it to occur in the medium and at the frequencies we used. Additional added noise to the data illustrated that mainly the conductivity results in terms of quantitative values are affected, while the differences in permittivity tomograms were minor. Therefore, to avoid errors in the conductivity models, we recommend pre-processing noisy data to increase the SNR before applying the FWI. Finally, we were able to show the better reconstruction of small-scale variations in the model by using higher frequencies with a stochastic variant of our input model. Overall, we can conclude that the GPR FWI can reconstruct finer structures in the subsurface as using lower frequencies, but also special care is needed in the choice of the starting model definition approach and to minimize the errors on the effective source wavelet. We provided guideline by applying a frequency-hopping approach, which combines both the advantages of time and frequency domain, to better constrain the half-wavelength criterion to guarantee reliable FWI results.

In a later study we want to test this new approach for high-frequency data on real data. For this we first need to investigate the influence of the lysimeter geometry on the electromagnetic signal.

6. Towards 3D GPR FWI for lysimeter investigation

We demonstrated in a synthetic study (Chapter 5) that higher frequencies can be utilized to achieve a sub decimeter-scale soil parameter reconstruction. Some imaging studies using GPR data, have already been performed on lysimeters (e.g., Stoffregen et al., 2002; Schmalholz et al., 2004). Nevertheless, none of these considered GPR FWI. This is because there are problems with the 3D-to-2D conversion, as it has difficulties to account for the late time reflections from the lysimeter geometry. As we tried to perform the 2D GPR FWI with measured lysimeter data, we faced significant challenges, as the 2D forward model could not account for many of the wave types measured at the lysimeter. Therefore, we had issues especially with the ray-based inversions as well as the estimation of the source wavelet (Havas, 2019). The inversion results, e.g., Figure 6.1 based on data measured at the lysimeter shown in Figure A.3 look promising. However, they did not meet the four criteria established by Klotzsche et al. (2019b). There was insufficient data fit, significant remaining gradients, and convergence problems. Note that we also could not solve these problems with applying the frequency hopping on the 2D GPR FWI.

To further investigate the wave types caused by the lysimeter geometry that are relevant to FWI, we first conducted simplified synthetic studies (see Appendix A). The main results of the studies were the presence of direct waves, reflections from the top/bottom and sides, and air waves traveling around the lysimeter. The separation of waves can be challenging, and interferences are observed for dry soil conditions. Additionally, an overlap of air and direct wave can cause issues in estimating the first-arrival times, which is especially challenging with Tx and Rx near the top and bottom of the lysimeter.

Since the study in Appendix A only considered homogeneous soils, we want to extend our investigation to realistic synthetic lysimeter data, based on a heterogenous soil. For this we first have to be able to generate them with a 3D lysimeter model. After that we will investigate the influence of a decreasing model complexity with heterogeneoeus soil on the simulated data. Finally, we perform a 2.5D GPR FWI on realistic synthetic lysimeter data.

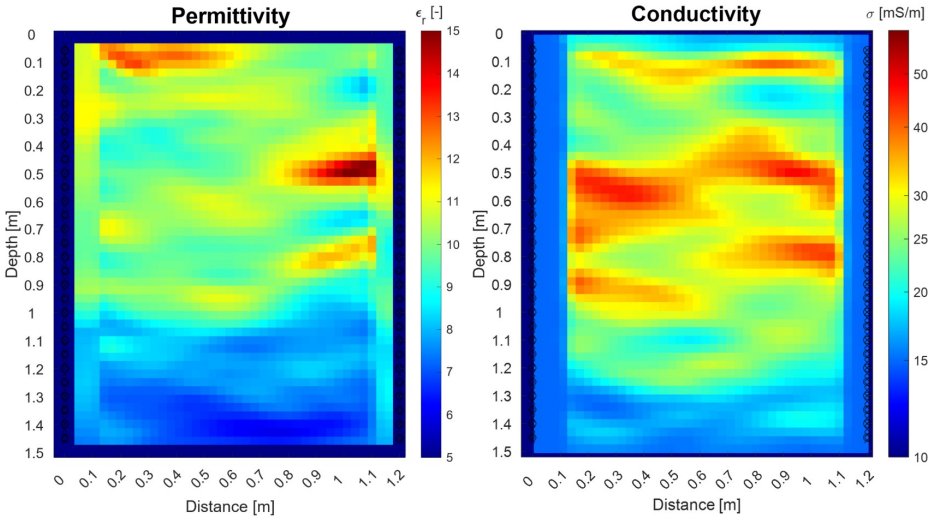

Figure 6.1.: Preliminary experimental lysimeter FWI results for a) ε_r and b) σ distribution for Kaldenkirchen soil. Tx and Rx locations are indicated with circles and crosses, respectively. Adapted from (Klotzsche, 2023).

6.1. 3D lysimeter model using 2.5D soil model

The model setup as introduced in Chapter 3 uses as the initial input a 2D permittivity and conductivity model (e.g., Figure 6.2a and b) that is extended in the perpendicular direction to build the entire 2.5D forward modeling domain. This is not suitable for the lysimeter because of its circular structure. However, it would be difficult to adjust the model building and expansion in the perpendicular plane to create a cylinder. Note that the resolution of the 2D models is 3 cm, resulting in an inversion cell size of 3 cm, while the cell size during the forward modeling is 1 cm. Furthermore, we use 180 cells in the perpendicular direction, to fully integrate the hollow cylinder that forms the lysimeter.

Therefore, we adapted the 2.5D model to realistic model lysimeter. We use the option of adding geometries, as used in Chapter 3 for borehole and antenna integration. For this we introduced the hollow cylinder as a new geometry class to gprMax, to insert hollow cylinder geometries in the 3D model (see Figure 6.2). With this it is possible to overwrite preexisting medium parameters from the general 2.5D model building process with polyvinyl chloride (PVC) ($\varepsilon_r = 4; \sigma = 1$ mS/m (derived from Mujal-Rosas et al., 2015), colored in blue in Figure 6.2c) as a lysimeter casing (3 cm thick) to form a soil column and also to add air (colored in grey in Figure 6.2c) around it. Note that we also added a 3 cm thick PVC plate

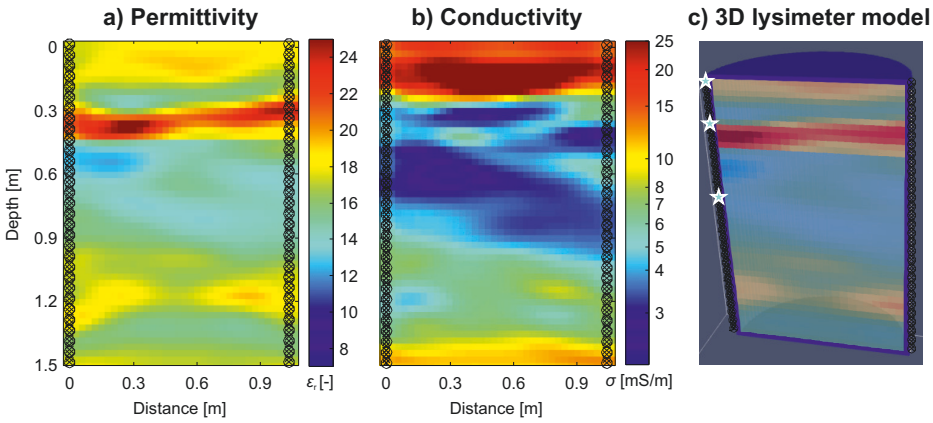

Figure 6.2.: Synthetic lysimeter model setup. a) 2D ε_r and b) 2D σ distributions for the synthetic lysimeter model, and c) a half of the 3D lysimeter model for permittivity. Note the logarithmic scale for the conductivity image. Input models are based on the Widen data (Klotzsche et al., 2012) and downscaled to the size of a lysimeter. The transmitter and receiver positions are marked with circles and crosses, respectively. We will later investigate GPR data from the transmitter marked with the star (Tx 1, Tx 6, and Tx 15).

at the top and bottom of the soil column, followed by air. We use a similar model as in Chapter 5 was used (see Figure 6.2a and b). Note that we show in Figure 6.2c only a 3D model for permittivity but we used the same setup for conductivity.

Similar to Chapter 5, the area enclosed by the Tx and Rx has a height of 1.5 m and a width of 1.08 m, and is located at the center of the modeling domain. In order to simplify the description in the model, we have set the coordinate origin to the position of the first transmitter. The Tx and Rx spacing are 0.05 m and 0.02 m for the multi-offset gathers, respectively, and we applied a reciprocal setup, which means that we switched the sides of Tx and Rx and repeated the measurements. Note that the Tx and Rx are placed in air and directly coupled to the PVC of the lysimeter. In the following, we consider a time window of 35 ns, while the time step of the forward simulation is 0.019258 ns.

6.2. Influence of decreasing model complexity

This section presents a brief study examining the impact of decreasing model complexity on the data. We tested three different scenarios which are all based on Figure 6.2a and b:

- Case A: This case represents the full lysimeter model. The 2D models have been extended to a 2.5D model, with the addition of a PVC hollow cylinder and PVC shields at the top and bottom. Air was added at the sides of the lysimeter as well as above and below (see Figure 6.3a).
- Case B: The 2D models have been extended to a 2.5D model, with the addition of a PVC hollow cylinder and PVC shields at the top and bottom. Air was added above and below the lysimeter (see Figure 6.3b).
- Case C: The 2D models have been extended to a 2.5D model, with the addition of PVC shields at the top and bottom. Air was added above and below the lysimeter (see Figure 6.3d).

Examining the GPR data, we can see that the air wave within the 9 ns to 14 ns time frame is only visible in case A(Figure 6.3a), marked with a rectangle. Additionally, there are noticeable differences between 20 ns and 25 ns, when we compare case a with case B and C (Figure 6.3b and d). While we still see a reflection from a depth of 0.4 m till 1.2 m, marked with an oval, in case B similar to case A, we can not see it in case C. Note that the reflection in case B is continuous, while it is interrupted in case A. Both cases A and B show two distinct reflections, marked by a starting at 20 ns and 25 ns up to 30 ns starting at the top and bottom moving towards the center, whereas in case C only one reflection is visible. Note that the reflections for case B are slightly earlier than for case A. Furthermore, the reflection of case C exhibits a higher amplitude. The plots in Figure 6.3c and e highlight the major differences in the model, particularly in the first-arrival between 15 ns and 20 ns, mostly in amplitude. A reflection from the top and bottom, forming a crossing, is especially visible in the difference for case C.

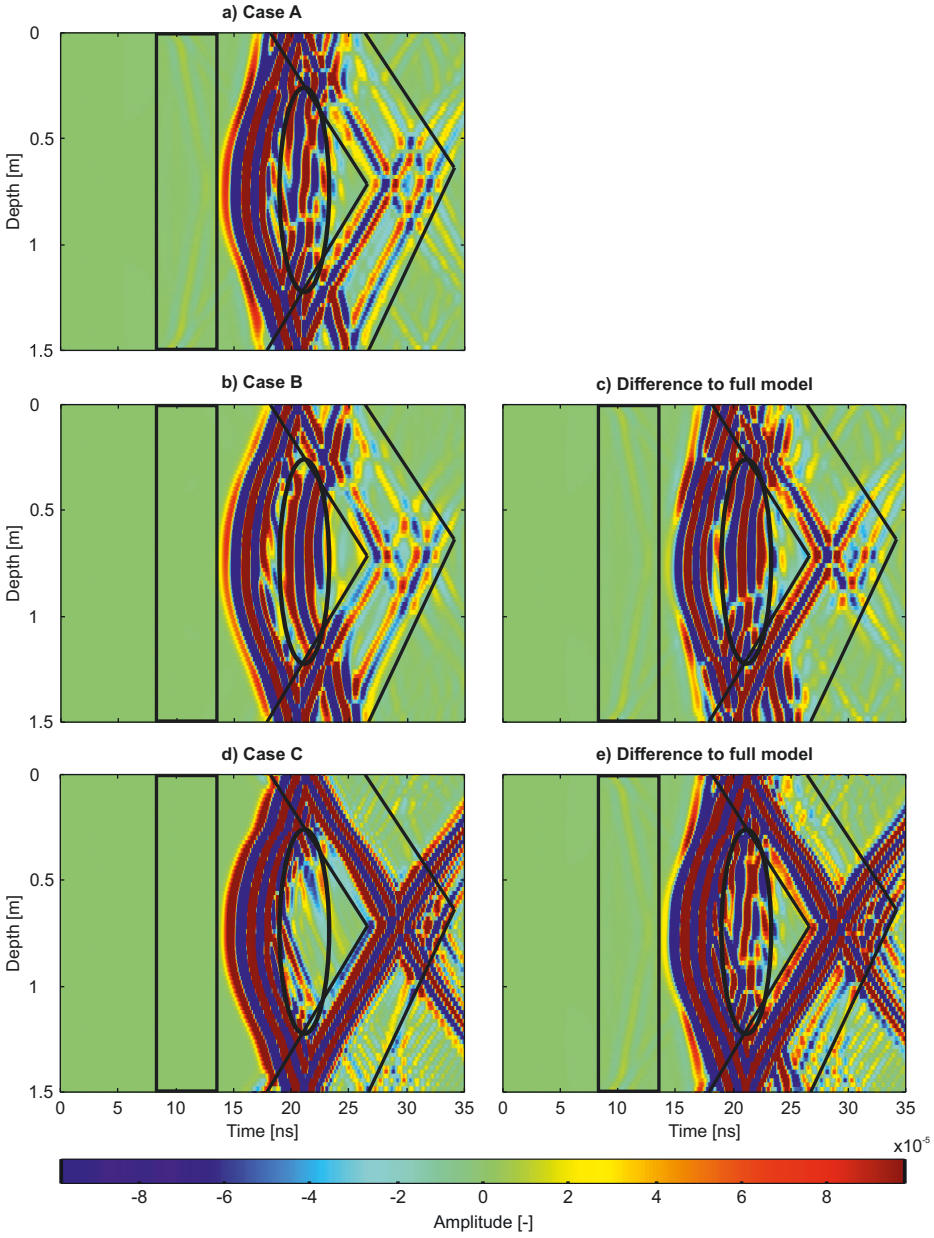

Figure 6.3.: Simulated data for transmitter 15 at the lysimeter using Figure 6.2 as a basis with reductions of model geometries. a) Case A, the full complex model, b) case B, only air at the top and bottom of the lysimeter with and including the full PVC casing, and d) case C, only air at the top and bottom of the lysimeter and only the top and bottom PVC shields. Additionally, the difference from b) and d) to a) in c) and e), respectively.

6.3. 2.5D FWI of synthetic 3D lysimeter data

As shown earlier, a full 3D model including the PVC and air around the lysimeter is necessary to reliably reproduce all wave types related to the geometry. Therefore, we included these geometries in the synthetic GPR data generation (based on Figure 6.2), as well as in the forward model of the 2.5D GPR FWI, to investigate its reconstruction capacities. We were able to show in Chapter 5, that a frequency hopping starting model may not be essential for a source wavelet with a center frequency of 450 MHz, but can improve the inversion results. Hence, we base the starting model for the 2.5D FWI on the RB-IT4 approach for 450 MHz of Section 5.2 (Figure 6.4). Note that we use the same source wavelet for the 2.5D GPR FWI as for the synthetic GPR data generation and a perturbation factor for the permittivity of 1e-2 and 5 for the conductivity.

The inversion results for permittivity show an exceptional good convergence of the starting model towards the original input model over the entire domain (Figure 6.4a). It is especially noticeable, that most parts of the curvature in the starting model between 0 m and 0.2 m has been reduced in the 2.5D FWI results. Note that the ϵ_r mean absolute error (MAE) is 0.2083 and the ϵ_r standard deviation of the absolute error (STDAE) is 0.2273. Both are exceptionally low, and no model reconstruction of the similar model without the lysimeter geometry in the 2D case (see Table 5.3) was able to yield such a good fit.

Also the conductivity shows a good convergence towards the original model (see Figure 6.4b). The thickness of the high conductivity layer between 0 m and 0.3 m depth at the center of the domain is accurately reconstructed, with a slight curve at the edges of the domain. The layer between 0.3 m and 0.9 m fits well in its intensity. In contrast to the 2D reconstructions for 450 MHz (see Figure 5.3), even at the borders of the domain. Note here, that the higher conductivity lenses in this layer at 0.4 m distance and at the right border are shifted to the left side of the domain, which could be related to the inversion parameters. The structures from 0.9 m to 1.5 m are reconstructed well. In particular, the lowest layer, between 1.4 m and 1.5 m, was better reconstructed than what was observed in the previously mentioned 2D results. Note that we use a gradient normalization. Furthermore, it is important to note the existence of vertical artefacts in the conductivity results between 0.3 m and 0.9 m at the left and right domain borders. Additional investigation is required to determine their origins. Nevertheless, the MAE for the conductivity is 1.4243 with a STDAE 2.8916. Similar to the values for permittivity, they are particularly low and yield better results than any 2D case (see Table 5.3).

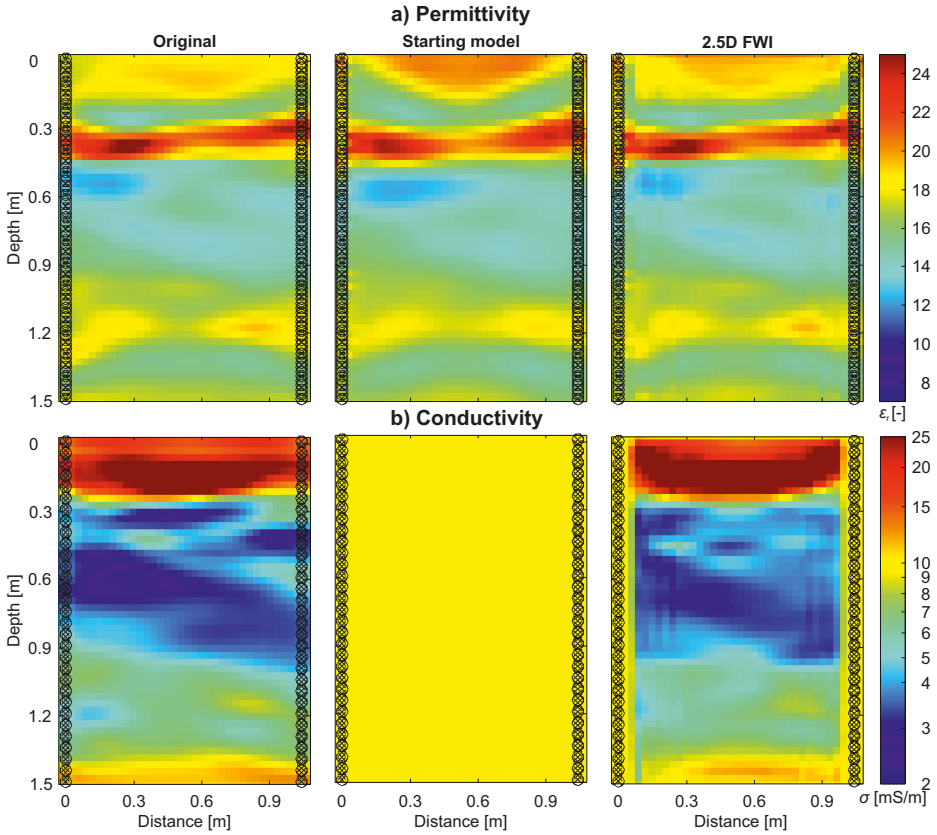

Figure 6.4.: Permittivity and conductivity distributions for the synthetic aquifer model (Original), 2.5D FWI starting model (based on Figure 5.3a) RB-IT4) and final 2.5D FWI inversion results (2.5D FWI). The transmitter and receiver positions are marked with circles and crosses, respectively.

To evaluate the model update, according to the four criteria for FWI by Klotzsche et al. (2019b), we compare the initial and final gradients (see Figure 6.5) and examine the remaining gradients. We see only a very small remaining permittivity gradient around 0.3 m depth in Figure 6.5b, indicating very good inversion results that explain the observed data well. In contrast, there is still a small gradient in the conductivity over the entire domain visible. The remaining conductivity gradient is highest between a depth of 0.3 m and 0.9 m, an area where we already saw in Figure 6.4 that there is still a mismatch with the structures. A possible explanation could be, that the perturbation factors for the step length of the conductivity are too low and should be increased to get a better convergence.

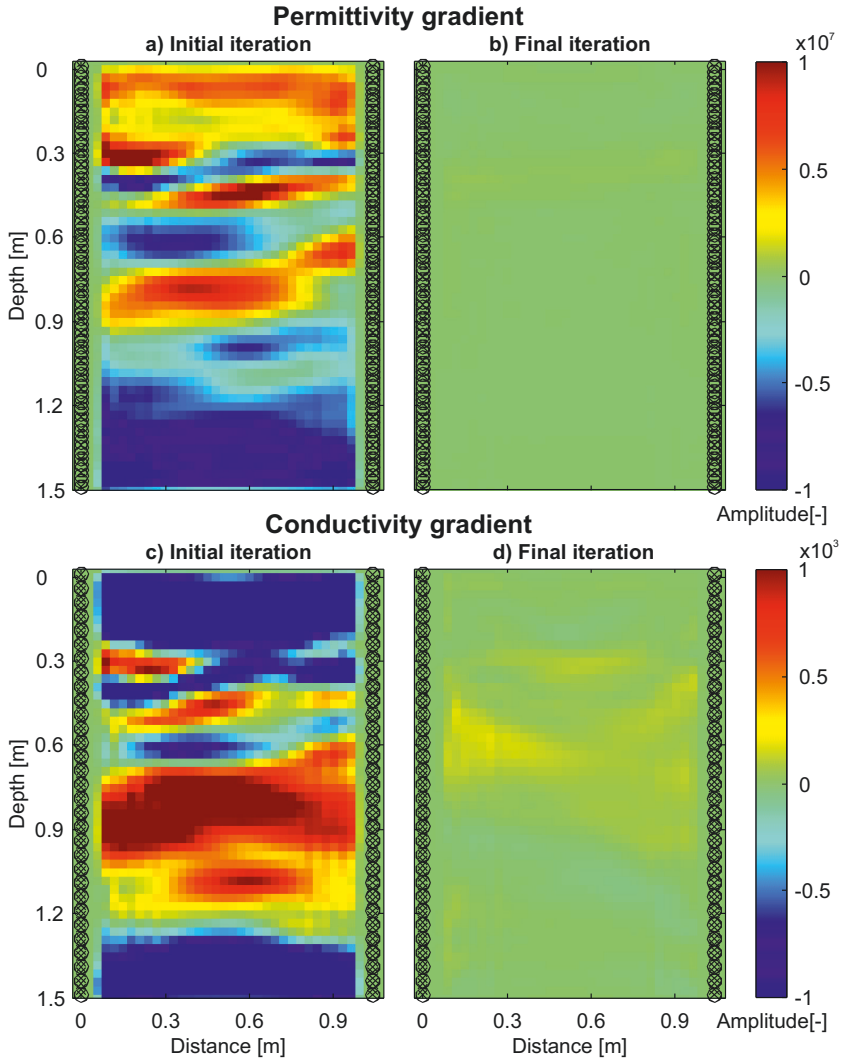

Figure 6.5.: Initial and final gradients of the 2.5D GPR FWI of a lysimeter, with the initial gradients of permittivity in a) and conductivity in c) and the final gradients of permittivity in b) and the conductivity in d). The transmitter and receiver positions are marked with circles and crosses, respectively.

The data comparison of the transmitter position at the top of the lysimeter, the high permittivity layer and the middle of the lysimeter shows the capacities of the 2.5D FWI (see Figure 6.6). The original data and the final 2.5D FWI results show a very small mismatch.

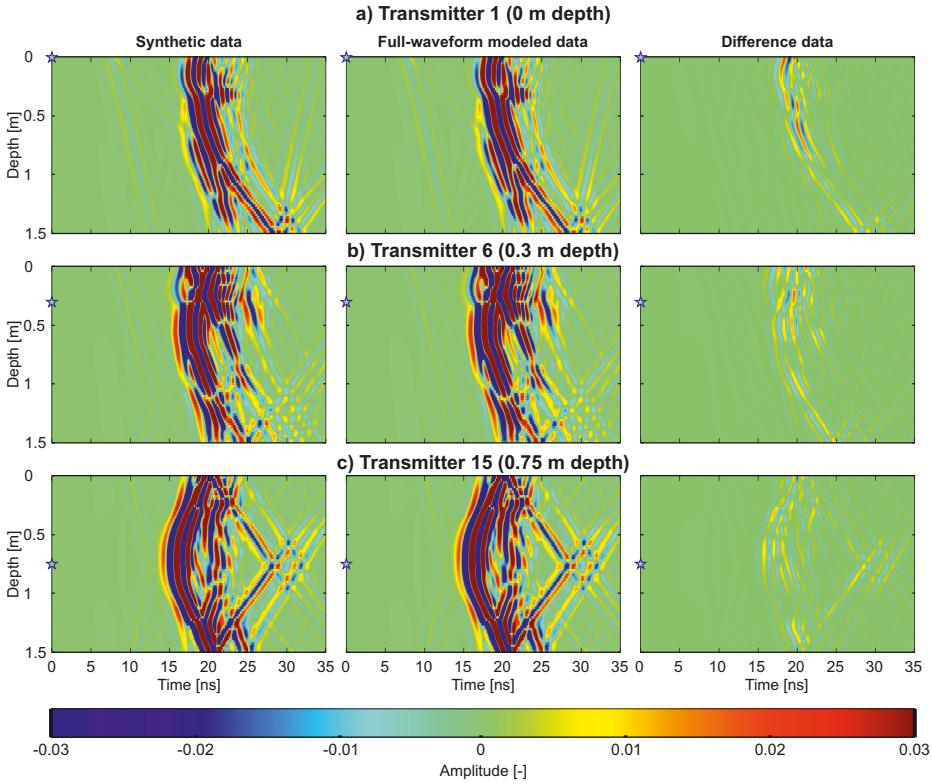

Figure 6.6.: Data comparison between original synthetic data and final 2.5D FWI data for transmitter 1, 6, and 15 (depth marked as a star). Additionally, the difference between the datasets was calculated.

Only the difference plots allow us to examine the difference of the results in detail. For all data, we can see no difference in the first 15 ns, indicating a good fit for the air waves (see Figure A.2). We can observe for both transmitter 1 and 6 (see Figure 6.6a and b) a higher error in the first 0.5 m. A detailed investigation for the individual traces (not shown) revealed, that the amplitude of the GPR data is fitted well, but there is a small time shift in the upper part, resulting in the higher errors. The errors below are based on a small amplitude difference. The overall error for transmitter 15 is low and mostly based on a small amplitude difference. Note that the coefficient of determination r^2, the square of the correlation coefficient, for all traces combined is 0.9985, indicating an exceptionally good fit to the data. This is also true for transmitters 1, 6, and 15, where it is also close to 1.

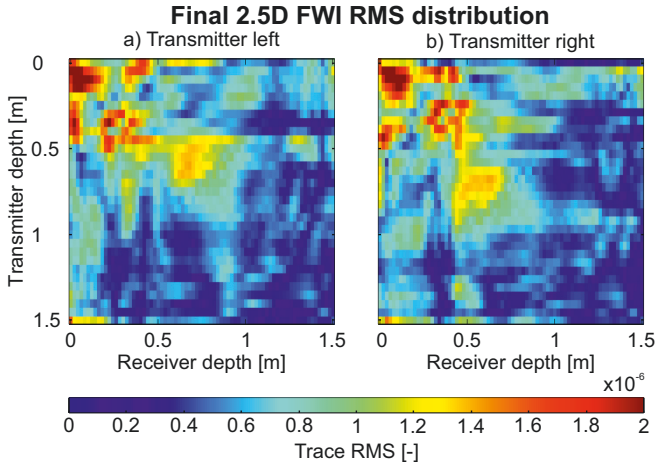

Figure 6.7.: Root mean square error (RMS) of the final 2.5D GPR FWI results for every transmitter receiver combination. a) The transmitter is placed on the left side of the lysimeter and b) on the right side.

Previously, higher errors were observed in the traces in the upper area for transmitter 1 and 6. This is also reflected in the RMS distribution (Figure 6.7), where we calculated the individual RMS for every transmitter receiver combination. A higher trace RMS is observed when the transmitter and receiver are within the first 0.5 m which can be traced back to a time shift, resulting from the permittivities. Furthermore, the trace RMS is higher and when Tx and Rx are placed more central. This can be attributed to more reflections from the top and bottom in the observed time window. A consistently lower trace RMS is visible for lower transmitter and receiver positions. Note that the total RMS was reduced from 6.71e-6 with the starting model to 5.36e-07 with the final inversion results, a reduction of 92%.

6.4. Conclusions and outlook

In this chapter, we have introduced a new geometric structure, the hollow cylinder, into the electromagnetic simulation software gprMax to invert for more complex geometrical structures like lysimeters. This allowed us to use our existing forward model construction routine and still create cylindrical objects. In addition, we successfully simulated various GPR waves, including air, direct, and reflected waves, at a lysimeter. Our results highlight

the importance of including the complete lysimeter structure, including both the PVC casing and the surrounding air, to capture all wave types. We also demonstrated the advancements with our new additions by performing a preliminary inversion of synthetic 3D GPR lysimeter data using a source wavelet with a center frequency of 450 MHz. While there is still some room for refinement in the inversion parameters, the results were remarkably good. Including the lysimeter structure within the forward model, we were able to reconstruct permittivity and conductivity in a way, that exceeded the reconstruction of the 2D case shown in Figure 5.3. Furthermore, it shows a good agreement with the data, including late arrival events.

In the future, we would like to apply this new development to measured lysimeter data, starting from low frequencies to high frequencies, demonstrating a sub decimeter-scale reconstruction with measured data. Since the forward modeling tool gprMax is based on a rectangular equidistant grid of Yee cells, a staircasing error is introduced during the implementation of cylindrical structures. Note that we have not investigated this effect on the reflection of electromagnetic waves in the cylindrical lysimeter in this study. To achieve a higher simulation accuracy, we want to investigate it and also possible solutions for implementing subgrids for the lysimeter walls to reduce these staircasing errors. Furthermore, antenna models could be included to enhance the source wavelet estimation and with this the inversion results. Finally, we want to develop this 2.5D inversion into a full 3D inversion. For this we need more GPR data, which could be acquired with antennas of a multi channel system placed all around the lysimeter (Steinbeck et al., 2023).

7. Conclusions and outlook

In this chapter, the main findings on the advances and applications of the newly developed 2.5D GPR FWI are presented, ranging from aquifer characterization to the now possible inversion of GPR data measured at lysimeters. The outlook summarizes possible future research topics.

7.1. Conclusions

This doctoral thesis presents a thorough investigation into the improvement of 2.5D ground penetrating radar (GPR) full-waveform inversion (FWI) techniques towards 3D. The goal is to enhance subsurface imaging resolution, particularly through the incorporation of complex geometries, e.g., borehole, antenna, and lysimeter geometries in forward modeling, and the use of high-frequency signals. It presents a series of focused studies to enhance our comprehension of high resolution subsurface characterization and tackles the difficulties that arise from such efforts.

In a first study, we use a synthetic 2.5D model with boreholes and antennas based on 2D GPR FWI crosshole results to test the capabilities of a newly developed 2.5D GPR FWI (Chapter 3). With this FWI it is now possible to include boreholes and finite-length antenna models in the forward model, which has become feasible through the use of subgrids implemented in the 3D forward solver gprMax. Source wavelet estimations based on different forward models showed a better reconstruction with more complex models. Therefore, we investigate the reconstructions of ray-based inversion (RBI), 2D FWI, 2.5D FWI, 2.5D FWI with borehole, and 2.5D FWI with borehole and antenna to determine the influence of borehole and antenna geometries and high angle data on the results. We observed limitations of the RBI reconstructions and negligible differences for the different FWI approaches when applied to high-angle data. The results showed that including detailed models of boreholes and antennas within the forward modeling significantly improved the reconstructions

at conductivity, particularly in zones of high contrast. Furthermore, the RMS reduction from the RBI starting model to the final FWI results increased continuously with the more complex FWI, from 72.80% up to 81.47%. However, the use of more complex models in 2.5D FWI results in significantly higher computational requirements, which highlights a necessary trade-off between computational requirements and gain in reconstruction accuracy. Therefore, we recommend using the 2.5D FWI BH for the time being if high-contrast zones are expected, particularly in conductivity. If the main interest is in investigating permittivity structures and high-contrast zones are not expected, the 2D FWI is adequate as it already provides significantly better results than the RBI and also requires significantly less computational resources than the 2.5D FWI BH.

Moving from a synthetic dataset to measured data, we evaluate in Chapter 4 the application of 2.5D modeling with boreholes and antennas and 2.5D FWI with these geometries. The dataset includes data measured in both unsaturated and saturated soils at the Widen test site in Switzerland. The analysis demonstrates an improvement in source wavelet estimations for transmitter-receiver positions where at least one of them is located in a water filled borehole. Nevertheless, we still observed challenges with the current antenna model to account for the different coupling in air filled boreholes. It indicates that it can be possible to perform 2.5D GPR FWI studies with only one source wavelet for all unsaturated-saturated transmitter-receiver position combinations if the forward model includes both boreholes and antenna models. Therefore, we peformed a 2.5D FWI with boreholes and antenna models and one source wavelet and observed similar reconstruction results as in previous studies for the saturated zone. To obtain reliable results in the unsaturated zone, it is necessary to update the antenna model.

We investigate in Chapter 5, how high-frequency GPR FWI can be used for higher resolution reconstructions. The higher frequency data causes a stricter definition of the starting model for the FWI. The RBI starting models are not sufficient anymore above 450 MHz, whereas frequency hopping approaches result in suitable starting models. We investigated the influence of time shifts in the source wavelet and observed, that time shifts of 5% cause the inversion to fail, while the results for +5% time shifts are still acceptable to obtain a good inversion result. Amplitude variations cause a lower electrical conductivity result for a lower amplitude and higher values for a higher amplitude, but still allow reliable conclusions to be drawn about the model. Finally, we were able to show an improved reconstruction of small-scale variations in the model by using higher frequencies with a stochastic variant of our input model. Overall, we can conclude that the GPR FWI can reconstruct

finer structures in the subsurface as using lower frequencies, but also special care is needed in the choice of the starting model definition approach and to minimize the errors on the effective source wavelet. We have provided guidelines for applying a frequency-hopping approach that combines the advantages of both time and frequency domains. With this approach, the half-wavelength criterion can be better constrained to ensure reliable FWI results when using high-frequency data.

In the last study, we present in Chapter 6 a novel approach for applying 2.5D GPR FWI to lysimeter data. We introduced the hollow cylinder as a new geometric structure in gprMax. By using a realistic soil model, we were able to demonstrate the ability to account for different reflections in a lysimeter. Then, we applied the 2.5D GPR FWI to realistic synthetic lysimeter data using a source wavelet with a center frequency of 450 MHz. The results demonstrate a remarkably good fit to the data. Furthermore, the results for permittivity and conductivity inversion exceeded those of the 2D case.

In conclusion, this thesis highlights the critical role of advanced 3D modeling and the use of high-frequency data in achieving high resolution GPR FWI subsurface reconstructions. This doctoral thesis offers convincing evidence to support the use of 2.5D FWI for measured data, because of superior inversion results that are critical for accurate modeling of critical zone processes. These insights provide a foundational framework for future research in high resolution subsurface imaging.

7.2. Outlook

Several key areas have been identified for future research and development. During the course of this doctoral thesis, it became clear that a particularly important goal is the optimization of the code of the 2.5D GPR FWI. A lower computational burden will increase the flexibility and scientific scope of GPR FWI and researchers without access to high-performance computers would be able to use GPR FWI. This optimization has to include a detailed evaluation of the current use of MPI and the potential of implementing more code with pre-compile techniques such as Cython or just-in-time compilation techniques such as numba. Additionally, we aim to adapt the FWI for GPUs to reduce computation time, taking advantage of gprMax's support for GPU-accelerated computation of the FDTD. It is important to note that the current GPU gprMax version does not support subgrids.

Currently, the inversion model is not updated in areas that include subgrids. This is because subgrids are currently only used around the antenna where a transmitter-receiver muting is regularly used. However, we plan to implement the inclusion of subgrids within the inversion domain. This would be particularly useful for refining the representation of the capillary fringe, which is currently simulated more abruptly than it is in reality. The use of a subgrid could reduce errors in the simulation by conducting a smoothing of the cells of the capillary fringe in the subgrid.

In all of the FWI studies conducted in this doctoral thesis and other related studies, we have observed challenges in determining appropriate perturbation factors for calculating the gradient step length. To overcome this problem, we want to investigate different seismic FWI approaches or even artificial intelligence based techniques for the step length calculation.

Another area of ongoing development is the refinement of antenna models. We plan to collaborate with experts in the field, such as Peter Annan, and explore optimization techniques for antenna models (e.g., Warren and Giannopoulos, 2011; Stadler and Igel, 2018; Giannakis et al., 2019; Stadler and Igel, 2022; Patsia et al., 2023). Specially, for a variable aquifer system (e.g., Chapter 4), we aim to achieve better results in the source wavelet estimation using an adapted antenna model. Furthermore, additional GPR data, such as those measured in the peatlands of the Ahlenmoor, Germany, can benefit from the more realistic forward model, as it increases the accuracy of the inversion.

We aim to expand the use of 2.5D GPR FWI beyond crosshole and lysimeter applications to include a wider range of GPR applications. For instance, it could be applied to horizontal boreholes, such as those installed in the minirhizotron facility in Selhausen, Germany. This facility utilized to monitor small-scale processes within the soil-plant-atmosphere continuum (Lärm et al., 2023; Lärm et al., 2024; Klotzsche et al., 2019a). Using GPR FWI would allow us to further investigate the different small-scale processes. Until now on-ground GPR data alone could not be used for a valid conductivity reconstruction with the 2D GPR FWI, as the far-field assumptions of the 3D-to-2D correction are invalid for this case. To address this issue, Domenzain et al. (2020) proposed a joint inversion of GPR and electrical resistivity data. The 3D forward model of the 2.5D GPR FWI can account for the near-field effects of the antenna. Therefore, we want to extend it for on-ground GPR data and test it on WARR measurements conducted near Thyrow, Germany. With these new options for GPR FWI, we hope to expand the use of the technique in geophysical exploration for a variety of research questions.

The next step in the development of the GPR FWI is to expand the 2.5D inversion to a full 3D inversion. This task involves investigating the calculation of new 3D GPR FWI gradients. We hope to gain more valuable information on 3D gradient calculation from the 3D seismic FWI and try to adapt it to GPR FWI. Given the increased model and inversion complexity of the full 3D GPR FWI, we expect significantly higher computational requirements. With the installation of exascale computers worldwide, we are excited to take advantage of this new computing power. The GPR data for a 3D GPR FWI could be acquired with antennas of a multichannel system placed around a lysimeter (Steinbeck et al., 2023). It is important to note that managing the large amount of measurement data using such a setup will require targeted big data management.

A. Appendix A: Using full-wave inversion techniques to image soil properties and water content at high spatial resolution at the lysimeter scale[1]

A.1. Abstract

Resource efficient management of soils requires detailed information on the spatial and temporal variation of soil properties and soil water content (SWC). Cropped lysimeters have been widely used to analyze and assess management practices. In this paper, we present theoretical analyses of using a full-wave form inversion (FWI) approach to analyze ground penetrating radar (GPR) data in order to obtain soil water content estimates in cylindrical lysimeters. We performed synthetic studies using a 3D finite-difference time-domain forward modeling tool to investigate wave propagation effects in order to better constrain FWI and thus soil water content estimates. Various scenarios of changing soil properties, different interfaces of the soil cylinder and antennae frequencies were considered. Our studies show that it is highly important to incorporate the geometry of the 3D lysimeter into the FWI inversion approach especially for dry soil conditions and higher frequency antennae. Comparison of the various wave types observed by the synthetic findings with experimental GPR data confirmed these observations. To achieve the best possible resolution of a 3D distribution of soil electromagnetic state variables (permittivity and electrical conductivity)

[1] adapted from A. Klotzsche, K. Havas, D. Hoven, and H. Vereecken (in preparation). "Using full-wave inversion techniques to image soil properties and water content at high spatial resolution at the lysimeter scale". Planned for publication in IEEE.

including soil water content, two aspects need to be considered. First the differentiation of the various wave types is easier for saturated than for dry soils, although the attenuation is increasing for saturated soils. Second, only FWI approaches can handle and use the interfering wave types and relate them to high resolution soil information such as permittivity and electrical conductivity.

A.2. Introduction

Ground penetrating radar (GPR) has the potential to provide images of soil properties and state variables such as water content (SWC) with the highest resolution among commonly near surface geophysical methods because of the applied high frequencies. This becomes important especially since the near surface stores many resources such as soil water, soil carbon and nutrients which play a key role in the challenges related to the environmental changes (Binley et al., 2015). Therefore, an improved understanding of soil structures and heterogeneities is necessary to improve agricultural management and soil-root models, among other things. Especially, small-scale heterogeneities within the subsurface can have a significant impact on, for example, flow and transport processes (Haarder et al., 2011; Binley et al., 2015) or soil-plant interactions (Blanchy et al., 2020; Lärm et al., 2024). Geophysical techniques, including electromagnetic induction (EMI), electrical resistivity tomography (ERT), and GPR have been extensively studied and employed to improve subsurface characterization and to better understand associated processes (Binley et al., 2015). Thereby, GPR utilizes high frequency electromagnetic (EM) pulses, allowing it to provide subsurface details with higher resolution compared to most other geophysical methods (Klotzsche et al., 2018). GPR can provide both the velocity v and the attenuation of the EM waves in the investigated medium through which the waves travels. The velocity and attenuation of the waves can be converted into the relative electromagnetic permittivity ϵ_r and electrical conductivity σ (Jol, 2009). These properties are strongly linked to hydrogeological parameters such as porosity, clay content and soil state variables such was water content. The versatility of GPR makes it applicable to a range of engineering, geological, hydrological and environmental applications (e.g., Paz et al. (2017)). In the past ten years, the GPR full-waveform inversion (FWI) approach has advanced significant, with widespread usage of time-domain crosshole applications. A detailed overview of these developments and the various applications is provided by Klotzsche et al. (2019b). Most of these studies used frequencies ranging between 100 and 250 MHz and successfully resolved structures on the

decimeter scale. Although FWI is a promising inversion method, it still has limitations. Data collection needs to be accurate, and the data pre-processing must be done carefully to provide sufficiently good initial models. For using higher frequencies for the FWI, especially, the starting model definition will be challenging because of the decreased wave length of the data and the requirements to meet the half- wavelength criterion. Hoven et al. (2023) performed for the first time 2D GPR FWI to investigate the opportunities and challenges associated with high frequency FWI by using synthetic data of up to 700 MHz. It was shown that defining the initial values of permittivity and electrical conductivity tomograms and the estimation of the effective source wavelet become more difficult for the higher frequency data, leading to a greater likelihood of errors in the tomograms.

Thus, in order to apply the FWI to higher frequency experimental GPR data needed to obtain the highest possible spatial resolution, we need to understand effects that should be considered in the starting model definition to guarantee a stable FWI. Therefore, a comprehensive investigation of relevant wave phenomena resulting from the setup and the requirement for 3D modeling is necessary. Lysimeters (soil columns) are frequently utilized in soil research to study flow and transport processes in soils. Lysimeters can either be packed or filled with undisturbed soil and enable flow and transport studies under controlled conditions (Pütz et al., 2018). Applications of geophysical methods to lysimeters are limited in number and were mainly focused on ERT (e.g., Garré et al. (2010) and Bechtold et al. (2011)). GPR data typically are acquired either on top of the lysimeter constructed in metal or around the lysimeter when the material is PVC. Stoffregen et al. (2002) was one of the first performing GPR measurements on top of a lysimeter to investigate SWC variations. While Schmalholz et al. (2004) employed 1000 MHz multi-offset data around the lysimeter and showed flow and transport patterns with a resolution of $\sim$0.3 m using ray-based inversion. Instead of using lysimeters, rectangular tanks, in which wave propagation can be investigated more easily, have been employed to study alterations in soils. For example, Mangel et al., 2012 applied multi-offset GPR to monitor infiltration fronts, whereas Glaser et al. (2012) examined the influence of ethanol releases on the GPR signals.

Here, we aim to investigate and improve our understanding of the factors to be considered prior to employing high frequency GPR using cylindrical lysimeter and utilizing the FWI for 3D soil characterization with a spatial resolution down to the centimeter-scale. We focus on understanding wave propagations effects related to the lysimeter casing, as well as disentangling the various types of waves and their interaction in the GPR data caused by the cylindrical setup. Therefore, high resolution 3D forward modeling is conducted to examine

the impact of varying frequencies, soil conditions (dry with no SWC and fully saturated), antenna placement, PVC casing, and model configuration. A comparison with field data demonstrates the necessity of enhancing our understanding to untangle the effects of wave propagation and analyze the data for realistic soil permittivity and electrical conductivity distributions.

A.3. Materials and methods

A.3.1. Realistic synthetic lysimeter model setup

Since full-waveform inversion approaches require a large amount of data with many different angles, this study focus on multi-offset GPR data, where the transmitter antenna (Tx) and the receiver antenna (Rx) are positioned on opposite sides of a lysimeter. While the Tx remains fixed at a specific location on the lysimeter wall, the Rx is subsequently moved to the next location with a defined constant spacing. This kind of measurement setup is referred to multi offset gather (MOG). This procedure is repeated for many different Tx locations, resulting in a dense ray-coverage in the soil column. This technique is similar to the multi-offset gathers used in crosshole scenarios (Paz et al., 2017; Klotzsche et al., 2019b). The synthetic data is simulated using the 3D EM forward model gprMax (additional details can be found in Warren et al. (2016)). This open-source software uses the FDTD technique to numerical model EM wave propagation using Maxwells equations. For the complex simulations the high-performance cluster JURECA is used (Jülich Supercomputing Centre, 2016).

The model domain is 2.2 m by 2.2 m by 2.5 m. To avoid numerical artefacts and signal dispersion especially for the high frequency data, a cell size of 0.002 m was used. A cylinder was placed in the middle of the model to represent the lysimeter with a height of 1.5 m and a diameter of 1.2 m (Figure A.1). For the tests using a PCV casing, the inner diameter was reduced to 1.16 m to consider a wall thickness of 2 cm (blue layer in Figure A.1). As permittivity of the PVC 3.3 was chosen (Table A.1). For the perfect matched layers, which are part of the model, 10 cell were implemented. As source pulse a Ricker wavelet was used with a center frequencies f_c of 500 MHz or 800 MHz. The Tx antennae was placed either at the middle or at the top part of the column's side. The Rx array is placed vertically at the opposite side of the examined soil body with a spacing of 0.02 m resulting in total in 76 traces. The material inside the lysimeter is described with either dry or saturated sand.

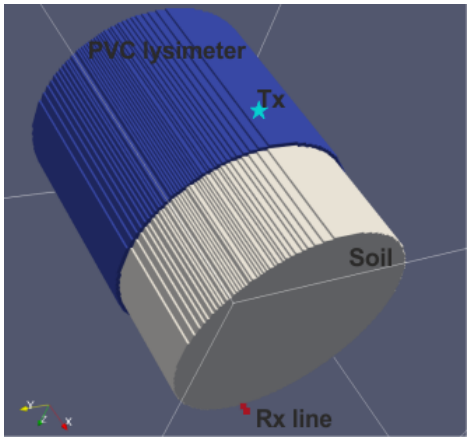

Figure A.1.: Schematic 3D geometry setup of the lysimeter model. Lysimeter height and diameter are 1.5 m and 1.2 m, respectively. PVC lysimeter's wall (blue) around the soil (grey) with transmitter (light blue star) and receiver line (red rectangular) on the opposite side of the column. Around the lysimeter the domain is filled with air and surrounded with PMLs on the side of the cube. Visualizations are performed with Paraview.

Table A.1 presents the related EM wave velocity in the soil v_{soil} and corresponding wavelength λ. As saturated sand has a higher permittivity than dry sand, the velocity and the wavelength of the EM wave are smaller. The wavelength of the saturated sand is twice that of the dry sand. Since resolution is directly related to the wavelength, structures can be better imaged in a saturated media. However, it has the drawback that the EM waves attenuate faster (Jol, 2009). Note that in experimental data for the same antennae frequency often a reduction of the f_c with increasing permittivity can be observed.

Table A.1.: Medium properties of the synthetic model. The wavelength is calculated using $\lambda = \frac{v_{soil}}{f}$ and frequency of 500 MHz.

	Dry sand	Saturated sand	PVC
ε_r [-]	4	20	3.3
v_{soil} [m/ns]	0.15	0.067	0.17
λ_{soil} [m]	0.3	0.134	0.33

A.3.2. Expected ray-paths and travel-times for the synthetic case

Considering the geometry of the lysimeter several wave types, i.e., air, direct and reflected waves, can be expected and should be investigated. Depending on the medium properties the expected travel times of the direct wave (DW), the air wave (AW), which can go around the lysimeter, and reflected waves (RW) measured at the Rx can interfere (Table A.2). The AW travels around the lysimeter walls and on top and bottom of the lysimeter in air. While the DW propagates directly through the media from the Tx to Rx, the RW are waves which are according to reflections coefficients back reflected into the medium and occur at layer boundaries such lysimeter walls. For different Tx and Rx location, the ray-path can be different and hence also reflections and refractions in the data can occur and interfere with the direct wave (Figure A.2). Note that the lysimeters wall is 2 cm PVC, which is currently not considered in these calculations of the travel times because of their minor impact but should be considered for experimental data (4 cm of PVC relates to approximately 0.24 ns). This will especially be challenging for the dry sand scenario, while for the saturated sand case the AW will arrive earlier then the DW and RW. The wave arrivals in the dry sand are earlier than in the saturated media (Table A.2). This can cause difficulties to detect the individual arrival times, as in the dry sand the direct wave (D1) arrives just a few ns later than the A1 and overlaps with A2. Reflections can occur either from the top/bottom (R1) or from the sides (R2) of the lysimeter and can also possible Interference with the DW. In most cases a back reflected direct wave should not interact with other events.

Table A.2.: Travel times TT of the expected ray-paths for a homogeneous sand model, when the Tx (500 MHz) is placed on the middle and the top part of the column. As air wave velocity 0.3 m/ns is used.

Tx position	**Wave type**	**Distance [m]**	**TT dry [ns]**	**TT saturated [ns]**
Middel	A1	1.91	6.37	6.37
Middel	A2	2.72	9.07	9.07
Middel	D1	1.22	8.13	22.33
Middel	R1	1.93	12.87	32.17
Middel	R2	1.73	11.53	28.83
Middel	3*D1	3.66	24.39	60.99
Top	A1	1.91	6.37	6.37
Top	A3	1.22	4.07	4.07
Top	D1	1.22	8.13	20.33
Top	R1	3.24	21.66	54.00
Top	R2	1.73	11.53	28.83
Top	3*D1	3.66	24.39	60.99

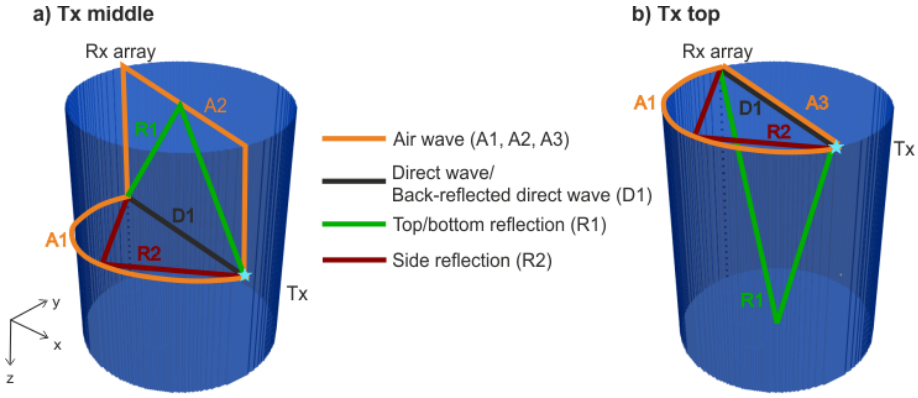

Figure A.2.: Possible ray-paths for the lysimeter setup. Ray-paths for the case when Tx and Rx are positioned a) at the middle and b) at the top of the lysimeter. Different wave types are indicated A1, A2 and A3 for the air waves, D1 as the direct wave (D1), and reflected waves from the bottom part of the column (R1) and from the side (R2).

A.3.3. Experimental design of a sand filled lysimeter

To compare our synthetic observations with experimental data, we acquired MOG GPR data acquired at an artificially sand filled lysimeter for both dry and saturated conditions. The experimental PVC lysimeter had identical dimensions as the synthetic counterpart. The bottom of the lysimeter was sealed using a perforated PVC plate to enable de-watering. The holes with roughly 1 cm diameter were arranged in three different circles. The entire construction was positioned on a PVC pallet with appropriate holes that matched those located at the bottom of the lysimeter. Geotextile was positioned at the bottom of the lysimeter to avoid loss of sand material. The lysimeter is positioned below a tent to protect it from precipitation. The lysimeter was packed with a mixture of different soil materials: fine sand (MF36, grain size 0.125-0.25 mm), coarse sand (MF31 grain size 0.25-0.5 mm), pebbles (between 2-5 cm) and moist loamy soil from Kaldenkirchen (Garré et al., 2010). Note that inside the lysimeter, two layers with structures and objects were installed. Layer 1 of objects is located at approximately a height of 0.75 m and Layer 2 at a height of 0.6 m (Figure A.3). The single objects were wrapped in geotextile to guarantee their form and locations during the filling process. Note that we chose such challenging structures to demonstrate the need

of FWI methods to construct them and to investigate the existence of complicated wave type interactions.

A.4. Results and discussion

A.4.1. Effects of lysimeter casing on wave propagation

To entangle the various wave types AW, DW and RW, models that systematically build up from a single soil column towards a full covered lysimeter are tested. Using the described model setup, homogeneous soils were tested, while three scenarios of casing were investigated to estimate the effects related to the PVC covers:

- sand column: sand column without casing and air around it,
- sand in PVC lysimeter: sand column within PVC lysimeter and air around it,
- sand in PVC ring: sand column within PVC ring (no bottom and top cover) and air around it.

A.4.1.1. Saturated sand scenario

As expected from the calculations, we see in all three scenarios first the arrival of the AW and the DW can clearly be separated (Figure A.4a-c). The AW arrives approximately after 8-9 ns, while the DW arrives at around 20-23 ns for all receivers. The different AW arrivals from different pathways are visible within the saturated system, e.g., there are diagonal small amplitude arrivals around 10-14 ns (Figure A.4). The differences between stand-alone column and PVC lysimeter or PVC ring can be seen on picture d) and e). The air waves are more pronounced as in the soil column case and slightly latter detected caused by the different permittivity of the PCV. The first detected direct wave arrives approximately 24 ns after the transmitter emitted the wave. The amplitudes are different in the middle part of the lysimeter, than on the sides, as in the middle there is a strong positive amplitude and on the upper and lower part of the column a negative amplitude can be noticed. For both cases two diminished amplitude zones in the height of approximately 0.4 and 1.2 m can be observed. This phenomenon is probably caused by negative wave interference effects, where two reflected wavefield interact and cancel each other. The reflections from the top and bottom are hardly visible (Figure A.4d-e black circles) and are only visible by

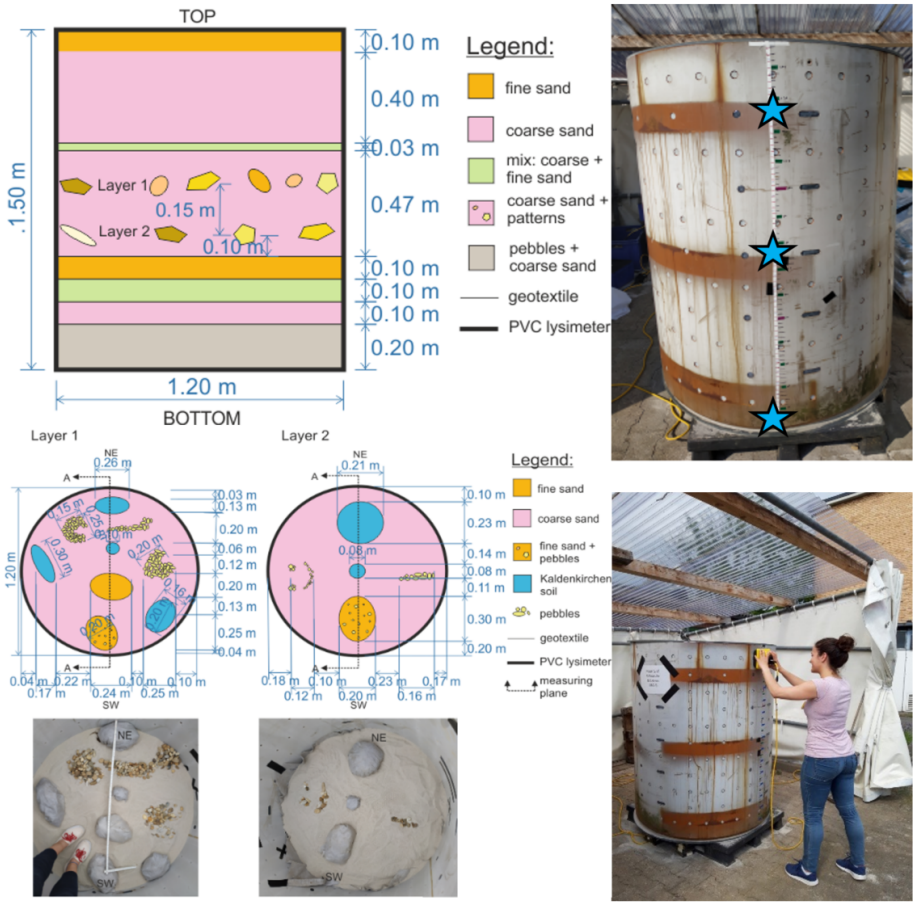

Figure A.3.: Experimental design of a lysimeter filled with different sand types and patterns. Left part shows the pattern placing in Layer 1 including a picture from the preparation of the different layers. GPR measurements were performed section AA' indicated by the dashed line. Right side shows the measurement locations marked on the outer side of the land along the entire height of the lysimeter. The distance between the marked lines is the outer diameter of the column with 1.2 m. Tx and Rx spacing for the MOG measurements was 0.05 m. Tx locations at 0.1, 0.75 and 1.05 m are indicated by blue stars.

subtracting both wavefield from each other (Figure A.4f). Thereby, we can notice that the diagonal reflections between 25-30 ns are stronger in the case of the lysimeter than without the PVC covers on the top and bottom.

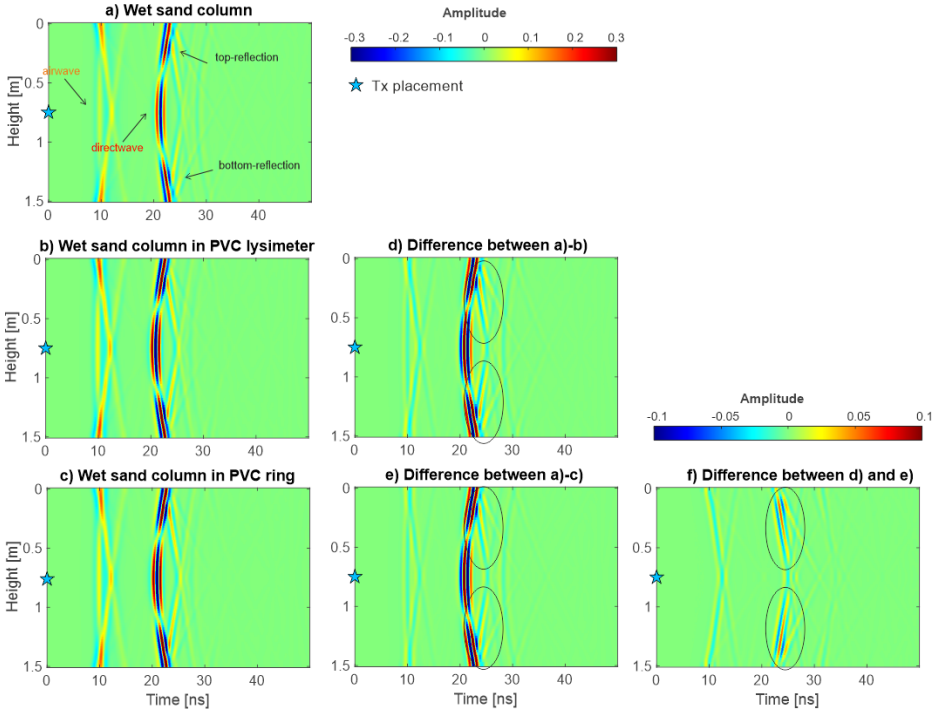

Figure A.4.: Synthetic GPR data for 500 MHz for a) saturated (wet) sand column, b) saturated sand column in PVC lysimeter and c) saturated sand column in PVC ring without bottom and top. Right side image plots represent differences in reflections between d) sand column minus sand column in PVC lysimeter and e) sand column minus sand column in PVC ring. f) Shows the difference between d) and e) to illustrate the effect of the bottom and top shielding. Tx is located at the center of the column and indicated by a star. Rx positions are placed at the opposite side using a spatial sampling interval of 0.02 m.

To understand the wave interference effects better, snapshots of the electromagnetic wave propagation for the saturated sand lysimeter case are generated (Figure A.5). The spherical shape of the wave-field can be clearly seen from time 1-21 ns, and the reflections from the side of the Tx and the bottom and top. At later times (16 and 21 ns) the interfering of the direct wave and the reflections from the boundaries are visible, where the reflections from the side and top coincide in the middle of the lysimeter with the DW. Interesting to

notice is the development of the complicated wavefield formed for the receiver close to the top and bottom of the lysimeter, where the DW, AW and RW interfere.

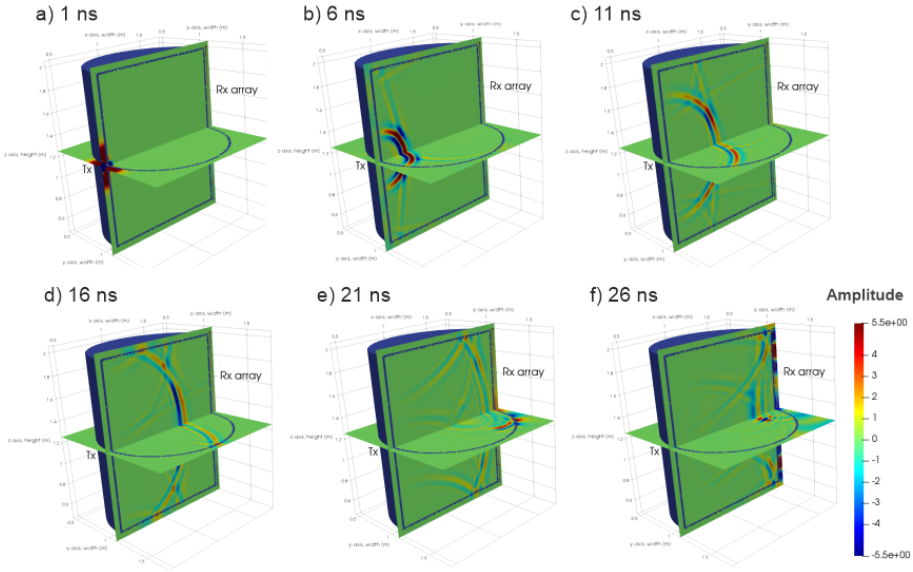

Figure A.5.: Snapshot of the electromagnetic wave propagation for the Ez field inside the saturated sand filled lysimeter using a Tx with 500 MHz positioned at the middle of the lysimeter. Snapshots are shown every 5 ns starting with 1 ns.

A.4.1.2. Dry sand scenario

While already several interferences of wave types have been observed for the saturated case, this becomes more complicated for the dry case (Figure A.6). The air waves arrive the same time as for the saturated case, but the other wave types arrive and interfere much earlier caused by the much higher velocity of the media. The AW and the DW are directly interfering and can hardly be separated. The reflections are hard to distinguish, as mostly the arrival times overlap, the air wave has the similar travel time as some of the direct waves. Only in the difference plot, the effect of the PVC covers, and casing are visible (Figure A.6d and e). The direct wave arrives around 12 ns for all three cases, as previously discussed the PVC has a small effect on the travel times (0.24 ns longer for the EM waves to travel through the PVC cover). The diagonals in Figure A.66d and e, starting from the top and bottom of the image plot at approximately 16 ns, are reflections from the lids (R1). The

amplitudes are stronger and more visible in the difference between the stand-alone dry sand column and the soil in PVC lysimeter, therefore it can be assumed that those reflections are caused by the PVC top and bottom covers (Figure A.6d). Comparing the amplitude of the stand-alone column and the soil with the PVC ring, the amplitudes of the diagonals are more reduced (Figure A.6e). The events after 28 ns are back reflected waves, which travel once back and forth in the lysimeter.

Investigating the wave propagation over time for the dry sand cases (Figure A.7), we can clearly see that the wave propagation is much faster as for the saturated case. Additionally, the first wavefront propagates between 1-10 ns, the spherical shape is nicely recognizable at the snapshot of 5 ns. The wave interferences are clearly visible at 11 and 16 ns, e.g., the interference of air waves with the direct wave. At around 21 ns, the back-reflected waves have already reached the Tx side of the lysimeter, with a reduced amplitude caused by the attenuation of the wave. Interesting to notice is the difference frequency content of the dry and saturated cases (compare Figures A.4 and A.6). The phases of the waveforms are longer for the dry cases and shorter for the saturated cases indicating a lower frequency content of the dry case data.

A.4.2. Antenna effects on wave propagation

A.4.2.1. Effects of transmitter position

This section compares the effect of the EM wave propagation and related reflections from two different Tx (500 MHz) positions. Both dry and saturated scenarios are considered for the lysimeter case with PVC cover all around the soil column (Figure A.8). When Tx is positioned in the center of the lysimeter at 0.75 m height, the reflections are symmetrical, and the earliest detected direct wave is clearly visible at the same height as the Tx (Figure A.8a and c). The upper and lower PVC covers cause the transverse reflections, which interfere with each other in the center of the lysimeter. When Tx is positioned at the top of the lysimeter, the reflected wave at the bottom exhibits diagonal arrival features (Figure A.8b and d). The air wave travels atop the lysimeter (A3), making it the fastest wave with a travel-time of about 8 ns. When Tx is positioned at the top, the bottom reflected waves have diagonal feature of arrival.

For the saturated scenario, it is possible to distinguish the AW and the RW, while in the dry scenario both wave types interfere. Especially in the dry case, it is difficult to

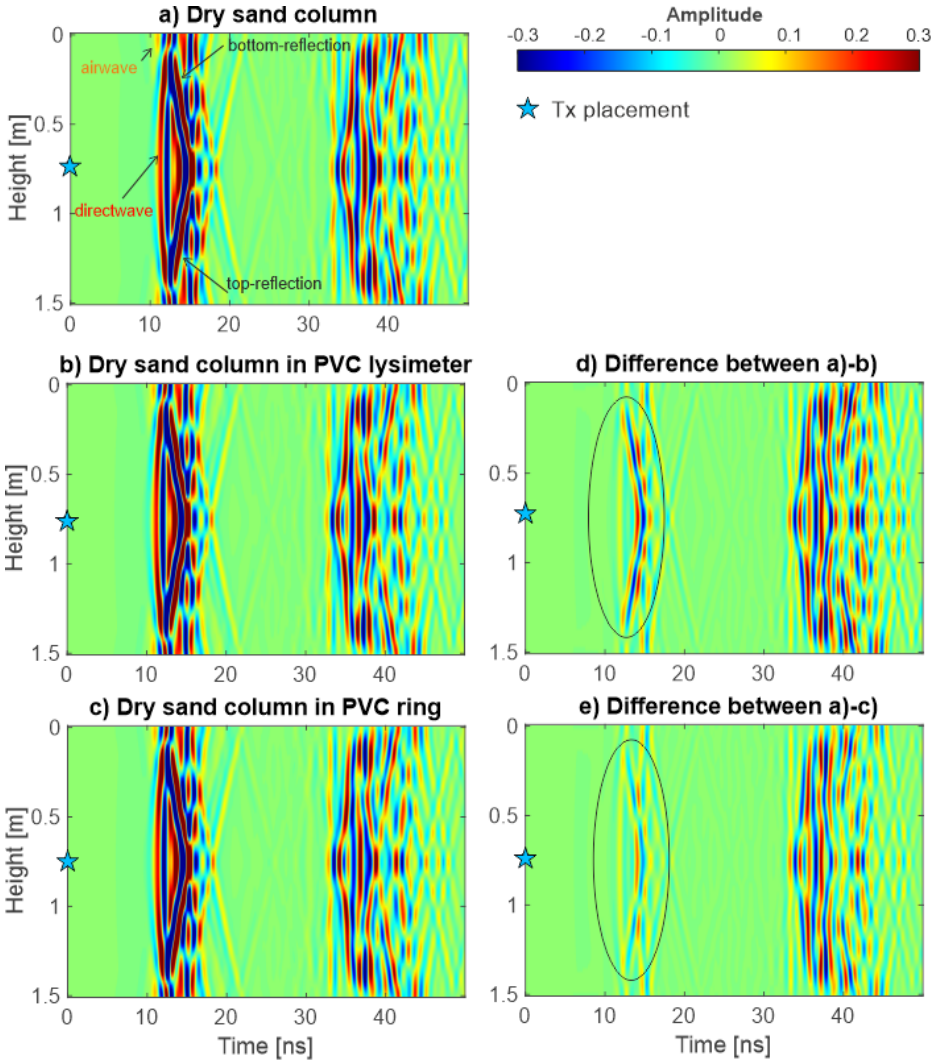

Figure A.6.: Synthetic GPR data for 500 MHz for a) dry sand column, b) dry sand column in PVC lysimeter and c) dry sand column in PVC ring without bottom and top. Right side image plots represent differences in reflections d) between sand column minus sand column in PVC lysimeter and e) between sand column minus sand column in PVC ring. Tx is located at the center of the column and indicated by a star. Rx positions are placed at the opposite side using a spatial sampling interval of 0.02 m.

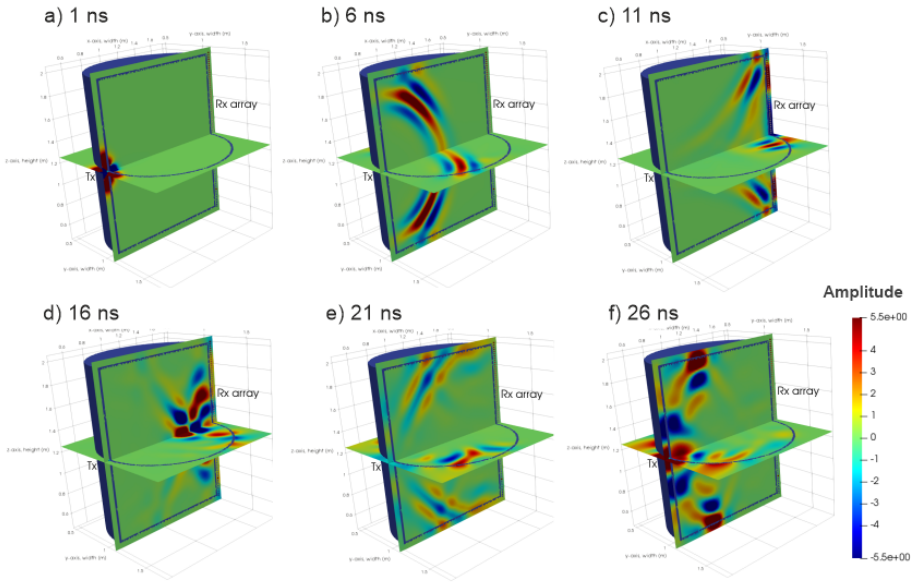

Figure A.7.: Snapshot of the electromagnetic wave propagation for the Ez field inside the dry sand filled lysimeter using a Tx with 500 MHz positioned at the middle of the lysimeter. Snapshots are shown every 5 ns starting with 1 ns.

distinguish between the direct wave and the AW below a height of 0.8 m. However, it is likely that the AW has a minor effect on the first breaks. The reflection caused by the bottom of the lysimeter starts at 15-17 ns. It could be difficult to determine the precise time since the direct and reflected waves from the bottom interfere in the lower part of the lysimeter. Furthermore, it can be observed that the reflected waves have reduced amplitudes. It is apparent that the reflected wave originated from the bottom diminishes as the Rx height increase. After 30 ns, the back-reflected waves in the dry scenario are again visible.

For the saturated scenario, it is possible to differentiate the waveforms and the first breaks from the Aws, and no impact should be anticipated. It is interesting to note that in saturated conditions, the receivers in the top 0.4 m receive signals with lower amplitudes compared to the receivers located below. Similar to the saturated case for the Tx position in the center, at a certain distance from Rx, the signals polarization switches from a positive to a negative amplitude. The same effect is visible when the Tx is situated at the top, at a height of approximately 0.5. A significant positive amplitude is detectable at the top of the lysimeter when the Tx is positioned there. The amplitude is a result of the different air

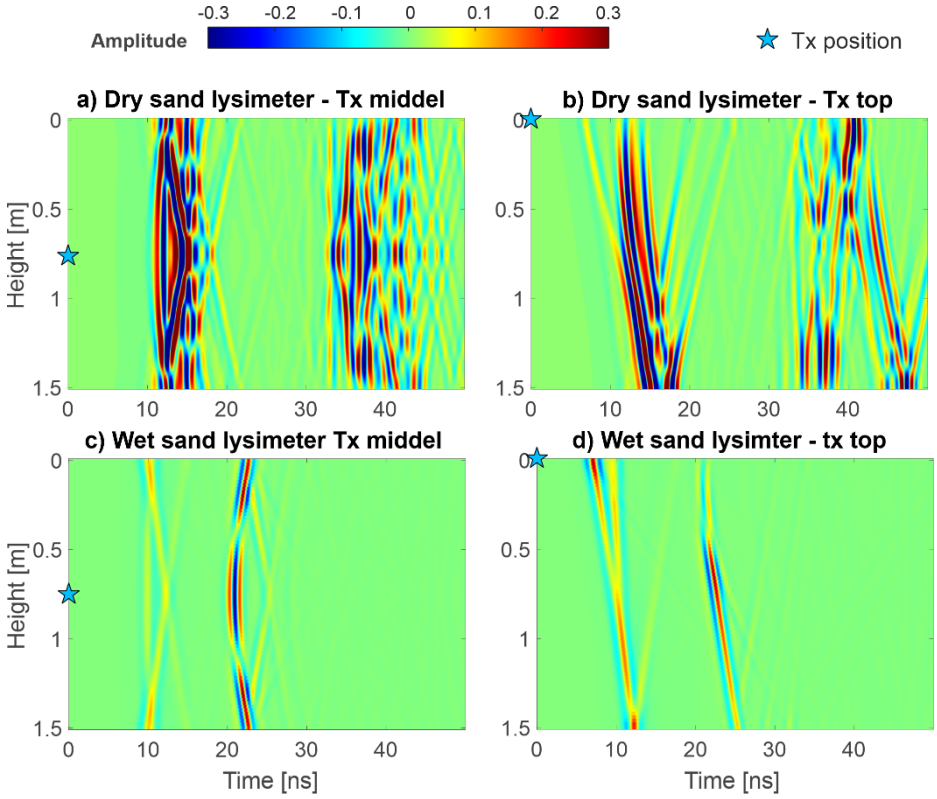

Figure A.8.: Comparison of the different locations of Tx. Synthetic GPR data for 500 MHz for a)-b) dry and c)-d) saturated conditions for Tx located in the middle and the top of the lysimeter, respectively. Rx positions are placed at the opposite side using a spatial sampling interval of 0.02 m.

wave interactions, which travel around and above the top of the lysimeter. This scenario highlights that the position of the antennae and the PVC cover have an effect of the GPR signals and must be considered.

A.4.2.2. Effects of transmitter frequency

Since we have noticed that the AW and DW cause interferences, particularly for dry scenarios, we explored whether using higher frequencies with shorter wavelengths can more effectively separate these interferences. Note that higher frequencies will result in images with higher spatial resolution. Therefore, we compare the dry and saturated scenarios for

the Tx located at the center of the lysimeter (Figure A.9). It should be noted that the different frequencies results in different delays in picking the first break caused by the slowly steady increase of the amplitude increase of the Ricker wavelet (not shown). For 500 MHz this delay is about 4 ns, and for 800 MHz, it is around 2.3 ns. Compared to the lower frequencies, higher frequencies result in a clearer distinction between the different wave events (compare Figure A.9 to Figure A.8a and c). Additionally, the AWs attenuate more quickly with the higher frequencies. When comparing the dry and saturated sand models, it can also be observed that the polarization change in the amplitudes is more pronounced in the saturated medium. In the saturated sand scenario, the gap in the data previously observed around heights of 0.4 and 1.2 m is less pronounced in the 800 MHz data due to the smaller wavelength. Overall, higher frequencies provide higher resolutions and more precise arrival-time distinction, however, they attenuate faster, especially in saturated conditions.

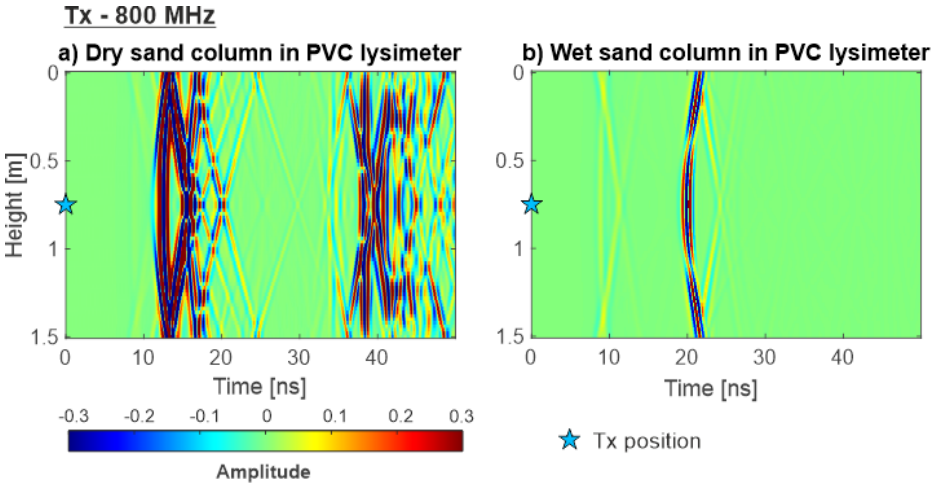

Figure A.9.: Comparison of a) dry and b) saturated wave propagation for a 800 MHz ricker wavelet for Tx located in the center of a PVC lysimeter. Rx positions are placed at the opposite side using a spatial sampling interval of 0.02 m.

A.4.3. Experimental data measured data at a sand filled lysimeter

The GPR data was measured with the 1000 MHz surface GPR antennae of the PulseEKKO Pro System (Sensors & Software Inc.). It turned out that the central frequency f_c of data was around 1500 MHz in air, therefore, we refer in the following to 1500 MHz antennae. Similar to the synthetic studies, we conducted MOG GPR measurements through NE-SW direction. First, the Tx was placed on the NE side and the Rx was moved along the SW side. Afterwards the sides were changed. The sampling rate dt was 0.05 ns and the time widow was 30 ns. Tx and Rx spacing were 0.05 and 0.02 m, respectively. These measurements were repeated for the dry sand and for a fully saturated sand scenario. To achieve this, tap water was infiltrated on top of the lysimeter for an entire day to guarantee a fully saturated system. To investigate the effects caused by the lysimeter setup, we will investigate three positions of the MOGs along the lysimeter, marked with blue start in Figure A.3. Note only a dewow filter and a time-zero correction was applied at the GPR data.

A.4.3.1. Wave propagation for dry and saturated conditions

To compare the experimental data with the modeling, we investigated the GPR data for three positions at a height of 0.1, 0.75 and 1.05 m height at the lysimeter for dry and saturated conditions (Figure A.10). Since the contrast permittivity between dry sand ($\varepsilon_r \sim 4$) and air ($\varepsilon_r \sim 1$) is relatively small, the first arrivals for the Tx located at 0.75 m exhibit a parabolic shape, similar to the synthetic study (Figure A.10b). The transverse reflections from the top and bottom PVC covers can be identified from time 14-16 ns, which also shows a good correspondence with Figure A.7a. Note that no clear air wave which travels around the lysimeter is visible. Only for receivers located close to the bottom and the top of the lysimeter show some features of air refracted waves. The corresponding frequency spectra show two peaks, the highest is at approximately 1400 MHz and a second at around 1100 MHz (Figure A.10h). This effect is probably caused by coupling effects and the design of the antennae (personal communication with Peter Annan).

For the fully saturated sand scenario, overall later wave arrivals and clear effects caused by the layering can be noticed. Since the different materials have varying porosities, the effect of water filled pores is causing higher contrast in the layers. For example, the boundary of the two coarse sand layers at around 1 m height can be seen (Figure A.10e), while the layer between 0.5 m and 1 m height can be related to higher velocities, which

is caused by the presence of the artificial designs. For the data of the Tx located at 0.1m height (Figure A.10f), the impact of the pebbles embedded in the coarse sand can be noticed for the Rx ranging from 0-0.2m. Additionally, amplitude changes related to the different layering are present. The fully saturated data has overall a lower f_c of the data of about 800 MHz, which is also indicated by the longer waveforms of the data (Figure A.10g-i). When the lysimeter is filled with dry sand and air in the pores, the attenuation of the waves is not as strong as in a saturated environment. Nevertheless, because of the higher relative permittivity of water of $\sim$80 than air of 1, differences in porosity are easier detectable in the saturated media. Note that for the saturated sand case the f_c of the data is almost half of that of the dry scenario. Considering the related medium properties (Table A.1), the resulting wavelength is approximately the same of about 0.1 m indicating a similar possible resolution.

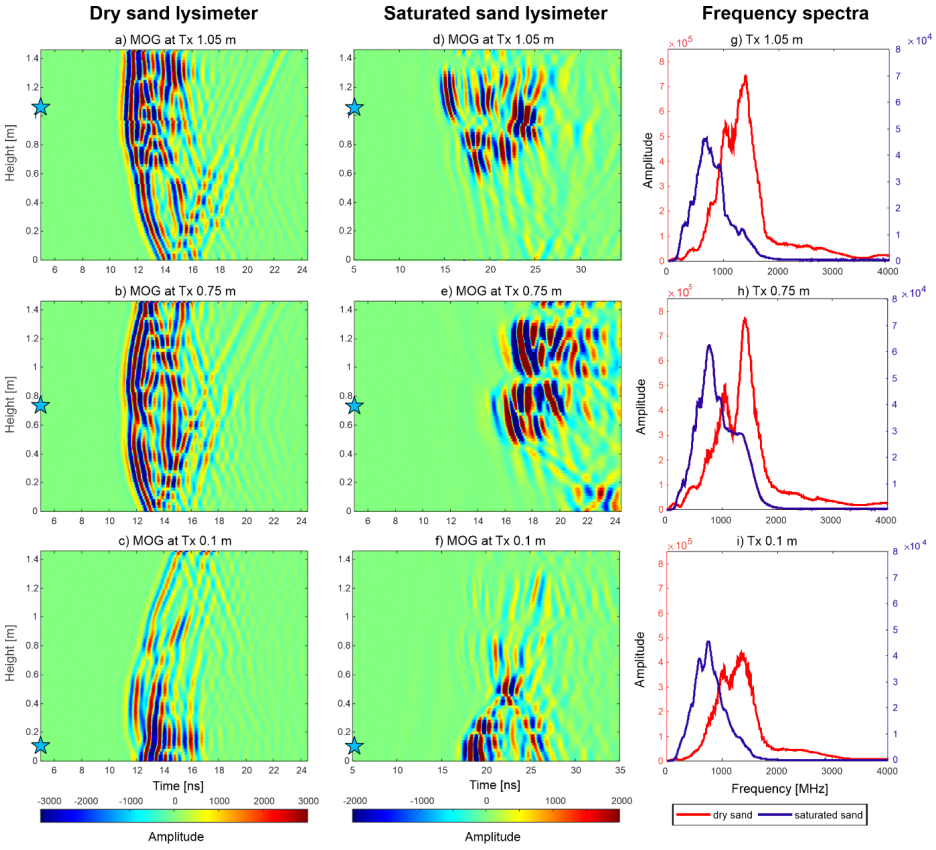

Figure A.10.: Overview of experimental GPR MOG data acquired for the dry and saturated sand lysimeter. For both scenarios three different Tx location at a), d) 1.0 m, b), e) 0.75 m and c), f) 0.1m height are plotted (indicated in Figure A.3) Corresponding frequency spectra are shown in g)-i).

A.5. Conclusion

Since the emerge of FWI approaches for GPR data, the boundaries of the resolution limits of the medium properties permittivity and electrical conductivity, and their reconstructions have been continuously improved. To explore the potential for achieving centimeter-scale high resolution and 3D reconstruction of these soil parameters and related soil states like soil water content, this study investigates the effects that should be considered while using high frequency GPR data acquired at a lysimeter. Because of the size and the possibility to position GPR antennas around the soil column, commercial surface high frequency

antennas up to the GHz range can be considered, resulting in signal wavelengths in the centimeter-scale. Caused by the special geometry of these lysimeters, various wave types are caused by reflections and refractions on the top and bottom of the lysimeter, but also from the surrounding walls. Next to direct waves, reflections from the top/bottom and sides, and air waves traveling around the lysimeter can be noticed. By investigating dry and saturated soil states, and GPR frequencies between 500 and 800 MHz, we observed that especially for dry conditions, a clear separation of wave types is difficult and wave interferences are observed. Not only that the direct wave and reflection are interfering, but also air wave and direct wave overlapping provide difficulties to estimate first arrival times of the wave which travel through the soil column.

Overall, the different wave types were best visible for fully saturated media because of the high contrast in permittivity between air and water. Furthermore, significant issues arise for the Tx and Rx near the top and bottom of the lysimeter, where airwaves which travel on top of the layer boundaries and direct waves interfere. While in high frequency data the different wave types can better be distinguished and more precise arrival-time distinction is possible, the attenuation of the waves is faster, especially in saturated conditions. This should be considered for real applications, especially for soil with an intermediate or higher electrical conductivity.

Considering all the interfering wave types it becomes evident that standard ray-based approaches, which mostly uses the first arrival times of the measured signals which travel through the soil, can hardly resolve medium properties in a lysimeter. This is especially the case for dry conditions, when the direct waves easily interference with the airwave. Note that in FWI approaches reflected and refracted waves can be fitted if the contrasts that cause these events are included in the model, e.g., top and bottom layer boundary conditions and PVC casing. Our studies show that it is very important to incorporate the lysimeter geometries in the 3D model for the FWI inversion approach to obtain reliable inversion results. Generally, AW are more critical to fit since the modeling considers a point source instead of a realistic antennae model. Therefore, a setup, where AW can clearly be distinguished between DW and RW is favorable. So far, we have utilized point sources and homogeneous media in the numerical study. Further work should investigate the integration of antennae models and more complex medium parameters, which needed to be rejected in this study because of the computational burdens (Giannakis et al., 2019; Mozaffari et al., 2022; Warren and Giannopoulos, 2011).

We performed first GPR measurements using a commercial 1000 MHz antennae for a dry and a fully saturated artificially filled sand lysimeter. Overall, similar direct and reflect waves can be noticed in the GPR data as shown by the synthetic studies. Excepts on the top and bottom of the lysimeter, the air wave which travel around the lysimeter is not visible, probably because of the shielding of the antennae. Especially for the saturated scenario, various layers and changing wavefronts can be noticed caused by the different porosities and hence soil water content of the sand. Additionally, these measurements show that with increasing saturation the center frequency of the data is reduced by a factor of 2. Although higher permittivity values related to a higher saturation are better to distingue between the wave types, the attenuation of the waves is also increases. Considering the related medium properties, for both cases a wavelength of approximately 10 cm is accepted. In a next step, these data set will be analyzed using full-waveform inversion approaches, which should be able to resolve features in the centimeter-scale with such datasets. Since high frequency data required adaption especially in terms of the starting model requirements to meet the half wavelength criterion (Hoven et al., 2023), this research is beyond the scope of this study, which should provide a first guideline how to handle GPR data acquired around lysimeter. To conclude, if we fully understand the resulting waveforms at the lysimeter scale, it will allow us to consider sophisticated inversion approaches such as full-waveform inversion, to improve the link between centimeter-scale high-resolution images, and, hydrological and biotic processes. For study showed that the geometry of the lysimeter should be considered in the modelling and soils with a certain saturation are more favorable to reconstruct the medium properties.

A.6. Acknowledgements

We would like to thank the John von Neumann Institute for Computing for the computing time granted on the supercomputer JURECA at the Jülich Supercomputing Centre.

Tools

The tools used in the process of writing this thesis are listed alphabetically below:

- ChatGPT
- DeepL
- iffTeX
- Matlab
- Perplexity

Acknowledgements

I am most grateful to Prof. Dr. Anja Klotzsche and Prof. Dr. Harrie-Jan Hendricks-Franssen for their roles in making my doctoral journey possible. Their profound knowledge, invaluable guidance, insightful feedback, and patient instructions have shaped my academic pursuits. I am very grateful for their generous investment of time and energy in my development as a researcher. They have helped me immensely through their support, motivation, and encouragement in times of struggle.

I would like to express my sincere appreciation to Prof. Dr. Harry Vereecken, Prof. Dr. Jan Vanderborght, and Prof. Dr. Jan van der Kruk for their precious advice, which broadened my research perspectives and helped me to overcome my research challenges.

This doctoral project would not have been possible without Dr. Craig Warren. During my visit to Newcastle upon Tyne in 2022, his expertise and support with gprMax were invaluable in coupling gprMax to the FWI. In addition, I am grateful for the opportunity to meet and form friendships with other doctoral students.

I find it challenging to fully articulate the depth of my gratitude towards everyone who devoted their time and expertise to meticulously reviewing my manuscript. Their efforts to enhance its clarity and ensure its consistency have been invaluable. Each suggestion and correction has contributed to the refinement of this thesis.

I would like to extend my sincere thanks to my friends and colleagues at IBG-3 (Forschungszentrum Jülich) for their tremendous support throughout my doctoral studies. Many delightful talks, refreshing walks, and lovely lunch breaks provided much-needed breaks from the seriousness of research.

Thanks should also go to Monika Marx for her assistance with LaTeX in enhancing the presentation of my doctoral thesis, and to Michael Knobloch for guiding me through the optimization of the 2.5D GPR FWI code using Score-P.

I would like to recognize the computational resources provided by the Jülich Supercomputing Centre (JSC, Forschungszentrum Jülich) on Jureca under grant no. cjicg41, as they were essential to my research.

I also acknowledge the ZEA-2 (Forschungszentrum Jülich) for their role in securing the Lukas funds that financed my research.

Above all, my deepest gratitude is reserved for my parents. Their constant faith in my choices and their boundless encouragement have been the foundation of my endurance and achievements.

List of Figures

List of Tables

Glossary

ABC absorbing boundary condition

CFL Courant-Friedrichs-Lewy

CMP common-midpoint

EMI electromagnetic induction

ERT electrical resistivity tomography

FDTD finite-difference time-domain

FIT finite integration technique

FWI full-waveform inversion

GPR ground-penetrating radar

GPU graphics processing unit

MAE mean absolute error

MOG multiple-offset gather

MPI message passing interface

MRAE mean relative absolute error

PDE partial differential equation

PML perfectly matched layer

PVC polyvinyl chloride

RBI ray-based inversion

RLA resistive-loaded antenna

RMS root-mean-square error

Rx Receiver

S/N signal-to-noise ratio

STDAE standard deviation of the absolute error

SWC soil water content

TDFEM time-domain finite element method

TDR time-domain reflectometry

TE transverse electric

TM transverse magnetic

Tx Transmitter

WARR wide-angle reflection and refraction

ZOP zero-offset profile

Bibliography

Amestoy, P., R. Brossier, A. Buttari, J.-Y. L'Excellent, T. Mary, L. Metivier, A. Miniussi, and S. Operto (2016). "Fast 3D frequency-domain full-waveform inversion with a parallel block low-rank multifrontal direct solver: Application to OBC data from the North Sea". In: *GEOPHYSICS* 81.6, R363–R383. DOI: 10.1190/GEO2016-0052.1.

Annan, A. P. (1993). *Practical Processing of GPR Data*. Proceedings of the Second Government Workshop on Ground Penetrating Radar. Columbus, Ohio, October 1993.

Banwart, S. A., J. Chorover, J. Gaillardet, D. Sparks, T. White, S. Anderson, A. Aufdenkampe, S. Bernasconi, S. Brantley, O. Chadwick, et al. (2013). "Sustaining Earths critical zone basic science and interdisciplinary solutions for global challenges". In: *The University of Sheffield, United Kingdom*.

Banwart, S. A., N. P. Nikolaidis, Y.-G. Zhu, C. L. Peacock, and D. L. Sparks (2019). "Soil Functions: Connecting Earth's Critical Zone". In: *Annual Review of Earth and Planetary Sciences* 47.1, pp. 333–359. DOI: 10.1146/annurev-earth-063016-020544.

Barrash, W. and P. S. Routh (Dec. 2006). "Boise Hydrogeophysical Research Site: Field-Scale Test Facility for Addressing Fundamental Questions of Environmental Science". In: *AGU Fall Meeting Abstracts*. Vol. 2006, H51D-0519, H51D–0519.

Bérenger, J.-P. (1994). "A perfectly matched layer for the absorption of electromagnetic waves". In: *Journal of Computational Physics* 114.2, pp. 185–200. ISSN: 0021-9991. DOI: 10.1006/jcph.1994.1159.

Bérenger, J.-P. (2006). "A Huygens Subgridding for the FDTD Method". In: *IEEE Transactions on Antennas and Propagation* 54.12, pp. 3797–3804. DOI: 10.1109/TAP.2006.886519.

Bechtold, M., S. Haber-Pohlmeier, J. Vanderborght, A. Pohlmeier, T. P. A. Ferré, and H. Vereecken (2011). "Near-surface solute redistribution during evaporation". In: *Geophysical Research Letters* 38.17. DOI: 10.1029/2011GL048147.

Beff, L., T. Günther, B. Vandoorne, V. Couvreur, and M. Javaux (2013). "Three-dimensional monitoring of soil water content in a maize field using Electrical Resistivity Tomogra-

phy". In: *Hydrology and Earth System Sciences* 17.2, pp. 595–609. DOI: 10.5194/hess-17-595-2013.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux (2008). "Velocity model building by 3D frequency-domain, full-waveform inversion of wide-aperture seismic data". In: *GEOPHYSICS* 73.5, VE101–VE117. DOI: 10.1190/1.2957948.

Binley, A., S. S. Hubbard, J. A. Huisman, A. Revil, D. A. Robinson, K. Singha, and L. D. Slater (2015). "The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales". In: *Water Resources Research* 51.6, pp. 3837–3866. DOI: 10.1002/2015WR017016.

Blanchy, G., C. W. Watts, J. Richards, J. Bussell, K. Huntenburg, D. L. Sparkes, M. Stalham, M. J. Hawkesford, W. R. Whalley, and A. Binley (2020). "Time-lapse geophysical assessment of agricultural practices on soil moisture dynamics". In: *Vadose Zone Journal* 19.1, e20080. DOI: 10.1002/vzj2.20080.

Bleistein, N. (1986). "Two-and-one-half dimensional in-plane wave propagation*". In: *Geophysical Prospecting* 34.5, pp. 686–703. DOI: 10.1111/j.1365-2478.1986.tb00488.x.

Blindow, N. (2005). "Bodenradar". In: *Geophysik*. Ed. by K. Knödel, H. Krummel, and G. Lange. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 389–424. ISBN: 978-3-540-26606-8. DOI: 10.1007/3-540-26606-2_6.

Boddupalli, B., T. A. Minshull, J. Morgan, G. Bayrakci, and D. Klaeschen (May 2021). "Comparison of 2-D and 3-D full waveform inversion imaging using wide-angle seismic data from the Deep Galicia Margin". In: *Geophysical Journal International* 227.1, pp. 228–256. ISSN: 0956-540X. DOI: 10.1093/gji/ggab164.

Bohlen, T. and F. Wittkamp (Feb. 2016). "Three-dimensional viscoelastic time-domain finite-difference seismic modelling using the staggered AdamsBashforth time integrator". In: *Geophysical Journal International* 204.3, pp. 1781–1788. ISSN: 0956-540X. DOI: 10.1093/gji/ggv546.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent (1995). "Multiscale seismic waveform inversion". In: *GEOPHYSICS* 60.5, pp. 1457–1473. DOI: 10.1190/1.1443880.

Busch, S., J. van der Kruk, J. Bikowski, and H. Vereecken (2012). "Quantitative conductivity and permittivity estimation using full-waveform inversion of on-ground GPR data". In: *GEOPHYSICS* 77.6, H79–H91. DOI: 10.1190/geo2012-0045.1.

Butzer, S. (2015). "3D elastic time-frequency full-waveform inversion". PhD thesis. DOI: 10.5445/IR/1000047328.

Cai, G., J. Vanderborght, A. Klotzsche, J. van der Kruk, J. Neumann, N. Hermes, and H. Vereecken (2016). "Construction of Minirhizotron Facilities for Investigating Root Zone Processes". In: *Vadose Zone Journal* 15.9, vzj2016.05.0043. DOI: 10.2136/vzj2016.05.0043.

Chew, W. and J. Lin (1995). "A frequency-hopping approach for microwave imaging of large inhomogeneous bodies". In: *IEEE Microwave and Guided Wave Letters* 5.12, pp. 439–441. DOI: 10.1109/75.481854.

Chorover, J., R. Kretzschmar, F. Garcia-Pichel, and D. L. Sparks (Oct. 2007). "Soil Biogeochemical Processes within the Critical Zone". In: *Elements* 3.5, pp. 321–326. ISSN: 1811-5209. DOI: 10.2113/gselements.3.5.321.

Coscia, I., S. A. Greenhalgh, N. Linde, J. Doetsch, L. Marescot, T. Günther, T. Vogt, and A. G. Green (2011). "3D crosshole ERT for aquifer characterization and monitoring of infiltrating river water". In: *GEOPHYSICS* 76.2, G49–G59. DOI: 10.1190/1.3553003.

Courant, R., K. Friedrichs, and H. Lewy (1928). "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Mathematische Annalen* 100, pp. 32–74. DOI: 10.1007/BF01448839.

Coyle, D. and L. Hampton (2024). "21st century progress in computing". In: *Telecommunications Policy* 48.1, p. 102649. ISSN: 0308-5961. DOI: 10.1016/j.telpol.2023.102649.

Dafflon, B., J. Irving, and W. Barrash (2011). "Inversion of multiple intersecting high-resolution crosshole GPR profiles for hydrological characterization at the Boise Hydrogeophysical Research Site". In: *Journal of Applied Geophysics* 73.4, pp. 305–314. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2011.02.001.

Domenzain, D., J. Bradford, and J. Mead (2020). "Joint inversion of full-waveform ground-penetrating radar and electrical resistivity data: Part 1". In: *GEOPHYSICS* 85.6, H97–H113. DOI: 10.1190/geo2019-0754.1.

Dubois, A., K. Belkebir, I. Catapano, and M. Saillard (2009). "Iterative solution of the electromagnetic inverse scattering problem from the transient scattered field". In: *Radio Science* 44.1. DOI: 10.1029/2007RS003765.

Economou, N. (2016). "Time-varying band-pass filtering GPR data by self-inverse filtering". In: *Near Surface Geophysics* 14.2, pp. 207–217. DOI: 10.3997/1873-0604.2015025.

Ellefsen, K. J., A. T. Mazzella, R. J. Horton, and J. R. McKenna (2011). "Phase and amplitude inversion of crosswell radar data". In: *GEOPHYSICS* 76.3, J1–J12. DOI: 10.1190/1.3554412.

Ernst, J. R., A. G. Green, H. Maurer, and K. Holliger (2007a). "Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data". In: *GEOPHYSICS* 72.5, J53–J64. DOI: 10.1190/1.2761848.

Ernst, J. R., H. Maurer, A. G. Green, and K. Holliger (2007b). "Full-Waveform Inversion of Crosshole Radar Data Based on 2-D Finite-Difference Time-Domain Solutions of Maxwell's Equations". In: *IEEE Transactions on Geoscience and Remote Sensing* 45.9, pp. 2807–2828. DOI: 10.1109/TGRS.2007.901048.

Feng, D., X. Wang, and B. Zhang (2021). "A Frequency-Domain Quasi-Newton-Based Biparameter Synchronous Imaging Scheme for Ground-Penetrating Radar With Applications in Full Waveform Inversion". In: *IEEE Transactions on Geoscience and Remote Sensing* 59.3, pp. 1949–1966. DOI: 10.1109/TGRS.2020.3004465.

Field, J. P., D. D. Breshears, D. J. Law, J. C. Villegas, L. López-Hoffman, P. D. Brooks, J. Chorover, G. A. Barron-Gafford, R. E. Gallery, M. E. Litvak, R. A. Lybrand, J. C. McIntosh, T. Meixner, G.-Y. Niu, S. A. Papuga, J. D. Pelletier, C. R. Rasmussen, and P. A. Troch (2015). "Critical Zone Services: Expanding Context, Constraints, and Currency beyond Ecosystem Services". In: *Vadose Zone Journal* 14.1, vzj2014.10.0142. DOI: 10.2136/vzj2014.10.0142.

Garré, S., M. Javaux, J. Vanderborght, L. Pagès, and H. Vereecken (2011). "Three-Dimensional Electrical Resistivity Tomography to Monitor Root Zone Water Dynamics". In: *Vadose Zone Journal* 10.1, pp. 412–424. DOI: 10.2136/vzj2010.0079.

Garré, S., J. Koestel, T. Günther, M. Javaux, J. Vanderborght, and H. Vereecken (2010). "Comparison of Heterogeneous Transport Processes Observed with Electrical Resistivity Tomography in Two Soils". In: *Vadose Zone Journal* 9.2, pp. 336–349. DOI: 10.2136/vzj2009.0086.

Giannakis, I., A. Giannopoulos, and C. Warren (2019). "Realistic FDTD GPR Antenna Models Optimized Using a Novel Linear/Nonlinear Full-Waveform Inversion". In: *IEEE Transactions on Geoscience and Remote Sensing* 57.3, pp. 1768–1778. DOI: 10.1109/TGRS.2018.2869027.

Giroux, B., E. Gloaguen, and M. Chouteau (2007). "bh_tomo-a Matlab borehole georadar 2D tomography package". In: *Computers & Geosciences* 33.1, pp. 126–137. ISSN: 0098-3004. DOI: 10.1016/j.cageo.2006.05.014.

Glaser, D., D. Werkema, R. Versteeg, R. Henderson, and D. Rucker (2012). "Temporal GPR imaging of an ethanol release within a laboratory-scaled sand tank". In: *Journal of Applied Geophysics* 86, pp. 133–145. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2012.07.016.

gprMax (Feb. 2024). *gprMax: spatial discretizazion*. https://docs.gprmax.com/en/latest/gprmodelling.html#spatial-discretization. Accessed: 10-02-2024.

Guo, L. and H. Lin (2016). "Critical Zone Research and Observatories: Current Status and Future Perspectives". In: *Vadose Zone Journal* 15.9, vzj2016.06.0050. DOI: 10.2136/vzj2016.06.0050.

Güting, N., T. Vienken, A. Klotzsche, J. van der Kruk, J. Vanderborght, J. Caers, H. Vereecken, and A. Englert (2017). "High resolution aquifer characterization using crosshole GPR full-waveform tomography: Comparison with direct-push and tracer test data". In: *Water Resources Research* 53.1, pp. 49–72. DOI: 10.1002/2016WR019498.

Haarder, E. B., M. C. Looms, K. H. Jensen, and L. Nielsen (2011). "Visualizing Unsaturated Flow Phenomena Using High-Resolution Reflection Ground Penetrating Radar". In: *Vadose Zone Journal* 10.1, pp. 84–97. DOI: 10.2136/vzj2009.0188.

Hartley, J., A. Giannopoulos, and C. Warren (2018). "A Huygens subgridding approach for efficient modelling of Ground Penetrating Radar using the Finite-Difference Time-Domain method". In: *2018 17th International Conference on Ground Penetrating Radar (GPR)*, pp. 1–6. DOI: 10.1109/ICGPR.2018.8441677.

Haruzi, P., J. Schmäck, Z. Zhou, J. van der Kruk, H. Vereecken, J. Vanderborght, and A. Klotzsche (2022). "Detection of Tracer Plumes Using Full-Waveform Inversion of Time-Lapse Ground Penetrating Radar Data: A Numerical Study in a High-Resolution Aquifer Model". In: *Water Resources Research* 58.5, e2021WR030110. DOI: 10.1029/2021WR030110.

Havas, K. (2019). "HIGH RESOLUTION IMAGING OF SOIL PROPERTIES INSIDE A LYSIMETER USING GPR FULL-WAVEFORM INVERSION". Masterarbeit, RWTH Aachen, 2019. Masterarbeit. RWTH Aachen, p. 103. URL: https://juser.fz-juelich.de/record/865927.

Holliger, K., M. Musil, and H. Maurer (2001). "Ray-based amplitude tomography for crosshole georadar data: a numerical assessment". In: *Journal of Applied Geophysics* 47.3. Ground Penetrating Radar, pp. 285–298. ISSN: 0926-9851. DOI: 10.1016/S0926-9851(01)00072-6.

Hoven, D., A. Mester, H. Vereecken, and A. Klotzsche (2023). "Evaluation of starting model approaches and effective source wavelet variations for high-frequency ground-penetrating radar full-waveform inversion". In: *GEOPHYSICS* 88.2, KS27–KS45. DOI: 10.1190/geo2021-0683.1.

Hoven, D., C. Warren, J. van der Kruk, H. Vereecken, and A. Klotzsche (in preparation). "Including the effects of borehole fillings and finite-length antenna models in 2.5D cross-

hole ground penetrating radar full-waveform inversion". Planned for publication in GEOPHYSICS.

Huisman, J. A., S. S. Hubbard, J. D. Redman, and A. P. Annan (2003). "Measuring Soil Water Content with Ground Penetrating Radar: A Review". In: *Vadose Zone Journal* 2.4, pp. 476–491. DOI: 10.2136/vzj2003.4760.

Igel, J., T. Günther, and M. Kuntzer (2013). "Ground-penetrating radar insight into a coastal aquifer: the freshwater lens of Borkum Island". In: *Hydrology and Earth System Sciences* 17.2, pp. 519–531. DOI: 10.5194/hess-17-519-2013.

Igel, J., S. Stadler, and T. Günther (2016). "High-resolution investigation of the capillary transition zone and its influence on GPR signatures". In: *2016 16th International Conference on Ground Penetrating Radar (GPR)*, pp. 1–5. DOI: 10.1109/ICGPR.2016.7572603.

Irving, J. D. and R. J. Knight (2005). "Effect of antennas on velocity estimates obtained from crosshole GPR data". In: *GEOPHYSICS* 70.5, K39–K42. DOI: 10.1190/1.2049349.

Jol, H. M. (2009). *Ground penetrating radar theory and applications*. elsevier. ISBN: 978-0-444-53348-7.

Jülich Supercomputing Centre (2016). "JURECA: General-purpose supercomputer at Jülich Supercomputing Centre". In: *Journal of large-scale research facilities* 2, A62. DOI: 10.17815/jlsrf-2-121.

Jülich Supercomputing Centre (2018). "JURECA: Modular supercomputer at Jülich Supercomputing Centre". In: *Journal of large-scale research facilities* 4, A132. DOI: 10.17815/jlsrf-4-121-1.

Keskinen, J., M. C. Looms, A. Klotzsche, and L. Nielsen (2021). "Practical data acquisition strategy for time-lapse experiments using crosshole GPR and full-waveform inversion". In: *Journal of Applied Geophysics* 191, p. 104362. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2021.104362.

Klotzsche, A. (2013). "Full-waveform inversion of crosshole GPR data for hydrogeological applications". Dissertation, RWTH Aachen, 2013. Dissertation. Jülich: RWTH Aachen, X. 164 S. ISBN: 978-3-89336-915-7. URL: https://juser.fz-juelich.de/record/139929.

Klotzsche, A. (2023). "Unlocking the potential of GPR for subsurface characterization by using full-waveform inversion". In: Virtual Near Surface Global Lecturer, online (Germany). URL: https://juser.fz-juelich.de/record/1019795.

Klotzsche, A., K. Havas, D. Hoven, and H. Vereecken (in preparation). “Using full-wave inversion techniques to image soil properties and water content at high spatial resolution at the lysimeter scale”. Planned for publication in IEEE.

Klotzsche, A., F. Jonard, M. C. Looms, J. van der Kruk, and J. A. Huisman (2018). “Measuring Soil Water Content with Ground Penetrating Radar: A Decade of Progress”. In: *Vadose Zone Journal* 17.1, p. 180052. DOI: 10.2136/vzj2018.03.0052.

Klotzsche, A., J. van der Kruk, G. Angelo Meles, J. Doetsch, H. Maurer, and N. Linde (2010). “Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland”. In: *Near Surface Geophysics* 8.6, pp. 635–649. DOI: 10.3997/1873-0604.2010054.

Klotzsche, A., J. van der Kruk, J. Bradford, and H. Vereecken (2014). “Detection of spatially limited high-porosity layers using crosshole GPR signal analysis and full-waveform inversion”. In: *Water Resources Research* 50.8, pp. 6966–6985. DOI: 10.1002/2013WR015177.

Klotzsche, A., J. van der Kruk, N. Linde, J. Doetsch, and H. Vereecken (Aug. 2013). “3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection”. In: *Geophysical Journal International* 195.2, pp. 932–944. ISSN: 0956-540X. DOI: 10.1093/gji/ggt275.

Klotzsche, A., L. Lärm, J. Vanderborght, G. Cai, S. Morandage, M. Zörner, H. Vereecken, and J. van der Kruk (2019a). “Monitoring Soil Water Content Using Time-Lapse Horizontal Borehole GPR Data at the Field-Plot Scale”. In: *Vadose Zone Journal* 18.1, p. 190044. DOI: 10.2136/vzj2019.05.0044.

Klotzsche, A., J. van der Kruk, G. Meles, and H. Vereecken (2012). “Crosshole GPR full-waveform inversion of waveguides acting as preferential flow paths within aquifer systems”. In: *GEOPHYSICS* 77.4, H57–H62. DOI: 10.1190/geo2011-0458.1.

Klotzsche, A., H. Vereecken, and J. van der Kruk (2019b). “Review of crosshole ground-penetrating radar full-waveform inversion of experimental data: Recent developments, challenges, and pitfalls”. In: *Geophysics* 84.6, H13–H28. ISSN: 0016-8033. DOI: 10.1190/geo2018-0597.1.

Klotzsche, A., H. Vereecken, and J. van der Kruk (2019c). “GPR full-waveform inversion of a variably saturated soil-aquifer system”. In: *Journal of Applied Geophysics* 170, p. 103823. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2019.103823.

Kuroda, S., M. Takeuchi, and H. J. Kim (2007). “Full-waveform inversion algorithm for interpreting crosshole radar data: a theoretical approach”. In: *Geosciences Journal* 11, pp. 211–217. ISSN: ISSN 1226-4806 1598-7477. DOI: 10.1007/BF02913934.

Lal, R., C. Monger, L. Nave, and P. Smith (2021). "The role of soil in regulation of climate". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 376.1834, p. 20210084. DOI: 10.1098/rstb.2021.0084.

Lanz, E., H. Maurer, and A. G. Green (1998). "Refraction tomography over a buried waste disposal site". In: *GEOPHYSICS* 63.4, pp. 1414–1433. DOI: 10.1190/1.1444443.

Lärm, L., F. M. Bauer, N. Hermes, J. van der Kruk, H. Vereecken, J. Vanderborght, T. H. Nguyen, G. Lopez, S. J. Seidel, F. Ewert, A. Schnepf, and A. Klotzsche (2023). "Multi-year belowground data of minirhizotron facilities in Selhausen". In: *Scientific data* 10.1, p. 672. ISSN: 2052-4436. DOI: 10.1038/s41597-023-02570-9.

Lärm, L., F. M. Bauer, J. van der Kruk, J. Vanderborght, S. Morandage, H. Vereecken, A. Schnepf, and A. Klotzsche (2024). "Linking horizontal crosshole GPR variability with root image information for maize crops". In: *Vadose Zone Journal* 23.1, e20293. DOI: 10.1002/vzj2.20293.

Lavoué, F., R. Brossier, L. Métivier, S. Garambois, and J. Virieux (Jan. 2014). "Two-dimensional permittivity and conductivity imaging by full waveform inversion of multi-offset GPR data: a frequency-domain quasi-Newton approach". In: *Geophysical Journal International* 197.1, pp. 248–268. ISSN: 0956-540X. DOI: 10.1093/gji/ggt528.

Lee, J.-F., R. Lee, and A. Cangellaris (1997). "Time-domain finite-element methods". In: *IEEE Transactions on Antennas and Propagation* 45.3, pp. 430–442. DOI: 10.1109/8.558658.

Liu, T., A. Klotzsche, M. Pondkule, H. Vereecken, Y. Su, and J. van der Kruk (2018). "Radius estimation of subsurface cylindrical objects from ground-penetrating-radar data using full-waveform inversion". In: *GEOPHYSICS* 83.6, H43–H54. DOI: 10.1190/geo2017-0815.1.

Liu, T. (2023). "Time-domain poroelastic full-waveform inversion of shallow seismic data". PhD thesis. Karlsruher Institut für Technologie (KIT). 127 pp. DOI: 10.5445/IR/1000156231.

Looms, M. C., K. H. Jensen, A. Binley, and L. Nielsen (2008). "Monitoring Unsaturated Flow and Transport Using Cross-Borehole Geophysical Methods". In: *Vadose Zone Journal* 7.1, pp. 227–237. DOI: 10.2136/vzj2006.0129.

Mangel, A. R., S. M. J. Moysey, J. C. Ryan, and J. A. Tarbutton (2012). "Multi-offset ground-penetrating radar imaging of a lab-scale infiltration test". In: *Hydrology and Earth System Sciences* 16.11, pp. 4009–4022. DOI: 10.5194/hess-16-4009-2012.

Marelli, S., H. Maurer, and E. Manukyan (2012). "Validity of the acoustic approximation in full-waveform seismic crosshole tomography". In: *GEOPHYSICS* 77.3, R129–R139. DOI: 10.1190/geo2011-0274.1.

Maurer, H., S. Greenhalgh, and S. Latzel (2009). "Frequency and spatial sampling strategies for crosshole seismic waveform spectral inversion experiments". In: *GEOPHYSICS* 74.6, WCC79–WCC89. DOI: 10.1190/1.3157252.

Maurer, H. and M. Musil (2004). "Effects and removal of systematic errors in crosshole georadar attenuation tomography". In: *Journal of Applied Geophysics* 55.3, pp. 261–270. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2004.02.003.

Meles, G., S. Greenhalgh, J. van der Kruk, A. Green, and H. Maurer (2011). "Taming the non-linearity problem in GPR full-waveform inversion for high contrast media". In: *Journal of Applied Geophysics* 73.2, pp. 174–186. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2011.01.001.

Meles, G. A., J. van der Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and A. G. Green (2010). "A New Vector Waveform Inversion Algorithm for Simultaneous Updating of Conductivity and Permittivity Parameters From Combination Crosshole/Borehole-to-Surface GPR Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 48.9, pp. 3391–3407. DOI: 10.1109/TGRS.2010.2046670.

Mozaffari, A., A. Klotzsche, C. Warren, G. He, A. Giannopoulos, H. Vereecken, and J. van der Kruk (2020). "2.5D crosshole GPR full-waveform inversion with synthetic and measured data". In: *GEOPHYSICS* 85.4, H71–H82. DOI: 10.1190/geo2019-0600.1.

Mozaffari, A., A. Klotzsche, Z. Zhou, H. Vereecken, and J. v. d. Kruk (2022). "3-D Electromagnetic Modeling Explains Apparent-Velocity Increase in Crosshole GPR Data-Borehole Fluid Effect Correction Method Enables to Incorporating High-Angle Traveltime Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 60, pp. 1–10. DOI: 10.1109/TGRS.2021.3107451.

Mujal-Rosas, R., M. Marin-Genesca, and J. Ballart-Prunell (2015). "Dielectric properties of various polymers (PVC, EVA, HDPE, and PP) reinforced with ground tire rubber (GTR)". In: *Science and Engineering of Composite Materials* 22.3, pp. 231–243. DOI: 10.1515/secm-2013-0233.

Oberröhrmann, M., A. Klotzsche, H. Vereecken, and J. van der Kruk (2013). "Optimization of acquisition setup for cross-hole: GPR full-waveform inversion using checkerboard analysis". In: *Near Surface Geophysics* 11.2, pp. 197–209. DOI: 10.3997/1873-0604.2012045.

Patsia, O., A. Giannopoulos, and I. Giannakis (2023). “Developing a realistic numerical equivalent of a GPR antenna transducer using global optimizers”. English. In: *NEAR SURFACE GEOPHYSICS*. DOI: 10.1002/nsg.12280.

Paz, C., F. J. Alcalá, J. M. Carvalho, and L. Ribeiro (2017). “Current uses of ground penetrating radar in groundwater-dependent ecosystems research”. In: *Science of The Total Environment* 595, pp. 868–885. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2017.03.210.

Pinard, H., S. Garambois, M. Métivier Ludovicand Dietrich, G. Sénéchal, and D. Rousset (2016). “Full-waveform inversion of GPR data acquired between boreholes in Rustrel carbonates”. In: *E3S Web of Conf.* 12, p. 01002. DOI: 10.1051/e3sconf/20161201002.

Pratt, R. G., C. Shin, and G. J. Hick (May 1998). “GaussNewton and full Newton methods in frequencyspace seismic waveform inversion”. In: *Geophysical Journal International* 133.2, pp. 341–362. ISSN: 0956-540X. DOI: 10.1046/j.1365-246X.1998.00498.x.

Pratt, R. G. and R. M. Shipp (1999). “Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data”. In: *GEOPHYSICS* 64.3, pp. 902–914. DOI: 10.1190/1.1444598.

Pratt, R. G. and M. H. Worthington (1990). “Inverse theory applied to multi-source cross-hole tomography.” In: *Geophysical Prospecting* 38.3, pp. 287–310. DOI: 10.1111/j.1365-2478.1990.tb01846.x.

Pütz, T., J. Fank, and M. Flury (2018). “Lysimeters in Vadose Zone Research”. In: *Vadose Zone Journal* 17.1, p. 180035. DOI: 10.2136/vzj2018.02.0035.

Qin, T. (2022). “Full-waveform inversion of ground-penetrating radar data and its indirect joint petrophysical inversion with shallow-seismic data”. PhD thesis. Karlsruher Institut für Technologie (KIT). 105 pp. DOI: 10.5445/IR/1000152587.

Qin, T., T. Bohlen, and N. Allroggen (Aug. 2022). “Full-waveform inversion of ground-penetrating radar data in frequency-dependent media involving permittivity attenuation”. In: *Geophysical Journal International* 232.1, pp. 504–522. ISSN: 0956-540X. DOI: 10.1093/gji/ggac319.

Rabbel, W. (2009). “Seismic methods”. In: *Groundwater Geophysics: A Tool for Hydrogeology*. Ed. by R. Kirsch. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 23–84. ISBN: 978-3-540-88405-7. DOI: 10.1007/978-3-540-88405-7_2.

Romero-Ruiz, A., N. Linde, T. Keller, and D. Or (2018). “A Review of Geophysical Methods for Soil Structure Characterization”. In: *Reviews of Geophysics* 56.4, pp. 672–697. DOI: 10.1029/2018RG000611.

Schmalholz, J., H. Stoffregen, A. Kemna, and U. Yaramanci (2004). "Imaging of Water Content Distributions inside a Lysimeter using GPR Tomography". In: *Vadose Zone Journal* 3.4, pp. 1106–1115. DOI: 10.2136/vzj2004.1106.

Sirgue, L. and R. G. Pratt (2004). "Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies". In: *GEOPHYSICS* 69.1, pp. 231–248. DOI: 10.1190/1.1649391.

Song, Z.-M. and P. R. Williamson (1995). "Frequencydomain acousticwave modeling and inversion of crosshole data: Part I2.5-D modeling method". In: *GEOPHYSICS* 60.3, pp. 784–795. DOI: 10.1190/1.1443817.

Stadler, S. and J. Igel (2018). "A numerical study on using guided GPR waves along metallic cylinders in boreholes for permittivity sounding". In: *2018 17th International Conference on Ground Penetrating Radar (GPR)*, pp. 1–4. DOI: 10.1109/ICGPR.2018.8441666.

Stadler, S. and J. Igel (2022). "Developing Realistic FDTD GPR Antenna Surrogates by Means of Particle Swarm Optimization". In: *IEEE Transactions on Antennas and Propagation* 70.6, pp. 4259–4272. DOI: 10.1109/TAP.2022.3142335.

Steinbeck, L., A. Mester, E. Zimmermann, A. Klotzsche, and S. Van Waasen (2023). "Calibration of a Stationary Multichannel GPR Monitoring System Using Internal Reflection Measurements". In: *IEEE transactions on geoscience and remote sensing* 61, pp. 1–10. ISSN: 0018-9413. DOI: 10.1109/TGRS.2023.3275191.

Stephan, S. M., N. Allroggen, and J. Tronicke (2022). "Experimental investigation of GPR signal quality using different stacking values". In: *19th International Conference on Ground Penetrating Radar*, pp. 143–146. DOI: 10.1190/gpr2022-049.1.

Stoffregen, H., T. Zenker, and G. Wessolek (2002). "Accuracy of soil water content measurements using ground penetrating radar: comparison of ground penetrating radar and lysimeter data". In: *Journal of Hydrology* 267.3, pp. 201–206. ISSN: 0022-1694. DOI: 10.1016/S0022-1694(02)00150-6.

Tabarovsky, L., M. Goldman, M. Rabinovich, and K.-M. Strack (1996). "2.5-D modeling in electromagnetic methods of geophysics". In: *Journal of Applied Geophysics* 35.4, pp. 261–284. ISSN: 0926-9851. DOI: 10.1016/0926-9851(96)00025-0.

Taflove, A. and S. C. Hagness (2005). *Computational electrodynamics : the finite difference time domain method*. 3rd ed. englisch. Boston: Artech House. ISBN: 1580538320.

Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". In: *GEOPHYSICS* 49.8, pp. 1259–1266. DOI: 10.1190/1.1441754.

The HDF Group, N., Q. Koziol, and U. O. of Science (2020). *HDF5-Version 1.12.0*. DOI: 10.11578/dc.20180330.1. URL: https://github.com/HDFGroup/hdf5.

Thomsen, R., V. H. Søndergaard, and K. I. Sørensen (2004). "Hydrogeological mapping as a basis for establishing site-specific groundwater protection zones in Denmark". In: *Hydrogeology Journal* 12.5, pp. 550–562. DOI: 10.1007/s10040-004-0345-1.

van der Kruk, J., N. Güting, A. Klotzsche, G. He, S. Rudolph, C. von Hebel, X. Yang, L. Weihermüller, A. Mester, and H. Vereecken (2015). "Quantitative multi-layer electromagnetic induction inversion and full-waveform inversion of crosshole ground penetrating radar data". In: *Journal of Earth Science* 26.6, pp. 844–850. DOI: 10.1007/s12583-015-0610-3.

Vanderborght, J., A. Englert, J. A. Huisman, J. van der Kruk, and H. Vereecken (Oct. 17, 2012). "The Krauthausen Test Site". In: 16ième rencontre H+, Rennes, 17 Oct 2012 - 18 Oct 2012. URL: https://juser.fz-juelich.de/record/127323.

Vereecken, H., A. Schnepf, J. Hopmans, M. Javaux, D. Or, T. Roose, J. Vanderborght, M. Young, W. Amelung, M. Aitkenhead, S. Allison, S. Assouline, P. Baveye, M. Berli, N. Brüggemann, P. Finke, M. Flury, T. Gaiser, G. Govers, T. Ghezzehei, P. Hallett, H. Hendricks Franssen, J. Heppell, R. Horn, J. Huisman, et al. (2016). "Modeling Soil Processes: Review, Key Challenges, and New Perspectives". In: *Vadose Zone Journal* 15.5, vzj2015.09.0131. DOI: 10.2136/vzj2015.09.0131.

Vidale, J. E. (1990). "Finitedifference calculation of traveltimes in three dimensions". In: *Geophysics* 55, pp. 521–526. URL: https://api.semanticscholar.org/CorpusID:3567805.

Virieux, J. and S. Operto (2009). "An overview of full-waveform inversion in exploration geophysics". In: *GEOPHYSICS* 74.6, WCC1–WCC26. DOI: 10.1190/1.3238367.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. tekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand (2013). "Anisotropic 3D full-waveform inversion". In: *GEOPHYSICS* 78.2, R59–R80. DOI: 10.1190/geo2012-0338.1.

Warren, C. and A. Giannopoulos (2011). "Creating finite-difference time-domain models of commercial ground-penetrating radar antennas using Taguchis optimization method". In: *GEOPHYSICS* 76.2, G37–G47. DOI: 10.1190/1.3548506.

Warren, C., A. Giannopoulos, and I. Giannakis (2016). "gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar". In: *Computer Physics Communications* 209, pp. 163–170. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2016.08.020.

Warren, C., A. Giannopoulos, A. Gray, I. Giannakis, A. Patterson, L. Wetter, and A. Hamrah (2019). "A CUDA-based GPU engine for gprMax: Open source FDTD electromagnetic simulation software". In: *Computer Physics Communications* 237, pp. 208–218. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2018.11.007.

Weiland, T. (1977). "A discretization model for the solution of Maxwell's equations for six-component fields". In: *Archiv Elektronik und Uebertragungstechnik* 31, pp. 116–120.

Weiland, T. (2003). "Finite Integration Method and Discrete Electromagnetism". In: *Computational Electromagnetics*. Ed. by P. Monk, C. Carstensen, S. Funken, W. Hackbusch, and R. H. W. Hoppe. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 183–198. ISBN: 978-3-642-55745-3.

Williamson, P. R. (1991). "A guide to the limits of resolution imposed by scattering in ray tomography". In: *GEOPHYSICS* 56.2, pp. 202–207. DOI: 10.1190/1.1443032.

Yang, X., J. van der Kruk, A. Klotzsche, and H. Vereecken (May 2014). "GPR full-waveform inversion using a combined frequency- and time-domain approach". In: *EGU General Assembly Conference Abstracts*. EGU General Assembly Conference Abstracts, 14046, p. 14046.

Yee, K. (1966). "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media". In: *IEEE Transactions on Antennas and Propagation* 14.3, pp. 302–307. DOI: 10.1109/TAP.1966.1138693.

Zhang, K. and D. Li (1998). "Chapter 7 Electromagnetic Waves in Dispersive Media". In: *Electromagnetic Theory for Microwaves and Optoelectronics*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 433–452. ISBN: 978-3-662-03553-5. DOI: 10.1007/978-3-662-03553-5_7.

Zhou, Z., A. Klotzsche, T. Hermans, F. Nguyen, J. Schmäck, P. Haruzi, H. Vereecken, and J. van der Kruk (2020). "3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole ground-penetrating radar amplitude analysis and full-waveform inversion". In: *GEOPHYSICS* 85.6, H133–H148. DOI: 10.1190/geo2020-0067.1.

Zhou, Z., A. Klotzsche, and H. Vereecken (2021). "Improving crosshole ground-penetrating radar full-waveform inversion results by using progressively expanded bandwidths of the data". In: *Near Surface Geophysics* 19.4, pp. 465–487. DOI: 10.1002/nsg.12154.

Band / Volume 630
Elucidation of Barocaloric Effect in Spin Crossover Compounds
H. Shahed (2024), x, 261 pp
ISBN: 978-3-95806-758-5

Band / Volume 631
Computational Investigation of Solvation Phenomena at Metal-Electrolyte Interfaces
O. Cheong (2024), xvii, 142 pp
ISBN: 978-3-95806-759-2

Band / Volume 632
Senkung zukünftiger Stickoxid- und Partikelemissionen in Nordrhein-Westfalen durch den Einsatz alternativer Energieträger und Antriebe
J. L. Breuer (2024), vii, 339 pp
ISBN: 978-3-95806-760-8

Band / Volume 633
Development of Model-Based Correction Methods for Temperature-Dependent Electromagnetic Induction (EMI) Measurement Errors in Soil Conductivity Estimations
T. M. Tchantcho Amin (2024), xx, 100 pp
ISBN: 978-3-95806-761-5

Band / Volume 634
Investigation and implementation of improved and degradation-tolerant fuel electrodes for solid oxide cells
A. Schwiers (2024), VI, 163, XIII pp
ISBN: 978-3-95806-762-2

Band / Volume 635
In Situ Time Calibration for Stationary Multichannel GPR Monitoring Systems
L. Steinbeck (2024), xvi, 98, xxxi pp
ISBN: 978-3-95806-767-7

Band / Volume 636
Erneuerbares Methanol als Ausgangsstoff für die Bereitstellung von flüssigen Kraftstoffen für den Transportsektor
F. Schorn (2024), VI, 275 pp
ISBN: 978-3-95806-769-1

Band / Volume 637
Investigation of Lower Boundary Conditions of Brominated Very Short-lived Species (VSLS)
S. Zheng (2024), 2, iii, 160 pp
ISBN: 978-3-95806-770-7

Band / Volume 638
Modellgestützte Analyse zukünftigen Mobilitätsverhaltens
J. P. Reul (2024), XVI, 291 pp
ISBN: 978-3-95806-771-4

Band / Volume 639
Insights into Mechanisms of Secondary Organic Aerosol Formation: Approaching Atmospherically Relevant Conditions in an Atmospheric Reaction Chamber
Y. Baker (2024), XVII, 122 pp
ISBN: 978-3-95806-776-9

Band / Volume 640
Advancing the representation of agricultural systems in Land Surface Models: systematic model evaluations and technical model developments
T. S. Boas (2024), xxi, 145 pp
ISBN: 978-3-95806-777-6

Band / Volume 641
Imaging spatial and temporal soil water content variations of the soil-plant continuum using ground penetrating radar
L. Lärm (2024), xii, 303 pp
ISBN: 978-3-95806-778-3

Band / Volume 642
Development of Iridium-based Nanostructures for Oxygen Evolution Reaction in PEM Water Electrolysis
S. Park (2024), 135 pp
ISBN: 978-3-95806-779-0

Band / Volume 643
Multi-dimensional GPR full-waveform inversion for small-scale hydrogeophysical soil characterization
D. Hoven (2024), IX, 163 pp
ISBN: 978-3-95806-781-3

Energie & Umwelt / Energy & Environment
Band / Volume 643
ISBN 978-3-95806-781-3

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum