# **@AGU**PUBLICATIONS

# Journal of Geophysical Research: Atmospheres

# **RESEARCH ARTICLE**

10.1002/2014JD021489

#### **Key Points:**

- New method to measure terrestrial rainfall from satellite soil moisture data
- Global-scale application by using three satellite soil moisture products
- The new satellite products estimate accurately the accumulated rainfall

#### **Supporting Information:**

Readme

## Figures S1–S8

#### Correspondence to:

L. Brocca, luca.brocca@irpi.cnr.it

#### Citation:

Brocca, L., L. Ciabatta, C. Massari, T. Moramarco, S. Hahn, S. Hasenauer, R. Kidd, W. Dorigo, W. Wagner, and V. Levizzani (2014), Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, *J. Geophys. Res. Atmos.*, *119*, 5128–5141, doi:10.1002/ 2014JD021489.

Received 13 JAN 2014 Accepted 14 APR 2014 Accepted article online 15 APR 2014 Published online 6 MAY 2014

# Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data

Luca Brocca<sup>1</sup>, Luca Ciabatta<sup>1</sup>, Christian Massari<sup>1</sup>, Tommaso Moramarco<sup>1</sup>, Sebastian Hahn<sup>2</sup>, Stefan Hasenauer<sup>2</sup>, Richard Kidd<sup>2</sup>, Wouter Dorigo<sup>2</sup>, Wolfgang Wagner<sup>2</sup>, and Vincenzo Levizzani<sup>3</sup>

<sup>1</sup>Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy, <sup>2</sup>Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria, <sup>3</sup>Institute of Atmospheric Sciences and Climate, National Research Council, Bologna, Italy

JGR

Abstract Measuring precipitation intensity is not straightforward; and over many areas, ground observations are lacking and satellite observations are used to fill this gap. The most common way of retrieving rainfall is by addressing the problem "top-down" by inverting the atmospheric signals reflected or radiated by atmospheric hydrometeors. However, most applications are interested in how much water reaches the ground, a problem that is notoriously difficult to solve from a top-down perspective. In this study, a novel "bottom-up" approach is proposed that, by doing "hydrology backward," uses variations in soil moisture (SM) sensed by microwave satellite sensors to infer preceding rainfall amounts. In other words, the soil is used as a natural rain gauge. Three different satellite SM data sets from the Advanced SCATterometer (ASCAT), the Advanced Microwave Scanning Radiometer (AMSR-E), and the Microwave Imaging Radiometer with Aperture Synthesis are used to obtain three new daily global rainfall products. The "First Guess Daily" product of the Global Precipitation Climatology Centre (GPCC) is employed as main benchmark in the validation period 2010–2011 for determining the continuous and categorical performance of the SM-derived rainfall products by considering the 5 day accumulated values. The real-time version of the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis product, i.e., the TRMM-3B42RT, is adopted as a state-of-the-art satellite rainfall product. The SM-derived rainfall products show good Pearson correlation values (R) with the GPCC data set, mainly in areas where SM retrievals are found to be accurate. The global median R values (in the latitude band  $\pm 50^\circ$ ) are equal to 0.54, 0.28, and 0.31 for ASCAT-, AMSR-E-, and SMOS-derived products, respectively. For comparison, the median R for the TRMM-3B42RT product is equal to 0.53. Interestingly, the SM-derived products are found to outperform TRMM-3B42RT in terms of average global root-mean-square error statistics and in terms of detection of rainfall events. The regions for which the SM-derived products perform very well are Australia, Spain, South and North Africa, India, China, the Eastern part of South America, and the central part of the United States. The SM-derived products are found to estimate accurately the rainfall accumulated over a 5 day period, an aspect particularly important for their use for hydrological applications, and that address the difficulties of estimating light rainfall from TRMM-3B42RT.

### 1. Introduction

Accurate estimates of rainfall are of vital importance for mitigation strategies of natural hazards such as floods and landslides [*Keefer et al.*, 1987; *Hong et al.*, 2007; *Wake*, 2013; *Hou et al.*, 2013] as well as for disease and famine prevention [*Scholthof*, 2006] and many other applications. Nowadays, satellite rainfall products such as the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis product (TMPA) [*Huffman et al.*, 2007] are the main tools to address rainfall estimation on a global scale, particularly in data-poor world areas where ground observations are sparse [*Kidd and Levizzani*, 2011]. The deployment of the Global Precipitation Measurement mission satellite constellation, whose main spacecraft was launched on 27 February 2014, represents a critical advance in these efforts (http://pmm.nasa.gov/GPM) [*Hou et al.*, 2013]. However, the achievement of the accuracy levels required by the hydrological applications has proven difficult [*Hossain and Anagnostou*, 2004; *Uijlenhoet and Berne*, 2008; *Serrat-Capdevila et al.*, 2012]. Currently, one of the major issues for rainfall retrieval from space is related to the estimation of light rainfall that causes a general underestimation of rainfall accumulations [*Kucera et al.*, 2013]. These could be the main reasons for the continued prominence of in situ gauge and precipitation radar observations in the majority of applications.

An underexplored possibility for estimating rainfall lies in the more efficient use of complementary spaceborne water cycle observations [McCabe et al., 2008]. Among them, soil moisture (SM) shares an obvious and strong physical connection with rainfall. On the one hand, on a regional scale, antecedent SM has a significant impact on the rainfall-triggering mechanism [Koster and the GLACE Team, 2004; Taylor et al., 2012], mainly for convective storms. On the other hand, rainfall is the main driver of SM temporal variability locally [e.g., Crow et al., 2009]. After a rainfall pulse, SM exhibits a sudden increase followed by a smooth recession limb driven by evapotranspiration and drainage [Porporato et al., 2004; Brocca et al., 2014]. Based on this concept, some recent studies have shown that satellite SM data can be effectively employed for correcting satellite [Pellarin et al., 2008; Crow et al., 2009; Pellarin et al., 2013] and reanalysis [Chen et al., 2012] rainfall products. These frameworks need a first guess rainfall estimate (typically given by the real-time TMPA rainfall product) as the proposed algorithms are only able to correct rainfall by using SM data. More recently, Brocca et al. [2013] developed an approach (called SM2RAIN) based on the inversion of the hydrological water balance, for estimating, and not correcting, rainfall from SM observations. The soil is regarded as a natural reservoir used for measuring the amount of rainfall. The capability of SM2RAIN method to estimate, and not only correct, rainfall might be particularly useful in those regions for which satellite rainfall data are affected by higher errors or not available as, for instance, for the higher latitudes (>50°) in the case of TMPA products. Moreover, through SM2RAIN a fully independent rainfall product is obtained characterized by a different error structure that will allow an optimal integration with other satellite-based rainfall products through data assimilation techniques. The first application of this approach to three sites in Italy (Umbria), Spain (Remedhus), and France (Valescure) with in situ and satellite SM data has demonstrated its capability to estimate rainfall satisfactorily [Brocca et al., 2013]. Specifically, when hourly in situ SM observations were used, the correlation coefficient, R, with daily rainfall observations ranged between 0.900 and 0.945 for the three sites. When satellite data were considered, the R values against 5 day rainfall data were found to be equal to 0.80 and 0.78 for the site in Italy and Spain, respectively. These results encourage the application of the SM2RAIN method on a global scale. This is made possible by the availability of several global satellite SM products developed in the last decades with good spatial-temporal resolution and accuracy [e.g., Owe et al., 2008; Kerr et al., 2012a; Wagner et al., 2013]. Through the application of the SM2RAIN algorithm, these products can be effectively used for rainfall estimation on a global scale.

On this basis, this study proposes a bottom-up rainfall measurement approach, SM2RAIN, as a new way of monitoring rainfall from a satellite perspective. This approach is opposed to the classical top-down perspective that is based on the inversion of the atmospheric signals reflected or radiated by atmospheric hydrometeors. SM2RAIN is applied for terrestrial rainfall estimates on a global scale by using SM observations. The capability of SM2RAIN to retrieve rainfall accumulations is expected to provide improvements in our overall capability to estimate rainfall, complementing more classical approaches for estimating rainfall from space. Three satellite SM products are used (1) the MetOp Advanced SCATterometer (ASCAT) [*Wagner et al.*, 2013], (2) the NASA Advanced Microwave Scanning Radiometer for EOS (AMSR-E) [*Owe et al.*, 2008], and (3) the Soil Moisture and Ocean Salinity (SMOS) Microwave Imaging Radiometer with Aperture Synthesis [*Kerr et al.*, 2012b]. From each SM data set, a new SM-derived rainfall product is obtained through the application of the SM2RAIN algorithm. The quasi-global (±50° north-south latitude band) TRMM-3B42RT product is used in the comparison as state-of-the-art rainfall product.

To evaluate the overall accuracy of the rainfall estimates, three rainfall products available on a global scale and at daily time resolution are considered: the First Guess Daily product of the Global Precipitation Climatology Centre (GPCC) which is based on ground observations only [*Schamm et al.*, 2014], the One-Degree Daily precipitation data set of the Global Precipitation Climatology Project (GPCP), based on ground plus satellite data [*Huffman et al.*, 2001], and the modeled 3 h rainfall data obtained from the ERA-Interim (ERAI) reanalysis [*Dee et al.*, 2011]. The combined use of three products will allow a comprehensive and fair evaluation of the rainfall estimates obtained from SM2RAIN. For ASCAT and AMSR-E (SMOS), the period 2007–2009 (2012) is used for the calibration of SM2RAIN algorithm, the common period 2010–2011 is used for the validation of the rainfall products.

### 2. Soil Moisture and Rainfall Data Sets

#### 2.1. Satellite Soil Moisture Observations

As mentioned above, three satellite SM products are used in this study. All the satellite SM products are characterized by a nearly daily global coverage when both ascending and descending orbits are considered.

The MetOp ASCAT SM data are obtained from C band (5.255 GHz, VV polarization) backscatter observations using the TU-Wien algorithm water retrieval package at version 5.5 release 1.1 [*Wagner et al.*, 2013]. The algorithm is based on a time series change detection approach previously developed for the ERS-1/2 scatterometer by *Wagner et al.* [1999]. In this approach, SM is considered to have a linear relationship to backscatter in the decibel space, the surface roughness is assumed to have a constant contribution in time, and the typical yearly vegetation cycle is assumed to be known. The output of the algorithm is a time series of relative SM percentage values for the first few centimeters of the soil (2–3 cm). The product is available since 2007 with a spatial resolution of 25 km (resampled at 12.5 km).

The AMSR-E SM data set is obtained through the application of the VUA-NASA Land Parameter Retrieval Method (LPRM) [*Owe et al.*, 2008] at version 5 to C band (6.9 GHz) brightness temperature data. The LPRM algorithm is a three-parameter retrieval model (SM, vegetation optical depth, and soil/canopy temperature) that uses the dual-polarized channel (either 6.9 or 10.6 GHz) for the retrieval of surface SM and vegetation optical depth while the land surface temperature is derived separately from the vertically polarized 36.5 GHz channel. The product is expressed in volumetric terms ( $m^3/m^3$ ) and is representative of a 2–3 cm deep surface soil layer. The product is available from 2002 to October 2011 with a spatial resolution of 75 × 43 km (regridded to 0.25° regular grid).

For the SMOS data set [*Kerr et al.*, 2012a], the reprocessed level 2 (L2) data generated with the latest available version of the L2 algorithms (V5.51) are taken. The L Band Microwave Emission of the Biosphere Model [*Wigneron et al.*, 2007] is used for inverting SM from dual-polarized multiangular brightness temperature observations of SMOS. The algorithm details can be found in *Kerr et al.* [2012b]. The product is expressed in volumetric terms (m<sup>3</sup>/m<sup>3</sup>) and is representative of a soil layer of ~5 cm. The product is available since 2010 with a spatial resolution of 43 km.

For all the products, the quality flags supplied with the data are used to discard observations characterized by frozen surface, temporary water and snow conditions, and high probability of radio frequency interference (RFI). Moreover, retrievals characterized by a soil temperature (obtained from ERA-Interim reanalysis data) below 3°C are masked out. After masking, the temporal resolution of satellite SM products is significantly reduced in some regions and, hence, the 5 day accumulated rainfall data are considered in all the analyses.

#### 2.2. Rainfall Data Sets

Four different rainfall data sets derived from satellite sensors, ground observations, and reanalysis are used. For the sake of comparison, all the products are resampled at 1° spatial resolution and aggregated at 1 day temporal resolution.

The baseline satellite rainfall product is the real-time version of the TMPA product [*Huffman et al.*, 2007], i.e., the TRMM-3B42RT, available from 1997 onward with 3 h temporal resolution and a spatial resolution of 0.25° for the ±50° north-south latitude band. Several studies [e.g., *Ebert et al.*, 2007; *Huffman et al.*, 2007] have recognized that TRMM-3B42RT represents the state-of-the-art of the currently available satellite rainfall products. Therefore, this product is used as comparison for assessing the possible benefits deriving from the new SM-derived rainfall products.

The new First Guess Daily product provided by GPCC [*Schamm et al.*, 2014], which is available since 1 January 2009 with a spatial sampling grid of 1°, was employed as main independent benchmark for the validation period 2010–2011. The GPCP Daily Precipitation Data Set, version 1.2 [*Huffman et al.*, 2001], which merges data from over 6000 rain gauge stations with various types of satellite observations into a daily 1° resolution rainfall product from October 1996 to the present, is used as benchmark in the calibration period 2007–2009 (2012) for ASCAT and AMSR-E (SMOS) (as the GPCC product is not available before 2009). Finally, the ERA-Interim (ERAI) reanalysis rainfall data set [*Dee et al.*, 2011], characterized by a spatial resolution of 0.703°, and 3 h temporal resolution is employed as an additional benchmark data set. For all rainfall products, rainfall and snowfall variables are not distinguished in this study.

#### 3. Methods

#### 3.1. SM2RAIN Algorithm

The SM2RAIN algorithm is based on the inversion of the soil water balance equation for retrieving rainfall from SM data. The soil is assumed to work as a natural rain gauge for measuring the amount of rainfall fallen



Figure 1. Flowchart for the application of the SM2RAIN algorithm to satellite soil moisture data.

into the ground. Specifically, the soil water balance equation for a layer depth Z[L] can be described by the following expression:

$$Zds(t)/dt = p(t) - r(t) - e(t) - g(t)$$
<sup>(1)</sup>

where s(t) [-] is the relative saturation of the soil or relative SM, t [T] is the time and p(t), r(t), e(t), and g(t) [L/T] are the precipitation, runoff, evapotranspiration, and drainage rate, respectively. Whenever it rains, the evaporation rate can be safely assumed as negligible (e(t) = 0). Moreover, by assuming that all precipitation infiltrates into the soil, the runoff rate is zero (r(t) = 0). For the drainage rate, the following relation may be adopted,  $g(t) = as(t)^b$ , where a [L/T] and b [-] are two parameters expressing the nonlinearity between drainage rate and soil saturation. The rearrangement of equation (1), with the described assumptions, yields to

$$p(t) \cong Zds(t)/dt + as(t)^{b}$$
<sup>(2)</sup>

This equation can be used for estimating the precipitation rate from the knowledge of relative SM, s(t), its fluctuations in time, ds(t)/dt, and three parameters (*Z*, *a*, and *b*) to be estimated through calibration. Even though the assumptions made for deriving equation (2) might introduce some errors, they allow obtaining a simple (but effective) method for the retrieval of rainfall from SM data [*Brocca et al.*, 2013] that can be easily applied on a global scale.

#### 3.2. Rainfall Estimation Through SM2RAIN

The procedure used for estimating, through the SM2RAIN algorithm, daily rainfall at 1° resolution from each SM data set is straightforward and is schematized as (see Figure 1)

- Regrid SM data at 1° resolution by computing the average of the retrievals contained in each grid point. The data are rescaled between 0 and 1 by using the maximum and minimum measurement values as relative SM data should be used as input into the SM2RAIN algorithm and linearly interpolated at 0000 UTC of each day, for obtaining daily rainfall products equally distributed in time. A threshold of 4 days is used for avoiding to interpolate the data when SM observations are not available.
- 2. Classify land points in several classes (12 for this study) in accordance with the amount of rainfall that is estimated from the benchmark data set in the calibration period (GPCP in this study).
- 3. Calibrate the three SM2RAIN parameters (*Z*, *a*, and *b*) for each class by selecting a number of points (equal to 400 for this study) not significantly affected by noise in the SM retrievals (i.e., excluding the area covered by forest, desert, and characterized by complex topography). The GPCP product is used here for the parameters calibration. The minimization of the root-mean-square error (RMSE), averaged for each point of the rainfall class, is used as objective function; and the accumulated rainfall over 5 days is considered.

- 4. Run the SM2RAIN algorithm for all the land points with the optimized parameters for the whole period for which SM data are available.
- 5. Mask rainfall retrievals when daily soil temperature is below a certain threshold and when the SM data set is flagged (frozen surface, temporary water and snow conditions, and radio frequency interference).

We note here that several attempts are made for distributing the SM2RAIN parameters. Specifically, soil (from the Harmonized World Soil Database v1.2), land use, and climatic regimes (Köppen-Geiger climate classification) maps are selected but the best results are obtained by using rainfall classes. For the sake of brevity, only this analysis is reported in the manuscript. Future investigations are planned to identify more in detail the correlation between SM2RAIN parameters and other relevant properties at a finer spatial resolution (e.g., soil, land use, topography, and climatic regimes).

#### 3.3. Performance Metrics

Three continuous metrics are used for the evaluation of the performance of the SM-derived rainfall products (and TRMM-3B42RT): the correlation coefficient, *R*, the root-mean-square error, RMSE, and the mean error, BIAS, computed as percentage error between the accumulated rainfall of "observed" (GPCC, GPCP, and ERAI) and estimated data in the investigated period.

Moreover, three categorical metrics, widely used for evaluating the performance of satellite rainfall data, are also used: FAR (False Alarm Ratio), POD (Probability Of Detection), and TS (Threat Score). FAR refers to the fraction of predicted events that are actually nonevents and POD to the fraction of all qualifying events correctly predicted; TS provides an integrated measure of the overall performance. Through the definition of a rainfall accumulation "event" as 5 day rainfall accumulation in excess of a given threshold, the categorical metrics are defined as

$$FAR = \frac{F}{H+F}$$
(3)

$$\mathsf{POD} = \frac{H}{H+M} \tag{4}$$

$$TS = \frac{H}{H + F + M}$$
(5)

where *H* is the number of successfully predicted events by a given rainfall product, *F* the number of nonevents erroneously predicted to occur, and *M* the number of actual events that are missed. In this evaluation, an event is defined as a 5 day rainfall accumulation that exceeds a given percentile threshold of all 5 day accumulations observed for a given 1° pixel in the analyzed period [*Chen et al.*, 2012].

#### 4. Results

Following the procedure schematized in Figure 1 and described above, the SM2RAIN parameters for each satellite SM product are obtained. First, the three SM data sets are resampled on a regular 1° grid, thus obtaining 13,237 points over land. Second, by using GPCP product as benchmark (only for the calibration period), the land points are subdivided in 12 (nonequally spaced) classes as a function of the rainfall amount. Specifically, the following thresholds, in terms of annual rainfall (in mm), are considered for deriving 12 classes: 0–133, 134–267, 268–367, 368–467, 468–567, 568–667, 668–767, 768–867, 868–1067, 1068–1333, 1333–1833, and >1834. The number of points of each class is nearly uniformly distributed and ranges between 900 and 1400. Third, the SM data sets are splitted into a calibration (2007–2009 for ASCAT and AMSR-E and 2012 for SMOS) and a validation (2010–2011, common to all products) period. Note that (results not shown) calibrating ASCAT in 2012 provides nearly the same results. For each class, 400 points are selected where the SM2RAIN parameters are calibrated (separately for each SM product). Fourth, the algorithm is run for the whole period with the parameters spatially distributed according to the previously defined classes. Finally, rainfall retrievals are masked by considering the quality flags of each SM product and a soil temperature threshold of 3°C (from ERA-Interim reanalysis data). After masking, on average the percentage of missing data is equal to 14%, 18%, and 21% for ASCAT, AMSR-E, and SMOS products, respectively.

For the sake of simplicity, in the sequel we will refer to the three SM-derived rainfall products by using only the name of the satellite SM data set (i.e., ASCAT, AMSR-E, and SMOS) and to the TRMM-3B42RT product with

 Table 1. SM2RAIN Parameter Values (Z, a, and b) Obtained From the Calibration Against the GPCP Product in the Calibration Period 2007–2009 for ASCAT<sup>a</sup>

 Rainfall Classes (mm/yr)

	0–133	134–267	268–367	368–467	468–567	568–667	668–767	768–867	868–1067	1068–1333	1333–1833	>1834
<i>Z</i> (mm)	26.6	32.7	27.7	29.9	36.3	35.3	33.9	35.4	46.3	53.8	60.9	56.5
<i>a</i> (mm/d)	1.0	2.4	4.2	7.2	9.3	9.3	10.4	12.0	20.3	19.9	24.7	23.3
b (—)	23.2	1.5	1.8	3.0	3.4	2.8	2.4	2.4	2.6	2.3	2.5	1.4

<sup>a</sup>Parameters are subdivided in 12 classes as a function of the annual rainfall amount obtained from GPCP.

TRMM-RT. Analogously, the three benchmark data sets are referred to with their acronyms: GPCC, GPCP, and ERAI.

As an example, Table 1 shows the SM2RAIN parameter values (*Z*, *a*, and *b*) obtained from the calibration against GPCP product in the calibration period 2007–2009 for ASCAT, subdivided for the 12 rainfall classes. As it can be seen, the values of the parameters *Z* and *a* increase with rainfall as it can be expected from equation (2) by considering that SM data are normalized between 0 and 1. The parameter *b* is found to be varying in a lower range (1.4–3.4 if the first class is excluded). The anomalous *a* and *b* values obtained for the first class (0–133 mm) are likely due to the noise of ASCAT SM retrievals, especially in desert areas, which produces erroneous rainfall even where the observations are actually very low. The continuous performance in the calibration period for the estimated rainfall from ASCAT is shown in Figure 2 and compared with those obtained through the TRMM-RT product. We underline that some of the satellite sensors used for developing the GPCP product are also used by the TRMM-RT algorithm. The ASCAT-derived SM product shows median *R* values equal to 0.48 and 0.52 in the comparison against the GPCP and ERAI product, respectively. When using ERAI, the correlation is slightly higher than that obtained with the TRMM-RT product (median *R* = 0.51). In terms of RMSE, the ASCAT-derived product performs better than TRMM-RT for both benchmark data sets while BIAS results are similar.

#### 4.1. Continuous Performance Assessment

After the calibration, all the SM-derived products are evaluated in the validation period from 14 January 2010 to 30 September 2011 against the GPCC rainfall product. We underline here that the selection of a different data set (GPCC rather than GPCP) and an independent time period maximizes the even-handedness of the comparison against the TRMM-RT product. However, it is evident that if the TRMM-RT product were calibrated on these data sets, better performance would be expected.

Figure 3 shows the corresponding global maps of correlation for the three rainfall products derived from the three satellite SM data sets and for the TRMM-RT product. Note that the 5 day accumulated rainfall data are considered in all the analyses made. The selection of the GPCC product (instead of GPCP) is motivated by the need to perform a fair comparison when considering the TRMM-RT product (for the sake of completeness, the comparison against the GPCP and ERAI products is shown in the supporting information, Figures S1 and S2). For each panel, the insets show the correlation values averaged as a function of the vegetation optical depth (VOD), a measure of vegetation density obtained from the AMSR-E sensor [*Owe et al.*, 2008]. Table 2 summarizes the main statistics of the scores used for the evaluation of the rainfall products.



**Figure 2.** Summary results in terms of correlation coefficient (*R*), root-mean-square error, RMSE, and BIAS for the estimated rainfall from ASCAT, and the TRMM-RT product, against GPCP and ERAI products in the calibration period 2007–2009. The results are computed for the  $\pm$ 50° latitude band (TRMM-RT coverage). In the box plots the min/max, 25/75° percentiles and the median values are shown.



**Figure 3.** Correlation, *R*, maps of the 5 day rainfall products derived from ASCAT, AMSR-E, and SMOS soil moisture (through SM2RAIN) and from TRMM-RT against the GPCC product used as benchmark in the validation period 2010–2011. The insets show the histogram of *R* averaged as a function of the Vegetation Optical Depth (VOD). The median *R* values are also shown (the value corresponding to the TRMM-RT data set, i.e., between  $\pm 50^{\circ}$  latitude, are reported in brackets). The *R* values greater than 0.20 (from yellow to red) are significant at the 0.01 confidence level.

Overall, the correlation between the estimated rainfall from SM data and the GPCC data set is remarkable and, on average, comparable with the TRMM-RT product when ASCAT is considered (R = 0.54, see Figure 3), but lower for AMSR-E- and SMOS-derived products. The lower performances of AMSR-E and SMOS are attributed to their lower temporal resolution (mainly due to radio frequency interference, RFI) which significantly affects the rainfall retrieval through SM2RAIN. Moreover, prior to July 2010, SMOS was in the commissioning phase during which lower retrieval performance is expected [*Kerr et al.*, 2012a]. Note that the spatial correlation pattern shown in Figure 3 closely resembles the expected accuracy of the SM data sets [*Dorigo et al.*, 2010; *Leroux et al.*, 2013]. Lower correlation values have to be expected over tropical forests where the signal of all satellite sensors does not penetrate the dense vegetation cover (for SMOS these areas are masked out and no

**Table 2.** Summary of Performance Scores (Median and Standard Deviation in Brackets) Computed for 5 Day Accumulated Rainfall From the Three Soil Moisture-Derived Rainfall Products (ASCAT, AMSR-E, and SMOS) and for TRMM-RT Against the GPCC Product Used as Benchmark in the Validation Period 2010–2011<sup>a</sup>

Rainfall		R			RMSE (mm)		BIAS (—)			
Product	ALL	VOD <0.5	VOD >0.5	ALL	VOD <0.5	VOD >0.5	ALL	VOD <0.5	VOD >0.5	
ASCAT	0.47 (0.26)	0.48 (0.27)	0.42 (0.21)	11.76 (9.28)	10.38 (7.62)	19.62 (9.69)	0.14 (0.38)	0.14 (0.38)	0.18 (0.34)	
AMSR-E	0.26	0.29	0.19	11.88	10.64	22.26	-0.13 (0.43)	-0.15 (0.42)	0.12	
SMOS	0.24 (0.25)	0.28 (0.25)	0.09 (0.19)	11.57 (10.13)	10.80 (18.20)	8.94 (12.02)	-0.13 (0.42)	-0.14 (0.42)	-0.15 (0.44)	
TRMM-RT	0.53 (0.20)	0.53 (0.20)	0.55 (0.17)	16.36 (11.78)	14.47 (10.29)	25.38 (12.43)	0.04 (0.37)	0.06 (0.38)	0.07 (0.30)	

<sup>a</sup>The scores are shown for the whole domain (ALL) and also subdivided as a function of the vegetation optical depth, VOD. For the BIAS, the values outside the range ±1 are masked out. (*R*: correlation coefficient, RMSE: root-mean-square error (mm).)

retrieval is done). A general decrease of the correlation with vegetation density (VOD) is observed for the AMSR-E- and SMOS-derived products, while this is not the case for ASCAT. This finding is in accordance with the expected behavior of passive (AMSR-E and SMOS) and active (ASCAT) microwave retrievals [Dorigo et al., 2010]. Over desert areas, and particularly over the Sahara, the Namibian desert, and the Arabian Peninsula, the ASCAT-derived product provides the lowest performance likely due to volume scattering effects from dry subsurface soil layers [Wagner et al., 2013]. However, we note that in these regions also the gauge-based rainfall products are affected by significant uncertainties. Similar issues are found for regions with very complex topography as, for instance, the Andes and the Himalaya. At northern latitudes, the residual presence of snow, ice, and freezing conditions negatively affects the correlation values for all the products. Excluding these well-known challenging regions, covering less than 18% of the land points, a general very good performance is reached elsewhere, particularly with ASCAT (median R = 0.578), and mainly over the Australian continent, South Africa, Sahel, Brazil, Argentina, Mexico, central United States, Spain, India, and China. In these areas, all the SM-derived products perform well with correlations frequently higher than TRMM-RT (58% of cases for ASCAT). As to the comparison with the GPCP and ERAI products (Figures S1 and S2), results of the SM-derived products are quite similar both in quantitative terms and for the spatial correlation patterns thus corroborating their robustness (see also the supporting information). Moreover, a consistent spatial pattern can be seen comparing the results shown in Figure 3 for AMSR-E with those reported in Crow et al. [2010] (Figure 5) who used a similar, but different, approach for correcting rainfall through SM data.

The accuracy of the rainfall products is also evaluated in terms of absolute amounts. The RMSE and the BIAS are computed by using the GPCC (Figure 4 and Table 2) and the GPCP/ERAI products (Figure S3 in the supporting information) as benchmark. The most interesting result comes out from Figure 4 for the RMSE. The SM-derived rainfall products show very good performance with median RMSE equal to 11.8, 11.8, and 11.4 mm for ASCAT, AMSR-E, and SMOS, respectively, while TRMM-RT performs worst with median RMSE = 16.4 mm. A similar picture is found when the GPCP/ERAI products are adopted (Figure S3). In order to deeply investigate the motivations of these results, the temporal standard deviation of the 5 day rainfall time series of each product is analyzed (see Figure S4 in the supporting information). A highly consistent behavior between the three benchmark rainfall data sets (GPCC, GPCP and ERAI) is obtained with most land areas characterized by values on the order of 5–15 mm. The SM-derived products are found to be less temporally variable particularly for the SMOS and AMSR-E products while ASCAT shows a good agreement till to 80° percentile and above it tends to underestimate. On the other hand, the TRMM-RT product has found to overestimate the temporal variability and this might be an important reason for its highest RMSE values. This result is clarified from the Taylor diagram [Taylor, 2001] shown in Figure S5 (supporting information) where it is evident that the higher RMSE of TRMM-RT product is due to its higher temporal standard deviation while the SMOS and AMSR-E products have lower RMSE due to the tendency of the SM-derived product to overly suppress rainfall temporal variability.

In terms of BIAS, the rainfall products derived from AMSR-E and SMOS show a median value close to zero but higher values (>30%) in absolute terms. The ASCAT-derived product shows a general overestimation but a



**Figure 4.** Frequency histogram and box plots of the (left) root-mean-square error (RMSE) computed on 5 day accumulated rainfall, and (right) BIAS computed on the accumulated rainfall (in %) for the three soil moisture-derived rainfall products (ASCAT, AMSR-E, and SMOS) and for TRMM-RT against the GPCC data set used as benchmark in the validation period 2010–2011. The RMSE (BIAS) median values and the standard deviation (median of absolute values) are shown for each product. The values of the BIAS outside the range ±100% are masked out. In the box plots the min/max, 25/75° percentiles and the median values are shown.

slightly lower median absolute value. The accumulated rainfall for each product in the validation period is also computed and mapped in Figures S6 and S7 of the supporting information. It is obtained that all the SM-derived products are able to reproduce very well the overall rainfall pattern (with some local differences) even though the total amount might be quite different (see, for instance, at the higher latitudes for the AMSR-E- and SMOS-derived products, Figure S7).

#### 4.2. Categorical Performance Assessment

As mentioned above, the rainfall products are also evaluated in terms of categorical performance. Therefore, FAR, POD, and TS are computed for the different SM-derived rainfall products, and for TRMM-RT (see Figure 5 for GPCC and Figure S8 for GPCP and ERAI in the supporting information). In terms of FAR, TRMM-RT outperforms the other three products likely due to the positive noise in the SM time series that are interpreted by SM2RAIN as a low-intensity rainfall event (as already highlighted in *Crow et al.* [2011]). Conversely, in terms of POD, the SM-derived products tend to outperform TRMM-RT. This behavior is attributed to the difficulties of estimating light rainfall from TRMM-RT that appears to be addressed by the proposed approach. By considering the TS, ASCAT outperforms the other products for lower rainfall percentiles (<40 percentiles) while TRMM-RT is slightly better for the higher percentiles thus confirming the tendency of the SM2RAIN algorithm to underestimate the higher rainfall rates [*Brocca et al.*, 2013]. The comparison against the GPCP and ERAI products provides very similar results (Figure S8).



**Figure 5.** Global averages of categorical metrics, computed on 5 day accumulated rainfall, for the three soil moisturederived rainfall products (ASCAT, AMSR-E, and SMOS) and for TRMM-RT against the GPCC data set used as benchmark in the validation period 2010–2011, (left) False Alarm Ratio, (middle) Probability of Detection, and (right) Threat Score for a range of 5 day rainfall accumulation threshold. An event is defined as a 5 day rainfall accumulation that exceeds a given percentile threshold of all 5 day accumulations observed for a given 1° pixel over the whole analyzed period. For each percentile, the mean values  $\pm 1$  standard deviation are shown as solid and dashed lines, respectively.

**AGU** Journal of Geophysical Research: Atmospheres



Figure 6. As in Figure 3 but for the Threat Score (TS) computed considering a rainfall threshold of 1 mm.

Figure 6 shows the global map of TS for the different products against the GPCC product in the validation period. For this figure, a rainfall event is defined as a 5 day rainfall accumulation in excess of 1 mm. The maps clearly highlight the good performance of the SM-derived products to detect rainfall events. Considering the  $\pm$ 50° latitude region, the median TS is found to be equal to 0.808, 0.806, and 0.606 for ASCAT, AMSR-E, and SMOS, respectively. The corresponding value for TRMM-RT is equal to 0.660.

#### 4.3. Products Assessment at Monthly Time Scale

The SM-derived products are also evaluated at the monthly time scale by using the high-quality GPCC Full Data Reanalysis of monthly global land surface precipitation based on the 67,200 stations worldwide [Schneider et al., 2014]. As this product is available until December 2010, the comparison was made for the period 2007–2010 even though this means that for the SMOS-derived product only the year 2010 is

# **AGU** Journal of Geophysical Research: Atmospheres





considered. Figure 7 summarizes the results in terms of box plots for *R*, RMSE, and BIAS. As it can be seen, the previous results are confirmed also by using the high-quality GPCC monthly product. Specifically, for ASCAT the median *R* value is equal to 0.68 and close to those obtained with TRMM-RT (R = 0.70) while the median RMSE = 39.88 mm for ASCAT is significantly lower than those obtained with the other products (RMSE > 45 mm).

In order to visualize the rainfall time series, the monthly rainfall values from January 2007 to December 2009 are reported in Figure 8 for the ASCAT and TRMM-RT products. Specifically, 100 points (out of 13237 land points) evenly distributed on a quasi-global scale are displayed and ordered by latitude and longitude for



**Figure 8.** Monthly rainfall time series obtained from the ASCAT SM-derived rainfall products and from the TRMM-RT against the GPCC data set for 100 points nearly evenly distributed on a quasi-global scale (Lon = -90:150, Lat = -50:50) from January 2007 to December 2009. Note that the *y* axis is truncated at 350 mm for the sake of visualization.



Figure 9. Rainfall time series for 5 day accumulated rainfall obtained from the ASCAT SM-derived products and from the TRMM-RT against the ERAI data set for six points in Brazil, South Africa, Eastern Australia, United States, Spain, and China from 2007 to 2011.

giving a visual interpretation of the analysis. As expected, the ASCAT-derived product slightly underestimates rainfall in the Tropical and Equatorial bands (see middle rows). However, elsewhere a general agreement between the seasonal patterns for each point and the total amount is observed with a general better agreement for the high latitudes (first two rows).

#### 4.4. Time Series for Some Specific Locations

Finally, to visualize the data at 5 day temporal resolution, a time series plot for some specific locations worldwide (in Brazil, South Africa, Eastern Australia, United States, Spain and China) is shown in Figure 9 for the ASCAT and TRMM-RT products against the ERAI data set and by considering the 5 year period 2007–2011. It is evident the reliability of the ASCAT-derived rainfall product in reproducing the seasonal and long-term rainfall pattern with results comparable with the TRMM-RT product.

### 5. Conclusions

The results shown in this study offer compelling evidence that it is possible to estimate rainfall from satellite SM data on a global scale (also for high latitudes). Specifically, by considering the daily GPCC product as benchmark in the validation period 2010–2011 the following conclusions can be drawn:

- 1. the global correlation maps (Figure 3) shows that the SM-derived rainfall products provide good *R* values in the areas where SM retrievals are expected to be accurate. In the latitude band  $\pm$ 50°, the median *R* values are equal to 0.542, 0.281, and 0.313 for the ASCAT, AMSR-E, and SMOS products, respectively. The corresponding *R* value for the TRMM-RT product is equal to 0.534. The regions for which the SMderived products perform very well are Australia, Spain, South and North Africa, India, China, the Eastern part of South America, and the central part of the United States.
- the most interesting results are obtained in terms of RMSE (Figure 4). In fact, the median RMSE values are lower than 12 mm while for the TRMM-RT product a median value equal to 16.4 mm is obtained. In terms of BIAS, the results of the three SM-derived products are quite similar and slightly worse than the TRMM-RT product.
- 3. the categorical scores (Figures 5 and 6), and specifically the TS, show that the SM-derived product performs well in the detection of low to medium rainfall events.
- 4. the analysis at the monthly time scale (Figures 7 and 8) with the higher quality GPCC product confirms the previous results in terms of continuous performance scores.

Overall, these results demonstrate the validity of the proposed SM2RAIN method for improving the rainfall estimation on a global scale with a promising reduction of the RMSEs and a better detection of low to medium rainfall events with respect to the TRMM-RT product. However, we acknowledge that the reduction into the RMSE values of the SM-derived products might be linked to their tendency to overly suppress rainfall accumulation variability. Further improvements can be obtained by merging different satellite SM products, both in space and in time, and by refining the parameterization of the SM2RAIN. Additionally, new satellite missions are foreseen for the remote sensing of SM with expected improved accuracy and spatial-temporal resolution such as the Soil Moisture Active and Passive mission that will be launched in October 2014 [*Entekhabi et al.*, 2010; *Brown et al.*, 2013]. The proposed approach has thus still room for improvement and it will become an essential new tool for global-scale rainfall estimation over land surfaces. Moreover, the same approach can be used to evaluate the reliability of different satellite SM products on a regional and global scale by using rainfall data as benchmark [*Crow et al.*, 2010; *Tuttle and Salvucci*, 2014]. This would allow addressing the issue of the validation of coarse scale SM products due to the lack of dense in situ SM network [*Wagner et al.*, 2014].

#### Acknowledgments

The authors gratefully acknowledge support from the ESA Climate Change Initiative (CCI) (ESA/ESRIN contract 4000104814/11/I-NB, http://www.esasoilmoisture-cci.org/), EUMETSAT through the "Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF)" (http://hsaf.meteoam.it/), and the European Commission through the FP7 Project eartH2Observe. S. Camici, A. Mondini, M. Rossi, and A. Tarpanelli are thanked for their useful comments on an earlier version of this paper. We also thank Wade Crow and two anonymous reviewers for their useful suggestions that significantly improved the readability of the paper. The rainfall data sets obtained from ASCAT, AMSR-E, and SMOS at 1° resolution are available at http:// hydrology.irpi.cnr.it/people/l.brocca or by contacting the first author at luca.brocca@irpi.cnr.it.

#### References

- Brocca, L., F. Melone, T. Moramarco, and W. Wagner (2013), A new method for rainfall estimation through soil moisture observations, Geophys. Res. Lett., 40, 853–858, doi:10.1002/grl.50173.
- Brocca, L., S. Camici, F. Melone, T. Moramarco, J. Martinez-Fernandez, J.-F. Didon-Lescot, and R. Morbidelli (2014), Improving the representation of soil moisture by using a semi-analytical infiltration model, *Hydrol. Processes*, 28(4), 2103–2115.
- Brown, M. E., V. Escobar, S. Moran, D. Entekhabi, P. E. O'Neill, E. G. Njoku, B. Doorn, and J. K. Entin (2013), NASA's Soil Moisture Active Passive (SMAP) mission and opportunities for applications users, *Bull. Am. Meteorol. Soc.*, 94, 1125–1128.
- Chen, F., W. T. Crow, and T. H. Holmes (2012), Improving long-term, retrospective precipitation datasets using satellite-based surface soil moisture retrievals and the soil moisture analysis rainfall tool, J. Appl. Remote Sens., 6(1), 063604.
- Crow, W. T., G. F. Huffman, R. Bindlish, and T. J. Jackson (2009), Improving satellite rainfall accumulation estimates using spaceborne soil moisture retrievals, *J. Hydrometeorol.*, *10*, 199–212.
- Crow, W. T., D. G. Miralles, and M. H. Cosh (2010), A quasi-global evaluation system for satellite-based surface soil moisture retrievals, *IEEE Trans. Geosci. Remote Sens.*, 48(6), 2516–2527.
- Crow, W. T., M. J. van Den Berg, G. F. Huffman, and T. Pellarin (2011), Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART), *Water Resour. Res.*, 47, W08521, doi:10.1029/2011WR010576.
- Dee, D. P., et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. R. Meteorolog. Soc., 137, 553–597.

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

Ebert, E. E., J. Janowiak, and C. Kidd (2007), Comparison of near real time precipitation estimates from satellite observations and numerical models, *Bull. Am. Meteorol. Soc.*, 88, 47–64.

Entekhabi, D., et al. (2010), The Soil Moisture Active and Passive (SMAP) mission, Proc. IEEE, 98(5), 704-716.

- Hong, Y., R. F. Adler, and G. J. Huffman (2007), An experimental global prediction system for rainfall triggered landslides using satellite remote sensing and geospatial datasets, *IEEE Trans. Geosci. Remote Sens.*, 45, 1671–1680.
- Hossain, F., and E. N. Anagnostou (2004), Assessment of current passive-microwave and infrared-based satellite rainfall remote sensing for flood prediction, J. Geophys. Res., 109, D07102, doi:10.1029/2003JD003986.
- Hou, A. Y., R. K. Kakar, S. Neeck, A. A. Azarbarzin, C. D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi (2013), The Global Precipitation Measurement (GPM) mission, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-13-00164.1, in press.
- Huffman, G. J., R. F. Adler, M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind (2001), Global precipitation at one-degree daily resolution from multi-satellite observations, J. Hydrometeorol., 2, 36–50.

Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeorol., 8(1), 38–55.

Keefer, D. K., R. C. Wilson, R. K. Mark, E. E. Brabb, W. M. Brown, S. D. Ellen, E. L. Harp, G. F. Wieczoreck, C. S. Alger, and R. S. Zatkin (1987), Realtime landslide warning during heavy rainfall, *Science*, 238, 921–925.

Kerr, Y. H., et al. (2012a), The SMOS soil moisture retrieval algorithm, IEEE Trans. Geosci. Remote Sens., 50(5), 1384-1403.

Kerr, Y. H., J. Font, M. Martin-Neira, and S. Mecklenburg (2012b), Introduction to the special issue on the ESA's Soil Moisture and Ocean Salinity Mission (SMOS) - Instrument performance and first results, *IEEE Trans. Geosci. Remote Sens.*, 50(5), 1351–1353.

Kidd, C., and V. Levizzani (2011), Status of satellite precipitation retrievals, Hydrol. Earth Syst. Sci., 15, 1109–1116.

Koster, R. D., and the GLACE Team (2004), Regions of strong coupling between soil moisture and precipitation, *Science*, 305, 1138–1140.
Kucera, P. A., E. E. Ebert, F. J. Turk, V. Levizzani, D. Kirschbaum, F. J. Tapiador, A. Loew, and M. Borsche (2013), Precipitation from space:
Advancing earth system science. *Bull. Am. Meteorol. Soc.*, 94, 365–375.

Leroux, D. J., Y. H. Kerr, P. Richaume, and R. Fieuzal (2013), Spatial distribution and possible sources of SMOS errors at the global scale, *Remote Sens. Environ.*, 133, 240–250.

McCabe, M., E. F. Wood, R. Wojcik, M. Pan, J. Sheffield, H. Gao, and H. Su (2008), Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Remote Sens. Environ.*, 112, 430–444.

Owe, M., R. A. M. De Jeu, and T. R. H. Holmes (2008), Multi-sensor historical climatology of satellite-derived global land surface moisture, J. Geophys. Res., 113, F01002, doi:10.1029/2007JF000769.

Pellarin, T., A. Ali, F. Chopin, I. Jobard, and J.-C. Bergs (2008), Using spaceborne surface soil moisture to constrain satellite precipitation estimates over West Africa, *Geophys. Res. Lett.*, 35, L02813, doi:10.1029/2007GL032243.

Pellarin, T., S. Louvet, C. Gruhier, G. Quantin, and C. Legout (2013), A simple and effective method for correcting soil moisture and precipitation estimates using AMSR-E measurements, *Remote Sens. Environ.*, 136, 28–36.

Porporato, A., E. Daly, and I. Rodriguez-Iturbe (2004), Soil water balance and ecosystem response to climate change, Amer. Nat., 164(5), 625–632.

Schamm, K., M. Ziese, A. Becker, P. Finger, A. Meyer-Christoffer, U. Schneider, M. Schröder, and P. Stender (2014), Global gridded precipitation over land: A description of the new GPCC First Guess Daily product, *Earth Syst. Sci. Data*, *6*, 49–60.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2014), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theor. Appl. Climatol.*, 115(1–2), 15–40. Scholthof, K. B. G. (2006). The disease triangle: Pathogens. the environment and society. *Nat. Rev. Microbiol.*, 5(2), 152–156.

Serrat-Capdevila, A., J. B. Valdes, and Z. S. Eugene (2014), Water management applications for satellite precipitation products: Synthesis and recommendations, J. Am. Water Resour. Assoc., 50(2), 509–525.

Tapiador, F. J., et al. (2012), Global precipitation measurement: Methods, datasets and applications, Atmos. Res., 104–105, 70–97.

Taylor, C. M., R. A. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo (2012), Afternoon rain more likely over drier soils, *Nature*, 489(7416), 423–426.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106(D7), 7183–7192, doi:10.1029/2000JD900719.

Tuttle, S. E., and G. D. Salvucci (2014), A new approach for validating satellite estimates of soil moisture using large-scale precipitation: Comparing AMSR-E products, *Remote Sens. Environ.*, 142, 207–222.

Uijlenhoet, R., and A. Berne (2008), Stochastic simulation experiment to assess radar rainfall retrieval uncertainties associated with attenuation and its correction, *Hydrol. Earth Syst. Sci.*, *12*, 587–601.

Wagner, W., G. Lemoine, and H. Rott (1999), A method for estimating soil moisture from ERS scatterometer and soil data, *Remote Sens. Environ.*, 70, 191–207.

Wagner, W., et al. (2013), The ASCAT soil moisture product: Specifications, validation results, and emerging applications, *Meteorol. Z., 22*(1), 5–33.
 Wagner, W., et al. (2014), Clarifications on the "Comparison between SMOS, VUA, ASCAT, and ECMWF soil moisture products over four watersheds in U.S.", *IEEE Trans. Geosci. Remote Sens., 52*(3), 1901–1906.

Wake, B. (2013), Flooding costs, Nat. Clim. Chanae, 3, 778.

Wigneron, J. P., et al. (2007), L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields, *Remote Sens. Environ.*, 107(4), 639–655.