remote sensing

MDPI

Article

Random Forest-Based Soil Moisture Estimation Using Sentinel-2, Landsat-8/9, and UAV-Based Hyperspectral Data

Hadi Shokati [1,2,*], Mahmoud Mashal [1], Aliakbar Noroozi [3], Ali Akbar Abkar [4], Saham Mirzaei [5], Zahra Mohammadi-Doqozloo [6], Ruhollah Taghizadeh-Mehrjardi [2,7], Pegah Khosravani [2,8], Kamal Nabiollahi [2,9] and Thomas Scholten [2,10]

1 Department of Water Engineering, University of Tehran, Tehran 3391653755, Iran; mmashal@ut.ac.ir
2 Department of Geosciences, Soil Science and Geomorphology, University of Tübingen, 72076 Tübingen, Germany; ruhollah.taghizadeh-mehrjardi@mnf.uni-tuebingen.de (R.T.-M.); pegah.khosravani@uni-tuebingen.de (P.K.); kamal.nabiollahi@mnf.uni-tuebingen.de (K.N.); thomas.scholten@uni-tuebingen.de (T.S.)
3 Soil Conservation and Watershed Management Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Tehran 1985713133, Iran; noroozi@itc.nl
4 Geographic Information System and Remote Sensing of Agriwtach BV, 7542 SC Enschede, The Netherlands; ali.abkar@agriwatch.nl
5 Institute of Methodologies for Environmental Analysis, Italian National Research Council, 85050 Potenza, Italy; sahammirzaei@cnr.it
6 Department of Agricultural Machinery Engineering, University of Tehran, Tehran 3158777871, Iran; zahra.mohammadi76@ut.ac.ir
7 Faculty of Agriculture and Natural Resources, Ardakan University, Ardakan 9549189518, Iran
8 Department of Soil Science, University of Shiraz, Shiraz 7194684471, Iran
9 Department of Soil Science and Engineering, University of Kurdistan, Sanandaj 1517566177, Iran
10 Cluster of Excellence Machine Learning: New Perspectives for Science, University of Tübingen, 72076 Tübingen, Germany
* Correspondence: hadi.shokati@ut.ac.ir or hadi.shokati@uni-tuebingen.de

Abstract: Accurate spatiotemporal monitoring and modeling of soil moisture (SM) is of paramount importance for various applications ranging from food production to climate change adaptation. This study deals with modeling SM with the random forest (RF) algorithm using datasets comprising multispectral data from Sentinel-2, Landsat-8/9, and hyperspectral data from the CoSpectroCam sensor (CSC, licensed to AgriWatch BV, Enschede, The Netherlands) mounted on an unmanned aerial vehicle (UAV) in Iran. The model included nine bands from Landsat-8/9, 11 bands from Sentinel-2, and 1252 bands from the CSC (covering the wavelength range between 420 and 850 nm). The relative feature importance and band sensitivity to SM variations were analyzed. In addition, four indices, including the perpendicular index (PI), ratio index (RI), difference index (DI), and normalized difference index (NDI) were calculated from the different bands of the datasets, and their sensitivity to SM was evaluated. The results showed that the PI exhibited the highest sensitivity to SM changes in all datasets among the four indices considered. Comparisons of the performance of the datasets in SM estimation emphasized the superior performance of the UAV hyperspectral data (R^2 = 0.87), while the Sentinel-2 and Landsat-8/9 data showed lower accuracy (R^2 = 0.49 and 0.66, respectively). The robust performance of the CSC data is likely due to its superior spatial and spectral resolution as well as the application of preprocessing techniques such as noise reduction and smoothing filters. The lower accuracy of the multispectral data from Sentinel-2 and Landsat-8/9 can also be attributed to their relatively coarse spatial resolution compared to the CSC, which leads to pixel non-uniformities and impurities. Therefore, employing the CSC on a UAV proves to be a valuable technology, providing an effective link between satellite observations and ground measurements.

Keywords: soil moisture; random forest; CoSpectroCam; UAV; Sentinel-2; Landsat-8/9

Citation: Shokati, H.; Mashal, M.; Noroozi, A.; Abkar, A.A.; Mirzaei, S.; Mohammadi-Doqozloo, Z.; Taghizadeh-Mehrjardi, R.; Khosravani, P.; Nabiollahi, K.; Scholten, T. Random Forest-Based Soil Moisture Estimation Using Sentinel-2, Landsat-8/9, and UAV-Based Hyperspectral Data. *Remote Sens.* **2024**, *16*, 1962. https://doi.org/10.3390/rs16111962

Academic Editors: Xiaoling Wu, Chong Luo, Liujun Zhu and Xiaoji Shen

Received: 24 April 2024
Revised: 27 May 2024
Accepted: 28 May 2024
Published: 29 May 2024

1. Introduction

Soil moisture monitoring is crucial in various fields, including soil science [1], ecology [2], hydrology [3,4], and water resource management [5]. In agriculture, accurate monitoring of SM at different stages of crop growth is particularly important for water stress detection [6] and crop yield estimation [7]. In addition, SM is considered an important component of the hydrological cycle as it effectively controls the ratio of rainfall conversion to infiltration and runoff [8,9]. Consequently, spatial and temporal monitoring of SM changes can be an important prerequisite for climate change mitigation and adaptation [10], which is essential not only for the sustainability of agricultural systems, but also for the further development of precision agriculture and for ensuring food security [11–15].

Direct methods for estimating SM are limited to point-based approaches that provide localized SM measurements at specific depths and points. However, such point-based measurements do not represent distribution patterns of SM, as SM is highly variable both spatially and temporally [16,17]. Point-based methods also have disadvantages, such as the high cost of equipment, the high labor input for field-level measurements, and the time-intensive nature of soil sampling. Therefore, an alternative solution for SM estimation is the use of remote sensing technology [18–20]. This approach offers a cost-effective way to achieve both temporal and spatial coverage, e.g., for precision agriculture, and to avoid the above-mentioned problems of point-based soil sampling.

In general, an increase in SM leads to a decrease in soil reflectance [21]. This complex relationship can lead to significant errors in modeling and remote sensing-based estimates of SM. Numerous studies have addressed the modeling of SM using multispectral satellite data. A study from South Africa highlighted the high potential of Sentinel-1 and Sentinel-2 data for SM estimation [22]. The findings indicated that RF outperformed support vector regression (SVR) and simple linear regression algorithms in estimating SM. In another study, SM estimation from Sentinel-2 data was performed using a convolutional neural network (CNN) in Austria and Australia [23]. The results showed that the short-wave infrared-1 band (SWIR-1) had the highest sensitivity, while the blue band had the lowest sensitivity for SM. The significant potential of Landsat-8 data for SM estimation has been demonstrated in various studies [24–26]. Despite the considerable potential of multispectral satellite imagery for SM estimation, its limited spatial resolution hampers its applicability at smaller scales. The increasing use of UAVs has become prominent due to their enhanced temporal flexibility and higher spatial resolution compared with satellite data. Recent advances include the use of small digital sensors on UAVs that provide exceptional spectral and spatial resolution [27]. There is a growing body of research on SM prediction through the integration of UAVs and various sensors. Notable studies have used multispectral sensors on UAVs [28–30], while a few have integrated hyperspectral sensors [31,32]. Modeling SM from these datasets requires the extraction of relationships between SM content and surface reflectance and emission. To achieve this, machine learning algorithms prove helpful in establishing the intricate connections within the data.

The versatility of machine learning algorithms has been demonstrated on a variety of remote sensing data, including optical, radar, and UAV data [33]. Comparative analyses with other parametric approaches, including iterative and Bayesian methods, consistently show the superior performance of machine learning algorithms in numerous studies [34]. In contrast to linear mathematical models, machine learning is characterized by the detection of non-linear relationships between predictor and target variables, as has been observed in various studies [31]. A distinct advantage of machine learning methods is that they rely less on preprocessing data for modeling, especially due to their ability to process different datasets without extensive transformations [35,36]. The machine learning landscape includes several established methods such as the artificial neural network (ANN), RF, SVR, extreme learning machine (ELM), partial least squares regression (PLSR), and Gaussian process regression (GPR). Numerous studies have applied machine learning algorithms to estimate SM using various remote sensing data

and demonstrated their effectiveness [24,31,37–39]. In a study from China, RF, KNN, PLSR, and backpropagation neural network (BPNN) models were used to estimate SM in a maize field using multispectral, RGB, and thermal data collected by a UAV-mounted sensor [40]. The results showed that multispectral data consistently provided more accurate results than RGB and thermal data, regardless of the model used. Furthermore, the results emphasized the superior performance of the RF model in SM estimation compared to KNN, PLSR, and BPNN. In another study, UAV-based RGB data were used to estimate SM in a bare soil field in Iran using RF, M5 tree (M5P), multilayer perceptron (MLP), and sequential minimal optimization regression (SMOreg) algorithms [41]. The results showed that the RF model was superior to the other models in accurately estimating SM. In addition, further research addressed SM modeling using the general linear model (GLM) at depths of 2, 8, and 12 inches in South Georgia using UAV hyperspectral data [42]. The results showed a statistically significant correlation between hyperspectral data and ground data at a depth of 2 inches. However, the models tested for depths of 8 and 12 inches were not able to estimate SM with the same accuracy. Similarly, another study focused on developing a quantitative model to estimate SM using ELM and RF models [31]. A hyperspectral sensor mounted on the Matrice 600 UAV captured electromagnetic waves in the wavelength range from 400 to 1000 nm. The comparison of the two machine learning algorithms showed the superiority of the RF model over the ELM. Estimating SM from UAV thermal and multispectral sensors, Bertalan [43] used the RF, GLM, elastic net regression (ENR), and robust linear model (RLM) models. The results showed that the RF model, utilizing both thermal and multispectral data, outperformed the other models. Comparing the accuracy of machine learning models in estimating SM, de Oliveira [44] showed that the RF model outperformed the average neural network and weighted k-nearest neighbor (KNN) models in spatial and temporal modeling of SM. While several studies have focused on SM modeling using satellite or UAV data, few comparative analyses directly evaluate the accuracy and performance of UAV-based hyperspectral sensors versus satellite-based multispectral data in SM modeling. This research gap prevents a comprehensive understanding of the strengths and limitations of each data source in accurately estimating SM under different environmental conditions and at different scales. Such comparative analyses are critical for selecting the most appropriate data source for specific SM modeling applications, considering factors such as spatial resolution, spectral range, and temporal frequency.

The main objectives of this study are (1) to compare the performance of multispectral data from Sentinel-2 and Landsat-8/9 with hyperspectral data from the UAV-mounted CSC in SM estimation using the RF algorithm, (2) to investigate the relationships between all possible two-band combinations of each dataset with SM, and (3) to determine the most sensitive bands of each dataset in SM estimation.

2. Materials and Methods

2.1. The Study Area

The study was conducted on a research farm of the University of Tehran, located in Pakdasht of Iran (latitude 35°28′25″N, longitude 51°33′35″E), with an area of 100 hectares. This area is slightly sloped and has an average elevation of 985 m above sea level. The region receives an average of 217 mm annual precipitation. The predominant land uses in the study area include bare soil, alfalfa, wheat, and corn. This land use composition can vary from year to year, depending on the availability of water resources. Figure 1 shows an overview of the study area and the location of the soil samples collected for the individual land uses.

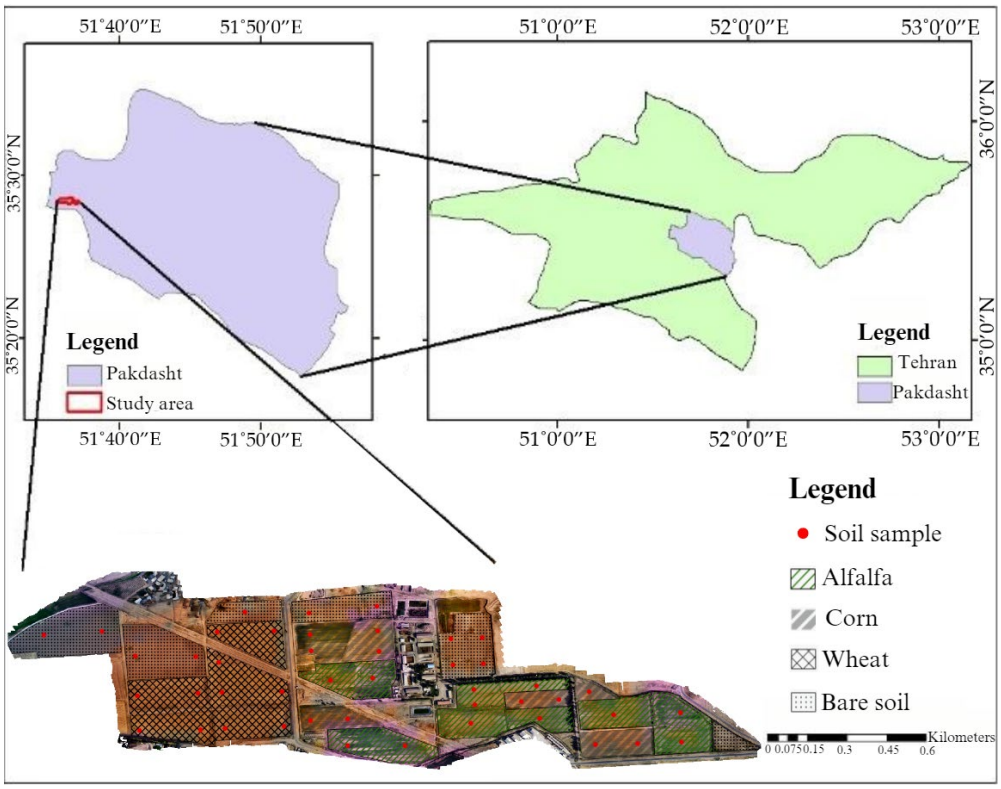

Figure 1. Overview of the study area and the location of the soil samples collected from the individual land uses.

2.2. Required Data

2.2.1. In Situ Soil Moisture Data

In each field campaign, 45 points (12 samples in alfalfa, 11 samples in wheat, 11 samples in corn, 11 samples in bare soil) were selected for in situ soil sampling. At each sampling point, five topsoil cores (0–10 cm depth) were combined to create composite samples (45×5 soil samples were taken from 45 points in each field campaign). Soil moisture content was measured using the gravimetric method [4].

2.2.2. Soil Properties Data

The spectral reflectance of the soil surface is influenced by various soil properties [45–47]. Therefore, 25 points were selected to investigate the spatial variations of electrical conductivity (EC), pH, soil texture, and soil color in the region. The hydrometric method proposed by Bouyoucos [48] was used to assess soil texture at the selected points. In addition, an EC meter was used to measure the EC value, a pH meter to determine the pH of the soil, and the Munsell color system to characterize the soil color. Notably, all samples showed remarkable homogeneity in terms of texture, with silty loam identified as the predominant soil texture in the field. In addition, the samples were uniform in color, pH, and salinity. The premise of this study is therefore that soil properties in the region are predominantly homogeneous, with the spectral response of the soil being primarily influenced by variations in SM.

2.2.3. Satellite Data

The Sentinel-2 Level-2A (L2A) surface reflectance images were obtained from the European Space Agency (ESA) website (https://scihub.copernicus.eu, accessed on 28 March 2017). The Level-2A product provides atmospherically corrected surface reflectance images derived from the corresponding Level-1C products. The Aerosol, Blue, Green, Red, Red Edge 1, Red Edge 2, Red Edge 3, NIR, Red Edge 4, SWIR-1, and SWIR-2 bands were selected for SM modeling as they are particularly sensitive to SM [23].

The specifications of the Landsat-8 and Landsat-9 satellites are similar; the only relevant difference is that the radiometric resolution of 12 bits for Landsat-8 has been upgraded to 14 bits for Landsat-9, and the overall signal-to-noise ratio has been slightly improved. Corrected Collection-2 Level-2 images of these two satellites were downloaded from the United States Geological Survey (USGS) website. The Aerosol, Blue, Green, Red, NIR, SWIR-1, SWIR-2, TIR-1, and TIR-2 bands were used for modeling SM.

2.2.4. UAV Hyperspectral Data Acquisition

In this study, data from a USB2000PLUS hyperspectral sensor (CSC) was used in addition to the satellite data. The CSC has spectroscopic capabilities in the wavelength range from 339.6 to 1028.8 nm, covering the visible and near-infrared (NIR) regions, and offers a spectral resolution of 0.35 nm (with a total of 2048 bands). The DJI Matrice 100 UAV served as the carrier for the CSC. Data acquisition with the CSC was performed on seven different days, synchronized with satellite passes over the study area. To mitigate atmospheric effects, data acquisition was done on sunny days. Figure 2 shows the Matrice 100 UAV and the CSC mounted on it. The availability of datasets in each data collection is shown in Figure 3.

Figure 2. CoSpectroCam installed on Matrice 100 UAV.

CoSpectrocam software served as an intermediary to establish a connection between the UAV and the CSC. Before the flight, the CSC was calibrated using a Spectralon White Diffuse Reflectance Standard. Equation (1) was used to measure the reflectance of the objects [49]:

$$\rho = \frac{L_{sam}}{L_{ref}} \times \rho_{ref} \quad (1)$$

where ρ is the reflectance of the sample, L_{sam} and L_{ref} are the measured radiance of the sample and the white reference, respectively, and ρ_{ref} is the reflectance of the white standard reference.

DJI Go software (version 3.1.61) was used to establish a connection between the UAV and the remote control. Drone Harmony software (version 2.3.0, Drone Harmony Company, Zurich, Switzerland) was used for the automatic navigation of the UAV. The flight parameters were set using Drone Harmony software, with the flight altitude set to 50 m. In addition, the overlap between the images was set to 70% and the movement speed of the UAV was set to 3 m/s. This software proved effective in mitigating the effects of the angle of solar radiation on the spectra by allowing the UAV to maintain a constant heading

during the round trip. To ensure optimal conditions for data collection, all data were collected on sunny days between 9 a.m. and 12 p.m. local time to minimize atmospheric effects. Finally, the CSC hyperspectral data were smoothed using a Savitzky–Golay filter with a polynomial degree of 2 and a window size of 25 (4.2 nm on each side).

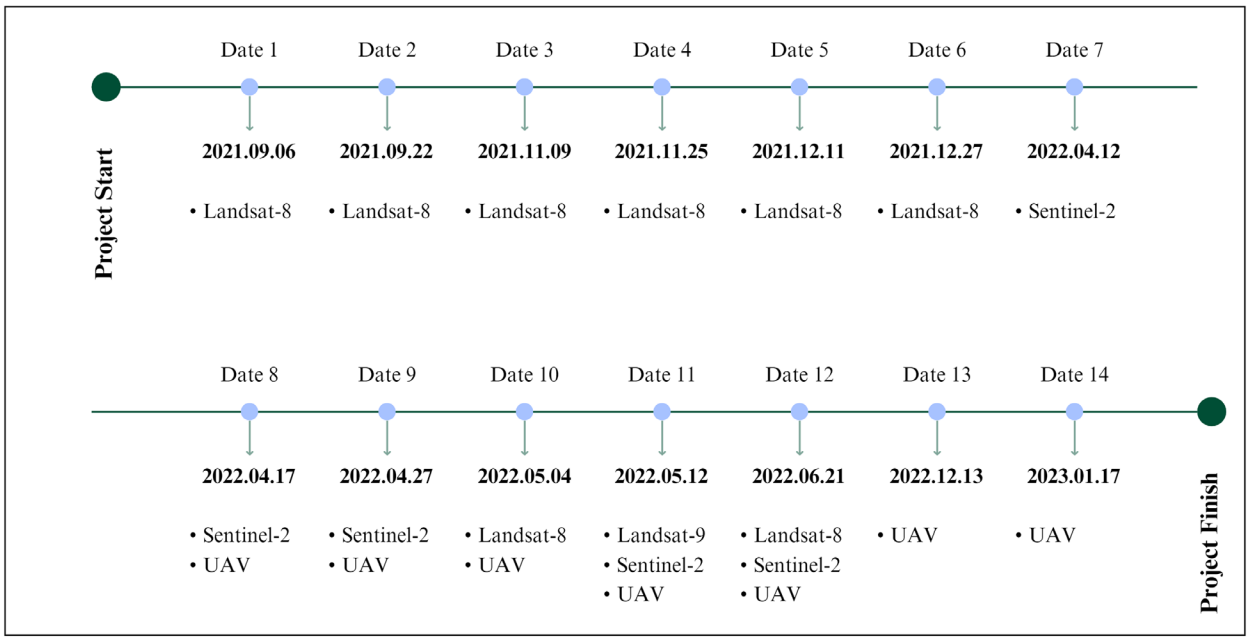

Figure 3. Availability of datasets in each data collection.

2.3. Methodology

2.3.1. Spectral Indices Extraction

To estimate the SM from optical data, vegetation indices are commonly used to reflect the condition of the plant cover at different SM contents. Since these indices can reflect the physiological state of plants, they serve as indicators of vegetation cover [50,51] and respond to changes in SM [52]. Therefore, several spectral indices, including the Normalized Difference Vegetation Index (NDVI) [53], the Normalized Difference Water Index (NDWI) [54], the Normalized Difference Moisture Index (NDMI) [55] and the Soil Adjusted Vegetation Index (SAVI) [56], were extracted from Landsat-8/9 and Sentinel-2 data using ArcMap (version 10.8.1). These indices were then analyzed for their correlation with in situ SM data.

2.3.2. Narrow Band Spectral Indices

Four indices, including the DI, RI, NDI, and PI were created using satellite bands and CSC bands. For this purpose, all possible two-band combinations were investigated in MATLAB R2023a using Equations (2)–(5) [31], and the correlations between all calculated indices and SM was evaluated.

$$DI_{(R_i,R_j)=R_i-R_j} \tag{2}$$

$$RI_{(R_i,R_j)=\frac{R_i}{R_j}} \tag{3}$$

$$NDI_{(R_i,R_j)=\frac{R_i-R_j}{R_i+R_j}} \tag{4}$$

$$PI_{(R_i,R_j)} = \frac{R_i - 0.4401R_j - 0.3308}{\sqrt{1+0.4401^2}} \tag{5}$$

where, R_i and R_j are the spectral reflectance of i and j bands, respectively.

2.3.3. Soil Moisture Modeling

In this study, the RF algorithm [57] was used for modeling SM. The RF algorithm is recognized as an ensemble learning method that is commonly used for regression and classification tasks. In the training phase, several decision trees are constructed. Each decision tree is trained on a random sample of the data with replacement. In addition, only a random subset of features is considered for splitting at each node of the decision tree, introducing randomness and diversity into the trees. During prediction, each tree in the RF independently generates its own prediction. For classification tasks, the class predicted by the majority of trees is chosen as the final prediction, while for regression tasks, the average prediction of all trees is taken.

The RF algorithm has been widely used in various soil studies related to soil properties such as soil texture and soil type, and its effectiveness in modeling SM with different datasets has been demonstrated in several studies [22,24,31,40,41,44]. The main advantages of RF in soil studies are its compatibility and high accuracy, especially for datasets with numerous predictors (features) and limited samples. The algorithm is characterized by its robustness to irrelevant features, and it eliminates the need for precise parameter settings [58]. Consequently, it is well suited for datasets with a large amount of data [59]. The ARTMO package [60] in MATLAB R2023a was used for SM modeling. After preprocessing the data, outliers and irrelevant data were excluded from the modeling process. The model included nine bands from Landsat-8/9, 11 bands from Sentinel-2, and 1252 bands from the CSC (covering the wavelength range between 420 and 850 nm).

In this study, the modeling was carried out independently for each dataset. Depending on the availability of datasets on the days of data collection, the total amount of in situ data for each dataset varied. Thus, a total of 213 data were collected for Sentinel-2, 324 data for Landsat-8/9, and 236 data for the CSC. Two-thirds of the in situ data were used to train the RF model for each dataset, and the remaining data were used to validate the trained models. Because of the small number of samples, a k-fold (k = 10) approach was used for cross-validation of the trained models. In this way, all the samples were used for training and validation. In addition, the results of the datasets on compatible days were compared with each other. Finally, the sensitive bands of each dataset to SM were identified, and uncertainty maps were created. Figure 4 shows the flowchart of the work.

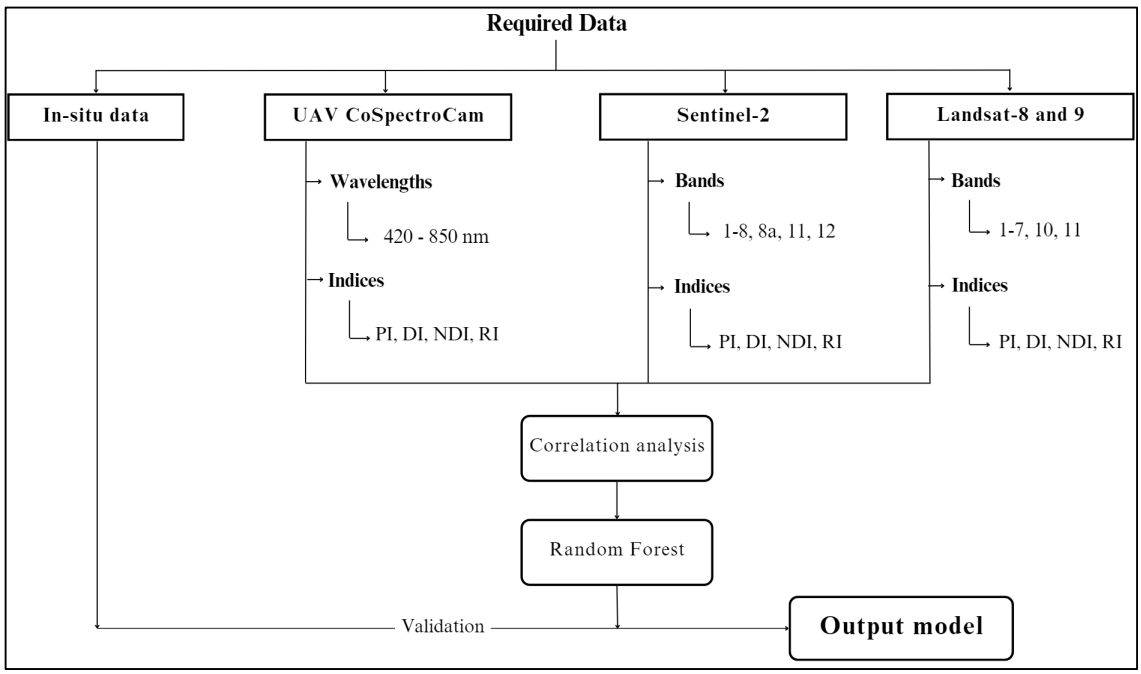

Figure 4. The flowchart of the work.

2.3.4. Validation

In this study, four statistical indices, including the coefficient of determination (R^2), the adjusted coefficient of determination (R^2_{adj}), the root mean square error (RMSE) [61], and the mean absolute error (MAE) [62], were used based on the following equations.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(SMC_{ob,i} - SMC_{p,i}\right)^2}{\sum_{i=1}^{n}\left(SMC_{ob,i} - \overline{SMC}\right)^2} \quad (6)$$

$$R^2{}_{adj} = 1 - \frac{n-1}{n-m-1}\left(1 - R^2\right) \quad (7)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(SMC_{p,i} - SMC_{ob,i}\right)^2}{n}} \quad (8)$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}\left|SMC_{p,i} - SMC_{ob,i}\right|}{n} \quad (9)$$

where SMC_p and SMC_{ob} are the predicted and observed SM, respectively, $\overline{SMC}$ is the average of the observed SM, n is the number of soil samples, and m is the number of independent variables in the model that are used to predict the SM.

3. Results and Discussion

3.1. Effect of Soil Moisture on Multispectral Data

Table 1 shows the correlations between different bands and indices of Sentinel-2 and Landsat-8/9 satellites and SM. Of the Sentinel-2 bands, the SWIR-1 band showed the highest correlation with in situ SM data, with a correlation coefficient of 0.66 [23,63–65]. In addition, the Blue, Green, SWIR-2, and Red bands had correlations of −0.65, −0.65, −0.64, and −0.64, respectively, showing only minor differences in their rankings. Conversely, the Red Edge bands showed minimal correlation with SM, despite their indirect link to vegetation water content, which is sensitive to changes in SM [66,67]. While the Red Edge 3 band is pivotal for estimating crop water content, its correlation with SM was remarkably low in this study. This discrepancy can be attributed to the prevalence of bare soils and irregular irrigation schedules in the study area, and the finding differs from the results of Hegazi [23], who focused on a vegetated region. In addition, the time interval between irrigation and the response of plants to SM changes also influences these correlations [68]. In the case of the Sentinel-2 data, the SWIR-1 band showed a stronger correlation with the ground SM data than the SWIR-2 band, which is consistent with the results of Hegazi [23] and Hashim [69]. However, this contrasts with the results reported by Sanchez [70]. In contrast to the results of Zheng [71] and Hegazi [23], the Blue band in this study showed a remarkable correlation with the in situ SM data.

Table 1. Correlation of the bands and indices of the Sentinel-2 and Landsat-8/9 with soil moisture.

Features	Landsat-8/9			Sentinel-2		
	Equation	**R ***	**Band**	**Equation**	**R ***	**Band**
Bands	y = −0.0144x + 33.4	−0.55	Aerosol	y = −0.007x + 81.01	−0.68	Aerosol
	y = −0.0113x + 31.7	−0.65	Blue	y = −0.0059x + 74.13	−0.68	Blue
	y = −0.0103x + 35.4	−0.65	Green	y = −0.0043x + 65.01	−0.68	Green
	y = −0.007x + 29.73	−0.64	Red	y = −0.003x + 52.69	−0.67	Red
	y = −0.0084x + 37.1	−0.61	Red Edge 1	y = −0.0009x + 23.98	−0.23	NIR
	y = −0.0036x + 23.9	−0.12	Red Edge 2	y = −0.0027x + 57.97	−0.67	SWIR-1

Table 1. *Cont.*

Features	Landsat-8/9			Sentinel-2		
	Equation	R *	Band	Equation	R *	Band
Bands	y = 0.0018x + 0.71	0.08	Red Edge 3	y = −0.0026x + 49.72	−0.69	SWIR-2
	y = 0.0018x + 0.73	0.07	NIR	y = −0.0006x + 34.41	−0.30	TIR-1
	y = 0.0032x − 5.81	0.14	Red Edge 4			
	y = −0.0101x + 51.3	−0.66	SWIR-1			
	y = −0.0071x + 34.8	−0.64	SWIR-2			
Indices	y = 54.345x + 6.56	0.57	NDMI	y = 65.453x + 7.94	0.43	NDMI
	y = 33.049x + 1.7	0.59	NDVI	y = 45.414x + 3.11	0.4	NDVI
	y = −42.71x − 2.94	−0.58	NDWI	y = −47.385x + 0.56	−0.32	NDWI
	y = 22.035x + 1.7	0.59	SAVI	y = 30.277x + 3.11	0.4	SAVI

* Pearson correlation.

The results of the Landsat-8/9 data showed that of the Landsat bands, the SWIR-2 band had the strongest correlation with the in situ SM data (r = −0.69). The effectiveness of the SWIR band in accurately estimating SM has been emphasized in previous studies [71–73]. This is closely followed by the Blue band, with a correlation coefficient of −0.68, which contrasts with the results of Zheng [71] and Hegazi [23]. The SWIR-2 band showed a higher correlation with in situ SM data than the SWIR-1 band, which agrees with Sanchez [70] but differs from Hegazi [23] and Hashim [69]. The lowest correlation with the in situ SM data was found for TIR-2, with a correlation coefficient of 0.07. In general, the Landsat-8/9 thermal bands showed a modest linear correlation with SM, possibly due to their lower spatial resolution (originally 100 m, later converted to 30 m by processing). The considerable spatial variability of land use in the region, combined with the large pixel size, resulted in the presence of different land uses within a pixel, contributing to pixel non-uniformity and consequently reduced band performance. In addition, the different land uses in the region complicated the relationship between thermal bands and SM [74,75].

The correlation of both satellites with in situ SM data in the NIR bands was remarkably low due to water vapor absorption effects at a wavelength of 825 nm [76]. Moreover, the SWIR bands of both satellites showed similar performance [76] and are very suitable for estimating surface soil moisture in drylands [77]. For both satellites, the SWIR bands showed a higher correlation with the in situ SM data compared to the NIR bands [71,78,79].

As shown in Figure 5a, the NDVI values of the Landsat-8/9 were lower than those of the Sentinel-2 images for collected soil samples. Because of the presence of less green vegetation, the single band of Landsat-8/9 showed higher capability for SM retrieval, compared to Sentinel-2.

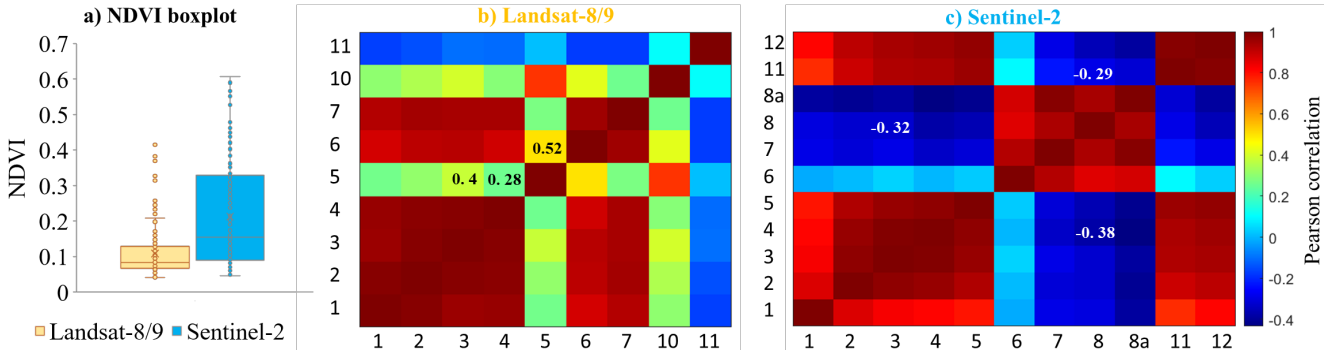

Figure 5. (**a**) NDVI boxplot and between-bands Pearson correlation of the (**b**) Landsat-8/9 and (**c**) Sentinel-2 images for collected soil samples.

As shown in Table 1, the Sentinel-2 indices showed a higher correlation with SM compared to Landst-8/9. Apart from the NDWI, all indices showed a direct relationship with SM, meaning that as SM increased, the values of these indices also increased. The NDVI and SAVI were calculated using bands 4 and 5 of Landsat-8/9 and pair-bands 4 and 8 of Sentinel-2 images. The between-bands correlations for Landsat-8/9 and Sentinel-2 for the NDVI and SAVI were 0.28 and −0.38, respectively. The NDMI was calculated between pair-bands 5 and 6 of Landsat-8/9 and pair-bands 8 and 11 of Sentinel-2 images. The between-bands correlation for Landsat-8/9 and Sentinel-2 for the NDMI were 0.52 and −0.29, respectively. The NDWI was calculated between pair-bands 3 and 5 of Landsat-8/9 and pair-bands 3 and 8 of Sentinel-2 images. The between-bands correlation for Landsat-8/9 and Sentinel-2 for the NDWI were 0.4 and −0.32, respectively. The presence of more green vegetation in Sentinel-2 images can be proven by a positive Pearson correlation between pair-bands for Landsat-8/9-derived indices (Figure 5b) and a negative Pearson correlation between pair-bands for Sentinel-2-derived indices (Figure 5c). The between-bands correlation of Landsat-8/9 bands, achieved by increasing the data redundancy, reduced the performance of indices calculated from these bands for SM retrieval, whereas combining the pair-bands of Sentinel-2 led to improved accuracy of SM estimation.

3.2. *Effect of Soil Moisture on Hyperspectral Data*

In this study, bands with wavelengths below 420 nm and above 850 nm, which are characterized by noise, were excluded from the processing. The correlation between the different CSC bands and the in situ SM data is shown in Figure 6. The results showed that the effect of SM on different wavelengths is not the same. The lowest correlation was observed for the first bands in the 420 to 430 nm range. After this range, the correlation values increased sharply and peaked at the wavelength of 480 nm. Within the 480 to 770 nm range, the correlation of the different bands with the in situ SM data remained relatively constant, with the highest correlation coefficient recorded at 527 nm (r = −0.85). The correlation patterns observed in this study are consistent with the results of Jiang [80] and Mirzaei [21]. Beyond the wavelength of 770 nm, the correlation values gradually decreased.

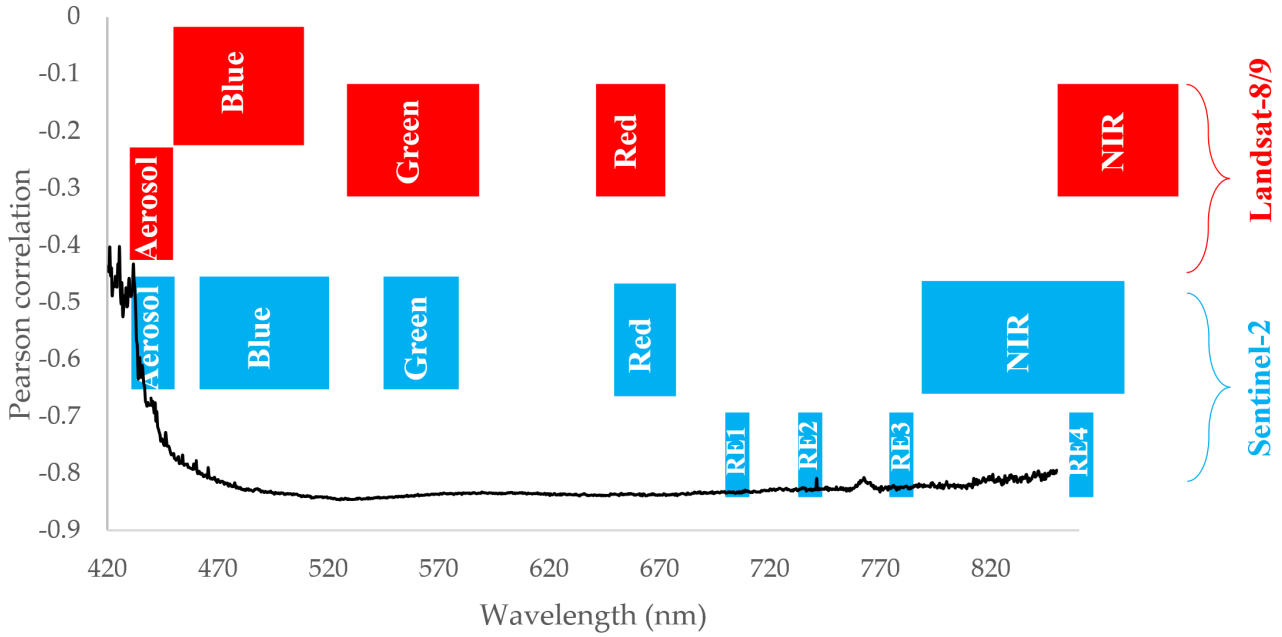

Figure 6. Pearson correlation between CoSpectroCam bands and in situ soil moisture data. To make comparison easier with the achievements reported in Table 1, the positions of the Landsat-8/9 and Sentinel-2 overlap bands with those of the CoSpectroCam are shown in red and blue, respectively.

3.3. Soil Moisture Estimation Using Indices

Figure 7 shows the 2D maps illustrating the coefficients of determination of the indices derived from the different bands of Sentinel-2, Landsat-8/9, and the CSC, for the estimation of SM. The results showed that the PI had an increased sensitivity to SM variations in all datasets [31,41]. Furthermore, the results showed that all CSC-derived indices performed better compared to the indices derived from Landsat-8/9 and Sentinel-2 data for SM estimation. In particular, the accuracy of the Landsat-derived indices exceeded that of the Sentinel-2-derived indices.

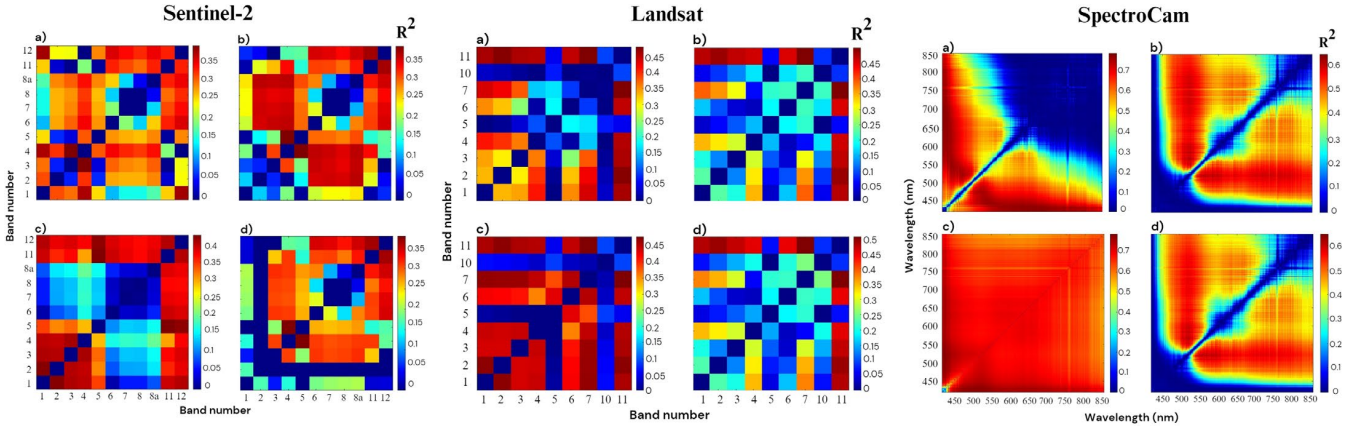

Figure 7. 2D plot of R^2 of (**a**) DI, (**b**) NDI, (**c**) PI, and (**d**) RI for soil moisture prediction using Sentinel-2 (**left**), Landsat-8/9 (**middle**), and CoSpectroCam (**right**) datasets.

The optimal bands for each index, yielding the highest accuracy in SM estimation, are given in Table 2. In the case of the Sentinel-2 data, the PI created with Red Edge 1 and SWIR-1 had the highest accuracy in estimating SM, with an R^2 value of 0.42. This result confirms the claims of Liu [73] that the combination of Red Edge and SWIR bands significantly improves the accuracy of SM estimation. In the context of the DI, the optimal spectral bands for SM estimation were identified as Green and Red. In addition, the most effective band combination for creating the NDI and RI was identified as a combination of the Red and Red Edge 1 bands.

Table 2. Optimal pair-bands for indices.

Index	Optimal Bands/Wavelengths			R^2		
	Sentinel-2	Landsat-8/9	CoSpectroCam (nm)	Sentinel-2	Landsat-8/9	CoSpectroCam
DI	Green, Red	SWIR-2 and TIR-2	427.5, 678.0	0.36	0.48	0.78
NDI	Red, Red Edge 1	SWIR-2 and TIR-2	525.4, 562.1	0.38	0.50	0.65
PI	Red Edge 1, SWIR-1	SWIR-2 and TIR-2	431.6, 523.6	0.42	0.48	0.79
RI	Red, Red Edge 1	SWIR-2 and TIR-2	525.4, 562.1	0.38	0.51	0.65

For the Landsat-8/9 data, the PI performed consistently better for SM estimation. In particular, the RI, which was formulated from SWIR-2 and TIR-2 bands, showed the best performance for SM estimation, with an R^2 of 0.51. The defined bands were also simultaneously identified as optimal for the DI, NDI, and PI and showed robust SM estimation abilities, with R^2 values of 0.48, 0.5, and 0.48, respectively.

In the CSC data, the PI showed higher sensitivity to SM changes compared to the other three indices. In particular, the PI achieved a remarkable R^2 of 0.79 in SM estimation when wavelengths of 431.6 and 523.6 nm were used. The optimal wavelengths for creating the DI, NDI, and RI were found to be 427.5 and 678, 431.6 and 523.6, and 525.4 and 562.1 nm, respectively. These wavelengths showed remarkable accuracy in estimating SM, yielding R^2 values of 0.78, 0.65, and 0.65, respectively.

3.4. Soil Moisture Retrieval Using Random Forest Algorithm

The data from all datasets were used as inputs to the RF model, and the results were validated with in situ SM data by 10-fold cross-validation (Table 3). The results underlined the superior performance of the CSC hyperspectral data in SM estimation compared to Sentinel-2 and Landsat-8/9 multispectral data [42]. In particular, the Landsat-8/9 data outperformed the Sentinel-2 data in SM estimation, partly due to the lack of a thermal band in Sentinel-2. The thermal bands in Landsat-8/9 had a high sensitivity to SM changes, contributing to its better performance. The R^2 for the model created with the Sentinel-2 data was 0.49, which, in contrast to the results of Liu [73], indicates a relatively modest potential for SM estimation; this discrepancy can be attributed to the limited size of the study area.

Table 3. The accuracy of the random forest algorithm for soil moisture estimation using Sentinel-2, Landsat-8/9, and CoSpectroCam data.

Dataset	R	R^2	R^2_{adj}	RMSE (%)	MAE
Sentinel-2	0.70	0.49	0.48	5.95	3.43
Landsat-8/9	0.81	0.66	0.65	4.61	2.42
CoSpectroCam	0.93	0.87	0.86	2.60	1.93

In contrast, the model built with Landsat-8/9 data yielded an R^2 of 0.66, confirming the significant potential of Landsat-8/9 data for SM estimation [26,81,82]. Impressively, the model created with CSC data achieved an R^2 of 0.87, highlighting the great potential of CSC hyperspectral data for SM estimation (See Supplementary Information, Figures S1 and S2 for SM maps predicted using Sentinel-2 and Landsat-8/9 data).

Figure 8 shows the predicted and observed SM levels for all datasets, with each date shown in a different color. The results showed a consistent trend wherein the models produced tended to overestimate the SM content in regions with less than 8% SM, while there was a tendency to underestimate the SM content in regions with more than 8% SM [41]. The degree of over- and underestimation of the SM was found to be more pronounced when using satellite data compared to when using the CSC. This can be attributed to the lower spatial resolution of satellites compared to the CSC, which leads to the non-uniformity and impurity of pixels [83]. A pixel may include both bare soil and vegetation while providing only a single value. Other factors contributing to the higher accuracy of CSC data are the higher spatial and spectral resolution and the use of preprocessing techniques such as noise removal and smoothing filters for CSC data.

SM maps derived from Sentinel-2 and Landsat-8/9 data on common overpass days (see Supplementary Information, Figure S3) revealed a consistent pattern in SM estimation for both satellites. Both satellites showed higher SM values in regions with vegetation compared to regions with bare soil, a phenomenon that can be attributed to the increased preservation of SM in vegetated areas.

Following the modeling of SM using satellite data, an uncertainty map for the SM estimation was generated based on the standard deviation of the observed and predicted values for each date (see Supplementary Information, Figures S4 and S5). The results indicated that the accuracy of both satellites decreased when estimating SM in vegetated areas, leading to an increased level of uncertainty in these regions. A comparison of the uncertainty map produced with Sentinel-2 and Landsat-8/9 data on days when both satellites had the same transit time over the study area (12 May 2022, and 21 June 2022) showed a higher degree of uncertainty in the maps derived from Sentinel-2 data. Nevertheless, the uncertainty patterns in SM estimation were similar for both satellites, with lower performance observed in areas with vegetation.

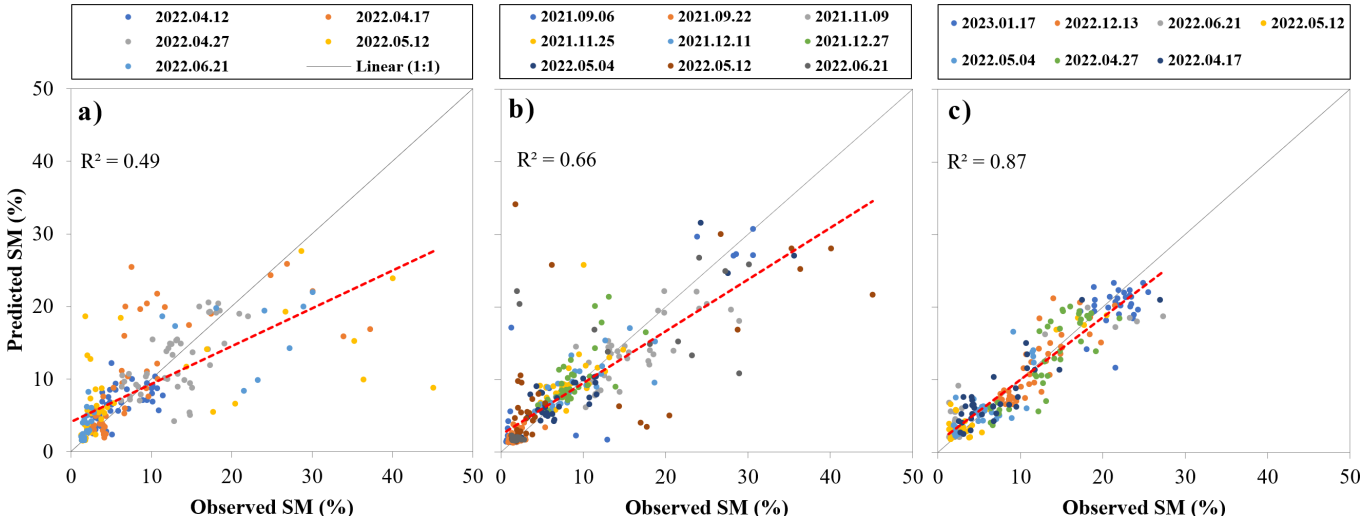

Figure 8. Observed against predicted soil moisture using the random forest (RF) algorithm for (**a**) Sentinel-2, (**b**) Landsat-8/9, and (**c**) CoSpectroCam data.

3.5. Band Sensitivity Analysis

The sensitive wavelengths of the CSC and the sensitive bands of the satellites in each iteration of the modeling process are shown in Figure 9. The findings showed that the Red Edge1 band of Sentinel-2 scored highest in 5 of the 10 replicates. In addition, the Red Edge 2 and Red Edge 4 bands secured the top position in three and two replicates, respectively. These results emphasize the importance and sensitivity of the Red Edge Sentinel-2 bands to variations in SM. However, as indicated in Table 1, the linear correlation between these bands and SM was comparatively lower. This discrepancy may be attributed to the inherent limitation of linear correlation in accurately representing the complicated relationship between these bands and SM, which was extracted non-linearly by the RF algorithm. The sensitivity of the Red Edge bands to SM changes is consistent with the findings of the study by Paloscia [73].

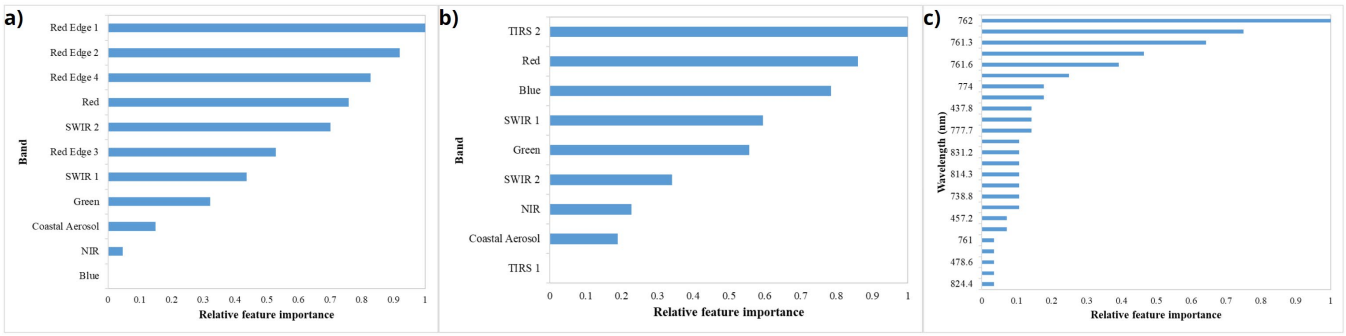

Figure 9. The relative feature importance of (**a**) Sentinel-2 bands, (**b**) Landsat-8/9 bands, and (**c**) CoSpectroCam wavelengths in soil moisture estimation using the random forest (RF) algorithm.

In the Landsat-8/9 data, the TIR-2 band secured the top position in nine of the ten replicates, highlighting its remarkable sensitivity to SM changes. However, as indicated in Table 1, the linear correlation between this band and SM was comparatively lower. The noticeable difference in sensitivity is probably due to the inadequacy of linear correlation in accurately describing the relationship between the dependent and independent variables. Since this relationship is not linear, it could be better elucidated by the RF algorithm. In the only remaining replicate, the Red band was identified as the most sensitive of the bands considered.

The analysis results of the sensitive bands in the CSC showed that the wavelengths most sensitive to SM variations are predominantly in the Blue and Red Edge ranges [31].

4. Conclusions

In this study, we aimed to fill the gap of previous studies on SM modeling by comparing the potential of multispectral and hyperspectral data for improved SM estimation. Our comparative analysis, using multispectral data from Sentinel-2, Landsat-8/9, and hyperspectral data from a CSC mounted on the Matrice 100 UAV, showed the superiority of hyperspectral UAV data for accurate SM modeling. This finding is significant as it enables the acquisition of high-resolution, site-specific SM information that is critical for optimizing irrigation practices and resource management in precision agriculture. This study not only contributes to a better understanding of SM distribution, but also highlights the potential of UAV-based hyperspectral data collection for improved SM monitoring. Despite the promising results in the accurate estimation of SM, there are still open questions that require further investigation. First, further insights into the influence of environmental factors such as land use change and climate variability on the accuracy of SM modeling need to be gained. Second, the integration of multiple datasets (e.g., the integration of UAV and satellite data) should be explored to improve the accuracy and spatial resolution of SM predictions. Third, investigating the potential of different machine learning algorithms for SM estimation could further improve modeling techniques. Finally, future research efforts should aim to evaluate the capability of radiative transfer models (e.g., MARMIT) for SM estimation.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/rs16111962/s1, Figure S1: Predicted soil moisture maps (%) using Sentinel-2 data for data collection days; Figure S2: Predicted soil moisture maps (%) using Landsat-8/9 data for data collection days; Figure S3: Predicted soil moisture maps (%) for 12 May 2022 using (a) Sentinel-2 and (b) Landsat-8/9 data and for 21 June 2022 using (c) Sentinel-2 and d) Landsat-8/9 data; Figure S4: Uncertainty of predicted soil moisture maps (%) using Sentinel-2 data for data collection days; Figure S5: Uncertainty of predicted soil moisture maps (%) using Landsat-8/9 data for data collection days.

Author Contributions: Conceptualization, H.S.; methodology, H.S. and S.M.; software, H.S., A.N. and A.A.A.; validation, H.S. and Z.M.-D.; formal analysis, H.S.; investigation, H.S. and S.M.; resources, Z.M.-D.; data curation, H.S.; writing—H.S.; writing—review and editing, T.S., R.T.-M. and K.N.; visualization, P.K.; supervision, M.M. and T.S.; project administration, M.M. and T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are available by reasonable email request to the corresponding author at hadi.shokati@ut.ac.ir. The data are not publicly available due to ethical reasons.

Conflicts of Interest: Author Ali Akbar Abkar was employed by the Geographic Information System and Remote Sensing of Agriwtach Company. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Baumann, F.; He, J.-S.; Schmidt, K.; Kühn, P.; Scholten, T. Pedogenesis, Permafrost, and Soil Moisture as Controlling Factors for Soil Nitrogen and Carbon Contents across the Tibetan Plateau. *Glob. Chang. Biol.* **2009**, *15*, 3001–3017. [CrossRef]
2. Zhu, M.; Kong, F.; Li, Y.; Li, M.; Zhang, J.; Xi, M. Effects of Moisture and Salinity on Soil Dissolved Organic Matter and Ecological Risk of Coastal Wetland. *Environ. Res.* **2020**, *187*, 109659. [CrossRef]
3. Pauwels, V.R.N.; Hoeben, R.; Verhoest, N.E.C.; De Troch, F.P.; Troch, P.A. Improvement of TOPLATS-Based Discharge Predictions through Assimilation of ERS-Based Remotely Sensed Soil Moisture Values. *Hydrol. Process.* **2002**, *16*, 995–1013. [CrossRef]
4. Robinson, D.A.; Campbell, C.S.; Hopmans, J.W.; Hornbuckle, B.K.; Jones, S.B.; Knight, R.; Ogden, F.; Selker, J.; Wendroth, O. Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone J.* **2008**, *7*, 358–389. [CrossRef]

5. Dobriyal, P.; Qureshi, A.; Badola, R.; Hussain, S.A. A Review of the Methods Available for Estimating Soil Moisture and Its Implications for Water Resource Management. *J. Hydrol.* **2012**, *458–459*, 110–117. [CrossRef]
6. Darvishi Boloorani, A.; Ranjbareslamloo, S.; Mirzaie, S.; Bahrami, H.A.; Mirzapour, F.; Abbaszadeh Tehrani, N. Spectral Behavior of Persian Oak under Compound Stress of Water Deficit and Dust Storm. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *88*, 102082. [CrossRef]
7. Holzman, M.E.; Rivas, R.; Piccolo, M.C. Estimating Soil Moisture and the Relationship with Crop Yield Using Surface Temperature and Vegetation Index. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 181–192. [CrossRef]
8. Anderson, S.H.; Udawatta, R.P.; Seobi, T.; Garrett, H.E. Soil Water Content and Infiltration in Agroforestry Buffer Strips. *Agroforest. Syst.* **2009**, *75*, 5–16. [CrossRef]
9. Laiolo, P.; Gabellani, S.; Campo, L.; Silvestro, F.; Delogu, F.; Rudari, R.; Pulvirenti, L.; Boni, G.; Fascetti, F.; Pierdicca, N.; et al. Impact of Different Satellite Soil Moisture Products on the Predictions of a Continuous Distributed Hydrological Model. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *48*, 131–145. [CrossRef]
10. Forgotson, C.; O'Neill, P.E.; Carrera, M.L.; Bélair, S.; Das, N.N.; Mladenova, I.E.; Bolten, J.D.; Jacobs, J.M.; Cho, E.; Escobar, V.M. How Satellite Soil Moisture Data Can Help to Monitor the Impacts of Climate Change: SMAP Case Studies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 1590–1596. [CrossRef]
11. Zareie, A.; Amin, M.S.R.; Amador-Jiménez, L.E. Thornthwaite Moisture Index Modeling to Estimate the Implication of Climate Change on Pavement Deterioration. *J. Transp. Eng.* **2016**, *142*, 04016007. [CrossRef]
12. Drusch, M. Initializing Numerical Weather Prediction Models with Satellite-Derived Surface Soil Moisture: Data Assimilation Experiments with ECMWF's Integrated Forecast System and the TMI Soil Moisture Data Set. *J. Geophys. Res. Atmos.* **2007**, *112*, D3. [CrossRef]
13. Narasimhan, B.; Srinivasan, R. Development and Evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for Agricultural Drought Monitoring. *Agric. For. Meteorol.* **2005**, *133*, 69–88. [CrossRef]
14. Xu, L.; Baldocchi, D.D.; Tang, J. How Soil Moisture, Rain Pulses, and Growth Alter the Response of Ecosystem Respiration to Temperature. *Glob. Biogeochem. Cycles* **2004**, *18*, 4. [CrossRef]
15. Pastor, J.; Post, W.M. Influence of Climate, Soil Moisture, and Succession on Forest Carbon and Nitrogen Cycles. *Biogeochemistry* **1986**, *2*, 3–27. [CrossRef]
16. Engman, E.T. Applications of Microwave Remote Sensing of Soil Moisture for Water Resources and Agriculture. *Remote Sens. Environ.* **1991**, *35*, 213–226. [CrossRef]
17. Wood, E.F.; Lettenmaier, D.P.; Zartarian, V.G. A Land-Surface Hydrology Parameterization with Subgrid Variability for General Circulation Models. *J. Geophys. Res. Atmos.* **1992**, *97*, 2717–2728. [CrossRef]
18. Wang, L.; Qu, J.J. Satellite Remote Sensing Applications for Surface Soil Moisture Monitoring: A Review. *Front. Earth Sci. China* **2009**, *3*, 237–247. [CrossRef]
19. Petropoulos, G.P.; Ireland, G.; Barrett, B. Surface Soil Moisture Retrievals from Remote Sensing: Current Status, Products & Future Trends. *Phys. Chem. Earth Parts A/B/C* **2015**, *83–84*, 36–56. [CrossRef]
20. Korres, W.; Reichenau, T.G.; Fiener, P.; Koyama, C.N.; Bogena, H.R.; Cornelissen, T.; Baatz, R.; Herbst, M.; Diekkrüger, B.; Vereecken, H.; et al. Spatio-Temporal Soil Moisture Patterns—A Meta-Analysis Using Plot to Catchment Scale Data. *J. Hydrol.* **2015**, *520*, 326–341. [CrossRef]
21. Mirzaei, S.; Darvishi Boloorani, A.; Bahrami, H.A.; Alavipanah, S.K.; Mousivand, A.; Mouazen, A.M. Minimising the Effect of Moisture on Soil Property Prediction Accuracy Using External Parameter Orthogonalization. *Soil Tillage Res.* **2022**, *215*, 105225. [CrossRef]
22. Gangat, R.; van Deventer, H.; Naidoo, L.; Adam, E. Estimating Soil Moisture Using Sentinel-1 and Sentinel-2 Sensors for Dryland and Palustrine Wetland Areas. *S. Afr. J. Sci.* **2020**, *116*, 1–9. [CrossRef] [PubMed]
23. Hegazi, E.H.; Samak, A.A.; Yang, L.; Huang, R.; Huang, J. Prediction of Soil Moisture Content from Sentinel-2 Images Using Convolutional Neural Network (CNN). *Agronomy* **2023**, *13*, 656. [CrossRef]
24. Adab, H.; Morbidelli, R.; Saltalippi, C.; Moradian, M.; Ghalhari, G.A.F. Machine Learning to Estimate Surface Soil Moisture from Remote Sensing Data. *Water* **2020**, *12*, 3223. [CrossRef]
25. Ghasemloo, N.; Matkan, A.A.; Alimohammadi, A.; Aghighi, H.; Mirbagheri, B. Estimating the Agricultural Farm Soil Moisture Using Spectral Indices of Landsat 8, and Sentinel-1, and Artificial Neural Networks. *J. Geovisualization Spat. Anal.* **2022**, *6*, 19. [CrossRef]
26. Hezarian, F.; Khalilimoghadam, B.; Zoratipour, A.; Nejad, M.F.; Yusefi, A. Assessment of the Capability of Satellite Images in Determining the Topsoil Moisture Content in the Dust Hotspot of Southeastern Ahvaz in Iran. *Eurasian Soil Sc.* **2022**, *55*, 1576–1590. [CrossRef]
27. Baluja, J.; Diago, M.P.; Balda, P.; Zorer, R.; Meggio, F.; Morales, F.; Tardaguila, J. Assessment of Vineyard Water Status Variability by Thermal and Multispectral Imagery Using an Unmanned Aerial Vehicle (UAV). *Irrig Sci* **2012**, *30*, 511–522. [CrossRef]
28. de Lima, R.S.; Li, K.-Y.; Vain, A.; Lang, M.; Bergamo, T.F.; Kokamägi, K.; Burnside, N.G.; Ward, R.D.; Sepp, K. The Potential of Optical UAS Data for Predicting Surface Soil Moisture in a Peatland across Time and Sites. *Remote Sens.* **2022**, *14*, 2334. [CrossRef]
29. Hassan-Esfahani, L.; Torres-Rua, A.; Jensen, A.; McKee, M. Assessment of Surface Soil Moisture Using High-Resolution Multi-Spectral Imagery and Artificial Neural Networks. *Remote Sens.* **2015**, *7*, 2627–2646. [CrossRef]

30. Aboutalebi, M.; Allen, L.N.; Torres-Rua, A.F.; McKee, M.; Coopmans, C. Estimation of Soil Moisture at Different Soil Levels Using Machine Learning Techniques and Unmanned Aerial Vehicle (UAV) Multispectral Imagery. In Proceedings of the Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV, Baltimore, MD, USA, 14 May 2019; SPIE: Bellingham, WA, USA; Volume 11008, pp. 216–226.
31. Ge, X.; Wang, J.; Ding, J.; Cao, X.; Zhang, Z.; Liu, J.; Li, X. Combining UAV-Based Hyperspectral Imagery and Machine Learning Algorithms for Soil Moisture Content Monitoring. *PeerJ* **2019**, *7*, e6926. [CrossRef]
32. Ge, X.; Ding, J.; Jin, X.; Wang, J.; Chen, X.; Li, X.; Liu, J.; Xie, B. Estimating Agricultural Soil Moisture Content through UAV-Based Hyperspectral Images in the Arid Region. *Remote Sens.* **2021**, *13*, 1562. [CrossRef]
33. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sens.* **2015**, *7*, 16398–16421. [CrossRef]
34. Paloscia, S.; Pampaloni, P.; Pettinato, S.; Santi, E. A Comparison of Algorithms for Retrieving Soil Moisture from ENVISAT/ASAR Images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3274–3284. [CrossRef]
35. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013; ISBN 978-1-4614-6848-6.
36. Hengl, T.; Nikolić, M.; MacMillan, R.A. Mapping Efficiency and Information Content. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *22*, 127–138. [CrossRef]
37. Bousbih, S.; Zribi, M.; El Hajj, M.; Baghdadi, N.; Lili-Chabaane, Z.; Gao, Q.; Fanise, P. Soil Moisture and Irrigation Mapping in A Semi-Arid Region, Based on the Synergetic Use of Sentinel-1 and Sentinel-2 Data. *Remote Sens.* **2018**, *10*, 1953. [CrossRef]
38. Ahmad, S.; Kalra, A.; Stephen, H. Estimating Soil Moisture Using Remote Sensing Data: A Machine Learning Approach. *Adv. Water Resour.* **2010**, *33*, 69–80. [CrossRef]
39. 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. A Low-Cost Approach for Soil Moisture Prediction Using Multi-Sensor Data and Machine Learning Algorithm. *Sci. Total Environ.* **2022**, *833*, 155066. [CrossRef]
40. Cheng, M.; Jiao, X.; Liu, Y.; Shao, M.; Yu, X.; Bai, Y.; Wang, Z.; Wang, S.; Tuohuti, N.; Liu, S.; et al. Estimation of Soil Moisture Content under High Maize Canopy Coverage from UAV Multimodal Data and Machine Learning. *Agric. Water Manag.* **2022**, *264*, 107530. [CrossRef]
41. Shokati, H.; Mashal, M.; Noroozi, A.; Mirzaei, S.; Mohammadi-Doqozloo, Z. Assessing Soil Moisture Levels Using Visible UAV Imagery and Machine Learning Models. *Remote Sens. Appl. Soc. Environ.* **2023**, *32*, 101076. [CrossRef]
42. Finn, M.P.; Lewis, M.; Bosch, D.D.; Giraldo, M.; Yamamoto, K.; Sullivan, D.G.; Kincaid, R.; Luna, R.; Allam, G.K.; Kvien, C.; et al. Remote Sensing of Soil Moisture Using Airborne Hyperspectral Data. *GIScience Remote Sens.* **2011**, *48*, 522–540. [CrossRef]
43. Bertalan, L.; Holb, I.; Pataki, A.; Négyesi, G.; Szabó, G.; Kupásné Szalóki, A.; Szabó, S. UAV-Based Multispectral and Thermal Cameras to Predict Soil Water Content—A Machine Learning Approach. *Comput. Electron. Agric.* **2022**, *200*, 107262. [CrossRef]
44. de Oliveira, V.A.; Rodrigues, A.F.; Morais, M.A.V.; Terra, M.d.C.N.S.; Guo, L.; de Mello, C.R. Spatiotemporal Modelling of Soil Moisture in an Atlantic Forest through Machine Learning Algorithms. *Eur. J. Soil Sci.* **2021**, *72*, 1969–1987. [CrossRef]
45. Ben-Dor, E.; Banin, A. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. *Soil Sci. Soc. Am. J.* **1995**, *59*, 364–372. [CrossRef]
46. Hummel, J.W.; Sudduth, K.A.; Hollinger, S.E. Soil Moisture and Organic Matter Prediction of Surface and Subsurface Soils Using an NIR Soil Sensor. *Comput. Electron. Agric.* **2001**, *32*, 149–165. [CrossRef]
47. Roy, S.K.; Shibusawa, S.; Okayama, T. Textural Analysis of Soil Images to Quantify and Characterize the Spatial Variation of Soil Properties Using a Real-Time Soil Sensor. *Precis. Agric* **2006**, *7*, 419–436. [CrossRef]
48. Bouyoucos, G.J. Hydrometer Method Improved for Making Particle Size Analyses of Soils1. *Agron. J.* **1962**, *54*, 464–465. [CrossRef]
49. Oltra-Carrió, R.; Baup, F.; Fabre, S.; Fieuzal, R.; Briottet, X. Improvement of Soil Moisture Retrieval from Hyperspectral VNIR-SWIR Data Using Clay Content Information: From Laboratory to Field Experiments. *Remote Sens.* **2015**, *7*, 3184–3205. [CrossRef]
50. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a Green Channel in Remote Sensing of Global Vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]
51. Rastogi, A.; Stróżecki, M.; Kalaji, H.M.; Łuców, D.; Lamentowicz, M.; Juszczak, R. Impact of Warming and Reduced Precipitation on Photosynthetic and Remote Sensing Properties of Peatland Vegetation. *Environ. Exp. Bot.* **2019**, *160*, 71–80. [CrossRef]
52. Damm, A.; Paul-Limoges, E.; Haghighi, E.; Simmer, C.; Morsdorf, F.; Schneider, F.D.; van der Tol, C.; Migliavacca, M.; Rascher, U. Remote Sensing of Plant-Water Relations: An Overview and Future Perspectives. *J. Plant Physiol.* **2018**, *227*, 3–19. [CrossRef]
53. Rouse, J.W.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; NASA/GSFC: Greenbelt, MD, USA, 1974.
54. McFEETERS, S.K. The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
55. Gao, B. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]
56. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
57. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
58. Díaz-Uriarte, R.; Alvarez de Andrés, S. Gene Selection and Classification of Microarray Data Using Random Forest. *BMC Bioinform.* **2006**, *7*, 3. [CrossRef] [PubMed]

59. Inoubli, R.; Bennaceur, L.; Jarray, N.; Ben Abbes, A.; Farah, I.R. A Comparison between the Use of Machine Learning Techniques and the Water Cloud Model for the Retrieval of Soil Moisture from Sentinel-1A and Sentinel-2A Products. *Remote Sens. Lett.* **2022**, *13*, 980–990. [CrossRef]
60. Caicedo, J.P.R.; Verrelst, J.; Muñoz-Marí, J.; Moreno, J.; Camps-Valls, G. Toward a Semiautomatic Machine Learning Retrieval of Biophysical Parameters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 1249–1259. [CrossRef]
61. Peng, J.; Shen, H.; He, S.W.; Wu, J.S. Soil Moisture Retrieving Using Hyperspectral Data with the Application of Wavelet Analysis. *Environ. Earth Sci.* **2013**, *69*, 279–288. [CrossRef]
62. Shokati, H.; Mashal, M.; Noroozi, A.A.; Mirzaei, S. Evaluating the Accuracy of Precipitation Products Over Utah, United States, Using the Google Earth Engine Platform. *Desert* **2023**, *28*, 145–162. [CrossRef]
63. Sadeghi, M.; Jones, S.B.; Philpot, W.D. A Linear Physically-Based Model for Remote Sensing of Soil Moisture Using Short Wave Infrared Bands. *Remote Sens. Environ.* **2015**, *164*, 66–76. [CrossRef]
64. Yang, X.; Yu, Y.; Li, M. Estimating Soil Moisture Content Using Laboratory Spectral Data. *J. For. Res.* **2019**, *30*, 1073–1080. [CrossRef]
65. Lobell, D.B.; Asner, G.P. Moisture Effects on Soil Reflectance. *Soil Sci. Soc. Am. J.* **2002**, *66*, 722–727. [CrossRef]
66. Jackson, T.J.; Chen, D.; Cosh, M.; Li, F.; Anderson, M.; Walthall, C.; Doriaswamy, P.; Hunt, E.R. Vegetation Water Content Mapping Using Landsat Data Derived Normalized Difference Water Index for Corn and Soybeans. *Remote Sens. Environ.* **2004**, *92*, 475–482. [CrossRef]
67. Yilmaz, M.T.; Hunt, E.R.; Jackson, T.J. Remote Sensing of Vegetation Water Content from Equivalent Water Thickness Using Satellite Imagery. *Remote Sens. Environ.* **2008**, *112*, 2514–2522. [CrossRef]
68. Shafian, S.; Maas, S.J. Improvement of the Trapezoid Method Using Raw Landsat Image Digital Count Data for Soil Moisture Estimation in the Texas (USA) High Plains. *Sensors* **2015**, *15*, 1925–1944. [CrossRef] [PubMed]
69. Hashim, B.M.; Sultan, M.A.; Attyia, M.N.; Al Maliki, A.A.; Al-Ansari, N. Change Detection and Impact of Climate Changes to Iraqi Southern Marshes Using Landsat 2 MSS, Landsat 8 OLI and Sentinel 2 MSI Data and GIS Applications. *Appl. Sci.* **2019**, *9*, 2016. [CrossRef]
70. Sánchez, N.; Alonso-Arroyo, A.; Martínez-Fernández, J.; Camps, A.; González-Zamora, A.; Pablos, M.; Herrero-Jiménez, C.M.; Gumuzzio, A. Multisensor Experiments over Vineyard: New Challenges for the GNSS-R Technique. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-7-W3*, 1299–1303. [CrossRef]
71. Zheng, X.; Feng, Z.; Li, L.; Li, B.; Jiang, T.; Li, X.; Li, X.; Chen, S. Simultaneously Estimating Surface Soil Moisture and Roughness of Bare Soils by Combining Optical and Radar Data. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *100*, 102345. [CrossRef]
72. Paloscia, S.; Pettinato, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Reppucci, A. Soil Moisture Mapping Using Sentinel-1 Images: Algorithm and Preliminary Validation. *Remote Sens. Environ.* **2013**, *134*, 234–248. [CrossRef]
73. Liu, Y.; Qian, J.; Yue, H. Comprehensive Evaluation of Sentinel-2 Red Edge and Shortwave-Infrared Bands to Estimate Soil Moisture. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7448–7465. [CrossRef]
74. Muller, E.; Décamps, H. Modeling Soil Moisture–Reflectance. *Remote Sens. Environ.* **2001**, *76*, 173–180. [CrossRef]
75. Zhang, D.; Zhou, G. Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review. *Sensors* **2016**, *16*, 1308. [CrossRef] [PubMed]
76. Wang, Q.; Li, J.; Jin, T.; Chang, X.; Zhu, Y.; Li, Y.; Sun, J.; Li, D. Comparative Analysis of Landsat-8, Sentinel-2, and GF-1 Data for Retrieving Soil Moisture over Wheat Farmlands. *Remote Sens.* **2020**, *12*, 2708. [CrossRef]
77. Wang, Q.; Li, P.; Pu, Z.; Chen, X. Calibration and Validation of Salt-Resistant Hyperspectral Indices for Estimating Soil Moisture in Arid Land. *J. Hydrol.* **2011**, *408*, 276–285. [CrossRef]
78. Sadeghi, M.; Babaeian, E.; Tuller, M.; Jones, S.B. The Optical Trapezoid Model: A Novel Approach to Remote Sensing of Soil Moisture Applied to Sentinel-2 and Landsat-8 Observations. *Remote Sens. Environ.* **2017**, *198*, 52–68. [CrossRef]
79. Babaeian, E.; Sadeghi, M.; Franz, T.E.; Jones, S.; Tuller, M. Mapping Soil Moisture with the OPtical TRApezoid Model (OPTRAM) Based on Long-Term MODIS Observations. *Remote Sens. Environ.* **2018**, *211*, 425–440. [CrossRef]
80. Jiang, X.; Luo, S.; Ye, Q.; Li, X.; Jiao, W. Hyperspectral Estimates of Soil Moisture Content Incorporating Harmonic Indicators and Machine Learning. *Agriculture* **2022**, *12*, 1188. [CrossRef]
81. Mobasheri, M.R.; Amani, M. Soil Moisture Content Assessment Based on Landsat 8 Red, near-Infrared, and Thermal Channels. *J. Appl. Remote Sens.* **2016**, *10*, 026011. [CrossRef]
82. Zhang, Y.; Liang, S.; Zhu, Z.; Ma, H.; He, T. Soil Moisture Content Retrieval from Landsat 8 Data Using Ensemble Learning. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 32–47. [CrossRef]
83. Foody, G.M.; Cox, D.P. Sub-Pixel Land Cover Composition Estimation Using a Linear Mixture Model and Fuzzy Membership Functions. *Int. J. Remote Sens.* **1994**, *15*, 619–631. [CrossRef]