



Evaluating the IMERG precipitation satellite product to derive intensity-duration-frequency curves in Colombia

Evaluación del producto satelital de precipitación IMERG para derivar curvas intensidad-duración-frecuencia en Colombia

Erasmus Rodríguez ¹, Camila García-Echeverri ¹, Ana González ¹, John Sandoval ¹, Manuel Patarroyo-González ¹, Daniela Estefanía Agudelo-Duque ¹

¹Departamento de Ingeniería Civil y Agrícola, Facultad de Ingeniería, Universidad Nacional de Colombia. Carrera 45 # 26-85. C. P 111321. Bogotá, Colombia.

CITE THIS ARTICLE AS:

E. Rodríguez, C.
García-Echeverri, A. González,
J. Sandoval, M.
Patarroyo-González, D. E.
Agudelo-Duque, "Evaluating
the IMERG precipitation
satellite product to derive
intensity-duration-frequency
curves in Colombia", *Revista
Facultad de Ingeniería
Universidad de Antioquia*, no.
110, pp. 31-47, Jan-Mar, 2024
[Online]. Available: <https://www.doi.org/10.17533/udea.redin.20230212>

ARTICLE INFO:

Received: April 12, 2021
Accepted: February 06, 2023
Available online: February 07,
2023

KEYWORDS:

Precipitation; Remote Sensing;
Colombia

Precipitación; Sensoramiento
Remoto; Colombia

ABSTRACT: This article explores the potentialities of the IMERG V06B FINAL product for estimating Intensity-Duration-Frequency curves in Colombia, using the in-situ data available for 110 rain gauges. From observed data for 76 of these stations, we validated the satellite IMERG precipitation data for the period 2001-2019, at daily, monthly, and annual resolutions. For 60 stations, better results were obtained for the monthly time aggregation, followed by the yearly and daily scales, suggesting that seasonality is the main rainfall characteristic captured by the product. Concerning the occurrence of daily precipitation, results indicate that both the probability of detection and the probability of false detection are high. In general terms, the comparison between intensities from existing IDF curves and those derived from IMERG showed underestimations of the rainfall intensities for the short durations studied (0.5 and 1 h), with mean relative errors in the range [-69%,+56%], and overestimations for the large durations of 2 and 6 h, with mean relative errors in the range [-61%,+171%]. Results also suggest that the IMERG product at this moment is not able to capture the daily rainfall distribution in most of the stations. Nevertheless, for almost 20% of the rain gauges, located mainly in the Amazon, Orinoco, and Pacific Regions, the analysis showed that the maximum intensities derived from IMERG are within +/-25% relative error, compared with the ones calculated using the traditional approach.

RESUMEN: El artículo investiga las potencialidades del producto de precipitación satelital IMERG V06B FINAL para estimar intensidades de curvas intensidad-duración-frecuencia en Colombia, a partir de información de 110 estaciones. Utilizando datos para 76 de estas estaciones en el período 2001-2019, se validó el producto IMERG a nivel diario, mensual y anual. En 60 de estas estaciones se ha encontrado que el producto obtiene mejores resultados en la escala mensual, seguida de las escalas anual y diaria, lo cual indica que la estacionalidad es la principal característica de la lluvia adecuadamente capturada por el producto. En general, la comparación entre curvas IDF existentes y las derivadas a partir de IMERG mostró subestimaciones para las intensidades de precipitación correspondientes a las duraciones más cortas (0.5 y 1 h), con errores relativos medios en el rango [-69%, +56%], y sobreestimaciones para las duraciones más grandes de 2 y 6 h, con errores medios relativos en el rango [-61%, +171%]. Los resultados obtenidos sugieren que el producto IMERG en la actualidad no es capaz de capturar la distribución de la lluvia diaria en la mayoría de las estaciones. A pesar de esto, para cerca del 20% de las estaciones, ubicadas principalmente en las regiones Orinoquía, Amazonía y Pacífica, los análisis muestran que las intensidades

* Corresponding author: Erasmo Rodríguez

E-mail: erodriguezs@unal.edu.co

ISSN 0120-6230

e-ISSN 2422-2844



máximas derivadas de IMERG se encuentran dentro de +/-25% de error relativo, en comparación con las deducidas utilizando el método tradicional.

1. Introduction

Intensity-Duration-Frequency (IDF) curves summarize the extreme rainfall characteristics for short time durations at a point in space [1]. They are a fundamental input for hydrological design as they provide data for estimating flows in small ungauged basins and for constructing design storms in medium and large-size watersheds [2]; they are also a basis for flood risk assessments in mountainous areas [3] and for studying rainfall triggering landslides [4, 5] among others. They are calculated from rainfall data using either manual, semi-automatic or automatic procedures that extract information from pluviograph strips or digital records [6]. The traditional parametric method for constructing IDF curves is based on a statistical analysis of annual maximum series [7, 8]. The method consists of calculating annual maximum intensities for different durations (usually 15, 30, 60, 120, and 360 minutes), followed by the fitting of a probability distribution that usually corresponds to one from the GEV (Generalized Extreme Value) family [9, 10]. This distribution is used to calculate intensities for selected return periods (usually 2, 5, 10, 25, 50, and 100 years) [11–13]. When pluviograph information is not available, as usual in several parts of the world, synthetic or generalized IDF curves can be estimated from rainfall information using regionalization, scaling, or other methods based on rainfall characteristics [14–17].

Recent products that derive rainfall intensity information from meteorological satellite observations complement the in-situ data and are beginning to be used for several purposes, including, for example, the validation and intercomparison of satellite rainfall estimates [18], the analysis of flash floods [19, 20], hydrological modeling [21–25], downscaling of rainfall extremes [26], and the estimation of IDF curves [27–30]. Among these products is the Integrated MultisatellitE Retrievals from GPM-IMERG (or in short IMERG) produced by NASA, a quasi-global dataset (60°N - 60°S), calibrated using monthly rain gauge precipitation values, which provides rainfall intensities at 0.5 h intervals, with a spatial resolution of 0.1° (approximately 11 km at the equator). As part of the GPM-era satellites [31], IMERG is the successor of the former GPM-based rainfall estimates v3, v4, and v5, which use a multi-satellite fusion algorithm to generate precipitation data. The latest IMERG V06B product, additionally incorporates data coming from the TRMM mission, extending the period with available data back to 2001. As the GPM mission, TRMM was a research satellite mission launched in 1997, developed

by NASA and JAXA, that came to an end in 2015. Using the TRMM multi-satellite precipitation analysis (TMPA), gridded products at different spatiotemporal scales were developed by NASA and JAXA, including versions 3B42 (monthly) and 3B43 (daily), both with a resolution of 1°, which are the TMPA satellite products usually assessed for hydrometeorological purposes.

We have selected the IMERG V06B product for conducting this research based on four considerations: i) IMERG has been applied and validated in a vast number of regions around the world [32], showing that in several areas, the product overestimates rainfall in lowlands and underestimates rainfall at high altitudes. Evaluations of the IMERG product in South American countries like Brazil [33–35], Chile [27, 36] Colombia [37], and Ecuador [38], have shown that the IMERG product correctly represents the spatial pattern of orographic precipitation in the tropical Andes; that IMERG captures nicely the spatial and temporal rainfall characteristics in the rainiest region in Colombia, and that better performances of the product are obtained at regional rather than at grid scales ii) Comparisons of IMERG with other precipitation products such as TRMM [37], TMPA [38] or GSMaP [39] in regions of South America suggest that both IMERG and TRMM perform well in the Choco area in Colombia. Regarding TMPA, results indicate that IMERG is a better product than TMPA in detecting and quantifying rainfall in areas in Ecuador and Peru, especially in the high Andes, and that IMERG shows a better estimation of in-situ observed rainfall than TMPA. Concerning GSMaP, evaluations in a flat area in Brazil indicate that IMERG outperforms GSMaP at annual and monthly scales, but it is slightly worse at daily resolution iii) Compared with other satellite rainfall products used to derive IDF curves, IMERG outperforms TRMM and GSMaP [40] iv) IMERG represents the state of the art of freely available homogeneous and consistent rainfall products with temporal resolutions compatible with those required when constructing IDF curves.

Various studies, conducted with different versions of the IMERG products and in different parts of the world, have evaluated and reported its use. A study in southern Austria [41] describes the evaluation of the EARLY, LATE and FINAL IMERG V03 products for 2014. They found that the rainfall IMERG FINAL estimations outperformed the EARLY and LATE products and that the IMERG FINAL half-hourly estimates correspond approximately to 25 minutes of observed accumulations, with an offset of about +40 minutes. Another study in Iran [42] describes the evaluation of the IMERG dataset for the period 2014-2017 using data from 370 stations, reporting that the best performance is obtained with the FINAL product and the worst one with the EARLY product. On the island of Bali in Indonesia, a study [43] tested, at multiple time scales, the

use of the IMERG FINAL product for the period 2014-2019. They found that, in general, rainfall estimations with IMERG were lower than observations. Besides this, IMERG estimations were highly accurate at monthly and seasonal scales, but less exact at daily time resolution, although the probability of detection of rainfall events was correct. In Spain [44] were investigated the effects of orographic biases of IMERG V06B-FINAL estimates in the Ebro River basin for the period 2014-2018. Results showed that the performance of IMERG depends highly on altitude (with discrepancies increasing with height), on the precipitation regime, and on the rainfall gauge density considered to create the observed gridded product, used as ground validation data.

A comprehensive study in Canada [45] made an evaluation of the IMERG V06B FINAL product for the period 2014-2018. They found that over the coastal eastern and western parts of the country, IMERG tends to overestimate the hourly precipitation intensities by approximately 25% and that the discrepancies with ground truth data are larger in areas with larger rainfall. Overall, the results of this study showed that IMERG is a good product for investigating rainfall at high spatiotemporal resolutions. Best performances occurred in the plain areas, with lower uncertainties during warm months. Researchers in Pakistan [46] investigated the performance of the IMERG products using 62 rain gauge stations and found that on a daily scale, the FINAL product performs well in terms of frequency and intensity and that this product captures relatively well extreme precipitation events. At monthly time resolution, they found that the IMERG FINAL product produces good results in plain and medium elevation regions but has limitations at higher altitudes.

Additional studies have investigated the performance of IMERG and/or TMPA products in Colombia. In the western part of Colombia, researchers evaluated the performance of the IMERG V06B product for the period 2014-2017 over complex terrain in the Chocó Region by using daily time series of rainfall from 185 stations [37]. The results showed that IMERG represents well the spatial and temporal variations in the mean daily precipitation in the study area. Overestimations appeared for rainfall in the relatively low precipitation and medium-to-high altitude areas, and underestimations for mean daily precipitation in areas with very high precipitation and medium-to-low altitude. A study for the whole country, analyzed the performance of the TRMM 3B43 V7 precipitation product for the 1998-2015 period, at a monthly scale, using 1.180 rain gauge stations [47]. Their findings suggest that the product performs well in the plain areas of the Amazon, Orinoco, and Caribbean regions. Over the complex relief of the Andes region, the product tends to overestimate precipitation, while in the wet Pacific region, precipitation

is largely underestimated. Also, the performance of TRMM 3B43 V7 decreases during wet seasons. Another finding of this research indicates that the product frequently misses light rainfall events and less frequent but very heavy storms, which causes rainfall overestimations in the Andes region and underestimations in the Pacific region.

Several studies described above have shown that sub-daily rainfall information is the basis for calculating IDF curves using a traditional approach [1, 10, 17]. However, the development of IDF curves continues to be a challenge in the whole world, especially in countries from the Global South that have limitations, if not lack of this type of information [48]. Additionally, in Colombia and several parts of the world, the number of rainfall stations providing this information has decreased since the 1980s [24]. Considering that IDF curves are the main input for the hydrological design of hydraulic structures and that they represent an important investment for the governments, they need to be rigorously constructed and updated continuously, as they are derived from extreme rainfall that are non-stationary [10]. Constructing and updating IDF curves from analog information, like in Colombia, is a time-consuming and costly process [17]. Although satellite rainfall data may be used to develop IDF curves, few studies have reported this. Given all these reasons, the aim of this study is to contribute to the knowledge of the use of satellite-based information to construct IDF curves, bridging some of these gaps [26-30].

Accordingly, in this study, we first validated at annual, monthly, and daily resolutions the IMERG V06B FINAL product using ground validation data for 76 rain gauge stations in Colombia for the period 2001-2019, using a pixel-point methodology and calculating thirteen different statistical and contingency metrics. Then, we estimated and validated the maximum intensities of IDF curves at 110 stations obtained from the IMERG V06B FINAL product, using the corresponding IDF curves at the same locations, constructed with the traditional method and in-situ data [17].

2. Study area

This study uses IDF curves, calculated up to the year 2010 using the traditional approach, from 110 stations in Colombia. Due to daily rainfall information availability, only 76 out of the 110 rain gauges have been used for evaluating the IMERG V06B product. Figure 1a shows the spatial distribution of the stations, which is heterogeneous within the country, with most of the gauges (70%) located in the central (Andes), 14% in the northern region (Caribbean), 9% in the eastern Orinoco region, 4% in the Amazon southeastern region, and 3% in the Pacific western region.

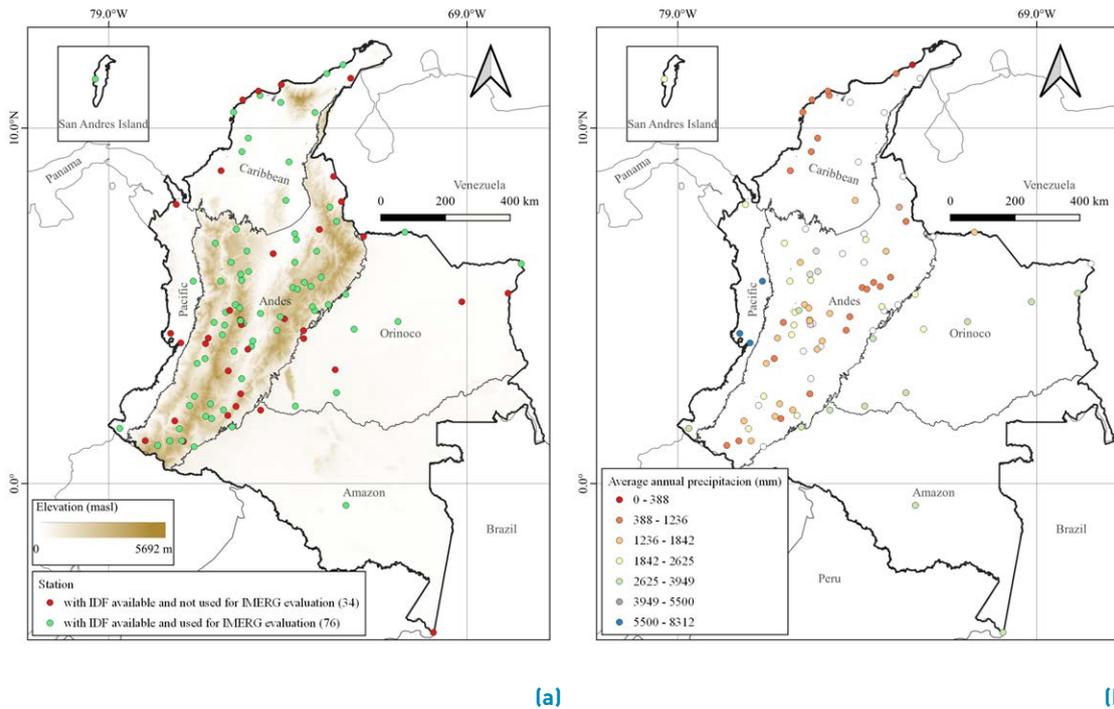


Figure 1 (a) Spatial distribution of the rain gauges investigated within the five different regions in Colombia. (b) Mean annual precipitation for each station

Interactions between several macroscale and mesoscale meteorological phenomena, the Pacific, Atlantic, and Caribbean Oceans, the Andes ranges and the Amazon rainforest result in a complex spatiotemporal precipitation pattern in Colombia [49]. On the interannual scale, Colombia's dominant driver of climate and weather variability is the El Niño Southern Oscillation (ENSO), which brings rainfall above/below normal during the La Niña/El Niño phases [50, 51]. On the annual scale, the movement of the Intertropical Convergence Zone (ITCZ) exerts a strong control on the rainfall seasonality in Colombia, with a bimodal rainfall regime in the Andes region, and unimodal rainfall regimes in the Caribbean and Amazon regions [49, 52]. Precipitation in the Pacific region is mainly associated with a low-level westerly jet (Chocó jet) that brings rainfall most of the year, following a unimodal precipitation pattern [51].

In the Orinoco region, rainfall is produced mainly by the South American low-level jet that brings Mesoscale Convective Systems (MCS) from the Atlantic Ocean and the Amazon rainforest, with a rainfall pattern of high precipitation during the middle part of the year [49]. Table 1 summarizes the rainfall characteristics of Colombia's five major natural regions. The complex terrain and diverse rainfall climatology make the analysis of precipitation in the study area challenging with either

point or gridded observed data or satellite products such as IMERG.

3. Data

3.1 Ground-based observations

Daily time series of precipitation for 76 rain gauge stations in Colombia, during the period 2001-2019, recorded by the *Instituto de Hidrología, Meteorología y Asuntos Ambientales* (IDEAM) were obtained from the hydrometeorological data portal DHIME [53]. The list of stations used in this study, their characteristics, and a summary of the statistical and contingency criteria calculated are included in the supplementary material (see supplementary material). Rainfall intensity Tables and Equations for the IDF curves for 110 rain gauges were also obtained from IDEAM [17].

3.2 IMERG data

EARLY and LATE (near real-time data) and FINAL (post-processed data) IMERG products are freely available online. IMERG EARLY, LATE, and FINAL datasets are computed every 4 hours, 14 hours, and 3.5 months after observation time, respectively. EARLY and LATE datasets are computed an hour at a time, while FINAL is computed a month at a time. According to the IMERG developers

Table 1 Rainfall characteristics of the five major natural regions in Colombia

Region	Code	Elevation Ranges (MASL)	Main Climatic Drivers	Rainfall Regime	Annual Rainfall (mm)	# Gauges with IDF /daily rainfall data
Caribbean	1	0-1400	ITCZ / Tropical Storms/ ENSO/ MCS	Unimodal MJJ SON	1500	16/11
Andes	2	50-5700	ITCZ/Chocó Jet ENSO/ MCS	Bimodal MAM OND	1500	72/53
Orinoco	3	100-250	ITCZ MCS	Unimodal MJJA	2800	12/7
Amazon	4	100-300	ITCZ MCS	Unimodal DJF MAM	3500	4/3
Pacific	5	0-700	Chocó Jet ITCZ ENSO/MCS	Unimodal AMJJASOND	5500	5/2

and several authors, the FINAL product provides the most suitable estimations for research purposes [31].

IMERG V06B FINAL time series of rainfall intensity at 0.5 h intervals in (mm/h) and at a resolution of 0.10° (≈ 11 km) were downloaded for the period 2001-2019 from the Giovanni NASA repository [54]. For 10 stations, the EARLY, LATE, and FINAL IMERG products were also downloaded for the same period, but initial validations of the three products (results are included in the supplementary material) showed that the FINAL product systematically outperformed the other two, as reported by several authors. For this reason, in the analysis reported here, only the FINAL product was considered.

3.3 Quality control and data preparation

For the IDEAM rain gauge daily data, missing information (blank data) was replaced by N.A. Stations with more than 25% of missing data, or identified outliers, were discarded from the analysis. In total, 76 stations were considered in the validation of the IMERG product. IMERG data in GMT, were converted to GMT-5 in order to obtain IMERG estimations in local time. Then IMERG data were converted from intensity to rainfall depth and then aggregated for validation, at daily, monthly, and yearly temporal resolutions, using an algorithm developed in R.

3.4 IDEAM IDF Curves

For each of the 110 rain gauges studied, there are IDF curves available, updated to the year 2010, with different starting years, as early as 1972 [17]. The development of these updated IDF curves was based on data of mass curves of precipitation for selected events at 1-min resolution, covering either 24 h or one week, depending on

the type of pluviograph strip chart available. We compared maximum intensities for selected durations and return periods coming from IDEAM with those estimated from IMERG data. Perhaps, different periods for comparison can induce some differences in the curves that were not considered in the analysis. However, UNAL [17] reports that when updating the IDF curves up to the year 2010, the differences in rainfall intensities for the new period (updated) and the ones derived using the old period are almost negligible, except for the short 15-minute duration. Due to the IMERG temporal resolution (30 minutes), the 15 minutes duration was not investigated; then, we assumed here that there are no significant differences in IDF curves, associated with the different periods used in their construction.

4. Methodology

4.1 Ground Validation

Similar studies where satellite products such as IMERG, TMPA, APHRODITE, and PERSIANN-CDR are verified, have adopted validation metrics and detection rates such as Pearson's correlation coefficient (r), mean absolute error (MAE), root-mean-square error (RMSE), Nash-Sutcliffe efficiency (NSE), relative bias (bias), relative error (RE), and probability of detection (POD), false alarm ratio (FAR), equitable threat score (ETS), frequency bias index (FBI), critical success index (CSI), respectively [26, 55-59]. These metrics have been key for the verification of the potential of satellite precipitation products. Validation of the IMERG V06B FINAL product was performed on a pixel-point basis by comparing the IMERG gridded data with in-situ information (IDEAM) at three different time resolutions: daily, monthly, and yearly. Although several studies

have reported on the use of the pixel-point approach [37, 45, 47], it is clear that a comparison between the satellite gridded information and the rain gauge data is not direct, as the properties of rainfall change with the spatial scale investigated. In this regard, a study in the USA [26] proposes a downscaling procedure to tackle the problem of scale disparity between the two sources of information.

Equations 1 to 5 show the five statistical metrics used in the validation process to investigate the magnitude of the error, the correlation, and the skill of the IMERG product to reproduce the observed rainfall information in each rain gauge station. The metrics include the normalized root mean square error (nRMSE), used to describe the error, with an optimum value equal to zero; the dimensionless Pearson correlation coefficient (r) that establishes the relationship between the covariance and variance of the IMERG estimation and the IDEAM data, with an optimum value equal to 1; the normalized mean absolute error (nMAE), that describes the magnitude of the overestimation or underestimation of the IMERG data, with a perfect value equal to 0; the bias, in percentage, that represents the magnitude of the underestimation or overestimation of the satellite data, with an optimum value equal to 0; and the dimensionless Nash-Sutcliffe efficiency (NSE) with a perfect value equal to the unity, used to determine the relative magnitude of the residual variance compared to the measured data variance. We have used the normalized statistics for RMSE and MAE to facilitate comparisons between stations with large precipitation variability.

To support the analysis, we produced maps for each temporal resolution and statistical criteria. Additionally, for each rain gauge and both the IMERG and the IDEAM datasets, we made comparisons for the mean rainfall at daily and monthly intervals, during the period 2001-2019. Besides, to study the variability of the metrics among regions, we also created boxplots.

$$nRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (I_{IMERG_i} - I_{IDEAM_i})^2}}{\sum_{i=1}^N I_{IDEAM_i}} \quad (1)$$

$$r = \frac{N \sum_{i=1}^N (I_{IMERG_i} * I_{IDEAM_i}) - \sqrt{N \sum_{i=1}^N I_{IMERG_i}^2 - \left(\sum_{i=1}^N I_{IMERG_i}\right)^2} \sqrt{N \sum_{i=1}^N I_{IDEAM_i}^2 - \left(\sum_{i=1}^N I_{IDEAM_i}\right)^2}}{\sum_{i=1}^N I_{IMERG_i} * \sum_{i=1}^N I_{IDEAM_i}} \quad (2)$$

$$nMAE = \frac{\frac{1}{N} \sum_{i=1}^N |I_{IMERG_i} - I_{IDEAM_i}|}{\sum_{i=1}^N I_{IDEAM_i}} \quad (3)$$

$$\text{bias}(\%) = \frac{\frac{1}{N} \sum_{i=1}^N (I_{IMERG_i} - I_{IDEAM_i})}{\sum_{i=1}^N I_{IDEAM_i}} * 100 \quad (4)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (I_{IMERG_i} - I_{IDEAM_i})^2}{\sum_{i=1}^N (I_{IDEAM_i} - \overline{I_{IDEAM}})^2} \quad (5)$$

At daily resolution, we calculated four dimensionless contingency metrics to identify the ability of the IMERG product to detect the occurrence of rainfall. Description and application of these categorical or contingency statistics for comparing satellite-based precipitation with observed data have been given in many references [37, 42, 43, 45]. These criteria, originally proposed in two studies [60, 61], are shown in Equations 6 to 9 and include the Probability of Detection (POD), the Probability of False Detection (POFD), the False Alarm Ratio (FAR) and the Critical Success Index (CSI). In Equations 6 to 9, hits represent the occurrence of daily rainfall in both, IMERG and IDEAM datasets; misses represent daily rainfall observed by IDEAM but not detected by the IMERG dataset; false alarm represents daily rainfall detected by IMERG but not observed by IDEAM, and correct negatives are cases where neither of the two products detects rainfall. The results for the four contingency criteria are dimensionless and vary between 0 and 1. For POD and CSI, the perfect score is 1 and for POFD and FAR, it is 0. For each contingency criterion, we constructed a map, and boxplots classified by regions to facilitate the analysis.

$$POD = \frac{\text{hits}}{\text{hits} + \text{misses}} \quad (6)$$

$$POFD = \frac{\text{false alarm}}{\text{correct negatives} + \text{false alarm}} \quad (7)$$

$$FAR = \frac{\text{false alarm}}{\text{hits} + \text{false alarm}} \quad (8)$$

$$CSI = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarm}} = \frac{1}{\frac{1}{(1-FAR)} + \frac{1}{(1-POD)} - 1} \quad (9)$$

We also produced quantile-quantile plots for the IMERG and the IDEAM datasets to investigate, for each of the 76 rain gauges, how the two rainfall distributions are compared at a daily scale and how the extreme daily events are captured by the IMERG product in each rain gauge station.

Finally, to investigate the explained variance regarding the

statistical, contingency and IDF's error metrics, and reduce the dimensionality of the data, we implemented a Principal Component Analysis (PCA). We also explored relationships of the error metrics with five different proxies, including rain gauge latitude, altitude, geographic region, mean annual rainfall, and precipitation regime.

For this purpose, in the principal components space, we plotted values for the Objective Function (OF) shown in Equation 10 (optimum value 0), which includes each of the metrics described above and the four Relative Errors (ERR REL DUR), calculated as the ratio of the difference between the IMERG intensity and the IDEAM intensity, and the IDEAM intensity, for durations DUR of 30, 60, 120 and 360 min. This approach helped us to identify common physiographic and climatological characteristics of the stations with similar performance levels of the IMERG product, both at validation and at the reproduction of intensities of IDF curves.

4.2 Intensity estimation using IMERG

For each of the 110 rain gauge stations, and during the period 2001-2019, we aggregated the 0.5 h IMERG intensity series at 1, 2, and 6 h. We used no other duration to allow the direct comparison of rainfall intensities from the IDEAM IDF curves constructed at 15, 30, 60, 120, and 360 minutes. Unfortunately, the resolution of the IMERG product allowed no comparisons at 15 minutes. After aggregation of the data, for each rain gauge station and the four durations (0.5, 1, 2, and 6 h), we constructed the annual maximum rainfall intensity time series. Afterward, we performed a frequency analysis using the lmoms R-library by fitting the Gumbel probability distribution and obtaining intensities for the selected return periods (2, 5, 10, 25, and 50 years). We used this probability distribution as it was the one used to construct the IDEAM IDF curves, and it has been reported as the one with the best fit to annual maximum intensities from IMERG data. We calculated no intensities for 100 years as results for this return period are highly uncertain due to the IMERG time-series length of only 19 years (2001-2019).

Rainfall intensities obtained for the selected durations and return periods were compared with the corresponding values from the IDEAM IDF curves using a graphical methodology. Relative errors in (%) for each duration and return period were also calculated using Equation 11. The relative errors for each duration were averaged over all the return periods considered and plotted in maps. To analyze the variability of the relative errors and the other metrics among natural regions, boxplots were also created.

5. Results

5.1 Validation process

Figure 2 shows the spatial distribution of statistical measurements, including nRMSE, r , and nMAE, for the three temporal resolutions investigated (yearly, monthly, and daily). Despite the low number of stations in the Orinoco and Amazon regions, maps in Figure 2 indicate that better results are obtained for stations in these two zones. These regions are characterized by flat terrain. Results for all the statistical criteria but r suggest that most of the stations obtained better results at the yearly scale, followed by monthly and daily resolutions. For r , better results are obtained at the monthly scale, followed by the yearly and daily time resolutions (see box plots in Figure 3). This may be a consequence of the global calibration of the IMERG product at monthly resolution. The bottom panels in Figure 2 depict the contingency measurements POD, FAR, and POFD on a daily scale. In general, POD is high in all stations, FAR is quite variable and PFOD is even more variable. The summary of the results, consolidated in the boxplots in Figure 3, shows mixed tendencies. Some stations located in the Pacific region, such as Aeropuerto El Caraño, obtained an adequate performance of the IMERG product for detecting the occurrence of daily rainfall (POD = 0.94, FAR = 0.06, CSI = 0.88).

$$OF = \frac{(1 - POD) + POFD + FAR + (1 - CSI) + RMSE + |bias| + MAE + (1 - r) + (1 - NSE) + |ERR REL 30| + |ERR REL 60| + |ERR REL 120| + |ERR REL 360|}{13} \quad (10)$$

$$RELATIVE ERROR = ((I_{IMERG} - I_{IDEAM}) * 100 / I_{IDEAM}) \quad (11)$$

Others such as Aeropuerto Benito Salas, located in the Andes Region, obtained acceptable performances (POD = 0.87, FAR = 0.40, CSI = 0.55), and others, such as Manaure, located in the Caribbean Region, had inferior results (POD = 0.74, FAR = 0.70, CSI = 0.27). Overall, results suggest that the occurrence of rainfall is better detected in stations with high precipitation (> 2,400 mm), reasonably well detected in stations with medium precipitation (1,200 - 2,400 mm), and poorly detected in stations with low precipitation (< 1,200 mm).

The supplementary material includes detailed results for each rain gauge station, aggregation scale, and statistical and contingency criteria. Throughout the spatial domain investigated, results emphasize the mixed tendencies. Yet considering the uneven distribution of rain

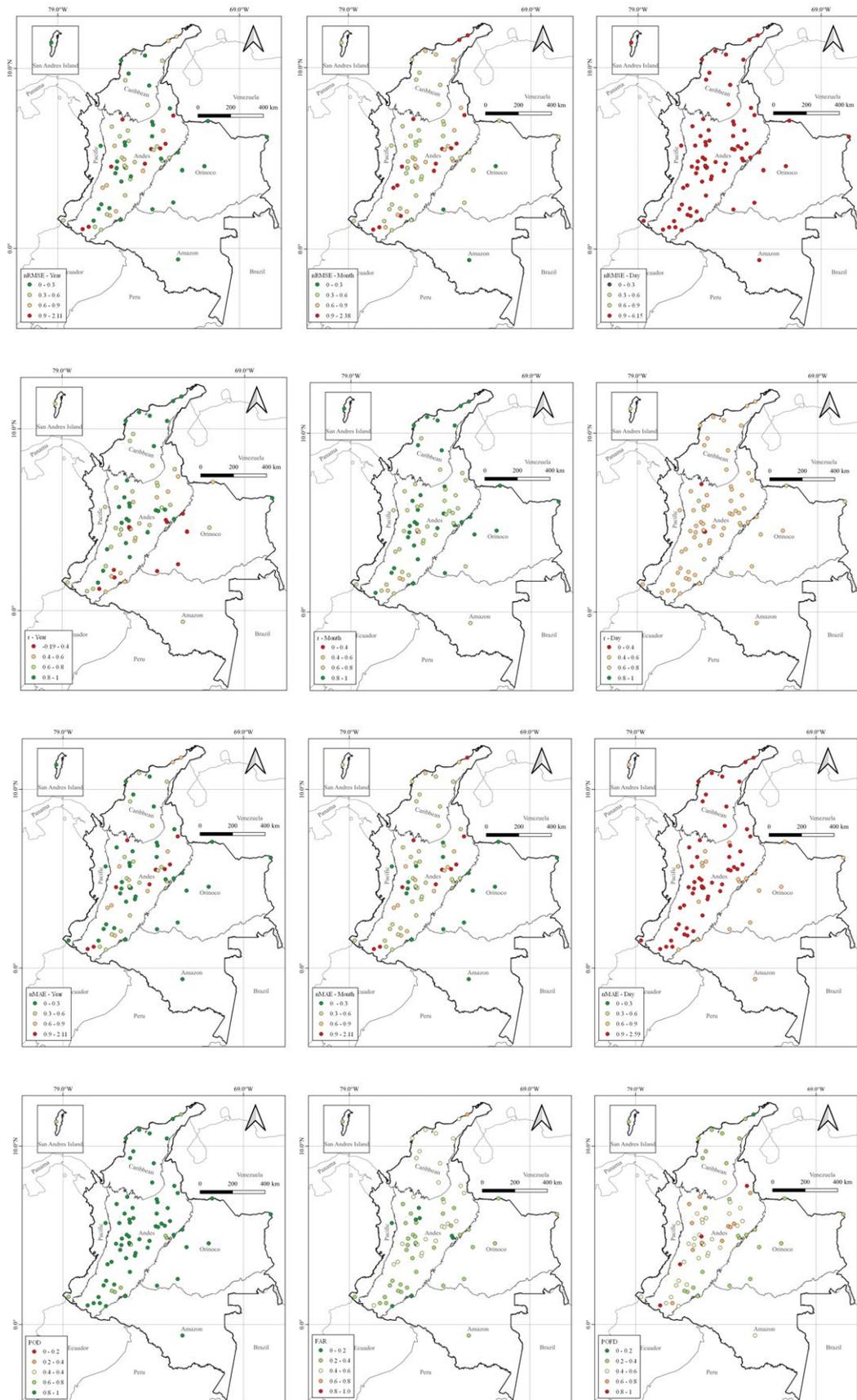


Figure 2 Maps showing the spatial distribution of statistical measurements nRMSE (-), r (-), nMAE (-) at yearly, monthly and daily scales, and contingency metrics POD (-), FAR (-) and POFD (-) calculated at daily time resolution.

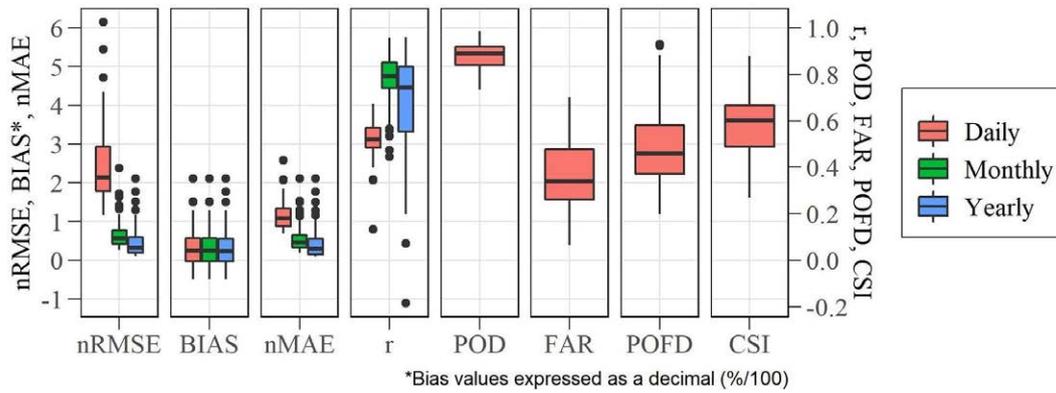


Figure 3 Box plots for the statistical metrics nRMSE, bias, nMAE and r were used to investigate IMERG performance at yearly, monthly and daily time resolutions, and contingency metrics POD, FAR, POFD and CSI calculated at a daily time resolution

Table 2 Summary of the performance of the IMERG product in capturing daily extreme rainfall events in the investigated stations as a result of the quantile-quantile analysis

Region	No. stations	Underestimation %	Overestimation %	Correct %
Caribbean	11	45	36	19
Andes	53	26	48	26
Orinoco	7	29	57	14
Amazon	3	33	67	-
Pacific	2	-	50	50
All	76	29	48	23

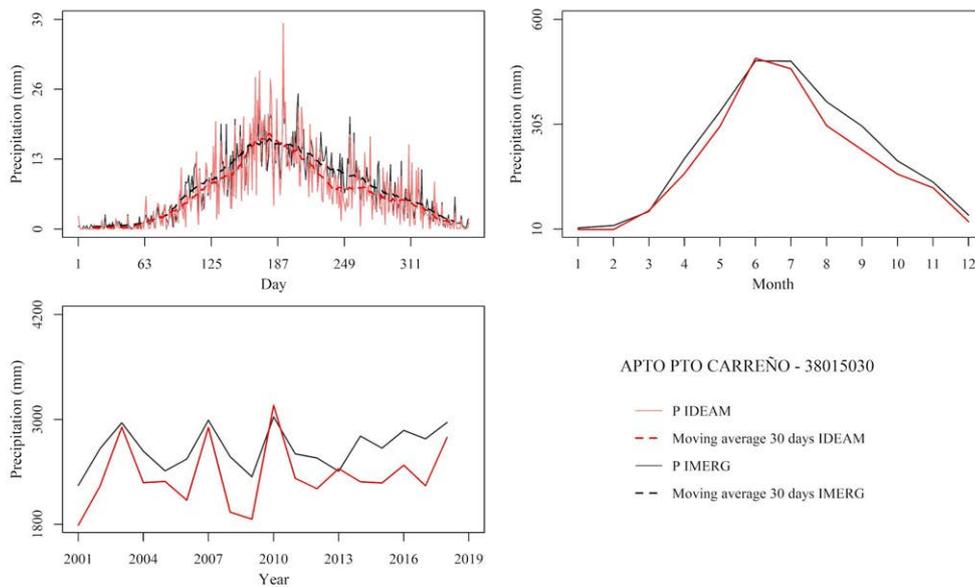


Figure 4 a) Comparison of the mean multiannual daily precipitation, b) the mean monthly precipitation, and c) the yearly precipitation in the Apto Pto Carreño - Orinoco Region rain gauge station for the IMERG and IDEAM datasets

gauge stations in the five regions, in general, the Pearson correlations calculated depend on the time scale of analysis. For stations located in the Orinoco and Amazon regions, Pearson correlation shows high values when the monthly time scale is evaluated; reasonable results are

obtained with annual and monthly evaluations for stations located in the Pacific and Caribbean Regions. In the case of the Andes region, results are mixed between low and high. From the analysis of the r metric at the regional scale (see box plots in the supplementary material) results

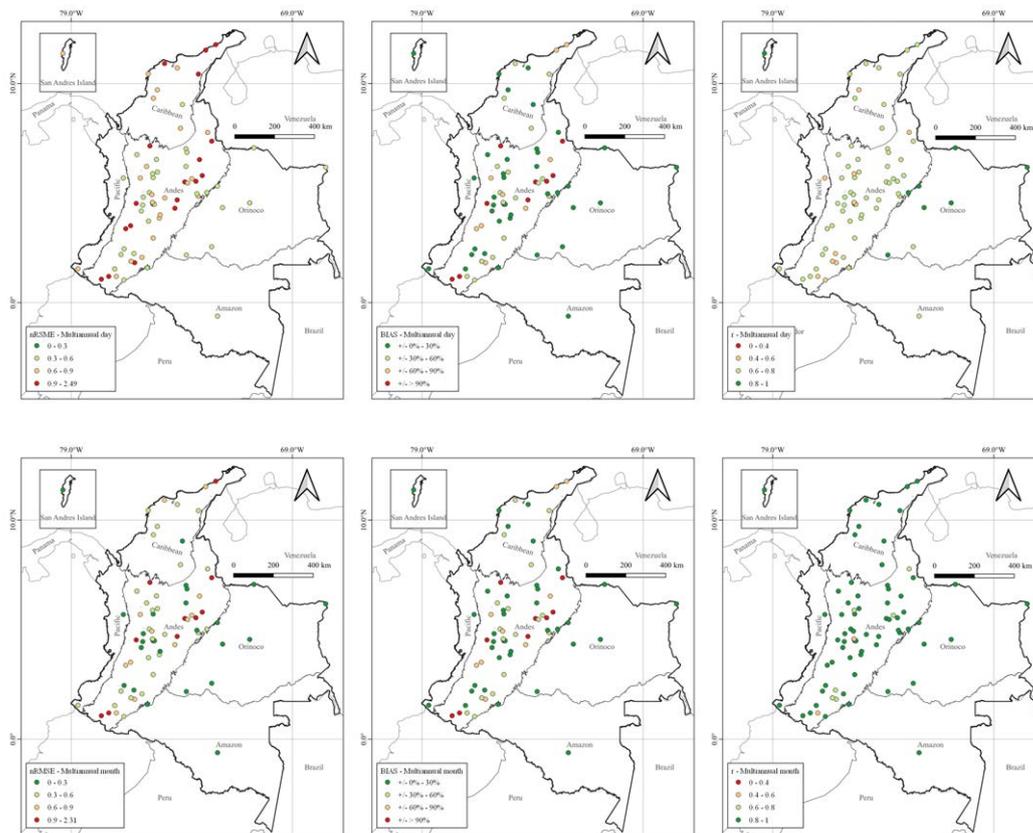


Figure 5 Maps showing the nRMSE (left panels), percent bias (%) (center panels) and the Pearson correlation coefficient (r) (-) (right panels), for the mean multiannual daily precipitation (top panels) and the mean monthly precipitation (bottom panels), as estimated using the IDEAM dataset as ground-validation

show that at the daily time resolution, higher median values are obtained for the Orinoco ($r=0.65$), followed by the Pacific (0.60) and Amazon ($r=0.55$) regions. This is associated with the low number of stations in these three regions, compared to the number of stations in the Caribbean ($r=0.45$) and Andean ($r=0.50$) regions.

Regarding the multiannual analysis conducted for validation of the IMERG product, Figure 4 shows an example (rain gauge station Apto. Pto Carreño - Orinoco) of the mean daily multiannual precipitation, the mean monthly multiannual precipitation, and the yearly rainfall, as calculated using the IMERG and IDEAM datasets. For the mean daily precipitation, 30-day moving averages were also plotted, and showed a very good agreement between both datasets. A detailed observation of Figure 4a shows that for this station, extreme daily rainfall is underestimated, and overall, there is a small overestimation of the multiannual mean daily and monthly rainfall, which is larger at the annual scale. Error metrics are nRMSE = 0.50, bias = + 13.47%, and $r = 0.88$ for the mean daily data, and nRMSE = 0.18, bias = + 13.47%, and $r = 0.99$ for the mean monthly data. These results show

that the IMERG product well captures the seasonality of daily and monthly average rainfall. Results for other rain gauge stations are included in the supplementary material and spatialized in Figure 5, where the percent bias and the Pearson correlation coefficient are plotted for each station and temporal resolutions, showing again mixing tendencies. The best results were obtained for stations in the Orinoco, Amazon, and Pacific Regions. Better results are obtained for the multiannual monthly precipitation, than for the multiannual daily rainfall, thus reinforcing the findings described before during the analysis of the whole time series.

To complement the validation of the performance of the IMERG product at a daily scale, we constructed quantile-quantile plots for all 76 rain gauge stations. Examples of the results are presented in Figure 6 and summarized for all stations in Table 2. Results showed mixed tendencies for the daily extreme rainfall events, with overestimations in most of the cases (48% overall), followed by underestimations (29% overall) and good agreements (23% overall).

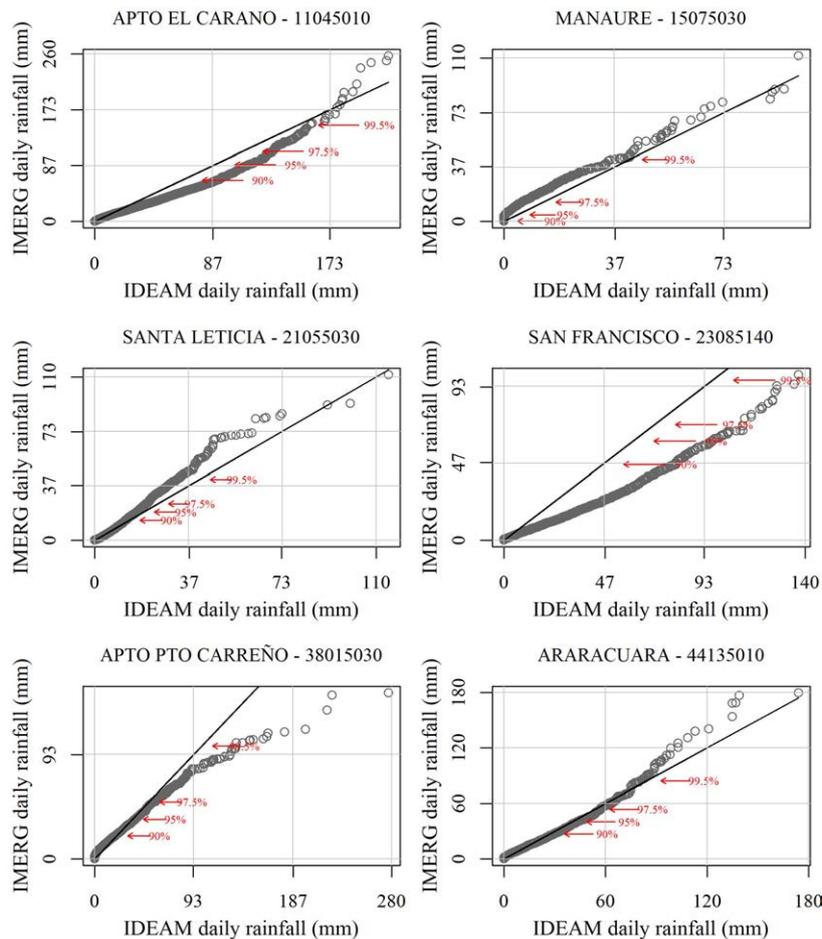


Figure 6 Quantile-quantile plots of daily rainfall (IMERG and IDEAM) for the selected stations. The arrows and numbers show the location of quantiles

As an illustration, in Figure 6, results for station Apto El Caraño - Pacific indicate that daily low precipitation (< 25 mm) is well captured, but there are underestimations for the moderate rain (25 -180 mm) and overestimations of the heavy rains (> 180 mm). For station Manaure - Caribbean, results indicate that light and moderate daily rains are slightly overestimated, while heavy rainfall (> 70 mm) has mixed patterns with under and overestimations. For station Araracuara - Amazon, light and moderate rains (< 70 mm) are well captured, and heavy rains (> 70 mm) are overestimated. For station Apto Pto. Carreño - Orinoco, results indicate that low and moderate rains (< 70 mm) are well captured, and heavy rainfall are underestimated. For station Santa Cecilia in the Andes Region, results show that light rains are well captured, and moderate and heavy rains have mixed patterns.

For the San Francisco rain gauge, also in the Andes region, results show that there is an underestimation of the daily rainfall events for all types of rainfall. Quantile-quantile plots for all 76 rain gauge stations are

included in the supplementary material. Results highlight the limitations of the IMERG product in capturing daily extreme rainfall events in the majority of the cases (only a fifth of them with good agreement). These extreme events are the most used to derive IDF curves and, in this sense, limitations in capturing extreme daily rainfall could be transferred into weaknesses of the IMERG product for reproducing maximum intensities in IDF curves in Colombia.

The selected results of the Principal Component Analysis implemented are shown in Figure 7. Additional results are included in the supplementary material. For the 13-error metrics calculated at a daily scale, the three principal components (PC1, PC2, and PC3) explain in total 83.4% of the data variance, with individual values of 41.3%, 29.5%, and 12.6%, respectively. In the PC1-PC2 space, besides the trivial correlations (CSI is inversely correlated with FAR and POFD; ERR_REL_30, ERR_REL_60, ERR_REL_120, ERR_REL_360 are highly correlated; normalized statistical measurements 1-NSE, nMAE, nRMSE and bias are highly

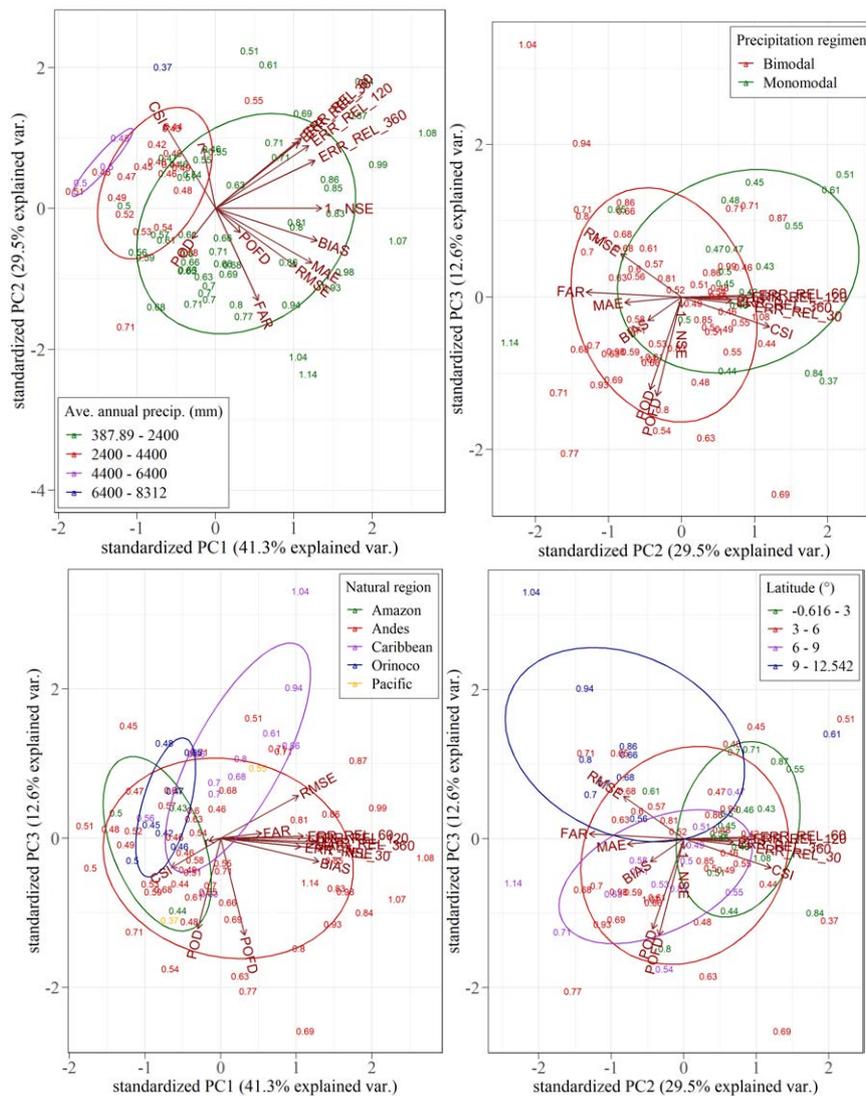


Figure 7 Selected results of the Principal Component Analysis. For each rain gauge station, values in colors indicate the results of the objective function as formulated in Equation 10 for the corresponding proxy investigated (mean annual precipitation, rainfall regime, natural region, latitude, and altitude)

correlated), we found an inverse relationship between POD and the average relative errors in the estimated intensities of IDF curves [ERR_REL 30, ERR_REL_60, ERR_REL_120, ERR_REL 360]. This means that, in general, the best results in the estimation of intensities of IDF curves are obtained for stations with a high probability of detection of daily rainfall (POD), such as El Rancho - Andes (POD=0.97), and poorer results are obtained for rain gauges with lower POD values, such as station Manaure - Caribbean (POD=0.74). In the PC1-PC3 and PC2-PC3 spaces, it is clear that statistical metrics and errors in IDF curve estimations are highly correlated and that CSI and RMSE are inversely correlated. Although counterintuitive, in these two spaces, we found that POD and POFD are highly correlated for stations located in the Orinoco, Amazonas, and Pacific Regions. However, these results were obtained

with very few stations in these three areas, so they cannot be generalized for the whole regions. Additionally, the best performances of the IMERG FINAL product occur for plain areas with annual precipitation larger than 2,400 mm, experiencing a monomodal rainfall regime. We found no correlation between IMERG performance and altitude.

Numbers presented in the PCA spaces in Figure 7 show the values of the objective function formulated in Equation 10, which helped us to investigate relationships between the IMERG performance and physiographic and climatological characteristics. In general, we found that better results are obtained for stations located mainly in the Orinoco, Amazon, and Pacific Regions with a monomodal rainfall regime, and mean annual precipitation above 2,400 mm. Regarding latitude, it seems that IMERG results improve

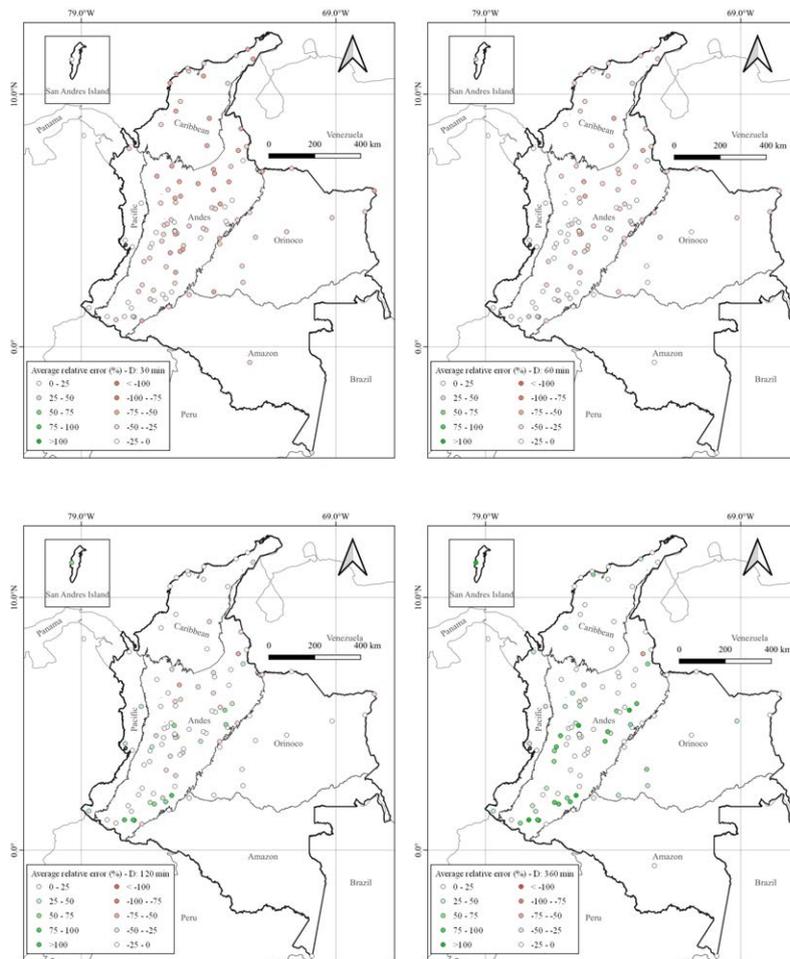


Figure 8 Average relative error [%] $\left(\frac{\text{IMERG} - \text{IDEAM}}{\text{IDEAM}} \times 100\right)$ in rainfall intensities for all return periods (T_r) and for durations 0.5h (top left panel), 1h (top right panel), 2h (bottom left panel), 6h (bottom right panel)

towards the south. No clear relationship was evidenced between the error metrics and the altitude.

5.2 Comparison of intensities of IDF curves

Figure 8 shows for the four durations investigated (0.5, 1, 2, 6 h) the magnitude and sign of the average relative errors for all return periods in the 110 stations. Maps for each duration and return period are included in the supplementary material. Results show that, for short durations (0.5 and 1 h), there is a systematic underestimation of the rainfall intensities in almost all stations and that for large durations (2 and 6 h) the opposite occurs, with a general overestimation of the rainfall intensities. Figure 9 compares the intensities of IDF curves for return periods (T_r) of 2 and 50 years, for the selected stations in each of the five regions in Colombia. In general, results show that there are underestimations of the intensities for short durations (0.5 and 1 h) and

overestimations for large durations (2 and 6 h). For rain gauge stations El Rancho (Andes) and Mercaderes (Pacific), the IDF curves are quite similar, with low errors that become larger for the 2 and 6 h durations. For the Aeropuerto El Caraño rain gauge (Pacific), results show that there is an underestimation of the intensities (maximum of -22%) for the 0.5 h duration, while there is an overestimation of the intensities for the other durations and return periods, which is smaller for the lower return periods (+25%) and larger for the others (+31%).

The Manure rain gauge station, in the Caribbean region, shows one of the poorest results overall. There are important underestimations of the intensities for short durations (0.5 and 1 h) with a maximum of -55%. For large durations, the observed and estimated IDFs are similar, with maximum relative errors of $\pm 20\%$. For the Araracuara rain gauge station (Orinoquia), results show maximum underestimations of -41% and maximum

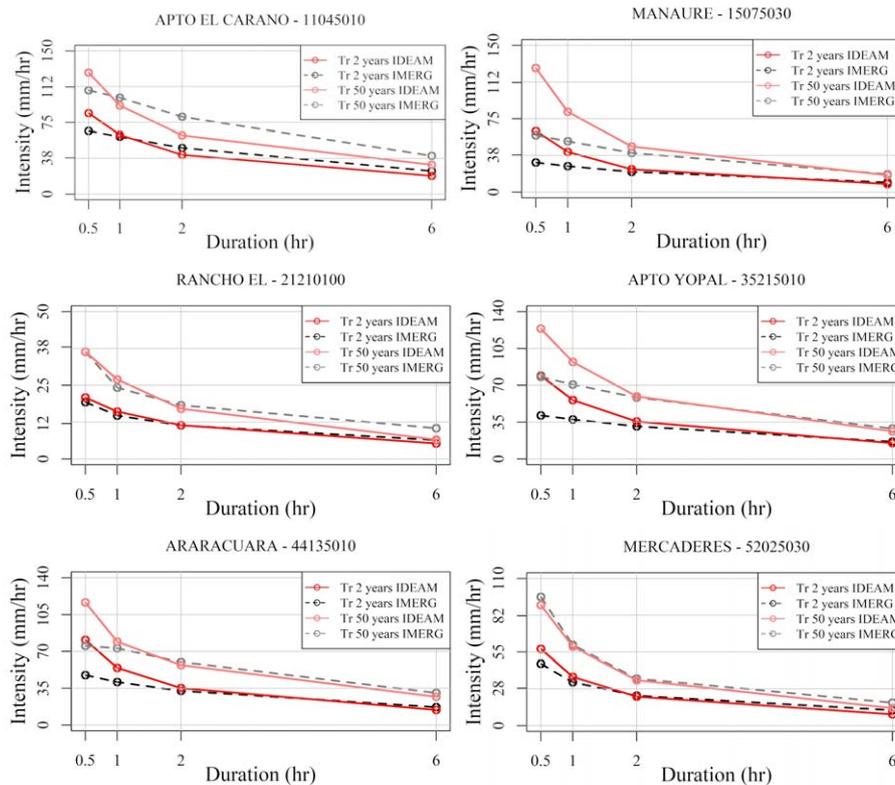


Figure 9 Comparison between IDEAM and IMERG intensities from IDF curves for selected stations. To facilitate the comparison, results are only shown for the minimum (2 years) and maximum (50 years) return periods (Tr).

overestimations of +17%. The comparisons of intensities for all IDF curves are included in the supplementary material and show mixed results, similar to the ones obtained for the six stations here analyzed.

6. Conclusions

In this study, we validated and explored the capabilities of the IMERG V06B FINAL product for reproducing intensities of IDF curves in 110 rain gauges in Colombia using in-situ and satellite data for the period 2001-2019. Based on the methods and analysis implemented, we found the following conclusions:

- The performance of the IMERG FINAL product in capturing rainfall distribution at yearly, monthly, and daily scales in the period 2001-2019 is quite variable and depends on the physiographic and climatological characteristics of the places where the stations are located. However, aggregated analysis for all stations in the five natural regions in Colombia shows that the best results are obtained yearly, followed by monthly, and daily time resolutions. The regional analysis of the errors and contingency metrics highlights that the best results are obtained for the Orinoco,

Amazon, and Pacific regions. Although this could be associated with the low number of stations in these three regions, the comparisons between IDEAM and IMERG maximum intensities reinforce these findings. Nevertheless, caution is recommended when extrapolating these results.

- The best agreements between intensities of IDF curves derived from IMERG and those available from IDEAM occur for the rain gauge stations located in the Orinoco, Amazon, and Pacific regions with relative errors in the rainfall intensities in the range +/- 2%. These rain gauges represent around a fifth of the total. For other stations, results are inferior, with larger relative errors in rainfall intensities in the range -69, +17%.
- In the majority of the stations for the IMERG IDF curves, there are underestimations of the rainfall intensities for short durations (0.5 and 1 h) and overestimations for the large durations (2 and 6 h), with relative errors in the ranges [-69, +5%] and [-61, +171%], respectively.
- In general terms, the best results for the intensities of IMERG IDF curves were obtained for stations with a high probability of detection of daily rainfall (POD). However, results also show that the IMERG product possibly reaches a high POD at the expense

of also a high probability of false detection (POFD). This represents an opportunity to improve the IMERG product by its developers.

- The results reported in this study used all the 110 rain gauge stations in Colombia with available IDF curves, updated to the year 2010. In this sense, our results are limited and may be considered cautiously. The inclusion of more stations in the analysis and the continuous update of the existing IDF curves are necessary to improve the results and enlarge the period of comparison, which was short in this study (2001-2019). An invitation to IDEAM was made to support these research activities.

7. Declaration of competing interest

We declare that we have no significant competing interests, including financial or non-financial, professional, or personal interests interfering with the full and objective presentation of the work described in this manuscript.

8. Acknowledgements

We acknowledge IDEAM for supporting the development of the IDF curves. Also, many thanks to the two anonymous reviewers for their thoughtful comments that allowed us to improve the quality of the manuscript.

9. Funding

This work was supported by Universidad Nacional de Colombia.

10. Author contributions

ER provided the idea for the study and helped write the first version of the article. CG, AG, JS, MP, and DA compiled the data and developed the codes for processing the information, producing the figures, maps, and tables. They also contributed to the writing of the final version of the paper.

11. Data availability Statement

The authors confirm that all the data, codes and results supporting the findings of this study are available in the supplementary material. <https://data.mendeley.com/datasets/bvbxpm4fsm/2>

References

- [1] D. Koutsoyiannis, D. Kozonis, and A. Manetas, "A mathematical framework for studying rainfall intensity-duration-frequency relationships," *Journal of Hydrology*, vol. 206, no. 1-2, apr. 1998. [Online]. Available: [https://doi.org/10.1016/S0022-1694\(98\)00097-3](https://doi.org/10.1016/S0022-1694(98)00097-3)
- [2] Z. Şen, "Annual daily maximum rainfall-based idf curve derivation methodology," *Earth Systems and Environment*, vol. 3, Sep. 24, 2019. [Online]. Available: <https://doi.org/10.1007/s41748-019-00124-x>
- [3] N. Bezak, M. Šraj, and M. Mikoš, "Copula-based idf curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides," *Journal of Hydrology*, vol. 541, Mar. 08, 2016. [Online]. Available: <https://doi.org/10.1016/j.jhydrol.2016.02.058>
- [4] E. V. Aristizábal-Giraldo, E. F. García-Aristizábal, R. J. Marín-Sánchez, F. Gómez-Cardona, and J. C. Guzmán-Martínez, "Rainfall-intensity effect on landslide hazard assessment due to climate change in north-western colombian andes," B.A. thesis, Universidad de Antioquia, Facultad de Ingeniería, Medellín, Colombia, 2022.
- [5] B. D. Avila-Rangel and H. F. Avila-Rangel, "Spatial and temporal estimation of the erosivity factor r based on daily rainfall data for the department of atlántico, colombia," B.A. thesis, Universidad Nacional, Ingeniería e investigación, Medellín, Colombia, 2015.
- [6] K. Kirby and J. Stratton, "Hidfun una herramienta para la extracción y análisis de pluviogramas," XXVIII Congreso latinoamericano de hidráulica, Buenos Aires, Argentina, 2018.
- [7] S. M. Papalexiou and D. Koutsoyiannis, "Battle of extreme value distributions: A global survey on extreme daily rainfall," *Water resources Research*, vol. 49, no. 1, Jan. 15, 2013. [Online]. Available: <https://doi.org/10.1029/2012WR012557>
- [8] F. Serinaldi and C. G. Kilsby, "Rainfall extremes: Toward reconciliation after the battle of distributions," *Water resources Research*, vol. 50, no. 1, Jan. 04, 2014. [Online]. Available: <https://doi.org/10.1002/2013WR014211>
- [9] S. Emmanouil, A. Langousis, E. I. Nikolopoulos, and E. N. Anagnostou, "Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records," *Journal of Hydrology*, vol. 589, May. 01, 2020. [Online]. Available: <https://doi.org/10.1016/j.jhydrol.2020.125151>
- [10] D. Grajales-Cardona and L. F. Carvajal-Serna, "Nonstationary intensity-duration-frequency curves for medellin river basin," *DYNA*, vol. 86, no. 208, Jan-Mar. 2019. [Online]. Available: DOI:<https://doi.org/10.15446/dyna.v86n208.69300>
- [11] A. Langousis and D. Veneziano, "Intensity-duration-frequency curves from scaling representations of rainfall," *Water resources Research*, vol. 43, no. 2, Feb. 21, 2007. [Online]. Available: <https://doi.org/10.1029/2006WR005245>
- [12] H. V. de Vyver and G. R. Demarée, "Construction of intensity-duration-frequency [idf] curves for precipitation at lubumbashi, congo, under the hypothesis of inadequate data," *Hydrological Sciences Journal*, vol. 55, no. 4, Nov. 02, 2009. [Online]. Available: <https://doi.org/10.1080/02626661003747390>
- [13] H. Tyrallis and A. Langousis, "Estimation of intensity-duration-frequency curves using max-stable processes," *Stochastic Environmental Research and Risk Assessment*, vol. 33, Jul. 06, 2018. [Online]. Available: <https://doi.org/10.1007/s00477-018-1577-2>
- [14] C. Cheng-lung, "Rainfall intensity duration frequency formulas," *Journal of Hydraulic Engineering*, vol. 109, no. 12, Dec. 01, 1983. [Online]. Available: [https://doi.org/10.1061/\(ASCE\)0733-9429\(1983\)109:12\(1603\)](https://doi.org/10.1061/(ASCE)0733-9429(1983)109:12(1603))
- [15] S. Urbanek, M. Vargas, and R. Díaz-Granados, "Curvas sintéticas regionalizadas de intensidad-duración-frecuencia para colombia." in *Seminario de Hidráulica e Hidrología, Sociedad Colombiana de Ingenieros*, Cali, Colombia, 1998, pp. 119-128.
- [16] E. G. Pulgarín-Dávila, "Fórmulas regionales para la estimación de curvas intensidad - frecuencia - duración basadas en las propiedades de escala de la lluvia (región andina colombiana),"

- M.S. thesis, Departamento de Geociencias y Medio Ambiente, Universidad Nacional, Medellín, Colombia, 2009.
- [17] *Estudio de Actualización y Análisis Comparativo de las Curvas Intensidad, Duración y Frecuencia Disponibles en el IDEAM*, Convenio Interadministrativo No. 113 de 2016, IDEAM-UNAL, Bogotá, D.C., 2016.
- [18] T. Dinku, F. Ruiz, S. J. Connor, and P. Ceccato, "Validation and intercomparison of satellite rainfall estimates over colombia," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 5, May, 01, 2010. [Online]. Available: <https://doi.org/10.1175/2009JAMC2260.1>
- [19] M. Ma, H. Wang, P. Jia, G. Tang, D. Wang, and *et al.*, "Application of the gpm-imerg products in flash flood warning: A case study in yunnan, china," *Remote Sensing*, vol. 12, no. 12, Jun. 15, 2020. [Online]. Available: <https://doi.org/10.3390/rs12121954>
- [20] W. Guachamín, S. Páez-Bimos, and N. Horna, "Evaluation of imerg v03 and tmpa v7 products in the detection of flood events in cañar river basin," *Revista Politécnica*, vol. 42, no. 2, Nov-Jan. 2019. [Online]. Available: <http://tinyurl.com/mvm8ubst>
- [21] R. Zubieta1, A. Getirana, J. C. Espinoza, W. Lavado-Casimiro, and L. Aragon, "Hydrological modeling of the peruvian-ecuadorian amazon basin using gpm-imerg satellite-based precipitation dataset," *Communications Week*, vol. 21, no. 7, Jul. 14, 2017. [Online]. Available: <https://doi.org/10.5194/hess-21-3543-2017>
- [22] R. Junqueira, M. R. Viola, J. da S. Amorim, C. Camargos, and C. R. de Mello, "Hydrological modeling using remote sensing precipitation data in a brazilian savanna basin," *Journal of South American Earth Sciences*, vol. 115, Mar. 13, 2022. [Online]. Available: <https://doi.org/10.1016/j.jsames.2022.103773>
- [23] F. Satgé, D. Ruelland, M. P. Bonnet, J. Molina, and R. Pillco, "Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the lake titicaca region," *Hydrology and Earth System Sciences*, vol. 23, no. 1, Jan. 14, 2019. [Online]. Available: <https://doi.org/10.5194/hess-23-595-2019>
- [24] E. Rodríguez, I. Sánchez, N. Duque, P. Arboleda, C. Vega, and *et al.*, "Combined use of local and global hydro meteorological data with hydrological models for water resources management in the magdalena - cauca macro basin - colombia," *Communications Week*, vol. 34, Mar. 19, 2019. [Online]. Available: <https://doi.org/10.1007/s11269-019-02236-5>
- [25] P. López-López, W. W. Immerzeel, E. A. Rodríguez-Sandoval, G. Sterk, and J. Schellekens, "Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the magdalena river basin in colombia," *Communications Week*, vol. 6, Jun. 08, 2018. [Online]. Available: <https://doi.org/10.3389/feart.2018.00068>
- [26] E. Zorretto and M. Marani, "Downscaling of rainfall extremes from satellite observations," *Water Resources Research*, vol. 5, no. 1, Nov. 26, 2018. [Online]. Available: <https://doi.org/10.1029/2018WR022950>
- [27] M. Zambrano-Bigiarini, C. Soto-Escobar, and O. M. Baez-Villanueva, "Spatially-distributed idf curves for center-southern chile using imerg," in *EGU General Assembly 2020*, Munich, Alemania, 2020.
- [28] M. Noor, T. Ismail, S. Shahid, M. Asaduzzaman, and A. Dewan, "Evaluating intensity-duration-frequency (idf) curves of satellite-based precipitation datasets in peninsular malaysia," *Atmospheric Research*, vol. 248, no. 19, Aug. 18, 2020. [Online]. Available: <https://doi.org/10.1016/j.atmosres.2020.105203>
- [29] K. Venkatesh, R. Maheswaran, and J. Devacharan, "Framework for developing idf curves using satellite precipitation: a case study using gpm-imerg v6 data," *Earth Science Informatics*, vol. 15, Oct. 02, 2021. [Online]. Available: <https://doi.org/10.1007/s12145-021-00708-0>
- [30] M. Ombadi, P. Nguyen, S. Sorooshian, and K. Hsu, "Developing intensity-duration-frequency (idf) curves from satellite-based precipitation: Methodology and evaluation," *Water Resources Research*, vol. 54, no. 10, Sep. 04, 2018. [Online]. Available: <https://doi.org/10.1029/2018WR022929>
- [31] G. J. Huffman, E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan. [2019] Gpm imerg final precipitation l3 half hourly 0.1 degree x 0.1 degree v06 (gpm-3imerghh). [Online]. Available: <https://doi.org/10.5067/GPM/IMERG/3B-HH/06>
- [32] R. K. Pradhan, Y. Markonis, M. R. Vargas-Godoy, A. Villalba-Pradas, K. M. Andreadis, and *et al.*, "Review of gpm imerg performance: A global perspective," *Remote Sensing of Environment*, vol. 268, Oct. 12, 2021. [Online]. Available: <https://doi.org/10.1016/j.rse.2021.112754>
- [33] A. Almagro, P. T. Sanches-Oliveira, and L. Brocca, "Assessment of bottom-up satellite rainfall products on estimating river discharge and hydrologic signatures in brazilian catchments," *Journal of Hydrology*, vol. 603, Agu. 28, 2021. [Online]. Available: <https://doi.org/10.1016/j.jhydrol.2021.126897>
- [34] J. R. Rozante, D. A. Vila, J. Barboza-Chiquetto, A. D. A. Fernandes, and D. Souza-Alvim, "Evaluation of trmm/gpm blended daily products over brazil," *Remote Sensing*, vol. 10, no. 6, Apr. 11, 2018. [Online]. Available: <https://doi.org/10.3390/rs10060882>
- [35] E. da S. Freitas, V. H. R. Coelho, Y. Xuan, D. de C. D. Melo, A. N. Gadelha, and *et al.*, "The performance of the imerg satellite-based product in identifying sub-daily rainfall events and their properties," *Journal of Hydrology*, vol. 589, May. 28, 2020. [Online]. Available: <https://doi.org/10.1016/j.jhydrol.2020.125128>
- [36] Y. Rojas, J. R. Minder, L. S. Campbell, A. Massmann, and R. Garreaud, "Assessment of gpm imerg satellite precipitation estimation and its dependence on microphysical rain regimes over the mountains of south-central chile," *Atmospheric Research*, vol. 253, no. 19, Jan. 04, 2021. [Online]. Available: <https://doi.org/10.1016/j.atmosres.2021.105454>
- [37] S. Palomino-Ángel, J. A. Anaya-Acevedo, and B. A. Botero, "Evaluation of 3b42v7 and imerg daily-precipitation products for a very high-precipitation region in northwestern south america," *Communications Week*, vol. 217, Oct. 23, 2018. [Online]. Available: <https://doi.org/10.1016/j.atmosres.2018.10.012>
- [38] D. Delgado, M. Sadaoui, W. Ludwig, and W. Méndez, "Spatio-temporal assessment of rainfall erosivity in ecuador based on rusle using satellite-based high frequency gpm-imerg precipitation data," *Communications Week*, vol. 219, Aug. 21, 2022. [Online]. Available: <https://doi.org/10.1016/j.catena.2022.106597>
- [39] B. Manz, S. Páez-Bimos, N. Horna, W. Buytaert, B. Ochoa-Tocachi, and *et al.*, "Comparative ground validation of imerg and tmpa at variable spatiotemporal scales in the tropical andes," *Journal of Hydrometeorology*, vol. 18, no. 9, Sep. 01, 2017. [Online]. Available: <https://doi.org/10.1175/JHM-D-16-0277.1>
- [40] A. Kyaw-Kyaw, S. Shahid, and X. Wang, "Remote sensing for development of rainfall intensity-duration-frequency curves at ungauged locations of yangon, myanmar," *Water*, vol. 14, no. 11, May. 23, 2022. [Online]. Available: <https://doi.org/10.3390/w14111699>
- [41] S. OH, U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, and *et al.*, "Evaluation of gpm imerg early, late, and final rainfall estimates using wegenernet gauge data in southeastern austria," *Hydrology and Earth System Sciences*, vol. 21, no. 12, Nov. 13, 2017. [Online]. Available: <https://doi.org/10.5194/hess-21-6559-2017>
- [42] F. F. Maghsood, H. Hashemi, S. Hasan-Hosseini, and R. Berndtsson, "Ground validation of gpm imerg precipitation products over iran," *Remote Sensing*, vol. 12, no. 1, Dec. 18, 2019. [Online]. Available: <https://doi.org/10.3390/rs12010048>
- [43] I. W. Andi-Yuda1, R. Prasetya1, A. R. As-syakur, T. Osawa, and Masahiko-Naga, "An assessment of imerg rainfall products over bali at multiple time scale," *Environmental Sciences*, vol. 153, no. 12, Feb. 17, 2020. [Online]. Available: <https://doi.org/10.1051/e3sconf/202015302001>
- [44] A. Navarro, E. García-Ortega, A. Merino, J. L. Sánchez, and F. J. Tapiador, "Orographic biases in imerg precipitation estimates in the ebro river basin (spain): The effects of rain gauge density and altitude," *Atmospheric Research*, vol. 244, no. 19, May. 22, 2020. [Online]. Available: <https://doi.org/10.1016/j.atmosres.2020.105068>
- [45] S. Moazami and M. R. Najafi, "A comprehensive evaluation of gpm-imerg v06 and mrms with hourly ground-based precipitation observations across canada," *Journal of Hydrology*, vol. 594, Dec. 20, 2020. [Online]. Available: <https://doi.org/10.1016/j.jhydrol.2020.125929>
- [46] M. Arshad, X. Ma, J. Yin, W. Ullah, G. Ali, and *et al.*, "Evaluation of gpm-imerg and trmm-3b42 precipitation products over pakistan,"

- Atmospheric Research*, vol. 249, Oct. 24, 2020. [Online]. Available: <https://doi.org/10.1016/j.atmosres.2020.105341>
- [47] S. M. Vallejo-Bernal, V. Urrea, J. M. Bedoya-Soto, A. O. D. Posada, and *et al.*, "Ground validation of trmm 3b43 v7 precipitation estimates over colombia. part i: Monthly and seasonal timescales," *International Journal of Climatology*, vol. 41, no. 1, May. 14, 2020. [Online]. Available: <https://doi.org/10.1002/joc.6640>
- [48] S. Blenkinsop, H. J. Fowler, R. Barbero, S. C. Chan, S. B. Guerreiro, E. Kendon, G. Lenderink, and *et al.*, "The intense project: using observations and models to understand the past, present and future of sub-daily rainfall extremes," *Advances in Science and Research*, vol. 15, Jun. 05, 2018. [Online]. Available: <https://doi.org/10.5194/asr-15-117-2018>
- [49] O. D. Álvarez Villa, J. I. Vélez, and G. Poveda, "Improved long-term mean annual rainfall fields for colombia," *International Journal of Climatology*, vol. 31, no. 14, Oct. 07, 2010. [Online]. Available: <https://doi.org/10.1002/joc.2232>
- [50] G. Poveda, "La hidroclimatología de colombia: una síntesis desde la escala inter-decadal hasta la escala diurna," *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, vol. 28, no. 107, Jun. 2004. [Online]. Available: <https://tinyurl.com/4wsv9fec>
- [51] O. Mesa, L. F. Carvajal, and G. Poveda, *Introducción al clima de Colombia*. Medellín, Colombia: Universidad Nacional de Colombia, Facultad de Minas, 1997.
- [52] J. Pabón, J. Eslava, and R. Gómez, "Características de gran escala del clima de la américa tropical, 4, 47-53," *Meteorol*, vol. 4, pp. 47-53, 2001.
- [53] DHIME. (2021) Consulta y descarga de datos hidrometeorológicos. IDEAM. [Online]. Available: <https://dhime.ideam.gov.co/webgis/home/>
- [54] G. N. repository. (2021) lmerg v06b final time series of rainfall intensity at 0.5 h intervals in (mm/hr) and at a resolution of 0.10°. NASA Official. [Online]. Available: <https://giovanni.gsfc.nasa.gov/giovanni/>
- [55] Z. Li, D. Yang, B. Gao, Y. Jiao, Y. Hong, and *et al.*, "Multiscale hydrologic applications of the latest satellite precipitation products in the yangtze river basin using a distributed hydrologic model," *Journal of Hydrometeorology*, vol. 16, no. 1, Feb. 01, 2015. [Online]. Available: <https://doi.org/10.1175/JHM-D-14-0105.1>
- [56] B. Peng, J. Shi, W. Ni-Meister, T. Zhao, and D. Ji, "Evaluation of trmm multisatellite precipitation analysis (tmpa) products and their potential hydrological application at an arid and semiarid basin in china," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote*, vol. 7, no. 9, Sep. 2014. [Online]. Available: <https://doi.org/10.1109/JSTARS.2014.2320756>
- [57] X. Guan, J. Zhang, Q. Yang, X. Tang, C. Liu, and *et al.*, "Evaluation of precipitation products by using multiple hydrological models over the upper yellow river basin, china," *Remote Sensing*, vol. 12, no. 24, Dec. 06, 2020. [Online]. Available: <https://doi.org/10.3390/rs12244023>
- [58] Z. Wang, J. Chen, C. Lai, R. Zhong, X. Chen, and *et al.*, "Hydrologic assessment of the tmpa 3b42-v7 product in a typical alpine and gorge region: the lancang river basin, china," *Hydrology Research*, vol. 49, no. 6, Dec. 01, 2018. [Online]. Available: <https://doi.org/10.2166/nh.2018.024>
- [59] B. Zhu, Y. Huang, Z. Zhang, R. Kong, J. Tian, and *et al.*, "Evaluation of tmpa satellite precipitation in driving vic hydrological model over the upper yangtze river basin," *Water*, vol. 12, no. 11, Nov. 16, 2020. [Online]. Available: <https://doi.org/10.3390/w12113230>
- [60] R. J. Donaldson, R. M. Dyer, and M. J. Kraus, "An objective evaluator of techniques for predicting severe weather events," in *Preprints Ninth Conference on Severe Local Storms, Norman, Amer*, 1975.
- [61] C. A. D. III, R. Davis-Jones, and D. L. Keller, "On summary measures of skill in rare event forecasting based on contingency tables," *Weather and Forecasting*, vol. 5, no. 4, 1990. [Online]. Available: [https://doi.org/10.1175/1520-0434\(1990\)005<0576:OSMOSI>2.0.CO;2](https://doi.org/10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2)