https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


Open Access Earth System Science Data Discussions

# GSSM-10 (Global 10-m Surface Soil Moisture) Derived from Multi-Sensor Data and Ensemble Learning

**Nuo Xu[1], Andre Daccache[1*], Arman Ahmadi[2]**

*Corresponding author. Email: adaccache@ucdavis.edu

[1]Department of Biological and Agricultural Engineering, University of California, Davis, 95616, USA

[2]Department of Environmental Science, Policy, and Management, University of California, Berkeley, 94720, USA

**Abstract**

Satellite-driven soil moisture monitoring systems currently available fail to meet the spatial resolution requirement for a wide range of applications. This limitation is particularly critical for agricultural water management, assessing risks associated with extreme events, and hydrological modeling. This work aims to address the spatial limitations of satellite soil moisture remote sensing by developing GSSM-10, a global 10-meter resolution surface soil moisture dataset, using multi-sensor datasets integrated within an ensemble machine learning framework. These datasets encompass diverse data types—active microwave, multispectral, thermal infrared, and land elevation—offering a robust and comprehensive approach to estimating surface soil moisture (SSM). The ensemble model incorporates TabNet, Random Forest (RF), and Extreme Gradient Boosting (XGBoost). The model was trained on ground-truth data collected from the International Soil Moisture Network (ISMN). The ensemble model demonstrated robust performance, achieving an $R^2$ of 0.8344, a bias of –0.0001, an RMSE of 0.0433 $m^3/m^3$, and an ubRMSE of 0.0433 $m^3/m^3$ in 5-fold cross-validation. When evaluated on a held-out test set, the model achieved similar levels of accuracy, with an $R^2$ of 0.8591, a bias of –0.0002 $m^3/m^3$, and an RMSE/ubRMSE of 0.0401 $m^3/m^3$. An interactive web platform has been developed for data access, visualization, and download, enabling broad adoption by researchers, practitioners, and policymakers. By providing globally consistent, high-resolution SM estimates, GSSM-10 represents a significant advancement in satellite-based soil moisture monitoring for environmental and agricultural applications.

**Keywords:** soil moisture, remote sensing, data fusion, machine learning.

## 1. Introduction

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


Soil moisture plays a pivotal role in the Earth's water and energy cycles, influencing climate, hydrology, and ecosystem dynamics (Klein and Taylor, 2020; Porkka et al., 2024; Sun et al., 2025; Zhang et al., 2020). It governs the partitioning of rainfall into runoff or infiltration and modulates evapotranspiration, thereby affecting weather and climate patterns at regional to global scales (Denissen et al., 2022; Seneviratne et al., 2010; Sun et al., 2025). Recognized as an essential climate variable, soil moisture data are indispensable for climate research (Humphrey et al., 2021; Liu et al., 2023; Qiao et al., 2023; Seneviratne et al., 2010; Soares and Lima, 2022), hydrological modeling (Droppers et al., 2024; Eini et al., 2023; Fatima et al., 2024; Leonarduzzi et al., 2021; Mei et al., 2023), drought and flood forecasting (Lesinger and Tian, 2025; Qing et al., 2023; Wasko and Nathan, 2019; Wyatt et al., 2020; Yao et al., 2023), and agricultural management (Chatterjee et al., 2022; Li et al., 2022; Martínez-Fernández et al., 2016; McNairn et al., 2012; Zhou et al., 2021).

Compared to traditional in-situ point measurement soil moisture (SM) sensors, satellite remote sensing offers cost-effective and large-scale SM monitoring solutions (Chaudhary et al., 2022; Cheng et al., 2022; Dubois et al., 2021). NASA's Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2010), EUMETSAT's Advanced Scatterometer (ASCAT) (Wagner et al., 2013), and ESA's Soil Moisture and Ocean Salinity (SMOS) (Kerr et al., 2001) have demonstrated the power of spaceborne sensors to map surface soil moisture globally. However, their low spatial resolutions are not pertinent to many applications, particularly to agriculture and field-level scale applications (Babaeian et al., 2021; Nguyen et al., 2022). For example, SMAP's radiometer measures moisture at ~36 km resolution (with an enhanced 9 km product), ASCAT scatterometer products are available at ~12.5 km, and SMOS provides data at ~25 km. Even multi-sensor merged datasets like the ESA Climate Change Initiative (CCI) soil moisture (which merges multiple satellites) are typically gridded at 0.25° (~25 km) (Dorigo et al., 2017). At these scales, crucial fine-scale heterogeneity is lost. Model-based products such as ERA5-Land (Muñoz-Sabater et al., 2021), which provides hourly global soil moisture data at ~9 km resolution, and GLDAS (Global Land Data Assimilation System), which produces 3-hourly estimates at 0.25° resolution (Syed et al., 2008; Zawadzki and Kędzior, 2016), offer valuable temporal continuity but remain too coarse for capturing sub-field heterogeneity. While these datasets are essential for large-scale hydrological modeling and climate analysis, their spatial granularity is insufficient for localized applications(Liu et al., 2019; Xu et al., 2021). This mismatch between the scale of observation and the scale of decision-making severely constrains the usability of current global datasets for applications such as irrigated agriculture, flood forecasting, and wildfire risk assessment (Gebrechorkos et al., 2023; He et al., 2023; Peng et al., 2021; Sabaghy et al., 2020).

A variety of global SM datasets have been developed using remote sensing, land surface models, and machine learning. Recently, several products with a resolution of ~1 km have emerged, marking a significant step toward finer spatial detail. For

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


example, Fan et al. (2025) introduced a global 1-km SM product derived from Sentinel-1 SAR observations. Zhang et al. (2023) generated a daily 1-km global surface SM dataset for 2000–2020 by integrating multi-source satellite-driven information (albedo, land temperature, leaf area index) and reanalysis data using an ensemble learning (XGBoost) model. Zheng et al., 2023 achieved a similar 1-km global product by downscaling the 0.25° ESA CCI satellite soil moisture with Random Forest, producing a gap-free daily 21-year record (2000–2020). Han et al., 2023 developed another 1-km global dataset (surface top 5 cm layer) using a physics-informed machine learning approach, achieving high accuracy (correlation coefficient of ~0.9) from 2000 to 2020. These high-resolution global datasets represent significant progress in capturing soil moisture at much finer scales than earlier global products. Despite their improved resolution, the ~1 km global datasets still fall short for applications demanding sub-field-scale detail. A 1-km pixel (~100 ha) averages over heterogeneous terrain and management units and is too coarse for field-level precision applications.

To meet fine-scale needs, several regional datasets push spatial resolutions to tens or hundreds of meters, albeit over limited areas. Vergopolan et al. (2021) developed SMAP-HydroBlocks, a 30-m resolution surface 5cm soil moisture dataset for the conterminous U.S. (2015–2019) by integrating high-resolution land surface modeling with downscaled satellite observations and machine learning. Baghdadi et al., (2017) developed the S2MP (Sentinel-1/Sentinel-2-derived Soil Moisture Product), a regional 10 m resolution soil moisture dataset. However, its coverage is presently limited to specific regions (e.g., parts of southern France). In summary, high spatial resolution satellite-derived soil moisture is available at regional scales, but current datasets are constrained by both limited geographic coverage and relatively short temporal records (Table 1). Therefore, there is a need to develop a soil moisture product that combines global coverage with high spatial resolution.

Table 1 Information of common soil moisture products.

| Category | Datasets | Scale | Period | Spatial Resolution | Temporal Resolution | Depth (cm) |
|---|---|---|---|---|---|---|
| Satellite data | SMAP | Global | 2015 - present | 9 km | 2 - 3 days | 0-5; 0-100 |
| | ASCAT | Global | 2000 - present | 12.5 / 25 km | Twice per day | 0-2 |
| | SMOS | Global | 2000 - present | 35 - 50 km | 2 - 3 days | 0-5 |
| Artificially developed products | ERA5 | Global | 1950 - present | 11.132 Km | hourly / daily | 0-7, 7-28, 28-100, and 100-200 |
| | GLDAS | Global | 1948 - present | 27.830 Km | daily | 0-10,; 10-40, 40-100, and |

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


Earth System Science Data Open Access Discussions

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 100-200 |
| Skulovich and Gentine, 2023 | Global | 2002 - 2020 | 25 km | 3days | 0-5 |
| Han et al., 2023 | Global | 2000 - 2020 | 1 km | daily | 0-5 |
| Zheng et al., 2023 | Global | 2000 - 2020 | 1 km | daily | 0-5 |
| Zhang et al., 2023 | Global | 2000 - 2020 | 1 km | daily | 0-5 |
| O. and Orth, 2021 | Global | 2000 - 2019 | 0.25° (27.75 km) | | 0-10, 10-30, and 30-50 |
| SMRFR (Liu et al., 2025) | Global | 2000 - 2023 | 9 km | daily | 0-5, 5-10, 10-30, 30-50, and 50-100 |
| SMAP-HydroBlocks (Vergopolan et al., 2021) | U.S. | 2015 - 2019 | 30 m | 6 hours | 0-5 |
| Cui et al., 2019 | Tibet Plateau | 2002 - 2015 | 0.25° (27.75 km) | daily | 0-5 |
| Song et al., 2022 | China | 2003 - 2019 | 1 km | daily | 0-10 |
| S2MP (Lozac'h et al., 2020) | parts of France, Morocco, Germany, U. S. A | 2017 - 2024 | field scale | Synchronize with Sentinel-1/2 | 0-5 |

To address existing limitations in soil moisture monitoring, this research introduces a global 10-m surface soil moisture dataset developed using multi-sensor remote sensing data, including active microwave, multispectral, thermal, and elevation inputs, and advanced ensemble machine learning techniques. The dataset provides continuous near real-time updates and includes historical records dating back to January 2016. We employed a variety of ensemble models, such as TabNet (Arik and Pfister, 2021), Random Forest (Belgiu and Drăguţ, 2016), and XGBoost (Chen and Guestrin, 2016), to reduce individual model bias and leverage their complementary strengths. This work represents a significant advancement in high-resolution soil moisture mapping, with broad applications in precision agriculture, high-resolution disaster risk assessment, sub-catchment hydrological modeling, and (micro-) climate research.

## 2. Methodology

### 2.1 Data

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025

Open Access Earth System Science Data Discussions

### 2.1.1 Remote Sensing Data

In this study, we utilized a variety of datasets, each with a unique role in soil moisture analysis (Table 2). Sentinel-1 with Synthetic Aperture Radar is capable of emitting microwaves that penetrate the surface, with varying degrees of signal reflection indicating different levels of moisture in the soil, even under vegetation cover.

Sentinel-2 and Landsat-8/9 provide multispectral imagery highlighting information about vegetation and moisture using various vegetation indices such as the Normalized Difference Vegetation Index (NDVI, eq. (1)) (Pettorelli et al., 2005; Xu et al., 2025) and the Normalized Difference Moisture Index (NDMI, eq. (2)) (Mkhwenkwana et al., 2025; Monteiro et al., 2024). Both datasets were used jointly to enrich the spectral feature space. Differences in sensor characteristics, spectral band configurations, and observation geometries provide complementary information that enhances the model's ability to generalize across varying land surface and atmospheric conditions. Moreover, land topography obtained from ALOS DSM and STRM DEM has a significant impact on soil moisture distribution. For example, areas at higher elevations may experience more runoff, resulting in drier soils, while areas at lower elevations may retain more moisture. ALOS DSM is the primary elevation data used. In areas where ALOS DSM is unavailable, the SRTM DEM is used as a substitute.

$$NDVI = \frac{(NIR-Red)}{(NIR+Red)} \tag{1}$$

$$NDMI = \frac{(NIR-SWIR)}{(NIR+SWIR)} \tag{2}$$

Table 2. Summary of the datasets used in the study.

| Sensor | Spatial resolution | Temporal resolution | Features | Time |
|---|---|---|---|---|
| Sentinel-1 | 10 m | 3-5 days | VV, VH, incident angle | 1/1/2016-12/31/2023 |
| Sentinel-2A/B | 10 m | 2-3 days | B1 - B12, NDVI, NDMI | 1/1/2016-12/31/2023 |
| Landsat-8/9 | 30 | 8 days | B1 - B11, NDVI, NDMI | 1/1/2016-12/31/2023 |
| ALOS DSM / SRTM DEM | 30 m | Static | Elevation | 2006/2000 |
| ISMN | Point-based (station) | Daily | Soil moisture, latitude, longitude | 1/1/2016-12/31/2023 |

### 2.1.2 Ground-based soil moisture dataset

We used ground data from the International Soil Moisture Network (ISMN, https://ismn.earth/en/) for model training and validation. ISMN is a global in-situ data hosting service that consolidates soil moisture measurements from various sources and networks. We filtered the stations to include only those with available data from

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


January 1, 2016, to December 31, 2023, and provided surface soil moisture measurements for the top 5 cm of soil depth. In total, data from 699 stations were used in this study. The spatial distribution of the selected ISMN stations used in this study is shown in Figure 1. While a large proportion of the selected stations are located within the United States, the selected stations span diverse climatic conditions and land cover types.

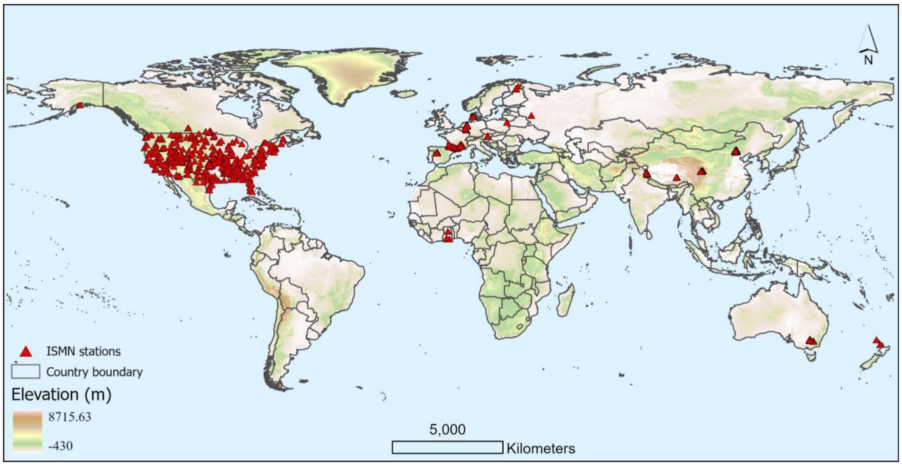



Fig. 1 Spatial distribution of the selected ISMN stations used in this study.

2.2.1 Water cloud model

To correct for vegetation effects on the radar backscatter prior to soil moisture retrieval, a dynamic Water Cloud Model (WCM) approach was applied (Attema and Ulaby, 1978). The WCM relates the observed radar backscatter ( $\sigma_{total}$ ) to contributions from vegetation and soil (Li et al., 2024; Nijaguna et al., 2023).

For each Sentinel-1 acquisition date, the closest available cloud-free Sentinel-2 image within a ±3-day window was identified to capture the most representative vegetation conditions at the time of radar observation. From the selected Sentinel-2 image, the NDVI was computed and sampled at the station location. The sampled NDVI value was then used to dynamically calculate the WCM parameters, specifically the vegetation scattering coefficient (A) and attenuation coefficient (B), for both VV and VH polarizations. Using these dynamically adjusted parameters, vegetation effects were removed from the Sentinel-1 VV and VH backscatter according to the standard WCM formulation (Attema and Ulaby, 1978) Finally, the corrected soil surface backscatter ($\sigma_{soil}$) for both polarizations was retained for subsequent analysis and soil moisture retrieval.

$$\sigma_{total} = \sigma_{veg} + \tau^2 \times \sigma_{soil} \tag{3}$$

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025




where $\sigma_{veg}$ is the direct vegetation scattering contribution, $\tau^2$ is the two-way attenuation factor through the vegetation layer, and $\sigma_{soil}$ is the soil surface backscatter.

The two-way attenuation $\tau^2$ was modeled as:

$$\tau^2 = \exp\left(-2B \times V \times sec\theta\right) \tag{4}$$

and the vegetation scattering $\sigma_{veg}$ was modeled as:

$$\sigma_{veg} = A \times V \times cos\theta \times (1 - \tau^2) \tag{5}$$

Where $V$ is the vegetation descriptor, here taken as NDVI, $\theta$ is the local incidence angle (in radians), and $A$ and $B$ are empirical coefficients dependent on vegetation density.

The soil backscatter was then retrieved by rearranging the WCM equation:

$$\sigma_{soil} = \frac{\sigma_{total} - \sigma_{veg}}{\tau^2} \tag{6}$$

All backscatter computations were performed on a linear scale (i.e., reflectivity coefficient, unitless), with Sentinel-1 backscatter converted from decibels (dB) to a linear scale prior to WCM correction and then converted back to dB afterward.
Unlike conventional methods that use static WCM parameters, this study dynamically adjusted the WCM scattering (A) and attenuation (B) coefficients based on NDVI at each Sentinel-1 acquisition time (Baghdadi et al., 2019, 2017; Rawat et al., 2021). This allowed the vegetation contribution to vary spatially and temporally in response to actual plant growth stages.

For VV polarization:

$$A_{VV} = \begin{cases} 0.12 \times NDVI & if\ NDVI \leq 0.8 \\ 0.095 \times NDVI & if\ NDVI > 0.8 \end{cases} \tag{7}$$

$$B_{VV} = \begin{cases} 0.70 \times NDVI & if\ NDVI \leq 0.8 \\ 0.56 \times NDVI & if\ NDVI > 0.8 \end{cases} \tag{8}$$

For VH polarization:

$$A_{VH} = \begin{cases} 0.05 \times NDVI & if\ NDVI \leq 0.8 \\ 0.04 \times NDVI & if\ NDVI > 0.8 \end{cases} \tag{9}$$

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


$$B_{VH} = \begin{cases} 1.45 \times NDVI & if\ NDVI \leq 0.8 \\ 1.16 \times NDVI & if\ NDVI > 0.8 \end{cases} \quad (10)$$

Distinct WCM parameters were applied for NDVI values above and below the 0.8 threshold. These dynamic parameterizations allowed the vegetation scattering and attenuation effects to vary smoothly with vegetation conditions.

2.2.2 Ensemble learning

The ISMN and remote sensing data were compiled into a structured tabular format, where each row represents a specific location and acquisition time, and each column corresponds to a predictor variable or the observed soil moisture value. To effectively capture the complex, nonlinear relationships between multi-source remote sensing features and in-situ soil moisture observations, this study employed three complementary machine learning models: TabNet, Random Forest, and XGBoost. TabNet is a deep learning architecture specifically designed for tabular datasets (Arik and Pfister, 2021). It employs attention-based feature selection and sparse representation to enhance interpretability and effectively model complex patterns in high-dimensional tabular data derived from remote sensing (Khaliq et al., 2025; Triana-Martinez et al., 2025). Random Forest (Breiman, 2001), an ensemble of decision trees, offers high robustness to overfitting, effective modeling of nonlinear relationships, and strong performance in the presence of noise and collinearity (Belgiu and Drăguţ, 2016; Liu et al., 2025). XGBoost (Chen and Guestrin, 2016), a regularized gradient boosting framework, has demonstrated superior predictive accuracy across a wide range of applications. Its advantages include efficient handling of missing data, built-in regularization, and the capacity to capture complex feature interactions through additive model training (Aydin and Ozturk, n.d.; Deng and Lumley, 2024; Karthikeyan and Mishra, 2021).

The SM estimation workflow is illustrated in Figure 2. We trained TabNet, Random Forest, and XGBoost models using remote sensing features and ISMN ground observations, with 80% of the entire dataset randomly allocated for training and 20% for testing. A 5-fold cross-validation was performed on the entire training dataset.

To optimize the hyperparameters of each model, we employed Optuna, an open-source hyperparameter optimization framework that uses a flexible and efficient sampling-based approach to automate tuning (Akiba et al., 2019). In addition to individual model optimization, Optuna was used to determine the optimal ensemble weights for combining the outputs of TabNet, XGBoost, and Random Forest, thereby integrating TabNet's deep learning feature extraction, XGBoost's gradient boosting, and Random Forest's handling of nonlinear relationships.

Model accuracy was assessed using the coefficient of correlation ($R^2$, eq(11)), bias (eq(12)), root mean square error (RMSE, eq(13)), and unbiased root mean square error (ubRMSE, eq(14)). Finally, the optimized ensemble model was used to generate

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


10-m resolution SM maps from January 2016 onward, with real-time updates.

$$R^2 = 1 - \frac{\sum (Y_i - \hat{Y}_i)^2}{\sum (Y_i - \bar{Y}_i)^2} \tag{11}$$

$$Bias = \frac{1}{n}\sum_{i=1}^{n} (\hat{Y}_i - Y_i) \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2} \tag{13}$$

$$ubRMSE = \sqrt{RMSE^2 - Bias^2} \tag{14}$$

Where, $\hat{Y}_i$ is the predicted value, $Y_i$ is the observed value, $\bar{\hat{Y}}_i$ is the mean of predicted values, $\bar{Y}$ is the mean of observed values, n is the number of observations.

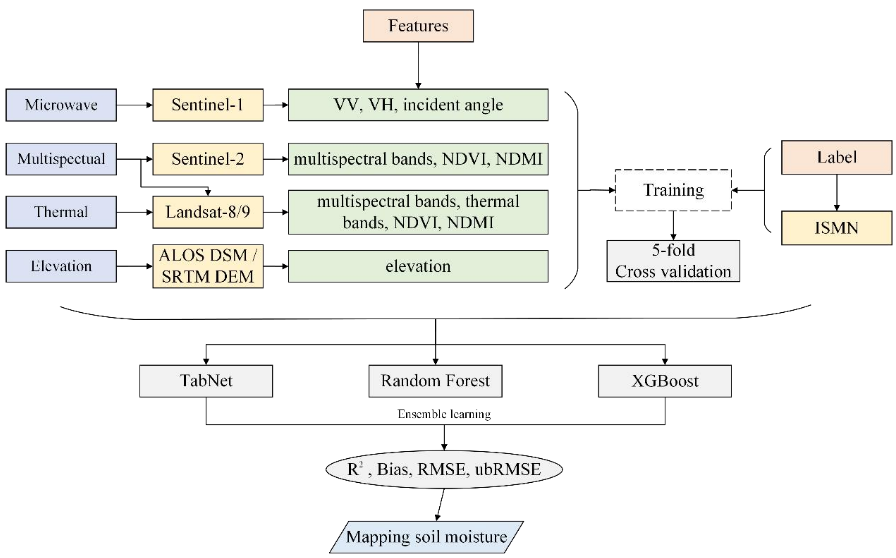



Fig. 2. Workflow of SM estimation from multi-source datasets.

## 2. Results

### 3.1 Model performance

Table 3 presents the average accuracy metrics obtained from 5-fold cross-validation for each model. Among the individual models, TabNet achieved the highest performance, with an $R^2$ of 0.7756 and the lowest RMSE (0.0506 m³/m³), followed closely by Random Forest ($R^2$ = 0.7702, RMSE = 0.0510 m³/m³). XGBoost showed comparatively lower accuracy, with an $R^2$ of 0.6792 and a higher RMSE of 0.0603 m³/m³. In contrast, the ensemble model outperformed all individual models, achieving the highest $R^2$ (0.8344) and the lowest RMSE (0.0433 m³/m³), as well as near-zero bias and the lowest unbiased RMSE (ubRMSE). These results highlight the advantage

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025

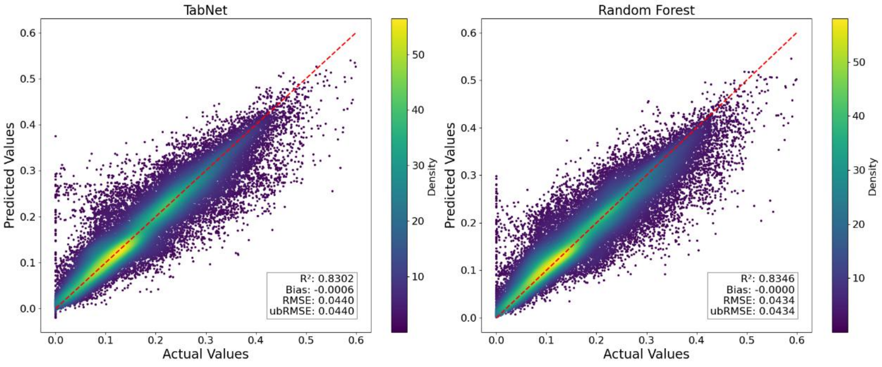
Open Access Earth System Science Data Discussions

of ensemble learning in integrating the complementary strengths of deep learning (TabNet), decision trees (Random Forest), and gradient boosting (XGBoost) to enhance the accuracy and robustness of soil moisture predictions.

Table 3. Average accuracy metrics from 5-fold cross-validation for each model.

| Model | $R^2$ | Bias | RMSE ($m^3/m^3$) | ubRMSE ($m^3/m^3$) |
|---|---|---|---|---|
| TabNet | 0.7756 | -0.0003 | 0.0506 | 0.0506 |
| RF | 0.7702 | 0.0002 | 0.0510 | 0.0510 |
| XGBoost | 0.6792 | 0.0001 | 0.0603 | 0.0603 |
| Ensemble | 0.8344 | -0.0001 | 0.0433 | 0.0433 |

Predicted versus actual soil moisture values on the 20% test set demonstrate strong agreement across all models, with the ensemble model showing the highest accuracy (Figure 3). The ensemble predictions were generated using a weighted combination of the three models, with weights of 0.56 for Random Forest, 0.43 for TabNet, and 0.26 for XGBoost. The ensemble model achieved an $R^2$ of 0.8591, a near-zero bias (–0.0002 $m^3/m^3$), and the lowest RMSE and ubRMSE (both 0.0401 $m^3/m^3$), indicating excellent predictive performance with minimal systematic error. XGBoost followed closely with an $R^2$ of 0.8586 $m^3/m^3$ and RMSE of 0.0401 $m^3/m^3$, while Random Forest and TabNet yielded slightly lower accuracies ($R^2$ = 0.8346 $m^3/m^3$ and 0.8302 $m^3/m^3$; RMSE = 0.0434 $m^3/m^3$ and 0.0442 $m^3/m^3$, respectively). However, the ensemble model exhibited the tightest clustering and lowest dispersion, suggesting that it effectively leveraged the complementary predictive capabilities of its constituent algorithms.



https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025

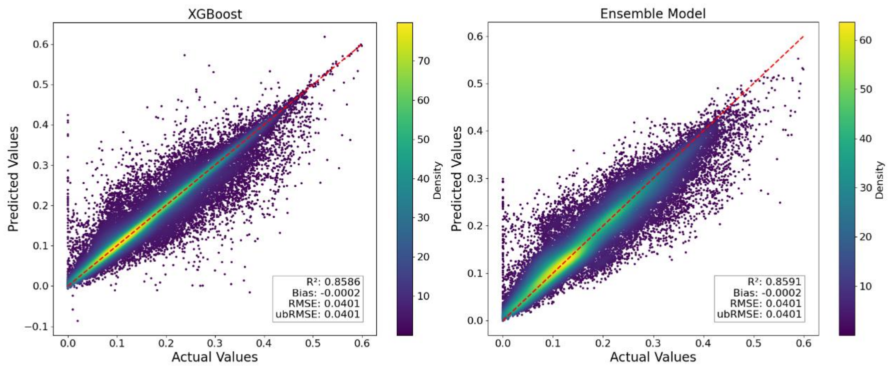




Fig. 3 Scatter density plots of predicted versus actual surface soil moisture on the test set for (a) TabNet, (b) Random Forest, (c) XGBoost, and (d) the Ensemble model.

## 3.2 Feature importance

Feature importance analysis was conducted across the three individual models—Random Forest, TabNet, and XGBoost—as well as their ensemble to better understand the contributions of each input variable to soil moisture prediction (Figure 4). Across all models, the Digital Surface Model (DSM) emerged as the most influential predictor, highlighting the dominant role of topography in governing soil moisture dynamics. Longitude and latitude were also consistently ranked among the top features, particularly in the ensemble and Random Forest models. These spatial coordinates likely serve as proxies for geospatial trends associated with climate zones, soil texture, and land management practices.

The ensemble model combined predictions from TabNet, Random Forest, and XGBoost using weighted contributions. It emphasized DSM and geolocation (longitude and latitude) as the three most important features, suggesting strong influence from landscape-driven and region-specific moisture patterns. Following these, vegetation-related indicators such as Landsat-based NDMI and NDVI were ranked highly, capturing the relationship between plant water content and soil moisture. A range of spectral reflectance bands from both Landsat (e.g., B3, B4, B11) and Sentinel-2 (e.g., B4, B8, B11) were also assigned substantial importance, reflecting their utility in detecting surface wetness and vegetation vigor. While SAR-derived features (e.g., VV and VH backscatter) were generally ranked lower in the ensemble, their inclusion still contributed valuable information in certain contexts, especially under low vegetation cover.

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025



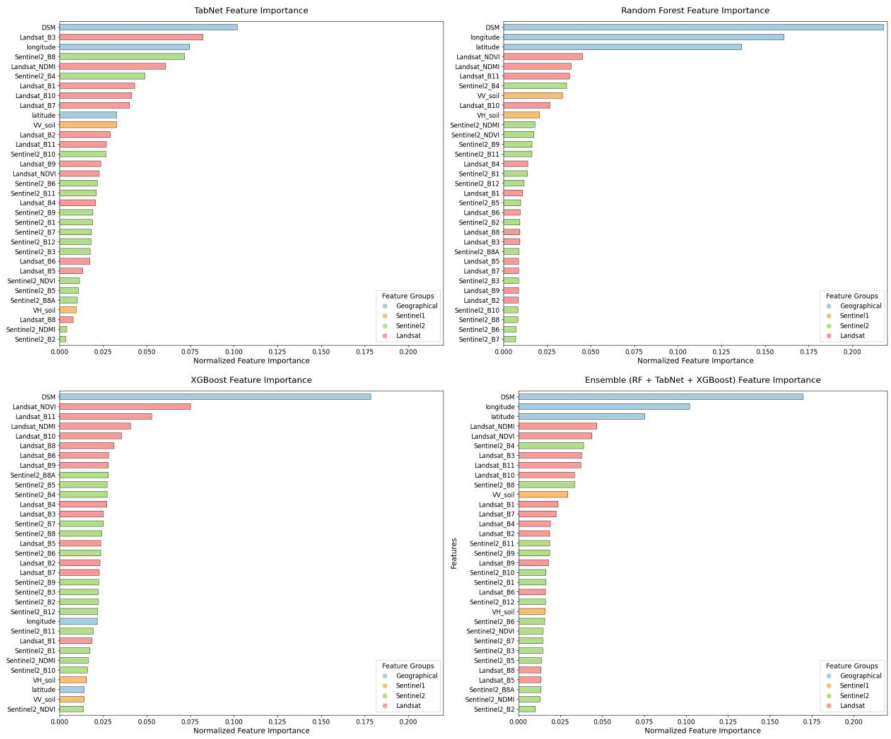



Fig. 4 Feature importance of each model.

## 3.3 Comparison to other soil moisture datasets

### 3.3.1 Compare to SMAP-HydroBlocks

The SMAP-HydroBlocks (SMAP-HB) dataset is a high-resolution, satellite-based surface soil moisture product developed at 30 m spatial resolution for the conterminous United States from 2015 to 2019. It was generated by integrating high-resolution land surface modeling, radiative transfer modeling, machine learning, and SMAP satellite microwave observations through a scalable cluster-based merging framework. The dataset was validated using measurements from 1,192 in situ observational sites. SMAP-HB achieved a median temporal correlation of $0.73 \pm 0.13$ and a median Kling-Gupta Efficiency (KGE) of $0.52 \pm 0.20$, indicating marked improvements relative to the standard SMAP Level 3 products.

To assess the relative accuracy of the proposed GSSM-10 dataset, both GSSM-10 and SMAP-HB were evaluated against independent in situ soil moisture measurements obtained from the ISMN. The evaluation metrics summarized in Table 5. The GSSM-10 dataset demonstrated strong agreement with ground observations, achieving an $R^2$ of 0.8601, a near-zero bias of $-0.0003\ m^3/m^3$, and RMSE and ubRMSE values of $0.0406\ m^3/m^3$. In contrast, the SMAP-HB dataset exhibited poor

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


correspondence with the in-situ measurements, yielding a negative $R^2$ of −0.4253, a bias of −0.0075 m³/m³, and RMSE and ubRMSE values of 0.1296 m³/m³ and 0.1293 m³/m³, respectively.

These findings suggest that GSSM-10 offers substantially improved predictive performance over SMAP-HB when benchmarked against independent field observations. The low bias and minimal random error observed in GSSM-10 indicate its robustness in capturing surface soil moisture variabilitys.

Table 4. Validation of SMAP-HB and GSSM-10 using 2569 in situ soil moisture observations from ISMN.

| Location | $R^2$ | Bias (m³/m³) | RMSE (m³/m³) | ubRMSE (m³/m³) |
|---|---|---|---|---|
| GSSM-10 | 0.8601 | -0.0003 | 0.0406 | 0.0406 |
| SMAP-HB | -0.4253 | -0.0075 | 0.1296 | 0.1293 |

### 3.3.1 Compared to S²MP

The S²MP (Sentinel-1/Sentinel-2-derived soil moisture product), developed by Baghdadi et al. (2017), was designed to estimate surface soil moisture at the plot scale by coupling radar backscatter from Sentinel-1 with vegetation indices derived from Sentinel-2 imagery using a neural network inversion approach. The product has been validated against in situ measurements collected across several regions, including parts of France, Morocco, Germany, and the United States. It achieved a RMSE of approximately 5 vol.%, demonstrating a high level of accuracy in agricultural regions with vegetation cover. In addition to ground-based validation, S²MP has been compared with other widely used soil moisture products, including SMOS, SMAP, ASCAT, and Copernicus-SSM. These comparisons revealed that S²MP consistently outperforms other products in terms of accuracy when benchmarked against ground observations. Furthermore, S2MP exhibited strong spatial and temporal consistency with precipitation data from the Global Precipitation Mission (GPM), suggesting that it captures realistic hydrological patterns.

Given the small sample size ($n = 14$), the $R^2$ was not computed, as it is sensitive to data distribution and may not yield statistically meaningful results under such conditions. Instead, three error metrics were used: bias, RMSE, ubRMSE.

The results of the validation are summarized in Table 5. The S2MP product exhibited a slight negative bias of −0.0180 m³/m³, with an RMSE of 0.0373 m³/m³ and a ubRMSE of 0.0326 m³/m³. In comparison, GSSM-10 showed a small positive bias of 0.0246 m³/m³, an RMSE of 0.0404 m³/m³, and a slightly lower ubRMSE of 0.0320 m³/m³. These findings suggest that both products demonstrate reasonable agreement with ground observations and are capable of capturing surface soil

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025



moisture dynamics with similar levels of accuracy, even under a limited number of validation instances. However, the spatial and temporal applicability of the two products differs significantly. S2MP is limited in geographic scope, covering only selected agricultural plots within parts of France, Morocco, Germany, and the United States. In contrast, GSSM-10 is a globally available product, providing surface soil moisture estimates at 10-meter resolution from 2016 to present. This extensive spatial and temporal coverage makes GSSM-10 more suitable for operational applications in regions where in situ data are sparse and where S2MP is unavailable, thus offering broader utility for global soil moisture monitoring and large-scale environmental assessments.

Table 5. Validation of S2MP and GSSM-10 using 14 in-situ soil moisture observations from USCRN.

| Location | Bias ($m^3/m^3$) | RMSE ($m^3/m^3$) | ubRMSE ($m^3/m^3$) |
|---|---|---|---|
| GSSM-10 | -0.0180 | 0.0373 | 0.0326 |
| SMAP-HB | -0.0246 | 0.0404 | 0.0320 |

# 4. Applications

## 4.1 After-fire Assessment

GSSM-10 offers valuable support for wildfire monitoring, assessment, and post-fire recovery planning by capturing high-resolution surface moisture dynamics. For instance, the soil moisture maps of a wildfire-affected region in northern Los Angeles demonstrate a clear contrast between November 10, 2024 and March 22, 2025. The pre-fire condition is illustrated in the Sentinel-2 true color image (Figure 5a) and the corresponding soil moisture map (Figure 5b). The post-fire condition, captured after the Eaton Fire that ignited on January 7, 2025, is shown in the true color image (Figure 5c) and soil moisture map (Figure 5d). The fire burned approximately 14,000 acres in the forested foothills of Los Angeles County, causing extensive environmental damage. The post-fire maps exhibit extensive areas of low soil moisture (reddish-brown tones), highlighting a widespread and persistent surface dryness following the wildfire disturbance. Such depletion of soil moisture can exacerbate erosion risks, delay vegetation recovery, and signal lasting ecosystem stress, especially on steep, fire-exposed slopes prone to debris flows. GSSM-10 can also be used for post-fire impact analysis, restoration monitoring, and land management decisions, providing critical support for assessing ecosystem recovery and mitigating secondary hazards in fire-prone regions.

## 4.2 Agriculture

The 10-meter resolution of GSSM-10 offers significant utility for irrigated agriculture,

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025

Open Access Earth System Science Data Discussions

enabling spatially explicit assessments of field-scale soil moisture conditions. This supports more efficient irrigation management, early detection of crop water stress, and optimized resource allocation. Figure 5 presents two representative case studies in agricultural regions. The first site, located in California's Central Valley, is shown in the Sentinel-2 true color image (Figure 5e) and the corresponding soil moisture map (Figure 5f) from November 3, 2016. The second site, located in South Africa, is depicted in the true color image (Figure 5g) and soil moisture map (Figure 5h) from October 20, 2023. The true color imagery delineates field boundaries and management zones, while the corresponding soil moisture maps reveal considerable intra- and inter-field variability in surface moisture conditions.

This spatial variability reflects differences in irrigation and management practices, soil types, vegetation status, and topography—factors that are difficult to capture using coarse-resolution satellite products. The high level of spatial detail provided by GSSM-10 enables applications such as detecting irrigated areas, planning variable-rate irrigation, targeting fertilizer application, detecting early water stress, and evaluating water conservation strategies.

### 4.3 Flood monitoring

The 10-meter resolution global soil moisture dataset developed in this study provides fine-scale information for analyzing the hydrological impacts of extreme weather events. Figure 5(i) shows a Sentinel-2 true color image, and Figure 5(j) presents the corresponding 10-m resolution soil moisture map for an agricultural area near Ravenna, Italy, on May 6, 2023, shortly after severe rainfall and regional flooding in the Emilia-Romagna region. The soil moisture map reveals widespread saturation, with values ranging from 0.25 to 0.32 $m^3/m^3$, consistent with the persistent rainfall that affected the region in early May. In contrast, Figure 5(k) displays a true color image and Figure 5(m) the corresponding soil moisture map for June 27, 2023, following a regional heatwave. This later image reveals markedly drier soils, reflecting the high atmospheric demand and reduced surface moisture after the extreme heat event, during which daily maximum temperatures exceeded 35 °C, rising 8 - 10 °C above long-term average. These observations highlight the dataset's ability to resolve intra-seasonal hydrological variability, effectively capturing both flood-induced soil saturation and subsequent surface drying.

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025

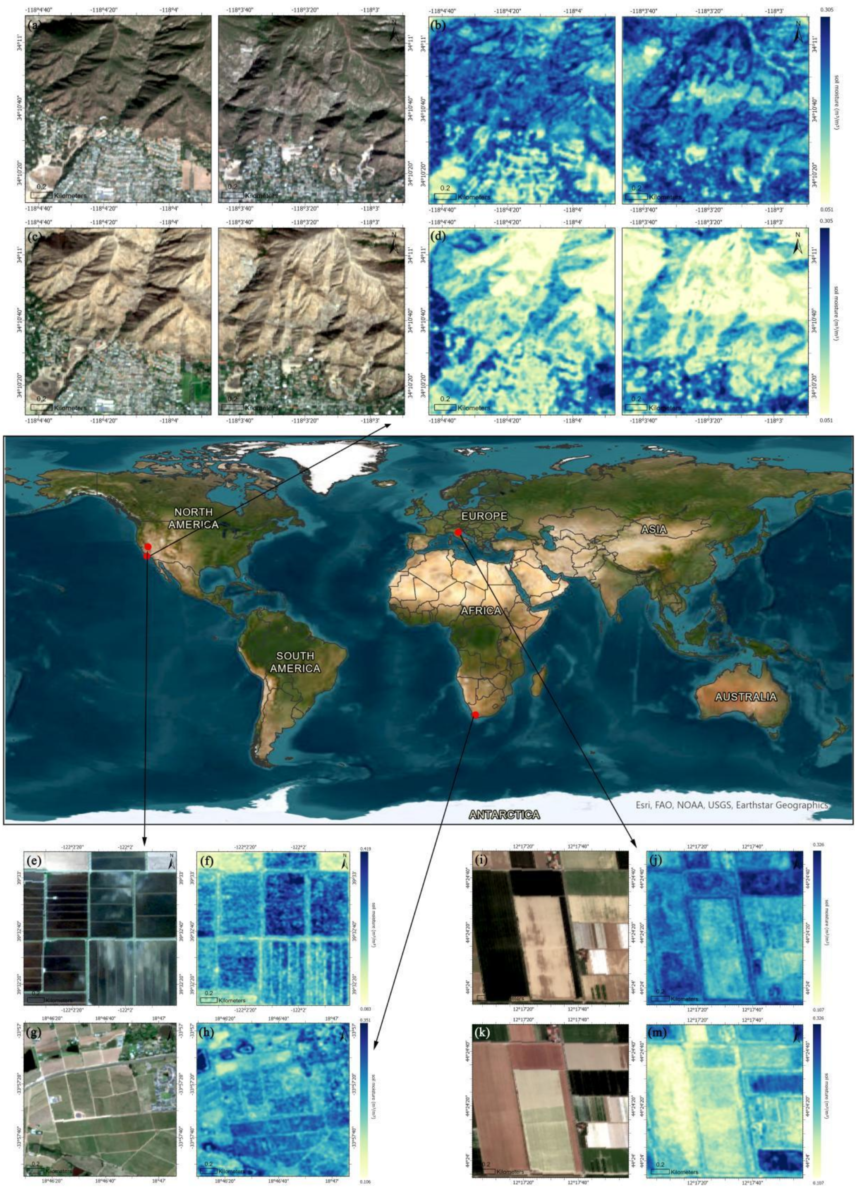




Fig. 5. (a) Sentinel-2 true-color image and (b) corresponding 10-m resolution soil moisture map of northern forest of Los Angeles on Nov. 10th, 2024 (before Eaton Fire); (c) Sentinel-2 true color image and (d) corresponding 10-m resolution soil moisture map of northern forest of Los Angeles on Mar. 22nd, 2025 (after Eaton Fire).

(e) Sentinel-2 true color image and (f) corresponding 10-m resolution soil moisture map of an agricultural field in California's Central Valley on Nov 3rd, 2016.
(g) True color image and (h) corresponding 10-m resolution soil moisture map of irrigated

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


fields in South Africa on Oct 20th, 2023.

(i) True color image and (j) corresponding 10-m resolution soil moisture map near Ravenna, Italy on May 6, 2023, shortly after extreme rainfall and regional flooding; (k) True color image and (m) corresponding 10-m resolution soil moisture map on June 27, 2023, following a regional heatwave.

## 5. Code and Data Availability

The source code and datasets associated with this research are publicly accessible. The code repository, titled Global-10-m-Surface-Soil-Moisture-Maps, is available on GitHub: https://github.com/RSNuo/Global-10-m-Surface-Soil-Moisture-Maps.git
.
In addition, the dataset and code have been archived on Zenodo to ensure long-term accessibility: https://doi.org/10.5281/zenodo.16956743 (Xu et al., 2025). The Zenodo record is published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Users are encouraged to access the repository and archive to reproduce the results and apply the models to new geographic regions or temporal periods.

## 6. Conclusion

This study presents the development of a global 10-meter resolution surface soil moisture (GSSM-10) dataset using a multi-sensor, ensemble machine learning framework. By integrating active microwave, multispectral, thermal, and geographical data from Sentinel-1, Sentinel-2, Landsat-8/9, and ALOS DSM, and by leveraging ensemble learning models including TabNet, Random Forest, and XGBoost, we produced a high-resolution product that substantially advances spatial detail and predictive accuracy of global SM mapping.

Our ensemble model demonstrated superior performance compared to individual models, achieving an $R^2$ of 0.8344 and an RMSE of 0.0433 $cm^3/cm^3$ during cross-validation, and an $R^2$ of 0.8591 and RMSE of 0.0401 $cm^3/cm^3$ on the test set. Feature importance analysis highlighted the key roles of geographical features in shaping soil moisture patterns. The GSSM-10 dataset was validated against in situ observations and showed comparable or superior accuracy to S2MP and SMAP-HydroBlocks. Unlike these regionally limited products, GSSM-10 offers global coverage, 10-m spatial resolution, and near real-time updates, making it a robust and scalable tool for global soil moisture monitoring and environmental applications.

The high spatial resolution of GSSM-10 enables a wide range of applications

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


previously hindered by the coarseness of existing SM products. We demonstrated the utility of this product in wildfire monitoring, irrigated agriculture, and flood analysis, highlighting its ability to resolve fine-scale hydrological variability associated with both natural and anthropogenic disturbances. The dataset also holds promise for supporting ecosystem restoration, climate resilience planning, and precision water management.

## Author contribution

N.X. designed the study, processed the satellite and ground datasets, developed the ensemble machine learning framework, and carried out the experiments. A.D. supervised the project, provided critical feedback on methodology and interpretation, and contributed to writing and revising the manuscript. A.A. supported the design of the modeling strategy, contributed to feature analysis and interpretation, and provided comments and revisions on the manuscript. All authors discussed the results and contributed to the final version of the paper.

## Reference

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: A Next-generation Hyperparameter Optimization Framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Presented at the KDD '19: The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, ACM, Anchorage AK USA, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701

Arik, S.Ö., Pfister, T., 2021. TabNet: Attentive Interpretable Tabular Learning. AAAI 35, 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

Attema, E.P.W., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio Science 13, 357–364. https://doi.org/10.1029/RS013i002p00357

Aydin, Z.E., Ozturk, Z.K., n.d. Performance Analysis of XGBoost Classifier with Missing Data.

Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V.K., Tuller, M., 2021. Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning. Remote Sensing of Environment 260, 112434. https://doi.org/10.1016/j.rse.2021.112434

Baghdadi, N., El Hajj, M., Zribi, M., Bousbih, S., 2017. Calibration of the Water Cloud Model at C-Band for Winter Crop Fields and Grasslands. Remote Sensing 9, 969. https://doi.org/10.3390/rs9090969

Baghdadi, N., Hajj, M.E., Zribi, M., 2019. An Operational High Resolution Soil Moisture Retrieval Algorithm Using Sentinel-1 Images, in: 2019 PhotonIcs & Electromagnetics Research Symposium - Spring (PIERS-Spring). Presented at the 2019 PhotonIcs & Electromagnetics Research Symposium - Spring (PIERS-Spring), IEEE, Rome, Italy, pp. 4086–4092. https://doi.org/10.1109/PIERS-Spring46901.2019.9017477

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24–31.

https://doi.org/10.1016/j.isprsjprs.2016.01.011
Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324
Chatterjee, S., Desai, A.R., Zhu, J., Townsend, P.A., Huang, J., 2022. Soil moisture as an essential component for delineating and forecasting agricultural rather than meteorological drought. Remote Sensing of Environment 269, 112833. https://doi.org/10.1016/j.rse.2021.112833
Chaudhary, S.K., Srivastava, P.K., Gupta, D.K., Kumar, P., Prasad, R., Pandey, D.K., Das, A.K., Gupta, M., 2022. Machine learning algorithms for soil moisture estimation using Sentinel-1: Model development and implementation. Advances in Space Research 69, 1799–1812. https://doi.org/10.1016/j.asr.2021.08.022
Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Presented at the KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco California USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785
Cheng, M., Li, B., Jiao, X., Huang, X., Fan, H., Lin, R., Liu, K., 2022. Using multimodal remote sensing data to estimate regional-scale soil moisture content: A case study of Beijing, China. Agricultural Water Management 260, 107298. https://doi.org/10.1016/j.agwat.2021.107298
Cui, Y., Zeng, C., Zhou, J., Xie, H., Wan, W., Hu, L., Xiong, W., Chen, X., Fan, W., Hong, Y., 2019. A spatio-temporal continuous soil moisture dataset over the Tibet Plateau from 2002 to 2015. Sci Data 6, 247. https://doi.org/10.1038/s41597-019-0228-x
Deng, Y., Lumley, T., 2024. Multiple Imputation Through XGBoost. Journal of Computational and Graphical Statistics 33, 352–363. https://doi.org/10.1080/10618600.2023.2252501
Denissen, J.M.C., Teuling, A.J., Pitman, A.J., Koirala, S., Migliavacca, M., Li, W., Reichstein, M., Winkler, A.J., Zhan, C., Orth, R., 2022. Widespread shift from ecosystem energy to water limitation with climate change. Nat. Clim. Chang. 12, 677–684. https://doi.org/10.1038/s41558-022-01403-8
Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S.I., Smolander, T., Lecomte, P., 2017. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. REMOTE SENSING OF ENVIRONMENT 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001
Droppers, B., Rakovec, O., Avila, L., Azimi, S., Cortés-Torres, N., De León Pérez, D., Imhoff, R., Francés, F., Kollet, S., Rigon, R., Weerts, A., Samaniego, L., 2024. Multi-model hydrological reference dataset over continental Europe and an African basin. Scientific Data 11, 1009. https://doi.org/10.1038/s41597-024-03825-9
Dubois, A., Teytaud, F., Verel, S., 2021. Short term soil moisture forecasts for potato crop farming: A machine learning approach. Computers and Electronics in Agriculture 180, 105902. https://doi.org/10.1016/j.compag.2020.105902
Eini, M.R., Massari, C., Piniewski, M., 2023. Satellite-based soil moisture enhances the reliability of agro-hydrological modeling in large transboundary river basins. Science of The Total

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025



Environment 873, 162396. https://doi.org/10.1016/j.scitotenv.2023.162396

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proc. IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Fan, D., Zhao, T., Jiang, X., García-García, A., Schmidt, T., Samaniego, L., Attinger, S., Wu, H., Jiang, Y., Shi, J., Fan, L., Tang, B.-H., Wagner, W., Dorigo, W., Gruber, A., Mattia, F., Balenzano, A., Brocca, L., Jagdhuber, T., Wigneron, J.-P., Montzka, C., Peng, J., 2025. A Sentinel-1 SAR-based global 1-km resolution soil moisture data product: Algorithm and preliminary assessment. Remote Sensing of Environment 318, 114579. https://doi.org/10.1016/j.rse.2024.114579

Fatima, E., Kumar, R., Attinger, S., Kaluza, M., Rakovec, O., Rebmann, C., Rosolem, R., Oswald, S.E., Samaniego, L., Zacharias, S., Schrön, M., 2024. Improved representation of soil moisture processes through incorporation of cosmic-ray neutron count measurements in a large-scale hydrologic model. Hydrol. Earth Syst. Sci. 28, 5419–5441. https://doi.org/10.5194/hess-28-5419-2024

Gebrechorkos, S., Leyland, J., Slater, L., Wortmann, M., Ashworth, P.J., Bennett, G.L., Boothroyd, R., Cloke, H., Delorme, P., Griffith, H., Hardy, R., Hawker, L., McLelland, S., Neal, J., Nicholas, A., Tatem, A.J., Vahidi, E., Parsons, D.R., Darby, S.E., 2023. A high-resolution daily global dataset of statistically downscaled CMIP6 models for climate impact analyses. Scientific Data 10, 611. https://doi.org/10.1038/s41597-023-02528-x

Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., Su, B., 2023. Global long term daily 1 km surface soil moisture dataset with physics informed machine learning. Sci Data 10, 101. https://doi.org/10.1038/s41597-023-02011-7

He, K., Zhao, W., Brocca, L., Quintana-Seguí, P., 2023. SMPD: a soil moisture-based precipitation downscaling method for high-resolution daily satellite precipitation estimation. Hydrol. Earth Syst. Sci. 27, 169–190. https://doi.org/10.5194/hess-27-169-2023

Humphrey, V., Berg, A., Ciais, P., Gentine, P., Jung, M., Reichstein, M., Seneviratne, S.I., Frankenberg, C., 2021. Soil moisture–atmosphere feedback dominates land carbon uptake variability. Nature 592, 65–69. https://doi.org/10.1038/s41586-021-03325-5

Karthikeyan, L., Mishra, A.K., 2021. Multi-layer high-resolution soil moisture estimation using machine learning over the United States. Remote Sensing of Environment 266, 112706. https://doi.org/10.1016/j.rse.2021.112706

Kerr, Y.H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans. Geosci. Remote Sensing 39, 1729–1735. https://doi.org/10.1109/36.942551

Khaliq, A., Khan, A., Jan, S., Umair, M., Gulshair, A., Ali, A., Ali Shah, U., 2025. AI-Driven Smart Agriculture: An Integrated Approach for Soil Analysis, Irrigation, and Crop-Fertilizer Recommendations. IEEE Access 13, 141124–141138. https://doi.org/10.1109/ACCESS.2025.3594162

Klein, C., Taylor, C.M., 2020. Dry soils can intensify mesoscale convective systems. Proc. Natl. Acad. Sci. U.S.A. 117, 21132–21137. https://doi.org/10.1073/pnas.2007998117

Leonarduzzi, E., Maxwell, R.M., Mirus, B.B., Molnar, P., 2021. Numerical Analysis of the Effect

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025



of Subgrid Variability in a Physically Based Hydrological Model on Runoff, Soil Moisture, and Slope Stability. Water Resources Research 57, e2020WR027326. https://doi.org/10.1029/2020WR027326

Lesinger, K., Tian, D., 2025. Skillful subseasonal soil moisture drought forecasts with deep learning-dynamic models. Nature Communications 16, 7461. https://doi.org/10.1038/s41467-025-62761-3

Li, W., Migliavacca, M., Forkel, M., Denissen, J.M.C., Reichstein, M., Yang, H., Duveiller, G., Weber, U., Orth, R., 2022. Widespread increasing vegetation sensitivity to soil moisture. Nat Commun 13, 3959. https://doi.org/10.1038/s41467-022-31667-9

Li, Z., Yuan, Q., Yang, Q., Li, J., Zhao, T., 2024. Differentiable modeling for soil moisture retrieval by unifying deep neural networks and water cloud model. Remote Sensing of Environment 311, 114281. https://doi.org/10.1016/j.rse.2024.114281

Liu, Y., Yang, Y., Song, J., 2023. Variations in Global Soil Moisture During the Past Decades: Climate or Human Causes? Water Resources Research 59, e2023WR034915. https://doi.org/10.1029/2023WR034915

Liu, Y., Zha, Y., Ran, G., Zhang, Y., Shi, L., 2025. SMRFR: A global multilayer soil moisture dataset generated using Random Forest from multi-source data. Scientific Data 12, 1170. https://doi.org/10.1038/s41597-025-05511-w

Liu, Yongwei, Liu, Yuanbo, Wang, W., 2019. Inter-comparison of satellite-retrieved and Global Land Data Assimilation System-simulated soil moisture datasets for global drought analysis. Remote Sensing of Environment 220, 1–18. https://doi.org/10.1016/j.rse.2018.10.026

Lozac'h, L., Bazzi, H., Baghdadi, N., Hajj, M.E., Zribi, M., Cresson, R., 2020. Sentinel-1/Sentinel-2-Derived Soil Moisture Product At Plot Scale ($S^2$ MP), in: 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS). Presented at the 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), IEEE, Tunis, Tunisia, pp. 168–171. https://doi.org/10.1109/M2GARSS47143.2020.9105210

Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A., Herrero-Jiménez, C.M., 2016. Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index. Remote Sensing of Environment 177, 277–286. https://doi.org/10.1016/j.rse.2016.02.064

McNairn, H., Merzouki, A., Pacheco, A., Fitzmaurice, J., 2012. Monitoring Soil Moisture to Support Risk Reduction for the Agriculture Sector Using RADARSAT-2. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing 5, 824–834. https://doi.org/10.1109/JSTARS.2012.2192416

Mei, Y., Mai, J., Do, H.X., Gronewold, A., Reeves, H., Eberts, S., Niswonger, R., Regan, R.S., Hunt, R.J., 2023. Can Hydrological Models Benefit From Using Global Soil Moisture, Evapotranspiration, and Runoff Products as Calibration Targets? Water Resources Research 59, e2022WR032064. https://doi.org/10.1029/2022WR032064

Mkhwenkwana, A., Matongera, T.N., Blaauw, C., Mutanga, O., 2025. A critical review on the applications of Sentinel satellite datasets for soil moisture assessment in crop production. International Journal of Applied Earth Observation and Geoinformation 141, 104647. https://doi.org/10.1016/j.jag.2025.104647

Monteiro, A.T., Arenas-Castro, S., Punalekar, S.M., Cunha, M., Mendes, I., Giamberini, M., Marques Da Costa, E., Fava, F., Lucas, R., 2024. Remote sensing of vegetation and soil moisture content in Atlantic humid mountains with Sentinel-1 and 2 satellite sensor data. Ecological Indicators 163, 112123. https://doi.org/10.1016/j.ecolind.2024.112123
Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021
Nguyen, T.T., Ngo, H.H., Guo, W., Chang, S.W., Nguyen, D.D., Nguyen, C.T., Zhang, J., Liang, S., Bui, X.T., Hoang, N.B., 2022. A low-cost approach for soil moisture prediction using multi-sensor data and machine learning algorithm. Science of The Total Environment 833, 155066. https://doi.org/10.1016/j.scitotenv.2022.155066
Nijaguna, G.S., Manjunath, D.R., Abouhawwash, M., Askar, S.S., Basha, D.K., Sengupta, J., 2023. Deep Learning-Based Improved WCM Technique for Soil Moisture Retrieval with Satellite Images. Remote Sensing 15, 2005. https://doi.org/10.3390/rs15082005
O., S., Orth, R., 2021. Global soil moisture data derived through machine learning trained with in-situ measurements. Sci Data 8, 170. https://doi.org/10.1038/s41597-021-00964-1
Peng, J., Tanguy, M., Robinson, E.L., Pinnington, E., Evans, J., Ellis, R., Cooper, E., Hannaford, J., Blyth, E., Dadson, S., 2021. Estimation and evaluation of high-resolution soil moisture from merged model and Earth observation data in the Great Britain. Remote Sensing of Environment 264, 112610. https://doi.org/10.1016/j.rse.2021.112610
Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., Stenseth, N.Chr., 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution 20, 503–510. https://doi.org/10.1016/j.tree.2005.05.011
Porkka, M., Virkki, V., Wang-Erlandsson, L., Gerten, D., Gleeson, T., Mohan, C., Fetzer, I., Jaramillo, F., Staal, A., Te Wierik, S., Tobian, A., Van Der Ent, R., Döll, P., Flörke, M., Gosling, S.N., Hanasaki, N., Satoh, Y., Müller Schmied, H., Wanders, N., Famiglietti, J.S., Rockström, J., Kummu, M., 2024. Notable shifts beyond pre-industrial streamflow and soil moisture conditions transgress the planetary boundary for freshwater change. Nat Water 2, 262–273. https://doi.org/10.1038/s44221-024-00208-7
Qiao, L., Zuo, Z., Zhang, R., Piao, S., Xiao, D., Zhang, K., 2023. Soil moisture–atmosphere coupling accelerates global warming. Nat Commun 14, 4908. https://doi.org/10.1038/s41467-023-40641-y
Qing, Y., Wang, S., Yang, Z.-L., Gentine, P., 2023. Soil moisture−atmosphere feedbacks have triggered the shifts from drought to pluvial conditions since 1980. Communications Earth & Environment 4, 254. https://doi.org/10.1038/s43247-023-00922-2
Rawat, K.S., Singh, S.K., Ray, R.L., 2021. An integrated approach to estimate surface soil moisture in agricultural lands. Geocarto International 36, 1646–1664. https://doi.org/10.1080/10106049.2019.1678674
Sabaghy, S., Walker, J.P., Renzullo, L.J., Akbar, R., Chan, S., Chaubell, J., Das, N., Dunbar, R.S., Entekhabi, D., Gevaert, A., Jackson, T.J., Loew, A., Merlin, O., Moghaddam, M., Peng, Jian, Peng, Jinzheng, Piepmeier, J., Rudiger, C., Stefan, V., Wu, X., Ye, N., Yueh, S., 2020. Comprehensive analysis of alternative downscaled soil moisture products. REMOTE

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025




SENSING OF ENVIRONMENT 239. https://doi.org/10.1016/j.rse.2019.111586

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews 99, 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Skulovich, O., Gentine, P., 2023. A Long-term Consistent Artificial Intelligence and Remote Sensing-based Soil Moisture Dataset. Sci Data 10, 154. https://doi.org/10.1038/s41597-023-02053-x

Soares, P.M.M., Lima, D.C.A., 2022. Water scarcity down to earth surface in a Mediterranean climate: The extreme future of soil moisture in Portugal. Journal of Hydrology 615, 128731. https://doi.org/10.1016/j.jhydrol.2022.128731

Song, P., Zhang, Y., Guo, J., Shi, J., Zhao, T., Tong, B., 2022. A 1 km daily surface soil moisture dataset of enhanced coverage under all-weather conditions over China in 2003–2019. Earth Syst. Sci. Data 14, 2613–2637. https://doi.org/10.5194/essd-14-2613-2022

Sun, W., Zhou, S., Yu, B., Zhang, Y., Keenan, T., Fu, B., 2025. Soil moisture-atmosphere interactions drive terrestrial carbon-water trade-offs. Commun Earth Environ 6, 169. https://doi.org/10.1038/s43247-025-02145-z

Syed, T.H., Famiglietti, J.S., Rodell, M., Chen, J., Wilson, C.R., 2008. Analysis of terrestrial water storage changes from GRACE and GLDAS. Water Resources Research 44, 2006WR005779. https://doi.org/10.1029/2006WR005779

Triana-Martinez, J., Álvarez-Meza, A., Castellanos-Dominguez, G., 2025. Enhancing agricultural data interpretability and visualization with TabNet-driven feature extraction and Local Biplots. Results in Engineering 27, 106672. https://doi.org/10.1016/j.rineng.2025.106672

Vergopolan, N., Chaney, N.W., Pan, M., Sheffield, J., Beck, H.E., Ferguson, C.R., Torres-Rojas, L., Sadri, S., Wood, E.F., 2021. SMAP-HydroBlocks, a 30-m satellite-based soil moisture dataset for the conterminous US. Sci Data 8, 264. https://doi.org/10.1038/s41597-021-01050-2

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., De Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., 2013. The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. metz 22, 5–33. https://doi.org/10.1127/0941-2948/2013/0399

Wasko, C., Nathan, R., 2019. Influence of changes in rainfall and soil moisture on trends in flooding. Journal of Hydrology 575, 432–441. https://doi.org/10.1016/j.jhydrol.2019.05.054

Wyatt, B.M., Ochsner, T.E., Krueger, E.S., Jones, E.T., 2020. In-situ soil moisture data improve seasonal streamflow forecast accuracy in rainfall-dominated watersheds. Journal of Hydrology 590, 125404. https://doi.org/10.1016/j.jhydrol.2020.125404

Xu, L., Chen, N., Zhang, X., Moradkhani, H., Zhang, C., Hu, C., 2021. In-situ and triple-collocation based evaluations of eight global root zone soil moisture products. Remote Sensing of Environment 254, 112248. https://doi.org/10.1016/j.rse.2020.112248

Xu, N., Daccache, A., Ahmadi, A., Houtz, D., Perez-Barquero, F.P., 2025. Soil moisture estimation with microwave remote sensing: a systematic review and meta-analysis. International

https://doi.org/10.5194/essd-2025-511
Preprint. Discussion started: 16 October 2025


Journal of Digital Earth 18, 1–27. https://doi.org/10.1080/17538947.2025.2468413

Xu, N. (2025). GSSM-10 (Global 10-m Surface Soil Moisture) Derived from Multi-Sensor Data and Ensemble Learning. Zenodo. https://doi.org/10.5281/zenodo.16956743

Yao, Y., Liu, Y., Zhou, S., Song, J., Fu, B., 2023. Soil moisture determines the recovery time of ecosystems from drought. Global Change Biology 29, 3562–3574. https://doi.org/10.1111/gcb.16620

Zawadzki, J., Kędzior, M., 2016. Soil moisture variability over Odra watershed: Comparison between SMOS and GLDAS data. International Journal of Applied Earth Observation and Geoinformation 45, 110–124. https://doi.org/10.1016/j.jag.2015.03.005

Zhang, Y., Liang, S., Ma, H., He, T., Wang, Q., Li, B., Xu, J., Zhang, G., Liu, X., Xiong, C., 2023. Generation of global 1km daily soil moisture product from 2000 to 2020 using ensemble learning. Earth System Science Data 15, 2055–2079. https://doi.org/10.5194/essd-15-2055-2023

Zhang, Y., Parazoo, N.C., Williams, A.P., Zhou, S., Gentine, P., 2020. Large and projected strengthening moisture limitation on end-of-season photosynthesis. Proc. Natl. Acad. Sci. U.S.A. 117, 9216–9222. https://doi.org/10.1073/pnas.1914436117

Zheng, C., Jia, L., Zhao, T., 2023. A 21-year dataset (2000-2020) of gap-free global daily surface soil moisture at 1-km grid resolution. SCIENTIFIC DATA 10. https://doi.org/10.1038/s41597-023-01991-w

Zhou, K., Li, J., Zhang, T., Kang, A., 2021. The use of combined soil moisture data to characterize agricultural drought conditions and the relationship among different drought types in China. Agricultural Water Management 243, 106479. https://doi.org/10.1016/j.agwat.2020.106479