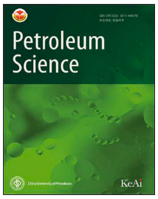
Petroleum Science 23 (2026) 1908–1928

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science



Original Paper

Multi-task U-net inversion of synthetic look-ahead logging-while-drilling data

Shun Zhang[a,b], Wen-Xiu Zhang[a,b,*], Wen-Xuan Chen[a,b], Peng-Fei Liang[a,b], Wen-Yang Wang[a,b], Xing-Han Li[a,b]

[a] *Key Laboratory of Deep Oil and Gas Theory and Intelligent Exploration and Production, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, 100029, China*
[b] *College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China*

ARTICLE INFO

Article history:
Received 16 February 2025
Received in revised form
14 December 2025
Accepted 14 December 2025
Available online 18 December 2025

Edited by Jia-Jia Fei

Keywords:
Multi-task U-net
Look-ahead
Anisotropy
Multiple components
Inversion

ABSTRACT

Electromagnetic look-ahead logging while drilling instruments detect the electrical characteristics of undrilled formations, enabling proactive decision-making. Real-time geological insight ahead of the drill bit is critical for effective geosteering. This study introduces a multi-task U-net neural network that simultaneously inverts multiple formation parameters real-time. Six datasets, each corresponding to different electromagnetic components, were used to train six specialized neural networks. All networks exhibited rapid convergence and successfully inverted 60,000 sample in 15 s, satisfying real-time requirements. Residual and relative error analyses reveal that the multi-component network delivers the highest accuracy. Sensitivity analysis shows that coaxial and coplanar components are more sensitive to conductivity variations, whereas coaxial and cross-components excel at resolving interface positions. The *yy* component displays the strongest sensitivity to anisotropy. Compared with the traditional Levenberg-Marquardt algorithm, the proposed method demonstrates improved accuracy and efficiency. Moreover, the Levenberg-Marquardt inversion with the neural network output as initial models further enhances accuracy. Benchmark comparisons reveal that the multi-task U-net outperforms various mainstream machine learning and deep learning models, including LSTM, FCN, ResNet, and XGBoost, in both inversion accuracy and generalization. Moreover, sensitivity analyses to noise and near-bit geological complexity reveal that, while the proposed model experiences some performance degradation under high noise levels or highly heterogeneous backgrounds, it maintains strong robustness under moderate noise conditions and achieves reliable inversion results in two-layer geological settings. These results establish the multi-task U-net as a fast, accurate, and robust tool for real-time electromagnetic look-ahead inversion in geosteering applications.

1. Introduction

Electromagnetic (EM) logging-while-drilling (LWD) technology plays a crucial role in geosteering for oil and gas development. These tools provide real-time identification and characterization of reservoirs, which is essential for optimizing well placement and trajectory planning, ultimately improving recovery efficiency and economic return (Hawkins et al., 2015; Wang and Fan, 2019; Wu et al., 2022; Zhu et al., 2023). Over the past few decades, EM LWD tools have evolved from non-azimuthal logging system (Coope and Yearsley, 1986; Clark et al., 1988; Fredericks et al., 1989) to azimuthal EM logging (Prensky, 2006; Zhang et al., 2008; Yang et al., 2025), and subsequently to azimuthal deep-reading EM logging (Ndokwu et al., 2018; Durdyyev, 2021). Recently, advancements have enabled EM look-ahead capabilities (Ma et al., 2022), allowing detection of geological features ahead of the drill bit (Bittar et al., 2021). For instance, Schlumberger introduced early prototypes in 2016 for detecting boundaries with 12¼ to 14-inch wellbores (Constable et al., 2016), followed by the commercial IriShpere service in 2019, which offered look-ahead detection up to 30 m (Seydoux et al., 2019). In 2022, Halliburton released the BrightStar tool (Halliburton, 2023), and the Institute of Geology

---

* Corresponding author.
E-mail address: zhangwenxiu@mail.iggcas.ac.cn (W.-X. Zhang).
Peer review under the responsibility of China University of Petroleum (Beijing).

https://doi.org/10.1016/j.petsci.2025.12.023
1995-8226/

and Geophysics of the Chinese Academy of Sciences also developed a prototype (Liang et al., 2023).

Despite their potential, EM look-ahead tools typically record data in the form of attenuation (Att) and phase shift (PS), which do not directly reflect geological properties. Inversion of these measurements into electrical properties is required for actionable interpretation.

Early EM LWD inversions primarily used cross-plot corrections to estimate conductivity, boundaries, and dip angles (Gianzero et al., 1994; Jackson and Hagiwara, 1998). This technique is intuitive and easy to implement. However, as the detection depth of EM LWD tools increases, the demand for multi-parameter inversion across multiple formations has emerged. The cross-plot correction method is inherently limited, as it cannot simultaneously invert multiple parameters. Multi-parameter inversion is an ill-posed and highly nonlinear problem, typically formulated as a nonlinear least-squares optimization that requires iterative refinement to converge to an optimal solution. To address this challenge, researchers introduced iterative inversion techniques for EM LWD measurements. Common iterative methods include gradient-based approaches, such as Gauss-Newton, Levenberg-Marquardt, and conjugate gradient methods (Dong and Zhao, 2020; Zhang et al., 2022; Li et al., 2023), as well as stochastic methods, such as differential evolution, Bayesian algorithms, and particle swarm optimization (Mohd Aris et al., 2021; Rammay et al., 2022). Iterative methods offer high flexibility and adaptability, making them widely used in EM LWD data inversion. However, these methods are highly sensitive to the choice of initial models and require substantial computational resources for forward modeling, derivative calculations, or extensive sampling during the iterative process, leading to significant time consumption. Therefore, there is an urgent need for an efficient inversion method to perform real-time inversion of EM look-ahead LWD data.

In recent years, deep learning (DL) has made remarkable progress and has been increasingly adopted in geophysical applications, owing to its strong ability to model nonlinear relationships and its end-to-end data-driven processing capabilities (Yu and Ma, 2021). Notably, DL has shown great promise in the inversion of EM LWD data. Existing deep learning-based inversion networks for EM LWD data can be broadly categorized into two types. The first involves long short-term memory (LSTM) networks, which have been effectively employed for the inversion of azimuthal EM LWD measurements (Fan et al., 2022; Kang et al., 2023a). The second category includes convolutional neural network (CNN), which, compared with LSTM models, exhibit stronger spatial feature extraction capabilities and have seen broader adoption in EM LWD inversion tasks. Among CNN architectures, U-net and residual neural networks (ResNets) have emerged as particularly effective due to their powerful feature extraction ability and training stability, making them suitable for inverting both azimuthal and deep-reading azimuthal EM measurements (Noh et al., 2021; Kang et al., 2023b; Sun et al., 2024). To enhance the interpretability and physical plausibility of DL-based inversion, recent studies have introduced physics-informed training strategies. These approaches incorporate prior physical knowledge into the learning process, resulting in better generalization and physically consistent outputs (Jin et al., 2019; Zhao et al., 2024a, 2024b). Such physics-driven DL inversion frameworks are predominantly based on CNN architectures, including fully convolutional networks (FCNs) and ResNets. Collectively, these studies demonstrate that deep learning offers a powerful alternative for EM LWD inversion, combining high computational efficiency with the advantage of eliminating the need for initial model specification.

While DL has demonstrated considerable success in the inversion of conventional EM LWD and azimuthal/deep-reading azimuthal measurements, its application to EM look-ahead LWD data remains largely unexplored. EM look-ahead tools are capable of sensing the electrical properties of unpenetrated formations ahead of the drill bit, which plays a vital role in proactive geosteering and drilling risk mitigation. To bridge this research gap, the present study focuses on DL-based inversion of EM look-ahead LWD measurements. Specifically, we propose a multi-task U-net architecture designed to improve both inversion accuracy and computational efficiency. The study begins by constructing anisotropic geological models featuring four stratigraphic interfaces. For each type of EM look-ahead signal component, synthetic datasets are generated using the generalized reflection coefficient method. These datasets are then employed to train separate multi-task U-net models, yielding task-specific inversion networks optimized for different signal components. The trained networks are subsequently evaluated on simulated test data to assess their accuracy, robustness, and look-ahead detection capability. Finally, the inversion performance of the proposed multi-task U-net approach is benchmarked against that of the conventional Levenberg-Marquardt algorithm to further demonstrate its effectiveness.

2. Methods

2.1. EM look-ahead LWD instrument

This study focuses on the instrument developed by the Chinese Academy of Sciences (Liang et al., 2023). A schematic of its antenna system is illustrated in Fig. 1. The instrument features a single-transmitter, dual-receiver configuration, in which both the

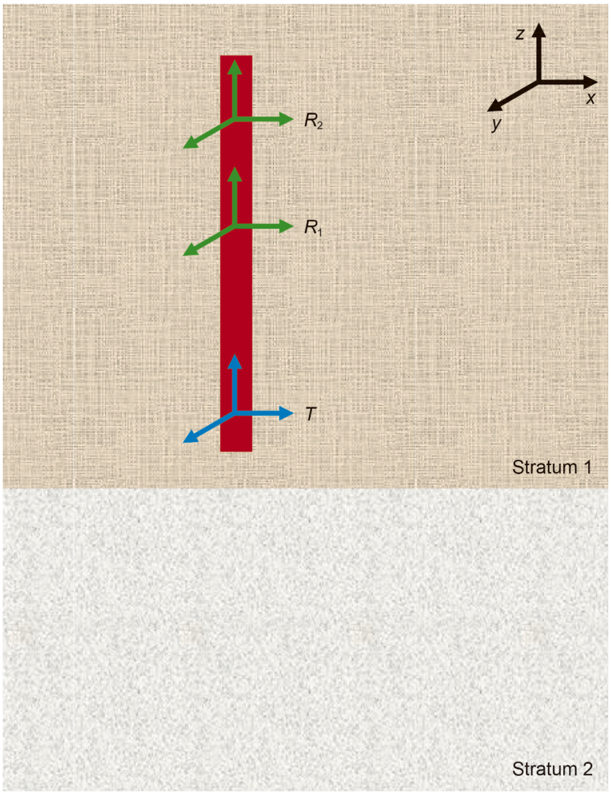


Fig. 1. Schematic illustration of the EM look-ahead LWD instrument.

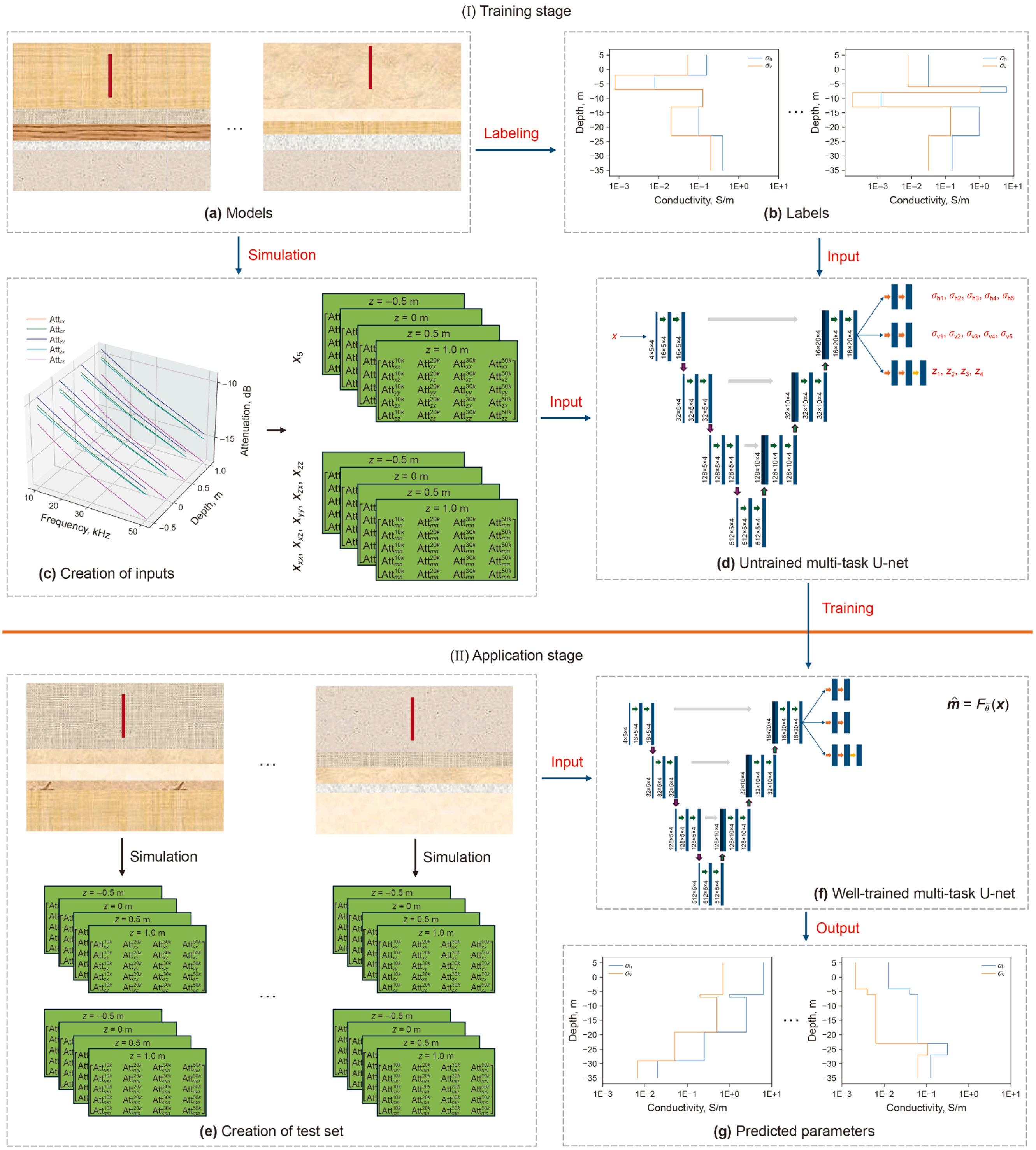


Fig. 2. Workflow of multi-task U-net for inverting EM look-ahead LWD data to extract geological information ahead of the drill bit. The process includes two phases: the training phase (**a**–**d**) and the application phase (**e**–**g**).

transmitter and receivers utilize three-axis orthogonal coils that can be modeled as magnetic dipoles. In Fig. 1, the blue orthogonal arrows represent the transmitting magnetic dipoles (T), while the green arrows denote the receiving magnetic dipoles (R_1 and R_2). The dipole orientations are indicated by black arrows. The distances between the transmitter and the two receivers are $TR_1 = 10$ m and $TR_2 = 14$ m, respectively.

The three-axis orthogonal transmitting and receiving coils allow for measurements of nine components of the induced voltage, forming a 3 × 3 voltage response tensor $\boldsymbol{V}$, defined as:

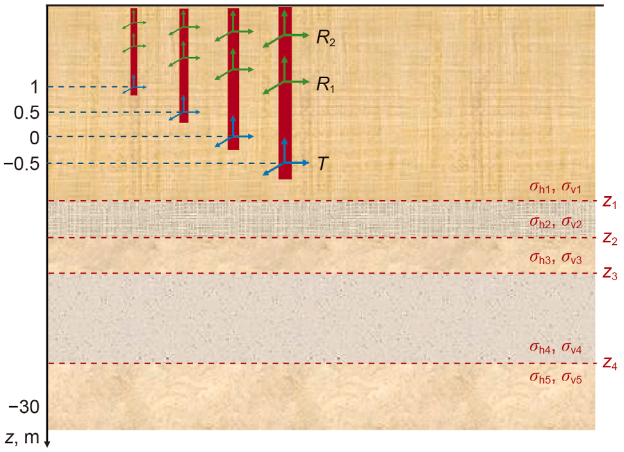


Fig. 3. Schematic diagram of the four-interface geological model structure.

Table 1
Dataset split.

Training set	Validation set	Test set
486,000	54,000	60,000

$$\boldsymbol{V} = \begin{bmatrix} V_{xx} & V_{xy} & V_{xz} \\ V_{yx} & V_{yy} & V_{yz} \\ V_{zx} & V_{zy} & V_{zz} \end{bmatrix}, \tag{1}$$

where the first subscript denotes the orientation of the receiving coil and the second denotes that of the transmitting coil. Based on this, the attenuation ratio Att_{ij} between the two receiver positions R_1 and R_2 can be calculated for each component as:

$$\mathrm{Att}_{ij} = -20\,\mathrm{lg}\frac{\left|V_{ij}^{R_1}\right|}{\left|V_{ij}^{R_2}\right|}, i,j \in \{x,y,z\}, \tag{2}$$

where $|\cdot|$ denotes the magnitude of a complex number, and the superscripts R_1 and R_2 indicate the receiver positions.

2.2. Inversion theory

The Att of the EM look-ahead LWD instrument does not directly represent the formation properties ahead of the drill bit and must be inverted to extract subsurface information. The inversion is inherently ill-posed and highly nonlinear, often admitting multiple solutions. The goal of inversion is to estimate the subsurface electrical properties $\boldsymbol{m}$ from the measurements $\boldsymbol{d}$, as defined by:

$$\widehat{\boldsymbol{m}} = I(\boldsymbol{d}), \tag{3}$$

where $I(\cdot)$ denotes the inversion operator. The objective is to ensure that the estimated model $\widehat{\boldsymbol{m}}$ closely approximates the true formation parameters $\boldsymbol{m}$.

EM look-ahead LWD data inversion typically involves estimating multiple subsurface parameters, which can be formulated as a classical nonlinear least-squares optimization problem. To address this, conventional inversion algorithms employ iterative schemes to progressively approximate the true solution. Among these methods, the LM algorithm is a widely used and effective approach for solving nonlinear least-squares problems, as it combines the robustness of gradient descent with the efficiency of the Gauss-Newton method. When the current model estimate is far from the true solution, the LM algorithm behaves predominantly like gradient descent, yielding stable updates. As the estimate progressively approaches the optimal solution, it transitions

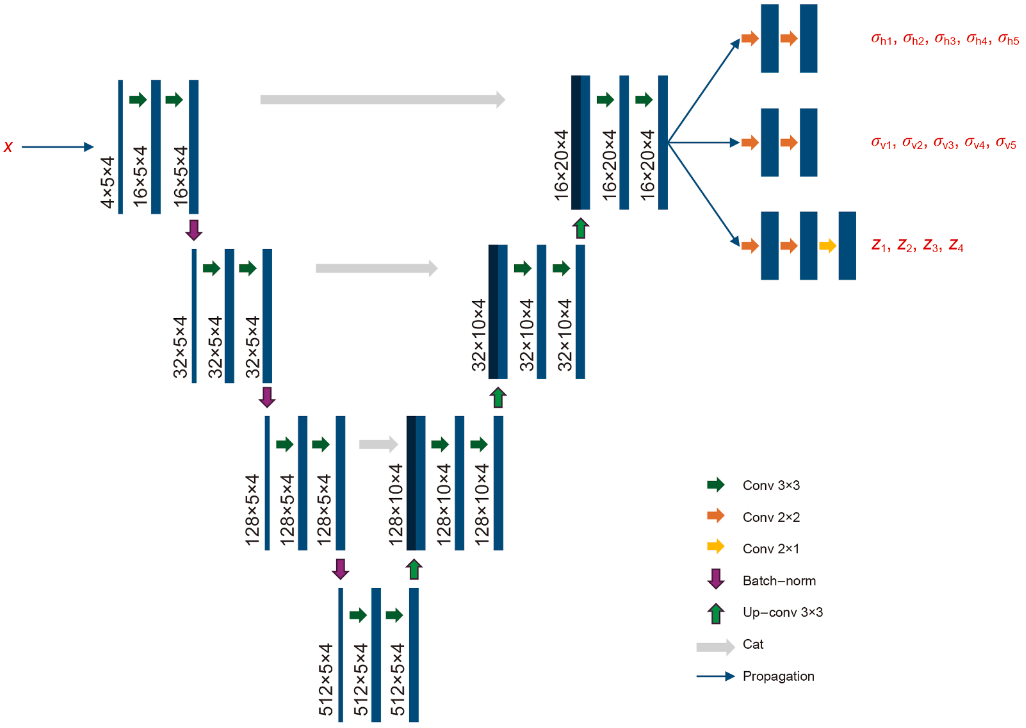


Fig. 4. Structure diagram of the multi-task U-net neural network.

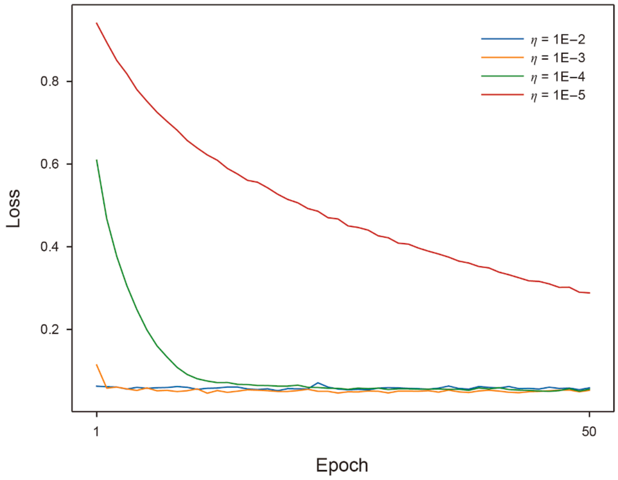


Fig. 5. Loss function decay curves under different learning rates. η represents the learning rate.

toward a Gauss-Newton update, thereby accelerating convergence and improving precision.

The objective function is defined as the weighted sum of squared differences between the observed data $\boldsymbol{d}$ and the simulated data $\widehat{\boldsymbol{d}}$, often referred to as the Chi-squared error criterion (Gavin, 2019). It is expressed as:

$$\chi^2(\widehat{\boldsymbol{m}}) = \sum_{i=1}^{n}\left[\frac{d_i - \widehat{d}_i}{\sigma_{d_i}}\right]^2 = (\boldsymbol{d}-\widehat{\boldsymbol{d}})^T \boldsymbol{W}(\boldsymbol{d}-\widehat{\boldsymbol{d}}), \tag{4}$$

$$\widehat{\boldsymbol{d}} = F(\widehat{\boldsymbol{m}}). \tag{5}$$

Here, n represents the length of both the observed data vector $\boldsymbol{d}$ and the simulated data vector $\widehat{\boldsymbol{d}}$, and σ_{d_i} denotes the standard deviation of the measurement error associated with the i th data point. The weighting matrix $\boldsymbol{W}$ is a diagonal matrix with entries $W_{ii} = \sigma_{d_i}^2$. The function F represents the forward modeling process. The objective of the LM method is to minimize the Chi-squared error χ^2, which necessitates iterative computations. In each iteration, the goal is to determine a perturbation $\boldsymbol{h}$ to the parameter $\widehat{\boldsymbol{m}}$ such that χ^2 is reduced. The LM method typically computes the perturbation $\boldsymbol{h}$ using the following equation (Gavin, 2019):

$$\left[\boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{J}+\lambda I\right]\boldsymbol{h} = \boldsymbol{J}^{\mathrm{T}}\boldsymbol{W}(\boldsymbol{d}-\widehat{\boldsymbol{d}}), \tag{6}$$

where $\boldsymbol{J}$ is the Jacobian matrix, and λ is the regularization term. During the kth iteration, if the updated inversion parameters satisfy the acceptance criteria, the parameter $\widehat{\boldsymbol{m}}$ is updated as follows: $\widehat{\boldsymbol{m}} = \widehat{\boldsymbol{m}} + \boldsymbol{h}$.

The DL inversion algorithm is a data-driven approach, distinct from iterative inversion methods. The deep neural network is represented as $F_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ denotes the parameters of the network that need to be learned. The objective of the neural network is to update the parameters $\boldsymbol{\theta}$ such that $F_{\boldsymbol{\theta}}$ can closely approximate the true and complex mapping G between the parameters $\boldsymbol{m}$ and the data $\boldsymbol{d}$. During the training phase, for a mini-batch dataset $\Omega = \{(\boldsymbol{d}_i, \boldsymbol{m}_i), i = 1, 2, ..., N\}$, which contains N samples, the model parameters are updated using the chain rule. The update rule is expressed by the following equation:

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \eta \frac{1}{N}\sum_{i=1}^{N}\frac{\partial L(F_{\boldsymbol{\theta}}(\boldsymbol{d}_i), \boldsymbol{m}_i)}{\partial \boldsymbol{\theta}}, \tag{7}$$

where η represents the learning rate, and L is the loss function to be minimized. Once the neural network is trained, it can perform end-to-end inversion of the data as follows:

$$\widehat{\boldsymbol{m}} = F_{\overline{\boldsymbol{\theta}}}(\boldsymbol{d}), \tag{8}$$

where $\overline{\boldsymbol{\theta}}$ represents the optimal neural network parameters learned during the training process. In contrast to iterative methods, the DL approach does not require iterations during inversion, and the inversion results can be obtained directly.

3. Deep learning inversion

3.1. Inversion process

The overall framework for inverting the electrical properties of the forward formation using a multi-task U-net neural network is illustrated in Fig. 2. It comprises two primary stages: training and application. In the training phase, a set of synthetic geological models is first constructed to represent various formation scenarios, forming the model dataset (Fig. 2(a)). From each model, the interface positions, along with the horizontal and vertical conductivities of all layers, are extracted as labeled data (Fig. 2(b)). Forward modeling is then performed to compute the Att of different electromagnetic components, producing six types of input data: five single-component inputs and one multi-component input (Fig. 2(c)). These inputs, along with the corresponding labels, are used to train the neural network (Fig. 2(d)). Through this process, the network learns the complex nonlinear relationships between multi-parameter geological properties and amplitude ratios, resulting in a trained inversion model (Fig. 2(f)). In the application phase, the trained network can directly predict formation parameters from new input data without the need for iterative optimization, enabling efficient, end-to-end multi-parameter inversion (Fig. 2(g)).

3.2. Error measurement metrics

To quantitatively assess the accuracy of the inversion results obtained from the trained neural network, we employ a set of residual-based error metrics tailored to the different types of predicted parameters. These residual metrics are defined as:

$$\mathrm{res}_{\sigma_j} = \frac{1}{5}\sum_{i=1}^{5}\lg\left(\sigma_{\mathrm{t}}^{i}\right) - \lg\left(\sigma_{\mathrm{p}}^{i}\right), \tag{9}$$

$$\mathrm{res}_{z} = \frac{1}{4}\sum_{i=1}^{4} z_{\mathrm{t}}^{i} - z_{\mathrm{p}}^{i}, \tag{10}$$

$$\mathrm{res}_{\lambda} = \frac{1}{5}\sum_{i=1}^{5} \lambda_{\mathrm{t}}^{i} - \lambda_{\mathrm{p}}^{i}, \tag{11}$$

In addition to these absolute metrics, we also compute relative errors to quantify the normalized deviation between predicted and true values. The relative errors are defined as:

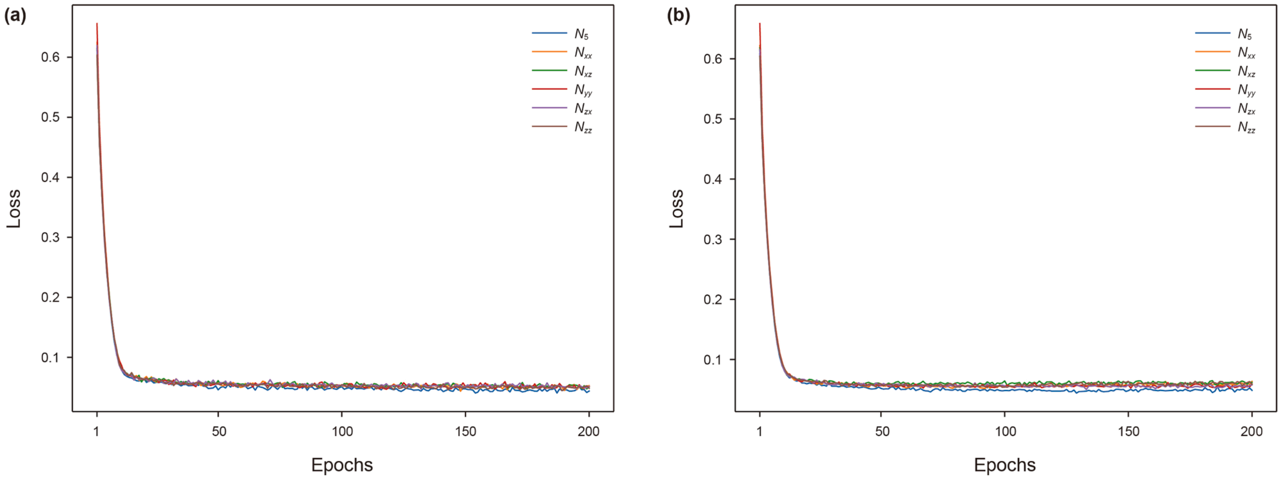


Fig. 6. Loss function error decay curves. (**a**) training set, and (**b**) validation set.

Table 2
Comparisons of training time.

Training time, h					
N_5	N_{xx}	N_{xz}	N_{yy}	N_{zx}	N_{zz}
1.19	1.42	1.41	1.42	1.40	1.45

Table 3
The mean of the residual of parameters.

	$\mathrm{res}_{\sigma_\mathrm{h}}$, S/m	$\mathrm{res}_{\sigma_\mathrm{v}}$, S/m	res_λ	res_z, m
N_5	0.0237	0.0179	−0.1003	−0.0443
N_{xx}	0.0340	0.0433	−0.1395	−0.2866
N_{xz}	0.0030	0.0007	−0.1125	−0.0853
N_{yy}	−0.0261	−0.0311	−0.1042	−0.0405
N_{zx}	0.0017	−0.0029	−0.1062	0.1084
N_{zz}	0.0183	0.0202	−0.1223	−0.0353

Table 4
Residual distribution statistics (Unit: %).

	$\|\mathrm{res}_{\sigma_\mathrm{h}}\|$				$\|\mathrm{res}_{\sigma_\mathrm{v}}\|$			
	≤ 0.1	≤ 0.2	≤ 0.4	≤ 0.6	≤ 0.1	≤ 0.2	≤ 0.4	≤ 0.6
N_5	38.9	68.7	94.5	99.3	33.4	61.4	90.9	98.5
N_{xx}	33.7	61.3	91.0	98.8	29.7	55.2	86.5	97.3
N_{xz}	31.3	58.0	89.4	98.6	27.7	52.3	84.4	96.6
N_{yy}	34.0	61.2	90.8	98.8	29.8	55.6	86.9	97.5
N_{zx}	31.7	58.5	89.8	98.7	27.6	52.3	84.6	96.8
N_{zz}	33.2	61.2	91.2	98.7	28.7	54.0	86.1	97.2
	$\|\mathrm{res}_\lambda\|$				$\|\mathrm{res}_z\|$			
	≤ 0.1	≤ 0.2	≤ 0.4	≤ 0.6	≤ 1	≤ 2	≤ 3	≤ 5
N_5	24.1	46.2	79.4	95.2	38.2	67.9	87.4	99.8
N_{xx}	22.3	43.0	75.8	93.3	36.4	65.3	85.2	99.5
N_{xz}	22.9	44.5	77.4	94.1	36.7	65.6	85.4	99.7
N_{yy}	23.25	45.1	78.1	94.4	36.4	65.5	85.2	99.7
N_{zx}	22.9	44.8	77.8	94.3	36.7	65.6	85.5	99.7
N_{zz}	22.7	43.9	76.7	93.8	37.5	66.7	86.4	99.8

$$\mathrm{RE}_{\sigma_j} = \frac{1}{5}\sum_{i=1}^{5}\frac{\lg\left(\sigma_\mathrm{t}^i\right) - \lg\left(\sigma_\mathrm{p}^i\right)}{\lg\left(\sigma_\mathrm{t}^i\right)} \times 100\%, \tag{12}$$

$$\mathrm{RE}_z = \frac{1}{4}\sum_{i=1}^{4}\frac{z_\mathrm{t}^i - z_\mathrm{p}^i}{z_\mathrm{t}^i} \times 100\%, \tag{13}$$

$$\mathrm{RE}_\lambda = \frac{1}{5}\sum_{i=1}^{5}\frac{\lambda_\mathrm{t}^i - \lambda_\mathrm{p}^i}{\lambda_\mathrm{t}^i} \times 100\%, \tag{14}$$

where σ_j represents the horizontal ($j = \mathrm{h}$) or vertical ($j = \mathrm{v}$) conductivity, z denotes the interface depth, and λ denotes the anisotropy coefficient. The subscript t and p refer to the true and predicted model parameters, respectively. The superscript i indicates the layer index.

4. Neural network training

4.1. Creation of the dataset

The first critical step in DL training is the construction of a representative and diverse dataset. In this study, we design a four-interface geological model to approximate the stratigraphy ahead of the drill bit. The logging tool is oriented perpendicular to the formation boundaries, as illustrated in Fig. 3. A total of 14 geological parameters are inverted, defined as follows:

$$\boldsymbol{y} = [\sigma_{\mathrm{h}1}, \sigma_{\mathrm{h}2}, \sigma_{\mathrm{h}3}, \sigma_{\mathrm{h}4}, \sigma_{\mathrm{h}5}, \sigma_{\mathrm{v}1}, \sigma_{\mathrm{v}2}, \sigma_{\mathrm{v}3}, \sigma_{\mathrm{v}4}, \sigma_{\mathrm{v}5}, z_1, z_2, z_3, z_4]. \tag{15}$$

To ensure generalizability, the values of all 14 parameters are randomly sampled for each model instance. In EM look-ahead LWD applications, signal contributions from formations close to the bit are substantially stronger than those from distant formations. To reflect this, a high-resolution sampling strategy is employed near the bit, with layer interfaces spaced every 1 m

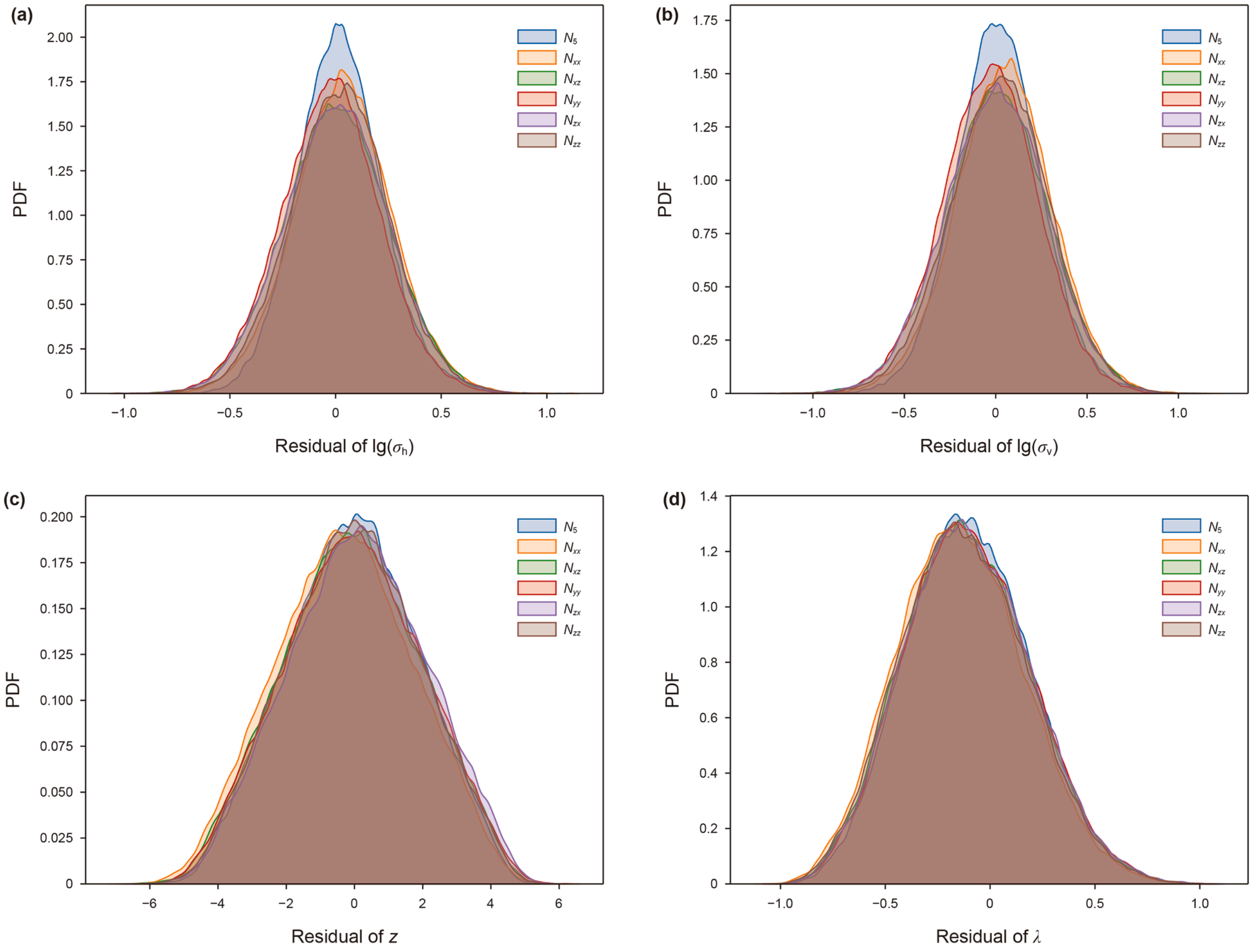


Fig. 7. Probability density distribution of residuals for four parameters. (**a**) horizontal conductivity, (**b**) vertical conductivity, (**c**) layer interface position, and (**d**) anisotropy.

within a 10-m depth interval. In this near-bit region, the logarithmic values of both vertical and horizontal conductivities span from −3 to 1, with a step size of 0.1. Beyond the 10-m range, a coarser sampling scheme is adopted, with interface intervals of 2 m. In this region, conductivity values range from −2 to 1 (log scale), also with a step size of 0.1. The squared anisotropy coefficient is defined as $\lambda^2 = \sigma_h/\sigma_v$, and is sampled from the range [1, 10].

The instrument operates at emission frequencies of 1×10^4, 2×10^4, 3×10^4, and 5×10^4 Hz. The generalized reflection coefficient method (Hong, 2009) and the global amplitude propagator matrix algorithm (Wang et al., 2023) can be used to simulate the three-axis orthogonal induced voltage tensor received by the receivers R_1 and R_2, and the Att is calculated according to Eq. (2). To further improve inversion accuracy, a sliding window inversion approach is adopted, as illustrated in Fig. 3. Four measurement points are used as input, with vertical (z-axis) coordinates of −0.5, 0, 0.5, and 1 m. To analyze the look-ahead capabilities of different components, six types of input data are constructed: one multi-component input $\boldsymbol{x}_5$, and five single-component inputs $\boldsymbol{x}_{xx}$, $\boldsymbol{x}_{xz}$ $\boldsymbol{x}_{yy}$, $\boldsymbol{x}_{zx}$, and $\boldsymbol{x}_{zz}$, defined as follows:

$$\boldsymbol{x}_5 = \{\boldsymbol{x}_i\}_{i=1}^{5}, \tag{16}$$

$$\boldsymbol{x}_{mn} = \left\{\boldsymbol{x}_j\right\}_{j=1}^{5}, \tag{17}$$

Since a sliding window technique is employed, the input data dimension is (4, 5, 4), where i and j denote the measurement points, and $\boldsymbol{x}_i$ or $\boldsymbol{x}_j$ represents the tool response at the corresponding point. The specific forms are given by:

$$\boldsymbol{x}_i = \begin{bmatrix} \mathrm{Att}_{xx}^{10\mathrm{k}} & \mathrm{Att}_{xx}^{20\mathrm{k}} & \mathrm{Att}_{xx}^{30\mathrm{k}} & \mathrm{Att}_{xx}^{50\mathrm{k}} \\ \mathrm{Att}_{xz}^{10\mathrm{k}} & \mathrm{Att}_{xz}^{20\mathrm{k}} & \mathrm{Att}_{xz}^{30\mathrm{k}} & \mathrm{Att}_{xz}^{50\mathrm{k}} \\ \mathrm{Att}_{yy}^{10\mathrm{k}} & \mathrm{Att}_{yy}^{20\mathrm{k}} & \mathrm{Att}_{yy}^{30\mathrm{k}} & \mathrm{Att}_{yy}^{50\mathrm{k}} \\ \mathrm{Att}_{zx}^{10\mathrm{k}} & \mathrm{Att}_{zx}^{20\mathrm{k}} & \mathrm{Att}_{zx}^{30\mathrm{k}} & \mathrm{Att}_{zx}^{50\mathrm{k}} \\ \mathrm{Att}_{zz}^{10\mathrm{k}} & \mathrm{Att}_{zz}^{20\mathrm{k}} & \mathrm{Att}_{zz}^{30\mathrm{k}} & \mathrm{Att}_{zz}^{50\mathrm{k}} \end{bmatrix}, \tag{18}$$

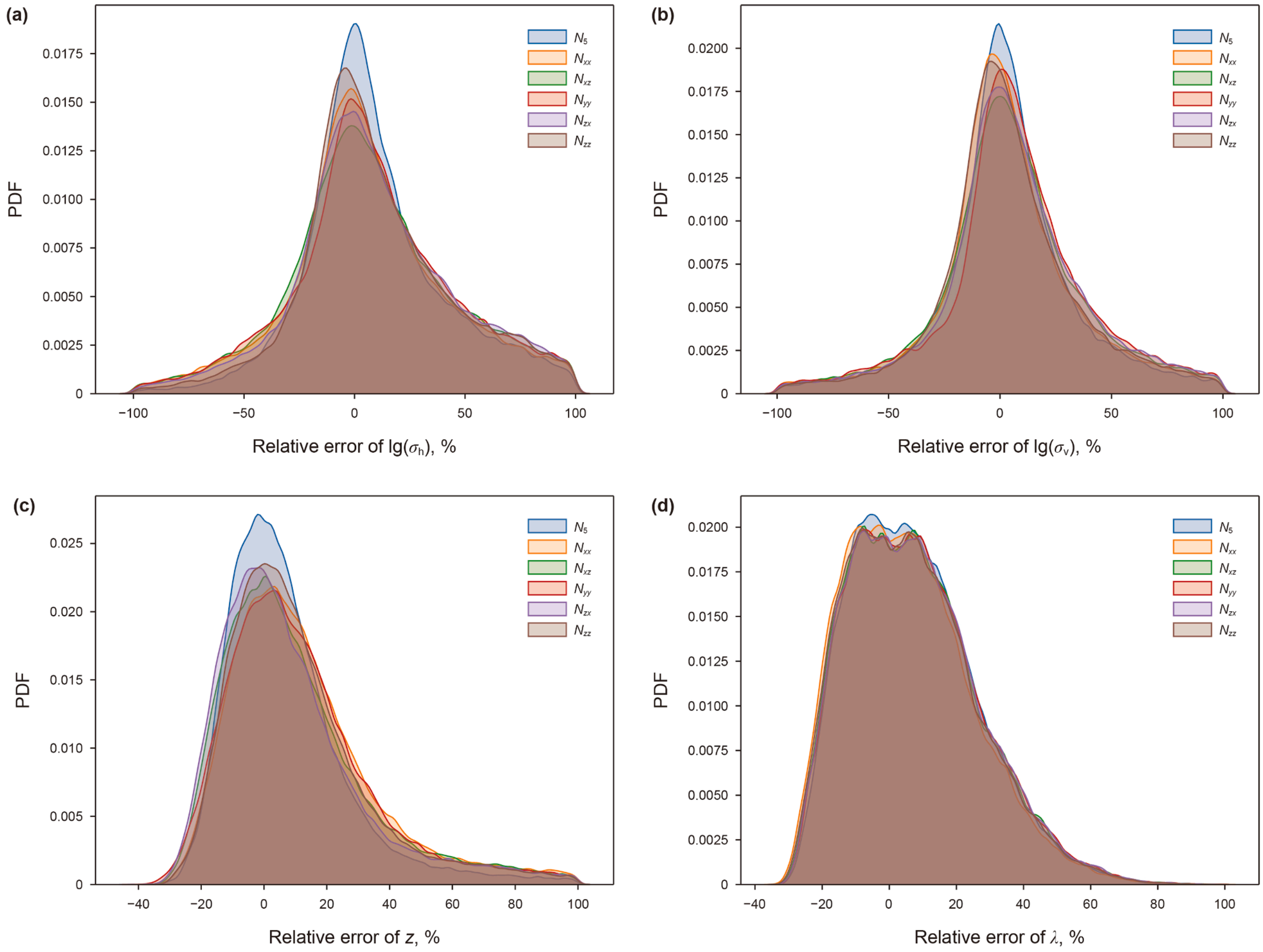


Fig. 8. Probability density distribution of relative error for four parameters. (**a**) horizontal conductivity, (**b**) vertical conductivity, (**c**) layer interface position, and (**d**) anisotropy.

$$\boldsymbol{x}_j = \begin{bmatrix} \mathrm{Att}_{mn}^{10\mathrm{k}} & \mathrm{Att}_{mn}^{20\mathrm{k}} & \mathrm{Att}_{mn}^{30\mathrm{k}} & \mathrm{Att}_{mn}^{50\mathrm{k}} \\ \mathrm{Att}_{mn}^{10\mathrm{k}} & \mathrm{Att}_{mn}^{20\mathrm{k}} & \mathrm{Att}_{mn}^{30\mathrm{k}} & \mathrm{Att}_{mn}^{50\mathrm{k}} \\ \mathrm{Att}_{mn}^{10\mathrm{k}} & \mathrm{Att}_{mn}^{20\mathrm{k}} & \mathrm{Att}_{mn}^{30\mathrm{k}} & \mathrm{Att}_{mn}^{50\mathrm{k}} \\ \mathrm{Att}_{mn}^{10\mathrm{k}} & \mathrm{Att}_{mn}^{20\mathrm{k}} & \mathrm{Att}_{mn}^{30\mathrm{k}} & \mathrm{Att}_{mn}^{50\mathrm{k}} \\ \mathrm{Att}_{mn}^{10\mathrm{k}} & \mathrm{Att}_{mn}^{20\mathrm{k}} & \mathrm{Att}_{mn}^{30\mathrm{k}} & \mathrm{Att}_{mn}^{50\mathrm{k}} \end{bmatrix}. \tag{19}$$

Here, the superscript denotes the emission frequency in Hz, and the subscript represents the component. The *mn* pairs include *xx*, *xz*, *yy*, *zx*, and *zz*. For each type of input, six datasets are generated, each containing 600,000 samples.

It is worth noting that when both the input and label data are scaled to lie within the range of 0–1, the neural network typically demonstrates improved convergence and stability (Singh and Singh, 2020). Therefore, normalization is applied to the data prior to training. The normalization procedure for the input data is described as follows:

$$x_{\mathrm{norm}}^{ij} = \frac{x^{ij} - x_{\min}^{ij}}{x_{\max}^{ij} - x_{\min}^{ij}}, i = 1,2,3,4,5; \ j = 1,2,3,4, \tag{20}$$

where x^{ij} is an element in the input data $\boldsymbol{x}$, and x_{norm}^{ij} is the corresponding element in the normalized input data. *i* and *j* represent the row and column indices, respectively.

For the label data, due to the significantly different orders of magnitude between conductivity and boundary positions, distinct normalization formulas are applied. Specifically, for conductivity, the normalization formula is as follows:

$$\sigma_{\mathrm{norm}}^{k} = \frac{\lg\left(\sigma^{k}\right) - \lg\left(\sigma_{\min}^{k}\right)}{\lg\left(\sigma_{\max}^{k}\right) - \lg\left(\sigma_{\min}^{k}\right)}, k = 1,2,3,4,5. \tag{21}$$

The normalization formula for the layer interface position is given by:

$$z_{\mathrm{norm}}^{k} = \frac{z^{k} - z_{\min}^{k}}{z_{\max}^{k} - z_{\min}^{k}}, k = 1,2,3,4. \tag{22}$$

where the superscript *k* represents the interface index. The subscripts min and max denote the minimum and maximum values, respectively. The subscript norm indicates the normalized data.

To effectively mitigate overfitting and assess the neural network's ability to generalize to unseen data, the dataset is divided into three subsets: a training set, a validation set, and a test set. The training set is used to learn the network's parameters, the validation set is employed to monitor and avoid overfitting during training, and the test set is reserved for the final evaluation of the model's generalization performance. The sample size ratio for the three subsets is 0.81: 0.09: 0.1. The number of samples in each subset is presented in Table 1.

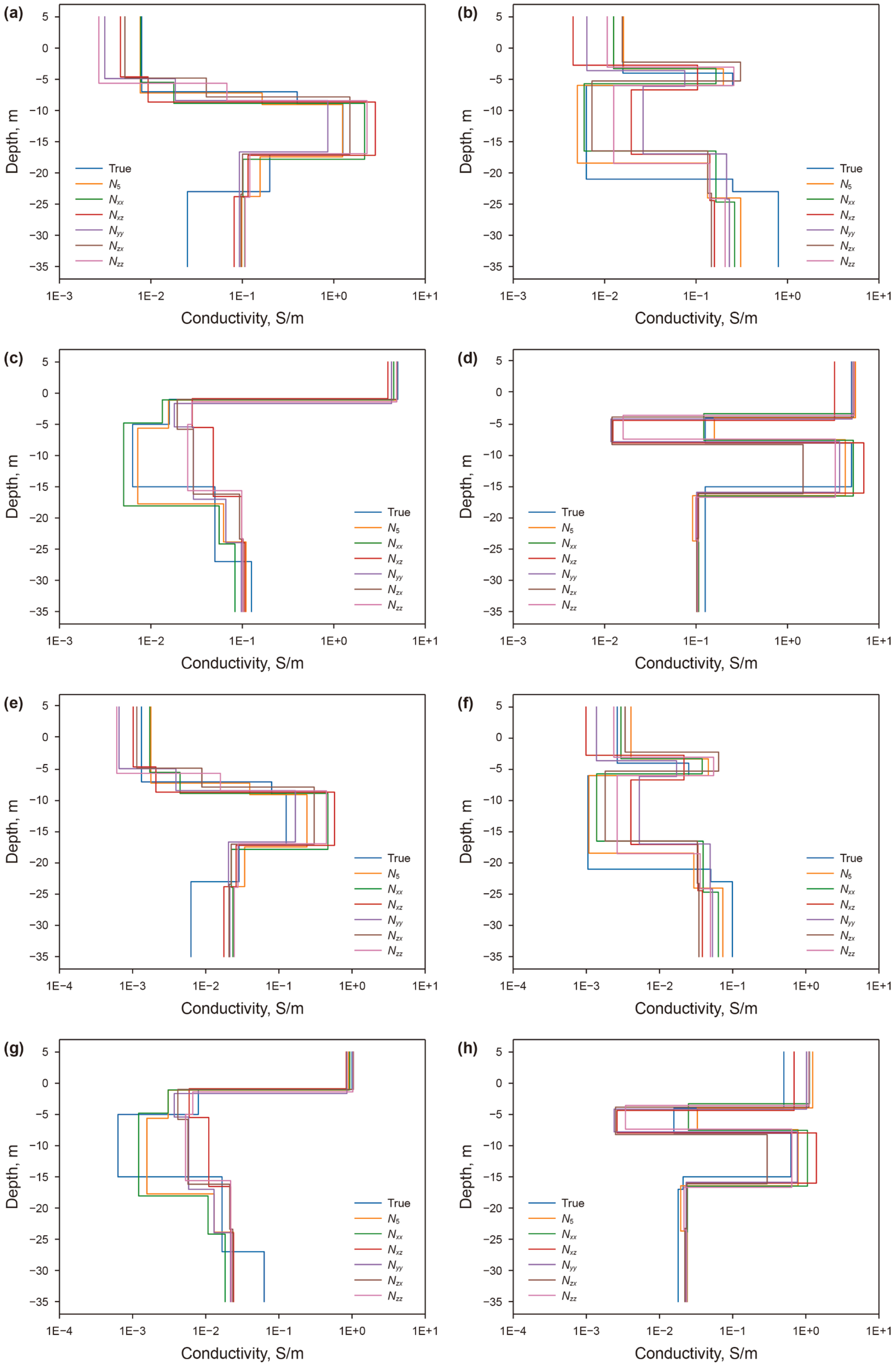


Fig. 9. Comparison of horizontal (**a**–**d**) and vertical (**e**–**h**) conductivity inversion results. Subfigures in the same column depict the same model.

Prior to the interpretation of prediction results, the normalized outputs from the neural network require denormalization to reconstruct the physical measurements. The denormalization procedures for formation conductivity and bed boundary positions are implemented as:

$$\lg\left(\hat{\sigma}^k\right) = \lg\left(\hat{\sigma}^k_{\mathrm{out}}\right)\left(\lg\left(\sigma^k_{\mathrm{max}}\right) - \lg\left(\sigma^k_{\mathrm{min}}\right)\right) + \lg\left(\sigma^k_{\mathrm{min}}\right), \tag{23}$$

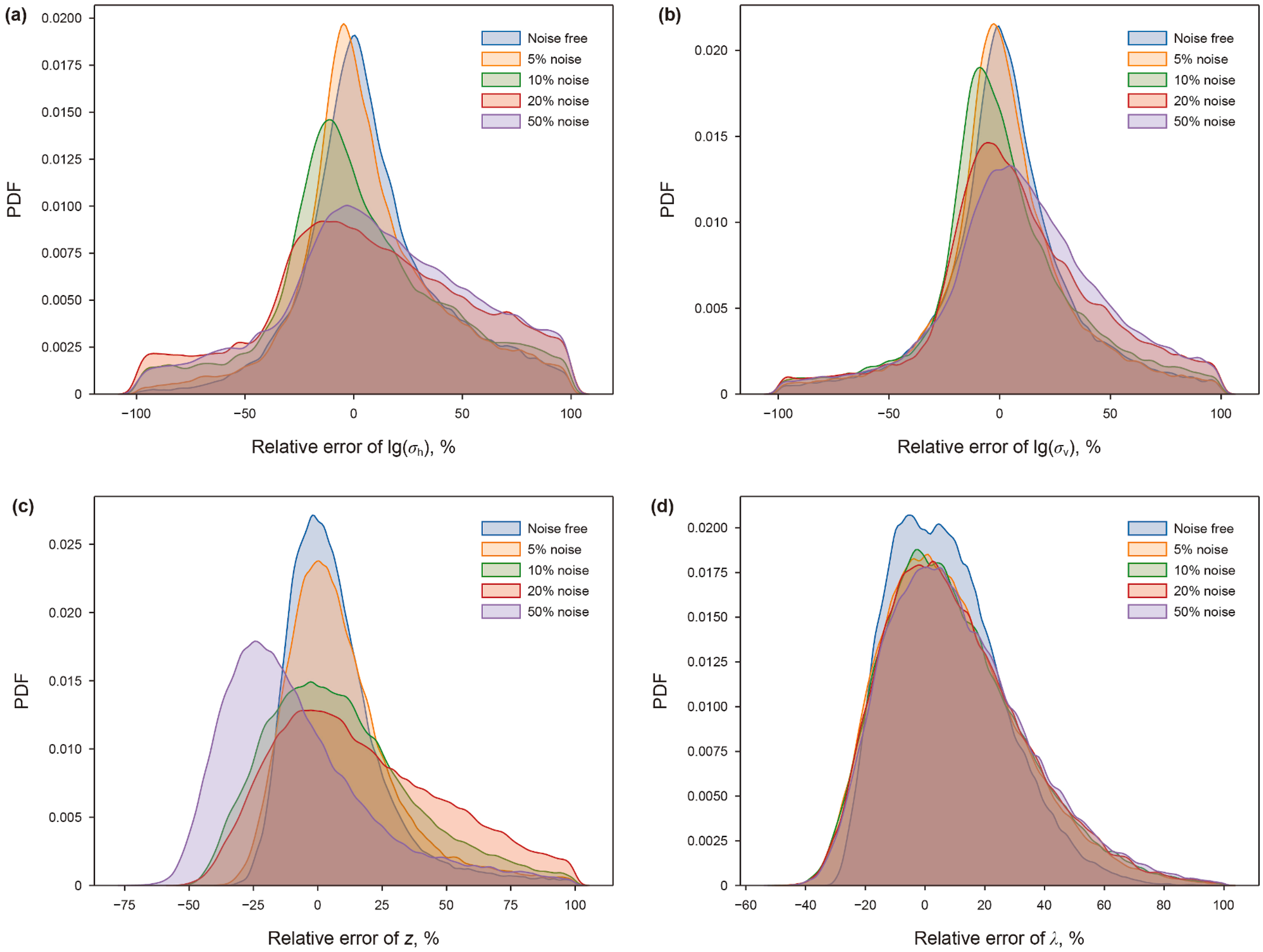


Fig. 10. Probability density distribution of relative error for four parameters. (**a**) horizontal conductivity, (**b**) vertical conductivity, (**c**) layer interface position, and (**d**) anisotropy.

$$\hat{z}^k = \hat{z}^k_{\text{out}}\left(z^k_{\max} - z^k_{\min}\right) + z^k_{\min}, \tag{24}$$

where $\hat{\sigma}^k_{\text{out}}$ and $\hat{z}^k_{\text{out}}$ represent the conductivity and interface depth directly output by the neural network, respectively. The terms $\hat{\sigma}^k$ and $\hat{z}^k$ correspond to their denormalized counterparts, representing the final inversion results.

4.2. Structure of the multi-task U-net

The model's inversion parameters $\boldsymbol{y}$ are defined in Eq. (12), consisting of three distinct physical quantities: horizontal conductivity, vertical conductivity, and layer boundary position. Due to their fundamentally different physical characteristics, we developed a multi-task U-net architecture for simultaneous inversion of these parameters. As shown in Fig. 4, the network's core structure adopts the U-net framework, while its sub-task component comprises three dedicated neural pathways each specifically handling the inversion of horizontal conductivity, vertical conductivity, and layer interface positions, respectively. The network thus produces three corresponding outputs as follows:

$$\boldsymbol{y}_1 = [\sigma_{\text{h}1}, \sigma_{\text{h}2}, \sigma_{\text{h}3}, \sigma_{\text{h}4}, \sigma_{\text{h}5}], \tag{25}$$

$$\boldsymbol{y}_2 = [\sigma_{\text{v}1}, \sigma_{\text{v}2}, \sigma_{\text{v}3}, \sigma_{\text{v}4}, \sigma_{\text{v}5}], \tag{26}$$

$$\boldsymbol{y}_3 = [z_1, z_2, z_3, z_4], \tag{27}$$

The overall output $\boldsymbol{y}$ of the neural network can be represented as

$$\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3] \tag{28}$$

4.3. Determination of hyperparameters

To determine the optimal hyperparameters for the neural network, the method of controlling variables was utilized. Taking the learning rate determination as an example, the process is explained as follows. By keeping the neural network structure and other hyperparameters fixed, we adjusted only the learning rate and observed the decay curves of the loss function under different learning rates. The selected learning rates were 1×10^{-2}, 1×10^{-3}, 1×10^{-4}, and 1×10^{-5}. The loss function decay curves for these learning rates are shown in Fig. 5. The results indicate that when the learning rate is set to 1×10^{-2} and 1×10^{-3}, the loss function fails to converge. However, as the learning rate decreases, better convergence behavior is observed when a learning rate is 1×10^{-4}. Further reduction of the learning rate results in slower convergence. Therefore, we chose 1×10^{-4} as the final learning rate.

Other hyperparameters were determined using the same method. The batch size is 512, and the number of training epochs is 200. The multi-task U-net uses the hyperbolic tangent function

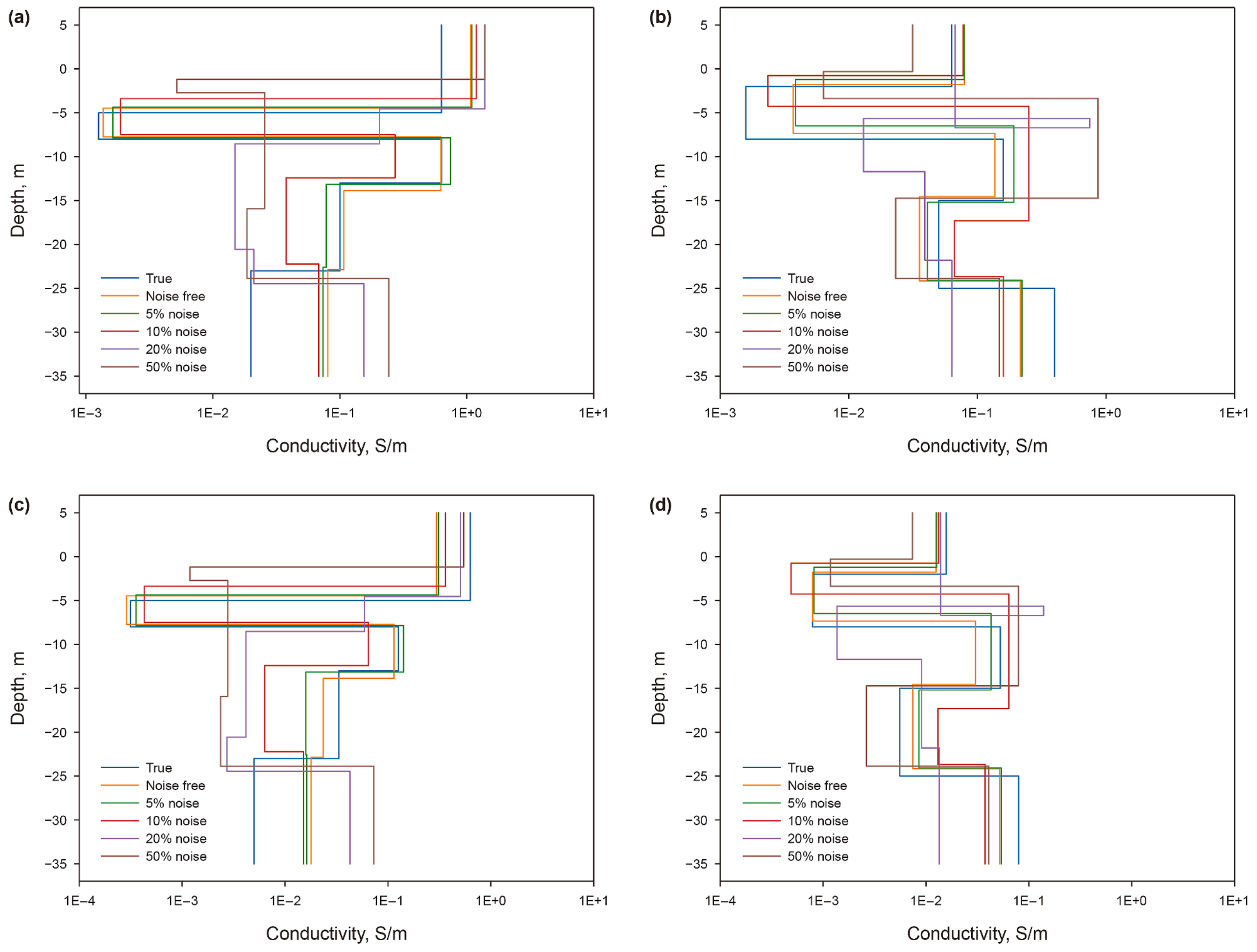


Fig. 11. Comparison of horizontal (**a**, **b**) and vertical (**c**, **d**) conductivity inversion results under different noise levels. Subfigures in the same column depict the same model.

Table 5
The mean of the residual of parameters.

	res_{σ_h}, S/m	res_{σ_v}, S/m	res_{λ}	res_z, m
LM	−0.5288	0.4204	1.7928	1.8000
N_5	0.0236	0.0199	−0.1044	−0.0004

(Tanh) as the activation function. To prevent overfitting and enhance the model's generalization ability, a regularization term is added to the L2 loss function. For a normalized data-label pair $\left\{\left(\boldsymbol{x}^i_{\mathrm{norm}}, \boldsymbol{y}^i_{\mathrm{norm}}\right), i = 1, \ldots, n-1, n\right\}$ with batch size n, the loss function is defined as:

$$L = \frac{1}{3}\frac{1}{n}\frac{1}{p}\sum_{j=1}^{3}\sum_{i=1}^{n}\sum_{k=1}^{p}\left[\left(\widehat{\boldsymbol{y}}^i_{\mathrm{norm}}[j][k] - \boldsymbol{y}^i_{\mathrm{norm}}[j][k]\right) + \frac{\alpha}{2}\|\boldsymbol{\theta}\|_2^2\right],$$
$$p = 4, 5, \tag{29}$$

where j represents the output of the jth sub-task, and k denotes the number of parameters in the sub-task output (either 4 or 5). $\boldsymbol{\theta}$ represents the parameters of the neural network. α is introduced to prevent overfitting. $\widehat{\boldsymbol{y}}^i_{\mathrm{norm}} = F_\theta\left(\boldsymbol{x}^i_{\mathrm{norm}}\right)$ is the output of the neural network, where $F_{\boldsymbol{\theta}}$ denotes the mapping defined by the neural network. In addition, the Adam optimization algorithm is used to improve the training efficiency of the neural network (Kingma, 2014). The neural network construction is carried out using the PyTorch-GPU framework (Paszke et al., 2019), with training and testing performed on an Nvidia GeForce RTX 3090 GPU.

4.4. Neural network training process and results

The neural network is trained using the error backpropagation algorithm. The pseudocode of the multi-task U-net is described in Algorithm 1.

Algorithm 1. Pseudocode of multi-task U-net.

```
Require: Training set $\mathbf{D}_{\mathrm{train}} = \left\{\left(\boldsymbol{x}^i_{\mathrm{norm}}, \boldsymbol{y}^i_{\mathrm{norm}}\right), i = 1, \ldots, N\right\}$
Require: Learning rate $\eta$, neural network parameters $\boldsymbol{\theta}$, batch size $n$, loss
    function $L$ (Eq. (29)).
Initialize the parameters $\boldsymbol{\theta}$.
While the stopping criterion is not met do
    Sample a minibatch of $n$ examples from the training set $\{\boldsymbol{x}^1_{\mathrm{norm}}, \cdots, \boldsymbol{x}^n_{\mathrm{norm}}\}$
        with corresponding labels $\{\boldsymbol{y}^1_{\mathrm{norm}}, \cdots, \boldsymbol{y}^n_{\mathrm{norm}}\}$.
    Compute loss: $L$.
    Compute gradient: $g = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial L\left(F_\theta\left(\boldsymbol{x}^i_{\mathrm{norm}}\right), \boldsymbol{y}^i_{\mathrm{norm}}\right)}{\partial \boldsymbol{\theta}}$.
    Compute update: $\Delta\boldsymbol{\theta} = -\eta g$.
    Update parameters: $\boldsymbol{\theta} = \boldsymbol{\theta} + \Delta\boldsymbol{\theta}$.
End
Return $\boldsymbol{\theta}$
```

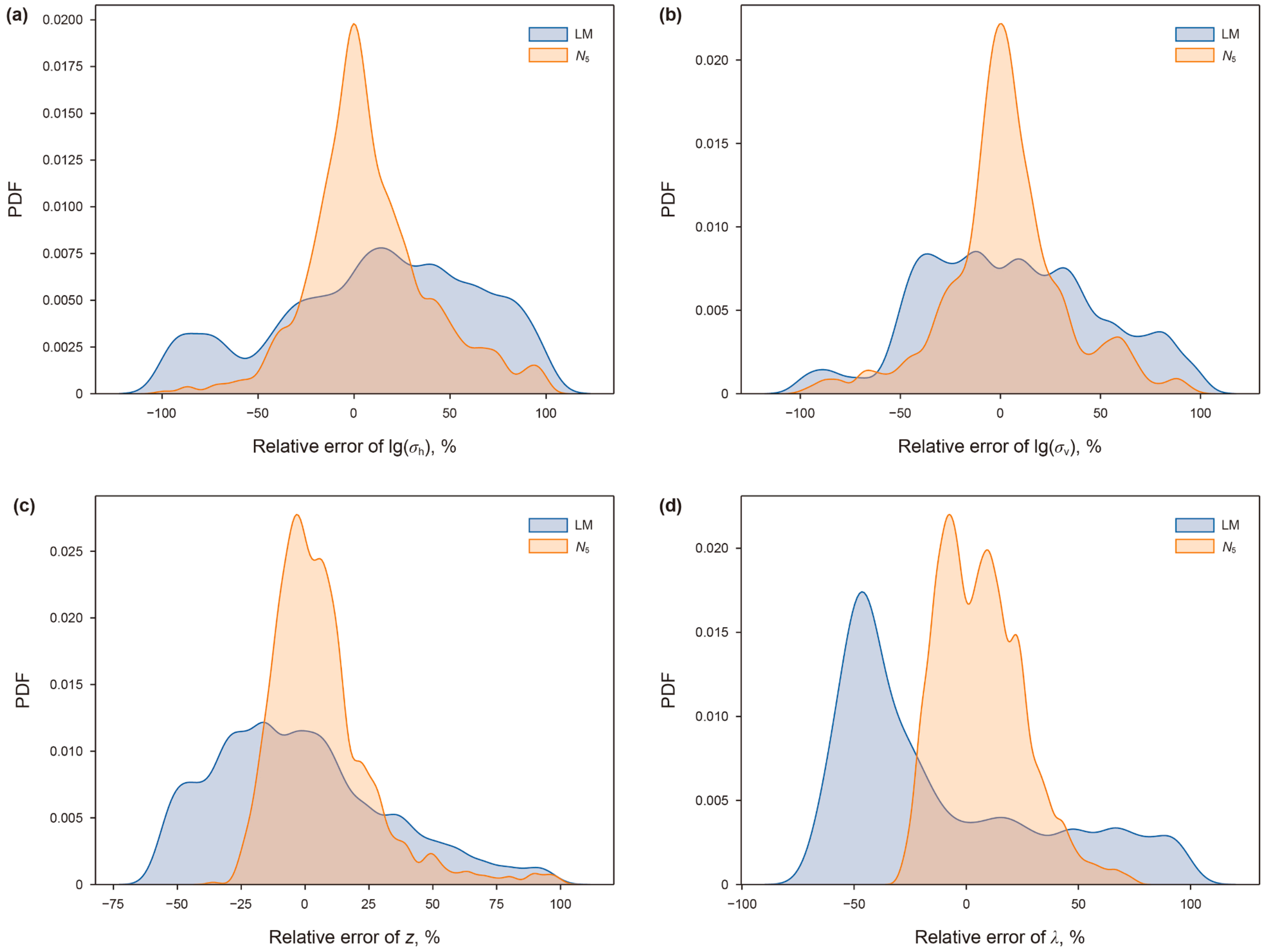


Fig. 12. Probability density distribution of relative error for LM method and DL inversion method. (**a**) horizontal conductivity, (**b**) vertical conductivity, (**c**) layer interface position, (**d**) anisotropy.

Six different datasets were used to train the same initial multi-task U-net, resulting in six distinctly trained networks, denoted as N_{xx}, N_{xz}, N_{yy}, N_{zx}, N_{zz}, and N_5. Fig. 6(a) and (b) depict the loss evolution of six independently trained neural networks on the training and validation datasets, respectively. In both subplots, all networks exhibit a rapid loss decrease within the first 20 epochs, followed by a stable convergence phase. This sharp initial drop indicates that the models effectively captured the dominant patterns of the data early in training. After approximately 50 epochs, the loss values plateau, suggesting that the networks entered a fine-tuning stage where minor adjustments are made to refine performance. Training and validation curves remain closely aligned over 200 epochs, with no divergence, confirming good generalization and absence of overfitting. Overall, the process is stable and effective, ensuring reliable performance in inversion tasks.

The training durations of the six neural networks are shown in Table 2, and their training times are approximately equivalent.

4.5. Neural network inversion performance

During the testing phase, the six trained neural networks were evaluated on the same test set containing 60,000 samples. The inversion times for all neural networks range from 14 to 15 s. This indicates that multi-task U-net is highly efficient in processing large-scale data and can meet the real-time inversion requirements of geological guidance tasks.

To comprehensively evaluate the inversion accuracy of the neural networks, we calculated the absolute errors between the predicted parameters and the true parameters for the six neural networks. The error calculation formulas are given in Eqs. (9–11). First, a statistical analysis of residuals was performed for different parameters. The mean residuals for each neural network are presented in Table 3. As shown, the mean residuals for all models are close to zero, indicating accurate inversion performance. Additionally, Table 4 presents the cumulative distribution of inversion residuals within different intervals. A higher percentage within a given range suggests greater overall inversion accuracy. By comparing the distributions, the inversion results can be categorized into three levels. The best performance across all parameters is achieved by N_5. For conductivity, N_{xx}, N_{yy}, and N_{zz} outperform N_{xz} and N_{zx}, indicating that the coaxial and co-planar are more sensitive to the formation conductivity ahead of the drill bit compared to the cross-components. However, for the anisotropic parameter, the inversion accuracy of N_{yy} is higher than that of the other four neural networks, suggesting that the strata the *yy* component is more sensitive to the anisotropy of the strata. For interface depth estimation, the inversion accuracy of the cross-components and coaxial components is superior to that of the co-planar components. Since N_5 incorporates all components, it effectively

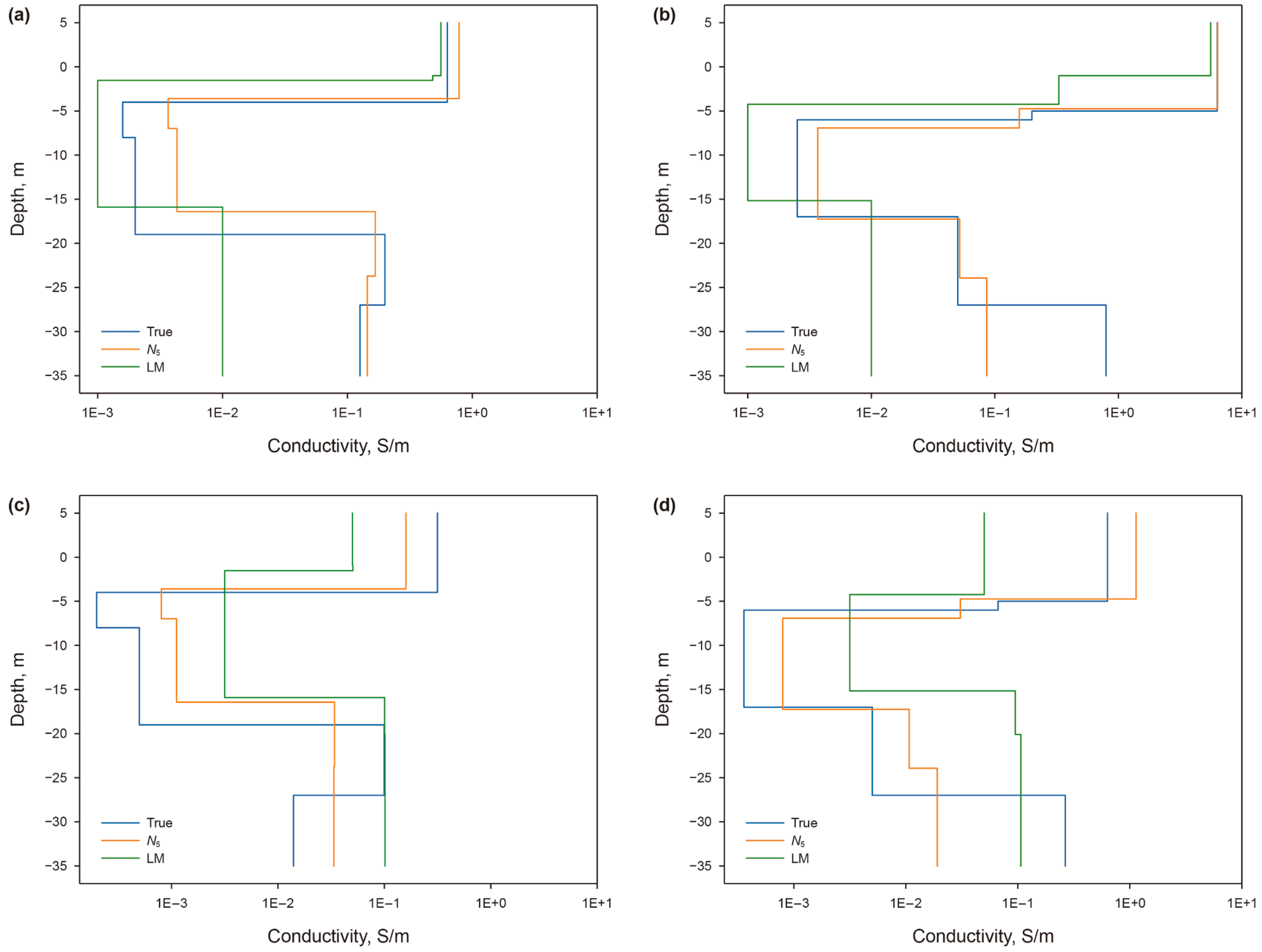


Fig. 13. Comparison of horizontal (**a**, **b**)) and vertical (**c**, **d**) conductivity inversion results between the DL method and LM method. Subfigures in the same column depict the same model.

leverages their combined sensitivities, resulting in the most accurate inversion outcomes.

To more comprehensively display the overall residual distribution of the 60,000 samples, we calculated the probability density function (PDF) of the residuals. The calculation formula is given as follows:

$$f(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{30}$$

$$K(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}} \tag{31}$$

$$h = 1.06\sigma n^{-\frac{1}{5}} \tag{32}$$

where $f(x)$ represents the probability density function, n is the number of sample points, and x_i is the ith data point. The bandwidth h is determined using Silverman's Rule of Thumb (Silverman, 2018), $K(\cdot)$ denotes the Gaussian kernel, and σ is the sample standard deviation. The PDF provides a more detailed representation of the overall distribution of the inversion residuals.

Fig. 7 shows that the inversion residuals approximate a standard normal distribution, indicating the accuracy and stability of the neural network inversion. By comparing the probability density at a residual of 0 for the different neural networks, we can see that the best inversion results for conductivity are obtained by neural network N_5, followed by N_{xx}, N_{yy}, and N_{zz}, with the poorest results from N_{zz} and N_{xz}. This phenomenon is consistent with the previous analysis.

Furthermore, a relative error analysis is provided. The formulas for calculating the relative errors of different parameters are given in Eqs. (12–14). The probability density distributions of the relative errors are shown in Fig. 8. From Fig. 8, we can clearly observe that for conductivity parameters, the inversion results obtained by N_5 are the most accurate, followed by N_{xx}, N_{yy}, and N_{zz}, with N_{xz} and N_{zx} performing the worst. For interface positions, N_5 again yields the best results, followed by N_{xz}, N_{zx}, and N_{zz}, while N_{xx} and N_{yy} show the highest relative errors. For anisotropy, the inversion results of N_5 remain the best, while N_{yy} shows a higher inversion accuracy than the other five neural networks.

To provide a more intuitive display of the neural network inversion results, some inversion results are presented by comparing the predicted parameters from the six neural networks with the true parameters. The results are shown in Fig. 9. It shows that all six neural networks are able to invert the variation trends of horizontal and vertical conductivity with depth. However, the inversion neural network N_5 is the most accurate one in revealing the layer interfaces and electrical parameters.

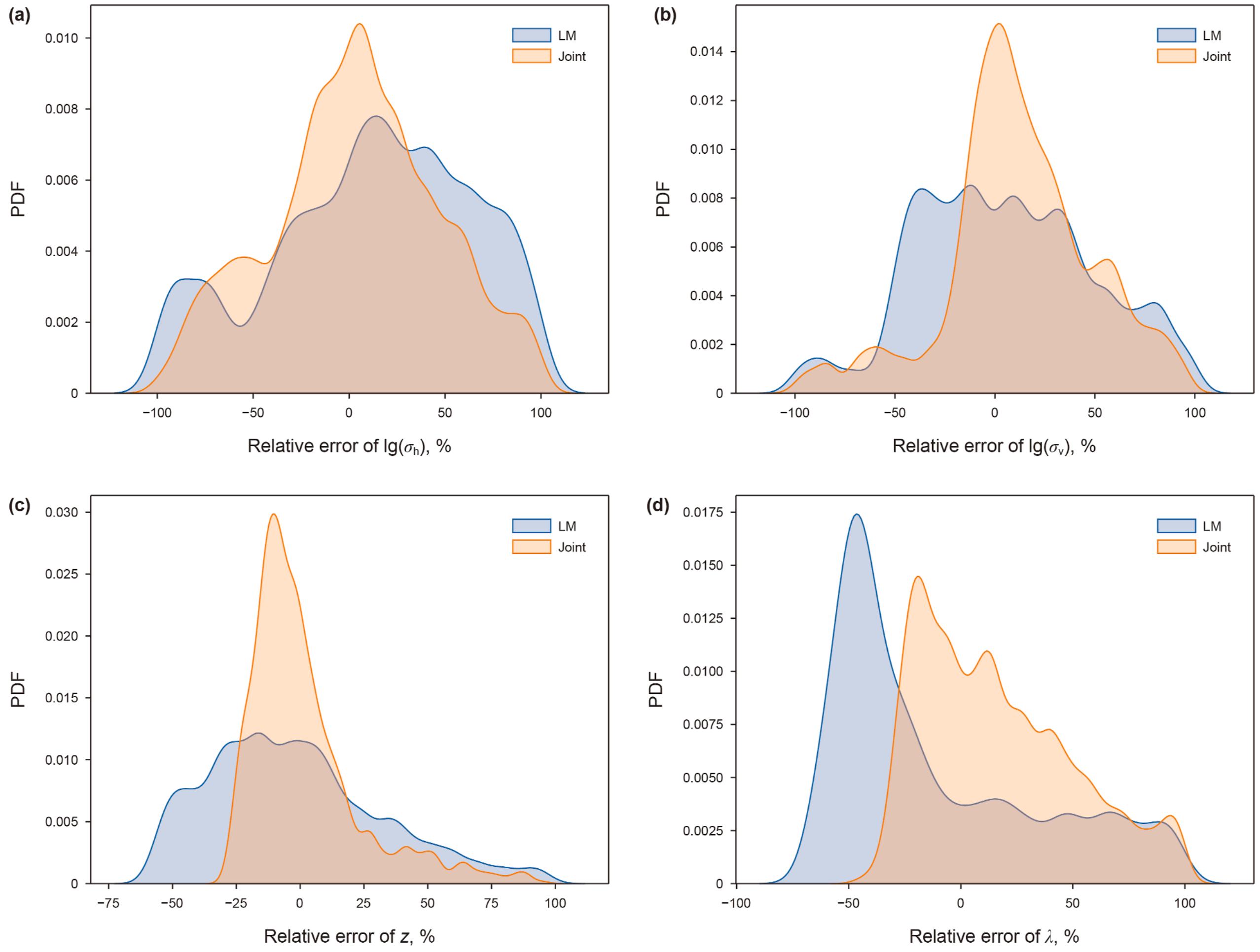


Fig. 14. Probability density distribution of relative error for LM method and joint inversion method. (**a**) horizontal conductivity, (**b**) vertical conductivity, (**c**) layer interface position, (**d**) anisotropy.

5. Discussion

5.1. Performance under varying noise levels

During the drilling process, EM look-ahead LWD tools are subject to significant noise interference. Therefore, it is essential to assess the noise sensitivity of the neural network. To this end, we added random Gaussian noise at four different levels—5%, 10%, 20%, and 50%—to the original test dataset, resulting in four distinct noisy test sets. The trained neural network N_5 was then applied to both the noise-free and noisy datasets for inversion, and the relative errors of the inversion results were computed, as shown in Fig. 10. The results indicate that at a noise level of 5%, the inversion accuracy remains largely unaffected. When the noise level increases to 10%, a moderate decline in accuracy is observed. At 20% noise, the inversion performance degrades more significantly. Notably, compared to conductivity parameters, the estimated interface depths exhibit greater sensitivity to noise, suggesting that structural boundaries are more vulnerable to degradation in measurement quality.

Representative inversion results under different noise levels are shown in Fig. 11. As observed, when the noise level is 5% or 10%, the neural network is still able to accurately reconstruct the variation of conductivity with depth. As the noise level increases beyond 20%, the network can still capture conductivity variations near the drill bit, but its ability to resolve features farther away from the bit significantly deteriorates.

5.2. Comparison with the LM method

To demonstrate the effectiveness of the multi-task inversion neural network, we perform a quantitative comparison between the inversion results of the neural network N_5 and those obtained from the LM method. The input data used is $\boldsymbol{x}_5$. We create a new dataset containing 1000 sample data points and perform inversions using both the LM method and the neural network N_5. The initial model for the LM method is a homogeneous medium with a horizontal conductivity of 0.1 S/m and a vertical conductivity of 0.01 S/m. To enable a more meaningful comparison, the deep learning inversion algorithm was run on a single CPU core with a single thread, while the LM algorithm was accelerated using all 24 CPU cores in multithreaded mode. The processor used was an AMD EPYC 7402 24-Core CPU, with a base clock of 2.8 GHz and a maximum Turbo Boost frequency of 3.35 GHz. The neural network inversion requires only 0.81 s, compared to 22824.31 s for the LM method, thereby fully satisfying the real-time demands of geosteering.

Subsequently, we performed a comparative analysis of the overall residuals of the two inversion results. First, the mean residuals of the inversion results are shown in Table 5. The residuals of the N_5 inversion results are smaller by 1–3 orders of magnitude

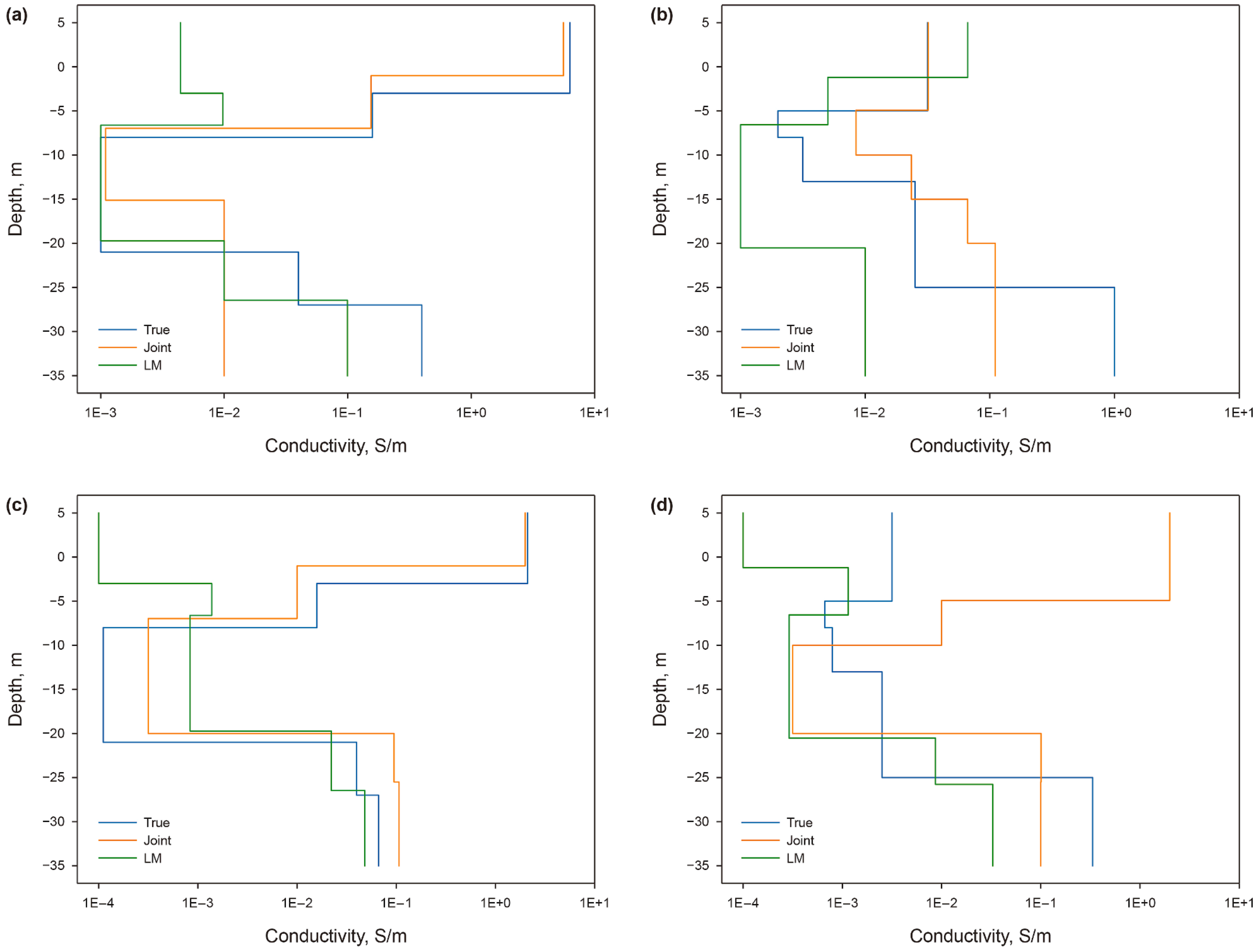


Fig. 15. Comparison of inversion results for horizontal (**a**, **b**) and vertical (**c**, **d**) conductivity using joint inversion and LM method. Subfigures in the same column depict the same model.

Table 6
Parameters of different networks.

Network	Input's dimensions	Output's dimensions	Number of layers	Number of parameters
LSTM	$(4,20)$	$(1,14)$	9	4,498,958
FCN	$(4,5,4)$	$(1,14)$	9	3,939,902
ResNet	$(4,5,4)$	$(1,14)$	26	5,198,343
MT-U-net	$(4,5,4)$	$\begin{bmatrix}(1,4)\\(1,5)\\(1,5)\end{bmatrix}$	34	4,104,229

Table 7
Architectures of different networks.

FCN		ResNet		LSTM	
Layer	Parameters	Layer	Parameters	Layer	Hidden size
conv	$3\times3,16$	conv	$3\times3,16$	LSTM	256
conv	$3\times3,32$	pool	max	LSTM	256
conv	$3\times3,64$	block1	$[3\times3,16]\times6$	LSTM	256
conv	$3\times3,128$	block2	$[3\times3,32]\times8$	LSTM	256
conv	$3\times3,256$	block3	$[3\times3,128]\times12$	LSTM	256
conv	$3\times3,512$	block4	$[3\times3,256]\times6$	LSTM	256
conv	$3\times3,512$	pool	average	LSTM	256
pool	average	linear	256×14	LSTM	256
conv	$1\times1,14$			LSTM	256

Table 8
Structure of the XGBoost model.

Input's dimensions	Output's dimensions	Number of estimators	Max depth	Subsample	Number of nodes
(1, 80)	(1, 14)	100	15	0.8	5,198,343

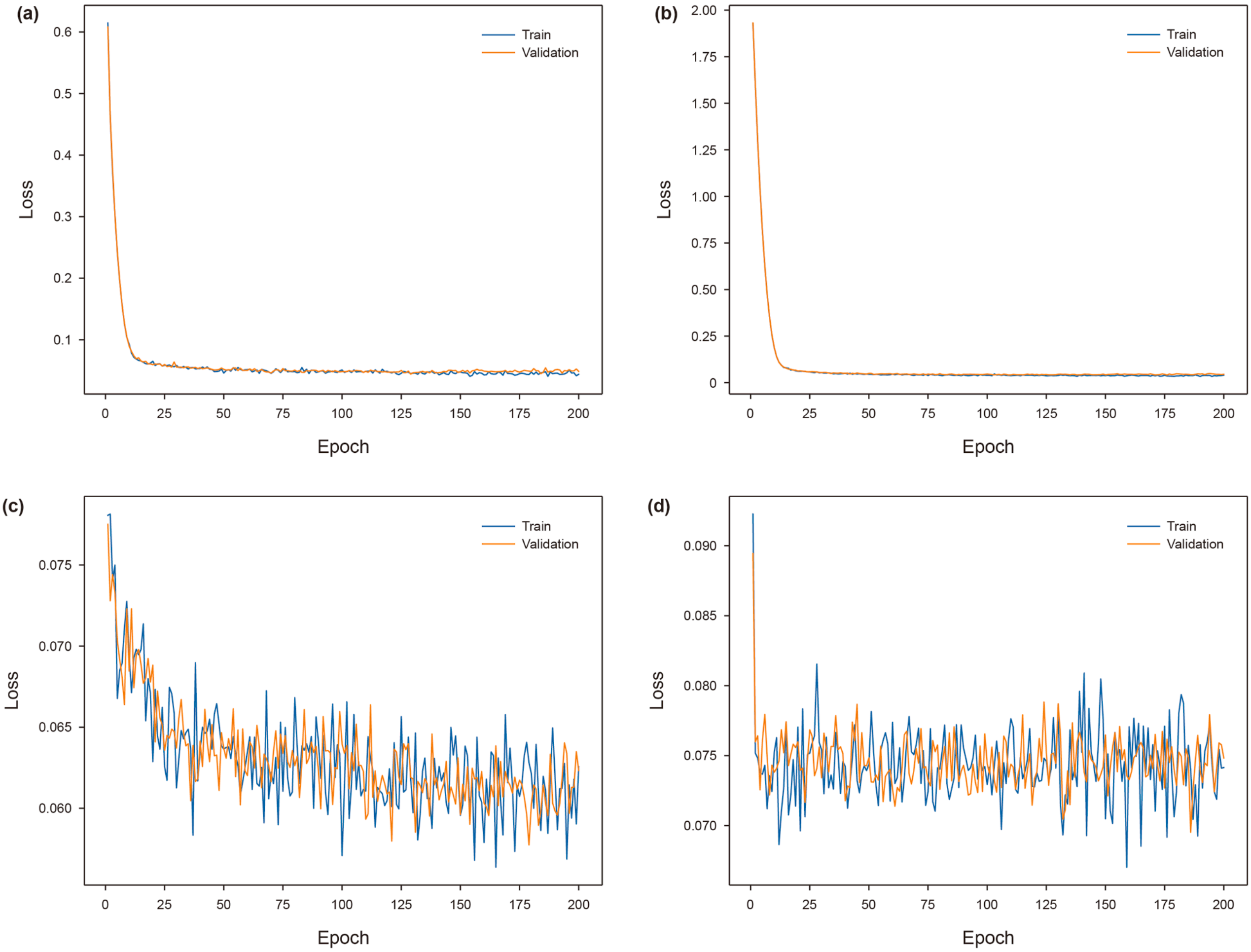


Fig. 16. Loss function decay curves for different neural networks: (**a**) multi-task U-net, (**b**) ResNet, (**c**) FCN, and (**d**) LSTM.

compared to the LM method. The larger residuals in the LM method are due to the initial model being far from the true model, making it susceptible to getting stuck in local minima.

Furthermore, we plotted the probability density distribution of the relative error for the four parameters, as shown in Fig. 12. It is evident that the relative error of N_5 is closer to a standard normal distribution. This indicates that the MT-U-net inversion method has higher adaptability and accuracy compared to the LM method.

We select two inversion results from the dataset (Fig. 13). The comparison shows that both the DL method and the LM algorithm are capable of capturing the overall conductivity variation trends. However, the deep learning approach demonstrates superior capability in delineating layer boundaries, particularly in accurately resolving thin layers.

To address the problem of initial value dependence in iterative methods, we employ the inversion results from the DL method as the initial model for the LM inversion. This approach effectively performs a joint inversion that combines the strengths of DL and LM methods. The overall distribution of the relative error for 1000 models is shown in Fig. 14. After using DL inversion results as the initial model, the probability density distribution of relative error becomes more concentrated and symmetric about zero. Specifically, the distribution under joint inversion is narrower and more peaked than that of the pure LM inversion across all parameters.

This improvement can be explained by noting that using the DL inversion result as the initial model provides a reasonably accurate starting point for the LM algorithm. With a more reliable initial estimate, LM is able to converge more directly toward the true solution and avoid becoming trapped in poor local minima. As a result, the inversion process becomes more stable and accurate, yielding results that are less sensitive to initial conditions and less prone to large deviations.

We randomly present inversion results for two geological models, as shown in Fig. 15. It is evident that the joint inversion method further improves the accuracy of the inversion results compared to the LM algorithm, particularly in enhancing the estimation of conductivity and interface positions near the drill bit.

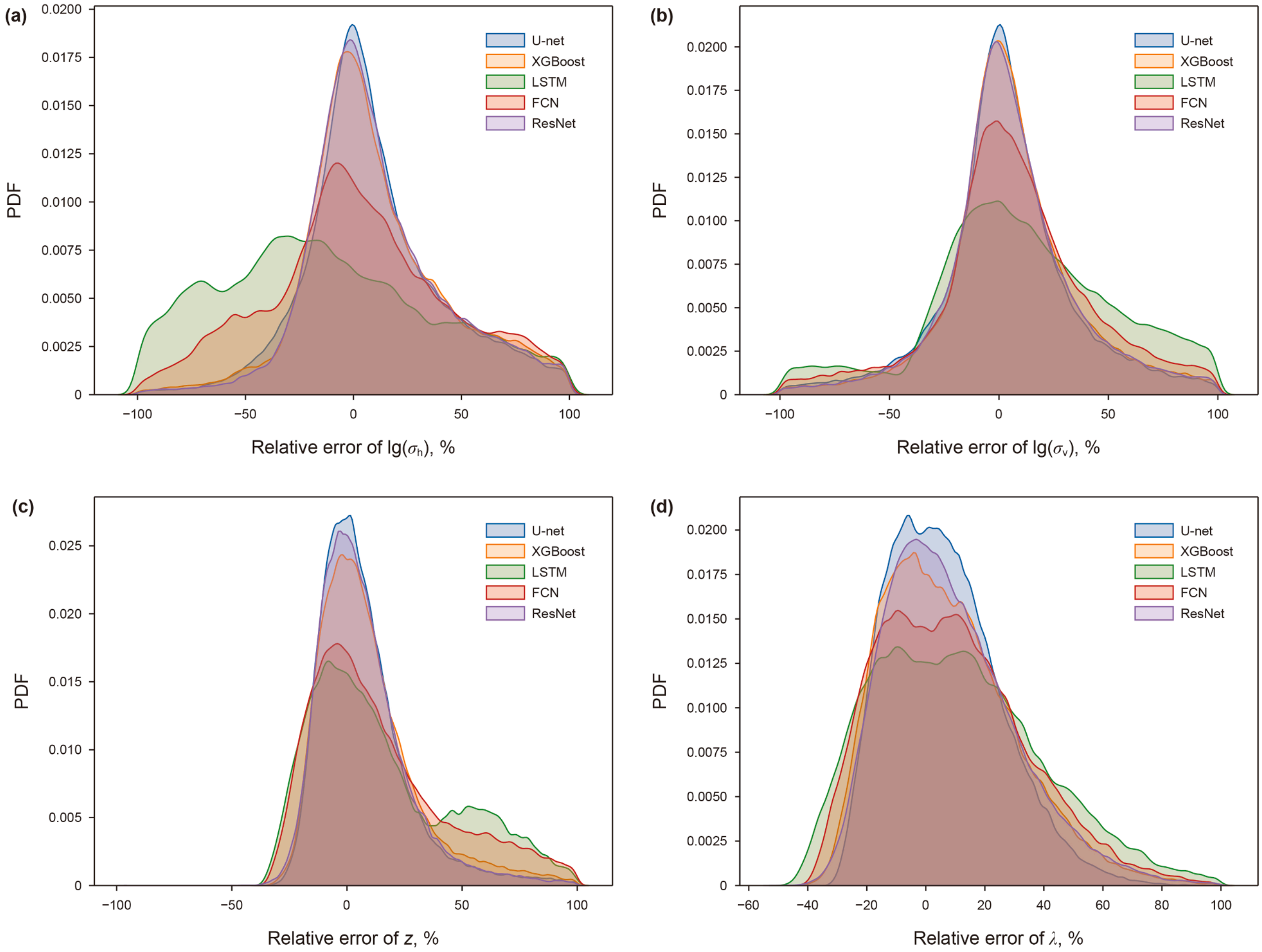


Fig. 17. PDFs of relative inversion errors for different machine learning methods applied to a common test set (60,000 samples). Panels (**a**), (**b**), (**c**), and (**d**) depict the error distributions for lg (σ_h), lg(σ_V), z, and λ, respectively.

5.3. Comparison with baseline models

To validate the effectiveness and superiority of the proposed multi-task U-net architecture in electromagnetic inversion tasks, we conducted a comprehensive comparison with several representative deep learning and machine learning models. These baseline models include LSTM (Hochreiter and Schmidhuber, 1997; Fan et al., 2022), fully convolutional networks (FCN) (Long et al., 2015), residual networks (ResNet) (He et al., 2015), and the eXtreme gradient boosting algorithm (XGBoost) (Chen and Guestrin, 2016). To ensure a fair comparison, all models were trained and evaluated on the same datasets as those used for training the proposed N_5 model. Furthermore, the number of trainable parameters for each neural network was kept at a comparable level to control for model complexity. All inversion methods used identical input data and target labels, with only minor differences in input dimensionality: convolution-based models (FCN, ResNet, and MT-U-net) received input tensors of shape $(4, 5, 4)$; the LSTM model accepted inputs of shape $(4, 20)$; and XGBoost, which only supports flat vector inputs, used reshaped inputs of size $(1, 80)$. All models produced outputs of shape $(1, 14)$, except for the multi-task U-net, which yielded three output branches corresponding to different geophysical properties. Detailed architectural configurations of LSTM, FCN, and ResNet are provided in Table 7, and the number of trainable parameters is summarized in Tables 6 and 8. In Table 7, "conv" denotes 2D convolutional layers, "pool" refers to pooling layers, "block" indicates residual blocks, and "linear" corresponds to fully connected layers.

During the training of neural networks, the hyperparameter settings for all models were kept consistent with each other. Fig. 16 illustrates the trends of training and validation loss for four neural network architectures (U-net, ResNet, FCN, and LSTM) under identical training datasets and hyperparameter configurations. It is evident that these networks exhibit markedly different loss behaviors throughout the training and validation processes. Both the multi-task U-net and ResNet demonstrate strong convergence properties, stabilizing at relatively early iterations while maintaining low training and validation errors. Their overall convergence speeds and final error levels are comparable, indicating robust fitting capacity and generalization performance. In contrast, the FCN and LSTM show significantly slower convergence rates with more pronounced fluctuations in loss curves. Notably, the LSTM displays a sluggish decrease in error and poor stability throughout training, suggesting limited modeling and representational capabilities for this type of electromagnetic inversion task.

To comprehensively assess the inversion capabilities of various machine learning algorithms, five distinct methods were employed on an identical test dataset comprising 60,000 samples. Fig. 17 illustrates the PDF distributions of relative errors associated with the inversion of different parameters. The results reveal that the multi-task U-net model consistently achieves the most

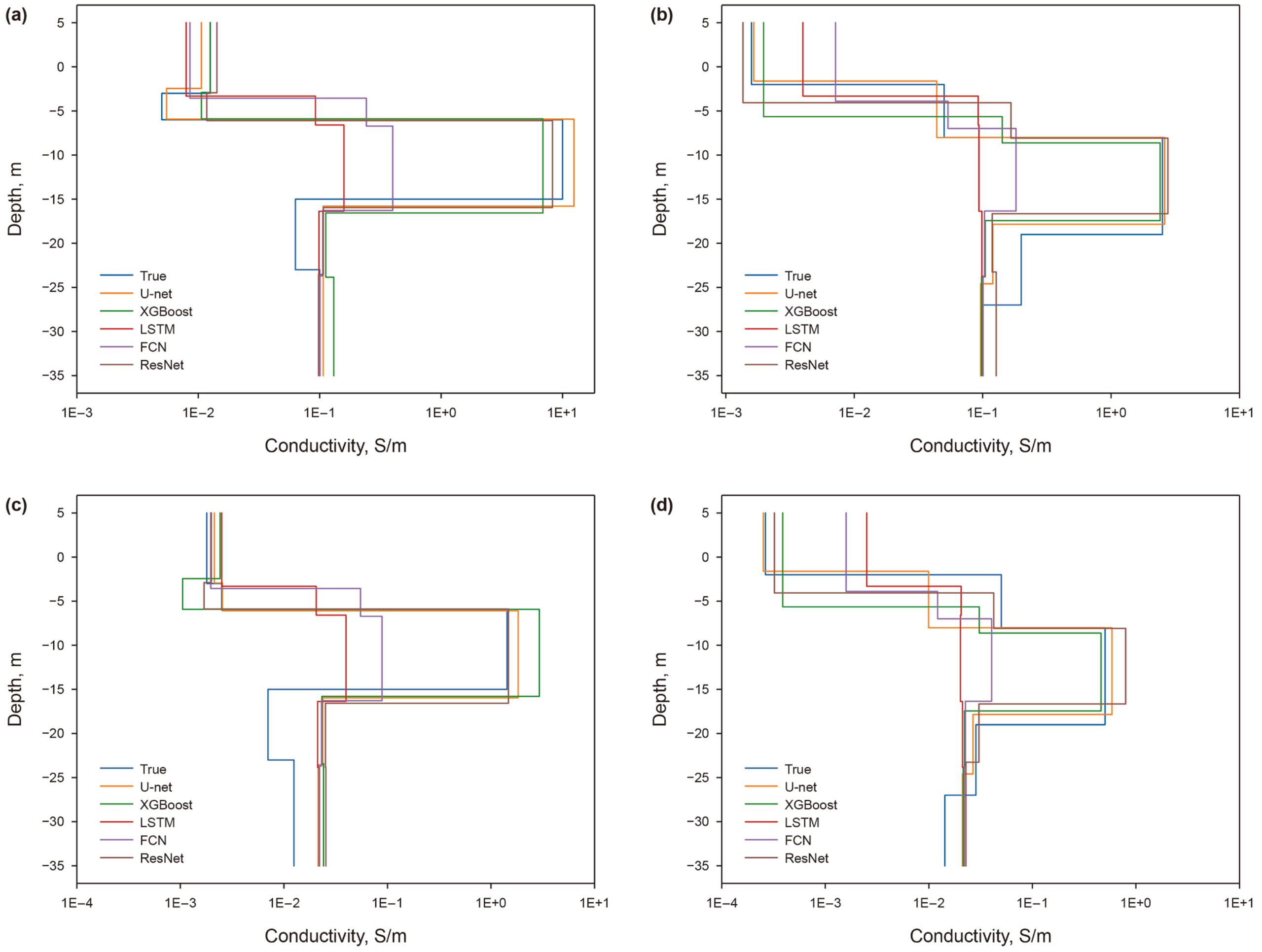


Fig. 18. Presents representative inversion results from five different algorithms. Across these multiple inversion cases, the multi-task U-net consistently delivers the most stable and accurate reconstructions, effectively delineating multilayer formation interfaces and capturing variations in conductivity values with high spatial resolution. In contrast, ResNet and XGBoost also provide satisfactory inversion outcomes, reasonably approximating the true profiles, especially in terms of interface localization and conductivity trends. However, the FCN and LSTM models demonstrate comparatively inferior performance; although they occasionally identify primary formation boundaries, their conductivity estimations suffer from substantial deviations, resulting in overall inversion outputs that deviate notably from the ground truth.

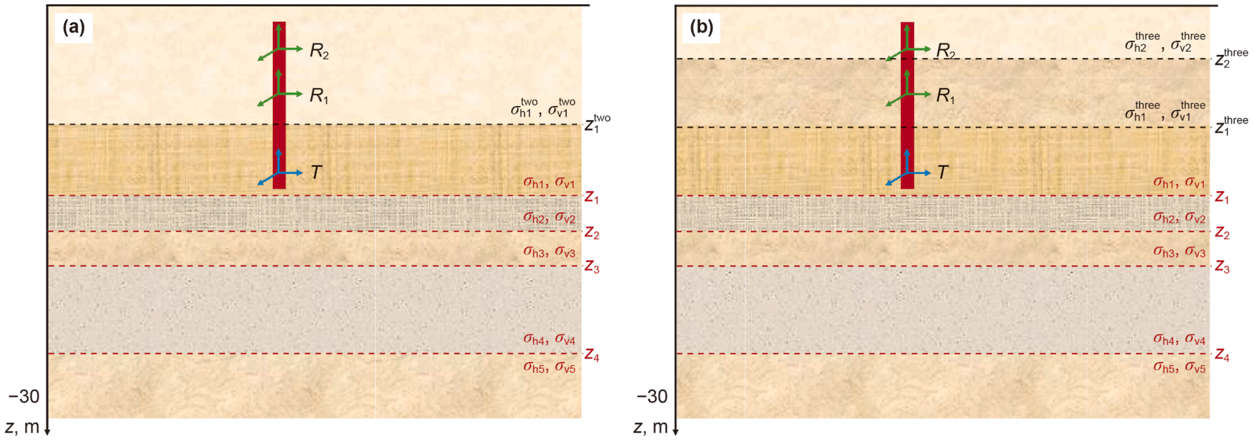


Fig. 19. Schematic of the extended geological models. (**a**) The tool is located within a two-layer formation; (**b**) The tool is located within a three-layer formation.

concentrated error distributions across all parameters, characterized by pronounced peak values and narrow variance. This performance underscores its superior predictive accuracy and robustness. ResNet and XGBoost also yield relatively concentrated error profiles, although with moderately wider distributions and reduced peak densities compared to U-net, indicating comparatively lower but still competitive modeling capacity in feature representation and nonlinear mapping. Conversely, the FCN and

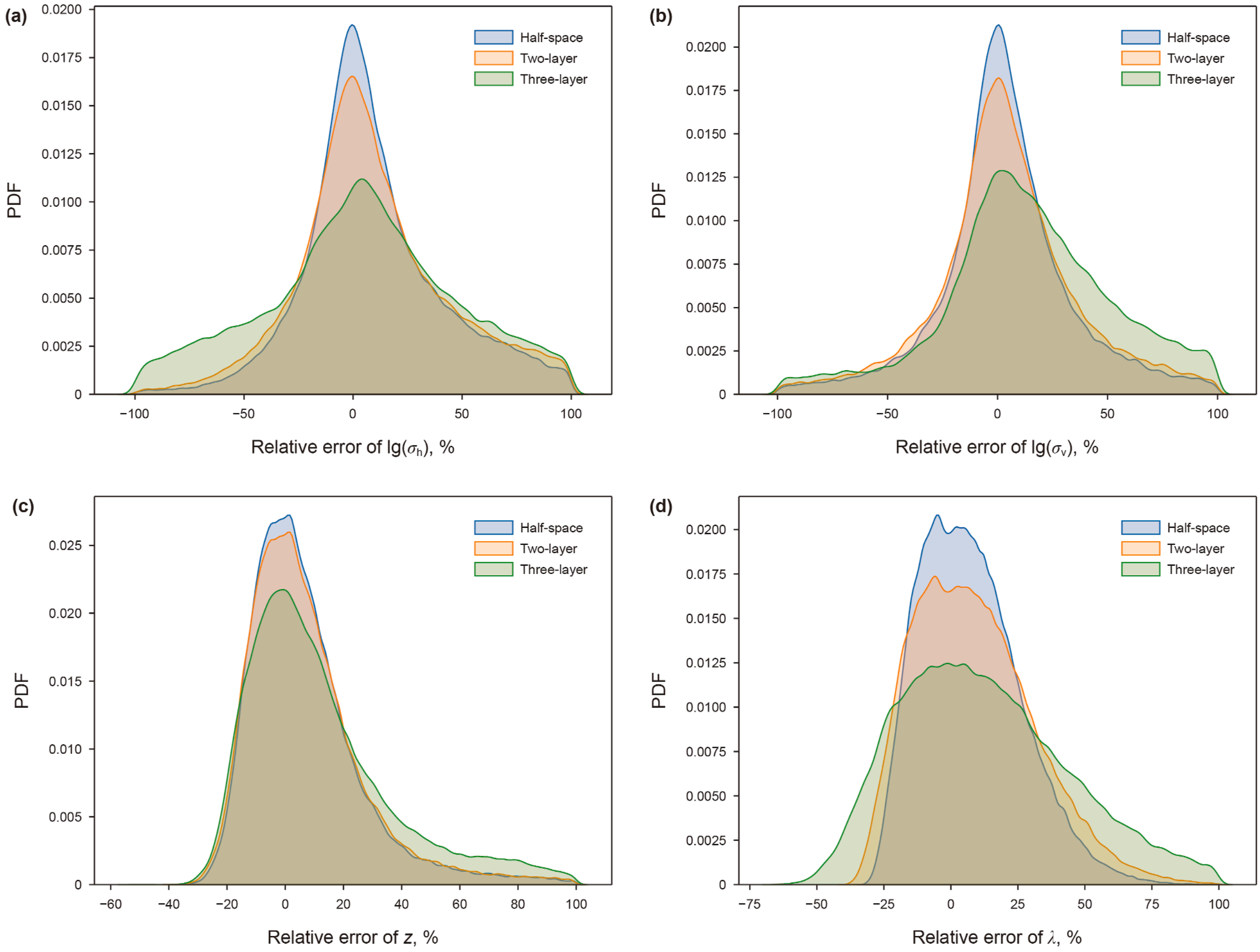


Fig. 20. Relative error distributions of inversion results on three test sets. The terms half-space, two-layer, and three-layer indicate that the tool is located in a formation modeled as a homogeneous half-space, a two-layer structure, and a three-layer structure, respectively.

LSTM models exhibit markedly inferior inversion performance, evidenced by broader error distributions and significant long-tail behavior. Notably, LSTM displays a systematic bias in error across all parameters, highlighting its limitations in capturing the complex nonlinearities and spatial dependencies present in subsurface electromagnetic data. Collectively, these findings demonstrate that the multi-task U-net's deep hierarchical feature fusion and encoder-decoder architecture confer a substantial advantage for multi-parameter inversion tasks. ResNet and XGBoost can be considered viable baseline models, whereas FCN and LSTM show suboptimal performance under the current experimental framework (see Fig. 18).

Using the trained neural network N_5, we conducted inversion on the T_{semi}, T_{two}, and T_{three} test sets (see Fig. 19) and examined the relative error distributions of the predicted parameters ($\lg(\sigma_{\mathrm{h}})$, $\lg(\sigma_{\mathrm{v}})$, z, and λ), as depicted in Fig. 20. The results show that transitioning from a single-layer half-space (T_{semi}) to a two-layer structure (T_{two}) results in only a slight broadening of the error distributions, with inversion accuracy remaining largely stable. This suggests robust network performance under moderate geological complexity. Conversely, in the three-layer model (T_{three}), the error distributions exhibit significant widening with heavier tails, indicating a substantial decline in inversion accuracy and generalization capability in highly complex geological settings.

These findings confirm that while the U-net-based model exhibits robustness in moderately complex geological settings, its performance declines in highly heterogeneous environments. To improve the adaptability of the inversion framework to realistic subsurface conditions, future efforts will focus on incorporating multi-layer conductivity structures as prior knowledge into the neural network and developing more sophisticated geological models to enhance the generalization capability of the deep learning framework. These advancements aim to improve the accuracy, stability, and applicability of look-ahead inversion in complex geological scenarios.

6. Conclusions

A multi-task U-net inversion method for inverting five-layer anisotropic formation using the LWD electromagnetic look-ahead data has been developed. By constructing datasets of six different components ($\boldsymbol{x}_{xx}$, $\boldsymbol{x}_{xz}$, $\boldsymbol{x}_{yy}$, $\boldsymbol{x}_{zx}$, $\boldsymbol{x}_{zz}$, and $\boldsymbol{x}_5$), six neural network models were trained and tested, and their inversion performance was comprehensively evaluated by using residual and relative error analysis. Among them, the multi-component neural network N_5 demonstrates optimal inversion accuracy and stability, proving the effectiveness of combining multiple components for inversion. Further analysis indicates that the coaxial and co-planar components are most sensitive to conductivity, while cross-components and coaxial components are most sensitive to the layer interface location. The *yy* component is most sensitive to anisotropic parameters. Also, the inversion results show that the multi-task U-net neural network can complete inversions within millisecond without relying on initial models, providing a significant advantage in time consumption.

Robustness analyses reveals that the proposed model has reliable inversion performance under moderate noise conditions and two-layer geological backgrounds, although its accuracy noticeably declines when facing high noise levels (>20%) or complex three-layer structures.

Compared with the traditional Levenberg-Marquardt (LM) iterative inversion method, the multi-task U-net exhibits robust advantages in both inversion accuracy and computational efficiency. Furthermore, a hybrid approach that uses the U-net predictions as initialization for the LM algorithm improves performances in recovering formation conductivities, suggesting strong complementarity between data-driven and physics-based methods. Another comparison with mainstream machine learning models—including LSTM, FCN, ResNet, and XGBoost—further confirms that the proposed network consistently achieves high accuracy and faster convergence.

CRediT authorship contribution statement

Shun Zhang: Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wen-Xiu Zhang:** Writing – review & editing, Supervision, Funding acquisition. **Wen-Xuan Chen:** Writing – review & editing. **Peng-Fei Liang:** Writing – review & editing, Supervision. **Wen-Yang Wang:** Writing – review & editing. **Xing-Han Li:** Writing – review & editing.

Declaration of competing interest

All authors declare that there are no competing interests.

Acknowledgements

This work is supported by the National Key Research and Development Program of China under Grant 2019YFA0708301 and the Youth Innovation Promotion Association Program of CAS under Grant 2020069.

References

Bittar, M., Wu, H.H., Ma, J., Pan, L., Fan, Y., Griffing, M., Lozinsky, C., 2021. First LWD co-located antenna sensors for real-time anisotropy and dip angle determination, yielding better look-ahead detection. Petrophysics 62 (3), 296–310. https://doi.org/10.30632/PJV62N3-2021a4.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

Clark, B., Liiling, M.G., Jundt, J., Ross, M., David, B., 1988. A dual depth resistivity measurement for fewd. In: SPWLA 29th Annual Logging Symposium, San Antonio, Texas.

Constable, M.V., Antonsen, F., Stalheim, S.O., Olsen, P.A., Fjell, Ø.Z., Dray, N., Eikenes, S., Aarflot, H., Haldorsen, K., Digranes, G., Seydoux, J., Omeragic, D., Thiel, M., Davydychev, A., Denichou, J., Salim, D., Frey, M., Homan, D., Sarwa, T., 2016. Looking ahead of the bit while drilling: from vision to reality. In: SPWLA Annual Logging Symposium.

Coope, D.F., Yearsley, E.N., 1986. Formation Evaluation Using EWR Logs. International Meeting on Petroleum Engineering, Beijing, China. https://doi.org/10.2118/14062-MS.

Dong, C.J., Zhao, H., 2020. Attitude estimation based on gradient descent algorithm and adaptive parameters. Chin. J. Sensors Actuat. 33 (7), 997–1002 (in Chinese).

Durdyyev, G., 2021. New Technologies and Protocols Concerning Horizontal Well Drilling and Completion. Doctoral Dissertation. Politecnico di Torino.

Fan, J., Zhang, W., Chen, W., Li, X., 2022. Inversion based on deep learning of logging-while-drilling directional resistivity measurements. J. Petrol. Sci. Eng. 208, 109677. https://doi.org/10.1016/j.petrol.2021.109677.

Fredericks, P.D., Hearn, F.P., Wisler, M.M., 1989. Formation evaluation while drilling with a dual propagation resistivity tool. In: SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/SPE-19622-MS.

Gavin, H.P., 2019. The Levenberg-Marquardt Algorithm for Nonlinear Least Squares curve-fitting Problems. Department of Civil and Environmental Engineering Duke University. https://api.semanticscholar.org/CorpusID:113404737.

Gianzero, S., Merchant, G.A., Haugland, M., Strickland, R.W., 1994. New developments in 2 MHz electromagnetic wave resistivity measurements. In: SPWLA Annual Logging Symposium.

Halliburton, 2023. Halliburton Introduces BrightStar Look-Ahead Resistivity Service. https://www.halliburton.com/en/about-us/press-release/halliburton-introduces-brightstar-look-ahead-resistivity-service. (Accessed 10 January 2025).

Hawkins, D., Phettongkam, N., Nakchamnan, N., 2015. Optimizing well placement in geosteering using an azimuthal resistivity tool in complex thin bed reservoirs in the Gulf of Thailand. In: SPWLA Annual Logging Symposium. SPWLA. SPWLA-2015.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Hong, D.C., 2009. Study on the Data Processing Method of Triaxial Induction Logging. Doctoral Dissertation. Jilin University (in Chinese).

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neur. Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Jackson, C., Hagiwara, T., 1998. A new simultaneous anisotropy and dielectric correction algorithm for LWD resistivity measurements. In: SPWLA Annual Logging Symposium. SPWLA.

Jin, Y., Wu, X., Chen, J., Huang, Y., 2019. Using a physics-driven deep neural network to solve inverse problems for LWD azimuthal resistivity measurements. In: SPWLA Annual Logging Symposium. SPWLA. https://doi.org/10.30632/T60ALS-2019_IIII.

Kang, Z.M., Qin, H.J., Zhang, Y., Li, X., Ni, W.N., Li, F.B., 2023a. Data inversion of azimuthal electromagnetic wave logging while drilling based on LSTM neural network. Petrol. Drill. Techn. 51 (2), 116–124. https://doi.org/10.11911/syztjs.2023047.

Kang, Z.M., Zhang, Y., Qin, H., Gan, W., Chen, G., 2023b. An intelligent inversion method for azimuth electromagnetic logging while drilling measurements. IEEE Access 11, 79285–79294. https://doi.org/10.1109/ACCESS.2023.3298972.

Kingma, D.P., 2014. Adam: A method for stochastic optimization arXiv preprint arXiv:1412.6980. https://doi.org/10.48550/arXiv.1412.6980.

Li, X.H., Zhang, W.X., Chen, P., Chen, W.X., Wang, J., Zheng, J., You, J.Q., Zhang, Y.L., Li, H., 2023. Accuracy analysis and real-time inversion and imaging logging while drilling azimuth electromagnetic wave resistivity measurements. Chin. J. Geophys. 66 (9), 3990–3998. https://doi.org/10.6038/cjg2023Q0412 (in Chinese).

Liang, P.F., Di, Q.Y., Chen, W.X., Liu, R.M., Li, X.H., 2023. An EM LWD tool for deep reading looking-ahead. IEEE Access 11, 142601–142610. https://doi.org/10.1109/ACCESS.2023.3339777.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440. https://doi.org/10.48550/arXiv.1411.4038.

Ma, J., Riofrio, K., Clegg, N., Sinha, S., Wu, H.H., Kotwicki, A., Skorve, T., Johansen, Y.B., 2022. Successful geostopping using a recently developed LWD look-ahead ultra-deep EM resistivity tool. In: Abu Dhabi International Petroleum Exhibition and Conference. SPE. https://doi.org/10.2118/211613-MS.

Mohd Aris, M.N., Daud, H., Mohd Noh, K.A., Dass, S.C., 2021. Stochastic process-based inversion of electromagnetic data for hydrocarbon resistivity estimation in seabed logging. Mathematics 9 (9), 935. https://doi.org/10.3390/math9090935.

Ndokwu, C., Brock, P., Wharton, P., Okowi, V., Foekema, N., 2018. Extra-Deep Azimuthal Resistivity in Deep Offshore Niger Delta. Second EAGE/SPE Geosteering and Well Placement Workshop. European Association of Geoscientists & Engineers, pp. 1–5. https://doi.org/10.3997/2214-4609.201803119.

Noh, K., Pardo, D., Torres-Verdín, C., 2021. 2.5-D deep learning inversion of LWD and deep-sensing EM measurements across formations with dipping faults. IEEE Geosci. Rem. Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2021.3128965.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z.M., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J.J., Chintala, S., 2019. Pytorch: An imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. 32. https://doi.org/10.48550/arXiv.1912.01703.

Prensky, S., 2006. Recent advances in LWD/MWD and formation evaluation. World Oil 227 (3), 69.

Rammay, M.H., Alyaev, S., Elsheikh, A.H., 2022. Probabilistic model-error assessment of deep learning proxies: An application to real-time inversion of borehole electromagnetic measurements. Geophys. J. Int. 230 (3), 1800–1817. https://doi.org/10.1093/gji/ggac147.

Seydoux, J., Denichou, J.M., Amir, I., Wibowo, V., Bauch, T., Salim, D., Viandante, M., Han, Y.S., Wang, C., Cuadros, G., Hamada, M., Tan, S., Feng, Y., Leveque, S., 2019. Real-time EM look-ahead: A maturing technology to decrease drilling risk in low inclination wells. In: SPWLA Annual Logging Symposium. SPWLA. https://doi.org/10.30632/T60ALS-2019_GGGG.

Silverman, B.W., 2018. Density Estimation for Statistics and Data Analysis. Routledge.

Singh, D., Singh, B., 2020. Investigating the impact of data normalization on classification performance. Appl. Soft Comput. 97, 105524. https://doi.org/10.1201/9781315140919.

Sun, Q.F., Ni, H.S., Yue, X.Z., Zhang, P.Y., Gong, F.M., 2024. Inversion of azimuthal electromagnetic wave resistivity lwd based on deep residual network. Petrol. Drill. Techniq. 52 (5), 97–104. https://doi.org/10.11911/syztjs.2024089 (in Chinese).

Wang, L., Fan, Y., 2019. Fast inversion of logging-while-drilling azimuthal resistivity measurements for geosteering and formation evaluation. J. Petrol. Sci. Eng. 176, 342–351. https://doi.org/10.1016/j.petrol.2019.01.067.

Wang, L., Qiao, P., Li, Z.Q., Zhang, P., Deng, S.G., Fan, Y.R., 2023. A new semi-analytical algorithm for rapid simulation of triaxial electromagnetic logging responses in multilayered biaxial anisotropic formations. Geophysics 88 (2), D115–D129. https://doi.org/10.1190/geo2021-0714.1.

Wu, Z., Li, H., Han, Y., Zhang, R., Zhao, J., Lai, Q., 2022. Effects of formation structure on directional electromagnetic logging while drilling measurements. J. Petrol. Sci. Eng. 211, 110118. https://doi.org/10.1016/j.petrol.2022.110118.

Yang, K.L., Wang, L., Fang, H., Ai, W.P., Wang, N., Zeng, Z.C., 2025. A new logging-while-drilling azimuthal electromagnetic measurement for highly resistive coal mines. J. Geophys. Eng. 22 (4), 1017–1025. https://doi.org/10.1093/jge/gxaf056.

Yu, S., Ma, J., 2021. Deep learning for geophysics: current and future trends. Rev. Geophys. 59 (3), e2021RG000742. https://doi.org/10.1029/2021RG000742.

Zhang, S., He, Y.F., Lun, Z.M., 2022. Hybrid Newton bisection algorithm for solving Lee-Kesler equation. Sci. Technol. Eng. 22 (18), 7859–7865 (in Chinese).

Zhang, Z.Y., Gonguet, C., Rajani, V., Roeterdink, R., 2008. Directional LWD resistivity tools and their business impacts. In: SPWLA Annual Logging Symposium. SPWLA. https://doi.org/10.1190/1.3063898.

Zhao, N., Li, N., Xiao, Z., Wang, X., Qin, C., 2024a. Physics-driven deep learning pixel-based inversion of logging-while-drilling in anisotropic formation. IEEE Trans. Geosci. Rem. Sens. 62, 1–12. https://doi.org/10.1109/TGRS.2024.3355463.

Zhao, N., Shen, S.N., Li, N., Hu, H.T., Qi, C., Qin, C., 2024b. Physics-driven deep learning inversion for azimuthal lwd elctromagnetic wave measurement. Oil Geophys. Prospect. 59 (5), 1069–1079. https://doi.org/10.13810/j.cnki.issn.1000-7210.2024.05.014 (in Chinese).

Zhu, R.X., Jin, Z.J., Di, Q.Y., Yang, C.C., Chen, W.X., Tian, F., Zhang, W.X., 2023. Research and progress of intelligent drilling technology system and related theories. Chin. J. Geophys. 66 (1), 1–15. https://doi.org/10.6038/cjg2022Q0730 (in Chinese).