Comments from the Reviewer #2 and the corresponding revision

SM2RAIN-ASCAT (2007–2018): GLOBAL DAILY SATELLITE RAINFALL FROM ASCAT SOIL MOISTURE

by Luca Brocca, Paolo Filippucci, Sebastian Hahn, Luca Ciabatta, Christian Massari, Stefania Camici, Lothar Schüller, Bojan Bojkov, Wolfgang Wagner

Anonymous Referee #2

An update of a satellite soil moisture-based rainfall dataset (SM2RAIN-ASCAT) is presented. The paper is fairly well written but paints an overly rosy picture of the dataset. Both the dataset and the validation exhibit a number of serious issues which must be addressed before the paper can be published.

We thank the reviewer for the valuable suggestions that helped us to clarify and improve the manuscript. A detailed answer to each comment is reported in the sequel.

As general reply to all the reviewers, we would like to underline that the paper goal is to present and describe the SM2RAIN-ASCAT global rainfall dataset and to perform a preliminary assessment of the product with respect to other state-of-the-art global rainfall products. We do not want to show a comprehensive assessment of the product. Indeed, we believe that the validation of the dataset should be performed by researchers other than the product developers (indeed the dataset is made freely available and a first paper was already published: Paredes-Trejo et al., 2019; doi:10.3390/rs11091113). Even better, we stress the importance of performing the validation by using the dataset in hydrological or agricultural applications (e.g., flood prediction and agricultural water management). The comparison with raingauges or any reference dataset could be misleading, mainly when the rainfall products include the ground observed information used for their derivation.

(1) The peak underestimation issue has not been resolved in the current release of the dataset, as revealed by both the low STDRATIO values (Figure 3) and the time series comparison (Figure 4). This major issue has been highlighted in two large precipitation dataset evaluations that have been ignored in the present study (https://www.hydrol-earth-syst-sci.net/21/6201/2017/ and https://www.hydrolearth-syst-sci.net/23/207/2019/). It is important that previously identified issues are addressed or at least discussed.

The reviewer is right; SM2RAIN-ASCAT has underestimation issue that has not been resolved completely. It will be clearly underlined in the revised manuscript. However, we want to stress that the climatological correction partly addresses this issue. A more specific CDF correction can be used for addressing the target (e.g., daily CFD matching), but we have preferred not to implement to avoid a strong dependency between SM2RAIN-ASCAT rainfall data and the reference dataset (indeed, also the reference might be wrong, particularly in poorly gauged regions).

(2) The CDF correction is based on the REF data and is thus not independent, giving the dataset an unfair advantage compared to GPM-ER in Figure 3.

In figure 3, we have used GMP-ER as a state-of-the-art reference, not to perform a comparison between the datasets. As mentioned above, the paper is not intended to perform a comprehensive assessment of SM2RAIN-ASCAT dataset, or its comparison in terms of accuracy with respect to other products. We only want to show that SM2RAIN-ASCAT is performing similarly to state-of-the-art products and, hence, can be a valuable alternative for applications using rainfall observations as input. It will be specified better in the revised manuscript.

(3) The RMSE metric should not be used for the evaluation of precipitation datasets at daily time scales as it yields misleading results (makes it seem datasets with underestimated peaks such as SM2RAIN are better). This is due to the high skewness of the precipitation distribution and the prevalence of temporal mismatches between estimated and observed precipitation peaks. The problem is illustrated in the paper by Figure 3, which shows a higher RMSE value (i.e., "worse" performance) for the bias- and CDF-corrected SM2RAIN product (BC-CDF) than for any of the uncorrected SM2RAIN products.

We agree with the reviewer that the RMSE statistic has some limitations in evaluating precipitation datasets. Indeed, we have used different statistics and in the revised manuscript we will perform the evaluation with multiple statistics also for Figure 6. Anyhow, RMSE is used in many papers evaluating precipitation datasets (and we don't believe they are all wrong), and suffers from similar limitation as any single score used for assessing a dataset. As mentioned above, we believe that the real validation should be performed using the rainfall products in the hydrological or agricultural applications. These aspects will be underlined in the revised manuscript.

(4) Only correlation and RMSE statistics are presented for the performance evaluation in Figure 6. Please remove the RMSE for the previously mentioned reason and add other metrics, such as variability ratio, bias, hit/miss ratio, frequency of wet days, peak magnitude, etc. for a more thorough performance evaluation.

In the revised manuscript we will add multiple statistics in Figure 6, similarly to Figure 3.

(5) The TC evaluation only takes into account the monthly correlation – just one aspect of dataset performance (monthly temporal dynamics). Hence the TC evaluation alone cannot be used to conclude whether a particular dataset is better or worse (as is done in the last paragraph of the abstract: "SM2RAIN-ASCAT dataset provides better performance better than GPM and GPCC in the data scarce regions of the world"). Other aspects should also be considered.

TC analysis is performed at daily time scale, not monthly time scale. Therefore, we believe TC analysis provides some information on the accuracy of the different rainfall products at daily time scale, it will be clarified in the revised manuscript.

(6) "The recent "bottom up" approach that uses satellite soil moisture observations for estimating rainfall through the SM2RAIN algorithm is suited to build long-term and consistent rainfall data record as a single polar orbiting satellite sensor is used." If this is true, why does the dataset span such a short

period (2007-2018)? All datasets listed in Table 1 (excluding IMERG) span a longer period. This statement should be revised.

The statement will be revised as we intended to say that a long-term SM2RAIN-ASCAT dataset, starting from 2007, and it is ensured until mid-2040s, can be built based the proposed approach. Sorry for the misunderstanding that will be corrected in the revised manuscript.

(7) On a related note, the evaluation of https://www.hydrol-earth-systsci.net/21/6201/2017/ (co-authored by the first author of the present study) shows that SM2RAIN-ASCAT performs worst among all precipitation datasets in terms of trend, due to the combination of data from different ASCAT sensors. So are the different ASCAT sensors consistent with each other or not? Has this trend issue been resolved in this SM2RAIN-ASCAT release? If so, this should be shown. If not, this should be communicated to the reader.

The trend issue has been solved as in the previous delivery of the SM2RAIN-ASCAT dataset (preliminary distribution) we didn't consider appropriately the availability of two ASCAT sensors (Metop-A and -B) after 2013. The dual calibration performed in this study (see lines 400-402) has been carried out exactly to address this issue. It will be clarified in the revised manuscript.

(8) In the interest of transparency the abstract should mention that the presented SM2RAIN dataset i) is limited to liquid precipitation (snowfall is not present in the dataset), ii) exhibits spurious drizzle, iii) underestimates extremes (as demonstrated by Figures 3 and 4 of the paper), and iv) potentially suffers from intercalibration issues (see comment (7)). If any of these problems have been fixed in the current release of SM2RAIN-ASCAT, this should be shown in the paper.

As suggested by the reviewer, we will clearly communicate the limitations of SM2RAIN-ASCAT dataset in the abstract of the revised manuscript. Limitations and strengths of the SM2RAIN-ASCAT dataset will be clearly illustrated. In the interest of transparency, we have made the SM2RAIN-ASCAT product freely available, and also the dataset at 1009 points that we have used for selecting the best configuration to develop the product. All the input and test datasets used in the paper are freely available and the analysis can be easily performed by the reader (note that also SM2RAIN code is made available on GitHub).

(9) "The limitations of the bottom up approach are the possibility to estimate only terrestrial rainfall and its dependence on land characteristics (e.g., low accuracy for dense vegetation coverage and complex topography, Broccaet al., 2014)." The other limitations (spurious drizzle, underestimation of extremes, and intercalibration issues) should also be mentioned here.

Limitations and strengths of the SM2RAIN-ASCAT dataset will be clearly illustrated in the revised manuscript.

(10) To my understanding the regional evaluation is performed using daily accumulations, while the triple collocation (TC) analysis is performed using monthly accumulations – correct? To avoid

confusion, please state the time scale of each specific evaluation/analysis in both the abstract and the captions of all figures.

All the analyses have been performed at daily time scale and it will be clarified in the revised manuscript.

(11) Version numbers should be assigned to the different SM2RAIN-ASCAT releases, to avoid confusion. I know there have been at least two releases. Which one is this?

The first official version of the SM2RAIN-ASCAT dataset can be considered the one presented in this paper. Indeed, the dataset has been published on Zenodo and a DOI (digital object identifier) has been assigned to the dataset to avoid confusion.

(12) Please add ERA5 to Figure 3 and make it easier to see the differences among the boxes, either by reducing the range of the y-axes or by exanding the size of the y-axes.

In the revised manuscript we will add ERA5 and we will also improve figure readability.

(13) The intro/methods part of the abstract is a bit too long, while the results/discussion part is a bit too short (just three sentences).

The abstract of the revised manuscript will be revised accordingly.

(14) "the surface runoff rate, i.e., the water that does not infiltrate into the soil and flows at the surface to the watercourses, is much lower than the rainfall rate, mainly if equation (1) is applied at coarse spatial resolution (20 km), i.e., with satellite soil moisture data." This statement does not make sense to me. Runoff can be equal to rainfall if the soil is saturated, at all scales – from hillslope to catchment.

We are saying that surface runoff is expected to be negligible at larger spatial scales due to the possibility that locally generated surface runoff (e.g., over impervious surfaces) can re-infiltrate into more permeable areas in the same pixel. Of course, this hypothesis can be not valid in some areas, but we have indirectly validated this hypothesis as we have hardly seen the ASCAT soil moisture signal to be completely saturated for more than one day. Therefore, surface runoff due to saturated soil is expected to occur very rarely at 20 km scale. This aspect will be clarified in the revised manuscript showing the number of times ASCAT soil moisture signal is saturated for more than one day.