

## Performance of precipitation products obtained from combinations of satellite and surface observations

José Roberto Rozante , Enver Ramirez Gutierrez , Alex de Almeida Fernandes & Daniel A. Vila

To cite this article: José Roberto Rozante , Enver Ramirez Gutierrez , Alex de Almeida Fernandes & Daniel A. Vila (2020) Performance of precipitation products obtained from combinations of satellite and surface observations, International Journal of Remote Sensing, 41:19, 7585-7604, DOI: [10.1080/01431161.2020.1763504](https://doi.org/10.1080/01431161.2020.1763504)

To link to this article: <https://doi.org/10.1080/01431161.2020.1763504>



Published online: 16 Jul 2020.



Submit your article to this journal [↗](#)



Article views: 9



View related articles [↗](#)



View Crossmark data [↗](#)



## Performance of precipitation products obtained from combinations of satellite and surface observations

José Roberto Rozante , Enver Ramirez Gutierrez, Alex de Almeida Fernandes and Daniel A. Vila



Center for Weather Forecast and Climate Studies, National Institute for Space Research, Cachoeira Paulista, Brazil

### ABSTRACT

Knowing the spatiotemporal distribution of precipitation is undoubtedly important for planning various economic/social activities, such as agriculture, livestock, and energy production. The coarse observation density over certain regions may significantly compromise the quality of precipitation products interpolated by only surface observations. To minimize the lack of observations over certain regions, the Centre for Weather Forecast and Climate Studies (CPTEC) of National Institute for Space Research (INPE) developed two types of blended precipitation products, namely, the Combined Scheme (CoSch) and MERGE, which combine observed precipitation data with satellite estimates on a daily scale. To understand how different blending methodologies impact the final results, a comparison of each algorithm with independent rain gauges was performed with a focus over the Brazilian territory. Both products were generated at a 10-km horizontal resolution using input data from the Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG-Early) for product (Version 5) in conjunction with surface observations from Surface Synoptic Observations (SYNOP), data collection platforms (DCPs) and data from regional meteorological centres. The cumulative 24-hour precipitation was evaluated for the period from June 2014 to June 2017. The results show that both products reliably characterize the precipitation regimes over most of the study regions, although MERGE and CoSch tend to over- and underestimate the amount of precipitation, respectively. However, the magnitude of the Bias achieved by MERGE is smaller than that achieved by CoSch. Overall, MERGE outperforms CoSch when analysing rain/no rain and light to moderate rainfall (0.5 to 20.0 mm). For heavy precipitation (>35.0 mm), the performance of both products is similar. The most significant differences between the two products occur over the Northeast Region of Brazil (R3 and R4), where CoSch tends to encounter difficulties characterizing the precipitation regime during the northeastern wet period (April – November). In R3 and R4, MERGE relies more on surface observations, whereas CoSch relies on GPM-IMERG-Early, which could be associated with the deficiency of GPM-IMERG-Early in estimating the amount of precipitation associated with warm clouds.

### ARTICLE HISTORY

Received 13 August 2019  
Accepted 2 April 2020

**CONTACT** José Roberto Rozante  [roberto.rozante@inpe.br](mailto:roberto.rozante@inpe.br)  Center for Weather Forecast and Climate Studies, National Institute for Space Research, Rodovia Presidente Dutra, Km 40, SP-RJ, Cachoeira Paulista, SP CEP: 12630-000, Brasil

## 1. Introduction

Understanding the spatiotemporal distribution of precipitation is important not only for scientific studies but also for making decisions related to the social and economic aspects of a country. Over South America, many regions (e.g. those composed predominantly of rain forest, mountains, desert, or ocean) suffer from limited access to precipitation data, and hence, the characterization of the ground-based rainfall regime in these regions is compromised. Although remote sensing techniques are useful for filling gaps associated with such limited access, large uncertainties are still encountered. To minimize such uncertainties, techniques that combine satellite estimates with rain gauge data are widely used at the global scale (Huffman et al. 1997; Xie and Arkin 1997) as well as the regional scale, for example, in China (Chen et al. 2002), Ethiopia (Dinku et al. 2014), South America (Vila et al. 2009; Rozante et al. 2010), Australia (Chappell et al. 2013), and India (Mitra et al. 2013).

Satellites have long been valuable tools for measuring atmospheric parameters at regular intervals. The first generation of satellite observation systems was initiated during the 1960 s with the launch of the first weather satellite, namely, the first of the Television Infrared Observation Satellites series (TIROS-1), which detected cloud systems. Ten years after TIROS, the first precipitation estimates using the visible channel were made available. The method known as the cloud index was later improved by Follansbee (1973), who configured this index to estimate the daily mean precipitation rate. With numerous technological advances, various improvements to satellite-derived products have been implemented. By the end of the 1990 s, several onboard satellite sensors were in use to successfully estimate global precipitation. Currently, a variety of satellite precipitation products are available for operations and research, including the following: a) Climate Prediction Centre Morphing technique (CMORPH) (Joyce et al. 2004); b) Tropical Rainfall Measuring Mission (TRMM) – Multi-satellite Precipitation Analysis (TMPA) (Huffman et al. 2007); c) Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) (Hong et al. 2004); d) Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) (Huffman et al. 2015); and e) Global Satellite Mapping of Precipitation (GSMaP) (Aonashi et al. 2009).

Other products combine rain gauge-based estimates with estimates from TMPA at the regional level (Vila et al. 2009; Rozante et al. 2010; Woldemeskel, Sivakumar, and Sharma 2013; Chappell et al. 2013; Mitra et al. 2013); however, as TRMM-TMPA is scheduled to end soon, these products will be discontinued. With this consideration, the Global Precipitation Measurement (GPM hereafter) satellite was launched at the beginning of 2014 with the objective of producing a new generation of precipitation and snow estimates at high spatial ( $0.1^\circ \times 0.1^\circ$ ) and temporal (30 minutes) resolutions. The GPM mission is considered the natural continuation of the TRMM mission, and the precipitation product corresponding to TRMM-TMPA is IMERG. Considered the next generation of satellite-derived precipitation products, IMERG includes previous products, such as TMPA, CMORPH and PERSIANN. To date, several studies have shown that IMERG outperforms TMPA retrievals (Prakash et al. 2016; Tang et al. 2016; Rozante et al. 2018).

The Centre for Weather Forecast and Climate Studies (CPTEC) of the National Institute for Space Research (INPE) has developed and continued to maintain two different blended precipitation products, namely, the Combined Scheme (CoSch – Vila et al. (2009)) and MERGE (Rozante et al. (2010)). Both of these products combine observed

precipitation data with satellite estimates on a daily scale. Although a variety of similar products are available on the global scale (e.g., IMERG-Final, TMPA-V7, and GSMaP-Gauge), in general, MERGE and CoSch perform better over South America (Rozante et al. 2010; Brahm et al. 2019), which is probably related to the larger databases accessed by CPTec, some of which are not available through the Global Telecommunication System (GTS). In addition, both CPTec products are available with latencies on the order of hours after the data are received (i.e. in near real time). Although the IMERG-Early product is also available in near real time, this product does not incorporate pluviometric data, whereas IMERG-Final has a latency of months.

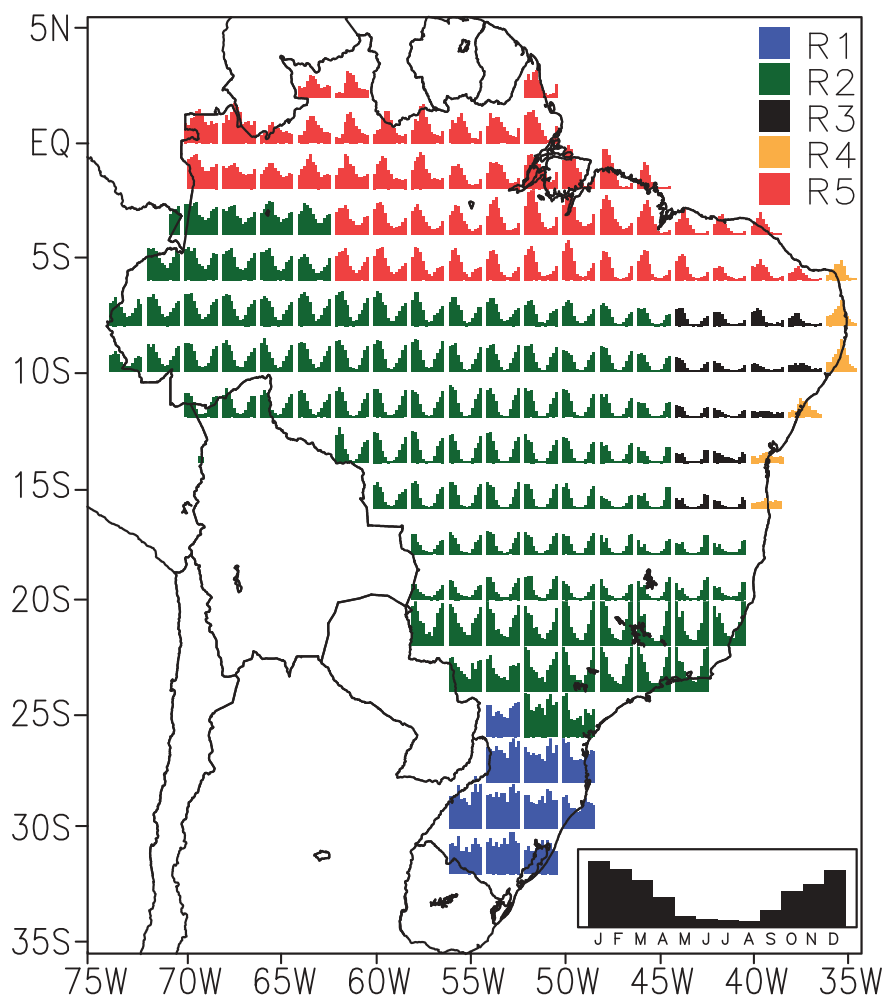
MERGE and CoSch impact a relatively diverse base of users who are not necessarily confined to an academic environment. For example, a series of reports were published by the Brazilian Geological Service, superintendency of Minas Gerais, on the use of MERGE to monitor the sediment plume after the rupture of the Brumadinho dam in January 2019. MERGE has also been used to monitor the drought in northeastern Brazil, as this product can be found via public disclosure by the World Bank (Martins et al. 2015). On the other hand, a recent paper by Brahm et al. (2019) based on CoSch and a weather-related disaster database showed that the former can satisfactorily capture insurance-relevant losses on the ground. Thus, both MERGE and CoSch are consolidated products that fulfil a niche not addressed by other available products. Nevertheless, despite the success of both MERGE and CoSch, the present paper depicts, for the first time, a more in-depth comparison of both CPTec blended precipitation products. In addition, regarding the technical aspects of the present research and the complexity of the system, CPTec needs to lead this type of comparison. Furthermore, although both products are generated through the same input data, differences have been discovered in the final results due to differences in the methodologies applied to merge satellite-based information with rain gauge estimates. In essence, CoSch prioritizes an unbiased approach to obtain the final product (Vila et al. 2009), while MERGE proposes a blending approach that favours observational data over regions featuring a high density of observations (Rozante et al. 2010). Initially, MERGE and CoSch were developed using the TMPA product as an input; the modifications applied to the algorithms will be described in the methodology section.

The present study is designed to provide a more in-depth comparison of these two CPTec products and to understand the differences and evaluate the performance of the abovementioned products over different regions of Brazil. The remainder of this paper is organized as follows: [Section 2](#) addresses the data used and the applied statistics as well as the area of study and the regionalization according to distinguished precipitation regimes. In [Sections 3](#) and [4](#), the results and discussions are presented, and in [Section 5](#), the main conclusions and summary of this work are highlighted.

## **2. Data and methodology**

### **2.1. Area of study and characterization of the precipitation regimes**

Brazil is characterized by its large area and broad diversity of climates, topographies, and precipitation regimes. In this sense, to better study the behaviours of CoSch and MERGE, it is necessary to distinguish among the various regions and pluviometric regimes



**Figure 1.** Spatial distribution of the precipitation climatology (1998–2016) based on MERGE data (Rozante et al. 2018); the climatology is computed for each grid box (approximately 2°). The indicated regions (R1–R5) characterize different precipitation regimes.

throughout the country. The present study uses the criteria defined by Rozante et al. (2018), who applied data spanning 18 years to define five different pluviometric regions (R1–R5, see Figure 1). The southernmost region R1 (blue) presents well-distributed precipitation year-round. R2 (green) covers a large area of Brazil characterized by a large amount of precipitation during austral summer (December January and February (DJF)) and little precipitation during winter (June July and August (JJA)). R3 (black), located in northeastern Brazil, does not present a well-defined pluviometric regime, which is probably because it is situated within a transition region. R4 (orange) is located on the coast of northeastern Brazil and presents precipitation during JJA with minimum precipitation during DJF. Finally, R5 (red) covers the northern part of the study domain and presents more intense precipitation during January February and March (JFM).

## **2.2. Data sources**

The data used for this study are divided into observational and estimated data, and they are described below.

### **2.2.1. Observational data**

The pluviometer data employed herein are from the Global Telecommunication System (GTS), the Brazilian National Meteorological Service (INMET), the National Water Agency (ANA), the Paraná Meteorological System (SIMEPAR), the Minas Gerais Energy Company (CEMIG), and the Agronomy Institute of Campinas (IAC), in addition to other regional centres. On average, considering the whole studied domain, approximately 3300 pluviometers were used to generate the products for each day (Figure 3). The process used to construct 24-hour accumulations of data follows the recommendation of the World Meteorological Organization (WMO), i.e. the accumulated precipitation between 12 Greenwich Mean Time (GMT) of the previous day and 12 GMT of the current day (WMO 2009).

The quality control phase has two stages: a) an objective treatment that identifies and marks potentially spurious data and b) a subjective analysis that determines whether to accept or reject the marked data. Further details about the procedure used by CPTEC/INPE can be found in Rozante et al. (2018).

### **2.2.2. Satellite precipitation estimates**

The IMERG Core Observatory Satellite (Huffman et al. 2015) launched on 27 February 2014 is used for the precipitation estimates. The estimation algorithm was built to calibrate, combine and interpolate satellite-derived precipitation data (e.g., microwave and infrared) and worldwide observational data. IMERG is available in near real time for operational purposes and with a two-month delay for research purposes. IMERG provides two near-real-time precipitation estimate options: early and late. The IMERG-Early product provides a quick estimate (with a 4-hour lag) considering only data that have been collected at that moment, while the IMERG-Late product has a 12-hour lag (after more data have arrived) and is obviously more precise. The IMERG-Early product used in this study was obtained from <ftp://arthurhou.pps.eosdis.nasa.gov/gpmallversions/V05/> with a temporal resolution of 30 minutes and a spatial resolution of  $0.1^\circ \times 0.1^\circ$ . IMERG-Early covers most of the globe: all surface areas between  $60^\circ\text{N}$  and  $60^\circ\text{S}$ , corresponding to 87% of the Earth's surface. According to the WMO guidelines, IMERG-Early data are also accumulated over 24 hours.

### **2.2.3. MERGE**

MERGE (Rozante et al. 2010) is a technique intended to minimize uncertainties in precipitation data associated with interpolations over regions with a low density of pluviometers. The observational data are combined with GPM-IMERG-Early satellite estimates (Huffman et al. 2015). MERGE was described in Rozante et al. (2010) using the TMPA product as the satellite data. With TRMM-TMPA being discontinued and replaced by GPM-IMERG, the only necessary adaptation to the algorithm was to remove a relatively large number of points close to each observation station to preserve the action radius of that station. The MERGE product relies on the following steps. First, the pluviometric data are georeferenced onto the satellite grid. Then, for sites near pluviometric stations, a  $5 \times 5$

square grid box of satellite data is removed. Finally, all the pluviometric data are interpolated, and any remaining satellite estimates are performed.

#### **2.2.4. CoSch**

CoSch (Vila et al. 2009) facilitates a high-resolution analysis based on a combination of pluviometers with satellite precipitation estimates over continental South America. The CoSch methodology is based on additive and multiplicative corrections to obtain a smaller Bias of the product with respect to pluviometric data. This product was also developed during the TRMM-TMPA era; however, no adaptation was needed due to changes in the satellite database. The GPM-IMERG-Early product (Huffman et al. 2015) is used as a first estimate, and daily allowed rain gauge data are used to correct the first estimate.

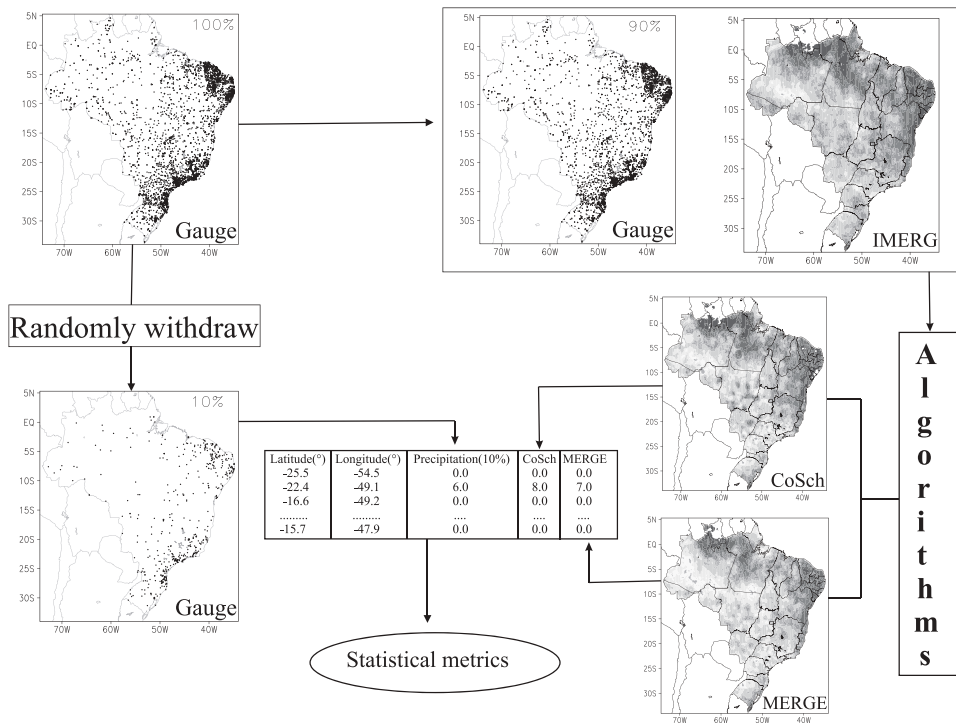
### **2.3. CoSch – MERGE evaluation methodology**

Both products are studied within a 3-year period (26 June 2014 to 25 June 2017), and although CoSch and MERGE are both generated for South America at a resolution of 10 km, only regions over the territory of Brazil are studied because of the broad availability of data over this country. Cross-validation tests (Chen et al. 2002, 2008) are then applied to quantify the performance of the algorithms.

At each time step (one day), the rain gauge data are randomly split into two parts, with 90% used for the CoSch and MERGE algorithms (using the IMERG-Early product as input data) and 10% reserved for validation. Figure 2 illustrates the data splitting scheme. Due to the different precipitation regimes described in section 2.1, all statistics are calculated for each individual region (R1 through R5). It is worth noting that randomly selecting rain gauges from the whole region every day might lead to variability in the number of stations used to adjust and validate the products for each region. Indeed, recent work by Prakash et al. (2019) assessed the uncertainties of gridded gauge-based rainfall products and suggested that variability in the rain gauge density impacts the mean values and thresholds for precipitation. However, due to the large sample of days (spanning a 3-year period), the mean value of the removed gauges is very close to 10% for every region (see Figure 3); therefore, the significance and robustness of the statistics are not compromised.

### **2.4. Evaluation metrics**

For evaluation purposes, seven commonly used metrics are computed, namely, three statistical index and four categorical classifications. The three statistical metrics include the root mean square error (RMSE), the mean error (ME) and the Pearson correlation coefficient ( $r$ ) (see Table 1), whereas the categorical index used in this study include the equitable threat score (ETS) (Mesinger 2008), probability of detection (POD), relative Bias (Bias), and false alarm ratio (FAR). All of the values of these categorical indexes are obtained from a contingency table (Table 2), and their equations and optimal values are summarized in Table 3. In this research, daily precipitation data are grouped into eight categories with different rain intensities (0.5, 2.0, 5.0, 10.0, 15.0, 20.0, 35.0 and 50.0 mm day<sup>-1</sup>). To facilitate further discussion, these categories are grouped into rain/no rain, light, moderate and heavy rain classes, as shown in Table 4.



**Figure 2.** Illustrative scheme of the processes to evaluate the CoSch and MERGE products.

The POD, FAR, and Bias results are analysed through Roebber's performance diagram (Roebber 2009), which synthesizes several metrics within a geometric plot, where the abscissa represents the success measured between 0 and 1 (0 representing a completely false alarm and 1 representing a complete success) and the ordinate represents the POD, which is also measured between 0 and 1 (1 corresponding to the maximum POD). Strait lines represent the relative Bias; a 45° angle corresponds to a perfect Bias, while larger (shorter) angles represent overestimation (underestimation). In addition, there are hyperbolic curves representing the critical success index (CSI); the curve closest to the right upper corner of the diagram represents the better CSI.

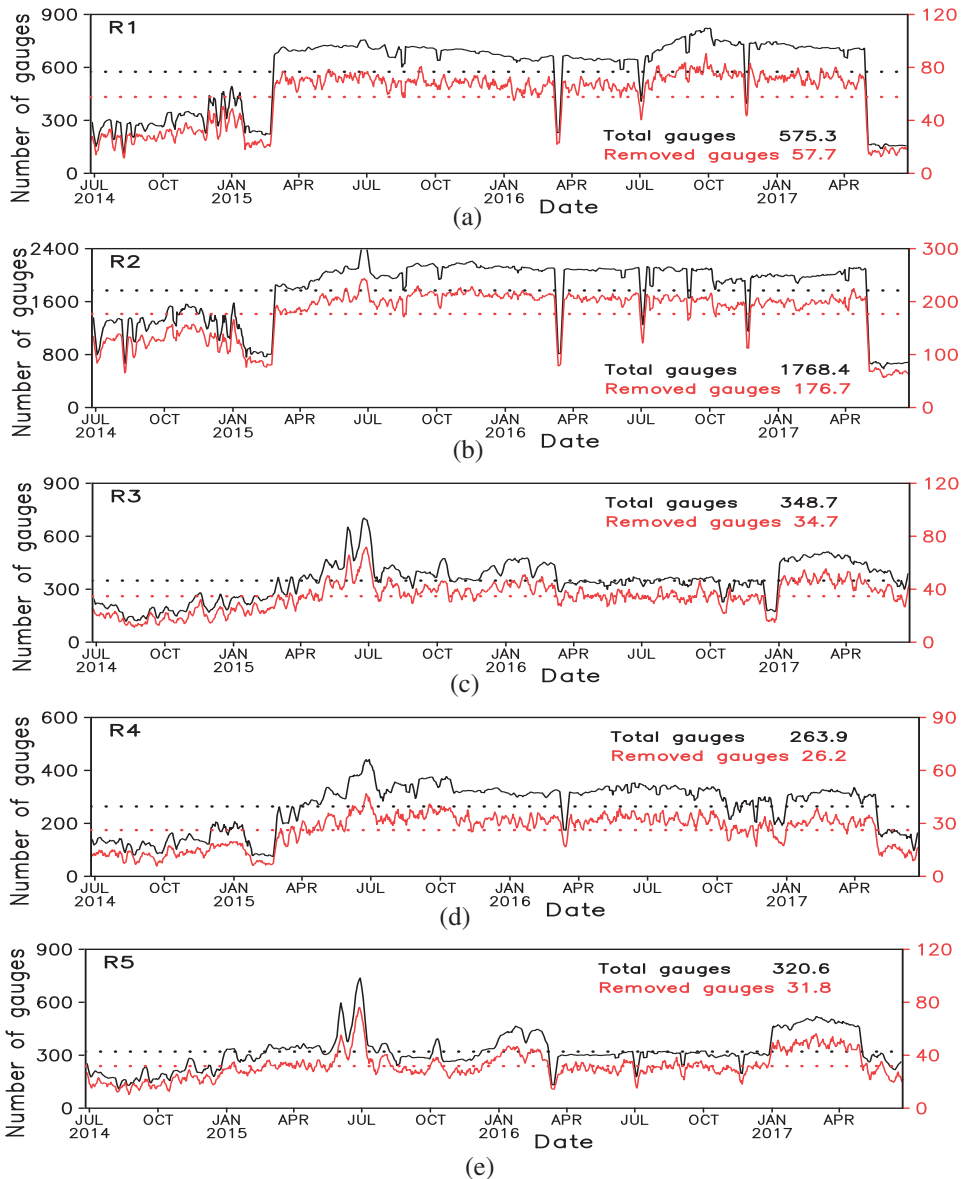
### 3. Results

#### 3.1. Temporal evolution

In this section, the temporal evolution of the mean precipitation value for each region on a daily scale and the values of the three continuous statistical indexes (ME, RMSE and  $r$ ) for each algorithm are presented. The mean precipitation values are calculated only for the points reserved for validation (10%) and do not represent the areal average of the corresponding region.

Figure 4 depicts the results for R1. Figure 4(a) shows the temporal evolution of the mean precipitation value for each day; evidently, both products represent the observed behaviour relatively well. Analysis of the ME (Figure 4(b)) shows that MERGE tends to





**Figure 3.** Total number of rain gauges for each region (black) and number of randomly removed gauges (red) for the regions R1(a), R2(b), R3(c), R4(d) and R5(e). Dotted lines represent the mean values for the whole period.

overestimate the amount of precipitation by approximately  $0.11 \text{ mm day}^{-1}$ , while CoSch is prone to underestimating the amount of precipitation by approximately  $-0.21 \text{ mm day}^{-1}$ . The RMSE evolution is depicted in Figure 4(c), demonstrating that both products produce very similar results. The largest RMSE values are found in January. An analysis of  $r$  shows that the correlations between MERGE and the rain gauges are slightly higher than those for CoSch for the whole analysed period (Figure 4(d)). In addition, R2 (Figure 5(a)) depicts results very similar to those obtained for R1, that is, very similar temporal evolution characteristics

**Table 1.** Statistical index.

Statistical index	Root mean square error (mm day <sup>-1</sup> )	Mean error (mm day <sup>-1</sup> )	Pearson correlation coefficient
Equation	$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_i - O_i)^2}{N}}$	$ME = \frac{1}{N} \sum_{i=1}^N (P_i - O_i)$	$r = \frac{\sum_{i=1}^N (O_i - \bar{O}) \cdot (P_i - \bar{P})}{\sqrt{((\sum_{i=1}^N (O_i - \bar{O})^2) (\sum_{i=1}^N (P_i - \bar{P})^2))}}$
Optimum value	0	0	1

Where;

$P_i$  – Precipitation products (MERGE or CoSch);

$O_i$  – Observational data (10% of the available pluviometric data are randomly removed);

$i$  – Point of station;

$N$  – Number of stations.

**Table 2.** Contingency table.

	Rain (gauge)	No rain (gauge)	Total
Rain (Precipitation products)	$a = H$	$b$	$p = (a + b)$
No rain (Precipitation products)	$c$	$d$	$(c + d)$
Total	$o = (a + c)$	$(b + d)$	$n = (a + b + c + d)$

Where,

$a$  – hit ( $H$ ) – an event estimated to occur, and it did occur;

$b$  – false alarm – an event estimated to occur, but it did not occur;

$c$  – miss – an event estimated not to occur, but it did occur;

$d$  – correct negative – an event estimated not to occur and it did not occur;

$H$  – Number of hits;

$p$  – Number of precipitation products;

$o$  – Number of observations;

$n$  – Total number.

**Table 3.** Categorical index used.

Categorical index	Equation	Optimum value
Adjusted equitable threat score (Mesinger 2008)	$ETS = \frac{(H_a - \frac{o^2}{n})}{(p + o + H_a - \frac{o^2}{n})}$ where; $H_a = o \left( 1 - \left( \frac{o-H}{o} \right)^{\frac{o}{p}} \right)$	1
Probability of detection	$POD = \frac{H}{o}$	1
False alarm ratio	$FAR = \frac{p-H}{p}$	0
Bias	$Bias = \frac{p}{o}$	1

Where,

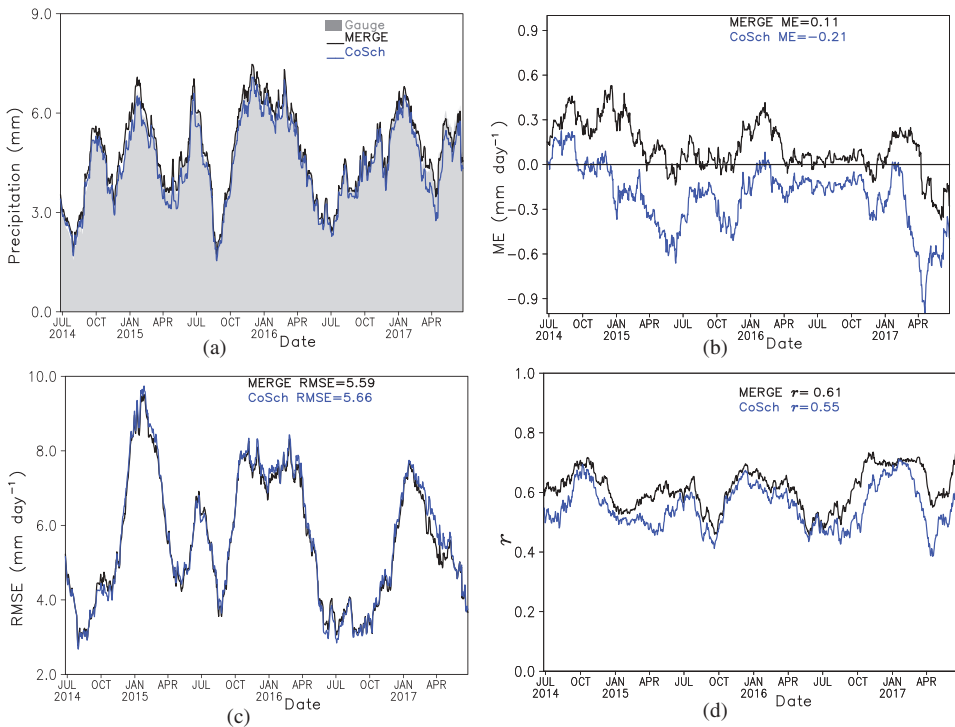
$H_a$  is the bias adjusted value of  $H$ ;

$p$ ,  $o$ ,  $H$  and  $n$  was defined in Table 2.

**Table 4.** Rain classification and thresholds.

Rain intensity classification	Precipitation thresholds (mm)
Rain/no rain	0.5
Light	2.0 to 5.0
Moderate	10.0 to 20.0
Heavy	35.0 to 50.0

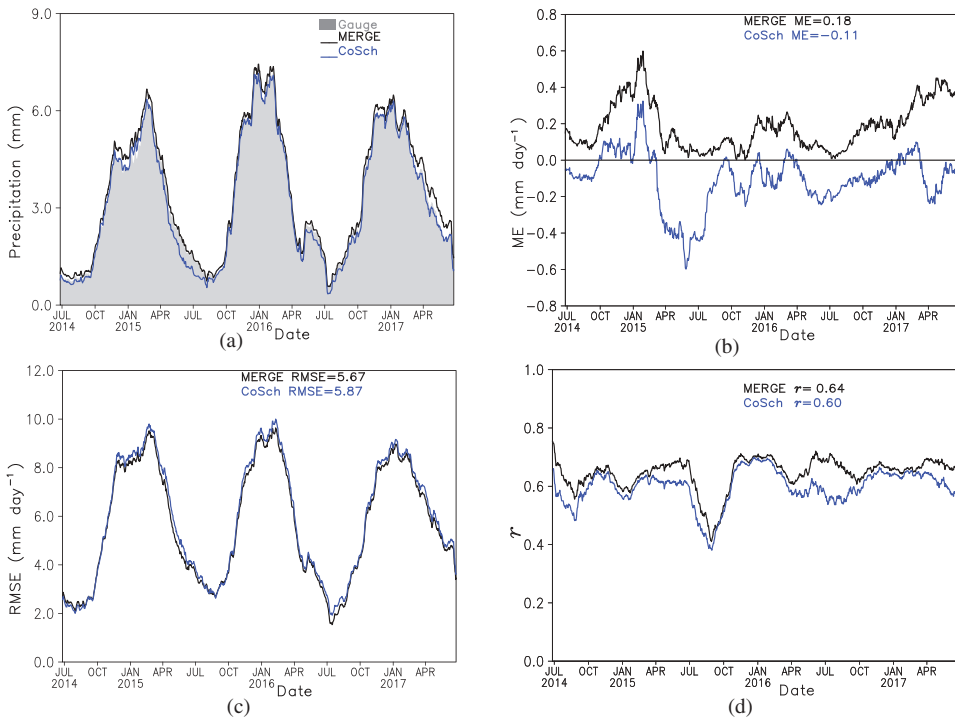
between CoSch or MERGE and the observed precipitation. However, MERGE tends to overestimate the amount of precipitation by approximately 0.18 mm day<sup>-1</sup>, while CoSch tends to underestimate the amount of precipitation by approximately -0.11 mm day<sup>-1</sup> (see Figure 5(b)). Nevertheless, the RMSE for MERGE is slightly better than that for CoSch (difference of approximately 0.2 mm day<sup>-1</sup>, see Figure 5(c)), and the correlations are slightly better when using MERGE (Figure 5(d)).



**Figure 4.** Time series of (a) the evolution of the precipitation spatial average, (b) Mean Error (ME), (c) Root Mean Square Error (RMSE) and (d) Pearson correlation coefficient ( $r$ ) for R1 (MERGE is the black line, and CoSch is the blue line). The mean values for the ME, RMSE and  $r$  are depicted close to the figure titles.

An analysis of R3 and the temporal evolution of the observed precipitation (Figure 6(a)) shows that CoSch underestimates the pluviometric regime between April and October; for the other months, both products are close to the observed values. The same underestimation tendency (approximately  $-0.23 \text{ mm day}^{-1}$ ) is noted in the ME (Figure 6(b)); note that in terms of the ME, MERGE overestimates the amount of precipitation with a mean of approximately  $0.08 \text{ mm day}^{-1}$ . Some significant differences are detected in the RMSE (Figure 6(c)) between the products both between August and December 2015 and between June and September 2016. However, the  $r$  values (Figure 6(d)) indicate slightly larger correlations with the observations for MERGE than for CoSch, with a mean correlation difference of approximately 0.13.

The precipitation time series for R4 (Figure 7(a)) indicates that CoSch does not perform well in northeastern Brazil and that this deficiency persists over the whole analysed period, with special emphasis during the wet months. In contrast, MERGE follows the observations quite closely for the whole analysed period. For this region (Figure 7(b)), CoSch also exhibits frequent underestimation (mean of approximately  $-1.01 \text{ mm day}^{-1}$ ), whereas the ME for MERGE is close to zero. Figure 7(c) displays the RMSE obtained for CoSch, with a mean difference between MERGE and CoSch of approximately  $0.46 \text{ mm day}^{-1}$ . Furthermore, the calculated values of  $r$  are also depicted in Figure 7(d); here, CoSch presents a mean correlation of approximately 0.39, while MERGE presents a mean correlation of 0.57.

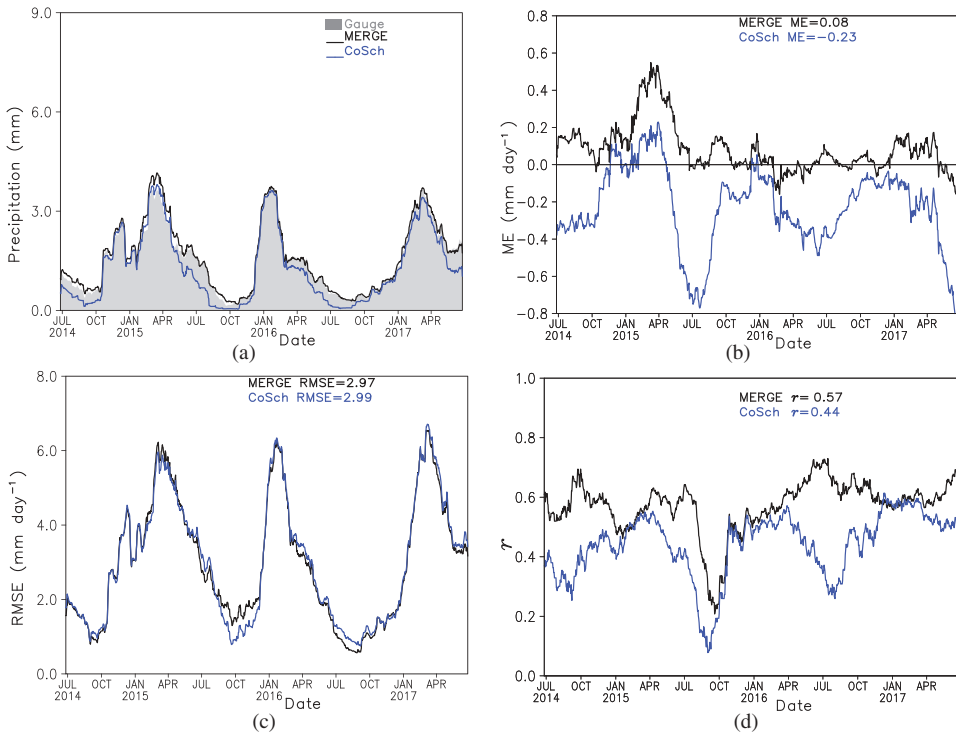


**Figure 5.** Time series of (a) the evolution of the precipitation spatial average, (b) Mean Error (ME), (c) Root Mean Square Error (RMSE) and (d) Pearson correlation coefficient ( $r$ ) for R2 (MERGE is the black line, and CoSch is the blue line). The mean values for the ME, RMSE and  $r$  are depicted close to the figure titles.

The results for R5 are very similar to those obtained for R1 and R2; that is, both products represent the pluviometric regime for the whole period quite well (Figure 8(a)), with MERGE tending to overestimate (mean of  $0.18 \text{ mm day}^{-1}$ ) and CoSch tending to underestimate (mean of  $-0.14 \text{ mm day}^{-1}$ ) the precipitation (Figure 8(b)). The RMSE (Figure 8(c)) also indicates a very similar behaviour between the two products, which present small differences for certain days during the analysed period. The values of  $r$  (Figure 8(d)) show that MERGE presents slightly larger correlations than CoSch during the analysed period.

### 3.2. Quantitative precipitation estimates (QPEs)

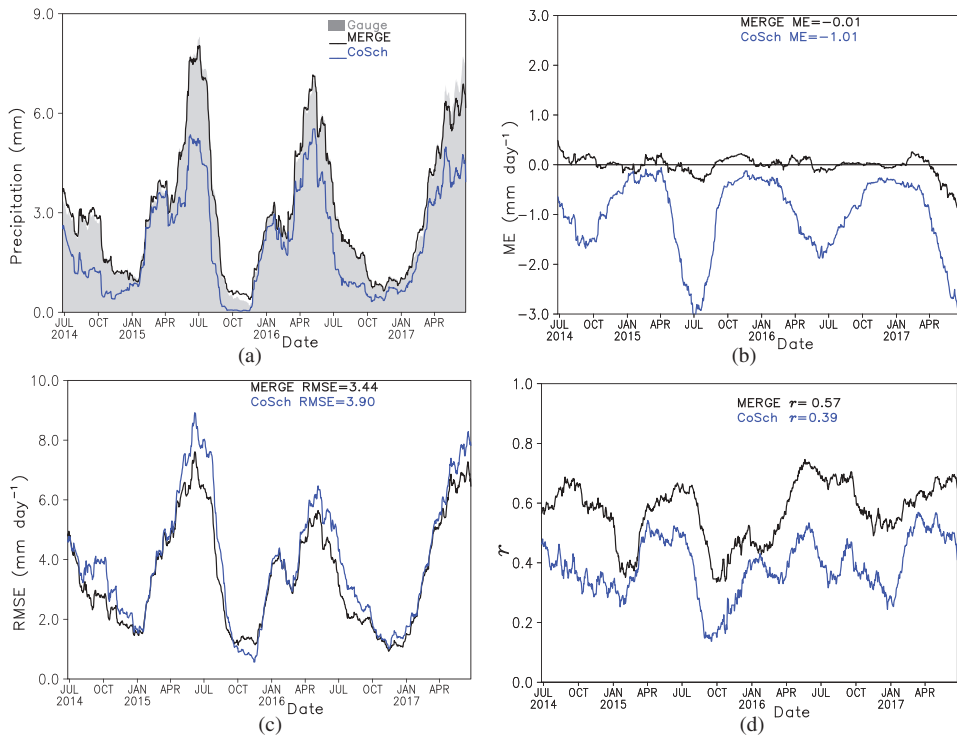
In this section, the Quantitative precipitation estimates (QPEs) are analysed for eight different thresholds (0.5, 2.0, 5.0, 10.0, 15.0, 20.0, 35.0, 50.0 mm). The results are based on a contingency table (Table 2), and the precipitation intensity is defined according to Table 4. The CSI is presented in the diagram (Roebber 2009). However, because the behaviour of the CSI is similar to the ETS, it does not need to be discussed separately. Indeed, more attention has been paid to the ETS since the CSI has a relatively low accuracy (Mesinger 2008). The POD, FAR and Bias are also presented in Roebber's performance diagram.



**Figure 6.** Time series of (a) the evolution of the precipitation spatial average, (b) Mean Error (ME), (c) Root Mean Square Error (RMSE) and (d) Pearson correlation coefficient ( $r$ ) for R3 (MERGE is the black line, and CoSch is the blue line). The mean values for the ME, RMSE and  $r$  are depicted close to the figure titles.

Figure 9(a) shows that for R1, MERGE exhibits better performance for several thresholds: rain/no rain (0.5 mm), light precipitation (2.0 to 5.0 mm) and moderate precipitation (10.0 to 15.0 mm). For heavy precipitation (greater than 20.0 mm), the performance of both products is very similar. However, both products exhibiting decreasing performance for moderate to heavy precipitation. According to Roebber's performance diagram (Figure 9(b)), the POD presents higher values for MERGE under rain/no rain and light to moderate precipitation. For thresholds above 20.0 mm, CoSch and MERGE display similar performance. In terms of the Bias, CoSch presents better results for thresholds between 0.5 and 15.0 mm with multiplicative Bias values very close to 1 (ideal case), whereas MERGE tends to overestimate the amount of precipitation. For the categories above these thresholds, both products tend to underestimate the amount of precipitation. The FAR tends to increase with increasing precipitation and is slightly smaller for CoSch between 0.5 and 10.0 mm; above these thresholds, the results are very similar between both products.

For R2, the ETS (Figure 9(c)) shows that MERGE boasts better performance than CoSch for all thresholds. However, the performance of both products drops with an increase in the precipitation threshold, except for precipitation between 0.5 and 2.0 mm. In terms of the POD and Bias (Figure 9(d)), the results for R2 are very close to those for R1, with MERGE having more opportunities to detect precipitation up to 20.0 mm and CoSch depicting an

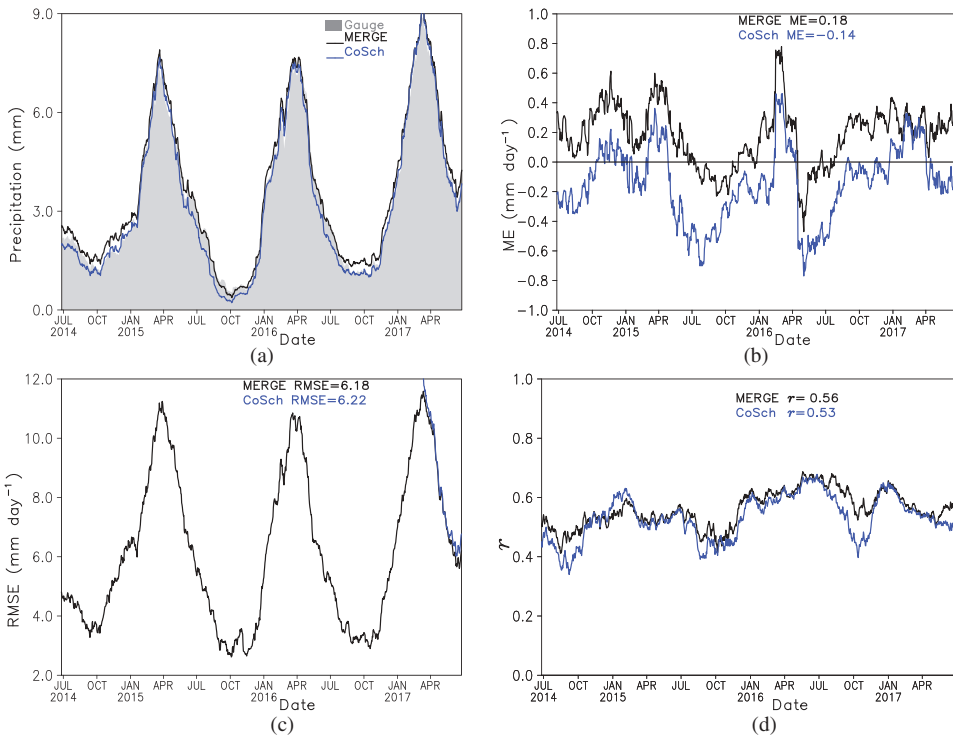


**Figure 7.** Time series of (a) the evolution of the precipitation spatial average, (b) Mean Error (ME), (c) Root Mean Square Error (RMSE) and (d) Pearson correlation coefficient ( $r$ ) for R4 (MERGE is the black line, and CoSch is the blue line). The mean values for the ME, RMSE and  $r$  are depicted close to the figure titles.

almost perfect Bias score. [Figure 9\(d\)](#) also indicates that for the categories between rain/no rain and light precipitation, MERGE has a relatively large FAR; in contrast, for heavy precipitation, CoSch has a relatively large FAR, whereas for moderate precipitation, the products have similar FAR values.

R3 represents the interior of northeastern Brazil; in this region, MERGE presents better results for the categories between rain/no rain and light precipitation. For the other thresholds, the performance of either product is very similar and drops considerably as the threshold increases ([Figure 9\(e\)](#)). In this case, the underestimation tendency of CoSch is revealed by the Bias results ([Figure 9\(f\)](#)). MERGE overestimates precipitation in the categories from rain/no rain to light precipitation. For moderate to heavy precipitation, both products have similar Bias results. The POD shows that MERGE detects more events than CoSch for thresholds between 0.5 and 15.0 mm; however, MERGE also has larger FAR values ([Figure 9\(f\)](#)).

In the eastern band of northeastern Brazil (R4, [Figure 9\(g\)](#)), MERGE outperforms CoSch in several categories except for very heavy precipitation (greater than 50.0 mm). Moreover, the performance of MERGE is almost constant (approximately 0.5) among the different precipitation categories. On the other hand, CoSch presents a performance increase as the precipitation threshold increases – from 0.28 for rain/no rain to 0.5 for heavy precipitation which results in CoSch exhibiting better performance than MERGE. Considering the Bias



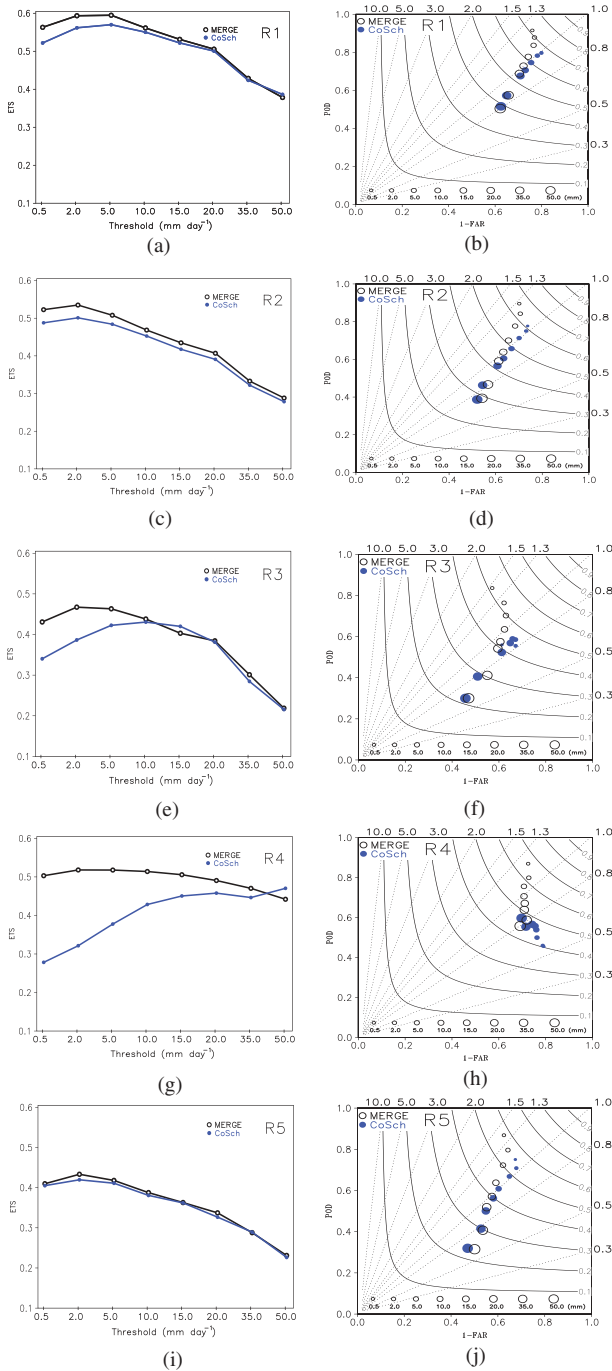
**Figure 8.** Time series of (a) the evolution of the precipitation spatial average, (b) Mean Error (ME), (c) Root Mean Square Error (RMSE) and (d) Pearson correlation coefficient ( $r$ ) for R5 (MERGE is the black line, and CoSch is the blue line). The mean values for the ME, RMSE and  $r$  are depicted close to the figure titles.

(Figure 9(h)), CoSch tends to underestimate the observed precipitation in all categories, while MERGE displays a tendency to overestimate for thresholds between 0.5 and 5.0 mm and underestimate for thresholds above 15.0 mm. Similar to R3, the POD and FAR values associated with MERGE are larger than those obtained through CoSch (Figure 9(h)).

In R5, which represents northern Brazil, very similar ETS values are obtained for both products (Figure 9(i)). A slight increase in performance is observed between 0.5 and 2.0 mm, and a reduction is observed in the other thresholds. The Bias results (Figure 9(j)) indicate that CoSch presents a multiplicative bias very close to the ideal value (one), with the exception of heavy precipitation, which CoSch tends to underestimate. MERGE overestimates precipitation from rain/no rain to light precipitation. For moderate to heavy precipitation, MERGE and CoSch show similar behaviours. Figure 9(j) also depicts results for the POD and FAR that are very similar to those found for R2 and R4, that is, large values of the POD and FAR for the categories between rain/no rain and light precipitation. Finally, the results for moderate to heavy precipitation are similar between the different products.

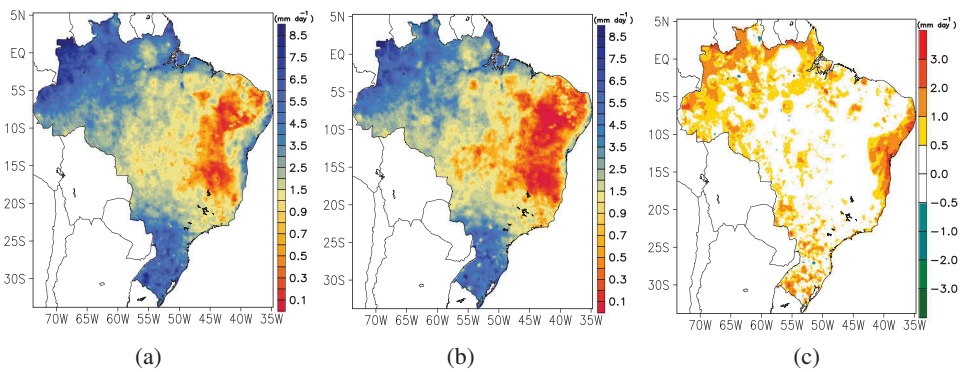
### 3.3. Spatial distributions

The spatial distributions of MERGE and CoSch are analysed in consideration of both the mean value over the whole study period in every grid point the differences between the



**Figure 9.** Adjusted equitable threat score (left panel) and the performance diagram (right panel) for MERGE (black) and CoSch (blue) for regions R1 (a and b), R2 (c and d), R3 (e and f), R4 (g and h), and R5 (i and j). The circles represent the eight precipitation thresholds. The smallest circle represents the rain/no rain threshold (0.5 mm), and the largest circle represents the threshold above 50.0 mm.





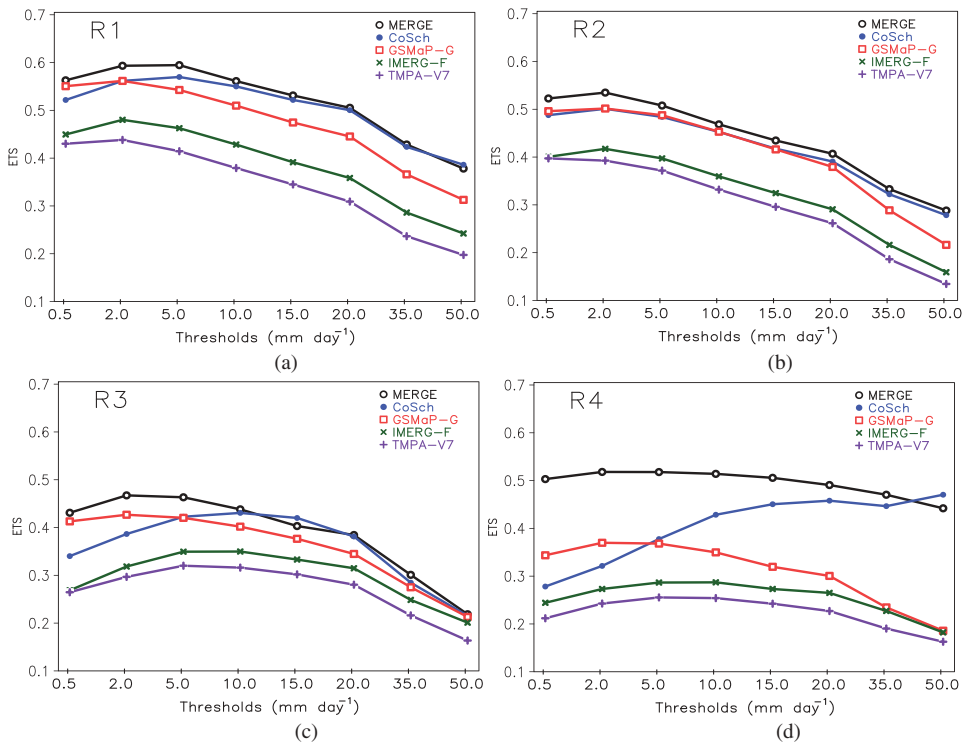
**Figure 10.** Spatial distribution of the mean precipitation (whole period) for (a) MERGE and (b) CoSch, (c) difference between (a) and (b).

products (Figure 10). In general, the spatial patterns are very similar, except for those in northeastern Brazil, where the precipitation regimes R3 and R4 are located (Figure 10(a)). CoSch (Figure 10(b)) depicts an inability to characterize precipitation regimes that present values smaller than the observed values. The differences between MERGE and CoSch (Figure 10(c)), reveal that MERGE depicts consistently larger values of precipitation all over Brazil, and the largest differences occur over the north and northeastern parts of the country.

#### 4. Discussion

Blended regional products, which have been developed for several regions (China, Australia, India, Ethiopia, and Brazil, among others), have shown numerous advantages over both satellite estimates and gauge-only interpolations (Wu et al. 2018; Li and Shao 2010; Chappell et al. 2013; Mitra et al. 2013; Dinku et al. 2014; Vila et al. 2009; Rozante et al. 2010). In particular, the two CPTC products examined herein capture the observed precipitation time series well. However, differences are noted between the products. In regions R3 and R4, CoSch displays a relatively poor performance compared to MERGE. Rozante et al. (2018) found similar errors for regions R3 and R4 when analysing global precipitation products (IMERG-Final, TMPA-V7, and GSMaP-Gauge); thus, the poor performance of CoSch could be associated with its method that prioritizes satellite estimates. In terms of QPEs, both regional CPTC products perform better than the global products (Figure 11). Additionally, better skills are obtained for regions R1, R2 and R4. Among them, in region R4, MERGE depicts a stable and satisfactory prediction skill across the different precipitation thresholds, while CoSch is characterized by decaying performance in region R4, which is associated with the deficiencies of CoSch in the reproduction of rain. Moreover, for the  $< 2.0$  mm threshold in region R4, CoSch performs even worse than GSMaP-Gauge but is still superior to both IMERG-Final and TMPA-V7.

In a recent study, Rozante et al. (2018) compared global products for the same regions used in the present study and found that in terms of QPEs, GSMaP-Gauge performs better than the other global products. The results of the present paper show



**Figure 11.** Adjusted equitable threat score for MERGE (black), CoSch (blue), GSMaP-G (red), IMERG-F (green), TMPA-V7 (purple) for regions R1 (a), R2 (b), R3 (c), R4 (d). Figure adapted from Rozante et al. (2018).

that MERGE and CoSch generally perform even better than GSMaP-Gauge in all the analysed regions (Figure 11). Blended satellite-gauge products are expected to improve independent deficiencies in disparate product satellite estimates and observed gauge data. Satellites are spatially comprehensive yet less accurate, while observational data are more accurate but spatially restricted (Chappell et al. 2013). Thus, the initial hypothesis is satisfactorily achieved with the two CPTec products, which perform better and with shorter latency than global products. The evaluations of five different regions allow us to better understand the performance of both products in different precipitation regimes. Furthermore, regionalized evaluations avoid limitations associated with large-scale averages. Renzullo et al. (2011) used a stochastic approach to obtain spatially explicit rainfall uncertainties, and Prakash et al. (2019) demonstrated that variability in the rain gauge density impacts the mean values and thresholds of precipitation. However, the random selection of removed gauges reaches nearly 10% for every region (see Figure 3); therefore, the significance and robustness of the statistics are not compromised, and thus, the regional characteristics of the products can be reliably assessed.

## 5. Conclusions

The performance of the precipitation estimates obtained from the CoSch and MERGE algorithms are evaluated for five regions in Brazil characterized by distinct precipitation

regimes. The period of analysis is four years from 26 June 2014, to 25 June 2017. The 24-hour accumulated precipitation for regions R1, R2 and R5 is well represented by both products. However, for R3 and R4, where the observational network is dense, only MERGE represents the precipitation well over the study period; in contrast, CoSch cannot effectively capture the precipitation patterns from April to November, during which the precipitation regime is dominated by warm clouds. Warm clouds are underestimated by satellite sensors (Rozante et al. 2018; Zeng et al. 2018) and consequently directly affect the CoSch results since the CoSch algorithm relies on the removal of both additive and multiplicative biases. MERGE does not seem to be affected by such a satellite deficiency over regions with a high observational density, which is likely related to the fact that MERGE retrieves information from sensors in the neighbourhood of existing observations and interpolates only observational data. Upon analysing various statistical indexes, the ME shows that MERGE presents a tendency to overestimate the amount of precipitation, while CoSch underestimates the amount of precipitation. This behaviour is observed over all the analysed regions, with R3 and R4 presenting larger differences between MERGE and CoSch. In accordance with the ME, the RMSE and  $r$  values indicate better results for MERGE, mainly over R3 and R4.

The QPEs reflect that MERGE performs better for rain/no rain to light precipitation (0.5 to 5.0 mm) for all the analysed regions. For moderate rain (10.0 to 20.0 mm), MERGE performs better in R1, R2 and R4. However, the results in R3 and R5 are very similar. Concerning heavy precipitation (35.0 to 50.0 mm), the performance of both products is very similar, with the exception of R4, where MERGE performs better than CoSch for the 35.0 mm threshold. For the 50.0 mm threshold, CoSch is superior to MERGE. The QPEs show that for all the analysed regions, MERGE overestimates the precipitation in the rain/no rain to light precipitation categories and tends to underestimate it in the heavy precipitation category. On the other hand, CoSch depicts an almost perfect Bias for rain/no rain to moderate precipitation in regions R1, R2 and R5, whereas it underestimates the amount of precipitation in all the categories for R3 and R4. Furthermore, for the heavy precipitation category, the PODs of both products are similar, while for the rain/no rain and light and moderate precipitation categories, MERGE performs better than CoSch.

The results obtained thus far show that CPTEC's product performs better than the global products and with a decreased latency. In specific terms, MERGE offers slightly better results than CoSch, especially for R3 and R4. Based on the present study, actions can be taken to improve the CoSch and MERGE algorithms.

## Acknowledgements

The authors would like to acknowledge Fundação de Amparo à Pesquisa de São Paulo (FAPESP) - project 2018/11160-2 "Sistema de Monitoramento Hidrometeorológico (SMH) Baseado em Produtos de Sensoriamento Remoto".

## Disclosure statement

No potential conflict of interest was reported by the authors.

**ORCID**

José Roberto Rozante  <http://orcid.org/0000-0002-1105-6988>

**References**

- Aonashi, K., J. Awaka, M. Hirose, T. Kozu, T. Kubota, G. Liu, S. Shige, et al. 2009. "GSMaP Passive Microwave Precipitation Retrieval Algorithm : Algorithm Description and Validation." *Journal of the Meteorological Society of Japan* 87A :119–136. doi:10.2151/jmsj.87A.119.
- Brahm, M., D. Vila, S. Martinez Saenz, and D. Osgood. 2019. "Can Disaster Events Reporting Be Used to Drive Remote Sensing Applications? A Latin America Weather Index Insurance Case Study." *Meteorological Applications*. doi:10.1002/met.1790.
- Chappell, A., L. J. Renzullo, T. H. Raupach, and M. Haylock. 2013. "Evaluating Geostatistical Methods of Blending Satellite and Gauge Data to Estimate near Real-Time Daily Rainfall for Australia." *Journal of Hydrology* 493 (June): 105–114. doi:10.1016/j.jhydrol.2013.04.024.
- Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin. 2002. "Global Land Precipitation: A 50-Yr Monthly Analysis Based on Gauge Observations." *Journal of Hydrometeorology* 3 (3): 249–266. doi:10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.
- Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. Wayne Higgins, and J. E. Janowiak. 2008. "Assessing Objective Techniques for Gauge-Based Analyses of Global Daily Precipitation." *Journal of Geophysical Research* 113: D4. doi:10.1029/2007JD009132.
- Dinku, T., K. Hailemariam, R. Maidment, E. Tarnavsky, and S. Connor. 2014. "Combined Use of Satellite Estimates AND Rain Gauge Observations to Generate High-Quality Historical Rainfall Time Series over Ethiopia: Combining Satellite and Ground Observations." *International Journal of Climatology* 34 (7): 2489–2504. doi:10.1002/joc.3855.
- Follansbee, W. A. 1973. "Estimation of Average Daily Rainfall from Satellite Cloud Photographs." *NOAA Technical Memorandum NESS*; 44. Washington, DC.
- Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao. 2004. "Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System." *Journal of Applied Meteorology* 43 (12): 1834–1853. doi:10.1175/JAM2173.1.
- Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Guojun, Y. Hong, K. P. Bowman, and E. F. Stocker. 2007. "The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales." *Journal of Hydrometeorology* 8 (1): 38–55. doi:10.1175/JHM560.1.
- Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider. 1997. "The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset." *Bulletin of the American Meteorological Society* 78 (1): 5–20. doi:10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.
- Huffman, G. J., D. T. Dan Braithwaite, R. Bolvin, J. K. Hsu, and P. Xie. 2015. "Nasa Global Precipitation Measurement (Gpm) Integrated Multi-satellite retrievals for Gpm (Imerg)." In algorithm Theor Basis Document (Atbd) Version, 2015. Nasa, Greenbelt, Maryland
- Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie. 2004. "CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution." *Journal of Hydrometeorology* 5 (3): 487–503. doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.
- Li, M., and Q. Shao. 2010. "An Improved Statistical Approach to Merge Satellite Rainfall Estimates and Raingauge Data." *Journal of Hydrology* 385 (1–4): 51–64. doi:10.1016/j.jhydrol.2010.01.023.
- Martins, E. S. P. R., E. D. Nys, C. Molejón, B. Biazeto, R. F. V. S. Silva, and N. B. M. Engle. 2015. *Monitor de Secas do Nordeste, em busca de um novo paradigma para a gestão de secas*. 1ª Edição. Brasília. ISBN: 124p.; ISBN 978-85-88192-16-4.
- Mesinger, F. 2008. "Bias Adjusted Precipitation Threat Scores." *Advances in Geosciences* 16 (April): 137–142. doi:10.5194/adgeo-16-137-2008.

- Mitra, A. K., I. M. Momin, E. N. Rajagopal, S. Basu, M. N. Rajeevan, and T. N. Krishnamurti. 2013. "Gridded Daily Indian Monsoon Rainfall for 14 Seasons: Merged TRMM and IMD Gauge Analyzed Values." *Journal of Earth System Science* 122 (5): 1173–1182. doi:10.1007/s12040-013-0338-3.
- Prakash, S., A. Seshadri, J. Srinivasan, and D. S. Pai. 2019. "A New Parameter to Assess Impact of Rain Gauge Density on Uncertainty in the Estimate of Monthly Rainfall over India." *Journal of Hydrometeorology* 20: 821–832. doi:10.1175/JHM-D-18-0161.1.
- Prakash, S., A. K. Mitra, D. S. Pai, and A. Amir. 2016. "From TRMM to GPM: How Well Can Heavy Rainfall Be Detected from Space?" *Advances in Water Resources* 88 (February): 1–7. doi:10.1016/j.advwatres.2015.11.008.
- Renzullo, L.J., Chappell, A., Raupach, T., Dyce, P., Li, M., Shao, Q. 2011: An Assessment of Statistically Blended Satellite-Gauge Precipitation Data for Daily Rainfall Analysis in Australia, in proceedings of the 34th International Symposium on Remote Sensing of Environment, Sydney, Australia.
- Roebber, P. J. 2009. "Visualizing Multiple Measures of Forecast Quality." *Weather and Forecasting* 24 (2): 601–608. doi:10.1175/2008WAF2222159.1.
- Rozante, J., D. Vila, J. B. Chiquetto, A. Fernandes, and D. S. Alvim. 2018. "Evaluation of TRMM/GPM Blended Daily Products over Brazil." *Remote Sensing* 10 (6): 882. doi:10.3390/rs10060882.
- Rozante, J. R., D. S. Moreira, L. G. G. de Goncalves, and D. A. Vila. 2010. "Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America." *Weather and Forecasting* 25 (3): 885–894. doi:10.1175/2010WAF2222325.1.
- Tang, G., M. Yingzhao, D. Long, L. Zhong, and Y. Hong. 2016. "Evaluation of GPM Day-1 IMERG and TMPA Version-7 Legacy Products over Mainland China at Multiple Spatiotemporal Scales." *Journal of Hydrology* 533 (February): 152–167. doi:10.1016/j.jhydrol.2015.12.008.
- Vila, D. A., L. G. G. de Goncalves, D. L. Toll, and J. R. Rozante. 2009. "Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America." *Journal of Hydrometeorology* 10 (2): 533–543. doi:10.1175/2008JHM1048.1.
- WMO (World Meteorological Organization). 2009. *Manual on Codes. Volumes 1.2 & 1.2 Volumes 1.2 & 1.2*. Geneva, Switzerland: Secretariat of the World Meteorological Organization.
- Woldemeskel, F. M., B. Sivakumar, and A. Sharma. 2013. "Merging Gauge and Satellite Rainfall with Specification of Associated Uncertainty across Australia." *Journal of Hydrology* 499 (August): 167–176. doi:10.1016/j.jhydrol.2013.06.039.
- Wu, Z., Y. Zhang, Z. Sun, Q. Lin, and H. Hai. 2018. "Improvement of a Combination of TMPA (Or IMERG) and Ground-Based Precipitation and Application to a Typical Region of the East China Plain." *Science of the Total Environment* 640–641 (November): 1165–1175. doi:10.1016/j.scitotenv.2018.05.272.
- Xie, P., and P. A. Arkin. 1997. "Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs." *Bulletin of the American Meteorological Society* 78 (11): 2539–2558. doi:10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.
- Zeng, Q., Y. Wang, L. Chen, Z. Wang, H. Zhu, and Li, B. 2018. "Inter-comparison and Evaluation of Remote Sensing Precipitation Products over China from 2005 to 2013." *Remote Sensing* 10 (3): 168. doi: 10.3390/rs10020168.