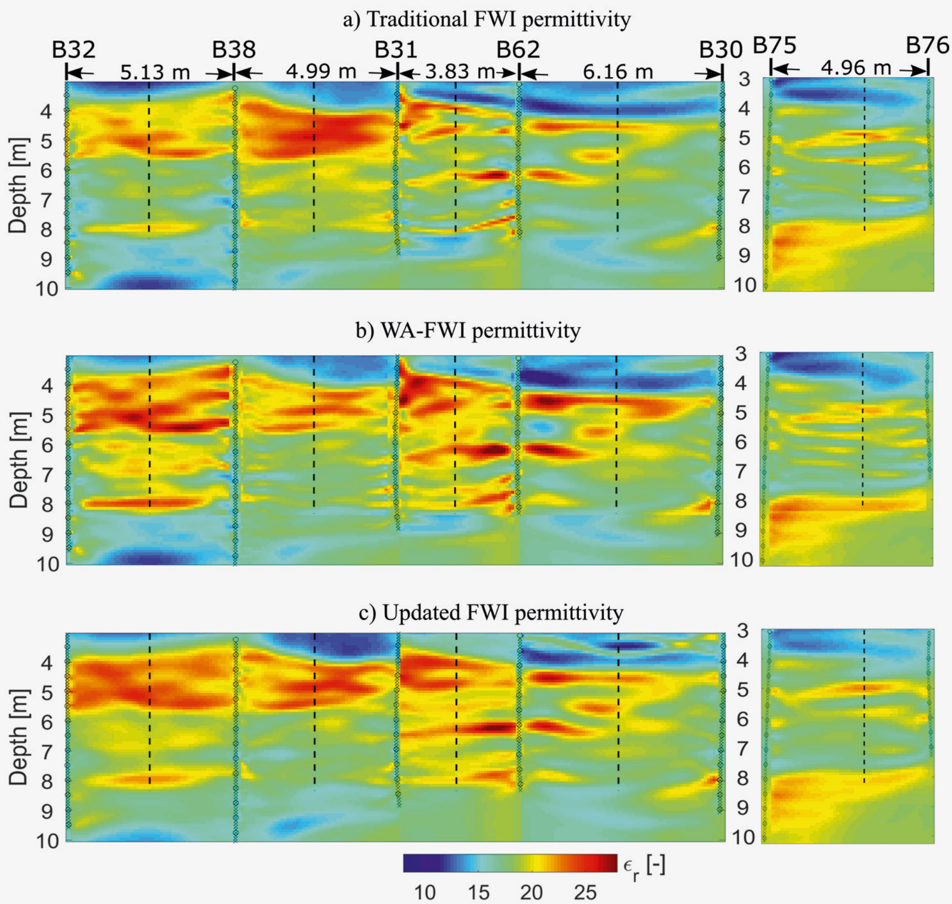


Enhanced crosshole GPR full-waveform inversion to improve aquifer characterization

Zhen Zhou

Energie & Umwelt / Energy & Environment
Band / Volume 512
ISBN 978-3-95806-500-0

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 512

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

Enhanced crosshole GPR full-waveform inversion to improve aquifer characterization

Zhen Zhou

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment Band / Volume 512

ISSN 1866-1793 ISBN 978-3-95806-500-0

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 2020

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 512

D 82 (Diss. RWTH Aachen University, 2020)

ISSN 1866-1793
ISBN 978-3-95806-500-0

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

Abstract

Complex heterogeneities in the aquifers are critical and challenging to be detected and can have a significant effect on subsurface flow and transport. Thereby, reliable prediction of groundwater flow and solute transport is important for the protection of drinking water, and the remediation of contaminants. Small-scale high resolution images of the subsurface can help to improve the understanding of complex and heterogeneous aquifer structures that effect hydrological properties and processes. To successfully obtain hydrological parameters distributed in a 2D cross-section with high resolution, we can apply crosshole ground penetrating radar (GPR). Crosshole GPR uses high-frequency electromagnetic pulses that are emitted from a dipole-type antenna in a borehole and recorded by a receiver antenna in a second borehole. The received electromagnetic wave with its arrival time and amplitude contains information about the subsurface medium properties through which it travelled. Thereby, crosshole GPR is able to provide two electromagnetic parameters the dielectric permittivity and the electrical conductivity at the same time. Conventional inversion approaches for crosshole GPR data are generally based on geometrical ray theory, which provide relatively smooth images with a resolution that scales approximately with the diameter of the first Fresnel zone. In contrast, the crosshole GPR full-waveform inversion (FWI) provides decimeter-scale high resolution images, because it considers the fully recorded waveform information and the inversion is based on solving the full Maxwell's equations. However, the crosshole GPR FWI approach also includes some limiting factors and requires several detailed processing and inversion steps. If these steps are not carefully applied, the inversion can be trapped in a local minimum. In order to minimize the influence of at least some of these factors, appreciate FWI starting models and an accurate estimation of the effective source wavelet are required.

To precisely describe aquifers, high porosity layers and clay lenses, that can strongly effect flow and transport processes, need to be considered. In the framework of this thesis, we first extend this amplitude analysis approach to identify two different types of low-velocity waveguides, caused by an increased porosity and/or by a higher electrical conductivity. The obtained information about extension and dimension of such wave guiding structures is considered to improve the starting models of the FWI. Moreover, we estimate an updated effective source wavelet based on the updated permittivity starting model. To verify the presented scheme, nine GPR cross-sections were measured and analyzed at the Hermalle-sous-Argenteau test site near Liege in Belgium. Consistent structures between different cross-sections show the robustness of the updated amplitude analysis approach and the FWI results. In addition, the aquifer structures obtained from the new FWI results agree with the crosshole electrical resistivity tomography (ERT) monitoring

results of a previous conducted heat tracer experiment and is able to explain a precisely obtained plume distribution in more detail.

To further improve the reconstruction of the subsurface properties, we combine the standard FWI results of crosshole GPR data with Cone Penetration Test (CPT) data. In previous studies, the FWI results were compared with the low wavenumber information of nearby measured CPT data and found that the low wavenumbers information of GPR FWI models was reliable. Therefore, we introduce an improved scheme that is able to enhance the effective source wavelet by combining the standard FWI permittivity results with the porosity data derived from the CPT measurements. Firstly, we converted the FWI permittivity values to porosity values, which correspond to water content in the saturation zone. Secondly, a lowpass filter based on the porosity values of the two approaches was constructed, and used to amplify the FWI permittivity results within the full 2D cross-section. By using these wavenumber amplified permittivities in the forward modeling an updated effective source wavelet can be obtained, that not only contains the CPT information but also possesses a larger bandwidth than the traditional effective source wavelet. Finally, the new updated FWI was performed using the ray-based starting models and the updated source wavelet. The new scheme was applied to a realistic synthetic data set and to experimental data of the Krauthausen aquifer in Germany. A comparison between the traditional FWI and the updated FWI permittivity results illustrates a higher and more reliable resolution of the updated FWI.

In addition, another approach that improves the FWI results for the crosshole GPR based on a progressively expanded bandwidth approach is presented. To tame the non-linearity problem of the FWI process, we consider using the progressively expanded bandwidths of the modeled data and the measured GPR data (PEBDD) to construct a better permittivity starting model. Therefore, we designed tapered bandpass filters and applied them to the standard effective source wavelet and the GPR data. The first FWI uses the sub-data with the smallest bandwidth. The generated FWI results replace the previous starting models while the bandwidth of the sub-data is progressively expanded. This procedure is repeated until the sub-data with the selected maximum bandwidth and the resulting FWI permittivity result as the new starting model for the final FWI with the full bandwidth data. To test the FWI with PEBDD approach, synthetic GPR data, that is generated using a stochastic model of the Krauthausen test site, and field experimental GPR data from the same site are used. We compared the standard FWI results with the new updated FWI results by computing the root mean square error and correlation coefficient with the CPT data and concluded that the new PEBDD approach is able to improve the results of the GPR FWI. In addition, by the usage of the PEBDD scheme the detailed work to construct good starting models for experimental GPR data can be reduced and the application to field data will be much easier.

These enhanced 2D GPR full-waveform inversion schemes provide detailed hydrogeological aquifer characterization, and can be applied in other Geophysics applications, thus improving our ability to detect small-scale structures in the subsurface.

Zusammenfassung

Komplexe Heterogenitäten in den Grundwasserleitern sind kritisch und schwierig zu erkennen und können einen erheblichen Einfluss auf die unterirdische Strömung und den Transport haben. Eine zuverlässige Vorhersage der Grundwasserströmung und des Transports gelöster Stoffe ist entscheidend für den Schutz des Trinkwassers und für die Schadstoffsanierung. Hochauflösende Darstellungen des oberflächennahen Untergrundes können dazu beitragen, komplexe und heterogene Strukturen in Grundwasserleitern, und deren Einfluss auf hydrologische Eigenschaften und Prozesse, besser zu verstehen. „Crosshole“ (Bohrloch zu Bohrloch) Bodenradar (englisch: Ground Penetrating Radar; GPR), kann genau für solche Fragestellungen eingesetzt werden und die erzeugten 2D-Querschnitte können erfolgreich hydrologische Parameter abbilden. Dadurch ist das Crosshole-GPR in der Lage gleichzeitig zwei elektromagnetische Parameter, die dielektrische Permittivität und die elektrische Leitfähigkeit, zu ermitteln. Dazu liefert die Crosshole-GPR Wellenform-Inversion (englisch: Full-Waveform Inversion; FWI) hochauflösende Bilder im Dezimeterbereich, da sie die gesamte Form der aufgezeichneten Wellen berücksichtigt und auf der Lösung der vollständigen Maxwell-Gleichungen basiert. Allerdings beinhaltet der Crosshole-GPR FWI Ansatz auch einige limitierende Faktoren und erfordert mehrere detaillierte Vorbearbeitungs- und Inversionsschritte. Wenn diese Schritte nicht sorgfältig angewendet werden, kann die Inversion in einem lokalen Minimum stagniert. Um den Einfluss zumindest einiger dieser Faktoren zu minimieren, sind gute FWI-Startmodelle und eine genaue Schätzung des effektiven Quellsignals erforderlich.

Zur genauen Beschreibung von Grundwasserleitern müssen Schichten mit hoher Porosität und Tonlinsen, welche Strömungs- und Transportprozesse stark beeinflussen können, berücksichtigt werden. Deshalb erweitern wir den Amplitudenanalyse-Ansatz, um zukünftig beide zuvor beschriebenen Typen von Niedriggeschwindigkeits-Wellenleitern zu identifizieren, die durch eine erhöhte Porosität und/oder durch eine höhere elektrische Leitfähigkeit verursacht werden. Die gewonnenen Informationen über die Abemessungen solcher wellenleitenden Strukturen werden zur Verbesserung der FWI Startmodelle benutzt. Zusätzlich ermitteln wir, auf der Grundlage des aktualisierten Permittivitäts-Startmodells, ein aktualisiertes effektives Quellsignal. Zur Verifizierung des vorgestellten Schemas, wurden neun GPR-Querschnitte auf dem Testgelände Hermalle-sous-Argenteau bei Lüttich in Belgien gemessen und analysiert. Konsistente Strukturen zwischen verschiedenen Querschnitten zeigen die Robustheit des aktualisierten Amplitudenanalyse-Ansatzes und der aktualisierten FWI-Ergebnisse.

Um die Rekonstruktion der Untergrundeigenschaften weiter zu verbessern, kombinieren wir die FWI-Standardergebnisse der Crosshole-GPR-Daten mit Daten aus Drucksondierungsversuchen (englisch: Cone Penetration Test; CPT). Wir führen ein verbessertes Schema ein, das in der Lage ist das effektive Quellsignal

zu verbessern, indem es die Ergebnisse der Standard-FWI mit den CPT-Daten kombiniert. Zuerst wurden die FWI Permittivitätswerte zu Porositätswerten konvertiert, die in der gesättigten Zone dem Wassergehalt entsprechen. Dann wurde ein Tiefpassfilter basierend auf den Porositätswerten der beiden Ansätze konstruiert und zur Verstärkung der FWI Permittivitäts-Ergebnisse innerhalb des gesamten 2D-Querschnitts genutzt. Durch Verwendung dieser wellenzahlverstärkten Permittivitäten in der Vorwärtsmodellierung kann ein aktualisiertes effektives Quellsignal ermittelt werden, das nicht nur die CPT-Informationen enthält, sondern auch eine größere Bandbreite als das traditionelle effektive Quellsignal aufweist. Schließlich wurde eine aktualisierte FWI unter Verwendung der strahlenbasierten Startmodelle und des aktualisierten Quellsignals durchgeführt. Das neue Vorgehen wurde auf einen realistischen synthetischen Datensatz und auf experimentelle Daten eines Grundwasserleiters des Testgeländes Krauthausen in Deutschland angewandt. Ein Vergleich zwischen den traditionellen FWI- und den aktualisierten FWI-Permittivitätsergebnissen zeigt eine höhere und zuverlässigere Auflösung der aktualisierten FWI.

Darüber hinaus wird ein weiterer Ansatz vorgestellt, der die FWI-Ergebnisse für das Crosshole-GPR auf der Grundlage eines progressiv erweiterten Bandbreitenansatzes verbessert. Um das Nicht-Linearitätsproblem des FWI-Prozesses zu bewältigen, versuchten wir die progressiv erweiterten Bandbreiten der modellierten und der gemessenen GPR-Daten (PEBDD) zu verwenden, um ein besseres Permittivitäts-Startmodell zu konstruieren. Dafür konstruierten wir konisch zulaufende Bandpassfilter, die auf das effektive Standard-Quellsignal und die GPR-Daten angewendet wurden. Die erste FWI verwendet die Teildaten mit der kleinsten Bandbreite. Die generierten FWI-Ergebnisse ersetzen die früheren Startmodelle, während die Bandbreite der Teildaten schrittweise erweitert wird. Dieses Verfahren wird so lange wiederholt, bis zu den Teildaten mit der maximalen Bandbreite und die daraus resultierenden FWI-Permittivitätsergebnisse werden als neues Startmodell für die endgültige FWI mit den Daten der vollen Bandbreite eingesetzt. Um den FWI mit dem PEBDD-Ansatz zu testen, werden synthetische GPR-Daten, die mit Hilfe eines stochastischen Modells des Versuchsstandorts Krauthausen generiert wurden, und experimentelle GPR-Felddaten vom gleichen Standort verwendet. Darüber hinaus kann durch die Verwendung des PEBDD-Schemas die notwendige Arbeit zur Erstellung guter Startmodelle für experimentelle GPR-Daten reduziert werden und die Anwendung auf Felddaten kann wesentlich erleichtert werden.

Diese verbesserten 2D-GPR Wellenform-Inversionsmethoden mit ermöglichen eine detaillierte hydrogeologische Grundwasserleiter-Charakterisierung und können auch in anderen geophysikalischen Anwendungen angewandt werden, wodurch die Erkennung kleinräumiger Strukturen im Untergrund verbessert wird.

Contents

Chapter 1

Introduction

With the development of society economic, the demand for Earth natural resources continues to increase. Freshwater resources are extremely important for the progress and development of the human civilization. Groundwater is the most important freshwater resource on Earth that constitutes over 95% of the Earth's unfrozen freshwater (Wada et al., 2010). The shallow subsurface of the earth is an extremely important geological zone that contains much of our available freshwater, supports our agriculture, and ecosystems, and influences our climate (Hubbard et al., 2005). With growing population, contaminants associated with daily life and industrially polluted groundwater increase the threat to human life. Therefore, it is very urgent to improve our understanding of the shallow subsurface heterogeneities that we are able to accurately predict groundwater flow and contaminant transport (e.g., Bear and Cheng, 2010; Binley et al., 2015).

The heterogeneity and the spatial variability in the subsurface aquifers have a great influence on the prediction of groundwater storage, the determination of priority flow direction, and the spread of contaminants. Traditional hydrogeological methods to characterize estimate hydrological parameters in near-surface aquifers include, e.g., coring, pumping tests, flowmeter tests and permeameter tests (e.g., Colwell et al., 1992; Bohling at al., 2002; Le Borgne at al., 2007). Such methods have either an average response over a large volume with a lack of detailed characterization at smaller-scale (Landon et al., 2001), or, they represent small spatial sampling volume with a poor lateral expansion (Brauchler et al., 2011). Especially for the spatial connectivity of subsurface structures at the field scale, classic hydrogeological methods are not able to provide information with higher resolution. To overcome the difficulty of estimation of the subsurface spatial heterogeneity at the point scale (e.g., high contrast layers linked to preferential flow paths or clay lenses) with a high spatial resolution, Cone Penetration Test (CPT) surveys with a fast, accurate, and minimally invasive manner have become popular in the last decades to investigate shallow unconsolidated sediments (e.g., Fejes and Jósa, 1990; Lunne et al., 1997; Tillmann et al., 2008). The CPT approach can be used to collect subsurface lithological and soil data and quantitatively investigate the data properties in detail. For example, CPT can provide mechanical resistance, natural gamma activity, bulk density, matrix density, water content, and electrical resistivity data (Tillmann et al., 2008). However, the drawback of the CPT approach is its one-dimensional nature, and therefore the distribution of lateral information is poor.

In the last decades, geophysical methods have been widely applied for the subsurface modeling and imaging, which can be used to provide detailed models of the near-surface and the critical zone (e.g., Binley et al., 2015). Generally, these geophysical methods are able to close the gap between small-scale investigations with high resolution (e.g., coring) and large-scale zones mapping (e.g., flowmeter tests). Seismics (e.g., Doetsch et al., 2010a), electrical resistivity tomography (ERT; e.g., Coscia et al., 2012) and ground penetrating radar (GPR; e.g., Klotzsche et al., 2010) are popular to gain 2D and 3D continuous models of subsurface structures. Some research groups have successfully used one or a combination of these methods to characterize the near-subsurface and to derive hydrological relevant properties (Binley et al., 2015). For example, Feng et al. (2017) provided a joint inversion of crosshole seismic and GPR data in the frequency domain to map the shallow subsurface structures. Hermans et al. (2015a) investigated the ability of crosshole ERT to monitor a heat tracer experiment in a complex heterogeneous alluvial aquifer. Looms et al. (2008) introduced an approach to monitor unsaturated flow and transport using cross-borehole GPR data. These results improved the understanding of how different geophysical methods can be used to obtain detailed aquifer structures. By using empirical relationships or specific mixing models, the measured geophysical parameters can be converted to the hydrological parameters and further image the subsurface heterogeneity in detail. To improve the characterization of hydrogeological parameters (e.g., porosity, hydraulic conductivity or dispersivity), time-lapse tracer experiments can be applied (e.g., Linde et al., 2006; Saar, 2011; Hermans et al., 2015a). Commonly applied tracers are salt or heat tracers, which can be monitored with ERT (e.g., LaBrecque et al., 1996b; Hermans et al., 2012b) or GPR traveltime tomography (Kowalsky et al., 2004). For example, to investigate flow and transport processes, time-lapse crosshole GPR measurements with a high spatial and temporal resolution can be linked to soil hydrological parameters such as hydraulic conductivity (e.g., Doetsch et al., 2010b). Müller et al. (2010) applied a negative and positive salt tracer and monitored the time–lapse changes with crosshole ERT to characterize flow and transport processes.

During the last two decades, GPR has undergone a rapid development and has shown a great potential to non-or/minimally invasively map and characterize aquifers with a higher resolution than ERT and seismic investigations (e.g., Huisman et al., 2001; Klotzsche et al., 2018). The method is using high frequency electromagnetic (EM) waves to obtain the near-subsurface parameters dielectric permittivity ε (related to velocity v of the EM wave) and electrical conductivity σ (related to attenuation α of the EM wave). In this thesis, the relative permittivity is used, which is calculated with the formula $\varepsilon_r = \varepsilon / \varepsilon_0$, where ε_0 is the dielectric permittivity of free space with 8.8542×10^{-12}. Because of the large difference between the relative permittivity of air (ε_r=1) and pure water (ε_r = 80 at 20°C), ε_r can be linked to the soil water content or porosity in the subsurface saturated media using appreciate empirical equations or mixing models (e.g., Birchak et al., 1974; Eisenberg and Kauzmann, 2005; Steelman and Endres, 2011; Gueting et al., 2015; Carmichael, 2017). Meanwhile, σ can be linked to hydrological relevant variables such as water salinity, clay content, and lithological variations (e.g., Davis & Annan, 1989; Tronicke et al., 2004; Busch et al., 2012).

The most common GPR measurement techniques include surface GPR and cross-borehole GPR measurements (Huisman et al., 2001). The center frequency of most GPR antennae is in the high frequency range 20-1000 MHz with a corresponding dominant wavelengths of 5 - 0.1 m for common earth materials. Surface GPR is a non-invasive measurement method in which both transmitter (Tx) and receiver (Rx) antennae are located on the surface. Thereby, two types of measurements can be conducted. The common midpoint (CMP) or the wide-angle reflection and refraction (WARR) setup are applied to derive subsurface velocity distribution at the point scale. The CMP or WARR sounding method are the EM equivalent to seismic refraction and wide-angle reflection, where the transmitter and receiver spacing are varied (Jol, H.M. ed., 2008). The second type of measurements setup is the common-offset profiling (COP) that can be applied at larger scale along profiles or grids. For COP mode, which is commonly applied in the surface GPR data, are acquired with a fixed spacing between the transmitter and the receiver antennae (e.g., Liu et al., 2018; Klotzsche et al., 2018). In contrast to the CMP or WARR measurements, the COP is much faster but can only provide information about lateral changes. To derive accurate depth information from COP data, CMP or WARR measurements are necessary to link the structures to velocity distributions.

The crosshole GPR measurement technique emits electromagnetic pulses from the transmitter antennae in one borehole, which are received from receiver antennae in a second borehole. Commonly two types of measurement setup are possible for crosshole GPR. The zero-offset profiling (ZOP) approach is a quick and simple survey method that can locate velocity anomalies or attenuation zones (Gilson et al., 1996; Binley et al., 2001) by synchronously moving the two antennae up or down in the two different boreholes with a constant spacing (Figure 1.1a). However, the results obtained from the ZOP measurements cannot distinguish the horizontal heterogeneity in the subsurface. To receive more scattered waves caused by reflections and refractions in a heterogeneous subsurface (Figure 1.1c), the more advanced measurement technique of multi-offset-gathers (MOG) measurement (Figure 1.1b) is needed. In this surveying mode, the transmitter antenna stays fixed in one borehole, while the receiver antenna is moved down or up with a constant spacing in a second borehole. This is repeated for multiple transmitter locations and can additionally be done in the other borehole by interchanging of the boreholes positions for transmitter and receiver. The resulting data set has a large number of rays with a larger number of angles that cover the domain between the two boreholes (Huisman et al., 2001; Klotzsche et al., 2019a). In contrast to surface techniques, crosshole applications of GPR are well suited to characterize the shallow subsurface and are beneficial for monitoring of aquifers (e.g., Paz et al., 2017). Because of the known borehole distance between the two antennae, the dense ray-coverage, and relatively small acquisition errors, crosshole GPR is a good constraint for sophisticated inversion schemes (e.g., Axtell et al., 2016; Klotzsche et al., 2019b). Note that both GPR measurement techniques are able to receive reflected, refracted, and direct waves.

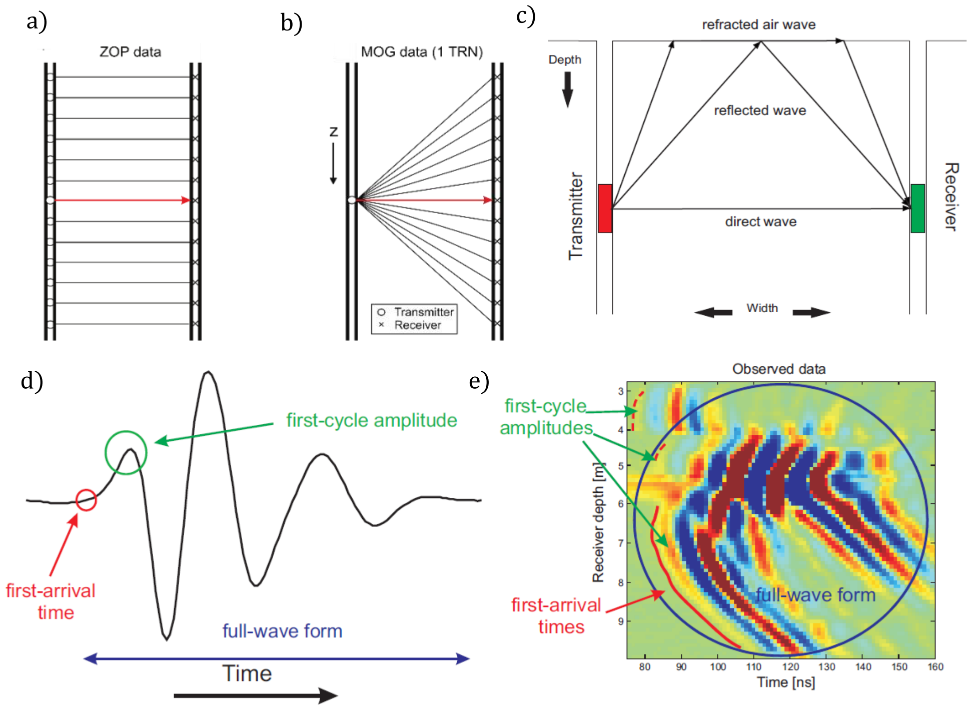


Figure 1.1. a) Illustration of a transillumination zero-offset profiling (ZOP). b) The setup of one MOG data set for one transmitter (TRN) position. c) Illustrations of direct, refracted, and reflected waves that can be measured between two boreholes. d) And e) show differences between the ray-based method (red and green) and the full-waveform inversion (blue) input parameters. Red and green arrows indicate first arrived times and first-cycle amplitudes locations, respectively. In which d) trace of one receiver and e) all receiver traces for one transmitter position of MOG data. Figure 1.1a and b are adapted from Klotzsche et al. (2019b). Figure 1.1c, d and e are adapted from Klotzsche (2013).

To derive the distribution of the subsurface structures, tomographic inversion algorithms can be applied to the crosshole GPR data. The most commonly applied inversion algorithms are based on the ray-based approaches by utilizing the first-arrival time and first-cycle amplitude information of the data (indicated by red and green in Figure 1.1d and e). Thereby, damping and smoothing constraints are necessary to stabilize the inversion (e.g., Holliger et al., 2001; Maurer & Musil, 2004). The information of first-arrival time can be used to calculate the velocity distribution of the subsurface (or permittivity), while the damping of the EM waves (conductivity) distribution in aquifers is derived from first-cycle amplitude information (e.g., Tronicke et al., 2001, 2004; Irving and Knight, 2005; Clement and Barrash, 2006; Musil et al., 2006; Paasche and Tronicke, 2007). However, the spatial resolution of a standard ray-based inversion for crosshole GPR data is limited, because it uses only a small amount of the data and the resolution scales approximately with the diameter of the first Fresnel zone $\sqrt{\lambda L}$ (Williamson, 1991), where λ and L are the dominant wavelength and the wave propagation path length, respectively. Most often ray-based imaging is not able to resolve targets with a size smaller than the dominant wavelength of the antennae (e.g., Holliger et al., 2001; Irving et al., 2007). Furthermore, ray-based approaches only consider data until a certain angle to avoid an increasing apparent-velocity for increasing ray path angles (Peterson, 2001), therefore only a limited angular coverage of the target is available.

To improve the resolution of crosshole GPR inversions, a number of waveform-type approaches have been developed, such as weak-scattering iterative methods based on integral representations of Maxwell's equations (e.g., Wang and Chew, 1989), the approximate wave-equation traveltime (e.g., Cai et al., 1996), Fresnel-volume (e.g., Johnson et al., 2005), and diffraction tomography methods (e.g., Cui and Chew, 2000 and 2002). Inspired by the full-waveform inversion (FWI) approach that was first developed and applied in the seismic community (e.g., Tarantola, 1984, 2005; Shin & Cha, 2008; Virieux & Operto, 2009), the 2D time-domain FWI for crosshole GPR data was implemented by Ernst et al. (2007a, b) and Kuroda et al. (2007). Meles et al. (2010) improved the method of Ernst et al. (2007a) by including the vector characteristics of the EM fields and introduced a simultaneous update of the permittivity and electrical conductivity parameters. Additionally, the vector properties enabled combining surface GPR and borehole GPR measurements in the inversion. To tame the non-linearity issue of GPR data FWI caused by high contrast media, Meles et al. (2011) introduced an approach using the progressive bandwidth expansion synthetic data (PBED) by applying bandpass filter to the effective source wavelet. The 2D time domain FWI with vector properties based on Meles et al. (2010) was successfully applied to experimental data sets from different test sites, for example to the River Thur test site in Switzerland (Klotzsche et al., 2010, 2012, and 2013) and to the Boise Hydrogeophysics Research Site in the USA (Yang et al., 2013; Klotzsche et al., 2014). Additionally, the measurement efficiency and computational costs of the FWI were improved by changing from a dense one-sided to a semi-reciprocal measurement setup (Klotzsche et al., 2010; Oberröhrmann et al., 2013) and Yang et al. (2013) introduced normalized gradients that are independent from the number of transmitters and receivers. Klotzsche et al. (2012) were able to identify a preferential flow path within an aquifer using the FWI results together with logging data. Klotzsche et al. (2014) combined information gained from an amplitude analysis approach with the starting models of the full-waveform inversion and showed the potential to detect small scale contrast structures in aquifers with both methods. At Krauthausen test site in Germany, Gueting et al. (2015) employed CPT measurements to verify GPR FWI results and introduced a clustering approach of the data to identify lithological structures of the aquifer. This study indicates that combining the 2D FWI results with the 1D CPT data can help to improve FWI imaging results along the GPR cross-section. Furthermore, by stitching 15 GPR tomograms together from multiple adjacent crosshole planes, they applied the 2D FWI clusters to derive a 3D larger scale facies model of the entire Krauthausen site. Thereby, they were able to explain a plume splitting that was previously observed during a salt tracer test experiment (Gueting et al., 2017).

Generally, FWI approaches can be implemented in time and frequency domain. Both methods have pros and contrasts. While frequency approaches can highly minimize the calculation costs and improve the cycle skipping problem, the time domain approaches are more mature because it provides the most flexible framework to apply time windowing of arbitrary geometries. This is especially true for 3D problems (Virieux & Operto, 2009). Similar to seismics, also frequency-domain FWI approaches for GPR data were implemented, because of certain benefits such as a few discrete frequencies of data are benefit for frequency-dependent medium properties and a wider range of misfit functions that can be implemented (e.g., Lavoue et al., 2014). Ellefsen et al. (2011) inverted measured crosshole GPR data from a laboratory tank by using a 2.5D frequency-domain FWI approach. Furthermore, a 2D frequency-domain quasi-Newton approach for multioffset GPR data was implemented to a synthetic model (Lavoue et al., 2014) and carbonate rocks data (Pinard et al., 2016). Although frequency-domain FWI seems to have certain benefits compared to time-domain FWI, for many experimental GPR data, the low frequency data is missing or shows a low signal-to-noise ratio. Until now almost all successful applications to experimental data have been performed using the time-domain FWI approach for GPR data (Klotzsche et al., 2019b).

The 2D time-domain crosshole GPR FWI based on Ernst at al. (2007a) and Meles at al. (2010) uses a conjugate-gradient algorithm (Polak et al., 1969) to optimize the misfit function between the measured and modeled data. This inversion approach is with the inevitable ill-posed and non-linear problems due to multiple scattering in the heterogeneous subsurface (e.g., Mora, 1987). Both of these problems are easily causing the FWI trapped into local minima convergence. Especially if the starting models differ too much from the true models and they exceed the half-wavelength criteria, the inverted results fail to converge to the true solution. The half-wavelength criterion requires that the modeled data based on the starting models are within half a wavelength of the measured data. To tame these problems, a number of solutions were proposed. The frequency hopping method was used in frequency domain for microwave and seismics (Chew and Lin, 1995; Pratt et al., 1998; Zhou and Greenhalgh, 2003; Dubois et al., 2009; Maurer et al., 2009). To mitigate the cycle skipping issue in the FWI of seismic data, some researchers have proposed to generate artificial low frequencies information for measured seismic data lack of low frequency (Shin et al., 2008; Choi et al., 2018). They inverted the zero-frequency damping wavefield to obtain low-wavenumbers structures of the starting models. In addition, the artificial low-frequency components hidden in the seismic envelope were able to be estimated based on the modulation signal model (Wu et al., 2014). Recently, Meng et al. (2019) have proposed an adaptive Laplace domain waveform inversion to build more suitable starting models. However, for many field experimental GPR data, the low frequency data was a low signal-to-noise ratio because of noises contaminated. To avoid the local minima convergence for the crosshole GPR data FWI, Meles et al. (2011) presented a combined frequency-time-domain approach that uses the progressive bandwidth expansion of the modeled data (PBED) as iterations proceed, while the observed data keep the full bandwidth. This approach worked very well for synthetic data, but for experimental data the choice of the frequency bands and the inversion parameters such the perturbation factors hindered a successful application so far.

In the presence of special subsurface structures such as high contrast small-scale heterogeneities, the 2D time-domain crosshole GPR FWI can easily result in incorrect inverted structures because of ill-posed and non-linear problems. Meanwhile, these small-scale high contrast layers, which are often caused by, e.g., changes in porosity (higher permittivity) or clay content (higher electrical conductivity), are critical for improving our ability to detect and visualize for hydrological processes in aquifers. Klotzsche et al. (2014) introduced the amplitude analysis approach that could identify and map the boundaries of sub-wavelength high contrast zones caused by an increase in permittivity/porosity already in the measured data. Such zones can act as low-velocity EM waveguides for the GPR data (e.g., Arcone, 1984; Arcone et al., 2003; van der Kruk et al., 2009, 2010; Klotzsche et al., 2012, 2013; Strobach et al., 2013) causing characteristic wave propagation behavior like late arrival high amplitude elongated wave trains in the data. However, waveguides related to clay content changes are more difficult to detect in the GPR data because of lacking obvious features caused by the higher damping of the waves and hence diminished amplitudes.

In addition, some preprocessing steps are necessary for experimental crosshole GPR data such as an accurate time zero estimation, defining good starting models, considering a 3D to 2D conversion of the data, and estimating an effective source wavelet. Although many studies show the potential of FWI for experimental GPR data, to apply the FWI to experimental data is still challenging due to the higher inverted resolution, data acquisition with noises, and detailed preprocessing and inversion of the data need to be performed carefully (e.g., Klotzsche et al., 2019b).

Thesis objectives and outline

The primary objective of this thesis is to improve the reconstruction of crosshole GPR FWI results and to make the inversion more robust against processing errors. This is achieved by extending, for example the amplitude analysis approach for high contrast structures of the subsurface from measured GPR data, including additional information of CPT data to the FWI improvement, and using the adaption progressive frequency-bandwidth expansion approach for both modeled data and observed data.

After the introduction part, the second Chapter of this thesis presents the fundamentals of the electromagnetic wave propagation (Maxwell's equations), the ray-based approach, and the full-waveform inversion algorithm. Some preprocessing steps of the full-waveform inversion for experimental crosshole GPR data with focuses on the 3D to 2D transformation of the measured data, the estimation of the starting models and the effective source wavelet estimated are explained in detail.

The third Chapter of this thesis presents a combination of the amplitude analysis and the standard GPR FWI. Due to the difficult to detect clay lens in the measured data by using the amplitude analysis approach, we extend the amplitude analysis to identify two different types of low-velocity waveguides either caused by an increased porosity or and a higher electrical conductivity. The standard FWI is performed according to the ray-based starting models and the standard effective source wavelet. To enhance the FWI results, we use the standard low iterations FWI permittivity results (iteration=10), which are able to provide approximate waveguide structures locations, as a new permittivity starting model. After updating the effective source wavelet, the updated FWI results obtained a better reconstruction of the model. The new approaches were tested at the Hermalle-sous-Argenteau test site, where nine boreholes GPR data sets were acquired. We inverted each 2D cross-section separately and integrated these cross-sections together in pseudo 3D view. Consistent structures in the nine permittivity and conductivity cross-sections showed the robustness of the updated inversion results and allowed the interpretation of a previously performed heat tracer experiment.

In Chapter 4, we improve the FWI results by using 1D CPT data. Firstly, a bandpass filter was generated and applied to the crosshole 2D FWI permittivity data. After that, we updated a new effective source wavelet based on the 2D wavenumber amplified FWI permittivity model by using the previous filter. Using this updated effective source wavelet and the standard ray-based starting models, the new FWI was performed. In further, the traditional and the updated FWI results were compared with the CPT data. Finally, the updated permittivity FWI results showed better reconstructions. To further verify the amplified approach with CPT data, we test one synthetic model and five measured crosshole GPR data sets from a test site in Krauthausen, Germany.

In Chapter 5, we present a scheme to tame the nonlinearity problem for the 2D time-domain crosshole GPR full-waveform inversion. We propose an idea of progressively expanding the bandwidths of both modeled and observed data (PEBDD). This new scheme is able to avoid some defects of the approach in Meles et al. (2011), such as the difficult determination of perturbation factors. In addition, for experimental GPR data, the approach of Meles et al. (2011) is unavailable. To verify our new scheme, we test two different synthetic case studies and applied the approach to experimental crosshole GPR data sets of the Krauthausen test site. Both synthetic and experimental data showed that the PEBDD scheme improved the FWI results by taming the local minima convergence problem.

Final conclusions and an outlook for further work are presented in Chapter 6. The outlook provides topics for future research. One is to recover low-wavenumber information of starting models based on high-frequency observed data by applying the approach of angle difference identity for Cosine (Wang et al., 2019), and the second one is to improve the computation of gradient directions in the process of FWI using the seismic staining algorithm (Chen and Jia, 2014; Li and Jia, 2017).

Chapter 2

Theory

In this Chapter, firstly Maxwell's equations are introduced, which mathematically describe the EM wave propagation and provide the foundations of achieving the subsurface medium information through using the EM waves traveling. Afterwards the ray-based inversion schemes including first-arrival travel time inversion and first-cycle amplitude inversion are described. Finally, the details of full-waveform inversion for experimental crosshole GPR data are presented based on Klotzsche et al. (2019b).

2.1 MAXWELL'S EQUATIONS OF ELECTRODYNAMICS

GPR is based on the fundamentals of the electromagnetic wave propagation. It is well known that the electric field originates from electric charges and the magnetic field from current loops. According to Maxwell's equations, we are able to find relations between the electric field $\mathbf{E}$, the magnetic field $\mathbf{H}$, time t, space $\mathbf{x}$, and material related equations such as Ohm's law. The propagation and affiliation of the electric and magnetic fields are applied for investigating the interactions of the fields with objects (i.e., the structures in the subsurface). The Earth's near-surface structures are able to be interpreted through analyzing these received electric and magnetic fields signals. By simulating electromagnetic wave propagation, we can invert and update the related material properties. According to Meles et al. (2010), the partial differential system of Maxwell's equations is defined at any point in time t and space $\mathbf{x}$ by:

$$\begin{cases} \partial_t \mathbf{B}(\mathbf{x},t) + \nabla \times \mathbf{E}(\mathbf{x},t) = 0; \\ \nabla \times \mathbf{H}(\mathbf{x},t) - \partial_t \mathbf{D}(\mathbf{x},t) - \mathbf{J}_{ex}(\mathbf{x},t) - \mathbf{J}_c(\mathbf{x},t) = 0; \\ \nabla \cdot \mathbf{D}(\mathbf{x},t) = \rho(\mathbf{x},t); \\ \nabla \cdot \mathbf{B}(\mathbf{x},t) = 0; \end{cases} \tag{2.1}$$

with

$\mathbf{E}(\mathbf{x},t)$ = Electric field intensity (V/m),

$\mathbf{H}(\mathbf{x},t)$ = Magnetic field intensity (A/m),

$\mathbf{D}(\mathbf{x},t)$ = Electric displacement ($\mathrm{C/m^2}$),

$\mathbf{B}(\mathbf{x},t)$ = Magnetic induction ($\mathrm{W/m^2}$),

$\rho(\mathbf{x},t)$ = Volume charge density ($\mathrm{C/m^3}$),

$\mathbf{J}_c(\mathbf{x},t)$ = Conduction (induced) current density ($\mathrm{A/m^2}$),

$\mathbf{J}_{ex}(\mathbf{x}, t)$ = External (source) current density $(\mathrm{A/m^2})$.

To connect the electric and magnetic fields, constitutive relationships are constructed based on the three fundamental bulk electromagnetic properties of material media, which are the dielectric permittivity ε, the magnetic permeability μ, and the electrical conductivity σ. To simply show the relationship between the electric field $\mathbf{E}(\mathbf{x}, t)$ and bulk electromagnetic properties (similar equations apply for $\mathbf{H}(\mathbf{x}, t)$, $\mathbf{D}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$), a damped wave equation with the partial differential form is presented:

$$\nabla^2 \mathbf{E}(\mathbf{x}, t) = \mu\sigma\partial_t \mathbf{E}(\mathbf{x}, t) + \mu\varepsilon\partial_{tt}\mathbf{E}(\mathbf{x}, t). \tag{2.2}$$

For most GPR applications, magnetic field $\mathbf{H}$ is not considered, and therefore a simplified formulation of the electric field $\mathbf{E}$ is given by (Meles et al., 2010)

$$\mathbf{E}^s = \mathbf{G}\,\mathbf{J}^s, \tag{2.3}$$

where the superscript s is used for the particular source, and $\mathbf{G}$ represents the Greens operator and describes the propagation of the electrical field through the medium. The explicit formulation of Equation 2.3 for a specific time–space point $(\mathbf{x}, t)$ is given by

$$\mathbf{E}^s(\mathbf{x}, t) = \int_V dV(\mathbf{x}') \int_0^{T_{max}} dt'\, \mathbf{G}(\mathbf{x}, t, \mathbf{x}', t')\; \mathbf{J}^s(\mathbf{x}', t'), \tag{2.4}$$

where $\mathbf{E}^s(\mathbf{x}, t)$ is the electric field generated by the source $\mathbf{J}^s(\mathbf{x}', t')$ at a specific time and space point ($\mathbf{x}$, t) and T_{max} is the maximum observation time. V represents the model space. $\mathbf{G}(\mathbf{x}, t, \mathbf{x}', t')$ is the Green's tensor that acts on the source term $\mathbf{J}^s(\mathbf{x}', t')$, which can be defined as $\mathbf{J}^s = \delta(\mathbf{x} - \mathbf{x}_s).\mathbf{S}(\omega)$, where $\mathbf{S}$ is the source wavelet and the δ function shows the position.

Generally, for medium with high frequencies, low loss, and non-magnetic property, the GPR wave velocity v in medium can be calculated by using the real part of the bulk dielectric permittivity ε in nature medium and the air wave velocity c with 0.29979 m/ns by

$$v = \frac{c}{\sqrt{\varepsilon}}. \tag{2.5}$$

Typically for GPR applications, the relation between the electrical conductivity σ and the amplitude attenuation α for GPR data for the same high frequency low loss attenuation medium structures is used ($\mu \approx 1$ with $\mu = \mu_0 . \mu_r$ with $\mu_0 = 4 . \pi . 10^{-7}$) that is described by

$$\sigma = 2\alpha\sqrt{\frac{\varepsilon}{\mu}}. \tag{2.6}$$

2.2 RAY-BASED METHODS

Conventional crosshole GPR tomographic inversion is based on geometrical ray theory (e.g., Maurer & Musil, 2004; Irving et al., 2007; Dafflon et al., 2011, 2012). Thereby, the first-arrival times and the first-cycle amplitudes of measured GPR traces are considered in the inversion process to derive velocity and attenuation models. In this thesis, we applied the finite-difference Eikonal algorithm (Vidale, 1990) for the forward modeling of the ray-based approach, which demonstrated to be more suitable for the heterogeneous mediums.

The subsurface velocity model is derived based on picked first-arrival times for crosshole GPR data. To solve the velocity distributions in the subsurface, it is necessary to simulate the ray-paths between the transmitting and receiving antennae. By varying input parameters and performing forward modeling according to a finite difference implementation of the Eikonal Equation (Vidale, 1990), a velocity distribution can be obtained by minimizing the misfit function between the observed t_{sr}^{obs} and the calculated travel times t_{sr}^{syn}. According to Lanz et al. (1998) and Rabbel (1996), the travel time misfit function C_{TT} can be described by

$$C_{TT} = \sum_s \sum_r \frac{(t_{sr}^{obs} - t_{sr}^{syn})^2}{n}, \tag{2.7}$$

where n, s and r represent the data points, the source and receiver numbers, respectively. t_{sr}^{obs} and t_{sr}^{syn} are the observed and calculated first-arrival times at the locations with transmitter x_s and receiver x_r, respectively.

To solve the calculated first-arrival times t_{sr}^{syn} based on the present model value parameters, we describe the propagating time along a ray path S between the transmitter and receiver as (e.g., Lanz et al., 1998)

$$t = \int_S \mathbf{u}(r(\mathbf{x}, \mathbf{z}))\, dr, \tag{2.8}$$

where $\mathbf{u}(r(\mathbf{x}, \mathbf{z}))$ is the slowness field with $r(\mathbf{x}, \mathbf{z})$ location vector. In further, the slowness field $\mathbf{u}(r(\mathbf{x}, \mathbf{z}))$ can be approximated by using m equidimensional cells, each having a constant slowness u_k (k=1…m), so the ith traveltime of n observations can be described as

$$t_i = \sum_{k=1}^{m} l_{ik}\, u_k = \mathbf{L}_i \mathbf{u}, \tag{2.9}$$

where l_{ik} denotes the ith portion of the ray path for in the kth cell. Determination of matrix **L** requires calculating traveltimes in 2D media.

In most cases for heterogeneous medium, regularizations with smoothing and damping constrains are necessary to solve the travel time inversion by using the finite-difference Eikonal solver. Based on Musil et al. (2003), Equation 2.9 can changed as

$$\begin{bmatrix} t \\ 0 \\ u_0 \end{bmatrix} = \begin{bmatrix} \mathbf{L} \\ \mathbf{A} \\ \mathbf{I} \end{bmatrix} \mathbf{u} \ , \tag{2.10}$$

where u_0 is a damping constraints vector, **A** is a smoothing matrix and the identity matrix is represented by **I**. A more compact form can be shown as

$$\mathbf{d} = \mathbf{Du}. \tag{2.11}$$

The misfit function C_{TT} is minimizing during the inversion process by searching minimum and is non-linear because $\mathbf{L}$ relies on the unknown slowness field $\mathbf{u}$. Therefore, Equation 2.11 is iteratively solved. Considering $\mathbf{D}$ is a sparse matrix, some sparse matrix solvers, such as a least squares (LSQR) approach based on Paige and Saunders (1982), can be applied to solve the sparse least squares problems (Lanz et al., 1998).

For the picked first-cycle amplitudes from crosshole GPR data, we can estimate the subsurface electrical conductivity distribution from the amplitude attenuation inversion based on ray-based methods. The inversion of the amplitudes needs a priori assumptions about the GPR antennae radiation patterns. The most common approach is that the radiation patterns of antennae are infinitesimal dipoles that correspond to the prevailing in a homogeneous medium (e.g., Holliger et al., 2001). The corresponding ray-based paths based on the picked first-arrival times are necessary for the amplitude inversion. Note that we take a homogeneous conductivity starting model to replace the inverted first-cycle amplitudes results in this thesis, thereby much more details can be referenced from Holliger et al. (2001) and Maurer & Musil (2004).

For the ray-based methods some critical shortcomings associated with the inherent high frequency limitations appear in the tomographic results. For instance, the ray-based tomographic inversion is only relatively smooth images and the resolution scales approximately with the diameter of the first Fresnel zone (Williamson, 1991; Williamson and Worthington, 1993). In addition, ray-based approaches only consider data until a certain angle to avoid an increasing apparent-velocity for increasing ray path angles (Peterson, 2001), therefore only a limited angular coverage of the target is available. The third point is that damping and smoothing constraints are necessary to stabilize the application of ray-based inversion (e.g., Holliger et al., 2001; Maurer & Musil, 2004). Although some inevitable shortcomings exist, for experimental crosshole GPR data, ray-based inversion results (especially for relative permittivity) are used to provide starting models for the full-waveform inversion (Klotzsche et al., 2019b).

2.3 FULL-WAVEFORM INVERSION

In contrast to ray-based methods, full-waveform inversion (FWI) takes the entire waveform of GPR data into account, which includes secondary events like scattered and refracted waves (e.g., Meles et al., 2010; Klotzsche et al., 2010). Therefore, the FWI promises a better imaging capability. However, it is well known that the FWI problems include both ill-posed and non-linear. Therefore, good starting models and an accurate effective source wavelet are important. As mentioned before, the starting models together with the effective source wavelet need to yield modeled data within half a wavelength of the measured data in the entire inversion domain to avoid cycle skipping and trapping of the inversion in a local minimum. Generally, the FWI scheme is very consuming computational costs, thus the main limitation factor for FWI development was computer computing power in the early years. With the developments of parallel programming tools, and the availability of high performance cluster, the limitation has been solved. Inspired by the FWI works in seismic domain, FWI was developed and progressively applied in EM domain based on Maxwell’s equations by using finite-difference methods. In this thesis, we applied the FWI approach of Meles et al. (2010), who further improved the approach of Ernst et al. (2007a, b) with modifications from a stepped (cascaded) inversion scheme to simultaneously updating parameter scheme.

2.3.1 Pre-processing

For experimental crosshole GPR data, some pre-processing steps are critical. First of all, the raw GPR data are filtered to remove low-frequency noise. The raw data is affected by a slowly decaying low frequency 'wow', caused by signal saturation due to early wave arrivals, the electrical properties of the ground, inductive coupling effects, and/or the proximity of the transmitter and receiver. This low frequency 'wow' superposes the high frequency signal, and needs to be filtered out by using a highpass filter, called 'dewow'. The dewow method applies a running average filter to each trace, and is able to pass the transmitted signal spectral peak for the specific antenna center frequency and to suppress the low frequency wow in the data. In addition, accurate borehole deviation information is indispensable and artefacts occurred when uncorrected coordinates are used (Maurer & Green, 1997).

During the MOG crosshole GPR measurement, time-shifts or jumps in the first arrival time of the wave occur that are caused by thermal drift, electronic instability, cable length differences and variations in antenna coupling (Nobes, 1999; Olhoeft, 2000). Therefore, the exact starting time of the wave, which is called time zero, needs to be determined and the GPR traces need to be adjusted. In this thesis, we apply the method of time-zero correction by a cross-correlation of a ZOP with corresponding horizontal traces of each MOG based on Oberröhrmann et al. (2013). In the first step, the ZOP traces, which are additionally measured after all MOGs are acquired, need to be time corrected with a subsequently calibration CMP or WARR measurement in air. The calibration measurement is performed by placing the antennae at the surface and performing a CMP or WARR measurement to measure the air wave. Using the known air wave velocity, the time shift between the actual starting time and current time can be estimated. This time shift is applied to the ZOP data. Afterwards, for each ZOP trace, a corresponding MOG trace, which has traveled the same path between the same transmitter and receiver positions (see red arrow in Figure 1.1a and b), can be found. It can be assumed that the highest cross-correlation of the two traces shows the relative time shift that occurred in the MOG trace, caused by the time zero changes. Finally, we can correct the time-zero of MOG traces according to shifted times. After the data is filtered, time zero corrected and accurate deviations data is available, first-arrival travel times and first-cycle amplitudes can be picked and used in the ray-based inversion. Note that the same data is used in the FWI. The estimated velocity and attenuation tomograms of the subsurface based on ray-based methods are considered to obtain permittivity and conductivity distributions according to Equations 2.5 and 2.6 as starting models for the FWI.

Because the forward modeling and the inversion are performed in the 2D domain (data measured in 3D), it is necessary to understand that the 3D radiation characteristics of electromagnetic wave propagation are different in 2D. Therefore, the experimental GPR data needs to be convert from 3D to 2D using the approach by Bleistein (1986) to reduce the influence of the 3D wave propagation phenomena. The main reason is the geometrical spreading differences in 2D and 3D. For the 3D case, the wave front propagates spherical and the amplitude decay is proportional to the traveled distance. In contrast, the wave energy is spread over the perimeter of the circle in the 2D case. Therefore, multiplying the 3D data with $\sqrt{t}$ is necessary. Furthermore, a phase shift must be introduced to compensate an assumption with infinitive long extended sources in 2D that is not true in reality (e.g., Klotzsche, 2013). Similar to Ernst et al. (2007b) and Bleistein (1986), we use a scheme that compensate for differences in geometrical spreading and pulse shape. The transformation uses a phase shift of $\pi/4$ and a scaling factor of $1/\sqrt{\omega}$ in the frequency domain. The expressed equation is shown as following:

$$\hat{\mathbf{E}}^{2D}(\mathbf{x}_{trn}, \mathbf{x}_{rec}, \omega) = \hat{\mathbf{E}}^{obs}(\mathbf{x}_{trn}, \mathbf{x}_{rec}, \omega)\sqrt{\frac{2\pi t(\mathbf{x}_{trn}, \mathbf{x}_{rec})}{-j\omega\varepsilon^{mean}\mu}}, \tag{2.12}$$

where $\hat{\mathbf{E}}^{2D}(\mathbf{x}_{trn}, \mathbf{x}_{rec}, \omega)$ is the corrected 2D data for a transmitter at location $\mathbf{x}_{trn}$ and a receiver at $\mathbf{x}_{rec}$. $\hat{}$ indicates parameters are in frequency domain. $\hat{\mathbf{E}}^{obs}(\mathbf{x}_{trn}, \mathbf{x}_{rec}, \omega)$ represents 3D original GPR data. $t(\mathbf{x}_{trn}, \mathbf{x}_{rec})$ is the travel time; and ε^{mean} is the mean dielectric permittivity of the media. A good agreement can be achieved between 3D and corresponding pure 2D data in the far-field, which has been verified by Ernst et al. (2007b). The main reason of only valid in the far field and in slowly changing medium is the filter assumes that the high amplitude/energy is associated with the first break. The high contrast layers problem probably causes elongated late arrived amplitudes.

2.3.2 Forward Problem

A forward modeling tool is necessary in the process of both an effective source wavelet estimation and the full-waveform inversion. In this thesis, we applied a 2D FDTD solution of the Maxwell Equations in Cartesian coordinates based on Ernst et al. (2007a) and Meles et al. (2010). The main idea is to find a numerical solution of time-dependent differential Equation 2.1 by finite differences with grid-based in differential time domain. The generalized perfect matched layers (GPML) are implemented to reduce back-refection artefacts at the model boundaries (Berenger, 1994; Ernst et al., 2006). In the FWI process, we need to compute four times the forward modeling to compute the updating direction and step-lengths. Considering most crosshole cases, the standard vertical antenna orientations are used to measure the vertical component of the electrical field, we employ the transverse electrical (TE) model of the Maxwell Equations in this thesis.

2.3.3 Source wavelet estimation and correction

The estimation of an effective source wavelet is a critical step for the full-waveform inversion. The traditional effective source wavelet is based on ray-based starting models and the deconvolution method (e.g., Klotzsche et al., 2010). First, an initial source wavelet s_k is estimated from horizontally travelling waves of each transmitter, where only the shape of the wavelet is determined without considering any amplitude information. Because the electrical field is proportional to the time derivative of the current density source (multiplication of $i\omega$ in the frequency domain), we need to divide the averaged Fourier transformed pulse by $i\omega$ in the frequency domain to obtain the effective source wavelet (more details can be found in Klotzsche et al., 2019b). Second, we correct the initial source wavelet s_k for shape and amplitude based on a deconvolution approach. Thereby, synthetic data based on the starting models and initial source wavelet s_k are derived. Considering the modeled data can be expressed mathematically in the time domain as a convolution or in the frequency domain as a multiplication of the source wavelet and the Green's function. Thereby, the effective source wavelet s_{k+1} is solved based on the Green's function and the measured GPR data by using the deconvolution approach. The computational details can be found in the following deconvolution equations:

$$\hat{\mathbf{G}}(f) = \hat{\mathbf{E}}^{syn}(f)[\hat{s}_k(f) + \eta_D]^{-1}, \tag{2.13}$$

and

$$\hat{s}_{k+1}(f) = \left[\hat{\mathbf{G}}(f) + \eta_I\right]^{-1}\hat{\mathbf{E}}^{obs}(f)\,, \tag{2.14}$$

where s_k represents the initial source wavelet. $\mathbf{E}^{syn}$ is the forward modeled GPR data. $\mathbf{G}$ is the Green's function and $\mathbf{E}^{obs}$ is the measured GPR data. s_{k+1} the effective source wavelet with an optimized phase and amplitude. η_D and η_I are prewhitening factors that are applied to stabilize the solution and avoid dividing by zero. ^ indicates frequency domain. Note that we can refine the time-domain effective source wavelet $s_{k+1}(t)$ by repeating the deconvolution approach until the updated source wavelet fulfills the FWI converge before or during the FWI if necessary (e.g., Klotzsche et al., 2019b).

2.3.4 Inversion algorithm

The basic method of FWI is to seek an accurate model of the earth responds in terms of certain parameters (e.g., ε_r and σ) by minimizing the differences between the observed and the modeled data. Here, we apply a conjugate-gradient type algorithm for the time-domain crosshole GPR FWI to optimize the misfit function C between the measured and modeled data (Polak et al., 1969; Meles et al., 2010). The following squared misfit norm $C(\varepsilon, \sigma)$ describes the misfit between the observed $\mathbf{E}^{obs}$ and modeled synthetic $\mathbf{E}^{syn}$ GPR data over the number of transmitters s, receivers r, and the observation time τ (Tarantola et al., 2005):

$$C(\varepsilon, \sigma) = \frac{1}{2} \sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon, \sigma) - \mathbf{E}^{obs}]_{r,\tau}^T \, \delta\left(\mathbf{x} - \mathbf{x}_r, t - \tau\right) [\mathbf{E}^{syn}(\varepsilon, \sigma) - \mathbf{E}^{obs}]_{r,\tau}, \tag{2.15}$$

here T denotes the transpose operator. Each of the fields is locally defined at any point of space $\mathbf{x}$ and time t. The multiplication with the Dirac delta δ function selects from the entire wavefield the used receiver locations and observation times. Additionally, the gradients of the misfit function with respect to permittivity ∇C_ε and conductivity ∇C_σ are calculated by a zero-lag cross-correlation of the synthetic wavefield with the back-propagated residual wavefield $\mathbf{R}^S$.

$$\begin{bmatrix} \nabla C_\varepsilon(\mathbf{x}') \\ \nabla C_\sigma(\mathbf{x}') \end{bmatrix} = \sum_s \begin{matrix} (\delta\,(\mathbf{x} - \mathbf{x}')\partial_t \mathbf{E}^{syn})^T \hat{\mathbf{G}}^T \mathbf{R}^S \\ (\delta\,(\mathbf{x} - \mathbf{x}')\mathbf{E}^{syn})^T \hat{\mathbf{G}}^T \mathbf{R}^S \end{matrix} \tag{2.16}$$

with

$$\mathbf{R}^S = \sum_r \sum_\tau \delta\left(\mathbf{x} - \mathbf{x}_r, t - \tau\right) [\mathbf{E}^{syn}(\varepsilon, \sigma) - \mathbf{E}^{obs}]_{r,\tau} = \sum_r \sum_\tau \Delta\,[\mathbf{E}^{syn}]_{r,\tau}, \tag{2.17}$$

The spatial delta function $\delta\,(\mathbf{x} - \mathbf{x}')$ in Equation 2.16 corresponds to the spatial components of the gradients and reduces the inner product to a zero-lag cross-correlation in time (Meles et al., 2010). Note that a time derivative is the only difference between the gradients of ε and σ. The medium parameters can be updated to reduce the misfit function Equation 2.15 by updating all locations gradients values in Equations 2.16 and 2.17. These gradient values determine the updated directions for permittivity and conductivity models. To search how far to update in a direction, the step-length ζ is necessary to update the current model matrixes using

$$[\varepsilon_{upd}] = [\varepsilon] - \zeta[\nabla C_\varepsilon], \tag{2.18}$$

and

$$[\sigma_{upd}] = [\sigma] - \zeta[\nabla C_\sigma]. \tag{2.19}$$

Thereby, we need to choose the step-length ζ in a way to avoid overshooting of the full-waveform inversion or to reduce truncation errors (Meles et al., 2010). Using appreciate perturbation factor κ, the step-length ζ can be solved by:

$$\zeta = \kappa \frac{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon+\kappa\nabla C_\varepsilon,\sigma+\kappa\nabla C_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[\mathbf{E}^{syn}(\varepsilon,\sigma)-\mathbf{E}^{obs}]_{r,\tau}}{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon+\kappa\nabla C_\varepsilon,\sigma+\kappa\nabla \boldsymbol{C}_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[\mathbf{E}^{syn}(\varepsilon+\kappa\nabla C_\varepsilon,\sigma+\kappa\nabla C_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}}. \tag{2.20}$$

Considering large differences between permittivity and conductivity sensitivities the simultaneous inversion of ε and σ, with one step-length is not possible for complex subsurface structures. A stepped (cascaded) inversion approach was applied by Ernst et al. (2007a), in which the permittivity was inverted firstly, while the conductivity was kept constant for a certain number iterations. After that, the permittivity was fixed and the conductivity was inverted. In contrast, a simultaneous update of the permittivity and conductivity models in each iteration was introduced by Meles et al. (2010) using two different step-length calculations:

$$[\varepsilon_{upd}] = [\varepsilon] - \zeta_\varepsilon[\nabla C_\varepsilon], \tag{2.21}$$

and

$$[\sigma_{upd}] = [\sigma] - \zeta_\sigma[\nabla C_\sigma], \tag{2.22}$$

whereas the individual step-lengths $\zeta_\varepsilon$and ζ_σ can be calculated using:

$$\zeta_\varepsilon = \kappa_\varepsilon \frac{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon+\kappa\nabla C_\varepsilon,\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[\mathbf{E}^{syn}(\varepsilon,\sigma)-\mathbf{E}^{obs}]_{r,\tau}}{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon+\kappa_\varepsilon\nabla C_\varepsilon,\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[(\mathbf{E}^{syn}(\varepsilon+\kappa_\varepsilon\nabla C_\varepsilon,\sigma))-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}}, \tag{2.23}$$

and

$$\zeta_\sigma = \kappa_\sigma \frac{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon,\sigma+\kappa\nabla C_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[\mathbf{E}^{syn}(\varepsilon,\sigma)-\mathbf{E}^{obs}]_{r,\tau}}{\sum_s \sum_r \sum_\tau [\mathbf{E}^{syn}(\varepsilon,\sigma+\kappa_\sigma\nabla C_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}^T\ \delta(\mathbf{x}-\mathbf{x}_r,t-\tau)[\mathbf{E}^{syn}(\varepsilon,\sigma+\kappa_\sigma\nabla C_\sigma)-\mathbf{E}^{syn}(\varepsilon,\sigma)]_{r,\tau}}. \tag{2.24}$$

Note that the perturbation factors are critical and needed to be chosen carefully. Too large values possibly cause the perturbed mode to deviate the linearity range and can cause overshooting of the inversion; too small values can produce round-off errors of the computer system when dealing with small numbers (Meles et al., 2010).

Finally, the permittivity and conductivity are updated with the obtained updated directions and step-lengths using Equations 2.21 and 2.22. In this thesis, we stop the inversion when the root-mean-squared error (RMSE) change between the modeled and measure data is less than 0.5% between to subsequent iterations. Furthermore, these results are considered as optimal solution if no significant gradient for both parameters is present for this iteration, a good correlation between the measured and modeled data is reached, and the inversion converged.

2.3.5 Implementation details

Parametrization of the physical system

For the unknown parameters ε and σ, we applied a logarithmic scaled version based on Ernst at al. (2007a) via using the following Tarantola's approach

$$\tilde{\varepsilon} = log\frac{\varepsilon}{\varepsilon_0} = \log(\varepsilon_r), \tag{2.25}$$

and

$$\tilde{\sigma} = log\frac{\sigma}{\sigma_0} = \log(\sigma), \tag{2.26}$$

where σ_0 is 1 S/m and ε_0 is the dielectric permittivity of free space. It is necessary to adopt such as logarithmic scaling to ensure the models space with a linear structure and further to provide positive values for two parameters within a wider range of the values. In further, we can rewrite the gradient function combining with Equations 2.25 and 2.26 by using

$$C\big(\varepsilon(\tilde{\varepsilon}), \sigma(\tilde{\sigma})\big) = C'(\tilde{\varepsilon}, \tilde{\sigma}). \tag{2.27}$$

Normalized gradients

During the first applications of the FWI to experimental data for different test sites and setups, it was observed that the values of gradients depend on the number of transmitters and receivers. Therefore, to allow a direct comparison between different setups, normalized gradients for the permittivity and conductivity was introduced by Yang et al. (2013), which are independent of the number of sources N_S and receivers N_r. The details can be described by the following equations:

$$\nabla\bar{C}_\varepsilon(\mathbf{x}') = \frac{\nabla C_\varepsilon(\mathbf{x}')}{N_S N_r}, \tag{2.28}$$

and

$$\nabla \bar{C}_\sigma(\mathbf{x}') = \frac{\nabla C_\sigma(\mathbf{x}')}{N_S N_r}. \tag{2.29}$$

Furthermore this normalized gradients equation allows a direct comparison of the gradients and step-lengths for different crosshole survey layouts and results in fewer testing to define the optimal perturbation factors for the inversion (Yang et al., 2013; Klotzsche et al., 2019b).

Gradient preconditioning

Considering that the gradients are highly sensitivity close to transmitter and receiver positions, inversion artifacts can easily arise close to boreholes. The approach of Meles at al. (2010) applied a muting zone to avoid numerical artifacts close to the boreholes during the inversion. This muting zone is normally chosen to be two cells next both boreholes, which hindered a direct comparison with independent geophysical and hydrologic logging data. To minimize these inversion artifacts in the vicinity of the boreholes, the approach of Kurzmann et al. (2013) is applied using a gradient preconditioning (van der Kruk et al., 2015). To overcome this issue, the gradient preconditioning operator $\mathbf{P}^k$ for the updating domain $\mathbf{x}$ is defined and applied to the gradients using the maximum values of the forward propagated field and the back-propagated residual field:

$$\mathbf{P}^k(\mathbf{x}) = \frac{\mathbf{b}(\mathbf{x})}{\max_{\mathbf{x}} \mathbf{b}(\mathbf{x})}, \tag{2.30}$$

$$\mathbf{b}(\mathbf{x}) = \frac{1}{a(\mathbf{x}) + C_{stab}\bar{a}}, \tag{2.31}$$

and

$$a(\mathbf{x}) = \max_t |\mathbf{E}^{syn}| + \max_t |\mathbf{R}^S|, \tag{2.32}$$

where C_{stab} represents the stabilization factor that is selected as between 0 and 100. The terms $\mathbf{E}^{syn}$ and $\mathbf{R}^S$ are determined by space $\mathbf{x}$ and time t. The spatial average of $a(\mathbf{x})$ is indicated by $\bar{a}$. The first term and the second term on the right side of Equation 2.32 are considered as the maximum of the synthetic wavefield and the maximum of the residual wavefield, respectively.

Chapter 3

3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole GPR amplitude analysis and full-waveform inversion[1]

In this chapter, we explore the GPR amplitude analysis and the crosshole GPR full waveform inversion using a set of nine crosshole GPR datasets from a test site in Hermalle-sous-Argenteau near Meuse River in Belgium. We investigate the datasets to characterize the aquifer within a decimeter scale resolution and to improve the understanding of a previously performed heat tracer experiment. Thereby, we extend the amplitude analysis to identify two different types of low-velocity waveguides either caused by an increased porosity or/and a higher electrical conductivity. Combining the GPR amplitude analysis for low-velocity waveguide zones with the standard FWI results, provided information on waveguide zones which modified the starting models and enabled to improve the FWI results further. Moreover, an updated effective source wavelet is estimated based on the updated permittivity starting model. In comparison with the traditional FWI results, the updated FWI results present smaller gradients and smaller the root mean squared error values in the final inversion results. The nine crosshole sections are used to generate a 3D image of the aquifer and allowed a detailed analysis of the porosity distribution along the different sections. Consistent structures of the permittivity and electrical conductivity show the robustness of the updated FWI results. The aquifer structures obtained by the FWI results agree with previous results of the heat tracer experiment and are able to explain the heat tracer plume splitting in more detail.

[1]adapted from Zhou, Z., A. Klotzsche, T. Hermans, F. Nguyen, J. Schmäck, P. Haruzi, H. Vereecken, and J. van der Kruk, 2020a. 3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole GPR amplitude analysis and full-waveform inversion: Geophysics, under review.

3.1 FIELD SITE AND GPR MEASUREMENT SETUP

The study site is located on the alluvial plain of the Meuse River at Hermalle-sous-Argenteau near the city of Liege, Belgium (Figure 3.1a). In the saturated zone, between 3.0 m and 10.0 m depth, the aquifer is composed of gravel and pebbles in a sandy matrix (Hermans et al., 2015a; Lesparre et al., 2019). This layer can be divided in two main units: the upper aquifer from 3.0 m to 6.0 m depth, consisting of sandy gravels and the lower aquifer between 6.0 m and 10.0 m depth, which is characterized by coarser and cleaner gravels. The water table lies at approximately 3.2 m depth. Below 10.0 m depth, the aquitard consists of folded shales and sandstones. According to Hermans et al. (2015a), we can expect an electrical conductivity change of 5 mS/m to 10 mS/m between boreholes Pz13 and Pz17 (see Figure 3.1c). The heated water experiment of Hermans et al. (2015a), hot water was injected from Pz09 (red triangle in Figure 3.1b). The ERT results revealed the ability of the time-lapse ERT to monitor the variations of temperature in the aquifer as shown exemplary for an ERT plane after 30 hours of injection in Figure 3.1d, where the temperature variations clearly display two anomalies as a result of a heterogeneous flow field (Klepikova et al., 2016; Hoffmann et al., 2019).

Crosshole GPR measurements were performed at nine cross-sections at the same investigation area as the ERT monitoring (different color lines in Figure 3.1b) using 200 MHz PulseEKKO borehole antennae (Sensors & Software Inc.) in September 2018. A semi-reciprocal MOG acquisition setup was used with transmitter and receiver spacings of 0.2 m and 0.1 m, respectively, as proposed by Klotzsche et al. (2010, 2019b). We chose as 0.0 m depth the casing of borehole Pz09 similar to Hermans et al. (2015a). To avoid critical angle reflections and refractions of the GPR waves from the interface between the saturated domain and the groundwater table, the first antenna position was located at least 0.4 m deeper than the water table (Klotzsche et al., 2019a). Up to 60 transmitters and 120 receivers positions were used for each plane (see Table 3.1 for detailed information). Note that the plane between the boreholes Pz10-Pz17 was measured two months later than the others. One critical step in the FWI data pre-processing is the estimation of an accurate time zero of the signal (see Klotzsche et al., 2019b for more details). To reduce travel time errors, time-zero was determined using a cross-correlation between MOG and zero-offset (ZOP) data as proposed by Oberröhrmann et al. (2013). Firstly, wide-angle reflection and refraction (WARR) measurements in air before and after the multiple-offset (MOG) measurements are acquired. Secondly, a ZOP measurement is performed before the last WARR measurement (used for correcting the ZOP data); the individual time shifts of each MOG are obtained by cross correlating the ZOP traces with the corresponding horizontally travelled rays within the MOGs traces. To reduce borehole geometry errors in the measurements, borehole dip and azimuth deviation data were collected using a magnetic inclinometer tool (QL40-DEV from Mount Sopris Instrument Co.).

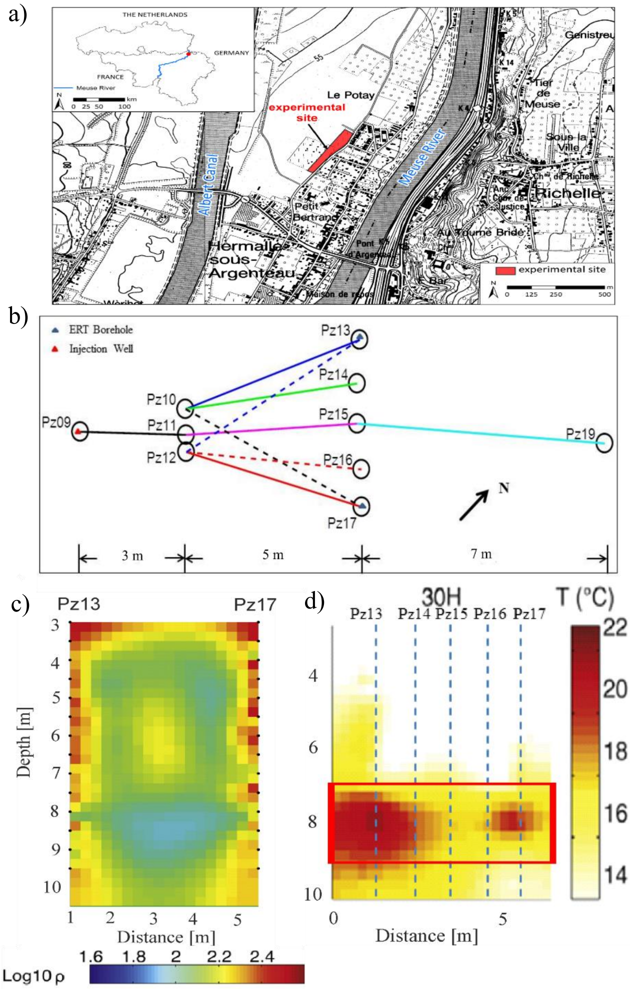


Figure 3.1. a) The site of Hermalle-sous-Argenteau is located at the northern part of the Meuse River in Belgium (modified from Hermans et al., 2015a). The red area indicates the experimental site. b) Schematic setup of the Hermalle-sous-Argenteau test site in Belgium indicating the boreholes used for the crosshole GPR measurements with Pz-numbers. The colored lines indicate the nine GPR cross-sections (modified from Klepikova et al., 2016). c) Background inverted ERT section showing the resistivity values between the boreholes Pz13-Pz17 (indicated by blue triangles in Figure 3.1b). d) ERT-estimated temperature between Pz13-Pz17 30 hours following hot water injection from Pz09 (modified from Hermans et al., 2015a). The location of the heat plume intrusion is indicated with a red rectangle.

3.2 GPR AMPLITUDE ANALYSIS RESULTS

3.2.1 Amplitude analysis approach

Guided electromagnetic waves can be formed in the near surface when a low-velocity layer with a thickness smaller than the electromagnetic wavelength is embedded between higher velocity layers (van der Kruk et al., 2009). Such small-scale high contrast layers can often be linked to high porosity layers because they have a lower velocity (high permittivity) than the surrounding medium. Such so-called waveguides cause most of the energy of the emitted wave to be trapped within the low-velocity layer if the propagating wave angle at the interface is larger than or close to the Snell critical angle. Klotzsche et al. (2014) observe that when a transmitter in a crosshole setup is located within such a waveguide zone, multiple internal reflections interfere constructively, causing late arrival high amplitudes in the observed GPR data. Furthermore, for transmitter positions outside the waveguide zone, a clear diminished amplitude gap could be observed close to the boundary of the waveguide zone. Using these characteristic wave-propagation features, they proposed an amplitude analysis approach that can detect low-velocity waveguides from the measured data. This amplitude analysis approach detects local maxima and minima positions within the trace energy (squared amplitude of the trace) spectra of each transmitter to first detect and second map low-velocity waveguides, respectively. In the first step, clear maxima positions (at least one order of magnitude higher) in the trace energy spectra are picked to identify a low-velocity waveguide. Secondly, in the trace energy spectra for transmitter which do not show maxima, local minima positions close to the previously found maxima position are estimated. Finally, the estimated maxima and minima picks are plotted against receiver depth to indicate the dimensions of the waveguide zone. Note that the maxima positions can only be observed if the electrical conductivity of the high contrast layers is not too large, which is the case for example for thin higher porosity layers. Until now this approach was only applied to experimental data to detect small-scale high contrast zones related to higher porosity layers. To distinguish between waveguides caused by porosity and clay content changes, the image plots of measured GPR data should be carefully analyzed for elongated wave trains and amplitude gaps in the data. If only amplitude gaps and significant minima positions are present in the GPR data, this could be an indicator for a low-velocity waveguide caused by an increase of electrical conductivity caused e.g., by a higher amount of clay. To sum up, we propose to distinguish in the first step on visual inspection of the data between two types of waveguides that could be identified with the amplitude analysis approach:

- Waveguide type I (WGT I): High permittivity and low/intermediate electrical conductivity that could be related to high porosity zones. Elongated wave trains and zones with diminished amplitudes can be found in the GPR data.

- Waveguide type II (WGT II): High permittivity and high electrical conductivity that could be related to layers with high clay content. No elongated wave trains can be detected due to higher attenuation of the electromagnetic wave in the waveguide zone, but zones with diminished amplitudes are present.

By combining the amplitude analysis with FWI results and independently measured logging data, the origin of such zones could be validated.

3.2.2 Detailed analysis for crosshole plane Pz10-Pz13

We applied the amplitude analysis to the measured GPR data from the Hermalle-sous-Argenteau site. We will explain the associated steps and the results in detail for one plane Pz10-Pz13 and summarize the finding for all planes. Plane Pz10-Pz13 was considered due to that both waveguide types can nicely be identified and explained in this dataset. In the first step, we analyze the measured data for the characteristic wave propagation features that could indicate waveguide zones (Figure 3.2). For transmitter locations at 8.09 m and 9.09 m in borehole Pz10 (Tx nr. 23 and 28), and, at 7.73 m and 9.13 m in borehole Pz13 (Tx nr. 20 and 27) between 7.7 m and 9.2 m depth in both boreholes very clear elongated wave trains can be observed (red circles). Therefore, we see indicators for at least two low-velocity waveguides with a higher permittivity and lower electrical conductivity (WGT I). Interestingly around the two elongated wave train features, zones with diminished amplitudes can be observed probably related to the other waveguides (see green circles with numbers in Figure 3.2e – h). Transmitter locations around 6.5 m depths (Tx nr. 15 in Pz10; Tx nr. 14 in Pz13) in both boreholes show no significant features in the data. Some minor diminished amplitude changes could be sensed at 4.0 m, 6.0 m, and 7.0 m depth for transmitter at 6.5 m in Pz13 (blue and light blue circles Figure 3.2d). For the transmitter positions at 3.9 m depth (Tx Pz13=1) in borehole Pz13 a zone at around 4.0 m can be observed, which shows diminished amplitudes in contrast to the surrounding. Such features with no clear elongated wave trains close by could be an indicator for a low-velocity waveguide with a higher permittivity and a higher electrical conductivity (WGT II).

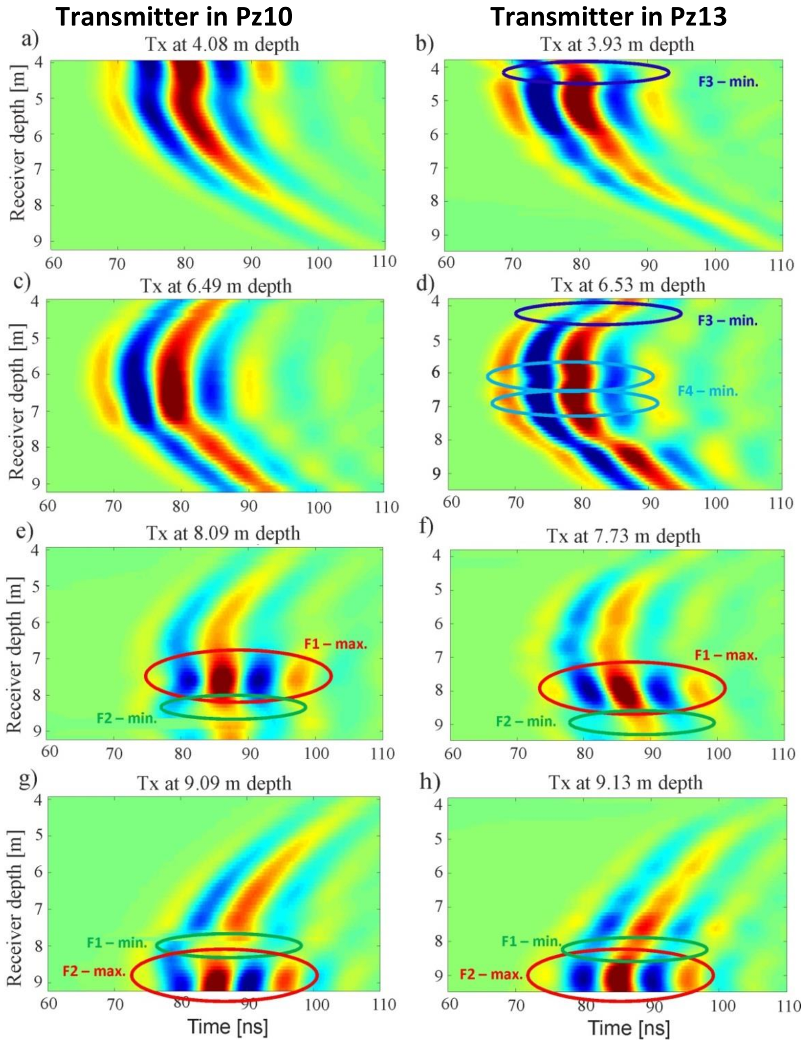


Figure 3.2. Image plots of the measured data of the cross-section Pz10-Pz13 for different transmitter (Tx) locations in both boreholes. Different waveguide features are indicated with F1-F4, while WGT I features are marked with red and green circles and WGT II features with blue and light blue circles. Amplitudes of each image plot are normalized to the maximum value of the amplitudes for the cross-section and range from -7×10^{-1} to 7×10^{-1}.

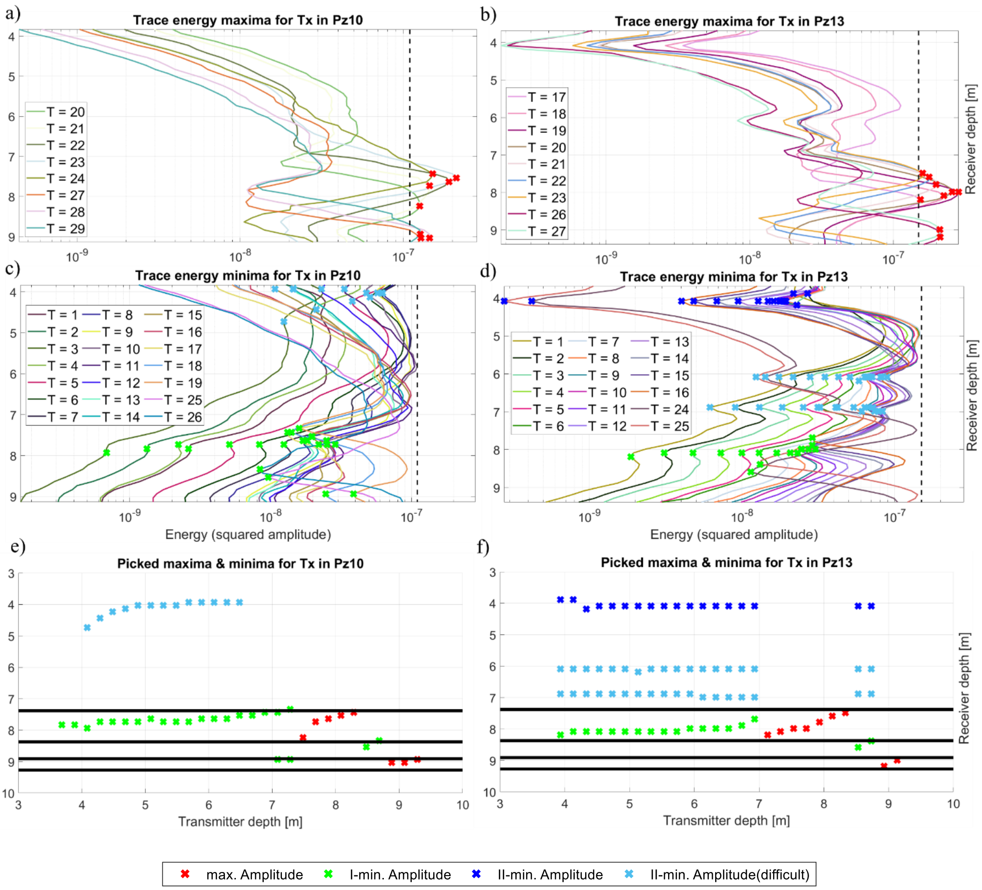


Figure 3.3. Trace energy profiles of the measured data of Pz10 and Pz13 that show clear a) - b) maxima and c) - d) minima. Each transmitter has a certain color affiliation in a) - d) and the vertical dashed lines represent the selected energy thresholds. Picked positions of the trace energy spectra for the Tx data in e) Pz10 and f) Pz13. The red and green crosses indicate the position of the maxima and local minima of the energy caused by waveguides WGT I. The blue (and light blue) crosses represent the local minima energy locations caused by possible WGT II. The black solid lines indicate the boundaries of waveguide structures of WGT I using amplitude analysis.

In the second step, we calculate the energy spectra for each transmitter gather and plot them against receiver depth. First, spectra that show clear maxima beyond a certain threshold (1.1×10^{-7} for transmitter in Pz10 and 1.5×10^{-7} for transmitter in Pz13) are identified for both boreholes (Figure 3.3a and b). We used the logarithmic scale along the energy direction to better show the energy changes. The peak of the maximum is picked (red crosses) for each transmitter position that is used to identify a low-velocity waveguide. For the borehole pairs Pz10 and Pz13 two waveguide zones of the WGT I can be identified. Second, all remaining spectra are plotted (Figure 3.3c and d) and distinct local minimum are picked with green crosses around the previously found maxima positions. Note that in the plots that shows the maxima in the spectra also the minima of the other features can be observed (not shown). Next to the two defined waveguide structures below 7.5 m depth, a very clear zone with minima (blue crosses in Figure 3.3d) can be observed at around 4.0 m depth for the transmitter in Pz13. Since no elongated wave trains (Figure 3.2b) and no maxima at this location can be observed, these minima could be an indicator for a zone with a higher permittivity and conductivity of a WGT II. Between 6.0 m and 7.0 m depth in Pz13 two zones with minor minima positions can be identified. These minima structures are only observed in one borehole, which indicates that the detected structures are not continues between the boreholes as demonstrated by Klotzsche et al. (2014). The obtained positions of maxima and minima are plotted against receiver depth to estimate the dimension of the different waveguide zones (Figure 3.3e and f). Note that only the boundaries of the waveguides close to the borehole can be obtained.

3.2.3 Amplitude analysis for all measurement planes

Similar to the amplitude analysis for boreholes Pz10-Pz13, the amplitude analysis was employed to the other crosshole planes to define the approximate locations and dimensions of the low-velocity waveguides of both types WGT I and II. Thereby, each plane is separately analyzed and the different types of waveguides are identified. In Figure 3.4, we present the waveguide locations of WGT I marked by red and brown boxes. Blue and light blue boxes represent the approximate locations of WGT II with both high permittivity and high conductivity clay layers. Different colors mean waveguide structures at different depths that might be discontinuous. In general, the locations of WGT I are present for all boreholes between 7.3 m to 9.1 m. Furthermore, indicators of WGT II can be found nearby 4.0 m depth at most of the boreholes. In addition, in the vicinity of 6.0 m, 7.0 m and 9.0 m depths, there are some discontinuous features that could be related to events caused by WGT II. Next, we will use the obtained information to modify the permittivity starting models of the updated FWI, and with this, we will verify the location of the waveguiding structures.

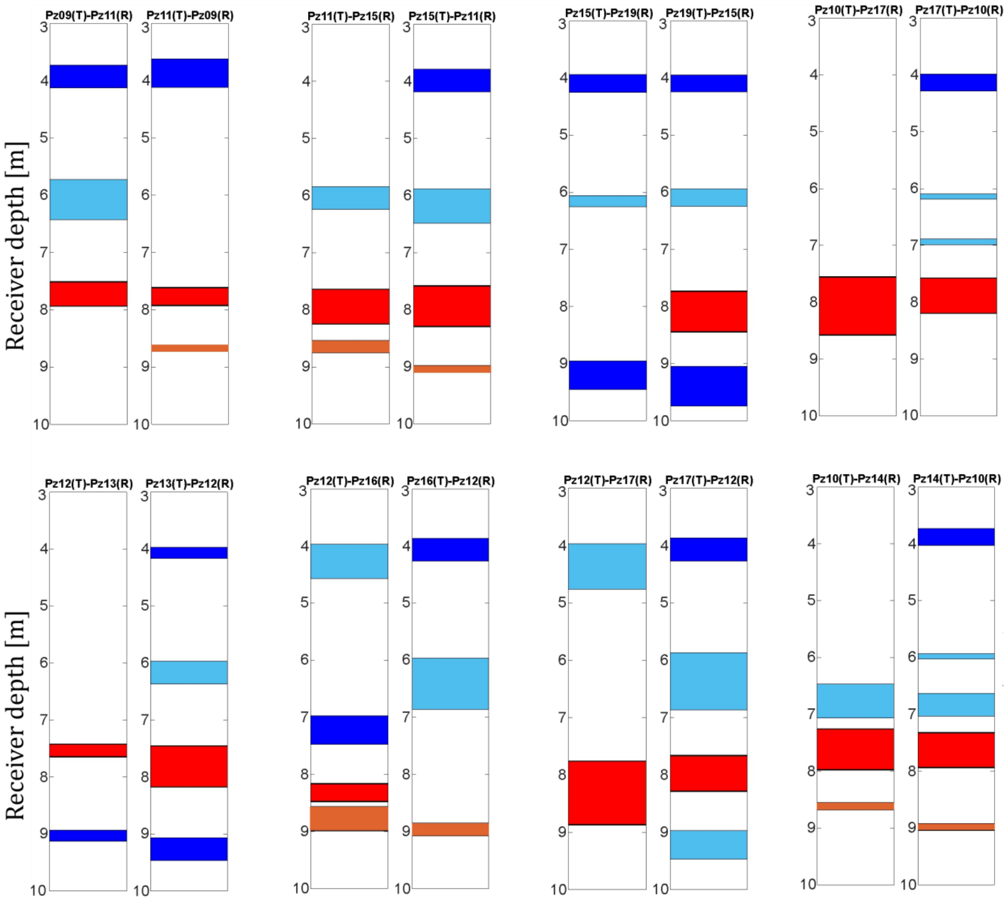


Figure 3.4. Approximate waveguide structures based on the amplitude analysis for the eight cross-sections. Red and brown (difficult picking) indicate possible waveguide locations caused by WGT I, while blue and light blue (difficult picking) indicate possible waveguide locations caused by WGT II. The depths of different color shades indicate possible discontinuous waveguide structures.

.

3.3 FULL-WAVEFORM INVERSION RESULTS

3.3.1 Full-waveform inversion method

We consider the FWI converges when the root mean squared error RMSE of observed and model traces changes less than 0.5% between two subsequent iterations. Furthermore, we investigate the behavior of the remaining gradient values of ε_r and σ. In the framework of this study, we applied a muting zone around the boreholes to avoid the effect of nested wells in the vicinity of GPR boreholes (Klepikova et al., 2016). Note that gradient normalization as proposed by van der Kruk et al. (2015) leads to artifacts close to the boreholes due to the presence of more than one tubes in some of the boreholes. Each of the planes was independently analyzed and inverted with the FWI following the guideline of Klotzsche et al. (2019b).

3.3.2 Short distance boreholes

During this study, we note that the ray-based results of short distance boreholes (approximate 3.0 m) provided ε_r starting models that were not within half a wavelength of the measured data. Such borehole pairs are probably affected by the presences of the 10 cm diameter water-filled boreholes and could create 3D effects. Generally, the permittivity values are overestimated in comparison to neighboring planes with larger offsets (Peterson, 2001). Reducing the angle of the measurements did not improve the results. To solve this problem, we combined the results of the short distance boreholes with close larger distance boreholes, and then compared the mean velocity based on both ZOP data sets. We assume that the aquifer is isotropic and that the mean velocity should not change much between neighboring boreholes. For example, the plane between Pz09-Pz11 has a width of 3.02 m and the close by plane Pz11-Pz15 has an offset of 4.88 m. In Figure 3.5a, the calculated mean velocities (black lines) and relative permittivity (red lines) based on the ZOP data for cross-section Pz09-Pz11 (solid lines) and Pz11-Pz15 (dashed lines) show a clear difference of about 2 - 4 in ε_r. Pz09-Pz11 shows lower mean velocity (higher mean ε_r) along the vertical depth axis. If we assume that in the horizontal direction the medium behaves similarly and therefore the mean velocities should be in the same range, we can speculate that the lower velocity for short distance borehole pairs is caused by the increased proportion of the borehole fillings compared to the whole measured distance, an effect that should be corrected. Considering the boreholes distance between Pz09 and Pz11 is 3.02 m, we calculated a delay time of approximately 3.3 ns ($\Delta\varepsilon_r$= 2.9) for the short distance boreholes ($3.3 = \frac{3.02}{\bar{V}_{Pz09-11}} - \frac{3.02}{\bar{V}_{Pz11-15}}$). Using this delay time, we corrected the picked first arrival times of the MOG data for the short plane and performed a new ray-based inversion. For the unsaturated zone above the water table, we modeled a homogeneous layer with a relative permittivity of ε_r= 4.4 (not shown, same for all

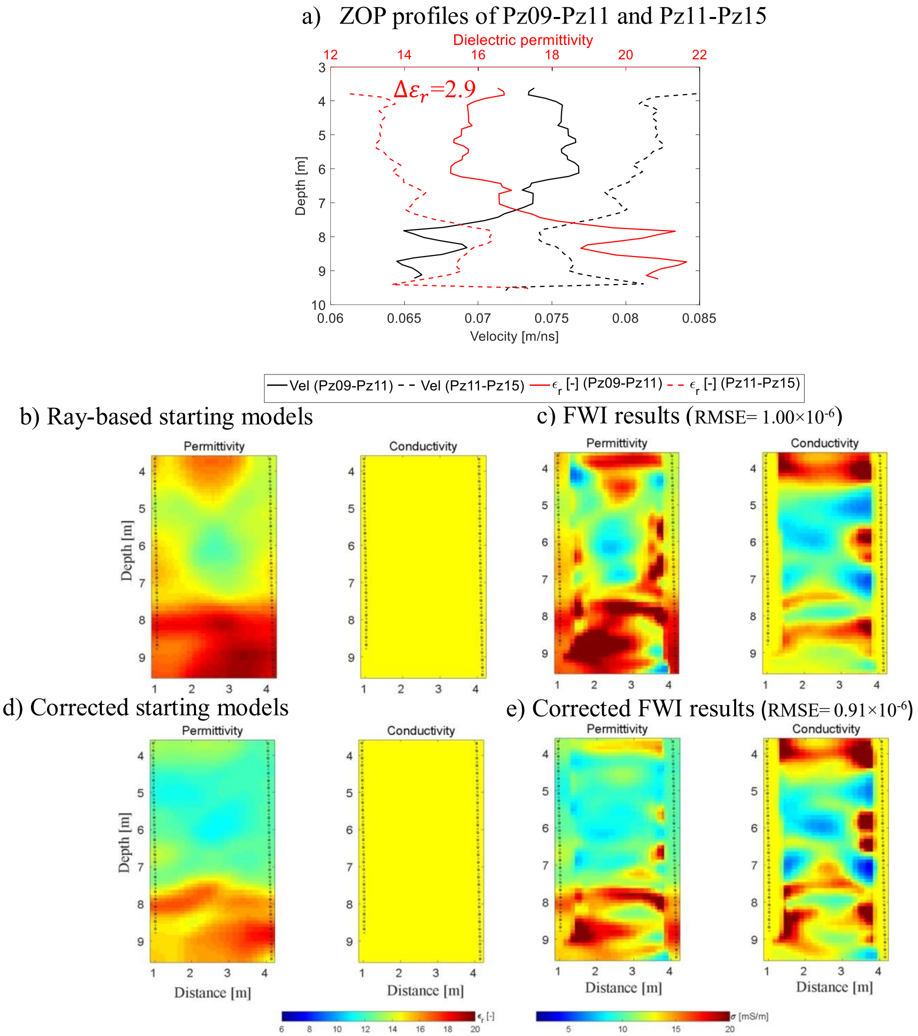


Figure 3.5. a) Velocity (black lines) and ε_r (red lines) comparison over depth based on the ZOP data for boreholes Pz09-Pz11 (solid lines) and Pz11-Pz15 (dashed lines). b) and d) indicate the ray-based ε_r results and the updated starting models based on the ZOP analysis, respectively. Conductivity starting model is homogeneous with 13 mS/m. c) and e) show the corresponding FWI results based on the ray-based starting models and the updated starting models, respectively. The RMSE value is indicating the root-mean-square error between the measured and model GPR data.

following inversions) and for the σ starting model, we selected a homogeneous model of 13 mS/m (mean of the first cycle amplitude inversion results). By comparing the FWI results of both uncorrected and corrected ray-based starting models for the FWI, we note that the corrected FWI results of Pz09-Pz11 show generally lower permittivity results that the uncorrected data (Figure 3.5). This is also in agreement with the ZOP results of the neighboring borehole pair with the larger offset. Although the final RMSE is similar for both inversions, the data fit and the remaining gradient were better for the corrected data.

3.3.3 Starting model test for waveguide zones

From the amplitude analysis we know the approximate locations of possible waveguide structures. Instead of adding homogeneous layers at the expected locations of the waveguide zones as shown by Klotzsche et al. (2012), we investigate another strategy that considers different stages/iterations of the traditional FWI. Details are investigated for the crosshole plane Pz10-Pz13 and will then be applied to the other planes. First, we perform the standard FWI using the ray-based starting models including the water table contrast (Figure 3.6a and b). The obtained FWI results show higher resolution images than the ray-based results, and the FWI modeled and measured data are in a good agreement in phase and amplitude (not shown), and, show only minor differences also indicated by a correlation coefficient of 0.9652 (Table 3.1). Analyzing the results in more detail, we note that the permittivity results with only ten iterations provide clear indications of the waveguide locations, which were consistent with the amplitude analysis results. Therefore, we used the permittivity result of the tenth iteration as a new starting model (Figure 3.6c) and updated the effective source wavelet (the red one of intersect in Figure 3.6c). Note that standard applications of the crosshole GPR FWI normally consider the ray-based results also with updated effective source wavelets. Here, we apply results of a previous FWI iteration as an updated permittivity starting model and a corresponding updated effective source wavelet. The updated FWI results based on this new effective source wavelet and updated starting model show clearer structures with more details and provide a lower RMSE in the final iteration (RMSE=0.90×10^{-6} instead of RMSE=1.11×10^{-6} for the traditional FWI). The green boxes in Figure 3.6d indicate the waveguide structures that are probably caused by high porosity waveguides (WGT I), while the blue boxes indicate waveguide zones that could be caused by a higher clay content yielding a higher electrical conductivity (WGT II). Generally, we see a good correlation between the waveguide zones identified with the amplitude analysis and the final FWI inversion results. Similar to previous studies (e.g., Klotzsche et al., 2014), the WGT I structure below 7.0 m depth shows a higher permittivity than the surrounding media indicating an increase in porosity. This is consistent with the heat tracer was flowing preferentially at the bottom of the aquifer, which has a larger hydraulic conductivity. Furthermore, the FWI confirmed the hypothesis that the WGT II structures are caused by an increase of electrical conductivity that could be caused by an increase in clay content. To further confirm the updated FWI results, we present image plots of the FWI modeled data for boreholes Pz10-Pz13 in Figure 3.7, which can be compared to the measured data in Figure 3.2. The modeled data show a good fit to the measured data in the entire aquifer domain indicating that the FWI solution explains the measured data well. The features that were identified and marked as possible WGT I (red and green circles) and WGT II (dark and light blue) features in the measured data correspond to the features observed in the FWI result.

An erroneous permittivity starting model also affects the high electrical conductivity zone and vice versa. If the permittivity estimate is inaccurate and the modeled data cannot fit the shape of the traces, the conductivity model tries to compensate for this and yields erroneous structures. Therefore, it is highly important to confirm that the inverted permittivity results are reliable. Numerous studies show that it is necessary that first the permittivity (shape of traces) are updated before the conductivity (amplitude of the traces) can be optimized in more detail (Klotzsche et al., 2019a). In previous studies, higher homogenous permittivity layers were added into the permittivity starting model (Klotzsche et al., 2014) to ensure that the starting models fulfill the half-wavelength criteria.

To verify the higher permittivity and conductivity zone near 4.0 m close to Pz10, we performed additional starting model tests. In these tests, we investigate the possibility that the high electrical conductivity zone is caused by an erroneous permittivity starting model and the possibility that the waveguide is caused by an increased permittivity instead of conductivity. Note that due to a reduced ray coverage in the upper part of the investigation domain, it is possible that the permittivity starting model based on the ray-based approach does not fulfil the half-wavelength criterion. In such a case, it can happen that the electrical conductivity results compensate for an erroneous permittivity model. As proposed by Klotzsche et al. (2014), we added different scenarios of a higher homogeneous ε_r layer (ε_r=16) in the ray-based ε_r model (Figures 3.8a-d) and repeated the FWI analysis. We kept the effective source wavelet and the conductivity starting model unchanged. All results of the different starting model tests confirmed the previously obtained structure, but some discontinuity in the inversion results is visible indicating difficulties of the inversion to find the global minimum. Therefore, we can assume that our FWI results based on the updated permittivity starting model (Figure 3.6c) without the homogeneous layers produced reliable results and that the FWI indicates that the minimum locations in the amplitude analysis approach are caused by an increased electrical conductivity (Figure 3.8, WGT II).

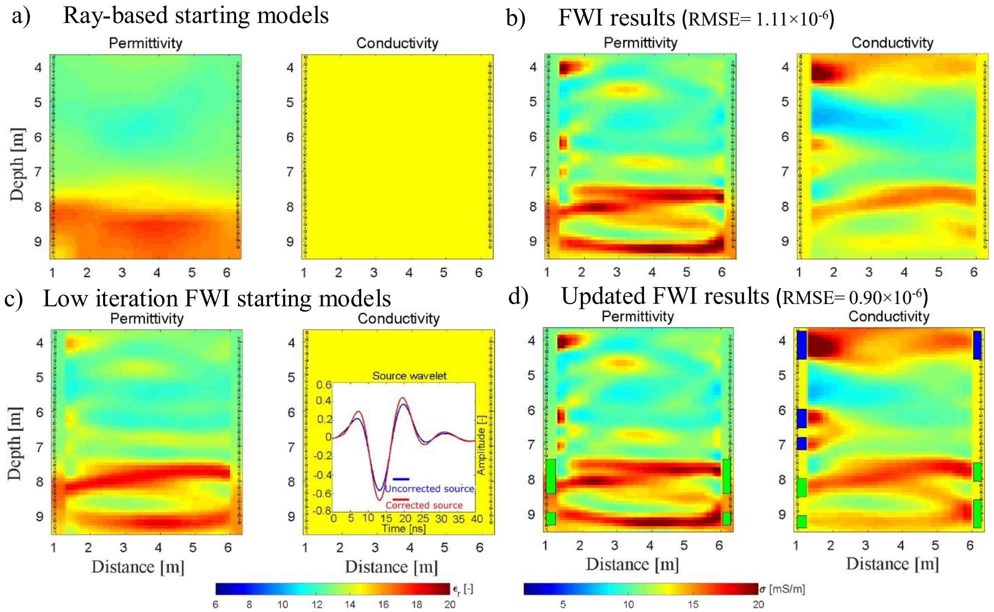


Figure 3.6. a) Ray-based starting models and b) corresponding FWI results for crosssection Pz10-Pz13. c) Updated starting models based on a lower iteration number results of b) and d) correponding FWI results. Inset in Figure 3.6 c) shows the estimated effective source wavelets based on different starting models. The green and blue boxes in d) indicate the locations of waveguides of WGT I and II, respectively

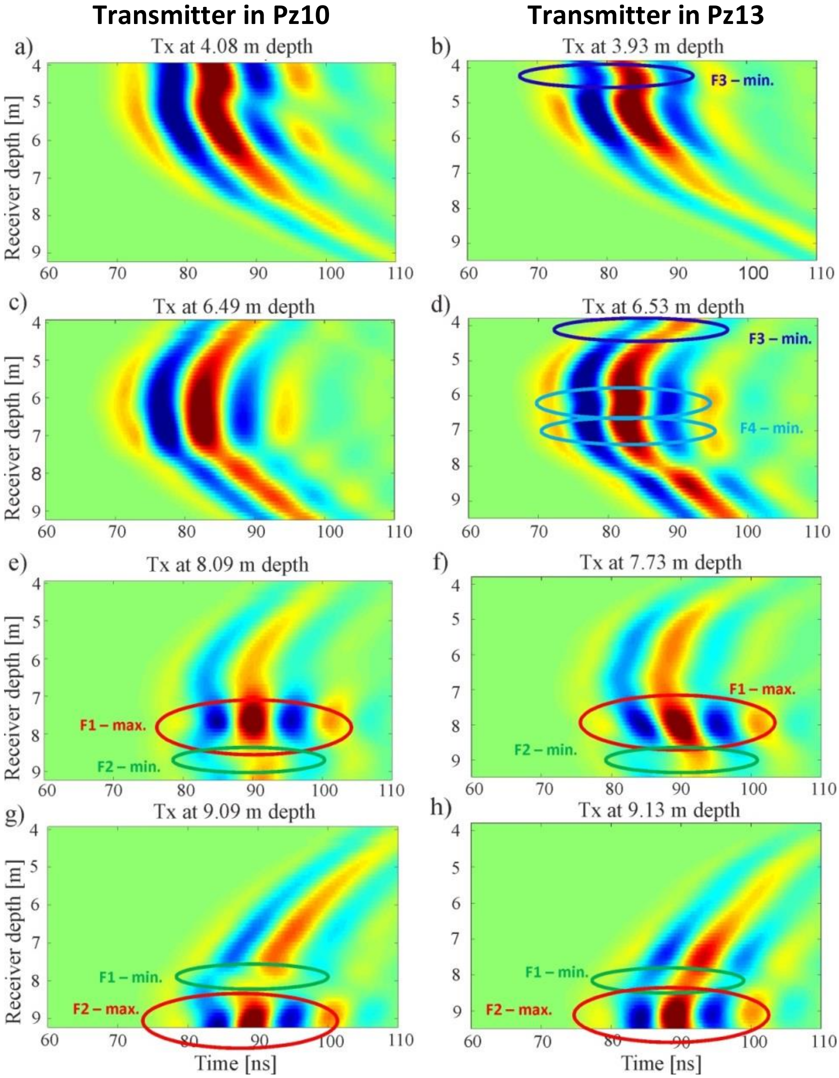


Figure 3.7. Image plots of the modeled data based on updated FWI results for the cross-section Pz10-Pz13 (Figure 3.6d). Different waveguide features are indicated with F1-F4, while WGT I features are marked with red and green and WGT II features with blue and light blue, respectively. Note that we have normalized the amplitudes values to the maximum value of amplitudes for the cross-section (range from -7×10^{-1} to 7×10^{-1}).

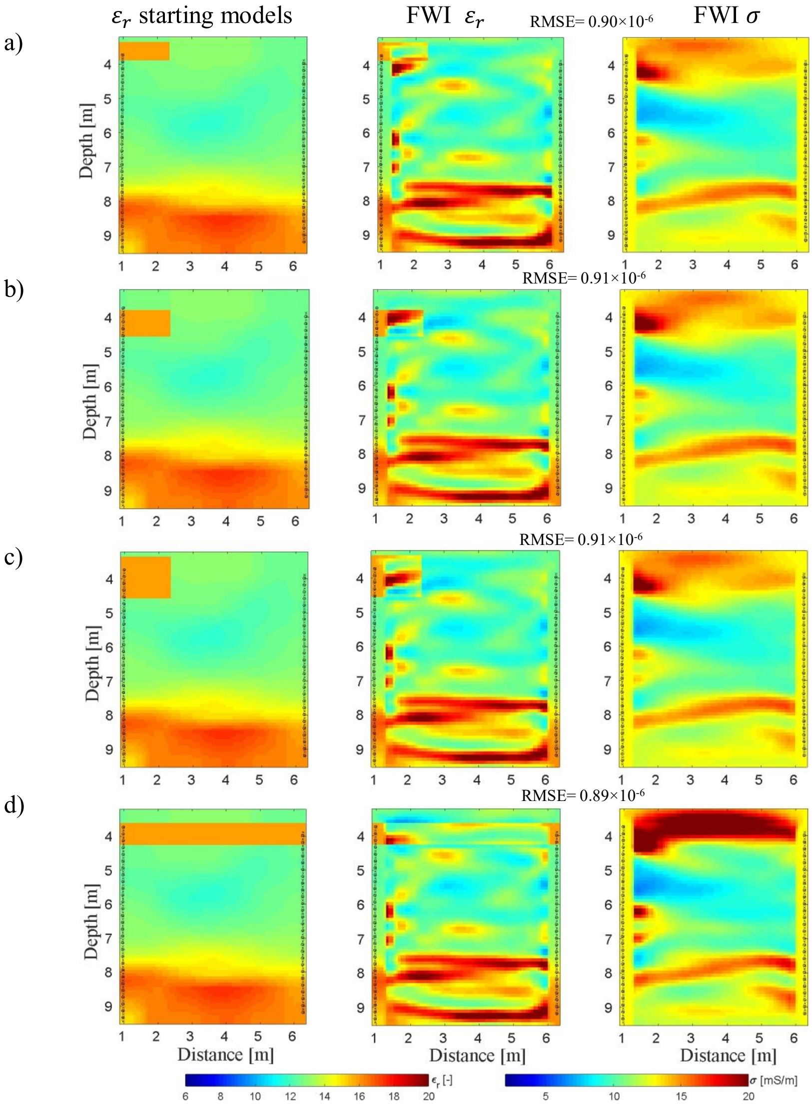


Figure 3.8. Starting model tests for ε_r are to determine the FWI σ reliability. The first column shows the ray-based starting models of ε_r with added a high ε_r value zones (ε_r=16). Conductivity starting models are homogenous with 13 mS/m (not shown). The second and third columns show the corresponding FWI ε_r and σ results, respectively. Note the effective source wavelet in Figure 3.6c (red) is used to perform the FWIs.

3.3.4 Combined FWI results of the Hermalle-sous-Argenteau

For the other crosshole planes, we followed the same approach and updated the effective source wavelets based on the iteration ten FWI ε_r field. The effective source wavelet of a cross-section is depending on the borehole fillings and couplings, and anything is not included in the forward model (e.g., finite length antennae). The obtained updated effective source wavelets are also depending on the borehole distance as shown in Figure 3.9. The crosshole pair with the largest distance shows the highest amplitude in time and frequency domain. Furthermore, we observe a shift in center frequency from a higher center frequency for small distance pairs towards a lower center frequency for larger offset combinations (Figure 3.9b). Note that boreholes Pz09-Pz11 and Pz15-Pz19 show the largest ($f_c = 80$ MHz) and the smallest ($f_c = 57$ MHz) center frequency, respectively. Except for the effect in amplitude values, similar shapes are observed for all cross-sections using the 200 MHz antennae. The effective center frequency is significantly lower than the nominal center frequency of the antennae in air (200 MHz). This is caused by the fact that the antennae are electrically longer in high-permittivity media and emit lower frequencies than in air (e.g., Klotzsche et al., 2013). Therefore, considering that the signals travel longer in the subsurface for larger offset datasets, the center frequency of effective source wavelet decreases with increasing distance.

The ray-based permittivity results and the final FWI results of the nine crosshole planes combined together to generate a 3D image of the aquifer (Figure 3.10). The ray-based results show generally a three-layer model with lower and intermediate permittivity values above 7.5 m depth and higher permittivity values between 7.5 m to 9.5 m depth. In contrast, the FWI results show higher resolution images for both permittivity and conductivity and more structures can be observed. Although the inversions are performed independently, consistent structures at the borehole locations are observed. For simplification, the locations of waveguides with high porosity from amplitude analysis are marked at the cross-sections of Pz09-Pz11, Pz11-Pz15 and Pz15-Pz19 using green boxes (WGT I). Additionally, the waveguides caused by WGT II are indicated with blue boxes in the conductivity image. In the FWI ε_r results, the higher permittivity values located from 7.5 m to 9.5 m comprise waveguide structures of WGT I and a higher continuous conductivity zone is clearly shown in the FWI σ results around 4.0 m depth.

For all inversion planes a good fit between the measured and modeled data was found (not shown). Table 3.1 shows a comprehensive comparison between the traditional and the updated FWI for all planes. The final RMSE, the absolute mean gradient (AMG) values for permittivity and conductivity, and the correlation coefficients (R) between observed and modeled data indicate that the updated FWI results are better than the traditional FWI results.

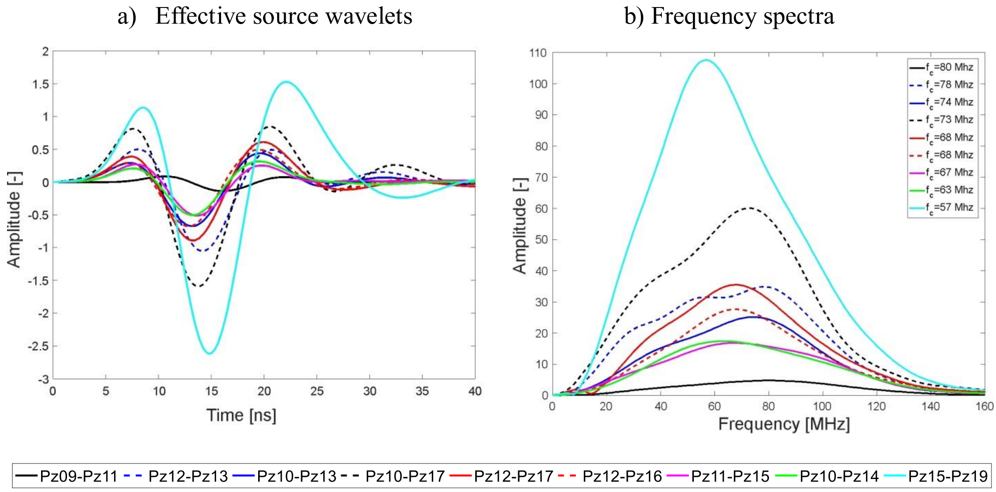


Figure 3.9. a) Estimated effective source wavelets for the nine cross-sections in time domain using the updated low iteration FWI results as starting models. b) Corresponding frequency spectra of the nine different effective source wavelets. Each color represents a different plane. The legend values in b) indicate these effective center frequencies for different planes.

The updated FWI results show a good consistency at the different borehole cross-sections as indicated by the R at the intersections of 15 cross-sections (Table 3.2). For all cross-sections, we used the mean of two inversion cells to compute R values (Klotzsche et al., 2013). For the planes that have a separation of approximately 5.0 m the R for ε_r shows values from 0.52 to 0.89. Note that the cross-sections Pz09-Pz11 (3.02 m) and Pz15-Pz19 (7.07 m) are different from others. Further, a lower R of 0.26 for the permittivity intersection between Pz09-Pz11 and Pz11-Pz15 is calculated, which shows the FWI results are not very consistent at the borehole Pz11. Similarly, Pz15 shows a weak correlation of the intersections. The R for ε_r between Pz10-Pz13 and Pz10-Pz17 is 0.24, and R for ε_r between Pz10-Pz14 and Pz10-Pz17 is 0.35. The reason for this weak correlation could be that the plane between the borehole Pz10 and Pz17 was measured two months later than the other crosshole sections and seasonal fluctuations could have caused this effect in the permittivity. Comparing these values with the R (ε_r) values derived from the crossing point of non-borehole (Pz10-Pz17) zones, which are 0.69, 0.89 and 0.90, we believe the later measured GPR cross-section is affected by the vicinity of borehole Pz10. Note, that there is a second well very close by to Pz10, which could affect the results of the planes connected to this borehole. Finally, although a lower mean R (ε_r) was achieved for some locations, for most cases an acceptable value above 0.61 for permittivity and 0.85 for conductivity indicate consistent and reliable results.

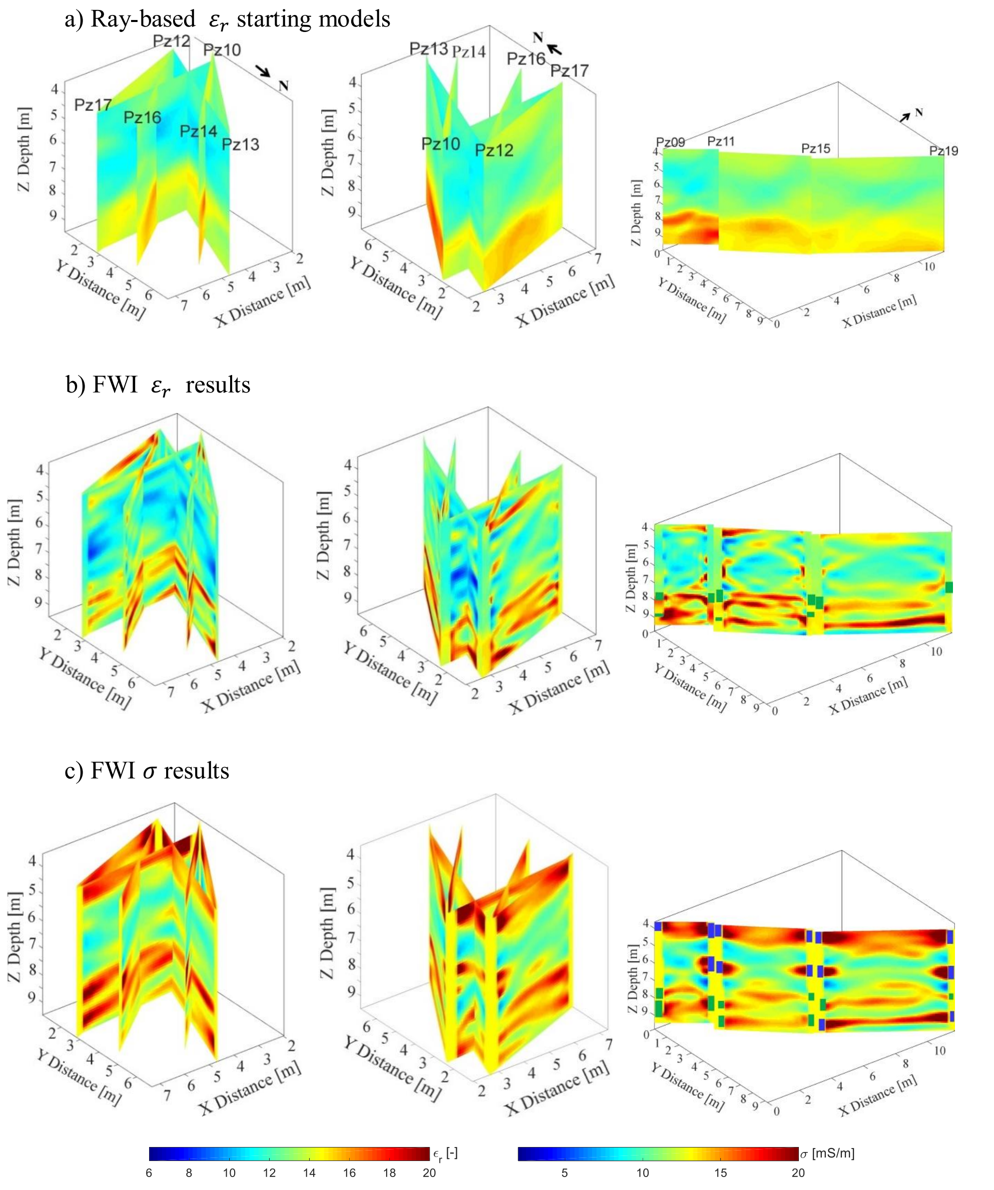


Figure 3.10. a) Ray-based results of ε_r for all crossections for different viewing angles. b) and c) FWI ε_r and σ results, respectively, using the traditional low (10) iteration FWI results as starting models and a homogenous model with 13 mS/m for the σ (not shown). Green and blue boxes next to the cross-sections image boundaries along Pz09 to Pz19 indicate the boundaries of the waveguide structures of WGT I and II obtained from the amplitude analysis, respectively.

Table 3.1. A comparison of FWI results between the traditional FWI and the updated FWI. Percentages in parentheses indicate the ratio of the updated FWI final RMSE to the traditional FWI RMSE.

Different boreholes	**Pz10-Pz13**	**Pz11-Pz15**	**Pz12-Pz17**	**Pz15-Pz19**	**Pz09-Pz11**	**Pz10-Pz14**	**Pz12-Pz13**	**Pz10-Pz17**	**Pz12-Pz16**
Traditional FWI									
Iterations	26	30	24	30	27	27	30	30	30
Mean abs. Gradient (ε_r) (10^5)	2.08	0.88	1.52	1.77	0.94	0.90	0.72	1.82	0.54
Mean abs. Gradient (σ) (10^{-4})	1.58	1.21	2.38	1.95	1.55	2.02	0.99	2.12	1.38
***R* for observed and modeled data**	0.9652	0.9618	0.9774	0.9478	0.9761	0.9669	0.9702	0.9726	0.9677
RMSE (10^{-6})	1.11	1.06	1.23	1.47	0.91	1.02	0.90	1.45	1.17
Updated FWI									
Iterations	29	26	30	27	30	30	30	29	28
Mean abs. Gradient (ε_r) (10^5)	0.28	0.31	0.55	0.66	0.87	1.08	0.59	1.21	0.52
Mean abs. Gradient (σ) (10^{-4})	0.88	1.15	1.46	1.22	0.96	1.45	0.80	1.87	1.25
***R* for observed and modeled data**	0.9768	0.9660	0.9798	0.9593	0.9763	0.9743	0.9729	0.9749	0.9695
RMSE (10^{-6})	0.90	0.99	1.16	1.28	0.91	0.89	0.86	1.39	1.13
	(81.1%)	(93.4%)	(94.3%)	(87.1%)	(100%)	(87.3%)	(95.6%)	(95.9%)	(96.6%)

Table 3.2. Correlations coefficient (R) for the updated full-waveform inversion permittivity and conductivity results at the ten boreholes. The mean values of R for the permittivity and conductivity planes were 0.61 and 0.85, respectively.

Borehole (Pz)		10		11		12		13	15	17	The crossing of diagonal (intersecting) planes				
Planes	10-13	10-13	10-14	09-11	12-16	12-16	12-13	10-13	11-15	10-17*	11-15	12-16	11-15	10-17*	10-14
	10-14	10-17*	10-17*	11-15	12-13	12-17	12-17	12-13	15-19	12-17	10-17*	10-17*	12-13	12-13	12-13
***R* (ε_r)**	0.52	0.24	0.35	0.26	0.67	0.89	0.75	0.74	0.38	0.67	0.69	0.89	0.58	0.90	0.62
***R* (σ)**	0.79	0.93	0.75	0.70	0.88	0.89	0.90	0.92	0.87	0.91	0.85	0.78	0.81	0.88	0.83

Pz10-Pz17* was measured later than other boreholes and close by to Pz10 is a second well installed which is water filled.

3.4 PETROPHYSICAL INTERPRETATION

To improve the understanding and to explain the heat tracer experiment of Hermans et al. (2015a), which described quantitative temperature monitoring at the Hermalle-sous-Argenteau field site, we analyzed the porosity of the saturated zone based on the new GPR FWI ε_r results using the complex refractive index model (CRIM) similar to Gueting et al. (2015):

$$\emptyset = \frac{\sqrt{\varepsilon_r} - \sqrt{\varepsilon_s}}{\sqrt{\varepsilon_f} - \sqrt{\varepsilon_s}}, \tag{3.1}$$

for the fluid permittivity ε_f we considered 84 for a water temperature of 10 ℃ and for the solid permittivity ε_s we used 4.5 based on literature values of quartz (e.g., Birchak et al., 1974; Eisenberg and Kauzmann, 2005; Carmichael, 2017). To evaluate the porosity values between Pz13 and Pz17 (ERT cross-section in Figure 3.1d), we considered five cross-section porosity results in Figure 3.11a (3D) and Figure 3.11b (2D). Higher permittivity zones of the FWI results (Figure 3.10) result in higher porosity values (Figure 3.11). Porosity values reach 30-35% in the high porosity zones between 7.0 m to 9.0 m depth. The higher porosity at these depths is consistent with the classification from ERT and borehole log results (Hermans, 2014; Hermans et al., 2015a, b, 2017) with coarse gravel at the bottom and gravel in a sandy matrix on the top. The gravel at the bottom is very coarse (pebbles) with almost no matrix, and is referred to as the swimming pool due to its large hydraulic conductivity. In addition, the coarse gravel has a lower resistivity compared to the sandy gravel due to a larger water content. The high porosity layers are more continuous between the boreholes P10-Pz13 and Pz10-Pz14, while for P12-Pz16 and Pz12-Pz17 more discontinuous structures can be observed, which is consistent with the splitting of the thermal plume (Figure 3.1d). By analyzing the mean porosity close to the boreholes crossing the intersection of Pz13 and Pz17 between 7.0 m and 9.5 m depth (dashed rectangles in Figure 3.11b), we see that the mean porosity values in the selected 2D zones become smaller from Pz13 towards Pz17 (Figure 3.11c). Such a behavior of the porosity structures could explain the heat tracer distribution observed by Hermans et al. (2015a). Zones with a higher porosity and continuous structures such as observed close to Pz13 and Pz14 could also cause preferential flow paths for heat when hot water was injected from borehole Pz09.

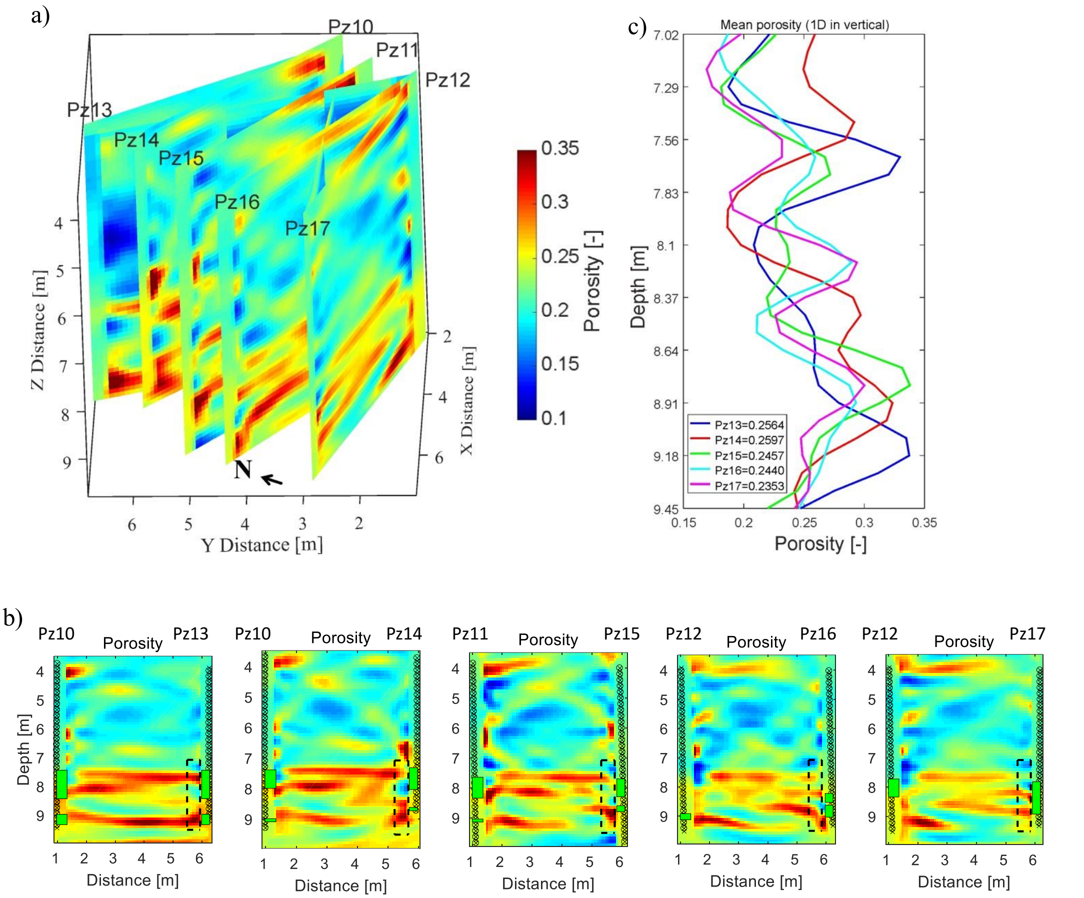


Figure 3.11. a) Estimated porosity distributions in 3D based on CRIM to related to the ERT measured profile Pz13-Pz17 (Figure 3.1). b) Detailed porosity distribution of the five relevant crosshole sections. The green boxes close to antennae indicate the boundaries of the wave-guiding structures WGT I obtained from the amplitude analysis. Dashed black rectangles indicate the zones close to boreholes to compute the mean porosity profiles in c) along the vertical direction.

3.5 CONCLUSIONS AND OUTLOOK

We applied an extended amplitude analysis approach and the crosshole GPR FWI to the Hermalle-sous-Argenteau test site located in the alluvial aquifer of the Meuse River in Belgium. The results of the GPR FWI provided high-resolution images of the aquifer within decimeter-scale resolution and allowed a detailed characterization of the porosity structures. The amplitude analysis, which can be used to identify low-velocity guided wave structures, is extended to detect two waveguide types (WGT) with different origins. To further improve traditional FWI results, the estimation of different starting models for permittivity was proposed. First, for short distance cross-sections, zero-offset profile data were used to correct the travel-time inversion results and to update the starting model of the permittivity. Second, the FWI results were updated by using low iteration number traditional FWI results as starting models and performing new effective source wavelet corrections. Comparisons of the RMSE, the mean remaining gradients for ε_r and σ, and correlation R of the measured and the final FWI data, indicated an improvement of the FWI results using these low iteration starting models in contrast to the traditional results.

By using the amplitude analysis, the approximate locations of WGT I caused by higher permittivity (porosity) and low/intermediate electrical conductivity that produced elongated wave trains in the measured GPR data were detected and confirmed by FWI. For the first time, WGT II structures that are caused by an increase in permittivity and electrical conductivity were identified in the measured data using the amplitude analysis approach. Such waveguides are difficult to detect from the measured data and the amplitude analysis due to the absences of the elongated wave trains (higher wave attenuation in the waveguide zone). The FWI inversion results of the WGT II zones confirmed the existence of these structures probably caused by an increase of clay content.

For the 3D characterization of the alluvial aquifers, nine 2D intersecting GPR planes were separately inverted. We evaluated the updated 2D FWI results by computing correlation coefficient (R) for consistent structures, where acquisition planes intersect. In most of the cases, the FWI results in combination with the amplitude analysis presented consistent structures between the intersections for both permittivity and conductivity images. From the 3D FWI image, high-permittivity layers between 7.5 m to 9.5 m depth for all cross-sections were observed (WGT I). Further, continuous higher conductivity zones above 4.0 m are detected in all cross-sections that aligned with the WGT II features. Compared with ERT tomography results, the updated GPR FWI results confirmed the presence of a high hydraulic conductivity zone with high porosity, but provided higher resolution images of the subsurface and much more details about the lateral heterogeneity. This refined description of the heterogeneity allows a better explanation of the spatial distribution (plume splitting) of the heat tracer test performed on the experimental site. Further research will investigate the potential to use the FWI results to construct a full 3D aquifer model including small-scale heterogeneity and to discriminate geological scenarios as a basis for starting models. In addition, the approach of source-independent time-domain waveform inversion of cross-hole GPR data will be considered in future studies.

Chapter 4

Improvement of GPR Full-waveform inversion images using Cone Penetration Test data[1]

In this chapter, a new approach that uses CPT data to enhance the final FWI relative permittivity resolution via updating an effective source wavelet applied in the inversion process *has been presented*. The updated source wavelet possesses a priori CPT information and a larger bandwidth. Using the same starting models, a synthetic model comparison between the conventional FWI and the updated FWI results demonstrate that the updated FWI method provides reliable and more consistent structures. For field-observed GPR data of the same test site, five GPR cross-sections results were analyzed. Both synthetic and experimental results indicate the potential of improving the reconstruction of subsurface aquifer structure by combining conventional 2D FWI results and 1D CPT data.

[1]adapted from Zhou, Z., A. Klotzsche, J. Schmäck, H. Vereecken, and J. van der Kruk, 2020b. Improvement of GPR Full-waveform inversion images using Cone Penetration Test data: Geophysics, under review.

4.1 UPDATING WAVELET USING AMPLIFIED FWI

Generally, to avoid the overfitting of between the observed and the modeled data, the inversion of experimental data is stopped if the change of the misfit function value is less than 0.5% between two subsequent iterations. In this chapter, we considered the remaining mean gradient values of permittivity and the root mean squared error (RMSE) values between observed and modeled radar traces to judge the different inversion levels. The FWI results with optimal iteration should include the smallest normalized ε_r gradient and RMSE values. Considering that the gradients are highly sensitive close to the transmitter and receiver positions, inversion artifacts can easily arise close to the boreholes. To minimize these inversion artifacts the approach of Kurzmann et al. (2013) is applied using a gradient preconditioning (van der Kruk et al., 2015).

4.1.1 Wavenumber filter

The CPT data presents a high spatial resolution along the vertical 1D profile. Transforming the spatial CPT data to amplitude-wavenumber domain by a fast Fourier transformation (FFT), we can obtain a broader bandwidth of the CPT data than the FWI models bandwidth (e.g., Yang et al., 2013). Therefore, it is feasible to improve the FWI resolution by expanding the bandwidth of the FWI amplitude values using the CPT data in the amplitude-wavenumber domain. In the first step, we convert the relative dielectric permittivity of the FWI results into porosity $\emptyset$ by using the complex refractive index model (CRIM) for the saturated zone (e.g., Birchak et al., 1974) (Equation 3.1).

In the second step, the selected 1D vertical porosity-model of the CPT data (located close by or at the GPR cross-section) and FWI permittivity results were interpolated to receive enough data points to generate the filter in wavenumber domain. Here, we used a mean value of the selected data and two cosine functions (represented as tapers) to expand the initial data. The final expanded data (data points = 512) with tapers are transferred into wavenumber domain using the 1D FFT. Note that the two endpoints of the resampled data should satisfy periodicity to avoid trapping into Gibbs jumps (e.g., Li et al., 2018). The resampled process can be expressed by

$$\begin{cases} Exp_{\emptyset} = p \times mean\,(Ori_{\emptyset}), & (x < Lidx - Tlen \;or\; x > Ridx + Tlen) \\ Exp_{\emptyset} = C_1 \times \left(1 - \cos\left(\frac{\pi}{Tlen} \times \left(x - (Lidx - 1 - Tlen)\right)\right)\right) + p \times mean\,(Ori_{\emptyset}), & (Lidx - Tlen \leq x < Lidx) \\ Exp_{\emptyset} = C_2 \times \left(1 - \cos\left(\frac{\pi}{Tlen} \times \left((Ridx + 1 + Tlen) - x\right)\right)\right) + p \times mean\,(Ori_{\emptyset}), & (Ridx < x \leq Ridx + Tlen) \end{cases} \tag{4.1}$$

where $Ori_{\emptyset}$ and $Exp_{\emptyset}$ represent the original data and the final expanded data points, respectively. p is 0.8, which is always used in this paper. $mean\ (Ori_{\emptyset})$ shows a mean value of original data. x represents point position in the final expanded data. $Lidx$ and $Ridx$ represent the left and right positions of original data in the expanded data. $Tlen$ shows the taper length that is 0.3 times of the original data length. C_1 and C_2 are chosen parameters, which are used to smooth connection points between the taper and the original data two points.

In the third step, a smooth function is applied to flatten the highly fluctuating amplitudes of both interpolated data set, which are caused by the interference of the real and imaginary parts for the data in wavenumber domain as explained by Yang et al. (2013). These smooth results are estimated by

$$SA(x) = \begin{cases} A(x) & ,x = 1 \\ smooth(A(x), span) & ,x > 1 \end{cases} \tag{4.2}$$

where A and SA represent amplitude and smooth amplitude values in the wavenumber domain, respectively, and, x indicates the wavenumber sample up to the selected maximum wavenumber threshold. For the smoothing of amplitude values in amplitude wavenumber domain, we applied standard function $smooth$ of MATLAB (MathWorks, Inc. 2016b). For this function a span value needs to be defined that is used in the smooth function. Note that the smooth function starts with the second sample (Equation 4.2) because of unusual zero-frequency values and ends with an appropriate maximum threshold wavenumber value. In general, the selected maximum threshold value is determined using an empirical rule that keeps the generated filter to be approximate monotonically increasing or to be fluctuated around the amplitude value with 1 (Zhou et al., 2019). Finally, a filter is designed with a ratio factor that was calculated in the wavenumber domain to amplify the FWI results in the whole 2D domain. This filter was implemented with:

$$Filter = \frac{SA_{CPT}}{SA_{FWI}}\ , \tag{4.3}$$

where SA_{CPT} and SA_{FWI} represent the smooth CPT data and the smooth FWI results from one to the maximum threshold in wavenumber domain, respectively. After this filter has been calculated, it is multiplied with the conventional FWI results in the wavenumber domain to generate the 2D wavenumber amplified FWI (WA-FWI) permittivity results, which are back-transformed into the spatial domain using an Inverse Fast Fourier transformation (IFFT).

Since the real emitted source wavelet of experimental GPR data cannot directly be obtained, it is important to estimate an effective source wavelet for the FWI. Different from the traditional deconvolution approach that uses the ray-based starting models or later iterations of the FWI results, we employed the 2D WA-FWI ε_r results to replace the ray-based ε_r model. Therefore, similar to the standard procedure, synthetic data is generated with forward modeling using the standard effective source wavelet and the WA-FWI ε_r model (σ model is the same). Using the Greens function (synthetic data divided by the conventional source wavelet in frequency domain) and the observed data, an updated effective source wavelet SW_{WA-FWI} can be obtained that contains the high wavenumber information. After obtaining the updated source wavelet SW_{WA-FWI}, an updated FWI was performed using the same starting models of the conventional FWI. Generally, a second-updated source wavelet is necessary that can be computed based on the deconvolution method, when the previous updated source wavelet replaces the standard source wavelet and the new FWI ε_r results replaces the WA-FWI ε_r models. The updated processing including generating the filter, updating the effective wavelet and performing the updated FWI, is summarized in a workflow diagram (Figure 4.1).

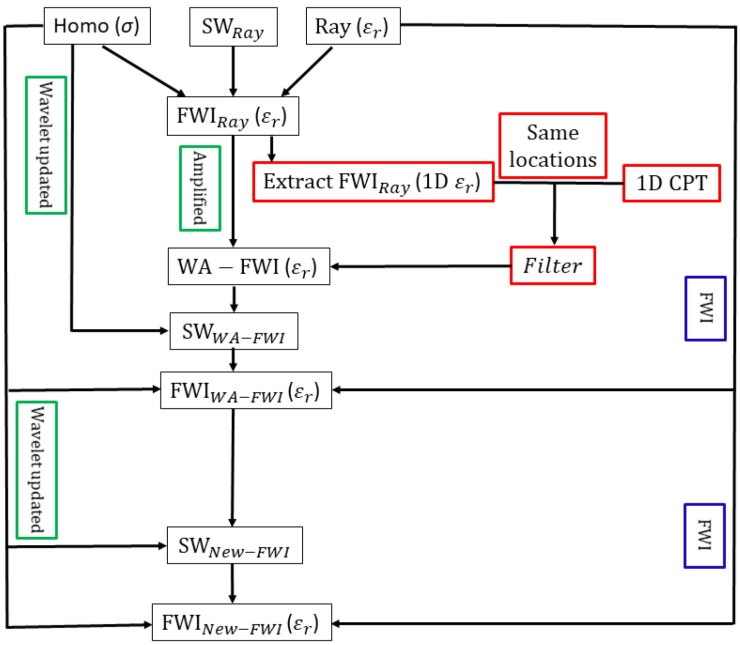


Figure 4.1. Workflow chart of the updating strategy of the effective source wavelet based on WA-FWI results and of the performance of the new FWI. The red boxes represent data in wavenumber domain.

4.2 SYNTHETIC CASE STUDIES

4.2.1 Stochastic aquifer models

To verify the approach of improving the resolution of GPR FWI results using the CPT data, a hydrological model based on experimental hydrological and geophysical data of the well-known Krauthausen test site (Figure 4.2) was used to derive GPR data (Haruzi et al., 2018). We constructed realistic synthetic models of relative dielectric permittivity and electrical conductivity using a stochastic simulation (Sequential Gaussian Simulation; e.g., Bortoli et al., 1993). For the simulation, the aquifer facies model was divided into 3 facies based on Tillmann et al. (2008): sand, sandy gravel and gravel (Figure 4.2a). The simulation of each facies was performed separately. The mean and variance values for permittivity and conductivity were calculated from the traditional GPR FWI results of the Krauthausen test site. Correlation lengths of both ε_r and σ are the same and are adapted from hydraulic conductivity values estimated from high spatial resolution CPT analysis (Tillman et al., 2008). The input parameters (mean, variance, horizontal and vertical correlation lengths) for the variogram model are summarized in Table 4.1.

Before computation of the forward synthetic modeling result, the boundaries of the stochastic models needed to be enlarged to use the same borehole geometries as experimental GPR boreholes (B38-31 in Figure 4.2d) and to avoid interactions with the inversion domain boundaries. Here, we employed a uniform value, which is close to the boundaries within the stochastic models (shadowed areas in Figure 4.3a). For the unsaturated zone above the water table, we chose a homogenous layer with a relative permittivity of ε_r= 4.4 (not shown, same for all following inversions). A semi-reciprocal acquisition setup was used for the models with transmitter TRN and receiver REC spacing of 0.5 m and 0.1 m, respectively. Black circles (TRN=27) and crosses (REC=129) show the exact transmitter and receiver positions within the boreholes. The effective source wavelet used to generate synthetic data is similar to the effective source wavelet of previous measurements performed in the borehole pair 38-31 of the Krauthausen test site (Figure 4.2d, Gueting et al., 2015). Realistic synthetic GPR trace data (called observed data) hereafter without noise based on the stochastic models were generated using 2D finite-difference time-domain (FDTD) modeling. The vertical dashed line (Figure 4.3a) indicates the selected locations of stochastic CPT (Sto-CPT) data that was used to calculate the wavenumber filter.

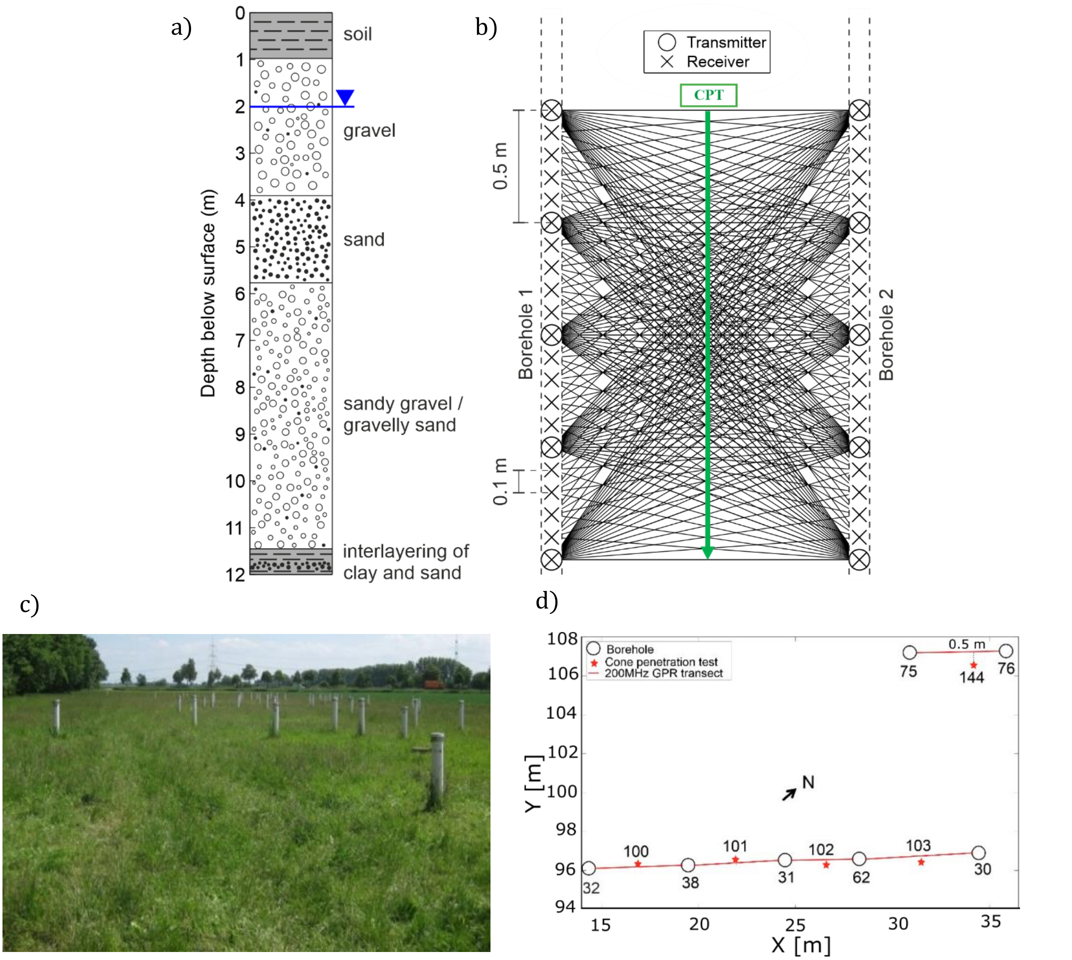


Figure 4.2. a) Generalized cross-section of the uppermost aquifer based on Tillmann et al. (2008). b) Schematic sketch of the crosshole GPR acquisition setup, in which the green arrow indicates the location of CPT data. c) Picture of the Krauthausen test site and d) location of boreholes (circles) and cone penetration tests (asterisk), in which the distance from the CPT 144 to the corresponding cross-section is about 0.5 m. Figure 4.2 a) and b) are adapted from Gueting et al. (2015).

Table 4.1. Parameters for stochastic simulation of permittivity and conductivity based on data of the Krauthausen test site (Tillman et al., 2008). Parameters ε and σ are mean values for different facies. s_ε^2 and s_σ^2 represent variance values for permittivity and conductivity, respectively. Parameters $\lambda_{\varepsilon,h}$ and $\lambda_{\varepsilon,v}$ are the horizontal and vertical correlation lengths fitted with an exponential model for permittivity. And the horizontal and vertical correlation lengths of conductivity are shown by $\lambda_{\sigma,h}$ and $\lambda_{\sigma,v}$, respectively.

		Sand (1)	**Sandy gravel (2)**	**Gravel (3)**
Permittivity	ε	21.52	17.82	13.89
	s_ε^2	9.83	8.71	8.68
	$\lambda_{\varepsilon,h}[m]$	5	1.75	0.3
	$\lambda_{\varepsilon,v}[m]$	0.19	0.2	0.41
Electrical conductivity	$\sigma\left[\frac{mS}{m}\right]$	15	10.4	9.6
	$s_\sigma^2\left[\left(\frac{mS}{m}\right)^2\right]$	4.32	17.68	4.48
	$\lambda_{\sigma,h}[m]$	5	1.75	0.3
	$\lambda_{\sigma,v}[m]$	0.19	0.2	0.41

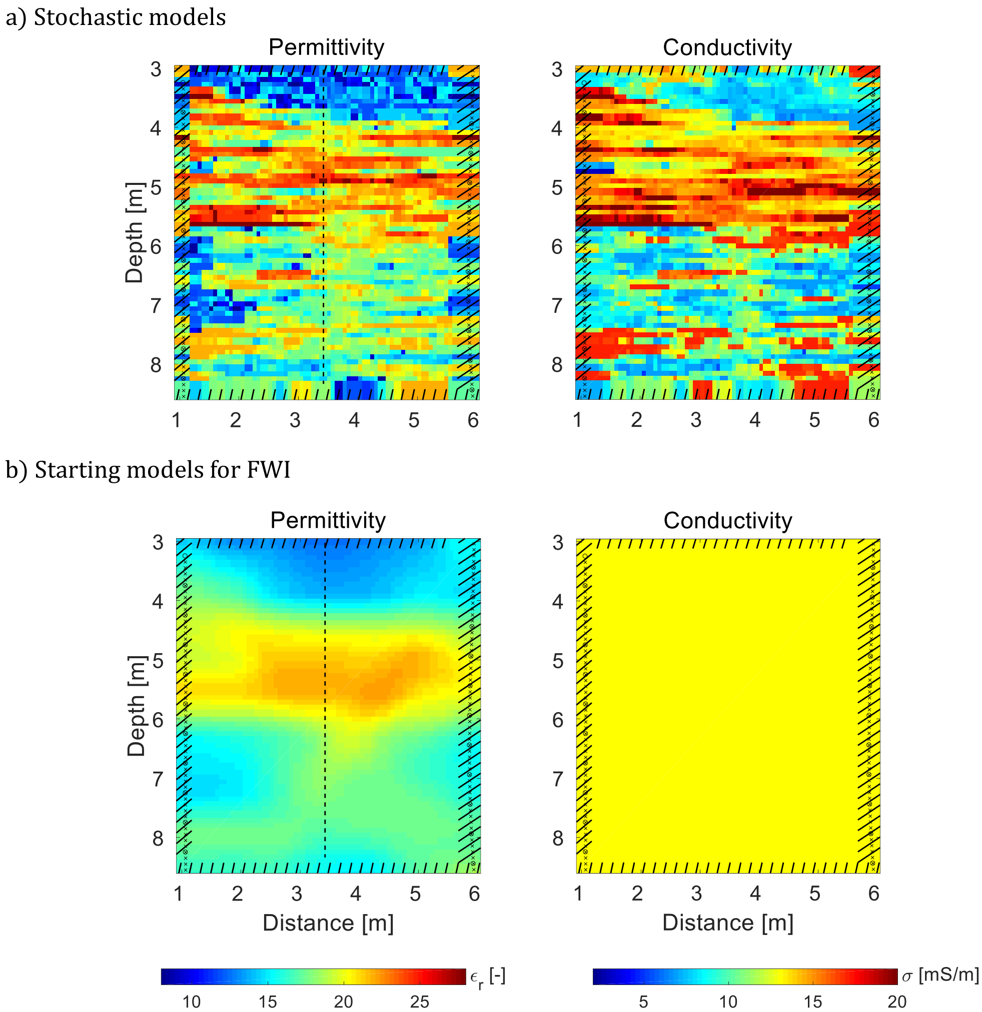


Figure 4.3. a) Results of ε_r and σ based on the stochastic simulation that are used to generate the realistic synthetic GPR data. The shadow zones at the boundaries indicate the extended domain of the inversion. The vertical dashed line indicates the selected Sto-CPT location used to compute the filter and to amplify the wavenumber of the FWI results. b) FWI starting models. Ray-based result for ε_r using the GPR data based on a) and a uniform starting model for σ.

4.2.2 Conventional FWI

First, we applied the ray-based method to generate the relative permittivity starting model for the FWI (Figure 4.3b). For the electrical conductivity starting model, a homogeneous model with 13 mS/m was used. This is consistent with previous inversion results of experimental GPR data from this test site. Using the ray-based starting model, the standard effective source wavelet SW_{Ray} was computed using the deconvolution approach. To determine the optimal iteration of the final FWI results, the normalized remaining gradient values of the ε_r results and the normalized RMSE were analyzed (Figure 4.4a). Thereby, iteration 28 was selected as the optimal iteration of the FWI considering the stopping criterions (Klotzsche et al., 2019b), and the FWI ε_r and σ results are shown in Figure 4.4b. A comparison of the ray-based results (Figure 4.3b), the FWI results (Figure 4.4b) and the real stochastic models (Figure 4.3a) indicates that the FWI results show higher resolution images and more details in the tomograms than the ray-based results, but still a certain mismatch to the real models can be noticed. Note that a good fit between the modeled traces based on the FWI results and observed data was achieved and almost no remaining gradient was present (not shown) for chosen iteration.

4.2.3 Construction of the wavenumber filter

To obtain a generalized filter in the wavenumber domain for the synthetic data set, we applied Equation 4.1 to smoothen highly fluctuating amplitudes of the selected and interpolated 1D FWI permittivity (dashed line in Figure 4.4b) and the stochastic CPT (Sto-CPT) data (Figure 4.3a). Note that both data sets are transformed into porosity using Equation 3.1. To find the optimal span value of the smoothing function, the 1D wavenumber-amplified FWI results and the filtered 1D stochastic CPT data are compared in wavenumber amplitude domain by computing the root mean square error (RMSE) and correlation coefficients (R) for different span values (Figure 4.5a). A final span value of 21 was selected because it provided a high R and a low RMSE values. Using this value, we transformed and smoothed the three different results (ray-based, FWI and Sto-CPT) in the wavenumber domain (Figure 4.5b). Here, the selected maximum threshold wavenumber was 2.00 m^{-1} (vertical dashed line) so that the generated filter still provided approximately monotonically increasing results. Finally, the filter was calculated by dividing the smooth Sto-CPT by the smooth FWI results. To intuitively show the resolution differences of the different methods along the selected vertical profile, we analyze the porosity value distribution along the depth direction from 3 m to 8.28 m separately (Figure 4.6). The comparison of the full wavenumber information for three different results along this vertical profile indicate that the resolution differences of the three approaches (Figure 4.6a). In addition, comparisons of low wavenumber parts of different results indicate WA-FWI results are better fitting with the filtered Sto-CPT data, which means the calculated filter is valid along the 1D profile (Figure 4.6b). A quantitative comparison of the results can be found in Table 4.2.

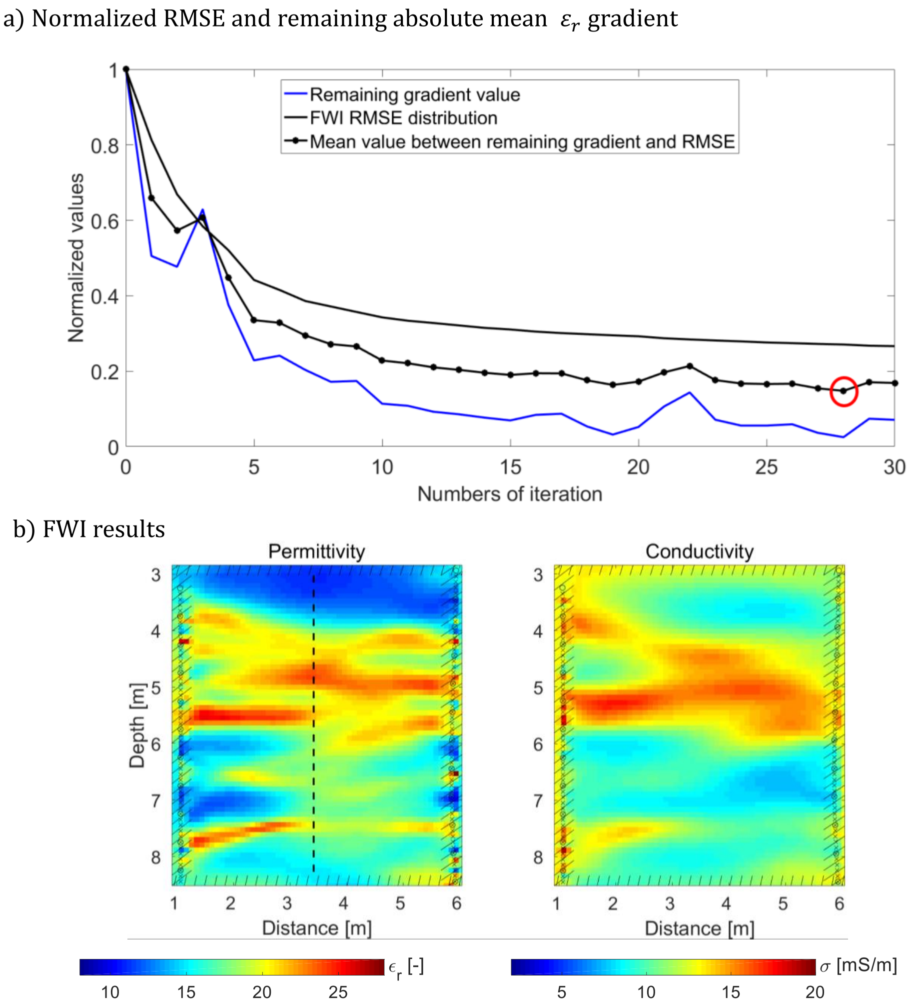
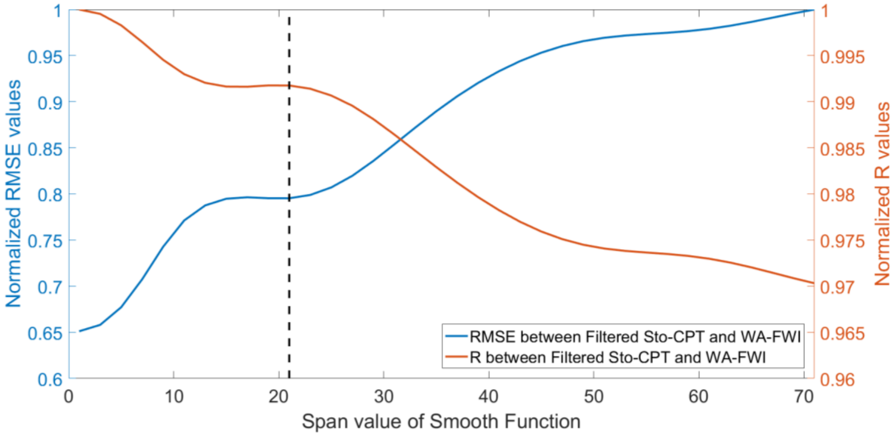


Figure 4.4. a) Evolution of the FWI RMSE misfit (black line) and the remaining absolute mean ε_r gradient (blue line) over number of iterations. Black asterisk line indicates the average value between the normalized remaining gradient values and the normalized RMSE. The red circle shows the iteration with the optimal value. b) The standard FWI permittivity and conductivity results after iteration 28. The dashed vertical line indicates the selected FWI profile to generate the amplifying filter.

a) Correlation between normalized RMSE, R and smooth span along Sto-CPT profile



b) Porosity spatial fourier domain along Sto-CPT profile

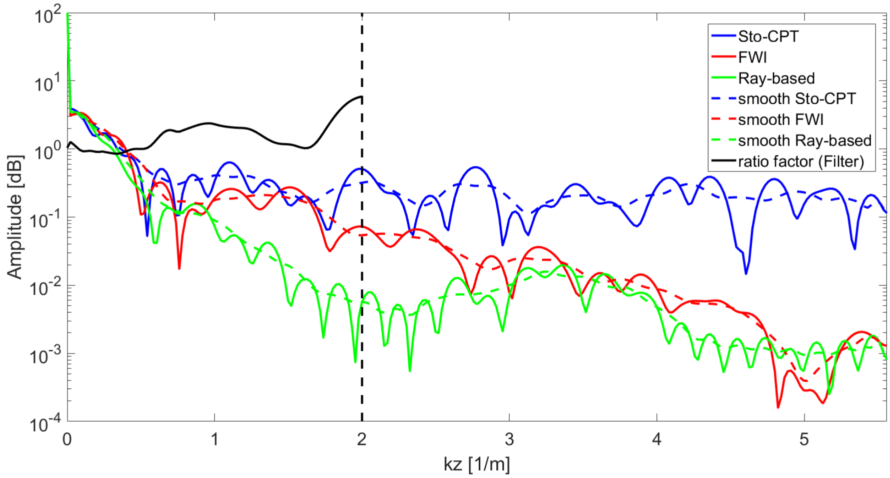


Figure 4.5. The distributions of RMSE and R values with changing the smooth function span values for the selected range of 0 to 71. The dashed line indicates the optimal span value 21. b) A comparison of the spatial wavenumber spectra of Sto-CPT data (blue), FWI (red) and ray-based (green) results. The filter is indicated by the black solid line, which is derived from the ratio between the smooth Sto-CPT and the smooth FWI (smooth span is 21). The dashed black line shows the maximum wavenumber for the filter.

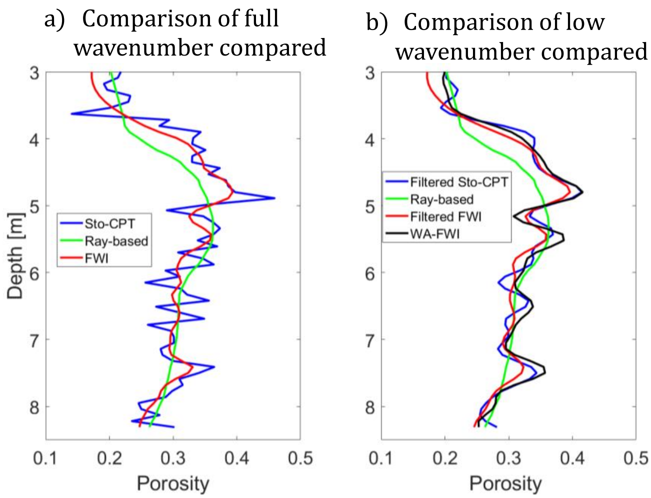


Figure 4.6. Comparisons of the a) full and b) low wavenumber information for Sto-CPT (blue), ray-based (green) and FWI (red) porosity results.

Table 4.2. Comparisons of the correlation coefficient *R* and the root-means square error RMSE of the filtered Sto-CPT, filtered FWI and wavenumber amplified FWI (WA-FWI) results given the maximum wavenumber, the suitable span value and the optimal FWI iteration value. *R* is Pearson's Correlation Coefficient between two variables (same for all following tables). The percentage in parentheses indicates the improvement of the WA-FWI RMSE to the filtered FWI RMSE.

Considered parameter	ε_r
Max. wavenumber for filter (m^{-1})	2.00
Span value of smooth function	21
Optimal iteration of FWI	28
***R* (Filtered FWI: Filtered Sto-CPT)**	0.9562
***R* (WA-FWI: Filtered Sto-CPT)**	0.9655
RMSE (Filtered FWI: Filtered Sto-CPT)	1.0907
RMSE(WA-FWI: Filtered Sto-CPT)	0.8830

4.2.4 Updating effective source wavelet and FWI

Although the developed filter was based on 1D vertical information, it was employed for the entire 2D domain of the conventional FWI permittivity model. Thereby, for some locations, especially those which are furthest away from the CPT profile location, higher wavenumber information appeared that was not consistent with the true model. To remove this inconsistent noise, we used an approach inspired by spectrum whitening deconvolution (Li et al., 2009). In particular, we replaced the traditional ray-based permittivity model with the 2D wavenumber-amplified FWI (WA-FWI) results and used the deconvolution method to generate an updated effective source wavelet (Zhou et al., 2019). To analyze and investigate which source wavelet strategy provides the most accurate final FWI results, we tested six different effective source wavelets based on different input models in the deconvolution approach.

The effective source wavelet used to generate the observed data is named real source wavelet SW_{Real}. The effective source wavelet based on the ray-based ε_r and a homogeneous σ (13 mS/m) is referred as SW_{Ray} and is used to generate the conventional FWI results. For comparison to the standard procedure without CPT data, this effective source wavelet is updated with the final conventional FWI results providing SW_{FWI}. As mentioned before, the source wavelet SW_{WA-FWI} is based on the wavenumber-amplified FWI (WA-FWI) permittivity results and SW_{Ray}. Similar to the conventional approach, also this wavelet is once updated with the final FWI results using SW_{WA-FWI}, which provides $SW_{New-FWI}$. For a complete comparison of all cases, an ideal source wavelet SW_{Sto} is estimated based on the real subsurface structures of the stochastic ε_r model (homogeneous σ model). Note that for better comparisons of these effective source wavelets, all source wavelets are normalized to their minimum in the provided figure (Figure 4.7). Comparing the six different effective source wavelets, a similar shape can be observed although a minimal time difference of the pulses is visible. Except for SW_{Ray} (blue line), all the wavelets show similar amplitude spectra in the frequency domain (Figure 4.7b). Note that the bandwidth for SW_{Ray} is smaller compared to the other wavelets suggesting FWI results with a lower resolution using SW_{Ray}. The bandwidth of $SW_{New-FWI}$ (cyan line) is slightly larger than the bandwidth of SW_{WA-FWI} (red line). As expected the bandwidth of SW_{Sto} and SW_{Real} are showing the largest bandwidth. By analyzing the unwrapped phases, we find that the phase of $SW_{New-FWI}$ is closest to the phase of SW_{Real}, especially for high frequency parts (Figure 4.7c) indicating that $SW_{New-FWI}$ should provide the most optimal effective source wavelet when the real models are unknown.

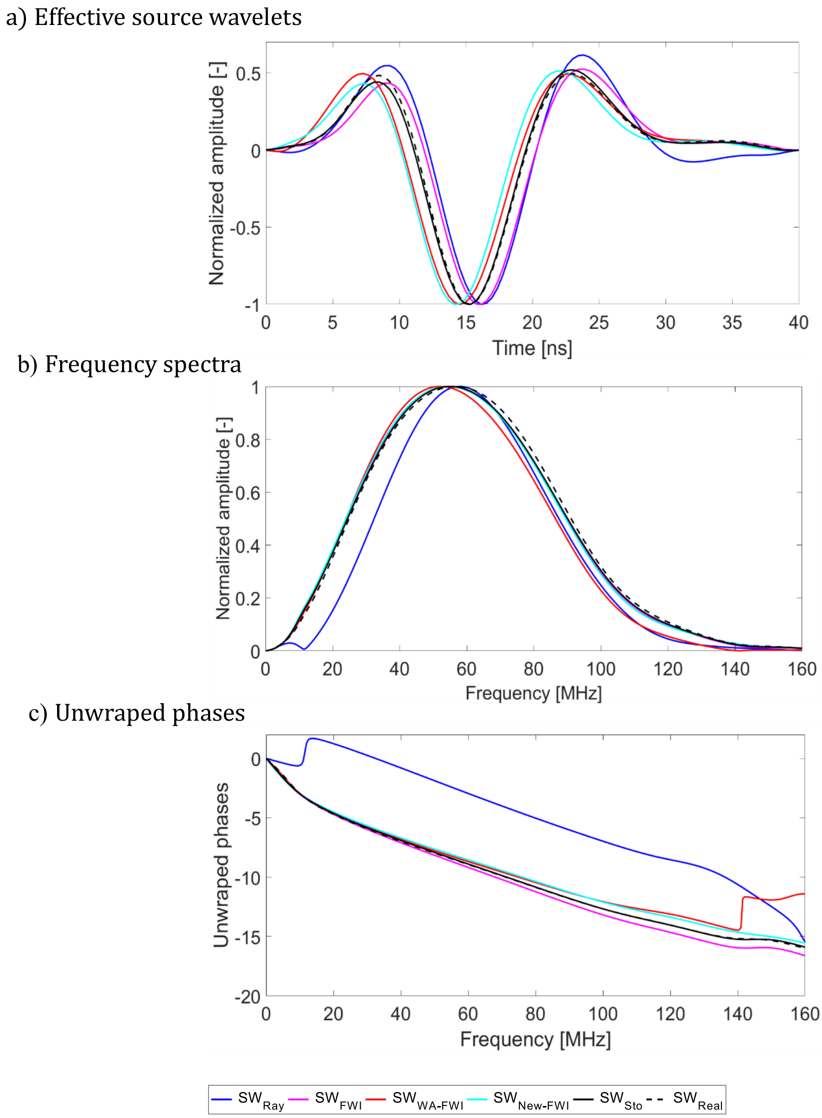


Figure 4.7. Comparisons of different effective source wavelets in a) time domain, b) corresponding frequency spectra from 0 to 160 MHz, and c) phase spectra based on the different processing steps indicated in Figure 4.1. Note that all source wavelets are estimated for different ε_r models, while σ models were the same for all steps with a homogenous model of 13 mS/m. Amplitudes of a) and b) are normalized to their corresponding minimum and maximum for a better comparison, respectively.

Table 4.3. Mean RMSE and R of different ε_r model comparisons for the entire 2D domain. F-Stochastic and F-FWI (to keep the same wavenumber information as WA-FWI) are filtered Stochastic and filtered FWI permittivity models, respectively. The F-FWI results with $SW_{New-FWI}$ are the optimal choice because of the lower RMSE and the higher R value.

Compared models (ε_r)	**Mean RMSE**	**Mean *R***
F-Stochastic and Ray-based	2.4836	0.7237
F-Stochastic and F-FWI (SW_{Ray})	1.4886	0.9222
F-Stochastic and F-FWI (SW_{FWI})	1.3234	0.9409
F-Stochastic and WA-FWI	2.1754	0.8713
F-Stochastic and F-FWI (SW_{WA-FWI})	1.6465	0.8907
F-Stochastic and F-FWI ($SW_{New-FWI}$)	1.3660	0.9315
F-Stochastic and F-FWI (SW_{Sto})	1.2029	0.9445

All source wavelets were tested using the same starting models based on the ray-based ε_r results and a homogenous σ model with 13 mS/m to verify the relationship between source wavelet bandwidth and the accuracy of the FWI results. Note that $SW_{New-FWI}$ is generated with a FWI ε_r model that includes the full wavenumber information. The final FWI results for the five different source wavelets (except for SW_{Real}) are shown in Figure 4.8a and 4.8b. RMSE values were computed based on the filtered stochastic permittivity model and the filtered FWI permittivity model in 2D domain. First thing to notice is that all FWI results show more details and structures as the ray-based results. Further, it is interesting to notice that although the final RMSE for the FWI results of SW_{FWI} and $SW_{New-FWI}$ are similar, for $SW_{New-FWI}$ more consistent structures close to the boreholes can be seen that match the input model better. In addition, except for the FWI conductivity results with SW_{Ray}, the other conductivity tomograms are very similar as expected because the wavenumber filter should only change the resolution of the FWI permittivity. Finally, using the stochastic permittivity model as starting model for SW_{Sto} did not significantly improve the FWI results compared to $SW_{New-FWI}$. A similar behavior can be observed by analyzing the vertical distribution of the correlation coefficient R and the RMSE for the filtered 2D permittivity models (Figure 4.8c and Table 4.3). Comparisons between WA-FWI and other FWI results show that the WA-FWI results have larger differences between 1 m to 3 m along the horizontal distance, which indicate that the filter is not valid in these zones due to overamplification. As expected, while all FWI results are slightly better resolved in the middle regions of the tomograms, FWI results are degraded in the vicinity of the boreholes due to the acquisition strategy in crosshole applications. Furthermore, the FWI results of $SW_{New-FWI}$ show a higher R and smaller RMSE values than the results of SW_{WA-FWI}, SW_{FWI} and SW_{Ray}, especially in vicinity of the left borehole. As expected the best FWI results are obtained using SW_{Sto}, which can only be obtained in synthetic model studies. In the absence of complete knowledge on the subsurface, the FWI results based on $SW_{New-FWI}$ showed the best results.

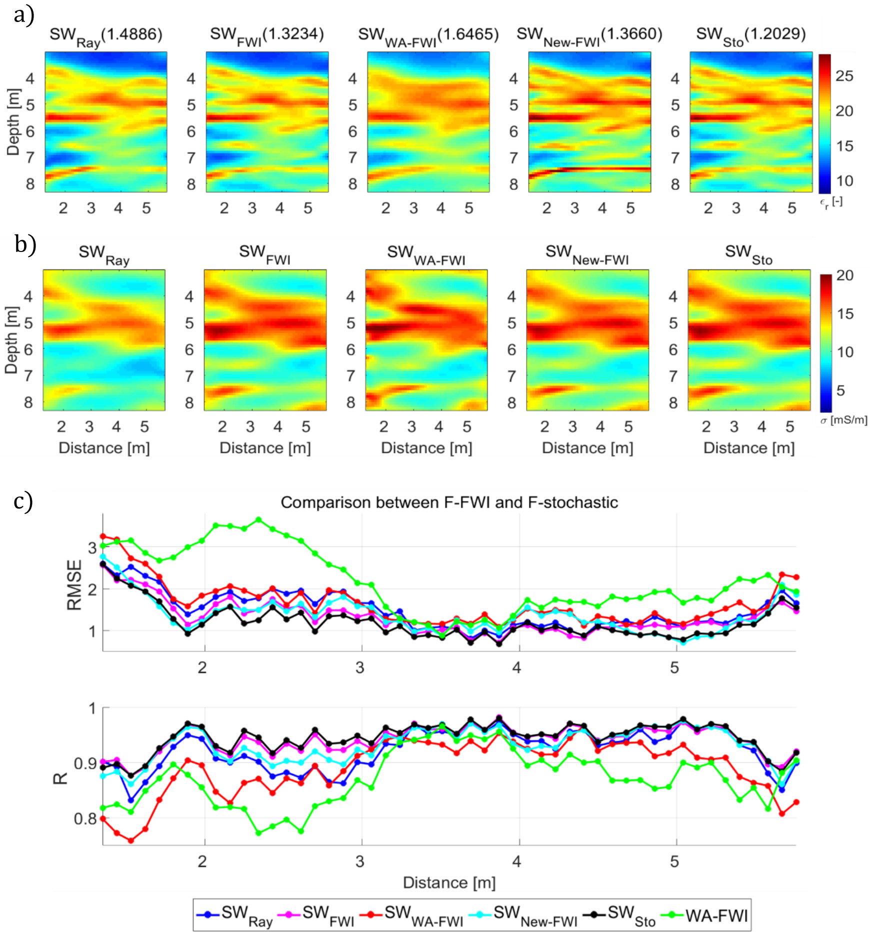


Figure 4.8. Comparisons of FWI a) permittivity and b) conductivity results using different effective source wavelets (Figure 4.7). Values in parentheses indicate mean RMSE between filtered FWI permittivity models and the filtered stochastic permittivity model in the entire 2D domain (see Table 4.3 for more details). c) Quantitative comparisons of the RMSE and R between filtered stochastic permittivity model and different filtered FWI permittivity results (same wavenumber as WA-FWI) along the vertical profile.

4.3 EXPERIMENTAL GPR DATA STUDIES

At the Krauthausen test site in Germany (Figure 4.2c), we measured the uppermost aquifer using crosshole GPR with 200 MHz antennae between several boreholes (red lines in Figure 4.2d). A detailed description of the site was provided by Vereecken et al. (2000). The measured uppermost aquifer can be broadly divided into three layers (Figure 4.2a): A poorly sorted gravel layer extending from 1 m to 4 m depth; the middle sand layer extending from 4 m to 6 m depth; and a bottom layer including sandy and gravely grains extending from 6 m to 11.5 m depth (Tillmann et al., 2008). For the acquisition of the experimental data, a semi-reciprocal acquisition setup (Figure 4.2b) was used with transmitter and receiver spacing of 0.5 m and 0.1 m, respectively. The water table was approximately at 2 m depth during the measurements. Therefore, GPR measurements started below 3 m depth. The CPT profiles that are closest to the crosshole sections are shown in Figure 4.2d (red asterisk). To improve the crosshole GPR FWI results, we analyzed five GPR cross-sections and the corresponding CPT profiles. For five CPT locations, the CPT probe was pushed into the ground to record cone resistance, electrical resistivity, natural gamma, gamma-gamma and neutron activity values. Tillmann et al. (2008) changed the neutron log into water content by calibration. In contrast to Gueting et al. (2015), we reanalyzed the FWI results following the suggestion given by the Corrigendum to the paper (Klotzsche et al., 2020). Reanalysis of the time zero correction of the GPR data showed that there was an error in the automatic picking routine, which is now updated. Therefore, the conventional FWI results are different to the results of Gueting et al. (2015) and show generally higher permittivities and lower electrical conductivities, while the structures are similar.

In the first step, the porosity information of five 1D vertical CPT profiles was compared to the corresponding FWI porosities, and the wavenumber filter for each borehole pair was estimated separately (Figure 4.9). Note that the original CPT data was used (Tillmann et al., 2008) without applying a shift as proposed by Gueting et al. (2015). For the experimental GPR data, a smooth function span value of 27 was selected for all cross-sections and the maximum threshold wavenumber of the filters was 2.31 m^{-1} (to keep the generated filters approximately monotonically increasing). The 1D porosity amplitude values along the CPT profile locations in the amplitude-wavenumber domain clearly show that the CPT values contain the largest bandwidth, whereas the FWI results have a reduced bandwidth, and the ray-based data have the lowest bandwidth for all five borehole pairs (Figure 4.9a-e). By comparing the five obtained filters (Figure 4.9f) in the wavenumber range of 0 to 2.31 m^{-1}, it can be noticed that the filter of profile 103 between boreholes 62-30 differs from the other four filters near 0.5 m^{-1}. Note that the cross-section distance between the boreholes 62-30 is 6.16 m, which is larger than others (Figure 4.2d).

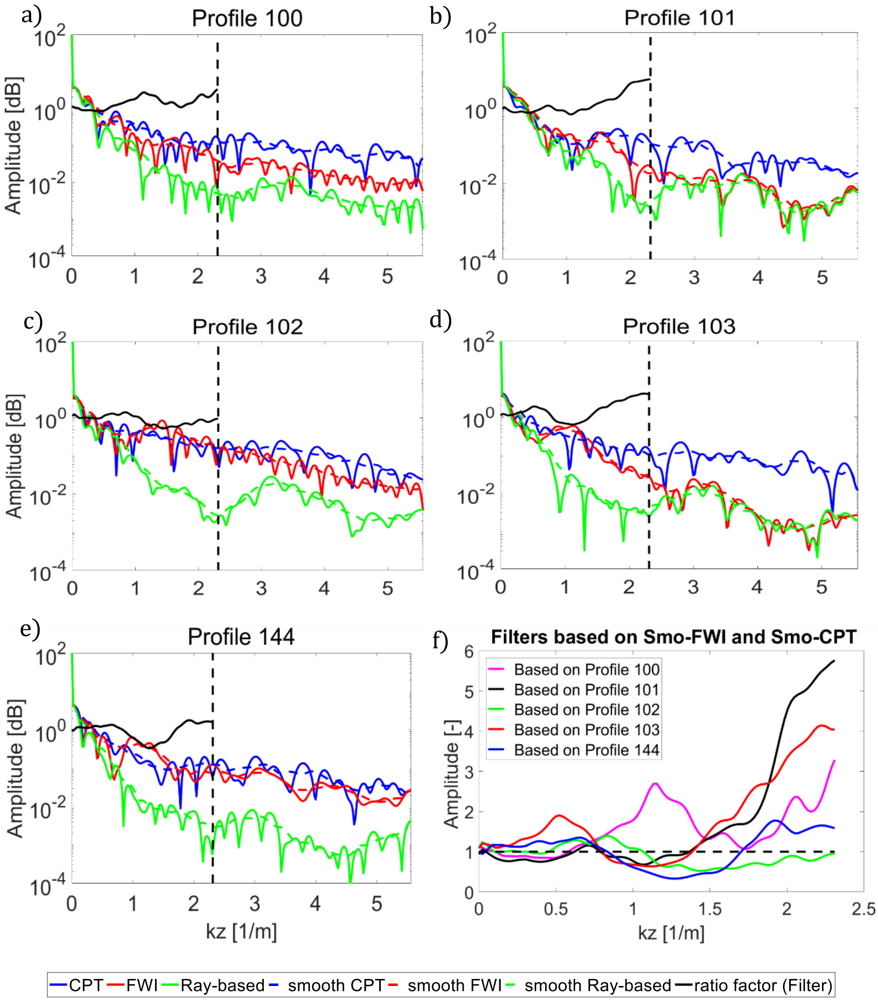


Figure 4.9. a) - e) Comparisons of spatial frequency spectra of the CPT data (blue), the ray-based (green) and the conventional FWI (red) results in the wavenumber domain for different profiles (see Figure 4.2d for the locations of the profiles). The wavenumber filter is indicated by the black solid line for each profile. f) Comparisons of the five filters, where a clear difference of profile 103 to the other profiles near 0.5 m^{-1} is noticeable.

In the next step, these five wavenumber filters are applied to derive WA-FWI results between each borehole pair and the corresponding updated effective source wavelets $SW_{New-FWI}$ (Figure 4.10). Similar to the synthetic case study, these wavelets were generated by updating the standard effective source wavelets with the deconvolution approach and the WA-FWI results (σ starting model homogenous with 13 mS/m) to the source wavelet SW_{WA-FWI} and then updating these wavelets to $SW_{New-FWI}$ (see flow diagram in Figure 4.1). Note that we only show the permittivity FWI results based on $SW_{New-FWI}$, since this source wavelet provided satisfying results in the synthetic study. The final effective source wavelets show similar shapes with slight shifts in time (Figure 4.10a) and similar bandwidth in the frequency spectra (Figure 4.10b and c). Note that the effective source wavelet SW_{WA-FWI} for the cross-section 62-30 is solved based on "WA-FWI subtract 1" because the WA-FWI results are too large to solve an effective source wavelet. One possible reason is that the larger borehole distance causes the lower FWI resolution and then the solved filter over amplified the WA-FWI values.

The traditional FWI ε_r results (Figure 4.11a) using SW_{Ray} are used to derive the WA-FWI results (Figure 4.11b). The updated FWI results (Figure 4.11c) are derived using the corresponding updated source wavelets $SW_{New-FWI}$ and the ray-based ε_r results as starting models. Similar to the synthetic studies, the WA-FWI results show over amplified features close to the boreholes. The vertical dashed lines indicate the locations of CPT data for each pair of boreholes. The updated FWI results show more consistent structures in the individual planes and at the crossings of the boreholes in comparison to the conventional FWI permittivity results. Generally, better RMSE values and R factors are obtained for the updated FWI results than for the conventional FWI (Table 4.5).

Finally, to verify the updated FWI results, we computed and compared the FWI porosity results using Equation 3.1 with the CPT porosity values (Figure 4.12). Thereby, we first compared the wavenumber-amplified FWI results with the filtered CPT (Figure 4.12a) similar to the synthetic case study (Figure 4.6b). Note that we selected the same depth of the CPT data from 3 m to 8.28 m for five different measurements to calculate the filters and compared with different FWI results. Both Figure 4.12a and Table 4.4 show the generated filters along respective 1D CPT profile. Comparisons of full wavenumbers information for CPT (blue), ray-based results (green), conventional FWI (red) and updated FWI (black) along each CPT profile are shown in Figure 4.12b (quantitative comparison in Table 4.5). An improved fit between the CPT and updated FWI results in contrast to the conventional FWI porosity results is visible. By comparing the computed R and RMSE between the CPT data and the 1D different FWI results, we conclude that the updated effective source wavelets, which incorporate the CPT information, improved the FWI permittivity results for all planes.

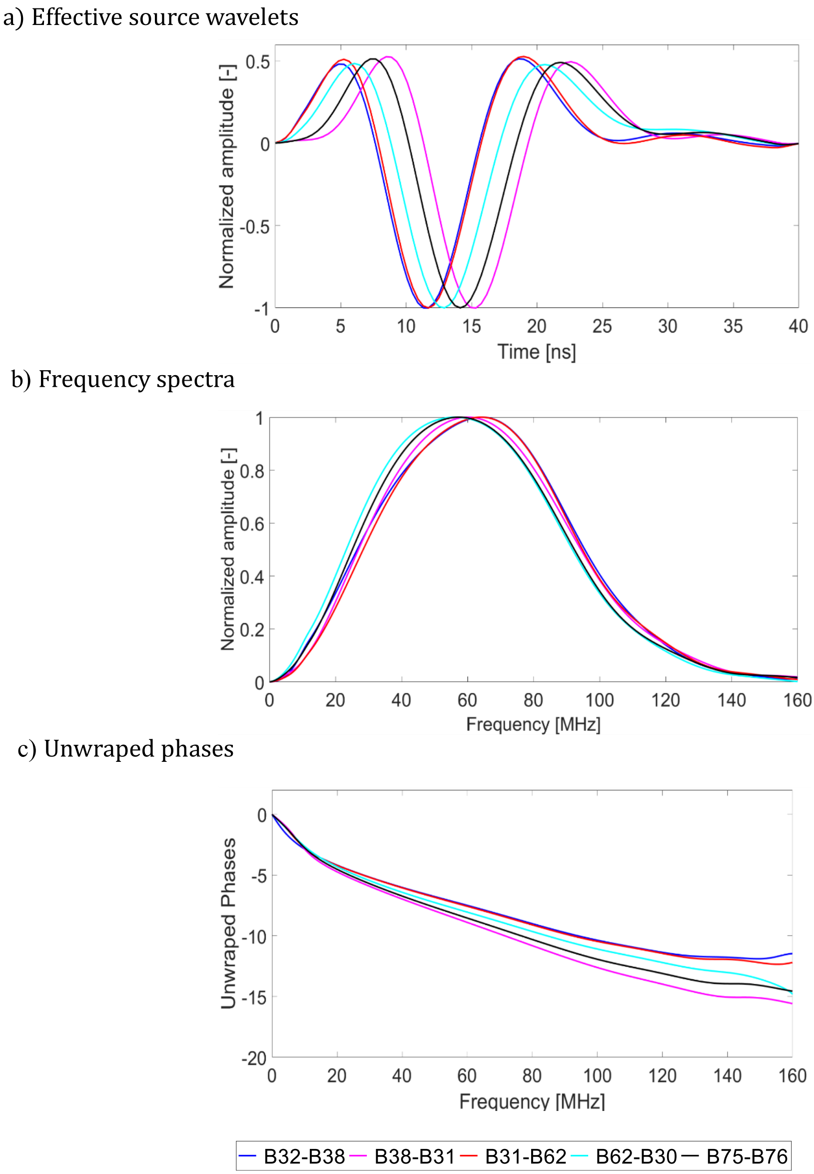


Figure 4.10. Comparison of the updated effective source wavelets of the five cross-sections used for the experimental study in a) time domain, b) frequency and c) phase spectra. Amplitudes of a) and b) are normalized to their corresponding minimum and maximum for a better comparison, respectively.

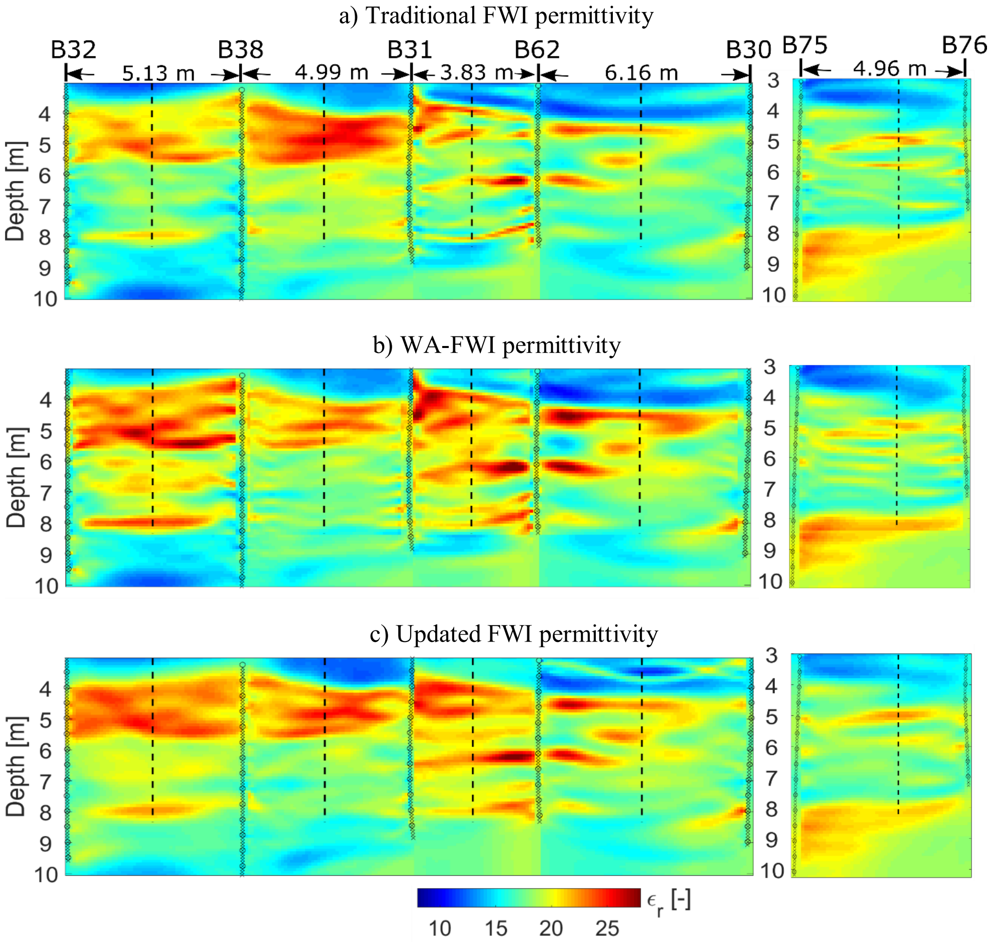


Figure 4.11. a) Traditional permittivity FWI results using SW_{Ray} for the five cross-sections. Circles and crosses indicate the transmitter and receiver locations, respectively. Dashed lines present the locations of the CPT profiles. b) Shows the permittivity images of the wavenumber-amplified FWI using the filters shown in Figure 4.9. c) Updated FWI results using the updated effective source wavelets as shown in Figure 4.10.

Table 4.4 Comparisons between filtered CPT, filtered FWI and wavenumber amplified FWI (WA-FWI) porosity results of the experimental data set from the Krauthausen site. Percentages in parentheses indicate the improvement of the WA-FWI RMSE to the filtered FWI RMSE.

	32-38 (5.13 m)	**38-31 (4.99 m)**	**31-62 (3.83 m)**	**62-30 (6.16 m)**	**75-76 (4.96 m)**
Profiles of CPT	100	101	102	103	144
Max. wavenumber for filter ($\mathbf{m^{-1}}$)	2.31	2.31	2.31	2.31	2.31
Span value of smooth function	27	27	27	27	27
Optimal iteration of FWI	30	22	30	15	26
***R* (Filtered FWI: Filtered CPT)**	0.7851	0.9278	0.8414	0.8711	0.8340
***R* (WA- FWI: Filtered CPT)**	0.7604	0.9031	0.8771	0.9054	0.9062
RMSE (Filtered FWI: Filtered CPT)	0.0436	0.0308	0.0386	0.0349	0.0269
RMSE(WA- FWI: Filtered CPT)	0.0291	0.0197	0.0210	0.0210	0.0207
	(33.3%)	(36.0%)	(45.6%)	(39.8%)	(23.0%)

Table 4.5. Comparisons of the full wavenumber CPT and FWI porosity results using different effective source wavelets. *R* and RMSE are calculated based on 1D full wavenumber profile data. Percentages in parentheses indicate the improvement of the New-FWI RMSE to the traditional FWI RMSE.

	32-38 (5.13 m)	**38-31 (4.99 m)**	**31-62 (3.83 m)**	**62-30 (6.16 m)**	**75-76 (4.96 m)**
***R* (FWI: CPT)**	0.7576	0.9153	0.8049	0.8564	0.8149
***R* (New-FWI: CPT)**	0.7701	0.9189	0.8312	0.8569	0.8636
RMSE (FWI: CPT)	0.0448	0.0316	0.0410	0.0360	0.0285
RMSE (New-FWI : CPT)	0.0296	0.0249	0.0249	0.0272	0.0255
	(33.9%)	(21.2%)	(39.3%)	(24.4%)	(10.5%)

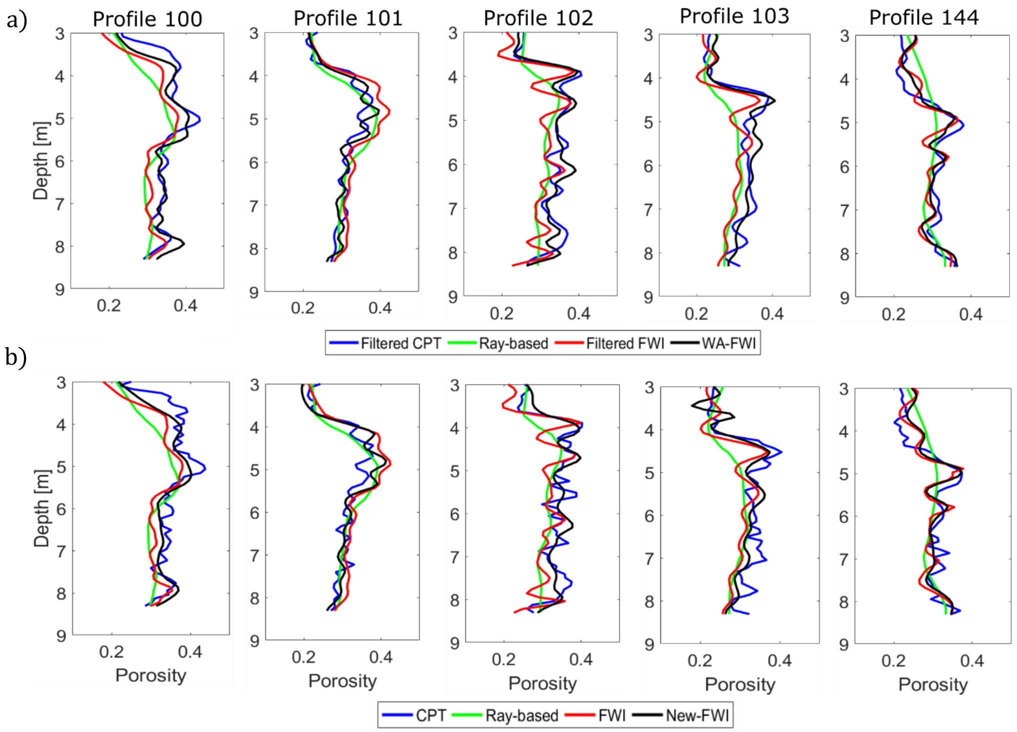


Figure 4.12. a) Porosity comparisons of the filtered CPT (blue), the ray-based results (green), the filtered FWI results (red) and the wavenumber-amplified FWI (black) along each vertical profile. b) Indicates the full wavenumber porosity results comparison of the CPT, ray-based, the FWI results (using SW_{Ray}) and the updated FWI results (using $SW_{New-FWI}$).

To analyze the behavior of the model updated for permittivity and conductivity over the iterations of two FWI together with two FWI converged RMSE curves, we defined the differences $\Delta\varepsilon_r$ and $\Delta\sigma$ between the models for different iterations (iter) to the starting models according to the approach of Klotzsche et al.(2019b). The computing details can be found from the following Equations 4.4 and 4.5:

$$\Delta\varepsilon_r(iter) = \sum_{i=1}^{nx}\sum_{j=1}^{nz}\left|\varepsilon_r{}^{iter} - \varepsilon_r{}^{start}\right| / \sum_{i=1}^{nx}\sum_{j=1}^{nz}\left|\varepsilon_r{}^{start}\right|, \tag{4.4}$$

$$\Delta\sigma(iter) = \sum_{i=1}^{nx}\sum_{j=1}^{nz}\left|\sigma^{iter} - \sigma^{start}\right| / \sum_{i=1}^{nx}\sum_{j=1}^{nz}\left|\sigma^{start}\right|, \tag{4.5}$$

where nx and nz show the cells number in the horizontal and vertical directions of the inversion domain, respectively.

By analyzing the developments of two FWI tomograms and the differences in the model update for permittivity and conductivity (Figure 4.13), it can clearly be noted that the starting synthetic data with the updated source wavelets are more much differences between the observed data than the tradtional starting synthetic data except borheoles 38-31(Figure 4.13b) because higher dashed green curves values (iter=0). The main reason is the amplitude differences because the updated source wavelets are corrected according to WA-FWI permittivity models. The second point is the conductivity updated are more efficient than the permittivity for both FWI approaches in the early interations.

Considering the small perturbation factors are important to determine the step-lengths for permittivity and conductivity in the inversion process, therefore we need to be carefully to chose these values. These perturbation factors are small enough to guarantee the perturbed model is in the linearity range to avoid overshooting. Meanwhile, to avoid round-off errors of the computer system, the perturbation factors should be large enough (Meles et al., 2010). In addtion, different perturbation factors can effect the final results in Figure 4.13. Finally, we summarized the details of parameters in the inversion for two FWI in Table 4.6.

Table 4.6. Comparisons of selected iterations, FWI converged RMSE, and perturbation factors for two FWI approaches.

	32-38 (5.13 m)	**38-31 (4.99 m)**	**31-62 (3.83 m)**	**62-30 (6.16 m)**	**75-76 (4.96 m)**
			Traditional FWI		
Optimal Iteration	30	22	30	15	26
RMSE (10^{-7})	8.5055	9.4156	9.1763	7.0538	8.8673
Perturbation factor (ε_r)	10^{-2}	10^{-2}	10^{-2}	10^{-1}	10^{-1}
Perturbation factor (σ)	10^{0}	10^{0}	10^{0}	5×10^{0}	10^{0}
			Updated FWI		
Optimal Iteration	25	30	18	24	24
RMSE (10^{-7})	8.2225	6.4862	9.5057	5.6149	7.2403
Perturbation factor (ε_r)	10^{-2}	10^{-2}	10^{-2}	10^{-1}	10^{-1}
Perturbation factor (σ)	10^{-1}	10^{0}	5×10^{-1}	2×10^{0}	10^{0}

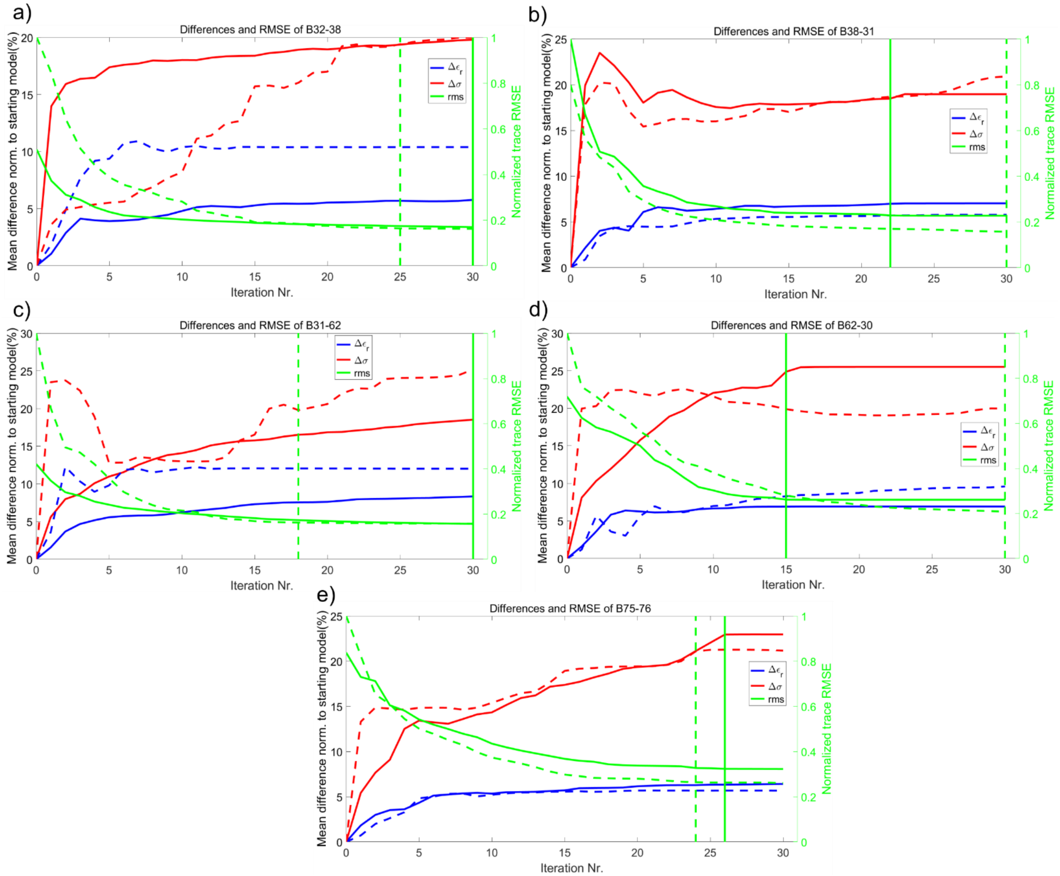


Figure 4.13 a) - e) Summations of the differences between the FWI permittivity (blue) and conductivity (red) to the starting models over the number of iterations for 5 boreholes. Green curves indicate the corresponding FWI converged RMSE. The vertical green lines indicate the thresholds of selected optimal iterations. The traditional and updated results are shown by solid and dashed curves, respectively.

4.4 CONCLUSIONS AND OUTLOOK

We demonstrated a new approach to improve the permittivity FWI results by incorporating additional information from CPT data. The novel approach was tested and verified at a realistically synthetic case study and applied to an experimental data set from the Krauthausen test site. To improve the FWI results, we proposed to design a 1D wavenumber filter based on CPT data and apply this filter to the 2D conventional FWI results. To verify the approach of updating the source wavelet based on the CPT data, we generated a stochastic simulated model of the Krauthausen test site. Combining the conventional FWI permittivity results and Sto-CPT data, we generated a filter that was applied in 2D FWI domain and obtained the WA-FWI results. To remove the inconsistent high wavenumber data of the wavenumber-amplified FWI results, we estimated an effective source wavelet SW_{WA-FWI} based on the WA-FWI results. Further, we used five different effective source wavelets to perform FWIs to determine the best effective source wavelet. The synthetic studies indicate that by applying an additional source wavelet correction cycle with the deconvolution approach and obtaining an enhanced source wavelet $SW_{New-FWI}$, the FWI results could be improved even further. Thereby, the comparisons of updated FWI permittivity results and the CPT data show improved results and more consistent structures in contrast to the conventional FWI.

The new approach for optimizing the effective source wavelet with the CPT data, was tested at experimental GPR datasets of five cross-boreholes sections. Similar to the synthetic study, the updated FWI results based on the effective source wavelet $SW_{New-FWI}$ show an improved consistency of the images in the entire domain and improved FWI ε_r results. The comparison of the updated FWI and CPT porosities confirmed the improvement in contrast to the conventional FWI results. In the following research, we will try to tame the non-linearity problem by gradually expanding the bandwidth of the updated effective source wavelet.

Chapter 5

Improvement of crosshole GPR FWI results by using progressively expanded bandwidths of the data[1]

In this chapter, we introduce a new approach that improves the starting model problematic and is able to enhance the reconstruction of the subsurface medium properties. The new approach tames the non-linearity issue caused by high contrast complex media in the inversion by applying different designed bandpass filters, which are progressively expanded to the full-bandwidth effective source wavelet and the observed GPR data. The resulting permittivity FWI model with the progressively expanded bandwidths of both modeled and observed data (PEBDD) is used in the next step as updated starting model and is applied to update the effective source wavelet. The following FWI with the full bandwidth data (FBD) and the second-updated effective source wavelet is able to enhance the reconstruction of the permittivity and electrical conductivity results in contrast to the standard FWI results. The new approach has been applied to two synthetic case studies and an experimental dataset, while the field data was additionally compared to cone penetration test data for validation.

---

[1]adapted from Zhou, Z., A. Klotzsche, H. Vereecken, and J. van der Kruk, 2020c. Improving the resolution of crosshole GPR FWI results by using progressively expanded bandwidths of the data: Journal of Applied Geophysics, submitted.

5.1 NOVEL PEBDD FULL-WAVEFORM INVERSION SCHEME

Our new updated approach proposes to tame the non-linearity issue of the time-domain FWI, which can be considered as extensions of the standard FWI procedure to improve the characterization of small-scale subsurface structures. Therefore, we consider the idea of frequency-domain approach that use longer wavelengths with lower frequency in the beginning of the inversion to avoid the cycle skipping problem, and that the bandpass filters are defined according to the center frequency of an effective source wavelet. As a first step, we construct a series of bandpass filters according to different high cut frequencies, while keeping the lowest cut frequency constant. Figure 5.1a shows an example for an effective source wavelet with a center frequency of 65 MHz and a bandwidth of 12 MHz to 140 MHz. For such a wavelet we would select the lowest cut frequency with 12 MHz, while the maximum cut frequency is considered larger than the center frequency of the effective source wavelet, which is in our case 68 MHz. To smooth the bandpass, we assigned two tappers with lengths of 12 MHz and 10 MHz for the starting and ending frequencies points, respectively (Figure 5.1a). These tapered bandpass filters are applied in the next step to observed GPR data and the effective source wavelet, which results after the FDTD in sub-modelled data with the sub-source wavelet that has a similar frequency spectra as the sub-observed data under using the same filter (Figure 5.1b green box).

Secondly, the FWI with the progressively expanded bandwidths of the modelled data and observed data (PEBDD) is performed (Figure 5.1b green box loop 1 and 2). The first starting models for the FWI are based on the ray-based results $Model_{Ray}$ ("Ray (ε_r) and Homo (σ)" in Figure 5.1b green box). Using the starting models with the sub-source wavelet and sub-observed data with the same smallest bandwidth, we perform a certain number iterations of the FWI (5 iterations are selected in Figure 5.1b green box loop 1: $n = 0$; n represents the number of bandwidth expansions). Note that it is also possible to use different number of iterations (Meles et al., 2011). The perturbation factors for the inversion, which are necessary to define the step-lengths for the gradient approach, are kept the same as for the standard FWI. The ε_r and σ FWI results after this 5th iterations are considered as the next new starting models for the next sub-data with progressively expanded bandwidth, while the maximum cut frequency was increased by 4 MHz in our case (Figure 5.1b green box loop 2). These steps are repeated until the selected maximum cut frequency is reached (Figure 5.1b green box loop 1: $n = n_{\max}$). Until this point all data (including source wavelet and observed data) used for the inversion are progressive bandwidth expanded.

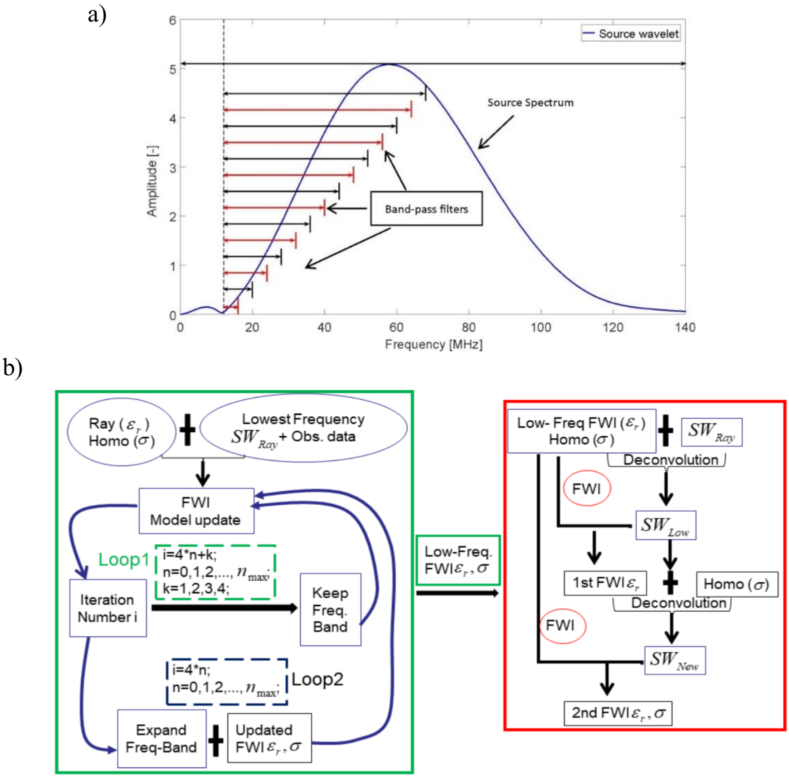


Figure 5.1. a) Example for an effective source wavelet amplitude spectrum and corresponding step wise increased bandpass filters (shown schematically by the black and red horizontal bars). These filters are progressively expanded until the center frequency is reached and applied to the observed data and the effective source wavelet. The bandpass filters start at the lowest frequency of 12 MHz (taper length 12 MHz) indicated by a vertical dashed line. The high cut frequency is stepwise expanded every five iterations of the FWI and has a taper of 10 MHz. After the highest corner frequency of 68 MHz is reached (center frequency of 65 MHz), all subsequent iterations use the full bandwidth of an updated source wavelet and the observed data. b) Flowchart of the new FWI PEBDD approach. In the first part (green box) the progressively expanded effective source wavelet and the observed data are used, in which i, n, and k represent all iterations using all sub-data, the numbers of filters, and individual iterations of each group sub-data, respectively. $n_{\max}$ is related to the center frequency of the wavelet. In this application, we select $n_{\max} = 13$. In the second part (red box) the full bandwidth data (FBD) FWI is performed, where two effective source wavelets corrections are performed and used during FWI. The final result is indicated by 2nd FWI with the SW_{New} as effective source wavelet.

Thirdly, the FWI with the full bandwidth data (FBD) is calculated (Figure 5.1b red box). From these results only the PEBDD permittivity FWI results with the maximum cut frequency are considered as new permittivity starting model in the next step (see "Low-Freq FWI (ε_r)" in Figure 5.1b). Together with the conductivity starting model with the same as the traditional one (Figure 5.1b "Homo (σ)"), we construct the new starting models $Model_{New}$. Note that tests indicated that also using the conductivity results with the maximum cut frequency as starting models did not improve the final results. The new starting models $Model_{New}$ and the traditional effective source wavelet SW_{Ray} are used in the next step to generate an updated effective source wavelet SW_{Low} with the full bandwidth data using the deconvolution approach (Equation 2.14, Figure 5.1b red box). After obtaining the updated source wavelet SW_{Low}, an updated FWI is performed with the new starting models $Model_{New}$ and the wavelet SW_{Low}. Generally, a second-updated source wavelet SW_{New} is necessary to further improve the FWI results. Thereby, the effective source wavelet SW_{Low} is updated to SW_{New} using the deconvolution method with the wavelet SW_{Low} and the permittivity model from the updated FWI results with SW_{Low}. The final updated FWI results with FBD are solved based on the second-updated source wavelet SW_{New} and the new starting models $Model_{New}$ (details can be found in Figure 5.1b).

5.2 SYNTHETIC CASE STUDIES

5.2.1 Synthetic case study I: Stochastic input models and ray-based starting models

To verify the aforementioned new FWI scheme for improving the GPR FWI results, we construct realistic synthetic models of relative dielectric permittivity and electrical conductivity using a stochastic simulation (Sequential Gaussian Simulation). For the simulation existing data set of the Krauthausen aquifer test site in Germany are used to generate a facies model with certain parameters (based on Tillmann et al. (2008) and more details can be found in Zhou at al. (2020b)). The construed aquifer consistent of a 3-layered structure similar to the Krauthausen aquifer: sandy layer between 1.0 m to 4.0 m, sandy gravel between 4.0 m to 5.5 m and below coarse gravel (Left side in Figure 5.3 and Figure 5.8). For the unsaturated zone above water table at 2.0 m, we chose a homogenous layer with a relative permittivity of ε_r = 4.4 (not shown, same for all following inversions). To generate the realistic synthetic GPR data (called observed data), we used a source wavelet (SW_{Real}; dashed curves in Figure 5.2a and b) based on available experimental data from the cross-section B38-31 of the Krauthausen test site (Gueting et al., 2015). Similar to the acquisition of experimental data, a semi-reciprocal acquisition setup was used with transmitter and receiver spacing of 0.5 m and 0.1 m, respectively. The realistic synthetic crosshole GPR data are modelled with the same time-domain 2D FDTD approach as used for the FWI. Similar to experimental data applications, we defined the ε_r starting model for the FWI by picking the first arrival times of the synthetic data and performed a ray-based inversion (1st column of Figure 5.3a; e.g., Dafflon et al., 2012). Similar to previous studies, we choose for a homogenous σ starting model of 13 mS/m (not shown) that was defined by testing various different homogenous models and used the results of the first cycle amplitude inversion of the experimental data.

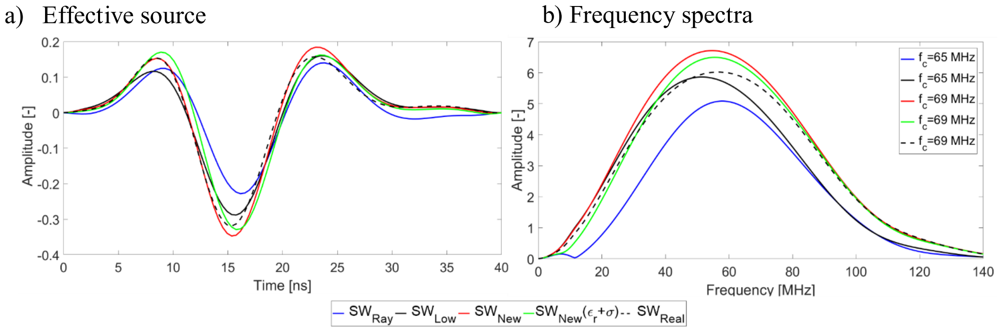


Figure 5.2. Comparisons of the a) effective source wavelets and the b) corresponding frequency spectra of the different steps of the new PBEDD approach used for synthetic case study I. The legend values in b) indicate center frequencies for different source wavelets. The source wavelet used to generate the 'observed data' is indicated with a dashed black line. Note that $SW_{New}(\varepsilon_r + \sigma)$ is the effective source wavelet based on the results shown in Figure 5.6.

To keep it as realistic as for experimental data applications, we used the ray-based starting models to estimate an effective source wavelet (SW_{Ray}; blue curves in Figure 5.2a and b) and performed the standard FWI (1st column of Figure 5.3b and c). As expected the FWI results show higher resolution images as the ray-based results within decimeter-scale resolution. Because of the known input models, we can calculate the mean absolute error MAE according to Equation 5.1 between the resolved FWI models and the true input models based on the stochastic simulation (Shown in the 1st column titles of Figure 5.3d and e). The final results were obtained after 28 iterations and the inverted results fulfilled the criteria for a reliable inversion. The corresponding RMSE curve that is calculated between the observed and modelled radar data is indicated by the blue curve in Figure 5.4 with the final RMSE of 6.9945×10^{-7}. Generally, most of the distinct features of the stochastic input models are resolved and good fit between the observed and the modelled data is obtained (see Figure 5.5a and b). By comparing the differences between observed data and modelled data for one exemplary data set (Figure 5.5c), we find most of the regions in the tomograms a small misfit is visible, but for some domains a still increased difference can be noticed. Considering that the real source wavelet has a center frequency of 69 MHz and the stochastic model has an approximate average permittivity value of 18, the corresponding wavelength of the GPR signal is 1.03 m. Comparing the model cell size 0.09 m with the wavelength 1.03 m, it is hard to match all the features of the stochastic model especially when the contrast is relatively high. For synthetic case studies, the mean absolute error (MAE) in the 2D domain can be described as:

$$MAE = \sum_{i=1}^{k} \sum_{j=1}^{m} (|RM_{i,j} - FWI_{i,j}|) \ /(k \times m) \ , \qquad (5.1)$$

where $RM_{i,j}$ and $FWI_{i,j}$ represent the input and the FWI models located at the cell of i, j. k and m show the cell numbers in 2D domain along horizontal and vertical directions, respectively.

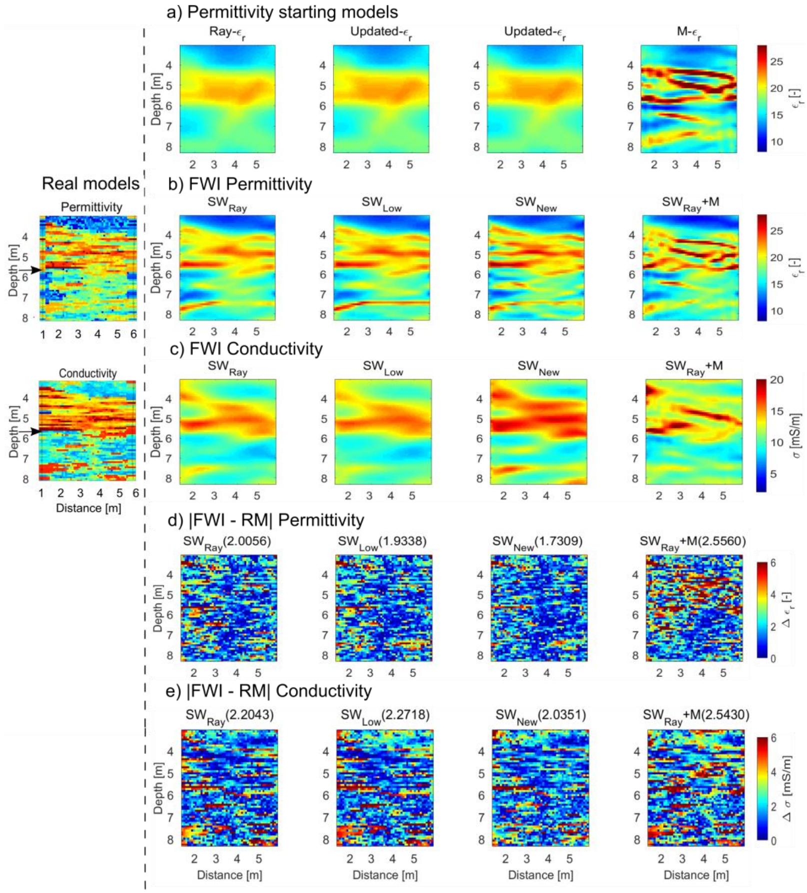


Figure 5.3. Overview of the FWI results using the different approaches and starting models. Left side shows the real input models based on the stochastic simulation (Zhou et al. 2020b). a) Permittivity starting models and corresponding FWI results for b) ε_r and c) σ for the different FWI approaches. The applied effective source wavelets are named in titles of the Figures. Image plots of the absolute error between the real input models and final FWI results are shown for d) ε_r and e) σ for the different approaches. The mean absolute error MAE of the entire 2D domain are shown in parentheses of titles.

Following our new introduced approach of the progressively expanded bandwidths of the modelled and observed data (PEBDD), we applied the different bandpass filters as shown in Figure 5.1a to the standard effective source wavelet SW_{Ray} and the observed data. For this study, we choose the optimal 5 iterations for the sub-data FWI by comparing with other iterations values. The first sub-data FWI started from the filtered sub-source wavelet and filtered sub-observed data with bandwidth 12-16 MHz and the ray-based starting models $Model_{Ray}$. Every 5 iterations the bandpass is increased by 4 MHz until the final cut off frequency of 68 MHz is reached. The final permittivity FWI results using the maximum bandwidth sub-data are shown in Figure 5.3a (title is "Updated- ε_r"). Combining the conductivity starting model with a homogenous value of 13 mS/m, we construct the new starting models $Model_{New}$ for the following FWI results and the updated source wavelet. Following the flowchart (Figure 5.1b red box), the effective source wavelet SW_{Low} is updated using the new starting models $Model_{New}$ and the wavelet SW_{Ray}. The effective source wavelet SW_{Low} together with the starting models $Model_{New}$ is used to calculate the FWI as shown in Figure 5.3b and c (second column) using the full dataset. By comparing the RMSE curve behavior of this inversion (black graph in Figure 5.4; black graph is covered by red graph before 71 iterations), we can notice that first the RMSE is stepwise increased until the full data are used and the RMSE curve with the full data is decreased after 23 iterations to a final value of 6.9749 $\times 10^{-7}$. The misfit between the input and resolved tomograms (Figure 5.3d and e) and the data misfit (Figure 5.5c) are slightly better than using the standard FWI approach.

These updated FWI results are in the last step considered to update the effective source wavelet a last time (SW_{New} in Figure 5.2a and b, red source wavelet) by using the new FWI ε_r results with SW_{Low} in the deconvolution approach. The corresponding final FWI results (Figure 5.3 third column) are derived with the SW_{New} and $Model_{New}$ as the starting models. Note that we also tested the new approach with the both permittivity and conductivity starting models based on the maximum bandwidth sub-data (Figure 5.6a). The generated FWI results (Figure 5.6b) indicate the results with SW_{New} ($\varepsilon_r + \sigma$) did not improve significantly than the FWI results with the SW_{New} and $Model_{New}$ as the starting models, therefore, we consider in further applications only the permittivity updated starting model. For comparisons, we performed the approach of Meles et al. (2011) that only uses the sub-source wavelets in the FWI, which means using the progressively bandwidth expanded modelled data (PBED), while the observed data includes the full bandwidth (Figure 5.3 fourth column). Note that the generated FWI results (including ε_r and σ) with the maximum bandwidth sub-source wavelet are the new starting models (only show the ε_r starting model with the title "$M - \varepsilon_r$" in Figure 5.3a) for the following FWI with FBD.

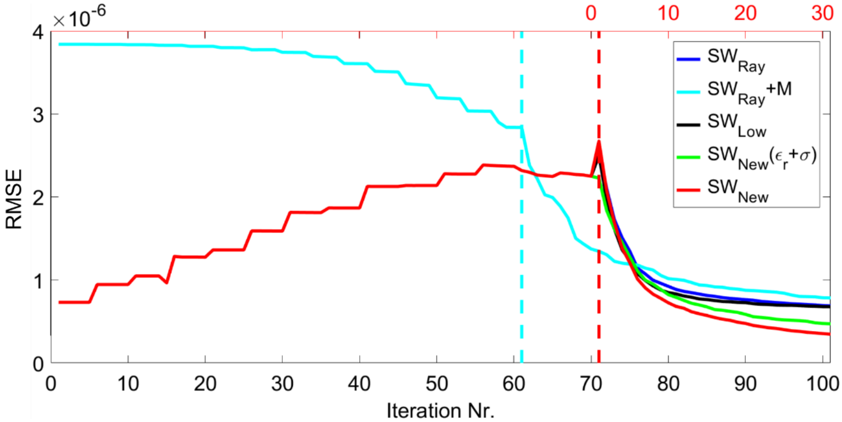


Figure 5.4. RMSE misfit curves of the different FWI approaches of the synthetic case study I. Blue graph represents the RMSE of the standard FWI using the ray-based starting models (iteration numbers are shown at the top with a red label). The cyan graph represents the RMSE behavior for the PBED inversion scheme based on Meles et al. (2011). Black and red graphs represent the new PEBDD inversion scheme using the 1st updated and 2nd updated source wavelet, respectively. The green graph shows the new FWI RMSE according to the updated starting models including ε_r and σ (Figure 5.6) by using PEBDD inversion scheme. Note that on the left side of the dashed lines, the progressively expansion of the bandwidths of effective source wavelets and observed data were used, while on the right side the FWI was performed considering the full bandwidth of all data.

The FWI results using the SW_{New} show the smallest misfit of the resolved tomograms and the smallest final RMSE indicating that this approach is able to resolve the input tomograms best. Note that we are able to improve the FWI results for the permittivity and the conductivity by 13.7% and 7.7% in comparisons to the standard FWI models using SW_{Ray} by comparing the mean absolute error MAE (Equation 5.1) values with two approaches, respectively (Table 5.1). Especially, the small-scale structures close to the boreholes are clearer and more accurately resolved as by the standard method. By comparing the final RMSE curves (only show the max iterations until 101 or 31; the same for Figure 5.9) of the five different FWI results (green graph based on results shown in Figure 5.6), the second-updated FWI results with the wavelet SW_{New} indicate the best convergence behavior resulting in the smallest residual value of 2.8996×10^{-7} after 50 iterations (red curve in Figure 5.4). Comparisons of 4 different FWI ε_r and σ results in 2D domain (Figure 5.3b and c) show the FWI results with the approach of Meles et al. (2011) are unrealistic. By computing the absolute errors between these different FWI results and the real input models in 2D domain (Figure 5.3d and e), we can find that the second-updated FWI results with SW_{New} are close to the input models indicated by the smallest mean absolute error MAE values (Table 5.1). By investigating the fit between the observed and the FWI modelled GPR data, a good fit can be observed for all the FWI results generally (one example shown for transmitter depth 5.69 m in Figure 5.5a and b), but analyzing the differences (Figure 5.5c) in more detail, we notice that the FWI using SW_{New} shows the smallest misfit.

In order to describe the regional differences of different FWI models, we computed the mean absolute errors MAE between the input stochastic models and the FWI models results along the horizontal direction (Figure 5.5d) and the vertical direction (Figure 5.5e) for ε_r and σ, respectively. The smallest MAE of the horizontal direction can be observed in the central part of the tomograms between 3.0 m and 5.0 m (Figure 5.5d). For all results the MAE of the horizontal direction increases towards the boundaries of the inversions domain, meaning the boreholes of the cross-section (Figure 5.5d). The MAE in horizontal direction and an increase in MAE towards the boreholes is related to the acquisition geometry and the related distribution of the ray coverage between the boreholes. Oberröhrmann et al. (2013) showed already that the resolution is highly effect by the acquisition geometry and depends on the ray-coverage. The MAE along the vertical direction shows more fluctuations around 2, while the results in horizontal direction are smoother. Interestingly to notice is that all the FWI results (except the approach based on Meles at al., 2011) show improved results with a smaller MAE in the vertical direction between 4.0 m and 5.5 m depth in comparison to other depths, where the small-scale structures are located. A small increase of the MAE at 5.5 m depth can be noticed at the lower boundary of the small-scale high permittivity zone. For both horizontal and vertical directions the FWI results based on the SW_{New} shows the lowest MAE in contrast to the other FWI results. Comparing the FWI results and the related MAE, we can conclude that our new updated PEBDD scheme is effective to enhance the complicated synthetic models FWI results for permittivity and conductivity in contrast to the standard FWI approach.

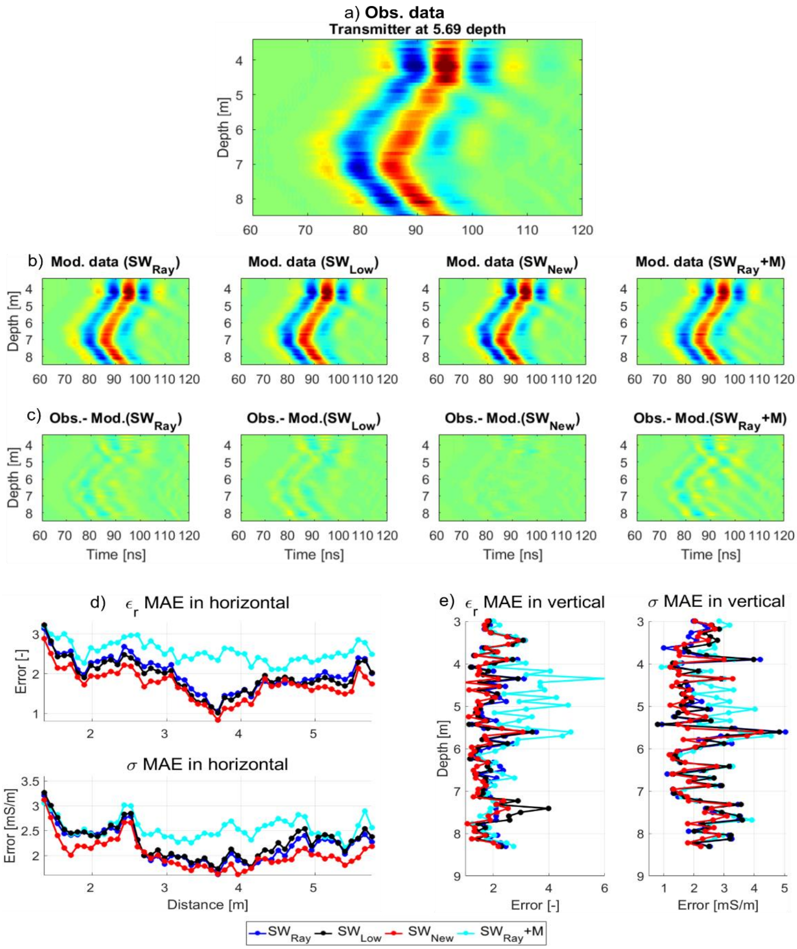


Figure 5.5. a) Observed data, b) modelled data based on the FWI results, and c) differences between the observed and modelled data for one exemplary data set of transmitter location at 5.69 m depth (see input models and black arrow for the transmitter location in Figure 5.3). Note the amplitudes in a), b) and c) are normalized to the maxima amplitude of the real observed data (shown range from -7×10^{-1} to 7×10^{-1}). The mean absolute error MAE of the permittivity and conductivity between the real input and the different FWI results are shown along d) horizontal cross-section and e) vertical direction.

For the synthetic case study I, we constructed a second set of starting models $Model_{New}(\varepsilon_r + \sigma)$, which were derived from the final FWI ε_r and σ results using the maximum bandwidth sub-data (Figure 5.6a). Following the flowchart (Figure 5.1b), an effective source wavelet $SW_{Low}(\varepsilon_r + \sigma)$ (not shown) is updated using the new starting models $Model_{New}(\varepsilon_r + \sigma)$ and the wavelet SW_{Ray}. This updated wavelet $SW_{Low}(\varepsilon_r + \sigma)$ together with the starting models $Model_{New}(\varepsilon_r + \sigma)$ are used in the next step to calculate the FWI results using the full dataset (not shown). These updated FWI results (including ε_r and σ) are in the last step considered to update the effective source wavelet $SW_{New}(\varepsilon_r + \sigma)$ (Figure 5.2a and b, green source wavelet). The corresponding final FWI results (Figure 5.6b) are derived with the $SW_{New}(\varepsilon_r + \sigma)$ and $Model_{New}(\varepsilon_r + \sigma)$ as the starting models. By comparing the final RMSE curves after 45 iterations (green graph in Figure 5.4), the second-updated FWI results with the wavelet SW_{New} (RMSE = 2.8996×10^{-7}; red curve) indicate a better convergence behavior and a smaller RMSE than the new updated results with the $SW_{New}(\varepsilon_r + \sigma)$ (RMSE = 3.8620×10^{-7}; green curve). Computing the absolute errors between the FWI results and the input models of the 2D domain (Figure 5.6c), it can be noticed that the new updated FWI results with $SW_{New}(\varepsilon_r + \sigma)$ show a less good fit than the second-updated FWI results with SW_{New} (Table 5.1). Therefore, we abandon this approach of using both permittivity and conductivity results derived by the FWI using the maximum bandwidth sub-data and consider for all other studies the homogeneous σ starting model with 13 mS/m.

Table 5.1. Comparison of the different FWI approaches for synthetic case study I using the mean absolute error MAE between the real input models and the different FWI results for the entire 2D domain, and the root-mean-squared errors RMSE between observed and modelled radar traces represents residual values. Percentages in parentheses indicate the ratio of the single FWI RMSE to the standard FWI RMSE with SW_{Ray}, while a decreased value means the higher improvement efficiency.

Real models	**MAE (ε_r)**	**MAE (σ)**	**RMSE (10^{-7})**
FWI (SW_{Ray})	2.0056	2.2043	6.9945 (100%)
FWI (SW_{Low})	1.9338	2.2718	6.9749 (99.7%)
FWI (SW_{New})	1.7309	2.0351	2.8996 (41.5%)
FWI ($SW_{Ray} + M$)	2.5560	2.5430	8.6499 (123.7%)
FWI ($SW_{New}(\varepsilon_r + \sigma)$)	1.8500	2.1851	3.8620 (55.2%)

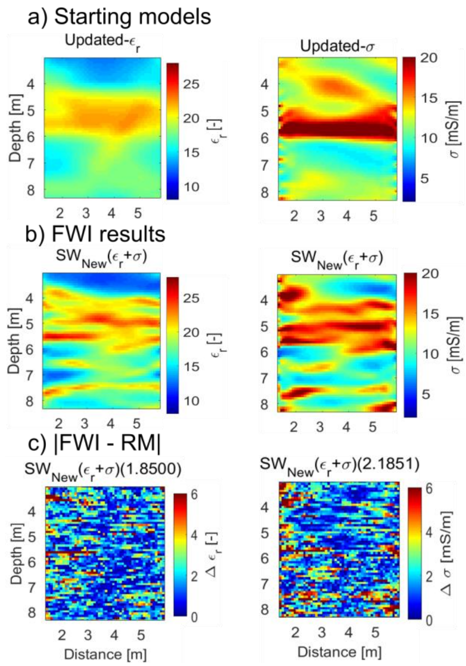


Figure 5.6. a) The permittivity and conductivity starting models based on the FWI results using the maximum expended bandwidth sub-data. b) Corresponding FWI results using the updated starting models and the updated effective source wavelet. Effective source wavelet is named in titles. c) Image plots of the absolute error between the stochastic input models and the final FWI ε_r and σ results. The mean absolute errors MAE of the entire 2D domain are shown in parentheses of titles.

5.2.2 Synthetic case study II: Permittivity starting model beyond the half-wavelength criteria

In the presence of high contrast layers ray-based results are often erroneous and need to be updated to be fulfilled the half-wavelength starting model criteria for the standard FWI. Here, we want to demonstrate the potential of the PEBDD scheme, which allows also starting models that are beyond the half-wavelength criteria. Therefore, we perform a second synthetic case study with the same stochastic input models as before and enforce the permittivity starting model to provide modelled data more than half a wavelength away from the measured data by enforcing a smaller ε_r model. The changed ε_r starting model was obtained by subtracting three from the normal ray-based ε_r model in the entire domain (Figure 5.8a, title is "Ray- ε_r(-3)"). The conductivity starting model is unchanged with a homogenous value of 13 mS/m. We follow the same approach as in the synthetic study case I and obtained the different updated effective source wavelets (Figure 5.7) and FWI results (Figure 5.8).

Firstly, we estimate the effective source wavelet $SW_{Ray(-3)}$ based on the erroneous starting models $Model_{Ray-3}$. We can clearly notice that the $SW_{Ray(-3)}$ based on the erroneous starting models is shifted in time to the right (Figure 5.7a). This is indicating that the permittivity starting model is currently too far away from the input models and that the wavelet is compensating for this by starting later in time (blue curves in Figure 5.7a and b). Note that a good effective source wavelet needs to start at 0 ns (Klotzsche et al., 2019b). The FWI results (first column in Figure 5.8b and c) based on these starting models and the effective source wavelet $SW_{Ray(-3)}$ fulfill the stopping criteria after 24 iterations and no remaining gradient is present. The FWI RMSE curve by using the source wavelet $SW_{Ray(-3)}$ is shown with the blue curve in Figure 5.9. Generally, the resulting FWI results show a lower ε_r than the input model demonstrating that the ε_r results trapped in a local minimum of the inversion process (Figure 5.8b, first column). This is also indicated by the differences between the input model and FWI results in Figure 5.8d, where many regions are present that show differences with more than 3 in permittivity. Note that this inversion still fulfills most of the criteria for a good inversion process. The only indicators that the inversion is too far away from the input model are provided by the effective source wavelet that is shifted in time and the high MAE (not available for experimental data applications). To fulfill the criteria that the effective source wavelet needs to start at 0 ns, we shifted in the next step the source wavelet by - 4 ns in the time domain $SW_{Ray(-3shift)}$ (cyan curves in Figure 5.7a and b) to ensure that it starts at 0 ns. The corresponding FWI results (Figure 5.8b and c, second column) show a better reconstruction of the permittivity results than the previous inversion and the final RMSE value after 30 iterations is 8.3326×10^{-7} (cyan curve in Figure 5.9). Meanwhile the FWI σ results are better than the FWI σ results with $SW_{Ray(-3)}$ (Figure 5.8c, columns 1st and 2nd).

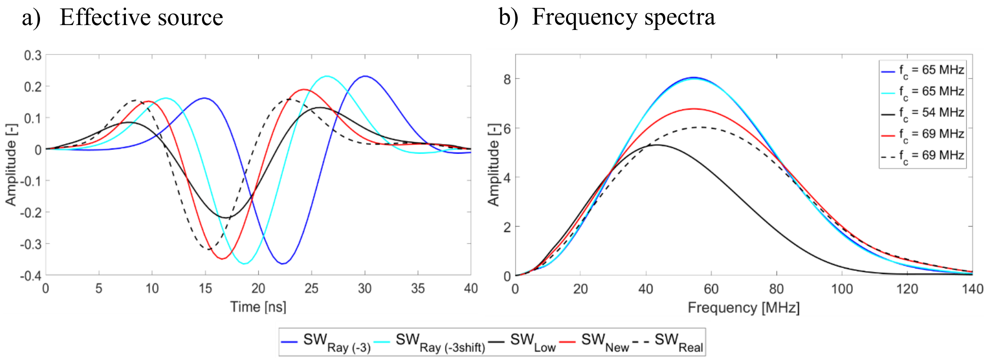


Figure 5.7. a) Comparisons of the different effective source wavelets used for synthetic case study II. The source wavelet used to generate the 'observed data' is indicated with a dashed graph. b) Frequency spectra comparisons from 0 to 140 MHz for the five different effective source wavelets. The legend values in b) indicate these effective center frequencies for different steps.

Similar to the previous synthetic case I, we apply the same bandpass filters for the same number of iterations to the shifted effective source wavelet $SW_{Ray(-3shift)}$ to enhance the reconstruction of the FWI results with the PEBDD approach. For the FWI with the different bandpass filters, the first starting models $Model_{Ray-3}$ are used. The final FWI ε_r results using the maximum bandwidth sub-data (Figure 5.8a, Updated- ε_r (-3)) and the homogenous σ model with 13 mS/m are used as the new updated starting models $Model_{New-3}$. It can be noticed that the updated permittivity starting model shows generally higher values in the entire domain and the model is much closer to the ray-based model used in the synthetic study I indicating that the PEBDD approach is able to enhance the erroneous starting model (compare Figure 5.3a and 5.8a). The updated starting models are considered to derive SW_{Low} before performing the FWI.

The updated FWI results using SW_{Low} show improved permittivity results and more continues structures, although the conductivity performs less good indicated by the images differences (Figure 5.8b and c, third column) and the higher RMSE value for the final iteration (Figure 5.9, black graph). The final permittivity results of this inversion and SW_{Low} are used to update the effective wavelet a last time to SW_{New} (Figure 5.7a and b, red source wavelet). The corresponding FWI results (Figure 5.8b and c, fourth column) after 50 iterations are derived with the SW_{New} and $Model_{New-3}$ as the starting models. These results show an improved reconstruction of the permittivity and the conductivity results in comparison to the previous FWI results. Furthermore, comparisons of the absolute error MAE tomograms for different FWI ε_r and σ results in 2D domain (Figure 5.8d and e) show that the second-updated FWI results with the wavelet SW_{New} are the most accurate reconstruction of the input models.

The MAE tomograms and behavior of the RMSE curves are the smallest for the second-updated FWI results with the wavelet SW_{New} and are indicating the best FWI results (more details in Table 5.2). Analyzing the misfit between the measured and FWI modelled data shows that the second-updated FWI results provides the best fit and indicates that this inversion obtained a model that describe the data well and best, while for the other three FWI results a significant misfit can be observed (Figure 5.10a and b, exemplary for transmitter at 5.69 m depth). To describe the regional differences of different FWI models, we compute the MAE between the stochastic input models and the FWI models results along the horizontal direction (Figure 5.10c) and the vertical direction (Figure 5.10d) for ε_r and σ, respectively. Thereby, we can notice that the FWI ε_r results with $SW_{Ray(-3)}$ display the largest MAE values for both directions. The FWI σ results with SW_{Low} (black curves) have a large MAE, which implies that they cannot be resolved by only using the first-updated effective source wavelet in time when the starting models are too far away from the real models. Finally, we can conclude that the progressively expanded bandwidth scheme can not only improve the FWI results, but also that it is able to retrieve accurate FWI results for starting models more than a half-wavelength away from the measured data. Therefore, a lot of previous detailed work to construct good starting models for experimental GPR data can be reduced and the application to field data could be much easier.

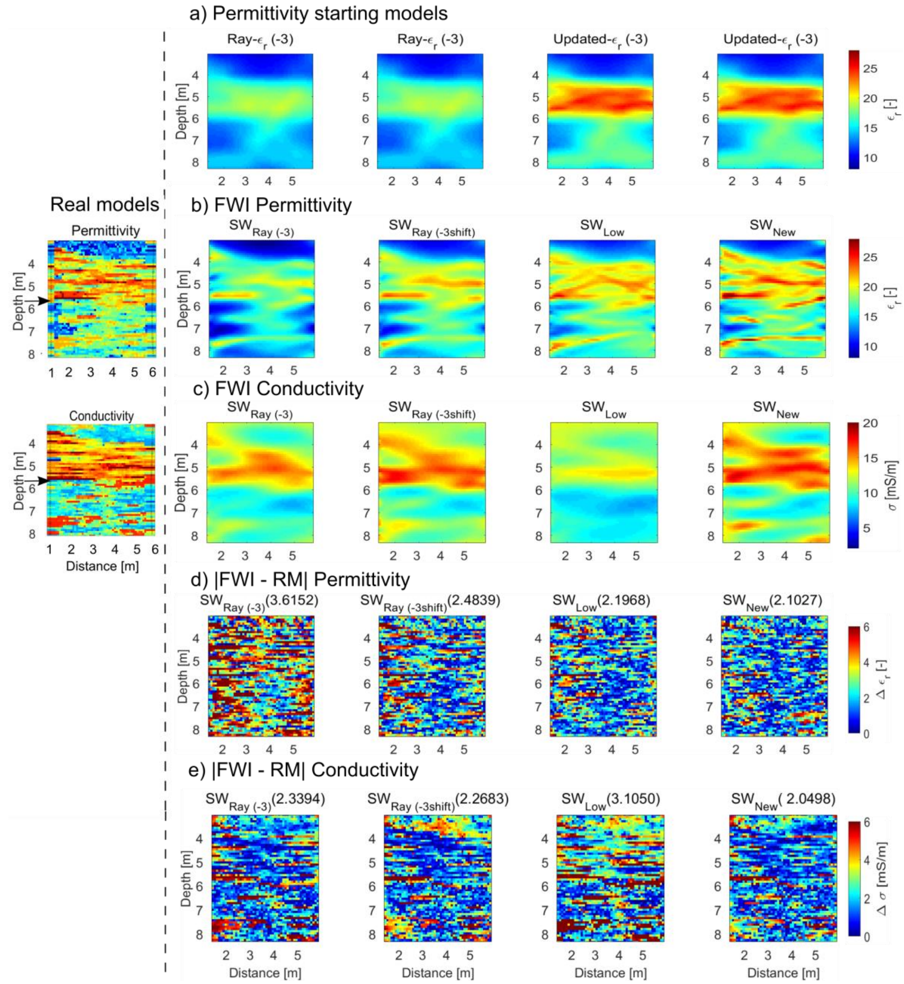


Figure 5.8. Overview of the FWI results using different approaches. Left side real input models for the stochastic simulation. a) Permittivity starting models, and corresponding FWI results for b) ε_r and c) σ for the different FWI approaches. The applied effective source wavelets are named in titles of the Figures. Image plots of the absolute error between the input models and final FWI results for d) ε_r and e) σ for the different approaches. The mean absolute error MAE of the entire 2D domain are shown in parentheses of titles.

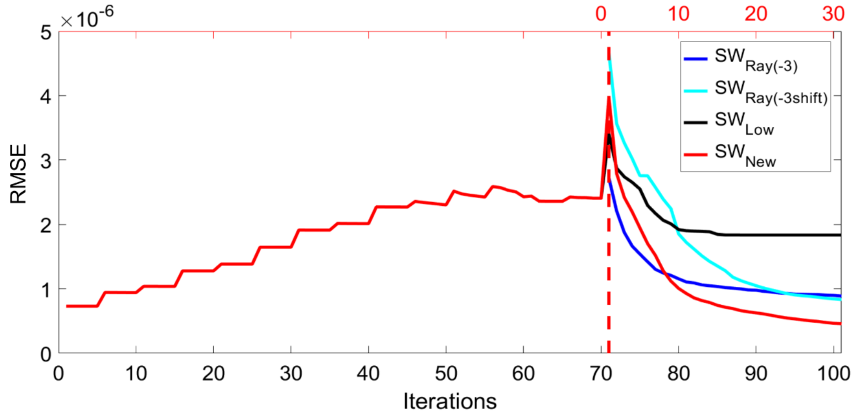


Figure 5.9. RMSE misfit curves of the different FWIs for synthetic case study II. The blue and cyan graphs represent the FWI RMSE convergence behaviors using the source wavelets $SW_{Ray\,(-3)}$ and $SW_{Ray\,(-3\,shift)}$ based on enforcing a smaller ε_r starting model (iterations are from 1 to 31 at red label on the top), respectively. Black and red graphs represent the new PEBDD inversion scheme using the 1st updated and 2nd updated effective source wavelets, respectively. Note that before the dashed lines, the progressively expanded of the bandwidths of source wavelets and observed data are used, while on the right side the FWI with the full bandwidth of all data is performed.

Table 5.2. The mean absolute error MAE between the real input models and the different FWI models for the entire 2D domain (synthetic case study II). RMSE represents residual values between observed and modelled radar traces. Percentages in parentheses indicate the ratio of the other FWI RMSE to the standard FWI RMSE with $SW_{Ray\,(-3)}$, the lower value means the higher improvement efficiency.

Real models	**MAE (ε_r)**	**MAE (σ)**	**RMSE (10^{-7})**
FWI ($SW_{Ray\,(-3)}$)	3.6152	2.3394	9.1843 (100%)
FWI ($SW_{Ray\,(-3shift)}$)	2.4839	2.2683	8.3326 (90.7%)
FWI (SW_{Low})	2.1968	3.1050	18.366 (200.0%)
FWI (SW_{New})	2.1027	2.0498	3.6288 (39.5%)

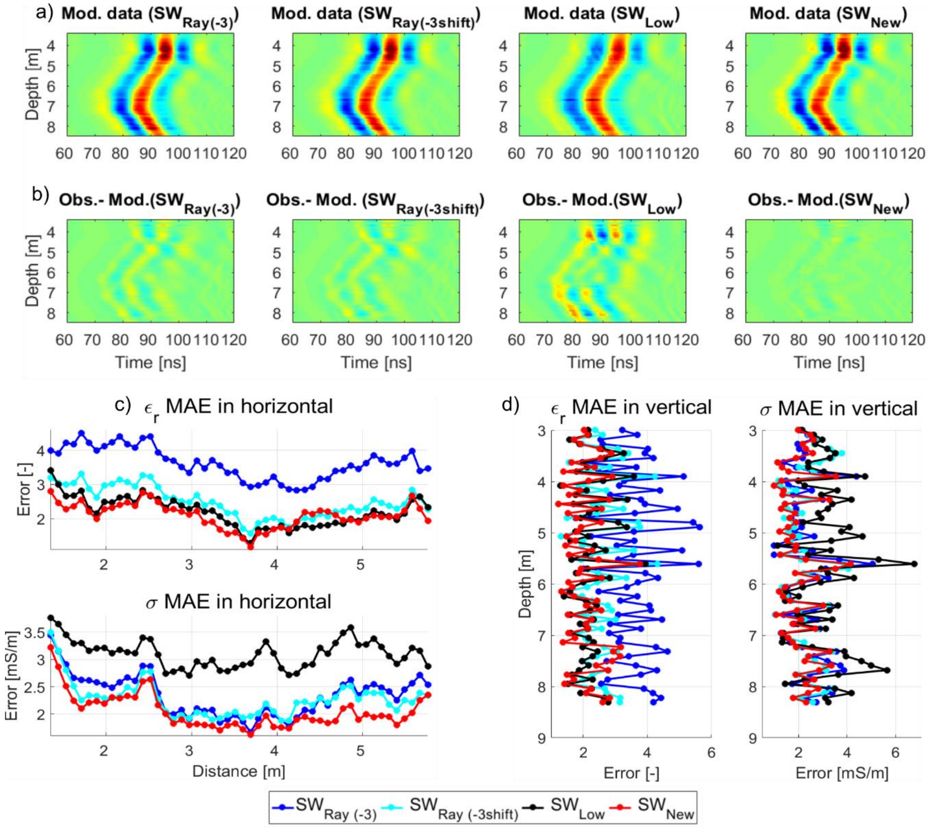


Figure 5.10. a) Modelled data based on the FWI results, and b) differences between the observed and modelled data for the transmitter at 5.69 m depth (see input models and black arrow for the transmitter location in Figure 5.8). Note the amplitudes in a) and b) are normalized to the maxima amplitude of the real observed data (shown range from -7×10^{-1} to 7×10^{-1}). The mean absolute error MAE of the permittivity and conductivity between the input models and the different FWI results are shown along d) horizontal cross-section and e) vertical direction.

5.3 EXPERIMENTAL GPR DATA STUDIES

As final test, we applied the new PEBDD approach to an experimental dataset from the Krauthausen test site in Germany. The Krauthausen study site is located approximately 10 km northwest of the city of Düren in Germany and a detailed description of the site can be found in Vereecken et al. (2000). In last decades, many hydrological and geophysical field techniques have been applied in this site to study the aquifer spatial distribution and flow characteristics, including flowmeter tests (Li et al., 2008), tracer experiments (Vereecken et al., 2000; Vanderborght and Vereecken, 2001), cone penetration tests (Tillmann et al., 2008), and GPR measurements (e.g., Gueting et al., 2015). Here, we utilized the measured crosshole GPR data of Gueting et al. (2015) and Zhou et al. (2020b) using 200 MHz antennae (Sensors & Software) to test the PEBDD FWI scheme for 4 cross-sections between 5 boreholes. The GPR data were acquired using a semi-reciprocal acquisition setup with a transmitter and receiver spacing of 0.5 m and 0.1 m, respectively.

Before the full-waveform inversion of experimental GPR data can be performed, some pre-processing steps are necessary (more details can be found in Klotzsche et al. 2019b). After applying a standard dewow filter and defining the borehole coordinates for the antennae positions, we need to determine an accurate time-zero of the received radar signals. Here, we employed an improved time-zero correction method based on the ZOP-MOG cross-correlation method (Oberröhrmann et al., 2013). The ray-based method was applied for the four GPR cross-sections to obtain the ray-based permittivity starting models (Figure 5.11a). Similar to previous studies, the electric conductivity starting models were chosen to be homogenous with 13 mS/m in the entire domain (not shown). In addition, to reduce the influence of 3D wave propagation phenomena, the 3D data needs to be transformed into 2D according to the approach by Bleistein (1986). The approach is available under assuming both the subsurface and scattering are invariant in one coordinate direction, antennae are line sources, and polarization affects are not properly accounted for (Watson et al., 2016). The effective source wavelet SW_{Ray} (Figure 5.12, dashed curves) is estimated based on the ray-based starting models according to the standard procedure. Using the wavelet SW_{Ray} and the corresponding starting models, the standard FWI is performed (Figure 5.13a and c). Note that in contrast to Gueting et al. (2015) the time-zero correction of the GPR data was improved and corrected (more details in the Corrigendum Klotzsche et al. 2020). Therefore, the final standard FWI results show generally higher permittivity of about 3 - 4 and lower electrical conductivity results, while the structures are similar.

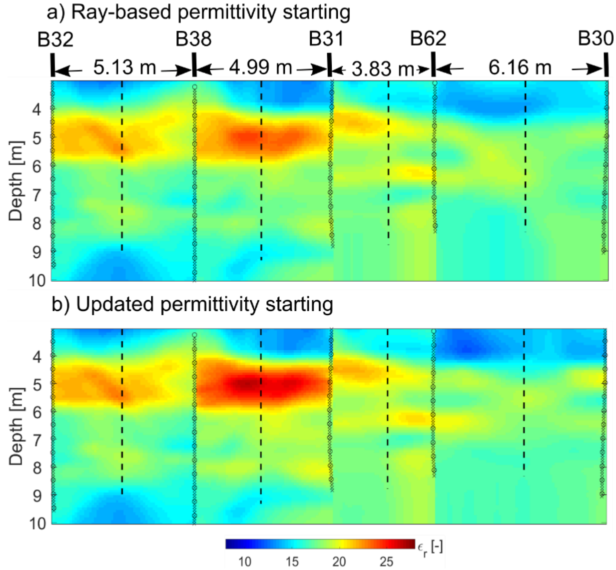


Figure 5.11. a) Ray-based ε_r results and b) FWI ε_r results using the PEBBD approach. Note that results shown in b) are used as starting models for the full bandwidth inversion of the experimental Krauthausen data. And a homogeneous conductivity starting model of 13 mS/m was used for all inversions.

Similar to the previous synthetic cases studies, we filtered the effective source wavelets and corresponding measured GPR data by using the PEBDD scheme. The smallest filtered sub-data bandwidths are 12-16 MHz for four cross-sections. The bandpass widths are increased every 5 iterations by 4 MHz until the maximum bandwidth sub-data is reached. Note that we selected different highest cut corner frequencies according to center frequencies of the different cross-section data sets. For the cross-sections B38-31 and B62-30 we used 68 MHz and for the datasets of B32-38 and B31-62 we applied 72 MHz. The final FWI ε_r results using the maximum bandwidth sub-data (Figure 5.11b) and the homogenous σ model were considered as the updated starting models and applied in a following FWI under the full bandwidth data. Following the flowchart in Figure 5.1b, the effective source wavelet SW_{Low} is updated using the new starting models and the effective source wavelet SW_{Ray}. The updated source wavelets SW_{Low} together with the new starting models are used to calculate the FWI results (not shown). Finally the second-updated effective source wavelets SW_{New} are derived using the new FWI ε_r results and SW_{Low} in the deconvolution approach (Figure 5.12, solid curves). The corresponding updated final FWI results (Figure 5.13b and d) show very similar structures as the standard FWI results, but the structures are more continues closer to the boreholes. The RMSE values between the observed and modelled data based on selected different FWI models show that the updated FWI results are smaller than of the standard FWI (Table 5.3) indicating improved results of the updated FWI.

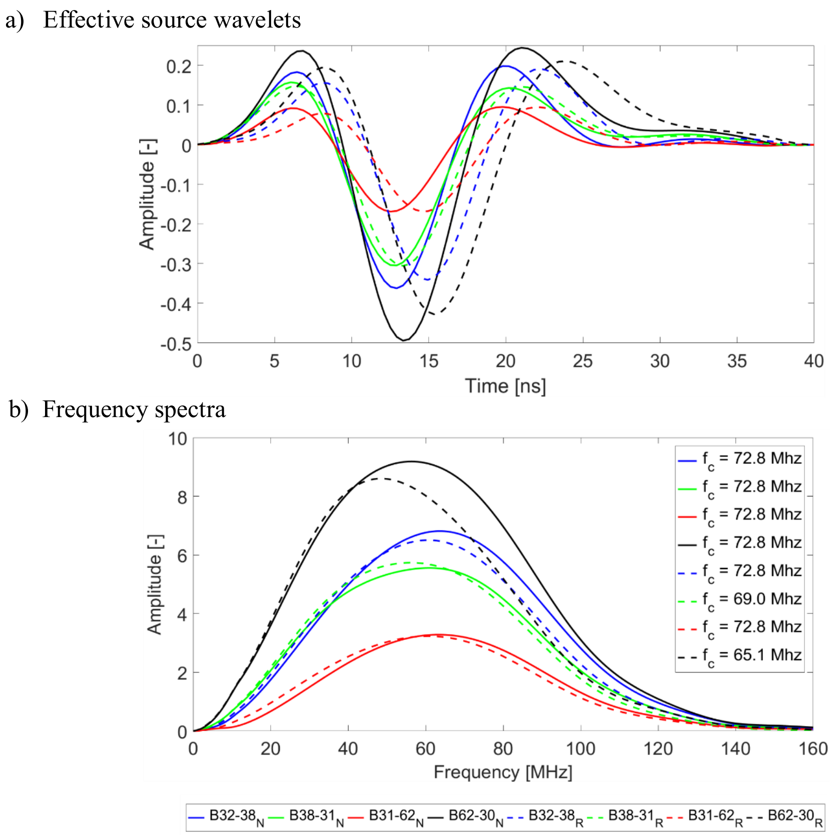


Figure 5.12. a) Comparisons of the effective source wavelet estimation based on ray-based models (dashed graphs) and the updated source wavelets of the PEBBD scheme (solid graphs) for the four cross-sections in time domain for the experimental data of the Krauthausen test site. b) Corresponding frequency spectra of standard and updated effective source wavelets. The legend values in b) indicate these effective center frequencies for different cross-sections.

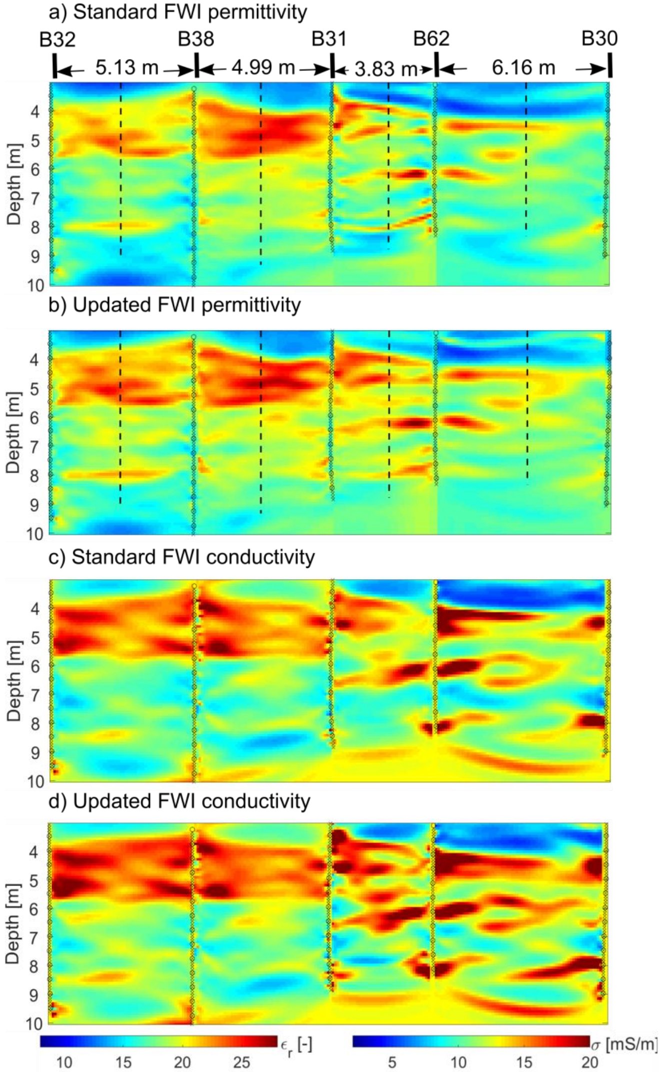


Figure 5.13. Standard FWI results of a) ε_r and c) σ of the four cross-sections. Updated FWI results of b) ε_r and d) σ based on the new ε_r starting models (Figure 5.11b) and corresponding updated new effective source wavelets (solid graphs in Figure 5.12). Dashed lines mark the CPT data locations between the boreholes.

To validate the different FWI results, we compared them with porosity data based on cone-penetrations test (CPT) measurements (Zhou et al., 2020b). Note that the CPT data was acquired mostly in the center of the cross-sections. First of all, we converted the porosity values $\emptyset$ of the CPT to the relatively permittivity using the three phase complex refractive index model (CRIM) in saturated zone (e.g., Gueting et al., 2015),

$$\varepsilon_r = \left[\emptyset \times \left(\sqrt{\varepsilon_f} - \sqrt{\varepsilon_s}\right) + \sqrt{\varepsilon_s}\right]^2 . \tag{5.2}$$

Thereby, ε_s and ε_f are the permittivity of the solid and the fluid, respectively. Considering a water temperature of 10℃ we use a ε_f of 84 and we apply a ε_s value with 4.5 based on literature values of quartz (e.g., Eisenberg and Kauzmann, 2005; Carmichael, 2017). We compared the two FWI ε_r results with the CPT data along the vertical profile locations (dashed lines in Figure 5.13). Comparisons of these 1D vertical results indicate that the updated FWI ε_r results (Figure 5.14, red lines) are closer to the CPT data (Figure 5.14, black lines) than the standard FWI ε_r results (Figure 5.14, blue lines). Quantitative comparisons are described in Table 5.3, where we computed RMSE and correlation coefficient *R* with larger than 0.8 between the CPT and two FWI results. In general, the updated FWI results with the PEBDD scheme are better than the standard FWI results indicating that the new PEBDD approach improved the experimental GPR FWI results. Note that the RMSE value (for 1D CPT) of the updated FWI for boreholes 38-31 is larger than the standard FWI (green value in Table 5.3). A possible explanation could be that the standard FWI was already very good because many studies have been already performed with this data set and many optimizations have been performed with the standard approach. For the other data sets less intense tests have been applied for the standard FWI.

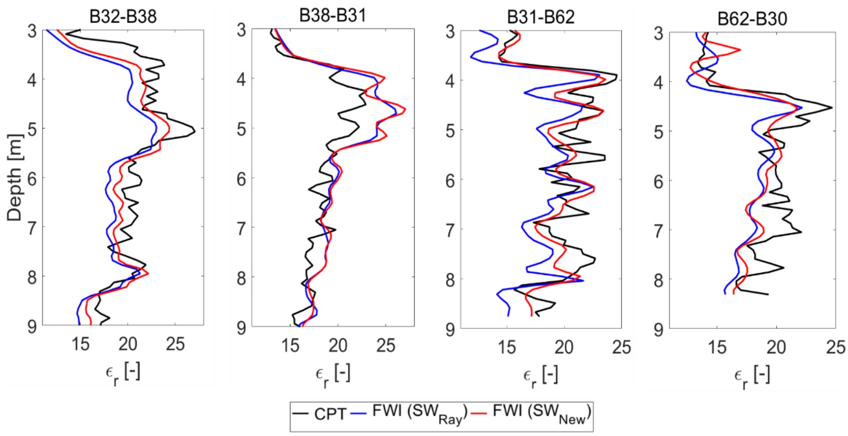


Figure 5.14. Permittivity comparisons derived from cone-penetration test (CPT) data (black), standard FWI ε_r results (blue) and the updated new FWI ε_r (red) results along CPT profiles between boreholes (see Figure 5.13 for locations).

Table 5.3. RMSE (10^{-7}) represents residual values between observed and modelled radar traces. Percentages in parentheses indicate the ratio of the new FWI RMSE to the standard FWI RMSE with SW_{Ray}, the lower value means the higher improvement efficiency. Correlation coefficient R and RMSE between 1D ε_r FWI and the CPT data (CPT porosity has been converted into ε_r) represent the reliability of the FWI results, the larger R and the lower RMSE values mean the FWI results are closer to the real values.

	32-38 (5.13 m)	**38-31 (4.99 m)**	**31-62 (3.83 m)**	**62-30 (6.16 m)**
Standard FWI RMSE (10^{-7}) (SW_{Ray})	8.5055	9.4156	9.1763	7.0538
New FWI RMSE (10^{-7}) (SW_{New})	6.8210	7.1789	7.8503	5.4136
	(80.2%)	(76.2%)	(85.5%)	(76.7%)
***R* (Sta-FWI:CPT)**	0.8105	0.9032	0.8301	0.8830
***R* (New-FWI:CPT)**	0.8409	0.8973	0.8408	0.8723
RMSE (Sta-FWI:CPT)	2.6807	1.9307	2.5131	2.1026
RMSE (New-FWI:CPT)	1.9054	2.1503	1.5224	1.7333

5.4 CONCLUSIONS

We have presented a new crosshole GPR FWI approach using progressively expanded bandwidth of the measured and the modelled data (PEBDD) with frequency bandpass filters to improve the reconstruction of the subsurface properties. The new PEBDD scheme was applied to two synthetic case studies and an experimental data set from the Krauthausen test site in Germany. Thereby, we have demonstrated that the new scheme is able to improve the FWI results by taming the problem of the inversion to be easily trapped in local minima caused by non-linear and ill-posed problems. The introduced PEBDD scheme applies designed bandpass filters to the effective source wavelet and the observed data. Thereby, the different sub-data sets with various bandwidths are progressively expanded until the center frequency of the data is reached. The FWI results of this PEBDD provide an updated ε_r starting model, which is used to update an effective source wavelet in the deconvolution approach and to perform the FWI with full bandwidth data. In further, we update the effective source wavelet SW_{New}. Considering this updated effective source wavelet and the starting models based on the PEBDD scheme, the resulting FWI results show a better reconstruction of the medium properties and the final root-mean-square values are decreased in contrast to the standard FWI.

Two synthetic cases have proven that the PEBDD scheme is robust and can also reconstruct the permittivity and electrical conductivity results for starting models that are more than half a wavelength away from the measured data. To fulfill the starting models criteria that the modelled data need to be within half a wavelength of the measured data, normally need an accurate processing of the data and a detailed understanding of trained user. Problems arising from erroneous starting models are sometimes hard to notice and therefore an improved approach such as the PEBDD can significantly improve the applicability of the FWI. For the experimental GPR data, we have computed the standard and the updated FWI results using the PEBDD scheme for four cross-sections. We compared the different FWI results with CPT data from the center of the cross-sections, where the updated FWI results provided the best correlation to the CPT data. Finally, we concluded that the progressively expanded bandwidth approach can not only improve the FWI results, but also that it is able to retrieve more accurate FWI results also for starting models more than a half wavelength away from the measured data. Therefore, a lot of previous detailed work to construct good starting models for experimental GPR data can be reduced and the application to field data will be much easier. In future work, we will also investigate the possibility to apply the new PEBDD approach to other frequencies such as for 100 MHz antennae and to different test sites.

Chapter 6

Conclusions and Outlook

In this chapter, we summarized the main findings from each chapter and presented some general conclusions for the application of crosshole GPR full-waveform inversion with the improved FWI reconstruction. Some future research directions and ideas are shown in the outlook.

6.1 CONCLUSIONS

During this thesis, our main goal was to improve the full-waveform inversion for crosshole GPR data. For this, we focused on two aspects: the relative permittivity ε_r starting model and the effective source wavelet. The most applied approach is to generate a starting ε_r model based on the ray-based inversions. In special cases, the ray-based inaccurate starting ε_r model possible causes the FWI trapped into a local minimum. To improve starting models, an updated starting ε_r model using a low iteration FWI result that is confirmed by the amplitude analysis results is considered in the FWI. Furthermore, the approach to progressively expand the bandwidths of both modeled and observed data (PEBDD) is to enhance the starting ε_r model. The second aspect is to improve the effective source wavelet estimation. The conventional effective source wavelet has a small bandwidth and is estimated and corrected using ray-based starting models. Expanding the bandwidth of an effective source wavelet is an approach of improving the FWI resolution. In this thesis, we applied Cone Penetration Test data (CPT), and the PEBDD scheme to enhance the effective source wavelet in the forward modeling and the inversion.

In Chapter 3, experimental crosshole GPR data at an alluvial aquifer located near the Meuse River in Belgium was analyzed by amplitudes analysis. The results displayed the approximate locations of waveguides with high porosity that were indicated by the energy maxima of GPR trace profiles (WGT I). Meanwhile, local minima of GPR trace energy profiles were either caused by waveguide boundaries due to EM wave total reflection or by the high conductivity of clay (WGT II) that causes an increased attenuation. For the FWI we updated the ε_r starting model using a low iteration (10) FWI result, which included the waveguide structures that corresponded to the results of the amplitude analysis. Meanwhile, we kept the traditional σ starting model with a homogeneous value (13 mS/m). By using the deconvolution method, an updated effective source wavelet, which contains waveguides information, was determined. Finally, the FWI results using the updated ε_r starting model and the new effective source wavelet presented the waveguide structures with a higher resolution. We inverted nine 2D intersecting GPR planes individually, and combined them to characterize the alluvial aquifer in 3D domain. By computing the correlation coefficient R for intersecting acquisition planes, we showed that the updated FWI provides results that are more consistent. The 3D FWI results were able to show the spread and location of the high-permittivity layers at depths from 7.5 m to 9.5 m in all nine cross-sections (WGT I). Further, continuous higher conductivity zones above 4 m were detected in all cross-sections that aligned with the WGT II features. Compared to the previous heat tracer test by ERT tomography, the updated FWI results provided better consistency with a high hydraulic conductivity zone with high porosity and gave some reasonable explanations for the plume splitting that occurred in the heat tracer test.

In Chapter 4, we improved the ε_r FWI results by using 1D vertical CPT data. First, we compared filtered CPT data and filtered ε_r FWI results, which were filtered by a lowpass filter, and found a satisfying consistency. Meanwhile, some differences between two filtered results implied the possibility to improve the FWI results resolution. Therefore, we constructed a filter based on the two filtered 1D results, then we used the filter to amplify the ε_r FWI results in the whole 2D domain. Further, we updated the effective source wavelet by replacing the ray-based ε_r results with the 2D wavenumber-amplified FWI ε_r (WA-FWI) results. The updated source wavelet had a broader bandwidth and was used to perform the new FWI combining with the traditional starting models. In most cases, we needed to repeat the process of effective source wavelet updating and the FWI conduction by using the updated ε_r FWI results to achieve an even better fit between measured and modeled data. The new scheme was verified using a synthetic stochastic model of the Krauthausen test site, which was constructed based on CPT parameters. In addition, we applied our new approach to field experimental GPR data of five pair boreholes. Similar to the synthetic results, the updated effective source wavelets showed a broader bandwidth than traditional source wavelets. Finally, we compared the CPT data and the updated ε_r FWI results, both of them are full wavenumber information, and were able to obtain smaller RMSE values. These results demonstrated the improvement of the updated ε_r FWI results under neglecting conductivity.

In Chapter 5, inspired by Meles et al. (2011), we focused on the non-linearity problem of full-waveform inversion for complicated synthetic models and experimental crosshole GPR data. Different from the approach of Meles et al. (2011), we used an approach named progressive expansion bandwidth of the modeled data and observed data (PEBDD) to construct the new starting models and a new effective source wavelet for the 2D full-waveform time-domain inversion. The approach can be divided into two processes. First, we utilized designed bandpass filters to divide the traditional effective source wavelet and observed data. Thereby, we obtained different sub-source wavelets and sub-observed data with various bandwidths. The FWI was performed starting from the smallest bandwidth data. With the sub-data bandwidth increasing until the selected maximum bandwidth, we computed various FWI results (ε_r and σ) when we considered the previous bandwidth FWI results as the next starting models. Note starting models are updated for both ε_r and σ under various sub-data in the first process. The second process is the performance of the FWI using the full frequency data (FBD) information. Using the deconvolution method, a new effective source wavelet was generated by being updated twice. The low and high frequency parts of the effective source wavelet were corrected with the first and second updated, respectively. Note that the starting σ model is always a homogenous value with 13 mS/m in the second process when we correct the source wavelet and perform the FWI. Two synthetic case studies indicate that the progressively expanded bandwidth scheme can not only improve the FWI results, but also that it is able to retrieve accurate FWI results also for starting models more than a half-wavelength away from the measured data. Therefore, a lot of previous detailed work to construct good starting models for experimental GPR data can be reduced and the application to field data will be much easier. In addition, we applied the PEBDD scheme to four experimental GPR cross-sections of the Krauthausen test site. The comparison of RMSE and R as the traditional FWI results, the updated FWI results, and the CPT data showed that the updated FWI results with the PEBDD scheme achieve a better fit with the real data.

In conclusion, this thesis demonstrates that crosshole GPR data with the full-waveform inversion method is an effective tool to construct the subsurface aquifer with sub-wavelength resolution for both permittivity and conductivity. And the starting models and the effective source wavelet are important for the final FWI resolution. Combining the FWI and the amplitude analysis results, we distinguish two types of waveguide structures. Through using the CPT data, the updated effective source wavelet with a larger bandwidth enable enhancing the FWI results. This scheme possible extends to seismic domain in future. To tame the non-linearity issue of the FWI algorithm, the PEBDD scheme is applied for complicated synthetic models and experimental GPR data. To compare the standard FWI, the updated FWI using the CPT data, and the updated FWI with the PEBDD scheme, we draw three different FWI results for boreholes 38-31 in Figure 6.1. In which, the standard FWI with final RMSE value after 22 iterations is 9.42×10^{-7}, while the updated FWI using the CPT data with the smallest final RMSE value after 30 iterations is 6.49×10^{-7}. The FWI results using the PEBDD scheme show that final RMSE value after 22 iterations is 7.18×10^{-7}. The final goal is to improve our ability to detect small-scale structures in the subsurface, in further to better forecast and protect underground water resources.

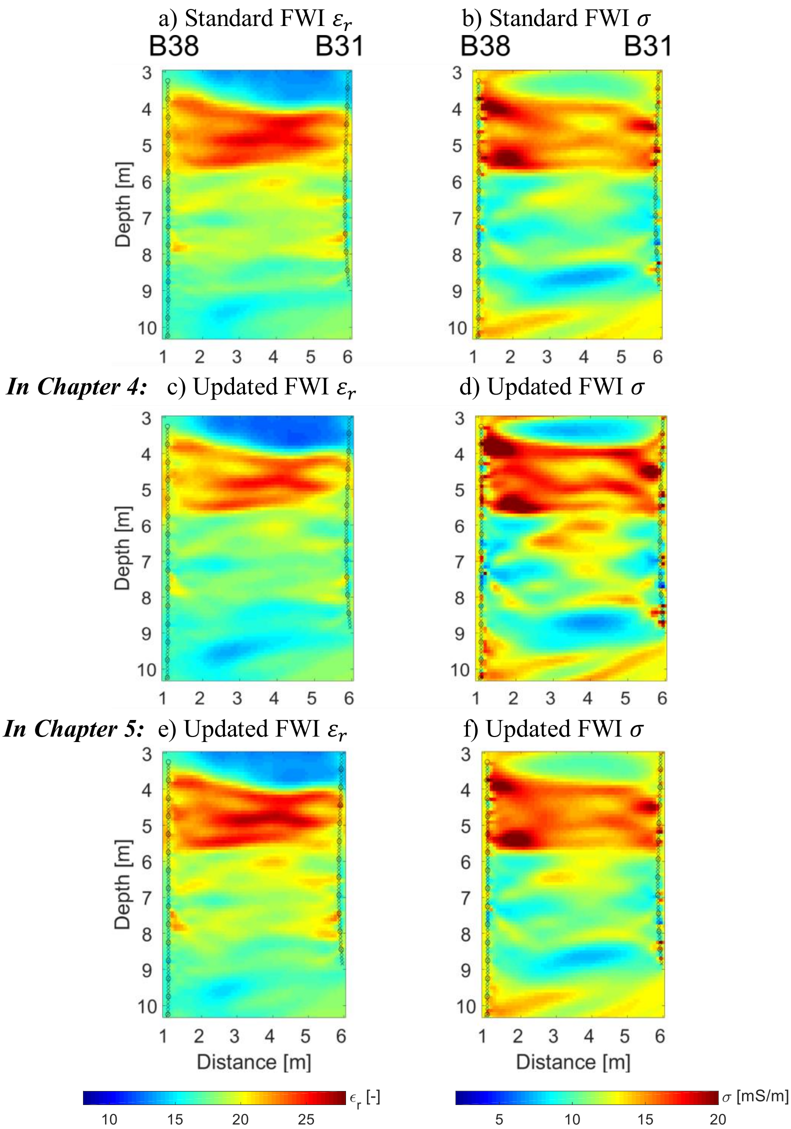


Figure 6.1. Standard FWI results of a) ε_r and b) σ for the cross-section of B38-31. Updated FWI results of c) ε_r and d) σ based on the amplified approach according to the CPT data in Chapter 4. e) and f) represent the updated FWI ε_r and σ results using the PEBDD scheme in Chapter 5, respectively.

6.2 OUTLOOK

Following with the aforementioned improved approach, we will also investigate the possibility to apply the new PEBDD approach to other frequencies such as for 100 MHz antennae and to different test sites in the future work. In addition, we should improve the starting models by using other approaches, such as CPT, ERT, and seismic data. Furthermore, the previous work on the crosshole GPR data FWI inspires two ideas for the future work. First of all, recovering low-wavenumber information of the starting models based on high-frequency observed data by applying the approach of angle difference identity for cosine (ADIC) (Wang et al., 2019). The second one is to improve the computation of gradient directions in the FWI process triggered by the approach of seismic staining algorithm (Chen and Jia, 2014; Hu et al., 2016; Li and Jia, 2017).

6.2.1 Improvement of starting models by applying ADIC

The lack of low-frequency information in the crosshole GPR data causes the GPR FWI to suffer from local minima convergence and serious nonlinearity. When the saturated aquifers include high contrast layers, missing low frequency information especially affects the FWI results due to the ray-based inversion starting models differ a lot from the true values. To enhance the low-frequency information, Wang et al. (2019) applied the approach of angle difference identity for cosine in the seismic data. By building an internal connection between high- and low-frequency signals, a plausible recovery of the low-wavenumber velocity can be obtained from the high-frequency information (Wang et al., 2019). For crosshole GPR data, because the interfering noises in the field measurement process, the low frequency information in measured data is easy to be contaminated. In addition, the electromagnetic pluses are common high frequency, which causes more pronounced phenomena of high-frequency suppression of low-frequency.

To recovery the low-wavenumber permittivity starting model from crosshole GPR observed data, we should invert the low-wavenumber model with a new shifted effective source wavelet. The amplitude spectrum of the new source wavelet should be moved on the left some suited frequencies to improve the low frequencies amplitude values. The results of an updated FWI by using the shifted source wavelet provide the full new starting models with enough low-wavenumber information. In further, we improve our ability to detect small high contrast layers in the saturated aquifers.

6.2.2 Improvement of FWI results by using staining algorithm

The gradient of the misfit function for the FWI for crosshole GPR data is estimated by correlating the forward propagating synthetic wavefield and the backwards propagating residual wavefield (Taratola, 1984). While the forward modeling data and the back propagating receiver data are correlated in the reverse time migration (RTM) (Hu et al., 2016). Therefore, there are many similarities when the back propagating receiver data are replaced by the backwards propagating residual wavefield. In addition, the staining algorithm has proven to practically effective and enhance the RTM images in local areas with amplitude-preserved. In other words, we are able to improve the RTM images in some blind zones without illuminations. Therefore, we would suggest for the future to use the staining algorithm to improve local areas gradients and thereby enhance FWI results. Considering the limitation of 2D time-domain FWI of crosshole GPR data, which can only provide rough FWI results near boreholes, top, and bottom areas, we possible acquire high inversion resolution in these zones using the staining algorithm. In addition, the staining algorithm with amplitude-preserved probably improve the FWI conductivity results.

Appendix A

Corrigendum to “Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests” [1]

The authors regret that an error has occurred and the following corrections need to be recorded regarding the above cited paper. After reanalyzing the crosshole GPR data, we found an error in the automatic picking routine for estimating the time zero of the GPR data. After correction, the GPR data was shifted in time affecting the calculated permittivities ε_r and electrical conductivities σ (Original Figure 3 and 4). Using the corrected time zero, the permittivity and conductivity results of the full-waveform inversion were updated. The comparison between the original and correct tomograms now show that the permittivity and electrical conductivity results are approximately 4 higher and 10 mS/m lower, respectively (see Figure A.1). Using this correction the full-waveform inversion results are in a better agreement with the CPT data (Figure A.2). It is Interesting to note that the constant shift of -0.08 that was applied previously to align the porosity CPT data with the FWI results is not necessary anymore. The porosity values based on the updated FWI results are now in a very good agreement with the original values of Tillman et al. (2008) indicated by a correlation coefficient of 0.91, which was before 0.80 (Figure A.2a). Furthermore, the updated electrical conductivity FWI results are closer to the electrical conductivity results based on the CPT data (Figure A.2b). We expect only minor changes in the results using the cluster analysis to derive the facies of the aquifer indicating that the main conclusions of the paper remain valid.

[1]adapted from Klotzsche, A., Z. Zhou, J. Schmäck, J. van der Kruk, J. Vanderborght, and H. Vereecken, 2020, Corrigendum to “Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests”, Journal of Hydrology, submitted.

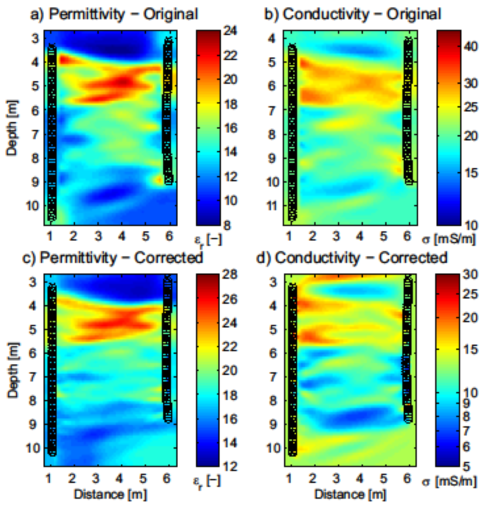


Figure A.1. Comparison for the exemplary transect B38-B31 of the original a) permittivity and b) electrical conductivity full-waveform inversion results, and, the corrected c) permittivity and d) electrical conductivity results using the corrected time zero estimation. Please note the different color scales of the tomograms.

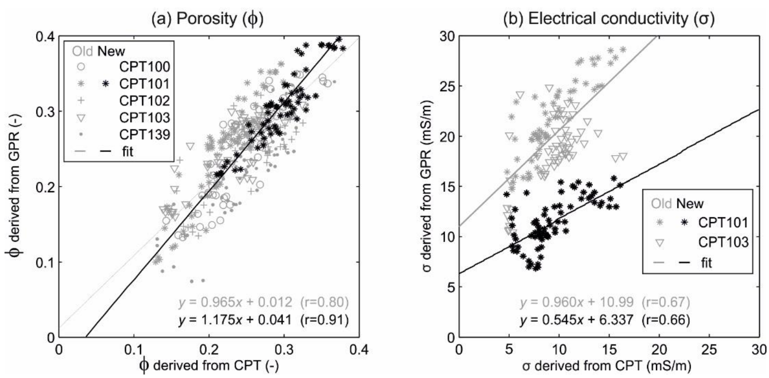


Figure A.2. Comparison of the old (grey) and corrected (black) cross-plots between the FWI results and CPT data. Cross-plots of a) porosities and b) electrical conductivities derived from CPT and GPR data. Results based on the corrected FWI are shown in black for the exemplary transect B38-B31 for which co-located porosity and electrical conductivity data of the CPT 101 exist. Data based on Gueting et al. (2015) presented in grey for the a) five and b) two profiles and corresponding co-located CPT porosity and electrical conductivity data, respectively. Regression lines through all data points are depicted in grey and black for old and corrected data, respectively. The corresponding straight-line equations are given at the bottom of the cross plot, r is the correlation coefficient.

Appendix B

Improved resolution of ground penetrating radar full-waveform inversion by using cone penetration test data: A synthetic study[1]

Crosshole ground penetrating radar (GPR) full-waveform inversion (FWI) has shown to be a powerful tool providing decimeter-scale high resolution images of the subsurface. To enhance this resolution even more, we present here a new approach that uses the cone penetration test (CPT) data acquired between the two boreholes to improve the bandwidth of the effective source wavelet by using wavenumber-amplified FWI models. A synthetic model is generated using a stochastic simulation based on measured parameters at Krauthausen. After generating forward modeling data (called true data) based on the synthetic models, white noise is added in true GPR data. The standard full-waveform inversion using a ray-based start model and the updated results based on the CPT data are compared and show that the new effective source wavelet based on CPT data can improve the resolution of the FWI results.

B.1 INTRODUCTION

Crosshole GPR has often been used in hydrogeological investigations to obtain cross-sectional information about porosity, soil water content and connectivity of structures (Klotzsche et al., 2018). Compared with conventional crosshole GPR tomographic inversions based on geometrical ray theory (e.g., Dafflon et al., 2011 and 2012), full-waveform inversion (FWI), which uses the entire waveform, provides higher resolution images of the subsurface properties such as relative permittivity ε_r and electrical conductivity σ (Ernst et al., 2007a). In contrast to crosshole GPR, CPT measurements provide accurate and high resolution porosity results along the 1D vertical probing location (Tillmann et al., 2008). Combining the results and resolution of both methods can enhance the understanding of the investigated aquifer.

[1]adapted from Zhou, Z., Klotzsche, A., Güting, N., Haruzi, P., Vereecken, H. and Kruk, J.V.D., 2019. Improved resolution of ground penetrating radar full-waveform inversion by using cone penetration test data: A synthetic study. In SEG Technical Program Expanded Abstracts 2019 (pp. 2898-2902). Society of Exploration Geophysicists. https://doi.org/10.1190/segam2019-3215765.1

Here, we present an approach that combines the standard FWI with the CPT data to enhance the resolution of the FWI results. First, we evaluate the reliability of the obtained FWI images by comparing them with CPT data acquired between the boreholes. The CPT data provides water content values or for a fully saturated porous media the porosity. The FWI derived permittivity values can be transformed into porosity using the Complex Refractive Index Model (CRIM) (e.g., Gueting et al., 2015). Due to the differences in resolution, the overlapping wavenumber information of the CPT and GPR FWI data are compared (see Yang et al., 2013 for more details) along the vertical locations where the CPT data coincides with the GPR FWI results. Similar to the approach of Yang et al. (2013), we compute a wavenumber filter based on the CPT data and apply it to the FWI results in wavenumber domain. Note that the filter was only computed in the lower vertical wavenumbers domain and higher spatial frequency information is neglected. In the process of computing the filter, a smoothing function was used to smooth out any fluctuating amplitudes.

After applying the wavenumber domain filter to the FWI results to obtain a wavenumber-amplified FWI (WA-FWI) model that includes a broad bandwidth, the effective source wavelet is updated such that only consistent high frequency information remains. To quantitatively compare the conventional FWI with the new FWI results using the updated effective source wavelet, we compute the root mean squared error (RMSE) and the correlation coefficient (R) along each vertical profile between the boreholes. We test this new approach on GPR data modeled for a stochastic representation of the Krauthausen test site in Germany.

B.2 STOCHASTIC SIMULATION MODEL STUDY

We constructed realistic models (Figure B.1a and b) of relativity permittivity ε_r and electrical conductivity σ using a stochastic simulation (Sequential Gaussian Simulation) based on Haruzi et al. (2018). A semi-reciprocal acquisition setup was used with transmitter and receiver spacing of 0.5 m and 0.1 m, respectively. In Figure B.1, black circles and crosses show the exact transmitter and receiver positions within the boreholes (see Figure B.1a and b).

Realistic synthetic GPR data are generated by performing the finite-difference time-domain (FDTD) modeling using the stochastic models and a real wavelet. The real wavelet is similar as the wavelet of GPR data measured by 200 MHz antennas at Krauthausen, which has an approximate center frequency of 70 MHz (not shown). Note that we added *10%* white noise to the true synthetic GPR data (see Figure B.1g).

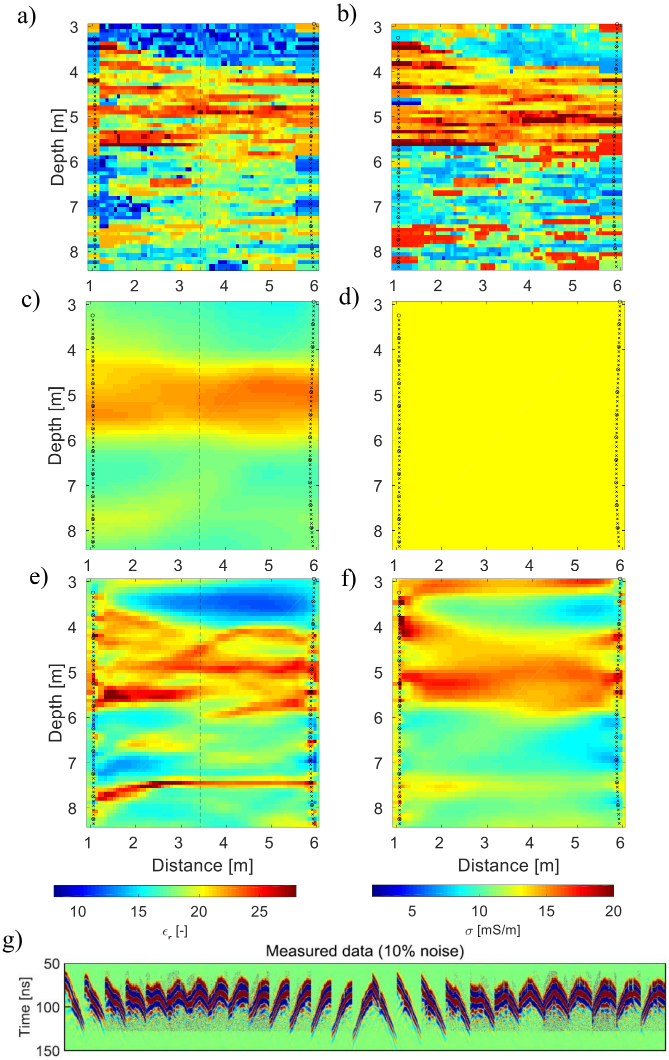


Figure B.1. a) And b) show the input models based on stochastic simulation for ε_r and σ, respectively, to generate the true synthetic GPR data. c) And d) show the ray-based results for ε_r and a homogeneous σ model. Here, the ray-based results were inverted based on true GPR data including white noise. e) And f) show the FWI results (iterations=30) based on ray-based models and true GPR data with noise. Dashed lines indicate the location of the CPT data, which will be used to compute the filter and to amplify wavenumber of FWI results. g) Shows the true synthetic GPR data containing *10%* white noise.

B.3 FULL-WAVEFORM INVERSION RESULTS

The FWI for crosshole GPR data is based on a 2D FDTD solution of Maxwell's equations (e.g., Klotzsche et al., 2012, 2013). The ray-based method has been used to obtain a starting model for ε_r (Figure B.1c) for the FWI, whereas for the conductivity we choose a homogenous model with 13 mS/m (Figure B.1d, similar to Gueting et al., 2015). During the inversion process, the misfit function is computed by subtracting the synthetic data from the observed data. The gradient is determined by using a zero-lag cross-correlation of the forward propagated synthetic wavefield with the backward propagated residual wavefield. To minimize inversion artefacts in the vicinity of the boreholes, the approach of 3 cells (each cell: 9 cm) with a gradient preconditioning (van der Kruk et al., 2015) is employed in this paper.

Before performing the FWI, an effective wavelet is estimated. Using an approximated wavelet $\hat{s}_k(f)$ obtained by analyzing horizontal rays (Klotzsche et al., 2010); synthetic data $\hat{E}^{syn}(f)$ is calculated using FDTD. Then the Greens function $\hat{G}(f)$ is obtained using Equation B.1, followed by the calculation of an updated effective wavelet $\hat{s}_{k+1}(f)$, SW_{Ray}, as shown by Equation B.2.

$$\hat{G}(f) = \hat{E}^{syn}(f)[\hat{s}_k(f) + \eta_D]^{-1}, \tag{B.1}$$

$$\hat{s}_{k+1}(f) = \left[\hat{G}(f) + \eta_I\right]^{-1}\hat{E}^{obs}(f), \tag{B.2}$$

where η_D and η_I are prewhitening factors that are applied to stabilize the solution and avoid dividing by zero, and ^ indicates frequency domain.

The conventional FWI results when using the ray-based inversion results as start model are shown in Figure B.1e, and f. The reconstructed ε_r and σ images have a higher resolution than the ray-based results but have still a reduced resolution compared to the true data shown in Figure B.1a, and b.

B.4 WAVENUMBER INFORMATION IN RAY-BASED, FWI AND CPT

Along the dashed lines in Figure B.1a, c and e, porosity data are calculated as if a virtual CPT measurement was carried out in between the two boreholes. To investigate the spatial bandwidth present in the true data (Sto-CPT), the ray-based and FWI reconstructed data, a 1D FFT was carried out. Figure B.2a shows the spatial wavenumber information for all three cases. It can be clearly seen that the Sto-CPT data contains the largest bandwidth, whereas the FWI data has a reduced bandwidth, and the ray-based data has the lowest bandwidth.

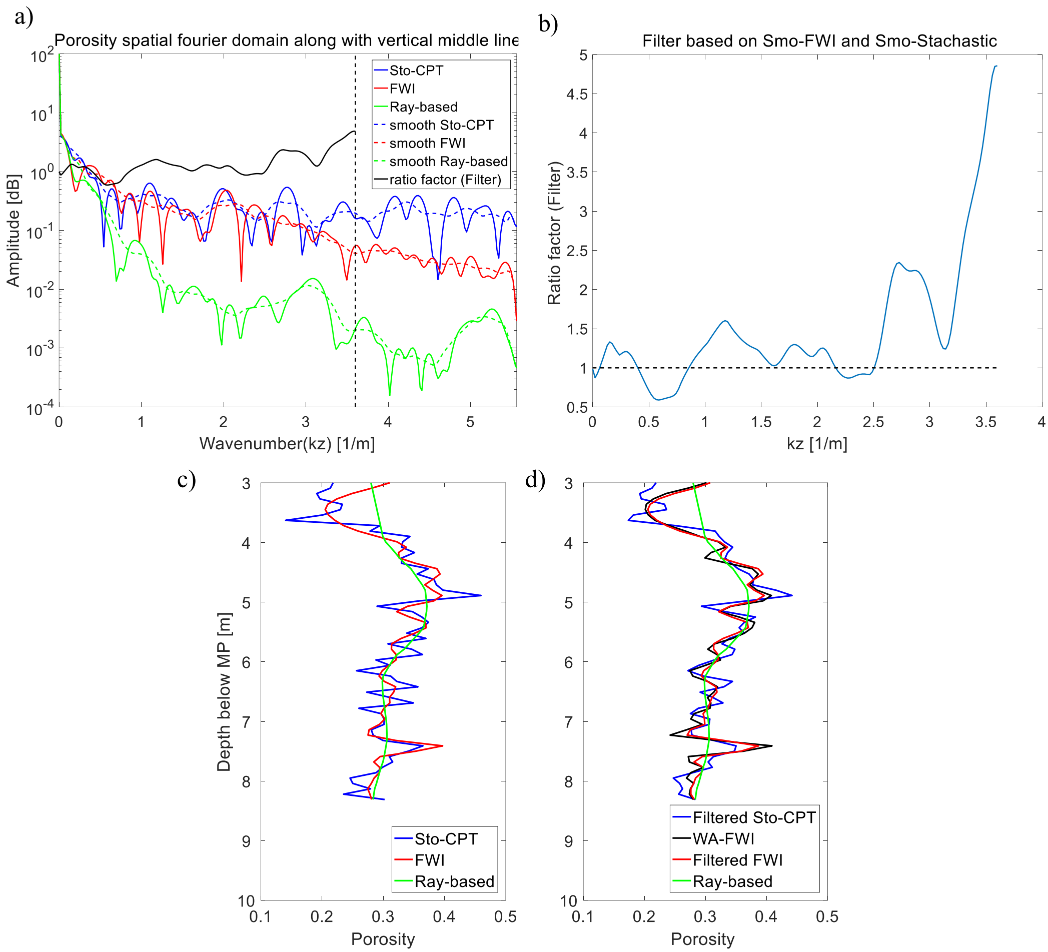


Figure B.2. a) Comparison of the wavenumber spectra of Sto-CPT data (blue), FWI results (red) and ray-based (green) in wavenumber domain. Filter is indicated by black solid line. b) Filter based on smooth FWI results and smoothed Sto-CPT. c) Porosity comparison for Sto-CPT, FWI, and ray-based results along the vertical dashed lines in Figure B.1a, e and c, respectively. d) Porosity comparison between filtered Sto-CPT, wavenumber-amplified FWI, filtered FWI and ray-based results up to a maximum $K_Z = 3.60\ \mathrm{m}^{-1}$.

B.5 WAVENUMBER FILTER ESTIMATION USING CPT DATA

To use the higher spatial bandwidth present in the true data (Sto-CPT), a 1D FFT filter function is introduced that amplifies the FWI reconstructed data up to a threshold wavenumber, that was estimated as K_Z=3.60 m^{-1}(the vertical dashed line in Figure B.2a). First, a smoothing function with a span of 21 is used to smooth the wavenumber data after which we compute the ratio factor that is used as a filter to increase the spatial bandwidth up to a selected maximum wavenumber. This filter in frequency-wavenumber domain is implemented as follows:

$$\begin{cases} SA(1) = A(1)\,, & \text{(B.3)} \\ SA(x) = smooth(A(x), span)\ \ (1 < x)\,, & \text{(B.4)} \end{cases}$$

$$Filter = \frac{SA_{CPT}}{SA_{FWI}}\,, \qquad \text{(B.5)}$$

where A and SA represent amplitude and smooth amplitude values in wavenumber domain, respectively, and x means wavenumber samples up to the maximum wavenumber. The filter is described by Equation B.5. Figure B.2b shows the filter in Cartesian coordinate system, which keeps approximate incremental trend. Figures B.2c and d show the full-wavenumber information and filtered results comparisons for different methods. From Table B.1, we find WA-FWI is closer to filtered Sto-CPT, although the increased value is small.

Table B.1 Comparisons between filtered Sto-CPT, filtered FWI and WA-FWI results are listed.

Filter	ε_r
Max wavenumber for filter (m^{-1})	3.6
Span value of smoothing function	21
Optimal FWI iteration	30
R (Filtered FWI: Filtered Sto-CPT)	0.8653
R (WA-FWI: Filtered Sto-CPT)	0.8688
RMSE (Filtered FWI: Filtered Sto-CPT)	1.6679
RMSE(WA-FWI: Filtered Sto-CPT)	1.6585

B.6 UPDATING EFFECTIVE SOURCE WAVELET BASED ON WA-FWI RESULTS

Although the filter is based on the information present along the vertical dashed lines in Figure B.1a and e, it is employed for all the FWI data along all positions. Probably, high wavenumber information is generated that is not consistent with the true data. To remove this inconsistent high wavenumber data, we estimate an updated effective wavelet where we use the wavenumber-amplified FWI data as $\hat{E}^{syn}(f)$ and the initial effective wavelet ($SW_{Ray}\ as\ \hat{s}_k(f)$) in Equation B.1, which means the synthetic data ($\hat{E}^{syn}$) will be forward modeled based on the initial wavelet (SW_{Ray}) and WA-FWI models. Using the $\hat{G}(f)$ in Equation B.2, that now contains higher wavenumbers information; an improved effective source wavelet $\hat{s}_{k+1}$ is calculated that contains a higher bandwidth.

We perform the FWI by using different effective source wavelets. Figure B.3a-c shows the ray-based, wavenumber-amplified FWI, and stochastic models that are used to perform a forward modeling, and subsequent generation of the ray-based, wavenumber-amplified FWI, and stochastic effective source wavelets. Figure B.3d-f shows the FWI results using the corresponding wavelets while using the same start models shown in Figure B.1c and d. Note that here the results contain wavenumbers up to K_Z = 3.60 m^{-1}, similar to the WA-FWI results. In Figure B.3g and h, the RMSE and R values over the vertical profiles for each horizontal position are shown. It can be seen that the FWI results using SW_{Sto} return erroneous FWI results, especially in the top location between distances from 2 to 4.1 m. The main reason is probably the white noise that was added in the true GPR data and the broadband effective wavelet causing the inversion results being trapped in a local minimum value. The SW_{Ray} results return better R values and smaller RMSE values than the WA-FWI results; this is probably due to the fact that for the WA-FWI results all spatial bandwidth has been amplified. By estimating an effective source wavelet from the WA-FWI data and performing a FWI, only the consistent large waveband data remains which is indicated by the smallest RMSE and the highest R values for SW_{WA-FWI}. Table B.2 gives an overview of the RMSE and R values obtained over the 2D domain, which confirms the former statements and indicates that the best results are obtained by the SW_{WA-FWI} approach and can improve the resolution for FWI results when CPT data is available.

Table B.2 Mean RMSE and mean R between filtered stochastic (F-Sto) and filtered FWI (F-FWI) results over full 2D domain.

Different Cases (ε_r)	RMSE	R
F-Sto and Ray-based	2.9988	0.6434
F-Sto and F-FWI (SW_{Ray})	2.4706	0.7737
F-Sto and WA-FWI	2.6892	0.7462
F-Sto and F-FWI (SW_{WA-FWI})	2.3733	0.7904
F-Sto and F-FWI (SW_{Sto})	3.0358	0.6532

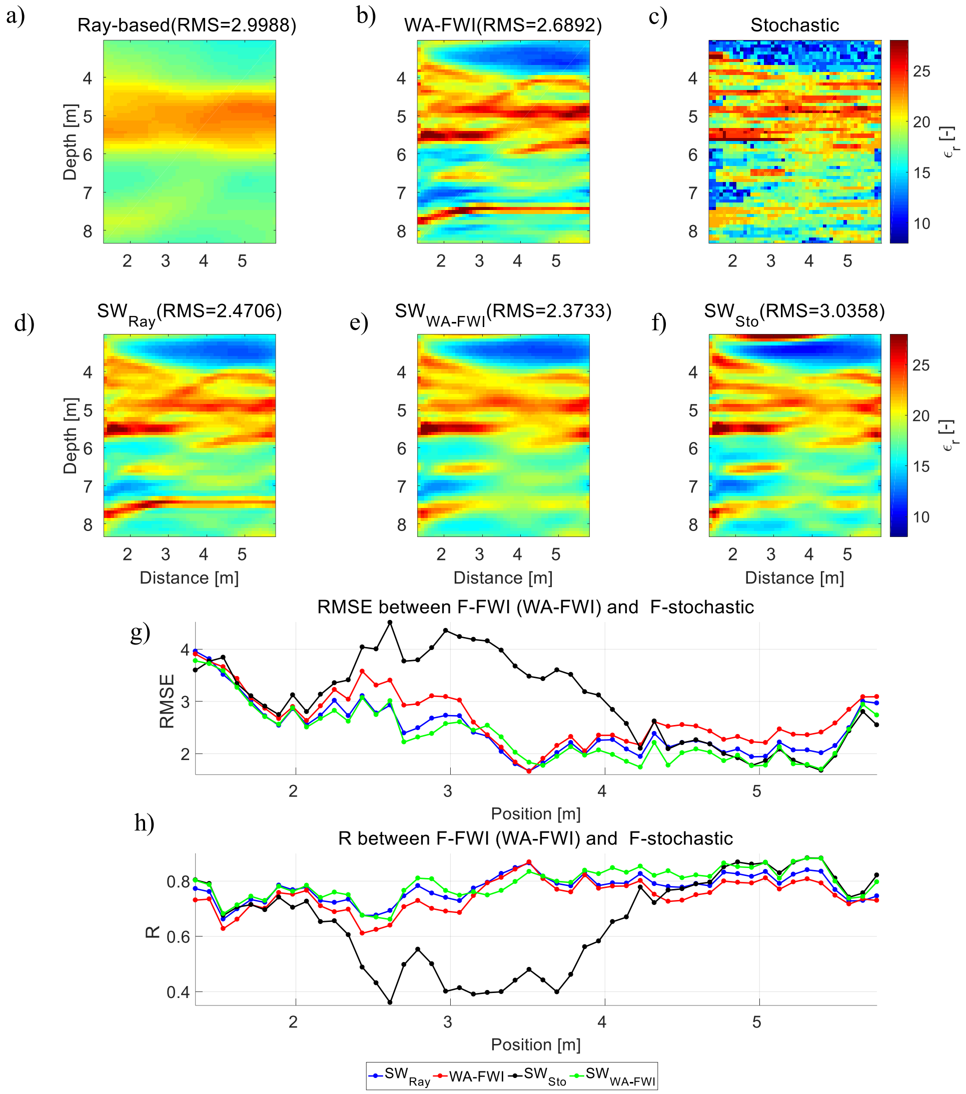


Figure B.3. a)-c) Show the starting permittivity models (ray-based, wavenumber-amplified FWI, and stochastic) for forward modeling, and subsequent generation of the ray-based, wavenumber-amplified FWI, and stochastic source wavelets. d)-f) Show the FWI results using the corresponding wavelets. The indicated RMSE values are the mean RMSE within the 2D domain, which are also given in Table B.2 including the *R* values. g) And h) show RMSE and *R* values illustrate quantitative comparison between F-stochastic models and different F-FWIs (or WA-FWI) along vertical profiles between boreholes by computing RMSE and *R* values.

B.7 CONCLUSIONS

In this paper, we present an approach to improve the resolution of GPR FWI imaging results using 1D CPT data. By using 1D vertical CPT data, we can obtain a filter based on 1D FWI permittivity results and CPT data. Then we amplify wavenumber for the whole 2D FWI cross-section and obtain wavenumber-amplified FWI results. However, WA-FWI results have been amplified for all spatial bandwidth, which generate inconsistent high wavenumber data with true models data. To remove the inconsistent high wavenumber data, we estimate an updated effective wavelet based on conventional effective source wavelet (SW_{Ray}) and WA-FWI permittivity results. In addition, the stochastic source wavelet (SW_{Sto}) is also estimated according to the SW_{Ray} and true stochastic models. Compared with the traditional source (SW_{Ray}) and the stochastic source (SW_{Sto}), the new source wavelet (SW_{WA-FWI}) has a higher bandwidth and can improve the resolution and reliability for the FWI results, especially for GPR data including white noise. Final results illustrate the feasibility and efficiency of the new source wavelet when combining FWI results and CPT data. Future work is testing the presented approach to field measured GPR data.

Acknowledgments

It is hard to imagine how to finish my PhD studies without helps and encouragements of many special people; I am at the end of my dissertation because of selfless helps and supports from these people. I want to express gratitude to who participate in my doctoral research and life.

Foremost, for the professor Dr. Jan van der Kruk, who is my doctoral supervisor and introduces me to the field of Hydrogeophysics, I would like to express thanks with my sincere gratitude. I am always remembering the scene where we first met in the AGU meeting of 2014. I was full of joy and gratitude when he promised me to join his group. Even when I postponed for half a year to arrive in Germany for personal reasons, he still kept a position for me. When I started the GPR work, Jan often spends a lot of time to help me to learn some basic concepts, to discuss some good ideas, and help me to improve my English. His open field of vision and broad knowledge were very helpful for my research direction. Besides his rigorous attitudes for academic research deeply influence and infect me. I appreciate him for all those bits of help and wish him to have a great life in the future.

Many thanks for Dr. Anja Klotzsche. She paid much time for my project, including detailed discussions of ideas, updating papers and her confidence for my abilities and my work. Also, she is one of the smartest persons I have encountered. As we knew, it is always easy and enjoyable to work together with a smart partner. And I also want to thank Professor Harry Vereecken, who provided many good ideas and feedbacks for me during my PhD reports and oral presentations. Especially my thesis defense probably not proceeds as scheduled if no help and support from Harry.

Many thanks to Prof. Klaus Reicherter and Prof. Florian Wellmann at RWTH Aachen University for their valuable support and discussion on my doctoral exams and research work.

I would also like to thank Tao Liu; a guy seems to know everything about GPR. He gave me a lot of helps at the begging of my PhD studies. And for some PhD students and post-doc in our group, I would like to say it was my pleasure to cooperate with you: Christian, Nils, Igor, Pasha, Manuela, Peleg, Yi Yu, Jessica, and Lena. Many thanks to you. Special thanks to Jessica, she helps me so much to correct papers and this thesis. Furthermore, Jessica provided many supports in the process of field measurement and dealing with observed data.

It was fortunate things to be friends with other PhD students at Juelich. Cheng Sheng, Taihong Huang, Yi Wang, Ying Xing, Yajie Sun, Hancong Xu, JiHuan Wang, Tarig, Satochi, Yueling Ma and so on. Because of your appearances, my life was become richly colorful in Germany. Special thanks to JiHuan Wang for having many lunches together.

Special thanks to my parents and my sister because of your silent supports and encouragements. And my parents-in-law and my wife's sister provided me much help in my life. The most special thanks should be received by the best important person in my life, my wife: Huiqian Tao. If no supports from you, I could never accomplish the PhD degree.

Finally, the financial support from China Scholarship Council (Project No. 201506340136) is acknowledged. Further, we gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC).

Bibliography

Arcone, S.A., 1984. Field observations of electromagnetic pulse-propagtion in dielectric slabs, Geophysics, 49, 1763–1773.

Arcone, S.A., Peapples, P.R. & Liu, L.B., 2003. Propagation of a ground penetrating radar (GPR) pulse in a thin-surface waveguide, Geophysics, 68, 1922–1933.

Axtell, C., Murray, T., Kulessa, B., Clark, R.A. and Gusmeroli, A., 2016. Improved accuracy of cross-borehole radar velocity models for ice property analysis. Geophysics, 81(1), pp.WA203-WA212.

Bear, J. and Cheng, A.H.D., 2010. Modeling groundwater flow and contaminant transport (Vol. 23). Springer Science & Business Media.

Becker, M.W., 2006. Potential for satellite remote sensing of ground water. Groundwater, 44(2), pp.306-318.

Berenger, J.P., 1994. A perfectly matched layer for the absorption of electromagnetic waves. Journal of computational physics, 114(2), pp.185-200.

Binley, A., Winship, P., Middleton, R., Pokar, M. and West, J., 2001. High-resolution characterization of vadose zone dynamics using cross-borehole radar. Water Resources Research, 37(11), pp.2639-2652.

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. Water Resour. Res. 51:3837–3866.

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

Bleistein, N., 1986. 2–1/2 dimensional inplane wave-propagation. Geophys. Prospect. 34 (5), 686–703.

Bohling, G.C., Zhan, X., Butler Jr, J.J. and Zheng, L., 2002. Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. Water Resources Research, 38(12), pp.60-1.

Bortoli, L.J., Alabert, F., Haas, A. and Journel, A., 1993. Constraining stochastic images to seismic data. In Geostatistics Tróia'92 (pp. 325-337). Springer, Dordrecht.

Brauchler, R., Hu, R., Dietrich, P. and Sauter, M., 2011. A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography. Water Resources Research, 47(3).

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

Cai, W., Qin, F. and Schuster, G.T., 1996. Electromagnetic velocity inversion using 2-D Maxwell's equations. Geophysics, 61(4), pp.1007-1021.

Carmichael, R.S. ed., 2017. Handbook of Physical Properties of Rocks (1984) (Vol. 3). CRC press.

Chen, B. and Jia, X., 2014. Staining algorithm for seismic modeling and migration. Geophysics, 79(4), pp.S121-S129.

Chew, W.C., Lin, J.H., 1995. A frequency-hopping approach for microwave imaging of large inhomogeneous bodies. IEEE Microwave Guided Wave Lett. 5, 439–441.

Choi, Y. and Alkhalifah, T., 2018. Time-domain full-waveform inversion of exponentially damped wavefield using the deconvolution-based objective function. Geophysics, 83(2), pp.R77-R88.

Clement, W.P. and Barrash, W., 2006. Crosshole radar tomography in a fluvial aquifer near Boise, Idaho. Journal of Environmental & Engineering Geophysics, 11(3), pp.171-184.

Colwell, F.S., Stormberg, G.J., Phelps, T.J., Birnbaum, S.A., McKinley, J., Rawson, S.A., Veverka, C., Goodwin, S., Long, P.E., Russell, B.F. and Garland, T., 1992. Innovative techniques for collection of saturated and unsaturated subsurface basalts and sediments for microbiological characterization. Journal of microbiological methods, 15(4), pp.279-292.

Coscia, I., Linde, N., Greenhalgh, S., Vogt, T. and Green, A., 2012. Estimating traveltimes and groundwater flow patterns using 3D time-lapse crosshole ERT imaging of electrical resistivity fluctuations induced by infiltrating river water. Geophysics, 77(4), pp.E239-E250.

Cui, T.J. and Chew, W.C., 2000. Novel diffraction tomographic algorithm for imaging two-dimensional targets buried under a lossy earth. IEEE Transactions on Geoscience and Remote Sensing, 38(4), pp.2033-2041.

Cui, T.J. and Chew, W.C., 2002. Diffraction tomographic algorithm for the detection of three-dimensional objects buried in a lossy half-space. IEEE Transactions on Antennas and Propagation, 50(1), pp.42-49.

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, J. Appl. Geophys., 73(4), 305–314.

Dafflon, B., Barrash, W., 2012. Three-dimensional stochastic estimation of porosity distribution: benefits of using ground-penetrating radar velocity tomograms in simulated-annealing-based or Bayesian sequential simulation approaches. Water Resour. Res.48.

Davis, J.L. & Annan, A.P., 1989. Ground-penetrating radar for highresolution mapping of soil and rock stratigraphy, Geophys. Prospect, 37, 531–551.

Doetsch, J., Linde, N., Coscia, I., Greenhalgh, S.A. and Green, A.G., 2010a. Zonation for 3D aquifer characterization based on joint inversions of multimethod crosshole geophysical data. Geophysics, 75(6), pp.G53-G64.

Doetsch, J., Linde, N. and Binley, A., 2010b. Structural joint inversion of time - lapse crosshole ERT and GPR traveltime data. Geophysical research letters, 37(24).

Dubois, A., Belkebir, K., Catapano, I., Saillard, M., 2009. Iterative solution of the electromagnetic inverse scattering problem for the transient scattered field. Radio Sci. 44 (RS1007).

Eisenberg, D. and Kauzmann, W., 2005. The structure and properties of water. Oxford University Press on Demand.

Ellefsen, K. J., A. T. Mazzella, R. J. Horton, and J. R. McKenna, 2011, Phase and amplitude inversion of crosswell radar data: Geophysics, 76, no. 3, J1–J12.

Ernst, J.R., Holliger, K., Maurer, H. and Green, A.G., 2006. Realistic FDTD modelling of borehole georadar antenna radiation: methodolgy and application. Near Surface Geophysics, 4(1), pp.19-30.

Ernst, J.R., Maurer, H.,Green,A.G., Holliger, K., 2007a. Full-waveform inversion of crosshole radar data based on 2D finite-difference time-domain solutions of Maxwell's equations. IEEE Trans. Geosci. Remote Sens. 45, 2807–2828.

Ernst, J.R., Green, A.G., Maurer, H., Holliger, K., 2007b. Application of a new 2D timedomain full-waveform inversion scheme to crosshole radar data. Geophysics 72, J53–J64.

Fejes, I. and Jósa, E., 1990. The engineering geophysical sounding method: principles, instrumentation, and computerised interpretation. Geotechnical and environmental geophysics, 2, pp.321-331.

Feng, X., Ren, Q., Liu, C. and Zhang, X., 2017. Joint acoustic full-waveform inversion of crosshole seismic and ground-penetrating radar data in the frequency domainJoint FWI of seismic and GPR data. Geophysics, 82(6), pp.H41-H56.

Gilson, E.W., Redman, J.D., Pilon, J. and Annan, A.P., 1996, January. Near surface applications of borehole radar. In Symposium on the Application of Geophysics to Engineering and Environmental Problems 1996 (pp. 545-553). Society of Exploration Geophysicists.

Gueting, N., Klotzsche, A., van der Kruk, J., et al., 2015. Imaging and Characterization of Facies Heterogeneity in an Alluvial Aquifer Using GPR Full-Waveform Inversion and Cone Penetration Tests. Journal of Hydrology, 524:680– 695.

Gueting, N., T. Vienken, A. Klotzsche, J. van der Kruk, J. Vander borght, 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, Water Resour. Res., 53, 49–72.

Haruzi, P., Gueting, N., Klotzsche, A., Vanderborght, J., Vereecken, H. and van der Kruk, J., 2018. Time-lapse ground-penetrating radar full-waveform inversion to detect tracer plumes: A numerical study. In SEG Technical Program Expanded Abstracts 2018 (pp. 2486-2490). Society of Exploration Geophysicists.

Hermans, T., Vandenbohede, A., Lebbe, L. and Nguyen, F., 2012b. A shallow geothermal experiment in a sandy aquifer monitored using electric resistivity tomography. Geophysics, 77(1), pp.B11-B21.

Hermans, T., 2014. Integration of near-surface geophysical, geological and hydrogeological data with multiple-point geostatistics in alluvial aquifers.

Hermans, T., Wildemeersch, S., Jamin, P., Orban, P., Brouyère, S., Dassargues, A., Nguyen, F., 2015a. Quantitative temperature monitoring of a heat tracing experiment using cross-borehole ERT. Geothermics 53, 14–26.

Hermans, T., Nguyen, F. and Caers, J., 2015b. Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: Workflow and case study. Water Resources Research, 51(7), pp.5332-5352.

Hermans, T. and Irving, J., 2017. Facies discrimination with electrical resistivity tomography using a probabilistic methodology: effect of sensitivity and regularisation. Near Surface Geophysics, 15(1), pp.13-25.

Hoffmann, R., Dassargues, A., Goderniaux, P. and Hermans, T., 2019. Heterogeneity and prior uncertainty investigation using a joint heat and solute tracer experiment in alluvial sediments. Frontiers in Earth Science-Hydrosphere, 7.

Holliger, K., Musil, M. and Maurer, H.R., 2001. Ray-based amplitude tomography for crosshole georadar data: A numerical assessment. Journal of Applied Geophysics, 47(3-4), pp.285-298.

Hubbard, S.S. and Rubin, Y., 2005. Introduction to hydrogeophysics. In Hydrogeophysics (pp. 3-21). Springer, Dordrecht.

Hu, C., Li, Q. and Jia, X., 2016. Staining algorithm for full-waveform inversion. In SEG Technical Program Expanded Abstracts 2016 (pp. 1444-1448). Society of Exploration Geophysicists.

Huisman, J.A., Sperl, C., Bouten, W. and Verstraten, J.M., 2001. Soil water content measurements at different scales: accuracy of time domain reflectometry and ground-penetrating radar. Journal of Hydrology, 245(1-4), pp.48-58.

Irving, J.D. and Knight, R.J., 2005. Effect of antennas on velocity estimates obtained from crosshole GPR data. Geophysics, 70(5), pp.K39-K42.

Irving, J.D., Knoll, M.D., Knight, R.J., 2007. Improving crosshole radar velocity tomograms: a new approach to incorporating high-angle travel time data. Geophysics 72, J31–J41.

Johnson, T.C., Routh, P.S. and Knoll, M.D., 2005. Fresnel volume georadar attenuation-difference tomography. Geophysical Journal International, 162(1), pp.9-24.

Jol, H.M. ed., 2008. Ground penetrating radar theory and applications. elsevier.

Klepikova, M., Wildemeersch, S., Hermans, T., Jamin, P., Orban, P., Nguyen, F., Brouyère, S., Dassargues, A., 2016. Heat tracer test in an alluvial aquifer: field experiment and inverse modelling. J. Hydrol. 540, 812–823.

Klotzsche, A., van der Kruk, J., Meles, G.A., Doetsch, J., Maurer, H., Linde, N., 2010. Full waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland. Near Surf. Geophys. 8, 635–649.

Klotzsche, A., van der Kruk, J., Meles, G. A., et al., 2012. Crosshole GPR Full-Waveform Inversion of Waveguides Acting as Preferential Flow Paths within Aquifer Systems. Geophysics, 77(4): H57–H62.

Klotzsche, A., van der Kruk, J., Linde, N., et al., 2013. 3D Characterization of High-Permeability Zones in a Gravel Aquifer Using 2D Crosshole GPR Full-Waveform Inversion and Waveguide Detection. Geophysical Journal International, 195(2).

Klotzsche, A., 2013. Full-waveform Inversion of Crosshole GPR Data for Hydrogeological Applications. Forschungszentrum Jülich.

Klotzsche, A., van der Kruk, J., Bradford, J., et al., 2014. Detection of Spatially Limited High-Porosity Layers Using Crosshole GPR Signal Analysis and Full-Waveform Inversion. Water Resources Research, 50(8): 6966–6985.

Klotzsche, A., Jonard, F., Looms, M.C., van der Kruk, J. and Huisman, J.A., 2018. Measuring soil water content with ground penetrating radar: a decade of progress. Vadose Zone Journal, 17(1).

Klotzsche, A., Vereecken, H. and van der Kruk, J., 2019a. GPR full-waveform inversion of a variably saturated soil-aquifer system. Journal of Applied Geophysics, 170, p.103823.

Klotzsche, A., Vereecken, H. and van der Kruk, J., 2019b. Review of Crosshole GPR Full-waveform Inversion of Experimental Data: Recent Developments, Challenges and Pitfalls. Geophysics, 84(6), pp.1-66.

Klotzsche, A., Z. Zhou, J. Schmäck, J. van der Kruk, J. Vanderborght, and H. Vereecken, 2020, Corrigendum to "Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests", Journal of Hydrology, submitted.

Kowalsky, M.B., Finsterle, S. and Rubin, Y., 2004. Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. Advances in Water Resources, 27(6), pp.583-599.

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

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

LaBrecque, D.J., Miletto, M., Daily, W., Ramirez, A. and Owen, E., 1996b. The effects of noise on Occam's inversion of resistivity tomography data. Geophysics, 61(2), pp.538-548.

Landon, M.K., Rus, D.L. and Harvey, F.E., 2001. Comparison of instream methods for measuring hydraulic conductivity in sandy streambeds. Groundwater, 39(6), pp.870-885.

Lanz, E., Maurer, H. and Green, A.G., 1998. Refraction tomography over a buried waste disposal site. Geophysics, 63(4), pp.1414-1433.

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

Le Borgne, T., Bour, O., Riley, M.S., Gouze, P., Pezard, P.A., Belghoul, A., Lods, G., Le Provost, R., Greswell, R.B., Ellis, P.A. and Isakov, E., 2007. Comparison of alternative methodologies for identifying and characterizing preferential flow paths in heterogeneous aquifers. Journal of hydrology, 345(3-4), pp.134-148.

Lesparre, N., Robert, T., Nguyen, F., Boyle, A., & Hermans, T. 2019. 4D electrical resistivity tomography (ERT) for aquifer thermal energy storage monitoring. Geothermics, 77, 368-382.

Li, G., H. Zhou, and C. Zhao, 2009, Potential risks of spectrum whitening deconvolution compared with well-driven deconvolution: Petroleum Science, 146–152.

Li, J., Innanen, K.A. and Tao, G., 2018. A 3D pseudo-spectral method for SH wave simulation in heterogeneous media. In SEG Technical Program Expanded Abstracts 2018 (pp. 4005-4009). Society of Exploration Geophysicists.

Li, Q. and Jia, X., 2017. Generalized staining algorithm for seismic modeling and migration. Geophysics, 82(1), pp.T17-T26.

Li, W., Englert, A., Cirpka, O.A., Vanderborght, J. and Vereecken, H., 2007. Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests. Water Resources Research, 43(4).

Li, W., Englert, A., Cirpka, O.A. and Vereecken, H., 2008. Three-dimensional geostatistical inversion of flowmeter and pumping test data. Groundwater, 46(2), pp.193-201.

Linde, N., Finsterle, S. and Hubbard, S., 2006. Inversion of tracer test data using tomographic constraints. Water Resources Research, 42(4).

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

Looms, M.C., Jensen, K.H., Binley, A. and Nielsen, L., 2008. Monitoring unsaturated flow and transport using cross-borehole geophysical methods. Vadose Zone Journal, 7(1), pp.227-237.

Lunne, T., Robertson, P.K. and Powell, J.J.M., Cone Penetration Testing in Geotechnical Practice., 1997. Blackie Academic and Professional, London.

MathWorks, Inc, 2016b. Signal Processing Toolbox for Use with MATLAB: User's Guide. The MathWorks.

Maurer, H. and Green, A.G., 1997. Potential coordinate mislocations in crosshole tomography: Results from the Grimsel test site, Switzerland. Geophysics, 62(6), pp.1696-1709.

Maurer, H. and Musil, M., 2004. Effects and removal of systematic errors in crosshole georadar attenuation tomography. Journal of Applied Geophysics, 55(3-4), pp.261-270.

Maurer, H., Greenhalgh, S.A., Latzel, S., 2009. Frequency and spatial sampling strategies for acoustic crosshole full waveform inversion experiments. Geophysics 74 WCC79-WCC89.

Meles, G.A., van der Kruk, J., Greenhalgh, S.A., Ernst, J.R., Maurer, H., Green, A.G., 2010. A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination Crosshole/ Borehole-to-surface GPR data. IEEE Trans. Geosci. Remote Sens. 48 (12), 4329.

Meles, G., Greenhalgh, S., Van der Kruk, J., Green, A. and Maurer, H., 2011. Taming the non-linearity problem in GPR full-waveform inversion for high contrast media. Journal of Applied Geophysics, 78, pp.31-43.

Meng, X., Liu, S., Xu, Y. and Fu, L., 2019. Application of Laplace Domain Waveform Inversion to Cross-Hole Radar Data. Remote Sensing, 11(16), p.1839.

Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data. Geophysics 52, 1211–1228.

Musil, M., Maurer, H.R. and Green, A.G., 2003. Discrete tomography and joint inversion for loosely connected or unconnected physical properties: Application to crosshole seismic and georadar data sets. Geophysical Journal International, 153(2), pp.389-402.

Musil, M., Maurer, H., Hollinger, K. and Green, A.G., 2006. Internal structure of an alpine rock glacier based on crosshole georadar traveltimes and amplitudes. Geophysical Prospecting, 54(3), pp.273-285.

Müller, K., Vanderborght, J., Englert, A., Kemna, A., Huisman, J.A., Rings, J. and Vereecken, H., 2010. Imaging and characterization of solute transport during two tracer tests in a shallow aquifer using electrical resistivity tomography and multilevel groundwater samplers. Water Resources Research, 46(3).

Nobes, D.C., 1999. Geophysical surveys of burial sites: a case study of the Oaro urupa. Geophysics, 64(2), pp.357-367.

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

Olhoeft, G.R., 2000. Maximizing the information return from ground penetrating radar. Journal of Applied Geophysics, 43(2-4), pp.175-187.

Paasche, H. and Tronicke, J., 2007. Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis. Geophysics, 72(3), pp.A35-A39.

Paige, C.C. and Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse least squares. ACM Transactions on Mathematical Software (TOMS), 8(1), pp.43-71.

Paz, C., Alcala, F.J., Carvalho, J.M. and Ribeiro, L., 2017. Current uses of ground penetrating radar in groundwater-dependent ecosystems research. Science of the Total Environment, 595, pp.868-885.

Peterson, J. E. 2001. Pre-inversion corrections and analysis of radar tomographic data. 488 Journal of Environmental and Engineering Geophysics, 6(1), 1–18.

Pinard, H., S. Garambois, L. Metivier, M. Dietrich, G. Senechal, and D. Rousset, 2016, Full-waveform inversion of GPR data acquired between boreholes in Rustrel carbonates: I-Dust-Inter-Disciplinary Underground Science and Technology.

Polak, E. and Ribiere, G., 1969. Note on Convergence of Conjugate Direction Methods. French Journal of Computer Science and Operations Research. Red series, 3 (16), pp.35-43.

Pratt, R.G., Shin, C., Hicks, G.J., 1998. Gauss–Newton and full Newton methods in frequency-space seismic waveform inversion. Geophys. J. Int. 133, 341–362.

Rabbel, W. 1996: Groundwater Geophysics: Seismic Methodes (Springer Berlin Heidelberg).

Saar, M.O., 2011. Geothermal heat as a tracer of large-scale groundwater flow and as a means to determine permeability fields. Hydrogeology Journal, 19(1), pp.31-52.

Shin, C. and Cha, Y.H., 2008. Waveform inversion in the Laplace domain. Geophysical Journal International, 173(3), pp.922-931.

Steelman, C.M. and Endres, A.L., 2011. Comparison of petrophysical relationships for soil moisture estimation using GPR ground waves. Vadose Zone Journal, 10(1), pp.270-285.

Strobach, E., Harris, B.D., Dupuis, J.C. and Kepic, A.W., 2013. Waveguide properties recovered from shallow diffractions in common offset GPR. Journal of Geophysical Research: Solid Earth, 118(1), pp.39-50.

Tarantola, A., 1984. Inversion of seismic-reflection data in the acoustic approximation, Geophysics, 49, 1259–1266.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, PA.

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

Tronicke, J., Tweeton, D.R., Dietrich, P. and Appel, E., 2001. Improved crosshole radar tomography by using direct and reflected arrival times. Journal of Applied Geophysics, 47(2), pp.97-105.

Tronicke, J. and Holliger, K., 2004. Effects of gas-and water-filled boreholes on the amplitudes of crosshole georadar data as inferred from experimental evidence. Geophysics, 69(5), pp.1255-1260.

Vanderborght, J. and Vereecken, H., 2001. Analyses of locally measured bromide breakthrough curves from a natural gradient tracer experiment at Krauthausen. Journal of contaminant hydrology, 48(1-2), pp.23-43.

van der Kruk, J., Steelman, C.M., Endres, A.L. and Vereecken, H., 2009. Dispersion inversion of electromagnetic pulse propagation within freezing and thawing soil waveguides. Geophysical Research Letters, 36(18).

van der Kruk, J., Jacob, R.W. and Vereecken, H., 2010. Properties of precipitation-induced multilayer surface waveguides derived from inversion of dispersive TE and TM GPR data. Geophysics, 75(4), pp.WA263-WA273.

van der Kruk, J., Gueting, N., Klotzsche, A., He, G., Rudolph, S., von Hebel, C., Yang, X., WeihermŘller, L., Mester, A. and Vereecken, H., 2015. Quantitative multi-layer electromagnetic induction inversion and full-waveform inversion of crosshole ground penetrating radar data. Journal of Earth Science, 26(6), pp.844-850.

Vereecken, H., Döring, U., Hardelauf, H., Jaekel, U., Hashagen, U., Neuendorf, O., Schwarze, H., Seidemann, R., 2000. Analysis of solute transport in a heterogeneous aquifer: the Krauthausen field experiment. J. Contam. Hydrol.45, 329–358.

Vidale, J.E., 1990. Finite-difference calculation of traveltimes in three dimensions. Geophysics, 55(5), pp.521-526.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, WCC1–WCC26.

Wada, Y., Van Beek, L.P., Van Kempen, C.M., Reckman, J.W., Vasak, S. and Bierkens, M.F., 2010. Global depletion of groundwater resources. Geophysical research letters, 37(20).

Wang, Y.M. and Chew, W.C., 1989. An iterative solution of the two-dimensional electromagnetic inverse scattering problem. International Journal of Imaging Systems and Technology, 1(1), pp.100-108.

Wang, G., Yuan, S. and Wang, S., 2019. Retrieving Low-Wavenumber Information in FWI: An Efficient Solution for Cycle Skipping. IEEE Geoscience and Remote Sensing Letters.

Watson, F., 2016, May. Towards 3D full-wave inversion for GPR. In 2016 IEEE Radar Conference (RadarConf) (pp. 1-6). IEEE.

Williamson, P.R., 1991. A guide to the limits of resolution imposed by scattering in ray tomography. Geophysics, 56(2), pp.202-207.

Williamson, P.R. and Worthington, M.H., 1993. Resolution limits in ray tomography due to wave behavior: Numerical experiments. Geophysics, 58(5), pp.727-735.

Wu, R.S., Luo, J. and Wu, B., 2014. Seismic envelope inversion and modulation signal model. Geophysics, 79(3), pp.WA13-WA24.

Yang, X., Klotzsche, A., Meles, G., Vereecken, H., van der Kruk, J., 2013. Improvements in crosshole GPR full-waveform inversion and application on data measured at the Boise hydrogeophysics research site.J.Appl.Geophys.99, 114 –124.

Zhou, B., Greenhalgh, S.A., 2003. Crosshole seismic inversion with normalized full waveform amplitude data. Geophysics 68, 1320–1330.

Zhou, Z., Klotzsche, A., Güting, N., Haruzi, P., Vereecken, H. and Kruk, J.V.D., 2019. Improved resolution of ground penetrating radar full-waveform inversion by using cone penetration test data: A synthetic study. In SEG Technical Program Expanded Abstracts 2019 (pp. 2898-2902). Society of Exploration Geophysicists.

Zhou, Z., A. Klotzsche, T. Hermans, F. Nguyen, J. Schmäck, P. Haruzi, H. Vereecken, and J. van der Kruk, 2020a. 3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole GPR amplitude analysis and full-waveform inversion: submitted to Geophysics.

Zhou, Z., A. Klotzsche, J. Schmäck, H. Vereecken, and J. van der Kruk, 2020b. Improvement of GPR Full-waveform inversion images using Cone Penetration Test data: submitted to Geophysics.

Zhou, Z., A. Klotzsche, H. Vereecken, and J. van der Kruk, 2020c Improvement of crosshole GPR FWI results by using progressively expanded bandwidths of the data: submitted to Journal of Applied Geophysics.

List of Figures

List of Tables

Curriculum Vitae

Zhen Zhou

Linnicher Str. 53 52428 Juelich Germany

Tel: +49 (0)2461613019

Mobile: +49 (0)15905869384

Email: z.zhou@fz-juelich.de

zzhen12@mail.ustc.edu.cn

Personal:

Date of birth December 2nd, 1986

Place of birth Huai Bei, An hui, China

Citizenship Chinese

Education:

03/2016 – 08/2020

PhD of Geophysics, *Forschungszentrum Juelich and RWTH Aachen University (Germany)*,

PhD thesis: *Enhanced crosshole GPR full-waveform inversion to improve aquifer characterization,*

Supervisors: Prof. Dr. Jan van der Kruk; Dr. Anja Klotzsche.

09/2012 – 07/2015

Master of Geophysics, *University of Science and Technology of China (China),*

Master thesis: *GPU-accelerated element-free reverse-time migration with Gauss points partition,*

Supervisors: Prof. Xiaofeng Jia.

09/2005 – 07/2009

Bachelor of Geophysics, *Chang'an University (China),*

Bachelor thesis: *Application of Direct Current Sounding on Detection of Collapsing Karst Column in Coal Mine,*

Supervisors: Prof. Bing Feng.

Publications:

Zhou, Z., A. Klotzsche, H. Vereecken, and J. van der Kruk, 2020c. *Improvement of crosshole GPR FWI results by using progressively expanded bandwidths of the data:* Journal of Applied Geophysics, submitted.

Zhou, Z., A. Klotzsche, J. Schmäck, H. Vereecken, and J. van der Kruk, 2020b. *Improvement of GPR Full-waveform inversion images using Cone Penetration Test data*: Geophysics, under review.

Zhou, Z., A. Klotzsche, T. Hermans, F. Nguyen, J. Schmäck, P. Haruzi, H. Vereecken, and J. van der Kruk, 2020a, *3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole GPR amplitude analysis and full-waveform inversion:* Geophysics, under review.

Klotzsche, A., **Z. Zhou**, J. Schmäck, J. van der Kruk, J. Vanderborght, and H. Vereecken, 2020, Corrigendum to “Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests”, Journal of Hydrology, submitted.

Mozaffari Amirpasha**, Zhen Zhou**, Anja Klotzsche, Harry Vereecken, Jan van der Kruk, *3D electromagnetic modeling explains apparent velocity increase in crosshole GPR data - borehole fluid effect correction method enables to incorporating high-angle traveltime data:* submitted to Geophysics.

Zhen Zhou, Xiaofeng Jia and Xiaodong Qiang, 2018. *GPU-accelerated element-free reverse-time migration with Gauss points partition*, Journal of Geophysics and Engineering, 15: 718-728.

Zhen Zhou and Xiaofeng Jia, 2016. *OpenMP-accelerated element-free reverse-time migration,* Computing Techniques for Geophysical and Geochemical Exploration, 38 (6): 821-831 (Chinese).

Conference Proceedings & Abstracts:

Zhen Zhou, Anja Klotzsche, Nils Güting, Peleg Haruzi, Harry Vereecken, Jan van der Kruk, 2019, *Improved resolution of ground penetrating radar full-waveform inversion by using cone penetration test data: A synthetic study*. In SEG Technical Program Expanded Abstracts 2019 (pp. 2898-2902). Society of Exploration Geophysicists. **Oral presentation** in San Antonio, USA.

Zhen Zhou and Xiaofeng Jia, 2014, *Element-Free Method with Gauss Points Partition*, CPS/SEG 2014 International Geophysical Conference & Exposition, Expanded Abstracts, 677-680. **Oral presentation** in Beijing, China.

October 2018, **Poster report**, *TERENO International Conference* 2018, in Berlin, Germany.

April 2018, **Poster report**, *Terrestrial Systems Research conference* 2018, in Bonn, Germany.

December 2014, **Poster report**, *American Geophysical Union meeting* 2014, in San Francisco, USA.

Band / Volume 499
WEGE FÜR DIE ENERGIEWENDE
Kosteneffiziente und klimagerechte Transformationsstrategien für das deutsche Energiesystem bis zum Jahr 2050
M. Robinius et al (2020), VIII, 141 pp
ISBN: 978-3-95806-483-6

Band / Volume 500
Mechanical Behavior of Solid Electrolyte Materials for Lithium-ion Batteries
G. Yan (2020), x, 139 pp
ISBN: 978-3-95806-484-3

Band / Volume 501
Retrieval of atmospheric quantities from remote sensing measurements of nightglow emissions in the MLT region
Q. Chen (2020), 208 pp
ISBN: 978-3-95806-485-0

Band / Volume 502
Auswirkungen der Energiewende auf das deutsche Gastransportsystem
B. Gillessen (2020), XVII, 186
ISBN: 978-3-95806-487-4

Band / Volume 503
Lagrangian Simulation of Stratospheric Water Vapour: Impact of Large-Scale Circulation and Small-Scale Transport Processes
L. Poshyvailo (2020), 124 pp
ISBN: 978-3-95806-488-1

Band / Volume 504
Water Management in Automotive Polymer-Electrolyte-Membrane Fuel Cell Stacks
S. Asanin (2020), XVIII, 172 pp
ISBN: 978-3-95806-491-1

Band / Volume 505
Towards a new real-time irrigation scheduling method: observation, modelling and their integration by data assimilation
D. Li (2020), viii, 94 pp
ISBN: 978-3-95806-492-8

Band / Volume 506
Modellgestützte Analyse kosteneffizienter CO_2-Reduktionsstrategien
P. M. Lopion (2020), XIV, 269 pp
ISBN: 978-3-95806-493-5

Band / Volume 507
Integration of Renewable Energy Sources into the Future European Power System Using a Verified Dispatch Model with High Spatiotemporal Resolution
C. Syranidou (2020), VIII, 242 pp
ISBN: 978-3-95806-494-2

Band / Volume 508
Solar driven water electrolysis based on silicon solar cells and earth-abundant catalysts
K. Welter (2020), iv, 165 pp
ISBN: 978-3-95806-495-9

Band / Volume 509
Electric Field Assisted Sintering of Gadolinium-doped Ceria
T. P. Mishra (2020), x, 195 pp
ISBN: 978-3-95806-496-6

Band / Volume 510
Effect of electric field on the sintering of ceria
C. Cao (2020), xix, 143 pp
ISBN: 978-3-95806-497-3

Band / Volume 511
Techno-ökonomische Bewertung von Verfahren zur Herstellung von Kraftstoffen aus H_2 und CO_2
S. Schemme (2020), 360 pp
ISBN: 978-3-95806-499-7

Band / Volume 512
Enhanced crosshole GPR full-waveform inversion to improve aquifer characterization
Z. Zhou (2020), VIII, 136 pp
ISBN: 978-3-95806-500-0

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

Energie & Umwelt / Energy & Environment
Band / Volume 512
ISBN 978-3-95806-500-0

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum