AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019498

Key Points:
- Highly resolved tomographic images of the subsurface are obtained from GPR full-waveform inversion
- Logistic regression is used to translate GPR electrical properties into hydrogeological facies
- The distribution of GPR facies explains the plume splitting observed in a tracer experiment

Correspondence to:
N. Gueting,
n.gueting@fz-juelich.de

Citation:
Gueting, N., T. Vienken, A. Klotzsche, J. van der Kruk, J. Vanderborght, J. Caers, H. Vereecken, and A. Englert (2017), High resolution aquifer characterization using crosshole GPR full-waveform tomography: Comparison with direct-push and tracer test data, *Water Resour. Res.*, *53*, 49–72, doi:10.1002/2016WR019498.

Received 13 JUL 2016
Accepted 20 NOV 2016
Accepted article online 23 NOV 2016
Published online 10 JAN 2017

High resolution aquifer characterization using crosshole GPR full-waveform tomography: Comparison with direct-push and tracer test data

Nils Gueting [1], Thomas Vienken [2], Anja Klotzsche[1,3], Jan van der Kruk [1,3], Jan Vanderborght [1,3], Jef Caers [4], Harry Vereecken [1,3], and Andreas Englert [5]

[1]Agrosphere, IBG-3, Forschungszentrum Jülich GmbH, Jülich, Germany, [2]Department Monitoring and Exploration Technologies, UFZ - Helmholtz Centre for Environmental Research, Leipzig, Germany, [3]Centre for High-Performance Scientific Computing in Terrestrial Systems (HPSC-TerrSys), Geoverbund ABC/J, Jülich, Germany, [4]Geological Sciences, Stanford University, Stanford, California, USA, [5]Earth Sciences Department, Ruhr University Bochum, Bochum, Germany

Abstract Limited knowledge about the spatial distribution of aquifer properties typically constrains our ability to predict subsurface flow and transport. Here we investigate the value of using high resolution full-waveform inversion of cross-borehole ground penetrating radar (GPR) data for aquifer characterization. By stitching together GPR tomograms from multiple adjacent crosshole planes, we are able to image, with a decimeter scale resolution, the dielectric permittivity and electrical conductivity of an alluvial aquifer along cross sections of 50 m length and 10 m depth. A logistic regression model is employed to predict the spatial distribution of lithological facies on the basis of the GPR results. Vertical profiles of porosity and hydraulic conductivity from direct-push, flowmeter and grain size data suggest that the GPR predicted facies classification is meaningful with regard to porosity and hydraulic conductivity, even though the distributions of individual facies show some overlap and the absolute hydraulic conductivities from the different methods (direct-push, flowmeter, grain size) differ up to approximately one order of magnitude. Comparison of the GPR predicted facies architecture with tracer test data suggests that the plume splitting observed in a tracer experiment was caused by a hydraulically low-conductive sand layer with a thickness of only a few decimeters. Because this sand layer is identified by GPR full-waveform inversion but not by conventional GPR ray-based inversion we conclude that the improvement in spatial resolution due to full-waveform inversion is crucial to detect small-scale aquifer structures that are highly relevant for solute transport.

1. Introduction

Predicting subsurface flow and transport is challenging due to the complex heterogeneity found in most geologic media. While it has been shown that flow processes can be reasonably well described using spatially averaged (or upscaled) properties, which are easily available from classical hydrogeological methods (e.g., pumping test), this approach fails for the description of transport processes, which inherently depend on the distribution of local properties [*Dagan*, 1989]. Because the detailed spatial characterization of aquifers is unfeasible using classical hydrogeological methods, novel techniques are needed that are able to characterize aquifers with higher spatial resolution.

Recent advances in geophysical measurement and inversion techniques allow imaging the subsurface with unprecedented resolution and coverage [*Binley et al*., 2015]. A particularly promising method with regard to the detailed characterization of porous aquifers is full-waveform inversion of cross-borehole GPR data [*Ernst et al*., 2007; *Klotzsche et al*., 2014; *Gueting et al*., 2015]. This method uses a transmitting and a receiving antenna which are placed in two adjacent boreholes, several meters apart from each other. The transmitting antenna emits an electromagnetic wave which propagates through the subsurface and is detected by the receiving antenna. If the wave encounters contrasting electrical properties in the subsurface, it will be reflected, refracted or scattered [*Davis and Annan*, 1989]. The waveform that arrives at the receiving antenna thus contains information about the medium the wave has passed. By repeating the measurement for multiple transmitter-receiver configurations, and inverting the measured data, two-dimensional tomograms of the subsurface electrical properties (dielectric permittivity and electrical conductivity) can be obtained.

Because the attenuation of GPR waves depends on electrical conductivity [*Davis and Annan*, 1989], GPR performance is optimal for settings where the subsurface electrical conductivity is relatively low, such as coarse grained porous aquifers with low groundwater salinity. If significantly higher electrical conductivities are encountered (due to clay-rich sediments or high groundwater salinity), the use of GPR for subsurface characterization may be limited because of strong signal attenuation.

Traditional inversion of crosshole GPR data focuses on the reconstruction of first arrival times, and ray-based forward simulators are used to approximate the propagation of radar waves in the ground [e.g., *Holliger et al*., 2001]. Ray-based models, however, strongly simplify the propagation of electromagnetic waves, which can lead to biased inversion results [*Linde and Vrugt*, 2013]. In addition, first arrival times constitute only a small portion of the information contained in the measured waveforms. Recently, full-waveform inversion approaches, which are commonly used in seismic inversion, have been introduced to crosshole GPR inversion [*Ernst et al*., 2007; *Meles et al*., 2010]. Full-waveform inversion aims at reconstructing not only first arrival times but entire waveforms, and employs a numerical solution of Maxwell's equations to more accurately simulate the electromagnetic wave propagation. As a consequence, this approach exploits more accurately the information contained in the GPR data and allows characterizing the subsurface with a considerably improved spatial resolution [*Belina et al*., 2009; *Yang et al*., 2013; *Klotzsche et al*., 2013, 2014; *Gueting et al*., 2015; *van der Kruk et al*., 2015].

Although the potential benefit of utilizing high resolution geophysical data for hydrogeological site characterization has been demonstrated [e.g., *Scheibe and Chien*, 2003; *Linde et al*., 2006; *Dafflon et al*., 2011], its routine application is hindered by the lack of direct and universal relationships between geophysical and hydrogeological properties. Even though strong and linear relationships seem to exist at some sites, which has motivated direct estimation of hydraulic conductivity from e.g., seismic slowness [*Hyndman et al*., 2000] or from radar velocity [*Hubbard et al*., 2001; *Linde et al*., 2008], relationships at other sites are essentially unknown and may or may not exist [*Binley et al*., 2015].

Rather than relying on explicit petrophysical relationships, other studies have instead exploited the geometrical information obtained from geophysical data [*Hyndman and Gorelick*, 1996; *Tronicke et al*., 2004; *Schmelzbach et al*., 2011]. The underlying assumption is that geophysical variations delineate lithological zonation, which is in turn related to hydrogeological properties. As no explicit petrophysical relationship is needed, zonation approaches alleviate the difficulty of defining one specific quantitative relationship required for the direct estimation of hydraulic parameters from geophysical data. *Hyndman and Gorelick* [1996] jointly inverted seismic and tracer data to estimate the geometry and hydraulic conductivity of lithological zones. With the approximation that each zone can be represented by a constant hydraulic conductivity, they obtained a reasonably good match between observed and numerically modeled tracer breakthrough curves. *Wainwright et al*. [2014] developed a Bayesian hierarchical approach to jointly invert multiscale geophysical data sets and used this approach to estimate reactive facies at a contaminated site by integrating high resolution crosshole and low resolution surface seismic data. In a synthetic study, *Linde et al*. [2006] used crosshole GPR data to delineate GPR facies and inverted tracer data to infer pixel-wise hydraulic conductivities from radar velocity, allowing a different petrophysical relationship between radar velocity and hydraulic conductivity for each GPR facies. *Dogan et al*. [2011] and *Schmelzbach et al*. [2011] employed surface GPR measurements to define subsurface facies and derived individual hydraulic conductivity distributions for each facies from colocated direct-push hydraulic conductivity profiles.

In the present study, we follow a conceptually similar approach and employ crosshole GPR to characterize the heterogeneous architecture of an alluvial aquifer. But in contrast to the aforementioned studies, we analyze the GPR data by full-waveform inversion, which enables us to image the subsurface with increased spatial resolution. Full-waveform inversion of crosshole GPR data has been applied previously to characterize alluvial aquifers [*Oberröhrmann et al*., 2013; *Klotzsche et al*., 2013, 2014; *Gueting et al*., 2015], but these studies considered only relatively small spatial scales (i.e., one or a few crosshole planes). Here we investigate a considerably larger domain by analyzing a total number of 15 crosshole planes that are connected across an aquifer volume of 50 m $\times$ 25 m $\times$ 10 m.

The main contribution of this study, however, lies in the comparison of the GPR results with hydrogeological data, which provides an excellent basis to assess the actual benefit of GPR full-waveform inversion for hydrogeological site characterization. Thereto, the GPR results are compared with the tracer breakthrough

observed in a field experiment and with vertical profiles of porosity and hydraulic conductivity from direct-push, flowmeter and grain size data.

The remaining paper is arranged as follows: In section 2, we describe measurement and inversion of the GPR data, and we describe direct-push injection logs and direct-push slug tests which were conducted to characterize the vertical variations in hydraulic conductivity at selected profiles along the GPR transects. Because the newly conducted GPR and direct-push measurements are compared with data from previous studies at this site, we briefly revisit some of the data retrieved in previous studies to facilitate the later comparisons. In section 3, we first explore the direct correlation between GPR results, porosity and hydraulic conductivity using crossplots and linear regression; then we derive a lithological zonation from the GPR tomograms using logistic regression. The resulting GPR facies are characterized in terms of porosity and hydraulic conductivity using colocated direct-push, flowmeter and grain size data, and the facies distribution along the GPR transects is compared with the breakthrough of a tracer observed in a field experiment. Section 4 summarizes the main findings.

2. Material and Methods

2.1. Study Site

The Krauthausen test site is situated in the Lower Rhine Embayment, approximately halfway between the cities of Cologne and Aachen. It was set up by the research center Jülich in 1993 with the goal to study the in-situ distribution of subsurface properties and their link to groundwater flow and solute transport. A detailed description of the test site and its instrumentation is given by *Döring* [1997] and *Vereecken et al.* [2000]. The test site has an extent of 200 × 50 m with its longitudinal axis approximately aligned to the regional groundwater flow direction (Figure 1). The site is equipped with 76 observation wells reaching to a depth of 10–13 m which corresponds to the base of the uppermost aquifer. The aquifer material is

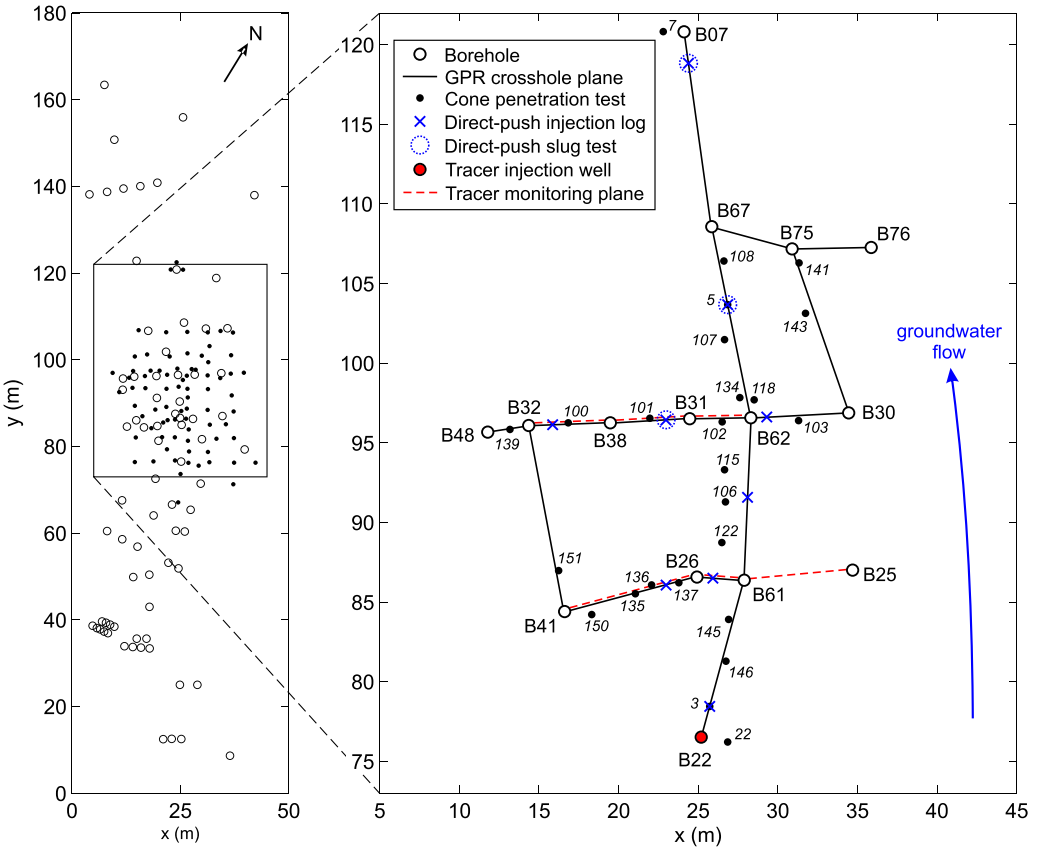

Figure 1. Map of the Krauthausen test site. For clarity, the close-up (right) shows only those boreholes and cone penetration tests that are located along the GPR transects.

composed of alluvial terrace sediments, deposited by the Rur river, a local braided river system, on top of older Rhine and Maas sediments [*Englert*, 2003]. The aquifer is characterized by sandy to gravely grain size with only small amounts of clay. The mean aquifer porosity, based on laboratory analysis of core samples is 26% [*Vereecken et al*., 2000]. The averaged hydraulic conductivity derived from a large scale pumping test is 3×10^{-3} m/s [*Englert*, 2003].

Over the last decades, the Krauthausen site has been intensively investigated using a broad spectrum of methods, including tracer experiments [*Vereecken et al*., 2000; *Vanderborght and Vereecken*, 2001; *Kemna et al*., 2002; *Müller et al*., 2010], flowmeter and pumping tests [*Li et al*., 2007, 2008], geophysical imaging methods [*Kemna et al*., 2002; *Müller et al*., 2010; *Gueting et al*., 2015], cone penetration tests [*Tillmann et al*., 2008] and laboratory characterization of core samples [*Döring*, 1997]. These studies have brought together a rich set of information, which provides excellent conditions to test and validate novel methods.

2.2. Crosshole Ground Penetrating Radar (GPR) Tomography

2.2.1. Data Acquisition

GPR data were acquired along 15 crosshole planes with individual borehole distances from 3 to 12 m (Figure 1). Crosshole planes with borehole distances smaller than 7 m were measured with 200 MHz antennas, crosshole planes with borehole distances larger than 7 m were measured with 100 MHz antennas because the signal detected with 200 MHz antennas was too weak. To obtain sufficient ray coverage at affordable acquisition times, measurements were carried out in a semi-reciprocal acquisition setup with transmitter and receiver spacing of 0.5 m and 0.1 m, respectively [*Oberröhrmann et al*., 2013]. Measurements were conducted from approximately 3 m depth (about 1 m below the groundwater level during the measurements) down to approximately 10 m depth, depending on the depth of the boreholes. To avoid inversion artifacts due to imprecisely known transmitter and receiver positions [*Linde et al*., 2006], borehole deviations from the vertical were accurately determined by deviation log, prior to the GPR measurements.

2.2.2. Full-Waveform Inversion (FWI)

A full-waveform inversion approach was applied to infer the relative dielectric permittivity, ϵ, and the electrical conductivity, σ, from the measured GPR data. A brief description of the inversion scheme is given below, for a more detailed description of the inversion procedure and its implementation the reader is referred to *Meles et al*. [2010] and *Klotzsche et al*. [2010].

We used the inversion scheme introduced by *Ernst et al*. [2007] and modified by *Meles et al*. [2010]. This approach employs a conjugate gradient method to iteratively minimize the misfit between observed and model-predicted radar traces, the latter being generated by a forward model based on a 2-D finite difference time domain (FDTD) solution of Maxwell's equations. After each iteration, the parametrization of the forward model (i.e., the spatial distribution of ϵ and σ) is updated using local gradients derived from a cross-correlation of the back-propagated residual wavefield and the model-predicted wavefield [*Meles et al*., 2010]. To avoid overfitting, the inversion is stopped when the root-mean squared (RMS) error between observed and model-predicted radar traces changes less than 1% from one iteration to the next.

An important requirement for the full-waveform inversion are adequate starting models for ϵ and σ. In particular, the ϵ starting model is of critical importance because proper convergence of the full-waveform inversion requires that for each transmitter-receiver combination the initial offset between observed and modeled wavefield is not greater than half the dominant wavelength, otherwise the inversion will converge to a local minimum [*Klotzsche et al*., 2014]. To derive adequate ϵ starting models, we conducted a ray-based traveltime inversion [*Holliger et al*., 2001] for each crosshole plane, and used the result of the ray-based inversion as starting model for the full-waveform inversion. For σ, we used one homogenous starting model ($\sigma = 13$ mS/m) for all crosshole planes. In principle, individual σ starting models for each crosshole plane could be obtained from ray-based amplitude inversion [*Maurer and Musil*, 2004]. However, we found the ray-based σ models obtained for a series of adjacent crosshole planes to show large discrepancies in absolute values, which is unlikely for adjacent planes measured on consecutive days on the same site. In fact, it has been shown that the estimation of σ using ray-based methods is not always straight-forward and, in particular, absolute σ-values can be biased [*Holliger et al*., 2001; *Maurer and Musil*, 2004]. We therefore took a different approach and tested, for a few crosshole planes, the convergence of the full-waveform inversion for a broad range of homogeneous σ starting models. The model which showed the best convergence was selected as starting model and used for all crosshole planes.

To minimize inversion artifacts in the vicinity of the boreholes, we adopted the approach by *Kurzmann et al.* [2013] and implemented a gradient preconditioner that reduces the high sensitivity of the inversion close to transmitter and receiver positions. Details on the implementation are given in *van der Kruk et al.* [2015].

2.3. Complex Refractive Index Model (CRIM)

Petrophysical models can be applied to translate the electrical properties obtained from GPR into hydrologically more relevant properties such as soil moisture or porosity. Among the various empirical, semi-empirical and theoretical models that have been proposed in the literature [e.g., *Topp et al.*, 1980; *Dobson et al.*, 1985; *Steelman and Endres*, 2011], we opted for the CRIM relationship [*Birchak et al.*, 1974], which is based on the intuitive idea that the bulk dielectric permittivity, ϵ_b, of an *n*-phase medium can be expressed as the sum of the dielectric permittivities of the individual phases, ϵ_i, weighted by their volume fraction, V_i,

$$\epsilon_b^\alpha = \sum_{i=1}^{n} V_i \epsilon_i^\alpha. \tag{1}$$

The exponent α describes the spatial arrangement of the individual phases and takes a value of 0.5 for an isotropic phase arrangement [*Birchak et al.*, 1974]. Although the CRIM does not account for the full complexity of natural porous media [*Dobson et al.*, 1985; *Brovelli and Cassiani*, 2008], the CRIM has been found to yield robust porosity estimates for various types of sediments [*Roth et al.*, 1990]. For a two-phase fully saturated porous medium with isotropic phase arrangement, equation (1) can be written as

$$\phi = \frac{\sqrt{\epsilon_b} - \sqrt{\epsilon_s}}{\sqrt{\epsilon_f} - \sqrt{\epsilon_s}}, \tag{2}$$

where ϕ is the porosity, ϵ_b is the bulk dielectric permittivity (obtained e.g., from GPR), and ϵ_s and ϵ_f are the dielectric permittivities of the solid grain matrix and of the pore fluid, respectively.

2.4. Logistic Regression

Just like linear regression, logistic regression investigates the dependency between a response variable and one or more explanatory variables. But while linear regression applies to continuous response variables, logistic regression applies to binary or multinomial response variables. Given this difference, logistic regression may be able to identify relations between variables when linear regression does not. For example, *Tesoriero et al.* [1998] showed that, while it was not feasible to use linear regression to predict quantitative values for nitrate concentrations in groundwater from well depth data, it was still possible to use logistic regression to predict the probability of the concentration exceeding a critical threshold.

A brief description of the principle of logistic regression is given below, detailed discussions can be found in *Hosmer and Lemeshow* [2000] or in *Helsel and Hirsch* [2002]. In the field of geosciences, logistic regression has been used to predict landslide susceptibility on the basis of geomorphological parameters [*Ayalew and Yamagishi*, 2005; *van den Eeckhaut et al.*, 2006], and groundwater vulnerability on the basis of land use data [*Eckhardt and Stackelberg*, 1995; *Nolan*, 2001]. The basic assumption of logistic regression is that the logit of the probability of an observation being in a response category can be modeled using a linear combination of explanatory variables

$$logit(\tau) = \ln\left(\frac{\tau}{1-\tau}\right) = \beta_0 + \beta X, \tag{3}$$

where τ is the probability, X is the vector of explanatory variables, and β_0 and β are the intercept and the vector of regression coefficients, respectively. For a multinomial response function with k categories and n explanatory variables, equation (3) becomes a set of $k-1$ equations

$$\ln\left(\frac{\tau_j}{\tau_k}\right) = \beta_{j0} + \beta_{j1}X_1 + \beta_{j2}X_2 + \ldots + \beta_{jn}X_n, \qquad j = 1, \ldots, k-1, \tag{4}$$

where τ_j/τ_k represents the relative probability of belonging to category j versus belonging to the reference category k (here by definition the last category). The regression coefficients β can be determined using maximum likelihood estimation [*Helsel and Hirsch*, 2002]. Once β is determined, the probabilities of an observation belonging to each category can be calculated

$$\tau_j = \frac{e^{\beta_{j0}+\beta_{j1}X_1+\beta_{j2}X_2+\ldots+\beta_{jn}X_n}}{1+\sum_{j=1}^{k-1} e^{\beta_{j0}+\beta_{j1}X_1+\beta_{j2}X_2+\ldots+\beta_{jn}X_n}}, \qquad j=1,\ldots,k-1, \tag{5}$$

$$\tau_k = 1-\sum_{j=1}^{k-1}\tau_j. \tag{6}$$

The individual model components (explanatory variables) can be tested for statistical significance using *t*-statistics, which are obtained by dividing each component of β by its standard error. To identify the 'best' model among a set of models with different (combinations of) explanatory variables, Akaike's Information Criterion (AIC) can be applied [*Helsel and Hirsch*, 2002]. The AIC calculates as AIC = 2*p*– 2*L*, with *p* the number of estimated model parameters and *L* the log-likelihood of the fitted model [*Akaike*, 1974]. Among a set of different models, the model that minimizes the AIC is the preferred model. The AIC, thus, rewards a high likelihood and penalizes a high number of model parameters.

In the present study, we applied a logistic regression of GPR results (explanatory variables) and lithological facies derived from grain size data (response categories). GPR and grain size data were available at colocated locations along 3 boreholes. The basic idea is to use the colocated data to derive an empirical relationship (i.e., the fitted logistic regression model), and utilize this relationship to predict the distribution of grain size facies at locations where no grain size data exist, solely on the basis of GPR. The following workflow was applied: prior to the logistic regression, the grain size data were classified into lithological facies using simple k-means clustering [*MacQueen*, 1967]. The idea behind this is to transform the multivariate grain size data into a categorical variable, which can be used in logistic regression. The logistic regression was conducted using the clustered grain size data as categorical response variable and the bivariate GPR results (dielectric permittivity and electrical conductivity) as explanatory variables. Maximum likelihood estimates for β were obtained using an iteratively weighted least squares algorithm. With β determined, equations (5) and (6) were applied to calculate, for each GPR pixel, the probability to belong to each grain size facies.

To evaluate the benefit of GPR full-waveform inversion instead of classical GPR ray-based inversion, we compared the outcome of the logistic regression conducted with the full-waveform inversion results, and with the ray-based inversion results as explanatory variables.

2.5. Direct-Push Hydraulic Measurements

2.5.1. Direct-Push Injection Log (DPIL)

The DPIL method [*Dietrich et al*., 2008] allows to assess local variations in hydraulic conductivity. The method consists of advancing a small diameter direct-push rod, equipped with a filter screen, into the subsurface. At selected depths, the pushing is stopped, water is injected through the screen, and the injection flow rate and the injection pressure are recorded. The ratio of injection rate and pressure, corrected for head loss in the tube, gives a proxy, K_{DPIL}, for the aquifer's local hydraulic conductivity [*Dietrich et al*., 2008]. Through calibration with absolute hydraulic conductivity data from direct-push slug tests (or any other method), K_{DPIL} can subsequently be converted into absolute hydraulic conductivities [*Dietrich et al*., 2008; *Lessoff et al*., 2010; *Dietze and Dietrich*, 2012]. At Krauthausen, DPIL profiles were measured at 9 locations (Figure 1). Since the focus is on a comparison with the GPR results, the DPIL measurements were conducted at locations where crosshole GPR results were available. DPIL measurements were conducted from 3 to 9 m depth with a vertical sampling interval of 20 cm. At each depth, 3 repeat measurements with different flow rates were conducted for quality control (e.g., to identify potential effects caused by opening of preferential flowpaths due to the water injection or water flow along the probing rods). During the advancement of the DPIL probe from one sampling depth to the next, water was continuously injected in order to prevent clogging of the screen.

2.5.2. Direct-Push Slug Test (DPST)

In order to obtain absolute values of hydraulic conductivity, multilevel DPST [*Butler et al*., 2002] were performed at 3 profile locations (Figure 1). Because of the multitude of repeat measurements at each measurement interval that were performed for quality control [see *Butler*, 1997], the DPST measurements were relatively time consuming. Therefore, only 2 or 3 depth intervals were tested in each profile, resulting in a total number of 7 individual DPST. To enable calibration of the DPIL data, the DPST were carried out in close

proximity (30 cm lateral distance) to DPIL locations; and depth intervals with fairly constant DPIL ratios were selected. Slug tests were performed in temporarily installed direct-push monitoring wells with a screen length of 1 m. During well installation, the screen was protected inside the rods to prevent clogging of the screen. At the selected depths, the screen was exposed and water was pumped until the pumped water was free of fines. The slug tests were initiated pneumatically [*Butler et al*., 2002], by pressurizing the air column above the water to create the initial head displacement and rapidly depressurizing the air column to initiate the tests. The recorded water table responses were analyzed using the approach of *Butler et al*. [2003], which is based on matching aquifer response type-curves to the measured head recovery. The obtained absolute hydraulic conductivities, $K_{DPST,}$ were subsequently used as calibration data to translate the relative DPIL ratios, $K_{DPIL,}$ into absolute hydraulic conductivities.

2.6. Previous Studies at the Krauthausen Site

To facilitate the later comparisons, we briefly revisit some of the data from previous studies that will be compared with the newly conducted crosshole GPR and direct-push hydraulic measurements.

2.6.1. Cone Penetration Tests (CPT)

More than 70 cone penetration tests have been conducted at Krauthausen and are described in detail by *Tillmann et al*. [2008]. At each CPT location, measurements of cone resistance and electrical resistivity were recorded while the CPT probe was continuously pushed into the ground. Subsequently, additional measurements of natural gamma, gamma-gamma and neutron activity were conducted. By calibration, gamma-gamma and neutron log were related to bulk density and water content, respectively [*Tillmann et al*., 2008]. Recently, *Gueting et al*. [2015] reviewed this calibration by comparing the water contents derived from CPT neutron log with the saturated porosity of core samples determined in the laboratory. They found that the CPT water contents overestimated the laboratory measurements by a constant value and applied a simple constant-shift correction to compensate for the overestimation. In the present study, we used the corrected CPT water contents, the electrical resistivity and the mechanic cone resistance for comparison with the GPR results. We focus only on those CPT that are located close to the GPR transects; the locations and ID numbers of the CPT considered in the present study are shown in Figure 1.

2.6.2. Hydraulic Conductivity From Flowmeter and Grain Size

Previous studies at Krauthausen determined local variations in hydraulic conductivity from flowmeter measurements and grain size [*Vereecken et al*., 2000; *Li et al*., 2008; *Tillmann et al*., 2008]. The flowmeter method consists of measuring, as function of depth, the vertical flow in a screened well subject to a constant pumping rate. From the increase in cumulative flow rate, ΔQ_i, along some depth interval, z_i, the hydraulic conductivity, K_i, of that interval can be estimated [*Molz et al*., 1989]

$$K_i = \frac{\Delta Q_i}{Q}\frac{b}{z_i}\bar{K}, \tag{7}$$

where Q is the total pumping rate, b is the screened thickness of the aquifer and $\bar{K}$ is the depth-averaged hydraulic conductivity obtained from a pumping test carried out in the same well.

Grain size based estimation of hydraulic conductivity is an indirect method (i.e., no flow process is measured) that relies on empirical or semi-empirical relationships between hydraulic conductivity and grain size distribution. Various relationships have been proposed in the literature [see e.g., *Vienken and Dietrich*, 2011]. In the present study, we follow *Tillmann et al*. [2008], who employed the empirical formula of *Bialas and Kleczkowski* [1970] to derive hydraulic conductivity estimates for the Krauthausen sediments and found the obtained estimates to be of the same order of magnitude compared with the hydraulic conductivity derived from a pumping test. The *Bialas and Kleczkowski* [1970] formula is given by

$$K = 0.0036\ d_{20}^{2.3}, \tag{8}$$

with K the hydraulic conductivity in (m/s) and d_{20} the 20th percentile of the sediment particle size in (mm). Due to the empirical nature of this relationship the units are not balanced in equation (8).

2.6.3. Tracer Experiments

Two consecutive tracer tests were carried out in the years 2002 and 2003 at the Krauthausen site [*Müller et al*., 2010]. In both experiments, tracer was injected into borehole B22 and tracer breakthrough was monitored along two downgradient detection planes using time-lapse electrical resistivity tomography (ERT) and multilevel groundwater samplers. The overall setup is shown in Figure 1. ERT electrodes were installed in

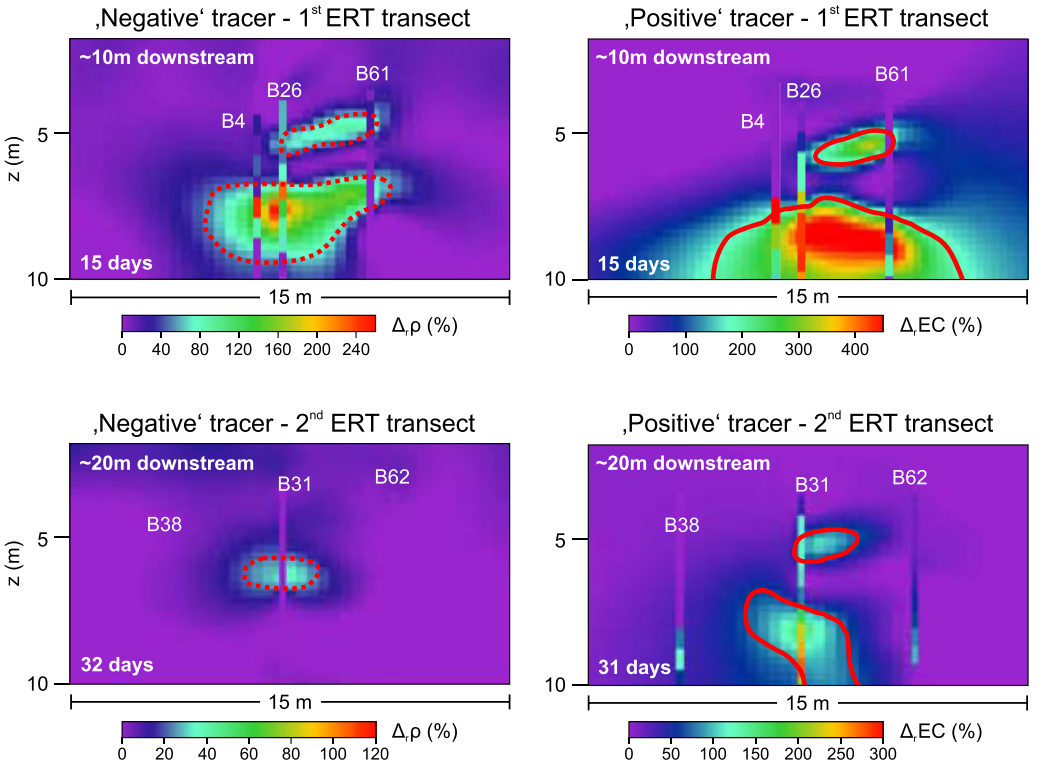

Figure 2. ERT observed tracer plume splitting (modified from *Müller et al.* [2010]). Colorbars refer to changes in bulk resistivity (left) and bulk electrical conductivity (right), related to the breakthrough of a resistive (negative) and a conductive (positive) tracer. Tracer plumes are outlined by the red dashed and solid lines. Vertical bars show results from multilevel groundwater samplers.

boreholes B41, B26, B61 and B25 (first ERT transect), and in boreholes B48, B32, B31 and B62 (second ERT transect). Additional electrodes were placed at the surface [*Müller et al.*, 2010]. In the first experiment, a 'positive' tracer with a higher electrical conductivity than the ambient groundwater was applied. In the second experiment, a 'negative' tracer with a lower electrical conductivity than the ambient groundwater was applied. The idea behind this was that if the two experiments yield consistent results, density effects related to the tracer application can be neglected. A consistent observation in both experiments was a splitting of the tracer plume in the first ERT transect (Figure 2, top). This splitting was also observed in the second ERT transect for the positive tracer experiment, but not for the negative tracer experiment (Figure 2, bottom). The reason why the negative tracer experiment did not show a plume splitting in the second ERT transect could be a smaller signal-to-noise ratio in the negative tracer experiment which possibly inhibited seeing detailed breakthrough patterns in the ERT images. For a detailed discussion of the results, the reader is referred to *Müller et al.* [2010].

In the present study, we compared the ERT observed tracer breakthrough with crosshole GPR data acquired along the same transects, in order to evaluate if the GPR tomograms give evidence about subsurface structures that caused the splitting of the tracer plume. Moreover, we derived estimates of hydraulic conductivity from the tracer arrival times in order to compare them with the hydraulic conductivity derived from direct-push, flowmeter and grain size data. Thereto, hydraulic conductivity estimates were derived from tracer velocities

$$K = \frac{v\,n}{i}, \tag{9}$$

with v the pixel-wise equivalent (streamtube) tracer velocities, that were determined by *Müller et al.* [2010] by fitting the solution of the 1D convection-dispersion equation to locally measured breakthrough curves (i.e., to the tracer breakthrough detected within a pixel of the ERT cross sections). n and i are the aquifer porosity and the hydraulic gradient, respectively.

3. Results and Discussion

3.1. GPR Full-Waveform Inversion Results

A full-waveform inversion was applied to infer the subsurface dielectric permittivity and electrical conductivity for a total number of 15 individual GPR crosshole planes. Depending on the depths of the boreholes, a single crosshole data set comprised approximately 1500–2000 individual radar traces. Each crosshole data set was inverted separately, i.e., independent from adjacent crosshole planes. A grid cell size of 3 cm × 3 cm was used for the forward modeling to ensure numerical accuracy [*Holliger and Bergmann*, 2002]. For the inversion, 9 forward modeling cells (3 × 3) were combined into one inversion cell [*Ernst et al*., 2007]. Convergence of the full-waveform inversion was typically reached after 15–20 iterations. For each crosshole plane the full-waveform inversion lead to a reduction of the initial RMS misfit (obtained with the ray-based permittivity starting model and the uniform conductivity starting model) by about 50–70%.

Figure 3 shows a 3-D plot of the full-waveform inversion results for all crosshole planes. Close-ups of ray-based and full-waveform dielectric permittivity along the main 2-D transects are shown in Figure 4. Compared with the ray-based results the full-waveform tomograms contain spatial structures at greater detail. This is most notable in the longitudinal transect (Figures 4a and 4b), where the full-waveform inversion identifies several thin layers with high permittivity, that are not resolved by ray-based inversion.

An important finding is that the tomograms in Figures 3 and 4 show relatively consistent spatial structures across boreholes, which indicates robust inversion results. For some boreholes, however, the tomograms left and right of the boreholes depart and/or the spatial structures in the tomograms can be observed to "fade out" close to the boreholes. This effect can most clearly be seen in Figure 4 along the boreholes B62, B61 and B38, both for ray-based and full-waveform inversion results. In the following, we will briefly discuss potential causes for these inconsistencies as well as possible strategies against it. In fact, inversion artifacts close to boreholes are an issue commonly encountered in crosshole GPR (and seismic) tomography [e.g., *Maurer and Green*, 1997; *Becht et al*., 2004; *Irving et al*., 2007]. A well-known source of such artifacts are imprecisely known transmitter and receiver positions within the boreholes [*Linde et al*., 2006]. Therefore, it

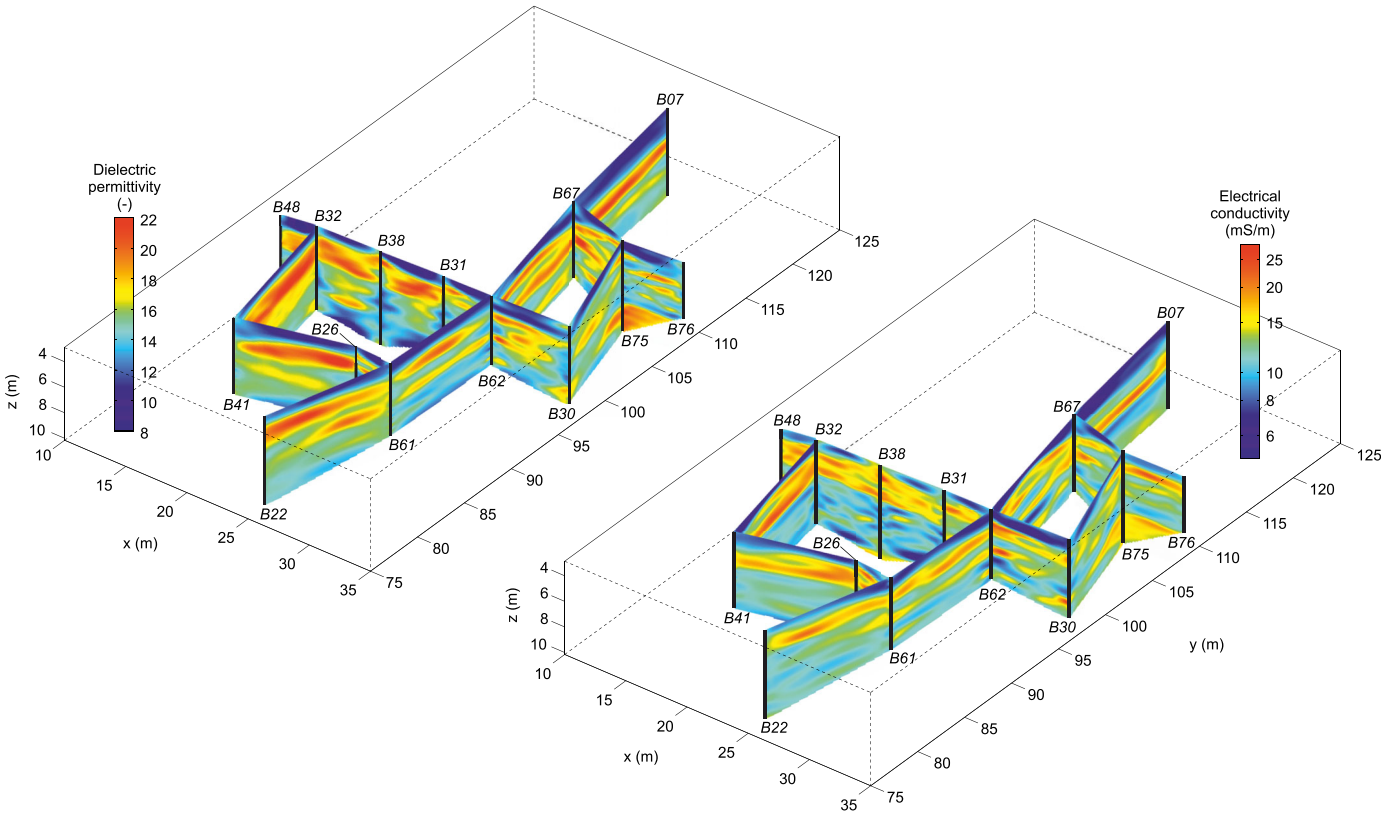

Figure 3. 3-D plot of (left) dielectric permittivity and (right) electrical conductivity derived from full-waveform inversion of crosshole GPR data. Black vertical lines represent boreholes.

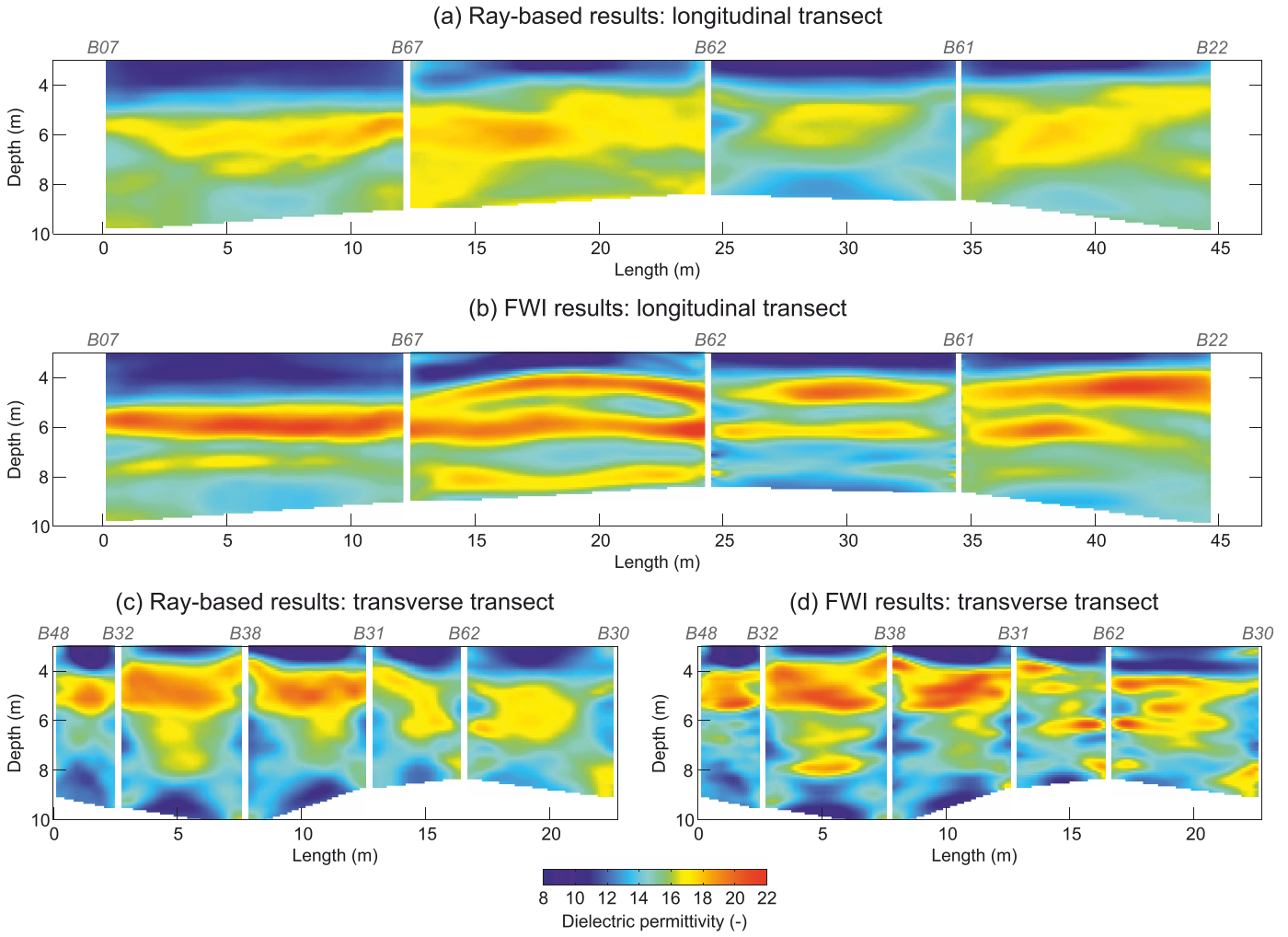

Figure 4. Comparison of (a, c) ray-based and (b, d) FWI-derived dielectric permittivity along the two main transects (from borehole B07 to B22 and from B48 to B30; see Figure 1 for reference).

is common practice to perform borehole deviation log measurements prior to the crosshole measurements [e.g., *Doetsch et al*., 2010; *Dafflon et al*., 2011]. This is also what was done in the present study. However, *Maurer and Green* [1997] noted that measured borehole deviations can be flawed, and demonstrated that relatively minor coordinate errors can produce considerable inversion artifacts. As a strategy to account for potential errors in the measured coordinates they proposed to treat transmitter and receiver coordinates as uncertain and include them as additional inversion parameters. Applied to seismic data from the Grimsel site their approach was able to strongly reduce the inconsistencies between two adjacent tomograms.

Dafflon et al. [2011] analyzed radar traveltimes from a total number of 31 crosshole planes at the Boise Hydrogeophysical Research Site and showed that the consistency of intersecting tomograms could be greatly improved by simultaneously inverting all crosshole data sets together instead of separately. Unfortunately, they also observed that by inverting all data sets jointly, a significantly smoother subsurface model was obtained (i.e., a model containing less details) than when each data set was inverted separately. As our primary goal is to image subsurface structures with the highest possible resolution we decided to invert each crosshole data set separately. An unwanted side effect of that might be that the obtained tomograms are not completely consistent with each other.

It should also be mentioned that the drilling of the boreholes may have locally disturbed the soil structure. In that case, discontinuities in the GPR tomograms would reflect actual zones of disturbance rather than inversion artifacts.

3.2. Comparison of GPR Full-Waveform Inversion Results With CPT Neutron and Electrical Resistivity Logs

As several CPT profiles are located in close proximity to the GPR transects (Figure 1), they provide an opportunity to validate the GPR results. It should be noted, however, that also the CPT profiles are not expected to be flawless, and thus, discrepancies between GPR and CPT could result from errors in either or both techniques. In the following, we compare the porosity from CPT neutron logs with porosity estimates derived from GPR; and we compare the electrical conductivity from CPT resistivity logs with the electrical conductivity derived from GPR. Altogether, 17 CPT porosity profiles and 9 CPT electrical conductivity profiles were available at locations colocated with the GPR transects. To obtain porosity estimates from GPR, the CRIM relation (equation (2)) was applied with $\epsilon_f = 84$, which represents the dielectric permittivity of water at 10 °C [*Eisenberg and Kauzmann*, 1969], and with $\epsilon_s = 4.5$, which represents the dielectric permittivity of quartz [*Carmichael*, 1982], i.e., no fitting was performed. Figure 5a shows a comparison of CPT- and GPR-derived porosity along CPT profile 141. Overall, the porosity from CPT and GPR shows a strong correlation. Compared with the ray-based inversion results the full-waveform inversion improves the reconstruction of small scale structures, such as the high porosity layer at about 4 m depth. The ray-based results give a good estimate of the basic trends in porosity, though.

Crossplots of CPT porosity versus GPR FWI porosity along all CPT profiles are shown in Figures 6a–6c. Note that the data from crosshole planes measured with different GPR antenna frequencies are plotted both separately (Figures 6a and 6b) and combined (Figure 6c). As expected, the agreement between GPR and CPT is better (higher correlation coefficient, regression line closer to the 1:1 line) for the GPR data acquired with the 200 MHz antennas (Figure 6a), than for the GPR data acquired with the 100 MHz antennas (Figure 6b). This can be attributed to the higher spatial resolution that can be obtained from shorter wavelengths.

Figures 6d–6f show crossplots of electrical conductivity from CPT and GPR FWI. Although the correlation is not as strong as it is for the porosity, GPR and CPT results still correlate fairly well. For the electrical conductivity data, a better correlation of the 200 MHz data compared to the 100 MHz data cannot be recognized, but this could also be due to the limited amount of data (due to technical problems during the CPT survey, the electrical conductivity was measured only at approximately every second CPT profile). Generally, it can be observed that the majority of data points plot above the 1:1 line in Figures 6d–6f, which indicates that the electrical conductivity obtained from GPR is slightly higher than the electrical conductivity obtained from CPT. Figure 5b shows a comparison of CPT and GPR-derived electrical conductivity along CPT profile 135. Along this profile the CPT and GPR results can additionally be compared with the ERT results from *Müller et al.* [2010]. Overall, the three different methods identify fairly similar trends. However, some notable differences exist, particularly in the lower part of the profile. It is very interesting to see how the small scale variability is most pronounced in the CPT profile, less pronounced in the GPR profile, and virtually absent in the ERT profile, which highlights the differences in spatial resolution of the individual methods.

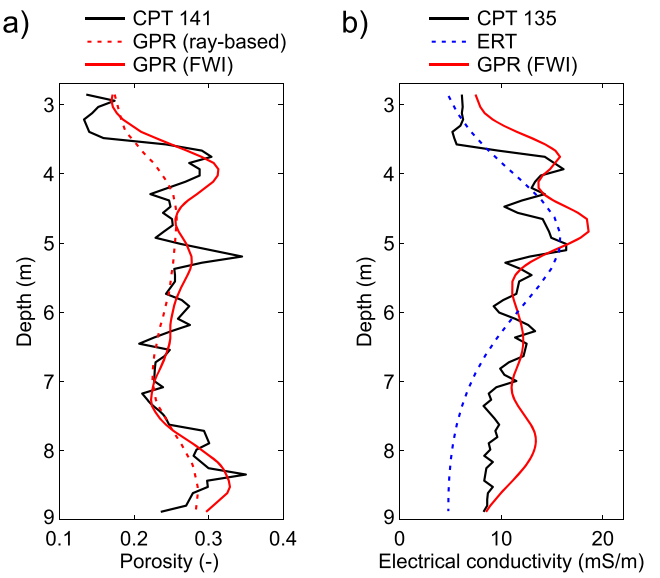

Figure 5. (a) CPT porosity profile and comparison with colocated porosity estimated from GPR ray-based inversion and from GPR full-waveform inversion. (b) CPT electrical conductivity profile and comparison with colocated electrical conductivity estimated from ERT and from GPR full-waveform inversion.

3.3. Comparison of GPR Full-Waveform Inversion Results With Direct-Push Hydraulic Conductivity

Prior to comparing the direct-push derived hydraulic conductivity with the GPR full-waveform inversion results, we converted the relative K_{DPIL} values into absolute hydraulic conductivity using the few K_{DPST} values as calibration data. Following *Dietrich et al.*

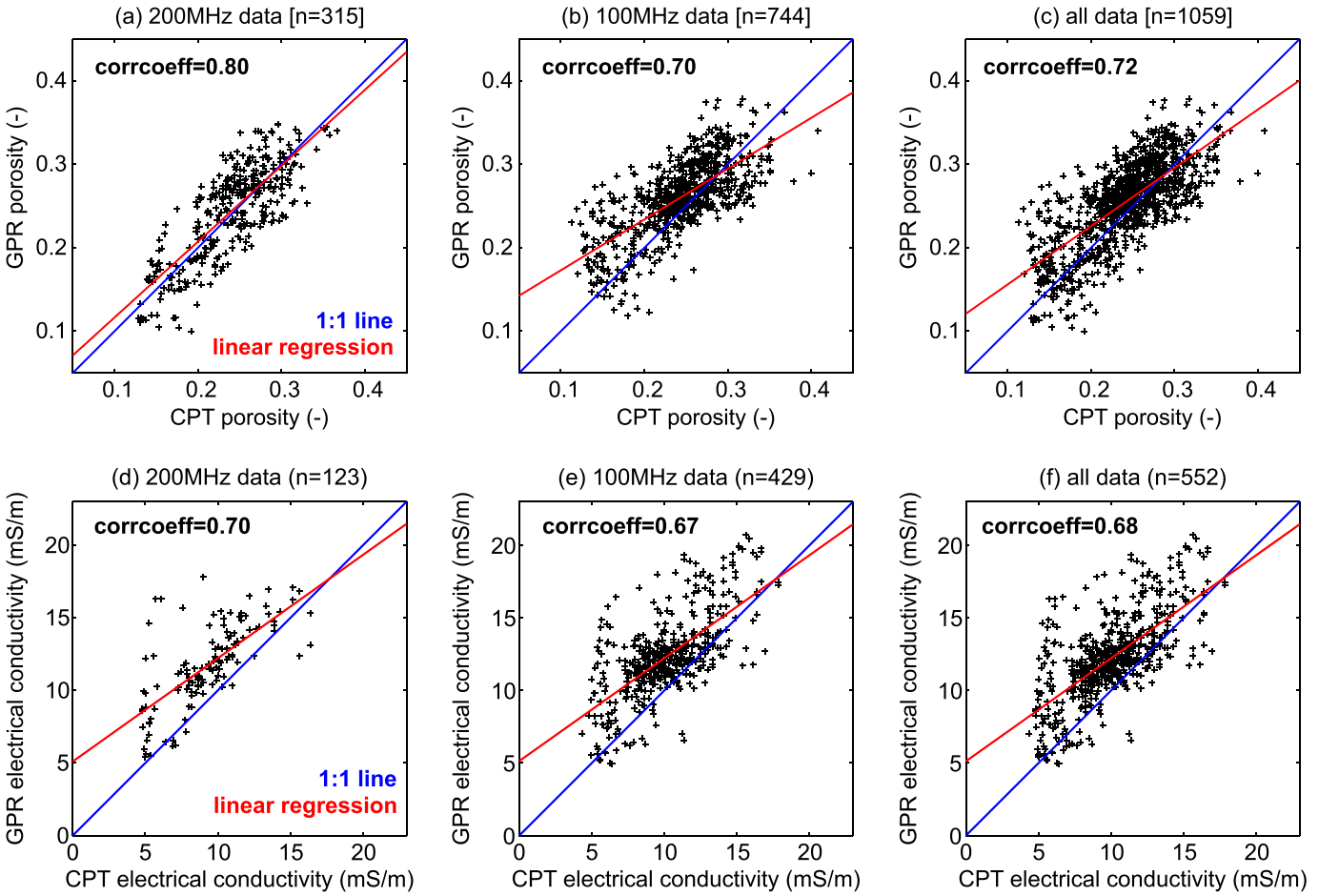

Figure 6. (a–c) Crossplots of porosity from CPT and GPR FWI; (d–f) Crossplots of electrical conductivity from CPT and GPR FWI. Data from crosshole planes measured with different GPR antenna frequencies are plotted separately and combined (a, d) 200 MHz; (b, e) 100 MHz, (c, f) together). *n* is the number of data points.

[2008], we conducted a linear regression of colocated log K_{DPIL} and log K_{DPST} values and used the obtained regression equation to translate K_{DPIL} values into hydraulic conductivity (Figure 7). Because the DPST were conducted with a filter screen length of 1 m, and the DPIL were measured with a vertical sampling interval of 20 cm, there are five DPIL measurements colocated with each DPST measurement. Assuming horizontal layering, we calculated an averaged (or upscaled) K_{DPIL} value for each DPST interval, by taking the arithmetic mean of the five associated K_{DPIL} values. The regression was conducted with the arithmetic means of the K_{DPIL} values and the corresponding K_{DPST} values. The high correlation coefficient of 0.92 suggests that the regression equation allows robust estimation of absolute hydraulic conductivity from K_{DPIL}.

The absolute hydraulic conductivity derived from DPIL and DPST were used to investigate if a correlation exists between hydraulic conductivity and GPR attributes. Figure 8 shows vertical profiles of hydraulic conductivity and GPR FWI results along DPIL profiles 2 and 5. Scatterplots of hydraulic conductivity (entire DPIL and DPST data set) versus GPR attributes are presented in Figures 9a and 9b. The scatterplots were constructed by averaging the values from GPR cells located along the DPST and the DPIL depth intervals, respectively. To reduce noise, a three-point moving average was applied to the DPIL data. Overall, both GPR attributes show a negative correlation with hydraulic conductivity (note that in order to make it easier to see the negative correlation in Figure 8 the log K axis was reversed). The negative correlation is stronger for the (few) DPST hydraulic conductivity estimates than for the DPIL hydraulic conductivity estimates (Figures 9a and 9b).

Interestingly, the scatterplots for the two GPR attributes versus hydraulic conductivity look quite similar. The reason for this similarity is probably that the clay content at the Krauthausen site is so small that both the dielectric permittivity and the electrical conductivity primarily depend on porosity. According to grain size

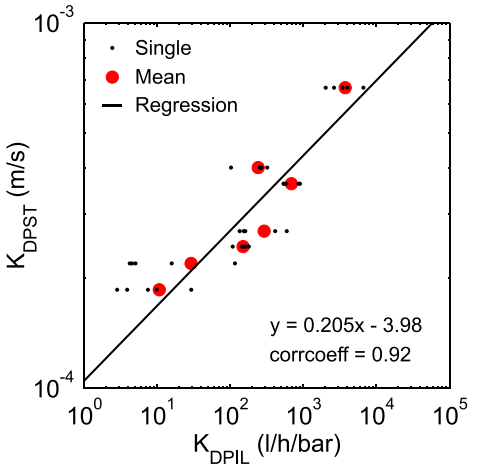

Figure 7. Linear regression of log K_{DPST} and log K_{DPIL}. The equation of the regression line and the correlation coefficient are given at the bottom of the crossplot.

analyses, the average clay content of the part of the aquifer covered by GPR measurements is only 2%. If changes in clay content are minor, the bulk electrical conductivity of saturated media is controlled by the electrical conductivity of the pore fluid. Under such conditions, both the dielectric permittivity and the electrical conductivity obtained from GPR can be expected to depend mainly on saturated porosity, which is also reflected in the strong correlation of the two GPR attributes (Figure 9c).

It is interesting that our data suggest a negative correlation between porosity and hydraulic conductivity, although intuitively one would expect a positive relationship. However, several studies showed that the relation between porosity and hydraulic conductivity is strongly site-dependent and a correlation between the two parameters can be positive, negative or absent [*Nelson*, 1994; *Hubbard et al*., 2001; *Morin*, 2006]. The reason for this is that the hydraulic conductivity is not a function of the total pore volume alone but also of the structure, size, tortuosity and connectivity of the individual pores within the pore network [*Koltermann and Gorelick*, 1996]. At the Krauthausen site, it appears that the structural features of the pore network that lead to a higher hydraulic conductivity are negatively correlated to the total pore volume. According to *Tillmann et al*. [2008] and *Gueting et al*. [2015], the Krauthausen aquifer comprises different lithological facies that are characterized by different average grain sizes (ranging from sand to gravel). In addition, these facies are characterized by different sorting, namely the sand-dominated facies is significantly better sorted than the gravel-dominated facies. This leads to a smaller total porosity of the gravel facies compared to the sand facies. Despite its smaller total porosity, however, we suspect that the gravel is characterized by (fewer but) larger individual pores which compensate for its smaller total porosity and lead to a higher hydraulic conductivity of the gravel facies compared to the sand facies. We believe that this is the reason why we observe a negative correlation between hydraulic conductivity and GPR attributes.

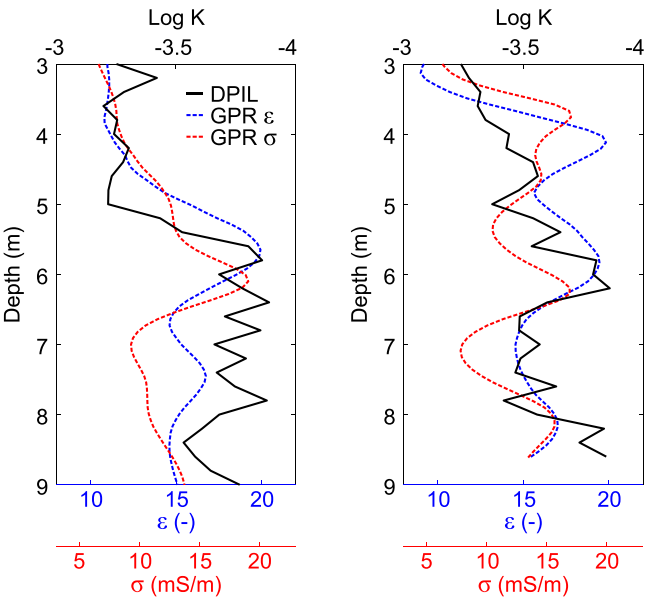

Figure 8. DPIL-derived hydraulic conductivity (note the reversed axis for log K) and colocated dielectric permittivity and electrical conductivity from GPR full-waveform inversion.

Yet despite this overall negative correlation, the relationship between DPIL-derived hydraulic conductivity and GPR results is characterized by considerable scatter (Figures 9a and 9b). There are several possible explanations for that: (1) The small-scale variations in DPIL and/or GPR data are measurement scatter. (2) The GPR and DPIL estimates are not perfectly colocated so that a combination of a small spatial shift with large small-scale variations leads to considerable noise on the correlation between GPR and DPIL estimates. (3) The porosity does not explain the variations in hydraulic conductivity completely, which appears reasonable considering that the total porosity does not seem to exert primary control over hydraulic conductivity at the Krauthausen site, as was discussed above.

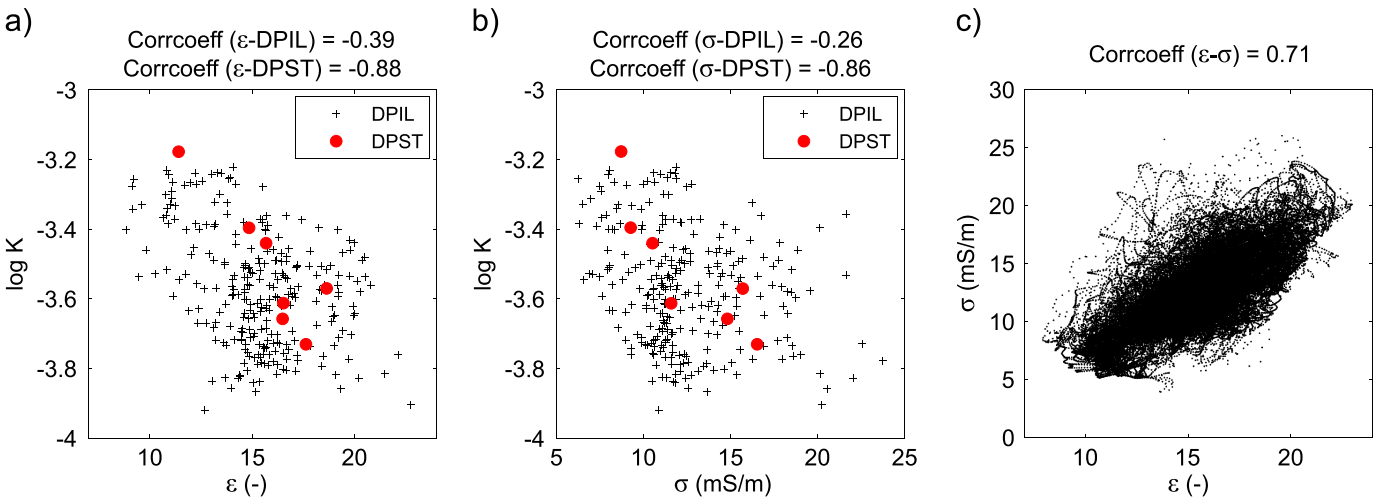

Figure 9. Scatterplots of direct-push hydraulic conductivity and GPR FWI attributes.

In conclusion, the analysis of colocated GPR full-waveform inversion results and direct-push hydraulic conductivity data suggests that an overall negative correlation exists between porosity and hydraulic conductivity at the Krauthausen site. As such, this relationship provides the potential to estimate hydraulic conductivity on the basis of GPR attributes which are sensitive to porosity. Indeed the hydraulic conductivity derived from DPST correlates strongly with the GPR attributes. The weak correlation between the DPIL-derived hydraulic conductivity and the GPR attributes, however, suggests that the use of GPR to estimate fine scale variations in hydraulic conductivity is limited at the Krauthausen site. Note that also a multilinear regression approach, using both GPR attributes (or their principal components) simultaneously, did not lead to a significantly improved correlation between GPR and DPIL data (not shown here). In the following, we present an alternative facies-based approach in which the GPR attributes are not used to directly estimate hydraulic conductivity but instead to distinguish different aquifer units (facies), whose hydraulic conductivity is derived independently, based on direct-push, flowmeter and grain size data.

3.4. Facies Classification

Lithological facies were defined on the basis of grain size data, available along 3 boreholes (B7, B22 and B32). During the drilling of these boreholes, soil samples were collected with a vertical sampling interval of 10 or 20 cm and analyzed for grain size distribution [*Vereecken et al*., 2000]. Colocated with the GPR data, a total number of 140 grain size distribution curves are available (Figure 10a).

To extract lithological facies from the ensemble of grain size distribution curves, we applied a simple k-means cluster analysis [*MacQueen*, 1967]. In this way, grain size distribution curves that are relatively similar to each other (i.e., belong to the same facies) are grouped together in clusters. Figures 10b–10d show the results from clustering the multivariate grain size data (12 sieve fractions for each sediment sample) into three clusters. The division into three clusters was chosen based on the findings of *Döring* [1997] and *Tillmann et al*. [2008], who analyzed borehole logs and CPT profiles and found that the Krauthausen aquifer can be broadly divided into three lithological units. A similar finding was reported by *Gueting et al*. [2015], who analyzed multivariate GPR and CPT data and showed that both data sets were best represented by a division into three facies. It should be noted, however, that these facies are not expected to form clearly separated spatial structures in the Krauthausen aquifer, because sedimentation conditions in alluvial systems often change gradually [*Miall*, 1996], which leads to gradual instead of abrupt changes in subsurface properties (e.g., fining upward sequences). Indeed, it can be observed that the obtained clusters (Figures 10b–10d) represent different subsets of a continuous ensemble rather than clearly separated units. Nevertheless, the facies based approach may be used to represent part of the spatial variability in subsurface properties since the variability of the grain size distribution within a certain cluster is reduced compared to the variability of the ensemble of all grain size distribution curves (Figure 10). Table 1 gives characteristic

properties for each grain size facies. The main differences are reflected in the sand-to-gravel ratio and sorting.

As described in section 2.4, we conducted a logistic regression of the clustered grain size data (response variable) and the colocated GPR data (explanatory variables). The results are given in Table 2. Note that we constructed four different logistic regression models with different sets of explanatory variables. The variables contained in each model are listed in the second column of Table 2. The first model ("Intercept-Only") contains only a constant term, i.e., no GPR data at all. As a consequence, the probabilities derived from this model are constant, simply corresponding to the relative amounts of the individual facies in the grain size data. The reason why we included this most basic model is that any improvement due to incorporating GPR data can be seen by comparison with this basic model. The second model ("Ray-based") contains, in addition to a constant term, the GPR dielectric permittivity obtained from ray-based inversion. In the third model ("FWI (ϵ only)"), the ray-based dielectric permittivity is replaced by the full-waveform dielectric permittivity. The fourth model ("FWI") contains additionally the electrical conductivity obtained from full-waveform inversion.

The maximum likelihood coefficient estimates, β, obtained for each model (Table 2) represent matrices with as many rows as there are explanatory variables and as many columns as there are categories (not including the reference category, i.e., facies 3, for which the coefficients are by definition zero). The first column of β thus corresponds to the distinction between facies 1 and facies 3, while the second column corresponds to the distinction between facies 2 and facies 3 (compare equation (4)). The p-values, computed from the t-statistics of the individual components of β, suggest that all coefficients are statistically significant (i.e., $p<0.05$). The only exception is a p-value of 0.34 for FWI σ, which suggest that FWI σ does not provide additional significant information to distinguish between facies 2 and 3. But note that FWI σ is significant to distinguish between facies 1 and 3 (indicated by $p<0.05$, one position further left in Table 2).

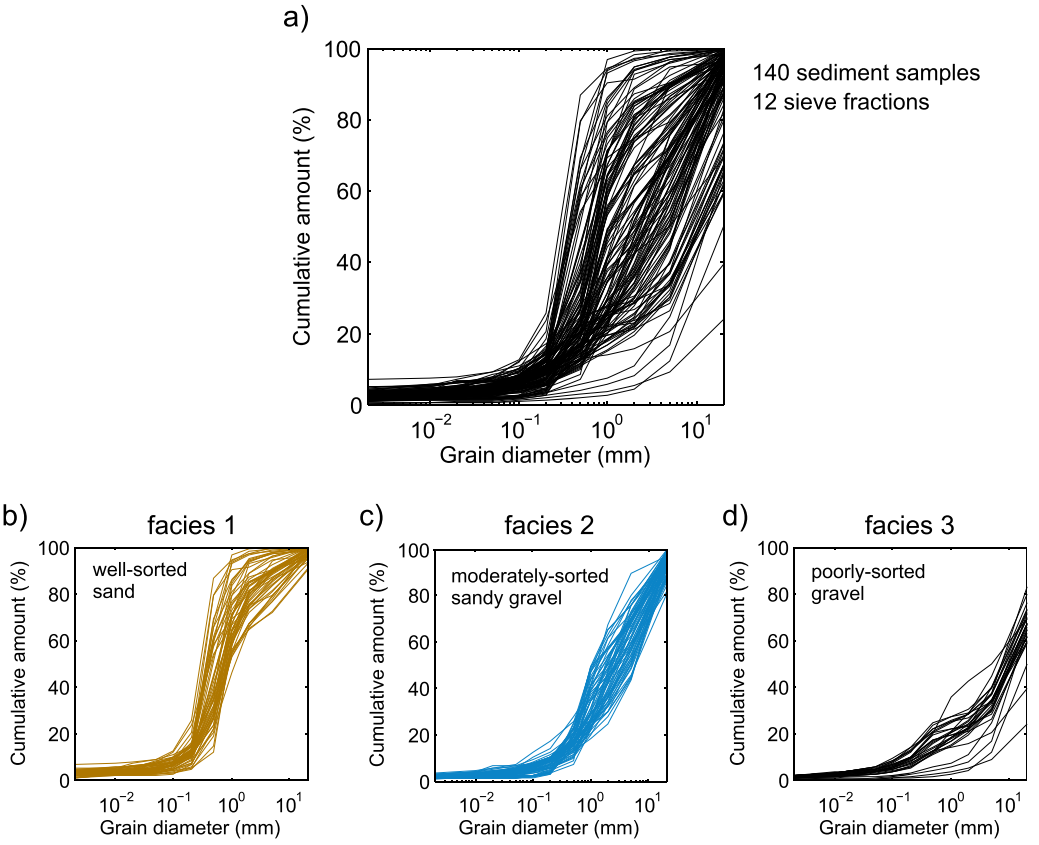

Figure 10. (a) Ensemble of 140 grain size distribution curves. (b–d) Results from k-means clustering.

Table 1. Characteristic Properties of Grain Size Facies With Standard Deviations (in Brackets)

Facies	Clay (%)	Silt (%)	Sand (%)	Gravel (%)	U[a]
1	3.1 (1.1)	2.8 (1.3)	76 (11)	18 (10)	5 (2)
2	2.0 (0.5)	2.4 (1.5)	41 (11)	55 (10)	19 (11)
3	1.4 (0.5)	2.9 (1.2)	18 (7)	78 (8)	60 (35)

[a]$U = d_{60}/d_{10}$, the coefficient of uniformity.

The last two columns in Table 2 (GSD and AIC) give global measures that describe how well the different models explain the clustered grain size data. Based on these two metrics, the models listed in Table 2 get better from top to bottom, the FWI model best explains the grain size data. As expected, all models that contain GPR data as explanatory variable perform much better than the basic Intercept-Only model. The improvement using the GPR full-waveform instead of the ray-based inversion results seems only minor. As we will see later, however, this small improvement is quite important.

3.5. Comparison of GPR Predicted Facies Distribution With CPT Cone Resistance Logs

Figure 11 shows for one GPR transect the facies distribution predicted by the Ray-based, the FWI (ϵ only) and the FWI logistic regression model. Information about the reliability of the facies classification is included by scaling the color saturation proportional to the probability of facies assignment predicted by the logistic regression model (equations (5) and (6)). Full color saturation refers to a probability of 1, zero color saturation refers to a probability of 0.33 (the lowest possible value for three facies).

Overall, the facies distribution obtained from the three different logistic regression models is quite similar for facies 3, but considerable differences exist regarding the distribution of facies 1 and 2. The GPR predicted facies distribution can be compared with CPT cone resistance logs (black lines in Figure 11) which were measured in close proximity to the GPR transect (Figure 1). While the CPT profiles show only very little correlation with the ray-based predicted facies distribution, they correlate quite well with the FWI predicted facies distribution: low cone resistance values coincide with the depth intervals occupied by facies 1; intermediate and high cone resistance values coincide with the depth intervals occupied by facies 2 and 3, respectively. Note that this order is also what we would expect based on the gravel contents of the individual grain size facies (Table 1).

It is noteworthy that the greatest part of the improvement of the FWI model relative to the ray-based model is already obtained when using the FWI (ϵ only) model. This shows that the improvement is not only due to the additional information from electrical conductivity but also due to the higher resolution of the dielectric permittivity obtained from full-waveform inversion. But including the electrical conductivity improves the consistency of facies across boreholes B62 and B61 (Figure 11c). This shows that the electrical conductivity contributes valuable information for facies estimation.

3.6. Comparison of GPR Predicted Facies Distribution With Tracer Breakthrough

Figures 12a and 12b show, respectively, the ray-based and the FWI predicted facies distribution in the part of the aquifer where the two tracer experiments were conducted. The transects parallel to the x-axis are the same transects that were monitored with ERT during the tracer tests (section 2.6.3). The red dashed and solid lines show the outlines of the ERT observed tracer breakthrough for the negative and positive tracer experiment, respectively (compare with Figure 2).

Table 2. Results of the Logistic Regression of GPR and Grain Size Data

Model	Variables	β[a]		t-Statistic		p-Value		GSD[b]	AIC[c]
Intercept-Only	Constant	0.68 (0.15)	0.93 (0.14)	4.56	6.52	<0.05	<0.05	46%	780
Ray-based	Constant	−23.17 (2.29)	−9.16 (1.42)	−10.13	−6.43	<0.05	<0.05	69%	562
	Ray-based ϵ	1.60 (0.15)	0.73 (0.10)	10.43	7.02	<0.05	<0.05		
FWI (ϵ only)	Constant	−20.42 (1.98)	−8.81 (1.43)	−10.32	−6.14	<0.05	<0.05	73%	558
	FWI ϵ	1.41 (0.13)	0.70 (0.10)	10.48	6.73	<0.05	<0.05		
FWI	Constant	−19.43 (1.99)	−8.75 (1.47)	−9.73	−5.97	<0.05	<0.05	76%	545
	FWI ϵ	1.16 (0.15)	0.77 (0.12)	7.60	6.30	<0.05	<0.05		
	FWI σ	0.24 (0.12)	−0.10 (0.10)	2.08	−0.96	<0.05	0.34		

[a]Maximum likelihood estimates of regression coefficients and standard errors (in brackets).
[b]Proportion of cells with colocated GPR and grain size data, where the modeled facies classification (i.e., highest probability predicted by the logistic regression model) matches the observed facies classification (clustered grain size data along boreholes).
[c]Akaikes Information Criterion (AIC) = 2 p − 2 L, with p the number of model parameters and L the log-likelihood of the fitted model.

The main difference in Figure 12 between ray-based and FWI results is the distribution of facies 1. Note that facies 1 is the facies with the finest grain size distribution which can therefore be expected to be less hydraulically conductive than facies 2 and 3. The ray-based approach shows a thick facies 1 layer between approximately 5 and 8 m depth which is interrupted at some places by facies 2. The FWI approach, in contrast, suggests that this layer is split up between boreholes B62 and B22 into two thinner layers of facies 1, and that between these two layers, there is a thin facies 2 layer at approximately 5–6 m depth that is laterally more or less continuous. At the bottom of this facies 2 layer, there is again a thin facies 1 layer found. The consequence of this facies structure derived from FWI for transport of a tracer that is injected into borehole B22 is that part of the tracer plume will be transported through the upper continuous facies 2 layer (between 5 and 6 m depth) whereas the lower facies 1 layer will lead to a split of the tracer plume (indicated by red arrows in Figure 12). In contrast, the facies structure derived from the ray-based approach would suggest tracer transport mainly through facies 2 below 7 m depth. This finding suggests that the improvement in resolution due to full-waveform inversion is crucial to detect small-scale aquifer structures which are highly important for solute transport.

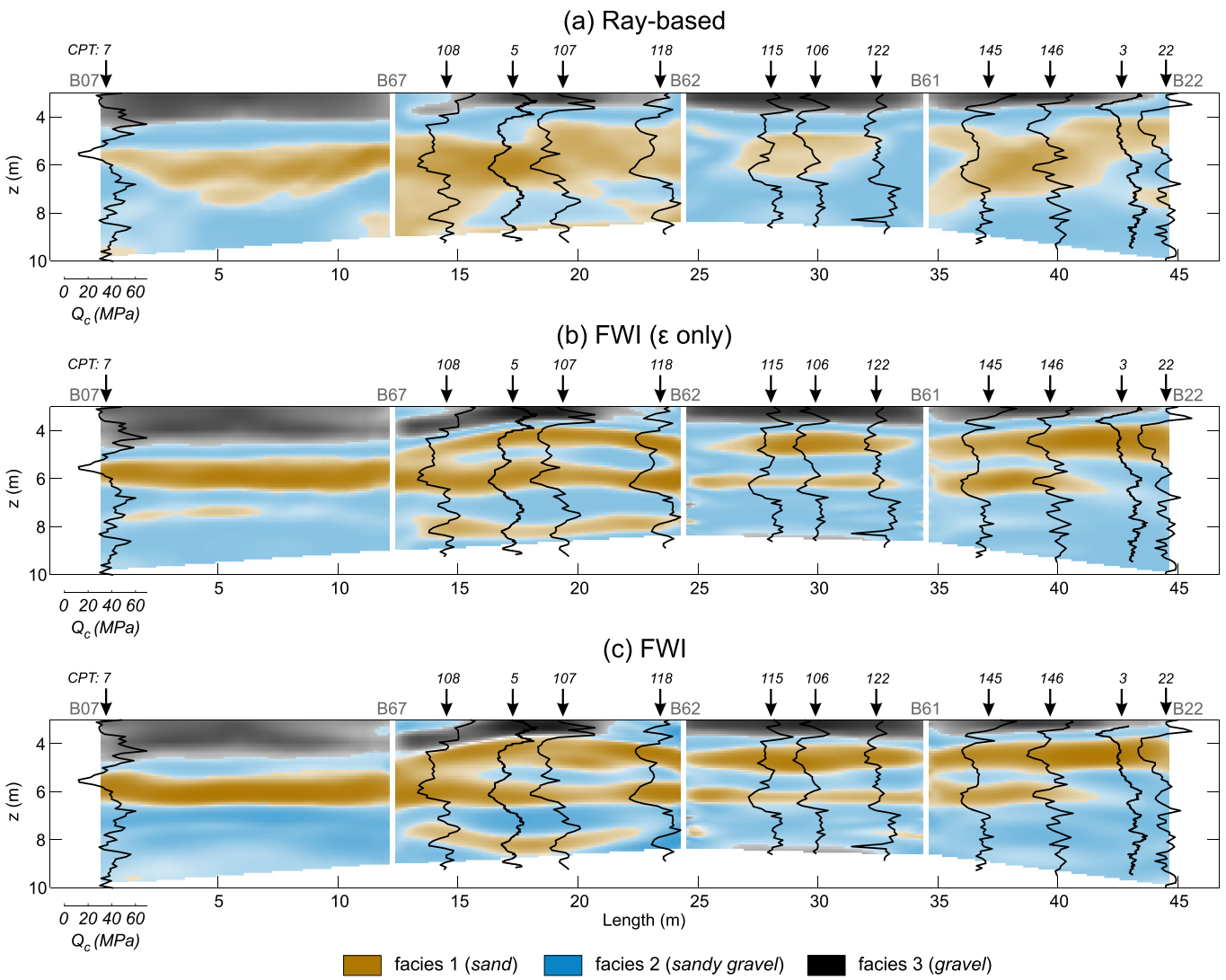

Figure 11. GPR predicted facies distribution using (a) ray-based dielectric permittivity, (b) full-waveform dielectric permittivity, (c) full-waveform dielectric permittivity and electrical conductivity (see Table 2). Color saturation is proportional to the probability of facies assignment predicted by the logistic regression model. Vertical gaps represent boreholes. Black lines represent CPT cone resistance logs.

The correlation between the tracer breakthrough and the FWI predicted facies distribution suggests that the derived facies are meaningful with regard to flow and transport. In the following, we will test this hypothesis by investigating the porosity and the hydraulic conductivity of the GPR FWI facies.

3.7. Porosity of GPR Facies

The porosity of the FWI predicted GPR facies was estimated from CPT neutron logs and from GPR dielectric permittivity (see section 3.2). The porosity of each facies, based on GPR and CPT data, is shown in Figure 13. Note that the portion of the aquifer sampled by GPR and CPT is different, because CPT porosity is available along 17 vertical profiles, whereas GPR porosity is available for each and every GPR pixel. Nevertheless, the porosity estimates from GPR and CPT are relatively consistent. Both methods suggest that the porosity decreases from facies 1, to facies 2, to facies 3, and that the porosity distribution of each facies is distinct (i.e., different from the porosity distribution of the other two facies).

3.8. Hydraulic Conductivity of GPR Facies

The hydraulic conductivity of the FWI predicted GPR facies was estimated from the DPST and DPIL profiles; and from flowmeter (equation (7)) and grain size data (equation (8)). Colocated with the GPR transects, flowmeter data were available along boreholes B22, B26, B31 and B32; and grain size data were available along boreholes B7, B22 and B32. Unfortunately, there were no flowmeter data available at locations colocated with facies 3, because the flowmeter measurements were conducted only below approximately 4 m depth, and facies 3 occurs mainly above 4 m depth (Figure 12).

Figure 14 shows the hydraulic conductivity of each facies according to DPST, DPIL, flowmeter and grain size. Overall, it can be noted that the different methods show a fairly similar trend, i.e., an increase in hydraulic conductivity from facies 1, to facies 2, to facies 3. This suggests that the GPR-derived facies are meaningful with regard to hydraulic conductivity, although the overlap of the hydraulic conductivity distributions for individual facies indicates that the GPR facies are not as well constrained by hydraulic conductivity as they are by porosity (section 3.7).

Interestingly, the highest hydraulic conductivity is obtained for facies 3, although no tracer breakthrough was observed above 4 m depth, where facies 3 primarily occurs (Figure 12b). This is most likely a consequence of the tracer injection into borehole B22, which is equipped with a filter screen ranging from 2.9 to 10.9 m depth. As facies 3 seems to occur only at the very top of the screened interval (the uppermost GPR pixels in Figure 12 correspond to a depth of 2.86 m depth), it can be expected that the amount of tracer injected into this uppermost layer was only small, which explains that no tracer mass was detected in this layer farther downstream.

The lowest hydraulic conductivity is obtained for facies 1, which is in agreement with the hypothesis that the facies 1 layer caused the splitting of the tracer plume (section 3.6). It is intriguing, however, that the distinction of facies 1 and 2 (the two transport relevant facies) is not very clear when looking at their hydraulic conductivity distributions which overlap (in particular the DPIL-derived distributions). Nevertheless, the difference in the average properties of the two facies was apparently large enough to have an important impact on transport in the aquifer because the tracer tests do show a clear and consistent effect of the facies structure on the observed tracer transport. This indicates that the structure and especially the connectivity of the facies is important even though the measured variability of local hydraulic properties within a facies is large.

Although the hydraulic conductivity derived from direct-push, flowmeter and grain size data agree fairly well regarding the trend, they show quite some disagreement regarding the absolute values (note the different x-axes in Figure 14). In order to evaluate which method best explains the observed tracer arrival times, we derived hydraulic conductivity estimates from the tracer test data, by applying equation (9) with $i = 0.002$, the mean hydraulic gradient during the tracer experiments [*Müller et al.*, 2010], and $n = 0.25$, the mean porosity of GPR FWI facies 2 (because the tracer breakthrough occurred primarily through facies 2 (Figure 12b)).

Figure 15 shows a comparison of the tracer test-derived hydraulic conductivity and the hydraulic conductivity of facies 2 according to direct-push, flowmeter and grain size data. It can be seen that in order to explain the observed tracer arrival times, log hydraulic conductivities of around -3 are required. This is approximately half an order of magnitude higher than the direct-push derived hydraulic conductivity, and

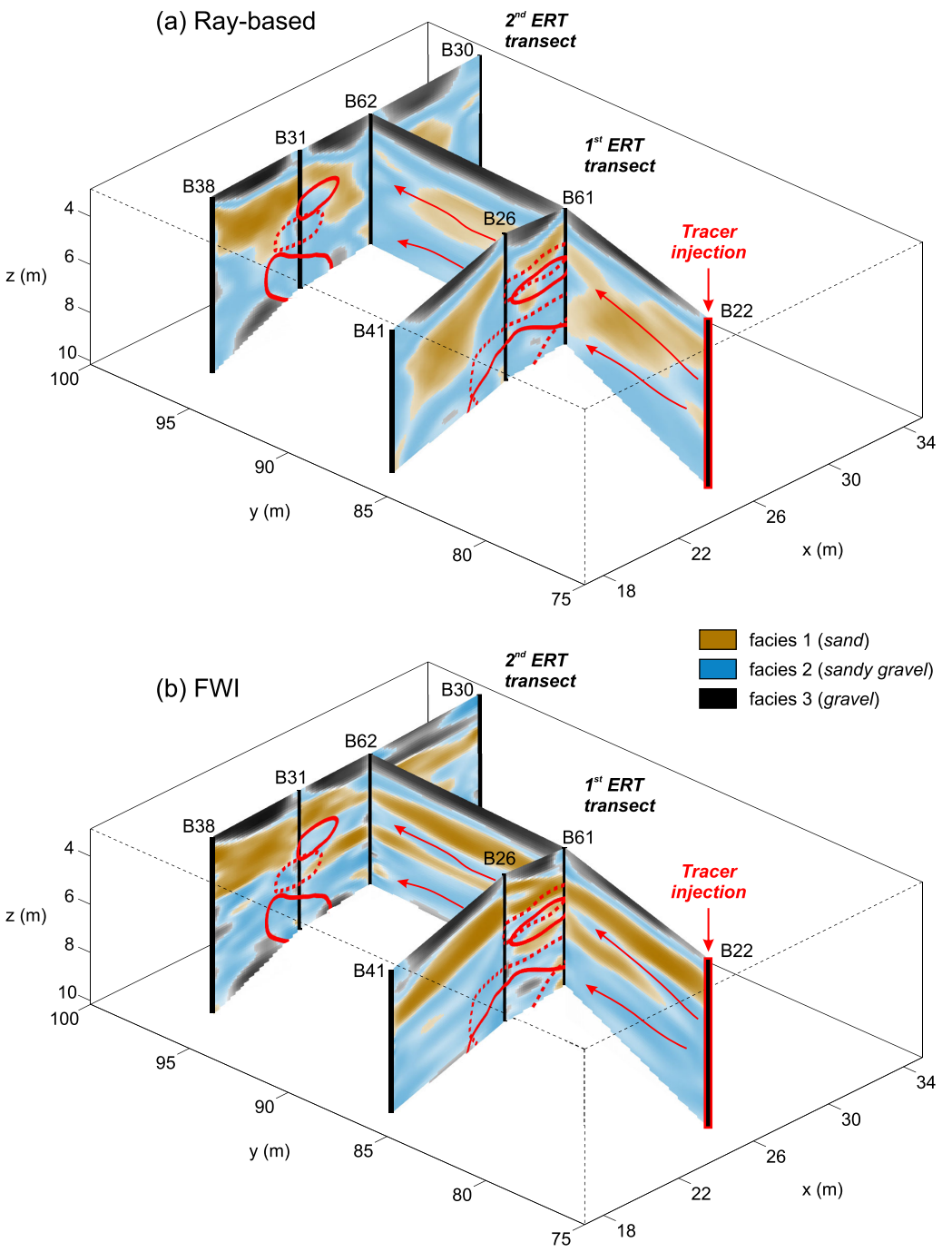

Figure 12. Comparison of ERT observed tracer breakthrough and GPR predicted facies distribution using (a) ray-based dielectric permittivity, (b) full-waveform dielectric permittivity and electrical conductivity. Color saturation is proportional to the probability of facies assignment predicted by the logistic regression model. Red dashed and solid lines show the outlines of the ERT observed tracer plumes projected on the GPR tomograms (compare with Figure 2).

approximately half an order of magnitude lower than the flowmeter-derived hydraulic conductivity. The grain size based estimates lie between the direct-push and flowmeter results and show the best agreement with the tracer test-derived hydraulic conductivity.

The differences in absolute hydraulic conductivities obtained from the different methods may be attributed to different sensitivities, support volumes and biases of the different measurements techniques. Particularly at strongly heterogeneous sites, different methods have been reported to give deviating results. For

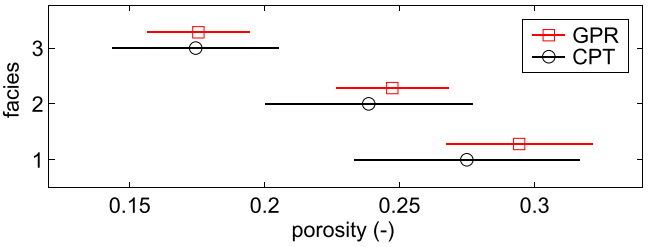

Figure 13. Porosity of the FWI predicted GPR facies according to CPT neutron logs and GPR dielectric permittivity. Horizontal lines represent two times the standard deviation.

example, *Rehfeldt et al.* [1992] statistically analyzed hydraulic conductivity at the MADE site on the basis of flowmeter, grain size, slug tests and laboratory permeameter tests and found that the different methods yielded hydraulic conductivity distributions with comparable covariance parameters but largely varying mean values. *Bohling et al.* [2012] compared flowmeter and direct-push derived hydraulic conductivity and found that the two methods identified similar structures in the subsurface, but the direct-push estimates were on average about a factor 5–10 smaller compared with the flowmeter results.

One explanation for lower hydraulic conductivity measured with direct-push is that direct-push tools may locally compress the subsurface structure and thereby decrease hydraulic conductivity. Another practical difficulty consists in keeping the injection screen free from clogging because this can be misinterpreted as a decrease in hydraulic conductivity [*Dietrich et al.*, 2008]. Although it was found that the continuous injection of water during DPIL tool advancement [*Dietrich et al.*, 2008] and the use of shielded screens for DPST measurements [*Butler et al.*, 2002] can mitigate or prevent screen clogging, under unfavorable conditions screen clogging may still occur and lead to an underestimation of hydraulic conductivity. *Butler et al.* [2002] discussed the implication of well development on DPST results and noted that insufficient development of the direct-push installed well can lead to significant underestimation of hydraulic conductivity by DPST.

Flowmeter-derived hydraulic conductivities, on the other hand, are well known to be sensitive to skin effects, filter pack, ambient flow in the well and head losses in well and screen [*Rehfeldt et al.*, 1989; *Molz et al.*, 1994]. *Ruud et al.* [1999] demonstrated numerically that the combined effect of flowmeter-induced hydraulic head losses and cross-flow through the filter pack can easily lead to an overestimation of the true hydraulic conductivity by a factor 10.

Finally, we want to note that the method-based discrepancy in absolute hydraulic conductivity observed in the present study (approximately one order of magnitude) appears less dramatic if one considers the enormous variability (several orders of magnitude) of hydraulic conductivity in the subsurface.

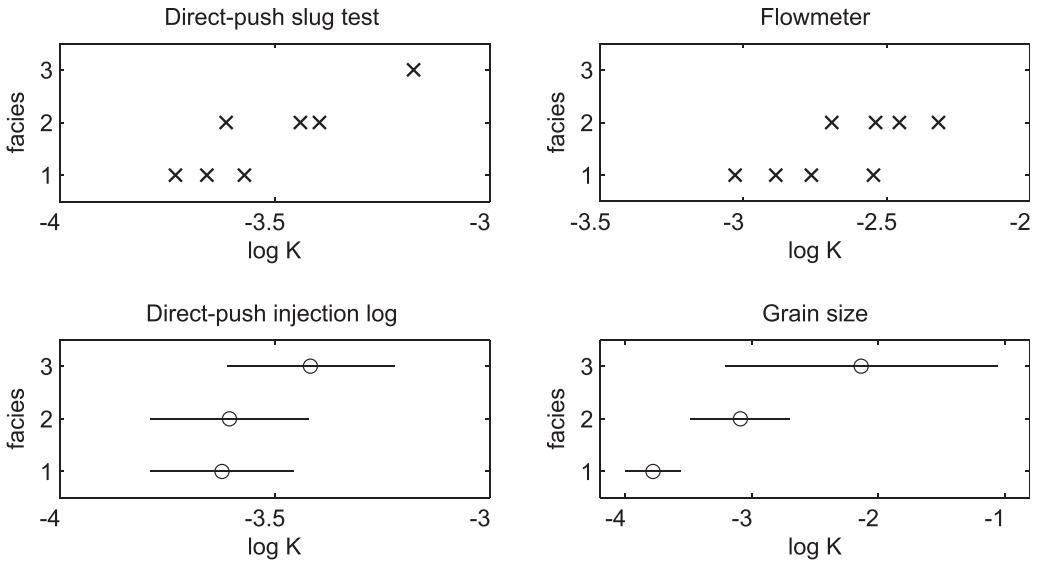

Figure 14. Hydraulic conductivity of the FWI predicted GPR facies according to direct-push slug test, direct-push injection log, flowmeter and grain size data (note the different x-axes!). Horizontal lines represent two times the standard deviation.

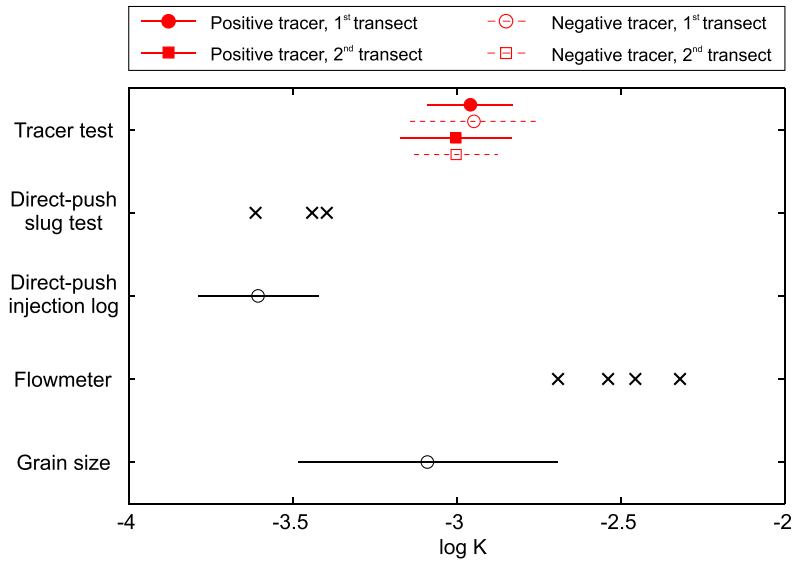

Figure 15. Tracer test-derived hydraulic conductivity and comparison with the hydraulic conductivity of GPR FWI facies 2 according to direct-push slug test, direct-push injection log, flowmeter and grain size data. Horizontal lines represent two times the standard deviation.

4. Conclusions

High resolution tomographic images of the electrical properties of an alluvial aquifer were obtained from full-waveform inversion of cross-borehole GPR data. By analyzing multiple crosshole planes and stitching together the tomograms from adjacent planes, we were able to image, with a decimeter scale resolution, the dielectric permittivity and electrical conductivity along several cross sections up to 50 m in length and 10 m in depth.

Porosity estimates derived from the GPR results using the Complex Refractive Index Model (CRIM) showed a strong correlation with porosity estimates derived from CPT neutron logs (correlation coefficient of 0.8). A less pronounced correlation was observed between the GPR results and the hydraulic conductivity from direct-push injection logs (correlation coefficient of −0.4).

By employing a logistic regression model that relates the electrical properties obtained from GPR to lithological facies obtained from grain size data, we utilized the GPR full-waveform inversion results to predict the spatial distribution of lithological facies along the GPR transects. Vertical profiles of porosity and hydraulic conductivity from direct-push, flowmeter and grain size data suggest that the GPR predicted facies are meaningful with regard to porosity and hydraulic conductivity, even though the distributions of individual facies show some overlap. The hydraulic conductivity estimates from the different methods (direct-push, flowmeter, grain size) were found to identify similar trends but deviating absolute values. Taking tracer test-derived hydraulic conductivity as a reference, the direct-push derived hydraulic conductivities were approximately half an order of magnitude lower, and the flowmeter-derived hydraulic conductivities were approximately half an order of magnitude higher. Somewhat surprisingly, the indirect (i.e., purely empirical) estimation of hydraulic conductivity from grain size data was found to best explain the observed tracer arrival times.

Comparison of the GPR predicted facies distribution with tracer test data suggests that the plume splitting observed in a tracer experiment was caused by a hydraulically low-conductive sand layer with a thickness of only a few decimeters. Because this sand layer was detected by GPR full-waveform inversion but not by conventional GPR ray-based inversion we conclude that the improvement in spatial resolution due to full-waveform inversion is crucial to detect small-scale aquifer structures that are highly important for solute transport.

We hope that the results presented here will stimulate further applications of GPR full-waveform inversion. Admittedly, the results from intensively studied test sites, such as the Krauthausen site, must be regarded as best case scenarios of what information a method can provide. Yet full-waveform inversion of crosshole

GPR data has been tested under different conditions at several sites now [*Klotzsche et al*., 2013, 2014; *Gueting et al*., 2015], and we believe that hydrological site characterization could often benefit from its application (provided that the fundamental requirement for any GPR application is fulfilled, i.e., signal attenuation due to electrical conductivity is not too high).

A particularly interesting application in our view would be to use tomograms from GPR full-waveform inversion as "training images" for multiple-point-statistical (MPS) simulations [e.g., *Mariethoz and Caers*, 2014]. In this framework, GPR full-waveform inversion would be used to collect site-specific high resolution information about the geological structures that exist at a site; and MPS simulations would be used to stochastically derive 3-D subsurface models based on the structural information provided by the GPR tomograms. The resulting models could be parametrized for flow and transport simulations.

Acknowledgments
This work was supported in part by the TERrestrial ENvironmental Observatories (TERENO) and in part by the Transregional Collaborative Research Centre 32 (TR32) Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation. We want to thank Carly Hyatt, Manuel Kreck and Helko Kotas whose assistance during the field measurements is highly appreciated. We also want to thank Lee Slater and three anonymous reviewers whose suggestions strongly improved this paper. Data are available from the authors upon request.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, *19*(6), 716–723.
Ayalew, L., and H. Yamagishi (2005), The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan, *Geomorphology*, *65*, 15–31.
Becht, A., J. Tronicke, E. Appel, and P. Dietrich (2004), Inversion strategy in crosshole radar tomography using information of data subsets, *Geophysics*, *69*(1), 222–230.
Belina, F., J. Ernst, and K. Holliger (2009), Inversion of crosshole seismic data in heterogeneous environments: Comparison of waveform and ray-based approaches, *J. Appl. Geophys.*, *68*, 85–94.
Bialas, Z., and A. Kleczkowski (1970), Über den praktischen Gebrauch von einigen empirischen Formeln zur Berechnung des Durchlässigkeitskoeffzienten, *K. Archivum Hydrotechniki, Warschau*.
Binley, A., S. Hubbard, J. Huisman, A. Revil, D. Robinson, K. Singha, and L. Slater (2015), The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales, *Water Resour. Res.*, *51*, 3837–3866, doi:10.1002/2015WR017016.
Birchak, J., C. Gardner, J. Hipp, and J. Victor (1974), High dielectric constant microwave probes for sensing soil moisture, *Proc. IEEE*, *62*(1), 93–98.
Bohling, G. C., G. Liu, S. J. Knobbe, E. C. Reboulet, D. W. Hyndman, P. Dietrich, and J. J. Butler (2012), Geostatistical analysis of centimeter-scale hydraulic conductivity variations at the MADE site, *Water Resour. Res.*, *48*, W02525, doi:10.1029/2011WR010791.
Brovelli, A., and G. Cassiani (2008), Effective permittivity of porous media: A critical analysis of the complex refractive index model, *Geophys. Prospect.*, *56*, 715–727.
Butler, J., E. Garnett, and J. Healey (2003), Analysis of slug tests in formations of high hydraulic conductivity, *Ground Water*, *41*(5), 620–630.
Butler, J. J. (1997), *The Design, Performance, and Analysis of Slug Tests*, A. F. Lewis, Boca Raton, Fla.
Butler, J. J., J. M. Healey, G. W. McCall, E. J. Garnett, and S. P. Loheide (2002), Hydraulic tests with direct-push equipment, *Ground Water*, *40*(1), 25–36.
Carmichael, R. S. (1982), *Handbook of Physical Properties of Rocks*, CRC Press, Boca Raton, Fla.
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*, 305–314.
Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, New York.
Davis, J., and A. Annan (1989), Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy, *Geophys. Prospect.*, *37*, 531–551.
Dietrich, P., J. J. Butler, and K. Faiß (2008), A rapid method for hydraulic profiling in unconsolidated formations, *Ground Water*, *46*(2), 323–328.
Dietze, M., and P. Dietrich (2012), Evaluation of vertical variations in hydraulic conductivity in unconsolidated sediments, *Ground Water*, *50*(3), 450–456.
Dobson, M., F. Ulaby, M. Hallikainen, and M. El-rayes (1985), Microwave dielectric behavior of wet soil-part II: Dielectric mixing models, *IEEE Trans. Geosci. Remote Sens.*, *GE-23*(1), 35–46.
Doetsch, J., N. Linde, I. Coscia, S. Greenhalgh, and A. Green (2010), Zonation for 3D aquifer characterization based on joint inversions of multimethod crosshole geophysical data, *Geophysics*, *75*(6), G53–G64.
Dogan, M., R. L. Van Dam, G. C. Bohling, J. J. Butler, and D. W. Hyndman (2011), Hydrostratigraphic analysis of the MADE site with full-resolution GPR and direct-push hydraulic profiling, *Geophys. Res. Lett.*, *38*, L06405, doi:10.1029/2010GL046439.
Döring (1997), Transport der reaktiven Stoffe Eosin, Uranin und Lithium in einem heterogenen Grundwasserleiter, PhD thesis, Christian Albrecht Univ. Kiel, Kiel, Germany.
Eckhardt, D. A. V., and P. E. Stackelberg (1995), Relation of ground-water quality to land use on Long Island, New York, *Ground Water*, *33*(6), 1019–1033.
Eisenberg, D. S., and W. Kauzmann (1969), *The Structure and Properties of Water*, Oxford Univ. Press, New York.
Englert, A. (2003), Measurement, estimation and modelling of groundwater flow velocity at Krauthausen Test Site, PhD thesis, RWTH Aachen, Aachen, Germany.
Ernst, J., H. Maurer, A. Green, and K. Holliger (2007), Full-waveform inversion of crosshole radar data based on 2-D finite-difference time-domain solutions of Maxwell's equations, *IEEE Trans. Geosci. Remote Sens.*, *45*(9), 2807–2828.
Gueting, N., A. Klotzsche, J. van der Kruk, J. Vanderborght, H. Vereecken, and A. Englert (2015), Imaging and characterization of facies heterogeneity in an alluvial aquifer using GPR full-waveform inversion and cone penetration tests, *J. Hydrol.*, *524*, 680–695.
Helsel, D., and R. Hirsch (2002), *Statistical Methods in Water Resources*, U.S. Geol. Surv, Reston, Va.
Holliger, K., and T. Bergmann (2002), Numerical modeling of borehole georadar data, *Geophysics*, *67*, 1249–1257.
Holliger, K., M. Musil, and H. Maurer (2001), Ray-based amplitude tomography for crosshole georadar data: A numerical assessment, *J. Appl. Geophys.*, *47*, 285–298.
Hosmer, D. W., and S. Lemeshow (2000), *Applied Logistic Regression*, Wiley-Interscience, New York.
Hubbard, S., J. Chen, J. E. Peterson, E. L. Mayer, K. H. Williams, D. J. Swift, B. Mailloux, and Y. Rubin (2001), Hydrogeological characterization of South Oyster Bacterical Transport Site using geophysical data, *Water Resour. Res.*, *37*(10), 2431–2456.

Hyndman, D. W., and S. M. Gorelick (1996), Estimating lithologic and transport properties in three dimensions using seismic and tracer data: The Kesterson aquifer, *Water Resour. Res.*, *32*(9), 2659–2670.
Hyndman, D. W., J. M. Harris, and S. M. Gorelick (2000), Inferring the relation between seismic slowness and hydraulic conductivity in heterogeneous aquifers, *Water Resour. Res.*, *36*(8), 2121–2132.
Irving, J., M. Knoll, and R. Knight (2007), Improving crosshole radar velocity tomograms: A new approach to incorporating high-angle traveltime data, *Geophysics*, *72*(4), J31–J41.
Kemna, A., J. Vanderborght, B. Kulessa, and H. Vereecken (2002), Imaging and characterisation of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models, *J. Hydrol.*, *267*, 125–146.
Klotzsche, A., J. van der Kruk, G. Meles, J. Doetsch, H. Maurer, and N. Linde (2010), Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the Thur River, Switzerland, *Near Surf. Geophys.*, *8*, 635–649.
Klotzsche, A., J. van der Kruk, N. Linde, J. Doetsch, and H. Vereecken (2013), 3-D characterization of high-permeability zones in a gravel aquifer using 2-D crosshole GPR full-waveform inversion and waveguide detection, *Geophys. J. Int.*, *195*, 932–944.
Klotzsche, A., J. van der Kruk, J. Bradford, and H. Vereecken (2014), Detection of spatially limited high-porosity layers using crosshole GPR signal analysis and full-waveform inversion, *Water Resour. Res.*, *50*, 6966–6985, doi:10.1002/2013WR015177.
Koltermann, C., and S. Gorelick (1996), Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches, *Water Resour. Res.*, *32*(9), 2617–2658.
Kurzmann, A., A. Przebindowska, D. Kohn, and T. Bohlen (2013), Acoustic full waveform tomography in the presence of attenuation: A sensitivity analysis, *Geophys. J. Int.*, *195*(2), 985–1000.
Lessoff, S. C., U. Schneidewind, C. Leven, P. Blum, P. Dietrich, and G. Dagan (2010), Spatial characterization of the hydraulic conductivity using direct-push injection logging, *Water Resour. Res.*, *46*, W12502, doi:10.1029/2009WR008949.
Li, W., A. Englert, O. A. Cirpka, J. Vanderborght, and H. Vereecken (2007), Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests, *Water Resour. Res.*, *43*, W04433, doi:10.1029/2006WR005333.
Li, W., A. Englert, O. A. Cirpka, and H. Vereecken (2008), Three-dimensional geostatistical inversion of flowmeter and pumping test data, *Ground Water*, *46*(2), 193–201.
Linde, N., and J. A. Vrugt (2013), Distributed soil moisture from crosshole ground-penetrating radar travel times using stochastic inversion, *Vadose Zone J.*, *12*(1), doi:10.2136/vzj2012.0101.
Linde, N., S. Finsterle, and S. Hubbard (2006), Inversion of tracer test data using tomographic constraints, *Water Resour. Res.*, *42*, W04410, doi:10.1029/2004WR003806.
Linde, N., A. Tryggvason, J. E. Peterson, and S. S. Hubbard (2008), Joint inversion of crosshole radar and seismic traveltimes acquired at the South Oyster Bacterial Transport Site, *Geophysics*, *73*(4), G29–G37.
MacQueen, J. B. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, edited by L. M. L. Cam and J. Neyman, pp. 281–297, Univ. of Calif. Press, Berkeley, Calif.
Mariethoz, G., and J. Caers (2014), *Multiple-Point Geostatistics: Stochastic Modeling With Training Images*, John Wiley, Hoboken, N. J.
Maurer, H., and A. Green (1997), Potential coordinate mislocations in crosshole tomography: Results from the Grimsel test site, Switzerland, *Geophysics*, *62*(6), 1696–1709.
Maurer, H., and M. Musil (2004), Effects and removal of systematic errors in crosshole georadar attenuation tomography, *J. Appl. Geophys.*, *55*, 261–270.
Meles, G. A., J. van der Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and A. G. Green (2010), A new vector waveform inversion algorithm for simultaneous updating of conductivity and permittivity parameters from combination crosshole/borehole-to-surface GPR data, *IEEE Trans. Geosci. Remote Sens.*, *48*(9), 3391–3407.
Miall, A. (1996), *The Geology of Fluvial Deposits*, Springer, New York.
Molz, F., R. Morin, A. Hess, J. Melville, and O. Güven (1989), The impeller meter for measuring aquifer permeability variations: Evaluation and comparison with other tests, *Water Resour. Res.*, *25*(7), 1677–1683.
Molz, F., G. Boman, S. Young, and W. Waldrop (1994), Borehole flowmeters: Field application and data analysis, *J. Hydrol.*, *163*, 347–371.
Morin, R. H. (2006), Negative correlation between porosity and hydraulic conductivity in sand-and-gravel aquifers at Cape Cod, Massachusetts, USA, *J. Hydrol.*, *316*, 43–52.
Müller, K., J. Vanderborght, A. Englert, A. Kemna, J. A. Huisman, J. Rings, and H. Vereecken (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*, W03502, doi:10.1029/2008WR007595.
Nelson, P. (1994), Permeability-porosity relationships in sedimentary rocks, *Log Anal.*, *3*, 38–62.
Nolan, B. T. (2001), Relating nitrogen sources and aquifer susceptibility to nitrate in shallow ground waters of the United States, *Ground Water*, *39*(2), 290–299.
Oberröhrmann, M., A. Klotzsche, H. Vereecken, and J. van der Kruk (2013), Optimization of acquisition setup for cross-hole GPR full-waveform inversion using checkerboard analysis, *Near Surf. Geophys.*, *11*, 197–209.
Rehfeldt, K. R., P. Hufschmied, L. Gelhar, and M. Schaefer (1989), *Measuring hydraulic conductivity with the borehole flowmeter*, *Tech. Rep.* EN-6511, Electr. Power Res. Inst., Palo Alto, Calif.
Rehfeldt, K. R., K. M. Boggs, and L. W. Gelhar (1992), Field study of dispersion in a heterogeneous aquifer, 3. Geostatistical analysis of hydraulic conductivity, *Water Resour. Res.*, *28*(12), 3309–3324.
Roth, K., R. Schulin, H. Flühler, and W. Attinger (1990), Calibration of time domain reflectometry for water content measurement using a composite dielectric approach, *Water Resour. Res.*, *26*(10), 2267–2273.
Ruud, N. C., Z. J. Kabala, and F. J. Molz (1999), Evaluation of flowmeter-head loss effects in the flowmeter test, *J. Hydrol.*, *224*, 55–63.
Scheibe, T. D., and Y. J. Chien (2003), An evaluation of conditioning data for solute transport prediction, *Ground Water*, *41*(2), 128–141.
Schmelzbach, C., J. Tronicke, and P. Dietrich (2011), Three-dimensional hydrostratigraphic models from ground-penetrating radar and direct-push data, *J. Hydrol.*, *398*, 235–245.
Steelman, C. M., and A. L. Endres (2011), Comparison of petrophysical relationships for soil moisture estimation using GPR ground waves, *Vadose Zone J.*, *10*, 270–285.
Tesoriero, A. J., E. L. Inkpen, and F. D. Voss (1998), Assessing ground-water vulnerability using logistic regression, in *Proceedings for the Source Water Assessment and Protection 98 Conference*, pp. 157–165, Dallas, Tex.
Tillmann, A., A. Englert, Z. Nyari, I. Fejes, J. Vanderborght, and H. Vereecken (2008), Characterization of subsoil heterogeneity, estimation of grain size distribution and hydraulic conductivity at the Krauthausen test site using Cone Penetration Test, *J. Contam. Hydrol.*, *95*, 57–75.

Topp, G., J. Davis, and A. Annan (1980), Electromagnetic determination of soil water content: Measurements in coaxial transmission lines, *Water Resour. Res.*, *16*(3), 574–582.

Tronicke, J., K. Holliger, W. Barrash, and M. D. Knoll (2004), Multivariate analysis of cross-hole georadar velocity and attenuation tomograms for aquifer zonation, *Water Resour. Res.*, *40*, W01519, doi:10.1029/2003WR002031.

van den Eeckhaut, M., T. Vanwalleghem, J. Poesen, G. Govers, G. Verstraeten, and L. Vandekerckhove (2006), Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium), *Geomorphology*, *76*, 392–410.

Vanderborght, J., and H. Vereecken (2001), Analyses of locally measured bromide breakthrough curves from a natural gradient tracer experiment at Krauthausen, *J. Contam. Hydrol.*, *48*, 23–43.

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

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

Vienken, T., and P. Dietrich (2011), Field evaluation of methods for determining hydraulic conductivity from grain size data, *J. Hydrol.*, *400*, 58–71.

Wainwright, H., J. Chen, H. Hardelauf, D. Sassen, and S. Hubbard (2014), Bayesian hierarchical approach and geophysical data sets for estimation of reactive facies over plume scales, *Water Resour. Res.*, *50*, 4564–4584, doi:10.1002/2013WR013842.

Yang, X., A. Klotzsche, G. Meles, H. Vereecken, and J. van der Kruk (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.