Block coordinate descent 3D inversion method for AEM data interpretation

Bo Zhang, Changchun Yin*, Kelin Qu, Yinfeng Wang, Luyuan Wang, Yunhe Liu, Xiuyan Ren, Yang Su
[1]College of Geoexploration Science and Technology, Jilin University, Changchun, China.
yinchangchun@jlu.edu.cn

SUMMARY

The airborne electromagnetic (AEM) method usually has large datasets, and the high demand for computing resources limits the application of three-dimensional (3D) AEM inversions. In the process of 3D AEM inversion, the step that consumes most computing resources is the solution of the large matrix equations. To avoid this disadvantage, the local mesh inversion method has been proposed, which significantly reduces the computational cost of 3D inversion. However, when the AEM survey area is very large, a large-scale inversion matrix equation will be generated during the inversion. Solving this large-scale matrix equations and storing the sensitivity matrix of all data points will consume a lot of memory. To solve this problem, we develop a 3D AEM data inversion method based on block coordinate descent technique (BCD). This method divides the large-scale inversion problem into many small inversion problems, avoiding the waste of computational resources in solving large equation systems. The experimental results indicate that the BCD method can obtain inverse models that are very similar to traditional inversions, but the memory cost is only one-third of the conventional 3D AEM inversion.

Keywords: 3D inversion, airborne electromagnetic (AEM), block coordinate descent

INTRODUCTION

The AEM exploration usually has a large dataset. Because of the high calculation resource requirement, in the past decades, the one-dimensional (1D) inversion has been widely used in AEM data interpretation (Ahl, 2003; Asif et al., 2022; Auken et al., 2005; Yin et al., 2007). The 1D inversion assumes that the model has a layered resistivity distribution which conflicts with the real earth. This assumption will bring large errors to the AEM data interpretation result. The three-dimensional (3D) inversion builds the model without any assumption which can accurately describe the earth's resistivity distribution. However, limited by the high calculation cost, the 3D inversion has not been widely applied to AEM data interpretation. To solve this problem, many 3D AEM data inversion acceleration methods have been developed. Among them, Cox et al (2007) proposed the local mesh inversion method that divides the numerical simulation for large model into many small models for numerical simulation. The local mesh method greatly decreases the calculation cost and makes the large model 3D AEM data inversion possible. Ren et al (2018) proposed the secondary field-based AEM inversion method that decreases the numerical simulation scale of each local mesh. Yang et al (2014) proposed an adaptive-sounding data selection inversion method that inverts only a part of the data during each inversion iteration. All of these works focused on decreasing the calculation of AEM responses and improved the speed of 3D AEM data simulation and sensitivity calculation. The AEM exploration usually has a large survey region which generates large-scale matrix equations in 3D AEM inversion. Solving this large-scale matrix equations and storing the sensitivity matrix for all data points will cost lots of memory. None of the previous researches on decreasing the scale of the model inversion equations. In this paper, we propose a 3D AEM data inversion method based on block coordinate descent technique (BCD). This method divides the large model inversion problem into many small inversions without generating any false anomaly on the sub-region boundaries but can greatly decrease the inversion memory cost. The synthetic data inversion demonstrates that the BCD method can obtain an inversion model very similar to that generated from traditional inversion. However, the memory cost is only one-third of the general 3D AEM data inversion.

METHODS

The 3D forward and adjoint forward modeling is the basis of the 3D AEM data inversion. The forward modelling calculates the AEM responses which determines the data misfit of the inversion model. The adjoint forward modeling calculates the sensitivity matrix which is the main component of the in an inversion process.

For the 3D forward modeling, we implement the unstructured finite-element method. The frequency-domain secondary electrical field $\mathbf{E}^s$ abides the

following double curl equation (Newman and Alumbaugh, 1995), i.e.

$$\nabla\times\nabla\times\mathbf{E}^s-\kappa\sigma\mathbf{E}^s=\kappa\sigma^s\mathbf{E}^p, \quad (1)$$

where $\mathbf{E}^p$ is the primary electrical field in the free air space, $\sigma^s=\sigma-\sigma^0$ is the abnormal conductivity, σ and σ^0 are the conductivity of the earth and the air, $\kappa=i\omega\mu$, $i=\sqrt{-1}$, ω is the angular velocity and μ is the permeability. Applying the Galerkin method to equation 1, we get the weak form of the double curl equation as

$$\int_\Omega(\nabla\times\mathbf{N}\cdot\nabla\times\mathbf{E}^s-\kappa\sigma\mathbf{N}\cdot\mathbf{E}^s)dV=\int_\Omega\kappa\sigma^s\mathbf{N}\cdot\mathbf{E}^p dV, \quad (2)$$

where $\mathbf{N}$ is the vector shape functions. The finite element method divides the calculation domain into small elements. By calculating the integral of equation 2 in each element, we can get a small matrix equation for each element. Assembling these small matrix equations, we can get a large linear matrix equations for AEM forward modeling as

$$\mathbf{K}\mathbf{e}^s=\mathbf{B}, \quad (3)$$

where the matrix $\mathbf{K}$ and vector $\mathbf{B}$ are the numerical integral of equation 2, $\mathbf{e}^s$ is the numerical solution vector of the electrical field. Solving equation 3, we can get the secondary field on each edge. Applying Faraday's law, we finally obtain the secondary magnetic fields.

For the Jacobean matrix calculation, we use the joint forward modeling method. The Jacobean matrix $\mathbf{G}$ can be defined as

$$\mathbf{G}_j=\boldsymbol{\psi}_j^{\mathrm{T}}\frac{\partial\mathbf{E}_j^s}{\partial\mathbf{m}}, \quad (4)$$

where $\boldsymbol{\Psi}$ is the interpolation matrix, $\mathbf{m}$ is the model vector, and the subscript j denotes the data point. From equation 1, we have

$$\mathbf{K}\frac{\partial\mathbf{E}^s}{\partial\mathbf{m}}+\frac{\partial\mathbf{K}}{\partial\mathbf{m}}\mathbf{E}^s=\frac{\partial\mathbf{B}}{\partial\mathbf{m}}. \quad (5)$$

Then, the Jacobean matrix becomes

$$\mathbf{G}_j=\boldsymbol{\psi}_j^{\mathrm{T}}\mathbf{K}^{-1}\frac{\partial\mathbf{K}_j^s}{\partial\mathbf{m}}\mathbf{E}_j. \quad (6)$$

Finally, the Jacobean matrix can be calculated by solving the joint forward modeling problem of

$$\mathbf{K}_i^{\mathrm{T}}\mathbf{V}^{\mathrm{T}}=\boldsymbol{\psi}_i \quad (7)$$

where $\mathbf{V}=\boldsymbol{\psi}_i^{\mathrm{T}}\mathbf{K}_i^{-1}$. Substituting equation 7 into equation 4, we can get the Jacobean matrix.

AEM data inversion is essentially to find a model whose responses match the observed data. This scheme can be implemented by searching for a model that makes the objective function minimum. The objective function can be defined as

$$\varphi=\left\|\mathbf{W}_d(\mathbf{d}_{obs}-\mathbf{d}_{for})\right\|_{L_2}+\lambda\left\|\mathbf{W}_m(\mathbf{m}-\mathbf{m}^{ref})\right\|_{L_2}, \quad (8)$$

where $\mathbf{d}_{obs}$ denotes the observed data, $\mathbf{d}_{for}$ is responses calculated from equation 3, λ is the regularization factor, $\mathbf{W}_d$ and $\mathbf{W}_m$ are the data and model covariance matrix, $\mathbf{m}^{ref}$ is the reference model. Making a Taylor expansion to the objective function and neglecting the high-order terms, we can get the equation for the Gauss-newton inversion, i.e.

$$\begin{aligned}&(\mathbf{G}^T\mathbf{W}_d^T\mathbf{W}_d\mathbf{G}+\lambda\mathbf{W}_m^T\mathbf{W}_m)\Delta\mathbf{m}=\\&\mathbf{G}^T\mathbf{W}_d^T\mathbf{W}_d\mathbf{r}^k-\lambda\mathbf{W}_m^T\mathbf{W}_m(\mathbf{m}^k-\mathbf{m}^{ref})\end{aligned}, \quad (9)$$

where $\mathbf{r}=\mathbf{d}_{obs}-\mathbf{d}_{for}$ is the data misfit vector, $\Delta\mathbf{m}$ is the model modification vector. Solving equation 9 and modifying the model using $\Delta\mathbf{m}$, we can get a new model. This process is repeated until the data fitting is less than a preset threshold.

Usually, the traditional inversion equation 9 has a very large scale. Solving this matrix equations and storing the Jacobean matrix for the entire survey area will be very costly. The BCD inversion renews the large model resistivity block by block and replace the solution of large linear matrix equation by many small matrix equations. This scheme can greatly save the memory of 3D AEM inversions. According to Liu et al. (2020), the numerical optimization problem:

$$\begin{aligned}&\min_{m\in M}F(m_1,\ m_2,\ \cdots\ ,m_S)=\\&f(m_1,\ m_2,\ \cdots\ ,m_S)+\sum_{i=1}^{S}r_i(m_i)\end{aligned} \quad (10)$$

can be solved by calculating many small problems of

$$\begin{aligned}&f_i^k(m_i)=\\&f(m_1^k,\ \cdots\ m_{i-1}^k,\ m_i^k,\ m_{i+1}^{k-1},\ \cdots\ m_S^{k-1})\end{aligned}, \quad (11)$$

where M is the feasible domain of the optimization problem, f is a differentiable function, r_i denotes a convex function, $m_1, m_2, \cdots, m_S$ denotes the sub-model of the entire parameters x, S denotes the number of sub-models. For the BCD inversion, equation 9 can be divided into the small Gauss-newton inversion equations of

$$\begin{aligned}&(\mathbf{G}_i^T\mathbf{W}_{d_i}^T\mathbf{W}_{d_i}\mathbf{G}_i+\lambda\mathbf{W}_{m_i}^T\mathbf{W}_{m_i})\Delta\mathbf{m}_i=\\&\mathbf{G}_i^T\mathbf{W}_{d_i}^T\mathbf{W}_{d_i}\mathbf{r}_i^k-\lambda_i\mathbf{W}_{m_i}^T\mathbf{W}_{m_i}(\mathbf{m}_i^T-\mathbf{m}_i^{ref})\end{aligned},\quad(12)$$

where the subscript i denotes the number of sub-models. Solving the matrix equation 12, we can get the model update block by block by

$$\mathbf{m}'=\begin{bmatrix}\mathbf{m}_1+s\Delta\mathbf{m}_1\\\vdots\\\mathbf{m}_i+s\Delta\mathbf{m}_i\\\vdots\\\mathbf{m}_S\end{bmatrix},\quad(13)$$

where s is the step length of the model modification. During the k^{th} iteration of the i^{th} sub-model, for the 1st to i-1$^{\text{th}}$ sub-model, we use the k^{th} updated resistivity, while for the i^{th} to S^{th} sub-model, we use the k-1$^{\text{th}}$ resistivity. This scheme treats the survey region as an entire model while solving the inversion matrix equation as small sub-models.

RESULTS

To demonstrate that our BCD inversion can get as good inversion results as the traditional method, we first test on a synthetic model and invert the synthetic data using both the traditional inversion and the BCD and compare the memory cost and inversion results. The synthetic model and inversion results are given in Figure 1. It is seen that the BCD inversion can get similar results as the traditional inversion. Both the BCD and traditional inversion methods can well recover the anomalous bodies buried under the topographic earth. However, the BCD inversion spends only 7.27GB of memory. In contrast, the traditional inversion spends 12.28 GB of memory. This indicates that the BCD inversion saves 40% of memory for the 3D AEM data inversion without generating any false anomaly.

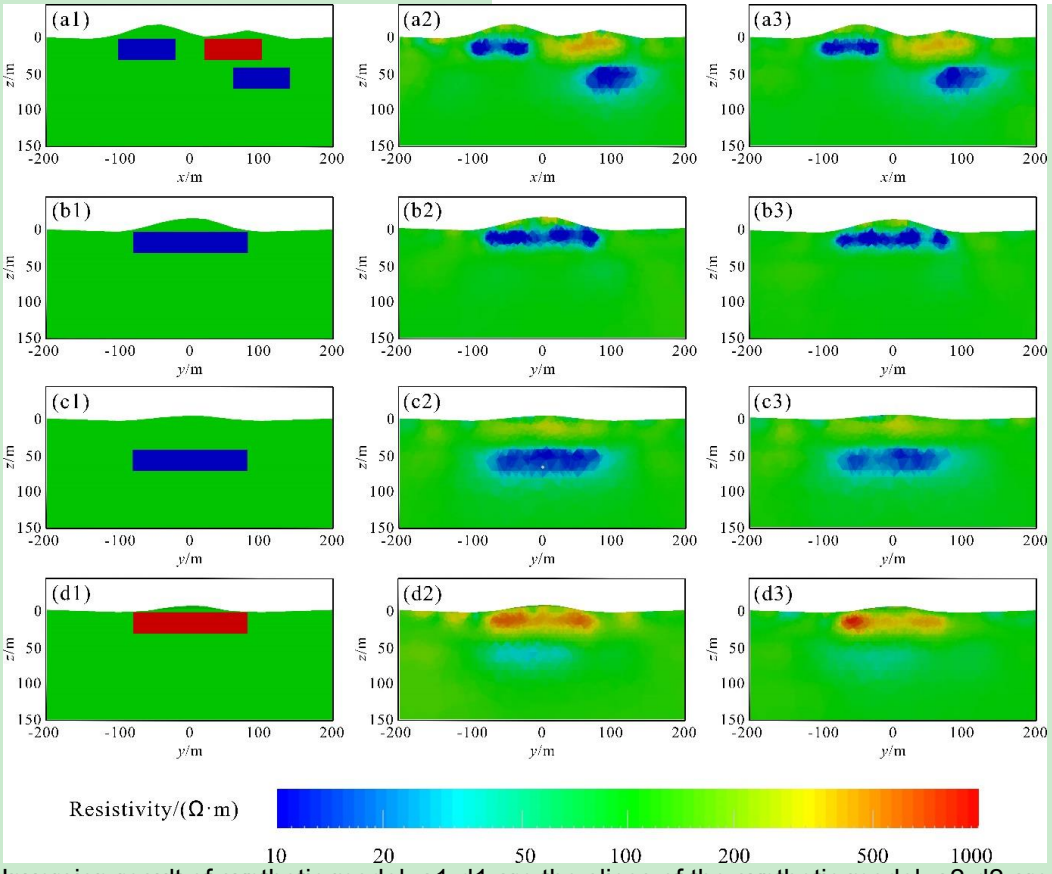

Figure 1. Inversion result of synthetic model. a1-d1 are the slices of the synthetic model; a2-d2 are the slices of the traditional inversions; a3-d3 are the slices of the BCD inversions. a1-a3 are the model slices at y=0m; b1-b3 are the model slices at x=-60m; c1-c3 are the model slices at x=40m; d1-d3 are the model slices at x=105m.

Conclusions

We have successfully developed a new 3D AEM data inversion scheme based on the BCD scheme. The BCD method inverts the survey data as an entire model but calculates the model modification vectors by solving the Gauss-Newton matrix equations in many small sub-regions. This avoids generation of false anomalies at the boundaries of each sub-region. The numerical example demonstrated that both the traditional and BCD inversion can well recover the topographic model with anomalous bodies embedded. However, the BCD inversion can save around 40% of memory compared with the traditional inversions. Limited by the length of the abstract, we only show one synthetic data inversion as example. More synthetic and field data inversions can be found in our future presentation.

Acknowledgements

This paper is financially supported by the National Natural Science Foundation of China (42174167, 42030806).

References

Cox L. and Zhdanov M. S. (2007) Large scale 3D inversion of HEM data using a moving footprint. Seg Technical Program Expanded Abstracts, San Antonio, Texas, 3124

Ren X., Yin C., Macnae J.et al. (2018) 3D time-domain airborne electromagnetic inversion based on secondary field finite-volume method. Geophysics, 83(4), 1-10

Yang D. K., Oldenburg D. W., Haber E. (2014) 3-D inversion of airborne electromagnetic data parallelized and accelerated by local mesh and adaptive soundings. Geophysical Journal International, 196(3), 1492-1507

Newman, G. A., and Alumbaugh D. L. (1995) Frequency-domain modeling of airborne electromagnetic responses using staggered finite differences. Geophysics, 43, 1021–1042

Liu H., Hu J., Li Y., Wen W., (2020) Optimization: modeling algorithm and theory, Higher Education Press (Chinese), pp. 367-378

Ahl A., (2003) Automatic 1D inversion of multifrequency airborne electromagnetic data with artificial neural networks: discussion and a case study. Geophysical Prospecting, 51(2), 89-98

Asif M. R., Bording T. S., Maurya P. K. et al. (2022) A Neural Network-Based Hybrid Framework for Least-Squares Inversion of Transient Electromagnetic Data. IEEE Transactions on Geoscience and Remote Sensing, 60, 1-10

Auken E., Christiansen A. V., Jacobsen L., and Sørensen K. I. (2005) Laterally Constrained 1D‐Inversion of 3D TEM Data. Symposium on the Application of Geophysics to Engineering and Environmental Problems 2005

Yin C. C., Greg H. (2007) Simulated annealing for airborne EM inversion. Geophysics, 72(4),189-195.