remote sensing

MDPI

Article

An Approach for Downscaling SMAP Soil Moisture by Combining Sentinel-1 SAR and MODIS Data

Jueying Bai [1], Qian Cui [2], Wen Zhang [1] and Lingkui Meng [1,*]

[1] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China; baijueying@whu.edu.cn (J.B.); wen_zhang@whu.edu.cn (W.Z.)

[2] Information Center (Hydrology Monitor and Forecast Center), Ministry of Water Resources, Beijing 100053, China; cuiqian@mwr.gov.cn

* Correspondence: lkmeng@whu.edu.cn

Received: 31 August 2019; Accepted: 19 November 2019; Published: 21 November 2019

Abstract: A method is proposed for the production of downscaled soil moisture active passive (SMAP) soil moisture (SM) data by combining optical/infrared data with synthetic aperture radar (SAR) data based on the random forest (RF) model. The method leverages the sensitivity of active microwaves to surface SM and the triangle/trapezium feature space among vegetation indexes (VIs), land surface temperature (LST), and SM. First, five RF architectures (RF1–RF5) were trained and tested at 9 km. Second, a comparison was performed for RF1–RF5, and were evaluated against in situ SM measurements. Third, two SMAP-Sentinel active–passive SM products were compared at 3 km and 1 km using in situ SM measurements. Fourth, the RF5 model simulations were compared with the SMAP L2_SM_SP product based on the optional algorithm at 3 km and 1 km resolutions. The results showed that the downscaled SM based on the synergistic use of optical/infrared data and the backscatter at vertical–vertical (VV) polarization was feasible in semi-arid areas with relatively low vegetation cover. The RF5 model with backscatter and more parameters from optical/infrared data performed best among the five RF models and was satisfactory at both 3 km and 1 km. Compared with L2_SM_SP, RF5 was more superior at 1 km. The input variables in decreasing order of importance were backscatter, LST, VIs, and topographic factors over the entire study area. The low vegetation cover conditions probably amplified the importance of the backscatter and LST. A sufficient number of VIs can enhance the adaptability of RF models to different vegetation conditions.

Keywords: soil moisture; downscaling; random forest; SMAP; SAR data; optical/infrared data

1. Introduction

Soil moisture (SM) is an important component in global water and energy cycles, and it plays a crucial role in driving hydrological and land surface processes [1]. Therefore, SM monitoring is valuable in many applications including drought monitoring [2,3], flood warnings [4–6], evapotranspiration estimations [7–9], and water resource management [10,11].

SM can be obtained from station-based measurements, data assimilation products based on land surface models, and remote sensing monitoring data. In situ SM observations cannot represent the SM status at the regional scale. In areas with sparsely distributed stations, data assimilation products by means of station-based measurements cannot fully reflect the spatial and temporal variations of surface SM [12,13]. Among the remote sensing technologies, microwave remote sensing can penetrate the soil surface to directly detect the surface SM content by utilizing the large difference in dielectric properties between dry soil and liquid water. Passive microwaves can generate accurate surface SM estimates because they are less affected by vegetation, soil surface roughness, topography, and water content [14,15]. At present, the soil moisture active passive (SMAP) [14] and soil moisture and ocean

salinity (SMOS) satellites [10] based on L-band passive microwaves can provide high-accuracy global daily SM products [16]. However, coarse-resolution passive microwave SM data cannot reflect the detailed distribution of surface SM; therefore, many researchers have downscaled coarse-resolution passive microwave SM data based on fine-resolution auxiliary data [17–23].

Some downscaling methods are based on active or passive microwave data including downscaling coarse-resolution microwave brightness temperature (TB) data or coarse-resolution SM data based on microwave backscatter data [22,24–30], and downscaling low-frequency passive microwave data based on high-frequency passive microwave data [31–33]. Active microwave techniques offer higher spatial resolutions than passive microwave techniques. Downscaling methods based on active microwaves usually leverage the linear or near-linear relationship between microwave TB or SM data and microwave backscatter data in a time series. Vegetation and surface roughness are generally considered time-invariant during long-time series [26]. However, active microwaves are greatly affected by vegetation and surface roughness [34]; therefore, their applications in most areas are limited [35]. Another assumption is that invariance of vegetation and surface roughness occurs in adjacent observations (several days); thus, these techniques require frequent co-located observations. Additionally, a few single-pass retrieval models have been developed because the invariance of vegetation and surface roughness may be invalid for most existing synthetic aperture radar (SAR) missions with limited revisit frequencies (dozens of days) [25,27]. A single-pass retrieval model has been applied in the SMAP mission [25]. The 1 km and 3 km SMAP SM products are obtained by disaggregating the coarse-resolution SMAP TB or SM data using Sentinel-1 SAR data; however, a trade-off is observed between the finer spatial resolution and noise associated with SAR data [25]. Overall, reducing the uncertainty caused by complex surface scattering conditions on vegetation and surface roughness is critical for SM downscaling methods depending on active microwaves [25,27].

Some downscaling methods based on fine-resolution optical/infrared data mainly use triangular/trapezoidal feature space (non-linear fitting methods with VIs and LST [15,36–39], Peng's methods [40,41], the University of California, Los Angeles (UCLA) method [42]), the soil evaporation process (evaporative fraction (EF), soil evaporative efficiency (SEE), disaggregation based on physical and theoretical scale change (DisPATCh) [43–45])), and thermal inertia theory [46]. In the downscaling process, a SM proxy variable is usually used to construct the relationship model with SM such as Peng's methods [40,41], UCLA, EF, SEE, and DisPATCh. However, Peng et al. [47] and Malbéteau et al. [45] showed that simple SM proxy variables could not easily provide accurate findings in terms of the temporal and spatial variations in SM. The thermal inertia theory is more commonly applied to bare soil conditions [18]. Additionally, the triangular/trapezoidal feature space of LST and VIs [48–50] is the main theoretical basis of the non-linear relationship between land surface parameters and SM status. Zhao et al. indicated that there are large variations in the determination coefficient (with a range of 0.2–0.9) based on fitting the optical/infrared observations and AMSR-E surface SM using polynomial regression [51,52]. Compared with polynomial fitting, machine learning methods (e.g., RF and neural network) can better express the non-linear relationship between different variables and SM [32,34,38,39,51–53]. Overall, optical/infrared data can provide fine-resolution land surface parameters to downscale coarse SM, but this can only indirectly reflect surface SM values [18]. Additionally, the attenuation of solar radiation via opaque covering media (e.g., atmosphere, vegetation, etc.) is a factor for reducing the accuracy of the downscaled SM data.

Several researchers have combined SAR data and optical data as inputs with a machine learning method for fine-resolution SM mapping. The combination of SAR backscatter and the normalized difference vegetation index (NDVI) has been used for surface SM mapping in machine learning methods (e.g., neural network [54] and support vector regression [55]). Moreover, El Hajj et al. [56] developed a neural network incorporating SAR backscatter, NDVI, and weather forecasts (e.g., precipitation and temperature) as the main input variables for SM retrieval. The SAR and NDVI data were generated from the calibrated water cloud model (WCM) and integral equation model (IEM). NDVI can explain the contribution of vegetation, and the weather forecast information about the SM status can constrain

the range of SM values to reduce the error caused by soil surface roughness [56]. Although only a single vegetation index (VI) and air temperature were used in the above SM retrieval, the results indicated that the machine learning method with SAR data, VIs, and LST as input variables might be effective for downscaling coarse SM data. Qiu et al. investigated the suitability of different VIs for the parameterization of vegetation water content in SAR-based SM retrievals [57]. This study indicated that the combination of different VIs might be necessary to express different vegetation conditions. Additionally, Amazirh et al. [58] constructed a relationship model among radar backscatter, NDVI, and LST by polynomial fitting for the disaggregation of MODIS LST data in low vegetation cover areas. Therefore, the machine learning models combining optical/infrared data and SAR data may be feasible for downscaling passive microwave SM.

In light of the above insights, we proposed an RF model combining SAR data and optical/infrared data for downscaling passive microwave SM. In this study, the input variables were from SAR data, optical/infrared data, and topography data, and the output variable was SM. The VV polarization backscatter may be an important input variable due to the sensitivity of SAR data to surface SM, especially in bare soil and low vegetation cover conditions. The VIs can explain the contribution of vegetation to SAR data. Additionally, the triangular/trapezoidal feature space among LST, VIs, and SM can also contribute to the relationship model. Although some RF models that are mainly dependent on LST/VIs feature space have been developed, feature space is relatively unstable in low vegetation cover conditions [53,59,60]. The surface elevation and slope related to land surface energy and surface backscatter were used for constructing models [38,53]. In the training process, the output variable was from 9 km enhanced SMAP SM data, and the input variables were from 9 km aggregated SAR data, optical data, and topography data. We assumed that the relationship models were spatial-scale independent [30,36,37,53,58,61,62]. The trained RF models at a 9 km resolution were then directly applied to the input variables at 3 km and 1 km resolutions to obtain the corresponding 3 km and 1 km SM data. In the present study, three sets of comparative tests (RF1 vs. RF2, RF1 vs. RF3, and RF4 vs. RF5) were first established to show the effects of backscatter at VV polarization, elevation, and other auxiliary input variables (another VIs and slope), respectively. RF1 and RF3 were mainly based on the triangular/trapezoidal feature space among LST, NDVI, and SM. Second, the 3 km and 1 km downscaled SM from RF1–RF5 were compared using in situ SM. Third, the two SMAP-Sentinel active–passive SM products were compared using in situ SM measurements at 3 km or 1 km. Fourth, the RF5 model simulations were compared with the SMAP L2_SM_SP product based on the optional algorithm at 3 km and 1 km.

2. Materials and Methodology

2.1. Study Area and Data

2.1.1. Study Area and In Situ Soil Moisture (SM) Measurements

In this study, an area in northern China was selected as the study area (Figure 1). The elevation is highly variable, ranging from approximately 27 to 3455 m. The terrain and landforms are complex, with mountains, plateaus, and basins. The region presents a temperate continental climate with little rainfall. Most of the area is located in a semi-arid region, and the SM content is usually at a low level, especially under the colder and drier conditions of winter. Vast dry land and complex and scattered land cover types including dry land, woodland, and sandy land are present in the eastern part of the study area. Thus, fine-resolution SM monitoring is particularly important.

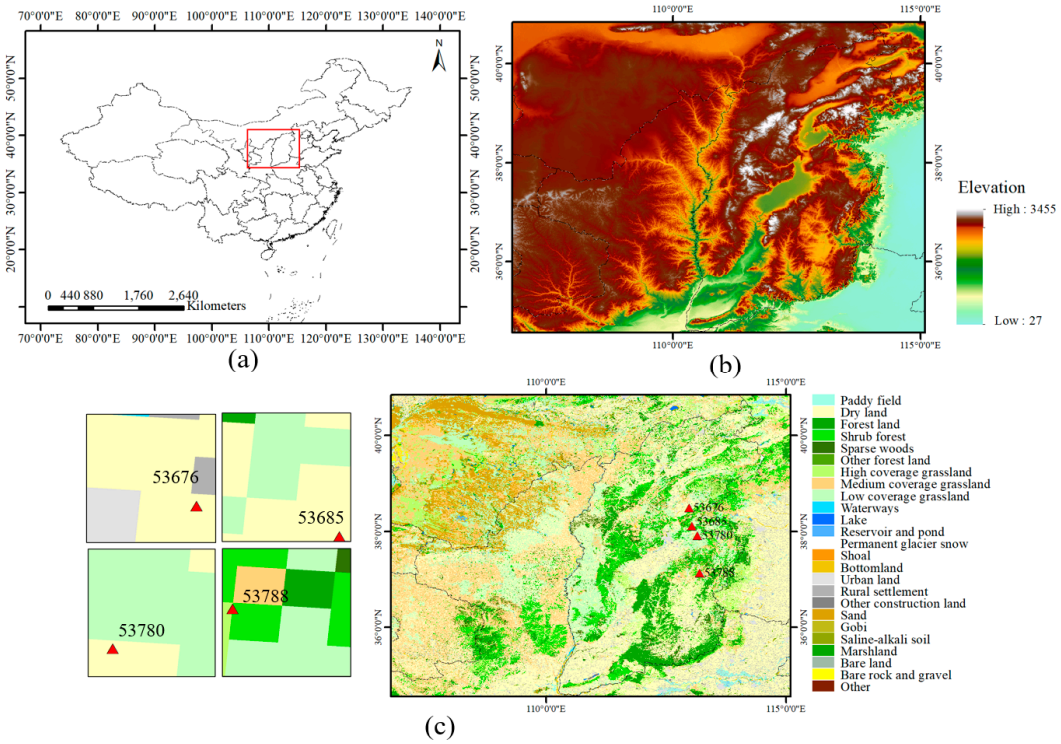

Figure 1. The study area in northern China with the geographical locations of four soil moisture (SM) stations. (**a**) Location of the study area. (**b**) Elevation from the shuttle radar topography mission (SRTM) digital surface model (DSM). (**c**) Land cover map and the 3 km footprints over the four stations: 53676, 53685, 53780, and 53788. Details of the background maps are described in Section 2.1.5.

In situ SM measurements of four representational stations were collected to evaluate the more overlapping observations of Sentinel-1 and moderate-resolution imaging spectroradiometer (MODIS) satellites in agricultural regions. The stations (53676, 53685, 53780, and 53788) operated by the China Meteorological Administration (CMA) monitor SM content at depths of 10, 20, and 40 cm at 8:00 a.m. and 3:00 p.m. local time. In this study, we evaluated the remote sensing SM data depending on the in situ SM at 3:00 p.m. The land cover types of stations 53676, 53685, and 53780 are dry land, and the land cover type of station 53788 is shrubland (Table 1). The four stations are at a high elevation with a moderately steep slope.

Table 1. Attributes of the stations with SM observations.

ID	Latitude & Longitude	Elevation (m)	Land Cover
53676	112°58′32″N, 38°29′42″E	757	Dry land
53685	113°02′16″N, 38°06′47″E	1236	Dry land
53780	113°09′06″N, 37°54′37″E	1084	Dry land
53788	113°12′00″N, 37°07′59″E	1389	Shrubland

Due to the lack of SM monitoring stations at 5 cm in China, SM values at 10 cm were used instead, as in past research [41,63,64]. Although the measurement depths are inconsistent, there is a strong correlation between the SM values of the two continuous soil layers [13]. In theory, the measuring depth inconsistency should have little effect on the correlation between the SMAP SM and in situ SM values.

2.1.2. Soil Moisture Active Passive (SMAP) SM Products

The SMAP satellite was launched by the National Aeronautics and Space Administration (NASA) in January 2015, and it is operated within a 685 km near-polar orbit [14]. The satellites, with ascending

(6:00 p.m.) and descending (6:00 a.m.) modes and carrying an L-band radar (1.41 GHz) and radiometer (1.26 GHz), are dedicated to monitoring global surface SM and freeze–thaw states. The SMAP can cover the entire globe in 2–3 days with three spatial resolutions of 3, 9, and 36 km. SMAP SM products are defined as L2, L3, and L4, where L2 is semi-orbit data, L3 is a composite of daily SM estimates, and L4 is the model assimilation products. Unfortunately, the L-band radar became inoperable in July 2015, and only the L-band radiometer is still operating, providing radiometer-based SM data at a coarse resolution (36 km). To restore the finer spatial resolution of SMAP products, the Backus-Gilbert interpolation technique was applied to antenna temperature in the original SMAP L1B TB product [29] to generate a SMAP-enhanced TB product on a 9 km Equal-Area Scalable Earth version 2 (EASEv2) grid; then, the 9 km enhanced passive SM product (L3_SM_P_E) was obtained [65]. Then, Sentinel-1 C-band SAR backscatter data were used as background values to disaggregate 9 km SMAP TB and SM data (L3_SM_P_E), to generate 3 km/1 km L2_SM_SP based on the baseline algorithm (L2_SM_SP$_{\text{baseline}}$), and L2_SM_SP based on the optional algorithm (L2_SM_SP$_{\text{optional}}$) [66]. Both L2_SM_SP$_{\text{baseline}}$ and L2_SM_SP$_{\text{optional}}$ are SMAP-Sentinel active–passive L2 SM products (L2_SM_SP).

In this study, 9 km SMAP L3_SM_P_E ascending data (Version 2) and 3 km/1 km L2_SM_SP data (Version 2) from 1 August, 2017 to 1 August, 2018 were downloaded from NASA's Earth Observing System Data and Information System (EOSDIS) [67]. The 9 km SMAP L3_SM_P_E ascending data were used in this study because they have more similar observation times to the available Sentinel-1 data in most areas of northern China.

2.1.3. Sentinel-1 Data

Sentinel-1 is operated in a polar sun-synchronous orbit of 693 km, and it carries a C-band (5.405 GHz) SAR. The similar orbit configurations of the Sentinel-1 and SMAP satellites are key to the synergy of SMAP with Sentinel-1, and these satellites have adequately overlapping image scenes and minimal overpass time differences.

In this study area, Sentinel-1A interferometric wide swath (IW) ground range detected (GRD) observations with a pixel spacing of 10 m × 10 m in ascending mode were recorded. The 136 Sentinel-1A images with VV and vertical-horizontal (VH) polarizations between 1 August, 2017 and 1 August, 2018 were acquired by the Copernicus Open Access Hub [68]. The postprocessing of Sentinel-1 SAR data was performed in ENVI SARscape 5.5 Basic Module including speckle filtering (5 × 5 refined Lee filter [69]), geocoding, radiometric calibration, and normalization. The local incidence angle (LIA) from a 30 m resolution SRTM DSM was used to compute the scattering area for radiometric calibration. The radiometric normalization was performed by the semi-empirical method to empirically correct the effects of the unequal incidence angles and the incidence angles of all Sentinel-1A images were normalized to 40 degrees. In the semi-empirical method, a linear regression between the cosine of the LIA and the backscattering coefficient in logarithmic form was performed. Finally, the backscattering coefficients at VV polarization (σ^0_{vv}) in a linear unit were averaged within 9 km, 3 km, and 1 km grids in the WGS 1984 coordinate system [26,56]. The 9, 3, and 1 km aggregated Sentinel-1 SAR data were used to build the model and assess its accuracy.

2.1.4. Moderate-resolution Imaging Spectroradiometer (MODIS) Products

The MODIS satellites Terra and Aqua operate on a solar synchronous polar orbit. The Terra (10:30 a.m. local time (descending)/10:30 p.m. local time (ascending)] and Aqua (1:30 a.m. local time (descending)/1:30 p.m. local time (ascending)] satellites have different overpass times. MODIS data contain many land surface parameters including VIs, LST, and surface albedo (ALB).

For this study, MODIS products for 1 August, 2017 through 1 August, 2018 including the 1 km daily LST product (MYD11A1), 1 km 16-day NDVI and enhanced vegetation index (EVI) product (MOD13A2), 500 m 8-day leaf area index (LAI) product (MCD15A2H) and 500 m 16-day albedo product (MCD43A3) were downloaded from NASA's EOSDIS [50,67,70–72]. In addition, the 500 m 8-day surface reflectance product (MOD09A1) was downloaded to calculate the normalized difference water

index (NDWI) [73]. The LST data were measured during the ascending mode of the Aqua satellite (1:30 p.m.). The overpass times of the Aqua ascending data (1:30 p.m.) and SMAP ascending data (6:00 p.m.) are similar; thus, errors caused by rainfall and the change in surface temperature can be minimized. Furthermore, the daytime LST, which exhibits large variations, is sensitive to surface SM [47,53,74]. Finally, cloud-free scenes in MODIS images were averaged over 9 km, 3 km, and 1 km grids, respectively; the LST data below zero degrees centigrade were removed.

2.1.5. Other Geospatial Data

A DSM, land cover data, precipitation data, and soil map were also used. The SRTM DSM with 30 m spatial resolution can be downloaded from the United States Geological Survey (USGS) Earth Explorer [75]. The land cover data with 1 km resolution, namely, WESTDC_Land_Cover_Products1.0, were acquired from the Cold and Arid Regions Science Data Centre at Lanzhou [76,77]. The precipitation data were provided by NASA's Global Precipitation Measurement (GPM) mission [78].

2.2. RF-Based Downscaling Method

The random forest (RF) technique [79] is a widely used machine learning model for solving classification, regression, and other tasks. RF is an ensemble learning model that leverages multiple weak classifiers (decision trees) to improve the generalization and reduce over-fitting phenomena. In a regression, the mean predicted values of all independent decision trees are regarded as the RF model outputs. The adaptive, randomized, and decorrelated features make RF suitable for complex and highly non-linear relationship models. RF is simple and flexible, and it is less affected than other models by hyper-parameters. Previous studies have also shown that the RF model is effective in complex non-linear fitting and that it can represent a feasible SM downscaling model [39,53].

Downscaling SMAP SM utilizes detailed information from fine-spatial-resolution auxiliary data. Figure 2 shows the general process of downscaling SMAP SM, and the text description is as above (Section 1). Surface SM from the 9 km SMAP L3_SM_P_E is the target variable. The σ^0_{vv} from Sentinel-1 SAR data; the VIs, LST and ALB from MODIS optical/infrared data; and the elevation and slope from topographic data are all input variables.

The 9 km SM data can provide a more detailed SM distribution and more adequate SM values than the 36 km SMAP SM data [80–82]. Although the SMAP SM products with the unbiased root mean square error (ubRMSE) requirement of the SMAP mission (0.04 cm^3/cm^3) have a satisfactory accuracy, the products are still affected by many complex factors (e.g., topography, vegetation, surface temperature,' and RFI). At the same time, RF is a suitable model because it is less affected by sample errors and has a strong ability to resist over-fitting.

The σ^0_{vv} from Sentinel-1 SAR data, the common parameters (i.e., NDVI, EVI, NDWI [73], LAI, LST, and ALB) from MODIS optical/infrared data and topographic data (i.e., elevation and slope) are all input variables. First, C-band radar can penetrate the soil surface to detect the surface SM content directly. Compared with VH polarization, VV polarization is more sensitive to SM and less affected by vegetation and soil surface roughness [56]. Furthermore, VIs based on optical data can describe the contribution to and attenuation of vegetation on the total backscattering coefficients from Sentinel-1 SAR data [56]. The LST of the SM state can constrain the range of SM values to reduce the error caused by surface roughness. Second, the triangular/trapezoidal feature space of the VIs (i.e., NDVI, EVI, NDWI, and LAI) and LST is also a pivotal physical basis for downscaling SM [70]. In addition, surface SM has an impact on ALB [42]. Third, the effect of surface elevation and slope associated with land surface energy and surface backscatter are also used as input variables due to the wide range and variations in the surface elevations and slopes in this study area [23]. The slope data were calculated from elevation using ENVI software.

Nevertheless, it is necessary to outline the limitations of the RF models in this study. First, the depths of the soil layer contributing to the L-band radiometer signal (~5 cm) and the C-band SAR backscatter signal (~2 cm) are different [14,20,36]. The penetration depths of microwaves mainly

depend on microwave frequency, soil texture, and SM content [83]. Nonetheless, Sentinel-1 SAR data may provide valuable information for downscaling coarse-resolution SMAP SM [26,32]. Second, the MODIS daily LST values could be significantly different from LST values at the time of SMAP acquisition. SM status may be different at the two acquisition times depending on the soil type [36]. Additionally, the sensing depths of the SMAP L-band for bare soil and the MODIS thermal infrared band are ~5 cm and ~1 mm, respectively [14,36]. The thermal regime of the two depths may be quite different. Nonetheless, MODIS LST data may provide valuable information for downscaling coarse-resolution SMAP SM [36,39,56,58].

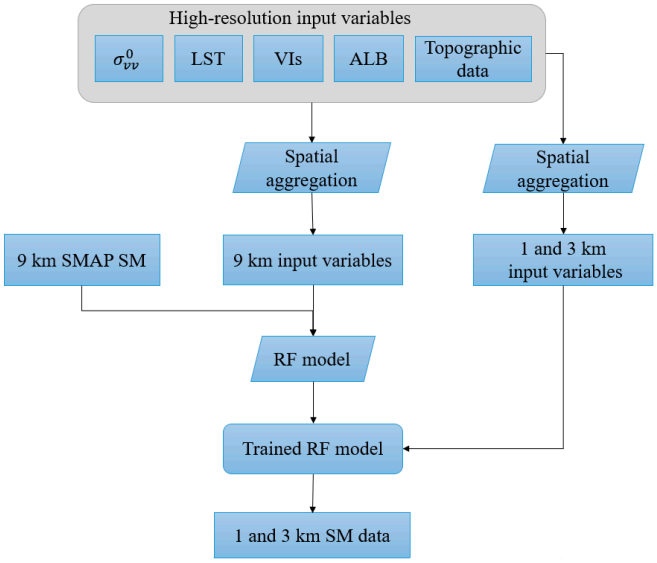

Figure 2. Schematic flow of downscaling coarse-resolution SM.

Therefore, the RF1–RF5 models were established for three sets of comparative tests. RF1 vs. RF2, RF1 vs. RF3, and RF4 vs. RF5 were used to show the effects of σ^0_{vv}, elevation, and other auxiliary input variables (another VI and slope), respectively. Equations (1)–(5) show the five RF models:

$$\mathrm{SM} = RF1\ (\mathrm{LST},\ \mathrm{NDVI}) \quad (1)$$

$$\mathrm{SM} = RF2\left(\mathrm{LST},\ \mathrm{NDVI},\ \sigma^0_{vv}\right) \quad (2)$$

$$\mathrm{SM} = RF3\ (\mathrm{LST},\ \mathrm{NDVI},\ \mathrm{elevation}) \quad (3)$$

$$\mathrm{SM} = RF4\left(\mathrm{LST},\ \mathrm{NDVI},\ \mathrm{elevation},\ \sigma^0_{vv}\right) \quad (4)$$

$$\mathrm{SM} = RF5\left(\mathrm{LST},\ \mathrm{NDVI},\ \mathrm{EVI},\ \mathrm{NDWI},\ \mathrm{LAI},\ \mathrm{ALB},\ \mathrm{elevation},\ \mathrm{slope},\ \sigma^0_{vv}\right) \quad (5)$$

3. Results

3.1. Evaluation of the Random Forest (RF) Models at 9 km Resolution

In this study, one year (1 August, 2017 to 1 August, 2018) of available overlapping SMAP, MODIS, and Sentinel-1 acquisitions in the entire study area were used to evaluate the RF1–RF5 models at 9 km. In the model training and testing process, there were 14,629 sets of samples with a 9 km resolution, and a 10-fold cross-validation method was adopted for parameter adjustment to ensure the robustness and generalizability of the RF models. In other words, the 14,629 sets of samples were randomly and evenly divided into 10 parts, and then the training and testing were circulated 10 times with nine parts as the training dataset and the remaining part as the testing dataset. In addition, the number of trees was set to approximately 120 by the trial-and-error method.

First, the importance of the different variables was analyzed based on the range and distribution of these samples at 9 km (Figure 3). The results revealed an increase in the mean squared error (%) (%IncMSE), which indicates the corresponding increase in mean square error (MSE) when one variable is set to a random number. Overall, the input variables in decreasing order of importance were σ^0_{vv}, LST, VIs, and topographic factors according to the scores. Among the VIs, NDVI and LAI were more important than NDWI and EVI.

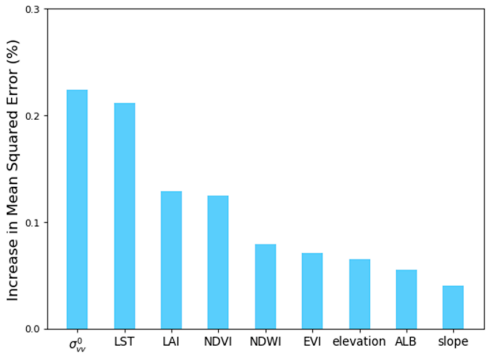

Figure 3. The variable importance scores.

Second, the performances of RF1–RF5 generally agreed with the variable importance (Figure 4). The 10-fold cross-validation method was adopted; thus, the entire testing dataset with a 9 km resolution (14,629 sets of samples) was used for the evaluation of the RF models. For the entire testing dataset, the RF1–RF5 models captured SM variations with R values of approximately 0.855, 0.912, 0.901, 0.931, and 0.955, and ubRMSE values of approximately 0.023, 0.018, 0.019, 0.016, and 0.013 cm^3/cm^3. Among the RF models, RF5, with σ^0_{vv} and more input variables from the optical/infrared data, performed best. The RF2 model (with input variables of NDVI, LST, and σ^0_{vv}) exhibited significant improvement over the RF1 model (with input variables of NDVI and LST). The elevation can optimize the model according to the difference between RF3 and RF1. The comparison between RF4 and RF5 indicates that the RF model with more VIs and topographic information can increase the accuracy further to obtain a more robust RF model. In addition, the slopes were all less than 1, which is similar to [84].

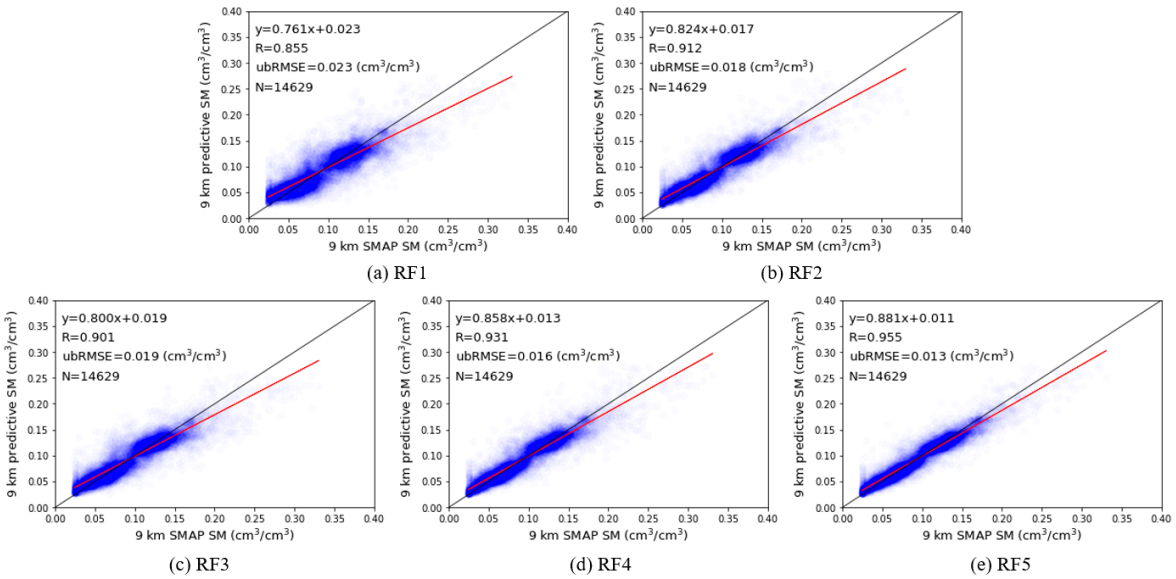

Figure 4. Comparison between the 9 km predictive SM data based on the RF1–RF5 models and 9 km SMAP L3_SM_P_E SM data. N is the number of samples.

3.2. Evaluation of the Downscaled SM

To compare the RF1–RF5 models, the downscaled SM estimates at 3 km and 1 km spatial resolutions were evaluated by the point-scale observations at four stations. The overlapping Sentinel-1 and cloud-free MODIS data between 1 August, 2017 and 1 August, 2018 were used to calculate 3 km and 1 km SM over the four stations. Figure 5, and Tables 2 and 3 show the assessment of the 3 km and 1 km downscaled SM data from RF1–RF5, respectively. The performance of RF1 in Figure 5a indicated that a slight underestimation would occur over dry land. Although relatively large spatial heterogeneity was observed over station 53788 (Figure 1c), the performance of the 3 km and 1 km SM estimates over station 53788 were satisfactory. According to the difference between RF1 and RF2, the added input variable (σ^0_{vv}) increased the average R from 0.51 to 0.69 and decreased the corresponding average ubRMSE from 0.033 cm^3/cm^3 to 0.023 cm^3/cm^3 (Tables 2 and 3). The elevation can slightly increase the accuracy of the downscaled SM according to the difference between RF3 and RF1. The RF5 model, with more VIs and slope data, performed better than the RF4 model; therefore, adequate land surface parameters from optical/infrared data are needed to obtain a robust non-linear relationship. In general, the RF5 model, with σ^0_{vv} and several common land surface parameters from optical/infrared data, performed the most stably at both the 3 km and 1 km spatial resolutions.

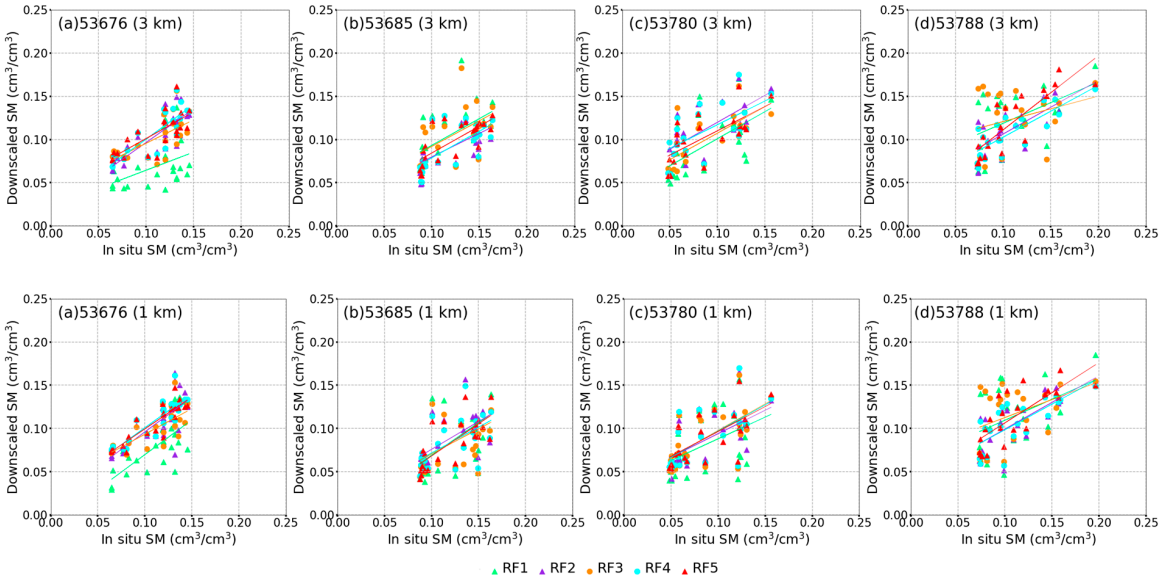

Figure 5. Scatter plots of downscaled SM estimates based on RF1–RF5 at 3 km (first row) and 1 km (second row) vs. point-scale in situ SM measurements. The results are shown for each station: 53676 (first column), 53685 (second column), 53780 (third column), and 53788 (fourth column).

Table 2. Results comparison summary for the 3 km downscaled SM based on RF1–RF5 vs. in situ SM. The number of samples at stations 53676, 53685, 53780, and 53788 is 22, 21, 20 and 20, respectively. Avg is short for average.

ID	RF1		RF2		RF3		RF4		RF5	
	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)
53676	0.43	0.028	0.80	0.017	0.66	0.021	0.80	0.016	0.78	0.017
53685	0.53	0.029	0.59	0.024	0.53	0.030	0.67	0.021	0.81	0.016
53780	0.60	0.031	0.68	0.025	0.72	0.023	0.66	0.026	0.70	0.024
53788	0.46	0.036	0.74	0.023	0.32	0.039	0.80	0.020	0.83	0.020
Avg	0.50	0.031	0.70	0.022	0.59	0.028	0.72	0.021	0.78	0.019

Table 3. Results comparison summary for the 1 km downscaled SM based on RF1–RF5 vs. in situ SM. The number of samples at stations 53676, 53685, 53780, and 53788 is 22, 21, 20 and 20, respectively. Avg is short for average.

ID	RF1		RF2		RF3		RF4		RF5	
	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)
53676	0.71	0.022	0.74	0.020	0.72	0.018	0.83	0.015	0.89	0.012
53685	0.39	0.047	0.65	0.023	0.36	0.043	0.57	0.026	0.74	0.020
53780	0.46	0.035	0.59	0.029	0.64	0.027	0.64	0.027	0.67	0.025
53788	0.47	0.036	0.70	0.024	0.41	0.036	0.77	0.021	0.79	0.021
Avg	0.51	0.035	0.67	0.024	0.53	0.031	0.70	0.022	0.77	0.020

Although the RF models exhibited satisfactory performance according to the evaluation based on the testing dataset at 9 km, the accuracy of the downscaled SM at 3 and 1 km may decrease. In general, the errors of downscaled SM mainly arise from three factors: inherent uncertainty in the original coarse SM data, model fitting error, and spatial differences among the coarse SM data, fine-resolution SM data, and point-scale SM measurements [51].

3.3. Evaluation of SMAP SM Products

To compare the performance of the two L2_SM_SP products at 3 km or 1 km, the overlapping observations of L2_SM_SP$_{baseline}$ and L2_SM_SP$_{optional}$ between 1 August, 2017 and 1 August, 2018 were evaluated against in situ SM measurements at four sites (Figure 6 and Table 4). The closest SMAP data in ascending/descending modes were spatially matched with Sentinel-1 SAR data in the ascending mode to generate L2_SM_SP. Although the two fine-resolution L2_SM_SP products can monitor the SM variations, their performances were different. The L2_SM_SP$_{optional}$ product showed a moderate performance (3 km L2_SM_SP$_{optional}$ product with an average R of 0.68 and an average ubRMSE of 0.033 cm^3/cm^3 and 1 km L2_SM_SP$_{optional}$ product with an average R of 0.43 and an average ubRMSE of 0.046 cm^3/cm^3), whereas the L2_SM_SP$_{baseline}$ product performed worse (3 km L2_SM_SP$_{baseline}$ product with an average R of 0.53 and an average ubRMSE of 0.052 cm^3/cm^3 and 1 km L2_SM_SP$_{baseline}$ product with an average R of 0.29 and an average ubRMSE of 0.074 cm^3/cm^3). Clearly, L2_SM_SP$_{optional}$ performed significantly better than L2_SM_SP$_{baseline}$. Overall, although the combinations of low-spatial-resolution SM/TB data and high-spatial-resolution SAR data can achieve higher-resolution SM retrievals, these gains came at the cost of degradation in the temporal statistics of disaggregated TB and retrieved SM data [66]. Additionally, note that more missing data existed in the L2_SM_SP$_{baseline}$ product when compared with the L2_SM_SP$_{optional}$ product.

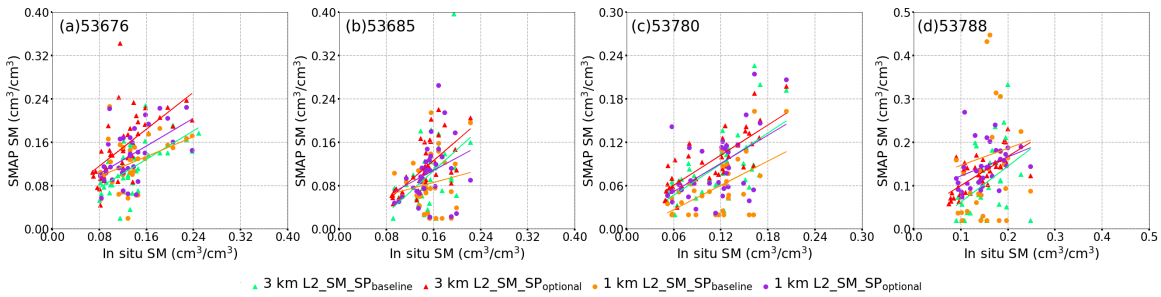

Figure 6. Scatterplots of the SMAP L2_SM_SP products (3 km and 1 km) vs. the in situ SM at the four SM stations.

Table 4. Results comparison summary of the L2_SM_SP products (3 km and 1 km) vs. in situ SM. The number of samples at stations 53676, 53685, 53780, and 53788 was 36, 37, 34, and 33, respectively. Avg is short for average.

ID	$L2\_SM\_SP_{baseline}$ (3 km)		$L2\_SM\_SP_{optional}$ (3 km)		$L2\_SM\_SP_{baseline}$ (1 km)		$L2\_SM\_SP_{optional}$ (1 km)	
	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)	R	ubRMSE (cm^3/cm^3)
53676	0.61	0.038	0.61	0.046	0.39	0.047	0.52	0.042
53685	0.39	0.041	0.60	0.035	0.17	0.051	0.31	0.049
53780	0.62	0.037	0.78	0.025	0.54	0.035	0.51	0.042
53788	0.49	0.069	0.73	0.025	0.18	0.164	0.50	0.050
Avg	0.53	0.052	0.68	0.033	0.29	0.074	0.43	0.046

3.4. Comparison between the Downscaled SM and SMAP SM

The RF1–RF5 models and SMAP L2_SM_SP products were evaluated separately because more in situ SM in one year can be used for assessing them. According to the above results, RF5 performed the best among the RF models and $L2\_SM\_SP_{optional}$ performed better than $L2\_SM\_SP_{baseline}$. Then we further compared the downscaled SM estimates from RF5 with the $L2\_SM\_SP_{optional}$ product. Although only cloud-free overlapping SMAP, MODIS, and Sentinel-1 scenes were available for the comparison, it was still meaningful.

First, a temporal comparison between the downscaled SM from RF5 and SMAP $L2\_SM\_SP_{optional}$ data at 3 km and 1 km resolutions was conducted (Figures 7 and 8, and Table 5). The overlapping observations of $L2\_SM\_SP_{optional}$ and SM estimates based on RF5 at 3 km or 1 km were extracted for further temporal comparison. As shown in Figure 8 and Table 5, a comparable performance was exhibited between the 3 km SM estimates and 3 km $L2\_SM\_SP_{optional}$ due to similar R and ubRMSE values. More outliers were observed in the 1 km $L2\_SM\_SP_{optional}$ than in the 1 km SM based on RF5, likely caused by SAR data (Figure 7 and Table 5). In addition, it appeared that the higher the spatial resolution, the more significant the effect of noise related to SAR data (Figure 7). The fine resolution of the $L2\_SM\_SP_{optional}$ product came at the cost of degradation in the temporal statistics of disaggregated SM data [66]. Overall, although optical/infrared data can only indirectly reflect surface SM values, RF5 with less noise related to radar data can generate SM data with a stable performance at both 3 km and 1 km. Compared with $L2\_SM\_SP_{optional}$, RF5 is more superior at 1 km. Note that an extremely uncorrelated relationship was obtained between the 1 km $L2\_SM\_SP_{optional}$ product and in situ SM data (Figure 8 and Table 5) because the overlapping SM observations were mainly distributed in the dry season and the effect of vegetation and surface roughness was heavier at the low SM content level [85].

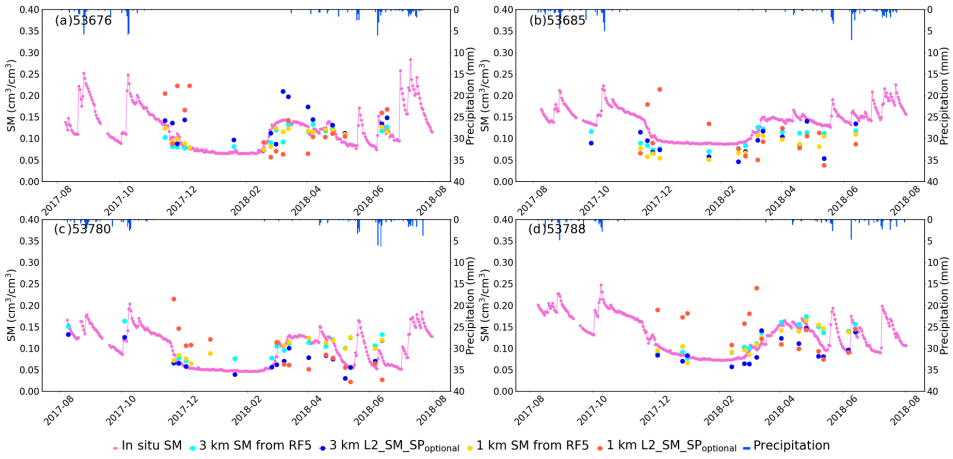

Figure 7. Time series of in situ SM, 3 km, and 1 km SM estimates calculated by RF5, 3 km, and 1 km $L2\_SM\_SP_{optional}$ data and daily precipitation at four stations.

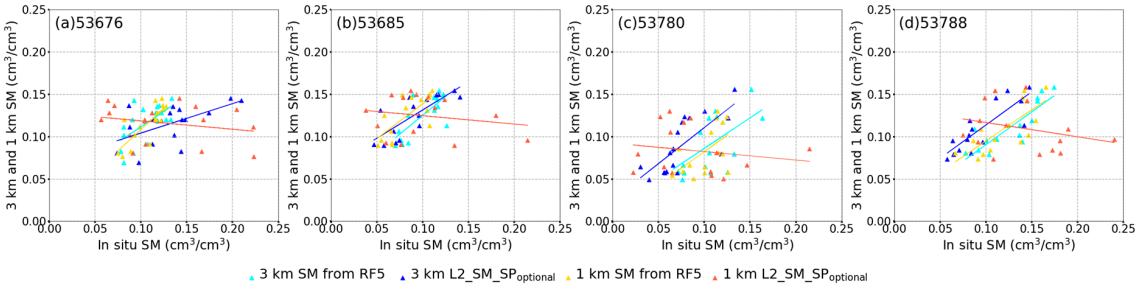

Figure 8. Scatterplots of SM estimates from RF5 and L2_SM_SP$_{optional}$ product (3 km and 1 km) vs. in situ SM at the four SM stations.

Table 5. Results comparison summary of the overlapping observations of L2_SM_SP$_{optional}$ and SM estimates based on RF5 at 3 km or 1 km vs. in situ SM. Avg is short for average.

ID	RF5 (3 km)		L2_SM_SP$_{optional}$ (3 km)		RF5 (1 km)		L2_SM_SP$_{optional}$ (1 km)	
	R	ubRMSE (cm³/cm³)	R	ubRMSE (cm³/cm³)	R	ubRMSE (cm³/cm³)	R	ubRMSE (cm³/cm³)
53676	0.62	0.018	0.57	0.030	0.80	0.013	−0.25	0.063
53685	0.75	0.016	0.79	0.017	0.73	0.016	−0.19	0.056
53780	0.65	0.025	0.74	0.022	0.55	0.024	−0.17	0.059
53788	0.86	0.016	0.83	0.017	0.76	0.019	−0.27	0.059
Avg	0.72	0.019	0.73	0.022	0.71	0.018	−0.22	0.059

Second, a spatial comparison between the SM based on RF5 and L2_SM_SP$_{optional}$ product was conducted. Only the comparison at the 1 km resolution is shown due to the similar performances at 3 km and 1 km. The 9 km L3_SM_P_E product is shown as a reference. Two regions under different conditions were selected, namely, regions A and B (Figure 9). Region A mainly has a relatively flat terrain covered with sand and grass as well as mountains covered with forest in the southeast. Region B is mountainous with interlaced dry land and forests. Additionally, according to the performance of RF1–RF5 in regions A and B, RF1 and RF2 also exhibited relatively significant differences from RF5. The SM values from RF3 and RF4 showed similar spatial distributions to those from RF1 and RF2, respectively.

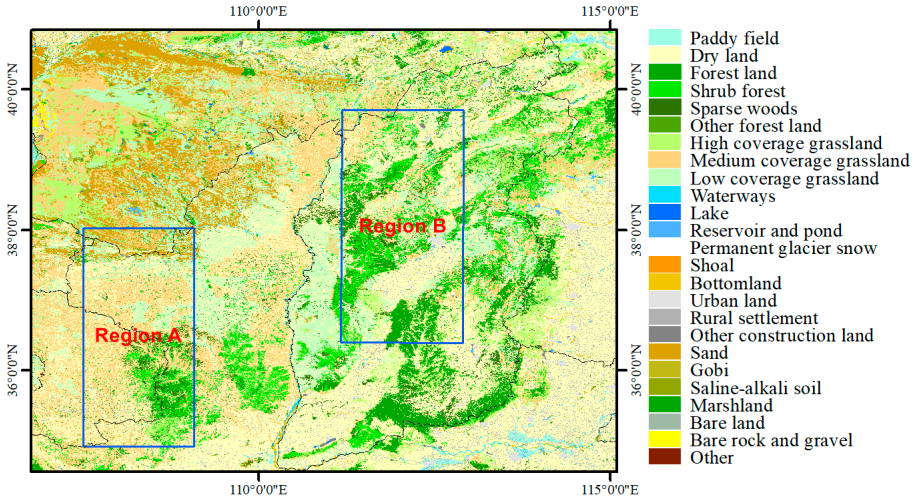

Figure 9. The locations of regions A and B in the study area.

The downscaled SM based on the RF models and L2_SM_SP$_{optional}$ product were compared at 1 km in region A (Figure 10). The SM from the RF models exhibited smoother SM spatial distribution

and fewer outliers than that from L2_SM_SP$_{optional}$. The flat terrain, low vegetation cover, and 9 km SMAP SM data of good quality may contribute to the good performance of the RF models in region A. The parameters that account for the effects of vegetation/roughness and heterogeneity within the coarse resolution are from a relationship between SMAP radiometer TB and Sentinel-1 SAR data [66]. The quality flag of 1 km disaggregated TB observations at vertical (V) polarization can offer some explanations for L2_SM_SP$_{optional}$ (Figure 11). Although the TB data at V polarization had acceptable quality, L2_SM_SP$_{optional}$ appears to be of poor quality due to noise related to radar measurements. Furthermore, compared with RF5, RF1 and RF2 without sufficient VIs clearly underestimated the SM values in areas covered with forest (Figure 10). Overall, RF5 exhibited a smoother and better SM spatial distribution than L2_SM_SP$_{optional}$ and other RF models at 1 km in region A.

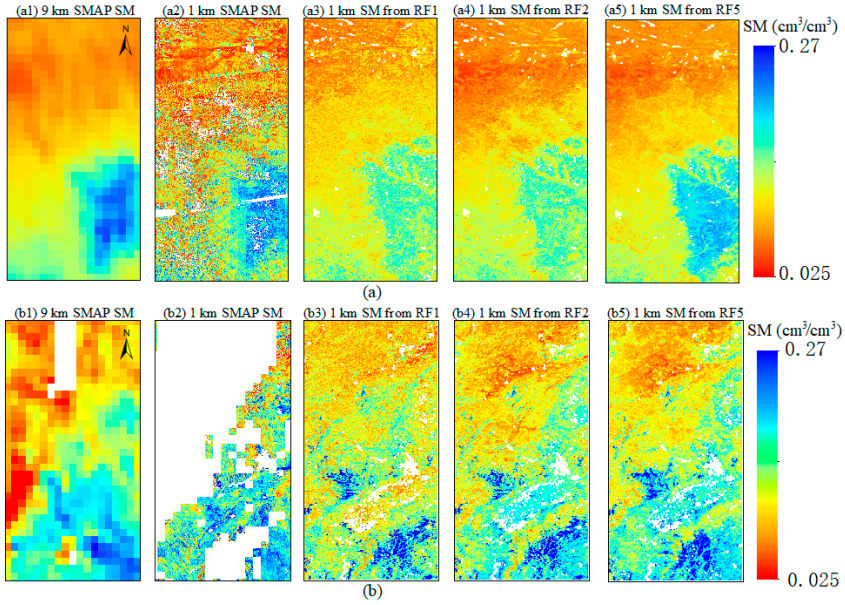

Figure 10. Spatial patterns of 9 km SMAP passive SM data (first column), 1 km L2_SM_SP$_{optional}$ data (second column), and 1 km downscaled SM data based on RF1 (third column), RF2 (fourth column) and RF5 (fifth column) in (**a**) region A (DOY = 2018, 71) and (**b**) region B (DOY = 2017, 342).

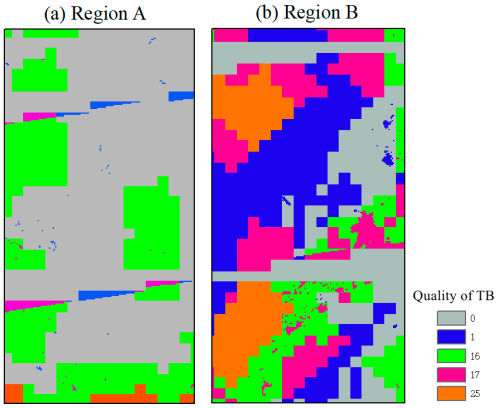

Figure 11. The quality flag of 1 km disaggregated brightness temperature (TB) data at V polarization for (**a**) region A (DOY = 2018, 71) and (**b**) region B (DOY = 2017, 342). (0: disaggregated TB data has acceptable quality. 1: unable to disaggregate TB data into cells; 16: significant levels of RFI were detected, and the TB data was repaired because of the effects of RFI; 17: significant levels of RFI were detected, and unable to disaggregate TB data into cells; 25: significant levels of RFI were detected, unable to disaggregate TB data into cells, and some V polarization TB input used for SM retrieval were questionable or of poor quality.).

The downscaled SM data based on RF models and the L2_SM_SP$_{optional}$ product were compared at 1 km in region B (Figure 10). Compared with the L2_SM_SP$_{optional}$ product, the RF models showed a smoother SM spatial distribution and featured fewer missing values. Large numbers of missing values in 1 km L2_SM_SP$_{optional}$ occurred in region B because the TB data could not be disaggregated into 1 km resolution cells, probably due to poor-quality SMAP TB observations heavily contaminated by RFI (the cells flagged with 17 and 25 in Figure 11b) and variations in the DSM (the cells flagged with 1 in Figure 11b). Furthermore, although the TB data was repaired because of the effect of RFI in the cells flagged with 16, many SM values were flagged as anomalously high, as shown in the retrieval quality flag of the L2_SM_SP$_{optional}$ product. Additionally, RF1 without σ^0_{vv} slightly underestimated SM values in dry land, which was consistent with the underestimation of RF1 in Figure 5a. Overall, RF5 exhibited a smoother and better SM spatial distribution than L2_SM_SP$_{optional}$ and other RF models at 1 km in region B.

Overall, RF5 with σ^0_{vv} and a sufficient number of VIs appeared to be superior to other RF models. RF5 was comparable with L2_SM_SP$_{optional}$ at 3 km and expressed more accurate SM values in the time series and smoother SM spatial distributions than L2_SM_SP$_{optional}$ data at 1 km. RFI, noise related to radar measurements, and variations in the DSM appear to have a greater impact on L2_SM_SP$_{optional}$ than on the RF models, leading to a large number of missing values.

4. Discussion

Different methods for surface SM mapping perform differently. The characteristics of the RF-based downscaling method with optical/infrared and Sentinel-1 SAR data and the discrepancies among the RF models and L2_SM_SP products should be discussed.

4.1. Analysis of Input Variables in the RF Models

The variable importance scores, the test at the 9 km resolution, and the evaluation of the 3 km and 1 km downscaled SM products indicate that the input variables in decreasing order of importance are σ^0_{vv}, LST, VIs, and topographic factors.

First, σ^0_{vv} plays a more significant role than the other variables because active microwaves can penetrate the soil surface to detect the surface SM content directly. The better performance of RF2 versus RF1 in the time series (Section 3.2) and the underestimation of RF1 without σ^0_{vv} in dry land (Figure 10b) indicate the high sensitivity of σ^0_{vv} to surface SM. Additionally, the low vegetation cover conditions and bare soil probably amplified the importance of the backscatter. σ^0_{vv} is very appropriate for northern China with vast areas of low vegetation or sparse vegetation. The RF model from Zhao et al. [53] that is dependent on the LST/VIs feature space was proposed for relatively high vegetation cover conditions, and the importance scores of VIs in their study were very high. However, the feature space performed relatively poorly in areas with NDVI values ranging from 0 to 0.30 [59,60], and the LST with low sensitivity to surface SM in winter would reduce the accuracy of the downscaling SM [53]. Thus, σ^0_{vv} with a high sensitivity to surface SM plays a more significant role than the other variables in this study. Amazirh et al. used σ^0_{vv} instead of SM proxies from optical/infrared data to express the SM variability for the disaggregation of LST in areas of low vegetation cover [58]. Overall, σ^0_{vv} may be able to serve as an input variable for downscaling coarse SM data.

Second, LST plays a crucial role because of the constraint on the combination of SAR data and VIs and the LST/VIs feature space. Although the combination of SAR data and VIs has been used for constructing non-linear models for SM retrievals, an error is introduced when ignoring vegetation type and surface roughness. Many combinations of characteristics (e.g., surface SM and surface roughness) contribute to the same backscattering coefficient, and LST seems to be an effective constraint on SM conditions, which is comparable with [56]. Additionally, LST is very sensitive to surface SM in this study area because the SM status is the main variable controlling the LST variability in areas of low vegetation cover [58]. Thus, LST is a pivotal feature in large regions with low vegetation cover.

Third, VIs play an important role because they can compensate for the influence of vegetation on the total backscattering coefficients and the LST/VIs feature space [56]. The suitability of different VIs for the explanation of vegetation information varies [57]. Among the VIs, both LAI and NDVI exhibit remarkable importance (Figure 3), which indicates that LAI has a similar vegetation cover representativeness as NDVI [47,53]. EVI and NDWI are less important, likely because these two variables are more sensitive in areas of high vegetation cover. EVI is usually used to compensate for the easily saturated NDVI, and the NDWI represents the vegetation water content. Thus, the RF models without appropriate VIs underestimate the SM in forest areas and are less stable than the RF5 model (Figure 10a).

Fourth, the effects of the elevation and slope were weak when compared with the effects of other input variables (Figure 3), possibly because the differences in elevation and slope were not large for most of this study area and the value ranges at the 9 km resolution narrowed because of the smoothing effect caused by spatial aggregation. The RF3 model performed slightly better than RF1 (Tables 2 and 3), which is consistent with the findings for variable importance.

In short, the input variables in decreasing order of importance are σ^0_{vv}, LST, VIs, and topographic factors. The backscatter at VV polarization and LST appear to be pivotal in areas with low vegetation cover. Appropriate VIs can enhance the adaptability of RF models to different vegetation cover levels. The RF5 model can obtain a more stable performance due to its inclusion of σ^0_{vv} and a sufficient number of VIs.

4.2. Analysis of the Differences between RF Models and L2_SM_SP Products

The discrepancies between the RF models and L2_SM_SP products, mainly including the noise associated with SAR data, uncertain auxiliary data, and dependence on original SMAP SM, should be discussed.

First, L2_SM_SP products are more affected by the noise associated with SAR data than the RF models, which is likely attributable to the complex surface scattering conditions on vegetation and surface roughness. The SM and TB data from the 9 km L3_SM_P_E product were disaggregated using the Sentinel-1 SAR backscatter with VV and VH polarizations to generate the $L2\_SM\_SP_{optional}$ and $L2\_SM\_SP_{baseline}$ data, respectively. Common parameters account for heterogeneity and the effects of vegetation/roughness in the baseline and optional approach, and are derived from a relationship between SMAP radiometer TB and Sentinel-1 SAR data [66]. The complex surface scattering conditions caused by vegetation and surface roughness can bring more uncertainty for the L2_SM_SP products, especially at a 1 km spatial resolution. Rainfall events, cultivation activities, and the inconsistent observation time between SMAP and Sentinel-1 satellites can further introduce additional outliers for the L2_SM_SP data [20,86]. The RF5 model is less affected by noise related to SAR data through a combination with other land surface features (e.g., VIs and LST), and is capable of generating 3 km and 1 km SM data with satisfactory performance.

Second, more uncertain auxiliary data are used for generating $L2\_SM\_SP_{baseline}$, which reduces the accuracy of $L2\_SM\_SP_{baseline}$. The auxiliary data (e.g., approximately 25 km soil surface temperature (SST) and coarse soil texture data) have resolutions coarser than 3 km [66]. China is a mountainous country with complex topography. Additionally, the four in situ stations are in a basin surrounded by mountains with relatively complex topography, and the land surface parameters may present complex heterogeneity. Thus, $L2\_SM\_SP_{baseline}$ may exhibit more uncertainty in temporal and spatial variations than $L2\_SM\_SP_{optional}$. In these situations, the SM data from the synergistic use of fine-resolution SAR and MODIS data and $L2\_SM\_SP_{optional}$ without coarse auxiliary data may be superior to $L2\_SM\_SP_{baseline}$.

Third, the performance of L2_SM_SP data is more dependent on the original SMAP SM than the RF models. The L2_SM_SP data are directly generated from the original SMAP SM/TB data, whereas the RF-based downscaled SM data are calculated using Sentinel-1 SAR data, MODIS data, and the trained RF models. The SMAP TB data with poor quality and noise associated with SAR data can lead to a large number of missing values and many outliers in L2_SM_SP products. RFI has a significant impact on SMAP TB observations because the L-band is severely contaminated by RFI

in China [87,88]. The RF-based downscaling method seems to be a suitable method, especially for areas easily contaminated by RFI. Additionally, the SMAP SM data with the closest observed time to Sentinel-1 SAR data were matched with SAR data; thus, the inconsistent observation time between SMAP and Sentinel-1 satellites has an impact on L2_SM_SP.

In general, compared with L2_SM_SP, the RF5 model may be less affected by noise related to SAR data and the accuracy of the original SMAP SM data, which is superior at both 3 km and 1 km resolutions. The $L2\_SM\_SP_{optional}$ data are better than the $L2\_SM\_SP_{baseline}$ calculated with low-resolution auxiliary data.

5. Conclusions

A method was presented to produce downscaled SMAP SM data by combining optical/infrared data with SAR data based on the RF technique. The method leverages the sensitivity of C-band radar backscatter to SM and the triangular/trapezoidal feature space of the VIs and LST.

In this study, five architectures (RF1–RF5) were first trained and tested at a 9 km resolution. SM from the 9 km SMAP L3_SM_P_E product was the target variable. Second, the model simulations from RF1–RF5 were compared using the in situ SM. RF5, with σ^0_{vv} and a sufficient number of VIs, performed best among the five RF models in the time series. Third, the two SMAP-Sentinel active-passive products at 3 km and 1 km were compared using in situ SM measurements. The $L2\_SM\_SP_{optional}$ product performed significantly better than the $L2\_SM\_SP_{baseline}$ product, possibly because coarse-resolution auxiliary data were used in $L2\_SM\_SP_{baseline}$. Fourth, the RF5 model and $L2\_SM\_SP_{optional}$ were further compared at 3 km and 1 km. The SM data from RF5 was comparable with $L2\_SM\_SP_{optional}$ at 3 km and showed more satisfactory accuracy at temporal and spatial scales than $L2\_SM\_SP_{optional}$ at 1 km.

In conclusion, the synergy of the Sentinel-1 backscatter signal and MODIS optical/infrared data can provide reliable temporal and spatial variability in semi-arid areas with relatively low vegetation cover. The input variables in decreasing order of importance were σ^0_{vv}, LST, VIs, and topographic factors over the entire study area. The low vegetation cover conditions probably increased the importance of the radar backscatter and LST. The presence of sufficient VIs can enhance the adaptability of RF models to different vegetation covers. The RF5 model, with σ^0_{vv} and sufficient VIs, could obtain stable performance at both 3 km and 1 km. Compared with L2_SM_SP, RF5 was more superior at the 1 km resolution.

Author Contributions: J.B. designed the research and analyzed the results with the assistance of Q.C.; J.B. performed the experiments with the help of W.Z.; and L.M. and Q.C. provided comments for the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (No. 2017YFC0405806), the National Natural Science Foundation of China (41601401), and the Land Water Resource Satellite System.

Acknowledgments: The authors wish to thank the anonymous reviewers for their detailed and constructive comments and also wish to express their sincere thanks to all data providers. The SMAP data were provided by the National Snow and Ice Data Center (NSIDC). Sentinel-1A data were provided by the European Space Agency (ESA)'s Copernicus Project. The in situ SM data were provided by the CMA and the Information Center of the Ministry of Water Resources.

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

References

1. Ochsner, T.E.; Cosh, M.H.; Cuenca, R.H.; Dorigo, W.A.; Draper, C.S.; Hagimoto, Y.; Kerr, Y.H.; Larson, K.M.; Njoku, E.G.; Small, E.E. State of the Art in Large-Scale Soil Moisture Monitoring. *Soil Sci. Soc. Am. J.* **2013**, *77*, 1888–1919. [CrossRef]
2. Sánchez, N.; Ángel, G.-Z.; Martínez-Fernández, J.; Piles, M.; Pablos, M. Integrated remote sensing approach to global agricultural drought monitoring. *Agric. For. Meteorol.* **2018**, *259*, 141–153. [CrossRef]

3. Martinez-Fernandez, J.; Gonzalez-Zamora, A.; Sanchez, N.; Gumuzzio, A.; Herrero-Jimenez, C.M. Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index. *Remote Sens. Environ.* **2016**, *177*, 277–286. [CrossRef]
4. Kim, S.; Paik, K.; Johnson, F.M.; Sharma, A. Building a Flood-Warning Framework for Ungauged Locations Using Low Resolution, Open-Access Remotely Sensed Surface Soil Moisture, Precipitation, Soil, and Topographic Information. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 375–387. [CrossRef]
5. Tekeli, A.E.; Fouli, H. Reducing False Flood Warnings of TRMM Rain Rates Thresholds over Riyadh City, Saudi Arabia by Utilizing AMSR-E Soil Moisture Information. *Water Resour. Manag.* **2017**, *31*, 1243–1256. [CrossRef]
6. Du, J.; Kimball, J.S.; Galantowicz, J.; Kim, S.-B.; Chan, S.K.; Reichle, R.; Jones, L.A.; Watts, J.D. Assessing global surface water inundation dynamics using combined satellite information from SMAP, AMSR2 and Landsat. *Remote Sens. Environ.* **2018**, *213*, 1–17. [CrossRef]
7. Purdy, A.J.; Fisher, J.B.; Goulden, M.L.; Colliander, A.; Halverson, G.; Tu, K.; Famiglietti, J.S. SMAP soil moisture improves global evapotranspiration. *Remote Sens. Environ.* **2018**, *219*, 1–14. [CrossRef]
8. Walker, E.; García, G.A.; Venturini, V. Evapotranspiration estimation using SMAP soil moisture products and bouchet complementary evapotranspiration over Southern Great Plains. *J. Arid Environ.* **2019**, *163*, 34–40. [CrossRef]
9. Martens, B.; Miralles, D.; Lievens, H.; Fernández-Prieto, D.; Verhoest, N.E.C. Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *48*, 146–162. [CrossRef]
10. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.J.; Font, J.; Reul, N.; Gruhier, C. The SMOS Mission: New Tool for Monitoring Key Elements of the Global Water Cycle. *Proc. IEEE* **2010**, *98*, 666–687. [CrossRef]
11. Koster, R.D.; Crow, W.T.; Reichle, R.H.; Mahanama, S.P. Estimating Basin-Scale Water Budgets with SMAP Soil Moisture Data. *Water Resour. Res.* **2018**, *54*, 4228. [CrossRef] [PubMed]
12. Zawadzki, J.; Kędzior, M. Soil moisture variability over Odra watershed: Comparison between SMOS and GLDAS data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 110–124. [CrossRef]
13. Kędzior, M.; Zawadzki, J. Comparative study of soil moisture estimations from SMOS satellite mission, GLDAS database, and cosmic-ray neutrons measurements at COSMOS station in Eastern Poland. *Geoderma* **2016**, *283*, 21–31. [CrossRef]
14. Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J. The soil moisture active passive (SMAP) mission. *Proc. IEEE* **2010**, *98*, 704–716. [CrossRef]
15. Piles, M.; Sánchez, N.; Vall-llossera, M.; Camps, A.; Martínez-Fernández, J.; Martínez, J.; González-Gambau, V. A downscaling approach for SMOS land observations: Evaluation of high-resolution soil moisture maps over the Iberian Peninsula. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3845–3857. [CrossRef]
16. Ma, H.; Zeng, J.; Chen, N.; Zhang, X.; Cosh, M.H.; Wang, W. Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations. *Remote Sens. Environ.* **2019**, *231*, 111215. [CrossRef]
17. Peng, J.; Loew, A.; Merlin, O.; Verhoest, N.E. A review of spatial downscaling of satellite remotely sensed soil moisture. *Rev. Geophys.* **2017**, *55*, 341–366. [CrossRef]
18. Sabaghy, S.; Walker, J.P.; Renzullo, L.J.; Jackson, T.J. Spatially enhanced passive microwave derived soil moisture: Capabilities and opportunities. *Remote Sens. Environ.* **2018**, *209*, 551–580. [CrossRef]
19. Bindlish, R.; Barros, A.P. Subpixel variability of remotely sensed soil moisture: An inter-comparison study of SAR and ESTAR. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 326–337. [CrossRef]
20. Li, J.; Wang, S.; Gunn, G.; Joosse, P.; Russell, H.A.J. A model for downscaling SMOS soil moisture using Sentinel-1 SAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 109–121. [CrossRef]
21. Wu, X.; Walker, J.; Rudiger, C.; Panciera, R.; Ying, G. Medium-Resolution Soil Moisture Retrieval Using the Bayesian Merging Method. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6482–6493. [CrossRef]
22. Lievens, H.; Reichle, R.H.; Liu, Q.; Lannoy, G.J.M.D.; Wagner, W. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates: Sentinel-1 and SMAP soil moisture. *Geophys. Res. Lett.* **2017**, *44*, 6145–6153. [CrossRef] [PubMed]
23. Han, J.; Mao, K.; Xu, T.; Guo, J.; Zuo, Z.; Gao, C. A soil moisture estimation framework based on the CART algorithm and its application in China. *J. Hydrol.* **2018**, *563*, 65–75. [CrossRef]

24. Akbar, R.; Moghaddam, M. A Combined Active–Passive Soil Moisture Estimation Algorithm with Adaptive Regularization in Support of SMAP. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3312–3324. [CrossRef]
25. Das, N.N.; Entekhabi, D.; Dunbar, S.; Kim, S.; Yueh, S.; Colliander, A.; Jackson, T.J.; O'Neill, P.E.; Cosh, M.; Caldwell, T.; et al. *Assessment Report for the L2_SM_SP Beta Release Data Products*; Jet Propulsion Laboratory: Pasadena, CA, USA, 2017.
26. He, L.; Hong, Y.; Wu, X.; Ye, N.; Walker, J.P.; Chen, X. Investigation of SMAP Active–Passive Downscaling Algorithms Using Combined Sentinel-1 SAR and SMAP Radiometer Data. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4906–4918. [CrossRef]
27. Jagdhuber, T.; Baur, M.; Akbar, R.; Das, N.N.; Link, M.; He, L.; Entekhabi, D. Estimation of active-passive microwave covariation using SMAP and Sentinel-1 data. *Remote Sens. Environ.* **2019**, *225*, 458–468. [CrossRef]
28. Montzka, C.; Jagdhuber, T.; Horn, R.; Bogena, H.R.; Hajnsek, I.; Reigber, A.; Vereecken, H. Investigation of SMAP fusion algorithms with airborne active and passive L-band microwave remote sensing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3878–3889. [CrossRef]
29. Hosseini, M.; McNairn, H. Using multi-polarization C- and L-band synthetic aperture radar to estimate biomass and soil moisture of wheat fields. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 50–64. [CrossRef]
30. Das, N.N.; Entekhabi, D.; Njoku, E.G.; Shi, J.J.C.; Johnson, J.T.; Colliander, A. Tests of the SMAP Combined Radar and Radiometer Algorithm Using Airborne Field Campaign Observations and Simulated Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2018–2028. [CrossRef]
31. Santi, E.; Paloscia, S.; Pettinato, S.; Brocca, L.; Ciabatta, L.; Entekhabi, D. Integration of microwave data from SMAP and AMSR2 for soil moisture monitoring in Italy. *Remote Sens. Environ.* **2018**, *212*, 21–30. [CrossRef]
32. Santi, E.; Paloscia, S.; Pettinato, S.; Brocca, L.; Ciabatta, L.; Entekhabi, D. On the synergy of SMAP, AMSR2 AND SENTINEL-1 for retrieving soil moisture. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *65*, 114–123. [CrossRef]
33. Santi, E. An application of the SFIM technique to enhance the spatial resolution of spaceborne microwave radiometers. *Int. J. Remote Sens.* **2010**, *31*, 2419–2428. [CrossRef]
34. Millard, K.; Richardson, M. Quantifying the relative contributions of vegetation and soil moisture conditions to polarimetric C-Band SAR response in a temperate peatland. *Remote Sens. Environ.* **2018**, *206*, 123–138. [CrossRef]
35. Piles, M.; McColl, K.A.; Entekhabi, D.; Das, N.; Pablos, M. Sensitivity of Aquarius Active and Passive Measurements Temporal Covariability to Land Surface Characteristics. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4700–4711. [CrossRef]
36. Piles, M.; Camps, A.; Vall-Llossera, M.; Corbella, I.; Panciera, R.; Rudiger, C.; Kerr, Y.H.; Walker, J. Downscaling SMOS-derived soil moisture using MODIS visible/infrared data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3156–3166. [CrossRef]
37. Piles, M.; Petropoulos, G.P.; Sánchez, N.; González-Zamora, Á.; Ireland, G. Towards improved spatio-temporal resolution soil moisture retrievals from the synergy of SMOS and MSG SEVIRI spaceborne observations. *Remote Sens. Environ.* **2016**, *180*, 403–417. [CrossRef]
38. Alemohammad, S.H.; Kolassa, J.; Prigent, C.; Aires, F.; Gentine, P. Global Downscaling of Remotely-Sensed Soil Moisture using Neural Networks. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, *22*, 5341–5356. [CrossRef]
39. Im, J.; Park, S.; Rhee, J.; Baik, J.; Choi, M. Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environ. Earth Sci.* **2016**, *75*, 1120. [CrossRef]
40. Peng, J.; Niesel, J.; Loew, A.; Zhang, S.; Wang, J. Evaluation of Satellite and Reanalysis Soil Moisture Products over Southwest China Using Ground-Based Measurements. *Remote Sens.* **2015**, *7*, 15729–15747. [CrossRef]
41. Peng, J.; Loew, A.; Zhang, S.; Wang, J.; Niesel, J. Spatial downscaling of satellite soil moisture data using a vegetation temperature condition index. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 558–566. [CrossRef]
42. Kim, J.; Hogue, T.S. Improving spatial soil moisture representation through integration of AMSR-E and MODIS products. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 446–460. [CrossRef]
43. Merlin, O.; Chehbouni, A.; Walker, J.P.; Panciera, R.; Kerr, Y.H. A simple method to disaggregate passive microwave-based soil moisture. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 786–796. [CrossRef]
44. Merlin, O.; Rudiger, C.; Bitar, A.A.; Richaume, P.; Kerr, Y.H. Disaggregation of SMOS Soil Moisture in Southeastern Australia. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1556–1571. [CrossRef]
45. Malbéteau, Y.; Merlin, O.; Molero, B.; Rüdiger, C.; Bacon, S. DisPATCh as a tool to evaluate coarse-scale remotely sensed soil moisture using localized in situ measurements: Application to SMOS and AMSR-E data in Southeastern Australia. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 221–234. [CrossRef]

46. Senanayake, I.P.; Yeo, I.Y.; Tangdamrongsub, N.; Willgoose, G.R.; Hancock, G.R.; Wells, T.; Fang, B.; Lakshmi, V.; Walker, J.P. An in-situ data based model to downscale radiometric satellite soil moisture products in the Upper Hunter Region of NSW, Australia. *J. Hydrol.* **2019**, *572*, 820–838. [CrossRef]
47. Peng, J.; Niesel, J.; Loew, A. Evaluation of soil moisture downscaling using a simple thermal based proxy the REMEDHUS network (Spain) example. *Hydrol. Earth Syst. Sci.* **2015**, *9*, 8505–8551. [CrossRef]
48. Carlson, T.N.; Gillies, R.R.; Perry, E.M. A Method to Make Use of Thermal Infrared Temperature and NDVI measurements to Infer Surface Soil Water Content and Fractional Vegetation Cover. *Remote Sens. Rev.* **1994**, *9*, 161–173. [CrossRef]
49. Carlson, T. An Overview of the "Triangle Method" for Estimating Surface Evapotranspiration and Soil Moisture from Satellite Imagery. *Sensors* **2007**, *7*, 1612–1629. [CrossRef]
50. Mu, Q.; Heinsch, F.A.; Zhao, M.; Running, S.W. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.* **2007**, *111*, 519–536. [CrossRef]
51. Zhao, W.; Li, A. A comparison study on empirical microwave soil moisture downscaling methods based on the integration of microwave-optical/IR data on the Tibetan Plateau. *Int. J. Remote Sens.* **2015**, *36*, 4986–5002. [CrossRef]
52. Zhao, W.; Li, A.; Jin, H.; Zhang, Z.; Bian, J.; Yin, G. Performance evaluation of the triangle-based empirical soil moisture relationship models based on landsat-5 TM data and in situ measurements. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2632–2645. [CrossRef]
53. Zhao, W.; Sanchez, N.; Lu, H.; Li, A. A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression. *J. Hydrol.* **2018**, *563*, 1009–1024. [CrossRef]
54. 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]
55. Holtgrave, A.-K.; Förster, M.; Greifeneder, F.; Notarnicola, C.; Kleinschmit, B. Estimation of Soil Moisture in Vegetation-Covered Floodplains with Sentinel-1 SAR Data Using Support Vector Regression. *PFG J. Photogramm. Remote Sens. Geoinf. Sci.* **2018**, *86*, 85–101. [CrossRef]
56. El Hajj, M.; Baghdadi, N.; Zribi, M.; Bazzi, H. Synergic Use of Sentinel-1 and Sentinel-2 Images for Operational Soil Moisture Mapping at High Spatial Resolution over Agricultural Areas. *Remote Sens.* **2017**, *9*, 1292. [CrossRef]
57. Qiu, J.; Crow, W.T.; Wagner, W.; Zhao, T. Effect of vegetation index choice on soil moisture retrievals via the synergistic use of synthetic aperture radar and optical remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *80*, 47–57. [CrossRef]
58. Amazirh, A.; Merlin, O.; Erraki, S. Including Sentinel-1 radar data to improve the disaggregation of MODIS land surface temperature data. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 11–26. [CrossRef]
59. Xiang, D.; Liu, L.; Wang, Q.; Yang, N.; Han, T. Evaluation of data quality and drought monitoring capability of FY-3A MERSI data. *Adv. Artif. Intell.* **2010**, *2010*, 1–10. [CrossRef]
60. Chen, L.; Yan, G.; Ren, H.; Li, A. A modified vegetation index based algorithm for thermal imagery sharpening. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010; pp. 2444–2447.
61. Piles, M.; Entekhabi, D.; Camps, A. A change detection algorithm for retrieving high-resolution soil moisture from SMAP radar and radiometer observations. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 4125–4131. [CrossRef]
62. Das, N.N.; Entekhabi, D.; Njoku, E.G. An algorithm for merging SMAP radiometer and radar data for high-resolution soil-moisture retrieval. *IEEE Trans. Geosci. Remote Sens.* **2010**, *49*, 1504–1512. [CrossRef]
63. Sun, Y.; Huang, S.; Ma, J.; Li, J.; Li, X.; Wang, H.; Chen, S.; Zang, W. Preliminary Evaluation of the SMAP Radiometer Soil Moisture Product over China Using In Situ Data. *Remote Sens.* **2017**, *9*, 292. [CrossRef]
64. Bai, J.; Cui, Q.; Chen, D.; Yu, H.; Mao, X.; Meng, L.; Cai, Y. Assessment of the SMAP-Derived Soil Water Deficit Index (SWDI-SMAP) as an Agricultural Drought Index in China. *Remote Sens.* **2018**, *10*, 1302. [CrossRef]
65. Chan, S.K.; Bindlish, R.; O'Neill, P.; Jackson, T.; Njoku, E.; Dunbar, S.; Chaubell, J.; Piepmeier, J.; Yueh, S.; Entekhabi, D.; et al. Development and assessment of the SMAP enhanced passive soil moisture product. *Remote Sens. Environ.* **2018**, *204*, 931–941. [CrossRef]

66. Das, N.; Entekhabi, D.; Dunbar, R.S.; Kim, S.; Yueh, S.; Colliander, A. *SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Soil Moisture, Version 2*; NASA National Snow and Ice Data Center Distributed Active Archive Center: Boulder, CO, USA, 2018.
67. NASA's Earth Observing System Data and Information System (EOSDIS). Available online: https://search.earthdata.nasa.gov/search (accessed on 19 November 2019).
68. European Space Agency's Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu (accessed on 19 November 2019).
69. Lee, J.S.; Grunes, M.R.; Grandi, G.D. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans. Geosci. Remote Sens.* **2002**, *37*, 2363–2373.
70. Choi, M.; Hur, Y. A microwave-optical/infrared disaggregation for improving spatial representation of soil moisture using AMSR-E and MODIS products. *Remote Sens. Environ.* **2012**, *124*, 259–269. [CrossRef]
71. Liang, S. Narrowband to broadband conversions of land surface albedo I algorithms. *Remote Sens. Environ.* **2001**, *76*, 213–238. [CrossRef]
72. Lucht, W.; Roujean, J.L. Considerations in the parametric modeling of BRDF and albedo from multiangular satellite sensor observations. *Remote Sens. Rev.* **2000**, *18*, 343–379. [CrossRef]
73. Gao, B.-C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]
74. Pablos, M.; Piles, M.; Sánchez, N.; Vall-llossera, M.; Martínez-Fernández, J.; Camps, A. Impact of day/night time land surface temperature in soil moisture disaggregation algorithms. *Eur. J. Remote Sens.* **2016**, *49*, 899–916. [CrossRef]
75. United States Geological Survey (USGS) Earth Explorer. Available online: https://earthexplorer.usgs.gov (accessed on 19 November 2019).
76. Ran, Y.; Li, X.; Lu, L. Evaluation of four remote sensing based land cover products over China. *Int. J. Remote Sens.* **2010**, *31*, 391–401. [CrossRef]
77. Ran, Y.; Li, X.; Lu, L. *Land Cover Products of China*; Cold and Arid Regions Science Data Center at Lanzhou: Lanzhou, China, 2010; Available online: http://westdc.westgis.ac.cn (accessed on 19 November 2019).
78. NASA's Precipitation Measurement Missions. Available online: https://pmm.nasa.gov (accessed on 19 November 2019).
79. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
80. Li, C.; Lu, H.; Yang, K.; Han, M.; Wright, J.; Chen, Y.; Yu, L.; Xu, S.; Huang, X.; Gong, W. The Evaluation of SMAP Enhanced Soil Moisture Products Using High-Resolution Model Simulations and In-Situ Observations on the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 535. [CrossRef]
81. Colliander, A.; Jackson, T.J.; Chan, S.K.; O'Neill, P.; Bindlish, R.; Cosh, M.H.; Caldwell, T.; Walker, J.P.; Berg, A.; McNairn, H. An assessment of the differences between spatial resolution and grid size for the SMAP enhanced soil moisture product over homogeneous sites. *Remote Sens. Environ.* **2018**, *207*, 65–70. [CrossRef]
82. Zhang, R.; Kim, S.; Sharma, A. A comprehensive validation of the SMAP Enhanced Level-3 Soil Moisture product using ground measurements over varied climates and landscapes. *Remote Sens. Environ.* **2019**, *223*, 82–94. [CrossRef]
83. Owe, M.; Griend, A.A.V. Comparison of soil moisture penetration depths for several bare soils at two microwave frequencies and implications for remote sensing. *Water Resour. Res.* **1998**, *34*, 2319–2327. [CrossRef]
84. Wei, Z.; Meng, Y.; Zhang, W.; Peng, J.; Meng, L. Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan Plateau. *Remote Sens. Environ.* **2019**, *225*, 30–44. [CrossRef]
85. Baghdadi, N.; El Hajj, M.; Zribi, M.; Bousbih, S. Calibration of the water cloud model at C-band for winter crop fields and grasslands. *Remote Sens.* **2017**, *9*, 969. [CrossRef]
86. Zhu, L.; Walker, J.P.; Tsang, L.; Huang, H.; Ye, N.; Rüdiger, C. Soil moisture retrieval from time series multi-angular radar data using a dry down constraint. *Remote Sens. Environ.* **2019**, *231*, 111237. [CrossRef]
87. Wang, X.X.; Wang, X.; Han, Z.; Yang, J.H. Radio frequency interference detection and characteristic analysis based on the L band stokes parameters remote sensing data. *Chemosphere* **2015**, *70*, 1699–1706.
88. Yan, S.; Khazaal, A.; Cabot, F.; Kerr, Y.H. An RFI Index to Quantify the Contamination of SMOS Data by Radio-Frequency Interference. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1577–1589.