Stochastic inversion of 2D magnetotelluric data using sharp boundary parameterization

Jinsong Chen, Lawrence Berkeley National Laboratory, Michael G. Hoversten, Chevron Energy Company, Kerry Key, Scripps Institution of Oceanography, and Gregg Nordquist, Chevron Geothermal and Power Operations*

Summary

Stochastic approaches for inverting geophysical data have many advantages over deterministic inversion methods in terms of finding global solutions and quantifying their associated uncertainty. Since stochastic methods often need to run forward models more times than deterministic algorithms, their applications currently are limited to 1D inverse problems, where forward models can be run very fast. However, due to the rapid growth in computing power, especially parallel computing techniques, and the recent development of new numerical simulation methods, stochastic inversion of 2D geophysical data has become feasible. In this study, we explore the use of Markov chain Monte Carlo based Bayesian models for inverting 2D magnetotelluric (MT) data. To minimize the number of parameters, we adopt sharp boundary parameterization, in which we consider the locations and the resistivity of regions formed by the interfaces as unknowns. We use a parallel, adaptive finite-element algorithm to forward simulate frequency-domain MT responses of 2D conductivity structure. Synthetic case studies show the developed stochastic model is effective for estimating the interface locations and resistivity; most importantly, it provides detailed uncertainty information on each unknown parameter.

Introduction

Current approaches for geophysical inverse problems are mainly deterministic methods (e.g., Gauss-Newton, conjugate gradient, and steepest decent techniques). The solutions obtained using these methods often depend on the choice of initial values and thus are local rather than global. Additionally, deterministic methods provide very limited uncertainty information about solutions. Stochastic inversion methods have been recognized as a powerful approach for solving geophysical inverse problems; many successful applications can be found in the literature (e.g., Mosegaard and Tarantola, 1995; Miguel, 1999; Eide *et al.*, 2002; Tarantola, 2005). Stochastic methods have many 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 and EM data (Chen *et al.*, 2007). 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 (manuscript in preparation) recently developed a parallel algorithm for 2D EM that computes 2D MT and CSEM responses in a few seconds to tens of seconds when run on large cluster computers. New capabilities such as this now make it feasible to implement stochastic inversion methods for 2D MT and CSEM data sets.

In this study, we explore the use of stochastic inversion methods for inverting 2D MT data. We develop a Bayesian model based on sharp boundary parameterization and use the newly developed parallel, adaptive finite-element algorithm to forward simulate MT responses. We estimate the locations of sharp boundaries and resistivity of regions formed by the boundaries. The spatial dependence of unknown parameters is accounted by geostatistics model through prior distribution. We solve the inverse problem using Markov chain Monet Carlo (MCMC) sampling methods. We use a synthetic study to test the developed Bayesian model. This is an ongoing study and the model will be applied to a geothermal field in Indonesia later.

Method

<u>Sharp Boundary Parameterization and Bayesian Model</u>

We develop a Bayesian model using sharp boundary methods as used by Hoversten *et al.* (1998) and Smith *et al.* (1999). Figure 1 shows a schematic of the parameterization, where we assume that the domain size and the location of ground surface are predetermined. For a given inverse problem, this parameterization gives us a minimal number of unknown

© 2010 SEG

parameters. Let n be the total number of sharp boundaries (or interfaces) and m be the total number of inner estimation nodes. The number of inner nodes can vary from one interface to another interface. The unknown parameters associated with each estimation node are logarithmic thickness s_{ij} and resistivity r_{ij}, where $i = 1, 2, \cdots, n$, and $j = 1, 2, \cdots, m$. The data used for our inversion are apparent resistivity and phase defined by Vozoff (1991) for complex impedance Z_{xy} and Z_{yx}. For ease of description, we introduce four index numbers to show data types, among which 1 and 2 are referred to as apparent resistivity and phase of impedance Z_{xy} and 3 and 4 are referred to as apparent resistivity and phase of impedance Z_{yx}. We let d_{ijk} be the *k-th* type of MT data for the *i-th* frequency at the *j-th* site, where $k = 1, 2, 3, 4$, $i = 1, 2, \cdots, n_f$, and $j = 1, 2, \cdots, n_s$. Variables n_f and n_s represent the total number of frequencies and sites, respectively.

Let vector $\mathbf{d}$ represent all the MT data arranged by frequencies, sites, and data types. Let vector $\mathbf{s}$ represent all the unknown thickness arranged by estimation nodes and interfaces. Let vector $\mathbf{r}$ represent all the unknown resistivity arranged first by nodes and then by interfaces. Thus, we have the following Bayesian model:

$$f(\mathbf{s},\mathbf{r}\,|\,\mathbf{d}) \propto f(\mathbf{d}\,|\,\mathbf{s},\mathbf{r})\, f(\mathbf{s},\mathbf{r}). \qquad (1)$$

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 the likelihood function of 2D MT data. The second term is the prior probability distribution of thickness and resistivity as random fields.

Numerical Forward Model and Likelihood Function

We used a new, adaptive finite-element algorithm developed by Key and Ovall (manuscript in preparation) to simulate frequency-domain MT responses of 2D electrical resistivity structure. The model domain is discretized using unstructured triangular element grids that can easily accommodate arbitrarily complex structures. An automated, adaptive grid refinement algorithm is implemented in which the finite-element solution is computed iteratively on successively refined grids. Grid refinement is guided by an a-posteriori error estimator based on a recently developed dual-weighted residual operator. This algorithm builds off of earlier adaptive finite element works by Key and Weiss (2006) and Li and Key (2007), but uses a residual based error estimator that is computed using a hierarchical basis, rather than the gradient recovery error estimator used in the early works. For the heterogeneous conductivity models typically found by inversions, the residual based error estimator has been found to be more robust than the gradient recovery method. Furthermore, the new algorithm has been parallelized so that it can computes accurate, adaptively refined solutions very quickly. For small clusters, it can compute MT responses for models of the size described here in only a few tens of seconds, while on larger clusters (>50 nodes) it can compute responses in a few seconds. This makes stochastic inversion of 2D MT data feasible.

The 2D MT data used for the inversion are apparent resistivity and phase recorded for each frequency at a number of sites along a profile. In this stage, we assume that errors in data are independent and have Gaussian distribution with zero mean and given standard deviation. Thus, we can obtain the following likelihood model:

$$f(\mathbf{d}\,|\,\mathbf{s},\mathbf{r}) = \prod_{i=1}^{n_f}\prod_{j=1}^{n_s}\prod_{k=1}^{4}\frac{1}{\sqrt{2\pi}\sigma_{ijk}}\left(\frac{d_{ijk}^{obs}-d_{ijk}^{calc}}{\sigma_{ijk}}\right)^2, \qquad (2)$$

where σ_{ijk} is the standard error, and d_{ijk}^{obs} and d_{ijk}^{calc} are observed and calculated MT responses of the *k-th* type for the *i-th* frequency at the *j-th* site. The independency and normality assumptions are not critical and can be modified by assuming those errors are correlated and have other types of probability distributions.

Spatial Correlation and Geostatistical Prior Model

The Bayesian framework allows us to explicitly incorporate information from other sources into the inversion through the use of prior models. In this study, we assume that thickness and resistivity vectors are independent of each other. For the sharp boundary parameterization, we also assume that thickness and resistivity on different interfaces are independent. This assumption permits sharp changes in thickness and resistivity across interfaces. However, on each interface, we assume that both thickness and resistivity are spatially correlated and the spatial structure can be characterized using a geostatistical model with the exponential variogram. Thus, we have the following prior model for thickness:

$$f(\mathbf{s}) \propto \prod_{j=1}^{n-1}\exp\left\{-(\mathbf{s}_j-\mathbf{m}_j)^T\mathbf{C}_j^{-1}(\mathbf{s}_j-\mathbf{m}_j)\right\}, \qquad (3)$$

where $\mathbf{s}_j$ and $\mathbf{m}_j$ are thickness and prior mean thickness vectors on interface *j*, and $\mathbf{C}_j^{-1}$ is the inverse covariance matrix of the thickness vector on interface *j*. The prior mean vector is calculated from prior ranges of each parameter and

© 2010 SEG

the covariance matrix is calculated from both prior ranges and variogram model with given correlation length by using the methods given by Chen *et al.* (2009). We use the same method to define the prior model of resistivity vector.

Case Study

Synthetic True Model

We first apply the Bayesian model to a synthetic model that closely simulates field conditions. Figure 2 shows the true 2D resistivity cross-section. There are four interfaces, which form four layers with the last layer being bedrock. The thickness of the first layer increases from 200 m at the left edge to 800 m at the right edge, and its resistivity varies between 10Ωm (left edge) and 200 Ωm (middle) . The thickness of the second layer is constant and equal to 200 m, but its resistivity linearly increases from 1 Ωm at the left edge to 3Ωm at the right edge. The thickness of the third layer increases from 2600 m at the edge to 3300 m in the middle, and its resistivity also increases from 20Ωm at the edges to 500 Ωm in the middle. The resistivity of bedrock is assumed to be constant and equal to 50Ωm. The lateral range (or x -coordinate) is from -10 km to 10 km. We pick up 21 nodes for each interface, starting from -10 km and increasing with increment of 1 km. To avoid boundary effects, we choose very large domain sizes for forward simulation, which are from -350 km to 350 km along both the X-coordinate and Z-coordinate.

Synthetic Data and Prior Bounds

We generated noisy synthetic data using the adaptive finite-element numerical forward model by following the processes below: (1) Calculate the true complex impedances for the given true 2D resistivity structure as shown on Figure 2; (2) Add zero-mean Gaussian random noise whose standard deviation is equal to 5-percent of their corresponding amplitude; (3) Convert 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 thickness are the values that are 40-percent smaller and larger than their corresponding true values. The prior lower and upper bounds of resistivity from the first to the last layer are (1, 250) Ωm, (0.5, 100) Ωm, (10, 800) Ωm, and (10, 100) Ωm, respectively.

Inversion of 2D Synthetic MT Data

Figures 3 shows the estimated 2D resistivity structure using the medians of each unknown parameter. Compared to the true resistivity cross-section given in Figure 2, we can see that the estimated results are very close to the true model, but the quality of estimation varies. For the first layer, both estimated resistivity and thickness are very close to the true values. For the second layer, the estimated results are also consistent with the corresponding true values. For the third layer, however, both estimated resistivity and thickness fluctuate around the true values. Table 1 summarizes quantitative comparison between the estimated and true values and their associated uncertainty information. Based on the relative root-mean-squares (RMS) of the estimated and true values and the relative half-width (RHW) of 95% predictive intervals. We can see again that the inversion provides the most reliable estimates of resistivity and thickness at the first layer and the worst estimates at the third layer, especially for resistivity, which is worse than that at the last layer (bedrock).

Conclusions

The results of the synthetic study show the developed Bayesian model is effective for inverting MT data for 2D resistivity cross-section, given proper information on large structure and prior ranges of unknown parameters. The stochastic method can provide not only the estimates of each unknown parameter but also sufficient information on uncertainty.

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.

© 2010 SEG

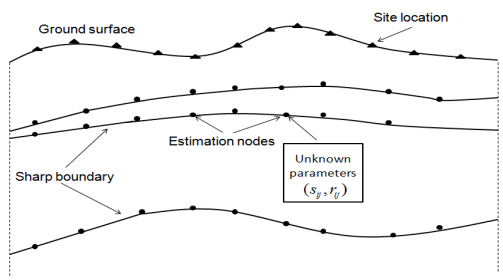

Figure 1: Schematic of sharp boundary parameterization

Table 1: Comparison between the true and estimated medians and uncertainty information

Parameters		RMS	RHW(%)
Thickness	Layer-1	0.1253	1.59
	Layer-2	0.2122	0.39
	Layer-3	0.2487	17.33
Resistivity	Layer-1	0.1549	7.18
	Layer-2	0.3452	7.83
	Layer-3	1.3479	33.29
	Layer-4	0.5326	14.40

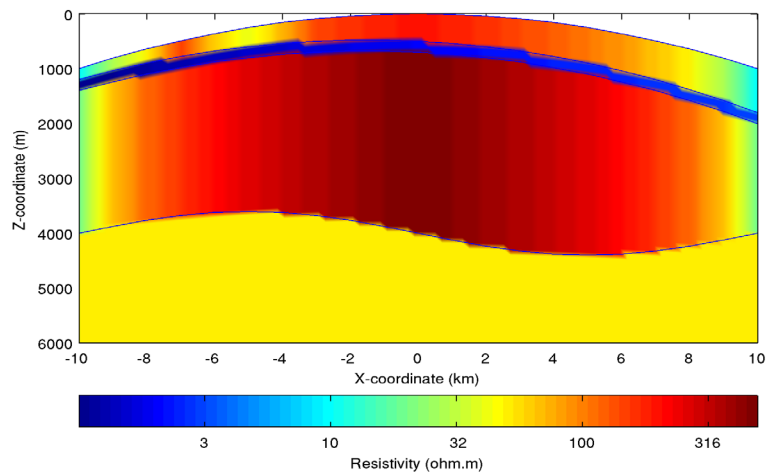

Figure 2: True synthetic 2D resistivity structure

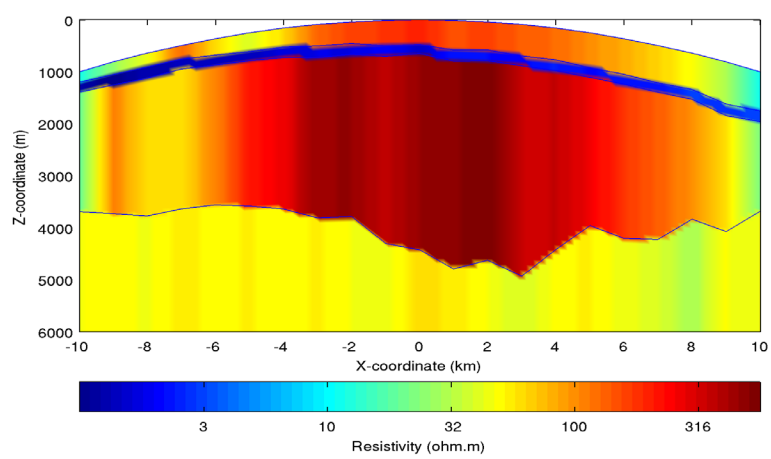

Figure 3: Estimated synthetic 2D resistivity structure

© 2010 SEG

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2010 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Bosch, M., 1999, Lithologic tomography: From plural geophysical data to lithology estimation: Journal of Geophysical Research, **104**, B1, 749–766, doi:10.1029/1998JB900014.

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 Resources Research, **45**, no. 8, W08420, doi:10.1029/2008WR007698.

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, doi:10.1190/1.2976115.

Chen, J., 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, doi:10.1190/1.2435082.

Eide, A. L., H. Omre, and B. Ursin, 2002, Prediction of reservoir variables based on seismic data and well observations: Journal of the American Statistical Association, **97**, no. 457.

Gunning, J., and M. E. Glinsky, 2004, Delivery: An open-source model-based Bayesian seismic inversion program: Computer and Geosciences, 30, no. 6.

Hoversten, G. M., H. F. Morrison, and S. Constable, 1998, Marine Magnetotellurics for Petroleum Exploration Part 2: Numerical analysis of subsalt resolution: Geophysics, **63**, 826–840, doi:10.1190/1.1444394.

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.

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.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, B7, 12431, doi:10.1029/94JB03097.

Smith, J. T., G. M. Hoversten, E. Gasperikova, and H. F. Morrison, 1999, Sharp boundary inversion of 2D magnetotelluric data: Geophysical Prospecting, **47**, no. 4, 469–486, doi:10.1046/j.1365-2478.1999.00145.x.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: SIAM, Philadelphia.

Vozoff, K., 1991, The magnetotelluric method, *in* M. N. Nabighian, ed., Electromagnetic Methods, in Applied Geophysics, 2: SEG, 641-707.