Estimation of soil moisture from GPR data using artificial neural networks

Nairit Barkataki*, Sharmistha Mazumdar*, Banty Tiru† and Utpal Sarma*
*Dept of Instrumentation & USIC, Gauhati University, Guwahati, India
†Dept of Physics, Gauhati University, Guwahati, India

***Abstract*—Soil moisture estimation is essential for understanding the water cycle and its impact on weather and climate. GPR based soil moisture estimation is non-invasive in nature and provides quicker results as compared to standard laboratory approaches. Deep learning algorithms have been shown to be an effective method for extracting characteristics from GPR data. To assess soil moisture, this paper offers a deep artificial neural network (ANN) model. The data set was created using finite-difference-time-domain (FDTD) simulation and is used to train and validate the ANN model. The suggested model predicts soil moisture from GPR A-Scans well, with R^2 = 0.998 and mean average percentage error (MAPE) of 1.23.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Soil moisture is considered to be a key variable in environmental studies such as meteorology, hydrology, agriculture and climate change. The estimation of soil moisture provides better understanding of the soil dynamics for making proactive decisions in the field of agriculture.

Soil moisture measurement is generally classified into direct and indirect techniques. Direct methods usually involve digging of holes to collect soil samples and then separating water from the collected samples. The separation of water from soil can be achieved by heating the soil samples or by chemical reactions [1]. These methods gives highly accurate results for a particular depth but they are expensive, time consuming and destructive [2].

Indirect methods mainly involve measurement of physical or chemical properties of soil that are dependent on the amount of soil moisture [1]. Previous studies shows that the electromagnetic (EM) properties of soil is correlated to its moisture content [3]. Time domain reflectrometry (TDR), ground penetrating radar (GPR) and some other techniques use EM properties for measurement of soil moisture.

Calamita et al. studied the variation in soil moisture with soil resistivity in the Vallaccia catchment (central Italy) using portable TDR, along with frequency domain reflectrometer sensors. The results were compared by applying various resistive methods and showed high correlation between resistivity and soil moisture [4]. However, TDR probes limits the depth of investigation to a few meters [5].

Brocca et al. explored the reliability soil moisture estimation from satellite data. Correlation values of relative soil moisture increase upto 0.81 upon application of different filters but decrease with increase in vegetation [6].

Salam et al. developed a real time, in-situ method based on wireless underground communications (WUC) for estimation of soil permittivity and moisture. To obtain highly accurate results, studies were carried out in the frequency range of 100–500 MHz with the antenna buried at 10 cm-40 cm depths in various soils [7].

Compared to TDR, GPR technique has the advantage of giving data from greater spatial regions. GPR has been widely used in sedimentary research [8] and for estimating water content of soil [9]–[11].

GPR is widely used for object size detection [12], [13]. However, the size of a target observable with a GPR is dependent on the antenna's centre frequency [14]. The frequency useful for different applications depends on the soil properties such as soil moisture [2] and soil type [15]. Thus, prior knowledge about the soil moisture content of the soil helps in determining the appropriate frequency to be used for a particular application.

Hubbard et al. examined the potential of estimating soil moisture over large areas with varying hydrological conditions using GPR reflections [16]. Benedetto et al. proposed an approach to estimate soil moisture for pavement inspection using GPR. The results were validated using capacitance probes [17].

However, the interpretation of GPR data is still a manual process. It is done by people who have extensive experience in GPR data interpretation. Of late, machine learning algorithms have been implemented for processing and interpretation of GPR data. Rahman et al. predicted 11 soil series using machine learning techniques in order to find appropriate crops for cultivation [18]. Senanayake et al. developed an neural network model to estimate catchment scale soil moisture fom remote sensing and GPR data [19]. A long short-term memory (LSTM) predictive framework integrated with Boruta-random forest (BRF) optimiser was proposed by Ahmed et al. to estimate soil moisture under global warming scenarios [20]. An SVM model was proposed by Malajner et al. to classify soil types [21]. After determining the soil type, another model for estimating soil moisture was employed. Ultra-wideband radio frequency modules and antennas were used to capture the data. Recently, various regression and machine learning models were proposed by Teneja et al. to predict soil organic matter and soil moisture content using images taken from cell phones [22].

Although various remote sensing techniques are utilised to estimate soil moisture, further research is needed to investigate the feasibility of assessing soil moisture content using GPR. This paper mainly focuses on

1) Creation of a synthetic GPR database for estimation of soil moisture.
2) Development of a regression model to estimate soil moisture with a high degree of accuracy.

II. GPR Data

GprMax, an open source electromagnetic simulator, is used to generate the database. It uses FDTD method for simulation of GPR model. Each simulated model has dimensions of $1000\ mm \times 400\ mm \times 600\ mm$ $(X \times Y \times Z)$. The top 200 mm of the model's overall height, i.e. 600 mm, is an air layer, while the remaining 400 mm is a soil layer. Soil moisture is varied randomly between 1% and 10% and the corresponding values of relative permittivity (ε_r) and conductivity (σ) are used in simulation. Table I shows the soil permittivity and conductivity as a function of moisture content [23]. For fractional values of soil moisture, the corresponding values of ε_r and σ are intrapolated.

TABLE I
Relative permittivity and conductivity corresponding to moisture content in soil

S. No.	Soil Moisture	Relative Permitivity ε_r	Conductivity σ(S/m) $\times 10^-6$
1	1%	3.2343	9.9
2	2%	3.4686	18.8
3	3%	3.7029	27.7
4	4%	3.9372	36.6
5	6%	4.4058	54.5
6	7%	4.6401	63.4
7	8%	4.8744	72.3
8	9%	5.1087	81.2
9	10%	5.343	90.1

A cylindrical object of aluminium (ε_r = 10.8 and $\sigma = 3.5 \times 10^7$ S/m) is placed underneath the soil surface. The object depth varies from 25 mm to 365 mm from the top of the soil layer and the object radius has values between 9 mm to 20 mm. The depth and radius values are chosen randomly using a program. A simulated model with a cylindrical object buried in the soil is shown in Figure 1.

The simulation time window must be long enough for the EM waves to travel through the medium to the cylindrical object and then reflect back to the receiver. However, the time window depends on the relative permittivity of the soil. In this case, the value of time window is taken such that it is suitable for all the different values of soil moisture. The different parameters used while simulating the GPR models are given in Table II. A simulated A-Scan having 6% of moisture in soil is shown in Figure 2.

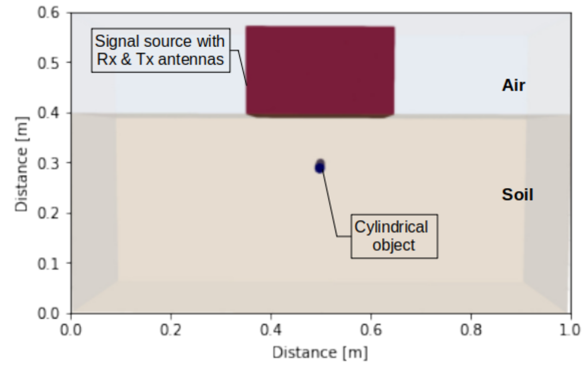

Fig. 1. Simulated model

TABLE II
Simulation parameters

Sl. No.	Simulation Parameter	Values
1	Excitation waveform	Gaussian
2	Frequency used	400 MHz
3	Spatial resolution	2 mm
4	Time window	12 ns

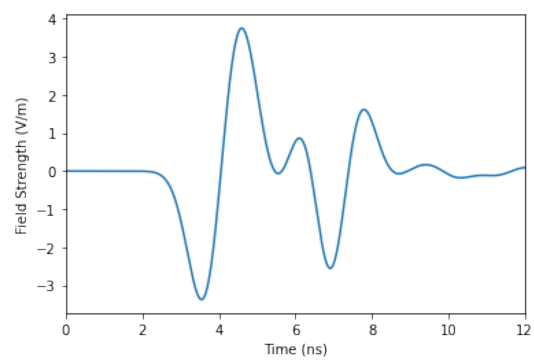

Fig. 2. Simulated A-Scan

A total of 2690 A-Scans are generated, each having 3113 data points along the depth of the model. The dataset contains around 8 million data points. NVIDIA GPU RTX3090 is used to accelerate the simulation and generate the dataset. 85% of the dataset is used to train and validate the model and 15% of it is kept for testing the model. The dataset is normalised before training the model.

III. Methodology

In this work, an ANN model is used. An ANN model consists of multiple layers, each of which has one or more neurons or units. Each of these neurons is linked to every other neuron in the following layer. In ANN, feature extraction is performed hierarchically, with increments after each layer.

Forward propagation and back propagation are the two stages of ANN. Forward propagation entails assigning weights

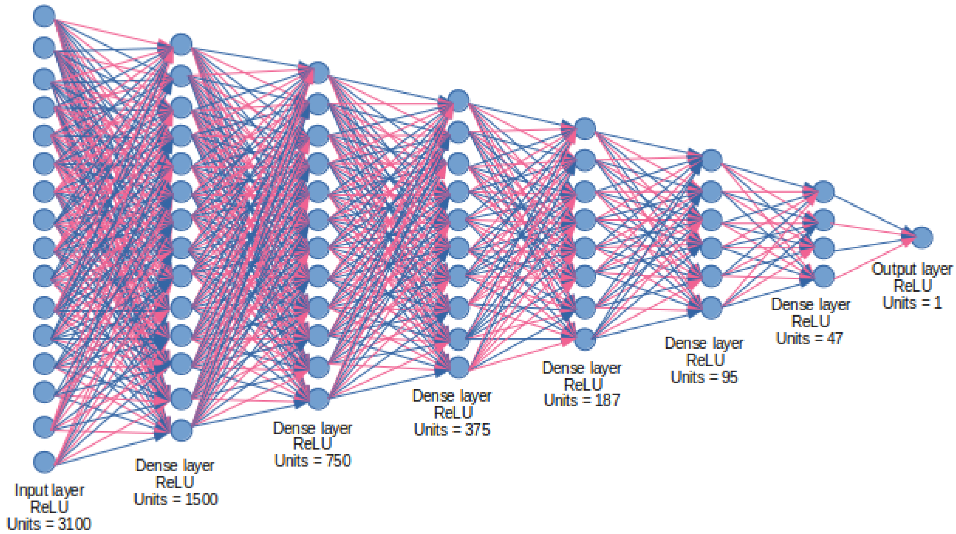

Fig. 3. Proposed ANN architecture

and biases to each neuron in a layer, as well as applying an activation function to each neuron. A loss function is used to calculate the error between the true and predicted values. Using an optimisation function, backward propagation assists in determining optimal parameter values for the model, by minimising the loss function.

A. Proposed model

Initially, 3113 neurons were used in the input layer. Multiple hidden layers were added after the input layer which had half the number of neurons than the previous layers. After tuning the number of neurons through multiple training runs, the optimised model is found to have 3100 units in the input layer, followed by 6 hidden layers having 1500, 750, 375, 187, 95 and 47 units respectively, as shown in Figure 3. The output layer contains a single unit that produces the predicted values. All the layers in the neural network implement the function:

$$output = activation(dot(input, kernel) + bias) \quad (1)$$

where *activation* is the activation function for the layer, *kernel* is the weights matrix created by the layer, *bias* is the bias vector created by the layer.

Each layer's activation function is a Rectified Linear Unit (ReLU), which can be described mathematically as shown in 2

$$f(x) = max(0, x) \quad (2)$$

,where x is a neuron's input.

If x is less than zero, the activation function f(x) returns 0; otherwise, it returns x (input). To train a deep learning model, the input data is fed to the model and predictions are generated. The predicted values are compared to the actual values and the loss is calculated using a loss function. The loss function used in this study is Mean Absolute Percentage Error (MAPE). Adaptive gradient (AdaGrad) algorithm is used to optimise the model with an initial learning of 1×10^{-5}. The training and validation losses over 300 epochs is shown in Figure 4. The GPR data processing pipeline is written in python along with Keras and Tensorflow frameworks for application of deep learning algorithms.

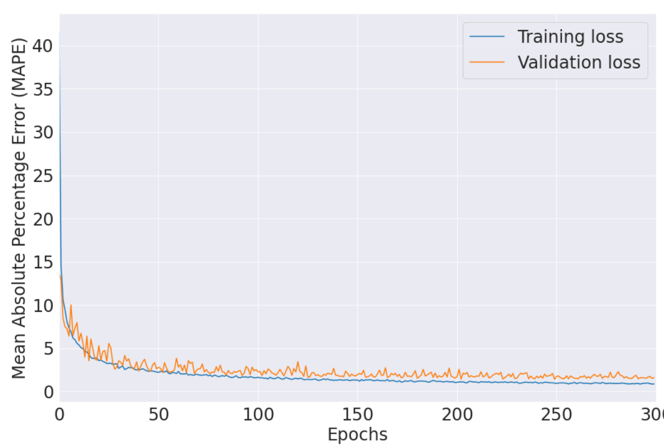

Fig. 4. Loss curve during training and validation of the model

IV. Results

429 A-Scans are used for final performance evaluation of the proposed model. The best and worst predicted values corresponding the different values of soil moisture are given in Table III. Moreover, different metrics such as mean squared error (MSE), mean absolute error (MAE), mean squared logarithmic error (MSLE), mean absolute percentage error (MAPE) and R-squared are used to analyse the performance of the ANN model as given in Table IV.

TABLE III
PREDICTED VALUES OF SOIL MOISTURE BY THE ANN MODEL

Sl. No.	True Value (%)	Best Predicted Value (%)	Worst Predicted Value (%)	Median of Predicted Values (%)
1	1	0.999	1.467	0.999
2	2	2.002	1.908	1.992
3	3	2.999	3.451	3.008
4	4	3.999	4.244	4.005
5	6	6.000	6.342	6.008
6	7	7.001	7.355	7.009
7	8	8.001	7.737	8.014
8	9	8.999	9.536	9.033
9	10	10.000	9.488	9.927

TABLE IV
PERFORMANCE METRICS OF THE ANN MODEL

Metrics	Performance of the ANN model
MSE	9.562×10^{-7}
MAE	0.00055
MAPE	1.232291
MSLE	8.344×10^{-7}
R-squared	0.998

V. CONCLUSION

This study presented a ANN model for estimating soil moisture from GPR A-Scan data. The overall performance of the model was evaluated using the metrics listed in Table IV. The most popular statistical metric for a regression problem is R-squared whose value lies between 0 and 1. In order for a regression model to be reliable, the value should be greater than 0.95. The proposed model achieves an R-squared value of 0.998 and MAPE of 1.23 on the test set, demonstrating its high reliability.

A comparison with previously reported literature work is shown in Table V. Most of the techniques used to estimate soil moisture are invasive in nature. Moreover, soil moisture estimation was done using satellite data and laboratory-based methods, which rarely produced results in real time. The proposed work uses GPR, non-invasive method, to collect data. This will lead to faster data collection and is capable of producing results in real time.

TABLE V
COMPARISON OF PRESENT WORK WITH PAST STUDIES

Sl. No.	Authors	Technique used	Real time/ Offline	R^2
1	Lee et al. (2018) [24]	Deep neural network applied on remotely sensed satellite data (non-invasive)	Offline	0.79
2	Koley et al. (2020) [25]	Surface soil moisture estimated using LST-NDVI feature space from satellite data (non-invasive)	Offline	0.8
3	Taneja et al. (2021) [22]	Machine learning algorithms applied on soil sample images collected using cell phone. (invasive)	Offline	0.95
4	**Present work**	ANN applied on GPR data (non-invasive)	Real time	0.998

However, the proposed model is trained and tested in dataset with limited variation of soil moisture. In real scenarios, different environmental factors might lead to large variations in soil moisture. To improve the proposed model, the authors plan to implement the proposed model on real data.

VI. ACKNOWLEDGEMENT

The authors are grateful to Google for providing Google Colaboratory, a computational platform, used for training and evaluation of the proposed model.

REFERENCES

[1] M. R. Carter and E. G. Gregorich, *Soil sampling and methods of analysis*. CRC press, 2007.

[2] X. Liu, X. Dong, and D. I. Leskovar, "Ground penetrating radar for underground sensing in agriculture: a review," *International Agrophysics*, vol. 30, no. 4, 2016.

[3] R. Porretta and F. Bianchi, "Profiles of relative permittivity and electrical conductivity from unsaturated soil water content models," *Annals of Geophysics*, vol. 59, no. 3, p. 0320, 2016.

[4] G. Calamita, L. Brocca, A. Perrone, S. Piscitelli, V. Lapenna, F. Melone, and T. Moramarco, "Electrical resistivity and tdr methods for soil moisture estimation in central italy test-sites," *Journal of Hydrology*, vol. 454, pp. 101–112, 2012.

[5] G. C. Topp, G. Parkin, T. P. Ferré, M. Carter, and E. Gregorich, "Soil water content," *Soil sampling and methods of analysis*, pp. 939–962, 2008.

[6] L. Brocca, S. Hasenauer, T. Lacava, F. Melone, T. Moramarco, W. Wagner, W. Dorigo, P. Matgen, J. Martínez-Fernández, P. Llorens *et al.*, "Soil moisture estimation through ascat and amsr-e sensors: An intercomparison and validation study across europe," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3390–3408, 2011.

[7] A. Salam, M. C. Vuran, and S. Irmak, "Di-sense: In situ real-time permittivity estimation and soil moisture sensing using wireless underground communications," *Computer Networks*, vol. 151, pp. 31–41, 2019.

[8] A. Neal, "Ground-penetrating radar and its use in sedimentology: principles, problems and progress," *Earth-science reviews*, vol. 66, no. 3-4, pp. 261–330, 2004.

[9] J. A. Huisman, S. S. Hubbard, J. D. Redman, and A. P. Annan, "Measuring soil water content with ground penetrating radar: A review," *Vadose zone journal*, vol. 2, no. 4, pp. 476–491, 2003.

[10] A. Klotzsche, F. Jonard, M. C. Looms, J. van der Kruk, and J. A. Huisman, "Measuring soil water content with ground penetrating radar: A decade of progress," *Vadose Zone Journal*, vol. 17, no. 1, pp. 1–9, 2018.

[11] M. Zhang, M. Bano, X. Feng, N. Lesparre, J. Girard, O. Razakarisoa, B. Belfort, F. Lehmann, and P. Friedmann, "Estimating water content of unsaturated sandy soils by gpr during a drainage experiment," in *IOP Conference Series: Earth and Environmental Science*, vol. 660, no. 1. IOP Publishing, 2021, p. 012018.

[12] N. Barkataki, S. Mazumdar, R. Talukdar, P. Chakraborty, B. Tiru, and U. Sarma, "Prediction of size of buried objects using ground penetrating radar and machine learning techniques," in *2020 International Conference on Computational Performance Evaluation (ComPE)*. IEEE, 2020, pp. 781–785.

[13] I. Giannakis, A. Giannopoulos, and C. Warren, "A machine learning scheme for estimating the diameter of reinforcing bars using ground penetrating radar," *IEEE Geoscience and Remote Sensing Letters*, 2020.

[14] N. Barkataki, B. Tiru, and U. Sarma, "Performance investigation of patch and bow-tie antennas for ground penetrating radar applications," *International Journal of Advanced Technology and Engineering Exploration*, vol. 8, no. 79, p. 753, 2021.

[15] N. Barkataki, S. Mazumdar, P. B. D. Singha, J. Kumari, B. Tiru, and U. Sarma, "Classification of soil types from GPR B scans using deep learning techniques," in *2021 6th IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT 2021)*, August 2021, pp. 840–844.

[16] I. Lunt, S. Hubbard, and Y. Rubin, "Soil moisture content estimation using ground-penetrating radar reflection data," *Journal of hydrology*, vol. 307, no. 1-4, pp. 254–269, 2005.

[17] A. Benedetto, F. Tosti, B. Ortuani, M. Giudici, and M. Mele, "Soil moisture mapping using gpr for pavement applications," in *2013 7th International Workshop on Advanced Ground Penetrating Radar*. IEEE, 2013, pp. 1–5.

[18] S. A. Z. Rahman, K. C. Mitra, and S. M. Islam, “Soil classification using machine learning methods and crop suggestion based on soil series,” in *2018 21st International Conference of Computer and Information Technology (ICCIT)*. IEEE, 2018, pp. 1–4.

[19] I. Senanayake, I.-Y. Yeo, J. Walker, and G. Willgoose, “Estimating catchment scale soil moisture at a high spatial resolution: Integrating remote sensing and machine learning,” *Science of The Total Environment*, vol. 776, p. 145924, 2021.

[20] A. M. Ahmed, R. C. Deo, A. Ghahramani, N. Raj, Q. Feng, Z. Yin, and L. Yang, “Lstm integrated with boruta-random forest optimiser for soil moisture estimation under rcp4. 5 and rcp8. 5 global warming scenarios,” *Stochastic Environmental Research and Risk Assessment*, pp. 1–31, 2021.

[21] M. Malajner, D. Gleich, and P. Planinsic, “Soil type characterization for moisture estimation using machine learning and uwb-time of flight measurements,” *Measurement*, vol. 146, pp. 537–543, 2019.

[22] P. Taneja, H. K. Vasava, P. Daggupati, and A. Biswas, “Multi-algorithm comparison to predict soil organic matter and soil moisture content from cell phone images,” *Geoderma*, vol. 385, p. 114863, 2021.

[23] J. Padarian, B. Minasny, and A. McBratney, “Using deep learning to predict soil properties from regional spectral data,” *Geoderma Regional*, vol. 16, p. e00198, 2019.

[24] C. S. Lee, E. Sohn, J. D. Park, and J.-D. Jang, “Estimation of soil moisture using deep learning based on satellite data: A case study of south korea,” *GIScience & Remote Sensing*, vol. 56, no. 1, pp. 43–67, 2018.

[25] S. Koley and C. Jeganathan, “Estimation and evaluation of high spatial resolution surface soil moisture using multi-sensor multi-resolution approach,” *Geoderma*, vol. 378, p. 114618, 2020.