# Soil Moisture Prediction and Benchmark Reliability: A Review with Multi-Site SMAP–ISMN Analysis

Kamrul Hasan, Cliff Mutegi, Arnold Muiruri, Puneeth Kumar Amudala, and Andrew Njiyo
College of Computing
Grand Valley State University
Allendale, MI, USA
e-mail: {hasanka | mutegic | muiruria | amudalap | njiyoa}@mail.gvsu.edu

***Abstract*—Unsustainable irrigation practices are accelerating groundwater depletion worldwide, making accurate soil moisture monitoring increasingly important for agricultural and water resource management. However, a fundamental challenge remains: the benchmark datasets used to validate soil moisture models often contain uncertainties that are not fully understood. This paper addresses this gap through two complementary contributions. First, we present a structured review of soil moisture prediction approaches, including physics-based land surface models, machine learning methods, and hybrid frameworks. We also provide a critical assessment of commonly used benchmark datasets, including ISMN, SCAN, FLUXNET, and satellite-derived products such as SMAP, SMOS, and ASCAT, together with recent advances in data fusion techniques. Second, we present an empirical case study comparing NASA SMAP satellite estimates with ISMN ground observations at two sites in the United States from January 2023 to January 2025. The results reveal strong site-dependent differences in performance. The Michigan site exhibits reasonable agreement for root-zone soil moisture estimates ($r = 0.63$, RMSE = $0.02\ \mathrm{m}^3/\mathrm{m}^3$), whereas the Indiana site shows near-zero correlation and errors more than five times larger (RMSE = $0.11\ \mathrm{m}^3/\mathrm{m}^3$). Moreover, SMAP retrievals are largely insensitive to the dynamic range of in-situ variability, a finding with important implications for model validation practices. Residual analysis further reveals structured, moisture-dependent bias patterns at both sites, suggesting that environmental factors such as row-crop agriculture, tile drainage systems, and sub-pixel heterogeneity are not adequately captured by current retrieval algorithms. Overall, these findings provide strong evidence against single-benchmark validation approaches and support the adoption of integrated, multi-source evaluation frameworks. We conclude by discussing the implications for precision agriculture, downscaling algorithm development, and the design of future soil moisture monitoring networks.***

***Keywords-soil moisture prediction; remote sensing; international soil moisture network; data fusion; machine learning.***

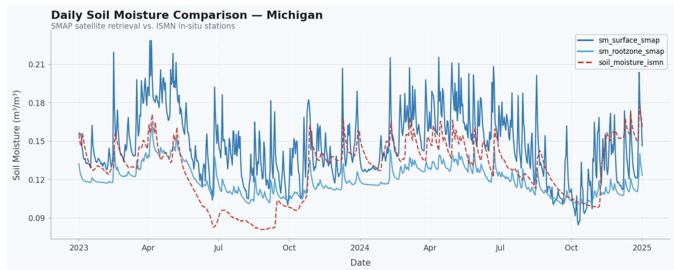


Figure 1. Average daily soil moisture by SMAP (surface-level and root-zone) vs. ISMN at Gaylord-9-SSW (Michigan, U.S.). Null values were imputed through forward-fill (rolling average, window-size $= 3$).

## I. Introduction

Groundwater levels are declining at an alarming rate across the globe due to various factors, with excessive irrigation practices being one of the primary ones [1][2][3]. According to the 2018 U.S. Census of Agriculture, approximately 50% of irrigated land in the United States depends exclusively on groundwater, while an additional 16% relies on a combination of groundwater and surface water [4]. Moreover, groundwater levels have declined substantially across many U.S. aquifers since 1980, highlighting increasingly unsustainable patterns of water use [5].

To address this growing crisis, it is imperative to optimize agricultural water consumption. An ongoing research project at Grand Valley State University (GVSU), Michigan, conducted under the Precision Agriculture Research Lab, aims to address this challenge. The focus of the project is on predicting soil moisture by integrating data from sparsely distributed in-situ moisture sensors with satellite-based observations, such as NASA's Soil Moisture Active Passive (SMAP) mission [6] and the European Space Agency's Climate Change Initiative (ESA CCI) [7]. An initial study from this project [1] introduced the problem of benchmark dataset reliability in soil moisture prediction and provided a preliminary review of key datasets and their limitations. The present paper substantially extends that earlier work in three respects: it broadens the review to encompass recent advances in machine learning, deep learning, and data fusion techniques; it introduces a formal research framework situating these contributions; and, most critically, it presents an original empirical case study comparing NASA SMAP satellite estimates against ISMN ground observations at two contrasting U.S. sites, providing quantitative evidence for the benchmark reliability concerns raised in [1].

Soil moisture monitoring and prediction can play a pivotal role not only in minimizing water waste but also in enabling informed decision-making for farmers and policy makers. Effective soil moisture management supports long-term soil health, prevents erosion, and ensures sustained agricultural productivity. In addition to precision agriculture, it enables better drought monitoring, flood forecasting, and land-atmosphere interaction modeling [8][9]. Although soil moisture prediction has been widely investigated, the development of consistent and reliable benchmark datasets remains an ongoing challenge.

Figure 1 presents the aggregated daily average soil moisture

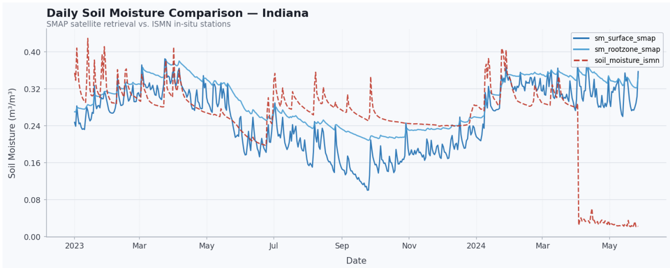


Figure 2. Average daily soil moisture by SMAP (surface-level and root-zone) vs. ISMN at Bedford-5-WNW (Indiana, U.S.). Null values were imputed through forward-fill (rolling average, window-size = 3).

measurements from January 2023 to January 2025 at the Gaylord-9-SSW station in Michigan, USA, an example site within the International Soil Moisture Network (ISMN), a publicly accessible global database that consolidates in-situ soil moisture observations from numerous monitoring networks. By offering standardized data formats and automated quality control protocols, the ISMN serves a vital role in validating satellite-derived soil moisture products and land surface models, and has become a widely adopted reference in hydrological and climate research due to its comprehensiveness and accessibility [10].

To assess the consistency between ground-based and satellite-derived soil moisture measurements, we compare average daily values from NASA's SMAP products with corresponding ISMN data. As illustrated in Figure 1, both datasets exhibit similar seasonal trends, with the primary differences occurring in magnitude rather than overall pattern. A comparable analysis at a second ISMN site, Bedford-5-WNW in Indiana, is shown in Figure 2. In this case, the discrepancy between SMAP and ISMN measurements is more pronounced than at the Gaylord-9-SSW station. These findings raise important questions regarding the reliability of these data as a benchmark for soil moisture modeling: *To what extent can ISMN be trusted for model evaluation? What are its inherent strengths and limitations? And are there viable alternatives that offer improved consistency or coverage?*

This paper addresses these questions through the following contributions, which together extend the preliminary investigation in [1]:

- A structured review of recent advances and key challenges in soil moisture prediction, spanning physics-based models, machine learning, deep learning, and hybrid frameworks.
- A critical evaluation of existing benchmark datasets for soil moisture monitoring and prediction, including ISMN, SCAN, FLUXNET, and satellite-derived products, with particular emphasis on their strengths and limitations in supporting robust model development.
- An original empirical case study comparing SMAP satellite estimates against ISMN observations at two contrasting U.S. sites, providing quantitative, site-specific evidence of the benchmark reliability concerns identified in [1] and motivating the need for multi-source validation frameworks.

The remainder of this paper is organized as follows. Section II presents the research framework guiding the study. Section III discusses recent advances and key challenges in soil moisture prediction. Section IV provides a review of ISMN data, highlighting its strengths, limitations, and applications. Section V examines alternative soil moisture datasets and data fusion approaches. Section VI presents a comparative analysis of SMAP and ISMN soil moisture data for two representative sites introduced in Section I: Gaylord-9-SSW (Michigan, U.S.) and Bedford-5-WNW (Indiana, U.S.). Finally, Section VII summarizes the key findings and provides concluding remarks.

## II. Research Framework

To situate the contributions of this paper within a coherent analytical structure, Figure 3 presents an overview of the research framework guiding this work. The framework comprises four interconnected components: a critical review of soil moisture prediction approaches and benchmark datasets, an empirical comparative analysis of SMAP and ISMN observations, a rigorous statistical evaluation of agreement and bias structure, and a synthesis of findings into actionable recommendations for multi-source validation practice. It is worth noting that these components are not strictly sequential; the review informed site selection for the case study, while the empirical results in turn strengthen and substantiate the review's critique of ISMN as a universal benchmark.

### *Component 1: Literature Review*

The first component systematically examines existing soil moisture prediction methodologies across three categories: physics-based land surface models (LSMs) such as the Noah LSM and Community Land Model; machine learning and deep learning approaches including random forests, support vector machines, convolutional neural networks (CNNs), and recurrent neural networks (RNNs); and hybrid frameworks that integrate physical modeling with data-driven learning. Alongside these prediction approaches, this component critically evaluates benchmark datasets used for model training and validation, including ground-based repositories (ISMN, SCAN, FLUXNET), satellite-derived products (SMAP, SMOS, ASCAT), and regional networks such as OzNet and REMEDHUS. Recent advances in data fusion techniques that integrate satellite and in-situ observations are also reviewed to establish the conceptual context within which the empirical case study is situated.

### *Component 2: Empirical Case Study*

The second component presents an original comparative analysis drawing on two years of daily soil moisture observations from January 2023 to January 2025 at two contrasting sites in the United States. The sites were deliberately selected to represent opposing land surface conditions: Gaylord-9-SSW in Michigan, characterized by a comparatively stable, naturally vegetated temperate landscape, and Bedford-5-WNW in Indiana, dominated by intensive row-crop agriculture, artificial tile drainage infrastructure, and spatially variable soil

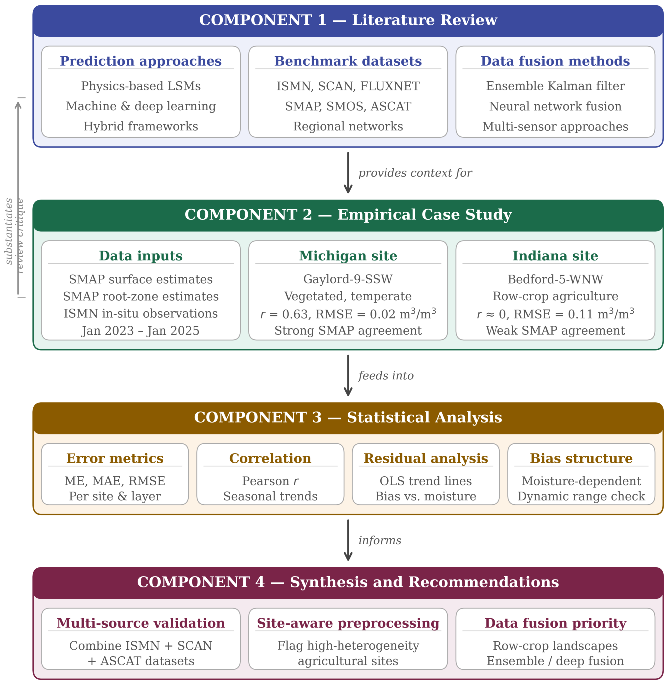


Figure 3. Research framework illustrating the four interconnected components of this study: literature review, empirical case study, statistical analysis, and synthesis of recommendations. Arrows indicate the primary direction of analytical flow; note that Components 1 and 2 also share a feedback relationship, as the review informed site selection and the empirical results substantiate the review's critique of ISMN benchmark reliability.

organic matter. For each site, NASA SMAP satellite-derived estimates at both surface and root-zone layers are compared against corresponding ISMN in-situ measurements, enabling a controlled contrast between a favorable and a challenging retrieval environment.

*Component 3: Statistical Analysis*

The third component applies a suite of statistical metrics to quantify the agreement between SMAP and ISMN observations across both sites and soil layers. Metrics include mean error (ME), mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation coefficient ($r$). Critically, the analysis extends beyond aggregate summary statistics to include ordinary least squares (OLS) residual analysis, which examines bias as a function of in-situ moisture magnitude rather than treating error as uniformly distributed. This approach reveals moisture-dependent bias structures that aggregate metrics alone would obscure, and provides the empirical basis for distinguishing systematic retrieval limitations from random measurement noise.

*Component 4: Synthesis and Recommendations*

The fourth component integrates insights from the review and the case study to evaluate the reliability of ISMN as a universal benchmark and to propose concrete, actionable guidance for practitioners. Specific recommendations address three areas: the adoption of multi-source validation frameworks that combine ISMN with complementary datasets such as SCAN and ASCAT; the application of site-aware preprocessing protocols that flag high-heterogeneity agricultural environments for additional scrutiny; and the prioritization of data fusion approaches for row-crop dominated landscapes where passive microwave retrievals alone are insufficient. These recommendations are grounded directly in the empirical findings of Component 3 rather than derived from general principles alone.

Table I
MAJOR SOIL MOISTURE PREDICTION APPROACHES AND THEIR CHARACTERISTICS.

| Method | Strengths | Limitations | Refs. |
|---|---|---|---|
| Physics-based LSMs | Physically interpretable; models water and energy fluxes; widely used | Parameter uncertainty; scale mismatch; computational cost | [11]–[13] |
| Traditional ML | Captures nonlinear relationships; flexible data fusion; good predictive accuracy | Large training data requirements; limited extrapolation; low interpretability | [14], [15] |
| Deep Learning | Learns complex spatial and temporal patterns; highly scalable | Data- and computation-intensive; black-box behavior | [16] |
| Hybrid Physics–ML | Combines physical and data-driven models; improved generalization | Complex integration; limited evaluation frameworks | [17], [18] |
| Satellite-based Estimation | Global coverage; long-term observations; large-scale monitoring | Coarse resolution; vegetation and atmospheric effects | [19]–[21] |

## III. Soil Moisture Prediction: Approaches, Advances, and Open Challenges

Soil moisture monitoring and prediction research has evolved significantly over the past two decades, primarily driven by advances in remote sensing, data assimilation, and machine learning. Traditional approaches generally relied on physics-based land surface models (LSMs), such as the Noah LSM and the Community Land Model (CLM), to simulate water and energy fluxes at the land-atmosphere interface [11][12]. These models use meteorological inputs and land surface parameters, but their performance is often constrained by uncertainties in input data, parameterization, and the scale mismatches between model outputs and observational datasets [13].

Machine Learning (ML) and Deep Learning (DL) methods have recently emerged as powerful alternatives or complements to traditional models. Data-driven algorithms, including random forests, support vector machines, and artificial neural networks, have been employed to estimate soil moisture from remote sensing and meteorological data [14][15]. Deep learning architectures, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated strong capabilities in modeling complex spatiotemporal patterns in soil moisture dynamics [16]. Additionally, hybrid approaches that integrate physical modeling with ML have gained attention for improving generalizability and interpretability [17][18]. To illustrate these evolutions, Table I compares the key approaches.

Satellite missions such as NASA's SMAP, ESA's Soil Moisture and Ocean Salinity (SMOS), and the AMSR series have facilitated the development of predictive models at multiple spatial scales, contributing to applications from global hydrological assessment to localized precision farming [19][20]. However, most satellite-derived products are available at coarse spatial resolutions (e.g., 1 km or greater), limiting their usefulness in field-level agricultural decision-making [21].

Despite these technological advancements, several key challenges hinder the development of accurate and reliable soil moisture prediction models. A major issue is the scarcity and spatial sparsity of high-quality ground truth data, which is critical for both model training and validation [22]. The heterogeneity of environmental variables, such as soil properties, vegetation cover, land use patterns, and topography, further complicates model generalization across different regions [23]. Equally critical are the challenges associated with data collection. In-situ soil moisture measurements, such as those provided by ISMN, offer valuable ground truth but are often spatially sparse and unevenly distributed, particularly in under-monitored regions [24]. Variations in sensor type, calibration, and installation practices introduce inconsistencies, while sensor failure or communication issues can lead to temporal gaps. Satellite-based data, while offering broader coverage, are impacted by cloud cover, vegetation, and surface roughness, reducing measurement reliability in many settings [25][26]. Arid and semi-arid regions, where accurate soil moisture monitoring is most crucial, are particularly affected due to low signal-to-noise ratios [27].

Addressing these multifaceted challenges calls for multi-disciplinary strategies involving improved sensor networks, data harmonization, uncertainty quantification, and interpretable modeling frameworks. The integration of adaptive machine learning algorithms with heterogeneous data sources is critical to developing high-resolution, accurate soil moisture predictions that can transform sustainable water resource management and data-driven agriculture.

## IV. The International Soil Moisture Network: Capabilities, Limitations, and Applications

The ISMN has emerged as a critical resource for collecting and harmonizing in-situ soil moisture data across global observation networks. It serves as a foundational resource for validating, calibrating, and benchmarking satellite- and model-derived soil moisture datasets. Its importance lies in the harmonized collection and open dissemination of in-situ soil moisture data from a wide array of monitoring networks across different climate zones, land cover types, and soil structures [10][24]. The ISMN enables intercomparison of remote sensing products (e.g., SMAP, SMOS, AMSR2) by providing a global standard against which these data sources can be evaluated [22]. It also supports the assessment and development of downscaling algorithms and machine learning models by offering high-quality ground truth measurements [28]. Moreover, the temporal consistency and metadata richness of ISMN facilitate long-term hydrological studies and trend detection, which are crucial for climate resilience planning and agricultural decision-making. By improving the accuracy and robustness of predictive models, ISMN plays a critical role in the advancement of soil moisture science and its practical applications in water resource management, agriculture, and disaster mitigation.

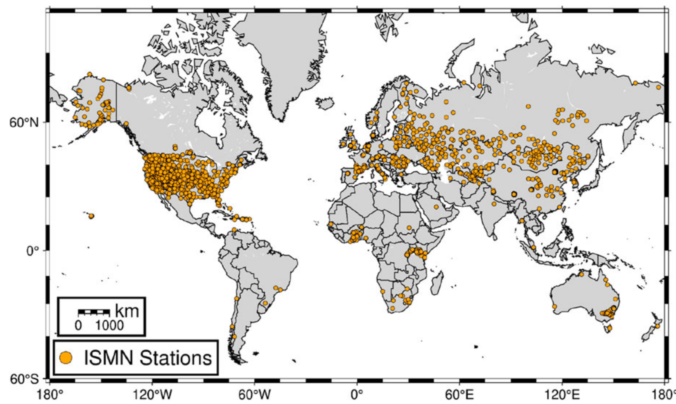


Figure 4. Worldwide distribution of ISMN Stations – an extract from [29].

ISMN aggregates soil moisture measurements from a variety of sources, standardizing and applying quality control procedures to improve accessibility and usability [10]. However, the ISMN data can still exhibit inconsistencies due to differences in sensor types, installation depths, and environmental heterogeneity [22]. Draper et al. emphasized the importance of preprocessing ISMN data before using it for validation or modeling tasks [30].

Despite its value, ISMN presents several challenges when used in predictive soil moisture modeling. The spatial distribution of ISMN stations, as shown in Figure 4, is highly uneven, with denser coverage in North America and Europe and sparse representation in Africa, Asia, and South America. This limits global-scale modeling and regional calibration, especially in underrepresented ecosystems. Station metadata, including soil depth, vegetation, and land use, is sometimes incomplete or inconsistent, complicating efforts to standardize data inputs for machine learning and physical models [22]. Discrepancies also arise from heterogeneity in sensor types, calibration protocols, and measurement depths across networks, introducing uncertainty into inter-station comparisons and satellite validation studies [24]. Moreover, data gaps due to sensor maintenance or environmental interference pose problems for time series continuity. These limitations necessitate pre-processing steps such as harmonization, gap-filling, and filtering, which introduce additional complexity into model development pipelines. Despite these challenges, ISMN remains a critical benchmark for validating satellite retrievals and downscaling methods, though its shortcomings highlight the importance of complementing it with other data sources and standardization frameworks.

The increasing availability of ISMN data has enabled its integration into machine learning and deep learning models for high-resolution soil moisture estimation. Xu et al. [31] used ISMN data to validate a wide and deep neural network that improved the spatial resolution of SMAP satellite data across the U.S. Similarly, Celik et al. [32] and Lee et al. [33] developed deep learning models incorporating ISMN observations to improve performance in heterogeneous landscapes by reducing dependency on physical modeling assumptions. In the agricultural domain, Custódio and Prati [34] applied ensemble machine learning models to IoT-supported irrigation systems, using soil moisture as a key variable. Their results, validated with real-time field data, support the use of AI for operational water resource management.

While the ISMN is the most prominent repository for in-situ soil moisture measurements, several alternative datasets and platforms also play crucial roles in soil moisture research and modeling. One key alternative is the USDA Soil Climate Analysis Network (SCAN), which provides high-resolution, near-real-time soil moisture data across agricultural zones in the United States [35]. Similarly, the FLUXNET network offers point-based data through eddy covariance towers, which include soil moisture as part of broader ecosystem flux measurements [36]. In terms of satellite-derived products, SMAP and ESA's SMOS missions provide global, gridded soil moisture datasets at regular intervals [37]. The Advanced Scatterometer (ASCAT) onboard EUMETSAT MetOp satellites also offers a long-term record of soil moisture estimations with relatively high temporal resolution [38]. Additionally, regional in-situ networks such as the OzNet (Australia), REMEDHUS (Spain), and ARM Southern Great Plains (USA) serve as valuable sources for local model calibration and validation. These alternatives, while often complementary to ISMN, highlight the diversity of data sources available for soil moisture modeling and reinforce the importance of integrated approaches that combine satellite, in-situ, and model-based observations.

## V. Alternative Datasets and Data Fusion Strategies for Soil Moisture Monitoring

Recent studies have evaluated the performance of alternative soil moisture datasets relative to the International Soil Moisture Network (ISMN). These datasets provide complementary advantages in spatial coverage, temporal continuity, and measurement accuracy. For example, the USDA Soil Climate Analysis Network (SCAN) often demonstrates lower root mean square error (RMSE), approximately $0.04\ \mathrm{m^3/m^3}$, in agricultural validation studies across the United States compared to global ISMN averages of approximately $0.06\ \mathrm{m^3/m^3}$. This improvement is largely attributed to denser station coverage and standardized sensor installations [39]. FLUXNET provides valuable integration of soil moisture measurements with ecosystem flux observations, enabling analysis of land–atmosphere exchanges. However, its global network consists of roughly 200 sites, significantly fewer than the approximately 2200 stations in ISMN, which limits its applicability for high-resolution spatial modeling [40].

Satellite-derived products such as ASCAT offer improved temporal continuity, typically providing daily observations compared with the occasional temporal gaps present in in-situ datasets. Nevertheless, these products operate at coarser spatial resolutions (approximately 25 km), with validation studies reporting unbiased RMSE values around $0.05\ \mathrm{m^3/m^3}$ when compared with in-situ observations [39]. Regional soil moisture networks, including OzNet and REMEDHUS, can achieve high site-specific accuracy, with validation coefficients of determination ($R^2$) exceeding 0.85. However, their geographic coverage remains limited, restricting global applicability [40]. Table II summarizes key alternative soil moisture datasets and their complementary roles relative to ISMN.

Table II
Comparison of representative soil moisture datasets relative to ISMN.

| Dataset | Coverage | Resolution | Strengths | Limitations |
|---|---|---|---|---|
| USDA SCAN | U.S. | Point-based | Standardized sensors and rich metadata | Limited geographic coverage |
| FLUXNET | Global | Point-based | Combines ecosystem flux and soil moisture measurements | Sparse station distribution |
| SMAP / SMOS | Global | 9–36 km | Broad spatial coverage and frequent observations | Coarse resolution and vegetation effects |
| ASCAT | Global | $\sim$25 km | Long-term record and high temporal frequency | Limited to surface soil moisture |
| Regional Networks | Local / Regional | Point-based | High-quality site-specific measurements | Limited scalability and coverage |

Leveraging these complementary datasets can help mitigate spatial biases in ISMN, particularly in underrepresented environments such as arid and semi-arid regions where satellite products such as ASCAT may provide more reliable temporal coverage [39]. Recent research increasingly focuses on data fusion techniques to integrate satellite observations with in-situ measurements. Approaches such as ensemble Kalman filtering and artificial neural networks have demonstrated substantial improvements in soil moisture estimation accuracy. For instance, integrating SMAP satellite observations with ground measurements has reduced RMSE by approximately 31–50% in semi-arid grassland ecosystems [41].

More advanced techniques incorporating deep learning, ensemble modeling, and multi-sensor satellite fusion have further enhanced spatial resolution and predictive performance [42]–[44]. Emerging distributed learning frameworks, including federated learning approaches, also show promise for soil moisture prediction in large-scale agricultural Internet-of-Things networks [45]. These developments highlight the growing importance of integrating heterogeneous datasets and advanced machine learning techniques to improve soil moisture monitoring and prediction capabilities [46]. These advances align closely with ongoing efforts at the Precision Agriculture Research Lab at GVSU to develop integrated predictive frameworks for precision agriculture applications.

## VI. Multi-Site Case Study: SMAP Satellite Estimates Versus ISMN Ground Observations

This section extends the comparative analysis of SMAP and ISMN soil moisture data for the two sites introduced

Table III
COMPARISON OF STATISTICAL METRICS BETWEEN SMAP AND ISMN SOIL MOISTURE OBSERVATIONS AT THE MICHIGAN AND INDIANA SITES. ME: MEAN ERROR; MAE: MEAN ABSOLUTE ERROR; RMSE: ROOT MEAN SQUARE ERROR; $r$: PEARSON CORRELATION COEFFICIENT. ALL ERROR METRICS IN M$^3$/M$^3$.

| Site | Layer | ME | MAE | RMSE | $r$ |
|---|---|---|---|---|---|
| Michigan | Surface | 0.019 | 0.024 | 0.031 | 0.48 |
| Michigan | Root-zone | -0.009 | 0.017 | 0.020 | 0.63 |
| Indiana | Surface | -0.002 | 0.077 | 0.111 | -0.01 |
| Indiana | Root-zone | 0.039 | 0.070 | 0.112 | -0.09 |

in Section I: Gaylord-9-SSW (Michigan, U.S.) and Bedford-5-WNW (Indiana, U.S.). The objective is to evaluate the consistency, bias, and variability between satellite-derived and in-situ measurements across multiple temporal scales, and to identify systematic discrepancies arising from spatial resolution differences, measurement techniques, and environmental heterogeneity.

### A. Cross-Site Statistical Overview

Table III summarizes the key statistical metrics for both sites and soil layers. A clear contrast emerges between the two locations: Michigan exhibits strong agreement between SMAP and ISMN observations, particularly for the root-zone product, while Indiana shows substantially weaker agreement, with higher errors and near-zero or negative correlations across both layers.

The site-dependent performance of SMAP visible in Table III is most evident in two contrasts. First, root-zone RMSE at the Indiana site (0.112 m$^3$/m$^3$) is more than five times larger than at the Michigan site (0.020 m$^3$/m$^3$), indicating that retrieval error is not uniformly distributed across environments but is strongly conditioned by local land surface characteristics. Second, while Michigan yields meaningful positive correlations at both layers ($r = 0.48$ surface, $r = 0.63$ root-zone), Indiana produces near-zero or slightly negative values ($r = -0.01$ surface, $r = -0.09$ root-zone), suggesting that SMAP fails to track day-to-day moisture variability at that site entirely, irrespective of the depth layer considered.

These contrasts are attributable to the distinct land surface characteristics of the two sites [23]. The Indiana site is heavily influenced by extensive row-crop agriculture, highly variable soil organic matter, and the presence of artificial tile drainage infrastructure. Together, these factors introduce sub-pixel spatial heterogeneity that significantly complicates passive microwave moisture retrievals [25][26]. The Michigan site, by contrast, presents a comparatively stable, naturally vegetated landscape that aligns more closely with the parameterizations underlying SMAP's retrieval algorithms, yielding higher correlations and lower overall error [13].

It is important to note that the near-zero correlation observed at Indiana does not imply that SMAP is globally unreliable; rather, it underscores the strong site dependency of satellite soil moisture retrievals and cautions against using a single ground-based benchmark such as ISMN to draw general conclusions about retrieval performance. This distinction motivates the site-specific residual analyses presented next.

### B. Michigan Site: Gaylord-9-SSW

For the Michigan site ($n = 732$ daily observations), SMAP and ISMN exhibit strong seasonal consistency, with both datasets capturing wet and dry periods effectively (Figure 5).

At the daily scale, SMAP surface soil moisture shows a positive bias ($ME \approx 0.019$), indicating overestimation relative to ISMN observations. Error levels are moderate ($MAE \approx 0.024$, $RMSE \approx 0.031$). In contrast, the root-zone product demonstrates improved accuracy with lower errors ($MAE \approx 0.017$, $RMSE \approx 0.020$) and a slight negative bias ($ME \approx -0.009$). **Correlation analysis** further confirms stronger agreement for the root-zone product ($r \approx 0.63$) compared to the surface product ($r \approx 0.48$), suggesting that deeper-layer estimates better capture soil moisture variability. **Residual analysis** (Figures 6 and 7) indicates that surface estimates exhibit greater variability and occasional extreme deviations, whereas root-zone residuals are more stable and tightly distributed. Temporal aggregation (Figure 5) reduces high-frequency variability and further improves agreement, indicating that discrepancies are partly driven by daily-scale variability and spatial mismatch. Overall, SMAP performs well at this site, particularly for root-zone soil moisture, demonstrating strong capability in capturing temporal dynamics.

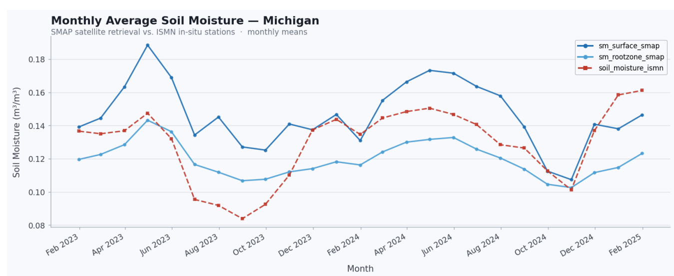


Figure 5. Monthly average soil moisture comparison between SMAP (surface and root-zone) and ISMN observations at the Michigan site.

The residual analysis in Figure 6 for the Michigan site reveals that SMAP surface retrievals exhibit a consistent positive wet bias relative to ISMN in-situ measurements, with a mean bias of $+0.0186$ m$^3$/m$^3$ and an RMSE of $0.0313$ m$^3$/m$^3$. Critically, the residuals are not randomly distributed around zero, the OLS trend line (slope $= -0.463$, $R^2 = 0.185$) indicates a statistically meaningful moisture-dependent bias structure, whereby SMAP most strongly overestimates soil moisture at low in-situ values (approximately $0.08$–$0.10$ m$^3$/m$^3$) and progressively converges toward, and ultimately crosses, zero bias at higher moisture levels near $0.17$–$0.18$ m$^3$/m$^3$. The density-coloured scatter shows that the bulk of observations cluster between $0.12$ and $0.15$ m$^3$/m$^3$, where SMAP and ISMN are in closest agreement, and the $\pm 1$ SD band of $0.0251$ m$^3$/m$^3$ confirms moderate but non-negligible spread. This moisture-dependent overestimation at drier conditions is consistent with known SMAP retrieval limitations in temperate, vegetated

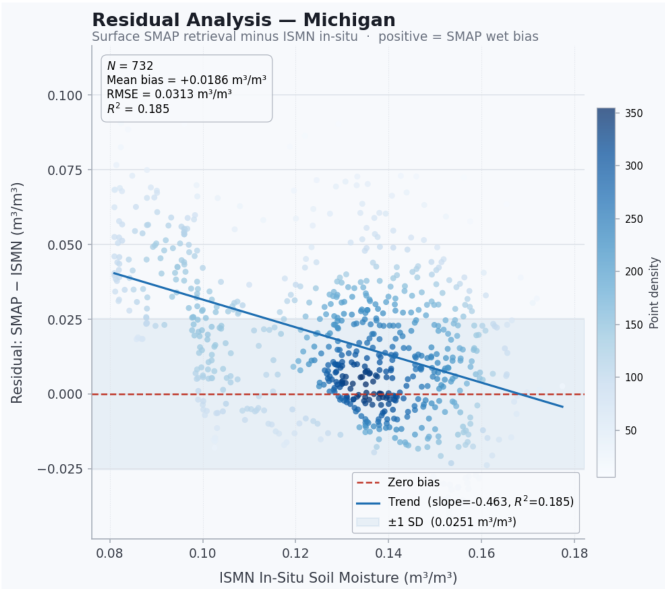


Figure 6. Residuals (SMAP surface minus ISMN) over time for Michigan.

landscapes such as Michigan, where canopy opacity and soil heterogeneity can inflate passive microwave-derived surface moisture estimates when true soil moisture is low.

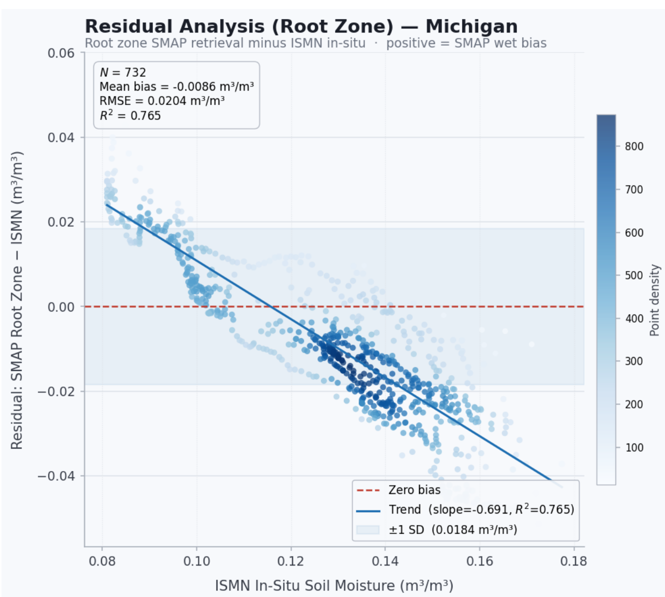


Figure 7. Residuals (SMAP root-zone minus ISMN) over time for Michigan.

The root zone residual analysis for the Michigan site in Figure 7 reveals a strongly moisture-dependent bias structure between SMAP root zone retrievals and ISMN in-situ measurements, with the OLS trend line (slope $= -0.691$, $R^2 = 0.765$) explaining over three-quarters of the residual variance, a markedly stronger relationship than observed for the surface layer. The mean bias is small and slightly negative ($-0.0086$ m$^3$/m$^3$), indicating that SMAP root zone estimates are near-unbiased on average, yet this aggregate statistic masks a pronounced systematic pattern: SMAP overestimates at low in-situ moisture values (roughly 0.08–0.10 m$^3$/m$^3$) and increasingly underestimates as in-situ moisture rises above approximately 0.10 m$^3$/m$^3$, with the zero-crossing occurring near 0.10 m$^3$/m$^3$. The tight clustering of high-density points along the trend line, particularly visible between 0.12 and 0.15 m$^3$/m$^3$, and the relatively narrow $\pm 1$ SD band of 0.0184 m$^3$/m$^3$ together confirm that this is a systematic, structured bias rather than random noise. The RMSE of 0.0204 m$^3$/m$^3$ is notably lower than that of the surface layer (0.0313 m$^3$/m$^3$), suggesting that while the root zone product is more precise, its strong moisture-dependent slope points to a scaling or model assumption within the SMAP root zone algorithm that does not fully capture the dynamic range of soil moisture conditions observed across Michigan's in-situ network.

### C. Indiana Site: Bedford-5-WNW

For the Indiana site ($n = 732$ daily observations), SMAP captures the general seasonal patterns observed in ISMN data (Figure 8), but with substantially weaker agreement compared to Michigan.

At the daily scale, SMAP exhibits smoother temporal behavior, reflecting spatial averaging effects. The surface product shows a negligible bias ($ME \approx -0.002$), but significantly larger errors ($MAE \approx 0.077$, $RMSE \approx 0.111$) compared to Michigan. The root-zone product shows a positive bias ($ME \approx 0.039$) with similarly high error levels ($MAE \approx 0.070$, $RMSE \approx 0.112$). Unlike Michigan, **correlation** between SMAP and ISMN is extremely weak, with $r \approx -0.01$ (surface) and $r \approx -0.09$ (root-zone). This indicates that SMAP fails to capture day-to-day variability at this site, despite reproducing general seasonal trends. **Residual analysis** (Figures 9 and 10) reveals structured discrepancies, suggesting the influence of environmental factors such as precipitation events, vegetation, and soil properties. Temporal aggregation (Figure 8) improves visual agreement, indicating that discrepancies are most pronounced at finer temporal scales. Overall, SMAP performance at the Indiana site is notably weaker, with higher errors and poor correlation, highlighting strong site dependency in satellite soil moisture accuracy.

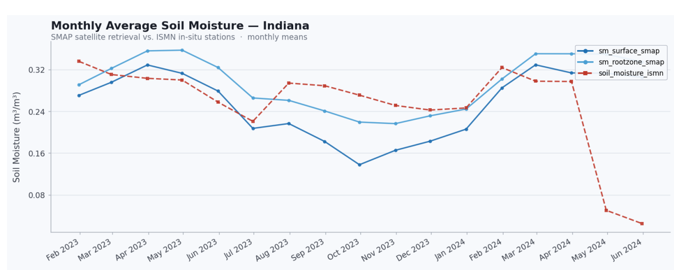


Figure 8. Monthly average soil moisture comparison between SMAP and ISMN at the Indiana site.

The surface residual analysis for Indiana in Figure 9 reveals a pattern of SMAP retrieval error that closely mirrors the root zone results for the same state, yet with important distinctions.

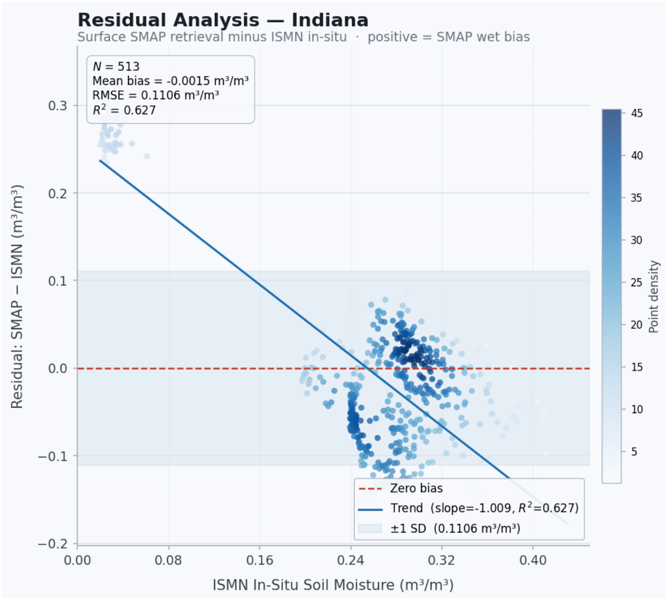


Figure 9. Residuals (SMAP surface minus ISMN) over time for Indiana.

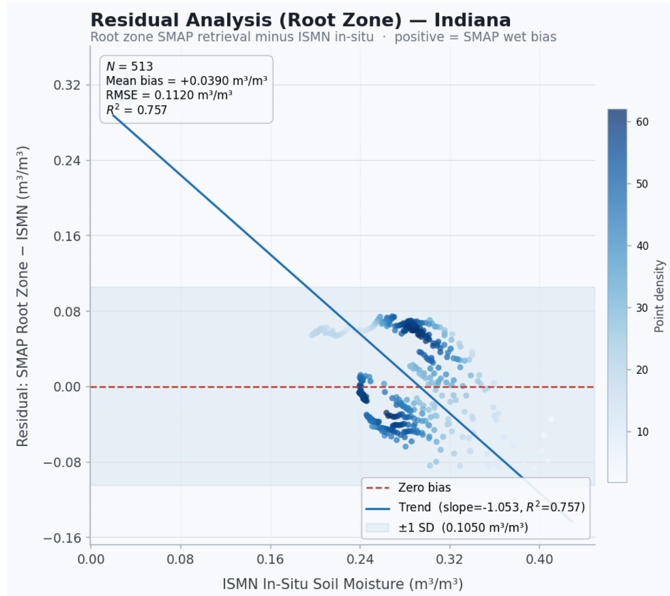


Figure 10. Residuals (SMAP root-zone minus ISMN) over time for Indiana.

The mean bias is near-zero at $-0.0015$ m$^3$/m$^3$, suggesting that SMAP surface retrievals are essentially unbiased on average over Indiana, a deceptively favourable statistic that obscures a severe moisture-dependent structure captured by the OLS trend line (slope $= -1.009$, $R^2 = 0.627$). A slope of approximately $-1$ implies that SMAP retrievals are almost entirely insensitive to the dynamic range of in-situ moisture variability, effectively returning a near-constant value while the ground truth varies substantially across the 0.24–0.38 m$^3$/m$^3$ range observed by the ISMN network. The RMSE of 0.1106 m$^3$/m$^3$ is essentially identical to the root zone RMSE (0.1120 m$^3$/m$^3$) and more than five times larger than the Michigan surface RMSE (0.0313 m$^3$/m$^3$), confirming that Indiana represents a persistently challenging retrieval environment at both depth layers. The scatter plot further reveals a structurally bimodal residual distribution, with a dense high-positive cluster near 0.24 m$^3$/m$^3$ and a secondary negative cluster near 0.30 m$^3$/m$^3$, which together drive the wide $\pm 1$ SD band of 0.1106 m$^3$/m$^3$ and point toward systematic station-level or land-surface heterogeneity, likely tied to Indiana's extensive row-crop agriculture, artificial tile drainage, and spatially variable soil organic matter, that the SMAP surface retrieval algorithm does not adequately resolve.

The root zone residual analysis for the Indiana site in Figure 10 exposes a severe and systematic disagreement between SMAP root zone retrievals and ISMN in-situ measurements that is qualitatively distinct from the Michigan results. The OLS trend line (slope $= -1.053$, $R^2 = 0.757$) carries a slope exceeding unity in magnitude, indicating that the residual grows faster than the in-situ moisture signal itself, a hallmark of a fundamental dynamic range mismatch rather than a simple offset. The mean bias of $+0.0390$ m$^3$/m$^3$ is considerably larger than Michigan's ($-0.0086$ m$^3$/m$^3$), and the RMSE of 0.1120 m$^3$/m$^3$ is more than five times higher (0.0204 m$^3$/m$^3$ for Michigan), reflecting the much wider scatter visible in the figure. Notably, the in-situ observations are confined to a narrow and relatively high moisture range of approximately 0.24–0.38 m$^3$/m$^3$, yet the residuals span nearly $\pm 0.10$ m$^3$/m$^3$ within that window, and the data appear to cluster into two distinct lobes straddling the zero-bias line, suggestive of station-level heterogeneity, multi-depth sampling artefacts, or land cover discontinuities across the Indiana network. The wide $\pm 1$ SD band of 0.1050 m$^3$/m$^3$ further underscores the large unexplained variability. Taken together, these results suggest that the SMAP root zone algorithm performs substantially worse over Indiana than Michigan, possibly reflecting differences in soil texture, agricultural land use, or tile drainage infrastructure that are not adequately parameterized in the SMAP retrieval model.

## VII. Conclusions and Recommendations

Accurate soil moisture estimation is essential for sustainable water resource management, yet the reliability of the benchmark datasets used to evaluate predictive models has received relatively limited attention. This study addressed this gap through a combined review and empirical analysis, providing insights with direct implications for the development, validation, and deployment of soil moisture prediction systems.

The review showed that recent advances in machine learning, deep learning, and data fusion have substantially improved soil moisture prediction capabilities. However, progress has been constrained in part by limitations in the benchmark datasets themselves. Although ISMN remains an indispensable source of global ground observations, it suffers from several well-documented challenges, including uneven spatial distribution, sensor heterogeneity, inconsistent metadata, and temporal gaps. These issues represent structural limitations that may systematically bias model evaluation, particularly when a single dataset is used as the sole benchmark.

The empirical case study further reinforced this concern. At the Michigan site, SMAP root-zone estimates exhibited reasonable agreement with ISMN observations ($r = 0.63$, RMSE $= 0.02$ $\mathrm{m}^3/\mathrm{m}^3$), and residual analysis revealed moisture-dependent biases consistent with known retrieval limitations in vegetated temperate environments. In contrast, performance at the Indiana site was considerably poorer, with both surface and root-zone estimates exhibiting RMSE values exceeding $0.11$ $\mathrm{m}^3/\mathrm{m}^3$ and near-zero correlation with in-situ measurements. Moreover, the ordinary least squares residual slope approached $-1$, indicating that SMAP retrievals were largely insensitive to variations in ground moisture conditions and instead produced nearly constant estimates while observed moisture values varied substantially. This behavior reflects a fundamental dynamic-range mismatch rather than a simple calibration offset. Environmental characteristics such as extensive row-crop agriculture, subsurface tile drainage, and heterogeneous soil organic matter likely contribute to these discrepancies and are not adequately represented in current retrieval algorithms.

Taken together, these findings suggest three practical recommendations for researchers and practitioners:

1) **Adopt multi-source validation as standard practice.** Reliance on ISMN alone, or any single benchmark dataset, may lead to misleading conclusions regarding model performance. Instead, multiple complementary datasets should be incorporated. For example, SCAN can provide agricultural measurements, ASCAT offers temporally continuous observations, and regional networks such as OzNet and REMEDHUS provide valuable site-specific information. Agreement across independent sources provides stronger evidence of model robustness.
2) **Apply site-aware preprocessing before using ISMN for validation.** Sites characterized by intensive agriculture, artificial drainage systems, or high sub-pixel heterogeneity should receive additional scrutiny. Residual analysis, which examines bias as a function of soil moisture magnitude rather than relying solely on aggregate metrics, can reveal systematic errors that may be overlooked by summary statistics.
3) **Prioritize data fusion approaches in heterogeneous agricultural landscapes.** The Indiana case study suggests that passive microwave retrievals alone may be insufficient in row-crop environments. Integrating SMAP observations with higher-resolution ancillary information, including soil texture maps, crop calendars, and drainage infrastructure data, through ensemble Kalman filtering or deep learning-based fusion frameworks represents a promising direction for improving field-scale soil moisture estimation.

Future work at the Precision Agriculture Research Lab at GVSU will extend this analysis to additional ISMN sites encompassing diverse land cover types and climatic regimes. Planned efforts include the development of site-classification protocols to guide benchmark selection and the construction of integrated predictive frameworks that combine satellite, in-situ, and model-based observations for precision agriculture applications. Ultimately, the reliability of soil moisture prediction systems depends not only on the sophistication of the underlying models, but also on the rigor with which the datasets used for training and evaluation are characterized and understood.

## References


[1] K. Hasan and A. Muiruri, "Predictive modeling of soil moisture: A review of benchmark datasets, their strengths, and limitations", *The Second International Conference on Sustainable and Regenerative Farming (CSRF)*, vol. 978-1-68558-327-9, pp. 9–12, 2025.

[2] L. F. Konikow, "Groundwater depletion in the united states (1900–2008)", *US Geological Survey Scientific Investigations Report*, vol. 2013, no. 5079, p. 63, 2005.

[3] S. Foster and J. Chilton, "Groundwater: The processes and global significance of aquifer degradation", *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1440, pp. 1957–1972, 2003.

[4] U.S. Department of Agriculture, National Agricultural Statistics Service, "2018 irrigation and water management survey", USDA-NASS, Tech. Rep., 2019.

[5] S. Jasechko, D. Perrone, K. M. Befus, *et al.*, "Rapid groundwater decline and some cases of recovery in aquifers globally", *Nature*, vol. 625, pp. 715–721, 2024. DOI: 10.1038/s41586-023-06879-8.

[6] D. Entekhabi *et al.*, "The soil moisture active passive (SMAP) mission", *Proceedings of the IEEE*, vol. 98, no. 5, pp. 704–716, 2010. DOI: 10.1109/JPROC.2010.2043918.

[7] W. Dorigo *et al.*, "Esa cci soil moisture for improved earth system understanding: State-of-the art and future directions", *Remote Sensing of Environment*, vol. 203, pp. 185–215, 2017.

[8] I. Eniang, A. Edet, and E. Anwan, "Importance of soil moisture content in agriculture: A review", *International Journal of Agriculture Innovations and Research*, vol. 7, no. 2, pp. 2319–1473, 2018.

[9] L. Alletto, Y. Coquet, P. Benoit, D. Heddadj, and E. Barriuso, "Soil moisture monitoring for optimizing irrigation management in a maize cropping system", *Agricultural Water Management*, vol. 148, pp. 1–13, 2015.

[10] W. A. Dorigo *et al.*, "The international Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements", *Hydrology and Earth System Sciences*, vol. 15, no. 5, pp. 1675–1698, 2011. DOI: 10.5194/hess-15-1675-2011.

[11] M. Rodell *et al.*, "The global land data assimilation system", *Bulletin of the American Meteorological Society*, vol. 85, no. 3, pp. 381–394, 2004.

[12] K. W. Oleson *et al.*, "Technical description of version 4.0 of the community land model (clm)", NCAR, Tech. Rep., 2010.

[13] T. Enemark, I. Sandholt, R. Fensholt, and K. H. Jensen, "Evaluating the skill of land surface models to represent the soil moisture variability over africa", *Hydrology and Earth System Sciences*, vol. 24, no. 7, pp. 3717–3738, 2020.

[14] Y. Zhang, Z. Li, H. Shen, and E. Chen, "Machine learning for downscaling remotely sensed soil moisture using multisource data", *Remote Sensing*, vol. 10, no. 3, p. 431, 2018.

[15] D. Jeong, H. Lee, and S. Lee, "Machine learning-based soil moisture prediction using remote sensing and weather data", *Environmental Modelling & Software*, vol. 148, p. 105 248, 2022.

[16] S. Khaki *et al.*, "Cnn-rnn deep learning framework for spatiotemporal soil moisture estimation", *Water Resources Research*, vol. 56, no. 2, 2020.

[17] M. Reichstein *et al.*, "Deep learning and process understanding for data-driven earth system science", *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.

[18] J. Yin, Q. Zhuang, H. Tian, and M. Pan, "A physics-informed machine learning framework for hydrological modeling: Application to the mississippi river basin", *Water Resources Research*, vol. 57, no. 8, 2021.

[19] S. Chan, P. O'Neill, E. Njoku, T. Jackson, and R. Bindlish, "Soil moisture active passive (smap) enhanced level 3 radiometer global daily 9 km ease-grid soil moisture, version 1", *NASA National Snow and Ice Data Center Distributed Active Archive Center*, 2018.

[20] M. Saberi, H. Moradkhani, W. Bardsley, and X. He, "Multiscale soil moisture modeling for agricultural drought monitoring using remote sensing and land surface modeling data", *Agricultural and Forest Meteorology*, vol. 308, p. 108 548, 2021.

[21] F. Greifeneder, L. Brocca, T. Pellarin, and W. Wagner, "Machine learning in soil moisture retrievals: Progress, challenges, and future directions", *Reviews of Geophysics*, vol. 59, no. 2, 2021.

[22] A. Gruber, W. T. Crow, W. A. Dorigo, and W. Wagner, "Characterizing coarse-scale representativeness of in situ soil moisture measurements from the international soil moisture network", *Vadose Zone Journal*, vol. 12, no. 2, 2013.

[23] N. Nicolai-Shaw, L. Gudmundsson, M. Hirschi, and S. I. Seneviratne, "Improving soil moisture modeling in heterogeneous landscapes", *Hydrology and Earth System Sciences*, vol. 23, pp. 1217–1235, 2019.

[24] W. A. Dorigo *et al.*, "The international soil moisture network: Serving earth system science for over a decade", *Hydrology and Earth System Sciences*, vol. 25, no. 2, pp. 5749–5804, 2021. DOI: 10.5194/hess-25-5749-2021.

[25] S. Kim, T. J. Jackson, R. Bindlish, M. H. Cosh, and V. Lakshmi, "Validation of smos soil moisture retrievals over the continental us using ground-based observations", *Remote Sensing of Environment*, vol. 121, pp. 383–398, 2012.

[26] Y. Ma *et al.*, "Validation of smap soil moisture products with in situ observations from the international soil moisture network", *Remote Sensing*, vol. 8, no. 6, p. 524, 2016.

[27] H. Wu, X. Li, and J. Peng, "Evaluating soil moisture retrievals from amsr2 using ground-based observations over semiarid regions in china", *Remote Sensing*, vol. 9, no. 5, p. 499, 2017.

[28] Y. Liu, R. M. Parinussa, W. A. Dorigo, R. A. M. de Jeu, and W. Wagner, "A comparative study of four algorithms for downscaling remotely sensed soil moisture over the contiguous us", *Remote Sensing of Environment*, vol. 179, pp. 1–19, 2016.

[29] S. K. Do, T.-N.-D. Tran, M.-H. Le, J. Bolten, and V. Lakshmi, "A novel validation of satellite soil moisture using sm2rain-derived rainfall estimates", *Frontiers in Remote Sensing*, vol. 5, p. 1 474 088, 2024.

[30] C. S. C. Draper *et al.*, "Comparison of near-surface soil moisture from smos and ascat with in situ measurements from the u.s. scan network", *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3070–3088, 2011. DOI: 10.1016/j.rse.2011.03.026.

[31] L. Xu, X. Zhang, J. Wang, and Y. Zhu, "Downscaling SMAP soil moisture using wide and deep learning with spatial feature engineering", *Remote Sensing*, vol. 14, no. 4, p. 987, 2022. DOI: 10.3390/rs14040987.

[32] M. Celik, Y. Yilmaz, and O. Kose, "Multisource deep learning model for soil moisture prediction in diverse climates", *Environmental Modelling & Software*, vol. 155, p. 105 430, 2022.

[33] S. Lee, M. Choi, and E. Kim, "A deep learning framework for smap soil moisture retrieval without radiative transfer models", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.

[34] P. Custódio and A. Prati, "Adaptive ensemble models for soil moisture estimation in iot-supported irrigation", *Agricultural Water Management*, vol. 295, p. 108 579, 2024.

[35] G. L. Schaefer, M. H. Cosh, and T. J. Jackson, "Usda soil climate analysis network (scan)", *Journal of Atmospheric and Oceanic Technology*, vol. 24, no. 12, pp. 2073–2077, 2007.

[36] D. Baldocchi *et al.*, "FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities", *Bulletin of the American Meteorological Society*, vol. 82, no. 11, pp. 2415–2434, 2001. DOI: 10 . 1175 / 1520 - 0477(2001 ) 082<2415:FANTTS>2.3.CO;2.

[37] Y. H. Kerr *et al.*, "The SMOS mission: New tool for monitoring key elements of the global water cycle", *Proceedings of the IEEE*, vol. 98, no. 5, pp. 666–687, 2010. DOI: 10.1109/JPROC.2010.2043032.

[38] W. Wagner *et al.*, "ASCAT soil moisture product: Operational validation, hydrometeorological applications, and beyond", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 3, pp. 1452–1466, 2013. DOI: 10.1109/JSTARS.2013.2262336.

[39] A. C. Teixeira, M. Bakon, D. Lopes, A. Cunha, and J. J. Sousa, "A systematic review on soil moisture estimation using remote sensing data for agricultural applications", *Science of Remote Sensing*, p. 100 328, 2025.

[40] C. Zheng, L. Jia, and T. Zhao, "A 21-year dataset (2000–2020) of gap-free global daily surface soil moisture at 1-km grid resolution", *Scientific Data*, vol. 10, no. 1, p. 139, 2023.

[41] Y. Zhu, L. Zhang, F. Li, J. Xu, and C. He, "Comparison of data fusion methods in fusing satellite products and model simulations for estimating soil moisture on semi-arid grasslands", *Remote Sensing*, vol. 15, no. 15, p. 3789, 2023.

[42] P. Harani, S. Gautam, S. K. Joshi, L. G. Asirvatham, and B. L. Rakshith, "Advancements in soil moisture estimation through integration of remote sensing and artificial intelligence techniques", *Science of The Total Environment*, vol. 1001, p. 180 503, 2025.

[43] A. Singh and K. Gaurav, "Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images", *Scientific Reports*, vol. 13, no. 1, p. 2251, 2023.

[44] M. S. Tahmouresi, M. H. Niksokhan, and A. H. Ehsani, "Enhancing spatial resolution of satellite soil moisture data through stacking ensemble learning techniques", *Scientific Reports*, vol. 14, no. 1, p. 25 454, 2024.

[45] S. Zakzouk and L. A. Said, "Federated learning for soil moisture prediction: Benchmarking lightweight cnns and robustness in distributed agricultural iot networks", *Machine Learning and Knowledge Extraction*, vol. 7, no. 4, p. 132, 2025.

[46] S. O. Ihuoma, C. A. Madramootoo, and M. Kalacska, "Integration of satellite imagery and in situ soil moisture data for estimating irrigation water requirements", *International journal of applied earth observation and geoinformation*, vol. 102, p. 102 396, 2021.