Time-lapse ground penetrating radar full-waveform inversion to detect tracer plumes: Numerical study

Peleg Haruzi[1], Jessica Schmäck[1], Zhen Zhou[1], Jan van der Kruk[2], Harry Vereecken[3], Jan Vanderborght[2], and Anja Klotzsche[2]

[1]Forschungszentrum Jülich
[2]Forschungszentrum Juelich
[3]Forschungszentrum Jülich GmbH

November 26, 2022

Abstract

The movement and spreading of contaminated groundwater plumes and their mixing with non-contaminated water is strongly influenced by the heterogeneity of the aquifer properties, which may vary strongly over small spatial scales. Thus, imaging these small-scale features and monitoring transport of tracer plumes at a fine resolution is of interest to characterize transport processes in aquifers. Full-waveform inversion (FWI) of crosshole ground penetrating radar (GPR) measurements can provide an aquifer characterization at decimeter-scale resolution. The method produces images of both relative dielectric permittivity (ϵ_ρ) and bulk electrical conductivity (σ_β), which related to hydraulic aquifer properties and tracer distributions. To test the potential of time-lapse GPR FWI for imaging tracer plumes, we conducted a numerical experiment of tracer transport in a heterogeneous aquifer. Concentration was converted to saline and desalinated tracers, which changed σ_β, and to ethanol, which changed both ϵ_ρ and σ_β. The simulated ϵ_ρ and σ_β distributions in a crosshole plane were considered to simulate GPR data. These data were subsequently used to reconstruct ϵ_ρ and σ_β distributions using the crosshole 2D GPR FWI. Tracer concentrations were retrieved from the inverted ϵ_ρ and σ_β models using information about petrophysical parameters. GPR FWI ϵ_ρ images could recover preferential paths of ˜0.2 m width, while the σ_β images resolved structures up to ˜ 0.2-0.3 m. The results highlight that changes in ϵ_ρ, e.g., ethanol and hot water, can be used to image transport processes with high resolution by time-lapse GPR FWI, while the accuracy of the recovery of σ_β is limited.

Supporting Information for

Time-lapse ground penetrating radar full-waveform inversion to detect tracer plumes: Numerical study

P. Haruzi [1,2*], J. Schmäck[1,2], Z. Zhou [1,2], J. van der Kruk[1,2], H. Vereecken[1,2], J. Vanderborght [1,2], A. Klotzsche[1,2]

[1]Agrosphere Institute, IBG-3, Forschungszentrum Jülich GmbH

[2]Centre for High-Performance Scientific Computing in Terrestrial Systems (HPSC-TerrSys), Geoverbund ABC/J.

Corresponding author: Peleg Haruzi (p.haruzi@fz-juelich.de)

Contents of this file

Figures S1 to S6

Introduction

This supporting information provides complementary information for the paper and the supporting figures are referred from the body of the text.

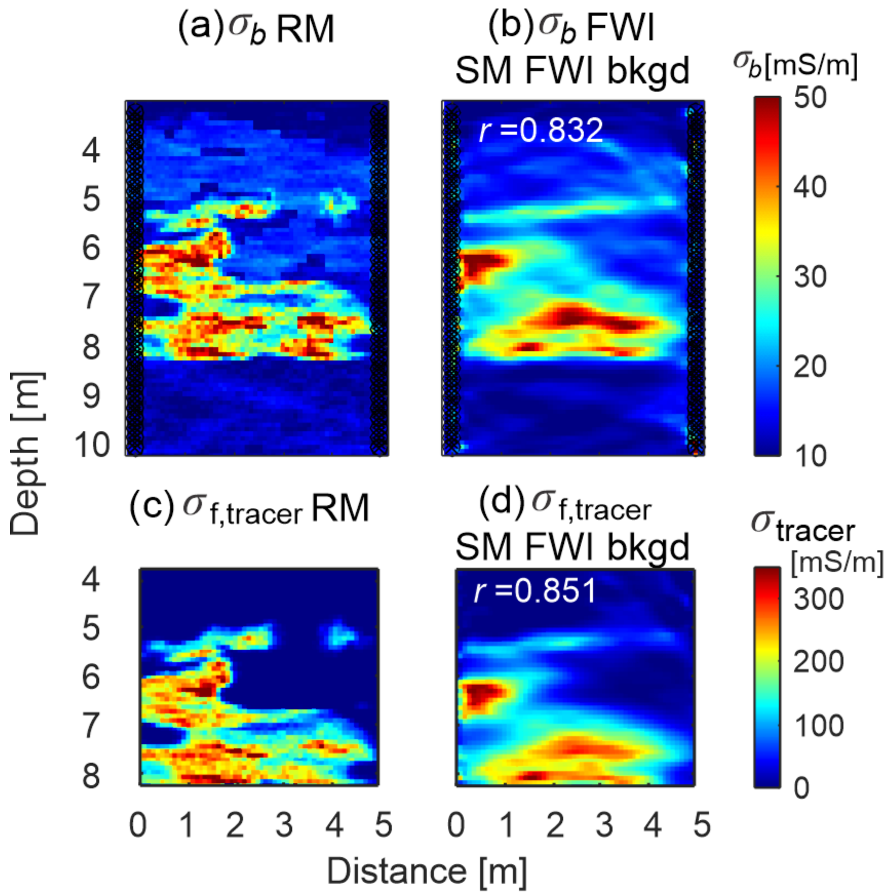


Figure S1. Tracer recovery of salt tracer at day 13. Bulk conductivity (top row) and tracer fluid conductivity (bottom) from (a,c) true model, (b,d) FWI of noise-added data with SM from FWI background.

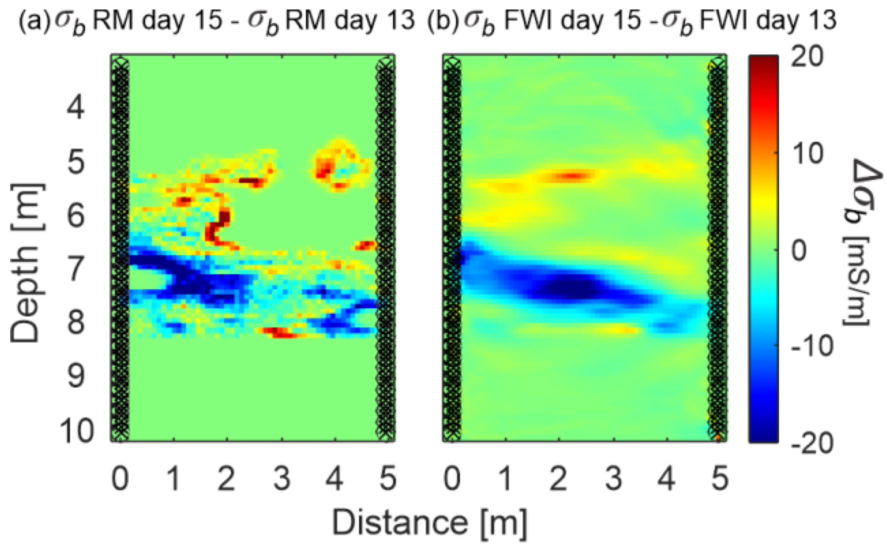


Figure S2. Changes in bulk conductivity between day 15 and day 13 based from (a) real models and (b) FWI recovered models.

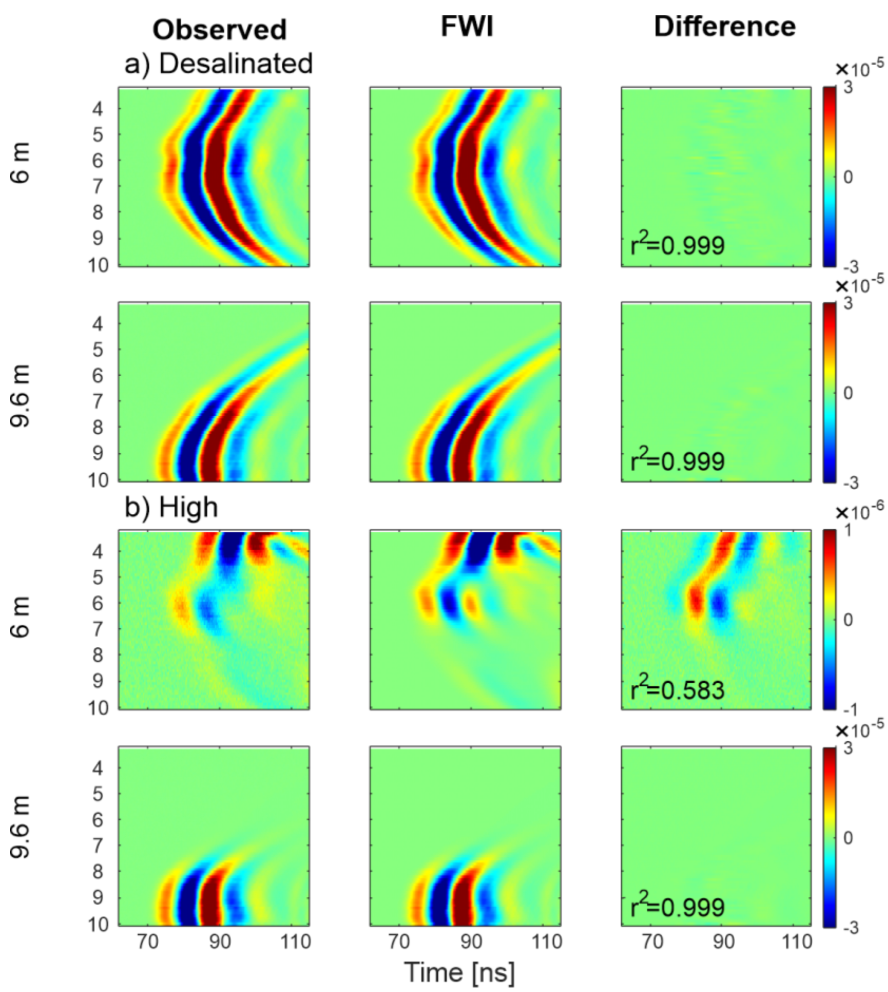


Figure S3. *Observed* (real data noise-added), FWI inverted and the difference between inverted and observed data for transmitters at the depth of major tracer intrusion (6 m) and at depth where no intrusion occurs (9.6 m). This figure is complementary to Figure 13 in the paper. Data is presented for the (a) Desalinated (tracer case in Figure 10 a) and (b) High (tracer case in Figure 10 e) tracer cases. Note that for the High case in (b) for the transmitter at 6 m depth panel where the signal is weaker because it the wave travels through the increased σ of the tracer, the color scale is 30 times smaller. r^2 quantifies the correlation between FWI inverted and the observed data. The standard deviation of the Gaussian ambient noise was 4.6[?]10^{-8} in all cases. Figure S3 (in Supplements) show Desalinated and High tracer cases.

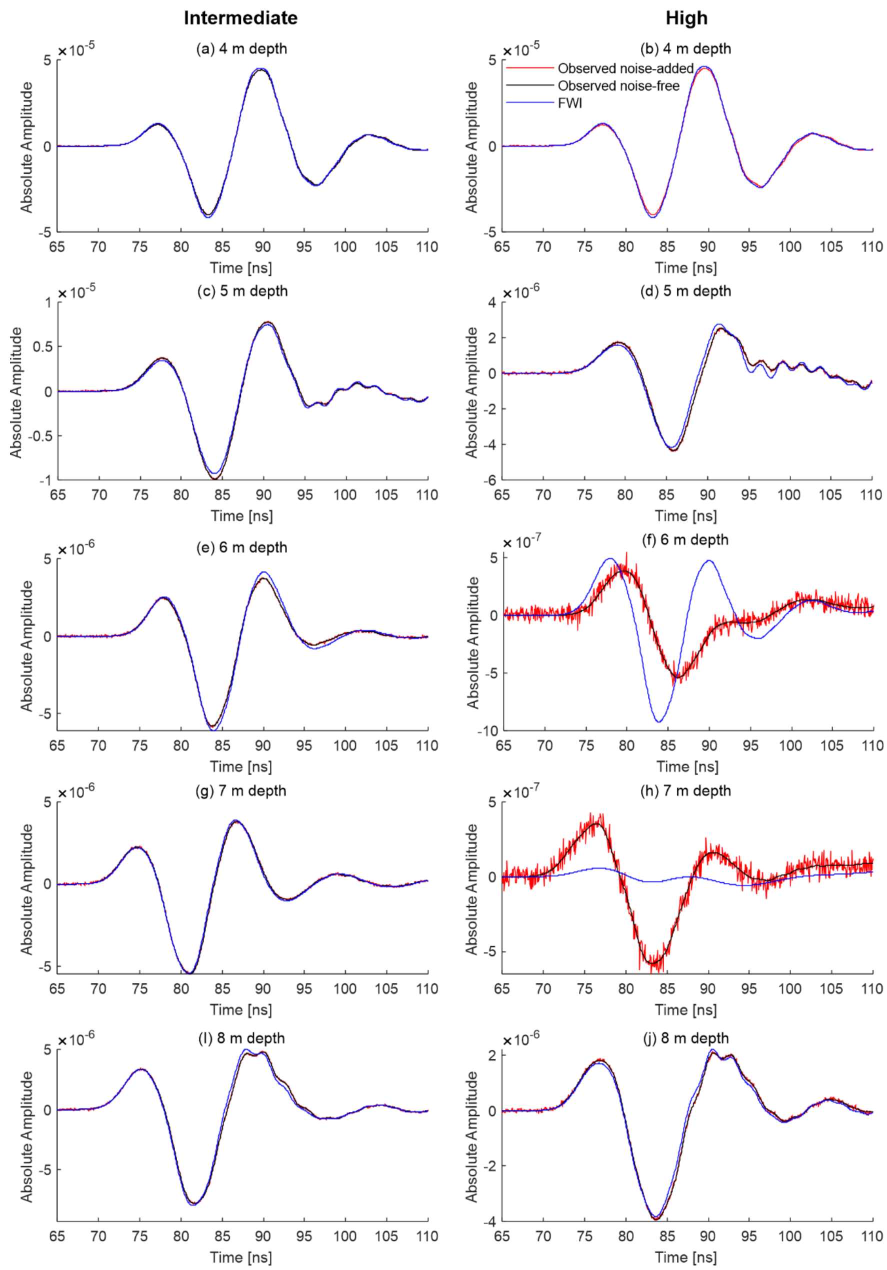


Figure S4. Fit of FWI traces. Traces of Intermediate (left) and High cases (right) for ray paths travelling parallel to surface at depths where the plume intrudes in the crosshole plane. The observed and FWI

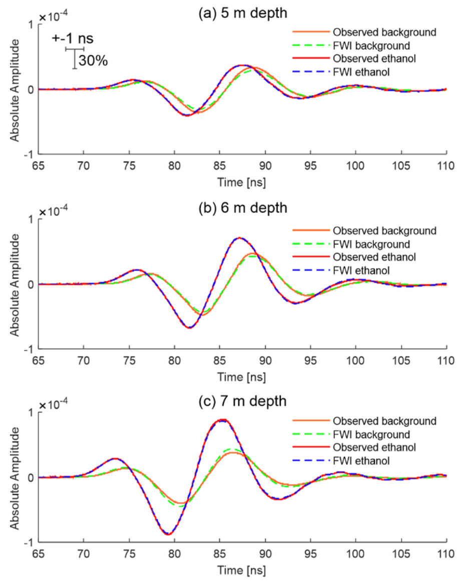
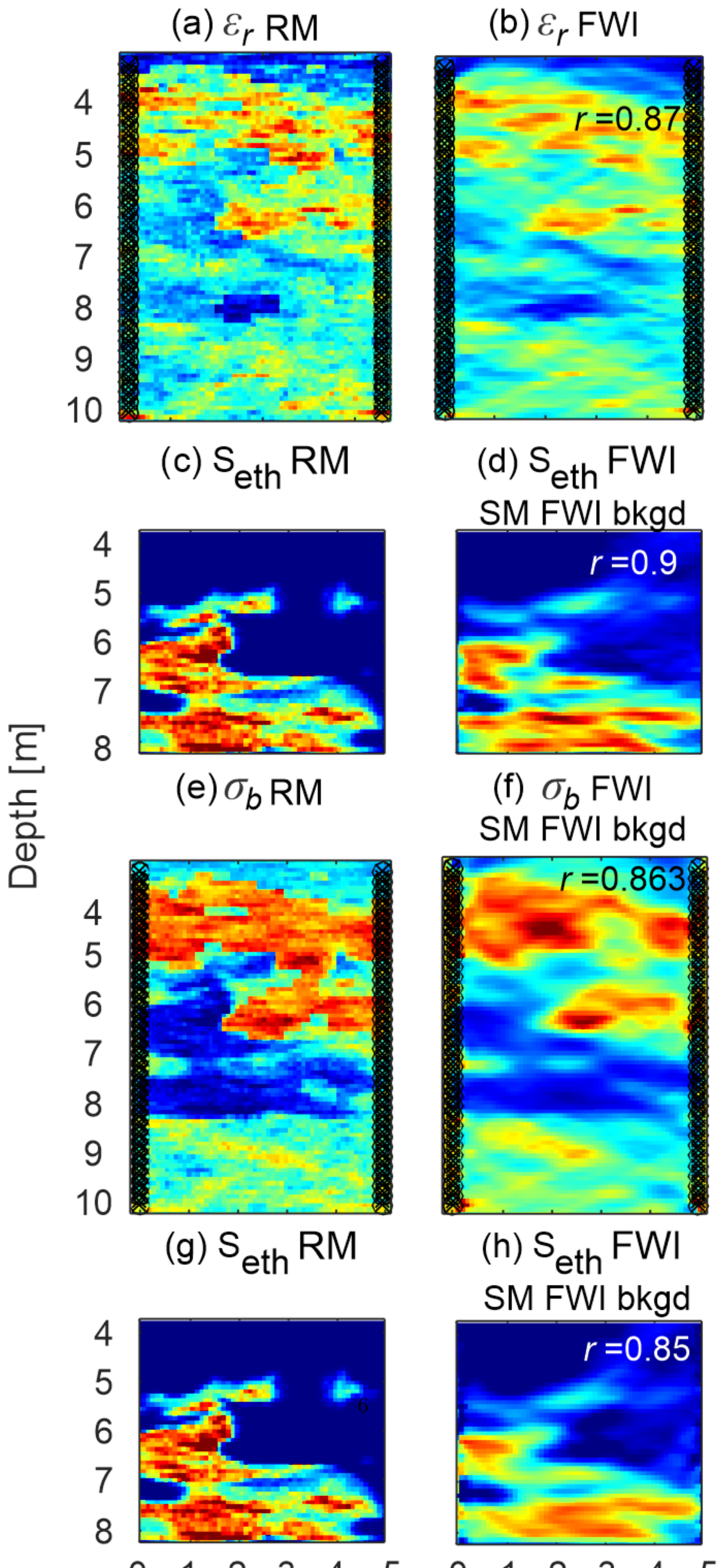


Figure S5. Observed and modelled FWI traces for background and ethanol (At day 15, using starting model from day 13) models at the main intrusion (a) 5 m, (a) 6 m and (c) 7 m.



(a) ε_r RM
(b) ε_r FWI
r =0.87
(c) S_{eth} RM
(d) S_{eth} FWI SM FWI bkgd
r =0.9
(e) σ_b RM
(f) σ_b FWI SM FWI bkgd
r =0.863
(g) S_{eth} RM
(h) S_{eth} FWI SM FWI bkgd
r =0.85
Depth [m]

Figure S6. Tracer recovery of ethanol at day 13. Top row: permittivity, second row: ethanol volumetric concentration from permittivity, third row: bulk conductivity, bottom row: ethanol volumetric concentration from permittivity. Left column: the real models, right column the recovered FWI models using SM of background FWI.

Time-lapse ground penetrating radar full-waveform inversion to detect tracer plumes: Numerical study

P. Haruzi [1,2*], J. Schmäck[1,2], Z. Zhou [1,2], J. van der Kruk[1,2], H. Vereecken[1,2], J. Vanderborght [1,2], A. Klotzsche [1,2]

[1]Agrosphere Institute, IBG-3, Forschungszentrum Jülich GmbH

[2]Centre for High-Performance Scientific Computing in Terrestrial Systems (HPSC-TerrSys), Geoverbund ABC/J.

Corresponding author: Peleg Haruzi (p.haruzi@fz-juelich.de)

Key Points:

Hydrogeophysics, GPR, full-waveform inversion, time-lapse, tracer experiment, petrophysical relations, effective medium models, stochastic simulation, transport modeling.

Abstract

The movement and spreading of contaminated groundwater plumes and their mixing with non-contaminated water is strongly influenced by the heterogeneity of the aquifer properties, which may vary strongly over small spatial scales. Thus, imaging these small-scale features and monitoring transport of tracer plumes at a fine resolution is of interest to characterize transport processes in aquifers. Full-waveform inversion (FWI) of crosshole ground penetrating radar (GPR) measurements can provide an aquifer characterization at decimeter-scale resolution. The method produces images of both relative dielectric permittivity (ε_r) and bulk electrical conductivity (σ_b), which related to hydraulic aquifer properties and tracer distributions. To test the potential of time-lapse GPR FWI for imaging tracer plumes, we conducted a numerical experiment of tracer transport in a heterogeneous aquifer. Concentration was converted to saline and desalinated tracers, which changed σ_b, and to ethanol, which changed both ε_r and σ_b. The simulated ε_r and σ_b distributions in a crosshole plane were considered to simulate GPR data. These data were subsequently used to reconstruct ε_r and σ_b distributions using the crosshole 2D GPR FWI. Tracer concentrations were retrieved from the inverted ε_r and σ_b models using information about petrophysical parameters. GPR FWI ε_r images could recover preferential paths of ~0.2 m width, while the σ_b images resolved structures up to ~ 0.2-0.3 m. The results highlight that changes in ε_r, e.g., ethanol and hot water, can be used to image transport processes with high resolution by time-lapse GPR FWI, while the accuracy of the recovery of σ_b is limited.

1 Introduction

Ever increasing water demands and anthropogenic pollution lead to depletion of clean groundwater resources. Detailed knowledge of the flow and transport processes, which control migration of fluids, particles, and solutes in the subsurface (hereafter tracers), is necessary, e.g., to protect groundwater pumping wells from pollution and operate remediation measures [Maliva, 2016]. Important transport characteristics that need to be known are the tracer velocity, the tracer plume spreading, and the tracer dilution by mixing with groundwater. These transport characteristics depend strongly on the heterogeneity of hydraulic aquifer properties [Cheng and Bear, 2016], which are difficult to observe directly because of the intrinsic inaccessibility of the subsurface. Tracer experiments that monitor tracer plumes in aquifers can be used to determine transport characteristics and infer the underlying hydraulic aquifer properties and their spatial variability [e.g., Vereecken et al., 2000].

Traditional techniques for hydrologic characterization, such as pumping tests, provide data on large-scale aquifer hydraulic properties but with low spatial resolution [e.g., Li et al., 2007]. Other well established techniques provide fine-scale information in the vertical direction, such as borehole measurements [Englert, 2003], cone penetration tests [Tillman et al., 2008], and measurements on sediment cores [Vereecken et al., 2000], but cannot characterize spatial variability in the horizontal (flow) direction with high spatial resolution. Geophysical imaging techniques such as electrical resistivity tomography (ERT) and GPR can close this gap in observation capabilities and provide information on an appropriate scale (up to ~10 m) and with high spatial resolution in both vertical and horizontal direction, while being minimally intrusive [e.g., Looms et al., 2008, Binley et al., 2015]. Geophysical imaging methods enable to image the subsurface by sensing changes in the physical parameters of a porous medium. Specifically, relative dielectric permittivity (ε_r) and electrical conductivity (σ) of porous media or an aquifer vary in space and time [Everett, 2013]. ε_r is mainly dominated by the water content and its temperature, while σ depends on the salinity and temperature of the pore water and on the clay content [Everett, 2013]. Migration of a tracer through the aquifer changes these aquifer properties so that imaging these changes in a time-lapse manner using dedicated geophysical methods, such as GPR [Klotzsche et al., 2019a, Looms et al., 2008], and ERT [Kemna et al., 2002, Singha et al., 2005, Hermans et al., 2015], can be used to image the tracer plume. Whereas

ERT measurements are made using direct current and provide bulk electrical conductivity (σ_b), GPR operates at high-frequencies range (typically 10-2600 MHz) and uses the propagation of the electromagnetic (EM) wave in resistive earth materials. In contrast to ERT, GPR can provide both ε_r and σ_b. While the velocity of the EM wave can be linked to ε_r , the attenuation of the EM wave provides information about the σ_b [Annan, 2009]. The used high-frequency of the GPR systems allow higher imaging resolution of the subsurface that scales with the wavelength (λ) of the measured signal. For a typically used frequency spectra of 10-200 MHz (the range used in this study) of the EM signal and a ε_r of 12-25 of the media, the wavelength scales between 0.3 and 8.5 m [Annan, 2009]. Especially, GPR acquisition in a wave transmission configuration with transmitters in one borehole and receivers in another (crosshole) [Huisman et al., 2003, Klotzsche et al., 2019b] allows a good subsurface illumination with dense ray-coverage and relatively small acquisition errors [Axtell et al., 2016]. Time-lapse crosshole GPR monitoring of fluid transport was successful in illuminating preferential pathways from either signal attenuation due to a saline tracer [Day-Lewis et al., 2003], or wave velocity changes due to soil water content changes [Looms et al., 2008].

Crosshole GPR data is measured mainly in multi-offset gather (MOG) measurements and commonly imaged with ray-based tomography. Velocity distribution (e.g., Dafflon et al., [2011]), from which ε_r images are derived, are obtained from the first arrival travel times of the wave signals, and, attenuation tomograms of the subsurface, from which σ_b images are estimated, are obtained from first-cycle amplitudes [Holliger et al., 2001]. Unlike the ray-based approach, which uses only specific features of the recorded waveform, GPR full-waveform inversion (FWI) uses the full information content of the received signal, what ultimately improves the resolution of the ε_r and σ_b images [Klotzsche et al., 2019b] . Time-domain crosshole GPR FWI was applied in the last decade to more than 40 different datasets from various test sites and demonstrated the possibility to characterize aquifers within decimeter-scale resolution including important small-scale structures like high porosity zones and impermeable clay lenses (overview provided by Klotzsche et al. [2019b]). Thereby, an amplitude analysis approach and the FWI was able to detect and localize zones of higher permittivity (intermediate σ_b), which act as low velocity electromagnetic waveguide and which were linked to zones of preferential water flow with higher hydraulic conductivity [Klotzsche et al., 2013]. A similar study, investigated the possibility to map zones with higher σ_b associated with increased clay content indicating clay lenses in the aquifer [Zhou et al., 2020]. Gueting et al., [2015, 2017] demonstrated that 2D crosshole GPR FWI results improved large-scale characterization of aquifer heterogeneity and could identify aquifer layers of a few decimeters thickness. This high-resolution reconstruction of layers allowed to explain a previously observed tracer plume transport and in particular tracer plume splitting [Müller et al., 2010] that was caused by the presence of a thin layer with a lower hydraulic conductivity.

Generally, the time-domain crosshole GPR FWI is an iterative approach to simultaneously estimate ε_r and σ_b by minimizing the misfit function between measured and modeled GPR data with a gradient-type approach (for more details we refer to Meles et al. [2010] and Klotzsche et al. [2019b]). Thereby, a 2D finite-difference time-domain (FDTD) algorithm is used that solves the full Maxwell equations and allows predicting the EM wave propagation through the heterogeneous medium. In order to prevent the misfit function to converge to a local minimum, a ε_r starting model (SM) is required that yields synthetic waveforms that match all the observed data within less than half of the wavelength and avoids cycle-skipping [Meles et al., 2010]. Normally, ray-based inversion results can provide such starting models. In the presence of high contrasts, such as a water table or high permittivity zones, ray-based SM often need to be updated to meet these criteria [Klotzsche et al., 2012]. Local invasion of tracer may generate small-scale high contrasts in ε_r and σ_b over short distances, which cannot be resolved by ray-based inversions. SM based on a ray-based inversion may differ too much from the true distribution so that a local gradient based optimization algorithm may not find the global minimum of the misfit function.

Next to spatial resolution, another problem in geophysical imaging is the translation of imaged parameters (ε_r and σ_b in GPR) to the property of interest, the tracer or substance concentration. Since the petrophysical relations between them depend on aquifer properties like porosity [Birchak et al., 1974], pore structure [Archie, 1942], surface charge density of the mineral surfaces [Rhoades et al., 1981], which are caused by spatially variable aquifer heterogeneity, this relation is spatially variable and site dependent (e.g., Müller et al. [2010)]. The translation of the imaged electric property distributions to concentration distribution is therefore afflicted by this spatial variability. Utilizing high-resolution GPR FWI before and after the tracer can be used to reduce the uncertainty in petropysical relations.

In this study, we analyze the potential of time-lapse crosshole GPR FWI for imaging tracer tests in heterogeneous aquifers, using a numerical experiment. The setup of the experiment is based on the properties of the aquifer at the Krauthausen test site, which consists of heterogeneous alluvial sandy-gravel sediments showing preferential flow paths with thicknesses of ~0.2 m [Gueting et al., 2017]. Different tracer scenarios were analyzed using a salt and an ethanol tracer. Through petrophysical relations the plume concentrations I) of positive/negative salt tracers were converted to increases/decreases in σ_b and II) of an ethanol tracer to decreases in both σ_b and ε_r. Note, commonly, only changes in σ_b from salt [Kemna et al., 2002] and heat [Hermans et al., 2015] tracers are imaged with methods like ERT. Since the GPR FWI is able to provide both high-resolution ε_r and σ_b images, we want to evaluate the potential of imaging small-scale tracer distributions from both parameters. Crosshole GPR data were generated before and during the plume intrusion. We tested the ability of FWI to recover the tracer distribution for different tracer concentrations that generated different changes and contrasts in σ_b and ε_r. To monitor tracer experiments, starting models for ε_r and σ_b based on a high resolution FWI model of a previous time step may be beneficial [Zhang and Huang, 2013, Asnaashari et al., 2015]. That because When changes in electrical properties due to changes in tracer concentrations between two time steps are moderate and smaller than the changes from the background distribution, a starting model that uses ε_r and σ_b distributions from the first time step rather than from the background might be closer to the global minimum of the misfit function of the second data. Therefore, we tested two starting model strategies: I) Using the recovered FWI background models and, II) considering the recovered FWI models from GPR measurements from a previous time step when the plume concentration distribution is relatively similar to the actual one.

2 Realistic hydrological aquifer model domain

To realistically model time-lapse GPR data and perform the FWI for the different tracer scenarios, we developed a realistic hydrological model domain of an aquifer in which we simulated flow and transport processes. To achieve this, we used the detailed database and knowledge from the Krauthausen test site in Germany (see Tillmann et al. [2008] and Gueting et al., [2017] for more details).

2.1 Krauthausen test site as aquifer model domain

The Krauthausen aquifer is an alluvial sandy-gravel aquifer with a silt and clay content that varies between 0.5 - 7.5% [Vereecken et al., 2000]. With respect to GPR, the aquifer is well suited due to its low to intermediate electrical conductivity between 5 and 20 mS/m [Zhou et al, 2021]. The test site extent is 200 x 70 m and has more than 70 wells used for pumping and water sampling which reach to 9 - 12 m depth. The ground water level varies through the year between 1-3 m below surface [Englert, 2003]. The site was investigated in multiple studies using hydrogeological characterization by cone penetration test (CPT) [Tillman et al., 2008] and tracer tests [Vereecken et al., 2000, Vanderborght and Vereecken, 2001], using soil and water sampling [Englert et al., 2000], borehole velocity measurements [Englert, 2003], geophysical imaging methods of ERT [Kemna et al., 2002, Müller et al., 2010], GPR [Oberröhrmann et al., 2013, Gueting et al., 2017, Zhou et al., 2020], spectral induced polarization [Kelter et al., 2018]), and pumping tests [Li et al., 2008]. The hydraulic conductivity (K) and porosity ($\boldsymbol{\phi}$, derived from neutron activity in cone penetration tests – CPT [Tillman et al., 2008]) datasets were used for flow and transport modeling, and ε_r (from $\boldsymbol{\phi}$ using CRIM model [Birchak et al., 1974]) and σ_b datasets for GPR modeling. Heterogeneity of K that mostly influences preferential path thickness was derived from lnK histograms with variance of 0.6 and correlation lengths (I) in the vertical I_v=0.18 m [Englert et al., 2003] and in the horizontal directions of I_h=1.75 m direction [Tillmann et al., 2008].

As a first step, we adopted the 3D facies model (Figure 1a) from Gueting et al. [2017 and 2018], which was generated based on adjoint tomograms from 2D GPR full-waveform inversions, and subsequently expanded to a 3D cube using multiple-point statistics. This model is composed of three facies: sandy gravel, sand, and gravel. Secondly, we generated the distributions of four aquifer parameters (K, ϕ, ε_r, σ_b), in each of the three facies using stochastic Gaussian simulation (SGSIM) based on variogram modeling [SGeMS software, Remy et al., 2009]. For each property and facies, the simulation was performed over the entire model domain, and then they were integrated to one aquifer model domain (“cookie-cutter”) based on the 3D facies model [Gueting et al., 2018]. K and ϕ models were simulated independently with no spatial crosscorrelation, ε_r was calculated directly from the ϕ model using the CRIM model [Birchak et al., 1974], and σ_b was simulated using sequential Gaussian co-simulation (COSGSIM) based on the ε_r spatial distribution as a secondary information with correlation r=0.5 in each facies [Gueting et al., 2015]. For all the stochastic simulations an exponential variogram model was used with a nugget of 0. Tables 1 and 2 summarize the parameter values and references used to set up the different property distributions. The final models are visualized in Figure 1, with a grid that covers a domain size of 20.07 x 30.15 x 4.68 m, from 3.58 to 8.26 m depth and this is composed of cubic cells with edge size of 0.09 m.

Table 1: Mean values, variance, horizontal and vertical correlation lengths (I_h, I_v) of aquifer properties: the porosity ϕ, log hydraulic conductivity lnK, relative permittivity ε_r and electrical conductivity σ_b of the Krauthausen test site used for stochastic simulation within Facies 1-3.

	*ϕ	**lnK *[K in m/d]*	***ε_r *[-]*	****σ_b *[mS/m]*
Facies 1 - sand				
Mean	35. 8	-7.69	21	17.2
Variance	7.7	0.1	2.94	2
I_v [m]	0.13	0.6	0.13	0.13
I_h [m]	0.56	5	0.56	0.56
Facies 2 - sandy-gravel				
Mean	31.9	-6.4	18.6	12
Variance	3.6	0.6	1.24	1.75
I_v [m]	0.12	0.18	0.12	0.12
I_h [m]	0.39	1.75	0.39	0.39
Facies 3 - gravel				
Mean	25.7	-5.78	15.2	10.3
Variance	6	1	1.97	1.5
I_v [m]	0.12	0.4	0.12	0.13
I_h [m]	0.6	0.3	0.6	0.6

* The mean variance of the porosity was calculated from water content point measurements from neutron logs at CPT locations [Tillmann et al., 2008]. The correlation lengths I_h and I_v were

calculated (using semi-variogram analysis) after conversion from ε_r tomograms at multiple GPR FWI planes [Gueting et al., 2017, Zhou et al., 2021] using the CRIM model.

** Mean value of lnK was adopted from a K model based on grain size distribution (GSD) [Bialas and Kleczkowski, 1970, Gueting et al., 2017]. Variance and I_h of lnK were calculated based on a dense grid of vertical CPT with ~1.5 m horizontal separation distance [Tillmann et al., 2008], using a calibrated correlation between GSD and CPT geophysical properties (mechanical resistance, natural gamma activity and bulk density), where co-located data was available. I_v=0.18 m of lnK in the main Facies 2 of the model (green facie in Figure 1a) was adopted from borehole groundwater velocity measurement [Englert, 2003].

*** ε_r was calculated directly from the porosity model using CRIM model.

**** Mean value of σ_b and variance, I_h and I_v were calculated from σ_b tomograms at multiple GPR FWI planes [Gueting et al., 2017, Zhou et al., 2021].

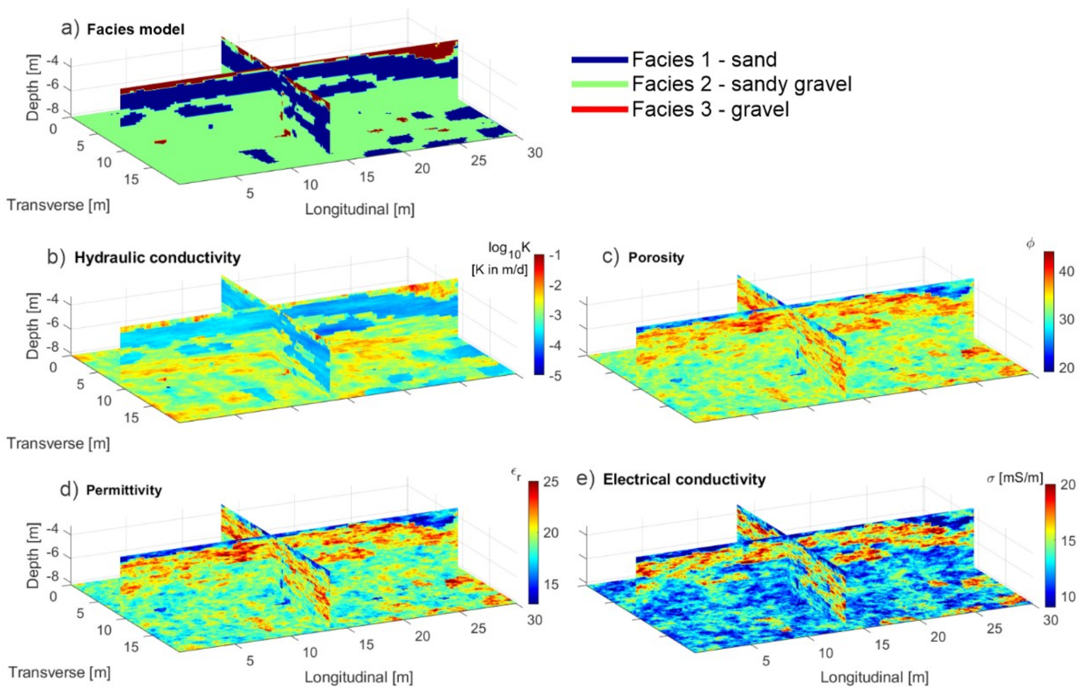


Figure 1. Aquifer model domains used for the hydrological flow and transport modeling. (a) Facies model, (b) log-conductivity, (c) porosity model, (d) background relative permittivity, and (e) electrical conductivity. The 3D facies model was adopted from Gueting et al. [2018].

2.2 Tracer transport simulation

We used a flow and transport model based on the designed aquifer model domain to imitate a past positive saline tracer test performed by Müller et al. [2010]. The results of this tracer test were used in our numerical study as a reference for the synthetic plume fate reconstruction. The 3D flow equation of the transport model was solved using TRACE [Vereecken et al., 1994] and the transport equation was solved using a random-walk particle-tracking algorithm PARTRACE [Bechtold et al., 2011]. We simulated a tracer injection for 7 days using a uniform water influx source of 20 m^3/day between 3.58 - 8.26 m depth in the borehole, and a particle injection source of $1.429 \cdot 10^7$ "conservative or non-reactive" particles injected per day, resulting in an injection concentration of $7.15 \cdot 10^5$ particles/m^3. We modelled the borehole (diam. 50.8 mm, slots 0.5 mm) by a vertical column of grid cells (cubic, edge of 0.09 m) assigned with K = 267 m/d [Klotz, 1990] and porosity of 1. The borehole was surrounded by a gravel pack that fills the well (diam. 0.328 m), modelled by 8 grid cell columns with K = 2246 m/d [Klotz, 1977] and porosity 0.4. To solve for the total head and velocity distributions in the heterogeneous aquifer, we adopted a natural hydraulic gradient in the aquifer of 0.002 m/m [Vereecken et al., 2000] implemented by pressure head boundary conditions at the up and downstream boundaries, and zero flux condition at the lateral, top and bottom boundaries. During the injection phase, we used the flow velocity field that was simulated considering the water injection in the well for the transport simulation. After the injection phase, this field was instantaneously shifted to the flow velocity field that was simulated for the natural background hydraulic gradient. The heterogeneity of the simulated plume was controlled by the stochastically-generated *lnK* and porosity (Table 1), which generated a variable fluid velocity. To account for the effect of velocity fluctuations on solute transport at the grid-cell scale, we used longitudinal and transverse dispersivities of 0.003 and 0.001 m, respectively.

The results of the transport simulation are shown in plan, side and front view (Figure 2a, c, d), for an exemplary snapshot at 15 days after the start of the injection. In each plane view, the distribution of mass represent the sum of particle mass along the perpendicular axis to that plane. A substantial part of the plume was transported over a large distance in the lower part of the aquifer (between 7 and 8 m depth), whereas a second part of the plume was moving slower between roughly 5 and 6.5 m depth. This was also apparent in breakthrough curves that showed an earlier plume arrival at 7.5 than at 5.5 m depth both at 10 and 20 m downstream of the injection well (Figure 2b). The plume is mainly transported in facies 2, which has a higher hydraulic conductivity than facies 1, whereas facies 3 (with the highest conductivity) was hardly present within the range of depths where the tracer was injected. Figure 3 illustrates the simulated tracer distribution in the monitoring plane at 15 days after the start of the injection, the facies distribution, and the distribution of a salt tracer that was imaged in this plane using ERT during a real tracer test carried out under the same conditions (injection well, injection rates) in the Krauthausen test site [Müller et al., 2010]. The simulated plume splitting corresponded with the observed one and the correspondence of the simulated and observed tracer distribution indicated that the reconstructed facies distribution represented the real distribution quite well. The simulated tracer distribution is characterized by thin horizontal lenses of 0.1 m thickness with high concentrations (e.g., at 6.2 m depth, Figure 3b), which corresponds with the vertical correlation length of I_v = 0.18 m of the hydraulic conductivity in facies 2. Note that the ERT images did not resolve these small-scale tracer concentration variations and the results are more smoothed (Figure 3c).

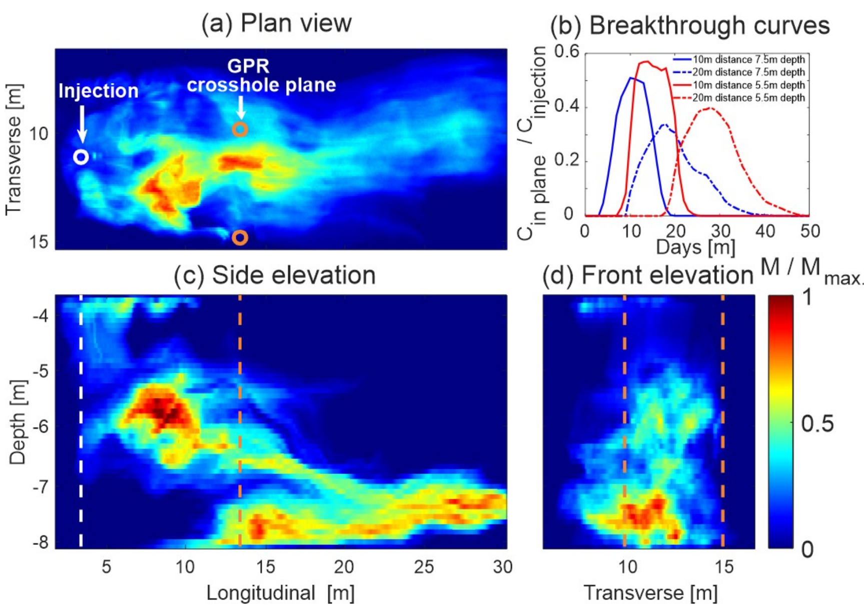


Figure 2. Normalized particles mass in plan (a), side (c) and front (d) view at day 15 after beginning of the tracer injection. The mass shown is a sum of mass at line of cells perpendicular to the view. Color map is normalized to the maximum mass for each view. Boreholes, which are used to derive the GPR data, are located 10 m down gradient from the injection well and are illustrated by orange circles and dashed lines. (b) Breakthrough curves at two down-gradient positions, 10 and 20 m, and at two depths, 5.5 and 7.5 m.

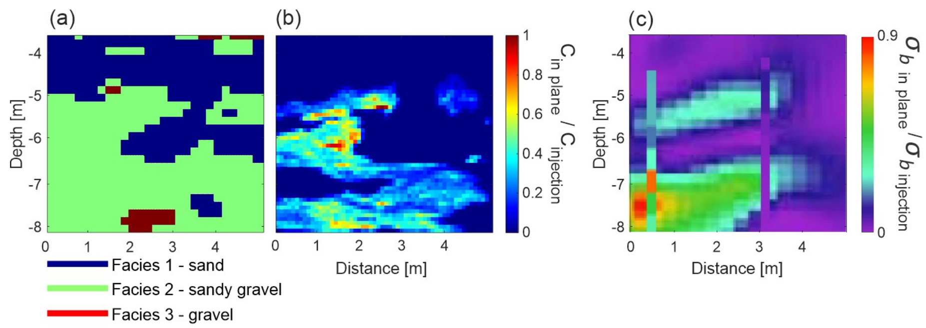


Figure 3. Tracer intrusion in the monitoring plane. (a) Facies model, (b) concentration distribution, and (c) ERT image (modified from Müller et al. [2010], Figure 5) for the domain of the plane between the two GPR monitoring boreholes 10 m downgradient from the injection borehole (see Figure 2). The concentration represented the distribution after day 15 from the transport simulation normalized to injected tracer distribution. The x-axis relates to the distance between boreholes (4.95 m). The electrical conductivity image is derived from the ERT tomogram and borehole loggers from a previous tracer test at day 15 after beginning the injection. Note that the logger data was obtained in two boreholes (seen as vertical anomalies) with vertical intervals of 0.35 m. The color scale represents the bulk electrical conductivity difference.

3 Tracer types and petrophysical relations

3.1 Change only in electrical conductivity: Salt and desalinated water tracer

The concentration of electrolytes in the groundwater determines the electrical conductivity of the fluid phase, whereas it has only minor influence on the permittivity [Sreenivas et al., 1995, Hagrey and Mülller, 2000]. For pore fluid conductivities that are smaller than ~15 S/m, the fluid electrical conductivity is proportional to the equivalent electrical charge concentration [Sreenivas et al., 1995]. Salt tracers with a higher concentration are not often imaged with GPR in transmission mode because of the high attenuation of the EM wave [exception in Day-Lewis, 2003].

3.1.1 Implementation of the salt tracer simulation

We simulated the solute transport using particle tracking and one particle was associated with a certain equivalent additional charge compared to the background charge concentration in the groundwater. If a tracer solution with a lower electrical conductivity than the background groundwater conductivity was injected (desalinated water), particles were associated with a 'negative' additional charge. Assuming that the electrical conductivity of the injected tracer solution $\sigma f\_tracer\_injection$ is constant during the injection and that the background fluid conductivity in the aquifer $\sigma f\_background$ is constant in space, the electrical conductivity of the fluid at time t in a grid cell centered at a 3D coordinate $\boldsymbol{x}$, $\sigma f\,(,t)$, was related to the number of particles in grid cell $np(\boldsymbol{x},t)$ at time t, the volume $Vcell$ and porosity $\boldsymbol{\phi}$ ($\boldsymbol{x}$) in the grid cell, the total number of particles injected $np,injection$, and the total volume of water that was injected $Vinjection$ as:

$$\sigma f(\boldsymbol{x},t) = (\sigma f\_tracer\_injection - \sigma f\_background) \cdot C(\boldsymbol{x},t) / Cinjection + \sigma f\_background \qquad \text{Eq. (1)}$$

with

$$C(\boldsymbol{x},t) = np(\boldsymbol{x},t)/(Vcell \cdot \boldsymbol{\phi}(\boldsymbol{x})) \text{ and} \qquad \text{Eq. (2)}$$

$$Cinjection = np\ injection/Vinjection\ . \qquad \text{Eq. (3)}$$

$C(\boldsymbol{x},t)$ and $Cinjection$ are the particle concentrations in a cell and in the injected tracer solution, respectively. The background pore fluid conductivity was equal to $\sigma f\_background$ = 93.7 mS/m. For the electrical conductivity of the injected salt tracer, we considered four cases:

I. Injection of water with an electrical conductivity smaller than the background (negative tracer, Desalinated case $\sigma f\_tracer\_injection$ = 69.6 mS/m),

II. Injection with a conductivity slightly higher (positive tracer) than the background (*Low* conductivity case: $\sigma f\_tracer\_injection$ = 117.8 mS/m),

III. Injection with an *Intermediate* conductivity ($\sigma f\_tracer\_injection$ = 610 mS/m), and

IV. Injection with a *High* conductivity ($\sigma f\_tracer\_injection$ = 1525 mS/m).

The Low electrical conductivity case adds the same magnitude of tracer fluid electrical conductivity that Desalinated subtracts, and the high case adds 2.5 times the tracer fluid conductivity of the Intermediate case. The background pore water conductivity and the negative and intermediate tracer conductivities were adopted from the tracer experiments carried out by Müller et al. [2010].

3.1.2 Salt tracer – electrical conductivity petrophysical relations

The bulk electrical conductivity σb at each cell of the grid is calculated using Archie's Law [Archie, 1942]:

$$\sigma b(\boldsymbol{x},t)=\sigma f(\boldsymbol{x},t)/F(\boldsymbol{x})+\sigma surf(\boldsymbol{x}), \qquad \text{Eq. (4)}$$

where $F(\boldsymbol{x})$ is the formation factor of saturated soil and $\sigma surf(\boldsymbol{x})$ is the surface conductivity. $F(\boldsymbol{x})$ is linked to the complex geometry of the pore channels and is smaller for a larger porosity and smaller tortuosity of the pore network [Archie, 1942; Jackson et al., 1978]. $\sigma surf(\boldsymbol{x})$ is controlled by the specific surface area, surface charge density, and effective ionic mobility in the electrical double layer around the charged surface [Johnson et al., 1986]. For low fluid conductivities, $\sigma surf$ depends in a non-linear way on the fluid conductivity σf. But, for sufficiently large σf, $\sigma surf$ reaches a constant value so that the relation between σb and σf is linear, which we assume further in this study.

We adopted an average $\sigma surf$ of 1.2 mS/m derived from laboratory experiments in packed columns with sediments from the Krauthausen aquifer (Müller et al., [2010], sampled in B70 at 5.5 - 6 m). Using this information, we generated a random field of $\sigma surf(x)$ using SGSIM of normally disturbed $\sigma surf \sim N$ (1.2,0.32) mS/m with the same correlation lengths as σb (Table 1), but with no spatial correlation between $\sigma surf$ and σb [Müller et al., 2010]. Then, we calculated $F(\boldsymbol{x})$ using Equation 4 with $\sigma f(\boldsymbol{x},t)= \sigma f\_background$. The distribution of the obtained $F(\boldsymbol{x})$ (Figure 8d) shows a range from 4.5 to 14.5 and is bimodal, reflecting the different distributions in the two main facies 1 and 2 (Figure 3a). The well sorted sand facies 1 has a mean porosity of $\boldsymbol{\phi}$1_mean=0.36 and a mean $F(\boldsymbol{x})$ of approximately 6. The sandy gravel facies 2 shows a mean porosity $\boldsymbol{\phi}$2_mean=0.31 and a mean formation factor of 8.5. The gravel facies 3 has a mean porosity $\boldsymbol{\phi}$3_mean= 0.25 and a mean $F(\boldsymbol{x})$ value is about 11, which is larger than the laboratory measured value for the disturbed samples of Müller et al. [2010] of 4.56 to 6.63 which excluded larger stones.

The flowchart in Figure 4 illustrates the steps to obtain the tracer concentrations from GPR FWI σ_b images. First, $F(\boldsymbol{x})$ is recovered from background GPR FWI σ_b (Figure 4a) using $\sigma_{f,background}$ and assuming a constant $\bar{\sigma}_{surf}$, which represents the average of $\sigma_{surf}(\boldsymbol{x})$ derived from lab measurements. Second, σ_f is estimated from time-lapse GPR FWI σ_b (Figure 4c). Last, the tracer $\sigma_{f,tracer}$ is calculated by subtracting $\sigma_{f,background}$ from σ_f.

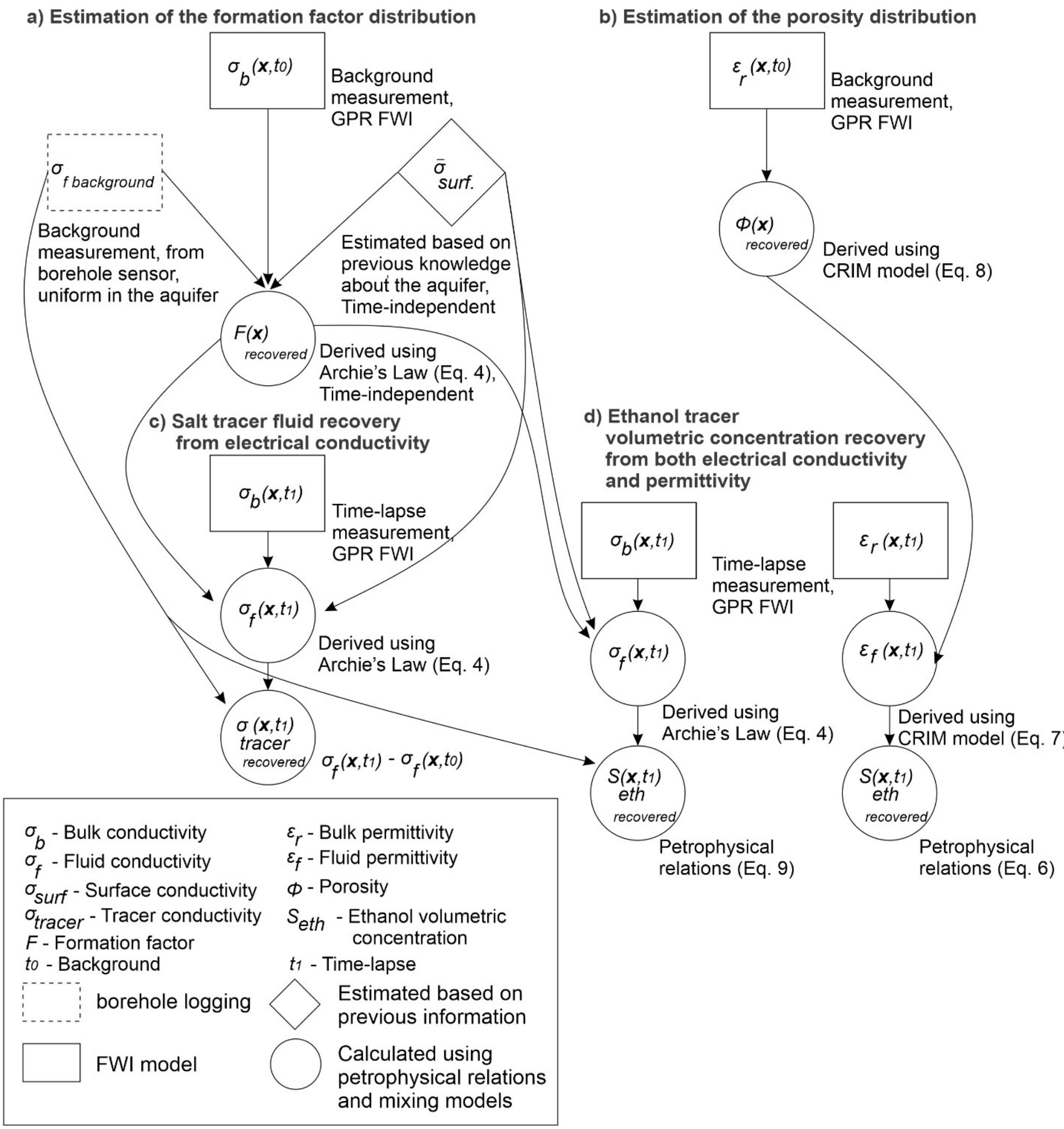


Figure 4. Flowchart presenting the recovery of solute (salt and desalinated water) and ethanol tracers. a) Formation factor and b) porosity recovery from background GPR FWI. c) Salt and d) ethanol recovery from time-lapse GPR FWI.

3.2 Change in permittivity and electrical conductivity: Ethanol tracer

Ethanol is commonly used as an additive in gasoline blends [McDowell et al., 2003, Spalding et al., 2011], and is currently treated as an emerging environmental contaminant [Gomez and Alvarez, 2009]. The dielectric properties of ethanol differ from water and these differences can be used to detect ethanol in water-saturated conditions in a sand matrix with GPR [Glaser et al., 2012]. Pure ethanol has a relative permittivity of 26.7 at 10 °C and an electrical conductivity of 0.025 mS/m [Petong et al., 2000, Glaser et al., 2012]. Note that the properties of the (ground)water at the Krauthausen test site at 10 °C are εr = 84 [Malmberg and Maryott, 1956] and σ ~ 90 mS/m [Müller et al., 2010]. Water-EtOH mixtures are miscible in all proportions as they are both dipolar liquids [Lide, 2004]. Ethanol experiences polarization relaxation at central frequency of about 1 GHz and dispersive behavior becomes effective from about f > 200 MHz, lower than those of water: 25 GHz and 1GHz, respectively [Petong et al., 2000]. Thus, dispersive behavior is expected for high GPR frequency ranges, but were not considered in study using low frequencies between 10-200 MHz with central frequency of 69 MHz. Regarding transport properties, ethanol has a lower density and a higher viscosity than water, and it is microbiologically degraded. However, we neglected density, degradation, and temperature effects on ethanol transport, because we focused on the ability to retrieve the distribution of the tracer from time-lapse GPR FWI parameter changes (εr, σb). Therefore, we simulated the ethanol plume migration with the same particle tracking method and using the same transport parameters (velocity, dispersivity) as the ones used for the salt tracer.

3.2.1 Implementation of the ethanol tracer simulation

We produced heterogeneous ethanol plumes that have the same structure as the salt tracer plumes. For the ethanol plume simulations, a particle represents a certain volume of ethanol *Veth*, and therefore the volumetric concentration of ethanol in a cell $Seth(\boldsymbol{x},t)$ is:

$$Seth(\boldsymbol{x},t) = Seth\ injection \cdot (np\ injection\ /\ Vinjection\ )\text{-}1 \cdot np(\boldsymbol{x},t)/\ (Vcell \cdot \boldsymbol{\phi}(\boldsymbol{x})). \qquad \text{Eq. (5)}$$

where *Seth injection* is the volume concentration of ethanol in the injected solution, which was considered to be 0.5. The injection volume *Vinjection* and the duration of the injection were identical to those of the saline tracer simulations.

3.2.2 Ethanol-permittivity petrophysical relations

An effective mixing model for permittivity of water ethanol mixtures, εr_f, at 10°C was derived from fitting a second-order polynomial to experimental data [Wyman, 1931] (Figure 5):

$$\varepsilon r\_f\,(\mathbf{x},t) = 84.05 - 42.6\cdot Seth(\mathbf{x},t) - 15.7\cdot Seth(\mathbf{x},t)2 \qquad \text{Eq. (6)}$$

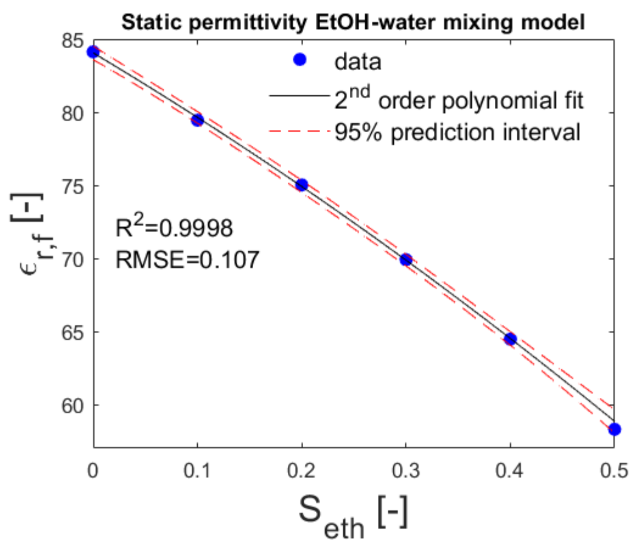


Figure 5. Relative permittivity $\varepsilon r,f$ of the fluid at 10°C as a function of ethanol volumetric concentration in ethanol-water mixture (adapted from Wyman [1931]).

To derive the bulk relative permittivity εr of the mixture-soil system, we used the Complex Refractive Index Model (CRIM) [Birchak et al., 1974]

$$\varepsilon r\,(\mathbf{x},t)=(\,\boldsymbol{\phi}(\mathbf{x})\sqrt{\varepsilon r,f}\,\,(\mathbf{x},t)\cdot + (1-\boldsymbol{\phi}(\mathbf{x}))\sqrt{\varepsilon r,s}\,)\;2 \qquad \text{Eq. (7)}$$

where $\varepsilon r,s = 4.5$ is the relative permittivity of the solid grains [Carmichael, 1988]. In order to retrieve $\varepsilon r,f$ and eventually $S(x,t)$ from a GPR FWI permittivity model (Figure 4d), the porosity must be recovered from εr background measurements (Figure 4b):

$$\boldsymbol{\phi}\mathrm{recovered}(\boldsymbol{x}) = (\sqrt{\varepsilon r(\mathbf{x},t0)}-\sqrt{\varepsilon r,s})\;/\;(\sqrt{\varepsilon r,f}-\sqrt{\varepsilon r,s}) \qquad \text{Eq. (8)}$$

3.2.2 Ethanol-conductivity petrophysical relations

We modeled the electrical conductivity of the ethanol-water fluid mixture σ_f with the Lichtenecker–Rother (L–R) model [Guéguen and Palciauskas 1994] by Personna et al. [2013]:

$$\sigma_f(\boldsymbol{x},t) = \left(S_{eth}(\boldsymbol{x},t) \cdot \sigma_{eth}^{\alpha} + (1 - S_{eth}(\boldsymbol{x},t)) \cdot \sigma_{f,background}^{\alpha}\right)^{1/\alpha} \qquad \text{Eq. (9)}$$

with $\alpha = 0.3$ for ethanol volumetric concentration $S_{eth} \leq 0.5$.

In order to retrieve $\sigma_f(\boldsymbol{x},t)$ and eventually $S_{eth}(\boldsymbol{x},t)$ from a GPR FWI bulk electrical conductivity model (Figure 4c), the formation factor must be recovered from $\sigma_b(\boldsymbol{x},t_0)$ background measurements (Figure 4a, Equation 2).

4 GPR modeling

Synthetic GPR data were calculated in a crosshole setup at 10 m distance from the injection borehole and perpendicular to the main flow direction (Figure 2, Figure 3a, b). The distance between the boreholes was 4.95 m. GPR data were derived between 3.2 - 10 m depth, which is below the water table (2.4 m depth). We added a realistic ambient noise level to the synthetic waveforms to evaluate its effect on the inversion performance (Appendix A1). To realistically include reflection and refractions of the GPR data, we describe the unsaturated zone above 2.4 m depth with $\varepsilon_r = 4.7$ [Daniels, 2004]. A semi-reciprocal acquisition setup was used with 35 transmitters and 69 receivers on each side, spaced 0.2 and 0.1 m, respectively, similar to previously performed measurements at the test site [Oberröhrmann et al., 2013]. With this setup a high ray coverage that improves the electrical conductivity reconstruction is obtained. We considered for our modeling a constant source wavelet (SW) with a central frequency of 69 MHz for the background and time-lapse cases (adopted from a previous FWI studies [Gueting et al., 2015]). It has been shown in experimental studies that for this operating frequency GPR FWI can obtain models with a vertical resolution as small as 0.2 m [Zhou et al., 2020].

5. GPR FWI results

For the inversion and forward modeling, we considered a cell size of the models of 0.09 m and 0.03 m, respectively. Note that the inversion grid has the same size of the of the transport simulation (Figure 3b). For the crosshole GPR FWI we considered the approach based on van der Kruk et al. [2015] that allows an update of the medium properties close to the boreholes and we follow the criteria given by Klotzsche et al. [2019b] to define the final inversion results.

5.1 Background models

The starting models (SM) for the FWI were derived from ray-based inversion results (Figure 6 b, g). The permittivity starting model is based on the first arrival time inversion. For the σ_b starting model, we considered a uniform σ_b of 15 mS/m based on the mean of the first-cycle amplitude inversion [Holliger et al., 2001]. The ε_r and σ_b results of FWI (Figure 6 c, d, h, i) of the noise-free and noise-added datasets visually show the same structures with decimeter-scale resolution and a smoothed recovery of the real input models. As expected, the inversion of the noise-free data performed slightly better than the noise added inversion (see Table 2 for performance evaluation), therefore from here on we consider only the noise-added dataset. Model errors (Figure 6e, j) are larger at locations of high contrasts. The RMSE of the background models are 0.86 for ε_r and 1.64 mS/m σ_b. The ε_r FWI models resolved the fine features better than the FWI σ_b models, as indicated in horizontal and vertical 1D profiles and by spectral analysis (Figure 7). The illumination of the domain using crosshole acquisition results in a better resolution of the vertical than the horizontal structures [Meles et al., 2010]. In vertical direction the ratio of the spectral densities of the FWI to RM starts decreasing for wavenumbers larger than υ = 3 m^{-1} and υ = 1.13 m^{-1} (equiv. to wavelength λ of 0.33 m and 0.88 m, $\lambda = 1/\upsilon$) for ε_r and σ_b, respectively. In the horizontal direction, this ratio starts increasing for wavenumbers larger than υ = 0.77 m^{-1} (λ = 1.3 m) and υ=0.51m^{-1} (λ = 2 m) for ε_r and σ_b, respectively.

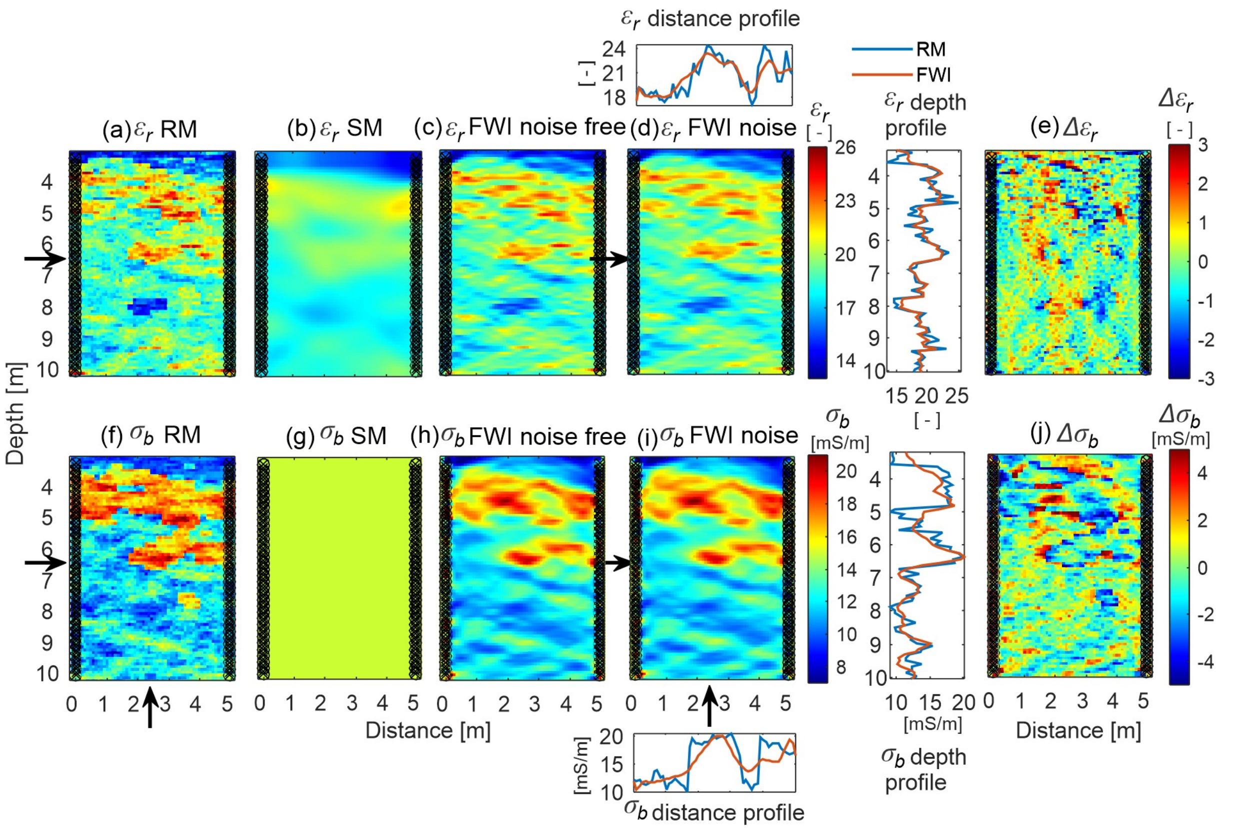


Figure 6. Background permittivity (top row) and electrical conductivity (bottom) models. (a, f) real models, (b, g) starting models based on ray-based inversion results, and (c, h) FWI models of the noise free dataset and (d, i) with noise. (e, j) Show the difference between real and FWI models with noise. Transmitter and receiver positions are located on black circles and crosses, respectively near the panel side boundaries. Above and right to (d), plots compare the real (blue) with FWI (orange) permittivity models (a, d) along vertical profile at 2.5 m distance and horizontal profile along 6.5 m depth (indicated by arrows). Below and right to (i), plots compare the real (blue) with FWI (orange) electrical conductivity models (f, i).

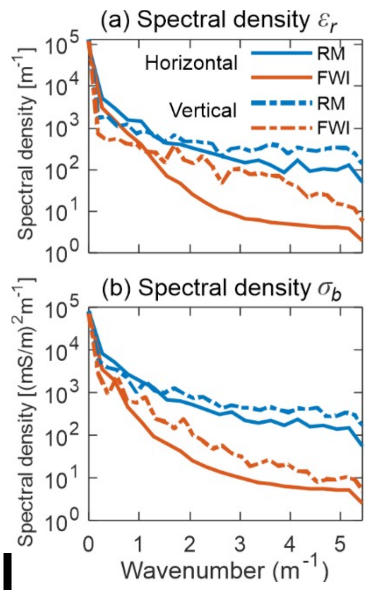


Figure 7. Spectral density of (a) permittivity and (b) electrical conductivity. Curves compare between the real models (RM, blue) and reconstructed FWI (orange) models for horizontal (line) and vertical (dashed line) directions.

Table 2: Performance evaluation of the FWI of the background data and models.

		Noise-free	Noise
GPR FWI data	Iteration[a]	67	59
	rms[b] $\cdot 10^{-7}$	1.26	1.42
	R^2	0.9986	0.9983
	MAE[c] $\cdot 10^{-7}$	4.94	5.41
	RMSE $\cdot 10^{-7}$ (% $RMSE_u$)[d]	1.77 (99.1)	1.96 (98.7)
FWI model ε_r	R^2	0.768	0.742
	MAE	0.659	0.714
	RMSE (%$RMSE_u$)	0.862 (72.8)	0.936 (66.5)
FWI model σ_b	R^2	0.648	0.698
	MAE [mS/m]	1.24	1.27
	RMSE [mS/m] (%$RMSE_u$)	1.64 (71.3)	1.68 (70.3)

[a] - Number of FWI iterations.
[b] - Root-mean squared error of the misfit function between real and modelled data
[c] - Mean absolute error.
[d] - See Appendix B Equations 16-19.

5.2 Porosity and formation factor estimation of the background models

The background FWI results (Figure 6 d, i) are used to derive the porosity $\phi recovered$ and the formation factor $Frecovered$ distributions using Equations 4 and 8 (Figure 8). $\phi recovered$ shows a better correlation with the real model porosity ϕRM than $Frecovered$ with FRM. The mismatch for both parameters is related to the unresolved structures and deviations between FWI and real ε_r and σ_b models (Table 2, Figure 6e, i). We see that low values of $\phi recovered$ (<0.28) overestimate ϕRM, while high values of $\phi recovered$ underestimate ϕRM (Figure 8c), which is a bias originating from FWI results. For higher values of $F_{\mathrm{recovered}}$ (>10) we can notice larger scatter. Locations with a high F correspond with locations where $\sigma_{b,bakground}$ is low. Errors in the recovered $\sigma_{b,background}$ and deviations between the local $\sigma surf$ and the mean $\sigma surf$, which is used to recover F, lead to a larger scatter for high values of $Frecovered$.

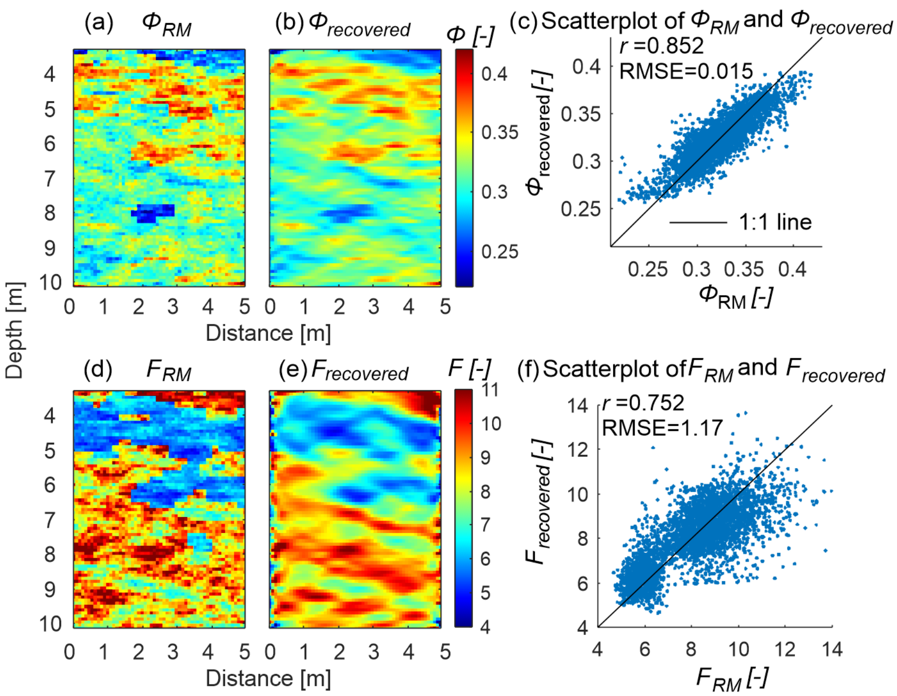


Figure 8. Porosity (top row) and formation factor (bottom row) distribution calculated from permittivity and electrical conductivity models, respectively: a) and d) show the real models, and, b) and e) the FWI models. Correlation plots and coefficient of correlation are presented in c) and f).

5.3 Time-lapse GPR FWI results of salt tracer

We chose to investigate day 15 after the tracer injection of the different tracer scenarios in the defined monitoring plane. Note that permittivity models are unchanged [Sreenivas et al., 1995] for all scenarios of the salt tracer test and therefore only the σb real models are used (Figure 9). However, in the simultaneous inversion nature of the FWI, the permittivity and conductivity updates are influenced and linked with each other, thus also the permittivity is updated.

For injections with Intermediate and High tracer conductivity, the bulk electrical conductivity distribution is predominantly determined by the distribution of the saline tracers, which generate larger variations in σb than spatial variations of $\sigma b\_background$ (Figure 9d and e). For the Desalinated (negative) and Low (positive) conductivity tracer, the changes in σb are in the same order of magnitude as the spatial variation in $\sigma b\_background$.

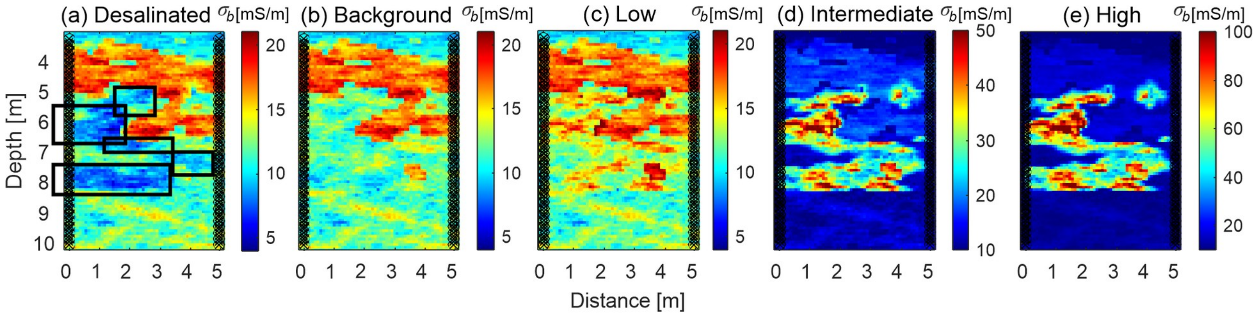


Figure 9. Bulk electrical conductivity real models in the monitoring plane for different cases of injected salt tracer concentrations: (a) Desalinated water, (b) Background, (c) Low, (d) Intermediate and, (e) High. The intrusion is shown for 15 days after the injection, and the main location of the tracer intrusion are emphasized by rectangles in (a). Note that panels (a-c) have the same colorbar scale, and (d) and (e) have different ones. Transmitter and receiver positions are located on black circles and crosses, respectively near the panel side boundaries.

We obtained time-lapse FWI models using two strategies of the starting models: *I*) using reconstructed FWI models of the background, and, *II*) reconstructed FWI models from a previous day. Note that we also investigated the use of ray-based starting models of each day, but the performance was less good than the other two strategies. We chose to use as a SM the FWI results from 2 days before (13 days after the tracer injection, Figure S1 in Supporting Information).

We show the difference in FWI σ_b models between day 15 and the background ($\Delta\sigma_b$) from noise-added data for the *Intermediate* tracer conductivity case (Figure 10a-c), together with the recovery of tracer fluid conductivity distribution, σ_{f_tracer} (Figure 10d-f) at depths within the transport model domain, from 3.58 to 8.26 m. The σ_{f_tracer} calculated from σ_b using petrophysical relation conversion which includes $F_{recovered}$ (Figure 8e) that contains an additive error (0.75 correlation between F_{FWI} and F_{RM}), recovered the structure of the plume (Figure 10f) and could resolve σ_b anomalies up to ~0.2-0.3 m in thickness with lower values than real one (Figure 10d), but shows slightly more smoothed structure than $\Delta\sigma_b$ (e.g. at location 7.5-8 m depth, compare with Figure 10c). An unexpected slightly better recovery evaluation for σ_{f_tracer} than $\Delta\sigma_b$ (higher R^2 in Figure 10, and other measures in Table 3) may be explained by the spatial distribution in the errors of σ_b background (see Figure 6j) and of day 15, and the distribution of the tracer concentration at the location of these errors. Inversions using the SM from FWI distribution from 2 days before, yielded slightly better inversion results than the SM using FWI of the background (Table 3), evaluated by a larger R^2 (0.728 compared with 0.698) and also visible by less smearing of thin preferential paths, e.g., at 5 m depth between 2 - 2.5 m distance. Subtracting σ_b at day 13 (FWI model and the recovered plume are shown Figure S1 in Supporting Information) from that in day 15 shows that time-lapse GPR FWI can follow the solute changes (Figure S2 in Supporting Information).

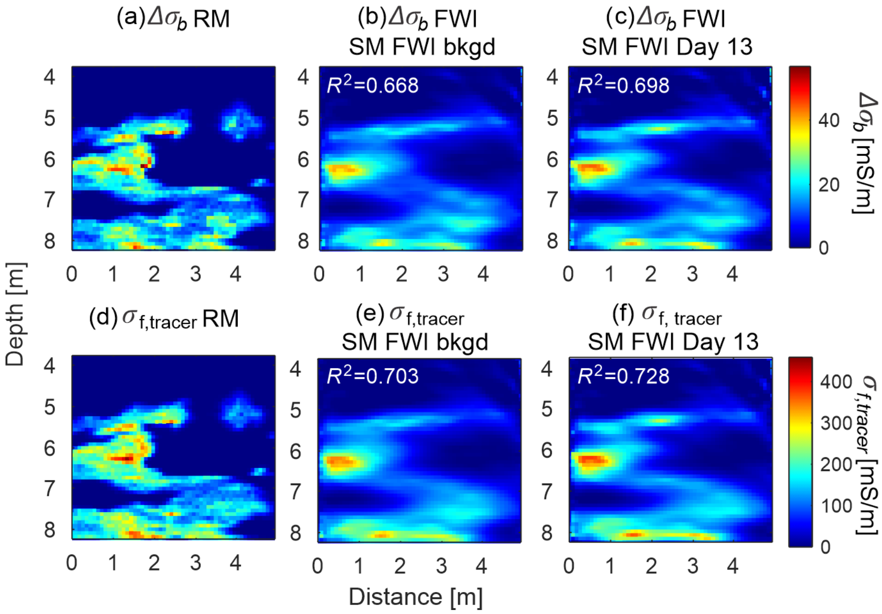


Figure 10. Tracer recovery of the intermediate salt tracer at day 15. Difference in bulk conductivity between day 15 and the background (top row) and tracer fluid conductivity (bottom) from (a,d) true model, (b,e) FWI of noise-added data with SM from FWI background, and FWI of noisy-data and (c,f) with SM from FWI day 13.

Table 3: Performance evaluation of FWI modelled data, FWI model parameters ε_r and σ_b, and tracer fluid conductivity of the salt tracer, σ_f, at day 15 for four cases: Desalinated (negative conductivity change), Low, Intermediate and High. For each scenario, results are given for noise-free and noise-added datasets using the background FWI as SM, and noise-included dataset using the day 13 FWI as SM.

		GPR FWI modelled data					FWI model $\Delta\sigma_b$		Tracer σ_f			
		Iter.	$rms \cdot 10^{-8}$	R^2	MAE $(\cdot 10^{-7})$	$RMSE \cdot 10^{-7}$ (%$RMSE_u$)	R^2	RMSE [mS/m] (%$RMSE_u$)	R^2	MAE [mS/m]	MBE [mS/m]	RMSE [mS/m] (%$RMSE_u$)
Desalinated	Noise-free	15	14.9	0.9987	0.609	2.11 (99.2)	0.65	0.864 (57.7)	0.704	5.07	0.43	6.85 (56)
	Noise	24	16.2	0.9986	0.648	2.24 (98.8)	0.629	0.89 (58.3)	0.687	5.22	0.454	7.05 (56.6)
	SM day 13 noise	32	15.2	0.9988	0.603	2.1 (98.9)	0.645	0.87 (63.5)	0.708	5.12	0.314	6.81 (61.9)
Low	Noise-free	32	10.5	0.9987	0.39	1.49 (99.3)	0.672	0.836 (71.1)	0.727	4.85	-0.577	6.59 (68.5)
	Noise	23	12.2	0.9983	0.441	1.67 (99.1)	0.655	0.858 (62.4)	0.703	4.84	-0.578	6.87 (61.8)
	SM day 13 noise	19	11.2	0.9987	0.39	1.47 (99.4)	0.643	0.873 (68)	0.697	5.13	-0.659	6.93 (64.7)
Intermediate	Noise-free	51	7.93	0.9985	0.243	1.12 (99.6)	0.703	5.91 (63.7)	0.734	32.1	-4.21	48.2 (65)
	Noise	36	9.52	0.9981	0.281	1.27 (99.1)	0.668	6.02 (61.8)	0.703	32.9	-4.43	48.3 (64.6)
	SM day 13 noise	34	8.76	0.9984	0.253	1.15 (99.2)	0.698	5.95 (64.8)	0.728	32.7	-4.42	48.8 (65.8)
High	Noise-free	66	8.8	0.9979	0.275	1.24 (98.8)	0.386	21.2 (49.7)	0.46	116	-14.6	172 (48.1)

	Noise	46	10.7	0.9971	0.323	1.45 (98.9)	0.294	22.8 (44)	0.397	125	-10.6	181 (41.7)
	SM day 13 noise	30	9.59	0.9977	0.284	1.28 (99.7)	0.363	21.6 (50)	0.447	117	-10.1	173 (47.6)

Recovery of FWI $\Delta\sigma_b$ model and tracer intrusion *σf_tracer* for the four tracer conductivity cases, using FWI models of two days before as SM, are shown in Figure 11a - d (only shown for the noise-added data), in comparison to the real models (Figure 9a and c-e). Desalinated, Low and Intermediate recovered FWI *σb* model cases show a good recovery of the plume structure, whereas the High case in general captures the structure, but contains internal errors inside the anomaly distribution. The best $\Delta\sigma_b$ and *σf_tracer* recoveries are obtained for the Intermediate case (shown by R^2 and *%RMSEu,* Table 3), while Desalinated and Low cases show lower performance. This is a consequence of the spatial variability in *σsurf(**x**)*, which is approximated by a constant value in the petrophysical relations (Equation 2) and has a larger impact on the *σf* recovery for lower *σb* values.

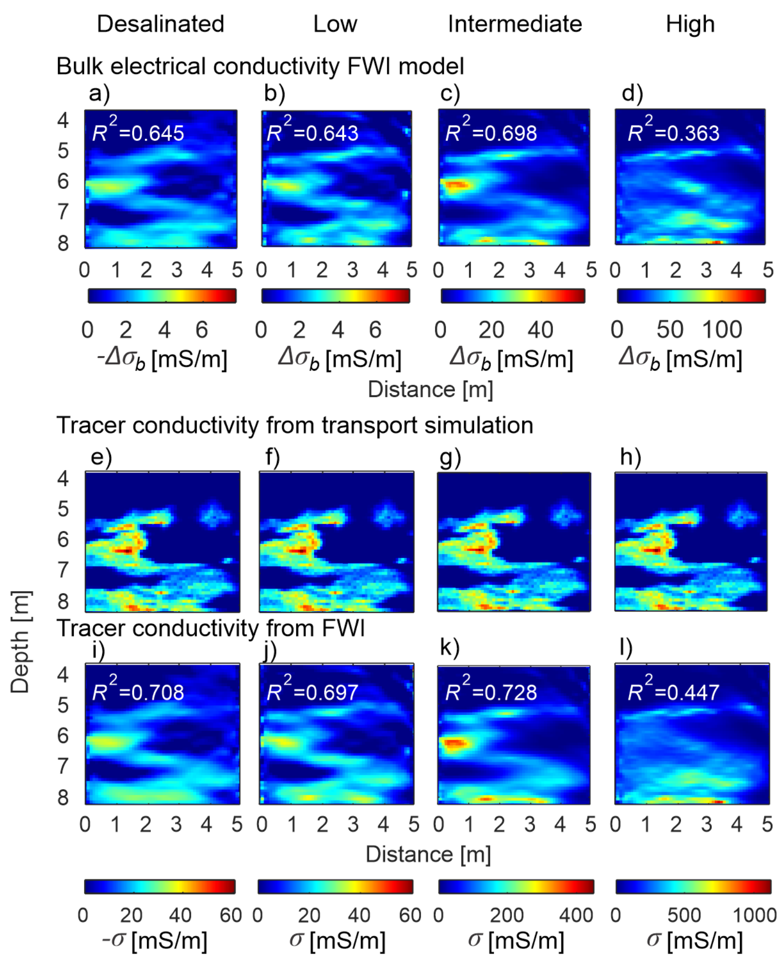


Figure 11. Recovered salt tracer distributions for four different conductivity magnitude cases from FWI (noise-added datasets) using FWI from day 13 as SM. Difference in FWI bulk conductivity between day 15 and the background, real and recovered tracer conductivity distribution are in top, middle and bottom rows, respectively. Note the different colorbars for each case including negative values colorbar for Desalinated case.

Investigating the data fit between the simulated and FWI modeled traces for the different scenarios (Figure 12), exemplary shown for one transmitter position at the depth of major tracer intrusion (6 m) and one with no intrusion (9.6 m), we notice generally a good overlap of the traces and that the FWI can resolve most details of the traces. A higher σ_b entails lower amplitudes causing gradually decreased amplitudes from Background to Intermediate cases (compare the *Observed* data for the transmitter shot gathers at 6 m of different cases). We observed a gradual decrease in R^2 between the true and inverted traces from Background to Intermediate cases, associated with decreasing signal-to-noise ratio. The models seems to have difficulties to represent the very large difference in signal amplitude in different regions (e.g. between 6 m and 9.6 m in Intermediate case) of the domain rather than to distinguish the signal from the noise. This is even more pronounced for High tracer case.

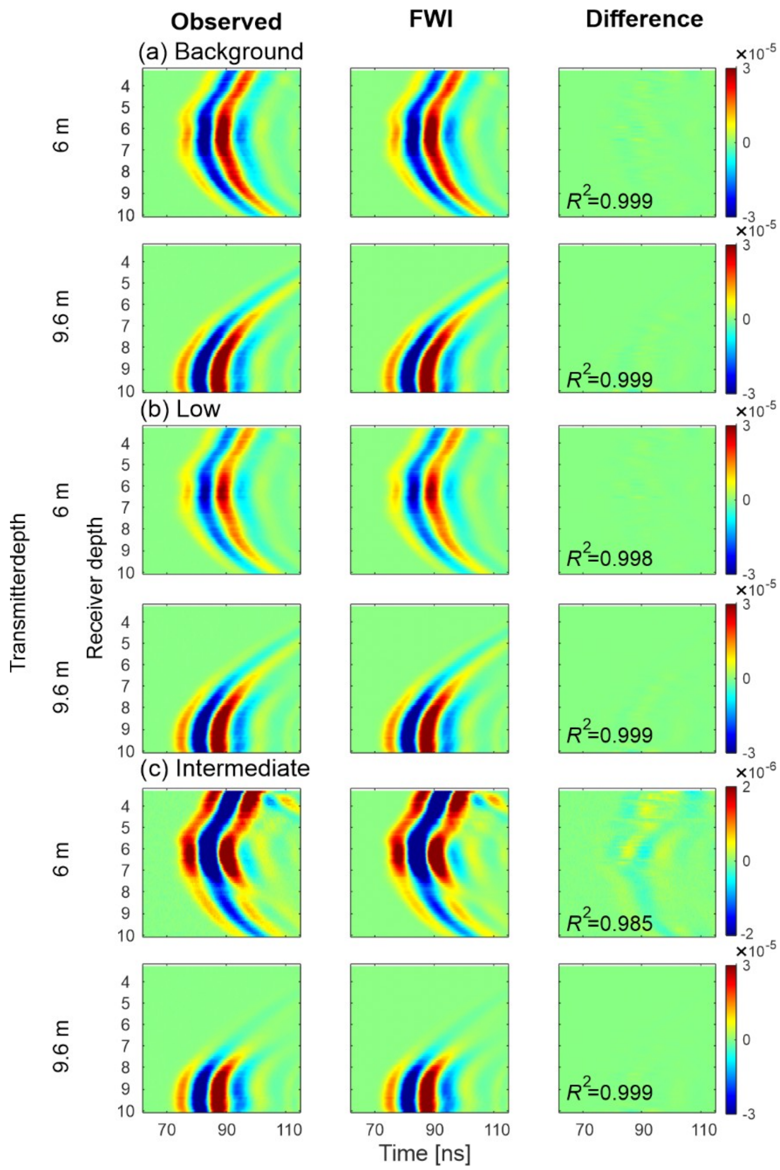


Figure 12. Observed (real data noise-added), FWI inverted and the difference between inverted and observed data for transmitters at the depth of major tracer intrusion (6 m) and at depth where no intrusion occurs (9.6 m). Data is presented for the (a) background (tracer case in Figure 9 b) and (b,c) Low and Intermediate (tracer case in Figure 9 c-d) tracer cases. Note that for the Intermediate case in (c) for the transmitter at 6 m depth panel where the signal is weaker because the wave travels through the increased σ of the tracer, the color scale is 15 times smaller. R^2 quantifies the correlation between FWI inverted and the observed data. The standard deviation of the Gaussian ambient noise was $4.6{\cdot}10^{-8}$ in all cases.

Figure S4 (in Supporting Information) shows observed and FWI traces of Intermediate and High cases for ray paths travelling parallel to the surface. At the central part of the plume, the High FWI waveform at 6 and 7 m show a bad fit (Figure S4 f,h), with erroneous amplitudes and phase shifts. The FWI traces with a large misfit are related to a *i)* a further decrease in signal-to-noise ratio, and *ii)* large σ_b parameter anomalies in combination with trade-offs between ε_r and σ_b on trace attenuations that result in local optima in the objective function far from the true optimum where local and gradient based optimization schemes converge to [Klotzsche et al., 2019b].

5.4 Time-lapse GPR FWI of ethanol tracer

The real models of ε_r and σb in the monitoring plane 15 days after the ethanol tracer injection show a decrease in both ε_r and σ_b (Figure 13a, g) compared with the background (Figure 6a, e), with maximum changes of $\Delta\varepsilon r$ = -3.35 and $\Delta\sigma b$ = -11.95 mS/m at 6.3 m depth. This corresponds to the GPR traces for the ethanol tracer distributions with increased amplitude and which are shifted to earlier times by about 1-1.5 ns (examples of ethanol FWI and corresponding RM traces at the main intrusion depth are shown in Figure S5). We used the same SM strategies as for the salt tracer. Difference in FWI permittivity between day 15 and the background ($\Delta\varepsilon_r$) models using the two different SM show (Figure 13b, c compared with Figure 13a) FWI using a SM of a reconstructed FWI model from a previous day is better than using the background SM (Table 4). FWI $\Delta\sigma_b$ recovery (Figure 13h, i compared with Figure 13g) is more smoothed than that of ε_r, and is related to a higher sensitivity to fit the phase than to fit the amplitude. Modelled FWI traces (blue dashed lines in Figure S5) for the ethanol case show a good fit to the observed traces.

Volumetric concentration of ethanol S_{eth} distributions calculated from ε_r (Figure 13e, f compared with Figure 13d) using SM from FWI of a previous day recovers a sharper distribution, a finer size of the preferential paths, and a better reconstruction of the quantitative values than SM from FWI of the background (Table 4). A more accurate recovery of S_{eth} distribution is derived from ε_r rather than from σ_b (Figure 13k, l compared with Figure 13j) because of the better recovery of FWI ε_r, and due to uncertainty in $\sigma_{\mathrm{surf}}(x)$ that propagates in the derivation of S_{eth} from σ_b. In addition, since the uncertainty of $\boldsymbol{\Phi}_{recovered}$ is smaller than that of $F_{recovered}$ (Figure 8), less errors propagate in the recovery of S_{eth} from ε_r than from σ_b.

Table 4: Performance evaluation of FWI modelled data at day 15, difference in permittivity ($\Delta\varepsilon_r$) and conductivity ($\Delta\sigma_b$) between day 15 and the background, and the volume concentration of the *ethanol* tracer at day 15. Report is on FWI datasets with 1) SM background noise-free, 2) SM background noise added, and 3) SM using day 13 noise added.

	GPR FWI modelled data					**FWI model $\Delta\varepsilon_r$**		**FWI model $\Delta\sigma_b$**	
	Iter	*Rms $\cdot 10^{-8}$*	R^2	*MAE ($\cdot 10^{-7}$)*	*RMSE $\cdot 10^{-7}$ (%RMSE$_u$)*	R^2	*RMSE [-] (%RMSE$_u$)*	R^2	*RMSE [mS/m] (%RMSE$_u$)*
SM bkgd	37	20.4	0.9985	0.805	2.71 (98.8)	0.777	0.3 (65.8)	0.629	1.55 (67.9)

noise-free									
SM bkgd noise	31	22.2	0.998	0.909	3.11 (99.6)	0.712	0.34 (58.4)	0.61	1.59 (62.8)
SM day 13 noise	31	18.9	0.9986	0.778	2.64 (99)	0.776	0.3 (78.4)	0.636	1.53 (61.9)

	Tracer S_{eth} from ε_r				Tracer S_{eth} from σ_b			
	R^2	*MAE*	*MBE*	*RMSE [-] (%RMSE$_u$)*	R^2	*MAE*	*MBE*	*RMSE (%RMSE$_u$)*
SM bkgd noise-free	0.806	0.0284	0.0019	0.04 (59.3)	0.663	0.0386	-0.0067	0.054 (57.9)
SM bkgd noise	0.747	0.0318	0.0014	0.0469 (53.7)	0.642	0.0395	-0.0083	0.0557 (51.6)
SM day 13 noise	0.812	0.0281	0.0043	0.0404 (69.5)	0.668	0.038	-0.0073	0.0536 (48.3)

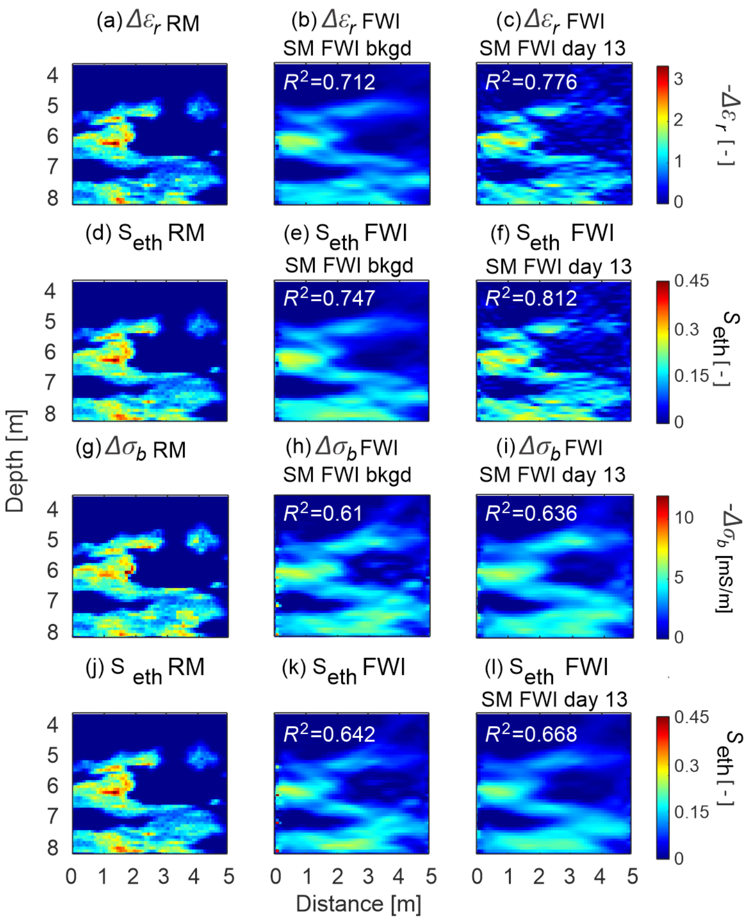


Figure 13. Tracer recovery of ethanol at day 15. Top row: difference in permittivity between day 15 and the background, second row: ethanol volumetric concentration from permittivity, third row: difference in bulk conductivity between day 15 and the background, bottom row: ethanol volumetric concentration from permittivity. Left column: the real models, middle column: the recovered FWI models using SM of background FWI, right column: the recovered FWI models using SM of day 13 FWI. Note that the change in both permittivity and conductivity is negative.

Further, subtracting ε_r at day 13 (FWI model of the recovered plume are shown Figure S6b in Supporting Information) from that in day 15 (Figure 14a,b) shows that time-lapse GPR FWI can image the solute concentration changes based on time lapse ε_r images with about 0.2 m resolution, and better than from time lapse σ_b images (Figure 14c,d, after subtracting ε_r in day 13 in Figure S6f from that in day 15 in Figure 13i).

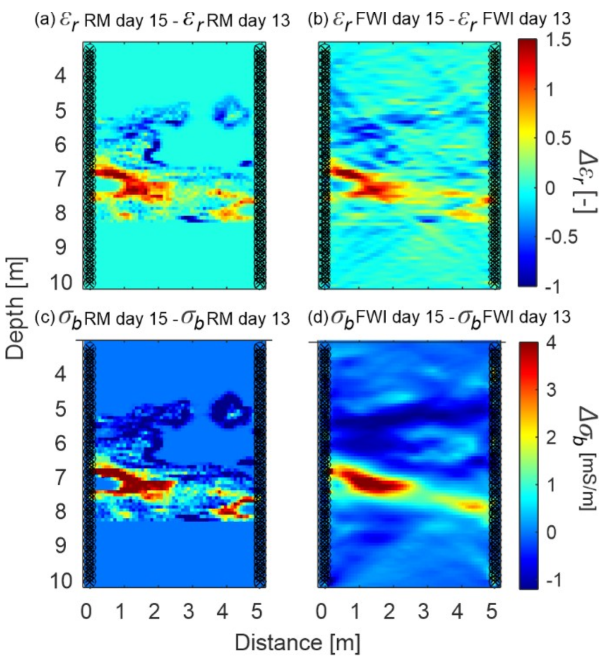


Figure 14. Changes between day 15 and day 13 based on permittivity: a) Real and b) FWI recovered models, and based on bulk conductivity: c) Real and d) FWI recovered models.

6 Conclusions and outlook

In this study, we tested the reconstruction of a tracer plume with crosshole GPR FWI in a numerical experiment. Realistic dimensions of the plume and the electrical aquifer properties influenced by it were derived from available aquifer and tracer test datasets from previous studies at the Krauthausen test-site. We used this information to generate aquifer and transport simulations. We tested the GPR FWI to reconstruct the plume from a noise-free and a noise-added dataset for a saline tracer, which changed σ_b, and an ethanol tracer, which changed both ε_r and σ_b. We retrieved petrophysical models for ethanol for conversion between ethanol concentrations and relative permittivity and electrical conductivity. For the salt tracer experiment, large increase in σ_b from tracer intrusion can cause trade-offs between ε_r and σ_b as they both depend on the GPR trace amplitude. Plume reconstruction of ethanol from GPR FWI crosshole permittivity changes showed an improvement compared to reconstructions from electrical conductivity changes, because GPR data is intrinsically more sensitive to ε_r than to σ_b anomalies [Lavoué et al., 2014]. Fine plume fingers with a thickness of ~ 0.2 m could be resolved by GPR FWI from permittivity changes.

Our research showed that the selection of the starting model is important to adequately recover time-lapse FWI models. A starting model from a previous day FWI recovered model was found to perform better than using the FWI of the background. This however depends on the magnitude of the tracer concentration changes. It requires further investigation how such a time lapse approach can be designed optimally. Too large differences between the starting model and the recovered model should be avoided, to fulfil the FWI requirement of a half-wavelength criteria [Klotzsche et al., 2019b]. Note that when consecutive distributions do not differ enough, changes may not be detected correctly. It is unclear how these errors will propagate through a series of distributions that do not differ a lot consecutively, but which depict significant changes over longer times.

For field experiments, the quality of the recovered distributions of electrical aquifer properties from GPR FWI is reduced with sparse data sampling, and increased noise level in the data, which is affecting the electrical conductivity results more [Oberröhrmann et al., 2013]. We carefully need to analyze the influence of the different processing steps such as data processing [Peterson, 2001, Axtell et al., 2016], time zero correction [Oberröhrmann et al., 2013], 3D to 2D data conversion [Mozaffari et al., 2020] and the estimation of the unknown source wavelet, which is affected by different borehole fillings [Klotzsche et al., 2013, Klotzsche et al., 2019b]. A crucial point will be the nature of the tracer plume being 3D [Vanderborght et al., 2005], when the GPR wave travels in 3D but the inverted crosshole tomogram is in 2D.

As an outlook, by utilizing the high-resolution of ε_r, GPR FWI has a potential to monitor temperature changes [Seyfried and Grant, 2007]. Thus, heat can be used as tracer to investigate zones of low hydraulic conductivity, like immobile water regions [Dassargues, 2018] and thin rock fractures where heat is smeared over a volume larger than that influenced by the fluid in the fracture [de La Bernardie et al., 2018].

APPENDIX A

Noise in GPR synthetic data

GPR data are contaminated with ambient white noise that originated from the electronics of the transmitter, receiver and cables [Annan, 2003]. In order for events in the data to be detectable, the power at the receiver must be in excess of the noise level at the receiver. At the presence of a salt tracer, the signal power at the receiver will lose energy, and events in the data may be overshadowed by the ambient noise. Eventually, this will lead to incorrect FWI modelled data and FWI reconstructed parameter models. Other sources of uncertainty in the reconstructed FWI models that originated from the instrument time drift, antennae spatial positioning [Axtell et al., 2016], the effect of the effective source wavelet [Belina et al., 2012] and data processing [Peterson, 2001] are neglected in this study. Nevertheless, in GPR datasets the quality of recorded data and the acquisition setup differ and each of the sources of errors may be the one of the largest impact on the quality of the reconstructed model, in this study we concentrate on the effect of the ambient noise.

We added a realistic ambient noise level to the synthetic waveforms to evaluate its effect on the inversion performance. We assumed that the ambient noise level (of the tracer before the first rise) is independent of the conductivity of the tracer. This then leads to a different *signal* to noise ratio for different configurations. We obtained realistic ambient noise levels from real GPR traces that were acquired from the Krauthausen site using the same cross borehole distance [Gueting et al., 2015]. Based on this data, we calculated a relative ambient noise level with:

$$Ambient\ noise\ ratio\ min.[\%]=\frac{std.(noise\ recorded\ traces)}{max.|A(recorded\ traces)|}\cdot 100\,\%\,, \qquad \text{Eq. (10)}$$

where $std.(noise\ recorded\ traces)$ is the standard deviation of the ambient noise in recorded traces, which was calculated from amplitudes at the time range before the wave first arrival time, and $max.|A(recorded\ traces \in entire\ dataset)|$ is the maximum absolute amplitude of the recorded traces over the entire dataset.

From Equation 10, the minimal ambient noise ratio is 0.062%. For short distances (4.95 m, parallel ray path) the signal dominates over the noise (Figure A1 a,c), with typical noise level ~ 0.2%. For a typical wide-angle long-distance ray path trace (8.4 m, 54^0) the noise has a larger footprint in the recorded trace (Figure A1 6e, g) with noise level ~ 3%.

Then, we generated the absolute ambient noise as Gaussian random error $G\left(0, std.^2_{noise \in synthetics}\right)$ with a standard deviation based on the $Ambient\ noise\ ratio\ min.$

$$std._{noise \in synthetic\ traces}=\frac{max.|A(noise\ free\ simulated\ traces)|\cdot Ambient\ noise\ ratio\ m}{100\,\%} \qquad \text{Eq. (11)}$$

where $max.|A(noise\ free\ simulated\ traces)|$ is the maximum absolute amplitude of the simulated traces over the entire dataset before adding noise.

Finally, we added the same absolute ambient noise to the synthetic simulated waveforms of the background, salt and ethanol tracer cases:

$$data(noisy\ added\ simulated\ traces)=data(noise\ free\ simulated\ traces)+G\left(0, std.^2\right. \qquad \text{Eq. (12)}$$

The noise-added traces (in blue) eventually show a similar ambient noise level to the real traces for short and long ray paths (see Figure A1).

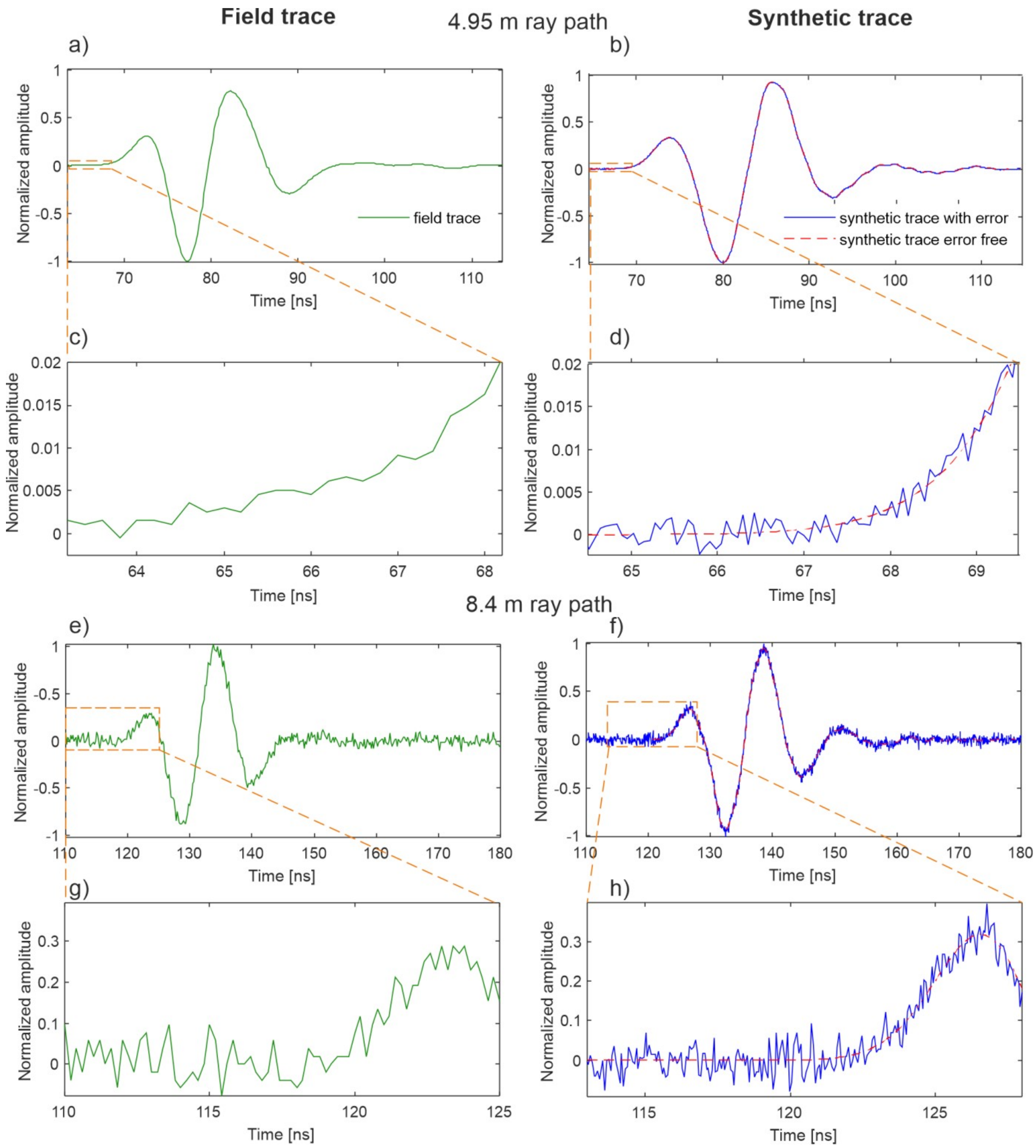


Figure A1: Realistic noise added to synthetic waveforms based on ambient noise from experimental data. Examples for two different waveforms: a - d) the shortest ray path of 4.95 m and e-h) 54° wide-angle ray path of 8.4 m. Panels c, d, g, h zoom in to view the noise. Note the scale is a normalized amplitude to the maximal absolute amplitude of the same trace. Sampling rate of the experimental GPR data is 0.2 ns, whereas the sampling rate of the synthetic data is 0.063 ns.

APPENDIX B

FWI model performance evaluation:

In a synthetic study the information about the true subsurface model allows to evaluate the results from GPR FWI. To evaluate the *correlation* between the true subsurface tomogram (***x***) in the space domain and the FWI modelled tomogram (***y***) the coefficient of determination R^2 is used:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - x_i\right)^2}{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2} \quad \text{Eq. (13),}$$

where n is the number of variables in the tomogram, x_i and y_i are the variables in the true and modelled tomograms, respectively, and $\overline{x}$ is the true tomogram mean. The coefficient of determination R^2 provides the percentage of the total variation in the FWI modelled tomogram ***y*** that can be explained by the theoretical linear relationship ***y***=***x*** between the FWI modelled and true tomograms.

However, correlation measure like R^2 is not consistently related to the *accuracy* and *precision* of prediction model, i.e. the degree that the predicted variables approach the *magnitude* and a *linear function* of their true variables counterpart, respectively [Willmott, 1982]. The evaluation is also based on the difference measures between counterpart variables. Therefore, bias can be described by the mean bias error (MBE):

$$MBE = n^{-1} \sum_{i=1}^{n} \left(x_i - y_i\right) \quad \text{Eq. (14).}$$

Average difference is quantified by the mean absolute error (MAE) or the root mean square error (*RMSE*):

$$MAE = n^{-1} \sum_{i=1}^{n} \left|x_i - y_i\right| \quad \text{Eq. (15).}$$

$$RMSE = \left[n^{-1} \sum_{i=1}^{n} \left(x_i - y_i\right)^2 \right]^{0.5} \quad \text{Eq. (16).}$$

Nevertheless, RMSE does not illuminate the type of errors. Systematic and unsystematic types of errors with respect to the expected linear relation between x_i and y_i, are quantified by the systematic and unsystematic root mean squared errors ($RMSE_s$ and $RMSE_u$):

$$RMSE_s = \left[n^{-1} \sum_{i=1}^{n} \left(\hat{y}_i - x_i \right)^2 \right]^{0.5} \quad \text{Eq. (17)}$$

$$RMSE_u = \left[n^{-1} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2 \right]^{0.5} \quad \text{Eq. (18),}$$

where $\hat{y}_i$ is calculated from the ordinary least-squares (OLS) regression $\hat{y}=a\boldsymbol{x}+b$ between the FWI modelled and true tomograms.$RMSE_s$ describes the linear bias produced by the FWI tomogram and $RMSE_u$ is used to interpret the precision of the FWI tomogram.

Equations A3 and A4 represent all partitions of the errors in $RMSE$:

$$RMSE^2 = RMSE_s{}^2 + RMSE_u{}^2 \quad \text{Eq. (19).}$$

The relative fraction of $RMSE_s{}^2$ and $RMSE_u{}^2$ in $RMSE^2$ is then used to estimate the extent of each type of error in the FWI tomogram.

Large-range persistence of processes in the space domain can be analyzed in the spectral domain using power spectral density (PSD). It measures the wavenumber content of a process. Detailed technical explanation of calculation of PSD can be found in Witt and Malamud. A process can be defined with persistence in the space domain if for a range of wavenumbers, PSD depends in a power law on wavenumber υ:

$$\mathrm{PSD}(\upsilon) \sim \upsilon^{-\beta} \quad \text{Eq. (20),}$$

where parameter β is the strength of range persistence.

Acknowledgments

This study is part the Enigma ITN program (European training Network for in situ imaGing of dynaMic processes in heterogeneous subsurfAce environments). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No 722028. The measurements we used to generate the aquifer model were funded by the “TERrestrial ENvironmental Observatories” (TERENO). The computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC). Hydrological and geophysical datasets from Krauthausen test site are open in public domain of the TERENO database (file identifier ad404c9f-419a-4b14-b6e0-6ee9acd8f80e).

References

Annan, A. P. (2009). Electromagnetic principles of ground penetrating radar. *Ground penetrating radar: theory and applications*, *1*, 1-37.

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

Asnaashari, A., Brossier, R., Garambois, S., Audebert, F., Thore, P., & Virieux, J. (2015). Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: which strategy?. Geophysical prospecting, 63(1), 78-98.

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

Bechtold, M., Vanderborght, J., Ippisch, O., & Vereecken, H. (2011). Efficient random walk particle tracking algorithm for advective−dispersive transport in media with discontinuous dispersion coefficients and water contents. Water Resources Research, 47(10).

Bialas, Z., and A. Kleczkowski (1970), Über den praktischen Gebrauch von einigen empirischen Formeln zur Berechnung des Durchlässigkeitskoeffizienten, K. Archivum Hydrotechniki, Warschau.

Belina, F., Irving, J., Ernst, J., & Holliger, K. (2012). Analysis of an iterative deconvolution approach for estimating the source wavelet during waveform inversion of crosshole georadar data. *Journal of applied geophysics*, *78*, 20-30.

Binley, A., Hubbard, S. S., Huisman, J. A., Revil, A., Robinson, D. A., Singha, K., & Slater, L. D. (2015). The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. *Water resources research*, *51*(6), 3837-3866.

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

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

Carmichael, R. S. (1988). Practical Handbook of Physical Properties of Rocks and Minerals . CRC press.

Cheng, A. H., & Bear, J. (2016). Modeling Groundwater Flow and Contaminant Transport. Springer Publishing Company, Incorporated.

Dafflon, B., Irving, J., & Barrash, W. (2011). Inversion of multiple intersecting high-resolution crosshole GPR profiles for hydrological characterization at the Boise Hydrogeophysical Research Site. *Journal of Applied Geophysics*, *73*(4), 305-314.

Daniels, D. J. (2004). Ground Penetrating Radar, the Institution of Electrical Engineers, *London, UK*.

Day−Lewis, F. D., Lane Jr, J. W., Harris, J. M., & Gorelick, S. M. (2003). Time−lapse imaging of saline−tracer transport in fractured rock using difference−attenuation radar tomography. Water Resources Research, 39(10).

Dassargues, A. (2018). *Hydrogeology: groundwater science and engineering*. CRC Press.

de La Bernardie, J., Bour, O., Le Borgne, T., Guihéneuf, N., Chatton, E., Labasque, T., Le Lay, H. & Gerard, M. F. (2018). Thermal attenuation and lag time in fractured rock: Theory and field measurements from joint heat and solute tracer tests. Water Resources Research, 54(12), 10-053.

Englert, A., Hashagen, U., Jaekel, U., Nitzsche, O., Schwarze, H., & Vereecken, H. (2000). Transport von gelösten Stoffen im Grundwasser–Untersuchungen am Testfeld Krauthausen. *Grundwasser*, *5*(3), 115-124.

Englert, A. (2003). Measurement, estimation and modelling of groundwater flow velocity at Krauthausen test site (Doctoral dissertation, Bibliothek der RWTH Aachen).

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

Friedman, S. P. (1998). Simulation of a potential error in determining soil salinity from measured apparent electrical conductivity. Soil Science Society of America Journal, 62(3), 593-599.

Glaser, D. R., Werkema, D. D., Versteeg, R. J., Henderson, R. D., & Rucker, D. F. (2012). Temporal GPR imaging of an ethanol release within a laboratory-scaled sand tank. Journal of Applied Geophysics, 86, 133-145.

Gomez, D. E., & Alvarez, P. J. (2009). Modeling the natural attenuation of benzene in groundwater impacted by ethanol−blended fuels: Effect of ethanol content on the lifespan and maximum length of benzene plumes. *Water Resources Research*, *45*(3).

Guéguen, Y., & Palciauskas, V. (1994). *Introduction to the Physics of Rocks*. Princeton University Press.

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

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

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

Gueting, N., Caers, J., Comunian, A., Vanderborght, J., & Englert, A. (2018). Reconstruction of three-dimensional aquifer heterogeneity from two-dimensional geophysical data, Mathematical Geosciences, 50(1), 53-75.

Hagrey, A. (2000). GPR study of pore water content and salinity in sand [Link]. *Geophysical Prospecting*, *48*(1), 63-85.

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

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

Huisman, J. A., Hubbard, S. S., Redman, J. D., & Annan, A. P. (2003). Measuring soil water content with ground penetrating radar: A review. Vadose zone journal, 2(4), 476-491.

Jackson, P. D., Smith, D. T., & Stanford, P. N. (1978). Resistivity-porosity-particle shape relationships for marine sands. *Geophysics, 43(6)*, 1250-1268.

Johnson, D.L., Koplik, J., Schwartz, L.M., 1986. New pore-size parameter characterizing transport in porous media. Phys. Rev. Lett. 57, 2564–2567.

Kemna, A., Vanderborght, J., Kulessa, B., & Vereecken, H. (2002). Imaging and characterisation of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models. Journal of Hydrology, 267(3-4), 125-146.

Kemna, A., Binley, A., & Slater, L. (2004). Crosshole IP imaging for engineering and environmental applications. Geophysics, 69(1), 97-107.

Kelter, M., Huisman, J. A., Zimmermann, E., & Vereecken, H. (2018). Field evaluation of broadband spectral electrical imaging for soil and aquifer characterization. *Journal of applied geophysics*, *159*, 484-496.

Klotz, D., Berechnete Durchlässigkeiten handelsüblicher Brunnenfilterrohre und Kunststoff-Kiesbelagfilter, Tech. rep., GSF, 1990, GSF-Bericht Nr. 35.

Klotz, D., Berechnung der Filtergeschwindigkeit einer Grundwasserströmung aus Tracerverdünnungsversuchen in einem Filterpegel, Tech. rep., GSF, 1977, GSF-Bericht R149.

Klotzsche, A., van der Kruk, J., Meles, G., & Vereecken, H. (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., Doetsch, J., & Vereecken, H. (2013). 3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection. *Geophysical Journal International*, *195*(2), 932-944.

Klotzsche, A., Lärm, L., Vanderborght, J., Cai, G., Morandage, S., Zörner, M., Vereecken, H., & van der Kruk, J. (2019a). Monitoring Soil Water Content Using Time-Lapse Horizontal Borehole GPR Data at the Field-Plot Scale. Vadose Zone Journal, 18(1).

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

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

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

Lide, D. R. (Ed.). (2004). CRC handbook of chemistry and physics (Vol. 85). CRC press.

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

Malmberg, C. G., & Maryott, A. A. (1956). Dielectric Constant of Water from 00 to 1000 C. *Journal of research of the National Bureau of Standards*, 56, 1.

Maliva, R. G. (2016). *Aquifer characterization techniques*. Berlin: Springer.

McDowell, C. J., Buscheck, T., & Powers, S. E. (2003). Behavior of gasoline pools following a denatured ethanol spill. *Ground Water*, *41*(6), 746.

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 Transactions on geoscience and remote sensing*, *48*(9), 3391-3407.

Mozaffari, A., Klotzsche, A., Warren, C., He, G., Giannopoulos, A., Vereecken, H., & van der Kruk, J. (2020). 2.5 D crosshole GPR full-waveform inversion with synthetic and measured data. *Geophysics*, *85*(4), 1-51.

Müller, K., Vanderborght, J., Englert, A., Kemna, A., Huisman, J. A., Rings, J., & 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).

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.

Personna, Y. R., Slater, L., Ntarlagiannis, D., Werkema, D., & Szabo, Z. (2013). Electrical signatures of ethanol–liquid mixtures: Implications for monitoring biofuels migration in the subsurface. *Journal of contaminant hydrology*, *144*(1), 99-107.

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

Petong, P., Pottel, R., & Kaatze, U. (2000). Water− ethanol mixtures at different compositions and temperatures. A dieletric relaxation study. The Journal of Physical Chemistry A, 104(32), 7420-7428.

Remy, N., Boucher, A., & Wu, J. (2009). *Applied geostatistics with SGeMS: a user's guide*. Cambridge University Press.

Rhoades, J. D. (1981). Predicting bulk soil electrical conductivity versus saturation paste extract electrical conductivity calibrations from soil properties. *Soil Science Society of America Journal*, *45*(1), 42-44. (2009). Applied geostatistics with SGeMS: a user's guide. Cambridge University Press.

Seyfried, M. S., & Grant, L. E. (2007). Temperature Effects on Soil Dielectric Properties Measured at 50 MHz. *Vadose Zone Journal*, *6*(4), 759-765.

Singha, K., & Gorelick, S. M. (2005). Saline tracer visualized with three−dimensional electrical resistivity tomography: Field−scale spatial moment analysis. Water Resources Research, 41(5).

Spalding, R. F., Toso, M. A., Exner, M. E., Hattan, G., Higgins, T. M., Sekely, A. C., & Jensen, S. D. (2011). Long−term groundwater monitoring results at large, sudden denatured ethanol releases. *Groundwater Monitoring & Remediation*, *31*(3), 69-81.

Sreenivas, K., Venkatarantan, L., & Rao, P. N. (1995). Dielectric properties of salt-affected soils. *International Journal of Remote Sensing*, *16*(4), 641-649.

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

Vanderborght, J., Kemna, A., Hardelauf, H., & Vereecken, H. (2005). Potential of electrical resistivity tomography to infer aquifer transport characteristics from tracer studies: A synthetic case study. *Water Resources Research*, *41*(6).

van der Kruk, J., Gueting, N., Klotzsche, A., He, G., Rudolph, S., von Hebel, C., Yang, X., WeihermŘller, L., Mester, A. & 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), 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. *Journal of Contaminant Hydrology*, *45*(3-4), 329-358.

Vereecken, H., G. Lindenmayr, O. Neuenedorf, U. Döring, and R. Seidemann, 1994, TRACE: A mathematical model for reactive transport in 3D variable saturated porous media: ICG-4 Internal Report 501494, Forschungszentrum Jülich GmbH.

Willmott, C. J. (1982). Some comments on the evaluation of model performance. Bulletin of the American Meteorological Society, 63(11), 1309-1313.

Witt, A., & Malamud, B. D. (2013). Quantification of long-range persistence in geophysical time series: conventional and benchmark-based improvement techniques. *Surveys in Geophysics*, *34*(5), 541-651.

Wyman, J. (1931). The dielectric constant of mixtures of ethyl alcohol and water from-5 to 40. *Journal of the American Chemical Society*, *53*(9), 3292-3301.

Zhou, Z., Klotzsche, A., Hermans, T., Nguyen, F., Schmäck, J., Haruzi, P., Harry Vereecken, and van der Kruk, J. (2020). 3D aquifer characterization of the Hermalle-sous-Argenteau test site using crosshole GPR amplitude analysis and full-waveform inversion. *Geophysics*, *85*(6), H133-H148.

Zhou, Z., Klotzsche, A., & Vereecken, H. Improving crosshole ground penetrating radar full−waveform inversion results by using progressively expanded bandwidths of the data. *Near Surface Geophysics*. doi: 10.1002/nsg.12154.