JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2024JH000536

Special Collection:
Advanced machine learning in solid earth geoscience

Rapid Bayesian Imaging of Large-Scale Transient Electromagnetic Data Using Probabilistic Neural Networks

Jian Chen[1], Stefano Galli[1], Alessandro Signora[1], Nicole Anna Lidia Sullivan[2], Bo Zhang[3], and Gianluca Fiandaca[1,2]

[1]The EEM Team for Hydro & eXploration, Department of Earth Sciences "Ardito Desio", University of Milano, Milano, Italy, [2]The EEM Team Spin-Off Company, Milano, Italy, [3]Institute of Earth Exploration Science and Technology, Jilin University, Changchun, China

Key Points:

- We estimated posterior resistivity distribution of electromagnetic data using a probabilistic neural network, considering data uncertainty
- Utilizing the posterior distribution to extract useful information, such as a smooth constrained model and depth of investigation
- We verified this network with 200 km survey data, in which bathymetric data support the quality of the subsurface reconstruction

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:
J. Chen,
jian.chen@unimi.it

Citation:
Chen, J., Galli, S., Signora, A., Sullivan, N. A. L., Zhang, B., & Fiandaca, G. (2025). Rapid Bayesian imaging of large-scale transient electromagnetic data using probabilistic neural networks. *Journal of Geophysical Research: Machine Learning and Computation*, *2*, e2024JH000536. https://doi.org/10.1029/2024JH000536

Received 27 NOV 2024
Accepted 12 APR 2025

Author Contributions:
Conceptualization: Jian Chen, Gianluca Fiandaca
Funding acquisition: Gianluca Fiandaca
Investigation: Jian Chen, Stefano Galli, Alessandro Signora
Methodology: Jian Chen
Resources: Gianluca Fiandaca

Abstract Transient electromagnetic (TEM) can conduct efficient large-scale geological surveys on the near-surface. Exploring fast and comprehensive interpretation strategies for electromagnetic data is crucial for geoscientists to make high-quality decisions on site. In this study, we proposed a probabilistic neural network (PNN) structure to estimate the posterior probability density function (PDF) of model parameters. The feature of this structure is that it uses noisy data and the data standard deviation information as inputs of the training data set, and the model parameters retrieved through deterministic inversion are used as labels. Such a structure enables the posterior PDF output by the PNN to take into account the uncertainty information of the input data itself, and allows us to add existing field data to the training data set to continuously enrich reasonable prior information. Additionally, we aim to extract useful information from the posterior PDF, including smooth models similar to those obtained through laterally or spatially constrained inversion, as well as the estimation of the depth of investigation of the imaging results. The PNN structure was verified using 200 km of waterborne TEM survey data. The results shows that the PNN network efficiently delineated the subsurface electrical property distribution of a large-scale lake water system, and the lake depth and depth-uncertainty extracted from the imaging results demonstrated good consistency with the sonar bathymetric data. Besides, the smooth model extracted from the resistivity posterior PDF estimated by PNN not only improves the smoothness of the model but also reduces the data misfit.

Plain Language Summary Due to the nonunique solution and noise interference in geophysical electromagnetic data, there are multiple interpretation results for the electromagnetic response data. Probabilistic neural networks (PNNs), such as mixture density networks (MDNs), can predict the posterior distribution of the model and have unique advantages in solving the non-unique inversion problem. However, how to obtain a representative posterior distribution while considering the existence of the uncertainty of the data itself remains a major challenge. To address this issue, we have designed a PNN structure that can provide Bayesian probabilistic imaging results and employ traditional deterministic inversion to construct the training data set. This training strategy takes into account the uncertainty inherent in the data and allows for the subsequent addition of field-measured data to enrich the training data set. We also discussed how to utilize the obtained posterior distribution to estimate useful information, including extracting a smooth model analogous to the results of laterally constrained inversion and estimating the depth of investigation, both of which can provide rapid decision-making support for field crews. The PNN structure successfully estimated the electrical parameters of a large-scale lake groundwater system within tens of seconds.

1. Introduction

As mobile electromagnetic detection technologies such as airborne transient electromagnetics (ATEM) (Chandra et al., 2021; Silvestri et al., 2019), ground-towed transient electromagnetics (Auken et al., 2019), and waterborne transient electromagnetics (FloaTEM) (Maurya, Christensen, et al., 2022) have matured, transient electromagnetic (TEM) methods are now able to conduct high-efficiency, large-scale geological surveys across terrestrial, lacustrine, and aerial environments. For instance, the tTEM technology facilitates mobile geological surveys using all-terrain vehicles or boats, offering a mobility similar to ATEM counterparts. The tTEM system operates at speeds up to 20 km per hour (Grombacher et al., 2021), enabling efficient and cost-effective geological surveys over extensive areas, with a typical depth of investigation (DOI) ranging from 80 to 100 m. Due to its smaller footprint and increased data acquisition density, the tTEM offers superior horizontal and vertical resolution in

Software: Nicole Anna Lidia Sullivan, Gianluca Fiandaca
Supervision: Gianluca Fiandaca
Validation: Jian Chen
Visualization: Nicole Anna Lidia Sullivan
Writing – original draft: Jian Chen
Writing – review & editing: Bo Zhang, Gianluca Fiandaca

shallow subsurface detection, particularly within the first 20 m (Maurya et al., 2023), compared to the ATEM method. Currently, electromagnetic exploration technology has successfully been implemented in large-scale geological surveys for applications including mapping of groundwater hydrology systems (Grombacher et al., 2021), assessing groundwater vulnerability (Sandersen et al., 2021), and probing deep geothermal fluids (Finn et al., 2022).

When conducting large-scale survey work, obtaining rapid and comprehensive inversion interpretation results is crucial for enabling geophysical engineers to accurately assess geological features and make informed decisions. In the existing interpretation methods, TEM still primarily relies on smooth constrained deterministic inversion (including laterally or spatially constrained inversion (Auken et al., 2008)) based on one-dimensional (1D) forward operators to provide fast interpretation (Maurya, Grombacher, et al., 2022). The smooth constrained inversion method uses neighboring measurement points as constraints and generally provides better quality than conventional single-point inversion methods (Viezzoli et al., 2009). For deterministic inversion methods, the Jacobian matrix can be employed in conjunction with the model covariance matrix to estimate the uncertainty of the current model parameters (Fiandaca et al., 2013). However, the presence of data noise, model nonlinearities, and inherent imperfections leads to the existence of an infinite number of models capable of fitting the observed data (Chen et al., 2021; Xue et al., 2020). Consequently, deterministic inversion approaches are generally unable to properly account for such uncertainties (Hansen & Finlay, 2022; Zaru et al., 2024). In contrast, Bayesian inversion methods based on Markov chain Monte Carlo (MCMC) sampling provide a more robust approach to uncertainty estimation. By generating a large number of models that fit the observed data, Bayesian inversion enables the computation of a posterior probability density function (PDF) of resistivity and has been widely adopted in the inversion of geophysical electromagnetic data (Blatter et al., 2018; Chen et al., 2022a; Killingbeck et al., 2020). Nevertheless, the computational cost of such sampling-based methods is substantial due to the extensive sampling required, making them time-intensive. To address this issue, Hansen (2021) proposed an extended rejection sampler strategy, which involves utilizing a pre-calculated lookup tables to estimate the posterior PDF of frequency domain airborne electromagnetic data. This strategy eliminates the need for additional forward modeling computations compared to traditional Bayesian inversion, thereby significantly improving computational efficiency. However, for large-scale electromagnetic data sets, achieving near real-time model uncertainty analysis in field applications remains a significant challenge.

In recent years, numerous studies have demonstrated the feasibility of deep learning algorithms for real-time imaging of TEM data, including work by Colombo et al. (2021), Puzyrev (2019), Puzyrev et al. (2021), Shi and Cao (2022), Li et al. (2022), Wu et al. (2021a, 2021b), Chen et al. (2022b), and Asif et al. (2022). These studies principally leverage deep neural network (DNN) frameworks to either establish a specific mapping between TEM data and resistivity parameters or learn partial derivatives to speed up the inversion process. This type of single-output DNN is capable of rapidly generating model parameters akin to those derived from deterministic inversion. However, the above studies did not provide estimates of uncertainties since for each input data vector their neural networks only predict single-output model parameters vector. Additionally, the inherent nonunique mapping relationships within the constructed geophysical training data set present significant challenges for training single-output DNN networks.

Given these training challenges and the impact on the reliability of deep learning networks in interpreting TEM data, probabilistic neural networks (PNNs) have emerged as an effective solution for addressing the nonlinear inversion problems encountered in geophysics. Devilee et al. (1999) were among the early researchers to utilize PNNs for the inversion of seismic wave velocities, successfully producing a maximum-likelihood map of crustal thickness across Eurasia. One of the classical structures within PNNs is the Mixture Density Network (MDN). MDNs leverage the capability to map an input vector to an n-dimensional conditional probability distribution, parameterizing this distribution as a Gaussian mixture model (GMM) to learn and model arbitrary probability distributions. Earp and Curtis (2020) employed a MDN to perform a two-dimensional Bayesian seismic travel time tomography, demonstrating that the prior information used to construct the training data set is crucial for the quality of network training. Mosher, Eilon, et al. (2021) and Mosher, Audet, and Gosselin (2021) further applied this method to the inversion of shear wave velocity structures in marine seismology, showing that MDN inversion is consistent with traditional MCMC inversion techniques, which facilitates the standardization of geophysical inversions. Additionally, MDNs have exhibited excellent performance in the inversion of electromagnetic data (Yu et al., 2024), surface wave dispersion data (Earp et al., 2020) and geoacoustic data (Wu et al., 2021e). However, the training of MDN networks can be unstable in high-dimensional problems, which may lead to

suboptimal performance in multimodal predictions (Earp & Curtis, 2020; Zhang & Curtis, 2021). To address this issue, Alyaev and Elsheikh (2022) adopted the Multiple-Trajectory-Prediction (MTP) strategy to improve the design of the MDN loss function, effectively avoiding mode collapse.

Another typical PNN is the Invertible Neural Network (INN). INN structures are designed to perform bidirectional mapping between inputs and outputs, estimating the posterior probability density function (PDF) through the introduction of latent variables on the output side, which is crucial for accurate Bayesian inference (Ardizzone et al., 2018). Zhang and Curtis (2021) compared the performance of INNs, MDNs, and traditional MCMC methods in estimating one-dimensional surface wave dispersion inversion and two-dimensional travel time tomography. Their results indicated that INN networks have a significant advantage in inferring the correlations between model parameters. Wu et al. (2023) applied the INN method to the Bayesian inversion of electromagnetic data, demonstrating that INN methods can obtain marginal distributions similar to those derived from MCMC methods. However, since not all types of layers can be designed to be invertible, INN networks typically require specific designs, which can limit their flexibility. Moreover, compared to MDN networks, INN networks have higher memory requirements and training costs, making them less feasible for very high-dimensional problems such as full waveform electromagnetic or seismic inversion (Zhang & Curtis, 2021). With the diversification of neural network architectures, many new neural network frameworks have been applied in geophysical probabilistic inversion. For instance, Li et al. (2024) utilized Variational Physics-Informed Neural Networks via Variational Inference for seismic petrophysical inversion. Jiang et al. (2025) combined Transformer models with semi-supervised training strategies to achieve joint inversion and uncertainty estimation of surface waves and receiver functions.

Overall, geophysical probabilistic inversion based on deep learning networks has developed rapidly in recent years, and the importance of assessing the uncertainty of inversion parameters has become widely recognized. However, in the aforementioned studies, more attention has been focused on evaluating whether the posterior PDFs obtained by deep learning networks are representative, with few studies exploring how to extract more valuable information from these posterior PDFs. In actual field work, the direct guidance value of quickly obtained posterior PDFs of model parameters remains limited, especially for on-site personnel who are not experts in this field.

This study focuses on the interpretation needs of large-scale electromagnetic survey data, specifically investigating whether more valuable inversion information can be extracted from the outputs of PNNs and whether existing large-scale data sets can be effectively integrated into training data sets to enrich prior geological information. The MDN can directly predict the posterior PDFs of model parameters, which is beneficial for post-processing. Therefore, this paper designs a Bayesian imaging structure for large-scale electromagnetic data based on the MDN framework. We propose a training data set construction framework that enables us to incorporate existing actual electromagnetic data sets and their inversion models into the training data set. Moreover, we successfully extracted smooth constraint models from the obtained posterior PDFs of resistivity, which are particularly useful in geological systems with good geological continuity. By using the divergence between the posterior PDF and prior knowledge, we can also estimate the DOI of model parameters. Finally, we discuss the impact of adding field measured data and the number of mixed Gaussian kernels on network training performance.

2. Method

2.1. Data Generation

As shown in Figure 1, the Bayesian imaging framework based on the PNN mainly includes three stages: data generation, network construction and training, and imaging result output.

In the data generation stage, the training data set needs to be designed specifically. For electromagnetic data inversion, the uncertainty of model parameters is mainly composed of the noise existing in the observation data, the non-uniqueness of the inversion, and imperfections such as modeling errors caused by using a 1D layered earth approximation (Bai et al., 2021; Deleersnyder et al., 2024). While modeling errors can be mitigated through three-dimensional forward simulations, the 1D approximation remains the preferred approach in both industry and academia due to its computational efficiency (Zaru et al., 2024). Thus, in this study, we still use 1D forward modeling for data set construction. Traditional inversion training data sets typically utilize forward modeling data

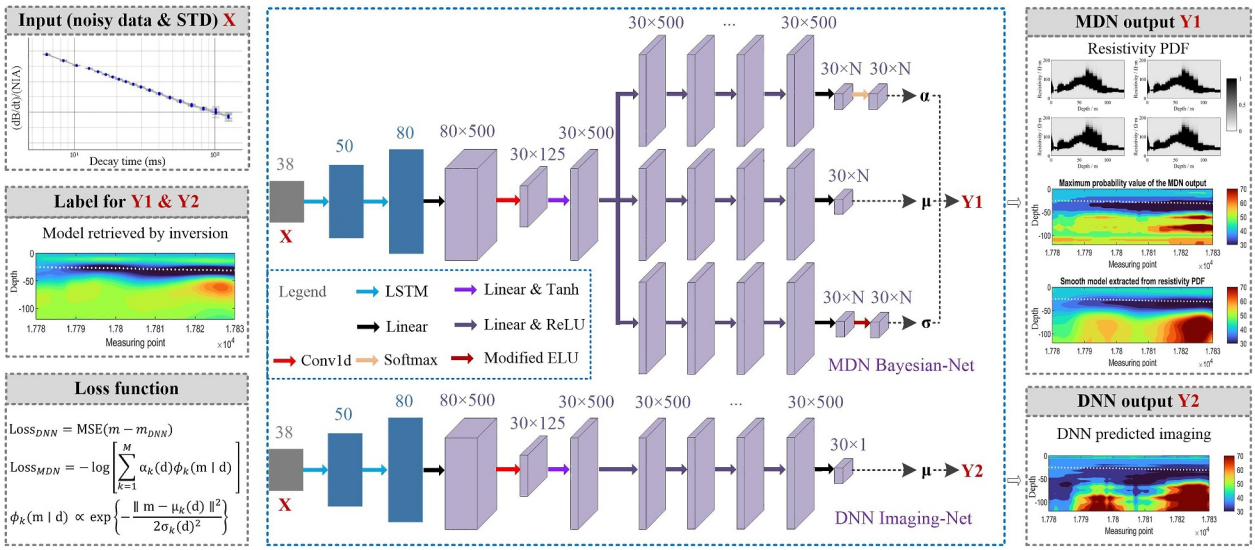

Figure 1. Diagram of the probabilistic neural network (PNN) and the single-output deep neural network (DNN) structures. Both architectures employ transient electromagnetic data and corresponding standard deviation values as input. The target labels are deterministic inversion resistivity models generated using the EEMverter platform (Fiandaca et al., 2024). The DNN outputs the predicted resistivity model, whereas the PNN outputs the mixture density network-derived posterior resistivity distribution function, from which the maximum probability, mean, stochastic, and smooth models are readily derived.

as input and theoretical resistivity models as training labels. Since there are no errors in the data or labels in this kind of training data set, the model uncertainty obtained after training mainly reflects the non-uniqueness of the model, and it is difficult to well reflect the uncertainty caused by data noise. In this research, we utilize both the forward modeling noisy data and the data standard deviation (STD, see Equation 4) information as joint inputs, with the geological models obtained through deterministic inversion serving as training labels.

The specific difference lies in the fact that the construction of typical data sets only requires generating resistivity models and performing forward modeling. In contrast, our data set construction method involves four steps: (a) generating resistivity models and performing forward modeling to obtain noiseless response data; (b) adding noise according to the method described in Section 2.3 and calculating the STD of the data; (c) performing deterministic inversion to obtain inversion model parameters and recording the inversion misfit; (d) using the model parameters with an inversion misfit below the set threshold (a misfit threshold of 1 is used in this study) as the output labels of the training data set, while the corresponding noisy data and STD information are used as the input of the training data set.

$$\text{Misfit} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{\left(d_{\text{obs},i} - d_{fwr,i}\right)^2}{\left(\text{STD}\cdot d_{obs,i}\right)^2}}. \quad (1)$$

Where d_{obs} is the field data, d_{fwr} is the forward data, and N is the number of data time gates. A data misfit near 1 means that the forward data can fit the field data within the errorbar. This data set construction method is more time-consuming than the traditional approach. However, the advantage is that the training labels, which are the model parameters obtained from deterministic inversion, take into account the STD information inherent in the input data. This facilitates the transmission of data uncertainty information to the posterior PDF. Additionally, this data set construction method closely mimics the inversion process of real-world field data, thus allowing us to continuously add field-measured data to enrich the prior information of the training data set.

During the generation of random resistivity models, we set the number of layers for the forward modeling to a fixed 30 layers, with the layer thickness increasing logarithmically over a range of 120 m. The last layer is assumed to be a semi-infinite half-space, thus we only need to define the resistivity values for these 30 layers. Prior information consistent with geological knowledge is very important for a good estimate of the posterior PDF in probabilistic inversion (Hansen & Minsley, 2019). Considering that the electromagnetic instrument system (tTEM and FloaTEM) is primarily applied in hydrogeological surveys, where the range of resistivity variations is

wide and the model labels in the training data set will be derived from regularized deterministic inversion, we imposed implicit prior knowledge in the form of smoothness constraints during model generation. This means that the resistivity parameters of the 30 randomly generated layers vary smoothly along the depth direction. The model resistivity range is from 0.1 Ω·m (such as seawater or saline water) to 10,000 Ω·m (such as rock). We refer to the smooth model generation strategy proposed by Wu et al. (2021d), which mainly includes the following steps: First, 2–4 anomalous layer positions are randomly selected within the 30-layer subsurface space (i.e., within 120 m below the surface), ensuring a minimum separation of 15 m between any two anomalous layers. Second, random values of logarithmic resistivity between −1 and 4 are assigned to these anomalous positions. For the resistivity values of the remaining layers, a smooth model along the depth direction is obtained using cubic spline interpolation based on the resistivity values of the anomalous layers.

This study's training data set comprises a forward modeling data set (405,000 sets of data) and a field-measured data set (15,000 sets of data). The inputs in the field-measured data set are the field data and its corresponding STD, and the labels are the corresponding deterministic inversion models (data with inversion misfit greater than the set threshold are removed). We did not specially design the proportion of field data/models in the training data set, as the amount of field data is always more limited compared to synthetic data, and the field-measured data added here does not include the FloaTEM data used for subsequent testing. We will discuss the potential benefits of adding field measured data in Section 4.

2.2. Network Construction

In the stage of network construction and training, considering that the TEM data used for inversion processing is sparsified by gate integration (Neven et al., 2021), and there exists a complex non-unique mapping relationship between the input data and the output resistivity model, we hope to design a network that can extract the high-dimensional features of the TEM data before entering the MDN Bayesian-Net, so as to accurately evaluate the uncertainty of the resistivity parameters. Based on the time decay characteristics of TEM data and the superior performance of long short-term memory (LSTM) in processing time series data (Wu et al., 2021c), this study adopts an upscaling architecture to design LSTM-Net. The output of LSTM-Net is directly combined with the subsequent MDN Bayesian-Net. To facilitate subsequent comparison, we designed a single-output DNN imaging network using a structure consistent with the PNN predicted mean output. The difference between MDN and single-output DNN is that MDN outputs a resistivity distribution function, while DNN outputs direct resistivity parameters. This characteristic allows MDN to model arbitrary distributions via GMM, offering flexibility in accommodating the varied shapes of posterior distributions that can arise from the data. The number of Gaussian kernels directly influences the complexity of the parameters predicted by the MDN. In this study, we set the number of Gaussian kernels to 3. Section 4 will discuss the impact of different numbers of Gaussian kernels on the training performance of the MDN network.

Suppose we have N training data sets $R = \{(\mathbf{d}_i, \mathbf{m}_i): i = 1, \ldots, N\}$, where $\mathbf{d}$ and $\mathbf{m}$ represent the input space of TEM data and the output space of resistivity model parameters, respectively. Given an input $\mathbf{d}_i$, if the trained resistivity model set $\mathbf{m}_i$ satisfies a prior PDF distribution, the structure of a conventional neural network will output the corresponding $\mathbf{m}_i$ by minimizing the sum of squared errors on the set R. This output result will approximate the mean solution of the Bayesian posterior distribution $p(\mathbf{m}|\,\mathbf{d})$ (Earp et al., 2020). In contrast, MDN can directly output an estimate of the Bayesian posterior distribution $p(\mathbf{m}|\,\mathbf{d})$. For details on MDN principle and loss function design, see Text S1 in Supporting Information S1.

2.3. Noisy Data Generation and Definition of STD

EEMverter (Fiandaca et al., 2024) was utilized for forward calculations on the resistivity model to obtain electromagnetic response data. For the synthetic data obtained by forward modeling, we need to add noise to it and define STD. Referring to the work of Vignoli et al. (2015) and Auken et al. (2008) the perturbed data can be expressed as:

$$V_{\text{resp}} = V + G(0,1)\left[\text{STD}_{\text{uni}}^2 + \left(\frac{V_{\text{noise}}}{V}\right)^2\right]^{\frac{1}{2}} \cdot V. \tag{2}$$

Where V_{resp} is the noisy data, V is the forward modeling noiseless data, G(0,1) is a zero-mean Gaussian distribution with STD 1, STD_{uni} is uniform noise STD and its value is set to 0.03 (i.e., relative percent deviation), which is consistent with the default benchmark STD value of the tTEM/FloaTEM instrument. V_{noise} is the background noise contribution. If the surrounding noise is white (i.e., random and uniformly distributed across all frequencies), applying logarithmic gating leads to an average noise reduction that is proportional to $t^{-1/2}$ (Munkholm et al., 1996). We can write V_{noise} as:

$$V_{noise} = b \cdot \left(\frac{t}{1 \cdot 10^{-3}}\right)^{\frac{-1}{2}}. \tag{3}$$

Where t is gate time, b is the noise level at 1 ms and its value is set to 1 nV/m^2 according to the experience from Auken et al. (2008). The STD of the synthetic data is defined as:

$$STD = \left[STD_{uni}^2 + \left(\frac{V_{noise}}{V}\right)^2\right]^{\frac{1}{2}}. \tag{4}$$

After obtaining the STD of the forward modeling data, we applied the EEMverter software for resistivity inversion of both synthetic and field measured data. During the inversion process, the STD of the data was taken into account, and a constraint term was applied along the depth direction (with a constraint factor set to three, meaning that a penalty would be imposed if the resistivity change between adjacent layers exceeded threefold). This ensures that the inversion results exhibit relatively smooth variations with depth, which is consistent with our prior knowledge in constructing the resistivity model. The inversion objective function is as follows:

$$Q_{inv} = \left[\frac{1}{N_d + N_m - 1}\left(\sum_{i=1}^{N_d}\frac{(d_{obs,i} - d_{forward,i})^2}{(STD_i \cdot d_{obs,i})^2}\right) + \sum_{i=1}^{N_m-1}\frac{(m_i - m_{i+1})^2}{\sigma_m^2}\right]. \tag{5}$$

Where N_d is number of time gates, N_m is number of layers, d_{obs} is the target data to be inverted, $d_{forward}$ is the forward modeling data, m is the inversion parameter, and σ_m is the constraint STD in the vertical direction of the model. Ultimately, the TEM data, the data STD, and the resistivity models obtained from inversion are each subjected to logarithmic transformation and normalization. The normalized data are combined with the STD to form the input data set, while the inverted models are utilized as training labels, thus completing the construction of the training set. The settings of training parameters are shown in Text S2 and Table S1 in Supporting Information S1.

3. Data and Results

3.1. Data

The PNN inversion was tested through a large-scale FloaTEM survey conducted on the south shore of lake Iseo in Italy (see Figure 2), aimed at studying lake-groundwater interaction. This involved a waterborne tTEM survey spanning approximately 200 km with nearly 35,000 survey points. Furthermore, to measure bathymetry accurately, both a sonar sounding device and a GPS device were installed on the boat. Figure 2 illustrates the quality of the FloaTEM survey line data, which we briefly processed to align the data time gates with the PNN training data set. Notably, the data quality is high and the STD is minimal in areas away from the shore. Figure S1 in Supporting Information S1 depicts the data quality near the shore, where it notably degrades. Increased STD in the late time gate data is attributed to interference from nearby human-made structures and power lines. Additionally, significant local changes in data amplitude are observed. This is linked to the lakeshore's geological connections with adjacent mountains. Rocks extending from these mountains into the lake bed introduce high-resistance anomalies, attenuating the electromagnetic response.

3.2. PNN Inversion Reveals the Interaction Between Lake Water and Groundwater

The inversion results of the PNN network were compared with those obtained from the laterally constrained inversion based on the EEMverter modeling platform (Fiandaca et al., 2024). The number of inversion layers and

10.1029/2024JH000536

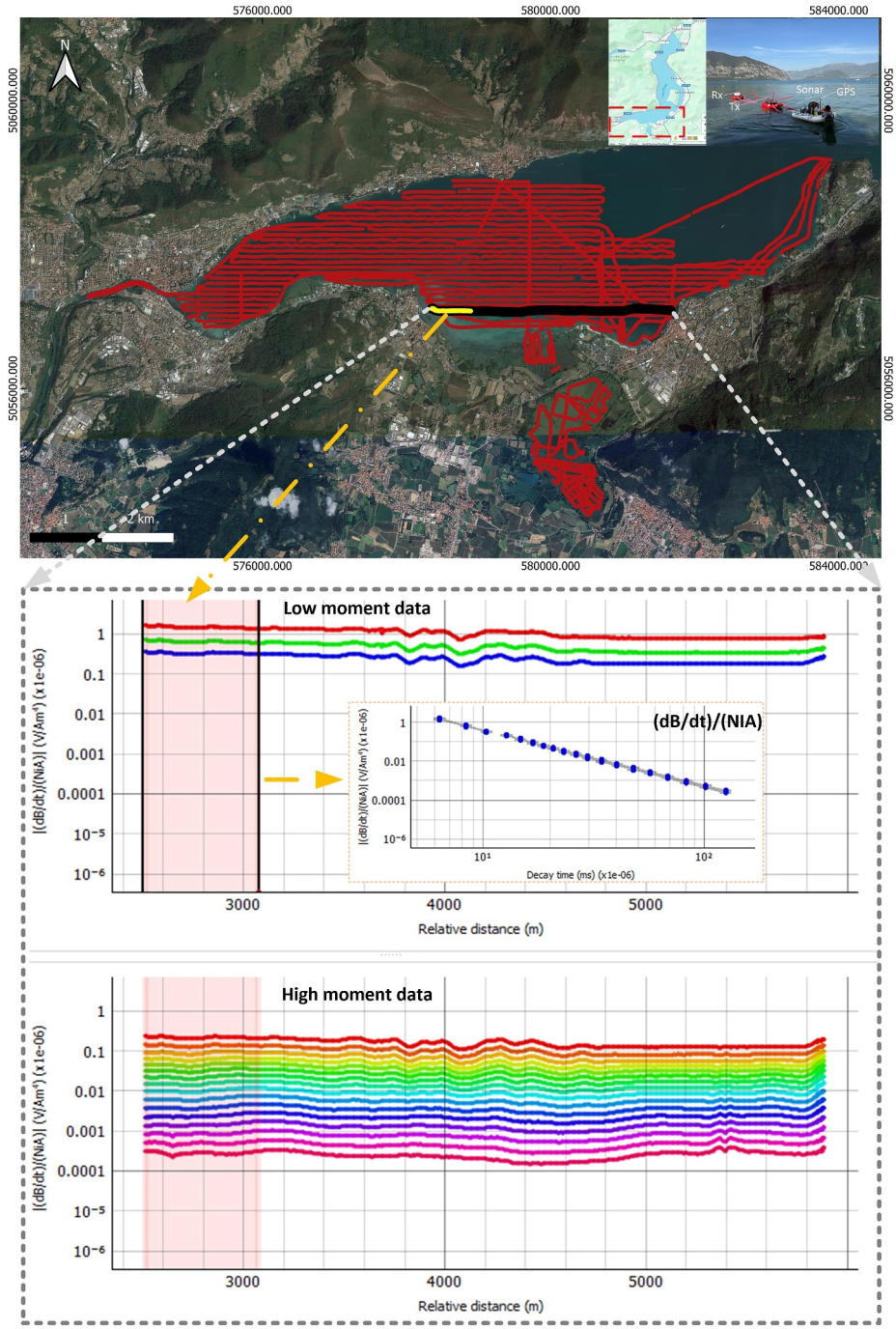

Figure 2. The distribution of waterborne transient electromagnetics survey lines of the survey carried out on the south shore of the lake Iseo, together with an image of the acquisition and a map of the lake.

layer thicknesses were consistent with the training parameters of the deep learning model. The number of inversion layers and layer thickness is consistent with the deep learning training parameters. The MDN-Net, in contrast to the DNN-Net that delivers direct resistivity values, outputs the posterior resistivity PDFs of each layer. In the subsequent comparative analysis of inversion presented this section, we use the maximum probability (MP) value model extracted from the PDF as the representative model for discussion.

The left panel of Figure 3 compares the inversion results of the deterministic inversion, DNN imaging, and MDN MP value imaging. The inversion model reveals some intriguing subsurface characteristics in the lake area, as

10.1029/2024JH000536

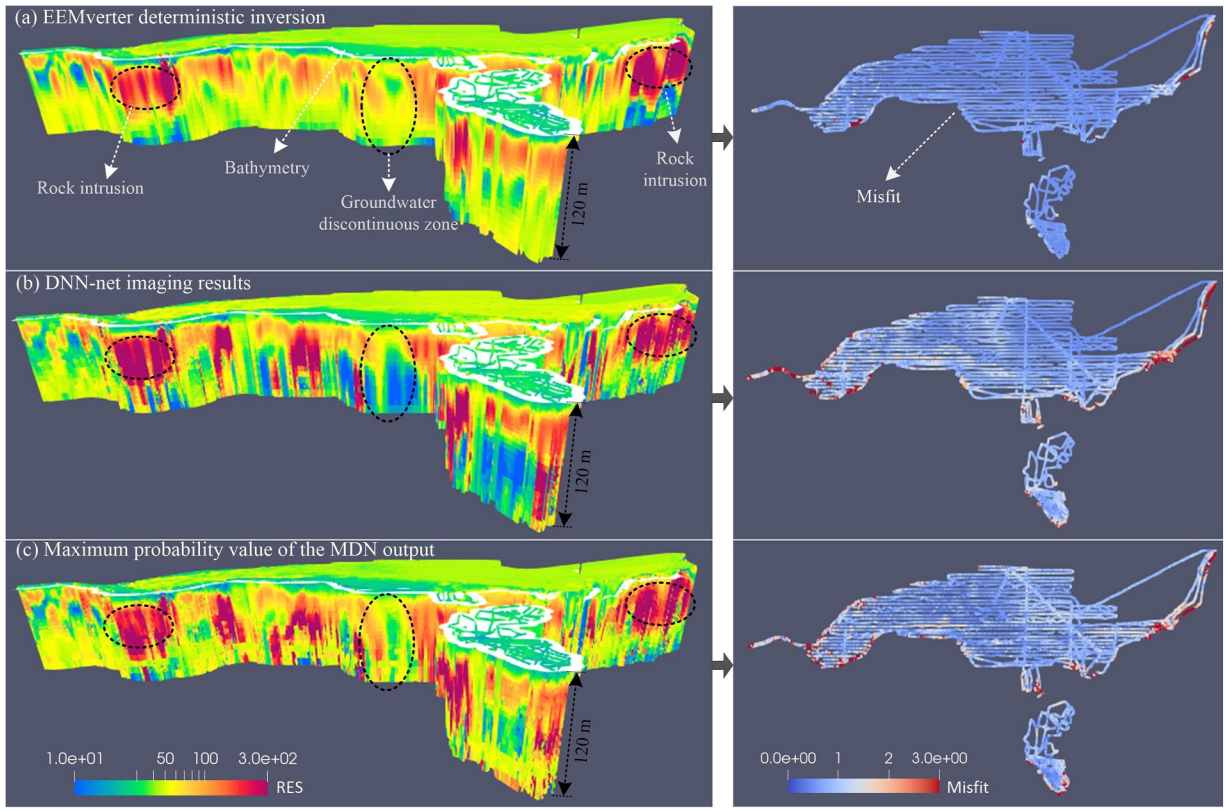

Figure 3. Waterborne transient electromagnetics data inversion and misfit analysis for lake Iseo. From top to bottom: (a) Deterministic inversion results using EEMverter; (b) deep neural network imaging results; (c) Maximum probability value from mixture density network output. The white grid in the figure is the bathymetry information of the lake. For enhanced clarity in visual representation, (a), (b), and (c) are presented with consistent X and Y scales, while the Z scale is expanded tenfold.

delineated by the black dashed lines. First, there are zones of high-resistance anomalies underlying the western and far eastern regions, indicating the presence of rocks extending from the surrounding mountains into the lake. Second, the southern region exhibits anomalies around 100 Ω-meters, representing discontinuous aquifers beneath the lakebed that directly connect with deeper groundwater in some areas. All three methods exhibit good overall consistency in imaging characteristics; however, the deterministic inversion results are generally smoother. This smoothness is attributed to the laterally constrained inversion, which can spread the model information detected at adjacent measurement points, thereby improving the continuity of the model and enhancing the resolution of weakly resolved parameters. Although the trained neural networks can quickly predict model parameters for large-scale data sets, their output results are relatively independent and do not consider the geological correlations between data. Consequently, both single-output DNN and MDN imaging results show significant roughness. Notably, compared with DNN imaging, the MDN imaging results are more closely aligned with the deterministic inversion in some details, such as areas with discontinuous groundwater features.

An important parameter to evaluate the quality of the inversion model is the misfit between the forward response of the inversion model and actual field data. The right panel of Figure 3 shows the misfit calculation results corresponding to the three inversion models. The misfit values for the EEMverter inversion model are generally low, with only some larger misfit values appearing in the regions near the lakeshores. In comparison, the misfit values for the DNN and MDN models are notably higher than those for the EEMverter model, but the misfit values in most areas are less than 1, indicating that both predictive models are reliable. The regions with high misfit values for both the DNN and MDN models exhibit similar distributions, primarily concentrating along the lakeshores of the survey area. As shown in Figure S1 in Supporting Information S1, these areas are close to residential zones, where the data are seriously disturbed by noise.

Figure 4 presents the overall distribution of the subaqueous clay layers extracted from the inversion results. Both the deep learning inversion and EEMverter inversion revealed the existence of two layers of low-resistivity clay

10.1029/2024JH000536

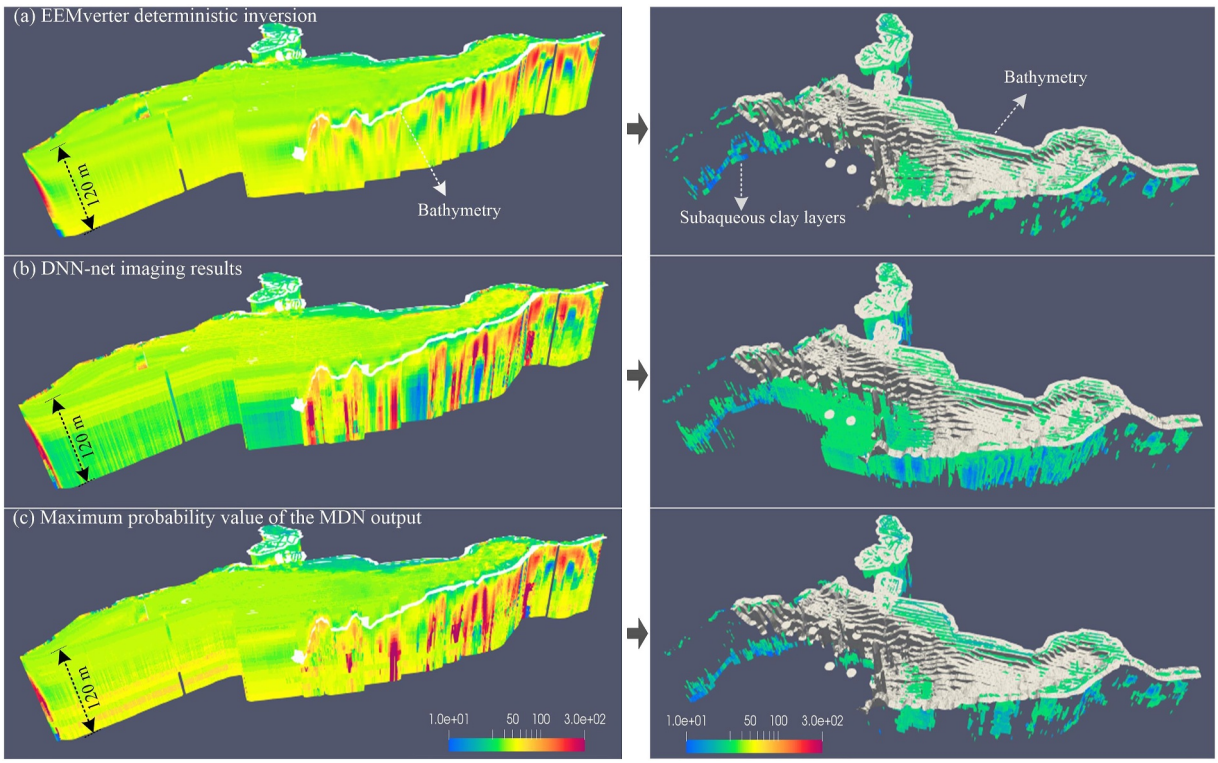

Figure 4. Distribution of the subaqueous clay layers extracted from the inversion results. From top to bottom: (a) Deterministic inversion results using EEMverter; (b) deep neural network imaging results; (c) Maximum probability value from mixture density network output. The gray grid in the figure is the bathymetry information of the lake. For enhanced clarity in visual representation, (a), (b), and (c) are presented with consistent X and Y scales, while the Z scale is expanded tenfold.

underground. The first layer of clay exhibits good continuity, situated between the lake water and aquifer, basically coinciding with the sonar bathymetry. It effectively prevents the lake water from percolating down into the aquifer, thereby protecting the aquifer from direct contamination by surface pollutants. The second clay layer, located at the bottom of the aquifer beneath the lake, is discontinuously distributed. This corroborates the localized connectivity between the sub-lake aquifer and deeper groundwater systems. However, as can be clearly observed from the comparative results on the right side of Figure 4, there are localized discrepancies between the clay layer distribution extracted by the DNN model and the sonar bathymetry results. These discrepancies are manifested in the central region where sonar bathymetry data is missing (due to water depth exceeding the sonar's measurement limit), while the DNN model locally displays unreasonable clay layer representations. Both the EEMverter inversion results and the MP model of MDN demonstrate good agreement with the sonar bathymetry results. This substantiates that the MDN network slightly outperforms the single-output DNN network in handling non-unique geophysical inversion problems.

Additionally, Figures 5a–5c present a comparison between lake depth data extracted from three inversion results and sonar-detected bathymetric data. The extraction strategy details are provided in Text S3 in Supporting Information S1. The lake water depths derived from the inversions are observed to closely approximate the bathymetric values, showing a significant positive correlation. To ensure a fair comparison, the Pearson correlation between the overlapping water depth points extracted by the three types of inversions and the bathymetric data was calculated. This approach mitigated the potential biases associated with varying extraction methodologies. The correlation coefficients for all three sets of extraction outcomes exceeded 0.95, underscoring the reliability of leveraging inversion results for the delineation of lake water depth features. Notably, the correlation results for MDN and EEMverter were remarkably close, achieving a high correlation coefficient of 0.98, outperforming the DNN result of 0.96. Figure 5d presents the credible distribution range of lake water depth derived from the posterior PDF of resistivity output by the MDN. The results demonstrate that the uncertainty in water depth extraction is relatively lower in shallow water regions compared to deeper areas, which is consistent with the law that the resolution of TEM method gradually decreases from shallow to deep. Furthermore, the process of

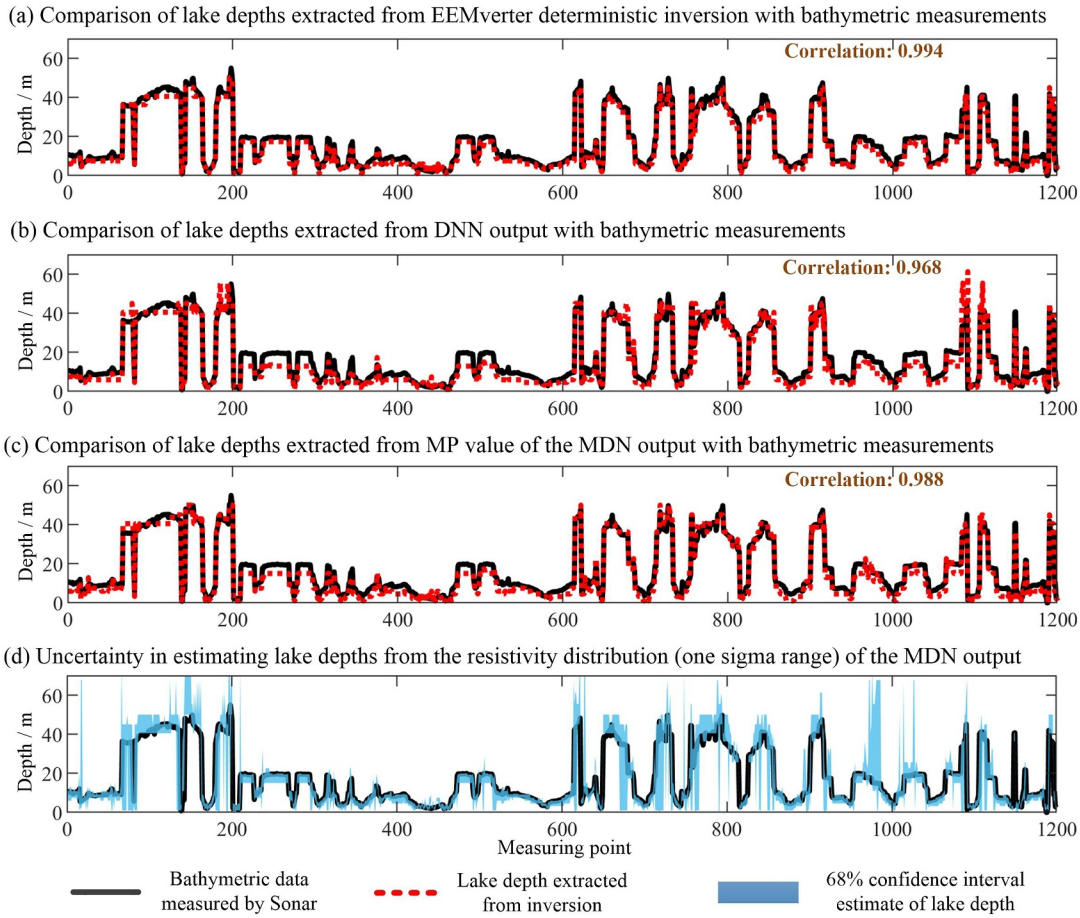

Figure 5. Lake depths extracted from (a) EEMverter, (b) deep neural network, and (c) mixture density network (MDN) imaging models are compared with sonar bathymetry data and Pearson correlation coefficients are calculated. (d) 68% credible interval of lake depth inferred from resistivity distribution output by MDN.

interpreting geophysical observation data to estimate the uncertainty of lake water depth represents a critical transition from geophysical data analysis to practical engineering applications. The ability to rapidly obtain large-scale predictions of lake water depth, along with their associated uncertainties, provides valuable insights for hydrogeological management of lake systems and facilitates timely disaster early warning. The PNN approach can greatly accelerate the decision-making process for environmental monitoring and risk assessment. Despite the inevitable errors in extracted water depths from inversions, these results still demonstrate the good imaging accuracy of the PNN. The PNN takes about 36 s to invert Iseo data in the laptop, while the EEMverter inversion based on the server platform takes approximately 6,500 s. The configuration parameters of the laptop and server are given in Text S2 in the Supporting Information S1.

3.3. Extract Smooth Model

As analyzed in Section 3.2, when neural networks generate large-scale predictive models, the outputs are independent and do not consider the continuity of models corresponding to adjacent geological observations. This results in the models produced by neural networks often being rougher compared to those generated by traditional inversion methods. In this study, the MDN-Net obtained the posterior PDF of the resistivity for each layer, allowing for the estimation of the possible range of layer resistivity through distribution parameters. In this context, we can easily extract thousands of possible models from these distribution functions. Then, the question arises: For lake systems with strong continuity characteristics of geological parameter distribution, is it possible to extract a smooth model that is more in line with geological characteristics, and to quickly obtain models similar to those of laterally constrained inversions (Auken et al., 2008; Viezzoli et al., 2009)? The answer is affirmative. By

10.1029/2024JH000536

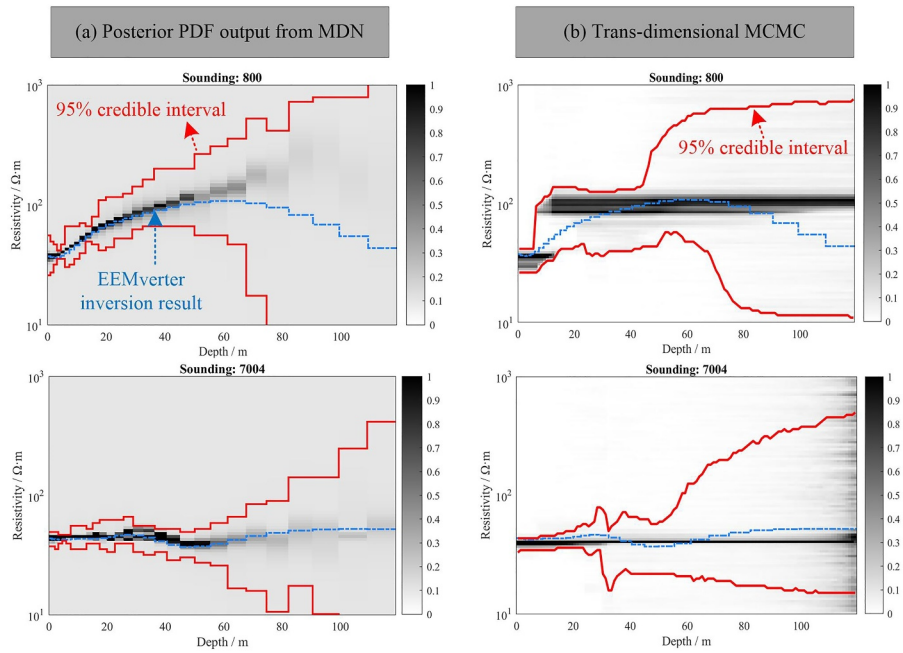

Figure 6. Comparison of posterior model credible intervals between mixture density network (MDN) prediction and Markov chain Monte Carlo (MCMC) inversion at the 800th and 7,004th survey soundings. (a) MDN results; (b) Trans-dimensional MCMC results.

searching for models within the potential distribution range of resistivities, and using an optimization algorithm that identifies models with similar adjacent resistivities, we can readily obtain smooth models. This strategy offers a high degree of flexibility, and we can even design sharp constraints (Vignoli et al., 2015) to search for models that are smooth locally (Chen et al., 2023). In this respect, it is paramount to verify the reliability of the resistivity distribution functions obtained by the PNN and to ensure the credibility of any model extracted therein. In a prior research work we were involved in, the reliability of the resistivity distribution function under the mixture Gaussian model has been confirmed through simulations and field cases (Yu et al., 2024). In this study, we first compared the posterior 95% credible intervals of resistivity predicted by the MDN with the credible intervals obtained by classical MCMC sampling. We adopted the trans-dimensional MCMC inversion strategy proposed by Blatter et al. (2018), using three Markov chains, each sampling 200,000 times, with a burn-in period of 5,000 samples. We used the same laptop computer and the MCMC inversion took about 6 hr per sounding.

Figure 6 shows the comparison results of two random survey points in lake Iseo. The red lines in the figure show the 95% credible interval distribution while the green lines indicate the deterministic inversion results from EEMverter. The comparative results demonstrate that the 95% credible intervals of the posterior models obtained from both the MDN and MCMC methods exhibit consistency across two sounding points (Soundings 800 and 7004). Observations from the results highlight that in shallower depths, the 95% credible interval tends to be narrower for both the MDN and MCMC approaches, indicating higher model certitude and lesser uncertainty. However, with the increase in depth, the uncertainty manifests itself by demonstrating the broadened credible interval widths, suggesting greater uncertainty and complexity in estimating the resistivity in deeper layers. Furthermore, the MDN method exhibits a more smooth output for the posterior PDF than the MCMC method when deriving the posterior distribution. This smoother output could be attributed to the MDN's capability of incorporating and learning the prior knowledge of smooth variations in resistivity from the training data set, resulting in a posterior PDF with more gradual and smooth changes in resistivity high probability density regions with depth progression.

We extracted 100 sets of observational data from lake Iseo, calculated the mean and STD from the mixture of Gaussian functions using the law of total variance, randomly sampled 50 models within the one-sigma range of the distribution function, and computed the average misfit between the forward modeling data and the observational data for all models. Figure 7 presents a comparative analysis of the average misfit between the MDN

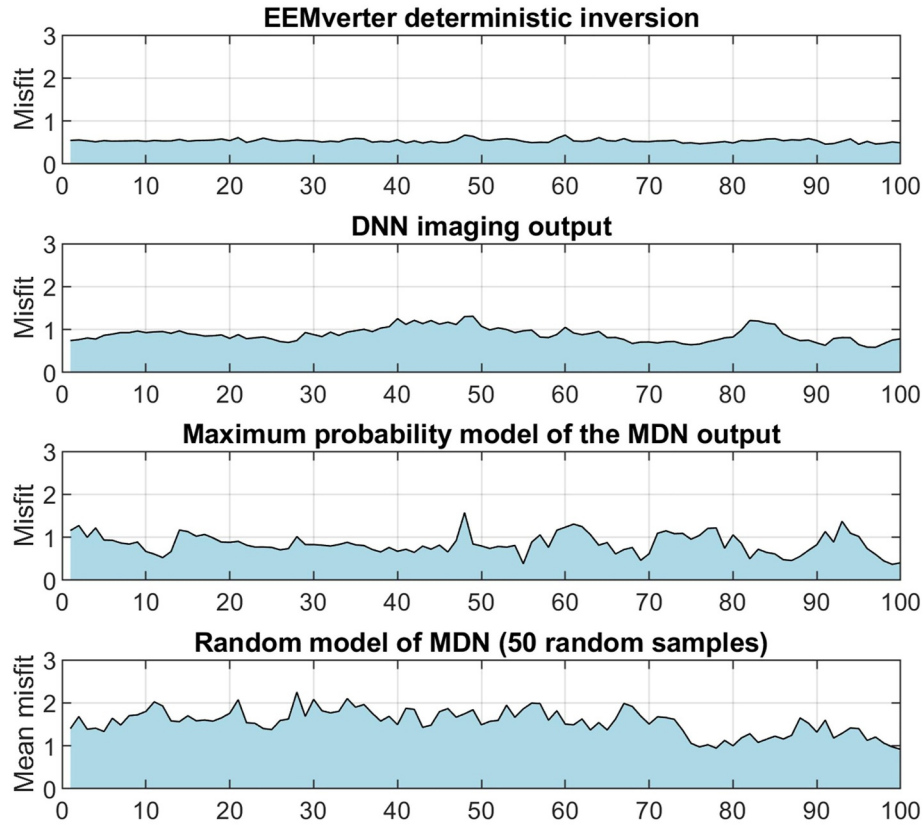

Figure 7. Evaluation of the reliability of resistivity probability density function using 100 random data sets from lake Iseo.

random models and the inversion misfits of EEMverter, DNN, and MDN-MP models. Notably, the average misfit of the random sampling model is significantly inferior to that of the MDN-MP solution, with the majority of errors ranging between 1.5 and 2. Given that this model is randomly sampled from a 30-layer resistivity distribution function, such errors are reasonable due to the inherent roughness and randomness between models. In scenarios where additional geological prior knowledge is unavailable, mobile electromagnetic detection systems with large-scale, high-density acquisition often exhibit smooth geological variations in adjacent data. Consequently, identifying smooth models from the posterior distribution of MDN may enhance the overall imaging resolution. Figures 8a and 8b illustrate the results of smooth models extracted from the MDN posterior PDFs generated from two sets of lake Iseo survey line data. Compared to the MDN-MP and DNN imaging models, the MDN smooth models significantly enhance the smoothness of the imaging results, producing models that closely resemble those obtained from EEMverter deterministic laterally constrained inversion. This makes imaging models more representative in hydrogeological environments with good continuity. Furthermore, Figures 8c and 8d present the misfit between the forward responses of different models and the observed data. It is evident that the smooth model significantly enhances the quality of data fitting compared to the MDN-MP model, with the average misfit for the two survey lines decreasing by 24.8% and 22.2%, respectively. This demonstrates that extracting models from the posterior PDF that align with prior geological knowledge not only improves model resolution but also enhances the model reliability. In fact, you can also design a special regularization function to extract a model that meets the model constraints from the posterior resistivity distribution. These results provide a practical approach for leveraging deep learning techniques to rapidly achieve constrained geophysical inversion.

3.4. Estimate of Depth of Investigation

The DOI is a critical parameter for evaluating the reliability of inversion models at different depth ranges. Building upon the global DOI concept defined by Christiansen and Auken (2012) and the method of estimating the DOI of inversion models using Kullback–Leibler (KL) divergence proposed by Blatter et al. (2018),we provide a deep learning-based DOI estimation strategy leveraging resistivity distribution functions obtained from

10.1029/2024JH000536

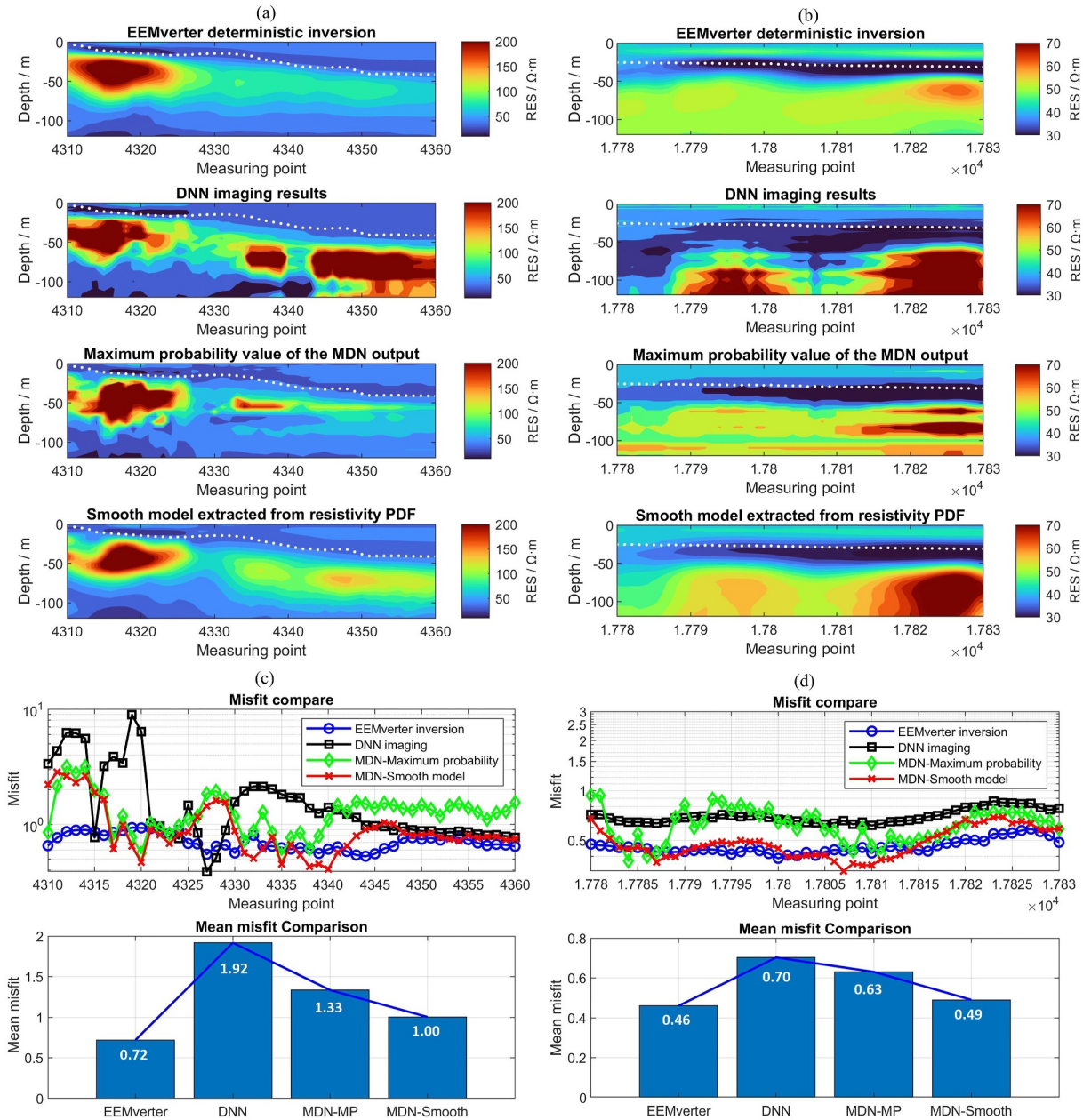

Figure 8. Comparison of the smooth model extracted by MDN-Net with the model obtained by EEMverter inversion and the MDN-MP model. Left column from top to bottom: (a) Inversion models corresponding to the 4,310–4,360th survey point data of lake Iseo and their corresponding (c) misfit comparison. Right column from top to bottom: (b) Inversion models corresponding to the 17,780–17,830th survey point data of lake Iseo and their corresponding (d) misfit comparison. The white dots in the model diagram represent sonar bathymetric data.

the MDN. Essentially, our strategy involves computing the global cumulative sensitivity by analyzing the KL divergence information of the resistivity posterior PDF. The KL divergence quantifies the difference between two probability distributions, where a higher divergence value indicates that the posterior distribution has extracted substantial information from the data, signifying higher sensitivity in identifying the resistivity of that layer. Consequently, this allows us to compute the cumulative sensitivity from deeper to shallower regions. We then establish a sensitivity threshold, which represents the minimum sensitivity required to produce a meaningful impact on the observed data. When the calculated cumulative sensitivity exceeds this threshold, the corresponding depth is designated as the DOI for the deep learning-based inversion. For details, see Text S4 in Supporting Information S1.

10.1029/2024JH000536

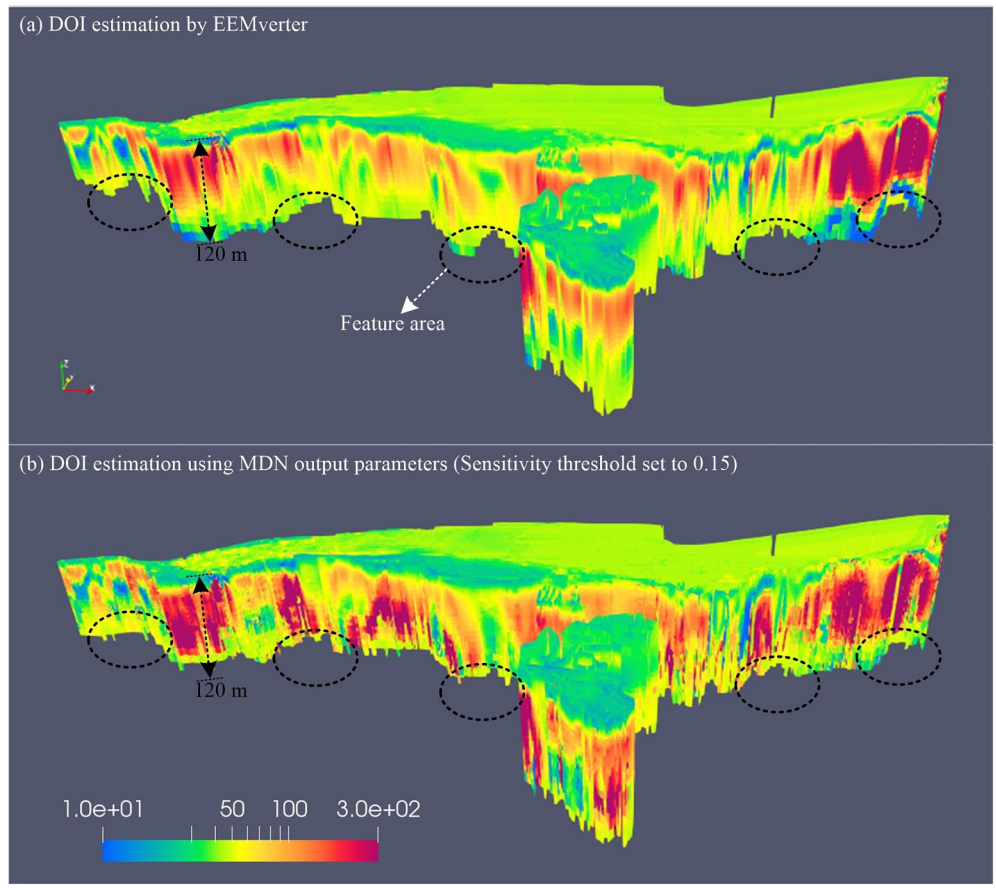

Figure 9. Estimation of depth of investigation (DOI). (a) DOI estimation from EEMverter inversion, (b) DOI estimation from probabilistic neural network using cumulative Kullback–Leibler divergence.

Figure 9 illustrates the comparative analysis of the DOI region as determined by EEMverter inversion and the DOI region identified by the PNN network. This comparison reveals that, under suitable threshold conditions, there is a good agreement between the DOI region delineated by the PNN and that identified through classical inversion techniques. Moreover, conventional DOI estimation strategies are directly related to model parameters. Affected by the non-uniqueness of inversion, different models that can satisfy data fitting may cause fluctuations in DOI (Blatter et al., 2018), while cumulative KL divergence uses the resistivity posterior PDF for estimation, and its estimation results are applicable to any model extracted from this distribution. The described method provides a feasible strategy for DOI estimation based on geophysical data inversion that utilizes deep learning techniques.

4. Discussions

In this study, the training data set construction framework allows us to add field data, but whether adding field data will improve the performance of the MDN network remains to be evaluated. To this end, we designed a comparative experiment to assess the network's performance under three different data sets. Dataset-*a* consists exclusively of synthetic data, Dataset-*b* is the same as the training set used in the previous section of this paper, with 15,000 sets of field data added (excluding the data from lake Iseo), and Dataset-*c* further incorporates 3,000 sets of field data from lake Iseo (approximately 8.57% of the total data from lake Iseo). The training parameters for these three data sets are identical, specifics are provided in Table S1 of Supporting Information S1.

Figure 10 shows the comparison of the misfit of the MDN-MP model output on the Iseo data and the posterior PDF of the two soundings (same sounds in Figure 6) under three training data sets. The posterior PDF features of the two soundings exhibit similarity across the three training data sets, while the output misfit distributions shows

10.1029/2024JH000536

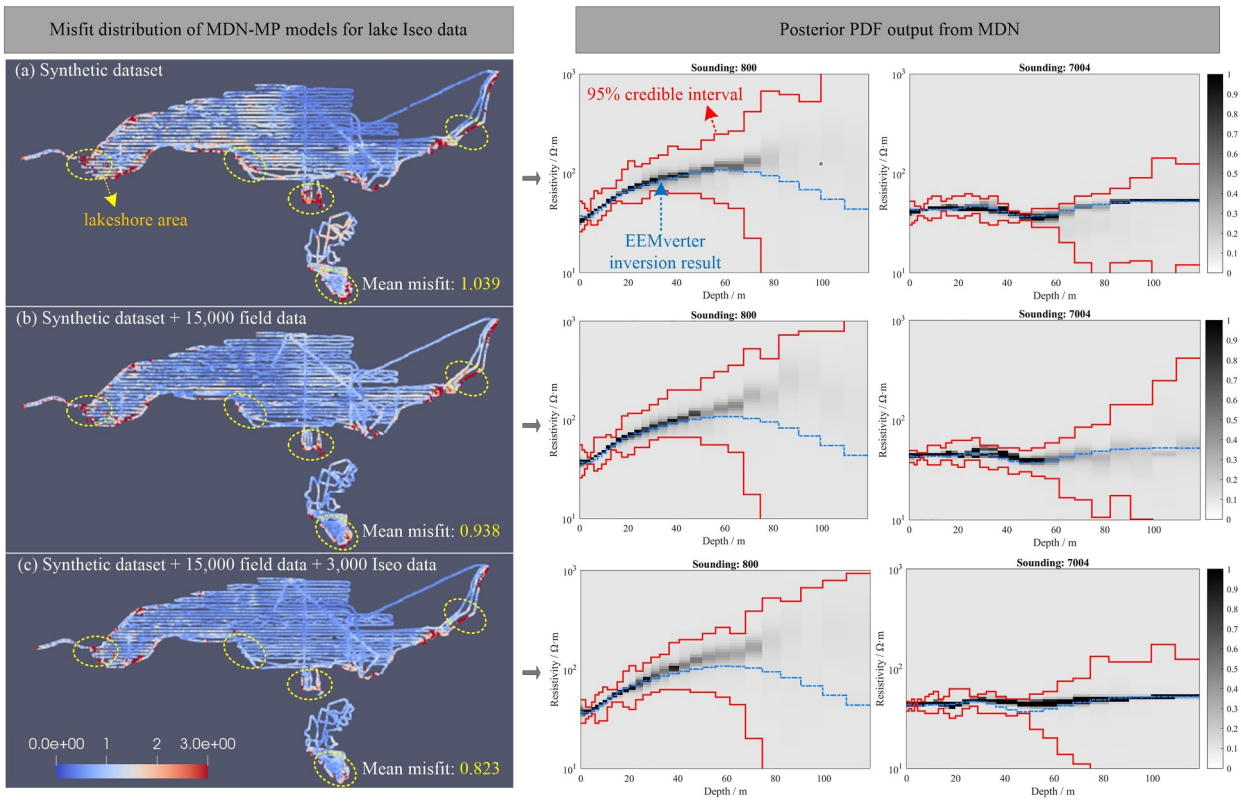

Figure 10. Impact of training data sets with different compositions on the performance of the mixture density network. (a) Training data set composed solely of synthetic data, (b) training data set composed of the data in set (a) with the addition of 15,000 sets of field data, (c) training data set composed of the data in set (b) with the addition of 3,000 sets of field data collected at lake Iseo. Consistent with the way of constructing label models for synthetic data sets, the models corresponding to the field data are all obtained by deterministic inversion.

obvious variability. Given that misfit is a critical parameter for evaluating the reliability of the MDN network prediction model, we computed the mean misfit across all soundings for each model to provide a quantitative basis for comparison. In Figure 10, the yellow dashed lines mark several areas along the lake shore where the data are relatively more disturbed by noise. Compared with the purely synthetic data set, the model predicted by the training set with field data added performs better in these areas, which may be related to the similarities between the noise model contained in the added field data and the noise pattern on the lake shore. Overall, the average misfit decreases from 1.039 to 0.938 after adding field data, representing an improvement in data fitting quality by approximately 10.77%. Figure 10c shows the network performance after adding some measured data from lake Iseo. Since this method is equivalent to adding prior knowledge to the training set, it can significantly improve the reliability of the MDN network prediction model. The overall misfit is further decreased from 0.938 to 0.823, which means that the data fitting quality improved by 12.26%. These results suggest that incorporating filed data into the training set can enhance the training quality of PNN, particularly in improving the reliability of prediction models.

Furthermore, in Figure S2 of the Supporting Information S1, we discuss the impact of the number of Gaussian kernels on the performance of the MDN. We compare the testing results of the MDN on the Iseo data set for Gaussian kernel counts of 1, 3, 5, and 7. The comparative analysis reveals that while the MDN with a single Gaussian kernel (i.e., a simple Gaussian distribution) also yields a relatively favorable average misfit value, the utilization of a Gaussian mixture distribution allows for a more comprehensive description of the multimodal characteristics inherent in the data. Within a certain range, an appropriate increase in the number of Gaussian kernels enhances the reliability of the predictive model generated by the MDN. In our study case, when the number of Gaussian kernels increased from 3 to 5, the testing performance of the MDN network for the prediction model can be further improved, with the misfit decreasing from 0.938 to 0.849. However, as the number of Gaussian kernels continues to increase, the misfit of the MDN network prediction model begins to rise, possibly

due to the increased complexity of the neural network from the excessive number of Gaussian kernels, leading to overfitting. The test results under different data sets and different numbers of Gaussian kernels show that although changing the network parameters and training sets will affect the accuracy of the network prediction model, the posterior PDF distribution characteristics of their predictions show relatively good robustness. This is important, as it suggests that the MDN network has learned basic rules for evaluating model uncertainty from the training data set. Consequently, we do not need to overly pursue perfect network training performance, because the strategy proposed in this study for finding smooth models from the posterior distribution allows us to identify models that are more consistent with geological knowledge from an imperfect prediction model, by setting appropriate constraints. In this process, the correct information from geological knowledge is diffused in the found model through the constraints, thereby improving the roughness and reliability of the model.

5. Conclusions

In this study, we proposed a PNN structure that can directly convert large-scale observational TEM data into a resistivity model and estimate its posterior resistivity distribution. Through testing on the large-scale FloaTEM data of lake Iseo in Italy, we discuss the feasibility of extracting a smooth model from the posterior PDF predicted by the PNN and estimating the DOI of the prediction model by calculating the cumulative KL divergence. Thanks to the unique design of the inversion training data set, the PNN framework allows us to incorporate field data to enrich the geological model library, thereby enhancing the quality of network training. In the discussion section, we compare the network performance under different data sets. Adding field data to the training data set can improve the reliability of the network's model prediction for data with strong noise interference. However, there are still some limitations: on the one hand, the length of the input data of the current PNN structure is fixed, which means that we still need to preprocess the field data to match the length of the training data set. On the other hand, we have not yet considered the induced polarization phenomenon in the TEM data (Maurya, Grombacher, et al., 2022) and the modeling error caused by 1D approximate (Bai et al., 2021; Deleersnyder et al., 2024). These imitations will be the focus of future research efforts. Overall, compared with the single-output DNN network, the PNN can rapidly provide a comprehensive inversion assessment for large-scale data, aiding geophysical engineers in gaining a more thorough understanding of subterranean geological systems and in making informed decisions.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The FloaTEM data, the bathymetry data compared with inversion depth file data, the PNN code, and the inversion results of this study are available at https://doi.org/10.5281/zenodo.13952357 (Chen et al., 2024).

Acknowledgments
This study has been carried out within the Horizon Europe project SEMACRET (Grant Agreement no. 101057741). It was also funded under the 'Progetto di Eccellenza 2023–2027: Le Georisorse per la Transizione Ecologica e lo Sviluppo Territoriale' by the Italian Ministry of University and Research. The FloaTEM data have been acquired in the LakEMaging project, funded by Acque Bresciane S.r.l. SB. Open access publishing facilitated by Universita degli Studi di Milano, as part of the Wiley - CRUI-CARE agreement.

References

Alyaev, S., & Elsheikh, A. H. (2022). Direct multi-modal inversion of geophysical logs using deep learning. *Earth and Space Science*, *9*, e2021EA002186. https://doi.org/10.1029/2021EA002186

Ardizzone, L., Kruse, J., Wirkert, S., Rahner, D., Pellegrini, E. W., Klessen, R. S., et al. (2018). Analyzing inverse problems with invertible neural networks. *arXiv*. https://arxiv.org/abs/1808.04730

Asif, M. R., Bording, T. S., Maurya, P. K., Zhang, B., Fiandaca, G., Grombacher, D. J., et al. (2022). A neural network-based hybrid framework for least-squares inversion of transient electromagnetic data. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 4503610. https://doi.org/10.1109/TGRS.2021.3076121

Auken, E., Christiansen, A. V., Jacobsen, L. H., & Sørensen, K. I. (2008). A resolution study of buried valleys using laterally constrained inversion of TEM data. *Journal of Applied Geophysics*, *65*(1), 10–20. https://doi.org/10.1016/j.jappgeo.2008.03.003

Auken, E., Foged, N., Larsen, J. J., Lassen, K. V. T., Maurya, P. K., Dath, S. M., & Eiskjær, T. T. (2019). tTEM—A towed transient electromagnetic system for detailed 3D imaging of the top 70 m of the subsurface. *Geophysics*, *84*(1), E13–E22. https://doi.org/10.1190/GEO2018-0355.1

Bai, P., Vignoli, G., & Hansen, T. M. (2021). 1D stochastic inversion of airborne time-domain electromagnetic data with realistic prior and accounting for the forward modeling error. *Remote Sensing*, *13*(19), 3881. https://doi.org/10.3390/rs13193881

Blatter, D., Key, K., Ray, A., Foley, N., Tulaczyk, S., & Auken, E. (2018). Trans-dimensional Bayesian inversion of airborne transient EM data from Taylor Glacier, Antarctica. *Geophysical Journal International*, *214*(3), 1919–1936. https://doi.org/10.1093/gji/ggy255

Chandra, S., Tiwari, V. M., Vidyasagar, M., Raju, K. B., Choudhury, J., Lohithkumar, K., et al. (2021). Airborne electromagnetic signatures of an ancient river in the water-stressed Ganga Plain, Prayagraj, India: A potential groundwater repository. *Geophysical Research Letters*, *48*(23), e2021GL096100. https://doi.org/10.1029/2021GL096100

Chen, J., Galli, S., Signora, A., Sullivan, N. A. L., Zhang, B., & Fiandaca, G. (2024). Instantaneous Bayesian imaging for TEM [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.13952357

Chen, J., Jia, W., Zhang, Y., & Lin, J. (2021). Integrated TEM and GPR data interpretation for high-resolution measurement of urban underground space. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–9. https://doi.org/10.1109/TIM.2021.3134995

Chen, J., Zhang, Y., & Lin, J. (2022a). Fast transdimensional Bayesian transient electromagnetic imaging for urban underground space detection. *Measurement*, *187*, 110300. https://doi.org/10.1016/j.measurement.2021.110300

Chen, J., Zhang, Y., & Lin, T. (2023). High-resolution quasi-three-dimensional transient electromagnetic imaging method for urban underground space detection. *IEEE Transactions on Industrial Informatics*, *19*(3), 3039–3046. https://doi.org/10.1109/TII.2022.3176890

Chen, J., Zhang, Y., & Lin, T. T. (2022b). Transient electromagnetic machine learning inversion based on pseudo wave field data. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–10. https://doi.org/10.1109/TGRS.2022.3187021

Christiansen, A. V., & Auken, E. (2012). A global measure for depth of investigation. *Geophysics*, *77*(4), 171–177. https://doi.org/10.1190/GEO2011-0393.1

Colombo, D., Turkoglu, E., Li, W., Curiel, E. S., & Rovetta, D. (2021). Physics-driven deep-learning inversion with application to transient electromagnetics. *Geophysics*, *86*(3), E209–E224. https://doi.org/10.1190/GEO2020-0760.1

Deleersnyder, W., Dudal, D., & Hermans, T. (2024). A multidimensional AI-trained correction to the 1D approximate model for Airborne TDEM sensing. *Computers & Geosciences*, *188*, 105602. https://doi.org/10.1016/j.cageo.2024.105602

Devilee, R., Curtis, A., & Roy-Chowdhury, K. (1999). An efficient, probabilistic neural network approach to solving inverse problems: Inverting surface wave velocities for Eurasian crustal thickness. *Journal of Geophysical Research*, *104*(B12), 28841–28857. https://doi.org/10.1029/1999jb900273

Earp, S., & Curtis, A. (2020). Probabilistic neural network-based 2D travel-time tomography. *Neural Computing & Applications*, *32*(22), 17077–17095. https://doi.org/10.1007/s00521-020-04921-8

Earp, S., Curtis, A., Zhang, X., & Hansteen, F. (2020). Probabilistic neural network tomography across Grane field (North Sea) from surface wave dispersion data. *Geophysical Journal International*, *223*(3), 1741–1757. https://doi.org/10.1093/gji/ggaa328

Fiandaca, G., Chen, J., & Zhang, B. (2024). Closing the gap between galvanic and inductive methods: EEMverter, a new inversion tool for electric and electromagnetic data with focus on induced polarization. In *Australian society of exploration geophysicists extended abstracts, volume 2024, 1st ASEG DISCOVER symposium, Hobart*. https://doi.org/10.5281/zenodo.13883805

Fiandaca, G., Ramm, J., Binley, A., Gazoty, A., Christiansen, A. V., & Auken, E. (2013). Resolving spectral information from time domain induced polarization data through 2-D inversion. *Geophysical Journal International*, *192*(2), 631–646. https://doi.org/10.1093/gji/ggs060

Finn, C. A., Bedrosian, P. A., Holbrook, W. S., Auken, E., Bloss, B. R., & Crosbie, J. (2022). Geophysical imaging of the Yellowstone hydrothermal plumbing system. *Nature*, *603*(7902), 643–647. https://doi.org/10.1038/s41586-021-04379-1

Grombacher, D., Maurya, P. K., Lind, J. C., Lane, J., & Auken, E. (2021). Rapid mapping of hydrological systems in Tanzania using a towed transient electromagnetic system. *Groundwater Series*, *60*, 1–9. https://doi.org/10.1111/gwat.13130

Hansen, T. M. (2021). Efficient probabilistic inversion using the rejection sampler—Exemplified on airborne EM data. *Geophysical Journal International*, *224*(1), 543–557. https://doi.org/10.1093/gji/ggaa491

Hansen, T. M., & Finlay, C. C. (2022). Use of machine learning to estimate statistics of the posterior distribution in probabilistic inverse problems—An application to airborne EM data. *Journal of Geophysical Research: Solid Earth*, *127*(11), e2022JB024703. https://doi.org/10.1029/2022JB024703

Hansen, T. M., & Minsley, B. J. (2019). Inversion of airborne EM data with an explicit choice of prior model. *Geophysical Journal International*, *218*(2), 1348–1366. https://doi.org/10.1093/gji/ggz230

Jiang, Y., Ma, J., Ning, J., Li, J., Wu, H., & Bao, T. (2025). One-fit-all transformer for multimodal geophysical inversion: Method and application. *Journal of Geophysical Research: Machine Learning and Computation*, *2*, e2024JH000432. https://doi.org/10.1029/2024JH00043

Killingbeck, S. F., Booth, A. D., Livermore, P. W., Bates, C. R., & West, L. J. (2020). Characterisation of subglacial water using a constrained transdimensional Bayesian transient electromagnetic inversion. *Solid Earth*, *11*(1), 75–94. https://doi.org/10.5194/se-11-75-2020

Li, P., Grana, D., & Liu, M. (2024). Bayesian neural network and Bayesian physics-informed neural network via variational inference for seismic petrophysical inversion. *Geophysics*, *89*(6), M185–M196. https://doi.org/10.1190/GEO2023-0737.1

Li, R., Wu, X., Tian, H., Yu, N., & Wang, C. (2022). Hybrid memetic pretrained factor analysis-based deep belief networks for transient electromagnetic inversion. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–20. https://doi.org/10.1109/TGRS.2022.3208465

Maurya, P. K., Christensen, F. E., Kass, M. A., Pedersen, J. B., Frederiksen, R. R., Foged, N., et al. (2022). Technical note: Efficient imaging of hydrological units below lakes and fjords with a floating, transient electromagnetic (FloaTEM) system. *Hydrology and Earth System Sciences*, *26*(11), 2813–2827. https://doi.org/10.5194/hess-26-2813-2022

Maurya, P. K., Foged, N., Madsen, L. M., & Christiansen, A. V. (2023). Comparison of towed electromagnetic with airborne electromagnetic and electrical resistivity tomography in a hydrogeophysical context. *Geophysical Journal International*, *235*(1), 817–830. https://doi.org/10.1093/gji/ggad276

Maurya, P. K., Grombacher, D., Lind, J., Lane, J. W., & Auken, E. (2022). Inversion of induced polarization-affected towed-transient electromagnetic data in a lateritic regolith geology: A case study from western Tanzania. *Geophysics*, *87*(4), B247–B254. https://doi.org/10.1190/GEO2021-0396.1

Mosher, S. G., Audet, P., & Gosselin, J. M. (2021b). Shear-wave velocity structure of sediments on Cascadia's continental margin from probabilistic inversion of seafloor compliance data. *Geochemistry, Geophysics, Geosystems*, *22*(9), e2021GC009720. https://doi.org/10.1029/2021GC009720

Mosher, S. G., Eilon, Z., Janiszewski, H., & Audet, P. (2021a). Probabilistic inversion of seafloor compliance for oceanic crustal shear velocity structure using mixture density neural networks. *Geophysical Journal International*, *227*(3), 1879–1892. https://doi.org/10.1093/gji/ggab315

Munkholm, M. S., & Auken, E. (1996). Electromagnetic noise contamination on transient electromagnetic soundings in culturally disturbed environments. *Journal of Environmental & Engineering Geophysics*, *1*(2), 119–127. https://doi.org/10.4133/JEEG1.2.119

Neven, A., Maurya, P. K., Christiansen, A. V., & Renard, P. (2021). tTEM20AAR: A benchmark geophysical data set for unconsolidated fluvioglacial sediments. *Earth System Science Data*, *13*(6), 2743–2752. https://doi.org/10.5194/essd-13-2743-2021

Puzyrev, V. (2019). Deep learning electromagnetic inversion with convolutional neural networks. *Geophysical Journal International*, *218*(2), 817–832. https://doi.org/10.1093/gji/ggz204

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

Sandersen, P. B. E., Kallesøe, A. J., Mølle, I., Høyer, S., Jørgensen, F., Pedersen, J. B., & Christiansen, A. V. (2021). Utilizing the towed transient Electromagnetic method (tTEM) for achieving unprecedented near-surface detail in geological mapping. *Engineering Geology*, *288*, 106125. https://doi.org/10.1016/j.enggeo.2021.106125

Shi, M., & Cao, H. (2022). An ATEM 1D inversion based on K-Means clustering and MLP deep learning. *Journal of Geophysics and Engineering*, *19*(4), 775–787. https://doi.org/10.1093/jge/gxac050

Silvestri, S., Christensen, C. W., Lysdahl, A. O. K., Anschütz, H., Pfaffhuber, A. A., & Viezzoli, A. (2019). Peatland volume mapping over resistive substrates with airborne electromagnetic technology. *Geophysical Research Letters*, *46*(12), 6459–6468. https://doi.org/10.1029/2019GL083025

Viezzoli, A., Auken, E., & Munday, T. (2009). Spatially constrained inversion for quasi 3D modelling of airborne electromagnetic data – An application for environmental assessment in the Lower Murray Region of South Australia. *Exploration Geophysics*, *40*(2), 173–183. https://doi.org/10.1071/EG08027

Vignoli, G., Fiandaca, G., Christiansen, A. V., Kirkegaard, C., & Auken, E. (2015). Sharp spatially constrained inversion with applications to transient electromagnetic data. *Geophysical Prospecting*, *63*(1), 243–255. https://doi.org/10.1111/1365-2478.12185

Wu, G., Dong, H., Song, J., & Zhang, J. (2021e). Bayesian geoacoustic inversion using mixture density network. *arXiv*. https://arxiv.org/abs/2008.07902

Wu, S. H., Huang, Q., & Zhao, L. (2021a). Instantaneous inversion of airborne electromagnetic data based on deep learning. *Geophysical Research Letters*, *49*(10), e2021GL097165. https://doi.org/10.1029/2021GL097165

Wu, S. H., Huang, Q., & Zhao, L. (2021c). De-noising of transient electromagnetic data based on the long short-term memory-autoencoder. *Geophysical Journal International*, *224*(1), 669–681. https://doi.org/10.1093/gji/ggaa424

Wu, S. H., Huang, Q., & Zhao, L. (2021d). Convolutional neural network inversion of airborne transient electromagnetic data. *Geophysical Prospecting*, *69*(8–9), 1–12. https://doi.org/10.1111/1365-2478.13136

Wu, S. H., Huang, Q., & Zhao, L. (2023). Fast Bayesian inversion of airborne electromagnetic data based on the invertible neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–11. https://doi.org/10.1109/TGRS.2023.3264777

Wu, X., Xue, Q., Zhao, Y., Lv, P., Zhou, Z., & Shi, J. (2021b). A deep learning estimation of the earth resistivity model for the airborne transient electromagnetic observation. *Journal of Geophysical Research: Solid Earth*, *127*(3), e2021JB023185. https://doi.org/10.1029/2021JB023185

Xue, G., Li, H., He, Y., Xue, J., & Wu, X. (2020). Development of the inversion method for transient electromagnetic data. *IEEE Access*, *8*, 146172–146181. https://doi.org/10.1109/ACCESS.2020.3013626

Yu, S. B., Shen, Y. H., Meng, F. Z., Chen, J., & Zhang, Y. (2024). An efficient transient electromagnetic uncertainty 1D inversion method based on mixture density network. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–13. https://doi.org/10.1109/TGRS.2024.3357650

Zaru, N., Silvestri, S., Assiri, M., Bai, P., Hansen, T. M., & Vignoli, G. (2024). Probabilistic petrophysical reconstruction of Danta's Alpine peatland via electromagnetic induction data. *Earth and Space Science*, *11*, e2023EA003457. https://doi.org/10.1029/2023EA00345

Zhang, X., & Curtis, A. (2021). Bayesian geophysical inversion using invertible neural networks. *Journal of Geophysical Research: Solid Earth*, *126*(7), e2021JB022320. https://doi.org/10.1029/2021JB022320

References From the Supporting Information

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *arXiv*. https://arxiv.org/abs/1412.6980

Prechelt, L. (1998). Early stopping—But when? In *Neural networks: Tricks of the trade* (pp. 55–69). https://doi.org/10.1007/3-540-49430-8_3