

Electromagnetic Full Waveform Inversion based on Bayesian Markov-chain Monte-Carlo Method

Lei Yue, Jianhua Yue, Zhixin Liu, Xuemei Qi
School of Resources and Geosciences
China University of mining and technology
Xuzhou, China
yuelei@cumt.edu.cn

Knud Skou Cordua,
Niels Bohr Institute
University of Copenhagen
Copenhagen, Denmark

Abstract—Straight-ray-based inversion technique to estimate attenuation rate from electromagnetic tomography in coal mine has been available to geophysicists for over twenty years. This method gives a good computational efficiency but not a satisfy resolution. On account of the increasing computational power, accurate forward modeling can be included in advanced inversion approaches such that the full-waveform content can be exploited. Conventional full-waveform inversion methods are referred to as deterministic and are based on the minimization of an error term between the forward responses and the observed waveforms at each trace location. It is commonly used to give a ‘best estimate’ or ‘most likely’ case, regardless of the attendant uncertainties—which is the nature of most geophysical problems. Deterministic method is also easy to be trapped in a local minimum if there is not a good start model. To address this limitation, we present a probabilistic full-waveform inversion method in Bayesian formulation. In this formulation, solution to inverse problem is a probability density function refers to as posteriori distribution which describes all information available. Make use of Bayesian theorem combined with Markov-chain Monte-Carlo (MCMC) sampling, we can generate stochastic realizations from the posteriori distribution of model parameters. Bayesian-MCMC methods can incorporate any information that can be expressed in terms of probabilities and provide more precise model parameter even with an arbitrary initial model. In case study, we explore the performance of electromagnetic full-waveform inversion with MCMC through a simple synthetic tomographic example in coal mine, dielectric permittivity values of a moisture anomaly in coal seam can be obtained with a good resolution. Results demonstrate the feasibility of our statistical inversion method.

Keywords—radio image method; full waveform inversion; Bayesian; Monto-Carlo simulation; Markov-chain ;prior;posterior

I. INTRODUCTION

Safety production of coal mining operations could be greatly affected by geologic anomalies in coal seam. Since most of the geologic anomalies in coal seams effect the attenuation rate of electromagnetic signal that passes through them, the radio imaging method (RIM), using a signal in the kilohertz range, is capable of locating the zones of geologic anomalies in underground coal mines with the help of tomographic reconstruction programs. RIM is the most promising geophysical tool for identify and locate the geologic anomalies ahead of mining. Straight-ray-based inversion of

amplitudes of RIM data provides estimates of the signal attenuation distribution in coal seam region. This tomographic images obtained by ray-theory are limited in resolution because of straight-ray approximation and relatively small part of waveform information being included in the inversion.

To get a better indication of the anomaly in coal seam, waveform-type approaches should be considered. The first full-waveform inversion method appeared in Tarantola’s pioneering papers^[9], then it was widely used for seismic data but rarely for electromagnetic data. Ernst et al.^[7] firstly developed a full-waveform inversion scheme for crosshole GPR data. Using adjoint-gradient approach to minimize a misfit function between observed and modeled waveforms, sub-wavelength features can be resolved with few iteration times. This adjoint-based method providing a unique inversion result is referred as deterministic inversion. Algorithm convergence of this method largely depends on a good initial model guess obtained by ray-based program. The convergence process is a global optimization problem and therefore is always in risk of getting trapped in a local minimum.

Cordua et al.^[1] proposed a GPR crosshole full waveform inverse algorithm in Bayesian formulation to deal with the limitation above. In Bayesian framework, solution to inverse problem is described by a posteriori distribution provided by observed data and priori information. The posteriori realizations are solutions to the inverse problem that satisfies both the waveform data within their uncertainty and at the same time satisfies the priori information about the solution. Many different possible solutions are obtained by sample from the posteriori distribution, which indicate the uncertainty of the solutions. It is a probabilistic inversion method.

In this paper, we will adopt the latter method and investigate the performance of this probabilistic full-waveform inversion with MCMC in RIM.

II. METHODOLOGY

Consider that the coal seam can be described by a discrete set of model parameters m , and that a data set d , as measurements data. The model m describes some physical properties (i.e. dielectric permittivity, electrical conductivity) of the coal seam that influences the data observations. The forward relationship between the model parameters m and observations d is given as a finite-difference time-domain (FDTD) solution of Maxwell’s equations. The inverse problem

The Fundamental Research Funds for the Central Universities (China University of Mining and Technology) 2014QNB50

is to infer properties of m based on a set of observations d , and possibly some priori information about m .

RIM tomography is a typical under-determined, ill-posed inversion problem due to the coverage of observation system. That is to say, solution of this problem is not unique. In a Bayesian probabilistic formulation, the solution to this kind of inverse problem can be given as a posteriori probability density over the model parameters^[11]:

$$\sigma_M(m) = k \rho_M(m) L(m) \quad (1)$$

Where k is a normalization factor, always constant, $\rho_M(m)$ is the priori probability density function, and $L(m)$ is the likelihood function. $\rho_M(m)$ describes the probability that the model satisfies the priori information. The priori information is some expectation about the parameters of coal seam. $L(m)$ is a probabilistic measure of how well the modeled parameter explains the observed data. Therefore, $\sigma_M(m)$, the combined states of information provided by the data and the priori information, describes the probability that a certain model is a solution to the inverse problem. But the FDTD forward relation used in electromagnetic full waveform inversion is highly nonlinear; hence the posteriori distribution cannot be obtained analytically. That is to say there is not a closed-form mathematical expression of the posteriori distribution can be found. The Markov-chain Monte-Carlo (MCMC) algorithm makes it possible to sample from the posteriori distribution as the solution to the non-linear inverse problem. In MCMC method, a Markov chain is constructed over the state space of model parameters in order to generate samples from a target distribution (posteriori). Proper transition probability for the Markov-chain is chosen according to detailed balance equation. Then, with enough times of iteration, the Markov-chain is guaranteed to equilibrate to the target distribution no matter what the starting state is. The extended Metropolis algorithm is one of MCMC approaches developed by Hansen et al.^[6]. In this method, given an algorithm that can sample the priori distribution, the posteriori will be sampled as following:

- The priori sampler proposes a sample as initial model $m_{current}$, compute $L(m_{current})$;
- Perturb $m_{current}$, propose a new model $m_{propose}$;

- The proposals are stochastically accepted at each iteration with probability (Metropolis rule)

$$P_{accept} = \min \left(1, \frac{L(m_{propose})}{L(m_{accept})} \right) \quad (2)$$

- If the proposed model is accepted, $m_{current} = m_{propose}$, otherwise, $m_{current}$ is reused. This iterative procedure is repeated until a desirable number of models have been accepted.

In this procedure, the FFT Moving Average generator (FFTMA)^[7] algorithm serve as the priori sampler, which is

very efficient for generating unconditional realizations from a multivariate Gaussian model, which produces samples from a priori probability density defined by priori distribution- a bimodal Gaussian distribution discussed below (background and anomalies).

The likelihood function is defined below as a Gaussian distribution:

$$L(m) = c \prod_{k=1}^K \exp \left[-\frac{1}{2} (f(m)^k - d_{obs}^k)^T C_D^{-1} (f(m)^k - d_{obs}^k) \right] \quad (3)$$

where $f(m)^k$ and d_{obs}^k represent the simulated (through FDTD) and observed waveform traces related to the k th transmitter and receiver pairs. C is a normalization factor, K is the total trace numbers. C_D is the data covariance matrix.

$m_{propose}$ is obtained by perturbing $m_{current}$ using geostatistical algorithms as following:

A square subarea of $m_{current}$, with side-length step (exploration step size), is randomly selected in every perturbation and set to be unknown, then the values of this unknown area are re-drawn from a conditional probability distribution over these parameters conditional to the rest of the

model parameters, therefore a perturbed new model $m_{propose}$ is obtained. This procures are repeated multiple times in order to obtain a random walk in the space of model parameters. This sampling strategy was described in the paper by Hansen et al.^[5]. And it was named sequential Gibbs sampling due to its similarity to Gibbs sampling^[3]. Changing the number of parameters to be re-drawn in each step controls the step-length of the random walk. The size of step-length controls the exploration nature of the random walk. Larger step size leads to lower accept ratio (in equation 2). On the other hand, smaller step size gives a higher accept probability. But, accepted realizations of the Metropolis algorithm become statistically dependent with small step size. Hence, the size of the step-length is automatically adapted according to the accept ratio.^{[1][6]} Large step is used in the initial phase, and then gradually smaller and smaller step is used. At some point the step-size is fixed (see Fig.2 right). This is a demand for the algorithm to work correctly in the sampling part.

III. RESULT

We test the methodology outlined above on a synthetic tomographic RIM full-waveform data set. Wavefield simulations of the RIM signals (i.e., the forward relation) are obtained using FDTD calculations of Maxwell's equations in TE mode^[3].

Fig. 1(left) shows a synthetic reference image that simulates a matrix of coal seam with embedded collapse column (refer as moisture anomaly). Electromagnetic signals in coal seam waveguide are sensitive to the dielectric permittivity and the electrical conductivity of the materials. In this study we only

consider the influence of the dielectric permittivity, which is sensitive to moisture content. Dielectric permittivity of collapse column is often high compared to coal seam. Therefore, the dielectric permittivity of the collapse column is set to a relative dielectric permittivity of $\epsilon_r \approx 80$ and the coal seam is set to $\epsilon_r \approx 10$. The electrical conductivity is set to a constant value of 0.1 mS/m and is assumed known in the following discussion.

For the priori information, a bimodal distribution with increased probability of values around 10 and 80 is defined according to the reference model. Realistic priori information in field experiments can be obtained by field observations and geological expert knowledge. Fig. 1 (right) is an unconditional sample of the priori obtained using FFTMA.

A full waveform synthetic data set is calculated using the FDTD algorithm. Then, white noise is added to the synthetic data to simulate the observation waveforms. A Ricker wavelet with a central frequency of 1 MHz is used as source pulse. The source pulse is assumed known during the inversion. The transmitter and receiver positions are separated by 40 m and 10 m respectively. (See Fig. 1). Data acquired with a transmitter-receiver angle larger than 45 degrees from horizontal are omitted. This leads to a total of 116 data observations (i.e. recorded waveforms).

Realizations accepted by the Metropolis rule constitute the Markov-chain. The chain has two periods: one is burn-in period and then is stationary period. The end of the burn-in period can be located where the log-likelihood curve flattens out. In this case, burn-in was reached after approximately 5000 iterations. (see Fig.2 left)

In the stationary period, the chain converges to the target distribution; a sample from this distribution is then the solution to the inverse problem, which means after the Markov-chain reached a stationary state, realizations accepted by the Metropolis rule are representative samples of the posteriori distribution. It is not a single model being obtained as a solution, but many different possible solutions. This tells us

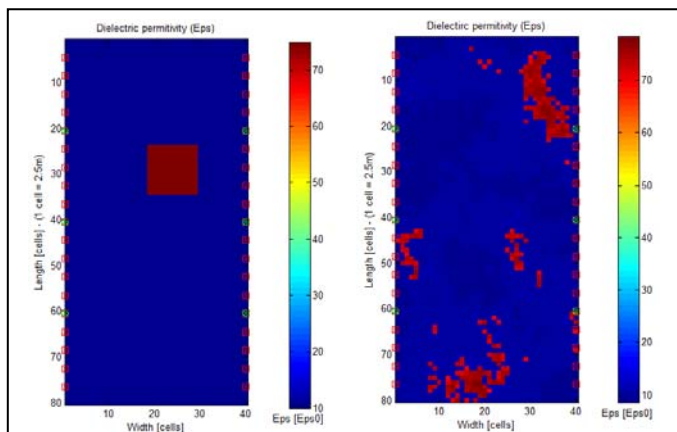


Fig. 1 (left) synthetic reference model which mimic coal seam with embedded collapse column. Green sign show transmitter positions and red sign show receiver positions. Right) The initial guess model used as input for the inversion.

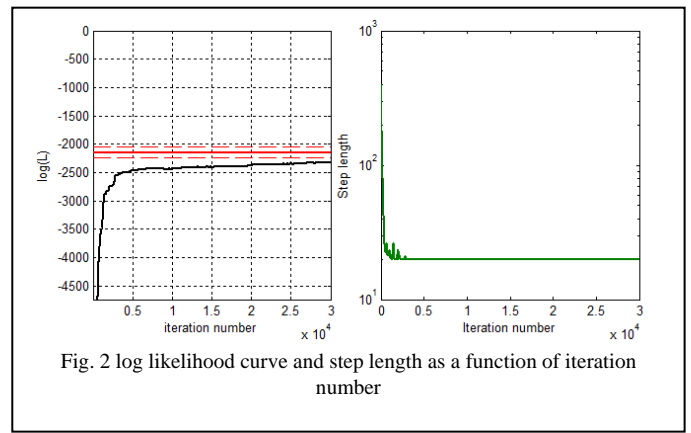


Fig. 2 log likelihood curve and step length as a function of iteration number

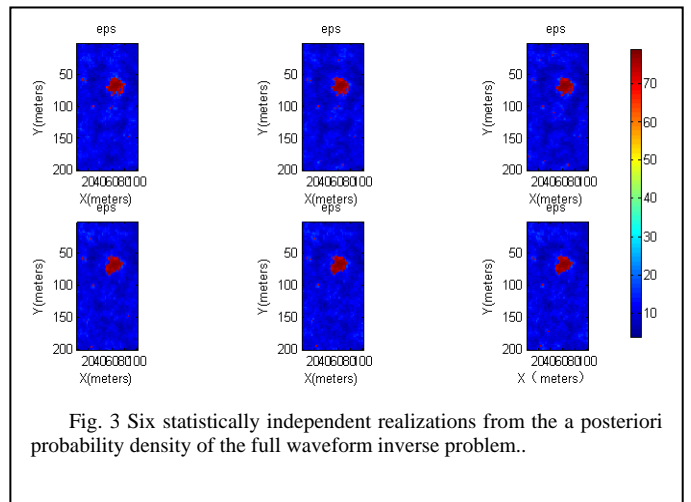


Fig. 3 Six statistically independent realizations from the a posteriori probability density of the full waveform inverse problem..

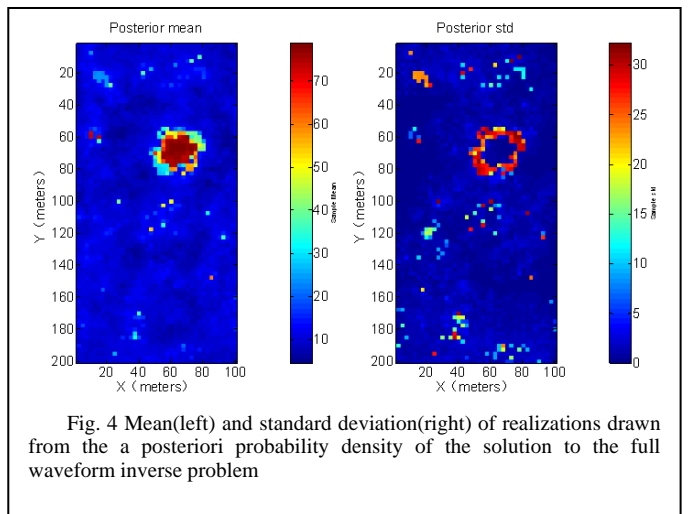


Fig. 4 Mean(left) and standard deviation(right) of realizations drawn from the a posteriori probability density of the solution to the full waveform inverse problem

about the uncertainty of the solution. The more different the individual realizations (i.e., solutions from the sampler) are, the worse is the resolution. On the other hand, if all solutions look almost the same, it means that we have a good resolution. Fig. 3 shows six accepted samples after burn-in period. Only a slight deviation between the individual samples is seen, which means that this inversion strategy gives a good resolution.

Fig. 4 shows the a posteriori mean and standard deviation based on samples from the a posteriori probability density during the stationary period. From the posteriori standard deviation it shows that the overall structure of the model is

recovered (standard deviation close to or equal to zero) whereas the higher standard deviation along the edges between collapse column and coal seam. A comparison between the reference model (Fig. 1 left) and the mean of the samples (Fig. 3 left) confirms that the full waveform inversion is able to recover collapse column structures well. Moreover, it should be noted that the algorithm is able to reach these results even though it is started in a model very different with the reference model (compare Fig. 1 right and left).

IV. CONCLUSION

We have demonstrated the potential of tomographic full waveform inversion together with MCMC sampling. It is a combination of geophysical inverse problems and geostatistics. This probabilistic methodology provides a means of evaluating the posteriori uncertainty, which is not provided using deterministic full waveform inversion. Moreover, the present approach is robust even with an arbitrarily initial guess of the solution. Finally, the challenge of this algorithm is reducing the computational expense of the Monte Carlo strategy.

ACKNOWLEDGMENTS

SIPPI was used for sampling the posteriori distribution while mGstat was used for geostatistical simulation; special thanks will be given to ETH for sharing FDTD code.

REFERENCES

[1] Cordua, K.S., Hansen, T.M. & Mosegaard, K., 2012. Monte Carlo full waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information, *Geophysics*, 77(2), H19–H31.

[2] Ernst, J. R., A. G. Green, H. Maurer, and K. Holliger, 2007a, Application of a new 2D time-domain fullwaveform inversion scheme to crosshole radar data: *Geophysics*, 72, J53 – J64.

[3] Ernst, J.R., H. Maurer, A. G. Green, and K. Holliger, 2007b, Full-waveform inversion of crosshole radar data based on 2- D finite-difference time-domain solutions of maxwell's equations: *IEEE Transactions on Geoscience and Remote Sensing*, 45, 2807 – 2828.

[4] Hansen, T.M., K. Mosegaard, and K. S. Cordua, 2008a. Using geostatistics to describe complex a priori information for inverse problems. In: J. M. Ortiz, and X. Emery, eds., *Geostatistics 2008: Chile*, vol. 1.

[5] Hansen, T. M., K. C. Cordua, M. C. Looms, and K. Mosegaard (2013a), SIPPI: a Matlab toolbox for sampling the solution to inverse problems with complex prior information: Part 1 - Methodology, *Comput. Geosci.*, 52, 470–480, doi:10.1016/j.cageo.2012.09.004.

[6] Hansen, T. M., K. C. Cordua, M. C. Looms, and K. Mosegaard (2013b), SIPPI: a Matlab toolbox for sampling the solution to inverse problems with complex prior information Part 2 - Application to crosshole GPR tomography, *Comput. Geosci.*, 52, 481–492, doi:10.1016/j.cageo.2012.10.001.

[7] Le Ravalec, M., Noetinger, B., Hu, L.Y., 2000. The FFT moving average (FFT-MA) generator: an efficient numerical method for generating and conditioning Gaussian simulations. *Mathematical Geology* 32 (6), 701–723.

[8] Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point statistics. *Math Geol* 34:1–21.

[9] Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: *Geophysics*, 51, 1893–1903, doi: 10.1190/1.1442046.

[10] Tarantola, A., 1988, Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation: *Pure and Applied Geophysics*, 128, 1-2, 365–399, doi: 10.1007/BF01772605.

[11] Tarantola, A., 2005, *Inverse problem theory and methods for model parameter estimation*: Society of Industrial and Applied Mathematics.