A sampling-based Bayesian model for gas saturation estimation using seismic AVA and marine CSEM data

Jinsong Chen, Michael Hoversten, and Don Vasco

Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Yoram Rubin and Zhangshuan Hou

University of California, Berkeley, CA 94720

Submission to Geophysics

ABSTRACT

We develop a sampling-based Bayesian model to jointly invert seismic amplitude versus angles (AVA) and marine controlled-source electromagnetic (CSEM) data for layered reservoir models. The porosity and fluid saturation in each layer of the reservoir, the seismic P- and S-wave velocity and density in the layers below and above the reservoir, and the electrical conductivity of the overburden are considered as random variables. Pre-stack seismic AVA data in a selected time window and real and quadrature components of the recorded electrical field are considered as data. We use Markov chain Monte Carlo (MCMC) sampling methods to obtain a large number of samples from the joint posterior distribution function. Using those samples, we obtain not only estimates of each unknown variable, but also its uncertainty information. The developed method is applied to both synthetic and field data to explore the combined use of seismic AVA and EM data for gas saturation estimation. Results show that the developed method is effective for joint inversion, and the incorporation of CSEM data reduces uncertainty in fluid saturation estimation, when compared to results from inversion of AVA data only.

INTRODUCTION

Deepwater gas exploration is challenging and subject to a large degree of uncertainty. Seismic imaging techniques, such as seismic amplitude versus angles (AVA), can provide good information about the physical location and porosity of potential gas-bearing sands, but cannot discriminate between economic and non-economic gas concentrations because seismic velocity and density have low sensitivity to gas saturation (Castagna and Backus, 1993; Debski and Tarantola, 1995; Plessex and Bork, 2000). Marine controlled-source electromagnetic (CSEM) methods have the ability to discriminate between non-economic and economic gas saturation because electrical resistivity of reservoir materials is highly sensitive to gas saturation through the link to water saturation. However, estimating gas saturation using marine CSEM data alone is impractical because EM data has low spatial resolution.

In addition to applications in gas exploration, the addition of EM data can facilitate the use of time-lapse seismic data in predicting changes in pressure and fluid saturation by providing a third, independent source of data. Predictions of changes in pore pressure and water saturation (Tura and Lumley, 1999; Landro, 2001; Lumley et al., 2003) can be done when there is only oil and water saturation to consider, since there are only two independent variables, pressure and water or oil saturation, to be derived from two types of data (acoustic impedance and shear impedance). The presence of gas complicates the problem by introducing a third independent variable (gas saturation) which causes, for example, the change in oil saturation as a function of the change in shear and acoustic impedance to become non-unique.

Seismic AVA and EM methods are sensitive to different physical properties of reservoir materials. Seismic AVA data are functions of the seismic P- and S-wave velocity and density of reservoir materials (Shuey, 1985). EM data are functions of the electrical resistivity of reservoir materials and the overburden. Since both elastic and electrical properties of oil and gas reservoirs are physically related to fluid saturation and porosity through rock physics models (Archie, 1942; Gassmann, 1951; Mavko et al., 1998), joint inversion of seismic AVA and EM data has the potential of providing better estimates of gas saturation and porosity than inversion of individual data sets. However, joint inversion of seismic AVA and EM data for a three-dimensional reservoir is very difficult because forward simulations of seismic AVA and CSEM data are computationally intensive.

In this study, we develop a sampling-based Bayesian model based on layered reservoir models, where the response can be calculated quickly. We apply the developed approach to explore the combined use of seismic AVA and EM data for fluid saturation and porosity estimation. This is a simplified representation of gas exploration in the deepwater of the Gulf of Mexico, where the spatial variability of fluid saturation and porosity changes only along the vertical direction. In addition, we assume that rock-physics models for linking elastic and electrical properties to fluid saturation and porosity are obtainable from nearby borehole logs.

METHODOLOGY

Bayesian Model

The Bayesian model is developed according to an exploration scenario in deepwater, such as the Gulf of Mexico (GOM) and the North Sea. As shown in Figure 1, we consider a layered reservoir model that includes gas, oil, and water. Seismic data are pre-stack common midpoint gathers containing several incident angles over a predefined time window that covers the reservoir. The time window can be determined from check shots or sonic log calculations of time-depth pairs. We invert for water and gas saturation and porosity within the reservoir, but invert for seismic P- and S-wave velocity and density in the zones outside the reservoir. Inversion for seismic parameters outside of the reservoir is done because well logs necessary for deriving rock-physics models to link water and gas saturation and porosity to seismic AVA data are typically only recorded within the reservoir.

Marine CSEM data are the real (in-phase) and quadrature (out-of-phase) components of the recorded electrical field from many receivers located on the seafloor. Those data are the response to the electrical conductivity of the entire half space, which includes the seawater, the overburden above the reservoir, the reservoir, and the sediments below the reservoir. Since the electrical conductivity in seawater and overburden often affects estimates of fluid saturation in the reservoir, we also consider them as unknown parameters in this model.

Let vectors $\mathbf{S_w}$, $\mathbf{S_g}$, and $\boldsymbol{\varphi}$ be water saturation, gas saturation, and porosity in a reservoir, respectively. Let vector $\boldsymbol{\sigma}$ be the electrical conductivity at the overburden. Let vectors $\mathbf{V_p}$, $\mathbf{V_s}$, and $\boldsymbol{\rho}$ be the seismic P- and S-wave velocity and density above and below the gas reservoir. Let matrices $\mathbf{R}$ and $\mathbf{E}$ represent seismic AVA and CSEM data accordingly. We assume that $\mathbf{R}$ and $\mathbf{E}$ are independent of each other. Based on the Bayes' theorem (Bernardo and Smith, 1994), the Bayesian model is given by:

$$f(\mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \boldsymbol{\sigma}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho} \mid \mathbf{R}, \mathbf{E}) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) \\ f(\mathbf{E} \mid \mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma}) f(\mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \boldsymbol{\sigma}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) \quad (1)$$

Equation 1 defines a joint posterior probability distribution function of all unknown parameters, which is known up to a normalizing constant. The first term on the right side of the equation is referred to as the likelihood function of seismic AVA data, the second term on the right side is referred to as the likelihood function of EM data, and the last term on the right side is referred to as the prior distribution of unknown parameters.

Likelihood Models

We determine the likelihood functions of seismic AVA and EM data using different methods according to their error structures in data acquisition and processing. Seismic AVA reflectivity is a direct function of seismic P- and S-wave velocity and density in the reservoir and in the zones outside the reservoir. In our application, we use the Zoeppritz equations to model the reflectivity. Seismic P- and S-wave velocity and density can be related to water and gas saturation and porosity in the reservoir using rock-physics models. Let seismic AVA data $\mathbf{R} = \{r_{ij}\}$, where $i = 1, 2, \cdots, m_a$, and m_a is the number of incident angles, and $j = 1, 2, \cdots, m_d$, and m_d is the number of time samples. Thus,

$$r_{ij} = M_{ij}^a(\mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) + \varepsilon_{ij}^a, \quad (2)$$

where M_{ij}^a is the $ij-th$ component of the seismic AVA forward model and ε_{ij}^a represents its corresponding measurement error. We follow here the common assumption that the measurement errors have a Gaussian distribution with zero mean and are uncorrelated to each other (Mosegaard and Tarantola, 1995; Malinverno, 2002; Buland and Omre, 2003), and thus we obtain the likelihood function of seismic data as follows:

$$f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) = \prod_{\mathrm{i=1}}^{\mathrm{m_a}} \prod_{\mathrm{j=1}}^{\mathrm{m_d}} \frac{1}{\sqrt{2\pi}} \tau_i^{1/2} \exp\left(-\frac{\tau_i}{2} (r_{ij} - M_{ij}^a(\mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}))^2 \right), \quad (3)$$

where τ_i is the inverse variance of measurement errors along the $i-th$ incident angle. The above assumptions can be changed to allow seismic reflectivity for a given incident angle to be spatially correlated in time. In this case, the multivariate Gaussian distribution function may be used for obtaining the likelihood function. If the error structure is other than Gaussian and can be modeled, then a more sophisticated likelihood function may be used.

We determine the likelihood function of the EM data using relative errors instead of absolute errors, because the amplitudes of the recorded electrical field span several orders of magnitude. The EM data used in this study include in-phase and out-of-phase components of the recorded electrical fields at several offsets using different frequencies. Let EM data matrix $\mathbf{E} = \{e_{ijk}\}$, where $i = 1, 2, \cdots, n_f$, represent different frequencies of EM sources, $j = 1, 2, \cdots, n$, represent different offsets, and $k = 1, 2$ represent in-phase and out-of-phase components, respectively. Thus,

$$e_{ijk} = M_{ijk}^e(\mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma})(1 + \varepsilon_{ijk}^e), \quad (4)$$

where M_{ijk}^e is the $ijk-th$ component of the results of the EM forward model, and ε_{ijk}^e is its corresponding relative error. Similar to seismic AVA data, we assume that the relative errors of EM data have a Gaussian distribution with zero mean, and the data collected from different frequencies and offsets are independent of each other. Therefore, we can obtain the following likelihood function:

$$f(\mathbf{E} \mid \mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma}) = \prod_{i=1}^{n_f} \prod_{j=1}^{n_o} \prod_{k=1}^{2} \frac{1}{\sqrt{2\pi}} \tau_{ijk}^{1/2} \exp\left\{ -\frac{\tau_{ijk}}{2} \left(\frac{e_{ijk} - M_{ijk}^{e}(\mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma})}{M_{ijk}^{e}(\mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma})} \right)^2 \right\}, \qquad (5)$$

where τ_{ijk} represents the inverse variance of EM data. As for seismic AVA data, we can develop a more sophisticated likelihood function according to the error structure of EM data.

Prior Model

The prior distribution is determined using prior knowledge and other information about the unknown parameters, such as data from nearby boreholes, which may be subjective and site specific. In this study, we assume that the unknown parameters in the reservoir (i.e. $\mathbf{S}_w, \mathbf{S}_g, \boldsymbol{\varphi}$) are independent of the ones outside the reservoir (i.e. $\boldsymbol{\sigma}, \mathbf{V}_p, \mathbf{V}_s, \boldsymbol{\rho}$) and water and gas saturation ($\mathbf{S}_w, \mathbf{S}_g$) are independent of porosity ($\boldsymbol{\varphi}$). We also assume that the electrical conductivity ($\boldsymbol{\sigma}$) in the thick overburden is independent of seismic velocity and density ($\mathbf{V}_p, \mathbf{V}_s, \boldsymbol{\rho}$) in the thin layers above and below the reservoir and seismic velocity and density ($\mathbf{V}_p, \mathbf{V}_s, \boldsymbol{\rho}$) are independent of each other. Consequently, we can write the prior distribution function as follows:

$$f(\mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \boldsymbol{\sigma}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) = f(\mathbf{S_w}, \mathbf{S_g}) f(\boldsymbol{\varphi}) f(\boldsymbol{\sigma}) f(\mathbf{V_p}) f(\mathbf{V_s}) f(\boldsymbol{\rho}). \qquad (6)$$

The prior distribution functions of water and gas saturation are determined jointly because they are dependent of each other. Let (a_1, b_1), (a_2, b_2), and (a_3, b_3) be the prior bounds of water, gas, and oil saturation. As shown in Figure 2, the inversion domain of water and gas saturation is a joint set given by

$$D = \{(S_w, S_g):\ S_w \in (a_1, b_1),\ S_g \in (a_2, b_2),\ S_o \in (a_3, b_3),\ S_w + S_g + S_o = 1\}.$$

We assume that the prior distribution of water and gas saturation is uniform in the domain D. For all other parameters, we assume that their prior distributions are uniform within their corresponding ranges.

SAMPLING-BASED METHODS

We use Markov chain Monte Carlo (MCMC) sampling methods to obtain estimates of unknown parameters from the Bayesian model defined in Equation 1. Unlike optimization-based methods seeking single optimal solutions of unknown parameters, MCMC sampling-based methods draw many samples from the joint posterior distribution. Using those samples, we can make inferences about the marginal distributions of each parameter, such as its mean, variance, and predictive intervals.

MCMC sampling methods have been found recently to be useful for inverting complex geophysical data set by numbers of authors, such as Bosch (1999), Malinverno (2002), and Buland et al. (2003). The main steps for using MCMC methods entail: (1) deriving conditional probability functions given all the data and other unknown variables, which are referred to as full conditional distribution functions; (2) generating samples using suitable algorithms; (3) making inferences about each unknown. In the following, we first show the full conditional distribution functions of unknown vectors given in Equation 1 and then describe the sampling algorithms used in this study, which include the Metropolis-Hasting methods (Hasting, 1970) and the slice sampling methods (Neil, 2003).

Full Conditional Distributions

A full conditional distribution function is proportional to the joint posterior distribution function shown in the left side of Equation 1. By retaining only those terms related to the variable of interest, we can obtain the full conditional distribution function of the variable. For example, the full conditional probability distribution function of porosity is given by

$$f(\boldsymbol{\varphi} \mid \cdot) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\mathbf{E} \mid \mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma}) f(\boldsymbol{\varphi}) . \quad (7)$$

The dot in $f(\boldsymbol{\varphi} \mid \cdot)$ of Equation 7 represents all the data and other unknown variables. Similarly, we can obtain full conditional distribution functions of other unknown variables, which are given below:

$$\begin{aligned}
& f(\mathbf{S_w}, \mathbf{S_g} \mid \cdot) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\mathbf{E} \mid \mathbf{S}_w, \boldsymbol{\varphi}, \boldsymbol{\sigma}) f(\mathbf{S_w}, \mathbf{S_g}) \\
& f(\boldsymbol{\sigma} \mid \cdot) \propto f(\mathbf{E} \mid \mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma}) f(\boldsymbol{\sigma}) \\
& f(\mathbf{V_p} \mid \cdot) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\mathbf{V_p}) \\
& f(\mathbf{V}_s \mid \cdot) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\mathbf{V}_s) \\
& f(\boldsymbol{\rho} \mid \cdot) \propto f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\boldsymbol{\rho}).
\end{aligned} \quad (8)$$

Each conditional distribution function shown in Equation 8 can only be evaluated numerically for given samples of unknown parameters.

Metropolis-Hasting Sampling Methods

We use different methods to draw samples of fluid saturation, porosity, overburden conductivity, seismic velocity, and density from the joint posterior distribution function shown in Equation 1. For water and gas saturation, to take account of dependence between water and gas saturation, we use the multivariate Metropolis-Hasting method. Suppose we start from an initial value $(\mathbf{S}_w^{(0)}, \mathbf{S}_g^{(0)})$, which can be any vector in domain D shown in Figure 2. We want to obtain a new sample of water and gas saturation $(\mathbf{S}_w^{(1)}, \mathbf{S}_g^{(1)})$. The steps of the procedure are:

(1) Generate a random vector $(\mathbf{S}_w^{*}, \mathbf{S}_g^{*})$ uniformly from domain D shown in Figure 2.

(2) Calculate the following ratio:

$$\alpha = \min\left(1, \frac{f(\mathbf{R} \mid \mathbf{S}_w^{*}, \mathbf{S}_g^{*}, \varphi^{(0)}, \mathbf{V}_p^{(0)}, \mathbf{V}_s^{(0)}, \boldsymbol{\rho}^{(0)})}{f(\mathbf{R} \mid \mathbf{S}_w^{(0)}, \mathbf{S}_g^{(0)}, \varphi^{(0)}, \mathbf{V}_p^{(0)}, \mathbf{V}_s^{(0)}, \boldsymbol{\rho}^{(0)})} \cdot \frac{f(\mathbf{E} \mid \mathbf{S}_w^{*}, \varphi^{(0)}, \boldsymbol{\sigma}^{(0)})}{f(\mathbf{E} \mid \mathbf{S}_w^{(0)}, \varphi^{(0)}, \boldsymbol{\sigma}^{(0)})}\right) \tag{9}$$

(3) Generate a random value u uniformly from interval (0,1)

(4) If $\alpha < u$, let $\mathbf{S}_w^{(1)} = \mathbf{S}_w^{*}$ and $\mathbf{S}_g^{(1)} = \mathbf{S}_g^{*}$; otherwise, let $\mathbf{S}_w^{(1)} = \mathbf{S}_w^{(0)}$ and $\mathbf{S}_g^{(1)} = \mathbf{S}_g^{(0)}$.

Repeating steps from (1) to (4) by replacing index (0) with index (1), we can obtain many samples of water and gas saturation as follows: $\left\{(S_w^{(k)}, S_g^{(k)}) : k = 0, 1, \cdots, n\right\}$. From the procedure, we can see that the value $(S_w^{(k+1)}, S_g^{(k+1)})$ depends solely on the value $(S_w^{(k)}, S_g^{(k)})$, not on the value $\left\{(S_w^{(k-1)}, S_g^{(k-1)}) : k = 1, 2, \cdots, n\right\}$. Therefore, these samples form a Markov chain. This chain has been shown to be converged to the joint true posterior distribution of random vector (S_w, S_g) under weak conditions when the sample size is large (Gilks et al., 1996).

For porosity, overburden conductivity, seismic velocities, and density, we use mixing methods, which include single-variable Metropolis-Hasting methods, multivariate Metropolis-Hasting methods, single-variable slice sampling methods, and multivariate slice sampling methods. Since the Metropolis-Hasting methods are very similar to the one just given above, in the following we describe only the single-variable slice sampling methods for generating samples of porosity from the joint distribution function. Similar methods can be used to obtain samples of other variables.

Slice Sampling Methods

Slice sampling methods are described in details by Neil (2003). To use single-variable slice sampling methods to obtain samples of porosity in a given layer (say, layer-i), we first transform porosity φ_i from a given finite interval (c, d) to an infinite domain $(-\infty, +\infty)$, where c and d are the lower and upper bounds of porosity, using the following formula:

$$x = \log\left(\frac{\varphi_i - c}{d - \varphi_i}\right),$$

where variable x is the transformation of porosity φ_i defined on $(-\infty, +\infty)$. To simplify description, we let

$$f(x) = f(\mathbf{R} \mid \mathbf{S_w}, \mathbf{S_g}, \boldsymbol{\varphi}, \mathbf{V_p}, \mathbf{V_s}, \boldsymbol{\rho}) f(\mathbf{E} \mid \mathbf{S_w}, \boldsymbol{\varphi}, \boldsymbol{\sigma}) f(\boldsymbol{\varphi}) \tag{10}$$

Notice that all variables on the right side of Equation 10 are vectors, and vector $\boldsymbol{\varphi}$ includes component φ_i and therefore is a function of x. Function $f(x)$ is the marginal posterior probability density function (pdf) of x. Our goal is to make inferences about this pdf using sampling methods.

Figure 3 shows a three-step procedure given by Neil (2003) to obtain a new value x_1 from the current value x_0.

(1) Draw a value y, which is uniformly distributed on $(0, f(x_0))$. The value y defines a horizontal "slice": $S = \{x : y < f(x)\}$, shown as thick lines in Figure 3(a). Note that x_0 is always within S.

(2) Find an interval, $I = (L, R)$ around the value x_0 that contains all or much of the slice, where L and R represent the lower and upper bounds of the interval.

(3) Draw the new point x_1 from the part of the slice within the interval.

Steps (2) and (3) can be implemented in several ways. In this study, we use "stepping out" methods to find the interval I and "shrinkage" methods to draw a new value from the interval. Figure 3(b) shows the main idea of "stepping out" methods. We step out in both directions from the value x_0 with a given interval width w for a given maximum number of iteration m until both ends are outside the slice. We then uniformly pick a new value from the interval. If the point picked is inside the slice, we consider it as the new value x_1; otherwise we use the point to shrink the interval. Neil (2003) shows how this procedure guarantees that the obtained chains are converged to pdf $f(x)$. The parameters w and m are specified beforehand. They do not affect sampling results, but do affect speed of convergence of the chains.

Strategies for Speeding Convergence

The success of MCMC methods depends on the efficiency of the sampling methods used. If a sampling method is inefficient, we may need to run a very long chain, and thus computational efforts are very large. Typically, the raw sampling methods (for both Metropolis-Hasting and slice sampling) are not very efficient. We need to tune those parameters that control chain movements. Unfortunately, the efficiency of a sampling method is often problem-specific. In this study, we use a mixing method to obtain samples. At each sampling step, we randomly pick one of the following four methods, single-variable Metropolis-Hasting methods, multivariate Metropolis-Hasting methods, single-variable slice sampling methods, and multivariate slice sampling methods. This strategy has been shown to be very efficient for solving our joint inversion problems.

SYNTHETIC EXAMPLE

The following synthetic case study is designed to show the capability and flexibility of our joint inversion approach for integrating different types of information, and to demonstrate the benefits resulting from joint inversion of seismic AVA and CSEM data for gas saturation estimation.

Synthetic True Model

The synthetic reservoir includes five layers with a thickness of 25 m and zero oil saturation. From the upper to the bottom layers, the gas saturation of the reservoir is 0.05, 0.95, 0.4, 0.9, and 0.1, and porosity is 0.15, 0.25, 0.15, 0.1, and 0.05. Above the gas reservoir is an overburden with a thickness of 2,000 m and electrical conductivity of 1.0 S/m, above the overburden is 1,000 m of seawater with conductivity of 3.2 S/m. To account for uncertainty in selecting suitable time windows for seismic AVA data, we add two 25 m thick layers above the reservoir and one 25 m thick layer below the reservoir, which have Vp, Vs and density that are also considered as inversion parameters.

Synthetic Data

The seismic AVA data are NMO-corrected angle gathers generated by convolving a wavelet with the angle-dependent reflectivity for each layer interface. The seismic velocities and density in the reservoir are calculated from porosity and fluid saturation using the rock-physics models given in Table 1 and described in Hoversten et al. (2006).

A 28 Hz Ricker wavelet is use for incident angles of 5, 10, 15, 20, 25, and 30 degrees. Zoeppritz equations (Shuey, 1985) are used to calculate the angle-dependent reflectivity. We assume that the synthetic seismic data include Gaussian random noises. The variances of the Gaussian noises are angle dependent, which are determined by the assigned signal-to-noise ratios (SNRs). For example, we can assign the signal-to-noise ratio of the six incident angles as 6, 5, 4, 3, 2, and 1 from the near to far offsets.

The synthetic EM data mimic commercial EM field data collected using controlled-source electromagnetic (CSEM) techniques. The marine CSEM system consists of a ship-towed electric dipole source and a number of seafloor-deployed recording instruments capable of recording orthogonal electric (and optionally magnetic) fields. A common configuration consists of an electric dipole transmitter, 100–300 m in length, towed in a neutrally buoyant configuration approximately 50 m off the seafloor to avoid collision with stationary receiver systems (Ellingsrud et al., 2001). A square wave of electric current is sent into the transmitter at a variable fundamental frequency between 0.01 and 10 Hz. The earth response, along with the primary field from the transmitter, is measured at the array of receivers. In this study, we use EM sources with five different frequencies (0.1, 0.25, 0.5, 0.75, and 1 Hz) and six different offsets (4, 5, 6, 7, 8, and 10 km). The relation between electrical resistivity and water saturation and porosity is given by Archie's law using coefficients given in Table 2. We added 5–15% relative Gaussian random noises to the synthetic EM data from the near to far offsets.

Inversion Results

Inversion using seismic AVA data:

We first demonstrate the capability of the developed stochastic model for distinguishing high gas saturation layers from low gas saturation layers, using seismic AVA data only for the five-layer reservoir model. We consider two levels of noise, which correspond to signal-to-noise ratios from 6 to 1 and from 12 to 2 from near to far offsets. We assume that oil saturation is zero in each layer, and thus the unknown parameters in Equation 1 are porosity and gas saturation in the five layers. We assume no prior information about gas saturation, and thus the prior distribution of gas saturation is uniform on [0, 1]. The prior distribution of porosity is uniform on [0, 0.35].

Figure 4 shows the estimated gas saturation and porosity using seismic AVA data with the two levels of noise. The back and red curves are the estimated marginal pdfs of gas saturation and porosity in the five layers using seismic AVA data with SNRs from 6 to 1 and from 12 to 2, respectively. The blue straight lines are the true values. From the figure, we notice that seismic data provide (1) good estimates of porosity in each layer, (2) good estimates of gas saturation in layers 1 and 2, and (3) poor estimates of gas saturation in layers 3, 4, and 5. With the decrease of seismic noise, uncertainty in gas saturation and porosity decrease, but in each case both gas-rich layer 4 and gas-poor layer 5 are poorly resolved by seismic data. Table 3 shows the root-mean-squares (RMS) of the differences between the true values and the estimated values using seismic data with SNRs from 6 to 1 and from 12 to 2. Based on comparison between the true values and the estimated means, medians, and modes, the improvement in accuracy for both gas saturation and porosity is small when the SNRs of seismic data are increased by a factor of two.

Inversion using seismic AVA and CSEM data:

To demonstrate the benefits of adding CSEM data for gas saturation estimation, we jointly invert seismic AVA data given in the proceeding section and CSEM data with relative noise between 5% and 15% from the near to far offsets. Figure 5 shows the estimated marginal pdfs of gas saturation and porosity in the two situations. Compared to Figure 4, we see that joint inversion of seismic AVA and CSEM data reduces uncertainty in gas saturation estimation, especially for layers 4 and 5. Most importantly, both high gas concentrations (layers 2 and 4) and low gas concentrations (layers 1 and 5) are clearly identified by the major modes of their corresponding pdfs. Table 4 shows the RMS of the differences between the true values and the estimated values using both seismic and EM data. Compared to Table 3, the incorporation of EM data improves the estimates of gas saturation and porosity, and the improvement in gas saturation estimation is significant.

Inversion with unknown oil concentration:

In the preceding inversion, we assume that oil saturation in the reservoir is zero or given. However, in reality, oil saturation in each layer may be another parameter that needs to be estimated. To explore the effects of our prior knowledge about oil saturation on the joint inversion, we assume that oil saturation may lie in the ranges of [0, 0.1], [0, 0.3], and [0, 0.5], respectively. The true oil saturation remains zero in all layers, but we allow oil saturation to take values between the above ranges when we invert seismic AVA and EM data.

Figure 6 compares the estimated water and gas saturation using both seismic AVA and EM data when the allowed ranges of oil saturation are [0, 0.1], [0, 0.3], and [0, 0.5]. The black, red, and blue curves represent the estimated pdfs of water and gas saturation using the prior bounds of oil saturation [0, 0.1], [0, 0.3], and [0, 0.5], and the solid blue lines represent the true values. Overall, uncertainty in water and gas saturation estimation increases as the prior bounds of oil saturation increases, especially for Layer 4. In terms of the modes of the estimated pdfs, the joint inversion provides fair estimates of water saturation for layers 2–5. For Layer 1, when the prior bound of oil saturation is [0, 0.5], two modes exist and the second one is close to the true value. For gas saturation estimation, the joint inversion provides good estimates of gas saturation in each layer when the prior bound of oil saturation is [0, 0.1]. As the prior bounds of oil saturation increase, the joint inversion cannot identify the second gas-rich layer.

Table 5 shows the RMS of the differences between the true values and the estimated values obtained from the join inversion. It is evident that with the increasing of the prior bounds of oil saturation, the misfits between the true and estimated values increase for both water and gas saturation estimation. We also notice that the modes of the estimated pdfs provide much better estimates of the corresponding true values than other statistics.

FIELD EXAMPLE

We apply the developed Bayesian model to field data obtained from the Troll site in the North Sea. We compare estimates of gas saturation with the measurements at a nearby borehole to evaluate the benefits of joint inversion of seismic AVA and CSEM data for gas and oil saturation estimation under field conditions. In the following, we first briefly describe seismic and EM data used in this study and then focus on the inversion results and their comparisons with the borehole logs. Detailed information on this site was given in Hoversten et al. (2006).

Seismic AVA data

Many types of geophysical surveys have been carried out over the past 30 years. For this study, we choose seismic AVA and marine CSEM data near a well referred to as 31/2-1. Seismic AVA data were collected from CDP locations within 50 m of EM receivers sit on the seafloor. At each CMP location, there are six incident angles (7.2, 13.5, 19.8, 25.6, 31.1, and 36.3 degrees). Figure 7 shows the pre-stack NMO corrected data at CMP 1147, which is near the marine EM receiver Rx16. From the figure, we can see a strong reflector near 1,500 ms, which may correspond to the top of the gas reservoir. Similarly, we can approximately determine the bottom of the reservoir, which is around 1,800 ms. Consequently, we chose the pre-stacked data between 1,418 ms and 1,816 ms for seismic AVA inversion. Angle-dependent wavelets were also derived by matching seismic data at a well 1.5 km from this site.

Marine CSEM Data

Marine EM surveys measure the electromagnetic responses of electrical resistivity in the entire half-space under the surface of ocean, which includes seawater, overburden, gas reservoir, and bedrock. The recorded EM data are the in-phase and out-of-phase electrical fields at several different frequencies. In this study, we use data collected at three frequencies (0.25 Hz, 0.75 Hz, and 1.25 Hz). To be consistent with the seismic AVA data, we use the EM data obtained from receiver Rx16 for eight different transmitters, whose source-receiver distances are 775, 1,700, 2,500, 3,300, 4,100, 4,900, 5,700, and 6,500 m, respectively. The relative errors of those EM data are estimated to be 10%.

Rock-Physics Models

The rock-physics models used for this application include petrophysical models that link fluid saturation and porosity to seismic velocities and density, and the empirical relationship that links fluid saturation and porosity to electrical resistivity. The models for tying seismic elastic properties to reservoir parameters for North Sea sandstones are described by Dvorkin and Nur (1996). The fitted results for joint inversion of seismic AVA and CSEM data using data collected from borehole 31/2-1 are given in Hoversten et al. (2006) and shown in Table 1. Archie's law is used to model the link between electrical resistivity and porosity and water saturation with the constants given in Table 2.

Prior Distribution

We use different methods to determine prior distributions of unknown parameters in the potential reservoir and in the zones outside the reservoir. For the unknown parameters in the zones outside the reservoir, including seismic P- and S-wave velocities, density, and overburden electrical resistivity, we assume that those variables are uniformly distributed between 70% and 130% of their corresponding borehole logs collected from 31/2-1. For the unknown parameters in the reservoir, we assume that those variables are also uniformly distributed with given bounds. We first determine reference values for water and gas saturation. The values for water saturation range from zero to one; values for gas saturation range from one to zero from the top to the base of the reservoir. The bounds for water and gas saturation are the reference values plus or minus 0.3. The lower bounds of oil saturation are zero for all the layers, and the upper bound is 0.1 for depths less than 1,544.5 m, below 1,544.5 m, the upper bounds of oil saturation linearly decrease from 0.7 to 0.1 at the base of the reservoir to allow oil where it was originally present.

Inversion of Seismic AVA Data

For inversion of seismic data, we divide the potential reservoir into 16 layers with a thickness of 20 m and consider porosity, water saturation, and gas saturation in each layer to be unknown parameters. To account for uncertainty in the time-depth function that provides the time window for the seismic AVA data, we also include five 20 m thick layers above the reservoir and one 20 m thick layer below the reservoir. We jointly invert seismic P- and S-wave velocity and density in those layers.

Figure 8 shows the comparison between the inversion results using seismic AVA data and borehole logs collected from Well 31/2-1. The thin black lines show borehole logs, the solid red lines show the modes of the inverted parameters, and the red dotted lines show the 95% predictive intervals. We notice that the seismic AVA data provide good estimates of water saturation for layers 1–6, where water saturation is low. Since we only allow oil saturation to change within 0.0 and 0.1, we obtain good estimates of gas saturation. However, for layers 7–13, inversion of seismic AVA data only provides poor estimates of water saturation. This is because seismic data have low sensitivity when water saturation in the range of 0.3–0.8, as found by Hoversten et al. (2006).

Joint Inversion of Seismic AVA and CSEM Data

To jointly invert seismic AVA and EM data, we need to take account the effects of electrical overburden above the reservoir. In this study, we divide the overburden (including seawater) above the reservoir into 13 layers according to the resistivity logs collected from borehole 31/2-1, and consider electrical conductivity in each of those layers as unknown parameters. Figure 9 shows the prior bounds (dashed lines) and initial values (solid lines) of overburden conductivity derived from borehole 31/2-1. We assume that the unknown overburden conductivity parameters are uniformly distributed in the given ranges.

Figure 10 shows the inversion results using both seismic AVA and CSEM data. Similar to Figure 7, the thin back lines show borehole logs, the solid red lines show the modes of the inverted parameters, and the red dotted lines show the 95% predictive intervals. For layers 1–6, the joint inversion provides similar estimates of water and gas saturation, both of which are close to borehole logs. For layers 7–13, the joint inversion generally provides slightly better estimates of water saturation than those obtained from the inversion of seismic data only, but uncertainty in both estimations is large. Figure 11 from the top to the bottom compares the pdfs of estimated water and gas saturation for layers 7–12 using seismic data only (black lines) and using both seismic and EM data (red lines). The blue lines show the results from the nearby borehole. Based on the comparison between the estimated modes and the borehole logs, we can see that the joint inversion is better. Table 6 shows the RMS of the differences between the averaged well-log values and the estimates values (modes) from the Troll data sets. Again, in terms of the misfits, the joint inversion is also slightly better.

Data Misfits

To show how the estimated model (modes) of inversion fits the data, we compare the seismic AVA and CSEM responses of the estimated model with their corresponding seismic AVA and EM data. Figure 12 compares the in-phase and out-of-phase electrical fields of the estimated model with the EM data at three frequencies over eight different offsets. We found that the estimated model fits the electrical field at frequency 0.25 Hz well. However, for high frequencies (0.75 Hz and 1.25 Hz), the fitting is not good when the offsets are larger than 3,500 m. The misfits for large offsets at high frequencies could have many different causes. One possible reason for misfits is the assumption of the layered reservoir model. The higher-frequency EM data typically have higher resolution and therefore make the three-dimensional localized features of the gas reservoir easier to detect. Figure 13 compares the misfits of seismic AVA data for the models obtained from the joint inversion of seismic and EM data. We can see that joint inversion fits the seismic data well.

DISCUSSION AND CONCLUSIONS

In this paper we describe a sampling-based Bayesian model that we have developed to estimate gas saturation using seismic AVA and marine CSEM. We also demonstrate the capability of this model for solving nonlinear and non-unique inverse problems using synthetic five-layer reservoir models. Compared to deterministic inverse methods, which typically provide single-valued estimations and have difficulty finding global solutions, the stochastic methods provide us marginal pdfs for unknown parameters, which may include multiple modes. The derived pdfs allow us to evaluate the mean, variance, mode, and predictive intervals, all of which are useful in quantifying the uncertainty associated with inversion.

We also use a five-layer synthetic model to show the benefits of joint inversion of seismic AVA and CSEM data for gas saturation estimation under several conditions. Using seismic AVA data only, even with very high resolution, it is still difficult to distinguish high or low gas concentrations in deep layers, because seismic properties are insensitive to gas concentrations. With the inclusion of CSEM data, uncertainty in gas saturation estimation is decreased, and the ability to identify high gas concentrations in deep layers is enhanced. The improvement becomes less prominent when the errors of CSEM data and uncertainty in overburden conductivity are large, and when the reservoirs include unknown concentrations of oil. In addition, the effects of rock-physics models and approximations of one-dimensional reservoir models should be considered.

Finally, we applied the developed Bayesian model to real-life data sets collected from the Troll site. Although seismic waveforms and rock-physics models are estimated from borehole logs with uncertainty, and both seismic AVA and CSEM data are three-dimensional data, a comparison between the estimated results using seismic AVA data only and the estimated results using both seismic AVA and CSEM data shows that joint inversion of seismic and EM data provides better estimation of gas saturation.

We notice that the benefits of combining seismic AVA and CSEM data are more striking in synthetic tests than in the field-data example presented here. Part of the difference is almost certainly a result of the large number of unknown noise sources inherent in the field data. These may include noise in the estimated angle-dependent wavelets and the possible presence of correlated (non-Gaussian) noise in both seismic AVA and CSEM data sets. The saturation and porosity logs themselves, assumed to be ground truth, can also be in error. In addition, the one-dimensional model may not accurately represent the actual earth. This is more likely to be a problem for the CSEM data, which has a larger spatial footprint, than it is for the seismic AVA modeling, although the assumption that all multiples have been removed and that true relative amplitudes have been recovered in the seismic data may also not be strictly valid. In any case, we believe that our synthetic examples provide sufficient evidence of the possible improvements when seismic AVA and CSEM data are combined so as to induce others to improve on our efforts with field data.

ACKNOWLEDGMENTS

The work is funded by the Research Partnership to Secure Energy for America (RPSEA) and the U.S. DOE under Contract No. DE-AC03-76SF00098.

REFERENCES

Archie, G. E., 1942, The electrical resistivity log as an aid in determining some reservoir characteristics: Trans. Am. Inst. Mech. Eng., **146**, 54-62.

Bernardo, J. M. and Smith, A. F., 1994, Bayesian theory: John Wily and Sons, New York.

Bosch, M., 1999, Lithologic tomography: from plural geophysical data to lithology estimation: Journal of Geophysical Research, **104**, 749-766.

Buland, A., and Omre, H., 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185-198.

Castagna, J. P., AVO Analysis – Tutorial and Review, 1993, in Offset-Dependent Reflectivity – Theory and Practice of AVO Analysis., J. P. Castagna and M. M. Backus eds, Investigations in Geophysics no. 8, Soc. of Exp. Geophysics, pp. 3-36.

Debski, W. and Tarantola, A., 1995, Information on elastic parameters obtained from the amplitudes of reflected waves: Geophysics, **60**, 1426-1436.

Drufuca, G., and Mazzotti, A., 1995, Ambiguities in AVO inversion of reflections from a gas sand: Geophysics, **60**, 134-141.

Dvorkin, J., and Nur, A., 1996, Elasticity of high-porosity sandstones: Theory for two North Sea datasets: Geophysics, **61**, 1363-1370.

Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M.C., MacGregor, L. M., Constable, S., 2002, Remote sensing of hydrocarbon layers by seabed logging (SBL): Results from a cruise offshore Angola: The Leading Edge, **21**, 972-982.

Gassmann, F., 1951, Elastic waves through a packing of spheres: Geophysics, **16**, 673-685.

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

Hoversten, G., Cassassuce, F., Gasperikova, E., Newman, G., Chen, J., Vasco, D., Rubin, Y., and Hou, Z., 2006, Direct reservoir parameter estimation using joint inversion of marine seismic AVA and CSEM data: Geophysics, in press.

Landro, M., 2001, Discrimination between pressure and fluid saturation changes from time-lapse seismic data: Geophysics Soc. of Expl. Geophys., **66,** 836-844.

Lumley, D., Meadows, M., Cole, S. and Adams, D., 2003, Estimation of reservoir pressure and saturations by cross-plot inversion of 4D seismic attributes: 73rd Ann. Internat. Mtg., Soc. of Expl. Geophys., 1513-1516.

Maliverno, A., 2002, Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem: Geophysical Journal International, **151**, 675-688.

Mavko, G., Mukerji, T. and Dvorkin, J., 1998, The rock physics handbook: Tools for seismic analysis in porous media, Cambridge University Press, New York, USA.

Mosegaard, K. and Tarantola, A., 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research, **100**, 12431-12447.

Neal, R. M., 2003, Slice sampling: The Annuals Statistics, **31**, 705-767.

Plessix, R. and Bork, J., 2000, Quantitative estimation of VTI parameters from AVA responses: Geophysical Prospecting, **48**, 87-108.

Shuey, R. T., 1985, A simplification of the Zoeppritz Equations: Geophysics, **50**, 609-614.

Tura, A. C. and Lumley, D., 1999, Estimating pressure and saturation changes from time-lapse AVO data: 69th Ann. Internat. Mtg: Soc. of Expl. Geophys., 1655-1658.

TABLES

TABLE 1. Rock-physics models derived using data collected from borehole 31/2-1

Parameters	Values
Critical porosity	0.38
Number of grain contacts	9.0
Grain shear modulus (Gpa)	22.5
Grain Poisson ratio	0.349
Grain density (kg/m^3)	2567.3
Oil API gravity	28.5
Brine salinity	0.007
Gas G ratio to air	0.59
Sg factor	1.0

TABLE 2. Archie's law coefficients obtained using data collected from borehole 31/2-1

Parameters	Fitted Values
Archie's Law constant	0.788
Water saturation exponent	-1.3091
Porosity exponent	-0.14429

TABLE 3. Root-mean-squares of the differences between the true values and the estimated values using seismic AVA data with signal-to-noise ratios from 6 to 1 and from 12 to 2

	SNRs	Estimated Means	Estimated Medians	Estimated Modes
Gas saturation	6 to 1	0.2351	0.2128	0.1717
	12 to 2	0.2149	0.1908	0.1613
Porosity	6 to 1	0.0026	0.0031	0.0033
	12 to 2	0.0021	0.0024	0.0028

TABLE 4. Root-mean-squares of the differences between the true values and the estimated values using both seismic AVA and EM data

	SNRs	Estimated Means	Estimated Medians	Estimated Modes
Gas saturation	6 to 1	0.1314	0.1097	0.0272
	12 to 2	0.0342	0.0306	0.0510
Porosity	6 to 1	0.0017	0.0018	0.0061
	12 to 2	0.0006	0.0008	0.0011

TABLE 5. Root-mean-squares of the differences between the true values and the estimated modes when the prior bounds of oil saturation are [0, 0.1], [0,0.3], and [0, 0.5]

	Oil bounds	Estimated Means	Estimated Medians	Estimated Modes
Water saturation	[0,0.1]	0.1095	0.0797	0.0421
	[0,0.3]	0.1766	0.1570	0.0787
	[0,0.5]	0.1898	0.1690	0.1020
Gas saturation	[0,0.1]	0.1088	0.0799	0.0313
	[0,0.3]	0.1805	0.1672	0.0957
	[0,0.5]	0.1963	0.1938	0.1873

TABLE 6. Root-mean-squares of the differences between the averaged well-log values and the estimated modes for Troll data sets

	Water Saturation	Gas Saturation	Oil Saturation	Porosity
Seismic AVA Data only	0.1877	0.1760	0.0965	0.0431
Seismic AVA and marine CSEM Data	0.1398	0.1650	0.1112	0.0442