WWW.IJRPR.COM

**International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# AI-POWERED SOIL MOISTURE PREDICTION FOR PRECISION IRRIGATION

***Prof. Monali Jadhav[1], Shubham Gadekar[2], Harshada Pansare[2], Sanika Kharge[2]***

[1] Seminar Guide; [2] Students, Department of Artificial Intelligence and Machine Learning, ISBM College of Engineering, Nande, Pune, Maharashtra, India

**ABSTRACT :**

Root-zone soil moisture (RZSM) is an important variable for irrigation planning because it reflects the water available to plant roots. Direct measurement across large agricultural areas can be costly, while satellite observations commonly provide surface moisture rather than a direct view of moisture at different soil depths. This paper presents a proposed AI-powered framework for estimating and predicting RZSM by combining satellite-derived surface soil moisture, local sensor readings, weather variables and relevant soil or crop information. The workflow includes data quality checks, spatial and temporal alignment, feature selection, sequence modelling and probabilistic integration. A dynamic multilayer perceptron (D-MLP) is proposed for selecting informative auxiliary features, followed by a hybrid Transformer with dual attention to represent temporal patterns and relationships between sensor locations. Bayesian averaging is proposed to combine location-level estimates and represent uncertainty. The predictions may be used by an irrigation decision-support module, while local thresholds, pump time limits and manual override remain in place. The paper reviews related research, outlines the methodology and defines an evaluation plan using measures such as MAE, RMSE, Nash-Sutcliffe Efficiency and normalized mean bias error. No implementation results are claimed; model performance and possible water savings require independent testing with documented data.

**Keywords:** root-zone soil moisture; precision irrigation; machine learning; hybrid Transformer; satellite data; feature selection; Bayesian averaging; smart agriculture

## INTRODUCTION

Irrigation decisions depend on the amount of water available in the crop root zone, not only on the moisture measured at the soil surface. Insufficient moisture can limit plant growth, whereas excessive irrigation can waste water and energy and may move nutrients below the active root zone. Soil moisture changes with soil texture, crop development, rainfall, temperature, wind, drainage and irrigation history. A single fixed schedule or an isolated sensor reading may therefore be insufficient for planning irrigation over time and across a field.

Internet of Things (IoT) devices can collect frequent local measurements, while satellite products and weather records provide wider spatial coverage and environmental context. Machine-learning models can learn relationships among these data sources and estimate soil moisture at times or locations where direct measurements are unavailable. The proposed work focuses on root-zone soil moisture estimation and prediction, with irrigation scheduling treated as a downstream decision-support use rather than as a substitute for agronomic judgement.

### *Problem Statement*

Directly measuring root-zone soil moisture at many locations and depths is resource-intensive, while surface satellite measurements alone do not directly describe vertical soil-moisture conditions. A practical prediction framework is needed to combine available satellite, weather and in-situ information, select useful inputs, model temporal and cross-location dependencies, and estimate RZSM with measurable uncertainty. The resulting estimates must be validated against suitable reference observations before being used to guide irrigation.

### *Objectives*

The objectives are to: (1) study relationships between environmental variables and root-zone soil moisture; (2) prepare a data workflow for satellite, weather and in-situ observations; (3) investigate feature selection using a dynamic multilayer perceptron; (4) outline a hybrid Transformer with dual attention for temporal and cross-location relationships; (5) evaluate Bayesian averaging for combining estimates from multiple locations; and (6) define a validation and decision-support process using appropriate prediction and resource-use metrics.

### *Scope and Contribution*

The proposed scope includes data preparation, indirect RZSM estimation, model comparison and an irrigation decision-support stage. Candidate data include satellite-derived surface soil moisture, weather variables, historical soil-moisture observations and available soil or crop descriptors. A local capacitive moisture sensor and an ESP32-based unit may be used for ground checks or a small prototype, but the central estimation approach aims to reduce dependence on dense in-situ sensor networks. The contribution is a clear, testable workflow connecting multi-source data integration with RZSM estimates, uncertainty-aware evaluation and safe irrigation decisions.

## LITERATURE SURVEY

Research on soil-moisture prediction spans IoT-based irrigation, conventional machine learning, time-series forecasting and satellite-assisted root-zone estimation. Goap et al. (2018) presented an IoT-based irrigation management system that combines field sensor readings with weather forecasts and uses predicted moisture to support irrigation scheduling. Their work illustrates how prediction can be linked to an irrigation workflow, although the usefulness of such a system depends on the coverage and quality of the input data. Adeyemi et al. (2018) used dynamic neural networks to predict soil-moisture content from previous moisture, precipitation and climate measurements. Their study also examined predictive irrigation scheduling, indicating the value of modelling moisture change rather than relying only on a current threshold.

For short-term forecasting, Togneri et al. (2022) evaluated data-driven soil-moisture prediction across multiple fields, farms and crop types. They reported strong performance from boosting-based methods and highlighted poor data quality as a practical challenge. Yu et al. (2021) developed a hybrid convolutional neural network and gated recurrent unit (CNN-GRU) model for maize root-zone moisture prediction at different depths. The study combines feature extraction with temporal memory, demonstrating why both recent measurements and weather variables can matter when forecasting root-zone conditions.

Spatial and depth variation is particularly important where dense sensor installations are costly. Chandrappa et al. (2023) proposed spatiotemporal modelling of soil moisture across time and soil depth, using lower-depth observations and environmental data to support estimates at other depths. Carranza et al. (2021) studied Random Forest for RZSM estimation and found that soil and land-cover properties were important to spatial estimation. These works suggest that a model trained only on recent moisture readings may miss relevant differences between locations and soil conditions.

More recent applied studies compare several models on sensor datasets. Challa et al. (2024) evaluated decision trees, Random Forest, multiple linear regression, XGBoost, K-nearest neighbours and artificial neural networks in an IoT-based polyhouse irrigation system; Random Forest and XGBoost performed strongly in their reported test setting. Madhukumar et al. (2024) proposed a hybrid Transformer network for indirect RZSM estimation using downscaled satellite moisture, ground observations, auxiliary variables, feature selection, dual attention and Bayesian averaging. They reported improvements over the sequential models included in their comparison and potential water and cost savings at their selected sites. Such numerical findings belong to those studies and are not results of the present proposed framework. Yadav et al. (2023; available online in 2024) described an IoT and machine-learning workflow for classifying soil-moisture conditions from sensor and environmental inputs, providing another example of the connection between sensing, prediction and irrigation action.

### *Research Gap and Design Rationale*

The reviewed literature demonstrates that sensor-based forecasting, conventional machine learning, temporal networks and satellite-assisted RZSM estimation are established research directions. Accordingly, this paper does not claim that RZSM prediction or Transformer-based modelling is new in itself. The practical design focus is to specify an end-to-end workflow that can be tested with the data available to the project: validate and align inputs, select features, compare the proposed hybrid architecture with transparent baselines, quantify uncertainty, and verify whether improvements remain when testing different periods or locations. A model should not be considered successful solely because it performs well on one dataset; transfer to other fields requires separate validation.

## PROPOSED METHODOLOGY

The proposed workflow contains seven stages: (1) data acquisition; (2) quality checking and calibration; (3) spatial and temporal alignment; (4) feature construction and selection; (5) model training; (6) RZSM estimation and uncertainty aggregation; and (7) validation and irrigation decision support. Data availability will determine the initial scope. Satellite and weather inputs should be included only when their timestamps, locations, units and spatial resolutions can be matched appropriately with the reference moisture observations.

### *Input Data and Feature Representation*

The data record may contain satellite-derived surface soil moisture, in-situ moisture at available depths, precipitation, air temperature, relative humidity, wind speed, soil properties, crop information, location and time. Lagged measurements and rolling summaries can represent recent trends. Table 1 summarizes the main input groups and their intended roles. Missing values, duplicates, implausible readings and sensor faults should be flagged. Data splits should respect time order for forecasting and should include location-based holdouts when the aim is to estimate moisture at unseen sites.

**Table 1. Proposed input data and their roles**

| Input group | Examples | Purpose |
|---|---|---|
| Moisture observations | Satellite surface moisture; in-situ readings at available depths | Target reference, history and spatial context |
| Weather variables | Rainfall, temperature, humidity, wind speed | Represent atmospheric drivers of moisture change |
| Soil and crop data | Soil texture, land cover, crop type or growth stage | Represent differences between locations and plants |
| Spatial and temporal data | Coordinates, observation time, seasonal indicators | Align observations and represent site/time effects |
| Optional device data | Capacitive sensor, ESP32 logger, timestamped readings | Provide local monitoring and ground-validation data |

***Feature Selection and Candidate Models***

A dynamic multilayer perceptron (D-MLP) is proposed to identify informative auxiliary variables from the aligned ground and satellite dataset. The selected features can then be supplied to a hybrid Transformer neural network. Its temporal component models changes over successive observations, while a dual-attention design is intended to represent both dependencies within a location and relationships among candidate sensor locations. A Bayesian averaging stage may combine estimates from multiple locations using their conditional contributions, producing a final RZSM estimate and an uncertainty indicator. These elements are proposed design choices; their benefit must be established through comparison.

The main comparison should include a persistence baseline (the latest observed value), a linear-regression baseline, Random Forest or XGBoost, and an appropriate sequence model such as LSTM or CNN-GRU. The proposed hybrid Transformer should be retained only if the available dataset supports its training and testing without data leakage. Model selection should be based on performance on held-out observations, complexity, interpretability and the intended prediction horizon.

***Proposed System Architecture***

**Satellite surface moisture + weather + in-situ / soil-crop data**
**↓ Quality control, calibration and space-time alignment**
**↓ Feature construction and D-MLP feature selection**
**↓ Hybrid Transformer with temporal and cross-location dual attention**
**↓ Bayesian averaging → RZSM estimate + uncertainty**
**↓ Validation against reference observations → irrigation decision support**

**Figure 1. Conceptual workflow of the proposed RZSM prediction framework. It represents a planned design, not a completed deployment.**

***Irrigation Decision Support***

A predicted RZSM value can be compared with crop- and soil-specific operating limits to flag possible water stress or a need for inspection. Irrigation should not be triggered by an AI prediction alone. Any prototype that operates a pump should retain a local safety controller, sensor-fault checks, maximum run time, water-availability checks and manual override. Basic safe operation should continue during network or cloud-service interruptions. The model is intended to support a decision, not certify crop water needs without context.

***Technology Requirements***

A prototype workflow can be developed in Python using NumPy, Pandas and scikit-learn for data preparation and baseline models, with PyTorch or TensorFlow for neural-network experiments. Jupyter Notebook or Visual Studio Code can support development. Spatial data may require GeoPandas or a similar library; model outputs can be stored in CSV or SQLite and shown in a lightweight dashboard. A capacitive soil-moisture sensor and ESP32 are optional for collecting local reference measurements. Satellite products, weather sources and hardware are to be selected according to availability, resolution, calibration needs and project resources.

## PROTOTYPE EVALUATION PLAN

Evaluation should compare all models on the same held-out observations and report results by forecast horizon and, where possible, by site, season and moisture range. A random split alone may overstate performance when nearby or future records appear in both training and test sets. Tests should document sensor calibration, missing-data treatment, feature construction, model settings and the dates or locations excluded from training. Table 2 lists the proposed measures. No numerical performance or water-saving result is claimed for the current paper.

**Table 2. Proposed evaluation measures**

| Measure | Interpretation | Evaluation purpose |
|---|---|---|
| MAE | Mean absolute difference between predicted and observed RZSM | Measures average prediction error |
| RMSE | Square root of mean squared prediction error | Penalizes larger errors more strongly |
| NSE | $1 - \text{sum}((y - \hat{y})^2) / \text{sum}((y - \text{mean}(y))^2)$ | Compares predictions with the observed-mean baseline |
| NMBE | $100 \times \text{sum}(\hat{y} - y) / \text{sum}(y)$ | Shows the direction and relative size of systematic bias |
| Generalization | Performance on a held-out site or time period | Checks transfer across locations and seasons |
| Resource use | Measured irrigation volume and energy under a defined baseline | Tests water/energy benefit only in a controlled comparison |

For numerical evaluation, let $y$ be the reference moisture value and $\hat{y}$ the prediction for each of $n$ test observations. MAE and RMSE quantify average error; NSE indicates whether the prediction improves on using the observed mean as a reference. NMBE should be reported with the exact convention used and interpreted alongside the data scale. Any water or energy saving claim requires actual measurements and a fair baseline under comparable crop, soil and weather conditions.

## DISCUSSION

The main potential benefit of the framework is to estimate root-zone conditions using complementary data rather than depending only on dense, direct measurements at every depth. Surface satellite products can provide spatial coverage, local sensors can provide frequent readings, and weather variables can help explain changes over time. Feature selection and attention mechanisms may assist the model in identifying useful variables and relationships, while Bayesian averaging provides a way to combine multiple estimates. These benefits are hypotheses to evaluate, not guaranteed outcomes.

The principal challenges are data availability, mismatched spatial and temporal resolutions, calibration differences, gaps in reference observations and shifts between soil types or growing seasons. Satellite measurements may be affected by retrieval limits and should not be treated as direct measurements of the full root zone. Deep-learning models may require substantial data and computation, and a complicated model can underperform simpler methods when the dataset is small. A baseline comparison, uncertainty reporting and independent site validation are therefore central parts of the proposed methodology.

### *Potential Applications*

Potential applications include irrigation planning in precision agriculture, monitoring soil-water conditions in greenhouses or parklands, identifying zones that may require field inspection, and providing estimates where installation of deep sensors is impractical. In a later prototype, RZSM estimates could be combined with crop-specific thresholds and local safety rules to support drip-irrigation scheduling.

### *Limitations and Ethical Considerations*

The framework depends on authorized, dependable data sources and documented preprocessing. Location-linked farm records should be protected and used only for the defined purpose. The model may be less reliable in regions, soils or seasons that are not represented in training data. Uncertainty and missing input data should be visible to users, and the system should not present estimates as guaranteed measurements. Human review and agronomic requirements remain important for irrigation decisions.

## CONCLUSION AND FUTURE WORK

This paper has described a proposed AI-powered workflow for soil-moisture prediction with an emphasis on root-zone estimation for precision irrigation. The design combines satellite-derived surface moisture, in-situ and environmental data, D-MLP-based feature selection, a hybrid Transformer with dual attention, and Bayesian averaging. It connects the resulting estimates to an evaluation plan and a safety-aware irrigation decision-support stage. The literature shows promising results from sensor-based machine learning and more advanced spatiotemporal models, but these findings cannot be assumed to transfer directly to a new dataset or deployment.

Future work should begin with obtaining and documenting aligned satellite, weather and reference moisture data. The proposed model should then be compared with persistence, regression, tree-based and sequence-model baselines using time- and location-aware testing. Calibration, uncertainty, generalization and resource-use outcomes should be recorded. A small sensor-based prototype may later provide ground-validation data and demonstrate safe irrigation alerts. Claims about accuracy, water savings or cost reduction should be made only after the relevant experiments have been completed.

## REFERENCES:

1. A. Goap, D. Sharma, A. K. Shukla, and C. R. Krishna, “An IoT based smart irrigation management system using machine learning and open source technologies,” Computers and Electronics in Agriculture, vol. 155, pp. 41–49, 2018, doi: 10.1016/j.compag.2018.09.040.
2. O. Adeyemi, I. Grove, S. Peets, Y. Domun, and T. Norton, “Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling,” Sensors, vol. 18, no. 10, Art. no. 3408, 2018, doi: 10.3390/s18103408.

1. R. Togneri, D. F. dos Santos, G. Camponogara, H. Nagano, G. Custódio, R. Prati, S. Fernandes, and C. Kamienski, "Soil moisture forecast for smart irrigation: The primetime for machine learning," Expert Systems with Applications, vol. 207, Art. no. 117653, 2022, doi: 10.1016/j.eswa.2022.117653.
2. J. Yu, X. Zhang, L. Xu, J. Dong, and L. Zhangzhong, "A hybrid CNN-GRU model for predicting soil moisture in maize root zone," Agricultural Water Management, vol. 245, Art. no. 106649, 2021, doi: 10.1016/j.agwat.2020.106649.
3. V. Y. Chandrappa, B. Ray, N. Ashwatha, and P. Shrestha, "Spatiotemporal modeling to predict soil moisture for sustainable smart irrigation," Internet of Things, vol. 21, Art. no. 100671, 2023, doi: 10.1016/j.iot.2022.100671.
4. C. Carranza, C. Nolet, M. Pezij, and M. van der Ploeg, "Root zone soil moisture estimation with Random Forest," Journal of Hydrology, vol. 593, Art. no. 125840, 2021, doi: 10.1016/j.jhydrol.2020.125840.
5. L. P. Challa, C. D. Singh, K. V. R. Rao, A. Subeesh, and M. Srilakshmi, "Prediction of soil moisture using machine learning techniques: A case study of an IoT-based irrigation system in a naturally ventilated polyhouse," Irrigation and Drainage, vol. 73, no. 3, pp. 1138–1150, 2024, doi: 10.1002/ird.2933.
6. N. Madhukumar, E. Wang, Y. Everingham, and W. Xiang, "Hybrid Transformer Network for Soil Moisture Estimation in Precision Irrigation," IEEE Access, vol. 12, pp. 48898–48909, 2024, doi: 10.1109/ACCESS.2024.3378257.
7. D. Yadav, S. P. Gangwar, G. Verma, A. K. Mishra, R. K. Srivastava, S. K. Srivastava, M. Chaurasia, and A. K. Srivastava, "IoT and Machine Learning-Based Prediction of Smart Soil Moisture Monitoring and Irrigation System," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 9, pp. 5107–5114, 2023. Available: https://www.ijritcc.org/index.php/ijritcc/article/view/10194.