INTERNATIONAL JOURNAL OF PHARMACY WITH MEDICAL SCIENCES
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3068-2738
Vol.6, No. 1(2026)
www.ijpams.com

ENHANCING PRECISION AGRICULTURE THROUGH MACHINE LEARNING-BASED GPR SOIL MOISTURE ANALYSIS

N. RADHAMMA[1], GANDLA NIKHITHA[2], BENDRAM SURENDHAR[3], DORAPALLY CHARANTEJA[4], GANJI MANOJ[5]

ASSISTANT PROFESSOR[1], UG SCHOLAR[2,3,4&5]

DEPARTMENT OF CSE, NARSIMHA REDDY ENGINEERING COLLEGE (UGC- AUTONOMOUS) MAISAMMAGUDA (V), KOMPALLY, SECUNDERABAD, TELANGANA-500100

ABSTRACT

Precision agriculture plays a crucial role in improving crop productivity and resource management in large-scale farming. Accurate assessment of soil moisture, particularly in the root zone, is essential for efficient irrigation planning and sustainable agricultural practices. Traditional soil moisture monitoring methods often rely on manual sampling or basic sensor-based techniques, which may provide limited spatial coverage and require significant labor and time. This study proposes an advanced approach that integrates Ground Penetrating Radar (GPR) with machine learning techniques to enhance soil moisture analysis in precision agriculture. GPR technology enables non-destructive and large-scale subsurface soil analysis by capturing electromagnetic signal reflections from different soil layers. However, interpreting GPR signals can be challenging due to noise, soil heterogeneity, and environmental factors. To address these challenges, machine learning algorithms are employed to process and analyze GPR signal data, enabling accurate estimation of root zone soil moisture levels. The proposed system includes data preprocessing, feature extraction from GPR signals, and predictive modeling using machine learning methods such as Random Forest, Support Vector Machine, or Neural Networks. These models learn complex relationships between GPR signal characteristics and soil moisture conditions.The integration of machine learning with GPR technology improves the accuracy, efficiency, and scalability of soil moisture assessment in large agricultural fields. The proposed approach supports data-driven irrigation management, optimized water usage, and improved crop health monitoring, contributing to sustainable and intelligent precision agriculture systems.

INTRODUCTION

Agriculture plays a vital role in global food security, and the increasing demand for agricultural productivity has led to the adoption of **precision agriculture technologies**. Precision agriculture focuses on optimizing farming practices through advanced technologies such as sensors, remote sensing, and data analytics. One of the most important factors influencing crop growth and yield is **soil moisture**, particularly within the **root zone**, where plants absorb water and nutrients. Accurate monitoring of soil moisture helps farmers manage irrigation efficiently, conserve water resources, and improve crop productivity. Traditional soil moisture measurement techniques, such as **gravimetric sampling and soil moisture sensors**, are often limited in spatial coverage, labor-intensive, and time-consuming. These methods may not provide a comprehensive understanding of soil moisture distribution across large agricultural fields or mega farms. To overcome these limitations, modern sensing technologies like **Ground Penetrating Radar (GPR)** have emerged as effective tools for non-destructive subsurface analysis. GPR uses electromagnetic waves to detect variations in soil properties, making it possible to estimate soil moisture content within the root zone. However, interpreting GPR signals can be complex due to noise, signal attenuation, and variability in soil composition. To address these challenges, **machine learning techniques** are increasingly being integrated with GPR data analysis. Machine learning algorithms can process large volumes of radar data, identify patterns, and accurately estimate soil moisture levels by learning from historical datasets. This combination enables more reliable and automated soil moisture assessment. By integrating **machine learning models with GPR signal processing**, precision agriculture systems can provide real-time insights into soil moisture distribution across large farming areas. This allows farmers to implement **data-driven irrigation management, reduce water wastage, and improve crop health and yield**. The proposed approach enhances the efficiency of agricultural operations while supporting sustainable farming practices. Therefore, the integration of **machine learning and GPR technology** represents a promising solution for accurate root zone soil moisture monitoring in precision agriculture, particularly in large-scale farming environments.

LITERATURE REVIEW

Precision agriculture focuses on improving crop productivity and resource efficiency by using advanced sensing technologies, data analytics, and automation. Among the key parameters influencing crop growth is soil moisture, which directly affects irrigation scheduling, plant health, and yield. Recent research has increasingly explored the integration of **Ground Penetrating Radar (GPR)** with **Machine Learning (ML)** techniques to estimate soil moisture more accurately and non-destructively. This literature review summarizes significant contributions in this domain.

Importance of Soil Moisture Monitoring in Precision Agriculture

Soil moisture plays a crucial role in plant growth, nutrient transport, and irrigation management. Traditional measurement techniques such as gravimetric and volumetric sampling are labor-intensive, time-consuming, and provide limited spatial coverage. These limitations have motivated the development of automated and remote sensing–based soil moisture estimation methods. Machine learning has emerged as a promising solution because it can analyze large datasets and model complex relationships between environmental variables and soil moisture levels.

Advanced predictive models can support irrigation scheduling, reduce water wastage, and increase crop yield. Studies show that AI-driven irrigation systems can significantly improve agricultural productivity while conserving water resources.

Ground Penetrating Radar (GPR) for Soil Moisture Measurement

Ground Penetrating Radar is a non-destructive geophysical technique that uses electromagnetic waves to detect subsurface features. Variations in soil moisture influence the dielectric properties of soil, which in turn affect the reflected radar signals. This relationship allows GPR data to be used for estimating soil moisture content.

Recent studies demonstrate that GPR provides high-resolution subsurface data that can be processed with machine learning algorithms to extract moisture-related features. Histogram features derived from radar signals have been used as inputs for regression models to predict soil moisture at multiple depths.

Furthermore, GPR systems integrated with other sensing technologies, such as hyperspectral sensors or UAV-based platforms, enhance spatial coverage and enable detailed monitoring of root-zone soil moisture across agricultural fields.

Machine Learning Approaches for Soil Moisture Prediction

Various machine learning algorithms have been applied to soil moisture estimation and prediction. Commonly used models include:

- **Support Vector Machines (SVM)**
- **Random Forest (RF)**
- **Artificial Neural Networks (ANN)**
- **Gradient Boosting**
- **Deep Learning models such as CNN and LSTM**

These algorithms can capture nonlinear relationships between environmental variables and soil moisture levels, leading to more accurate predictions compared to traditional empirical models.

Deep learning models have recently gained attention for their ability to process complex datasets such as radar signals, satellite imagery, and hyperspectral data. For example, convolutional neural networks (CNNs) have been used to extract meaningful features from GPR signals before feeding them into predictive models, improving estimation accuracy.

Integration of GPR and Machine Learning

The combination of GPR with machine learning offers a powerful approach for soil moisture analysis in precision agriculture. In such systems, GPR sensors collect subsurface signal data, which are then processed using ML algorithms to estimate soil moisture distribution across different depths.

Studies have demonstrated that ensemble models such as boosted trees outperform simple regression models in predicting soil moisture from GPR data, achieving higher prediction accuracy and better generalization.

Similarly, hybrid models combining CNN feature extraction with neural networks or gradient boosting have shown superior performance for estimating soil moisture at multiple depths, particularly within the crop root zone.

Challenges in Existing Research

Despite significant progress, several challenges remain in ML-based soil moisture prediction systems:

- **Limited training datasets** for diverse soil types and climatic conditions.
- **Signal noise and environmental interference** affecting GPR measurements.
- **Complex soil heterogeneity**, which makes universal modeling difficult.
- **Integration challenges** between sensor systems, remote sensing data, and predictive algorithms.

Researchers emphasize the need for hybrid frameworks, multi-sensor fusion, and physics-informed machine learning models to improve reliability and scalability.

Research Gap and Future Directions

Recent literature suggests that future research should focus on:

- Integration of **IoT-enabled sensors with GPR systems** for real-time soil monitoring.
- Development of **deep learning models for automatic feature extraction from radar signals**.
- Application of **transfer learning and adaptive ML models** for different soil and climate conditions.
- Combining **satellite, UAV, and ground-based sensing data** for large-scale precision agriculture applications.

These advancements can significantly enhance the accuracy and scalability of soil moisture prediction systems in modern agricultural environments.

SYSTEM ANALYSIS

EXISTING SYSTEM

Traditional soil moisture monitoring systems in precision agriculture mainly rely on conventional sensors and manual data collection techniques. Farmers typically use soil moisture sensors, weather data, or irrigation scheduling methods to estimate the moisture level in the root zone of crops. Although these methods provide useful information, they have several limitations.

Most existing approaches depend on point-based soil moisture sensors, which measure moisture only at a specific location in the field. This results in limited spatial coverage and may not accurately represent the moisture variation across large agricultural lands. In addition, installing multiple sensors across large farms can increase cost and maintenance requirements.

Another commonly used approach is remote sensing using satellite imagery, which provides large-scale monitoring but often lacks sufficient resolution to capture detailed subsurface soil moisture conditions. Environmental factors such as cloud cover and atmospheric interference may also reduce the accuracy of satellite-based observations.

Furthermore, traditional methods usually rely on basic statistical analysis rather than advanced predictive techniques. As a result, they may not effectively capture complex relationships between soil properties, crop conditions, and environmental variables.

INTERNATIONAL JOURNAL OF PHARMACY WITH MEDICAL SCIENCES
Peer Reviewed, Referred & Indexed Journal
E-ISSN:3068-2738
Vol.6, No. 1(2026)
www.ijpams.com

This limits their ability to provide accurate and real-time irrigation recommendations for farmers.

Overall, the existing systems suffer from limitations such as low spatial resolution, limited subsurface insight, high sensor deployment cost, and lack of intelligent predictive capabilities, which restrict their effectiveness in modern precision agriculture.

PROPOSED SYSTEM

The proposed system introduces an advanced precision agriculture framework that integrates Ground Penetrating Radar (GPR) with Machine Learning techniques to analyze soil moisture distribution in the crop root zone more accurately and efficiently.

Ground Penetrating Radar is used to non-invasively scan the soil and collect subsurface electromagnetic signals that reflect different soil characteristics. These signals contain valuable information about soil moisture levels, soil texture, and underground structures. The collected GPR data is then processed and analyzed using machine learning algorithms to identify patterns and estimate soil moisture content.

Machine learning models such as Random Forest, Support Vector Machines, or Neural Networks are employed to interpret complex GPR signal patterns and accurately predict moisture variations across the field. This approach enables high-resolution mapping of soil moisture, which helps farmers understand water distribution within the root zone of crops.

The system can also integrate additional data sources such as weather conditions, soil properties, and crop growth information to improve prediction accuracy. Based on the analyzed data, the system provides intelligent irrigation recommendations, enabling farmers to optimize water usage and enhance crop productivity.

By combining GPR technology with machine learning, the proposed system offers several advantages including non-destructive soil analysis, wider area coverage, higher prediction accuracy, and efficient water management. This solution supports sustainable farming practices and significantly improves decision-making in precision agriculture.

IMPLEMENTATION

MODULES

Data Acquisition Module

- Collects **soil moisture data using Ground Penetrating Radar (GPR)** sensors.
- GPR signals penetrate the soil and capture reflections based on soil dielectric properties.
- Additional environmental parameters such as **temperature, humidity, and soil type** may also be recorded.
- Data is stored for further preprocessing and analysis.

Signal Processing and Preprocessing Module

- Raw GPR signals contain **noise and interference** that must be cleaned.
- Applies techniques such as:
  - Noise filtering
  - Signal normalization
  - Feature extraction from radar reflections
- Converts radar signals into **usable datasets for machine learning models**.

Feature Extraction and Selection Module

- Extracts meaningful characteristics from processed GPR data.
- Features may include:
  - Signal amplitude
  - Reflection time delay
  - Dielectric constant estimation
  - Soil layer characteristics
- Feature selection techniques help identify **the most relevant parameters affecting soil moisture prediction**.

Machine Learning Model Training Module

- Uses machine learning algorithms to learn patterns from extracted features.
- Possible algorithms include:
  - Random Forest
  - Support Vector Machine (SVM)
  - Artificial Neural Networks (ANN)
  - Gradient Boosting models
- The model is trained using **historical soil moisture data and GPR features**.

Soil Moisture Prediction Module

- The trained model predicts **soil moisture levels in the root zone**.
- Helps farmers understand:
  - Soil water availability
  - Irrigation requirements
  - Crop water stress conditions.

Decision Support and Irrigation Recommendation Module

- Converts predictions into **actionable recommendations for farmers**.
- Suggests:

- Optimal irrigation scheduling
- Water usage optimization
- Soil management strategies.

Visualization and Monitoring Module

- Displays results through **graphs, dashboards, and soil moisture maps**.
- Farmers or agricultural managers can monitor:
  - Soil moisture distribution across fields
  - Real-time predictions
  - Historical trends.

8. System Evaluation and Performance Analysis Module

- Evaluates model performance using metrics such as:
  - Accuracy
  - Mean Squared Error (MSE)
  - Root Mean Square Error (RMSE)
  - R^2 score
- Ensures the system provides **reliable soil moisture predictions for precision agriculture applications**.

ALGORITHMS

DECISION TREE CLASSIFIERS

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, …, Ck is as follows:

Step 1. If all the objects in S belong to the same class, for example Ci, the decision tree for S consists of a leaf labeled with this class

Step 2. Otherwise, let T be some test with possible outcomes O1, O2,…, On. Each object in S has one outcome for T so the test partitions S into subsets S1, S2,… Sn where each object in Si has outcome Oi for T. T becomes the root of the decision tree and for each outcome Oi we build a subsidiary decision tree by invoking the same procedure recursively on the set Si.

GRADIENT BOOSTING Gradient boosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees.[1][2] When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest.A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

K-NEAREST NEIGHBORS (KNN)

- Simple, but a very powerful classification algorithm
- Classifies based on a similarity measure
- Non-parametric
- Lazy learning
- Does not “learn” until the test example is given
- Whenever we have a new data to classify, we find its K-nearest neighbors from the training data

Example

- Training dataset consists of k-closest examples in feature space
- Feature space means, space with categorization variables (non-metric variables)
- Learning based on instances, and thus also works lazily because instance close to the input vector for test or prediction may take time to occur in the training dataset

LOGISTIC REGRESSION CLASSIFIERS

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name *logistic regression* is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name *multinomial logistic regression* is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar. Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does. This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

NAÏVE BAYES

The naive bayes approach is a supervised learning method which is based on a simplistic hypothesis: it assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature .Yet, despite this, it appears robust and efficient. Its performance is comparable to other supervised learning techniques. Various reasons have been advanced in the literature. In this tutorial, we highlight an explanation based on the representation bias. The naive bayes classifier is a linear classifier, as well as linear discriminant analysis, logistic regression or linear SVM (support vector machine). The difference lies on the method of estimating the parameters of the classifier (the learning bias).While the Naive Bayes classifier is widely used in the research world, it is not widespread among practitioners which want to obtain usable results. On the one hand, the researchers found especially it is very easy to program and implement it, its parameters are easy to estimate, learning is very fast even on very large databases, its accuracy is reasonably good in comparison to the other approaches. On the other hand, the final users do not obtain a model easy to interpret and deploy, they does not understand the interest of such a technique. Thus, we introduce in a new presentation of the results of the learning process. The classifier is easier to understand, and its deployment is also made easier. In the first part of this tutorial, we present some theoretical aspects of the naive bayes classifier. Then, we implement the approach on a dataset with Tanagra. We compare the obtained results (the parameters of the model) to those obtained with other linear approaches such as the logistic regression, the linear discriminant analysis and the linear SVM. We note that the results are highly consistent. This largely explains the good performance of the method in comparison to others. In the second part, we use various tools on the same dataset (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b and RapidMiner 4.6.0). We try above all to understand the obtained results.

RANDOM FOREST

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for decision trees' habit of overfitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance. The first algorithm for random decision forests was created in 1995 by Tin Kam Ho[1] using the random subspace method, which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg. An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).The extension combines Breiman's "bagging" idea and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance. Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SVM

In classification tasks a discriminant machine learning technique aims at finding, based on an *independent and identically distributed* (*iid*) training dataset, a discriminant function that can correctly predict labels for newly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminant classification function takes a data point x and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminant approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric perspective, learning a classifier is equivalent to finding the equation for a multidimensional surface that best separates the different classes in the feature space. SVM is a discriminant technique, and, because it solves the convex optimization problem analytically, it always returns the same optimal hyperplane parameter—in contrast to *genetic algorithms* (*GAs*) or *perceptrons*, both of which are widely used for classification in machine learning. For perceptrons, solutions are highly dependent on the initialization and termination criteria. For a specific kernel that transforms the data from the input space to the feature space, training returns uniquely defined SVM model parameters for a given training set, whereas the perceptron and GA classifier models are different each time training is initialized. The aim of GAs and perceptrons is only to minimize error during training, which will translate into several hyperplanes' meeting this requirement.

CONCLUSION

This study develops a novel framework that integrates GPR with ML to estimate root-zone soil moisture and subsurface depths, significantly surpassing traditional methods. Our approach, particularly through RF and NN, demonstrates better accuracy in soil moisture and the associated depth predictions by employing a synthetic GPR, innovative feature extraction methods, and dataset enrichment using our real-world collected data. The broader implications of our research are substantial, providing a precise tool for soil moisture mapping that is vital for designing irrigation strategies, conserving water, and promoting soil health. Consequently, this contributes to advancing sustainable agricultural practices and food security. The versatility of our framework across various soil conditions highlights its potential to support sustainable farming and efficient water management globally. Looking ahead, we aim to redesign our intelligent GPR model considering diverse, realistic soil texture, and composition. This will increase the temporal soil moisture assessment accuracy, critical to precision irrigation. Extensive field validations across various agricultural settings will be essential to gauge our framework's real-world applicability and impact. The intelligent method introduced in this article can also be used for soil nutrition management and pest control [64]. In addition, we also aim to explore other environmental information, such as humidity, temperature, soil texture, soil composition, air pressure, and sensory information, such as satellite, IR, and optical imaging, which can be integrated with GPR data through a multimodal ML technique to increase the accuracy of the proposed approach. Accurate soil moisture and depth estimations, combined with meteorological data and satellite imagery, can be applied to multimodal ML models. This integration produces multidisciplinary insights, improving agricultural and subsurface assessments

REFERENCES

[1] N. J. R. Fernandez, J. M. Sabater, P. Richaume,A.A.Yaari, andY.H.Kerr, "SMOS soil moisture retrieval over grasslands: SMOS L2 algorithm or a new neural network?," Geophysical Res. Lett., 2015. [Online]. Available: https://doi.org/10.1002/2015GL063388

[2] D. Zhang and G. Zhou, "Estimation of soil moisture from optical and thermal remote sensing:Areview," Sensors, vol. 16, no. 8, 2016, Art. no. 1308.

[3] H. Namdari, M. Moradikia, D. T. Petkie, R. Askari, and S. Zekavat, "ComprehensiveGPRsignal analysis via descriptive statistics andmachine learning," in Proc. IEEE Int. Conf.Wireless Space Extreme Environ., 2023, pp. 127–132.

[4] V. Filardi et al., "Data-driven soil water content estimation at multiple depths using SFCW GPR," in Proc. IEEE Int. Opportunity Res. Scholars Symp., 2023, pp. 86–90.

[5] FAO, "Coping with water scarcity—An action framework for agriculture and food security," 2012. [Online]. Available: http://www.fao.org/
3/i3015e/i3015e.pdf

[6] J. Wallace, "Increasing agricultural water use efficiency to meet future food production," Agriculture Ecosystems Environ., vol. 82, no. 1-3, pp. 105–119, 2000.

[7] K. C. Kornelsen and P. Coulibaly, "Root-zone soil moisture estimation using data-driven methods," Water Resour. Res., vol. 50, no. 4,
pp. 2946–2962, 2014.

[8] V.Komarov, S.Wang, and J.Tang, "Permittivity andmeasurements," 2005.

[9] I. Rodriguez-Iturbe, P. D'odorico, A. Porporato, and L. Ridolfi, "On the spatial and temporal links between vegetation, climate, and soilmoisture," Water Resour. Res., vol. 35, no. 12, pp. 3709–3722, 1999.

[10] M. Rasol et al., "GPR monitoring for road transport infrastructure: A systematic review and machine learning insights," Construction Building Mater., vol. 324, 2022, Art. no. 126686.

[11] C. Albergel et al., "From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations," Hydrol. Earth Syst. Sci., vol. 12, no. 6, pp. 1323–1337, 2008.

[12] A. M. Peterson,W. D. Helgason, and A.M. Ireson, "Estimating field-scale root zone soil moisture using the cosmic-ray neutron probe," Hydrol. Earth Syst. Sci., vol. 20, no. 4, pp. 1373–1385, 2016.

[13] 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 Proc. Int. Conf. Recent Trends Electron. Inf. Commun. Technol., 2021, pp. 840–844.

[14] N. N. Das and B. P. Mohanty, "Root zone soil moisture assessment using remote sensing and vadose zone modeling," Vadose Zone J., vol. 5, no. 1, pp. 296–307, 2006.

[15] E. G. Njoku and D. Entekhabi, "Soil moisture remote sensing: Stateof- the-science," Proc. IEEE, vol. 85, no. 8, pp. 1349–1375, 1996, doi: 10.1109/5.535743.