ANNALS OF GEOPHYSICS, 69, 4, DM453, 2026; doi:10.4401/ag-9500
OPEN ACCESS

Enhanced generalization of a deep learning framework for 1-D magnetotelluric inversion

Mehdi Rahmani Jevinani[1], Banafsheh Habibian Dehkordi[*,1], Ian J. Ferguson[2]

(1) Institute of Geophysics, University of Tehran, Tehran, Iran
(2) Dept. of Geological Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

Article history: November 24, 2025; accepted May 27, 2026

Abstract

Widespread and successful application of deep learning-based methods to solve inverse problems in geophysics- particularly in magnetotellurics (MT) – remains limited largely due to poor generalization of trained models to unseen and noisy field data. In order to contribute to addressing this issue, we present a robust one-dimensional (1-D) MT inversion framework in which the deep learning model has increased capacity to generalize. First, by employing a randomized conductivity model construction scheme, constrained by geophysical resolution limitations and geologically realistic scenarios, a comprehensive and diverse training dataset representing a wide variety of subsurface conductivity models is generated. To further ensure a representative sampling of the model space, Gaussian noise is added to part of the generated data. Second, a clustering-based data splitting strategy is introduced as an alternative to the random approach, in which each cluster proportionally contributes to the training, validation, and test subsets, to preserve consistent statistical distribution across them. Qualitative and quantitative evaluations confirm that the combined approach performs robustly and reliably when exposed to a broad range of synthetic data as well as a sample of real field MT data. Resistivity models constructed using a modified convolutional U-Net architecture, trained with a combination of clean and noisy data generated in the stated manner, are compared to those derived from deterministic methods and show high levels of accuracy.

Magnetotellurics; Deep Learning; Inversion; Generalization; Resistivity

1. Introduction

The inherently ill-posed and non-linear magnetotelluric (MT) inverse problem aims at inferring the subsurface electrical resistivity distribution from surface electromagnetic impedance measurements, usually expressed as apparent resistivity and phase. Conventional deterministic inversion methods typically rely on iterative linearized schemes to minimize a combination of data misfit and model complexity- often based on first- or second-order smoothness constraints to achieve a predefined value. While one-dimensional (1-D) MT inversion provides limited spatial resolution, it remains a crucial initial step in characterizing subsurface structures and serves as a foundation for multidimensional analyses.

Deep learning (DL) methods have emerged, in recent years, as promising alternatives to deterministic geophysical inversion approaches (Liu et al., 2020; Liu et al., 2021b; Liu et al., 2022; Liu et al., 2024; Lopez-Alvis et al., 2022; Oh et al., 2020; Pintea et al., 2022; Puzyrev, 2019; Puzyrev and Swidinsky, 2021; Shahriari et al., 2020; Yadav et al., 2021). They can yield near real-time, interpretable solutions, offering high computational efficiency once trained. The principal challenge, however, lies in generalization capability. DL models trained on noise-free synthetic data tailored to specific geological settings and frequency ranges, tend to overfit and often fail to perform reliably when exposed to diverse field data (Zhang and Pavlick, 2025).

To address this limitation, this study introduces a generalizable and robust DL-based framework for 1-D MT inversion. There are numerous examples of one-dimensional magnetotelluric inversion based on deep learning (Liao et al., 2022; Liu et al., 2021a; Liu et al., 2022; Wang et al., 2019; Wang et al., 2022), where some type of techniques to reduce over-fitting and strategies for training data generation are mainly used to improve generalization. Therefore contributions of this study need to be clearly highlighted. A probabilistic approach-described in detail in the next section- is employed to assign resistivity values over a wide range to 35 subsurface delineated layers in which the layer thicknesses increases in accordance with the decrease in MT resolution with depth. This model enables, through forward calculations, the generation of training datasets that cover a large part of frequency and apparent resistivity range of MT responses and captures both smooth and sharp boundaries, improving the diversity and representativeness of the training data. Consequently, the trained model can perform accurate, rapid, and flexible 1-D MT inversions without requiring prior information such as the number of layers or interpreter intervention. Additionally, the robustness of the trained model to noise is investigated by introducing a mixture of noisy and noise-free samples into the training process. Our results show that incorporating noisy data into the training process substantially enhances network performance and generalizability. As another contribution, random data splitting into training, validation, and test subsets is replaced with a clustering-based strategy for partitioning the data. By improving convergence stability and smoothing validation loss, this approach further enhances generalization capability.

Following our previous study (Rahmani Jevinani et al., 2024), the layer thicknesses are considered variable and part of the deep learning model output; we also incorporate MT impedance phase values into the input features, justified by the fact that because of possible static shift effects on the amplitude of the MT response, the phase is the most robust measure of the contrast. This is different from previous studies on deep learning-based 1-D MT inversion, where fixed layer boundaries are assumed and most of them rely only on apparent resistivity as input. At a glimpse, Section 2 introduces methodology. Section 3 is devoted to model performance evaluation and discussion. Conclusions are given in Section 4.

2. Methodology

In magnetotellurics, natural electromagnetic field fluctuations are simultaneously recorded on the Earth's surface and used to explore the subsurface geoelectric structure. The impedance tensor (Z), as the main MT transfer function, expresses the relationship between the horizontal components of the electric (E) and magnetic (H) fields:

$$\mathrm{Z} = \frac{\mathrm{E}}{\mathrm{H}} \tag{1}$$

In a one-dimensional environment, impedance can be treated as a scalar quantity, i.e., estimates relating orthogonal components of the horizontal an magnetic field do not depend on the coordinate system. The response is often expressed in terms of apparent resistivity ρ_a and impedance phase φ:

$$\rho_a = \frac{1}{\mu_0 \omega} |\mathrm{Z}|^2 \tag{2}$$

$$\varphi = \arg(Z) \tag{3}$$

2.1 Data Preparation

2.1.1 Synthetic Model Design

Designing resistivity models that generate comprehensive, flexible and realistic responses requires considering both the actual geological structures and the resolution limitations inherent to the MT method. The model considered here starts with a thickness of 50 m for the first layer and the thickness of each subsequent layer increases by 20%. A total of 35 layers is included, extending down to a depth of 150 km. The 35 layers are categorized into seven groups, each consisting of five layers. Within each group, individual layer thicknesses are randomly selected between the groups' specified lower and upper bounds; while the total thickness of each group remains unchanged, and the total thickness of the 35 layers is 150 km (Fig. 1). Although the thickness of individual layers is variable, the thickness of the composite group layers retains an increase with depth match the theoretical MT resolution.

When assigning resistivity values randomly to the layers, layers with smaller resistivity values are considered more likely to occur, reflecting MT data higher sensitivity to conductive structures, where currents concentrate. Moreover, practical scenarios can involve sharp as well as smooth boundaries. Consequently, we proceed with the

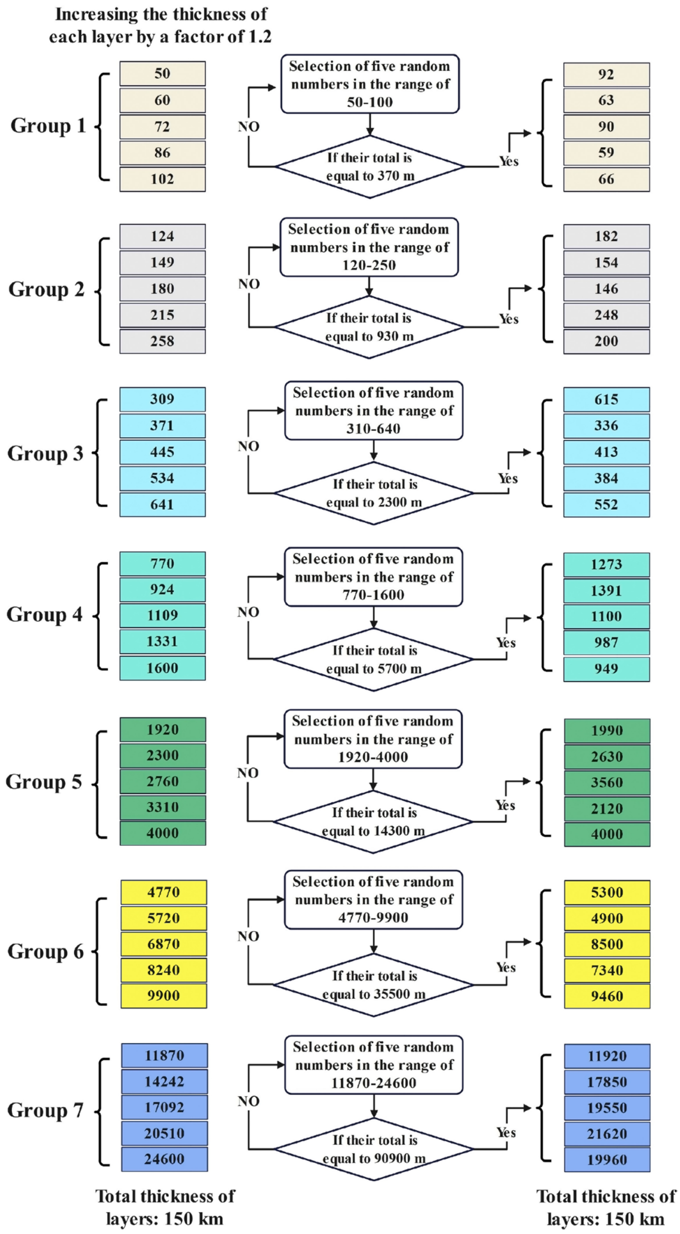

Figure 1. An illustration of determining the thickness of layers of a 35-layer earth to a depth of 150 km.

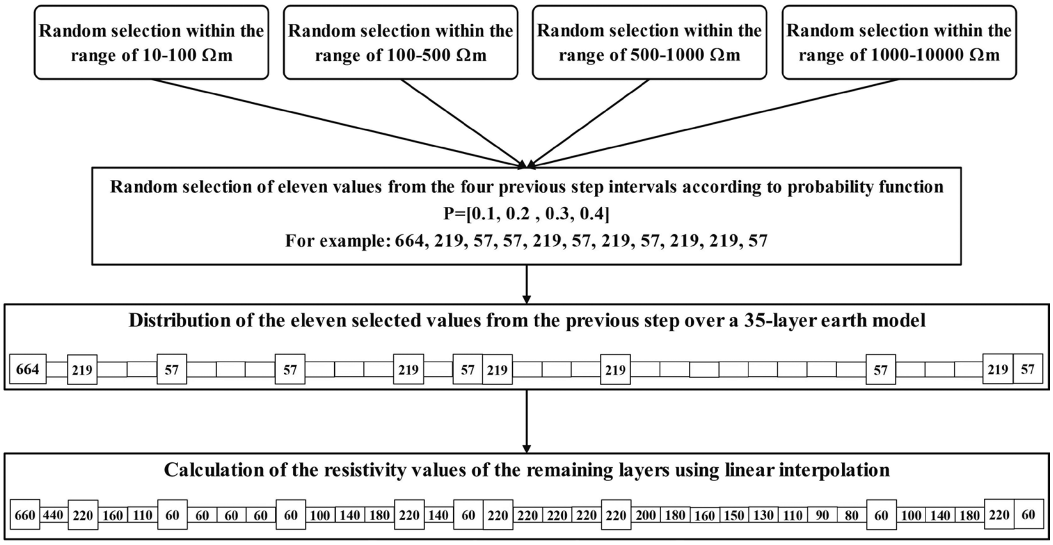

Figure 2. An illustration of probability-based assignment of electrical resistivity to layers.

following steps: Four resistivity ranges are defined ([10, 100], [100, 500], [500, 1000] and [1000, 10000]), and one value is randomly selected from each range, assuming a uniform linear distribution of values over the range. Subsequently, eleven resistivity values are sampled from these four initial values so that the probabilities of choosing a value from the first to the fourth range are set to 0.4, 0.3, 0.2, and 0.1, respectively. The key idea is to use a probabilistic approach to assign resistivity values to different layers, with the first range being the most likely and the fourth range being the least likely to occur. The resulting eleven values are randomly placed in 35 layers, provided that the first and last layers must have values. The remaining layers are filled by interpolation with a linear function. This procedure is repeated from the first step for the generation of each synthetic model. This scheme (Fig. 2) naturally allows for both smooth gradients and sharp contrasts.

The sample dataset is generated through forward modeling based on the recursive relationships between the impedances of successive layers as derived from Maxwell's equations. After generating five million geoelectric models, apparent resistivity and phase values are calculated at 42 logarithmically spaced frequencies in the range of $[8.3923 \times 10^{-4} - 4.2969 \times 10^{2}]$ Hz. Both gradient- and distance-based algorithms are influenced by the scale of the data. By properly pre-processing the data, the algorithm's performance is prevented from being biased towards a specific feature.

2.1.2 Data Augmentation through Noise Addition

The addition of noise to a portion of the training data serves as a form of regularization (Bishop, 1995), which can accelerate learning and enhance generalization. By adding noise into the data as a tunable hyper-parameter (Goodfellow et al., 2016), new samples are generated in the vicinity of the original ones, effectively smoothing the input space and facilitating the learning of the underlying non-linear mapping (An, 1996).

Although noise addition is recognized as a common data augmentation technique, a standardized methodology for incorporating it within the deep learning frameworks for MT data is yet to be established. Given that real MT measurements inherently contain noise, it is essential to ensure the robustness of the trained models against noise. Two perspectives exist in this context: training with noise-free data and utilizing noisy datasets in the test phase to evaluate model robustness and generalizing it to noisy observations (Liu et al., 2022; Rahmani et al., 2024); and training directly with noisy data to enhance stability and generalization (Liao et al., 2022).

In this study, two datasets are generated: one noise-free and another with two percent zero-mean random Gaussian noise added to the calculated impedance data at all frequencies and applied randomly to ten percent of the samples. This approach enables an assessment of how noise influences network performance and generalization capability. Addition of noise to the training data is done here for the purpose of inversion regularization (Bishop, 1995) rather than representation of practical MT data. For reference, the 2% level used corresponds approximately to the noise level of high quality MT data, and in many inversion studies, response noise is implicitly considered to have a Gaussian distribution (e.g., Rodi and Mackie, 2012). In practice, the distribution of response noise may be more complex, and depends on the statistics of the recorded time-series, the impedance estimation method, and the noise estimation method (e.g., Chave, 2012).

2.1.3 Clustering-based Data Splitting

Following data generation, an essential step in the deep learning workflow is to partition the data into train, validation, and test subsets. A purely random splitting as the common method, however, does not guarantee that the three subsets share statistically similar distributions, and using non-homogeneous subsets can adversely affect the generalization of the final models. Here, both noise-free and noisy datasets are initially divided into train, validation, and test subsets in a 98:1:1 ratio randomly. To evaluate the statistical consistency among the subsets, the non-parametric Kolmogorov-Smirnov (KS) test (Smirnov, 1948) is performed. As shown in Fig. 3, the KS statistics for

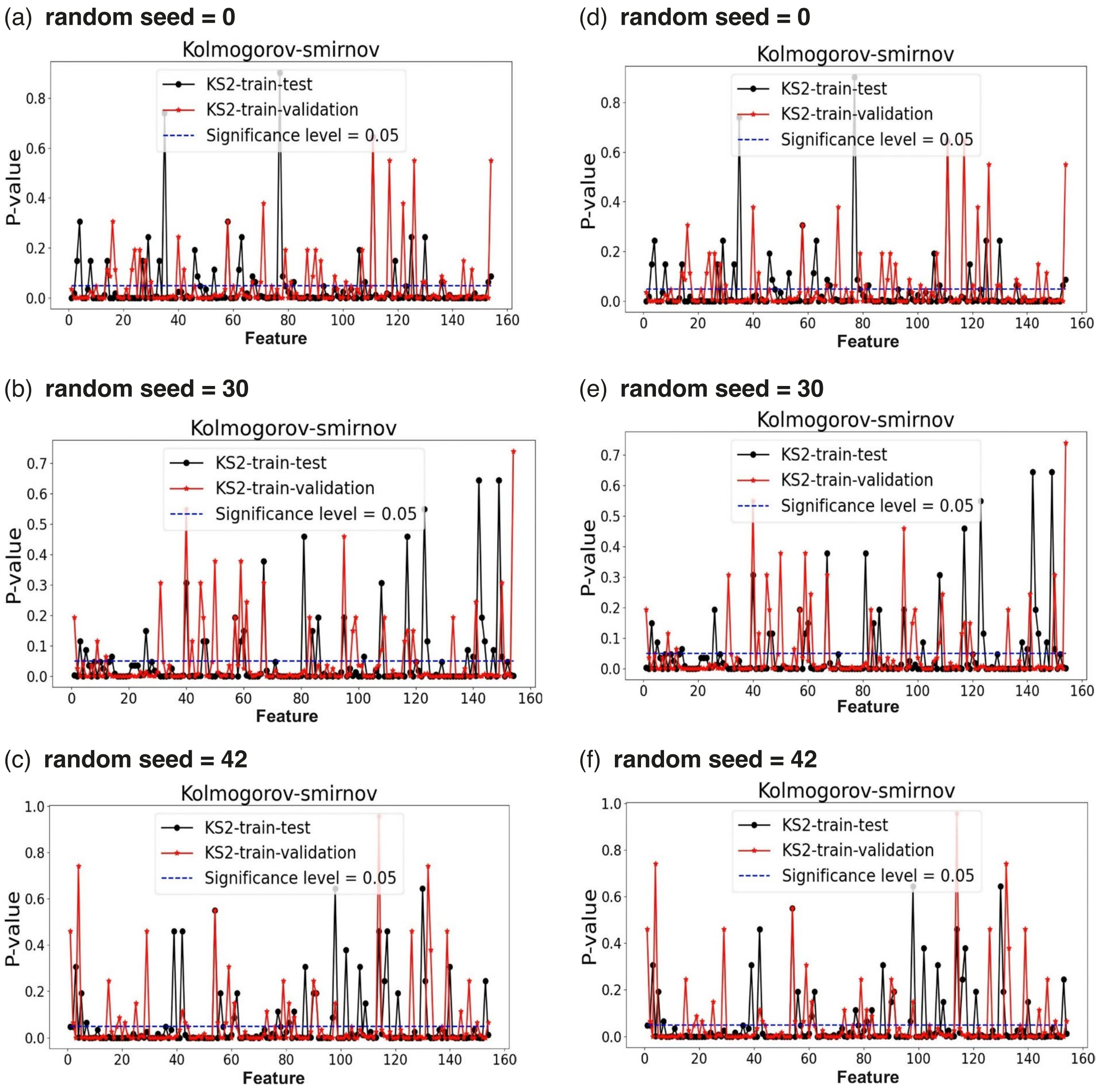

Figure 3. KS test results for assessing the similarity of the statistical distribution among the training, validation, and test datasets with random splitting and a splitting ratio of 98/1/1. Right and left columns: noise-free and noisy data, respectively.

various random seeds are below the 0.05 significance level, indicating that the three subsets do not originate from identical distributions in either the noisy or noise-free cases. For each data sample, the feature vector is defined to include both data-space attributes- comprising apparent resistivity and phase at each frequency- together with the 1-D model parameters, namely layer resistivities and thicknesses. This comprehensive representation ensures the KS-based consistency evaluation among the three subsets is performed in terms of both data space and models space, preventing potential bias that could arise if only one domain were considered.

To overcome this limitation, a clustering-based data splitting strategy is implemented. The K-means algorithm (MacQueen, 1967), one of the most widely used unsupervised clustering methods, is applied to the generated MT responses to cluster the data prior to splitting. The optimal number of clusters (K) is determined using the elbow method, which plots the sum of squares of intra-cluster distances as a function of K (Kodinariya and Makwana, 2013). Based on Fig. 4, the optimal cluster numbers are considered to be 32 and 33 for the noisy and noise-free datasets, respectively.

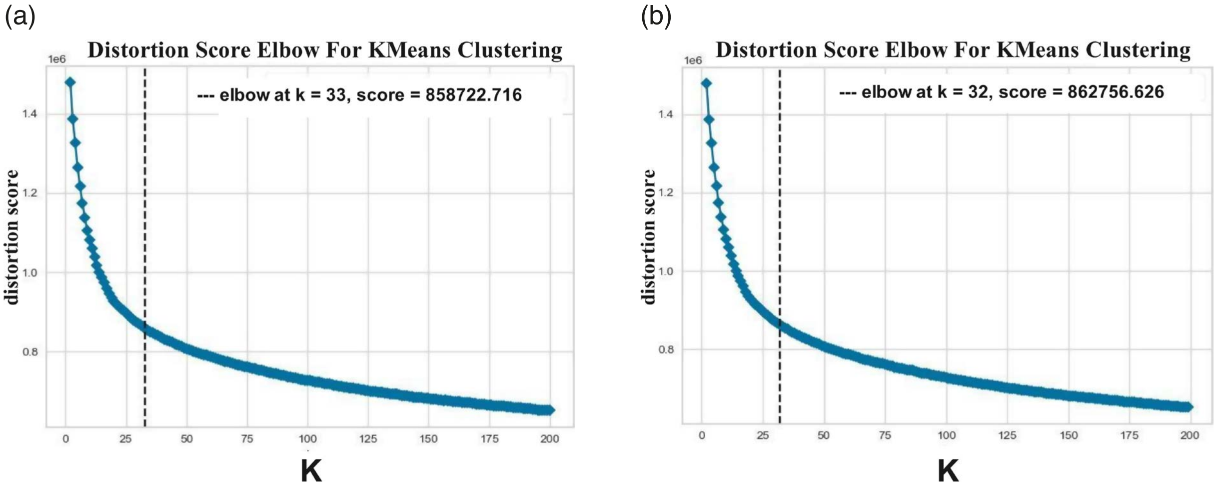

Figure 4. Elbow curve for a) noise-free data with optimal k equal to 33; b) noisy data with optimal k of 32.

After clustering, data from each cluster are proportionally assigned to the train, validation, and test subsets to ensure consistent statistical representation across all three sets. The KS test is again applied, and the resulting p-values exceed 0.05 for almost all of the features, confirming the similarity of distributions among the subsets (Fig. 5). In this framework, the input data consist of apparent resistivity and phase values, and the target parameters are layer resistivities and thicknesses. Each data sample can be therefore represented as a two-dimensional matrix, and the entire dataset can be expressed as a three-dimensional tensor.

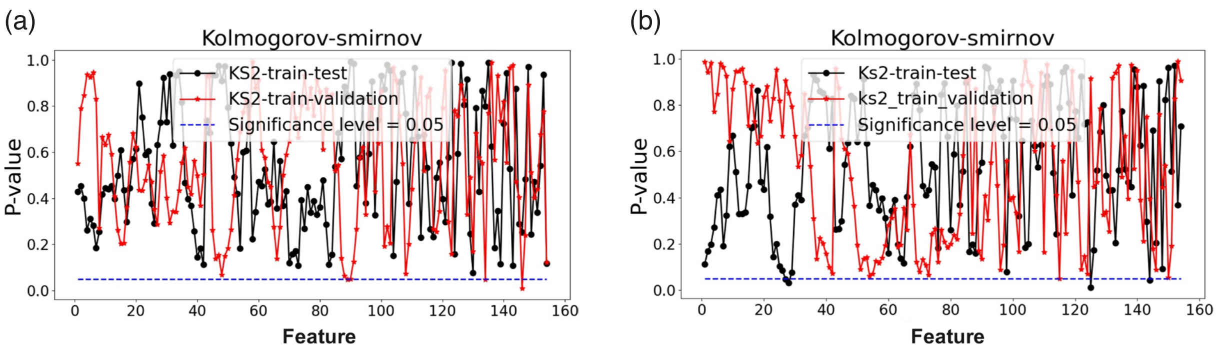

Figure 5. KS test results for assessing the similarity of the statistical distribution among the training, validation, and test datasets with clustering-based splitting and a splitting ratio of 98/1/1; (a) and (b): noise-free and noisy data, respectively.

2.2 Network Specification and Training Strategy

A specified deep learning model based on the U-Net (Ronneberger et al., 2015) architecture is employed to perform one-dimensional MT inversion. The network consists of 57 layers, each comprising of a set of filters, as illustrated in Fig. 6. The number of filters starts at 32 and doubles after each downsampling layer, enabling the model to progressively capture more complex and higher level representations. In the expansive path, the number of filters decreases symmetrically following the same scaling pattern. Each convolutional block incorporates batch normalization, the ReLU activation function, and max-pooling operation, which collectively enhance training stability, accelerate convergence, and improve feature extraction efficiency. A fully connected layer is included at the final stage of the network, which is a common design for regression tasks. The choice of the U-Net is motivated by a comparative investigation in which U-Net, Residual Network (ResNet) (He et al., 2016), and Variational Autoencoder (AE) (Kingma and Welling, 2014) were evaluated in terms of accuracy and generalizability for magnetotelluric inversion problem (Rahmani Jevinani et al., 2024).

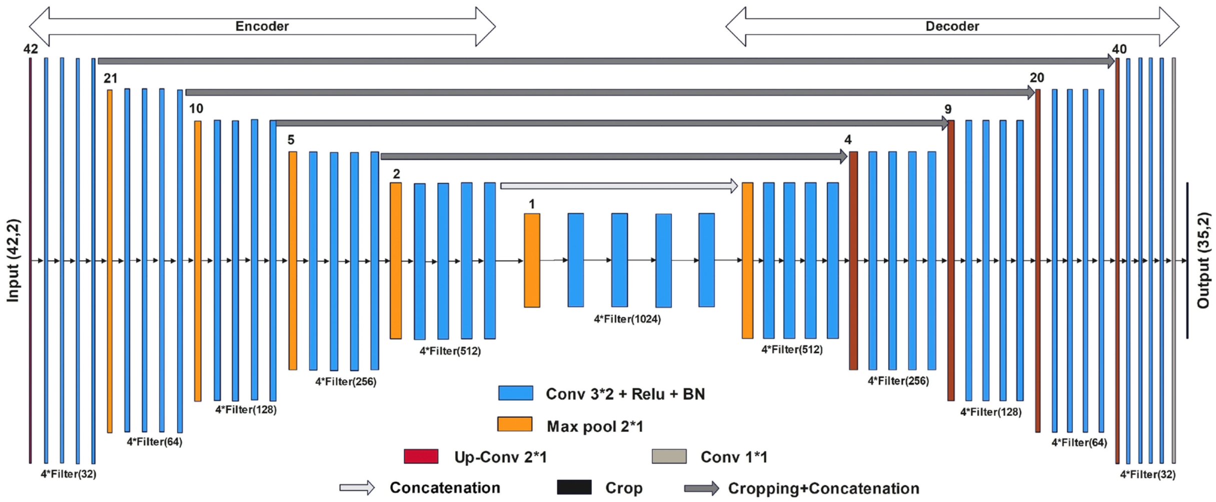

Figure 6. Modified U-Net architecture used for MT inversion.

The proposed U-Net framework effectively captures the main features of the input data through iterative weight optimization. The initial weights are randomly initialized and subsequently updated during the training process. The effectiveness of the training procedure highly depends on the hyperparameters, which are carefully selected prior to training and, when necessary, refined based on the behavior of the validation loss curve (Kumar et al., 2024). The Adam optimizer is utilized, which combines momentum-based updates with adaptive learning rates for Individual parameters, providing both stability and rapid convergence. In practice, values close to the default exponential decay rates for the first- and second-moment estimates (β_1 and β_2)are generally found to be appropriate and require little adjustment. The parameters are set to $\beta_1 = 0.85$ and $\beta_2 = 0.949$. Learning rate, batch size, and number of epochs are tuned around commonly accepted values through controlled trial and error procedures, guided primarily by the behavior of the validation loss. Indeed, monitoring validation loss serves as a central mechanism for hyperparameter tuning, helping to balance convergence speed, model stability, and generalization performance. Accordingly, the network is trained for 40 epochs with a batch size of 32 and a learning rate of 1×10^{-6}. The Huber loss function is chosen to achieve a balance between robustness to outliers and smooth convergence. By combining quadratic behavior for small errors with linear behavior for large deviations, it provides stable convergence while reducing the influence of extreme residuals and potential non-Gaussian noise in the in the real MT data during training.

The loss curves, representing the variation of accuracy or error during the parameter update, are diagnostic tools for evaluating the training and generalization behavior of the model. In a properly trained model, both the training and validation loss curves show a consistent decreasing trend with a minimal separation, indicating that the network has leaned the underlying data distribution without over-fitting. However, if the statistical properties of the training and validation datasets differ significantly, this balance is disturbed, leading to irregular validation loss patterns. The monitoring of training and validation loss curves is particularly essential for regression problems, as it provides an insight into the learning dynamics and the generalization capability of the model (Goodfellow et al., 2016). Since the validation dataset typically contains fewer samples than the training set, differences in their statistical distributions can cause the validation loss to fluctuate while the training loss continues to decrease monotically. As illustrated in Fig. 7, the loss curves confirm that over-fitting does not occur. After 40 epochs, the final training loss reaches 0.00061 for the case of noisy data and 0.00293 for the noise-free case. This result clearly indicates that the model trained with also noisy data achieved a more stable convergence and superior generalization capability, highlighting the robustness of the proposed deep learning framework for MT inversion.

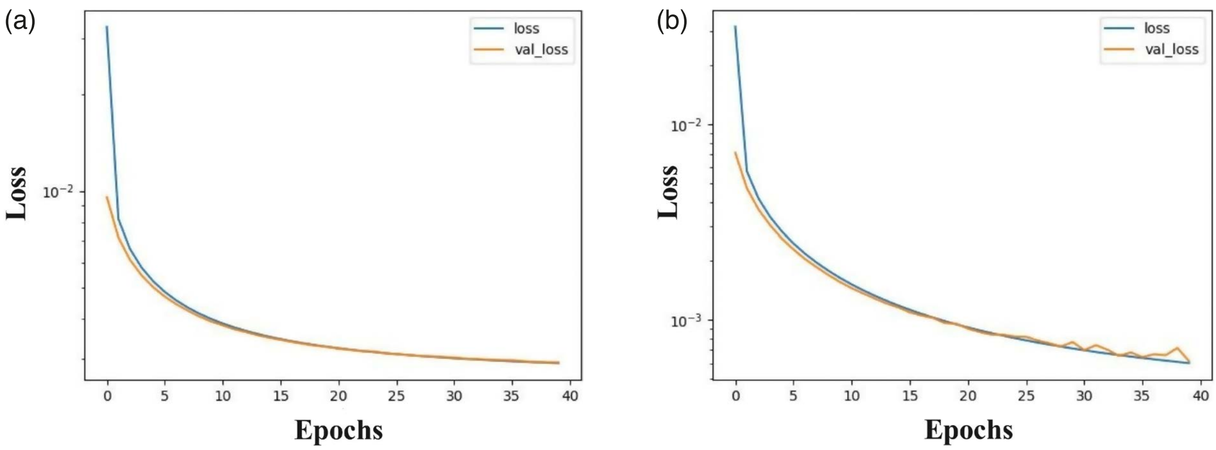

Figure 7. Training and validation loss curves for a) noise-free and b) noisy data.

3. Discussion on the Results

3.1 Qualitative Evaluation

To qualitatively evaluate the inversion results, noise-free and noisy data are applied separately to the two models: one trained with noise-free data and the other with noisy data. For the first six samples of the noise-free test set, the predictions obtained from the network trained with noisy data exhibit higher accuracy and better consistency, characterized by smaller errors and improved convergence (Fig. 8). This observation confirms the superior convergence behavior inferred previously from Fig. 7.

The generality of the proposed inversion approach is demonstrated by the diversity of 1-D resistivity structures that have been created by the model generation procedure and fitted well by noise-trained inversions and in most cases by the noise-free trained inversion. Even in just the first six models shown in Fig. 6 the method has recovered relatively conductive and relatively resistive surface layers, embedded thin conductive and thin resistive layers, and embedded thick conductive and thick resistive layers. These structures include model parameters that in deterministic inversions can be resolved by the MT data alone, including the resistivity of surface layers, the depth to the top of conductive layers (Jones, 1982, 1993; Kalscheuer and Pedersen, 2007) and parameters of relatively thick conductive and resistive layers, as well as structures that require the introduction of a priori model information to be defined, including the depth to the base of thin conductive layers (Kalscheuer and Pedersen, 2007) and the presence and parameters of thin resistive layers.

For the next evaluation of the inversion procedure it is applied to test data contaminated with 2% Gaussian noise (Fig. 9). In this case, the predictions from the network trained with noisy data are significantly more accurate that those obtained from the network trained with noise-free data. As emphasized earlier, adding noise during training acts similarly to regularization mechanism, thereby enhancing the generalization capability of the network. In contrast, predictions made by the network trained with noise-free data under noisy test conditions are highly unreliable and deviate considerably from the true model.

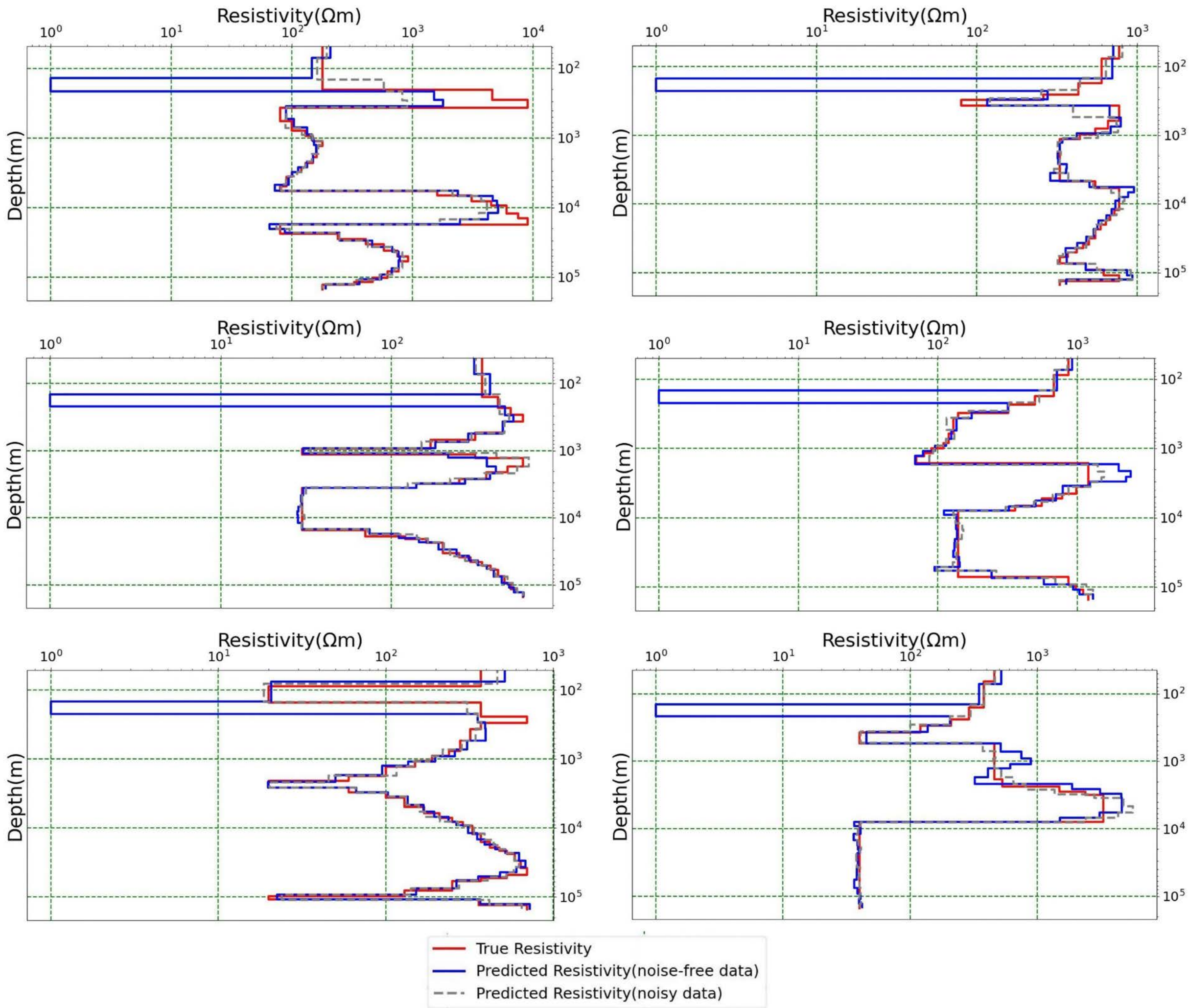

Figure 8. Comparison of predictions made by networks trained on noise-free and noisy data for the first six samples of the noise-free test data.

To further verify the robustness and flexibility of the network, MT responses of models with fewer than 35 layers are calculated, contaminated with 2% Gaussian noise, and subsequently provided as input into the trained network. The prediction results for various layer numbers (Fig. 10) indicate a good agreement with the corresponding true models, confirming the network's adaptability to structural variability.

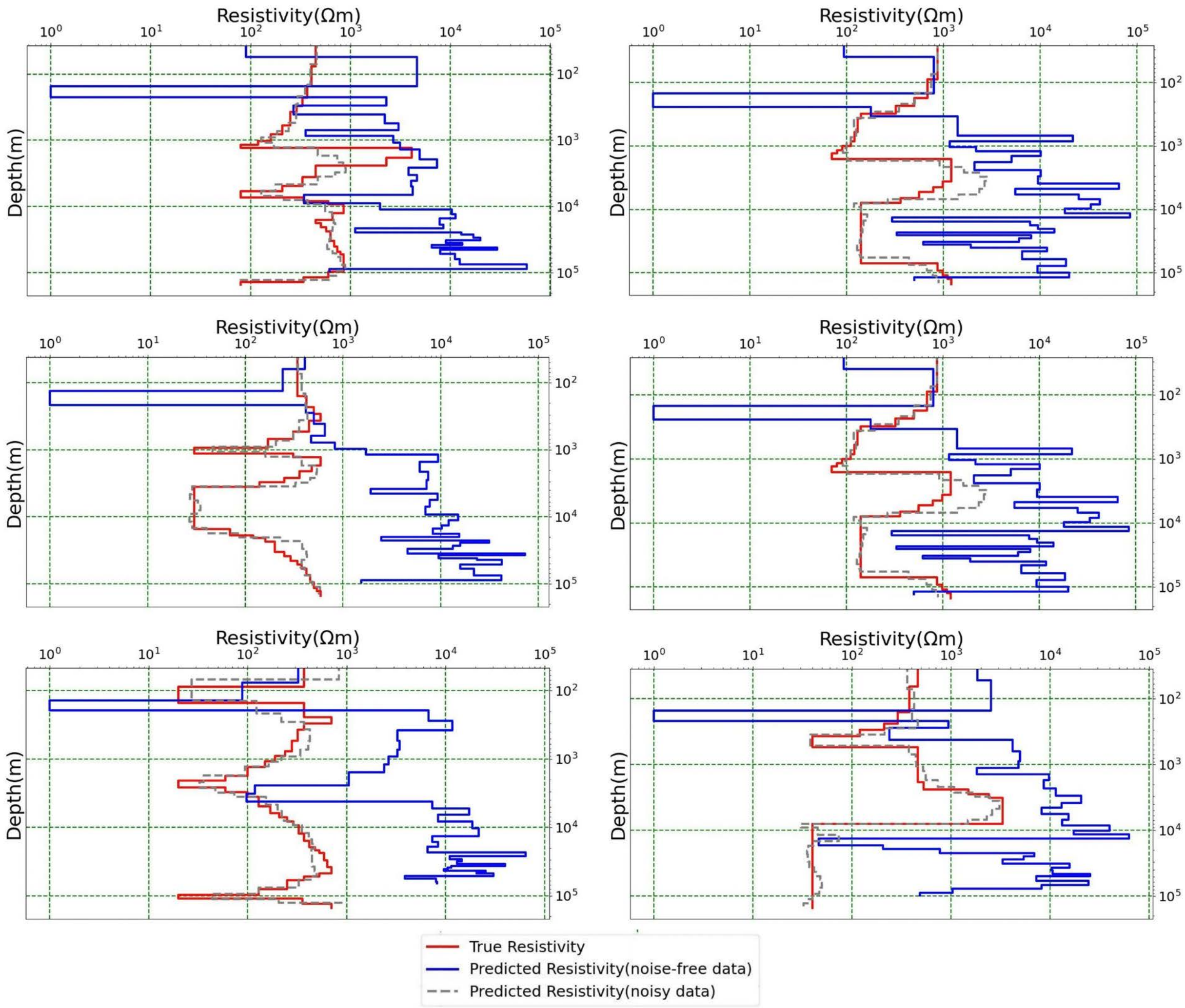

Figure 9. Comparison of predictions made by networks trained on noise-free and noisy data for the first six samples of the noisy test data.

3.2 Quantitative Evaluation

To quantitatively evaluate the inversion performance and measure the degree of dissimilarity between the predicted true models, several distance-based metrics can be employed. Among them, Euclidean, Manhattan, and Minkowski distances are commonly used, with the latter representing a generalized form of the first two. In the case of numerical predictions such as the estimation of subsurface resistivity and thickness in an N-layer earth model, the Euclidean distance (E.D) is widely adapted and can be calculated as follows:

$$\mathrm{E.D} = \sqrt{\sum_{i=1}^{N}\left[\left(\rho_i^{\mathrm{True}} - \rho_i^{\mathrm{predict}}\right)^2 + \left(h_i^{\mathrm{True}} - h_i^{\mathrm{predict}}\right)^2\right]} \tag{4}$$

Since this metric is scale-variant, it is important to use normalized values for resistivity and thickness parameters, as these features have different dimensions and physical nature. Figures 11 and 12 present the histograms of the average normalized Euclidean distance for the predictions obtained from networks trained separately with clean and noisy data for 50000 test samples, including both noise-free and noisy samples. As shown in Fig. 11, histograms (a) and (b) exhibit a high degree of similarity. In other words, except for the relatively poor recovery of the third layer's resistivity and thickness by the network trained on noise-free data, there is no significant difference-based on the Euclidean metric between the predictions of the two networks for clean test data. Conversely, according to Fig. 12b,

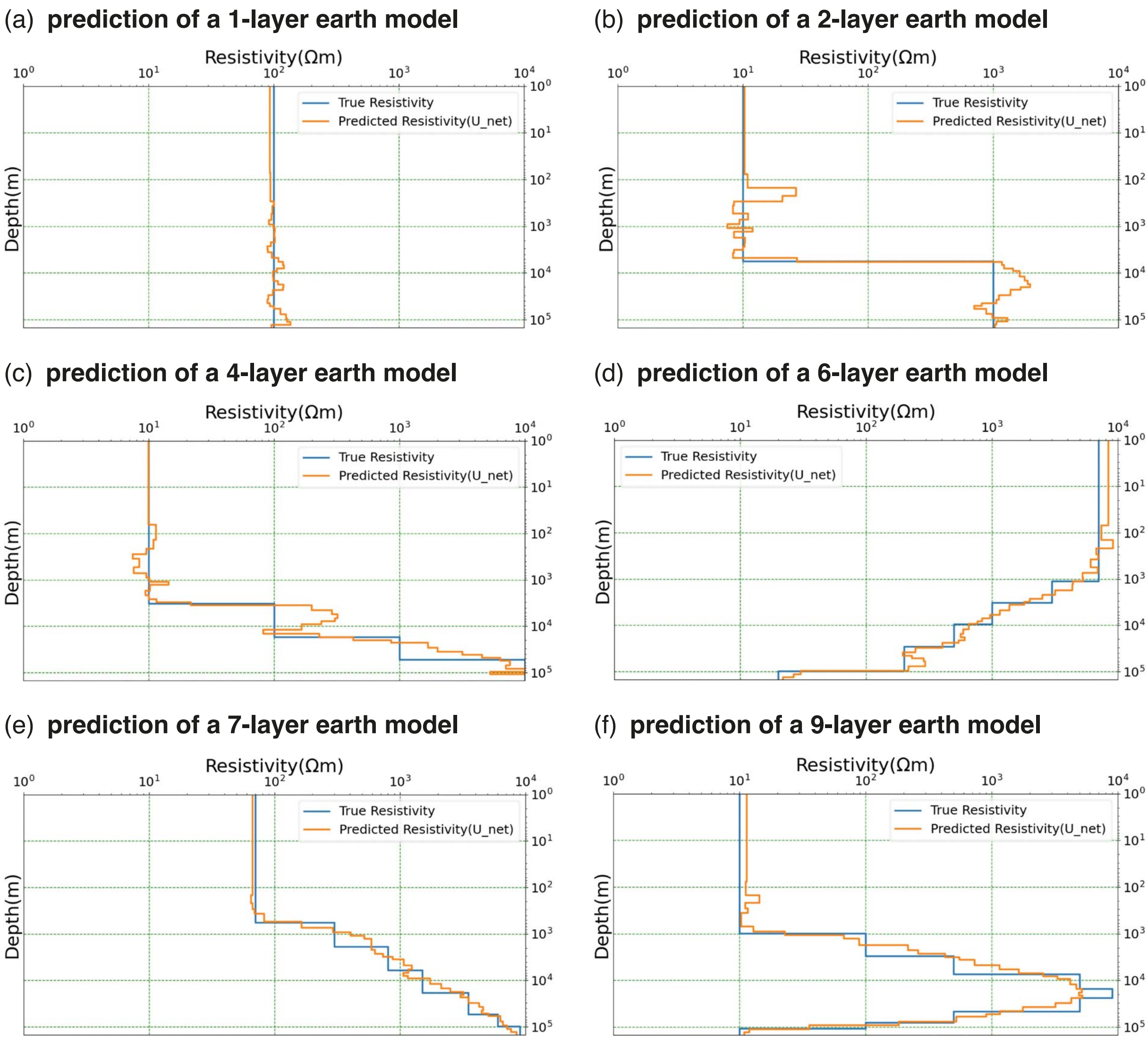

Figure 10. Deep learning model predictions for layered Earth models having less than 35 layers.

about 76% of the test samples show Euclidean distances smaller than 0.1, indicating a noticeable improvement in predictive accuracy in the inversion of noisy data. This observation is also quantitatively consistent with the results displayed in Fig. 9.

The second quantitative criterion employed to assess the network performance is the Root Mean Square (R.M.S) error, a conventional metric that quantifies the misfit between the observed data and the model response derived from forward modeling of the predicted parameters. To avoid excessive detail in this section, the corresponding R.M.S results are presented in electronic supplement. These results reconfirm the relatively poor performance of the network trained with noise-free data when applied to noisy test samples, whereas the network trained with noisy data demonstrates stable and satisfactory performance.

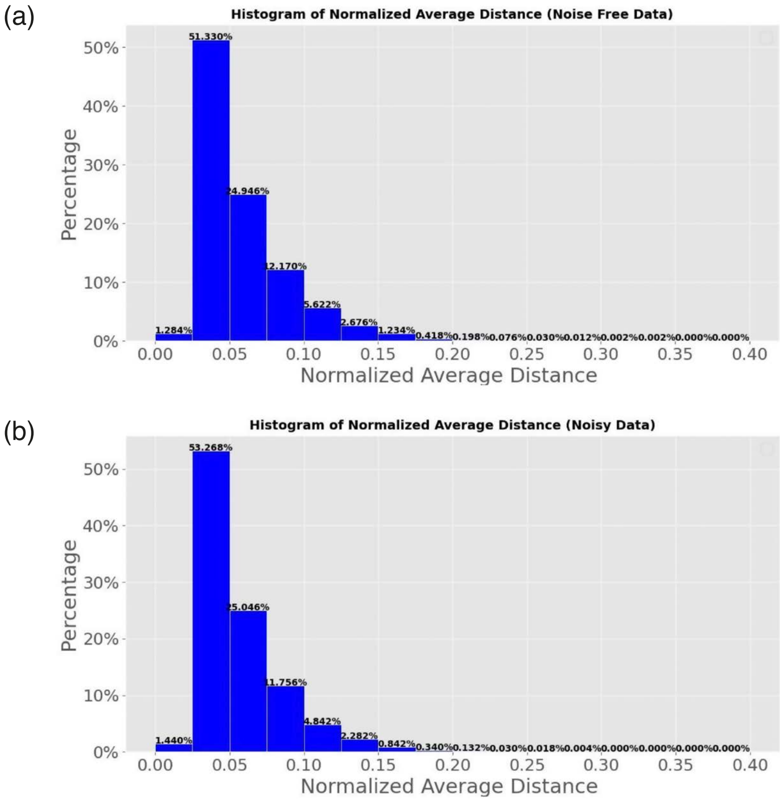

Figure 11. Average Euclidean Distance histograms for 50000 noise-free test samples; obtained by the network trained with a) noise-free and b) noisy data.

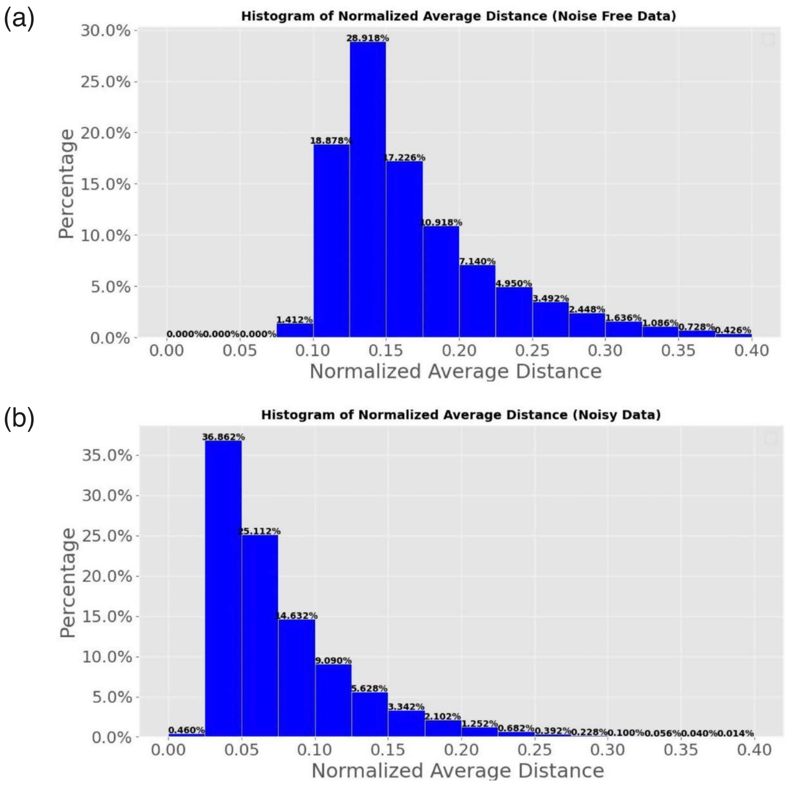

Figure 12. Average Euclidean Distance histograms for 50000 noisy test samples; obtained by the network trained with (a) noise-free and (b) noisy data.

3.3 Comparison to the deterministic methods

In the first stage of comparison of the proposed inversion approach with a deterministic approach, the Newton algorithm with Tikhonov regularization is implemented using the IPI2WinMT package (Bobachev et al., 2001). The algorithm fits the data assuming the minimum number of layers required for a stable inversion and utilizes a nonlinear inversion approach based on the Levenberg-Marquardt optimization method, where the model parameters are iteratively refined until the RMS misfit between measured and calculated data falls below a predefined threshold. Although an option for automatic inversion is available, an experienced interpreter can utilize prior geological information to regularize the minimization process and select the most geologically reasonable model. Figure 13 presents a comparison among the true models, network predictions, and automatic inversion results obtained from IPI2WinMT for six randomly selected test samples. Only in cases (a) and (b), the IPI2WinMT inversion results are consistent with both the true model and the network predictions, whereas in the other cases, significant discrepancies are observed. In cases (c)-(f), spurious layers appear at various depths in the IPI2WinMT inversion results with the resistivity of some thin layers overestimated by more than a factor of 10, which should be disregarded by interpreters relying on complementary information such as geological and other geophysical data. The deep learning-based model, without requiring such auxiliary inputs, demonstrates superior robustness and accuracy in reconstructing the true subsurface resistivity distribution.

In the second comparison, the Occam inversion code (Constable et al., 1987) implemented in the WinGLink software is applied. The initial settings includes a starting resistivity of 100 Ω·m, 30 iterations, 42 layers, and a depth range between 50 and 150,000 m. Figure 14 shows a comparison among the network predictions, the true models, and

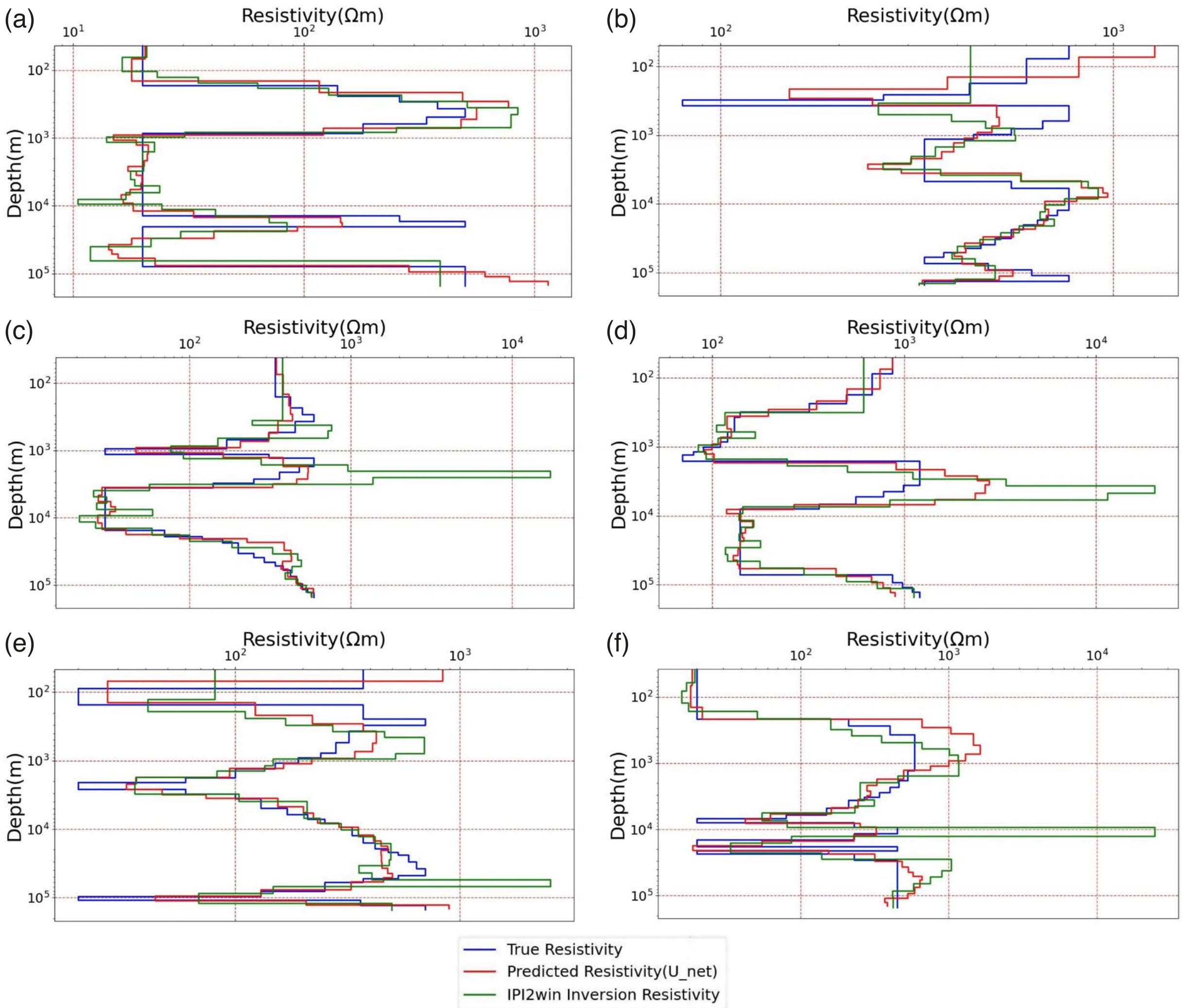

Figure 13. Comparison of deep learning model predictions, the true model, and IPI2Win inversion results.

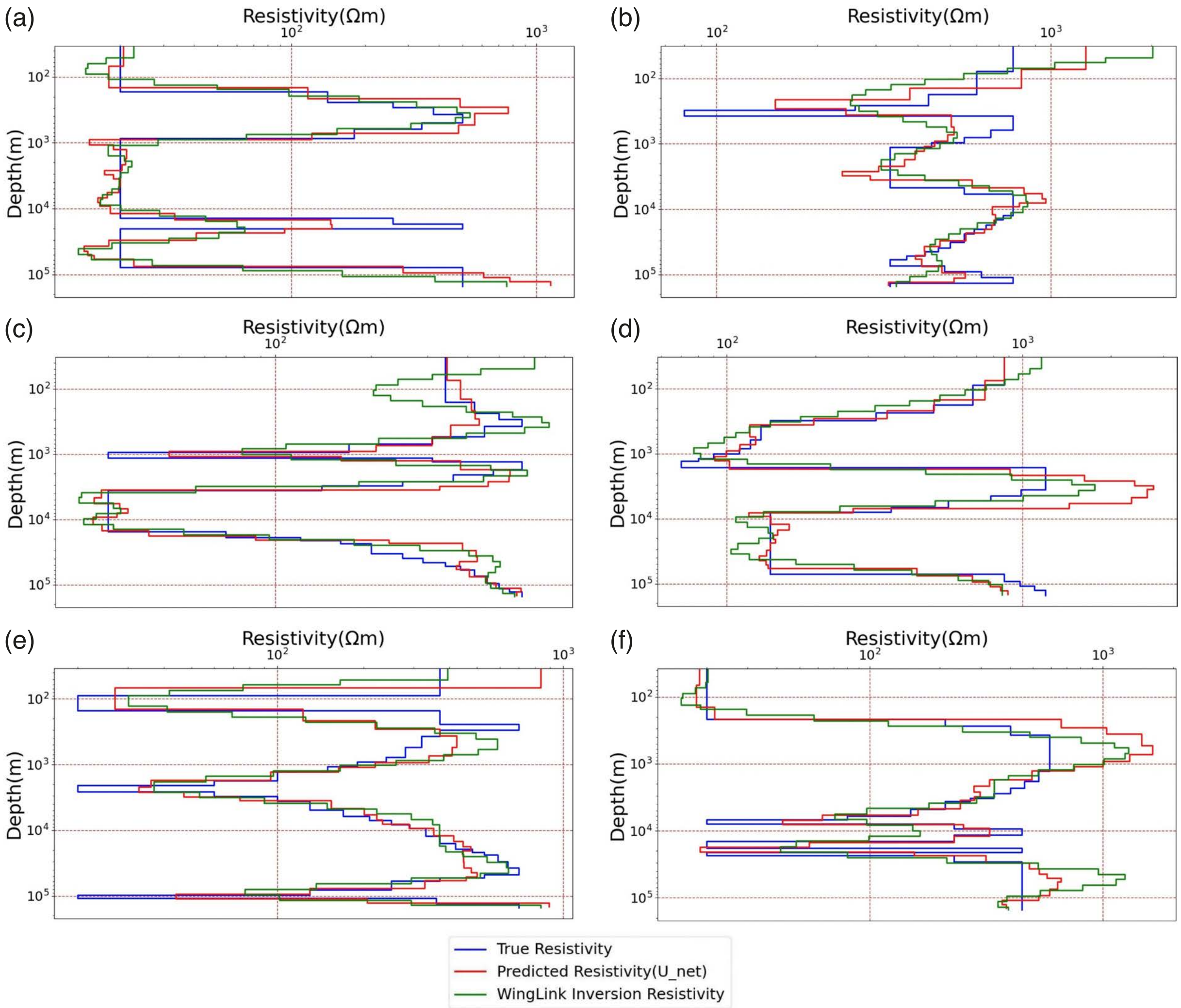

Figure 14. Dimensionality analysis for a typical site obtained from WALDIM code (Martí et al., 2009).

the inversion results obtained using the Occam code. A relatively good agreement is observed among these models; however, the Occam approach exhibits more fluctuations in resistivity values. Overall, the deep learning-based method provides a more accurate one-dimensional subsurface model without requiring the manual adjustment of parameters and prior assumptions, which-if chosen incorrectly-may adversely affect the inversion results. It is notable that the proposed deep learning method is here compared with a subset of deterministic inversions that involve Tikhonov regularization based on model gradient.

3.4 Application to Field Data

To assess the flexibility and generalization capability of the trained deep learning model, it is applied to a field MT dataset acquired for geothermal exploration in the Burwash Landing Region, Canada (Tschirhart et al., 2022). Dimensionality analysis is performed using the rotational WAL invariants (Weaver et al., 2000), implemented through the WALDIM code (Martí et al., 2009), to identify those sites presenting predominantly one-dimensional behavior. A representative example of the dimensionality analysis is illustrated in Fig. 15. At periods shorter than approximately 1 s, the MT responses are mainly one-dimensional, as supported by the near-circular phase tensor ellipses and the small values of the λ parameter (Caldwell et al., 2004), not shown here. The inversion results produced by the deep learning model for the xy and yx impedance polarizations of the same site used in Fig. 15 are presented in Fig. 16. For depths greater than approximately 10 km corresponding to long periods where two and three-dimensional effects dominate, the two models do not fully coincide. However, for depths shallower than 10 km, the models show a high degree of consistency. The deep learning inversion models agree well with 3-D deterministic inversion results presented in Tschirhart et al. (2022) with the conductive layer at ~200-250 m in the 1-D models corresponding to the C1 conductor in the 3-D model that is attributed to graphitic parts of the Kluane Schist and with the increase in conductivity at ~7000 m depth in the 1-D models corresponding to a similar increase In the 3-D model. These results indicate that the proposed deep learning-based inversion presents reliable generalization when applied to real MT data under valid 1-D assumption.

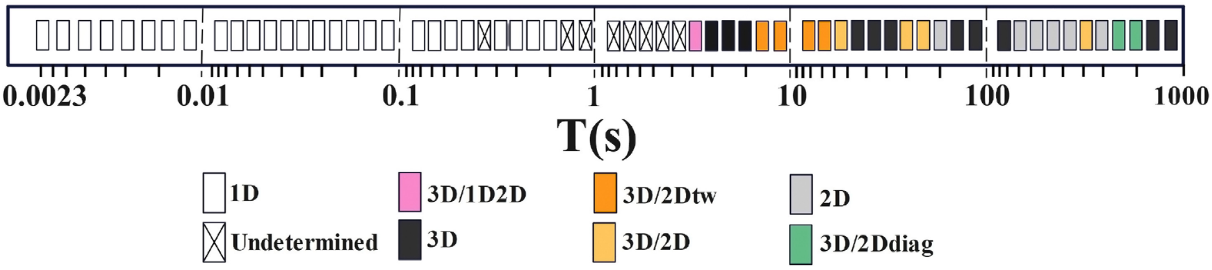

Figure 15. Comparison of deep learning model predictions, the true model, and Occam's algorithm inversion results.

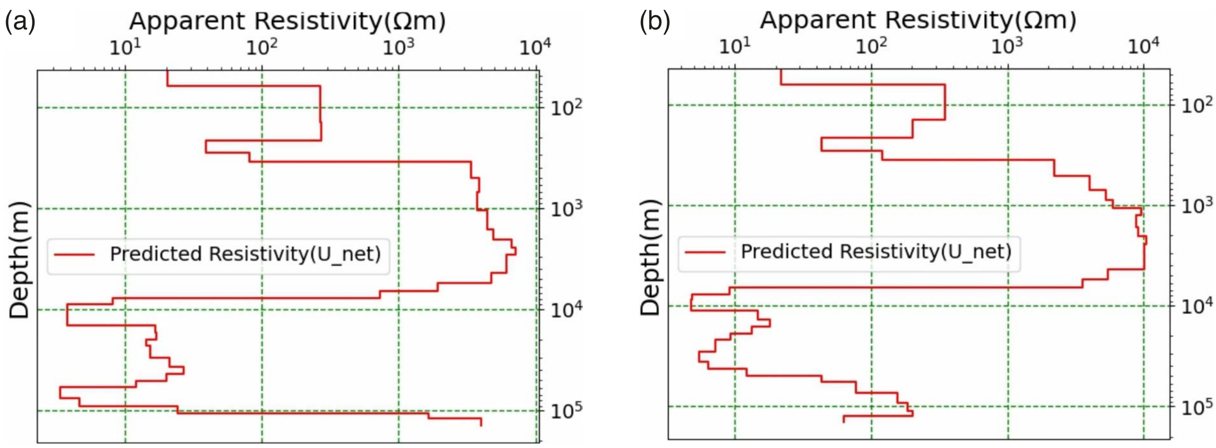

Figure 16. Resistivity model derived from deep learning-based inversion of (a) xy and (b) yx polarization data.

4. Conclusions

In this study, a randomized scheme through a probabilistic approach is employed to generate diverse layered resistivity models, enabling the construction of a comprehensive training dataset that effectively captures a wide range of geological scenarios. By incorporating Gaussian noise and applying clustering-based data splitting, a deep learning framework is established with enhanced generalization capability when faced with variabilities and inaccuracies in real MT measurements.

The performance of the proposed network is assessed through both qualitative and quantitative metrics. To further verify its robustness, the trained network is tested on resistivity structures that are not included in original designed set. The high degree of agreement between the predicted and true models confirms the reliability and stability of the proposed inversion strategy.

Comparison with deterministic inversion approaches suggests that the deep learning framework implicitly performs an effective form of regularization. This implicit regularization arises from the statistical distribution of resistivity layers in the synthetic models used during training and guides the network to the generation of geologically reasonable models, while mitigating overfitting.

This study has demonstrated the proposed deep learning method with enhanced generalization capabilities provides correct and efficient MT inversions of a diverse range of 1-D resistivity structures, as well as providing accurate inversion of a sample of real MT data. Future studies include comparison of the proposed method with a broader suite of deterministic inversions, such as those incorporating model constraints, comparison with a selection of deep learning inversions, and application to other real MT datasets. It will also be of value to apply the generalized approach to multi-dimensional MT data and to other forms of geophysical data.

Acknowledgements. The authors would like to express their sincere appreciation to Dr. James Craven from Geological Survey of Canada Central Division for generously providing the magnetotelluric dataset used in this study.

References

An, G. (1996). The Effects of Adding Noise During Backpropagation Training on a Generalization Performance, Neural Computat., 8, 3, 643-674, doi:10.1162/neco.1996.8.3.643.

Bishop, C. M. (1995). Training with noise is equivalent to Tikhonov regularization, Neural Computat., 7, 1, 108-116, doi:10.1162/neco.1995.7.1.108.

Bobachev, A. A, I. N. Modin and V. A. Shevnin (2001). IPI2Win v.2 for VES data interpretation: User's manual. Moscow State University, Department of Geophysics.

Caldwell, T. G., H. M. Bibby and C. Brown (2004). The magnetotelluric phase tensor, Geophys. J. Int 158, 457-469, doi:10.1111/j.1365-246X.2004.02281.x.

Chave, A. D. (2012). Estimation of the magnetotelluric response function. In Chave, A.D. and Jones, A.G. (eds.) The Magnetotelluric method: Theory and practice, Cambridge University Press, 347-420.

Constable, S. C., R. L. Parker and C. G. Constable (1987). Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, Geophysics 52, 3, 289-300, doi:10.1190/1.1442303.

Goodfellow, I., Y. Bengio and A. Courville (2016). Deep Learning. MIT Press.

He, K., X. Zhang, S. Ren and J. Sun (2016). Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778, doi:10.1109/CVPR.2016.90.

Jones, A. G. (1982). On the electrical crust – mantle structure in Fennoscandia: no Moho, and the asthenosphere revealed? Geophysical Journal International, 68, 371-388., doi:10.1007/BF01453996.

Jones, A. G. (1993). Electromagnetic images of modern and ancient subduction zones. Tectonophysics, 219, 1-3, 29-45, doi:10.1016/0040-1951(93)90285-R.

Kalscheuer, T. and L. B. Pedersen (2007). A non-linear truncated SVD variance and resolution analysis of two-dimensional magnetotelluric models. Geophys. J. Int., 169, 2, 435-447. doi:10.1111/j.1365-246X.2006.03320.x.

Kingma, D. P. and M. Welling (2014). Auto-encoding variational bayes. In: 2nd International Conference on Learning Representation, ICLR, doi:10.48550/arXiv.1312.6114.

Kodinariya, T. M. and P. R. Makwana (2013). Review on determining number of Cluster in K-Means Clustering, Int. J. Adv. Res. Compu. Sci. Manag. Studies, 1, 6, 90-95.

Kumar, J., N. Dalal and M. Sethi (2024). Hyperparameters in Deep Learning: A Comprehensive Review, Int. J. Intelligent Syst. Applic. Engin. 12, 4, 4015-4023.

Liao, X., Z. Zhang, O. Yan, Z. Shi, K. Xu and D. Jia (2022). Inversion of 1-Dmagnetotelluric data using CNN-LSTM hybrid network, Arabian J. Geosci. 15, 1430, doi:10.1007/s12517-022-10687-1.

Liu, B., Q. Guo, S. Li, B. Liu et al. (2020). Deep Learning Inversion of Electrical Resistivity Data, IEEE Trans. Geosci. Remote Sensing 99, 1-14, doi:10.1109/TGRS.2020.2969040.

Liu, Z., H. Chen, Z. Ren, J. Tang et al. (2021a). Deep learning audio magnetotellurics inversion using residual-based deep convolutional neural network, J. Appl. Geophys., 188, 104309, doi:10.1016/j.jappgeo.2021.104309.

Liu, W., Z. Xi, H. Wang and R. Zhang (2021b). Two-dimensional deep learning inversion of magnetotelluric sounding data, J. Geophys. Engin., 18, 627-641, doi:10.1093/jge/gxab040.

Liu, W., H. Wang, Z. Xi, R. Zhang and X. Huang (2022). Physics-Driven Deep Learning Inversion with Application to Magnetotelluric, Remote Sensing 14, 3218, doi:10.3390/rs14133218.

Liu, X., J. A. Craven, V. Tschirhart and S. E. Grasby (2024). Estimating Three-Dimensional Resistivity Distribution with Magnetotelluric Data and a Deep Learning Algorithm, Remote Sensing 16, 18, 3400, doi:10.3390/rs16183400.

Lopez-Alvis, J., F. Nguyen, M. C. Looms and T. Hermans (2022). Geophysical inversion using a variational auto encoder to model an assembled spatial prior uncertainty, J. Geophys. Res.: Solid Earth 127, e2021JB022581, doi:10.1029/2021JB022581.

MacQueen, J. (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, 281-297, https://projecteuclid.org/euclid.bsmsp/1200512992.

Martí, A., P. Queralt and J. Ledo (2009). WALDIM: A code for the dimensionality analysis of magnetotelluric data using the rotational invariants of the magnetotelluric tensor, Compu. Geosci., 35, 2295-2303, doi:10.1016/j.cageo.2009.03.004.

Oh, S., N. Kyubo, S. J. Seol and J. Byun (2020). Cooperative deep learning inversion of controlled-source electromagnetic data for salt delineation, Geophysics 85, 4, E121-E137, doi:10.1190/GEO2019-0532.1.

Pintea, S. L., S. Sharma, F. C. Vossepoel, J. C. van Gemert, M. Loog and D. J. Verschuur (2022). Seismic inversion with deep learning, Computat. Geosci., 26, 351-364, doi:10.1007/s10596-021-10118-2.

Puzyrev, V. (2019). Deep learning electromagnetic inversion with convolutional neural networks, Geophys. J. Int., 218, 817-832, doi:10.1093/gji/ggz204.

Puzyrev, V. and A. Swidinsky (2021). Inversion of 1D frequency- and time domain electromagnetic data with convolutional neural networks, Compu. Geosci., 149, 104681, doi:10.1016/j.cageo.2020.104681.

Rahmani Jevinani, M., B. Habibian Dehkordi, I. J. Ferguson and M. H. Rohban (2024). Deep learning-based 1-D magnetotelluric inversion: performance comparison of architectures, Earth Sci. Informatics, 17, 1663-1677, doi:10.1007/s12145-024-01233-6.

Rodi, W. L. and R. L. Mackie (2012). The inverse problem. In Chave, A. D. and A. G. Jones (eds.) The magnetotelluric method: Theory and practice, Cambridge University Press, 347-420.

Ronneberger, O., P. Fischer and T. Brox (2015). U-Net: Convolutional networks for biomedical image segmentation, Lecture Notes Compu. Sci., 9351, Springer, Cham, doi:10.1007/978-3-319-24574-4_28.

Shahriari, M., D. Pardo, A. Picon, A. Galdran et al. (2020). A deep learning approach to the inversion of borehole resistivity measurements, Computat. Geosci. 24, 971-994, doi:10.1007/s10596-019-09859-y.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions, Ann. Math. Stat., 19, 2, 279-281, doi:10.1214/aoms/1177730256.

Tschirhart, V., M. Colpron, J. Craven, F. H. Ghalati and R. J. Enkin (2022). Geothermal exploration in the Burwash Landing Region, Canada, using three-dimensional inversion of passive electromagnetic data, Remote Sensing 14, 5963, doi:10.3390/rs14235963.

Wang, H., W. Liu and Z. Xi (2019). Nonlinear inversion for magnetotelluric sounding based on deep belief network, J. Central South University, 26, 9, 2482-2494, doi:10.1007/s11771-019-4188-2.

Wang, H., Y. Liu, C. Yin, J. Li et al. (2022). Stochastic inversion of magnetotelluric data using deep reinforcement learning, Geophysics 87, 1, E49-E61, doi:10.1190/geo2020-0425.1.

Weaver, J. T., A. K. Agarwal and F. E. M. Lilley (2000). Characterization of the magnetotelluric tensor in terms of its invariants, Geophys. J. Int., 141, 321-336, doi:10.1046/j.1365-246x.2000.00089.x.

Yadav, A., K. Yadav and S. Anirbid (2021). Feedforward Neural Network for joint inversion of geophysical data to identify geothermal sweet spots in Gandhar, Gujarat, India, Energy Geoscience 2, 189-200, doi:10.1016/j.engeos.2021.01.001.

Zhang, L. and E. Pavlick (2025). Does Training on Synthetic Data Make Models Less Robust?. In: cs.CL. doi:10.48550/arXiv.2502.07164.

Appendix A. R.M.S misfit error

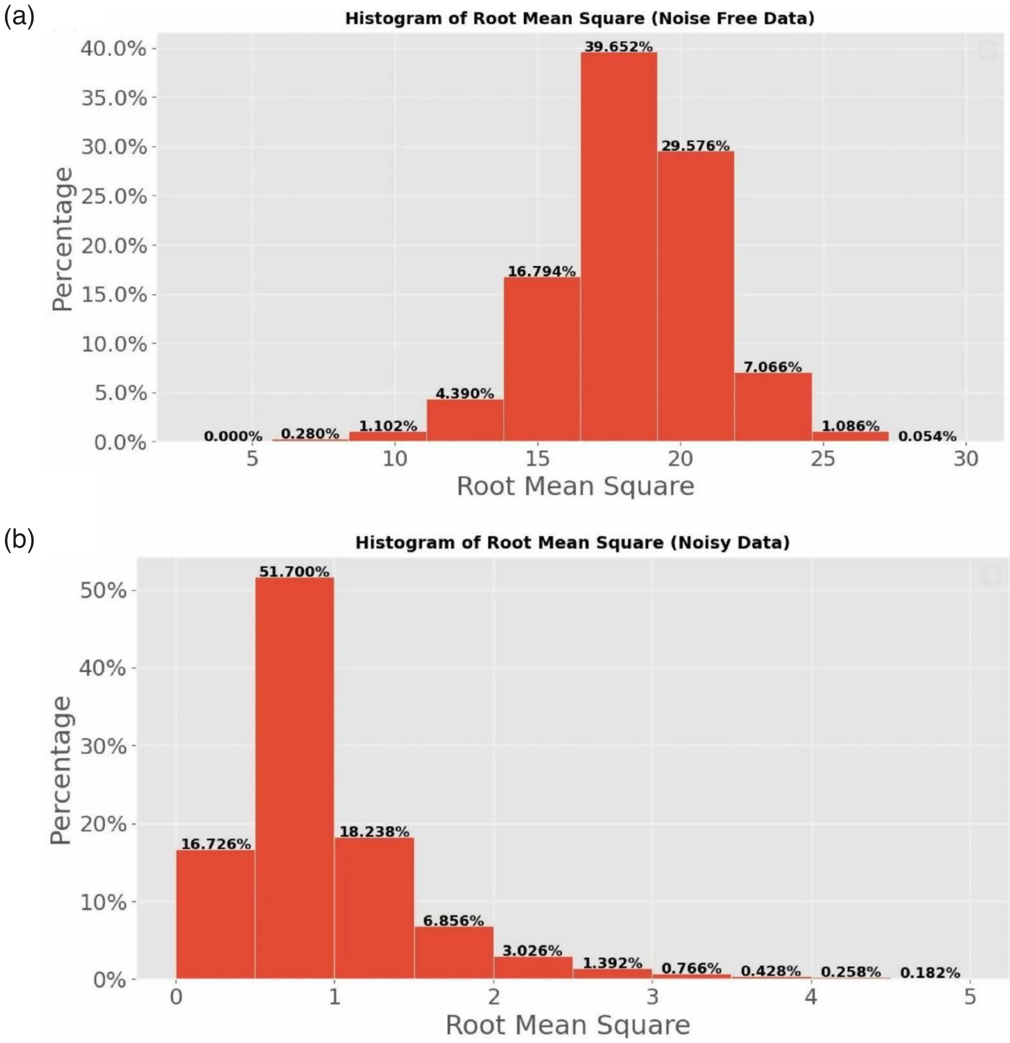

Figure A1. R.M.S error for 50000 noise-free test samples; obtained by the network trained with (a) noise-free and (b) noisy data.

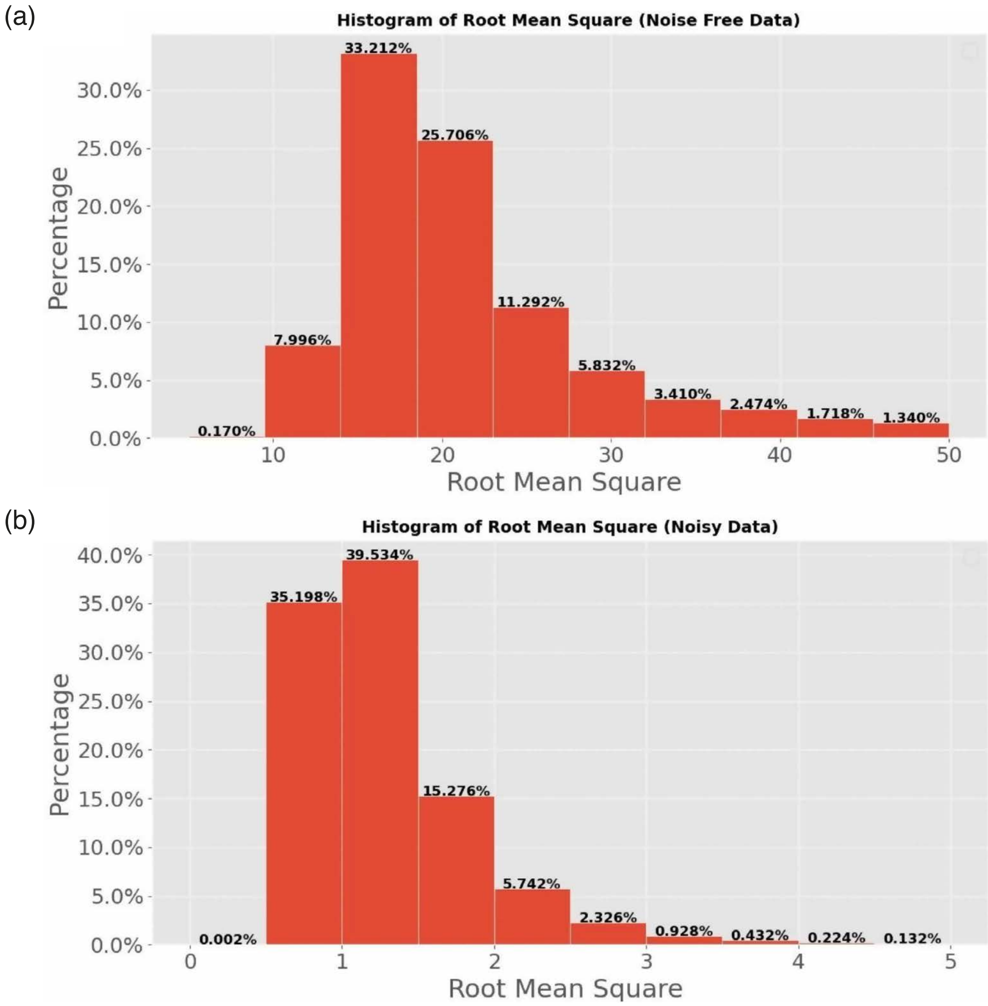

Figure A2. R.M.S error for 50000 noisy test samples; obtained by the network trained with (a) noise-free and (b) noisy data.

***CORRESPONDING AUTHOR:** Banafsheh **HABIBIAN DEHKORDI**,
Institute of Geophysics, University of Tehran, Tehran, Iran
e-mail: bhabibian@ut.ac.ir