2-D MT gradient prediction with the transformer + Unet network: TM case

Kunpeng Wang[1], Chongxin Yuan[2] and Wei Luo[3]
[1]Key Laboratory of Earth Exploration and Information Technology of Ministry of Education, Chengdu University of Technology, Chengdu 610059, xfnwkp@163.com
[2]School of Big Data and Statistics, Sichuan Tourism University, Chengdu 610100, ycx@sctu.edu.cn
[3]Sichuan Natural Resources Investigation and Design Group Company Ltd., Chengdu 610051, luovitas@qq.com

SUMMARY

The calculation of magnetotelluric (MT) data gradients can be used for data sensitivity analysis, which is of great significance for actual sensitive areas. However, such calculations are highly complex and time consuming and therefore not conducive for the implementation of rapid analysis. To address this issue and improve computational efficiency, we propose a solution based on the transformer + Unet (T-Unet) neural network model to predicate 2-D MT gradient. Our research was conducted using TM polarization. First, we create a seven-channel dataset corresponding to the gradient label and then obtain the neural network weight model through network training and iteration and predict the gradient value rapidly and accurately on this basis. The experimental results indicate that, the T-Unet network not only significantly reduces computation time but also ensures high gradient prediction accuracy. This research demonstrates the potential of gradient fast prediction in sensitivity analysis and accelerated MT inversion.

Keywords: Gradient, MT, Transformer + Unet.

INTRODUCTION

There are two main methods to calculate sensitivity in commonly used regularized MT inversion. One is to solve the sensitivity matrix through, for example, the Occam (de and Constable 2004), Gauss–Newton method (Nádasi et al. 2022), and the other is to directly solve the objective function gradient vector to obtain the model modification direction through, for example, the non-linear conjugate gradient method (NLCG) (Newman and Alumbaugh 2000) or LBFGS (Lu et al. 2020). Among these methods, the inversion method based on gradient information is widely used. However, the gradient calculation is remains complex and time consuming. In this study, we propose a new method to predict gradient values using transformer-Unet (T-Unet) network, which combines the strengths of the transformers and U-Net architectures.

GRADIENT CALCULATION

In MT inversion, the objective function is usually given by

$$\varphi = (\mathbf{d}^{obs} - F(\mathbf{m}))^{\mathrm{T}} \mathbf{C}_{\mathbf{d}}^{-1} (\mathbf{d}^{obs} - F(\mathbf{m})) + \lambda (\mathbf{m} - \mathbf{m}_0)^{\mathrm{T}} \mathbf{C}_{\mathbf{m}}^{-1} (\mathbf{m} - \mathbf{m}_0) \quad (1)$$

where $\mathbf{d}^{obs}$ is the observation data (apparent resistivity and phase), F is the forward operator, λ is regularization parameter. **m** and $\mathbf{m}_0$ are the model parameters and a priori modes, respectively. $\mathbf{C_d}$ and $\mathbf{C_m}$ are the covariance matrix of the data and model, respectively.

In order to stabilize the inversion, the parameter m is transformed as follows (Egbert and Kelbert A 2012):

$$\tilde{\mathbf{m}} = \mathbf{C}_{\mathbf{m}}^{-\frac{1}{2}} (\mathbf{m} - \mathbf{m}_0) \quad (2)$$

Based on the inversion objective function (1), the gradient vector of the objective function is:

$$\nabla \varphi = -2 \mathbf{C}_{\mathbf{m}}^{\frac{1}{2}} \mathbf{J}^{\mathrm{T}} \mathbf{C}_{\mathbf{d}}^{-1} \left(\mathbf{d}^{\mathrm{obs}} - F(\mathbf{m}) \right) + 2\lambda \tilde{\mathbf{m}} \quad (3)$$

Where **J** is the sensitive matrix. In order to avoid direct solution of the sensitivity matrix, the general calculation scheme is to calculate $\mathbf{J}^{\mathrm{T}} \mathbf{C}_{\mathbf{d}}^{-1} \left(\mathbf{d}^{\mathrm{obs}} - F(\mathbf{m}) \right)$. According to (Newman and Alumbaugh 2000), this can be transformed into:

$$\mathbf{J}^{\mathrm{T}} \mathbf{C}_{\mathbf{d}}^{-1} \left(\mathbf{d}^{\mathrm{obs}} - F(\mathbf{m}) \right) = \mathbf{G}^{\mathrm{T}} \mathbf{A}^{-1} \mathbf{Q}^{\mathrm{T}} \mathbf{C}_{\mathbf{d}}^{-1} \left(\mathbf{d}^{\mathrm{obs}} - F(\mathbf{m}) \right) \quad (4)$$

Where, **A** is the forward discrete equation, **G** is the matrix related to the model parameters and **A**, and **Q** is the interpolation matrix (remaining unchanged in the calculation).

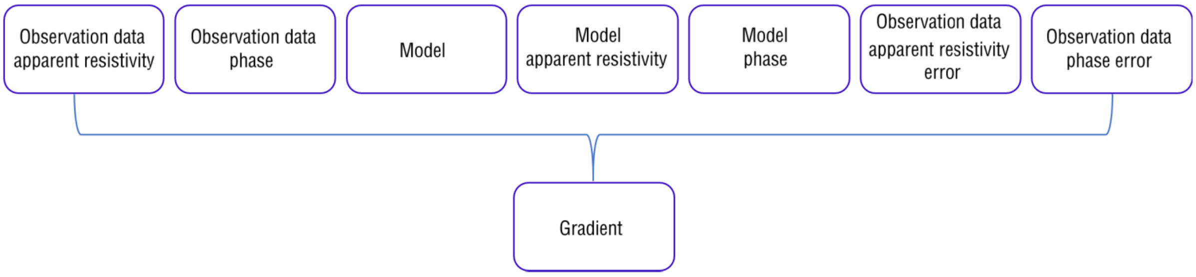

Figure 1. Related variables for predicting gradient

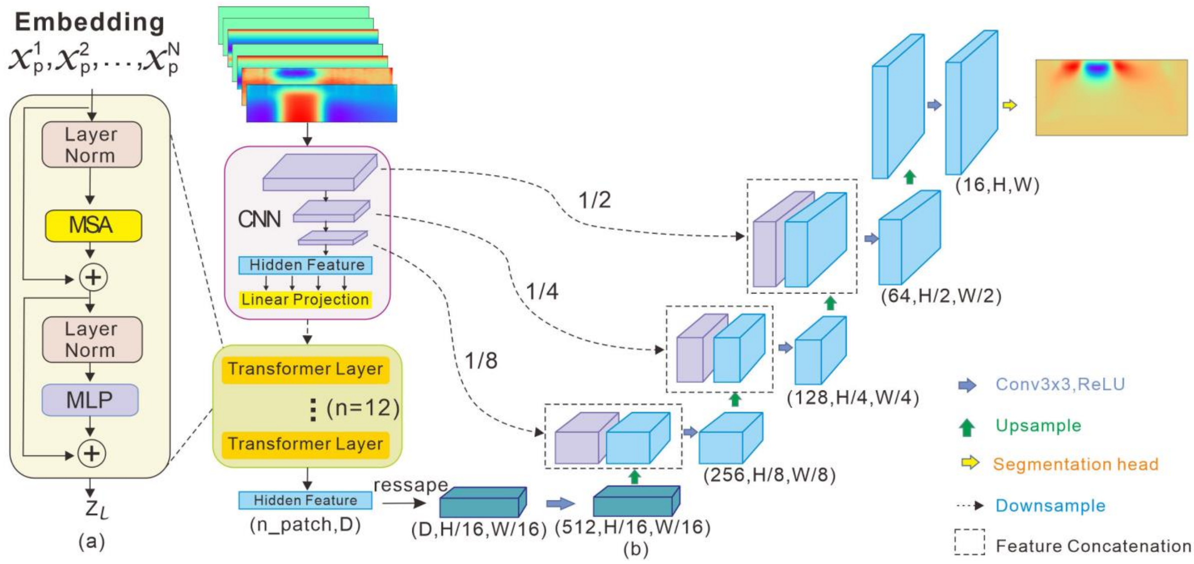

Figure 2. Overview of the framework. (a) Schematic of the transformer layer. (b) Architecture of the proposed T-Unet.

Deep Learning for TM Gradient Vector

Based on (3)–(4) and related descriptions, we know that the gradient calculation is related to: observation data (apparent resistivity and phase), model and corresponding forward modeling data (geoelectric model, which corresponds to apparent resistivity and phase), and observation data error (apparent resistivity error and phase error) (Fig. 1). For the task of calculating MT gradients, we choose the T-Unet architecture to train the network. Its structure is shown in Fig. 2.

In the training process stage, the inputs are observation data apparent resistivity, observation data phase, input forward model, model forward apparent resistivity, model forward model phase, observation data apparent resistivity error, and observation data phase error, i.e., a total of seven channel data. label: gradient value, so that the corresponding dataset can be created. Once the neural network is trained and meets predetermined conditions, we save the trained model for prediction. This trained DL neural network model captures the relationship between the seven channels of input data and corresponding gradient values.

During the prediction stage, we input the new seven-channel data into the pretrained neural network model to obtain the rapidly predicted gradient values from the network model.

Experiments and Results

In real applications, the distribution of subsurface resistivity may be more complex, and we cannot consider every case in the training dataset. Here, we generate test data for assessing the model's generalization capability based on more intricate combinations of basic shapes and layered structures found in the training dataset. Detailed information regarding the input types is as follows.

1) Two or three resistivity geoelectric models of different shapes and locations (background resistivity is 100 $\Omega \cdot \mathrm{m}$ and block anomaly resistivity is arbitrarily selected between 500–2000 $\Omega \cdot \mathrm{m}$; observation data (apparent resistivity and phase) contain 2% noise synthetic data; input model; input model corresponding apparent resistivity and phase; apparent resistivity error and phase error of observation data [Fig. 3(a)].

2) Three-layer geoelectric models of different shapes (different from G-Database, the resistivity changes greatly. The resistivity of the first layer varies between 50 and 100$\Omega \cdot \mathrm{m}$ and the middle and third layers are in select arbitrarily within the range of 600–1500 $\Omega \cdot \mathrm{m}$); observation data (apparent resistivity and phase) contain 2% noise

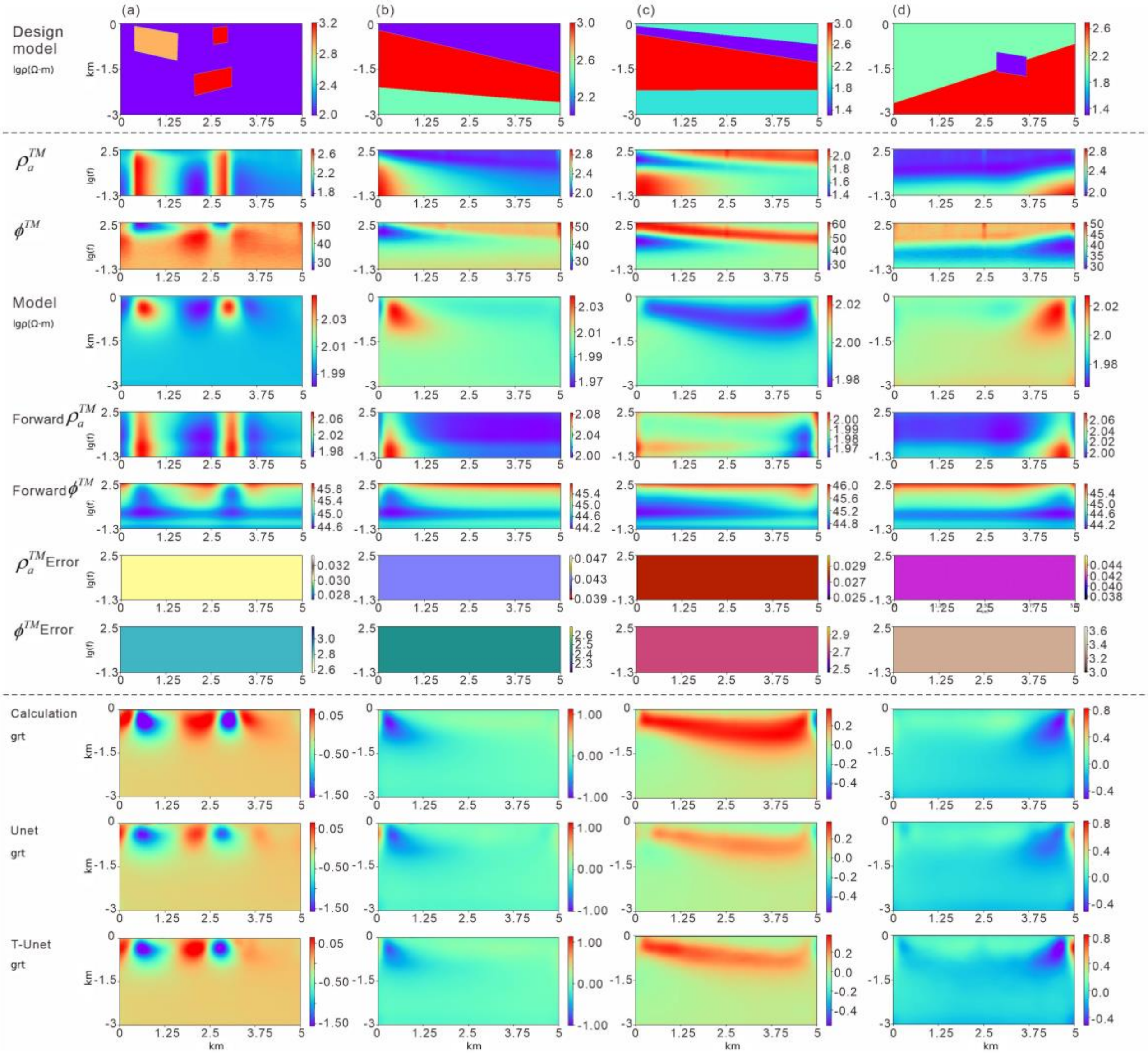

Figure 3. Comparison of T-Unet and Unet prediction results and calculated gradient results (a)–(d) are four different scenarios. The first row of models is designed to generate forward data. The second to eighth rows are seven-channel data used for gradient prediction, respectively. The ninth to eleventh rows are calculating gradients, Unet predicting gradients, and T-Unet predicting gradients, respectively. “grt” means gradient.

synthetic data; input the model; input the apparent resistivity and phase corresponding to the model; the apparent resistivity error and phase of the observation data error (the error size is random at 2%–7%) [Fig. 3(b)].

3) Four layers of geoelectric models with different shapes (the resistivity of the first layer varies between 80 and 100 $\Omega \cdot \mathrm{m}$, the resistivity of the second layer is arbitrarily selected within the range of 5–70 $\Omega \cdot \mathrm{m}$, the resistivity of the third layer ranges from 800 to 1500 $\Omega \cdot \mathrm{m}$ and the fourth layer can be selected arbitrarily within the range of 50–200 $\Omega \cdot \mathrm{m}$); observation data (apparent resistivity and phase) contain 2% noise synthetic data; input the model; input the apparent resistivity and phase corresponding to the model; apparent resistivity error and phase error of observation data [Fig. 3(c)].

4) Layered and block geoelectric models of different shapes and locations (the resistivity of each layer ranges from 80 to 1000 $\Omega \cdot \mathrm{m}$, block anomalies are embedded between layers and the resistivity ranges from 5 to 50 $\Omega \cdot \mathrm{m}$); observation data (apparent resistivity and phase) contain 2% noise synthetic data; input model; input model corresponding apparent resistivity and phase; apparent resistivity error and phase error of observation data [Fig. 3(d)].

Fig. 3 shows the prediction results of the T-Unet and Unet networks, as well as a detailed comparison with the calculated solutions. Regardless of how the observation data and input model change, it can be seen from the comparison results in the figure that the gradient predicted by T-Unet is relatively consistent with the calculation results.

CONCLUSIONS

In this study, we propose a new method to predict gradient values using T-Unet. We successfully trained the T-Unet model of gradient calculation. In our test, T-Unet took 2.45 s to calculate 200 gradient values, while the traditional gradient calculation of the CPU took 126.875 s. T-Unet significantly improves computational speed, surpassing traditional gradient solving methods.

We believe that our work can bring new possibilities for improving the efficiency of inversion.

ACKNOWLEDGEMENTS

This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFB3905004, in part by the China Postdoctoral Science Foundation under Grant 2022M723684, in part by the Sichuan Science and Technology Program under Grant 2022JDRC0066, in part by the Sichuan-Tibet Synergism and Innovation Center of Tourism Industry Competitiveness under Grant 71-000316, and in part by the Financial Research Foundation of Sichuan Province under Grant 51000024Y000010956258.

REFERENCES

de Groot-Hedlin C, Constable S (2004) Inversion of magnetotelluric data for 2D structure with sharp resistivity contrasts. Geophysics 69: (1), 78-86.

Nádasi E, Gribenko AV, Zhdanov MS (2022) Large-scale inversion of magnetotelluric data using regularized Gauss–Newton method in the data space. Pure and Applied Geophysics 179: (10), 3785-3806.

Newman GA, Alumbaugh DL (2000) Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. Geophysical journal international 140: (2), 410-424.

Lu L, Wang K, Tan H, Li Q (2020) Three-dimensional magnetotelluric inversion using L-BFGS. Acta Geophysica 68: 1049-1066.

Egbert GD, Kelbert A (2012) Computational reci -pes for electromagnetic inverse problems. G-eophysical Journal International, 189: (1), 251 -267.