APPLICATION OF NEW MONTE CARLO METHOD FOR INVERSION OF PRESTACK SEISMIC DATA

Yang Xue[1], Mrinal K. Sen[1] and Zhiwen Deng[2]

[1] Institute for Geophysics, The University of Texas at Austin
[2]BGP, CNPC

ABSTRACT

Global optimization methods such as very fast simulated annealing (VFSA) and a multiple VFSA (MVFSA) inversion have been applied to seismic waveform inversion and uncertainty characterization respectively. Here we address some of the limitations of MVFSA by developing a new stochastic method, named greedy annealed importance sampling (GAIS). GAIS combines very fast simulated annealing (VFSA) and greedy importance sampling (GIS), which uses a greedy search in the important regions located by VFSA in order to attain fast convergence and provide unbiased estimation. We demonstrate the performance of GAIS with application to pre-stack seismic waveform inversion of angle stack traces. Initial models containing the fractal behavior as the well logs are used to gain high resolution estimates of reservoir elastic properties. The results indicate that GAIS can provide better estimation of both models and their uncertainties than using VFSA alone. A gas layer can be easily identified using GAIS. The new MC method demonstrated here could also be applied to reservoir characterization and reservoir monitoring for accurate uncertainty estimation.

INTRODUCTION

The goal of exploration geophysics is to simulate a log of rock properties as a function of two-way vertical travel time or depth. The most accurate and high resolution logs can be generated from well measurements, however, only in vertical direction. Highly heterogeneous spatial distribution of reservoir elastic properties is not resolvable by seismic data: this causes multiple models to fit the measurements. A deterministic inversion based on minimization of the error between the observation and forward modeling only offers one of the best fit models and is usually band-limited. A complete solution should include both models and their uncertainties, which requires drawing samples from the posterior distribution. To assess this problem, stochastic inversion of seismic data with well logs is fairly common.

Conventional stochastic inversion methods based on Monte Carlo Markov Chain (MCMC) are either computationally intensive or provide biased estimation. The most common MCMC method is Metropolis-Hastings sampler (Metropolis and Ulam 1949, Hastings 1970), which generates random samples from a proposal distribution and rejects proposed moves based on Metropolis criterion. Although this method is proven to asymptotically converge to a stationary distribution, it is generally very slow. Very fast simulated annealing (VFSA) provides a fast

approximation of the expectation value. However, the uncertainty estimation by this method is biased because of continuous change of proposal distribution with iteration.

Schuurmans and Southey (2000) presented Greedy Importance sampling (GIS), which is a simple variation of importance sampling and shows an improved inference quality than other stochastic inference methods. They applied GIS for conducting Monte Carlo inference in a graphical model and proved that the technique yields unbiased estimates.

Based upon GIS, we further speed up the convergence of Monte Carlo inference by combining it with VFSA, which is named greedy annealed importance sampling (GAIS). MVFSA with models drawn from a prior distribution and with varying starting temperature are employed in order to locate the important regions of the model space which are further explored by GIS. GAIS is expected to provide an optimal balance between computational efficiency and accuracy of estimation. Furthermore, fractal initial model (Srivastava and Sen 2009, 2010) is applied to expand the frequency band of traditional seismic inversion results. The feasibility of GAIS is tested using prestack seismic from Hampson Russell Strata demo data.

METHODOLOGY

Here we first review the Bayesian formulation of a geophysical inverse problem and then discuss how to draw samples from a multi-dimensional posterior probability density (PPD) in model space (Tarantola 1987, Sen and Stoffa 1995) using different stochastic inversion methods.

Bayesian formulation

In Bayesian framework, the degree of belief can be updated by evidence. Represented in mathematical form, the PPD is proportional to the product of the prior and the likelihood:

$$p(\mathbf{m} \mid \mathbf{d_{obs}}) = \frac{p(\mathbf{m}) \cdot p(\mathbf{d_{obs}} \mid \mathbf{m})}{p(\mathbf{d_{obs}})}, \tag{1}$$

where $\mathbf{m}$ and $\mathbf{d_{obs}}$ represent model parameter and data vectors; $p(\mathbf{m})$ is a priori probability density function (pdf), $p(\mathbf{d_{obs}} \mid \mathbf{m})$ is the likelihood function and $p(\mathbf{m} \mid \mathbf{d_{obs}})$ is the conditional pdf of $\mathbf{m}$ given the data $\mathbf{d}$. Denominator $p(\mathbf{d_{obs}})$ is pdf of observation, or marginal evidence. It is a constant and independent on $\mathbf{m}$. Normally the likelihood function dominates the much larger subspaces of the model space than prior pdf. The choice of the likelihood functions depends on the noise distribution. Gaussian noise is the most common assumption for noise statistics. Therefore we normally have the likelihood function as shown in equation:

$$l(\mathbf{d_{obs}} \mid \mathbf{m}) \propto \exp(-E(\mathbf{m})), \tag{2}$$

where $E(\mathbf{m})$ is the error function given by

$$E(\mathbf{m}) = \left((\mathbf{d_{obs}} - g(\mathbf{m}))/2\right)^T \mathbf{C_D^{-1}} (\mathbf{d_{obs}} - g(\mathbf{m})), \tag{3}$$

where $g(\mathbf{m})$ is the forward modeling operator and $\mathbf{C_D}$ is called the data covariance matrix, which consists observation error and theory error. The marginal PPD of a particular model parameter, the posterior mean model and the covariance matrix are:

$$\sigma(m_i \mid \mathbf{d}_{obs}) = \int dm_1 ... \int dm_{i-1} \int dm_{i+1} ... \int dm_M \, \sigma(\mathbf{m} \mid \mathbf{d}_{obs}), \tag{4}$$

$$\langle \mathbf{m} \rangle = \int \mathbf{dm} \cdot \mathbf{m} \cdot p(\mathbf{m} \mid \mathbf{d_{obs}}), \tag{5}$$

$$\mathbf{C_M^{'}} = \int \mathbf{dm} \cdot (\mathbf{m} - \langle \mathbf{m} \rangle) \cdot (\mathbf{m} - \langle \mathbf{m} \rangle)^T \cdot p(\mathbf{m} \mid \mathbf{d_{obs}}). \tag{6}$$

All of the three equations can be expressed in the form of an integral

$$I = \int d\mathbf{m} f(\mathbf{m}) p(\mathbf{m} \mid \mathbf{d_{obs}}). \tag{7}$$

For multi-dimensional problems, evaluation of those integrals is computationally very expensive; some of the methods for evaluating these integrals are described below.

Simulated annealing

Simulated Annealing (SA) is a global optimization method drawing analogy between the model parameters and particles in an idealized physical system (Sen and Stoffa 1995). All the particles are distributed randomly in a liquid phase after being heated to a certain temperature. The crystallization, or the minimum energy state, occurs if annealing process follows a slow cooling schedule. Thermal equilibrium is required at each temperature with the probability :

$$P(E_i) = \frac{\exp(-E_i /(KT))}{\sum_{j \in S} \exp(-E_j /(KT))}, \tag{8}$$

where E is the energy function. In our application, the set S consists of all possible states (models) and K is Boltzmann's constant, which is set equal to 1 in our problem. The energy or the error function is given by

$$E(\mathbf{m}) = 1/2(\mathbf{d_{obs}} - g(\mathbf{m}))^T \mathbf{C_D^{-1}} (\mathbf{d_{obs}} - g(\mathbf{m})), \tag{9}$$

where $g(\mathbf{m})$ is the forward modeling operator and $\mathbf{C_D}$ is called the data covariance matrix, which consists observation error and theory error. We can rewrite the Gibbs distribution as PPD of model parameters:

$$P(m_i) = \frac{\exp(-E(m_i)/T)}{\sum_{j \in S} \exp(-E(m_j)/T)}. \tag{10}$$

.

Several modifications are employed by Ingber (1989) to improve the speed of SA, namely very fast simulated annealing (VFSA). Instead of building NM-dimensional Cauchy distribution, NM-product of one-dimensional Cauchy distribution is applied (Sen and Stoffa 1995).

The problem is: both SA and VFSA attempt to reach the global minimum using a temperature-ladder, which has the conflict with the requirement of MCMC that all the states stay at the same temperature. To be more specific, the proposal distribution becomes sharper and sharper along the cooling schedule defined by VFSA, which bring biasness to the estimation due to the short tail (Sen and Stoffa 1996). The mean value of samples is not the true expectation value, although they are very close to each other. Furthermore, the variance of estimation is usually underestimated. Greedy importance sampling, described in the following section, avoids biased sampling by taking fixed steplength and using independent sampling method.

Greedy Importance Sampling

Importance sampling draws independent samples from a simpler "proposal" distribution *Q*, and then weighs these points by the ratio of posterior pdf *P* and prior pdf *Q* to obtain a "fair" representation of PPD. However, if *Q* misses high probability regions of *P*, the resulting estimator will have high variance because the sample will almost always contain unrepresentative points but is sometimes dominated by a few high weight points (Schuurmans and Southey 2000).

To overcome this problem, Schuurmans and Southey (2000) presented Greedy Importance sampling (GIS) based on a simple variation of importance sampling. Starting with independent sampling from a given prior distribution *Q*, GIS expands each individual sample to a block of points by explicitly searching for important regions of the target distribution *P*. Due to the trend based search algorithm, each sample block will contain at least one or two important points from the posterior distribution. The advantage is that even if *Q* misses high probability regions of *P*, the weighted samples from *Q* still be able to demonstrate a "fair" representation of *P*. The procedure of GIS is illustrated below (fig. 1):

- Step 1: draw samples $\mathbf{m}$ independently from Q.
- Step 2: For each m_i, let $m_{i,1} = m_i$.
  - Compute block $B_i = \{m_{i,1}, m_{i,2}, ..., m_{i,n}\}$ by taking local search steps in the direction of maximum $| f(m)P(m) |$ until a local maximum or $n-1$ steps.
  - Weigh each point $m_k \in B_i$ with a weight $w_i(m_k) = P(m_k)\alpha_{i,k} / Q(m_i)$, where $\alpha_{ik} = 1/(b_{i+1}b_{i+2}\cdots b_{i+k}(k+1)(k+2))$ with b_i an inward branching factor and $k \in [0, n-1]$.
- Step 3: Create the final sample from the blocks of points $m_{1,1}, ..., m_{1,n}, m_{2,1}, ..., m_{2,n}, ..., m_{q,1}, ..., m_{q,n}$.
- Step 4: For a random variable f, estimate $\mathrm{E}_{P(m)} f(m)$ using the formula $\hat{f} = \sum_{i=1}^{q}\sum_{k=1}^{n} f(m_{i,k})w_i(m_k)$.

Figure 1. Workflow of greedy importance sampling (from Shuurmans and Southey 2000).

Therefore we expect smaller uncertainty of the true estimation and faster convergence than a general importance sampling. Furthermore, compared with VFSA, the prior distribution of GIS is non-temperature dependent and can be as large as possible, which means a better quantification of uncertainty than VFSA.

The crucial issue is where to start the local greedy search initialed by $m_{i,1}$. Sampling from a uniform distribution is perhaps logical but it is time consuming to generate blocks with large samples considering forward modeling is required at each step. An efficient way is to start from a region near the global minimum error, which can be located by VFSA with a small number of iterations. We describe this approach in the next section.

Greedy Annealed Importance Sampling (GAIS)

GAIS avoids being trapped into a local minimum and enables faster convergence by employing VFSA before GIS. Uniformly drawn temperatures are used as initial temperatures, for multiple parallel VFSA. In a Bayesian framework, we can consider the temperature to be a hyper-parameter. Parallel VFSA starting from these temperatures with a small number of iterations (about 200) is employed to locate the regions near the global minimum error. Then blocks will be expanded according to the second step of GIS to explore important regions

Seismic Modeling

One of the goals of seismic inversion is to estimate subsurface impedance models (Zp, Zs). In order to improve the resolution of our estimation, we use fractal based initial models (Srivastava and Sen 2009, 2010), which have the same frequency band as that of the well log. Fatti's approximation (Fatti. *et. al.*1994) is employed to calculate the angle dependent reflectivity:

$$R_{\mathrm{pp}}(\theta) = (1+\tan^2(\theta))[\ln Z_p(i+1) - \ln Z_p(i)]/2 - 8\left(\frac{Z_s}{Z_p}\right)^2 \sin^2\theta \cdot [\ln Z_s(i+1) - \ln Z_s(i)]$$
$$-\left[\frac{\tan^2(\theta)}{2} - 2\left(\frac{Z_s}{Z_p}\right)^2 \sin^2\theta\right]\frac{\Delta\rho}{\rho}, \qquad (11)$$

where $Z_{p,s}$ are compressive and shear wave impedance, ρ is the density, and θ is the reflection angle.

Synthetic seismogram using the convolution of the wavelet and reflectivity are constructed. The error is evaluated using L_2 Norm of the misfit between forward modeling and seismic observation together with L_2 Norm of the misfit between velocity model and well log statistics.

EXAMPLE

Test with Hampson Russell Strata demo data

Here we demonstrate the feasibility of GAIS in seismic inversion based on a 2D line of PP prestack data set, given as a demonstration data in Hampson and Russell software, which is available in STRATA module. For the preliminary test and quality control, we employed GAIS inversion on the seismic traces located at the well location within an angle range of 3° to 24°. Before inversion, well tie, wavelet extraction and angle gather generation were carried out using Hampson Russell software package. The well log measurements do not contain shear wave velocity.

At first, 20 temperatures are randomly drawn from 1 to 100 and treated as initial temperatures for VFSA. 20% deviation from the well logs is employed as prior constraints for velocity models. Starting with each initial temperature, parallel VFSA with 200 iterations each is

applied to update Zp, Zs, ρ from fractal based initial models. Due to limitation of large angles, it is impossible to estimate reliable density, thus we do not update density models after this stage. Next we use the best fit modelss of (Zp, Zs) from each VFSA as starting model and expand those to sample blocks using the criterion defined in step 2 in the workflow of GIS (fig. 1). Specifically, at each grid point (Zp, Zs), we define four candidates with perturbations of $(\Delta Zp, \Delta Zs)$, $(\Delta Zp, -\Delta Zs)$, $(-\Delta Zp, \Delta Zs)$ and $(-\Delta Zp, -\Delta Zs)$, as the square's end demonstrated in fig. 2 (right) in blue circle. Only the one with minimize misfit (red circle in fig. 2 right) will be chosen for the next grid point. This is repeated 30 times with a fixed step length of $\Delta Zp = \Delta Zs = 50 m/s * g/cc$.

To summarize, as shown in fig. 2 (left), the middle point locates the best fit model after 200 iterations of VFSA and the green point locates the end model after 30 steps of greedy search. The magenta points indicate the direction of searching. In the end, all the samples are weighed and summed to estimate the expectation values of (Zp, Zs).

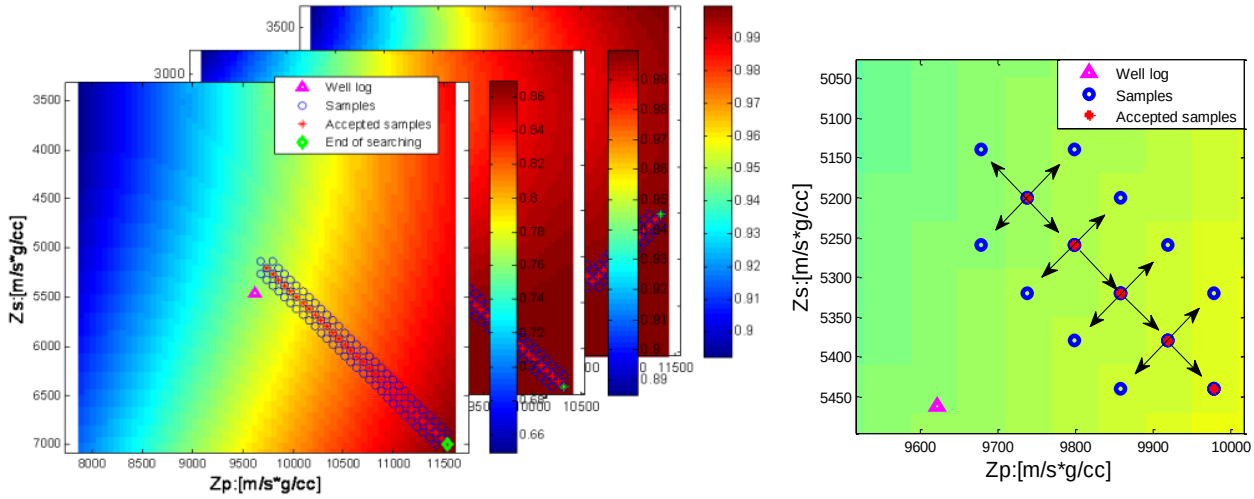


Figure 2. Three marginal probability maps of one layer at the well location (left) with each map corresponding to one greedy search after independent VFSA. Zoomed one marginal probability map (right).

RESULTS

Qualify control plot at the well location is shown in fig. 3 and fig. 4. The mean value of inverted P and S impedances as well as the variance of all accepted impedance samples derived from GAIS are compared with those derived from the best fit model after 1000 iterations of VFSA (fig. 3). The same fractal based initial models are used for both methods for comparison.

Based on the estimated impedance model, Fatti's approximation (Fatti. *et. al.*1994) is applied for forward modeling of reflectivity. Synthetic angle gathers using convolution of the reflectivity with the wavelet are compared with the real measurement (fig. 4).

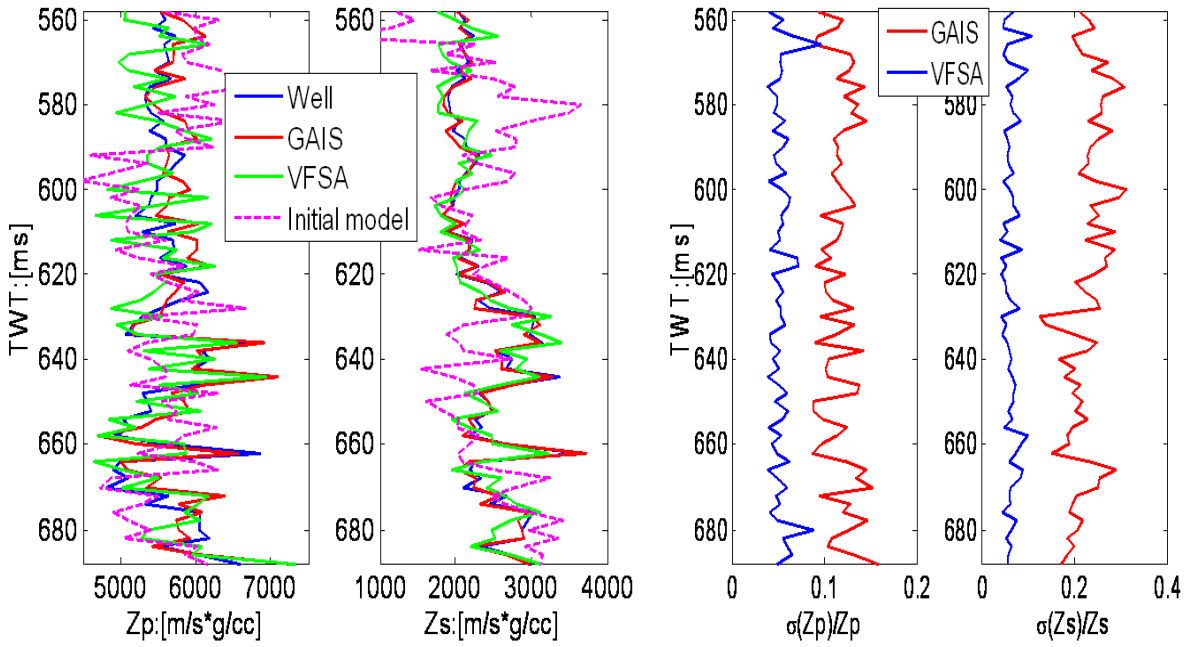
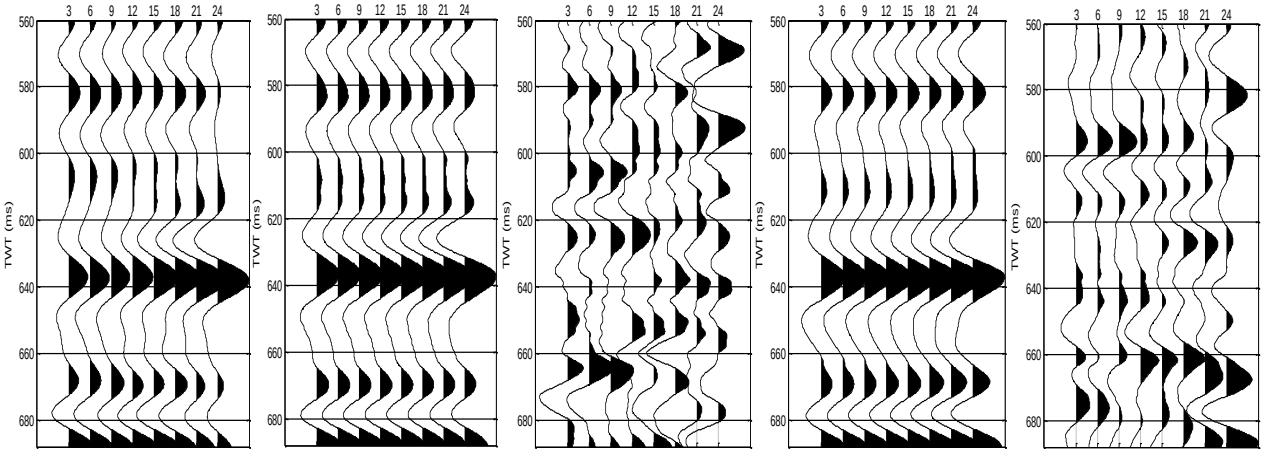


Figure 3. Left two columns: Impedance from well logs (blue), initial model (magenta-), the best fit model from VFSA (green-) with 1000 iterations and the mean model derived from 10 realizations of GAIS (red). Right two columns: relative variance of all samples generated from VFSA (blue) and GAIS (red).



(left 2) **(left 1)** **(middle)** **(right 1)** **(right 2)**

Figure 4. Seismic angle gathers of 3° to 24° from the measurements (left 2), synthetic angle gather generated from best fit model of VFSA with 1000 iterations (left 1) and their residuals (middle), synthetic angle gather derived from 10 realizations of GAIS (right 1) and their residuals (right 2).

After quality control at the well location, we further employ GAIS along the 2D line to run inversion trace by trace. Fig. 5, fig. 6 and fig. 7 show the inverted Zp, Zs and ZpZs ratio near the well location (CDP number 71) derived from GAIS and from the mean model of VFSA with 800 iterations.

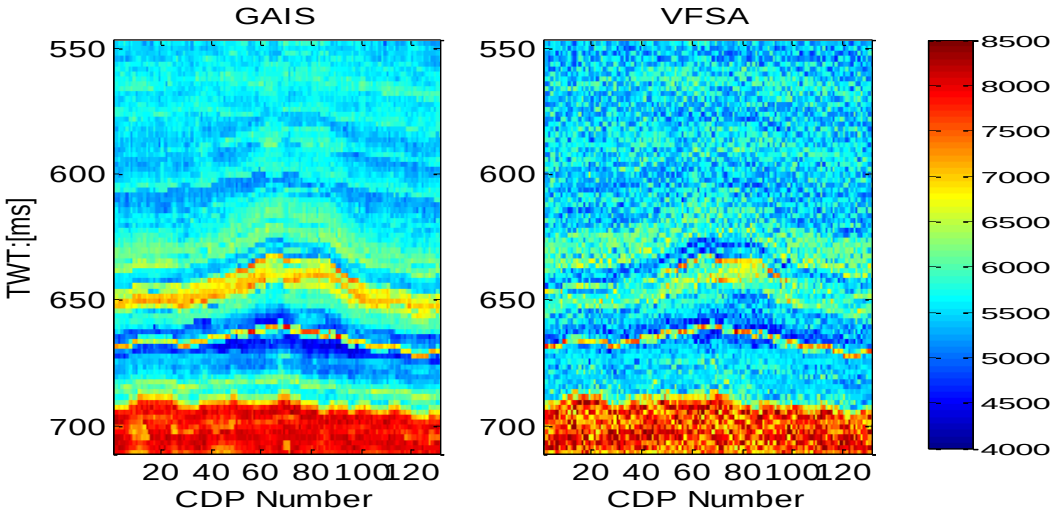


Figure 5. inverted Zp using GAIS (left) and compared with mean model of VFSA with 800 iterations (right).

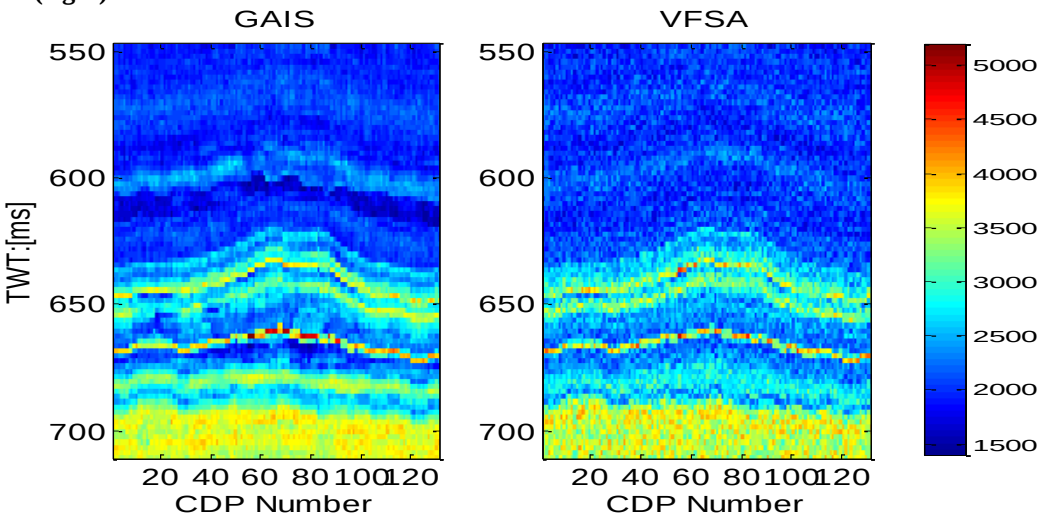


Figure 6. inverted Zs using GAIS (left) and compared with mean model of VFSA with 800 iterations (right).

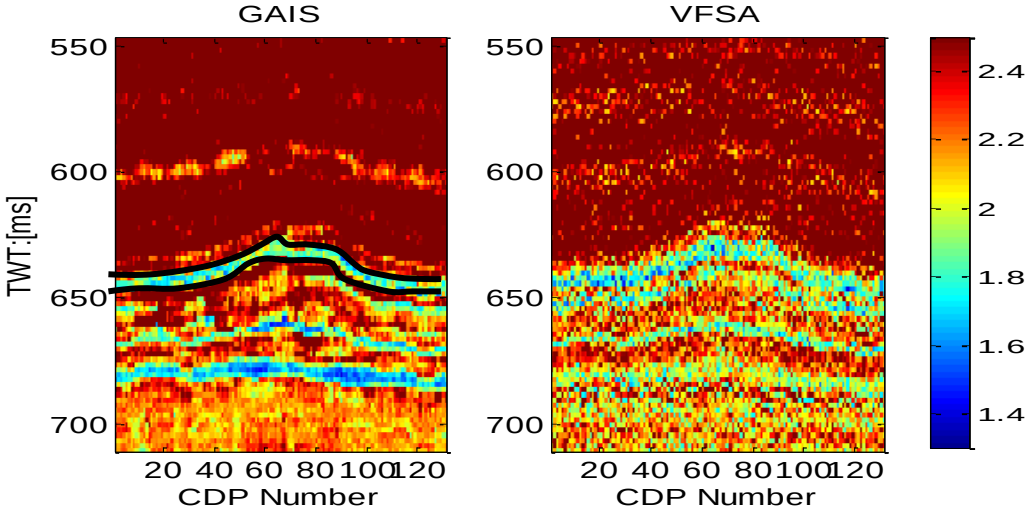


Fig. 7: ZpZs ratio derived from GAIS (left) and compared with the mean ZpZs value from VFSA with 800 iterations (right).

DISCUSSION

Based on our test using Hampson Russell Strata demo data, we note that the inverted P and S impedances from GAIS follow the well log trend better than the best fit model from 1000 iterations of VFSA (fig. 3). Furthermore, as we expected, the variance from VFSA with 1000 iterations is smaller than GAIS because VFSA is temperature dependent and the sampling is biased towards the best fit model, which results an underestimated variance (Sen and Stoffa 1995). GAIS addresses this problem by further expanding the sample space in the direction of important region with fixed small step length and assigning the weights based on the ratio between posterior distribution and prior distribution.

Inverted ZpZs ratio along a 2D seismic line (fig. 7) using GAIS shows more continuity than the mean model from VFSA with 800 iterations, which helps to identify the gas layer (marked using black curves) away from the well location more easily and accurately.

CONCLUSIONS

In this paper we investigated the applicability and accuracy of a newly developed GAIS algorithm to seismic inversion. GAIS starts to seek important regions starting with models that are close to the important regions already located by VFSA and estimate the expectation value very accurately. Furthermore, the blocks of samples generated using GIS around the global minimum error region provides reliable uncertainty of the estimation, which assess the problem of under estimated variance resulted from typical VFSA. Our test using Hampson Russell Strata demo data demonstrates a superior performance of GAIS than using VFSA alone.

ACKNOWLEDGMENTS

We thank EDGER forum of the University of Texas at Austin for supporting this research, Son Phan and Thomas Hess from Institute of Geophysics – University of Texas at Austin for helping with Hampson Russell software.

REFERENCES

Fatti, J. L., G. C. Smith, P. J. Vail, P. J. Strauss and P. R. Levitt, 1994, Detection of gas in sandstone reservoirs using AVO analysis: Geophysics, **59**,1362–1376.

Gassmann, F., 1951, Elastic waves through a packing of spheres, Geophysics **16**, 673–85.

Hastings, W. K., 1970, Monte Carlo methods using Markov chains and their applications, Biometrika, **57**, 97-109.

Ingber, L., 1989, Very fast simulated annealing, Mathematical Computer Modeling, **12**, 967-993.

Metropolis, N. and S. Ulam, 1949, The Monte Carlo method, J. Acous. Soc. Am., **44**, 335-341.

Schuurmans, D. and F. Southey, 2000, Monte Carlo inference via greedy importance sampling, in Proceedings UAI.

Sen, M. K. and P. L. Stoffa, 1995, Global optimization methods in geophysical inversion, Elsevier.

Sen, M. K. and P. L. Stoffa, 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion, Geophysical Prospecting, **44**, 313-350.

Srivastava, R. P. and M. K. Sen, 2009, Fractal-based stochastic inversion of poststack seismic data using very fast simulated annealing, Journal of Geophysics, **6**, 412-425.

Srivastava, R. P. and M. K. Sen, 2010, Stochastic inversion of prestack seismic data using fractal-based initial models, Geophysics, **75**, No. 3, R47-R59.

Tarantola, A. ,1987, Inverse Problem Theory. Elsevier Science.