Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

Earth System Science Data
Open Access Discussions

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

Luca Ciabatta[1,4], Christian Massari[1], Luca Brocca[1], Alexander Gruber[2], Christoph Reimer[3], Sebastian Hahn[2], Christoph Paulik[2], Wouter Dorigo[2], Richard Kidd[3], Wolfgang Wagner[2]

[1] Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy
[2] Department of Geodesy and Geoinformation, TU Wien, Vienna, Austria
[3] Earth Observation Data Centre, Vienna, Austria
[4] Department of Civil and Environmental Engineering, University of Perugia, Italy

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

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

ABSTRACT

Accurate and long-term rainfall estimates are the main inputs for several applications, spanning 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° 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-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 modelled data set (ERA-Interim). The assessment is carried out on a global scale at 1° of spatial sampling and 5-day 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) and BIAS (median value <-14.44%) during the evaluation period. The validation has been also carried out at original resolution (0.25°) over Europe, Australia and other 5 areas worldwide to test the capabilities of the data set to correctly identify rainfall events under different climate and precipitation regimes.

The CCI-SM derived rainfall data set is freely available at http://www.esa-soilmoisture-cci.org 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 modelling (Dinku et al., 2007).

Ground stations provide very accurate local estimates of rainfall (Villarini et al., 2008) and are considered the most accurate source of rainfall data for modelling and process monitoring. However, two main issues limit their usefulness. Firstly, they are characterized by a non-homogenous coverage (Kidd et al., 2016) throughout the globe and, secondly, 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 radars, which are able to provide measurements that are more representative of the actual rainfall spatial variability. However, also ground meteorological radars are 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 raingauges 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° every 30 minutes, 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 a., 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 provides continuous rainfall estimates since 1983. Despite the relative advantages of having an estimate of rainfall in every place of 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 at 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 using satellite surface SM 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, SMOS, etc… (Brocca et al., 2013, 2014, 2016; Ciabatta et al., 2015, 2017; Koster et al., 2016, Massari et al. 2017).

With the aim of facing and monitoring climate change, 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 (ECV). 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, Liu et al., 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., 2016). The analysis is performed on a global scale at 1° spatial sampling, during the period 1998-2015. In addition, a regional scale analysis at 0.25° 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, satellite-based and ground-based observations are intercompared with the new SM2RAIN-CCI data set. In particular, the two following products are considered as benchmark:

GPCC-FDD, available at 1° spatial sampling during the period 1988-2013 (ground based data set) at daily temporal resolution, used for calibrating SM2RAIN;

MSWEP, available from 1st January 1979 to 31st December 2015 at 0.25° spatial sampling on a daily basis (combination of models, ground measurements and satellite observations), used as independent benchmark for the yearly global analysis.

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 1st March 2000 to present at 0.25° spatial resolution for the ± 50° latitude band every 3 hours (only satellite);
2) CMORPH raw data (hereinafter referred to as CMORPH), available from 1st January 2000 to present at 0.25° spatial resolution for the ± 60° latitude band every 3 hours (only satellite);
3) ERA-Interim reanalysis product, available from 1st January 1978 to present at 0.77° 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 ± 50° 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 60000 stations are used. GPCC-FDD is provided on a global scale over a grid with 1° spatial sampling and on a daily basis. The product is available for the period 1988-2013. Here, GPCC-FDD is used as benchmark 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 rainfall 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°. More information about MSWEP can be found in Beck et al. (2016).

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 ± 50° latitude band over a grid with a 0.25° spacing every 3 hours. Daily accumulated rainfall is computed by summing up all rainfall estimates within one 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

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

overpasses. The product is considered at daily temporal resolution over the 0.25° sampling TRMMRT grid for the ± 60° latitude band. In this study, CMORPH raw data version 1.0 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°. The data used have been downloaded from the ECMWF FTP (http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/) resampled over the 0.25° 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; Dorigo et al., in review). The products are provided on a global scale with a spatial sampling of 0.25° 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 Special Sensor Microwave Imager (SSM/I, operating at 19.35 GHz, Owe et al., 2008), the TRMM Microwave Imager (TMI, operating at 10.65 GHz and above, Gao et al., 2006), the Advanced Microwave Scanning Radiometer - Earth Observing System (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 frequency 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., accepted). In this study, we consider the ESA CCI SM product at version v03.2. For further details regarding the ESA CCI SM product development, sensors availability and performances the reader is referred to Liu et al., (2011; 2012), Dorigo et al., (2015; in review) and Wagner et al., (2012).

2.2 ESA CCI Soil Moisture pre-processing

Before applying SM2RAIN algorithm the following preprocessing steps are applied to the ESA CCI SM data sets. A static mask (Figure 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 Section 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 with 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. A maximum data gap of three days is considered for the temporal interpolation. Data gaps larger than three 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 preprocessed by using routines developed in Python® 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.

2.3 SM2RAIN algorithm and SM2RAIN-CCI rainfall product generation

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

$$p(t) = Z^{*} ds(t)/dt + as(t)^{b} \tag{1}$$

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 non-linearity 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, 2007-2013 on a pixel-by-pixel basis. We considered 5-day 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 also applied to the COMBINED SSM data set, but we observed a reduction of performance with respect to the individual ACTIVE and PASSIVE products (not shown), hence the COMBINED SSM data set is not considered here. In order to match the different spatial resolutions of the considered data sets, GPCC-FDD was regridded to the 0.25° CCI grid by using the *griddata* function implemented in MATLAB® 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:

$$P_{SM2RAIN-CCI} = kP_{ACT} + (1-k)P_{PAS} \tag{2}$$

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

$$k = \frac{\rho_{AB} - \rho_{AP} * \rho_{AB}}{\rho_{PB} - \rho_{AP} * \rho_{AB} + \rho_{AB} - \rho_{AP} * \rho_{PB}} \tag{3}$$

Where ρ is the Pearson correlation coefficient between two data sets with the subscript A, P and B denoting the ACTIVE, the PASSIVE and the benchmark (GPCC-FDD in this case) rainfall estimates, respectively. When one of the two data sets (P_{ACT} or P_{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 Figure 2. The data are available in netCDF format via the CCI SM FTP server. The rainfall data are provided in mm per day, over land at 0.25° 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 1st of January 1998 to 31st December 2015 with daily temporal resolution. The data are provided over a 0.25° grid on a global scale, given the native spatial resolution of SM observation of 25 and 50 km. The spatio-temporal coverage is reported in Figure 3. As it 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 satellites. 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 it 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 both in terms of correlation and RMSD, as drawn in Figure 4 for the ±50° latitude bands during the three calibration periods at 1° of spatial resolution, in order to perform a fair comparison with the benchmark. 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 it can be seen in Figure 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. 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 in Figure 5. The Figure displays the 1°x1° (± 50°) correlation maps of 5 day of accumulated rainfall (on the left) and the differences in the mean annual rainfall (on the right) 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 both in terms of R and mean annual rainfall. Non-negligible differences can be observed over the Sahara Desert, Eastern US, South America, the tropical area and over Europe, where SM2RAIN-CCI provides a smaller amount of rainfall than the other rainfall data sets. On the other hand, very good performance can be observed over Africa, Brazil, western US, India and Australia, both in terms of R and mean annual rainfall.

Seven macro-regions worldwide have been selected to check the capability of the SM2RAIN-CCI in estimating rainfall under different climatic conditions. Therefore, Mean monthly rainfall (MMR) was computed from GPCC-FDD and SM2RAIN-CCI during the period 1998-2013 within these regions, illustrated as green boxes in Figure 6. From Figure 6, one can see that the temporal rainfall patterns agree well in all considered macro-regions. 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 (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 ± 50° 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 macro-areas 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° scale.

The comparison over Europe is carried out by considering the so-called E-OBS rainfall data set (Haylock et al., 2008) as benchmark. This data set provides daily rainfall estimates over the European area at 0.25° spatial resolution starting from 1950. The estimates are obtained by interpolating via a three-step kriging procedure rainfall values from gauge stations over Europe. For this analysis, we consider the region between -9.875°W and 24.875°E longitude and between 28.125°N and 59.875°N latitude. Due to the TRMM orbit geometry, the considered TRMMRT data set covers only the area between 28.125°N and 49.875°N latitude. The analysis is carried out during the period 2002-2015, in order to avoid to consider partly the data calibrated during the period 1998-2001. Figure 8 shows R and RMSD statistics against E-OBS for 5 days 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 approximately between 5 and 20 mm. ERA Interim provided very good performance, both in terms of R and RMSD, due to the use of dense meteorological networks in Europe that guarantees good performance of the ERA-Interim reanalysis product. It is worth noting that ERA-Interim does not use raingaug data, but only other meteorological variables. MSWEP provided the best performance over Europe, due to the merging 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 (IMD). The considered region spans from 70°E to 90°E longitude and from 5°N to 25°N latitude. As can be seen in Figure 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 Himalaya.

Over Australia, the Australian Water Availability Project (AWAP) observed rainfall data during the period 2010-2013 is used as benchmark. The analysis box spans from 120°E to 160°E longitude and from 10°S to 40°S latitude. The analysis shows very good results both in terms of R and RMSD (Figure 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 macro-region. 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 mm 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 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 (Dorigo et al., 2017, in review) during the period 1998-2015, named SM2RAIN-CCI. 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 calibrations periods are considered: 1) 1998-2001; 2) 2002-2006 and 3) 2007-2013. The minimization of the RMSD between 5-day accumulated rainfall from GPCC-FDD and SM2RAIN applied to both the ACTIVE and PASSIVE ESA CCI SM data sets is used as objective function. After the calibration, the two rainfall products derived from the ACTIVE and the PASSIVE ESA CCI SM data sets are merged into the final SM2RAIN-CCI data set. 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 3 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 Figure 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 Figure 7).

The product is further evaluated over three macro-areas (Europe, India and Australia) where it provides satisfactory results, both in terms of R and RMSD when compared to spatially interpolated high-density rain gauge measurement networks (see Figure 8). Higher errors are found over India and Australia due to the larger total rainfall amounts of precipitation. However, the analysis showed relatively good results also over 5 other considered macro-regions (see Figure 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 new perspectives and opportunities in the use of satellite rainfall products over developing countries or in remote areas with non-existing 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. 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).

ACKNOWLEDGEMENT

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 eartH2Observe project (European Union's Seventh Framework Programme, Grant Agreement No. 603608).

REFERENCES

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

Ashouri, H., Hsu, K.L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. Bull. Amer. Meteor. Soc., 96, 6983. Doi: http://dx.doi.org/10.1175/BAMS-D-13-00068.1

Beck, H.E., van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., de Roo, A. (2016). MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, Hydrology and Earth System Sciences, 21, 589–615, doi:10.5194/hess-21–589-2017, 2017.

Brocca, L., Melone, F., Moramarco, T., Wagner, W. (2013). A new method for rainfall estimation through soil moisture observations. Geophys. Res. Lett., 40(5), 853-858.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., Levizzani, V. (2014). Soil as a natural raingauge: estimating rainfall from global satellite soil moisture data. J. Geophys. Res., 119(9), 5128–5141.

Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zuecco, G., Pianezzola, L., Borga, M., Matgen, P., Martínez-Fernández, J. (2015). Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of SM2RAIN algorithm. Journal of Hydrology and Hydromechanics, 65, 3, 201-209.

Brocca, L., Pellarin, T., Crow, W.D., Ciabatta, L., Massari, C., Ryu, D., Su, C.H., Rüdiger, C., Kerr, Y. (2016), Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia, J. Geophys. Res. Atmos., 121, doi:10.1002/2016JD025382.

Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Gabellani, S., Puca, S., Rinollo, A., Wagner, W. (2015). Integration of satellite soil moisture and rainfall observations over the Italian territory. *Journal of Hydrometeorology*, 16(3), 1341-1355.

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

Ciabatta, L., Marra, A.C., Panegrossi, G., Casella, D., Sanò, P., Dietrich, S., Massari, C., Brocca, L. (2017). Daily precipitation estimation through different microwave sensors: verification study over Italy. Journal of Hydrology, 545:436-450.

Crow, W.T., Huffman, G.F., Bindlish, R., Jackson, T.J. (2009). Improving satellite rainfall accumulation estimates using spaceborne soil moisture retrievals. J. Hydrometeorol., 10, 199- 212.

Crow, W.T., Van Den Berg, M.J., Huffman, G.J., Pellarin, T. (2011). Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). Water Resources Research, 47 (8).

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., Vitard, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quartely Journal of the Royal Meteorological Society, 137 (656A): 553-597.

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S.J. Connor, C.F. Ropelewsky, 2007. Validation of satellite rainfall products over East Africa's complex topography. Int. J. Remote Sens., 28 (7), 1503–1526.

Dorigo, W.A., Scipal, K., Parinussa, R.M., Liu, Y.Y., Wagner, W., de Jeu, R.A.M., Naeimi, V. (2010). Error characterisation of global active and passive microwave soil moisture datasets. Hydrology and Earth System Sciences, 14-12, 2605-2616.

Dorigo, W. A., R. A.M. de Jeu, D. Chung, R. M. Parinussa, Y. Y. Liu, W. Wagner, and D. Fernandez-Prieto (2012), Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture, Geophys. Res. Lett., VOL. 39, L18405, 7 PP. 2012 doi:10.1029/2012GL052988.

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., Kidd, R. (2015). Evaluation of the ESA CCI soil moisture product using ground-based observations. Remote Sensing of Environment, doi: 10.1016/j.rse.2014.07.023.

Dorigo, W.A., Wagner, W., Albergel, C., Ablrecht, 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., Lecomte, P.. (in review). ESA CCI Soil Moisture for improved Earth system understating: state-of-the-art and future decision. Submitted to Remote Sensing of Environment.

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2, 150066. doi:10.1038/sdata.2015.66 2015.

Gao, H., Wood, E. F., Jackson, T. J., Drusch, M., and Bindlish, R., 2006: Using TRMM/TMI to retrieve surface soil moisture over the southern United States from 1998 to 2002. J. Hydrometeorology, 7(1), 23-38, DOI: 10.1175/JHM473.1

Gruber, A., Dorigo, W. Crow, W., Wagner, W. (accepted). Triple collocation based merging of satellite soil moisture retrievals". IEEE Transactions on Geoscience and Remote Sensing.

Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, M. New. 2008: A European daily high-resolution gridded dataset of surface temperature and precipitation. J. Geophys. Res (Atmospheres), 113, D20119, doi:10.1029/2008JD10201

Herold, N., Behrangi, A., Alexander, L., V. (2017). Large uncertainties in observed daily precipation extremes over land. Journal of Geophysical Research (Atmospheres), 122-2, 668-981.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T. (2014). The Global Precipitation Measurement (GPM) mission. *Bull. Amer. Meteor. Soc.*, 95(5), 701-722.

Hsu, K., Gao, X., Sorooshian, S., Gupta, H.V. (1997). Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks, Journal of Applied Meteorology, Vol. 36, No. 9, 1176-1190.

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

Joyce R. J., Janowiak J. E., Arkin P. A., Xie P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, J. Hydrometeorol., 5(3), 487–503.

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.-P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S.E., Leroux, D., Mialon, A., Delwart, S. (2012). The SMOS Soil Moisture Retrieval Algorithm. IEEE Trans. Geosci. Remote Sens., 50(5), 1384-1403.

Kidd C., Matsui T., Chern J., Mohr K., Kummerow C. D., and Randel D. (2016). Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme, J. Hydrometeorol., 17(1), 383–400.

Kim, S., Parinussa, R. M., Liu, Y. Y., Johnson, F. M., Sharma, A. (2015). A framework for combining multiple soil moisture retrievals based on maximizing temporal correlation. Geophysical Research Letters, 42(16), 6662-6670.

Koster, R.D., Brocca, L., Crow, W.T., Burgin, M.S., De Lannoy, G.J.M. (2016). Precipitation Estimation Using L-Band and C-Band Soil Moisture Retrievals. submitted to Water Resources Research. Water Resour. Res., 52, doi:10.1002/2016WR019024.

Kucera, P.A., E.E. Ebert, F.J. Turk, V. Levizzani, D. Kirschbaum, F.J. Tapiador, A. Loew, M. Borsche, 2013. Precipitation from space: Advancing earth system science. Bull. Amer. Meteor. Soc., 94, 365-375.

Li, L., Gaiser, P.W., Gao, B.C., Bevilacqua, R.M., Jackson, T.J., Njoku, E.G., Rudiger, C., Calvet, J.C., Bindlish, R. (2010). WindSat global soil moisture retrieval and validation. IEEE Transactions on Geoscience and Remote Sensing, 48 (5), 2224-2241.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., van Dijk, A. I. J. M., McCabe, M. F., Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. Hydrology and Earth System Sciences, 15, 425-436, doi:10.5194/hess-15-425-2011.

Liu, Y.Y., Dorigo, W.A., Parinussa, R.M., de Jeu, R.A.M., Wagner, W., McCabe, M.F., Evans, J.P., van Dijk, A.I.J.M. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280-297, doi: 10.1016/j.rse.2012.03.014.

Massari, C., Crow, W.D., Brocca, L., (2017). An assessment of the accuracy of global precipitation estimates without gauge observations. Hydrol. Earth Science Discuss, doi:10.5194/hess-2017-163, 2017.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., & Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. Geophysical Research Letters, 41(17), 6229–6236. https://doi.org/10.1002/2014GL061322.

Owe, M., de Jeu, R.A.M., Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, IEEE Trans. Geosci. Remote Sens., 39, 1643–1694.

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

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., Porcù, F. (2016). Use of the GPM constellation for monitoring heavy precipitation events over the Mediterranean region. IEEE Trans. Geosci. Remote Sens., in press, doi:900 10.1109/JSTARS.2016.2520660.

Parinussa, R., Holmes, T.R.H., de Jeu, R. (2012). Soil moisture retrievals from the Windsat spaceborne polarimetric microwave radiometer. IEEE Transactions on Geoscience and Remote Sensing 50(7): 2683-2694.

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

Schamm, K., Ziese, M., Raykova, K., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U. (2015). GPCC Full Data Daily Version 1.0 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_D_V1_100.

Stoffelen, A. (1998): Toward the true near-surface wind speed: Error modeling and calibration using triple collocation, Journal of Geophysical Research, 103, 7755, doi:10.1029/97JC03180.

Tarpanelli A., Massari C., Ciabatta L., Filippucci P., Amarnath G., Brocca L. (...) Exploiting a constellation of satellite soil moisture sensors for accurate rainfall estimation. Submitted to Advances in Water Resources.

Trenberth, K.E., Asrar, G.R., 2014. Challenges and opportunities in water cycle research: WCRP contributions. Surv. Geophys. 35(3), 515–532. doi:10.1007/s10712-012-9214-y.

Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J.: Rainfall and sampling uncertainties: A rain gauge perspective, Journal of Geophysical Research, 113, D11 102, doi:10.1029/2007JD009214, http://doi.wiley.com/10.1029/2007JD009214, 2008.

Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., Ertl, M. (2012). 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 (ISPRS Annals), Volume I-7, XXII ISPRS Congress, Melbourne, Australia, 25 August-1 September 2012, 315-321.

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., Rubel, F. (2013). The ASCAT soil moisture product: Specifications, validation results, and emerging applications. *Meteorologische Zeitschrift*, 22(1), 5-33.

Wanders, N., Pan, M., Wood, E.F. (2015). Correction of real-time satellite precipitation with multi-sensor satellite observations of land surface variables. *Remote Sensing of Environment*, 160, 206-221.

Zhan, W., Pan, M., Wanders, N., Wood, E.F. (2015). Correction of real-time satellite precipitation with satellite soil moisture observations. *Hydrol. Earth Syst. Sci.*, 19, 4275-4291.

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

TABLES

Sensor (Active/Passive)	Temporal interval
AMI-WS / SSMI & TMI	1998-01-01 to 2002-06-30
AMI-WS / AMSR-E	2002-07-01 to 2006-12-31
ASCAT-A & ASCAT B / AMSR-E & Windsat & SMOS & AMSR2	2007-01-01 to 2013-12-31
AMI-WS / SSMI & TMI & AMSR-E	1992-01-01 to 2006-12-31
ASCAT-A & ASCAT-B / AMSR-E & Windsat & SMOS & AMSR2	2007-01-01 to 2013-12-31

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

FIGURES

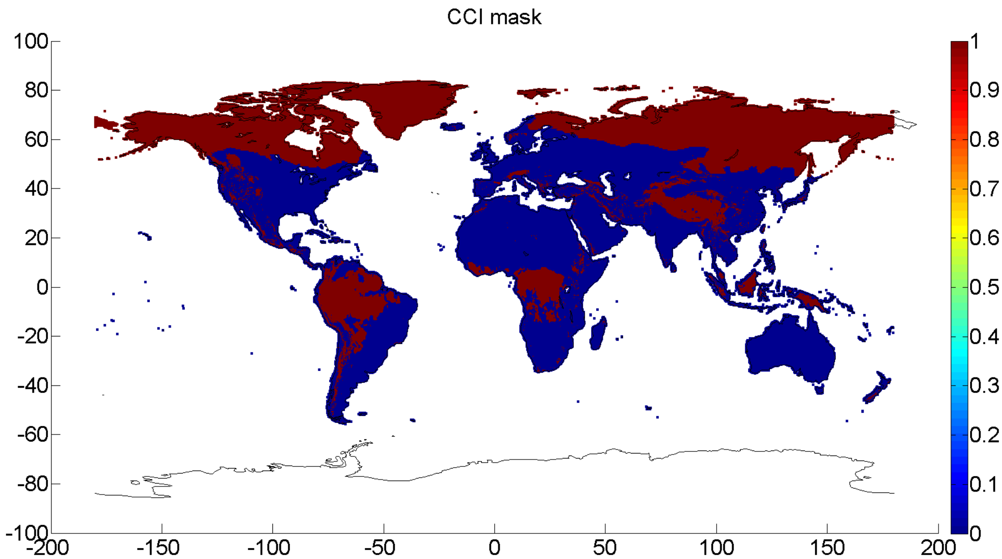
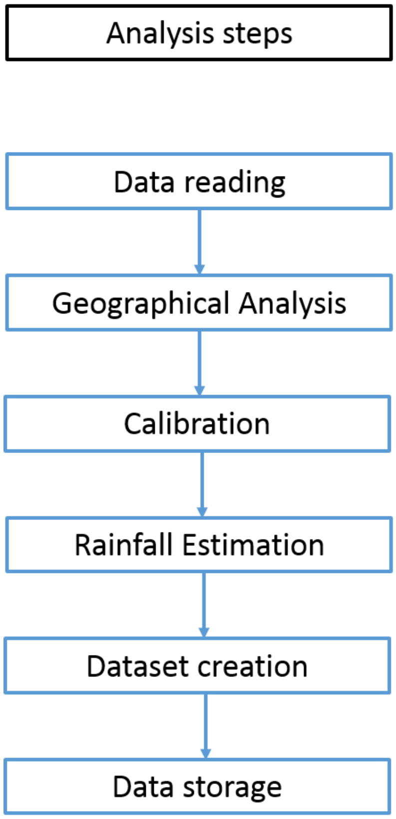


Figure 1 – Data mask used for remove areas (red areas) characterized by issues in the soil moisture retrieval.

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017



515

516 **Figure 2 - Analysis framework.**

517

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

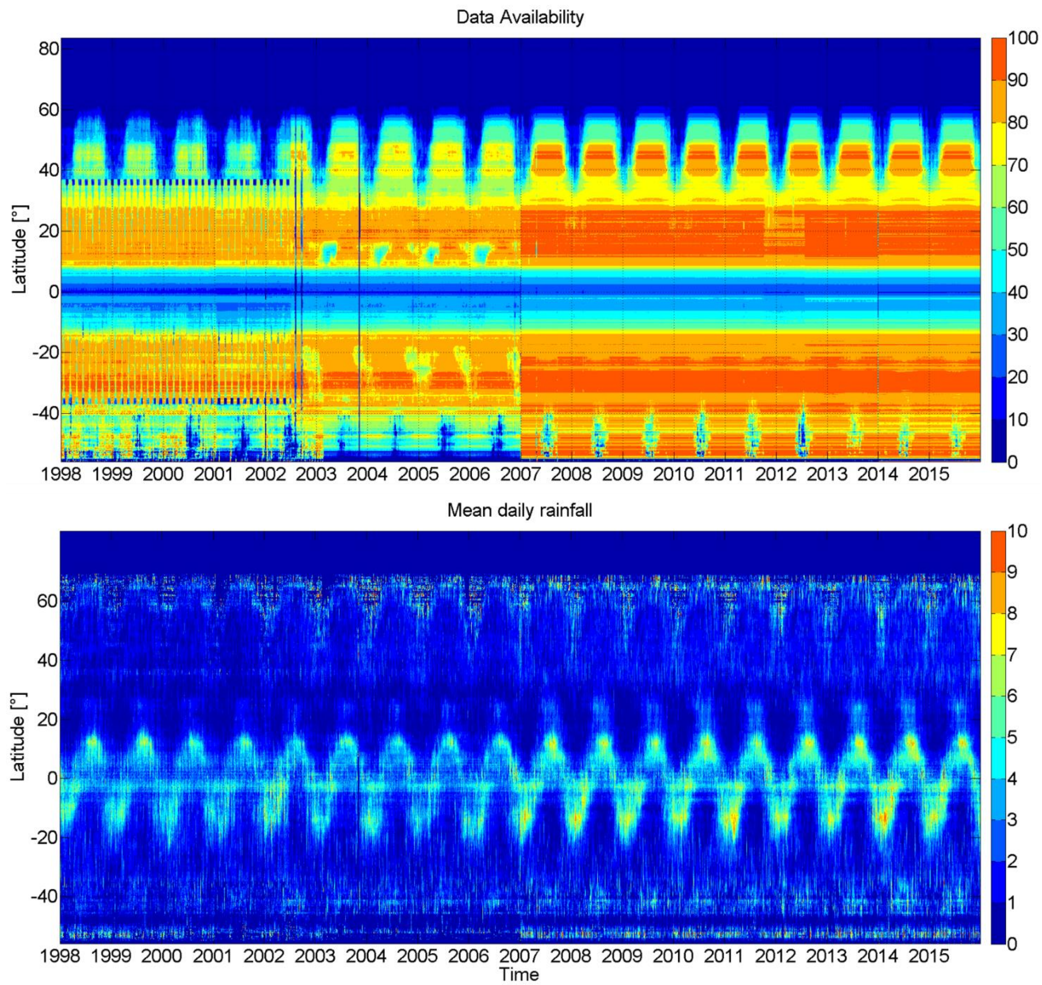


Figure 3 – Hovmoller plot showing the spatial-temporal data availability, in percentage of the total annual available data (upper panel) and the mean daily rainfall (lower panel) of the SM2RAIN-CCI rainfall data set for different latitude bands.

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

Open Access Earth System Science Data Discussions

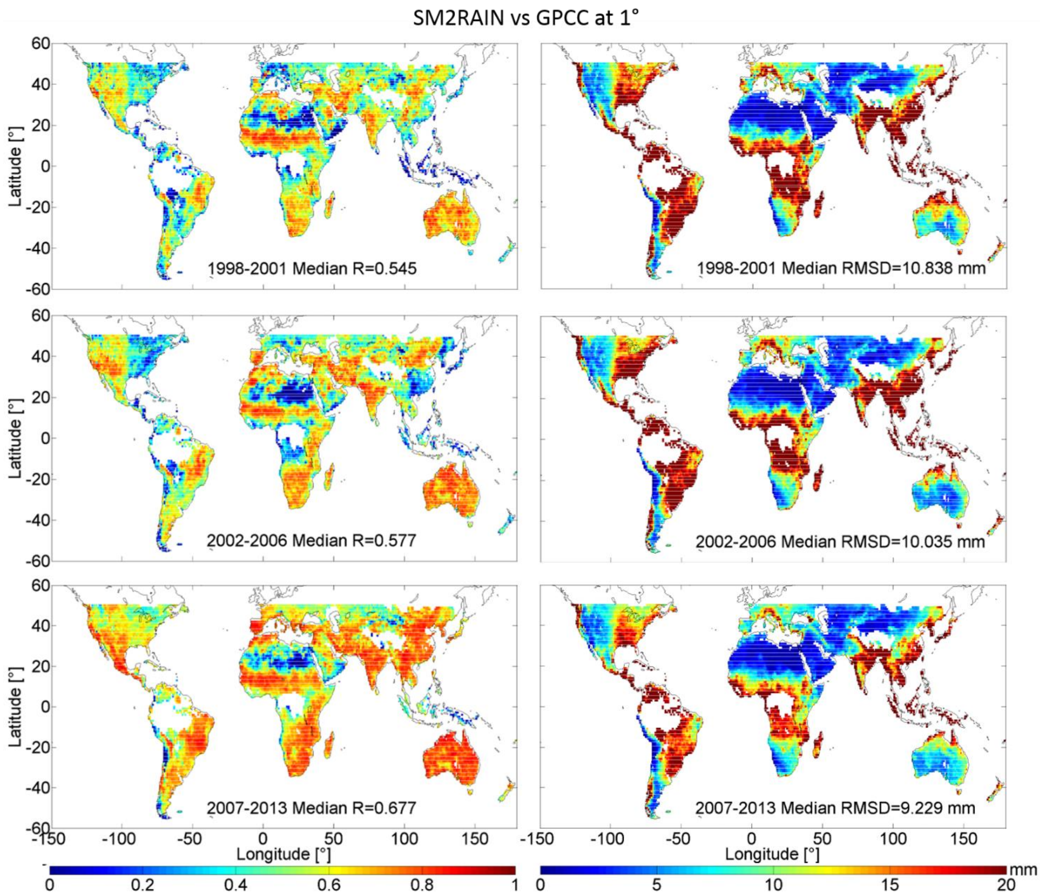
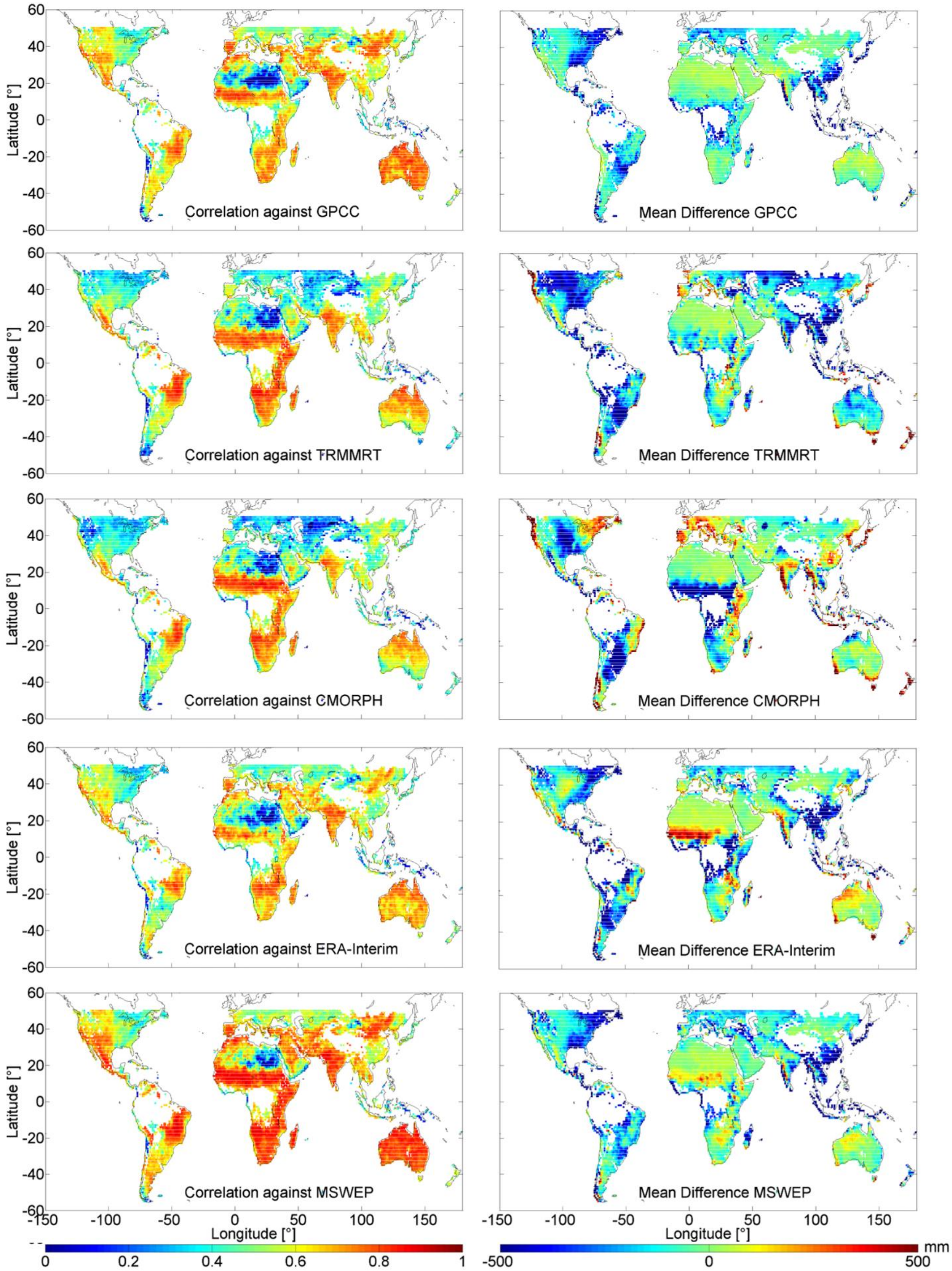


Figure 4 – Global Pearson correlation (left) and Root Mean Square Difference (right) maps obtained between GPCC-FDD and the SM2RAIN-CCI rainfall data set for 5-day accumulated rainfall during the periods 1998-2001 (upper panel), 2002-2006 (middle panel) and 2007-2013 (lower panel).

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017



527

528 **Figure 5 – Correlation maps for 5 days of accumulated rainfall (left column) and differences in mean annual rainfall (right**
529 **column) obtained by comparing (from top to bottom) SM2RAIN-CCI and GPCC (a), SM2RAIN-CCI and TRMMRT (b),**
530 **SM2RAIN-CCI and CMORPH (c), SM2RAIN-CCI and ERA-Interim (d) and SM2RAIN-CCI and MSWEP (e) at 1° of**
531 **spatial resolution.**

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

Open Access Earth System Science Data Discussions

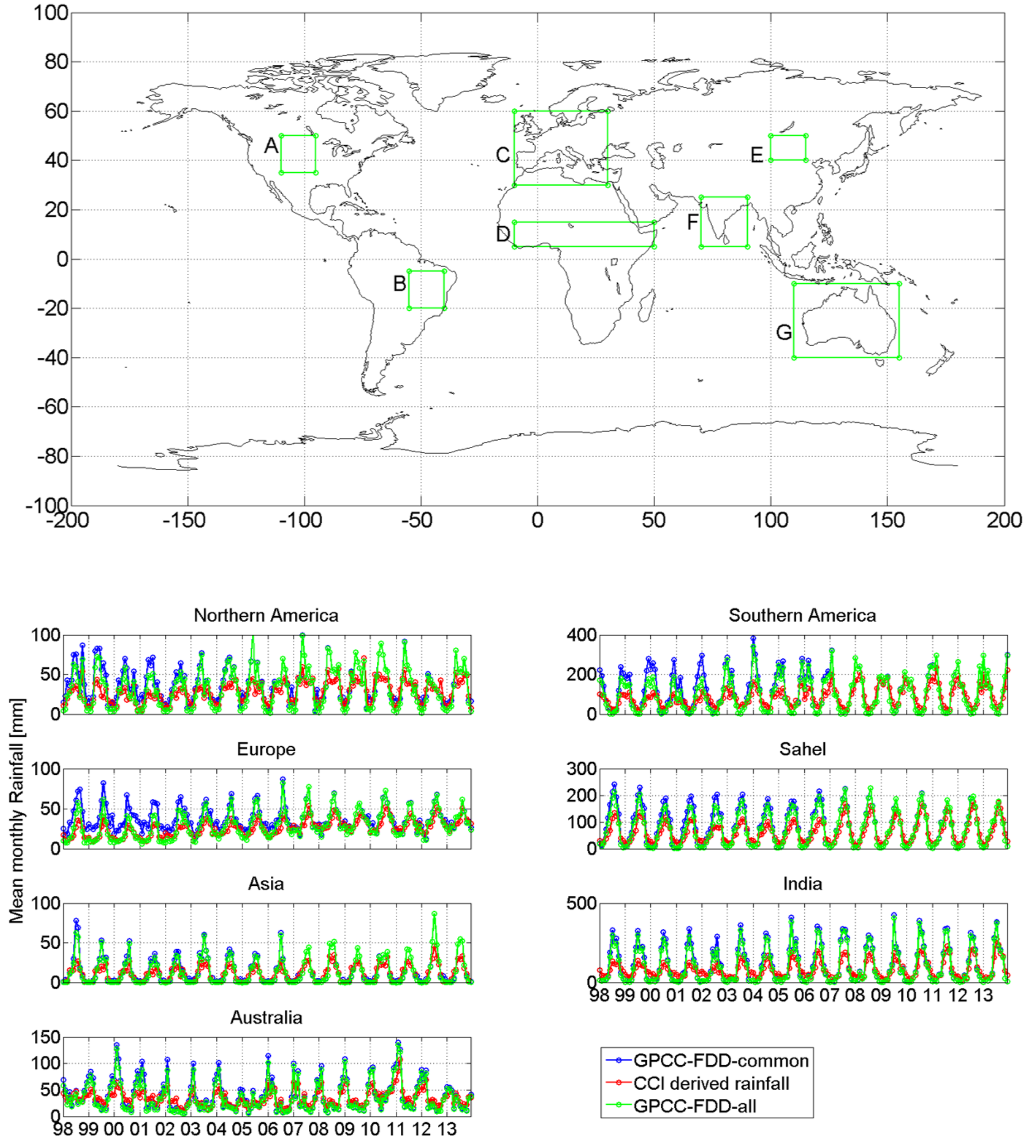


Figure 6 – Mean Monthly Rainfall estimated by GPCC-FDD (blue line) and the new CCI-derived rainfall data set (red line) over the six analysis boxes throughout North America (A), South America (B), Europe (C), Sahel (D), Asia (E), India (F) and Australia (G) during the period 1998-2013. The blue lines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

Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-86
Manuscript under review for journal Earth Syst. Sci. Data
Discussion started: 21 September 2017

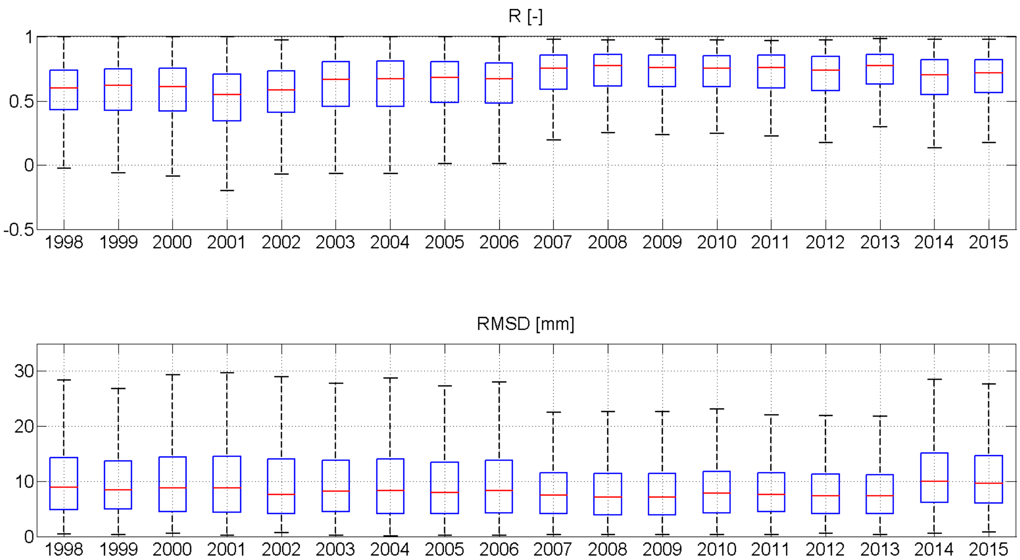
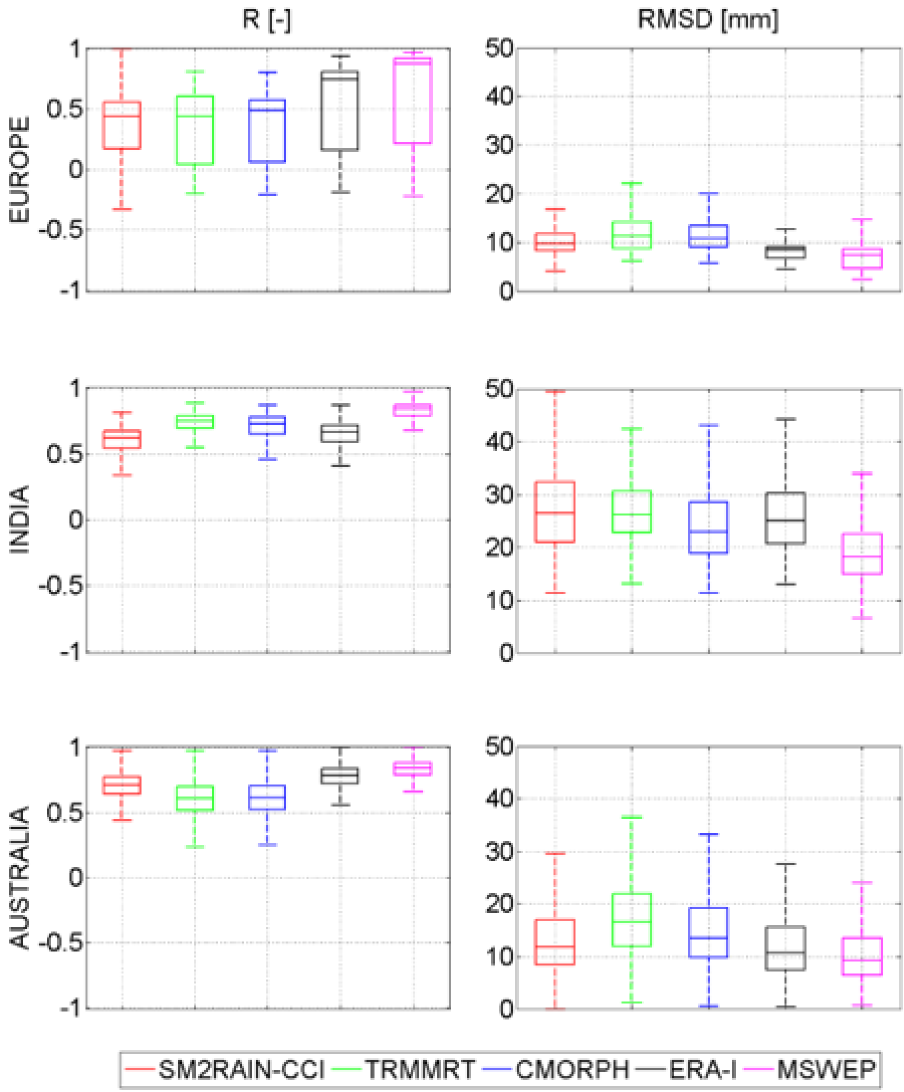


Figure 7 – Yearly boxplots for the correlation coefficients (R) and Root Mean Square Differences (RMSD, in mm) between SM2RAIN-CCI and MSWEP obtained on a global scale at 0.25° spatial resolution during the period 1998-2015. For each box, the red line represents the median values, the blue box the 25th and 75th percentile, while the black dotted whiskers extend to the most extreme data points.



544

545 **Figure 8 – Correlation coefficient (left) and Root Mean Square Difference (RMSD, right) box plots obtained by comparing**
546 **SM2RAIN-CCI (in red), TRMMRT (in green), CMORPH (in blue), ERA-Interim (in black) and MSWEP (in magenta) with**
547 **gauge-based data sets over Europe (top), India (middle) and Australia (bottom).**

548