

Article

Machine Learning-Based Soil Moisture Inversion in Agro-Pastoral Areas Using Sentinel Data

Teng Ma ^{1,*}, Ling Han ², Zhihui Yuan ¹ , Quanming Liu ¹ and Danyun Zhao ³

¹ College of Water Conservancy and Civil Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China; yuanzhihui@imau.edu.cn (Z.Y.); nndlqm@imau.edu.cn (Q.L.)

² School of Land Engineering, Chang'an University, Xi'an 710064, China; hanling@chd.edu.cn

³ School of Civil and Transportation Engineering, Qinghai Minzu University, Xining 810007, China; 2025023@qhmu.edu.cn

* Correspondence: mateng@imau.edu.cn

Highlights

What are the main findings?

- Machine learning algorithms, especially boosting methods, estimated soil moisture in agro-pastoral areas with moderate accuracy.
- Vertically polarized Sentinel-1 backscattering coefficients are the most sensitive predictors, while optical remote sensing data and topographic factors also contribute substantially to soil moisture inversion.

What are the implications of the main findings?

- The proposed framework offers an effective, simplified approach for soil moisture inversion over heterogeneous and complex surfaces, where conventional methods often fail.
- By leveraging freely available Sentinel imagery and machine learning, the proposed method enables practical and efficient soil moisture monitoring across large spatial scales, thereby supporting agricultural and ecological management.

Abstract

Soil moisture is a critical regulator in Earth's energy cycle, exerting substantial influence on hydrological, ecological, and thermal processes. However, the spatio-temporal heterogeneity of soil moisture poses significant challenges for remote-sensing inversion, and conventional methods are constrained under complex surface conditions. In this study, Sentinel-1 and Sentinel-2 imagery were integrated with terrain data, and machine learning algorithms were employed to develop a soil moisture estimation model for the agro-pastoral transition zone. The primary objectives were to evaluate the accuracy of machine learning approaches for remote-sensing estimation of soil moisture under complex surface conditions, and to assess the suitability of different machine learning algorithms for this application. The results showed that ensemble methods, specifically boosting approaches, delivered low inversion errors and high R^2 values, whereas linear models showed limited accuracy. The vertically polarized microwave backscattering coefficients emerged as the parameters most sensitive to soil moisture variations, with optical bands and topographic factors also making substantial contributions. The overall inversion yielded a root mean square error (RMSE) of $0.066 \text{ cm}^3 \cdot \text{cm}^{-3}$. This study provides an effective approach for soil moisture inversion on complex surfaces, simplifies the modeling process, and offers a tool for soil moisture monitoring across large spatial scales.



Academic Editor: Konstantinos X. Soulis

Received: 5 September 2026

Revised: 15 October 2026

Accepted: 16 October 2026

Published: 18 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: soil moisture; machine learning; agro-pastoral transition zone; remote sensing

1. Introduction

Soil moisture plays a crucial role in governing hydrological, ecological, and thermal processes [1,2]. It directly affects biological productivity, soil erosion, and land surface energy-water exchanges. Over the past several decades, a global trend of soil moisture divergence has emerged, characterized by an increasing disparity in both temporal and spatial dimensions [3]. Climate change has resulted in two significant phenomena: an intensification of aridity in dryland areas and an elevated risk of waterlogging in humid regions. Seasonal drought conditions are deteriorating, as evidenced by more prominent warm-season moisture deficits and stronger interannual fluctuations [4,5]. In addition, the “soil memory” effect generates important atmospheric feedback loops. In the arid regions of Asia, spring droughts frequently exacerbate subsequent summer heatwaves, posing long-term challenges to food security, water resource management, and ecosystem stability [6–8].

Currently, monitoring or estimating soil moisture across spatial scales remains a significant challenge. Global or regional monitoring networks can provide continuous and highly precise measurements; however, their effectiveness is constrained by the significant spatial heterogeneity of soil texture [9,10]. This heterogeneity hinders the accurate acquisition of area-average soil moisture observations. Microwave remote sensing satellites have the potential for large-scale soil moisture estimation. Among microwave remote sensing techniques, synthetic aperture radar (SAR) outperforms passive microwave sensors [11,12]. Nevertheless, the accuracy of SAR-based soil moisture retrieval remains susceptible to factors including surface vegetation coverage, surface roughness, soil texture, and topographic conditions, all of which can introduce distortions in the retrieved soil moisture data. Optical remote sensing can partially characterize auxiliary surface properties, including vegetation and soil texture, but its integration with SAR data often leads to overly complex fusion models and may degrade overall retrieval performance [13,14].

The agro-pastoral transition zone has considerable spatial heterogeneity in surface roughness, vegetation composition, and soil texture, which hinders soil moisture retrieval using SAR [15]. To date, numerous SAR-based retrieval models have been developed. These models incorporate a variety of surface parameters, including root mean square height, correlation length, vegetation height, canopy water content, soil texture, and organic matter, to build functional relationships between backscattering coefficients and soil moisture. These models enable high-precision soil moisture estimation at relatively fine spatial scales. However, these models rely on a large number of hard-to-obtain surface parameters, which inevitably reduces the efficiency of the retrieval process [16]. In recent years, change-detection algorithms based on SAR backscattering coefficients have emerged as an active research area in soil moisture retrieval. This approach leverages the temporal variation in SAR backscattering coefficients to detect changes in soil moisture, which can partially mitigate the influence of multi-scale surface roughness and systematic bias in microwave remote sensing systems, thereby simplifying the parameterization of the inversion model. As a result, it has become an effective method for soil moisture retrieval under relatively complex surface conditions [17–19].

Machine learning algorithms demonstrate the capability to effectively extract diverse remote sensing features. This enables them to overcome the inherent limitations of conventional models, including inadequate predictive performance and restricted generalizability [20–22]. Consequently, they have emerged as a valuable approach for quantitative

remote sensing inversion. In recent years, numerous researchers have studied how to use machine learning algorithms combined with remote sensing data to estimate soil moisture. Reference [23] utilized Long Short-Term Memory (LSTM) networks in combination with Random Forests (RF) to estimate the soil moisture content in agricultural regions of northern Ghana, achieving favorable results. Reference [24] thoroughly evaluated the performance of several machine learning models, specifically RF, eXtreme Gradient Boosting (XGBoost), Support Vector Regression (SVR), Gaussian Process Regression (GPR), Partial Least Squares (PLS), and Multiple Linear Regression (MLR), for soil moisture retrieval from remote sensing data. The results demonstrated that RF and XGBoost achieved the best performance among all models. Collectively, these studies underscore the potential of machine learning algorithms in soil moisture inversion. Reference [25] compared three algorithms: RF, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The results showed that the RF model trained with the IM12 combination outperformed SVM and ANN in retrieving soil moisture information. Reference [26] used Sentinel-1 synthetic aperture radar (SAR) data in combination with climate and terrain data to construct a large-scale (Europe) surface soil moisture (SM) estimation model via the RF algorithm. After comparing the RF model's predictions with in situ measurements from the International Soil Moisture Network (ISMN) and other data sources, the model achieved a Pearson correlation coefficient (R) of 0.847, which indicates significantly higher accuracy than the official SMAP/Sentinel-1 products at 1-km and 3-km spatial resolutions. Reference [27] proposed a knowledge-guided deep learning method, which integrates the physical principles of the Water Cloud Model (WCM) into a Long Short-Term Memory (LSTM) network and uses Sentinel-1 SAR data to estimate field-scale soil moisture. The results indicate that this method reduces the uncertainty of SM inversion by $0.02 \text{ m}^3/\text{m}^3$. The aforementioned studies demonstrate that applying machine learning algorithms to remote sensing-based soil moisture estimation can significantly improve estimation accuracy. However, current research still has several limitations. First, most existing studies focus on areas with a single land cover type, such as farmland or bare land, while there is relatively little research on soil moisture in regions with diverse land cover types, such as an agro-pastoral area. Differences in land cover types can cause significant changes in the physical and chemical properties of soil, which may in turn reduce the efficiency and accuracy of soil moisture estimation. Second, there is a lack of systematic cross-comparison between different machine learning algorithms, meaning the optimal algorithm for this task has not yet been clearly determined.

To evaluate and compare the applicability of different machine learning algorithms for remote sensing-based soil moisture inversion in the agro-pastoral transition zone, this study selected the Shandian River Basin in China as the study area and integrated Sentinel-1 imagery, Sentinel-2 imagery, and auxiliary topographic data including digital elevation models, slope, and aspect. We then applied machine learning algorithms to construct models for estimating soil moisture content in the agro-pastoral transition zone. The performance of these models was evaluated against in situ measurements to assess their suitability under complex surface conditions. Additionally, the contribution of individual input features to model predictions was analyzed to clarify the influence of various surface factors on estimation performance. Overall, this study aimed to provide an effective methodology for soil moisture inversion over heterogeneous surfaces and to offer insights for large-scale and fine-resolution mapping.

2. Materials and Methods

2.1. Study Area

The study area is located in the Shandian River sub-basin of the Luan River Basin, spanning longitudes 115.6°E to 116.4°E and latitudes 41.7°N to 42.4°N (Figure 1). Specifically, a total of 34 soil moisture monitoring stations affiliated with the International Soil Moisture Network (ISMN) [28] are distributed across the Shandian Basin region, covering both cropland and grassland ecosystems. These stations perform hourly measurements of volumetric soil moisture at depths ranging from 0.03 m to 0.5 m.

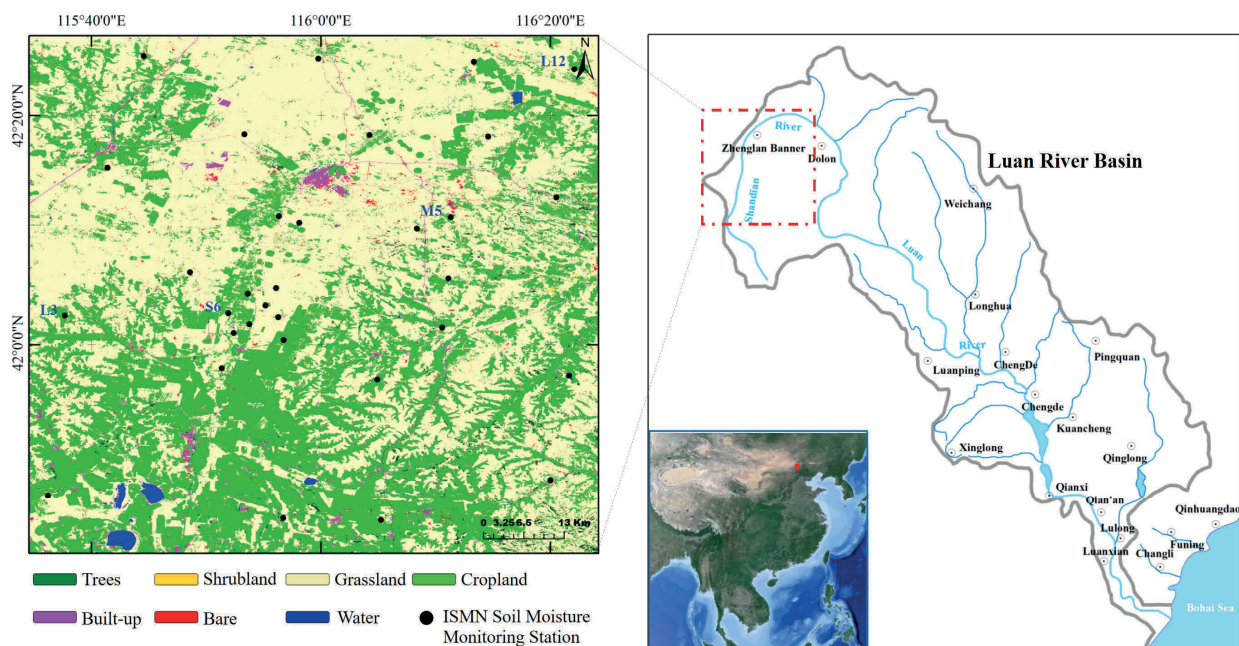


Figure 1. Geographical Location of the study area. The land cover map of the study area is derived from the ESA WorldCover 10 m 2020 v100 product.

Considering the substantial topographic relief in this mountainous region and aiming to minimize backscattering coefficient variations caused by changes in the SAR incidence angle, we selected Sentinel-1 scenes acquired along the same relative orbit that completely cover the study area. Sentinel-1 IW GRD and Sentinel-2 L2A images from September 2018 to December 2019, which covered all four seasons, were acquired to ensure the reliability of soil moisture inversion under various surface conditions. Data preprocessing was performed using the Sentinel Application Platform (SNAP, version 10.0.0, ESA). The processing sequence for Sentinel-1 data included the following: (1) application of the orbit file; (2) removal of thermal noise; (3) radiometric calibration and terrain flattening; (4) filtering of speckle noise; and (5) range-Doppler terrain correction. To alleviate residual speckle noise effects, the images were resampled to a coarser spatial resolution of 100 m $\times$ 100 m. Since Sentinel-2 images were already provided as Level-2A products, they were only resampled to match the spatial resolution of the Sentinel-1 images.

2.2. Auxiliary Data Preprocessing

Enhancing the efficiency of soil moisture retrieval from SAR data necessitates the generation of auxiliary datasets, specifically including (1) elevation; (2) local incidence angle (LIA); (3) slope; and (4) aspect.

All auxiliary data were derived from the Shuttle Radar Topography Mission (SRTM) 1-s DEM or calculated from it, and were resampled to match the spatial resolution of the Sentinel-1 and Sentinel-2 imagery.

2.3. Change Detection Algorithm

Since SAR backscattering coefficients are governed by surface roughness, slope, aspect, dielectric constant, soil texture, and organic matter content, most of which remain essentially invariant over extended periods at a given location, differencing of images acquired along identical orbital paths on different dates effectively cancels out these long-term factors. Given that the soil is frozen in winter and demonstrates relatively low backscattering coefficients, the Sentinel-1 image acquired on 9 January 2019 was chosen as the reference image, and all other images were subtracted from it to produce the change series.

$$\begin{cases} \Delta\gamma_{VV}^i = \gamma_{VV}^i - \gamma_{VV}^0 \\ \Delta\gamma_{VH}^i = \gamma_{VH}^i - \gamma_{VH}^0 \end{cases} \quad (1)$$

Here, γ_{VV}^i , γ_{VH}^i are the backscattering coefficients for the i -th period. γ_{VV}^0 , γ_{VH}^0 are the backscattering coefficients of the reference image on 9 January 2019. $\Delta\gamma_{VV}^i$, $\Delta\gamma_{VH}^i$ are the changes in the backscattering coefficients for the i -th period.

In light of the substantial spatial heterogeneity in soil types, terrain conditions, and land-use patterns throughout the study area, it is possible that identical differences in backscattering coefficients may correspond to varying magnitudes of soil moisture change. To address this ambiguity, we utilized the vertical polarization backscattering coefficients, the cross-polarization backscattering coefficients, and Sentinel-2 imagery to characterize the aforementioned surface properties, which subsequently facilitated the accurate inversion of soil moisture from backscattering coefficients. Nevertheless, we avoided applying differential processing to slope, aspect, incidence angle, or Sentinel-2 data. This is because these variables are expected to remain relatively stable during the study period, and their incorporation into the differencing scheme would impose an unnecessary computational burden without enhancing the retrieval performance.

2.4. In Situ Data

The soil moisture network in the Shandian River was established within the framework of the Soil Moisture Experiment in the Shandian River Basin of China, which was conducted from 18 July to 28 September 2018 [29]. This network covers an area of approximately 10,000 km², spanning the geographical coordinates of 115.5°E–116.5°E and 41.5°N–42.5°N. The 34 stations are arranged across three distinct spatial sampling scales: 100 km (large-scale), 50 km (medium-scale), and 10 km (small-scale). Each station is equipped with Decagon 5TM soil moisture sensors, which are installed at five specific measurement depths: 3 cm, 5 cm, 10 cm, 20 cm, and 50 cm. Additionally, among the 34 stations, 20 are equipped with HOBO tipping bucket rain gauges. This deployment is intended to facilitate land surface modeling and the validation of multiple satellite products, including those from SMAP, SMOS, and AMSR2, alongside downscaled soil moisture datasets. The terrain within the network area is relatively flat, and the land cover is predominantly composed of grassland and cropland.

2.5. Machine Learning Algorithms

When constructing and optimizing the machine learning model, we used PyCaret (version 3.3.2), an open-source low-code machine learning framework developed in Python (version 3.11.7). Its core function is to automate key workflow stages of machine learning, enabling researchers to quickly and efficiently build and deploy end-to-end projects with minimal coding. This library supports more than 25 regression algorithms, including Adaptive Boosting (AdaBoost), Light Gradient Boosting Machine (LGBM), Decision Tree (DT), RF, and XGBoost.

2.6. Training and Validation Datasets

In the process of model training and testing, in situ measurements from all 34 ground stations were utilized. Both Sentinel-1 data and field-observed soil moisture data were acquired on 11 September, 23 September, 5 October, 17 October, 29 October, 10 November, 22 November, 4 December, and 28 December 2018, as well as 9 January, 22 March, 3 April, 9 May, 21 May, 2 June, 13 August, 18 September, 24 October, 17 November, and 11 December 2019. The feature set contained 16 variables, including $\Delta\gamma_{VV}^0$, $\Delta\gamma_{VH}^0$, the LIA of Sentinel-1, and elevation, slope, and aspect. Sentinel-2 reflectance data from all bands except Bands 1, 9, and 10 were acquired on 5 September, 5 October, 29 October, 10 November, 24 November, 29 November 2018, as well as 8 January, 24 March, 3 April, 13 May, 23 May, 21 August, 25 September, 30 October, 19 November, 24 November 2019. We filtered out samples where Sentinel-2 images were contaminated by clouds or ground stations failed to obtain valid measurements. We then matched the closest available Sentinel-2 images with Sentinel-1 data to ensure all samples included both valid soil moisture measurements and complete image features. The samples from September 2018 to May 2019 of the 23-station dataset were used for training and 3-fold cross-validation. The processed dataset ultimately yielded 240 training samples. The training sample set was randomly split into three non-overlapping subsets (folds), each containing 80 samples. In each iteration, two of the three folds were used to train the model, and the remaining fold was reserved for validation. After the optimal model was identified via 3-fold cross-validation, all 240 samples were used for the final training and optimization of the model.

To evaluate the model's performance under various seasonal conditions, the trained model was applied to eight dates in 2019: 2 June, 13 August, 25 August, 18 September, 30 September, 24 October, 17 November, and 11 December. The retrieval results were tested using the 11 independent ground stations that were never used during model training, yielding a total of 88 samples.

Figure 2 presents the overall workflow of this study.

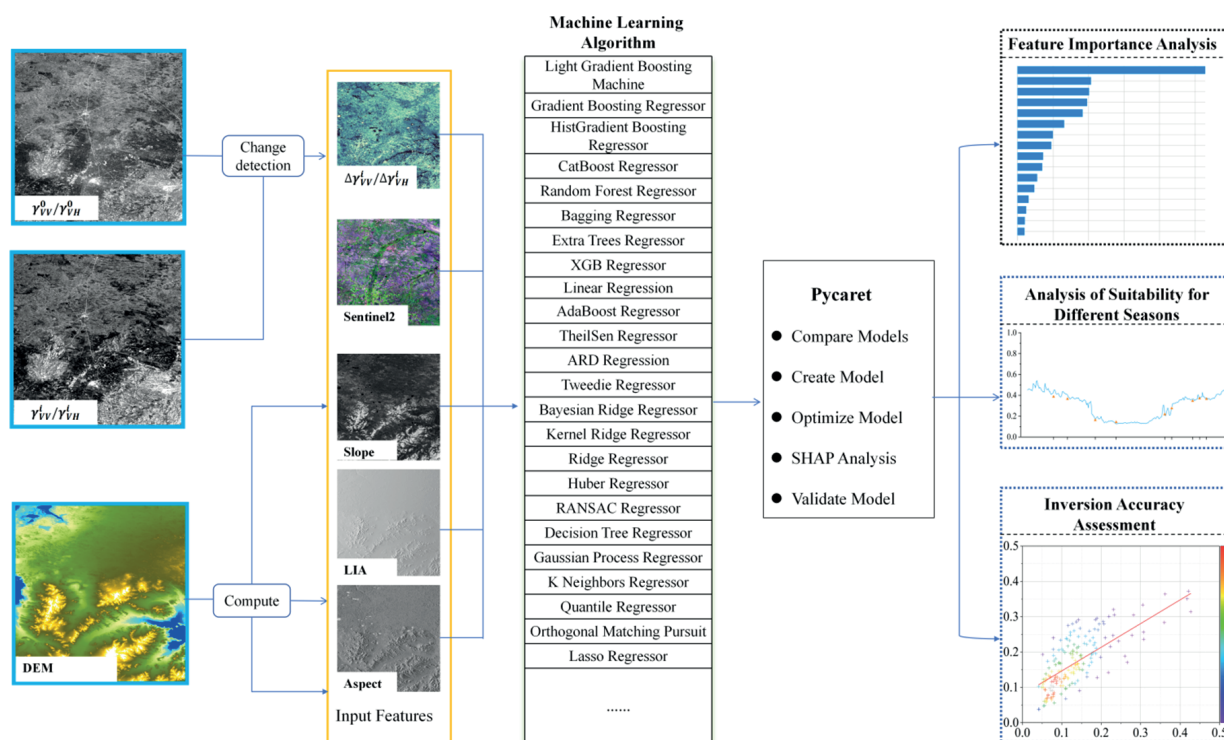


Figure 2. Workflow of the soil moisture retrieval.

2.7. Model Evaluation Criteria

The assessment of model performance was conducted through the utilization of three metrics: the mean absolute error (MAE), the root mean square error (RMSE), and R^2 . The conceptual definition of the MAE metric is presented as follows:

$$\text{MAE} = \frac{\sum |y_i - y'_i|}{n} \quad (2)$$

Here, n represents the total number of samples, y_i represents the i -th inverted value, and y'_i represents the i -th real value.

RMSE can be expressed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum (y_i - y'_i)^2}{n}} \quad (3)$$

R^2 can be expressed as follows:

$$R^2 = 1 - \frac{\sum (y_i - y'_i)^2}{\sum (y_i - \bar{y})^2} \quad (4)$$

Here, $\bar{y}$ represents the mean of the true values.

3. Results

3.1. Machine Learning Model Selection Results

The machine learning methods presented in Table 1 can be systematically classified into four principal categories: linear models and their derivatives, bagging-based ensemble methods, boosting-based ensemble methods, and others. The complete performance comparison of all tested models is provided in Table S1 of the Supplementary Materials. Linear models, such as linear regression, ridge regression, and lasso regression, determine target values through the identification of a linear combination of input features. Bagging-based ensemble methods, such as RF and Extra Trees Regressor (ETR), train multiple base learners and integrate their outputs. These approaches aim to reduce variance, thereby improving the model's stability and generalization performance. In contrast, boosting-based ensemble methods, such as LGBM, CatBoost, and AdaBoost, train weak learners sequentially. Each subsequent model attempts to correct the errors of its predecessor, thus reducing bias and enhancing the overall predictive capability. The remaining model types include decision tree (DT) and kernel methods.

Model performance comparisons show that the boosting ensemble methods achieved the lowest inversion error: their RMSEs were lower than those of other tested regression models, and their R^2 values were significantly higher.

Performance assessment of the models showed that boosting methods achieved relatively low inversion errors and high R^2 values, while linear models showed significantly lower accuracy. Specifically, boosting methods achieved almost identical average performance ($\text{MAE} = 0.034 \text{ cm}^3 \cdot \text{cm}^{-3}$, $\text{RMSE} = 0.042 \text{ cm}^3 \cdot \text{cm}^{-3}$, $R^2 = 0.623$), whereas linear models performed substantially worse ($\text{MAE} = 0.049 \text{ cm}^3 \cdot \text{cm}^{-3}$, $\text{RMSE} = 0.065 \text{ cm}^3 \cdot \text{cm}^{-3}$, $R^2 = 0.033$). Other model types, such as decision trees, also exhibited inversion performance similar to that of linear models.

Table 1. Performance metrics of machine learning models for soil moisture retrieval.

Model	MAE (cm ³ .cm ^{−3})	RMSE (cm ³ .cm ^{−3})	R ²	Algorithm Type
Light Gradient Boosting Machine	0.032	0.041	0.646	Ensemble Methods—Boosting
HistGradient Boosting Regressor	0.033	0.042	0.638	Ensemble Methods—Boosting
XGB Regressor	0.033	0.042	0.637	Ensemble Methods—Boosting
Gradient Boosting Regressor	0.033	0.042	0.627	Ensemble Methods—Boosting
AdaBoost Regressor	0.035	0.043	0.608	Ensemble Methods—Boosting
CatBoost Regressor	0.036	0.045	0.583	Ensemble Methods—Boosting
Bagging Regressor	0.035	0.045	0.580	Ensemble Methods—Bagging
Linear Regression	0.038	0.050	0.480	Linear Methods
Huber Regressor	0.039	0.051	0.455	Linear Methods
ARD Regression	0.040	0.051	0.453	Linear Methods
Quantile Regressor	0.039	0.052	0.435	Linear Methods
Random Forest Regressor	0.042	0.052	0.434	Ensemble Methods—Bagging
Bayesian Ridge Regressor	0.042	0.053	0.405	Linear Methods
Decision Tree Regressor	0.041	0.055	0.383	Decision Tree Methods
K Neighbors Regressor	0.043	0.055	0.367	Instance-Based Learning
SVR Regressor	0.049	0.058	0.297	Kernel Methods
Ridge Regressor	0.052	0.062	0.177	Linear Methods
Gaussian Process Regressor	0.050	0.063	0.184	Gaussian Processes
Extra Trees Regressor	0.052	0.063	0.177	Ensemble Methods—Bagging
Kernel Ridge Regressor	0.053	0.067	0.066	Kernel Methods
Tweedie Regressor	0.054	0.067	0.057	Linear Methods
Orthogonal Matching Pursuit	0.055	0.068	0.041	Linear Methods

3.2. Cross-Validation Results

The optimal model identified through comparative evaluation, specifically the LGBM, was designated as the soil moisture inversion model. Its performance was validated using 3-fold cross-validation, with results presented in Figure 3. For all three validation subsets, the RMSE remained below 0.05 cm³.cm^{−3}, while the R² exceeded 0.60 for each fold. These findings indicated that the model exhibited favorable estimation performance. The error distribution histogram demonstrated that the residuals of the retrieved soil moisture primarily fell within the range of −0.15 to 0.20 cm³.cm^{−3}, exhibiting an approximately normal distribution. This suggests that the principal factors contributing to the inversion errors are largely free from systematic bias.

3.3. Feature Importance Analysis

Figure 4 illustrates the feature importance analysis, which is divided into two subplots. Figure 4a presents the overall distribution of SHAP values, revealing that $\Delta\gamma_{VV}^i$ is the most significant feature for detecting changes in soil moisture content. Other relatively important features include $\Delta\gamma_{VH}^i$, Sentinel-2 bands B4, B6, and B12, as well as the slope derived from the DEM. The density scatter plot in Figure 4b further elucidates the SHAP distribution characteristics of each feature. The findings indicated that $\Delta\gamma_{VV}^i$ and $\Delta\gamma_{VH}^i$ exerted a positive influence on soil moisture increments, suggesting a directly proportional relationship between these variables and the rise in soil moisture. The band values of B4 and B6 demonstrated sensitivity to soil texture and crop growth. Specifically, an increase in the B4 band value generally corresponded to a decrease in soil moisture. The B12 band value, which corresponds to the short-wave infrared, showed a clear negative correlation with soil moisture content. Topographic variables, namely slope and aspect, were also negatively correlated with soil moisture increments, implying that increases in soil moisture in mountainous areas are comparatively lower than those in plains.

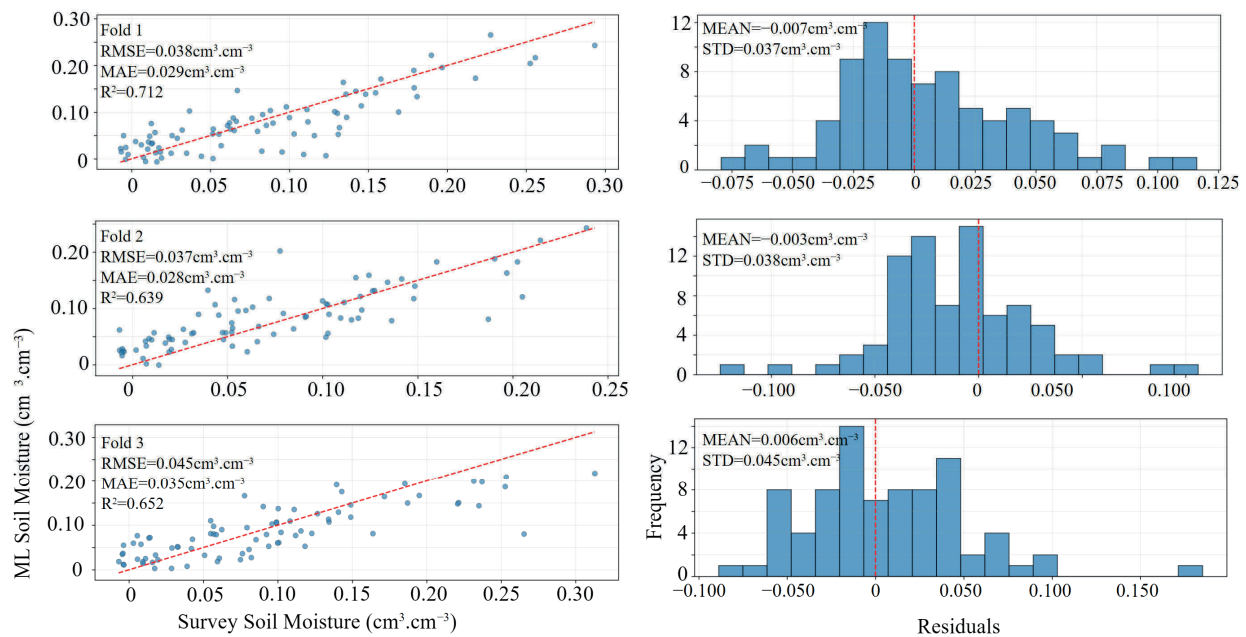


Figure 3. Three-fold cross-validation results and residual distribution.

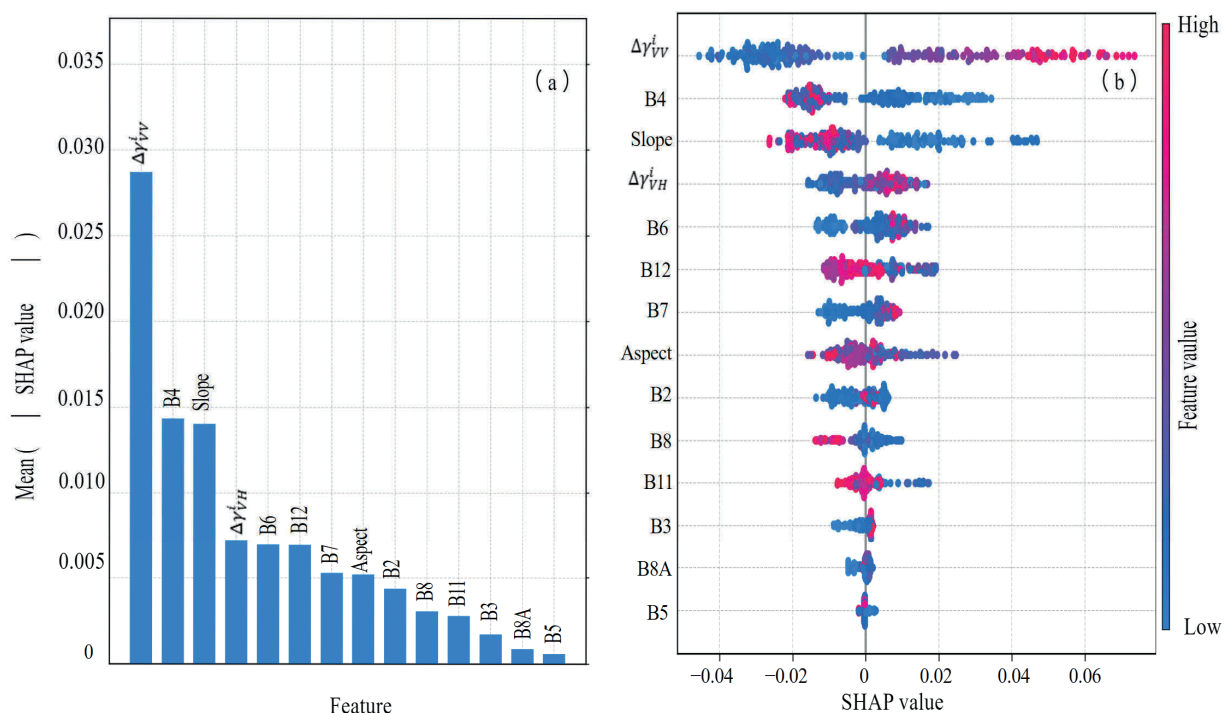


Figure 4. SHAP distribution of overall features. (a) SHAP Value Ranking; (b) SHAP summary plot.

3.4. Overall Inversion Error

Retrieval of soil moisture content was conducted for the validation set within the study area, with the results illustrated in Figure 5. The retrieval process yielded a mean absolute error (MAE) of $0.054 \text{ cm}^3.\text{cm}^{-3}$, a root mean square error (RMSE) of $0.066 \text{ cm}^3.\text{cm}^{-3}$, and an R^2 of 0.528. Figure 6 illustrates the error distribution for validation points categorized by different soil types. The study area demonstrates considerable diversity in soil types and notable spatial heterogeneity. Soil type represents a fundamental factor contributing to variations in soil texture and organic matter content. These validation points were not employed in the model training process, thus providing an unbiased assessment of the

model's generalization capability. The resulting RMSE and MAE values were comparable to those obtained from the cross-validation results, indicating that the model maintained a similar level of predictive accuracy for both training data and unseen data.

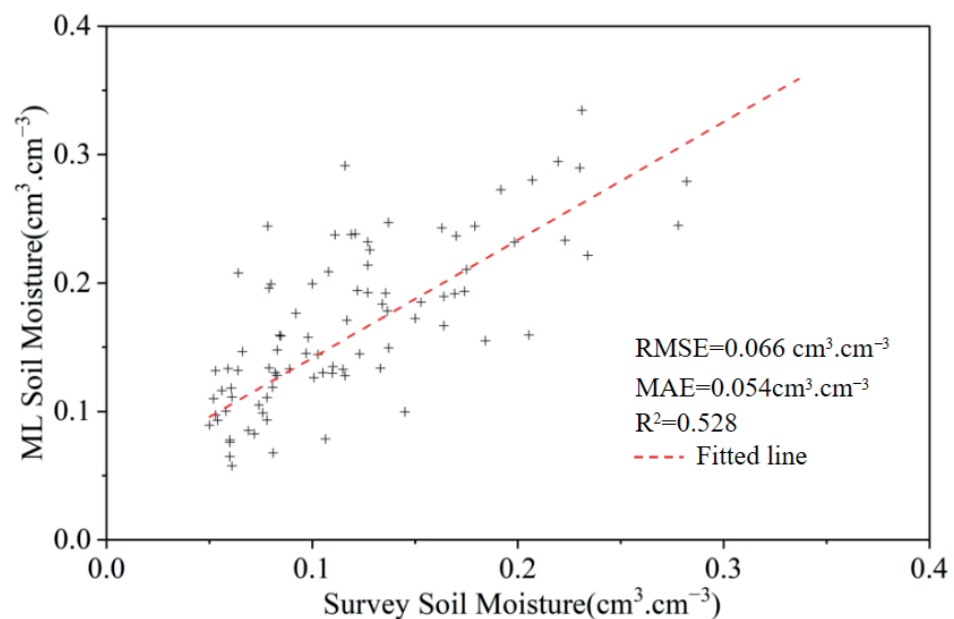


Figure 5. Comparison between machine learning and survey-based soil moisture.

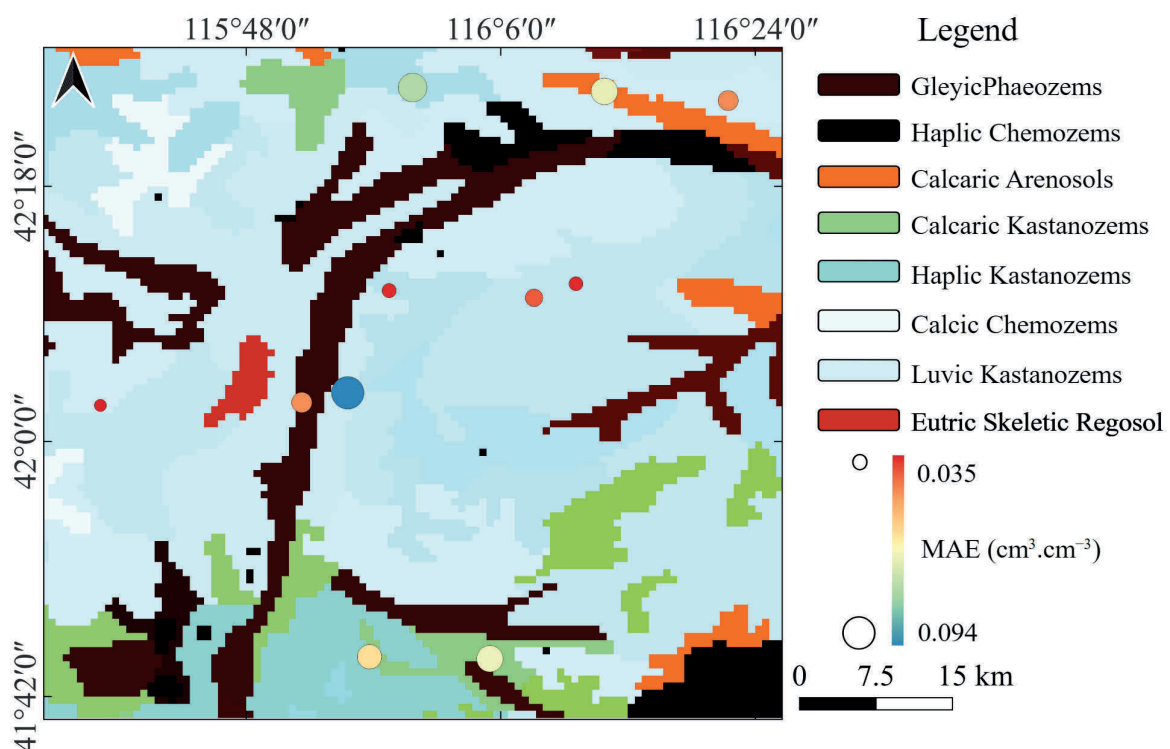


Figure 6. Map of the soil type and soil moisture retrieval errors. The soil texture map of the study area is derived from the HWSO version 2.1.

Figure 6 illustrates the spatial distribution of soil types within the study area. Kastanozems constitute the dominant soil group, encompassing the calcaric, haplic, and luvic subgroups. Furthermore, along the floodplains of the Shandian River, the deposition of fluvial sediments has resulted in the formation of Phaeozems, Chernozems, and Arenosols.

Phaeozems and Chernozems generally exhibit higher organic matter content and bulk density, and consequently, they possess a greater water-holding capacity. In contrast, Kastanozems and Arenosols demonstrate moderate water-holding capacity. As indicated by the distribution of test set errors in Figure 6, the MAE values do not show any obvious clustering. This indicates that when combined with optical data, SAR data can still deliver reliable soil moisture retrieval results across different soil types.

3.5. Inversion Accuracy in Different Seasons

Figure 7 illustrates the annual variation in the normalized difference vegetation index (NDVI) for four representative sampling sites. Site M5 is located within a cultivated area, whereas the remaining three sites are located in grassland regions. For all four sites, the NDVI values attain their maximum levels between August and September, with the values fluctuating between approximately 0.5 and 0.8, followed by a reduction to near zero between February and March. Throughout the entire study period, the variations in ground surface temperature generally exhibited trends parallel to the NDVI patterns, indicating a positive correlation between vegetation growth and temperature variations. Furthermore, within the study area, the soil annually enters a frozen condition in early November and transitions to a thawed state in early April of the subsequent year, remaining frozen during the intervening months.

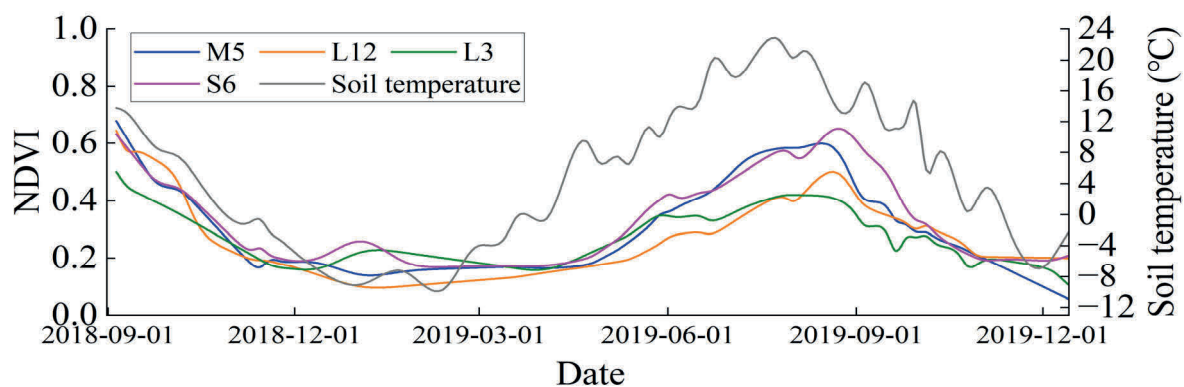


Figure 7. Trends of average soil temperature and typical sample point NDVI in the study area over the entire study period.

Figure 8 depicts the time series of retrieved soil moisture for four representative sampling sites. The retrieved values demonstrated a high level of consistency with in situ measurements during the winter and autumn seasons. The consistency becomes less pronounced during summer. During summer and early autumn, vegetation growth was observed at all four sites, with NDVI reaching its peak in autumn and exceeding 0.6. This suggests that the presence of surface vegetation, especially grass, might introduce a certain amount of error into the retrieval process. Although microwaves are generally considered capable of penetrating low-lying vegetation, our findings indicate that the influence of vegetation is non-negligible. In the study area, the surface soil remains frozen from mid-October to mid-April of the following year. This frozen period significantly affects the SAR backscattering coefficients; however, because the retrieval targets the unfrozen water fraction, this effect is explicitly represented in the model. The inversion results show that the retrieved unfrozen water content agrees closely with in situ measurements, demonstrating the effectiveness of the proposed retrieval of unfrozen water during winter.

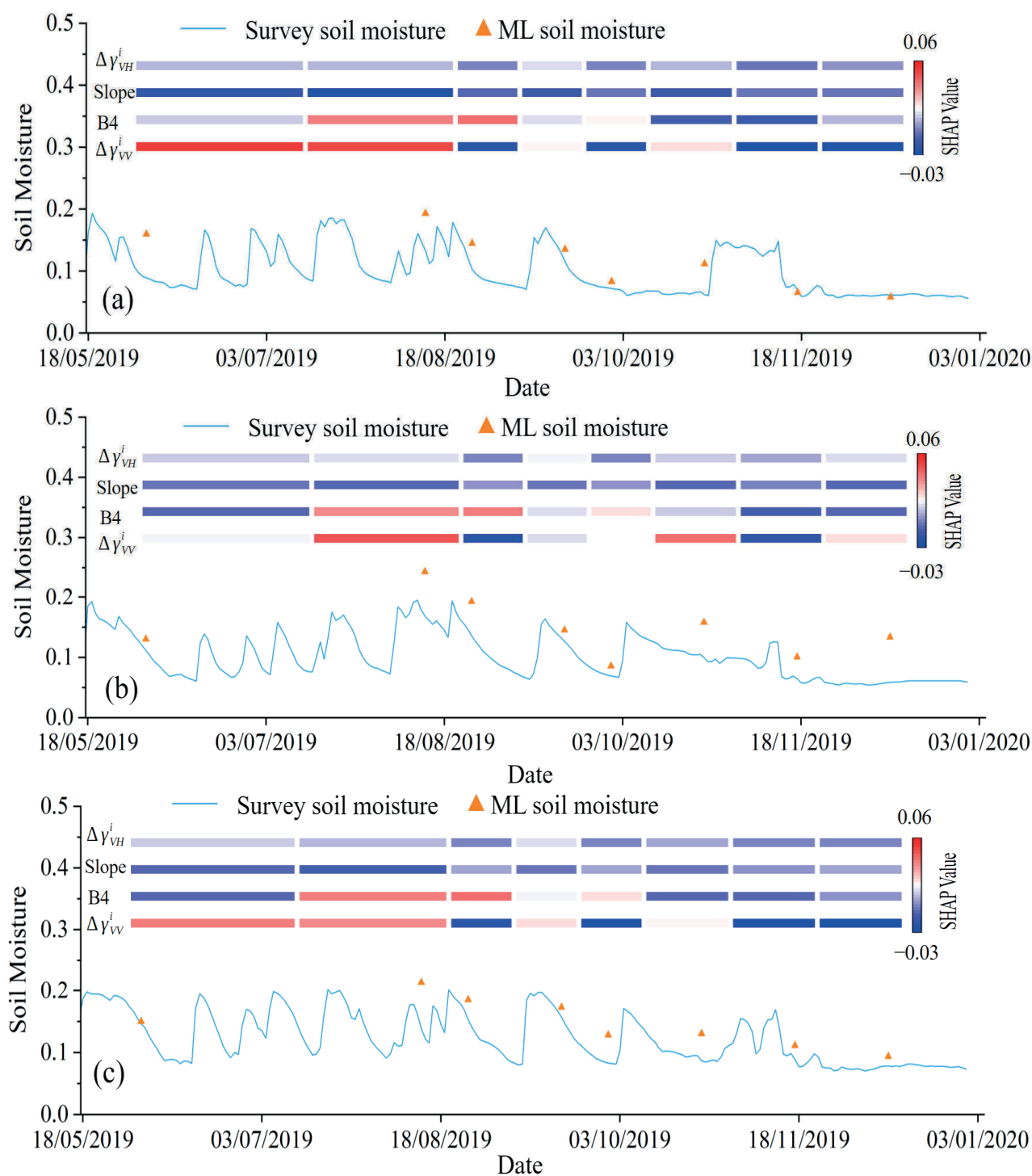


Figure 8. Cont.

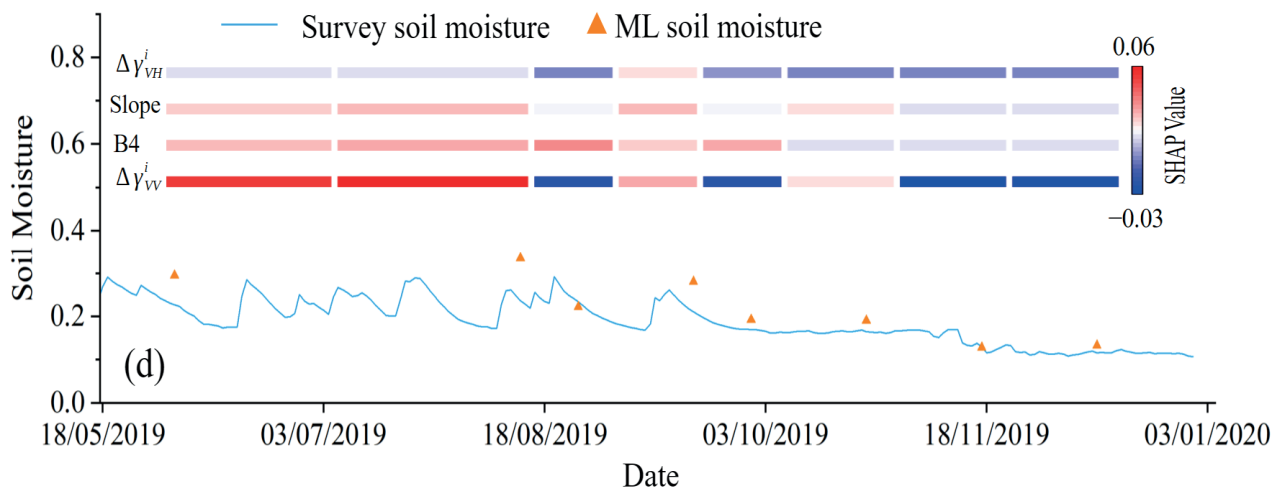


Figure 8. Soil moisture inversion results and SHAP value changes throughout the entire study period. (a) M5. (b) L12. (c) L3. (d) S6.

Figure 8 additionally illustrates the SHAP values of four essential features— $\Delta\gamma_{VV}^i$, Sentinel-2 band B4, $\Delta\gamma_{VH}^i$, and slope—in relation to soil moisture content. Among these features, the SHAP value trends of $\Delta\gamma_{VV}^i$ demonstrate consistency across the four sampling sites. Moreover, $\Delta\gamma_{VV}^i$ exhibits the most substantial variation with soil moisture, and a positive correlation is observed between its SHAP values and moisture changes. Conversely, the positive trend of $\Delta\gamma_{VH}^i$ is less pronounced. Under low moisture conditions or during winter, its SHAP values are negative, and as the moisture level increases to a high state, these values rise to slightly above zero. This phenomenon indicates that both $\Delta\gamma_{VV}^i$ and $\Delta\gamma_{VH}^i$ increase as soil moisture rises, yet the increase of $\Delta\gamma_{VV}^i$ is greater than that of $\Delta\gamma_{VH}^i$.

The slope parameter, which remains temporally invariant at each sampling site, exhibits values of 5.81, 3.54, 2.82, and 1.31 at sites M5, L12, L3, and S6, respectively. This indicates that site M5 is located on steeper terrain and site S6 is on gentler terrain. A comparison of slope SHAP values across the four sites revealed that S6, characterized by gentler slopes, demonstrated higher SHAP contributions than M5. L12 and L3 have medium slopes, with their slope SHAP values higher than those of M5 but lower than those of S6. Consistent with the findings in Figure 4, the slope made a positive contribution to increases in soil moisture when the gradient is small, whereas a negative contribution became evident for steeper slopes. Throughout the entire observation period, the SHAP values for the slope remained relatively stable. Nevertheless, during periods of high soil moisture, a significant decreasing trend in slope SHAP values was observed for M5, while S6 showed a marked increasing trend. This pattern suggests that soils with shallower gradients possess stronger water-holding capacity.

The B4 SHAP values at each sampling site exhibit relatively smooth temporal variations, generally oscillating around zero. During periods of elevated soil moisture, the SHAP values typically demonstrate an upward trend, which indicates that the B4 band reflectance responds to soil moisture fluctuations. As illustrated in Figure 4, an inverse correlation exists between the B4 band reflectance and soil moisture. The SHAP values of the B4 band at sampling site S6 exceed those at other sites, suggesting that the soil moisture content at S6 is higher, a finding confirmed by in situ measurements. In addition, during the peak NDVI period, the SHAP value of B4 is noticeably higher, which indicates that B4 responds to vegetation. This peak period also coincides with the time when soil moisture reaches its highest level in the study area.

4. Discussion

This study aimed to explore the viability of integrating machine learning algorithms with microwave and optical remote sensing data for the estimation of soil moisture in agro-pastoral transitional zones. The performance of four types of machine learning approaches, namely linear models, bagging ensembles, boosting ensembles, and other methods, represented by decision trees, was compared within the framework of soil moisture estimation. The findings indicated that boosting methods have better inversion effects, while linear methods generally performed worse. Moreover, feature importance and SHAP values were examined to clarify the contributions of individual features to soil moisture estimation. The analysis revealed that $\Delta\gamma_{VV}^i$, B4, slope, and $\Delta\gamma_{VH}^i$ were the four most significant features; among them, $\Delta\gamma_{VV}^i$ demonstrated significantly greater sensitivity to soil moisture than the rest.

Evaluation of the bagging algorithm showed that, except for specific inversion methods such as bagging regression whose inversion accuracy is comparable to that of boosting algorithms, the inversion accuracy of other bagging-based algorithms is lower than that of boosting algorithms. The bagging algorithm operates through the generation of multiple training subsets via bootstrap sampling. This approach demonstrates high robustness against noise and outliers, yielding reliable results under conditions of limited sample sizes or heterogeneous data quality. In contrast, boosting algorithms sequentially integrate multiple weak learners in the context of nonlinear problems, such as estimating soil moisture from remote sensing data. This approach presents notable advantages.

Previous studies on the SAR inversion of soil moisture have demonstrated that the vertical polarization channel displays the highest sensitivity to variations in soil moisture. Cross-polarization also reacts to these changes; however, it is more prone to interference from surface vegetation and roughness [30–32]. The feature importance analysis conducted in this study validated these findings: the vertically polarized backscattering coefficients exhibit a significant positive correlation with soil moisture, and the cross-polarization coefficients also show a positive correlation, albeit a weaker one. This phenomenon can be ascribed to the increase in the soil dielectric constant as the moisture content rises, which strengthens the microwave reflectivity and, in turn, boosts the backscattering power.

Assessment of slope impact on soil moisture dynamics represents a significant aspect in understanding hydrological processes. The most consistent finding in recent research indicates that soil moisture content typically exhibits a decreasing trend with increasing slope gradient [33–35]. On gentle slopes, the prolonged retention of precipitation facilitates infiltration, consequently resulting in elevated soil moisture levels. Conversely, on steep slopes, accelerated surface runoff occurs, which constrains infiltration capacity and ultimately leads to lower soil moisture levels. Our research also confirmed this conclusion, and we found that, compared with slope, aspect has a limited impact on soil moisture.

In the current study, a negative correlation was observed between the Sentinel-2 B4 band and soil moisture. This is mainly because chlorophyll strongly absorbs red light during the summer growing season, which reduces the reflectance in this band. This finding is consistent with numerous previous reports [36–39].

A multitude of studies have shown that soil moisture is influenced by diverse soil physicochemical properties, including texture, vegetation cover, and bulk density [40–44]. In the present study, optical remote sensing bands were employed as Supplementary Data to characterize these soil physical and chemical attributes. Additionally, terrain factors, specifically slope, aspect, and elevation, were incorporated into the model to enhance simulation accuracy. The study area is characterized by a complex mixture of soil types and substantial grass cover during the summer season. The inversion results indicated an overall R^2 of 0.528 and an RMSE of $0.066 \text{ cm}^3 \cdot \text{cm}^{-3}$. This accuracy is comparable to

that reported in previous studies over similar regions, suggesting that machine learning algorithms are suitable for this type of research. Although our study covered soil freeze–thaw cycles and the crop growing season, the sample size was limited, which remains a limitation. Future work will include more in situ data and SAR data from multiple orbits to further improve the robustness of the model.

Several assumptions underlying our approach should be acknowledged. Our method, which used a frozen winter image as a reference, assumes that soil roughness remains constant over the study period. This assumption may be violated in tilled croplands, where tillage operations and rainfall can alter surface roughness significantly. Previous studies have reported that while tillage can cause abrupt roughness changes, the subsequent temporal evolution of roughness under natural rainfall tends to stabilize relatively quickly [45]. Moreover, machine learning algorithms can learn to account for roughness effects through training, and VH polarization is sensitive to and can characterize surface roughness variations [46,47]. In our study, the validation accuracy ($\text{RMSE} = 0.066 \text{ cm}^3 \cdot \text{cm}^{-3}$) is consistent with the expected retrieval performance. Furthermore, a visual inspection of the retrieved soil moisture time series does not show systematic anomalies that could be attributed to tillage events. This suggests that the residual roughness effect is secondary to the soil moisture signal for our study area and period.

Regarding frozen periods, the retrieved variable represents the unfrozen (liquid) water content in frozen soil, rather than total soil water. Unfrozen water coexists with ice and strongly influences the soil dielectric properties, so microwave observations remain sensitive to this liquid fraction [48]. Previous studies have successfully retrieved unfrozen water content from satellite observations during frozen periods, with accuracy meeting mission requirements [49]. Therefore, the frozen-period values reported here are physically meaningful and valid for cold-region hydrological analysis.

5. Conclusions

The remote sensing inversion of soil moisture constitutes a complex systemic issue regulated by multiple interrelated parameters. These parameters include not only inherent soil properties but also terrain characteristics and vegetation structure. Conventional approaches for soil moisture retrieval have predominantly focused on the construction of simple linear or nonlinear models. Such models generally exhibit satisfactory performance in regions with homogeneous soil or relatively uniform surface conditions. However, their adaptability is constrained in areas with heterogeneous surface features. To address this challenge, the present study proposed the utilization of machine learning methods for soil moisture inversion, explicitly considering diverse factors such as soil type, topography, and vegetation cover. Furthermore, we undertook a systematic evaluation of the adaptability of various machine learning algorithms within this context.

This study indicated that boosting algorithms achieved the highest level of accuracy, demonstrating their potential for soil moisture inversion in complex terrain conditions. In contrast, linear methods exhibited poor performance and are inadequate for soil moisture retrieval in heterogeneous environments. The vertical polarization backscattering coefficients of SAR were identified as the most sensitive parameters to soil moisture variations. Nevertheless, optical remote sensing (e.g., bands B4, B6, and B12) and terrain factors, including slope, also contribute significantly to retrieval accuracy. Integration of optical and microwave imagery with terrain data enhances retrieval performance, as optical bands provide insights into soil texture characterizing large-scale soil moisture dynamics.

The combination of microwave and optical remote sensing data for soil moisture inversion typically resulted in relatively low retrieval accuracy during summer months. This limitation could be attributed to the presence of vegetation cover on the soil surface.

Although the vegetation types within the study area exhibit a degree of uniformity, further investigation was required to develop soil moisture models capable of functioning across diverse phenological phases and varying degrees of vegetation coverage. Moreover, despite the considerable diversity of soil types within the study area, the organic matter content is generally characterized by low levels. Consequently, the performance of machine learning algorithms in regions with elevated organic matter content necessitates additional validation.

Overall, this study demonstrated the potential of machine learning algorithms in estimating soil moisture over complex surfaces. Through the application of machine learning approaches, the requirement for comprehensive parameterization of soil and vegetation properties can be circumvented, thereby simplifying the modeling process and enhancing its flexibility. This provides a practical tool for soil moisture monitoring at large spatial scales.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/rs18203578/s1>. Table S1: Complete comparison of all tested machine learning models.

Author Contributions: Conceptualization, T.M.; methodology, T.M.; software, T.M.; validation, L.H.; formal analysis, D.Z.; investigation, Q.L.; resources, Q.L.; data curation, T.M.; writing—original draft preparation, T.M.; writing—review and editing, T.M.; visualization, Z.Y.; supervision, Z.Y.; project administration, Z.Y.; funding acquisition, T.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Program for Improving the Research Ability of Young Teachers of Inner Mongolia Agricultural University (grant number: BR240102); Natural Science Foundation of Inner Mongolia Autonomous Region of China (grant number: 2025MS05044); and Inner Mongolia Agricultural University High-level and Outstanding Doctoral Talent Introduction Research Startup Project (grant number: NDYB2022-57).

Data Availability Statement: Data presented in this study were derived from the following publicly available resources: Sentinel-1 and Sentinel-2 datasets (<https://www.sentinel-hub.com/>, accessed on 29 September 2025); soil moisture data (<https://ismn.earth/en/data/>, accessed on 5 January 2025); soil type (<https://iiasa.ac.at/models-tools-data/hwsd>, accessed on 18 November 2025).

Acknowledgments: During the preparation of this manuscript, the authors used DeepSeek-V4-Pro for language polishing and refinement of the manuscript's logical structure. This tool was not used for data generation, data analysis, interpretation of results, or formulation of conclusions. The authors reviewed and edited all AI-assisted outputs and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Vereecken, H.; Amelung, W.; Bauke, S.L.; Bogaen, H.; Brüggemann, N.; Montzka, C.; Vanderborght, J.; Bechtold, M.; Blöschl, G.; Carminati, A.; et al. Soil hydrology in the Earth system. *Nat. Rev. Earth Environ.* **2022**, *3*, 573–587. [\[CrossRef\]](#)
2. Rahmati, M.; Amelung, W.; Brogi, C.; Dari, J.; Flammini, A.; Bogaen, H.; Brocca, L.; Chen, H.; Groh, J.; Koster, R.D.; et al. Soil moisture memory: State-of-the-art and the way forward. *Rev. Geophys.* **2024**, *62*, e2023RG000828. [\[CrossRef\]](#)
3. Liu, Y.; Yang, Y.; Song, J. Variations in global soil moisture during the past decades: Climate or human causes. *J. Geophys. Res. Atmos.* **2023**, *128*, e2022JD037348.
4. Reinsch, S.; Robinson, D.A.; van Soest, M.A.J.; Keith, A.M.; Parry, S.; Tye, A.M. Temperate soils exposed to drought—Key processes, impacts, indicators, and unknowns. *Land* **2024**, *13*, 1759. [\[CrossRef\]](#)
5. Chen, S.; Yuan, X. The timing of detectable increases in seasonal soil moisture droughts under future climate change. *Earth's Future* **2024**, *12*, e2023EF004174. [\[CrossRef\]](#)
6. Wang, C.; Li, Z.; Chen, Y.; Ouyang, L.; Li, Y.; Sun, F.; Liu, Y.; Zhu, J. Drought-heatwave compound events are stronger in drylands. *Weather Clim. Extrem.* **2023**, *42*, 100632. [\[CrossRef\]](#)

7. Seo, Y.W.; Ha, K.J. Changes in land-atmosphere coupling increase compound drought and heatwaves over northern East Asia. *npj Clim. Atmos. Sci.* **2022**, *5*, 100. [\[CrossRef\]](#)
8. Ullah, I.; Mukherjee, S.; Syed, S.; Mishra, A.K.; Ayugi, B.O.; Aadhar, S. Anthropogenic and atmospheric variability intensifies flash drought episodes in South Asia. *Commun. Earth Environ.* **2024**, *5*, 267. [\[CrossRef\]](#)
9. Peng, C.; Zeng, J.; Chen, K.-S.; Ma, H.; Letu, H.; Zhang, X.; Shi, P.; Bi, H. Spatial representativeness of soil moisture stations and its influential factors at a global scale. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 4402915. [\[CrossRef\]](#)
10. Song, P.; Li, X.; Zhao, T.; Niu, S.; Sun, Y.; He, L. Quantification and correction of SMAP soil moisture retrieval biases induced by subpixel moisture and texture heterogeneity over China. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 5300816. [\[CrossRef\]](#)
11. Akash, M.; Mohan Kumar, P.; Bhaskar, P.; Deepthi, P.R.; Sukhdev, A. Review of estimation of soil moisture using active microwave remote sensing technique. *Remote Sens. Appl. Soc. Environ.* **2024**, *33*, 101118. [\[CrossRef\]](#)
12. Rahmati, M.; Balenzano, A.; Bechtold, M.; Brocca, L.; Fluhrer, A.; Jagdhuber, T.; Karamvavis, K.; Mengen, D.; Reichle, R.; Kim, S.-B.; et al. Soil moisture retrieval from Sentinel-1: Lessons learned after more than a decade in orbit. *Remote Sens. Environ.* **2026**, *333*, 115146. [\[CrossRef\]](#)
13. Teixeira, A.C.; Bakon, M.; Lopes, D.; Cunha, A.; Sousa, J.J. A systematic review on soil moisture estimation using remote sensing data for agricultural applications. *Sci. Remote Sens.* **2025**, *12*, 100328. [\[CrossRef\]](#)
14. Muhuri, A.; Goita, K.; Magagi, R. Soil moisture retrieval during crop growth cycle using satellite SAR time series. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 9517–9534. [\[CrossRef\]](#)
15. Nguyen, H.H.; Cho, S.; Choi, M. Spatial soil moisture estimation in agro-pastoral transitional zone based on synergistic use of SAR and optical-thermal satellite images. *Agric. For. Meteorol.* **2022**, *312*, 108719. [\[CrossRef\]](#)
16. Verhoest, N.E.C.; Lievens, H.; Wagner, W.; Álvarez-Mozos, J.; Moran, M.S.; Mattia, F. On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. *Sensors* **2008**, *8*, 4212–4248. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Montzka, C.; Brocca, L.; Chen, H.; Das, N.N.; Dasgupta, A.; Rahmati, M.; Jagdhuber, T. AI in soil moisture remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2026**, *146*, 104500. [\[CrossRef\]](#)
18. Li, Z.; Zhao, L.; Wang, L.; Zou, D.; Liu, G.; Hu, G.; Du, E.; Xiao, Y.; Liu, S.; Zhou, H.; et al. Retrieving soil moisture in the permafrost environment by Sentinel-1/2 temporal data on the Qinghai–Tibet Plateau. *Remote Sens.* **2022**, *14*, 5966. [\[CrossRef\]](#)
19. Nativel, S.; Ayari, E.; Rodriguez-Fernandez, N.; Baghdadi, N.; Madelon, R.; Albergel, C.; Zribi, M. Hybrid methodology using Sentinel-1/Sentinel-2 for soil moisture estimation. *Remote Sens.* **2022**, *14*, 2434. [\[CrossRef\]](#)
20. Graldi, G.; Zardi, D.; Vitti, A. Retrieving soil moisture at the field scale from Sentinel-1 data over a semi-arid Mediterranean agricultural area. *Remote Sens.* **2023**, *15*, 2997. [\[CrossRef\]](#)
21. Wang, Y.; Zhao, H.; Fan, J.; Wang, C.; Ji, X.; Jin, D.; Chen, J. A review of Earth's surface soil moisture retrieval models via remote sensing. *Water* **2023**, *15*, 3757. [\[CrossRef\]](#)
22. Abbes, A.B.; Jarray, N.; Farah, I.R. Advances in remote sensing based soil moisture retrieval: Applications, techniques, scales and challenges for combining machine learning and physical models. *Artif. Intell. Rev.* **2024**, *57*, 224. [\[CrossRef\]](#)
23. Tefera, M.L.; Zeleke, E.B.; Pirastru, M.; Melesse, A.M.; Seddaiu, G.; Awada, H. Satellite-based machine learning for soil moisture prediction and land conservation practice assessment in West African drylands. *Remote Sens.* **2025**, *17*, 3651. [\[CrossRef\]](#)
24. Ali, A. Machine learning approaches for soil moisture prediction: Enhancing agricultural water management with integrated data. *Model. Earth Syst. Environ.* **2025**, *12*, 16. [\[CrossRef\]](#)
25. Hou, C.; Tan, M.L.; Li, L.; Zhang, F. Comparison of three machine learning algorithms for retrieving soil moisture information from Sentinel-1A SAR data in northwest Shandong plain, China. *Adv. Space Res.* **2024**, *74*, 75–88. [\[CrossRef\]](#)
26. Bulut, U.; Mohammadi, B.; Duan, Z. Estimation of surface soil moisture from Sentinel-1 synthetic aperture radar imagery using machine learning method. *Remote Sens. Appl. Soc. Environ.* **2024**, *36*, 101289. [\[CrossRef\]](#)
27. Yu, Y.; Filippi, P.; Bishop, T.F.A. Field-scale soil moisture estimated from Sentinel-1 SAR data using a knowledge-guided deep learning approach. In Proceedings of the IGARSS 2025-2025 IEEE International Geoscience and Remote Sensing Symposium, Brisbane, Australia, 3–8 August 2025; pp. 193–198.
28. Dorigo, W.; Himmelbauer, I.; Aberer, D.; Schremmer, L.; Petrakovic, I.; Zappa, L.; Preimesberger, W.; Xaver, A.; Annor, F.; Ardö, J.; et al. The International Soil Moisture Network: Serving Earth system science for over a decade. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 5749–5804. [\[CrossRef\]](#)
29. Zhao, T.; Shi, J.; Lv, L.; Xu, H.; Chen, D.; Cui, Q.; Jackson, T.J.; Yan, G.; Jia, L.; Chen, L.; et al. Soil moisture experiment in the Luan River supporting new satellite mission opportunities. *Remote Sens. Environ.* **2020**, *240*, 111680. [\[CrossRef\]](#)
30. Bousbih, S.; Zribi, M.; Lili-Chabaane, Z.; Baghdadi, N.; El Hajj, M.; Gao, Q.; Mougenot, B. Potential of Sentinel-1 radar data for the assessment of soil and cereal cover parameters. *Remote Sens.* **2017**, *9*, 1177.
31. Blanchard, A.J.; Newton, R.W.; Tsang, L.; Jean, B.R. Volumetric effects in cross-polarized airborne radar data. *IEEE Trans. Geosci. Remote Sens.* **1982**, *20*, 36–41. [\[CrossRef\]](#)
32. Zhou, X.; Wang, J.; Shan, B.; He, Y.; Xing, M. Sensitivity of multi-frequency and multi-polarization SAR to soil moisture at different depths in agricultural regions. *J. Hydrol.* **2025**, *660*, 133513. [\[CrossRef\]](#)

33. Bao, H.; Lann, T.; Ao, X.; Yang, L.; Lan, H.; Peng, J. Evolution characteristic of soil water in loess slopes with different slope angles. *Geoenviron. Disasters* **2024**, *11*, 43. [\[CrossRef\]](#)
34. Shen, Q.; Zhang, Y.; Yan, Y.; Dong, H.; Lei, W. Experimental study on infiltration characteristics of shallow rainwater in expansive soil slopes at different gradients. *Water* **2025**, *17*, 642. [\[CrossRef\]](#)
35. Hu, Y.; Tang, C.; Chen, X.; Zhao, Y.; He, H.; Li, M.; Zhang, J. Spatiotemporal variability of soil water content and its influencing factors on a microscale slope. *Agronomy* **2023**, *13*, 2035. [\[CrossRef\]](#)
36. Lin, R.; Chen, H.; Wei, Z.; Li, Y.; Zhang, B.; Sun, H.; Cheng, M. Improved surface soil moisture estimation model in semi-arid regions using the vegetation red-edge band sensitive to plant growth. *Atmosphere* **2022**, *13*, 930. [\[CrossRef\]](#)
37. Rabiei, S.; Jalilvand, E.; Tajrishy, M. A method to estimate surface soil moisture and map the irrigated cropland area using Sentinel-1 and Sentinel-2 data. *Sustainability* **2021**, *13*, 11355. [\[CrossRef\]](#)
38. Attarzadeh, R.; Amini, J.; Notarnicola, C.; Greifeneder, F. Synergetic use of Sentinel-1 and Sentinel-2 data for soil moisture mapping at plot scale. *Remote Sens.* **2018**, *10*, 1285. [\[CrossRef\]](#)
39. Castaldi, F.; Chabrilat, S.; Don, A.; van Wesemael, B. Soil organic carbon mapping using LUCAS topsoil database and Sentinel-2 Data: An approach to reduce soil moisture and crop residue effects. *Remote Sens.* **2019**, *11*, 2121. [\[CrossRef\]](#)
40. Cardoso, E.J.B.N.; Vasconcellos, R.L.F.; Bini, D.; Miyauchi, M.Y.H.; dos Santos, C.A.; Alves, P.R.L.; de Paula, A.M.; Nakatani, A.S.; Pereira, J.d.M.; Nogueira, M.A. Soil health: Looking for suitable indicators. What should be considered to assess the effects of use and management on soil health? *Sci. Agric.* **2013**, *70*, 274–289. [\[CrossRef\]](#)
41. Smirnova, M.A.; Kozlov, D.N. Soil properties as indicators of soil water regime: A review. *Eurasian Soil Sci.* **2023**, *56*, 306–320. [\[CrossRef\]](#)
42. Zhu, Q.; Liao, K.; Xu, Y.; Yang, G.; Wu, S.; Zhou, S. Monitoring and prediction of soil moisture spatial–temporal variations from a hydrogeological perspective: A review. *Soil Res.* **2013**, *50*, 625–637. [\[CrossRef\]](#)
43. Xu, N.; Daccache, A.; Ahmadi, A.; Houtz, D.; Perez-Barquero, F.P. Soil moisture estimation with microwave remote sensing: A systematic review and meta-analysis. *Int. J. Digit. Earth* **2025**, *18*, 1–27. [\[CrossRef\]](#)
44. Rao, S.; Gupta, V.; Murugesan, S.; Zribi, M.; Muddu, S. An information proxy framework that explains why soil-specific radar calibration succeeds in sandy textures but fails in clay. *Remote Sens. Lett.* **2026**, *17*, 416–428. [\[CrossRef\]](#)
45. Callens, M.; Verhoest, N.E.C.; Davidson, M.W.J. Parameterization of tillage-induced single-scale soil roughness from 4-m profiles. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 878–888. [\[CrossRef\]](#)
46. Zeng, L.; Liu, Q.; Jing, L.; Lan, L.; Feng, J. Using Generalized Regression Neural Network to Retrieve Bare Surface Soil Moisture From Radarsat-2 Backscatter Observations, Regardless of Roughness Effect. *Front. Earth Sci.* **2021**, *9*, 646787. [\[CrossRef\]](#)
47. Balenzano, A.; Mattia, F.; Satalino, G.; Davidson, M.W.J. Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *4*, 439–450. [\[CrossRef\]](#)
48. Zhang, L.; Zhao, T.; Jiang, L.; Zhao, S. Estimate of phase transition water content in freeze-thaw process using microwave radiometer. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4248–4255. [\[CrossRef\]](#)
49. Wang, C.; Yang, N.; Zhao, T.; Xue, H.; Peng, Z.; Zheng, J.; Pan, J.; Yao, P.; Gao, X.; Yan, H.; et al. All-Season Liquid Soil Moisture Retrieval from SMAP. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 8258–8270. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.