Boletim de Ciências Geodésicas
ISSN: 1413-4853
ISSN: 1982-2170
Universidade Federal do Paraná

Nova, Raquel Arcoverde Vila; Gonçalves, Rodrigo Mikosz; Ferreira, Lígia Albuquerque de Alcântara; Lima, Fábio Vinícius Marley Santos

THE INFLUENCE OF THE REMOTELY SENSED RAINFALL PRODUCTS' SPATIAL RESOLUTION TO UNMASK EXTREME EVENTS IN NORTHEAST BRAZIL

Boletim de Ciências Geodésicas, vol. 27, no. 3, e2021023, 2021
Universidade Federal do Paraná

DOI: https://doi.org/10.1590/s1982-21702021000300023

Available in: https://www.redalyc.org/articulo.oa?id=393968997005

DOI 10.1590/s1982-21702021000300023

ORIGINAL ARTICLE

THE INFLUENCE OF THE REMOTELY SENSED RAINFALL PRODUCTS' SPATIAL RESOLUTION TO UNMASK EXTREME EVENTS IN NORTHEAST BRAZIL

A influência da resolução espacial dos produtos de chuva remotamente detectados para revelar eventos extremos no nordeste do Brasil

Raquel Arcoverde Vila Nova[1] - ORCID: 0000-0003-2055-0387

Rodrigo Mikosz Gonçalves[1] - ORCID: 0000-0002-5066-1910

Lígia Albuquerque de Alcântara Ferreira[1] - ORCID: 0000-0001-6117-2365

Fábio Vinícius Marley Santos Lima[1] - ORCID: 0000-0001-6799-5733

[1] Universidade Federal de Pernambuco, Programa de Pós-Graduação em Ciências Geodésicas e Tecnologias da Geoinformação, Recife- PE, Brasil.

E-mail: raquel.vilanova@ufpe.br; rodrigo.mikosz@ufpe.br; ligia.alcantara@ufpe.br; fabio.vinicius@ufpe.br

Received in 3rd May 2021

Accepted in 16th July 2021

Abstract:

This work presents the influence of the spatial resolution on precipitation samples to understand extreme events in the Agreste region of Pernambuco, northeast of Brazil. Among the materials used, the following sources of precipitation data (1998 to 2019) can be cited: The Tropical Rainfall Measuring Mission (TRMM), the Climatic Research Unit (CRU), and weather stations. In the process of validating the precipitation time series with the weather stations, the TRMM data showed a strong Pearson correlation (0.86 - 0.90) and the CRU data a moderate one (0.71 - 0.76). The relative bias (RB) and the standard deviation of observation ratio (RSR) were also calculated to identify the data's trend, which showed an overestimation for both sources. The extreme events were identified through the calculation of the Standardized Precipitation Index (SPI), where the TRMM with strong correlation (0.80 - 0.91) obtained a better performance than the CRU data. The TRMM data were selected to understand the extreme drought events in the study area, where the cities with altitudes above 500m obtained maximum values of probability of occurrence with 19%. Conversely, for extreme humidity events, the maximum was 14% for those with altitudes below 200m.

Keywords: Precipitation; Climate extremes; TRMM; CRU; SPI.

How to cite this article: VILA NOVA, R.A.; GONÇALVES, R.M.; FERREIRA, L.A.A.; LIMA, F.V.M.S. The influence of the remotely sensed rainfall products' spatial resolution to unmask extreme events in northeast Brazil. *Bulletin of Geodetic Sciences*. 27(3): e2021023, 2021.

1. Introduction

When using precipitation products, it is necessary to firstly consider the temporal resolution of the data, regarding which, naturally, the greater the historical series and the availability of information becomes the more attractive for behavior analysis, detection and projections of extreme event patterns over time (Mpelasoka et al., 2018). Secondly, it is also necessary to consider a crucial factor, which is the spatial resolution of such products. The lowest spatial resolution can sometimes be designated for global studies, taking into account the scope and influence of several water basins (Biemans et al., 2009; Schneider et al., 2017). On the other hand, for regional studies, the highest spatial resolution information is desired (Nastos et al., 2016; Tao et al., 2016).

Studies comparing the spatial resolution of multi-sensor images are very common, for instance, as Nouri et al. (2020) and Ozkan et al. (2017). However, the influence of the spatial quality precipitation products and their limitations are not usually questioned, as it is taken into account to initially supply the difficulty of availability of information. Often, due to the absence of alternative options for data acquisition with the necessary characteristics for the study, it is common to work with what is available, in terms of both temporal and spatial scope.

In practice, there are some ways to measure precipitation, for example: (i) using weather stations; (ii) through the use of remote sensors installed on artificial satellites, such as the Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Measurement (GPM) missions (Liu et al., 2012; Maggioni et al., 2016); or (iii) by means of global data that use precipitation temporal series to assess climatic extremes such as the GPCC (Global Precipitation Climatology Center) and CRU (Climatic Research Unit). The data obtained remotely, (ii) and (iii), provide subsidies for research in developing countries with limited investment when it comes to in situ stations, and places of difficult approach such as the Himalayas, the Amazon rainforest and the semiarid region of Brazil. (Almeida et al., 2015; Paredes-Trejo et al. (2018), Amini et al., 2019; Shukla et al., 2019). Amid the space missions developed with the intent of observing precipitation estimates, TRMM's Multi-Satellite Precipitation Analysis products have been widely used in order to understand climatic events in view of the great spatial and temporal potential (Chen et al., 2018; Fang et al., 2019). In the precipitation studies, once the temporal series of a given region is gathered, the next step is to assess the water crisis events.

A series of works have been carried out with the aim of evaluating the water crises that plague the world, through different techniques and sets of associated temporal samples. For instance, (a) Ndehedehe et al. (2020) that evaluated the evolutionary patterns of drought (1901 - 2014) in the Sahel of West Africa, comparing the Standardized Precipitation Evapotranspiration index (SPEI) with the SPI, as well as their teleconnections with low frequency climatic oscillations, identifying that the oscillations have a significant influence in the spatial/temporal evolution of these drought indicators; (b) In Brazil, Awange et al. (2016) used the monthly precipitation time series extracted from global data sets such as CRU and GPCC, in addition to the sea surface temperature data, in order to identify the probability of occurrence, intensity, duration and extent of droughts that happened during the period of 1901-2013; (c) Lima et al. (2020) combined in situ and remotely sensing data in order to disclose extreme drought events using SPI at different time scales for the Hydrological Region of the Northeast Atlantic East, where it was possible to identify the occurrence of the last severe drought in the time period from 2012 to 2017; (d) occurrences of water crisis have been afflicting several regions of Brazil, which face the aftereffects of a regime of low rainfall, such as the one that settled in the state of São Paulo and directly affected the population and the local economy during the years 2014 and 2015 (Nobre et al., 2016; Soriano et al., 2016).

In order to make up for the lack of the influence of different spatial resolutions of precipitation data while identifying extreme events, this article presents a case of study at a regional scale, considering the Agreste region, in state of Pernambuco, northeast of Brazil, with the following objectives: (i) to use the multi-satellite precipitation estimation time series TRMM and the CRU set for the time period (1998-2019) applying the SPI in both sources, aiming to detect extreme events; (ii) to validate both time series with in-situ samples, thus verifying

Pearson's correlation, the relative bias (RB), and the ratio between the average standard error and the standard deviation of the observations (RSR) between the sample sets. These last two have the goal of identifying the trend of the estimated data in relation to those observed; (iii) to compare and analyze the products classified by SPI taking into consideration the different spatial resolutions and present their differences, similarities and limitations and (iv) to detail the extreme scenario found for the study area, according to the sample set with the spatial resolution that came closest to the regional situation, and thus determining its probability of occurrence throughout the assessed time.

2. Regional Setting

According to the IBGE (Brazilian Institute of Geography and Statistics) (2017), the state of Pernambuco, located in the northeast of Brazil, comprises five Geographical Mesoregions defined through three dimensions that qualified its regional identity, which are (i) the social process; (ii) the natural framework and (iii) the communication network and places. The state's Mesoregions were classified as: Metropolitana of Recife, Mata Pernambucana, the Agreste of Pernambuco, Sertão Pernambucano and São Francisco Pernambucano. The study area is the Mesoregion of the Agreste of Pernambuco, located between the Mata of Pernambuco and the Sertão, with approximately 24,400 km², distributed in 71 cities, whose estimated population is around 2,412,364 inhabitants (IBGE, 2019). The climate is considered to be semi-arid, with an average precipitation regime that does not exceed 800mm/year (APAC, 2015). The Agreste region of the state of Pernambuco is marked by a hilly terrain which leads to an irregular distribution of precipitation. Figure 1 (a) shows the location of the state of Pernambuco in relation to Brazil and the countries of South America. Figure 1 (b) presents a map containing the division of the 5 mesoregions of the state of Pernambuco where the highlighted mesoregion is the study area. The in situ weather stations inserted in the region of the Agreste of the state of Pernambuco are indicated in Figure 1 (c), where the gray circles indicate the set of weather stations located in the study area. Even though this figure shows a high density of stations, most of them do not present a series of continuous data during the studied period proposed for this work (from 1998 to 2019), a fact that confirms the importance of remote sources of precipitation data for climate studies. Still on Figure 1 (c), the yellow triangles indicate the location of the stations that were used in the SPI step (item 3.2.2), which have a data set encompassing the years from 1998 to 2019, and also displayed gaps that can be corrected.

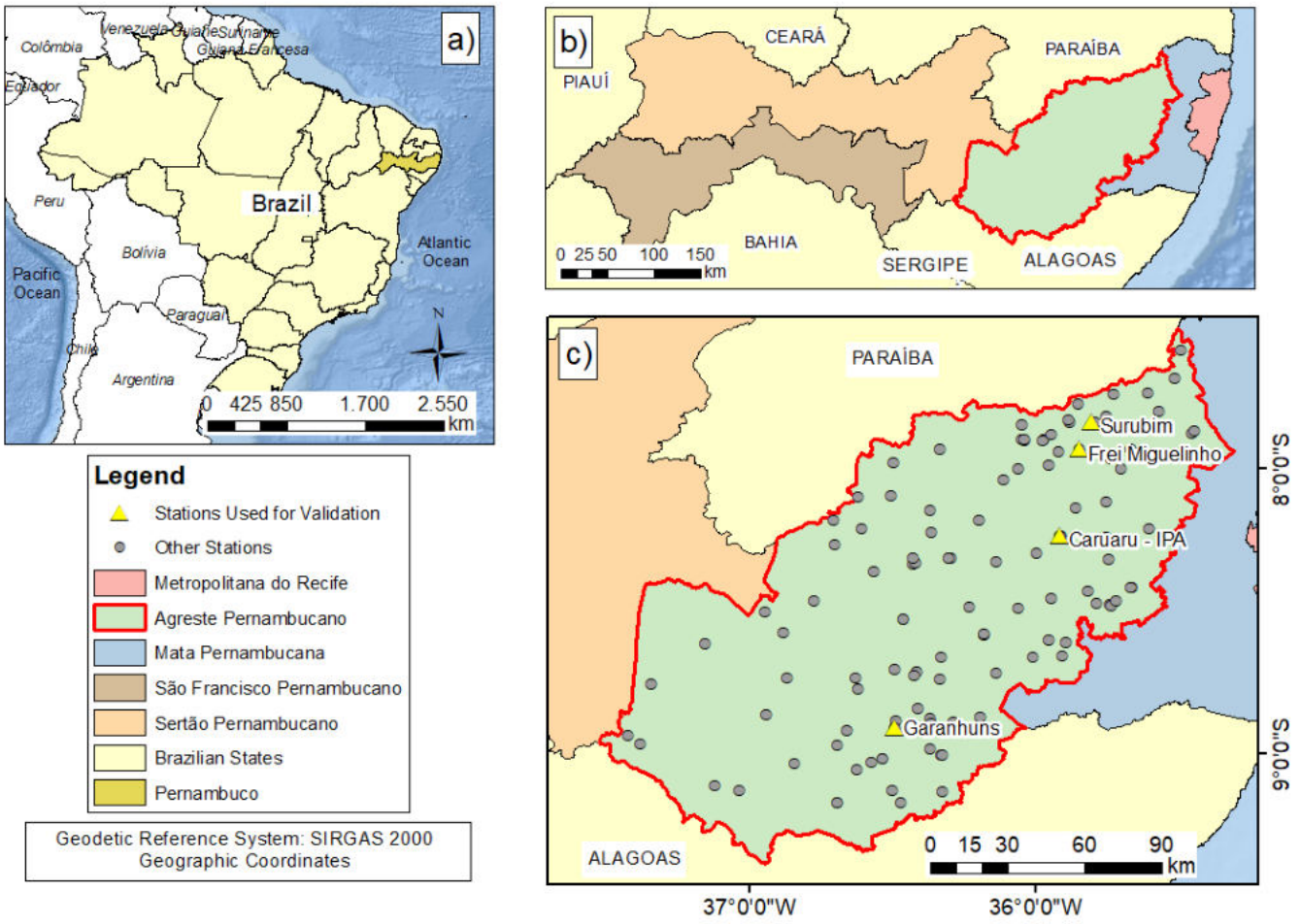

Figure 1: (a) Pernambuco state highlighted in Brazil, (b) Pernambuco state mesoregions division (c) meteorological stations distribution at Agreste Pernambucano.

3. Data and methods

3.1 Data

(i) TRMM 3B43

The artificial satellite precipitation data were gathered through TRMM, a joint mission of NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency), for the period from 1998 to 2019. The TRMM mission ended in 2015, but while it was working, it was possible to develop precipitation estimation algorithms based on observations from sensors on board the TRMM satellite, such as TMPA. Even with the end of the mission, the TMPA products continued to be released based on input data from other satellites in the constellation (Chen et al., 2018; Yong et al., 2015). For the development of this work, the TRMM 3B43 was used, which refers to the combined monthly precipitation estimates, with spatial resolution of 0.25 ° x 0.25 ° in a geographical coverage of 50 ° S to 50 ° N. The data were obtained in *NetCDF* format on the website of the Data Center for Information and Earth Sciences *Goddard* (GES DISC) <https://disc.gsfc.nasa.gov/>, converted through the free *Qgis* software to *raster* data containing the observations of precipitation estimates for the study area.

(ii) CRU TS v. 4.04

The Time-Series (TS) is a set of climatic data available in a grid with a spatial resolution of 0.5°, developed by the Climatic Research Unit (CRU) of the University of East Anglia, and supported by funding agencies. The database covers the time period from 1901 to 2019, with available data for all land areas, excluding Antarctica. The data are based on a time series with observations from more than 4000 stations (Belda et al., 2014) with a methodology of interpolation and data combination thoroughly studied by Mitchell and Jones (2005). There are six climatic variables available in this version and among them there are the rainfall data and the average temperature data. For the development of this research, only rainfall data from the time period of 1998 to 2019 were used. The database is available in *.kml* format at the link <https://crudata.uea.ac.uk/cru/data/hrg/#formats>, which was selected according to the area of study.

(iii) Weather stations

The in situ time series were used to validate the precipitation estimate data obtained in 3.1.1 and 3.1.2. The weather stations were selected according to the availability criteria of at least 25 months' worth of data within the studied period (from 1998 to 2019). These weather stations were used in the validation step of the precipitation data obtained remotely, in order to use the largest number of data for consistent validation. The in situ precipitation information was obtained through the Agência Pernambucana de Águas e Clima (APAC - Pernambuco Water and Climate Agency), where 168 weather stations are identified in the study area, in which many of them have long periods with gaps in their database. There are several causes for the lack of data which may be related to the measuring equipment, to the operator and even to a lack of maintenance investment in these stations. Thinking about using the largest number of data for a consistent validation, it was decided to use only the stations that had, at least, 10 percent of the data in the studied period; using this criterion, we reduced the number of stations to 115.

In the SPI step, data from consistent in situ stations were used for the results validation process. Among the stations made available by APAC, only the stations of Caruaru - IPA and Frei Miguelinho presented a concise historical series with few failures during the entire period from 1998 to 2019. Another source of precipitation data was the Instituto Nacional de Meteorologia (INMET - National Meteorological Institute) where, between the 3 conventional stations present in the study area, the Garanhuns and Surubim stations were selected, on account of having a complete historical series. The information failures or lack of information regarding the precipitation data in the time series, due to some problem in their gathering, were interpolated using the simple linear regression model, with R^2 average of 53.1%.

3.2 Methods

Figure 2 shows the methods used. The first frame brings the input data followed by: (i) validation process; (ii) the assessment of the Standardized Precipitation Index (SPI); (iii) influence of spatial resolution and finally (iv) extreme events analysis.

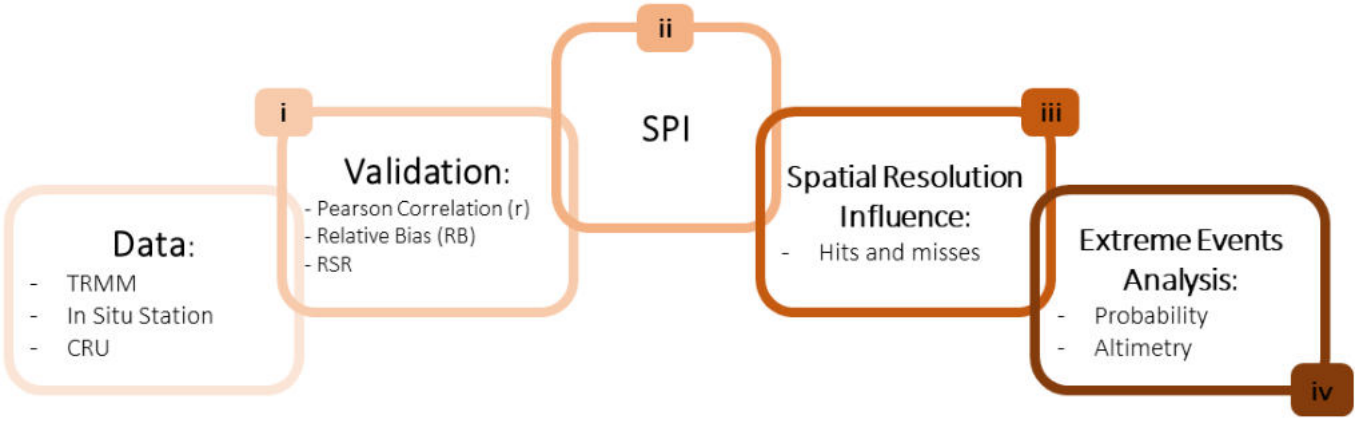

Figure 2: Methods: (i) Validation, (ii) SPI, (iii) Spatial Resolution Influence and (iv) Extreme Events Analysis.

(i) Validation

The validation procedure is used to evaluate the performance of TRMM and CRU products for the study area. The comparisons were made from the homologous geographical comparison of the in situ temporal series, with the respective TRMM and CRU temporal series. The statistics used for this comparison are: a) Pearson's correlation coefficient (r) with a 95% confidence interval, which aims to quantify the type of correlation and dependence among the time series according to equation 1:

$$r = \frac{\sum_{i=1}^{n}\left((O_i - \bar{O}) \times (P_i - \bar{P})\right)}{\sqrt{\left(\sum_{i=1}^{n}(O_i - \bar{O})^2\right) \times \left(\sum_{i=1}^{n}(P_i - \bar{P})^2\right)}} \quad (1)$$

where: P_i is equal to the estimated precipitation (mm) in the time interval i, and $\bar{P}$ is the average of the estimated precipitation; O_i is the observed precipitation (mm) of the weather station in question in time interval i and $\bar{O}$ is the average of the observed precipitation, n is the total number of data analyzed.

b) Relative Bias (RB) presented in Equation 2 that identifies the average tendency of remote data to be overestimated (positive RB values), or underestimated (negative RB values), in relation to the values observed in situ, where the ideal value for bias is 0 (Chen and Li, 2016; Chen et al., 2018; Erazo et al., 2018; Xu et al., 2017).

$$RB = \frac{\sum_{i=1}^{n}(P_i - O_i)}{\sum_{i=1}^{n} O_i} \quad (2)$$

c) the RSR is the ratio between the average standard error and the standard deviation of observations. Equation 3 shows that the RSR is calculated as it is the ratio of the mean squared error of the estimated precipitation data and those observed in situ, on the standard deviation of the data observed in situ. The RSR values vary from 0, which indicates that the estimated model is perfect; up to positive values indicating that the higher the RSR value, the worse the performance of the estimated measures is (Moriasi et al., 2007).

$$RSR = \frac{\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}} \quad (3)$$

(ii) Standardized Precipitation Index (SPI)

The SPI is calculated for all spatial resolutions used (CRU, TRMM and weather stations). The calculation basis for the SPI consists of the adjustment of a gamma probability density function and a temporal series of precipitation data, using the methodology endorsed by Zarch et al. (2015). The index was processed using the algorithms provided by the scientists Vicente-Serrano et al. (2010) in the free software environment R. The classification thresholds

adopted in this work is according to Mckee, Doesken and Kleist (1993) considering +2.00 and above = Extreme wet; +1.50 to +1.99 = Very wet; +1.00 to +1.49 = Moderately wet; -0.99 to +0.99 = Near normal; -1.00 to -1.49 = Moderate drought; -1.50 to -1.99 = Severe drought and -2.00 or less = Extreme drought. The scale used was 12 months taking into account the intention to find the climatic extremes considered long-term (Ali et al., 2019).

(iii) Spatial resolution influence

Hits and misses were calculated from the algebra of map (Sampaio, 2012) in order to identify the percentage of success obtained from the SPI data set for different resolutions. The reference mask used the TRMM grid to obtain the CRU values. In this way, the data from different fonts kept the respective spatial resolution.

(iv) Extreme events analysis

In the fifth step, an extreme event analysis is presented, in which these events are identified and considered as moisture events for SPI ≥ 1.00 and drought events for SPI ≤ -1.00. According to Magalhães (2006), the classic definition of probability refers to equiprobable unit subsets. In the case of the present work there are only three possibilities of events which are: extremes of drought, extremes of humidity, or normal. The probability *(P_A)* of occurrence of these events is calculated based on equation 4 described below, where A represents the number of events and T the total number of events, either of drought or humidity.

$$P_A = \frac{A}{T} \times 100 \quad (4)$$

The probability measure was based on the 21 years studied and was obtained per pixel, having been calculated with the data source that presented the best statistical performance when compared with the measurements in situ. To improve the analyzes, terrain information was added from the Shuttle Radar Topography Mission (SRTM) obtained through the website of the Brazilian Agricultural Research Corporation - Embrapa, with spatial resolution of 30m.

4. Results and discussions

4.1 Assessment and analysis of the rainfall products

Figure 3 shows the spatial distribution of the metrics (r, RB and RSR) used in the validation of the precipitation data using the remotely sensed rainfall products and in-situ rainfall data, which are taken as ground truth. The comparison of these data was carried out in relation to the common time periods. The statistical results were obtained punctually for each station used, and then the data were interpolated and extrapolated by QGIS using the Inverse Distance Weighted (IDW) method.

Figures 3a and b show Pearson's correlation coefficients (r) for CRU and TRMM respectively. The CRU data (Figure 3a) indicates that almost all the area of study is classified as moderately positive (0.5 ≤ r <0.8), and small areas are classified with a weak positive correlation (0.1 ≤ r <0.5). On the other hand, Figure 3b indicates that the data further to the north and south of the Agreste region of the state of Pernambuco, show a strong positive correlation (0.8 ≤ r <1.0) and that the rest of the area is classified as moderate positive. For this statistical parameter, the TRMM data have a higher quality when compared to the result obtained for the CRU data.

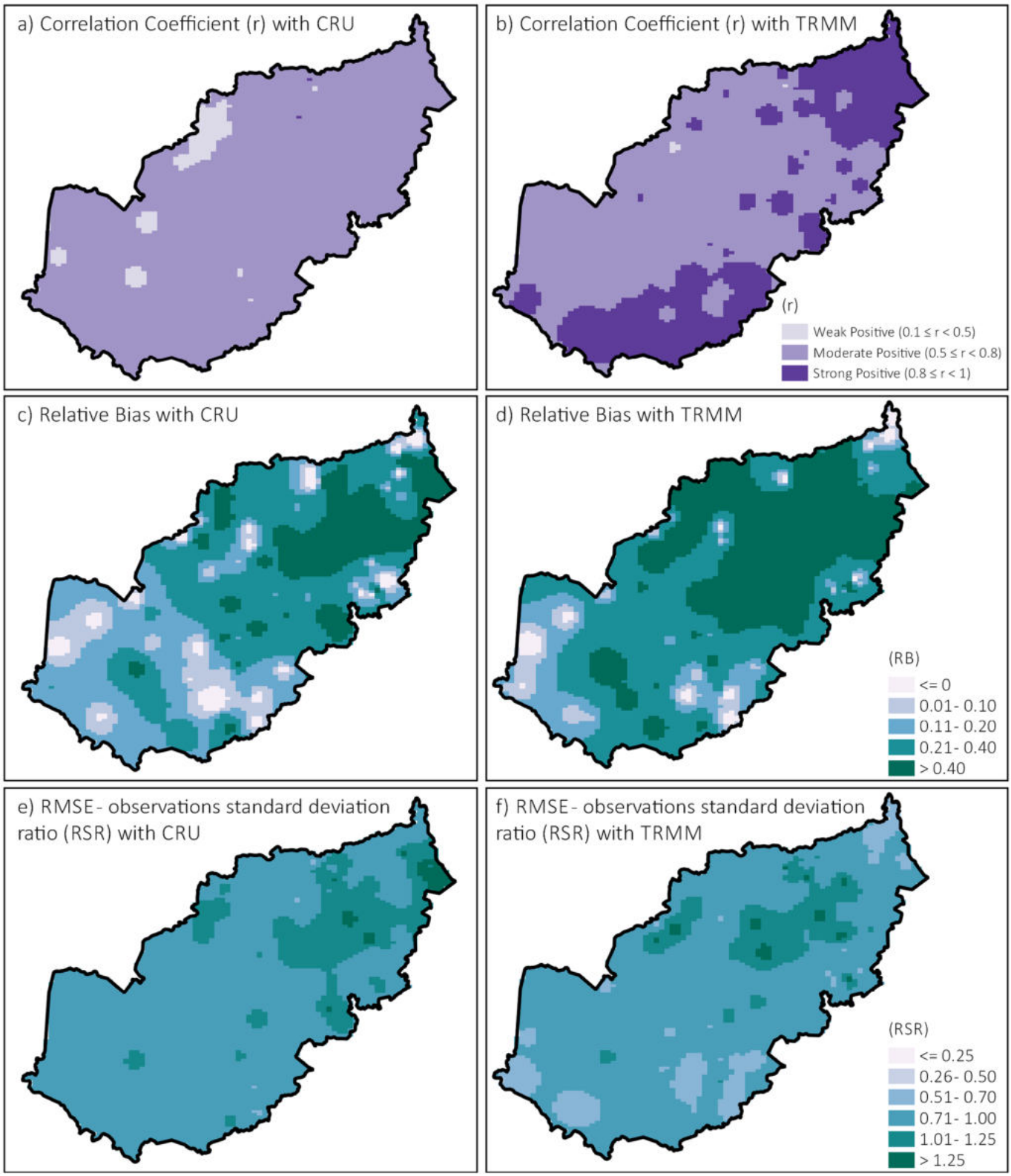

Figure 3: Validation. a), c) and e) are the results of r, RB and RSR respectively for the CRU data; b), d) and f) are the results of r, RB and RSR respectively for the TRMM data.

Figures 3a and b show Pearson's correlation coefficients (*r*) for CRU and TRMM respectively. The CRU data (Figure 3a) indicates that almost all the area of study is classified as moderately positive ($0.5 \leq r < 0.8$), and small areas are classified with a weak positive correlation ($0.1 \leq r < 0.5$). On the other hand, Figure 3b indicates that the data further to the north and south of the Agreste region of the state of Pernambuco, show a strong positive

correlation (0.8 ≤ *r* <1.0) and that the rest of the area is classified as moderate positive. For this statistical parameter, the TRMM data have a higher quality when compared to the result obtained for the CRU data.

The results obtained for the relative bias (RB) are shown in Figures 3c and 3d, where the ideal value of the parameter is zero. The positive values indicate that the precipitation products are overestimated when analyzed in relation to the precipitation measurements of the stations. The southern region of the area of the study was identified with lower values of RB, indicating that this region showed less overestimation of the data, when observing the northern areas of the Agreste region of the state of Pernambuco. Both sources collected high positive results (> 0.40), however the TRMM (Figure 3d) showed a larger area in this range. Xu et al. (2017) also found RB values greater than 0.40 for TRMM data, when the influence of elevation on the precipitation estimation products from TRMM and GPM was studied in the southern Tiberian plateau in China.

The RSR values obtained for both sources, based on the precipitation measurements of the stations, identified in Figure 3e for the CRU data and Figure 3f for the TRMM data correspond to the performance of the estimated data, where the perfect model indicative value is equal to zero. For the Agreste region of the state of Pernambuco, this parameter performed well for both data sources, showing a slight variation in the result obtained for the TRMM, which contains areas with values closer to the ideal value of the statistic (RSR = 0). According to Moriasi et al. (2007) the RSR values ≤ 0.60 can be considered good, indicating that the data present a good performance in relation to the data of the stations, which can be identified in the data from the Agreste region of Pernambuco, for both studied sources, where the TRMM data obtained better performance when compared to CRU data.

In order to bring specific statistical values, Table 1 presents the validation results obtained based on the weather stations that were used, to later validate the results obtained by the SPI.

Table 1: CRU and TRMM validation.

Meteorological stations	CRU		TRMM	
	r	RB	r	RB
Surubim	0.76	0.32	0.90	0.44
Frei Miguelinho	0.71	0.35	0.87	0.47
Caruaru (IPA)	0.72	0.45	0.86	0.43
Garanhuns	0.73	0.01	0.87	0.02

The TRMM data have a strong positive linear correlation with the data from the weather stations, ranging from *r* 0.86 to 0.90. On the other hand, in the validation of the data obtained from the CRU, a moderate positive correlation was found, with *r* values ranging from 0.71 to 0.76. Regarding the relative bias metric (RB) the CRU data performed slightly better than the TRMM data. However, Garanhuns was the station that obtained an excellent performance, close to the ideal value of RB = 0 for both data sets. The precipitation data from TRMM and CRU were overestimated for the other stations.

In general, from the statistics presented in Figure 3 and Table 1, it is possible to observe that the precipitation estimation data from the TRMM were superior to those obtained through the CRU considering the study area. Erazo et al. (2018) in a similar validation process, applied to a region in Ecuador, obtained a value of *r* = 0.82 and 0.66 for TRMM and CRU respectively. This comparison confirms the influence of spatial resolution on the precipitation temporal series, that is, the data obtained by the TRMM tend to have a stronger correlation when compared to the CRU data. Regarding the RB, small values were found for the Ecuadorian region, -0.03 for the TRMM and -0.02 for the CRU, with a slight variation among the data sets.

4.2 SPI with different sources

Figure 4 presents the results containing the behavior of the SPI according to the three sources of precipitation data: TRMM, CRU and the weather stations. The temporal series was classified with SPI values ranging from Extreme Wet (2.00) to Extreme Drought (-2.00), thus showing the moments of occurrence of extreme weather events when exceeding thresholds above ± 1.00.

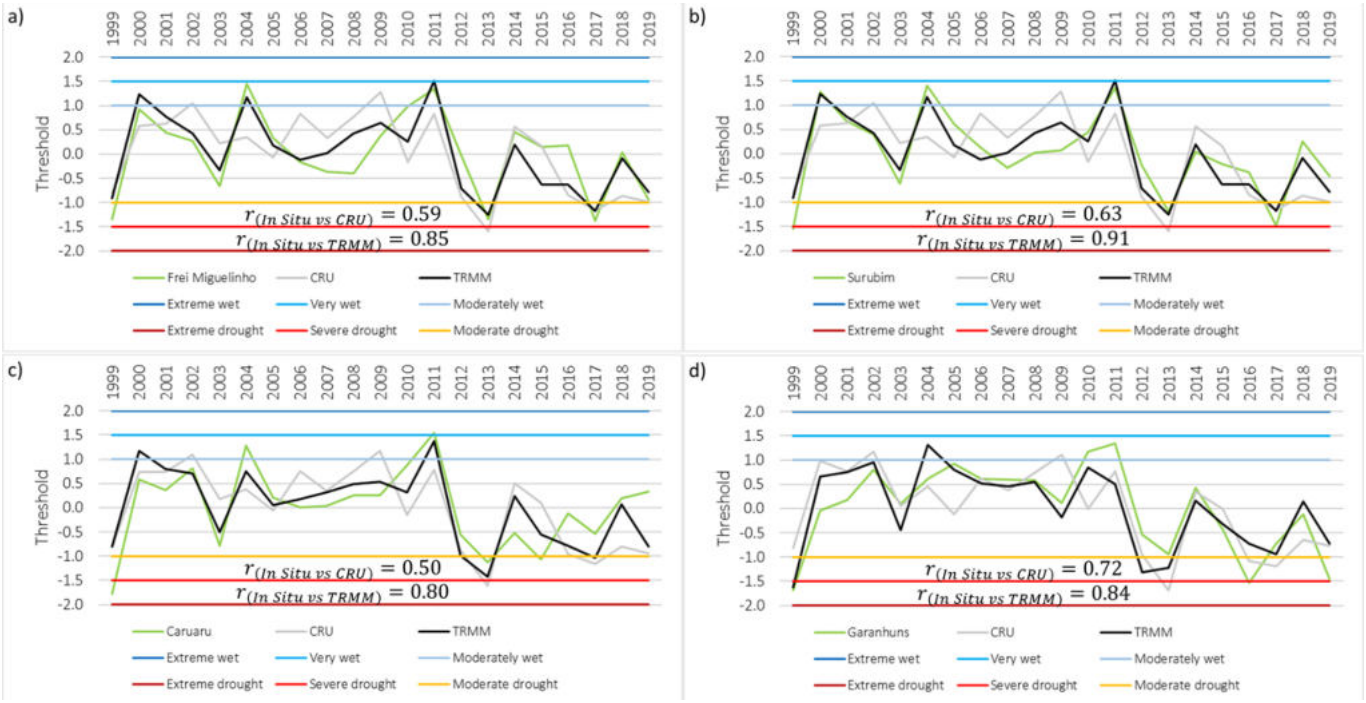

Figure 4: SPI behavior regarding different data sources. (a) Frei Miguelinho weather station, (b) Surubim weather station, (c) Caruaru weather station and (d) Garanhuns weather station.

Figures 4a, 4b, 4c and 4d show the results generated for the locations of the stations of Frei Miguelinho, Surubim, Caruaru, and Garanhuns respectively. Pearson's correlation coefficient *r* was generated based on the SPI results for the weather stations data, in relation to the SPI results for the TRMM and CRU data in the different locations.

In Figure 4a, the divergence between the CRU data and the other two data sources is noticeable, what happened in some years, such as the cases of 2002, 2004 and 2009, when the SPI value presented a large divergence in the index in relation to the TRMM. The SPI value for the year 2002 was 1.04 for the CRU, 0.42 for the TRMM and 0.26 for the Frei Miguelinho station. In the year 2004 the values obtained were 0.34 (CRU), 1.16 (TRMM) and 1.44 for the station, while in 2009 they were 1.28 (CRU), 0.64 (TRMM) and 0.39 for the station. All these years cited showed SPI values with CRU that were far from the reality found when using TRMM or stations.

As for the years 2000 to 2011, the series had an almost normal behavior in relation to the index, presenting an increase from the year 2010 to 2011. The period from 2011 to 2013 presented an abrupt decrease in the SPI, passing from the humidity classification to moderate to severe drought, corroborating to the results shown in the Lima et al. (2020) study.

Through the correlation values *r*, it was possible to identify the strong correlation of the SPI that exists between TRMM vs stations, presenting values from 0.80 (Fig. 4c) to 0.91 (Fig. 4b). Brasil Neto et al. (2021) also showed the extreme rainfall events in Paraíba State obtained a good TRMM accuracy compared to rain gauge-measurements. The

results obtained for the CRU set show a moderate correlation, ranging from 0.50 (Fig. 4c) to 0.72 (Fig. 4d). Suliman et al. (2020) also analyzed the performance of the SPI from the TRMM in Iraq, where the values of *r* in the range of 0.89 - 0.91 were found, thus corroborating to the fact that the precipitation time set obtained by the TRMM, also showed a better adherence to the data in situ. On the other hand, Rodrigues et al. (2020) highlighted the influence about the quality estimative considering three elements: seasonality, geographic location and time scale influence.

The period from 2012 to 2017 was considered the most severe with the longest period of drought in the region, which was also identified by Brito et al. (2018), when they evaluated the frequency, severity and duration of drought events in the Northeastern Semi-arid region of Brazil, using only in situ data, and by Marengo et al. (2017), who brought a general and historical overview of the drought in the Northeast of Brazil. These same years observed with the presence of extreme drought events showed by Santos et al. (2019) using data sets of at least 30 years of precipitation in the Northeast of Brazil. The positive SPI values were also observed by Santos et. al. (2019), where the years 2004 and 2009 stood out as moderate humidity. Such works reaffirm the results obtained regarding SPI in this study.

4.3 CRU spatial variability

Figure 5a shows the spatial variation of the SPI for the data obtained from the CRU, with the original spatial resolution (0.5 ° x 0.5 °). The processing was done by using the same time frame as the TRMM data (1998 to 2019), in order to rule out the inconsistency regarding the period studied. It is possible to identify from Figure 5a, the climatic extremes of drought for the years 2012 (moderate drought), 2013 (severe drought), 2016 (moderate drought) and 2017 (severe drought). The year 2019 presented a small area classified as drought, highlighting the year of 2013 as the most critical year in the historical series studied, identifying the region of study as severe drought in almost its entirety. Regarding the extremes of rainfall for the years 2002 and 2009, the moderate humidity indexes were highlighted for almost the whole area. On the other hand, for the years 2000 and 2011, a small portion with moderate humidity indexes was found.

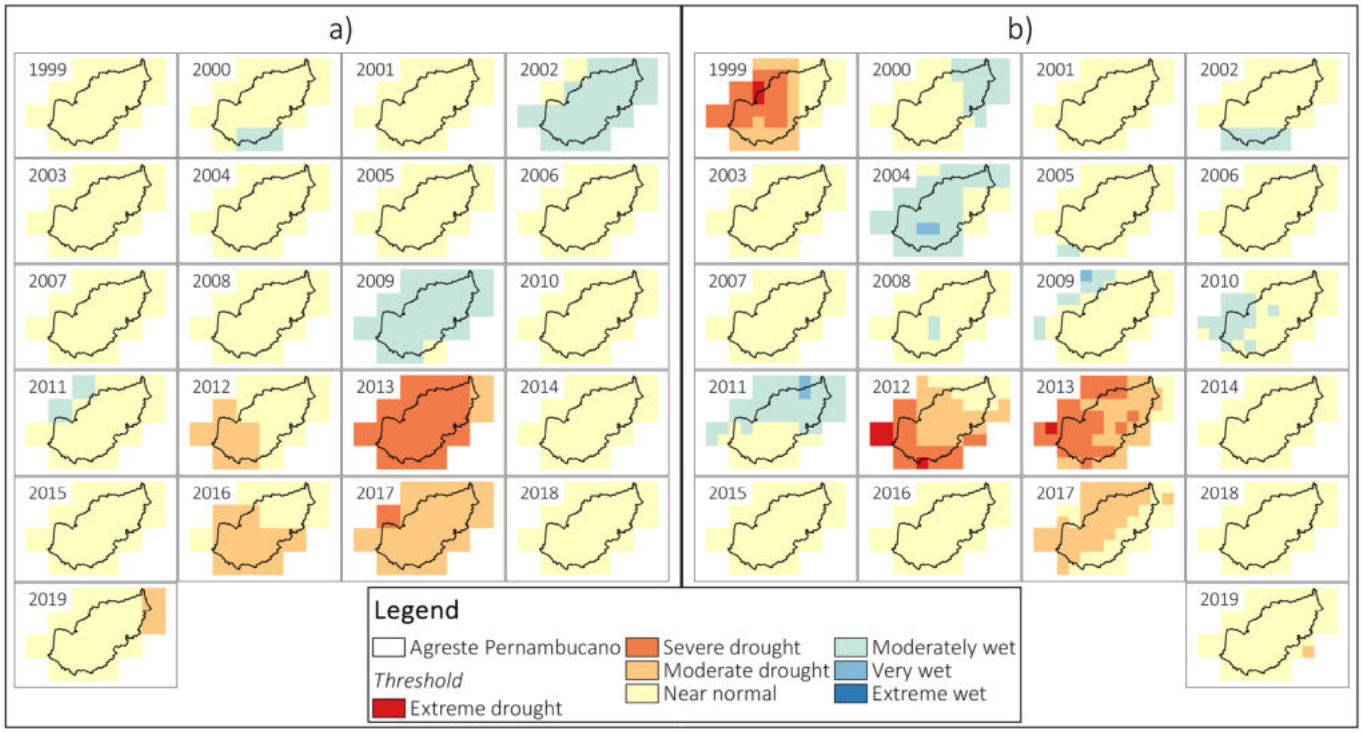

Figure 5: (a) 12-month time scale SPI generated from CRU and (b) Standardized Precipitation Index (SPIs) on the 12-month time scale generated from the TRMM data.

Figure 5b shows the spatial variation of the SPI for the TRMM precipitation estimation data (0.25 ° x 0.25 °). The results show a greater spatial, temporal and class variety identified as extreme humidity events when compared to Figure 4. Among them, the years 2000, 2002 and 2010 stand out, containing areas classified as moderately wet, and the years 2004, 2009 and 2011 presenting areas classified as moderately wet and very wet.

Conversely, the drought events that could be pinpointed in the years 1999, 2012, 2013 and 2017, with areas classified as extreme drought for the first three years cited (2000, 2002 and 2010). From the year 2011 to 2012, it is possible to identify a downward trend in terms of SPI (Figure 4). The same trend can be specially seen in the results obtained for the TRMM data in Figure 5b, where no more years are identified with extreme humidity events from the year 2012.

4.4 TRMM spatial variability

With the results gathered for the SPI with the TRMM data, the sensitivity of the data source in identifying extreme events is notorious, not only for humidity but also for drought, in which areas were able to be classified within the maximum limits of the adopted classification, with the exception of the extreme wet class, taking into account that it was historically expected that the study region would not have high levels of humidity.

Amini et al. (2019) also used TRMM precipitation data to assess the drought and its effects on water resource instability, in an Iranian province, and identified that remote data were highly accurate when applied to the region. Chen et al. (2018) also used the SPI to monitor and assess drought periods using TRMM data with weather stations. Through a comparative analysis they identified that the two data sources were highly correlated, which shows that the TRMM data can be used as a complement to the specific observations, made available by the weather stations, in view of their quality in terms of resolution and correlation with the data in situ.

4.5 Spatial resolution comparison

Figure 6a shows areas of hits and misses in relation to the classification of the SPI for both datasets, TRMM and CRU, in the same spatial resolution (0.25° x 0.25°). They represent the pixels that are in the same class. The values presented for each year identify the percentage of matches. The years with the lowest percentages were 1999 (29.41%), 2002 (29.41%), 2004 (20.59%), 2009 (17.65%), 2012 (27.94%) and 2016 (47.06%), all of them with values below 50% of accuracy.

Considering the results for 1999, the SPI through the TRMM identified drought events at the three levels (moderate, severe and extreme), whereas the SPI obtained from the CRU data classified the entire area of study as almost normal, not identifying extremes for this year. As the CRU data are interpolated using a set of global weather stations, the precipitation data may have been affected by neighboring regions (Belda et al., 2014; Naumann et al., 2012), so the percentage of accuracies for this year is low as is also the case for 2002, 2009 and 2016. The years that show drought events for TRMM data are identified with intensity and distinct areas from the events shown in Figure 5a for the CRU data.

For the year 2012, both sources identified extreme drought events, but the percentage of accuracies was less than 50%, due to variation in the intensity of the drought and also because of the resolution of the data, taking into consideration that the precipitation measures are obtained for each pixel in question and do not suffer interference from neighboring pixels. This may also be the cause of the low values of accuracies in the years previously mentioned.

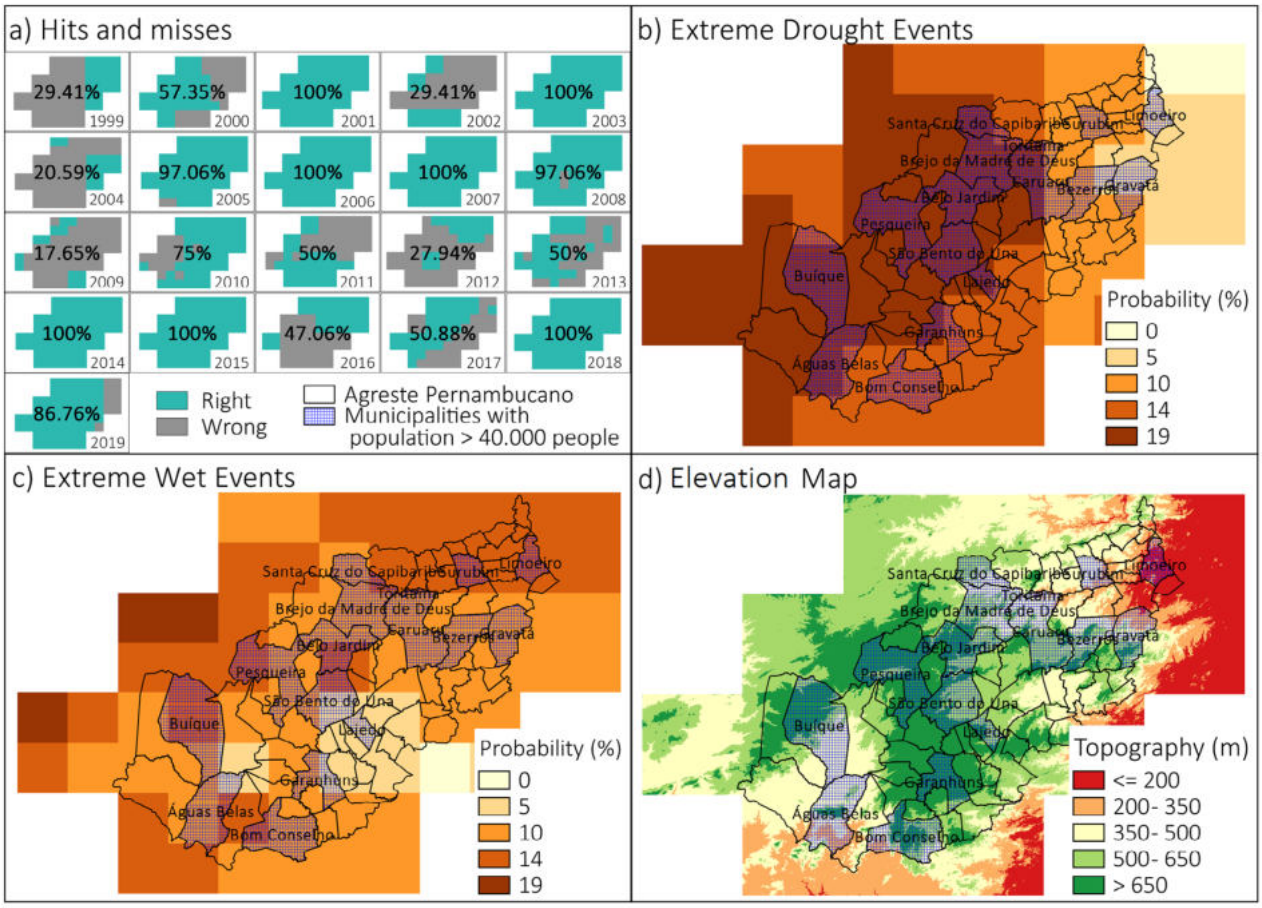

Figure 6: (a) Hits and misses applied to SPI results, (b) Extreme Drought Events Probability, (c) Extreme Wet Events Probability, (d) Togography.

A justification for the disparity in the number of accuracies could be seen through the difference in obtaining the precipitation estimate data, taking into account that the data gathered through satellites take into consideration the amount of rainfall that occurred for the pixel area in question. As for the CRU data, for example, they are defined by an interpolation with data from weather stations in the area (Naumann et al., 2012). Therefore, the resolution variation also influences in the quality of the results, considering that a smaller area of pixel receives less influence when the average for a given pixel is generated.

4.6 Extreme events impacts

This section presents the results and analyses regarding the probability of occurrence of extreme weather events, using the SPI results obtained through the precipitation estimate from the TRMM, which presented the best statistical performance in the previously described step, thus providing subsidies to identify the more vulnerable areas to extreme events which are located in the Agreste region of the state of Pernambuco.

The cities with more than 40,000 inhabitants (IBGE, 2019), are shown in Figures 6 (b, c and d). Figure 6b shows the probability of occurrence of extreme events regarding the drought. Out of the 16 most populous cities in the region, just Limoeiro and Gravatá, which are further to the northeast of the Agreste region, and consequently closer to the mesoregion of Zona da Mata of Pernambuco, showed a 5% probability of drought occurring throughout the studied period. The cities that showed 19% of probability, can be considered quite vulnerable to the drought

situation and need mitigating actions concerning these events. Among them the following can be listed: Santa Cruz do Capibaribe, Brejo da Madre de Deus, Belo Jardim, Caruaru, Pesqueira, São Bento do Uma, Lajedo, Garanhuns, Águas Belas and Buíque, being the latter two ones very close to the mesoregion of the Sertão of Pernambuco.

Extreme events in terms of humidity were less likely to occur in the Agreste region (Figure 6c). For these events, the cities that obtained the highest probability values, (14%) of occurrence, were Limoeiro and Surubim, both located in the northeast of the region. No area of the Agreste region of Pernambuco showed a probability of 19% for extreme humidity events. The city of Águas Belas, for instance, was classified with areas of 5%, 10% and 14% of probability of humidity events (Figure 6c). In contrast, Figure 6b presents almost its entire classification with a 19% of probability for occurrence of drought extremes.

The elevation map for the study area is shown in Figure 6d, where the areas below 200m of altitude, marked in red, are identified further east of the region where only the city of Limoeiro, out of the cities with more than 40,000 inhabitants, is identified with such altitudes. Also, it presented low probability of drought events (5%), and high probability for humidity events (14%). The city of Lajedo, which is located in area with altitudes above 500m (light green), presented a high probability for drought events (19%) and low for humidity events (5%).

As for the region of the Agreste of Pernambuco, it is possible to observe by the results gathered in this study, that the cities which are closer to the region of the Mata of Pernambuco (Figure 1c) and, consequently, having lower altitudes, presented less probabilities regarding drought extremes, and greater probabilities for events of extreme humidity. As for cities with higher altitudes and more to the north of the region were classified with a high probability of extreme drought events.

5. Conclusion

In this study, monthly precipitation data obtained from three different sources were assessed to verify the influence of different spatial resolutions in the identification of extreme weather precipitation events. One of the limitations in studying climatic extremes in terms of precipitation is the scarcity of data in situ. It is the case the Agreste region in Pernambuco state, where only 4 weather stations had a data set without considerable gaps. An alternative to this problem is the use of other data sources with a robust database such as the TRMM and the CRU database. The main results found show:

I. Through the statistical analysis of precipitation data, based on the calculation of Pearson's correlation coefficient *r*, a better performance was seen when using precipitation data from the TRMM. These showed positive strong linear correlation coefficients. As for the CRU data, it showed a moderately positive performance, indicating that there are quantitative differences in both data sets, and this may influence analyzes at regional scales. The metrics used in order to identify the trend of the precipitation data, RB and RSR, pointed out that there was an overestimation for both sets of data. Although the relative bias had a better response for the CRU data, the RSR classified the TRMM data with lower values and closer to a perfect model;

II. Through the correlation of the SPI applied to the data of the weather stations (Figure 4), it was possible to identify that the region obtained a strong correlation with the SPI generated by the TRMM, showing that the climatic extremes identified demonstrate an excellent performance for this end. However, in the case of the SPI correlation for the CRU data, it was moderate.

III. The years of greatest divergence between the two remote data sources, regarding the SPI, were the years of 1999, 2002, 2004, 2009, 2012 and 2016; all of them showed values below 50% of accuracy. The variation in the accuracy rate could be due to the difference in spatial resolution and data collection process;

IV. The results obtained through the precipitation estimation measures from the TRMM demonstrate that the sensitivity regarding the identification of extreme events can be attributed to the resolution of the data, in which the variation and gathering of the data occur in a more scattered way per pixel. Therefore, for this case study, the TRMM data are more accurate for the identification of extreme weather events when compared to the CRU data;

V. Despite the results obtained, It is highlighted the usage of a new higher spatial resolution dataset, such as the GPM, considered a target for future research.

ACKNOWLEDGEMENT

The first and third authors thank the research grants supported by the project 23076.033077 / 2018-58 / FADE / UFPE / City Hall of Caruaru, the second author thanks the support for the project 310452 / 2018-0 / PQ / CNPq and the last author thanks master's scholarship supported by CAPES.

AUTHOR´S CONTRIBUTION

Definition of research problem (R. Mikosz and L. Alcântara), Literature review (R. Vila Nova and F. Lima), Definition of methodological procedures (R. Mikosz and L. Alcântara), Data collection (R. Vila Nova and F. Lima), Data processing (R. Vila Nova), Analysis and interpretation of data (R. Vila Nova, R. Mikosz, L. Alcântara and F. Lima), Manuscript writing (R. Vila Nova, R. Mikosz, L. Alcântara and F. Lima).

REFERENCES

Ali, M.; Deo, R. C.; Maraseni, T.; Downs, N. J. 2019. Improving SPI-derived drought forecasts incorporating synoptic-scale climate indices in multi-phase multivariate empirical mode decomposition model hybridized with simulated annealing and kernel ridge regression algorithms. *Journal of Hydrology*, v. 576, n. June, p. 164–184.

ALMEIDA, C. T. de; Delgado, R. C.; Junior, J. F. de O.; Gois, G.; Cavalcanti, A. S. 2015. Avaliação das estimativas de precipitação do produto 3B43-TRMM do estado do Amazonas. *Floresta e Ambiente*, v. 22, n. 3, p. 279–286.

Amini, A.; Kolahchi, A. A.; Al-ansari, N.; Moghadam, M. K.; Mohammad, T. 2019. Application of TRMM Precipitation Data to Evaluate Drought and Its Effects on Water Resources Instability. *Applied Sciences*, v. 9, n. 24, p. 5377.

Agência Pernambucana de Águas e Clima. 2015. Report: Organograma. Pernambuco: APAC.

Awange, J. L.; Mpelasoka, F.; Goncalves, R. M. 2016. When every drop counts: Analysis of Droughts in Brazil for the 1901-2013 period. *Science of the Total Environment*, v. 566–567, p. 1472–1488.

Belda, M.; Holtanová, E.; Halenka, T.; Kalvová, J. 2014. Climate classification revisited: From Köppen to Trewartha. *Climate Research*, v. 59, n. 1, p. 1–13.

Biemans, H.; Hutjes, R. W. A.; Kabat, P.; Strengers, B. J.; Gerten, D.; Rost, S. 2009. Effects of precipitation uncertainty on discharge calculations for main river basins. *Journal of Hydrometeorology,* v. 10, n. 4, p. 1011–1025.

Brasil Neto, R.M.; Santos, C. A. G.; Silva, J.F.C.B.dC.; da Silva, R. M., Dos Santos, C. A. C., & Mishra, M. Evaluation of the TRMM product for monitoring drought over Paraíba State, northeastern Brazil: a trend analysis. *Sci Rep* 11, 1097 (2021).

Brito, S. S. B.; Cunha, A. P. M. A.; Cunningham, C. C.; Alvalá, R. C.; Marengo, J. A.; Carvalho, M. A. 2018. Frequency, duration and severity of drought in the Semiarid Northeast Brazil region. *International Journal of Climatology*, v. 38, n. 2, p. 517–529.

Chen, C.; Chen, Q.; Duan, Z.; Zhang, J.; Mo, K.; Li, Z.; Tang, G. 2018. 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*, v. 10, n. 6, p. 1–18.

Chen, F.; Li, X. 2016. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China. *Remote Sensing*, v. 8, n. 6, p. 1–18.

Erazo, B.; Bourrel, L.; Frappart, F.; Chimborazo, O.; Labat, D.; Dominguez-Granda, L.; Matamoros, D.; Mejia, R. 2018. Validation of satellite estimates (Tropical Rainfall Measuring Mission, TRMM) for rainfall variability over the Pacific slope and Coast of Ecuador. *Water (Switzerland)*, v. 10, n. 2, p. 1–23.

Fang, J.; Yang, W.; Luan, Y.; Du, J.; Lin, A.; Zhao, L. 2019. Evaluation of the TRMM 3B42 and GPM IMERG products for extreme precipitation analysis over China. *Atmospheric Research*, v. 223, n. September 2018, p. 24–38.

Instituto Brasileiro de Geografia e Estatística. 2017. Report: Censo Demográfico. Brasil: IBGE.

Instituto Brasileiro de Geografia e Estatística. 2019. Report: Censo Demográfico. Brasil: IBGE.

Lima, F. V. M. S.; Gonçalves, R. M.; Castro, H. D. M.; Vila Nova, R. A. 2020. Assessment of Hydrological Mass Losses in the Northeast Atlantic Eastern Hydrographic Region, Brazil. *Boletim de Ciências Geodésicas*, v. 26, n. 3, p. 1–20.

Liu, Z.; Ostrenga, D.; Teng, W.; Kempler, S. 2012. Tropical rainfall measuring mission (TRMM) precipitation data and services for research and applications. *Bulletin of the American Meteorological Society*, v. 93, n. 9, p. 1317–1325.

Magalhães, M. N. *Probabilidade e Variáveis Aleatórias.* 2a edição ed. São Paulo, 2006.

Maggioni, V.; Meyers, P. C.; Robinson, M. D. 2016. 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.

Marengo, J. A.; Torres, R. R.; Alves, L. M. 2017. Drought in Northeast Brazil—past, present, and future. *Theoretical and Applied Climatology*, v. 129, n. 3–4, p. 1189–1200.

Mckee, T. B.; Doesken, N. J.; Kleist, J. The relationship of drought frequency and duration to time scales. *Eighth Conference on Applied Climatology*. Anais... , 1993.

Mitchell, T. D.; Jones, P. D. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, v. 25, n. 6, p. 693–712.

Moriasi, D. N.; Arnold, J. G.; Van Liew, M. W.; Bingner, R. L.; Harmel, R. D.; Veith, T. L. 2007. Model Evaluation Guidelines For Systematic Quantification of Accuracy in Watershed Simulations. *American Society of Agricultural and Biological Engineers*, v. 50, n. 3, p. 885–900.

Mpelasoka, F.; Awange, J.; Goncalves, R. M. 2018. Accounting for dynamics of mean precipitation in drought projections: A case study of Brazil for the 2050 and 2070 periods. *Science of the Total Environment*, v. 622, p. 1519-1531.

Nastos, P. T.; Kapsomenakis, J.; Philandras, K. M. 2016. Evaluation of the TRMM 3B43 gridded precipitation estimates over Greece. *Atmospheric Research*, v. 169, p. 497–514.

Naumann, G.; Barbosa, P.; Carrao, H.; Singleton, A.; Vogt, J. 2012. Monitoring drought conditions and their uncertainties in Africa using TRMM data. *Journal of Applied Meteorology and Climatology*, v. 51, n. 10, p. 1867–1874.

Ndehedehe, C. E.; Agutu, N. O.; Ferreira, V. G.; Getirana, A. 2020. Evolutionary drought patterns over the Sahel and their teleconnections with low frequency climate oscillations. *Atmospheric Research*, v. 233, p. 104700.

Nobre, C. A.; Marengo, J. A.; Seluchi, M. E.; Cuartas, L. A.; Alves, L. M. 2016. Some Characteristics and Impacts of the Drought and Water Crisis in Southeastern Brazil during 2014 and 2015. *Journal of Water Resource and Protection*, v. 08, n. 02, p. 252–262.

Nouri, H.; Nagler, P.; Borujeni, S. C.; Munez, A. B.; Alaghmand, S.; Noori, B.; Galindo, A.; Didan, K. 2020. Effect of spatial resolution of satellite images on estimating the greenness and evapotranspiration of urban green spaces. *Hydrological Processes*, v. 34, n. 15, p. 3183–3199.

Ozkan, U. Y.; Ozdemir, I.; Demirel, T.; Saglam, S.; Yesil, A. 2017. Comparison of satellite images with different spatial resolutions to estimate stand structural diversity in urban forests. *Journal of Forestry Research*, v. 28, n. 4, p. 805–814.

Paredes-Trejo, F.; Barbosa, H. A.; Rossato Spatafora, L. 2018. Assessment of SM2RAIN-Derived and State-of-the-Art Satellite Rainfall Products over Northeastern Brazil. *Remote Sens,* v. *10*, p. 1093.

Rodrigues, D. T.;Gonçalves, W. A.; Spyrides, M. H. C.; Silva, C. M. S. 2020. Spatial and temporal assessment of the extreme and daily precipitation of the Tropical Rainfall Measuring Mission satellite in Northeast Brazil. *International Journal of Remote Sensing*, v. 41, n. 2, p. 549-572.

Sampaio, T. V. M. 2012. Diretrizes e procedimentos metodológicos para a cartografia de síntese com atributos quantitativos via álgebra de mapas e análise multicritério. *Boletim de Geografia*, v. 30, n. 1, p. 121–131.

Santos, S. R. Q. do; Cunha, A. P. M. do A.; Ribeiro-Neto, G. G. 2019. Avaliação De Dados De Precipitação Para O Monitoramento Do Padrão Espaço-Temporal Da Seca No Nordeste Do Brasil. *Revista Brasileira de Climatologia*, v. 25, p. 80–100.

Schneider, U.; Finger, P.; Meyer-Christoffer, A.; Rustemeier, E.; Ziese, M.; Becker, A. 2017. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, v. 8, n. 3, p. 1–17.

Shukla, A. K.; Ojha, C. S. P.; Singh, R. P.; Pal, L.; Fu, D. 2019. Evaluation of TRMM Precipitation Dataset over Himalayan Catchment: The Upper Ganga Basin, India. *Water*, v. 11, n. 3, p. 613.

Soriano, É.; Londe, L. de R.; Di Gregorio, L. T.; Coutinho, M. P.; Santos, L. B. L. 2016. Water crisis in são paulo evaluated under the disaster's point of view. *Ambiente e Sociedade*, v. 19, n. 1, p. 21–42.

Suliman, A. H. A.; Awchi, T. A.; Al-Mola, M.; Shahid, S. 2020. Evaluation of remotely sensed precipitation sources for drought assessment in Semi-Arid Iraq. *Atmospheric Research*, v. 242, n. January, p. 105007.

Tao, H.; Fischer, T.; Zeng, Y.; Fraedrich, K. 2016. Evaluation of TRMM 3B43 precipitation data for drought monitoring in Jiangsu Province, China. *Water (Switzerland),* v. 8, n. 6, p. 1–13.

Vicente-Serrano, S. M.; Beguería, S.; López-Moreno, J. I. 2010. A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, v. 23, n. 7, p. 1696–1718.

Xu, R.; Tian, F.; Yang, L.; Hu, H.; Lu, H. Hou, A. 2017. 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*, v. 122, n. 2, p. 910–924.

Yong, B.; Liu, D.; Gourley, J. J.; Tian, Y.; Huffman, G. J.; Ren, L.; Hong, Y. 2015. Global view of real-time TRMM multisatellite precipitation analysis: Implications for its successor global precipitation measurement mission. *Bulletin of the American Meteorological Society*, v. 96, n. 2, p. 283–296.

Zarch, M. A. A.; Sivakumar, B.; Sharma, A. 2015. Droughts in a warming climate: A global assessment of Standardized precipitation index (SPI) and Reconnaissance drought index (RDI). *Journal of Hydrology*, v. 526, p. 183–195.