APPLICATION OF THE FINITE ELEMENT METHOD TO GENERATE TRAINING DATA FOR AI-BASED APPARENT RESISTIVITY PREDICTION IN TWO-LAYER SOIL MODELS

ỨNG DỤNG PHƯƠNG PHÁP PHẦN TỬ HỮU HẠN ĐỂ TẠO DỮ LIỆU HUẤN LUYỆN CHO MÔ HÌNH AI DỰ ĐOÁN ĐIỆN TRỞ SUẤT BIỂU KIẾN CỦA LỚP ĐẤT PHÂN TẦNG

Nguyen Xuan Phuc[1],* Nguyen Thi Thuy Duong[1], Ninh Van Nam[2] and Tran Thi Ngoat[1]
[1]School of Electrical & Electronic Engineering, Hung Yen University of Technology and Education
[2]School of Electrical & Electronic Engineering, Hanoi University of Industry
Received: July 02, 2025 Accepted: April 20, 2026

Abstract:

This study presents a novel approach for predicting apparent resistivity in two-layered soil systems using artificial intelligence (AI). A dataset consisting of 7,600 simulated apparent resistivity values was generated by applying the Finite Element Method (FEM) to model subsurface resistivity in a two-layered soil configuration. The dataset incorporates various combinations of soil resistivity values (ρ1, ρ2) and layer thicknesses (l1, l2), enabling the creation of a diverse and representative training set. A machine learning model based on a Multi-Layer Perceptron (MLP), a type of Artificial Neural Network (ANN), was trained on this dataset to predict the apparent resistivity from the input parameters. The results demonstrate that the proposed MLP model can predict apparent resistivity with high accuracy, thereby providing a more computationally efficient alternative to traditional field-based methods. This work highlights the potential of FEM and AI in geophysical applications, providing a robust framework for predicting subsurface resistivity in layered soil systems.

Key words: Apparent resistivity; multilayer soil; FEM; ANN; MLP.

Tóm tắt:

Nghiên cứu này đề xuất một phương pháp mới để dự đoán điện trở suất biểu kiến của hệ đất hai lớp bằng cách ứng dụng trí tuệ nhân tạo (AI). Bộ dữ liệu gồm 7.600 giá trị điện trở suất biểu kiến được tạo ra thông qua phương pháp phần tử hữu hạn (Finite Element Method – FEM) nhằm mô phỏng đặc tính điện trở suất dưới bề mặt trong cấu trúc đất hai lớp. Bộ dữ liệu bao gồm nhiều tổ hợp khác nhau của điện trở suất các lớp đất (ρ1, ρ2) và chiều dày các lớp (l1, l2), từ đó xây dựng tập dữ liệu huấn luyện đa dạng và có tính đại diện cao. Mô hình học máy dựa trên mạng nơ-ron nhân tạo nhiều lớp (Multi-Layer Perceptron – MLP), một dạng của mạng nơ-ron nhân tạo (ANN), được huấn luyện để dự đoán điện trở suất biểu kiến từ các thông số đầu vào. Kết quả cho thấy mô hình MLP đề xuất có khả năng dự đoán điện trở suất biểu kiến với độ chính xác cao, đồng thời giảm đáng kể chi phí tính toán so với các phương pháp khảo sát và tính toán truyền thống. Nghiên cứu khẳng định tiềm năng kết hợp giữa FEM và AI trong các ứng dụng địa vật lý, đồng thời cung cấp một khuôn khổ hiệu quả để dự đoán điện trở suất dưới bề mặt trong các hệ đất phân lớp.

Từ khóa: Điện trở suất biểu kiến; lớp đất phân tầng; FEM; ANN; MLP.

1. INTRODUCTION

The study of apparent resistivity is a fundamental aspect of geophysical exploration, particularly for subsurface investigations in applications such as mineral exploration, environmental monitoring, and geotechnical engineering. Apparent resistivity values are typically determined through surface-based measurements, but accurate prediction of these values requires a comprehensive understanding of the subsurface resistivity structure, which can vary with soil composition, layer thickness, and moisture content.

Traditional methods for calculating apparent resistivity often involve analytical or empirical models, which may be limited in flexibility and computational efficiency, particularly for complex, layered subsurface structures. Numerous studies have investigated resistivity modeling and inversion techniques for layered soil systems. For example, Yu et al. [1] demonstrated that conventional analytical methods fail to capture magnetotelluric responses in rugged terrain, and improved both accuracy and efficiency by introducing a hybrid FEM mesh of triangular prisms and tetrahedrons. Wu e tal. [2] showed through FEM simulations of laminated shaly sandstones that shale distribution and laminae homogeneity strongly affect resistivity, and proposed a new water saturation equation more accurate than Archie-type models. Jesenik and Trbušić [3] highlighted the limitations of horizontally layered soil models, employing metaheuristic optimization to calibrate multilayer parameters and validating results with FEM simulations. Their findings indicated that appropriately optimized simplified three-layer models can still provide reliable accuracy. Collectively, these studies confirm that analytical and empirical approaches form a necessary foundation but are often insufficient for accurately characterizing resistivity in complex or heterogeneous soils, motivating the growing reliance on finite element modeling and advanced numerical optimization.

Numerical methods, such as the Finite Element Method (FEM), offer powerful tools for simulating the behavior of electric fields in heterogeneous materials, enabling more accurate and detailed models of subsurface resistivity [4]. FEM allows for the consideration of complex boundary conditions and varying resistivity values across different soil layers. However, the success of FEM-based models in practical geophysical applications heavily relies on the availability of accurate data, which traditionally comes from field measurements or simplified laboratory models.

Meanwhile, recent advancements in machine learning (ML) have opened new opportunities for predicting soil resistivity with higher efficiency. Zamanian et al. [5] applied deep learning models to a dataset of 842 clayey soil samples, incorporating geotechnical parameters such as water content, dry unit weight, plasticity index, and fine fraction, demonstrating the dominant influence of water content and unit weight. Rashid Mustafa and Ansari [6] conducted a comprehensive experimental and machine learning study on soil electrical conductivity using hybrid

models with inputs such as water content, dry density, degree of saturation, porosity, and applied voltage. Furthermore, Şehmusoğlu et al. [7] employed deep learning to predict soil liquefaction, proving the effectiveness of AI techniques for soil investigation. In addition, Carbonari et al. [8] proposed a machine learning-based framework for predicting subsurface electrical resistivity from petrophysical and thermodynamic data, further demonstrating the capability of data-driven models to approximate complex subsurface electrical behavior. These studies demonstrate that AI-based approaches are effective tools for capturing the complex, non-linear relationships involved in soil resistivity prediction.

Despite the growing interest in ML methods for soil resistivity prediction, many traditional approaches are constrained by their reliance on laboratory or field data, and fail to capture complex, heterogeneous subsurface conditions. Recent advancements in the Finite Element Method (FEM) have enabled more accurate and detailed simulations of resistivity profiles in multi-layered soils. These advancements, when coupled with ML techniques, hold significant promise for improving real-time predictions and optimizing geophysical assessments.

This study aims to fill this gap by leveraging the Finite Element Method (FEM) to generate a large dataset of apparent resistivity values under various conditions, such as different soil resistivities ($\rho1$, $\rho2$) and layer thicknesses ($l1$, $l2$), for a two-layered soil model. The dataset, consisting of 7600 data points, serves as the foundation for training an ANN model to predict apparent resistivity based on the input parameters. The novelty of this research lies in the generation of comprehensive training data using FEM and its application to ANN - based resistivity prediction. While previous studies have employed machine learning models for resistivity prediction, none have combined FEM-generated data in such a structured manner, offering an efficient and scalable solution for real-time geophysical assessments.

The results of this work demonstrate the potential of ANN models to provide accurate predictions of apparent resistivity. This approach offers a more efficient alternative to traditional field-based methods, paving the way for faster and more accurate geophysical investigations. This paper is organized as follows: Section 2 provides a detailed description of the FEM-based model used to generate the training data, Section 3 presents the ANN based prediction models, Section 4 discusses the results of model training and prediction accuracy, and Section 5 concludes with a summary and suggestions for future work.

In this study, we leverage the Finite Element Method (FEM) to generate a large dataset of apparent resistivity values under various conditions, such as different soil resistivities ($\rho1$, $\rho2$) and layer thicknesses ($l1$, $l2$), for a two-layered soil model. This dataset, consisting of 7600 data points, serves as the foundation for training the ANN model to predict apparent

resistivity based on the input parameters. The aim of this work is to explore the potential of MLP-ANN techniques to provide accurate predictions of apparent resistivity, offering an efficient and scalable alternative to traditional methods.

This paper is organized as follows: Section 2 provides a detailed description of the FEM-based model used to generate the training data, Section 3 presents the MLP-ANN based prediction models, Section 4 discusses the results of model training and prediction accuracy, and Section 5 concludes with a summary and suggestions for future work.

2. METHODOLOGY

This section describes the methodology used in this study, which consists of two main components: the COMSOL-based finite element model used to generate the apparent resistivity training data, and the artificial intelligence (AI) model employed for resistivity prediction. These two components are integrated to create a data-driven approach for accurate and efficient apparent resistivity prediction in two-layered soil systems.

2.1. COMSOL Model for Apparent Resistivity Calculation

The first step in the methodology involves creating a detailed finite element model in COMSOL Multiphysics® to simulate the apparent resistivity of a two-layered soil system. The model consists of two distinct layers with different resistivity values, denoted as ρ1 and ρ2, and varying layer thicknesses, denoted as l1 and l2.

The geometry of the model was set up in COMSOL, where the two layers are defined with the following parameters:

Layer 1: resistivity ρ1, thickness l1

Layer 2: resistivity ρ2, thickness l2

In the COMSOL model, the current electrodes have a depth of 0.5 m below the ground surface. The electrodes are arranged along a straight line with a spacing of 10 m between adjacent electrodes (Figure 1). This configuration allows the measurement of the electrical potential difference between the voltage electrodes, which is used to calculate the apparent resistivity.

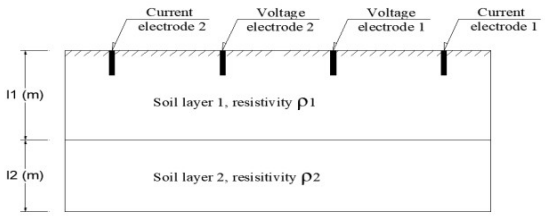

Figure 1. COMSOL model

Boundary conditions are defined at the surface and the interface between the two layers to simulate the electrical potential distribution within the system under a DC current. A current source is applied at the current electrodes, while the reference electrode is grounded. The injected current is set to I = 1 A. A voltage source is applied at one boundary, while the other boundary is grounded. The model is solved using the AC/DC module in COMSOL, which solves the electric field and calculates the potential distribution in the soil layers.

The Quasi-Static, Electrical Mode (EC) of the AC/DC module in COMSOL is used to

compute the potential at the voltage electrodes, the governing equations in this mode are:

$$J = \sigma E + J^e \quad (1)$$

$$E = -gradV \quad (2)$$

where J is the current density, σ is the conductivity, E is the electric field, Je is the externally generated current density, and V is the potential.

The ground surface is set at insulation boundary where n.J = 0. The computational domain has dimensions of 50 m × 50 m × 35 m (length × width × depth) to avoid boundary effects and simulate an effectively infinite medium. In order to extend the simulation domain to a far enough limit, the hemisphere domain is surrounded by an infinite element region with a thickness of 5 m. The current flowing into the current electrodes is set to be I = 1 A.

For each simulation, a range of resistivity and layer thickness values was used as input parameters. The apparent resistivity is calculated from the potential difference measured at various points in the model, considering the geometry and material properties of the layers. The dataset used for training the AI model consists of 7600 data points, each corresponding to different combinations of resistivity values ρ1, ρ2, and layer thicknesses l1, l2, along with the corresponding apparent resistivity.

The simulations in COMSOL were performed for different configurations to generate a wide variety of training data. The model used for training includes both the potential measurements at specific locations and the associated resistivity values for the two soil layers, making it possible to train the AI model to predict apparent resistivity based on these input parameters.

2.2. Model for Apparent Resistivity Prediction

The second component of the methodology involves using an ANN to predict the apparent resistivity based on the training data generated from the FEM simulations in COMSOL. In this study, a MLP model is employed, which is a type of feed-forward artificial neural network consisting of multiple layers of neurons. This model learns the complex relationship between the input parameters and the output apparent resistivity, enabling fast and accurate predictions.

The model is trained using the dataset of 7600 samples, which includes the following input parameters:

- ρ1: Resistivity of Layer 1 (in Ω.m)
- ρ2: Resistivity of Layer 2 (in Ω.m)
- l1: Thickness of Layer 1 (in meters)
- l2: Thickness of Layer 2 (in meters)

These input features are combined to predict the apparent resistivity ρapp, which is the output of the model.

The architecture of the MLP model consists of the following components:

1. Input Layer: This layer accepts the four input parameters (ρ1, ρ2, l1, l2).

2. Hidden Layers: Two fully connected dense layers, each with 128 neurons and ReLU (Rectified Linear Unit) activation functions,

are used to capture the complex non-linear relationships between the inputs and output.

3. Output Layer: A single neuron that outputs the predicted apparent resistivity value ρapp.

The model is compiled using the Adam optimizer and mean squared error (MSE) as the loss function, which is suitable for regression tasks like apparent resistivity prediction. The performance of the model is evaluated using the mean absolute error (MAE), a metric that measures the average magnitude of errors in the predictions, and MSE.

To prevent overfitting and ensure that the model generalizes well to unseen data, early stopping is applied during the training process. Early stopping monitors the validation loss and halts the training process if the validation loss does not improve for a specified number of epochs. This allows the model to retain the best weights achieved during training.

The training process involves splitting the dataset into training and testing subsets. The training data is scaled using StandardScaler to normalize the input features and improve convergence during training. The dataset is divided into 80% for training and 20% for validation. The model is trained over a maximum of 200 epochs, with a batch size of 32, and the early stopping callback ensures that the model's performance is optimized.

After training, the model is evaluated on the testing set to assess its predictive performance. The results, including test loss and MAE, are reported to demonstrate the effectiveness of the MLP model in predicting the apparent resistivity. Additionally, the trained model is used to make predictions on new data, providing fast and reliable estimates for apparent resistivity based on the input parameters.

3. RESULTS AND DISCUSSION

In this study, a comprehensive approach was developed to predict the apparent resistivity of a two-layered soil model using a Finite Element Method (FEM) model and a Multi-Layer Perceptron (MLP) artificial neural network (ANN). The FEM model was used to generate a dataset of 7600 data points, representing various conditions of soil resistivities (ro1, ro2) and layer thicknesses (11, 12). This dataset served as the foundation for training the MLP model, which was designed to predict the apparent resistivity based on these input parameters.

3.1. FEM Model Results

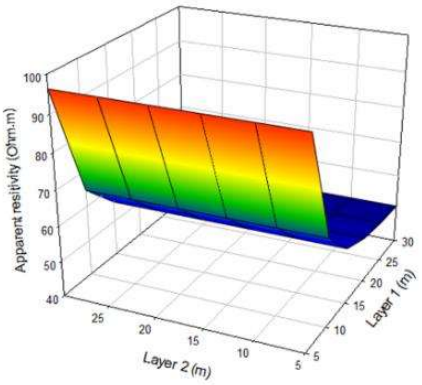

Figure 2. Apparent resistivity computed from COMSOL

The dataset generated by the FEM model is composed of the key parameters that influence the apparent resistivity, including the resistivity of the top layer (ρ1), the resistivity of the

bottom layer (ρ2), the thickness of the top layer (l1), and the thickness of the bottom layer (l2). The data points were simulated under various combinations of these parameters to ensure a diverse training set for the MLP model, Fig. 2. These data points were used to train the model, allowing the system to learn the relationships between the input parameters and the resulting apparent resistivity values.

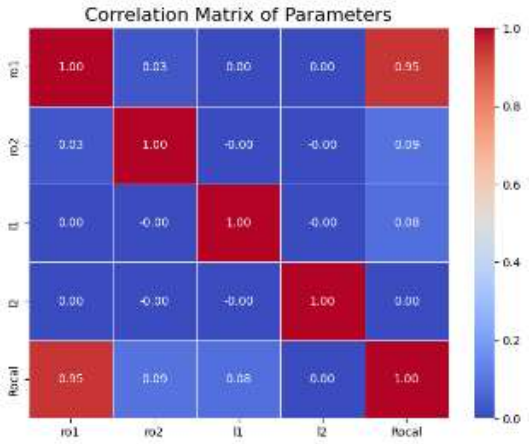

Figure 3. Correlation matrix of the dataset parameters

The input dataset comprising 7,600 simulated geophysical samples was generated by COMSOL. Correlation analysis reveals consistent physical relationships, with a strong positive correlation between primary resistivity and apparent resistivity, confirming the simulation's accuracy. (Figure 3 and Figure 4).

This balanced distribution and clear physical correlation provide a robust foundation for training the neural network. Ultimately, the dataset is highly structured, effectively capturing the complex non-linear characteristics essential for accurate predictive modeling.

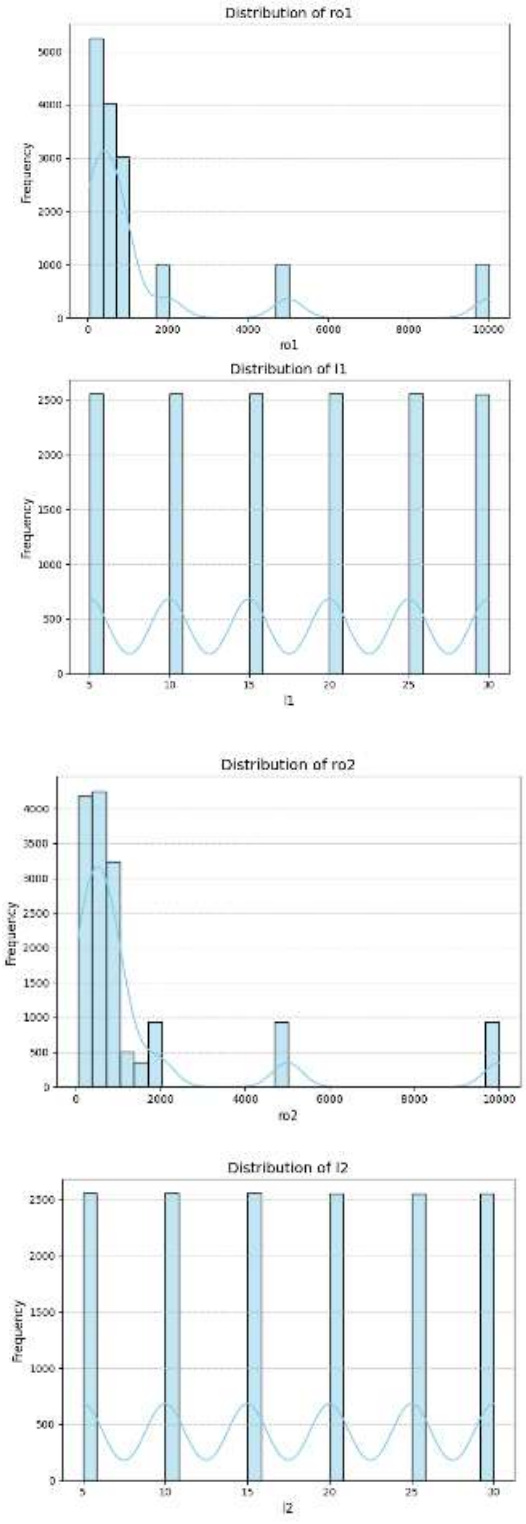

Figure 4. Statistical distribution of the input parameters.

3.2. MLP Model Training and Performance

The MLP model, which consists of several fully connected layers, was trained using the generated dataset. The training was conducted using the Mean Squared Error (MSE) loss function, with 200 epochs and early stopping to prevent overfitting. The results of the training process are shown in the loss curve (Figure 5), which indicates a significant reduction in both training and validation losses over the course of the training.

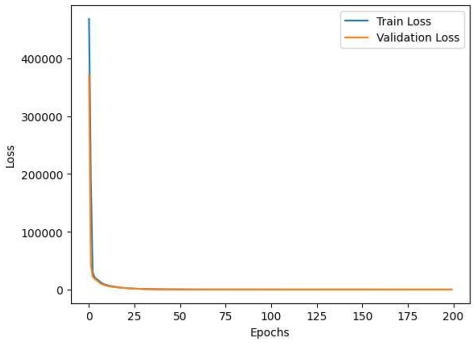

Figure 5. Training and Validation Loss

To assess generalization, three regression diagnostics are included. The predicted versus true scatter (Figure 6) closely follows the ideal y = x reference, a linear fit yields a slope approximately equal to 1 and near-zero intercept, indicating negligible global bias.

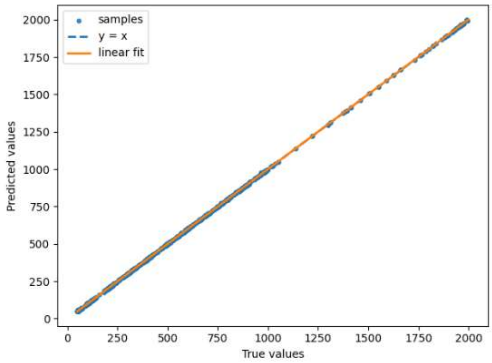

Figure 6. Predicted vs. True values

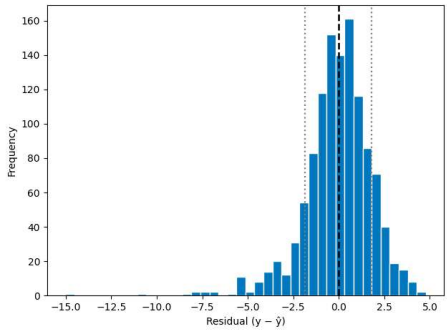

Figure 7. Error Distribution

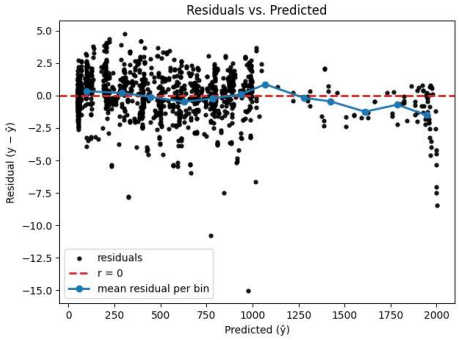

Figure 8. Residuals vs. Predicted

The error distribution (Figure 7) is centered near zero, with 73.5% and 93.0% of test residuals within ±1σ and ±2σ, respectively, suggesting small, symmetric errors with few outliers. The residuals versus predicted plot (Figure 8) displays residuals evenly scattered around the r=0 line; the mean residual in each bin remains close to zero, suggesting no heteroscedasticity.

The model achieved a test Mean Absolute Error (MAE) of approximately 1.6219, and the test loss was found to be 5.0609, which indicates a good performance for predicting the apparent resistivity. These results demonstrate the potential of using machine learning techniques like MLP to provide accurate predictions of apparent resistivity for geophysical applications.

The trained model is available at the following link.

3.3. Model Evaluation

The model was evaluated by comparing the predicted apparent resistivity values with the actual results from the FEM simulations. The predictions closely matched the true values, as shown by the small discrepancy between the predicted and actual resistivity values in the test set. The MLP model's ability to generalize across unseen data indicates that it has learned the underlying relationships between the input parameters and the apparent resistivity.

The prediction for one of the test cases showed that the model estimated the apparent resistivity as 206.55554 Ω.m (predicted) compared to the actual value of 207.602968 Ω.m (true), with only a small error margin, confirming the model's ability to make accurate predictions.

4. DISCUSSION

The results of this study demonstrate the effectiveness of combining FEM-generated data with an MLP model for the prediction of apparent resistivity. The model achieved an satisfactory predictive performance in terms of both accuracy and computational efficiency. This approach presents a promising alternative to traditional field-based measurements, which can be time-consuming and costly.

Furthermore, the ability of the MLP model to process large datasets and provide real-time predictions makes it highly suitable for applications in environmental monitoring, geotechnical engineering, and mineral exploration, where rapid decision-making is essential. Future improvements may include fine-tuning the model architecture and exploring different neural network approaches to further enhance prediction accuracy.

In conclusion, this research highlights the significant potential of AI-based methods, particularly MLP, in revolutionizing the way we predict and model subsurface resistivity. By utilizing FEM-generated data, we have demonstrated a scalable and efficient method for predicting apparent resistivity, offering a valuable tool for various geophysical applications.

5. CONCLUSION

In this study, we demonstrated the application of the Finite Element Method (FEM) to generate a large-scale dataset for the prediction of apparent resistivity in two-layered soil models. The dataset, consisting of 7600 data points with varying soil resistivities (ρ1, ρ2) and layer thicknesses (l1, l2), was used to train a MLP model for the efficient and accurate prediction of apparent resistivity. The results highlight the effectiveness of the MLP model in providing precise predictions, offering a more scalable and faster alternative to traditional field-based resistivity measurement methods.

Our work bridges the gap between FEM simulations and machine learning, providing a framework for generating comprehensive training data that can be used for real-time geophysical assessments. The integration of FEM-generated data with MLP techniques has shown great promise in advancing geophysical exploration, enabling faster decision-making and more accurate predictions of subsurface

properties. This approach can be particularly beneficial in applications such as mineral exploration, environmental monitoring, and geotechnical engineering, where precise resistivity predictions are essential for effective decision-making.

The MLP model achieved promising results, with a test Mean Absolute Error (MAE) of approximately 1.6219 and a test loss value of 5.0609, confirming the model's capability to make accurate predictions. These results demonstrate the potential of AI-driven techniques to replace traditional methods, offering a computationally efficient solution for resistivity prediction.

Despite the promising results, several areas remain for further development. Future work could focus on expanding the dataset to include more complex subsurface structures, such as multi-layered and heterogeneous soil models. Additionally, exploring the use of more advanced machine learning techniques, such as deep learning models or ensemble methods, could further improve the accuracy and robustness of the predictions. Incorporating real-world field data into the training process would enhance the generalization of the MLP model, making it more adaptable to diverse geological conditions.

Furthermore, the integration of the MLP driven approach with real-time geophysical measurement systems could open new possibilities for on-site, automated resistivity mapping. By combining real-time data acquisition with AI-powered prediction models, it would be possible to achieve immediate, actionable insights for various geophysical applications.

In conclusion, this study paves the way for the efficient use of MLP-based ANN models in geophysical applications, providing a solid foundation for future research aimed at improving the speed, accuracy, and applicability of resistivity prediction models in subsurface investigations. The integration of machine learning with FEM-generated data represents a significant advancement in geophysical modeling, with promising implications for real-time, large-scale applications in subsurface investigations.

REFERENCES

[1]. Yu, N., Dong, X., Gao, H., and Li, Y., Hybrid mesh for magnetotelluric forward modeling based on the finite element method. Scientific Reports, January 2023, 13.1: 532, DOI:10.1038/s41598-023-27758-2.

[2]. Wu, F., Cong, L., Ma, W., Liu, J., Dai, J., Nie, X., Shi, X., Luo, Y., and Chen, C. , Finite element method-based resistivity simulation and water saturation calculation of irregular laminated shaly sandstone, Geomechanics and Geophysics for Geo-Energy and Geo-Resources, February 2023, Volume 9, article number 2.

[3]. Jesenik, A., and Trbušić, M., The Approach of Using a Horizontally Layered Soil Model for Inhomogeneous Soil, by Taking into Account the Deeper Layers of the Soil, and Determining the Model's Parameters Using Evolutionary Methods, Applied Sciences, 2025, 15.3: 1420, https://doi.org/10.3390/app15031420.

[4]. Shuaiying Qiao, Tiaojie Xiao, Junjun Zhou, Chunye Gong, Bo Yang, Jie Liu, Yun Wang, Qinglin Wang, Parallel finite element forward modeling of 3-D magnetotelluric conductivity and permeability anisotropy with coupled PML boundary conditions, Computers & Geosciences, Volume 207, 2026, 106064, ISSN 0098-3004, https://doi.org/10.1016/j.cageo.2025.106064.

[5]. Zamanian, M., Asfaw, N., Chavda, P., & Shahandashti, M., Deep learning for exploring the relationship between geotechnical properties and electrical resistivities, Journal of the Transportation Research Board, 2024, Volume 2678, Issue 10, https://doi.org/10.1177/036119812412349.

[6]. Rashid Mustafa, and Alauddin Ansari, Comprehensive analysis of soil electrical conductivity: an experimental and machine learning approach, Discover Civil Engineering 1(82), September 2024, DOI: 10.1007/s44290-024-00086-8.

[7]. Şehmusoğlu, E.H., Kurnaz, T.F. & Erden, C. Estimation of soil liquefaction using artificial intelligence techniques: an extended comparison between machine and deep learning approaches. Environ Earth Sci 84, 130, 2025. https://doi.org/10.1007/s12665-025-12116-4.

[8]. Carbonari, R., Salone, R. & Di Maio, R. Predicting Electrical Resistivity in Natural Degassing Geological Systems Through Petrophysical and Thermodynamic Data: A Machine Learning Approach. Pure Appl. Geophys. 183, 1177-1201 (2026). https://doi.org/10.1007/s00024-025-03882-0.