

Article

Field-Scale Soil Moisture Retrieval Using PALSAR-2 Polarimetric Decomposition and Machine Learning

Xiaodong Huang ^{1,*}, Beth Ziniti ¹, Michael H. Cosh ², Michele Reba ³, Jinfei Wang ⁴ and Nathan Torbick ¹

¹ Applied Geosolutions, Durham, NH 03824, USA; bziniti@appliedgeosolutions.com (B.Z.); ntorbick@appliedgeosolutions.com (N.T.)

² Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705, USA; Michael.Cosh@usda.gov

³ Agricultural Research Service, United States Department of Agriculture, Jonesboro, AR 72401, USA; michele.reba@usda.gov

⁴ Department of Geography, Western University, London, ON N6A 3K7, Canada; jfwang@uwo.ca

* Correspondence: xhuang@appliedgeosolutions.com

Abstract: Soil moisture is a key indicator to assess cropland drought and irrigation status as well as forecast production. Compared with the optical data which are obscured by the crop canopy cover, the Synthetic Aperture Radar (SAR) is an efficient tool to detect the surface soil moisture under the vegetation cover due to its strong penetration capability. This paper studies the soil moisture retrieval using the L-band polarimetric Phased Array-type L-band SAR 2 (PALSAR-2) data acquired over the study region in Arkansas in the United States. Both two-component model-based decomposition (SAR data alone) and machine learning (SAR + optical indices) methods are tested and compared in this paper. Validation using independent ground measurement shows that the both methods achieved a Root Mean Square Error (RMSE) of less than 10 (vol.%), while the machine learning methods outperform the model-based decomposition, achieving an RMSE of 7.70 (vol.%) and R^2 of 0.60.

Keywords: soil moisture; PALSAR-2; model-based decomposition; machine learning



Citation: Huang, X.; Ziniti, B.; Cosh, M.H.; Reba, M.; Wang, J.; Torbick, N. Field-Scale Soil Moisture Retrieval Using PALSAR-2 Polarimetric Decomposition and Machine Learning. *Agronomy* **2021**, *11*, 35. <https://dx.doi.org/10.3390/agronomy11010035>

Received: 10 November 2020

Accepted: 24 December 2020

Published: 26 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soil moisture is identified as an essential climatic variable by the Global Observing System for Climate given its role in energy flux and climate–land feedback [1,2]. Within the agricultural community, the Group on Earth Observations (GEO) Global Agricultural Monitoring (GEOGLAM) initiative has further amplified soil moisture as an Essential Agricultural Variable (EAV) given its role in driving processes such as drought, irrigation, and crop production. Currently, programs such as NASA's Soil Moisture Active Passive (SMAP) and ESA's Soil Moisture Ocean Salinity (SMOS) missions have advanced science for land surface monitoring and agricultural decision support tools. These missions have made tremendous advances in synoptic monitoring and assessment of soil moisture.

Current challenges center on automating field scale soil moisture monitoring and assessment with robust and more physical or scalable approaches [3–6]. The spatial resolution of standard SMAP and SMOS products are at the spatial scale of tens of kilometers; thus, they are too coarse for agricultural decision making at the field scale. Typically, field scale is the spatial unit of decision making for operational management such as irrigation, crop rotations, tillage, and nutrient applications. Approaches to improve the spatial resolution of SMAP/SMOS scale products are to use downscaling routines that combine the spatially coarse resolution SMAP/SMOS information with fine (4 m) or moderate spatial resolution (30–100m) covariates such as Synthetic Aperture Radar (SAR) imagery or with process-based land surface models [7,8]. However, the typical product output resolution of these downscaling approaches is 3 km or 1 km, which is still too coarse [3,9,10] for farm operations. Further, a theme in these downscaling studies is that as the spatial resolution increases, the accuracy decreases [3].

Many research studies have investigated the potential to estimate field-scale soil moisture using moderate spatial resolution (<30 m) SAR imagery and these applications are historically empirical approaches often tied to a specific study region or temporal period [11–13]. Scaling these approaches towards higher Application Readiness Levels (ARLs) for wider deployment and fusion techniques, blending moderate resolution SAR and optical observations are critical next steps for the community. Radiation Transfer (RT) modeling is extensively employed [14–16] for physically retrieving soil moisture from SAR signals and is an option for driving decision support tools. However, rigorous first order RT modeling, such as Michigan Microwave Canopy Scattering (MIMICS) [17], requires extensive parameterization of phenological and surface attributes for thorough descriptions of surface structure, which limits operational execution and transferability. As a result, simplified RT models have been developed that generalize the vegetative structural components or ignore some scattering components such as double-bounce and multiple scatterings. One of the extensively used and less complex RT models has been the Water Cloud Model (WCM) [14], which simply treats the vegetation canopy as a water cloud layer given that the dielectric constant of water in the target vegetation is much higher than that of the dry vegetative matter. Numerous research applications have estimated soil moisture across a range of conditions using the WCM [18–24]. Generally, retrieval accuracy is largely influenced by the choice of the surface scattering model and vegetation descriptors [25]. Therefore, this approach tends to be scene- or image-based and the transferability of this approach or enhancing the ARL of field-scale retrieval has remained a challenge.

Without the assistance of the external optical data, quad polarized SAR data have shown potential for retrieving soil moisture using model-based decomposition. The model-based decomposition was first proposed by Freeman and Durden [26] and its application for soil moisture estimation has evolved. Hajnsek et al. [27] compared different model-based decompositions and found the results often underestimate soil moisture levels due to the inadequate estimation of the vegetation layer within agricultural fields, and a mean Root Mean Square Error (RMSE) of 10 (vol.%) was achieved for all methods. Jagdhuber et al. [28] investigated the multi-angular polarimetric decomposition to estimate soil moisture with high inversion rate and low RMSE for fully polarimetric L-band SAR data. Ballester-Berman et al. [29] presented a two-component polarimetric decomposition model for sparse vineyards using C-Band RADARSAT-2 data, but no measured ground truth data were used for validation. A hybrid decomposition method combining both the model-based and eigen-based decompositions was presented by Jagdhuber et al. [30] and showed a very high inversion rate for L-band data. However, the surface scattering models (X-Bragg model) referenced above are only applicable to characterize rough surfaces with small slopes. Therefore, Huang et al. [4] proposed a two-component model-based decomposition method to retrieve soil moisture in which the calibrated Integral Equation Model (IEM) is employed to describe more variable and rough surfaces. Recently, a two-component model-based decomposition was also proposed by Wang et al. [31] using the Oh surface scattering model to estimate the soil moisture, achieving a high accuracy with an RMSE of less than 7.5 (vol.%).

In addition to the evolution of physical RT models and polarimetric model-based decomposition, machine learning (ML) methods have gained in popularity in recent years [23,32]. For example, Paloscia et al. [32] applied an artificial neural network using C-band Sentinel-1 observations to estimate soil moisture and achieve an RMSE of less than 7 (vol.%). Similarly, El Hajj et al. [23] inverted soil moisture using X-band TerraSAR-X and ML techniques to achieve a relatively high accuracy with an RMSE of 6 (vol.%). In this research application, we build on this lineage to advance the retrieval of field-scale soil moisture using L-band polarimetric Phase Array type L-band Synthetic Aperture Radar-2 (PALSAR-2) and compare an improved two-component model-based decomposition (SAR alone) to a machine learning approach (SAR and optical indices).

2. Materials and Methods

2.1. Data Preparation and Processing

2.1.1. Study Area

The study area is located outside Osceola, Arkansas, USA (Figure 1). Major crops in this region include rice, corn, cotton, soybean, winter wheat, sorghum, and peanuts. Ten (10) soil moisture stations were installed for the entire crop season (field preparation through harvest) at select locations with local producers within agricultural fields. The Stevens Water Hydro probes were installed with soil moisture and temperature measurements obtained at 5 and 10 cm depths. This provides a quantification of volumetric soil moisture at integrated depths of 3–7 cm and 8–12 cm. Local micrometeorological stations and regional gridded weather data provided ancillary measurements to evaluate algorithm performance and dynamic responses. Here, we used the Parameter-elevation Regressions on Independent Slopes Model (PRISM) precipitation data products to evaluate space–time responses of derived metrics. From SoilGrids250m (<https://www.soilgrids.org/>), most of the soil class is Luvisols with high clay content, and has average clay, sand, and silt contents of 300 g/kg, 350 g/kg, and 350 g/kg over this region.

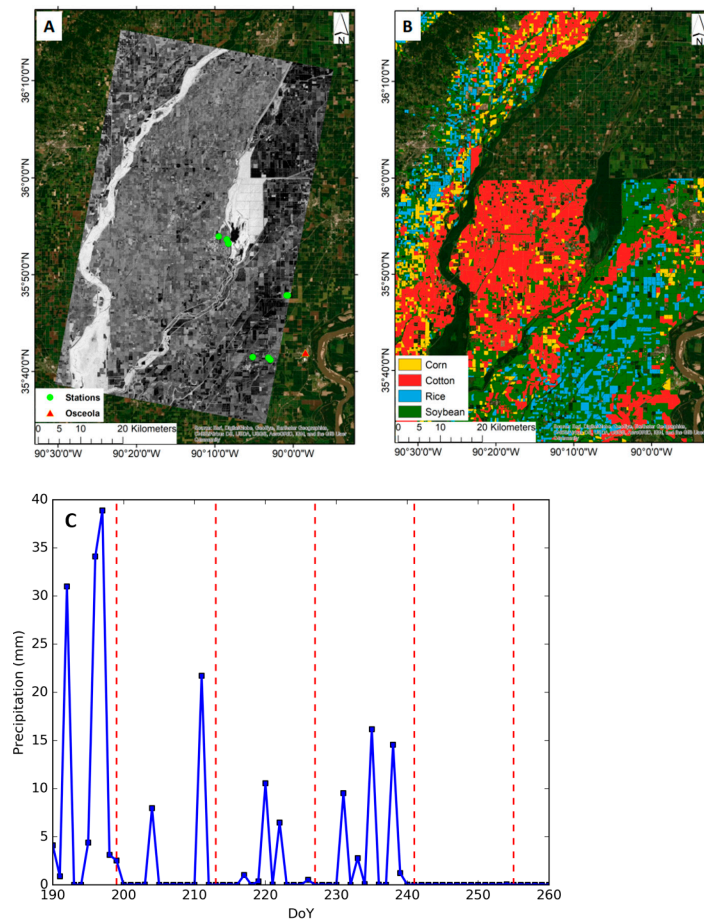


Figure 1. Illustration of (A) moderate spatial resolution (30 m) L-band volume scattering on 18 July 2019 with soil moisture probe locations shown as green dots. (B) USDA Cropland Data Layer (CDL) shows major crops in this agricultural production hot spot with dense patches of major row crops. (C) Parameter-elevation Regressions on Independent Slopes Model (PRISM) precipitation (blue line = study domain mean) shows diverse conditions and provided assistance in evaluating algorithm performance in relation to derived space–time patterns (PALSAR-2 time series satellite overpass dates = red lines).

2.1.2. SAR and Optical Data Processing

This study used time series fine beam quad polarization PALSAR-2 data acquired on 20190718, 20190801, 20190815, 20190829, and 20190912, and only those five quad polarization datasets acquired by JAXA in 2019 are used in this study. Our team had the study area designated as a research “super site” and time series, fine beam, quad polarization observations are historically very challenging tasks. The nominal incidence angle centers on 30 degrees over the study footprint with only slight, negligible variation in incident angle in the range direction. The polarimetric PALSAR-2 SAR data were first converted to a three by three coherency matrix, and a boxcar filter was applied to reduce the inherent speckle noise with a five by five window size. The filtered coherency matrix was then terrain geocoded to the Earth’s surface with a resultant 30 m spatial resolution using an external 90 m SRTM Digital Elevation Model (DEM). Generally, the landscape is extremely flat with relatively little slope relative to the scale of a pixel. The two-component model-based decomposition is applied to the geocoded coherency matrix to physically retrieve the soil moisture at moderate spatial resolution. Each period is processed independent from other time periods with no dependencies. In addition, the horizontal transmitting and vertical receiving (HH), vertical transmitting and horizontal receiving (HV), horizontal transmitting and vertical receiving (VH), and vertical transmitting and vertical receiving (VV) backscattering coefficients were generated to use as input parameters for the machine learning approach.

This study used Harmonized Landsat-8 Sentinel-2 (HLS) data, to generate optical time series surface reflectance, the Normalized Difference Vegetation Index (NDVI) [33], and cloud masks [34] to eliminate pixels potentially contaminated with atmospheric attenuation. Additional quality assurance quality control routines were applied to treat potentially poor-quality pixels (i.e., missed clouds) using a multi-temporal module [35] that allows for efficient custom pixel-wise time series calculations on image stacks. In this case, the multi-temporal module was implemented to remove additional observations (reflectance and NDVI) that fall between a decrease and increase larger than 0.01 over an 8-day period. However, periods of extended missed clouds (two or more) might remain. Therefore, weekly composite NDVI and reflectance products were then derived by linearly interpolating the remaining non-flagged values and applying a smoother. The smoother selected in this case was a least square penalized smoothing spline similar to the Whittaker smoother [36].

2.2. Methods

2.2.1. Two-Component Model-Based Decomposition

The framework of the two-component model-based decomposition developed by Huang et al. [4] is written as

$$T_{33} = f_s \times T_s + f_v \times T_v \quad (1)$$

in which T_{33} is the three by three coherency matrix, and f_s and f_v represent the scattering intensity of the surface and volume scattering components, while T_s and T_v are the coherency matrices of the surface and volume scattering. The volume scattering model used here leverages Huang et al. [4], which is written as

$$T_v(n) = \begin{bmatrix} T_{v11} & T_{v12} & 0 \\ T_{v12} & T_{v22} & 0 \\ 0 & 0 & T_{v33} \end{bmatrix} \quad (2)$$

where

$$T_{v11} = 1, T_{v12} = \frac{n}{2} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2} + 2\right)}, T_{v22} = \frac{(n^2 + 2n + 4)}{4} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2} + 3\right)}, T_{v33} = 2 \cdot \frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n}{2} + 3\right)} \cdot \frac{\Gamma\left(\frac{n}{2} + 1\right)}{\Gamma\left(\frac{n+1}{2}\right)}$$

The n is the randomness factor characterizing the probability distribution of the scatterer ranging from 0 to infinity. When $n = 0$, it represents the uniform distribution which is the Freeman volume scattering [26]. When $n = 1$, it is equal to the Yamaguchi volume scattering model [37]. Therefore, the improved volume scattering model would be a generalized model to characterize the scattering from the canopy. To determine the contribution of the volume scattering, the Non-Negative Eigenvalue Decomposition (NNED) is used [38] and the randomness factor n will be determined by minimizing the Radar Vegetation Index (RVI) derived from the PALSAR-2 data and the RVI [39] from the improved volume scattering model in Equation (2). Removing the contribution of the volume scattering, the remaining coherency matrix dominated by the surface scattering will be converted to the covariance matrix, in which the HH and VV backscattering coefficient will be extracted and applied to the calibrated IEM for L-band [40] to estimate the soil moisture under the minimization of the cost function Δ in Equation (3).

$$\Delta = \sqrt{(\sigma_{Mhh}^0 - \sigma_{Shh}^0)^2 + (\sigma_{Mvv}^0 - \sigma_{Svv}^0)^2} \quad (3)$$

where Δ represents the least square difference between the measured σ_{Mpp}^0 from the covariance matrix and the simulated σ_{Spp}^0 backscatter coefficients using the calibrated IEM [40]. More details about this volume scattering model and its determination can be found in [4].

2.2.2. Random Forest

We used the Random Forest (RF) [41,42] approach as a baseline machine learning method. RF is an ensemble learning method for classification and regression, in which a number of de-correlated decision trees are constructed during calibration. Before each split, m variables are selected randomly as candidates for splitting while $m \leq p$, where p is the total number of input features. Typically, values for m are the square root of p or even as low as 1 depending on the value of p . It should be noted that when the number of variables is large, but the fraction of relevant variables small, RF is likely to perform poorly with a small m . Compared with other classifiers, such as support vector machine and neural networks, RF uses the Out-Of-Bag (OOB) error samples, and once the OOB errors stabilizes, the training can be terminated and the number of trees are determined. RF also uses the OOB samples to measure the variable importance that represents the prediction strength of each variable, which might be used for the optimization of the choice of the input parameters. In this study, the OOB is used for the determination of the number of trees and set as $m = \sqrt{p}$. During the training processing of the RF, parameters were derived from both SAR and optical data are used as the inputs for the RF calibration, consisting of the HH, HV, and VV from SAR and the NDVI, and reflectance of green, red, Near Infrared (NIR), Short Wave Infrared 1 (SWIR1), and SWIR2 channels from the optical data.

3. Results

To gain insight, a three-component model-based decomposition proposed in [43] was used to study the surface, double-bounce, and volume scattering components as well as the randomness factor over time across the crop landscape. Figure 2 shows that the scattering components change over time for all crops, and, generally, the double-bounce scattering power is less than that of surface and volume scattering, which confirms our assumption of the weak double-bounce scattering over agricultural fields. One exception is the soybean field, in which the double-bounce scattering and the volume scattering have similar scattering power over time, while both remain dominated by surface scattering and the influence of the underlying ground conditions. It is also evident that in September (Day of Year (DoY) around 255), both the double-bounce and volume scattering over corn fields decrease sharply which might be because the SAR signal can penetrate the dry corn canopy easily when it was harvested at the end of the August and in early September. However, the surface scattering is still at a high level, likely due to the underlying soil moisture and surface roughness. In terms of the randomness factor, Figure 2 depicts that

its value is ranging from 0 to 6, which demonstrates that both Freeman and Yamaguchi volume scattering models cannot fully describe the crop canopy since their randomness factors are constant (i.e., 0 and 1), as explained in Section 2.2.1. In addition, the randomness factor of soybean tends to be higher than that of other crops, and for the corn at its late growth stage, the randomness factor is increased due to the majority of scattering being from the underlying ground.

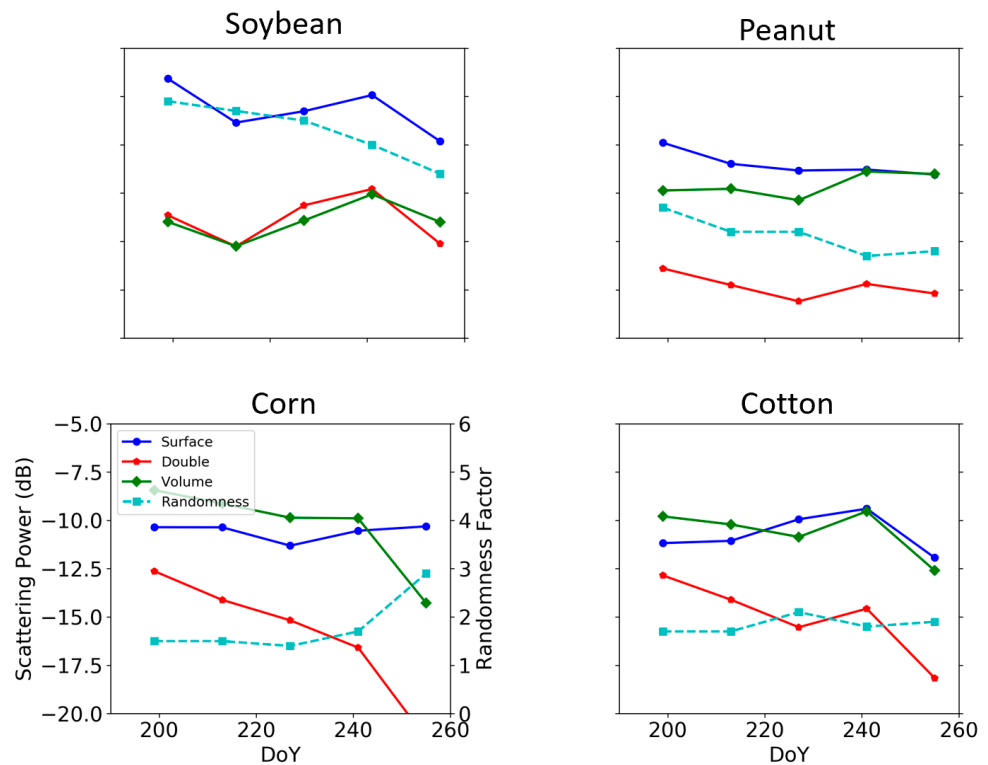


Figure 2. Temporal scattering components and randomness over the agricultural fields. DoY: Day of Year. Note that the scattering power is the average value.

The derived temporal soil moisture from the model-based decomposition is shown in Figure 3. Generally, it is wet on 18 July, and then becomes very dry on 12 September. The rice field (blue in Figure 1) show higher soil moisture than that of other crops due to its strong double-bounce backscattering caused by the interaction between the rice stem and underlying standing water. The cotton (red in Figure 1) shows high soil moisture on 18 July, but is very dry on 12 September. This trend is consistent with the precipitation data shown in Figure 1, when on 18 July, there is rain, while there is no rain at all on 12 September. Figure 4 shows that the RMSE of the retrieved soil moisture is about 9.34% and the R^2 is around 0.37.



Figure 3. Both the 30 m spatial resolution-derived soil moisture using the two-component model-based decomposition and “zoomed in” (black quadrilateral) example using the random forest machine learning method show the time series trending toward dryer conditions as harvest approaches.

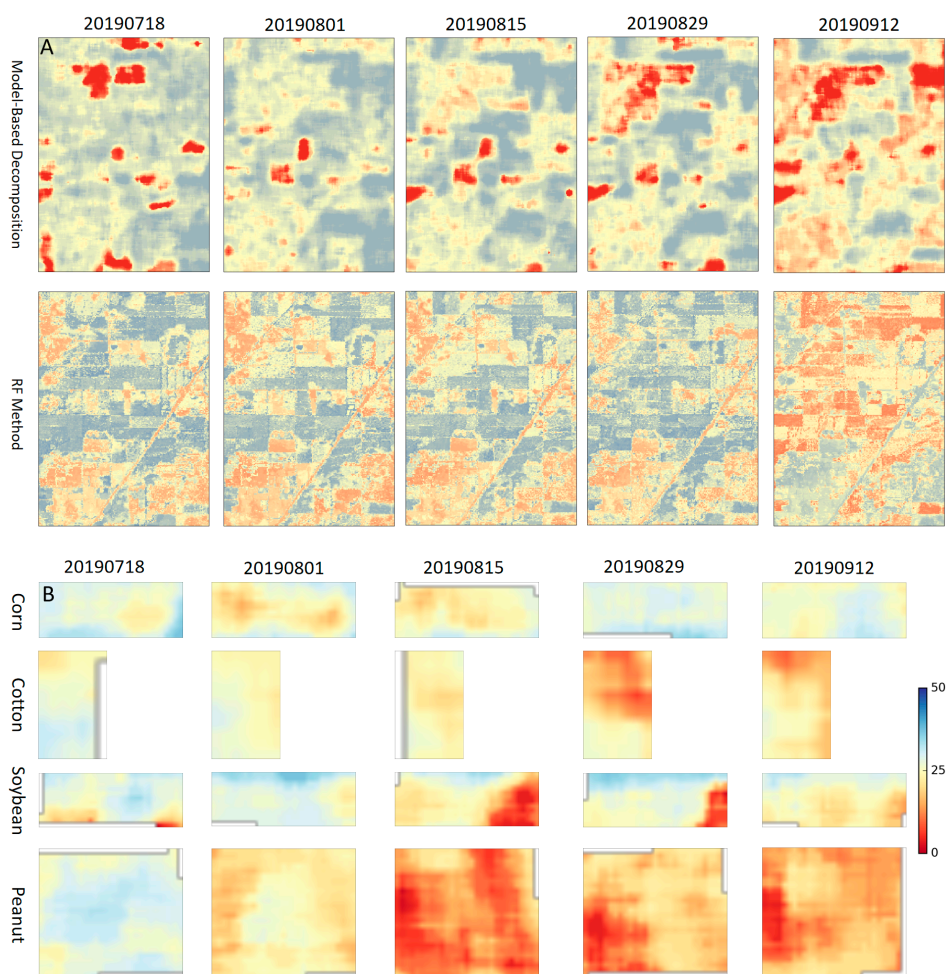


Figure 4. (A) Temporal soil moisture over the region outlined in the red dashed rectangular area in Figure 3 and (B) over the crop fields, showing the decrease in the soil moisture over time. Note: soil moisture depicted in Figure 4B is from the model-based decomposition method.

The temporal soil moisture from the RF method shown in Figure 3 shows that the soil moisture becomes low on 12 September, similar to the model-based decomposition method. In addition, the soil moisture of the rice field that is inundated from the beginning of June and middle of September is still higher than that of other crops, such as cotton and corn, due to the strong backscattering coefficient of HH and VV. The zoomed in region (Figure 4A) that is outlined in the red dashed rectangular area in Figure 3 shows the decrease in the soil moisture over time, and the similar spatial pattern between the model-based decomposition and RF methods. In addition, individual crop fields also show the decrease in the soil moisture over time, as depicted in Figure 4B. Due to the limited number of ground measurements, the leave-one-out cross validation is used for the accuracy assessment, showing that the RMSE is around 7.70% and R^2 is around 0.60 (Figure 5). Since the RF method requires both the SAR and optical data, only the common part of those two footprints (SAR and HLS) are used for the soil moisture estimation, which leads to chunked maps, as shown in Figure 3. Compared with the soil moisture from the model-based decomposition, the ML method show high accuracy, as more information is added in the RF, while the model-based decomposition makes use of the SAR data alone. For both the model-based decomposition and the RF methods, the underestimation is observed when the soil moisture is greater than 40 (vol.%). This is because the soil might be saturated by water when the soil moisture is higher than 40 (vol.%), which is difficult to be detected by the SAR signals.

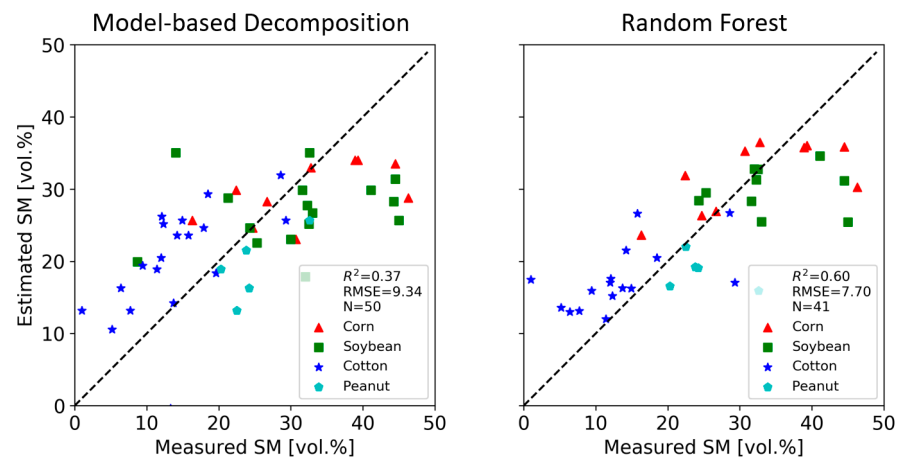


Figure 5. The Random Forest (RF) machine learning approach achieves lower Root Mean square Error (RMSE) by an order of magnitude compared to the model-based decomposition which requires no training data.

In addition, the importance of the input parameters of RF is also derived and shown in Figure 6, indicating the green, HH, red, and VV are the major contributions of the soil moisture estimation, showing the importance of the PALSAR polarizations in the soil moisture estimation. Instead, the HV shows less importance because it is primarily sensitive to the vegetation canopy other than the underlying surface.

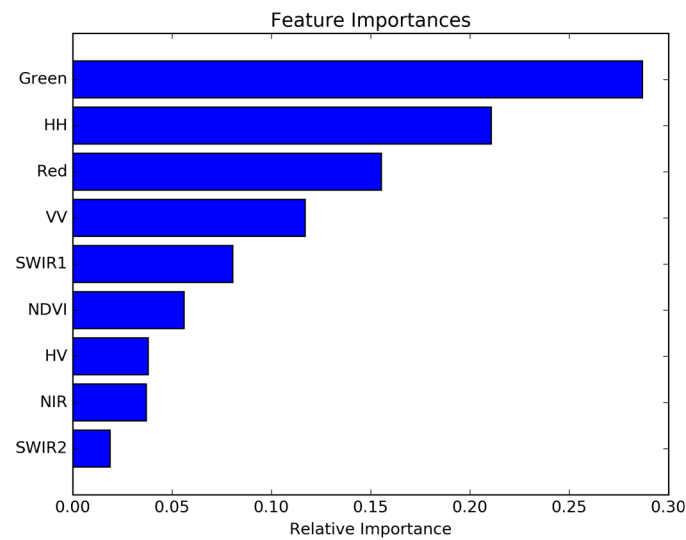


Figure 6. Importance of the input parameters of RF method. NIR: Near Infrared, SWIR: Short Wave Infrared.

Finally, instead of the RMSE and R^2 , an example of the comparison of temporal soil moisture over the soybean field using both the model-based decomposition and RF methods is shown in Figure 7. It shows that the estimated SM from the RF shows a high agreement with the ground measurement when the SM is less than 40 (vol.%), which is consistent with Figure 4. In addition, the SM from the model-based decomposition has the mean value of 30 (vol.%) over time, which is lower than the ground measurement with the mean value of 35 (vol.%). This comparison shows that the RF method outperforms the model-based decomposition, which might be because more parameters from both SAR and optical data are used.

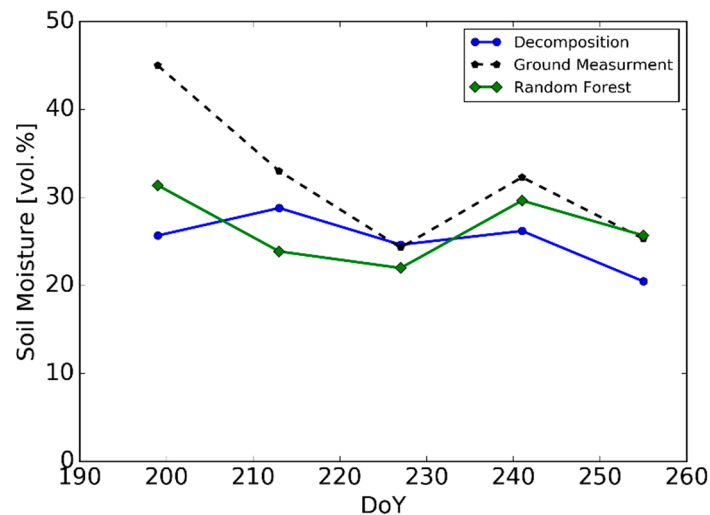


Figure 7. Example of the temporal soil moisture over a selected soybean field using the model-based decomposition and RF methods showing a high agreement of the estimated SM from the RF with the ground measurement when the SM is less than 40%.

4. Discussion

This research application aimed to advance the ARL of field-scale soil moisture mapping for agricultural monitoring. The mapping science focused on scaling physical and machine learning algorithms and fusing SAR and optical datacubes with moderate (30 m) spatial resolution. Getting robust, transparent, and cost-effective field-scale soil moisture

products will enable many new science applications, such as assessing irrigation regimes, nutrient management, tillage practices, yield, carbon cycling, and water quantity, that are all typically managed at field scales. The satellite data used quad polarization L-band, which provides an opportunity to extract several potentially useful scattering parameters using a relatively longer wavelength (24 cm) compared to using dual polarization C-band (5.5 cm) Sentinel-1, for example. In addition, the community will soon have access to operational and open moderate resolution L-band with the launch of NASA's NISAR mission as well as growing archives, such as SAUCOM.

In this study, a model-based decomposition that uses no training data was enhanced and applied to a cropland production region while comparing it to a popular machine learning RF method that can robustly integrate multi-modal inputs, such as optical indices like NDVI and SAR derivatives, like surface scattering, to estimate soil moisture conditions. Both the model-based decomposition and RF machine learning showed potential for estimating soil moisture at the field scale across the cropland production region. Generally, each had moderate accuracy, while RF achieved higher overall accuracy (R²) and precision (RMSE) when pooling all crops together into a single soil moisture prediction. Each showed the ability to capture landscape structure (fields and within-field variability) and temporal change using the time series satellite remote sensing datacubes. For brevity, we succinctly summarize the differences and strengths and limitations of the approaches.

The model-based decomposition method used in this research application requires assumptions about scattering properties and does not consider the contribution of the double-bounce scattering since it is rather weak over the agricultural fields [4,29,31]. The double-bounce scattering assumption is reasonable in this effort given that this scattering pathway has the least power (Figure 2). However, for some crops or potentially during specific growth stages, the double-bounce scattering might have an equivalent contribution or perhaps be even greater than the other scattering component, such as volume scattering over soybean fields. This can be influenced by the complexity of the canopy or changes in plant structure, such as large stalks that can develop in corn. As the influence of double-bounce scattering increases, the decomposition method enhanced in this research application might decrease in sensitivity and precision for SM estimation.

Another potential limitation of the model-based decomposition is that the attenuation effect is neglected. However, for the long wavelength L-band, the attenuation should be rather weak over agricultural fields since sometimes the attenuation can even be neglected over forestry regions [44]. Therefore, this is likely not a major driver of soil moisture retrieval performance in this application.

Third, while the decomposition is more automated compared to the machine learning approach, the IEM requires some calibration to determine the relationship between the surface roughness and correlation length [40]. These can be used from other studies; however, the transferability of those parameters might be a limiting factor. Ultimately, the decomposition is a semi-empirical method that requires tuning for ground conditions which might be landscape dependent, such as soil texture, for example. On the other hand, the RF machine learning approach was relatively robust in terms of integrating diverse data and taking advantage of statistical patterns. As the growth of in-field sensors continues and low-power wide-area technologies are adopted, these ML approaches will have more opportunities to source training data. A caution is to promote the consideration of meaningful, physical parameters of models to understand the underlying drivers.

Overall, the model-based decomposition with its noted limitations shows the ability to capture landscape features, such as fields with similar conditions and an RMSE of 9.34 (vol.%). The model-based decomposition method also achieves higher accuracy than that of the RF method, except for soybean. However, due to the limited number and variation of the observations for each crop, a single "best" recommendation is inappropriate. Conversely, the RF method shows higher accuracy than that of the model-based decomposition here, in part due to the inclusion of the optical parameters used in combination with the SAR parameters. Over time, the soil moisture from the model-based decomposition

tends to decrease from July to September, which is also similar to outcomes derived from the RF method, and so are their spatial patterns. The importance of the input parameters shows that the HH and VV show high importance, as shown in Figure 5, demonstrating the influence of SAR parameters on the soil moisture estimation. The transferability of these parameters and the RF machine learning method is an area for the next steps in scaling the ARL of these tools.

5. Conclusions

This paper studies the SM retrieval using the polarimetric model-based decomposition method and machine learning RF method, showing that both methods show high potential in the soil moisture retrieval with an RMSE of 9.34 (vol.%) and 7.70 (vol.%), and R^2 of 0.37 and 0.60, respectively. The RF method shows higher accuracy than that of the model-based decomposition due to the addition of the optical parameters. However, the RF method requires ground measurements for the calibration, while the model-based decomposition does not need external auxiliary data. The major limitation of the model-based decomposition is neglecting the double-bounce scattering, which might not be applicable to stem-based crops such as corn. To scale the RF method to a broad region and crops, more training samples with a high level of variety will be required. In future, as more ground measurements are collected, the model-based decomposition will be studied on more vegetation types, such as wheat, bare ground, forest, etc. In addition, the RF method can also be trained with more training samples with high variations and broad crop types.

Author Contributions: Conceptualization, X.H. and N.T.; methodology, X.H.; software, X.H.; validation, X.H. and B.Z.; data curation, X.H. and B.Z.; writing—original draft preparation, X.H.; writing—review and editing, N.T., J.W., M.R. and M.H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by NASA NISAR Award #80NSSC19K1511 and the NASA Harvest Consortium (No. 80NSSC17K0625).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The ALOS-2 quad pol PALSAR data are provided by Japan Aerospace Exploration Agency (JAXA) through the 6th ALOS-2 Research Announcement (PI No. P3141002). USDA is an equal opportunity employer and provider.

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

References

1. Koster, R.D.; Dirmeyer, P.A.; Guo, Z.; Bonan, G.; Chan, E.; Cox, P.; Gordon, C.T.; Kanae, S.; Kowalczyk, E.; Lawrence, D.; et al. Regions of Strong Coupling Between Soil Moisture and Precipitation. *Science* **2004**, *305*, 1138–1140. [[CrossRef](#)] [[PubMed](#)]
2. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Rev.* **2010**, *99*, 125–161. [[CrossRef](#)]
3. 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](#)]
4. Huang, X.; Wang, J.; Shang, J. An Integrated Surface Parameter Inversion Scheme Over Agricultural Fields at Early Growing Stages by Means of C-Band Polarimetric RADARSAT-2 Imagery. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2510–2528. [[CrossRef](#)]
5. Huang, X.; Wang, J.; Shang, J. An Adaptive Two-Component Model-Based Decomposition on Soil Moisture Estimation for C-Band RADARSAT-2 Imagery Over Wheat Fields at Early Growing Stages. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1–5. [[CrossRef](#)]
6. Xing, M.; He, B.; Ni, X.; Wang, J.; An, G.; Shang, J.; Huang, X. Retrieving Surface Soil Moisture over Wheat and Soybean Fields during Growing Season Using Modified Water Cloud Model from Radarsat-2 SAR Data. *Remote Sens.* **2019**, *11*, 1956. [[CrossRef](#)]
7. 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](#)]
8. Felfelani, F.; Pokhrel, Y.; Guan, K.; Lawrence, D. Utilizing SMAP Soil Moisture Data to Constrain Irrigation in the Community Land Model. *Geophys. Res. Lett.* **2018**, *45*, 12892–12902. [[CrossRef](#)]

9. Lievens, H.; Reichle, R.H.; Liu, Q.; De Lannoy, G.; Dunbar, R.S.; Kim, S.B.; Das, N.N.; Cosh, M.H.; Walker, J.P.; Wagner, W. Joint Sentinel-1 and SMAP data assimilation to improve soil moisture estimates. *Geophys. Res. Lett.* **2017**, *44*, 6145–6153. [[CrossRef](#)]
10. Das, N.N.; Entekhabi, D.; Dunbar, R.S.; Colliander, A.; Chen, F.; Crow, W.; Jackson, T.J.; Berg, A.; Bosch, D.D.; Caldwell, T.G.; et al. The SMAP mission combined active-passive soil moisture product at 9 km and 3 km spatial resolutions. *Remote Sens. Environ.* **2018**, *211*, 204–217. [[CrossRef](#)]
11. Wang, J.; Hsu, A.; Shi, J.; O'Neill, P.; Engman, E. A comparison of soil moisture retrieval models using SIR-C measurements over the little Washita River watershed. *Remote Sens. Environ.* **1997**, *59*, 308–320. [[CrossRef](#)]
12. Zribi, M.; Dechambre, M. A new empirical model to retrieve soil moisture and roughness from C-band radar data. *Remote Sens. Environ.* **2003**, *84*, 42–52. [[CrossRef](#)]
13. Baghdadi, N.; Aubert, M.; Cerdan, O.; Franchistéguy, L.; Viel, C.; Eric, M.; Zribi, M.; Desprats, J.F. Operational Mapping of Soil Moisture Using Synthetic Aperture Radar Data: Application to the Touch Basin (France). *Sensors* **2007**, *7*, 2458–2483. [[CrossRef](#)] [[PubMed](#)]
14. Attema, E.P.W.; Ulaby, F.T. Vegetation modeled as a water cloud. *Radio Sci.* **1978**, *13*, 357–364. [[CrossRef](#)]
15. Stamenković, J.; Ferrazzoli, P.; Guerriero, L.; Tuia, D.; Thiran, J.P. Joining a Discrete Radiative Transfer Model and a Kernel Retrieval Algorithm for Soil Moisture Estimation From SAR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 3463–3475. [[CrossRef](#)]
16. Kim, S.-B.; Huang, H.; Liao, T.-H.; Colliander, A. Estimating Vegetation Water Content and Soil Surface Roughness Using Physical Models of L-Band Radar Scattering for Soil Moisture Retrieval. *Remote Sens.* **2018**, *10*, 556. [[CrossRef](#)]
17. Ulaby, F.T.; Sarabandi, K.; McDonald, K.; Whitt, M.; Dobson, M.C. Michigan microwave canopy scattering model. *Int. J. Remote Sens.* **1990**, *11*, 1223–1253. [[CrossRef](#)]
18. Prévot, L.; Champion, I.; Guyot, G. Estimating surface soil moisture and leaf area index of a wheat canopy using a dual-frequency (C and X bands) scatterometer. *Remote Sens. Environ.* **1993**, *46*, 331–339. [[CrossRef](#)]
19. Gherboudj, I.; Magagi, R.; Berg, A.; Toth, B. Soil moisture retrieval over agricultural fields from multi-polarized and multi-angular RADARSAT-2 SAR data. *Remote Sens. Environ.* **2011**, *115*, 33–43. [[CrossRef](#)]
20. He, B.; Xing, M.; Bai, X. A Synergistic Methodology for Soil Moisture Estimation in an Alpine Prairie Using Radar and Optical Satellite Data. *Remote Sens.* **2014**, *6*, 10966–10985. [[CrossRef](#)]
21. 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](#)]
22. Bai, X.; He, B.; Li, X.; Zeng, J.; Wang, X.; Wang, Z.; Zeng, Y.; Su, Z. First Assessment of Sentinel-1A Data for Surface Soil Moisture Estimations Using a Coupled Water Cloud Model and Advanced Integral Equation Model over the Tibetan Plateau. *Remote Sens.* **2017**, *9*, 714. [[CrossRef](#)]
23. El Hajj, M.; Baghdadi, N.; Zribi, M.; Belaud, G.; Cheviron, B.; Courault, D.; Charron, F. Soil moisture retrieval over irrigated grassland using X-band SAR data. *Remote Sens. Environ.* **2016**, *176*, 202–218. [[CrossRef](#)]
24. 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](#)]
25. Kumar, K.; Rao, H.P.S.; Arora, M.K. Study of water cloud model vegetation descriptors in estimating soil moisture in Solani catchment. *Hydrol. Process.* **2014**, *29*, 2137–2148. [[CrossRef](#)]
26. Freeman, A.; Durden, S.L. A three-component scattering model for polarimetric SAR data. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 963–973. [[CrossRef](#)]
27. Hajnsek, I.; Jagdhuber, T.; Schon, H.; Papathanassiou, K.P. Potential of Estimating Soil Moisture Under Vegetation Cover by Means of PolSAR. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 442–454. [[CrossRef](#)]
28. Jagdhuber, T.; Hajnsek, I.; Bronstert, A.; Papathanassiou, K.P. Soil Moisture Estimation Under Low Vegetation Cover Using a Multi-Angular Polarimetric Decomposition. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 2201–2215. [[CrossRef](#)]
29. Ballester-Berman, J.D.; Vicente-Guijalba, F.; Lopez-Sanchez, J.M. POLARIMETRIC SAR MODEL FOR SOIL MOISTURE ESTIMATION OVER VINEYARDS AT C-BAND. *Prog. Electromagn. Res.* **2013**, *142*, 639–665. [[CrossRef](#)]
30. Jagdhuber, T.; Hajnsek, I.; Bronstert, A.; Papathanassiou, K.P. An Iterative, Generalized, Hybrid Decomposition on Fully Polarimetric SAR Data for Soil Moisture Retrieval under Vegetation. In Proceedings of the 10th European Conference on Synthetic Aperture Radar, Berlin, Germany, 3–5 June 2014; pp. 818–821.
31. Wang, H.; Magagi, R.; Goita, K. Potential of a two-component polarimetric decomposition at C-band for soil moisture retrieval over agricultural fields. *Remote Sens. Environ.* **2018**, *217*, 38–51. [[CrossRef](#)]
32. 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](#)]
33. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
34. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [[CrossRef](#)]
35. Braswell, B.; Fisk, J.; Cooke, I. Efficient, Chainable Time Series Processing of Raster Stacks. Applied-GeoSolutions/multitemporal. Available online: <https://github.com/Applied-GeoSolutions/multitemporal> (accessed on 26 April 2018).

36. Howlett, J.; Whittaker, E.; Robinson, G.; Pettofrezzo, A.J.; John, F. The Calculus of Observations. *Math. Gaz.* **1969**, *53*, 107. [[CrossRef](#)]
37. Yamaguchi, Y.; Moriyama, T.; Ishido, M.; Yamada, H. Four-component scattering model for polarimetric SAR image decomposition. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1699–1706. [[CrossRef](#)]
38. Van Zyl, J.J.; Arii, M.; Kim, Y. Model-Based Decomposition of Polarimetric SAR Covariance Matrices Constrained for Nonnegative Eigenvalues. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3452–3459. [[CrossRef](#)]
39. Kim, Y.; Jackson, T.; Bindlish, R.; Lee, H.; Hong, S. Radar Vegetation Index for Estimating the Vegetation Water Content of Rice and Soybean. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 564–568. [[CrossRef](#)]
40. Baghdadi, N.; Zribi, M.; Paloscia, S.; Verhoest, N.E.C.; Lievens, H.; Baup, F.; Mattia, F. Semi-Empirical Calibration of the Integral Equation Model for Co-Polarized L-Band Backscattering. *Remote Sens.* **2015**, *7*, 13626–13640. [[CrossRef](#)]
41. Breiman, L. Random forests. *Machine Learning* **2001**, *45*, 5–32. [[CrossRef](#)]
42. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2002.
43. Huang, X.; Wang, J.; Shang, J. Simplified adaptive volume scattering model and scattering analysis of crops over agricultural fields using the RADARSAT-2 polarimetric synthetic aperture radar imagery. *J. Appl. Remote Sens.* **2015**, *9*, 096026. [[CrossRef](#)]
44. Cloude, S.; Papathanassiou, K. Three-stage inversion process for polarimetric SAR interferometry. *IEE Proc. Radar Sonar Navig.* **2003**, *150*, 125. [[CrossRef](#)]