Estimating the spatiotemporal distribution of geochemical parameters associated with biostimulation using spectral induced polarization data and hierarchical Bayesian models

Jinsong Chen,[1] Susan S. Hubbard,[1] Kenneth H. Williams,[1] Adrián Flores Orozco,[2] and Andreas Kemna[2]

Received 1 June 2011; revised 26 March 2012; accepted 17 April 2012; published 30 May 2012.

[1] We developed a hierarchical Bayesian model to estimate the spatiotemporal distribution of aqueous geochemical parameters associated with in-situ bioremediation using surface spectral induced polarization (SIP) data and borehole geochemical measurements collected during a bioremediation experiment at a uranium-contaminated site near Rifle, Colorado (USA). The SIP data were first inverted for Cole-Cole parameters, including chargeability, time constant, resistivity at the DC frequency, and dependence factor, at each pixel of two-dimensional grids using a previously developed stochastic method. Correlations between the inverted Cole-Cole parameters and the wellbore-based groundwater chemistry measurements indicative of key metabolic processes within the aquifer (e.g., ferrous iron, sulfate, uranium) were established and used as a basis for petrophysical model development. The developed Bayesian model consists of three levels of statistical submodels: (1) data model, providing links between geochemical and geophysical attributes, (2) process model, describing the spatial and temporal variability of geochemical properties in the subsurface system, and (3) parameter model, describing prior distributions of various parameters and initial conditions. The unknown parameters were estimated using Markov chain Monte Carlo methods. By combining the temporally distributed geochemical data with the spatially distributed geophysical data, we obtained the spatiotemporal distribution of ferrous iron, sulfate, and sulfide, and their associated uncertainty information. The obtained results can be used to assess the efficacy of the bioremediation treatment over space and time and to constrain reactive transport models.

Citation: Chen, J., S. S. Hubbard, K. H. Williams, A. Flores Orozco, and A. Kemna (2012), Estimating the spatiotemporal distribution of geochemical parameters associated with biostimulation using spectral induced polarization data and hierarchical Bayesian models, *Water Resour. Res.*, *48*, W05555, doi:10.1029/2011WR010992.

1. Introduction

[2] In-situ bioremediation is often considered as a key approach for subsurface environmental remediation, and effective monitoring and understanding of biogeochemical processes are critical for success of the treatment [*Anderson et al.*, 2003; *Vrionis et al.*, 2005; *Yabusaki et al.*, 2007]. Geophysical methods, especially spectral induced polarization (SIP), have been shown to be very useful for providing remote information about bioremediation processes. This is because the accumulation of mineral precipitates, reactive ions, and biomass that are induced through the bioremediation process may collectively affect the electrical response [*Ntarlagiannis et al.*, 2005]. Although it has been recognized that SIP methods can be used to characterize and identify contaminants at the field scale [*Olhoeft*, 1992; *Vanhala*, 1997; *Sogade et al.*, 2006], the value of SIP methods for monitoring remediation-induced solid phase end-products has mostly been demonstrated through laboratory experiments. For example, *Williams et al.* [2005] established that SIP responses track the onset and space-time distribution of bioremediation-induced FeS precipitates using laboratory column experimental data. *Slater et al.* [2007] showed the SIP signatures are diagnostic of porescale geometrical changes associated with FeS bioremediation by sulfate reducing bacteria. *Personna et al.* [2008] used laboratory column experimental data to track the onset of anaerobic conditions and the reoxidation to aerobic conditions through SIP's sensitivity to iron sulfide precipitation and dissolution. *Chen et al.* [2009] developed a state-space Bayesian model that allowed quantitative estimation of the evolution of bioremediation-induced FeS precipitates and associated permeability reduction using time-lapse SIP data collected in a laboratory column experiment.

[3] However, monitoring the evolution of solid-phase end-products is not a standard approach for evaluating the

[1]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.
[2]Department of Geodynamics and Geophysics, University of Bonn, Bonn, Germany.

Corresponding author: J. Chen, Earth Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., MS 90-1116, Berkeley, CA 94720, USA. (jchen@lbl.gov)

efficacy of bioremediation at the field scale. Instead, the onset and extent of bioremediation is typically inferred by means of monitoring changes in geochemical parameters from groundwater samples, such as electron acceptor consumption, various final products (e.g., dissolved iron or sulfides), and concentrations of dissolved hydrogen based on aqueous wellbore measurements [e.g., *Lovley et al.*, 1994]. The geochemical measurements are also often used to investigate the underlying biogeochemical processes at the borehole sampling locations and to constrain reactive transport models [e.g., *Li et al.*, 2010].

[4] Several recent studies have qualitatively illustrated the potential of SIP imaging for tracking subsurface geochemical changes associated with bioremediation at the field scale. *Williams et al.* [2009, 2011] showed that the phase response in SIP images was associated with changes in groundwater geochemistry accompanying stimulated iron and sulfate reduction and sulfide mineral precipitation. *Commer et al.* [2011] performed a three-dimensional (3-D) inversion of time-lapse surface SIP data collected during a bioremediation and documented the change in phase and resistivity associated with the treatment. *Johnson et al.* [2010] developed a parallel distributed-memory forward and inverse modeling algorithm for analyzing resistivity and time domain induced-polarization (IP) data and applied it to the Brandywine field site in Maryland for monitoring bioremediation. *Flores Orozco et al.* [2011] reported a good correlation between the increase in the SIP phase response and the increases in Fe(II) and precipitation of metal sulfides following biostimulation, reflecting preservation of geochemically reduced conditions within the aquifer, for the SIP measurements collected over two years of monitoring and for different experiments.

[5] Our goal in this study is to provide quantitative information on the spatiotemporal distribution of remediation-induced changes in aqueous chemical species that are indicative of redox status (i.e., ferrous iron or Fe2, sulfate and sulfide) through integrating 2-D SIP data with sparse wellbore aqueous geochemical measurements. The primary challenge of the estimation is to handle properly the discrepancy between the temporal sampling frequency and the spatial resolution of the different types of measurements. We can measure aqueous geochemical concentrations over time (dense in time) as done at the laboratory during bioremediation treatments. Because drilling boreholes is invasive and costly, we can only sample them at very small numbers of locations (sparse in space). Surface SIP data have a large spatial coverage (tens of meters); hence, the measurement support scale is much larger than that associated with a typical wellbore sample. However, surface SIP data typically have lower spatial resolution relative to the wellbore geochemical measurements.

[6] In this study we develop a hierarchical Bayesian model based on time-lapse surface SIP and wellbore geochemical data collected at the Department of Energy (DOE) Integrated Field Research Challenge (IFRC) Site near Rifle, Colorado to estimate the spatiotemporal distribution of aqueous geochemical parameters associated with a subsurface biostimulation experiment. We combine the borehole geochemical measurements, having high temporal resolution but being spatially sparse, with the surface SIP data, having large coverage but low temporal resolution, within the Bayesian framework. We use Markov chain Monte Carlo (MCMC) sampling methods to explore the joint posterior probability distribution.

[7] The remainder of this paper is organized as follows. Section 2 describes the Rifle site and results of data analysis, which provide basis for development of the hierarchical Bayesian model. Section 3 describes the hierarchical Bayesian framework and the MCMC sampling approach for estimating the spatiotemporal distribution using borehole geochemical time series and SIP data. The estimation results are given in section 4 and discussion and conclusions are provided in sections 5.

2. Geophysical and Geochemical Measurements and Data Analysis

2.1. In-Situ Bioremediation

[8] Numerous in-situ bioremediation experiments have been carried out at the DOE IFRC Rifle site from 2002 to 2009 [*Williams et al.*, 2011]. The experiments have been conducted within an unconfined fluvial aquifer that includes sandy gravely unconsolidated sediments with variable clay content and that is underlain (at 5.9–7.0 m below the ground surface) by a relatively impermeable, regional aquitard known as the Wasatch formation [*Williams et al.*, 2011], which consists of the variegated mudstones and conglomeratic sandstones [*Lorenz*, 1982]. X-ray diffraction (XRD) analysis of the clay-sized fraction (<2 μm) of Rifle alluvium identified the primary clay minerals to be smectite, illite, and kaolinite, with smectites most abundant (K. H. Williams, unpublished data, 2011). The water table is located around 3.5 m below the ground surface, with fluctuations (less than 1 m) along the year. At the Rifle site, IFRC investigators have repeatedly demonstrated the ability to remove uranium rapidly from the tailings-contaminated groundwater by stimulating the activity of iron- and sulfate-reducing bacteria through acetate amendment [e.g., *Anderson et al.*, 2003; *Vrionis et al.*, 2005; *Williams et al.*, 2011].

[9] The transition from iron reduction to sulfate reduction at the Rifle site occurs at relatively predictable times following the amendment injection [*Yabusaki et al.*, 2007; *Li et al.*, 2010]. Injection of acetate (electron donor) is initially going to stimulate the activity of iron-reducing bacteria (Geobacter) commonly found in Rifle sediments [*Williams et al.*, 2011]. Iron reduction is expected to reduce iron (hydro-) oxide to ferrous iron Fe(II) or Fe2, simultaneously with aqueous U(VI) to immobile U(IV). After the depletion of "bioavailable" iron, sulfate is expected to be reduced by sulfate reducers (sulfate reducing proteobacteria), which should lead to the accumulation of aqueous S(-II) and eventually the formation of amorphous FeS(am).

[10] Our study focuses on a 36-days acetate injection experiment from 22 July 2009 to 27 August 2009 (see Figure 1). During the field experiment, the site groundwater was amended with sodium acetate, with target concentrations of 15 mM, and sodium bromide, with target concentrations of 1.3 mM. The acetate and bromide amended groundwater were injected to the aquifer via ten boreholes (G51–G60). Groundwater samples were collected throughout the experiment from three up gradient monitoring wells (U01–U03) and 12 down gradient monitoring wells (D01–D12).

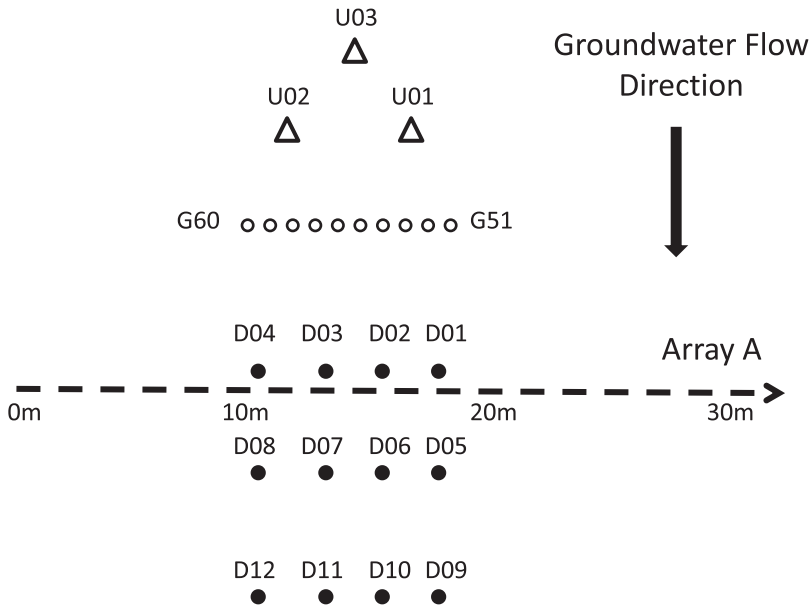

Figure 1. Schematic plan view of the 2009 bioremediation experiment. The open circles (G51–G60) are the ten injection boreholes and the solid circles (D01–D12) are the 12 down gradient monitoring wells. The open triangles (U01–U03) are the three up gradient monitoring wells and the dashed line is the survey profile of surface spectral induced polarization (SIP) used for this study.

2.2. Surface SIP Data and Cole-Cole Parameter Estimation

[11] To monitor and understand the in-situ bioremediation processes, surface spectral induced polarization (SIP) data were collected along array A by using a Zonge GDP32(II) with 15 channels on three days, i.e., 22 July 2009 (before injection), 10 August 2009 (during injection), and 29 August 2009 (after injection). The measurements were made by deploying 30 $Cu/CuSO_4$ nonpolarizing electrodes ($\sim$4 cm in radius) with a separation of 1 m between electrodes. We used coaxial cables to connect the electrodes with the receivers to reduce electromagnetic coupling effects, as demonstrated by Flores Orozco et al. (Time-lapse spectral induced polarization imaging of stimulated uranium bioremediation, submitted to *Near Surface Geophys*.). The measurement frequencies were selected based on the equipment constraints (i.e., 0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128, and 256 Hz). We repeated the measurements twice for frequencies below 2 Hz and repeated 4, 8, 16, 32, 64, 128, 256 times for higher frequencies (increasing by power of two with the increasing of the acquisition frequencies). We measured contact resistances before every survey and found a good contact between the electrodes and the ground, with values of $\sim$700 Ω. We injected electric currents between the electrodes with a constant voltage of 55 V, which results in current densities of $\sim$150 mA.

[12] The phase and amplitude estimates were obtained from the recorded data using the least-squares based inversion method developed by *Kemna* [2000] and detailed processing for the Rifle data were given by *Flores Orozco et al.* [2011]. Figure 2 shows the phases in milliradians (mrad) as a function of frequencies on 10 August 2009 along the cross section from depth $z = 0$ m to $z = 10$ m and horizontal distance from $x = 0$ m to $x = 30$ m, with grid sizes of $dx = dz = 0.5$ m. The domain that we focus on in this study nearly traverses the sampling boreholes D1, D2, D3, and D4 (see Figure 1). We can see that the spatial distribution of the phase changes over frequencies between 0.0625 and 32 Hz and almost no change is observed for frequencies beyond 32 Hz. As observed in previous studies performed at the site [*Williams et al.*, 2009; *Flores Orozco et al.*, 2011], for the frequency range analyzed in our study ($<$32 Hz), the polarization response at the Rifle site during biostimulation is mainly controlled by charge transfer processes taking place at the interface between pore water and the surface of the precipitated semiconductive (metallic) minerals.

[13] We use the obtained amplitude and phase values and a stochastic inversion method developed by *Chen et al.* [2008] to estimate Cole-Cole parameters, such as DC resistivity, chargeability, time constant, and dependence factor. Since the inclusion of amplitude data makes the fitting of the phase data significantly worse, we use only the phase data in this study. Figure 3 shows the data misfits of the phase data at the four nearby boreholes (i.e., D1, D2, D3, and D4) on the three survey days (i.e., 22 July, 10 August, and 29 August) using a simple or double Cole-Cole model, depending on data sets. Even in this case where we only have measurements at ten different frequencies, these figures show that the fitted curves generally follow the phase data well. For some data sets, we can see two peaks, which may suggest a better fitting by the superposition of a double Cole-Cole model as done by *Cosenza et al.* [2007]. Figure 4 shows the histograms of root-mean-squares (RMS) of the differences between the measured and calculated phase responses for 5000 random samples. We generally have smaller misfits for those data collected at borehole D1 and on 10 August 2009 (during injection).

[14] We can similarly fit the SIP data at other pixels of the 20×20 grids to get Cole-Cole model parameters along

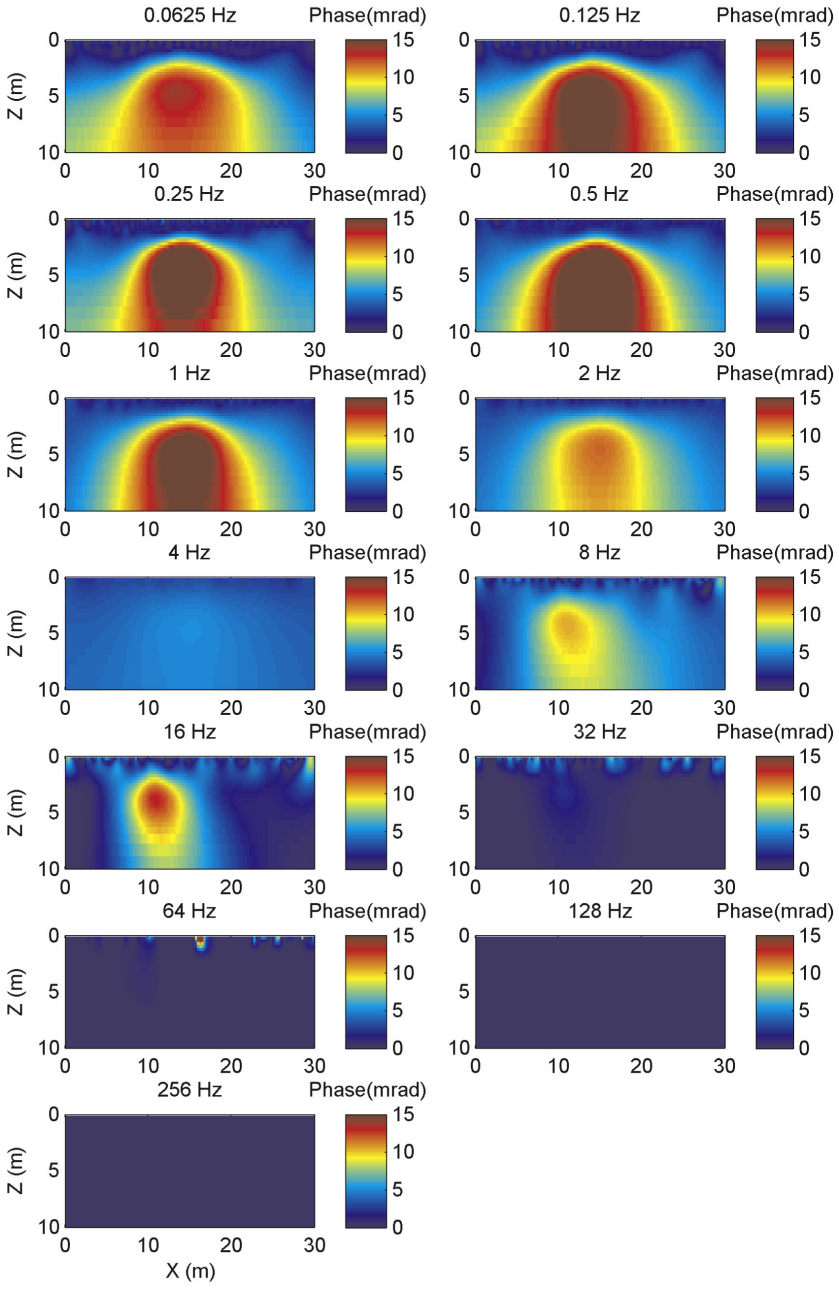

Figure 2. Phases (mrad) along array A at 13 different frequencies for data collected on 10 August 2009.

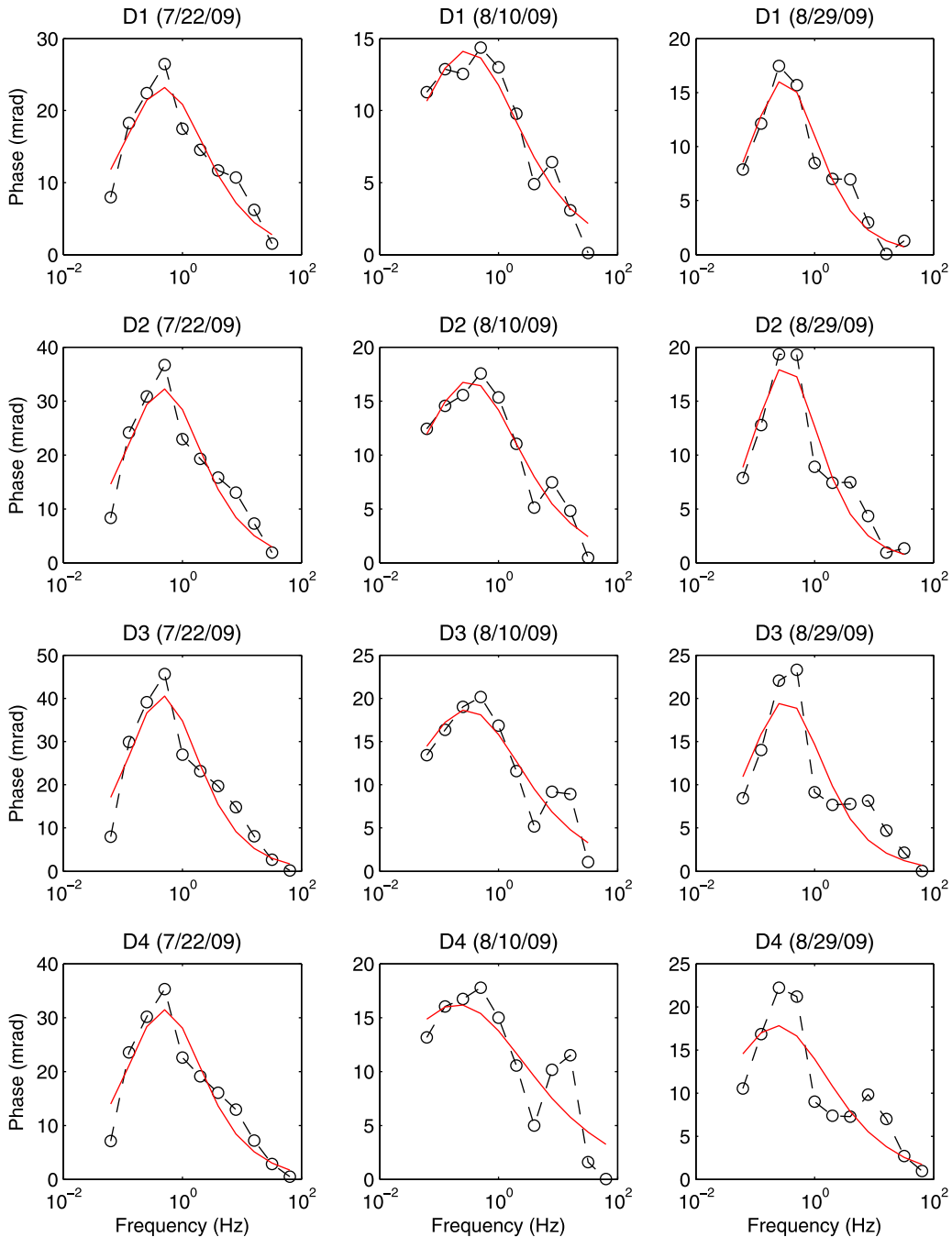

Figure 3. Comparison between the phase data and the calculated phases using the estimated medians of Cole-Cole parameters at the four geochemical sampling locations for the three time steps. The red lines are the fitted models and the circles are the phase data.

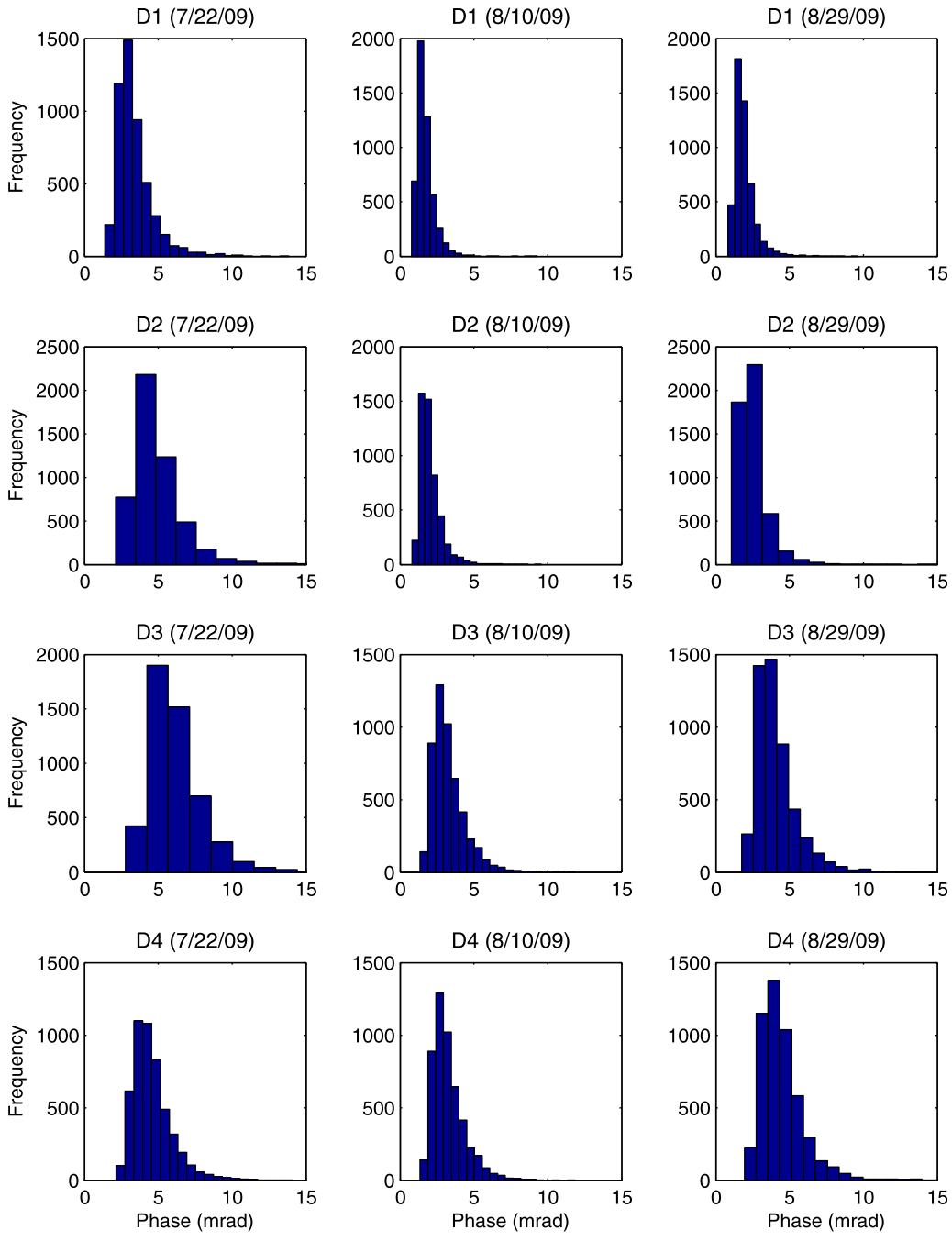

Figure 4. Histograms of the root-mean-squares (RMS) of the differences between the measured and calculated phases for 5000 random samples after Markov chains are converged.

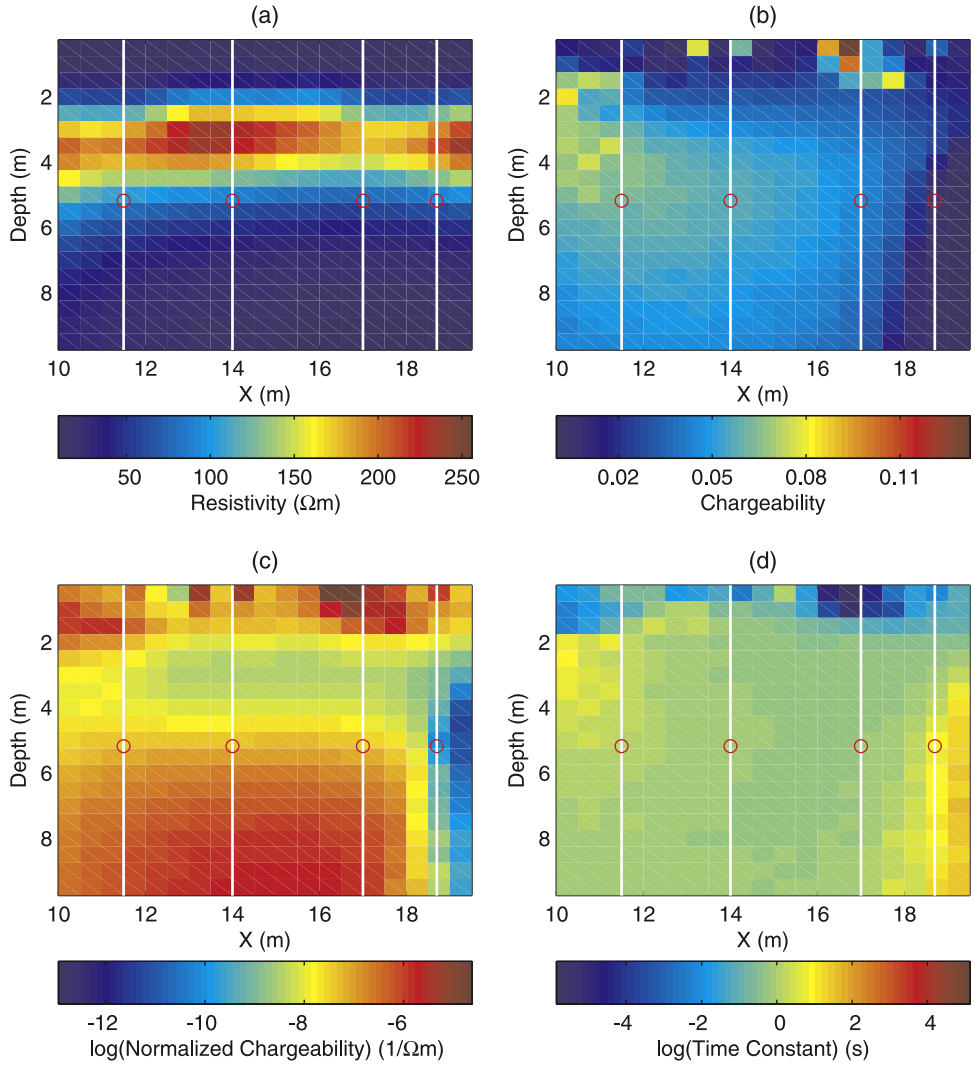

Figure 5. Estimated Cole-Cole parameters (a) DC resistivity ρ (Ωm), (b) chargeability m, (c) normalized chargeability (1/Ωm, log-scale), and (d) time constant τ (s, log-scale) along the cross section using SIP data collected on 10 August 2009. The white vertical segments show boreholes D1, D2, D3, and D4 from the right to the left, respectively. The red circles show the sampling locations.

the entire 2-D cross section. As an example, Figure 5 shows the estimated DC resistivity, chargeability, normalized chargeability, and time constant along array A using the SIP data collected on 10 August 2009. The white vertical line segments show the locations of boreholes D1, D2, D3, and D4 from the right to the left, respectively; the red circles show the groundwater sampling locations within the boreholes.

2.3. Borehole Aqueous Geochemistry Data and Time Series Analysis

[15] Groundwater samples were collected for geochemistry analysis, including Fe(II), sulfate, sulfide, acetate, uranium, chloride, and bromide concentrations, at ∼5 m depth within the four boreholes, starting from 10 July 2009 and ending on 8 December 2009. The fluid samples are representative of the groundwater conditions in the range between approximately 0.15 m above and below the sampling locations [*Williams et al.*, 2011]. Figure 6 shows the time series of logarithmic concentrations of Fe(II) in the unit of mg L^{-1} (black curves with triangles) collected from boreholes D1, D2, D3, and D4, respectively, where day 0 corresponds to 10 July 2009 and the initiation of the biostimulation is on day 12.

[16] Time series analysis was performed for each of the geochemical constituents quantified at the four borehole sampling locations. As an example, Figure 6 shows the fitting of Fe(II) concentrations at each of boreholes D1, D2, D3, and D4 using the second-order autoregressive (AR) model according to the Akaike Information Criterion (AIC) [*Akaike*, 1973], which is a good measure of the relative goodness of statistical model fit. The criterion makes a tradeoff between the accuracy and complexity of the model. The red curves in the figure show the calculated logarithmic Fe(II) concentrations using the estimated coefficients of the AR(2) model, whose values and standard errors are given in Table 1. On the basis of the fitting (Figure 6) and the coefficients (Table 1), we can see that

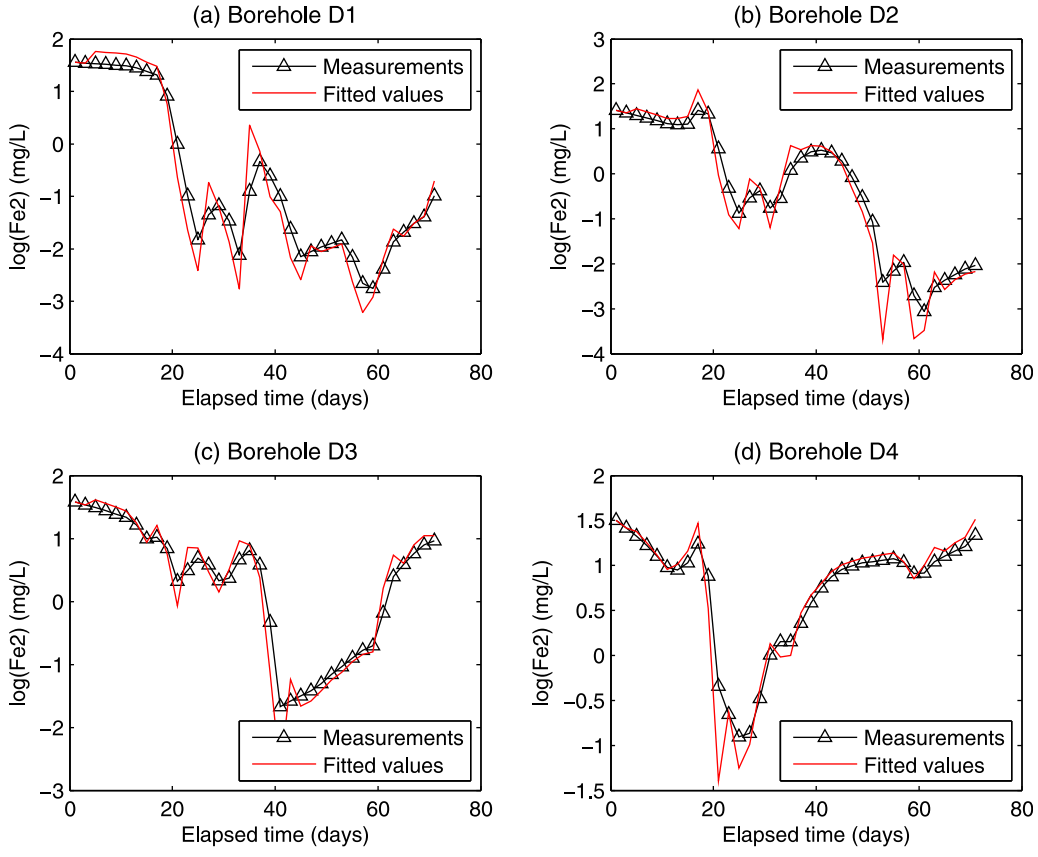

Figure 6. Data misfits of log(Fe(II)) (mg L^{-1}) at (a) borehole D1, (b) borehole D2, (c) borehole D3, and (d) borehole D4, using the second-order autoregressive model (AR(2)). The elapsed time 0 corresponds to 10 July 2009.

even if the time series of Fe(II) concentrations look different at different boreholes, especially at later time, they can all be described using AR(2) models. Therefore, we will use this temporal model for our estimation.

2.4. Cross-Correlation Between the Inverted Cole-Cole Parameters and Borehole Geochemistry Measurements

[17] The surface geophysical and borehole geochemical data complement each other in terms of spatial and temporal coverage and resolution. Geophysical data provide 2-D information along the entire cross section (rather than at a few locations) but are indirect and only available on three time steps. Borehole geochemical data were collected every few days throughout the entire experiment period (rather than at a few time steps); these data are direct but only available at four sampling locations. Since the emplacement of boreholes may have modified local field conditions, the obtained geochemical and geophysical data may both be subject to unknown biases or uncertainty. In addition, since the 2-D geophysical data were obtained from a regularized inversion, the estimated results may be constrained by the smoothing.

Table 1. Estimated Coefficients of AR(2) Models Using Fe(II), Sulfate, and Sulfide Data at the Four Borehole Sampling Locations

Geochemical Data	Boreholes	Beta-1	Beta-2	Sigma Squares
Fe(II)	D1	1.1842	−0.2900	0.3331
	D2	1.1391	−0.2434	0.3324
	D3	1.2385	−0.3584	0.1619
	D4	1.2499	−0.4126	0.0825
Sulfate	D1	1.4402	−0.5987	0.0974
	D2	1.2345	−0.3485	0.1653
	D3	1.3653	−0.5900	0.0340
	D4	1.2496	−0.5649	0.0518
Sulfide	D1	0.9282	−0.1903	2.7106
	D2	0.7978	−0.2778	3.3594
	D3	0.4517	0.0549	1.3664
	D4	1.1579	−0.2767	0.7418

[18] To explore the value of integrating geochemical and geophysical data sets for improved understanding of the changes in groundwater chemistry following the onset of biostimulation, we first examine the correlation between the geochemical data and the inverted Cole-Cole parameters that are "colocated" (or closely related) at four sampling locations and contemporary at three different times.

[19] Table 2 summarizes the site-specific correlations among various Cole-Cole parameters and geochemical measurements. Overall, chargeability shows good correlations with aqueous geochemical data and, in particular, has a good positive correlation with Fe(II) (cor = 0.73) and a negative correlation with sulfide (cor = −0.77). These trends are consistent with the field observations of *Flores*

Table 2. Correlation Coefficients Between Cole-Cole Parameters and Geochemical Data Based on Colocated Data at the Four Borehole Sampling Locations

Aqueous Geochemistry	Resistivity	Chargeability	Normalized Chargeability	Time Constant
Fe(II)	0.5186	0.7321	0.6321	−0.2562
Sulfide	−0.3356	−0.7751	−0.7843	0.1485
Acetate	−0.6522	−0.6722	−0.4846	0.0300
Bromide	−0.5291	−0.7423	−0.6389	0.2461
Sulfate	0.5466	0.8336	0.7354	−0.1243
Chloride	0.3184	0.7832	0.7944	−0.2263
Uranium	0.3811	0.8441	0.8401	−0.2383

Orozco et al. [2011] and the laboratory studies of *Williams et al.* [2009]. Although we have only a total number of 12 data points for the analysis, the consistence with other studies gives us confidence to apply a linear relationship to link Cole-Cole parameters to geochemical data in the later study.

3. Hierarchical Spatiotemporal Bayesian Model

[20] We develop a Bayesian model based on the results of data analysis presented in section 2 to estimate the spatiotemporal distributions of Fe(II), sulfate, and sulfide concentrations using the chargeability data, as those geochemical parameters are important indicators of the status of the bioremediation treatment. Although the chargeability data may be linked to those geochemical parameters as a multivariate variable, in this study we choose to estimate each of the geochemical parameters separately from the chargeability data because we cannot derive a reliable multivariate relationship between chargeability and geochemical data given the limited colocated and contemporary data points.

[21] We follow the hierarchical approach outlined by *Wikle et al.* [2003] and break the complex spatial and temporal Bayesian model into three simpler statistical submodels. They include: (1) data model, which links the chargeability data to geochemical parameters; (2) process model, which describes the spatial and temporal variability of the geochemical properties; and (3) parameter model, which describes the prior distributions of various parameters and initial conditions. Figure 7 is a flowchart showing the detailed procedures of the developed methodology. We use Markov chain Monte Carlo (MCMC) sampling methods to obtain the spatial and temporal distribution of the unknown geochemical properties and other associated parameters.

[22] We use regression-based statistical models, rather than physically based petrophysical models as used by *Karaoulis et al.* [2011] and *Revil and Skold* [2011] to link geophysical properties to geochemical parameters due to the complex field conditions. The development of statistical relationships is built upon extensive recent laboratory and field research that has demonstrated empirical relationships between chargeability and geochemical variables

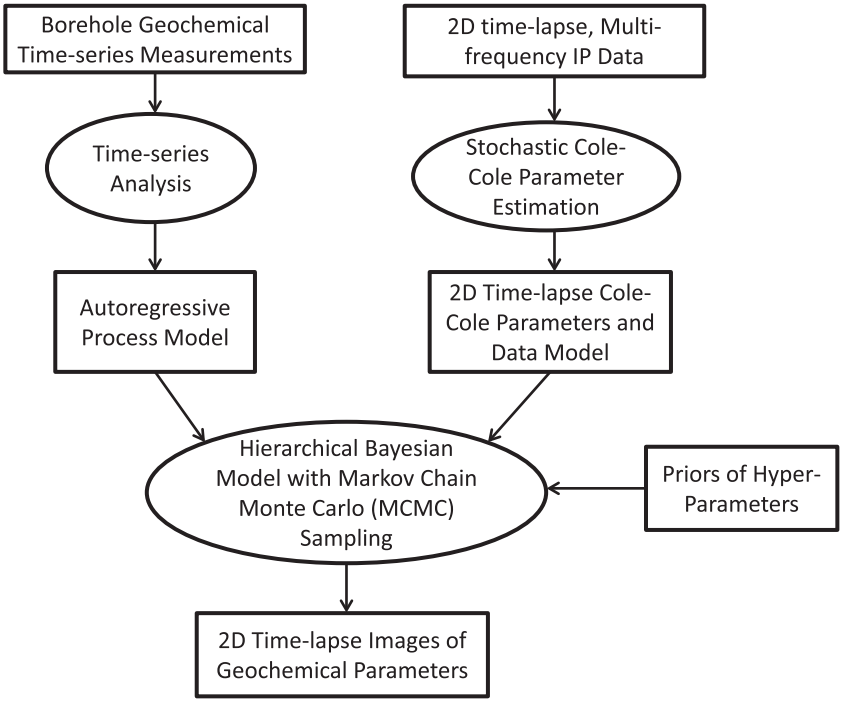

Figure 7. Flowchart showing the procedures of our data processing and hierarchical Bayesian model.

[*Ntargialiannis et al.*, 2005; *Slater et al.*, 2006, 2007; *Personna et al.*, 2008; *Williams et al.*, 2005, 2009, 2011; *Chen et al.*, 2009; *Wu et al.*, 2011; *Flores Orozco et al.*, 2011]. We acknowledge that the use of regression-based statistical models is a limitation of the current study because they are site specific and may be biased. Finding physically based petrophysical models is an active area of research in the community, and it is not the focus of this paper. However, the developed methodology can be extended to the cases where physically based models are available.

3.1. Data Model

[23] The data model links the inverted chargeability to geochemical concentrations at each pixel. Although we may use the correlation coefficients obtained through regression of the colocated data, we consider them as unknowns for more general cases. Let $u_t(s)$ be the unknown concentration at grid $s \in D = \{1, 2, \ldots, m\}$, where m is the total number of sites, and at time step $t \in T = \{1, 2, \ldots, n\}$, where n is the total number of time steps. Let $m_t^{\text{obs}}(s)$ be the chargeability data, where $s \in D$ and $t \in T_g$, which is a subset of set T and represents the three SIP survey days. From regression analysis of the colocated and contemporary geochemical and geophysical data, we can assume that there is a linear relationship between the aqueous geochemical and chargeability data. The linear assumption is derived from field data rather than from theory. At the early stage of our understanding of the SIP responses to geochemical heterogeneity, we feel more comfortable to use field-derived empirical relationships instead of theoretical relationships for the estimation. Consequently, we have $m_t^{\text{obs}}(s) = \alpha_1 + \alpha_2 u_t(s) + \varepsilon_m$, where α_1 and α_2 are unknown coefficients, and ε_m is the random error that accounts for uncertainty from multiple sources, such as errors in the petrophysical model and in the inverted chargeability data.

[24] Since the chargeability data were obtained from fitting the inverted surface IP data, which typically have a spatially variable resolution, we may allow the coefficients (i.e., α_1 and α_2) to be varied over the space. However, in this study, because we only use the data on a small subdomain, ranging from $x = 10$ m to $x = 20$ m and $z = 2$ m to $z = 6$ m, on which the coefficients of variation of sensitivity are 1% laterally and 5% vertically, we ignore such variations. Therefore, we assume that both coefficients α_1 and α_2 are same at each pixel and the errors at different locations are independent. We address the spatial variability by the additive error ε_m, which is assumed to have a Gaussian distribution with the inverse variance of τ_m.

[25] We can describe the data model using vectors for conciseness. Let $\mathbf{m}_t^{\text{obs}} = \{m_t^{\text{obs}}(s),\ s \in D\}$ and $\mathbf{u}_t = \{u_t(s),\ s \in D\}$. We thus have the following conditional probability distribution for geophysical data:

$$[\{\mathbf{m}_t^{\text{obs}},\ t \in T_g\} | \{\mathbf{u}_t, t \in T_g\}, \alpha_1, \alpha_2, \tau_m] = \prod_{t \in T_g} [\mathbf{m}_t^{\text{obs}} | \mathbf{u}_t, \alpha_1, \alpha_2, \tau_m], \tag{1}$$

where we use the bracket to denote probability distribution in equation (1) following the annotation provided by *Gelfand and Smith* [1990]. We can similarly obtain a data model for borehole geochemical measurements. However, because the errors in the borehole data are typically much smaller than those in the regression model in equation (1), we consider them as true values at the borehole locations in this study. Let $u_t^{\text{obs}}(s)$ be the direct measurements of geochemical concentrations at time $t \in T$ and site $s \in D_b$, which is a subset of set D and represents the four borehole sampling locations. We thus have $u_t(s) = u_t^{\text{obs}}(s)$.

3.2. Process Model

[26] We use a statistical rather than mechanistic (or reactive transport) model to simulate the evolution of geochemical processes. Since we can fit the geochemical time series using the second order autoregressive models as suggested by data analysis (see Table 1), we model the evolution processes using the following relationship: $u_t(s) = \beta_1(s) u_{t-1}(s) + \beta_2(s) u_{t-2}(s) + \varepsilon_u(s)$, where $\beta_1(s)$ and $\beta_2(s)$ are the coefficients at site $s \in D$. The spatially variable coefficients allow us to take account for the heterogeneity of geochemical properties as recognized by *Li et al.* [2009]. We consider $\varepsilon_u(s)$ as a stationary spatial process, and it has the inverse variance of τ_{pu} and the correlation matrix determined by the following exponential variogram function:

$$r(s_i, s_j) = \exp\left\{ -\sqrt{\left(\frac{x(s_i) - x(s_j)}{\lambda_x}\right)^2 + \left(\frac{z(s_i) - z(s_j)}{\lambda_z}\right)^2} \right\}, \tag{2}$$

where $(x(s_i), z(s_i))$ and $(x(s_j), z(s_j))$ are the 2-D coordinates of s_i and s_j grids, respectively. Symbols λ_x and λ_z are the spatial correlation lengths along lateral and vertical directions. Since we do not have direct information to determine the spatial correlation lengths at this time, we assume that they have similar spatial structure to the permeability field. Therefore, we pick $\lambda_x = 2$ m and $\lambda_z = 0.5$ m, derived from the permeability fields of *Englert et al.* [2009] and *Li et al.* [2010]. Let $\mathbf{R} = \{r(s_i, s_j),\ s_i, s_j \in D\}_{m \times m}$ be the known correlation matrix and let vector $\boldsymbol{\varepsilon}_u = (\varepsilon_u(1),\ \varepsilon_u(2), \ldots, \varepsilon_u(m))^T$. We assume that $\boldsymbol{\varepsilon}_u$ has the multivariate Gaussian distribution with zero mean and the inverse covariance matrix of $\tau_{pu}\mathbf{R}^{-1}$, i.e., $\boldsymbol{\varepsilon}_u \sim N(0,\ \tau_{pu}\mathbf{R}^{-1})$.

[27] We can simplify the process model by letting $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ be the diagonal metrics with $\beta_1(s)$ and $\beta_2(s)$, $s \in D$, being the corresponding diagonal terms. Thus, we have $\mathbf{u}_t = \boldsymbol{\beta}_1 \mathbf{u}_{t-1} + \boldsymbol{\beta}_2 \mathbf{u}_{t-2} + \boldsymbol{\varepsilon}_u$. Using the same notation, we assume that the initial geochemical concentrations $\mathbf{u}_1$ and $\mathbf{u}_2$ have multivariate Gaussian distributions, i.e., $\mathbf{u}_1 \sim N(\mu_{u1}\mathbf{e},\ \tau_{u1}\mathbf{R}^{-1})$ and $\mathbf{u}_2 \sim N(\mu_{u2}\mathbf{e},\ \tau_{u2}\mathbf{R}^{-1})$, where $\mathbf{e} = (1,\ 1,\ \ldots,\ 1)^T_{m \times 1}$ and μ_{u1}, μ_{u2}, τ_{u1}, and τ_{u2} are hyperparameters having priors determined by borehole logs. By combining the above information, we have a spatial and temporal process model for geochemical parameters as given below:

$$\begin{aligned}[\{\mathbf{u}_t, t \in T\} | \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}, \mu_{u1}, \tau_{u1}, \mu_{u2}, \tau_{u2}] \\ = [\mathbf{u}_1 | \mu_{u1}, \tau_{u1}][\mathbf{u}_2 | \mu_{u2}, \tau_{u2}] \prod_{k=3}^{n} [\mathbf{u}_k | \mathbf{u}_{k-1}, \mathbf{u}_{k-2}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}].\end{aligned} \tag{3}$$

[28] By combining the data (equation (1)) and process (equation (3)) models, we obtain the following full joint posterior distribution:

$$\begin{aligned}&\left[\{\mathbf{u}_t, t\in T\}, \alpha_1, \alpha_2, \tau_m, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}, \mu_{u1}, \tau_{u1}, \mu_{u2}, \tau_{u2} | \{\mathbf{m}_t^{\text{obs}}, t\in T_g\}\right]\\ &\propto \left[\{\mathbf{m}_t^{\text{obs}}, t\in T_g\} | \{\mathbf{u}_t, t\in T_g\}, \alpha_1, \alpha_2, \tau_m\right]\\ &[\mathbf{u}_1|\mu_{u1}, \tau_{u1}][\mathbf{u}_2|\mu_{u2}, \tau_{u2}] \times \prod_{k=3}^{n}\left[\mathbf{u}_k|\mathbf{u}_{k-1}, \mathbf{u}_{k-2}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}\right]\\ &\left[\alpha_1, \alpha_2, \tau_m, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}, \mu_{u1}, \tau_{u1}, \mu_{u2}, \tau_{u2}\right].\end{aligned} \tag{4}$$

3.3. Priors on the Parameters

[29] We need to specify the joint prior distribution (last term in equation (4)) so that the joint posterior probability distribution is defined. First, we assume that each of those parameters is independent of others and thus the joint prior can be written as the product of each individual prior. Second, we assume that each of those parameters is uniformly distributed on the ranges determined from borehole geochemical and surface IP data and the results are given in Table 3. For coefficients α_1 and α_2, we first fitted the colocated and contemporary chargeability and geochemical data and then use the mean values subtracting two-times standard deviation as lower bounds and the mean values adding two-times standard deviation as upper bounds. Similarly, we can obtain the lower and upper bounds of other parameters.

3.4. MCMC Sampling Methods

[30] We use the Gibbs sampler [*Gelman and Gelman*, 1984] to draw many samples from the joint posterior distribution given in equation (4) because we can derive full conditionals of each unknown. In the following we only show the derivation of the full conditional probability distributions of $\mathbf{u}_t$ and $\boldsymbol{\beta}_1$; the full conditionals of other parameters are provided in Appendix A.

[31] We adopt a block sampling strategy for $\mathbf{u}_t$, $t \in T$. If $t \in T_g$, the unknown vector $\mathbf{u}_t$ is related to the inverted chargeability vector $\mathbf{m}_t^{\text{obs}}$, together with the unknown vectors $\mathbf{u}_{t-2}$, $\mathbf{u}_{t-1}$, $\mathbf{u}_{t+1}$, and $\mathbf{u}_{t+2}$ through the process model. By dropping all the terms that are not directly related to $\mathbf{u}_t$ in equation (4), we have

$$\begin{aligned}[\mathbf{u}_t|\cdot] &\propto [\mathbf{m}_t^{\text{obs}}|\mathbf{u}_t, \alpha_1, \alpha_2, \tau_m][\mathbf{u}_t|\mathbf{u}_{t-2}, \mathbf{u}_{t-1}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}]\\ &\times [\mathbf{u}_{t+1}|\mathbf{u}_{t-1}, \mathbf{u}_t, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}][\mathbf{u}_{t+2}|\mathbf{u}_t, \mathbf{u}_{t+1}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}].\end{aligned} \tag{5}$$

If $t \notin T_g$, vector $\mathbf{u}_t$ is not related to the chargeability data and we just drop it from equation (5). Since both the data and process models have multivariate Gaussian distributions, the conditional distribution of $\mathbf{u}_t$ also has a multivariate Gaussian distribution as shown by *Carlin et al.* [1992] with the mean vector of $\boldsymbol{\Sigma}_u \mathbf{v}_t$ and the covariance matrix of $\boldsymbol{\Sigma}_u$, where

$$\boldsymbol{\Sigma}_u^{-1} = \alpha_2^2 \tau_m \mathbf{I}_m + \tau_{pu}\mathbf{R}^{-1} + \tau_{pu}\boldsymbol{\beta}_1^T\mathbf{R}^{-1}\boldsymbol{\beta}_1 + \tau_{pu}\boldsymbol{\beta}_2^T\mathbf{R}^{-1}\boldsymbol{\beta}_2, \tag{6}$$

$$\begin{aligned}\mathbf{v}_t &= \alpha_2\tau_m(\mathbf{m}_t^{\text{obs}} - \alpha_1\mathbf{e}) + \tau_{pu}\mathbf{R}^{-1}(\boldsymbol{\beta}_1\mathbf{u}_{t-1} + \boldsymbol{\beta}_2\mathbf{u}_{t-2})\\ &+ \tau_{pu}\boldsymbol{\beta}_1\mathbf{R}^{-1}(\mathbf{u}_{t+1} - \boldsymbol{\beta}_2\mathbf{u}_{t-1}) + \tau_{pu}\boldsymbol{\beta}_2\mathbf{R}^{-1}(\mathbf{u}_{t+2} - \boldsymbol{\beta}_1\mathbf{u}_{t+1}).\end{aligned} \tag{7}$$

[32] We sample each component of $\boldsymbol{\beta}_1$, i.e., $\beta_1(s)$, $s \in D$, individually. Let $\mathbf{L}$ be the lower triangle matrix obtained from the Cholesky decomposition of the correlation matrix, i.e., $\mathbf{R} = \mathbf{L}\mathbf{L}^T$. Thus, $\mathbf{R}^{-1} = (\mathbf{L}^{-1})^T\mathbf{L}^{-1}$. Let $\mathbf{Q} = \mathbf{L}^{-1} = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m)$, where $\mathbf{q}_i$ is the ith column of matrix $\mathbf{Q}$. By dropping all the terms that are not directly related to $\beta_1(s)$ in equation (4) and assuming $\beta_1(s)$ is uniform on $(a_{\beta_1}, b_{\beta_1})$, we have

$$\begin{aligned}[\beta_1(s)|\cdot] &\propto I(\beta_1(s)\in(a_{\beta_1}, b_{\beta_1}))\prod_{k=3}^{n}[\mathbf{u}_k|\mathbf{u}_{k-1}, \mathbf{u}_{k-2}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \tau_{pu}]\\ &\propto I(\beta_1(s)\in(a_{\beta_1}, b_{\beta_1}))\exp\left\{-0.5\tau_{pu}\sum_{k=3}^{n}(\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2})^T\right.\\ &\left.\mathbf{R}^{-1}(\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2})\right\}\\ &\propto I(\beta_1(s)\in(a_{\beta_1}, b_{\beta_1}))\exp\left\{-0.5\tau_{pu}\sum_{k=3}^{n}(\mathbf{Q}(\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2}))^T\right.\\ &\left.(\mathbf{Q}(\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2}))\right\}\\ &\propto I(\beta_1(s)\in(a_{\beta_1}, b_{\beta_1}))\exp\left\{-0.5\tau_{pu}\sum_{k=3}^{n}(\mathbf{c}_k(s)\beta_1(s) - \mathbf{d}_k(s))^T\right.\\ &\left.((\mathbf{c}_k(s)\beta_1(s) - \mathbf{d}_k(s))\right\}\\ &\propto N_T(\mu_{\beta_1(s)}, \tau_{\beta_1(s)}, a_{\beta_1}, b_{\beta_1}),\end{aligned} \tag{8}$$

where

$$\begin{aligned}\tau_{\beta_1(s)} &= \sum_{k=3}^{n}\mathbf{c}_k^T(s)\mathbf{c}_k(s),\\ \mu_{\beta_1(s)} &= \sum_{k=3}^{n}\mathbf{c}_k^T(s)\mathbf{d}_k(s)/\tau_{\beta_1(s)},\end{aligned} \tag{9}$$

Table 3. Prior Ranges of Various Model Parameters for Fe(II), Sulfate, and Sulfide Estimation Based on Data From Boreholes

	Log(Fe(II))		Log(Sulfate)		Log(Sulfide)	
Parameters	Lower Bounds	Upper Bounds	Lower Bound	Upper Bounds	Lower Bounds	Upper Bounds
α_1	−1.2376	−1.1198	−1.2874	−1.1897	−1.2319	−1.0630
α_2	0.0179	0.0838	0.0579	0.1459	−0.0798	−0.0086
$\boldsymbol{\beta}_1$	1.1335	1.2555	1.2242	1.4505	0.4447	1.1649
$\boldsymbol{\beta}_2$	−0.4211	−0.2349	−0.6112	−0.3360	−0.2812	0.0582
μ_{u1}	1.4017	1.5874	−0.0116	1.7410	−3.9598	0.9106
μ_{u2}	1.3493	1.5486	0.2482	1.7830	0.2757	1.6297

$$\mathbf{c}_k(s) = u_{k-1}(s)\mathbf{q}_s,$$
$$\mathbf{d}_k(s) = \sum_{i=1}^{m}(u_k(i) - \beta_2(i)u_{k-2}(i))\mathbf{q}_i - \sum_{i\neq s}^{m}\beta_1(i)u_{k-1}(i)\mathbf{q}_i, \quad (10)$$

and $I()$ is the indicator function with the value of 1 if the condition is satisfied and 0 otherwise. The symbol N_T represents the truncated normal distribution.

4. Estimation Results

[33] We apply the developed Bayesian model and sampling strategy to estimate the spatiotemporal distribution of Fe(II), sulfate, and sulfide concentrations. Since the methodology and procedures for estimating different types of geochemical parameters are similar, we only show the details for Fe(II) estimation but provide results for estimation of other parameters.

4.1. Priors and Sensitivity of Model Parameters

[34] Since the total number of unknowns is much larger than the total number of data points, we use informative priors based on the results of data analysis. We set the lower and upper bounds of α_1 and α_2 as their corresponding regression estimates, minus and plus two-times standard errors. We set the lower and upper bounds of $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, μ_1, and μ_2 as the minimum and maximum values obtained from four boreholes with extension by 5% from their original ranges. Table 3 lists the actual lower and upper bounds used in the study for the Fe(II), sulfate, and sulfide estimation.

[35] Figures 8(a)–8(d) shows the posterior probability densities of α_1, α_2, μ_{u1}, and μ_{u2}, relative to their corresponding priors. We can see that the posterior distributions of those parameters do not depend on their corresponding prior ranges given in Table 3; all the posterior distributions get updated by conditioning to the borehole and the 2-D SIP data. Figure 9 shows the estimated medians of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ along the cross section. We can see that those coefficients are not equal at different sites, which is consistent with our assumption that the AR(2) model parameters vary over space. In addition, we can see that most medians are close to the mean of the priors; their posterior distributions are close to the corresponding priors. This is because at the sites away from borehole sampling ports we only have SIP data at three different time steps. Better resolution of those parameters would require SIP data collection at more frequent time steps.

[36] We treat the four inverse variances (i.e., τ_{u1}, τ_{u2}, τ_{pu}, and τ_m) differently because they play different roles in the estimation. We set the lower and upper bounds of τ_{u1} and τ_{u2} as 0.1 and 100, both of which control the spatial variation of the initial concentrations of Fe(II). We set the lower and upper bounds of τ_{pu} as the minimum and maximum of values at the four boreholes and set the lower and upper bounds of τ_m as 0.1 and 100, which is the value obtained from regression analysis. We found that the estimates of parameters are insensitive to the priors of τ_{u1}, τ_{u2}, and τ_{pu} as shown in Figures 8(e)–8(g). However, the estimation results are very sensitive to the upper bound of τ_m and the posterior estimate of τ_m always approaches the upper limit that we set. This is because at each site away from the boreholes we have only three chargeability values but more than three parameters to estimate. Therefore, for the current study, we fix the value of τ_m as the one directly obtained from the regression analysis of the colocated geochemical and chargeability data. We consider it as unknown only when SIP data are available at more time steps.

4.2. Cross Validation

[37] Given the fact that we only have geochemical measurements at four sampling ports and SIP data at three time steps, it is difficult to perform rigorous statistical cross validation of our geophysically obtained estimates of groundwater geochemical concentrations. In this study we just apply the leave-one-out method [*Kohavi*, 1995] to provide an indication of the estimation procedure accuracy. We use geochemical data at three boreholes to estimate geochemical concentrations at the fourth borehole, and compare the estimated results with the borehole measurements at that location. Notice that at each validation location, geophysical data are always available at the three monitoring time steps. We first compare the estimated results at the three time steps when the geophysical data are available, and this provides a test of the consistency of the petrophysical models. Figure 10 compares the true measurements (circles) with the estimated 95% high probability domain (HPD) (vertical solid line segments). We can see that most of the borehole measurements are within the predictive intervals.

[38] We also compare the estimated results with the borehole measurements over the entire injection period, which includes the time steps with and without geophysical data. Figure 11 shows the estimated medians (red triangles), the 95% lower and upper bounds (blue dashed lines), and the borehole measurements (black circles). In general, the estimated medians follow the main trends of the borehole measurements. Since about 83% of the Fe(II) borehole measurements are in the 95% predictive bounds, our estimated uncertainties are little over-optimistic. This possibly is caused by the small variance that we used in the data model. As discussed in section 4.1, we can resolve this problem when SIP data are available at more time steps.

[39] As a comparison, we also put the estimates (i.e., green lines with squares) of geochemical data based on ordinary kriging [*Kitanidis*, 1997] on Figure 11 by following the leave-one-out cross-validation procedure. Similar to the Bayesian model developed in section 2, we use the exponential variogram with the vertical correlation length of 0.5 m, the lateral correlation length of 2 m, and the estimated median of the inverse variance τ_{u2} ($\sim$3). The kriging estimates at each location are equal to the linear combination of the values at other three locations. As we can see from Figure 11, the kriging estimates of Fe(II) at D1 is very close to the borehole measurements at D2, and the kriging estimate at D4 is close to the borehole measurements at D3. Overall, the medians of the Bayesian estimates are better than those obtained from ordinary kriging. To measure the goodness of predicting the measurements quantitatively at the testing wells, we calculate the root-mean-squares (RMS) of the differences between the estimated results and the measurements. The RMS for the Bayesian and kriging estimates are 0.83 and 1.18, respectively. One of the main drawbacks of kriging is that it leads to over-smoothed 2-D images because the lateral correlation length

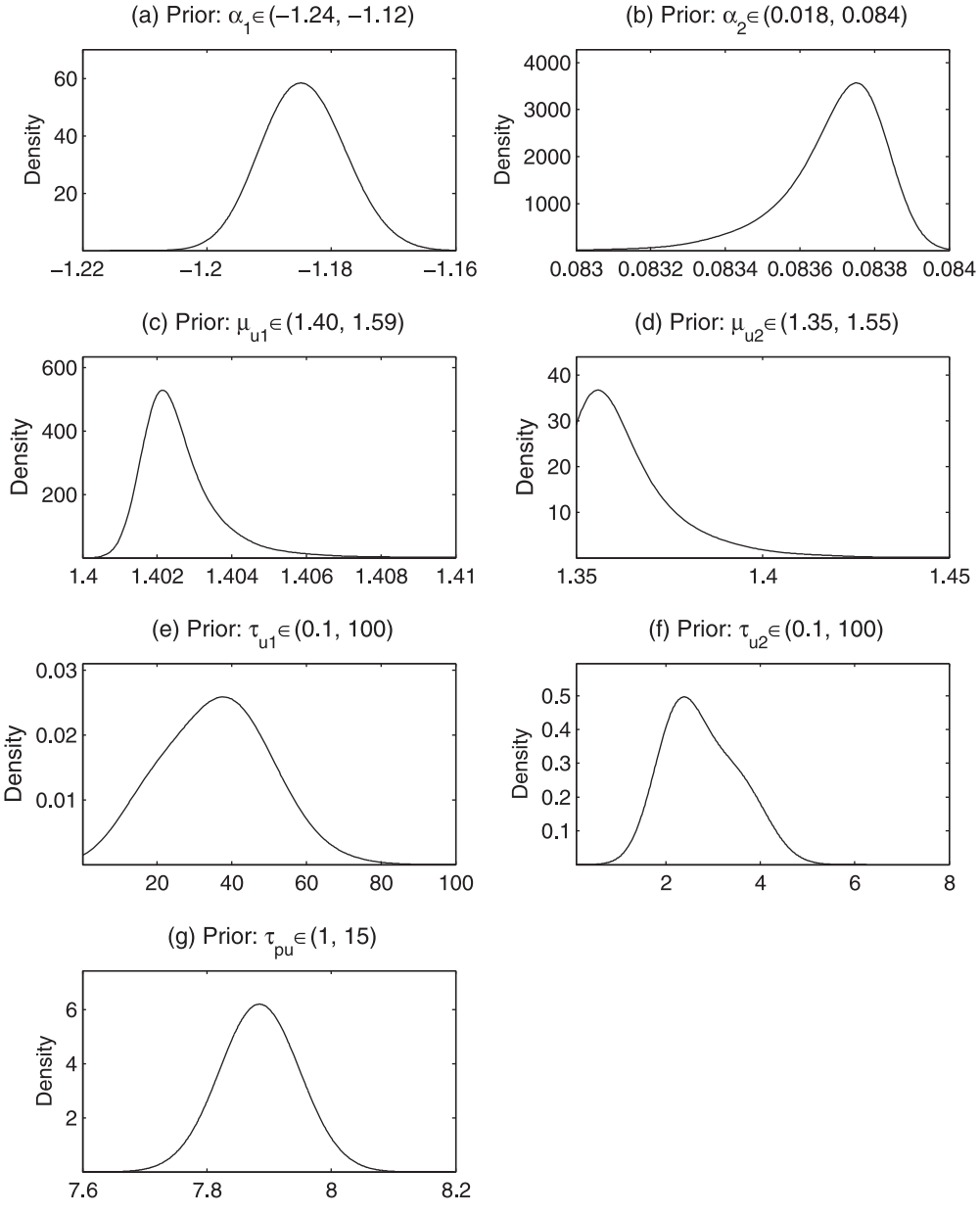

Figure 8. Estimated posterior probability densities of hyperparameters.

is 2 m, which is just 1/5 of the horizontal length of the target domain. As shown in Figure 12, the kriging estimates yield an almost spatially uniform distribution of estimated Fe(II) over the 2-D domain.

4.3. Estimated Spatiotemporal Distribution of Fe(II) Concentrations

[40] The developed Bayesian model allows us to estimate the spatial distribution of Fe(II) concentrations at each time step by conditioning to the 2-D geophysical data and the time series borehole geochemical measurements. Figure 13 shows the medians of Fe(II) concentrations from day 0 (starting of injection) to day 38 (after injection), obtained using the developed Bayesian procedure and both geochemical and geophysical data sets. The 2-D geophysical data are temporally sparse and available only at three time steps (i.e., day 0, day 20, and day 38), and the geochemical borehole measurements are spatially sparse and available only at four sampling ports showing as circles in Figure 13. The estimated spatiotemporal distribution shows the evolution of Fe(II) concentrations over time during biostimulation. We can see that in the early stage of injection (before day 10), the Fe(II) concentrations are approximately uniform on the cross section. On day 14, the concentrations near the upper-left portion of the imaging region start increasing and become most apparent by day 20, a day when 2-D geophysical data are available. Starting on day 30, the Fe(II) concentrations near the right side of the domain decrease (blue colors) and at the end of injection, the concentration near the lower right region of the imaging domain is much lower relative to surrounding regions. The estimated results provide more information than the geochemical borehole data alone (i.e., Figure 12) for understanding and monitoring the field-scale bioremediation processes.

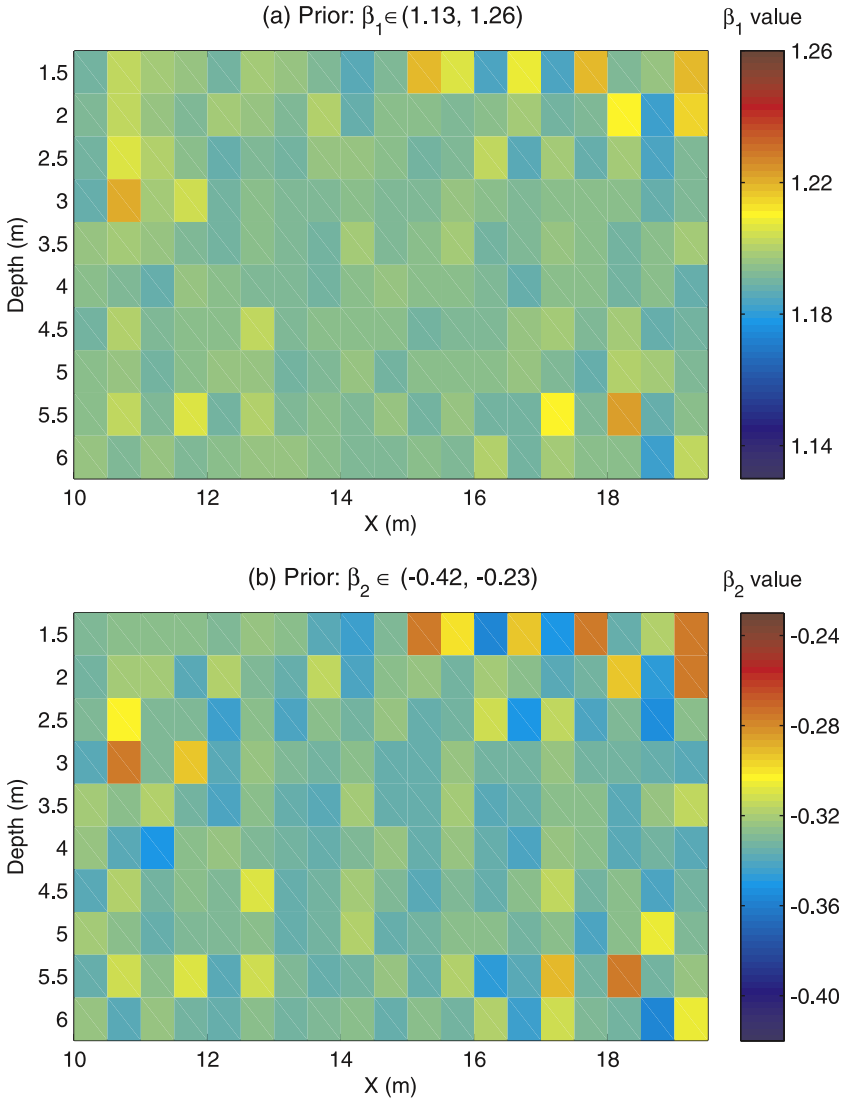

Figure 9. Estimated medians of the process model coefficients along the 2-D cross section: (a) β_1 and (b) β_2.

4.4. Estimated Spatiotemporal Distribution of Sulfate and Sulfide Concentrations

[41] To estimate the spatiotemporal distributions of sulfate and sulfide concentrations, we follow a similar procedure as was described for Fe(II). We first fit the geochemical wellbore time series data using AR(2) models; associated coefficients of those fits are also provided in Table 1. We can see that the variations in the AR(2) models of borehole sulfate measurements are very similar to those of Fe(II). But the variations of borehole sulfide data are significantly larger than those of Fe(II) and sulfate. The prior bounds for sulfate and sulfide estimation are given in Table 3. We additionally note that the spatial variations of initial sulfate and sulfide concentrations are significantly larger than those of Fe(II).

[42] Figures 14(a) and 14(b) show the cross validation of sulfate and sulfide during the time that geophysical data are available. We can see that the predictive ranges of sulfate concentrations are smaller than those of Fe(II) and most of the measured sulfate concentrations are near the upper bounds of the ranges. This may be caused by the large spatial variation in the initial sulfate concentrations. For sulfide we can see that the predictive ranges of sulfide concentrations are considerably larger than those of Fe(II) and sulfate and that the Bayesian model underestimates sulfide values for large concentrations. This is possibly caused by the large variance in fitting AR(2) models as shown in Table 1.

[43] Figures 15 and 16 show the cross validation for sulfate and sulfide, respectively, over the entire injection period. Again, the estimated medians of sulfate and sulfide follow the main trends of the measurements and slightly better than the results of kriging. For sulfide, the Bayesian estimates are considerably smaller than their corresponding borehole measurements.

[44] Figures 17 and 18 show the spatiotemporal distributions of sulfate and sulfide concentrations along the 2-D cross section, obtained using the developed Bayesian approach that honors both geochemical and geophysical data sets. In general, the spatial and temporal patterns of sulfate concentrations are similar to those of the Fe(II) concentrations, but

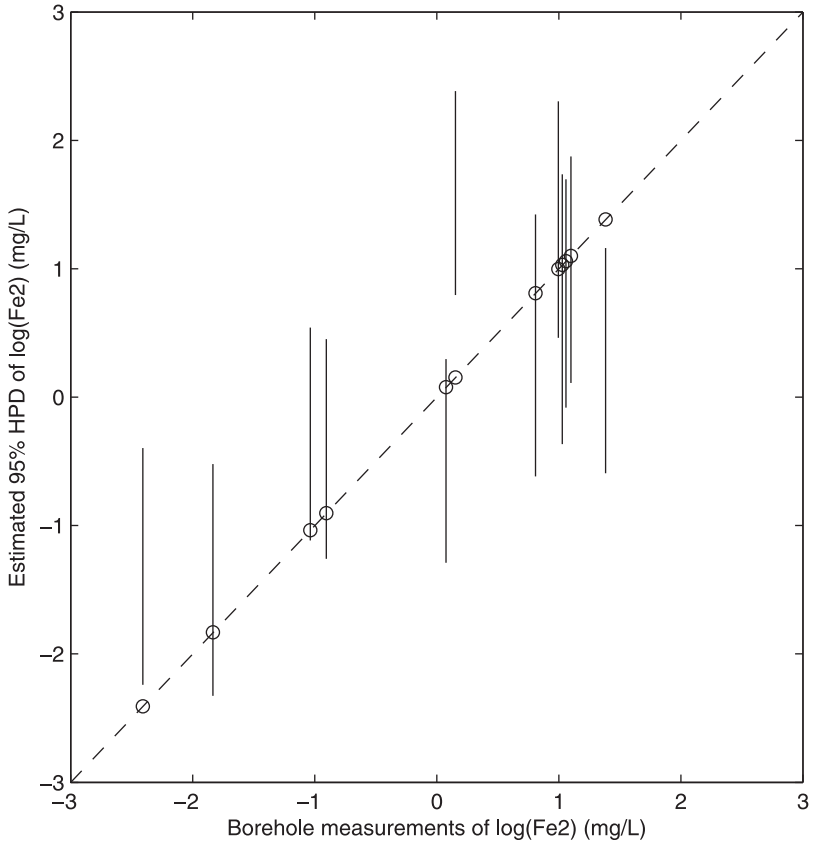

Figure 10. Cross validation: comparison between the estimated and measured log(Fe(II)) at the time step that SIP data are available. The vertical segments are the estimated 95% highest probability domains (HPD), and the circles are the borehole measurements.

the patterns of sulfide concentrations are opposite. Such an effect is expected, given that sulfate and sulfide—the metabolic by-product of sulfate reduction—are negatively correlated. Similarly, in locations where Fe(II) is in excess (e.g., upper left of Figure 13, images from days 16–38), threshold concentrations of sulfide will be largely removed from solution, as observed at similar locations in Figure 18. The figures reveal very detailed information about the spatial and temporal evolution of the geochemical groundwater species.

5. Discussion and Conclusions

[45] We have developed a hierarchical Bayesian model for estimating the spatiotemporal distribution of Fe(II), sulfate, and sulfide by integrating 2-D geophysical data and "point" geochemical borehole measurements. Geophysical data have large spatial coverage but are available only at a few time steps, whereas the borehole time series of aqueous geochemistry data are temporally dense but are available only at a few locations. We have shown that integration of the geophysical and geochemical data using the developed Bayesian approach significantly improves the estimates of the evolution of groundwater chemistry associated with a bioremediation experiment relative to what is available based on wellbore data alone. The developed Bayesian model is very flexible and can be applied to estimate other geochemical parameters, particularly those having a strong correlation with Cole-Cole parameters. The method can also be extended for use with 3-D SIP data.

[46] We chose to estimate aqueous geochemical concentrations of ferrous iron, sulfate, and sulfide because these species are commonly used to assess the onset and evolution of bioremediation treatments. This work advances previous estimation studies, where we focused on using SIP data to estimate the evolution of remediation induced solid phase end-products (such as FeS) and their impact on hydraulic parameters.

[47] The cross validation using the leave-one-out method shows that the majority of true measurements are within the 95% predictive intervals and the estimated medians approximately follow the main treads of the corresponding true values. But the point-to-point comparison of the estimated and measured time series is not impressive. Especially for sulfide estimation, the estimated results are substantially smaller than the corresponding borehole measurements. The spatial sparseness of the geochemical data and temporal sparseness of the geophysical data are interpreted to be the main culprits for this unsatisfactory comparison. However, the benefit of incorporating geophysical data in the estimation of geochemical parameters was substantial, as was illustrated through comparison of the Bayesian estimates (Figure 13) with kriging wellbore (Figure 12) estimates, the latter of which were almost spatially uniform and provided little information about the evolution of the groundwater chemistry following biostimulation.

[48] We made several key assumptions on the priors and error structures in the current Bayesian model based on the current available data sets and some may become less critical

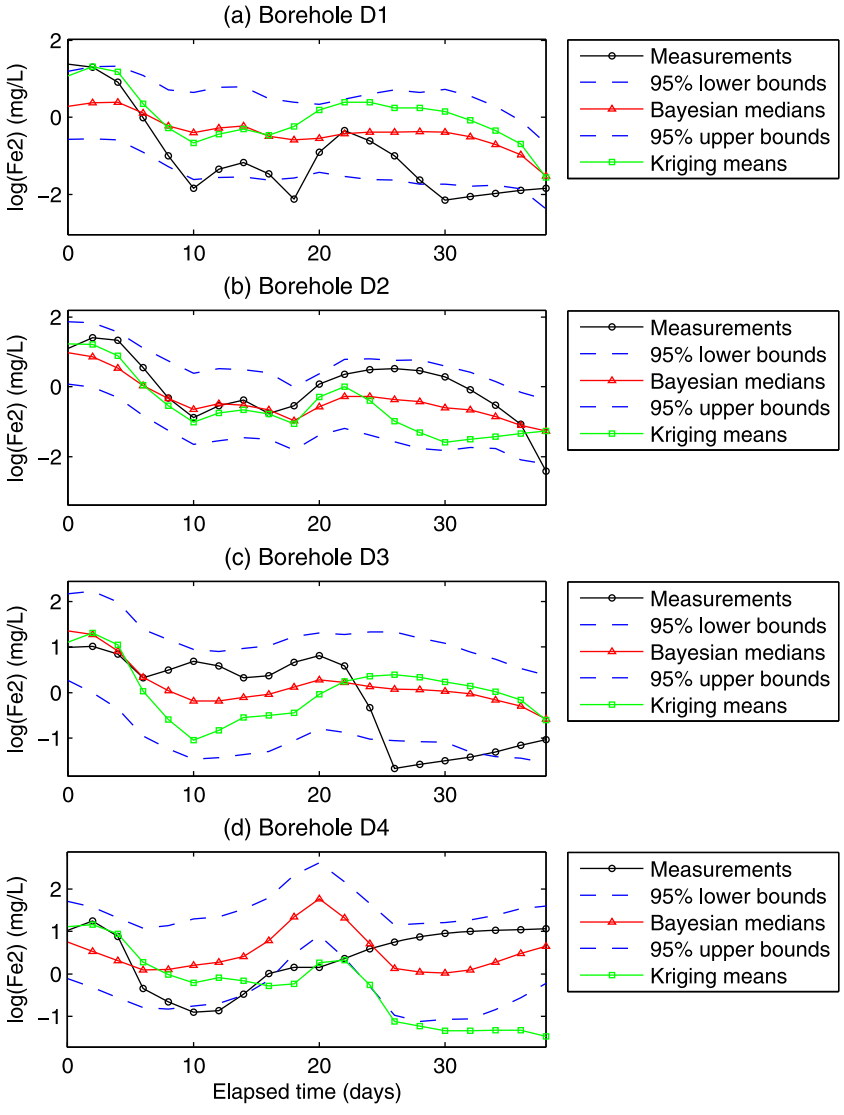

Figure 11. Cross validation: comparison between the estimated and measured log(Fe(II)) (mg L^{-1}) at (a) borehole D1, (b) borehole D2, (c) borehole D3, and (d) borehole D4 over the injection period. The RMS of the differences between the measured and the estimated values for Bayesian and kriging methods are 0.83 and 1.18, respectively.

with increasing number of data sets. In the current study we have to use informative prior bounds for the autoregressive model parameters because they cannot be reliably estimated from the SIP data available at only three time steps. If we have geophysical data at more time steps, we could release the constraints. In addition, the current estimation results depend on the variance of regression analysis, which is derived from data at only four borehole sampling ports and thus subject to large uncertainty. If more borehole measurements are available, improved cross correlation and thus estimation results are expected.

[49] Our study provides a methodology capable of providing estimates of groundwater geochemical evolution and associated parameter uncertainty through jointly honoring sparse yet direct wellbore measurements with indirect yet spatially extensive surface SIP data sets. Implementation of the method at the Rifle IFRC suggests the significant potential of the method for providing information about the spatiotemporal distribution of geochemical parameters, which are valuable for assessing the effectiveness of bioremediation treatment, constraining reactive transport models, and helping to understand the constitutive geochemical reaction networks in the presence of subsurface heterogeneity.

[50] We believe that the experiment successfully demonstrated the ability to remove dissolved uranium from groundwater through the injection of acetate amendments. Long-term monitoring of biostimulation experiments have also demonstrated that favorable conditions for uranium removal are maintained even for periods after cessation of acetate injections. Results in this study also revealed an ever-improving ability to derive quantitative estimates from SIP data sets, which permitted a better understanding of processes in the subsurface associated with changes in groundwater chemistry. The monitoring and understanding of such processes is of critical relevance to better assess fate and transport of contaminants in groundwater, as required

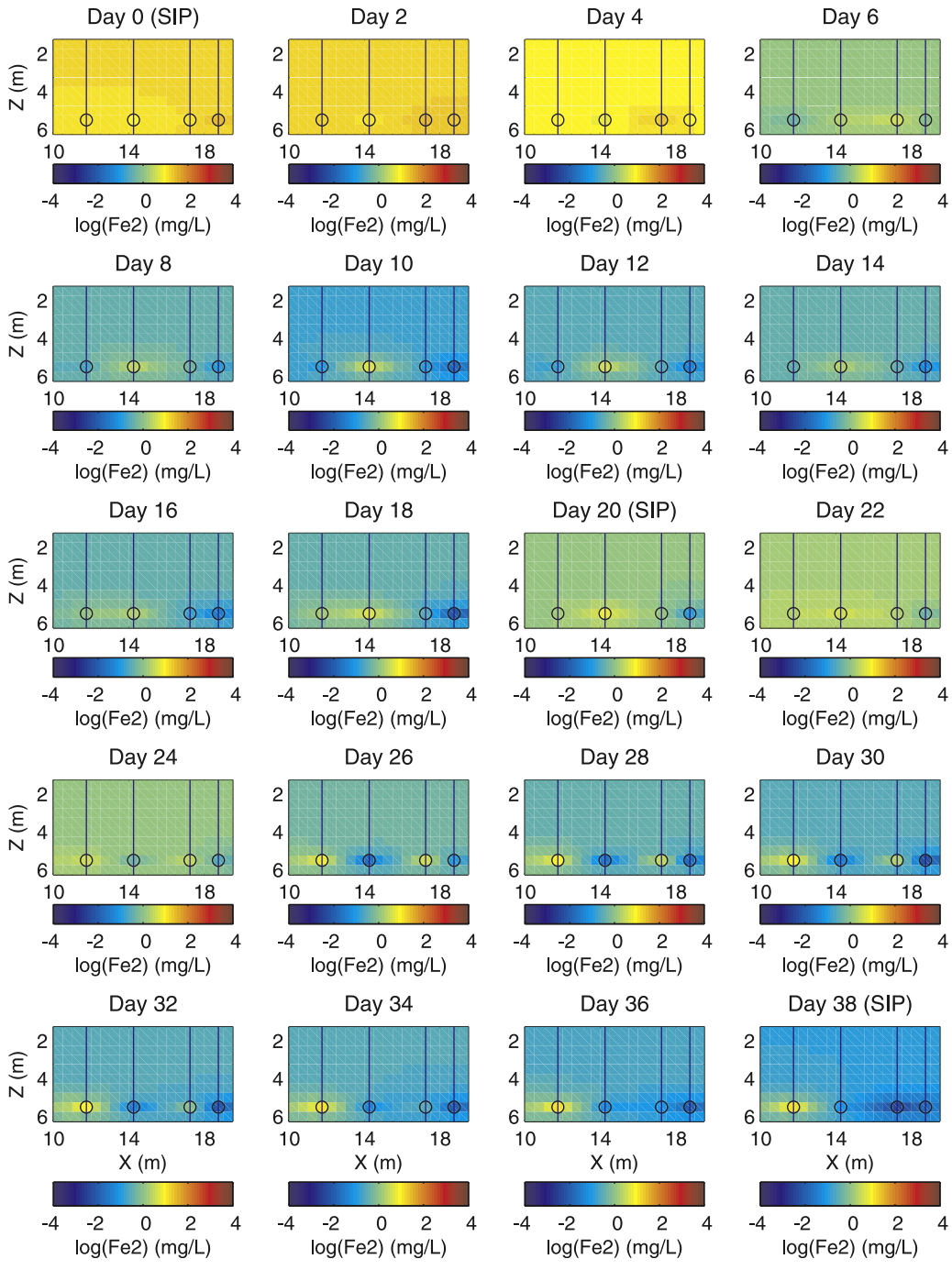

Figure 12. Estimated mean of log(Fe(II)) (mg L^{-1}) along the survey cross section over the injection period using the ordinary kriging. The vertical lines are borehole D1, D2, D3, and D4 from the right to the left, respectively. The circles are the locations where geochemical samples were collected.

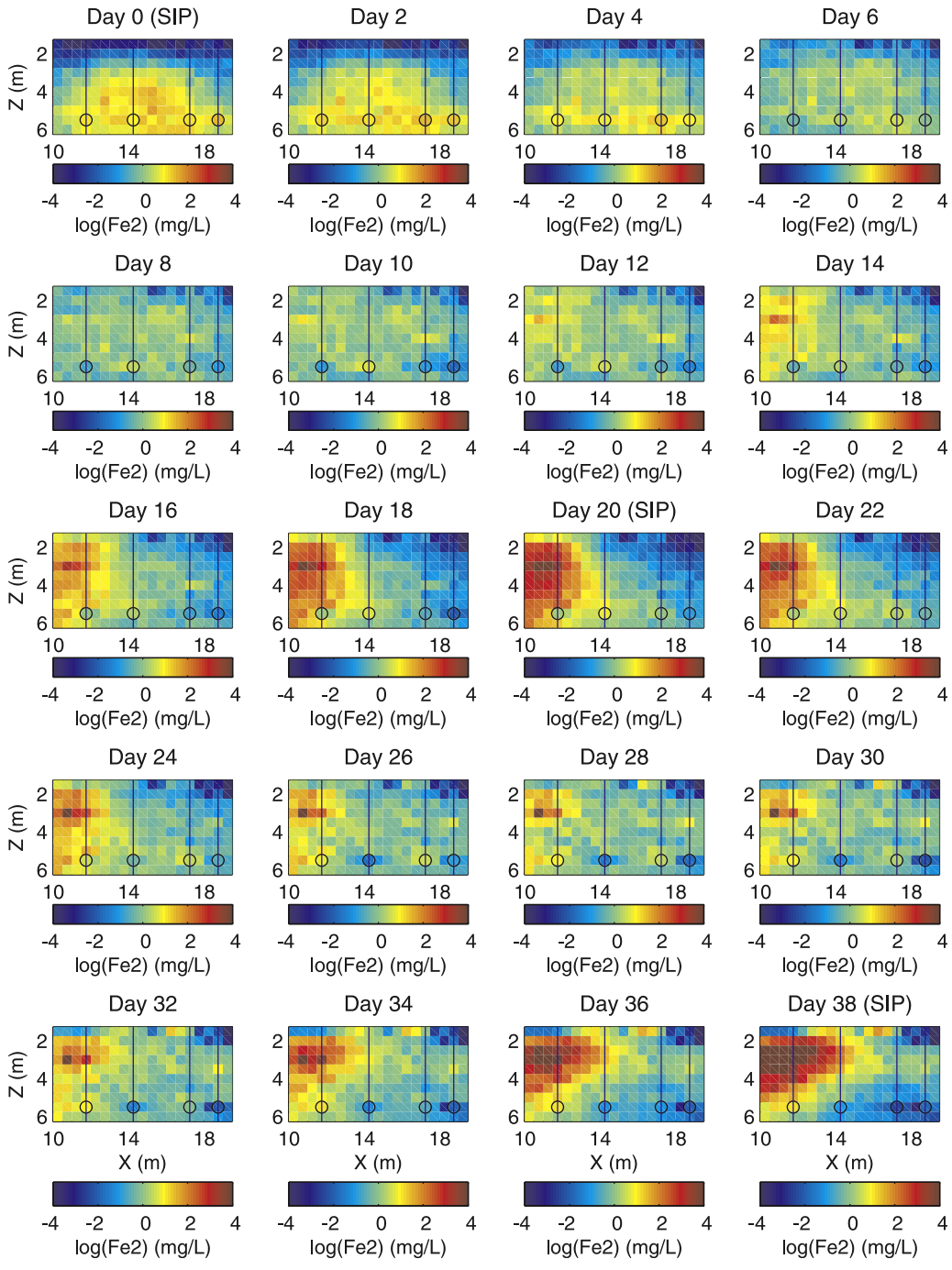

Figure 13. Estimated medians of log(Fe(II)) (mg L^{-1}) along the survey cross section over the injection period using the Bayesian model. The vertical lines are borehole D1, D2, D3, and D4 from the right to the left, respectively. The circles are the locations where geochemical samples were collected.

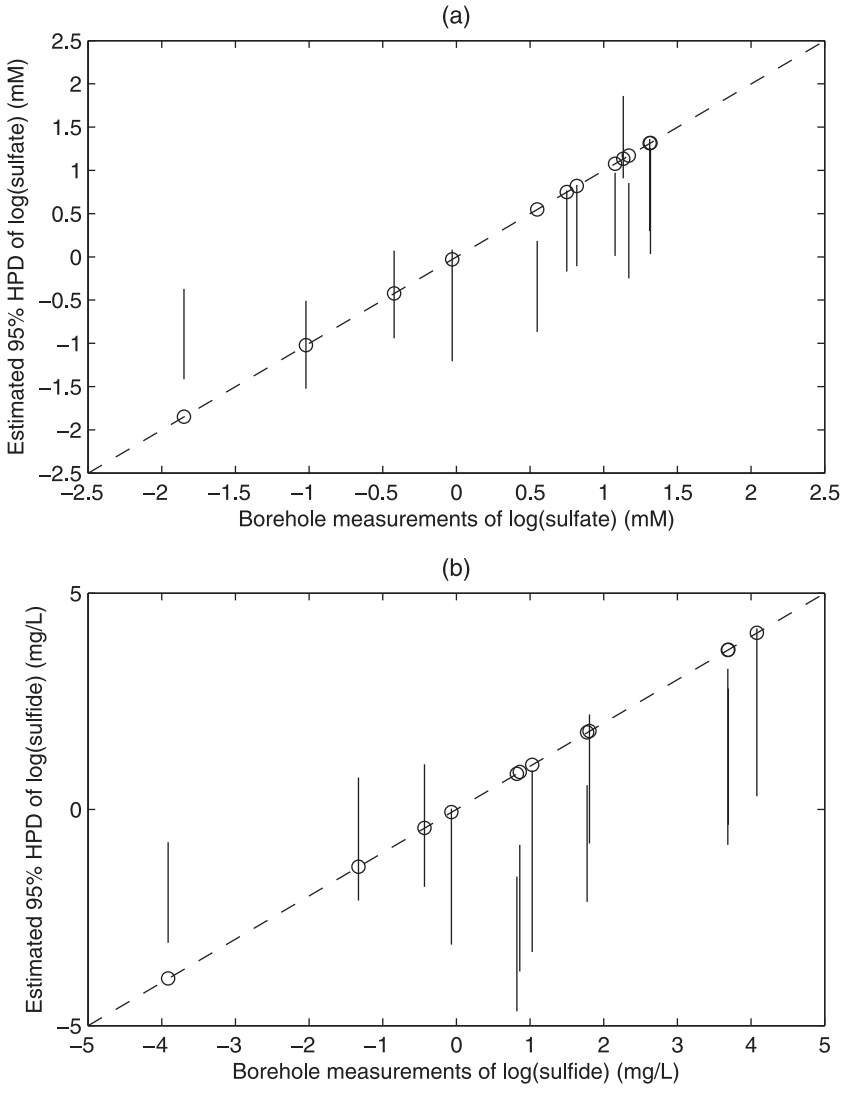

Figure 14. Cross validation: comparison between the estimated and measured (a) log(sulfate) (mM) and (b) log(sulfide) (mg L^{-1}) at the time step that SIP data are available. The vertical segments are the 95% highest probability domains (HPD) and the circles are the borehole measurements.

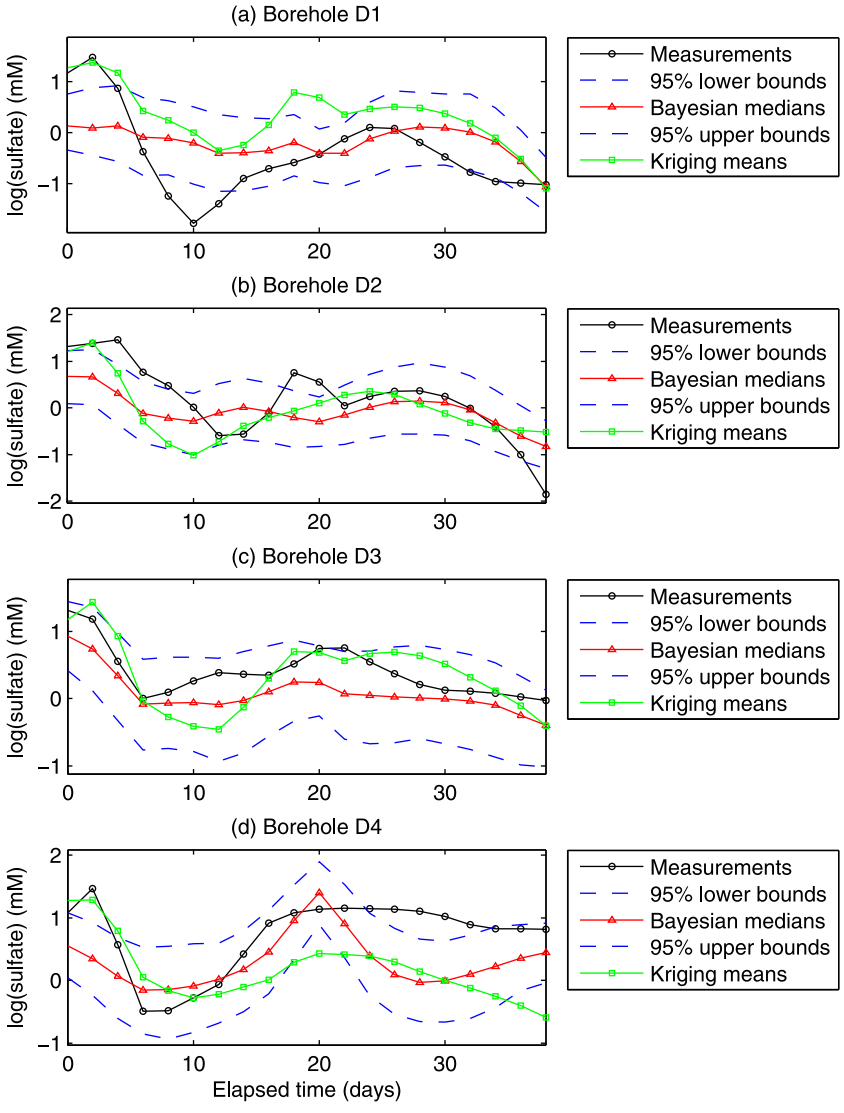

Figure 15. Cross validation: comparison between the estimated and measured log(sulfate) (mM) at (a) borehole D1, (b) borehole D2, (c) borehole D3, and (d) borehole D4 over the injection period. The RMS of the differences between the measured and the estimated values for Bayesian and kriging methods are 0.57 and 0.69, respectively.

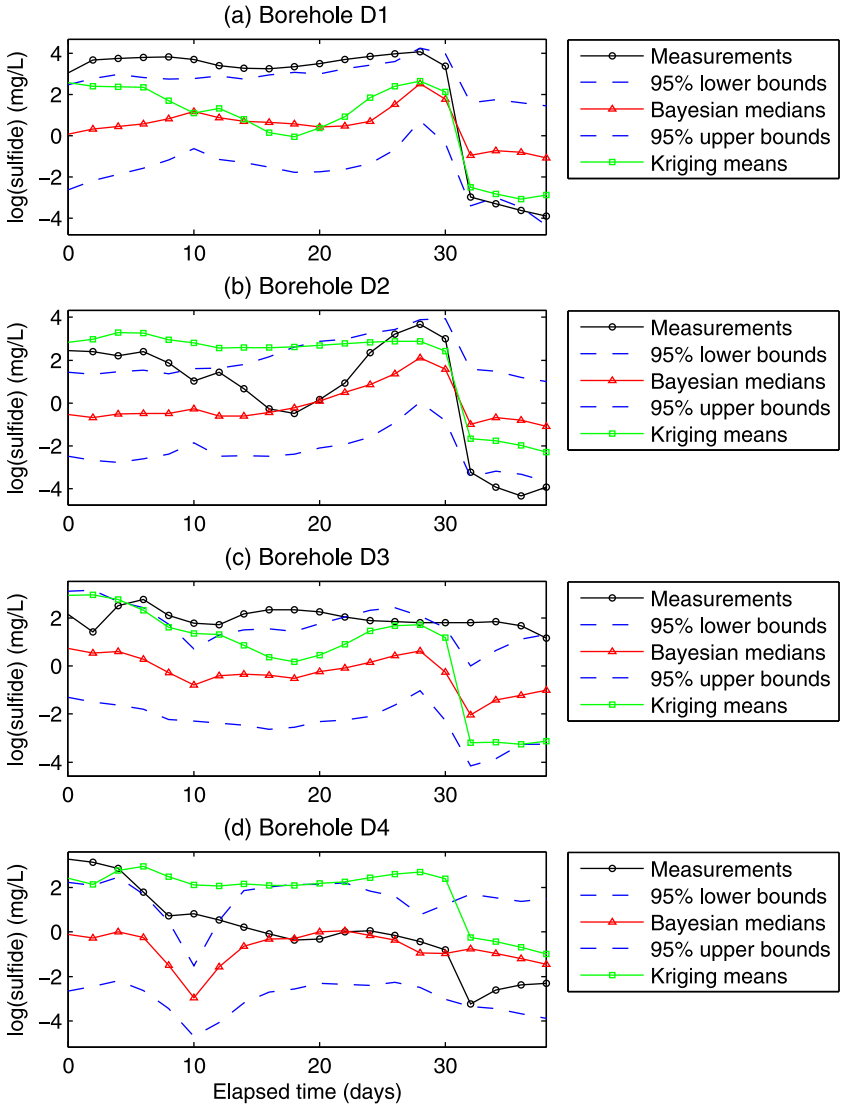

Figure 16. Cross validation: comparison between the estimated and measured log(sulfide) (mg L^{-1}) at borehole D1, (b) borehole D2, (c) borehole D3, and (d) borehole D4 over the injection period. The RMS of the differences between the measured and the estimated values for Bayesian and kriging methods are 1.25 and 1.73, respectively.

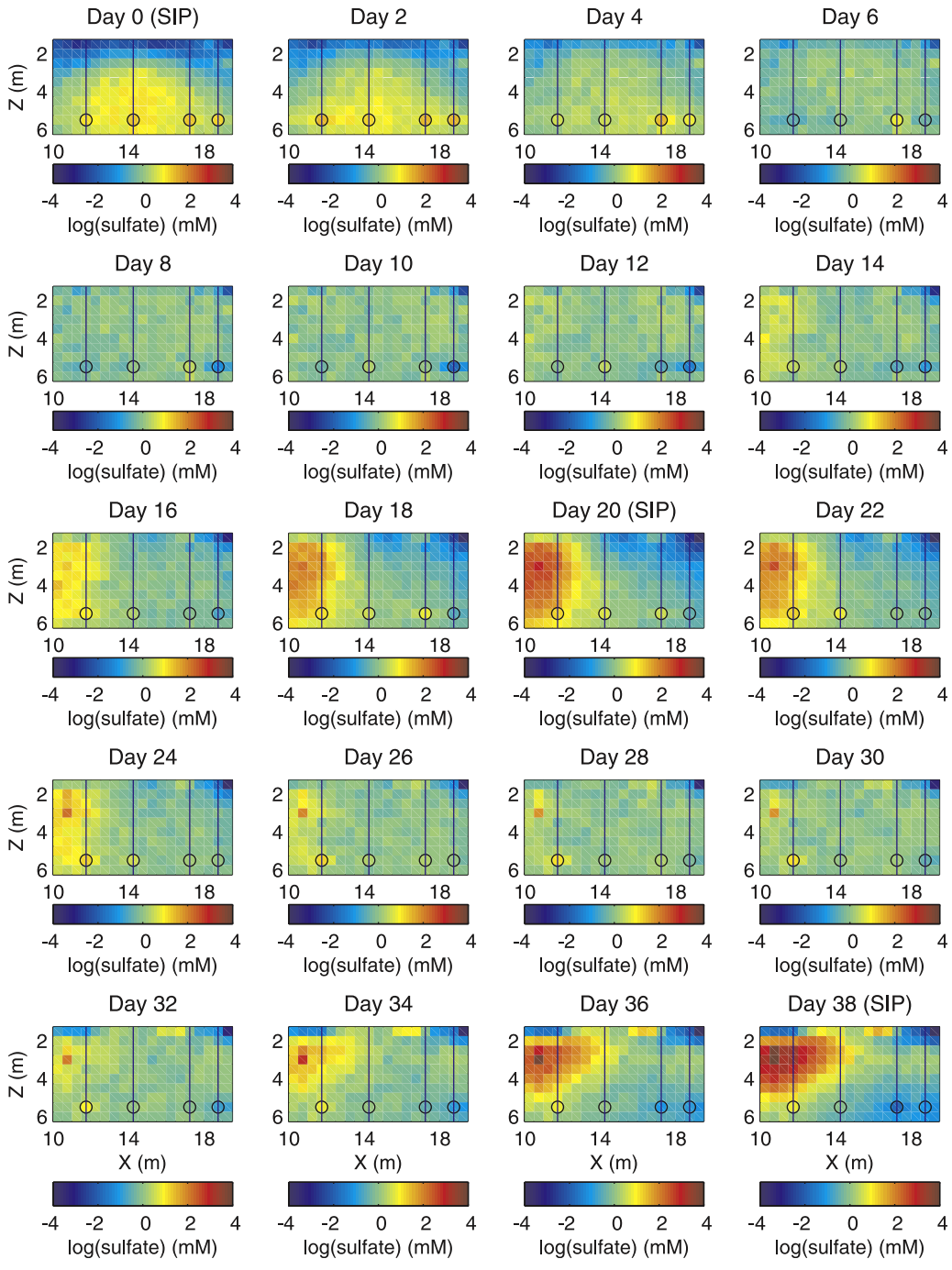

Figure 17. Estimated medians of log(sulfate) (mM) along the survey cross section over the injection period using the Bayesian model. The vertical lines are borehole D1, D2, D3, and D4 from the right to the left, respectively. The circles are the locations where geochemical samples were collected.

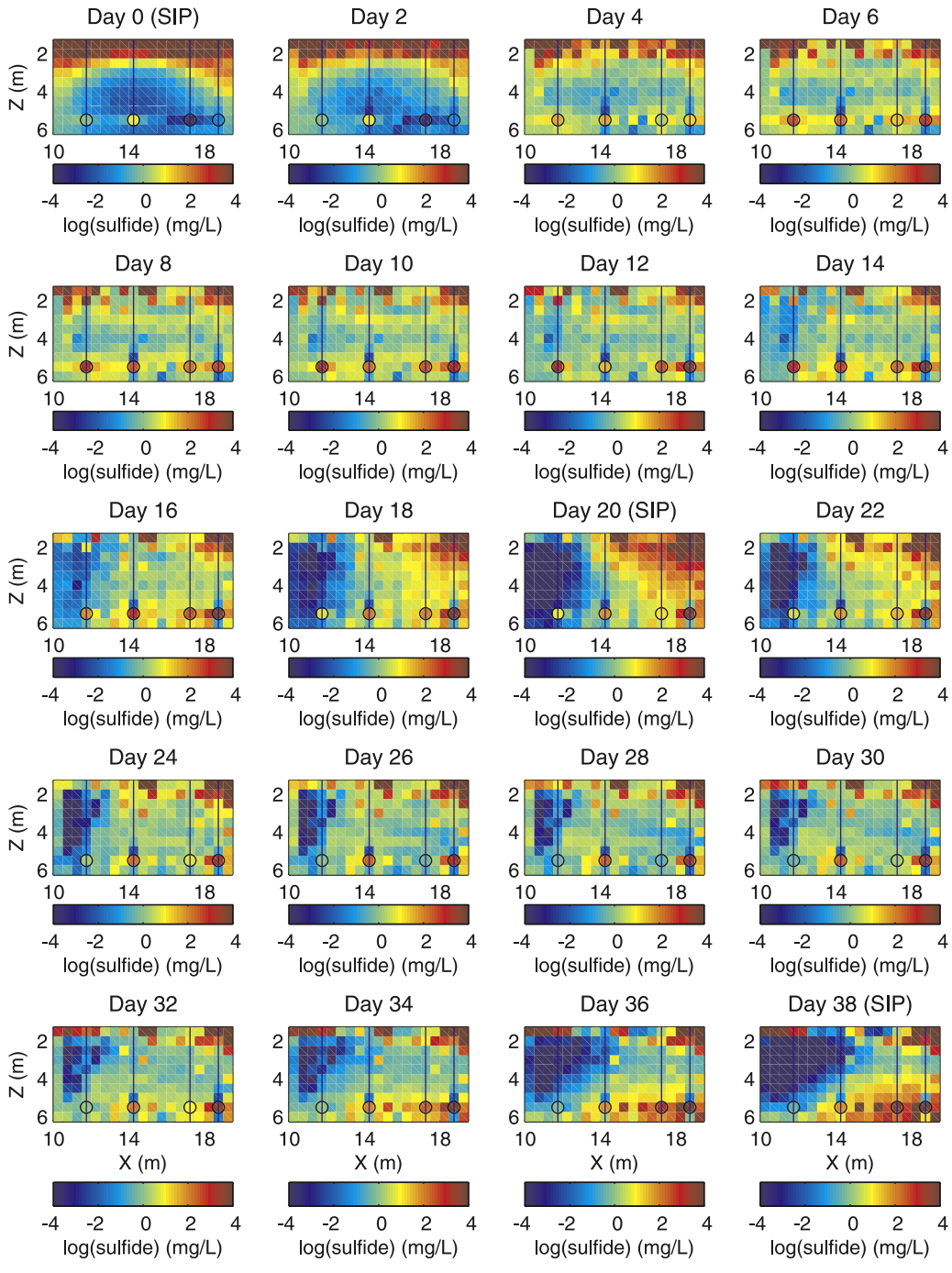

Figure 18. Estimated medians of log(sulfide) (mg L^{-1}) along the survey cross section over the injection period using the Bayesian model. The vertical lines are borehole D1, D2, D3, and D4 from the right to the left, respectively. The circles are the locations where geochemical samples were collected.

in modern hydrogeological studies. In particular, this study shows the possibility to solve for Cole-Cole parameters required for the application of petrophysical models as developed in laboratory studies.

Appendix A: Full Conditional Probability Distributions

[51] We can derive full conditional probability distributions of all other unknowns by following similar procedures for $[\mathbf{u}_t|\cdot]$ and $[\beta_1(s)|\cdot]$. In the following we list those full conditionals:

$$[\alpha_1|\cdot] \propto I(\alpha_1 \in (a_{\alpha_1}, b_{\alpha_1}))[\{\mathbf{m}_t^{\text{obs}}, t \in T_g\}|\{\mathbf{u}_t, t \in T_g\}, \alpha_1, \alpha_2, \tau_m]$$
$$\propto I(\alpha_1 \in (a_{\alpha_1}, b_{\alpha_1}))\exp\left\{-0.5\tau_m \sum_{t\in T_g} (\mathbf{m}_t^{\text{obs}} - \alpha_2\mathbf{u}_t)^T (\mathbf{m}_t^{\text{obs}} - \alpha_2\mathbf{u}_t)\right\} \sim N_T(\mu_{\alpha_1}, \tau_{\alpha_1}, a_{\alpha_1}, b_{\alpha_1}), \tag{A1}$$

where

$$\tau_{\alpha_1} = mn_g\tau_m,$$
$$\mu_{\alpha_1} = \left\{\tau_m \sum_{t\in T_g} (\mathbf{m}_t^{\text{obs}} - \alpha_2\mathbf{u}_t)^T\mathbf{e}\right\}/\tau_{\alpha_1}.$$

$$[\alpha_2|\cdot] \sim N_T(\mu_{\alpha_2}, \tau_{\alpha_2}, a_{\alpha_2}, b_{\alpha_2}), \tag{A2}$$

where

$$\tau_{\alpha_2} = \tau_m \sum_{t\in T_g} \mathbf{u}_t^T\mathbf{u}_t,$$
$$\mu_{\alpha_2} = \left\{\tau_m \sum_{t\in T_g} (\mathbf{m}_t^{\text{obs}} - \alpha_1\mathbf{e})^T\mathbf{u}_t\right\}/\tau_{\alpha_2}.$$

$$[\beta_2(s)|\cdot] \sim N_T(\mu_{\beta_2}(s), \tau_{\beta_2}(s), a_{\beta_2}, b_{\beta_2}), \tag{A3}$$

where

$$\tau_{\beta_2(s)} = \sum_{k=3}^{n} \mathbf{c}_k^T(s)\mathbf{c}_k(s),$$
$$\mu_{\beta_2(s)} = \sum_{k=3}^{n} \mathbf{c}_k^T(s)\mathbf{d}_k(s)/\tau_{\beta_2(s)},$$

and

$$\mathbf{c}_k(s) = u_{k-2}(s)\mathbf{q}_s,$$
$$\mathbf{d}_k(s) = \sum_{i=1}^{m}(u_k(i) - \beta_1(i)u_{k-1}(i))\mathbf{q}_i - \sum_{i\neq s}^{m}\beta_2(i)u_{k-2}(i)\mathbf{q}_i.$$

$$[\mu_{u1}|\cdot] \sim N_T(\mu_{u1}^*, \tau_{u1}^*, a_{u1}, b_{u1}), \tag{A4}$$

where

$$\tau_{u1}^* = \tau_{u1}(\mathbf{e}^T\mathbf{R}^{-1}\mathbf{e}),$$
$$\mu_{u1}^* = \{\tau_{u1}(\mathbf{u}_1^T\mathbf{R}^{-1}\mathbf{e})\}/\tau_{u1}^*.$$

$$[\mu_{u2}|\cdot] \sim N_T(\mu_{u2}^*, \tau_{u2}^*, a_{u2}, b_{u2}), \tag{A5}$$

where

$$\tau_{u2}^* = \tau_{u2}(\mathbf{e}^T\mathbf{R}^{-1}\mathbf{e}),$$
$$\mu_{u2}^* = \{\tau_{u2}(\mathbf{u}_2^T\mathbf{R}^{-1}\mathbf{e})\}/\tau_{u2}^*.$$

Similarly, we can obtain conditional probability distributions of all the inverse variances

$$[\tau_m|\cdot] \sim G_T(0.5mn_g + 1, 0.5S_m, a_{\tau_m}, b_{\tau_m}), \tag{A6}$$

where G_T represents the truncated Gamma distribution within given bounds and

$$S_m = \sum_{t\in T_g} (\mathbf{m}_t^{\text{obs}} - \alpha_1\mathbf{e} - \alpha_2\mathbf{u}_t)^T(\mathbf{m}_t^{\text{obs}} - \alpha_1\mathbf{e} - \alpha_2\mathbf{u}_t).$$

$$[\tau_{u1}|\cdot] \sim G_T(0.5m + 1, 0.5S_{u1}, a_{\tau_{u1}}, b_{\tau_{u1}}), \tag{A7}$$

where

$$S_{u1} = (\mathbf{u}_1 - \mu_{u1}\mathbf{e})^T\mathbf{R}^{-1}(\mathbf{u}_1 - \mu_{u1}\mathbf{e}).$$

$$[\tau_{u2}|\cdot] \sim G_T(0.5m + 1, 0.5S_{u2}, a_{\tau_{u2}}, b_{\tau_{u2}}), \tag{A8}$$

where

$$S_{u2} = (\mathbf{u}_2 - \mu_{u2}\mathbf{e})^T\mathbf{R}^{-1}(\mathbf{u}_2 - \mu_{u2}\mathbf{e}).$$

$$[\tau_{pu}|\cdot] \sim G_T(0.5m(n-2) + 1, 0.5S_{pu}, a_{\tau_u}, b_{\tau_u}), \tag{A9}$$

where

$$S_{pu} = \sum_{k=3}^{n} (\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2})^T\mathbf{R}^{-1}(\mathbf{u}_k - \boldsymbol{\beta}_1\mathbf{u}_{k-1} - \boldsymbol{\beta}_2\mathbf{u}_{k-2}).$$

[52] **Acknowledgments.** Funding for this study was provided by the U.S. Department of Energy, Biological and Environmental Research Program under Award Number DE-AC02-05CH11231 to the LBNL Sustainable Systems Subsurface Science Focus Area (SFA). We thank the associate editor Fred Day-Lewis, Andre Revil, Dimitris Ntarlagiannis, and one anonymous reviewer for their constructive comments.

References

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–281, Akademiai Kiado, Budapest.

Anderson, R. T., et al. (2003), Stimulating the in-situ activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer, *Appl. Environ. Microbiol.*, *69*(10), 5884–5891.

Carlin, B. P., N. G. Polson, and D. S. Stoffer (1992), A Monte Carlo approach to non-normal and nonlinear state-space modeling, *J. Am. Stat. Assoc.*, *87*(418), 493–500.

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.

Commer, M., G. A. Newman, K. H. Williams, and S. S. Hubbard (2011), 3D induced-polarization data inversion for complex resistivity, *Geophysics*, *76*(3), F157–F171.

Cosenza, P., A. Ghorbani, N. Florsch, and A. Revil (2007), Effects of drying on the low-frequency electrical properties of Tournemire argillites, *Pure Appl. Geophys.*, *164*(10), 2043–2066.

Englert, A., S. S. Hubbard, K. H. Williams, L. Li, and C. I. Steefel (2009), Feedbacks between hydrological heterogeneity and bioremediation induced biogeochemical transformations, *Environ. Sci. Technol.*, *43*, 5197–5204.

Flores Orozco, A., K. H. Williams, P. E. Long, S. S. Hubbard, and A. Kemna (2011), Using complex resistivity imaging to infer biogeochemical processes associated with bioremediation of an uranium contaminated aquifer, *J. Geophys. Res.*, *116*, G03001, doi:10.1029/2010JG001591.

Gelfand. A. E., and F. M. Smith (1990), Sampling-based approaches to calculating marginal densities, *J. Am. Stat. Assoc.*, *85*(410), 398–409.

Gelman, S., and D. Gelman (1984), Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intel.*, *6*, 721–741.

Johnson, T. C., R. J. Versteeg, A. Ward, F. D. Day-Lewis, and A. Revil (2010), Improved hydrogeophysical characterization and monitoring through high performance electrical geophysical modeling and inversion, *Geophysics*, *75*(4), WA27–WA41.

Karaoulis, M., A. Revil, D. D. Werkema, B. J. Minsley, W. F. Woodruff, and A. Kemna (2011), Time-lapse three-dimensional inversion of complex conductivity data using an active time constrained (ATC) approach, *Geophys. J. Int.*, *187*, 237–251.

Kemna, A. (2000), Tomographic inversion of complex resistivity: Theory and application, Ph.D. thesis, Ruhr-University, Bochum.

Kitanidis, P. K. (1997), *Introduction to Geostatistics: Applications in Hydrogeology*, Cambridge University Press, New York.

Kohavi, R. (1995), A study of cross-validation and bootstrap for accuracy estimation and model selection, *International Joint Conference on Artificial Intelligence*, Morgan Kaufmann Publ., San Francisco, Calif.

Li, L., C. I. Steefel, M. B. Kowalsky, A. Englert, and S. Hubbard (2010), Effects of physical and geochemical heterogeneities on mineral transformation and biomass accumulation during biostimulation experiments at Rifle, Colorado, *J. Contam. Hydrol.*, *112*, 45–63.

Lorenz, J. C. (1982), Sedimentology of the Mesaverde formation at Rifle gap, Colorado and implications for gas-bearing intervals in the subsurface, *Report SAND82-0604*, Sandia National Laboratory, Albuquerque, N.M., 43 pp.

Lovley, D. R., F. H. Chapelle, and J. C. Woodward (1994), Use of dissolved H2 concentrations to determine distribution of microbially catalyzed redox reactions in anoxic groundwater, *Environ. Sci. Technol.*, *28*, 1205–1210.

Ntarlagiannis, D., K. H. Williams, L. Slater, and S. S. Hubbard (2005), Low-frequency electrical response to microbial induced sulfide precipitation, *J. Geophys. Res.*, *110*, G02009, doi:10.1029/2005JG000024.

Olhoeft, G. (1992), *Geophysical detection of hydrocarbon and organic chemical contaminants*, 587–594, SAGEEP, proceedings, Oakbrook, IL.

Personna, Y. R., D. Ntarlagiannis, L. Slater, N. Yee, M. O'Brien, and S. S. Hubbard (2008), Spectral induced polarization and electrodic potential monitoring of microbially mediated iron sulfide transformations, *J. Geophys. Res.*, *113*, G02020, doi:10.1029/2007JG000614.

Revil, A., and M. Skold (2011), Salinity dependence of spectral induced polarization in sands and sandstones, *Geophys. J. Int.*, *187*, 813–824.

Slater, L., D. Ntarlagiannis, and D. Wishart (2006), On the relationship between induced polarization and surface area in metal-sand and clay-sand mixtures, *Geophysics*, *71*, A1–A5.

Slater, L., D. Ntarlagiannis, Y. R. Personna, and S. Hubbard (2007), Pore-scale spectral induced polarization signatures associated with FeS biomineral transformations, *Geophys. Res. Lett.*, *34*, L21404, doi:10.1029/2007GL031840.

Sogade, J. A., F. Scri-Scappuzzo, Y. Vichabian, W. Shi, W. Rodi, D. P. Lesmes, and F. D. Morgan (2006), Induced-polarization detection and mapping of contaminant plumes, *Geophysics*, *71*(3), B75.

Vanhala, H. (1997), Mapping oil-contaminated sand and till with the spectral induced polarization (SIP) method, *Geophys. Prospecting*, *45*, P303–P326.

Vrionis, H. A., R. T. Anderson, I. Ortiz-Bernad, K. R. O'Neill, C. T. Resch, A. D. Peacock, R. Dayvault, D. C. White, P. E. Long, and D. R. Lovley (2005), Microbiological and geochemical heterogeneity in an in situ uranium bioremediation field site, *Appl. Environ. Microbiol.*, *71*(10), 6308–6318.

Wikle, C. K. (2003), Hierarchical Bayesian models for predicting the spread of ecological processes, *Ecology*, *84*, 1384–1394.

Williams, K. H., D. Ntarlagiannis, L. D. Slater, A. Dohnalkova, S. S. Hubbard, and J. F. Banfield (2005), Geophysical imaging of stimulated microbial biomineralization, *Eviron. Sci. Technol.*, *39*, 7592–7600.

Williams, K. H., A. Kemna, M. Wilkins, J. Druhan, E. Arntzen, L. N′Guessan, P. E. Long, S. S. Hubbard, and J. F. Banfield (2009), Geophysical monitoring of coupled microbial and geochemical processes during stimulated subsurface bioremediation, *Environ. Sci. Technol.*, *43*(17), 6717–6723, doi:10.1021/es900855j.

Williams, K. H., et al. (2011), Acetate availability and its influence on sustainable bioremediation of uranium-contaminated groundwater. *Geomicrobiol. J.*, *28*, 519–539. doi:10.1080/01490451.2010.520074.

Wu, Y., J. Ajo-Franklin, N. Spycher, S. Hubbard, G. Zhang, K. H. Williams, J. Taylor, Y. Fujita, and R. Smith (2011), Geophysical monitoring and reactive transport modeling of ureolytically driven calcium carbonate precipitation, *Geochem. Trans.*, *12*(7), 1–20.

Yabusaki, S. B., et al. (2007), Uranium removal from groundwater via in-situ bioremediation: Field-scale modeling of transport and biological processes, *J. Contam. Hydrol.*, *93*, 216–235.