Stochastic Inversion of Electrical Resistivity Changes Using a Markov Chain, Monte Carlo Approach

Abelardo Ramirez, John Nitao, William Hanley, Roger Aines, Ronald Glaser, Sailes Sengupta, Kathy Dyer, Tracy Hickling, William Daily

September 27, 2004

Disclaimer

Stochastic Inversion of Electrical Resistivity Changes Using a Markov Chain, Monte Carlo Approach

A.L. Ramirez, J. J. Nitao, W. G. Hanley, R. Aines, R. E. Glaser, S. K. Sengupta, K. M. Dyer, T. L. Hickling, W. D. Daily

Lawrence Livermore National Laboratory, Livermore CA, 94550
September, 2004

Abstract

We describe a stochastic inversion method for mapping subsurface regions where the electrical resistivity is changing. The technique combines prior information, electrical resistance data and forward models to produce subsurface resistivity models that are most consistent with all available data. Bayesian inference and a Metropolis simulation algorithm form the basis for this approach. Attractive features include its ability to: 1) provide quantitative measures of the uncertainty of a generated estimate and, 2) allow alternative model estimates to be identified, compared and ranked. Methods that monitor convergence and summarize important trends of the posterior distribution are introduced. Results from a physical model test and a field experiment were used to assess performance. The stochastic inversions presented provide useful estimates of the most probable location, shape, and volume of the changing region, and the most likely resistivity change. The proposed method is computationally expensive, requiring the use of extensive computational resources to make its application practical.

Introduction

A fundamental earth sciences problem is to determine the properties of an object that we cannot directly observe. A variety of geophysical tomography techniques have been developed to provide detailed subsurface information. One such technique, electrical resistance tomography (ERT), is a relatively recent geophysical imaging scheme that provides 2-D and 3-D images of resistivity that are consistent with measurements made on an array of electrodes. With the increasing availability of computer controlled multi-electrode instruments and robust data inversion tools, ERT is becoming widely available. The value of ERT for monitoring dynamic subsurface processes has promoted new applications in a wide range of environments (e.g. Park and Van, 1991, Daily et al., 1992, Ellis and Oldenburg, 1994, Sasaki, 1994, LaBrecque et al, 1996, 1999, Binley et al., 1996, Morelli and LaBrecque, 1996, Park, 1998, Slater et al., 2000).

The goal of any ERT inversion method is to calculate the subsurface distribution of electrical resistivity from a large number of resistance measurements made from electrodes. A deterministic inversion procedure searches for a model (i.e., a spatially varying distribution of resistivity) that gives an acceptable fit to the data and satisfies any other prescribed constraints. A common solution minimizes an objective function consisting of a regularized, weighted least squares formulation. Typically, the search is

conducted using iterative, gradient-based methods (e.g., Park and Van, 1991, Ellis and Oldenberg, 1994, LaBrecque et al., 1996).

The ERT inversion problem is typically complicated by a non-linear relationship between data and the inverted parameters, state-space dimensionality, under/over determined systems, noisy and dependent data, etc. Hence, an exact inversion is rarely possible. It is common to use unrealistic simplifying assumptions to mitigate the severity of these problems when using classical optimization algorithms.

Previous Work— Stochastic Methods:
One alternative to the classical ERT inverse methodologies uses stochastic techniques that search for electrical resistivity models that best fit the collected electrical resistance measurements. The literature describes a variety of these methods and their application to geophysical problems. Zhang *et al.* (1995) suggest an inversion method that seeks to maximize a specified *a posteriori* probability density function of model parameters. In this case, maximizing the *a posteriori* density function is equivalent to minimizing the objective function in the classical inverse approach. Yang and Labrecque (1999) proposed an alternate solution that extends Zhang's approach by allowing a more efficient estimate of the parameter covariance matrix.

Mosegaard and Tarantola (1995) and Mosegaard and Sambridge (2002), describe an alternative stochastic inversion approach. The technique utilizes Markov Chain Monte Carlo (MCMC) methods, a class of importance sampling techniques that search for models that are most consistent with available data. In this approach, the inverse problem is formulated as a Bayesian inference problem. An importance sampling search algorithm is used to generate an empirical estimate of the *a posteriori* probability distribution based on available observations. Specifically, solutions are sampled at rates proportional to their posterior probabilities. This implies that models consistent with *a priori* information as well as observed data are sampled most often, while models that are incompatible with either prior information and/or observations are rarely sampled. This is the key difference between traditional Monte Carlo and MCMC: the former samples the space of possible models completely at random, while the latter moves through the models according to their posterior probabilities. This method yields an efficient sampling scheme that affords the user the flexibility to employ complex *a priori* information, and data with non-gaussian noise. Mosegaard and Tarantola show that this approach can be used to jointly invert disparate data types such as seismic and gravity data.

Kaipio et al. (2000) describe the application of the MCMC approach to medical imaging problems using electrical impedance tomography (EIT, for the purpose of this paper, ERT and EIT are synonymous). They considered a variety of non-differentiable priors including a minimum total variation prior, a second order smoothness prior and an "impulse" prior that penalizes the L^1 norm of the resistivity. Their approach estimates the posterior distribution of the unknown impedances conditioned on measurement data. From the posterior density, various estimates of the resistivity distribution and associated uncertainties are calculated.

Andersen et al. (2001a) describe an MCMC geophysical approach for the Bayesian inversion of electrical resistivity data. They used random, polygonal models to represent the layered composition of the earth, and demonstrate the performance of the method using field data. They analyze the posterior distribution by looking for the resistivity model that is most consistent with the data, and comparing it to the estimated posterior mean model. They also estimate the variability of the transitions between earth materials by comparing the standard deviation for each image pixel to its corresponding mean. Andersen et al. (2001b) describe another MCMC application aimed at the detection of cracks in electrically conductive media. Their approach assumes that the cracks are linear, non-intersecting and perfectly insulating. Using synthetic data, they demonstrate an updating scheme that assumes that the number of cracks is known *a priori*. Their approach estimates the posterior distribution of crack configurations and their associated variances.

Yeh et al., 2002, describe a sequential, geostatistical ERT approach that allows inclusion of prior knowledge of general geological structures through the use of spatial covariance. Their method also uses point electrical resistivity measurements (well logs) to further constrain the solution. They use the successive linear estimator approach to find an optimal model that consists of the "conditional effective electrical conductivity" (CEEC). They define CEEC as the parameter field that agrees with electrical resistivity measurements at core sample locations and that honors electrical potential measurements. They also compute conditional variances to estimate the uncertainty associated with their optimal CEEC model.

Our approach:

Here we describe a type of MCMC algorithm that incorporates resistance measurements, numerical forward simulators of subsurface electrical resistivity, and *a priori* knowledge to provide distributions of resistivity change that are likely to be present in a subsurface environment. This methodology produces resistivity models that show those subsurface plume configurations and resistivity values that are most consistent with the available data and forward models.

As with all MCMC approaches, Bayesian inference is driven by an importance-sampling algorithm that forms the basis for our methodology. There are two major components to the approach (refer to the flow diagram in Figure 1):

1) A base representation specifying the rules that the proposed resistivity models (alos referred to as states) of the system must obey (step B, Figure 1). These rules are based upon a priori knowledge.

2) A Markov Chain Monte Carlo (MCMC) simulation algorithm that generates samples according to the unknown posterior distribution. It uses a randomized decision rule to accept or reject the proposed states according to their consistency with the observed data. (steps C - F, Figure 1).

The main advantage of this approach is that it automatically identifies alternative models that are consistent with all available data, and ranks them according to their posterior probabilities and associated confidences. In geophysical ERT applications, the inverse problem is substantially under-constrained and ill-posed. Thus, the search for a solution that is unique and possesses a high degree of confidence is generally impossible. Hence, it is wise to consider approaches that are capable of generating alternative modelsand ranking them.

Base Representation - Subsurface Plumes

The base representation algorithm randomly generates neighboring models utilizing rules that are based upon available prior information. For example, the range of permissible resistivities is estimated using the properties of the injected fluid, and formation properties such as porosity and water saturation. We also make use of prior observations that these plumes tend to consist of one or more semi-contiguous regions and relatively simple shape.

The representation of a model space X (i.e., the individual resistivity models and their neighboring states) is critical to the overall effectiveness of the MCMC method. In those cases where generating samples from the model space is time intensive (e.g., executing forward models such as those used for 3D ERT), the model representation must be designed so that the number of degrees of freedom in the problem is well constrained. For this reason, we have chosen a categorical simulation approach where each category is associated with a discrete resistivity value. For example, if the set of categories is $\{A, B, C\}$, then $A \rightarrow R_1, B \rightarrow R_2, C \rightarrow R_3$, where the R_i represent distinct resistivity values.

A detailed outline of the algorithm that generates the resistivity models is presented in Appendix 1; this algorithm corresponds to step B in Figure 1. Examples of models generated by this method are shown in Figure 2. The samples generated are called proposal states and constitute possible solutions to the inverse problem. The search for models that are consistent with the data is controlled using a Markov chain **Q** designed to have a "stationary" distribution equal to the prior distribution $\rho(x)$. A stationary distribution is one in which the samples generated are representative of the true distribution because they are unaffected by the starting point of the Markov chain(s) and have explored sufficiently the distributional structure. Hence, as the Markov chain is executed, samples from $\rho(x)$ are generated. This means that exact knowledge of the properties $\rho(x)$ is not required because the properties are reflected by the samples from $\rho(x)$.

The base representation we developed is specific to our problem domain. It only applies to resistivity models associated with liquid plumes having relatively simple shapes and contained within porous media such as those created by subsurface tank leaks or fluid injection (e.g., steam, CO2, or water floods). The algorithm assumes that the system of interest consists of a zone of changing electrical resistivity embedded within an otherwise homogeneous volume. The changing volume can be described with two or more contiguous sub-volumes consisting of rectangular parallelepipeds. Each sub-volume has a

single resistivity value, and overlaps or is near to another sub-volume. These sub-volumes can have varying size, shape and resistivity properties. This approach limits the ability of the inversion to place spurious artifacts where the data has no sensitivity, and limits the anomalies to simple shapes.

We are interested in mapping temporal changes in resistivity such as those created by a liquid penetrating the subsurface. Our approach uses a ratio of two impedance datasets in the inversion. In this method, a new data vector, $\mathbf{d}_r$, is formed from:

$$\mathbf{d}_r = \frac{\mathbf{d}_t}{\mathbf{d}_b}\mathbf{f}(\sigma_{\text{hom}}), \tag{1}$$

where $\mathbf{d}_b$ is the data vector used as the baseline state, $\mathbf{d}_t$ is the data vector at some time t and $f(\sigma_{\text{hom}})$ is the forward solution for an arbitrarily chosen homogenous conductivity.

Inversion of the new dataset $\mathbf{d}_r$ in the usual manner then results in an image that will reveal changes relative to the reference value σ_{hom}. This approach works reasonably well in situations where the contrast in background resistivity is small. Park (1998) used a similar approach except he used $f(\sigma_{\text{heter}})$, where σ_{heter} is the heterogeneous conductivity for the baseline case. His approach is more general and well suited for situations where the contrast in background resistivity is relatively large.

Markov Chain Monte Carlo - Theory and Methodology
Our approach is a derivative of the Metropolis algorithm (Metropolis et. al., 1953) as described by Mosegaard and Tarantola (1995). It uses a Markov chain process to control the sampling of the space X of possible models. Within this framework, the solution to an inverse problem is an estimate of the posterior probability distribution defined over X. Then, for any potential solution $x_0 \in X$, the method will provide an estimate of the probability and confidence that state x_0 is the true state of the underlying system.

This MCMC approach is similar to classical inversion with the random model generator replacing the deterministic updating scheme based on a gradient search. In both cases, an initial model is chosen and responses are calculated with a forward solver. The calculated responses are compared to observed data. Finally, an updated model is chosen and the process repeats. The two approaches differ in how the updated model is chosen and the final result of the process. Specifically, MCMC produces a probability distribution defined over X; while, deterministic methods produce a single or a collection of states from X that best explain the data.)

The inverse problem under consideration may be described as follows. Let D denote the data space, and suppose that there exists a mapping G such that:

$$\underline{d} = G(\underline{x}) \tag{2}$$

where $\underline{x} \in X$ is a parameter vector describing the state of the system of interest and $\underline{d} \in D$ is a vector of measurements taken on that system. The inverse problem occurs when a vector of data values is observed, say $\underline{d}_0$, and we want to determine the value of the parameter vector $\underline{x}_0$ that gave rise to $\underline{d}_0$.

The sampling process can be viewed as consisting of two separate components: prior knowledge and measurements $\underline{d}$. In the previous section we introduced the first component, the generation of samples that are consistent with the available a-prior knowledge. We now discuss the second component: a decision process that either accepts or rejects these a-priori samples according to their consistency with the measurements $\underline{d}$ (Figure 1, steps C, D, E, F). Specifically, for each visited state, forward simulators are used to predict values of measurable quantities such as electrical resistance. These predictions are then compared to corresponding measurements to determine the likelihood $L(\underline{x})$ that the given state $\underline{x} \in X$ produced the observed data. An accept/reject decision based upon this likelihood is used to modify the prior sampling process. The result is a new Markov chain, **R**, which samples the posterior distribution, $P(\tilde{x})$ These samples provide the basis for estimating the posterior distribution and any subsequent inference concerning the true unknown state of the system.

Formally, Bayes rule relates the prior and posterior distributions as follows:

$$P(\underline{x}|\underline{d}) = k\, L(\underline{x})\rho(\underline{x}) \tag{3}$$

The likelihood $L(\underline{x})$ is a measure of the degree of fit between the data predicted assuming the model $\underline{x}$ and the observed data, and k is a normalizing constant. For this study, we assumed a likelihood function of the form:

$$L(\underline{x}) = k\, \exp\left(-\frac{1}{n}\sum_{i=1}^{N} \frac{\left| d(\underline{x})_{pred,i} - d_{0,i}\right|^n}{\sigma_i^n} \right) \tag{4}$$

where N is the number of data points, $d(\underline{x})_{pred,i}$ is the predicted data for a given model $\underline{x}$, $d_{0,i}$ is the vector of observed measurements, σ_i is the estimated data uncertainty, and $n \geq 1$. For the results described below, we assumed that $n = 2$. Eq. 4 assumes that the estimated data errors are uncorrelated; ERT surveys typically use the same electrodes for multiple measurements thereby increasing the probability that the data errors are correlated.

The decision to accept or reject a proposed state is made on the basis of likelihood comparisons (steps E and F, Figure 1). Suppose that the current state of the Markov chain is $\underline{x}^{(T)}$ and that a move to an adjacent state $\underline{x}^{(T+1)}$ is proposed. If these transitions were always accepted, then the simulation would be sampling from the prior distribution $\rho(\underline{x})$, i.e., the observed $\underline{d}_0$ would not influence the search. Instead, suppose that the decision to accept the proposed transition is made as shown by steps D, E and F in Figure 1. Note

that when the likelihood of the proposed state $L(\underline{x}^{(T+1)})$ is equal to or larger than that of the current state $L(\underline{x}^{(T)})$, the proposed transition is always accepted. If $L(\underline{x}^{(T+1)}) < L(\underline{x}^{(T)})$ but the two values are close to each other, the probability of acceptance is still around 1.0. Even when $L(\underline{x}^{(T+1)}) << L(\underline{x}^{(T)})$, the acceptance probability is not zero. Thus, this randomized rule allows a transition to a less likely state such that the process will move out of a local extremum. Theoretically, it will never get trapped in a region of locally high likelihood as long as the likelihood of the proposed state is greater than 0.0. Then, the randomized acceptance rule (step E, Figure 1) guarantees that the probability of accepting this transition will always be greater than 0.0.

This is the Metropolis algorithm, the best known of the importance sampling algorithms. Metropolis et al. (1953) proved that the samples generated through this three-step process will have a limiting distribution that is proportional to the desired posterior distribution $P(\underline{x}|\underline{d})$ -- the probability of model $\underline{x}$ being the true state of nature given that $\underline{d}$ has been measured. As a result of the randomized rule in step E (Figure 1), the search tends to hover in regions of space X containing states that better fit the prior information and ERT measurements. Because of this, space X is traversed more efficiently than with traditional Monte Carlo techniques.

The information contained in the ERT data determines whether the posterior distribution is a better representation of reality than the prior. When the data is "informative", i.e., is sufficiently sensitive to the characteristics of the target, the posterior distribution will be a better representation of reality. However, when the data is "uninformative" due to lack of sensitivity, measurement error, etc., then the prior and posterior distributions will be very similar, thereby indicating that the ERT data did not help to discriminate between models sampled from the prior distribution.

One desirable quality of the MCMC approach is that knowledge of the posterior distribution allows the uncertainty in any generated estimate of the true unknown state to be quantified. This provides the basis for: i) the objective assessment of competing hypotheses when the available information isn't sufficient to definitively identify the true system state, and ii) the propagation of uncertainty in modeling results through to follow-on predictions. Sources of uncertainty such as measurement error, contradictory data, lack of sensitivity or resolution, incomplete surveys, and non-unique relationship between measurements and inverted parameters can be addressed explicitly via this approach. Moreover, problems with many secondary extrema, a non-unique inverse, and/or contradictory or sparse data are mitigated.

Deterministic inversion methods are also able to address many of the problems listed above. A common approach makes use of different starting models to investigate solution uncertainty. For the models that converged, one can observe common features and gain confidence of their actual existence. This approach can be tedious to implement, and to the best of the authors' knowledge, not commonly used for ERT inversions. The MCMC approach described here performs the analysis automatically thereby sampling solution space more completely and reducing the probability of getting trapped on local extrema.

There are a variety of issues that must be addressed during the implementation of the MCMC methodology. The most fundamental concern is that the Markov chain must be designed so that it has a limiting stationary distribution. For this to happen, the transition probabilities must be specified so that the process is aperiodic (state transitions are not cyclical, state sequences do not repeat) and irreducible (it is possible to move from any given state to any other).

Convergence Analysis
As the Markov chain generates samples, it is important to verify that these samples are statistically representative of the posterior distribution $P(x|d)$. The chain is designed to have a long-run (i.e., stationary) distribution equal to the posterior distribution. Hence, after a sufficiently long "warm-up" period the sampling process will have forgotten its starting point and visited all of the modes (i.e., locations in space X at which a relative or absolute value occurs in the frequency distribution) of $P(x|d)$. After this point, the frequency of visits to a state will constitute a statistically reliable estimate of its posterior probability. More precisely, this means is that once the chain has gone a sufficient number of steps, T_0, the distribution of the generated states, $x^{(T)}$, at any step $T \geq T_0$ is unchanged and equals the posterior distribution, $P(x|d)$. We call T_0 the "burn-in" period. Hence, the MCMC process begins at a particular state that is selected at random and after the burn-in period, the chain has essentially forgotten where it started. At this point, the samples $x^{(T_0)}, x^{(T_0+1)},...$ constitute a representative random sample from $P(x|d)$ and can legitimately be used to perform posterior inference (e.g., generate estimates of the true state and their corresponding confidence).

Since the burn-in period must occur before samples can be reliably collected and used, metrics have been developed to help identify T_0. Several of the most widely used and effective metrics were employed during this study (Glaser, 2003).

The general idea is to compare the variability within each individual chain to the variability between the separate chains by using covariance matrices. Convergence of these metrics implies that the statistical difference between the collections of samples produced by the different chains is decreasing and collectively they are beginning to exhibit the same statistical behavior. Combining this trend with the observation that their individual variations are becoming stable, provides strong evidence that all chains are sampling from the same distribution, namely the posterior. In other words, the burn-in period has been completed and all subsequent samples are legitimately representative of the unknown posterior distribution $P(x|d)$.

The Gelman- Rubin Diagnostic.
Gelman and Rubin (1992) describe an approach that uses multiple Markov chains to estimate the burn-in period length T_0. To generate an accurate estimate, the method must address the difficulty caused by the properties and structure of $P(x|d)$ being unknown.

Specifically, the posterior distribution may contain multiple modes, or likelihood peaks that the Markov chain must effectively visit in order to produce a statistically representative sample. But since the number of significant modes is unknown a-priori, we can never be certain that a *single* chain has explored all critical structure of $P(\underline{x}|\underline{d})$. This apparent impasse is addressed through the use of multiple independent chains with individual, well-dispersed starting points. Although these chains start at different states, they share a common, but unknown, limiting distribution, $P(\underline{x}|\underline{d})$. The Gelman-Rubin diagnostic effectively detects when the variability between the sample sets produced by the individual chains settles down to a value that is expected when the chains are all sampling from a common distribution (i.e., the long-run stationary distribution $P(\underline{x}|\underline{d})$). When this behavior is detected, it is likely that the burn-in process is complete.

In our ERT problem, we are interested in describing contiguous subregions of specified resistivity values. For convergence analysis, each subregion is summarized by the triple $z = (z_1, z_2, z_3)$, where z_1 is the area, z_2 is the horizontal coordinate of the centroid, and z_3 is the vertical coordinate. By considering the largest contiguous subregion for each of say nine possible resistivity values, the dimensionality of the parameter of interest $\underline{Z}$ becomes $p = 9*3 = 27$. A multivariate version of the Gelman-Rubin diagnostic was used to track the behavior of several functions of the p-dimensional parameter vectors associated with the individual parallel chains for a moving and expanding window of steps (called iterations). In this study, a window is taken as a range of steps that can be characterized by a single parameter n. For example, $n = 50$ refers to the window of length 50 iterations ranging from iteration 51 through iteration 100, and in general, the window of size n considers each chain within the iteration sequence $n+1$, $n+2$, …, $2n$.

The quantities being tracked are correlated with the variability of $\underline{Z}$. They include a p-dimensional matrix W_n which estimates the within chain covariances for the window n, a p-dimensional matrix B_n/n which estimates the between chain covariances for the window n and the corresponding pooled p-dimensional matrix

$$V_n = \left(\frac{n-1}{n}\right)W_n + \left(1+\frac{1}{m}\right)\frac{B_n}{n}, \qquad (6)$$

which estimates of the covariance matrix of the posterior distribution of $\underline{Z}$. In this last expression, m is the number of Markov chains that are running in parallel. As n increases, i.e. the window moves and expands, the influence of the starting points on the individual chains diminishes, and the following trends begin to emerge:

- The within chain variation, summarized by the scalar $\det W_n$, stabilizes. Typically, $\det W_n$ increases as new modes in space X are encountered by the chains. $\det W_n$ settles to a limiting value once all significant modes are sufficiently sampled.
- The pooled chain variation, summarized by the scalar $\det V$, stabilizes, a result of the combined effect of the difference between chains, characterized by B/n, becoming negligible and the within chain variation stabilizing.

- The matrices V and W become "close" to one another. It is the "closeness" of V and W that generally indicates when the burn-in period has been achieved

The remaining challenge is how to assess the closeness of V and W. Brooks and Gelman address this issue by introducing a scalar measure of the distance between V and W:

$$R^p = \frac{n-1}{n} + \left\lceil \frac{m+1}{m} \right\rceil \lambda_1, \quad (7)$$

where λ_1 is the largest eigenvalue of the matrix $W^{-1}B/n$. Then, as n increases, the distance between V and W diminishes, the eigenvalue λ_1 decreases to 0, and R^p approaches 1. The Gelman-Rubin diagnostic, then, monitors R^p, $\det V$, and $\det W$, as a function of the window parameter n. For sufficiently large n, say $n \geq T_0$, the three conditions, R^p close to 1, $\det W$ approximately constant, and $\det V$ approximately constant, are satisfied. The nearness of R^p to 1 suggests burn-in has occurred by step T_0; while, stabilization of the determinants provides evidence that samples within the window starting at iteration T_o+1 are an adequate characterization of the stationary posterior distribution.

Suppose there are m chains. If the Gelman-Rubin diagnostics suggest a burn-in period of length T_0, then a total of mT_0 samples are discarded, and $m(N - T_0)$ samples are available for analysis of properties of the posterior.

In addition to the well-known Gelman-Rubin methodology, a variety of other convergence diagnostic methods have been proposed, developed and tested. These include a test of normality based upon the Central Limit Theorem (Robert et al.,1999), an examination of parameter quartiles (Raftery and Lewis, 1992, a and b), methods based upon renewal theory (Robert, 1995), etc. In fact, the literature focused upon the assessment of MCMC convergence is quite extensive and rich. Nevertheless, at this point in time, no suite of convergence diagnostics is capable of monitoring and identifying convergence with absolute certainty.

Posterior Analysis
After convergence has been verified and pre-burn-in models discarded, the resistivity models in the posterior distribution $P(x|d)$ can be analyzed. Our goal is to distill the relevant information in these models so that that we can infer the likely properties of the "true" resistivity model under study. The topography of $P(x|d)$ contains multiple hills whose heights are proportional to the likelihoods for each of its member resistivity models (shown schematically in Figure 3); each point in model space represents one resistivity model. The model corresponding to the peak of each hill is commonly referred to as its mode. Multiple hills indicate that the solution to the inverse problem is non-unique, the typical case for ERT. The distribution is called multi-modal when multiple hills are present and uni-modal when only a single hill is present. The width of each hill indicates that there is uncertainty in the model located at the mode; this variability may be due to factors such as measurement sensitivity or measurement error.

This complex, multi-modal structure provides a challenge when characterizing the distribution and extracting insight about the resistivity models included in $P(x|d)$. We use a clustering approach to extract this insight (Sengupta and Ramirez, 2003). Clustering is a standard data-mining technique used to extract structure from a collection of sampled data points – in this case sampled resistivity models. It segregates the models sampled from $P(x|d)$ into groups of models that exhibit similar properties. In our specific example, a cluster is a group of resistivity models that show similar spatial distribution of resistivity and similar resistivity values. The likelihood modes in Figure 3 represent these model clusters. The clustering process is accomplished by measuring the distance (in model space), between a model and a cluster center. A cluster's center is the model space location that best represents the central tendencies of all cluster members.

When deciding whether a resistivity model should be considered a member of a particular cluster, we measure the distance (in model space) between the candidate resistivity model and the cluster center. This distance is a measure of the dissimilarity of the sample relative to the central tendencies of all the models that are already members of the cluster. A cluster's central tendencies are represented by voxel-wise distribution of resistivity values. That is, for each voxel, we calculate histograms that show how frequently each of the possible resistivity values appears in all models included in the cluster; these frequencies are normalized to lie between 0 and 1. The following example should help clarify this method.

Suppose that each resistivity model contains three voxels and that the set of possible resistivities is: {10, 15, 20, 30}. Suppose further that there are 100 models in a cluster with the voxel-wise frequency distributions shown in Table 1. The frequencies for these are calculated by dividing the number of models showing a particular resistivity value by the total number of models in the cluster. Table 1 suggests that, for voxel 1 there are 30 samples with a resistivity of 10 (frequency is 0.3 = 30/100), 40 with a resistivity of 15, 20 with a resistivity of 20 and 10 with a resistivity of 30.

Table 1: Cluster frequencies

	Voxel		
resistivity value	1	2	3
10	0.3	0.3	0.3
15	0.4	0.2	0.2
20	0.2	0.3	0.4
30	0.1	0.2	0.1

Table 2: Resistivity model frequencies

	Voxel		
resistivity value	1	2	3
10	0.0	0.0	0.0

15	0.0	1.0	0.0
20	1.0	0.0	0.0
30	0.0	0.0	1.0

We also need to calculate frequency histograms for the resistivity model in question so that they can be compared to the cluster's histograms. Suppose that the resistivity model being considered has resistivity values of (20, 15, 30). The resistivity frequencies for this model are shown in Table 2; i.e., when one of the possible resistivity values is present in a voxel, the frequency for that resistivity value is set to 1.0 and frequencies for all other possible values are set to 0.0; this is repeated for all voxels.

We need to compare the frequencies in Tables 1 and 2 in order to calculate the model-cluster 'dissimilarity' *MCD*. For every Table 2 element where the frequency = 1.0, we subtract the corresponding element in Table 1. Thus, *MCD* = [(1.0 - 0.2) + (1.0 - 0.2) + (1.0 - 0.1)]/3 = 0.83. Large *MCD* values arise when the model anomaly is located in a different part of the 3D model, has different resistivity values, or both. When MCD approaches 0.0, the resistivity model under evaluation shows a resistivity distribution that is very similar to that of most resistivity models in the cluster.

The equation for *MCD* can be written as:

$$MCD_{m,n} = \frac{1}{nv} \sum_{i=1}^{nv} \sum_{j=1}^{nr} | mrf_{i,j}(mrf_{i,j} - crf_{i,j}) | \qquad (8)$$

where *m* and *n* are cluster and model numbers respectively, *nv* is the number of voxels in one resistivity model, *nr* is the number of possible resistivity values, *mrf* is the model's resistivity frequency and *crf* is the cluster's resistivity frequency.

The *MCD* can also be used to locate the "center state" (*CS*) for a given cluster. The *CS* is that resistivity model showing the minimum *MCD*; it is also the model that is closest to the cluster's "center of mass" and should be the one that best represents the cluster members. The *CS* and the mode refer to the same model when the cluster members are distributed symmetrically about the mode.

Dynamic K-Means clustering
We now describe the clustering algorithm; Figure 4 illustrates the flow diagram for this algorithm. A variation of the K-Means (Jain and Dubes, 1988) algorithm is obtained when starting from an initial value of K=1 clusters, we allow K to grow dynamically as new resistivity models are considered for assignment to an existing cluster (Pao, 1989). A new cluster is formed when a sample due for assignment to a cluster is considered as being too 'far' (i.e., has too large an *MCD*) from any existing cluster. Clearly a new parameter *td*, the threshold distance has to be introduced at this stage. Then, when the minimum of all distances between the model and existing cluster centers exceeds *td*, we create a new cluster with the sample under consideration as its first member. The process is repeated iteratively until a specified stability criterion is satisfied.

Suppose that *K* is the number of clusters, *u* the maximum number of clusters allowed, and *td* be the threshold distance.

Stage 1 (Cluster growing) Initially, set $K = 1$ with cluster #1 containing the first model and the cluster center located at the first model. At any given stage, repeat steps 1 and 2 where,

1. Get a new model and compute its *MCD's* from the existing cluster centers and their minimum(*MCD*) and locate a cluster C say, where this minimum occurs.
2. If minimum(*MCD*) is greater than *td*, then create a new cluster with the new model as the only cluster member and itself as the new cluster center.

Otherwise, assign the new model to cluster C and update the cluster center. Continue this process until either the maximum number of clusters has been attained or all of the models are assigned to one of the generated clusters.

Stage 2 (Re-clustering) At the end of the first stage, let there be $K \leq u$ clusters generated. Repeat steps 3 and 4 until a stability criterion *S* (defined below) is satisfied.

3. Replace the existing partition with a new one by assigning a group to each member based on the nearest cluster principle (minimum *MCD*).
4. Re-compute the cluster centers by determining the center-state for the new collection of members in each cluster.

Tests of this algorithm show that Stage 2 insures that the clustering results are independent of the order in which the models are introduced. The stability criterion *S* can be introduced in various ways. In our algorithm, *S* is reached when the updated cluster centers all remain within a small pre-assigned distance from the respective cluster centers computed in the previous step.

Two clusters may be merged to form one whenever the maximum of the intra-cluster distances within each cluster is smaller than the inter-cluster distance between the two clusters. Depending on the inter- and intra-cluster distance distributions, one may want to merge two or more clusters in the final clustering step. Suppose that the 'diameter' of a cluster is defined as the maximum of the intra-cluster distances between models within a cluster. Let us also define the 'inter-cluster distance' between two clusters as follows. Consider all pairs of models that can be formed by choosing one model from each of the two clusters. The largest distance between models within all such pairs is then defined as the 'inter-cluster distance'. In order that two clusters may be merged to form into one, we require that the sum of their diameters be less than the inter-cluster distance between the two. Depending on the inter- and intra-cluster distance distributions, one may want to merge two or more clusters in the final clustering step. This merger condition can be extended to accommodate this possibility. Moreover, this final merging step does not require any distributional assumptions or the specification of additional parameters to be implemented. The new cluster center and the corresponding cluster frequency are computed and cluster memberships for the merged clusters re-assigned.

Once the partitioning of the models is complete, the voxel-wise average resistivity, VAR_i, for all the models in a cluster is used to provide a reasonable estimate of the trends exhibited by most models in that cluster. Consider a column vector $\underline{u}$ with nv components, the components containing the voxel resistivities in a 3D resistivity model. If there are N column vectors corresponding to all the resistivity models in a given cluster, then:

$$VAR_i = 10^{(\frac{1}{N}\sum_{n=1}^{N}\log_{10}(u_{i,n}))} \qquad (9)$$

This is repeated for all components (i = 1, 2, … , nv). For non-ERT applications, one may choose to use an arithmetic mean instead of the geometric mean indicated by eq. 9.

Results and Discussion:

Physical Model Results
We have used physical models where we know the exact properties of the target in order to evaluate the performance of the MCMC approach. The physical model consisted of various targets immersed in a tank filled with water. The model included 4 vertical electrode arrays, each having 15 electrodes (refer to Figure 5). The arrays were submerged in water and a variety of solid and porous (sand-lead mixture encased in a nylon mesh) targets having different electrical resistivities were inserted at various locations between the electrodes. The water resistivity was 16 ohm-m. We will present results from the sand-lead mixture that had a resistivity of about 40 ohm-m.

Issues such as the accuracy of the inverted location, shape and change magnitude of the inversions were evaluated. Uncertainty arises from the inherent errors (measurement and modeling) and the non-unique relationship between inverted parameters and measurements.

We first needed to determine if the inversion had converged by using the convergence diagnostics described earlier. The Gelman-Rubin diagnostic is a metric indicating when T_0 (burn-in period length, stable value near 1.0) has been reached asymptotically. For the physical model results discussed below, the burn-in period ended around iteration $T_0 =$ 700. Also, stable values near 1.0 suggested that samples after T_0 had succeeded in visiting all the modes in $P(\underline{x}\,|\,\underline{d}_o)$. This diagnostic suggested that the Markov chains converged to the limiting posterior distribution $P(\underline{x}\,|\,\underline{d}_o)$, and that we could reliably employ these posterior samples to estimate the parameter(s) of interest.

Figure 6 shows the stochastic inversion results for the sand-lead physical model. The left image is a vertical section of the actual target showing its shape and location. Each image consists of 16 x 16 x 33 voxels representing a 6.8E-02 m^3; the volume of the model is 2.5E-05 m^3. The model has a relatively low-contrast (resistivity only 2.5 times as high as

the surrounding water). For these inversions, we assumed that the target could be sufficiently described by two contiguous or overlapping parallelepipeds and that the set of possible resistivity ratio values was {0.5, 1.0, 2.0, 4.0, 8.0, 16.0}. A value of 1.0 indicates that the target's resistivity equals that of the surrounding water, and values above 1.0 indicate that the target's resistivity is larger than the surrounding water. The inverse process searched for the most likely location, size, shape and contrast of the changing region.

To calculate the inversion results shown, we used all post-burn-in models (all $x^{(T)}$ such that $T > 700$). The Figure 6 top right frame shows the *VAR* for all the posterior models assuming that all models came from a uni-modal distribution (single cluster); i.e., no attempt was made to segregate the resistivity models in the posterior distribution into groups of models having similar spatial distribution of resistivity and similar resistivity values. The vertical slice shown through each model was positioned to show the maximum resistivity value present. The voxel-wise average results show similar shape and location as the actual target. The resistivity ratio magnitude lies in the range of 2.0 – 2.9, while the actual target value is 2.5.

We next considered the possibility that posterior models came from a multi-modal distribution. Using the dynamic K-means technique described earlier, $u = 10$, and $td =$ 0.04 we obtained the results shown by the lower row of Figure 6. The analysis produced 10 clusters of models; the three most probable ones (showing the highest frequency) are displayed in the bottom row of Figure 6. The results shown are cluster centroids (*VAR* of all the models in a given cluster). The frequency values have been normalized relative to the number of posterior distribution samples (4302) to obtain relative frequency values for each cluster. The bottom left image is the centroid for the most frequent cluster: four out of every five samples come from this part of the state space. Note that the location, and size of the cluster centroid are similar to the target. The maximum value resistivity ratio magnitude is about 2.48 while the target value is 2.51. The middle and right bottom row images represent centroids for low frequency clusters. The middle image shows two closely spaced anomalies that straddle the target's elevation and a resistivity ratio value (16.0) that is substantially larger than the target. The bottom right image also shows an anomaly shape that is significantly different from the target and has a maximum resistivity ratio value (16.0) that is substantially larger than the target.

We propose that the results shown in Figure 6 illustrate the value of clustering analysis. When the cluster analysis accounts for several likelihood hills in the posterior, the resistivity ratio for the most frequent cluster is about 2.48, close to the target value of 2.51. When a uni-modal distribution is assumed, a poorer match to the target is observed because all samples have lumped into a single cluster. The clustering analysis segregates posterior samples with lower and higher values into separate clusters, thereby improving the accuracy of the resistivity value corresponding to the most frequent cluster. Similar comments apply to the size, location and shape of the cluster anomalies.

Field results - Tank Leak Detection
Electrical resistance data was collected during a field experiment that simulated leakage from a large metallic tank. For testing, an electrical tracer (saline solution) was used instead of the real contaminant to preserve the environmental quality of the test site. The test site used for this work is part of the 200 East Area in the Hanford Site, located near Richland, Washington (additional details about the test site and testing are provided by Barnett et al., 2003). The near surface sediments at the test site consist primarily of fine to coarse-grained sand with an average porosity of 0.25 to 0.30. The field experiments were performed under a 15.2 m diameter steel tank mockup. Figure 7 shows the layout at the experimental site. This empty steel tank contained several built-in spill points (only the one used is shown). The bottom of the tank is located 1.5 m below ground surface. Sixteen boreholes with eight electrodes in each surrounded the tank. The electrodes were spaced every 1.52 m between the surface and 10.7 m depth. The water table was located well below the deepest electrode location. The diametrical distance between boreholes was 20.7 m.

Hypersaline solution was released from a point near the tank's center over a 52 hour period. The liquid consisted of a sodium-thiosulfate solution (36 wt%) with a conductivity of about 5 S/m (about 1.5 times that of seawater). This hypersaline fluid has similar electrical conductivity and density to the real Hanford tank liquids. ERT surveys were made before, during and after the brine release using a dipole-dipole approach.

Tank leak results are shown in Figure 8 for the case where 2160 liters had been released. Dimensions of the 3D image block are 28 m (along each horizontal axis), and 13 m (height). Each model consists of 22 x 22 x 17 voxels. For these inversions we assumed that the target could be sufficiently described by six contiguous or overlapping parallelepipeds. Their shape, location and resistivity contrast were allowed to vary. The set of possible resistivity ratio values was {1.0, 0.95, 0.90, … , 0.01}, for a total of 20 possible values. The posterior samples were analyzed using the "dynamic k-means" algorithm described earlier with $td = 0.04$ and $u = 20$. The clustering was done using all post-burn-in models (all $x^{(T)}$ with $T > T_0$, a total of 3802 models; $T_0 \approx 810$). The left column of graphs displays the three most likely cluster centroids with frequencies ranging from 0.12 to 0.38. These three clusters encompass 68 % of all posterior distribution samples. All of them suggest a roughly vertical anomaly directly below the release point to depths ranging from 8 – 10 m.

The left column of frames in Figure 8 shows the voxel-wise mean resistivity ratio for the three most probable clusters. The frequency values shown have been normalized relative to the number of posterior distribution samples in each cluster. Clusters A and B suggest that a strong vertical anomaly exists just East of the release point. Cluster A suggests the possibility of liquid ponding to the West of the release point at a depth of about 8m. Likewise, cluster C suggests a wider but somewhat weaker anomaly. An analyst could consider cluster A as the most likely thereby inferring a pillar-shaped vertical invasion zone with possible liquid ponding to the West of the release point at 8m depth. Considering clusters A and B, the analyst may also consider the possibility that the bulk of the contamination is located below and to the East of the release point. These

alternative models offer the analyst alternative models with model A being the most likely, and model C the least likely.

This analysis can be used to evaluate the propagation of uncertainty due to measurement error, or due to lack of sensitivity or resolution. For example, the mean resistivity ratios (left column of images in Figure 8) for all three clusters show significant variability directly below the release point. The highly conducting metal tank walls create a region of diminished sensitivity below the tank that is most severe just below the release point. We suggest that this diminished sensitivity explain is responsible for the variability observed.

Figure 8 also shows the center states for the clusters shown. Within a given cluster, the differences between the voxel-wise means (left column of images) and the center states (right column of images), provide the analyst some measure of variability (uncertainty) within the cluster. These center states are those posterior samples having the smallest *MCD*. We can use the center state as a way to evaluate a cluster's central tendencies. These center states are unaffected by the variability of such properties within a cluster; clearly, this variability affects the mean values shown in the left column of frames in Figure 8. The center states for clusters A and B show similar location and shape as the cluster mean image. However, the mean for cluster A suggests the possibility of liquid ponding to the West of the release point at a depth of about 8 m whereas the cluster center state does not show such features. This means that the possible ponding of liquid is not part of the central tendency for cluster A, but is a part of some of the states around the cluster center. The center state for cluster B suggests a stronger resistivity change (0.03) than the values near 0.2 indicated by the cluster's mean. This approach allows an analyst to consider alternative models that are consistent with the data to varying degrees. The frequency information determines how well each of the alternative models represents the available data.

A drill-back program that would have independently mapped the plume's characteristics was not carried out due to lack of funds. Instead, we are forced to use circumstantial evidence to evaluate the results. First, we visually compare the stochastic inversion to a deterministic tomograph calculated by the inversion algorithm described by LaBrecque et al (1999). We have confidence that this algorithm produces reasonable results because it has been used successfully in field applications (e.g., LaBrecque et al, 1999, Morelli and LaBrecque, 1996, Daily and Ramirez, 2000) as well as in controlled physical model experiments where the target characteristics are known exactly (Ramirez et al, 2003). Figure 9 shows the comparison of the two results. The stochastic result (left column of frames) shows the voxel-wise mean for the most likely cluster (cluster A, Figure 8 upper left hand image)The bottom row of images shows iso-surface views, where the white bar over the color bar indicates the range of transparent values used to render the iso-surfaces.

Figure 9 shows similarities between the classical inversion and MCMC results thereby suggesting that the stochastic results are reasonable. The location of the changing region is approximately the same. Also, both methods show that the zone of change extends

more to the East of the release point than to the West. There are also differences between the images. The stochastic image shows stronger resistivity changes reaching values of 0.1 (i.e., the invaded soil has 1/10 the resistivity of the pre-release value) and a smaller volume. The stochastic image also suggests that brine may be pooling on the western side of the image. In the absence of independent ground-truth data we cannot establish which of the two images is closer to reality.

We can compare the two results on the basis of how well either fits the resistivity measurements. We calculated the forward solution for the deterministic and stochastic inversions shown in Figure 9. Then, we calculated the root mean squared differences (RMS) between the forward solutions and the measurements. The RMS for the deterministic result is $1.6*10^{-7}$, substantially better than the MCMC result ($1.7*10^{-5}$). We speculate that the deterministic result fits the data better because the deterministic inversion allows a much wider range of resistivity values and anomaly shapes.
We can also compare the volume and average resistivities of the inversion anomalies to invaded zone estimates based on Archie's equation and independent observations of released brine resistivity (0.11ohm-m) and volume (2160 liters). We assume that average porosity is about 0.25, the resistivity of the pre-release pore water is about 5 ohm-m, and that the brine release increases the saturation from 0.4 to 0.8. We also assume that the brine displaces part of the original pore water so that pore water resistivity after release has decreased to 2.5 ohm-m. Given these assumptions, brine would change the bulk resistivity would from 316 ohm-m to about 40 ohm (resistivity ratio of 0.13) within a 10.8 m^3 volume of soil. For the deterministic inversion anomaly, the geometric-mean resistivity-ratio is 0.96 within an anomaly volume of 5511 m^3. For the MCMC inversion, the geometric-mean resistivity-ratio is 0.21 within an anomaly volume of 1387 m^3. If instead of the cluster mean, we choose the center state for cluster 4 (upper right image in Figure 8), then the geometric-mean resistivity-ratio is 0.19 within an anomaly volume of 391 m^3, the closest to the estimate

Clearly, both methods produce anomaly volumes that are grossly exaggerated probably due to the intrinsically low resolution of the ERT measurements and to decreased measurement sensitivity caused by shunting of electrical current through the tank's metal walls. The comparison suggests that the MCMC inversion produced an anomaly that is more compact and exhibits a larger resistivity change than the deterministic result. These differences are at least partly due to the regularization used by the deterministic algorithm that penalizes models with large "roughness" (the inverse of smoothness) (Park and Van, 1991, Shima, 1992, Ellis and Oldenburg, 1994, Sasaki, 1994, LaBrecque et al, 1996, Morelli and LaBrecque, 1996). It is well known that this approach tends to produce models that have reduced contrast and exaggerated extent.

In this work we have compared an MCMC method that is constrained by size, location, and resistivity limits with a deterministic method that is constrained by model roughness. The ideal comparison between the deterministic and stochastic methods would compare the results using the same set of constraints. We suggest that such a comparison be conducted as part of future research.

Figure 10 shows further evidence that the stochastic results are reasonable. The figure shows a sequence of MCMC inversions as the volume of released fluid increases from 340 to 2160 liters. Along a given column, released volume increases from the top row to the bottom row of images. The right column of images shows an iso-surface view where resistivity ratios from 0.9 to 1.0 are transparent. As expected, the voxel-wise average ratio decreases (i.e., the soil becomes more electrically conducting) from about 0.95 (top middle row, released volume = 0.34 m^3) to about 0.2 and below (bottom middle row, released volume = 2.16 m^3). Also, the vertical and horizontal extent of the anomaly grows with increasing released volume. The sequence shown suggests that the plume represented by these inversions is behaving as expected: anomaly volume grows and the resistivity ratio decreases as released brine volume increases.

Lastly, we suggest that the stochastic results in Figures 8 to 10 are consistent with plume characteristics determined independently by Ward and Gee (2001), and Gee and Ward (2001). Their field tests were conducted at a site located a few hundred yards from the Mock Tank leak test facility where our results were obtained; both sites are known to have similar geology. Their field tests, laboratory tests, and numerical simulations of the movement of hypersaline solutions through the vadose zone suggest that these plumes move along finger-like, vertical preferential flow paths due to the much higher density of the hypersaline solution. Their results also show that the plumes tend to be more compact, move deeper and show less lateral spreading than low ionic strength solutions. The anomalies shown in Figures 8 to 10 show similar characteristics to those observed by Ward and Gee, and Gee and Ward.

In summary, the similarity between the stochastic and deterministic results, the increasing stochastic anomaly volume as the released brine volume increases, and the consistency between stochastic anomaly characteristics and independent a priori expectations of the plume characteristics, leads us to believe that the MCMC results provide reasonable representations of reality.

Computational expense – parallel computing
The MCMC method we have used is computationally expensive. For example, a tank leak problem involving 28800 voxels and 128 transmitting electrodes and a typical work station with one CPU, will require about 45 days to accumulate a sufficient number of posterior samples (about 4000). Almost all of that time is used to solve the forward problem. In comparison, a deterministic inversion of the same problem using the same workstation may take 8 – 12 hours.

Clearly, parallel computation of the forward problem is required to make the MCMC approach practical. We have parallelized the problem in two ways: 1) Individual Markov chains are run on a separate processors. 2) The computational load for each chain is further distributed amongst multiple processors by computing the potential field due to each transmitter electrode on separate processors. The potential fields are then post-processed to compute the transfer resistances associated with individual readings. When 128 processors are used, this approach reduces the processing time to about 12 hours.

Processing times of 12 hours are acceptable for many real-life applications of the MCMC approach. Clearly, approach 2 could also be used to accelerate a deterministic run.

Step size (i.e. distance between neighboring models in space X) controls the characteristic changes one allows in model space. Step size has a large impact on computational expense; if the resolution is too fine or involves a high dimensional state vector, the convergence may be slowed beyond practical limits. If too small, movement through the state space will be slow, and it will take longer to move past local extrema. When the step size is too large, increased rejection ratios (number of states rejected/total number of states evaluated) are likely thereby slowing the convergence rate. By trial and error, we discovered that a randomized step size that sometimes took smaller or larger steps provided a reasonable solution to this dilemma.

Closely related to the step size is the topography of the likelihood surface under analysis. When the surface is exceptionally steep, the process will be slowed due to high rejection rates because the process attempts to move off a steep peak in likelihood space and ends up proposing states having much smaller relative likelihoods that are almost always rejected. This slow mixing process (i.e., how efficiently the process moves through space X) can be mitigated in several ways but that discussion is beyond the scope of this manuscript. Finally, the choice of the prior distribution may also significantly impact convergence because it controls the proposal of candidate states. In general, the closer the prior is to the posterior, the faster the process converges to $P(x|d_0)$.

Summary and conclusions:

We have discussed a stochastic methodology for the inversion of changing subsurface electrical resistivity data. This method is based upon Bayesian inference and is implemented via an MCMC algorithm. The inversion of electrical resistivity data is an ill-posed problem requiring regularization. Our approach makes use of prior information to sufficiently reduce the size of the space of feasible solutions in order to mitigate ill-posedness. The resistivity models consist of multiple blocky regions of resistivity change embedded within an unchanging volume. Additional information can include the sense of the change (increasing or decreasing resistivity), upper/lower bounds for the volume of the changing region, resistivity change magnitude and spatial relationships of the regions (e.g., requiring partially overlapping or contiguous blocks).

A key strength of MCMC is that solutions are sampled at a rate proportional to their consistency with available data. Hence, models that are most consistent with available data observations are sampled most often, while models that are incompatible with either prior information and/or observations are rarely sampled. As a result, the frequency of models in the posterior distribution can be used to determine the probability that a given model is the best explanation for the available data. The approach can be used to identify competing models when the available information isn't sufficient to definitively identify a single optimal model. Another strength is that it can be used to jointly invert disparate data types such as seismic and gravity as shown by Moosegard and Tarantola (1995).

We view the MCMC and the deterministic inversion methodologies as complementary approaches. The MCMC approach is similar to classical inversion with the random model generator replacing the deterministic updating scheme based on a gradient search. In both cases, an initial model is chosen and responses are calculated with forward solver. The calculated responses are compared to observed data. Finally, an updated model is chosen and the process repeats. The two approaches differ in how the updated model is chosen. The deterministic method is likely to be the preferred method when fast inversion times are required, and when then the regularization scheme produces sufficiently accurate models. The MCMC method may be most useful when inverting problems with many secondary extrema, when explicit estimates of solution uncertainty are required, and when alternative models are desired, ranked according to their consistency with available data.

Acknowledgments

References:

Aines, R., J. Nitao, R. Newmark, S. Carle, A. Ramirez, D. Harris, J. Johnson, V. Johnson, D. Ermack, G. Sugiyama, W. Hanley, S. Sengupta, W. Daily, R. Glaser, K. Dyer, Y. Zhang, Z. Yu, R. Levine, 2002,. *The Stochastic Engine Initiative: Improving Prediction of Behavior in Geologic Environments We Cannot Directly Observe*, UCRL-ID-148221, Lawrence Livermore National Laboratory, Livermore, CA

Barnett, B., G. Gee, and M. Sweeney, 2002,. Results of tank teak detection demonstration using geophysical techniques at the Mock Tank test site, PNNL-13818, Pacific Northwest National Laboratory, Richland, WA

Binley, A., S. Henry-Poulter and B. Shaw, 1996 Examination of solute transport in an undisturbed soil column using electrical resistance tomography, *Water Resour. Res.*, **32**(4), 763-769.

Binley, A., P. Winship, M. Pokar and J. West, 2001, Cross-borehole radar and resistivity tomography: A comparison of techniques in unsaturated sandstone, In: Proc. Symp. Applications of Geophysics to Engineering and Environmental Problems (SAGEEP2001), Environmental and Engineering Geophysical Society, Denver, CO.

Brooks, S.P., and Gelman, A., 1998, General Methods for Monitoring Convergence of Iterative Simulations, Journal of Computational and Graphical Statistics, vol. 7, no. 4, pp. 434 – 455.

Andersen, K.E., Brooks, S.P. and Hansen, M.B. 2001a, Bayesian inversion of geoelectrical resistivity data. Research Report R-01-2016. Department of Mathematical Sciences, Aalborg University, Denmark.

Andersen, K., S. Brooks, and M. Hansen, 2001b, A Bayesian approach to crack detection in electrically conducting media, *Inverse Problems,* vol. 17, p121-136.

Barnett, D., G. Gee, M. Sweeney, M. Johnson, V. Medina, D. Mendoza, B. Fritz, F. Khan, W. Daily, J. Fink, A. Ramirez, M. Levitt, A. Binley, and C. Balyga, 2003, Results of performance evaluation testing of electrical leak-detection methods at the Hanford Site mock tank – FY 2002-2003, PNNL-14192, Pacific Northwest National Laboratory, Richland, WA.

Cowles, M. K., and Carlin, B. C.,1996, "Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review", Journal of the American Statistical Association, 91, 883-904.

Daily, W. A. Ramirez, D. LaBrecque and J. Nitao, 1992, Electrical resistivity tomography of vadose water movement, *Water Resour. Res.*, **28**(5), 1429-1442.

Daily, W. D., W. Lin and T. Buscheck, 1987, Hydrological Properties of Topopah Spring Tuff: Laboratory Measurements, *J. Geophys. Res., ***92**, 7854-7864.

Daily, W., and A. Ramirez, 2000, Electrical imaging of engineered hydraulic barriers, *Geophysics*, 65, no. 1, p. 83-94.

Ellis, R.G. and D.W. Oldenburg, 1994, The pole-pole 3-D DC-resistivity inverse problem: a conjugate gradient approach, *Geophys. J. Int.*, **119**, 111-119.

Fox, C. and Nicholls, G. `Sampling Conductivity Images via MCMC'. In: The Art and Science of Bayesian Image Analysis, K.V. Mardia, C.A. Gill, R.G. Aykroyd eds, Proceedings of the Leeds Annual Statistical Research Workshop (LASR), 91-100, July 1997, Leeds University Press, 1997.

Gee, G. and, A. Ward, 2001,Vadose Zone Transport Field Study: Status Report, Pacific Northwest National Laboratory, PNNL-13679, Richland, WA.

Gelman, A., and Rubin, D. B. ,1992,, "Inference From Iterative Simulation Using Multiple Sequences", Statistical Science, 7, 457-511.

Glaser, R., 2003, Stochastic Engine convergence diagnostics, UCRL-ID-155140 Lawrence Livermore National Laboratory, Livermore, CA.

Jain A. and Dubes R., 1988, *Algorithms for clustering data*, Prentice Hall, Englewood Cliffs, NJ, 07632, 320 pp.

Kaipio, J., V.Kolehmainen, E. Somersalo, and M. Vauhkonen, 2000, Statistical inversion and Monte Carlo sampling in electrical impedance tomography, *Inverse Problems*, vol. 16, p 1487-1522.

Kemna, A, A.Binley, A.Ramirez and W.Daily, 2000, Complex resistivity tomography for environmental applications, *Chemical Engineering Journal*, **77**(1-2), p11-18.

Kolehmainen, V., E. Somersalo, P. Vauhkonen, M. Vauhkonen, and J. Kaipio, 1998, A Bayesian Approach and total variation priors in 3D electrical Resistance Tomography, In: *Biomedical Engineering Towards the Year 2000 and Beyond. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Hong Kong , China, October 29 - November 1, 1998, p. 1028-1031.

LaBrecque, D. J., Miletto, M., Daily, W. D., Ramirez, A. L., and Owen, E., 1996, The effects of noise on Occam's inversion of resistivity tomography data, *Geophysics*, **61**, 538-548.

LaBrecque, D., G. Morelli, W. Daily, A. Ramirez, and P. Lundegard, 1999, Occam's inversion of 3D Electrical Resistivity Tomography, in "Three-dimensional electromagnetics", eds. M. Oristaglio, B. Spies, and M.R. Cooper, Soc. Of Exploration Geophysicists, pages 575-590.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, 1953, Equation of state calculations by fast computing machines, J. Chem. Phys., **1**, no. 6, p. 1087-1092.

Morelli, G., and LaBrecque, D. J., 1996, Advances in ERT modeling, *Eur. J. Environ. Eng. Geophys.*, **1**, 171-186.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems, Journal of Geophysical Research, vol. 100, no. B7, p12431-12447.

Mosegaard, K., 1998, Resolution analysis of general inverse problems through inverse Monte Carlo sampling. *Inverse Problems*, v. 14, pp. 405-426.

Mosegaard, K., and M. Sambridge, 2002, Monte Carlo analysis of inverse problems. *Inverse Problems*, v. 18, pp. R29 - R54.

Mosegaard, K., and C. Rygaard-Hjalsted, 1999, Probabilistic analysis of implicit-inverse problems. *Inverse Problems,* v. 15, pp 573-583

Park, S.K. and G.P. Van, 1991, Inversion of pole-pole data for 3-D resistivity structure beneath arrays of electrodes, *Geophysics*, **56**, 951-960.

Park, S.K, 1998, Fluid migration in the vadose zone from 3D inversion of resistivity monitoring data, *Geophysics*, 63, no. 1, p. 41-51.

Pao, Y. H., 1989, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, Reading, MA

Ramirez, A., and W. Daily, 2003, Monitoring carbon dioxide floods using electrical resistance tomography (ERT): sensitivity studies, *J. Environmental and Engineering Geophysics ,* 08, no. 3, p.187-208.

Raftery, A., and Lewis, S., 1992a, How many iterations in the Gibbs Sampler? in *Bayesian Statistics 4* (J. M. Bernardo et al., editors), Oxford University Press, pp 763-773.

Raftery, A., and Lewis, S., 1992b, The number of iterations, convergence diagnostics, and generic Metropolis algorithms, Technical Report, Department of Statistics, University of Washington, Seattle.

Robert, C., 1995, Convergence control methods for Markov chain Monte Carlo Algorithms, *Statist. Sci,* vol. 10, no 3, pp 231-253.

Robert, C. P.,1998, Discretization and MCMC Convergence Assessment, New York: Springer.

Robert, C. P., Ryden, T., and Titterington, D. M.,1999, "Convergence Controls for MCMC Algotithms, With Applications to Hidden Markov Chains", Journal of Statistical Computation and Simulation, 64, 327-355.

Sasaki, Y, 1994, 3-D resistivity inversion using the finite-element method, *Geophysics*, **59**(11), 1839-1848.

Sengupta, S., and A. Ramirez, 2003, Exploratory global inference from posterior image samples in a Markov chain Monte Carlo simulation experiment: A high-dimensional data clustering approach with extensions to categorical data, Lawrence Livermore National Laboratory, UCRL-JRNL-200985, Livermore CA

Shima, H., 1992, , 2-D and 3-D resistivity image reconstruction using crosshole data, *Geophysics, 57*, 1270-1281.

Slater, L., A.Binley, W.Daily and R.Johnson, 2000, Cross-hole electrical imaging of a controlled saline tracer injection, *J.Applied Geophysics*, **44**, 85-102.

Ward, A. and, G. Gee, 2001,Vadose Zone Transport Field Study: FY 2001 Test Plan, Pacific Northwest National Laboratory, PNNL-13451, Rev. 1, Richland, WA.

Wong A. and Wang D., DECA, 1979, A discrete-valued data clustering algorithm, IEEE Trans. Pattern Analysis and Machine Intelligence**,** vol.1, 342-349.

Yang, X, D. LaBrecque, 1999, Comparison between Stochastic and Occam's Inversion of 3-D ERT data. Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems, Oakland, CA, March 14 – 18, 1999.

Yeh, T.C.J, S. Liu, R. Glass, K. Baker, J. Brainard, D. Alumbaugh, and D. LaBrecque, 2002, A geostatiscally based inverse model for electrical resistivity surveys and its applications to vadose zone hydrology, *Water Resources Research*, vol. 38, no. 12, 1278.

Zhang, J., Mackie, R.L. and Madden, T., 1995, 3-D resistivity forward modeling and inversion using conjugate gradients, *Geophysics*, **60**(5), 1313-1325.

Verify convergence (H)

Confirm that:
a) chains have sampled all important modes

b) collecting additional models will not change characteristics of posterior distribution

Analyze posterior samples: (I)

a) identify the posterior modes and identify the models associated with each

b) summarize the properties of all models associated with each mode

Add accepted model to posterior distribution (G)

Accept if P_{accept} > RN; otherwise reject (F)

Accept

Reject

Note: RN generated according to a uniformly random distribution such that $0 \le RN \le 1$

Compare likelihood $L(x^{(T+1)})$ to that of previous model $L(x^{(T)})$: (E)

calculate P_{accept}:
if $L(x^{(T+1)}) > L(x^{(T)})$
 $P_{accept} = 1$
else:
 $P_{accept} = L(x^{(T+1)})/L(x^{(T)})$

Calculate likelihood of proposed model $L(x^{(T+1)})$ (D)

Solve ERT forward problem (C)

pass model $x^{(T+1)}$ (A$_1$)

Perturb model according to rules based on prior information (base representation algorithm) (B)

pass model $x^{(T)}$ (A$_2$)

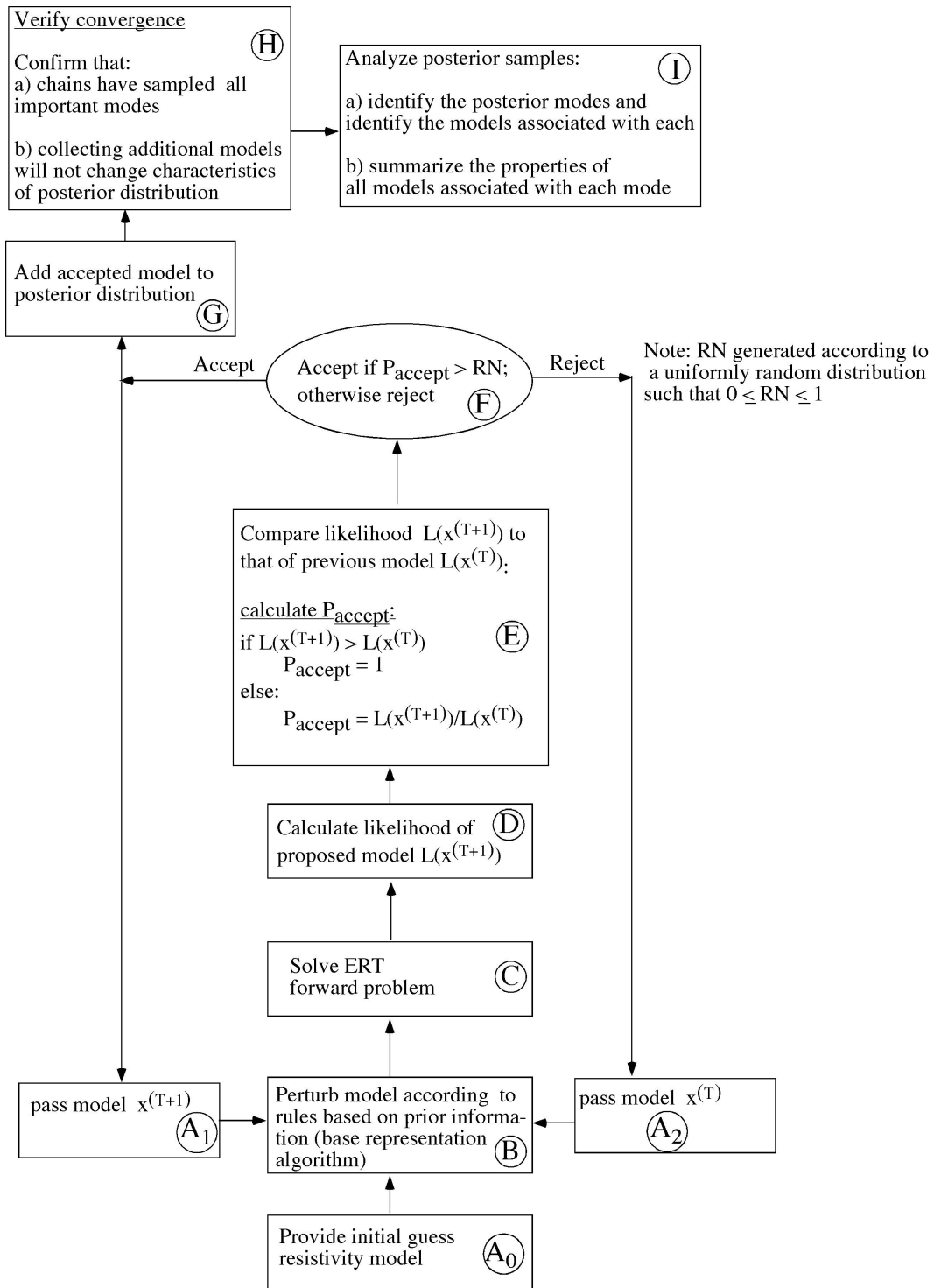Provide initial guess resistivity model (A$_0$)

Figure 1. Schematic diagram of the MCMC inversion process. The key differences between this approach and deterministic inversion are steps B (randomly propose inversion models) and E, F (control how the model is updated).
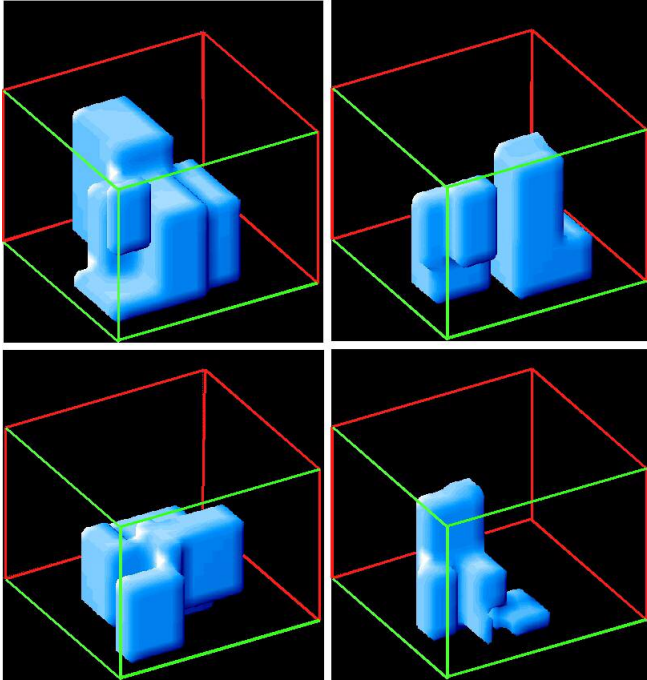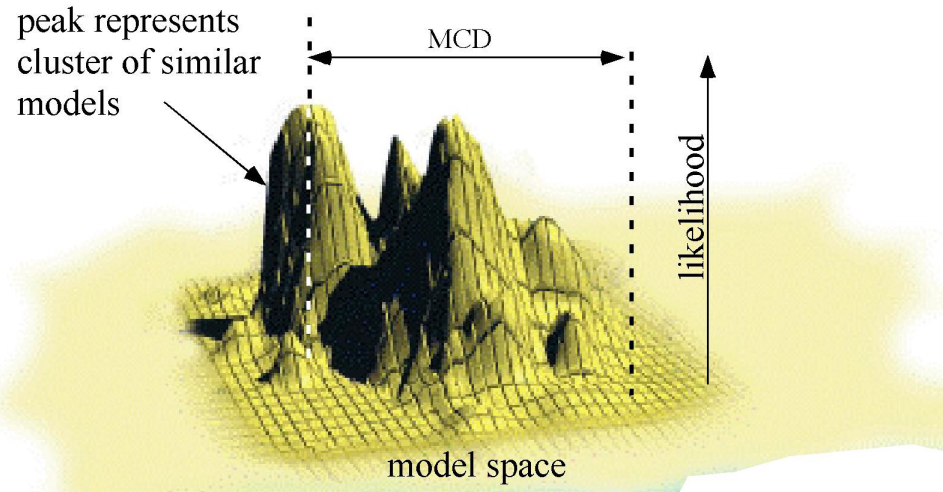
Figure 2. Examples of models generated by the base representation algorithm. A model consists of sub-volumes that can have varying size, shape and resistivity values. The top row models show that some of the sub-volumes can be separate from other subvolumes.

peak represents cluster of similar models

MCD

likelihood

model space

MCD = model to cluster dissimilarity

Figure 3 schematically shows the sector of model space included in the posterior distribution. Each grid node represents one resistivity model. Each hill represents a cluster of resistivity models having similar properties. Multiple peaks indicate that the MCMC inversion has produced non-unique results. The taller peaks identify regions containing models that are most consistent (i.e., most probable) with the observed data.
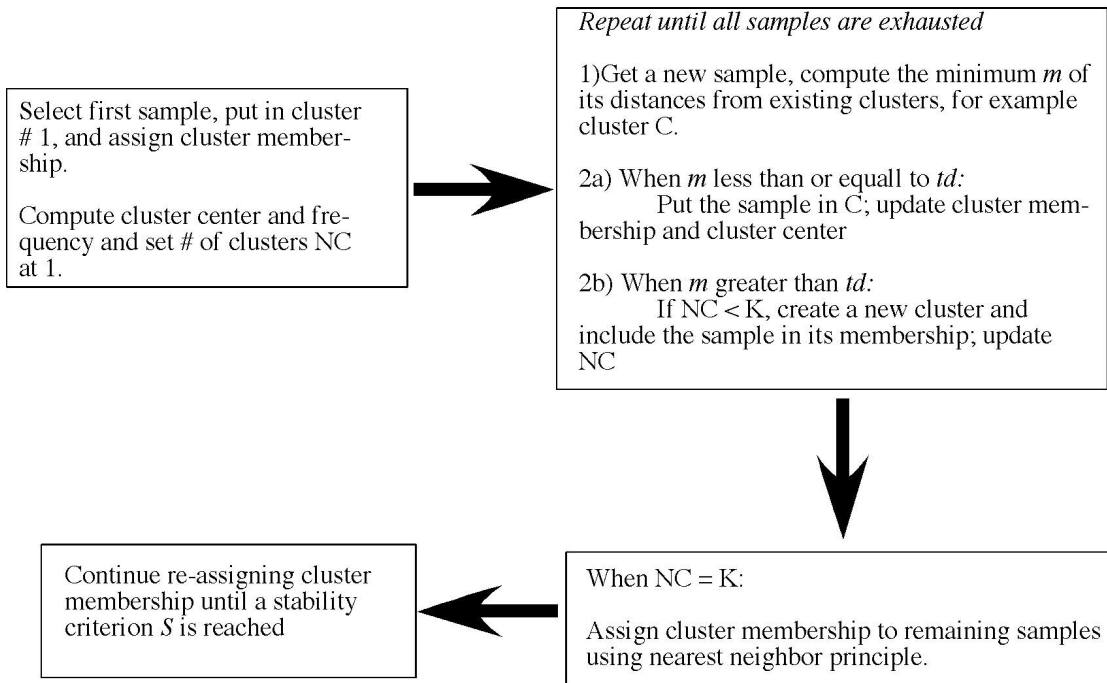
Figure 4. Flow diagram for the " dynamic k means" clustering algorithm used for this work.

The boxes in the flow diagram contain:

Box 1: Select first sample, put in cluster # 1, and assign cluster membership.

Compute cluster center and frequency and set # of clusters NC at 1.

Box 2: *Repeat until all samples are exhausted*

1)Get a new sample, compute the minimum m of its distances from existing clusters, for example cluster C.

2a) When m less than or equall to td:
        Put the sample in C; update cluster membership and cluster center

2b) When m greater than td:
        If NC < K, create a new cluster and include the sample in its membership; update NC

Box 3: When NC = K:

Assign cluster membership to remaining samples using nearest neighbor principle.

Box 4: Continue re-assigning cluster membership until a stability criterion S is reached

Plan
view

electrode array

outline of 3D image volume
shown in Figure 4

water
level

2.8 m

0.6 m

Side
view

0.20 m
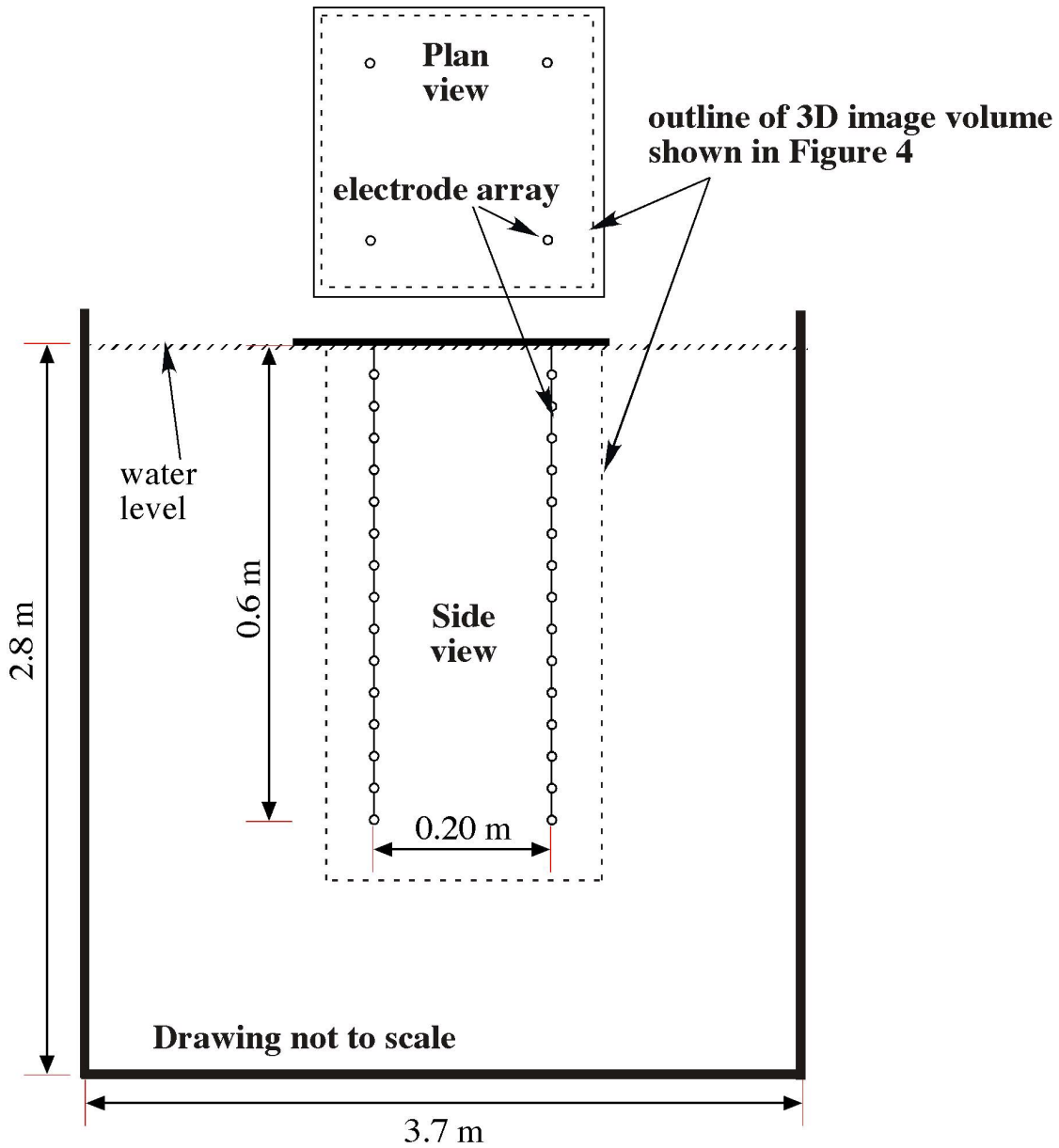
Drawing not to scale

3.7 m

Figure 5 shows a schematic representation of the physical model set up. Four vertical
electrode arrays were immersed in a fiberglass water tank. Various objects were inserted
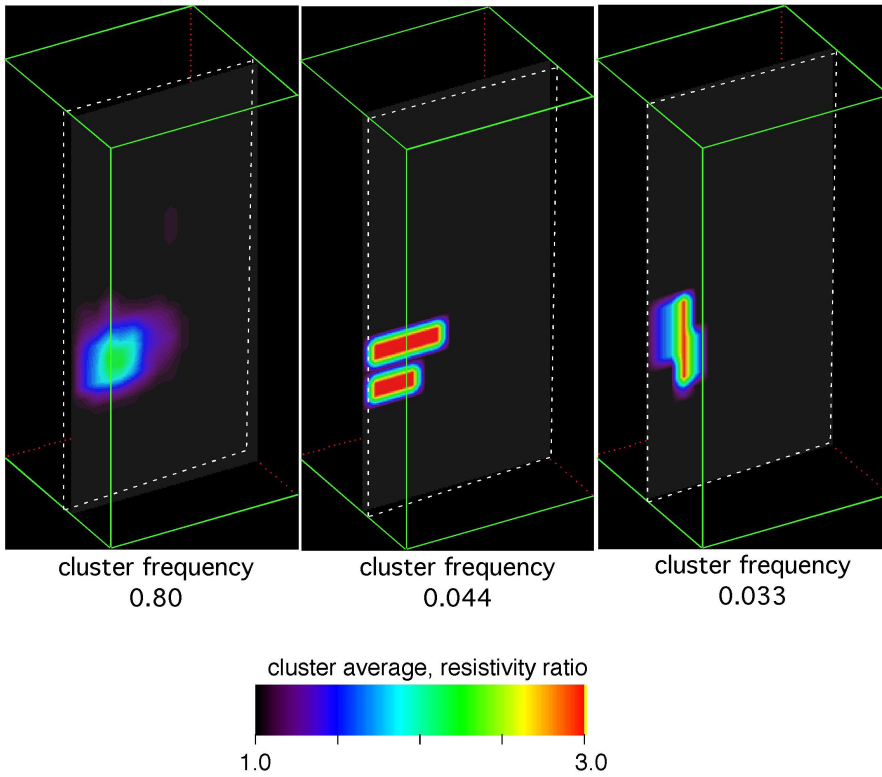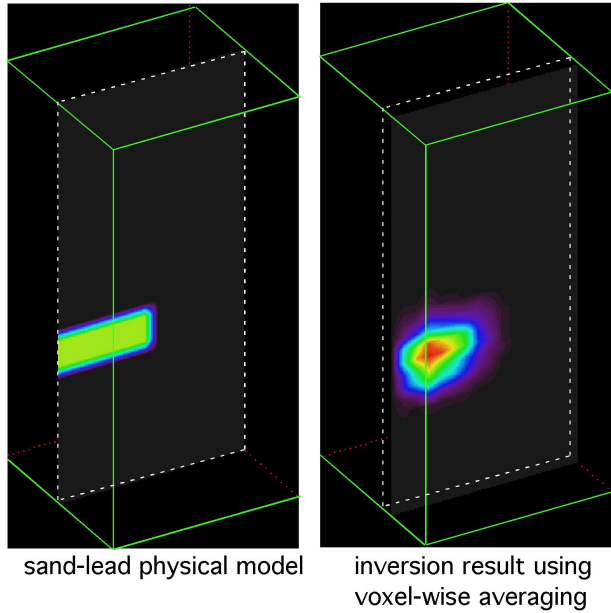between the arrays at a variety of locations and ERT data collected.

Figure 6 presents clustering analysis results for the case of the sand-lead target. The 3D block shown corresponds to the volume enclosed by the dashed lines in Figure 5. The top left frame shows a vertical slice through the target. The vertical slices were placed where the maximum resistivity ratio is observed. The right frame shows the voxel-wise average resistivity obtained when a uni-modal posterior distribution is assumed. The bottom frames shows the average resistivity ratio obtained when a multi-modal distribution is assumed. The three most probable clusters and their corresponding frequencies are shown.
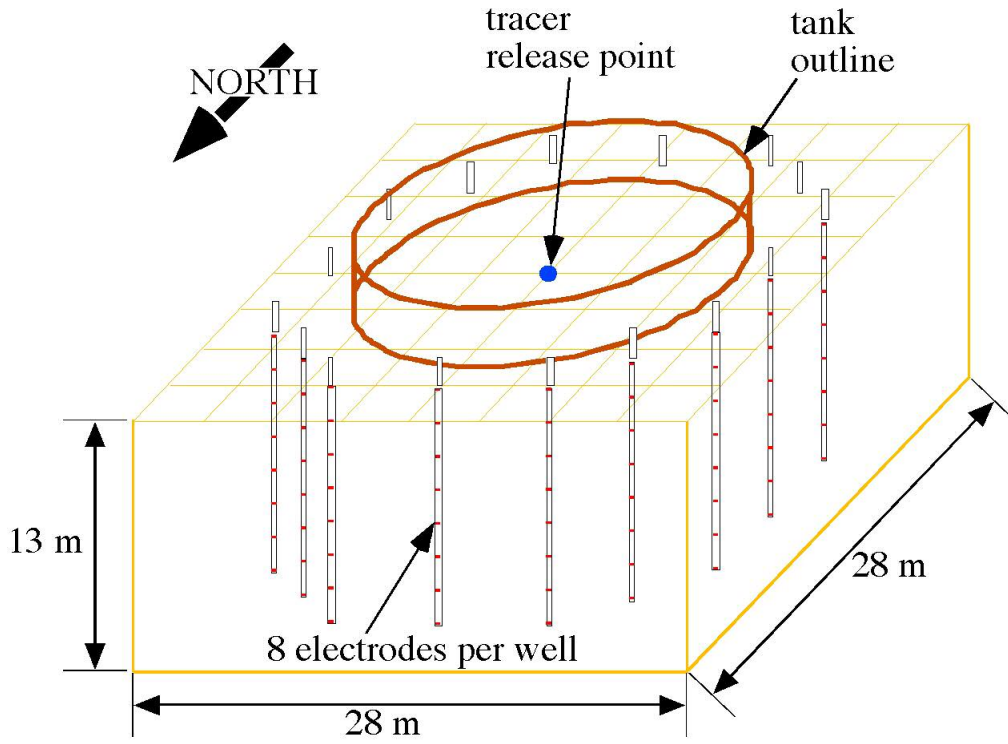
Figure 7 shows a schematic layout of the leaking tank site. Hypersaline brine solution was released from a point near the center of the tank's bottom. Sixteen vertical arrays of electrodes were used to monitor the infiltration process.

Mean value, cluster A, frequency = 0.38

Center state, cluster A

Mean value, cluster B, frequency = 0.18

Center state, cluster B

Mean value, cluster C, frequency = freq 0.12

Center state, cluster C

voxel-wise average resistivity ratio
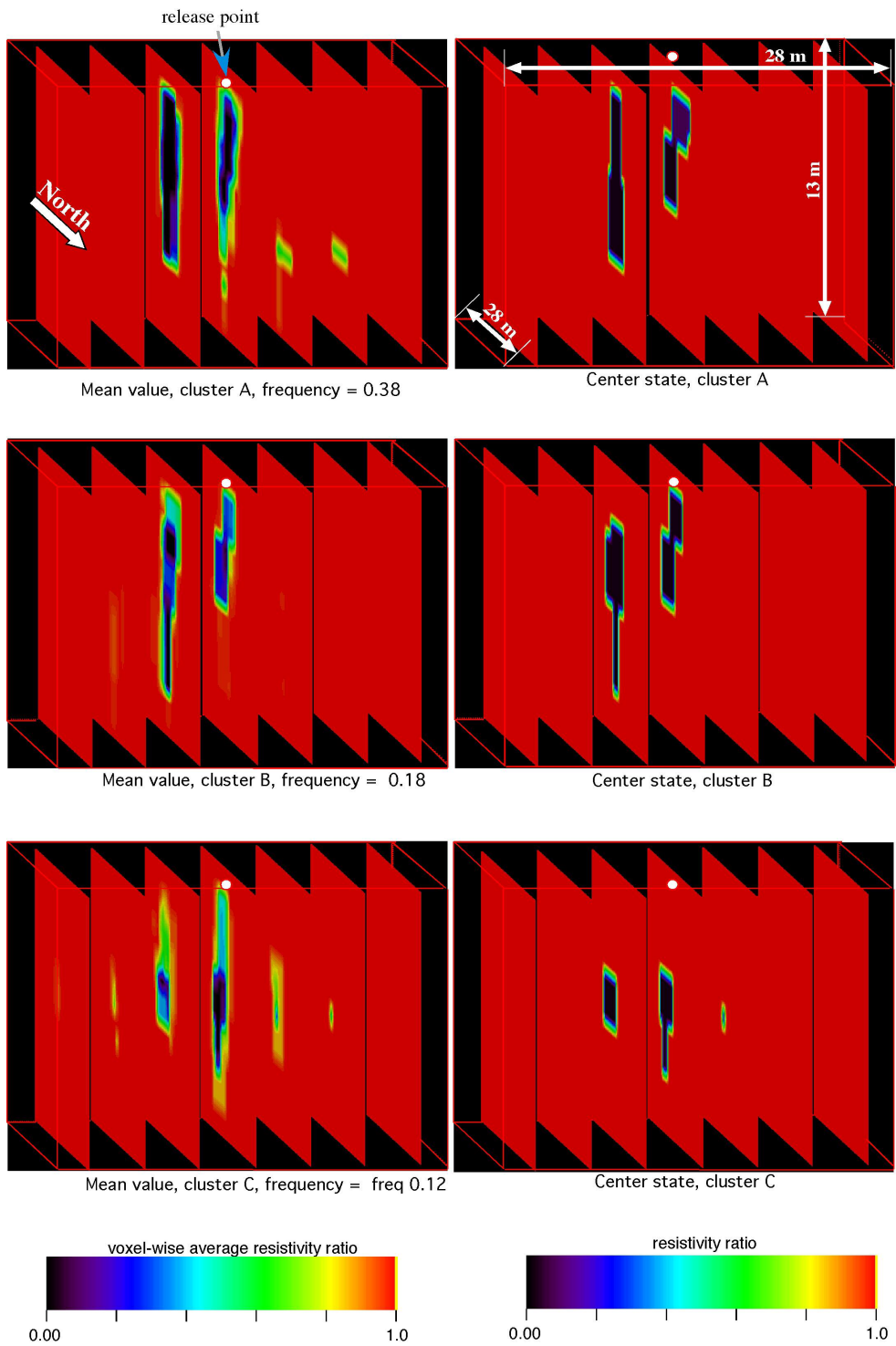
0.00   1.0

resistivity ratio

0.00   1.0

Figure 8 displays clustering analysis results corresponding to the brine release experiment. The 3D block shown here corresponds to the 3D block located beneath the tank in Figure 7. The left column of frames show the voxel-wise average resistivity ratio for the top three most probable clusters. The right column of frames shows the center state for the three most probable clusters.
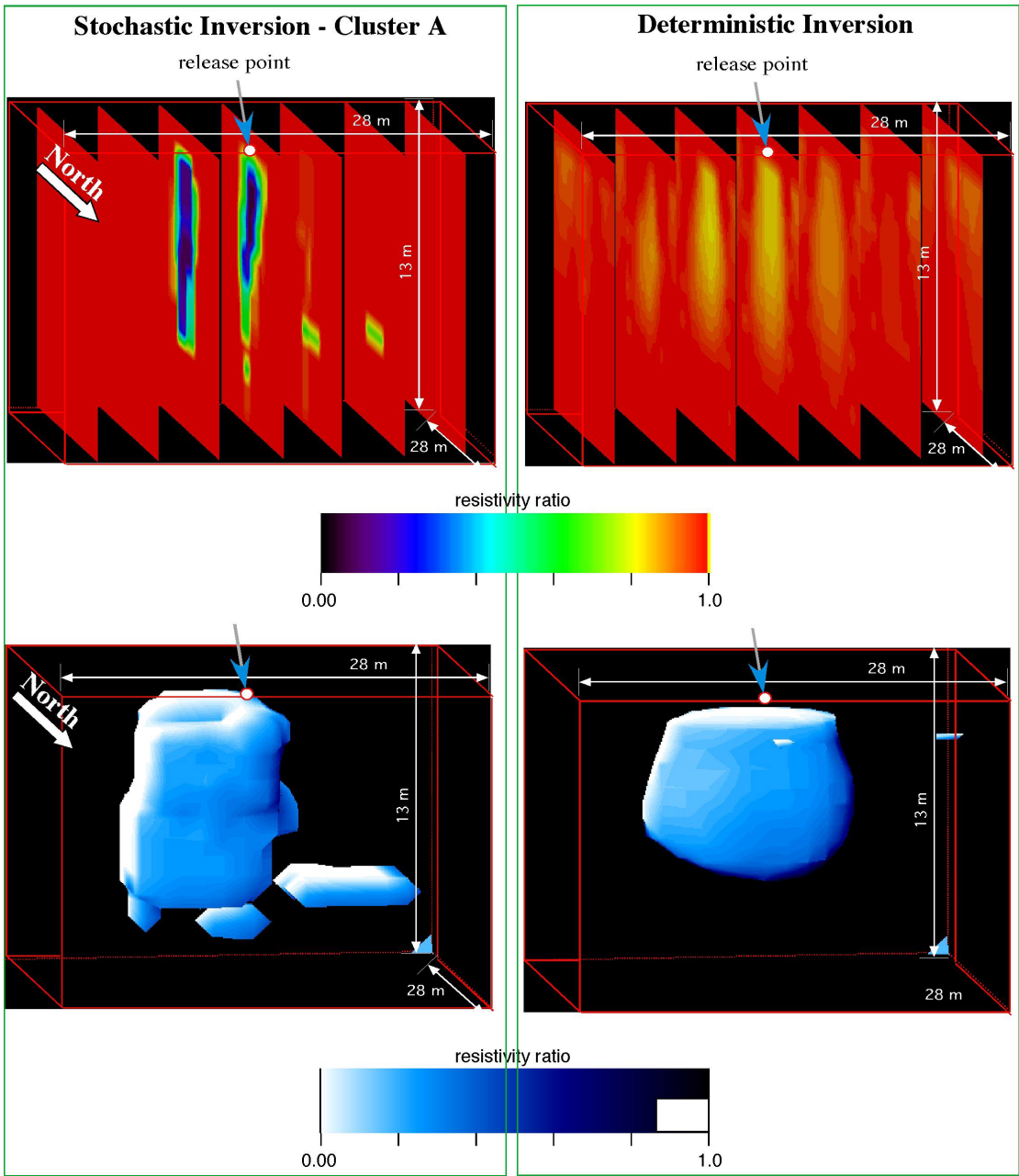
Figure 9 compares stochastic (left frames) and deterministic inversions (right frames). The top row of images consists of a series of vertical slices oriented parallel to North-South line. The bottom row shows iso-surfaces. For the iso-surfaces, all values > 0.9 are transparent.

Figure 10 displays a series of stochastic inversions corresponding to the brine release sequence; released brine volume increases from top to bottom. The left column shows the voxel-wise average resistivity ratio, and the right column shows the same results as iso-surfaces. For the iso-surfaces, all values > 0.9 are transparent.
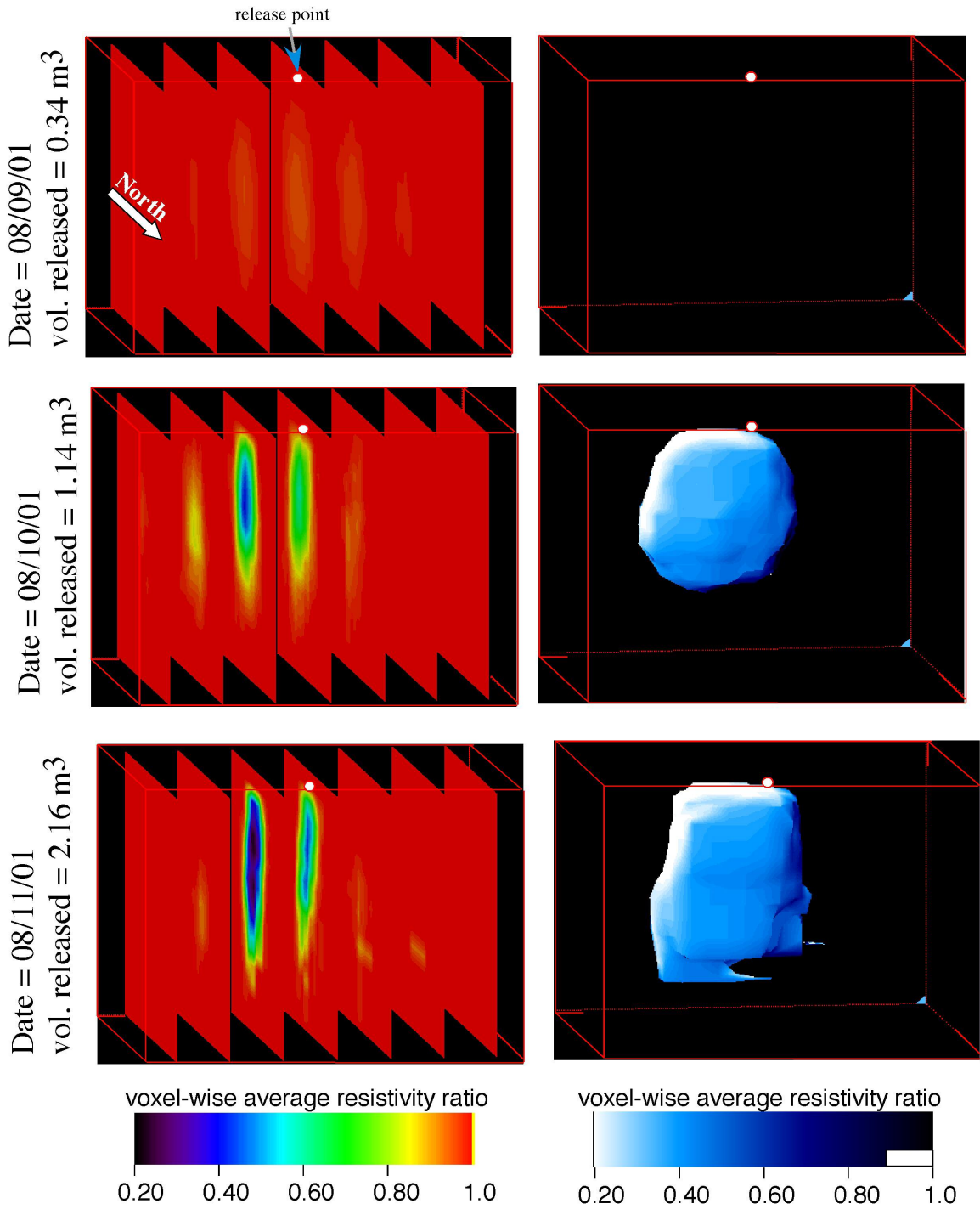
Appendix 1: Base representation algorithm: generate proposal models

Here we describe the algorithm used to generate the resistivity models. The algorithm assumes that the resistivity model consists of a mass of changing electrical resistivity embedded within an otherwise homogeneous volume. The changing volume is composed of two or more sub-volumes consisting of rectangular parallelepipeds. Each sub-volume has a single resistivity value selected from a user-specified set of values; other sub-volumes can have different resistivity values. A sub-volume has to overlap, be contiguous to or be near another sub-volume. These sub-volumes can have varying size, shape and resistivity properties.

Let x^T and x^{T+1} be two adjacent resistivity models in space X; x^T is the current model in the Markov chain and x^{T+1} will be the new model. Let *nvmax* – maximum number of sub-volumes that is allowed.

Algorithm:

For each sub-volume making up the mass of resistivity, propose new size, new location, and new resistivity values for x^{T+1} by perturbing the x^T values.

    1) For each sub-volume *sv* ($1 \leq sv \leq nvmax$) in x^{T+1}:
        1a) Propose the size in elements along the X, Y, Z directions by randomly perturbing the size of model x^T.

            If proposed size is outside acceptable range, flag the proposed size.

        1b) Propose a new location for the each sub-volume by randomly perturbing the location for x^T; do this for the X, Y, Z directions.

            Flag the proposed change if part or all of the sub-volume is outside the allowable region.

        1c) Randomly move through the set of allowable resistivity values to select a new resistivity values for each sub-volume. Suppose that the set of possible resistivity values = {1.0, 2.0, 8.0, 16.0, 32.0} and let the resistivity value for $x^T = 8.0$ (position 3 on the set).

            Choose add a random integer in the range –2 to 2 (or some other user-specified range) and add it to the position number corresponding to x^T. The resistivity value for x^{T+1} will be the value associated with this new position. Thus, if the random integer = 1, the resistivity value for x^{T+1} is the value associated with position 4, 16.0.

Check that proposed set element is one of the positions within the set; flag it if it is not.

1d) Check whether the proposed new size, location or resistivity value are outside the range of permissible values; if so, go back to step 1a and try again.

2) Check that each sub-volume is contiguous, overlaps or is acceptably close to at least one other sub-volume, do this check along the X, Y, Z directions.

If each sub-volume is NOT contiguous (or acceptably close) to at least one other sub volume along all three principal directions go back to step 1.

Save_previous_function:
Save the values of the x^T model (size, location, resistivity model). This function is called before proposing the values for the x^{T+1} model. Do this for all sub-volumes.

reset_function:
This function is used to reset the Markov chain back to the x^T model whenever the proposed x^{T+1} model is rejected by the Metropolis algorithm. This function sets all the values for proposed model x^{T+1} equal to those of x^T by using the values saved by "Save_previous_function".