UNIVERSIDADE FEDERAL DE ALAGOAS
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS HÍDRICOS E SANEAMENTO

WALLISSON MOREIRA DE CARVALHO

AVALIAÇÃO DE UM PRODUTO DE PRECIPITAÇÃO ESTIMADO POR SATÉLITE PARA O BRASIL E SUA APLICABILIDADE NA MODELAGEM HIDROLÓGICA DE UMA BACIA TROPICAL/SUBTROPICAL DE LARGA ESCALA

Maceió-AL

Dezembro de 2021

WALLISSON MOREIRA DE CARVALHO

Evaluation of Satellite Precipitation Product for Brazil and its Applicability for Hydrological Modeling in a Large-Scale Tropical/SubTropical Basin

Dissertação apresentada ao Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Centro de Tecnologia da Universidade Federal de Alagoas, como requisito para obtenção do título de Mestre em Recursos Hídricos e Saneamento.

Orientador: Prof. Dr. Carlos Ruberto Fragoso Júnior

Maceió-AL

Dezembro de 2021

Catalogação na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 – 1767

C331e Carvalho, Wallisson Moreira de.
Evaluation of Satellite Precipitation Product for Brazil and its applicability for hydrological modeling in a large-scale tropical/subtropical basin / Wallisson Moreira de Carvalho. – 2021.
58 f. : il. color.

Orientadora: Carlos Ruberto Fragoso Júnior.
Dissertação (Mestrado em Recursos Hídricos e Saneamento) – Universidade Federal de Alagoas. Centro de Tecnologia. Maceió.

Bibliografia: f. 45-49.
Apêndices: f. 50-58.

1. *Integrated Multi-satellite Retrievals for Global Precipitation Measurement*. 2. MGB. 3. Doce, Rio, Bacia. I. Título.

CDU: 556

Folha de Aprovação

WALLISSON MOREIRA DE CARVALHO

Avaliação de um produto de precipitação estimado por satélite para o Brasil e sua aplicabilidade na modelagem hidrológica de uma bacia tropical/subtropical de larga escala

Dissertação apresentada ao Programa de Pós-Graduação em Recursos Hídricos e Saneamento, Centro de Tecnologia da Universidade Federal de Alagoas, como requisito para obtenção do título de Mestre em Recursos Hídricos e Saneamento.

Em: 02/ 12 /2021

gov.br Documento assinado digitalmente
Carlos Ruberto Fragoso Junior
Data: 15/12/2021 18:15:23-0300
Verifique em https://verificador.iti.br

Prof. Dr. Carlos Ruberto Fragoso Júnior
(Orientador - PPGRHS/ CTEC/UFAL)

Banca examinadora:

gov.br Documento assinado digitalmente
Fabio Farias Pereira
Data: 15/12/2021 21:40:03-0300
Verifique em https://verificador.iti.br

Prof. Dr. Fábio Farias Pereira
(Examinador Interno – PPGRHS/ CTEC/UFAL)

gov.br Documento assinado digitalmente
REGINA CAMARA LINS
Data: 15/12/2021 18:19:01-0300
Verifique em https://verificador.iti.br

Prof.ª Dr.ª Regina Camara Lins
(Examinadora Externa – Campus Sertão/UFAL)

gov.br Documento assinado digitalmente
ANDERSON LUIS RUHOFF
Data: 15/12/2021 20:32:19-0300
Verifique em https://verificador.iti.br

Prof. Dr. Anderson Luis Ruhoff
(Examinador Externo – UFRGS)

AGRADECIMENTOS

Diante da crise sanitária, científica e ambiental que vivemos nos últimos dois anos, considero que tive diversos privilégios na minha trajetória na pós-graduação e que sem esses não seria possível chegar até aqui, e por isso, serei eternamente grato.

Agradeço aos meus pais José e Ana, e meu irmão, Wanderson, por todo amor e apoio incondicional, fazendo tudo isso ser possível. À Letícia, por todo companheirismo, carinho e compreensão. À minha segunda família Eliene, Isabela, Juliana e Paulo.

Ao meu orientador Ruberto, por todas as oportunidades e confiança a mim depositada. Não tenho dúvidas do quanto foram essenciais para a construção do profissional que sou hoje.

À Ágatha, Almir e Denis, por todos os bons momentos, pela amizade sincera, e por terem feito com que eu me sentisse em casa durante meu período em Lund.

Ao Cleber, pela disponibilidade em ajudar sempre que precisei. Aos amigos do *All Stars*, aos amigos de São Miguel e a todos, que de alguma forma contribuíram na minha formação científica, profissional e como cidadão.

Deus, muito obrigado.

RESUMO

Os dados de precipitação da missão Global Precipitation Measurement (GPM) fornecem uma nova fonte de informações com uma alta resolução espaço-temporal que supera as limitações das informações de uma rede de monitoramento pluviométrico convencional. Este estudo avalia o desempenho do produto Integrated Multi-satellitE Retrievals for GPM (IMERG) V06 Final Run sobre o Brasil e investiga a sua aplicabilidade de simulação hidrológica através de simulações realizadas numa bacia de grande escala submetida ao clima tropical/subtropical. Os resultados mostraram que o GPM-IMERG apresenta uma tendência para subestimar a quantidade de precipitação para maiores intensidades de precipitação, e que a sua capacidade de detecção é espacialmente variável e sensível às variações de longitude. Além disso, o GPM-IMERG mostrou melhores desempenhos em regiões sob climas subtropicais. No contexto do uso do GPM-IMERG para modelagem hidrológica, os resultados mostraram que o GPM-IMERG foi capaz de levar o modelo hidrológico a captar o padrão sazonal de precipitação e representar as variações espaciais e temporais das vazões.

Palavras-chaves: GPM-IMERG; MGB; Bacia do Rio Doce.

ABSTRACT

Rainfall data from the Global Precipitation Measurement (GPM) mission provide a new source of information with a high spatiotemporal resolution that overcomes the limitations of ground-based rainfall information worldwide. This study evaluates the performance of the Integrated Multi-satellitE Retrievals for GPM (IMERG) V06 Final Run product over Brazil and investigates its applicability hydrological simulation by performing in a large-scale tropical/subtropical basin. The results showed that GPM-IMERG presents a tendency to underestimate the amount of precipitation for higher precipitation intensities, and its detection ability is spatially variant and sensitive for longitude. Moreover, GPM-IMERG showed better performances in regions under subtropical climates. The GPM-IMERG application for hydrological modeling context showed that GPM-IMERG could drive the hydrological model to capture the seasonal rainfall pattern and represent the spatial and temporal streamflow variations.

Keywords: GPM-IMERG; MGB; Doce River Basin.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

SUMÁRIO

1 INTRODUCTION

Reliable information about the locations and extents of rain plays a fundamental role for water resources management. Besides, ground-based can be considered the reference data source for precipitation observation(TAPIADOR et al., 2012), and have been used for calibration and validation for other precipitation data sources, and applied for hydrological modeling (KALIN; HANTUSH, 2006; AMORIM et al., 2020). Despite its importance, obtaining consistent, continuously, and with an adequate density ground-based monitoring network still is a challenge (GADELHA et al., 2019).

In regional applications, where the spatial distribution of rainfall may play a crucial role, there are limitations for using ground-based measurements, especially in most developing countries (MAGHSOOD et al., 2020), which the rain gauges are insufficient and unevenly distributed (SALIO et al., 2015). To overcome these limitations, Satellite Precipitation Products (SPPs) have the potential to provide a solution for uninterrupted and spatially well-distributed precipitation measurements with nearly global coverage.

The strengths of the various groups of SPPs were combined to create the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM-IMERG) (HUFFMAN et al., 2014). This very high-resolution precipitation product is now available at 0.1° x 0.1° spatial and half-hourly temporal resolutions. Its most recent version has a temporal coverage from June 2000 until the present and introduces significant improvements to the previous versions (HUFFMAN et al., 2019).

Streamflow information is another main component of water resources management. In Brazil, the gauge network density for streamflow is even lower than for rain, leading to limitation of hydrological dynamics understanding over the country. Overall, the use of SPPs as inputs in hydrological models can overcome the imposed limitations from using rain gauge networks, providing a better representation of the spatial variability of precipitation (BITEW; GEBREMICHAEL, 2011; LE et al., 2020).

In this context, the SPPs can be a possible alternative to improve the lack of understanding of the hydrological dynamics in Brazil throughout the hydrological modeling. Thus, the first part of this study evaluates the GPM-IMERG product using ground-based precipitation stations over Brazil. Then, a comparison of the hydrological modeling for streamflow simulation performance using rain gauges and GPM-IMERG precipitation in a large-scale basin, which encompasses the main climate zones in Brazil.

1.1 Research questions and objectives

This study main objective was evaluate GPM-IMERG performance over Brazil and its applicability as input data for hydrological modeling in a large-scale tropical/subtropical basin.

The research questions of this thesis are:

1. How does the GPM-IMERG daily product precipitation generally perform over Brazil?
2. How does the latest GPM-IMERG daily precipitation product perform in different thresholds of precipitation intensity?
3. How is the GPM-IMERG precipitation performance spatially and temporally distributed?
4. Can GPM-IMERG serve as an alternative input to hydrological models?
5. The hydrological model using GPM-IMERG as precipitation input data could capture the hydrological cycle temporally and spatially?

1.2 Thesis Structure

This thesis is divided into 6 chapters. A general introduction, research rationale, and objectives are outlined in chapter 1. Then, the thesis is divided into two main parts (papers): Chapter 2) Long-term ground-based evaluation of GPM-IMERG V6 over Brazil, and Chapter 3) Evaluation of GPM-IMERG applicability for hydrological modeling in a large-scale tropical basin. In chapter 4 is presented the general conclusions of the study. Lastly, Chapters 5 and 6 are supplementary materials.

2 LONG-TERM GROUND-BASED EVALUATION OF GPM-IMERG V6 OVER BRAZIL

2.1 Introduction

Since precipitation is a fundamental input for a wide range of applications in Earth's science, reliable information about the locations and extents of rain is vital to support water resources managers and environmental planners. Despite that, obtaining a consistent, continuously, and with an adequate density ground-based monitoring network still is a challenge (GADELHA et al., 2019; HOBOUCHIAN et al., 2017), that can be associated with several factors, such as its cost of operation and maintenance (PARDO-IGÚZQUIZA, 1998), or the climate system of a region, where the large spatial and temporal variability makes challenging to measure or estimate the precipitation (FALCK et al., 2015).

Ground-based stations can be considered the reference data source for precipitation observation (TAPIADOR et al., 2012), and the most adequate for hydrological modeling (KALIN; HANTUSH, 2006; AMORIM et al., 2020). In regional applications, where the spatial distribution of rainfall may play a crucial role, however, there are limitations for using ground-based measurements, especially in most developing countries (MAGHSOOD et al., 2020), in which the rain gauges are insufficient and unevenly distributed (SALIO et al., 2015). Currently, Brazil has an average density of rain gauge per km^2 below the World Meteorological Organization (WMO) recommendation (GADELHA et al., 2019). Moreover, the gauges are non-uniformly distributed throughout the country, with a more acute lack in some basins. Also, frequently the available gauged stations include observation periods shorter than needed, and a few stations have uninterrupted data with most of the time series containing temporal gaps (MONTEIRO et al., 2016; GADELHA et al., 2019).

Weather radars are a common ground-based method to provide an indirect measurement of rainfall. The estimated precipitation from weather radars may have advantages compared to conventional rain gauge networks, such as extensive spatial coverage and high-resolution in space-time monitoring (TAPIADOR et al., 2012). However, a radar system network remains expansive due to high installment costs and maintenance demands (LE et al., 2020). Furthermore, radar rain rates may also be a source of errors and uncertainties that can result in low-quality hydrological forecasts (JORDAN; SEED; AUSTIN, 2000; FRANZ; HOGUE, 2011; CECINATI et al., 2017).

Satellite Precipitation Products (SPPs) have the potential to provide a solution for uninterrupted and spatially well-distributed precipitation measurements with nearly global coverage. The Tropical Rainfall Measuring Mission (TRMM) was the first satellite for space-borne precipitation measurements, which became operational in 1997. Several SPPs have been released since TRMM, and numerous studies have been evaluating them, both in terms of mean errors

and detection capabilities in different parts of the world encompassing diverse climatic and topographic conditions (MAGGIONI; MEYERS; ROBINSON, 2016; LE et al., 2020). The most frequent reported limitation is the underestimation of light precipitation and warm rain events, biased estimation over complex terrain (e.g., mountainous areas and orographic enhancement of rainfall) (MAGGIONI; MEYERS; ROBINSON, 2016).

The strengths of the various groups of SPPs were combined to create the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM-IMERG) (HUFFMAN et al., 2014). This high-resolution precipitation product is now available at 0.1° x 0.1° spatial and half-hourly temporal resolutions. Its most recent version has a temporal coverage from June 2000 until the present and introduces significant improvements to the previous versions (HUFFMAN et al., 2019).

The GPM-IMERG products can improve the understanding of the hydrological dynamics, especially in ungauged areas (ASONG et al., 2017; GADELHA et al., 2019). Although, still is necessary its evaluation concerning station gauge observations in different regions of the world (PRAKASH et al., 2018). In that way, to our knowledge, only two studies that investigate the performance of the GPM-IMERG product over Brazil at country level (ROZANTE et al., 2018; GADELHA et al., 2019), which used version 05. Moreover, none of them had used more than three years of data on their analyses.

Considering the above, the newly available GPM-IMERG products version (i.e., version 06) have not been explored for Brazil. This study aims to aggregate into the analyses of the GPM-IMERG Final Run (i.e., research product) over Brazil. Provides a better understanding of the performance of the GPM-IMERG product across the whole country can open opportunities to future studies regarding hydrological and hydrometeorological applications of these products.

2.2 Materials and Methods

2.2.1 Study Area

Brazil is the largest country in Latin America with continental dimensions, covering about 8,516,000 km^2, encompassing a vast diversity of landscapes, climate, topography, biodiversity, and precipitation regimes. By the Köppen's climate classification (ALVARES et al., 2013), Brazil has three main zones, classified as Tropical, Semiarid, and Humid Subtropical, which are subdivided into 12 different climate types (Figure 2.1a). On the other hand, it was identified (REBOITA et al., 2010; ROZANTE et al., 2018) four dominant precipitation regimes within Brazilian territory (Figure 2.1b). Region 1, located in southern Brazil, presents a well annually distributed precipitation with around 1,750-2,100 mm/year, mainly influenced by the South Atlantic Convergence Zone, cold fronts, mesoscale convective complexes, and low-level jetstream (VELASCO; FRITSCH, 1987; VERA et al., 2006; QUADRO et al., 2012). Region 2 encompasses most of the country, with a monsoon regime (ZHOU; LAU, 1998) featured by

seasonal rainfall variability with very rainy and dry seasons. Region 3 is influenced by the Inter-tropical Convergence Zone (ITCZ), and upper-level cyclonic vortexes (KOUSKY; GAN, 1981), and it is the driest region of the Brazilian territory, with totals of precipitation ranging around 200-500 mm/year. Region R4 comprises the northern region and the northeastern coast of Brazil, with maximum precipitation in winter and minimum in summer. The precipitation is above 2,000 mm/year in the northern region, while it is around 1,500 mm/year on the northeastern coast. The central systems in both regions are the ICTZ, the Trade Winds, upper-level cyclonic vortexes, and topical mesoscale convective system (KOUSKY; GAN, 1981; KOUSKY, 1988) The northeastern coast is also influenced by the easterly waves and sea breeze circulation (KOUSKY, 1988).

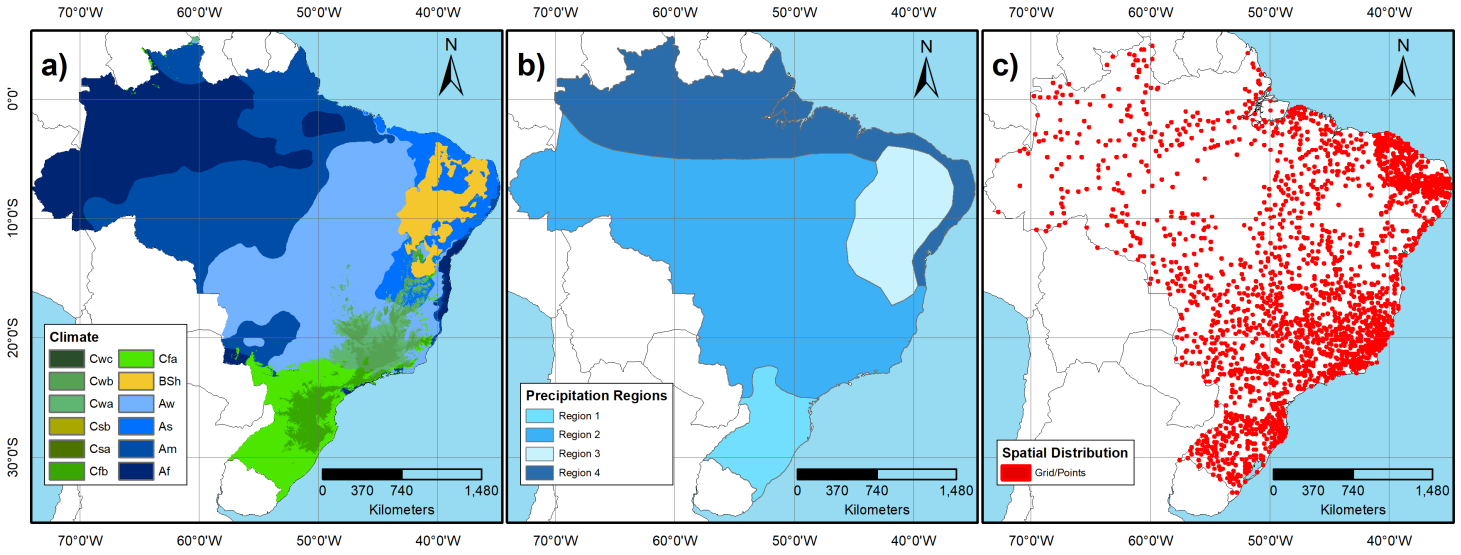

Figure 2.1 – (a) Köppen's climate classification map for Brazil according to Alvares et al. (2013), (b) Spatial distribution of precipitation climatology adapted from Reboita et al. (2010), (c) Spatial distribution of grid points which present rain gauge data frequency of at least 90% in the studied period.

2.2.2 Data Acquisition and Processing

2.2.2.1 Ground-based Precipitation

The ground-based daily precipitation data were obtained from the Brazilian Water Agency (ANA) database (*http://www.snirh.gov.br/hidroweb/*). For the last 20 years, ANA provides precipitation data for 5,336 rain gauges across the country, where some of these stations are not currently active. Besides the measured data, ANA provides classification regarding its consistency (i.e., if the data is already validated). In order to get more accurate statistical results, it was obtained only the validated data and it was selected the stations with high temporal overlap with the GPM-IMERG product, considering a maximum of 10% threshold of missing data between 01-June-2000 until 30-June-2020, which resulted in a total of 2,259 rain gauges (Figure 2.1c).

2.2.2.2 GPM-IMERG

The Global Precipitation Measurement (GPM) mission is a collaboration between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA), which was launched in February 2014. The CPM Cory Observatory carries a constellation with the most advanced sensors for measuring precipitation. From its publically available data, the GPM-IMERG products have a spatial resolution of 0.1º x 0.1º and multiples temporal resolutions ranging from half-hourly to monthly, which are categorized by Early, Late, and Final runs. The Early runs provide a near real-time estimate with a low-latency period (4-h). Late runs are available with a latency of 12-h, with better estimates as data from more partner satellites is merged. The IMERG-Final run provides the highest quality precipitation estimates with a latency of 3.5 months once it is presented after bias adjustment. More detailed information and an algorithm description can be found in Huffman et al. (2017).

The final run of daily GPM-IMERG V6 (HUFFMAN et al., 2019) from 01-June-2000 until 30-June-2020 was used in this study. In order to use a point-to-grid approach where the precipitation at rain gauges (i.e. point) is compared with the indirect precipitation measurement, it was selected, only the 0.1º x 0.1º pixels had a gauged station within to be compared. Thus, a database was formed that pairs the selected rain gauge dataset and the selected GPM-IMERG grids. The database has three dimensions: 7334 days on the rows, 107 gauged stations/grids on columns, and attributes on aisles (i.e., z or depth direction). The first and second attributes are selected rain gauge dataset and selected and selected GPM-IMERG grids.

2.2.2.3 Digital Elevation Model (DEM)

In consideration of the elevation is one of the factors that affect the precipitation pattern (HASHEMI et al., 2017), spatial attributes were included in the database in order to evaluate the relationship between the precipitation product and the topography. In that way, for each selected grid, it was included as variables the grid centroid latitude and longitude, and the elevation mean $(E_{mean})[m]$. The Shuttle Radar Topographic Mission (SRTM) images by 3 arc-seconds (approximately 90 meters, varies in different locations due to the curvature of the earth) were used to obtain these physiographic attributes (USGS, 2015).

2.2.2.4 Climate Classification

In order to encompass a climate variable in the assessment, Köppen's climate classification provided by Alvares et al. (2013) was also included in the database. This product (ALVARES et al., 2013) has twelve different Köppen's climate types, which are divided into three main zones: Tropical (Zone A), Semiarid (Zone B), and Humid Subtropical (Zone C).

2.2.3 Evaluation Procedures

Despite its limitations, the evaluation procedure for ground validation used in this study was a point-to-pixel approach (XU et al., 2017; GADELHA et al., 2019; MAHMOUD; HAMOUDA; MOHAMED, 2019; MAGHSOOD et al., 2020), due to the accuracy limitations for the methods that consider spatial interpolation methods in the evaluation procedure (STEPHENSON; PATRICK, 1990).

The statistical metrics were calculated using all the gauged stations/grids in a first approach. Then, it was performed a four steps evaluation: i) national scale (i.e., general evaluation); ii) spatial distribution of rainfall detection ability of the satellite sensors; iii) spatial and topographical performance analyses; and iv) temporal (i.e., monthly) performance by climatic zones.

As highlighted by Tian et al. (2018), to evaluate the performance of satellite rainfall products, it is fundamental to incorporate an assessment of differentiating precipitation events according to their intensities in the analyses. Table 2.1 presents an intensity rain classification based on daily thresholds following the classification suggested by Xu et al. (2017) adapted for Brazil based on Rozante et al. (2018), which allows exploring GPM-IMERG ability to detect different precipitation intensities ranging from light to extreme precipitation, where it was used on the national scale evaluation.

Rain Intensity Classification	Daily Precipitation (P) Thresholds (mm)
Rain/no-rain	$P \leq 0.5$
Light precipitation	$0.5 \leq P < 2$
Low moderate precipitation	$2 \leq P < 5$
Moderate precipitation	$5 \leq P < 20$
Heavy precipitation	$20 \leq P < 40$
Extreme precipitation	$40 \leq P < 90$
Torrential precipitation	$90 \leq P$

Table 2.1 – Rain classification and thresholds. Adapted from Xu et al. (2017) and Rozante et al. (2018).

2.2.3.1 Performance Measures

To assess the quality of the GPM-IMERG V06 product, seven statistical metrics divided into three main groups were used. As in Gadelha et al. (2019), all GMP-IMERG V06 and/or rain gauge data with daily values below 0.1 mm were treated as zero. This consideration is due to the temporal resolution of the rain gauge network (i.e., daily measures).

The first group of metrics is related to comparison of detecting the observed rainfall events (WILKS, 2011), including : (i) probability of detection (POD), which gives the fraction of rain occurrences correctly detected; (ii) false alarm ratio (FAR), which gives the fraction of rain occurrences estimated which are not detected in the ground-based data; and (iii) critical success

index (CSI), which combines the characteristics of false alarms and missed events.

$$POD = \frac{a}{a+c} \tag{2.1}$$

$$FAR = \frac{b}{a+b} \tag{2.2}$$

$$CSI = \frac{a}{a+b+c} \tag{2.3}$$

where a is the number of times that rainfall events was observed and correctly detected, b is the number of times that rainfall was detected but not observed, c is the number of times that rainfall was observed but not detected. The preferable values of POD and CSI are closer to one, while for FAR are nearest to zero.

The second group corresponds metrics to analyze the quantitative rainfall differences between the GMP-IMERG V06 detection and the rain gauged data. The comparison was performed using: (i) the Mean Absolute Error (MAE), which indicates the error distribution and mean magnitude of errors without considering the direction, providing a stable estimate of the differences in the temporal precipitation data (HASHEMI et al., 2017); (ii) the Root Mean Square Error ($RMSE$), which gives the sample standard deviation of the differences between the GPM-IMERG V06 product and the observed rainfall; (iii) the modified relative bias ($rBias_\epsilon$), which shows the deviation of the detected from the ground-observed rainfall defined by Hashemi et al. (2017) as the Equation 2.6, where the ϵ is used to regularize grids where very low values of precipitation can result in a very large relative bias. A value of $0.5mm/day$ for ϵ, can adequately large the huge relative bias for low-precipitation events (HASHEMI et al., 2017) and was set in this study.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|S_i - G_i| \tag{2.4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - G_i)^2} \tag{2.5}$$

$$rBias_\epsilon = 2\frac{S_i - G_i}{\epsilon - (S_i - G_i)} \tag{2.6}$$

where Gi and Si are the ground-base rain gauge and satellite-based precipitation data at pixel i and n is the total number of pixels.

The third group describes the agreement between GPM-IMERG V06 estimates and ground-based data, where was included the Pearson correlation coefficient (CC). CC gives the

degree of linear association between two variables, ranging from +1 to -1, where +1 represents the total positive linear correlation, and -1 is the total negative

$$CC = \frac{\sum_{i=1}^{n}(G_i S_i) - (\sum_{i=1}^{n} G_i)\,(\sum_{i=1}^{n} S_i)}{\sqrt{\left[(\sum_{i=1}^{n} G_i^2) - (\sum_{i=1}^{n} G_i)^2\right]\left[(\sum_{i=1}^{n} S_i^2) - (\sum_{i=1}^{n} S_i)^2\right]}} \tag{2.7}$$

where Gi and Si are the ground-base rain gauge and satellite-based precipitation data at pixel i and n is the total number of pixels.

2.3 Results and Discussion

2.3.1 General Evaluation

As a first analysis, a general evaluation of IMERG V06 was performed using all selected grid points lumped together, where was obtained a general overview of GPM-IMERG V06 accuracy over the whole country for the selected time frame (01 June 2000 to 30 June 2020).

The capability of GPM-IMERG to estimate the frequency of the ranging precipitation intensities is presented in Figure 2.2 by the empirical cumulative distribution function (ECDF). The figure shows that the GPM-IMERG underestimates the frequency of dry days, where the probability of precipitation less than 0.1 mm is 0.56 while the gauged precipitation is 0.72. The GPM-IMERG frequency underestimation decreases with higher precipitation values, where the frequency of precipitation starting from $5mm$ (0.81 for GPM-IMERG and 0.83 for gauged) is well-estimated. It is important to note that GPM-IMERG can accurately estimate the frequency of precipitation above 10 mm, referring to how often they might happen, but it might not be able to estimate their magnitude correctly. On the other hand, for the metrics $CORR$ and CSI (Table 2.2) show a poor agreement between the two datasets with 0.43 and 0.41. However, the $RMSE$ and MAE presents relatively low values, and the positive value for $rBIAS_\epsilon$ indicates that GPM-IMERG generally overestimates precipitation over Brazil.

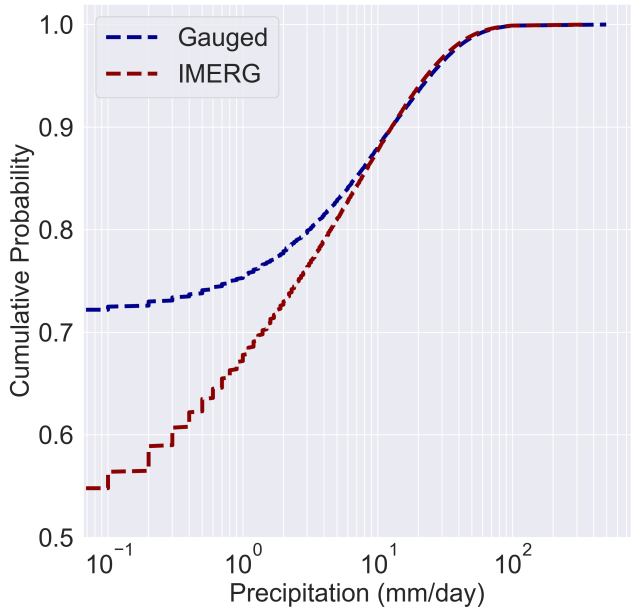

Figure 2.2 – Empirical cumulative distribution function of gauged data precipitation at stations and GPM-IMERG.

Perfomance measures	Values
$RMSE$	11.46
MAE	4.66
$rBIAS_{\epsilon}$	0.18
$CORR$	0.43
POD	0.77
FAR	0.53
CSI	0.41

Table 2.2 – Performance measures for the paired gauge-satellite dataset.

Performances measures for different precipitation thresholds are presented in Table 2.3. Except for the "Rain/no-rain"category, all days are rainy, which means that $FAR = 0$ and $CSI = POD$ hence FAR and CSI are not presented in Table 2.3. It is possible to note that while the probability of detection (POD) increases along with the precipitation intensity, the accuracy in estimate the amount of precipitation ($RMSE$, MAE, $rBIAS_{\epsilon}$) decreases. Moreover, GPM-IMERG could not capture the dynamics (i.e., $CORR$) for each precipitation intensity category presenting a worse performance relative to the validation where all data grid points were considered together (Table 2.2).

	Daily Precipitation (P) Thresholds (mm)						
Metrics	Rain/no-rain P≤0.5	Light 0.5≤P<2	Low moderate 2≤P<5	Moderate 5≤P<20	Heavy 20≤P<40	Extreme 40≤P<90	Torrential P≥90
$RMSE$	6.66	9.98	10.79	13.59	22.63	41.19	85.58
MAE	1.91	4.61	5.86	9.83	19.57	36.45	76.38
$rBIAS_{\epsilon}$	0.46	-0.06	-0.42	-0.72	-0.92	-1.01	-1.12
$CORR$	0.04	0.03	0.04	0.12	0.10	0.14	0.09
POD	0.53	0.63	0.70	0.80	0.89	0.93	0.95

Table 2.3 – Performance measures for the paired gauge-satellite dataset for different precipitation intensity thresholds.

Moazami e Najafi (2021) found that GPM-IMERG V06 tends to overestimate moderate to heavy precipitation events in Canada. In opposition to that, GPM-IMERG V06 increases its underestimation with higher amounts of precipitation in Brazil, as can be seen in Table 2.3.

2.3.2 Spatial Distribution of Rainfall Detection Ability

The rainfall detection ability of the satellite sensors can partly explain the agreement between the gauged and the estimated data (MAGHSOOD et al., 2020). In this section, the GPM-IMERG V06 was evaluated through spatial maps of the rainfall detection ability using the POD, FAR, and CSI metrics to investigate how the metrics vary for different parts of Brazil. The criteria indices were mapped using the Inverse of Distance Weighting method in $ArcMap$ 10.3.1 environment.

The spatial distribution of POD over Brazil (Figure 2.3 a), the satellite performance for precipitation detection shows high values ($POD > 0.8$) in most part of the country. Only along the eastern coast of the Brazilian North-East were observed values lower than 0.5. The poor POD over the eastern Brazilian coast has also been observed for GPM-IMERG V05 by Gadelha et al. (2019). Besides, lower values of POD in coastal areas is already expected (MAGHSOOD et al., 2020; GADELHA et al., 2019; PRAKASH et al., 2018).

According to Figure 2.3 (b), the GPM-IMERG V06 product exhibited higher false alarms at the central part of the NE, which is the Brazilian driest zone, and at the west part of the Midwest, these regions also showed a low number of rainy days during the study period, which can affect the FAR values, as was also noticed by Gadelha et al. (2019).

The critical success index (Figure 2.3 c) shows lower values in the Brazilian driest zone, which is consistent with the POD values, once the CSI combines the characteristics of FAR and the missed events.

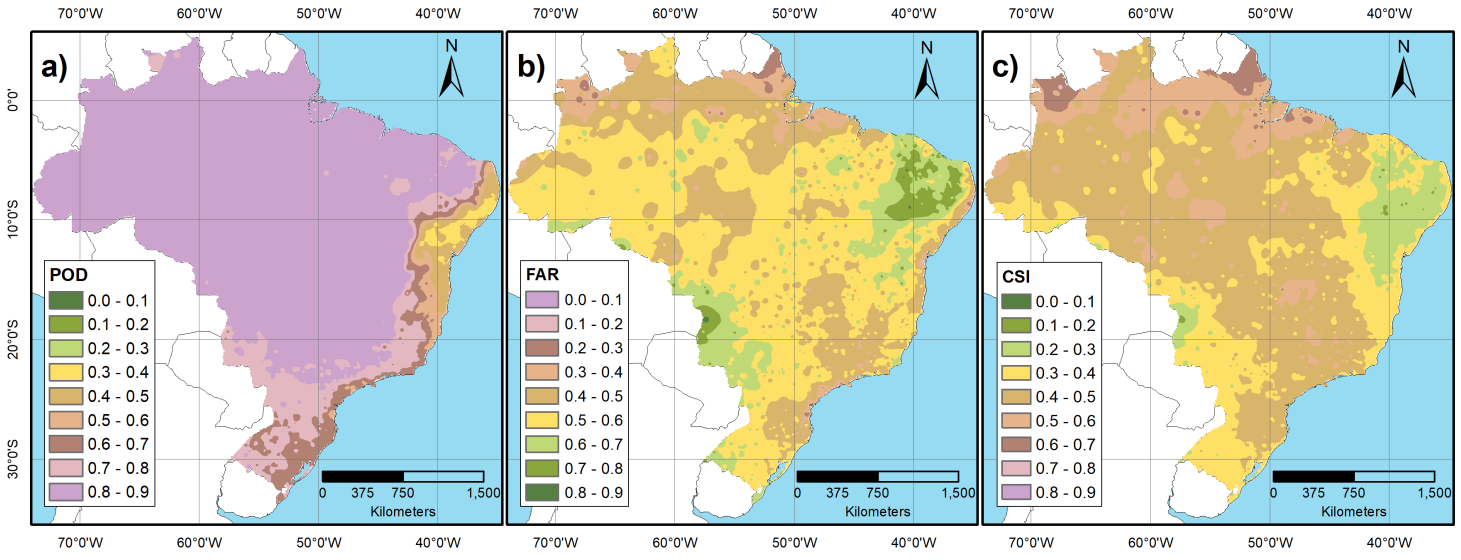

Figure 2.3 – Spatial distribution of (a) Probability of Detection - POD, (b) False Alarm Ration - FAR, and (c) Critical Success Index - CSI.

2.3.3 Spatial and Topographical Performance

To explore the impact of latitude, longitude, and topography on GPM-IMERG performance, the $RMSE$, MAE, $rBIAS_\epsilon$, and $CORR$ metrics were scatter plotted against these variables. Showing the distribution of the performance metrics allows identifying locations that with the worse GPM-IMERG perform.

Figure 2.4 shows that GPM-IMERG has a negative trend for the estimation errors (i.e., lower $RMSE$ and MAE) from south to north until around $15°S$, where there is a zone with higher dispersion of the performances (from $15°S$ to $5°S$), then the trend changes to positive. The zone from $15°S$ to $5°S$ encompasses three different precipitation zones (Figure 2.1 b), including the Brazilian driest region, which is the region with higher FAR. It is possible that the trends are associated with the amount of precipitation, where the regions with higher amounts tend to have higher $RMSE$ and MAE.

While in $rBIAS_\epsilon$, there is a low positive trend from south to north, it is not possible to see a clear trend for $CORR$, with its moving average ranging around 0.4.

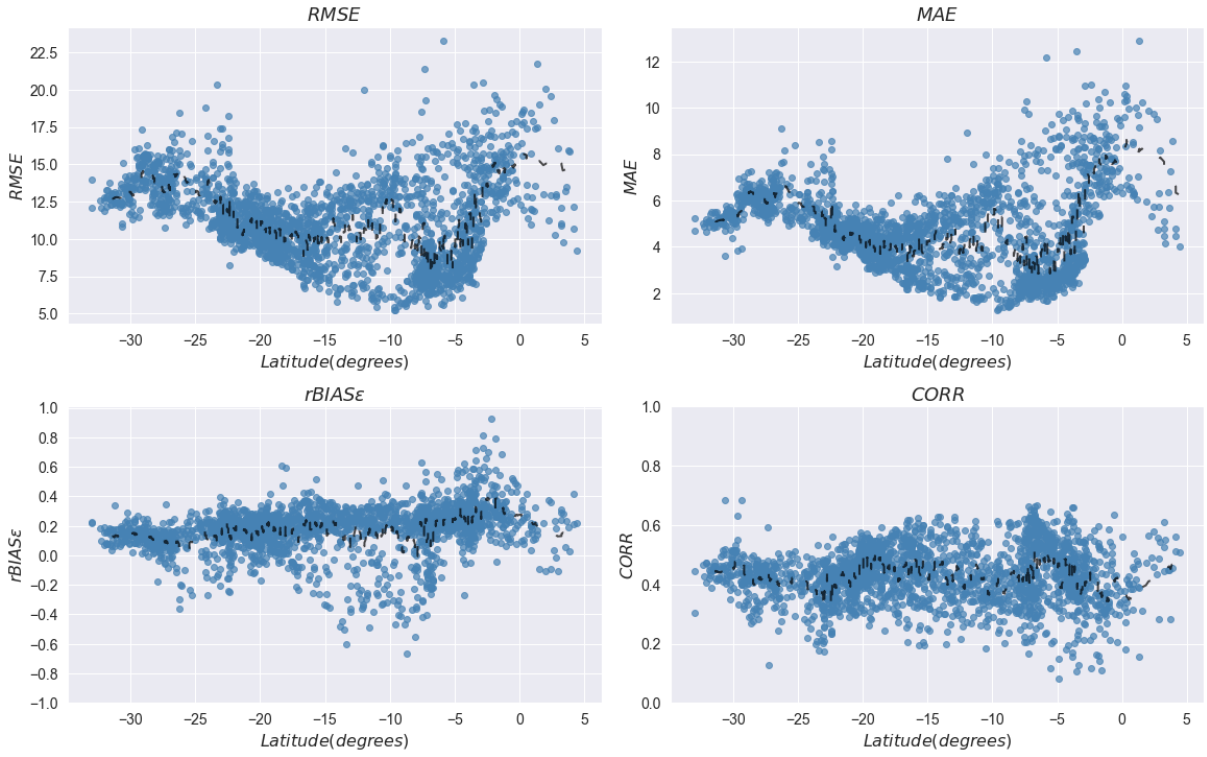

Figure 2.4 – Performance metrics of paired gauge-satellite data pixels plotted against the corresponding latitude of the pixel.

For metrics scatter plotted against Longitude (Figure 2.5) it is possible to see a clear negative trend for $RMSE$, MAE, and $rBIAS_\epsilon$, and a positive trend for $CORR$, from west to east, which means that GPM-IMERG V06 has better performance in the eastern parts of Brazil compared to western parts. The trends indicate that the performance of GPM-IMERG is more

sensitive to longitude compared to latitude.

Although some studies suggest that elevation as a predictor for poor performance (BHUIYAN et al., 2020; CHEN et al., 2018; LU et al., 2018), where the overestimation increases in highlands, Figure 2.6 shows a weak relationship between elevation and the performance metrics in Brazil, furthermore, the poor performance of GPM-IMERG for intense precipitation was found in this study, which corroborates with Hosseini-Moghari e Tang (2020) that argued the poor performance found in highlands is due to the precipitation intensity.

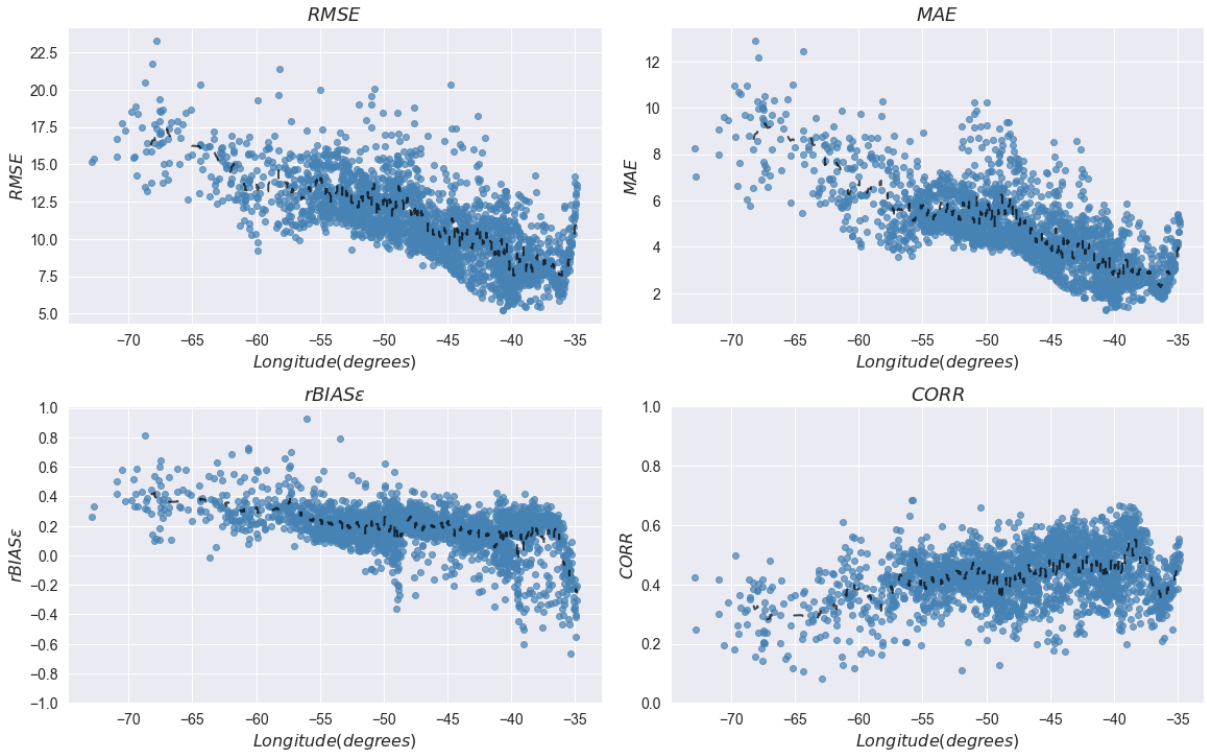

Figure 2.5 – Performance metrics of paired gauge-satellite data pixels plotted against the corresponding longitude of the pixel.

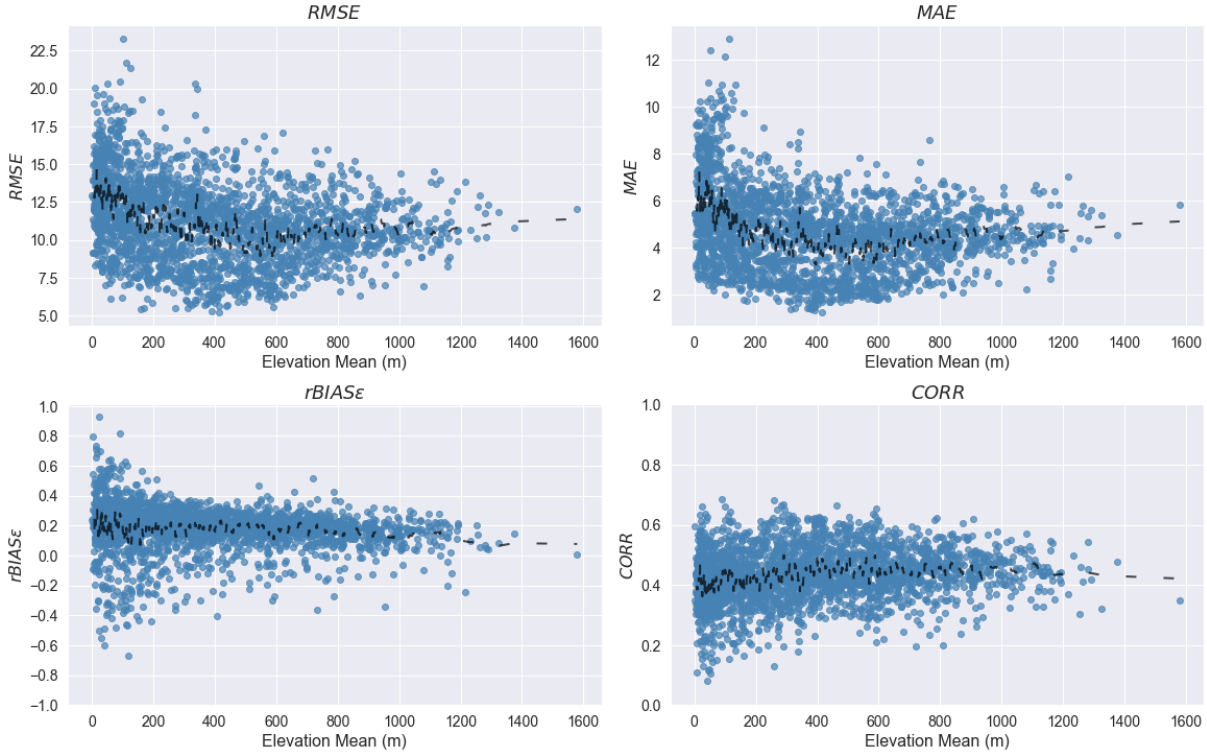

Figure 2.6 – Performance metrics of paired gauge-satellite data pixels plotted against the corresponding elevation mean of the pixel.

2.3.4 Temporal Performance by Climatic Zones

To explore the temporal characteristics of the calculated criteria and the effects of the climate zone in the GPM-IMERG performance, a monthly-based comparison was conducted using *boxplots* of the monthly precipitation average precipitation for the climatic zones (Figure 2.7) and using radar charts of performance indices for GPM-IMERG V06 daily product (Figure 2.8). The *boxplots* for all climate zones separately and the performance values used in the radar chart are displayed radar charts are separately displayed in the Supplementary Materials (Chapter 5).

In general, there is an association in temporal variance in the criteria values obtained for different months with the monthly accumulated precipitation averages, which means that the precipitation intensities is one of the main factors to determine the GPM-IMERG performance. As instance for climatic zone B, it is possible to note that the driest months (i.e., months with smaller accumulated precipitation) are from July to November, it presents lower $RMSE$, MAE and $rBIAS_\epsilon$ while the FAR increases and POD and CSI decreases.

The results show that the GPM-IMERG shows the worst criteria values in climatic zone B (Semiarid climate), the driest zone in Brazil, as mentioned in the previous discussions. The $rBIAS_\epsilon$ showed that the GPM-IMERG V06 performed better in regions under Subtropical (Zone C) climates. The Zone A showed the largest amount of *outliers* (Figure 2.7), which was already expected because it encompasses the largest climate area in Brazil, and it should be

pointed out that this zone shows a low density of rain gauges, which may lead to underestimated GPM-IMERG performance (TIAN et al., 2018).

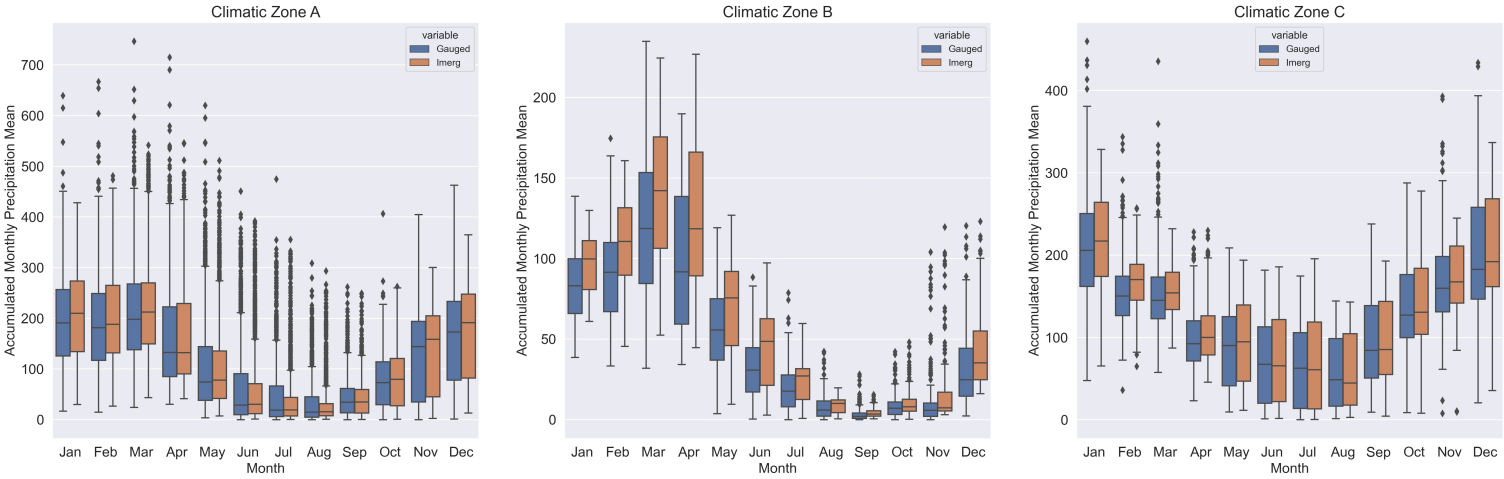

Figure 2.7 – Boxplots of monthly averages of the Gauged and IMERG data based on monthly accumulated precipitation for the climatic zones.

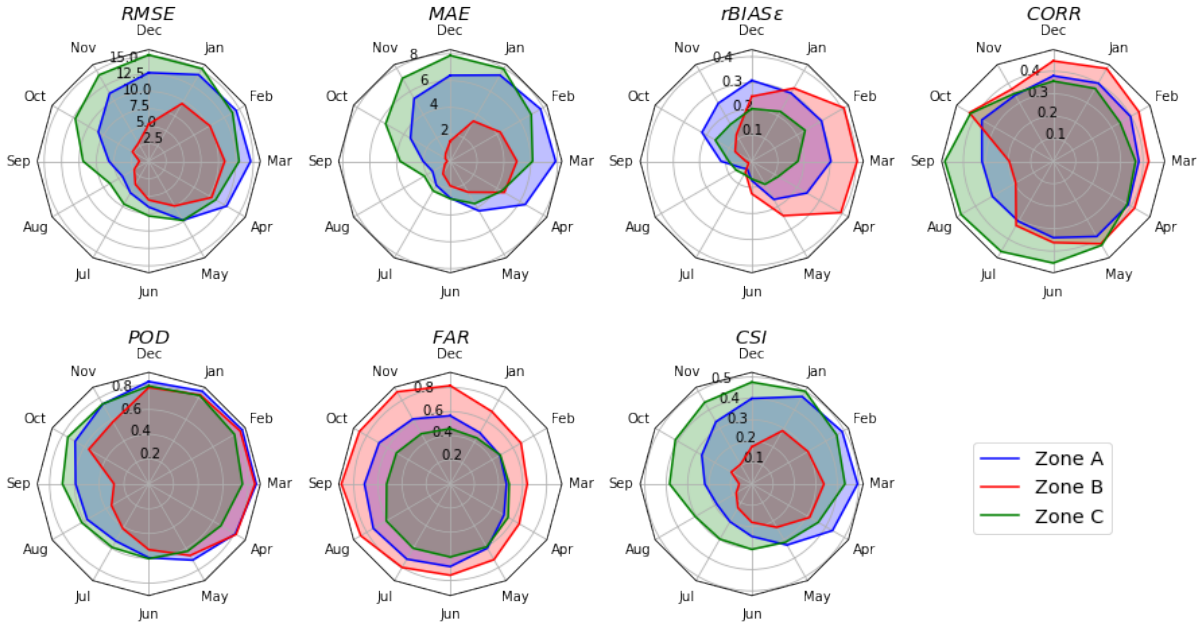

Figure 2.8 – Radar charts of the performance indices for IMERG daily products at the monthly time scale.

2.4 Conclusions

This study has analyzed the performance of the GPM-IMERG V6 final daily precipitation product over Brazil. Firstly, all available rain gauges across the country were selected, which were filtered to select the gauges with a maximum of 10% of missing data for the timeframe of 01 June 2000 to 30 June 2020, resulting in a total of 2,259 rain gauges for the ground-validation. The

selected rain gauge network represents a non-homogeneous density, where some regions have a relatively lower density, bringing an associated limitation as discussed by Tian et al. (2018) that found that lower-density gauge networks might conduct to underestimate the performance of the precipitation product from GPM-IMERG.

Another significant limitation is that this study used the precipitation measured from rain gauges as the *"ground truth"*. Nonetheless, rain gauges are susceptible to systematic error measurement, consisting of losses due to evaporation, wetting, and wind-induced losses, which can vary regardless of the types of measurement (SEVRUK; ONDRÁS; CHVÍLA, 2009). The World Meteorological Organization (WMO) recommends checking and correcting precipitation gauge measurements to eliminate these effects. To overcome this limitation, it was selected only the rain data classified as *consistent* by the Brazilian Water Agency, but it still is possible to have measured precipitation data with errors.

Despite the limitations, proceeding the study started by Gadelha et al. (2019), which have evaluated GPM-IMERG V05 product over Brazil for 2016, this study provided valuable information about the GPM-IMERG for an improved understanding of its strengths and weaknesses over Brazil, by a long-term based evaluation for the sixth version of GPM-IMERG. The main specific findings are summarised as follow:

- The precipitation frequency estimation was underestimated for dry days and well-estimated for precipitation above 5 mm by the GPM-IMERG V06;
- The performance of GPM-IMERG for estimating the amount of precipitation decreases along with the precipitation intensity, where the GPM-IMERG V06 present a tendency to underestimate the amount of precipitation;
- The detection ability of GPM-IMERG is spatially variant, and its estimation ability is more sensitive to longitude when compared to latitude. Also, higher values of FAR were detected in the driest zone, while most lower values of POD were found over the eastern coast. On the other hand, GPM-IMERG V06 showed batter values of $RMSE$, MAE, $rBIAS_{\epsilon}$ in the eastern parts of Brazil;
- The temporal evaluation by the climatic zones showed that GPM-IMERG presented better performances in the region under Subtropical climates and the worst performance in Semiarid climates. Besides, the monthly evaluation showed that the calculated criteria have an association with temporal variance, mainly due to associated precipitation intensities, where the driest months showed worst performances compared to rainy months.

3 EVALUATION OF GPM-IMERG APPLICABILITY FOR HYDROLOGICAL MODELING IN A LARGE-SCALE TROPICAL BASIN

3.1 Introduction

For hydrological modeling applications ground-based stations can be considered the reference data source for precipitation observation (TAPIADOR et al., 2012), and the most adequate for hydrological modeling (KALIN; HANTUSH, 2006; AMORIM et al., 2020). In regional applications, where the spatial distribution of rainfall may play a crucial role, however, there are limitations for using ground-based measurements, especially in most developing countries (MAGHSOOD et al., 2020), in which the rain gauges are insufficient and unevenly distributed (SALIO et al., 2015). Also, frequently the available gauged stations include observation periods shorter than needed, and a few stations have uninterrupted data with most of the time series containing temporal gaps (MONTEIRO et al., 2016; GADELHA et al., 2019).

The estimated precipitation from weather radar may have advantages when compared to conventional rain gauge networks, such as large spatial coverage, high-resolution in space-time monitoring (TAPIADOR et al., 2012). An important consideration is that the sensitivity of the hydrological response and thus the added value of higher resolution rainfall data increases for smaller catchment size, larger catchment spatial variability, smaller storm size, larger storm variability, and higher storm movement velocity (THORNDAHL et al., 2017). Thus, a balance between the benefit of higher accuracy and the required investment obtaining such accuracy is needed for applications. However, a radar system network remains expansive due to high installment costs and maintenance demands bringing significant deployment limitations (LE et al., 2020). On the other hand, Satellite Precipitation Products (SPPS) have been providing a solution for uninterrupted with high spatial resolution precipitation measurements with nearly global coverage. Overall, the use of SPPs as inputs in hydrological models can overcome the imposed limitations from using rain gauge networks, providing a better representation of the spatial variability of precipitation (BITEW; GEBREMICHAEL, 2011; LE et al., 2020).

The strengths of the various groups of SPPs were combined to create the Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM-IMERG) (HUFFMAN et al., 2014). This very high-resolution precipitation product is now available at 0.1° x 0.1° spatial and half-hourly temporal resolutions. Its most recent version has a temporal coverage from June 2000 until the present and introduces significant improvements to the previous versions (HUFFMAN et al., 2019). The performance of GPM-IMERG products through hydrologic simulation has been carried out in different basins of the world and have demonstrated a strong potential application for horological modeling (BITEW; GEBREMICHAEL, 2011; WANG et al., 2017; ZHANG et al., 2019; LE et al., 2020; AMORIM et al., 2020).

Currently, Brazil has an average density of rain gauge per km^2 below the World Mete-

orological Organization (WMO) recommendation (GADELHA et al., 2019). For streamflow observations, the density is even lower leading to limitation of hydrological dynamics understanding over the country. On the other hand, the Doce river basin is one of the most well-monitored large-scale river basin in Brazil, and it encompasses two of the three main climate zones in Brazil (ALVARES et al., 2013).

In this context, the application of SPPs for hydrological modeling purposes can be a possible alternative to improve Brazil's lack of streamflow data and hydrological dynamics understanding. According to that, the main objective of this study was to evaluate the applicability of GPM-IMERG as the precipitation data source for streamflow simulations by comparing with the simulations driven by gauged data from a spatially well-distributed network in a large-scale tropical/subtropical basin.

3.2 Materials and Methods

3.2.1 Study Area

The Doce river basin is located in the southeast part of Brazil, with an area of approximately 86,715 km^2, where 98% is inserted within the Atlantic Forest biome, with a strongly seasonal rainfall pattern (PINTO; LIMA; ZANETTI, 2015) and according to Köppen's climate classification (Figure 3.1 a), in general, the upper-part of the basin (approximately 59% of the basin's area) is submitted to a humid subtropical climate (i.e., Zone C) mainly with dry winter (i.e., Zone Cw) and a hot summer (i.e., Zone Cwa) or a temperate summer (i.e., Zone Cwb) while the lower-part (approximately 59% of the basin's area) is predominantly submitted to a tropical zone with dry winter (i.e., Zone Aw).

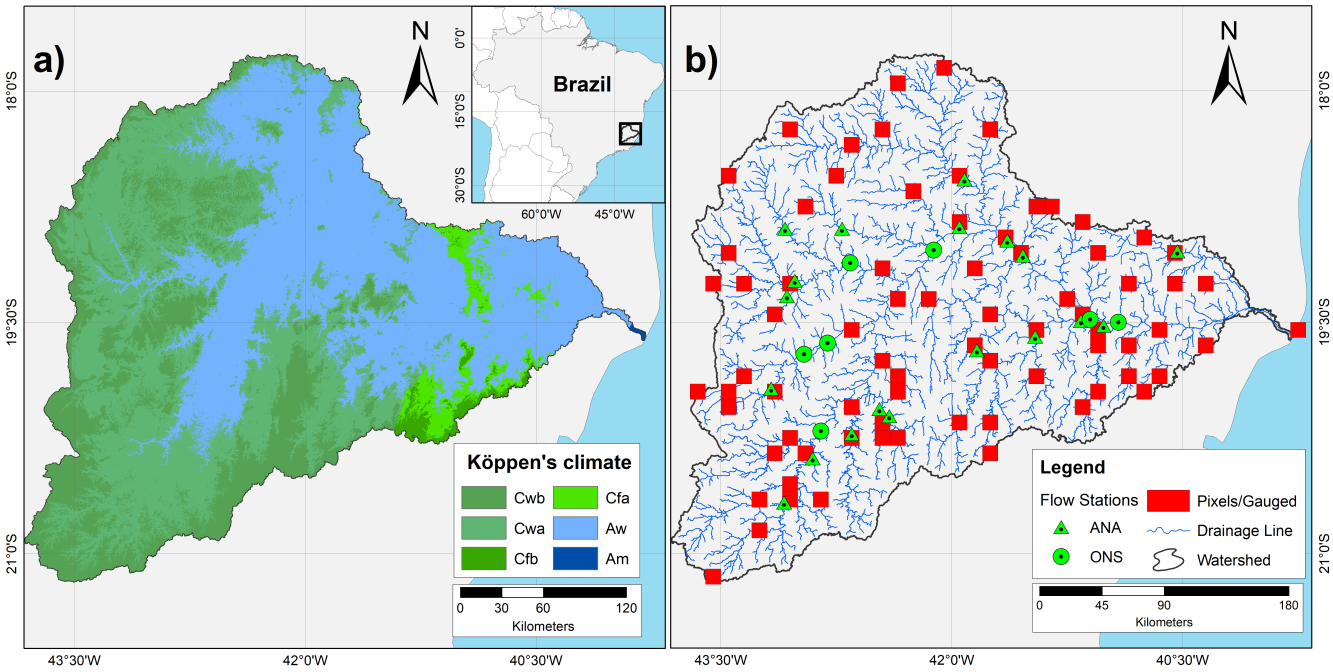

Figure 3.1 – a) Köppen's climate classification map for Doce river basin according to Alvares et al. (2013); b) Spatial distribution of grid points which present rain gauge data frequency of at least 90% in the studied period, and flow stations selected.

Doce river basin is currently one of the most well-monitored large-scale river basins in Brazil, leading a fair evaluation of the satellite precipitation product in the hydrological modeling application, once Tian et al. (2018) found that lower-density gauge networks might conduct to underestimate the performance of the precipitation product from GPM-IMERG.

3.2.2 Precipitation Datasets

The ground-based precipitation dataset at daily intervals over the study area was taken from a collection of 77 rainfall stations (Figure 3.1 b) Brazilian Water Agency (ANA) database (http://www.snirh.gov.br/hidroweb/). It was selected the stations with high temporal overlap with the GPM-IMERG precipitation product. The criteria adopted for selecting the gauges was a threshold of a maximum of 10% of missing data in the precipitation time series from June 2000 to December 2019.

The satellite precipitation product used in this study was the final run of daily GPM-IMERG V6. The IMERG-Final run provides the highest quality precipitation estimates with a latency of 3.5 months and a spatial resolution of 0.1° x 0.1°. It is provided by the Global Precipitation Measurement (GPM) mission, which is a collaboration between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) (HUFFMAN et al., 2019). As used by Amorim et al. (2020), the grids containing the GPM-IMERG data into the boundaries of the study area were transformed into points representing the centroid of each pixel, and the daily precipitation historical series obtained at

each point of the GPM-IMERG were used as input in the hydrological model.

3.2.3 GPM-IMERG Evaluation

In order to have a comparative overview of the GPM-IMERG V 06 performance for the study area, a ground validation by using the point-to-pixel approach was performed (XU et al., 2017; GADELHA et al., 2019; MAHMOUD; HAMOUDA; MOHAMED, 2019; MAGHSOOD et al., 2020). For the point-to-pixel approach, only the GPM-IMERG 0.1° x 0.1° pixels that had a gauged station within was selected to be compared. Thus, a database with the rain gauge data set pairs and the selected GPM-IMERG grids were formed. For this analysis were selected the time frame from 01 June 2000 to 30 June 2020, and 73 gauged stations/grids.

The evaluation was divided into two steps: i) the GPM-IMERG detection rainfall occurrence ability; and ii) GPM-IMERG quantitative differences and agreement with the ground-based data, both steps were performed using all selected grid points lumped together.

3.2.3.1 Rainfall Occurance

The first group of performance indexes applied is related to the effectiveness of GPM-IMERG rainfall events detection (WILKS, 2011). This group includes Probability of Detection (POD), which exhibits the fraction of rainfall events correctly detected; False Alarm Ratio (FAR), which gives the fraction of false alarm detected; and Critical Success Index (CSI), which combines the characteristics of false alarms and missed events, either rain occurrence of zero-rainfall. Table 3.1 shows the summary of these indices.

Performance Indices	Equation	Equation Meaning	Range	Optimal Score
POD	$POD = a/(a+c)$	Fraction of rain occurrences correctly detected	0–1	1
FAR	$FAR = b/(a+b)$	Fraction of false alarm detected	0–1	0
CSI	$CSI = a/(a+b+c)$	Fraction of success in events detection	0–1	1

where a is the number of times that rainfall events was observed and correctly detected; b is the number of times that rainfall was detected but not observed; c is the number of times that rainfall was observed but not detected.

Table 3.1 – Summary of the performance indices for events detection

3.2.3.2 Rainfall Quantitative Difference

A set of statistical metrics were used to compare differences between the estimated rainfall from GPM-IMERG with the rain gauge data. The metrics include: Mean error (ME), which represents the average magnitude of the satellite error; Root Mean Square Error ($RMSE$), which gives the sample standard deviation of the differences between the GPM-IMERG V06 product and the observed rainfall; modified relative Bias ($rBias_\epsilon$), which shows the deviation of

GPM-IMERG V06 from the ground-observed rainfall; and to describe the agreement between GPM-IMERG estimates and the ground-based data, the Pearson Correlation Coefficient (CC).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|S_i - G_i| \tag{3.1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - G_i)^2} \tag{3.2}$$

$$rBias_{\epsilon} = 2\frac{S_i - G_i}{\epsilon - (S_i - G_i)} \tag{3.3}$$

$$CC = \frac{\sum_{i=1}^{n}(G_i S_i) - (\sum_{i=1}^{n} G_i)(\sum_{i=1}^{n} S_i)}{\sqrt{\left[(\sum_{i=1}^{n} G_i^2) - (\sum_{i=1}^{n} G_i)^2\right]\left[(\sum_{i=1}^{n} S_i^2) - (\sum_{i=1}^{n} S_i)^2\right]}} \tag{3.4}$$

where Gi and Si are the ground-base rain gauge and satellite-based precipitation data at pixel i and n is the total number of pixels.

For the modified relative bias (Equation 3.3), Hashemi et al. (2017) defined ϵ as a value to be used to regularize grids where very low values of precipitation can result in a very large relative bias. The suggested value for ϵ is $0.5mm/day$, which can adequately large the huge relative bias for low-precipitation events, and was adopted in this study.

The CC (Equation 3.4) gives the degree of linear association between two variables, ranging from +1 to -1, where +1 represents the total positive linear correlation, and -1 is the total negative.

3.2.4 Hydrological Modeling

3.2.4.1 Model Description

The hydrological model applied in this study was the MGB-IPH, chosen by the successful application in South American watersheds for different study purposes (SILVA et al., 2007; SIQUEIRA et al., 2018; PONTES et al., 2017; SORRIBAS et al., 2016; PAIVA et al., 2013; GETIRANA et al., 2011). The MGB-IPH is a conceptual, semi-distributed, large-scale hydrological model composed of four calculation modules: soil water balance, evapotranspiration, flow calculation (surface, sub-surface and subterranean), and flow propagation in the drainage network. (COLLISCHONN et al., 2007). In the latest versions (PAIVA; COLLISCHONN; TUCCI, 2011; PONTES et al., 2017), basins are discretized into smaller units for simulation (i. e. unit catchments), each one has a unique drainage network, where the river routing process is performed, and combinations of soil type and land use categorized as Hydrological Response Units (HRU). The vertical water and energy balance are calculated independently for each HRU

of each unit catchment, and its results are propagated to downstream unit catchment using linear reservoirs. The vertical water and energy balance includes canopy interception estimated in terms of leaf area index, evapotranspiration calculated by the Penman-Monteith method, surface runoff produced using the variable contribution area concept, groundwater computed by linear functions, and subsurface flows that uses nonlinear equations.

The model parameters are composed of four fixed parameter values and six calibrated parameter values, where such values are based on the HRU characteristics. A more comprehensive description of the MGB-IPH model may be found in Collischonn et al. (2007). In the current version, flow routing through the drainage network can be computed using the Muskingum-Cunge method (COLLISCHONN et al., 2007) or the inertial method (PONTES et al., 2017). The Muskingum-Cunge method was applied in this work since it can present a lesser computational cost and was appropriate and able to represent the hydrological process effectively in the Doce river basin (FAGUNDES; FAN; PAIVA, 2019).

3.2.4.2 Model Setup

The physiographic information of the sub-basins was obtained using Geographic Information System (GIS), and all geoprocessing steps were conducted using the IPH-Hydro tools (SIQUEIRA et al., 2016). It was used the Shuttle Radar Topographic Mission (SRTM) images by 3 arc-seconds (approximately 90 meters) spatial resolution (USGS, 2015) as the Digital Elevation Model (DEM), which is sufficient to meet the demands related to the water balance simulated by the MGB-IPH model.

In order to support the calibration of the model, it was defined ten sub-basins (Figure 3.2 a), subdivided into 2,443 unit catchments. The composition of the type of land cover and soil type was defined (Figure 3.2 b) using the HRU product from Fan et al. (2015).

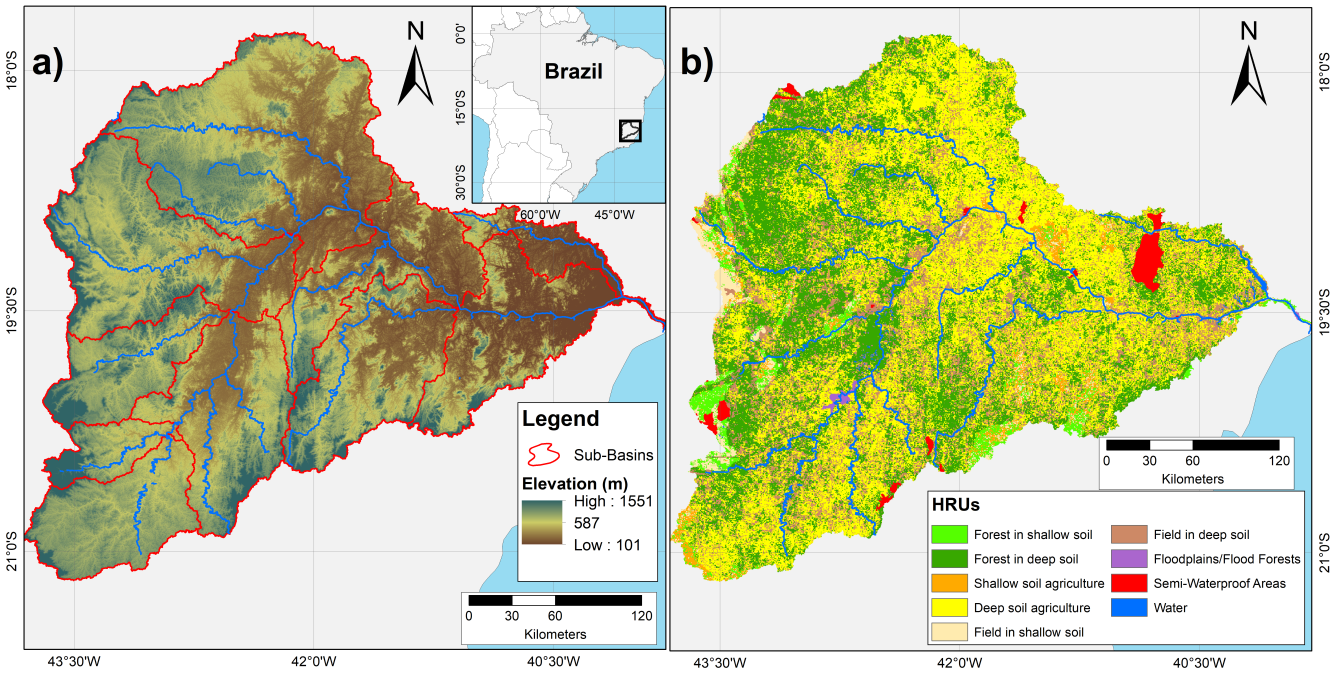

Figure 3.2 – (a) Sub-basins and and spatial distribution of elevation. (b) Spatial distribution of , adapted from Fan et al. (2015).

Flow data from stations in the Doce river basin were obtained from the ANA. It was selected only the stations without the influence of reservoirs and with a maximum of 10% threshold of missing data the time frame from June 2000 to December 2015, resulting in 19 stations. Seven stations downstream to the reservoirs with estimated naturalized flows from National Electric System Operator (ONS) were selected in addiction. These 26 stations (Figure 3.1 b) were used for comparison with the MGB-IPH simulation results. Code identification for the selected flow stations are displayed in Figure 6.1 (Supplementary Materials - Chapter 6).

The selected time frame (i.e., June 2000 to December 2015) was based on the temporal validity of the flow rating curves from ANA since the observed flow time series from its stations are estimated using rating curves applying stage measurements.

Naturalized streamflows can be defined as the flow that occurs if there were no anthropic actions that alter the hydrological regime, which means a river without reservoir operation, nor evaporation from artificial lakes, and nor withdrawals for water supply (GUILHON; ROCHA; MOREIRA, 2007). Naturalized streamflows have been succesfully used in a range of hydrological modeling studies in Brazil such as Amorim et al. (2020), Cassalho et al. (2020) and Nóbrega et al. (2011).

In order to reduce the subjective nature and have a fair comparison between the MGB-IPH simulation results using each source of precipitation as model input (i.e., using GPM-IMERG and gauged-data separately), the model was calibrated using the Multi-objective Complex Evolution–University of Arizona (MOCOM-UA) optimization algorithm developed by Yapo, Gupta e Sorooshian (1998). In order to have meaningful physically based parameters, the optimization

algorithm was bounded by a range of MGB-IPH calibrated parameters for similar basins in Brazil found by Siqueira et al. (2016).

This study ran the MGB-IPH model on daily time scale, selecting the period from June 2000 to December 2001 as the warm-up period, the next seven years (2002–2008) as the calibration period, and the last seven years (2009–2015) as the validation period. The calibration step was performed separately for both rain gauge-based and GPM-IMERG data.

3.2.4.3 Performance Evaluation Criteria

In this study, the same performance evaluation criteria were adopted for both calibration and validation periods as recommended by Moriasi et al. (2015). The performances at spatial and temporal were also observed in calibration and validation periods in order to identify the consistency of model computations for the selected time frames since substantial differences in performance during the calibration and validation periods may indicate the need for further calibration (MORIASI et al., 2015).

For optimization algorithm during calibration step, the performance evaluation criteria was driven by objective functions while for simulation results quantitative and qualitative analyses were utilized for both calibration and validation were utilized. The qualitative analyses, the simulated hydrograph was compared with the observed ones, while objective functions measured the quantitative performance.

The objective functions used in this study were the Nash-Sutcliffe efficiency (NSE) (NASH; SUTCLIFFE, 1970); the Nash-Sutcliffe efficiency calculated using logarithms (NS_{log}) in order to analyze the simulation efficiency by removing the weight of the maximum observed flows; the difference in volume in percentage ($PBIAS$), to quantify the proportion not represented by the model; and the Kling-Gupta efficiency (KGE) (GUPTA et al., 2009), which combines the three components of NSE (i.e., correlation, bias, the ratio of variances or coefficients of variation) in a more balanced way.

$$NS = 1 - \frac{\sum_{i=1}^{n} \left(Q_{obs} - Q_{sim}\right)^2}{\sum_{i=1}^{n} \left(Q_{obs} - \overline{Q_{obs}}\right)^2} \tag{3.5}$$

$$NS_{log} = 1 - \frac{\sum_{i=1}^{n} \left(\log(Q_{obs}) - \log(Q_{sim})\right)^2}{\sum_{i=1}^{n} \left(\log(Q_{obs}) - \log(\overline{Q_{obs}})\right)^2} \tag{3.6}$$

$$PBIAS = 100 * \frac{\sum_{i=1}^{n} Q_{sim} - \sum_{i=1}^{n} Q_{obs}}{\sum_{i=1}^{n} Q_{obs}} \tag{3.7}$$

where Q_{obs} and Q_{sim} are the observed and simulated daily discharge, respectively and $\overline{Q_{obs}}$ is

the average observed discharge.

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \tag{3.8}$$

where r the linear correlation coefficient between the simulated and observed flows; α is a measure of relative variability in the simulated and observed values; and β is the ratio between the mean simulated and mean observed flows.

3.3 Results and Discusssion

3.3.1 GPM-IMERG Evaluation Overview

The first analysis was made regarding the monthly average accumulated rainfall volume for the selected grids in the study area. Figure 3.3 shows *boxplots* of the monthly averages for the ground-based and GPM-IMERG data. We can see that the GPM-IMERG could capture the seasonal rainfall pattern variations of precipitation. However, in the first place, GPM-IMERG overestimates the amount of precipitation associated with a determined frequency. For instance, the median of the monthly accumulated precipitation was predominantly higher for GPM-IMERG than the ground-based data.

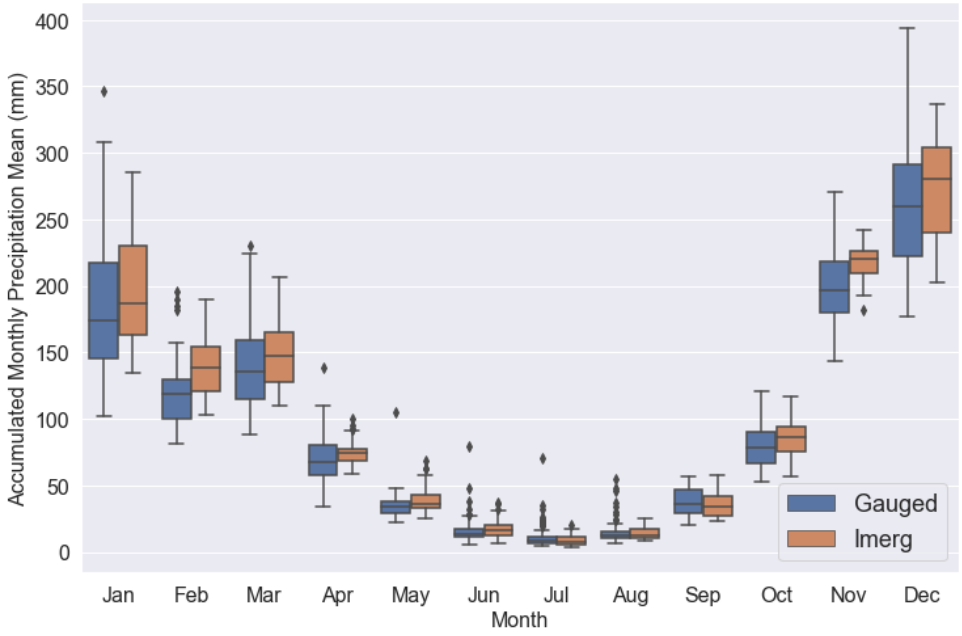

Figure 3.3 – Boxplots of monthly averages of the Gauged and GPM-IMERG data based on monthly accumulated precipitation for study area

The Figure 3.4 brings the rainfall detection ability of GPM-IMERG trough spatial maps using POD, FAR, and CSI metrics to investigate how the metrics vary for the study area. The criteria indices were mapped using the Inverse of Distance Weighting method in $ArcMap$ 10.3.1 environment.

According to the Figure 3.4 (a), the values of POD mainly were higher than 0.7 or around it showing a satisfactory precipitation detection by GPM-IMERG while the values of FAR (Figure 3.4 b) ranged between 0.37 and 0.6 with the upper-part of the basin showing lower values than the lower-part. This difference can be associated with the number of dry days during the study period since the upper part of the basin showed fewer dry days than the lower part, directly affecting the FAR values. As expected, a similar behavior was observed for CSI (Figure 3.4 c), once it combines the characteristics of FAR and the missed events.

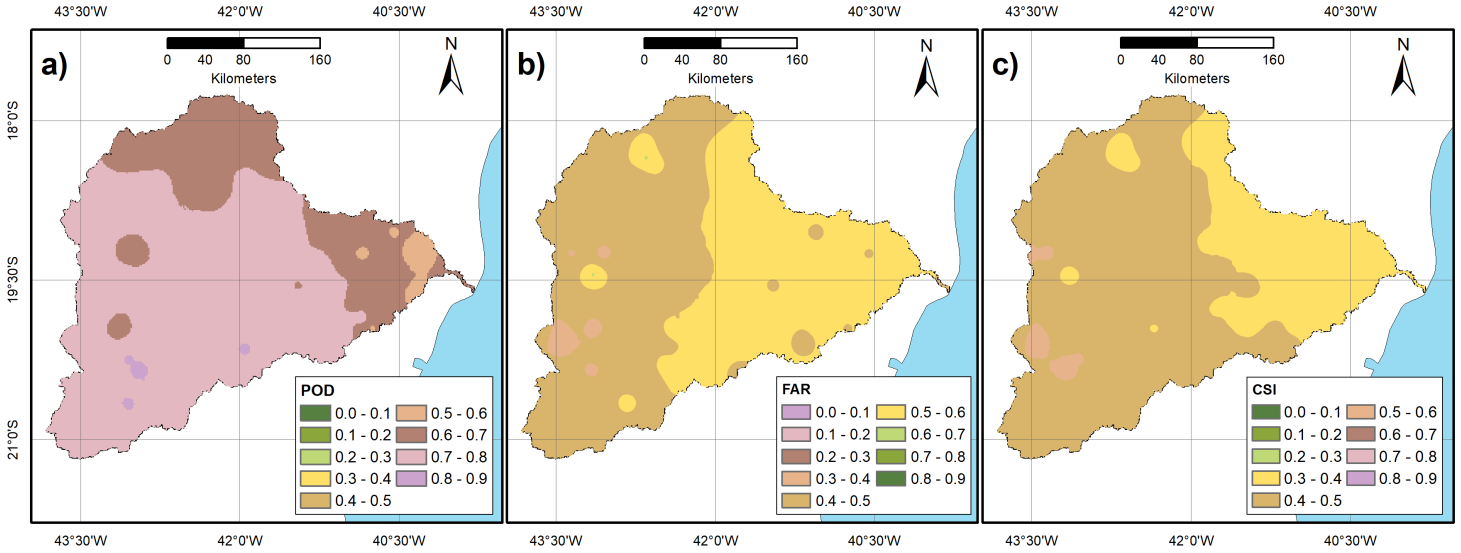

Figure 3.4 – Spatial distribution of (a) Probability of Detection - POD, (b) False Alarm Ration - FAR, and (c) Critical Success Index - CSI.

Figure 3.5 shows the empirical cumulative distribution function ($ECDF$) for GPM-IMERG and the gauged stations. GPM-IMERG underestimates the frequency of precipitation lower than 30 mm, with the GPM-IMERG frequency underestimation increasing with lower precipitation values. For precipitation above 30 mm, the frequency is well-estimated.

Table 3.2 presents overall performance measures of GPM-IMERG, while Table 3.3 brings performance measures for different precipitation thresholds following the classification suggested by (XU et al., 2017) and adapted for Brazil based on Rozante et al. (2018).

GPM-IMERG presents a positive $rBIAS_\epsilon$ when it is calculated using all the data, saying that its overestimates precipitation, but by Table 3.3 it is possible to note that GPM-IMERG underestimates precipitation for intensities higher than 0.5 mm. Besides, the $rBIAS_\epsilon$ decreases when the precipitation intensities decrease, which means that there is a higher underestimation for a higher amount of precipitation. It is possible to associate the GPM-IMERG underestimation for precipitation lesser 0.5 mm since there is an identified FAR of 0.49.

In the same way of $rBIAS_\epsilon$, due to the increased GPM-IMERG underestimation for higher amounts of precipitation, the $RMSE$ and MAE increases with the precipitation inten-

sities. On the other hand, the POD also increases with precipitation, which means there are greater chances of GPM-IMERG detecting higher precipitation amounts.

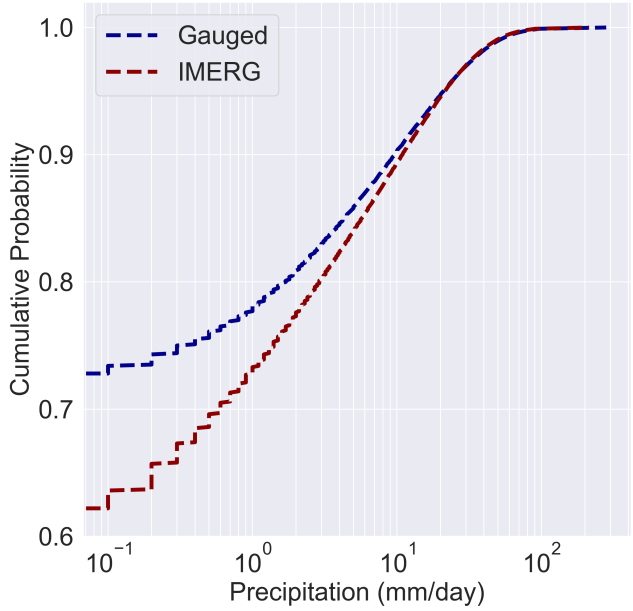

Figure 3.5 – Empirical cumulative distribution function of gauged data precipitation at stations and GPM-IMERG for study area.

Perfomance measures	Values
$RMSE$	9.92
MAE	3.70
$rBIAS_{\epsilon}$	0.12
$CORR$	0.50
POD	0.71
FAR	0.49
CSI	0.42

Table 3.2 – Performance measures for the paired gauge-satellite dataset.

	Daily Precipitation (P) Thresholds (mm)						
Metrics	Rain/no-rain	Light	Low moderate	Moderate	Heavy	Extreme	Torrential
	P≤0.5	0.5≤P<2	2≤P<5	5≤P<20	20≤P<40	40≤P<90	P≥90
$RMSE$	5.17	9.20	10.74	14.02	21.78	38.65	80.00
MAE	1.30	3.99	5.84	10.11	18.50	34.00	72.56
$rBIAS_{\epsilon}$	0.34	-0.25	-0.49	-0.63	-0.76	-0.91	-1.02
$CORR$	0.05	0.05	0.06	0.16	0.11	0.15	0.14
POD	0.43	0.53	0.65	0.80	0.92	0.96	0.98

Table 3.3 – Performance measures for the paired gauge-satellite dataset for different precipitation intensity thresholds.

3.3.2 Hydrological Modeling Perfomance

The performance measures for daily streamflow MGB-IPH simulation, driven by both precipitation datasets (i.e., gauged and GPM-IMERG) was analysed using *boxplots* for calibration and validation (Figure 3.6) and by a classified spatial distribution of performances (Figure 3.7 and Figure 3.8). Performance metrics values for each flow station for both calibration and validation steps using gauged and GPM-IMERG as precipitation input data are displayed in Supplementary Materials (Chapter 6).

Overall, based on the Figure 3.3, the rain gauge-based simulations exhibited better performances than IMERG-based simulations for NSE, NSE_{log} and KGE while the best $PBIAS$ values were found in IMERG-based simulations for both calibration and validation steps. The lower values of NSE and KGE for rain IMERG-based simulations are possibly associated with GPM-IMERG underestimations, leading the MGB-IPH simulations not to reach the high peaks of discharges while GPM-IMERG false alarm is probably impacting on NS_{log} by inputting precipitation during the low flow period.The lower performance for high peaks and low flow periods is likely to compensate and drive towards lower $PBIAS$ values for IMERG-based simulations. When the comparison is between the simulation steps, the model performance was better in the calibration than in the validation period, as already expected.

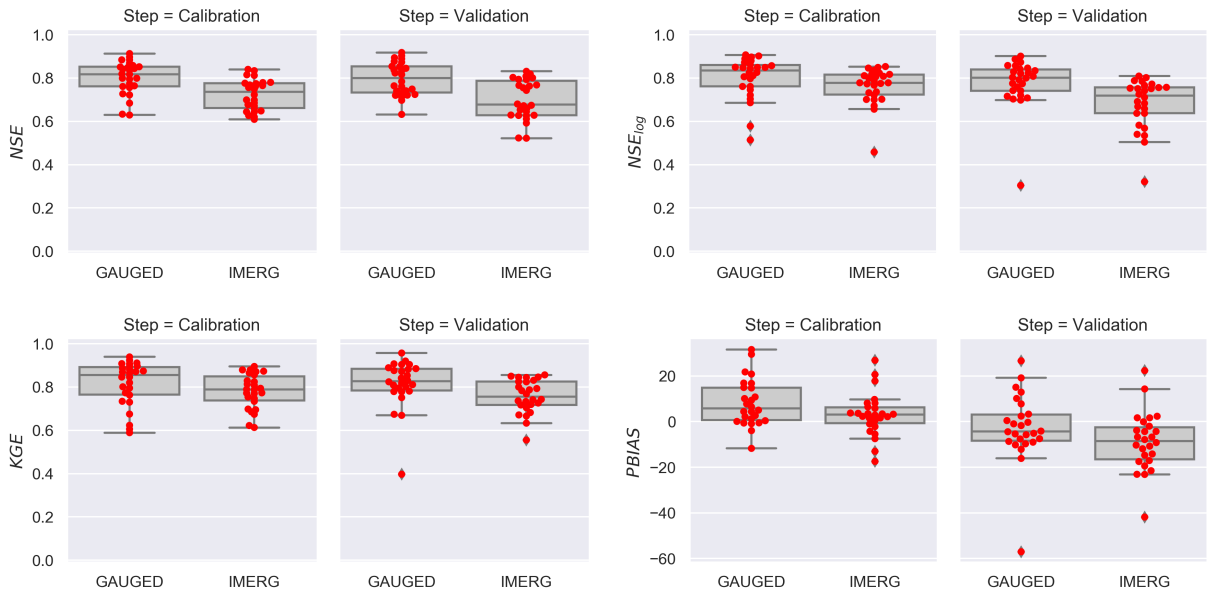

Figure 3.6 – MGB-IPH model performances for both gauged and IMERG as precipitation input data.

For rain gauged-based simulations the metrics median were mostly at *Very Good* levels (MORIASI et al., 2015) for both calibration and validation steps while for IMERG-based simulations the metrics median ranging from *Good* to *Satisfactory* levels (Table 3.4), which means that GPM-IMERG can serve as alternative input to enhance the performance of hydrological models.

Metric	Calibration		Validation	
	gauge-based	IMERG-based	gauge-based	IMERG-based
NSE	0.82	0.74	0.80	0.68
NSE_{log}	0.83	0.78	0.80	0.72
KGE	0.86	0.79	0.83	0.76
$PBIAS$	5.76	3.05	-4.3	-8.5

Table 3.4 – Performance measures median

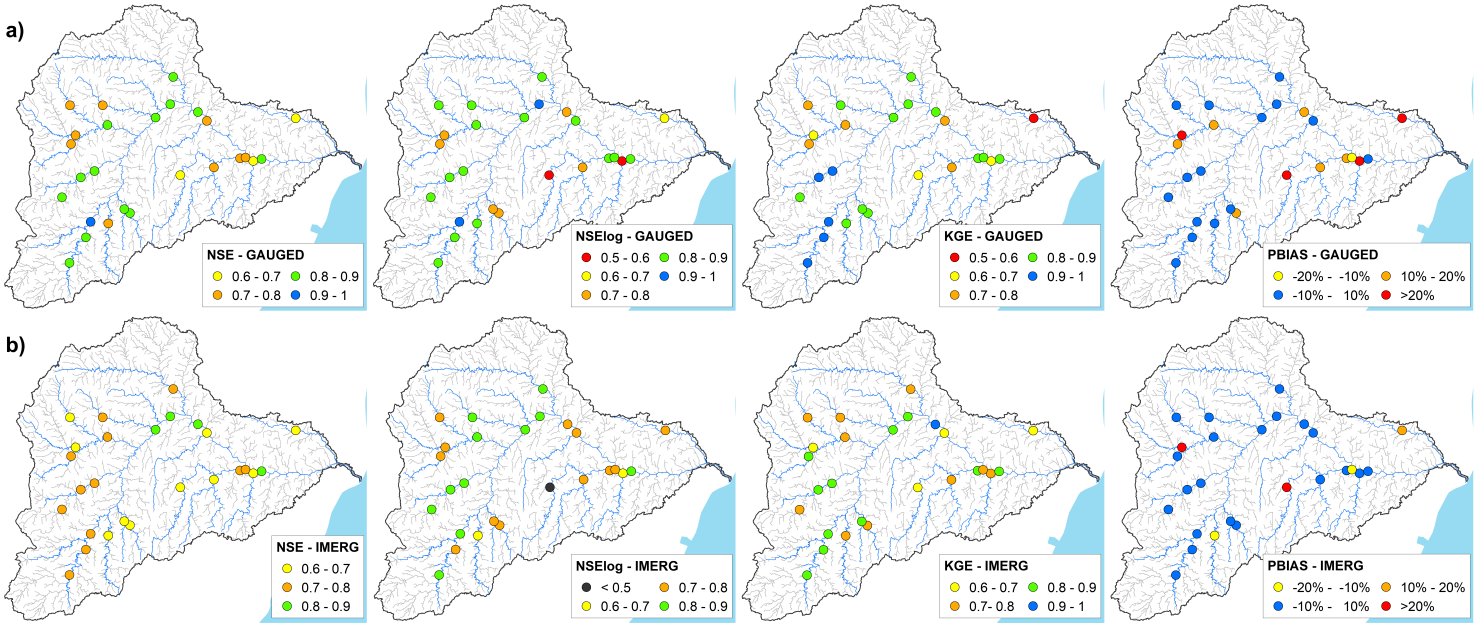

Figure 3.7 – MGB-IPH model spatial performances for both (a) gauged and (b) IMERG as precipitation input data for the calibration period.

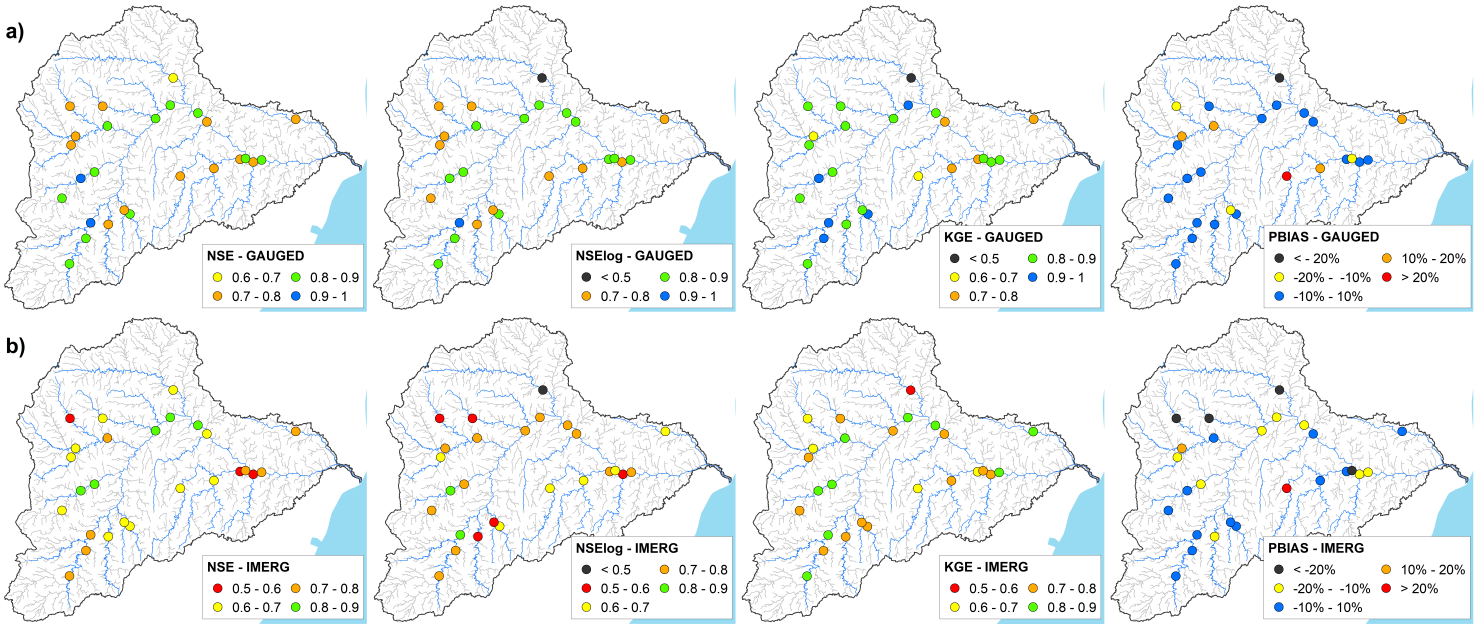

Figure 3.8 – MGB-IPH model spatial performances for both (a) gauged and (b) IMERG as precipitation input data for the validation period.

Figure 3.7 and Figure 3.8 shows metrics performances spatially distributed through the study area for calibration and validation steps, respectively. It was not possible to identify any spatial pattern for MGB-IPH simulation performances. Generally, the stations with lower metrics for rain gauge-based simulations also had lower metrics for IMERG-based simulations. The rain gauge network density used in this study was sufficient to capture the spatial distribution of rainfall (see Figure 3.1 b) at the study area leading to a fair comparison with GPM-IMERG. It is

possible to say that IMERG-based simulation could drive the hydrologic model to capture the hydrological cycle through the study area. These results support the GPM-IMERG applicability for hydrological modeling purposes.

Figure 3.9 shows the daily simulated discharge using both precipitation datasets for some representative stations, and the remaining stations are presented in the Supplementary Materials (Chapter 6. According to the hydrographs presented, there is a notable agreement most of the simulation time in both calibration and validation steps. There are no significant differences between the simulations using ground-based and IMERG-based precipitation for medium and low flows. On the other hand, differences in high flows were noted between the simulations since IMERG-based simulations underestimated the majority of high flows, likely associated with GPM-IMERG precipitation underestimation for high intensities.

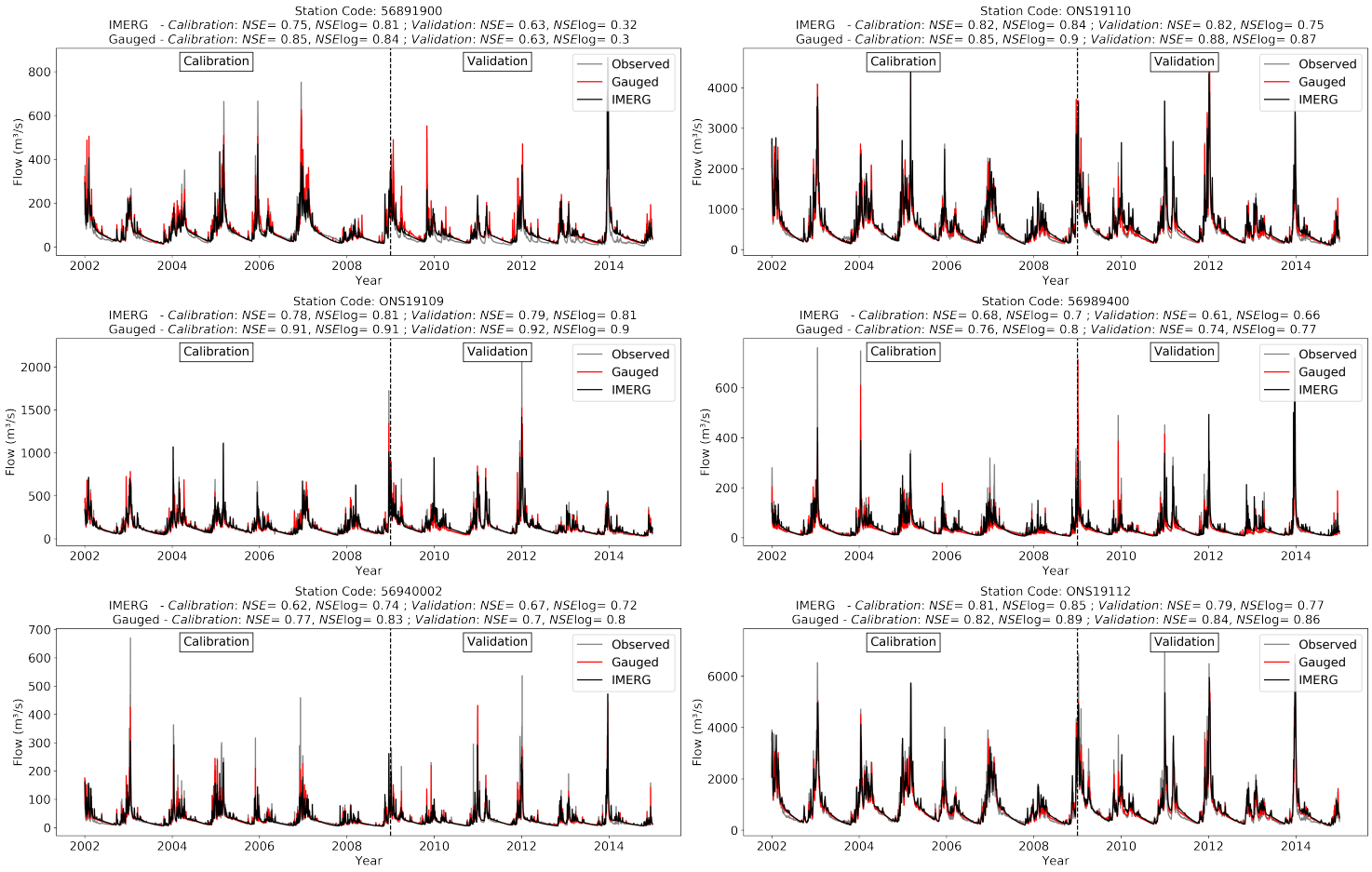

Figure 3.9 – Comparison between rain gauged-based (red) and IMERG-based (black) daily simulations with the observed (gray) discharge.

3.4 Conclusions

This study evaluated the GPM-IMERG applicability for hydrological modeling purposes in tropical/subtropical basins. The work was divided into two main parts: i) Comparative overview of GPM-IMERG precipitation estimation based on gauged-based stations, and ii) using gauged-

based and IMERG-based precipitation datasets to simulate daily discharges using the MGB-IPH model at 26 different stations in the Doce river basin.

In the context of using precipitation datasets for hydrological modeling applications, observed streamflow data plays a dominant role in model parameters calibration and validation of the simulated outputs, which means reliable streamflow observations have utmost importance for the trustworthiness of results. An important limitation of the study was the set of observed streamflow data used that was provided from ANA, which was estimated using rating curves, and ONS, which was estimated by flow naturalization methods. Both methods have limitations and bring estimation uncertainties with them. Moreover, flows from the rating curve bring highly variable uncertainty over time and stage within each gauge and between gauges, which can be propagated differently for practical applications (DOMENEGHETTI; CASTELLARIN; BRATH, 2012; TOMKINS, 2014).

Despite the limitations, this study could find valuable information about GPM-IMERG application for hydrological modeling purposes. The main specific findings can be summarised as follow:

- GPM-IMERG V06 was capable of capturing the seasonal rainfall pattern variations and showed satisfactory precipitation detection. On the other hand, the false alarm ratio was different from the upper part to the lower part of the study area, which can be associated with the number of dry days (i.e., more dry days, higher the false alarm);
- GPM-IMERG V06 presented an increasing trend of underestimation of e precipitation amount along with precipitation intensity;
- This study determines the ability of GPM-IMERG V06 to estimate rainfall and produce input data for streamflow simulations in a tropical/subtropical large-scale river basin;
- GPM-IMERG V06 could drive the hydrological model to capture the hydrological cycle to represent spatial and temporal streamflow variations.

4 CONCLUSIONS

This study has analyzed the performance of the GPM-IMERG V6 final daily precipitation product over Brazil and the GPM-IMERG applicability for hydrological modeling purposes in tropical/subtropical basins.

Despite the limitations encountered along with the study, valuable information about GPM-IMERG performance and application for hydrological modeling purposes was found. The main finds for GPM-IMERG performance over Brazil are; i) the GPM-IMERG present a tendency to underestimate the amount of precipitation for higher precipitation intensities; ii) the detection ability of GPM-IMERG is spatially variant and more sensitive to longitude when compared to latitude; iii) GPM-IMERG presented better performances in regions under subtropical climates and the worst performance in semiarid climates. For GPM-IMERG application for hydrological modeling purposes, the main finds were that GPM-IMERG could drive the hydrological model to capture the seasonal rainfall pattern and represent the spatial and temporal streamflow variations.

REFERÊNCIAS

ALVARES, C. A. et al. Köppen's climate classification map for brazil. *Meteorologische Zeitschrift*, Stuttgart, v. 22, n. 6, p. 711–728, 2013. Citado 6 vezes, páginas 8, 16, 17, 18, 30 e 31.

AMORIM, J. d. S. et al. Evaluation of satellite precipitation products for hydrological modeling in the brazilian cerrado biome. *Water*, Multidisciplinary Digital Publishing Institute, v. 12, n. 9, p. 2571, 2020. Citado 5 vezes, páginas 13, 15, 29, 31 e 35.

ASONG, Z. E. et al. Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against Ground Precipitation Observations: A Preliminary Assessment. *Journal of Hydrometeorology*, American Meteorological Society, v. 18, n. 4, p. 1033–1050, apr 2017. ISSN 1525-7541. Citado na página 16.

BHUIYAN, M. A. E. et al. Machine learning-based error modeling to improve gpm imerg precipitation product over the brahmaputra river basin. *Forecasting*, Multidisciplinary Digital Publishing Institute, v. 2, n. 3, p. 248–266, 2020. Citado na página 25.

BITEW, M. M.; GEBREMICHAEL, M. Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. *Water Resources Research*, Wiley Online Library, v. 47, n. 6, 2011. Citado 2 vezes, páginas 13 e 29.

CASSALHO, F. et al. Hydrologic validation of merge precipitation products over anthropogenic watersheds. *Water*, Multidisciplinary Digital Publishing Institute, v. 12, n. 5, p. 1268, 2020. Citado na página 35.

CECINATI, F. et al. Representing radar rainfall uncertainty with ensembles based on a time-variant geostatistical error modelling approach. *Journal of Hydrology*, Elsevier, v. 548, p. 391–405, 2017. Citado na página 15.

CHEN, C. et al. Multiscale comparative evaluation of the gpm imerg v5 and trmm 3b42 v7 precipitation products from 2015 to 2017 over a climate transition area of china. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 10, n. 6, p. 944, 2018. Citado na página 25.

COLLISCHONN, W. et al. The MGB-IPH model for large-scale rainfall-runoff modeling. *Hydrological Sciences Journal*, v. 52, p. 878–895, 2007. Citado 2 vezes, páginas 33 e 34.

DOMENEGHETTI, A.; CASTELLARIN, A.; BRATH, A. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrology and Earth System Sciences*, v. 16, n. 4, p. 1191–1202, 2012. ISSN 10275606. Citado na página 43.

FAGUNDES, H. d. O.; FAN, F. M.; PAIVA, R. C. D. de. Automatic calibration of a large-scale sediment model using suspended sediment concentration, water quality, and remote sensing data. *RBRH*, Associação Brasileira de Recursos Hídricos, v. 24, apr 2019. ISSN 1414-381X. Citado na página 34.

FALCK, A. S. et al. Propagation of satellite precipitation uncertainties through a distributed hydrologic model: A case study in the tocantins–araguaia basin in brazil. *Journal of Hydrology*, Elsevier, v. 527, p. 943–957, 2015. Citado na página 15.

FAN, F. M. et al. Um mapa de unidades de resposta hidrológica para a américa do sul. *XXI Simpósio Brasileiro de Recursos Hídricos, Novembro de*, 2015. Citado 3 vezes, páginas 8, 34 e 35.

FRANZ, K. J.; HOGUE, T. Evaluating uncertainty estimates in hydrologic models: borrowing measures from the forecast verification community. *Hydrology and Earth System Sciences*, v. 15, n. 11, p. 3367, 2011. Citado na página 15.

GADELHA, A. N. et al. Grid box-level evaluation of imerg over brazil at various space and time scales. *Atmospheric Research*, Elsevier, v. 218, p. 231–244, 2019. Citado 9 vezes, páginas 13, 15, 16, 19, 23, 28, 29, 30 e 32.

GETIRANA, A. et al. Assessment of different precipitation datasets and their impacts on the water balance of the Negro River basin. *Journal of Hydrology*, Elsevier, v. 404, n. 3-4, p. 304–322, jul 2011. ISSN 0022-1694. Citado na página 33.

GUILHON, L. G. F.; ROCHA, V. F.; MOREIRA, J. C. Comparação de métodos de previsão de vazões naturais afluentes a aproveitamentos hidroelétricos. *Revista Brasileira de Recursos Hídricos*, v. 12, n. 3, p. 13–20, 2007. Citado na página 35.

GUPTA, H. V. et al. Decomposition of the mean squared error and nse performance criteria: Implications for improving hydrological modelling. *Journal of hydrology*, Elsevier, v. 377, n. 1-2, p. 80–91, 2009. Citado na página 36.

HASHEMI, H. et al. Bias correction of long-term satellite monthly precipitation product (trmm 3b43) over the conterminous united states. *Journal of Hydrometeorology*, v. 18, n. 9, p. 2491–2509, 2017. Citado 3 vezes, páginas 18, 20 e 33.

HOBOUCHIAN, M. P. et al. Assessment of satellite precipitation estimates over the slopes of the subtropical andes. *Atmospheric Research*, Elsevier, v. 190, p. 43–54, 2017. Citado na página 15.

HOSSEINI-MOGHARI, S.-M.; TANG, Q. Validation of gpm imerg v05 and v06 precipitation products over iran. *Journal of Hydrometeorology*, v. 21, n. 5, p. 1011–1037, 2020. Citado na página 25.

HUFFMAN, G. et al. Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg) algorithm theoretical basis document (atbd) version 4.4. *Version*, v. 4, p. 26, 2014. Citado 3 vezes, páginas 13, 16 e 29.

HUFFMAN, G. et al. *Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation. NASA Tech. Doc., 54 pp*. 2017. Citado na página 18.

HUFFMAN, G. et al. *GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC)*. 2019. Citado 5 vezes, páginas 13, 16, 18, 29 e 31.

JORDAN, P.; SEED, A.; AUSTIN, G. Sampling errors in radar estimates of rainfall. *Journal of Geophysical Research: Atmospheres*, Wiley Online Library, v. 105, n. D2, p. 2247–2257, 2000. Citado na página 15.

KALIN, L.; HANTUSH, M. M. Hydrologic modeling of an eastern pennsylvania watershed with nexrad and rain gauge data. *Journal of Hydrologic Engineering*, American Society of Civil Engineers, v. 11, n. 6, p. 555–569, 2006. Citado 3 vezes, páginas 13, 15 e 29.

KOUSKY, V. E. Pentad outgoing longwave radiation climatology for the south american sector. *Revista Brasileira de Meteorologia*, Citeseer, v. 3, n. 1, p. 217–231, 1988. Citado na página 17.

KOUSKY, V. E.; GAN, M. A. Upper tropospheric cyclonic vortices in the tropical south atlantic. *Tellus*, Taylor & Francis, v. 33, n. 6, p. 538–551, 1981. Citado na página 17.

LE, M.-H. et al. Adequacy of satellite-derived precipitation estimate for hydrological modeling in vietnam basins. *Journal of Hydrology*, Elsevier, p. 124820, 2020. Citado 4 vezes, páginas 13, 15, 16 e 29.

LU, X. et al. Evaluation and correction of the trmm 3b43v7 and gpm 3imergm satellite precipitation products by use of ground-based data over xinjiang, china. *Environmental earth sciences*, Springer, v. 77, n. 5, p. 1–18, 2018. Citado na página 25.

MAGGIONI, V.; MEYERS, P. C.; ROBINSON, M. D. A review of merged high-resolution satellite precipitation product accuracy during the tropical rainfall measuring mission (trmm) era. *Journal of Hydrometeorology*, v. 17, n. 4, p. 1101–1117, 2016. Citado na página 16.

MAGHSOOD, F. F. et al. Ground validation of gpm imerg precipitation products over iran. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 12, n. 1, p. 48, 2020. Citado 6 vezes, páginas 13, 15, 19, 23, 29 e 32.

MAHMOUD, M. T.; HAMOUDA, M. A.; MOHAMED, M. M. Spatiotemporal evaluation of the gpm satellite precipitation products over the united arab emirates. *Atmospheric Research*, Elsevier, v. 219, p. 200–212, 2019. Citado 2 vezes, páginas 19 e 32.

MOAZAMI, S.; NAJAFI, M. R. A comprehensive evaluation of GPM-IMERG V06 and MRMS with hourly ground-based precipitation observations across Canada. *Journal of Hydrology*, Elsevier, v. 594, p. 125929, mar 2021. ISSN 0022-1694. Citado na página 22.

MONTEIRO, J. A. et al. Accuracy of grid precipitation data for brazil: application in river discharge modelling of the tocantins catchment. *Hydrological Processes*, Wiley Online Library, v. 30, n. 9, p. 1419–1430, 2016. Citado 2 vezes, páginas 15 e 29.

MORIASI, D. N. et al. Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*, American Society of Agricultural and Biological Engineers, v. 58, n. 6, p. 1763–1785, 2015. ISSN 21510032. Citado 2 vezes, páginas 36 e 40.

NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models part i—a discussion of principles. *Journal of hydrology*, Elsevier, v. 10, n. 3, p. 282–290, 1970. Citado na página 36.

NÓBREGA, M. et al. Uncertainty in climate change impacts on water resources in the rio grande basin, brazil. *Hydrology and Earth System Sciences*, Copernicus GmbH, v. 15, n. 2, p. 585–595, 2011. Citado na página 35.

PAIVA, R. C.; COLLISCHONN, W.; TUCCI, C. E. Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *Journal of Hydrology*, v. 406, n. 3-4, p. 170–181, sep 2011. ISSN 00221694. Citado na página 33.

PAIVA, R. C. D. et al. Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research*, v. 49, n. 3, p. 1226–1243, 2013. ISSN 00431397. Citado na página 33.

PARDO-IGÚZQUIZA, E. Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing. *Journal of Hydrology*, Elsevier, v. 210, n. 1-4, p. 206–220, 1998. Citado na página 15.

PINTO, W. d. P.; LIMA, G. B.; ZANETTI, J. B. Comparative analysis of models for times to series modeling and forecasting of scheme of average monthly streamflow of the doce river, colatina espirito santo, brazil. *Ciência e Natura*, v. 37, n. 3, p. 1–11, 2015. Citado na página 30.

PONTES, P. R. M. et al. MGB-IPH model for hydrological and hydraulic simulation of large floodplain river systems coupled with open source GIS. *Environmental Modelling and Software*, Elsevier Ltd, v. 94, p. 1–20, 2017. ISSN 13648152. Citado 2 vezes, páginas 33 e 34.

PRAKASH, S. et al. A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology*, Elsevier, v. 556, p. 865–876, jan 2018. ISSN 0022-1694. Citado 2 vezes, páginas 16 e 23.

QUADRO, M. F. L. d. et al. Análise climatológica da precipitação e do transporte de umidade na região da zcas através da nova geração de reanálises. *Revista Brasileira de Meteorologia*, SciELO Brasil, v. 27, p. 152–162, 2012. Citado na página 16.

REBOITA, M. S. et al. Precipitation regimes in South America: a bibliography review. *Revista Brasileira de Meteorologia*, Sociedade Brasileira de Meteorologia, v. 25, n. 2, p. 185–204, jun 2010. ISSN 0102-7786. Citado 3 vezes, páginas 8, 16 e 17.

ROZANTE, J. R. et al. Evaluation of TRMM/GPM Blended Daily Products over Brazil. *Remote Sensing 2018, Vol. 10, Page 882*, Multidisciplinary Digital Publishing Institute, v. 10, n. 6, p. 882, jun 2018. Citado 4 vezes, páginas 10, 16, 19 e 38.

SALIO, P. et al. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmospheric Research*, Elsevier Ltd, v. 163, p. 146–161, sep 2015. ISSN 01698095. Citado 3 vezes, páginas 13, 15 e 29.

SEVRUK, B.; ONDRÁS, M.; CHVÍLA, B. The WMO precipitation measurement intercomparisons. *Atmospheric Research*, Elsevier, v. 92, n. 3, p. 376–380, may 2009. ISSN 0169-8095. Citado na página 28.

SILVA, B. et al. Previsão Hidroclimática de Vazão de Curto Prazo na Bacia do Rio São Francisco. *Revista Brasileira de Recursos Hídricos*, v. 12, p. 31–41, 2007. Citado na página 33.

SIQUEIRA, V. A. et al. Ensemble flood forecasting based on operational forecasts of the regional Eta EPS in the Taquari-Antas basin. *Rbrh*, v. 21, n. 3, p. 587–602, 2016. Citado 2 vezes, páginas 34 e 36.

SIQUEIRA, V. A. et al. Toward continental hydrologic-hydrodynamic modeling in South America. *Hydrology and Earth System Sciences*, v. 22, n. 9, p. 4815–4842, 2018. ISSN 16077938. Citado na página 33.

SORRIBAS, M. V. et al. Projections of climate change effects on discharge and inundation in the Amazon basin. *Climatic Change*, Springer Netherlands, v. 136, n. 3-4, p. 555–570, jun 2016. ISSN 0165-0009. Citado na página 33.

STEPHENSON, D.; PATRICK, N. Spatial variation of rainfall intensities for short duration storms. *Hydrological Sciences Journal HSJODN,*, v. 35, n. 6, 1990. Citado na página 19.

TAPIADOR, F. J. et al. *Global precipitation measurement: Methods, datasets and applications*. 2012. 70–97 p. Citado 3 vezes, páginas 13, 15 e 29.

THORNDAHL, S. et al. Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences*, Copernicus GmbH, v. 21, n. 3, p. 1359–1380, mar 2017. ISSN 16077938. Citado na página 29.

TIAN, F. et al. How Does the Evaluation of the GPM IMERG Rainfall Product Depend on Gauge Density and Rainfall Intensity? *Journal of Hydrometeorology*, American Meteorological Society, v. 19, n. 2, p. 339–349, feb 2018. ISSN 1525-7541. Citado 4 vezes, páginas 19, 27, 28 e 31.

TOMKINS, K. M. Uncertainty in streamflow rating curves: methods, controls and consequences. *Hydrological Processes*, John Wiley & Sons, Ltd, v. 28, n. 3, p. 464–481, jan 2014. ISSN 1099-1085. Citado na página 43.

USGS. Shuttle radar topography mission (srtm) 1 arc-second global. *US Geological Survey*, 2015. Citado 2 vezes, páginas 18 e 34.

VELASCO, I.; FRITSCH, J. M. Mesoscale convective complexes in the americas. *Journal of Geophysical Research: Atmospheres*, Wiley Online Library, v. 92, n. D8, p. 9591–9613, 1987. Citado na página 16.

VERA, C. et al. The south american low-level jet experiment. *Bulletin of the American Meteorological Society*, American Meteorological Society, v. 87, n. 1, p. 63–78, 2006. Citado na página 16.

WANG, Z. et al. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmospheric Research*, Elsevier Ltd, v. 196, p. 151–163, nov 2017. ISSN 01698095. Citado na página 29.

WILKS, D. S. *Statistical methods in the atmospheric sciences*. [S.l.]: Elsevier Academic Press, 2011. v. 100. Citado 2 vezes, páginas 19 e 32.

XU, R. et al. Ground validation of gpm imerg and trmm 3b42v7 rainfall products over southern tibetan plateau based on a high-density rain gauge network. *Journal of Geophysical Research: Atmospheres*, Wiley Online Library, v. 122, n. 2, p. 910–924, 2017. Citado 4 vezes, páginas 10, 19, 32 e 38.

YAPO, P. O.; GUPTA, H. V.; SOROOSHIAN, S. Multi-objective global optimization for hydrologic models. *Journal of hydrology*, Elsevier, v. 204, n. 1-4, p. 83–97, 1998. Citado na página 35.

ZHANG, Z. et al. Hydrologic Evaluation of TRMM and GPM IMERG Satellite-based Precipitation in a Humid Basin of China. *Remote Sensing*, MDPI AG, v. 11, n. 4, p. 431, feb 2019. ISSN 2072-4292. Citado na página 29.

ZHOU, J.; LAU, K. Does a monsoon climate exist over south america? *Journal of climate*, v. 11, n. 5, p. 1020–1040, 1998. Citado na página 16.

5 APÊNDICE A: SUPPLEMENTARY MATERIALS - GROUND-BASED EVALUATION OF GPM-IMERG V6 OVER BRAZIL

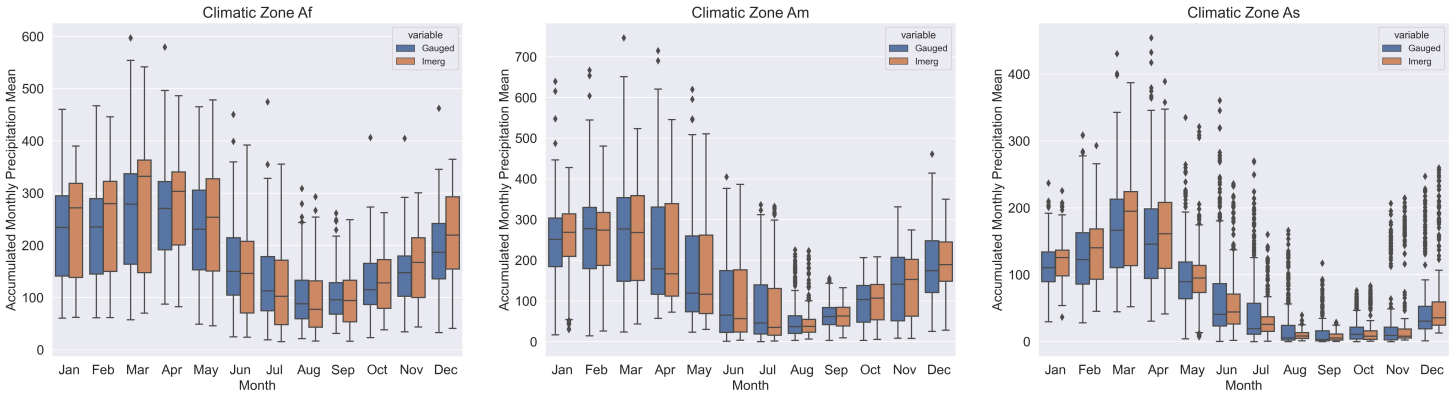

Figure 5.1 – Monthly averages of the Gauged and IMERG data based on monthly accumulated precipitation for the climatic zones Af, Am and As.

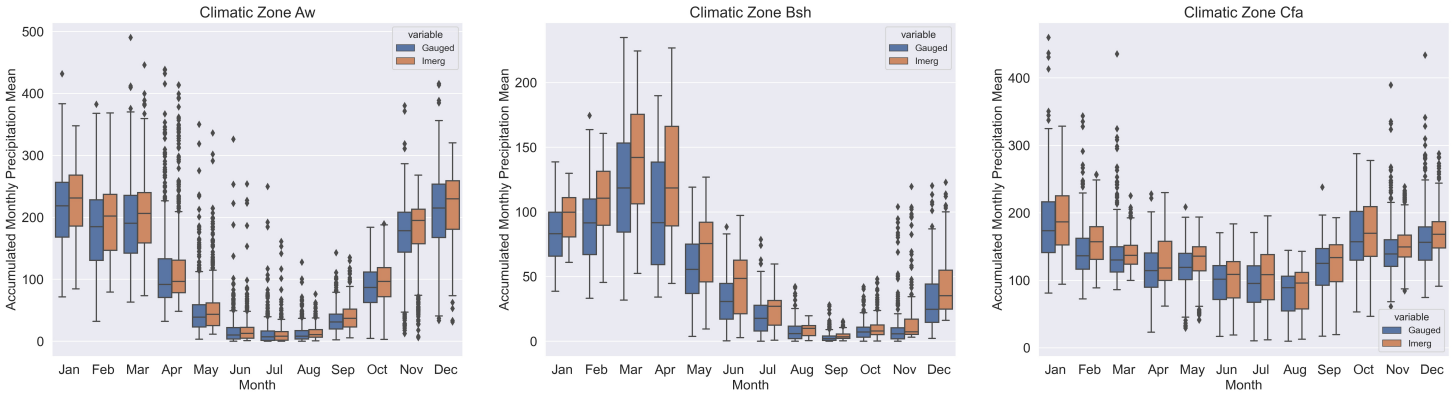

Figure 5.2 – Monthly averages of the Gauged and IMERG data based on monthly accumulated precipitation for the climatic zones Aw, Bsh and Cfa.

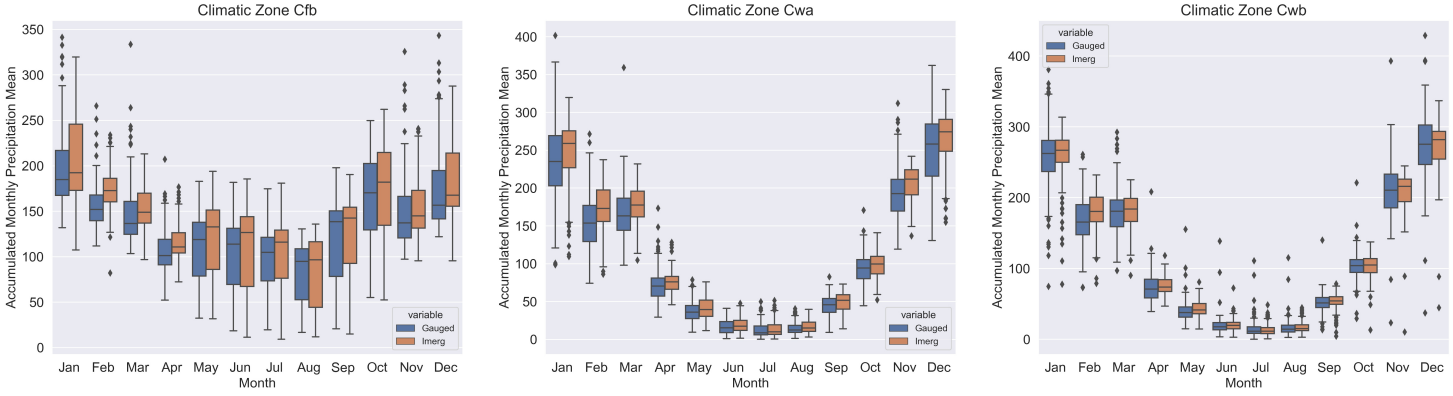

Figure 5.3 – Monthly averages of the Gauged and IMERG data based on monthly accumulated precipitation for the climatic zones Cfb, Cwa and Cwb.

Month	POD			FAR			CSI		
	Zone A	Zone B	Zone C	Zone A	Zone B	Zone C	Zone A	Zone B	Zone C
Jan	0.86	0.83	0.82	0.49	0.69	0.44	0.47	0.29	0.50
Feb	0.87	0.85	0.80	0.48	0.68	0.48	0.49	0.30	0.46
Mar	0.87	0.86	0.76	0.47	0.64	0.49	0.50	0.34	0.44
Apr	0.81	0.81	0.67	0.51	0.66	0.55	0.44	0.31	0.36
May	0.71	0.67	0.63	0.61	0.72	0.60	0.33	0.23	0.31
Jun	0.60	0.53	0.60	0.68	0.76	0.61	0.25	0.18	0.31
Jul	0.53	0.42	0.59	0.72	0.80	0.62	0.20	0.13	0.30
Aug	0.57	0.35	0.62	0.74	0.86	0.61	0.19	0.08	0.31
Sep	0.59	0.28	0.70	0.71	0.90	0.53	0.22	0.06	0.38
Oct	0.69	0.56	0.75	0.68	0.87	0.52	0.27	0.11	0.41
Nov	0.74	0.58	0.75	0.62	0.88	0.48	0.34	0.11	0.44
Dec	0.83	0.78	0.79	0.57	0.81	0.46	0.40	0.17	0.48

Table 5.1 – Summary of the results for the first group of statistical metrics for each month of the year grouped by the climate zones.

Month	$RMSE$			MAE			$rBIAS_\epsilon$		
	Zone A	Zone B	Zone C	Zone A	Zone B	Zone C	Zone A	Zone B	Zone C
Jan	14.32	9.55	15.32	7.41	3.45	7.94	0.30	0.32	0.22
Feb	14.50	10.13	13.88	7.77	4.31	6.97	0.31	0.41	0.24
Mar	14.62	10.90	12.97	7.85	4.97	6.14	0.30	0.40	0.18
Apr	12.89	10.41	11.08	6.45	4.66	4.44	0.24	0.39	0.12
May	9.77	7.41	9.80	4.28	2.67	3.65	0.17	0.24	0.10
Jun	6.58	5.57	7.85	2.75	1.84	2.80	0.07	0.12	0.07
Jul	5.25	3.84	7.11	2.07	1.09	2.55	0.03	0.05	0.06
Aug	4.36	2.41	6.41	1.52	0.47	2.23	0.05	0.01	0.07
Sep	5.72	1.36	9.42	2.03	0.23	3.74	0.12	0.02	0.09
Oct	8.36	2.71	12.17	3.42	0.44	5.56	0.22	0.07	0.16
Nov	11.23	2.98	14.27	5.39	0.68	7.13	0.26	0.12	0.16
Dec	12.69	5.31	15.24	6.40	1.50	7.87	0.31	0.25	0.20

Table 5.2 – Summary of the results for the second group of statistical metrics for each month of the year grouped by the climate zones.

Month	$CORR$		
	Zone A	Zone B	Zone C
Jan	0.40	0.48	0.37
Feb	0.40	0.44	0.34
Mar	0.38	0.43	0.37
Apr	0.39	0.42	0.38
May	0.39	0.42	0.43
Jun	0.34	0.36	0.45
Jul	0.31	0.33	0.46
Aug	0.31	0.19	0.47
Sep	0.32	0.20	0.48
Oct	0.37	0.43	0.43
Nov	0.34	0.37	0.35
Dec	0.38	0.45	0.36

Table 5.3 – Summary of the results for the third group of statistical metrics for each month of the year grouped by the climate zones.

6 APÊNDICE B: SUPPLEMENTARY MATERIALS - EVALUATION OF GPM-IMERG APPLICABILITY FOR HYDROLOGICAL MODELING IN A LARGE-SCALE TROPICAL BASIN

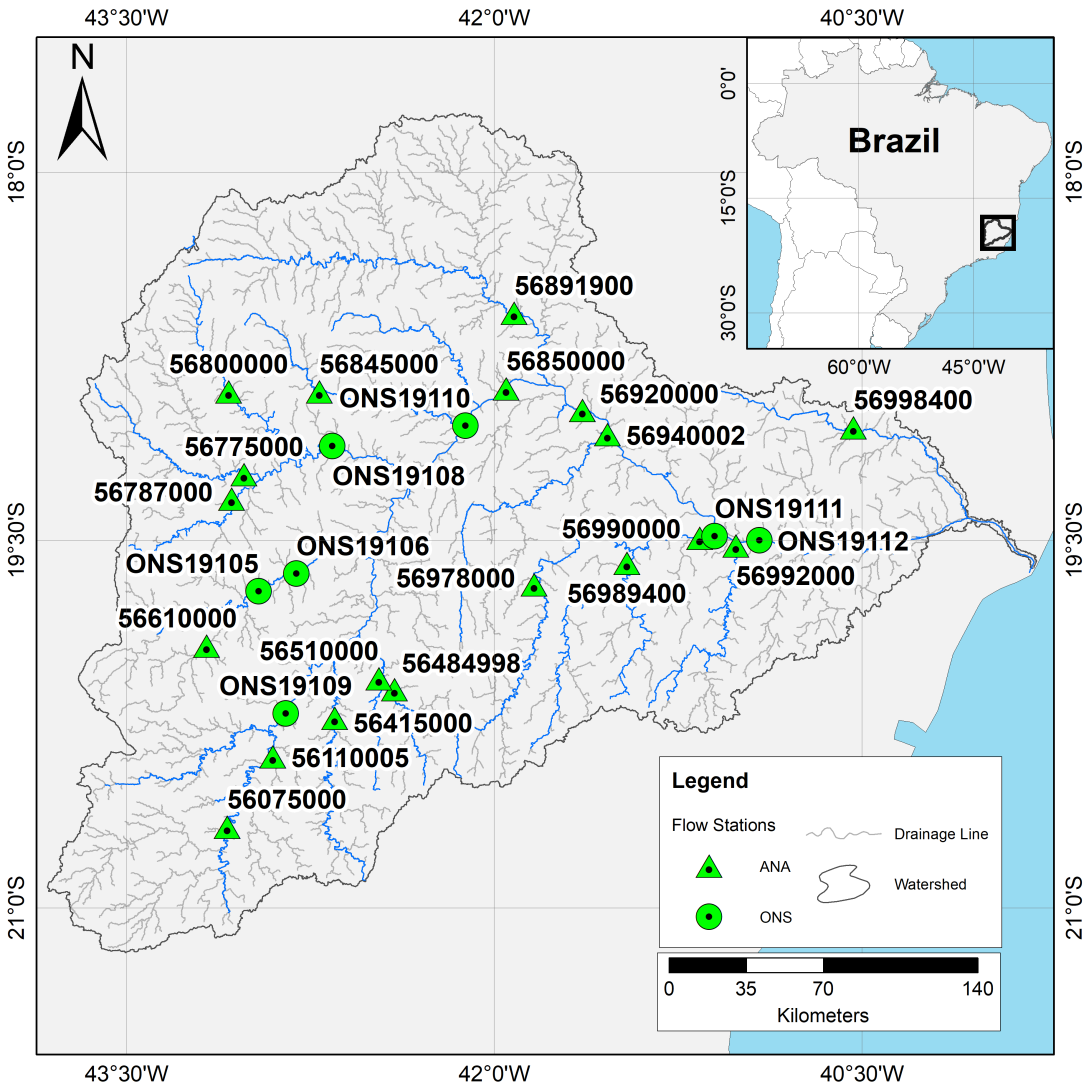

Figure 6.1 – Spatial distribution and code identification for selected flow stations.

Station	Calibration							
	Gauged				IMERG			
	NSE	NSE_{log}	KGE	$PBIAS$	NSE	NSE_{log}	KGE	$PBIAS$
56075000	0.841	0.88	0.907	-0.63	0.756	0.799	0.808	2.86
56110005	0.869	0.872	0.923	-0.55	0.768	0.778	0.824	1.61
56415000	0.761	0.806	0.871	4.42	0.609	0.656	0.773	-13.02
56484998	0.836	0.703	0.82	14.97	0.671	0.701	0.747	3.64
56510000	0.852	0.746	0.881	9.38	0.677	0.7	0.813	-2.18
56610000	0.819	0.849	0.864	5.04	0.699	0.811	0.695	8.12
56775000	0.727	0.761	0.675	20.86	0.643	0.778	0.622	20.65
56787000	0.722	0.721	0.73	16.9	0.723	0.775	0.794	3.38
56800000	0.764	0.811	0.773	7.76	0.648	0.767	0.698	2.31
56845000	0.799	0.848	0.801	6.49	0.698	0.819	0.754	8.1
56850000	0.897	0.902	0.895	0.4	0.834	0.853	0.871	-5.49
56891900	0.851	0.843	0.886	-3.95	0.749	0.808	0.734	5.85
56920000	0.858	0.761	0.874	10.11	0.84	0.773	0.895	6.46
56940002	0.773	0.826	0.795	2.16	0.62	0.736	0.685	0.43
56978000	0.684	0.514	0.623	31.61	0.659	0.458	0.613	26.88
56989400	0.756	0.797	0.734	14.75	0.685	0.702	0.772	2.17
56990000	0.795	0.849	0.846	10.81	0.774	0.721	0.879	1.54
56992000	0.629	0.578	0.605	21.78	0.627	0.672	0.77	3.24
56998400	0.633	0.686	0.589	29.53	0.644	0.732	0.674	17.81
ONS19105	0.884	0.85	0.908	2.67	0.78	0.817	0.829	3.47
ONS19106	0.884	0.862	0.911	0.09	0.779	0.823	0.856	-1.01
ONS19108	0.844	0.824	0.763	16.76	0.765	0.846	0.787	9.72
ONS19109	0.913	0.907	0.939	4.37	0.776	0.806	0.873	3.76
ONS19110	0.85	0.897	0.897	-0.89	0.815	0.836	0.865	-7.48
ONS19111	0.799	0.857	0.846	-11.74	0.761	0.788	0.789	-17.46
ONS19112	0.817	0.889	0.869	1.26	0.812	0.849	0.879	-4.26

Table 6.1 – Statistical metrics of MGB-IPH streamflow simulation using both gauged and IMERG as precipitation input data for the calibration period.

Station	Calibration							
	Gauged				IMERG			
	NSE	NSE_{log}	KGE	$PBIAS$	NSE	NSE_{log}	KGE	$PBIAS$
56075000	0.823	0.841	0.872	-8.68	0.763	0.757	0.8	-7.87
56110005	0.873	0.846	0.92	-4.45	0.766	0.741	0.783	-4.36
56415000	0.712	0.716	0.841	-5.76	0.629	0.504	0.704	-17.09
56484998	0.837	0.8	0.907	3.29	0.672	0.637	0.743	-0.09
56510000	0.741	0.775	0.83	-10.2	0.66	0.582	0.736	-6.77
56610000	0.816	0.792	0.853	-0.31	0.68	0.751	0.717	-2.01
56775000	0.749	0.759	0.674	19.18	0.645	0.751	0.633	14.32
56787000	0.721	0.741	0.802	7.78	0.626	0.683	0.768	-11.91
56800000	0.73	0.708	0.824	-12.15	0.522	0.534	0.682	-23.1
56845000	0.761	0.703	0.809	-9.73	0.627	0.54	0.726	-21.52
56850000	0.894	0.887	0.904	-5.14	0.831	0.788	0.825	-14.73
56891900	0.632	0.304	0.397	-57.07	0.628	0.321	0.555	-41.86
56920000	0.854	0.858	0.889	-7.61	0.808	0.756	0.831	-14.18
56940002	0.697	0.803	0.781	2.41	0.674	0.725	0.733	1.67
56978000	0.764	0.698	0.669	26.66	0.653	0.637	0.671	22.36
56989400	0.74	0.774	0.779	10.15	0.611	0.658	0.793	-4.32
56990000	0.784	0.829	0.778	-0.98	0.523	0.711	0.666	-9.2
56992000	0.724	0.742	0.811	-1.73	0.592	0.568	0.719	-17.5
56998400	0.72	0.72	0.751	15.07	0.743	0.668	0.856	2.35
ONS19105	0.903	0.834	0.904	-5.39	0.803	0.802	0.845	-6.75
ONS19106	0.889	0.824	0.876	-7.48	0.801	0.785	0.846	-10.3
ONS19108	0.853	0.821	0.795	12.93	0.768	0.772	0.824	1.66
ONS19109	0.918	0.901	0.957	0.45	0.794	0.809	0.85	-3.84
ONS19110	0.877	0.873	0.882	-9.02	0.819	0.753	0.787	-19.44
ONS19111	0.834	0.808	0.818	-16.15	0.755	0.691	0.741	-23.17
ONS19112	0.845	0.857	0.885	-4.16	0.793	0.766	0.844	-10.75

Table 6.2 – Statistical metrics of MGB-IPH streamflow simulation using both gauged and IMERG as precipitation input data for the validation period.

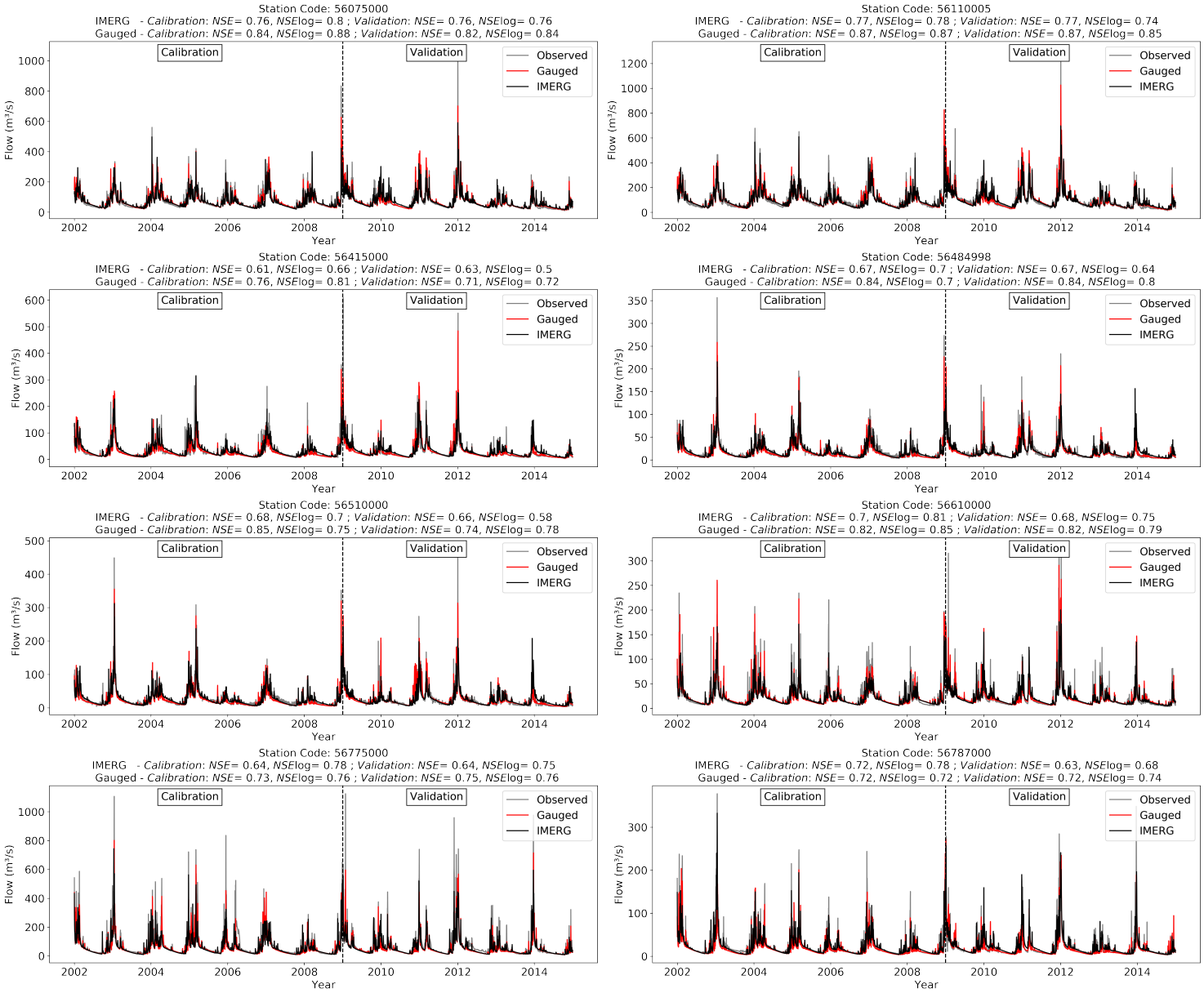

Figure 6.2 – Comparison between the simulated using gauged (red) and IMERG (black) as precipitation input data, and the observed (gray) discharge - Part 1.

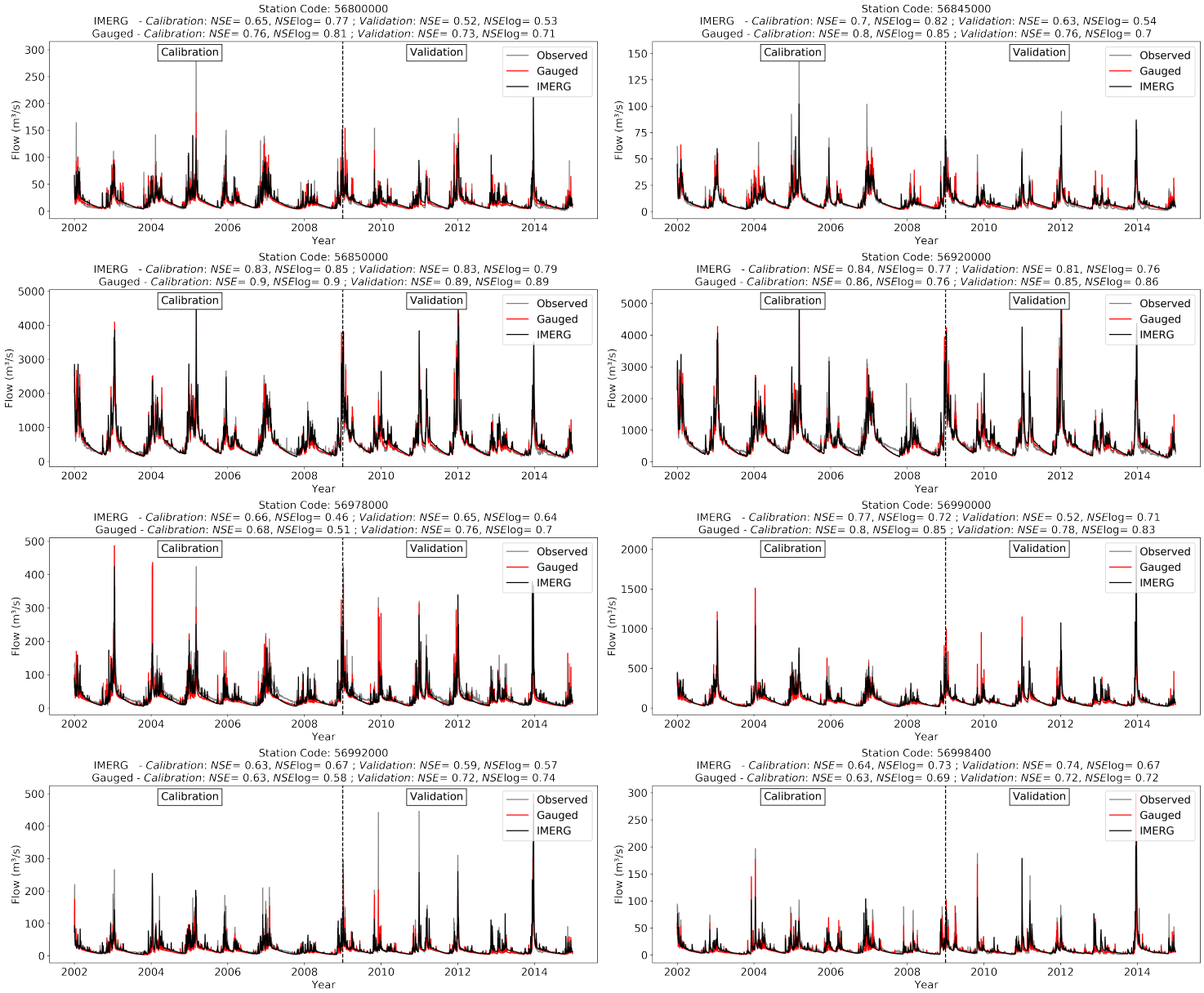

Figure 6.3 – Comparison between the simulated using gauged (red) and IMERG (black) as precipitation input data, and the observed (gray) discharge - Part 2.

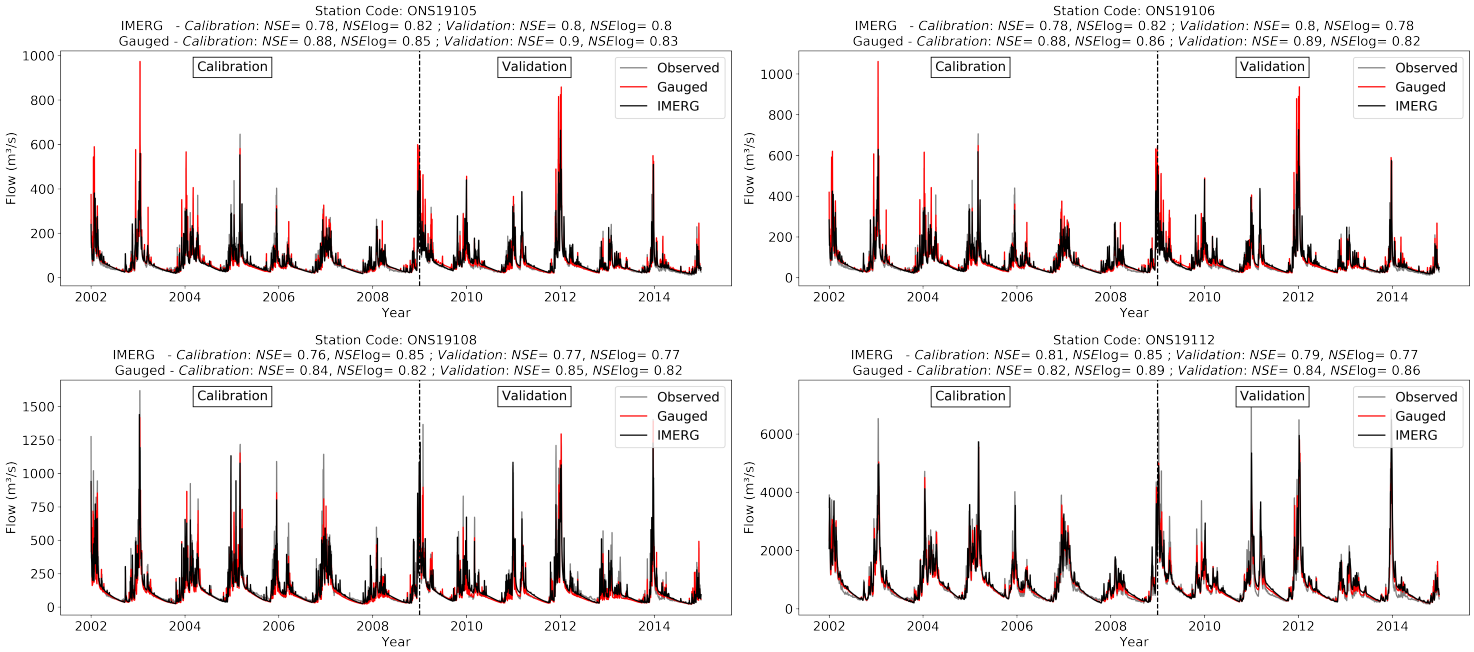

Figure 6.4 – Comparison between the simulated using gauged (red) and IMERG (black) as precipitation input data, and the observed (gray) discharge - Part 3.