Satellite-Based Precipitation Estimates Validation Using Surface Stations in the Central Region of the State of São Paulo, From 1981 to 2019

Bruno César dos Santos (bruno-unifal@hotmail.com)
Universidade de São Paulo https://orcid.org/0000-0001-8218-6803
Rafael Grecco Sanches
University of São Paulo
Talyson de Melo Bolleli
University of São Paulo
Paulo Henrique de Souza
Federal University of Alfenas
Vandoir Bourscheidt
Federal University of Sao Carlos

Research Article

Keywords: Precipitation, Missing data, CHIRPS, Statistical analysis

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1153248/v1

Abstract

With the advance of remote sensing technologies, meteorological satellites have become an alternative in the process of monitoring and measuring meteorological variables, both spatially and temporally. The present study brings some additional elements to the validation of satellite-based precipitation estimates by evaluating the CHIRPS (Climate Hazards Group Infra-Red Precipitation with Station) monthly product for the central region of the state of São Paulo, Brazil, in the period 1981-2019. Initially, the general relationship between satellite estimates and surface rainfall data is assessed using the linear adjustment and error analysis in both temporal and spatial perspectives, followed by a trend analysis using Laplace test. The monthly map analysis showed a better performance of CHIRPS during the dry period (April to August) than for the wet period (October to March). Finally, monthly trends showed, in general, the same pattern of variability in rainfall over 38 years and a prevalence toward the reduction of rainfall. In summary, CHIRPS product seems a reasonable alternative for regions that lack historical rainfall information.

1 Introduction

Rainfall variability has been a constant topic in scientific climate research, in view of its importance on local and regional scale, and due to the extent of rainfall fluctuations impact on the most diverse regions of the planet (Abreu et al. 2017; Ambrizzi et al. 2014; Cunha et al. 2014; Filho et al. 2019; Madsen et al. 2014; Sanches et al. 2020; Serrano et al. 1999; Teixeira and Satyamurty 2011; Zandonadi et al. 2016 ; Zilli et al. 2017).

In this sense, the need to find long historical series of precipitation usually becomes a major obstacle for carrying out studies on this matter (Madsen et al. 2014; Mekis; Vincent 2011; Piccarreta et al. 2004). It is common to find gaps in surface data, which limit the quality of statistical analysis that may help to understand the climatic behavior of a given region, such as the case of rainfall trends (Blain, 2013; Carvalho et al. 2004; Nasseri et al. 2013; Sanches 2019; Teixeira; Satyamurty 2011; and Zandonadi et al. 2016).

Missing data in historical series are often caused by limitations in the resources used to measure or transcribe the collected information. Conventional equipment requires daily manual readings, which can sometimes lead to errors or lack of information over long time periods (Gimenez and Nery 2017; Coutinho et al. 2018). To overcome these limitations, it is customary to use different statistical methods, aiming to improve the quality and assist in filling data gaps, which are recurrent in historical series of climatic data (Parmar et al. 2017; Rasouli et al. 2012; Ridwan et al. 2020; and Sachindra et al. 2018). However, in many cases, these methods involve complex tools or require other historical series in the fault-filling process.

On the other hand, with the advance of remote sensing technologies, meteorological satellites have become an alternative in the process of monitoring and measuring meteorological variables, both spatially and temporally, allowing a better understanding of the atmospheric dynamics. Several meteorological satellites present extensive time series datasets that allow users to carry out long term studies (Salio et al. 2015; Aires et al. 2016; Dembélé & Zwart 2016; Castelhano et al. 2017, Pereira et al. 2018; Silva et al. 2019; Calvalcante et al. 2020), especially in studies of extreme events like droughts and floods.

In this way, CHIRPS (Climate Hazards Group Infra-Red Precipitation with Station) product present more than three decades of rainfall information with a spatial resolution of about 5 km, combined in daily, monthly, quarterly, and annual intervals (Funk et al. 2015). This product has been widely validated and may be considered as an alternative of great potential for analyzing rainfall time series for regions that do not have data from surface stations, or even as a resource to fill in data gaps.

In light of this, the present study aims to evaluate the monthly precipitation estimates from CHIRPS product at the Southeastern Brazil, by comparing the satellite-based estimates against data from surface stations, seeking to verify the potential of this dataset as an alternative solution for filling gaps or as a solution for studies of rainfall variability. The following sections describe the procedures adopted in the present study.

2 Material And Methods

2.1 Study area

The study area is located in Southeastern Brazil, in the center-east of the state of São Paulo, with an area of 9,151.7 km², and comprising a total of 16 municipalities (Fig. 1).

According to Monteiro's (1973) climate classification, these municipalities fall into tropical regional climates with dry (April to September) and humid (October to March) seasons, with a predominance of type A2/Vb (Serra de São Carlos). Furthermore, according to Kopen’s classification, the area is classified as Cwb, i.e., Subtropical highland climate with dry winters (Alvares et al. 2013). This represents an important characteristic of the rainfall distribution, mainly related to the regional atmospheric circulation pattern (Moruzzi and Oliveirac 2009; Zilli et al. 2017; Sanches et al. 2018; Sanches 2019; Santos et al. 2017; 2018 and 2020), which eventually lead to exceptional episodes in the region.

In geomorphological terms, the area is in the transition between two morphostructures, the Western Plateau, formed in a large area of smooth relief composed of hills, low hills and mountains, with an average altitude of approximately 900m; and the Peripheral Depression, which largely consists of hillock and smooth terrain features, in addition to isolated hills and mountains with levels up to approximately 600m (Ross; Moroz 1997; Pinheiro and Queiroz Neto 2014). Details of the terrain can be seen in the transects shown in Figure 1.

Thus, the complex terrain features located in the north and south of the area have a fundamental role in the regional circulation, especially when it is influenced by different weather types that may contribute to the rain formation or intensification along the seasons (Santos et al. 2018).

2.2 Database

Thirty-one pluviometric stations were used in the study, with daily data, from historical series of 38 years (1979-2019) and that have the minimum possible failures. For access to rain data, the following sources were consulted: the online hidroweb platform, which belongs to the National Water Agency (ANA); the website of the Department of Water and Electricity (DAEE); and the Integrated Center for Agro-meteorological Information (CIIAGRO), belonging to the Secretariat of Agriculture and Supply.

In addition to the surface stations, monthly data from the satellite-based product called CHIRPS (Climate Hazards Group Infra-Red Precipitation with Station), were also used for the period between 1981 and 2019. As previously stated, this dataset consist of precipitation estimates from 1981 to the present day, with quasi-global coverage (between 50 ° N to 50 ° S), with a spatial resolution of 0.05 ° (Funk et al. 2015). The estimation process goes through a series of steps that help in the quality control of the data, so that CHIRPS users can find it useful. The validation process involves three main stages, as shown in Fig. 2. The first stage, called CHPclim, uses information from normal stations, satellite stations, elevation of the terrain, latitude, and longitude as a source of processing. The second stage uses satellite data collected by means of cold cloud duration (CCD) precipitation estimates, dividing with the average and resulting in percentage values of precipitation. The values resulting from these two steps are multiplied, thus generating the product called CHIRP. The last step is the use of data observed by the surface stations, going through statistical procedures (IDW and% Bias Correction) and multiplying their values with those of the CHIRP and, thus, validating the precipitation data for the platform called CHIRPS.

The use of CHIRPS has been validated in several studies through its comparison with ground stations, seeking to demonstrate the quality of its information, serving as an alternative database in different regions of the world and also within the Brazilian territory (Nogueira; Moreira; Volpato 2018; Costa et al. 2019; Santos; Cunha; Ribeiro-Neto 2019; Silva et al 2019; Pereira et al. 2018; Castelhano; Pinheiro; Roseghini 2017; Moctar Dembélé and Sander 2016; Bai et al. 2018; Tote et al. 2015; Aksu and Akgül 2020; Alejo et al. 2021; Ghozat et al. 2021).

3 Procedures And Data Analysis

From the data obtained, we initially sought to assess the relationship between satellite estimates and rainfall data from surface stations, similarly to the studies mentioned above.

For this purpose, monthly data were collected from both sources (surface and orbital). For CHIRPS, the values of the grid points (pixels) equivalent or closer to each station were selected. With these data, we proceeded with the verification and cross-validation against ground station using different statistical methods, certifying the equivalence between the data in the period of 38 years (1981-2019).

To evaluate the use of orbital data (CHIRPS) as an alternative to fill in the gaps in historical series, as well as for areas that absent precipitation information, a comparative analysis was carried out using all surface stations (without failures) and the values estimated by the satellite. Then, monthly averages of the different statistics were taken and spatialized into 12 monthly maps using the inverse distance weighting technique (Farias et al. 2017), to get a better view of the results in the study area and the influence that the landscape determines on rainfall records, as well as to evaluate the impact of seasonal variations on the estimates. Finally, a rainfall trend analysis was made on a monthly base for the two datasets. The calculations and systematization of the data through tables and graphs were performed using Microsoft Office Excel 2007. The statistical techniques used for all these analyses are described in detail below.

3.1 Mean Error (ME)

ME is used to determine the absolute difference between the value observed at the surface station and the value estimated by the satellite, and may be computed as:

$$ME = \frac{1}{n}\sum_{t=0}^{n}(\hat{y}_t - \bar{\hat{y}}_t) \text{ Equation (1)}$$

Where:

- n: number of samples in the time series
- t: time interval (month)
- $\hat{y}_t$: estimated value by the satellite in each month
- $\bar{\hat{y}}$: value observed at the surface station in each month

The ME represents the oscillation in the error of one dataset in relation to another by means of the amplitude of the differences between them, without taking into account the underestimation or overestimation of the error (Montgomery et al. 2008). The ideal value is that close to zero, indicating a better agreement between the datasets (Hallak and Filho 2011).

3.2 Root Mean Square Error (RMSE)

The RMSE is the average measure of estimated errors, with the smallest errors / best adjustments found as closer to zero this value is. It can be calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}[\hat{y}_t - \bar{\hat{y}}_t]} \text{ Equation (2)}$$

Where:

- n: number of samples in the time series
- t: time interval (months)
- $\hat{y}_t$: value observed at the surface station in each month

- $\bar{\hat{y}}$: value estimated by the satellite in each month

The RMSE determines the variability in the error of two time series. For this, the errors obtained from the square of the difference between the value observed in the pluviometric stations and the data estimated by the satellite are added (Montgomery et al. 2008). According to Hallak and Filho (2011), this statistical criterion is normally used to express the accuracy of the numerical results, with the advantage that this parameter presents error values in the same dimension as the analyzed variable.

3.3 Determination coefficient

The R^2 is a measure of statistical adjustment that aims to estimate the values of Y as a function of the value of X, and is given by:

$$R^2 = \frac{\left[\sum_{t=1}^{n}(y_t - \bar{y})(\hat{y}_t - \bar{\hat{y}})\right]^2}{\sum_{t=1}^{n}[y_t - \bar{y}]^2 \sum_{t=1}^{n}\left[\hat{y}_t - \bar{\hat{y}}\right]^2} \quad \text{Equation (3)}$$

Where:

- n: number of samples in the time series
- t: time interval (months)
- y_t: value observed at the surface station in each month
- $\bar{y}$: average of the observed values
- $\hat{y}_t$: value observed by the satellite in each month
- $\bar{\hat{y}}$: estimated mean values

The R^2 evaluates how CHIRPS data fits with the data observed in the pluviometric stations using a linear model. In other words, it indicates the proportion of the variation in CHIRPS satellite estimates that can be explained by the total variation in the observed data (Montgomery et al. 2008). Morettin and Bussab (2010) complement that this criterion varies between 0 and 1, so that the higher the value of linear R^2, the better the quality of adjustment of the precipitation estimates from CHIRPS to the values observed in the meteorological stations.

3.4 Laplace trend test

The method known as the 'Laplace Test' (Laplace trend factor) is applied to the data set to observe trends in the annual precipitation.

The trend is determined based on a given value u(t) and considering a given period [0, t], by the following equation (Kanoun; Martini; Souza 1991):

$$u(t) = \frac{\frac{\sum_{i=1}^{t}((i-1)n_i)}{N(t)} - \frac{(t-1)}{2}}{\sqrt{\frac{t^2-1}{12(N(t))}}} \quad \text{Equation (4)}$$

Where:

- t: represents the number of months
- ni: is the variable analyzed at time i (monthly precipitation at each station or pixel)

N(t): indicates the cumulative number in relation to the analyzed variable.

4 Results And Discussion

4.1 Analysis using monthly maps

Here, the results from the statistical analysis presented in the previous section for the 31 stations were interpolated, allowing a detailed spatial analysis of the variables under consideration.

4.1.1 Mean error (ME)

Figure 3 presents the average values of ME for each station as interpolated maps, which indicate the spatial variations in absolute errors between observed data and CHIRPS estimated rainfall accumulation.

Along the year, ME values vary between -39mm to 43mm over the study area. The months corresponding to the dry period (April to September) presented values closer to zero (optimal value) and slightly toward the underestimation, with emphasis on June, July, August and September. On the other hand, the months associated with the wet period (October to March) presented overestimation in general with the exception of January, which presented the biggest negative errors, that is, a greater underestimation of the values in relation to the observed values. From the spatial perspective, by comparing these maps with Figure 1, one could infer that the largest errors (sub and overestimation) are located in regions associated with the orographic pattern of the transition from peripheral depression to the western plateau of São Paulo and, also, in the surrounding the Serra de Itaqueri.

4.1.2 Root mean square error (RMSE)

The maps in Fig. 4 present the interpolated RMSE monthly values. It is noted, again, that the period between April and September (dry period), presented RMSE values closer to zero (optimum value), with emphasis on the months of June, July and August (9 to 20mm). In the same way, the biggest errors occur during the wet season (October to March), especially in December, January and February. In addition, areas in the south-central part of the map presented larger errors in general, which is possibly related to the orography of the region, as already stated for ME analysis. This is also the case of the singularity in the map of November.

4.1.3 Determination coefficient (R^2)

The spatialization of the monthly average values of R^2 is shown on the maps of Fig. 5. These maps allow to observe the areas (or stations) with the best level of statistical adjustment between the datasets according to a linear regression model, which can also be used to indicate high or low correlation between them within the study area. In general, the maps indicate good adjustments or a good correlation ($R2 > 0.7$), especially in May, June, July and September. In the other months of the year, some areas showed better adjustments, with the R^2 values varying between 0.5 to 0.9 within the study area.

The similarity between the values, especially during the dry season in the Tropical climate, when the participation of atmospheric systems that inhibit rain prevails in the Brazilian territory, may be related with the reduced occurrence of warm clouds/rain, which lead to a general underestimation of rainfall (Dinku et. al. 2018). For temperate regions, the inverse pattern is observed between summer and winter due to the limited capability of CHIRPS in accounting for snowfall (Bai et. al. 2018). Paredes-Trejo et. al. (2020) found the same pattern in the Northeastern Brazil, but including stations at the coastal region, for which the authors also suggest limitations due to the warm precipitating systems as well. Thus, the pattern found in the present study might be related with the location of the analyzed region: inside the continent and at latitudes that influence the occurrence of precipitation more effectively during summer than in the winter.

In the same way as for ME and RMSE, the general spatial variation suggests that the study area is influenced by the effect of the landscape, mainly due to the presence of a mountain range in the south portion of the area. During the rainy season, the joint performance of the atmospheric systems and the presence of the topography provide favorable conditions for the formation of cloudiness and rain, along its entire length. Thus, as can be seen in the south-central region of the map, the linear adjustment showed lower R^2 values, especially in the months of transition from the dry to the rainy period (October, November, December and March). The same is not true for the northernmost areas of the map, which showed slightly higher values between these months. As discussed by Paredes-Trejo et. al. (2020) and Dinku et. al. (2018), terrain features induce warm orographic rainfall that is occasionally misclassified as no precipitating due to the adoption of a fixed IRP CCD threshold value in the CHIRPS product.

Therefore, it is possible that low correlations / worse adjustments occur due to the algorithm adopted for satellite data, based on the so-called “cold cloud duration” (CCD), which, in view of the recurrent presence of (warm) cloudiness over the region, lead to a general overestimation of CHIRPS in relation to the surface data. However, in an anomalous way, the month of January, considered the wettest month, showed a better correlation of the data over the entire map in relation to the other months of the same period. This may be associated with the establishment of the South Atlantic Convergence Zone (SACZ), is characteristic at this period, implying in consecutive days of precipitation and high rainfall accumulations.

4.1.4 Laplace trend test.

Figure 6 summarizes the results for the trend analysis. Values above (+2) indicate increasing trends and values below (-2) indicate decreasing trends in precipitation over time. It is possible to observe that the variability of the trends of both data are very close in terms of the climatic rhythm, with quite similar values for some of the stations (P5, P6, P21, P22, P23, P30 and P31) along the 38 years. On the other hand, for the other stations the values were similar during the period from 1981 to 2000 and, later, the differences between trends for the different series increase, maintaining, however, the climatic rhythms the series over the years.

Those periods that presented the greatest discrepancy in the compared values, coincide with moments in which there was a reduction in the number of stations (surface) used for the adjustment and quality control of the data estimated by CHIRPS, throughout its time series (Funk et al. 2015; Nogueira et al. 2018). This analysis can be seen in Fig. 7, which shows the monthly density of the number of surface stations used by CHIRPS to adjust the trend / bias, starting in January 1981 in the study region and around it. As suggested by Funk et al (2015), the use of CHIRP product could possibly give better results than CHIRPS for this case, since it does not consider the ground stations in the estimates.

It is noted that between 1981 and 1997 the density of stations was quite high, with approximately 70 stations used in the estimates. After 1997, the density decreased significantly, dropping to less than 6, or even without stations in some years after 2015. The observed reduction in the number of stations, possibly associated with data gaps in the selected stations, implies in adjustments (using the Mean Field Bias or MFB method) based on stations farther from the investigated area, which ends up reducing the quality of the adjustment between the data.

Thus, the gradual reduction in the number of surface observations may have affected the correction stage of the CHIRPS product, which reflects in precipitation estimates with a moderate bias (Xavier et al. 2016; Paredes-Trejo et al. 2017). In addition, it is worth mentioning that information on terrain elevation, climatology, geographic location and station sources are also included in the quality control process in its historical database (Funk et al. 2015).

4.2 Spatialized trend of rainfall

The map in Fig. 8 shows the general trend of monthly rainfall for the two historical series compared (surface stations and satellite estimates). It is noted that both data show, for most seasons, a negative trend in monthly rainfall at the studied area. This pattern of negative trends was also observed in the region by Sanches et al. (2020), who pointed to a reduction in rainfall for the municipality of São Carlos-SP. In other study over a larger area (Paraná basin), Rafee et al (2021) have found general negative trends for annual rainfall in the north portion of the watershed (which includes the region of this study), and positive trends when moving toward the south of the area. Their results also indicate increasing trends in extreme rainfall (more than 50mm) episodes toward the south portion of the basin.

For surface data, about 90% of rainfall stations showed a reduction in monthly rainfall and for satellite estimates, about 74% of the used data (i.e., the pixel closest to the rainfall station) also showed a decrease in rainfall. In the case of stationary trends, these were found in 7% of the surface posts and 19% for the CHIRPS pixels. Only 3% of the stations showed positive trends in the analyzed period, with this value being 7% for the pixels of the satellite data.

It is important to stress that, at some points in the central region of the map the rainfall trends varies from negative (surface) to normality to positive (satellite). These inversions might be associated either with limitations in both products (surface stations and CHIRPS product) or even with gaps in the historical series used for the comparison process.

Finally, the existence of complex terrain features in the region may influence the rainfall values as well (Dinku et al. 2011, Rahman et al. 2009, Toté et al. 2015; Paredes-Trejo et al. 2015) and, with this, resulting in a different pluviometric behavior due to orographic effect (Santos et al. 2019; 2020).

5 Conclusion

In this work, different statistical methods were applied to compare the monthly precipitation estimates of the CHIRPS product against surface measurements of 31 stations located in the central region of the State of São Paulo. In general, the results were satisfactory. For the monthly analysis using maps, the influence of the tropical climate seasonality, showed a better performance of the CHIRPS during the winter/dry period (April to August), where atmospheric systems that impact on the reduction of rains (blocking highs and polar systems) are more frequent, and there is a reduction in the warm clouds and rain systems.

However, during the wet period (October to March) the performance of CHIRPS tends to decrease, as the atmospheric systems (frontal and convective) acting during the summer, combined with the orographic effects and the presence of warm rain, results in high levels of precipitation and in the overestimation of the CHIRPS values, with a certain level of saturation (with values not exceeding 550mm). It is noteworthy that over the year, the month of January (considered the rainiest) was the one that presented a greater underestimation of the values in relation to those observed on surface stations.

Finally, monthly trends showed, in general, the same pattern of variability in rainfall during the period from 1981 to 2019 (38 years). In addition, there was a prevalence of the reduction in monthly rainfall for most of the compared points, with the exception of some periods when opposite trends were found for the data observed on the surface and those estimated by CHIRPS. This difference may be associated with orographic effects or could be related to the reduction in the number of surface stations used in the CHIRPS adjustment, mainly after the 2000s.

In spite of this, the CHIRPS product reveals, in general, a pluviometric rhythm very similar to the one observed by the surface stations, being able to become an alternative database in the process of filling gaps in historical series, or even as a solution for regions that do not have historical rainfall information. Since there is no perfect solution to the serious flaw that many collection points have, the use of these technologies / resources could overcome these limitations and allow a complete analysis of the time series, bringing an important contribution to the research and the understanding of the behavior of the atmosphere in the various portions of the Earth's surface, especially in areas where human and financial needs make it difficult to collect data in a continuous way.

Declarations

Acknowledgments

The authors thank the Brazilian National Water Agency (ANA) and the Climate Hazards Group Infra-Red Precipitation with Station (CHIRPS) for providing the ground station and estimated precipitation data used in this work.

Funding

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel (CAPES/Brazil) - Finance Code 001.

Author's Contribution

B.C.S. and V.B. contributed to the formal analysis, methodology, investigations, and validation. B.C.S. and R.G.S. contributed to the data handling in software, examining the results, and writing—original draft preparation. V.B., P.H.S., and T.M.B. contributed to the visualization, verify results, and writing and editing. The authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All datasets used in the paper are freely available. For ground stations, data from ANA may be downloaded using the hidroweb platform (https://www.snirh.gov.br/hidroweb/). Data from DAEE is available on the website http://www.hidrologia.daee.sp.gov.br/. Finally, data from CIIAGRO may be downloaded using the website http://www.ciiagro.org.br/ema/. All platforms are in Brazilian Portuguese. CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) is registered with Creative Commons and all data may be downloaded at https://data.chc.ucsb.edu/products/CHIRPS-2.0/.

Code availability

Not applicable.

References

1. de Abreu FG, Sobrinha LA, Brandão JLB (2017) Análise da distribuição temporal das chuvas em eventos hidrológicos extremos. Engenharia Sanitaria e Ambiental 22:239–250. https://doi.org/10.1590/s1413-41522016146750
2. Aires URV, Neto J, de Mello OM, ESTIMATIVAS DE PRECIPITAÇÃO PLUVIAL DERIVADAS DO SENSOR TRMM PARA A BACIA HIDROGRÁFICA DO RIO PARAOPEBA (2017) Scientia Agraria 17:57–66. https://doi.org/10.5380/rsa.v17i2.46384
3. Aksu H, Akgül MA (2020) Performance evaluation of CHIRPS satellite precipitation estimates over Turkey. Theor Appl Climatol 142:71–84. https://doi.org/10.1007/s00704-020-03301-5
4. Alejo LA, Alejandro AS (2021) Validating CHIRPS ability to estimate rainfall amount and detect rainfall occurrences in the Philippines. Theor Appl Climatol 145:967–977. https://doi.org/10.1007/s00704-021-03685-y
5. Alvares CA, Stape JL, Sentelhas PC et al (2013) Köppen's climate classification map for Brazil. Meteorol Z 711–728. https://doi.org/10.1127/0941-2948/2013/0507
6. Ambrizzi T, Sacco MAL, Ferraz SET (2014) A IMPORTÂNCIA DAS TELECONEXÕES PARA A PREVISÃO SAZONAL. Ciência. https://doi.org/10.5902/2179460X13155. e Natura 36:
7. Bai L, Shi C, Li L et al (2018) Accuracy of CHIRPS Satellite-Rainfall Products over Mainland China. Remote Sensing 10:362. https://doi.org/10.3390/rs10030362
8. Blain GC (2013) Seasonal variability of maximum daily rainfall in Campinas, state of Sao Paulo, Brazil: trends, periodicities, and associated probabilities/ Variabilidade inter-sazonal da precipitacao pluvial maxima diaria em Campinas, estado de Sao Paulo: tendencias, periodicidades e probabilidades associadas. Acta Scientiarum Technology 35:557
9. Caparoci Nogueira SM, Moreira MA, Lordelo Volpato MM (2018) Evaluating Precipitation Estimates from Eta, TRMM and CHRIPS Data in the South-Southeast Region of Minas Gerais State—Brazil. Remote Sensing 10:313. https://doi.org/10.3390/rs10020313
10. Cavalcante RBL, Ferreira DB, da Pontes S et al (2020) Evaluation of extreme rainfall indices from CHIRPS precipitation estimates over the Brazilian Amazonia. Atmos Res 238:104879. https://doi.org/10.1016/j.atmosres.2020.104879
11. Costa JC, Pereira G, Siqueira ME, VALIDAÇÃO DOS DADOS DE PRECIPITAÇÃO ESTIMADOS PELO CHIRPS PARA O BRASIL (2019) Revista Brasileira de Climatologia 24.

https://doi.org/10.5380/abclima.v24i0.60237

12. Coutinho ER, da Silva RM, Madeira JGF et al (2018) Application of Artificial Neural Networks (ANNs) in the Gap Filling of Meteorological Time Series. Rev bras meteorol 33:317–328. https://doi.org/10.1590/0102-7786332013
13. da Cunha AC, Vilhena JEDS, Santos ESD et al (2014) Evento extremo de chuva-vazão na bacia hidrográfica do rio Araguari, Amapá, Brasil. Revista Brasileira de Meteorologia 29:95–110. https://doi.org/10.1590/0102-778620130051
14. Dembélé M, Zwart SJ (2016) Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa. Int J Remote Sens 37:3995–4014. https://doi.org/10.1080/01431161.2016.1207258
15. Dinku T, Funk C, Peterson P et al (2018) Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q J R Meteorol Soc 144:292–312. https://doi.org/10.1002/qj.3244
16. Farias OG, Francisco CN, Senna MCA, AVALIAÇÃO DE MÉTODOS DE INTERPOLAÇÃO ESPACIAL APLICADOS À PLUVIOSIDADE EM REGIÃO MONTANHOSO NO LITORAL SUL DO ESTADO RIO DE JANEIRO (2017) Revista Brasileira de Climatologia 21. https://doi.org/10.5380/abclima.v21i0.52065
17. Filho DFF, Bezerra PES, de Silva M, APLICAÇÃO DE TÉCNICAS DE INTERPOLAÇÃO PARA ESPACIALIZAÇÃO DE CHUVAS DA REDE HIDROGRÁFICA: ESTUDO DE CASO CALHA NORTE – PA (2019). Revista Brasileira de Climatologia 24:. https://doi.org/10.5380/abclima.v24i0.61516
18. Funk C, Peterson P, Landsfeld M et al (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2:150066. https://doi.org/10.1038/sdata.2015.66
19. Ghozat A, Sharafati A, Hosseini SA (2021) Long-term spatiotemporal evaluation of CHIRPS satellite precipitation product over different climatic regions of Iran. Theor Appl Climatol 143:211–225. https://doi.org/10.1007/s00704-020-03428-5
20. Gimenez DFS, Nery JT (2017) Aplicação das redes neurais artificiais no preenchimento de dados diários de chuva no estado de São Paulo. Os Desafios da Geografia Física na Fronteira do Conhecimento 1:1747–1755. https://doi.org/10.20396/sbgfa.v1i2017.2379
21. Hallak R, Pereira Filho AJ (2011) Metodologia para análise de desempenho de simulações de sistemas convectivos na região metropolitana de São Paulo com o modelo ARPS: sensibilidade a variações com os esquemas de advecção e assimilação de dados. Rev bras meteorol 26:591–608. https://doi.org/10.1590/S0102-77862011000400009
22. Kanoun K, Martini MR, de Souza B (1991) A method for software reliability analysis and prediction application to the TROPICO-R switching system. IEEE Trans Software Eng 17:334–344. https://doi.org/10.1109/32.90433
23. Madsen H, Lawrence D, Lang M et al (2014) Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. J Hydrol 519:3634–3650. https://doi.org/10.1016/j.jhydrol.2014.11.003

24. Mekis É, Vincent LA (2011) An Overview of the Second Generation Adjusted Daily Precipitation Dataset for Trend Analysis in Canada. Atmos Ocean 49:163–177. https://doi.org/10.1080/07055900.2011.583910
25. Monteiro CAF (1973) A Dinâmica Climática e as Chuvas no Estado de São Paulo: estudo geográfico sob a forma de atlas. USP/IG, São Paulo
26. Montgomery DC, Jennings CL, Kulahci M (2008) Introduction to Time Series Analysis and Forecasting, 2– edn. John Wiley & Sons, ed. New Jersey, p 469
27. Morettin PA, Bussab WO (2010) Estatística Básica. 6« ed. São Paulo: Editora Saraiva, 210 p
28. Moruzzi RB, de Oliveira SC (2009) Relação entre intensidade, duração e freqüência de chuvas em Rio Claro, SP: métodos e aplicação. 10
29. Nasseri M, Tavakol-Davani H, Zahraie B (2013) Performance assessment of different data mining methods in statistical downscaling of daily precipitation. J Hydrol 492:1–14. https://doi.org/10.1016/j.jhydrol.2013.04.017
30. Paredes-Trejo FJ, Barbosa HA, Lakshmi Kumar TV (2017) Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. J Arid Environ 139:26–40. https://doi.org/10.1016/j.jaridenv.2016.12.009
31. Parmar A, Mistree K, Sompura M (2017) Machine Learning Techniques For. A Review, Rainfall Prediction
32. Pereira G, Cardozo F, da Negreiros S, DA VARIABILIDADE DA PRECIPITAÇÃO PARA O ESTADO DE MINAS GERAIS (2018) ANÁLISE (1981-2017). Revista Brasileira de Climatologia 1:. https://doi.org/10.5380/abclima.v1i0.61028
33. Pereira G, Silva M, Moraes E, Cardozo F (2013) Avaliação dos Dados de Precipitação Estimados pelo Satélite TRMM para o Brasil. RBRH 18:139–148. https://doi.org/10.21168/rbrh.v18n3.p139-148
34. Piccarreta M, Capolongo D, Boenzi F (2004) Trend analysis of precipitation and drought in Basilicata from 1923 to 2000 within a southern Italy context. Int J Climatol 24:907–922. https://doi.org/10.1002/joc.1038
35. Pinheiro MR, Neto JP, de Q (2014) Reflexões sobre a gênese da serra geral e da depressão periférica paulista: o exemplo da região da Serra de São Pedro e do baixo Piracicaba, SP. Revista do Instituto Geológico 35:47–59. https://doi.org/10.5935/0100-929X.20140004
36. Rafee SAA, Freitas ED, Martins JA et al (2020) Spatial Trends of Extreme Precipitation Events in the Paraná River Basin. Journal of Applied Meteorology and Climatology 59:443–454. https://doi.org/10.1175/JAMC-D-19-0181.1
37. Rahman SH, Sengupta D, Ravichandran M (2009) Variability of Indian summer monsoon rainfall in daily data from gauge and satellite. Journal of Geophysical Research: Atmospheres 114. https://doi.org/10.1029/2008JD011694
38. Rasouli K, Hsieh WW, Cannon AJ (2012) Daily streamflow forecasting by machine learning methods with weather and climate inputs. J Hydrol 414–415:284–293. https://doi.org/10.1016/j.jhydrol.2011.10.039

39. Ridwan WM, Sapitang M, Aziz A et al (2020) Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia. Ain Shams Engineering Journal. https://doi.org/10.1016/j.asej.2020.09.011
40. Ross JLS, Moroz IC, MAPA GEOMORFOLÓGICO DO ESTADO DE SÃO PAULO (1996) Revista do Departamento de Geografia 10:41–58. https://doi.org/10.7154/RDG.1996.0010.0004
41. Sachindra DA, Ahmed K, Rashid MdM et al (2018) Statistical downscaling of precipitation using machine learning techniques. Atmos Res 212:240–258. https://doi.org/10.1016/j.atmosres.2018.05.022
42. Salio P, Hobouchian MP, García Skabar Y, Vila D (2015) Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. Atmos Res 163:146–161. https://doi.org/10.1016/j.atmosres.2014.11.017
43. Sanches RG, Neves GZ, de Santos F, dos BC et al (2018) Intense Rainfall in São Carlos/SP: Determination of Threshold Values Using Climate Indices and Their Spatio-Temporal Repercussion. American Journal of Climate Change 07:388. https://doi.org/10.4236/ajcc.2018.73023
44. Sanches RG, Santos BCD, Miani RS et al (2020) Analysis of Daily Rainfall in São Carlos/SP, Brazil over 1979-2017 Using Laplace Trend Test. Journal of Geoscience and Environment Protection 8:104–125. https://doi.org/10.4236/gep.2020.87006
45. Sanches RG, dos Santos BC, Silva MSD, Vecchia FAS (2017) As chuvas na região de São Carlos/SP: análise dos valores máximos diários na série histórica, 1993-2014. Os Desafios da Geografia Física na Fronteira do Conhecimento 1:2280–2292. https://doi.org/10.20396/sbgfa.v1i2017.2589
46. Santos BC, Sanches RG, Silva MSD, ANÁLISE DO EFEITO OROGRÁFICO POR MEIO DA INTERPOLAÇÃO DE ÍNDICES CLIMÁTICOS (2018) Revista de Geografia - PPGEO - UFJF 8:114–132. https://doi.org/10.34019/2236-837X.2018.v8.26005
47. Santos BC, Fontão PAB, de Souza PH (2020) O efeito do relevo nas chuvas na porção central do Estado de São Paulo em anos padrão extremos. Revista do Departamento de Geografia 40:132–147. https://doi.org/10.11606/rdg.v40i0.172973
48. Santos SRQD, Cunha APM, Ribeiro-Neto A, GG AVALIAÇÃO DE DADOS DE PRECIPITAÇÃO PARA O MONITORAMENTO DO PADRÃO ESPAÇO-TEMPORAL DA SECA NO NORDESTE DO BRASIL (2019). Revista Brasileira de Climatologia 25:. https://doi.org/10.5380/abclima.v25i0.62018
49. Silva CB, Silva MES, Ambrizzi T, PRECIPITAÇÃO NA AMÉRICA DO SUL – DADOS OBTIDOS POR ESTAÇÕES METEOROLÓGICAS AUTOMÁTICAS E POR SISTEMAS ORBITAIS (2019) Revista Brasileira de Climatologia 25. https://doi.org/10.5380/abclima.v25i0.58813
50. Teixeira M, da Satyamurty S P (2011) Trends in the Frequency of Intense Precipitation Events in Southern and Southeastern Brazil during 1960–2004. J Climate 24:1913–1921. https://doi.org/10.1175/2011JCLI3511.1
51. Toté C, Patricio D, Boogaard H et al (2015) Evaluation of Satellite Rainfall Estimates for Drought and Flood Monitoring in Mozambique. Remote Sensing 7:1758–1776. https://doi.org/10.3390/rs70201758

52. Xavier AC, King CW, Scanlon BR (2016) Daily gridded meteorological variables in Brazil (1980–2013). Int J Climatol 36:2644–2659. https://doi.org/10.1002/joc.4518
53. Zandonadi L, Acquaotta F, Fratianni S, Zavattini JA (2016) Changes in precipitation extremes in Brazil (Paraná River Basin). Theoret Appl Climatol 123:741–756. https://doi.org/10.1007/s00704-015-1391-4
54. Zilli MT, Carvalho LMV, Liebmann B, Silva Dias MA (2017) A comprehensive analysis of trends in extreme precipitation over southeastern coast of Brazil. Int J Climatol 37:2269–2279. https://doi.org/10.1002/joc.4840
55. Zilli MT, Carvalho LMV, Liebmann B, Silva Dias MA (2017) A comprehensive analysis of trends in extreme precipitation over southeastern coast of Brazil. Int J Climatol 37:2269–2279. https://doi.org/10.1002/joc.4840

Figures

Figure 1

Location of the study area (in red) within the São Paulo climate classification according to Monteiro (1973). The lines (horizontal and vertical) on the hypsometric map indicate the location of the transects shown at the bottom. The yellow dots indicate the location of the surface stations, and the black dots indicate the location of the satellite's pixels (CHIRPS). Source: Climate maps adapted from Alvares (2013) and Monteiro (1973).

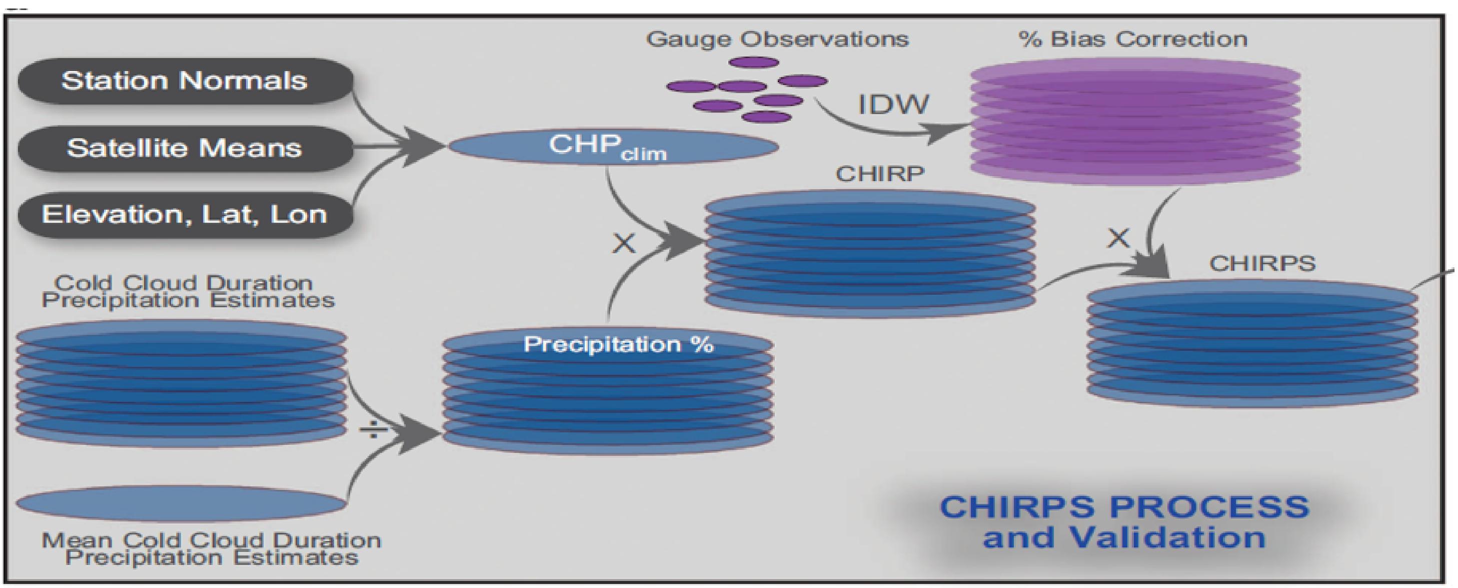

Figure 2

Steps taken in estimating and validating the precipitation data for the CHIRPS product. Source: Funk et al. (2015)

Figure 3

Monthly maps of the Mean Absolute Error (ME) varying between -39 mm (red) to 43 mm (green) in the period 1981-2019. Values close to zero (yellow) indicate a better agreement between observed and estimated rainfall data. Source: prepared by the authors

Figure 4

Monthly maps of the RMSE ranging from 0 mm (green) to 105 mm (red) in the period 1981-2019. The closer to zero, the better the agreement between the data observed by the surface stations (ANA) and the estimated data by satellite (CHIRPS). Source: prepared by the authors

Figure 5

Monthly spatialization of R^2 values between the data observed on the surface (ANA) and those estimated by satellite (CHIRPS). The closer the values are to 1, the greater the coefficient of determination and when the values are close to 0 the lower the coefficient. Source: Prepared by the authors

Figure 6

Analysis of the monthly trend of surface station data (light blue line) and estimated by CHIRPS (dark blue line) over 38 years (1981-2019). The intervals (+2 to -2) between the two red lines indicate the stability (normal range) of the series. Source: prepared by the authors

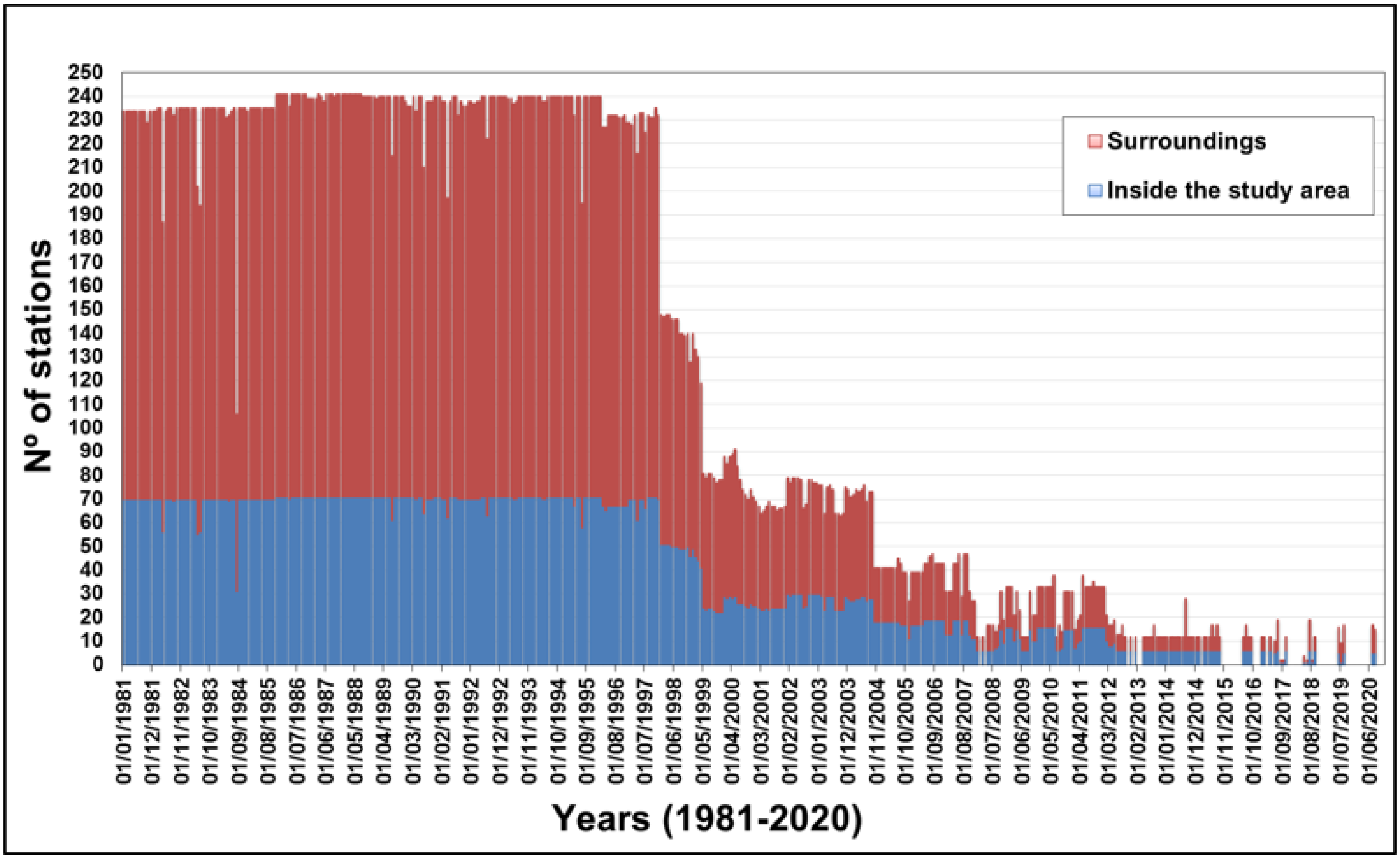

Figure 7

Monthly density of the number of surface stations used by CHIRPS inside (blue) and around (red) the study area, from 1981 to 2020. Source: Adapted from CHIRPS (2021)

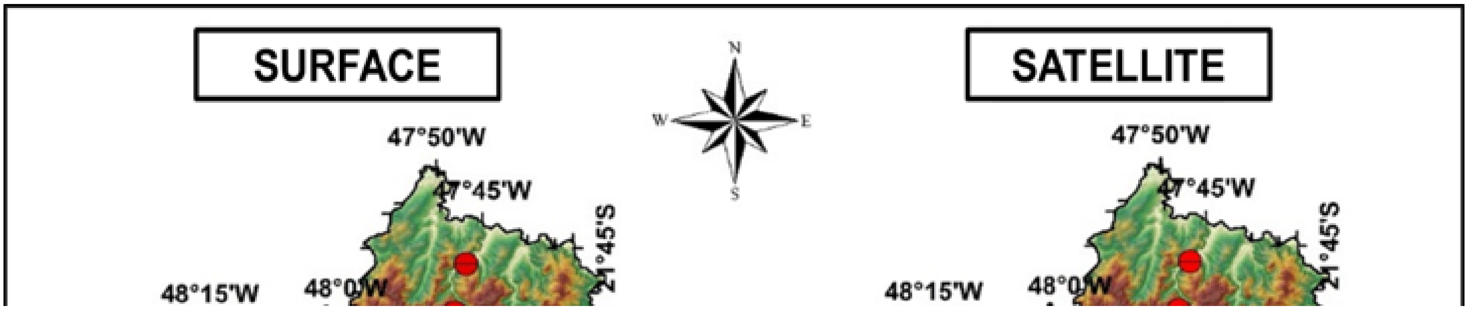

Figure 8

Monthly rainfall trend for data observed on the surface (ANA) and those estimated by satellite (CHIRPS) in the study area. Source: prepared by the authors