Evaluation of Diverse-based Precipitation Data Over the Amazon Region

Camila Sapucci (✉ camila.sapucci@usp.br)

University of Sao Paulo: Universidade de Sao Paulo <https://orcid.org/0000-0003-0127-5463>

Victor C. Mayta

University of Sao Paulo: Universidade de Sao Paulo

Pedro Silva Dias

University of Sao Paulo: Universidade de Sao Paulo

Research Article

Keywords: HYBAM rain-gauge observations, Climate Hazards Group InfraRed Precipitation, Multi-Source Weighted-Ensemble Precipitation, MSWEP

Posted Date: August 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-781813/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The skill of the diverse-based precipitation products is investigated in comparison with HYBAM rain-gauge observations. The performance of three remote sensing-based datasets (the Climate Hazards Group InfraRed Precipitation with Station, CHIRPS, the Multi-Source Weighted-Ensemble Precipitation, MSWEP, and the Tropical Rainfall Measuring Mission, TRMM) is evaluated considering different timescales for the Amazon Basin, an area with widely heterogeneous precipitation. The analysis considered seasonal, intraseasonal and diurnal timescales through the computation of the cluster analysis, the seasonality index, the Kling-Gupta Efficiency metric, spectral analysis and composing technique. CHIRPS has the lowest performance to represent the rainfall in the northwest portion of the basin, where it underestimated the mean precipitation compared to the other bases. In this region, the other remote sensing-based (TRMM and MSWEP databases) compared to HYBAM also showed considerable variability and misrepresentation of the intraseasonal rainfall. In general, all databases perform better in the north and eastern portions of the basin compared to HYBAM. The comparison of the diurnal rainfall cycle between remote sensing-based data and the field campaigns of TRMM-LBA and GoAmazon, and the Huayao station in the Andes was also evaluated. At the diurnal timescale, MSWEP predates the time of the rainfall peak, but represents the magnitude of the precipitation well compared with TRMM. This study is necessary to warn about the importance of a more complete and objective assessment of the data before considering it for applications in different precipitation studies, mainly in regions with high rainfall heterogeneity like the Amazon Basin.

1 Introduction

The Amazon Basin (AB) is considered the world's largest rainforest and hydrological area. The Amazon covers a broad area of about 6 million km² and includes seven countries in South America: Brazil, Peru, Bolivia, Colombia, Ecuador, Venezuela and Guyana (Fisch et al. 1998; Espinoza et al. 2009). The Amazon plays a fundamental role in the regulation of the regional and global climate, constituting an important heat source for the global atmospheric circulation, providing moisture for other regions in South America (Figuerola and Nobre 1990; Satyamurti et al. 1998; Marengo 2006; Nobre et al. 2009).

Precipitation over a broad area of the AB is strongly modulated by the South American Monsoon System (SAMS) (Zhou and Lau 1998; Vera et al. 2006; Marengo et al. 2012), characterized by a rainy period from October to April (Figs. 1a-d) and a dry period from May to September (Figs. 1e-h). The SAMS maximum convective activity occurs in the austral summer, peaking in the south-central region of the AB. The March-May period is considered the end of the rainfall activity, and during this period the precipitation migrates northward. June-August is considered the dry season in southern AB, when the maximum precipitation is located in the extreme north of the basin (Whang and Fu 2002). Annual precipitation over the basin varies from 1500 mm yr⁻¹ in the eastern portion, 2800 mm yr⁻¹ in the central and southern portions, and 3500 mm yr⁻¹ in the northwest and western regions. There are also regions with annual precipitation exceeding 6000 mm yr⁻¹, where the "hotspots" in the Andes-Amazon transition region are located (Figuerola and Nobre 1990; Satyamurty et al. 1998; Espinoza et al. 2015; Chavez and Takahashi

2017). The interactions between the large-scale moisture transport and the orography make this region one of the rainiest in the world. Even during the May-September “dry” period, the hotspot region receives about 1000 mm (Figs. 1e-h).

The interactions of large-scale physical and dynamical processes with local features make the precipitation in the AB a complex feature (Mayta et al. 2019), with a highly heterogeneous spatial pattern (Figs. 2 and 3). Previous studies, applying several statistical techniques, found different spatial precipitation patterns over the AB. Using gauge station data, for instance, Santos et al. (2015) found that three sub-regions are sufficient to separate the Brazilian Amazon into different patterns of precipitation, but six sub-regions are ideal for more detailed studies. Mayta et al. (2020) also found that six sub-regions reasonably describe the variety of the AB precipitation. These different precipitation regimes over the AB pointed out by previous studies are the result of interactions between short timescales, such as the diurnal and synoptic, and longer timescales, such as the intraseasonal and interannual. At the interannual timescale, the El Niño Southern Oscillation (ENSO) has been largely related to the occurrence of drought and flood extreme events in the AB (Marengo 2004; Shimizu et al. 2017). The Madden-Julian Oscillation (MJO, Madden and Julian 1971, 1994; Zhang 2005) is the dominant mode of the intraseasonal variability in the tropics, playing a key role, for instance, in the South Atlantic Convergence Zone (SACZ) position and intensity (Carvalho et al. 2004). During the active phase of the MJO over South America, a seesaw pattern is observed between the region of SACZ and the southeast of South America (SESA), which contributes to increased rainfall in the central-eastern portion of the continent, including the southern AB and suppression over SESA (Casarin and Kousky 1986; Nogués-Paegle and Mo 1997; Liebmann et al. 1999; Vera et al. 2018; Grimm 2019). In this phase, a second dipole pattern is also observed in the AB, with suppressed convection in the northwestern portion of the basin and increased convection over the southeast (Souza and Ambrizzi 2006; Mayta et al. 2019).

Precipitation in the AB is a heterogeneous and complex process due to the interaction between different atmospheric systems at different timescales. In addition, the density of the forest prevents the existence of a wide network of surface rain gauges, most of them situated along river courses (Fitzjarrald et al. 2008; Paiva et al. 2011), where the river breezes introduce small scale variability. Characterizing precipitation over the AB is still a challenging task and encourages the use of precipitation estimates based on remote sensing.

Several previous studies have evaluated and validated different precipitation databases for South America (e.g., Pinto et al. 2009), and, regionally, for the AB (e.g., Zubieta et al. 2019). Costa and Foley (1998) provided the first study of this type when comparing six precipitation datasets for the AB based on gauge-stations, satellite, and reanalysis, including the widely used Global Precipitation Climatology Project (GPCP; Janowiak and Arkin 1990) and the National Centers for Environmental Prediction (NCEP) reanalysis (Kalnay et al. 1996). They found that for the AB the databases agree with respect to the precipitation climatology, but there were discrepancies when the interannual variability was investigated. The authors therefore considered that some caution should be taken when applying different precipitation databases at different timescales for the AB. Juárez et al. (2009) compared rain gauge data

from the Climate Prediction Center (CPC, Silva et al. 2007) and the Global Precipitation Climatology Centre (GPCC; Rudolf and Schneider 2005; Schneider et al. 2008) and combined data (satellite and *in situ* observations) from the Tropical Rainfall Measuring Mission (TRMM, Huffman et al. 2007), the GPCP (Adler et al., 2003), and the CPC Merged Analysis of Precipitation (CMAP, Xie and Arkin, 1997). This study also documented that, despite the similarity between the seasonal pattern of the AB rainfall provided by different databases, uncertainties exist related to the rainfall spatio-temporal pattern at interannual and decadal timescales. For the AB, the TRMM data performance against the surface observations was also explored by Negri et al. (2002), Collischonn et al. (2008), Clarke et al. (2011), Buarque et al. (2011) and Condom et al. (2011). Similar analyses are also developed for the CPC Morphing technique database (CMORPH, Joyce et al. 2004) by Fitzjarrald et al. (2008) and Pereira Filho et al. (2010). Buarque et al. (2011) compared TRMM and CMORPH to *in situ* observations to explore some characteristics of the AB rainfall, like the annual mean precipitation and the number of wet days and wet extremes. The authors discussed the fact that CMORPH overestimated the observed precipitation more than TRMM.

Getirana et al. (2011) explored the potential and limitations of six precipitation products for applications in hydrological studies, such as water balance in the Negro River. They compared gauge observations from Hybam Observatory Precipitation (HYBAM, Espinoza et al. 2009) and CPC, satellite data from GPCP and TRMM, and reanalysis data from NCEP-2 (Kanamitsu et al. 2002) and ERA-40 (Uppala et al. 2005). TRMM and NCEP provided the least accurate precipitation estimates. The representation of the water cycle by these databases showed a close agreement, but the water discharge and the precipitation fields were overestimated. Gauge-based data still provided the best performance in representing the rainfall and the water cycle in the AB northern portion. Correa et al. (2017) also performed a similar analysis for hydrological purposes, comparing eight precipitation products, and pointed out that databases which merge reanalysis, satellite, and gauge observations, such as the Multi-Source Weighted-Ensemble Precipitation (MSWEP, Beck et al. 2017, 2019) and the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS, Funk et al. 2015), provide an accurate representation of the water discharge. Besides, MSWEP and CHIRPS showed best results for the interannual variability of drought and flood events compared to purely reanalysis-based datasets. CHIRPS was compared to HYBAM *in situ* observations by Espinoza et al. (2019) in order to evaluate the evolution of the rainfall intensity in the AB. CHIRPS was also considered by Paca et al. (2020) in their study on rainfall trends over the Amazon River Basin. Over other areas in South America, CHIRPS was also used to identify wet and dry conditions over Central-Western Argentina (Rivera et al. 2019). For the La Plata Basin (Cerón et al. 2020), CHIRPS showed a poor performance in the Andean portion of the basin. Fagundes et al. (2017) also found results in agreement for the MSWEP V1 (first version) estimation of the rainfall compared to 117 gauge stations over AB, with an overall overestimation tendency. Some studies have applied MSWEP V1 to statistical downscaling models (eg. Bettolli et al. 2021; Solman et al. 2021). Comparing the precipitation data from Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN, Ashouri et al. 2015), global data from CPC (Xie et al. 2010), CMORPH, MSWEP V1 and TRMM, Bettolli et al. (2021) observed that MSWEP provided one of the best precipitation estimates compared to observed data for

the South America CORDEX domain. Recently, Rosa et al. (2020) applied MSWEP V2 (Beck et al. 2019) to validate an algorithm for identification of continental and oceanic SACZ events.

Almost all of the studies mentioned above assess different precipitation databases for AB, including gauge-based, satellite, and reanalysis datasets. However, there is no study yet that evaluates the performance of these precipitation products at different timescales. Differences between datasets at different timescales are expected, as elucidated by Costa and Foley (1998) and Juárez et al. (2009). A strict quality control of the precipitation datasets is necessary before carrying out any study in order to ensure that these data represent as accurately as possible the full complexity of the AB regional precipitation. In this context, this study aims to (a) assess three rainfall databases (CHIRPS, TRMM and MSWEP V2) with respect to observed precipitation (HYBAM), at diurnal, seasonal and intraseasonal timescales for the AB domain, and (b) to investigate the usefulness of MSWEP V2 in the representation of the AB rainfall. The study is justified by the need to obtain detailed information on how each database behaves at different timescales. The paper is organized as follows: Section 2 describes the precipitation databases and the methodology applied to compare these datasets. Section 3 discusses the main results, and Section 4 gives the summary and the conclusions.

2 Data And Methodology

2.1 Precipitation Databases

The performance of CHIRPS, TRMM and MSWEP rainfall databases compared with HYBAM is evaluated for the AB region. The main aspects of these databases are summarized in Table 1.

Table 1
Summary of the main characteristics of the precipitation databases used for the AB region

Database	Temporal Resolution	Spatial Resolution	Period	References
HYBAM	Daily	1°x1°	1980–2009	Espinoza et al. (2009) Guimberteau et al. (2012)
CHIRPS	Daily	0.05°x0.05°	1981-present	Funk et al. (2015)
TRMM 3B42	Daily and 3-hourly	0.25°x0.25°	1998–2019	Huffman et al. (2007)
MSWEP V2	Daily and 3-hourly	0.1°x0.1°	1979-present	Beck et al. (2019)
TRMM-LBA	Hourly	□	Dec. 1998 - Feb. 1999	Silva Dias et al. (2002) Avisar et al. (2002) Marengo et al. (2004)
Go-Amazon	3-hourly	□	2014–2015	Machado et al. (2014) Schumacher et al. (2018)
Huayao Station	Hourly	□	2012–2020	Saavedra et al. (2020)

The HYBAM database (see <https://hybam.obs-mip.fr/>) was a collaborative action between the AB countries. It includes data from several agencies in Brazil, Peru, Bolivia, Colombia and Ecuador for hydrological and meteorological monitoring of the basin. The precipitation data of 756 stations were spatially interpolated to a one-degree grid and formed the HYBAM Observed Precipitation dataset, covering the period from 1980 to 2009, for the entire AB domain (Espinoza et al. 2009; Guimberteau et al. 2012).

In this study, the TRMM-3B42 V7 product is used, covering the area of 50°N-50°S, high spatial resolution (0.25°x0.25°) and with daily and 3-hourly temporal resolutions (Huffman et al. 2007; Huffman et al. 2010). The widely-used CHIRPS data is also used in this study. CHIRPS uses TRMM-3B42 V7 to calibrate Cold Cloud Duration (CCD) obtained from thermal infrared images (Funk et al. 2015). CHIRPS has a quasi-global coverage (50°N-50°S), with high spatial resolution of 0.05°x0.05° and daily, pentadal, and monthly temporal resolutions.

MSWEP V2 was developed by Beck et al. (2019) and provides a precipitation dataset with global coverage (including oceans) with 3-hourly and daily temporal resolutions and 0.1°x0.1° spatial resolution. MSWEP V2 combines data from more than 76000 gauge stations around the entire globe. This dataset passed through an extensive process of quality control and includes data from reanalysis (ERA-Interim, Dee et al. 2011; JRA-55, Kobayashi et al. 2015), satellite data from CMORPH, Gridded Satellite (GridSat, Knapp et al. 2011), Global Satellite Mapping of Precipitation (GSMaP, Ushio et al. 2009), and TRMM

3B42. The recently released MSWEP V2 dataset improved some aspects from MSWEP V1 (Beck et al. 2017). For instance, the introduction of cumulative density functions and rainfall corrections, the addition of rainfall estimates derived from GridSat thermal infrared imagery for the pre-TRMM era etc. In addition, the current version MSWEP V2 resolution has been improved from 0.25° to 0.1° .

It is important to emphasize that HYBAM is the only dataset considered totally based on gauge station observations, while the others (TRMM, CHIRPS and MSWEP) are estimates based on remote sensing. In this study, we interpolated CHIRPS and MSWEP data to a 0.25° grid resolution in order to describe rainfall variability in the entire AB. All the analyses described in the following sections are performed for the periods in common between the datasets (1998–2009 period).

The diurnal cycle analyses also used 3-hourly data from MSWEP and TRMM, compared to data from the TRMM-LBA and GoAmazon campaigns and data from the Huayao station. The Large-Scale Biosphere-Atmosphere Experiment in the Amazonian-Wet Season Atmospheric Mesoscale Campaign (LBA WET-AMC) together with the TRMM satellite validation campaign (TRMM-LBA) in southwestern Amazonia was conducted between December 1998 to February 1999. The LBA campaign aimed to explain some questions related to climate, carbon storage, atmospheric chemistry, interactions between the biosphere and atmosphere, hydrology, and impacts of land use/cover etc (Avisar et al. 2002). One of the main objectives of the TRMM-LBA campaign is the validation of TRMM satellite algorithms in the context of continental tropical convection (Silva Dias et al. 2002). In the present study the TRMM-LBA data from networks 2 (N2 at 10.75°S , 62.15°W), 3 (N3 at 10.65°S , 62.35°W) and 4 (N4 at 10.35°S , 62.55°W) in the Brazilian state of Rondônia are used.

The GoAmazon campaign was a field experiment that aimed to understand the interactions between aerosols and the cloud life cycle in the Amazon (Machado et al. 2014). This campaign extended from January 2014 to December 2015, focusing on the metropolitan region of Manaus and it included rainfall measurements by the S-band radar of the Amazon Protection System (SIPAM).

The Huayao station (12.04°S , 75.32°W) was also considered in the diurnal cycle analyses. This station is located in the Andes region at 3322m of altitude (Saavedra et al. 2020). This station was considered to verify how the databases represent the diurnal cycle in regions with strong orographic influence. The period analyzed for the Huayao station is from January 2014 to December 2015.

2.2 Cluster Analysis

The main goal of the cluster analysis is not to describe the precipitation annual cycle of each sub-regime of the AB, but to emphasize how each database identifies these precipitation regimes and to compare their spatial distribution. In this study, we grouped grid points that have similar precipitation characteristics by applying a non-hierarchical k-means cluster analysis. Two tests are computed: one with normalized and the other with non-normalized rainfall to characterize the annual precipitation cycle. The analysis followed the methodology described in Mayta et al. (2020).

2.3 Seasonality Index (SI)

The SI (Eq. 1) was proposed by Walsh and Lawler (1981) to provide an alternative way to analyze and subdivide areas in terms of their degree of seasonal variation in rainfall. This index varies from zero (when all months have the same amount of rainfall through the year) to 1.83 (when all year's rainfall occurs within a single month). The classification suggested by Walsh and Lawler (1981) is presented in Table S1 in the supplementary material. We calculated the SI for each year for the 1998 to 2009 period, considering all databases, and then we averaged the results to obtain a mean SI.

$$SI_i = \frac{1}{R_i} \sum_{n=1}^{12} |x_n - R_i/12|$$

1

where R_i represents the total annual precipitation for a specific year and x_n the total monthly rainfall during the month n (Walsh and Lawler 1981; Teegavarapu 2018). SI_i represents the SI for each year.

2.4 The Kling-Gupta Efficiency (KGE) Metric

KGE (Eq. 2) was proposed by Gupta et al. (2009) and modified by Kling et al. (2012). This metric combines Pearson's correlation (r), bias ratio (β , Eq. 3) and variability ratio (γ , Eq. 4) coefficients. KGE, r , β and γ have their optimum at unity, and the metric compares observed data (reference database) with estimated data (Kling et al. 2012). We chose HYBAM as the reference dataset. KGE is computed for the May-September and October-April periods (SAMS dry and wet periods, respectively) and also for each season of the year: December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON).

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

2

$$\beta = \mu_s / \mu_o$$

3

$$\gamma = (\sigma_s / \mu_s) / (\sigma_o / \mu_o)$$

4

where μ and σ represent mean and standard deviation, respectively, and subscripts s and o denote estimated (CHIRPS, TRMM or MSWEP) and observed (HYBAM) data, respectively. To ensure that the same grid points are compared in KGE metric, HYBAM data was first interpolated to a spatial resolution of $0.25^\circ \times 0.25^\circ$.

2.5 Spectral Analysis

Power spectra of filtered precipitation data are obtained for all databases for each AB sub-region identified in the cluster analysis. The spectra are calculated for each year of the 1998 to 2009 period and then averaged to obtain the mean power spectre, as in Liebmann et al. (1999) and Mayta et al. (2020). The focus of this analysis is to assess whether different rainfall datasets identify the same spectral power peaks.

2.6 Diurnal Variability

Diurnal variability of MSWEP and TRMM is assessed by using different datasets collected in different campaigns (Table 1). More specifically, the diurnal variability is analyzed in three stations around the AB: (i) Manaus city (3.15°S, 60°W), where rainfall measurements were obtained from the Doppler-S band radar during the GoAmazon campaign; (2) networks 2, 3 and 4 (see Marengo et al. (2004) for the networks' geographical position) in the Brazilian state of Rondônia during the TRMM-LBA campaign; and (iii) Huayao station (12.04°S, 75.321°W), located in the Andes region at 3322m of altitude.

3 Results

3.1 Precipitation spatial pattern

Figure 1 shows the precipitation climatology for the four databases. For wet (Figs. 1a-d) and dry (Figs. 1e-h) periods, CHIRPS data (Figs. 1b and 1f) underestimates the precipitation in the northwest portion of the AB: an area with lower precipitation (less than 5 mm day⁻¹) appears in the climatology between 3°S-7°S, 75°W-73°W. We can also notice differences in the Andes-Amazon transition region, where the hotspot rainfall regions are located. MSWEP (Fig. 1c) shows three centers of higher precipitation (> 15 mm day⁻¹) in this transitional area, while the other databases (Figs. 1a, 1b and 1d) characterize the precipitation as a narrow band, oriented in the northwest/southeast direction from 8°S to 18°S, approximately. This difference is most notable at the onset of the rainy season (SON) and its demise (MAM) (not shown).

Despite these differences in the precipitation climatology (Fig. 1), all databases identified the same rainfall AB sub-regimes through the cluster analysis (Fig. 2 and Fig. S1). As in previous studies (e.g., Mayta et al. 2020), six rainfall sub-regimes are defined: the North Amazon Basin (NAB, R1), the Northwest Amazon Basin (NWAB, R2), the Northeast Amazon Basin (NEAB, R3), the Central Amazon Basin (CAB, R4), the Southwest Amazon Basin (SWAB, R5), and the Southeast Amazon Basin (SEAB, R6).

For normalized precipitation (Fig. 2), the four databases identify the same sub-regimes, with the exception of slight differences in the spatial pattern in NAB and NWAB between HYBAM (Fig. 2a) and the others (Figs. 2b-d). Despite these differences, the annual cycle of daily precipitation (Figs. 2e-j) is almost the same for all datasets. For the non-normalized precipitation case (see Fig. S1 in the supplementary material), CHIRPS's misrepresentation of the rainfall in the AB northwestern portion, also identified in the climatology (Figs. 1b and 1f), is even more evident. CHIRPS (Fig. S1b) groups this region together with SWAB, while the other bases classify this region within the central basin sub-regime (R4, CAB).

Comparing the spatial distribution of the normalized (Figs. 2a-d) and non-normalized (Figs. S1a-d) precipitation sub-regimes, the largest differences occur for SWAB. This region covers a narrow area in the extreme west of the AB in the non-normalized case (Fig. S1), featuring less seasonality in the annual precipitation cycle (Fig. S1i), with a dry winter and a summer with lower precipitation ($\sim 5 \text{ mm day}^{-1}$). In the normalized case (Fig. 2), SWAB covers the entire region of hotspots and characterizes a typical monsoonal regime in the precipitation annual cycle (Fig. 2i), with a wet summer (DJF) and a dry winter (JJA). Another difference occurs for the SEAB region, which covers almost the entire southern portion of the basin in the non-normalized case (Fig. S1), although its annual precipitation cycle (Fig. S1j) is very similar to the normalized case (Fig. 2j).

Complementing the cluster analysis (Fig. 2 and Fig. S1), the SI pattern depicted in Fig. 3 also shows that the four rainfall databases capture the seasonality differences of the AB precipitation. In the NWAB, the SI indicates an area of equable rainfall regime (SI between 0.20 and 0.39), with no drier period defined (see NWAB boxplot in Fig. 2e). In the central, eastern and southwestern areas, SI is between 0.4 and 0.59, which indicates rather seasonal regimes, with short drier seasons. In fact, if we compare this result with the precipitation annual cycle obtained by the cluster analysis (Figs. 2g-i), we can observe that for these areas there is a dry period lasting only two to three months during the austral winter. For SEAB, SI indicates a seasonal regime (SI between 0.60 to 0.79), with well-defined wet and dry periods (Fig. 2j).

3.2 KGE comparison

KGE metric (Fig. 4) quantified the performance of each dataset based mainly on remote sensing (CHIRPS, MSWEP and TRMM) compared to observed-derived HYBAM data. Table 2 also shows the average KGE considering the entire basin and for different seasons. Overall, CHIRPS has the highest KGE values (Fig. 4a), considering the entire year ($\text{KGE} = 0.35$). A similar performance is observed for all seasons, except for the May-September and JJA period, in which MSWEP performs better. TRMM is the database with the worst performance (lowest correlation and highest variability ratio of the precipitation across the basin, Figs. 4c, 4f and 4i), with an average KGE of 0.14 considering the entire year.

Table 2
Mean KGE for each precipitation database for each season of the year

KGE	CHIRPS	MSWEP	TRMM
DJF	0.267	0.164	-0.002
MAM	0.308	0.217	0.089
JJA	0.134	0.210	-0.011
SON	0.315	0.227	0.087
Dry	0.224	0.245	0.075
Wet	0.326	0.222	0.090
All year	0.347	0.272	0.143
For each line, values in bold represent the highest KGE for each season.			

In general, Fig. 4 shows that all databases perform better (higher KGE values) for the AB eastern portion (Figs. 4a-c). It is interesting to note that CHIRPS (Fig. 4a) presents low KGE and underestimates the NWAB precipitation (Fig. 4g), which was already expected due to the previous results of the climatology (Fig. 1) and cluster analysis (Fig. S1).

KGE is also averaged (Table 3) for a representative area of each one of the six sub-regions identified in the cluster analysis. CHIRPS shows poor performance for the NWAB and the best representation of the observed precipitation for the NEAB and SEAB regions. The CHIRPS performance is better at the onset of the rainy season (SON) and for the entire rainy season. On the other hand, MSWEP has the best results for the NAB, NEAB and SEAB regions. For all databases, considering the entire year, the worst KGE values occur for the NWAB (0.073 for CHIRPS, 0.143 for MSWEP and 0.020 for TRMM). In this region, the correlation between CHIRPS, MSWEP and TRMM with HYBAM is low (Figs. 4d-f). CHIRPS (Fig. 4g) underestimates the precipitation, while MSWEP (Fig. 4h) and TRMM (Fig. 4i) overestimate it. In addition, there is a higher precipitation variability ratio for all databases at NWAB.

Table 3
Mean KGE for each AB sub-region

Sub-regions	NAB	NWAB	NEAB	CAB	SWAB	SEAB
All year						
CHIRPS	0.414	0.073	0.452	0.316	0.406	0.542
MSWEP	0.410	0.143	0.359	0.230	0.245	0.443
TRMM	0.195	0.020	0.164	0.090	0.125	0.279
Wet (Oct-Apr)						
CHIRPS	0.461	0.199	0.464	0.298	0.333	0.430
MSWEP	0.418	0.156	0.332	0.185	0.143	0.323
TRMM	0.216	0.028	0.149	0.062	0.017	0.134
Dry (May-Sep)						
CHIRPS	0.293	-0.084	0.391	0.236	0.296	0.452
MSWEP	0.321	0.101	0.373	0.229	0.236	0.384
TRMM	0.079	-0.013	0.147	0.025	0.069	0.179
For each column, values in bold represent the highest KGE values for each sub-region for each period						

Low KGE values are also found over the hotspot rainfall regions (SWAB sub-region). MSWEP and TRMM show high variability in this region, mainly during the austral summer. Both datasets also underestimate the precipitation over this region ($\beta < 1$ in Figs. 4h-i). Because MSWEP uses TRMM data as an input, the poor representation of the precipitation in this region by the satellite would justify the better performance of CHIRPS for the precipitation in the hotspots. But it is important to observe that CHIRPS also uses TRMM-3B42 in its calibration, but it gives more weight to the gauge data (Funk et al. 2015).

3.3 Rainfall spectral frequencies

Mayta et al. (2020) computed HYBAM precipitation power spectrum for each AB sub-region, with focus on the intraseasonal band (30 to 60 days). In this study, spectral analysis (Fig. 5) is computed for the representative areas of each sub-regime identified in the cluster analysis, considering the entire year. In general, we found similar spectral peaks to those documented in Mayta et al. (2020).

For the NAB region (Figs. 5a-d) only TRMM (Fig. 5d) exhibits an intense and significant peak at the 30 to 60 days intraseasonal band. For all databases, precipitation significantly peaks at a 5 day period, representing the influence of the synoptic scale in NAB. TRMM also depicts a small but statistically significant peak at the 1–3 days scale, which represents, for instance, the sea breeze influence and mesoscale convective complexes in the precipitation in the NAB (Reboita et al. 2010). It is worth

mentioning that the frequency of HYBAM is higher than the other databases for all sub-regions, so that the scales of the power versus frequency axis are different for HYBAM.

In the NWAB (Figs. 5e-h), only HYBAM and TRMM exhibit significant peaks at the intraseasonal band, with a maximum at 45 days period for TRMM (which is the mean life cycle of the MJO). Previous studies had already shown that intraseasonal variability induces a precipitation dipole between NWAB and SEAB regions (Mayta et al. 2019), and pronounced peaks in the intraseasonal spectral band are expected. CHIRPS, TRMM and MSWEP databases capture spectral peaks at the 5–10 days band, and this high-frequency peak is not pronounced for HYBAM data. Differences between these databases are also observed in the KGE analysis, which indicates some differences in the rainfall representation over the NWAB region between TRMM, CHIRPS and MSWEP compared to HYBAM (Fig. 4).

For the NEAB region (Figs. 5i-l), the intraseasonal peak barely reaches the 95% significance level (dashed black line) for all datasets. In contrast, all databases peak at the 5 day timescale. The TRMM and MSWEP peaks at the sub-monthly scale (10 to 30 days) also reach the significance curve, which is not observed for CHIRPS and HYBAM data in NEAB. For the CAB region (Figs. 5m-p), as also discovered by Mayta et al. (2020), there are more pronounced high frequency peaks (5–7 days), which were identified by MSWEP, but for HYBAM and TRMM these peaks are lower and barely reach the 95% curve.

In the SWAB (Figs. 5m-p), only HYBAM rainfall peaks significantly at the intraseasonal timescale. For the other databases, high frequency peaks are greater in amplitude than the intraseasonal. The SEAB (Figs. 5q-t) is also affected by the precipitation dipole triggered by the intraseasonal variability in the AB during the SAMS rainy season (Mayta et al. 2020). This region exhibits low-frequency (30 and 60 days) intraseasonal peaks for all databases. With the exception of HYBAM, all the other databases also capture significant sub-monthly peaks in the 20–30 day band. Comparing all AB sub-regions, the SEAB region shows similar spectral peaks for all databases. For SEAB, high frequency energy peaks are less intense in magnitude compared to the other sub-regions and almost not at all significant at the 95% level.

3.4 The intraseasonal precipitation pattern

The rainfall power spectra show significant peaks at the intraseasonal band for almost all AB regions, but differences in the magnitude of these peaks between the different datasets imply different spatial patterns of the intraseasonal precipitation. Figures 6 and 7 depict composites of band-pass filtered rainfall in the low frequency intraseasonal band (30–90 days) for wet and dry periods, respectively, for the OMI (Kiladis et al. 2014) convective index phases. This MJO index was chosen because it is better able to capture the MJO impacts on the AB precipitation, as shown by Mayta et al. (2020).

During the wet period (Fig. 6), for all databases there is a precipitation dipole between the NWAB (Figs. 6a-d) and the SEAB (Figs. 6i-l) for the OMI phases 8 and 1 and 4 and 5. According to Grimm (2019), the MJO active phase over South America (phases 8 and 1) induces a precipitation dipole between the SACZ region (which includes SEAB) and the SESA, favoring the rainfall in the southern portion of the AB. This pattern was detected by all databases. It is interesting to note that CHIRPS presents negative precipitation

anomalies during phases 8 and 1 over the NWAB (Fig. 6b). These anomalies are less intense than the anomalies for the other databases, which once again confirms the deficiency of CHIRPS in representing the precipitation in this AB portion. Through the KGE analysis, CHIRPS also presented an overall underestimation of the rainfall in the NWAB region (Fig. 4a).

For the inactive MJO phase in South America (phases 4 and 5, Figs. 6i-l) during the wet season, the pattern is of positive anomalies of outgoing long-wave radiation over the central-eastern portion of the continent (Vera et al. 2018; Grimm 2019). This pattern induces negative rainfall anomalies across the southern and eastern AB regions and positive ones in the NWAB. The extent of these positive precipitation anomalies over the NWAB is greater for HYBAM (Fig. 6i) and MSWEP (Fig. 6k), smaller for TRMM (Fig. 6l), and the anomalies are weaker for CHIRPS (Fig. 6j). For TRMM, the suppressed convection over the basin in phases 4 and 5 is represented by intense negative precipitation anomalies over almost the entire basin extension (with the exception of the NWAB).

In general, for the transition MJO phases (phases 2, 3, 6 and 7) in the wet season (Fig. 6e-h and Figs. 6m-p), the pattern of the anomalies is similar for all databases. However, there is an exception for a narrow area over the NWAB western edge, where CHIRPS and TRMM provide larger positive anomalies in phases 2 and 3 (Figs. 6f and 6h) and negative ones in phases 6 and 7 (Figs. 6n and 6p). For this region, HYBAM and MSWEP intraseasonal anomalies are weaker.

During the dry period (Fig. 7), the discrepancy between CHIRPS and TRMM anomalies at the northwest edge of the basin is observed when compared to HYBAM and MSWEP. CHIRPS provides intense negative precipitation anomalies for phases 4 and 5 (Fig. 7i) and phases 6 and 7 (Fig. 7n) in the NWAB region (and positive in phases 8 and 1, Fig. 7b, and phases 2 and 3, Fig. 7f), which is also observed for TRMM. It is worth mentioning that these CHIRPS results may also be a reflection of its low performance in NWAB for the dry period, as already indicated by the KGE analysis (Table 3).

With the exception of NWAB, in general, the databases agree closely as to the representation of the intraseasonal rainfall in all the AB regions for the dry period (Fig. 7). For the wet period (Fig. 6) there is a greater similarity between the intraseasonal precipitation pattern for the HYBAM and MSWEP data in all OMI phases.

3.5 Diurnal variability

To complete the comparisons of the precipitation between different databases at different timescales, the rainfall diurnal variability is compared for TRMM and MSWEP. Figure 8 shows the comparison of the precipitation diurnal cycle for three locations in the AB: Manaus (GoAmazon campaign), Rondônia state in Brazil (TRMM-LBA campaign), and a site located in the Andes region (Huayao station data).

At the GoAmazon station (Fig. 8b), the precipitation peak occurs around 18:00 UTC according to the GoAmazon radar. The precipitation rate peak in TRMM data also occurs at 18:00 UTC, but there the maximum intensity is underestimated. In general, the TRMM diurnal cycle underestimates the diurnal observed precipitation by about 0.55mm. MSWEP represents the magnitude of the rainfall peak well, but

its maximum occurs 3h before the observed peak by the radar. This fact contributed to a lower KGE between the estimated diurnal cycle by MSWEP compared to the radar (Fig. 8a). Although the accumulated precipitation of the MSWEP diurnal cycle is almost the same as that observed (the difference between the radar diurnal accumulated rainfall and that of MSWEP is only 0.03mm), the anticipation of the rainfall peak by MSWEP makes its KGE (KGE = 0.61) lower compared to TRMM (KGE = 0.71), which better represents the observed curve of the diurnal cycle (Fig. 8b).

For the LBA networks (Figs. 8c-e), the diurnal cycle, previously described by Marengo et al. (2004), is represented well by MSWEP and TRMM during the morning, as expected, due to the strong dependence of the MSWEP estimate on the TRMM results. However, considerable differences occur by the end of the afternoon, when the deep convection is well organized. For N2 site (Fig. 8c), MSWEP satisfactorily represents the observed rainfall curve, but again the precipitation peak occurs 3h in advance compared with the observations. The opposite occurs for TRMM, which overestimates the precipitation during the morning, but the peak of the rainfall coincides with that observed at 21:00 UTC. At N3 (Fig. 8d), the MSWEP precipitation rate is nearly two orders of magnitude larger than that observed at 18:00 UTC in the LBA data. In fact, the worst KGE result for MSWEP is for N3 (Fig. 8a). At N4 (Fig. 8e), a first low rainfall rate peak is observed in the early morning (9:00 UTC), which is registered at 6:00 UTC for TRMM and MSWEP. A second and more intense peak occurs in the afternoon (18:00 UTC), which is also observed in the MSWEP data, while TRMM overestimates by three times the maximum hourly precipitation and shows this peak 3h later than observed (at 21:00 UTC). For LBA data, in general, MSWEP represents the magnitude of the precipitation in the morning well, but it overestimates the rainfall in the afternoon and night, while TRMM overestimates rainfall rates throughout the day, but it “hits” the rainfall peak time.

At the Huayao station in the Andes region (Fig. 8f), as shown by Saavedra et al. (2020), the precipitation rate peaks around 00:00 UTC. This maximum is also observed in the TRMM data, which greatly underestimates the rainfall peak. MSWEP, on the other hand, represents the magnitude of the maximum hourly precipitation, but again indicates the peak 3h before the observed precipitation. From 09:00 UTC to 15:00 UTC, the Huayao station recorded almost no precipitation, similarly to TRMM and MSWEP.

4 Summary And Conclusions

In the present study, an objective comparison between four rainfall databases over the AB region was carried out. These databases included one rain-gauge based dataset (HYBAM) and three datasets based on remote sensing (CHIRPS, TRMM and MSWEP). We compared the performance of these datasets through different timescales: diurnal, seasonal and intraseasonal.

Rainfall precipitation for the dry and wet periods are similar for each one of the databases, but it is clear that CHIRPS has a deficiency to represent the rainfall in the northwest portion of the basin (Fig. 1). In this region, there is an area with lower precipitation in the CHIRPS climatology (Figs. 1b and 1f). Its performance measured by KGE in NWAB indicated higher variability ratio (Fig. 4j) and a low correlation (Fig. 4d) compared to HYBAM, and a poor representation of the intraseasonal precipitation variability

there (Figs. 6 and 7). However, for the rest of the basin, CHIRPS has the best performance compared to HYBAM (Fig. 4), especially during the wet period (Tables 2 and 3) and for the hotspot area.

The deficiency in representing the precipitation in NWAB is also present in the TRMM and MSWEP estimates. All remote sensing-based datasets considered in this study use TRMM-3B42 in their constitution. Thus, it is expected that they share similar discrepancies in the rainfall representation. This feature is also noticeable for the spectral analysis: NWAB is a region affected by the intraseasonal variability through a seesaw pattern with SEAB during the active phase of the MJO in the South America rainy season (Mayta et al. 2018; Grimm 2019). Thus, it is expected to find intraseasonal peaks of the rainfall in the NWAB and SEAB regions. However, this peak was only detected by HYBAM and TRMM in the NWAB (Fig. 5), while the other databases peaked significantly at high frequencies. The major differences of intraseasonal precipitation spatial distribution between the databases was also observed for the NWAB region (Figs. 6 and 7).

At the diurnal timescale (Fig. 8), MSWEP represents the magnitude of maximum rainfall rates well but not the timing of the rainfall peaks. On the other hand, TRMM, in general, matches the time of maximum rainfall but underestimates the precipitation. These timing mismatches could be a result of how each database computed the 24-hour precipitation. For MSWEP data, Beck et al. (2019) applied a technique to infer gauge reporting times and to apply daily gauge corrections, but they stated that these inferred reporting times are subject to an error of at most 1.5h, which could justify the rainfall timing anticipation for MSWEP. Besides, gauge reporting times may be not constant, which can invalidate this correction applied to MSWEP.

The present study has emphasized the differences between rainfall databases for the AB region and how much the performance of these datasets varies at different timescales. As elucidated previously by Costa and Foley (1998) and Juárez et al. (2009), this kind of study is extremely necessary in order to warn of the importance of a more complete and objective assessment of the data before considering it for applications in different precipitation studies, mainly in regions with high rainfall heterogeneity like the AB.

Our study also assessed the potential use of MSWEP for the AB, since this is a database with extremely high temporal and spatial resolution, in addition to providing a long period of data from 1979 to ~ 4 hours from real-time. In fact, MSWEP is able to represent the climatology of the precipitation in the AB (Fig. 1), its different precipitation regimes (Fig. 2), and its seasonality characteristics (Fig. 3). Furthermore, MSWEP can represent the spatial pattern of the intraseasonal precipitation in NWAB and SEAB (Fig. 6), which are the regions most affected by the precipitation seesaw due to the MJO activity. However, at the diurnal scale, caution must be taken when using MSWEP, as it does not identify the exact time of peaking rainfall, which is important information for analysis at the diurnal timescale.

It is suggested that future analysis includes other precipitation databases for the AB, such as CPC, PERSIANN and IMERG (Huffman et al. 2018). It is also recommended to extend the analysis to other timescales, such as the interannual, in order to evaluate how different datasets represent the impacts of

ENSO and the Tropical North Atlantic in the precipitation anomalies in the AB. These analyses should also be expanded to other regions, such as for the entire South America or other hydrological basins.

Declarations

Acknowledgements

This research was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and partially by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under grant 130629/2019-7. The authors would like to thank Dr. Hilk Beck for making the MSWEP V2 data available and the MASTER Laboratory of the Institute of Geophysics, Astronomy and Atmospheric Sciences at the University of São Paulo for the data from the LBA WET-AMC campaign. VM was supported by the National Science Foundation under grant AGS-1841559.

Conflict of Interest:

The authors declare no conflicting/competing interests.

Funding Statement:

Camila Sapucci was supported by CAPES and partially by CNPq under grant 130629/2019-7. Victor C. Mayta was supported by the National Science Foundation under grant AGS-1841559.

Author's Contribution:

Camila Sapucci: conceptualization, methodology, formal analysis, validation and writing the original draft. Victor C. Mayta: conceptualization, investigation, supervision and writing-review. Pedro L. Silva Dias: conceptualization, supervision and writing-review.

Availability of data and material:

All data and material published in this article are available upon reasonable request.

Code availability:

Most analyses have been carried out by Matlab and NCL codes, which are available upon request from the authors.

Ethics approval:

The authors confirm that this is an original paper and has not been published previously in any journal.

Consent to participate and for publication:

The authors agree to publish this paper.

References

1. Adler RF et al. (2003) The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.* (4): 1147–1167. [https://doi.org/10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2)
2. Ashouri H, Hsu KL, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Prat OP (2015) PERSIANN-CDR: daily precipitation climate data record from multi-satellite observations for hydrological and climate studies. *Bull Am Meteor Soc*, 96(1): 69–83. <https://doi.org/10.1175/BAMS-D-13-00068.1>
3. Avissar R, Silva Dias PL, Silva Dias MAF, Nobre C (2002) The Large-Scale Biosphere-Atmosphere Experiment in Amazonia(LBA): Insights and future research needs. *J Geophys Res* 107(D20):8086. <https://doi.org/10.1029/2002JD002704>
4. Beck HE, van Dijk AIJM, Levizzani V, Schellekens J, Miralles DG, Martens B, de Roo A (2017) MSWEP, 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol Earth Syst Sci Discuss* 21:1–38. <https://doi.org/10.5194/hess-21-589-2017>
5. Beck HE, Wood EF, Pan M, Fisher CK, Miralles DG, van Dijk AIJM, McVicar TR, Adler RF (2019) MSWEP V2 GLOBAL 3-HOURLY 0.1° PRECIPITATION: Methodology and Quantitative Assessment. *Bull Am Meteor Soc* 100:473–500. <https://doi.org/10.1175/BAMS-D-17-0138.1>
6. Bettolli ML, Solman SA, da Rocha RP et al (2021) The CORDEX Flagship Pilot Study in southeastern South America: a comparative study of statistical and dynamical downscaling models in simulating daily extreme precipitation events. *Clim Dyn* 56:1589–1608. <https://doi.org/10.1007/s00382-020-05549-z>
7. Buarque DC, Paiva RCD, Clarke RT, Mendes CAB (2011) A comparison of Amazon rainfall characteristics derived from TRMM, CMORPH and the Brazilian national rain gauge network. *J Geophys Res* 116:D19105. <https://doi.org/10.1029/2011JD016060>
8. Carvalho LMV, Jones C, Liebmann, B. (2002) Extreme precipitation events in southeastern South America and large-scale convective patterns in the South Atlantic Convergence Zone. *Journal of Climate*, 15, 2377–2394. [https://doi.org/10.1175/1520-0442\(2002\)015<2377:EPEISS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2377:EPEISS>2.0.CO;2)
9. Casarin DP, Kousky VE (1986) Anomalias de precipitação no sul do Brasil e variações na circulação atmosférica. *Rev Bras Meteorol* 2(83)

10. Cerón WL, Molina-Carpio J, Ayes Rivera I et al (2020) A principal component analysis approach to assess CHIRPS precipitation dataset for the study of climate variability of the La Plata Basin, Southern South America. *Nat Hazards* 103: 767–783. <https://doi.org/10.1007/s11069-020-04011-x>
11. Chavez SP, Takahashi K (2017) Orographic rainfall hot spots in the Andes-Amazon transition according to the TRMM precipitation radar and in situ data. *J Geophys Res Atmos* 122:5870–5882. <https://doi.org/10.1002/2016JD026282>
12. Clarke RT, Buarque DC, Paiva RCD, Collischonn W (2011) Issues of spatial correlation arising from the use of TRMM rainfall estimates in the Brazilian Amazon, *Water Resour. Res* (47): W05539. <https://doi.org/10.1029/2010WR010334>
13. Cohen JCP; Silva Dias MAF.; Nobre CA (1995) Environmental Conditions Associated with Amazonian Squall Lines: A Case Study. *Monthly Weather Review* 123: 3163–3174. [https://doi.org/10.1175/1520-0493\(1995\)123<3163:ECAWAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<3163:ECAWAS>2.0.CO;2)
14. Collischonn B, Collischonn W, Tucci CEM (2008) Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *J Hydrol* 360(1–4):207–216. <https://doi.org/10.1016/j.jhydrol.2008.07.032>
15. Condom T, Rau P, Espinoza JC (2011) Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007. *Hydrol Process* 25:1924–1933. <https://doi.org/10.1002/hyp.7949>
16. Correa SW, Paiva RCD, Espinoza JC, Collischonn W (2017) Multi-decadal Hydrological Retrospective: Case study of Amazon floods and droughts. *J Hydrol* 549:667–684. <https://doi.org/10.1016/j.jhydrol.2017.04.019>
17. Costa MH, Foley JA (1998) A comparison of precipitation datasets for the Amazon basin. *Geophys Res Lett* 25(2):155–158. <https://doi.org/10.1029/97GL03502>
18. Dee DP et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. <https://doi.org/10.1002/qj.828>
19. Duchon CE (1979) Lanczos filtering in one and two dimensions. *J. Appl. Meteorol.* 18(8): 1016–1022. [https://doi.org/10.1175/1520-0450\(1979\)018<1016:LFIOAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<1016:LFIOAT>2.0.CO;2)
20. Espinoza JC, Ronchail J, Guyot JL, Cochonneau G, Naziano F, Lavado W, Oliveira E, Pombosa R, Vauchel P (2009) Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *Int J Climatol* 29:1574–1594. <https://doi.org/10.1002/joc.1791>
21. Espinoza JC, Chavez S, Ronchail J, Junquas C, Takahashi K, Lavado W (2015) Rainfall hotspots over the southern tropical Andes: spatial distribution, rainfall intensity and relations with large-scale atmospheric circulation. *Water Resour Res* 51:3459–3475. <https://doi.org/10.1002/2014WR016273>
22. Espinoza JC, Ronchail J, Marengo JA et al (2019) Contrasting North–South changes in Amazon wet-day and dry-day frequency and related atmospheric features (1981–2017). *Clim Dyn* 52:5413–5430. <https://doi.org/10.1007/s00382-018-4462-2>
23. Fagundes HO, Correa SW, Paiva RCD (2017) Avaliação dos dados de precipitação estimados pelo MSWEP para a bacia Amazônica. *Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto*,

24. Ferreira RN, Rickenback TM, Herdies DL, Carvalho LMV (2003) Variability of South American convective cloud systems and tropospheric circulation during January–March 1998 and 1999. *Monthly Weather Review* 131(5), 961–973. [https://doi.org/10.1175/1520-0493\(2003\)131<0961:VOSACC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0961:VOSACC>2.0.CO;2)
25. Ferreira NJ, Ramírez MV, Gan MA (2009) Vórtices Ciclônicos de Altos Níveis que Atuam na Vizinhança do Nordeste Brasileiro. In: Cavalcanti IFA, Ferreira NJ, Silva MGAJ DIAS, MAFS (Org). *Tempo e Clima no Brasil*, 1st edn. Oficina de Textos, São Paulo, pp 43–60
26. Figueroa SN, Nobre CA (1990) Precipitations distribution over central and western tropical South America. *Climanálise: Boletim de Monitoramento e Análise Climática* 5(6): 36–45
27. Fitzjarrald DR, Sakai RK, Moraes OLL, Cosme de Oliveira R, Acevedo OC, Czikowsky MJ, Beldini T (2008) Spatial and temporal rainfall variability near the Amazon-Tapajós confluence. *J Geophys Res* 113(B11):G00. doi:10.1029/2007JG000596
28. Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2010) The climate hazards infrared precipitation with stations – a new environmental record for monitoring extremes. *Scientific Data* 2:150066. <https://doi.org/10.1038/sdata.2015.66>
29. Getirana ACV, Espinoza JCV, Ronchail J, Rotunno Filho OC (2011) Assessment of different precipitation datasets and their impacts on the water balance of the Negro River basin. *J Hydrol* 404(3–4):304–322. <https://doi.org/10.1016/j.jhydrol.2011.04.037>
30. Grimm AM (2019) Madden–Julian Oscillation impacts on South American summer monsoon season: precipitation anomalies, extreme events, teleconnections, and role in the MJO cycle. *Clim Dyn* 53:907–932. <https://doi.org/10.1007/s00382-019-04622-6>
31. Guimberteau M, Drapeau G, Ronchail J, Sultan B, Polcher J, Martinez J-M, Prigent C, Guyot J-L, Cochonneau G, Espinoza JC, Filizola N, Fraizy P, Lavado W, De Oliveira E, Pombosa R, Noriega L, Vauchel P (2012) Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets. *Hydrol Earth Syst Sci* 16:911–935. <https://doi.org/10.5194/hess-16-911-2012>
32. Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J Hydrol* 377:80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
33. Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007) The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J Hydrometeorol* 8:38–55. <https://doi.org/10.1175/JHM560.1>
34. Huffman GJ, Adler RF, Bolvin DT, Nelkin EJ (2010) The TRMM Multi-Satellite Precipitation Analysis (TMPA). In: Gebremichael M, Hossain F (eds) *Satellite Rainfall Applications for Surface Hydrology*. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-2915-7_1
35. Huffman GJ, Adler RF, Morrissey MM, Bolvin DT, Curtis S, Joyce R, McGavock B, Susskind J (2001) Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of*

- Hydrometeorology 2 (1):36–50. [https://doi.org/10.1175/1525-7541\(2001\)002<0036:GPAODD>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2)
36. Huffman GJ et al (2018) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD). Version 5.2
 37. Janowiak JE, Arkin PA (1991) Rainfall variations in the tropics during 1986–1989, as estimated from observations of cloud-top temperature. *J Geophys Res* 96(S01):3359–3373. [doi:10.1029/90JD01856](https://doi.org/10.1029/90JD01856)
 38. Jones C, Carvalho LMV (2002) Active and Break Phases in the South American Monsoon System. *Journal of Climate* 15, 905–914. [https://doi.org/10.1175/1520-0442\(2002\)015<0905:AABPIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0905:AABPIT>2.0.CO;2)
 39. Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.* 5: 487–503. [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2)
 40. Juárez RIN, Li W, Fu R, Fernandes K, Oliveira Cardoso A (2009) Comparison of Precipitation Datasets over the Tropical South American and African Continents. *J Hydrometeorol* 10(1):289–299. <https://doi.org/10.1175/2008JHM1023.1>
 41. Kalnay EM et al. (1996) The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Met. Soc.* (77): 437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)
 42. Kanamitsu M, Ebisuzaki W, Woolen J, Yang S-K, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP-DOE AMIP-II Reanalysis (R-2). *Bull Amer Meteor Soc* 83:1631–1643. <https://doi.org/10.1175/BAMS-83-11-1631>
 43. Kiladis GN, Dias J, Straub KH, Wheeler MC, Tulich SN, Kikuchi K, Weickmann KM, Ventrone MJ (2014) A comparison of OLR and circulation-based indices for tracking the MJO. *Mon Weather Rev* 142:1697–1715. <https://doi.org/10.1175/MWR-D-13-00301.1>
 44. Kling H, Fuchs M, Paulin M (2012) Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *J Hydrol* 424–425:264–277. <https://doi.org/10.1016/j.jhydrol.2012.01.011>
 45. Knapp KR et al (2011) Globally gridded satellite observations for climate studies. *Bull Amer Meteor Soc* 92:893–907. <https://doi.org/10.1175/2011BAMS3039.1>
 46. Kobayashi S et al (2015) The JRA-55 Reanalysis: General specifications and basic characteristics. *J Meteor Soc Japan* 93:5–48. <https://doi.org/10.2151/jmsj.2015-001>
 47. Liebmann B, Kiladis GN, Marengo J, Ambrizzi T, Glick JD (1999) Submonthly convective variability over South America and the South Atlantic convergence zone. *J Clim* 12(7): 1877–1891. [https://doi.org/10.1175/1520-0442\(1999\)012<1877:SCVOSA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1877:SCVOSA>2.0.CO;2)
 48. Liebmann B, Smith CA (1996) Description of a complete (interpolated) outgoing long-wave radiation dataset. *Bull Am Meteorol Soc* 77:1275–1277
 49. Machado LAT et al (2014) The Chuva Project: How Does Convection Vary across Brazil? *Bull Am Meteor Soc* 95(9):1365–1380. <https://doi.org/10.1175/BAMS-D-13-00084.1>

50. Madden RA, Julian PR (1971). Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *Journal of Atmospheric Science* 28: 702–708. [https://doi.org/10.1175/1520-0469\(1971\)028<0702:DOADOI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2)
51. Madden RA, Julian PR (1972). Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of Atmospheric Science* 29: 1109–1123. [https://doi.org/10.1175/1520-0469\(1972\)029<1109:DOGSCC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2)
52. Marengo JA (2004) Interdecadal variability and trends of rainfall across the Amazon basin. *Theor Appl Climatol* 78:78–96. <https://doi.org/10.1007/s00704-004-0045-8>
53. Marengo JA (2006) ON THE HYDROLOGICAL CYCLE OF THE AMAZON BASIN: A HISTORICAL REVIEW AND CURRENT STATE-OF-THE-ART. *Revista Brasileira de Meteorologia* 21(3): 1–19
54. Marengo JA, Fisch G, Morales C, Vendrame I, Dias PC (2004) Diurnal variability of rainfall in Southwest Amazonia during the LBA-TRMM field campaign of the Austral summer of 1999. *Acta Amaz* 34(4):593–603. <https://doi.org/10.1590/S0044-59672004000400011>
55. Marengo JA et al (2012) Recent developments on the South American monsoon system. *Int J Climatol* 32:1–21. <https://doi.org/10.1002/joc.2254>
56. Mayta VC, Ambrizzi T, Espinoza JC, Silva Dias PL (2019) The role of the Madden–Julian oscillation on the Amazon Basin intraseasonal rainfall variability. *Int J Climatol* 39(1):343–360. <https://doi.org/10.1002/joc.5810>
57. Mayta VC, Silva NP, Ambrizzi T, Silva Dias PL, Espinoza JC (2020) Assessing the skill of all-season diverse Madden–Julian oscillation indices for the intraseasonal Amazon precipitation. *Clim Dyn* 54:3729–3749. <https://doi.org/10.1007/s00382-020-05202-9>
58. Negri AJ, Xu X, Adler RF (2002) A TRMM-calibrated infrared rainfall algorithm applied over Brazil. *J Geophys Res* 107(D20):8048. <https://doi.org/10.1029/2000JD000265>
59. Nobre CA, Obregón GO, Marengo JA, Fu R, Poveda G (2009) Characteristics of Amazonian Climate: Main Features. In: Keller M, Bustamante M, Gash J, Silva Dias P (Eds). *Amazonia and Global Change. Geophysical Monograph Series 186*, Washington, pp 149–162
60. Nogués-Paegle J, Mo KC (1997) Alternating Wet and Dry Conditions over South America during Summer. *Monthly Weather Review* 125: 297–291. [https://doi.org/10.1175/1520-0493\(1997\)125<0279:AWADCO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0279:AWADCO>2.0.CO;2)
61. Paca VHM, Espinoza-Dávalos GE, Moreira DM, Comair G (2020) Variability of Trends in Precipitation across the Amazon River Basin Determined from the CHIRPS Precipitation Product and from Station Records. *Water* 12(5):1244. <https://doi.org/10.3390/w12051244>
62. Paiva RCD, Collischonn W, Tucci CEM (2011) Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *J Hydrol* 406(3–4):170–181. <https://doi.org/10.1016/j.jhydrol.2011.06.007>
63. Pereira Filho AJ, Carbone RE, Janowiak JE, Arkin P, Joyce R, Hallak R, Ramos CGM (2010) Satellite rainfall estimates over South America - Possible applicability to the water management of large

- watersheds. *Journal Of The American Water Resources Association* 46:344–360.
<http://dx.doi.org/10.1111/j.1752-1688.2009.00406.x>
64. Pinto LIC, Costa MH, Lima FZ, Diniz LMF, Sediya GC, Pruski FF (2009) Comparação de produtos de precipitação para a América do Sul. *Revista Brasileira de Meteorologia* 24(4):461–472.
<https://doi.org/10.1590/S0102-77862009000400008>
 65. Rao VB, Hada K (1990) Characteristics of rainfall over Brazil annual variations and connections with the Southern Oscillation. *Theoret Appl Climatol* 42:81–91. <https://doi.org/10.1007/BF00868215>
 66. Reboita MS, Gan MA, Rocha RP, Tércio A (2010) Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia* 25(2):185
 67. Rivera JA, Hinrichs S, Marianetti G (2019) Using CHIRPS Dataset to Assess Wet and Dry Conditions along the Semiarid Central-Western Argentina. *Advances in Meteorology*: Article ID 8413964.
<https://doi.org/10.1155/2019/8413964>
 68. Rosa EB, Pezzi LP, Quadro MFL, Brunsell N (2020) Automated Detection Algorithm for SACZ, Oceanic SACZ, and Their Climatological Features. *Front Environ Sci* 8:18.
<https://doi.org/10.3389/fenvs.2020.00018>
 69. Rozante JR, Moreira DS, Gonçalves LGG, Vila DA (2010) Combining TRMM and surface observations of precipitation: technique and validation over South America. *Weather Forecast* 25:885–894.
<https://doi.org/10.1175/2010WAF2222325.1>
 70. Rudolf B, Schneider U (2005) Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. *Proc. Second Workshop of the International Precipitation Working Group (IPWG), Monterey, CA, EUMETSAT*, 231–247
 71. Saavedra M, Junquas C, Espinoza JC, Silva Y (2020) Impacts of topography and land use changes on the air surface temperature and precipitation over the central Peruvian Andes. *Atmos Res* 234:104711. <https://doi.org/10.1016/j.atmosres.2019.104711>
 72. Santos EB, Lucio PS, Silva CMS (2015) Precipitation regionalization of the Brazilian Amazon. *Atmospheric Science Letters* 16:185–192. <https://doi.org/10.1002/asl2.535>
 73. Schamm K, Ziese M, Becker A, Finger P, Meyer-Christoffer A, Schneider U, Schröder M, Stender P (2014) Global gridded precipitation over land: a description of the new GPCC First Guess Daily product. *Earth Syst Sci Data* 6:49–60. <https://doi.org/10.5194/essd-6-49-2014>
 74. Schneider U, Fuchs T, Meyer-Christoffer A, Rudolf B (2008) Global Precipitation Analysis Products of the GPCC. Global Precipitation Climatology Centre (GPCC), DWD, Internet Publikation, 1–12
 75. Schumacher C, Funk A (2018) GoAmazon2014/5 Three-dimensional Gridded S-band Reflectivity and Radial Velocity from the SIPAM Manaus S-band Radar. United States: N. p., 2018.
doi:10.5439/1459573
 76. Shimizu MH, Amrizzi T, Liebmann B (2017) Extreme precipitation events and their relationship with ENSO and MJO phases over northern South America. *Int J Climatol* 37:2977–2989.
<https://doi.org/10.1002/joc.4893>

77. Solman SA, Bettolli ML, Doyle ME et al (2021) Evaluation of multiple downscaling tools for simulating extreme precipitation events over Southeastern South America: a case study approach. *Clim Dyn*. <https://doi.org/10.1007/s00382-021-05770-4>
78. Souza EB, Ambrizzi T (2006) Modulation of the intraseasonal rainfall over tropical Brazil by the Madden–Julian oscillation. *Int J Climatol* 26:1759–1776. <https://doi.org/10.1002/joc.1331>
79. Satyamurty P, Nobre CA, Silva Dias PL (1998) South America. In: Karoly DJ, Vincent DG (eds) *Meteorology of the Southern Hemisphere*. Meteorological Monographs, American Meteorological Society, Boston
80. Silva VBS, Kousky VE, Shi W, Higgins W (2007) An Improved Gridded Historical Daily Precipitation Analysis for Brazil. *J Hydrometeorol* 8(4):847–861. <https://doi.org/10.1175/JHM598.1>
81. Silva Dias MAF et al (2002) Cloud and rain processes in a biosphere-atmosphere interaction context in the Amazon Region. *J Geophys Res* 107(D20):8072. <https://doi.org/10.1029/2001JD000335>
82. Silva Dias MAF, Silva Dias PL, Longo M, Fitzjarrald DR, Denning AS (2004) River breeze circulation in eastern Amazonia: observations and modelling results. *Theor Appl Climatol* 78:111–121. <https://doi.org/10.1007/s00704-004-0047-6>
83. Teegavarapu RSV (2018) Changes and Trends in Precipitation Extremes and Characteristics: Links to Climate Variability and Change. In: Teegavarapu RSV (ed) *Trends and Changes in Hydroclimatic Variables*. Elsevier, pp 91–148
84. Uppala SM et al (2005) The ERA-40 re-analysis. *Quart J R Meteorol Soc* 131:2961–3012. <https://doi.org/10.1256/qj.04.176>
85. Ushio T et al (2009) A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J Meteor Soc Japan* 87A:137–151. <https://doi.org/10.2151/jmsj.87A.137>
86. Vera CS, Alvarez MS, Gonzalez PLM, Liebmann B, Kiladis GN (2018) Seasonal cycle of precipitation variability in South America on intraseasonal timescales. *Clim Dyn* 51(5–6):1991–2001. <https://doi.org/10.1007/s00382-017-3994-1>
87. Walsh RPD, Lawler DM (1981) RAINFALL SEASONALITY: DESCRIPTION, SPATIAL PATTERNS AND CHANGE THROUGH TIME. *Weather* 36:201–208. <https://doi.org/10.1002/j.1477-8696.1981.tb05400.x>
88. Wang H, Fu R (2002) Cross-equatorial flow and seasonal cycle of precipitation over South America. *J Clim* 15(13):1591–1608. [https://doi.org/10.1175/1520-0442\(2002\)015<1591:CEFASC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<1591:CEFASC>2.0.CO;2)
89. Wilks DS (2019) *Statistical Methods in the Atmospheric Sciences*. Elsevier, New York
90. Xie P, Arkin PA (1997). Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs, *Bulletin of the American Meteorological Society*, 78(11): 2539–2558. [https://doi.org/10.1175/1520-0477\(1997\)078<2539:GPAYMA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2)
91. Xie P, Chen M, Shi W (2010) CPC global unified gauge-based analysis of daily precipitation, Preprints, 24th conference on hydrology, Atlanta, GA, American Meteorological Society 2

92. Zhang C (2005) Madden-Julian Oscillation. *Reviews of Geophysics* 43(22): RG2003.
<https://doi.org/10.1029/2004RG000158>
93. Zhou J, Lau KM (1998) Does a Monsoon Climate Exist over South America? *Journal of Climate* 11: 1020–1040. [https://doi.org/10.1175/1520-0442\(1998\)011<1020:DAMCEO>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<1020:DAMCEO>2.0.CO;2)
94. Zubieta R, Saavedra M, Espinoza JC et al (2019) Assessing precipitation concentration in the Amazon basin from different satellite-based data sets. *Int J Climatol* 39:3171–3187.
<https://doi.org/10.1002/joc.6009>

Figures

Figure 1

Climatology of daily precipitation for the AB region for wet (October to April, a-d) and dry (May to September, e-h) periods for HYBAM, CHIRPS, MSWEP and TRMM databases

Figure 2

Precipitation sub-regimes for the AB obtained by the non-hierarchical (k-means) cluster analysis for the normalized precipitation. (a) to (d) represent the spatial distribution of each one of the six rainfall sub-regimes identified for HYBAM, CHIRPS, MSWEP and TRMM, respectively. The six sub-regions are namely: the North Amazon Basin (NAB, R1), the Northwest Amazon Basin (NWAB, R2), the Northeast Amazon Basin (NEAB, R3), the Central Amazon Basin (CAB, R4), the Southwest Amazon Basin (SWAB, R5) and the

Southeast Amazon Basin (SEAB, R6). (e) to (j) correspond to the precipitation annual cycle for each sub-region identified through the cluster analysis applied to HYBAM data. The annual precipitation cycle of each sub-region for CHIRPS, MSWEP and TRMM data is similar to HYBAM (not shown)

Figure 3

The SI calculated for each database: (a) HYBAM, (b) CHIRPS, (c) MSWEP, and (d) TRMM. $SI \leq 0.19$ describes regions with a very equable rainfall regime. SI between 0.20-0.39 represents an equable regime but with a well-defined wetter season. SI between 0.40-0.59 corresponds to a rather seasonal regime, with a short drier season. SI between 0.60-0.79 depicts a seasonal regime, with well-defined wet and dry seasons. SI between 0.80-0.99 describes a markedly seasonal regime, with a long drier season. SI between 1.0-1.19 represents a regime with most of the rain occurring in three or less months and $SI \geq 1.20$ corresponds to an extreme regime, with almost all the rainfall occurring in one or two months (Walsh and Lawler, 1981)

Figure 4

KGE (a-c), correlation (r , d-f), bias (β , g-i), and variability ratio (γ , j-l) metrics for the AB region for each one of the databases based on remote sensing. Here, HYBAM was considered as the reference dataset

Figure 5

Average power spectrum of HYBAM, CHIRPS, MSWEP and TRMM precipitation data for each AB sub-region. For the computation of the power spectrum, representative areas of each sub-region were considered: NAB (64.4°W-59.1°W, 0.1°S-4.1°N), NWAB (77.4°W-69.9°W, 2.4°S-1.6°N), NEAB (57.9°W-51.4°W, 4.1°S-1.1°N), CAB (73.6°W-62.3°W, 6.1°S-2.9°S), SWAB (71.6°W-65.4°W, 12.4°S-7.6°S) and SEAB (60.3°W-51.8°W, 14.4°S-9.4°S)

Figure 6

Wet period composing maps of 30-90 days filtered precipitation for HYBAM, CHIRPS, MSWEP and TRMM data for each OMI index phase

Figure 7

As in Fig. 6, but for the dry period

Figure 8

Diurnal cycle of MSWEP and TRMM 3-hourly data (in UTC) compared to the precipitation data collected in different campaigns. (a) represents the KGE of the diurnal cycle for the three stations analyzed. The circles represent the KGE of TRMM diurnal cycle and the squares are for MSWEP. For the TRMM-LBA data, the three consecutive circles (and squares) represents N2, N3 and N4 sites, respectively. (b) is the diurnal cycle comparison with the GoAmazon campaign radar data. (c), (d) and (e) represents the diurnal

cycle for N2, N3 and N4, respectively, for the TRMM-LBA campaign in the Brazilian state of Rondônia. (f) is the comparison for the Huayao station located in the Andes

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.docx](#)