



## SM2RAIN-CCI: a new global long-term rainfall data set derived from ESA CCI soil moisture

Luca Ciabatta<sup>1,4</sup>, Christian Massari<sup>1</sup>, Luca Brocca<sup>1</sup>, Alexander Gruber<sup>2,5</sup>, Christoph Reimer<sup>3</sup>, Sebastian Hahn<sup>2</sup>, Christoph Paulik<sup>2,6</sup>, Wouter Dorigo<sup>2</sup>, Richard Kidd<sup>3</sup>, and Wolfgang Wagner<sup>2</sup>

<sup>1</sup>Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy

<sup>2</sup>Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria

<sup>3</sup>Earth Observation Data Centre, Vienna, Austria

<sup>4</sup>Department of Civil and Environmental Engineering, University of Perugia, Perugia, Italy

<sup>5</sup>Department of Earth and Environmental Sciences, KU Leuven, Heverlee, 3001, Belgium

<sup>6</sup>Vandersat B.V., Haarlem, the Netherlands

**Correspondence:** Luca Ciabatta (luca.ciabatta@irpi.cnr.it)

Received: 3 August 2017 – Discussion started: 21 September 2017

Revised: 14 December 2017 – Accepted: 23 December 2017 – Published: 8 February 2018

**Abstract.** Accurate and long-term rainfall estimates are the main inputs for several applications, from crop modeling to climate analysis. In this study, we present a new rainfall data set (SM2RAIN-CCI) obtained from the inversion of the satellite soil moisture (SM) observations derived from the ESA Climate Change Initiative (CCI) via SM2RAIN (Brocca et al., 2014). Daily rainfall estimates are generated for an 18-year long period (1998–2015), with a spatial sampling of  $0.25^\circ$  on a global scale, and are based on the integration of the ACTIVE and the PASSIVE ESA CCI SM data sets.

The quality of the SM2RAIN-CCI rainfall data set is evaluated by comparing it with two state-of-the-art rainfall satellite products, i.e. the Tropical Measurement Mission Multi-satellite Precipitation Analysis 3B42 real-time product (TMPA 3B42RT) and the Climate Prediction Center Morphing Technique (CMORPH), and one modeled data set (ERA-Interim). A quality check is carried out on a global scale at  $1^\circ$  of spatial sampling and 5 days of temporal sampling by comparing these products with the gauge-based Global Precipitation Climatology Centre Full Data Daily (GPCC-FDD) product. SM2RAIN-CCI shows relatively good results in terms of correlation coefficient (median value  $> 0.56$ ), root mean square difference (RMSD, median value  $< 10.34$  mm over 5 days) and bias (median value  $< -14.44\%$ ) during the evaluation period. The validation has been carried out at original resolution ( $0.25^\circ$ ) over Europe, Australia and five other areas worldwide to test the capabilities of the data set to correctly identify rainfall events under different climate and precipitation regimes.

The SM2RAIN-CCI rainfall data set is freely available at <https://doi.org/10.5281/zenodo.846259>.

### 1 Introduction

Accurate estimation of rainfall is of paramount importance for many applications, e.g. natural hazards risk assessment and mitigation, famine and disease monitoring, water resources management, weather forecasting, and climate modeling (Dinku et al., 2007).

Ground stations provide accurate local estimates of rainfall (Villarini et al., 2008) and are considered the most accurate source of rainfall data for modeling and process moni-

toring. However, two main issues limit their usefulness. First, they are characterized by a nonhomogenous coverage (Kidd et al., 2016) throughout the globe and, second, they are only representative of a limited area around the gauge. These limitations impact the use of rain gauge data mainly over large and remote areas. Another source of rainfall information are ground meteorological radar instruments, which are able to provide measurements that are more representative of the actual rainfall spatial variability. However, ground meteorolog-

ical radar is also affected by issues that reduce the quality of the rainfall estimates such as beam blockage and frozen hydrometeors. In addition, ground-based observations are subjected to high costs of maintenance related to setting up, calibration and fixing of rain gauges and radars. These issues can limit the use of ground rainfall estimates, especially in developing countries.

Satellite rainfall estimates can offer a valuable alternative to ground-based observations and today provide measurements at an increased spatial and temporal resolution. For example, the recent NASA–JAXA joint Global Precipitation Measurement (GPM, Hou et al., 2014) mission delivers rainfall products in near-real time with a spatial sampling of  $0.1^\circ$  every 30 min, by using a constellation of satellite sensors. A large number of satellite rainfall products have been developed in the past, e.g. the near-real-time Tropical Rainfall Measurement Mission Multi-satellite Precipitation Analysis (TMPA 3B42RT, Huffman et al., 2007); the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN, Hsu et al., 1997); the Climate Prediction Center MORPHing technique, (CMORPH, Joyce et al., 2004); and the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS, Funk et al., 2015) products. These products are being used worldwide for several applications, such as drought and famine monitoring, weather forecasts and natural hazard risk mitigation. When providing a sufficiently long observation period, they are also used for climatological applications like the PERSIANN-Climate Data Record (Ashouri et al., 2015), which have provided continuous rainfall estimates since 1983. Despite the relative advantages of having an estimate of rainfall in every place on the Earth, satellite rainfall estimates, like ground observations, are not free of errors. In fact, the instantaneous satellite-based retrievals of precipitation, which is a process subject to high spatial and temporal variability, makes the reconstruction of the accumulated rainfall on longer temporal scales (e.g., daily accumulated rainfall) challenging (Trenberth and Asram, 2014). Another issue is related to the estimation of light rainfall, especially over land, which is impacted by the land surface emissivity (Kucera et al., 2013). These aspects negatively affect the rainfall estimates at the measurement area limiting their use, especially for operational purposes, like natural hazards assessment. The use of a constellation of satellites, as adopted in the GPM mission, is able to mitigate the issue of the accumulated rainfall estimation through more frequent satellite overpasses during a day, thus reducing the errors associated with the retrievals (Panegrossi et al., 2016).

A way to improve the quality of satellite rainfall estimates has been explored recently by means of different approaches and relies on the use of satellite surface soil moisture (SSM) data (Crow et al., 2009, 2011; Pellarin et al., 2013; Brocca et al., 2013, 2014; Wanders et al., 2015; Zhan et al., 2015). These approaches exploit the strong relationship between SSM and rainfall to correct and/or estimate rainfall by us-

ing satellite surface SM (soil moisture) data. Among these methods, SM2RAIN (Brocca et al., 2013) is the only technique that directly provides rainfall estimates from SSM observations while the others are correction-based techniques. SM2RAIN has been used to estimate precipitation from various single-sensor SSM products, e.g. from ASCAT and SMOS (Brocca et al., 2013, 2014, 2016; Ciabatta et al., 2015, 2017; Koster et al., 2016; Massari et al., 2017).

With the aim of providing valuable tools for climate change monitoring, the European Space Agency (ESA) has established the so-called Climate Change Initiative (CCI) project. The objective is to exploit Earth observation data sets for providing useful information to policy makers about several essential climate variables (ECVs). Within the CCI programme, three long-term SM products ( $> 37$  years) have been developed by merging SSM retrievals from both active and passive microwave instruments carried by various satellite platforms (Liu et al., 2011, 2012; Wagner et al., 2012). More specifically, the CCI SM project provides three different products, namely active (obtained by merging radar-estimated SM), passive (obtained by merging radiometer-estimated SM) and combined (obtained by merging the active and passive data sets). The availability of these SM data records opens up new opportunities for creating independent long-term rainfall data sets based on SM2RAIN.

The objective of this study is to present and evaluate a quasi-global long-term SM2RAIN-CCI rainfall data set obtained from the inversion of the ESA CCI SM via the SM2RAIN algorithm (Brocca et al., 2014). The SM2RAIN-CCI rainfall data set is compared against several precipitation products, e.g. TMPA 3B42RT, CMORPH, ERA-Interim, the Global Precipitation Climatology Centre Full Data Daily (GPCC-FDD) product (Schamm et al., 2015) and the recently developed Multi-Source Weighted-Ensemble Precipitation (MSWEP, Beck et al., 2017). The analysis is performed on a global scale at  $1^\circ$  spatial sampling, during the period 1998–2015. In addition, a regional-scale analysis at  $0.25^\circ$  spacing is performed by comparing SM2RAIN-CCI estimates against high-quality ground-based observations over Europe, India and Australia.

## 2 Data and methods

### 2.1 Data set generation

#### 2.1.1 State-of-the-art rainfall data sets

In this study, five state-of-the-art rainfall products including models and satellite-based and ground-based observations are intercompared with the new SM2RAIN-CCI data set. In particular, the following products are considered as benchmarks:

1. GPCC-FDD, available at  $1^\circ$  spatial sampling during the period 1988–2013 (ground based data set) at daily tem-

poral resolution, used for calibrating SM2RAIN and for quality check;

2. MSWEP v1.2, available from 1 January 1979 to 3 December 2015 at  $0.25^\circ$  spatial sampling on a 3-hourly basis (combination of models, ground measurements and satellite observations), used as an independent benchmark for the yearly global analysis;
3. E-OBS rainfall data set (Haylock et al., 2008), available from 1950 at  $0.25^\circ$  spatial resolution on a daily basis, used for validation over Europe;
4. rainfall data provided by the India Meteorological Department (IMD), available on a daily basis during the period 2002–2015, used for validation over India;
5. the Australian Water Availability Project (AWAP) observed rainfall data set during the period 2010–2013, used for validation over Australia.

Three rainfall data sets are additionally used for cross-comparison with the SM2RAIN-CCI product:

1. TMPA 3B42RT (hereinafter referred to as TRMMRT), available from 1 March 2000 to present at  $0.25^\circ$  spatial resolution for the  $\pm 50^\circ$  latitude band every 3 h (only satellite);
2. CMORPH raw data (hereinafter referred to as CMORPH), available from 1 January 2000 to present at  $0.25^\circ$  spatial resolution for the  $\pm 60^\circ$  latitude band every 3 h (only satellite);
3. ERA-Interim reanalysis product, available from 1 January 1978 to present at  $0.77^\circ$  spatial sampling on a daily basis (Dee et al., 2011) (reanalysis).

Due to the different spatial sampling and coverage (both in space and in time), the assessment is carried out during the period 1998–2013 for the  $\pm 50^\circ$  latitude band (due to data availability TRMMRT and CMORPH are considered starting from 2000).

The GPCC-FDD data set is a gauge-based product. The number of stations used in the data set varies throughout the years. In total, data from more than 60 000 stations are used. GPCC-FDD is provided on a global scale over a grid with  $1^\circ$  spatial sampling and on a daily basis. The product is available for the period 1988–2013. Here, GPCC-FDD is used for SM2RAIN calibration and quality check because it is completely independent from any satellite data and it does not contain any missing values (Herold et al., 2017). For further details, the reader is referred to Schamm et al. (2015).

MSWEP is a recently developed precipitation data set that combines precipitation information from several sources, including GPCC-FDD, TRMMRT, CMORPH and ERA-Interim. The estimates obtained through satellite sensors, global models and in situ stations are merged by the use

of integration weights. The product is available from 1979 to 2015 with a spatial sampling of  $0.25^\circ$ . More information about MSWEP can be found in Beck et al. (2017).

TRMMRT provides rainfall estimates by taking advantage of multiple satellite sensors, i.e., the TRMM Microwave Imager (TMI), the Special Sensor Microwave Imager (SSM/I), the Advanced Microwave Scanning Radiometer – Earth Observing System (AMSR-E) and the Advanced Microwave Sounding Unit B (AMSU-B). The microwave estimates are blended with infrared (IR) observations derived from sensors on board of Geostationary Earth Orbit (GEO) platforms to obtain rainfall estimates at higher temporal and spatial resolution. The product is provided for the  $\pm 50^\circ$  latitude band over a grid with a  $0.25^\circ$  spacing every 3 h. Daily accumulated rainfall is computed by summing up all rainfall estimates within 1 day. In this study the TMPA-3B42RT version 7 is used. For more details about the TRMMRT product, the reader is referred to Huffman et al. (2007).

CMORPH uses precipitation estimates derived from the same microwave sensors used for TRMMRT generation and uses GEO-IR data to propagate the microwave estimates at the times between two successive microwave satellite overpasses. The product is considered at daily temporal resolution over the  $0.25^\circ$  sampling TRMMRT grid for the  $\pm 60^\circ$  latitude band. In this study, CMORPH version 1.0 raw data are considered. The reader is referred to Joyce et al. (2004) for more details about CMORPH.

ERA-Interim is a reanalysis product provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is based on a global atmospheric model in which different types of observations are routinely assimilated. The product is available from 1979 with a spatial resolution of about  $0.77^\circ$ . The data used have been downloaded from the ECMWF API (<http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/>) and resampled over the  $0.25^\circ$  CCI grid. For further details about ERA-Interim, the reader is referred to Dee et al. (2011).

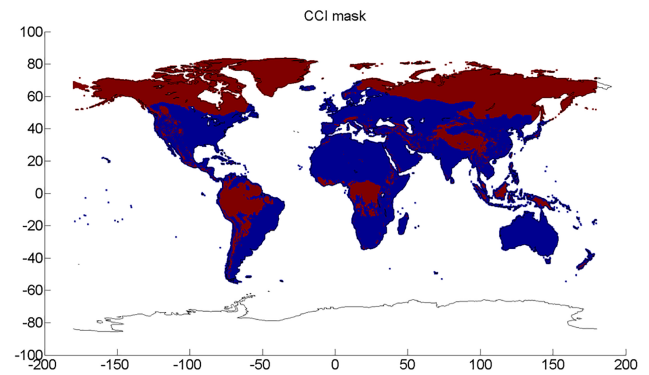
### 2.1.2 ESA CCI soil moisture

The ESA CCI (<http://www.esa-soilmoisture-cci.org/>) provides long-term SM data sets for the period 1978–2015 (Liu et al., 2011; Dorigo et al., 2015, 2017). The products are provided on a global scale with a spatial sampling of  $0.25^\circ$  with daily temporal sampling in three different configurations. The passive microwave product (hereinafter referred to as “PASSIVE”) is provided for the period 1978–2015 and it is generated by merging SM products derived from the Scanning Multichannel Microwave Radiometer (SMMR, operating at 6.6 and 10.7 GHz, Owe et al., 2001), the SSM/I (operating at 19.35 GHz, Owe et al., 2008), the TMI (operating at 10.65 GHz and above, Gao et al., 2006), the AMSR-E (operating at 6.9 and 10.65 GHz, Owe et al., 2008) and its successor AMSR2 (operating at 6.93, 7.3 and 10.65 GHz), WindSat (operating between 6.8 and 37 GHz, Li et al., 2010

and Parinussa et al., 2012), and the ESA Soil Moisture and Ocean Salinity mission (SMOS, Kerr et al., 2012). Although the PASSIVE data set is obtained by considering some of the sensors used for creating the TMPA products, this will not impact the comparison between TRMMRT and SM2RAIN-CCI as different microwave frequencies are taken into account for rainfall estimation. The active data set (hereinafter referred to as “ACTIVE”) is provided for the period 1991–2015 and it is generated by merging active microwave satellite retrievals from the European Remote Sensing satellites (ERS-AMI, operating at 5.3 GHz) and from the Advanced Scatterometer (ASCAT, operating at 5.255 GHz, Wagner et al., 2013) onboard the Metop-A and -B satellites. The third data set (hereinafter referred to as “COMBINED”) is obtained by merging the ACTIVE and PASSIVE products. The merging of the individual data sets is performed by means of a weighted averaging which is parameterized using a triple collocation (TC, Stoffelen, 1998) approach (Gruber et al., 2017). In this study, we consider the ESA CCI SM product at version v03.2. For further details regarding the ESA CCI SM product development, sensor availability and performances the reader is referred to Liu et al. (2011, 2012), Dorigo et al. (2015, 2017) and Wagner et al. (2012).

## 2.2 ESA CCI soil moisture pre-processing

Before applying the SM2RAIN algorithm the following pre-processing steps are applied to the ESA CCI SM data sets. A static mask (Fig. 1) is used to mask out periods with high frozen soil and snow probability, rainforest areas, and areas with high topographic complexity. The latter two are provided within the ESA CCI SM data portal. Notice that deserts are particularly challenging for SM retrieval from active instruments. Therefore, we use the passive data set only in such areas (see Sect. 2.3), which typically provides more reliable retrievals over desert areas (Dorigo et al., 2010). Moreover, a dynamic mask is applied to SSM data on a daily basis in order to remove observations characterized by issues in the retrieval (frozen soil, dense vegetation). This mask is provided alongside each of the ESA CCI SM products. After the application of the dynamic mask, many temporal gaps are found within the SM time series. In order to reduce the data gaps, the time series are interpolated to 00:00 UTC on a daily basis. A maximum data gap of 3 days is considered for the linear interpolation. Data gaps larger than 3 days are left empty, i.e., no rainfall estimation is carried out within these intervals. Prior to 1998, the SM data sets are characterized by a low temporal coverage and a reduced data quality (Dorigo et al., 2015). Thus, the SM2RAIN-CCI product is generated only for the period 1998–2015. The original ACTIVE and PASSIVE CCI SM data sets have been read and pre-processed by using routines developed in Python<sup>®</sup> language by the TUWIEN Remote Sensing Research Group (Ciabatta et al., 2016). After the preprocessing steps, the ESA CCI SM data are ready to be used as input in SM2RAIN.



**Figure 1.** Data mask used for remove areas (red areas) characterized by issues in the soil moisture retrieval.

## 2.3 SM2RAIN algorithm and SM2RAIN-CCI rainfall product generation

The SM2RAIN algorithm (Brocca et al., 2013, 2014) allows rainfall estimates to be derived from SM observations. It is based on the inversion on the following soil water balance equation:

$$p(t) = Z^* ds(t)/dt + a \cdot s(t)^b, \quad (1)$$

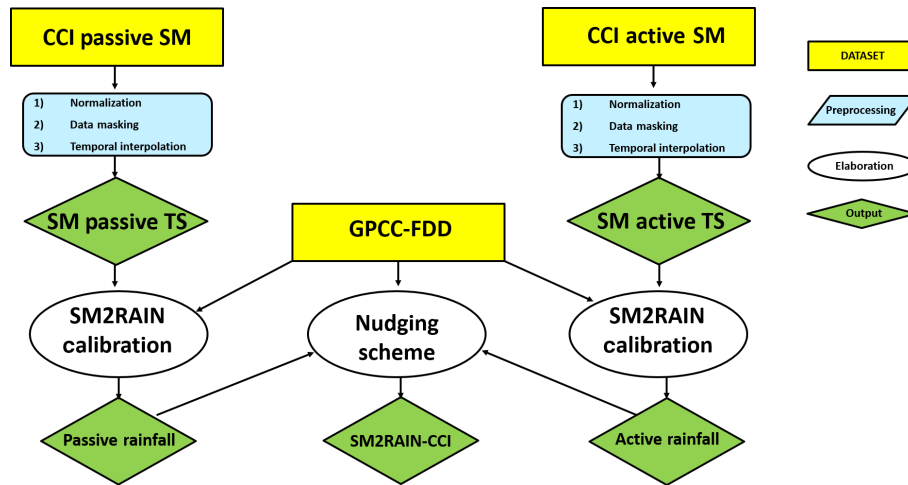
where  $p(t)$  is the estimated rainfall,  $Z^*$  is the soil water capacity (soil depth times soil porosity),  $s(t)$  is the relative soil saturation,  $t$  is the time, and  $a$  and  $b$  are two parameters describing the nonlinearity between soil saturation and drainage.  $Z^*$ ,  $a$  and  $b$  are estimated through calibration. The algorithm is based on the assumption that, during rainfall, evapotranspiration is negligible and surface runoff occurs only when the soil is fully saturated (Brocca et al., 2015). SM2RAIN has also the main limitation of not being able to estimate rainfall if the soil is close to saturation, since no SM variations can be observed after rainfall events in such conditions.

The algorithm has proved to accurately estimate rainfall both on a regional (Abera et al., 2016; Brocca et al., 2013, 2015, 2016; Ciabatta et al., 2015, 2017) and on a global scale (Brocca et al., 2014; Koster et al., 2016). For further details about the SM2RAIN formulation, the reader is referred to Brocca et al. (2013, 2014).

The SM2RAIN parameters are obtained by minimizing the root mean square difference (RMSD) between the 5-day estimated rainfall and the GPCC-FDD data during three calibration periods 1998–2001, 2002–2006 and 2007–2013 on a pixel-by-pixel basis. We considered 5 days of accumulation to reduce the amount of data and speed up the calibration step. The use of different calibration periods relies on the different data and sensors that we used for building the ACTIVE and PASSIVE SSM data sets (Table 1, see also Dorigo et al., 2012). The calibration is performed on a pixel-by-pixel basis separately for ACTIVE and PASSIVE. SM2RAIN was

**Table 1.** Available sensors and temporal intervals considered for the SM2RAIN algorithm application.

Sensor (active/passive)	Temporal interval
AMI-WS/SSMI & TMI	1 January 1998 to 30 June 2002
AMI-WS/AMSR-E	1 July 2002 to 31 December 2006
ASCAT-A & ASCAT B/AMSR-E & Windsat & SMOS & AMSR2	1 January 2007 to 31 December 2013
AMI-WS/SSMI & TMI & AMSR-E	1 January 1992 to 31 December 2006
ASCAT-A & ASCAT-B/AMSR-E & Windsat & SMOS & AMSR2	1 January 2007 to 31 December 2013

**Figure 2.** Analysis framework.

also applied to the COMBINED SSM data set, but we observed a slight reduction of performance with respect to the individual ACTIVE and PASSIVE products (see Table 2), and hence the COMBINED SSM data set is not considered here. The deterioration is due to the different merging techniques; while the ACTIVE and PASSIVE products are merged in one product at the rainfall level by maximizing the Pearson's correlation against the benchmark, the CCI COMBINED SSM is created by adopting a triple-collocation analysis to the original SM time series. In order to match the different spatial resolutions of the considered data sets, GPCCC-FDD was regridded to the  $0.25^\circ$  CCI grid by using the *grid-data* function implemented in MATLAB<sup>®</sup> R2012a, through linear interpolation. After the application of SM2RAIN to the ACTIVE and PASSIVE SM data sets, the two obtained rainfall products are integrated through the following:

$$P_{\text{SM2RAIN-CCI}} = k P_{\text{ACT}} + (1 - k) P_{\text{PAS}}, \quad (2)$$

where  $P_{\text{ACT}}$  and  $P_{\text{PAS}}$  are the two rainfall data sets obtained through the application of SM2RAIN to the ACTIVE and the PASSIVE SM data sets, respectively, and  $P_{\text{SM2RAIN-CCI}}$  is the final SM2RAIN-CCI rainfall data set. The integration weights ( $k$ ) are estimated through the following (Kim et al., 2015):

$$k = \frac{\rho_{AB} - \rho_{AP} \cdot \rho_{PB}}{\rho_{PB} - \rho_{AP} \cdot \rho_{AB} + \rho_{AB} - \rho_{AP} \cdot \rho_{PB}}, \quad (3)$$

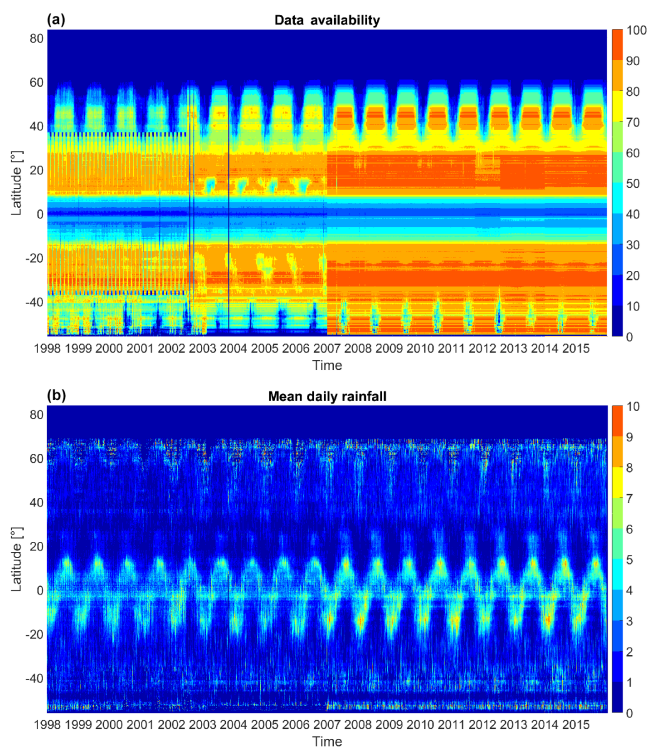
where  $\rho$  is the Pearson correlation coefficient between two data sets with the subscript A, P and B denoting the ACTIVE, the PASSIVE and the benchmark (GPCCC-FDD in this case) rainfall estimates, respectively. When one of the two data sets ( $P_{\text{ACT}}$  or  $P_{\text{PAS}}$ ) is not available at a certain location (e.g., due to unfavorable retrieval conditions), then only the available one is used for the generation of the combined rainfall product. The workflow is depicted in Fig. 2. The data are available in netCDF format via the CCI SM FTP server. The rainfall data are provided in millimeters per day, over land at  $0.25^\circ$  of sampling. The SM2RAIN-CCI data set temporal coverage will be extended when new ESA CCI SM updates will be released.

### 3 SM2RAIN-CCI performance

The SM2RAIN-CCI rainfall data set is available from 1 January 1998 to 31 December 2015 with daily temporal resolution. The data are provided over a  $0.25^\circ$  grid on a global scale, given the native spatial resolution of SM observation of 25 and 50 km. The spatiotemporal coverage is reported in Fig. 3. As can be seen, there is an increase of available data after 2002 and 2007, corresponding to the start of the AMSR-E and ASCAT operations, respectively. Before 2002, the ESA CCI SM products are characterized by a small amount of data, due to longer revisit times of the used satel-

**Table 2.** Statistical summary of correlation coefficient ( $R$ ) and root mean square difference (RMSD) for ACTIVE, PASSIVE and COMBINED rainfall data sets against GPCC-FDD during the three calibration periods 1998–2001, 2002–2006 and 2007–2013. For each score, the median, mean, minimum, maximum and standard deviation values are reported.

Data set	Period	$R$ (–)					RMSD (mm)				
		Median	Mean	Min.	Max.	SD	Median	Mean	Min	Max	SD
Active	1998–2001	0.56	0.50	–1.00	1.00	0.26	10.15	12.56	0.76	157.84	11.11
	2002–2006	0.56	0.50	–1.00	1.00	0.29	9.33	12.14	0.13	151.48	11.08
	2007–2013	0.63	0.57	–0.23	0.90	0.22	9.34	10.95	0.44	89.15	8.31
Passive	1998–2001	0.47	0.44	–1.00	1.00	0.23	10.40	12.94	0.13	126.65	10.90
	2002–2006	0.55	0.50	–1.00	1.00	0.22	10.12	12.59	0.44	122.01	10.56
	2007–2013	0.55	0.51	–0.44	1.00	0.21	10.26	12.58	0.34	117.55	10.61
Combined	1998–2001	0.52	0.49	–1	1	0.20	10.47	12.73	0.47	115.20	10.53
	2002–2006	0.57	0.53	–1	1	0.20	9.86	12.21	0.28	130.83	10.17
	2007–2013	0.63	0.60	–0.26	0.88	0.17	9.48	11.15	0.41	102.03	8.79



**Figure 3.** Hovmöller plot showing the spatial–temporal data availability, in percentage of the total annual available data (upper panel) and the mean daily rainfall (in millimeters per day, lower panel) of the SM2RAIN-CCI rainfall data set for different latitude bands.

lites. Before that date, the rainfall estimates obtained through SM2RAIN should be used with caution because of the likelihood of missing precipitation events. The lack of data over tropical areas and at high latitudes is due to the application of the mask described above. Figure 3 also shows the mean daily rainfall for the SM2RAIN-CCI data set during the analysis period. As can be seen, an increase in the daily values

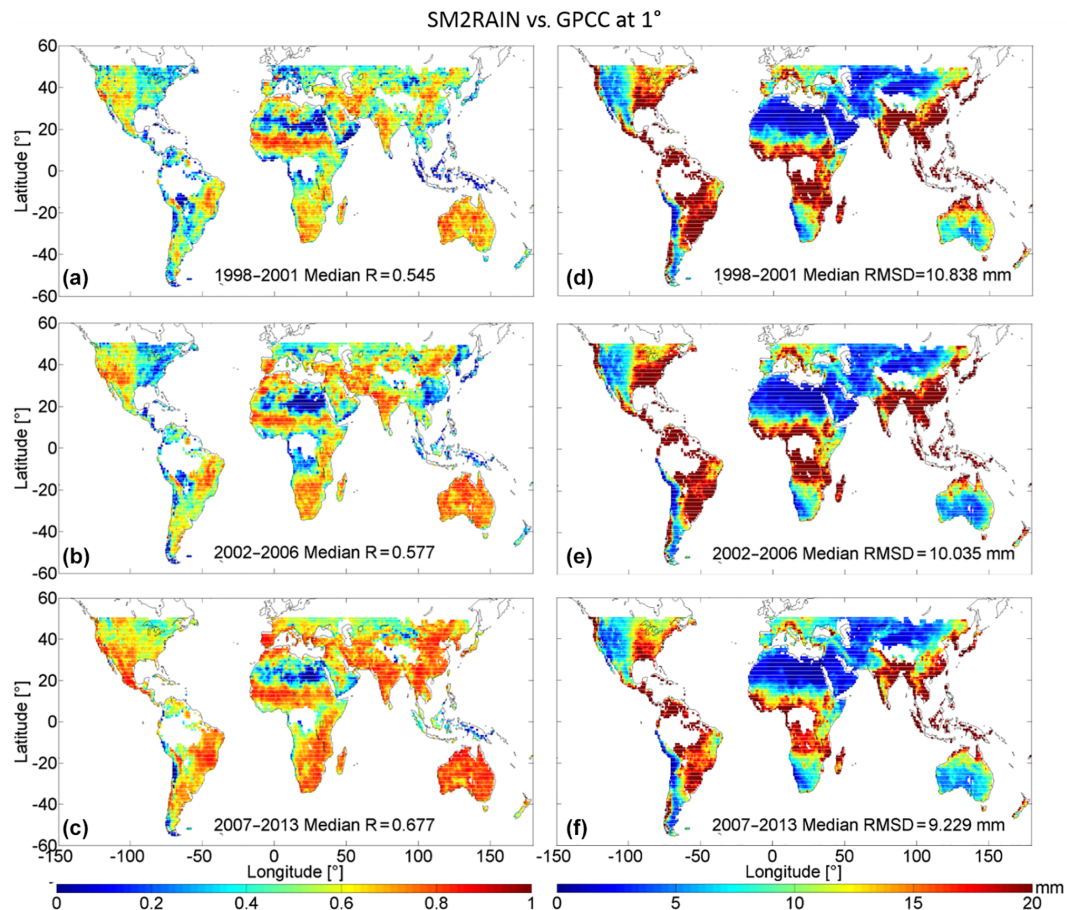
can be observed after 2007, especially over the tropical areas, where the seasonality is well reproduced, due to the higher number of satellite overpasses.

When compared to the GPCC-FDD rainfall data set, SM2RAIN-CCI shows relatively good performance for 5-day rainfall accumulation in terms of both correlation and RMSD, as drawn in Fig. 4 for the  $\pm 50^\circ$  latitude bands during the three calibration periods at  $1^\circ$  of spatial resolution, in order to check the impact of the different spatial resolution considered for the benchmark. The scores are summarized in Table 3. SM2RAIN-CCI rainfall shows relatively good agreement with GPCC-FDD, especially over Africa, Australia, India and South America in terms of correlation ( $R$ ). The RMSD pattern is related to the rainfall regimes. The highest values are located in those regions characterized by high total annual precipitation, e.g. tropical areas. The comparison also provides better performance for the 2007–2013 period than for the 1998–2001 and 2002–2006 periods due to the better temporal coverage of the ESA CCI SM products and their improved accuracy (Dorigo et al., 2015). As can be seen in Fig. 4, the median  $R$  (RMSD) obtained for the 1998–2001 calibration period is 0.54 (10.94 mm), while for the 2007–2013 period, a median value of 0.65 (9.6 mm) is obtained. Indeed, due to the nature of the SM2RAIN algorithm, more frequent satellite overpasses are expected to provide more reliable rainfall estimates. SM2RAIN-CCI shows a lower performance over the Sahara desert and at high latitudes, due to lower SM data quality over these regions. The performance of the parent products ACTIVE and PASSIVE are reported in Table 3. Figure 4 also displays the lower performances obtained for the eastern US. A similar performance pattern was also found by Massari et al. (2017), who calculated global correlation of different rainfall data sets by applying the extended TC (McColl et al., 2014) analysis.

A cross-comparison of SM2RAIN-CCI with GPCC, TRMM, CMORPH, ERA-Interim and MSWEP is reported

**Table 3.** Statistical summary of correlation coefficient ( $R$ ) and root mean square difference (RMSD) for SM2RAIN-CCI against GPCP-FDD during the three calibration periods 1998–2001, 2002–2006 and 2007–2013. For each score, the median, mean, minimum, maximum and standard deviation values are reported.

Period	$R$ (–)					RMSD (mm)				
	Median	Mean	Min.	Max.	SD	Median	Mean	Min.	Max.	SD
1998–2001	0.53	0.50	–1.00	1.00	0.20	10.29	12.71	0.54	125.22	10.65
2002–2006	0.56	0.52	–1.00	0.90	0.21	9.96	12.49	0.45	123.34	10.52
2007–2013	0.65	0.62	–0.15	0.91	0.17	9.16	10.81	0.47	90.45	8.50

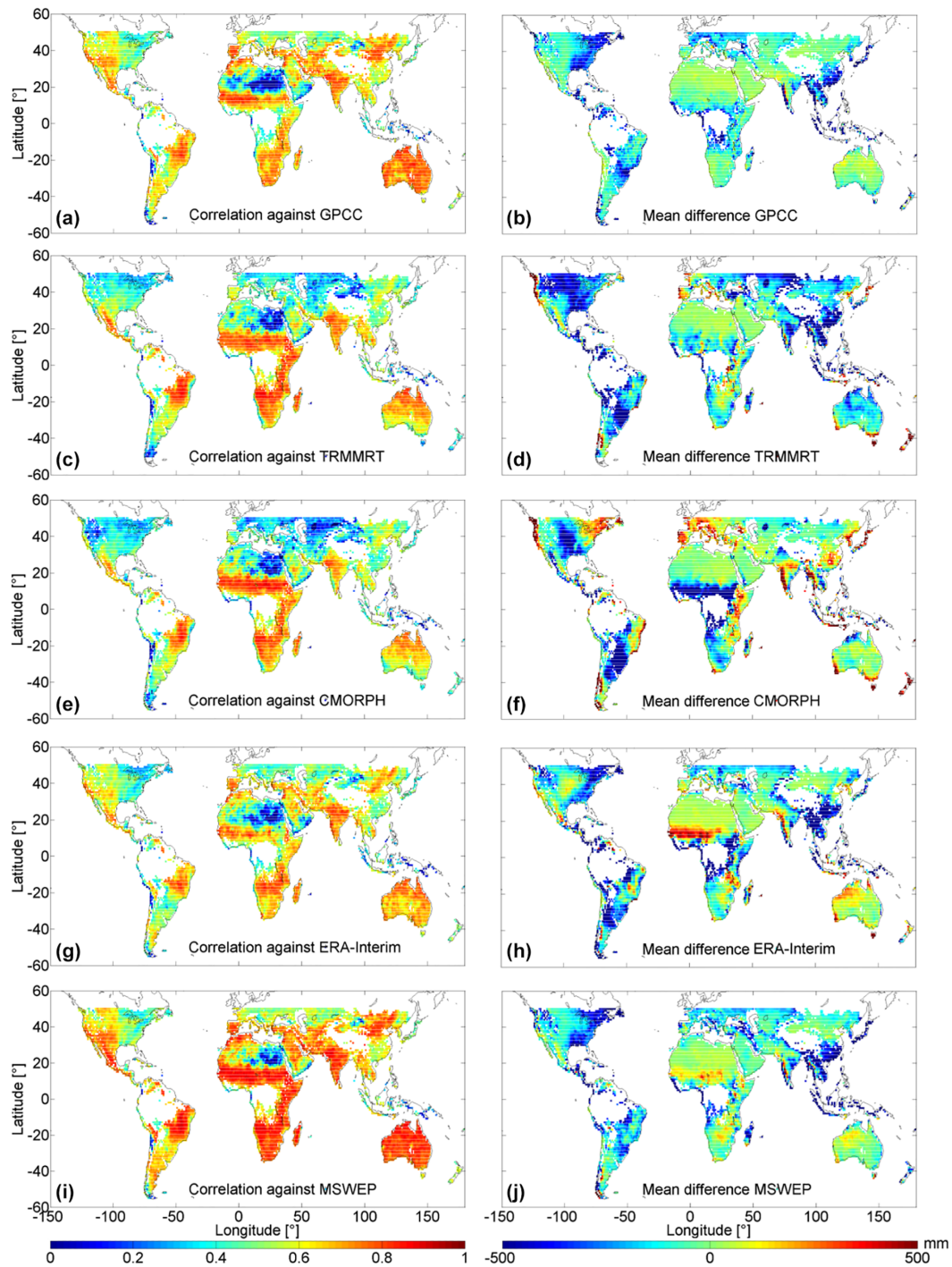


**Figure 4.** Global Pearson correlation (left column) and root mean square difference (right column) maps obtained between GPCP-FDD and the SM2RAIN-CCI rainfall data set for 5-day accumulated rainfall during the periods 1998–2001 (a, d), 2002–2006 (b, e) and 2007–2013 (c, f).

in Fig. 5. The figure displays the  $1^\circ \times 1^\circ$  ( $\pm 50^\circ$ ) correlation maps of 5 days of accumulated rainfall (left column) and the differences in the mean annual rainfall (right column) between SM2RAIN-CCI and the other rainfall data sets. The difference in the mean annual rainfall are calculated by subtracting the mean annual rainfall of each data set to the one provided by SM2RAIN-CCI. The analysis shows that SM2RAIN-CCI rainfall estimates are in good agreement with the state-of-the-art data sets in terms of both  $R$  and mean annual rainfall. Nonnegligible differences can be ob-

served over the Sahara desert, eastern US, South America, the tropical areas and over Europe, where SM2RAIN-CCI provides a smaller amount of rainfall than the other rainfall data sets. However, very good performance can be observed over Africa, Brazil, western US, India and Australia, in terms of both  $R$  and mean annual rainfall. A detailed summary of statistical scores is reported in Table 4.

Seven macroregions worldwide have been selected to check the capability of the SM2RAIN-CCI in estimating rainfall under different climatic conditions. Therefore, mean



**Figure 5.** Correlation maps for 5 days of accumulated rainfall (left column) and differences in mean annual rainfall (right column) obtained by comparing (from top to bottom) SM2RAIN-CCI and GPCC (**a, b**), SM2RAIN-CCI and TRMMRT (**c, d**), SM2RAIN-CCI and CMORPH (**e, f**), SM2RAIN-CCI and ERA-Interim (**g, h**), and SM2RAIN-CCI and MSWEP (**i, j**) at  $1^\circ$  spatial resolution during the period 1 March 2000–31 December 2013.

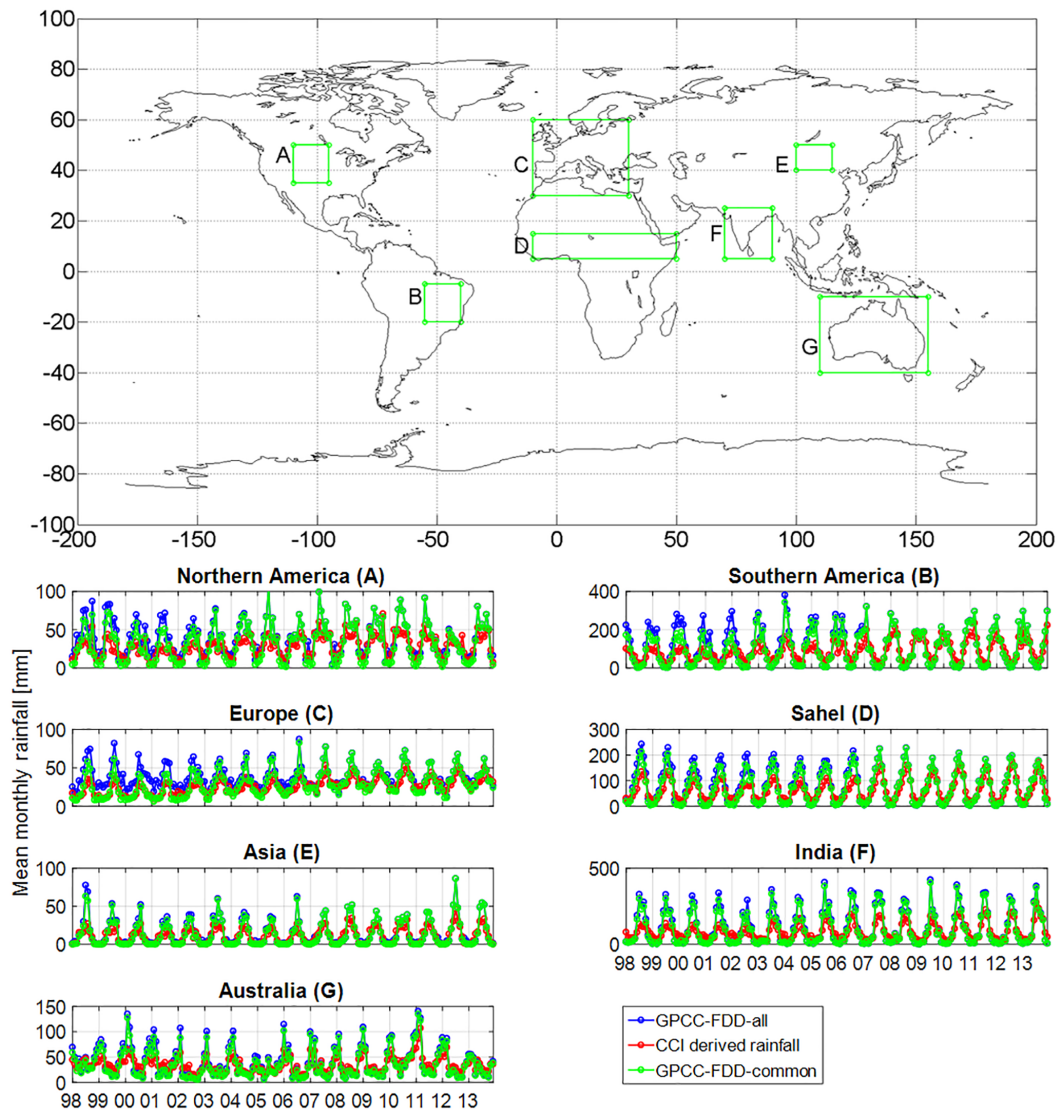
monthly rainfall (MMR) was computed from GPCC-FDD and SM2RAIN-CCI during the period 1998–2013 within these regions, illustrated as green boxes in Fig. 6. From Fig. 6, one can see that the temporal rainfall patterns agree

well in all considered macroregions. SM2RAIN-CCI provides a general underestimation before 2007, due to the increased number of data gaps. Indeed, if the GPCC-FDD MMR is estimated only when SM observations are available

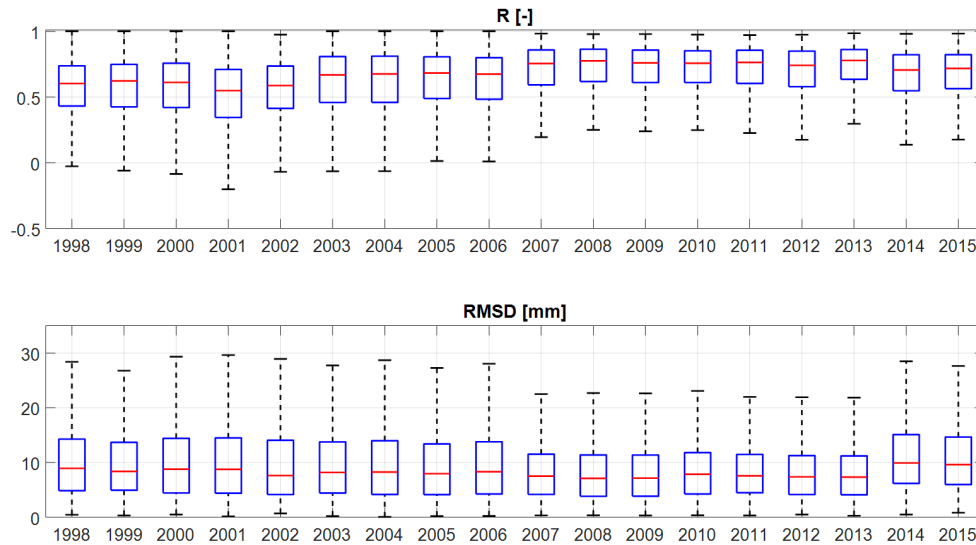


**Table 4.** Statistical summary of correlation coefficient ( $R$ ) and difference in mean rainfall for the analyzed data sets against SM2RAIN-CCI during the period 2000–2013. For each score, the median, mean, minimum, maximum and standard deviation values are reported.

Data set	$R$ (-)					Rainfall difference (mm)				
	Median	Mean	Min.	Max.	SD	Median	Mean	Min.	Max.	SD
GPCC-FDD	0.62	0.59	-0.06	0.87	0.16	-59.12	-127.67	-5474.30	360.66	257.67
TRMMRT	0.54	0.53	-0.06	0.89	0.18	-136.84	-176.16	-4944.90	1211.40	284.59
CMORPH	0.51	0.50	-0.29	0.89	0.19	-15.46	-55.31	-3678.20	1405.50	286.33
ERA-I	0.58	0.56	-0.17	0.85	0.16	-37.64	-136.91	-10 195.00	691.64	413.91
MSWEP	0.69	0.66	-0.05	0.91	0.16	-71.72	-132.13	-4633.60	1271.10	257.22



**Figure 6.** Mean monthly rainfall estimated by GPCC-FDD (blue line) and the new CCI-derived rainfall data set (red line) over the six analysis boxes throughout North America (A), South America (B), Europe (C), Sahel (D), Asia (E), India (F) and Australia (G) during the period 1998–2013. The blue lines draw the mean monthly rainfall estimated by GPCC-FDD when both a ground-based and a SM-derived rainfall estimate is available.



**Figure 7.** Yearly box plots for the correlation coefficients ( $R$ ) and root mean square differences (RMSDs, in mm) between SM2RAIN-CCI and MSWEP obtained on a global scale at  $0.25^\circ$  spatial resolution during the period 1998–2015. For each box, the red line represents the median values and the blue box represents the 25th and 75th percentile, while the black dotted whiskers extend to the most extreme data points.

(i.e. when both GPCC-FDD and SM2RAIN-CCI provide a rainfall estimate), the two estimates are very close to each other, for the entire analysis period.

### 3.1 SM2RAIN-CCI performance over time

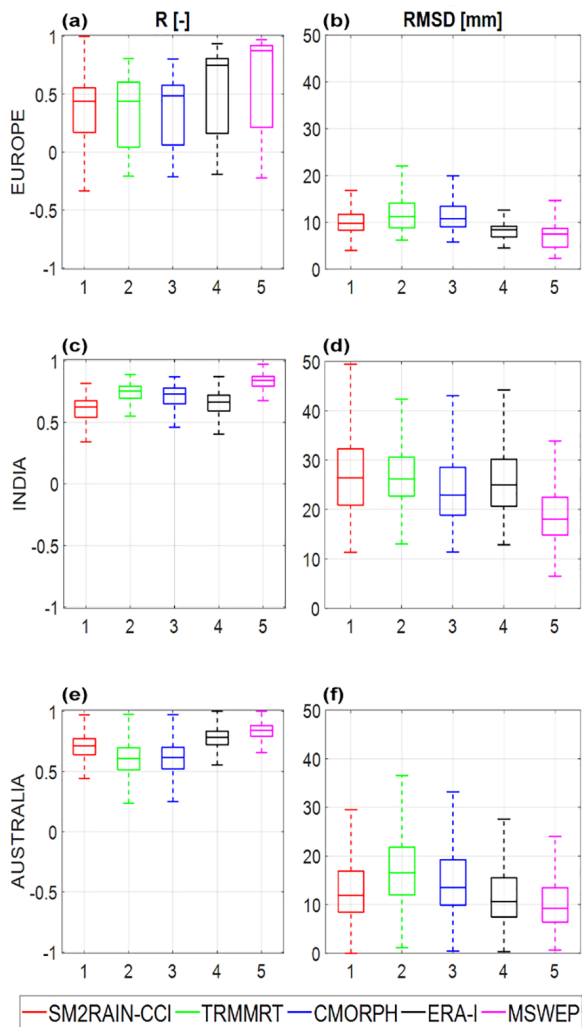
Figure 7 shows box plots of  $R$  and RMSD values between SM2RAIN-CCI and MSWEP on a yearly scale. The use of an independent benchmark removes the effect of the algorithm calibration against GPCC-FDD data set and (partly) the effect of in situ stations density on the benchmark reliability. The comparison is carried out over the  $\pm 50^\circ$  latitude band. The SM2RAIN-CCI rainfall product generally agrees well with MSWEP. An increasing trend in the performance can be observed over time during the analysis period, highlighting the impact of data availability on estimation uncertainty. The most significant improvements can be observed in 2003 and 2007, corresponding to the start of AMSR-E and ASCAT operations, respectively. Figure 7 shows that the SM2RAIN-CCI product provides the lowest  $R$  (0.57) during 2001 and the highest (0.80) during 2013. Similar patterns are found for the RMSD score. The improvements are not just recognizable in the median values, but also in the spread of  $R$  and RMSD values within each year.

### 3.2 Regional-scale assessment

For the regional-scale assessment, three macroareas with a high rain gauge station density are selected, which are Europe, India and Australia. SM2RAIN-CCI estimates are compared against data from these ground-based measurements on the  $0.25^\circ$  scale. Three ground-based rainfall data sets are

considered here to test the skills of SM2RAIN-CCI to identify precipitation events independently from GPCC-FDD, used for calibration.

The comparison over Europe is carried out by considering the so-called E-OBS rainfall data set (Haylock et al., 2008) as a benchmark. This data set provides daily rainfall estimates over the European area at  $0.25^\circ$  spatial resolution starting from 1950. The estimates are obtained by interpolating rainfall values from gauge stations over Europe via a three-step kriging procedure. For this analysis, we consider the region between  $-9.875^\circ$  W and  $24.875^\circ$  E longitude and between  $28.125$  and  $59.875^\circ$  N latitude. Due to the TRMM orbit geometry, the considered TRMMRT data set covers only the area between  $28.125$  and  $49.875^\circ$  N latitude. The analysis is carried out during the period 2002–2015, in order to avoid considering partly the data calibrated during the period 1998–2001. Figure 8 shows  $R$  and RMSD statistics against E-OBS for 5 days of accumulated rainfall. As can be seen, SM2RAIN-CCI provides a median  $R$  lower than 0.5, close to that provided by TRMMRT and CMORPH. All rainfall products show a large variability in terms of  $R$ , ranging between  $-0.4$  and almost 1. In terms of RMSD, all the products show median values close to 10 mm, with values ranging between approximately 5 and 20 mm. ERA-Interim provided very good performance, in terms of both  $R$  and RMSD, due to the stratiform-dominated precipitation regime over Europe which is well identified by atmospheric models, guaranteeing good performance of the ERA-Interim reanalysis product. It is worth noting that ERA-Interim does not use rain gauge data, but only other meteorological variables. MSWEP provided the best performance over Europe, due to the merging



**Figure 8.** Correlation coefficient (a, c, e) and root mean square difference (RMSD, b, d, f) box plots obtained by comparing SM2RAIN-CCI (in red), TRMMRT (in green), CMORPH (in blue), ERA-Interim (in black) and MSWEP (in magenta) with gauge-based data sets over Europe (a, b), India (c, d) and Australia (e, f).

of different rainfall products. In general, SM2RAIN-CCI performs quite well in southern Europe (Italy, Spain and southern France). In central and northern Europe, observations are subject to a high selective masking of frozen soil and snow, which reduces the temporal observation density and hence also the SM2RAIN retrieval accuracy.

The analysis over India is carried out during the period 2002–2015 using observed rainfall data provided by the India Meteorological Department. The considered region spans from 70 to 90° E longitude and from 5 to 25° N latitude. As can be seen in Fig. 8,  $R$  values are generally higher than those obtained over Europe, most likely due to the strong seasonal signal. The SM2RAIN-CCI data set shows a median  $R$  of 0.60, which is slightly lower than that achieved by TRMMRT, CMORPH, ERA-Interim and MSWEP. In terms

of RMSD, values are generally higher than over Europe, which result from the larger annual precipitation amount. SM2RAIN-CCI performs very well over India and is less reliable along the coast and in the northern parts of the country due to the impact of the Himalayas.

Over Australia, the Australian Water Availability Project rainfall data, observed during the period 2010–2013, are used as benchmarks. The analysis box spans from 120 to 160° E longitude and from 10 to 40° S latitude. The analysis shows very good results in terms both of  $R$  and RMSD (Fig. 8). SM2RAIN-CCI provides a median  $R$  of 0.71 which is higher than that obtained with TRMMRT and CMORPH. Moreover,  $R$  values are consistently higher than 0.5 in the entire macroregion. In terms of RMSD, median value of 11.90 mm is obtained for SM2RAIN-CCI, while TRMMRT and CMORPH provided median values of 16.56 and 13.52 mm, respectively. The large variability of errors is related to the different rainfall regimes in Australia, i.e. tropical climate in the northern sector and drier conditions in the inland part. In tropical rainfall regimes, the SM2RAIN algorithm is often subject to close-to-saturation soil conditions, which lead to a general underestimation of precipitation. Results are consistent with those of Tarpanelli et al. (2017), who applied the SM2RAIN algorithm to multiple satellite SM products over India.

#### 4 Data availability

The soil moisture derived rainfall dataset is freely available at <https://doi.org/10.5281/zenodo.846259>.

#### 5 Conclusions

This study presents a new rainfall data set obtained through the application of the SM2RAIN algorithm (Brocca et al., 2014) to the ACTIVE and the PASSIVE ESA CCI SM products, named SM2RAIN-CCI (Dorigo et al., 2017), during the period 1998–2015. The algorithm is calibrated using the GPCC-FDD data set. Due to the different characteristics of the satellite sensors used for creating the input SM data sets, three different calibration periods are considered: (1) 1998–2001, (2) 2002–2006 and (3) 2007–2013. SM2RAIN-CCI data set is available on a global scale (over land) with a daily temporal sampling on a 0.25° regular grid. A mask is applied to the data set in order to remove pixels and observations characterized by high topographic complexity, frozen soil and high snow probability.

The SM2RAIN-CCI data set is compared to three different global (or quasi-global) state-of-the-art rainfall products in order to check its capability in rainfall estimation. In general, the SM2RAIN-CCI shows relatively good performance in precipitation estimation, especially during the 2007–2013 period (see Figs. 4 and 5). On a global scale and for the entire analysis period, 5-day SM2RAIN-CCI rainfall estimates

provide a median  $R$  of 0.67 when compared to MSWEP (see Fig. 7).

The product is further evaluated over three macroareas (Europe, India and Australia) where it provides satisfactory results, in terms of both  $R$  and RMSD, when compared to spatially interpolated high-density rain gauge measurement networks (see Fig. 8). Higher errors are found over India and Australia due to the larger total rainfall amounts of precipitation. However, the analysis also showed relatively good results over five other considered macroregions (see Fig. 5) when compared to GPCC-FDD. In these regions, the impact of reduced temporal coverage on retrieval accuracy is clearly visible.

The multi-sensor data sets provided by ESA CCI and the application of SM2RAIN could open up new perspectives and opportunities in the use of satellite rainfall products over developing countries or in remote areas with nonexisting or spatially sparse ground monitoring networks. The new product is potentially suitable for several applications in the domains of climate (due to the long temporal coverage) and hydrology (due to good capabilities in accumulated rainfall estimation), complementing other state-of-the-art rainfall products. It is worth nothing that due to the calibration strategy, the data set is not suitable for trend analysis. Moreover, the SM2RAIN-CCI is completely independent from other existing state-of-the-art precipitation products, therefore offering an additional long-term data set that can be used for independently evaluating these global-scale precipitation products as shown by Massari et al. (2017).

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** The authors gratefully acknowledge the E-OBS data set from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project (<http://www.ecad.eu>). This work is supported by the ESA Climate Change Initiative (CCI, <http://www.esa-soilmoisture-cci.org/>, contract no. 400011226/14/I-NB) and the earthH2Observe project (European Union's Seventh Framework Programme, grant agreement no. 603608).

Edited by: David Carlson

Reviewed by: Hylke Beck and one anonymous referee

## References

Abera, W., Brocca, L., and Rigon, R.: Comparative evaluation of different satellite rainfall estimation products and bias correction in the Upper Blue Nile (UBN) basin, *Atmos. Res.*, 178–179, 471–483, 2016.

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatel-

lite Observations for Hydrological and Climate Studies, *B. Am. Meteorol. Soc.*, 96, 69–83, <https://doi.org/10.1175/BAMS-D-13-00068.1>, 2015

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, *Hydrol. Earth Syst. Sci.*, 21, 589–615, <https://doi.org/10.5194/hess-21-589-2017>, 2017.

Brocca, L., Melone, F., Moramarco, T., and Wagner, W.: A new method for rainfall estimation through soil moisture observations, *Geophys. Res. Lett.*, 40, 853–858, 2013.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural raingauge: estimating rainfall from global satellite soil moisture data, *J. Geophys. Res.*, 119, 5128–5141, 2014.

Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zuecco, G., Pianezzola, L., Borga, M., Matgen, P., and Martínez-Fernández, J.: Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of SM2RAIN algorithm, *J. Hydrol. Hydromech.*, 65, 201–209, 2015.

Brocca, L., Pellarin, T., Crow, W. D., Ciabatta, L., Massari, C., Ryu, D., Su, C. H., Rüdiger, C., and Kerr, Y.: Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia, *J. Geophys. Res.-Atmos.*, 121, 12062–12079, <https://doi.org/10.1002/2016JD025382>, 2016.

Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Gabellani, S., Puca, S., Rinollo, A., and Wagner, W.: Integration of satellite soil moisture and rainfall observations over the Italian territory, *J. Hydrometeorol.*, 16, 1341–1355, 2015.

Ciabatta, L., Massari, C., Brocca, L., Reimer, C., Hann, S., Paulik, C., Dorigo, W., and Wagner, W.: Using Python® language for the validation of the CCI soil moisture products via SM2RAIN, *PeerJ Preprints*, 4, e2131v2, <https://doi.org/10.7287/peerj.preprints.2131v2>, 2016.

Ciabatta, L., Marra, A. C., Panegrossi, G., Casella, D., Sanò, P., Dietrich, S., Massari, C., and Brocca, L.: Daily precipitation estimation through different microwave sensors: verification study over Italy, *J. Hydrol.*, 545, 436–450, 2017.

Crow, W. T., Huffman, G. F., Bindlish, R., and Jackson, T. J.: Improving satellite rainfall accumulation estimates using spaceborne soil moisture retrievals, *J. Hydrometeorol.*, 10, 199–212, 2009.

Crow, W. T., Van Den Berg, M. J., Huffman, G. J., and Pellarin, T.: Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART), *Water Resour. Res.*, 47, W08521, <https://doi.org/10.1029/2011WR010576>, 2011.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Holm, E. V., Isaksen, L., Kallberg, P., Kohler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thepaut, J. N., and Vitard, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, 2011.

- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., and Ropelewski, C. F.: Validation of satellite rainfall products over East Africa's complex topography, *Int. J. Remote Sens.*, 28, 1503–1526, 2007.
- Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., and Naeimi, V.: Error characterisation of global active and passive microwave soil moisture datasets, *Hydrol. Earth Syst. Sci.*, 14, 2605–2616, <https://doi.org/10.5194/hess-14-2605-2010>, 2010.
- Dorigo, W. A., de Jeu, R. A. M., Chung, D., Parinussa, R. M., Liu, Y. Y., Wagner, W., and Fernandez-Prieto, D.: Evaluating global trends (1988–2010) in harmonized multi-satellite surface soil moisture, *Geophys. Res. Lett.*, 39, L18405, <https://doi.org/10.1029/2012GL052988>, 2012.
- Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, *Remote Sens. Environ.*, 162, 380–395, <https://doi.org/10.1016/j.rse.2014.07.023>, 2015.
- Dorigo, W. A., Wagner, W., Albergel, C., Abrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understating: state-of-the-art and future decision, *Remote Sens. Environ.*, 203, 185–215, <https://doi.org/10.1016/j.rse.2017.07.001>, 2017.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes *Scientific Data*, 2, 150066, <https://doi.org/10.1038/sdata.2015.66>, 2015.
- Gao, H., Wood, E. F., Jackson, T. J., Drusch, M., and Bindlish, R.: Using TRMM/TMI to retrieve surface soil moisture over the southern United States from 1998 to 2002, *J. Hydrometeorol.*, 7, 23–38, <https://doi.org/10.1175/JHM473.1>, 2006.
- Gruber, A., Dorigo, W. Crow, W., and Wagner, W.: Triple collocation based merging of satellite soil moisture retrievals, *IEEE T. Geosci. Remote*, 55, 6780–6792, <https://doi.org/10.1109/TGRS.2017.2734070>, 2017.
- Haylock, M. R., N., Hofstra, A. M. G., Klein Tank, E. J., Klok, P. D., Jones, and New, M.: A European daily high-resolution gridded dataset of surface temperature and precipitation, *J. Geophys. Res.*, 113, D20119, <https://doi.org/10.1029/2008JD010201>, 2008.
- Herold, N., Behrangi, A., and Alexander, L. V.: Large uncertainties in observed daily precipitation extremes over land, *J. Geophys. Res.-Atmos.*, 122, 668–981, 2017.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement (GPM) mission, *B. Am. Meteorol. Soc.*, 95, 701–722, 2014.
- Hsu, K., Gao, X., Sorooshian, S., and Gupta, H. V.: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks, *J. Appl. Meteorol.*, 36, 1176–1190, 1997.
- Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeorol.*, 8, 38–55, 2007.
- Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, *J. Hydrometeorol.*, 5, 487–503, 2004.
- Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J.-P., Ferrazoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S. E., Leroux, D., Mialon, A., and Delwart, S.: The SMOS Soil Moisture Retrieval Algorithm, *IEEE T. Geosci. Remote*, 50, 1384–1403, 2012.
- Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C. D., and Randel, D.: Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme, *J. Hydrometeorol.*, 17, 383–400, 2016.
- Kim, S., Parinussa, R. M., Liu, Y. Y., Johnson, F. M., and Sharma, A.: A framework for combining multiple soil moisture retrievals based on maximizing temporal correlation, *Geophys. Res. Lett.*, 42, 6662–6670, 2015.
- Koster, R. D., Brocca, L., Crow, W. T., Burgin, M. S., and De Lannoy, G. J. M.: Precipitation Estimation Using L-Band and C-Band Soil Moisture Retrievals. submitted to *Water Resources Research*, *Water Resour. Res.*, 52, 7213–7225, <https://doi.org/10.1002/2016WR019024>, 2016.
- Kucera, P. A., Ebert, E. E., Turk, F. J., Levizzani, V., Kirschbaum, D., Tapiador, F. J., Loew, A., and Borsche, M.: Precipitation from space: Advancing earth system science, *B. Am. Meteorol. Soc.*, 94, 365–375, 2013.
- Li, L., Gaiser, P. W., Gao, B. C., Bevilacqua, R. M., Jackson, T. J., Njoku, E. G., Rudiger, C., Calvet, J. C., and Bindlish, R.: Wind-Sat global soil moisture retrieval and validation, *IEEE T. Geosci. Remote*, 48, 2224–2241, 2010.
- 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., and Evans, J. P.: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.*, 15, 425–436, <https://doi.org/10.5194/hess-15-425-2011>, 2011.
- Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sens. Environ.*, 123, 280–297, <https://doi.org/10.1016/j.rse.2012.03.014>, 2012.
- Massari, C., Crow, W., and Brocca, L.: An assessment of the performance of global rainfall estimates without ground-based observations, *Hydrol. Earth Syst. Sci.*, 21, 4347–4361, <https://doi.org/10.5194/hess-21-4347-2017>, 2017.
- McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, *Geophys. Res. Lett.*, 41, 6229–6236, <https://doi.org/10.1002/2014GL061322>, 2014.
- Owe, M., de Jeu, R. A. M., and Walker, J.: A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, *IEEE T. Geosci. Remote*, 39, 1643–1694, 2001.
- Owe, M., De Jeu, R. A. M., and Holmes, T. R. H.: Multi-sensor historical climatology of satellite derived global

- land surface moisture, *J. Geophys. Res.*, 113, F01002, <https://doi.org/10.1029/2007JF000769>, 2008.
- Panegrossi, G., Casella, D., Dietrich, S., Marra, A. C., Petracca, M., Sanò, P., Mugnai, A., Baldini, L., Roberto, N., Adirosi, E., Cremonini, R., Bechini, R., Vulpiani, G., and Porcù, F.: (2016). Use of the GPM constellation for monitoring heavy precipitation events over the Mediterranean region, *IEEE J. Sel. Top. Appl.*, 9, 2733–2753, <https://doi.org/10.1109/JSTARS.2016.2520660>, 2016.
- Parinussa, R., Holmes, T. R. H., and de Jeu, R.: Soil moisture retrievals from the Windsat spaceborne polarimetric microwave radiometer, *IEEE T. Geosci. Remote*, 50, 2683–2694, 2012.
- Pellarin, T., Louvet, S., Gruhier, C., Quantin, G., and Legout, C.: A simple and effective method for correcting soil moisture and precipitation estimates using AMSR-E measurements, *Remote Sens. Environ.*, 136, 28–36, 2013.
- Schamm, K., Ziese, M., Raykova, K., Becker, A., Finger, P., Meyer-Christoffer, A., and Schneider, U.: GPCC Full Data Daily Version 1.0 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data, [https://doi.org/10.5676/DWD\\_GPCC/FD\\_D\\_V1\\_100](https://doi.org/10.5676/DWD_GPCC/FD_D_V1_100), 2015.
- Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, *J. Geophys. Res.*, 103, 7755, <https://doi.org/10.1029/97JC03180>, 1998.
- Tarpanelli, A., Massari, C., Ciabatta, L., Filippucci, P., Amarnath, G., and Brocca, L.: Exploiting a constellation of satellite soil moisture sensors for accurate rainfall estimation, *Adv. Water Res.*, 108, 249–255, <https://doi.org/10.1016/j.advwatres.2017.08.010>, 2017.
- Trenberth, K. E. and Asrar, G. R.: Challenges and opportunities in water cycle research: WCRP contributions, *Surv. Geophys.*, 35, 515–532, <https://doi.org/10.1007/s10712-012-9214-y>, 2014.
- Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J.: Rainfall and sampling uncertainties: A rain gauge perspective, *J. Geophys. Res.*, 113, D11102, <https://doi.org/10.1029/2007JD009214>, 2008.
- Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., and Ertl, M.: Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-7, 315–321, <https://doi.org/10.5194/isprsannals-I-7-315-2012>, 2012.
- Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Zuger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Bloeschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., and Rubel, F.: The ASCAT soil moisture product: Specifications, validation results, and emerging applications, *Meteorol. Z.*, 22, 5–33, 2013.
- Wanders, N., Pan, M., and Wood, E. F.: Correction of real-time satellite precipitation with multi-sensor satellite observations of land surface variables, *Remote Sens. Environ.*, 160, 206–221, 2015.
- Zhan, W., Pan, M., Wanders, N., and Wood, E. F.: Correction of real-time satellite precipitation with satellite soil moisture observations, *Hydrol. Earth Syst. Sci.*, 19, 4275–4291, <https://doi.org/10.5194/hess-19-4275-2015>, 2015.