Hybrid Methodology Using Sentinel-1/Sentinel-2 for Soil Moisture Estimation

Simon Nativel 1, Emna Ayari 1,2,†, Nemesio Rodriguez-Fernandez 1, Nicolas Baghdadi 3, Remi Madelon 1, Clement Albergel 4 and Mehrez Zribi 1,*

1 CESBIO, CNES/CNRS/INRAE/UPS, Université de Toulouse, 18 Av. Edouard Belin, Bpi 2801, CEDEX 9, 31401 Toulouse, France; simon.nativel@univ-tlse3.fr (S.N.); emna.ayari@inat.u-carthage.tn (E.A.); nemesio.rodriguez-fernandez@univ-tlse3.fr (N.R.-F.); remi.madelon@univ-tlse3.fr (R.M.)
2 National Agronomic Institute of Tunisia, Carthage University, Tunis 1082, Tunisia
3 CIRAD, CNRS, INRAE, TETIS, University of Montpellier, AgroParisTech, CEDEX 5, 34093 Montpellier, France; nicolas.baghdadi@teledetection.fr
4 European Space Agency Climate Office, ECSAT, Harwell Campus, Oxfordshire, Didcot OX11 0FD, UK; Clement.albergel@esa.int
* Correspondence: mehrez.zribi@ird.fr; Tel.: +33-56155-8505
† LR17AGR01 InteGRatEd Management of Natural Resources: remoTE Sensing, Spatial Analysis and Modeling (GREEN-TEAM).

Abstract: Soil moisture is an essential parameter for a better understanding of water processes in the soil–vegetation–atmosphere continuum. Satellite synthetic aperture radar (SAR) is well suited for monitoring water content at fine spatial resolutions on the order of 1 km or higher. Several methodologies are often considered in the inversion of SAR signals: machine learning techniques, such as neural networks, empirical models and change detection methods. In this study, we propose two hybrid methodologies by improving a change detection approach with vegetation consideration or by combining a change detection approach together with a neural network algorithm. The methodology is based on Sentinel-1 and Sentinel-2 data with the use of numerous metrics, including vertical–vertical (VV) and vertical–horizontal (VH) polarization radar signals, the classical change detection surface soil moisture (SSM) index $I_{SSM}$, radar incidence angle, normalized difference vegetation index (NDVI) optical index, and the VH/VV ratio. Those approaches are tested using in situ data from the ISMN (International Soil Moisture Network) with observations covering different climatic contexts. The results show an improvement in soil moisture estimations using the hybrid algorithms, in particular the change detection with the neural network one, for which the correlation increases by 54% and 33% with respect to that of the neural network or change detection alone, respectively.

Keywords: soil moisture; Sentinel-1; Sentinel-2; change detection; artificial neural network

1. Introduction

Soil moisture is a key parameter for understanding different processes related to the transfer of the soil–vegetation–atmosphere flux [1–3]. It is also an essential parameter in the management of water resources, particularly for optimizing irrigation [4,5]. In this context, remote sensing has greatly contributed to allowing the spatial and temporal monitoring of this parameter at different spatial scales from global to local [6,7].

Most of the currently available operational surface soil moisture products are on a global scale with spatial resolutions of several kilometers. They are essentially based on active and passive microwave measurements [8–12]. In passive microwaves, these are mainly products based on SMOS [8] and SMAP [9] missions dedicated to monitoring soil moisture with L-Band measurements and other non-dedicated sensors using higher frequency bands. In active microwaves, these are measurements based on acquisitions with...
a scatterometer, particularly data acquired by the ASCAT/METOP satellite series [13]. The European Space Agency (ESA) Climate Change Initiative (CCI) soil moisture project also provides long time series by merging soil moisture estimations from active and passive sensors [14].

For soil moisture estimation at high spatial resolution, we identify products with an average resolution at approximately 1 km or at the plot scale [15–24]. There have been various studies that have developed methodologies based on low-resolution data disaggregation techniques, notably with measurements acquired in thermal infrared (MODIS) [25] or, more recently, data acquired by SAR sensors. The Synthetic Aperture Radar (SAR) technique offers a high spatial resolution estimate of the radar signal adapted to applications at agricultural field scale. The measured signal is dependent on the radar configurations (frequency, incidence angle, and polarization) and the dielectric and geometric properties of the surface. After numerous demonstration space missions (ERS, ASAR/ENVISAT, RADARSAT, etc.), the arrival of Sentinel-1 constellation [26] in the context of the Copernicus program has enabled exponential growth in the use of these signals for monitoring soil moisture and the dynamics of the vegetation cover. Other soil moisture products are then offered only based on Sentinel-1 data, with three types of methodologies: one based on the direct inversion of physical or semiempirical models [27–29]; one based on the application of machine learning approaches and particularly neural networks [30–32]; and one based on the change detection technique [33–35]. For example, at plot scale, El Hajj et al. [31] presented an Artificial Neural Network (ANN) approach with training using the coupling of the Integral Equation Model (IEM) and the Water Cloud Model (WCM) to provide an estimate of soil moisture at the scale of the agricultural plot. Gao et al. [35] also proposed an approach at the plot scale with greater consideration of the vegetation cover and its effect on the temporal variation of the radar signal. Bauer-Marschallinger et al. [33] proposed a change detection approach very close to the initial approach proposed with data from ASCAT scatterometers [13] at a 1 km scale.

For these products, which are highly useful for regional hydrology, the validation of existing products, despite the very interesting potential, still shows some limitations in different contexts, in particular that of dense vegetation covers but also in relatively complex contexts with strong heterogeneities in terms of land use and topography [36].

In this context, this study proposes to test hybrid approaches to soil moisture retrieval at a 1 km scale with the objective of improving the estimation accuracy of soil moisture. The approaches consider hybrid methodologies with a combination of a change detection approach with empirical modeling or machine learning.

Section 2 presents in the first subsection the database used in this study in terms of soil moisture data and satellite measurements. The second subsection presents the methodologies tested and proposed in this study. Section 3 illustrates the results. Section 4 includes the discussion of the proposed applications. The conclusions are presented in Section 5.

2. Materials and Methods

2.1. Database

2.1.1. ISMN Soil Moisture Data

The training and validation of the proposed methods are conducted based on data from the International Soil Moisture Network (ISMN) [37]. The data are available in conjunction with additional datasets of Koppen–Geiger climate classes, ESA’s CCI land cover, and soil characteristics. The upper soil layer (0–10 cm) moisture measurements are harmonized as fractional volumetric soil moisture ($m^3/m^3$) and converted into Coordinated Universal Time (UTC). After data quality verification, some ISMN networks suffer from a lack of measurements. Therefore, we considered 21 networks among a total of 71 spatially distributed as shown in Figure 1. The data of each station should cover a period of two years with at least 20 dates between the start and the end date of acquisitions—1 January 2015 and 19 August 2021, respectively. Consequently, in the same network, we retain only stations with valid dataset as detailed in Table 1.
of two years with at least 20 dates between the start and the end date of acquisitions—1 January 2015 and 19 August 2021, respectively. Consequently, in the same network, we retain only stations with valid dataset as detailed in Table 1.

Figure 1. The global distribution of the International Soil Moisture Network (ISMN).

Table 1. Overview of the considered ISMN networks.

<table>
<thead>
<tr>
<th>Network</th>
<th>Country</th>
<th>Number of Selected Stations</th>
<th>SM Sensors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMMA-CATCH</td>
<td>Benin, Niger</td>
<td>7</td>
<td>CS616</td>
<td>Cappelaere et al. [38]; De Rosnay et al. [39]; Lebel et al. [40]; Mougin et al. [41]; Pellarin et al. [42]; Galle et al. [43].</td>
</tr>
<tr>
<td>BIEBRZA-S-1</td>
<td>Poland</td>
<td>8</td>
<td>GS-3</td>
<td>Musial et al. 2016 [44]</td>
</tr>
<tr>
<td>COSMOS</td>
<td>USA</td>
<td>2</td>
<td>Cosmic-ray-Probe</td>
<td>Zreda et al. [45]; Zreda et al. [46]</td>
</tr>
<tr>
<td>HOBE</td>
<td>Denmark</td>
<td>3</td>
<td>Decagon-STE</td>
<td>Bircher et al. [47]; Jensen et al. [48]</td>
</tr>
<tr>
<td>FLUXNET-AMERIFLUX</td>
<td>USA</td>
<td>4</td>
<td>CS655, ThetaProbe-ML3, ThetaProbe-ML2X</td>
<td></td>
</tr>
<tr>
<td>FR-Aqui</td>
<td>France</td>
<td>3</td>
<td>ThetaProbe ML2X</td>
<td>Al-Yaari et al. [49]; Wigneron et al. [50]</td>
</tr>
<tr>
<td>GROW</td>
<td>UK</td>
<td>20</td>
<td>Flower-Power</td>
<td>Zappa et al. [51]; Xaver et al. [52]; Zappa et al. 2020 [53]</td>
</tr>
<tr>
<td>HOAL</td>
<td>Austria</td>
<td>32</td>
<td>SPADE-Time-Domain-Transmissivity</td>
<td>Vreugdenhil M. et al. [54]; Blöschl, Günter, et al. [55]</td>
</tr>
<tr>
<td>IPE</td>
<td>Spain</td>
<td>2</td>
<td>CS655, ThetaProbe-ML2X</td>
<td>Alday et al. [56]</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Network</th>
<th>Country</th>
<th>Number of Selected Stations</th>
<th>SM Sensors</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAQU</td>
<td>China</td>
<td>1</td>
<td>ECH20-EC-TM</td>
<td>Su et al. [57]; Dente et al. [58]</td>
</tr>
<tr>
<td>MOL-RAO</td>
<td>Germany</td>
<td>1</td>
<td>TRIME-EZ</td>
<td>Beyrich F. and Adam W.K. [59]</td>
</tr>
<tr>
<td>NAQU</td>
<td>China</td>
<td>5</td>
<td>5TM</td>
<td>Su et al. 2011 [60]</td>
</tr>
<tr>
<td>REMEDHUS</td>
<td>Spain</td>
<td>13</td>
<td>Stevens-Hydra-Probe</td>
<td>Gonzalez-Zamora et al. [61]</td>
</tr>
<tr>
<td>RISMA</td>
<td>Canada</td>
<td>5</td>
<td>Hydraprobe-II-Sdi-12</td>
<td>Canisius F. [62]; L'Heureux J. [63]; Ojo et al. [64]</td>
</tr>
<tr>
<td>RSMN</td>
<td>Romania</td>
<td>19</td>
<td>5TM</td>
<td></td>
</tr>
<tr>
<td>SCAN</td>
<td>USA</td>
<td>130</td>
<td>Hydraprobe-Sdi-12/Ana</td>
<td>Schaefer et al. [65]</td>
</tr>
<tr>
<td>SMOSMANIA</td>
<td>France</td>
<td>15</td>
<td>ThetaProbe ML2X</td>
<td>Calvet et al. [66]; Albergel et al. [67]; Calvet et al. [68]</td>
</tr>
<tr>
<td>SNOTEL</td>
<td>USA</td>
<td>84</td>
<td>Hydraprobe-Analog (2.5-Volt)</td>
<td>Leavesley et al. [69]</td>
</tr>
<tr>
<td>TAHMO</td>
<td>Ghana</td>
<td>3</td>
<td>TEROS10, TEROS12</td>
<td></td>
</tr>
<tr>
<td>TERENO</td>
<td>Poland</td>
<td>4</td>
<td>Hydraprobe-II-Sdi-12</td>
<td>Zacharias et al. [70]; Bogena et al. [71]; Bogena et al. [72]</td>
</tr>
<tr>
<td>USCRN</td>
<td>USA</td>
<td>77</td>
<td>Stevens-Hydraprobe-II-Sdi-12</td>
<td>Bell et al. [73]</td>
</tr>
</tbody>
</table>

2.1.2. Sentinel-1

The first S-1A satellite was launched on 3 April 2014 and was followed by the S-1B Sentinel satellite on 25 April 2016. This dual-satellite constellation offers a 6-day repeat frequency for all regions of the globe [74]. The SAR payloads use a C-band frequency of 5.4 GHz and have the following standard operating modes: stripe map (SM), interferometric wide swath (IW), extra wide swath (EW), and WaVe (WV). In the present study, IW S-1 images are analyzed. They are characterized by a 10 m × 10 m spatial resolution and dual VV and VH polarization measurements. All of the images were generated from the high-resolution, Level-1 Ground Range Detected (GRD) product. The calibration is designed to convert the digital values of the raw images into backscattering coefficients (σ0).

2.1.3. Sentinel-2

After the launch of Sentinel-2 A and B on 23 June 2015 and 7 March 2017, respectively, optical data became free and open access with a spatial resolution varying between 10 m × 10 m and 60 m × 60 m and a revisit time of up to 5 days in 13 spectral bands at visible and mid-infrared wavelengths. In the present study, we used Sentinel-2 surface reflectance products downloaded from the Theia site (https://www.theia-land.fr/, accessed on 16 May 2022), already orthorectified and atmospherically corrected with a mask of clouds and shadows owing to the MAJA algorithm [75]. On each acquisition date and using red visible and near infrared bands with center wavelengths of approximately 665 and 833 nm, respectively, we calculated the Normalized Difference Vegetation Index (NDVI) and averaged this index for each studied station as expressed in the following equation:

\[ NDVI = \frac{R_{NIR} - R_{Red}}{R_{NIR} + R_{Red}} \] (1)
where $R_{NIR}$ and $R_{Red}$ are the surface reflectance in the two bands, near infrared and red visible, respectively.

2.1.4. Satellite Data Processing

Both radar backscattering coefficients and NDVI time series are identified at each station. A temporal linear interpolation of NDVI data is proposed to estimate the NDVI at each radar acquisition date. In this averaging, a filter is applied to the optical pixels to only consider data between 0.15 and 0.8 of NDVI to avoid urban areas and water covers with low NDVI or strong NDVI corresponding mainly to dense forests.

The radar signal is averaged over a radius of 500 m around each station. For a given station, if more than 50% of the Sentinel-1 pixels are excluded, the processing of radar data is not considered for the analyzed data.

2.2. Methodology

2.2.1. Change Detection Algorithm

The classic change detection SSM index $I_{SSM}$ is defined as [76]:

$$ I_{SSM} = \frac{SSM_t - SSM_{min}}{SSM_{max} - SSM_{min}} = \frac{\sigma_{VV} - \sigma_{VVmin}}{\sigma_{VVmax} - \sigma_{VVmin}} $$

where $SSM_t$ is the soil moisture content at time $t$; $SSM_{min}$ and $SSM_{max}$ are the minimum and maximum values of in situ soil moisture, respectively; $\sigma_{VV}$ is the radar signal at time $t$; and $\sigma_{VVmin}$ and $\sigma_{VVmax}$ are the minimum and maximum values of the radar signal time series, respectively. An index equal to 1 corresponds to the wettest context, and an index equal to 0 corresponds to the driest context.

To convert this index $I_{SSM}$ to volumetric soil moisture at time $t$ $SSM_t$, we introduce [77]:

$$ SSM_t = I_{SSM} \times (SSM_{max} - SSM_{min}) + SSM_{min} $$

2.2.2. Improved Change Detection Approach

For the classic detection approach, radar signal change is linked to soil moisture change. It can be written as:

$$ \Delta VV = \alpha \Delta SSM $$

where the soil moisture changes and the radar signal change in VV polarization are expressed in Equations (5) and (6), respectively.

$$ \Delta SSM = SSM_t - SSM_{min} $$

$$ \Delta VV = \sigma_{VV} - \sigma_{VVmin} $$

This relationship is adapted from [35,78]. It considers as a hypothesis that the difference between two radar signals acquired on two different dates is mainly related to the change in the hydric state of the soil.

Here, we propose an improved change detection methodology by using a hybrid change detection and empirical approach in which the effect of the vegetation is taken into account thanks to a vegetation-related variable $V1$. Using this approach, the radar signal change is related to the soil moisture change by the following expression:

$$ \Delta VV = (\alpha - \beta V1) \Delta SSM $$

Unlike forward modeling approaches such as the WCM, the radar signal and the soil moisture are introduced as the difference between the radar signal at time $t$ and the minimum signal corresponding to the minimum moisture and the difference between the soil moisture and the minimum moisture value, respectively.
The main objective of introducing the change as a function of time is to reduce the dependency to other variables affecting the radar signal such as soil roughness, which can be very important, particularly in the context of strong topography or even important spatial changes in microtopography, that change little with time for a given site, in contrast to soil moisture. Two vegetation-related quantities were tested for the V1 parameter: the optical vegetation index NDVI estimated from Sentinel-2 data, as illustrated in Section 2.3, and the VH/VV ratio, considered to be strongly linked to the dynamics of the vegetation cover. This second option could be particularly interesting in the context of a humid climate with limited optical data.

### 2.2.3. Artificial Neural Network Hybrid Approach

The multilayer perceptron (MLP), which is a multilayer feed-forward ANN, is one of the most widely used ANNs, mainly in the field of water resources [79,80]. A multilayer perceptron has one or more hidden layers between its input and output layers. The neurons are organized in layers such that neurons of the same layer are not interconnected and that the connections are directed from lower to upper layers. Each neuron returns an output based on a weighted sum of all inputs and according to a nonlinear function called the transfer or activation function. The input layer, made up of different metrics from Sentinel-1 and Sentinel-2 data, is connected to the hidden layer(s), which is made up of hidden neurons. The final estimates of the ANN are given by an activation function associated with the final layer called the output layer, using a sum of the weighted outputs of the hidden neurons.

The ANN model architecture consists of three hidden layers of 20 neurons with a rectified linear function (ReLu) as activation functions and an output layer with a single neuron with a linear activation function. The mean square error was used as the loss function and the gradient backpropagation was carried out using a first order stochastic gradient-based optimizer (Adam).

Different predictors based on Sentinel-1 and Sentinel-2 were tested to estimate soil moisture: VV, VH, incidence angle, VH/VV, NDVI, and I_{SSM}.

1. The VV and VH signals are identified for their high sensitivity to soil moisture.
2. The classical change detection SSM index I_{SSM} is calculated as a function of radar backscattering coefficients in VV polarization to use it for soil moisture estimation.
3. The incidence angle has an effect on the contribution of soil and vegetation components on the radar signal.
4. The NDVI index is identified to take into account the effect of vegetation cover on the backscattering signal.
5. The VH/VV ratio is identified to take into account the effect of vegetation cover on the backscattering signal [81].
6. SSM_t estimated from the classic change detection approach described in Section 3.1, Equation (2) is also considered as input.

The ANN models were trained using in situ soil moisture measurements retrieved from the ISMN as target. The training of the ANN models was conducted using 70% of the data samples. Thirty percent were kept for validation.

### 2.3. Statistical Parameters for Accuracy Assessment

Datasets are randomly subdivided into two parts: 70% of the database for model calibration and 30% for validation. The training data are used to calculate the different parameters to be estimated in the empirical and semiempirical models.

The Bias, root mean square error (RMSE) and Pearson’s correlation (R) are considered to estimate the precision of the models.

\[ \text{Bias} = \frac{P_{\text{estimated}} - P_{\text{measured}}}{P_{\text{measured}}} \] (8)
\[ \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (p_{\text{estimated}}^i - p_{\text{measured}}^i)^2} \]  
(9)

where \( N \) is the number of data samples, \( p_{\text{estimated}}^i \) is the estimated value of sample \( i \), and \( p_{\text{measured}}^i \) is the measured value of sample \( i \).

\[ R = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2}}} \]  
(10)

where \( x_i \) and \( y_i \) are individual samples taken at points indexed with the variable \( i \).

### 3. Results

#### 3.1. Improved Change Detection Approach

The empirical improved change detection approach has a double objective, taking into account the effect of vegetation and limiting the effect of surface geometry. The calibration of \( \alpha \) and \( \beta \) parameters were conducted by using 70% of the dataset selected randomly (23869 samples). For the validation, the remaining 10,229 samples of the dataset were used.

Figure 2 illustrate the validations of the different algorithms described in Section 3.2 tested with ISMN data. The proposed results show an improvement in accuracy by considering the effect of vegetation cover in the tested relationships. The \( \text{RMSE} \) (\( R \)) values decrease (increase) from 0.074 m³/m³ (0.58) from the change detection approach (Equation (2), Figure 2a) to 0.073 m³/m³ (0.59) for the improved change detection approach (Equation (7)) when the VH/VV ratio is used as the \( V1 \) parameter and to 0.068 m³/m³ (0.63) when NDVI is used as the \( V1 \) (Figure 2b).

![Figure 2](image)

**Figure 2.** Scatterplots of the retrieved surface soil moisture (SSM) as a function of in situ SSM measurements colored according to NDVI value variation using two change detection approaches: (a) classic approach and (b) new approach expressed in Equation (7), where \( V1 \) is the NDVI.

#### 3.2. Neural Network Hybrid Approach

The different combinations of input metrics are tested to estimate soil moisture. Figure 3 illustrates the results of validations applied for 30% of the database, for each case of combination with the statistical parameters \( \text{RMSE} \) and \( R \).
3.1. Improved Change Detection Approach

The empirical improved change detection approach has a double objective, taking into account the effect of vegetation and limiting the effect of surface geometry. The calibration of (Equation (2)) as input to ANN allows a strong improvement in the accuracy of soil moisture estimates for the optimal case for different tested networks, where we represent the hybrid approach using multiple combinations of features: (a) VV, NDVI, the incidence angle, (b) VV, NDVI, the incidence angle, $I_{SSM}$, (c) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, VH/VV ratio, (d) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt, (e) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt, (f) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt, VH/VV ratio, (g) VV, NDVI, the incidence angle, $\Delta I_{VV}$, SSMt, VH/VV ratio, (h) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt, VH/VV ratio, (i) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt.

Figure 3. Scatterplots of the retrieved SSM as a function of in situ SSM measurements using the ANN approach using multiple combinations of features: (a) VV, NDVI, (b) VV, NDVI, the incidence angle, (c) NDVI, the incidence angle, $I_{SSM}$, (d) VV, NDVI, the incidence angle, $I_{SSM}$, (e) VV, NDVI, the incidence angle, $\Delta I_{VV}$, (f) VV, NDVI, the incidence angle, $\Delta I_{VV}$, (g) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, VH/VV ratio, (h) VV, NDVI, the incidence angle, $\Delta I_{VV}$, SSMt, (i) VV, NDVI, the incidence angle, $\Delta I_{VV}$, $I_{SSM}$, SSMt.
For the first six predictor combinations (Figure 3a–f), we observe relatively close precision with RMSE values in the range of 0.095 m$^3$/m$^3$ and 0.083 m$^3$/m$^3$ and $R$ of 0.3–0.6. The introduction of moisture estimated by the classic change detection algorithm (Equation (3)) as input to ANN allows a strong improvement in the accuracy of soil moisture estimation with an RMSE equal to 0.063 m$^3$/m$^3$ and $R = 0.76$ when we consider the predictors: VV, NDVI, the incidence angle, $\Delta V_V$, and SSM. By adding the VH/VV ratio and $I_{SSM}$, the RMSE value decreases to 0.062 m$^3$/m$^3$, and the correlation coefficient reaches a value of approximately 0.79. This result confirms the contribution of the hybrid approach to estimating soil moisture. This first estimated soil moisture strongly contributes to a better estimate of soil moisture by the ANN.

Figure 4 illustrates the accuracy of intercomparisons between in situ measurements and satellite estimates for the optimal case for different tested networks, where we represent the RMSE and $R$ parameters by blue and orange boxes. The RMSE values vary from 0.03 m$^3$/m$^3$ to 0.09 m$^3$/m$^3$, and $R$-values fluctuate between 0.37 and 0.84.

![Boxplot of statistical parameters (R and RMSE) of soil moisture retrieval as a function of ISMN-considered networks using the hybrid methodology of change detection and ANN.](image)

Good consistency is generally observed for networks such as AMMA-CATCH, COSMOS, MAQU, RSMN, HOAL, HOBE, IPE, BIEBRZA S-1, TAHMO, REMEDHUS and RISMA, and RMSE values are under or equal to 0.05 m$^3$/m$^3$. The NAQU network is characterized by the lowest RMSE value of 0.03 m$^3$/m$^3$ and $R$ value of 0.77.

Within the same soil moisture in situ network, the accuracy of soil moisture retrieval varies from one station to another. For the REMEHDUS case characterized by an RMSE equal to 0.05 m$^3$/m$^3$, RMSE values per station range between 0.03 m$^3$/m$^3$ and 0.09 m$^3$/m$^3$, and $R$ values vary between 0.34 and 0.69, as represented in Figure 5.
Figure 5. Scatterplots of the estimated soil moisture as a function of ISMN measurements in the REMEDHUS network per considered station.

4. Discussion

The proposed hybrid approaches have allowed more or less strong improvements compared to the initial estimates based on change detection or a separate ANN approach. With an improved change detection method, we observe a negligible contribution of the considered vegetation cover compared to a basic approach directly linking the radar signal to soil moisture. This can be explained by the highly diversified context at the scale of many soil moisture stations with very varied landscape contexts (crops, trees, bare soils, etc.) and different vegetation densities, which can generate significant noise in the modeling of the scattered signal that is difficult to take into account without a more precise description in terms of land use. This noise is particularly observed with the VH/VV index, which is very sensitive to the dynamics of the vegetation cover in a homogeneous context [81], but it could also mix different effects and particularly those of soil roughness [82].

To better analyze proposed results, we examined the time series of the in situ and retrieved soil moisture per station and network. Figure 6 displays the time series of the radar signal (VV), NDVI, and soil moisture $SSM_t$. The in situ soil moisture measurements are illustrated in blue, and the hybrid approach results are drawn in red. The intercomparison between the proposed approach performance within the LasBodegas and Canizal stations reveals RMSE values of 0.04 m$^3$/m$^3$ and 0.07 m$^3$/m$^3$, respectively. The two stations belong to the same climatic region of the arid steppe and characterize a clay fraction interval of approximately 35%. The performance difference may be induced by the land cover, where the Canizal station is occupied by shrubs and the LasBodegas station is covered by trees. The aforementioned land cover may impact the accuracy of soil moisture retrieval due to the vegetation volume impact on the radar signal in the C-band. Additionally, the measured soil moisture values are lower than 0.3 m$^3$/m$^3$ at the LasBodegas station, and higher values reach 0.4 m$^3$/m$^3$ at the Canizal station. Hence, the soil water content retrieval is more accurate in the first case due to the saturation of the C-band signal at high values of soil moisture.
water content retrieval is more accurate in the first case due to the saturation of the C-band signal at high values of soil moisture.

Figure 6. Scatterplots of the temporal evolution of radar signals in VV polarization, NDVI, and the predicted and in situ measurements of soil moisture using the hybrid methods within two stations of the REMEHDUS network: (a) LasBodegas station, (b) Canizal station.

However, the approach has difficulties for certain stations, as shown for FLUXNET-AMERFLUX, GROW, SNOTAL, and SMOSMANIA networks, where the RMSE values reach a maximum of $0.09 \text{ m}^3/\text{m}^3$. The analysis of these cases generally leads to contexts of dense vegetation cover that can induce a low sensitivity of the radar signal to soil moisture.

In Figure 7, we scatterplot the statistical parameters as a function of NDVI values. According to Figure 7a,b, we observe the increase of RMSE and Bias values as a function of the increase of NDVI values where RMSE can reach $0.10 \text{ m}^3/\text{m}^3$. The vegetation development may induce a Bias between $-0.06 \text{ m}^3/\text{m}^3$ and $0.04 \text{ m}^3/\text{m}^3$, where NDVI values exceed 0.5. This behavior may be explained by the C-band potential which is otherwise limited in dense canopies where NDVI values are higher than 0.5.
This is difficult to take into account in a general approach based on a neural network trained on stations with different types of surface conditions, such as the case of the FLUXNET-AMERFLUX network. The station land covers are a mixture of grasslands, temporary crops followed by harvest and bare soil periods, and woody savanna characterized by forest canopy cover between 30% and 60% and vegetation height exceeding 2 m. In this land cover context, the vegetation volume impacts the radar signal and complicates the soil moisture retrieval. We observe the vegetation impact within many stations in the SMOSMANIA network, such as the Mazan-Abbaye, Cabriers Avignon, and Ville Vieille stations occupied by trees or shrubs.

Furthermore, the use of NDVI as a vegetation descriptor may induce other limits, such as the availability of data in regions with temperate climates. The presence of clouds contaminates the surface reflectance, which damages the radiometric information. As a result, many time series suffer from gaps and lack data, which complicates the training and validation of the proposed model, such as the case of some stations of the USCRN network, where the mean RMSE value is equal to 0.06 m³/m³.

By considering the GROW network data, the RMSE is equal to 0.07 m³/m³. This relatively low accuracy in retrieving soil moisture may be linked to the predominant cold climate of the considered stations. This low-temperature climate may impact the radar signal, especially with the freeze–thaw phenomenon. This change in the physical state of the soil water content generates a fast variability in the Sentinel-1 signal, as discussed for agricultural plots in metropolitan France by Baghdadi et al. [83] and Fayad et al. [84].

5. Conclusions

Different approaches have been proposed for SSM estimation from space. The goal is to improve estimates by combining change detection logic with empirical or other approaches based on an ANN. The study is based on Sentinel-1 and Sentinel-2 data tested on the ISMN moisture network.

Relationships between temporal changes in radar signals and temporal changes in soil moisture are tested. Improved change detection relationships combine these effects with the contribution of vegetation through two optical and radar indices (NDVI and VH/VV ratio). The integration of the effect of vegetation slightly improves the precision with an RMSE that decreases slightly from 0.074 m³/m³ to 0.073 m³/m³ and 0.068 m³/m³ for VH/VV and NDVI, respectively.

Testing an ANN approach through numerous metrics based on radar and optical (VV, VH, VH/VV, NDVI, ΔV, incidence angle, etc.) time series illustrates precision within a 0.08 m³/m³–0.09 (m³/m³) range. These results are greatly improved with the integration of vegetation.
as input of soil moisture estimated from the change detection approach. Thus, we move on to precision below the bar of 0.07 m$^3$/m$^2$ for the different possible combinations of metrics. Thus, it seems highly useful to propose this combination to improve the precision of the estimated soil moisture. Despite this improvement, there are some limitations at some stations, particularly related to the vegetation density and presence of forests or extreme climates with cold conditions. In the future, it would be very useful to propose a spatialization of this approach by considering auxiliary information of soil properties and land use for a better application of the proposed algorithms and improvement of proposed precision. In fact, this allows us to distinguish effects due more precisely to vegetation for which volume and attenuation scattering are different from one cover to another. For a high-resolution scale, this aspect, which is generally not considered for a low-resolution scale, seems important.

**Author Contributions:** S.N., E.A. and M.Z. developed methods and analyzed the data; all authors contributed to the materials/analysis tools; and M.Z. and E.A. wrote the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by Projects; ESA No. 4000126684/19/I-NB “ESA CCI+” and TAPAS TOSCA/CNES.

**Data Availability Statement:** The data presented in this study are available in International Soil Moisture Network (ISMN) [https://ismn.geo.tuwien.ac.at/en/](https://ismn.geo.tuwien.ac.at/en/) and [https://scihub.copernicus.eu/sites](https://scihub.copernicus.eu/sites) (accessed on 16 May 2022).

**Acknowledgments:** The authors thank the International Soil Moisture Network (ISMN) and the supporting networks for the availability of soil moisture data.

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

**References**


Remote Sens. 2022, 14, 2434


22. Ma, C.; Li, X.; McCabe, M.F. Retrieval of High-Resolution Soil Moisture through Combination of Sentinel-1 and Sentinel-2 Data. *Remote Sens.* 2020, 12, 2303. [CrossRef]


47. Jensen, K.H.; Refsgaard, J.C. HOBE: The Danish Hydrological Observatory. *Vadose Zone J.* 2018, 17, 1–24. [CrossRef]
Remote Sens. 2022, 14, 2434


68. Calvet, J.-C.; Fritz, N.; Berne, C.; Piguet, B.; Maurel, W.; Meurey, C. Deriving pedotransfer functions for soil quartz fraction in southern franc from reverse modeling. Soil 2016, 2, 615–629. [CrossRef]


72. Bogaen, H.R. Tereno: German network of terrestrial environmental observatories. J. Large-Scale Res. Facil. JLSRF 2016, 2, 52. [CrossRef]


