Article

Soil Moisture Estimation Based on Polarimetric Decomposition and Quantile Regression Forests

Li Zhang 1,2,3, Xiaolei Lv 1,2,3,* and Rui Wang 1,2,3

1 Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China
2 Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
3 School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: lvxl@aircas.ac.cn

Abstract: The measurement of surface soil moisture (SSM) assists in making agricultural decisions, such as precision irrigation and flooding or drought predictions. The critical challenge for SSM estimation in vegetation-covered areas is the coupling between vegetation and surface scattering. This study proposed an SSM estimation method based on polarimetric decomposition and quantile regression forests (QRF) to overcome this problem. Model-based polarimetric decomposition separates volume scattering, double-bounce scattering, and surface scattering, while eigenvalue-based polarimetric decomposition provides additional parameters to describe the scattering mechanism. The combined use of these parameters explains the polarimetric SAR scattering information from multiple perspectives, such as vegetation, surface roughness, and SSM. As different crops differ in morphology and structure, it is essential to investigate the potential of varying polarimetric parameters to estimate SSM in areas covered by different crops. QRF, a regression method applicable to high-dimensional predictor variables, is used to estimate SSM from these parameters. In addition to the SSM estimates, QRF can also provide the predicted uncertainty intervals and quantify the importance of the different parameters in the SSM estimates. The performance of QRF in SSM estimation was tested using data from the soil moisture active passive validation experiment 2012 (SMAPVEX12) and compared with copula quantile regression (CQR). The SSM estimated by the proposed method was consistent with the in situ SSM, with the root-mean-square-error ranging from 0.037 cm$^3$/cm$^3$ to 0.079 cm$^3$/cm$^3$ and correlation coefficients ranging from 0.745 to 0.905. Meanwhile, the method proposed in this study can provide both the uncertainty of SSM estimation and the importance of different polarimetric parameters.

Keywords: polarimetric SAR; surface soil moisture; polarimetric decomposition; quantile regression forests; SMAPVEX12

1. Introduction

Surface soil moisture (SSM) occupies a critical position in agricultural production, which can be used to guide irrigation and predict droughts and floods [1,2]. If SSM is low during critical periods of vegetation growth, crops will wilt, and prolonged water shortages can cause this wilting to be irreversible [3]. In flood prediction, SSM is a crucial step in predicting flood sensitivity as an initial condition that indicates the location of river systems that may be affected by flooding [4]. Conventional in situ measurement methods provide SSM at discrete points; however, they cannot acquire SSM with an extensive spatial range and when it is continuous in space [5,6]. As opposed to in situ measurements, remote sensing provides a long-term, large-scale SSM monitoring method [6]. Remote sensing instruments such as optical, infrared, and microwave have all been used to estimate SSM [7–9]. Due to differences in penetration, optical and infrared are strongly influenced by weather and vegetation [10].
In comparison, microwave remote sensing with penetration has apparent advantages in estimating SSM [5,11]. Both passive and active microwave remote sensings are available for SSM estimation. Passive microwave remote sensing has been widely applied to the estimation of SSM on large scales, with the crucial targets being Soil Moisture and Ocean Salinity (SMOS) [12], and Soil Moisture Active Passive (SMAP) [13]. For agriculture-related applications, the resolution of passive microwave remote sensing SSM products is too coarse to meet field-scale needs [5]. Active microwave remote sensing can provide high resolution. Still, SSM estimations face significant challenges due to the influence of backscattering coefficients by vegetation, surface roughness, soil texture, and other factors [6].

In recent decades, numerous scholars explored estimating SSM using synthetic aperture radar (SAR). The SAR backscattering coefficient is mainly affected by SSM and surface roughness for bare ground. Based on this, scholars proposed theoretical, empirical and semi-empirical models to forward the SAR backscattering coefficients [13–20]. Since vegetation height, moisture content, morphology, and planting method all contribute to the SAR backscattering coefficients, the estimation of SSM in the vegetation-covered area is much more complicated than that of the bare surface [6,21]. With reasonable simplification and modeling of vegetation, scholars have developed empirical and theoretical models of vegetation scattering [22,23]. Traditional vegetation scattering models usually require vegetation parameters, which are usually obtained from field measurements or optical remote sensing [24,25]. Field measurements are often difficult to perform in practice, and optical data are strongly influenced by weather and cloud cover.

Beyond model-based estimation, several new methods have great potential for estimating SSM. These new methods can be roughly divided into four categories. The first category is change-detection-based SSM estimation, which primarily assumes that soil surface roughness and vegetation remain constant between two adjacent measurements [21,26,27]. The second category is based on statistics, mainly including Bayesian posterior estimation, copula probability map, and so on [28–30]. These methods can give an uncertainty range for SSM estimation. The third category is mainly SSM estimation based on polarimetric SAR data, which can estimate SSM without vegetation parameters from field measurements or optical [31–34]. The fourth category is the method based on machine learning and artificial neural network. It can effectively construct the nonlinear relationship between multiple parameters without considering the physical relationship between them [35–38]. In practice, vegetation changes in spring and autumn cannot be ignored, and there is a time difference between optical remote sensing and SAR remote sensing. Therefore, the change detection methods and vegetation model have difficulties estimating SSM. Polarimetric SAR data can separate surface and volume scattering, which has excellent potential for SSM retrieval in vegetated areas. Machine learning methods and statistical methods are weakly dependent on physical principles. They can flexibly capture the relationship between SSM and different SAR parameters, making it possible to use multiple polarimetric parameters to estimate SSM.

Polarimetric SAR can obtain various physical properties of surface scattering. In 2003, Hajnsek et al. [39] used three decomposition parameters based on eigenvalue decomposition, namely scattering entropy (H), scattering anisotropy (A), and alpha angle (α), to estimate SSM at the bare surface. This method estimates the SSM and roughness of bare surfaces from a principled level, making it possible to separate the contributions of surface roughness and SSM to backscattering. In 2016, Lian He et al. used a combination of adaptive non-negative eigenvalue decomposition and Oh model to estimate SSM in farmland, with root-mean-square-errors (RMSE) between 0.10 cm$^3$/cm$^3$ and 0.14 cm$^3$/cm$^3$ [40]. In 2017, Wang H. et al. compared three model-based polarimetric decompositions to estimate SSM on the surface covered by different crops, with estimated RMSE between 0.06 cm$^3$/cm$^3$ and 0.11 cm$^3$/cm$^3$ [33]. In addition, many scholars have used model-based polarimetric decomposition to decompose the correlation matrix or covariance matrix of polarimetric SAR data into surface scattering, volume scattering, and double-bounce scattering, and use surface scattering or double-bounce scattering to estimate SSM [11,41–44]. These methods
for estimating SSM using polarimetric data show great potential on vegetated surfaces without relying on field measurements of vegetation data and optical data. However, the estimation of SSM from polarimetric data is usually related to theoretical models such as X-Bragg, integral equation model (IEM), and Fresnel models. These models usually have complex expressions, can only be used in a limited range of roughness, and are sensitive to incident angles [33,39]. In addition, a few scholars have also studied the potential of different polarimetric decomposition parameters to estimate SSM [45,46]. They found that using various polarimetric decomposition parameters may help improve the accuracy of SSM estimation. Some scholars have also used multiple linear regression and support vector regression methods to estimate SSM from polarimetric decomposition parameters and obtained a satisfactory estimation effect [47–49]. However, all these studies considered only one vegetation cover and could not account for the differences in the importance of different polarimetric parameters in SSM estimation.

This experiment has three objectives: estimating SSM in different crops-covered areas using only polarimetric SAR data, giving the uncertainty range of SSM estimation, and quantifying the importance of each polarimetric parameter. For this reason, an SSM estimation method combining polarimetric decomposition and quantile regression forests (QRF) is proposed in this experiment. Random forests have been used to classify crops [50], estimate wheat parameters [51] and invert vegetation phenology [52], etc. Vijay Pratap Yadav et al. [53] have successfully estimated SSM in sparsely vegetated areas using random forest regression combined with a water cloud model from Sentinel-1 data. However, this estimation method requires additional optical data. To our knowledge, QRF is rarely combined with polarimetric decomposition to estimate SSM. In the current experiment, a new SSM estimation method is developed using the data from the soil moisture active passive validation experiment 2012 (SMAPVEX12). The experiments first extract various polarimetric parameters from airborne polarimetric SAR data using model-based and eigenvalue-based decomposition. Subsequently, SSM is estimated from polarimetric parameters and backscattering coefficients using QRF; finally, the estimated uncertainty range is given. The importance of different polarimetric parameters in estimating SSM under different vegetation coverage is analyzed. In addition, the experiment also compared QRF and copula quantile regression (CQR) to illustrate the advantages of QRF in SSM estimation.

The remainder of this paper is organized as follows. Section 2 describes the details of the study area, the uninhabited aerial vehicle synthetic aperture radar (UAVSAR) data, and field measurements. Section 3 mainly introduces the principles of two polarimetric decompositions, QRF and CQR. The estimation results are presented in Section 4, and the results are discussed in Section 5. Finally, the entire article is summarized in Section 6 and predicts subsequent learning.

2. Materials

2.1. Study Area

The data of this experiment are acquired from SMAPVEX12, which is located in the town of Elm Creek (98°0′23″W, 49°40′48″N) in the southern center of Manitoba, Canada, belonging to the Hong River Basin. Throughout the experimental area, the soil texture is diverse, and the terrain is flat or slightly undulating with slopes from 0% to 2% [54]. The region is dominated by mixed grassland agriculture, with wetlands and forest cover in the northwest. The current study only considers agricultural areas mainly covered by annual or perennial crops. These crops have a short growing season, typically seeding in April/May and harvesting in July/August. The data acquisition time used in this experiment was from June 17 to July 17, during which time all land was covered with crops. The geographical location of the experimental area and the vegetation classification map are given in Figure 1. The vegetation classification map was produced from a supervised classification of imagery acquired by SPOT-4, DMCii, and RADARSAT-2 [55], which can be downloaded at Earthdata. More information on experimental areas can be found in the
literature of McNairn, H. et al. [54], SMAPVEX12 experimental plan [56], and SMAPVEX12 database report [57].

Figure 1. The map of the geographical location of the experimental area, vegetation classification map, coverage of the four flights, and in situ SSM measurement points.

2.2. UAVSAR Data

The UAVSAR is an airborne L-band radar system operating at 1.26 GHz, equipped on the NASA Gulfstream-III aircraft. From 17 June 2012, to 17 July 2012, SMAPVEX12 collected 13 days of UAVSAR quad-polarimetric data on the same day as the agricultural SSM probe measurements. For each acquisition date, there are four flights, named 31603, 31604, 31605, and 31606. The UVASAR data collected SMAPVEX12 looked to the left of the direction of flight, with an incidence angle range from 26° to 65°. The coverage of the four flights is given in Figure 1, from which it can be seen that only 31606 covered all the SSM measurement points (black points in Figure 1). Therefore, 31606 5 × 5 resampled ground projected complex data are selected for this experiment. All UAVSAR data can be found on the Alaska Satellite Facility Data Search [58].

The pre-processing of UAVSAR images is accomplished by the polarimetric SAR data processing and education toolbox (PolSARpro) provided by the European Space Agency (ESA). The pre-processing steps include extracting the coherence matrix \( T \) and covariance matrix \( C \) of the polarimetric data and reducing speckle noise using a 7 × 7 refined Lee filter [11].

Furthermore, the UAVSAR backscattering coefficients with incident angles normalized to 40° are collected in this study. The study areas are covered with different crops, so it is necessary to analyze the sensitivity between the backscattering coefficients and in situ SSM. Figure 2 shows the sensitivity analysis between SAR backscattering coefficient and in situ SSM under different crops. The slope and correlation coefficient of the sensitivity analysis corresponding to Figure 2 are shown in Table 1.
Figure 2. Sensitivity analysis between SAR backscattering coefficients and in situ SSM under different crops. (a) HH polarization and in situ SSM; (b) VV polarization and in situ SSM; (c) HV polarization and in situ SSM.

Table 1. Slope and correlation coefficients (R) between in situ SSM and backscattering coefficients for different crop coverage areas.

| Sensitivity | Polarization | Canola | Corn | Grassland | Soybeans | Cereal |
|-------------|--------------|--------|------|-----------|----------|--------|
| slope       | HH           | 10.27  | -0.59| 12.17     | -6.32    | 9.90   |
|             | VV           | 9.18   | 18.61| 8.50      | 1.90     | 11.94  |
|             | HV           | 17.78  | 3.22 | 14.82     | -9.85    | 8.83   |

| R           | HH           | 0.42   | -0.01| 0.33      | -0.27    | 0.55   |
|             | VV           | 0.35   | 0.34 | 0.30      | 0.09     | 0.59   |
|             | HV           | 0.42   | 0.04 | 0.29      | -0.32    | 0.38   |

2.3. Ground Measurements

Given that this experiment focuses on SSM estimation in agricultural crop-covered areas, in situ SSM data of farm areas of SMAPVEX12 are used as the field measurement data. The in situ SSM collected by SMAPVEX12 originated from 55 agricultural fields, including 19 soybeans, 16 cereal fields (13 spring wheat, one oat, two winter wheat), eight corn, seven canola, and five grassland fields. All fields are at least 800 m × 800 m in size, each field is measured at 16 points, and each point is measured three times in different directions. The SSM 0–6 cm below the surface is measured at all measurement points through a Water Hydra Probe or Delta-T Theta Probe. The Water Hydra Probe is a frequency domain reflectometer sensor that measures the reflected voltage and converts this received signal to the actual dielectric constant and volumetric water content of the soil. The Delta-T sensor is an impedance probe directly converts the received voltage into soil volumetric water content. Those sensors were calibrated independently in each farm field, with a RMSE after calibration lower than 0.037 cm³/cm³. All in situ SSM measurement times are consistent with UVASAR acquisition times, starting at 7:30 AM and ending around noon [57,59]. More details related to field measurements of SSM can be found in Literature [54,56,57]. The locations of all farmlands are drawn in Figure 1. The experiment averaged three
measurements of each point and averaged 16 points of each farmland as the SSM of each field. For farmland covered by different vegetation, the statistical information such as the mean and variance of SSM are given in Table 2. In addition, SMAPVEX12 measured the height of the crops in the study sample fields for 11 days between 16 June and 18 July 2012. Table 3 gives the minimum, maximum, mean, and standard deviation of in situ measured heights for different crops.

Table 2. Statistical parameters for the minimum (MIN), mean (MEAN), maximum (MAX), standard deviation (STD), and the number of measurements (N) of the in situ SSM.

| Crop     | MIN cm$^3$/cm$^3$ | MEAN cm$^3$/cm$^3$ | MAX cm$^3$/cm$^3$ | STD cm$^3$/cm$^3$ | N   |
|----------|-------------------|--------------------|-------------------|-------------------|-----|
| soybeans | 0.052             | 0.220              | 0.588             | 0.105             | 201 |
| cereal   | 0.087             | 0.294              | 0.525             | 0.110             | 150 |
| corn     | 0.052             | 0.137              | 0.313             | 0.057             | 102 |
| canola   | 0.059             | 0.207              | 0.378             | 0.087             | 76  |
| grassland| 0.159             | 0.296              | 0.518             | 0.071             | 58  |

Table 3. Statistical parameters for the minimum (MIN), mean (MEAN), maximum (MAX) and standard deviation (STD) of the height of crops.

| Crop     | MIN cm | MEAN cm | MAX cm | STD cm |
|----------|--------|---------|--------|--------|
| soybeans | 7.0    | 26.6    | 75.9   | 17.5   |
| cereal   | 33.7   | 74.2    | 113.6  | 17.6   |
| corn     | 13.2   | 102.9   | 269.6  | 73.2   |
| canola   | 6.5    | 88.2    | 148.2  | 37.4   |
| grassland| 5.9    | 30.5    | 66.1   | 15.9   |

3. Methods

The main objective of this study is to estimate SSM in areas covered by various vegetation crops by combining the polarimetric decomposition parameters and the backscattering coefficients of different polarizations. The polarimetric decomposition methods used are Freeman-Durden decomposition and eigenvalue-based decomposition. The two decomposition methods extract various polarimetric parameters with different meanings from the physical and mathematical levels, which can explain polarimetric scattering from different perspectives. Subsequently, SSM is estimated from the decomposed parameters and backscattering coefficients by QRF. QRF is a regression approach suitable for high-dimensional predictor variables, which can estimate the value at any quantile of the dependent variable. Compared to other regression methods that can only calculate the conditional mean, QRF can give the prediction interval and thus determine the confidence of the prediction. In addition, the application of QRF quantifies the importance of different polarimetric parameters and backscattering coefficients for SSM estimation. To illustrate the effectiveness of QRF in SSM estimation, we have compared QRF and CQR concerning their ability to estimate SSM from polarimetric SAR data. The flow chart of the study is shown in Figure 3.
Figure 3. Flowchart of the algorithm for estimating soil moisture (SSM) from L-band UAVSAR data combining polarimetric decomposition and quantile regression forests (QRF).

3.1. Polarimetric Decompositions

3.1.1. Freeman-Durden Decomposition

The Freeman-Durden decomposition [60] is the most typical model-based polarimetric decomposition, which decomposes the polarimetric covariance matrix $C$ into three different scattering mechanisms: surface scattering $C_s$, double-bounce scattering $C_d$, and volume scattering $C_v$. The Freeman-Durden decomposition assumes that surface scattering is a first-order Bragg surface scatter, double-bounce scattering is a dihedral corner reflector, and volume scattering is a set of randomly oriented dipoles. There are five parameters to be determined for this decomposition process: volume scattering magnitude $f_v$, double-bounce scattering magnitude $f_d$, surface scattering magnitudes $f_s$, $\alpha$ and $\beta$. The Freeman-Durden decomposition suggests that the HV scattering associated with depolarization originates mainly from vegetation, so $f_v$ can be determined from cross-polarization. The values of $\alpha$ and $\beta$ depend on whether the dominant scattering is surface or double-bounce scattering. The definite equation between the total scattering matrix and the partial scattering matrix can calculate all other parameters.

Freeman-Durden decomposition is widely used in SSM estimation due to its clear principle and simple calculation. However, the disadvantage of the Freeman-Durden decomposition is that it tends to overestimate the volume scattering, resulting in a negative power of the decomposed surface scattering or double-bounce scattering. After decomposition of the UAVSAR images of the experimental area, it shows no negative power in the data used in this experiment. From the study of Wang H. et al. [33] in 2017, it has been found that Freeman-Durden decomposition resulted in optimal estimation results in corn and wheat and arrived at reasonable results in soybean and canola. The dataset used in this study is the same as that of Wang et al., so the Freeman-Durden decomposition is selected as the model-based decomposition. The powers of three scattering mechanisms of
surface, double-bounce, and volume are used to estimate SSM, denoted as Odd, Dbl, and Vol, respectively.

3.1.2. Eigenvalue-Based Decomposition

In contrast to model-based decomposition, eigenvalue-based decomposition [61] is based on mathematical theory, which does not depend on specific statistical distribution assumptions and is not constrained by physical models. Eigenvalue decomposition divides the $3 \times 3$ coherence matrix ($T$) into different types of scattering processes and corresponding relative magnitudes. Based on the eigenvalues of the coherence matrix $T$, eigenvalue-based decomposition constructs a variety of polarimetric parameters reflecting different scattering mechanisms. Cloude and Pottier [62] proposed three parameters to describe the fully polarimetric scattering. These parameters include: (1) polarimetric entropy ($H$) represents the randomness of the medium; (2) anisotropy ($A$) is an additional parameter to $H$ that represents the relative importance of the two smaller eigenvalues; (3) $\alpha$ angle indicates the physical mechanism of scattering. Hajnsek et al. have previously found that the $H$ and $\alpha$ parameters are mainly related to SSM, and the $A$ parameter can be used to estimate surface roughness [39]. In addition, the eigenvalue-based decomposition also derives some other new parameters. The radar vegetation index ($rvi$) is a parameter that can express vegetation growth dynamics, as shown in Equation (1). Target randomness ($Pr$) represents the randomness of the scattering process, as shown in Equation (2).

$$rvi = \frac{4\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$  \hspace{1cm} (1)

$$Pr = \sqrt{\frac{3}{2}} \sqrt{\frac{\lambda_2^2 + \lambda_3^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$  \hspace{1cm} (2)

where $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of $T$.

Furthermore, single eigenvalue relative difference ($SERD$) and double eigenvalue relative difference ($DERD$) can also be used to describe the characteristics of natural media [61]. The $SERD$ parameter is valid for regions with high $H$, where it can determine the characteristics and magnitude of the different scattering mechanisms. The $DERD$ parameter is sensitive to surface roughness and relatively less sensitive to SSM. These two parameters can be calculated from the eigenvalues of $T$ satisfying the reflection symmetry assumption. The eigenvalue corresponding to $\alpha_i < \pi/4$ is defined as $\lambda_s$, the eigenvalue corresponding to $\alpha_i \geq \pi/4$ is $\lambda_d$, and the remaining eigenvalue is $\lambda_3$. Where $\cos \alpha_i$ is the magnitude of the first component of the $i$th eigenvector $e_i$ of $T$. The expressions for $SERD$ and $DERD$ are shown in Equations (3) and (4).

$$SERD = \frac{\lambda_s - \lambda_3}{\lambda_s + \lambda_3}$$  \hspace{1cm} (3)

$$DERD = \frac{\lambda_d - \lambda_3}{\lambda_d + \lambda_3}$$  \hspace{1cm} (4)

3.2. Normalization of Polarimetric Parameters

In this study, the incidence angle of UAVSAR varies widely. The polarimetric decomposition parameters vary with the incidence angle, so the incidence angle needs to be unified to a single value. This study uses a histogram-based incidence angle normalization method proposed by Mladenova et al. [63]. The histogram method uses the mean ($\sigma^\varphi$) and variance ($\hat{\sigma}^\varphi$) to normalize the backscattering coefficients for an incident angle of $40^\circ$, as shown in Equation (5).

$$\sigma^\varphi_{norm,p,vc} = \sigma^\varphi_{ref,p,vc} + \hat{\sigma}^\varphi_{ref,p,vc} \left( \frac{\sigma^\varphi_{act,p,vc} - \sigma^\varphi_{act,p,vc}}{\sigma^\varphi_{act,p,vc}} \right)$$  \hspace{1cm} (5)
where $\sigma_{\text{norm}}$, $\sigma_{\text{ref}}$, and $\sigma_{\text{act}}$ represent the normalized, reference, and actual backscattering coefficients, respectively. In the subscript letters, $p$ represents the polarimetric mode, and $v_c$ represents the vegetation type. The backscattering coefficients are normalized in steps of 1° of incidence angle with the assistance of the vegetation classification map. In 2017, Wang H et al. [33] applied this approach to normalizing polarimetric decomposition parameters and successfully estimated SSM. Therefore this study used this method.

3.3. SSM Estimation

Although there is a physical relationship between SSM and surface scattering or double-bounce scattering, these physical models do not precisely coincide with the actual scattering. The model-based polarimetric decomposition decomposes scattering into several fixed scattering mechanisms without considering whether these scattering mechanisms are present in the actual situation. Meanwhile, it is almost impossible to estimate SSM in vegetated areas using only a polarimetric parameter or backscattering coefficient. Therefore, multiple polarimetric decomposition parameters are used to estimate SSM to consider the effects of various factors such as vegetation and surface roughness. QRF is used in this experiment, which is suitable for regression problems with high-dimensional predictor variables. Subsequently, a comparative experiment using CQR is performed to elucidate the validity of QRF.

3.3.1. Quantile Regression Forests (QRF)

Random forest regression (RFR) is an ensemble model that averages the outputs of multiple regression trees to approximate the conditional mean of a response variable. It has shown high potential for both vegetation phenology inversion [52] and SSM estimation [53]. Suppose there is a data set with training sample $(X_i, Y_i), i = 1, \ldots, n$, where the predictor variable $X$ is $p$-dimensional, denoted as $X \in \mathbb{R}^p$. For a decision tree $T(\theta)$ with parameter $\theta$ and $L$ leaf nodes, the prediction of the $T(\theta)$ for a new data point $X = x$ is determined by averaging the observations of the leaf nodes $l(x, \theta)$. For the training sample $X_i$, the weight vector $w_i(x, \theta)$ is shown in Equation (6).

$$w_i(x, \theta) = \frac{I(X_i \in \mathbb{R}(x, \theta))}{\sum_{j=1}^n I(X_j \in \mathbb{R}(x, \theta))}$$

where $I(\bullet)$ is the characteristic function, at this point, the prediction for a single tree is given by Equation (7).

$$\hat{u}_{\text{tree}}(x) = \sum_{i=1}^n w_i(x, \theta)Y_i$$

For a random forest with $k$ trees, the conditional mean $E(Y|X = x)$ is the average of the predictions of the $k$ trees, and its weight can be expressed as Equation (8).

$$w_i(x) = \frac{\sum_{t=1}^k w_i(x, \theta_t)}{k}$$

The predicted value of the random forest can then be calculated, as in Equation (9).

$$\hat{u}_{\text{forest}}(x) = \sum_{i=1}^n w_i(x)Y_i$$

In 2006, Meinshausen, N and Ridgeway, G [64] proposed QRF, which can reflect information about the complete conditional distribution of predictors, not just about the conditional mean. When predictor variable $X = x$ is given, the conditional distribution of the dependent variable $Y$ is given by Equation (10).
Following the description of RFR, the estimate of the conditional distribution can be written in the form of Equation (11).

\[
\hat{F}(y|X = x) = \frac{1}{n} \sum_{i=1}^{n} w_i(x) \cdot I(Y_i \leq y)
\]  

Further, the conditional quantile \( Q_\tau(Y) \) of \( Y \) can be estimated using sample data, as shown in Equation (12).

\[
Q_\tau(Y) = \inf \{ y : \hat{F}(y|X = x) \geq \tau \}
\]  

QRF is selected to present the uncertainty range of SSM estimation. Although the polarimetric decomposition parameters and the backscattering coefficients mentioned in Section 3.1 are used as input variables for the QRF, for each decision tree, different variables are randomly selected to train the model, which results in a low correlation between the decision trees. These properties also reduce the possibility of over-fitting in the QRF. In each category of crops, the SSM and SAR parameters measured in all agricultural fields are first randomly divided into two parts, 80% as the training dataset and 20% as the validation dataset. Subsequently, the training dataset is used to train the QRF model. Finally, the trained model is used to predict the SSM of the validation dataset, with the 0.5 quantiles as the predicted value and the 0.25 quantile and 0.75 quantiles as the upper and lower bounds of the prediction uncertainty interval. Simultaneously, CQR is used as a comparative experiment to demonstrate the effectiveness of QRF.

### 3.3.2. Copula Quantile Regression (CQR)

The copula function can model the marginal distribution of any two or more variables with a small number of parameters to establish a joint probability distribution between random variables [65]. Some scholars have demonstrated that the Archimedean copula function has the potential to estimate SSM [28,30]. This study compared QRF with Archimedean CQR to illustrate the effectiveness of QRF. The commonly used Archimedes copula functions are binary, so reducing the dimensionality of the prediction parameters composed of the polarimetric decomposition parameters and the backscattering coefficients is necessary. To maintain the correlation of SSM with the parameters after dimensionality reduction, supervised principal component analysis (SPCA) is used in this study. SPCA is a modified principal component analysis for regression analysis with many predictor variables, and the specific algorithm can be found in the literature of Bair, E et al. [66].

The copula function establishes the dependence between variables in two steps: estimating the marginal distribution of the variables and selecting the copula function. There are two commonly used methods for estimating sample distribution, parametric and non-parametric. Kernel density estimation (KDE) is a very effective nonparametric probability distribution method. This experiment uses the kernel density estimation method with the Gaussian kernel as the kernel function. Specific principles and methods can be found in references [28,30]. The commonly used Archimedean copula functions are Gumbel, Clayton, and Frank. The experiment selects the optimal copula function by comparing the theoretical and empirical copula distance. The Gumbel copula function is selected as the best fit for the data used in this study.

The copula function establishes the joint distribution between the SSM and the first principal component of the predictors reduced by SPCA. After the joint distribution is established, the conditional distribution of SSM under the condition of dimensionality-reduced predictors can be calculated by taking the partial derivative. The qth quantile of the SSM can be solved by the inverse function of the conditional distribution of the SSM and the inverse function of the probability distribution function of the SSM.
3.4. Evaluation Method

The estimation results are evaluated by considering four aspects, namely: estimation error, correlation, uncertainty range, and Goodness-of-fit. The estimation error of the proposed estimation method is evaluated by the mean relative error (MRE) and root-mean-square-error (RMSE). The Pearson’s correlation coefficient R is used to evaluate the correlation between the estimated SSM and the in situ SSM. The estimated uncertainty range is evaluated using the Relative Average Deviation (RDA). RDA mainly compares the difference between the mean of the boundary values of the prediction interval and the actual observed value to assess the uncertainty of the estimate [30]. In addition, the Kling-Gupta efficiency (KGE) is used to evaluate the Goodness-of-fit of the proposed SSM estimation method.

4. Results

4.1. Correlation Analysis

In vegetation-covered regions, soil surface scattering is often coupled with vegetation scattering, complicating the relationship between SSM and SAR backscattering coefficients. Figure 2 shows that the backscattering coefficients of most crops are very similar. In addition, the backscattering coefficient is insensitive to the variation of in situ SSM. This is because parameters such as the moisture content and height of the crops differ at different data acquisition times. The contribution of these crop parameters and SSM to the backscattering coefficients is coupled, so it is difficult to estimate the SSM based on the backscattering coefficients alone. Polarimetric decomposition can explain the scattering mechanism and separate the scattering of vegetation from the surface, which has excellent potential for SSM estimation of polarimetric data. The correlation coefficient matrices of in situ SSM and some commonly used polarimetric decomposition parameters under different crops are given in Figure 4. As shown in Figure 4, the correlation between the surface scattering component and in situ SSM is not the highest under all crops, so estimating the SSM in the experimental area with the first-order Bragg surface scatter results in more significant errors. The correlation between double-bounce scattering and in situ SSM is higher in corn, grassland, and cereal-covered areas, suggesting that dihedral angle scattering exists in these vegetation-covered areas. Typically, volume scattering is ignored in the estimation of SSM, but Figure 4 suggests a significant correlation between volume scattering and in situ SSM in areas covered by canola and cereal. In addition, $\alpha$, $A$, $H$, $DERD$, $rvi$, $Pr$, and $SERD$ all showed a relatively high correlation between in situ SSM under different crop covers.

Therefore, estimating SSM using only one or a few polarimetric decomposition parameters may result in significant uncertainties. A scheme that can take into account various polarimetric parameters is essential to estimate SSM accurately. QRF is a method that considers all polarimetric parameters and quantifies the importance of polarimetric parameters in SSM estimation.
Figure 4. Correlation coefficient matrices of in situ SSM and some commonly used polarimetric decomposition parameters under different crop coverage. (a) correlation coefficient matrix of soybeans; (b) correlation coefficient matrix of cereal; (c) correlation coefficient matrix of corn; (d) correlation coefficient matrix of canola; (e) correlation coefficient matrix of grassland.

4.2. SSM Estimation of Different Crop Coverage Areas

After analyzing the correlation between polarimetric parameters and SSM under different crops, QRF is used to estimate SSM, which is then compared with the performance of CQR. For each crop, four-fifths of the data are randomly selected to train the model and the remaining one-fifth to validate the model. The training and validation datasets of CQR are consistent with the QRF. The number of leaf nodes and decision trees of the random
forest is determined by the mean square error between the in situ and the estimated SSM values of the training data. Figure 5 compares the in situ SSM and the SSM estimated by the two methods in the validation dataset. In the figure, the expression of the error is: $\text{mv}_e - \text{mv}_i$, where $\text{mv}_e$ represents the estimated SSM and $\text{mv}_i$ represents the in situ SSM. As seen in the figures, all scatter points are distributed on both sides of the 1:1 straight line, indicating that both methods can potentially estimate SSM in these crop-covered areas. The histograms show that most of the estimated errors of SSM are between $-0.1 \text{ cm}^3/\text{cm}^3$ and $0.1 \text{ cm}^3/\text{cm}^3$, which shows that both estimation methods are robust. From Figure 5a,b, it can be concluded that the SSM values estimated by QRF in the area covered by soybean and cereal are significantly closer to the measured values than those estimated by CQR. Meanwhile, Figure 5b,d also illustrate that the error range of QRF estimation is smaller than that of CQR. These indicate that QRF is more advantageous than CQR for SSM estimation of soybean and cereal in the dataset used in this study. It is difficult to distinguish which method is more beneficial in the figures from the other three crops.

The evaluation parameters of the two methods for estimating SSM in areas covered by different crops are given in Table 4 to compare the two methods quantitatively. Table 4 suggests that the MRE estimated by the two methods is approximate in the area covered by five crops, which indicates that the estimated deviations of the two methods are similar. The use of QRF reduced the SSM estimated RMSE in the areas covered by crops soybean, cereal, and corn with 0.021 cm$^3$/cm$^3$, 0.033 cm$^3$/cm$^3$, and 0.006 cm$^3$/cm$^3$, respectively. RMSEs estimated by the two methods in canola and grassland-covered areas are not significantly different. Consistent with RMSE, the R and KEG parameters of QRF are considerably better than those of CQR in soybean, cereal, and corn-covered areas. In the canola-covered area, the R of QRF is higher than CQR by 0.025, but the KEG is 0.085 lower than CQR. Both R and KEG of CQR are better than those of QRF in the grassland areas. Combining Tables 2 and 4, we have found that with the increase in the number of field measurements, the SSM estimated by QRF gradually becomes better than CQR significantly better than CQR. This phenomenon may be because the random forest has little data to train each decision tree when the measurement data is small. The estimated value of SSM is not as accurate as that of the traditional copula method. Compared with the RMSE and R evaluated by Wang, H. et al. [33] with the same dataset in 2017 and 2019, the two methods of estimating SSM in this study have obtained more accurate results.

The SSM values for each in situ measurement and estimate accompanying the uncertainty range in the validation data set are given in Figure 6. The figures display that the dynamic range of the in situ SSM is more extensive than that of the estimated SSM for all the crop-covered areas. Most of the in situ SSM falls within the estimated uncertainty interval, and in situ SSM outside the uncertainty interval are usually peaks and valleys. These analyses illustrate the difficulty of both methods in predicting SSM for peaks and valleys, which is consistent with the conclusion of Nguyen, H.H. et al. in 2021 [30]. The evaluation parameters for the estimated uncertainty intervals are the RDA in Table 4, from which it is evident that moderate to good results are obtained for both methods in all crop-covered areas. Figure 6a–f suggest that QRF is more potential than CQR in capturing the peaks and valleys of in situ SSM in areas covered by three crops: soybean, cereal, and corn. While in the areas covered by the other two crops, no significant difference is found between the two methods.
Figure 5. Scatter plot and error histogram of the estimated and in situ SSM. (a) Scatter plot of soybeans; (b) Error histogram of soybeans; (c) Scatter plot of cereal; (d) Error histogram of cereal; (e) Scatter plot of corn; (f) Error histogram of corn; (g) Scatter plot of canola; (h) Error histogram of canola; (i) Scatter plot of grassland; (j) Error histogram of grassland.
Figure 6. Estimated and in situ SSM with the uncertainty intervals. (a) uncertainty intervals of QRF in soybeans; (b) uncertainty intervals of CQR in soybeans; (c) uncertainty intervals of QRF in cereal; (d) uncertainty intervals of CQR in cereal; (e) uncertainty intervals of QRF in corn; (f) uncertainty intervals of CQR in corn; (g) uncertainty intervals of QRF in canola; (h) uncertainty intervals of CQR in canola; (i) uncertainty intervals of QRF in grassland; (j) uncertainty intervals of CQR in grassland.
Table 4. Evaluation parameters of the error between estimated SSM and in situ SSM.

| Crop    | Method | MRE  | RMSE | R    | KGE  | RDA |
|---------|--------|------|------|------|------|-----|
| soybeans| QRF    | 0.101| 0.059| 0.745| 0.612| 0.107|
|         | CQR    | 0.109| 0.080| 0.481| 0.384| 0.121|
| cereal  | QRF    | 0.002| 0.053| 0.905| 0.732| 0.012|
|         | CQR    | −0.003| 0.086| 0.710| 0.491| 0.004|
| corn    | QRF    | 0.120| 0.037| 0.802| 0.577| 0.140|
|         | CQR    | 0.197| 0.043| 0.715| 0.458| 0.222|
| canola  | QRF    | 0.227| 0.079| 0.763| 0.288| 0.155|
|         | CQR    | 0.202| 0.075| 0.738| 0.373| 0.190|
| grassland| QRF   | 0.001| 0.037| 0.665| 0.431| −0.009|
|         | CQR    | −0.003| 0.036| 0.718| 0.488| −0.008|

4.3. Parameter Importance Assessment

In addition to the uncertainty interval of the estimate, the QRF can also evaluate the importance of the input parameters to the estimated result, making it possible to understand the contribution of different polarimetric decomposition parameters and backscattering coefficients to the SSM estimate. Figure 7 shows the ordered importance of polarimetric parameters and backscattering coefficients in soybean, cereal, canola, and grassland areas. These phenomena indicate that even for L-band SAR data, the scattering is strongly influenced by the vegetation structure, which is consistent with the conclusions of the correlation analysis in Figure 4. Wang H. et al. [52] found that the importance of different polarimetric parameters for estimating the phenology of varying vegetation varies greatly, which also confirms the influence of vegetation structure on polarimetric parameters.

For soybean, the backscattering coefficient of HH polarization is most important for SSM estimation. Volume scattering is most important in SSM estimates for cereal-covered areas and cannot be ignored in grassland and soybean-covered areas. The H parameter is the most important in SSM estimates for corn-covered areas, only to HH backscattering coefficient and Pr in soybean-covered areas. Surface scattering, often used to estimate SSM, occupies the second most important position in corn and canola-covered regions. The most advantageous parameter for SSM estimation in areas covered by grassland is the double-bounce scattering component, which has the potential to estimate SSM in all crop-covered areas except canola. In addition, in canola and grassland, the importance of some parameters became negative, which indicated that the participation of these parameters in SSM estimation would reduce the estimation accuracy. The main reason for the negative importance is that the out-of-bag importance evaluation method used by random forest is biased [67,68]. When predictors are correlated, randomly shuffled variables are replaced by correlated variables. The analysis of the importance of different parameters helps to select more relevant prediction parameters, simplifying the model in the subsequent development of the SSM estimation model. In summary, these parameters, which are essential in different crops, may have more significant advantages in subsequent SSM estimation studies.
Figure 7. Importance of the polarimetric parameters and backscattering coefficients for estimated SSM. (a) Importance of the parameters for estimated SSM in soybeans; (b) Importance of the parameters for estimated SSM in cereal; (c) Importance of the parameters for estimated SSM in corn; (d) Importance of the parameters for estimated SSM in canola; (e) Importance of the parameters for estimated SSM in grassland.
5. Discussion

The backscattering process in vegetation-covered areas is complex, so single-polarized SAR requires optical remote sensing or measured vegetation information to model vegetation scattering. Polarimetric SAR data can separate the contributions of vegetation scattering and surface scattering with the help of polarimetric decomposition, which makes it possible to estimate the SSM of vegetation-covered areas with SAR data alone. From the correlation coefficients in Figure 4 and the importance assessment in Figure 7, it can be found that volume scattering is not negligible in all crops, which also appeared in the study by Bai X. et al. [48]. This phenomenon may be because vegetation water content and growth status are related to SSM, which is the theoretical basis of some optical remote sensing estimation methods for SSM [7]. Meanwhile, the most significant disadvantage of model-based polarimetric decomposition is that the coherence matrix or covariance matrix is decomposed into definite scattering components, regardless of whether these components exist in the actual scattering. This disadvantage results in a weaker correlation between surface scattering and in situ SSM than volume or double-bounce scattering.

The parameters of the eigenvalue-based polarimetric decomposition can explain the scattering mechanism, which is used as auxiliary data in the SSM estimation. As in the study by Hajnsek, I. et al. [39], the H parameter and a parameter are highly correlated with the SSM of the bare surface, and the A parameter is highly correlated with the surface roughness. The study by Baghdadi, N. et al. [45] also showed that the a parameter has higher sensitivity in areas with high SSM, which can be used to correct the problem of SSM underestimation. Bai X. et al. [48] also found that the best results can be obtained when all the selected decomposition parameters are used for SSM estimation. From the error histogram in Figure 5, the error of the QRF estimate is about the same on both sides of 0, except for the cereal-covered areas. There is an underestimation of SSM in the cereal-covered areas, which may be caused by the inability of the model to capture the high SSM accurately. In other vegetation cover areas, there is no apparent overestimation or underestimation.

Comparing the evaluation parameters in Table 4 shows that the highest R and KEG are obtained for the QRF estimates in the areas covered by cereal. This result may be due to the small spacing between grains, which can be approximated as uniform coverage on the surface, so the correlation between backscattering and SSM is less affected by vegetation, as illustrated by the correlation coefficients in Figure 4. Grasslands, which have the same small plant spacing as cereals, are estimated with low bias by both methods. The possible reasons for the small R of the grassland are that the amount of data is too small to train an optimal model, or the variation range of in situ SSM in the validation dataset is too small. Despite the large spacing between corn plants, the density of corn is essentially the same in different fields in the experimental area, so a high R is also obtained. Unlike cereal, grassland and corn, there are differences in planting density in the areas covered by soybeans, with some producers using double seeding rows [54]. Figure 5a,b illustrate that the estimated SSM for the area covered by soybean has more points with more significant errors. This suggests that differences in planting methods and vegetation biomass impact SSM estimate significantly. Canola is also grown differently; the biomass varies more among fields than soybeans.

In addition to vegetation species and biomass, surface roughness is an important factor affecting SSM estimation. SMAPVEX12 measured the surface roughness of each field. Table 5 presents the statistical information of the surface roughness measured by SMAPVEX12 in areas covered by different vegetation to analyze the relationship between SSM estimation and surface roughness. Two parameters mainly measure surface roughness: the surface root mean square height (σ) and the correlation length (l). Table 5 illustrates that the surface roughness varies considerably in different fields. Typically, for wavelength λ, the roughness is scaled to ks and kl, where k = 2π/λ. Table 5 shows that the range of ks and kl are 0.09 ≥ ks ≤ 0.60 and 0.66 ≥ kl ≤ 6.73, which indicates that some fields do not satisfy the surface scattering model of Freeman-Durden decomposition [39]. This
analysis of Table 5 explains the low importance of surface scattering in SSM estimation from a roughness perspective. A comparison of the average roughness and the accuracy of SSM estimation for different crops suggests no significant relationship between them. This phenomenon indicates that the effect of vegetation on SSM estimation is more significant than the effect of roughness.

Table 5. Statistical parameters for the minimum (MIN), mean (MEAN), maximum (MAX) and standard deviation (STD) of the surface roughness parameters (SRP): surface root mean square height ($s$) and correlation length ($l$) (unit: cm).

| crop     | SRP | MIN | MEAN | MAX | STD |
|----------|-----|-----|------|-----|-----|
| soybeans | $s$ | 0.34| 0.88 | 1.89| 0.32|
|          | $l$ | 5.50| 11.88| 18.5| 3.58|
| cereal   | $s$ | 0.44| 1.14 | 2.26| 0.40|
|          | $l$ | 2.50| 11.03| 25.50| 5.02|
| corn     | $s$ | 0.60| 1.27 | 1.89| 0.44|
|          | $l$ | 6.50| 12.11| 22.50| 4.60|
| canola   | $s$ | 0.69| 1.22 | 1.71| 0.34|
|          | $l$ | 6.00| 9.96 | 15.50| 3.07|
| grassland| $s$ | 0.35| 1.07 | 1.76| 0.42|
|          | $l$ | 6.00| 14.00| 22.50| 6.21|

Figure 6 suggests that the method proposed in this study has a significant error in the capture of peaks and valleys, which has also appeared in the study of Nguyen, H.H. et al. [30]. Especially in the canola-covered areas, (g) and (h) of Figure 6 show that both methods are ineffective in capturing peaks and valleys, which may also be related to the differences in biomass and the structure. However, these peaks and valleys have important implications for predicting drought and flood. Typically, the peak in the time series occurs after rain or irrigation, and the valley appears long after rainfall. Therefore, heavy precipitation may have a greater impact on the accuracy of the method proposed in this experiment. Many scholars have explored the correlation between SSM and precipitation, and these studies show that SSM is an essential parameter for precipitation prediction, while precipitation affects SSM [69–71]. The effect of precipitation in SSM estimation is not considered because precipitation amount and time are not recorded in the data of this experiment. In the follow-up research, factors such as precipitation and the time difference between measurement and precipitation will be considered to improve the estimation accuracy of peak and valley values.

6. Conclusions

This study investigates SSM estimation methods for areas covered by different crops using L-band polarimetric parameters. QRF and CQR are used to estimate SSM from polarimetric parameters. The polarimetric parameters used in the study include the backscattering coefficients of the three polarizations and the parameters obtained from the model-based decomposition and the eigenvalue decomposition. The estimation results show that the QRF algorithm obtains a more accurate estimate with a minimum R of 0.665 and a maximum RMSE of 0.079 cm$^3$/cm$^3$. The reasonable and satisfactory results of this experiment indicate the potential of spaceborne polarimetric L-band SAR for SSM estimation in vegetation-covered areas, which means that SSM can be monitored periodically. However, there are still some problems. The algorithm proposed in this paper has unavoidable inaccuracy in capturing the peaks and valleys of SSM. In the subsequent study, the influence of precipitation and precipitation time will be considered in the algorithm of SSM estimation further to improve the capture ability of peak and valley values.
20. Oh, Y.; Sarabandi, K.; Ulaby, F.T. Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. IEEE Trans. Geosci. Remote Sens. 2002, 40, 1348–1355. [CrossRef]

21. Zhu, L.; Walker, J.P.; Ye, N.; Rüdiger, C. Roughness and vegetation change detection: A pre-processing for soil moisture retrieval from multi-temporal SAR imagery. Remote Sens. Environ. 2019, 225, 93–106. [CrossRef]

22. Attema, E.F.W.; Ulaby, F.T. Vegetation modeled as a water cloud. Radio Sci. 1978, 13, 357–364. [CrossRef]

23. 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]

24. Baghdadi, N.N.; El Hajj, M.; Zribi, M.; Fayad, I. Coupling SAR C-band and optical data for soil moisture and leaf area index retrieval over irrigated grasslands. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2015, 9, 1229–1243. [CrossRef]

25. 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]

26. Pathe, C.; Wagner, W.; Sabel, D.; Doubková, M.; Basara, J.B. Using ENVISAT ASAR global mode data for surface soil moisture retrieval over Oklahoma, USA. IEEE Trans. Geosci. Remote Sens. 2009, 47, 468–480. [CrossRef]

27. Kim, Y.; van Zyl, J.J. A time-series approach to estimate soil moisture using polarimetric radar data. IEEE Trans. Geosci. Remote Sens. 2009, 47, 2519–2527.

28. Pal, M.; Maity, R.; Suman, M.; Das, S.K.; Patel, P.; Srivastava, H.S. Satellite-based probabilistic assessment of soil moisture using C-band quad-polarized RISAT1 data. IEEE Trans. Geosci. Remote Sens. 2016, 55, 1351–1362. [CrossRef]

29. Ma, C.; Li, X.; Notarnicola, C.; Wang, S.; Wang, W. Uncertainty quantification of soil moisture estimations based on a Bayesian probabilistic inversion. IEEE Trans. Geosci. Remote Sens. 2017, 55, 3194–3207. [CrossRef]

30. Nguyen, H.H.; Cho, S.; Jeong, J.; Choi, M. A D-vine copula quantile regression approach for soil moisture retrieval from dual polarimetric SAR Sentinel-1 over vegetated terrains. Remote Sens. Environ. 2021, 255, 112283. [CrossRef]

31. 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]

32. Huang, X.; Wang, J.; Shang, J. An adaptive two-component model-based decomposition for soil moisture estimation for C-band RADARSAT-2 imagery over wheat fields at early growing stages. IEEE Geosci. Remote Sens. Lett. 2016, 13, 414–418. [CrossRef]

33. Wang, H.; Magagi, R.; Goïta, K. Comparison of different polarimetric decompositions for soil moisture retrieval over vegetation covered agricultural area. Remote Sens. Environ. 2017, 199, 120–136. [CrossRef]

34. Wang, H.; Magagi, R.; Goïta, 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]

35. Notarnicola, C.; Angiulli, M.; Posa, F. Soil moisture retrieval from remotely sensed data: Neural network approach versus Bayesian method. IEEE Trans. Geosci. Remote Sens. 2008, 46, 547–557. [CrossRef]

36. 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]

37. Santi, E.; Paloscia, S.; Pettinato, S.; Fontanello, G. Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. Int. J. Appl. Earth Obs. Geoinf. 2016, 48, 61–73. [CrossRef]

38. Kumar, P.; Prasad, R.; Choudhary, A.; Gupta, D.; Mishra, V.; Vishwakarma, A.; Singh, A.; Srivastava, P. Comprehensive evaluation of soil moisture retrieval models under different crop cover types using C-band synthetic aperture radar data. Geocarto Int. 2019, 34, 1022–1041. [CrossRef]

39. Hajnsek, I.; Pottier, E.; Cloud, S.R. Inversion of surface parameters from polarimetric SAR. IEEE Trans. Geosci. Remote Sens. 2003, 41, 727–744. [CrossRef]

40. He, L.; Panciera, R.; Tanase, M.A.; Walker, J.P.; Qin, Q. Soil moisture retrieval in agricultural fields using adaptive model-based polarimetric decomposition of SAR data. IEEE Trans. Geosci. Remote Sens. 2016, 54, 4445–4460. [CrossRef]

41. Jagdhuber, T.; Hajnsek, I.; Papathanassiou, K.P. An iterative generalized hybrid decomposition for soil moisture retrieval under vegetation cover using fully polarimetric SAR. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2014, 8, 3911–3922. [CrossRef]

42. Ponnurangam, G.G.; Jagdhuber, T.; Hajnsek, I.; Rao, Y. Soil moisture estimation using hybrid polarimetric SAR data of RISAT-1. IEEE Trans. Geosci. Remote Sens. 2015, 54, 2033–2049. [CrossRef]

43. Wang, H.; Magagi, R.; Goïta, K.; Jagdhuber, T. Refining a polarimetric decomposition of multi-angular UAVSAR time series for soil moisture retrieval over low and high vegetated agricultural fields. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2019, 12, 1431–1450. [CrossRef]

44. Gururaj, P.; Umesh, P.; Shetty, A. Assessment of surface soil moisture from ALOS PALSAR-2 in small-scale maize fields using polarimetric decomposition technique. Acta Geophys. 2021, 69, 579–588. [CrossRef]

45. Baghdadi, N.; Cresson, R.; Pottier, E.; Aubert, M.; Mehrez, M.; Jacome, A.; Benabdallah, S. A potential use for the C-band polarimetric SAR parameters to characterize the soil surface over bare agriculture fields. IEEE Trans. Geosci. Remote Sens. 2012, 50, 3844–3858. [CrossRef]

46. Bourgeau-Chavez, L.L.; Leblon, B.; Charbonneau, F.; Buckley, J.R. Evaluation of polarimetric Radarsat-2 SAR data for development of soil moisture retrieval algorithms over a chronosequence of black spruce boreal forests. Remote Sens. Environ. 2013, 132, 71–85. [CrossRef]
