Integrating the new variant OPTRAM and moisture-related indices for improved soil moisture prediction in Louisiana

Dorcas Twumwaa Gyan[1]*, Yaw A. Twumasi[1], Zhu H. Ning[1], Esi Dadzie[1], Joe Mensah[2], Jeff Dacosta Osei[1], Priscilla M. Loh[1], Kingsford Kobina Annan[1]

[1] Southern University and A&M College, Department of Urban Forestry, Environment, and Natural Resources, Baton Rouge, LA 70813, USA

[2] Mississippi State University, Department of Geosciences, Starkville, MS, 39762, USA

*Corresponding author: dorcas.gyan@sus.edu

Keywords: OPTRAM, Random Forest, Support Vector Machine, moisture-related indices

Abstract

Accurate soil moisture prediction is essential for advancing agricultural productivity, optimizing water resource management, and strengthening climate adaptation strategies, particularly in hydrologically vulnerable regions such as Louisiana. This study investigates the integration of the newly developed OPtical TRApezoid Model (OPTRAM) with conventional moisture-related vegetation indices to enhance soil moisture prediction. A suite of indices, including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Enhanced Vegetation Index (EVI), Structure Insensitive Pigment Index (SIPI), Atmospherically Resistant Vegetation Index (ARVI), and Normalized Difference Moisture Index (NDMI), was compared against OPTRAM to evaluate predictive capacity. Support Vector Machine with Radial Basis Function (SVM-RBF) kernel and Random Forest (RF) algorithms were employed using Sentinel-2 imagery coupled with Louisiana weather station records. Model performance was validated using statistical metrics (R^2, RMSE, and MAE). Results revealed that RF achieved MAE = 0.054, RMSE = 0.069, and R^2 = 0.539, while SVM-RBF achieved MAE = 0.065, RMSE = 0.079, and R^2 = 0.690. OPTRAM demonstrated superior performance with RF, whereas indices such as NDMI, EVI, SIPI, and NDWI outperformed OPTRAM in moisture prediction with SVM-RBF. These findings highlight the importance of integrating advanced optical models with machine learning approaches to improve soil moisture monitoring and support sustainable land management in Louisiana.

1.0 Introduction

Soil moisture is a fundamental variable in the hydrological cycle, tightly coupling water, energy, and carbon fluxes between the land surface and atmosphere (Pradhan et al., 2018; De Queiroz et al., 2020). It drives key ecological and agricultural processes such as evapotranspiration, plant growth, and runoff generation. Variations in soil moisture strongly influence yield stability, drought susceptibility, and water management practices (Wang et al., 2011; Ochsner et al., 2013). Although satellite missions like SMAP and SMOS provide global near-surface soil moisture data, their coarse spatial resolution (25-40 km) limits applicability in fine-scale agricultural and watershed studies (Entekhabi et al., 2010; Ambrosone et al., 2020).

To address this limitation, remote sensing techniques utilizing optical and thermal bands have been employed to infer surface soil moisture at higher spatial resolutions. Among these methods, the Optical Trapezoid Model (OPTRAM) has emerged as a promising approach, as it relates transformed shortwave-infrared reflectance (STR) and a vegetation index (e.g., NDVI) to derive moisture estimates without requiring thermal data (Sadeghi et al., 2017; Babaeian et al., 2019). Its newest variant further extends the capability to map water bodies and refine boundary definitions (Sadeghi et al., 2023). However, despite its promising performance, OPTRAM's accuracy depends strongly on correctly identifying the wet and dry edges and parameterization tailored to local conditions (Alavi et al., 2025). In parallel, traditional vegetation-based moisture indices, such as NDVI, NDWI, EVI, and NDMI, remain widely used owing to simplicity and interpretability (Acharya et al., 2022). Yet, the standalone predictive power of these indices is often limited by the confounding effects

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-M-10-2025-113-2026 | © Author(s) 2026. CC BY 4.0 License.

of soil background, canopy structure, and atmospheric interference.

To overcome these challenges, machine learning techniques like Random Forest (RF) and Support Vector Machine (SVM) offer robustness in modeling non-linear relationships and combining multiple predictors (e.g., indices, climatic variables) (Han et al., 2023; Lakra et al., 2025). By combining OPTRAM-derived variables with moisture-related indices within a learning framework, we hypothesize that soil moisture prediction can be enhanced in both accuracy and spatial detail. In Louisiana's heterogeneous landscapes, accurately estimating soil moisture at field-to-watershed scales remains a challenging task. OPTRAM offers a physically based model, but its local calibration is delicate, and traditional indices alone often fail to capture fine-scale hydrological variability. There is a gap in understanding how best to integrate OPTRAM with vegetation indices and leverage modern machine learning models to optimize soil moisture estimation performance in Louisiana. This study aims to 1) integrate the new OPTRAM variant with key moisture-related vegetation indices (NDVI, NDWI, EVI, SIPI, ARVI, NDMI) and weather station data for soil moisture prediction in Louisiana. 2) To compare the predictive accuracy of Random Forest and Support Vector Machine with Radial Basis Function (SVM-RBF) models through statistical metrics (R^2, RMSE, MAE) and determine the most effective combination of predictors for reliable soil moisture estimation across Louisiana's diverse landscapes.

2.0 Methods

2.1 Study Area

Louisiana's agricultural sector is a cornerstone of the state's economy, with nearly 25,000 farms covering roughly 8 million acres and contributing a gross farm value of approximately $7.40 billion in 2024, while value-added processing activities brought the total economic impact of agriculture, forestry, and fisheries to about $12.95 billion, underscoring its role as a major economic engine and employer in both urban and rural communities across the state (LSU AgCenter, 2025). The state's diverse production base, featuring key commodities such as sugarcane, rice, soybeans, corn, poultry, and seafood, supports local jobs and generates significant downstream economic activity, making agriculture integral not only to GDP but also to cultural identity and community resilience in Louisiana (LSU AgCenter, 2025).

This productive landscape operates within a humid subtropical climate characterized by long growing seasons and abundant rainfall, conditions that support high yields but also expose producers to hydrological extremes. In recent years, Louisiana has experienced both severe flooding and historic droughts; for example, the 2023 drought led to an estimated $290 million in agricultural losses and contributed to one of the state's worst wildfire seasons. Studies on Louisiana's changing climate show emerging trends of drier soils in some periods, more intense rainfall events, and shifting temperature regimes that affect corn, cotton, rice, and soybean yields, highlighting a growing vulnerability of the agricultural sector to water-related stress. These hydroclimatic swings translate directly into soil-moisture deficits or surpluses, with implications for planting decisions, pest and disease pressure, and overall system resilience.

At the same time, Louisiana's agricultural sector faces growing pressure from climate-driven shifts such as drought episodes, prolonged heat, and irregular rainfall patterns, all of which disrupt soil-moisture dynamics essential for healthy crop development. The impacts of recent extremes, including the 2023 drought that heavily affected farmers across the state, underscore the need for more precise and proactive tools to guide agricultural decision-making (Hauswirth, 2025). In this context, selecting effective technologies for soil-moisture assessment becomes increasingly valuable, as reliable predictions help producers anticipate water stress, optimize irrigation practices, and reduce vulnerability to weather fluctuations. The modeling framework developed in this study, combining OPTRAM, vegetation indices, and machine-learning methods, offers a targeted way to improve soil-moisture monitoring and enhance the long-term resilience of Louisiana's agricultural landscape.

2.2 Data Collection and Sources

This study utilized multi-temporal Sentinel-2 Level-2A surface reflectance imagery (COPERNICUS/S2_SR_HARMONIZED) and ground-based soil moisture records from the Monroe Weather Station, Louisiana, part of the U.S. Climate Reference Network (USCRN). Sentinel-2 imagery offers high-resolution multispectral data (10–60 m) across 13 bands, making it suitable for deriving vegetation and moisture-related indices (Drusch et al., 2012). Data were retrieved through Google Earth Engine (GEE), which provides cloud-based access to harmonized datasets and scalable geospatial computation (Gorelick et al., 2017). Satellite data were filtered between January 2019 and March 2025, with a cloud cover threshold of less than 20% to ensure data quality. Concurrent soil moisture measurements from the Monroe station served as ground-truth data for model calibration and validation.

2.3 Data Processing and Index Derivation

The preprocessing workflow was automated in GEE, following standard atmospheric and radiometric correction procedures. Sentinel-2 imagery was aggregated into monthly median composites using temporal aggregation to reduce cloud contamination and noise. Spectral bands (B2–B12) were scaled to surface reflectance using a 0.0001 factor. Six vegetation and moisture-related indices were computed using calibrated bands (Table 1).

Index	Formula	Description
NDVI	(B8 – B4) / (B8 + B4)	Vegetation greenness indicator
NDWI	(B8 – B11) / (B8 + B11)	Sensitive to water content in vegetation
NDMI	(B8 – B11) / (B8 + B11)	Reflects leaf and soil moisture conditions
EVI	2.5 × (B8 – B4) / (B8 + 6×B4 – 7.5×B2 + 1)	Reduces atmospheric influence
SIPI	(B8 – B2) / (B8 + B2)	Measures pigment ratios related to canopy stress
ARVI	(B8 – 2×B4 + B2) / (B8 + 2×B4 + B2)	Resistant to atmospheric scattering

Table 1: Vegetation and Moisture-related indices

Each monthly composite generated a stack of these indices, enabling temporal analysis of vegetation and soil water dynamics. Mean index values were extracted for the Monroe Weather Station’s coordinate and exported as CSV for further modeling in ArcGIS PRO.

2.4 OPTRAM-Based Soil Moisture Retrieval

To complement the vegetation indices, soil moisture was also estimated using the Optical Trapezoid Model (OPTRAM). The new OPTRAM variant, Optimized Trapezoid-Based Water Stress Index (Sadeghi et al., 2017; Babaeian et al., 2019), was implemented entirely in GEE. OPTRAM constructs a reflectance trapezoid using shortwave-infrared transformed reflectance (STR) and a vegetation index (typically NDVI) to estimate soil moisture fraction. In this study, Sentinel-2 data were filtered to the same spatial and temporal extent (2019–2025, <20% cloud cover). A custom GEE script computed STR using bands B8, B11, and B12, and subsequently applied the OPTRAM formulation. Monthly composites (n = 73) were produced, yielding an OPTRAM1B index representing the relative soil moisture condition.

2.5 Machine Learning Modeling and Analysis

The machine learning phase was performed in R (version 4.4) using the caret and DALEX frameworks. Two models, Random Forest (RF) and Support Vector Machine with Radial Basis Function kernel (SVM-RBF), were developed to estimate soil moisture using the vegetation indices and OPTRAM as predictor variables. The dataset was partitioned using stratified random sampling into 80% training and 20% testing subsets. Predictor variables (NDVI, NDWI, NDMI, EVI, SIPI, ARVI, and OPTRAM1B) were standardized using Z-score normalization based on training data statistics to avoid information leakage (Han et al., 2023). Two machine learning algorithms, Random Forest (RF) and Support Vector Machine with Radial Basis Function kernel (SVM-RBF), were implemented to model the relationship between vegetation indices, OPTRAM variables, and soil moisture. The RF model underwent hyperparameter tuning to identify the optimal number of predictors (mtry) using 5-fold cross-validation, with performance assessed through R^2, RMSE, and MAE (Breiman, 2001). Similarly, the SVM-RBF model optimized regularization (C) and kernel width (σ) parameters via grid search, leveraging the radial kernel to capture non-linear relationships while applying 5-fold cross-validation to ensure model robustness and prevent overfitting (Budiman, 2019). Model performance was compared across algorithms using R^2, RMSE, and MAE. Post-modeling, feature importance was analyzed using permutation-based methods in the DALEX package (Biecek, 2018). This enabled interpretation of predictor contributions (e.g., NDMI vs. OPTRAM) to overall model performance.

3.0 Results and Discussion

The comparative analysis of feature importance from the Random Forest (RF) and Support Vector Machine with Radial Basis Function (SVM-RBF) models demonstrates the varying sensitivity of machine learning algorithms to optical and vegetation-based predictors of soil moisture. In the RF model, OPTRAM1B_Soil_Moisture emerged as the dominant predictor, indicating that the newly developed OPTRAM variant effectively captures surface soil moisture variability across the Monroe Weather Station. This aligns with findings by Sadeghi et al. (2017) and Babaeian et al. (2019), who demonstrated that OPTRAM’s trapezoid structure efficiently relates optical reflectance to soil water content without relying on thermal bands. The relatively high contribution of NDWI and NDMI further suggests that water-sensitive shortwave infrared bands remain critical in representing the wetness gradient in vegetated areas, particularly under Louisiana’s humid subtropical conditions.

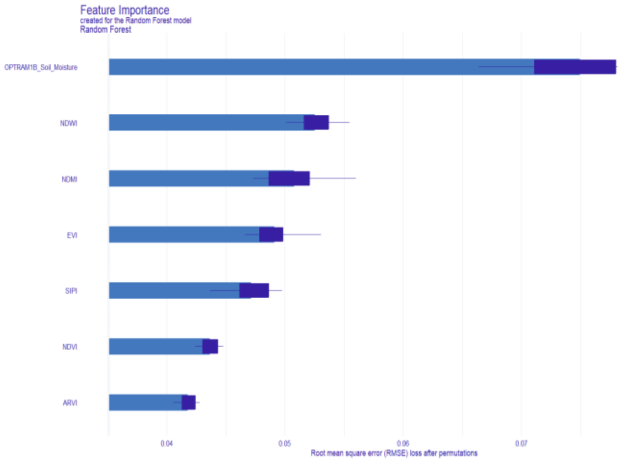

Figure 1: Random Forest

In contrast, the SVM-RBF model exhibited a different feature ranking, with NDMI and EVI contributing most significantly to soil moisture prediction, while OPTRAM1B played a comparatively lesser role (**Figure 2**). These discrepancies explain the fundamental difference in model mechanics. SVM-RBF prioritizes non-linear separability and kernel-based distance measures, which may have accentuated the influence of vegetation indices sensitive to canopy structure and chlorophyll concentration (Han et al., 2023). The prominence of NDMI corroborates studies by Wilson and Sader (2002) and Acharya et al. (2022), which identified NDMI as a robust proxy for vegetation water content and root-zone moisture. The higher predictive strength of EVI within the SVM framework also reflects its resistance to atmospheric effects and its ability to capture dense canopy dynamics, common in Louisiana's mixed cropland-forest mosaic.

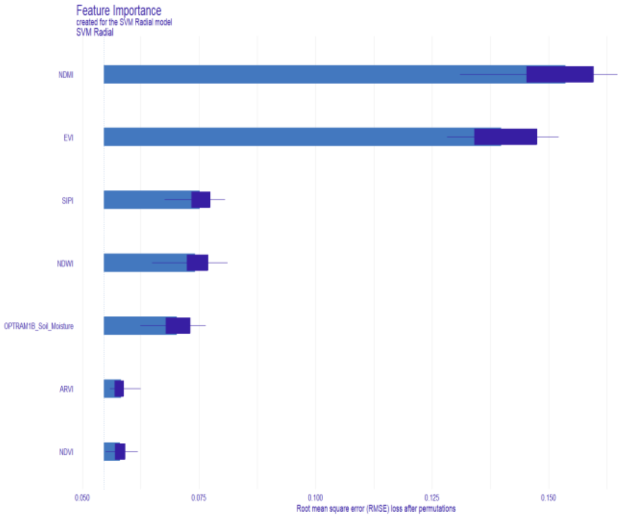

Figure 2: Support Vector Model (SVM) Radial Model

The comparative model evaluation highlights distinct predictive behaviors between the Random Forest (RF) and Support Vector Machine with Radial Basis Function kernel (SVM-RBF) models. Based on the validation results, the RF model achieved an RMSE of 0.069, MAE of 0.054, and R^2 of 0.539, whereas the SVM-RBF model attained slightly higher error metrics (RMSE = 0.079; MAE = 0.065) but a stronger goodness of fit ($R^2 = 0.690$). The higher R^2 value of SVM-RBF indicates a better ability to explain variance in soil moisture, suggesting that the non-linear mapping of the radial kernel effectively captured complex relationships between vegetation indices and soil moisture. Conversely, the lower RMSE and MAE values in RF imply that ensemble averaging in tree-based models produced more stable and less error-prone predictions, consistent with Breiman's (2001) assertion that RF performs well under noisy data conditions due to its bootstrap aggregation mechanism.

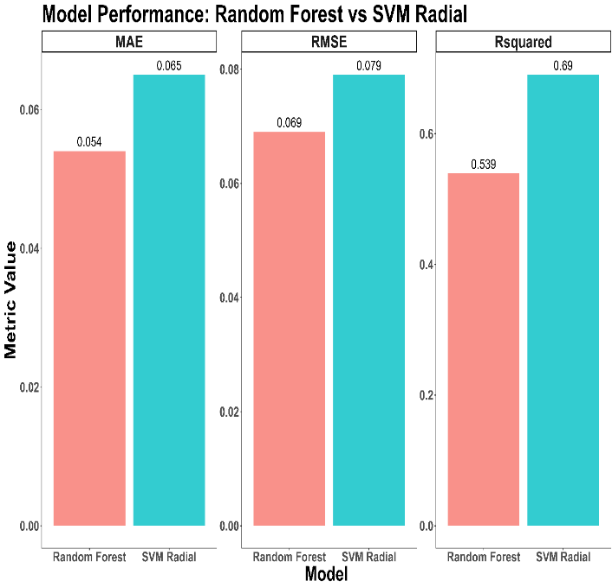

Figure 3: Model performance: SVM versus RF

The scatter plots of predicted versus observed soil moisture further confirm these distinctions (**Figure 4&5**). The SVM-RBF model demonstrated tighter clustering of data points along the 1:1 line at higher soil moisture ranges, reflecting its strength in modeling non-linear responses and localized fluctuations in canopy reflectance (Han et al., 2023). This superior fit suggests that soil moisture variability at finer spatial scales is more effectively modeled through non-linear transformations rather than the ensemble averaging used in RF. However, the RF model's lower RMSE suggests that it may generalize better across moderate soil moisture levels, particularly where vegetation indices such as NDMI, NDWI, and EVI respond linearly to soil wetness gradients (Wilson & Sader, 2002; Acharya et al., 2022). This balance between generalization (RF) and precision (SVM-RBF) reinforces the complementary nature of both models in environmental prediction tasks.

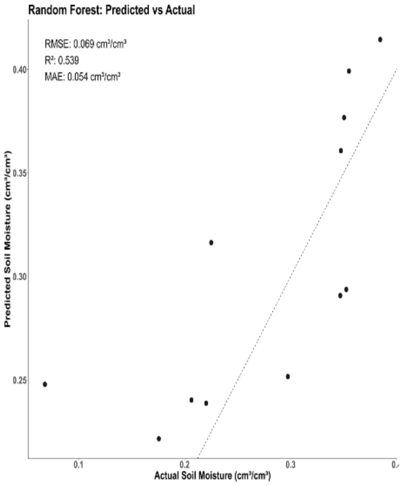

Figure 4: Random Forest: Predicted vs Actual

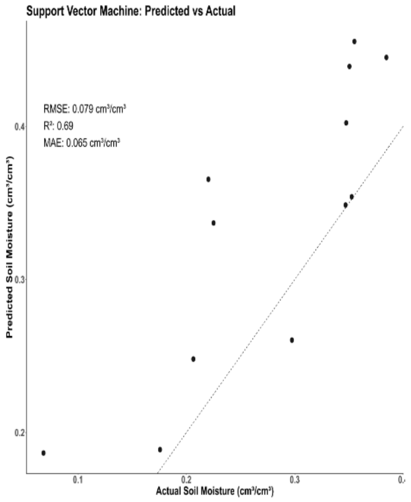

Figure 5: SVM: Predicted vs Actual

From a practical perspective, the results reveal the utility of hybrid modeling approaches in soil moisture prediction. The RF model's stronger error minimization complements the SVM-RBF's superior explanatory power, supporting prior findings that ensemble and kernel-based methods each offer unique advantages in biophysical modeling (Babaeian et al., 2019; Lakra et al., 2025). The integration of OPTRAM and vegetation indices enhanced both models' performance, confirming that combining physically based and spectral indicators yields a more comprehensive representation of surface moisture dynamics (Sadeghi et al., 2017). Overall, the SVM-RBF model's higher R² suggests greater predictive sensitivity, while RF's lower RMSE points to its reliability under varied landscape conditions, indicating that the optimal model choice may depend on whether accuracy or generalization is prioritized in regional soil moisture monitoring applications.

4.0 Conclusion

This study demonstrated the potential of integrating the new OPTRAM variant with moisture-related vegetation indices to improve soil moisture prediction across Louisiana's heterogeneous landscapes. Between the two models applied, the Random Forest (RF) approach emerged as the optimum model based on its lower Root Mean Square Error (RMSE = 0.069), indicating superior predictive stability and reduced residual variation. RMSE served as the decisive criterion because it provides a direct measure of how closely predicted values align with observed soil moisture levels. Lower values signifying more accurate predictions. While the Support Vector Machine with Radial Basis Function kernel (SVM-RBF) achieved a higher coefficient of determination (R² = 0.690), its slightly higher RMSE suggested over-sensitivity to localized fluctuations and potential overfitting in some instances. The RF model's ensemble learning structure allowed it to generalize better across varying soil and vegetation conditions, making it more reliable for operational soil moisture estimation and regional-scale monitoring. Overall, the findings highlight that integrating OPTRAM with vegetation indices within robust ensemble frameworks can significantly enhance soil moisture modeling accuracy and resilience under real-world environmental variability.

ACKNOWLEDGEMENT

The authors would like to acknowledge the USDA National Institute of Food and Agriculture (NIFA) McIntire Stennis Forestry Research Program-funded project with award number NI25MSCFRXXXG033 for providing financial support through Graduate Assistantships. Sincere appreciation also goes to the American Society for Photogrammetry and Remote Sensing (ASPRS) – the Imaging and Geospatial Information Society for the student grant and award to participate in the Fall 2025 virtual ASPRS International Technical Symposium

REFERENCES

Acharya, T. D., Subedi, A., & Lee, D. H. (2022). Evaluation of remote sensing-based vegetation indices for estimating soil moisture. *Remote Sensing Applications: Society and Environment, 27*, 100783. https://doi.org/10.1016/j.rsase.2022.100783

Ambrosone, M., Romano, F., Chirico, G. B., & Preti, F. (2020). Downscaling SMAP soil moisture over complex terrain using Sentinel-1 data and machine learning. *Remote Sensing, 12*(3), 448. https://doi.org/10.3390/rs12030448

Alavi, M., Nouraki, A., Homayouni, S., Albaji, M., Golabi, M., Naseri, A. A., & Célicourt, P. (2025). High-Resolution Crop Evapotranspiration Estimation Using the Automated OPTRAM-ETc Method. *Earth Systems and Environment*, 1-23.

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., & Tuller, M. (2019). Ground, proximal, and satellite remote sensing of soil moisture. *Reviews of Geophysics, 57*(2), 530–616. https://doi.org/10.1029/2018RG000618

Biecek, P. (2018). DALEX: Explainers for complex predictive models in R. *Journal of Machine Learning Research*, *19*(84), 1-5.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Budiman, F. (2019). SVM-RBF parameters testing optimization using cross validation and grid search to improve multiclass classification. *Научная визуализация*, *11*(1), 80-90.

De Queiroz, M. G., Silva, B. B., de Carvalho, F. C., & Gomes, H. B. (2020). Soil moisture estimation using remote sensing data: A review of methods and applications. *Environmental Earth Sciences, 79*, 1–18. https://doi.org/10.1007/s12665-020-09177-0

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., ... & Bargellini, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment, 120*, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., … Van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) mission. *Proceedings of the IEEE, 98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Han, Y., Zhao, W., & Ma, Y. (2023). Machine learning models for soil moisture estimation using multisource data. *Environmental Modelling & Software, 161*, 105573. https://doi.org/10.1016/j.envsoft.2023.105573

Huete, A. R., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*(1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Kaufman, Y. J., & Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing, 30*(2), 261–270. https://doi.org/10.1109/36.134076

Lakra, P., Dutta, S., & Kumar, R. (2025). Comparative assessment of machine learning models for soil moisture estimation using remote sensing indices. *Geocarto International, 40*(5), 1267–1281. https://doi.org/10.1080/10106049.2023.2234527

LSU AgCenter. (2025). *Louisiana summary: agriculture & natural resources*. Retrieved from https://www.lsuagcenter.com/articles/page1751297571587

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., … Zreda, M. (2013). State of the art in large-scale soil moisture monitoring. *Vadose Zone Journal, 12*(4), 1–14. https://doi.org/10.2136/vzj2013.03.0093

Penuelas, J., Filella, I., & Gamon, J. A. (1995). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist, 131*(3), 291–296. https://doi.org/10.1111/j.1469-8137.1995.tb03064

Pradhan, S. N., Anjum, M., & Jena, P. (2018). Estimation of soil moisture content by remote sensing methods: A review. *Journal of Pharmacognosy and Phytochemistry*, *7*, 1786-1792.

Sadeghi, M., Babaeian, E., Tuller, M., & Jones, S. B. (2017). The optical trapezoid model (OPTRAM) for soil moisture estimation: A novel approach for operational applications. *Remote Sensing of Environment, 198*, 52–65. https://doi.org/10.1016/j.rse.2017.05.040

Sadeghi, M., Babaeian, E., Jones, S. B., & Tuller, M. (2023). Enhancing OPTRAM soil moisture retrieval using dynamic boundary calibration and multi-temporal optical data. *Remote Sensing of Environment, 289*, 113422. https://doi.org/10.1016/j.rse.2023.113422

US Drought Monitor. (2025). *Drought conditions in Louisiana*. Retrieved from https://droughtmonitor.unl.edu/

Wang, L., Qu, J. J., Hao, X., & Zhu, Q. (2011). Estimating soil moisture from space: A review of optical, thermal, and microwave remote sensing techniques. *Sensors, 11*(3), 2809–2835. https://doi.org/10.3390/s110302809

Wilson, E. H., & Sader, S. A. (2002). Detection of forest harvest type using multiple dates of Landsat TM data. *Remote Sensing of Environment, 80*(3), 385–396. https://doi.org/10.1016/S0034-4257(01)00318-2