Editor

R: We thank the Editor for his decision. In the following, we provide point-by-point answers (in black) to the referees' comments (in red). We also report the changes made in the manuscript in green indicating also their position.

Referee 1

Summary: This study proposes to use a Triple Collocation (TC) method to asses uncertainties associated with rainfall estimates using different products at the daily/1degree scale. The study addresses important issues that are relevant to the HESS readership. I recommend this manuscript for publication after minor revisions.

R: We thank the reviewer for appreciating the value of the paper and for her/his valuable comments.

Page 1:

1. Title and throughout the text: I think there is some confusion regarding the difference between accuracy and precision (uncertainty vs. error, systematic vs. random). Accuracy refers to how close a measured value is to a standard value (i.e., the "true" value). Precision describes the statistical variability. Accuracy and precision are independent. It seems to me that the TC method provides a measure of precision (error variances and correlations), but no information regarding the accuracy, which would require a reference/benchmark. I urge the authors to clarify this difference in the text and revise the text where needed.

R: The reviewer makes a very good point here. If TC is to provide a true assessment of accuracy, it requires a perfectly-calibrated, bias-free scaling target which, of course, is not generally available in data-scarce regions. Instead, we utilize TC to provide an assessment of linear correlation against the "true" rainfall accumulation value. We agree that referring to this as "accuracy" is not correct. However, about the term "precision" can be still misleading since the term precision is used in describing the agreement of a set of results among themselves and is usually expressed in terms of the deviation of a set of results from the arithmetic mean of the set (e.g., standard deviation).

Therefore, to resolve this issue, we will use the word "performance" instead of "accuracy" and clarify early in the paper (at the end of the introduction section, pag.3, lines 24-27) that "performance" is defined in terms of correlation against "true" rainfall values:

"Thanks to the ability of TC to provide the correlation against the "unknown" truth (ETC, McColl et al. (2014)), in this study the assessment of the products will be carried out in terms of correlation against "true" rainfall values. As a result, the word "performance" and "TC results" are here onward referred to this index (additional clarifications are provided in section 2.3). "

Based on this strategy, the title has been also changed in:

"An assessment of the performance of global rainfall estimates without ground-based observations"

Additional changes have been made in the revised manuscript at lines (of the original manuscript) where the term resulted not appropriate:

• Abstract line 14. "Results convey the relatively high accuracy of the satellite rainfall estimates in Eastern North and South America, South Africa, Southern and Eastern Asia, Eastern Australia as well as Southern Europe and complementary performances…." That has been changed in

"Results convey the relatively high performance of the satellite rainfall estimates in Eastern North and South America, South Africa, Southern and Eastern Asia, Eastern Australia as well as Southern Europe and complementary performances…"

• Pag. 10 line 5 (now pag. 10 line 20) "Assessment of CPC accuracy". The section has been titled "Assessment of the CPC product"

• Pag. 10 line 29 (now pag. 11 line 15). "… Triplet C leads to an overly-optimistic assessment of their accuracy." has been changed to "Triplet C leads to an overly-optimistic assessment of their performance."

• Pag. 10 line 32 (now pag. 11 line 18). "It is often important to understand which is the best available rainfall product provides the highest relative accuracy." is changed to "It is often important to understand which is the best rainfall product among those available in a specific location".

• Pag. 12 line 7 (now pag 12 line 29) "particular TC configuration approach to assess the accuracy" has been changed to "particular TC configuration approach to assess the performance".

• Pag. 14 line 10 (now pag. 14 line 33). "Results convey the relatively high accuracy of daily rainfall accumulations" has been changed to "Results demonstrate the relatively high performance of daily rainfall accumulations…"

2. Line 1: remove "value".

R: We have removed it.

3. Line 4: remove "of this variable"

R: We have changed the sentence in (now pag 1 line 2):

"Recent studies have suggested the use of Triple Collocation (TC) to characterize uncertainties associated with rainfall estimates by using three collocated rainfall products."

4. Line 5: replace "among" with "with"

R: We have substituted it.

5. Line 18: is it really the best product? It is if precision is the chosen criterion, but it may not be the case if accuracy is considered to be more important.

R: Since we are not talking about the accuracy here the sentence is now adequate. See also the discussion to point 1.

Page 2:

6. Line 35: replace "provides" with provide"

R: We have replaced it.

Page 3:

7. Lines 3 and 6: add "the" before "TC"

R: We have removed "method" at line 3 and added "the" at line 6.

8. Line 9: rephrase as "if each product is affected by mutually-independent errors"

R: We think the sentence is correct.

9. Line 20: deleted "analysis" and "the"

R: We have removed "analysis" term.

10. Lines 20-25: CMORPH and TMPA 3B42RT are not completely independent. Can the authors explain what is the implication with using these two products in the TC analysis?

R: Cross-correlated error in CMORPH and TMPA 3B42RT have large implications for the application of TC. In fact, these implications provide the central motivation for the paper (particularly our proposed use of a soil moisture-based rainfall accumulation product).

This is an important point for us to convey to the reader. To make sure that it is clear, we have added some clarifying details at lines 32-34 pag. 3 about the possible consequences of applying TC to a triplet containing the two products to facilitate the understanding of this important point.

"Note that, given the number of common sensors shared by CMORPH and TMPA 3B42RT the application of TC to the triplet containing the two products will serve for demonstrating the impossibility to use both of them in the same triplet within the TC analysis."

In addition, note that we have already highlighted this point in the original manuscript at lines 8-13 page 3.

"TC can theoretically provide error and correlations of three products (a triplet) without using ground based observations – provided that each of the three products is afflicted by mutually-independent errors. However, given that state-of-the-art satellite rainfall products use a highly-overlapping set of common sensors for the retrieval of rainfall (see section 2.1 for further details), there is an inherent difficulty in obtaining triplets with mutually-independent errors."

11. Line 28: rephrase as "Section 2 presents datasets and methods;"

R: The sentence has been rephrased.

12. Line 31: rephrase as "Results and discussion are shown in Section 3 and the final remarks are presented in Section 4."

R: The sentence has been rephrased.

Page 4:

13. Line 13: delete "a"

R: "a" has been deleted.

Page 7

14. Lines 2-4: this is a fair model only to model the error for the "hit cases" when both sensor and ground truth are larger than zero or for cumulative rain over a long enough period of time. Otherwise, the multiplicative error model would assign zero anytime the sensor measure a zero.

R: In this section (2.2.1), we discuss thoroughly the multiplicative error assumption along with advantages and limitations when it is used within TC. As a results, we think this is a good place for debating the issue related to presence of zero values in the rainfall time series and the potential problems they can create when the log-transformation is applied.

We have added the reference of Tian et al. (2013).

Page 8

15. Line 6: shouldn't it be "utilizes"? What's the subject of that verb?

R: ρ_{t,X_i} refer the correlations of the products, with i=1,2,3 so we used the plural.

Page 9:

16. Line 16: replace "serve to" with "are used to"

R: We have replaced it.

Page 10:

17. Line 6: please replace "assessing" with "to assess"

R: We have replaced it.

18. Line 32: please rephrase (2 verbs).

R: We have rephrased it.

Page 11:

19. Line 9: drop the comma.

R: We have dropped it.

Page 12:

20. Lines 1-2: can the authors speculate on why this happens?

R: Some explanations are given at page 11 lines 30-35. In particular, the use of daily rainfall accumulations necessitates the removal of a lot of zero values for applying the log-transformation. This might shorten the dataset too much and decrease the robustness of the TC analysis.

21. Line 17: rephrase as "this corroborates what shown by…."

R: We have rephrased as suggested.

Page 13:

22. Line 2: replace "paper" with "study".

R: We have replaced it.

Referee 2

This paper presents a novel approach to estimating surface precipitation using retrieved soil moisture. The authors then apply their soil moisture estimates to understanding the uncertainties in satellite rainfall estimates, and indicate which potential rainfall products perform better in different regions of the CONUS and globally. The applicability to the hydrologic modeling community makes it appropriate for publication in HESS. I recommend publication with minor revisions, many of which deal with adding additional clarification for the reader.

R: We would like to thank the reviewer for appreciating the value of the paper and for her/his valuable comments.

Page 3:

1. Line 5: certainly not certainty
R: We have corrected the text.

Page 4:

2. Section 2.1.2: I think that the flow of the paper would be improved by including the description of SM2RAIN with the description of ASCAT (or as a sub- section to it) as opposed to the current arrangement of describing the instrument here and the product several sections later.

R: We thank the reviewer for this suggestion. However, we prefer to separate the methodology section from the data description. SM2RAIN can, in fact, be applied to any type of satellite soil moisture observation, and we would not like to not give the false impression that it can run only with ASCAT soil moisture data.

3. Section 2.1.3: Readers familiar with the 3B42 product will recognize that you are using the "Real Time" rather than the "Research" version. In the CMORPH description you mention using the raw version that lacks gauge information, this justification should be included as to why you use 3B42RT as well.

R: We have added the following clarification to the revised paper at page 3 lines 23-24 of the Introduction section:

"(note that both 3B42RT and CMORPH (raw version) do not include gauge information in their retrieval algorithms)"

Page 5:

4. line 8: SSM/I instruments are operated by the US Department of Defense, not NOAA.
R: We have corrected the text.

"It incorporates precipitation estimates derived from the PMW on board of the DMSP 13, 14 & 15

5. Line 23: 1st

R: We have corrected the text.

6. Line 11: should the second i in the square root also be subscripted?

R: Yes, it should be subscripted because it is referring to a diagonal element of the covariance matrix.

Page 9:

7. Line 19: "use" instead of "have utilize"

R: We have corrected the text.

8. Line 24: You indicate that equation (8) is only valid for liquid precipitation, and in the concluding remarks mention that the SM and combined satellite products are less reliable in cases of frozen precipitation/snow cover/frozen surfaces. Are you using the entire 2012-2015 time period, or only the warm seasons? If you are using the entire period, how are you dealing with the winter months?

R: We thank the reviewer for rising this important point. We used the entire period 2012-2015; however, we removed periods of snow cover/frozen soil by masking data where the surface state flag (SSF) of the ASCAT product indicates frozen (SSF=2), temporary melting/water on the surface (SSF=3) or permanent ice (SSF=4). In particular, given that the analysis was carried out at a 1-degree spatial resolution, grid cells were masked if more than 50% of their sub-grid areas consisted of ASCAT observations characterized by a SSF equal to 2, 3 or 4. Moreover, data points where we observed solid precipitation from ERA-Interim were also excluded. The latter, in addition to the consideration of the SSF, helped to reduce the probability of having snowy periods and consider only liquid precipitation. Thus, the results of the paper are not affected by snow.

This has been clarified in:

1. section 2.1.2 at lines 24-26:

"Prior to the application of SM2RAIN to ASCAT data, the points characterized by a surface state flag (SSF) of the ASCAT product that indicates frozen (SSF=2), temporary melting/water on the surface (SSF=3) or permanent ice (SSF=4) were excluded from the analysis."

2. Page 6 lines 5-6:

"Note that, we considered only liquid precipitation in the analysis. Solid precipitation were excluded by masking out periods experiencing snowfall (using the "large-scale snowfall" variable of ERA-Interim)."

3. And at page 9 lines 22-24:

"Finally, 1°x1°grid cells were masked if more than 50% of their sub-grid areas consisted of ASCAT observations characterized by a SSF equal to 2, 3 or 4."

9. Line 28: Remove the word "values"

R: We have removed it.

Page 10:

10. Line 19, "are", not "ae"

R: We have corrected the text.

11. Lines 27-28, and Page 12, Line 3: This may be arguing semantics a bit, but the results don't indicate that not using SM2RAIN yields unreliable results. The results indicate that not adhering to the assumptions of the TC method (specifically with respect to having estimates with uncorrelated errors) produces unreliable results. Table 1 indicates that triplets D and E do just as well without SM2RAIN.

R: The reviewer is right. The sentence is misleading. We have corrected the text as:

Lines 13-14 page 11

"This suggests those triplets not containing SM2RAIN (or CPC) provide unreliable results."

12. Line 32: Sentence needs revising

R: We will modify the sentence to read (Line 18 page 11):

"It is often important to understand which is the best rainfall product among those available in a specific location."

Page 11:

13. Lines 15-21: It would be nice to have some context as to why the statistics for the multiplicative error are different from the additive. This comes up a bit later (line 32), but could be more up front.

R: This has been clarified via new text added at lines 5-8 page 12:

"Here, SC values are generally lower than those obtained by TC (based on an assumed additive error model) likely due to necessity of removing zero rain days which modifies the original precipitation time series and reduces the sample size of TC calculations."

14. As a general comment, it might be interesting to look at the CMORPH and 3B42 with gauge-adjustment in the global comparison. Presumably, this would improve their results in data-rich

R: While we agree that this would be an interesting extension, it would require a substantial modification of the existing paper and would entail a substantial departure from the specific goal of this analysis (i.e., to demonstrate that the availability of independent SM2RAIN-based rainfall estimates enables rainfall validation without ground-based observations). However, we fully agree that this suggestion would be a valuable topic for future research

An assessment of the ~~accuracy~~performance of global rainfall estimates without ground-based observations

Christian Massari[1], Wade Crow[2], and Luca Brocca[1]

[1]Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy
[2]United States Department of Agriculture - Hydrology and Remote Sensing Laboratory, Beltsville, Maryland, USA

Correspondence to: Christian Massari (christian.massari@irpi.cnr.it)

Abstract. Satellite-based rainfall estimates over land have great potential ~~value~~ for a wide range of applications, but their validation is challenging due to the scarcity of ground-based observations of rainfall in many areas of the planet. Recent studies have suggested the use of Triple Collocation (TC) to characterize uncertainties associated with rainfall estimates by using three collocated rainfall products ~~of this variable.~~. However, TC requires the simultaneous availability of three products with mutually-uncorrelated

5 ~~uncorrelated~~ errors, a requirement ~~that~~which is difficult to satisfy ~~among~~with current global precipitation datasets.

In this study, a recently-developed method for rainfall estimation from soil moisture observations, SM2RAIN, is demonstrated to facilitate the accurate application of TC within triplets containing two state-of-the art satellite rainfall estimates and a reanalysis product. The validity of different TC assumptions are indirectly tested via a high quality ground rainfall product over the Contiguous United States (CONUS), showing that SM2RAIN can provide a truly independent source of rainfall accu-

10 mulation information which uniquely satisfies the assumptions underlying TC. On this basis, TC is applied with SM2RAIN on a global scale in an optimal configuration to calculate, for the first time, reliable global correlations (versus an unknown truth) of the aforementioned products without using a ground benchmark dataset.

The analysis is carried out during the period 2007-2012 ~~2015~~ using daily rainfall accumulation products obtained at 1°x1° spatial

resolution. Results convey the relatively high ~~accuracy~~performance of the satellite rainfall estimates in Eastern North and South ~~America,~~Amer-

15 ica, South Africa, Southern and Eastern Asia, Eastern Australia as well as Southern Europe and complementary performances between the reanalysis product and SM2RAIN, with the first performing reasonably well in the northern hemisphere and the second providing very good performance in the southern hemisphere.

The methodology presented in this study can be used to identify the best rainfall product for hydrologic models with sparsely-gauged areas and provide the basis for an optimal integration among different rainfall products.

1 Introduction

Thanks to the combined use of microwave and infrared sensors, the quality of available satellite rainfall estimates over land has ~~signif- icantly~~significantly increased in the few last decades. This strategy – also known as multi-sensor approach – has produced a number of different satellite rainfall products that either map infrared (IR) radiances to more direct passive microwave (PMW)

1

retrievals (generally termed "blended" algorithms) or morph PMW rainfall using IR measurements (generally termed "morphing" algo-

rithms). The new Global Precipitation Measurement Mission (GPM, Hou et al., 2014) has successfully expanded the concept of multi-sensor integration. Trough the Integrated Multi-satelliE Retrievals for GPM (IMERG) algorithm, rainfall estimates from the various precipitation-relevant satellite PMW and IR missions are intercalibrated, merged and interpolated with the GPM Combined Core Instrument product to produce rainfall accumulation estimates with an unprecedented accuracy. Despite these

5   technical advancements, the precipitation community still struggles to show a clear picture of the actual increased accuracy of satellite rainfall estimates in many areas of the world because validation studies rely upon the availability of high-quality (and sufficiently dense) ground-based rainfall instrumentation (e.g. rain gage and radars).

Many studies (e.g., Ebert et al., 2007; Sapiano and Arkin, 2009; Tian et al., 2007; Stampoulis and Anagnostou, 2012) have investigated ~~the~~ error associated with remotely-sensed precipitation products by comparing their estimates with those collected

10 ~~collected~~ by ground-based observations assuming they represent the zero-error rainfall. However, the physical characteristics of ~~precipitation~~precip- itation, particularly at finer spatial and temporal resolutions, necessitate frequent, systematic and sufficiently dense validation measurements – requirements that are often not met within data-scarce regions of Africa, Asia and South America. Indeed, despite their relative accuracy, the distribution of available gauges significantly varies around the world. Much of the land ~~masses~~surface (representing 25 – 30 % of the Earth's surface) have measurement networks, although those networks with good gauge

15 ~~gauge~~ densities are limited (Kidd et al., 2017).

The current networks of surface observations are therefore often ~~not sufficient~~insufficient for the quantitative assessment of the ~~er   ror associated~~error asso- ciated with satellite rainfall estimates. Moreover, despite the relatively higher accuracy of rainfall estimates that can be obtained by rain gauges, they are not error-free. (Peterson et al., 1998; Villarini et al., 2008). Therefore, evaluating the performance of different satellite rainfall products with ground based observations is challenging due to the scarcity of such observations and

20 ~~observations and~~ of the inherent error contained in their estimates.

Based on the work of Adler et al. (2009), Tian and Peters-Lidard (2010) estimated the uncertainties of satellite rainfall estimates by using the measurement spread of coincidental and collocated estimates from an ensemble of six different satellite-based datasets, thus providing a globally consistent methodology that does not require ground-based based validation data. The analysis yielded a lower bound estimate of the uncertainties, and a consistent global view of the error characteristics and their

25 regional and seasonal variations. However, the authors showed that the analysis is able to provide only a relative estimation of the measurement uncertainties because these data sets are not entirely independent measurements.

An alternative approach for assessing the quality of satellite rainfall products was proposed by Roebeling et al. (2012) and Alemohammad et al. (2015) based on the Triple Collocation (TC) method (Stoffelen, 1998). The first applications of TC concerned geophysical variables such as ocean wind speed and wave height (Stoffelen, 1998). More recently, it has been used

30 extensively to estimate errors in soil moisture (SM) products (Crow and Van Den Berg, 2010; Miralles et al., 2010; Dorigo et al., 2010; Draper et al., 2013; Su et al., 2014; Gruber et al., 2016). Given three estimates of the same variable, the main assumptions of the method are the: i) stationarity of the statistics, ii) linearity between the three estimates (versus the same target) across all time scales and iii) existence of uncorrelated error between the three estimates.

In the work of Roebeling et al. (2012), the authors determined the spatial and temporal error characteristics of three ~~precipita~~pre-

35 ~~tion~~cipitation datasets over Europe (a visible/near infrared, a weather radar and gridded rain gauge products) showing that it can ~~provides~~

provide realistic error estimates. The authors ensured a Gaussian distribution of the error by averaging the dataset over a ~~sufficiently~~ suffi- ciently long period (10 days) and re-gridding to a sufficiently low spatial ~~resolutions~~ resolution (0.25x0.25°). Alemohammad et al. (2015) applied TC ~~method~~ to 14 days cumulated rainfall estimates derived from satellite, ~~gauge, radar~~ gauges, radars and models in order to retrieve the error and the correlation of each dataset in United States. They also proposed the use of a logarithmic (i.e., multiplicative) error

5 ~~—~~ model which almost ~~certainty~~ certainly provides a more realistic description of rainfall accumulation errors at fine ~~scale~~ space/time scales. In addition, they calculated the theoretical correlation of each product with the unknown truth by using the Extended TC (ETC) (McColl et al., 2014) by analysing the covariance matrix of the three datasets.

TC can theoretically provide error and correlations of three products (a triplet) without ~~using~~ use of ground based observations – provided that each of the three products is afflicted by mutually-independent errors. However, given that state-of-the-art satellite

10 rainfall products use a highly-overlapping set of common sensors for the retrieval of rainfall (see section 2.1 for further details), there is an inherent difficulty in obtaining triplets with mutually-independent errors. Therefore, additional - highly independent - sources of rainfall accumulation estimates are needed.

Recently, Brocca et al. (2014) developed a method for estimating rainfall accumulation amounts directly from satellite SM observations based on the principle that the soil can be treated as a "natural raingauge". In contrast with classical satellite rainfall

15 products, this new bottom-up approach attempts to measure rainfall by calculating the difference between two successive SM measurements derived from a satellite SM product. In this respect, SM2RAIN offers a unique opportunity for applying the TC analysis because, being wholly independent from any other rainfall estimate, it can be used in place of a ground-based product. This opportunity has not yet been explored and could provide an appropriate basis for applying TC on a global scale without requiring the availability of ground-based rainfall accumulation data.

20 In this study, TC ~~analysis~~ is applied to the rainfall accumulation estimates derived from: 1) ERA-Interim (Dee et al., 2011),
2) SM2RAIN (Brocca et al., 2014) via inversion of Advanced SCATterometer (ASCAT, ~~(Wagner et al.,~~ (Wagner et al., ~~1999)~~ 1999) SM data, 3) the NOAA Climate Prediction Center morphing (CMORPH, raw version) (Joyce et al., 2004) and 4) the TRMM Multi-satellite Precipitation Analysis TMPA 3B42RT (Huffman et al., 2007) product over ~~the CONtiguous United States (CONUS). An assessment of the reliability of subsequent TC results, is conducted by direct comparison with the analogous evaluation results obtained via direct~~ CONUS (note that both 3B42RT and CMORPH do not include gauge information in their retrieval algorithms). Thanks to the ability of TC to provide the correlation against the

25 "unknown" truth (ETC, McColl et al. (2014)), the assessment of the products will be carried out in terms of correlation against "true" rainfall values. As a result, the word "performance" and "TC results" will be hereinafter referred to this correlation (additional clarification is provided in section 2.3).

An assessment of the reliability of subsequent TC results is conducted by direct comparison with the analogous evaluation results obtained via direct comparisons with the Climate Prediction Center (CPC) Unified Gauge-Based Analysis of Global

30 Daily Precipitation (hereafter as CPC) product. These assessments will be carried out with and without the use of SM2RAIN rainfall accumulation products to isolate the value of SM-based rainfall estimates for the evaluation of global rainfall products. Note that, given the number of common sensors shared by CMORPH and TMPA 3B42RT the application of TC to the triplet

containing both products will serve will serve to demonstrate the difficulties with using both of them in the same triplet within the TC analysis and evaluate the potential benefits of utilizing SM2RAIN-based accumulation products in a TC analysis.

The paper is organized as follows. Section 2 contains data and ~~method descriptions~~methods; in particular, the products used for the analysis are described in section 2.1; the theoretical background for TC is in sections 2.2 and 2.2.1; the description of the

30 performance scores used for the evaluation of the results is discussed in section 2.3, and sections 2.4 and 2.5 describe ~~SM2AIN~~SM2RAIN and the experiment setup. Results are presented and discussed in Section 3 ~~contains the results and discussion,~~ and final remarks are ~~contained~~presented in ~~section~~Section 4.

Data and Methods

2.1 Rainfall and soil moisture products

2.1.1 CPC

The 0.5°x0.5° gauge-based CPC product is used to evaluate the satellite-based rainfall estimates over CONUS and verify
5 evaluations provided by TC. Given the high raingauge density associated with this product across CONUS (Figure 1),1),
along

10 with the common practice of using ground-based rainfall data to validate satellite-based rainfall retrievals (Huffman et al.,
1997), CPC is expected to provide a reasonable proxy of true rainfall accumulation over CONUS. ThisNevertheless, this
assumption will be verified below. Figure 11 illustrates that the spatial density of CPC gauge coverage (calculated as average
number of rain gauge observations per day) during 2007-2012 is high in Eastern CONUS and along the western coast of
CONUS but relatively

10 relatively lower in many parts of the Central CONUS. CPC rainfall observations are aggregated to a 1°x1° spatial
resolution by simple

15 by simple averaging.

2.1.2 ASCAT data

a ASCAT (Bartalis et al., 2007) is a real-aperture radar instrument on boardonboard the MetOp satellites thatwhich measures
radar backscatter at C-band (5.255 GHz) and VV polarization. It has a spatial resolution of 25 km (resampled at 12.5 km),)
and it is available since

15 2007. The surface SM product (equivalent to a depth of 2-–3 cm of the soil) is calculated from the backscatter measurements
20 through the time-series-based change detection approach described in Wagner et al.Wagner et al. (1999).(1999). The SM is
measured in relative terms (degree of saturation) with respect to historical minimum and maximum values. Here, we used
the ASCAT dataset produced using the Soil Water Retrieval Package (WARP) (Naeimi et al., 2009) (v5.5) from Vienna
University of Technology (TU-Wien), providedand distributed as SM product H109 by the "EUMETSAT Satellite Application
Facility on Support to Operational Hydrology and Water Management (H-SAF)". Prior to the application of SM2RAIN to
ASCAT data, the points characterized

20 Hydrology and Water Management (H-SAF)". ASCAT data are used solely for rainfall estimation with SM2RAIN. For further
details25 by a surface state flag (SSF) of the ASCAT product that indicates frozen (SSF=2), temporary melting/water on the
surface (SSF=3) or permanent ice (SSF=4) were excluded from the analysis. For further details about the application of
SM2RAIN to ASCAT, the reader is referred to section 2.4.

2.1.3 TMPA 3B42RT

TMPA 3B42RT, version 7 (http://trmm.gsfc.nasa.gov), combines rainfall estimates from various satellite sensors. The mul-

30 tisatellite platform uses the TRMM Microwave Imager (TMI) on board of TRMM satellite, the Special Sensor Microwave

25

Imager (SSM/I) on board the Defense Meteorological Satellite Program (DMSP) satellites, the Advanced Microwave Scanning Radiometer for Earth observing system (AMSRE) on board the National Aeronautic and Space Administration (NASA) AQUA satellite, the Advanced Microwave Sounding Unit-B (AMSU-B) on board the National Oceanic and Atmospheric Administration (NOAA) satellite series and GEO IR rainfall estimates . The TMPA 3B42RT estimates are produced in three steps: 1) 5 1) the PMW estimates are calibrated with sensor-specific versions of the Goddard Profiling Algorithm (GPROF; Kummerow

30 et al., 1996) and combined, 2) IR rainfall estimates are created using the PMW estimates for calibration, and 3) PMW and IR estimates are then combined. The 3B42RT product is provided by NASA with a temporal resolution of 3 h and a spatial

10

resolution of 0.25°. The cumulated daily rainfall, available from March 2000, is obtained by simply summing the eight 3-h time windows for each day. The global coverage of the product is +50°/-50° latitude. To match the CPC spatial resolution, collocated TMPA 3B42RT estimates are aggregated to 1 ° spatial resolution by simple averaging.

2.1.4 CMORPH

~~5~~ CMORPH uses a Lagrangian approach to construct high-resolution global precipitation maps from the satellite IR and PMW observations (Joyce et al., 2004). This technique uses precipitation estimates that have been derived from PMW observations exclusively, and whose features are transported via spatial propagation information which is obtained entirely from IR data. ~~It incorporates precipitation estimates derived from the PMW on board the NOAA satellites (SSM/I, the NOAA satellite series and AMSU-B) as well as AMSR-E and TMI aboard NASA's Aqua and TRMM spacecraft, respectively. Precipitation~~ It incorporates precipitation estimates derived from the PMW on board of the DMSP 13,14 & 15 (SSM/I) and NOAA-15, 16, 17 & 18 (AMSU-B) satellites as well as AMSR-E and TMI aboard NASA's Aqua and TRMM spacecraft, respectively. Precipitation estimates are obtained as follows. First, advection vectors of cloud and precipitation systems are computed using consecutive geostationary IR images in 30 min intervals. These advection vectors are then applied to propagate the precipitating cloud systems observed by the PMW measurements along the advection vectors in both forward and backward directions toward the target time of the precipitation analysis. The final precipitation analysis value at a grid box is defined as the weighted mean of the estimates from the forward and backward propagations with the weights inversely proportional to the time separation between the target analysis time and the PMW observations. In this study, we used the daily (derived from 3-hourly aggregation) estimates of precipitation at 0.25 ° latitude/longitude resolution, distributed over the globe (+60°/-60° of latitude) by the NOAA Center for Weather and Climate Prediction. Note that the CMORPH version used in this study is the raw version which does not uses gauge information. To match the CPC spatial resolution, collocated CMORPH estimates are aggregated to 1 ° spatial resolution.

2.1.5 ~~20   2.1.5~~   ERA-Interim

The European Centre for Median-range Weather Forecasts (ECMWF) produces the ERA-Interim atmospheric, ocean and land reanalysis . ERA-Interim provides medium-range global forecasts for ~~some~~ environmental variables ~~that include~~ including soil temperature, ~~evaporation~~ evap-
oration, SM and rainfall. Products are available from 1th January 1979 to now. The forecast model ~~incor   porated~~ incorporated in the ERA-
Interim reanalysis is based on the ECMWF Integrated Forecast System (Cy31r2) forecast model (Dee et al.,

~~25   et al.,~~ 2011), with a ~~spectral~~ spec- tral horizontal resolution of about 80 km and 60 vertical levels. The ERA-Interim forecast ~~precipita-   tion~~ precipitation is the sum of two ~~components~~ com-   ponents which are computed separately in the model: large-scale stratiform precipitation ~~(Tomp   kins et al., 2007) and smaller~~ (Tompkins et al., 2007) and smaller

scale precipitation which originates solely from the parameterization of convection (Bechtold et al., ~~et al.,~~ 2004). Further information can be found at the ECMWF website (http://www.ecmwf.int). In this study, daily ~~precip itation~~precipitation values are obtained from the temporal aggregation of ERA-Interim 12-hourly precipitation accumulation estimates

~~30~~ (http://apps.ecmwf.int/datasets/) while co-location with CPC observations is determined by the nearest-neighbour method. Note that, we considered only liquid precip-

2.2 Triple Collocation analysis: general concepts

Here we apply the method of (McColl et al., 2014) to robustly estimate the correlation of a particular rainfall measurement system with the truth. Suppose we have three systems X_i, measuring the true variable t and afflicted by additive random error

$$X_i = X_j^t + \varepsilon_i = \alpha_i + \beta_i t + \varepsilon_i \tag{1}$$

10 where X_i ($i=1, 2, 3$) are collocated measurement systems linearly related to the true underlying value t with additive random

5 errors ε_i, and α_i and β_i are the ordinary least squares intercepts and slopes. Assuming that the errors from each system have zero mean ($E(\varepsilon_i)=0$) and), are mutually uncorrelated ($Cov(\varepsilon_i, \varepsilon_j) =0$, with $i \neq j$) and orthogonal with respect to t ($Cov(\varepsilon_i, t)=0$), the covariance between X_i is:

$$Q_{ij} = Cov(X_i, X_j) = \begin{cases} \beta_i \beta_j \sigma_t^2, & \text{for } i \neq j \\ \beta_i^2 \sigma_t^2 + \sigma_{\varepsilon_i}^2, & \text{for } i = j \end{cases} \tag{2}$$

By defining the new variable $\theta_i = \beta_i \sigma_t$, known as the sensitivity of the variable X_i, Eq. (2) becomes:

$$Q_{ij} = \begin{cases} \theta_i \theta_j, & i \neq j \\ \theta_i^2 + \sigma_{\varepsilon_i}^2, & \text{for } i=j \end{cases} \tag{3}$$

15 which is a system of six equations in six unknowns from which we derive (McColl et al., 2014):

$$\sigma_\varepsilon = \sqrt{\begin{pmatrix} Q_{11} - \dfrac{Q_{12} Q_{13}}{Q_{23}} \\ Q_{22} - \dfrac{Q_{12} Q_{23}}{Q_{13}} \\ Q_{33} - \dfrac{Q_{13} Q_{23}}{Q_{12}} \end{pmatrix}} \tag{4}$$

From Eq. (2), using the definition of the correlation and covariance we can write:

<center>**13**</center>

$$\theta_i = \rho_{t,x_i}{}^{\mathbf{j}} \ominus_{\overline{n}}$$

(5)

where ρ_{t,X_i} is the correlation coefficient between t and X_i. Since $\overline{Q_i i}$ is already estimated from the data, and we can solve for θ_i using Eq. (4), ρ_{t,X_i} (McColl et al., 2014):

$$\rho_{t,X} = \pm \left[\sqrt{\frac{Q_{12} Q_{13}}{Q_{11} Q_{23}}} \quad \sqrt{\frac{Q_{12} Q_{23}}{Q_{13} Q_{22}}} \quad \mathrm{sign}(Q_{13} Q_{-}) \quad \mathrm{sign}(Q_{12} Q_{23}) \sqrt{\frac{Q_{13} Q_{23}}{Q_{12} Q_{33}}} \right] \tag{6}$$

which provides the temporal correlation of each product with the unknown truth. Hereinafter, when talking about ρ_{t,X_i} or its squared value ρ_{t,X_i}^2, we will refer to the correlation of the product X_i with the unknown truth. ρ will be also used to refer to this variable but in more general terms.

155

~~2.2.1   Some considerations about the application of the Triple Collocation to rainfall estimates~~

2.2.1   Rainfall error model

It is generally accepted that a multiplicative model is more appropriate for describing errors in rainfall estimates (Hossain and Anagnostou, 2006~~).~~; Tian et al., 2013). Based on this assumption, ~~(~~Alemohammad et al.~~,~~ ~~2015)~~(2015) proposed the application of TC to the rainfall by introducing a multiplicative error model:

5-10 $R_i = a_i T^{\beta_i} e^{\varepsilon_i}$ (7)

in which R is the rainfall intensity estimate from product i, T is the true rainfall intensity and a_i is ~~the~~a multiplicative error. By transforming Eq. (7) in the log-space we obtain an equation equivalent to Eq. (1) where $X = log(R)$, $t = log(T)$ and $\alpha_i = log(a_i)$. In this way, the development of TC expressed in Eqs. 2–6, can be applied to the - potentially more relevant - case of multiplicative rainfall accumulation errors. The resulting log-RMSE can then be back-transformed into linear rainfall

10-15 accumulation errors by exploiting a Taylor series expansion of the logarithm operator(see Alemohammad et al., 2015 for further details).~~–~~

The main difficulty of this approach is its inability to consider the presence of zero values in the rainfall time series. To reduce their presence, Alemohammad et al. (2015) considered biweekly rainfall estimates and simply removed remaining zeros in this time series. This has two implications. First, the biweekly rainfall error may differ from the error of a shorter accumulation

20 period (e.g. daily) because the daily signal has a substantially different character with respect to the biweekly one due to the

15– higher presence of zero values. Second, the method may not be appropriate in very dry climates where even biweekly values of rainfall can contain a significant number of zero accumulation values.

For the reasons mentioned above, we apply TC in two different ways: i) to the rainfall time series using an additive error model, and ii) to log-transformed rainfall estimates using the multiplicative error model (by first removing rainfall accumulation

25 values equal to zero). Comparisons of these two different approaches will provide insights regarding the appropriateness of

various error model assumptions for rainfall estimates at a daily accumulation time scale.

20   various error model assumptions for rainfall estimates at a daily accumulation time scale.

2.3 Performance scores

In section 2.2, it has been demonstrated that TC can provide both error variances and correlation against an unknown truth for three collocated estimates of the same variable. When dealing with error variances, the products have to be rescaled to a common reference data space. However, such a rescaling imposes spatial patterns within the derived error metric which reflects the climatology of the chosen reference (Gruber et al., 2016). To this end, McColl et al. (2014) noted that correlation coefficients can provide important new information about the performance of the measurement systems with respect to the absolute error variances obtained via Eq. (4) with the added advantage of not requiring the arbitrary definition of one system as a scaling reference. Indeed, ρ^2 represents the unbiased signal to noise ratio, scaled between 0 and 1, which provides a measure of the relative similarity between two signals, independently from their phase differences. This was also underlined by Gruber et al. (2016), who showed that ρ^2 is the complement of the $fRMSE = \sigma^2/\sigma^2$ introduced by (Draper et al., 2013) ($\rho^2 = 1 - fRMSE$) which was used previously to remove the dependency of the error variance pattern on the spatial climatology of the chosen reference. Gruber et al. (2016) also pointed out that the absolute error variance provides only limited information

17

about the true dataset quality because a certain amount of noise can be either acceptable or unacceptable depending on the strength of the underlying signal (i.e., its variance). Therefore, we focus here only on ρ^2, or analogously, on its root square ρ,

15 i.e., Eq. (6).

As discussed above, a key goal is determining the relative accuracy of TC correlations obtained with and without the use 5 of SM2RAIN-based rainfall accumulation products. Assuming that R_{X_i} (or simply R) is the Pearson correlation coefficient between the product X_i and CPC, the main question is: how accurately can (TC-based) ρ_{t,X_i}, which utilize no ground obser- vations, reproduce spatial patterns ~~of~~in (CPC-based) R_{X_i}? We should expect a bias between the two (i.e., $R_{X_i} \leq \rho_{t,X_i}$) because

20 – while relatively accurate – CPC estimates still ~~contains~~contain representativeness errors (due to limitations in raingauge density) and ~~mea surement~~measurement errors due to wind and instrument inaccuracies~~, in~~. In contrast~~,~~. Eq.(6) provides the correlations with an error-free truth.

10 Nevertheless, if the TC hypothesis holds, the relative rank between the products predicted by TC should accurately reflect that obtained via direct comparisons with ground observations.

In order to evaluate the similarity between correlation-based maps of ρ_{t,X_i} and R_{X_i} a spatial correlation index SC was 25 calculated as the spatial Pearson correlation coefficient between maps of R_{X_i} and ρ_{t,X_i}. The closer SC is to one, the more spatially similar the two maps are and the more satisfied the assumptions of TC. In addition, based on the values of ρ_{t,X_i} and R_{X_i}, we

15 are able to sort the products according to their relative performance for each pixel in the analysis. That is, considering three products X_i, the rank value RK_i to be assigned to each product i will be 1 if ρ_{t,X_i} is the highest, 3 if it is the lowest and 2 if it is ~~comprised between the minimum and the maximum value.~~neither. If the same is done with R_{X_i}, the consistency of the resulting rank maps for each product provide feedback ~~maps for each product will provide feedback~~30 regarding the validity of assumptions underlying the application of TC. For the

quantification of the discrete maps, we also calculate the number of pixel providing ~~equal~~equivalent relative sorting of the products ~~obtained~~

20 based on R_{X_i} ~~and~~versus ρ_{t,X_i}.

2.4 SM2RAIN and its application to ASCAT data

SM2RAIN (Brocca et al., 2014) is a method of rainfall estimation which uses two successive SM retrievals ~~(i.e., at the current and at a past time step)~~ to estimate the rainfall accumulated ~~within the time interval~~ between the two retrievals. It exploits the soil water balance equation with appropriate simplifications valid only for liquid precipitation ~~(Tian et al.,~~(Tian et al., 2014):

$$Z^* ds(\tau)/d\tau = p(\tau) - r(\tau) - e(\tau) - g(\tau) \qquad (8)$$

where Z^* is the soil water capacity (soil depth times soil porosity); $s(\tau)$ is the relative saturation of the soil or relative SM; τ is the time and $p(\tau)$, $r(\tau)$, $e(\tau)$ and $g(\tau)$ are the rainfall, surface runoff, evapotranspiration and drainage rates, respectively. Under unsaturated soil conditions, and assuming negligible evapotranspiration rate during rainfall and Dunnian runoff, solving Eq. (8) for rainfall yields:

$$p(\tau) = Z^* ds(\tau)/d\tau + as(\tau)^b \qquad (9)$$

Note that in Eq. (9) the drainage rate has been expressed with a power law function of the type $g = as^b$ (Famiglietti and Wood, 1994), where a and b two model parameters. When the soil is fully saturated, no rainfall can be estimated from SM; however, at the scale of satellite pixel, the soil is rarely saturated (except in some exceptional places like tropical forests).

The SM2RAIN parameters a, b and Z^* can be estimated either by using a rainfall dataset as a reference or assigned based on soil properties. In this study, in order to maximize the independence of SM2RAIN predictions, SM2RAIN parameters were not calibrated and were instead assumed constant in space as in Koster et al. (2016). In particular, the drainage rate (the second

15 term in Eq. (9)) was assumed linearly related with SM ($b = 1$) and $a = 3.7$ mm/day and $Z^* = 62$ mm based on results obtained 5 in previous studies (Brocca et al., 2014). Note that, Z^* does not have a significant influence on the results because we are using a correlation based metric. In addition, it should be noted that, while maximizing the independence of SM2RAIN rainfall accumulation estimates, the use of this default calibration approach results in sub-optimal SM2RAIN performance. Superior SM2RAIN can easily be obtainable via calibration against existing satellite rainfall accumulation products.

20 Daily rainfall estimates from SM2RAIN were obtained by using linearly-interpolated (at at 00:00 UTC) ASCAT data with

10 a maximum allowable data gap of 5 days. The obtained 0.25 °x0.25 ° rainfall estimates were then aggregated to the 1 °x1° spatial resolution by through simple averaging of the collocated pixels with CPC. Finally, 1 °x1° grid cells were masked if more than 50% of their sub-grid areas consisted of ASCAT observations characterized by a SSF equal to 2, 3 or 4. Hereinafter, the thus obtained product is referred to as SM2RAIN for simplicity.

thus obtained product is referred to as SM2RAIN for simplicity.

2.5 25 2.5 Experimental setup

A TC analysis was carried out using five different daily rainfall accumulation triplets: 1) ERA-Interim-SM2RAIN-3B42RT 15 (Triplet A in the following) 2) ERA-Interim-SM2RAIN-CMORPH (Triplet B) 3) ERA-Interim-3B42RT-CMORPH (Triplet C), 4) ERA-Interim-3B42RT-CPC (Triplet D) and 5) ERA-Interim-CMORPH-CPC (Triplet E). Triplets A and B serve are used to assess the ability of SM2RAIN to provide meaningful TC results. Triplet C provides an alternative to triplets A and B which

contains two rainfall satellite products (with potentially cross-correlated errors). Triplets D and E serve only to evaluate the general performance of the CPC product (within CONUS) and to provide alternative triplets to A and B which do have utilize

20 use SM2RAIN. As a result, they will only be used for initial considerations about TC robustness and to evaluate the relative quality of the CPC product. Triplets A, B and C will be then used in the reminder of the paper to demonstrate the potential utility of SM2RAIN.

5 The analysis was carried out first across CONUS and then on a global scale using only ERA-Interim, 3B42RT, CMORPH and SM2RAIN during the period 2007-2012 2015. Over CONUS it was confirmed that the available sample size to apply TC was sufficient (about

25 within500) over the entire study domain (Gruber et al., 2016) while for the global analysis, grid cells with inadequate sample size were individually masked out of the analysis. The Extended TC analysis was applied for both additive and multiplicative error model assumptions. For the latter, we first removed zeros (they aredays with zero rainfall constituting about 80% of the rainfalldaily values that means aboutand leaving

10 approximately 450 non-zero daily values of non null rainfall) values from the signalin the 2007-2012 time series and then applied a log-transformation to the remaining daily rainfall estimates. This reduction in sample size may affect TC results by making the analysis with log-precipitation estimates statistically less robust.

30 **3 Results and discussion**

In this section, we present the results obtained from the application of TC (for both additive and multiplicative error models)
15 by following the subsequent methodological steps: 1) calculating TC based correlations (ρ_{t,X_i}) for Triplets A, B, C, D and

E over CONUS and providing an assessment of the CPC product (section 3.1); 2) understanding the adequacy of TC results based on the spatial similarity between (TC-based) ρ_{t,X_i} and (CPC-based) R_{X_i} (along with their relative rank) over CONUS in order to identify the optimal configuration for applying TC, and, 3) applying the optimal-configured TC on a global scale to calculate ρ_{t,X_i} globally for the selected rainfall products (section 3.3).

520 3.1 Assessment of the CPC ~~accuracy~~product

As described above, our first goal is ~~assessing~~to assess the relative performance of the CPC product. Table ~~1~~1 shows mean ρ_{t,X_i} (obtained via the spatial average of 0.25-degree CONUS grid-cells). Regardless of the triplet or error model applied, the TC analysis summarized in Table ~~1~~1 indicates that CPC is the most accurate product (mean TC-based correlation close to 0.9 for the additive

error model and close to 0.8 for the multiplicative one) which strengthens our assumption that within CONUS, CPC can be 1025 used as a benchmark to evaluate the optimal TC configuration for rainfall product evaluation. In addition, its correlation spatial pattern (not shown) provides very good performance almost everywhere except in the Central US where the spatial density of available rain gauges shown in Figure ~~1~~1 is relatively lower. Based on this, in the next section we will consider the CPC product as an appropriate benchmark for the selection of an optimal TC configuration which does not utilize a gauge-based precipitation product (and is therefore potentially applicable at a global scale).

15 3.2

3.2 Optimal Triple Collocation configuration

Figures 2̶2 (a, b, c and d) plot CPC-based Pearson correlation coefficients (i.e., R_{X_i}) for ERA-Interim, 3B42RT, CMORPH and SM2RAIN obtained with the assumption of additive error model (for multiplicative error model results the reader is referred to Figure A̶1̶A1 of the supplementary material). A comparison of these results with TC-based correlations (i.e., ρ_{t,X_i}) shows that

5   ρ_{t,X_i} a̶e̶are biased high with respect to R_{X_i}. This is expected given that CPC is not free of errors whereas TC should theoretically
20   p̶r̶o̶v̶i̶d̶e̶s̶provide the correlation with respect to t̶h̶e̶an error-free truth.

The spatial agreement between ρ_{t,X_i} and R_{X_i} is examined in Table 2̶2 and Figure 2̶.̶2. In particular, Figure 2̶2 shows that Triplets A (e, f, g in the figure) and B (h, i, l) accurately reproduce CPC-based results plotted in Figure 2̶2 (a, b, c and d), although they are characterized by higher values as underlined above (see section 2.3 for further details). This similarity is higher in Eastern

10   and Western US and lower in the Central US especially for ERA-Interim and SM2RAIN. This lower agreement in the Central

25   US is likely due to the lower rain gauge density of CPC here (see Figure 1̶)̶1) which degrades the quality of the CPC product as benchmark. However, in contrast, TC-results based on Triplet C predicts a substantial different behaviour with correlation patterns which differ substantially relative to CPC-based benchmark results in Figure 2̶2 (a, b, c and d). This suggests those triplets not containing SM2RAIN (or CPC) provide unreliable results. In particular, the simultaneous use of two satellite-based r̶a̶i̶n̶f̶a̶l̶l̶ ̶p̶r̶o̶d̶u̶c̶t̶s̶ ̶i̶n̶ ̶T̶r̶i̶p̶l̶e̶t̶ ̶C̶ ̶l̶e̶a̶d̶s̶ ̶t̶o̶ ̶a̶n̶ ̶o̶v̶e̶r̶l̶y̶ ̶o̶p̶t̶i̶m̶i̶s̶t̶i̶c̶ ̶a̶s̶s̶e̶s̶s̶m̶e̶n̶t̶ ̶o̶f̶ ̶t̶h̶e̶i̶r̶ ̶a̶c̶c̶u̶r̶a̶c̶y̶.̶ ̶T̶h̶i̶s̶ ̶i̶s̶ ̶l̶i̶k̶e̶l̶y̶ ̶d̶u̶e̶ ̶t̶o̶ ̶c̶r̶o̶s̶s̶-̶c̶o̶r̶r̶e̶l̶a̶t̶e̶d̶ ̶e̶r̶r̶o̶r̶s̶

3̶0̶15   rainfall products in Triplet C leads to an overly-optimistic assessment of their performance. This is likely due to cross-correlated errors in 3B42RT and CMORPH rainfall accumulation products which causes TC to misinterpret their mutually-consistency as an indication of high-accuracy (̶Y̶i̶l̶m̶a̶z̶ ̶a̶n̶d̶ ̶C̶r̶o̶w̶,̶(Yilmaz and Crow, 2014).

It is often important to understand which is the best a̶v̶a̶i̶l̶a̶b̶l̶e̶ rainfall product p̶r̶o̶v̶i̶d̶e̶s̶ ̶t̶h̶e̶ ̶h̶i̶g̶h̶e̶s̶t̶ ̶r̶e̶l̶a̶t̶i̶v̶e̶ ̶a̶c̶c̶u̶r̶a̶c̶y̶among those available in a specific location. As described in section 2.3, we ranked the products based upon how well they compare relative to each other using both R and ρ, respectively.

~~ρ, respectively.~~20 Figure ~~3~~3 shows the distribution – three products at time (d, e, f; k, l, m; r, s, t) – of the relative rank based on comparisons with the (CPC-based) R_{X_i} of each triplet, while subplots (a, b, c; g, h, i; n, o, p) of the same figure, provide similar information except that the relative rank is based on TC (i.e., ρ). The latter shows a very similar pattern with respect to CPC-based rank

~~to CPC based rank~~ for Triplets A and B; however, Triplet C yields again a distinct pattern with ERA-Interim being the worst
~~5~~ product and 3B42RT and CMORPH providing complementary performances. As in the comparisons discussed in Figure ~~2,~~2, this implies that triplets

25 containing SM2RAIN (i.e., Triplets A and B) provide more robust evaluation information than triplets utilizing ~~both~~ 3B42RT and CMORPH together.

The same analysis carried out with the assumption of multiplicative error model (see Figure ~~A2~~A2 in the supplementary material~~),~~) shows similar findings but larger differences between the spatial distribution of the rank obtained with CPC and the

~~40~~ one with TC, especially for Triplet B. To quantity this agreement, we have calculated the percentage of pixels which are ranked

30 the same in both TC- band CPC-results. (% of rank identified in Table ~~2).~~2). The table confirms the patterns observed in Figure ~~3~~3 and Figure ~~A2~~A2 of the supplementary material with Triplets A and B yielding the highest percentage of pixels with a common

rank – ranging from 65 to 81 % for the additive error model, and 48 to 71 % for the multiplicative error model. As discussed above, ~~worst~~inferior results are obtained in both cases for Triplet C (percentage of correct ranking between 5% and 60%).
~~15~~

A quantification of the agreement between the spatial variations of the correlations both for additive and multiplicative error models was also derived by the use of the spatial correlation SC in Table 2.2. The table shows that for Triplets A and B, when TC is used with the assumption of additive error model, SC is relatively high with values ranging from 0.61 to 0.84 while for Triplet C provides substantially lower SC for 3B42RT and CMORPH. A slightly different situation can be observed for the ~~multiplicative error model. Here, SC values are generally lower than those obtained by TC with the assumption of additive~~

20 ~~error model.~~5 multiplicative error model. Here, SC values are generally lower than those obtained by TC (based on an assumed additive error model) likely due to necessity of removing zero rain days which modifies the original precipitation time series and reduces the sample size of TC calculations. In particular, ERA-Interim provides the worst score. This is not clearly evident in the spatial distribution of R and

ρ (see Figure A1A1 in the supplementary material for further details) which show some similarities at least for Triplets A and B.

~~In summary, the application of TC to the different triplets shows that:~~

10   In summary, the application of TC to the different triplets shows that:

1. CPC product performs relatively well over CONUS with a TC-derived correlation versus truth of 0.9 (assuming an additive error model) demonstrating its relatively high quality here and supporting its application as a benchmark data set within CONUS.

25   ~~set within CONUS.~~

2. TC-based correlations are similar among the triplets except for Triplet C (i.e, ERA-Interim, 3B42RT and CMORPH).

~~2.~~   15   This is likely due to existence of non-negligible cross-correlated errors between 3B42RT and CMORPH.

~~3.~~ A comparison between ρ_{t,x_i} and R_{x_i} shows that ρ_{t,x_i} are biased ~~with higher values~~high with respect to R_{x_i}. In addition, the pattern of ρ_{t,x_i} and R_{x_i} is similar for all triplets except for Triplet C, which shows inconsistencies relative to the CPC

3. ~~30~~ benchmark for both the additive and multiplicative error model assumptions. The agreement, measured in term of spatial correlation

(Table ~~2),~~2), provides higher scores for ~~the~~an additive error model ~~assumptions with respect~~assumption relative to ~~the~~a multiplicative one. This is likely due

20   to a reduction of sampling power associated with the removal of daily rainfall accumulations equal to zero which are not acceptable in the log-transformation process. Therefore, it is possible that the observed differences in TC performance may shrink for larger sample sizes.

4. Retrieved spatial patterns of ρ_{t,X_i} for the triplets containing SM2RAIN (Figure 2)2) showsshow a higher degree of similarity with (CPC-based) R_{X_i} when we assume an additive (versus multiplicative) error model for daily rainfall accumulations.

25   On this basis, we can conclude that: i) TC results are unreliable unless SM2RAIN is used in the triplets and ii) the assumption of multiplicative error model in the application of TC at a daily time scale does not appear necessary.

3.3   5   3.3   Application of optimized TC approach

Based on the superior performance for Triplets A and B under the assumption of additive error model, we will apply this particular TC configuration approach to assess the accuracyperformance (in terms of ρ) of daily rainfall accumulation estimates derived

30   from 3B42RT, CMORPH, SM2RAIN and ERA-Interim first over CONUS (section 3.3.1 and Figure 2)2) and then on a global scale (section 3.3.2 and Figure 4).

10   3.3.1

3.3.1 CONUS

Over CONUS, ERA-Interim shows relatively better performance in Western and Eastern United States (US) with respect to the Central US where SM2RAIN is slightly superior. 3B42RT and CMORPH perform reasonably well in Eastern and along the ~~coast~~West Coast of ~~Western~~the US while ~~show~~demonstrating worse performance in the Central US ~~both in the south and in the north.~~. On the contrary, SM2RAIN performs

5 worse in Northern US probably due to the lower accuracy of the ASCAT data at high latitudes. The spatial

~~15~~ pattern of these correlations is similar to ~~the one~~those found in Gottschalck et al. (2005) and Ebert et al. (2007) who showed a general lower level of correlation of satellite-only rainfall products in the Central US due to ~~effect~~the effects of snow cover and frozen surface conditions. ~~They are also similar to the~~ This corroborates results ~~obtained by~~presented in Alemohammad et al. (2015) using TC, who found a similar pattern of correlation of 3B42RT in a box covering a large part of southeastern US (however, the authors here assumed a multiplicative error model and ~~biweekly rainfall accumulation estimates).~~

~~20~~ ~~3.3.2~~ ~~10~~ biweekly rainfall accumulation estimates).

3.3.2 Global

On a global scale, 3B42RT (Figure ~~4a)~~4a) shows relatively good performances in Eastern and Central South America, Southern and Central Africa, Southern and Eastern Asia, Eastern Australia, and Southern Europe while it performs relatively worse in Central Asia, Western Australia and in the southern part of the Sahel. The performance of CMORPH (Figure ~~4b)~~4b) is similar

15 to 3B42RT with slightly lower correlations in Australia, in the Horn of Africa and in Southern Asia. SM2RAIN (Figure ~~4c)~~

~~25~~4c) performs reasonably well in Africa (except in the tropical forest), Australia, Mexico, Eastern South America and India and generally in the southern hemisphere while worse results are obtained in the northern hemisphere, in the tropical forests and at high latitudes. On the contrary, ERA-Interim (Figure ~~4d)~~4d) provides much better results in the northern hemisphere with respect to the south of the planet (e.g., South America and South Africa) and performs relatively poor in Central and Northern Africa ~~as well as in the tropical forests.~~

~~30~~~~20~~ as well as in the tropical forests.

The results for 3B42RT and SM2RAIN are similar to those obtained in Brocca et al. (2014) who calculated the Pearson ~~cor relation~~correlation coefficient with the Global Precipitation Climatology Center (GPCC, (Schamm et al., 2014)) dataset. ~~Similar findings~~Similar find- ings are also presented in Yong et al. (2015) (Table 2 of their study) who compared different versions of the 3B42RT product against global CPC observations in the US, East Asia, Europe and Australia. In their study, the best results were obtained in

25 Australia and in East Asia (Europe showed slightly lower performance) while lower performances were obtained in US like in our analysis. Further comparisons can be also considered with the recent work of Beck et al. (2017) who, in attempting to create a high-

5 quality rainfall product specifically tailored for hydrological modelling, compared different satellite and modelledmod- elled products globally with the Global Historical Climatology Network-Daily (GHCN-D, Menne et al., 2012) database. Their results (in terms of spatial pattern of correlation) are consistent with the onesthose obtained in our study over the US, East Asia and

30 Middle East for CMORPH and 3B42RT, while less agreement is observed in Australia. For ERA-Interim, the results agree with our study in US, Europe and generally are better in the northern hemisphere, whereas they show some differences with SM2RAIN

10 results in Australia, Africa and in South America, although in these areas a low number of available rain gauges cannot provide a clear picture of the real performance of the analysed products. Substantial differences between our study and

the studies of ~~-~~Beck et al. (2017) and ~~Yong et al.~~Yong et al. ~~(2015)~~(2015) can likely be attributed to the quality of the benchmark dataset used for the ~~evalua   tions~~evaluations. This is the main limitation of rainfall validation studies relying upon ground based observations for assessment. With ~~the~~our proposed TC-based approach, this issue can be overcome because ground observations are no longer required.

~~15~~ An interesting feature of the global evaluation of the products (Figures ~~4a~~4a and ~~4b,~~4b, ~~4c~~4c and ~~4d),~~4d), but also over CONUS between

5   3B42RT (or CMORPH) and SM2RAIN (Figure ~~2~~2 triplets A and B), is the complementary nature of the ~~performance among the~~ products. Especially for Figures ~~4c~~4c and ~~4d,~~4d, it can be seen that ERA-Interim performs very well in the northern hemisphere and worse in the southern hemisphere whereas SM2RAIN is relatively good in the south and worse in the northern hemisphere. Similar findings can be seen between the two state-of-the art satellite rainfall products (i.e., 3B42RT and CMORPH) and

20   SM2RAIN over CONUS with the first performing better in Eastern US and the second in the Central and Western US. This opens up new possibilities

10   for the integration of ~~the~~multiple products ~~for obtaining~~to obtain a higher-quality merged rainfall estimate ~~–~~ as ~~also underlined~~outlined in Ciabatta et al. (2015) and in Beck et al. (2017).

4   Summary and conclusions

The assessment of the performance of satellite rainfall products on a global scale is challenging due to significant limitations
25   in the spatial coverage of high-quality, ground-based rain gauge observations. Provided that its underlying assumption are

15   respected (see section 2.2), TC provides an alternative approach for evaluating global rainfall products without reliance on ground-based observations. Here, we describe how a new method for rainfall estimation based on SM observations (i.e., SM2RAIN) provides a rainfall product that is uniquely suited to satisfy the error independent assumptions at the heart of the TC approach.

~~30~~ The extended version of TC introduced by McColl et al. (2014) ~~is able~~was applied to provide the correlation with the (unknown

20   error-free) truth for each of the products applied within a particular triplet. To assess the robustness of correlated-based results obtained with TC, we used an area characterized by a high quality rainfall product (CPC dataset over CONUS, see Figure 1)

with the assumption that it represents a good proxy of the true rainfall field. Therefore, if TC assumptions hold, Pearson correlation coefficients

1) computed against CPC should match those of TC – at least in terms of their relative values. Since we have two different error model options (i.e., additive and multiplicative) for the application of TC to rainfall data, we explored both.

25   both.

Results demonstrate that daily rainfall accumulations provided by the CPC product are indeed relatively high quality compared to competing products (Table 1),1), thus supporting the assumption that it provides an acceptable proxy of the true rainfall

5 field. Once it is established as a credible benchmark, CPC is used to evaluate 1) what type of triplets can be considered for a robust application of TC?, and 2) which model error assumption can be considered more appropriate? Triplets containing

30   SM2RAIN and assuming an additive error model (Table 2)2) appear to provide the most robust TC results. Based on this, an optimal TC configuration was applied (for the first time) to globally evaluate daily rainfall accumulation derived from the 3B42RT and CMORPH, ERA-Interim and SM2RAIN products (Figures 4a4a and 4b,4b, 4c4c and 4d)4d) without the use of any ground-based data. Results demonstrate the relatively high performance of daily rainfall accumulations derived from the satellite rainfall products

10 data. Results convey the relatively high accuracy of daily rainfall accumulations derived from the satellite rainfall products

(i.e., 3B42RT and CMORPH) in Eastern North and South America, South Africa, Southern and Eastern Asia, Eastern Australia as well as Southern Europe and complementary performances between ERA-Interim and SM2RAIN, with the first performing reasonably well in the northern hemisphere and the second providing very good performance in the southern hemisphere.

Based on the results obtained, we can therefore conclude that:

155   1. Despite the abundance of satellite rainfall estimates, their relative dependency impedes their use within the same triplet for the TC analysis, thus alternative independent products must be used for obtaining meaningful TC results. In particular, the use of two remotely-sensed rainfall products in a single triplet entails significant risk of a biased TC analysis.

   2. Wholly independent daily rainfall accumulation products obtained from SM2RAIN are uniquely valuable for obtaining robust global evaluation statistics in absence of ground-based gauge observations. This is not only important for simple validation purposes but also for hydrological studies and applications within developing countries where ground-based rain gauge networks are often limited or even totally absent and an alternative product has to be chosen.

2010

   3. At the time/space scales examined here, the assumption of additive error model provides reasonable and robust results and no advantage is observed for a log-transformation of the time series (which allows for the consideration of a multi-plicative error model). However, this result is likely to be scale-dependent and implies at the time/scale resolution of this analysis is sufficiently coarse such that averaging produces approximate additive/gaussian distributions (via the central limit theorem). Therefore, different results may be obtained at finer time/scales.

2515

   4. Both state-of-the art satellite rainfall estimates (i.e., 3B42RT and CMORPH) and SM-based rainfall estimates (i.e., SM2RAIN) performances are affected by the presence of snow cover and frozen soil conditions – thus these rainfall estimates may be unreliable at high latitudes and in mountainous regions. In these areas, a reanalysis product (i.e., ERA30 Interim) provides higher-quality rainfall estimates and should be considered in place of satellite-based estimates. SM-based rainfall estimates also work reasonably well in semi-arid climates (e.g., Sahel, Central Australia and Mexico) where the state-of-the-art satellite products report problems due to sub-cloud evaporation of hydrometeors (Ebert et al., 2007). Conversely, in wet climates (e.g, tropical forests) 3B42RT and CMORPH seem to be the only reliable option given that neither SM2RAIN nor ERA-Interim provide reasonable results.

3020

5.   25 5. Given the existence of complementary performances among the products, TC can potentially be a valuable tool for the characterization of their relative performances so as to be used for data fusion and assimilation experiments for obtaining more accurate rainfall estimates.

The question whether this analysis is valid for different spatio-temporal scales remain to be addressed and will be addressed 5 in future studies. Also, removing zeros for obtaining log-transformed rainfall may not be ideal for testing the validity of the 30 model error assumptions becausesince it shortens the sample size, thus providing less robust TC results. Other strategies should be considered.

References

Adler, R. F., Wang, J.-j., Gu, G., and Huffman, G. J.: A Ten-Year Tropical Rainfall Climatology Based on a Composite of TRMM Products, Journal of the Meteorological Society of Japan. Ser. II, 87, 281–293, doi:http://doi.org/10.2151/jmsj.87A.281, 2009.

Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., and Stoffelen, A.: Characterization of precipitation product errors across the United States using multiplicative triple collocation, Hydrology and Earth System Sciences, 19, 3489–3503, doi:10.5194/hess- 19-3489-2015, http://www.hydrol-earth-syst-sci.net/19/3489/2015/, 2015.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., and Anderson, C.: Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), Geophysical Research Letters, 34, L20 401, doi:10.1029/2007GL031088, http://doi.wiley.com/10.1029/2007GL031088, 2007.

Bechtold, P., Chaboureau, J.-P., Beljaars, A., Betts, A., Köhler, M., Miller, M., and Redelsperger, J.-L.: The simulation of the diurnal cy- cle of convective precipitation over land in a global model, Quarterly Journal of the Royal Meteorological Society, 130, 3119–3137, doi:10.1256/qj.03.103, http://doi.wiley.com/10.1256/qj.03.103, 2004.

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, Hydrology and Earth System Sciences Discussions, 21, 589–615, doi:10.5194/hess-2016-236, http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-236/, 2017.

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 rain gauge: Estimating global rainfall from satellite soil moisture data, Journal of Geophysical Research: Atmospheres, 119, 5128–5141, doi:10.1002/2014JD021489, http://dx.doi.org/10.1002/2014JD021489http://onlinelibrary.wiley.com/doi/10.1002/2014JD021489/full, 2014.

Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Puca, S., Rinollo, A., Gabellani, S., and Wagner, W.: Integration of Satellite Soil Moisture and Rainfall Observations over the Italian Territory, Journal of Hydrometeorology, 16, 1341–1355, doi:10.1175/JHM-D-14-0108.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-14-0108.1, 2015.

Crow, W. T. and Van Den Berg, M. J.: An improved approach for estimating observation and model error parameters in soil moisture data assimilation, Water Resources Research, 46, 1–12, doi:10.1029/2010WR009402, 2010.

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., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

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, Hydrology and Earth System Sciences, 14, 2605–2616, doi:10.5194/hess-14-2605-2010, http://www.hydrol-earth-syst-sci.net/14/2605/2010/, 2010.

Draper, C., Reichle, R., de Jeu, R., Naeimi, V., Parinussa, R., and Wagner, W.: Estimating root mean square errors in remotely sensed soil moisture over continental scale domains, Remote Sensing of Environment, 137, 288–298, doi:10.1016/j.rse.2013.06.013, http://dx.doi.org/10.1016/j.rse.2013.06.013, 2013.

Ebert, E. E., Janowiak, J. E., and Kidd, C.: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models, Bulletin of the American Meteorological Society, 88, 47–64, doi:10.1175/BAMS-88-1-47, http://journals.ametsoc.org/doi/abs/10.1175/BAMS-88-1-47, 2007.

Famiglietti, J. S. and Wood, E. F.: Multiscale modeling of spatially variable water and energy balance processes, Water Resources Research, 30, 3061–3078, doi:10.1029/94WR01498, http://doi.wiley.com/10.1029/94WR01498, 1994.

Gottschalck, J., Meng, J., Rodell, M., and Houser, P.: Analysis of Multiple Precipitation Products and Preliminary Assessment of Their Impact on Global Land Data Assimilation System Land Surface States, Journal of Hydrometeorology, 6, 573–598, doi:10.1175/JHM437.1, 2005.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W.: Recent advances in (soil moisture) triple collocation analysis, International Journal of Applied Earth Observation and Geoinformation, 45, 200–211, doi:10.1016/j.jag.2015.09.002, http://linkinghub.elsevier.com/retrieve/pii/S0303243415300258, 2016.

Hossain, F. and Anagnostou, E. N.: A two-dimensional satellite rainfall error model, IEEE Transactions on Geoscience and Remote Sensing, 44, 1511–1522, doi:10.1109/TGRS.2005.863866, 2006.

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 mission, Bulletin of the American Meteorological Society, 95, 701–722, doi:10.1175/BAMS-D-13-00164.1, 2014.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schnei derSchneider, U.: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset, Bulletin of the American MeteorologicalMeteorological Society, 78, 5–20, doi:10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2, http://journals.ametsoc.org/doi/abs/10.1175/ 1997. 1520-0477{%}281997{%}29078{%}3C0005{%}3ATGPCPG{%}3E2.0.CO{%}3B2, 1997.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, Journal of Hydrometeorology, 8, 38–55, doi:10.1175/JHM560.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM560.1, 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, Journal of Hydrometeorology, 5, 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D. B., Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, How Much of the Earth's Surface <i>Is</i> Covered by Rain Gauges?, Bulletin of the American Meteorological Society, 98, 69–78, doi:10.1175/BAMS-D-14-00283.1, http://journals.ametsoc. org/doi/10.1175/BAMS-D-14-00283.1, 2017.

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, Water Resources Research, 52, 7213–7225, doi:10.1002/2016WR019024, http://doi.wiley.com/10.1002/2016WR019024, 2016.

Kummerow, C., Oison, W. S., and Giglio, L.: A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors, IEEE Transactions on Geoscience and Remote Sensing, 34, 1213–1232, doi:10.1109/36.536538, 1996.

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, Geophysical Research Letters, 41, 6229–6236, doi:10.1002/2014GL061322, http://dx.doi.org/10.1002/2014GL061322, 2014GL061322, 2014.

35

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An Overview of the Global Historical Climatology Network-Daily Database, Journal of Atmospheric and Oceanic Technology, 29, 897–910, doi:10.1175/JTECH-D-11-00103.1, http://dx.doi.org/10.1175/JTECH-D-11-00103.1, 2012.

Miralles, D. G., Crow, W. T., and Cosh, M. H.: Estimating Spatial Sampling Errors in Coarse-Scale Soil Moisture Estimates Derived from Point-Scale Observations, Journal of Hydrometeorology, 11, 1423–1429, doi:10.1175/2010JHM1285.1, http://dx.doi.org/10.1175/2010JHM1285.1, 2010.

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W.: An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations, IEEE Transactions on Geoscience and Remote Sensing, 47, 1999–2013, doi:10.1109/TGRS.2008.2011617, 2009.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., ~~Vincent~~Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: a review, International Journal of ~~Climatology~~Climatology, 18, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.~~0.CO;2-T, http://dx.doi.org/10.1002/(SICI)-1097-0088(19981115)18:13{%}3C1493::AID-JOC329{%}3E3~~.0.CO;2-T, 1998.

~~Roebeling, R. a., Wolters, E. L. a., Meirink, J. F., and Leijnse, H.: Triple Collocation of Summer Precipitation Retrievals from SEVIRI over~~

Europe with Gridded Rain Gauge and Weather Radar Data, Journal of Hydrometeorology, 13, 1552–1566, doi:10.1175/JHM-D-11-089.1, 2012.

Sapiano, M. R. P. and Arkin, P. A.: An Intercomparison and Validation of High-Resolution Satellite Precipitation Estimates with 3-Hourly Gauge Data, Journal of Hydrometeorology, 10, 149–166, doi:10.1175/2008JHM1052.1, http://journals.ametsoc.org/doi/abs/10.1175/2008JHM1052.1, 2009.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schröder, M., and Stender, P.: Global gridded precipitation over land: A description of the new GPCC First Guess Daily product, Earth System Science Data, 6, 49–60, doi:10.5194/essd-6-49-2014, 2014.

Stampoulis, D. and Anagnostou, E. N.: Evaluation of Global Satellite Rainfall Products over Continental Europe, Journal of Hydrometeorology, 13, 588–603, doi:10.1175/JHM-D-11-086.1, http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-11-086.1, 2012.

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

Su, C.-h., Ryu, D., Crow, W. T., and Western, A. W.: Beyond triple collocation: Applications to soil moisture monitoring, Journal of Geophysical Research: Atmospheres, 119, 6419–6439, doi:doi:10.1002/2013JD021043, 2014.

~~Tian, Y. and Peters-Lidard, C. D.: A global map of uncertainties in satellite-based precipitation measurements, Geophysical Research Letters,~~ 37, L24 407, doi:10.1029/2010GL046008, http://doi.wiley.com/10.1029/2010GL046008, 2010.

Tian, Y., Peters-Lidard, C. D., Choudhury, B. J., and Garcia, M.: Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications, Journal of Hydrometeorology, 8, 1165–1183, doi:10.1175/2007JHM859.1, 2007.

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., and Wu, H.: Modeling errors in daily precipitation measurements: Additive or multiplicative?, Geophysical Research Letters, 40, 2060–2065, doi:10.1002/grl.50320, 2013.

Tian, Y., Liu, Y., Arsenault, K. R., and Behrangi, A.: A new approach to satellite-based estimation of precipitation over snow cover, International Journal of Remote Sensing, 35, 4940–4951, doi:10.1080/01431161.2014.930208, http://www.tandfonline.com/doi/abs/10.1080/01431161.2014.930208, 2014.

Tompkins, A. M., Gierens, K., and Rädel, G.: Ice supersaturation in the ECMWF integrated forecast system, Quarterly Journal of the Royal Meteorological Society, 133, 53–63, doi:10.1002/qj.14, http://doi.wiley.com/10.1002/qj.14, 2007.

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.

5   Wagner, W., Lemoine, G., and Rott, H.: A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, Remote Sensing of Environment, 4257, 1999.

5   Yilmaz, M. T. and Crow, W. T.: Evaluation of Assumptions in Soil Moisture Triple Collocation Analysis, Journal of Hydrometeorology, 15, 1293–1302, doi:10.1175/JHM-D-13-0158.1, http://dx.doi.org/10.1175/JHM-D-13-0158.1, 2014.

Yong, B., Liu, D., Gourley, J. J., Tian, Y., Huffman, G. J., Ren, L., and Hong, Y.: Global view of real-time TRMM multisatellite precipitation–
10   analysis: Implications for its successor global precipitation measurement mission, Bulletin of the American Meteorological Society, 96, 283–296, doi:10.1175/BAMS-D-14-00017.1, 2015.
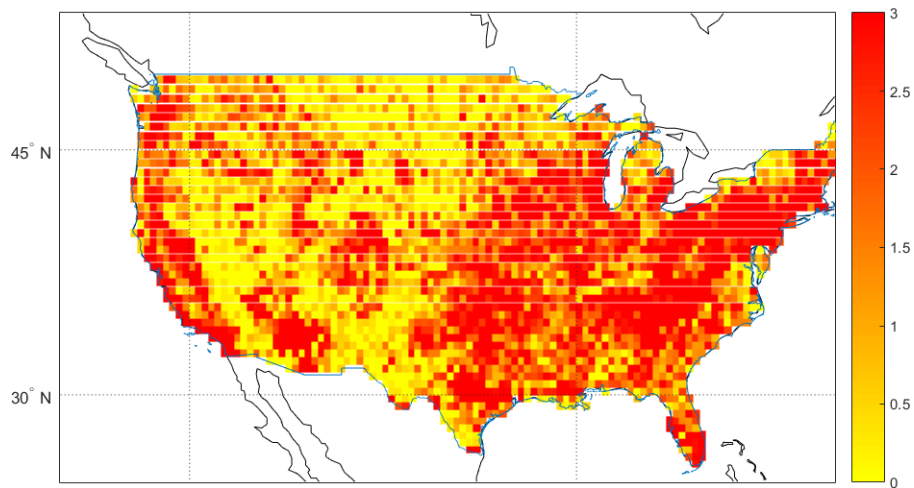
Figure 1. CPC gauge coverage during 2007-2012 expressed as average number of working raingauge per day within each 0.25-degree spatial grid cell.
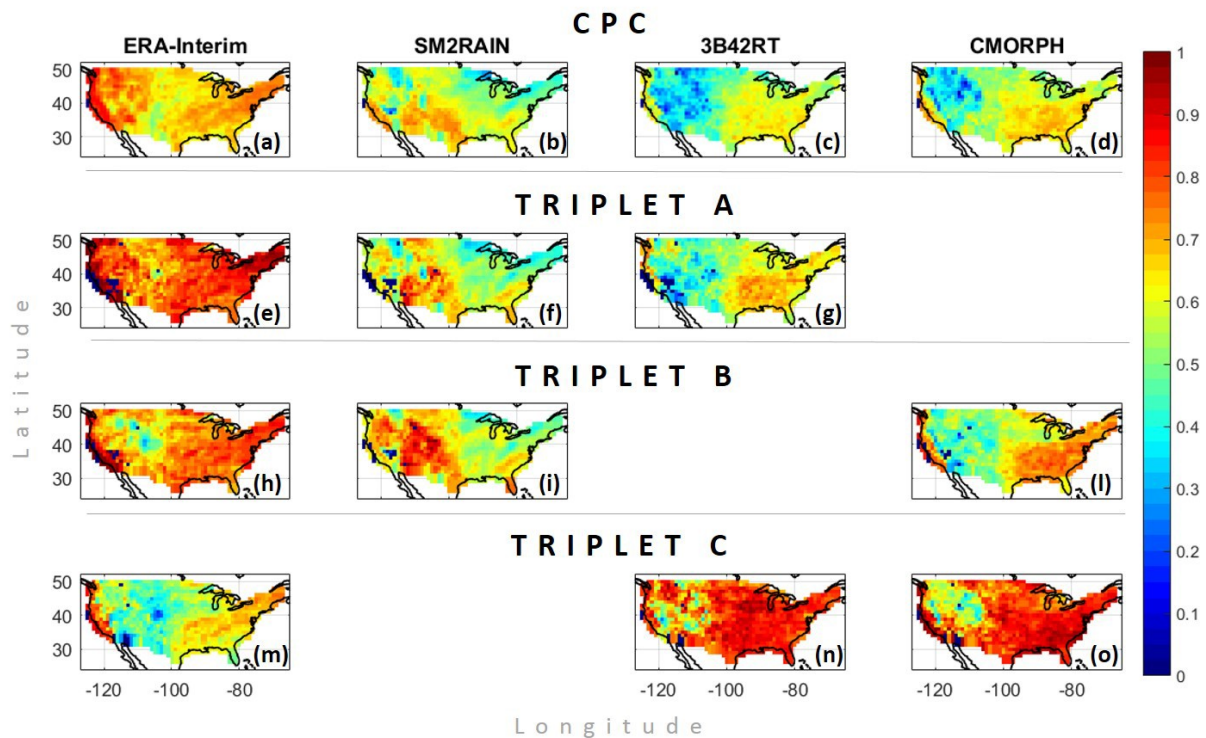

Figure 2. CPC-based (a, b ,c and d) and TC-based (e-o) correlation coefficient obtained for the triplets: i) ERA-Interim-SM2RAIN-3B42RT (Triplet A: e, f, g), ii) ERA-Interim-SM2RAIN-CMORPH (Triplet B: h, i, l) and iii) ERA-Interim-3B42RT-CMORPH (Triplet C: m, n, o) during the period 2007-2012 2015 using an additive error model.
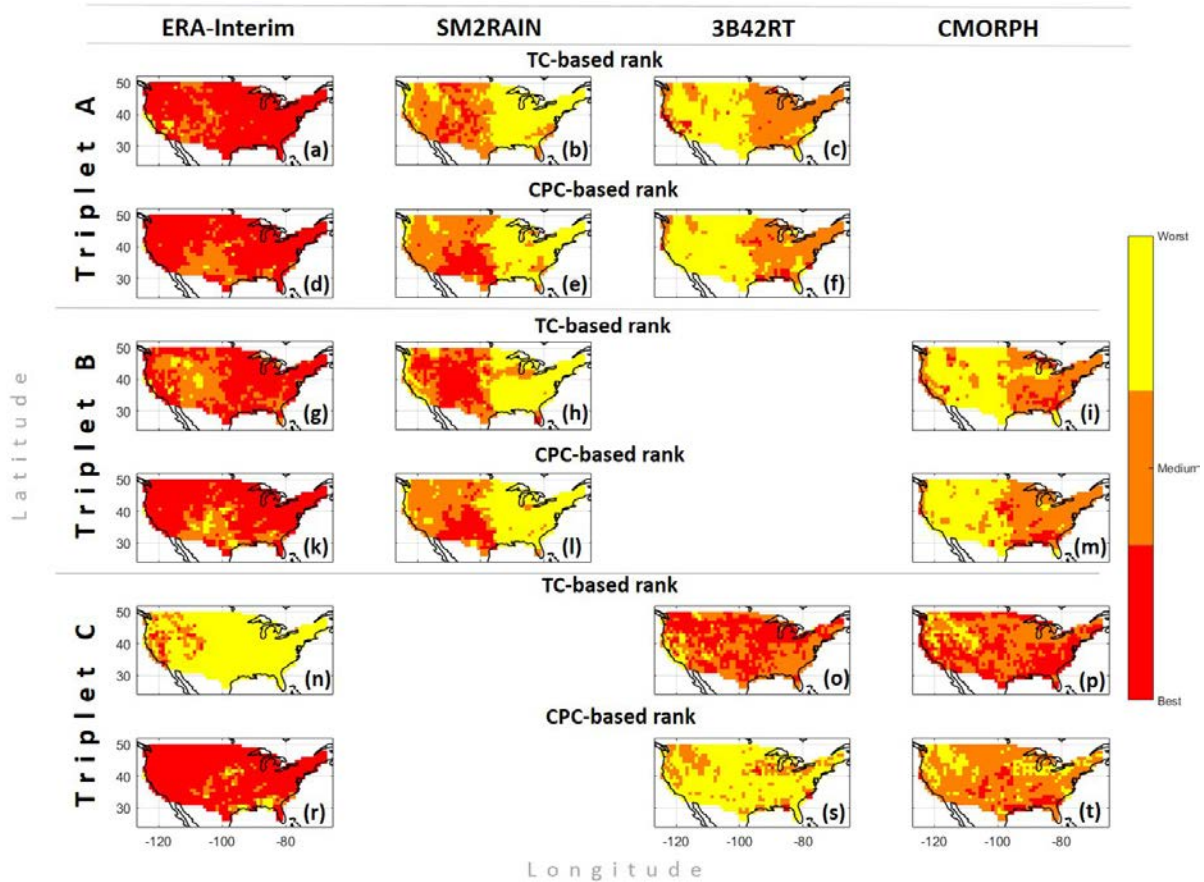
Figure 3. Rank based on CPC-based correlation (CPC-based rank in the figure) and TC-based correlation (TC-based rank in the figure) of the triplets: i) ERA-Interim-SM2RAIN-3B42RT (Triplet A: a, b, c for TC-based rank and d, e, f, g), for CPC-based rank) , ii) ERA-Interim-_SM2RAIN-CMORPH (Triplet B: g, h, i for TC-based rank and k, l, m for CPC-based rank) and iii) ERA-Interim-3B42RT-–CMORPH (m,Triplet_C: n, o, p for TC-based rank and r, s, t for CPC-based rank) during the period 2007-2012 2015 using an additive error model.

(a) 3B42RT

(b) CMORPH
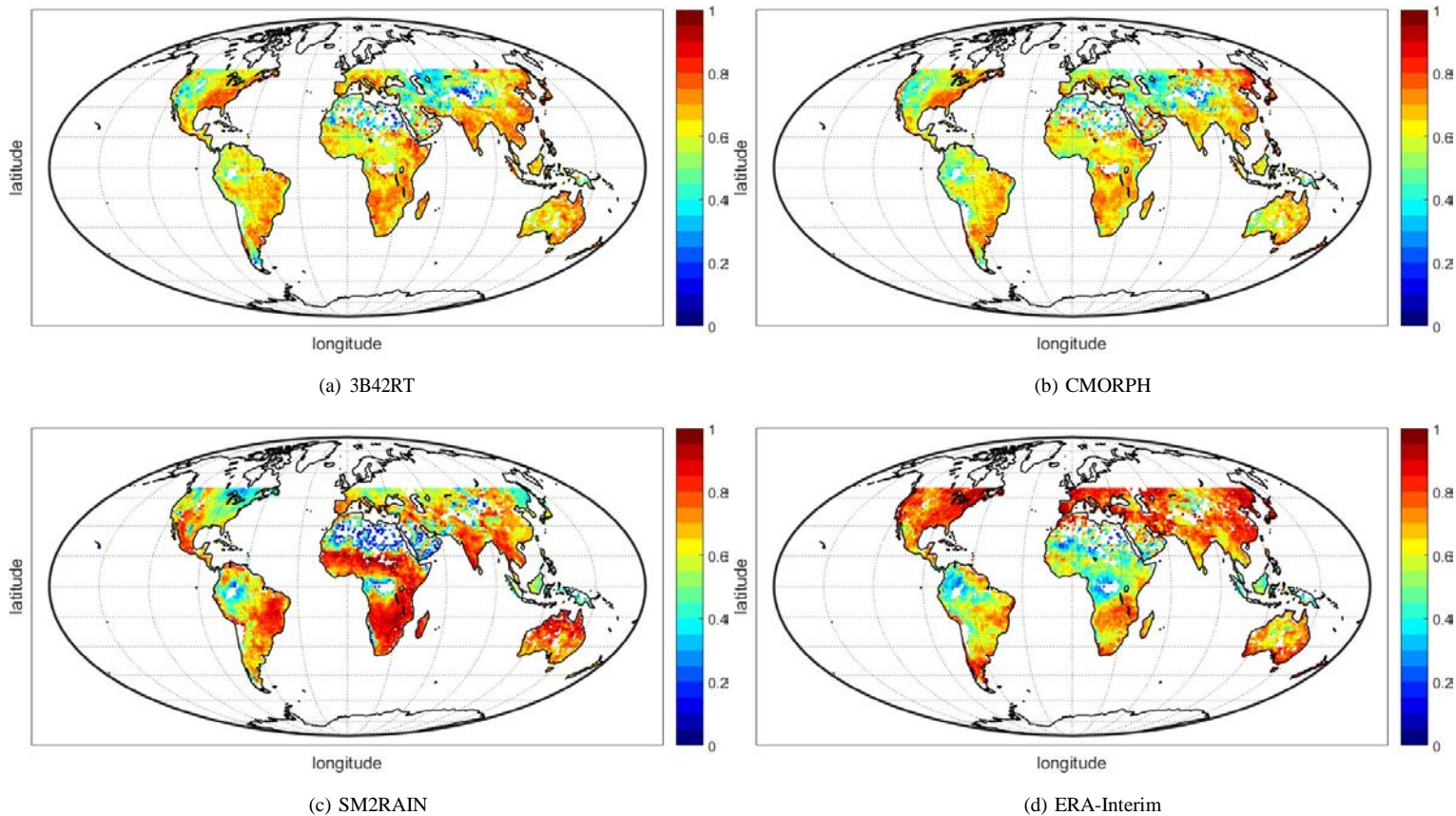
(c) SM2RAIN

(d) ERA-Interim

Figure 4. Global correlation of the 3B42RT (a), CMORPH (b), SM2RAIN (c) and ERA-Interim (d) products obtained by Triple Collocation using Triplet A (ERA-Interim-SM2RAIN-3B42RT) for 3B42RT, ERA-Interim and SM2RAIN and Triplet B (ERA-Interim-SM2RAIN-CMORPH) for CMORPH.

		ADDITIVE ERROR				
		Era-Interim	SM2RAIN	3B42RT	CMORPH	CPC
	Correlation with CPC	0.68	0.57	0.52	0.57	-
Triplet	**Products**	**Triple Collocation**				
A	**ERA - SM2RAIN - 3B42RT**	0.79	0.57	0.57	-	-
B	**ERA - SM2RAIN - CMORPH**	0.73	0.63	-	0.58	-
C	**ERA - 3B42RT - CMORPH**	0.43	-	0.68	0.76	-
D	**ERA - 3B42RT - CPC**	0.79	-	0.57	-	0.87
E	**ERA - CMORPH - CPC**	0.76	-	-	0.60	0.91
		MULTIPLICATIVE ERROR				
		Era-Interim	SM2RAIN	3B42RT	CMORPH	CPC
	Correlation with CPC	0.53	0.43	0.38	0.50	-
Triplet	**Products**	**Triple Collocation**				
A	**ERA - SM2RAIN - 3B42RT**	0.63	0.53	0.43	-	-
B	**ERA - SM2RAIN - CMORPH**	0.68	0.55	-	0.62	-
C	**ERA - 3B42RT - CMORPH**	0.43	-	0.68	0.76	-
D	**ERA - 3B42RT - CPC**	0.65	-	0.42	-	0.84
E	**ERA - CMORPH - CPC**	0.66	-		0.57	0.79

Table 1. Mean ~~CPC-based correaltion, R (~~correlation with CPC ~~in the table~~(R) and TC-based ~~correlations~~correlation (ρ) ~~between different~~for various triplets ~~for the~~assuming additive and multiplicative error models.

~~additive and for the multiplicative error models.~~ The "Triplet" column refers to the naming convention applied in the text.

Triplet	Products	Era-Interim	3B42RT	CMORPH	SM2RAIN
		SPATIAL CORRELATION			
		Additive error model			
A	**ERA - 3B42RT - SM2RAIN**	0.79	0.74	-	0.84
B	**ERA - CMORPH - SM2RAIN**	0.86	-	0.61	0.84
C	**ERA - 3B42RT - CMORPH**	0.96	0.28	0.07	-
		Multiplicative error model			
A	**ERA - 3B42RT - SM2RAIN**	0.380	0.751	-	0.648
B	**ERA - CMORPH - SM2RAIN**	0.265	-	0.798	0.570
C	**ERA - 3B42RT - CMORPH**	0.508	0.508	0.706	-
		% RANK IDENTIFIED			
		Additive error model			
A	**ERA - 3B42RT - SM2RAIN**	80%	81%	-	72%
B	**ERA - CMORPH - SM2RAIN**	65%	-	74%	65%
C	**ERA - 3B42RT - CMORPH**	6%	10%	41%	-
		Multiplicative error model			
A	**ERA - 3B42RT - SM2RAIN**	65%	71%	-	60%
B	**ERA - CMORPH - SM2RAIN**	48%	-	51%	67%
C	**ERA - 3B42RT - CMORPH**	11%	15%	50%	-

Table 2. Spatial correlation SC between ρ_{t,X_i} and R_{X_i} and percentage of rank correctly identified obtained ~~with the different~~for various triplets considered in the study. The "Triplet" column refers to the naming convention applied in the text.