Article

A Novel Scheme for Merging Active and Passive Satellite Soil Moisture Retrievals Based on Maximizing the Signal to Noise Ratio

B. G. Mousa [1,2,3], **Hong Shu** [1,2,*], **Mohamed Freeshah** [1,4] **and Aqil Tariq** [1]

[1]   State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,
      Wuhan University, Wuhan 430079, China; dr.bahaa1@azhar.edu.eg (B.G.M.);
      mohamedfreeshah@whu.edu.cn (M.F.); aqiltariq@whu.edu.cn (A.T.)
[2]   Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China
[3]   Faculty of Engineering, Al-Azhar University, Cairo 11884, Egypt
[4]   Surveying Engineering Department, Faculty of Engineering at Shoubra, Benha University,
      Benha 13518, Egypt
*   Correspondence: shu_hong@whu.edu.cn

check for updates

Abstract: In this research, we developed and evaluated a new scheme for merging soil moisture (SM) retrievals from both passive and active microwave satellite estimates, based on maximized signal-to-noise ratios, in order to produce improved SM products using least-squares theory. The fractional mean-squared-error (fMSE) derived from the triple collocation method (TCM) was used for this purpose. The proposed scheme was applied by using a threshold between signal and noise at fMSE equal to 0.5 to maintain the high-quality SM observations. In the regions where TCM is unreliable, we propose four scenarios based on the determinations of correlations between all three SM products of TCM at significance levels (i.e., p-values). The proposed scheme was applied to combine SM retrievals from Soil Moisture Active Passive (SMAP), Advanced Scatterometer (ASCAT), and Advanced Microwave Scanning Radiometer 2 (AMSR2) to produce SMAP+ASCAT and AMSR2+ASCAT SM datasets at a global scale for the period from June 2015 to December 2017. The merged SM dataset performance was assessed against SM data from ground measurements of international soil moisture network (ISMN), Global Land Data Assimilation System-Noah (GLDAS-Noah) and ERA5. The results show that the two merged SM datasets showed significant improvement over their parent products in the high average temporal correlation coefficients (R) and the lowest root mean squared difference (RMSE), compared with in-situ measurements over different networks of ISMN. Moreover, these datasets outperformed their parent products over different land cover types in most regions of the world, with a high overall average temporal R and the lowest overall average RMSE value with GLDAS and ERA5. In addition, the suggested scenarios improved SM performance in the regions with unreliable TCMs.

Keywords: soil moisture; merged datasets; the triple collocation; least-squares theory; remote sensing

1. Introduction

Soil moisture (SM) is a critical factor in atmosphere-land interactions, energy fluxes, water fluxes control, and carbon exchange. Therefore, an understanding of SM dynamics informs applications in agricultural research, hydrology, drought monitoring, and climate change monitoring [1–4]. The most widely used methods to obtain SM data are ground measurements, remotely sensed retrieval data, and values extracted from surface models [5,6].

In satellites, there are two common methods for retrieving SM, including active and passive microwave instruments working at frequencies ranging from 1 to 10 GHz [7]. In the last two decades, Active and passive microwave satellites have been widely used to retrieve surface SM variations with reliable performance at global and regional scales, and are sensitive to SM variations over different land covers. Passive satellites include Advanced Microwave Scanning Radiometer 2 (AMSR2), Soil Moisture and Ocean Salinity (SMOS), and Soil Moisture Active Passive (SMAP). The Advanced Scatterometer (ASCAT) is considered an active satellite, and is loaded on various platforms including the Metop-A, Metop -B, and Metop-C satellites [8–10].

In previous research, the passive and active microwave SM datasets were evaluated and validated against in-situ measurements, against land surface models, and by applying the triple collocation method (TCM) that accounts for different land covers. The results indicate that over sparsely and moderate vegetated regions, the passive microwave datasets showed higher performance on the reference datasets than the active products. Over densely vegetated areas, the active microwave SM product achieved the highest performance on reference datasets than passive SM products [11–15]. Therefore, a scheme to merge both active and passive microwave SM products may add value for an improved global SM product over various vegetation types.

Land surface models are available on a global scale, but their quality is not finely characterized, due to the errors in the modeling itself, requiring continuous updates when observations available [14,16]. In most of the existing methods to merge satellite SM products is implemented by algorithms that maximize correlation or decrease the error between satellite data and the specific land surface model. Therefore, the accuracy of the merged SM dataset is based on the accuracy of this model itself and shifts when another land surface model is deployed [13,17,18].

In relation to noise estimates to the SM sensitivity individually using fMSE, not only removes the error patterns based on the spatial sensitivity to the scaling reference, it also provides a quantitative quality assessment for actual data. The levels of error are usually highest in regions with a strong SM signal or when the SM dataset is more sensitive to changes within it. Conversely, the lower levels of error are expected in the regions with a very low SM sensitivity or very low SM variability, such as the desert areas. Error levels are considered important for certain applications; therefore, data quality cannot be investigated using absolute errors alone [19,20]. However, some previous studies merged satellite SM products based on error variance derived from the triple collocation method [21,22].

This study develops and evaluates a novel scheme for merging SM retrievals. Advanced Scatterometer (ASCAT), Soil Moisture Active Passive (SMAP), and Advanced Microwave Scanning Radiometer 2 (AMSR2) data for period from June 2015 to December 2017 were merged to produce SMAP+ASCAT and AMSR2+ASCAT SM datasets based on maximized signal to noise ratios using the weighted averaging method.

2. Datasets

In this research, in order to implement the proposed scheme, we used different datasets in the period from June 2015 to December 2017 include satellite SM datasets (SMAP, AMSR2, and ASCAT), in-situ SM observations (ISMN), reanalysis SM products (GLDAS and ERA5), land cover dataset. All the datasets were resampled to a resolution of 0.25° which is often adopted by the reanalysis datasets. The nearest-neighbor interpolation method was used in the resampling process [23–25]. The details about these datasets will be presented in the following subsections.

2.1. Satellite SM Datasets

In this work, the global daily SM retrievals of SMAP L3 Radiometer version 6 with grid resolution of 36 km was used as one of the parent SM products for the SMAP+ASCAT merged dataset. This data is available freely online (https://search.earthdata.nasa.gov). The National Aeronautics and Space Administration (NASA) launched the SMAP in January 2015 as the first Earth Observation (EO) satellite [26,27]. The SMAP mission was designed to covers the globe every three days and to retrieve

SM over moderate or low vegetation regions within an accuracy of 0.04 m^{3}m^{-3}. The overpass times for the equator, according to solar time, for this satellite are 6:00 p.m. in the ascending mode and 6:00 a.m. in the descending mode [28,29]. The SMAP soil moisture data set was filtered for SM values lower than 0.02 m^{3}m^{-3}and higher than 0.50 m^{3}m^{-3}, when surface soil temperature < 273.15, and based on the SM retrieval quality flag when was set as "recommended for retrieval" [13,30].

The global daily SM retrievals LPRM AMSR2 X-band with grid resolution a 25 was used as one of the parent SM products for AMSR2+ASCAT merged product. In May 2012, the Japan Aerospace Exploration Agency (JAXA) launched the AMSR2 SM sensor loaded on the GCOM-W1 platform [31,32]. The AMSR2 satellite mission follow-up of Aqua AMSR-E since October 2011 to maintain on long-temporal of satellite observations [33]. The overpasses times of the equator according to solar time for this satellite are 1:30 a.m. and 1:30 p.m. in descending and ascending modes, respectively. The AMSR2 SM datasets can be derived from two algorithms include the Land Parameter Retrieval Model (LPRM) and JAXA [15,34]. This data can be accessed freely through (https://search.earthdata.nasa.gov). The pixels of AMSR2 product are masked if the surface soil in a freezing state (i.e., soil temperature less than 273.15 K).

The ASCAT SM data record that available through the product H113 with grid resolution 12.5 km was used as one of the parent SM products for both SMAP+ASCAT and AMSR2+ASCAT merged datasets. This data is available freely online (http://hsaf.meteoam.it). The ASCAT is an active microwave radar working at C-band (5.255 GHz) using the TU Wien algorithm. The ASCAT was operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). At present, ASCAT loaded on several platforms include Metop-A (launched in October 2006), Metop-B (launched in September 2012), and Metop-C (launched in November 2018) satellites [35]. The H113 dataset was generated based on Metop-A and Metop-B Level 1b backscatter datasets [36]. The unit of ASCAT SM dataset was translated from the degree of saturation (%) to volumetric unit (m^{3}m^{-3}) by using soil porosity derived from the database of world soil for the top layer (0–0.40 m) (http://www.esa-soilmoisture-cci.org). The pixels of ASCAT SM Product were masked if the probability of snow cover/frozen and retrieval error > 50% using surface state flag data [5,14,37].

2.2. Reanalyzed SM Datasets

The ground SM observations can be considered the ideal reference dataset to evaluate the performance of SM datasets [38,39]. However, the ground observations have some challenges due to mismatches between it and the footprints from large scale satellite measurements and the limitations of spatial coverage [40–42]. To overcome these challenges, the SM retrievals can be assessed by a comparison with reanalyzed datasets [43].

The Global Land Data Assimilation System (GLDAS) was developed by the National Oceanic and Atmospheric Administration (NOAA) and NASA. The GLDAS-Noah model provides time series (3-h) SM data for four layers with different depths at a spatial resolution of 0.25° [44]. The SM of the top layer (0–10 cm) from GLDAS Noah version 2.1 was used as one of the triple collocation method (TCM) datasets and to assess the performance of merged and parent SM datasets. GLDAS Noah has been used as a reference dataset in several studies for blending satellite SM datasets [13,22,45]. Moreover in several studies, the top 10 cm layer of SM from reanalyzed products was used to validate and improve SM retrievals for different satellites [35,46].

The daily averaged of the ERA5 reanalysis SM dataset with spatial resolution 0.25 in the period from June 2015 to December 2017 was used to evaluate the performance of merged and parent SM datasets. The European Center for Medium-Range Weather Forecasts (ECMWF) provides a global ERA5 atmospheric reanalysis to the public. In previous studies, the evaluation of ERA5 on global scale proved that the performance of ERA5 has reliability in the absence of ground measurements [47–49]. More information about this product is available at https://apps.ecmwf.int/datasets/.

2.3. In-Situ SM Dataset

The ground SM measurements from various networks of the international soil moisture network (ISMN) were used to evaluate the performance of merged and parents SM products. The locations of ISMN ground stations used in this study are displayed in Figure 1.
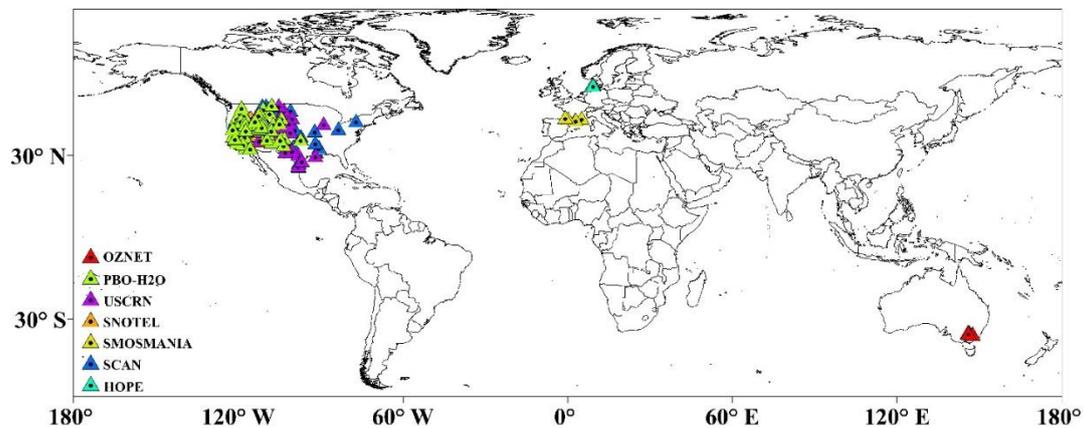

Figure 1. The locations of the ISMN validation stations used in this study at the global scale.

The ISMN is a website that stores and organizes the SM measurements form various in-situ stations of different global networks and make it freely through (https://ismn.geo.tuwien.ac.at). The ISMN data collected from approximately 1400 ground stations. The SM measurements of ISMN are important for evaluating the performance of SM retrievals from satellites and land surface models and for studying the climate [15,40,50]. The main features of ISMN validation sites used in this study are listed in Table 1.

Table 1. The main characteristics of ISMN validation sites were used in this study.

Network Name	Location	No. of Stations	Depth (cm)	Type of Sensor	References
PBO-H2O	USA	74	0–5	GPS	Larson et al. [51]
SCAN	USA	46	5	Hydra-Probe analog (2.5 Volt)	Schaefer et al. [52]
SNOTEL	USA	51	5	Hydra-Probe analog (2.5 Volt)	https://www.wcc. nrcs.usda.gov/
USCRN	USA	29	5	Stevens-Hydra-Probe II Sdi-12	Bell et al. [53]
SMOSMANIA	France	5	5	Theta-Probe ML2X	Albergel et al. [54]
HOBE	Denmark	20	0–5	Decagon-5TE-A	Jensen et al. [55]
OZNET	Australia	15	0–5	Stevens-Hydra-Probe	Smith et al. [56]

2.4. Land Cover Dataset

The quality of SM data from both satellite and land surface model changes according to land cover types [14,57]. In this study, The MODIS (Moderate Resolution Imaging Spectroradiometer) land cover product (MCD12Q1) with grid resolution 0.5 km was used to assess the performance of merged and parents SM products over different land covers on a global scale. MCD12Q1 product provides a yearly land cover on a global scale from 2001 to the present available freely for users through (https://search.earthdata.nasa.gov/) [58]. Figure 2 shows a global map of the MODIS land cover.
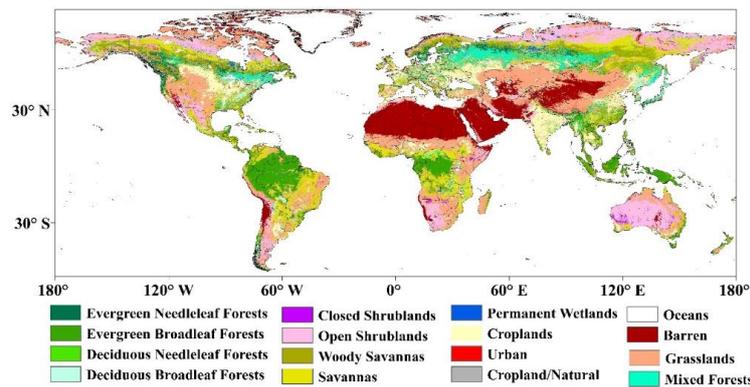
Figure 2. The global map of the MODIS land cover.

3. Methods

A flowchart of the proposed blending scheme of SM datasets is shown in Figure 3. The blending scheme is based on maximizing signal to noise ratios. The fractional mean-squared-error (fMSE) estimates derived by the TCM was used for weight estimation based on the least-squares method. The proposed scheme was applied by using a threshold between signal and noise at fMSE equal to 0.5 to maintain the high-quality SM observations. Based on the correlations determinations between all three SM datasets of TCM at significance levels (i.e., *p*-values), we proposed four scenarios in the regions of unreliable TCM. These scenarios include using passive SM product only or using an active SM dataset only or using an unweighted average or exclude the pixels without any retrieval skills. The proposed merge scheme was applied to SM retrievals from ASCAT, SMAP, and AMSR2 in the period from June 2015 to December 2017 to produce SMAP+ASCAT and AMSR2+ASCAT SM datasets. The two merged SM datasets were evaluated against SM estimates from in situ observations of ISMN and reanalysis datasets (GLDAS and ERA5). More details about the merging scheme in the following sub-sections.
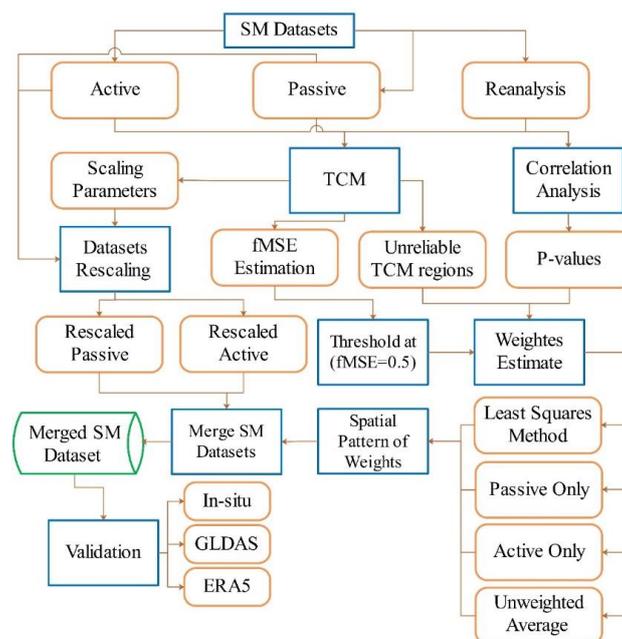

Figure 3. Flowchart of the proposed methodology for merging SM datasets.

3.1. Maximized SNR for Merging SM Datasets

There are different ways to define optimality. In this research, the least-squares method was used to determine the optimal weight. Least-squares is widely used for data assimilation, since it was developed by [59,60]. The least-squares method deploys a weighted average method for calculating the optimal weights for parent of the merged datasets [22,61,62]. The original formula of the weighted mean can be written as,

$$SM_c = \frac{\sum_{i=1}^{n} W_i * SM_i}{\sum_{i=1}^{n} W_i} \tag{1}$$

where SM_c is the averaged SM estimate; SM_i is the SM estimates for different datasets; and W_i are the weights given to SM estimates when the output product is blended as a linear combination of single products. To merge the independent uncorrelated observations of active and passive satellite SM, the original Equation (1) of merging datasets can be expressed as,

$$SM_c = W_{ac}SM_{ac} + W_{pa}SM_{pa} \tag{2}$$

where W_{ac} and W_{pa} are the weights of active and passive satellite SM datasets; and SM_{ac} and SM_{pa} are the values of active, and passive SM retrievals, respectively.

To overcome on the dependency of the relative estimated weights for merged product on a specific model, the weights are given for satellite datasets were calculated, based on maximizing signal to noise ratios of SM estimates. The fractional mean-squared-error (fMSE) was used for this purpose. Where fMSE has a normalized representation of the signal to noise ratio (SNR) and it has a specific range (i.e., from 0 to 1). A higher/lower fMSE indicates a noisier/clearer signal of the SM value. When the fMSE is equal to 0, it means that the SM observation is free from noise. When the fMSE is equal to 1, it is means the SM observations have only noise. Also, the SM signal becomes stronger than its noise when the fMSE value is lower than 0.5 [13,20]. The weights can be determined using the following equations,

$$W_{ac} = \frac{fMSE_{pa}}{fMSE_{ac} + fMSE_{pa}} \tag{3}$$

$$W_{pa} = \frac{fMSE_{ac}}{fMSE_{ac} + fMSE_{pa}} \tag{4}$$

where W_{ac} and W_{pa} are the weights of active and passive SM products respectively; $fMSE_{ac}$ and $fMSE_{pa}$ are the fractional mean squared errors of active and passive SM datasets respectively.

To maintain the high-quality SM observations, the proposed scheme was applied using a threshold between signal and noise at fMSE equal to 0.5. The different cases for using fMSE threshold to maintain high-quality SM retrieval data are listed in Table 2.

Table 2. The different cases for using fMSE threshold to maintain high-quality SM retrieval data. Note that, the threshold value was chosen at fMSE equal to 0.5.

Active	Passive	Decision
fMSE < 0.5	fMSE < 0.5	Weighted average method
fMSE > 0.5	fMSE > 0.5	Weighted average method
fMSE > 0.5	fMSE < 0.5	Passive only
fMSE < 0.5	fMSE > 0.5	Active only

The weighted average method was used only when the signal of both active and passive exceed the noise (i.e., fMSE < 0.5) or when the noise of both them exceed the signal (i.e., fMSE > 0.5). When the signal of one parent dataset exceeds the noise (i.e., fMSE < 0.5) and the noise for other dataset exceed the signal (i.e., fMSE > 0.5), we take only the dataset the strong signal.

Moreover, prior to merging datasets using the estimated weights, the systematic differences between the datasets were removed. This was achieved by rescaling SM products into a mutual data space. In this research, the fMSE estimates and rescaling coefficients were derived using TCM [20,63–65].

3.2. Triple Collocation Method

The triple collocation method (TCM) was utilized to determine the fMSE and rescaling coefficients of the active and passive SM datasets. TCM considers a promising method to validate remotely sensed SM datasets [63,65]. Where the random error variance and SNR of these datasets can be estimated on a larger scale without the need for ground reference information [66]. Furthermore, it provides an optimal solution for rescaling SM datasets take into consideration the individual random error properties and match the variability of the jointly observed signal [67]. Further details about TCM derivation are presented in previous research [20,68].

To remove the dependence of the scaled error pattern on the spatial climatology of the selected scaling reference, Draper et al. [68] suggested normalizing the estimates of the un-scaled error variance with the variance of corresponding datasets. The fMSE of active and passive SM datasets can be determined as,

$$\mathrm{fMSE_{ac}} = \frac{1}{1 + \mathrm{SNR_{ac}}} \tag{5}$$

$$\mathrm{fMSE_{pa}} = \frac{1}{1 + \mathrm{SNR_{pa}}} \tag{6}$$

where $\mathrm{SNR_{ac}}$ and $\mathrm{SNR_{pa}}$ are the signal to noise ratios of active and passive SM products respectively. The signal to noise ratios for active and passive SM datasets can be calculated using the following formulas,

$$\mathrm{SNR_{ac}} = \frac{\sigma_{ac,pa}\,\sigma_{ac,mo}}{\sigma^2_{ac}\sigma_{pa,mo}} \tag{7}$$

$$\mathrm{SNR_{pa}} = \frac{\sigma_{pa,ac}\,\sigma_{pa,mo}}{\sigma^2_{pa}\sigma_{ac,mo}} \tag{8}$$

where $\sigma_{ac,pa}$, $\sigma_{pa,mo}$ and $\sigma_{ac,mo}$ are the covariance of datasets; σ^2_{ac} and σ^2_{pa} are the variance of the active, and passive datasets, respectively.

Draper et al. [69] recommended normalization of the SM datasets to remove systematic differences between them, based on one reference dataset. To remove the systematic differences between SM datasets, the rescaling coefficients are estimated as,

$$\mathrm{B_{RESpa}} = \frac{\sigma_{ac,mo}}{\sigma_{pa,mo}} \tag{9}$$

$$\mathrm{SM_{RESpa}} = \mathrm{B_{RESpa}}(\mathrm{SM_{Pa}} - \overline{\mathrm{SM}_{Pa}}) + \overline{\mathrm{SM}_{ac}} \tag{10}$$

where $\mathrm{B_{RESpa}}$ is a coefficient used to linearly rescaling of the passive against the active SM dataset; $\overline{\mathrm{SM}_{Pa}}$ and $\overline{\mathrm{SM}_{ac}}$ are the average of SM retrieval for active and passive datasets respectively. The dataset used as a reference in rescaling process does not matter, as this does not affect the rescaled dataset or the merged time series.

3.3. Blending of Unreliable TCM Observations

TCM is deemed unreliable due to the limited numbers of observations (i.e., <100). This is caused by one or more of SM datasets in the TCM triplet having retrieval issues such as spatial coverage gaps due to masking procedures or radio frequency interference (RFI), the difference in overpass times, and the lower time-series coverage. TCM results can also be unreliable if one or more SM datasets of the TCM triplet are correlated insignificantly to other datasets as indicated by a student's

t-test [42,65]. The area of unreliability may be found in certain regions, such as over desert regions and at high latitudes.

In this work, we proposed four scenarios to increase the temporal and overall spatial coverage of merged datasets in regions where the TCM is unreliable. These scenarios include passive SM products only or active SM datasets only, an unweighted average, or excluded pixels. The selection of a suitable scenario is based on the retrieval quality of the SM dataset. The correlations between all three SM datasets in the TCM (i.e., passive and active products, active product and reanalysis, passive product and reanalysis) was calculated at significance levels (i.e., *p*-value < 0.05) as an indication for retrieval quality. The criteria for choosing the different scenarios in unreliable TCM regions are as follows:

- The active SM retrieval data were used only in the area where the passive product correlates insignificantly to both model and active product, while the active product is correlated significantly with the model. Moreover, the pixels of an active SM ware used in the regions where the correlations of the active product with both, passive product and model are significant, while the passive product correlates insignificantly with the model.
- The passive SM dataset was used only in the regions where the active correlates insignificantly with both the model and passive product, while the passive product is correlated significantly with the model. Also, a passive SM still the best choice, if the correlation of passive with both active product and model are significant, while the active product correlates insignificantly with the model.
- The unweighted average was used in the case of, the correlations of both active and passive SM datasets are significant with the reanalysis, but not with each other. Also, the unweighted average method outperformed than other methods, if the active and passive SM datasets are correlated insignificantly with the model but the correlation between active and model is significant.
- Excluded the pixels at which the correlations between all three SM of TCM are insignificant with each other. Moreover, these pixels showed an insignificant correlation against the independent reference dataset when applied in both scenarios using a single sensor and an unweighted mean method.

3.4. Hovmöller Diagrams

The SM strongly varies temporally and spatially. This variability mainly depends on latitude or longitude and season [23,70]. Hovmöller diagrams were used to study the spatiotemporal variability of merged and parents SM datasets [14,71,72]. A hovmöller diagram represents the time variability of spatial data. Where the time is presented on the *x*-axis and the average values (either overall latitudes or overall longitudes) of datasets are displayed on the *y*-axis [73]. In this research, the longitudinal averages for SM values were used to study the consistent between parents and merged SM products.

3.5. Error Statistics

In this research, the statistical metrics were used to evaluate the merged and parents SM datasets against reference (in-situ and reanalysis) datasets. These metrics include the Pearson correlation coefficients (R) and the root mean squared difference (RMSE) [74]. The R and RMSE between SM datasets and reference datasets are calculated by using the following equations,

$$R = \frac{E[(SM(t) - E[SM(t)])(SM_r(t) - E[SM_r(t)])]}{\sigma \sigma_r} \tag{11}$$

$$RMSE = \sqrt{E\left[(SM(t) - SM_r(t))^2\right]} \tag{12}$$

where E[·] represents the estimation of the average values, t is the time series of the dataset, SM(t) represents the SM observations of merged or parent dataset at time t; $SM_r(t)$ represents the SM

observations of reference dataset; σ represents the standard deviation of merged or parent SM dataset; σ$_r$ represents the standard deviation of the reference SM dataset.

4. Results

4.1. Spatiotemporal Variability of SM Datasets

The performance of merged datasets can be achieved by constructing hovmöller diagrams. Hovmöller diagrams of SMAP+ASCAT, AMSR2+ASCAT, SMAP, AMSR2, ASCAT, GLDAS, and ERA5 SM datasets on the global scale are presented in Figure 4. In these diagrams, time is represented on the abscissa from June 2015 to December 2017 and the latitude from 90°S to 90°N is displayed on the ordinate. These diagrams show that the overall global spatial patterns of all SM datasets agree. All datasets displayed the same seasonal SM dynamics and SM variations on a global scale. For example, around the equator, there are clear seasonal changes related to the intertropical convergence zone and monsoon changes. The seasonal change in the SM dynamic in the last half of 2015 was consistent with the same seasons in 2016 and 2017. The SMAP+ASCAT and AMSR2+ASCAT merged SM datasets showed the same spatiotemporal variability of parents (SMAP, ASCAT, and AMSR2) and reanalysis referenced (GLDAS and ERA5) SM products.
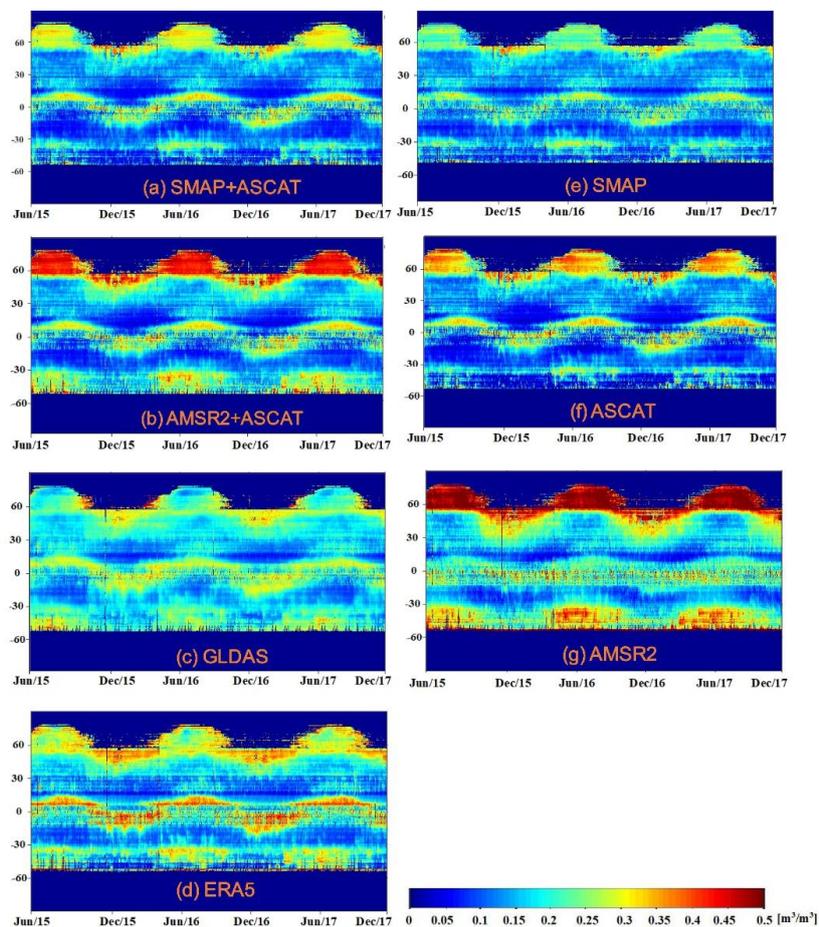

Figure 4. Hovmöller diagrams of merged and parents SM datasets in the period from June 2015 to December 2017: (**a**) SMAP+ASCAT; (**b**) AMSR2+ASCAT; (**c**) GLDAS; (**d**) ERA5; (**e**) SMAP; (**f**) ASCAT; and (**g**) AMSR2.

4.2. Optimal Weighting Factors

The optimal weights indicate the relative contribution of the remotely sensed SM datasets against each other. Also, the optimal weights provide evidence about the weakness and strengths of satellite SM products over different land covers. The relative optimal weights of parents SM products (SMAP, AMSR2, and ASCAT) for SMAP+ASCAT and AMSR2+ASCAT merged datasets are presented in Figures 5 and 6.
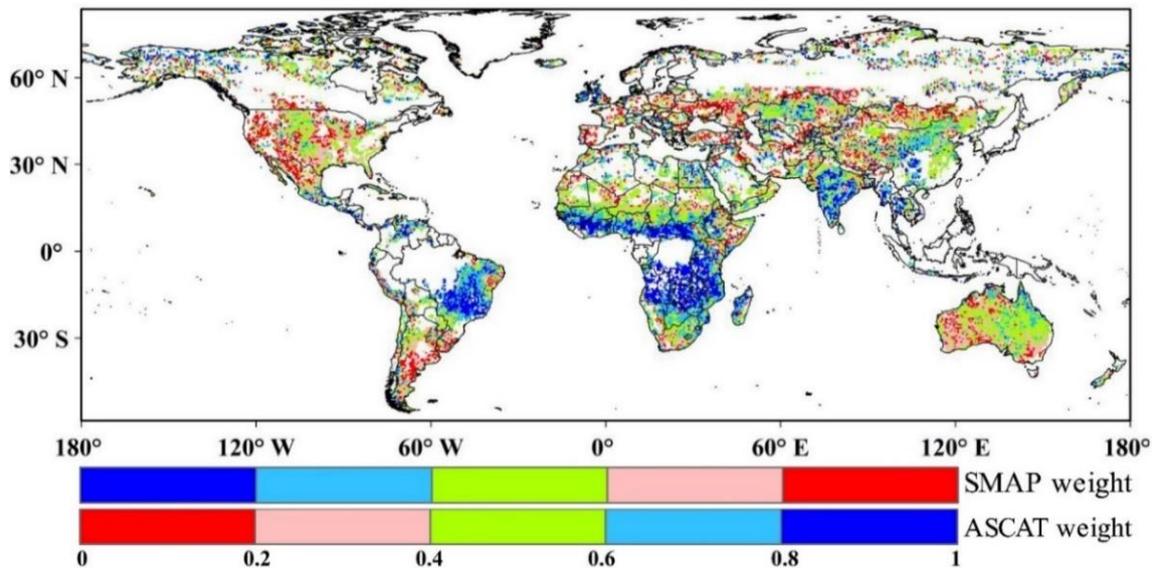

Figure 5. The spatial distribution of the global optimal weights for parent SM products (SMAP and ASCAT) of the SMAP+ASCAT merged dataset.
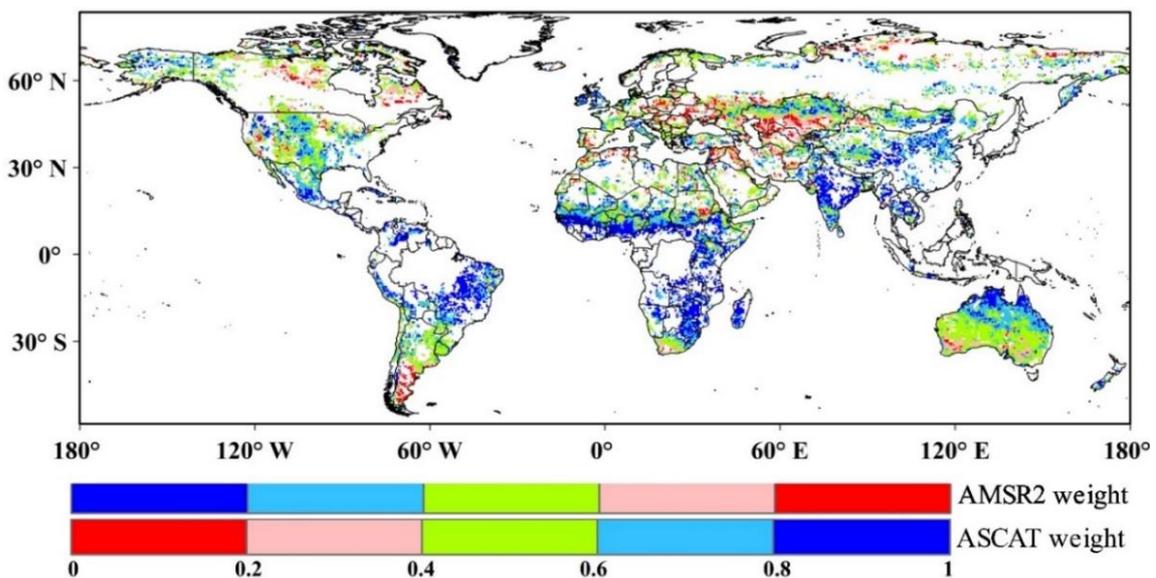

Figure 6. The spatial distribution of the global optimal weights for the parent SM products (AMSR2 and ASCAT) of the AMSR2+ASCAT merged dataset.

The weights are given for parent (SMAP, ASCAT, and AMSR2) SM products of SMAP+ASCAT and AMSR2+ASCAT datasets derived from the proposed scheme as discussed. In two merged SM datasets, the blue color indicates that the majority of the weight comes from the ASCAT product while the majority of SMAP and AMSER2 weights are represented with red colors. The ASCAT SM

product was given a high relative weight than SMAP and AMSR2 in their merged products over more dense vegetation regions (forests, savannas, and crops) (see land cover map Figure 2) which is mostly concentrated in high-temperature regions below and above the equator. While, the SMAP and AMSR2 are given high relative weight than ASCAT in SMAP+ASCAT and AMSR2+ASCAT merged datasets over the areas of moderate and less vegetation (grassland, shrub, and desert) (see land cover map in Figure 2). The analysis of the weights are given for both SMAP and AMSR2 in their merged datasets over various land cover we found that the SMAP was given a higher relative weight than AMSR2.

4.3. Evaluation of Merged SM Datasets

A key objective of the proposed scheme was to develop two merged SM products that outperform the parent products. Moreover, the selection of the SM dataset for some critical applications is based on their quality or accuracy. Therefore, the performance of SMAP+ASCAT and AMSR2+ASCAT SM merged products were assessed by a comparison with independent reference datasets include in-situ measurements, GLDAS, and ERA5 using R and RMSE metrics.

4.3.1. Validation of Merged Products against In-Situ Measurements

The SMAP+ASCAT and AMSR2+ASCAT products assessed against SM measurements of ISMN networks using correlations coefficients analysis and RMSE to fully understand how the merged products perform relative to parents products. Also, the parents (SMAP, AMSR2, and ASCAT) products were evaluated against the same ISMN networks which indicate the strengths points of these products and how contribute to merged products. The merged and parents datasets compared with the ground stations of different six networks available from ISMN and concentrated over America and Europe which include PBO-H2O, SNOTEL, SCAN, HOPE, SMOSMANIA, and USCRN. The results of overall averages of R and RMSE for the comparison of merged and parents SM datasets with the in-situ SM observations of different ISMN networks are listed in Table 3.

Table 3. The results of overall averages of R and RMSE for the comparison of merged (SMAP+ASCAT and AMSR2+ASCAT) and parents (SMAP, ASCAT, and AMSR2) SM datasets with the in-situ SM observations of different ISMN networks. Note that, p-value < 0.01.

Network Name	SMAP		ASCAT		AMSR2		SMAP+ASCAT		AMSR2+ASCAT	
	R	RMSE $(m^3 m^{-3})$	R	RMSE $(m^3 m^{-3})$	R	RMSE $(m^3 m^{-3})$	R	RMSE $(m^3 m^{-3})$	R	RMSE $(m^3 m^{-3})$
PBO-H2O	0.73	0.058	0.70	0.075	0.67	0.096	0.75	0.057	0.73	0.066
SCAN	0.69	0.068	0.62	0.090	0.53	0.132	0.70	0.067	0.66	0.089
SNOTEL	0.64	0.095	0.62	0.101	0.53	0.133	0.68	0.088	0.66	0.095
USCRN	0.72	0.083	0.64	0.095	0.57	0.126	0.73	0.079	0.69	0.089
SMOSMANIA	0.75	0.088	0.64	0.113	0.57	0.192	0.78	0.082	0.72	0.101
HOBE	0.73	0.097	0.63	0.143	0.60	0.233	0.74	0.084	0.68	0.131
OZNET	0.84	0.063	0.78	0.071	0.74	0.092	0.85	0.052	0.79	0.063
Average	0.73	0.079	0.66	0.098	0.60	0.143	0.75	0.073	0.70	0.091

The overall averages of correlation coefficients for all ISMN networks were 0.75, 0.70, 0.73, 0.66, and 0.60 for SMAP+ASCAT, AMSR2+ASCAT, SMAP, ASCAT, and AMSR2, respectively. In terms of overall average correlations values, the SMAP+ASCAT and AMSR2+ASCAT SM datasets showed higher overall average R values with in-situ measurements over all different six ISMN networks than either of the parents datasets (SMAP, AMSR2, and ASCAT).

The overall averages of RMSE for all ISMN networks were (0.073, 0.091, 0.079, 0.098, and 0.143) $m^3 m^{-3}$ for SMAP+ASCAT, AMSR2+ASCAT, SMAP, ASCAT, and AMSR2, respectively. In terms of overall average RMSE values, the SMAP+ASCAT and AMSR2+ASCAT SM products showed the lowest overall average RMSE values with in-situ measurements over all different six ISMN networks than either of the parents datasets (SMAP, AMSR2, and ASCAT).

To analysis the full performance of two merged datasets comparing with parent datasets we constructed the box plots. The overall performance of merged and parents SM products against the in-situ SM observations of different six ISMN networks are displayed on box plots of Figures 7 and 8.
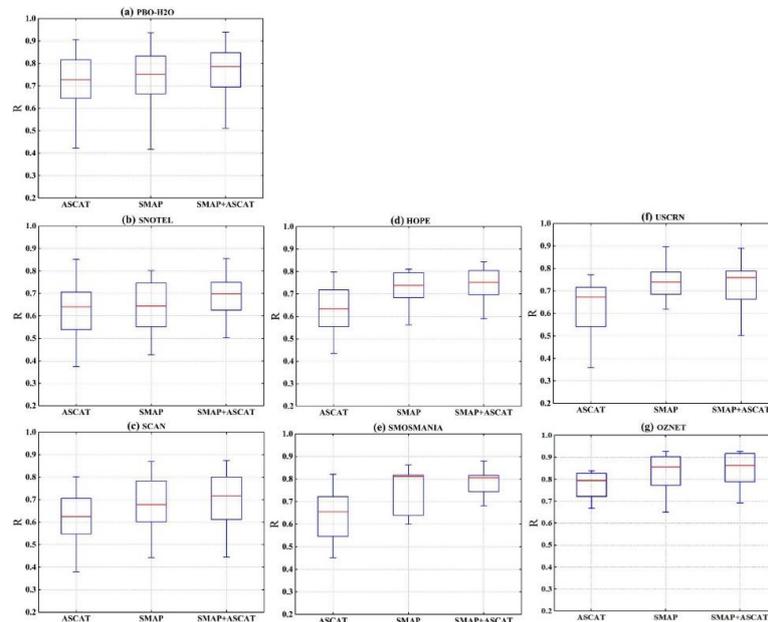

Figure 7. Box plots of correlation coefficients of SMAP+ASCAT, SMAP, and ASCAT SM products against the in-situ SM observations of different ISMN networks: (**a**) PBO-H2O; (**b**) SNOTEL; (**c**) SCAN; (**d**) HOPE; (**e**) SMOSMANIA; (**f**) USCRN; and (**g**) OZNET.
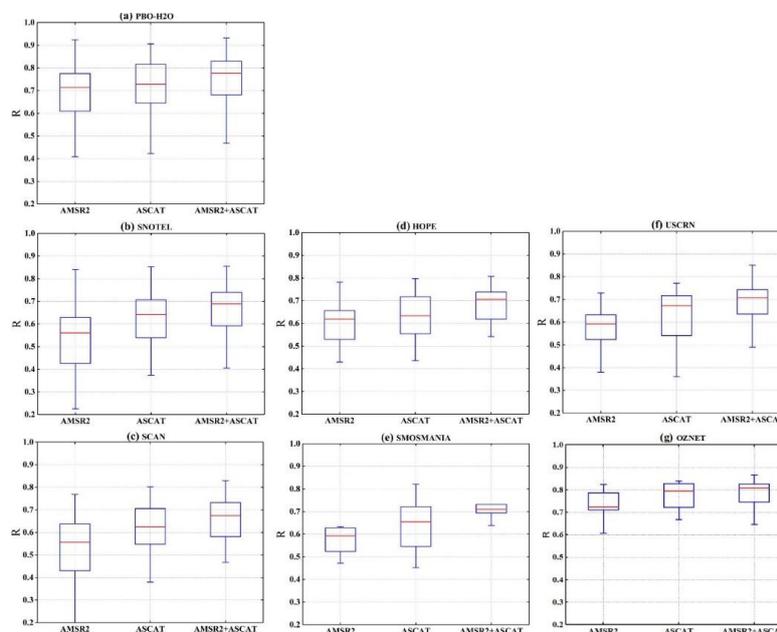

Figure 8. Box plots of correlation coefficients of AMSR2+ASCAT, AMSR2, and ASCAT SM datasets against the in-situ SM observations of different ISMN networks: (**a**) PBO-H2O; (**b**) SNOTEL; (**c**) SCAN; (**d**) HOPE; (**e**) SMOSMANIA; (**f**) USCRN; and (**g**) OZNET.

The overall performance of SMAP+ASCAT and AMSR2+ASCAT datasets outperformed the parents datasets (SMAP, ASCAT, and AMSR2) over different six ISMN networks. These improvements

are displayed in box plots of correlations coefficients for merged and parents' products against ground measurements (see the median, the 3rd quantile Q3, and the 1st quantile Q1).

4.3.2. Evaluation of Merged Products against Modeling Data

Although, SM measurements of ground stations consider more robust reliable methods to assess the performance of the satellite SM retrieval data, but still has limitations in the temporal and spatial coverage all over the world. Also, the ground observations have some challenges due to mismatches between it and the footprints from large scale satellite measurements and the limitations of spatial coverage [40–42]. Therefore, to overcome these problems and fully evaluate the merged and parents SM datasets over different land covers, the comparison with reanalysis becomes the alternative solution. The merged and parents SM products were assessed by comparing it with GLDAS and ERA5 reanalysis datasets using R and RMSE metrics. The spatial distribution of differences in correlation coefficients (against modeling datasets) between merged (SMAP+ASCAT and AMSR+ASCAT) and parents (SMAP, AMSR2, and ASCAT) SM datasets are displayed in Figures 9 and 10.
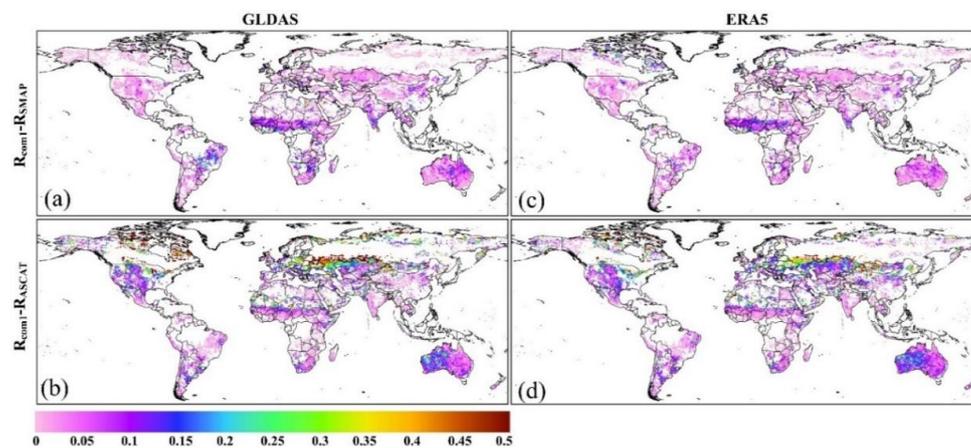

Figure 9. The spatial distribution of differences in correlation coefficients (against modeling datasets) between merged (SMAP+ASCAT) and parent (SMAP and ASCAT) SM datasets, where (Rcom1 minus RSMAP) (first row), (Rcom1 minus RASCAT) (second row): (**a**,**b**) for GLDAS, and (**c**,**d**) for ERA5. Note that, com1 refer to SMAP+ASCAT product.
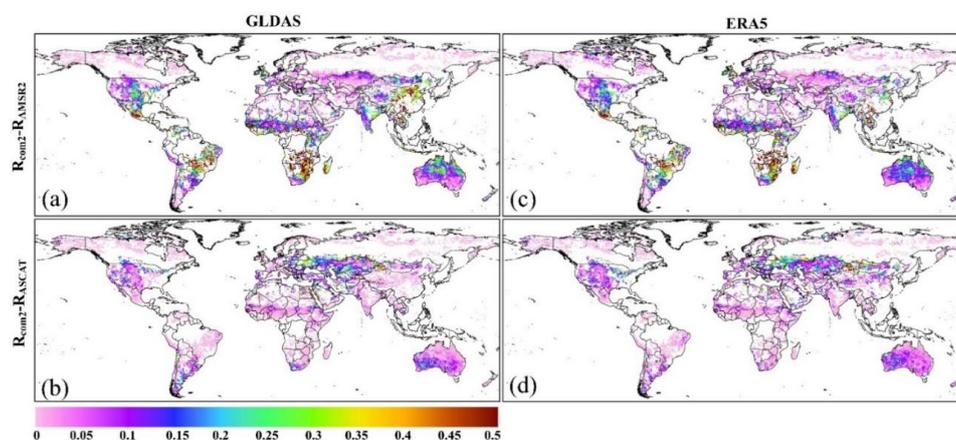

Figure 10. The spatial distribution of differences in correlation coefficients against modeling datasets between merged (AMSR2+ASCAT) and parent (AMSR2 and ASCAT) SM datasets, where (Rcom2 minus RAMSR2) (first row), (Rcom2 minus RASCAT) (second row): (**a**,**b**) for GLDAS and (**c**,**d**) for ERA5. Note that, com2 refers to the AMSR2+ASCAT dataset.

The SMAP+ASCAT and AMSR2+ASCAT are achieved improvement in temporal correlations coefficients with GLDAS and ERA5 outperforming the parents datasets overall the world. These correlations improvements spatially distributed over different land cover types of all most areas of the world according to the spatial weights patterns of parents' datasets. Over more dense vegetation regions (forests, savannas, and crops) the SMAP and AMSR2 are displayed high R differences with their merged datasets. Over these areas, the SMAP and AMSR2 were given the lower relative weights than ASCAT products for merged products. Conversely, over moderate and less vegetation regions (grassland, shrub, and desert) the ASCAT showed high R differences than their merged products. Over these regions the ASCAT was given the lowest relative weights than SMAP and AMSR2 products in their merged products. Although, the GLDAS is used as one of three datasets of TCM to calculate fMSE however the SMAP+ASCAT and AMSR2+ASCAT SM products showed better performance with ERA5 exceeds the performance with GLDAS. Box plots of correlation coefficients of merged and parents SM datasets against the modeling SM products are presented in Figures 11 and 12.
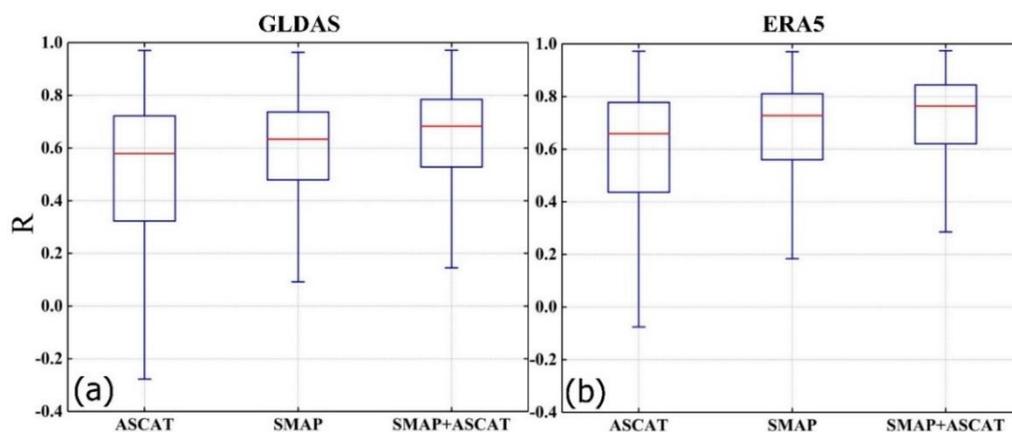

Figure 11. Box plots of correlation coefficients of merged (SMPA+ASCAT) and parent (SMAP and ASCAT) SM datasets against the modeling SM products (GLDAS and ERA5): (**a**) GLDAS; and (**b**) ERA5.
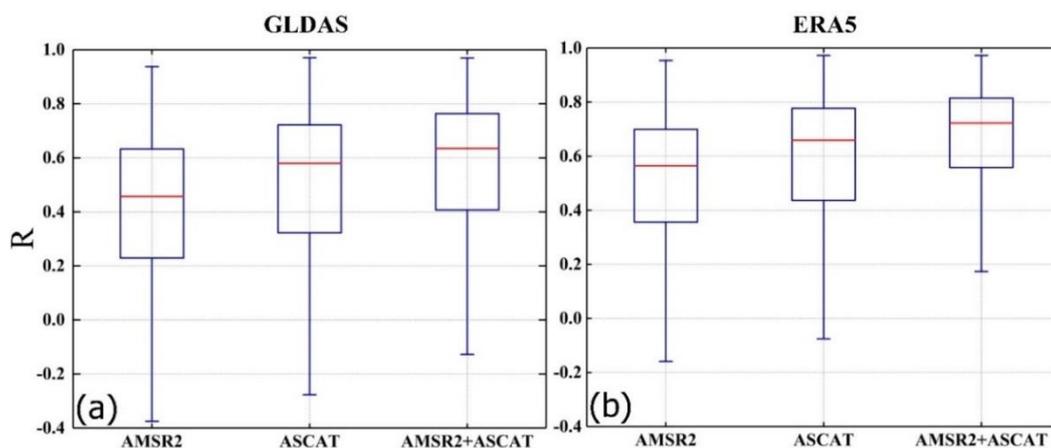

Figure 12. Box plots of correlation coefficients of merged (AMSR2+ASCAT) and parent (AMSR2 and ASCAT) SM datasets against the modeling SM products (GLDAS and ERA5): (**a**) GLDAS; and (**b**) ERA5.

Box plots of correlations coefficients for SMAP+ASCAT, AMSR2+ASCAT, SMAP, ASCAT, and AMSR2 products against GLDAS and ERA5 references datasets (see the median, the third quantile Q3, and the first quantile Q1) outperformed the merged datasets their parent datasets all over the world. The overall averages of temporal correlations coefficients for two merged and parents SM datasets against the modeling SM products are presented in bar graphs of Figure 13.
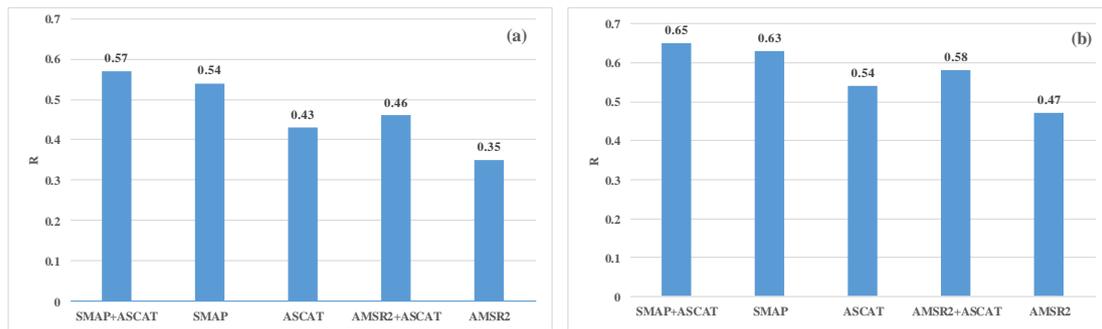
Figure 13. Bar graphs for the overall averages of temporal correlations coefficients for two merged and parents SM datasets against the modeling SM products: (**a**) GLDAS; and (**b**) ERA5.

In terms of overall average correlations coefficients values, the SMAP+ASCAT and AMSR+ASCAT merged SM datasets showed a high correlation with GLDAS and ERA5 which outperform the parents SM datasets (SMAP, AMSR2, and ASCAT).

In this study, the second statistics metric was used to assess the performance of merged products is RMSE. Box plots of RMSE of merged (SMAP+ASCAT and AMSR+ASCAT) and parents (SMAP, AMSR2, and ASCAT) SM datasets against the modeling (GLDAS and ERA5) SM products are presented in Figure 14.
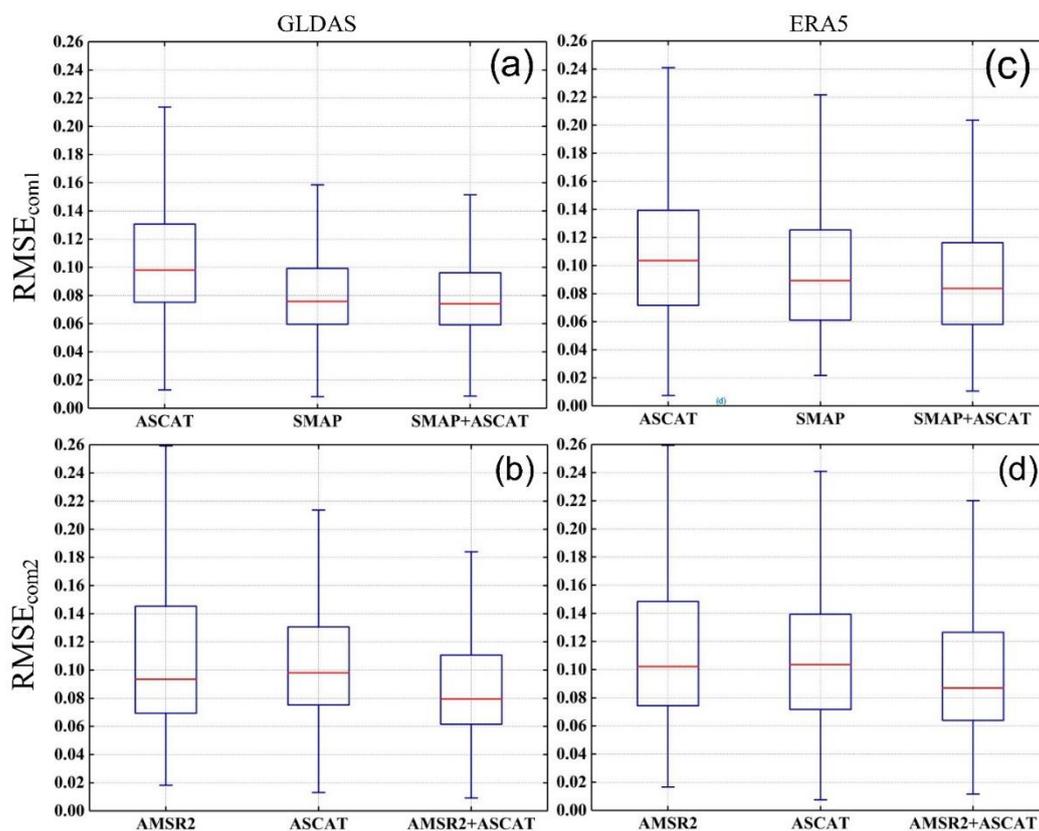

Figure 14. Box plots of RMSE of merged (SMAP+ASCAT and AMSR2+ASCAT) and parents (SMAP, AMSR2, and ASCAT) SM datasets against the modeling SM products (GLDAS and ERA5), where (SMAP+ASCAT) (first row), and (AMSR2+ASCAT) (second row): (**a**,**b**) for GLDAS, and (**c**,**d**) for ERA5. Note that, com1 and com2 refer to SMAP+ASCAT and AMSR2+ASCAT, respectively.

The overall averages RMSE values with GLDAS SM dataset were (0.081, 0.096, 0.083, 0.105, and 0.119) m^3m^{-3} for SMAP+ASCAT, AMSR2+ASCAT, SMAP, ASCAT, and AMSR2, respectively. While the overall averages RMSE values with the ERA5 SM dataset are (0.093, 0.101, 0.101, 0.109, and 0.121) m^3m^{-3} m^3m^{-3} for SMAP+ASCAT, AMSR2+ASCAT, SMAP, ASCAT, and AMSR2, respectively. In terms of overall average RMSE values, the merged SM datasets showed the lowest overall average RMSE, with two modeling datasets that outperform the parents SM products. Box plots of RMSE for SM products against GLDAS and ERA5 datasets displayed outperforming for the merged datasets on the parents datasets overall the world (see the median, the 3rd quantile Q_3, and the 1st quantile Q_1).

In the regions of unreliable triple collocation analysis, as mentioned, we proposed four scenarios based on the determinations of correlations between all three SM products of TCM at significance levels (i.e., *p*-values < 0.05). These scenarios include, either a passive SM dataset only, or using an active SM dataset only, or using an unweighted average method, or excluding the pixels without any retrieval skills. The independent ERA5 reference product was used to make sure from the reliability of the selected scenario according to the suggested criteria. The SMAP+ASCAT and AMSR2+ASCAT merged products produced by using the scenarios of using the single sensor (i.e., either active only or passive only) displayed a significant correlation with ERA5. Also, the SMAP+ASCAT and AMSR2+ASCAT datasets were produced by using an unweighted average scenario in the case of the significant correlations for both parent SM products with GLDAS, but not with each other, showed a high overall correlation with ERA5. At high latitudes, the two merged products produced by using an unweighted average scenario when both active and passive correlate insignificantly with GLDAS, but correlate significantly with each other, showed a much lower overall correlation value ERA5. The pixels at which the correlations between all three SM of TCM are insignificant with each other were excluded from SMAP+ASCAT and AMSR2+ASCAT products. Where these pixels also showed an insignificant correlation against the ERA5 dataset by applying the scenarios of both using a single sensor and an unweighted mean method.

5. Discussion

In this research, we introduce a new scheme that takes advantage of both active and passive satellite SM products to produce improved SM product. This scheme is based on maximizing signal to noise ratios by using least-squares theory. The proposed scheme does not require a specific land surface model in combination process like other previous studies, which calculated the relative weights of parent satellite SM datasets, based on increasing correlations or decreasing error with specific land surface model [13,17,18,75]. TCM enables us to estimates the sensitivity of the datasets to SM changes in two terms SNR or fMSE. The weights in the proposed scheme are calculated based on fMSE which has a specific range values (i.e., from 0 to 1) not like the SNR term [20,68]. The proposed scheme was applied to SM retrieval from SMAP, AMSR2, and ASCAT datasets to produce SMAP+ASCAT and AMSR2+ASCAT SM datasets.

The proposed scheme maintains the TCM assumptions and not violated it. These assumptions include the errors independency, the availability of datasets in the long temporal period, and stationarity of signal and error estimates [20,64]. To maintain on the assumption of independent errors we selected triplet items with different derivation because the similarly derived datasets may have partially correlated errors [13,14]. This was achieved by calculating TCM twice: Once with ASCAT, SMAP, and GLDAS and once with ASCAT, AMSR2 and GLDAS. While, the assumption of long time series is necessary to reduce the sampling errors of TCM estimates. It is noted that the better results of TCM estimates could be obtained with long time series datasets [75]. This was achieved in our study by applying the proposed scheme using datasets in the period from June 2015 to December 2017 (two years and a half). In this study, we calculated weights based on the stationarity of signal over the working time, and the merged datasets showed better results with reference datasets, as we demonstrated above. However, further studies are recommended, as working studies during different seasons may

give better results due to seasonal variations (i.e., rainfall and temperature changes over most regions of the world).

In the proposed scheme, the pixels of high retrieval SM quality are only used from the SMAP, AMSR2, and ASCAT SM products. This was achieved by masking these datasets according to retrieval quality flags as mentioned in Section 2 before merging it. Also, the used threshold at fMSE equal to 0.5 increased the performance of merged products, where this threshold helps us to select the pixels of high quality before the merging process.

The proposed scheme maintains on the variability and seasonal dynamics of SM for the parents and reanalysis references SM products. Where SMAP+ASCAT and AMSR2+ASCAT merged, the products showed the same spatiotemporal variability of the parents and reference datasets.

In SMAP+ASCAT and AMSR2+ASCAT merged products, the ASCAT product was given a high relative weight than SMAP over more dense vegetation (forests, savannas, and crops), which mostly spatially distributed in the high-temperature regions below and above the equator. This is because the ASCAT has a high signal to noise ratios than SMAP and AMSR2 products over these regions. Also, these results agree with previous studies that found the performance of C-band active SM retrievals exceeds the passive SM retrievals in the areas with dense vegetation. Moreover, in extensively high-temperature areas, the active sensors displayed less susceptibility to surface temperature than passive sensors. Therefore, over more dense vegetation regions, the ASCAT retrievals showed less sensitivity to surface temperature change [13,14,23,63,76]. On the other hand, the SMAP and AMSR2 datasets were given high relative weights in SMAP+ASCAT and AMSR2+ASCAT merged datasets than ASCAT product over the areas of moderate and less vegetation (grassland, shrub, and desert). That is because the SMAP and AMSR2 showed high signal to noise ratios over these areas than the ASCAT product. However, previous studies found that the passive SM datasets have better performance over less and moderate vegetation with achieving the highest signal to noise ratios and the lowest error variance than an active sensor [13,14]. The SMAP was given a high relative weight than AMSR2 in their merged datasets because the SMAP has better performance than AMSR2 with achieving a high signal to noise ratios. However, these results agree with previous studies that found that the SMAP SM retrievals are outperformed generally the AMSR2 SM data [2,13,15].

The SMAP+ASCAT and AMSR2+ASCAT products showed significant improvements with achieving the highest R and lowest RMSE values with ground SM measurements over different networks of ISMN than parents products. These improvements provide evidence about the reliability of the proposed scheme. The ground measurements have limitations in spatial and temporal coverage in the world. Therefore, the best alternative to a full evaluation of SM datasets in the world is comparing it with reanalysis datasets [43]. GLDAS and ERA5 atmospheric reanalysis datasets were used for this purpose. The SMAP+ASCAT and AMSR2+ASCAT datasets were achieved a high overall average temporal R-value and less overall average RMSE value with GLDAS and ERA5 references datasets, which exceeds the parents products s over different land cover types in most regions of the world. Demonstrating the proposed scheme assures improved performance. The correlations improvements spatially distributed over different land cover types of all most areas of the world according to the spatial weights patterns of parents' datasets. Over more dense vegetation regions, the SMAP and AMSR2 are given the lowest relative weight than ASCAT products in their merged products, therefore, displayed high R differences with their merged datasets. Conversely, over moderate and less vegetation regions the ASCAT was given the lowest relative weights than SMAP and AMSR2 products in their merged products. Therefore, high differences, R, with their merged products were demonstrated. These improvements change according to the weight differences between parents' datasets, which provide evidence for the accuracy of the estimated weights.

Although, the GLDAS is used as an item in three datasets of TCM to calculate fMSE, the SMAP+ASCAT and AMSR2+ASCAT products showed better performance with ERA5 exceeds the performance with GLDAS, which provides evidence about the independently of merged datasets. In recent studies [48,49,57], the different generations of commonly used global model products

were evaluated with observations on global and regional scales. The results showed that the ERA5 atmospheric reanalysis outperforms other models by achieving a high correlation and less RMSE with ground measurements. Therefore, the high correlation of merged SM datasets with ERA5 gives an advantage for the proposed scheme.

After applying the weighted-average method in the regions which achieving significant correlations between all three SM datasets of TCM we proposed four scenarios in the regions of unreliable triple collocation analysis, based on the determinations of correlations between all three SM products of TCM at significance levels (*p*-values < 0.05). These scenarios include, either using a passive SM dataset only, or using an active SM dataset only, using an unweighted average method or excluding the pixels without any potential skills. Since the errors of the reference product are not correlated with those of the satellite retrievals, hence, the quality of the reanalysis dataset does not affect relative calculated correlations [63,77]. Therefore, the independent ERA5 reference product was used to make sure from the reliability of the selected scenario according to the suggested criteria. The two merged products produced by using the scenarios of using the single sensors (i.e., either active only or passive only), or by using an unweighted average scenario in the case of the significant correlations for both parent SM products with GLDAS, but not with each other provides prove the proposed scenarios assures the performance, where these products displayed a significant correlation with ERA5. The two merged products were produced by using an unweighted average scenario when both active and passive correlate insignificantly with GLDAS, but where they correlated significantly with each other, they showed a much lower overall correlation value with ERA5 at high latitudes. The reasons for these may be due to the unreliability of GLDAS and ERA5 in these regions. In addition, it is known that the land surface models are characterized by poor quality at very high latitudes [78]. For these reasons, we kept the pixels at these regions and consider the significant correlation between active and passive products as an indication the parents products, perhaps still contain valuable SM retrieval data even though the very low correlation with the ERA5 reference dataset. Therefore, we applied an unweighted average scenario in these regions. To obtain high-quality merged SM products we excluded the pixels which don't have any potential skills. These are achieved by excluding the pixels from SMAP+ASCAT and AMSR2+ASCAT products at which the correlations between all three SM of TCM are insignificant with each other. In addition to, these pixels displayed an insignificant correlation against the ERA5 product by applying the scenarios of both using a single sensor and an unweighted mean method.

As discussed in this section the novel proposed scheme improved the performance of satellite SM retrieval over different land cover in most regions of the world.

6. Conclusions

In this study, we developed and evaluated a new merging scheme, based on maximizing signal-to-noise ratios that take the advantage of both active and passive satellite soil moisture (SM) products to produce an improved SM product. Where the fractional mean-squared-error (fMSE) is derived from the triple collocation method (TCM) is used to estimate weight, based on the least-squares theory. The proposed scheme was applied by using a threshold between signal and noise at fMSE equal to 0.5 to maintain the high-quality SM observations. In the regions where TCM is unreliable, we proposed four scenarios based on the determinations of correlations between all three SM products of TCM at significance levels (i.e., *p*-values < 0.05). The proposed scheme was applied to SM retrievals from ASCAT, SMAP, and AMSR2 at a global scale in the period from June 2015 to December 2017 to produce SMAP+ASCAT and AMSR2+ASCAT SM datasets. The performances of merged SM datasets were assessed against independent reference datasets include in-situ measurements, GLDAS, and ERA5. The proposed scheme maintains on the variability and seasonal dynamics of SM for the parents and reanalysis references SM products. The merged SM datasets showed better performance with a high average temporal R with in-situ measurements over different networks of ISMN than parents products. Moreover, these datasets showed significant improvements with achieving a high average temporal R value and less RMSE value with GLDAS and ERA5 datasets which exceeds the parents' datasets over

different land cover types in most regions of the world. In the regions of unreliable TCM, the suggested four scenarios improved the performance of SM retrievals. The proposed scheme has the potential to be applied existing microwave satellites as well as to new missions.

Further studies are recommended to calculate the relative weights of parent SM products, based on maximized signal-to-noise ratios, at different seasonal scales, and it maybe gives better results due to seasonal variations. Also, studies need to be performed on the possibilities of merging SM retrievals from more than two satellites. Finally, TCM has been applied successfully in different fields include oceanography, hydrometeorology, and ecology [79–82]. Therefore, our proposed demonstrated scheme for SM has flexibility and can be applied to other biogeophysical variables.

Author Contributions: Conceptualization, B.G.M.; methodology, B.G.M.; software, B.G.M; validation, B.G.M.; formal analysis, B.G.M.; investigation, B.G.M.; resources, B.G.M. and M.F.; data curation, B.G.M., and A.T.; writing—original draft preparation, B.G.M.; writing—review and editing, B.G.M., H.S., M.F. and A.T.; visualization, B.G.M.; supervision, B.G.M., H.S. All authors have read and agreed to the published version of the manuscript.

References

1. González-Zamora, Á.; Sánchez, N.; Pablos, M.; Martínez-Fernández, J. CCI soil moisture assessment with SMOS soil moisture and in situ data under different environmental conditions and spatial scales in Spain. *Remote Sens. Environ.* **2019**, *225*, 469–482. [CrossRef]

2. Zhang, X.; Zhang, T.; Zhou, P.; Shao, Y.; Gao, S.; Zhang, X.; Zhang, T.; Zhou, P.; Shao, Y.; Gao, S. Validation Analysis of SMAP and AMSR2 Soil Moisture Products over the United States Using Ground-Based Measurements. *Remote Sens.* **2017**, *9*, 104. [CrossRef]

3. Rajib, M.A.; Merwade, V.; Yu, Z. Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture. *J. Hydrol.* **2016**, *536*, 192–207. [CrossRef]

4. Shi, H.; Yang, J.; Li, P.; Zhao, L.; Liu, Z.; Zhao, J.; Liu, W. Soil moisture estimation using two-component decomposition and a hybrid X-Bragg/Fresnel scattering model. *J. Hydrol.* **2019**, *574*, 646–659. [CrossRef]

5. Al-Yaari, A.; Wigneron, J.P.; Dorigo, W.; Colliander, A.; Pellarin, T.; Hahn, S.; Mialon, A.; Richaume, P.; Fernandez-Moran, R.; Fan, L.; et al. Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements. *Remote Sens. Environ.* **2019**, *224*, 289–303. [CrossRef]

6. Li, Y.; Shu, H.; Mousa, B.G.; Jiao, Z. Novel soil moisture estimates combining the ensemble Kalman filter data assimilation and the method of breeding growing modes. *Remote Sens.* **2020**, *12*, 889. [CrossRef]

7. Wagner, W.; Dorigo, W.; de Jeu, R.; Fernandez, D.; Benveniste, J.; Haas, E.; Ertl, M. Fusion of Active and Passive Microwave Observations to Create AN Essential Climate Variable Data Record on Soil Moisture. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, 315–321. [CrossRef]

8. Kolassa, J.; Gentine, P.; Prigent, C.; Aires, F. Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 1: Satellite data analysis. *Remote Sens. Environ.* **2016**, *173*, 1–14. [CrossRef]

9. Njoku, E.G.; Jackson, T.J.; Lakshmi, V.; Chan, T.K.; Nghiem, S.V. Soil moisture retrieval from AMSR-E. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 215–228. [CrossRef]

10. Wu, X.; Lu, G.; Wu, Z.; He, H.; Scanlon, T.; Dorigo, W. Triple Collocation-Based Assessment of Satellite Soil Moisture Products with In Situ Measurements in China: Understanding the Error Sources. *Remote Sens.* **2020**, *12*, 2275. [CrossRef]

11. Cui, C.; Xu, J.; Zeng, J.; Chen, K.S.; Bai, X.; Lu, H.; Chen, Q.; Zhao, T. Soil moisture mapping from satellites: An intercomparison of SMAP, SMOS, FY3B, AMSR2, and ESA CCI over two dense network regions at different spatial scales. *Remote Sens.* **2018**, *10*, 33. [CrossRef]

12. Gruhier, C.; de Rosnay, P.; Hasenauer, S.; Holmes, T.; de Jeu, R.; Kerr, Y.; Mougin, E.; Njoku, E.; Timouk, F.; Wagner, W.; et al. Soil moisture active and passive microwave products: Intercomparison and evaluation over a Sahelian site. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 141–156. [CrossRef]

13. Kim, H.; Parinussa, R.; Konings, A.G.; Wagner, W.; Cosh, M.H.; Lakshmi, V.; Zohaib, M.; Choi, M. Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. *Remote Sens. Environ.* **2018**, *204*, 260–275. [CrossRef]

14. Mousa, B.G.; Shu, H. Spatial Evaluation and Assimilation of SMAP, SMOS, and ASCAT Satellite Soil Moisture Products Over Africa Using Statistical Techniques. *Earth Sp. Sci.* **2020**, *7*, 1–16. [CrossRef]

15. Ma, H.; Zeng, J.; Chen, N.; Zhang, X.; Cosh, M.H.; Wang, W. Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: A comprehensive assessment using global ground-based observations. *Remote Sens. Environ.* **2019**, *231*. [CrossRef]

16. Albergel, C.; de Rosnay, P.; Gruhier, C.; Muñoz-Sabater, J.; Hasenauer, S.; Isaksen, L.; Kerr, Y.; Wagner, W. Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. *Remote Sens. Environ.* **2012**, *118*, 215–226. [CrossRef]

17. Kim, S.; Parinussa, R.M.; Liu, Y.Y.; Johnson, F.M.; Sharma, A. A framework for combining multiple soil moisture retrievals based on maximizing temporal correlation. *Geophys. Res. Lett.* **2015**, *42*, 6662–6670. [CrossRef]

18. Khan, M.Z.K.; Mehrotra, R.; Sharma, A.; Sankarasubramanian, A. Global sea surface temperature forecasts using an improved multimodel approach. *J. Clim.* **2014**, *27*, 3505–3515. [CrossRef]

19. Su, C.H.; Ryu, D.; Crow, W.T.; Western, A.W. Stand-alone error characterisation of microwave satellite soil moisture using a Fourier method. *Remote Sens. Environ.* **2014**, *154*, 115–126. [CrossRef]

20. Gruber, A.; Su, C.-H.; Zwieback, S.; Crow, W.; Dorigo, W.; Wagner, W. Recent advances in (soil moisture) triple collocation analysis. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *45*, 200–211. [CrossRef]

21. Zeng, Y.; Su, Z.; Van Der Velde, R.; Wang, L.; Xu, K.; Wang, X.; Wen, J. Blending satellite observed, model simulated, and in situ measured soil moisture over Tibetan Plateau. *Remote Sens.* **2016**, *8*, 268. [CrossRef]

22. Zhuang, R.; Zeng, Y.; Manfreda, S.; Su, Z. Quantifying long-term land surface and root zone soil moisture over Tibetan plateau. *Remote Sens.* **2020**, *12*, 509. [CrossRef]

23. Al-Yaari, A.; Wigneron, J.P.; Ducharne, A.; Kerr, Y.H.; Wagner, W.; De Lannoy, G.; Reichle, R.; Al Bitar, A.; Dorigo, W.; Richaume, P.; et al. Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land). *Remote Sens. Environ.* **2014**, *152*, 614–626. [CrossRef]

24. Bankman, I.N. *Handbook of Medical Image Processing and Analysis*; Academic Press: San Diego, CA, USA, 2009.

25. Rüdiger, C.; Calvet, J.C.; Gruhier, C.; Holmes, T.R.H.; de Jeu, R.A.M.; Wagner, W. An intercomparison of ERS-Scat and AMSR-E soil moisture observations with model simulations over France. *J. Hydrometeorol.* **2009**, *10*, 431–447. [CrossRef]

26. Zeng, J.; Chen, K.-S.; Bi, H.; Chen, Q. A Preliminary Evaluation of the SMAP Radiometer Soil Moisture Product Over United States and Europe Using Ground-Based Measurements. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4929–4940. [CrossRef]

27. Colliander, A.; Jackson, T.J.; Bindlish, R.; Chan, S.; Das, N.; Kim, S.B.; Cosh, M.H.; Dunbar, R.S.; Dang, L.; Pashaian, L.; et al. Validation of SMAP surface soil moisture products with core validation sites. *Remote Sens. Environ.* **2017**, *191*, 215–231. [CrossRef]

28. Entekhabi, D.; Njoku, E.G.; O'Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The soil moisture active passive (SMAP) mission. *Proc. IEEE* **2010**, *98*, 704–716. [CrossRef]

29. Das, N.N.; Dunbar, R.S. *Level 3 Active/Passive Soil Moisture Product Specification Document*; California Institute of Technology: Pasadena, CA, USA, 2015.

30. Chan, S.K.; Bindlish, R.; O'Neill, P.E.; Njoku, E.; Jackson, T.; Colliander, A.; Chen, F.; Burgin, M.; Dunbar, S.; Piepmeier, J.; et al. Assessment of the SMAP Passive Soil Moisture Product. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4994–5007. [CrossRef]

31. Kachi, M.; Naoki, K.; Hori, M.; Imaoka, K. AMSR2 validation results. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Melbourne, Australia, 21–26 July 2013; pp. 831–834.

32. Imaoka, K.; Kachi, M.; Fujii, H.; Murakami, H.; Hori, M.; Ono, A.; Igarashi, T.; Nakagawa, K.; Oki, T.; Honda, Y.; et al. Global change observation mission (GCOM) for monitoring carbon, water cycles, and climate change. *Proc. IEEE* **2010**, *98*, 717–734. [CrossRef]

33. Imaoka, K.; Maeda, T.; Kachi, M.; Kasahara, M.; Ito, N.; Nakagawa, K. Status of AMSR2 instrument on GCOM-W1. In *Proceedings of the Earth Observing Missions and Sensors: Development, Implementation, and Characterization II*; Shimoda, H., Xiong, X., Cao, C., Gu, X., Kim, C., Kiran Kumar, A.S., Eds.; SPIE: Bellingham, WA, USA, 2012; Volume 8528, p. 852815.

34. van der Schalie, R.; de Jeu, R.A.M.; Kerr, Y.H.; Wigneron, J.P.; Rodríguez-Fernández, N.J.; Al-Yaari, A.; Parinussa, R.M.; Mecklenburg, S.; Drusch, M. The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E. *Remote Sens. Environ.* **2017**, *189*. [CrossRef]

35. Wagner, W.; Hahn, S.; Kidd, R.; Melzer, T.; Bartalis, Z.; Hasenauer, S.; Figa-Saldaña, J.; De Rosnay, P.; Jann, A.; Schneider, S.; et al. The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorol. Z.* **2013**, *22*, 5–33. [CrossRef]

36. *PUM Product User Manual (PUM) Metop ASCAT Soil Moisture CDR and Offline Products*; Issue/Revision 0.8; EUMETSAT: Darmstadt, Germany, 2018; p. 26.

37. Chen, F.; Crow, W.T.; Bindlish, R.; Colliander, A.; Burgin, M.S.; Asanuma, J.; Aida, K. Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. *Remote Sens. Environ.* **2018**, *214*, 1–13. [CrossRef] [PubMed]

38. Draper, C.; Mahfouf, J.-F.; Calvet, J.-C.; Martin, E.; Wagner, W. Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3829–3841. [CrossRef]

39. Brocca, L.; Melone, F.; Moramarco, T.; Wagner, W.; Naeimi, V.; Bartalis, Z.; Hasenauer, S. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1881–1893. [CrossRef]

40. Dorigo, W.A.; Wagner, W.; Hohensinn, R.; Hahn, S.; Paulik, C.; Xaver, A.; Gruber, A.; Drusch, M.; Mecklenburg, S.; van Oevelen, P.; et al. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1675–1698. [CrossRef]

41. Miralles, D.G.; Crow, W.T.; Cosh, M.H. Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. *J. Hydrometeorol.* **2010**, *11*, 1423–1429. [CrossRef]

42. Gruber, A.; Dorigo, W.A.; Zwieback, S.; Xaver, A.; Wagner, W. Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network. *Vadose Zone J.* **2013**, *12*. [CrossRef]

43. Leroux, D.J.; Kerr, Y.H.; Al Bitar, A.; Bindlish, R.; Jackson, T.J.; Berthelot, B.; Portet, G. Comparison Between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products Over Four Watersheds in U.S. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1562–1571. [CrossRef]

44. Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394. [CrossRef]

45. Liu, Y.Y.; Parinussa, R.M.; Dorigo, W.A.; De Jeu, R.A.M.; Wagner, W.; Van Dijk, A.I.J.M.; McCabe, M.F.; Evans, J.P. Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 425–436. [CrossRef]

46. Liu, Y.Y.; Dorigo, W.A.; Parinussa, R.M.; De Jeu, R.A.M.; Wagner, W.; McCabe, M.F.; Evans, J.P.; Van Dijk, A.I.J.M. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.* **2012**, *123*, 280–297. [CrossRef]

47. Beck, H.E.; Pan, M.; Miralles, D.G.; Reichle, R.H.; Dorigo, W.A.; Hahn, S.; Sheffield, J.; Karthikeyan, L.; Balsamo, G.; Parinussa, R.M.; et al. Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors. *Hydrol. Earth Syst. Sci. Discuss.* **2020**, 1–35. [CrossRef]

48. Li, M.; Wu, P.; Ma, Z. Comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis datasets. *Int. J. Climatol.* **2020**. [CrossRef]

49. Albergel, C.; Munier, S.; Calvet, J.; Dutra, E.; Munoz-Sabater, J.; DeRosnay, P.; Balsamo, G. ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? *Hydrol. Earth Syst. Sci.* **2018**, *22*. [CrossRef]

50. Zeng, J.; Li, Z.; Chen, Q.; Bi, H.; Qiu, J.; Zou, P. Evaluation of remotely sensed and reanalysis soil moisture products over the Tibetan Plateau using in-situ observations. *Remote Sens. Environ.* **2015**, *163*, 91–110. [CrossRef]

51. Larson, K.M.; Small, E.E.; Gutmann, E.D.; Bilich, A.L.; Braun, J.J.; Zavorotny, V.U.; Larson, C.; Small, E.E.; Gutmann, E.D.; Bilich, A.L.; et al. Use of GPS receivers as a soil moisture network for water cycle studies. *Geophys. Res. Lett.* **2008**. [CrossRef]

52. Schaefer, G.L.; Cosh, M.H.; Jackson, T.J. The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). *J. Atmos. Ocean. Technol.* **2007**, *24*, 2073–2077. [CrossRef]

53. Bell, J.E.; Palecki, M.A.; Baker, C.B.; Collins, W.G.; Lawrimore, J.H.; Leeper, R.D.; Hall, M.E.; Kochendorfer, J.; Meyers, T.P.; Wilson, T.; et al. US climate reference network soil moisture and temperature observations. *J. Hydrometeorol.* **2013**, *14*, 977–988. [CrossRef]

54. Albergel, C.; Rüdiger, C.; Pellarin, T.; Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Piguet, B.; Martin, E. From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 1323–1337. [CrossRef]

55. Jensen, K.H.; Illangasekare, T.H. HOBE: A Hydrological Observatory. *Vadose Zone J.* **2011**, *10*, 1–7. [CrossRef]

56. Smith, A.B.; Walker, J.P.; Western, A.W.; Young, R.I.; Ellett, K.M.; Pipunic, R.C.; Grayson, R.B.; Siriwardena, L.; Chiew, F.H.S.; Richter, H. The Murrumbidgee Soil Moisture Monitoring Network data set. *Water Resour. Res.* **2012**, *48*. [CrossRef]

57. Hagan, D.F.T.; Parinussa, R.M.; Wang, G.; Draper, C.S. An Evaluation of Soil Moisture Anomalies from Global Model-Based Datasets over the People's Republic of China. *Water* **2019**, *12*, 117. [CrossRef]

58. Tariq, A.; Shu, H. CA-Markov Chain Analysis of Seasonal Land Surface Temperature and Land Use Land Cover Change Using Optical Multi-Temporal Satellite Data of Faisalabad, Pakistan. *Remote Sens.* **2020**, *12*, 3402. [CrossRef]

59. Karl Friedrich Gauss, D.S. Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections. *Math. Comput.* **1964**, *18*, 525. [CrossRef]

60. Legendre, A.M. *Nouvelles Méthodes Pour la Détermination des Orbites des Comètes*; Chez FIRMIN DIDOT, Libraire Pour Les Mathématiques, la Marine, l'Architecture, et les Éditions Stéréotypes, Rue de Thionville: Paris, France, 1805.

61. Crow, W.T.; Su, C.-H.; Ryu, D.; Yilmaz, M.T. Optimal averaging of soil moisture predictions from ensemble land surface model simulations. *Water Resour. Res.* **2015**, *51*, 9273–9289. [CrossRef]

62. Yilmaz, M.T.; Crow, W.T.; Anderson, M.C.; Hain, C. An objective methodology for merging satellite- and model-based soil moisture products. *Water Resour. Res.* **2012**, *48*. [CrossRef]

63. Dorigo, W.A.; Scipal, K.; Parinussa, R.M.; Liu, Y.Y.; Wagner, W.; de Jeu, R.A.M.; Naeimi, V. Error characterisation of global active and passive microwave soil moisture datasets. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 2605–2616. [CrossRef]

64. Stoffelen, A. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *J. Geophys. Res. Ocean.* **1998**, *103*, 7755–7766. [CrossRef]

65. Scipal, K.; Holmes, T.; De Jeu, R.; Naeimi, V.; Wagner, W. A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophys. Res. Lett.* **2008**, *35*. [CrossRef]

66. Chen, F.; Crow, W.T.; Colliander, A.; Cosh, M.H.; Jackson, T.J.; Bindlish, R.; Reichle, R.H.; Chan, S.K.; Bosch, D.D.; Starks, P.J.; et al. Application of Triple Collocation in Ground-Based Validation of Soil Moisture Active/Passive (SMAP) Level 2 Data Products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*. [CrossRef]

67. Tugrul Yilmaz, M.; Crow, W.T. The optimality of potential rescaling approaches in land data assimilation. *J. Hydrometeorol.* **2013**, *14*, 650–660. [CrossRef]

68. Draper, C.; Reichle, R.; de Jeu, R.; Naeimi, V.; Parinussa, R.; Wagner, W. Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. *Remote Sens. Environ.* **2013**, *137*, 288–298. [CrossRef]

69. Draper, C.S.; Walker, J.P.; Steinle, P.J.; de Jeu, R.A.M.; Holmes, T.R.H. An evaluation of AMSR-E derived soil moisture over Australia. *Remote Sens. Environ.* **2009**, *113*, 703–710. [CrossRef]

70. Schlosser, C.A.; Milly, P.C.D. A model-based investigation of soil moisture predictability and associated climate predictability. *J. Hydrometeorol.* **2002**, *3*, 483–501. [CrossRef]

71. Miyaoka, K.; Gruber, A.; Ticconi, F.; Hahn, S.; Wagner, W.; Figa-Saldana, J.; Anderson, C. Triple Collocation Analysis of Soil Moisture from Metop-A ASCAT and SMOS Against JRA-55 and ERA-Interim. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2274–2284. [CrossRef]

72. Blunden, J.D.S.; Arndt, E. State of the climate in 2015. *Bull. Am. Meteorol. Soc.* **2016**, *97*. [CrossRef]

73. Hovmöller, E. The Trough-and-Ridge diagram. *Tellus* **1949**, *1*, 62–66. [CrossRef]

74. Entekhabi, D.; Reichle, R.H.; Koster, R.D.; Crow, W.T.; Entekhabi, D.; Reichle, R.H.; Koster, R.D.; Crow, W.T. Performance Metrics for Soil Moisture Retrievals and Application Requirements. *J. Hydrometeorol.* **2010**, *11*, 832–840. [CrossRef]

75. Hagan, D.F.T.; Wang, G.; Kim, S.; Parinussa, R.M.; Liu, Y.; Ullah, W.; Bhatti, A.S.; Ma, X.; Jiang, T.; Su, B. Maximizing temporal correlations in long-term global satellite soil moisture data-merging. *Remote Sens.* **2020**, *12*, 2164. [CrossRef]

76. Scipal, K.; Dorigo, W.; De Jeu, R. Triple collocation-A new tool to determine the error structure of global soil moisture products. *Int. Geosci. Remote Sens. Symp.* **2010**, *3*, 4426–4429. [CrossRef]

77. Gruber, A.; Crow, W.; Dorigo, W.; Wagner, W. The potential of 2D Kalman filtering for soil moisture data assimilation. *Remote Sens. Environ.* **2015**, *171*, 137–148. [CrossRef]

78. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A.; et al. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [CrossRef]

79. D'Odorico, P.; Gonsamo, A.; Pinty, B.; Gobron, N.; Coops, N.; Mendez, E.; Schaepman, M.E. Intercomparison of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. *Remote Sens. Environ.* **2014**, *142*, 141–154. [CrossRef]

80. Roebeling, R.A.; Wolters, E.L.A.; Meirink, J.F.; Leijnse, H. Triple collocation of summer precipitation retrievals from SEVIRI over europe with gridded rain gauge and weather radar data. *J. Hydrometeorol.* **2012**, *13*, 1552–1566. [CrossRef]

81. Fang, H.; Wei, S.; Jiang, C.; Scipal, K. Theoretical uncertainty analysis of global MODIS, CYCLOPES, and GLOBCARBON LAI products using a triple collocation method. *Remote Sens. Environ.* **2012**, *124*, 610–621. [CrossRef]

82. Caires, S.; Sterl, A. Validation of ocean wind and wave data using triple collocation. *J. Geophys. Res. Ocean.* **2003**, *108*, 1–16. [CrossRef]

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