GEOPHYSICS, VOL. 77, NO. 4 (JULY-AUGUST 2012); P. E265–E279, 16 FIGS., 2 TABLES.
10.1190/GEO2011-0430.1

Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site

Jinsong Chen[1], G. Michael Hoversten[2], Kerry Key[3], Gregg Nordquist[4], and William Cumming[5]

ABSTRACT

We developed a Bayesian model to invert magnetotelluric (MT) data using a 2D sharp boundary parameterization. We divided the 2D cross section into layers and considered the locations of interfaces and resistivity of the regions formed by the interfaces as random variables. We assumed that those variables are independent in the vertical direction and dependent along the lateral direction, whose spatial dependence is described by either pairwise difference or multivariate Gaussian priors. We used a parallel, adaptive finite-element algorithm to rapidly forward simulate frequency-domain MT responses of the 2D resistivity structure and used Markov chain Monte Carlo methods to draw many samples from the joint posterior probability distribution. We applied the Bayesian model to a synthetic case that mimics a geothermal exploration scenario. Our results demonstrated that the developed method is effective in estimating the resistivity and depths to interfaces and in quantifying uncertainty on the estimates. We also applied the developed method to the field MT data collected from the Darajat geothermal site in Indonesia. We compared our inversion results with those obtained from a deterministic inversion of 3D MT data; they are consistent even if the two inversion methods are very different and the amount of information used for inversion is different.

INTRODUCTION

Current approaches for geophysical inverse problems are mainly gradient-based deterministic methods, such as Gauss-Newton methods (Pratt et al., 1998), conjugate gradient methods (Newman and Alumbaugh, 2000), and steepest decent techniques (Roy, 2002). Those methods have been successfully used to solve a wide range of complex inverse problems with tens of millions of unknowns. However, the solutions obtained using the methods often depend on the choice of initial values and thus are local rather than global. Additionally, the deterministic methods provide very limited uncertainty information on the estimated parameters.

Stochastic inversion methods have been recognized recently as a powerful approach for solving geophysical inverse problems; many successful applications can be found in the literature (e.g., Mosegaard and Tarantola, 1995; Bosch, 1999, 2004; Maliverno, 2002; Eidsvik et al., 2004; Tarantola, 2005). Stochastic methods have several advantages over deterministic inversion methods. As demonstrated by Chen et al. (2008), stochastic inversion methods can provide extensive information about unknown parameters; inversion results are almost independent of initial values and therefore global and robust.

Because stochastic methods often need to run forward models tens of thousands of times, current applications are limited to problems where fast forward models are available, such as inversion of 1D seismic data (Gunning and Glinsky, 2004) and joint inversion of 1D seismic amplitude versus angles and controlled-sources electromagnetic (CSEM) data (Chen et al., 2007; Chen and Hoversten, 2012). Stochastic methods have also been applied to inversion of parameters on 2D domains when forward calculations are fast or can be approximated using simplified models. For example, Bodin and Sambridge (2009) develop a Bayesian model for seismic tomography inversion by using the fast-marching method (Sethian and Popovici, 1999) to calculate traveltimes and raypaths in a laterally heterogeneous medium. Irving and Singha (2010) develop a stochastic method to invert tracer test and electrical resistivity data for hydraulic conductivity tomography by using simplified models.

Manuscript received by the Editor 25 October 2011; revised manuscript received 7 March 2012; published online 26 June 2012.
[1]Lawrence Berkeley National Laboratory, Berkeley, California, USA. E-mail: jchen@lbl.gov.
[2]Chevron Energy Technology Company, San Ramon, California, USA. E-mail: hovg@chevron.com.
[3]Scripps Institution of Oceanography, San Diego, California, USA. E-mail: kkey@ucsd.edu.
[4]Chevron Geothermal and Power Operations, Bellaire, Texas, USA. E-mail: gnordquist@chevron.com.
[5]Cumming Geosciences, Santa Rosa, California, USA. E-mail: wcumming@wcumming.com.

Stochastic inversion of magnetotelluric (MT) and CSEM data on 2D or 3D media remains very challenging because of the lack of fast forward models. However, with the rapid growth in computing power, especially parallel computing techniques and recent development of forward simulation methods, many fast forward algorithms are becoming available. For example, Key and Ovall (2011) develop a fast, parallel algorithm that computes 2D MT and CSEM responses for complex models with realistic data parameters in a few to tens of seconds when it is run on large clusters (>50 nodes). These new capabilities make it feasible to implement 2D stochastic inversion for MT and CSEM data sets.

Several deterministic methods were developed to invert 2D MT data for mapping large scale structures for the purpose of petroleum and geothermal exploration in the past decades. For instance, de Groot-Hedlin and Constable (1990) develop an Occam's inversion method to generate smooth 2D models using a grid-base parameterization; Smith and Booker (1991) develop a rapid inversion method by approximation. To better estimate large structures from 2D MT data, Smith et al. (1999) and de Groot-Hedlin and Constable (2004) develop algorithms based on sharp boundary parameterizations. By carefully choosing the starting models, they could obtain inversion results that are supported by other sources of information. However, they find that the inversion results are subject to a large degree of uncertainty and the optimal solutions depend on the choice of the initial models.

We explore the use of stochastic inversion approaches for inverting 2D MT data in this study by using a newly developed parallel, goal-oriented, adaptive finite-element algorithm (Key and Ovall, 2011) to forward simulate MT responses. We start from a sharp boundary parameterization because it requires significantly fewer unknowns compared with grid-based parameterizations, although it is versatile enough to solve a wide range of problems. We develop a Bayesian model to estimate the locations of sharp boundaries and resistivity of the regions formed by the boundaries. The spatial dependence of unknown parameters is described by the pairwise difference or multivariate Gaussian priors, depending on prior information available. We solve the inverse problem using Markov chain Monte Carlo (MCMC) sampling methods (Gilks et al., 1996). We first use a synthetic study to test the developed Bayesian model and then apply the model to the field MT data sets collected from the Darajet geothermal site in Indonesia.

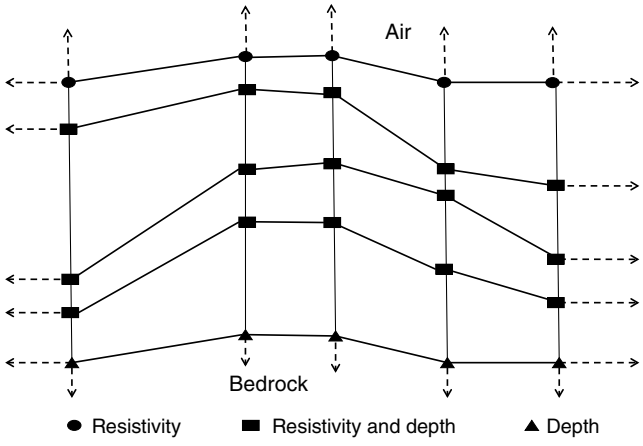

Figure 1. A schematic diagram of the sharp boundary parameterization. The dashed-lines with arrows point to the domain bounds far away from the focus area. The solid ovals, triangles, and rectangles are the nodes linked to unknown resistivity, depths, and both resistivity and depths, respectively.

METHODOLOGY

Sharp boundary parameterization and Bayesian model

We develop a Bayesian model using the sharp boundary method as used by Smith et al. (1999) and Auken and Christiansen (2004). The term "sharp boundary" is different from the one used with the L1 norm inversion. As pointed out by Auken and Christiansen (2004), although the L1 inversion method tends to give a more blocky appearance of model sections (Loke et al., 2001), the layer boundaries are still smeared. Figure 1 shows a schematic diagram of the parameterization, where the dashed arrows point to the bounds that are several hundreds of kilometers away from the focus area and the solid lines represent sharp boundaries (or interfaces). We assume that the maximum domain size and the ground surface locations are predetermined. In the figure, the solid ovals, rectangles, and triangles represent inner estimation nodes; we estimate resistivity and/or depths associated with each of those inner nodes, depending on the interface. For a given inverse problem, such a parameterization often has a minimal number of unknown parameters.

We obtain resistivity images from specifications of the 2D resistivity structure by linearly interpolating resistivity values at the inner nodes. We divide each trapezoid that is formed by the four adjacent nodes (see Figure 1) into several subblocks, typically three to six, depending on the sharpness of the lateral resistivity transition. We assume that resistivity in each subblock is constant and it is calculated by laterally interpolating the resistivity values at the two upper nodes of the trapezoid.

Let m be the total number of sharp boundaries, including the ground surface. Although the total number of inner estimation nodes can vary from one interface to another, we assume that it is constant for ease of description and let n be the number. We estimate logarithmic resistivity r_{ki} at each inner node starting from the ground surface, where $k = 1, 2, \cdots, m-1$, and $i = 1, 2, \cdots, n$. We also estimate logarithmic resistivity in the bedrock and denote it as r_0. To simplify description, we let vector $\mathbf{r}_k = (r_{k1}, r_{k2}, \cdots, r_{kn})^T$ represent the unknown resistivity on interface k, where T represents the transpose of the vector. We estimate logarithmic depth d_{ki} at each inner node from interface 2 to interface m, where $k = 2, 3, \cdots, m$, and $i = 1, 2, \cdots, n$. Similarly, we let vector $\mathbf{d}_k = (d_{k1}, d_{k2}, \cdots, d_{kn})^T$ represent the unknown depths on interface k.

The MT data used for our inversion are the apparent resistivity and phases of complex impedance Z_{xy} and Z_{yx} defined by Smith and Booker (1991), with x aligned with the strike of the structure, y perpendicular to the strike, and z positive downward. Impedance Z_{xy} is referred to as the MT data in the transverse electric (TE) mode in which the electric field is parallel to the strike of the structure, and impedance Z_{yx} is referred to as the MT data in the transverse magnetic (TM) mode in which the electric current flow perpendicular to the strike (Smith and Booker, 1991). Numerical 3D forward modeling shows that TM-mode MT data are more robust than the corresponding TE-mode data to non-2D effects, such as finite strike and static shifts (Wannamaker et al., 1984; Hoversten et al., 1998; Newsman et al., 2008). For ease of description, we introduce

four index numbers to show data types, among which 1 and 2 are referred to as the TE-mode apparent resistivity and phases, and 3 and 4 are referred to as the TM-mode apparent resistivity and phases. Let z_{stl}^{obs} be the lth type of MT data for the sth frequency at the tth site, where $s = 1, 2, \cdots, n_f$, $t = 1, 2, \cdots, n_s$, and $l = 1, 2, 3, 4$. Symbols n_f and n_s represent the total number of frequencies and sites, respectively.

Let vector $\mathbf{z}$ represent all the MT data arranged by frequencies, sites, and data types. Let vector $\mathbf{d} = (\mathbf{d}_2^T, \mathbf{d}_3^T, \cdots, \mathbf{d}_m^T)^T$ represent all the unknown depths and vector $\mathbf{r} = (r_0, \mathbf{r}_1^T, \mathbf{r}_2^T, \cdots, \mathbf{r}_{m-1}^T)^T$ represent all the unknown resistivity. We thus have the following Bayesian model

$$f(\mathbf{r}, \mathbf{d}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{r}, \mathbf{d}) f(\mathbf{r}) f(\mathbf{d}). \quad (1)$$

Equation 1 defines a joint posterior probability distribution function of all unknown parameters, which is known up to a normalizing constant. The first term on the right side of the equation is referred to as the likelihood function of MT data, which is the link between the unknown parameters and the MT data. The second and third terms are the prior probability distributions of resistivity and depths, respectively.

Numerical forward model and likelihood function

We use a parallel, adaptive finite-element algorithm developed by Key and Ovall (2011) to compute the frequency-domain MT responses of a given 2D electrical resistivity structure. The discretization of a model domain using an unstructured triangular element grid readily accommodates complex structures. Specifically, the finite-element grid is generated by finding a constrained conforming Delaunay triangulation for a given input polygonal boundary structure (Shewchuk, 1996); therefore, this is ideally suited for the sharp boundary parameterization. An automated, adaptive grid refinement algorithm is implemented in which the finite-element solution is computed iteratively on successively refined grids to ensure an accurate and efficient solution for a given model structure. In the present context, this is useful over traditional, static grid methods because the refinement ensures the solution accuracy is maintained as the model parameters (i.e., resistivity and depths to interfaces) change during the inversion sampling process.

The choice of elements for refinement is accomplished by using an a posteriori error estimator based on a recently developed dual-weighted residual operator (Key and Ovall, 2011). The dual-weighting method essentially applies a receiver-based sensitivity function to the error estimates, thereby allowing for a more efficient grid refinement than possible with simple global error estimators. The residual operator at the core of the error estimation is computed using a hierarchical basis, which has been found to be more robust for heterogeneous models than the gradient-based error estimator considered in earlier works by Key and Weiss (2006) and Li and Key (2007).

Furthermore, the new algorithm has been parallelized so that it can compute accurate, adaptively refined solutions quickly. For the MT computations, the algorithm runs in parallel in two ways. First, the adaptive refinement computations are performed independently in parallel for each frequency. This allows for a rapid solution and also ensures that each frequency has a unique adaptively refined mesh that is accurate for that frequency. The second aspect of the parallelization involves grouping the MT receivers into small subsets so that the adaptive refinement computations are run in parallel for each subset. This allows the adaptive refinement to converge more rapidly because each task only needs to find an adaptive mesh that is accurate for its subset of receivers. For small clusters, it can compute MT responses for models of the size described here in

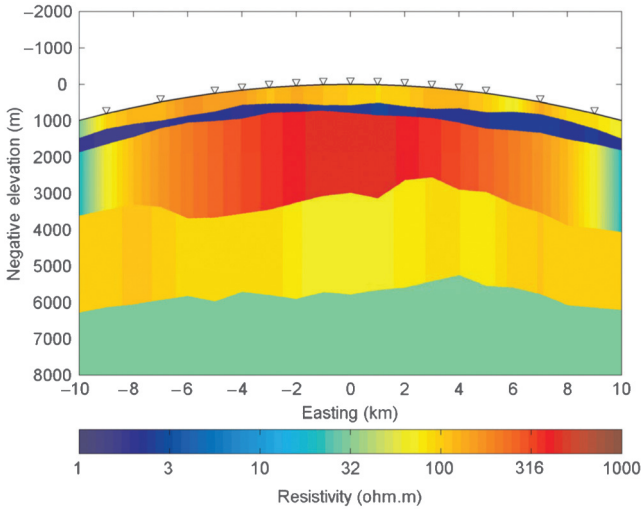

Figure 2. Synthetic 2D resistivity structure. The small triangles represent MT survey stations and the solid curves represent the boundaries with sharp contrasts in resistivity.

Table 1. Geostatistical parameters for generating the true resistivity cross section using the exponential variogram. The resistivity values at the reservoir layer are specified and the bedrock resistivity is assumed to be constant.

	Depth (m)			Resistivity		
Layers	Mean (m)	Standard deviation (m)	Correlation length (m)	Mean (ohm-m)	Standard deviation (ohm-m)	Correlation length (m)
Layer-1 (Overburden)				100.0	40.0	2000.0
Layer-2 (Clay Cap)	500.0	50.0	2000.0	2.0	0.4	4000.0
Layer-3 (Reservoir)	800.0	80.0	4000.0			
Layer-4 (Transition)	3000.0	300.0	6000.0	80.0	16.0	8000.0
Layer-5 (Bedrock)	5000.0	500.0	8000.0	30.0		

only a few tens of seconds, while on larger clusters (>50 nodes), it can compute responses in a few seconds. This makes 2D stochastic inversion of MT data feasible.

The likelihood function of MT data is determined by characteristics of the differences or total errors between the MT data and their corresponding simulation results. The determination of the likelihood function plays an important role in the parameter estimation because it is the only linkage between the unknown parameters and the MT data. The total errors are closely related to the errors in measurements and in forward modeling; correctly characterizing them is very difficult in practice because measurement errors are often affected by many factors, such as acquisition geometry and postprocessing, while errors in forward modeling are often affected by numerical methods and sometimes mathematical models. Because investigating the effects of the total errors on inversion results is not our focus in the current study, we replace the total errors with the measurement errors. Such simplification is valid as long as the errors in the numerical forward modeling are significantly smaller than the measurement errors, which is the case of the current study. Let ε_{stl} be the measurement error for MT data $z_{\text{stl}}^{\text{obs}}$ and let $\text{var}(\varepsilon_{\text{stl}})$ be the variance. By assuming that the errors are independent for different frequencies, sites, and data types, we have the following likelihood function

$$f(\mathbf{z}|\mathbf{r},\mathbf{d}) = \prod_{s=1}^{n_f}\prod_{t=1}^{n_s}\prod_{l=1}^{4}\frac{1}{\sqrt{2\pi\text{var}(\varepsilon_{\text{stl}})}}\exp\left\{-0.5\frac{\varepsilon_{\text{stl}}^2}{\text{var}(\varepsilon_{\text{stl}})}\right\}. \tag{2}$$

The independency and normality assumptions are not critical for the Bayesian model and they can be modified by assuming those errors are correlated and have other types of probability distributions, such as the multivariate t distribution as used by Chen et al. (2010).

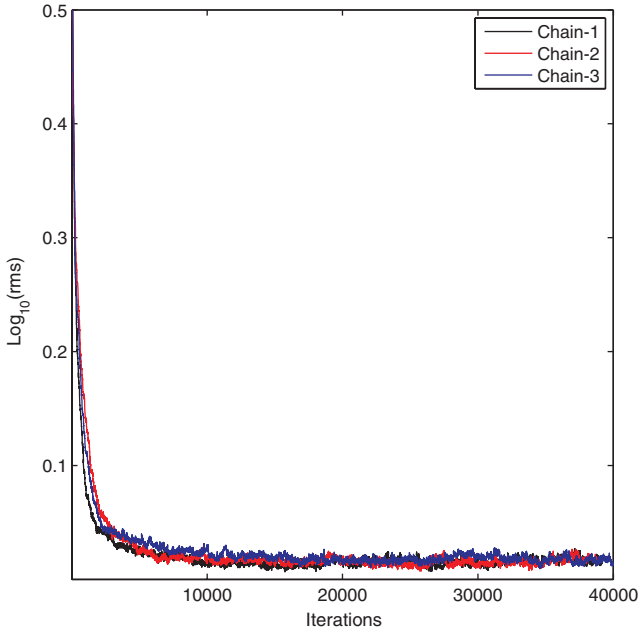

Figure 3. The rms of error-weighted misfits between the synthetic observed and calculated apparent resistivity and phases as a function of iterations by starting from the same initial model but different random seeds.

Prior probability distribution of resistivity

We can incorporate other types of information into the model through prior distributions, such as lower and upper bounds and spatial dependence or smoothness of the unknown parameters. Let a_0 and b_0 be the lower and upper bounds of the bedrock resistivity and let vectors $\mathbf{a}_k$ and $\mathbf{b}_k$ be the lower and upper bounds of the resistivity vector on interface k; hence, we have $r_0 \in (a_0, b_0)$ and $\mathbf{r}_k \in (\mathbf{a}_k, \mathbf{b}_k)$. We assume that resistivity within each layer is spatially independent in the vertical direction but dependent in the lateral direction as done by many other studies (Smith et al., 1999; Auken and Christiansen, 2004). If we further assume that r_0 is uniformly distributed on (a_0, b_0), we can write the joint prior distribution of resistivity in equation 1 as follows:

$$f(\mathbf{r}) = \text{Ind}[r_0 \in (a_0, b_0)] \times \prod_{k=1}^{m-1}\{\text{Ind}[\mathbf{r}_k \in (\mathbf{a}_k, \mathbf{b}_k)] \times f(\mathbf{r}_k)\}, \tag{3}$$

where Ind[A] represents the indicator function having the value of one if the condition A is satisfied and the value of zero, otherwise.

We define $f(\mathbf{r}_k)$ using either the first-order pairwise-difference priors as given by Besag et al. (1995) or multivariate Gaussian models, depending on applications and information available. The pairwise-difference prior is similar to the controls on the roughness or smoothness in the deterministic inversion (Smith et al., 1999), and it is given by

$$f(\mathbf{r}_k) \propto \exp\left(-\sum_{i\sim j} w_k^r(i,j)\frac{|r_{ki} - r_{kj}|}{\sigma_k^r}\right), \tag{4}$$

where σ_k^r is the standard deviation showing resistivity variation along the lateral direction and $w_k^r(i,j)$ is the weight that depends on the lateral distance between nodes i and j. The summation is over all pairs of inner nodes $i \sim j$ that are deemed to be neighbors on interface k. In this study, we let the weight to be the spatial correlation coefficient $\rho_k^r(i,j)$ between the two nodes; we calculate it using the exponential variogram as follows:

$$\rho_k^r(i,j) = \exp\left(-\frac{|x_{ki} - x_{kj}|}{\lambda_k^r}\right), \tag{5}$$

where x_{ki} and x_{kj} are the lateral coordinates of nodes i and j, and λ_k^r is the corresponding spatial correlation length.

We can use a multivariate Gaussian distribution as the prior if we have information about the mean vector $\boldsymbol{\mu}_k^r$, such as a reference model as used in Auken and Christiansen (2004). We may obtain the mean from their corresponding lower and upper bounds by $\boldsymbol{\mu}_k^r = (\mathbf{a}_k + \mathbf{b}_k)/2$. If the bounds are wide and not symmetric near the true value, the prior mean may be far away from the true values. We use the exponential variogram model to calculate the covariance matrix $\sum_k^r$, which is the same as the one used for calculating the spatial correlation coefficient in equation 5. The row i and column j component of the covariance matrix is given by $(\sigma_k^r)^2\rho_k^r(i,j)$.

Prior probability distribution of depth

We can similarly specify the prior probability distribution of depths by considering their lower and upper bounds and lateral spatial dependence. Let $\mathbf{g}_k$ and $\mathbf{h}_k$ be the lower and upper bounds of the depths to interface k from the ground surface. Let σ_k^d and λ_k^d be the standard deviation and the spatial correlation length of depths to interface k. We can define the first-order pairwise-difference and multivariate Gaussian prior models (see Appendix A).

We need to add two types of additional prior information for the current parameterization. The first is the constraints on the order of depths to avoid possible depth crossing. The depth of a shallower interface cannot be greater than that of the deeper interface, i.e., $\mathbf{d}_2 \le \mathbf{d}_3 \le \cdots \le \mathbf{d}_k \le \mathbf{d}_{k+1} \le \cdots \le \mathbf{d}_m$. The other constraint is required by the finite-element numerical forward modeling because of the use of automatic triangulations. If the angles between any two adjacent edges are less than a certain value, for example, 20° (Key and Weiss, 2006), the automatic triangulation may fail. Let $\varphi(\mathbf{d}_k)$ be the function and we have to maintain $\varphi(\mathbf{d}_k) \ge \alpha_0$, where α_0 is a threshold of the minimum angle. Consequently, we can write the prior distribution of depths in equation 1 as follows:

$$f(\mathbf{d}) = \prod_{k=2}^{m} \{\mathrm{Ind}[\mathbf{d}_k \in (\mathbf{g}_k, \mathbf{h}_k)] \times \mathrm{Ind}[\varphi(\mathbf{d}_k) \ge \alpha_0] \times f(\mathbf{d}_k)\} \times \prod_{k=2}^{m-1} \mathrm{Ind}[\mathbf{d}_{k+1} \ge \mathbf{d}_k]. \quad (6)$$

Conditional probability distributions of resistivity and depth

We can obtain the joint posterior probability distribution function (PDF) by combing the likelihood function in equation 1 and the prior distributions of resistivity and depths given in equations 3 and 6 as follows

$$f(\mathbf{r}, \mathbf{d}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{r}, \mathbf{d}) \times \mathrm{Ind}[r_0 \in (a_0, b_0)] \times \prod_{k=1}^{m-1} \{\mathrm{Ind}[\mathbf{r}_k \in (\mathbf{a}_k, \mathbf{b}_k)] \times f(\mathbf{r}_k)\} \times \prod_{k=2}^{m} \{\mathrm{Ind}[\mathbf{d}_k \in (\mathbf{g}_k, \mathbf{h}_k)] \times \mathrm{Ind}[\varphi(\mathbf{d}_k) \ge \alpha_0] \times f(\mathbf{d}_k)\} \times \prod_{k=2}^{m-1} \mathrm{Ind}[\mathbf{d}_{k+1} \ge \mathbf{d}_k]. \quad (7)$$

The joint posterior PDF in equation 7 is complicated, and it is not possible to obtain analytical solutions. Instead, we use MCMC methods to obtain many samples from the joint PDF. To implement MCMC algorithms, we first derive conditional probability distributions for resistivity and depths on each layer. For the bedrock resistivity r_0, we can obtain its conditional probability distribution by dropping all the terms in equation 7 that are not directly related to the bedrock resistivity, which is given as

$$f(r_0|\cdot) \propto f(\mathbf{z}|\mathbf{r}, \mathbf{d}) \times \mathrm{Ind}[r_0 \in (a_0, b_0)], \quad (8)$$

where $f(r_0|\cdot)$ represents the conditional probability distribution of r_0, given all other unknown parameters and data, and the centered dot represents all other unknowns and data.

Similarly, we can obtain the conditional probability distribution of resistivity vector $\mathbf{r}_k$ in layer k as follows:

$$f(\mathbf{r}_k|\bullet) \propto f(\mathbf{z}|\mathbf{r}, \mathbf{d}) \times \mathrm{Ind}[\mathbf{r}_k \in (\mathbf{a}_k, \mathbf{b}_k)] \times f(\mathbf{r}_k). \quad (9)$$

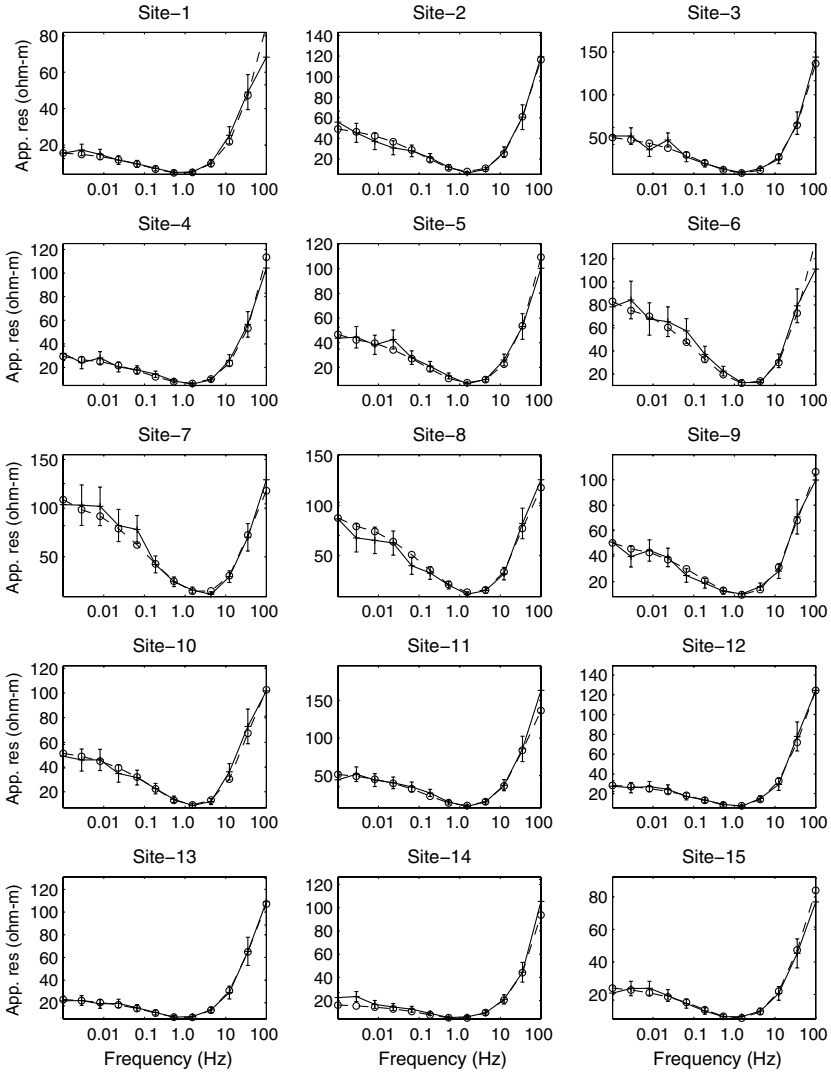

Figure 4. Comparison between the synthetic observed (solid lines) and calculated (dashed lines with circles) TM-mode apparent resistivity using the model formed from the estimated medians of individual parameters. The vertical segments show the 95% error bars around data.

Because the order constraints for depths at interface $k = 2$, $3 \le k \le m-1$, and $k = m$ are different, we need to separate them. For $3 \le k \le m-1$, we have

$$f(\mathbf{d}_k|\bullet) \propto f(\mathbf{z}|\mathbf{r},\mathbf{d}) \times \text{Ind}[\mathbf{d}_k \in (\mathbf{g}_k, \mathbf{h}_k)] \times f(\mathbf{d}_k) \times \text{Ind}[\mathbf{d}_k \ge \mathbf{d}_{k-1}] \times \text{Ind}[\mathbf{d}_{k+1} \ge \mathbf{d}_k] \times \text{Ind}[\varphi(\mathbf{d}_k) \ge \alpha_0]. \quad (10)$$

We need to drop $\text{Ind}[\mathbf{d}_k \ge \mathbf{d}_{k-1}]$ when $k = 2$ and $\text{Ind}[\mathbf{d}_{k+1} \ge \mathbf{d}_k]$ when $k = m$ in equation 10.

We can combine the order constraints with the range constraints (i.e., lower and upper bounds) to simplify equation 10 by introducing conditional lower bound $\mathbf{l}_k$ and upper bound $\mathbf{u}_k$

$$\mathbf{l}_k = \begin{cases} \mathbf{g}_k & \text{if } k = 2 \\ \max\{\mathbf{g}_k, \mathbf{d}_{k-1}\} & \text{otherwise} \end{cases}, \quad (11)$$

and

$$\mathbf{u}_k = \begin{cases} \mathbf{h}_k & \text{if } k = m \\ \min\{\mathbf{h}_k, \mathbf{d}_{k+1}\} & \text{otherwise} \end{cases}. \quad (12)$$

Consequently, equation 10 becomes the following:

$$f(\mathbf{d}_k|\bullet) \propto f(\mathbf{z}|\mathbf{r},\mathbf{d}) \times \text{Ind}[\mathbf{d}_k \in (\mathbf{l}_k, \mathbf{u}_k)] \times \text{Ind}[\varphi(\mathbf{d}_k) \ge \alpha_0] \times f(\mathbf{d}_k). \quad (13)$$

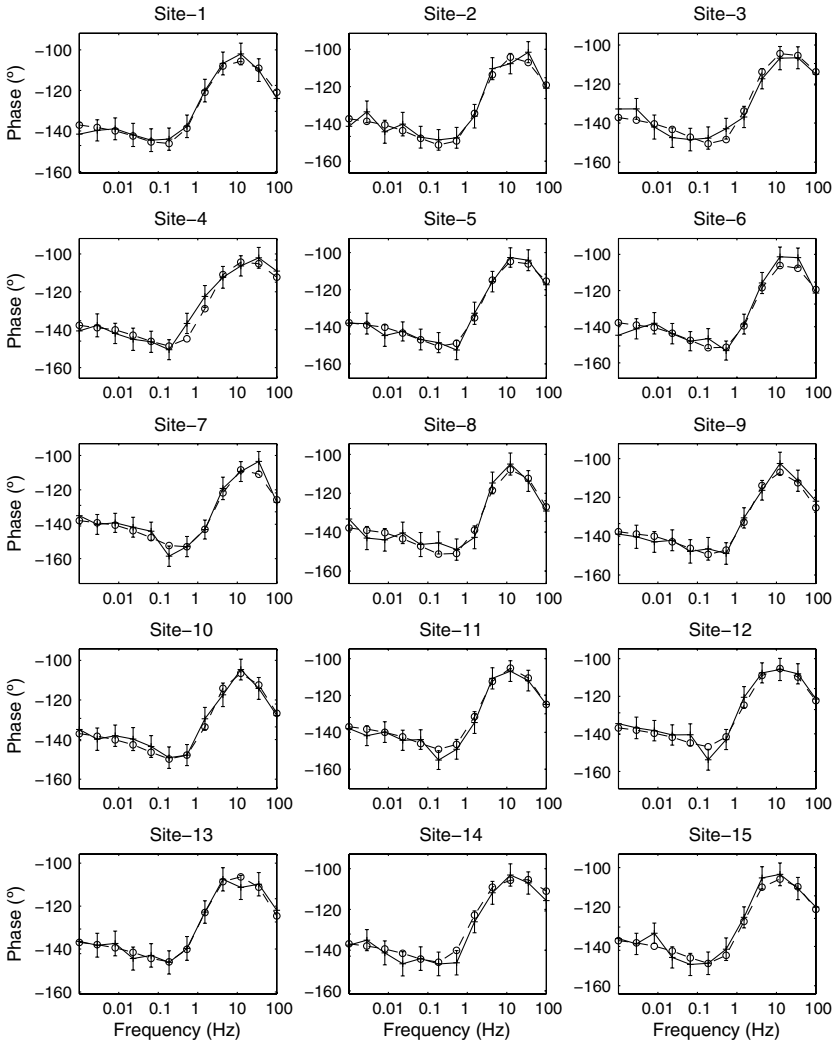

Figure 5. Comparison between the synthetic observed (solid lines) and calculated (dashed lines with circles) TM-mode phases using the model formed from the estimated medians of individual parameters. The vertical segments show the 95% error bars around data.

Markov chain Monte Carlo sampling strategies

We use a hybrid sampling method and block sampling strategies with random scans (Tienery, 1994) for resistivity and depth vectors. The hybrid method includes multivariate Metropolis-Hastings (MMH) (Metropolis, 1953; Hastings, 1970) and multivariate slice sampling methods (MSS) (Neal, 2003; Chen et al., 2007). At each iteration, we first randomly select an interface (e.g., the kth layer) and then update resistivity vector $\mathbf{r}_k$, depth vector $\mathbf{d}_k$, or both resistivity and depth vectors $\mathbf{r}_k$ and $\mathbf{d}_k$ at random using MMH or MSS sampling methods. In the following, we outline the MMH algorithm for sampling depths; the algorithms for sampling resistivity and for simultaneously sampling resistivity and depth are similar. The MSS algorithms for sampling resistivity, depths, or both resistivity and depths are similar to those given by Chen et al. (2007).

Step 1 At the ith iteration, randomly select an interface for updating (say the kth interface).

Step 2 Draw a candidate sample $\mathbf{d}_k^*$ from a multivariate Gaussian distribution truncated by the conditional bounds $(\mathbf{l}_k, \mathbf{u}_k)$ with mean at the current value $\mathbf{d}_k^{(i)}$ and covariance matrix $\sum_p$ (see Appendix B); thus $\mathbf{d}_k^* \in (\mathbf{l}_k, \mathbf{u}_k)$.

Step 3 Check $\varphi(\mathbf{d}_k^*)$ for a given cutoff angle α_0 (e.g., 20°). If $\varphi(\mathbf{d}_k^*) \ge \alpha_0$, go to step 4; go to Step 2, otherwise.

Step 4 Calculate the accepting ratio β by

$$\beta = \min\left\{1, \frac{f(\mathbf{z}|\mathbf{r},\mathbf{d}_2^{(i)},\mathbf{d}_3^{(i)},\cdots,\mathbf{d}_k^*,\cdots,\mathbf{d}_m^{(i)})}{f(\mathbf{z}|\mathbf{r},\mathbf{d}_2^{(i)},\mathbf{d}_3^{(i)},\cdots,\mathbf{d}_k^{(i)},\cdots,\mathbf{d}_m^{(i)})} \times \frac{f(\mathbf{d}_k^*)}{f(\mathbf{d}_k^{(i)})}\right\}.$$

Draw a sample u from the uniform distribution on (0, 1). If $u < \beta$, let $\mathbf{d}_k^{(i+1)} = \mathbf{d}_k^*$; otherwise, let $\mathbf{d}_k^{(i+1)} = \mathbf{d}_k^{(i)}$.

The above algorithm is a way to draw samples from the conditional distribution given in equation 13. Specifically, step 2 enforces the condition $\text{ind}[\mathbf{d}_k \in (\mathbf{l}_k, \mathbf{u}_k)] = 1$ whereas step 3 enforces the condition $\text{ind}[\varphi(\mathbf{d}_k) \ge \alpha_0] = 1$. Step 4 is the standard Metropolis-Hastings procedure to draw samples from the posterior distribution that is proportional to the product of the

likelihood function $f(\mathbf{z}|\mathbf{r},\mathbf{d})$ and the prior $f(\mathbf{d}_k)$. After many runs of the algorithm, the resulting depth vectors are approximately samples from the conditional distribution in equation 13 (Tienery, 1994). The detailed sampling procedures are given in Appendix C.

SYNTHETIC EXAMPLE

Synthetic true model

We first apply the Bayesian model to a synthetic case that mimics typical geothermal field conditions. Figure 2 shows the true 2D resistivity cross-section with five (i.e., the overburden, clay cap, reservoir, transition, and bedrock) layers, which spans 20 km laterally (i.e., from $x = -10$ km to $x = 10$ km). We use a geostatistical model based on the exponential variogram to generate depths to each interface and resistivity values in the overburden, clay cap, and transition layers. The geostatistical parameters for generating the 2D cross section are given in Table 1. For the reservoir layer, we let resistivity linearly increase from 20 ohm-m at the left (i.e., $x = -10$ km) and right (i.e., $x = 10$ km) edges to 460 ohm-m near the center (i.e., $x = -1$, 0, and 1 km). The bedrock resistivity is assumed to be constant and equal to 30 ohm-m.

We define 21 nodes for each interface, starting from $x = -10$ km and increasing with a fixed increment of 1 km. To minimize boundary effects, we choose a very large domain for forward simulations, which is from -350 to 350 km along both the x- and z-axes.

Synthetic MT data and prior information

We generated noisy synthetic data using the adaptive finite-element numerical forward model at 15 sites shown as small triangles in Figure 2, which correspond to $x = -9, -7, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 7$, and 9 km. We use 12 frequencies that are equally spaced in the logarithmic scale between 0.001 and 100 Hz. We first calculated the true complex impedances for the given true 2D resistivity structure, then added zero-mean Gaussian random noise whose standard deviation is equal to 5% of their

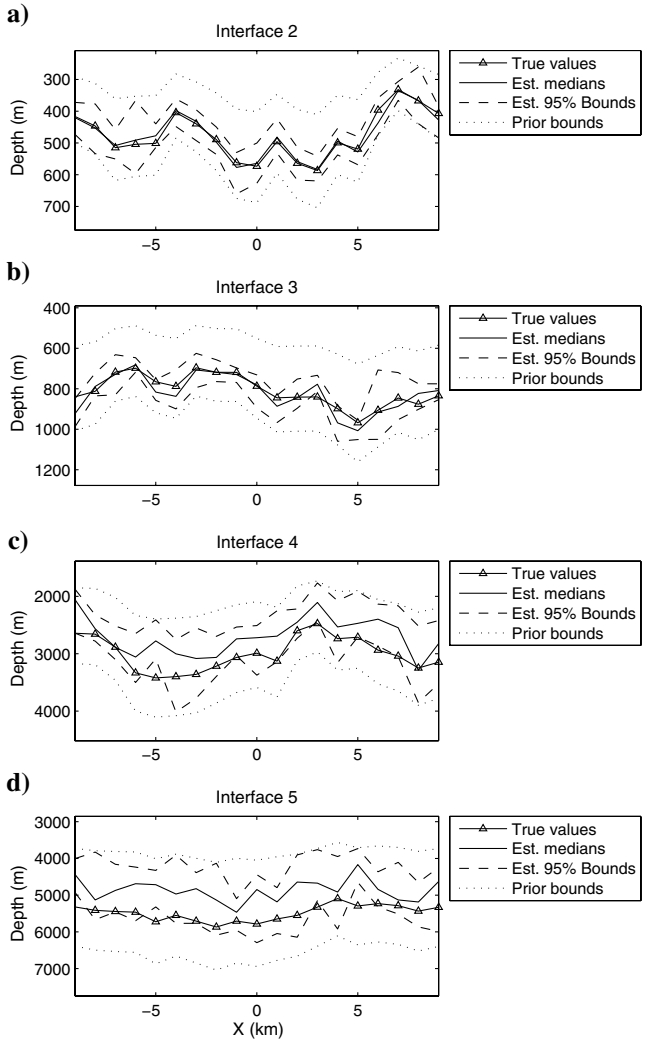

Figure 6. Comparison between the true depths (solid lines with triangles) and the estimated results. The solid and dashed lines are the estimated medians and 95% bounds, and the dotted lines are the prior bounds.

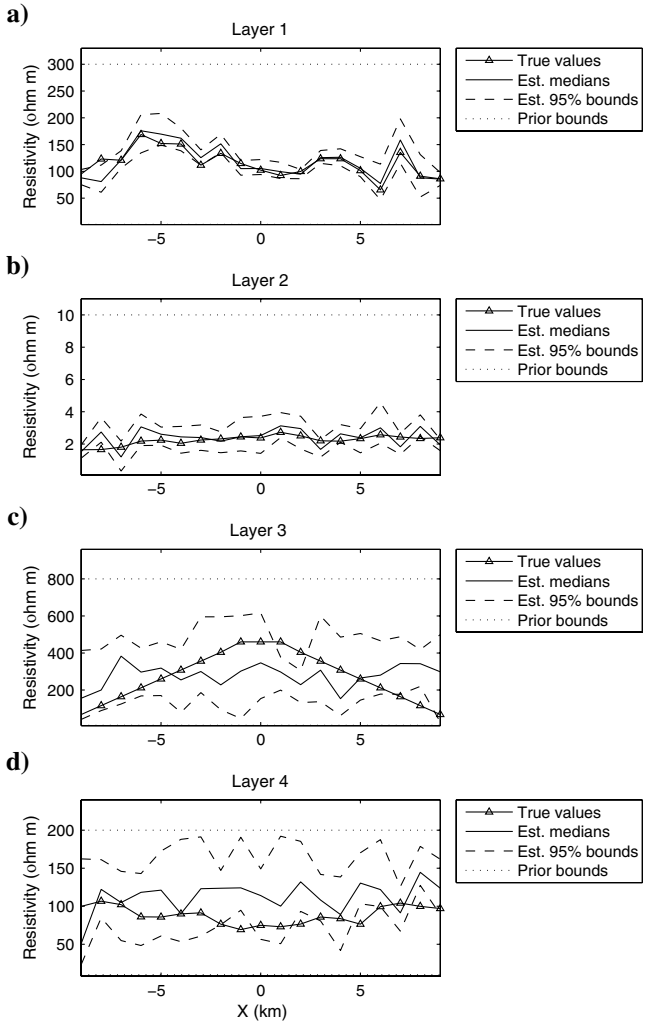

Figure 7. Comparison between the true resistivity (solid lines with triangles) and the estimated results. The solid and dashed lines are the estimated medians and 95% bounds, and the dotted lines are the prior bounds.

corresponding amplitudes, and finally converted the noisy complex impedances to apparent resistivity and phases.

We choose prior bounds for all unknown parameters as part of model definition based on our experience for typical geothermal fields. The prior lower and upper bounds of depths are the values that are 30% smaller and 20% larger than their corresponding true values. Such asymmetrical specifications can reduce the chance of starting from the true model parameters if we randomly pick the initial values using a Gaussian distribution. The prior lower and upper bounds of resistivity for the overburden, clay cap, reservoir, transition, and bedrock layers are (1, 300), (0.1, 10), (10, 800), (10, 200), and (10, 100) ohm-m, respectively. We use the first-order pairwise-difference priors (see equation 4) for the synthetic study because we do not have good information on the means of resistivity and depths. The standard deviations of resistivity and depths for each layer are assumed to be 20% of their corresponding means over all the nodes on the layer; the spatial correlation lengths are set to be the true values.

Table 2. Comparison between the true and estimated medians and uncertainty information.

Parameters		The rms of the relative differences between the true and estimated values		Relative widths of the predictive intervals	
		Mean	Maximum	Mean	Maximum
Depth	Layer 1	0.0325	0.1037	0.1183	0.2438
	Layer 2	0.0478	0.0996	0.0872	0.1874
	Layer 3	0.1179	0.2224	0.1546	0.2129
	Layer 4	0.1241	0.2130	0.1489	0.2413
Resistivity	Layer 1	0.1165	0.3415	0.1856	0.4489
	Layer 2	0.2479	0.6574	0.3314	0.7711
	Layer 3	1.0711	3.3905	0.6829	1.3792
	Layer 4	0.4327	0.7915	0.4535	1.3592

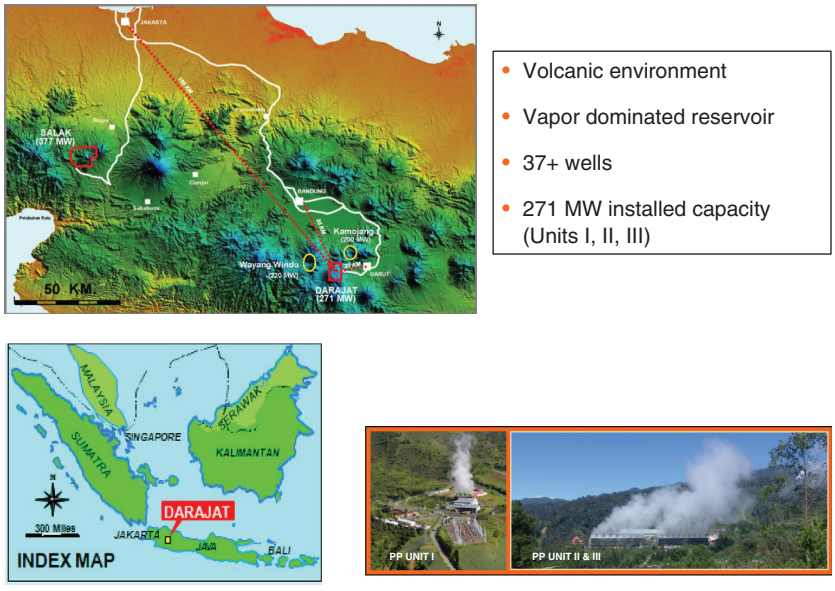

Figure 8. The location and main features of the Darajart geothermal site in Indonesia.

Inversion of 2D synthetic MT data

We use MCMC methods to draw many samples from the joint posterior distribution given by equation 7. We start from the medians of the prior bounds of resistivity and from the 25% quantiles of the prior bounds of depths. We run three chains using the same initial values but different random seeds for 80,000 iterations and only keep samples at every two iterations (i.e., a thinning of two) to save disk space. The use of different random seeds makes those chains move in different paths, which is equivalent to starting from different initial values.

We monitor the convergence of the chains using the rms of error-weighted differences between the data and the simulated results. To quantify the discrepancy among the three chains, we use the potential scale reduction factor (PSRF) defined by Gelman and Rubin (1992), which is a measure of the between-chain variability relative to the within-chain variability. If the scale reduction score is less than 1.2, the Markov chains are considered converged; otherwise, more runs are needed.

Figure 3 shows the rms of the differences between the synthetic data and the simulated results for the three chains as a function of iterations after thinning. Although we start from the same initial values, the rms of the three chains is very different in the early stage of simulations because they explore the joint posterior distribution in different paths. After about 20,000 iterations, the three chains visually are approaching the same distribution; they give the PSRF value of 1.13. Because the PSRF value is less than the threshold value of 1.2, we consider the three chains converged and use all those samples in the later part of the chains to estimate the medians and 95% predictive intervals of unknown parameters.

We choose to monitor the rms between the data and the simulated results to provide a balance between the reliability of inversion results and the cost for running long chains. The use of methods for monitoring PSRF of individual parameters (e.g., Gelman and Rubin, 1992) or for monitoring multivariate PSRF of all parameters (e.g., Brooks and Gelman, 1998) calls for very long chains, which is not feasible for this study. The convergence diagnostics of MCMC chains are very difficult, especially for high dimensional problems; they are active research areas that need breakthroughs both in theory and in practice (Cowles and Carlin, 1996).

Inversion results and uncertainty

Unlike deterministic inversion methods, which typically provide a single optimal solution under

some criteria, the stochastic approach provides 60,000 solutions (20,000 from each chain) that fit the data (see Figure 3). With those results, we can obtain detailed information on each unknown parameter, such as its mean, median, mode, standard deviation, and 95% predictive interval. We can use the means, medians, or modes of individual parameters to form a single solution.

Figures 4 and 5 compare the TM-mode apparent resistivity and phase data (solid lines) with those calculated from the estimated medians of individual parameters (dashed lines with circles) at the 15 receivers; the 95% error bars around data are shown as vertical line segments. The solution formed from the medians of individual parameters fits the synthetic data quite well, with the rms of 1.0337 for the TM-mode data and the rms of 1.0425 for the TE-mode data. Although the solution formed by the estimated individual medians may not be necessarily optimal, it often has good data misfits as shown in this study. In applications where a single optimal solution is needed, we can run a deterministic inversion algorithm by starting from the initial model formed from the estimated medians as suggested by Chen et al. (2008).

We compare the inversion results with the true resistivity and depths to demonstrate the effectiveness of the developed method for parameter estimation. Figures 6 and 7 show the true values (solid lines with triangles), the estimated medians (solid lines), the 95% lower and upper bounds (dashed lines), and the prior bounds (dotted lines), over the range covered by the 15 receivers (i.e., from $x = -9$ km to $x = 9$ km). Table 2 is a quantitative summary of the comparison, in which the 3rd and 4th columns show the rms and maximum of the differences between the estimated medians and true values, and the 5th and 6th columns show the averaged and maximal relative width of 95% predictive intervals.

We can see the accuracy of the parameter estimation varies from one layer to another. For the first (i.e., overburden) and the second (i.e., clay cap) layers, we have very good estimates of the true values with small uncertainty. For those layers, the estimated results are not sensitive to the prior bounds and other prior information, such as the standard deviations (σ_k^r, and σ_k^d) and spatial correlation lengths (λ_k^r, and λ_k^d) in the pairwise-difference priors. This is because the estimation results are primarily determined by the MT data.

For the third (i.e., reservoir) and the fourth (i.e., transition) layers, both the resistivity and depths get updated by conditioning to the MT data as their posterior uncertainty bounds (dashed lines) are narrower than the corresponding prior bounds (dotted lines); their estimated medians roughly follow the trends of the true values. However, the estimated results are less accurate and subject to a large degree of uncertainty. Such results are consistent with those found by Kumar et al. (2010). In their studies, they inverted a 3D MT data set using the least-square estimation method by starting from a half-space model, from a model formed using 1D stochastic inversion results, and from a model formed using 2D least-square estimation results. The differences in the estimates of the overburden and clay cap are small but in the estimates of the reservoir and deep layers are significant.

The increasing uncertainty with depth is likely caused by a combination of the decay of the electromagnetic fields as a function of skin depth and the cumulative effects of parameter uncertainty as energy propagates farther in the model. The uncertainty of a parameter at a given depth is a function of uncertainty of all the parameters above the depth because the energy must pass through the media before interacting with the parameter (either resistivity or thickness of a resistor).

The prior information on the parameters in the reservoir and transition layers is important for the estimation, especially the lower and upper bounds because MT data provide limited information on them. Informative priors from other sources of information, such as seismic, gravity, or electromagnetic surveys, may significantly improve the inversion results. For the bedrock layer, the estimated median provides a very accurate estimate of the true value (i.e., 30 ohm-m), with the estimated median of 29.77 ohm-m and the 95% predictive interval of (28.87, 30.74) ohm-m. This is because in the current model, we assume the bedrock resistivity is constant and it can get information from all the receivers.

FIELD EXAMPLE

Darajat geothermal site

The Darajat geothermal field is located about 150 km southeast of Jakarta and 35 km southeast of Bandung, the capital of West Java, Indonesia. The field is situated along the eastern side of a range of volcanic centers nearly 30 km in length, and lies about 9 km southwest of the producing Kamojang geothermal field and 10 km east of Wayang Windu geothermal field (see Figure 8). Darajat is a vapor dominated geothermal reservoir that currently supplies steam to generate 271 MW from three power plant units (Rejeki et al, 2010).

Resistivity methods dominate the geophysical tools applied during exploration for geothermal reservoirs because low resistivity is closely correlated with the low resistivity smectite clay that caps most geothermal systems (Cumming, 2009). Over the Darajat Geothermal Field, 90 MT stations have been deployed (Figure 9). The MT data collected over Darajat have provided important constraints on the interpretation of the areal extent and location of the Darajat field during the exploration and development of the field. These data when integrated with the surface geologic, geochemistry, and the subsurface drilling data, have provided important

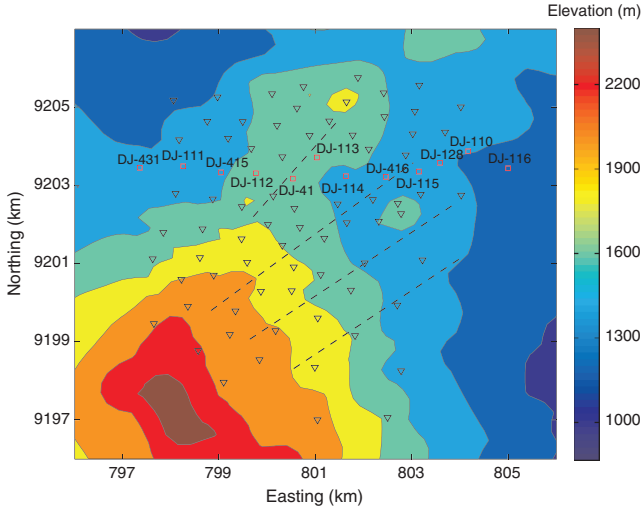

Figure 9. The locations of MT stations (squares and triangles) and the interpreted faults (dashed lines). The 2D profile for inversion starts from DJ-431 and ends at DJ-116.

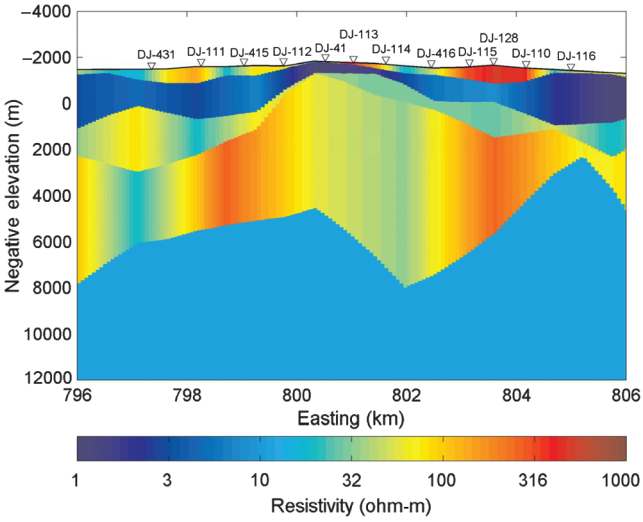

Figure 10. The 2D resistivity structure based on the estimated medians of parameters from the field TM-mode data. The small triangles show the locations of the MT receivers used for inversion.

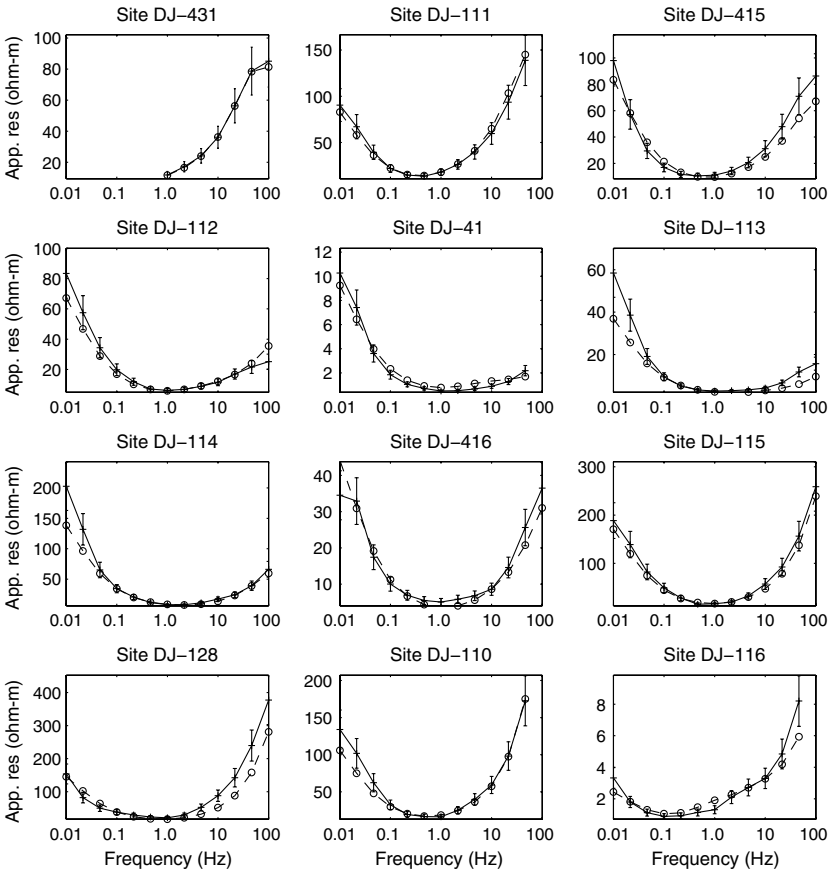

Figure 11. Comparison between the observed (solid lines) and calculated (dashed lines with circles) TM-mode apparent resistivity using the model formed from the estimated medians of individual parameters. The vertical segments show the 95% error bars around data.

constraints on the location and geometry of the clay cap and reservoir for the Darajat Geothermal system (Rejeki et al, 2010).

Such dense geophysical data allow us to estimate the underground image using 1D, 2D, or 3D inversion methods. For this study, we have selected MT data at the sites on or near the profile starting from DJ-431 and ending at DJ-116, as shown in Figure 9. This east–west profile includes 12 sites and is approximately perpendicular to the observed geoelectric strike direction. At each site, we have apparent resistivity and phases with frequencies from 0.01 to 100 Hz for both TE and TM modes. For the 2D inversion, we use only the TM-mode data because they are more robust than the TE-mode data to non-2D effects, such as finite strike and static shifts (Newman et al., 2008).

Inversion of field data

We use a similar parameterization to the synthetic study and assume that they have five interfaces and each has 27 inner estimation nodes. We use the same prior bounds as the synthetic study for resistivity. We use an initial model that was constructed from the 3D deterministic inversion results along the same profile.

Figure 10 shows the resultant 2D cross section based on estimated medians of resistivity. From the cross section, we can see the clearly defined resistivity layers. The estimated shapes and geometries of the layers have a good correspondence with other types of information collected at this site. For example, the locations and thickness of the low resistivity layer has a very good correspondence with the system clay cap and the location of the Darajat geothermal reservoir. The doming up of the high resistivity layer corresponds with high gravity that correlates with the intrusive rocks found beneath the field. Even for the deepest (i.e., fifth) layer, where the resistivity decreases again, the interface between the fourth and fifth layers corresponds well with the maximum depth of seismicity (or earthquakes) at the site (Rejeki et al, 2010).

Figures 11 and 12 show the data misfits for the inversion based on the estimated medians, with a mean rms of 1.88 and a standard deviation of 0.09. As we can see, most of the calculated MT responses from the estimated medians are within the 95% error bounds.

Comparison with 3D deterministic inversion

We visually compare the results with those obtained from a 3D deterministic inversion. The 3D inversion uses the Gauss-Newton method developed by Newman et al. (2008) to fit the MT data collected from 85 stations and with frequencies ranging from 0.01 to 100 Hz. The inversion starts from a half-space with the resistivity of 10 ohm-m; the off-diagonal elements of the impedance tensor were fit to the rms of 1.4. The 3D inversion method is very different from the stochastic inversion in terms of (1) parameterization (i.e., grid-based versus sharp boundary

parameterizations), (2) data used (i.e., TE and TM mode data at all the sites versus TM mode data at only 12 sites), (3) methods for finding solution (i.e., Gauss-Newton versus MCMC sampling based methods), and (4) smoothing (i.e., regularization versus prior models). However, as shown in Figure 13, where we superimpose the stochastically estimated interfaces on the 3D smoothing results, we can see that overall they are very similar. Specifically, we can see the laterally variable overburden, the persistent clay cap, the doming up transition and reservoir, and the deepest interface.

Quantification of uncertainty

One of the main advantages of stochastic inversion methods is that they provide extensive uncertainty information on the estimated parameters. Figures 14 and 15 show the estimated medians and 95% bounds of depths and resistivity, respectively. Because of large differences among those quantities, the bounds for smaller values are hardly visible. However, we can see the trends that uncertainty in the estimates increases with increasing of depths, and uncertainty also increases toward both ends, where data coverage is reduced.

Besides providing extensive uncertainty information on the parameters that we directly estimate, stochastic inversion also provides exhaustive information on other quantities of interest which are functions of the estimated parameters. For example, we may want to know the slopes, which are defined as angles relative to the horizontal direction and different from those defined for angle controls within the mesh, from one interface to another at a specific location. We can obtain the PDFs of the quantities by first calculating them for each MCMC sample and then applying a density estimation method to the resultant values (Gilks et al., 1996). Figure 16a and 16b shows the estimated PDFs of the slopes at $x = 800$ km and of the radii at the tipping point accordingly using the density estimation method given by Venables and Ripley (2002). As shown in Figure 16a, even if there are considerable uncertainty on those values, we can see that from interface 1 (solid curve) to interface 2 (dashed curve), the slope increases significantly. For geothermal applications, we may want to know the curvature or radius of the doming up part. As shown in Figure 16b, the radius at the tipping point decreases, starting from the ground surface. This means the doming of the tipping point increases with depth. This is a typical feature of geothermal reservoirs. In fact, we can make inferences on many other quantities of interest.

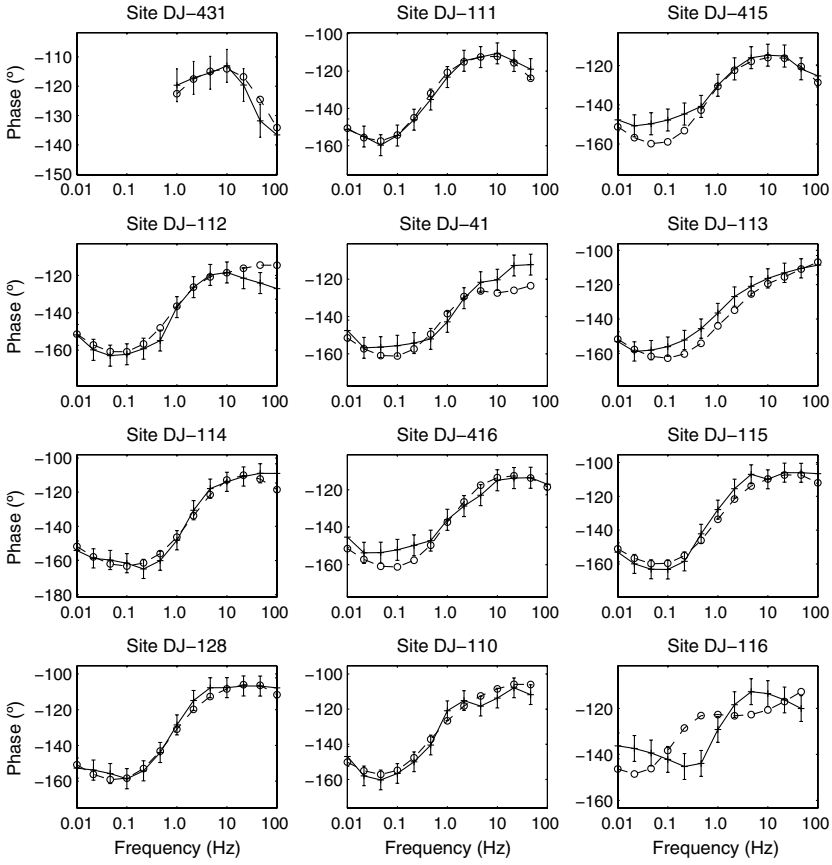

Figure 12. Comparison between the observed (solid lines) and calculated (dashed lines with circles) TM-mode phases using the model formed from the estimated medians of individual parameters. The vertical segments show the 95% error bars around data.

DISCUSSION

The developed stochastic method has several different features from the Gauss-Newton based iterative approaches by Smith et al. (1999) and de Groot-Hedlin and Constable (2004). The choice of initial models plays different roles in the two types of deterministic methods. For the deterministic inversion methods, the optimal

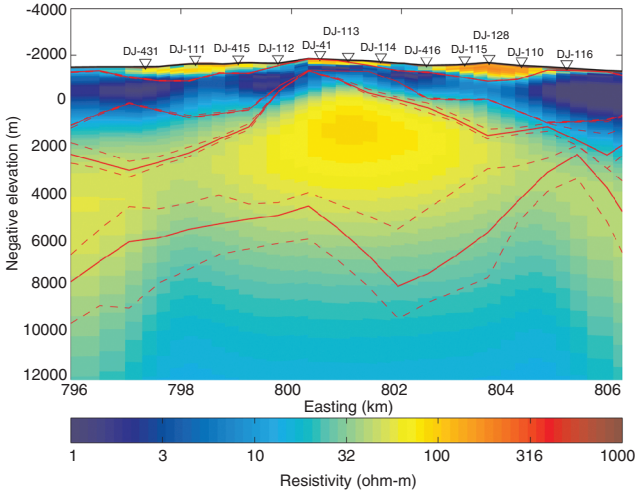

Figure 13. Comparison of estimated 2D resistivity structure using the 2D stochastic and 3D Gauss-Newton based deterministic inversion methods. The color image is plotted using the 3D inversion results; the solid and dashed lines show the estimated interfaces and their 95% bounds from the stochastic inversion.

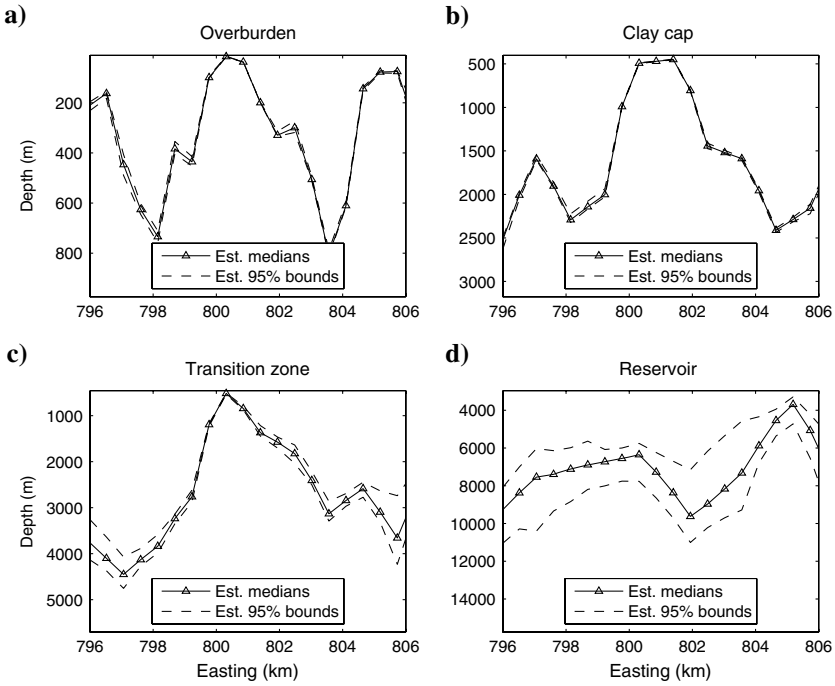

Figure 14. Estimated depths to each interface from the ground surface and their associated uncertainty. The solid lines with triangles and the dashed lines are the estimated medians and 95% bounds, respectively.

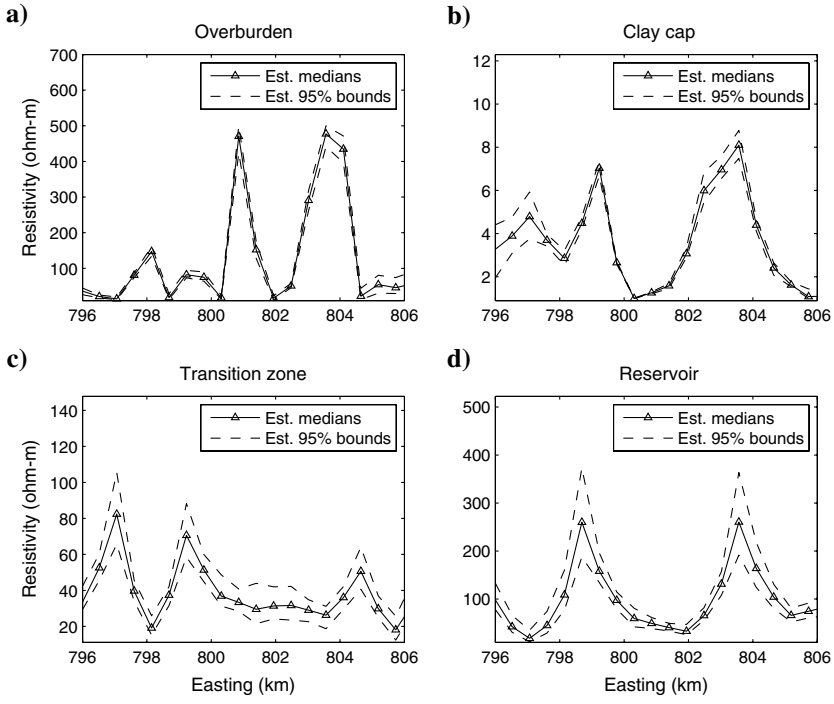

Figure 15. Estimated resistivity in each layer and their associated uncertainty. The solid lines with triangles and the dashed lines are the estimated medians and 95% bounds, respectively.

solution depends on the initial models; some initial models may lead to local optimal solutions. For the MCMC-based inversion method developed in this study, the choice of initial models only affects the speed of convergence to the underlying stationary distribution but not the inversion results. In fact, to detect false convergence, we suggest running at least three chains by starting from different initial values or from the same initial values using different random seeds, and running more chains if computing resources allow to do so.

Both stochastic and deterministic approaches provide uncertainty information on unknown model parameters, but they are different. Deterministic methods first calculate the covariance matrix of unknown parameters using the Jacobian derived at the optimal solution and then use the square root of the diagonal element as the measure of uncertainty on the parameters (Auken and Christiansen, 2004). This is a good approximation if the inverse problems are mildly nonlinear and the optimal solution found indeed is global. However, as pointed out by Tarantola (2005), for complex inverse problems, the above simple method of analyzing uncertainty does not make sense. The direct examination of the prior and posterior probability distribution is the only mean that we may have. The MCMC-based stochastic method estimates uncertainty by deriving marginal posterior probability density from many samples; this method is valid regardless of complexity of inverse problems. As demonstrated by Chen et al. (2008) and Trainor-Guitton and Hoversten (2012), the uncertainty derived from the covariance based methods are typically underestimate the true uncertainty.

The main limitation of stochastic inversion methods is the heavy demand for computing resources. The MCMC-based inversion method typically needs to run forward simulations tens of thousands times. For example, for the synthetic study, it takes about 140 hours to run 80,000 iterations using 11 processors with the speed of 3 GHz on a single modern desktop computer. This can be sped up by developing more efficient sampling methods to reduce the total number of iterations or using more processors to reduce the time for forward modeling. For example, to run 80,000 iterations for the synthetic study, it takes about 89 hours if using 21 processors and 70 hours if using 61 processors. However, compared with deterministic inversion methods, stochastic approaches still need significant more computing resources. For projects with limited resources, stochastic methods may not be the ones to use.

Another limitation of the stochastic method is that it does not provide a single solution that is

often needed in practice as deterministic methods. The solution formed from the median, mean, or mode of each individual parameter may not necessarily be the optimal solution. Therefore, we should use stochastic and deterministic inversion methods together, as suggested by Chen et al. (2008). For example, we can start from a range of different initial values using stochastic methods to obtain the medians, means, modes, and predictive intervals of unknown parameters and then run deterministic inversion methods by starting from the initial model formed from the medians, means, or modes obtained from the stochastic methods. We can also use deterministic approaches first to find the optimal solution and then use stochastic methods to verify whether the optimal solution found is global or not. If it is global, stochastic methods can be used to obtain uncertainty information. Otherwise, the stochastic inversion results may provide useful information for improving the deterministic inversion.

The sharp boundary parameterization makes stochastic inversion of 2D MT data tractable; it is also useful for applications where geologic units have sharp discontinuities (Smith et al., 1999) or resistivity structures present sharp contrasts along the vertical direction (Auken and Christiansen, 2004). However, it is limited for applications with large discontinuities along the lateral direction, such as interfaces with a large dip (e.g., faults) or even terminating layers. In addition, the results of inversions may be affected by the preset total number of layers and the total number and spacing of inner nodes in each layer.

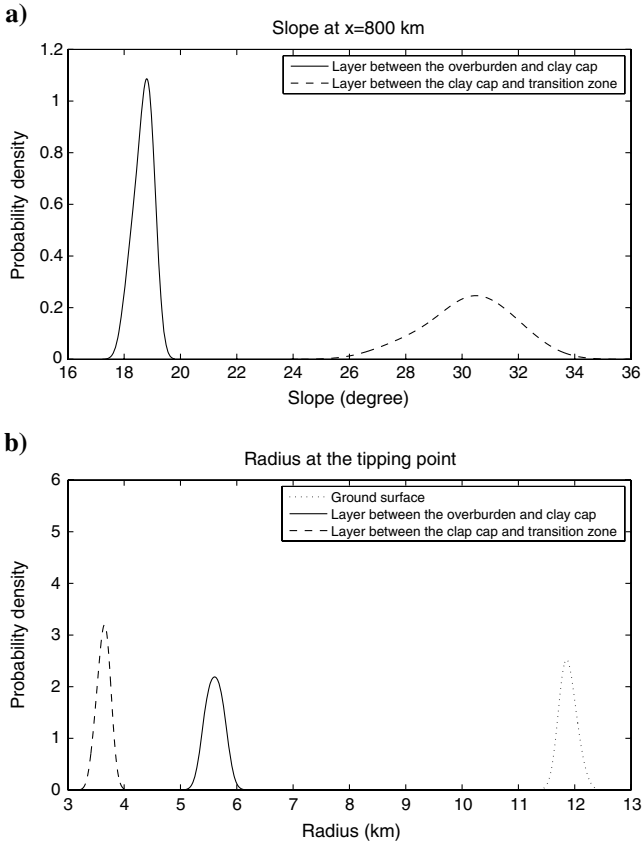

Figure 16. Estimated probability densities of (a) slopes at x = 800 km and (b) radii at the tipping point.

CONCLUSIONS

We developed a stochastic approach for inverting 2D MT data using a sharp boundary parameterization and applied it to both synthetic and field MT data sets. The synthetic case study shows that the developed model is effective in recovering the true interface locations and resistivity for first two layers and provides many samples of all unknown parameters. The obtained samples allow us to evaluate the means, variances, modes, predictive intervals, and marginal probability distributions of unknown parameters, all of which are useful for quantifying the uncertainty associated with inversion. Results of the field case study show that the developed Bayesian model is effective for estimating depths to interfaces and resistivity under the field conditions. Comparison with the results obtained from other sources of information shows that the stochastic inversion results are consistent. The stochastic inversion provides not only estimates of unknown parameters but also extensive information on the uncertainty. These allow for estimation of other types of parameters that are functions of the estimated unknown parameters.

Deterministic and stochastic approaches have their own advantages and disadvantages and can be used together for practical problems as deterministic methods can converge fast so that it can significantly shorten the burn-in period. Deterministic methods provide a single optimal solution under various constraints, whereas stochastic methods provide exhaustive information on uncertainty. The combined use of deterministic and stochastic methods can take advantage of both methods and thus provide more reliable information on unknown parameters than either approach alone.

We demonstrated the use of stochastic inversion methods for geophysical inverse problems beyond 1D using a sharp boundary parameterization. Because we have a fast forward model to simulate 2D EM fields, we can certainly develop stochastic methods for inverting 2D MT data using the grid-based parameterization as commonly used by deterministic inversion methods. However, because the total number of unknowns often causes slow convergence in MCMC sampling, we can use suitable prior models by incorporating other types of information to reduce the effective total number of unknowns as done by deterministic methods to use regularization or smoothing. Stochastic inversion of 3D MT data currently is still challenging and faces several challenges: (1) availability of fast forward modeling algorithms and codes; (2) suitable parameterizations. Reduced domain or approximation response surface methods may work. This is a direction for future research.

ACKNOWLEDGMENTS

The work is funded by Chevron Energy Technology Company. We thank Chevron for permission to publish this work. K. Key acknowledges support from the Seafloor Electromagnetic Methods Consortium at Scripps Institution of Oceanography. This work was also partially supported by the U.S. Department of Energy and LBNL under Contract No. DE-AC02-05CH11231.

APPENDIX A

PRIOR PROBABILITY DISTRIBUTION OF DEPTH

We follow the same way as resistivity to define $f(\mathbf{d}_k)$ using either the first-order pairwise-difference prior or multivariate Gaussian models. The pairwise-difference prior is given by

$$f(\mathbf{d}_k) \propto \exp\left(-\sum_{i\sim j} w_k^d(i,j)\frac{|d_{ki}-d_{kj}|}{\sigma_k^d}\right), \tag{A-1}$$

where σ_k^d is the standard deviation of depths to interface k and $w_k^d(i,j)$ is the weight that depends on the lateral distance between the ith and jth nodes. We assign the weight to the spatial correlation coefficient $\rho_k^d(i,j)$ between the ith and jth nodes, which are calculated using the exponential variogram as follows:

$$\rho_k^d(i,j) = \exp\left(-\frac{|x_{ki}-x_{kj}|}{\lambda_k^d}\right), \tag{A-2}$$

where x_{ki} and x_{kj} are the lateral coordinates of the ith and jth nodes and λ_k^d is the corresponding spatial correlation length.

We can use a multivariate Gaussian distribution as the prior with the mean vector $\boldsymbol{\mu}_k^d$, either given from a reference model or obtained from their corresponding lower and upper bounds. We use the exponential variogram model to calculate the covariance matrix $\sum_k^d$, which is the same as the one used for calculating the spatial correlation coefficient. The ith row and jth column component $\sum_k^d(i,j)$ of the matrix is given by $(\sigma_k^d)^2\rho_k^d(i,j)$.

APPENDIX B

COVARIANCE MATRIX OF PROPOSAL DISTRIBUTIONS

We determine the ith row and jth column of the covariance matrix Σ_p by combining the spatial correlation coefficient $\rho_p(i,j)$ with a small nugget for numerical stability and the standard deviations of depths at the ith and jth nodes $\delta_k^d(i)$ and $\delta_k^{(d)}(j)$. We define

$$\rho_p(i,j) = \begin{cases} 1.0 & \text{if } i=j \\ 0.8 * \exp(-|x_{ki}-x_{kj}|/\lambda_k^d) & \text{Otherwise} \end{cases}, \tag{B-1}$$

and

$$\begin{cases} \delta_k^d(i) = (u_k(i)-l_k(i))/n_d \\ \delta_k^d(j) = (u_k(j)-l_k(j))/n_d \end{cases}, \tag{B-2}$$

where n_d is a tuning parameter. The covariance matrix Σ_p is given by $\Sigma_p(i,j) = \rho_p(i,j)\delta_k^d(i)\delta_k^d(j)$.

APPENDIX C

MCMC SAMPLING PROCEDURES

The main procedures are summarized as follows:

1) Assign initial values to all the unknowns and referred them to as $r_0^{(0)}$, $\mathbf{r}_k^{(0)}$ $(k=1,2,\cdots,m-1)$, and $\mathbf{d}_k^{(0)}$ $(k=2,3,\cdots,m)$. Set $t=1$.
2) Select an interface uniformly and select one of three cases (i.e., resistivity, depth, or both resistivity and depth) and one of two sampling methods (i.e., MMH or MSS) with a preset probability for updating. Refer those samples to as $\mathbf{r}_k^{(t)}$ $(k=1,2,\cdots,m-1)$, and $\mathbf{d}_k^{(t)}$ $(k=2,3,\cdots,m)$. We keep the same values for those untouched variables.
3) Update bedrock resistivity r_0 using single variable Metropolis-Hastings or slice sampling methods and refer the drawn sample to as $r_0^{(t)}$.
4) If the preset total number of iterations has been reached, stop; otherwise, let $t=t+1$ and go to Step 2.

For implementation, we always evaluate the logarithmic likelihood to avoid possible numerical over- or under-flow.

REFERENCES

Auken, E., and A. V. Christiansen, 2004, Layered and laterally constrained 2D inversion of resistivity data: Geophysics, **69**, 752–761, doi: 10.1190/1.1759461.

Besag, J., P. J. Green, D. M. Higdon, and K. Mengersen, 1995, Bayesian computation and stochastic systems: Statistical Science, **10**, 3–41, doi: 10.1214/ss/1177010123.

Bodin, T., and M. Sambridge, 2009, Seismic tomography with the reversible jump algorithm: Geophysical Journal International, **178**, 1411–1436, doi: 10.1111/gji.2009.178.issue-3.

Bosch, M., 1999, Lithologic tomography: From plural geophysical data to lithology estimation: Journal of Geophysical Research, **104**, 749–766, doi: 10.1029/1998JB900014.

Bosch, M., 2004, The optimization approach to lithological tomography: Combining seismic data and petrophysics for porosity prediction: Geophysics, **69**, 1272–1282, doi: 10.1190/1.1801944.

Brooks, S. P., and A. Gelman, 1998, General methods for monitoring convergence of iterative simulations: Journal of Computational and Graphical Statistics, **7**, 434–455.

Chen, J., and G. M. Hoversten, 2012, Joint inversion of marine seismic AVA and CSEM data using statistical rock-physics models and Markov random fields: Geophysics, **77**, no. 1, R65–R80, doi: 10.1190/geo2011-0219.1.

Chen, J., G. M. Hoversten, D. Vasco, Y. Rubin, and Z. Hou, 2007, A Bayesian model for gas saturation estimation using marine seismic AVA and CSEM data: Geophysics, **72**, no. 2, WA85–WA95, doi: 10.1190/1.2435082.

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

Chen, J., A. Kemna, and 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**, no. 6, F247–F259, doi: 10.1190/1.2976115.

Cowles, M. K., and B. P. Carlin, 1996, Markov chain Monte Carlo convergence diagnostics: A comparative review: Journal of the American Statistical Association, **91**, 883–904.

Cumming, W. B., 2009, Geothermal resource conceptual models using surface exploration data, Proceedings 34th Workshop on Geothermal Reservoir Engineering Stanford University, 6.

de Groot-Hedlin, C., and S. Constable, 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data: Geophysics, **55**, 1613–1624, doi: 10.1190/1.1442813.

de Groot-Hedlin, C., and S. Constable, 2004, Inversion of magnetotelluric data for 2D structure with sharp contrasts: Geophysics, **69**, 78–86, doi: 10.1190/1.1649377.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji, and G. Mavko, 2004, Stochastic reservoir characterization using pre-stack seismic data: Geophysics, **69**, 978, doi: 10.1190/1.1778241.

Gelman, A., and D. B. Rubin, 1992, Inference from iterative simulation using multiple sequences: Statistical Science, **7**, 457–472, doi: 10.1214/ss/1177011136.

Gilks, W., S. Richardson, and D. Spiegelhalter, 1996, Markov chain Monte Carlo in practice: Chapman & Hall/CRC.

Gunning, J., and M. E. Glinsky, 2004, Delivery: An open-source model-based Bayesian seismic inversion program: Computers & Geosciences, **30**, 619–636, doi: 10.1016/j.cageo.2003.10.013.

Hastings, W. K., 1970, Monte Carlo sampling methods using Markov chains and their applications: Biometrika, **57**, 97–109, doi: 10.1093/biomet/57.1.97.

Hoversten, G. M., H. F. Morrison, and S. C. Constable, 1998, Marine magnetotellurics for petroleum exploration, Part II: Numerical analysis of subsalt resolution: Geophysics, **63**, 826–840, doi: 10.1190/1.1444394.

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

Key, K., and J. Ovall, A parallel goal-oriented adaptive finite element method for 2.5D electromagnetic modeling: Geophysical Journal International, **186**, 137–154, doi: 10.1111/gji.2011.186.issue-1.

Key, K., and C. Weiss, 2006, Adaptive finite-element modeling using unstructured grids: The 2D magnetotelluric example: Geophysics, **71**, no. 6, G291–G299, doi: 10.1190/1.2348091.

Kumar, D., G. M. Hoversten, G. Nordquist, and W. Cumming, 2010, Role of 1D MT inversion in a 3D geothermal field: 80th Annual International Meeting, SEG, Expanded Abstracts, 1107–1111.

Li, Y., and K. Key, 2007, 2D marine controlled-source electromagnetic modeling: Part 1 — An adaptive finite element algorithm: Geophysics, **72**, no. 2, WA51–WA62, doi: 10.1190/1.2432262.

Loke, M. H., I. Acworth, and T. Dahlin, 2001, A comparison of smooth and blocky inversion methods in 2-D electrical imaging surveys: 15th Geophysical Conference, Australian Society of Exploration Geophysicists, Preview, **93**, 90–91.

Maliverno, A., 2002, Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem: Geophysical Journal International, **151**, 675–688, doi: 10.1046/j.1365-246X.2002.01847.x.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953, Equations of state calculations by fast computing machines: Journal of Chemical Physics, **21**, 1087–1092, doi: 10.1063/1.1699114.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431–12447, doi: 10.1029/94JB03097.

Neal, R. M., 2003, Slice sampling: Annals of Statistics, **31**, 705–767, doi: 10.1214/aos/1056562461.

Newman, G. A., and D. L. Alumbaugh, 2000, Three-dimensional magnetotelluric inversion using non-linear conjugate gradients: Geophysical Journal International, **140**, 410–424, doi: 10.1046/j.1365-246x.2000.00007.x.

Newman, G. A., E. Gasperikova, G. M. Hoversten, and P. E. Wannamaker, 2008, Three-dimensional magnetotellurics characterization of the Coso geothermal field: Geothermics, **37**, 369–399, doi: 10.1016/j.geothermics.2008.02.006.

Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, doi: 10.1046/j.1365-246X.1998.00498.x.

Rejeki, S., D. Rohrs, G. Nordquist, and A. Fitriyanto, 2010, Geologic conceptual model update of the Darajat geothermal field, Indonesia: Proceedings World Geothermal Conference.

Roy, I. G., 2002, A robust descent type algorithm for geophysical inversion through adaptive regularization: Applied Mathematical Modeling, **26**, 619–634, doi: 10.1016/S0307-904X(01)00072-5.

Sethian, J., and A. Popovici1999, 3D traveltime computation using the fast matching method: Geophysics, **64**, 516–523, doi: 10.1190/1.1444558.

Shewchuk, J. R., 1996, Triangle: Engineering a 2D quality mesh generator and Delaunay triangulator, *in* M. C. Lin, and D. Manocha, eds., Applied computational geometry: Towards geometric engineering, Vol. 1148 of lecture notes in computer science: Springer-Verlag, 203–222.

Smith, J. T., and J. R. Booker, 1991, Rapid inversion of two- and three-dimensional magnetotelluric data: Journal of Geophysical Research, **96**, 3905–3922, doi: 10.1029/90JB02416.

Smith, J. T., M. Hoversten, E. Gasperikova, and F. Morrison, 1999, Sharp boundary inversion of 2D magnetotelluric data: Geophysical Prospecting, **47**, 469–486, doi: 10.1046/j.1365-2478.1999.00145.x.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: The Society for Industrial and Applied Mathematics (SIAM).

Tierney, L., 1994, Markov chains for exploring posterior distribution: Annals of Statistics, **22**, 1701–1728, doi: 10.1214/aos/1176325750.

Trainor-Guitton, W., and G. M. Hoversten, 2012, Stochastic inversion for electromagnetic geophysics: practical challenges and improving convergence efficiency: Geophysics, **76**, no. 6, F373–F386, doi: 10.1190/GEO2010-0223.1.

Venables, W. N., and B. D. Ripley, 2002, Modern applied statistics with S: Springer.

Wannamaker, P. E., G. W. Hohmann, and S. H. Ward, 1984, Magnetotelluric responses of three-dimensional bodies in layered earths: Geophysics, **49**, 1517–1533, doi: 10.1190/1.1441777.