ABSTRACT: Extreme precipitation events are a serious threat to societal well-being over rainy areas such as Bangladesh. The reliability of studies of extreme events depends on data quality and their spatial and temporal distribution, although these subjects remain with knowledge gaps in many countries. This work focuses on the analysis of four satellite-based precipitation products for monitoring intense rainfall events: the Climate Hazards Group Infrared Precipitation with Station Data (CHIRPS), the PERSIANN–Climate Data Record (PERSIANN-CDR), the Integrated Multisatellite Retrievals (IMERG), and the CPC morphing technique (CMORPH). Five indices of intense rainfall were considered for the period 2000–19 and a set of 31 rain gauges for evaluation. The number and amount of precipitation associated with intense rainfall events are systematically underestimated or overestimated throughout the country. While random errors are higher over the wetter and higher-elevation northeastern and southeastern parts of Bangladesh, biases are more homogeneous. CHIRPS, PERSIANN-CDR, and IMERG perform similar for capturing total seasonal rainfall, but variability is better represented by CHIRPS and IMERG. Better results were obtained by IMERG, followed by PERSIANN-CDR and CHIRPS, in terms of climatological intensity indices based on percentiles, although the three products exhibited systematic errors. IMERG and CMORPH systematically overestimate the occurrence of intense precipitation events. IMERG showed the best performance representing events over a value of 20 mm day$^{-1}$; CMORPH exhibited random and systematic errors strongly associated with a poor representation of interannual variability in seasonal total rainfall. The results suggest that the datasets have different potential uses and such differences should be considered in future applications regarding extreme rainfall events and risk assessment in Bangladesh.

KEYWORDS: Asia; Extreme events; Satellite observations; Climate services

1. Introduction

Both globally and in South Asia, Bangladesh is recognized as one of the most vulnerable countries to climate change and natural disasters. Extreme hydrometeorological events can lead to floods, prolonged waterlogging, and landslides (Bhowmik et al. 2021). These disproportionately affect rural livelihoods and can damage infrastructure (Eckstein et al. 2020). These events occur predominantly during the monsoon season, typically from June to September when rainfall is concentrated (Ahmed and Karmakar 1993). The impact of intense rainstorms can be enhanced by Bangladesh’s geographical characteristics including proximity of the sea, and generally low-elevation and flat terrain (Mirza 2011). Combined with one of the world’s highest population densities and given that nearly 60% of Bangladesh’s land is under agricultural land use (World Bank 2020a,b), these conditions significantly increase the risk of exposure, damage, and losses from weather-induced disasters.

The summer monsoon season corresponds to the primary season during which the country’s rainfed rice crop is grown. Monsoon rainfall is therefore crucial for land preparation and crop establishment, as farmers transplant only after sufficient rainfall has occurred (Shelley et al. 2016). However, where extreme rainfall events and consequent flooding are experienced within the monsoon, farmers can lose seedlings (Ismail et al. 2013) and experience “lodging”—the knocking down or bending of crop stems from wind storms and rainstorm events (Weng et al. 2017). This can increase work where farmers need to retransplant, while also reducing profitability and yield in the case of lodging, respectively.

These issues highlight the importance of accurate precipitation measurements to characterize the number, frequency, and variability of extreme events over a small-area country such as Bangladesh where rainfall exhibits high interannual and within-season variation (Shahid 2010; Kelley et al. 2020). Furthermore, daily extreme precipitation events are expected
to become more frequent in future climate scenarios over South Asia (Donat et al. 2016). While gauge-based observation networks are considered as a primary source of reliable data, their spatial coverage is still a limitation for studies in which dense points in space are required, or where short-lived, highly variable convective systems are the dominant precipitation mechanism (Sunilkumar et al. 2016). Apart from ground observation networks, other instruments are available for measuring precipitation, namely, radar (radio detection and ranging) and satellites. Weather radar systems have proven to be very useful in generating detailed information at regional levels, but there are few networks worldwide providing uniform and long-term data, limiting the geographic coverage necessary for studies of extreme events (Liang and Ding 2017). However, time and space continuous and kilometric information provided by gridded precipitation products merging multiple ground and satellite measurements have permitted the capture of complex spatial and temporal variability, allowing continuous measurements available for multiple applications in datasets (Sorooshian et al. 2011).

Several satellite-based, gridded precipitation datasets are currently available, combining different data sources and differing in resolution and spatial coverages (Beck et al. 2017). They have become available due to technical and methodological advances in satellite sensors and remote sensing algorithms. These products are generated from sensors that vary widely in their characteristics, from geostationary or geocentric and active to passive signals (Kidd and Levizzani 2011); consequently, retrieved precipitation can be heterogeneous and highly vulnerable to error (Sun et al. 2018). These uncertainties have been reduced with the growth of research to develop gridded products generated from the merging of direct satellite precipitation and ground observations (Xie et al. 2003). Some of the most commonly used operational satellite-derived precipitation products include the microwave-based Climate Prediction Center (CPC) morphing technique (CMORPH) (Joyce et al. 2004), the microwave- and infrared-based Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) (Ashouri et al. 2014), the National Aeronautics and Space Administration’s (NASA) Tropical Rainfall Measuring Mission (TRMM) (Huffman et al. 2007), and the recent Integrated Multisatellite Retrievals for the Global Precipitation Measurement (IMERG) product (Huffman et al. 2017). Additionally, the Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) (Funk et al. 2015) has become one of the most widely used products given its easy access and almost real-time availability, which is generated by merging satellite and rain gauge data. Their global coverage and in some cases near-real-time availability make them useful for multiple applications such as the monitoring of extreme rainfall events and associated impacts in agriculture and other sectors (Davenport et al. 2018).

However, the abovementioned products have been reported as leading to significant variability in extreme rainfall estimations accuracy, making their assessment necessary for applications (Huang et al. 2014; Jiang et al. 2019). Differences between products are the result of multiple factors contributing to varying results in terms of estimated rainfall. Most of current satellite precipitation products are generated using rainfall retrievals from passive microwave (PMW) measurements, among others. Conversely, multiple sources of error have been described as influencing the performance of PMW sensors. For instance, Tian and Peters-Lidard (2007) attributed systematic errors in TRMM and CMORPH datasets over water bodies to inaccuracies in the characterization of temperature and emissivity of water surfaces. More recently, Petković and Kummerow (2017) reported a significant influence of the state of the atmosphere (air water content, ice scattering) and the large-scale environment on passive microwaves measured by TRMM and consequently on retrieved rainfall measurements over land. Other sources of error are the often limited ground observations available to develop correction factors, for example, over complex terrain areas where observations are usually more limited (e.g., Vicente et al. 2002), weaknesses in the physics and assumptions of the algorithms (Shige et al. 2013), sampling errors due to satellite overpass frequency (Nijssen and Lettenmaier 2004), or a combination of factors generating random and systematic errors, as in the case of IMERG (Tan et al. 2016).

Various studies have been carried out to assess satellite precipitation in Bangladesh. However, these studies have not focused on extreme events but rather on other aspects of rainfall variability. For instance, Islam et al. (2005) evaluated the performance of TRMM precipitation to retrieve daily scale precipitation, and Islam and Uyeda (2008) assessed seasonal patterns in vertical profiles of rainfall intensity derived from TRMM. Similarly, TRMM precipitation was assessed by Islam and Cartwright (2020) in terms of accumulated totals. Nashwan et al. (2019) documented differences in spatial patterns of annual and seasonal precipitation trends captured by a set of precipitation products.

The overall goal of this study is to perform an evaluation of a set of high-resolution satellite-derived daily precipitation products using percentile and threshold-based rainfall indices characterizing their intensity during the monsoon season in Bangladesh using observational data from a set of rain gauges. We considered four satellite gridded products at daily scales, namely, CHIRPS, IMERG, PERSIANN, and CMORPH. Main features such as total monsoon rainfall, as well as the climatology of rainfall intensity indices, were analyzed. Our results can potentially improve the understanding of the spatiotemporal behavior of intense rainfall events in Bangladesh and also provide guidance for the use of satellite precipitation data to stakeholders such as national meteorological and agricultural extension services. These are in addition to other applications such as weather-index-based insurance efforts focused on triangulating losses and damage from extreme weather events.

2. Materials and methods
a. Study area

Geographically, Bangladesh is part of the Ganges–Brahmaputra–Meghna delta, and is characterized by its flat and low-elevation terrain. However, hilly areas in the southeast and northeast of the country, which occupy about 21% of the
country’s surface (Brammer 2017), cause the elevation to reach heights up to 1000 m above sea level (Fig. 1). Annual rainfall varies between 1500 and 4000 mm, with the northeast being the rainiest area in the country and one of the rainiest in the world (Mohsenipour et al. 2020). The monsoonal regime is strongly concentrated during the summer, when extreme meteorological events are recurrent (Dastagir 2015).

b. Datasets

1) Rain Gauges

Ground-truth daily precipitation provided by the Bangladesh Meteorological Department (BMD) was used to evaluate the performance of the selected satellite precipitation products. This dataset consists of daily precipitation observations from 31 stations spanning the period 2000–19 (20 years), which is the overlapping period of satellite data time series. The monsoon season was defined from June through September, when extreme hydrometeorological events tend to be concentrated. Importantly, although BMD maintains a larger number of meteorological Department gauge stations used (circles). Elevation data were generated from the 90-m Shuttle Radar Topography Mission (SRTM) digital elevation model: https://www2.jpl.nasa.gov/srtm/.

2) Satellite Precipitation Products

Four global satellite-based precipitation products were evaluated in their performance in capturing intensity-based indices. These products provide time series long enough for a climatology, and their spatial resolution can be considered as comparable to the BMD stations’ area coverage. The dataset corresponds to daily precipitation from 1) CHIRPS, 2) PERSIANN–Climate Data Record (PERSIANN-CDR), 3) IMERG product, and 4) CMORPH reprocessed and bias corrected version 1.0. The main features of these products are presented as follows and summarized in Table 1.

Developed by the Santa Barbara Climate Hazards Group at the University of California in association the U.S. Geological Survey Earth Resources Observation and Science Center, the latest CHIRPS V.2 version was used in this study. CHIRPS is a satellite-derived precipitation dataset provided at 0.05° and 0.25° spatial resolutions, and from daily to annual time scales from 1981 until the present (Funk et al. 2015). The data are obtained by combining infrared cold cloud duration data and TRMM precipitation to generate a pentad rainfall estimation product. These data are subsequently blended with ground measurements using an inverse distance weighting-based algorithm with a latency of about 3 weeks. In this study, the CHIRPS V2.0 version at 0.05° × 0.05° native resolution was used.

PERSIANN-CDR (hereafter PERSIANN) precipitation is generated using an artificial neural network model developed by the National Centers for Environmental Prediction (NCEP) to convert infrared brightness temperature measured by geostationary satellites into precipitation rates. The final product is a multisatellite high-resolution estimation of daily precipitation at 0.25° × 0.25°, calibrated using monthly Global Precipitation Climatology Project (GPCP) precipitation (Ashouri et al. 2014), and provided with a latency of about 3 months.

IMERG is a last generation quasi-global (60°N–60°S latitude) high-resolution rainfall product of the Global Precipitation Measurement (GPM) considered as the successor of the TRMM mission. IMERG is derived from infrared, passive microwave sensor and radar data, is available since June 2000, and provides rainfall data at a 0.1° × 0.1° spatial resolution and every half hour to daily depending on the data processing level. We used the Level 3 final run daily product, which is generated using rain-gauge-based data from the Global Precipitation Climatology Centre (GPCC) for calibration and has a latency of 2.5 months study.

Finally, the CMORPH with corrected bias (CMORPH-CRT, hereafter CMORPH) algorithm is generated using satellite passive microwave observations to estimate instantaneous precipitation that is propagated spatially using thermal infrared observations from geostationary satellites. The “raw” product has 8-km and 30-min resolutions covering between 60°S and 60°N from 2002 to 2017. We used the more recent 0.25° × 0.25° resolution CMORPH version 1.0 daily product of Xie et al. (2017), which is a reprocessed version that uses observations from the CPC and GPCP products, provided with a latency of about 6 months.

c. Indices of precipitation intensity

Multiple indices have been proposed in order to characterize extreme precipitation events, ranging from indices describing the intensity of single events, to those allowing their statistical
aggregation (Zhang et al. 2011). In this work, we used five intensity-based precipitation indices proposed by the Expert Team on Climate Change Detection and Indices (ETCCDI; Zhang et al. 2011; Herold et al. 2018), which are summarized in Table 1. However, since rainfall in Bangladesh strongly concentrates during the monsoon season, we have slightly modified the original definition recommended by ETCCDI, taking the monsoon season from June through September (JJAS) as our period of interest, instead of the whole year. The five indices are suitable to characterize high-intensity precipitation events in Bangladesh considering their importance as physical damage agents to crops, and recurring locally generated flood events, among other relevant impacts.

Four indices characterizing the seasonal climatology of precipitation intensity were used: the sum of the daily rainfall ($\geq 1$ mm day$^{-1}$) for estimates recorded during JJAS that are higher than the interannual JJAS top 5% and 1% rainfall events (herein referred to as R95p and R99p, respectively; see Table 2), and the contribution to total JJAS precipitation by very ($> 95$th percentile) and extremely wet days ($> 99$th percentile), herein referred to as R95pTOT and R99pTOT, respectively. In addition, we used the threshold-based index R20mm, which corresponds to the annual count of days when daily precipitation is $\geq 20$ mm. Considering that Bangladesh ranks as one of the rainiest countries (Eckstein et al. 2020), absolute indices describing the number of days with precipitation above a fixed threshold are not necessarily expected to well discriminate regional differences in heavy rainfall events. However, the average daily JJAS rainfall from the 31 BMD stations is 20 mm day$^{-1}$, therefore, a threshold of 20 mm day$^{-1}$ seems to be a good choice.

d. Dataset evaluation

The performance of the four satellite products in representing intense precipitation events in Bangladesh was assessed by comparing ETCCDI intensity indices against station observations. We calculated basic metrics such as correlations and standard deviations, and the absolute root-mean-square error (RMSE$^1$) and the Bias$^2$ estimator, which are a measure of the absolute difference between datasets being compared. The closest satellite grid point to each weather station was taken for the comparison defined by the smallest Euclidean distance. Although most of the compared points are relatively proximal, some points along the coast and south of the country are relatively distant, which can result in discrepancies that are not necessarily associated with satellite products per se in the case of the lower-resolution products (Fig. S1 in the online supplemental material).

In addition to the above evaluation metrics, the 31 time series for JJAS total rainfall and the five indices were averaged and then statistically summarized at the country-level using Taylor diagrams in terms of their RMSE, standard deviation, and correlation coefficient (Taylor 2001). The corresponding average BMD time series was taken as the reference dataset.

3. Results and discussion

a. Monsoon (JJAS) satellite precipitation climatology

Maps of 20-yr climatology and interannual variability of JJAS precipitation from satellite precipitation and stations are presented in Fig. 2. For the 31 stations combined, the country average JJAS rainfall was 1663 mm, ranging from 964 to

Table 1. Summary of satellite precipitation products used.

<table>
<thead>
<tr>
<th>Product acronym</th>
<th>Product full name</th>
<th>Temporal and spatial resolution</th>
<th>Instrumentation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHIRPS$^a$</td>
<td>Climate Hazards Group Infrared Precipitation with Station</td>
<td>Daily, 0.05° × 0.05° and 0.25° × 0.25°</td>
<td>Gauge data, infrared, TRMM-T3B42 and CFSR</td>
<td>Funk et al. (2015)</td>
</tr>
<tr>
<td>PERSIANN-CDR$^b$</td>
<td>Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Climate Data Record</td>
<td>Daily, 0.25° × 0.25°</td>
<td>Infrared, artificial neural network trained on gauge and radar data. Scaled to GPCP monthly satellite-gauge data.</td>
<td>Ashour et al. (2014)</td>
</tr>
<tr>
<td>IMERG$^c$</td>
<td>Integrated Multisatellite Retrievals for the Global Precipitation Measurement</td>
<td>Half hourly, daily, 0.1° × 0.1°</td>
<td>Infrared and microwave satellite precipitation, radar, GPCP rain gauges</td>
<td>Huffman et al. (2017)</td>
</tr>
<tr>
<td>CMORPH$^d$</td>
<td>Climate Prediction Center morphing technique</td>
<td>3-hourly and daily, 0.25° × 0.25°; half-hourly, 0.07° × 0.07° and 0.25° × 0.25°</td>
<td>Passive microwave data from low-orbit satellites. GPCP monthly precipitation.</td>
<td>Xie et al. (2019)</td>
</tr>
</tbody>
</table>

$^c$ https://www.ncei.noaa.gov/data/precipitation-persiann/access/.
$^d$ https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/access/daily/0.25deg/.

1 RMSE = \(\sqrt{\frac{1}{n}\sum_{k=1}^{n}(S_k - O_k)^2}\),
2 Bias = \(\frac{1}{n}\sum_{k=1}^{n}(S_k - O_k)\), where $S$ represents satellite-derived estimations, $O$ the observations, and $n$ the total number of $k$ elements of the data pairs being compared.
3609 mm. An overall agreement is observed in terms of higher precipitation over the southeast and northeast area as captured by the four gridded products, which appears as a noticeable feature in observations. CHIRPS precipitation exhibited the best performance, likely because several local station data are incorporated in its algorithm. However, the number of stations varies greatly from one year to another in the product generation. For Bangladesh it ranges from an annual monthly mean of 2 in 2000 to 10 stations in 2019, with a minimum of 2 stations in 2001 and a mean maximum of 17 in 2017 (Fig. S2).

Moreover, some stations used to retrieve precipitation over Bangladesh are located outside the country, mostly from India. On the other hand, the number of stations used to generate the GPCC data over Bangladesh, which in turns is used to calibrate IMERG, has undergone some changes over time: the total number being close to 32 for the periods 2000–08 and 2014–18, and close to 10 in average between 2009 and 2013 (Fig. S3), which generates uncertainties in the performance of the products. Similarly, the quality index of the GPCP (Huffman et al. 1997) product that is incorporated in both CMORPH and PERSIANN experienced a significant drop during 2009 and 2013 (Fig. S4).

Both country average RMSE and Bias in Fig. 2 indicate similar performance of CHIRPS, PERSIANN, and IMERG in the representation of total precipitation during the monsoon season. Greater error in relation to stations is observed for PERSIANN and CMORPH, which show a narrower amplitude in values as compared to CHIRPS and IMERG. The latter can be corroborated by the lower precipitation values in the northeast and southeast. Thus, the main differences between products are observed for maximum mean values, with CHIRPS and CMORPH showing the higher and lower maximum, respectively. Figure 2 also shows the RMSE and Bias between estimations and observations, both calculated between gauges and the closest grid cell. There are consistent differences in the performance of satellite products: the higher

---

**Fig. 2.** (a)–(d) Climatology (mean) (2000–19) and (e)–(h) interannual variability (standard deviation) of total JJAS precipitation for the four precipitation products and rain gauges. Spatial statistics and error (country average RMSE and Bias) of satellite products are displayed at the top and bottom of each map, respectively. RMSE and Bias at the bottom of each map represent the corresponding regional average (n = 31).
country-average RMSE is observed for CMORPH and PERSIANN, followed by CHIRPS and IMERG. While PERSIANN shows a slightly positive country-level Bias (48 mm), CMORPH exhibits a generalized underestimation of total JJAS precipitation (Bias = −240 mm). The latter might be explained by restrictions associated with microwave data to perform a full global coverage given the less complete sampling, which has been previously described for passive/active microwave sensors in countries such as India (Mondal et al. 2018; Prakash 2019). Similarly, Sunilkumar et al. (2015) found a negative bias of similar magnitude over Indian regions adjacent to Bangladesh which were associated with a smaller number of rainy days retrieved by CMORPH in relation to the India Meteorological Department observational gridded product (Rajeevan et al. 2006). Interannual variability, represented by the standard deviation (SD), is displayed in Figs. 2e–h. As expected, the areas with the highest accumulated JJAS precipitation are also those that show the highest interannual variability. Disagreements are observed among satellite products, being CHIRPS/CMORPH the one showing the lowest/highest SD and RMSE. While CMORPH overestimates interannual variability (Bias = 106 mm), especially over the rainiest areas in the northeast and southeast, CHIRPS and IMERG show a lower and more homogeneous variability along the country. In terms of error metrics, CHIRPS, PERSIANN, and IMERG tend to underestimate SD

<table>
<thead>
<tr>
<th>Time series</th>
<th>Product</th>
<th>Correlation coefficient</th>
<th>Standard deviation (mm)</th>
<th>RMSE (mm)</th>
<th>Bias (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>CHIRPS</td>
<td>0.88</td>
<td>220</td>
<td>128</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>0.89</td>
<td>213</td>
<td>109</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>IMERG</td>
<td>0.97</td>
<td>220</td>
<td>83</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>0.38</td>
<td>404</td>
<td>441</td>
<td>−239</td>
</tr>
<tr>
<td>SD</td>
<td>CHIRPS</td>
<td>0.54</td>
<td>99</td>
<td>120</td>
<td>−43</td>
</tr>
<tr>
<td></td>
<td>PERSIANN</td>
<td>0.46</td>
<td>57</td>
<td>419</td>
<td>−403</td>
</tr>
<tr>
<td></td>
<td>IMERG</td>
<td>0.88</td>
<td>98</td>
<td>175</td>
<td>−163</td>
</tr>
<tr>
<td></td>
<td>CMORPH</td>
<td>0.35</td>
<td>166</td>
<td>321</td>
<td>−273</td>
</tr>
</tbody>
</table>
in a similar magnitude (average negative Bias from \(-64\) to \(-87\) mm), the three products showing the area of higher variability toward the east, which is coinciding with the maximum amount of Figs. 2a–d.

Time series country-average total JJAS precipitation described by Fig. 3 presents the interannual variation of regional climatology (station-level annual average). CHIRPS, PERSIANN, and IMERG show a reasonable agreement with observations in terms mean values and interannual fluctuations (Fig. 3a), although discrepancies are observed in terms of bias for some years such as in 2006 where CHIRPS presents lower mean values. On the other hand, CMORPH largely underestimates JJAS precipitation in Bangladesh, except for the last years of the time series (Fig. 3a), whose magnitude is similar to the those reported by Mondal et al. (2018) over India. The spatial variability for each year, calculated as the SD of grid point for every year, and presented in Fig. 3b, describes variability from about 300 mm in the case of PERSIANN to almost 800 mm for CHIRPS. The latter product produced the closest values to ground observations, followed by IMERG. Error metrics including annual spatial RMSE and Bias (Figs. 3c and 3d, respectively) calculated for the 31 points and compared yearly, show a similar RMSE magnitude between 2007 and 2019 for PERSIANN and CMORPH, and also during the complete time period for the case of CHIRPS and IMERG. Interannual regional Bias results are similar for CHIRPS, PERSIANN, and IMERG, though they gradually change from around \(-600\) to 400 mm in the case of CMORPH.

Table 3 summarizes the comparative and descriptive statistics (correlation coefficient and SD) and error features of satellite products as compared to station data. Both correlations, SD, RMSE, and Bias suggest a good agreement between observations and CHIRPS, PERSIANN, and IMERG products in terms of mean JJAS total rainfall. The performance of these three products decreases when the time series of SD are compared. Notably, CMORPH presents a poorer performance for all the statistical indices used.

b. Intense precipitation indices for rain gauges

The climatological values of the five precipitation indices calculated from station data only are presented in Fig. 4. Figure S5 shows the statistical distribution of the indices for all 31 rain gauges. Precipitation associated with both very (R95p) and extremely (R99p) wet days exhibit a spatial pattern similar to total JJAS rainfall (Fig. 2), ranging from 268 to 852 mm and from 80 to 253 mm, respectively, the highest values observed over the east (Figs. 4a,b). The total contribution from the top 5\% rainy days (R95pTOT, Fig. 4c) averages 25\% at the country level, ranging from 23\% to 29\%. The contribution of extremely wet days (R99pTOT, Fig. 4d) ranges from 7\% to 9\%, averaging 8\%. Figures 4c and 4d show a slightly different spatial distribution for R95pTOT and R99pTOT in relation to R95p and R99p. While the rainy area in the northeast has the highest mean R95p and R99p, the total contribution of very and extremely wet days is relatively lower. In comparison, some stations in central and northern Bangladesh show a higher contribution from very rainy days. In addition, the respective averages of 25\% and 8\% are indicative that a small number of heavy precipitation events can account for a significant percentage of total precipitation during the rainy season. Interestingly, the highest values of R95pTOT and R99pTOT are observed for a station located in an area of complex terrain in the southeast with the highest elevation in our dataset, which suggests an orographic effect in the origin of intense precipitation events. As for the absolute threshold-based index R20mm (Fig. 4e), the climatological values show a spatial distribution that more closely resembles Rp95p and Rp99p than Rp95pTOT and Rp99pTOT, where the highest values, in this case the number of days of heavy rain, are mainly concentrated in the south/southeast of the country, and whose values range from 15 to 50 days.

Fig. 4. (a)–(e) Climatology (mean) and (f)–(j) interannual variability (standard deviation) of intense precipitation indices (2000–19) for rain gauges.
Considering the country-average R95p and R99p of 444 and 139 mm (Figs. 4a,b), respectively, interannual variation in total rainfall from the top 5% and 1% precipitation events equal 61% (274 mm) and 120% (168 mm) of the corresponding long-term mean, respectively (Fig. 4g). The standard deviation for R95pTOT and R99pTOT ranges from 9% to 17% and from 7% to 13%, respectively, with the spatial distribution of standard deviation exhibiting a pattern similar to the climatological values, with variability that tends to be higher for stations located in the center and north of the country (Figs. 4h,i). The standard deviation for R20mm (Fig. 4j) depicts values that are more homogeneously distributed around the country, ranging from 4 to 7 days on average.

c. Intense precipitation indices for satellite products

In this section, the performance of the four precipitation products is evaluated in representing the intense precipitation indices. In general, Figs. 5a–d shows a similar spatial distribution of R95p between products and to the previously described patterns in total precipitation (Fig. 2). CHIRPS and PERSIANN look similar in their climatologies, while IMERG and CMORPH show higher values along the country in general. The latter can be corroborated by RMSE values of Figs. 5e–h. The error over inland areas appears to be similar for the studied datasets, although it is higher for CMORPH depicting 100 mm on average. In addition, higher errors are observed over coastal areas, which might be due to the complexity of retrieving precipitation data in sea–land transition areas, a challenge that has been documented for microwave and infrared sensors and similar error magnitudes (Kim et al. 2017; Jiang et al. 2019). Additionally, Figs. 5e–g show that the error associated with CHIRPS, PERSIANN, and IMERG are quite similar.

The Bias statistics (Figs. 5i–l) shows a shift from an average underestimation of 6 mm by CHIRPS to an average overestimation of 131 mm by IMERG. CHIRPS overestimates R95p in the northeast and south areas of the country. In the case of PERSIANN, a gradient that goes from a general underestimation in the south to an overestimation of R95p in the north is observed. On the other hand, both TRMM and CMORPH
show a general overestimation of R95p throughout the country. These results indicate that although the spatial pattern of R95p is adequately captured by these precipitation products, the associated error varies considerably. The highest RMSE is presented by CMORPH, especially in coastal and southern areas. Furthermore, both CHIRPS, PERSIANN, and IMERG error do not differ substantially in terms of magnitude and spatial distribution. Satellite products present a varying performance in terms of Bias, which has been also highlighted when probability of detection metrics are used (Islam 2018).

Similarly, Fig. 6 shows the climatology of R99p. PERSIANN generally shows the lower R99p, followed by CHIRPS and CMORPH. IMERG has the highest values, which range between 123 and 372 mm. In the present case, PERSIANN shows the lowest mean RMSE (166 mm), slightly below CHIRPS (174 mm) and IMERG (183 mm), which range from 89 to 293 mm and from 99 to 383 mm, respectively, while CMORPH shows higher associated error. The Bias exhibits a general underestimation by CHIRPS and PERSIANN, and overestimation by IMERG and CMORPH, suggesting contrasting behavior in capturing total precipitation during extremely wet days.

Based on the above-presented results, it could be argued that the four products evaluated in this study provide negatively or positively biased results for R95p and R99p; as such, they might not adequately replace the usefulness of high-quality ground measurements. However, ground data collection is associated with high transactional costs, and observed systematic biases in the estimates could offer opportunities for the use of bias correction methods. These patterns agree with those obtained in terms of total precipitation (Fig. 2).

The average contribution of R95pTOT ranges from about 21% to 35% according to satellite products (Fig. 7). Unlike previous indices, the greatest values are observed inland. The coastal areas, the northeast and southeast, which represent hotspots of extreme event according to R95p and R99p, present a higher number of rainy events that make the contribution of intense events to total precipitation relatively lower. Figures 7a and 7b show a similar performance of CHIRPS and PERSIANN capturing R95pTOT. On the other
hand, both IMERG and CMORPH show the highest values and also similar results. As expected, the above translates into similar random errors according to RMSE (Figs. 7e–h). Moreover, the deviation from the observations (Figs. 7i–l) shows the systematic underestimation/overestimation obtained with CHIRPS/PERSIANN and IMERG/CMORPH. These results agree with the skills of these products in representing both total JJAS rainfall and R95p: similar values of R95pTOT for IMERG and CMORPH are the result of the overestimation of R95p, while JJAS total rainfall does not differ significantly from CHIRPS and PERSIANN products. Likewise, the contribution from the top 1% rainy days to JJAS rainfall, represented by R99pTOT, is shown in Fig. 8. Although not substantial differences are observed for the four products in terms of climatological values, the separation of CHIRPS/PERSIANN from IMERG/CMORPH is clear. CHIRPS and PERSIANN show a fairly homogeneous range of values, ranging from 5% to 10%. IMERG and CMORPH reproduce a higher contribution of R99pTOT up to 14% with a similar spatial distribution, and a RMSE of similar range than the index. The same shift from a generalized underestimation to overestimation by CHIRPS/PERSIANN and IMERG/CMORPH is observed. Although flat terrain is the prominent geographical feature of Bangladesh, the biases of the mountain areas in the southeast can be explained by small-scale topographic factors. This region presents higher biases, which might be associated with the smoothing effect of the spatial aggregation of satellite data. Moreover, rain gauges over complex terrain are usually installed at the bottom of the valleys, which can lead to discrepancies between measured and estimated precipitation (e.g., Ouyang et al. 2020).

Results from the absolute threshold index R20mm, displayed in Fig. 9, show a high correspondence with the JJAS total precipitation maps of Fig. 2. In this way, the northeast region, and part of the southeast, presents the highest number of rainy days. Interestingly, the highest values of R20mm are related to the spatial resolution of the product. Most notably, R20mm during the monsoon season for

![Fig. 7. (a)–(d) Climatology (2000–19) of R95pTOT for the four satellite-based precipitation products, including (e)–(h) RMSE and (i)–(l) Bias.](image-url)
CHIRPS, the highest-resolution product, can reach up to 65 days, followed by IMERG (52 days), PERSIANN (50 days), and CMORPH (35 days). The latter accounts for the smoothing effect as the size of the grid decreases, which appears as relevant for the monitoring extreme events and applications such as flood modeling and analysis. Regarding the random error represented by RMSE (Figs. 9e–g), it is clearly observed that IMERG provides the best results, with an average RMSE of 5 days at the country level, followed by CHIRPS, CMORPH, and PERSIANN, which present a similar performance (RMSE \( \approx 10 \) days). The Bias estimator (Figs. 9i–l) shows a spatial distribution of values that differs from the previous indices. First, it is observed that in this case both CHIRPS and PERSIANN show are positively biased, indicating a general overestimation of R20mm throughout the country, contrary to that observed for the indices based on the 95% and 99% percentiles. On the other hand, while CMORPH shows an underestimation of the number of very rainy days R20mm, IMERG presents a Bias that varies from negative to positive from south to north, ranging from \(-6\) to 8 days.

The Taylor diagrams shown in Fig. 10 summarize the interannual country-average performance of the four satellite rainfall products in representing both total JJAS rainfall and the five selected indices. According to these diagrams, IMERG performs better representing country-level JJAS total rainfall, R95p, R95pTOT, and R20mm, being R20mm for which the greatest differences between satellite products are observed, especially in RMSE and along the correlation axis (Fig. 10f). Notwithstanding, CHIRPS, PERSIANN and IMERG present a similar performance representing JJAS rainfall, R95p, R99p, R95pTOT, and R99pTOT. For these cases, a relatively high correlation is observed in JJAS rainfall and R95p, which decreases for R99p, R95pTOT and R99pTOT. The poorest overall performance of the satellite products is observed for R95pTOT and R99pTOT in terms of correlation and error (RMSE), indices that incorporate both a percentile and total rainfall of the calculation, which makes its estimation more
difficult. In terms of products, CMORPH clearly had the poorest performance, which is mainly characterized by high RMSE values and low correlation with observations from station data. These results summarize the previous findings and allow to highlight that the systematic errors found in the performance of the selected satellite products strongly depend on the extreme rainfall index selected, since, for example, IMERG would be the most suitable for indices based on fixed thresholds, such as R20mm, but this is not as clear when comparing IMERG, CHIRPS, and PERSIANN, which show similar overall performance for the other indices. This is relevant considering that the four products evaluated have different spatial resolutions, from $0.05^\circ \times 0.05^\circ$ to $0.25^\circ \times 0.25^\circ$.

4. Summary and conclusions

In this study, four satellite-derived precipitation products were evaluated in terms of their performance representing the variability of five indices describing intense rainfall events in Bangladesh. An improved understanding of and methods to assess intense rainfall events are important both for research and in the management of weather extremes that can cause losses and damage in agriculture and other sensitive sectors. We utilized 31 rain gauges suitably distributed throughout the country for this assessment. The analysis centered in climatological features using RMSE and Bias as the main error metrics and on interannual trends.

We found important relative errors in the products including their ability to reproduce general features such as total precipitation during the rainy season. These products are generated using different input data and algorithms, which result in important differences in their performance when compared to precipitation data collected from ground monitoring. In this sense, CHIRPS incorporates a time-variable number of daily ground-based observations in its algorithm. However, given that this number varies from one year to another, time-variant performance can be also expected. Similarly, PERSIANN, TRMM, and CMORPH incorporate
GPCP and GPCC observations, a product that is homogeneous but generated on a monthly scale.

The spatial distribution pattern of total JJAS is preserved when intense precipitation indices are calculated, and the four satellite products are able to capture the spatial distribution of the indices. Nevertheless, their performance is highly variable in terms of values and associated errors with respect to ground observations. First, there is a poorer representation of the indices over coastal and high precipitation areas, which can be associated with the complexity of retrieving precipitation from satellites (sensors and algorithms) over these areas that are also crucial for the management of climate hazards. Interestingly, the representation of total rainfall during the rainy season is similar for CHIRPS, PERSIANN, and IMERG, despite having considerably different spatial resolutions, which is relevant, for example, in hydrological or crop model forcing and the assessment of impacts of extreme events. The latter suggest the importance of the algorithms over coastal and high precipitation areas, which can be associated with the complexity of retrieving precipitation from satellites (sensors and algorithms) over these areas that are also crucial for the management of climate hazards.

FIG. 10. Taylor diagrams for the interannual (2000–19) country-averaged time series of (a) total JJAS precipitation, (b) R95p, (c) R99p, (d) R95pTOT, (e) R99pTOT, and (f) R20mm, for the four satellite-based precipitation products. RMSE is the root-mean-square error. See Table 2 for units.
rainfall indices is variable. Furthermore, although results suggest that these three products adequately capture the variability in indices based on percentiles, IMERG seems to be better suited to estimate events above a fixed daily precipitation threshold in Bangladesh. Additionally, our results suggest that the use of algorithms to correct biases in satellite precipitation using ground observations could be a very useful option, as has been demonstrated for other regions such as eastern Africa (e.g., Dinku et al. 2018), but has been little explored in the case of extreme events.

Acknowledgments. This research was carried out under the United States Agency for International Development (USAID) funded Climate Services for Resilient Development (CSRDF) Project in South Asia and the USAID and Bill and Melinda Gates Foundation (BMGF) supported Cereal Systems Initiative for South Asia (CSISA). CSRDF is aligned with the CGIAR Research Program on Climate Change, Agriculture, and Food Security (CCAFS) Flagship on Climate Services and Safety Nets. The results of this research do not necessarily reflect the views of USAID, the United States Government, BMGF, or CCAFS, and shall not be used for advertising purposes. This research was carried out under the United States Agency for International Development (USAID) funded Climate Services for Resilient Development (CSRDF) Project in South Asia and the USAID and Bill and Melinda Gates Foundation (BMGF) supported Cereal Systems Initiative for South Asia (CSISA). CSRDF is aligned with the CGIAR Research Program on Climate Change, Agriculture, and Food Security (CCAFS) Flagship on Climate Services and Safety Nets. The results of this research do not necessarily reflect the views of USAID, the United States Government, BMGF, or CCAFS, and shall not be used for advertising purposes. The authors gratefully acknowledge the Bangladesh Meteorological Department for providing rainfall data. The valuable comments from three anonymous reviewers are deeply acknowledged.

REFERENCES
Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation