Performance of TRMM TMPA 3B42 V7 RT in replicating daily rainfall and rainfall regimes in the Amazon Basin (2000-2013)

Véronique Michot [1,4]
Daniel Vila [2]
Damien Arvor[1]
Beatriz M. Funatsu[1]
Vincent Dubreuil[1]
Josyane Ronchail[3,4]

[1] University of Rennes 2, LETG-Rennes, UMR 6554 CNRS
Place Recteur Henri Le Moal - 35000 RENNES - France
veronique.michot, damien.arvor, beatriz.funatsu, vincent.dubreuil@univ-rennes2.fr

[2] Instituto Nacional de Pesquisas Espaciais - INPE
Rod. Pres. Dutra, s/n – Zona Rural - 12630-000 – Cachoeira Paulista - SP, Brasil
daniel.vila@cptec.inpe.br

[3] University Paris Diderot, Sorbonne Paris Cité
Case Courrier 7001 – 75205 – Paris Cedex 13 - France

[4] LOCEAN, Institut Pierre Simon Laplace
4 place Jussieu – 75005 – Paris – France
josyane.ronchail@locean-ipsl.upmc.fr

Abstract. In the context of climate change, knowledge and studies on precipitation in the Amazon Basin (AB) are determinant for environmental aspects such as hydrology, ecology, as well as for social aspects like agriculture, food security, or health issues. Rainfall data availability with a wide geographical coverage and a good resolution is thus crucial for these purposes. Remote sensing techniques can provide an alternative to ground-based rainfall data but it is imperative to assess the quality of these estimated data. Hence, this paper aims to describe the reliability of the Topical Rainfall Measurement Mission (TRMM) 3B42 Version 7 (V7) Real Time (RT) at daily scale over the Amazon Basin (AB). The study covers the period of March 2000 to July 2013. TRMM estimates were compared with observations from 205 quality-controlled rain gauges stations. First, the spatial distribution of mean daily rainfall across the whole AB was validated using quantitative (Bias, Relative RMSE) and categorical statistics (POD, FAR). Estimations are more accurate in the lowlands than in the highlands. Furthermore a contrast of quality in estimates within the Andean regions can be noted. Second, TRMM 3B42 V7 RT succeeds in replicating regional rainfall regimes across the AB. The Andean regions show the worse results. However neither relevant relationships nor correlations appear with the altitude. However, the location of a rain gauge and its exposure seem to be relevant explanations of overestimations by TRMM 3B42 V7 RT in these mountainous regions.

Keywords: TRMM 3B42 V7 RT, rain gauge, Amazon Basin, precipitation, regional rainfall regimes, daily scale.

1. Introduction

The knowledge of climatic patterns and the monitoring of precipitation data play a very important role for the management of water resources, agriculture, or for health monitoring for instance, but this knowledge depends on the availability of regular rainfall data. Climatic studies usually rely on long time series (e.g. at least thirty years) to carry out trend analysis. For this reason long time series observations from rain gauges are preferred to remote-sensing based estimates. But the rain gauges are scarce and unevenly distributed in the AB (Ronchail et al. 2002, Delahaye 2013). Furthermore the resolution of the information is very local. In contrast, estimated data from remote sensing are remarkable as they allow enhancing the spatial information. The TRMM 3B42 V7 RT offers such possibility. However, since these data are estimated, it is necessary to assess the adequacy of estimations to represent the observations, and more specifically, to verify whether TRMM 3B42 V7 RT reproduces the annual rainfall regimes in the different regions of the AB. Therefore, the purpose of this work is to compare rain gauges against TRMM 3B42 V7 RT at daily scale from March 2000 to July 2013, in order to investigate the performance of this product to estimate precipitations at the AB scale for the whole period, and secondly at regional scale for each month of the year.

2. Data and methodology

2.1 Study area

The AB is located between about 6°N and 20.5°S and 48.5°W and 80.5°W (Figure 1). It spreads over 6 countries (Brazil, Venezuela, Colombia, Ecuador, Peru and Bolivia). The Basin presents a contrast between the highlands of the Andean mountains in the western and southwestern parts and the lowlands in the central, northern and eastern parts.

2.1 Observed precipitations : Rain gauge data

For this work, the network of observed data used as a reference to validate estimated data consists in 205 rain gauges at daily temporal resolution, which have been quality controlled (Figure 1; Michot et al. 2014). These data were obtained from the National Water Agency (ANA) and the National Meteorological Institute (INMET) in Brazil, the National Meteorological and Hydrological Institute (INAMHI) in Ecuador, the Hydrological Meteorological and Environmental Studies Institute (IDEAM) in Colombia and the National Hydrological and Meteorological Service (SENAMHI) in Peru and Bolivia. Unfortunately no such data could be collected in Venezuela. Figure 1 shows that rain gauges are unevenly distributed.

2.2 Estimated precipitations : TRMM TMPA 3B42 version 7 Real Time daily product

The TRMM 3B42 V7 Real Time (hereafter 3B42 RT) daily product is computed with the TRMM Multisatellite Precipitation Analysis (TMPA) algorithm near real-time, developed by the NASA Goddard Space Flight Center (Huffman et al., 2007). It consists in a gridded precipitation product with a spatial horizontal resolution of 0.25° x 0.25° (25km x 25km), between 50°N to 50°S. Unlike TRMM 3B42 V7, 3B42 RT is not gauge-adjusted, so that the independence between estimated and observed data is assured. The TMPA algorithm combines multiple independent precipitation estimates from the TMI, Advanced Microwave Scanning Radiometer for Earth Observing Systems (AMSR-E), Special Sensor Microwave Imager (SSMI), Special Sensor Microwave Imager/Sounder (SSMIS), Advanced Microwave Sounding Unit (AMSU), Microwave Humidity Sounder (MHS), and microwave-adjusted merged geo-infrared (IR).

2.3 Intercomparison methodology

The comparison between observed and estimated data was based on a point-to-pixel approach, using the nearest pixel of the rain gauge. This approach was adopted to take into account the difference of spatial resolution between the two datasets, and since a smaller spatial scale is necessary to measure heavy rainfall events (Thiemig et al. 2012). The validation of 3B42 RT focus on a climatological description, as we show the spatial distribution of quantitative and categorical statistics over the whole AB.

ISBN: 978-85-17-00088-1

The quantitative statistics used here are the mean daily rainfall, the Bias (equation 1) and the Relative Root Mean Square Error (relative RMSE, equation 2) while the categorical statistics are the Probability Of Detection (POD, equation 3) and the False Alarm Ratio (FAR, equation 4). These statistics are defined as follows:

$$Bias = 100\frac{\sum_{i=1}^{N}(Pe,i - Po,i)}{\sum_{i=1}^{N} Po,i} \quad (1) \qquad Relative\ RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Po,i - Pe,i)^2}{n}} / \overline{Po} \quad (2)$$

$$POD = \frac{H}{H+M} \quad (3) \qquad FAR = \frac{F}{H+F} \quad (4)$$

where *Pe* is the estimation and *Po* the observation; H (hit) is a precipitation event observed at the rain gauge and also detected by 3B42 RT; M (miss) is a precipitation event observed at the rain gauge but not detected by 3B42 RT; F (false alarm) is a precipitation event detected by 3B42 RT but not observed in the rain gauge. The bias indicates the overestimation or the underestimation by 3B42 RT in percentage. The relative RMSE gives a mean of the error of 3B42 RT, in millimeter (mm). The POD measures the number of rainfall events correctly detected by the estimated product with values ranging between 1 (a perfect score) and 0. The FAR measures the fraction of wrong events detected by 3B42 RT with values ranging between 1 (the worse score) and 0. The precipitation threshold to determine the occurrence of a precipitation event is 0.1mm.

Furthermore, in order to assess how the estimate product reproduces the diversity of seasonal cycles within the AB, the performance of 3B42 RT was assessed using the standard statistical methods above mentioned, at regional scale and for each month of the year. The regional mean annual rainfall and regimes have been computed for five regions that have been defined using a spectral clustering method (Figure 1).

As the results of the comparisons between estimated and observed data may vary within a region, potential explanations of these differences are examined using the geographical characteristics of the stations (altitude, windward or leeward location).

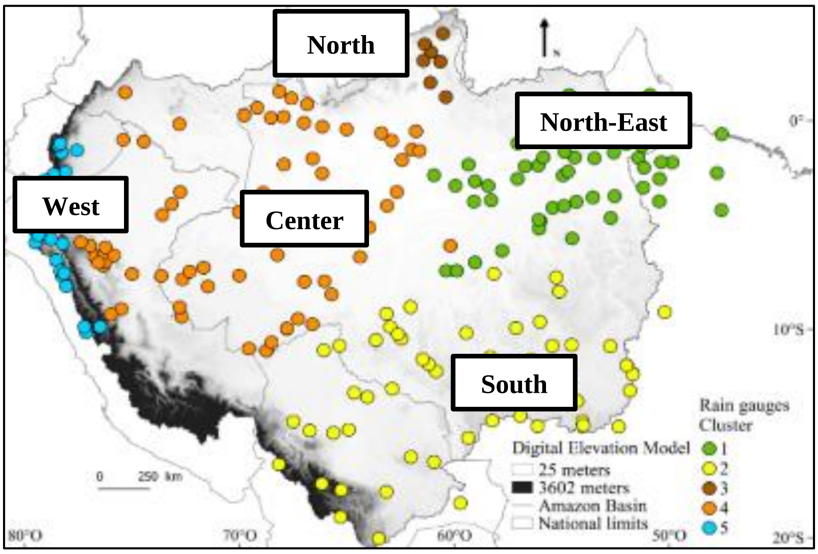

Figure 1. Rain gauge network used as ground reference. The colors of the stations and the labels allow identifying the climatic regions with similar regimes.

ISBN: 978-85-17-00088-1

3 Results and discussion
3.1 Large scale analysis

At large scale, for the whole AB, it can be noted that 3B42 RT well reproduces the spatial distribution of mean daily rainfall (Figure 2 a and b), with the highest rates of rainfall at the equatorial locations while rainfall decreases southward and northward in tropical latitudes. The lowest values are in mountainous regions as expected. But the estimations tend to underestimate rainfall in the lowlands. In the highlands the results are more contrasted. Indeed, even if the estimated precipitations are mainly underestimated, differences can be observed between the northern, middle and southern parts of the Andes. In the middle part of the Andes, the underestimation seems to be stronger than in the northern part, and rainfall is overestimated in the southern part of the Andes. This result is coherent with those found by Salio et al. (2015) for the Brazilian and Bolivian Amazon between 2008 and 2010.

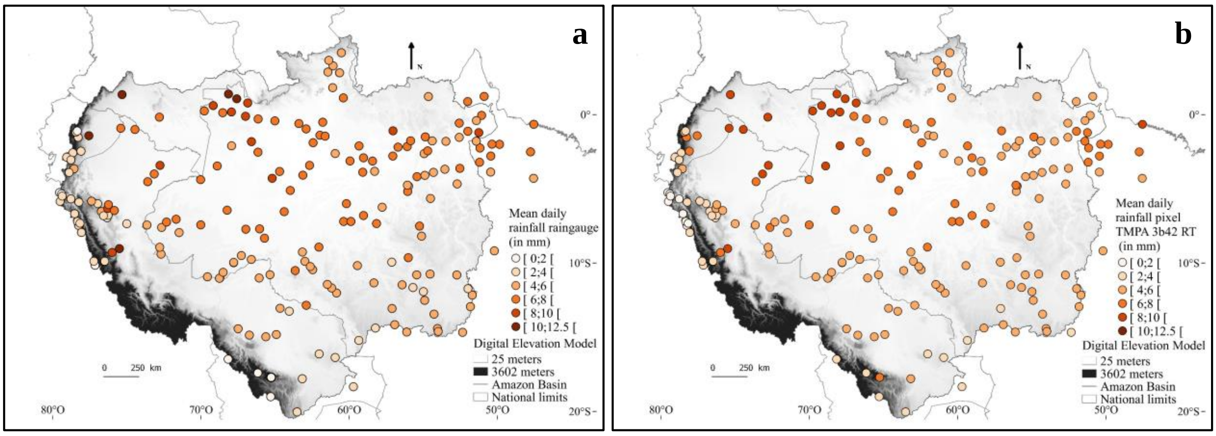

Figure 2. Spatial distribution of a) mean daily rainfall per rain gauge b) mean daily rainfall in the pixel 3B42 RT which is the closest to the rain gauge. Both are expressed in millimeters.

These behavior and differences between lowlands and highlands and within the highlands are also depicted by the other indicators. The most relevant is the bias (Figure 3 a) which shows a) a positive bias in most stations and b) greatest over and underestimation values in the Andean mountains while the lowest values are observed in the lowlands (with some exceptions in the North-East and South). In the Andes, the bias are positive, meaning that the estimated values are higher than the observations, in some stations of Ecuador (northern Andes) and Bolivia (southern Andes), while the bias is generally negative in Peru (Central Andes). This division between northern and Central Andes is also observed by Zulkafli et al. (2014). The heterogeneity of the bias values is also found by Salio et al. (2015) for the Brazilian and Bolivian Amazon and the order of magnitude for the underestimation is nearly the same.
The relative RMSE is used to quantify the error between estimated and observed data (Figure 3 b). In most stations the RMSE is lower than 2.55 mm. But in some Andean stations and in stations of the southern tropics the RMSE can reach values as high as 11 mm/day.
Better results are also observed in the lowlands with the categorical statistics. Indeed, the POD values indicate a good performance of the estimated product along a diagonal from northwestern to southeastern Amazon, along the mean position of the South Atlantic Convergence Zone that produce heavy rainfall in summer (Figure 4 a). There is also a correct detection of rainy events in the middle and southern Andes. The worse values are observed in the northern part of the Andes and in the Bolivian lowlands, but it has to be noted that the minimum of POD is not lower than 0.4.

ISBN: 978-85-17-00088-1

The false detection of rainy events (FAR) opposes the Andes and south of the AB with the worst results and the equatorial and northern locations that have pretty good results (Figure 4 b). Globally the results are fairly good with 149 over 205 points below 0.5, and intra-Andean differences are not as evident as they are with other indicators.

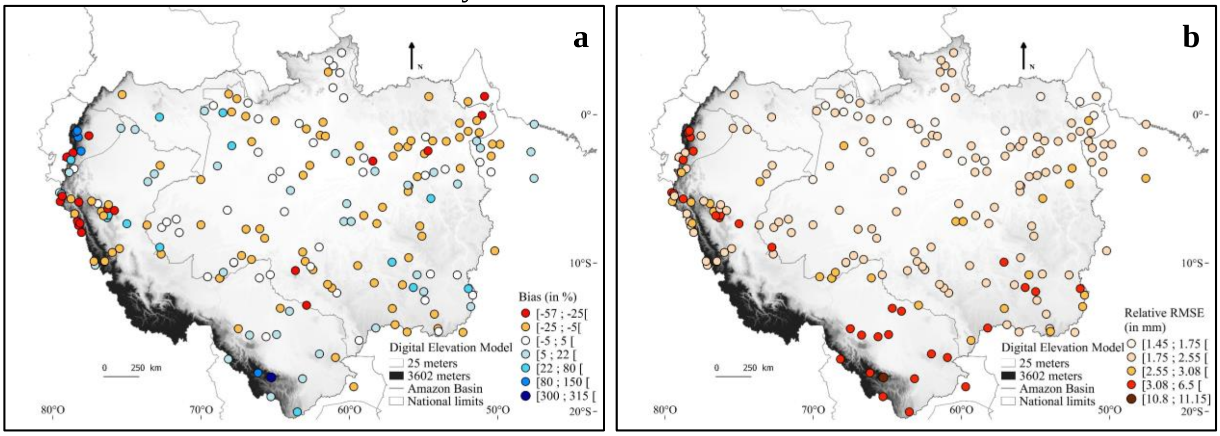

Figure 3. Spatial distribution of a) 3B42 RT Bias in percent, and b) 3B42 RT Relative RMSE, in millimeters. Both are calculated from 2000 to 2013 at daily scale.

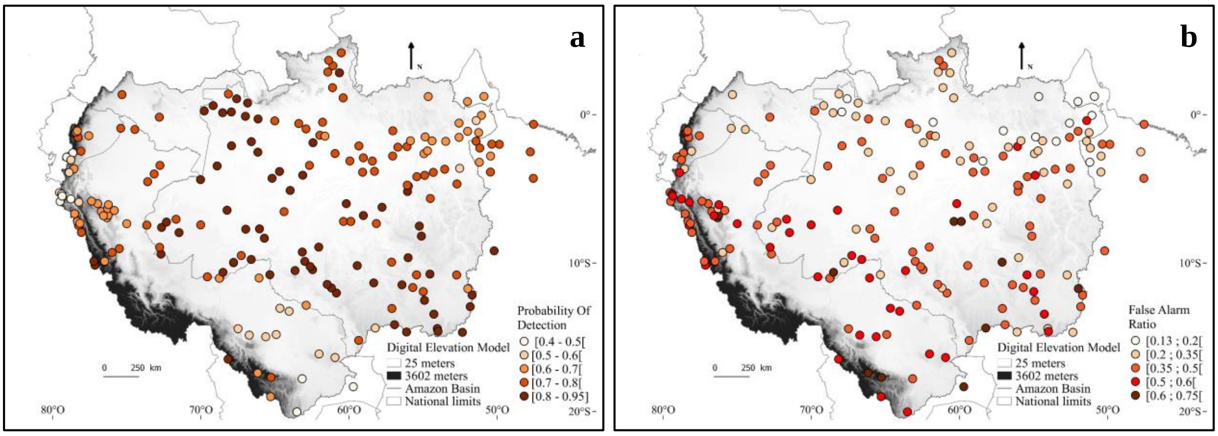

Figure 4. Spatial distribution of a) the POD and b) the FAR. Both are calculated from 2000 to 2013 at daily scale.

3.2 Regional analysis

Since the comparison between estimated and observed rainfall values differs throughout the AB, the statistical tools are computed at regional scale. Moreover, they are computed for each month to examine how 3B42 RT reproduces the annual cycle of precipitation. The results may be summarized as follows: the regions (Figure 1) may be divided in 2 groups, one with regions where rainy and dry seasons alternate (Figure 5, regions Northeast, South, North), and the other without pronounced dry or rainy season (Figure 5, Center, West).

Figure 5 shows a close reproduction of the annual precipitation cycle by 3B42 RT in all the regions. In the same figure the regions Northeast and Center present frequent underestimated months, which is coherent with the Bias results (Figure 3 a). The regions South and North (Figure 5) highlight the same number of months with under or overestimation. In the region South underestimation occur mainly during the first part of the hydrological year. The region North shows more contrasted behavior with the month of May showing greatest overestimation. Consistently with the results of part 3.1, overestimation is much more frequent in the region West, that is, in the Andean region (Figure 5).

ISBN: 978-85-17-00088-1

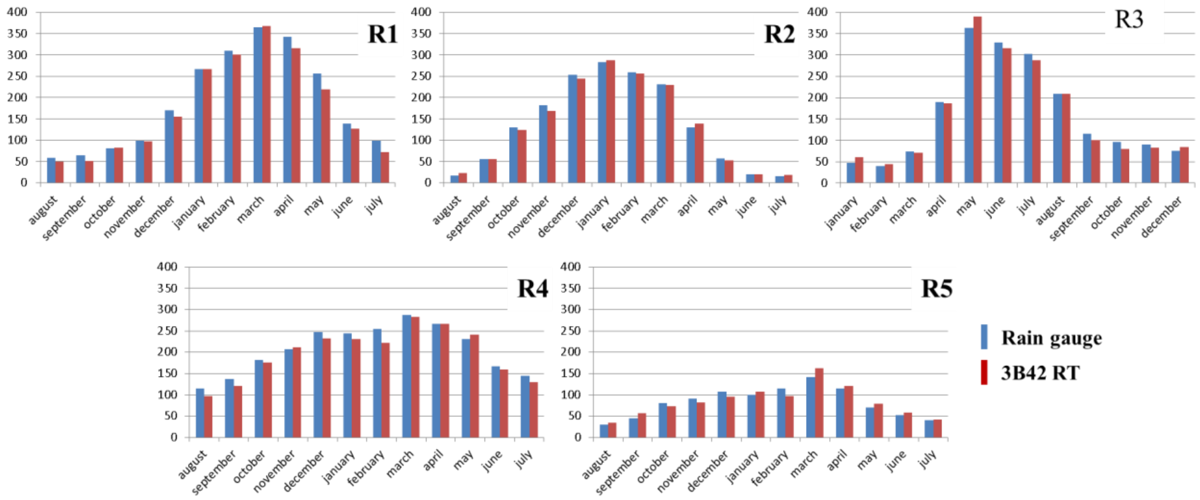

Figure 5. Annual rainfall regimes of the different regions defined in Figure 1 for the AB. The x axis represents each month of the year, y axis represents the rainfall in millimeters. Blue bars represent the observed rainfall and red ones the estimated rainfall.

Furthermore, a contrast between rainy and dry seasons can be noted (Figure 6 a). This could be explained by the type of cloud organization in summer and by the difference of spatial resolution of the two datasets (Salio et al. 2015; Thiemig et al 2012). Indeed, during this period the convection is more local, sparse and the rainy events are rarer than during the rainiest months. Because of its spatial resolution 3B42 RT can be able to detect a rainy event while the rain gauge cannot, but conversely 3B42 RT may not detect an event if the convection is too shallow. Secondly, because of the lower number of rainy events during the driest months the ratio of error may be higher.

The region West (Figure 6 b) presents the worst results, as already shown by figures 3 to 4. In this mountainous region the accuracy of the precipitation estimation tends to decrease with the height, as it has been already shown for example by Thieming et al. (2012) and Zulkafli et al. (2014). Then, with the aim to identify an altitude from which the estimation of precipitation hardly decreases, the correlation between all the statistical results and the altitude has been computed. The results show no relevant relationship and correlation between the altitude and the quality of the estimations of 3B42 RT. These results are consistent with Turko (2014) who has not found a correlation between altitude and the errors of TRMM estimations in the Bolivian Andes.

Another explanation of the poor results in the Andean regions could be the location of the rain gauge itself and more specifically, its exposure. Indeed, in region West, the stations where 3B42 RT overestimates the rainfall are located either at the bottom of a valley or are leeward (Figure 7). One potential interpretation is that when convection and rainy events occur, 3B42 RT is able to detect them because of its spatial resolution, while the rain gauge at the bottom of the valley cannot measure the rain that probably falls only at higher altitude. Similarly, in the case of leeward stations, the rain gauge does not record rainfall while the rainy event is detected by 3b42 RT.

Conversely, such relationship between the location of the station and an underestimation by 3B42 RT is more difficult to observe. Figure 7 b shows a wide range of situations, sometimes opposite, with shelter exposure or rain gauges located in open area. As underestimation occur in a diversity of situations, it is difficult to associate it with specific geographical contexts.

ISBN: 978-85-17-00088-1

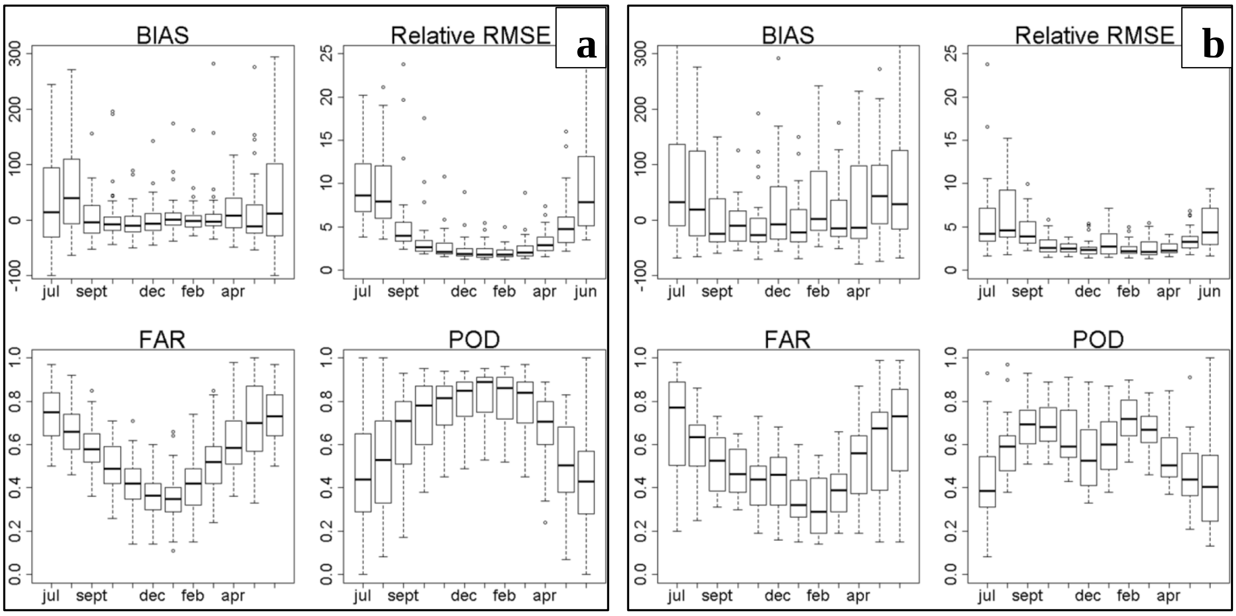

Figure 6. Statistical analysis of the difference at daily scale for each month between 3B42 RT and rain gauges for a) the region South, and b) the region West.

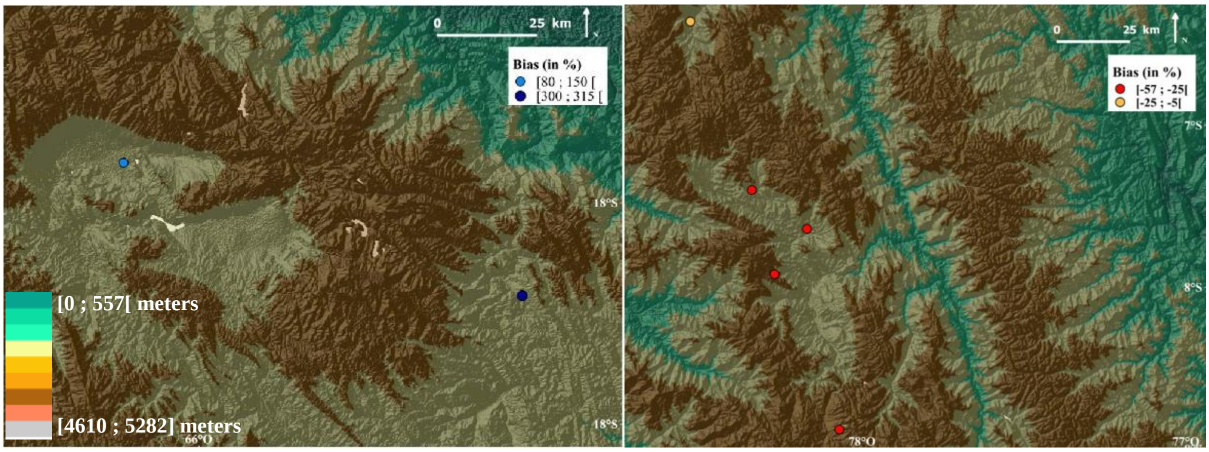

Figure 7. a) location of rain gauges with shelter effect and overestimation by 3B42 RT, b) location of rain gauges with underestimation by 3B42 RT.

4 Conclusions

This work shows that at daily scale, the performance of 3B42 RT is quite good in the lowlands of the AB. The quality of the estimation decreases in the Andean regions but with differences of results within these highlands. The differences between quantitative and categorical statistics showed that 3B42 RT performs better in the detection of rainy events than in the quantitative estimation of rainfall. The ability of 3B42 RT to accurately replicate the rainfall regimes in the different regions of the AB is generally good. But contrasts between months are observed that could be partially explained by the difference of spatial resolution of the two datasets. Furthermore, even if the quality of the estimation decreases in the mountainous regions, neither relevant relationships nor correlation can be noted between the altitude and the statistical results. However, it can be underlined that the location of the rain gauge and its exposure, at the bottom the valley or in a leeward position, may explain the overestimation errors by 3B42 RT.

ISBN: 978-85-17-00088-1

5 References

Delahaye, F. **Analyse comparative des différents produits satellitaires d'estimation des précipitations en Amazonie brésilienne.** 2013. 228p. Thesis (PhD Geography) - Université de Rennes 2 Haute Bretagne, 2013.

Huffman, G.J.; Adler, R.F.; Bolvin, D.T.; Gu, G.; Nelkin, E.J.; Bowman, K.P.; Hong, Y.; Stocker, E.F.; Wolff, D.B. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. **American Meteorological Society**, v. 8, p. 38-55, 2007.

Michot, V.; Dubreuil, V.; Ronchail, J.; Lucio, P. Constitution et analyse critique d'une base de données spatiale pour l'étude des saisons des pluies dans le bassin amazonien. In: Envibras 2014. **Actes du colloque Environnement et géomatique, Approches compares France - Brésil**, 2014, Rennes, p.237-244. Available at: <https://envibras2014.sciencesconf.org/conference/envibras2014/eda_pt.pdf>

Ronchail J.; Cochonneau G.; Molinier M.; Guyot J.L.; Goretti de Miranda Chaves A.; Guimarães V., de Oliveira E. : Rainfall variability in the Amazon Basin and SSTs in the tropical Pacific and Atlantic oceans. **International Journal of Climatology**, v.22, p.1663-1686, 2002.

Salio, P.; Hobouchian, M.P.; García Skabar, Y.; Vila, D. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. **Atmospheric Research**, v. 163, p 146-161, 2015.

Thiemig, V.; Rojas, R.; Zambrano-Bigiarini, M.; Levizzani, V.; De Roo, A. Validation of satellite-based precipitation products over sparsely gauged African River Basins. **Journal of Hydrometeorology**. v. 13, p. 1760-1783, 2012.

Turko, M. **Validation des données TRMM dans la région du Parapeti/Rio Grande (Chaco bolivien, sud-ouest du bassin de l'Amazone).** 2014. 50p. Thesis (Master Sciences de l'Univers, Environnement, Ecologie) – Université Pierre et Marie Curie, Ecole des Mines de Paris et Ecole Nationale du Génie Rural des Eaux et des Forêts, Paris, 2014.

Zulkafli, Z.; Buytaert, W.; Onof, C.; Manz, B.; Tarnavsky, E.; Lavado, W.; Guyot, JL. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean–Amazon River Basins. **Journal of Hydrometeorology**, v. 15, p. 581-592, 2014.