DOE/SC0002390 Technical Report

Entropy-Bayesian Inversion of Time-Lapse Tomographic GPR data for Monitoring Dielectric Permittivity and Soil Moisture Variations

Summary

In this study, we evaluate the possibility of monitoring soil moisture variation using tomographic ground penetrating radar travel time data through Bayesian inversion, which is integrated with entropy memory function and pilot point concepts, as well as efficient sampling approaches.

It is critical to accurately estimate soil moisture content and variations in vadose zone studies. Many studies have illustrated the promise and value of GPR tomographic data for estimating soil moisture and associated changes, however, challenges still exist in the inversion of GPR tomographic data in a manner that quantifies input and predictive uncertainty, incorporates multiple data types, handles non-uniqueness and nonlinearity, and honors time-lapse tomograms collected in a series. To address these challenges, we develop a minimum relative entropy (MRE)-Bayesian based inverse modeling framework that non-subjectively defines prior probabilities, incorporates information from multiple sources, and quantifies uncertainty. The framework enables us to estimate dielectric permittivity at pilot point locations distributed within the tomogram, as well as the spatial correlation range. In the inversion framework, MRE is first used to derive prior probability distribution functions (pdfs) of dielectric permittivity based on prior information obtained from a straight-ray GPR inversion. The probability distributions are then sampled using a Quasi-Monte Carlo (QMC) approach, and the

sample sets provide inputs to a sequential Gaussian simulation (SGSim) algorithm that constructs a highly resolved permittivity/velocity field for evaluation with a curved-ray GPR forward model. The likelihood functions are computed as a function of misfits, and posterior pdfs are constructed using a Gaussian kernel. Inversion of subsequent time-lapse datasets combines the Bayesian estimates from the previous inversion (as a memory function) with new data. The memory function and pilot point design takes advantage of the spatial-temporal correlation of the state variables. We first apply the inversion framework to a static synthetic example and then to a time-lapse GPR tomographic dataset collected during a dynamic experiment conducted at the Hanford Site in Richland, WA. We demonstrate that the MRE-Bayesian inversion enables us to merge various data types, quantify uncertainty, evaluate nonlinear models, and produce more detailed and better resolved estimates than straight-ray based inversion; therefore, it has the potential to improve estimates of inter-wellbore dielectric permittivity and soil moisture content and to monitor their temporal dynamics more accurately.

1. Introduction

The ability to monitor soil moisture in the vadose zone is crucial to meeting many societal needs. Agriculture, for instance, relies on current soil moisture data to guide irrigation practices and to predict the potential yield of crops (Shaxson & Barber, 2003). Weather forecasts rely on accurate estimates of soil moisture as input to climate models (Ni-Meister, Walker, & Houser, 2005). Soil moisture information is also a key component in predicting natural disasters, such as flooding and slope failure (Tohari, Nishigaki, & Komatsu, 2007). Soil moisture is also a key parameter in determining unsaturated hydraulic conductivity, which in turn affects solute mobility and contaminant fate and transport within the vadose zone (Murdoch, 2000).

Several techniques exist for monitoring soil moisture at a variety of spatial and temporal scales, including space-borne sensors, air-borne sensors, wireless sensor networks, and ground-based sensors (Vereecken et al., 2008). Within the class of ground-based sensors, ground penetrating radar (GPR) offers the unique advantage over other methods like time domain reflectometry (TDR) or capacitance probes by having intermediate spatial support volumes (centimeters to meters) (Vereecken et al., 2008) and temporal resolution (e.g., sub-daily), and being minimally invasive. Although GPR does not measure soil moisture directly, signal travel times recorded by GPR are dominantly regulated by variations in dielectric permittivity, a parameter controlled by the presence of moisture and related through a petrophysical model (Lunt, Hubbard, & Rubin, 2005). Particularly, tomographic GPR has emerged as a solution to long-term monitoring of spatial distribution of soil moisture within the vadose zone (Binley, Cassiani, Middleton, & Winship, 2002; Hubbard, Peterson, et al., 1997) and as a means of deriving other spatially heterogeneous soil physical properties, such as permeability (Binley et al., 2002; J. Chen, Hubbard, & Rubin, 2001; Dubreuil-Boisclair, Gloaguen, Marcotte, & Giroux, 2011; Hubbard, Rubin, & Majer, 1997; Hubbard et al., 2001; M. B. Kowalsky et al., 2005; Michael B. Kowalsky, Finsterle, & Rubin, 2004) and porosity (Clement & Barrash, 2006), through time-lapse, and/or joint inversion (Binley et al., 2002; M. B. Kowalsky et al., 2005; Michael B. Kowalsky et al., 2004).

The inversion process converting GPR traveltime tomography data to physical properties of interest is an area of continued investigation. Inverted images from tomograms are subject to great uncertainty due to non-uniqueness of solutions and variable spatial resolutions. Sine tomographic inversion problems tend to be underdetermined, and might have an infinite number of possible solutions, regularization (typically smoothing) can be applied to stabilize solutions, but the corresponding tomography inversion results tend to exaggerate the size and underestimate the magnitude of anomalies in the subsurface, and under-represent the true correlation structure (Day-Lewis, 2004). The optimum spatial resolution achievable by GPR is limited by survey geometry, measurement error, measurement physics, and the parameterization and/or regularization employed in the inversion (Day-Lewis, 2004). The variation in resolution within a tomogram makes pixel-specific inference of petrophysical properties uncertain.

Time-lapse GPR traveltime tomography datasets are a useful means for observing dynamic systems. Given that changes in soil moisture are hysteretic, it is sensible to capitalize on time-dependent correlation between time-lapse datasets. One useful technique is difference inversion (i.e., (LaBrecque & Yang, 2001)) wherein regularization is applied between time-lapse datasets. Such an approach is useful for highlighting changes to a system, but requires a reliable inversion of a background dataset. (Kim, Yi, Park, & Kim, 2009) and (Karaoulis et al., 2011) note that the subsurface can be treated as a space-time model, wherein the system is changing continuously on both space and time axes. The objective function therefore penalizes for changes in space-time.

It is widely recognized that quantifying uncertainty is a crucial aspect of inverting geophysical data (Cassiani, Strobbia, & Gallotti, 2004; J. Chen et al., 2001; Rubin & Hubbard, 2005; Yeh & Simunek, 2002), which cannot be done using deterministic approaches (such as least-square optimization). It has become increasingly common to invert data within a stochastic framework wherein parameters are treated in a probabilistic manner, although the approaches still lacks the ability to combine multiple data types. One solution is the Bayesian updating framework, where parameters retain their probabilistic structure throughout inversion and can be updated quantitatively with the addition of new information (for examples, see (J. Chen & Rubin, 2003; X. Chen, Rubin, Ma, & Baldocchi, 2008; Copty, Rubin, & Mavk, 1993; Dubreuil-Boisclair et al., 2011; Hou, Rubin, Hoversten, Vasco, & Chen, 2006; Hou & Rubin, 2005; Hubbard et al., 2001; M. B. Kowalsky et al., 2005; Michael B. Kowalsky et al., 2004; Lehikoinen, Huttunen, Finsterle, Kowalsky, & Kaipio, 2010) .

Bayesian inversion approaches offer a systematic way to quantitatively describe uncertainty and provide parameter estimates that can be conditionally updated when new data becomes available. However, the flexibility of Bayesian techniques can also raise concerns. First, there is the risk of subjectivity when defining prior distributions since the inversion hinge upon the choice of prior distributions (for an excellent discussion on subjectivity and options for choosing prior distributions, see (Berger, 2006)). Another critical component in Bayesian inversion is the weighting scheme. Suppose we have greater confidence in our data than we do in our prior information, we would seek to bias assign higher weights to the new data, for the situation with better instrumentation but lack of information to define the priors. Finally, the freedom of modeling within a Bayesian framework may result in excessive computational demand, in that increasing model complexity and number of parameters will result in more simulations and longer CPU time per simulation.

(Hou & Rubin, 2005) employ minimum relative entropy (MRE) within a Bayesian framework as a means of non-subjectively selecting prior probability distribution functions (pdfs) in a study estimating soil moisture from transient surface GPR, meteorological measurements, and TDR data. Briefly, MRE is a means of automatically updating parameter pdfs that minimizes the entropy between the existing pdfs and new constraints placed on that pdfs. The MRE-derived pdfs allow maximum flexibility in the inversion, without sacrificing loss of prior information from the constraints

(M. B. Kowalsky et al., 2005; Michael B. Kowalsky et al., 2004) used a Bayesian scheme for inversion of ground penetrating radar (GPR) tomography data to study infiltration processes occurring within the vadose zone. They developed a maximum a posteriori (MAP) Bayesian approach within a pilot point framework to infer dielectric permittivity and (hydraulic) permeability fields from transient borehole GPR and borehole saturation measurements. The method of pilot points uses only select points in tandem with a spatial covariance model to simulate parameter estimates over a refined spatial grid, and is an effective way of capturing effects of heterogeneity using a minimum number of parameters.

In this study, we adopt the MRE-Bayesian approach developed by (Hou & Rubin, 2005) and apply it to estimate spatial dielectric permittivity fields in the vadose zone from transient tomographic GPR data. Similarly to (Michael B. Kowalsky et al., 2004), we also incorporate the method of pilot points and sequential Gaussian simulation (SGSIM) to produce our dielectric permittivity fields. However, in our work, we treat each pilot point as a distinct Bayesian parameter defined by a unique MRE prior pdf and subsequent memory functions. Our target parameters for inversion are dielectric permittivity at several locations throughout the study area and the spatial correlation range. We demonstrate our method by applying it to a static synthetic example and a time-lapse field example, the latter collected at the Hanford nuclear site in Richland, WA.

2. GPR Background

Tomographic GPR is a borehole-based geophysical technique. It involves transmitting an electromagnetic (EM) pulse from a source in one borehole and recording the arrival of EM energy at a receiver position in a separate borehole. The source and receiver vertical locations are varied to collect a suite of data of signal arrival times and magnitude for various source-receiver pairs.

Inversion of the first arrival times of the EM energy is used to estimate the velocity and the dielectric permittivity (ϵ) distribution between the boreholes. For convenience, the relative dielectric permittivity (ϵ_r) is used, which is simply the dielectric permittivity normalized by the speed of light in a vacuum ($c = 0.3\ \mathrm{m/ns}$). At the high frequencies typically used for GPR (~50-1000MHz) and in low-loss environments (non-magnetic, low electrical conductivity), the dielectric constant (ϵ_r) of a soil can be related to EM velocity (v) by:

$$\epsilon_r \approx \left(\frac{c}{v}\right)^2, \qquad (1)$$

(Davis & Annan, 1989). It is clear that to estimate dielectric permittivity, information regarding the EM velocity structure between the boreholes is needed. Since we are interested in the spatial variation in EM velocity, we discretize the study area into n grid blocks with velocities $v_1 \ldots v_n$, and relate these velocities to travel time data through a forward model (**G**) that describes the propagation path (distance) travelled by the EM energy. The solution takes the form of an inversion problem:

$$\mathbf{G}(\mathbf{v}) = \mathbf{t}, \qquad (2)$$

where **v** is a vector of the velocities of the grid blocks, and **t** represents the vector of measured travel times.

GPR forward models generally fall into two categories: (1) full-waveform methods (for example, (Casper & Kung, 1996; Michael B. Kowalsky, Dietrich, Teutsch, & Rubin, 2001; Vasco, Peterson, & Lee, 1997)) and (2) ray-based methods (i.e., (Cai & Mcmechan, 1995; Peterson, 2001; Zhang, Rector, & Hoversten, 2005)). Full waveform methods compute solutions to Maxwell's equations, and offer thorough information on EM energy propagation. In a ray-based inversion approach, the total travel time from a source to a receiver is discretely represented as:

$$t = \sum_{i=1}^{n} \frac{l_i}{v_i}, \qquad (3)$$

where l_i is the distance travelled by the ray through the ith grid block. If straight-ray paths are assumed, the forward model matrix **G** is a ray-path matrix containing the lengths of the rays through each grid block for each measurement and the inversion problem becomes linear. Typically, solutions to model parameters (the grid blocks) are in terms of slowness, and can be computed via iterative reconstruction techniques (i.e., the algebraic reconstruction techniques (ART) (Peterson, Paulsson, & McEvilly, 1985; Peterson, 2001) or the simultaneous iterative reconstruction technique (SIRT) (Dines & Lytle, 1979). The straight ray approach is limited to situations with little variation in EM velocity (e.g., less than 10%

EM velocity variation (Day-Lewis, 2005), so curved-ray methods that account for ray-bending phenomena may be used in situations where significant heterogeneity is expected. The curved-ray method allows a GPR ray to follow physically realistic trajectories; the traveltimes along direct, reflected, refracted, and transmitted raypaths are computed and compared until the first arrival is found (Zhang et al., 2005)

3. Methodology

The complete inversion process consists of a pre-inversion step, an initial 'static' inversion, and subsequent time-lapse inversions when time-lapse data is available (Figure 1). The process is described in the following sections.

3.1. Pre-inversion

The pre-inversion process is designed to provide initial estimates for expected values (i.e., prior means) of the model parameters. This is achieved through a linear inversion of the data.

A ray-path matrix (**G**) is constructed analytically assuming straight ray paths, and the SIRT method is used to produce an initial estimate of the dielectric permittivity distribution within the study area. These estimates of dielectric permittivity represent weighted averages of the true permittivity (Day-Lewis, 2005).

Given initial estimates of the dielectric permittivity parameters, we are able to construct a variogram, from which we can estimate the correlation range, the distance beyond which dielectric permittivity is no longer correlated (Hengl, 2009). Due to the smearing effects of the SIRT inversion, we observe that this range likely overestimates the true value (Day-Lewis, 2004; Hubbard, Rubin, &

Majer, 1999). In our inversion study, we include the correlation range as an unknown and assign an upper bound corresponding to the diagonal length of the field, and a lower bound of 0, indicative of no correlation between the points. We do not assign a standard deviation to the range parameter.

3.2. Inversion: static case

In order to reduce the number of unknowns and make the inverse problem less ill-posed, we choose to invert dielectric permittivity at selected point locations and the associated spatial correlation range.

The prior pdfs for these parameters are derived using MRE method with given prior information, which includes physical lower and upper bounds, and statistical moments (mean, standard deviation). We use the SIRT inverted values as the prior mean values, and assign loose upper and lower bounds (e.g., ±4 from the mean values). These are loose bounds, considering that in general they represent a minimum 0.17 range in moisture content, according to Topp's petrophysical model (Topp, Davis, & Annan, 1980). Given that we have a reasonable confidence in the SIRT inversion, we opt to also impose a standard deviation of 2.

To construct the initial pdfs of dielectric constant values, we use the MRE algorithm originally developed by (Woodbury, 2004). Given a mean (μ), standard deviation (σ), and upper (U) and lower bounds (L), MRE selects a truncated normal distribution to represent the pilot point parameter pdfs:

$$p(x) = \begin{cases} \frac{1}{C} e^{-\gamma(x+\frac{\beta}{2\gamma})^2} & \text{if } L \leq x \leq U \\ 0 & \text{otherwise} \end{cases}, \qquad (4)$$

where *p(x)* represents the probability of any given value of *x* occurring. The coefficient β, γ, and C are determined by the constraints μ, σ, L and U. For a complete derivation of these solutions, see (Hou &

Rubin, 2005). Lacking a standard deviation, MRE selects a truncated exponential distribution for the parameter of correlation range:

$$p(x) = \begin{cases} \frac{\beta e^{-\beta x}}{e^{-\beta L} - e^{-\beta U}} & \text{if } 0 \le L \le x \le U \\ 0 & \text{otherwise} \end{cases}, \qquad (5)$$

Similarly, the coefficient β is determined by the constraints μ, L and U.

Now that each parameter is represented by a pdf, we are able to sample values from these distributions for numerical evaluations. Sampling technique is an important choice in our inversion, as the success of a numerical approach hinges upon evaluating all possibilities defined by the model space. Given a large number of dimensions (n_{par} = 49, in our case), systematic sampling techniques such as by Simpson's rule are not sufficient (Tarantola, 2005). Therefore, we turn to quasi-Monte Carlo (QMC) techniques, which incorporate deterministic sequences to guarantee good dispersion between sample points (Caflisch, 2008).

QMC requires a choice regarding input of a low-discrepancy sequence. It is widely acknowledged that Sobol sequences (Sobol, 1967) perform well for problems of greater than six dimensions, and avoid degradation effects observed in many other low-discrepancy sequences (Atanassov et al., 2010; Sobol & Shukhman, 2007; Wang & Sloan, 2008).

Therefore, we use a Sobol sequence to initialize our QMC sampling procedure, and produce adequate number of sample sets (q ~ 10^6) for numerical evaluation. Each sample set (**m**) contains discrete values for each of the dielectric constant pilot points, and a value for the range parameter. For each of the q sample sets, we estimate the prior mean, standard deviation, and correlation range. To generate a dielectric permittivity field, we discretize the study field into elements of equal area. Using the GSLIB sequential Gaussian simulation package SGSIM (Deutsch & Journel, 1997), we simulate spatially correlated dielectric permittivity values throughout the study region. These values are converted to EM velocity at each pixel through Equation (1).

Modeling of GPR first arrival times is accomplished through the method developed by (Zhang et al., 2005), which uses a curved-ray (eikonal solver) to compute EM first-arrival travel times (**T_{cal}**) from the modeled dielectric permittivity field.

The modeled first arrival times are compared to the observational data (**T_{obs}**) to compute residuals for each source-receiver pair. As each parameter set is processed, the standard deviation associated with each source-receiver pair (**r_{std}**) is updated by Welford's method (Welford, 1962).

After all parameter sets have been compared with the data, the residuals are normalized by their standard deviations. The normalized residuals are assumed to following standard normal distribution; therefore, the likelihood functions take the form of normal distributions. The likelihood function for each parameter set $\boldsymbol{m_j}$ is given by:

$$L(\boldsymbol{m}_j) = \exp\left[\sum_{i=1}^{n_{meas}} - \frac{\left[\frac{\left(Tobs_i - Tcal_i(\boldsymbol{m}_j)\right)}{r_{stdi}}\right]^2}{2}\right]. \qquad (6)$$

The posterior pdf is the product of the likelihood function in Equation 6 and the prior pdfs in Equations 4 or 5, expressed as weights (**w**) for each parameter set. Given the weights, the posterior mean (μ_{post}) and standard deviation (σ_{post}) of each parameter set are updated. We divide the full range for each parameter into 100 bins (n_{bins} = 100) and the posterior probability at each bin is calculated using the Gaussian kernel:

$$p(x_{bin})_{post} \sim \sum_{i=1}^{q} w_i \frac{1}{2\pi\sigma_{post}^2} e^{-\frac{(x_{bin}-x_i)^2}{2\sigma_{post}^2}}, \qquad (7)$$

3.3. Inversion: time-lapse data

Our purpose now is to monitor temporal variations in the dielectric permittivity field, but in order to take the advantage of temporal continuity, it is desired to use inversion results from previous stages as the new prior pdfs (actually intermediate estimates), which can be called ‘memory functions’ (Equation 7). To be able to capture temporal variation, it is reasonable to reset the bounds according to the updated means and standard deviations. We choose to reset the new bounds to the mean values ±2 standard deviations, the latter being a measure of spread of the distribution. Further, we limit the extension of these bounds beyond a global maximum and minimum, ranging from 3 to 25 for the pilot point dielectric constant values and 0.25 m to 15 m for the range parameter. This approach allows distributions to shift with time, which is reasonable given the time-lapse nature of the experiment; the distributions also serve as memory functions to honor spatial-temporal correlations of the state variables. We also have updated SIRT inversion results available. To combine these two input data on the model parameters, we transform the SIRT estimates into MRE pdfs but assign bounds corresponding to their global maxima and minima to guarantee compatibility with the Bayesian estimates. The SIRT-derived pdfs are considered to be independent from the corresponding Bayesian posterior pdfs, and are combined following the multiplicative law of joint probability. The joint probabilities of the SIRT-derived pdfs and the Bayesian estimates provide intermediate pdfs (memory functions) that act as the starting point (new prior probability) for the subsequent inversion.

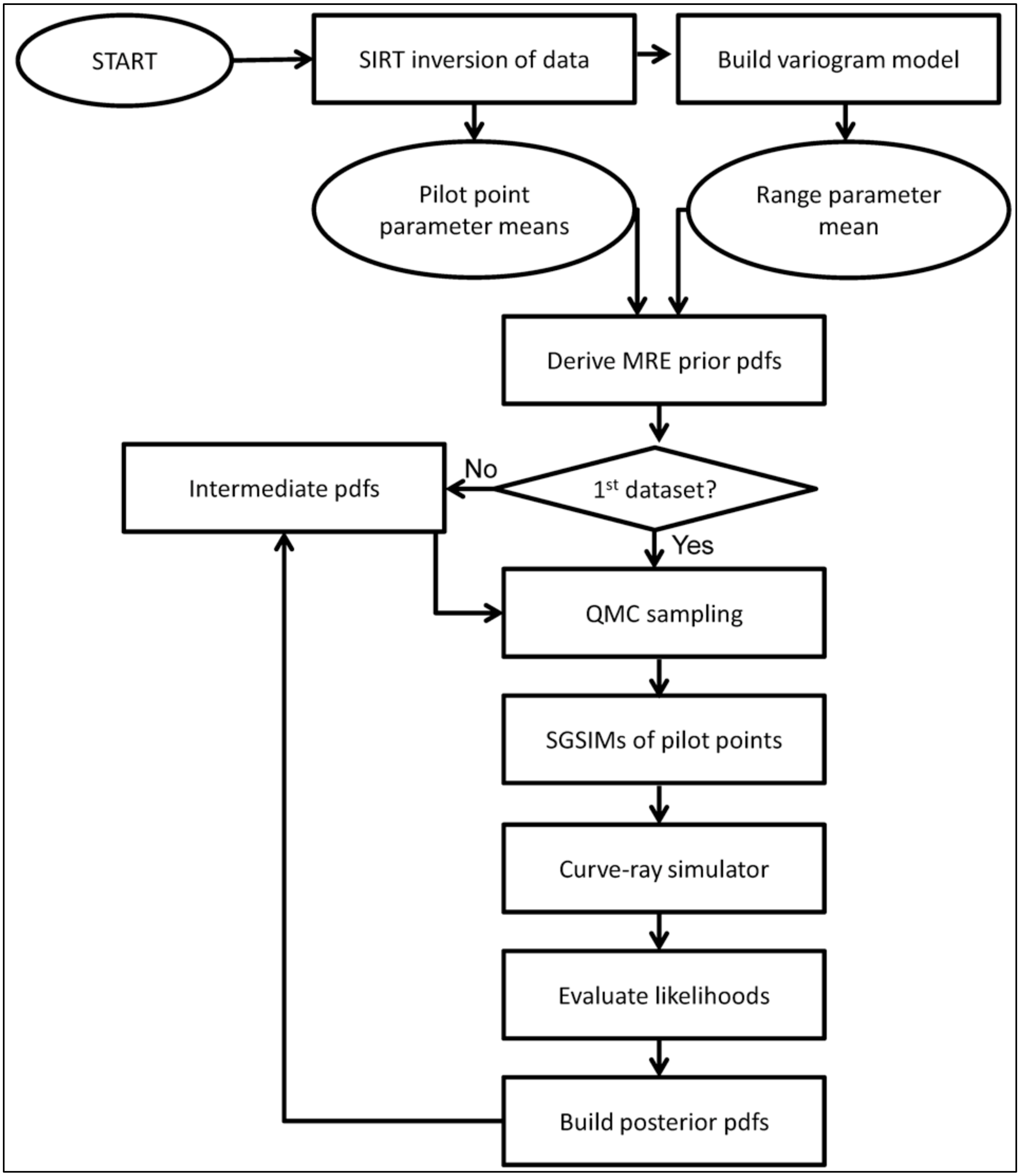

Figure 1: Flowchart showing the major steps in the MRE-Bayesian inversion process used in this report.

4. Synthetic Experiment

4.1. Forward modeling

The study area is 4 m width by 15 m depth (see figure 2). We consider 30 equally spaced source locations, and for each source location, data is collected at 30 evenly spaced receiver locations for a total of 900 observations (n_{obs} = 900), stored in a matrix $\mathbf{T_{obs}}$. We added two percent random noise to the traveltime observations as measurement errors.

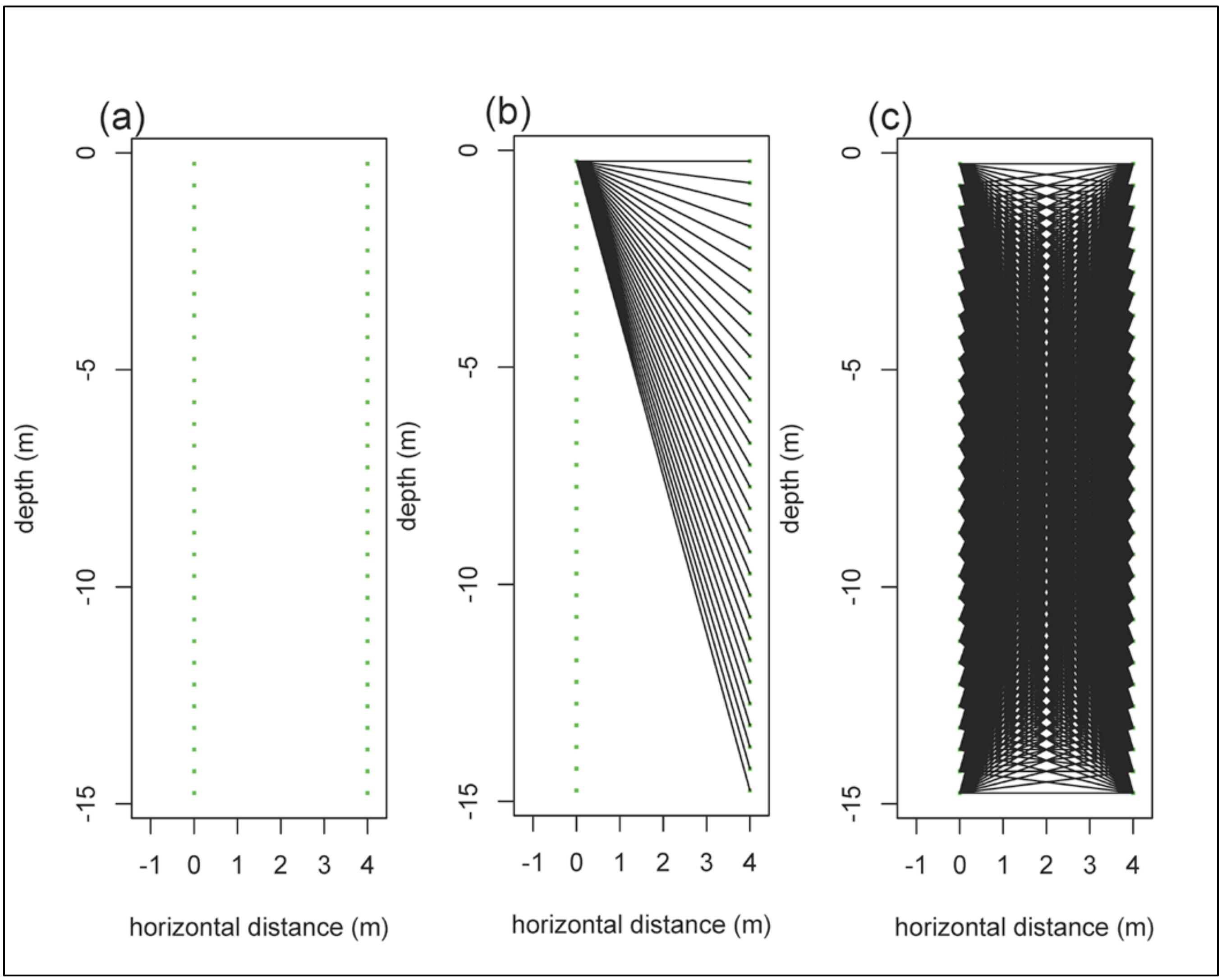

Figure 2: Geometry for the synthetic experiment. A tomographic GPR survey takes place between two parallel borehole with a separation of 4 m and depth of 15 m (a). For a given source location, each receiver location records the first arrival of EM energy (b). Tomography provides 2-dimensional coverage of the study area (c).

4.2. Results of the synthetic study

The prior MRE pdfs and Bayesian posterior pdfs for dielectric permittivity pilot points are plotted in Figures 3a and b, respectively, in correspondence to their relative location within the tomogram. When compared to their prior pdfs, most of the distributions have shifted more or less away from the prior values. This is not surprising since the SIRT-derived means are averaged representations of the true values of dielectric constant at the pilot point locations. Thus, our inversion approach provides improved resolution of extreme values and local anomalies, and thus makes the best use of available data.

To demonstrate this, Figure 4 shows the true dielectric constant values side by side with SIRT-derived estimates and a Gaussian simulation of the posterior mean values. We observe that the SIRT inversion provides a ‘smeared’ image of the study area, and that although higher and lower zones are generally captured, their values consistently under- or over-estimate the corresponding true values. Thus, these zones are not well resolved, and the range of values gives a false impression of the true field variations. The Gaussian simulated image, on the other hand, mimics the heterogeneity of the true image quite well, and the range of observed values more closely reflects reality than the SIRT image. However, we note that the edges of the field are much less reliable. This is consistent with the sparse sampling domain of tomographic GPR near the edges of the field.

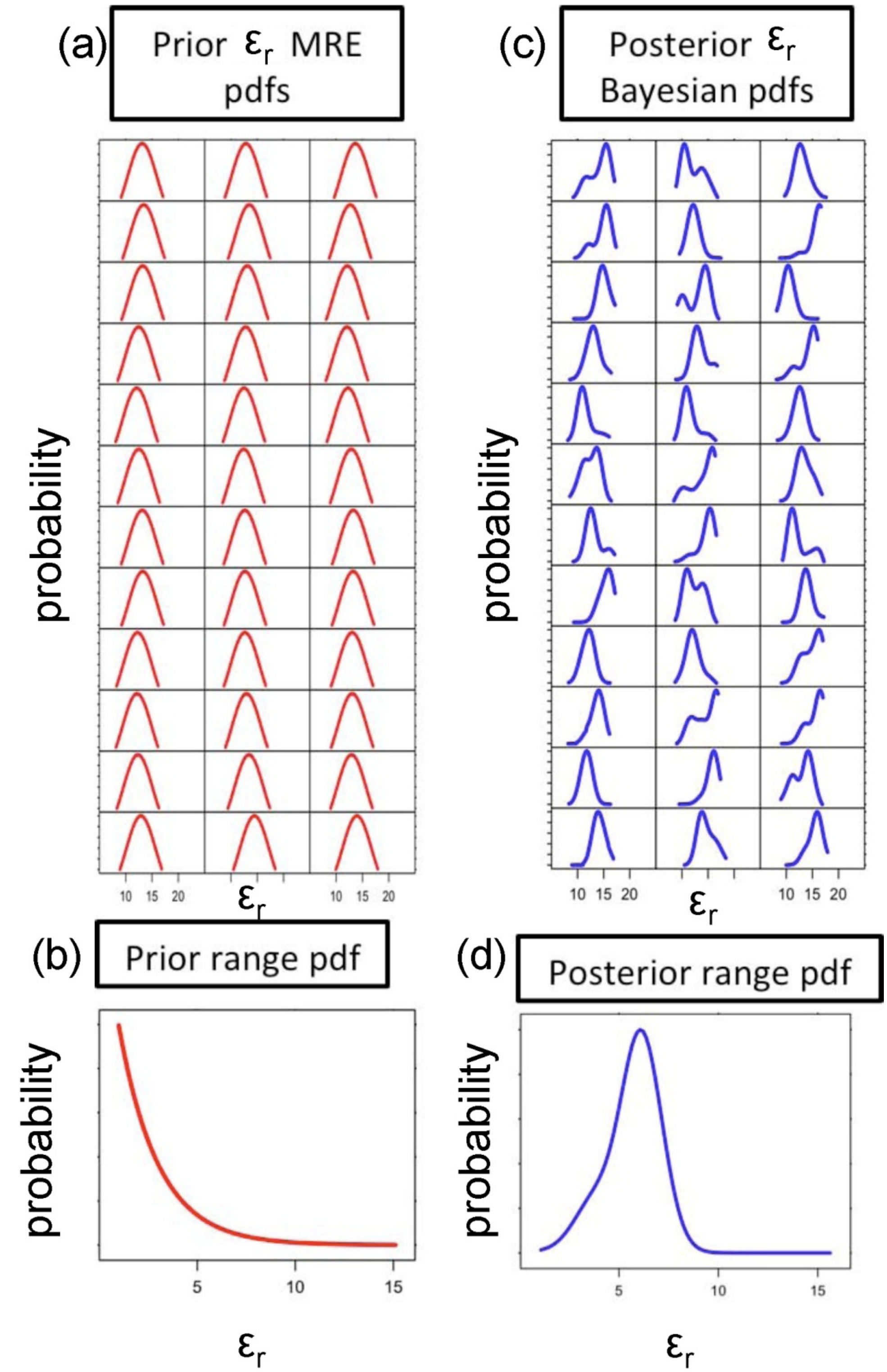

Figure 3: (a-b) are the MRE defined prior pdfs, and (c-d) are the Bayesian updated pdfs. Note particularly the update of the range parameter, which now displays a clear mode of about 6.

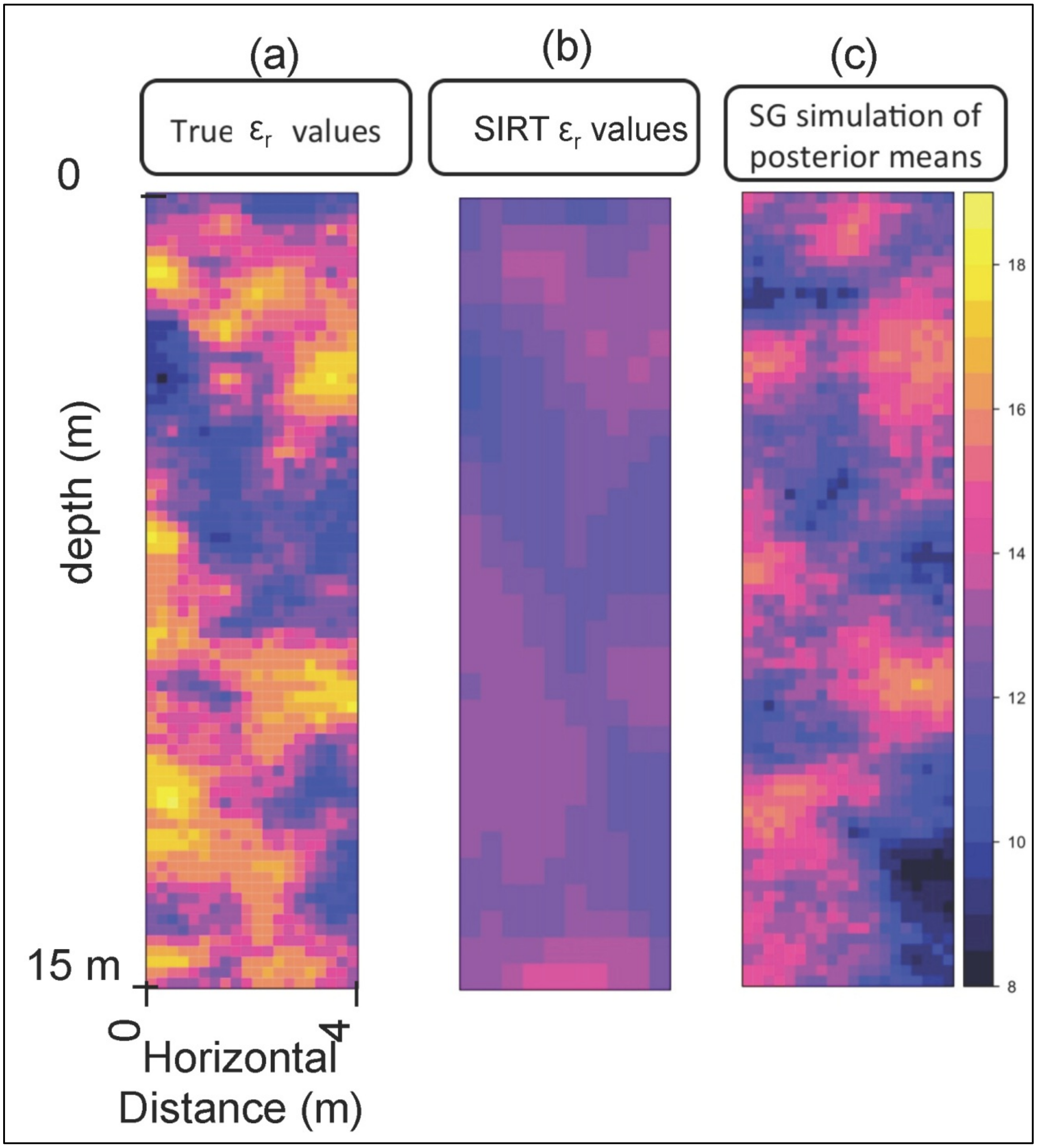

Figure 4: (a) The true relative dielectric permittivity values used in the synthetic study. (b) The SIRT estimated values, and (c) A sequential Gaussian simulation of the posterior Bayesian estimates. Although the SIRT inversion reproduces smeared, general features, it lacks to focus of the simulated values at right. Additionally, the simulated values depict a range of dielectric constant values close to reality.

Another advantage we observe is the quantitative handling of non-uniqueness, or the notion that many velocity fields that could produce similar data. Although the inversion will produce one particular set of parameters that most closely matches the data, this set may still deviate from the true values. Instead of seeking a single best parameter set, which in fact does not usually exist due to non-uniqueness issue and noise in the observations, we use likelihood functions as a weighting system to assign proper weights to each parameter set to allow for all possibilities in a reasonable way.

5. Field Experiment

5.1. Site description

Our data comes from a time-lapse desiccation monitoring experiment at the BC Cribs site, located south of the Hanford 200 area in the Columbia River Basin in Richland, Washington. The site lies above a thick (~ 100 m) vadose zone dominated by sands and gravels, but interspersed with thin, fine-grained layers (Rucker & Fink, 2007). Tomographic GPR and neutron probe data are collected from several of these wells over 4 campaigns dating October 2010 to February 2011, consisting of 1013 measurements each. Beginning at date November 22, 2010 a well in the vicinity of the radar boreholes was used to inject dry nitrogen gas into the soil and a down gradient well was used to extract fluids. The injection-extraction was performed in an effort to test the efficacy of relying on desiccation procedures to remove moisture from the soil, which can potentially decrease the downward mobility of contaminants of concern at the site toward the groundwater (Oostrom, Wietsma, Dane, Truex, & Ward, 2009).

5.2. Inversion

Inversion of the first BC Cribs dataset (October 19, 2010) utilizes the same process as the synthetic experiment (static case). Inversion of the second (December 2, 2010), third (December 16, 2010), and fourth (February 3, 2011) dataset utilize the time-lapse framework outlined in section 3.4.

5.3. Results of the field study

Figure 5 depicts the MRE-derived prior pdfs and the Bayesian posterior pdfs from the October 19, 2011 dataset. Again, we observe a shift in the posterior pdfs toward extreme values, suggesting that the numerical modeling is compensating for the averaging effect produced by the SIRT inversion. The pdf of the parameter correlation range shifts toward higher values, indicating an extended correlation range is more likely compared to the initial image from SIRT inversion.

Turning to the subsequent December 2, 2010 dataset, we see that the SIRT inversion again favors relatively low values for the range pdf (Figure 6). The joint probability of the previous Bayesian posterior and the current MRE-derived prior (Figure 6, center bottom) represents the combined probability of both inversion results. The updated pdfs of the dielectric permittivity pilot points are all shifted more or less toward extreme values, in reflection of these two sources of information (Figure 6, center top).

The December 16, 2010 and February 3, 2011 inversion results are shown in (Figures 7 and 8). We note similar trends in the pdf updating procedure for these datasets as with for the October 19, 2010 and December 2, 2010 datasets, although the shifts in the distributions from the intermediate joint pdfs to the Bayesian posteriors appears less pronounced. This may result from the greater peakedness of the later distributions. Although the distributions still, for the most part, have reasonable ranges from which to draw samples for numerical simulation, the associated probabilities for values near the bounds are extremely small. Although this approach allows for quantification of uncertainty, it cannot, at least at present, fully quantify the full relevance of a previous time-lapse inversion result to the current dataset. Therefore, concurrent with (Hou & Rubin, 2005), we stress the value in maintaining relatively relaxed constraints on parameter pdfs.

We choose realizations of the dielectric permittivity field from posterior estimates and plot the difference images from the Oct 19 baseline, as shown in Figure 9. The mean values for each image are -0.62, -0.29, and -1.2 respectively, the decreases in dielectric permittivity indicates an overall drying of the soil from baseline, in agreement with the desiccation experiment goal. More significantly, these images depict high-resolution heterogeneity and anomalous structure in dielectric permittivity field.

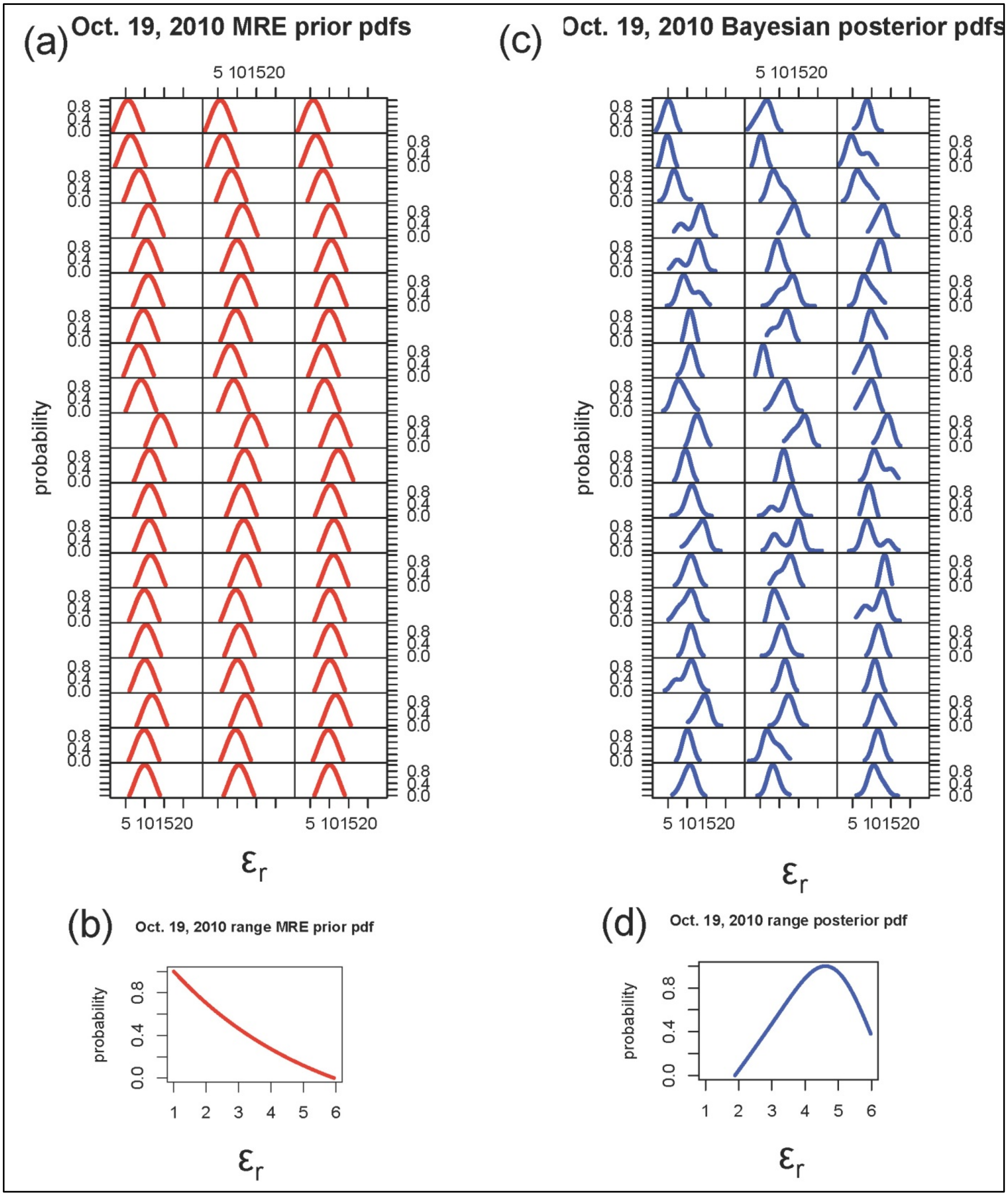

Figure 5: Prior (red, a-b) and posterior (blue, c-d) pdfs from the October 19, 2010 dataset. Note the shift in distributions toward more extreme values.

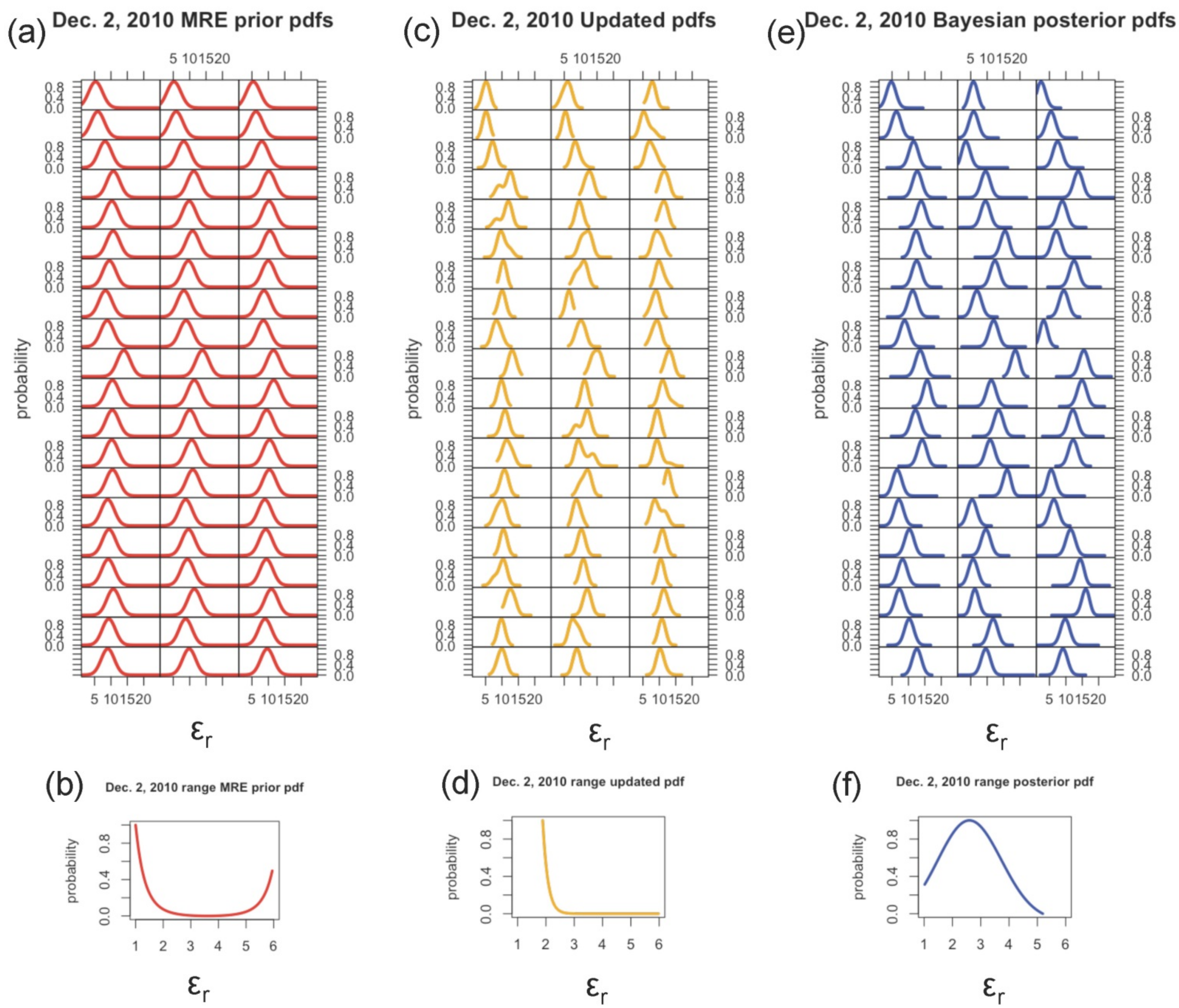

Figure 6: Prior (red, a-b), intermediate (orange, c-d)), and posterior (blue, e-f) pdfs from the December 2, 2010 dataset. The intermediate pdfs represent information from the previous Bayesian inversion, as well as updated information from a SIRT inversion of the current dataset.

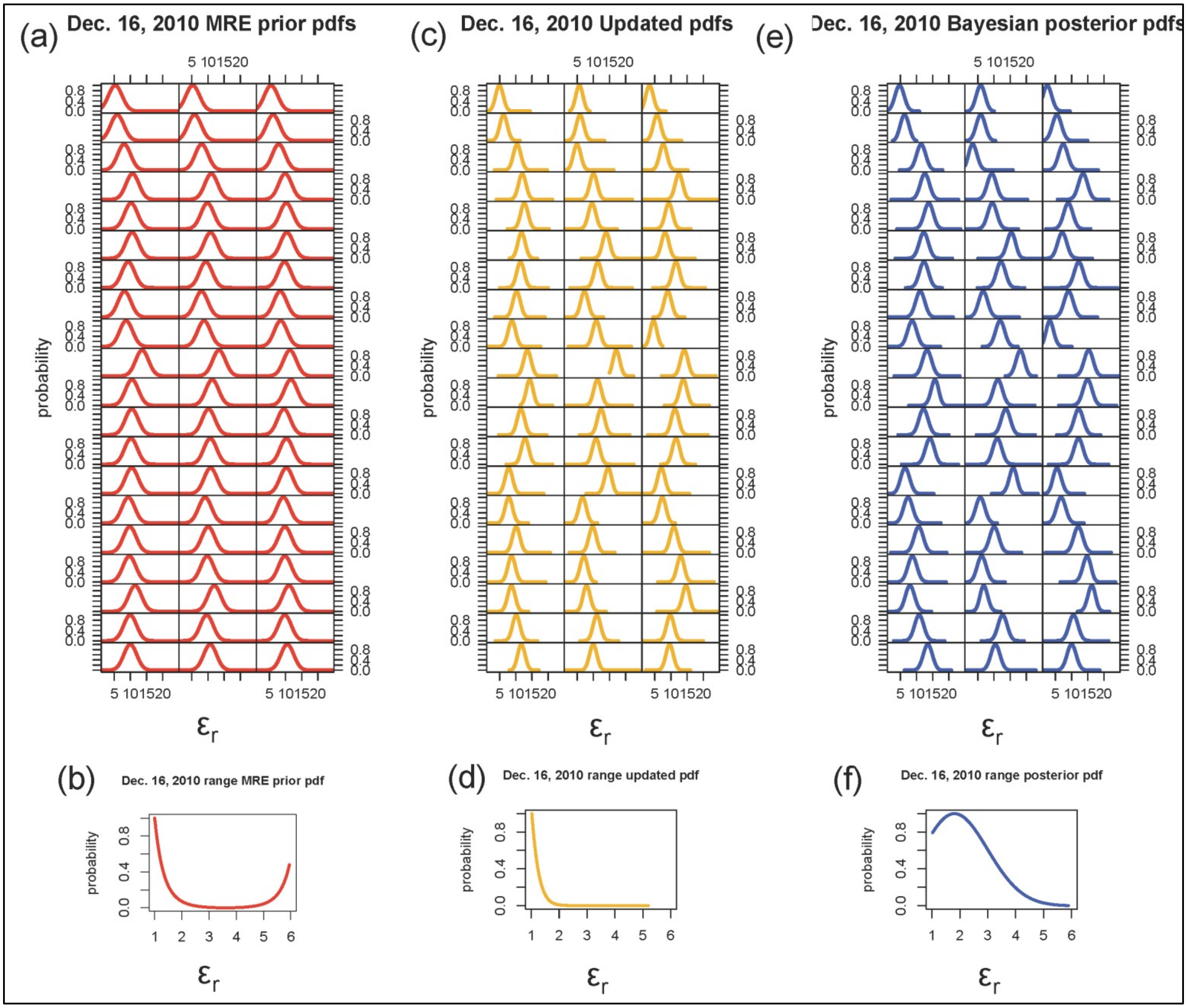

Figure 7: Inversion results for the December 16, 2010 dataset. (a) and (b) (red) are the MRE-derived prior pdfs, (c) and (d) (orange) are the intermediate pdfs reflecting the joint probability of the prior pdfs and the posterior pdfs from the previous time-lapse Bayesian inversion.

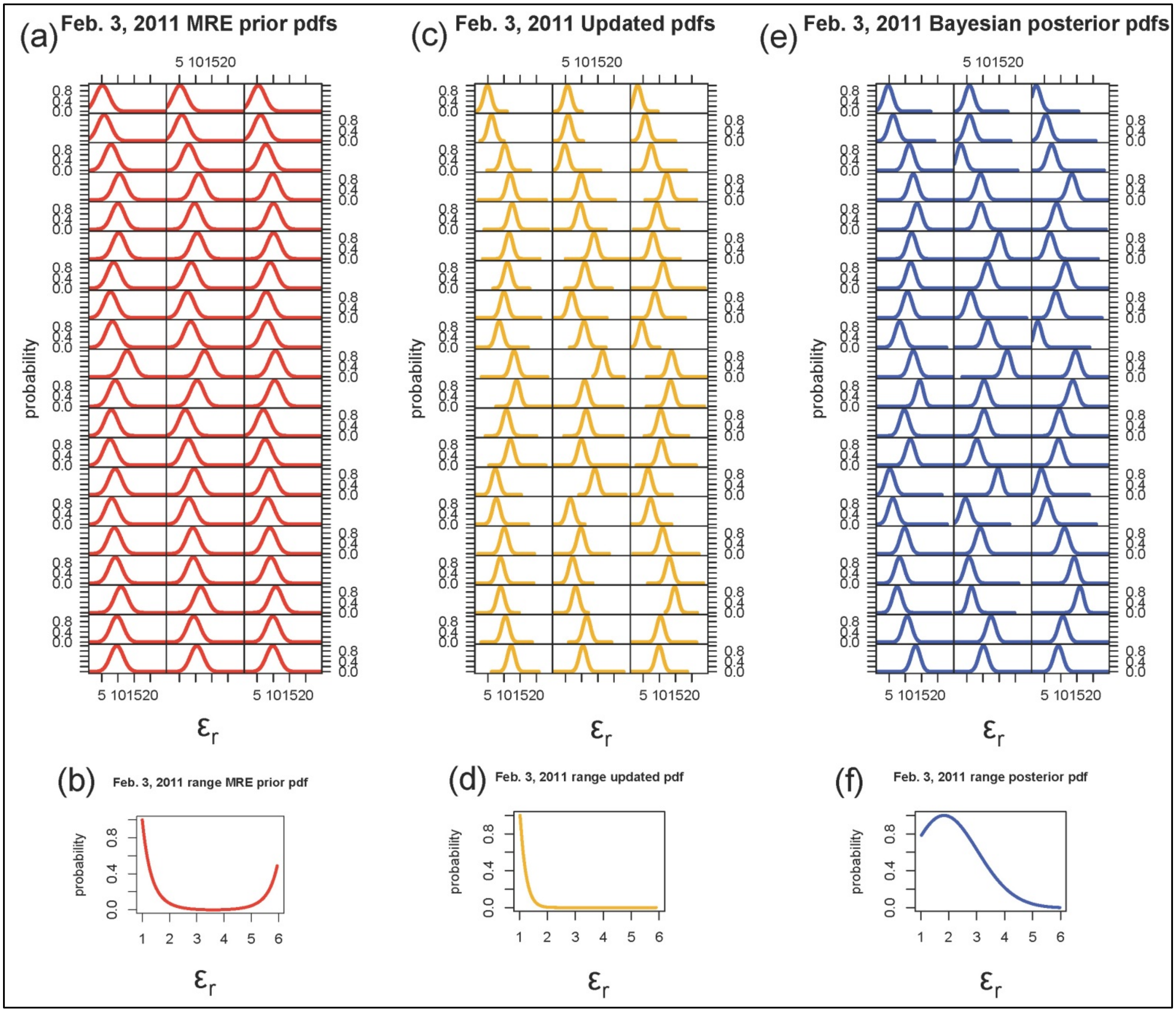

Figure 8: Inversion results for the February 3, 2011 tomography dataset.

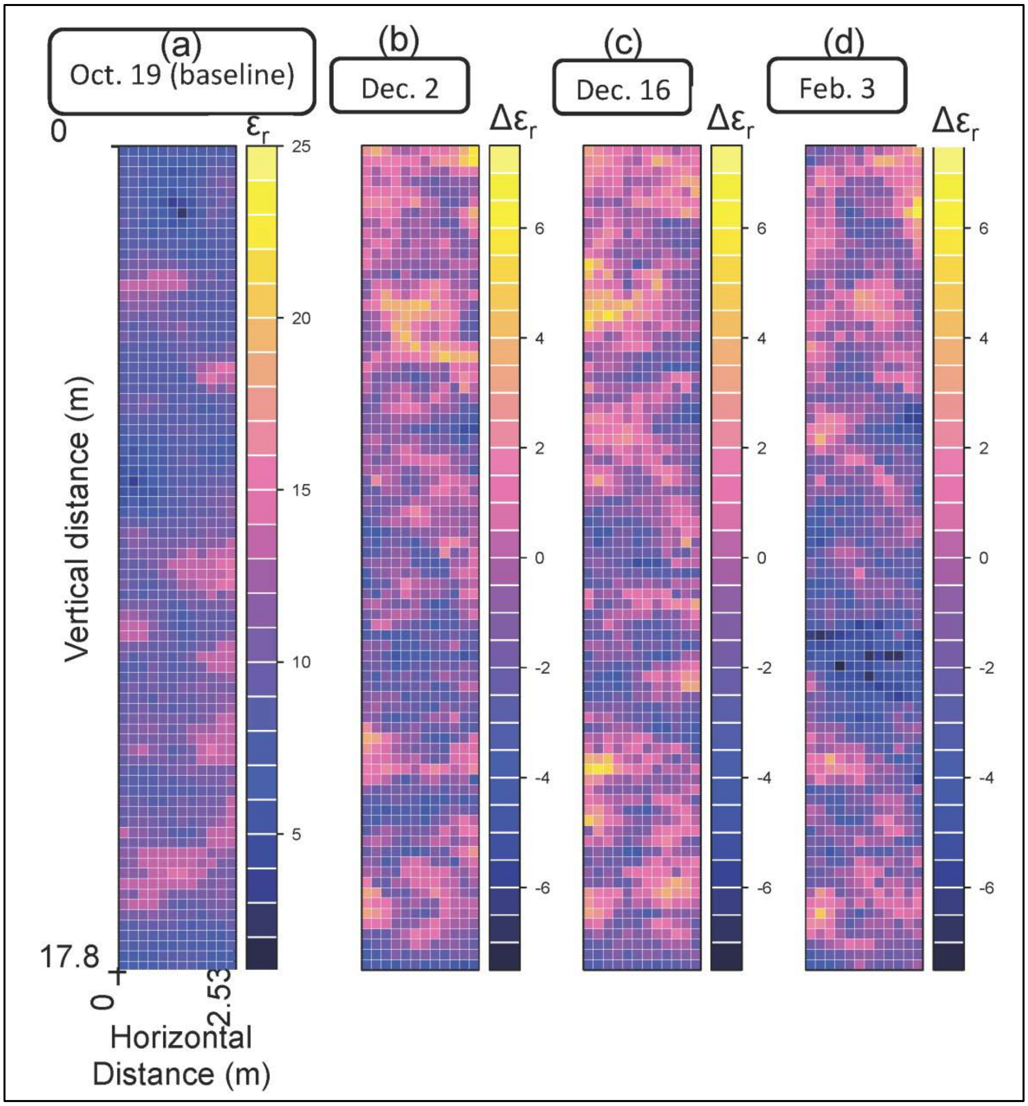

Figure 9: Sequential Gaussian simulations of dielectric constant produced from the Bayesian posterior modes from each time-lapse inversion. Image (a) (October 19, 2010) is the baseline image. (b) and (c) are difference images, representing the change in dielectric constant from the baseline values.

6. Conclusions

In this report, we demonstrate a novel approach to invert tomographic GPR data for dielectric permittivity estimation. It integrates the concepts of entropy, Bayesian updating, efficient sampling, pilot point, and geostatistics. There are several advanced features in the inversion approach: 1) it enables quantification of input and output uncertainty in the form of pdfs; 2) it permits fusion of multiple data types or data collected at different times and assigns proper weights using likelihood functions; 3) it takes advantage of potential spatial correlation in the subsurface state variables by introducing the concept of pilot points, thus reducing the number of parameters to be estimated; 4) it allows us to benefit from temporal correlation/continuity through a memory function, by updating pdfs with observations while quantitatively conditioning on the inverse results using dataset from previous stage. The approach can be easily adapted to a variety of inversion problems, providing a physically-based forward model exists.

The approach also has some limitations. For example, we adopt MRE to avoid the subjectivity in defining parameter priors; there is nevertheless no completely objective approach in assessing the relative value of various data with different quality, different resolution, and/or different amount of observations. Inevitably we also face the challenge of obtaining reliable natural normalization of the observations, and dealing with the prior incompatibility issues. The accuracy of inverse results can also be affected if strong side reflectors exist outside the 2D inversion domain.

A number of possibilities exist to further refine our inversion process. For example, a future effort could involve inverting for changes in dielectric permittivity field instead of estimating the permittivity image directly. Such an approach would require that the baseline estimates were well characterized with adequate resolution, but would enhance our ability to observe small-scale changes in dielectric permittivity. Another consideration is to introduce additional geostatistical parameters. This is particularly true for the Hanford datasets, where we expect relative strong spatial anisotropy (e.g., layering).

7. References

Atanassov, E., Karaivanova, A., Gurov, T., Ivanovska, S., Durchova, M., & Dimitrov, D. S. (2010). Quasi-Monte Carlo integration on the grid for sensitivity studies. *Earth Science Informatics*, *3*(4), 289–296. doi:10.1007/s12145-010-0069-9

Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, (3), 385–402.

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

Caflisch, R. E. (2008). Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, *7*, 1–49. doi:10.1017/S0962492900002804

Cai, J., & Mcmechan, G. A. (1995). Ray-based synthesis of bistatic ground-penetrating radar profiles. *Geophysics*, *60*(1), 87–96. doi:10.1190/1.1443766

Casper, D. A., & Kung, K.-J. S. (1996). Simulation of ground penetrating radar waves in a 2-D soil model. *Geophysics*, *61*(4), 1034–1049. doi:10.1190/1.1444025

Cassiani, G., Strobbia, C., & Gallotti, L. (2004). Vertical radar profiles for the characterization of deep vadose zones. *Vadose Zone Journal*, *3*(4), 1093–1105. doi:10.2136/vzj2004.1093

Chen, J., Hubbard, S. S., & Rubin, Y. (2001). Estimatingthe hydraulic conductivity at the South Oyster Site from geophysical tomographic data using Bayesian techniques based on the normal linear regression. *Water Resources Research*, *37*(6), 1603–1613.

Chen, J., & Rubin, Y. (2003). An effective Bayesian model for lithofacies estimation using geophysical data. *Water Resources Research*, *39*(5), 1118–1128. doi:10.1029/2002WR001666

Chen, X., Rubin, Y., Ma, S., & Baldocchi, D. (2008). Observations and stochastic modeling of soil moisture control on evapotranspiration in a Californian oak savanna. *Water Resources Research*, *44*(8), 1–13. doi:10.1029/2007WR006646

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

Copty, N., Rubin, Y., & Mavk, G. (1993). Geophysical-hydrological identification of field permeabilities through Bayesian updating. *Water Resources Research*, *29*(8), 2813–2825.

Davis, J. L., & Annan, A. P. (1989). Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophysical Prospecting*, *37*, 531–551.

Day-Lewis, F. D. (2004). Assessing the resolution-dependent utility of tomograms for geostatistics. *Geophysical Research Letters*, *31*(7), 4 p. doi:10.1029/2004GL019617

Day-Lewis, F. D. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations. *Journal of Geophysical Research*, *110*(B8), B08206. doi:10.1029/2004JB003569

Deutsch, C. V., & Journel, A. G. (1997). *GSLIB: Geostatistical Software Library and User's Guide (Applied Geostatistics)* (2nd ed., p. 384). New York: Oxford University Press, USA.

Dines, K. A., & Lytle, R. J. (1979). Computerized geophysical tomography. *Proceedings of the IEEE*, *67*(7), 1065–1073. doi:10.1109/PROC.1979.11390

Dubreuil-Boisclair, C., Gloaguen, E., Marcotte, D., & Giroux, B. (2011). Heterogeneous aquifer characterization from ground-penetrating radar tomography and borehole hydrogeophysical data using nonlinear Bayesian simulations. *Geophysics*, *76*(4), J13–J25. doi:10.1190/1.3571273

Hengl, T. (2009). *A practical guide to geostatistical mapping* (2nd ed., p. 270). Amsterdam: University of Amsterdam.

Hou, Z., & Rubin, Y. (2005). On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling. *Water Resources Research*, *41*(12), 1–13. doi:10.1029/2005WR004082

Hou, Z., Rubin, Y., Hoversten, G. M., Vasco, D., & Chen, J. (2006). Reservoir-parameter identification using minimum relative entropy-based Bayesian inversion of seismic AVA and marine CSEM data. *Geophysics*, *71*(6), O77–O88. doi:10.1190/1.2348770

Hubbard, S. S., Chen, J., Peterson, J. E., Majer, E. L., Williams, K. H., Swift, D. J., Mailliox, B., et al. (2001). Hydrogeological characterization of the South Oyster bacterial transport site using geophysical data. *Water Resources Research*, *37*(10), 2431–2456.

Hubbard, S. S., Peterson, J. E., Majer, E. L., Zawislanki, P. T., Williams, K. H., Roberts, J., & Wobber, F. (1997). Estimation of permeable pathways and water content using tomographic radar data. *The Leading Edge of Exploration*, *16*(1), 1623–1628.

Hubbard, S. S., Rubin, Y., & Majer, E. (1997). Ground-penetrating-radar-assisted saturation and permeability estimation in bimodal systems. *Water Resources Research*, *33*(5), 971–990.

Hubbard, S. S., Rubin, Y., & Majer, E. (1999). Spatial correlation structure estimation using geophysical and hydrogeological data. *Water Resources Research*, *35*(6), 1809–1825. doi:10.1029/1999WR900040

Karaoulis, M., Revil, A., Werkema, D. D., Minsley, B. J., Woodruff, W. F., & Kemna, A. (2011). Time-lapse three-dimensional inversion of complex conductivity data using an active time constrained (ATC) approach. *Geophysical Journal International*, *187*(1), 237–251. doi:10.1111/j.1365-246X.2011.05156.x

Kim, J.-H., Yi, M.-J., Park, S.-G., & Kim, J. G. (2009). 4-D inversion of DC resistivity monitoring data acquired over a dynamically changing earth model. *Journal of Applied Geophysics*, *68*(4), 522–532. doi:10.1016/j.jappgeo.2009.03.002

Kowalsky, M. B., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., et al. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resources Research*, *41*(11), 1–19. doi:10.1029/2005WR004237

Kowalsky, Michael B., Dietrich, P., Teutsch, G., & Rubin, Y. (2001). Forward modeling of ground-penetrating radar data using digitized outcrop images and multiple scenarios of water saturation. *Water Resources Research*, *37*(6), 1615–1625.

Kowalsky, Michael B., Finsterle, S., & Rubin, Y. (2004). Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. *Advances in Water Resources*, *27*(6), 583–599. doi:10.1016/j.advwatres.2004.03.003

LaBrecque, D. J., & Yang, X. (2001). Difference inversion of ERT data: a fast inversion method for 3-D in situ monitoring. *Journal of Environmental and Engineering Geophysics*, *6*(2), 83–89. doi:10.4133/JEEG6.2.83

Lehikoinen, A., Huttunen, J. M. J., Finsterle, S., Kowalsky, M. B., & Kaipio, J. P. (2010). Dynamic inversion for hydrological process monitoring with electrical resistance tomography under model uncertainties. *Water Resources Research*, *46*(4), W04513. doi:10.1029/2009WR008470

Lunt, I. A., Hubbard, S. S., & Rubin, Y. (2005). Soil moisture content estimation using ground-penetrating radar reflection data. *Journal of Hydrology*, *307*, 254–269. doi:10.1016/j.jhydrol.2004.10.014

Murdoch, L. (2000). Remediation of organic chemicals in the vadose zone. In B. B. Looney & R. W. Falta (Eds.), *Vadose Zone Science and Technology Solutions* (pp. 949–1156). Columbus, OH: Battelle Press.

Ni-Meister, W., Walker, J. P., & Houser, P. R. (2005). Soil moisture initialization for climate prediction: Characterization of model and observation errors. *Journal of Geophysical Research*, *110*(D13), 1–14. doi:10.1029/2004JD005745

Oostrom, M., Wietsma, T. W., Dane, J. H., Truex, M. J., & Ward, A. L. (2009). Desiccation of unsaturated porous media: intermediate-scale experiments and numerical simulation. *Vadose Zone Journal*, *8*(3), 643–650. doi:10.2136/vzj2008.0182

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

Peterson, J. E., Paulsson, B. N. P., & McEvilly, T. V. (1985). Applications of algebraic reconstruction techniques to crosshole seismic data. *Geophysics*, *50*(10), 1566–1580.

Rubin, Y., & Hubbard, S. S. (2005). *Hydrogeophysics*. (Y. Rubin & S. S. Hubbard, Eds.) (Vol. 50). Dordrecht: Springer. doi:10.1007/1-4020-3102-5

Rucker, D. F., & Fink, J. B. (2007). Inorganic plume delineation using surface high-resolution electrical resistivity at the BC Cribs and Trenches site, Hanford. *Vadose Zone Journal*, *6*(4), 946–958.

Shaxson, F., & Barber, R. (2003). *Optimizing soil moisture for plant production* (p. 107). Rome: Food and Agriculture Orginization of the United Nations.

Sobol, I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *Computational Mathematics and Mathematical Physics*, *7*(4), 784–802.

Sobol, I. M., & Shukhman, B. V. (2007). Quasi-random points keep their distance. *Mathematics and Computers in Simulation*, *75*(3-4), 80–86. doi:10.1016/j.matcom.2006.09.004

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation* (2nd ed., p. 342). Philadelphia: Society for Industrial and Applied Mathematics.

Tohari, A., Nishigaki, M., & Komatsu, M. (2007). Laboratory rainfall-induced slope failure with moisture content measurement. *Journal of Geotechnical and Geoenvironmental Engineering*, *133*(5), 575–587.

Topp, G. C., Davis, J. L., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water Resources Research*, *16*(3), 574–582.

Vasco, D. W., Peterson, J. E., & Lee, K. H. (1997). Ground-penetrating radar velocity tomography in heterogeneous and anisotropic media. *Geophysics*, *62*(6), 1758–1773. doi:10.1190/1.1444276

Vereecken, H., Huisman, J. a., Bogena, H., Vanderborght, J., Vrugt, J. a., & Hopmans, J. W. (2008). On the value of soil moisture measurements in vadose zone hydrology: A review. *Water Resources Research*, *44*, 1–21. doi:10.1029/2008WR006829

Wang, X., & Sloan, I. H. (2008). Low discrepancy sequences in high dimensions: How well are their projections distributed? *Journal of Computational and Applied Mathematics*, *213*(2), 366–386. doi:10.1016/j.cam.2007.01.005

Welford, B. (1962). Note on a method for calculating corrected sums of squares and products. *Technometrics*, *4*(3), 419–420.

Woodbury, A. D. (2004). A FORTRAN program to produce minimum relative entropy distributions. *Computers & Geosciences*, *30*(1), 131–138. doi:10.1016/j.cageo.2003.09.001

Yeh, T.-C. J., & Simunek, J. (2002). Stochastic Fusion of Information for Characterizing and Monitoring the Vadose Zone. *Vadose Zone Journal*, *1*(2), 207–221. doi:10.2113/1.2.207

Zhang, L., Rector, J. W., & Hoversten, G. M. (2005). Eikonal solver in the celerity domain. *Geophysical Journal International*, *162*(1), 1–8. doi:10.1111/j.1365-246X.2005.02626.x

8. Acknowledgements

The authors would like to thank Drs. Mike Truex, Tim Johnson, and Dawn Wellman for permission to use the BC Cribs tomography data for this study, and Dr. John Peterson for collecting and providing the technical data for the BC Cribs experiment. The data were collected under the EM Soil and Groundwater program to Susan Hubbard. The financial support for this project was provided by the DOE-SBR program to Zhangshuan Hou.