Geochemical Characterization Using Geophysical Data and Markov Chain Monte Carlo Methods: A Case Study at the South Oyster Bacterial Transport Site in Virginia

Jinsong Chen, Susan Hubbard, Yoram Rubin, Chris Murray, Eric Roden,

and Ernest Majer

J. Chen (jchen@lbl.gov), S. Hubbard, and E. Majer, Lawrence Berkeley National Laboratory, MS 90-1116, 1 Cyclotron Road, Berkeley, CA 94720, USA

Y. Rubin, Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720, USA

C. Murray, Pacific Northwest National Laboratory, P.O. Box 999, MS K6-81, Richmond, WA 99352, USA

E. Roden, Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 354872, USA

Abstract. The spatial distribution of field-scale geochemical parameters, such as extractable Fe(II) and Fe(III), influences microbial processes and thus the efficacy of bioremediation. Because traditional characterization of those parameters is invasive and laborious, it is rarely performed sufficiently at the field-scale. Since both geochemical and geophysical parameters often correlate to some common physical properties (such as lithofacies), we investigated the utility of tomographic radar attenuation data for improving estimation of geochemical parameters using a Markov Chain Monte Carlo (MCMC) approach. The data used in this study included physical, geophysical, and geochemical measurements collected in and between several boreholes at the DOE South Oyster Bacterial Transport Site in Virginia. Results show that geophysical data, constrained by physical data, provided field-scale information about extractable Fe(II) and Fe(III) in a minimally invasive manner and with a resolution unparalleled by other geochemical characterization methods. This study presents our estimation framework for estimating Fe(II) and Fe(III), and its application to a specific site. Our hypothesis—that geochemical parameters and geophysical attributes can be linked through their mutual dependence on physical properties—should be applicable for estimating other geochemical parameters at other sites.

1. Introduction

Field-scale bacterial transport in the subsurface plays an important role in the protection of groundwater supplies from contamination by pathogens and in the bioremediation of various metal and radionuclide contaminants [*Chapelle*, 2001]. Many studies have shown that both *in situ* physical and chemical heterogeneity control by diverse mechanisms field-scale bacterial movements in saturated aquifers [*DeFlaun et al.*, 2001; *Chen et al.*, 2002]. For those bacteria with low or neutral surface charges, the adhesion of bacteria to sediment is primarily determined by physical heterogeneity, specifically by grain size and pore-throat size distributions [*Dong et al.*, 2002]. However, for those bacterial strains bearing high negative surface charges and traveling through the heterogeneous subsurface, chemical heterogeneity becomes important because of electrostatic interactions. For example, in slightly acidic groundwater, quartz grains have negative charges and metal oxyhydroxides have positive charges, and thus bacterial strains with negative charges will be repelled by quartz grains that lack metal oxyhydroxide coatings, but will be attracted to the surfaces of quartz grains that are dominated by metal oxyhydroxides.

In aquifers containing Fe(III) oxides, the activity of dissimilatory iron-reducing bacteria (DIRB) has the potential to influence bacterial transport processes [*Caccavo et al.*, 1997]. The adsorption or precipitation of Fe(II) produced by Fe(III) reduction on Fe(III) oxides and on the surface of DIRB may reduce the adhesion of DIRB to sediment grains by reducing the bio-accessible Fe(III) surface areas. Fe(III) oxide reduction could also locally increase the pH of groundwater and thereby promote the desorption of DIRB by

reducing the positive surface charges of the metal oxyhydroxide minerals. In addition, the

desorption of daughter cells could increase net transport rates [*Roden et al.*, 2000].

Despite the importance of field-scale physical and geochemical heterogeneity for micro-

bial processes, characterization of physical and geochemical parameters remains a daunt-

ing task. Traditional methods for characterizing those parameters, such as laboratory

measurements of cores collected from boreholes, are invasive and laborious, and thus

are rarely performed sufficiently at the field-scale. In addition, core data are point-scale

measurements, which are usually insufficient for characterizing variability at the field-

scale where bioremediation occurs. In our previous study, we have successfully estimated

the high-resolution spatial distribution of field-scale hydraulic conductivity at the DOE

South Oyster Bacterial Transport Site in Virginia using geophysical data [*Chen et al.*,

2001; *Hubbard et al.*, 2001]. Comparison of field tracer experiment measurements and nu-

merical modeling predictions, based on geophysically obtained estimates, suggested that

the estimated hydraulic conductivity values provided information at a scale and resolution

that greatly improved the prediction of field-scale solute transport [*Scheibe and Chien*,

2003] and helped to improve our understanding of bacterial transport and attachment

[*Mailloux et al.*, 2003].

Motivated by the successful estimation of hydrogeological properties using geophysical

data, we explore the use of the same noninvasive and cost-effective approach for char-

acterization of geochemical heterogeneity, using data collected from the South Oyster

site. To our knowledge, this is the first effort to estimate field-scale geochemical param-

eters using geophysical data. We focus on estimation of solid-phase Fe(II) and Fe(III)

concentrations along a two-dimensional cross section by exploiting the site-specific mu-

tual dependence of crosshole radar attenuation and Fe(II) and Fe(III) concentrations on lithofacies. The geochemical parameters are known at the borehole locations, but are unknown between the boreholes. The two-dimensional spatial distributions of Fe(II) and Fe(III) concentrations are estimated from tomographic radar attenuation data using a Markov Chain Monte Carlo (MCMC) approach. These estimates should provide useful information for predicting field-scale microbial Fe(III) oxide reduction potential, and for understanding the field-scale bacterial transport experiments carried out at the South Oyster Focus Area [*Murray et al.*, 2001]. The approach developed should also be useful for estimating geochemical parameters at other locations where site-specific relationships between geophysical, physical, and geochemical parameters can be exploited.

The remainder of this paper is organized as follows. Section 2 describes site information and available data, and Section 3 describes the development of the statistical estimation method. Estimation results and cross-validation of the method are given in Section 4.

2. Site Information and Data

2.1. Site Information

The South Oyster site is located on the southern Delmarva Peninsula, which is situated on the eastern coast of the United States between Chesapeake Bay and the Atlantic Ocean. The surficial unconfined aquifer underlying the study area consists of unconsolidated to weakly cemented, well-sorted, medium- to fine-grained sands and pebbly sands, with the water table located $\sim$2 m below the ground surface. A field-scale experiment within the uncontaminated aquifer at the site was undertaken by a multidisciplinary research team to evaluate the importance of chemical and physical heterogeneity in controlling bacteria transport at the site [*DeFlaun et al.*, 2001; *Johnson et al.*, 2001]. Two focus areas exist

within the South Oyster Site: the Narrow Channel Focus Area and the South Oyster
Focus Area (SOFA). Forced gradient chemical and bacterial tracer test experiments were
performed at both focus areas.

This study focuses on data collected within a saturated aquifer at a site located $\sim$15
m along the geological strike from the SOFA transport site (Figure 1). Our goal was
to investigate the spatial heterogeneity of geochemical parameters as an analogue to the
immediately adjacent SOFA site. The available data included laboratory measurements
of physical, geological, and geochemical parameters from the cores retrieved from wells
D1, D2, and D3. The samples were taken between depths of 2.4 m and 8.7 m below
the ground surface, with a sampling interval of 0.15 m to 0.30 m [*Johnson et al.*, 2001].
The data also included high-resolution ground-penetrating radar (GPR) and seismic to-
mograms acquired along the cross sections between wells D1-D2 and D2-D3 (Figure 1).
We performed multivariate data analysis for all the datasets and found lithofacies and
GPR attenuation to be most informative for estimating extractable Fe(II) and Fe(III). As
such, in the following, we describe only the lithofacies, GPR attenuation, and solid-phase
Fe(II) and Fe(III) data, and explore their cross correlations.

2.2. Data Analysis

2.2.1. Lithofacies

Borehole lithofacies categorization was performed for the study during the logging and
core-sampling process, based on the visual grain-size estimation using a comparator chart
and on the soil color and texture description. In the original data, four lithofacies cat-
egories were identified: peat, mud, muddy-sand, and sand. Based on cluster analysis of
physical properties, we reduced the classifications from four to two categories, by grouping

peat and mud together, and referring to it as mud, and by grouping muddy-sand and sand together, and referring to it as sand.

The spatial structure of lithofacies was investigated using variograms, which are defined as the average squared difference of a quantity at two locations as a function of the measurement separation distance [*Rubin*, 2003]. We performed variogram analysis along the vertical direction by first computing the experimental variogram using the coded lithofacies data (sand=1 and mud=0) collected from wells D1, D2, and D3, and then fitting it using an exponential model with an integral scale of 0.5 m (Figure 2). Similarly, we attempted to fit a model to the experimental variogram in the horizontal direction. But with data from just three wells, no model could be reliably fit to the horizontal variogram, due to the sparse sample density in that direction. In this study, however, we assumed an anisotropy ratio of 5, borrowed from the nearby Narrow Channel Focus Area [*Hubbard et al.*, 2001], and used an integral scale of 2.5 m in the horizontal direction.

2.2.2. GPR Attenuation

GPR is a geophysical tool that has become increasingly popular as researchers across a variety of disciplines strive to better understand near-surface conditions. GPR uses electromagnetic energy at frequencies of 50–1500 MHz to probe the subsurface. At the frequencies used and under the low-loss conditions, the electromagnetic signals propagate primarily as waves, and the GPR attributes are functions of dielectric constant and electrical conductivity of the medium [*Davis and Annan*, 1989]. Radar velocities are influenced by the dielectric constant, which is sensitive to water content and porosity, and thus have been used for mapping soil water content in unsaturated systems [*Hubbard et al.*, 1997; *Binley and Beven*, 2003] and hydraulic conductivity in saturated systems [*Chen et al.*,

2001; *Hubbard et al.*, 2001]. Radar amplitudes (or attenuation), however, are influenced by both the dielectric constant and the electric conductivity of the medium [*Davis and Annan*, 1989]. In saturated aquifers, such as the one considered here, radar attenuation usually is dominated by electrical conductivity, and high electrical conductivity often leads to high radar attenuation. Because fine-grained soils (such as clay and silt) typically have much higher electrical conductivity (2–1000 mS/m for clay and 1–100 mS/m for silt) than that of coarse-grained soils such as sand (0.1–1 mS/m), higher radar attenuation is often associated with high clay content and silt fraction. Although the salinity of pore fluid may also influence GPR attenuation, it does not appear to be correlated with GPR attenuation in this study. Consequently, we hypothesize that both GPR attenuation and geochemical parameters are predominantly influenced by lithology, and focus herein on the use of GPR attenuation to estimate geochemical parameters by exploring the lithology link.

GPR tomographic data were collected along the cross sections between wells D1 and D2 and between wells D2 and D3, using borehole antennas having a central frequency of 100MHz. A typical crosshole tomographic geometry consists of two vertical boreholes separated by an interwell region of interest. Direct energy from a transmitting antenna in one borehole is recorded by the receiving antenna located in the other borehole. By moving the transmitting and receiving antennas in the boreholes, many ray paths can be recorded, which can be inverted to provide a tomographic image of the region between the boreholes [*Peterson et al.*, 1985]. The recorded data included the direct electromagnetic wave travel time from the transmitter to the receiver passing through the crosshole region and the amplitude of the direct arrivals. After dividing the interwell area into a grid of pixels (0.25 m x 0.25 m), an inversion algorithm was used to transform the recorded travel

time and amplitude information into estimates of the GPR velocity and attenuation at each pixel, following *Peterson* [2001]. Figure 3(a) shows a contoured GPR attenuation tomogram along the cross section between wells D1 and D3. GPR attenuation data "overlap" with borehole data at or near the borehole locations. These collocated data were used to develop site-specific relations between geophysical attributes and geochemical parameters. The developed relationships were then used in conjunction with the GPR tomographic data (Figure 3(a)) to extrapolate borehole geochemical measurements into the interwell areas.

Figure 4 shows the relationship between the sand and mud lithofacies and the natural logarithmic GPR attenuation. As expected from the previous discussions, the GPR amplitudes were more attenuated when passing through mud than through sand. Based on data collected at the three wells, the mean logarithmic GPR attenuation of sand was -0.74 1/m with a standard deviation of 0.23 1/m, while the mean logarithmic GPR attenuation of mud was -0.33 1/m with a standard deviation of 0.36 1/m.

2.2.3. Extractable Fe(II)

Extractable Fe(II) was measured by leaching triplicate 0.5–1.0 g subsamples, obtained from each depth interval for wells D1, D2, and D3, with 0.5 M HCl for one hour. The Fe(II) versus Fe(III) content of the extracts was determined using Ferrozine [*Roden and Lovely*, 1993]. The exact nature of the Fe(II)-bearing phases leached by the 0.5 M HCl is unknown, but may generally include Fe(II) from native iron-bearing minerals in the formation (e.g., carbonates or silicates), as well as Fe(II) phases produced from bacterial Fe(III) oxide reduction (e.g., siderite or Fe(II) sorbed to residual Fe(III) oxides or other mineral surfaces).

Figure 5 shows the histograms of the natural logarithmic extractable Fe(II) for sand and mud, based on data collected at wells D1, D2, and D3. The figure suggests that the distribution of logarithmic Fe(II) is symmetrical around the corresponding means for both sand and mud. However, the logarithmic Fe(II) of sand has a much larger logarithmic range (from -4 to 4) than that of mud (from 2 to 4). The logarithmic Fe(II) concentration also depends on lithofacies; mud has much higher concentrations of extractable Fe(II) than sand. The mean natural logarithmic Fe(II) of mud is 3.12 μmol/cc, with a standard deviation of 0.58 μmol/cc, whereas the mean natural logarithmic Fe(II) of sand is only 0.25 μmol/cc, with a standard deviation of 1.66 μmol/cc. The logarithmic Fe(II) distribution as a function of lithofacies is reasonable, because Fe(II) is usually sequestered in fine-grain sediments such as silt and clay [*Chapelle*, 2001], and our mud lithofacies included both silt and clay components.

As both logarithmic Fe(II) and logarithmic GPR attenuation display a correlation with lithofacies, we have physical justification for investigating the link between Fe(II) and GPR attenuation. The connection, however, may be affected by organic matter, grain size, porosity, and other physical parameters. Figure 6 shows a cross-plot of logarithmic Fe(II) versus logarithmic GPR attenuation based on data at wells D1, D2, and D3, where the circles represent sand and the triangles represent mud. This figure reveals that for both sand and mud, logarithmic Fe(II) linearly increases with increasing logarithmic GPR attenuation. For mud, however, the increase is not as apparent as for sand.

2.2.4. Extractable Fe(III)

Extractable Fe(III), including amorphous and crystalline Fe(III) oxides, was determined using a citrate-dithionite (CD) reagent (pH 4.8) [*Canfield*, 1989]. The 0.5 M HCl ex-

tractable Fe(II) content of parallel sediment extracts (see above) was subtracted from total CD extractable Fe to estimate Fe(III) abundance, as the CD reagent typically recovers the majority of solid-phase Fe(II) compounds.

Figure 7 shows the histograms of natural logarithmic Fe(III) for sand and mud. The logarithmic Fe(III) of mud seems to be symmetric around its mean, and the logarithmic Fe(III) of sand is slightly skewed towards the larger values. Although the mean values of the logarithmic Fe(III) for sand (2.12 μmol/cc) and for mud (-0.60 μmol/cc) are quite different, their ranges overlap considerably. In addition, the logarithmic Fe(III) has weak linear correlations with both logarithmic GPR attenuation ($R^2 = 0.05$) and logarithmic Fe(II) ($R^2 = 0.15$), but has a relatively good correlation with depth ($R^2 = 0.37$).

3. Statistical Model

This section describes the statistical model for estimating the spatial distribution of Fe(II) and Fe(III) concentrations along the cross section between wells D1 and D3. Within the statistical framework, unknown Fe(II) and Fe(III) concentrations and lithofacies at each location between the wells were considered as random variables, which are fully characterized by the joint conditional probability function given GPR tomographic data and borehole lithofacies logs. Estimation of those variables from the joint probability function was obtained using an MCMC approach, which will be described in the next section.

3.1. Model Setup

The developed statistical model is intended to meet conditions that are specific to our study site, even though its underlying concept is quite general and can be applied to other

sites. Figure 8 shows a discretization of the cross section (12 m x 6 m) between wells D1

and D3. There are a total number of 1225 pixels, each of which has dimensions of 0.25 m

x 0.25 m. Lithofacies, Fe(II), and Fe(III) concentrations are known at pixels along wells

D1, D2, and D3, but are unknown at pixels located between the wells. GPR attenuation

is considered to be known at all locations along the two-dimensional transect. Our goal

was to estimate all the unknown parameters given data available at the three wells and

along the cross section.

The estimation problem can be addressed in a stochastic framework, using a joint

conditional probability function. Let L_i be the indicator random variable representing

lithofacies at pixel-i, 1 for sand and 0 for mud, $i \in N$, where N is the index set of

all pixels for which lithofacies and Fe(II) and Fe(III) concentrations are unknown. Let

X_i and Y_i denote the unknown logarithmic Fe(II) and Fe(III) concentrations at pixel-i.

(For convenience, we shall refer to probability functions and probability density functions

as probability distributions in the later text.) Let a_i denote the known logarithmic GPR

attenuation at pixel-i and l_w denote the known lithofacies at the borehole pixels. Following

the convention suggested by *Gelfand and Smith* [1990], we use square brackets to denote

probability distributions. Consequently, the joint conditional distribution is given by:

$$[\{L_i\}, \{X_i\}, \{Y_i\}|\{a_i\}, \{l_w\}], \qquad i \in N \tag{1}$$

where { } denotes a set that includes all possible values of the variable. Our objective

is to obtain the marginal distribution functions of lithofacies and logarithmic Fe(II) and

Fe(III) at each pixel from the joint conditional distribution.

3.2. Local Conditional Distribution

We can simplify Equation 1 by considering only local dependence between logarithmic GPR attenuation, lithofacies, and logarithmic Fe(II) and Fe(III), using data collected at boreholes. We used the stepwise deletion technique [*Chen et al.*, 2001] to obtain the best regression models for logarithmic Fe(II) and Fe(III) estimation. We also fitted the logarithmic GPR attenuation as a linear function of the lithofacies indicator. Figure 9 is a schematic diagram of the local conditional relationships among lithofacies, logarithmic GPR attenuation, and logarithmic Fe(II) and Fe(III). We found that logarithmic Fe(II) concentrations depend on the co-located lithofacies and logarithmic GPR attenuation, and logarithmic Fe(III) concentrations depend on the co-located logarithmic Fe(II) concentrations and depth. Based on data analysis, we assumed each of the local conditional distributions to be Gaussian, as follows:

$$[a_i|L_i] \sim Normal(u_1 + u_2 L_i, \tau_1), \tag{2}$$

$$[X_i|a_i, L_i] \sim Normal(v_1 + v_2 L_i + v_3 a_i + v_4 L_i a_i, \tau_2), \tag{3}$$

$$[Y_i|X_i] \sim Normal(r_1 + r_2 X_i + r_3 d_i, \tau_3), \tag{4}$$

where d_i is depth at pixel-i, u_1, u_2, v_1, v_2, v_3, v_4, r_1, r_2, and r_3 are regression coefficients, and τ_1, τ_2, and τ_3 are the inverse variances of the conditional distributions of a_i, X_i, and Y_i, respectively. The coefficients and inverse variances, obtained from data analysis, are given in Table 1.

3.3. Joint Conditional Distribution

We can expand Equation 1 into several terms using Bayes' theorem [*Bernardo and Smith*, 1994] and subsequently simplify them using the local conditional distributions

given in Equation 2–4. We also assume that the values of logarithmic GPR attenuation and logarithmic Fe(II) and Fe(III) at each pixel are conditionally independent of the ones at other pixels. By incorporating Equations 2–4 into Equation 1, we obtain the form of the joint distribution as follows:

$$[\{L_i\}, \{X_i\}, \{Y_i\}|\{a_i\}, \{l_w\}] \propto [\{L_i\}, \{l_w\}] \cdot \prod_{i \in N} \{[Y_i|X_i][X_i|a_i, L_i][a_i|L_i]\}, \qquad (5)$$

where "$\propto$" represents "proportional to," which ignores the normalizing constant at the right side of Equation 5.

4. MCMC Sampling Method

This section outlines the method for obtaining estimates of Fe(II) and Fe(III) concentrations from the joint conditional distribution shown in Equation 5. Since conventional analytical approaches are not feasible because of the large number of unknown variables involved, we used instead an MCMC method, which has recently emerged as a powerful approach for solving complex statistical problems involving a large number of dependent, random variables. The MCMC method provides an efficient way to draw samples of unknown variables from their joint distributions by running a constructed Markov chain [*Gilks et al.*, 1996]. Using those samples, we can obtain the mean, variance, predictive intervals, and even probability function for each variable. Several applications of MCMC methods to hydrogeology have been found, including those documented in *Bosch* [1999] and *Michalak and Kitanidis* [2003].

We estimated Fe(II) and Fe(III) concentrations by following four basic steps:

1. Derive the conditional distribution for each unknown variable given all the data and all other variables, which is referred to as the full conditional distribution of the variable.

2. Sequentially draw samples from each of the conditional distributions.

3. Monitor convergence using the method developed by *Gelman and Rubin* [1992].

4. Make inferences about each variable using the generated samples.

The first three steps of the process are described below, and the final stage is given in Section 5.

4.1. Deriving Full Conditional Distributions

Full conditional distributions play a prominent role in applications of MCMC methods [*Gilks et al.*, 1996]. They are typically not the same as the local conditional distributions that are conditioned to all the data and subsets of all other variables. However, we can derive full conditional distributions by using the local conditional distributions. Theoretically, the full conditional distribution of each variable is proportional to the joint conditional distribution shown in Equation 5. Because MCMC methods require no information about normalizing constants, we can obtain the full conditional distribution of each variable by omitting the terms at the right side of Equation 5 not directly related to the variable.

4.1.1. Full conditional probability function of Y_i

Let $[Y_i|\cdot]$ denote the full conditional probability density function (PDF) of Y_i given all the data and other variables. After omitting all the terms in Equation 5 that are not directly related to Y_i, we obtain:

$$[Y_i|\cdot] \propto [Y_i|X_i]. \tag{6}$$

This is the same as the local conditional pdf of Y_i given X_i that was shown in Equation 4.

4.1.2. Full conditional probability function of X_i

Similarly, by omitting all the terms in Equation 5 not directly related to X_i, we obtain

the full conditional distribution of X_i as follows:

$$[X_i|\cdot] \propto [Y_i|X_i][X_i|a_i, L_i]. \tag{7}$$

By substituting $[X_i|a_i, L_i]$ and $[Y_i|X_i]$ with Equations 3 and 4, respectively, and after some

derivations (see Appendix A), we obtain:

$$[X_i|\cdot] \sim Normal(\frac{\tau_2(v_1 + v_2 L_i + v_3 a_i + v_4 L_i a_i) + r_2 \tau_3(Y_i - r_1 - r_3 d_i)}{\tau_2 + r_2^2 \tau_3}, \tau_2 + r_2^2 \tau_3) \tag{8}$$

4.1.3. Full conditional probability function of L_i

The MCMC method allows us to consider lithofacies as a spatially correlated random

variable. As a result, we can incorporate spatial correlation and borehole lithofacies

measurements into the estimation model. Similar to the derivation of $[Y_i|\cdot]$ and $[X_i|\cdot]$, we

omit those terms in Equation 5 that are not directly related to L_i and obtain:

$$[L_i|\cdot] \propto [a_i|L_i] \cdot [X_i|a_i, L_i] \cdot [L_i|\{L_j, j \neq i\}, \{l_w\}], \tag{9}$$

where $\{L_j, j \neq i\}$ is the set including lithofacies at all pixels except pixel-i. It is convenient

and reasonable to assume that lithofacies L_i depends only on the lithofacies at its adjacent

pixels [*Chen and Rubin*, 2003]. Let Set A_i be the index set of the adjacent pixels of pixel-i.

For those pixels not near boreholes, the conditional probability of L_i does not depend on

borehole lithofacies measurements, and thus it is given by

$$[L_i|\{L_j, j \neq i\}, \{l_w\}] = [L_i|L_j, j \in A_i] \sim Bernoulli(p_i^*), \tag{10}$$

where p_i^* is the probability of lithofacies being sand given the lithofacies at its surrounding

pixels, obtained using indicator kriging [*Rubin*, 2003] (see Appendix B). Although for

those pixels near boreholes, the conditional probability of L_i also depends on the borehole

lithofacies measurements $\{l_w\}$, we use a similar method to that as described in Appendix

B to obtain probability p_i^*.

By substituting $[a_i|L_i]$ and $[X_i|a_i, L_i]$ with Equations 2 and 3, respectively, and af-

ter some simplifications (see Appendix C), we obtain the full conditional probability of

lithofacies as follows:

$$[L_i|\cdot] \sim Bernoulli(\frac{p_i^* p_i}{1 - p_i^* + p_i^* p_i}) \tag{11}$$

where

$$p_i = \exp\{\tau_1 u_2(a_i - u_1 - 0.5u_2) + \tau_2(v_2 + v_4 a_i)(X_i - v_1 - 0.5v_2 - v_3 a_i - 0.5v_4 a_i)\}.$$

4.2. Sampling Full Conditional Distributions

The relationships shown in Equations 6, 8, and 11 represent the full conditional distribu-

tions of all variables of interest. The second step of the MCMC method is to sequentially

draw samples from those distributions. Two major algorithms can be used to draw sam-

ples from full conditional distributions: the Gibbs sampler [*Geman and Geman*, 1984] and

the Metropolis-Hastings method [*Metropolis et al.*, 1953; *Hastings*, 1970]. Because the full

conditional distributions of X_i and Y_i are Gaussian and the full conditional distribution

of L_i is Bernoulli, which are easily sampled, we used the Gibbs sampler. The main steps

are given as follows:

1. For each $i \in N$, assign 0 or 1 to L_i and refer to it as $L_i^{(0)}$, and assign real values to

X_i and Y_i and refer to them as $X_i^{(0)}$ and $Y_i^{(0)}$, respectively. Let $k = 1$.

2. Draw a sample from the Gaussian distribution $[Y_i|\cdot]$ (Equation 4) given $X_i^{(k-1)}$ and

refer to it as $Y_i^{(k)}$ for $i \in N$.

3. Draw a sample from the Gaussian distribution $[X_i|\cdot]$ (Equation 8) given $Y_i^{(k)}$ and $L_i^{(k-1)}$, and refer to it as $X_i^{(k)}$ for $i \in N$.

4. Draw a sample from the Bernoulli distribution $[L_i|\cdot]$ (Equation 11) given $X_i^{(k)}$, $\{L_j^{(k)}, j = 1, 2, \cdots, i - 1\}$, and $\{L_j^{(k-1)}, j = i + 1, \cdots, n\}$ and refer to it as $L_i^{(k)}$ for $i \in N$.

5. Let k=k+1. If $k > m$, where m is the maximum number of iterations allowed, stop; otherwise, go to step 2.

4.3. Monitoring Convergence of the Sampling

In the third step of the MCMC method, our goal is to determine the number of iterations needed to obtain samples for inferences using the Gibbs sampler. Samples obtained from the preceding algorithm may not be the samples of their marginal conditional distributions. However, theoretically, after a sufficiently long run (for example, t iterations), referred to as burn-in [*Gilks et al.*, 1996], samples $\{X_i^{(k)}, Y_i^{(k)}, L_i^{(k)} : k = t+1, \cdots, m, i \in N\}$ obtained from the algorithm are approximately samples from their corresponding true marginal conditional distributions [*Gelfand and Smith*, 1990]. In addition, as indicated by the ergodic theorem [*Gilks et al.*, 1996], the mean of any measurable function of those variables obtained using the generated samples after discarding burn-in samples asymptotically converges to its true expectation as $m \to +\infty$.

Many methods exist for monitoring convergence and finding the burn-in, such as the *Gelman and Rubin* [1992], *Geweke* [1992], and *Raftery and Lewis* [1992] methods. The most often used method is the *Gelman and Rubin* method. This method first entails running several Markov chains with very different initial values, followed by calculation of a criterion, referred to as the scale reduction score based on the multiple Markov chains [*Gelman and Rubin*, 1992]. If the scale reduction score is less than 1.2, the Markov chain

is considered to be converged; otherwise, more runs are needed. Using the preceding convergence diagnostics method, we found that a burn-in of 400 realizations is enough for all the unknown variables in this study. After the burn-in stage, we continued to run the chain for another 2,000 runs and used all those samples to make inferences about the unknown variables. The total computing time for running the 2,400 iterations is less than ten minutes on a Pentium-III personal computer.

5. Results and Discussion

At the final stage of the MCMC method, we summarizes the results of Fe(II) and Fe(III) estimation. We will first present the two-dimensional images of the mean logarithmic Fe(II) and Fe(III) concentrations, obtained using the previous MCMC method. We will then show the results of cross-validation analysis that demonstrated the effectiveness of the developed statistical model. Finally, we will provide a short discussion of our methodology.

5.1. Fe(II) and Fe(III) Estimation

Using the lithofacies and Fe(II) and Fe(III) data at wells D1, D2, and D3, and the GPR attenuation data along the cross sections from wells D1 to D3, we estimated extractable Fe(II) and Fe(III) PDFs along the cross section between wells D1 and D3, using the method developed in Section 4. Figure 3(b) shows the estimated mean logarithmic Fe(II) along the two-dimensional transect. The figure provides detailed spatial information about extractable Fe(II) on the cross section, which could not be accurately obtained from borehole measurements only. The estimated spatial pattern of Fe(II) is similar to that of GPR attenuation (Figure (3a)) because of the close correspondence of Fe(II) with GPR attenuation as shown in Figure 6. Figure 3(d) shows the two-dimensional image of

logarithmic Fe(III), which is much different from the GPR attenuation image. Comparison of Figures 3(c) and 3(d) reveals that extractable Fe(III) has higher concentrations beneath the mud layer than above the layer.

As the byproduct of the Fe(II) estimation, we obtained the probability of sand occurring at each pixel on the cross section between wells D1 and D3 (Figure 3(c)). Lithofacies on the cross section between wells D2 and D3 corresponds to GPR attenuation very well, with sand having lower attenuation and mud having higher attenuation. However, lithofacies on the lower part of the cross section between wells D1 and D2 does not correspond directly to GPR attenuation, where GPR attenuation is high but the estimated lithofacies is sand. One reason for the discrepancy is that lithofacies at each pixel were determined by both lithofacies measurements at boreholes and GPR attenuation along the cross section. Another reason is that our sand lithofacies classification included both pure sand and muddy-sand components, and lithofacies at those pixels in reality may be muddy-sand rather than pure sand, which would tend to have a higher GPR attenuation than pure sand. Figure 3(c) also suggests that there is a mud layer in the middle of the cross section, which passes through the three wells.

5.2. Cross-validation

Cross-validation is a model evaluation method, in which data are divided into two subsets: one for testing (referred to as the testing set) and the other for training (referred to as the training set). In this study, we split data available at wells D1, D2, and D3 into two subsets, and considered each individual well in turn as a testing well and the other two wells as training wells. As done previously using the entire dataset, in this exercise we derived cross correlations among various types of properties and spatial corre-

lation of lithofacies from the training dataset only. We then estimated logarithmic Fe(II) and Fe(III) concentrations at the testing locations, using data at the training wells and GPR attenuation data along the transect from wells D1 to D3. By comparing the estimated results with their corresponding true values at the testing well, we evaluated the effectiveness of the developed model for Fe(II) and Fe(III) estimation.

Figure 10 compares the estimated mean logarithmic Fe(II) concentrations (solid black lines), obtained during this cross-validation exercise, with their corresponding core measurements (red circles with red solid lines) at testing wells D1, D2, and D3, respectively. The dashed lines indicate the 95% predictive intervals. This figure suggests that the developed model is effective for logarithmic Fe(II) estimation. As shown in the figure, the mean estimates of logarithmic Fe(II) at well D1 closely follow the true measurements of logarithmic Fe(II), and the mean estimates of logarithmic Fe(II) at testing well D2 are in close agreement with the true values. Although the estimated results of logarithmic Fe(II) at testing well D3 do not match measured results as well as at other testing wells, most measurements are still within the 95% predictive intervals of the estimated values.

Figure 11 presents the estimated mean logarithmic Fe(II) obtained without using GPR attenuation data, to show the improvement offered by the GPR data. In this case, again, the solid black lines represent the estimated values and the red circles with red solid lines represent the true values at the testing wells. Although the three wells in the cross section were separated by only 6 m, the estimated results without using GPR attenuation are much worse than those obtained with the use of GPR attenuation. This is especially apparent at testing well D1. The mean estimates of Fe(II) are smooth, and

the 95% predictive intervals are much larger than those of the model that includes GPR attenuation.

Figure 12 shows the estimated mean logarithmic Fe(III) and their 95% predictive intervals at each testing well, obtained during the cross-validation. Again, Fe(III) estimates at each testing well were obtained using only data from the other two training wells, along with the GPR attenuation, and a comparison of the estimates with the true data is made to assess the validity of our developed model. We found that the measured logarithmic Fe(III) values followed the trends of the estimated mean, and that most of the estimates were within the 95% predictive intervals.

5.3. Discussion

We have developed an MCMC approach for estimating sediment Fe(II) and Fe(III) using GPR tomographic data and lithofacies borehole measurements. Through cross validation, we have shown that the developed approach is significant, given that currently no other methods can sufficiently predict the spatial distribution of field-scale extractable Fe(II) and Fe(III) at both a reasonable cost and in a minimally invasive manner. Although the model cannot accurately estimate Fe(II) and Fe(III) point values at the "core-scale", it provided estimates at the spatial resolution of the tomographic data (0.25m x 0.25m for this study), which was previously determined to be an effective characterization scale for understanding field-scale chemical and bacteria transport [*DeFlaun et al.*, 2001]. The estimated spatial distributions of Fe(II) and Fe(III) may be used further for the inference of microbial iron reduction potential [*Murray et al.*, 2001] and for the estimation of bacterial attachment or detachment parameters, which are needed as input for numerical predictions of subsurface bacterial transport.

Our goal in this study was to advance methodologies for linking sediment geochemical properties to geophysical data, rather than to provide a mechanistic explanation for the origin of the geochemical properties. Thus, it is not our intention to assess the extent to which the solid-phase Fe(III) and Fe(II) distributions are controlled by ongoing microbial reduction processes in the deposit, or by slow redistribution of iron during coupled transport/reaction processes acting on materials whose original composition was determined by short-term events taking place at the time of deposition.

The developed model was based on the data analysis results without considering field-scale geochemical and biogeochemical processes, and thus was site specific. However, our hypotheses that geochemical parameters and geophysical attributes can be linked through their mutual dependence on physical properties should be applicable to estimation of other geochemical parameters, such as organic matter and other metal oxyhydroxides, as well as to other field sites.

Appendix A: Derivation of $[X_i|\cdot]$

By substituting Equations 3 and 4 into Equation 7 and omitting all terms that are not directly related to X_i, we obtained $[X_i|\cdot]$ as follows:

$$
\begin{aligned}
[X_i|\cdot] \propto\ & \exp\{-\frac{\tau_3}{2}(Y_i - r_1 - r_2 X_i - r_3 d_i)^2\} \cdot \\
& \exp\{-\frac{\tau_2}{2}(X_i - v_1 - v_2 L_i - v_3 a_i - v_4 L_i a_i)^2\} \\
=\ & \exp\{-\frac{\tau_3}{2}(r_2 X_i - Y_i + r_1 + r_3 d_i)^2\} \cdot \\
& \exp\{-\frac{\tau_2}{2}(X_i - v_1 - v_2 L_i - v_3 a_i - v_4 L_i a_i)^2\} \\
\propto\ & \exp\{-\frac{\tau_2 + r_2^2 \tau_3}{2} X_i^2 + \\
& \tau_2(v_1 + v_2 L_i + v_3 a_i + v_4 L_i a_i)X_i + \tau_3 r_2(Y_i - r_1 - r_3 d_i)X_i\} \\
\sim\ & Normal(\frac{\tau_2(v_1 + v_2 L_i + v_3 a_i + v_4 L_i a_i) + r_2 \tau_3(Y_i - r_1 - r_3 d_i)}{\tau_2 + r_2^2 \tau_3}, \tau_2 + r_2^2 \tau_3),
\end{aligned}
\tag{A1}
$$

where v_1, v_2, v_3, v_4, r_1, r_2, r_3 are regression coefficients, τ_2 and τ_3 are the inverse variances, and Y_i, L_i, d_i, and a_i are the logarithmic Fe(III) concentration, lithofacies, depth, and logarithmic GPR attenuation at pixel-i, respectively.

Appendix B: Indicator Kriging

Let set A_i be the index set of the adjacent pixels of pixel-i. Let p_s be the unconditional probability of observing sand at any pixel. The simple kriging mean μ_{pi} of observing sand is given by:

$$
\mu_{pi} = p_s + \sum_{j \in A_i} \lambda_j (L_j - p_s).
\tag{B1}
$$

The coefficients λ_j $(j \in A_i)$ are determined by:

$$
\sum_{j \in A_i} \lambda_j C_{kj} = C_{ki}, \qquad k \in A_i,
\tag{B2}
$$

where C_{kj} and C_{ki} are the lithofacies covariances between pixel-k and pixel-j and between pixel-k and pixel-i, respectively. Both covariances are given by

$$C_{ij} = \sigma^2 \exp\left\{-\sqrt{(\frac{\Delta x}{I_h})^2 + (\frac{\Delta z}{I_v})^2}\right\}, \tag{B3}$$

where σ^2 is a constant, which does not affect estimation of the coefficients, and Δx and Δ_z are the distances along horizontal and vertical directions between pixels i and j. I_v and I_h are the integral lengths along vertical and horizontal directions, and were derived from borehole data as explained in Section 2. To ensure $\mu_{pi} \in [0,1]$, we let $p^* = min\{1, max\{0, \mu_{pi}\}\}$.

Appendix C: Derivation of $[L_i|\cdot]$

By substituting Equations 2 and 3 and using the identity $L_i^2 = L_i$, we can obtain the full conditional probability of lithofacies as follows:

$$[L_i|\cdot] \sim Bernoulli(p_i) \tag{C1}$$

where

$$p_i \propto [a_i|L_i][X_i|a_i, L_i][L_i|L_j, j \in A_i] \tag{C2}$$

$$\propto \exp\left\{-0.5\tau_1(a_i - u_1 - u_2 L_i)^2 - 0.5\tau_2(X_i - v_1 - v_2 L_i - v_3 a_i - v_4 L_i a_i)^2\right\}$$

$$\{L_i p^* + (1 - L_i)(1 - p^*)\}.$$

$$\propto \exp\left\{\tau_1 u_2 L_i(a_i - u_1 - 0.5 u_2) + \tau_2(v_2 + v_4 a_i)L_i(X_i - v_1 - 0.5 v_2 - v_3 a_i - 0.5 v_4 a_i)\right\}$$

$$\{L_i p^* + (1 - L_i)(1 - p^*)\}.$$

Normalization yields:

$$p_i = \frac{p^* p}{1 - p^* + p^* p}. \tag{C3}$$

where

$$p = \exp\left\{\tau_1 u_2(a_i - u_1 - 0.5u_2) + \tau_2(v_2 + v_4 a_i)(X_i - v_1 - 0.5v_2 - v_3 a_i - 0.5v_4 a_i)\right\} \quad (C4)$$

References

Bernardo, J. M., and A. F. Smith, *Bayesian Theory*, John Wiley and Sons, New York, 1994.

Binley, A., and K. Beven, Vadose zone flow model uncertainty as conditioned on geophysical data, *Ground Water*, *41*(2), 119–127, 2003.

Bosch, M., Lithologic tomography: From plural geophysical data to lithology estimation, *Journal of Geophysical Research*, *104*(B1), 749–766, 1999.

Caccavo, F., P. C. Schamberger, K. Keiding, and P. H. Nielsen, Isotope fractionation by natural populations of sulfate-reducing bacteria, *Applied & Environmental Microbiology*, *63*(10), 3837–3843, 1997.

Canfield, D. E., Isotope fractionation by natural populations of sulfate-reducing bacteria, *Geochimica et Cosmochimica Acta*, *65*(7), 1117–1124, 1989.

Chapellier, F. H., *Groundwater Microbiology and Geochemistry*, John Wiley & Sons, New York, 2001.

Chen, J., and Y. Rubin, An effective Bayesian model for lithofacies estimation using geophysical data, *Water Resources Research*, *39*(5), 2003.

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

Chen, J., S. Hubbard, Y. Rubin, C. Murray, E. Roden, and E. Majer, Geochemical characterization using geophysical data and Markov Chain Monte Carlo methods, *EOS Trans. AGU*, *83*(47), Fall Meet. Suppl., Abstract H52F-10, 2002.

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

DeFlaun, M. F., D. Balkwill, J. Chen, F. Dobbs, H. Dong, J. Fredrickson, M. Fuller, M. Green, T. Ginn, T. Griffin, W. Holben, S. Hubbard, W. Johnson, P. Long, B. Mailloux, E. Majer, M. Mclnerney, C. Murray, T. Onstott, T. Phelps, T. Scheibe, D. Swift, D. White, and F. Wobber, Breakthroughs in field-scale bacterial transport, *EOS Transactions*, *82*(38), 423–425, 2001.

Dong, H., T. C. Onstott, M. F. Deflaun, T. D. Scheibe, S. H. Streger, R. K. Rothmel, and B. J. Mailloux, Relative dominance of physical versus chemical effects on the transport of adhesion-deficient bacteria in intact cores from South Oyster, Virgina, *Environmental Sciences & Technology*, *36*(5), 891–900, 2002.

Gelfand, A. E., and A. F. M. Smith, Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, *85*(410), 398–409, 1990.

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

Geman, S., and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-6*(6), 721–741, 1984.

Geweke, J., Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, in *Bayesian Statistics 4*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and F. M. Smith, pp. 169–193, Oxford University Press, 1992.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, New York, 1996.

Hastings, W. K., Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *87*, 97–109, 1970.

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

Hubbard, S., J. Chen, J. Peterson, E. Major, K. Williams, D. J. Swift, B. Mailloux, and Y. Rubin, Hydrogeological characterization of the south oyster bacterial transport site using geophysical data, *Water Resources Research*, *37*(10), 2431–2456, 2001.

Johnson, W., P. Zhang, M. Fuller, T. Scheibe, B. Mailloux, T. Onstott, M. DeFlaun, S. Hubbard, J. Radtke, W. Kovacik, and W. Holbin, Ferrographic tracking of bacterial transport in the field at the Narrow Channel Focus Area, Oyster, VA, *Environmental Science & Technology*, *35*(1), 182–191, 2001.

Mailloux, B. J., M. E. Fuller, T. C. Onstott, J. Hall, H. Dong, M. F. DeFlaun, S. H. Streger, R. K. Rothmel, M. Green, D. J. P. Swift, and J. Radke, The role of physical, chemical, and microbial heterogeneity on the field-scale transport and attachment of bacteria, *Water Resources Research*, *39*(6), 2003.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, *21*, 1087–1091, 1953.

Michalak, A. M., and P. K. Kitanidis, A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification, *Water Resources Research*, *39*(2), 2003.

Murray, C. J., Y. L. Xie, E. E. Roden, and K. Overstreet, Microbial iron reduction potential in coastal plain sediments (abstract), in *Geological Society of American Annual Meeting*, Boston, MA, 2001.

Peterson, J. E., Pre-inversion processing and analysis of tomographic radar data, *Journal of Environmental and Engineering Geophysics*, 2001.

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

Raftery, A. E., and S. Lewis, How many iterations in the Gibbs sampler?, in *Bayesian Statistics 4*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and F. M. Smith, pp. 763–773, Oxford University Press, 1992.

Roden, E. E., and D. R. Lovely, Evaluation of Fe-55 as a tracer of Fe(iii) reduction in aquatic sediments, *Geomicrobiology Journal*, *11*(1), 49–56, 1993.

Roden, E. E., M. M. Urrutia, and C. J. Mann, Bacterial reductive dissolution of crystalline fe(iii) oxide in continuous-flow column reactors, *Applied & Environmental Microbiology*, *66*(3), 1062–1065, 2000.

Rubin, Y., *Applied Stochastic Hydrology*, Oxford University Press, New York, 2003.

Scheibe, T. D., and Y.-J. Chien, An evaluation of conditioning data for solute transport prediction, *Ground Water*, *41*(2), 128–141, 2003.
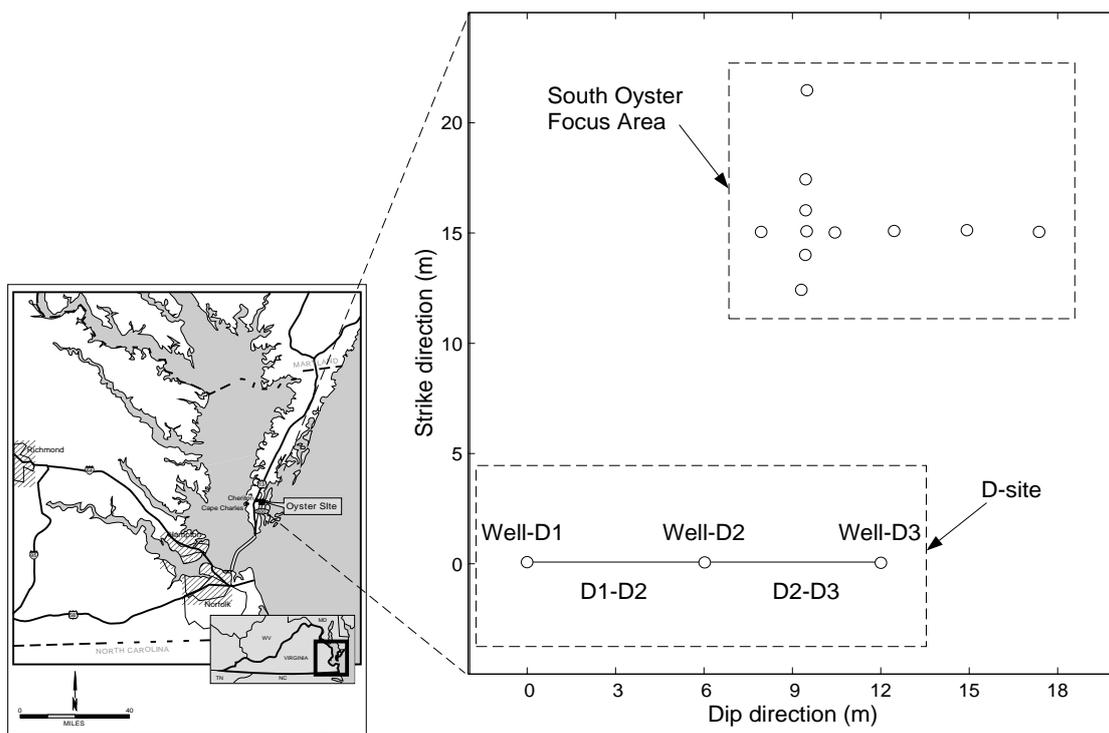
Figure 1. Locations of the South Oyster Bacterial Transport Site and our study area (D-site). Wellbore data were collected at wells D1, D2, and D3, and GPR attenuation tomograms were collected along transects D1-D2 and D2-D3.
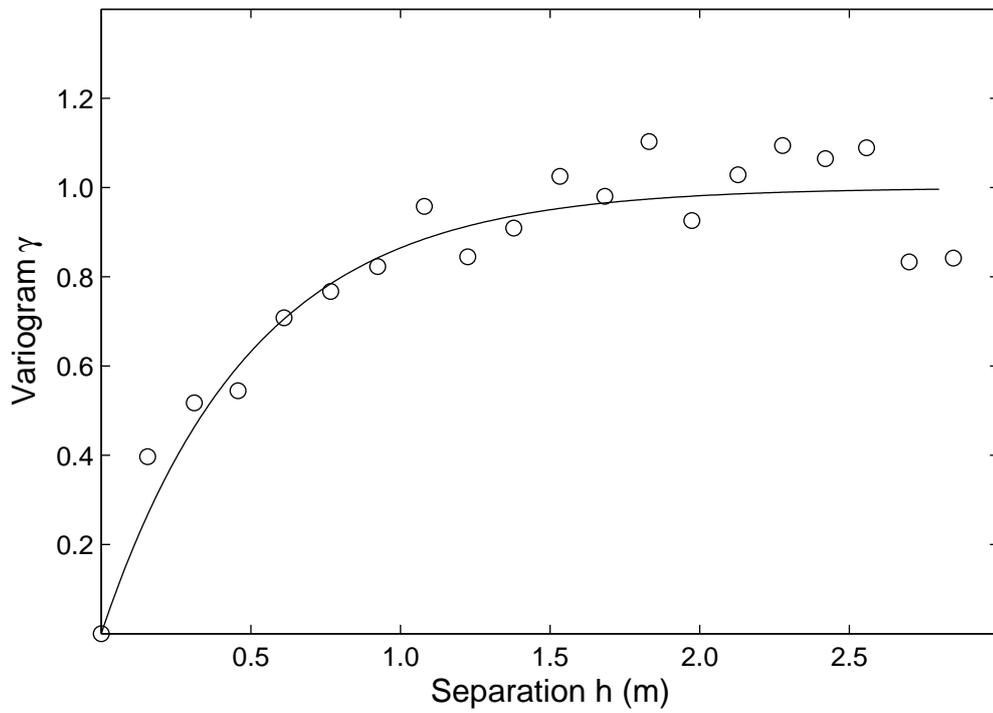
Figure 2.   Normalized variogram of the coded lithofacies (sand=1 and mud=0) along the

vertical direction. The solid line is the fitted exponential model, $\gamma(h) = 1 - \exp(-h/0.5)$, with

an integral length of 0.5 m. The circles are the experimental variogram calculated from data at

wells D1, D2, and D3.
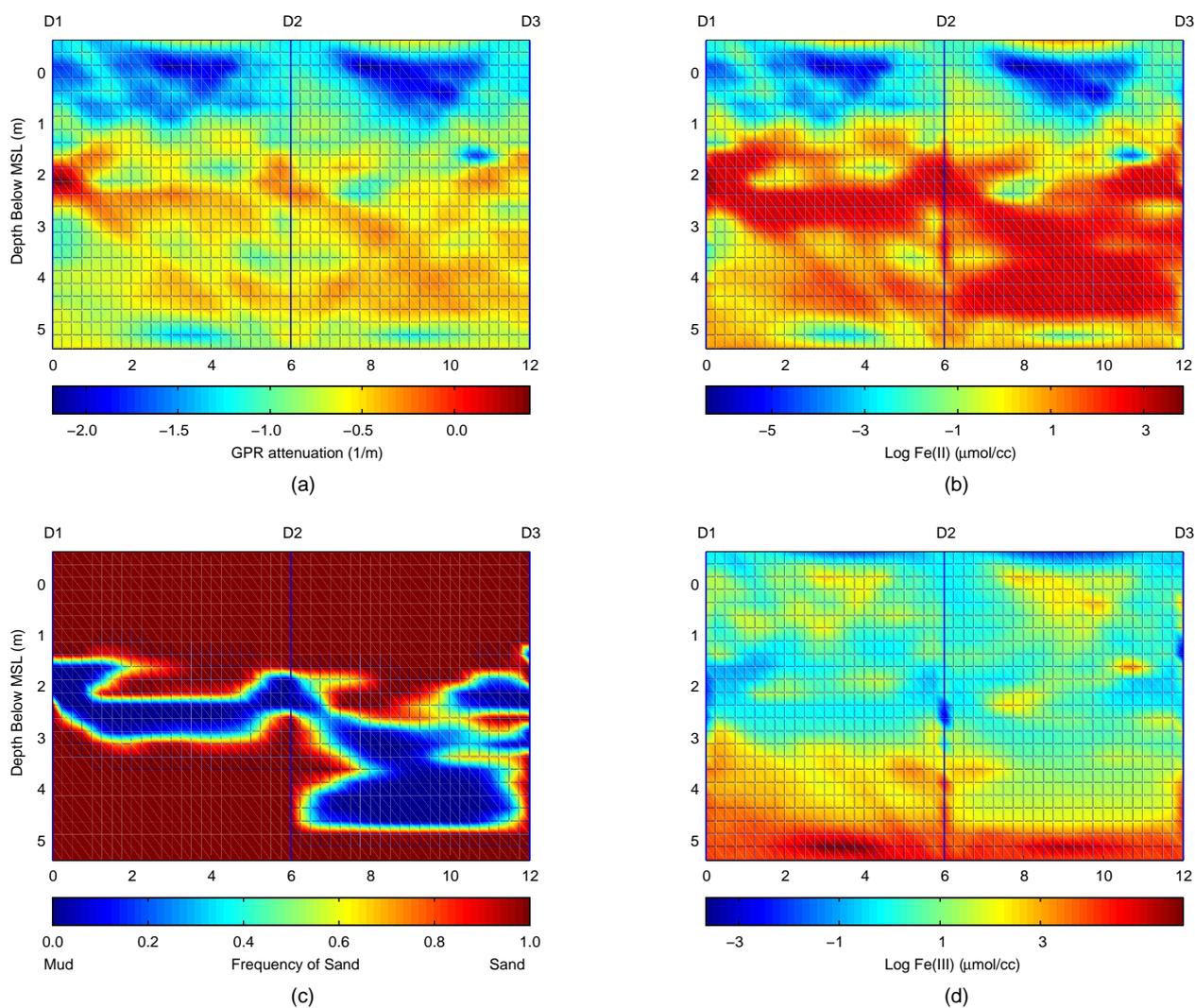
Figure 3.   (a) GPR attenuation; (b) Estimated mean natural logarithmic Fe(II); (c) Frequency of sand (a frequency of 0.0 implies that lithofacies is mud, whereas a frequency of 1.0 implies that lithofacies is sand); (d) Estimated mean natural logarithmic Fe(III).
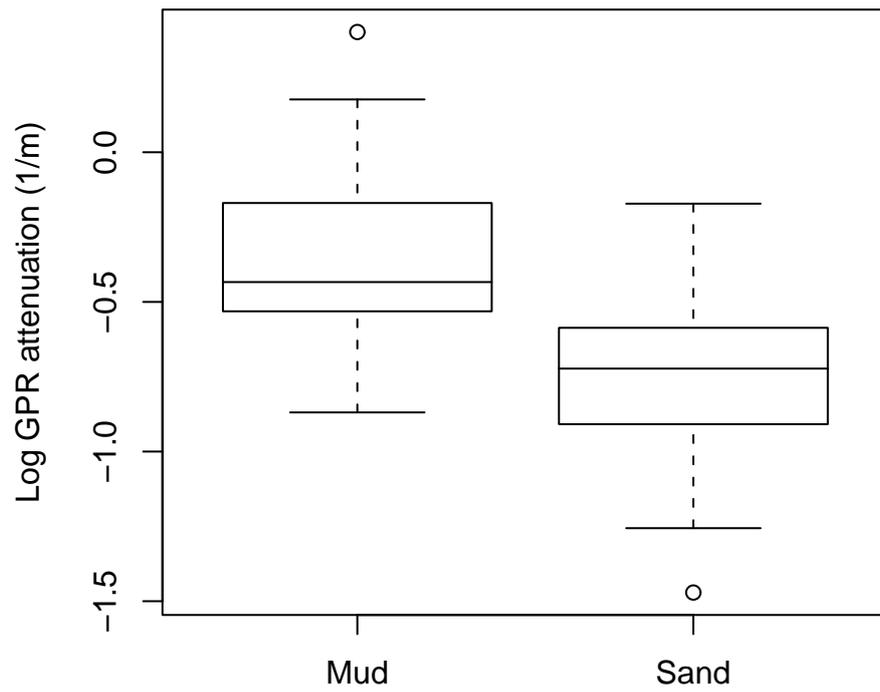
Figure 4. Boxplot of natural logarithmic GPR attenuation

(a)

(b)
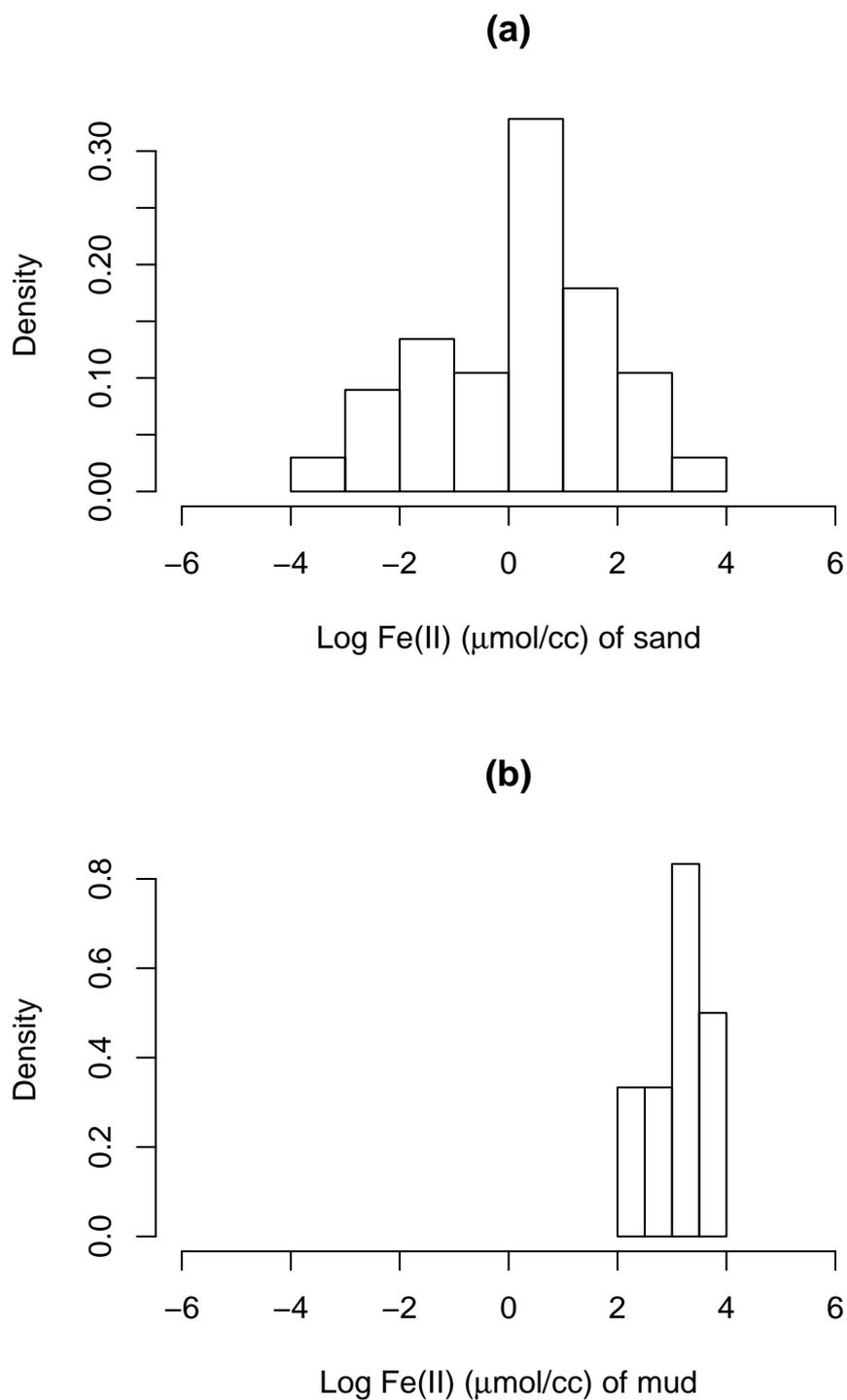
Figure 5.   Histograms of natural logarithmic Fe(II) as a function of lithofacies
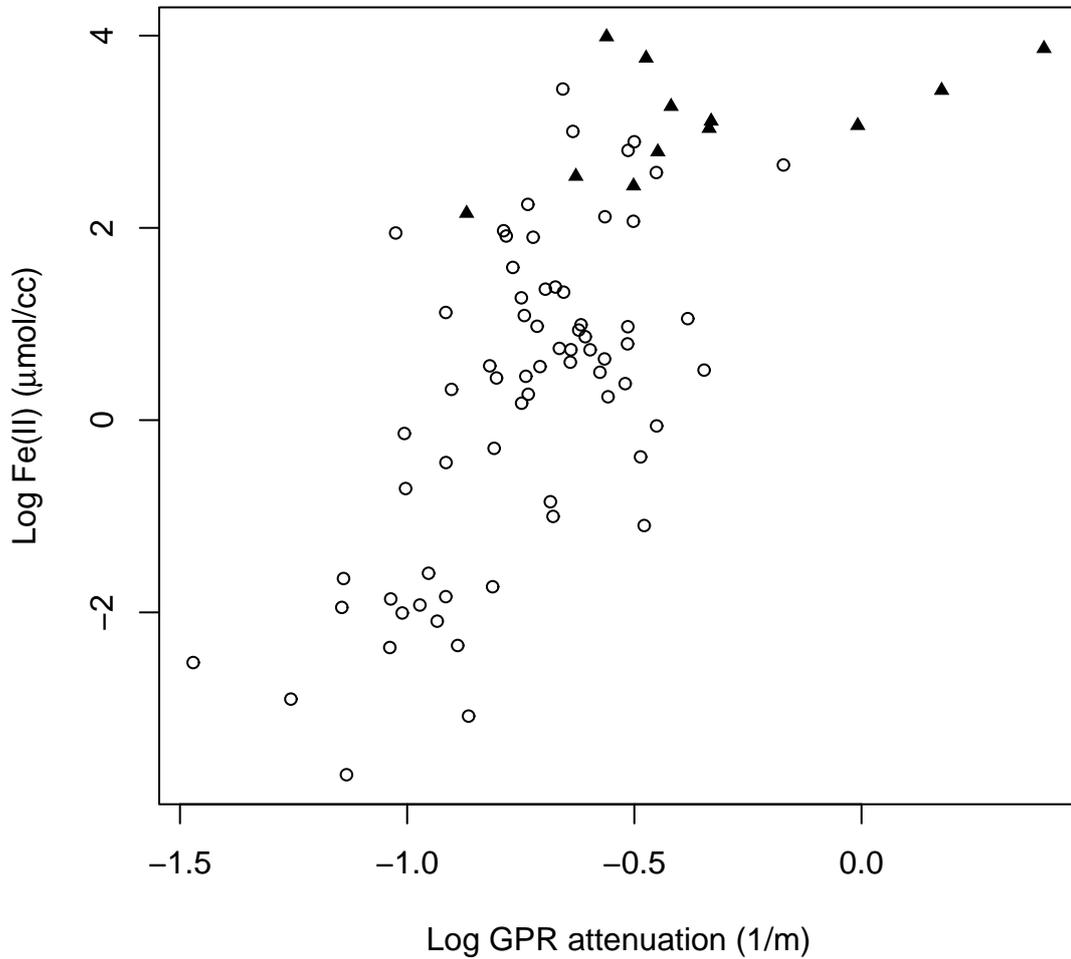
Figure 6.   Cross-plot of natural logarithmic Fe(II) versus natural logarithmic GPR attenuation, where the circles represent sand samples and the solid triangles represent mud samples

(a)

(b)
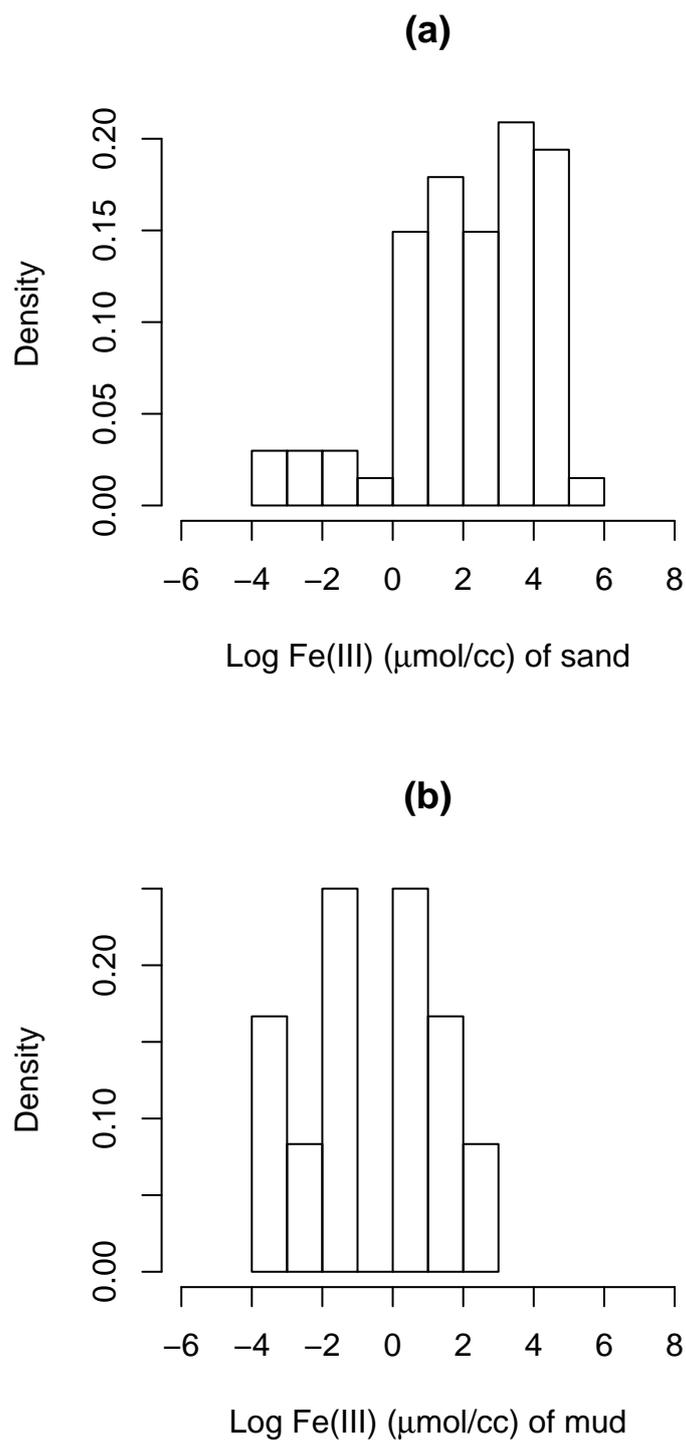
Figure 7. Histograms of natural logarithmic Fe(III).

Row=25, Column=49, Pixel size = 0.25m x 0.25m

Figure 8. Discretization of the cross section between wells D1 and D3. The total number of pixels is $25 \times 49 = 1225$.

Figure 9. Schematic diagram of the correlations among lithofacies, GPR attenuation, Fe(II), Fe(III), and depth. The rectangles mean given data, and the ellipses mean unknown variables.
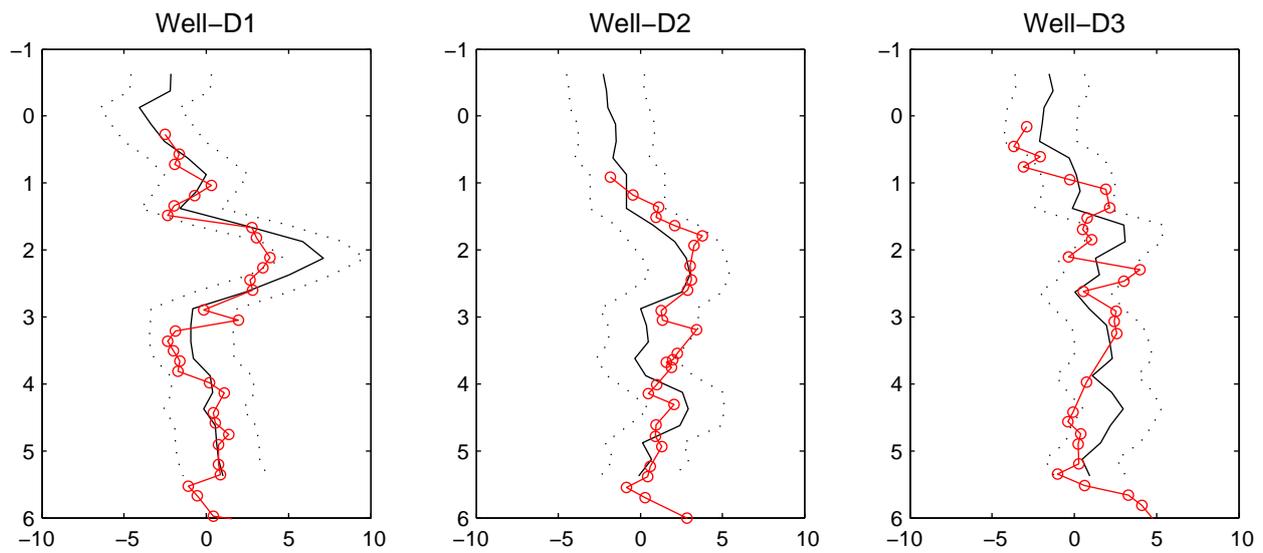
Figure 10. Comparisons of the true logarithmic Fe(II) and its corresponding estimated values obtained using GPR attenuation during the cross-validation exercise
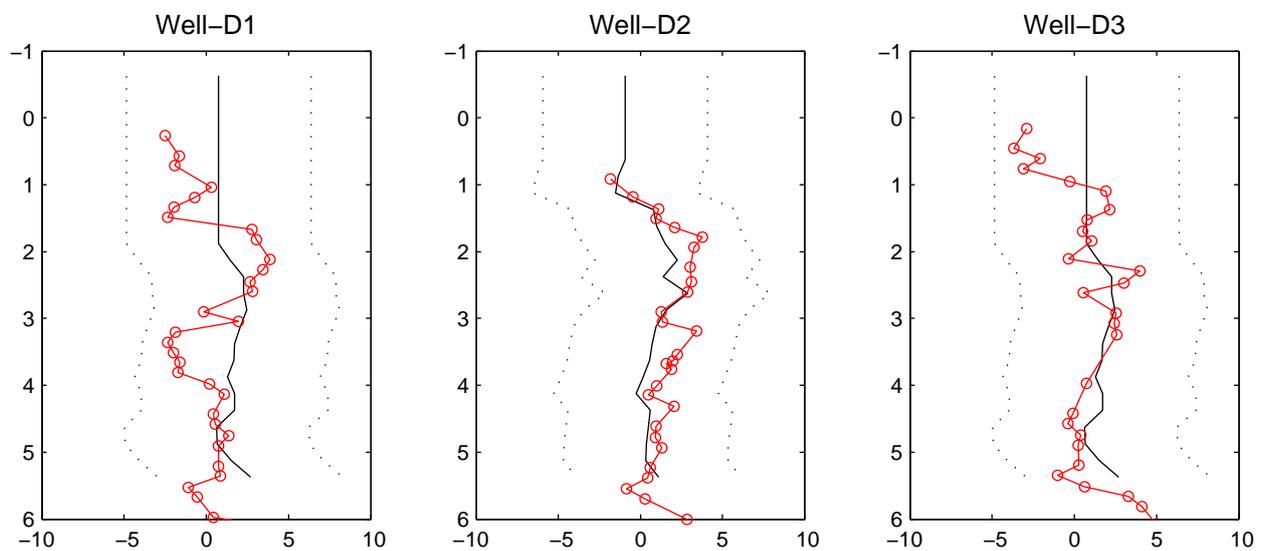
Figure 11. Comparisons of the true logarithmic Fe(II) and its corresponding estimated values obtained without using GPR attenuation during the cross-validation exercise
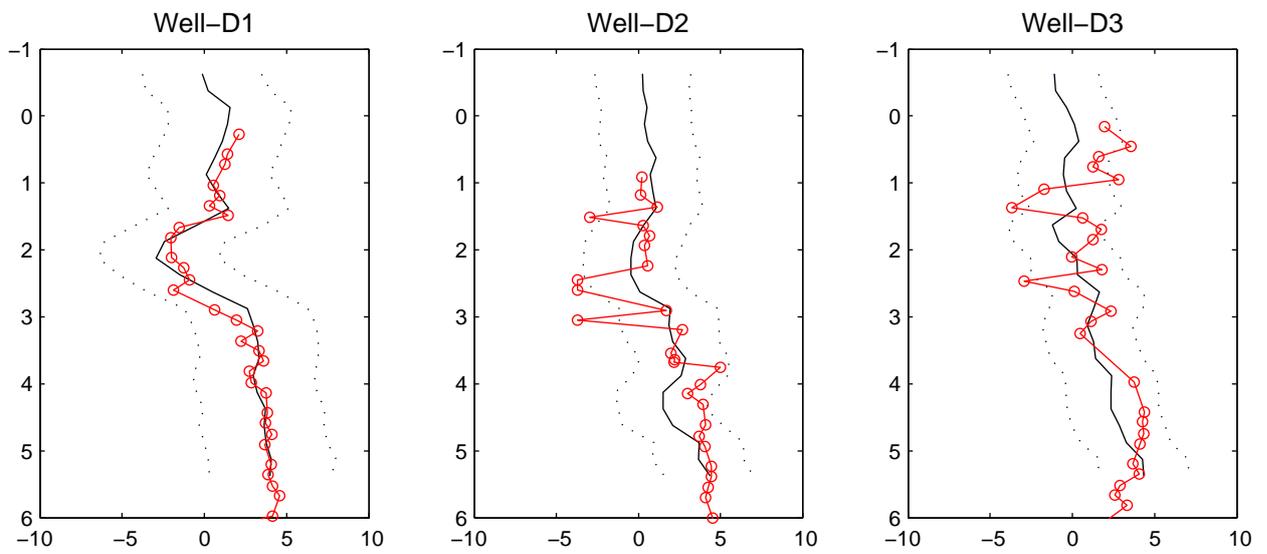
Figure 12. Comparisons of the true logarithmic Fe(III) and its corresponding estimated values obtained using GPR attenuation during the cross-validation exercise

Table 1.   Coefficients of cross correlations among lithofacies, geochemical parameters, and geophysical data corresponding to Equations 2–4

Models	Coefficients
Logarithmic GPR attenuation vs. lithofacies	$u_1 = -0.3332$, $u_2 = -0.4110$, $\tau_1 = 15.58$ $R^2 = 0.26$
Logarithmic Fe(II) vs lithofacies and logarithmic GPR attenuation	$v_1 = 3.4128$, $v_2 = 0.3085$, $v_3 = 0.8796$ $v_4 = 3.7870$, $\tau_2 = 0.70$, $R^2 = 0.60$
Logarithmic Fe(III) vs depth and logarithmic Fe(II)	$r_1 = -0.8813$, $r_2 = -0.5910$, $r_3 = 1.0026$ $\tau_3 = 0.45$, $R^2 = 0.60$