An adaptive reversible jump MCMC inversion algorithm for airborne time-domain electromagnetic data

Xiaodong Yu¹, Xuben Wang², Congde Lu², and Yuhang Wang³

¹ School of Information Science and Engineering, Chengdu University, Chengdu, Sichuan, China
² College of Geophysical, Chengdu University of Technology, Chengdu, Sichuan, China
³ Earthquake Administration of Sichuan Province, Chengdu, Sichuan, China
Email: yuxiaodong@cdu.edu.cn

Abstract. The reversible jump Markov Chain Monte Carlo algorithm is a trans-dimensional Bayesian inversion method, which can not only obtain the best inversion solution, but also provide the uncertainty information of inversion parameters, so as to effectively evaluate the reliability of inversion results. However, in trans-dimensional Bayesian inversion, the step size of model sampling has a great influence on the efficiency and precision of inversion. To solve the selection problem of sampling step size, this paper presents an adaptive reversible jump MCMC inversion algorithm for airborne time-domain electromagnetic data. In this method, the sampling step length is automatically adjusted in the sampling process through the data fitting error of each sampling model, so as to avoid the influence of unreasonable sampling step size and improve the acceptance rate of model sampling. The validity of the proposed method is verified by inversion test on the synthetic data and compared with the regularized inversion results. The inversion results demonstrate that the adaptive reversible jump MCMC inversion algorithm can improve the acceptance rate of model sampling and the accuracy of inversion results.

1. Introduction

For the inversion of airborne time-domain electromagnetic (ATEM) data, most existing inversion methods are optimization methods based on gradient descent technology, which generally rely on regularization strategy to construct the objective function, for instance, laterally constrained inversion [1], spatially constrained inversion [2-3], holistic inversion [4]. These methods can only provide a single best smooth model as a result, which does not provide the uncertainty information of model parameters. In order to obtain the uncertainty information on the inversion parameters, and evaluate the non-uniqueness and reliability of inversion results, the inversion methods based on Bayesian framework are proposed, which regard both the observed data and model parameters as random variables. Bayesian inversion method combines the prior probability of the model parameters and likelihood function (data misfit), and applies the Bayesian formula to calculate the posterior probability distribution (PPD) of the model parameters, which can not only obtain high precision inversion results but also quantitatively evaluate the uncertainty information of the inversion results.

Reversible jump Markov Chain Monte Carlo (MCMC) algorithm proposed by Green (1995) is a highly efficient algorithm for calculating the PPD of Bayesian inversion, which treat the number of unknown parameters as an inversion parameter. Malinverno (2002) applied the reversible-jump MCMC method to the inversion of one-dimensional DC resistivity layered media for the first time in the field.
of electromagnetic data inversion. Then, this method has also been applied to marine controlled source electromagnetic [5], airborne electromagnetic [6-7] and magnetotelluric [8].

For the reversible jump MCMC algorithm, the sampling step size will influence the acceptance rate and the sampling results. If the step size is too large, the acceptance rate is too low and lots of samples are being rejected as they land in low probability areas. If the step size is too small, the acceptance rate will be very high and lots of samples are accepted. This makes the algorithm not exploring the model space enough and will be slow to converge upon the PPD [9]. Therefore, how to choose the step size of model sampling is a problem that must be considered to improve the efficiency and accuracy of the reversible jump MCMC algorithm. In this paper, we present an adaptive reversible jump MCMC algorithm with sampling step adjustment automatically for the inversion airborne time-domain electromagnetic data.

2. Inversion Algorithm

According to Bayes’s theorem, the geophysical inversion problem is to estimate the posterior probability distribution of model parameters $p(m|d)$ using the observed data $d$ and the prior model parameters distribution $p(m)$

$$p(m|d) = \frac{p(m)p(d|m)}{p(d)} \quad (1)$$

where $p(d|m)$ is the likelihood function which is the probability of the data $d$ given the model $m$, its functional form depends on the statistic of the noise distribution. The evidence term $p(d)$ is an overall normalizing constant which ensures that $p(m|d)$ is properly normalized as a probability distribution on the $m$. Consequently it can be ignored [10].

It is difficult to obtain an analytical expression of the PPD for non-linear geophysical inversion problems, and it is not possible to exhaustively sample the model space for more than a few parameters, hence efficient sampling methods like the reversible jump MCMC method are usually used in practice. In the reversible jump MCMC method, the number of unknown model parameters $k$ is uncertain and is also considered as an inversion parameter. For the layered media, $k$ can represent the number of layers of the model. In this paper, we also take the height of transmitter and estimation data error as inversion parameters, thus the PPD for ATEM data can be expanded as [11]

$$p(m|d) \propto p(d|k,z,\sigma, h^e, d^e) p(k) p(z|k) p(\sigma|k,z) p(h^e) p(d^e) \quad (2)$$

where $p(d|k,z,\sigma, h^e, d^e)$ is the likelihood function, which is a measure of data fit. $p(k) , p(z|k) , p(\sigma|k,z) , p(h^e)$ and $p(d^e)$ describe the prior distribution of the unknown inversion parameters, namely, the number of layers for the model $k$, the interface depths $z$, the conductivities of each layer $\sigma$, the height of transmitter $h^e$ and the estimation data error $d^e$. Usually, the prior model has a simple distribution, such as uniform distribution or multivariate normal distribution.

In order to reduce the impact of sampling step size on the results, in this paper, we proposed an adjustment automatically method for sampling step through the change trend of data fitting error for each sampling model. The process of automatically adjusting the sampling step size is as follow: firstly, the allowable change range of step size is specified. Then, if the new proposed model is accepted and its fitting error is smaller than the fitting error of the current model, the sampling step size needs to be adjusted, otherwise the current sampling step is maintained. Finally, if the sampling step size needs to be adjusted, the percentage $p$ of the fitting error reduction is calculated, and then a random is generated from the one-dimensional normal distribution with the current sampling step size $s$ as the mean, the current sampling step multiplied by $p$ as the variance. This random is taken as the new sampling step size, i.e. $N(s, p \times s)$.

3. Examples

In this section we present the inversion results of two sets of synthetic data to test the adaptive reversible jump MCMC algorithm. The first set of synthetic data is generated by a forward modeling of a three-
layer geologic model (Figure 1(a)). The conductivity of these three layers are 0.01S/m, 0.1S/m, 0.002S/m, respectively, and the thickness are 60m, 40m, and infinity, respectively. We first added 5% Gaussian white noise to the theoretical response data, and then used the conductivity-depth imaging (CDI) results as the initial states of the Markov chain. The experiment run five MCMC chains in parallel, each MCMC chain was run 1 million sampling models, when the average data misfit of the continuous 5000 sampling models is less than 50, the burn-in period ended.

Figure 1. Inversion results of three-layer geologic model (a) CDI and regularization inversion result, (b) Reversible jump MCMC inversion result

Figure 2. The probability distribution of model parameters. (a) interface depth, (b) conductivity, (c) the number of layers, (d) the height of transmitter, (e) the estimation data error

The inversion results in Figure 1(b) show that because the inversion data contains noise, although the lowest misfit model is a three-layer model, there is a certain difference between the thickness and conductivity of the second layer and the true model; the mean model and the median model (50th percentiles) in the inversion results are closer to the true model. Figure 2 shows the normalized probability distribution of the inversion parameters for the accepted models in the sampling results after the burn-in period. It can be seen that in the probability distribution of interface depth and conductivity, the location of the highest probability (the peaks) is very close to the real value, and the model with three layers in the sampling results accounts for a large proportion. Moreover, the normalized probability
distribution of the height of transmitter and the estimation data error indicate that the mean and the median results are relatively consistent and are very close to the true value.

In addition, table 1 provides a summary of the acceptance rate of the reversible jump MCMC algorithm with the sampling step fixed and adjusted automatically, respectively. It illustrates that the acceptance rate can be greatly improved by introducing a sampling step size which is adjusted automatically.

| Sampling mode          | Chain 1 | Chain 2 | Chain 3 | Chain 4 | Chain 5 | Average   |
|------------------------|---------|---------|---------|---------|---------|-----------|
| Step size is fixed     | 16.55%  | 17.13%  | 16.88%  | 16.03%  | 15.65%  | 16.448%   |
| Step size is adjusted  | 29.66%  | 30.28%  | 31.95%  | 28.87%  | 29.09%  | 29.97%    |

Then, we generate the synthetic data for a 2D conductivity model using the 2.5D modeling program “ArjunAir_705”. The data contains 81 soundings located on a flight line with 25 m space. We add 5% Gaussian random noise to the data. The model is shown in Figure 3, and taken as a 2D model with three layers, the thickness of the intermediate layer was 40 m, the conductivity of the intermediate layer was 0.1 S/m, the conductivities of the background layers were 0.005 S/m, the depth of the interface between the first and second layer was 60 m on the left and 80 m on the right side of 0 m. During the inversion of this set of data, the inversion parameters are consistent with those of the first set of data.

![Figure 3. The 2D conductivity model](image-url)

![Figure 4. Inversion results of 2D conductivity model](image-url)
Figure 4 shows the results of the adaptive reversible jump MCMC inversion algorithm and regularization inversion method (Yu et al., 2018). As can be seen from the inversion results of the reversible jump MCMC algorithm, they can better reflect the conductivity distribution of the real model, in which the mean value result is the most stable and closest to the real model. Compared with the regularization inversion results (Figure 4(f)), the inversion results obtained by the reversible jump MCMC algorithm with automatic step sample adjustment show clearer demarcation between different media layers. In addition, figure 5 and figure 6 show the results of the mean and median of the height of the transmitter and the estimation data error. It can be seen from the change curve that the difference between their median and mean values is very small, and they both jump near the true value. Although there is a certain difference between the true value and the inversion value, the difference is small, and most of them are smaller than the true value.

Therefore, using the adaptive reversible jump MCMC method to invert airborne time-domain electromagnetic data can obtain relatively good inversion results, and can better estimate some parameters other than model conductivity and depth. It provides an effective and reliable solution for inaccurate data error estimation or inaccurate height measurement in the actual observation data inversion.

Figure 5. Inversion results of the height of the transmitter for 2D conductivity model using the adaptive reversible jump MCMC algorithm

Figure 6. Inversion results of the estimation data error for 2D conductivity model using the adaptive reversible jump MCMC algorithm

4. Conclusions
We have presented an adaptive reversible jump MCMC algorithm with sampling step adjustment automatically for the inversion airborne time-domain electromagnetic data, and tested on two sets of synthetic data. The inversion results demonstrate that the adaptive reversible jump MCMC algorithm can obtain more information about inversion parameters and analyze the uncertainty of inversion results better. Besides, the model parameters estimated by the mean and median are relatively stable and reasonable, and are closer to the true model. Moreover, by introducing a sampling step size which is adjusted automatically, the acceptance rate of model sampling can be greatly improved, and the
The demarcation between different media layers is clearer than the inversion result obtained by the regularization method.

Acknowledgments
This research was supported by the National Key Research and Development Program of China (2017YFC0601806).

References
[1] Brodie R C and Sambridge M 2006 A holistic approach to inversion of frequency-domain airborne EM data Geophysics 71 G301-G312
[2] Green P J 1995 Reversible jump Markov chain Monte Carlo computation and Bayesian model determination Biometrika 82 711-732
[3] Malinverno A 2002 Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem Geophys. J. Int 151 675-688
[4] Minsley B J 2011 A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data Geophys. J. Int 187 252-272
[5] Mandolesi E and Ogaya X 2018 A reversible-jump Markov chain Monte Carlo algorithm for 1D inversion of magnetotelluric data Computers and Geosciences 113 94-105
[6] Ray A and Keyb K 2012 Bayesian inversion of marine CSEM data with a trans-dimensional self parametrizing algorithm Geophys. J. Int 191 1135-1151
[7] Siemon B, Auken E and Christiansen A V 2009 Laterally constrained inversion of helicopter-borne frequency-domain electromagnetic data Journal of Applied Geophysics 67 259-268
[8] Viezzoli A, Christiansen A V and Auken E 2008 Quasi-3D modeling of airborne TEM data by spatially constrained inversion Geophysics 73 F105-F113
[9] Yin C C, Qi Y F and Liu Y H 2014 Trans-dimensional Bayesian inversion of frequency-domain airborne EM data Chinese J. Geophys. (in Chinese) 57 2971-2980
[10] Yin C C, Zhu J and Qiu C K 2018 Spatially constrained inversion for airborne EM data using quasi-3D models Chinese J. Geophys. (in Chinese) 61 2537-2547
[11] Yu X D, Wang X B and Lu C D 2018 A combining regularization strategy for the inversion of airborne time-domain electromagnetic data Journal of Applied Geophysics 155 110-121