

Implicit neural representation for 2D magnetotelluric inversion: A physics-constrained deep learning approach

Yanyi Wang^{a,b,c,e}, Xavier Garcia^b, Eric Attias^c, Boyao Zhang^{a,e}, Fansheng Xiong^d, Jianxin Liu^{a,e}, Zhenwei Guo^{a,e,*}

^a School of Geosciences and Info-Physics, Central South University, Changsha, 410083, China

^b Laboratory of Seafloor and Subseafloor Processes, Institute of Marine Sciences, CSIC, Barcelona, 08003, Spain

^c Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, 78712, USA

^d Beijing Institute of Mathematical Sciences and Applications, Beijing, 101408, China

^e Hunan Key Laboratory of Nonferrous Resources and Geological Hazard Exploration, Changsha, 410083, China

ARTICLE INFO

Keywords:

Magnetotelluric inversion
Deep learning
Implicit neural representations
Neural fields

ABSTRACT

Magnetotellurics (MT) inversion is essential for exploring subsurface electrical resistivity structures, but remains challenging due to nonlinearity and ill-posedness. Gradient-descent methods are often sensitive to local minima and initial models, whereas nonlinear methods are computationally expensive. Motivated by recent advances in deep learning, we propose a physics-constrained implicit neural representation (INR) framework for 2D MT inversion. Unlike data-driven neural networks, the proposed method constrains the reconstruction through a data-misfit loss evaluated using a Maxwell-equation-based MT forward operator, rather than requiring labelled training data. The continuous and compositional structure of the neural network provides an implicit regularization effect, enabling a smooth and continuous representation while retaining flexibility for complex variations. Block-model experiments confirm the robustness of the method across the tested noise levels, network sizes, and initializations. In the synthetic layered model, INR achieves an acceptable data fit comparable to MARE2DEM and yields improved model-domain metrics, reducing root mean square error (RMSE) from 0.3338 to 0.2445 and increasing structural similarity index measure (SSIM) from 0.3282 to 0.6378. Application to the COPROD2 field data further confirms the capability of the proposed 2D INR-MT inversion to recover complex subsurface structures. These results highlight INR as a promising foundation for next-generation MT inversion and motivate future extensions to 3D geophysical imaging, particularly for applications such as geological carbon dioxide storage monitoring and high-resolution characterization of complex subsurface structures.

1. Introduction

Magnetotelluric (MT) is a widely used geophysical method for probing the Earth's subsurface electrical structure by measuring natural variations in electric and magnetic (EM) fields at the surface (e.g., Constable, 2010; Di et al., 2019). These EM variations reflect the spatial distribution of subsurface resistivity, offering essential insights into the Earth's crust and upper mantle (e.g., Attias et al., 2017; Ogawa et al., 2014), mineral exploration (Di et al., 2020; Guo et al., 2020; Xue et al., 2019), hydrocarbon and geothermal reservoirs (Patro, 2017), and tectonic processes (Garcia et al., 2015, 2019; Wang et al., 2014).

Considerable research has focused on developing efficient and

accurate MT inversion methodologies. Linear inversion approaches, such as Occam and Gauss–Newton methods, rely on Tikhonov regularization theory (Loke and Dahlin, 2002; Siripunvaraporn and Egbert, 2007; Mojabi and LoVetri, 2010; Key, 2016). In these frameworks, the nonlinear inverse formulation is linearized, and the objective function is iteratively minimized using numerical optimization. While these methods successfully alleviate the inherent ill-posedness of MT inversion, their stability and accuracy depend strongly on the choice of initial model and the design of the regularization terms (Rung-Arunwan et al., 2022). Poorly configured models often cause instability or convergence to local minima, leading to biased imaging results. Additionally, nonlinear global-search strategies, including Bayesian optimization,

* Corresponding author. School of Geosciences and Info-Physics, Central South University, Changsha, 410083, China.

E-mail addresses: wang@icm.csic.es (Y. Wang), xgarcia@icm.csic.es (X. Garcia), attias@ig.utexas.edu (E. Attias), zhangboyao@csu.edu.cn (B. Zhang), hwx1334@126.com (F. Xiong), lxx6666@126.com (J. Liu), guozhenwei@csu.edu.cn (Z. Guo).

<https://doi.org/10.1016/j.cageo.2026.106277>

Received 25 February 2026; Received in revised form 11 July 2026; Accepted 27 September 2026

Available online 29 September 2026

0098-3004/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

simulated annealing, and genetic algorithms, have been applied to MT inversion (Guo et al., 2011; Chen et al., 2012). Although these techniques possess strong global exploration capabilities, they typically suffer from slow convergence, reduced computational efficiency, and extensive parameter tuning, which limit their applicability.

Recent advances in deep learning (DL) have opened new avenues for geophysical forward modeling, inversion, data processing, and interpretation (Yu and Ma, 2021). Leveraging their ability to approximate complex nonlinear mappings, supervised DL frameworks have been developed for end-to-end electromagnetic inversion (Puzyrev, 2019; Yao et al., 2024). While these approaches demonstrate the potential of data-driven inversion, their performance depends heavily on the quantity and quality of labelled training data—requirements that are difficult to meet in practical MT settings (Wu et al., 2024). As a result, they often exhibit limited generalization, weak robustness, and reduced interpretability.

Physics-constrained deep learning aims to embed physical laws directly into network training by incorporating governing equations or data residuals into the loss function (Colombo et al., 2021; Liu et al., 2023; Lin et al., 2023; Schuster et al., 2024). This strategy reduces dependence on labelled datasets and encourages the network to learn physically meaningful representations. Physics-guided recurrent neural networks (PGNN) have shown particular promise for ill-posed geophysical inversion tasks, including seismic waveform inversion (Sun et al., 2020) and EM inversion (Liu et al., 2022; Wu et al., 2024). More recent developments extend these ideas to fully unsupervised frameworks, such as unsupervised Physics-informed Full-Waveform Inversion (UPFWI) (Jin et al., 2021) and unsupervised 1D MT inversion (Goyes-Peñafiel et al., 2025), demonstrating gains in robustness and generalization using only observed data.

Implicit neural representations (INRs) have also emerged as a powerful unsupervised paradigm for geophysical inversion, offering continuous parameterizations of subsurface properties directly over coordinate space (Smith et al., 2025). Their introduction to seismic implicit full-waveform inversion (IFWI) (Song et al., 2021), followed by numerous enhancements (Du et al., 2024; Chen et al., 2025), highlights their ability to reconstruct complex structures while reducing sensitivity to initial models through intrinsic low-frequency bias. INR frameworks have since been extended to elastic and GPR FWI (Zhang et al., 2023; Sun et al., 2025), gravity inversion (Li et al., 2025), as well as resistivity inversion (Xu and Heagy, 2024, 2025), where they help mitigate near-electrode artifacts. These studies underscore the potential of INR as a stable, implicitly regularized approach for imaging subsurface electrical properties. More recently, Li et al. (2026) introduced INR into 2-D MT inversion by using a SIREN-based neural representation. Their method reparameterizes the residual between the subsurface resistivity model and a reference model, demonstrating the feasibility of coordinate-based neural parameterization for MT imaging. The sinusoidal activation in SIREN provides strong high-frequency representation capability. However, MT responses are diffusive in nature and usually provide limited resolution for small-scale resistivity variations, particularly in deep or weakly constrained regions. Therefore, an overly flexible high-frequency neural representation may introduce model variations that are not sufficiently supported by the data. The periodic nature of the sinusoidal activation may also make the recovered model more prone to oscillatory artifacts if no additional model constraint is imposed. Moreover, the performance of SIREN can be sensitive to the frequency scaling parameter, which controls the high-frequency expressive capability of the sinusoidal activation (Sitzmann et al., 2020). An inappropriate choice of frequency scaling parameter may either introduce oscillatory artifacts or weaken the recovery of localized structures, thereby increasing the difficulty of obtaining stable inversion results.

To address these issues, this study develops a Tanh-based INR-MT inversion framework that directly parameterizes the subsurface resistivity model itself. The tanh-based neural representation constructs a

smoother coordinate-to-resistivity mapping, which is more consistent with the diffusive nature and limited spatial resolution of MT responses. It also naturally couples neighboring spatial locations and introduces an implicit regularization effect through neural reparameterization, shared network weights, limited network capacity, and training dynamics. As a result, the proposed method reduces the dependence on a prescribed reference model and helps recover spatially coherent and geologically plausible resistivity structures without explicitly adding a model roughness regularization term.

We first assess the feasibility of the INR-based inversion using a simple 2D block model with different levels of noise. We then evaluate its performance using synthetic multilayer models and compare the results with those obtained from conventional 2D Occam inversion using MARE2DEM (Key, 2016; Key and Oval, 2011). Finally, we evaluate the method using the COPROD2 2D field data (Jones, 1993). The results demonstrate that the intrinsic implicit regularization of INR significantly enhances the stability, robustness, and fidelity of 2D MT imaging, enabling accurate recovery of complex structures without relying on explicit regularization terms. This study establishes a practical and effective 2D INR-based MT inversion framework and lays the groundwork for extending INR-EM inversion to fully 3D applications.

2. Methodology

2.1. MT forward problem

Forward modeling simulates observational data by numerically solving Maxwell's law for prescribed subsurface models and their electromagnetic properties. Accurate forward computations are indispensable for both conventional MT inversion and INR, as they provide the electromagnetic responses of trial models that are compared with the observed data. In the case of harmonic fields, Maxwell's law can be expressed as:

$$\begin{cases} \nabla \times \mathbf{E} = i\omega\mu\mathbf{H} \\ \nabla \times \mathbf{H} = \sigma\mathbf{E} \end{cases}, \quad (1)$$

where $\mathbf{E}$ and $\mathbf{H}$ refer to the electric and magnetic fields, respectively, i is the imaginary unit, σ denotes the electrical conductivity, ω the angular frequency, μ the permeability, and ∇ represents the gradient operator.

Based on the relative directions of the electric and magnetic fields, Maxwell's law in two-dimensional (2-D) magnetotellurics (MT) can be decomposed into two polarization modes: transverse electric (TE) and transverse magnetic (TM), each of which is sensitive to different subsurface conductivity structures. $\mathbf{E}_x$ and $\mathbf{H}_x$ of the two polarization modes satisfy the following electromagnetic field equation:

$$\frac{\partial}{\partial y} \left(\frac{1}{i\omega\mu} \frac{\partial \mathbf{E}_x}{\partial y} \right) + \frac{\partial}{\partial z} \left(\frac{1}{i\omega\mu} \frac{\partial \mathbf{E}_x}{\partial z} \right) + \sigma \mathbf{E}_x = 0, \quad (2)$$

$$\frac{\partial}{\partial y} \left(\frac{1}{\sigma} \frac{\partial \mathbf{H}_x}{\partial y} \right) + \frac{\partial}{\partial z} \left(\frac{1}{\sigma} \frac{\partial \mathbf{H}_x}{\partial z} \right) + i\omega\mu \mathbf{H}_x = 0. \quad (3)$$

To obtain a numerical solution, the computational domain is discretized into a mesh, and Eqs. (2) and (3) is solved using the finite difference method (FDM) under appropriate boundary conditions. This approach transforms the solution of the continuous field function into the solution of discrete field values at mesh nodes. Once the EM field components at each node are obtained, the MT impedance for the two polarization modes can be computed as:

$$Z_{TE} = \mathbf{E}_x / \left(\frac{1}{i\omega\mu} \cdot \frac{\partial \mathbf{E}_x}{\partial z} \right), \quad (4)$$

$$Z_{TM} = - \frac{1}{\sigma} \frac{\partial \mathbf{H}_x}{\partial z} / \mathbf{H}_x, \quad (5)$$

where Z_{TE} and Z_{TM} represent the wave impedance at the surface in TE

and TM polarizations, respectively.

Then, the impedance in the complex domain is transformed into two physical quantities in the real domain: the apparent resistivity ρ_s^{TE}, ρ_s^{TM} and the $\varphi^{TE}, \varphi^{TM}$, which are the electromagnetic response.

$$\rho_s^{TE} = \frac{1}{\omega\mu} |Z_{TE}|^2, \quad (6)$$

$$\varphi^{TE} = \text{atan } 2(\text{Im}(Z_{TE}), \text{Re}(Z_{TE})). \quad (7)$$

$$\rho_s^{TM} = \frac{1}{\omega\mu} |Z_{TM}|^2, \quad (8)$$

$$\varphi^{TM} = \text{atan } 2(\text{Im}(Z_{TM}), \text{Re}(Z_{TM})). \quad (9)$$

The forward modeling in this study is based on a finite-difference (FD) MT solver developed by Peng et al. (2023). It is implemented using a secondary-field formulation, where the 2D anomalous field is solved on an extended computational domain and homogeneous Dirichlet boundary conditions are imposed on the secondary field.

2.2. General MT inversion

In conventional Tikhonov-style MT inversion methods, the reconstruction of subsurface electrical properties is formulated as the minimization of an objective function comprising a data misfit term and a regularization term.

$$L = \|W(\mathbf{d}_{obs} - F(\mathbf{m}))\|^2 + \mu R(\mathbf{m}), \quad (10)$$

where $\mathbf{m}$ is the subsurface resistivity model, $\mathbf{d}_{obs} = [\rho_s^{TE}, \varphi^{TE}, \rho_s^{TM}, \varphi^{TM}]$ represents observations (i.e., apparent resistivity and phase of TE and TM modes), F is the forward operator (e.g., FDM), R is the regularization operator, and μ is the Lagrange multiplier that balances the trade-off between the data fit and the regularization terms. W represents a diagonal matrix with entries inversely proportional to the data errors.

Regularization plays a crucial role in conventional MT inversion by stabilizing the solution in the presence of noisy or incomplete data. By penalizing large deviations from a reference model or abrupt spatial variations, the regularization term ensures that the reconstructed subsurface models are geologically plausible and free from nonphysical oscillations. Typically, the L2 norm is employed to promote smooth solutions, particularly when prior geological information is limited (Rodi and Mackie, 2001). In contrast, the L1 norm is occasionally used to encourage compact or sparse structures (Farquharson and Oldenburg, 1998). Moreover, various application-specific regularization terms can be incorporated to integrate auxiliary information. Examples include reference-model constraints, spatially varying weighting, and structural constraints derived from prior geological knowledge (Mehanee and Zhdanov, 2002; Zhang et al., 2012).

Once the objective function is defined, an initial reference model, typically a uniform half-space, is provided. The model $\mathbf{m}$ is then iteratively updated to minimize the objective function. At each iteration, the updated model $\mathbf{m}$ is fed into a forward numerical simulation F , whose output $F(\mathbf{m}) = \mathbf{d}_{pred}$ is subsequently used to compute the data misfit term. The process stops either when the data misfit cannot be further reduced or when it reaches a specified threshold.

2.3. Implicit neural representation-based MT inversion

Assume a set of features can be interpreted as a continuous function Φ , with respect to the input $\mathbf{x} \in \Omega_k, k = 1, \dots, K$, which is implicitly defined by equations of the form:

$$\Phi : \mathbf{x} \in \Omega_k \rightarrow \Phi(\mathbf{x}). \quad (11)$$

In the context of two-dimensional magnetotelluric (MT) inversion, the subsurface electrical properties, such as conductivity or resistivity, can

be treated as a continuous function that maps spatial coordinates to physical parameter values. Specifically, for any point $\mathbf{r} = (y, z)$ in the spatial domain Ω , the electrical parameter σ or ρ can be represented as a continuous function $\Phi(\mathbf{r})$.

Based on this continuous-field representation, we use a coordinate-based neural network to parameterize the subsurface resistivity. The network takes spatial coordinates as input and outputs the corresponding resistivity value, as expressed as:

$$\rho = \Phi(\mathbf{r}) = \mathcal{N}_\Theta(\mathbf{r}), \quad (12)$$

where $\mathcal{N}_\Theta$ represents the deep neural network with trainable network parameters Θ , designed as described in Section 2.4. In this way, the resistivity model is represented by a shared set of Θ , rather than by independent values assigned to each inversion cell. This parameterization couples different spatial locations through the neural representation and provides a flexible way to describe continuous resistivity variations.

The network parameters are optimized under the guidance of Maxwell's equations by minimizing the discrepancy between the predicted and observed MT responses. In our implementation, the MT forward modeling is implemented as a differentiable finite-difference solver in PyTorch. The finite-difference system matrix, electromagnetic fields, impedance responses, apparent resistivities, phases, and data-misfit loss are all computed using PyTorch tensor operations, so that the complete forward calculation remains in the automatic-differentiation computational graph. The data-misfit gradient is therefore obtained by back-propagating through the differentiable forward solver and then through the coordinate-based neural representation using the chain rule. This allows the network parameters to be updated directly using gradient-based optimization while preserving the physics-based data constraint. The resulting optimization problem is formulated as follows:

$$L(\Theta) = \|W(F(\mathcal{N}_\Theta(\mathbf{r})) - \mathbf{d}_{obs})\|^2, \quad (13)$$

where F denotes the forward modeling operator of Maxwell's law.

Compared with Eq. (10), this objective function does not include an explicit regularization term, such as a Tikhonov-type roughness penalty. This is because the INR parameterization itself provides an implicit form of model constraint. Specifically, the resistivity model is reparameterized as the output of a coordinate-based MLP, rather than being treated as independent cell-wise parameters. When the network adopts smooth nonlinear activations, the coordinate-to-property mapping tends to generate continuous model variations. Moreover, the shared network weights naturally couple neighboring spatial locations, while the finite network capacity restricts the recovered model to the function space representable by the MLP and helps suppress arbitrary small-scale variations. In addition, the training dynamics tend to recover smooth, low-frequency structures before fitting high-frequency details. These factors jointly act as implicit regularization, enabling physics-based inversion without an explicit model roughness penalty.

In principle, the INR inversion framework does not require a starting model in the same way as conventional inversion, because the subsurface model is implicitly generated by the neural network. However, since the neural network's parameters are randomly initialized, and the initial model predicted by the network may deviate significantly from a reasonable resistivity distribution. This can increase the number of iterations required for convergence. Therefore, an optional pre-training stage can be introduced for better initialization, in which the network is first trained to represent a reference model. It should be noted that the reference model used in pre-training is only used to initialize the network parameters and is not imposed as a hard constraint during the subsequent inversion. This provides a more informed initialization for the network parameters and can accelerate the subsequent physics-guided inversion. The loss function used in the pre-training stage can be designed as:

$$L_{pre}(\Theta) = \|\mathcal{N}_\Theta(\mathbf{r}) - \mathbf{m}_{init}\|^2, \quad (14)$$

where m_{init} represents the initial model, commonly a uniform background in electromagnetic inversion.

The approach we use to solve this optimization problem is described in Algorithm 1, and the framework is illustrated in Fig. 1.

Algorithm 1. INR-based 2D MT inversion algorithm

Input:
 (x, z) : Cell spatial coordinates,
 d_{obs} : observed data (apparent resistivity and phase),
 m_{init} : the initial resistivity model,
 $\mathcal{N}$: the neural network,
 Θ : hyperparameters of $\mathcal{N}$,
 F : forward operator,
 T_{pre} : maximum epochs of pre-training,
 T_{main} : maximum epochs of main-training,
 T_{es} : patience of early stopping,
 T_{lr} : patience of learning-rate decay,
 P : loss average period,
 η : learning rate.

1. Step1: Prepare for inputs
 Normalize coordinates: $u = (u - u_{min}) / (u_{max} - u_{min}), u \in (x, z)$

2. Step2: Pre-training
 for $t = 1$ to T_{pre} do
 Evaluate network at all cell centers to obtain $m = \mathcal{N}_{\Theta}(u)$
 Compute $L_{pre}(\Theta)$
 Update Θ using $L_{pre}.backward()$ with the optimizer
 end for
 Return pre-trained network $\mathcal{N}_{\Theta}^{pre}$

3. Step3: Main-training
 Load pre-trained network $\mathcal{N}_{\Theta} = \mathcal{N}_{\Theta}^{pre}$
 for $t = 1$ to T_{main} do
 Evaluate network at all cell centers: $m = \mathcal{N}_{\Theta}(u)$
 Forward model: $d_{pred} = F(m)$
 Compute physics-constrained loss $L(\Theta)$
 Update Θ using $L.backward()$ with optimizer
 if $t \% P == 0$ then
 Compute mean loss over current period
 if mean loss did not decrease for T_{lr} then
 $\eta = \eta * 0.8$
 end if
 end if
 if mean loss did not improve for T_{es} or loss ≤ 1 then
 Terminate training
 end if
 end for
Output: Return final resistivity model m and trained network $\mathcal{N}_{\Theta}$.

2.4. Architecture of neural networks and optimization

To implement the physics-constrained optimization (Eq. (13)), the subsurface model m is reparameterized via a fully connected multilayer perceptron (MLP), denoted $\mathcal{N}_{\Theta}$. Fig. 1 presents a schematic diagram of this approach, illustrating how the network maps spatial coordinates to resistivity values.

In the 2D setting, the input $\mathcal{N}_{\Theta}$ consists of the spatial coordinates of the mesh cell centers (x, z) , scaled to $[0, 1]$. The output corresponds to the predicted resistivity at each coordinate. In the examples we present here, the network is designed with five hidden layers, each containing 256 neurons and employs the tanh activation function to introduce nonlinearity, as its smooth and continuous nature is well suited for modeling subsurface resistivity distributions that are typically smooth except at geological boundaries. Depending on the inversion target, the number of hidden layers, the number of neurons, and the choice of activation function can be adjusted to better suit the specific problem. During training, the optimizer is Adam, which is robust, converges quickly, and is suitable for large-scale hyperparameter optimization.

During main training, the learning rate determines the magnitude by which model parameters are updated in each iteration. A learning rate that is too high may cause oscillations or prevent convergence, while one that is too low may result in slow convergence. Therefore, a proper learning rate schedule is crucial for successful training.

Compared with a fixed-step decay schedule, the adaptive loss-based strategy provides a more robust mechanism for adjusting the learning rate. Fixed-step decay schemes reduce the learning rate at pre-determined epochs regardless of the model's convergence state, potentially leading to premature decay or unnecessary slowdowns. In contrast, monitoring average loss provides a more dynamic, data-driven control. When the loss stagnates over a given period, the learning rate is reduced to promote further convergence. This approach maintains a relatively large learning rate during the early stage for faster exploration and gradually switch to smaller steps as the model approaches a local minimum, improving both stability and convergence efficiency. In this study, we define a period as five epochs and calculate the mean loss for each period. If the mean loss failed to decrease within a specified number of epochs, which is set to 100 epochs in the following examples, the learning rate was multiplied by 0.8.

$$\eta = \eta_0 \times 0.8^a, \quad (15)$$

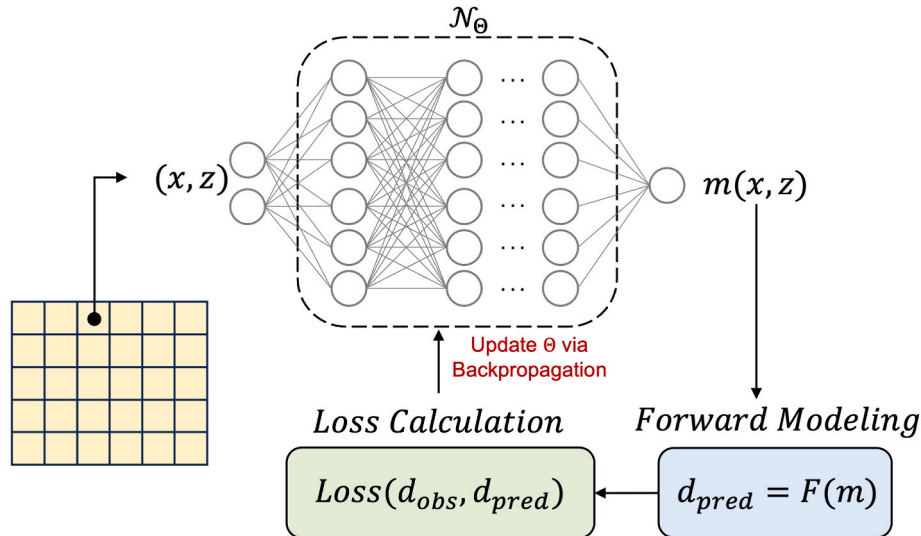


Fig. 1. Implicit neural representation (INR) based inversion framework. The yellow grids denote the subsurface computational mesh. The blue block represents the forward modeling operator (e.g., Maxwell's law), while the green block denotes the loss function used for network training. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

where, η_0 is the initial learning rate, here is 0.001, and n is the number of decay steps.

A stopping criterion was additionally implemented to terminate training once the loss did not improve over a specified number of epochs (150 in the following examples) or loss reaches 1.

This early stopping strategy not only prevents unnecessary computation but also helps avoid overfitting to small fluctuations in the loss, ensuring that the recovered subsurface model captures the main features of the data while remaining physically plausible.

2.5. Evaluation metrics

The accuracy of the inversion results can also be quantitatively evaluated using three statistical metrics: the root mean square error (RMSE), coefficient of determination (R^2), and structural similarity index measure (SSIM).

$$\text{RMSE} = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n (y - y')^2 \right)} \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \mu_y)^2} \quad (17)$$

$$\text{SSIM} = \frac{(2\mu_y\mu_{y'} + c_1)(2\sigma_{yy'} + c_2)}{(\mu_y^2 + \mu_{y'}^2 + c_1)(\sigma_y^2 + \sigma_{y'}^2 + c_2)}, \quad (18)$$

where y represents the reference model, and y' denotes the predicted model. μ_y and $\mu_{y'}$ correspond to the mean values of y and y' , respectively. σ_y and $\sigma_{y'}$ also represent the variances of y and y' , respectively. Finally, $\sigma_{yy'}$ signifies the covariance between y and y' . c_1 and c_2 are two constants determined by the numerical range of the image. Smaller RMSE values indicate a better fit between predicted and real data, while larger R^2 and SSIM values signify a better regression model fit.

3. Trials on synthetic examples

3.1. Case 1: simple block model

In this example, we consider a simple model consisting of two conductive blocks embedded in a uniform background (Fig. 2a) (Wang et al., 2026a). The tensor mesh size is 40×80 , and each cell is $40\text{m} \times 50\text{m}$. The two uniform blocks with a resistivity of $10 \Omega\text{m}$, each with dimensions of $400\text{m} \times 750\text{m}$, are embedded in a homogeneous $1000 \Omega\text{m}$ background. The shallower uniform is located at a depth of 400m , and another one is at a depth of 600m . Five logarithmically spaced padding cells were added on lateral sides and at the bottom, extending the lateral boundaries to two times the core-model half-width and the bottom boundary to ten times the core-model depth. Additionally, ten logarithmically spaced air cells were added above the surface. It should be noted that the air layer is also included in the forward modeling (not shown), similar to the padding zones.

In the forward simulations, responses are calculated over a frequency range from 10^{-4} Hz to 10^4 Hz, with a total of 31 frequency points. The 21 receivers are evenly distributed along the surface at 200m intervals from -2 km to 2 km (Fig. 2). To avoid the inverse crime, the synthetic MT data were generated using a finite-element forward solver implemented in MARE2DEM, whereas the forward operator employed during inversion is based on a finite-difference scheme.

To assess the robustness of the proposed method, zero-mean Gaussian random noise was introduced into the synthetic MT data at two noise levels. In the low-noise scenario, Gaussian noise with 2% standard deviation was added to the logarithm of apparent resistivity,

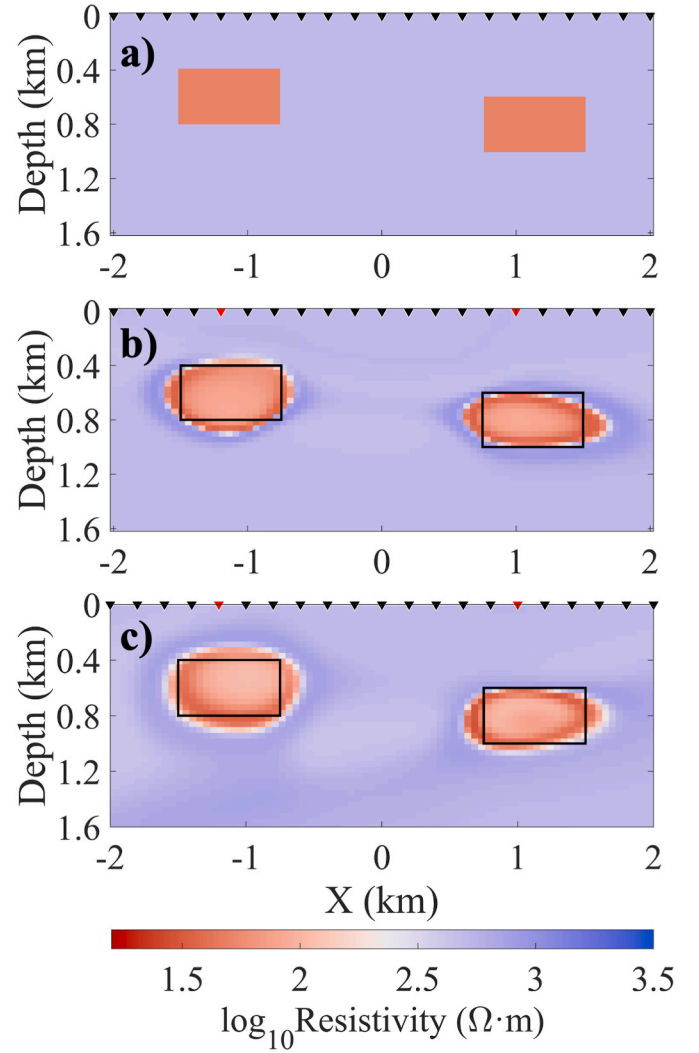


Fig. 2. a) Reference resistivity model; b) Predicted resistivity model from low-level noise observations; c) Predicted resistivity model from medium-level noise observations. The receivers are shown as triangles. Panels b) and c) show the true location of the conductive anomalies as boxes as a reference.

and Gaussian noise with 1° standard deviation was added to the phase data. In the medium-noise scenario, the standard deviation of noise levels was increased to 5% for apparent resistivity and 4° for the phase data. All other hyperparameters used during training are kept constant. For the pre-training stage, we use a homogeneous half-space with a resistivity of $500 \Omega\text{m}$.

Fig. 2 shows the recovered resistivity models for data with low- and medium-level noise added. As noise levels increase, the accuracy of delineating anomaly boundaries decreases. In the inversion results for medium-level noise, the boundary of the left anomaly extends outward, and the deeper part of the anomaly is overestimated, leading to an apparent deviation from results at lower noise levels (Fig. 2b and c). Correspondingly, the RMSE increases slightly, while the R^2 and SSIM values show a mild decrease with increasing noise (Table 1).

Table 1

Quantitative metric result of two-block model between reference and predicted models.

	RMSE	R^2	SSIM
Low-level	0.1901	0.6153	0.7314
Medium-level	0.2081	0.5674	0.6094

Nevertheless, these variations remain small overall, and both results accurately reproduce the background resistivity and yield satisfactory inversions of the shapes, positions, and amplitudes of the two block anomalies, demonstrating our method's robustness to fit noisy data within some error threshold.

Fig. 3 compares synthetic observations with low-level noise to the predicted responses of two receivers, demonstrating that the predictions generally fit the observations well. Figs. 4 and 5 present the corresponding data fit pseudosections along the survey line, illustrating the overall excellent agreement across all receivers and frequencies. Fig. 6 illustrates the dynamic adjustment of the learning rate during training using data with low-level noise. After 2250 epochs, the learning rate decreased from 0.001 five times, a 3-fold decrease, to approximately 0.0003. Fig. 7 shows the corresponding loss curves. Despite some short-term oscillations at the initial stage, the loss curve decreases as the number of training iterations increases, indicating that the model is gradually learning the data features. After ~ 1500 iterations, it levels off. And RMS finally reaches 1.007. Similar trends are also observed across medium noise levels, indicating the stability and convergence of the proposed scheme (see the Supplementary Material).

We further evaluated the sensitivity of the proposed method to the network architecture using the low-noise block-model example by varying the number of hidden layers and neurons (Fig. 8). For networks with 5 hidden layers, increasing the number of neurons from 64 to 128 improves the model-domain metrics, with the SSIM increasing from 0.7429 to 0.7749 and the R^2 increasing from 0.6544 to 0.7297. The final data RMS misfit also decreases slightly from 1.0348 to 1.0223. When the

number of neurons is further increased to 256, the model achieves the highest SSIM of 0.7992 and the lowest data RMS misfit of 1.0212, while requiring fewer training epochs. However, the R^2 and RMSE are not further improved compared with the 128-neuron case. In contrast, reducing the network depth to 3 hidden layers with 256 neurons per layer leads to poorer model-domain metrics, with an SSIM of 0.4726, an R^2 of 0.4617, and an RMSE of 0.2375, despite achieving a relatively low data RMS misfit of 1.0030.

Beyond the influence of network architecture, we also investigated the robustness of the inversion with respect to random initialization and the use of pre-training. To evaluate the stability of the proposed method, we performed five independent runs using different random seeds, namely 10, 100, 1000, 2000, and 3000, while keeping all other settings the same as those used in the block-model example. The mean and standard-deviation maps of the recovered resistivity models are shown in Fig. 9. With pre-training using a homogeneous 500 Ω m reference model, the mean model preserves the main resistivity structures of the block model, while the standard-deviation map shows that the variability among different runs is mainly concentrated near anomaly boundaries and resistivity transition zones (Fig. 9a and b). Averaged over the five runs, the inversion with pre-training requires 3269 training epochs and achieves a final data RMS misfit of 1.012.

In addition, we examined whether the proposed INR inversion remains stable without the pre-training stage. The five independent inversions without pre-training also recover consistent main resistivity structures, indicating that pre-training is not a necessary condition for obtaining stable inversion results (Fig. 9c and d). However, compared

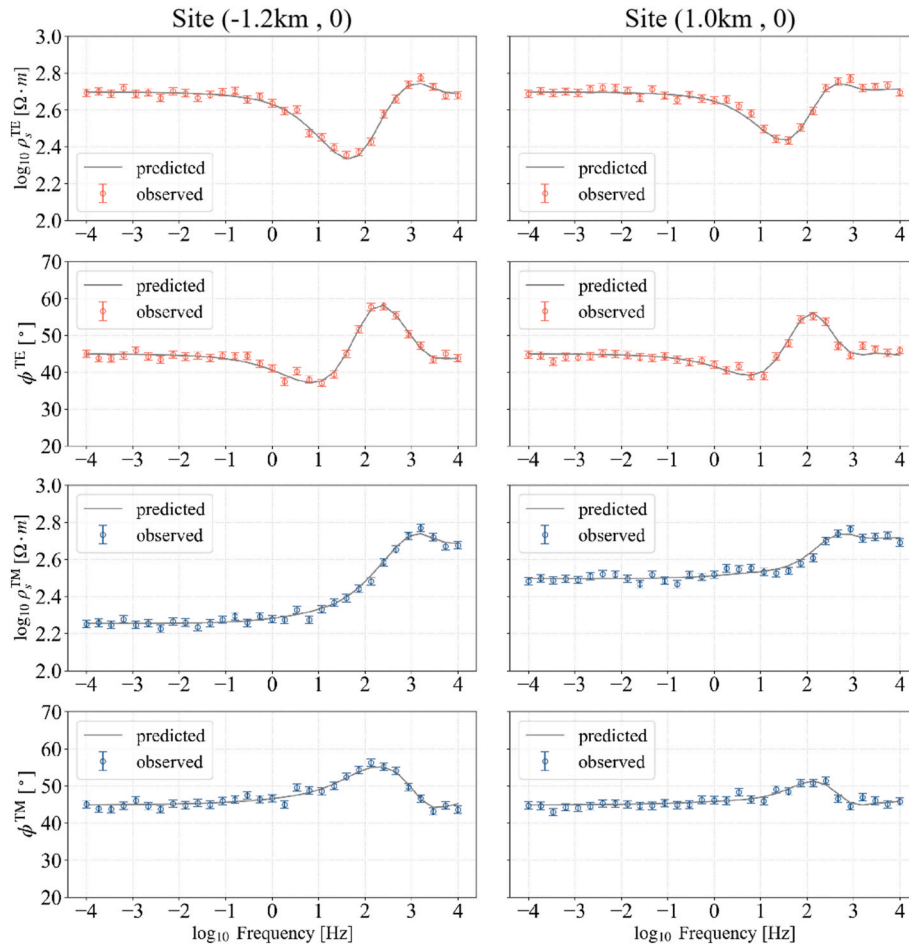


Fig. 3. Data comparison of the synthetic observations and predicted responses for two selected sites ($x = -1.2$ km and 1.0 km, red triangles in Fig. 2b) for the low-noise dataset. Dots with error bars denote measured data. Solid lines denote the response of the inverted model. Error bars shown represent 2% in apparent resistivity and 1° in phase. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

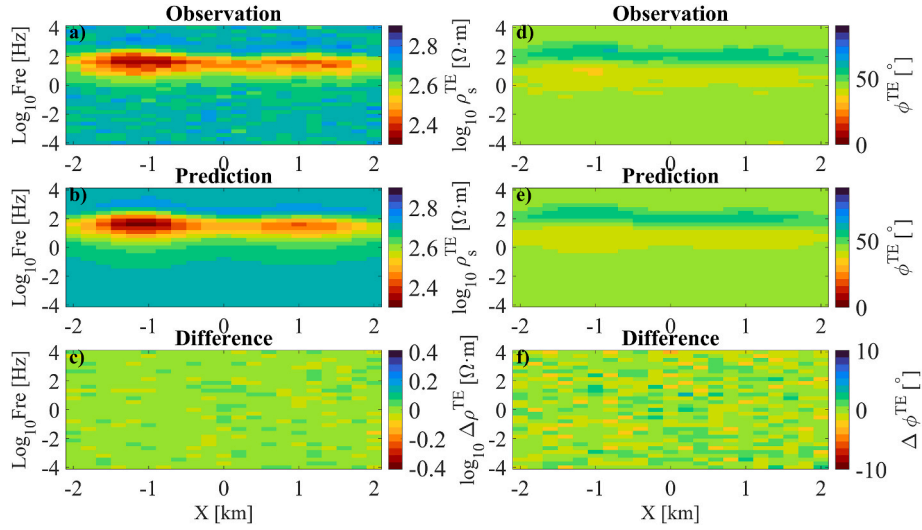


Fig. 4. TE Data-fit profiles for the low-noise dataset. Observation and predicted apparent resistivity and phase, together with their absolute data misfits. a) Observed apparent resistivity; b) Predicted apparent resistivity; c) Absolute difference between observation and prediction of apparent resistivity; d) Observed phase; e) Predicted phase; f) Absolute difference between observation and prediction of phase.

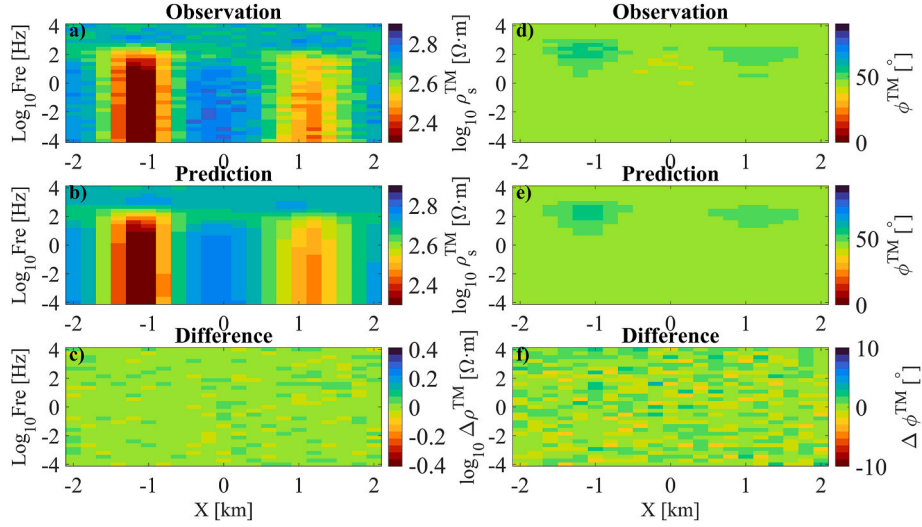


Fig. 5. TM Data-fit profiles for the low-noise dataset. Observation and predicted apparent resistivity and phase, together with their absolute data misfits. a) Observed apparent resistivity; b) Predicted apparent resistivity; c) Absolute difference between observation and prediction of apparent resistivity; d) Observed phase; e) Predicted phase; f) Absolute difference between observation and prediction of phase.

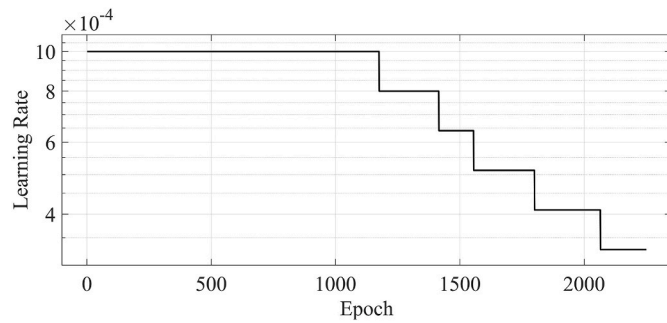


Fig. 6. Learning rate variation curve during NN training using observed data with low-level noise.

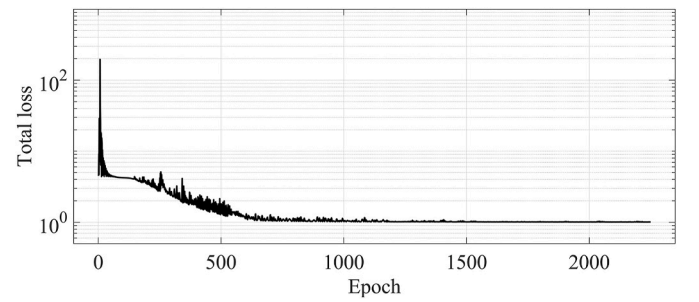


Fig. 7. Training Loss curves during NN training using observed data with low-level noise.

with the pre-training case, the inversions without pre-training require more training epochs, with an average of 4582 epochs, and obtain a

slightly higher average final data RMS misfit of 1.030.

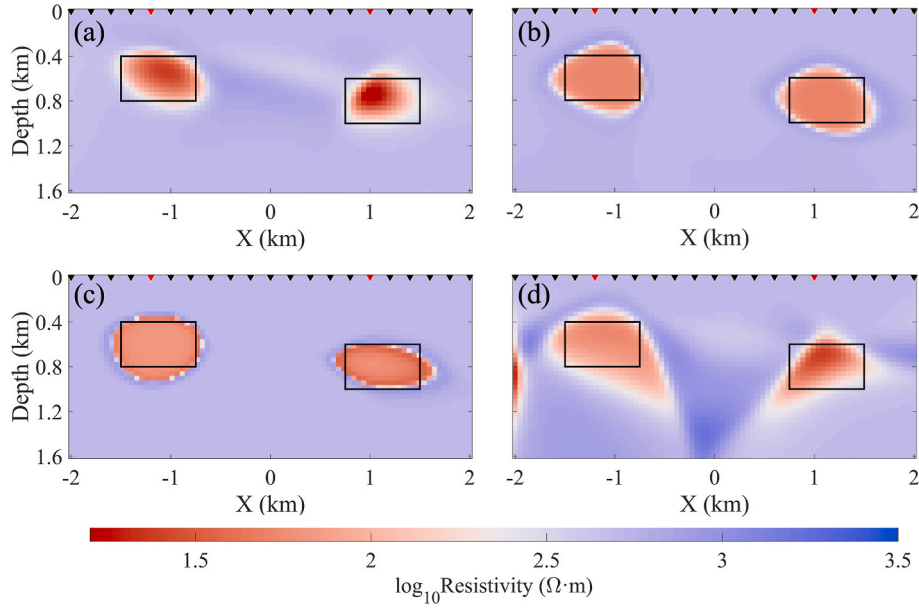


Fig. 8. Inversion results with different sizes of network. (a) 5 hidden layers with 64 neurons per layer. (b) 5 hidden layers with 128 neurons per layer. (c) 5 hidden layers with 256 neurons per layer. (d) 3 hidden layers with 256 neurons per layer.

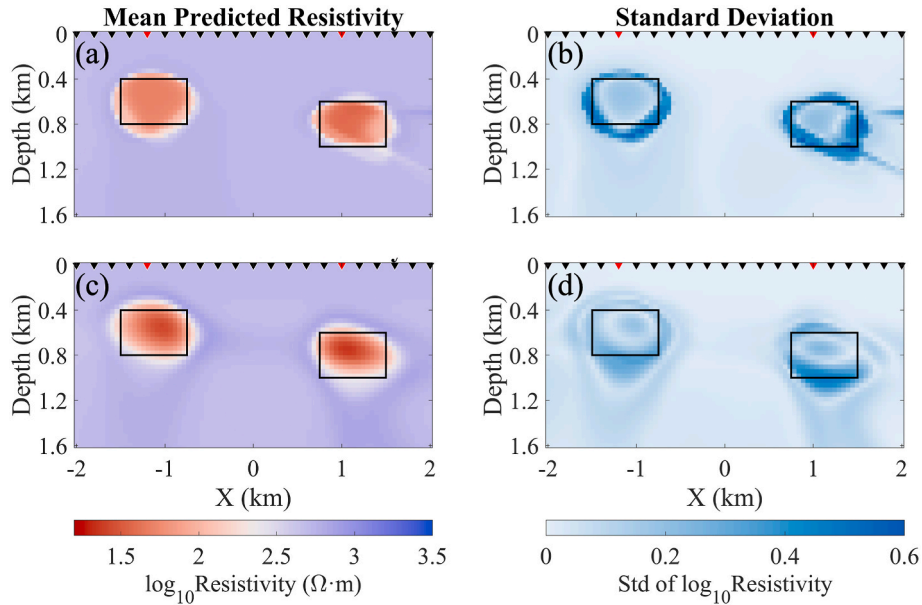


Fig. 9. Inversion results from five independent runs with different random seeds for block model. a) Mean predicted resistivity model with pre-training; b) standard deviation of the predicted resistivity models with pre-training; c) mean predicted resistivity model without pre-training; d) standard deviation of the predicted resistivity models without pre-training.

3.2. Case 2: layer model with two anomalies

In this section, the effectiveness of the proposed method is evaluated using synthetic examples. A four-layer synthetic model is constructed, in which resistivity decreases progressively with depth (Wang et al., 2026b). A shallow high-resistivity anomaly near the surface and a deeper low-resistivity anomaly within the second layer are designed to approximate real geological structures as exploration targets (Fig. 10a).

To further assess the inversion performance, the INR inversion results are compared with those obtained from the Occam inversion performed by MARE2DEM (Key, 2016; Key and Oval, 2011). The MT profile consists of 21 survey points spaced 2.5 km apart. The observed MT responses used for the inversion were synthetically generated by

MARE2DEM forward modeling over a frequency range of 10^{-4} – 10^4 Hz, and 2% Gaussian noise was added to simulate observational uncertainty.

For INR, we use a mesh with core cells of $250 \text{ m} \times 100 \text{ m}$, with 200 core-mesh cells in the x-direction and 100 cells in the z-direction. The mesh is further enlarged by five logarithmically stretched cells on both sides and beneath the model, where the lateral and bottom extensions reach four times the core-model half-width and ten times the core-model depth, respectively. An additional 10 air cells are placed above the surface. The INR inversion was implemented in PyTorch under the Python 3.12.0 environment and trained on a workstation equipped with an Intel(R) Xeon(R) Platinum 8383C CPU @ 2.70 GHz. For MARE2DEM inversion, the subsurface is meshed into about 24780 cells, a number comparable to that used in the INR model. The mesh was extended

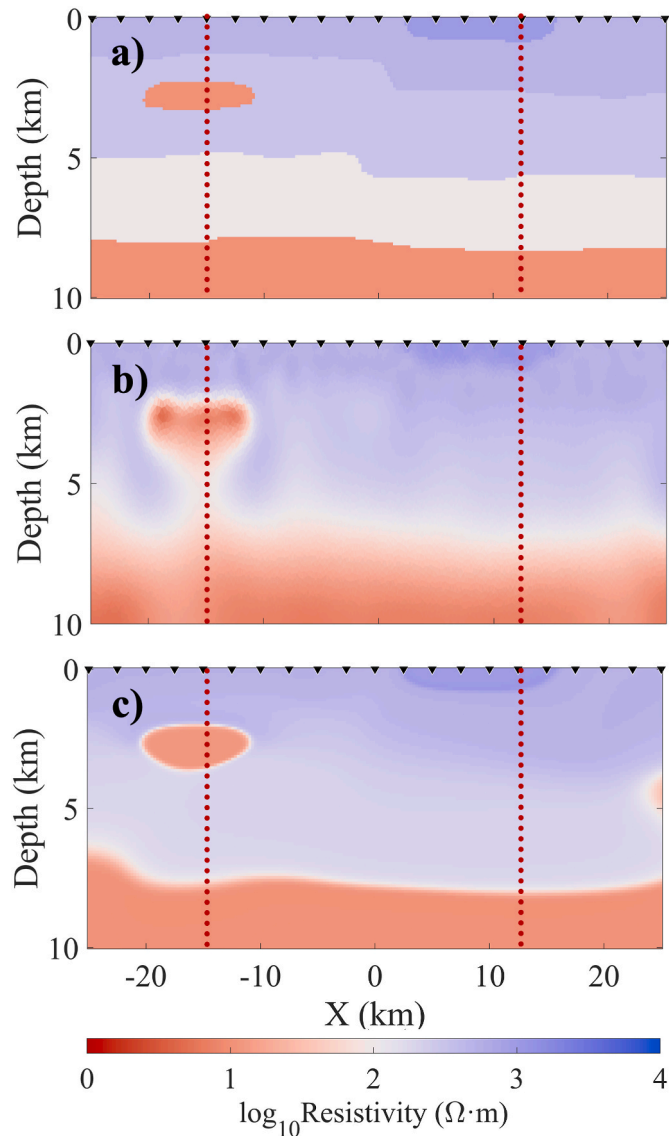


Fig. 10. a) Reference resistivity model; b) MARE2DEM inversion result; c) INR-MT inversion result. The red dotted lines correspond to the 1D profiles shown in Fig. 9. The receivers are shown as triangles. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

laterally and vertically to cover the same model domain as the INR mesh. The initial model used for MARE2DEM inversion was a 500 Ωm homogeneous half-space, which is the same for INR inversion pre-training. The resistivity model was constrained using gradient-based roughness regularization. The roughness weights in the horizontal and vertical directions were set to 3 and 1, respectively, and a penalty cut weight of 0.1 was applied. MARE2DEM Occam inversion was executed on a high-performance computing (HPC) cluster with 2 nodes and 48 cores.

The INR inversion was terminated after satisfying the early-stopping criterion, with 2614 training iterations and 2614 forward evaluations. And the final data RMS reaches 1.43. In contrast, MARE2DEM inversion satisfied the stopping criteria after 8 iterations, requiring 35 forward evaluations and reaching a final data RMS of 1.17. The runtime of the MARE2DEM and INR are 549.8s and 123306s, respectively.

Fig. 10b and c presents the inverted resistivity models obtained by INR and MARE2DEM. To further assess local accuracy, we extracted two 1D resistivity pseudo-well curves crossing the central anomaly regions at $x = -15\text{ km}$ and $x = 10\text{ km}$ (Fig. 11). From the 2D profiles, both methods

can resolve the two prominent anomalies and the large-scale layered structure. The MARE2DEM result (Fig. 10b) shows a relatively smooth distribution, which is consistent with its regularized inversion framework. It captures the broad resistivity variations and tends to smooth the layer boundaries and the anomaly boundaries. In comparison, the INR result preserves the main anomalous body and the layered structure more clearly in the 2-D section (Fig. 10c). The resistive anomaly is more compact and closer to the reference geometry. The vertical profiles in Fig. 11 provide a more detailed comparison of the local resistivity variations at the two selected locations. At $X = -15\text{ km}$, the two methods show different amplitude deviations in the middle and deep parts of the model (Fig. 11a). At $X = 10\text{ km}$, both methods capture the main vertical resistivity trend, with discrepancies mainly occurring near layer transitions and in deeper region, and the undulating interface between the middle and lower layers is better retained by INR (Fig. 11b).

The quantitative metrics in Table 2 further support these observations. INR achieves a lower RMSE of 0.2445, a higher R^2 , and a higher SSIM compared with MARE2DEM. These results suggest that, for this synthetic example, INR better preserves the overall structural similarity and resistivity contrast, while MARE2DEM provides a smoother reconstruction.

The accuracy of the inversion results can also be quantitatively evaluated using RMSE, SSIM, and R^2 . Table 2 summarizes the results for MARE2DEM and INR. INR achieves significantly lower RMSE, higher R^2 , and higher SSIM compared to MARE2DEM, indicating superior numerical accuracy, better agreement with the reference model, and improved structural fidelity.

4. Application to field data: COPROD2

To further verify the effectiveness of the INR method on field MT data, we applied it to invert MT data collected in southern Saskatchewan and Manitoba, Canada (Jones, 1988, 1993; Jones and Savage, 1986). It is an East–West profile across the Williston Basin in Canada. Along this profile, the surface is covered by low-resistivity sediments formed during the Paleozoic, beneath which lie high-resistivity Cambrian basement rocks. It passes through two large-scale conductive anomalous geo-features, namely, the North American Central Plains (NACP) and the Thompson Belt (TOBE) (Jones et al., 2005). The more prominent anomaly, located near the center of the profile, corresponds to NACP. In contrast, a second basement anomaly at the eastern end of the profile is associated with the TOBE.

The COPROD2 dataset is high-quality and consists of 40 frequencies, from 384Hz to 1/680Hz of data collected at 35 sites along a 407 km profile. We use the data set comprising the TE and TM mode data, which have already been corrected for static shift. We further inspected the apparent-resistivity and phase curves and excluded evident outliers or data points showing strong inconsistency with neighboring stations and frequencies, in order to reduce the influence of non-2-D responses. Data uncertainties are assigned as 10% for apparent resistivities and 2.9° for phases, consistent with the data fit levels reported by deGroot-Hedlin and Constable (1993). Both TE and TM components are included in the objective function without additional empirical mode-dependent weighting; their weights are controlled by the prescribed data uncertainties.

The inversion mesh for INR method comprises 90×120 cells in the vertical and horizontal directions, respectively. Horizontally, the core region is discretized into 100 uniform cells with a spacing of 4000 m, while 10 logarithmically spaced padding cells are added on each side. These padding cells extend the lateral boundaries to 1.5 times the half-width of the core model. Vertically, the subsurface core region extends to a depth of 60 km and is divided into 70 nonuniform cells, whose boundaries are defined using logarithmic spacing. An additional 10 logarithmically spaced padding cells are added beneath the core region, extending the bottom boundary to twice the core-model depth. Above the surface, 10 air cells are included, with their thicknesses increasing

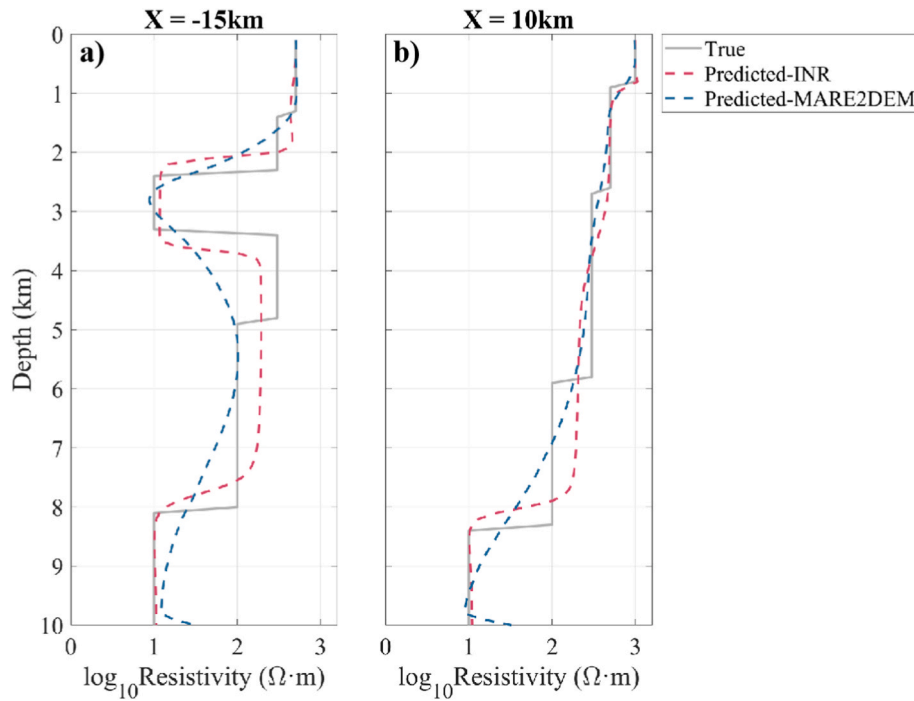


Fig. 11. 1D pseudo-well resistivity profiles crossing the anomaly regions at a) $x = -15$ km and b) $x = 10$ km.

Table 2

Quantitative comparison between the INR and MARE2DEM inversion results for the synthetic layered model case.

	RMSE	R^2	SSIM	RMS (data)	Forward evaluations	Time cost (s)
MARE2DEM	0.3338	0.6506	0.3282	1.17	35	549.8
INR	0.2445	0.8487	0.6378	1.43	2614	123306

progressively upward. The INR inversion is also implemented under the same software environment and hardware configuration with synthetic layered model case. It is terminated after 7869 epochs, corresponding to 7869 forward evaluations, with a runtime of 92408 s. The final data RMS misfit reaches 1.45 (Table 3).

We also inverted the same data set using MARE2DEM for comparison. The subsurface domain was discretized into approximately 15,244 cells, including the extension cells. The MARE2DEM mesh was also extended laterally and vertically to cover the same model domain as the INR mesh. The initial model used for MARE2DEM inversion was a 10 Ωm homogeneous half-space, which is the same for INR inversion pre-training. Other inversion settings, including the regularization strategy, roughness constraints, target misfit, fast Occam parameters, stopping criteria, and computational hardware configuration, were kept the same as those used in the synthetic layered-model case. The final accepted MARE2DEM model was obtained after 16 iterations, involving 335 forward evaluations, and achieved a final data RMS misfit of 1.39. And the total computational times were 7494s (Table 3).

We compare our solution with MARE2DEM results and previous studies to assess the recovery of the main conductivity structures. As shown in Fig. 12, both methods recover the principal conductive features of the profile, including the broad NACP anomaly, the TOBE conductor, and the shallow conductive layer interpreted as Paleozoic

sediments. The INR-MT model images the NACP as a continuous anticlinal-shaped, westward-dipping conductor with a horizontal extent of approximately -50 to 50 km, a top boundary at around 10 km depth, and a conductive core centered near 20 km depth. This overall geometry agrees with previous Occam 2-D inversion results reported by Degroot-Hedlin and Constable (1993).

The TOBE conductor is also recovered by both methods near the eastern part of the profile. For INR inversion, the TOBE structure is vertical and narrow, with its top located at the base of the Paleozoic sediments and bottom boundary at a depth of 8 km (Fig. 12). The resistivity values in both zones are below 10 Ωm , as described in previous studies (Romo et al., 2005; Zhao et al., 1993). The MARE2DEM result shows the corresponding conductor but with smoother boundaries and deeper extension. The resistivity values in both the NACP and TOBE zones are generally below 10 Ωm , consistent with previous studies by Romo et al. (2005) and Zhao et al. (1993). Both methods recover a laterally continuous, shallow low-resistivity layer, which is interpreted as Paleozoic sediments. In INR inversion, this conductive layer has a thickness of about 2.5 km, whereas it appears slightly thicker and more diffuse in the MARE2DEM result.

The comparison suggests that the proposed INR-MT method can recover the principal conductivity features of the COPROD2 benchmark in a manner broadly consistent with both MARE2DEM and previous studies. However, some differences can also be observed between the two inversion results in the background resistivity distribution and the detailed morphology of local conductive anomalies. The MARE2DEM result shows a smoother and more laterally continuous background, which is expected from the explicit Occam roughness regularization. In contrast, the INR-MT model exhibits stronger resistivity contrasts in some regions, reflecting the different implicit regularization behavior of

Table 3

Quantitative comparison between the INR and MARE2DEM inversion results for the COPROD2 dataset.

	RMS	Forward evaluations	Time cost (s)
MARE2DEM	1.39	335	7494
INR	1.45	7869	92408

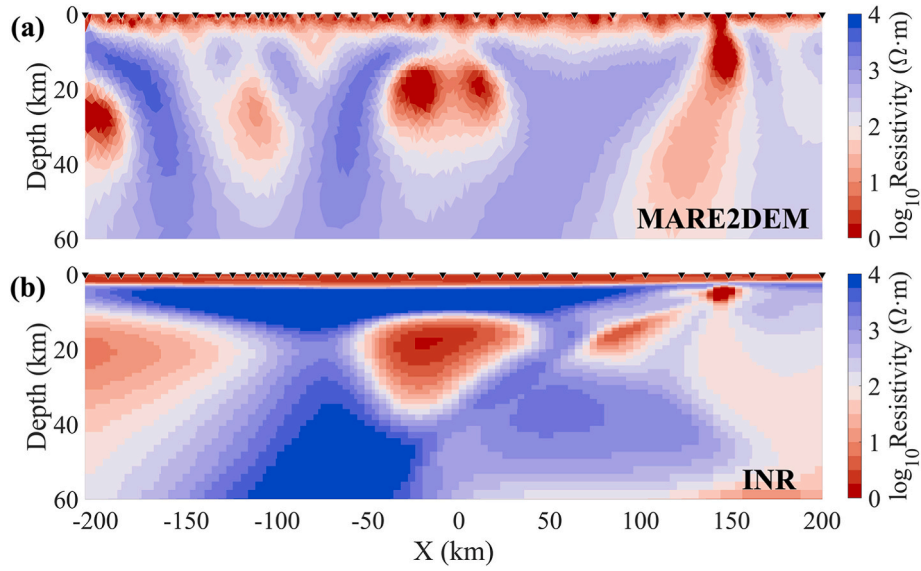


Fig. 12. Inversion models derived from the COPROD2 data set with (a) MARE2DEM; (b) INR. Location of the MT stations are shown as inverted black triangles along the surface.

the coordinate-based neural representation.

Figs. 13 and 14 present the resulting data fits of INR method, showing that the predicted responses reproduce the main trends of observed data across most frequencies and sites. Fig. 15a–d shows the RMS values calculated for each individual station. Most stations exhibit RMS values around 1, while a few reach up to 3, mainly due to misfits in the TM phase at low frequencies. The frequency-dependent RMS misfit in Fig. 15e shows that the residuals are mainly concentrated at the high- and low-frequency ends, while the intermediate frequency band is generally fitted better, with RMS values close to or below 1.0. The relatively high RMS values at high frequencies may be associated with shallow small-scale heterogeneity, residual static-shift effects, galvanic distortion, or local data noise, which are difficult to fully represent in a smooth 2-D model. The increased RMS values at low frequencies, particularly at the longest periods, may be related to lower signal-to-noise ratios, larger data uncertainty, deeper conductivity variations, or possible off-profile and along-strike structures beyond the 2-D assumption. Nevertheless, the overall RMS misfit of the INR-MT inversion

remains 1.45, which is comparable to the MARE2DEM result with an RMS misfit of 1.39, indicating that the proposed method reproduces the main data trends reasonably well, although localized residuals persist at the frequency-band edges.

The COPROD2 example demonstrates that the proposed INR-MT framework can be applied to field MT data and can recover the main conductivity features in a manner broadly consistent with MARE2DEM and previous studies. However, because localized residuals remain at several stations and frequency bands, and because possible non-2-D effects may still influence part of the dataset. The comparison also shows that, at the current stage, INR-MT is mainly valuable as an alternative neural parameterization and implicit-regularization framework, rather than as a more efficient replacement for mature conventional inversion codes. Future work will focus on improving the computational efficiency of the differentiable forward solver and extending the framework to more complex field settings.

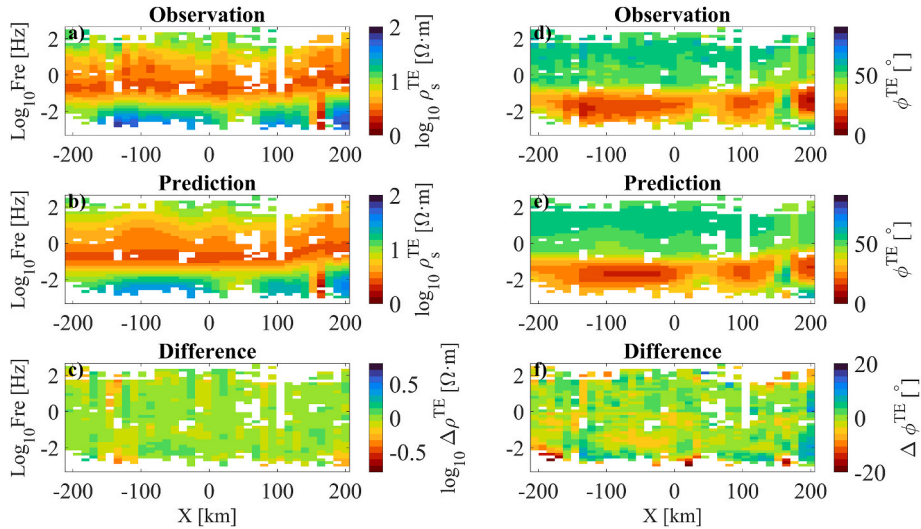


Fig. 13. Pseudosections of the observed, modelled MT data and absolute difference for TE polarization. a) Observed apparent resistivity; b) Predicted apparent resistivity; c) Absolute difference between observation and prediction of apparent resistivity; d) Observed phase; e) Predicted phase; f) Absolute difference between observation and prediction of phase.

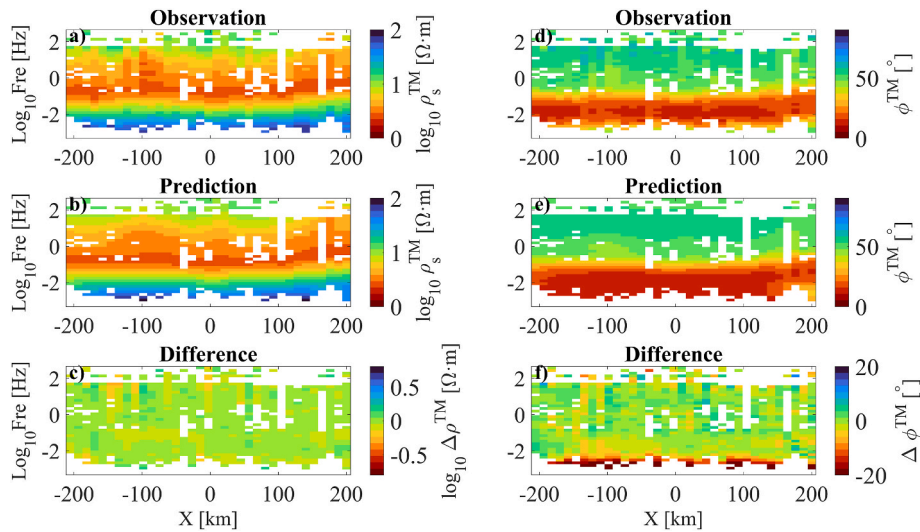


Fig. 14. Pseudosections of the observed, modelled MT data and absolute difference for TM polarization. a) Observed apparent resistivity; b) Predicted apparent resistivity; c) Absolute difference between observation and prediction of apparent resistivity; d) Observed phase; e) Predicted phase; f) Absolute difference between observation and prediction of phase.

5. Discussion

5.1. Robustness assessment and parameterization effects

The repeated-run results for block model and layered model indicate that the proposed INR-based inversion framework can consistently recover the main resistivity structures under different random initializations (Fig. 9 and Fig. S3). The relatively low variability within the interior of the main anomalies suggests that the large-scale resistivity distribution is stable across independent runs. In contrast, the higher variability near anomaly boundaries and resistivity transition zones indicates that sharp interfaces are less well constrained, which is consistent with the inherent non-uniqueness and smoothing tendency of regularized geophysical inversion. It should be noted that these repeated-run tests provide an empirical assessment of inversion stability rather than a full probabilistic uncertainty quantification. More rigorous uncertainty analysis, such as Bayesian sampling or ensemble-based posterior estimation, is beyond the scope of this study and will be investigated in future work.

In addition to random initialization, the inversion results are also influenced by the neural-network parameterization itself. The architecture-sensitivity test shown in Fig. 8 indicates that although the shallower network with 3 hidden layers achieves a relatively low data RMS misfit, its poorer model-domain metrics suggest that it lacks sufficient representational capability to recover geologically reasonable resistivity structures. In comparison, increasing the network width improves the reconstruction quality for the 5-layer networks, but the improvement is not strictly monotonic for all metrics. The architecture used in the main experiments, namely 5 hidden layers with 256 neurons per layer, provides a favorable balance among structural similarity, data fitting, and training efficiency in this example. Nevertheless, the optimal architecture may depend on the complexity of the subsurface model, data coverage, and noise level. Therefore, the network architecture should be selected by balancing model representation capability, data-fitting accuracy, and computational cost.

In addition to network depth and width, the supplementary activation-function test further shows that the choice of activation function also affects the performance of INR-based inversion and provides support for the use of Tanh in the main experiments (Fig. S4). This is because, in the proposed INR-based inversion, the neural network is not only an optimization tool, but also a continuous parameterization of the subsurface resistivity distribution. Therefore, the smoothness and

continuity of the activation function directly influence the smoothness and stability of the recovered resistivity model. Compared with ReLU, Tanh is smooth, bounded, and continuously differentiable, which makes it more suitable for representing gradual resistivity variations and smooth transitions in the subsurface. This is also confirmed by the noisy block-model test, where Tanh achieves higher SSIM and R^2 , lower RMSE and data RMS misfit, and fewer training epochs than ReLU under the same inversion settings (Table S3). Therefore, Tanh is adopted in the main experiments because it provides a more stable and effective representation for INR-based MT inversion in this study.

The role of pre-training should be interpreted as an initialization strategy rather than a necessary constraint on the final inversion model. The results without pre-training show that the proposed INR inversion can still recover the main resistivity structures from random initialization, indicating that the framework does not fundamentally rely on a prescribed starting model (Fig. 9c and d). However, when pre-training is used, the choice of the reference model can influence the subsequent optimization. A suitable homogeneous reference model can provide a more favorable initial neural representation, helping the inversion converge faster and improving the recovered model. The sensitivity test in the supplementary material shows that, among the tested homogeneous reference models, the 500 Ω m model provides the best overall model-domain metrics (Fig. S2). It suggests that pre-training is most beneficial when the reference model is reasonably selected. Although pre-training is not strictly required, it can improve the stability and efficiency of the inversion when appropriate prior information is available. Therefore, if reliable prior information is available, such as regional background resistivity, geological constraints, or results from previous inversions, it can be incorporated into the pre-training stage to provide a more informed initialization; otherwise, the inversion can still be performed without pre-training.

5.2. Comparison with other related methods

Compared with conventional purely data-driven deep-learning inversion methods, the proposed method reduces the dependence on large labelled training datasets. Purely data-driven methods usually learn a direct mapping from geophysical observations to subsurface models based on paired data-model samples. Once trained, they can provide fast predictions, but their performance strongly depends on the representativeness of the training dataset, and their generalization may be limited when the observed data differ from the training distribution.

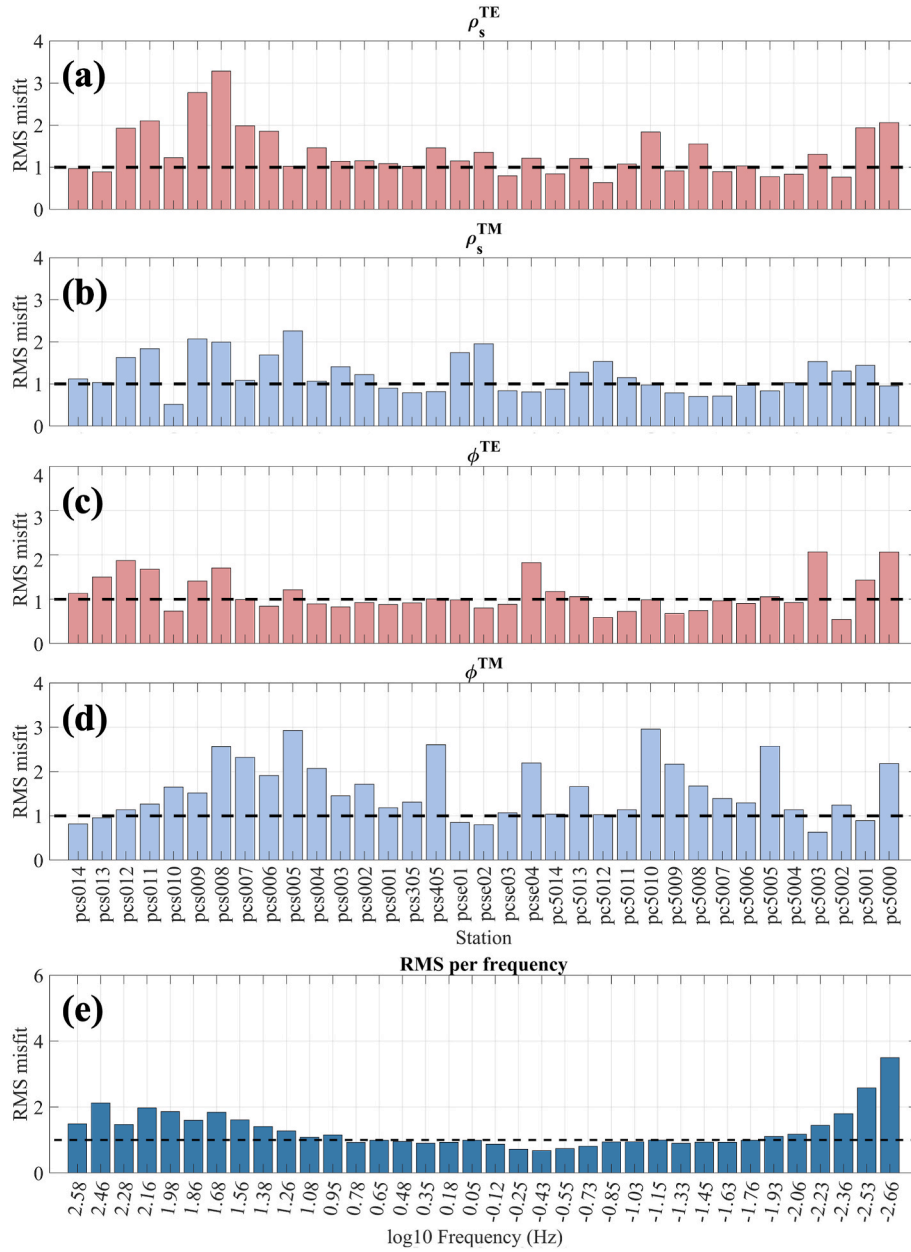


Fig. 15. a) TE apparent resistivity RMS for stations; b) TE phase RMS for stations; c) TM apparent resistivity RMS for stations; d) TM phase RMS for stations; e) RMS for frequencies.

To improve the physical consistency of data-driven inversion, several physics-guided or physics-driven deep-learning methods have been developed (Liu et al., 2022, 2024; Wu et al., 2024). These methods still generally rely on training datasets during the training stage, but introduce physical information into the loss function, or forward-modeling consistency. Therefore, these hybrid methods provide an important step beyond purely data-driven inversion by combining learned data-model relationships with geophysical constraints.

More recently, physically guided deep unsupervised inversion methods have been proposed to further reduce the dependence on labelled model datasets. For example, physically guided deep unsupervised inversion for 1-D MT models constrains the inversion using the data misfit calculated from the forward response, rather than relying on supervised labels (Goyes-Peñañiel et al., 2025). However, it learns a mapping from observed data to resistivity models using fully connected networks. Such a data-to-model mapping may face scalability challenges when extended to higher-dimensional MT inversion problems or more

complex survey configurations, because the network input and output dimensions are closely tied to the data and model discretization.

In contrast to the above deep-learning-based inversion methods, the proposed INR-based MT inversion does not learn a global data-to-model mapping from a labelled training dataset. Instead, the neural representation is optimized for each specific observed dataset through the MT data misfit. This avoids the need for large paired training datasets and decouples the network input from the dimension of the observed data. Compared with conventional grid-based MT inversion methods, the INR framework also provides a flexible coordinate-based representation of the subsurface model, in which the resistivity distribution is described as a continuous function rather than as independent cell-wise parameters. This continuous representation can reduce the dependence on a fixed inversion grid and may provide greater flexibility for representing complex resistivity variations. In this study, the forward responses are computed using a finite-difference solver; however, if an appropriate mesh-free forward-modeling algorithm is adopted, the INR framework

could potentially be extended toward a fully mesh-free inversion strategy (Huang and Alkhalifah, 2022).

Nevertheless, the proposed method also has limitations, particularly in terms of computational efficiency. Similar to conventional iterative MT inversion, each optimization step requires forward modeling, and the total computational cost is therefore largely controlled by the efficiency of the forward solver. For both the synthetic layered-model and COPROD2 examples, the INR-based inversion requires substantially more forward evaluations and longer runtimes than MARE2DEM, increasing from 35 to 2614 and from 335 to 7869 forward evaluations, respectively, while achieving slightly higher data RMS misfits (Tables 2 and 3). These comparisons indicate that the current INR implementation is substantially more computationally expensive than the mature conventional inversion algorithm used here. This higher computational cost mainly results from the test-time optimization strategy, in which the neural representation must be trained separately for each inversion problem, together with the repeated calls to the finite-difference forward solver during network optimization. In addition, the current forward solver has not been parallelized, whereas mature inversion codes such as MARE2DEM can efficiently exploit parallel computing resources. Therefore, the reported runtime mainly reflects the computational cost of the present implementation rather than the full efficiency potential of the INR-based framework. With further improvements in forward-modeling efficiency, parallel implementation, and optimization strategies, the proposed formulation could become practical for large-scale MT inversion while retaining the methodological advantages discussed above.

6. Conclusions

In this study, we propose a framework for MT inversion using implicit neural representations (INR) and implicit regularization that are inherently incorporated into the Neural Network. Unlike conventional inversion schemes that rely on explicit regularization terms (e.g., Tikhonov), the proposed INR-based inversion achieves smoothness and structural adaptivity through the neural network's architecture. The results of our synthetic block and complex-layer models with embedded sparse anomalies show strong recovery of sharp boundaries and fine structural details. The results further indicate that the proposed method maintains stable inversion performance under increased noise levels and provides improved reconstruction of the main anomalous structures compared with conventional L2-norm Tikhonov-regularized inversion methods. These findings suggest that the inherent inductive biases and continuity of the INR representation can serve as powerful forms of implicit regularization, stabilizing the inversion without the need for manually tuned penalty parameters. To further validate its practical applicability, we apply the proposed method to the inversion of field-measured data. INR performance on COPROD2 aligns well with previous studies' findings, further illustrating its feasibility for addressing real-world exploration objectives.

Our framework provides a flexible and extensible foundation for future geophysical inversion studies. Although the current implementation is computationally more expensive than mature conventional inversion algorithms and supervised deep-learning prediction after off-line training, its main value lies in reducing the dependence on labelled training datasets, providing implicit regularization through the coordinate-based neural representation, and offering a continuous and differentiable formulation for future extensions. This formulation may facilitate the development of three-dimensional and multi-scale inversion frameworks by incorporating spatial coordinates in a flexible manner. It also provides a natural way to integrate multiple types of geophysical data within a unified optimization process. Future work could focus on accelerating the inversion process by incorporating more efficient forward solvers, surrogate forward models, or adaptive modeling strategies. Combining the proposed INR representation with more efficient and potentially mesh-free forward-modeling algorithms

may further improve its applicability to large-scale and higher-dimensional geophysical inversion problems. These advances could further promote its application to geological CO₂ storage monitoring, where flexible and robust imaging of complex three-dimensional subsurface structures is crucial for tracking plume migration and evaluating storage integrity.

CRediT authorship contribution statement

Yanyi Wang: Writing – original draft, Visualization, Validation, Software, Methodology, Data curation, Conceptualization. **Xavier Garcia:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Formal analysis. **Eric Attias:** Writing – review & editing, Supervision, Resources, Formal analysis. **Boyao Zhang:** Writing – review & editing, Visualization, Validation, Data curation. **Fansheng Xiong:** Writing – review & editing, Conceptualization. **Jianxin Liu:** Writing – review & editing, Project administration. **Zhenwei Guo:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Code availability section

Name of the code/library: INR-MT.

Contact: wyy0213@csu.edu.cn.

Hardware requirements: Intel(R) Xeon(R) Platinum 8383C CPU @ 2.70 GHz.

Program language: Python.

Software required: PyTorch, Matplotlib, SciPy, NumPy.

Tested environment: Python 3.12.0, PyTorch 2.2.0, NumPy 1.26.4, SciPy 1.15.3, Matplotlib 3.9.2.

License: MIT License.

The source codes are available for downloading at the link: <https://github.com/wyy-csu/INR-MT.git>.

The released code includes the main differentiable inversion workflow used in this study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Alan Jones for making the processed COPROD2 dataset publicly available. And we also thank Zhong Peng for making the MT finite difference forward codes, and Kerry Key for making MARE2DEM publicly available.

The project was supported by the National Science and Technology Major Project for Deep Earth Probe and Mineral Resources Exploration of China (No. 2025ZD1008507). Financial support was also provided by the National Natural Science Foundation of China (Grant No. 42574195) and the Program of China Scholarship Council (Grant No. 202406370228).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cageo.2026.106277>.

Data availability

I have shared my code and dataset.

References

- Attias, E., Evans, R.L., Naif, S., Elsenbeck, J., Key, K., 2017. Conductivity structure of the lithosphere-asthenosphere boundary beneath the eastern North American margin. *G-cubed* 18, 676–696.
- Chen, J., Hoversten, G.M., Key, K., Nordquist, G., Cumming, W., 2012. Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site. *Geophysics* 77, E265–E279.
- Chen, R., Wu, B., Li, M., Luo, Y., 2025. Full waveform inversion with velocity model low-rank implicit neural representation. *IEEE Trans. Geosci. Rem. Sens.*
- Colombo, D., Turkoglu, E., Li, W., Sandoval-Curiel, E., Rovetta, D., 2021. Physics-driven deep-learning inversion with application to transient electromagnetics. *Geophysics* 86 (3), E209–E224.
- Constable, S., 2010. Ten years of marine CSEM for hydrocarbon exploration. *Geophysics* 75, 75A67–75A81.
- Degroot-Hedlin, C., Constable, S., 1993. Occam's inversion and the North American Central Plains electrical anomaly. *J. Geomagn. Geoelectr.* 45, 985–999.
- Di, Q., Xue, G., Zeng, Q., Wang, Z., An, Z., 2020. Magnetotelluric exploration of deep-seated gold deposits in the Qingchengzi orefield, Eastern Liaoning (China), using a SEP system. *Ore Geol. Rev.* 122, 103501.
- Di, Q., Zhu, R., Xue, G., Yin, C., Li, X., 2019. New development of the Electromagnetic (EM) methods for deep exploration. *Chin. J. Geophys.* 62, 2128–2138.
- Du, B., Sun, J., Jia, A., Wang, N., Liu, H., 2024. Physics-informed robust and implicit full waveform inversion without prior and low-frequency information. *IEEE Trans. Geosci. Rem. Sens.* 62, 1–12.
- Farquharson, C.G., Oldenburg, D.W., 1998. Non-linear inversion using general measures of data misfit and model structure. *Geophys. J. Int.* 134, 213–227.
- Garcia, X., Julià, J., Nemocón, A.M., Neukirch, M., 2019. Lithospheric thinning under the Araripe Basin (NE Brazil) from a long-period magnetotelluric survey: constraints for tectonic inversion. *Gondwana Res.* 68, 174–184.
- Garcia, X., Seillé, H., Elsenbeck, J., Evans, R., Jegen, M., Hölz, S., Ledo, J., Lovatini, A., Marti, A., Marcuello, A., 2015. Structure of the mantle beneath the A Ibora B asin from magnetotelluric soundings. *G-cubed* 16, 4261–4274.
- Goyes-Peñafiel, P., Bin Waheed, U., Argueld, H., 2025. Physically guided deep unsupervised inversion for 1D Magnetotelluric models. *IEEE Geosci. Remote Sens. Lett.*
- Guo, R., Dosso, S.E., Liu, J., Dettmer, J., Tong, X., 2011. Non-linearity in Bayesian 1-D magnetotelluric inversion. *Geophys. J. Int.* 185, 663–675.
- Guo, Z., Xue, G., Liu, J., Wu, X., 2020. Electromagnetic methods for mineral exploration in China: a review. *Ore Geol. Rev.* 118, 103357.
- Huang, X., Alkhalifah, T., 2022. PINNup: robust neural network wavefield solutions using frequency upscaling and neuron splitting. *J. Geophys. Res. Solid Earth* 127, e2021JB023703.
- Jin, P., Zhang, X., Chen, Y., Huang, S.X., Liu, Z., Lin, Y., 2021. Unsupervised Learning of full-waveform Inversion: Connecting CNN and Partial Differential Equation in a Loop arXiv preprint arXiv:2110.07584.
- Jones, A.G., 1988. Static shift of magnetotelluric data and its removal in a sedimentary basin environment. *Geophysics* 53, 967–978.
- Jones, A.G., 1993. Electromagnetic images of modern and ancient subduction zones. *Tectonophysics* 219, 29–45.
- Jones, A.G., Ledo, J., Ferguson, I.J., 2005. Electromagnetic images of the Trans-Hudson orogen: the North American Central Plains anomaly revealed. *Can. J. Earth Sci.* 42, 457–478.
- Jones, A.G., Savage, P.J., 1986. North American Central Plains conductivity anomaly goes east. *Geophys. Res. Lett.* 13, 685–688.
- Key, K., 2016. MARE2DEM: a 2-D inversion code for controlled-source electromagnetic and magnetotelluric data. *Geophys. J. Int.* 207, 571–588.
- Key, K., Ovall, J., 2011. A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling. *Geophys. J. Int.* 186, 137–154.
- Li, H., Liu, Ying, Sun, J., 2026. Implicit neural representation two-dimensional magnetotelluric inversion. *Chin. J. Geophys.* 69 (1), 396–410 (in Chinese).
- Li, X., Zhao, J., Zhou, S., 2025. Implicit neural representations for 3D gravity inversion. *Comput. Geosci.*, 106082.
- Lin, Y., Theiler, J., Wohlberg, B., 2023. Physics-guided data-driven seismic inversion: recent progress and future opportunities in full-waveform inversion. *IEEE Signal Process. Mag.* 40 (1), 115–133.
- Liu, B., Pang, Y., Jiang, P., Liu, Z., Liu, B., Zhang, Y., Cai, Y., Liu, J., 2023. Physics-driven deep learning inversion for direct current resistivity survey data. *IEEE Trans. Geosci. Rem. Sens.* 61, 1–11.
- Liu, L., Yang, B., Zhang, Y., 2024. Inverting magnetotelluric data using a physics-guided auto-encoder with scaling laws extension. *Front. Earth Sci.* 12, 1510962.
- Liu, W., Wang, H., Xi, Z., Zhang, R., Huang, X., 2022. Physics-driven deep learning inversion with application to magnetotelluric. *Remote Sens.* 14, 3218.
- Loke, M.H., Dahlin, T., 2002. A comparison of the Gauss–Newton and quasi-Newton methods in resistivity imaging inversion. *J. Appl. Geophys.* 49, 149–162.
- Mehanee, S., Zhdanov, M., 2002. Two-dimensional magnetotelluric inversion of blocky geoelectrical structures. *J. Geophys. Res. Solid Earth* 107, EPM 2-1-EPM 2-11.
- Mojabi, P., LoVetri, J., 2010. Comparison of TE and TM inversions in the framework of the Gauss–Newton method. *IEEE Trans. Antenn. Propag.* 58, 1336–1348.
- Ogawa, Y., Ichiki, M., Kanda, W., Mishina, M., Asamori, K., 2014. Three-dimensional magnetotelluric imaging of crustal fluids and seismicity around Naruko volcano, NE Japan. *Earth Planets Space* 66, 158.
- Patro, P.K., 2017. Magnetotelluric studies for hydrocarbon and geothermal resources: examples from the Asian region. *Surv. Geophys.* 38, 1005–1041.
- Peng, Z., Yang, B., Liu, L., Xu, Y., 2023. Rapid surrogate modeling of magnetotelluric in the frequency domain using physics-driven deep neural networks. *Comput. Geosci.* 176, 105360.
- Puzyrev, V., 2019. Deep learning electromagnetic inversion with convolutional neural networks. *Geophys. J. Int.* 218, 817–832.
- Rodi, W., Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion. *Geophysics* 66, 174–187.
- Romo, J.M., Gómez-Treviño, E., Esparza, F.J., 2005. Series and parallel transformations of the magnetotelluric impedance tensor: theory and applications. *Phys. Earth Planet. Inter.* 150, 63–83.
- Rung-Arunwan, T., Siripunvaraporn, W., Utada, H., 2022. The effect of initial and prior models on phase tensor inversion of distorted magnetotelluric data. *Earth Planets Space* 74, 51.
- Schuster, G.T., Chen, Y., Feng, S., 2024. Review of physics-informed machine-learning inversion of geophysical data. *Geophysics* 89 (6), T337–T356.
- Siripunvaraporn, W., Egbert, G., 2007. Data space conjugate gradient inversion for 2-D magnetotelluric data. *Geophys. J. Int.* 170, 986–994.
- Smith, L.T., Horrocks, T., Akhtar, N., Holden, E.-J., Wedge, D., 2025. Implicit neural representation for potential field geophysics. *Sci. Rep.* 15, 9799.
- Song, C., Alkhalifah, T., Waheed, U.B., 2021. Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks. *Geophys. J. Int.* 225, 846–859.
- Sun, J., Liu, Y., Lin, Y., Xing, L., Liu, H., 2025. Implicit multiparameter full waveform inversion of multioffset ground penetrating radar data. *Geophys. J. Int.* 240, 904–919.
- Sun, J., Niu, Z., Innanen, K.A., Li, J., Trad, D.O., 2020. A theory-guided deep-learning formulation and optimization of seismic waveform inversion. *Geophysics* 85, R87–R99.
- Wang, X., Zhang, G., Fang, H., Luo, W., Zhang, W., Zhong, Q., Cai, X., Luo, H., 2014. Crust and upper mantle resistivity structure at middle section of Longmenshan, eastern Tibetan plateau. *Tectonophysics* 619, 143–148.
- Wang, Y., Garcia, X., Attias, E., Zhang, B., Xiong, F., Liu, J., Guo, Z., 2026a. Blocks model example dataset for paper "Implicit Neural Representation for 2D Magnetotelluric Inversion: a Physics-Constrained Deep Learning Approach". Zenodo. <https://doi.org/10.5281/zenodo.18480852>.
- Wang, Y., Garcia, X., Attias, E., Zhang, B., Xiong, F., Liu, J., Guo, Z., 2026b. Layered model example dataset for paper "Implicit Neural Representation for 2D Magnetotelluric Inversion: a Physics-Constrained Deep Learning Approach". Zenodo. <https://doi.org/10.5281/zenodo.18497702>.
- Wu, S., Huang, Q., Zhao, L., 2024. Physics-guided deep learning-based inversion for airborne electromagnetic data. *Geophys. J. Int.* 238, 1774–1789.
- Xu, A., Heagy, L.J., 2024. A test-time learning approach to reparameterize the geophysical inverse problem with a convolutional neural network. *IEEE Trans. Geosci. Rem. Sens.*
- Xu, A., Heagy, L.J., 2025. Towards Understanding the Benefits of Neural Network Parameterizations in Geophysical Inversions: a Study with Neural Fields arXiv preprint arXiv:2503.17503.
- Xue, G., Chen, W., Cheng, J., Liu, S., Yu, J., Lei, K., Guo, W., Feng, X., 2019. A review of electrical and electromagnetic methods for coal mine exploration in China. *IEEE Access* 7, 177332–177341.
- Yao, H.M., Ng, M., Jiang, L., 2024. Deep learning electromagnetic inversion solver based on a two-step framework for high-contrast and heterogeneous scatterers. *IEEE Trans. Antenn. Propag.* 72, 5337–5342.
- Yu, S., Ma, J., 2021. Deep learning for geophysics: current and future trends. *Rev. Geophys.* 59, e2021RG000742.
- Zhang, L., Koyama, T., Utada, H., Yu, P., Wang, J., 2012. A regularized three-dimensional magnetotelluric inversion with a minimum gradient support constraint. *Geophys. J. Int.* 189, 296–316.
- Zhang, T., Sun, J., Trad, D., Innanen, K., 2023. Multilayer perceptron and Bayesian neural network-based elastic implicit full waveform inversion. *IEEE Trans. Geosci. Rem. Sens.* 61, 1–16.
- Zhao, G., Tang, J., Zhan, Y., 1993. Two-dimensional modelling by using COPROD2 magnetotelluric data. *J. Geomagn. Geoelectr.* 45, 963–968.