Stochastic estimation of aquifer geometry using seismic refraction data with borehole depth constraints

Jinsong Chen,[1] Susan S. Hubbard,[1] David Gaines,[2] Valeri Korneev,[1] Gregory Baker,[2] and David Watson[3]

Received 29 September 2009; revised 29 July 2010; accepted 25 August 2010; published 23 November 2010.

[1] We develop a Bayesian model to invert surface seismic refraction data with depth constraints from boreholes for characterization of aquifer geometry and apply it to seismic and borehole data sets collected at the contaminated Oak Ridge National Laboratory site in Tennessee. Rather than the traditional approach of first inverting the seismic arrival times for seismic velocity and then using that information to aid in the spatial interpolation of wellbore data, we jointly invert seismic first arrival time data and wellbore-based information, such as depths of key lithological boundaries. We use a staggered-grid finite difference algorithm with second-order accuracy in time and fourth-order accuracy in space to model seismic full waveforms and use an automated method to pick the first arrival times. We use Markov Chain Monte Carlo methods to draw many samples from the joint posterior probability distribution, on which we can estimate the key interfaces and their associated uncertainty as a function of horizontal location and depth. We test the developed method on both synthetic and field case studies. The synthetic studies show that the developed method is effective at rigorous incorporation of multiscale data and the Bayesian inversion reduces uncertainty in estimates of aquifer zonation. Applications of the approach to field data, including two surface seismic profiles located 620 m apart from each other, reveal the presence of a low-velocity subsurface zone that is laterally persistent. This geophysically defined feature is aligned with the plume axis, suggesting it may serve as an important regional preferential flow pathway.

Citation: Chen, J., S. S. Hubbard, D. Gaines, V. Korneev, G. Baker, and D. Watson (2010), Stochastic estimation of aquifer geometry using seismic refraction data with borehole depth constraints, *Water Resour. Res.*, *46*, W11539, doi:10.1029/2009WR008715.

1. Introduction

[2] Characterization of aquifer geometry is important for understanding and predicting subsurface flow and transport, as is needed for optimal management of water resources and environmental contaminants. However, such characterization is challenging due to the disparity of scales and mechanisms that contribute to the overall system flow behavior. Conventional techniques for characterizing subsurface flow properties typically rely on wellbore-based techniques, such as core sample analysis or pumping, slug, and flowmeter tests [e.g., *Butler et al.*, 2005]. Unfortunately, the small-scale data obtained using only wellbore-based methods may not provide information about key controls on overall system flow behavior.

[3] Estimation of large-scale flow properties directly from small-scale measurements has been recognized as one of critical unresolved problems in groundwater hydrology [*Anderson*, 2007]. In a recent special issue dedicated to the future of hydrology, *DeMarsily et al.* [2005] stated that "subsurface imaging is obviously an asset, as it can help describe the geometry of the heterogeneous system." Suggesting avenues for future research, they continued that "Geophysics could be used more systematically; so far, good examples of a successful use of geophysical images in the treatment of practical problems where heterogeneity plays a role is lacking."

[4] Although not yet well developed for routine use to treat practical problems, several examples provided by *Rubin and Hubbard* [2005] exemplify the state of the research in hydrogeophysics, illustrating how geophysical methods have been used to improve our understanding of subsurface heterogeneity and predictions of flow. However, the majority of quantitative hydrogeophysical studies have been performed at the local scale (~10 m), where the scale disparity between direct (wellbore) and indirect (geophysical) measurements is often not significant. In discussing hydrogeophysical frontier research areas, *Hubbard and Linde* [2010] described the pressing need to develop approaches for quantifying subsurface heterogeneity over scales relevant to the management of contaminant plumes and water resources. They reviewed recent studies that illustrate the potential of surface geophysical methods for larger-scale subsurface characterization and monitoring, which include the use of: seismic reflection for inferring 3-D subsurface stratigraphy at a contaminated

[1]Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.
[2]Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, Tennessee, USA.
[3]Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Copyright 2010 by the American Geophysical Union.
0043-1397/10/2009WR008715

site [*Addison et al.*, 2009], electrical resistivity methods for quantifying seasonal variations in subsurface moisture content [*Miller et al.*, 2008], and electromagnetic methods for monitoring freshwater-saltwater dynamics [*Falgas et al.*, 2009].

[5] The seismic refraction method is an appealing approach for characterizing subsurface aquifer geometry, because it is relatively inexpensive and can rapidly provide information about subsurface variability over large areas. The method has been used successfully for mapping the distribution of lithological layers in environments associated with petroleum exploration, mining, civil engineering, and deep crustal studies. Seismic refraction methods work by Snell's law, which governs the refraction of sound waves or light rays across the boundary between layers of different physical properties. If the velocity of the seismic waves increases with depth, the wave will be directed back to the ground surface where it can be recorded by geophones. Assessment of the travel times of the recorded refracted waves can provide information about the location of interfaces between units having different physical properties and are thus potentially useful for estimating groundwater aquifers that can be defined by one or more distinct units [*Haeni*, 1986].

[6] However, two problems exacerbate the use of surface seismic methods for delineating subsurface geometry. The first one is associated with the relative placement of low-velocity layers in the subsurface. In deeper subsurface environments, such as those common to ore mining or petroleum reservoirs, the consolidation of the geological layers (thus its seismic velocity) typically increases with depth. Under these conditions, the seismic waves are well behaved in that some energy is successively refracted to the subsurface while some continues to penetrate deeper as was described above. In the near subsurface, however, lower-velocity layers can be situated between layers having higher velocities. As described by *Chen et al.* [2006], low seismic velocity regions in shallow environments can be indicative of regions that are less consolidated, typically having higher porosity and hydraulic conductivity than surrounding layers. Traditional imaging of surface seismic refraction data sets, which rely on raypath assumptions, is challenging in these environments and thus the delineation of such low-velocity zones is subject to a large degree of uncertainty [*Sheehan et al.*, 2005a].

[7] The second key problem in using surface seismic methods to quantify subsurface heterogeneity is associated with the scale disparity between wellbore and geophysical-based measurements. Traditionally, a "two-step" procedure is performed, whereby the surface seismic refraction arrival times are inverted to obtain estimates of the distribution of subsurface seismic velocity along 2-D cross sections, which are then compared with wellbore data collected along the cross section. The 2-D velocity estimates are then used to aid in the interpolation of wellbore-based information (such as interfaces between key units). However, traditional inversion of surface seismic refraction data sets usually leads to a smoothed velocity distribution in the subsurface (rather than to distinct zones that can readily be identified with lithological changes). Additionally, the estimated distribution of seismic velocity can be significantly influenced by inversion choices made by the geophysicist [*Sheehan et al.*, 2005b].

[8] In this study, we develop a Bayesian model to jointly invert surface seismic refraction data with constraints available from wellbores and other sources of information. We first test the developed method using a synthetic data set and then apply it to the data sets collected from the Oak Ridge National Laboratory Integrated Field Research Center (ORNL IFC) for quantifying subsurface architecture at the scale that may impact regional flow and contaminant transport. Our method is distinct from currently used seismic refraction inversion methods in the following four aspects. First, we parameterize the inverse problem directly in terms of aquifer geometry, such as the interface between key geological layers, instead of the common approach of inverting for the distribution of seismic velocity. Second, we formulate the inverse problem within a Bayesian framework, which allows for incorporation of multisource and multiscale information (such as interface and water table locations) in a coherent and systematic way. Third, we adopt a more accurate forward seismic wave propagation model, i.e., the staggered-grid finite difference method with second-order accuracy in time and fourth-order accuracy in space [*Levander*, 1988], to numerically simulate seismic full waveforms with subsequent automated travel time picking. Finally, we use Markov chain Monte Carlo (MCMC) methods to explore the resulting joint posterior probability distribution function and obtain not only the estimates of aquifer geometry but also associated uncertainty information.

[9] The remainder of this paper is organized as follows. Section 2 describes the ORNL IFC site and available information. Section 3 describes the Bayesian framework and MCMC sampling approach for estimating architecture using surface seismic refraction and wellbore data sets. In section 4, we test the approach through synthetic case studies and evaluate the benefit of wellbore data sets for constraining the inversion. We apply the developed method to surface seismic refraction and wellbore data sets collected at the ORNL IFC in section 5. Discussion and conclusions are provided in section 6.

2. Oak Ridge National Laboratory Integrated Field Research Center

2.1. Site Information

[10] The Oak Ridge National Laboratory, located near Knoxville Tennessee (Figure 1a), hosts one of the Integrated Field Research Centers (IFC) of the Subsurface Biogeochemical Research (SBR) Program of the U.S. Department of Energy (DOE). The IFC site includes contaminated and uncontaminated background areas, both located in Bear Creek Valley, west of the Oak Ridge Y-12 National Security Complex [*Watson et al.*, 2005]. At the ORNL IFC, a groundwater plume exists that originated from the former S-3 Waste Disposal Ponds.

[11] The Ponds received 3.2×10^8 L of acidic, nitrate, and uranium-bearing waste for 32 years until the pond contents were neutralized, denitrified, and capped in 1988 [*Watson et al.*, 2004]. Although the ponds are capped, the vast majority of contaminant mass has migrated away from the ponds into the underlying geologic media where it has precipitated or adsorbed onto the solid phase or migrated into the matrix via diffusion, creating an extensive secondary source of contamination [*Jardine et al.*, 2002]. Groundwater interaction with this secondary source of contamination has resulted in a groundwater plume that extends over 4 km down the Bear Creek Valley. Total dissolved solid (TDS) content

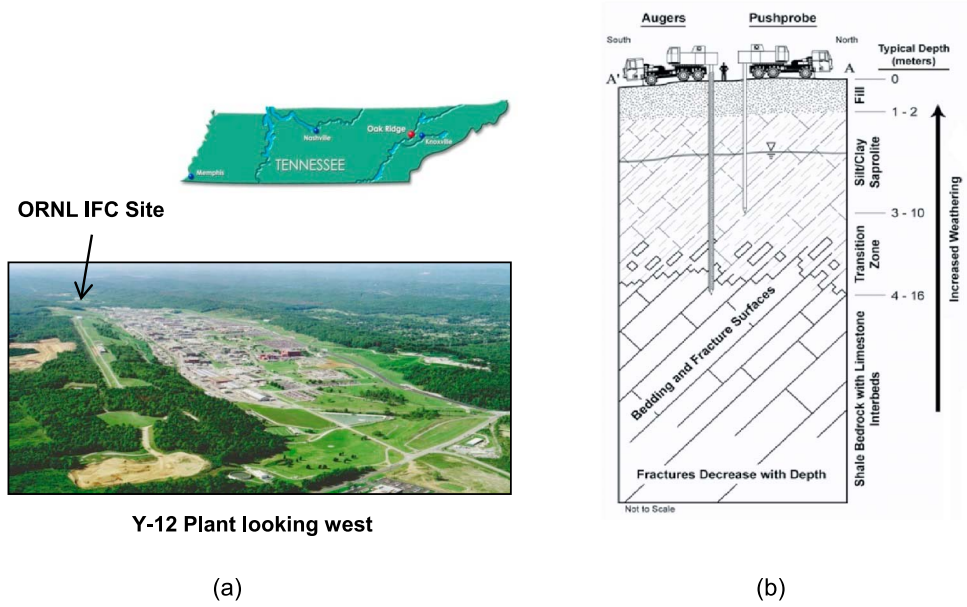

Figure 1. (a) Location of the Oak Ridge National Laboratory Integrated Field Research Center (ORNL IFC) site (USDOE). (b) Geological structure at the Oak Ridge site [after *Watson et al.*, 2005].

of the groundwater plume is greater than 40,000 mg/L. Contaminants in the groundwater plume also contain elevated levels of nitrate (up to 40,000 mg/L), sulfate (up to 3000 mg/L), uranium (up to 50 mg/L), technetium-99 (up to 40,000 pCi/L), and tetrachloroethylene (up to 9 mg/L) [*U.S. Department of Energy* (*USDOE*), 1997; *Watson et al.*, 2005]. Located within a humid regime, the subsurface transport processes are likely influenced by large annual rainfall inputs (>1400 mm/y).

[12] The site geology at the ORNL IFC is quite complex. Figure 1b shows a schematic structure that mimics the topography of Bear Creek Valley, with Appalachian Formations units dipping in the Valley and Ridge Province at a 37° angle to the southwest and aligned parallel to the valley [*Fienen*, 2005]. The primary direction of contaminant transport parallels the strike of bedding planes and the axis of valley groundwater and tends to be constrained within discrete packages of stratigraphic layers. The site is underlain by the Nolichucky Shale, although the top 4–15 m has weathered into unconsolidated saprolite that maintains remnant bedding structure. To a depth of approximately 10 m, the saprolite is clay-rich and has a low permeability [*Watson et al.*, 2005]. Between the shallow, low-permeability clay-rich saprolite and deeper, more competent bedrock is a transition zone that has been weathered to varying degrees. Some areas of the site were excavated and filled during facility construction activities. The fill can be 0–7 m thick with a large fraction consisting of poorly sorted limestone gravel mixed with lesser native saprolite. Groundwater flow near the S-3 Ponds source area is thought to occur within both the shale/saprolite and the fill [*Watson et al.*, 2005].

2.2. Previous Hydrogeological Characterization at the ORNL IFC

[13] Many types of information have been collected from ORNL IFC site over time and across a range of spatial scales to characterize the subsurface hydrogeology, including borehole pumping and flowmeter tests [*Fienen et al.*, 2004], tracer test experiments [*Luo et al.*, 2006; *Fienen et al.*, 2006], crosswell geophysical tomographic surveys [*Chen et al.*, 2006], and surface seismic refraction and electrical methods [*Watson et al.*, 2005; *Sheehan et al.*, 2005a]. Many of the wells were installed using push probe techniques; the depth of penetration of those wells is a function of the competency of the subsurface units and thus can also potentially be useful for subsurface characterization.

[14] Interpretations of both hydrological and geophysical data have suggested that the transition zone, described in section 2.1, serves as a preferential fast flow path. At the local scale (~10 m), *Fienen et al.* [2004] and *Luo et al.* [2006] used pumping and flowmeter tests, together with field-scale tracer test data collected near the S3 ponds, to interpret a fast flow path in the subsurface that appeared to control the overall hydraulic behavior of the system. At the same site, *Chen et al.* [2006] developed and implemented a joint stochastic inversion approach that used crosswell seismic travel times and borehole flowmeter test data to estimate the probability of encountering high hydraulic conductivity zones. They found that the joint inversion was successful at the local scale for identifying a high hydraulic conductivity zone (located between 10.5 and 13.5 m), which was coincident with a seismic low-velocity zone. Comparison of their obtained hydraulic zonation estimates with the tracer and biostimulation experimental results [e.g., *Wu et al.*, 2006] suggested that the estimation approach was very effective in defining the geometry of preferential flow zones at the local scale.

[15] *Watson et al.* [2005] used the surface seismic and electrical data sets, collected adjacent to the S-3 source ponds, to guide the installation of wellbores used for bioremediation tests. In their study, the seismic refraction travel times were first inverted for seismic velocity images using conventional deterministic methods; the velocity gradients inferred from the inversion results were used to indicate contrasts in lithology or rock competency. Our study is

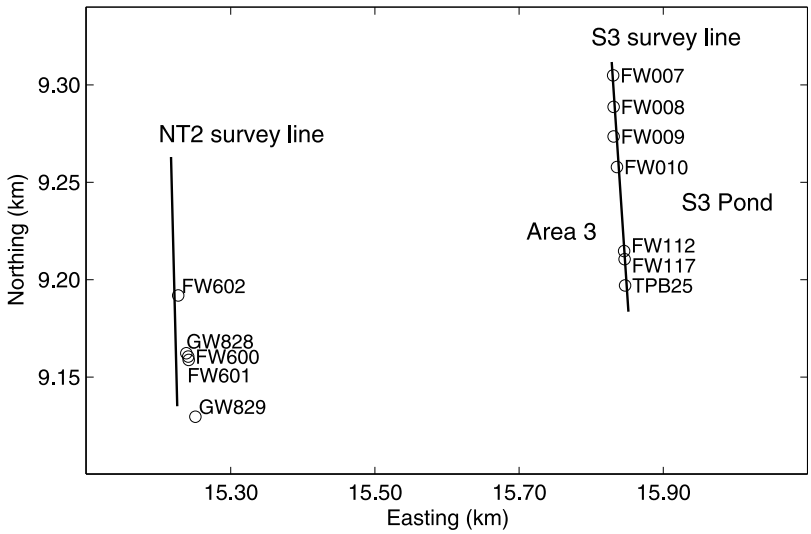

Figure 2. Seismic refraction survey lines and wellbores at the NT2 and S3 pond areas.

based on the premise that inversion of the seismic refraction data with constraints from those wellbores may provide significant information for characterizing aquifer geometry at the large scale. We test this concept by developing and implementing a Bayesian-based method using the two seismic refraction lines (i.e., NT-2 and S3 pond) as shown in Figure 2.

3. Bayesian Model for Aquifer Geometry Estimation

[16] In this section, we develop a Bayesian model based on the ORNL IFC site information. Bayesian methods provide a coherent and practical way to combine multiple sources of information; they have been successfully used in many applied fields as shown by *Gilks et al.* [1996]. In this study, we estimate aquifer geometry from seismic refraction data and borehole-based information. Within the Bayesian framework, the seismic refraction data are connected to unknown parameters (such as lithological interfaces) through likelihood functions and the borehole-based data affect the unknown parameters through prior distributions. The posterior estimates of unknown parameters are combination of both the seismic refraction data and the borehole-based information.

3.1. Parameterization of the Inverse Problem

[17] Since our focus in this study is on characterization of aquifer geometry in an environment where we expect the presence of a fast preferential flow path, we consider our 2-D velocity model as a superposition of three level submodels, each of which includes variable quality information from different sources. The first-level velocity submodel is the one based on gross hydrostratigraphy; it can be reliably derived from information, such as surface topography, groundwater table, and the thickness of man-emplaced fill (or the soil layer). The second-level submodel is the vertical global gradient of seismic velocity, which can be derived from the average of all the borehole velocity profiles, from assessing the conventionally inverted seismic refraction profile or from the joint inversion process. The third-level submodel is our primary focus, referred to as inversion submodel; this is the aquifer geometry that we strive to estimate.

[18] On the basis of the conceptual model of the ORNL IFC site [*Watson et al.*, 2004], we divide the inversion submodel into three distinct zones (Figure 3): (1) weathered saprolite (or unconsolidated bedrock), (2) consolidated bedrock, and (3) transition zone located between the saprolite and the bedrock. As shown in Figure 3, we strive to estimate

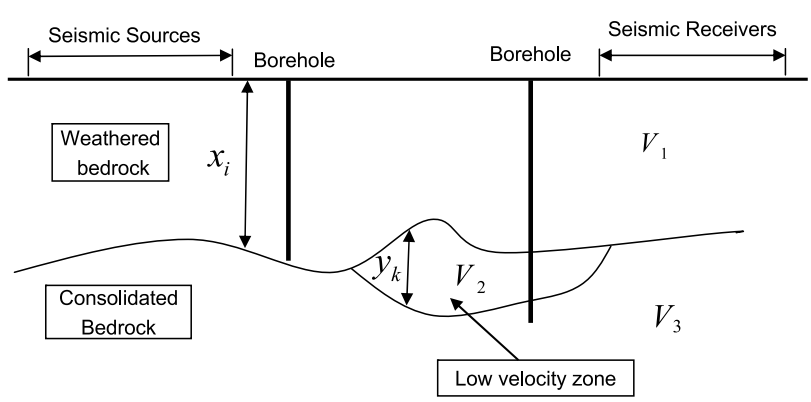

Figure 3. Two-dimensional cross section with inversion parameters.

seismic velocity (v_1, v_2, and v_3) in each zone, interface locations (x_i, i = 1,2,…,n) between the weathered and consolidated zones, and thickness (y_i, i = 1,2,…,n) of the less competent zone as a function of horizontal locations, where n is the total number of locations. The use of zonation concepts for characterizing aquifer geometry has also been used by *Hyndman and Harris* [1996] for inverting crosswell seismic travel times. To reduce the total number of unknowns, we choose to estimate parameters at 5 m intervals, which are referred to as pilot points. For seismic forward simulation, we linearly interpolate values at locations between those pilot points.

[19] The data used in the Bayesian model are the observed seismic refraction first arrival times for many source-receiver pairs. Let d_{ij}^{obs} be the first arrival time recorded at the jth receiver from the ith source. Let m_{s} and m_{r} be the total number of sources and receivers, respectively. For ease of description, we let vector $\mathbf{D}$ represent the seismic refraction data for all the source-receiver pairs, arranged first by sources and then by receivers. That is, $\mathbf{D} = (d_{11}^{\text{obs}}, d_{12}^{\text{obs}}, \ldots, d_{1m_{\text{r}}}^{\text{obs}}, \ldots, d_{m_{\text{s}}1}^{\text{obs}}, d_{m_{\text{s}}2}^{\text{obs}}, \ldots, d_{m_{\text{s}}m_{\text{r}}}^{\text{obs}})^T$, where the superscript denotes observation or measurements, the subscripts denote source-receiver pairs, and symbol T represents the transpose of the vector. We also use depth of refusal or penetration information to constrain unknown variables through prior models as given in section 3.3.

[20] We add two more unknown variables in this model besides those unknowns shown in Figure 3 to describe possible discrepancies between the recorded and numerically calculated seismic first arrival times. The first parameter is the bias between the first arrival time data and our automatic picking from the numerical forward modeling of the first arrival time. The parameter may also account for the systematic difference in picking first arrival time from real seismograms. The second parameter is the standard deviation of random measurement errors. Since the spatial correlation of the interfaces and thickness and the background velocity gradient are also often unknown, we include them in the estimation. Consequently, we can develop the following Bayesian model for the inversion,

$$f\left(\mathbf{X}, \mathbf{Y}, \mathbf{V}, \lambda_x, \lambda_y, g, s, \sigma | \mathbf{D}\right) \propto f(\mathbf{D}|\mathbf{X}, \mathbf{Y}, \mathbf{V}, g, s, \sigma) \cdot f\left(\mathbf{X}, \mathbf{Y}, \mathbf{V}, \lambda_x, \lambda_y, g, s, \sigma\right), \tag{1}$$

where $\mathbf{X} = (x_1, x_2, \ldots, x_n)^T$, $\mathbf{Y} = (y_1, y_2, \ldots, y_n)^T$, and $\mathbf{V} = (v_1, v_2, v_3)^T$ represent the interface locations, the thickness of low-velocity zones, and block seismic velocity, respectively. Variables λ_x and λ_y represent the correlation lengths of the interface locations and the thickness of low-velocity zones; variable g represents the background velocity gradient. Variable s represents the possible constant bias between the true and the picked first arrival times, and variable σ represents the standard deviation of measurement errors. The first term on the right side of equation (1) is referred to as the likelihood function, which is the link between the measured travel times and unknown variables; the second term is referred to as the prior distribution function, which summarizes all information besides the seismic refraction data.

3.2. Likelihood Model of Seismic First Arrival Times

[21] We assume that the first arrival times picked from seismograms include an unknown constant bias and random noise. Let G_{ij} be the first arrival times calculated from the staggered-grid numerical forward model for the ith source and jth receiver. Let ε_{ij} be the random error for the ith source and the jth receiver pair. Thus, we have

$$d_{ij}^{\text{obs}} = G_{ij}(\mathbf{X}, \mathbf{Y}, \mathbf{V}) + s + \varepsilon_{ij}. \tag{2}$$

We assume that error ε_{ij} has a t distribution with a given freedom of ν (i.e., ν = 4), mean of zero, and standard deviation of σ. The reason for choosing t distribution rather than the commonly used Gaussian distribution is that t distribution offers a more robust and realistic model of measurement errors because it has heavier tails.

[22] We model the joint likelihood function of the random errors for all the source-receiver pairs by considering them dependent but uncorrelated rather than independent as we used in other studies [*Chen et al.*, 2006, 2008, 2009]. Consequently, the likelihood function has a multivariate t distribution with the following probability density function, as given by [*Kotz and Nadarajah*, 2004]

$$f(\mathbf{D}|\mathbf{X}, \mathbf{Y}, \mathbf{V}, g, s, \sigma) \propto \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\nu^{1/2}(\pi^{1/2}\sigma)^{m_{\text{r}}m_{\text{s}}}} \cdot \left(1 + \frac{1}{\nu\sigma^2}\sum_{i=1}^{m_{\text{s}}}\sum_{j=1}^{m_{\text{r}}}\left(d_{ij}^{\text{obs}} - G_{ij}(\mathbf{X}, \mathbf{Y}, \mathbf{V}, g) - s\right)^2\right)^{-(\nu+m_{\text{s}}m_{\text{r}})/2}, \tag{3}$$

where Γ represents the gamma function. The multivariate t likelihood model given in equation (3) has advantages over a multivariate Gaussian likelihood model based on the independence assumption of measurement errors in dealing with real data sets in practice as shown by many studies [*Kelejian and Prucha*, 1985; *Joarder and Ahmed*, 1996]. The multivariate t distribution can be interpreted as a scale mixing of multivariate Gaussian distribution using a hierarchical model [*Pinheiro et al.*, 2001; *Marchev and Hobert*, 2004]. It is the distribution of a random vector created by dividing each component of a centered multivariate Gaussian random vector by a common independent gamma-distributed random variable [*Kotz and Nadarajah*, 2004].

3.3. Prior Models and Depth Constraints

[23] The prior distribution is determined from information other than seismic refraction data. We can simplify the prior model in equation (1) by assuming that the location of interfaces between unconsolidated and consolidated bedrock, the thickness of low-velocity zones, the block seismic velocity, the velocity gradient, the bias, and the standard deviation of unknown measurement errors are independent. This assumption is reasonable for this estimation problem. Consequently, we can write the prior distribution function as follows:

$$f\left(\mathbf{X}, \mathbf{Y}, \mathbf{V}, \lambda_x, \lambda_y, g, s, \sigma\right) = f(\mathbf{X}|\lambda_x) f\left(\mathbf{Y}|\lambda_y\right) f(\lambda_x) \cdot f\left(\lambda_y\right) f(\mathbf{V}) f(g) f(s) f(\sigma). \tag{4}$$

We specify prior models for each vector in equation (4) differently, depending on the characteristics of those variables. For the unknown bias, we assume it has the uniform distribution on the range between −3 and 3 ms based on

experience in first arrival time picking. For the standard deviation of errors, we assume its prior distribution has a uniform distribution, as suggested by *Gelman* [2006], with a range between 0.1 and 10 ms. For spatial correlation, we use the exponential variogram and assume the two correlation lengths are uniformly distributed between 1 and 20 m. Those assumptions are reasonable for the site based on other studies [e.g., *Chen et al.*, 2006]. For the background velocity gradient, we calculated the averaged gradient (82 m/s per meter depth) of the velocity profile given by *Watson et al.* [2005] and obtained from traditional inversion methods. We set the prior range from 60 to 100 m/s per meter depth.

[24] We specify the prior models of the interface locations and thickness of the low-velocity zones **X** and **Y** by taking account for both their bounds and spatial correlation. We assume that they have truncated multivariate Gaussian distributions, i.e., $\mathbf{X} \sim \mathrm{Ind}(\mathbf{X} \in D_X)\mathrm{MVN}(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_X)$, and $\mathbf{Y} \sim \mathrm{Ind}(\mathbf{Y} \in D_Y)\mathrm{MVN}(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_Y)$, where Ind() is the indicator function, taking the value of one if the condition is satisfied and zero, otherwise; MVN() is the multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}_X$ or $\boldsymbol{\mu}_Y$ and variance-covariance matrix $\boldsymbol{\Sigma}_X$ or $\boldsymbol{\Sigma}_Y$. Symbols D_X and D_Y are the allowed domains of vector **X** and **Y**, respectively.

[25] We define above vectors and matrices through specification of marginal bounds as done by *Chen et al.* [2006]. Let (l_{xi}, u_{xi}) and (l_{yi}, u_{yi}) be the bounds of the interface location and thickness of low-velocity zones at the ith pilot point. We define the ith component of mean vectors $\boldsymbol{\mu}_X$ and $\boldsymbol{\mu}_Y$ by $\mu_{xi} = (l_{xi} + u_{xi})/2$ and $\mu_{yi} = (l_{yi} + u_{yi})/2$, respectively. We define the ith marginal standard deviation of vectors **X** and **Y** by the quarter of their corresponding interval widths, i.e., $\sigma_{xi} = (\boldsymbol{\mu}_{xi} - l_{xi})/4$ and $\sigma_{yi} = (\boldsymbol{\mu}_{yi} - l_{yi})/4$. To account for the spatial correlation of each component of vectors **X** and **Y**, we use an exponential variogram with unknown correlation lengths λ_x and λ_y. The variance-covariance matrices matrix $\boldsymbol{\Sigma}_X$ and $\boldsymbol{\Sigma}_Y$ are determined by both the marginal standard deviations and the correlation matrices.

[26] At the control points where wellbores are available, we can use data from those wells to constrain prior ranges. At the ORNL site there are generally two types of wells: (1) shallow wells that have been drilled in the last decade to provide information on the lower bounds of the interface locations from geological descriptions; (2) sparse, older, but deeper wells that provide information on both interface and thickness of low-velocity zones. For a shallow well, if z_k^{ref} represents the depth of the well that corresponds to the kth control point, we let $l_{xk} = z_k^{\mathrm{ref}}$. For a deep well, suppose z_k^{deep} and δ_k^{deep} be the interface location and thickness of low-velocity zones, we can assign $l_{xk} = (1 - \alpha)\, z_k^{\mathrm{deep}}$, $u_{xk} = (1 + \alpha)\, z_k^{\mathrm{deep}}$, $l_{yk} = (1 - \alpha)\, z_k^{\mathrm{deep}}$, and $u_{yk} = (1 + \alpha)\, z_k^{\mathrm{deep}}$. Without loss of generality, we assume $\alpha = 0.1$ in this study.

3.4. Markov Chain Monte Carlo Sampling Methods

[27] We cannot obtain analytical solutions of the inverse problem defined by equation (1) because the joint posterior probability distribution function involves numerical forward modeling of seismic first arrival times and because the likelihood function given in equation (3) is a nonlinear function of the unknown parameters. Therefore, we use MCMC sampling strategies to draw many samples from the posterior joint distribution.

[28] To implement MCMC sampling strategies, we first need to derive the conditional probability distribution for each type of unknowns given all other unknowns by dropping those terms that are not directly related to it in equation (1). For example, for unknown vector **X**, we can have $f(\mathbf{X}|\cdot) \propto f(\mathbf{D}|\mathbf{X}, \mathbf{Y}, \mathbf{V}, g, s, \sigma) f(\mathbf{X}|\lambda_x)$. Similarly, we can obtain conditional probability distributions of other unknown vectors. We then need to draw samples using block-wise slice [*Neal*, 2003] or Metropolis-Hastings [*Metropolis et al.*, 1953; *Hastings*, 1970] sampling methods. Finally, we use the method developed by *Gelman and Rubin* [1992] to check the convergence of Markov chains.

4. Synthetic Case Studies

[29] In this section, we apply the developed Bayesian model to synthetic seismic refraction data generated by mimicking conditions at the ORNL IFC site. We use the synthetic study for two purposes: (1) to check if the developed Bayesian model and its associated MCMC sampling methods are effective for aquifer geometry estimation and (2) to investigate the effects of depth constraints on the estimation under relative simple conditions.

4.1. Synthetic True Model and Survey Geometry

[30] We developed a synthetic 2-D velocity model based on a conceptual model of the seismic velocity distribution from previous research [*Watson et al.*, 2005; *Chen et al.*, 2006]. As shown by Figure 4a, the 2-D cross section spans 130 m and extends to a depth of 25 m. We divide it into 520 × 100 intervals (i.e., $m = 520$, $n = 100$, and $\Delta x = 0.25$m, $\Delta z = 0.25$m). On the basis of typical geological structure at this site (see Figure 1b), we divide the cross section into four layers: (1) shallow topsoil with velocity of 600 m/s, (2) human-placed fill with velocity of 1000 m/s, (3) weathered bedrock (saprolite) with velocity of 1200 m/s, and (4) consolidated bedrock with velocity of 1400 m/s. Between the unconsolidated and consolidated bedrock is a low-velocity zone of 800 m/s. We add a gradient of 82 m/s per meter to the above 2-D block velocity model. To reduce the total number of unknowns, we pick 27 pilot points that are evenly spaced along the x axis with spacing of 5 m.

[31] We perform a forward simulation using the staggered-grid finite difference model to obtain seismic refraction travel time data. We use a Delta function pulse as input to the numerical system because it generates sharp first arrivals for which the automatic picking of first arrivals works robustly. Although the Delta pulse leads to contaminated waveforms by numerical dispersion in later arrivals, as we only use first arrival time in this study, there is no effects of the numerical artifacts [*Korneev et et al.*, 2008]. We follow similar acquisition procedures used in the field to collect real data. There are 64 shot sources, starting from $x = 1$ m with a spacing of 2 m, and 128 receivers, starting from $x = 1$ m with a spacing of 1 m. Each shot source and receiver are located in depth of 0.5 m. We assume that there are four boreholes nearby the survey lines. Two of those wells ($x = 10$ m and $x = 110$ m) are shallow and provide information about the depth of the interface between unconsolidated and consolidated bedrock; the other two wells ($x = 50$ m and $x = 90$ m) are deep and provide depth information on both the interface and the thickness of less competent zones.

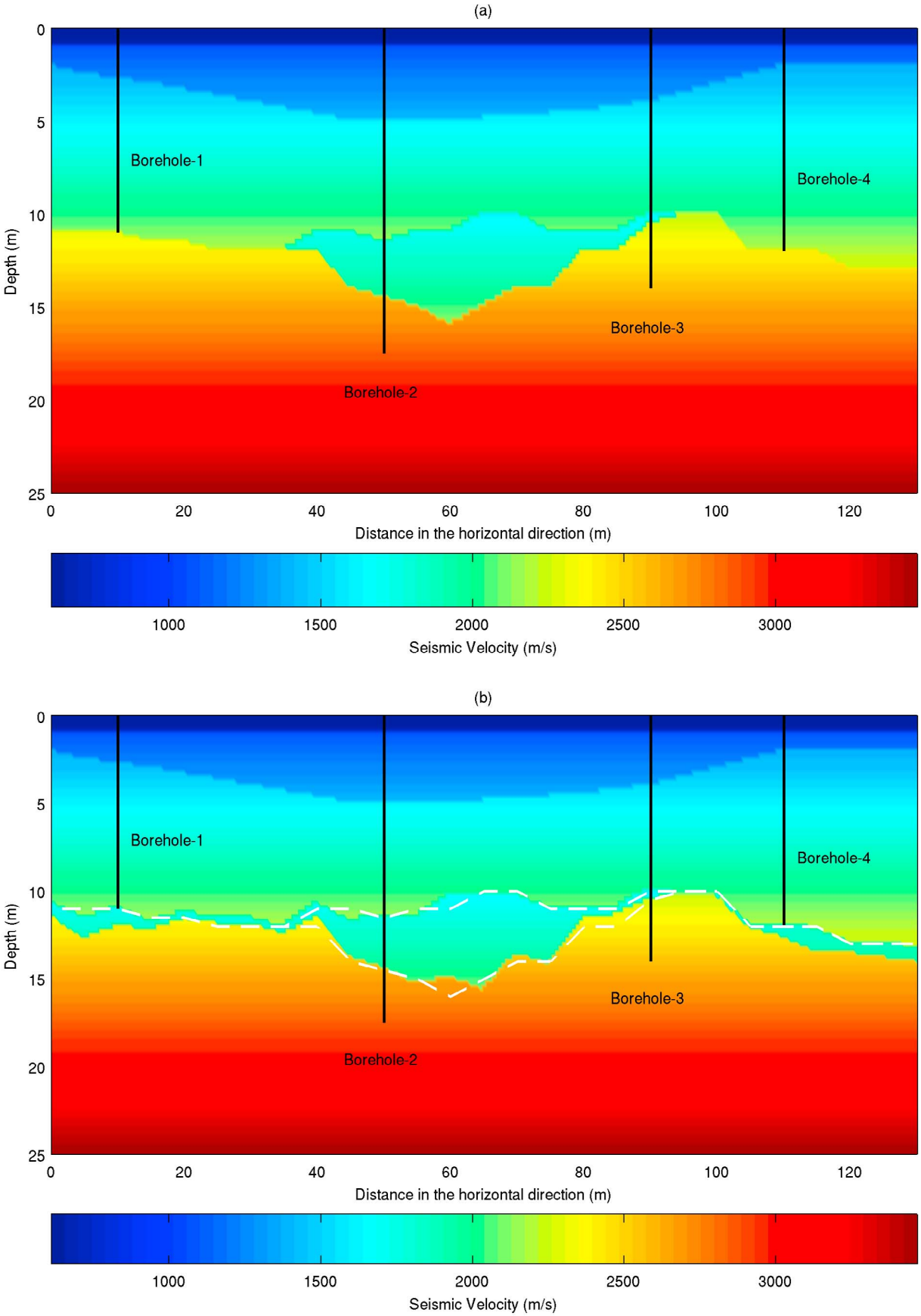

Figure 4. (a) Synthetic true velocity model. (b) Estimated velocity model using the stochastic inversion method.

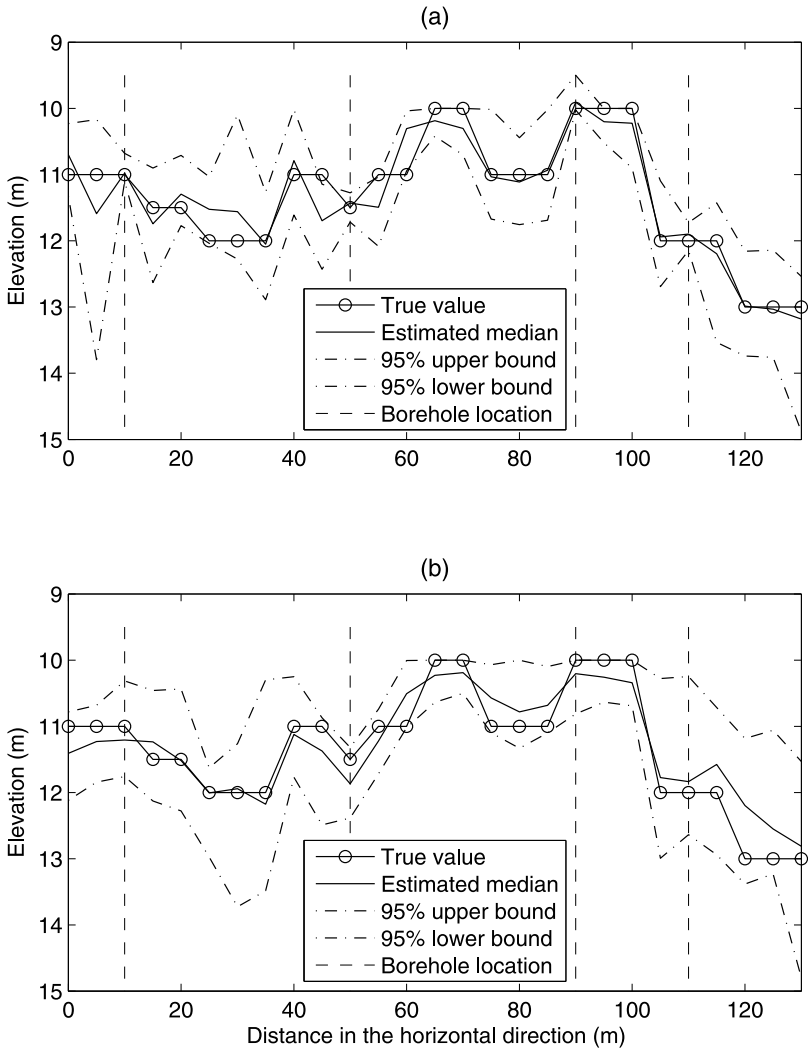

Figure 5. Estimated interface locations (a) with depth constraints and (b) without depth constraints.

[32] We generate synthetic first arrival times using the numerical forward model. We assume that there is a constant of 1 ms bias, which accounts for possible differences between the automatically and manually pickings of first arrival times. Moreover, we add Gaussian random noises with zero mean and 1 ms standard deviation. The added bias and random noises are reasonable for the Oak Ridge site data sets based on the error studies conducted by *Gaines et al.* [2008].

4.2. Stochastic Inversion of Synthetic Data Sets

[33] We invert the synthetic seismic data as described above using MCMC sampling methods for wide priors except at locations where wellbores are available. The prior ranges of three block velocities are [1000, 1400], [600, 1000], and [1200, 2000] m/s, respectively; the prior ranges of interface locations and thickness of the low-velocity zones are [8, 16] m and [0, 6] m. Comparison with the true synthetic model shows that block seismic velocities can be accurately estimated from the inversion because seismic first arrival times are mostly sensitive to changes in those parameters. This implies that when we invert field data for aquifer geometry, we can use wide prior ranges for block seismic velocity and the inversion results will not be very sensitive to the specification of the priors.

[34] The locations of interfaces between unconsolidated and consolidated bedrock can be resolved reasonably well. As shown in Figure 4b, we can see that the medians of the estimated interface locations (white dashed curves) are close to their corresponding true values, even at locations near both edges where seismic wave coverage is very low. This is because we assume that the interface locations are spatially correlated. However, as shown in Figure 5, uncertainty associated with the estimation is high, depending on whether we use or not use depth constraints.

[35] The performance of the stochastic inversion for estimating the thickness of low-velocity zones varies. As shown in Figure 6, the inversion method provides good estimates of thickness of low-velocity zones in the middle locations (40 m $< x <$ 100 m) along the x axis, but poor estimates at edge locations ($x <$ 40 m or $x >$ 100 m). The discrepancy is caused by a difference in the distribution of seismic waves. Similar to the estimates of interface locations, uncertainty in the estimates of the thickness is also large. The effects of depth constraints from wellbores on the interface and thickness will be further addressed in the subsequent section.

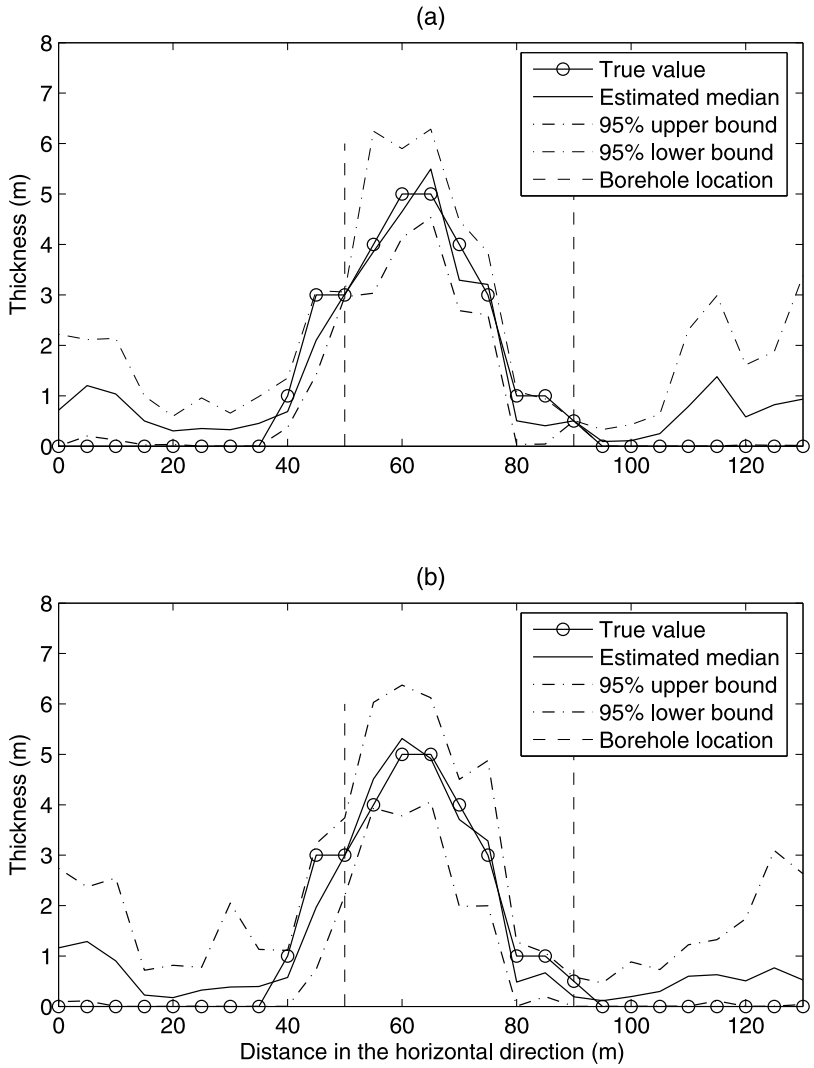

Figure 6. Estimated thickness of low-velocity zones (a) with depth constraints and (b) without depth constraints.

4.3. Effects of Depth Constraints on Estimation

[36] We invert the synthetic seismic data with and without wellbore information to investigate the effects of depth constraints on the estimation of aquifer geometry. Figures 5a and 5b compare the estimated medians and 95% predictive intervals of interface locations between unconsolidated and consolidated bedrock with and without depth constraints. We can see that the medians of both inversions are close to the true interface locations with the root mean squares (RMS) of difference from the true values of 0.31 m with depth constraints and of 0.32 m without depth constraints. Since the estimated medians of interface locations with depth constraints are very close to those without depth constraints, our focus here is on the comparison of uncertainty associated with the constrained and unconstrained estimation.

[37] We can see that overall the estimates of interface locations with depth constraints (Figure 5a) have smaller uncertainty than those without depth constraints (Figure 5b). The averaged width of highest probability domains (HPDs) with depth constrains has a value of 1.49 m while that without depth constraints has a value of 1.74 m. To see the detailed effects of depth constraints, we also show depth constraints at the four borehole locations (black vertical dashed lines). We found that at locations near the edges (x = 10 m or x = 110 m), the constraints reduce the uncertainty significantly and at locations close to the middle (x = 50 m or x = 90 m), the constraints reduce uncertainty slightly. This is because at locations in the middle, we have more wave coverage and thus the interfaces are better determined from seismic refraction data alone.

[38] Figures 6a and 6b show the thickness and location of the low-velocity zones, where the solid lines with circles are the true thickness and the solid and dash-dotted lines are the medians and the lower and upper bounds of 95% HPDs. We can see that in the location from x = 20 m to x = 100 m, the medians of both inversions are close to the true thickness with the root mean squares of difference from the true values of 0.42 m with depth constraints and of 0.40 m without depth constraints. Similar to the interface location, we compare uncertainty associated with the constrained and unconstrained estimation in the range and find that the averaged width of thickness with depth constraints (1.32 m) is much smaller than that without depth constraints (1.72 m). At the

two locations (x = 50 m and x = 90 m), where depth constraints are available, uncertainty in the estimates of thickness with depth constraints are significantly smaller than that without depth constraints.

[39] Different from the interface location, the medians of estimated thickness with and without depth constraints are significantly away from the corresponding true values at locations near both ends (i.e., $x < 20$ m or $x > 100$ m). This is partly because for the given survey geometry, we have much less wave coverage in the areas than that in the middle areas. Therefore, depth constraints at locations near edges play a far more important role than at locations in the middle area for improving estimates of thickness. Another reason is that the resolution of seismic refraction surveys typically decreases with increasing depth. For example, the RMS of misfits is about 0.3 m for interface location estimation and 0.4 m for thickness estimation. This again shows that the depth constraints from deeper wellbores are more important, especially at locations near edges.

5. Inversion of ORNL IFC Field Data Sets

[40] We apply the Bayesian model developed in section 3 to field seismic refraction and wellbore data sets collected down gradient in the NT-2 area and adjacent to the S3 ponds (see Figure 2). These profiles are located perpendicular to groundwater flow direction and are separated by 620 m. Although inversion of data collected from the S3 survey line using conventional deterministic methods indicated the possible existence of a low seismic velocity zone along the cross section [*Sheehan et al.*, 2005b; *Watson et al.*, 2005], given uncertainty in the two-step conventional approach described in section 1, the estimated location and geometry of the fast flow path zone are qualitative and subject to a large degree of uncertainty. To quantitatively estimate the aquifer geometry and associated uncertainty information, we apply the developed stochastic inversion approach to the seismic refraction data collected from both survey lines with depth constraints from local environments.

5.1. Inversion Using Data From the NT2 Survey Line

5.1.1. Survey Location and Geometry

[41] On the NT2 area, the ground surface decreases from the south (x = 0 m and elevation = 298 m) to the north (x = 120 m and elevation = 296 m). A total of 55 sources were used, starting from x = 0.5 m with increment of 4 m. A total of 221 receivers (or geophones) were used, starting from x = 0 m with increment of 0.5 m. We consider a rectangle inversion domain with a lateral extent of 120 m from the south to the north and a vertical extent of 25 m from elevation of 300 m to elevation of 275 m. For forward simulation, we use the same grid size for both horizontal and vertical directions, which is equal to 0.25 m. This renders 480 grids along the lateral direction and 100 grids along the vertical direction.

5.1.2. Depth Constraints From Wellbores and Other Types of Information

[42] Several wellbores are located near the NT2 survey profile as shown in Figure 2, and four of them are within the horizontal distance of 22 m from the profile. Wellbore GW828 has a horizontal distance of 18 m from the survey line and the total penetration depth of 51 m. Weathered rock (shale) was encountered at the depth of 2.3 m and fresh rock at the depth of 11.8 m [*USDOE*, 1995]. Wellbores FW600, FW601, and FW602 are also nearby and they are all shallow. The depths of refusal (or penetration) for wellbores FW600, FW601, and FW602 are 9.0, 6.1, and 10.2 m, respectively. Since FW600 and FW601 are very close, with a distance of 1.6 m, we only use information from wellbore FW600 in the inversion. The depths encountered fresh rock are likely beyond those depths.

[43] We also use information from the conventional inversion results [*Gaines et al.*, 2008] to build a prior velocity model for shallow subsurface. The three-layer background block velocity model has fixed thickness for each layer, which includes a very thin layer with a thickness of 0.5 m and a constant velocity of 300 m/s, a topsoil layer with a thickness of 1.0 m and a constant velocity of 500 m/s, and an unsaturated layer above the groundwater level with a thickness of 2.4 m and a constant velocity of 800 m/s.

5.1.3. Inversion of Seismic Refraction Data

[44] We invert seismic refraction data under several different conditions. The prior ranges of the bias, the standard deviation of measurement errors, the spatial correlation lengths, and the background gradient are given in section 3.3. The prior ranges of three block velocity are the same as those used in the synthetic studies for the Bayesian model with a background model, and they are [1000, 3000], [1000, 3000], and [2000, 4500] m/s for the Bayesian model without a background model. The prior ranges of the interface locations are [9, 18] m in depth; the prior range of the thickness is [0, 3] m for the narrow-bound case and [0, 6] m for the wide-bound case.

[45] We first invert seismic refraction data using the narrow bounds for thickness without a background gradient model. Figure 7 shows the 2-D velocity model calculated using the estimated medians of unknown parameters with depth constraints (Figure 7a) and without depth constraints (Figure 7b). Both images show the existence of two major low-velocity zones on the cross section: one is around x = 50 m and the other is around x = 80 m. The interface locations are also very similar on the right side ($x > 90$ m). However, the estimated aquifer geometry at the south end ($x < 40$ m) in Figure 7b is significantly different from that in Figure 7a due to the lack of depth constraints from wellbores. At wellbore GW828, the estimated interface at the wellbore location obtained from inversion with depth constraints is consistent with the observed depth that fresh bedrock was encountered (black circle in Figure 7a), but that obtained from inversion without depth constraints much shallower than the observation.

[46] The Bayesian model allows us to quantify uncertainty associated with the parameter estimation. Figure 8 shows the medians and the lower and upper bounds of 95% HPD for the interface locations (Figure 8a) and for the thickness (Figure 8b). We can see that uncertainty in the interface is smaller than that in thickness relative to their corresponding prior ranges ([9, 18] m for the interface and [0, 3] m for the thickness) because the resolution of seismic refraction survey typically decreases with increasing of depth. From Figure 8a, we can also see that uncertainty in the interfaces at the locations with depth constraints (i.e., x = 25, 30, and 60 m) is significantly smaller than that without depth constraints.

5.1.4. Effects of Prior Information on Estimation

[47] We vary the prior ranges of several types of unknown parameters on which we have less controls to investigate the effects of prior information on the estimation. We inverted

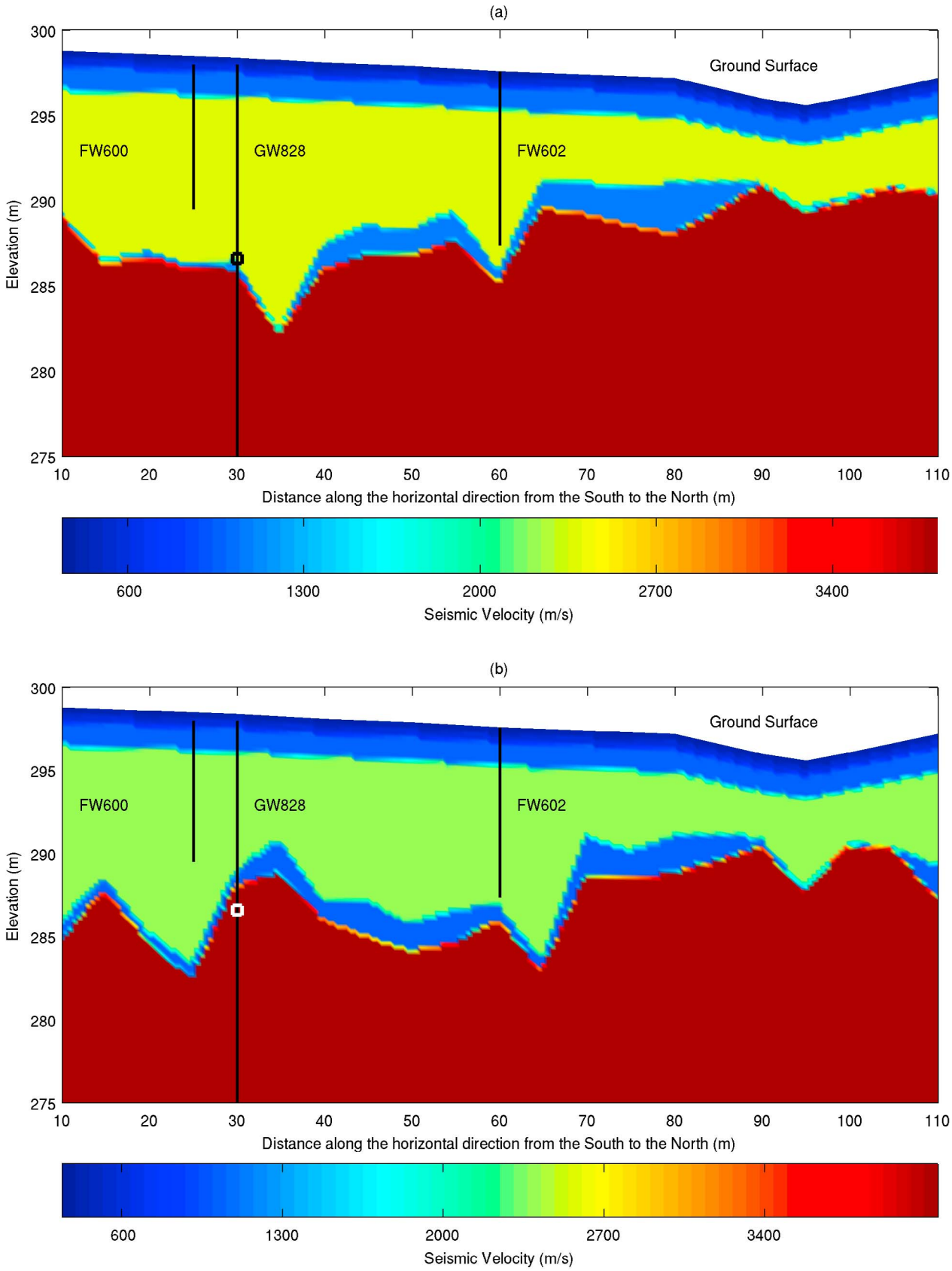

Figure 7. Estimated 2-D profile along the NT2 survey line (a) with depth constraints and (b) without depth constraints.

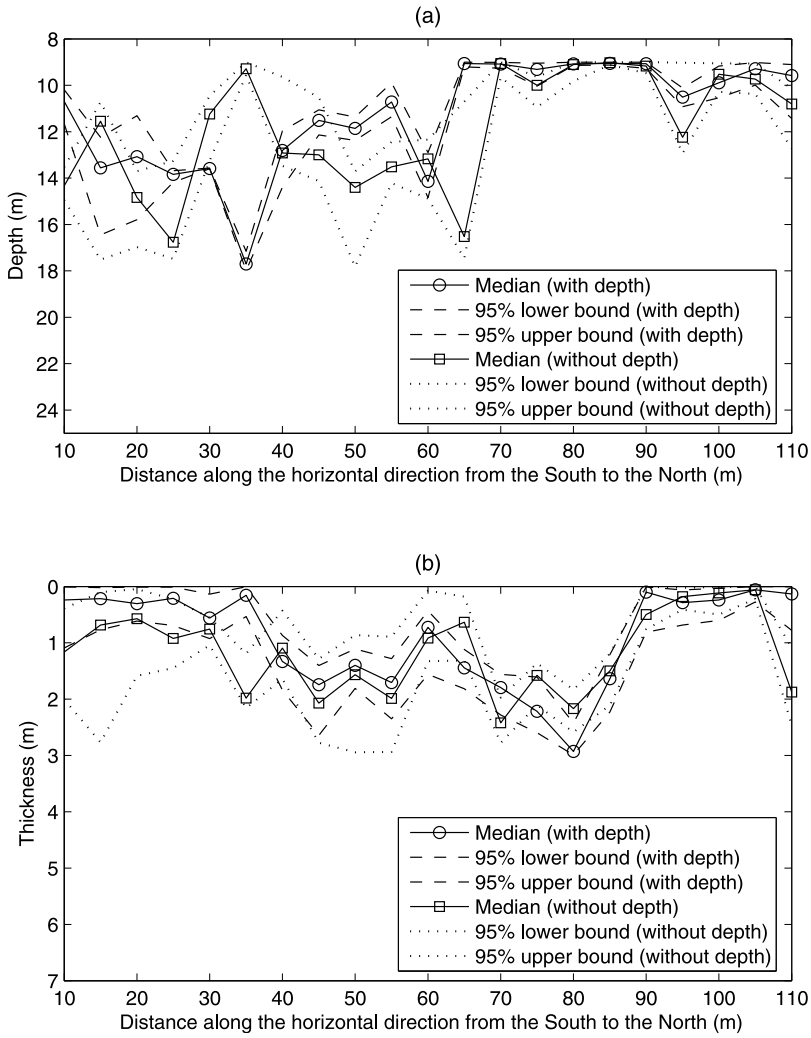

Figure 8. Uncertainty associated with inversion at the NT2 area for (a) interface locations and (b) thickness of low-velocity zones.

the seismic refraction data under the following six different situations: (1) using the narrow thickness bounds with depth constraints and without a background gradient model, referred to as "narrow-depth-block"; (2) using the narrow thickness bounds without depth constraints and a background gradient model, referred to as "narrow-nodepth-block"; (3) using the narrow thickness bounds with depth constraints and a background gradient model, referred to as "narrow-depth-gradient"; (4) using the wide thickness bounds with depth constraints and without a background gradient model, referred to as "wide-depth-block"; (5) using the wide thickness bounds without depth constraints and a background gradient model, referred to as "wide-nodepth-block"; (6) using the wide thickness bounds with depth constraints and a gradient background, referred to as "wide-depth-gradient."

[48] Figure 9 compares the estimated thickness of low-velocity zones for the above six cases. Although there are quite uncertain because of different prior information used for inversion, they all suggest there are low-velocity zones around location x = 50 m and x = 80 m. To obtain more accurate estimates of aquifer geometry, the use of more information is beneficial. Similarly, we analyze the effects of other types parameters, for instance, the prior ranges of three block velocity, and the prior ranges of the bias and the standard deviation of measurement errors. The effects of those parameters are less prominent.

5.2. Inversion Using Data From the S3 Survey Line

5.2.1. Survey Location and Geometry

[49] There are two survey profiles near the S3 pond. The first profile has 63 sources, starting from x = 35 m with increment of 2 m, and has 48 receivers, starting from x = 82 m with increment of 1 m. The second profile has 50 shot sources, starting from x = 28 m with increment of 2 m, and has 48 receivers, starting from x = 52 m with increment of 1 m. We consider a rectangle inversion domain with a lateral extent of 130 m from the south to the north and a vertical extent of 25 m downward from the ground surface. For forward simulation, we also use the same grid size for both horizontal and vertical directions, which is equal to 0.25 m. This gives us 520 grids along the lateral direction and 100 grids along the vertical direction.

5.2.2. Depth Constraints From Wellbores and Prior Model From Other Information

[50] Many wellbores have been drilled near the S3 pond survey lines in the past two decades. There are 22 wellbores

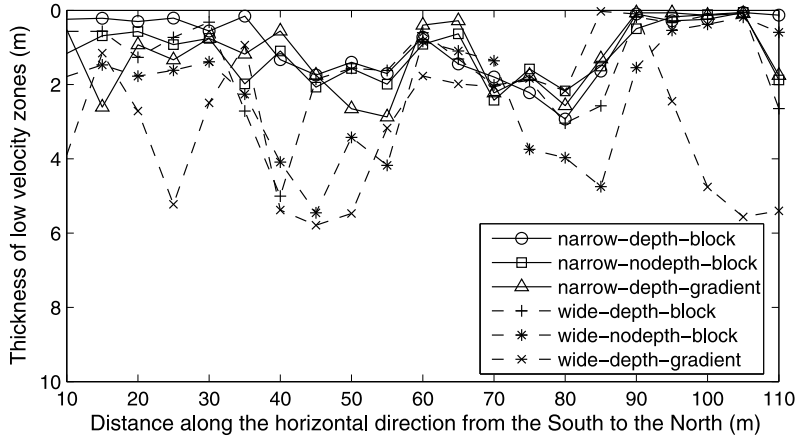

Figure 9. Effects of prior information on the estimates of thickness for the NT2 data set.

within a horizontal distance of less than 5 m from the survey line. However, among those wells, some are only a few feet deep and thus provide no information on the interface location and geometry of fast flow pathways. As shown in Figure 2, seven wellbores (i.e., FW007–FW010, FW112, FW117, and TPB25) are more than 10 ft deep and provide depth of refusal information, which may reflect information about rock competency and thus can be used to build prior models.

[51] Some deeper wells exist near the middle of the seismic survey line (Area 3 in Figure 2), within a vertical distance of about 10 m from the S3 survey line. At those wells, flowmeter and slug test data are available, and crosswell seismic data were also collected and inverted using joint stochastic methods for aquifer zonation [*Chen et al.*, 2006] as was described in section 2.3. The inversion results showed that there is a discrete, high hydraulic conductivity zone at depth, which has a thickness ranging from 0.5 to 3 m. Compared to large-scale (order of 100 m) seismic refraction surveys, the crosswell seismic surveys provide information of the aquifer in the small scale (~10 m). Such localized information also provides constraints on the interface and thickness of aquifer at the location near the middle of the survey line. The third type of information available near the S3 pond survey line is the inversion results using conventional methods [*Watson et al.*, 2005]. We build a background block velocity model using such information, which is the same as given in section 4 for synthetic case studies for joint inversion.

5.2.3. Inversion of Seismic Refraction Data

[52] Figure 10 shows a 2-D seismic velocity calculated from the estimated medians of unknown parameters, which can be explained as an average of many individual realizations. Figure 10 shows the existence of a low-velocity zone in the middle of the profile from x = 40 m to x = 90 m at depth around 15 m. These results are consistent with the findings by *Sheehan et al.* [2005b] in the previous studies at the site, who first suggested the presence of a low seismic velocity zone at the ORNL. The jointly inverted data clearly delineate the geometry of the low-velocity zone.

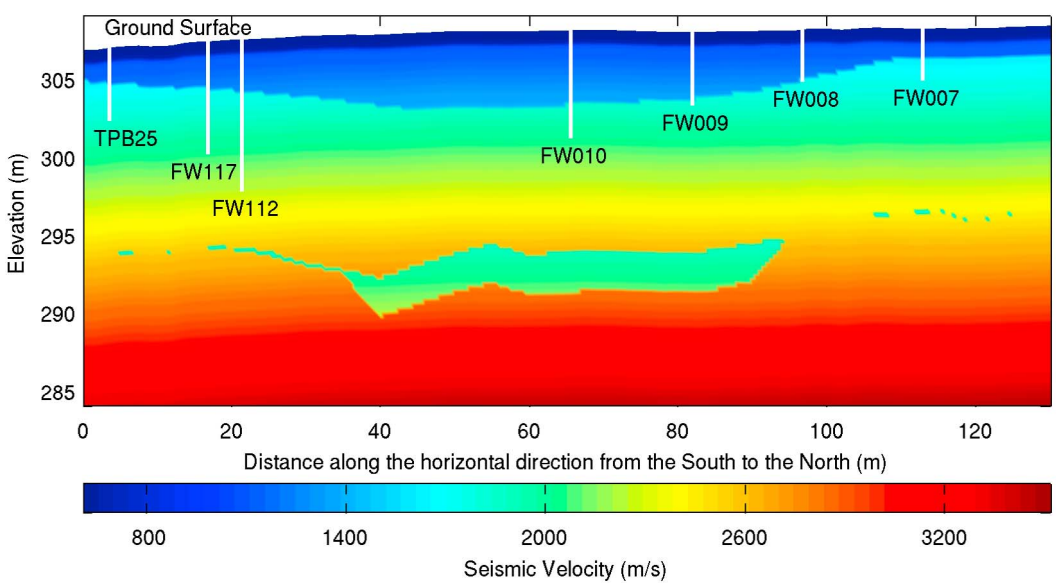

Figure 10. Estimated velocity model along the S3 pond profile.

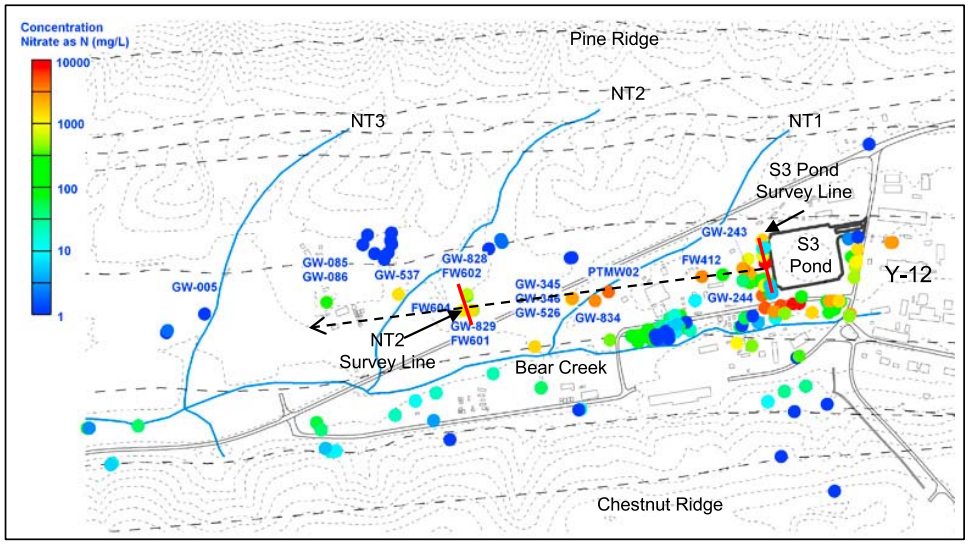

Figure 11. Distribution of the nitrate plume at the Oak Ridge National Laboratory site.

[53] Similar to the case of NT2 inversion, the estimated results shown in Figure 10 are also subject to uncertainty because of the specification of the background block velocity model and the prior distributions of unknown parameters. However, at the Oak Ridge site, given the fact that we have collected extensive information (see Figure 2) near the S3 pond, we think that our background model and prior distributions are reasonable.

5.3. Joint Interpretation of Seismic Refraction Inversion Results

[54] Inversion of seismic refraction data at the NT2 and S3 pond survey profiles suggests the presence of discrete low-velocity zones situated below the saprolite and above the bedrock. Comparison of the inversion results (Figures 7 and 10) suggests that these zones are parallel to the strike and about 40 m in length and 15 m in depth. Since the low-velocity zones are located at a similar location, this implies that the low-velocity zone is laterally persistent for at least 620 m from the source region down gradient. Our previous seismic research at the site (Area 3 in Figure 2) demonstrated a good correlation between low velocity and high hydraulic conductivity zonation at the local scale [*Chen et et al.*, 2006]. Additionally, the location of the laterally extensive zone is aligned well with the distribution of high concentrations of the nitrate plume (see black arrow in Figure 11). Together, these observations suggest that the interpreted low-velocity zone may serve as a preferential pathway at this study site.

6. Discussion and Conclusions

[55] We developed a Bayesian model to invert seismic refraction data with depth constraints offered by wellbore data sets using finite difference numerical forward modeling of seismic first arrival times and MCMC sampling methods. We applied the method to both synthetic and field seismic refraction data sets. Synthetic studies show that the developed method can be used to estimate the interface and geometry of low-velocity zones under suitable constraints. The synthetic studies also show that the depth constraints were important for reducing uncertainty in estimation of interface and geometry of aquifers. With the good depth constraints, we can reduce the uncertainty significantly.

[56] We applied the developed method to two sets of field data collected from the ORNL IFC site: one located near the plume source zone and the other located 620 m down gradient. Case studies on the NT2 survey profile show that estimation of aquifer geometry using field data is more challenging because many factors may affect the accuracy of inversion results, such as prior ranges of interface locations and thickness of low-velocity zones. In that case, depth constraints from wellbores become more important and the use of wellbore information can help to significantly improve the accuracy of the estimation.

[57] Although the data set near the S3 pond (i.e., source area) had much more prior information from wellbores relative to the down-gradient data set, both inversion revealed a similar subsurface geometry. The inversion results show that there is likely a low-velocity zone located at the same approximate depth below ground surface on the two profiles, which implies that the low-velocity zone is laterally persistent for 620 m from the source region down gradient. The zone is fairly extensive, with an average thickness of about 1.5 m and width of about 40 m. Our previous seismic research at the site demonstrated a good correlation between low velocity and high hydraulic conductivity zonation at the local scale. Additionally, the location of the laterally extensive zone is surprisingly coincident with the (narrow and long) plume distribution. These multiple lines of evidence all suggest that the geophysically estimated subsurface geometry, which contains a preferential flow path, plays a key role in transport at the ORNL contaminated site. Future efforts will include additional acquisition and inversion of surface seismic refraction data at the study site, validation of the interpretation through deep wellbore drilling, and the use of the geophysically obtained geometry to parameterize a site-wide flow and transport model.

[58] The current Bayesian method is computationally intensive. For instance, for a problem, such as inversion of seismic data along the NT2 profile with 55 sources, we need

to run multiple chains for several days on a cluster with 55 CPUs, each CPU having a speed of 3.6 GHz. However, because of the use of the staggered-grid finite difference numerical algorithm, we are able to calculate seismic refraction arrival time under complex subsurface environments with high accuracy. With ever expanding computing power, especially parallel computing techniques, computational limitations are expected to continue to diminish.

[59] The developed Bayesian model is advantageous because it allows us to parameterize the inverse problem in terms of aquifer geometry (rather than seismic velocity, as is the common approach), to simultaneously honor spatially extensive (yet indirect) geophysical data and sparser but direct wellbore "point" measurements, and to estimate associated uncertainty of the estimated interface locations. This method should be applicable in other shallow subsurface sites where both surface seismic refraction and wellbore data sets are available. Development of this methodology opens the door for quantification of aquifer geometries over spatial scales relevant to water resources and contaminant remediation efforts.

[60] **Acknowledgments.** Funding for this study was provided by the U.S. Department of Energy, Biological and Environmental Research Program as part of the Oak Ridge National Laboratory Integrated Field Research Center project. We wish to thank Jacob Sheehan from Battelle for providing seismic refraction data along the S3 survey profile and for offering constructive comments on this study. We thank David F. Aldridge from Sandia National Laboratory who suggested and demonstrated a validity of using the Delta function source waveform for numerical modeling, Guping Tang from the Oak Ridge National Laboratory for providing the figure of nitrate plume distribution, and three anonymous reviewers and the Editors for helpful and insightful suggestions.

References

Addison, A. D., M. G. Waddell, C. G. Knapp, D. T. Brantley, and J. M. Shafer (2009), Developing a robust geologic conceptual model using pseudo 3-D *P* wave seismic reflection data, *Environ. Geosci.*, *16*(1), 41–56.

Anderson, M. P. (2007), Introducing groundwater physics, *Phys. Today*, 42–47.

Butler, J. J., E. J. Garnett, and J. M. Healey (2005), Analysis of slug tests in formations of high hydraulic conductivity, *Ground Water*, *41*(5), 620–631.

Chen, J., S. Hubbard, J. Peterson, K. Williams, M. Fienen, P. Jardine, and D. Watson (2006), Development of a joint hydrogeophysical inversion approach and application to a contaminated fractured aquifer, *Water Resour. Res.*, *42*, W06425, doi:10.1029/2005WR004694.

Chen, J., A. Kemna, and S. S. Hubbard (2008), A comparison between Gauss-Newton and Markov Chain Monte Carlo based methods for inverting spectral induced polarization data for Cole-Cole parameters, *Geophysics*, *73*(6), F247–F259.

Chen, J., S. Hubbard, K. Williams, S. Pride, L. Li, C. Steefel, and L. Slater (2009), A state-space Bayesian framework for estimating biogeochemical transformations using time-lapse geophysical data, *Water Resour. Res.*, *45*, W08420, doi:10.1029/2008WR007698.

DeMarsily, G., F. Delay, J. Goncalves, P. renard, V. Teles, and S. Violette (2005), Dealing with spatial heterogeneity, *Hydrogeol. J.*, *13*, 161–183.

Falgas, E., J. Ledo, A. Marcuello, and P. Queralt (2009), Monitoring freshwater-seawater interface dynamics with audiomagnetotellurics data, *Near Surface Geophys.*, *7*(5–6), 391–399.

Fienen, M. N. (2005), The three-point problem, vector analysis and extension to the n-point problem, *J. Geosci. Educ.*, *53*(3), 257–262.

Fienen, M., P. Kitanidis, D. Watson, and P. Jardine (2004), An application of Bayesian inverse methods to vertical deconvolution of hydraulic conductivity in a heterogeneous aquifer at Oak Ridge National Laboratory, *Math. Geol.*, *36*(1), 101–126.

Fienen, M., J. Luo, and P. Kitanidis (2006), A Bayesian geostatistical transfer function approach to tracer test analysis, *Water Resour. Res.*, *42*, W07426, doi:10.1029/2005WR004576.

Gaines, D., G. Baker, S. Hubbard, D. Watson, and P. Jardine (2008), Enhancing monitoring of recharge-related environmental remediation processes using time-lapse seismic refraction, *Eos. Trans., AGU*, *89*(53), Fall Meet. Suppl., Abstract NS41A-04.

Gelman, A. (2006), Prior distributions for variance parameters in hierarchical models, *Bayesian Anal.*, *1*(3), 515–533.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, *7*, 457–511.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996), *Markov Chain Monte Carlo in Practice*, CRC Press, Boca Raton, Fla.

Haeni, F. P. (1986), Application of seismic refraction methods in groundwater modeling studies in New England, *Geophysics*, *51*(2), 236–249.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*, 97–109.

Hubbard, S., and N. Linde (2010), Hydrogeophysics, in *Treatise on Water Science*, vol. 2, *Hydrology*, edited by S. Uhlenbrook, Elsevier, Oxford, U. K., in press.

Hyndman, D. W., and J. M. Harris (1996), Traveltime inversion for the geometry of aquifer lithologies, *Geophysics*, *61*(6), 1728–1737.

Jardine, P. M., T. L. Mehlhorn, I. L. Larsen, W. B. Bailey, S. C. Brooks, Y. Roh, and J. P. Gwo (2002), Influence of hydrological and geochemical processes on the transport of chelated metals and chromate in fractured shale bedrock, *J. Contam. Hydrol.*, *55*(1–2), 137–159.

Joarder, A. H., and S. E. Ahmed (1996), Estimation of the characteristic roots of the scale matrix, *Metrika*, *44*, 259–267.

Kelejian, H. H., and I. R. Prucha (1985), Independent or uncorrelated disturbances in linear regression: An illustration of the difference, *Ecol. Lett.*, *19*, 35–38.

Korneev, V., A. Bakulin, and J. Lopez (2008), Imaging and monitoring with virtual sources on synthetic 3-D data set from Middle East, paper presented at 78th Annual International Meeting, Soc. of Explor. Geophys., Las Vegas, Nev.

Kotz, S., and S. Nadarajah (2004), *Multivariate T Distributions and Their Applications*, Cambridge Univ. Press, New York.

Levander, A. R. (1988), Fourth-order finite difference P-SV seismograms, *Geophysics*, *53*(11), 1425–1436.

Luo, J., O. A. Cirpka, M. N. Fienen, W. Wu, T. L. Mehlhorn, J. Carley, P. M. Jardine, C. S. Criddle, and P. K. Kitanidis (2006), A parametric transfer function concept for analyzing reactive transport in non-uniform flow, *J. Contam. Hydrol.*, *83*(1–2), 27–41.

Marchev, D., and J. P. Hobert (2004), Geometric ergodicity of van Dyk and Meng's algorithm for the multivariate Student's t model, *J. Am. Stat. Assoc.*, *99*(465), 228–238.

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

Miller, C. R., P. S. Routh, T. R. Brosten, and J. P. McNamara (2008), Application of time-lapse ERT imaging to watershed characterization, *Geophysics*, *73*(3), G7–G17.

Neal, R. M. (2003), Slice sampling, *Annu. Stat.*, *31*, 705–767.

Pinheiro, J. C., C. H. Liu, and Y. N. Wu (2001), Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution, *J. Comput. Graphical Stat.*, *10*(2), 249–276.

Rubin, Y., and S. Hubbard (2005), *Hydrogeophysics*, Springer, Netherlands.

Sheehan, J. R., W. E. Doll, and W. A. Mandell (2005a), An evaluation of methods and available software for seismic refraction tomography analysis, *J. Environ. Eng. Geophys.*, *10*(1), 21–34.

Sheehan, J. R., W. E. Doll, D. B. Watson, and W. A. Mandell (2005b), Detecting cavities with seismic refraction tomography: Can it be done?, paper presented at Symposium on the Application of Geophysics to Engineering and Environmental Problems (SAGEEP), Environ. and Eng. Geophys. Soc., Atlanta, Ga.

U.S. Department of Energy (USDOE) (1995), Fiscal year 1995 well installation program summary, Y-12 plant, *Rep. DOE/Y/SUB/95-99069C(Y24)/1*, Oak Ridge, Tennessee.

U.S. Department of Energy (USDOE) (1997), Report on the remedial investigation of Bear Creek Valley at the Oak Ridge Y-12 Plant, Oak Ridge, Tennessee, *Rep. DOE/OR/01–1455/V1&D2*, Oak Ridge, Tenn.

Watson, D. B., J. E. Kostka, M. W. Fields, and P. M. Jardine (2004), The oak ridge field research center conceptual model, research report, Oak Ridge Natl. Lab., U.S. Dep. of Energy, Oak Ridge, Tenn.

Watson, D. B., W. E. Doll, T. J. Gamey, J. R. Sheehan, and P. M. Jardine (2005), Plume and lithologic profiling with surface resistivity and seismic tomography, *Ground Water*, *43*(2), 169–177.

Wu, W., et al. (2006), Pilot-scale bioremedation of uranium in a highly contaminated aquifer: I. Conditioning of a treatment zone, *Environ. Sci. Technol.*, *40*(12), 3978–3985.

G. Baker and D. Gaines, Department of Earth and Planetary Sciences, University of Tennessee, 1412 Circle Dr., Knoxville, TN 37996-1410, USA.

J. Chen, S. S. Hubbard, and V. Korneev, Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. (jchen@lbl.gov)

D. Watson, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.