Estimating vadose zone hydraulic properties using ground penetrating radar: The impact of prior information

Marie Scholer,[1] James Irving,[1] Andrew Binley,[2] and Klaus Holliger[1]

Received 12 January 2011; revised 19 August 2011; accepted 19 August 2011; published 14 October 2011.

[1] A number of geophysical methods, such as ground-penetrating radar (GPR), have the potential to provide valuable information on hydrological properties in the unsaturated zone. In particular, the stochastic inversion of such data within a coupled geophysical-hydrological framework may allow for the effective estimation of vadose zone hydraulic parameters and their corresponding uncertainties. A critical issue in stochastic inversion is choosing prior parameter probability distributions from which potential model configurations are drawn and tested against observed data. A well chosen prior should reflect as honestly as possible the initial state of knowledge regarding the parameters and be neither overly specific nor too conservative. In a Bayesian context, combining the prior with available data yields a posterior state of knowledge about the parameters, which can then be used statistically for predictions and risk assessment. Here we investigate the influence of prior information regarding the van Genuchten-Mualem (VGM) parameters, which describe vadose zone hydraulic properties, on the stochastic inversion of crosshole GPR data collected under steady state, natural-loading conditions. We do this using a Bayesian Markov chain Monte Carlo (MCMC) inversion approach, considering first noninformative uniform prior distributions and then more informative priors derived from soil property databases. For the informative priors, we further explore the effect of including information regarding parameter correlation. Analysis of both synthetic and field data indicates that the geophysical data alone contain valuable information regarding the VGM parameters. However, significantly better results are obtained when we combine these data with a realistic, informative prior.

Citation: Scholer, M., J. Irving, A. Binley, and K. Holliger (2011), Estimating vadose zone hydraulic properties using ground penetrating radar: The impact of prior information, *Water Resour. Res., 47*, W10512, doi:10.1029/2011WR010409.

1. Introduction

[2] Water infiltration, evapotranspiration, solute transport, and soil erosion all take place in the vadose zone. These processes affect plant growth with respect to different atmospheric influences and the vulnerability of groundwater resources to pollution [e.g., *Haygarth and Jarvis*, 2002]. The accurate estimation of soil hydraulic parameters controlling vadose zone flow and transport processes is therefore of utmost importance. Conventional methods to estimate unsaturated soil hydraulic properties include laboratory measurements conducted on representative samples from the field as well as the monitoring of moisture content and/or pressure head in situ using techniques such as time domain reflectometry and tensiometery. A critical drawback of all of these methods is the small support volume of the measurements; that is, being made at essentially the point scale, the measurements are subject to significant variability and may not adequately represent larger-scale vadose zone processes [e.g., *Binley et al.*, 2002b]. Recently, much interest has been expressed in the use of geophysical methods in hydrology because they allow for the in situ estimation of subsurface properties at a larger, more relevant integral scale. Geophysical properties such as the high-frequency electromagnetic wave velocity and electrical conductivity are closely related to soil water content which means that, when combined with a hydrological process model, they may allow effective inference of vadose zone hydraulic properties [e.g., *Finsterle and Kowalsky*, 2008; *Looms et al.*, 2008a]. One commonly used geophysical method in this regard is ground-penetrating radar (GPR). *Parkin et al.* [2000] showed that moisture contents estimated from GPR data have a high reproducibility and accuracy and allow for the tracking of water front movement in the vadose zone. Both surface-based [e.g., *Greaves et al.*, 1996; *Huisman et al.*, 2002] and crosshole [e.g., *Binley et al.*, 2001; *Looms et al.*, 2008b; *Winship et al.*, 2006] GPR surveys have been used for vadose zone characterization.

[3] There is increasing interest in stochastic parameter estimation, or inversion, strategies in hydrology [e.g., *Vrugt et al.*, 2009b]. Stochastic inversion has the distinct advantage over deterministic inversion approaches of greater exploration of the space of model parameters, and is thus naturally suited to highly nonlinear and/or nonunique problems. Stochastic inversion can also provide, naturally, multiple sets of parameters that are consistent with the observed

[1]Institute of Geophysics, University of Lausanne, Lausanne, Switzerland.
[2]Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom.

0043-1397/11/2011WR010409

data, which facilitates the assessment of hydrological prediction uncertainty. The vast majority of stochastic inversion approaches that have been employed in hydrology operate within a Bayesian-style framework, whereby prior information about the model parameters is updated and refined on the basis of available measurements. In this regard, many researchers have performed hydrological parameter estimation in a formal Bayesian manner and used Markov chain Monte Carlo (MCMC) sampling to obtain multiple parameter realizations from the posterior probability distribution [e.g., *Hassan et al.*, 2009; *Irving and Singha*, 2010; *Vrugt et al.*, 2009a]. Others have used a pseudo-Bayesian strategy known as the generalized likelihood uncertainty estimation (GLUE) method, whereby multiple model configurations are proposed based on prior knowledge and are then accepted or rejected depending on some prescribed measure of acceptable data fit [e.g., *Beven and Binley*, 1992; *Freer et al.*, 1996; *Morse et al.*, 2003].

[4] The increased interest in both stochastic inversion and geophysical methods in vadose zone hydrology has led to a small number of recent studies where unsaturated zone parameter estimation based on geophysical data has been attempted in a stochastic manner. *Binley and Beven* [2003], for example, used zero-offset-profile (ZOP) crosshole GPR data to constrain the changes in moisture content caused by natural loading in the vadose zone during a 2-yr monitoring period. Assuming a homogeneous subsurface, the GPR-derived moisture content profiles were then used in a GLUE inversion strategy to identify "behavioral" sets of van Genuchten-Mualem (VGM) parameters [*Mualem*, 1976; *van Genuchten*, 1980] that all fit the data to within a prescribed degree. *Binley et al.* [2004] and *Cassiani and Binley* [2005] built on this work and used a similar strategy to estimate VGM parameters for a layered medium, where the layer boundaries were stochastically defined. *Looms et al.* [2008a] also used the GLUE methodology to assimilate crosshole GPR and electrical resistivity tomography (ERT) measurements, collected during a forced infiltration experiment, for the estimation of these parameters in a small number of subsurface layers. Finally, *Hinnell et al.* [2010] employed a Bayesian MCMC methodology to invert time-lapse electrical resistivity measurements for the VGM parameters in a synthetic example, their aim being to examine the potential benefits of directly coupling the geophysical and hydrological models in the inversion procedure. In all of these studies, it was found that the geophysical data allowed for some reduction in uncertainty regarding the VGM parameters. However, findings were largely inconclusive regarding the true utility of the geophysical measurements because of a lack of systematic testing on synthetic data and consideration of realistic prior knowledge in the inversion procedure.

[5] A critical part of any Bayesian or Bayesian-style inversion strategy is the appropriate definition and quantification of the prior information. Defining the prior probability distribution for a set of model parameters is not an easy task and is often done in a way that is mathematically most convenient and/or simple, rather than with the objective of representing our available knowledge in an optimal manner. In the stochastic inversion studies described above, for example, broad uniform prior distributions were chosen for the VGM parameters. The parameters were also assumed a priori to be statistically uncorrelated. Although these choices provided noninformative prior distributions that led to the posterior being shaped only by the geophysical measurements, they did not take into consideration a wealth of pertinent information regarding the basic nature of the VGM parameters that is available from soil property databases, the USDA ROSETTA database being one example [*Schaap et al.*, 2001]. Indeed, the analysis of a wide variety of soil samples has shown that, for a particular soil textural class, the VGM parameters tend to be markedly nonuniformly distributed and often strongly correlated [e.g., *Carsel and Parrish*, 1988; *de Rooij et al.*, 2004; *Mallants et al.*, 1996]. Given the strong influence of the prior on posterior uncertainties in Bayesian investigations, such information may allow for significantly improved parameter estimates and should be carefully considered [*Mertens et al.*, 2004; *Scharnagl et al.*, 2011]. Further, the use of a more informative prior distribution may also improve the efficiency of a stochastic inversion, as it limits the domain from which model parameters are sampled and tested [*Flores et al.*, 2010].

[6] In this paper, we investigate the effect of different types of prior knowledge regarding the subsurface VGM parameters on the stochastic inversion of ZOP crosshole GPR traveltime data, which are collected in the vadose zone under steady state conditions. In particular, we examine: (1) how much information regarding the VGM parameters is contained in the geophysical data through the use of a noninformative prior; and (2) whether posterior parameter estimates can be markedly improved using more informative, yet realistic, priors. Our work builds on the previously described research of *Binley and Beven* [2003], *Binley et al.* [2004], *Cassiani and Binley* [2005], and *Looms et al.* [2008a] involving the GLUE methodology for VGM parameter estimation from crosshole GPR data. However, we choose here to use a formal Bayesian MCMC inversion strategy, as was done by *Hinnell et al.* [2010], in order to take advantage of the computational benefits offered by the Markov chain sampling procedure and the stochastic rigor of this approach. We begin by outlining the hydrological and geophysical models used in this study, along with providing details regarding the inversion methodology. We then examine two synthetic examples of increasing complexity where we assess the information content of the geophysical data and the impact of prior information in the case where the "true" subsurface VGM parameters are known. Synthetic testing is a key part of validating any parameter estimation procedure and was lacking in most previous work in this domain. Finally, the corresponding methodologies are applied to crosshole GPR field data collected near Eggborough, Yorkshire, UK.

2. Methodological Background

2.1. Hydrological and Geophysical Models

[7] We assume in this work that vadose zone flow can be effectively described as only vertical, and that Richards' equation in 1-D can thus be used to predict the distribution of soil water content in the subsurface. To be consistent with much of the previous work in this domain, we also consider steady state conditions such that a single GPR-derived water content profile represents the long-term state of the system and can be used to estimate the soil hydraulic parameters [*Binley et al.*, 2002a; *Cassiani and Binley*,

2005]. Although it is clear that steady state conditions in the unsaturated zone are the exception, rather than the norm, these conditions provide a simplified, yet pertinent and illustrative, framework within which we can investigate the information content of the geophysical data and the value of prior information. The greater variability of water content in a transient problem could be expected to offer more information in the search for hydraulic parameters and will be examined in a future investigation. Note, however, that the relatively large support volume of borehole-based GPR measurements may limit our ability to observe such dynamic behavior in detail [*Cassiani and Binley*, 2005].

[8] In the steady state, Richards' equation in 1-D is given by

$$\frac{d}{dz}\left[K(h)\frac{dh}{dz}+K(h)\right]=0, \tag{1}$$

where K is the unsaturated hydraulic conductivity, h is the pressure head, and z is the elevation. We assume that $K(h)$ is dependent upon soil water content, $\theta(h)$, as parameterized by the VGM model

$$K(h)=K_sS_e^{\frac{1}{2}}\left[1-(1-S_e^{\frac{1}{m}})^m\right]^2, \tag{2}$$

where K_s is the saturated hydraulic conductivity and S_e is the effective saturation, given by

$$S_e=\frac{\theta(h)-\theta_r}{\theta_s-\theta_r}. \tag{3}$$

Here θ_r denotes the residual water content and θ_s the saturated water content or, equivalently, the porosity, and

$$\theta(h)=\theta_r+\frac{\theta_s-\theta_r}{(1+|\alpha h|^n)^m}, \tag{4}$$

where α, m, and n are empirical shape factors with $m=1-1/n$. A total of five model parameters (θ_s, θ_r, α, K_s, and n) therefore describe vadose zone hydraulic properties using the VGM model. To obtain the steady state, 1-D, water-content distribution corresponding to a given set of VGM parameters and specified boundary conditions, we use the program SS_INFIL [*Rockhold et al.*, 1997], which implements an integral solution to equation (1) for horizontally layered media. An arbitrary number of layers having different VGM parameters can be specified.

[9] To determine the set of ZOP crosshole GPR traveltimes corresponding to the 1-D water-content profile obtained with SS_INFIL, we require a relationship between water content and radar wave velocity. To this end, we use the complex refractive index model (CRIM) [*Roth et al.*, 1990] to predict the relative dielectric permittivity, κ, given the soil water content, θ:

$$\sqrt{\kappa}=(1-\theta_s)\sqrt{\kappa_s}+\theta\sqrt{\kappa_w}+(\theta_s-\theta)\sqrt{\kappa_a}, \tag{5}$$

where κ_s, κ_w, and κ_a are the relative permittivities of the dry matrix, water, and air components, respectively. The soil permittivity is then converted to GPR velocity through the following low-loss, high-frequency approximation that implicitly assumes a nonmagnetic subsurface and is valid in most environments amenable to GPR wave propagation [e.g., *Annan*, 2005]:

$$v\approx\frac{c}{\sqrt{\kappa}}, \tag{6}$$

where c is the free-space electromagnetic wave velocity. To determine the ZOP traveltimes based on the resulting 1-D velocity field, we solve the ray-based eikonal equation,

$$|\nabla t(\mathbf{r})|^2=s(\mathbf{r})^2, \tag{7}$$

where t is the traveltime of first-arriving energy from the transmitter to the receiver antenna at location $\mathbf{r}$ through the slowness field $s(\mathbf{r})=1/v(\mathbf{r})$. For ZOP measurements, the antennas are placed at the same depth in two adjacent boreholes and the traveltime between them is determined as a function of depth. We solve equation (7) using a MATLAB version of the PRONTO eikonal software of *Aldridge and Oldenburg* [1993], which accounts for bending of the radar wavefront at interfaces across which velocities change.

2.2. Bayesian MCMC Inversion

[10] As mentioned above, we consider a Bayesian MCMC stochastic inversion approach in this study. In an inversion context, Bayes' theorem can be written as follows [*Tarantola*, 2005]:

$$p(\mathbf{m}|\mathbf{d}_{\text{obs}})=\frac{p(\mathbf{d}_{\text{obs}}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d}_{\text{obs}})}, \tag{8}$$

where $p(\cdot)$ denotes a probability distribution, $\mathbf{m}$ is a vector containing the model parameters of interest, and $\mathbf{d}_{\text{obs}}$ is a vector containing the observed data. In equation (8), $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ is the conditional distribution for $\mathbf{m}$ given $\mathbf{d}_{\text{obs}}$ and hence the posterior probability we seek, $p(\mathbf{m})$ is the prior or marginal probability for the model parameters, and $p(\mathbf{d}_{\text{obs}}|\mathbf{m})$ is the conditional distribution for $\mathbf{d}_{\text{obs}}$ given $\mathbf{m}$, also known as the model likelihood. The marginal probability, $p(\mathbf{d}_{\text{obs}})$, can be treated as a normalization constant that ensures that $p(\mathbf{m}|\mathbf{d}_{\text{obs}})$ integrates to unity [*Tarantola*, 2005]. Assuming independent and identically normally distributed data residuals, the model likelihood in equation (8) can be described by the following equation:

$$p(\mathbf{d}_{\text{obs}}|\mathbf{m})=\frac{1}{(2\pi\sigma_r^2)^{N/2}}\exp\left[\frac{-\frac{1}{2}M}{\sigma_r^2}\right], \tag{9}$$

where N is the number of data, σ_r is the estimated standard deviation of the residuals, and M is the misfit, or summed squared residual, given by

$$M=(g[\mathbf{m}]-\mathbf{d}_{\text{obs}})^T(g[\mathbf{m}]-\mathbf{d}_{\text{obs}}). \tag{10}$$

Here $g(\cdot)$ represents the forward model linking the parameters of interest to the corresponding predicted data. In our case, it corresponds to the combined hydrological and geophysical models relating the VGM model parameters to the steady state ZOP GPR traveltime measurements.

[11] Because of the complexity of the forward models involved in hydrological and geophysical inverse problems, finding an analytical expression for the posterior distribution given by equation (8) is generally not possible. However, because the equation provides a means of calculating the posterior probability of occurrence of a particular set of model parameters, we can use MCMC methods to numerically generate samples from this distribution which can then be analyzed in terms of their posterior statistics. We base our MCMC inversion methodology on the seminal work of *Mosegaard and Tarantola* [1995] and refer the reader to that publication for full details. In the context of this study, the algorithm can be summarized as follows:

[12] 1. Draw randomly a starting set of values for the VGM model parameter vector, **m**, from the prior distribution for these parameters. Use the previously described hydrological and geophysical numerical models, along with equations (9) and (10), to compute the corresponding model likelihood, $p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m})$.

[13] 2. Perturb the current set of VGM parameters, **m**, to obtain a proposed set of parameters, $\mathbf{m}'$. This is done by drawing from a proposal density function, $Q(\mathbf{m}'|\mathbf{m})$, which depends on **m** and is chosen such that the size of the model perturbations allows for a reasonable rate of accepted transitions in the MCMC procedure, typically $\sim$30% [*Gilks et al.*, 1996]. For our work, we set $Q(\mathbf{m}'|\mathbf{m})$ to be a bounded uniform distribution that is centered on **m**.

[14] 3. Using the previously described hydrological and geophysical models, compute the likelihood of the proposed model, $p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m}')$.

[15] 4. Based on the prior probabilities of occurrence of **m** and $\mathbf{m}'$ and their corresponding likelihoods, decide to either accept or reject a transition to the proposed model using a stochastic decision rule. For this we use the Metropolis criterion [*Metropolis et al.*, 1953], whose probability of acceptance is given by

$$P_{\mathrm{acc}} = \frac{p(\mathbf{m}')p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m}')}{p(\mathbf{m})p(\mathbf{d}_{\mathrm{obs}}|\mathbf{m})}. \tag{11}$$

If the transition is accepted, then the next step of the Markov chain becomes $\mathbf{m}'$ and we set $\mathbf{m} = \mathbf{m}'$. If the transition is rejected, then the Markov chain remains at the current model, **m**.

[16] 5. Repeat steps 2–5.

[17] After a certain number of iterations, known as the burn-in period, the algorithm described above will reach equilibrium, beyond which steps in the Markov chain will effectively represent samples that have been drawn from the Bayesian posterior distribution for the model parameters. In other words, the MCMC procedure will become independent of the starting values for these parameters and samples generated after this point can then be collected and analyzed in terms of their posterior statistics. In this regard, the first critical step in the practical application of MCMC methods is to assess the convergence behavior. Although there is no way to judge the length of the burn-in period prior to running an MCMC inversion, a number of practical strategies exist to evaluate whether a Markov chain has reached this point. *Gelman and Rubin* [1992] define a metric based on the relative size of between-chain to within-chain variation for a series of parallel-running chains, in order to see whether the MCMC procedure has converged. *Hassan et al.* [2009] also evaluate burn-in using the results of a number of chains running in parallel, but do so graphically. That is, they plot the value of each model parameter versus iteration number for the different inversions in order to assess when the inversions have reached a similar equilibrium. We adopt the latter approach for our work. Once burn-in has been reached and the samples up to that point have been discarded, the next critical step in an MCMC inversion is to determine the number of iterations that are required to generate enough samples to properly characterize the Bayesian posterior distribution. This will depend on the dimension of the model parameter search space, the information content of the prior, and the degree of correlation between adjacent samples in the Markov chain which results from the bounded nature of the proposal density function. Indeed, if the posterior chain exhibits a long autocorrelation lag, then a greater number of iterations are required after burn-in to produce a sufficient number of independent posterior samples. In our work, we chose to use the methodology proposed by *Gilks et al.* [1996] to evaluate an appropriate length for the posterior chain after burn-in. This involves comparing ergodic averages for the various model parameters between a number of parallel chains and stopping when the averages are in good agreement. (For further details, please see *Gilks et al.* [1996].)

3. Tests on Synthetic Data

[18] In the following, we consider two synthetic examples in order to investigate the effect of including different degrees of prior information regarding the VGM parameters on the posterior distributions obtained from the stochastic inversion of ZOP crosshole GPR traveltime data. We first consider the simple case of a homogeneous subsurface medium. We then explore a more complex and realistic scenario where the subsurface is defined by four different layers. For consistency with the field study presented later, the probed domain for both synthetic examples is 6 m wide $\times$ 17 m deep and is assumed to consist of sandy fluvial deposits.

[19] To simulate the "observed" GPR traveltime data corresponding to each example, we first used the SS_INFIL code with a vertical cell discretization of 0.01 m. Constant flux boundary conditions were asssumed along the upper model surface with an infiltration rate of 3.5×10^{-4} m d^{-1}, whereas the lower model boundary at 17 m depth was defined to be the groundwater table. To convert the resulting steady state water content profile to GPR velocity, we used equations (5) and (6) with the following parameter values: $\kappa_s = 5$, $\kappa_w = 81$, $\kappa_a = 1$, and $\theta_s = 0.32$. Note that, for consistency with the field example presented in section 4.1, we assume a priori to know the value of the saturated water content, θ_s. It can be shown that this choice has little influence on the corresponding GPR data as the CRIM equation is highly insensitive to this parameter. For the subsequent simulation of GPR traveltimes with the PRONTO eikonal equation solver, the grid discretization was set to 0.1 m and the transmitter and receiver antennas were assumed to be moved at 0.25 m increments along the left- and right-hand model edges from 3 to 14 m depth. Finally,

Gaussian random noise with a standard deviation equal to 2% of the mean traveltime value was added to the data prior to the MCMC analysis.

[20] We examine the effect of three different priors for the VGM parameters in each synthetic example. In doing so, we progress from uninformed to moderately informed and then to highly informed conditions. The priors are the following:

[21] 1. The VGM parameters are assigned uniform prior distributions having bounds: $\theta_r \in U(0, 0.1)$, $\alpha \in U(0, 25)\ \mathrm{m}^{-1}$, $n \in U(1.5, 4)$, and $\log_{10}(K_s\ [\mathrm{m\ d}^{-1}]) \in U(-2, 2)$. These ranges are broad and consistent with previous work [e.g., *Binley and Beven*, 2003; *Cassiani and Binley*, 2005; *Looms et al.*, 2008a]. Most importantly, such uniform priors allow us to assess the information with regard to the VGM parameters contributed by the GPR traveltime data alone.

[22] 2. The VGM parameters are assumed to follow empirical prior distributions as determined by *Carsel and Parrish* [1988] for sandy soil based on the analysis of over 250 samples. In this case, the prior sampling domain is significantly restricted when compared to using the uniform priors. (Please see *Carsel and Parrish* [1988] for details on how to generate random sets of VGM parameters corresponding to these distributions.) For this particular prior, we ignore the joint relationships provided in their paper and thus assume that the parameters are statistically uncorrelated.

[23] 3. The VGM parameters are prescribed the same marginal prior statistics for sandy soil from *Carsel and Parrish* [1988] as given above, but we further restrict the sampling domain by now including the full parameter correlation information. Again, please see their paper for the relevant details.

[24] Figure 1 shows a series of bivariate scatterplots corresponding to the second and third prior scenarios described above, each of which was created from 50,000 randomly drawn VGM parameter sets. As mentioned, in going from the noninformative uniform distribution to the informative but uncorrelated distribution and finally to the informative correlated distribution, the prior samples become more likely to be drawn from a restricted region of the model space. In an inversion, using carefully chosen refined priors based on existing databases or field measurements can help to avoid sampling VGM parameter sets that may fit the observed data but are not consistent with the environment and/or material of interest.

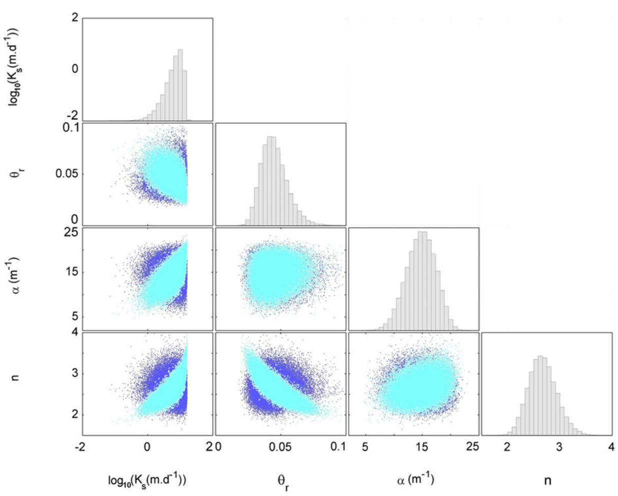

Figure 1. Bivariate scatter plots and marginal histograms for the refined prior VGM parameter distributions as defined by *Carsel and Parrish* [1988] for sand. The dark and light blue points indicate the uncorrelated and correlated distributions, respectively.

3.1. Homogeneous Medium

[25] In our first example, we assume that the subsurface is homogeneous and characterized by the following VGM parameter values, which are typical for a sandy soil: $\theta_r = 0.04$, $\alpha = 13\ \mathrm{m}^{-1}$, $n = 3$, and $\log_{10}(K_s\ [\mathrm{m\ d}^{-1}]) = 0.88$. These values were drawn randomly from the correlated distribution of *Carsel and Parrish* [1988] shown in light blue in Figure 1. To estimate the posterior distribution of the VGM parameters given the corresponding "observed" synthetic data, we ran the MCMC inversion procedure considering our three different prior scenarios. The residual uncertainty term, σ_r, in equation (9) was set equal to 2% of the mean GPR traveltime value in each case. Three independent parallel Markov chains with randomly chosen starting points were also initiated for each scenario such that burn-in could be graphically assessed. A sufficient burn-in period was determined to be 5000 iterations for the uniform prior and 1000 iterations for the uncorrelated and correlated refined priors, up until which all samples were discarded before analysis. After burn-in, 50,000 iterations for the uniform prior and 30,000 iterations for both refined priors were then determined to adequately sample the posterior parameter space. Running the inversions for the uniform and refined priors took 36 and 15 h on a standard 3.16 GHz desktop computer, respectively.

[26] Figures 2 and 3 show the marginal prior and posterior histograms of the different VGM parameters for the uniform and refined prior scenarios, respectively. The true values of the parameters are also shown. A simple but generally effective way to further assess the quality of the fit between each suite of prior and posterior realizations and the corresponding true values is to compute the root-mean-square error (RMSE), given by

$$\varepsilon_{\mathrm{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - m_{\mathrm{true}})^2}, \tag{12}$$

where N is the number of realizations, m_{true} is the true model parameter value, and m_i is the i-th model parameter value from the suite of realizations. Table 1 shows the corresponding RMSE values. Notice in Figure 2 that the posterior distributions for all VGM parameters are noticeably refined compared to the prior uniform distributions. Indeed, through the MCMC inversion of the GPR data, we obtain significant reductions in uncertainty for α, $\log_{10}(K_s)$, and n, and a smaller but still important reduction in uncertainty for θ_r. This is also observed in Table 1, where we have a significant decrease in the RMSE values in going from the uniform prior to the corresponding posterior distributions. These results clearly show that the GPR data alone contain valuable information about the subsurface VGM parameters.

[27] In Figure 3, where we consider the refined prior distributions based on the work of *Carsel and Parrish* [1988],

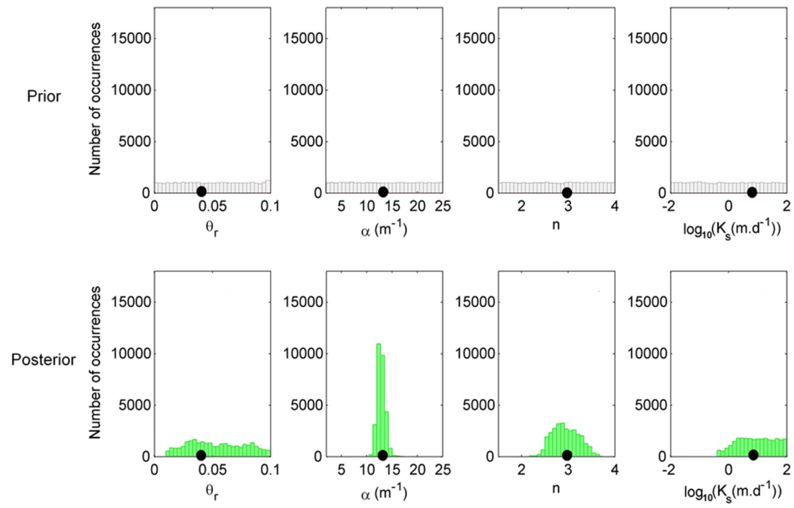

Figure 2. Uniform prior and corresponding posterior marginal histograms for the VGM parameters in the homogeneous medium example. The circles indicate the true parameter values.

we see a further reduction in uncertainty in the posterior distributions for all VGM parameters compared to the case of the uniform prior. Here the posterior distributions are more narrow than those in Figure 2, and the corresponding RMSE values in Table 1 are significantly smaller. This demonstrates how choosing a refined, yet realistic, prior distribution on the basis of soil property measurements and/or databases can be of great benefit to resolving subsurface VGM parameters in a stochastic inversion. We must emphasize, however, that there is still an important contribution of the GPR data in this case, even with the use of the refined prior distributions (i.e., the priors are not providing us with everything). Indeed, comparing the prior and posterior histograms in Figure 3, we see that the incorporation of the GPR data allows us to markedly reduce the uncertainty in the VGM parameter estimates in most cases. This again is reflected in Table 1, where we see a significant and consistent reduction in the posterior RMSE values compared to the prior distributions.

[28] Finally, we observe in Figure 3 a distinct difference between the posterior distributions obtained using the uncorrelated refined prior distribution and those obtained assuming that the VGM parameters are correlated. Assuming correlation between the parameters results in smaller posterior uncertainties for $\log_{10}(K_s)$ and θ_r, but larger posterior uncertainties for α and n, as compared to the uncorrelated case. This demonstrates that including more detailed information in the prior distribution (e.g., parameter correlation) does not always lead to better resolved posterior VGM parameter estimates, as one might intuitively assume. This in

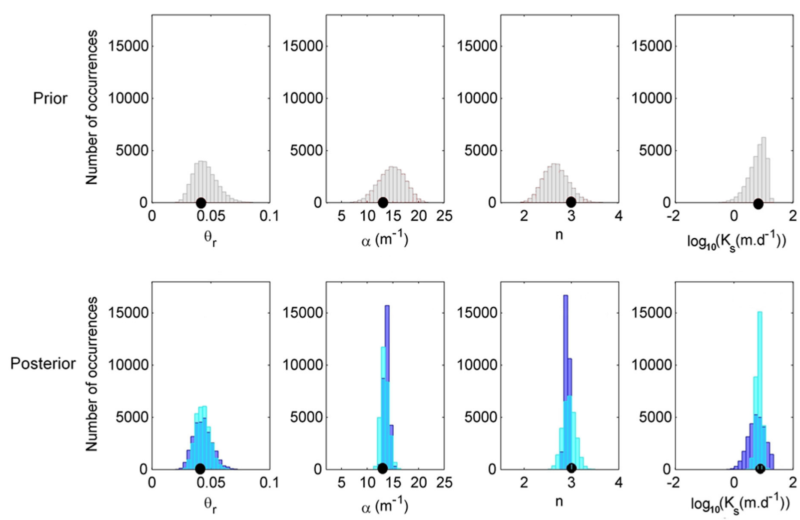

Figure 3. Refined prior and corresponding posterior marginal histograms for the VGM parameters in the homogeneous medium example. The dark and light blue colors indicate the results obtained for the uncorrelated and the correlated priors, respectively. Areas where the histograms overlap are shown as a medium blue color. The circles indicate the true parameter values.

Table 1. RMSE Values Quantifying the Difference Between the True VGM Parameters and the Corresponding Prior and Posterior Distributions for the Homogeneous Medium Synthetic Example

	Uniform Prior				Refined Prior (Uncorrelated)				Refined Prior (Correlated)			
Distribution	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])
Prior	0.030	6.66	0.76	1.45	0.012	3.08	0.42	0.32	0.012	3.08	0.42	0.32
Posterior	0.027	0.95	0.29	0.77	0.008	0.91	0.10	0.31	0.007	0.88	0.14	0.10

turn implies that using an inadequately informed prior, for example, by not accounting for parameter correlation when it exists, may actually lead to overconfidence with regard to the posterior estimates of certain model parameters.

3.2. Layered Medium

[29] We next consider a more realistic synthetic example involving a stratified medium consisting of four layers. The layers are defined as follows:

[30] 0 to 6 m: coarse sand (layer 1)

[31] 6 to 8 m: fine sand (layer 2)

[32] 8 to 12 m: coarse sand (layer 3)

[33] 12 to 14 m: fine sand (layer 4)

[34] The coarse sand material in layers 1 and 3 is characterized by the following VGM parameters: $\theta_r = 0.04$, $\alpha = 13$ m^{-1}, $n = 3$, and $\log_{10}(K_s$ [m d^{-1}]$) = 0.88$. The VGM parameters for the fine sand material in layers 2 and 4 are: $\theta_r = 0.05$, $\alpha = 10$ m^{-1}, $n = 2.2$, and $\log_{10}(K_s$ [m d^{-1}]$) = 0.50$. As in the previous example, we used these parameters and layer definitions to first simulate the corresponding "observed" ZOP GPR traveltime data under steady state conditions. These data were then stochastically inverted to estimate the VGM parameters in each layer assuming the same three prior scenarios outlined above. Please note that the same prior distributions were used for all of the layers in the MCMC inversion procedure. Three parallel independent MCMC inversions were again run for each type of prior, with σ_r in equation (9) set equal to 2% of the mean GPR traveltime value. A sufficient burn-in period was determined to be 15,000 iterations for the uniform prior and 5000 iterations for both refined priors. In this example, 100,000 iterations and 50,000 iterations were determined to adequately sample the posterior parameter space, respectively. The uniform prior inversion took 3 d on the same 3.16 GHz desktop computer, whereas each refined prior inversion took 1.5 d.

[35] Figures 4 and 5 show the marginal prior and posterior histograms of the different VGM parameters in each layer for the uniform and refined prior scenarios, respectively. Table 2 lists the corresponding RMSE values between these distributions and the true parameter values. In agreement with our findings in the case of a homogeneous model (Figure 2, Table 1), Figure 4 demonstrates that even for more complex and realistic subsurface structures, the GPR traveltime data alone contain a significant amount of information regarding the VGM parameters. Indeed, compared to the uniform prior distribution, the posterior distributions for the shape parameter, α, show significant refinement, those for $\log_{10}(K_s)$ and n show notable refinement, and those for θ_r are refined in layers 1 and 3. These findings are also reflected in the RMSE values. For the two refined prior scenarios considered in Figure 5, the inferred posterior distributions are generally better resolved for the correlated case than for the uncorrelated case. The most substantial improvement in terms of the resolution of these posterior distributions compared to the uniform prior scenario is seen for α. However, the posterior distributions of the other VGM parameters also benefit from the inclusion of a more informative prior distribution. In general, the results in Figures 4 and 5 illustrate that GPR traveltime data, in combination with well informed yet realistic prior distributions, should contain sufficient information to allow for an effective MCMC-based estimation of the VGM parameters in realistic layered media.

4. Application to Field Data

4.1. Field Site and Data

[36] We now apply the methodologies investigated above to observed data from a pertinent field site located near Eggborough, Yorkshire, UK. This field site was developed to study the unsaturated hydraulic characteristics of the Sherwood sandstone, which represents a major regional aquifer and consists of predominantly horizontally layered, medium- to fine-grained Triassic sandy fluvial deposits with occasional intercalations of siltstone [*Binley et al.*, 2002a, 2005; *Cassiani and Binley*, 2005; *West et al.*, 2003]. A total of 11 boreholes have been drilled at the site with six of these (R1, R2, R3, R4, A, and B) designed for GPR studies and the other five (C, E1, E2, E3, and E4) created for resistivity measurements (Figure 6). All of the boreholes were geophysically logged using electromagnetic induction and natural gamma sondes. One borehole (R5) was completely cored to 20 m depth, and the retrieved rock samples were analyzed in the laboratory to determine their hydraulic and electrical material properties [*Binley et al.*, 2005]. *West et al.* [2003] also carried out dielectric property measurements on other samples from the site which allowed them to estimate the dielectric constant of the sediment grains and the saturated water content.

[37] We seek to estimate subsurface VGM parameters from ZOP crosshole GPR data collected at the Eggborough site between 3 August 1999 and 20 February 2001. These data were acquired between boreholes R3 and R4, which are roughly 6 m apart and 16 m deep, using a Sensors and Software PulseEKKO GPR system with 50 MHz antennae and a vertical increment of 0.25 m in the observation interval from 3 to 14 m depth. During the monitoring period, the GPR measurements were repeated 15 times at regular intervals of ~6 wk. Interestingly, analysis of the data indicated that the temporal variations of subsurface water content in response to recharge events and subsequent

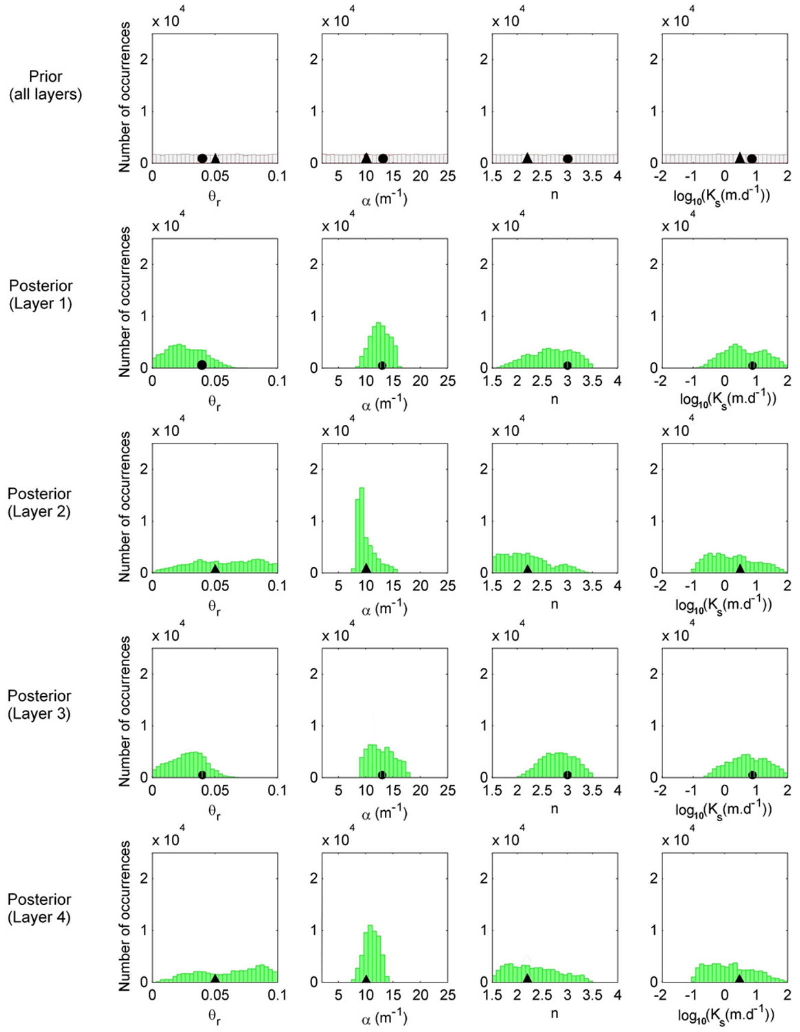

Figure 4. Uniform prior and corresponding posterior marginal histograms for the VGM parameters in each layer for the layered medium example. The circles indicate the true parameter values for Layers 1 and 3 whereas the triangles indicate the true parameter values for Layers 2 and 4.

moisture redistribution were minimal [*Cassiani and Binley*, 2005]. This in turn demonstrated that a steady state characterization approach is justified for the site. Consequently, we averaged the ZOP traveltimes measured in each GPR survey to obtain a mean traveltime versus depth curve, which was used in our analysis. Figure 7d shows this curve along with bounds corresponding to one standard deviation. In the numerical modeling codes to link subsurface VGM parameters to GPR traveltimes for the MCMC inversion, we used the same boundary conditions and discretization intervals that were used in both synthetic examples. Based on the extensive laboratory measurements on the available core material from borehole R5 [*Binley et al.*, 2005] and the additional measurements made by *West et al.* [2003], values of $\kappa_s = 5$ and $\theta_s = 0.32$ were assumed for the relative dielectric permittivity of the dry rock matrix and the saturated water content, respectively.

4.2. Model Parameterization

[38] A critical part of any inversion procedure is choosing the model parameterization. In the synthetic examples presented earlier, we had the advantage of knowing a priori the nature of the subsurface heterogeneity, which made this choice straightforward. For example, in the second example, we knew that the subsurface contained four layers and we also knew the locations of all layer boundaries. For field data, this kind of information must be gained from the data themselves as well as from complementary sources of information, and the model parameterization process is correspondingly uncertain and potentially nonunique. In this regard, Figure 7 shows some of the core and log data that have been collected at the Eggborough site. Figures 7a and 7b plot the median particle size and clay fraction versus depth in borehole R5, which were determined from core samples, whereas Figure 7c shows the natural gamma log acquired in neighboring borehole E4. These plots, along with the GPR traveltime curve in Figure 7d, suggest that the subsurface is by no means homogeneous, and comparison with other log data has indicated that a 1-D layered model is a reasonable approximation at the site. *Cassiani and Binley* [2005] performed a geostatistical analysis of the logs they collected in each borehole to obtain a stochastic

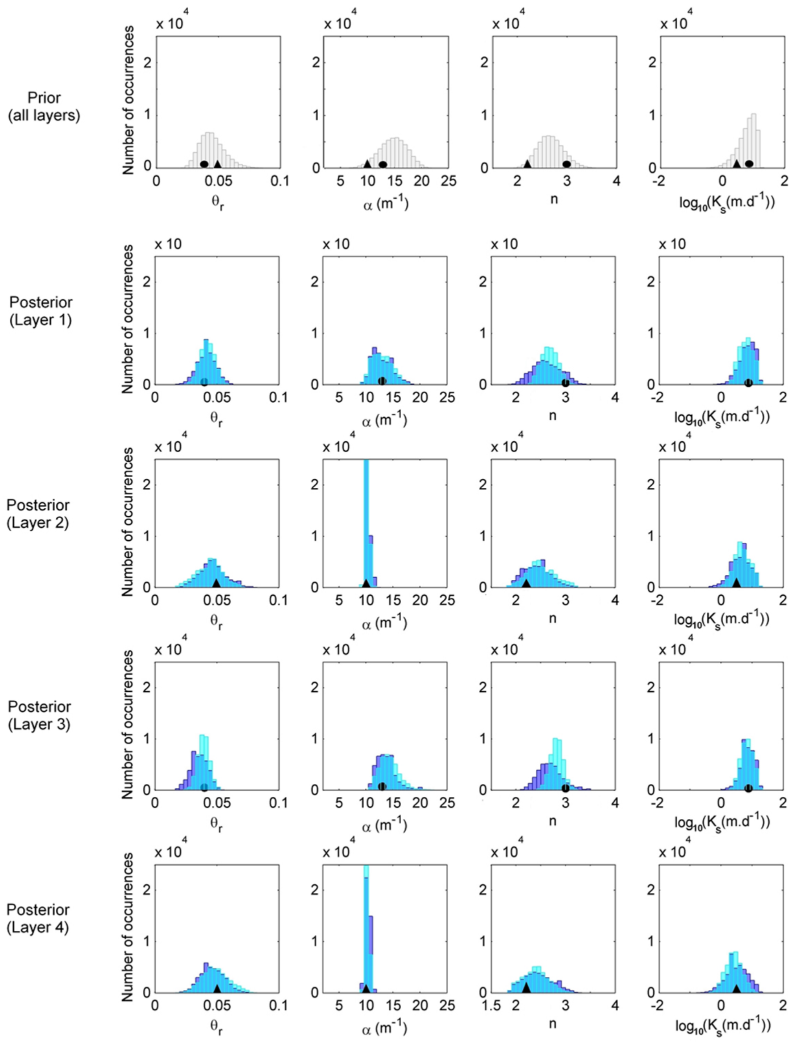

Figure 5. Refined prior and corresponding posterior marginal histograms for the VGM parameters in each layer for the layered medium example. The dark and light blue colors indicate the results obtained for the uncorrelated and the correlated priors, respectively. Areas where the histograms overlap are shown as a medium blue color. The circles indicate the true parameter values for Layers 1 and 3 whereas the triangles indicate the true parameter values for Layers 2 and 4.

description of this layering and, following *Binley et al.* [2004], they considered the vertical lithological profile as a binary system. In a subsequent study, *Linde et al.* [2006] carried out a joint inversion of crosshole electrical resistivity and multioffset GPR traveltime data collected at the site. Their results indicated that the subsurface petrophysical properties can be grouped into four distinct clusters and hence that, from a geophysical point of view, the site is adequately characterized by a four-unit model. The boundaries of these four units are shown in Figure 7 and they define

Table 2. RMSE Values Quantifying the Difference Between the True VGM Parameters and the Corresponding Prior and Posterior Distributions for the Layered Medium Synthetic Example

		Uniform Prior				Refined Prior (Uncorrelated)				Refined Prior (Correlated)			
Layer	Distribution	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])	θ_r	α (m^{-1})	n	Log$_{10}$ (K_s [m d^{-1}])
1	Prior	0.030	6.66	0.76	1.45	0.012	3.08	0.42	0.32	0.012	3.08	0.42	0.32
	Posterior	0.021	1.75	0.51	0.72	0.008	1.82	0.41	0.26	0.006	1.76	0.36	0.23
2	Prior	0.029	7.50	0.90	1.26	0.011	5.39	0.54	0.40	0.011	5.39	0.54	0.40
	Posterior	0.025	1.78	0.45	0.74	0.011	0.45	0.35	0.34	0.011	0.37	0.41	0.32
3	Prior	0.030	6.66	0.76	1.45	0.012	3.08	0.42	0.32	0.012	3.08	0.42	0.32
	Posterior	0.018	2.19	0.38	0.62	0.008	1.91	0.41	0.22	0.005	2.27	0.22	0.20
4	Prior	0.029	7.50	0.90	1.26	0.011	5.39	0.54	0.40	0.011	5.39	0.54	0.40
	Posterior	0.026	1.74	0.49	0.79	0.011	0.48	0.41	0.35	0.011	0.47	0.32	0.28

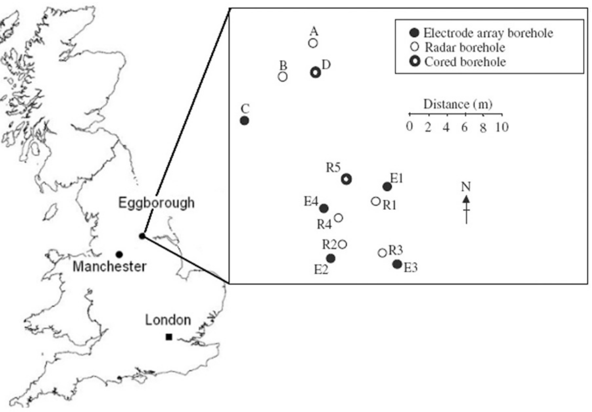

Figure 6. Location of the Eggborough field site and installed boreholes. Adapted from *Cassiani and Binley* [2005], with permission from Elsevier.

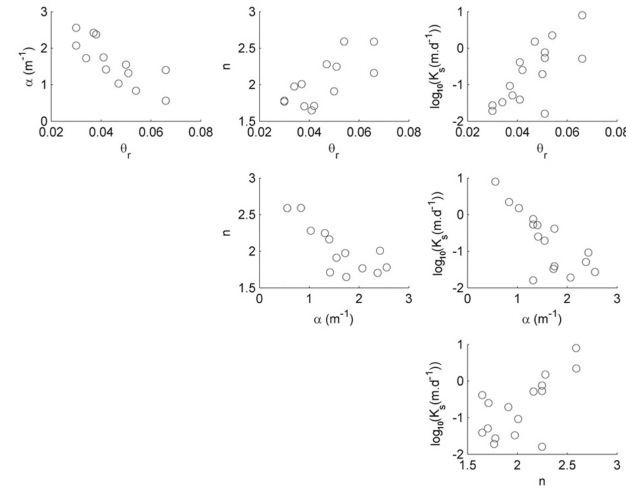

Figure 8. Bivariate scatter plots of the VGM parameters obtained from measurements on core samples from the Eggborough field site by *Binley et al.* [2005].

the layered parameterization that we use for our stochastic inversion. In other words, we invert the traveltime data in Figure 7d for the VGM model parameters in each of the different layers assuming that these units are internally homogeneous. (For further details regarding how this subsurface zonation was obtained, please see *Linde et al.* [2006].)

4.3. Prior Parameter Distributions

[39] For our field investigation at the Eggborough site, we define the prior distributions for the VGM model parameters in each lithological unit based on the hydraulic measurements made by *Binley et al.* [2005] on core samples retrieved from borehole R5. Mercury injection capillary pressure (MICP) tests were carried out on 19 of these samples and provided information with regard to the pore size distribution, which was then used to estimate an unsaturated moisture release curve and to infer the VGM parameters. Figure 8 shows a series of bivariate scatterplots of the results obtained. As in the synthetic examples, we again consider three prior distributions. First, we consider a noninformative uniform prior where the VGM parameter ranges are the following: $\theta_r \in U(0, 0.1)$, $\alpha \in U(0, 3)$ m^{-1}, $n \in U(1.5, 3)$, and $\log_{10}(K_s\ [\mathrm{m\ d}^{-1}]) \in U(-2, 1.5)$. Next, we consider a more refined but uncorrelated prior. The field samples allowed us to prescribe the following distributions: $\theta_r \in N(0.045, 0.011)$, $\alpha \in N(1.58, 0.58)$ m^{-1}, $n \in N(2.04, 0.31)$, and $\log_{10}(K_s\ [\mathrm{m\ d}^{-1}]) \in N(-0.74, 0.85)$. Finally, we consider that the VGM parameters are correlated. Although the limited number of samples implies that any statistical analysis is prone to significant uncertainties, the scatterplots in Figure 8 provide a relatively clear indication that moderate to strong correlation between the various VGM parameters exists, which is consistent with other previous work. On the basis of the available data, we inferred the following correlation matrix:

$$C = \begin{pmatrix} & \log_{10}(K_s) & \theta_r & \alpha & n \\ \log_{10}(K_s) & 1 & 0.752 & -0.764 & 0.682 \\ \theta_r & & 1 & -0.857 & 0.7834 \\ \alpha & & & 1 & -0.812 \\ n & & & & 1 \end{pmatrix}. \tag{13}$$

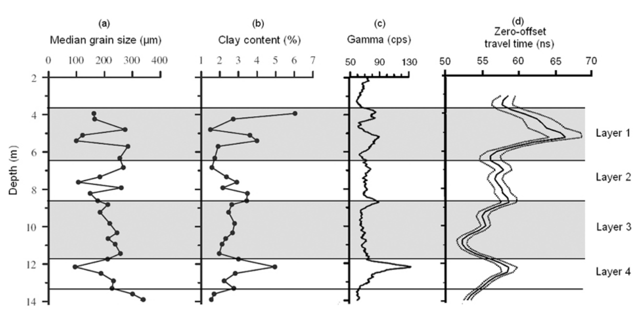

Figure 7. Example borehole data that were collected at the Eggborough field site. (a and b) median grain size and clay content measured on core samples from borehole R5, respectively, (c) gamma log taken in borehole R4, (d) mean and standard deviation of the ZOP GPR traveltimes measured between boreholes R3 and R4. Also shown is the layered subsurface parameterization used in this study. Adapted from *Linde et al.* [2006].

Please note that the same priors are assumed for each lithological unit in each case. Also note that field samples were necessary in this example to derive our informative prior distributions because the previously used distributions of *Carsel and Parrish* [1988] are for unconsolidated sandy soils, and not sandstone.

4.4. MCMC Inversion

[40] We performed the stochastic inversion of the GPR traveltime data in Figure 7d to estimate the posterior distributions of α, θ_r, n, and $\log_{10}(K_s)$ in each layer under the three different prior scenarios outlined above. As in the synthetic examples, three independent MCMC inversions were run in parallel for each case. A sufficient burn-in period was determined to be 20,000 iterations for the uniform

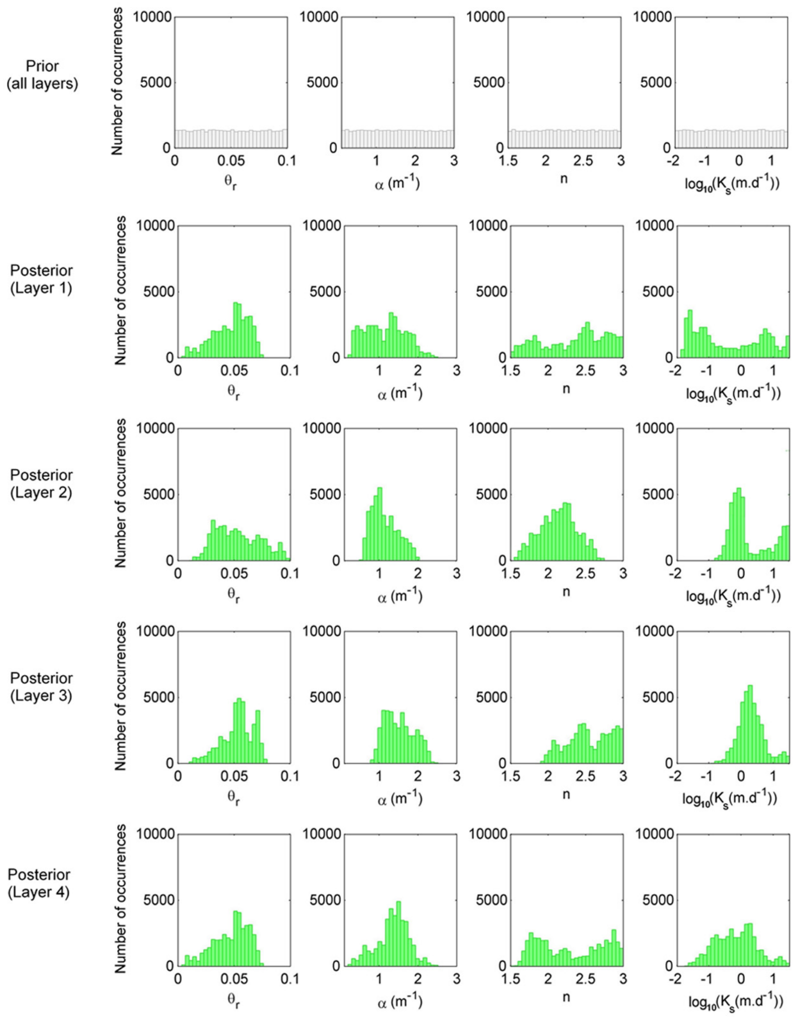

Figure 9. Uniform prior and corresponding posterior marginal histograms for the VGM parameters in each layer for the field example.

prior distribution and 5000 iterations for the refined prior distributions. A total of 100,000 iterations for the uniform prior and 50,000 iterations for the refined priors were found to adequately sample the posterior parameter space. The corresponding inversions took ~3 and 1.5 d, respectively. For this field example, σ_r in equation (9) was estimated based on the multiple repeated traveltime measurements that were used to obtain Figure 4d.

[41] Figures 9 and 10 show the prior and posterior histograms of the different VGM parameters in each layer for the uniform and refined prior scenarios, respectively. In agreement with previous work at the Eggborough site by *Binley and Beven* [2003] and *Cassiani and Binley* [2005], and in accordance with our synthetic examples, the results for the uniform prior indicate that an unequivocal estimation of the VGM parameters on the basis of the GPR traveltime data alone is not possible (Figure 9). However, it is clear that the GPR data contain important information regarding these parameters as we observe a significant reduction of uncertainty in going from the uniform prior to posterior distributions. In Figure 10, we again see that the use of more informative priors leads to significantly better resolved posterior distributions in each layer. Clearly, the refined priors provide much information, but it is important to note that they do not provide everything. Indeed, in most cases, the incorporation of the GPR data results in noticeably improved posterior parameter estimates as compared to the refined priors. Finally, note in Figure 10 that accounting for correlation between the VGM parameters in the definition of the prior results in more refined posterior distributions as compared to the case where the parameters are assumed to be uncorrelated.

5. Conclusions

[42] We have investigated how different degrees of prior information regarding vadose zone hydraulic parameters can affect the posterior distributions obtained for these parameters from the stochastic inversion of geophysical data. This was done in the context of estimating VGM parameters from ZOP crosshole GPR data assuming 1-D flow, steady state conditions, and knowledge of the pertinent boundary conditions. A number of previous studies found that the stochastic inversion of ZOP crosshole GPR data allowed for some refinement in VGM parameter estimates in homogeneous and layered media. However, the findings

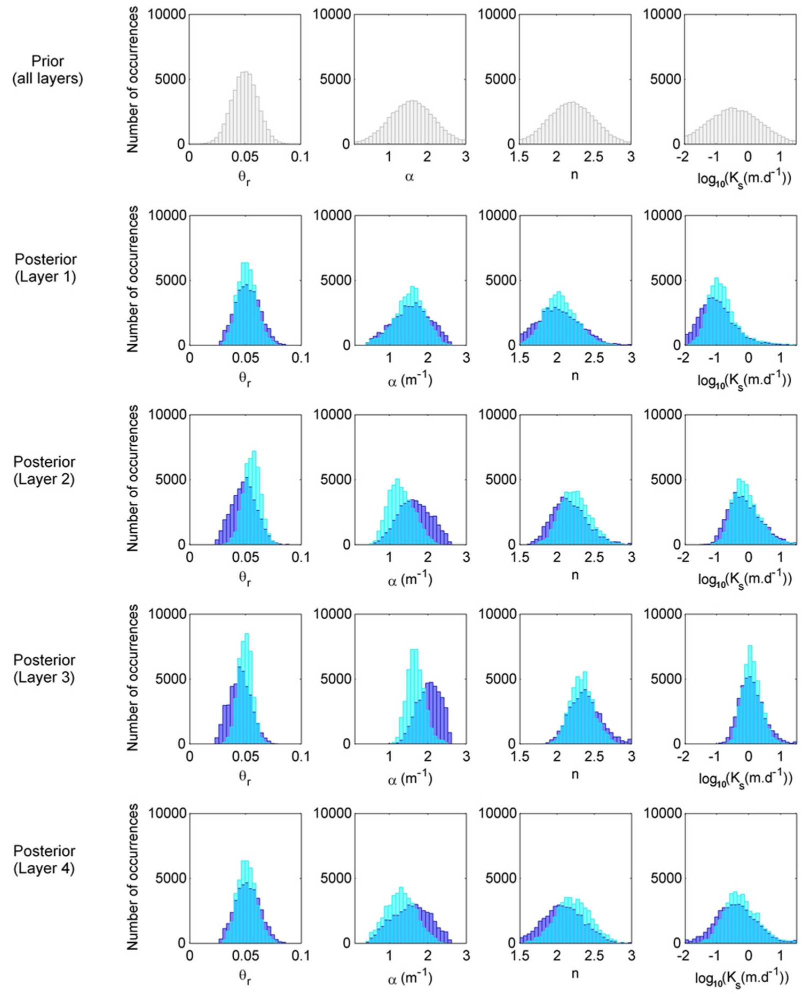

Figure 10. Refined prior and corresponding posterior marginal histograms for the VGM parameters in each layer for the field example. The dark and light blue colors indicate the results obtained for the uncorrelated and the correlated cases, respectively. Areas where the histograms overlap are shown as a medium blue color.

were inconclusive with regard to the true utility of the GPR measurements due to a lack of systematic testing on synthetic examples and consideration of realistic prior knowledge. Moreover, the previous work only involved the pseudo-Bayesian GLUE inversion methodology, and thus lacked the stochastic rigor of a formal Bayesian approach.

[43] In both our synthetic and field tests, we saw through the use of noninformative uniform prior distributions that the ZOP GPR data contain important information regarding the VGM parameters, even in the case where these parameters were estimated in a heterogeneous layered medium. However, not all parameters were well resolved from the GPR data alone, which is consistent with previous work. As a result, unequivocal estimation of the parameters using only the GPR data, at least in the steady state case, is not possible. Through the use of more realistic and informative priors based on soil property databases, we saw a significant improvement in the posterior distributions of the VGM parameters. This is encouraging and demonstrates that the combination of geophysical data with realistic prior knowledge may allow for the effective stochastic estimation of vadose zone hydraulic properties and their corresponding uncertainties. This in turn may be used to predict future hydrological behavior (e.g., groundwater recharge, contaminant transport) within a framework of uncertainty, the results of which might be used for risk assessment. Note, however, that as in any Bayesian study, much care needs to be taken when defining refined priors such that they are neither overly specific nor too conservative and that they adequately reflect the available a priori information. In our analysis of uncorrelated versus correlated prior distributions, for example, we saw that not accounting for parameter correlation when it exists may actually lead to overconfidence regarding a particular model parameter; that is, a seemingly conservative choice of prior may in fact yield overly optimistic posterior results.

[44] It is important to remember that, in any field study, a number of critical issues exist that are not present when dealing with synthetic models. In the case of the Eggborough data set, we made the assumption of 1-D steady state conditions and a four-layer subsurface parameterization based on existing site information. Choices of this kind are clearly approximations that will affect the results obtained

from an inversion. Future work should involve the study of errors in such aspects of the modelization process and their effects on the posterior statistics obtained. Another important topic for future work is the consideration of dynamic data. Dynamic geophysical measurements that are sensitive to changes in hydrological state variables, such as water content or salinity, are more strongly tied to subsurface hydrological properties and may thus offer much improved potential for stochastic parameter estimation.

[45] **Acknowledgments.** This research was supported by a grant from the Swiss National Science Foundation. Andrew Binley is grateful to Alan Green for supporting his period of leave at ETH Zurich, which allowed him to contribute to this work. We also thank Associate Editor Jasper Vrugt and four anonymous reviewers for their constructive comments which helped to greatly improve the quality of the final manuscript.

References

Aldridge, D. F., and D. W. Oldenburg (1993), Two-dimensional tomographic inversion with finite-difference traveltimes, *J. Seism. Explor.*, *2*, 257–274.

Annan, A. P. (2005), GPR methods for hydrogeological studies, in *Hydrogeophysics*, edited by Y. Rubin and S. S. Hubbard, 532 pp., Springer, Netherlands.

Beven, K. J., and A. M. Binley (1992), The future of distributed models—Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*(3), 279–298.

Binley, A. M., and K. J. Beven (2003), Vadose zone flow model uncertainty as conditioned on geophysical data, *Ground Water*, *41*(2), 119–127, doi:10.1111/j.1745-6584.2003.tb02576.x.

Binley, A. M., P. Winship, R. T. Middleton, M. Pokar, and J. West (2001), High-resolution characterization of vadose zone dynamics using cross-borehole radar, *Water Resour. Res.*, *37*(11), 2639–2652, doi:10.1029/2000WR000089.

Binley, A. M., G. Cassiani, R. T. Middleton, and P. Winship (2002a), Vadose zone flow model parameterisation using cross-borehole radar and resistivity imaging, *J. Hydrol.*, *267*(3–4), 147–159, doi:10.1016/S0022-1694(02)00146-4.

Binley, A. M., P. Winship, L. J. West, M. Pokar, and R. Middleton (2002b), Seasonal variation of moisture content in unsaturated sandstone inferred from borehole radar and resistivity profiles, *J. Hydrol.*, *267*(3–4), 160–172.

Binley, A. M., G. Cassiani, and P. Winship (2004), Characterization of heterogeneity in unsaturated sandstone using borehole logs and cross-borehole tomography, in *Aquifer Characterization* edited by J. S. Bridge and D. W. Hyndman, pp. 129–138, SEPM, Soc. Sedimentary Geology, Tulsa, Okla.

Binley, A. M., L. D. Slater, M. Fukes, and G. Cassiani (2005), Relationship between spectral induced polarization and hydraulic properties of saturated and unsaturated sandstone, *Water Resour. Res.*, *41*(12), W12417, doi:10.1029/2005WR004202.

Carsel, R. F., and R. S. Parrish (1988), Developing joint probability-distributions of soil-water retention characteristics, *Water Resour. Res.*, *24*(5), 755–769, doi:10.1029/WR024i005p00755.

Cassiani, G., and A. M. Binley (2005), Modeling unsaturated flow in a layered formation under quasi-steady state conditions using geophysical data constraints, *Adv. Water Resour.*, *28*(5), 467–477, doi:10.1016/j.advwatres.2004.12.007.

de Rooij, G. H., R. T. A. Kasteel, A. Papritz, and H. Fluhler (2004), Joint distributions of the unsaturated soil hydraulic parameters and their effect on other variates, *Vadose Zone J.*, *3*(4), 947–955.

Finsterle, S., and M. B. Kowalsky (2008), Joint hydrological-geophysical inversion for soil structure identification, *Vadose Zone J.*, *7*(1), 287–293, doi:10.2136/vzj2006.0078.

Flores, A. N., D. Entekhabi, and R. L. Bras (2010), Reproducibility of soil moisture ensembles when representing soil parameter uncertainty using a Latin Hypercube based approach with correlation control, *Water Resour. Res.*, *46*(4), W04506, doi:10.1029/2009WR008155.

Freer, J., K. J. Beven, and B. Ambroise (1996), Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, *Water Resour. Res.*, *32*(7), 2161–2173, doi:10.1029/95WR03723.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, *7*(4), 457–472.

Gilks, W. R., S. Richardson, and D. Spiegelhalter (1996), Introducing Markov chain Monte Carlo, in *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson, and D. Spiegelhalter, 512 pp., Chapman & Hall/CRC, London. U. K.

Greaves, R. J., D. P. Lesmes, J. M. Lee, and M. N. Toksoz (1996), Velocity variations and water content estimated from multi-offset, ground-penetrating radar, *Geophysics*, *61*(3), 683–695.

Hassan, A. E., H. M. Bekhit, and J. B. Chapman (2009), Using Markov chain Monte Carlo to quantify parameter uncertainty and its effect on predictions of a groundwater flow model, *Environ. Modell. Softw.*, *24*(6), 749–763, doi:10.1016/j.envsoft.2008.11.002.

Haygarth, P. M., and S. C. Jarvis (Ed.) (2002), *Agriculture, Hydrology and Water Quality*, 499 pp., CABI, Wallingford Oxon, U. K.

Hinnell, A. C., T. P. A. Ferre, J. A. Vrugt, J. A. Huisman, S. Moysey, J. Rings, and M. B. Kowalsky (2010), Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion, *Water Resour. Res.*, *46*, W00D40, doi:10.1029/2008WR007060.

Huisman, J. A., J. Snepvangers, W. Bouten, and G. B. M. Heuvelink (2002), Mapping spatial variation in surface soil water content: Comparison of ground-penetrating radar and time domain reflectometry, *J. Hydrol.*, *269*(3–4), 194–207, doi:10.1016/S0022-1694(02)00239-1.

Irving, J., and K. Singha (2010), Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities, *Water Resour. Res.*, *46*, W11514, doi:10.1029/2009WR008340.

Linde, N., A. Binley, A. Tryggvason, L. B. Pedersen, and A. Revil (2006), Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data, *Water Resour. Res.*, *42*(12), W12404, doi:10.1029/2006WR005131.

Looms, M. C., A. M. Binley, K. H. Jensen, L. Nielsen, and T. M. Hansen (2008a), Identifying unsaturated hydraulic parameters using an integrated data fusion approach on cross-borehole geophysical data, *Vadose Zone J.*, *7*(1), 238–248, doi:10.2136/vzj2007.0087.

Looms, M. C., K. H. Jensen, A. M. Binley, and L. Nielsen (2008b), Monitoring unsaturated flow and transport using cross-borehole geophysical methods, *Vadose Zone J.*, *7*(1), 227–237, doi:10.2136/vzj2006.0129.

Mallants, D., D. Jacques, M. Vanclooster, J. Diels, and J. Feyen (1996), A stochastic approach to simulate water flow in a macroporous soil, *Geoderma*, *70*(2–4), 299–324.

Mertens, J., H. Madsen, L. Feyen, D. Jacques, and J. Feyen (2004), Including prior information in the estimation of effective soil parameters in unsaturated zone modelling, *J. Hydrol.*, *294*(4), 251–269, doi:10.1016/j.jhydrol.2004.02.011.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, *21*(6), 1087–1092.

Morse, B. S., G. Pohll, J. Huntington, and R. Rodriguez Castillo (2003), Stochastic capture zone analysis of an arsenic-contaminated well using the generalized likelihood uncertainty estimator (GLUE) methodology, *Water Resour. Res.*, *39*(6), 1151, doi:10.1029/2002WR001470.

Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*(B7), 12431–12447, doi:10.1029/94jb03097.

Mualem, Y. (1976), A new model for predicting hydraulic conductivity of unsaturated porous-media, *Water Resour. Res.*, *12*(3), 513–522, doi:10.1029/WR012i003p00513.

Parkin, G. W., D. Redman, P. von Bertoldi, and Z. Zhang (2000), Measurement of soil water content below a wastewater trench using ground-penetrating radar, *Water Resour. Res.*, *36*(8), 2147–2154, doi:10.1029/2000WR900129.

Rockhold, M. L., C. S. Simmons, and M. J. Fayer (1997), An analytical solution technique for one-dimensional, steady vertical water flow in layered soils, *Water Resour. Res.*, *33*(4), 897–902, doi:10.1029/96WR03746.

Roth, K., R. Schulin, H. Fluhler, and W. Attinger (1990), Calibration of time domain reflectometry for water measurement using a composite dielectric approach, *Water Resour. Res.*, *26*(10), 2267–2273, doi:10.1029/WR026i010p02267.

Schaap, M. G., F. J. Leij, and M. T. van Genuchten (2001), ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, *J. Hydrol.*, *251*(3–4), 163–176.

Scharnagl, B., J. A. Vrugt, H. Vereecken, and M. Herbst (2011), Bayesian inverse modelling of in situ soil water dynamics: Using prior information about the soil hydraulic properties, *Hydrol. Earth Syst. Sci. Discuss.*, *8*(1), 2019–2063, doi:10.5194/hessd-8-2019-2011.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, 342 pp., SIAM, Philadelphia, Pa.
van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*(5), 892–898.
Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, B. A. Robinson, J. M. Hyman, and D. Higdon (2009a), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlin. Sci. Num*, *10*(3), 273–290.
Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson (2009b), Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?, *Stoch. Env. Res. Risk A*, *23*(7), 1011–1026, doi:10.1007/s00477-008-0274-y.
West, L. J., K. Handley, Y. Huang, and M. Pokar (2003), Radar frequency dielectric dispersion in sandstone: Implications for determination of moisture and clay content, *Water Resour. Res.*, *39*(2), 1026, doi:10.1029/2001WR000923.
Winship, P., A. M. Binley, and D. Gomez (2006), Flow and transport in the unsaturated Sherwood Sandstone: characterization using cross-borehole geophysical methods, *Geol. Soc., London, Special Publications*, *263*(1), 219–231, doi:10.1144/gsl.sp.2006.263.01.12.

A. Binley, Lancaster Environment Centre, Lancaster University, LA1 4YQ Lancaster, UK.
K. Holliger, J. Irving, and M. Scholer, Institute of Geophysics, University of Lausanne, CH-1015 Lausanne, Switzerland. (james.irving@unil.ch)