

Combining TRMM and Surface Observation Precipitation: Technique and Validation Over South American

José Roberto Rozante

Center of Weather Forecast and Climate Studies, CPTEC/INPE, Sao Paulo,
Brazil

Demerval Soares Moreira

Center of Weather Forecast and Climate Studies, CPTEC/INPE, Sao Paulo,
Brazil

Luis Gustavo G. de Goncalves

Hydrological Sciences Branch, NASA Goddard Space Flight Center,
Greenbelt, and Earth System Interdisciplinary Center, University of
Maryland, College Park, College Park, Maryland

Daniel A. Vila

Cooperative Institute of Climate Studies, and Earth System Interdisciplinary
Center, University of Maryland, College Park, College Park, Maryland

*Corresponding author address: José Roberto Rozante, Center of Weather Forecast and Climate Studies
CPTEC/INPE, Rodovia Presidente Dutra KM 40, Cachoeira Paulista, S.P., CEP 12 630-000 Brazil.
E-mail: roberto.rozante@cptec.inpe.br*

ABSTRACT

The measure of atmospheric models performance is highly dependent on the quality of the observations used in its evaluation process. In the particular case of operational forecast centers, large scale datasets should be made available in a time manner for continuous assessment of models results. Numerical models and surface observations usually work at distinct spatial scales (i.e. point measurements versus areal average in a regular grid), making direct comparison a hard task. Alternatively, interpolation methods are employed for mapping observational data to regular grids and vice-versa. A new technique (hereafter called MERGE) to combine TRMM satellite precipitation estimates with surface observations over the South American continent is proposed and its performance evaluated for the Summer and Winter trimesters of 2007. Two different approaches for the evaluation of the performance of this product against observations were tested: a cross-validation sub-sampling the entire continent and other sub-sampling only areas with sparse observations. Results show that over areas with high density observations the MERGE technique performance is equivalent to simply average stations over the grid boxes however, over areas with sparse observations MERGE has shown superior results.

1. Introduction

Among the routine activities in operational centers is to valuate numerical climate and weather forecast models in regularly spaced grid-points. Generally, surface observations are considered as “the truth” in such model validation. However, in most cases observations and numerical models output is presented at distinct spatial and temporal scales. Furthermore, the surface network of observations, irregularly spatially distributed, represents environment characteristics at a single point and at its near surroundings. Numerical models output, in other hand, are the result of averaged fields at regularly spaced grid boxes.

Therefore, atmospheric variables, such as precipitation when represented at regularly spaced grid points are valuable information when identifying regions of systematic errors in climate and weather forecast models results. South America represents a challenging region for numerical models precipitation evaluation with sparse and irregular observational network biased toward populated centers near the edge of the continent or along the major river courses. The low density of observations towards the center of the continent and non existent stations over the ocean makes precipitation interpolated to regular grids generally a poor quality product. Therefore, the low confidence in the interpolated observed datasets which in turn can not represent reality compromise the numerical models validation procedure. Remotely sensed estimates of precipitation inferred, for example, from infrared cloud-top temperatures may provide a means of filling the gaps between surface observations in remote regions. However, the calibration and validation of such remotely sensed estimates must be looked with care because ground-based observations are so sparse (de Goncalves et al., 2006).

The use of the precipitation estimation from the Tropical Rainfall Measuring Mission (TRMM) satellite has been extensively employed for numerical models evaluation over the continental South America (Valverde, 2003; de Goncalves, 2006; Rozante and Cavalcanti, 2008). TRMM is a joint project between the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploratory Agency (JAXA) launched on November 1997 with the specific objectives of study and monitoring of the tropical rainfall (Kummerov et al., 2000). Moreover only few merged products between TRMM and observations have been produced so far over South America (Huffman, 2007; Vila et al., 2009) although with good results, not entirely suitable for every day model evaluation at operational centers where information needs to be rapidly available.

Although the TRMM products are highly valuable for numerical models validation, systematic errors are verified, in particular in the Eastern shore of the Northeastern region of Brazil (precipitation is underestimated in that region due to warm clouds as shown by Huffman, 2007; Franchito et al., 2009 and Vila et al., 2009). Also the country borders between Argentina, Paraguay and Brazil (where precipitation is overestimated due to cold top clouds as shown by Rozante and Calcalcanti, 2008).

In order to minimize the interpolation problems (regions with low density observations network) and deficiencies in the TRMM product (underestimate/overestimate) a combination of raingauge datasets (GTS, automatic stations and various agencies in South America) and the real-time TRMM precipitation (3B42RT) is proposed. Thus higher quality gridded datasets are produced for operational numerical models evaluation.

2. Methodology

a. Datasets

The Center for Weather Forecast and Climate Studies (CPTEC from its acronym in Portuguese), a division of the Brazilian National Institute for Space Research (INPE from its acronym in Portuguese) maintain a database of: approximately 1500 surface stations that are regularly reported to the Global Telecommunication System (GTS); automatic surface observations stations (AOS) and observations routinely collected by regional agencies in Brazil. The special distribution for this observational network is very irregular as shown in Figure 1. There are large concentrations of surface stations towards the east portion of the continent over Brazil. Toward the center of the continent the network density decreases drastically including Northern and Southern regions.

The TRMM project produces various products through a combination of different satellite sensors i.e. TRMM Microwave Imager (TMI), Precipitation Radar (PR), and Visible and Infrared Radiometer (VIRS). The 3B42RT product uses microwave precipitation estimates from TMI sensor adjusted by cloud vertical structure obtained from the PR sensor. The A more detailed description for the 3B42RT product can be found in Huffman, 2003. 3B42RT is made available at 0.25 degrees spatial resolution every 3 hours. Over South America (Figure 1) a total of 52.528 grid points will be used to produce the merge precipitation product between 3B42RT and surface observations.

Prior to the merge precipitation product described in this study, the CPTEC/INPE evaluated its operational models using interpolated fields from surface stations only (approximately 1500 reporting stations every day) as shown on Figure 1. The

methodology for combining surface observations with TRMM precipitation estimates is presented next.

b. Merging technique

The technique (hereafter called just MERGE) is described as a sequence of steps as follow: First, TRMM grid boxes where surface observations are present are identified as shown on Figure 2. Secondly, TRMM precipitation is disregarded over those boxes where observations are present as well as surrounding grid boxes as shown in Figure 2b.

A list is then generated with the precipitation amount, geographic location (latitude/longitude) and an assigned station identifier (33333 has been chosen to represent satellite data) for the remaining TRMM grid boxes (see Figure 2b) followed by the surface observations (Table 1).

Surface observations and TRMM precipitation estimates are then interpolated into a regular grid using the Barnes objective analysis (Barnes, 1973).

c. Barnes Objective Analysis

The analysis process for the precipitation interpolation was made using successive corrections of the Barnes scheme (Barnes, 1973; Koch et al., 1983) which requires typically two steps.

The objective analysis scheme is depicted in Figure 3. Weights (W_n) are given to each station within a radius R as function of distance (X_n, Y_n) to a grid point (i, j)

$$W_n = \exp\left(-\frac{d_n^2}{R^2}\right) \quad (1)$$

Where d is the distance between the station and a grid point and R is the radius of influence.

The first guess is then calculated

$$g_1(i, j) = \frac{\sum_{n=1}^N W_n S(x_n, y_n)}{\sum_{n=1}^N W_n} \quad (2)$$

Where $S(x_n, y_n)$ are the observation values within the radius of influence. A correction term is then added to the first step by introducing a convergence parameter (γ) that controls the amount of smoothing determined by

$$g_2(i, j) = g_1(i, j) + \frac{\sum_{n=1}^N W'_n [S(x_n, y_n) - S^1(x_n, y_n)]}{\sum_{n=1}^N W'_n} \quad (3)$$

Where $S^1(x_n, y_n)$ is the value calculated at the observation point n through bilinear interpolation of the four adjacent grid points found in the same iteration. W' is the original weight corrected by γ and is given by

$$W'_n = \exp\left(-\frac{d_n^2}{\gamma R^2}\right) \quad (4)$$

Values for the convergence parameter (λ) can vary between 0 and 1, notwithstanding the values are generally assumed to vary between 0,2 and 0,5 according to many studies (Koch et al., 1983; Mills et al., 1997; Accadia, et al., 2003 and S. K. Sinha et al., 2006). Several tests were performed to determine the optimum range for the γ parameter to be used in this analysis. TRMM values at locations that coincided with the observations locations were interpolated to the TRMM grid points using 9

different gamma values between 0.1 and 1.0. The experiment was repeated for 30 consecutive days and the RMSE calculated for the entire period represented by:

$$\overline{RMSE} = \frac{1}{30} \sum_{d=1}^{30} \sqrt{\frac{1}{N} \sum_{n=1}^N (TRMM_{ori} - TRMM_{int})^2} \quad (5)$$

Where N represents the number of stations, “ori” stands for the initial dataset and “int” is the interpolated dataset through Barnes analysis.

The average RMSE as a function of gamma is shown in Figure 4 where, according to the literature, low errors are found within the range between 0.2 and 0.5. The gamma value that represents the lowest errors is 0.3 therefore, to be applied in this study.

3. Analysis of the Results

a. Cross-validation sub-sampling over the entire domain

This section investigates the potential improvements in the precipitation fields produced by the MERGE methodology when compared to interpolation of surface observations (OBS). Observations were interpolated following Caruso and Quarta (1998) after sub-sampling randomly 10% of the observations over the entire domain. The same procedure for the proposed MERGE (OBS+TRMM) technique was applied (Figure 5). Both fields (OBS90) and MERGE were evaluated at the locations where the 10% of observations were removed.

For this analysis, two trimesters representing austral summer (January, February and March) and winter (June, July and August) were selected. In general the precipitation regime over South America presents high amounts during the summer and relatively less

during the winter. Several precipitation indexes were used to validate the results namely: Equitable Threat Score (ETS – Messinger and Brill, 2004), BIAS, Probability of Detection (POD) and False Alarm Ratio (FAR) in addition to RMSE.

RMSE was computed for pentads (periods of 5 consecutive days) and its temporal evolution is shown on Figures 6a (Summer) and 6b (Winter). It is noticeable that OBS90 and MERGE errors show similar temporal pattern for both trimesters. During the summer (Fig. 6a) when the magnitude of the errors is larger when compared to the winter (Fig. 6b), due to the higher precipitation rates, MERGE show slightly lower errors. However during the winter MERGE errors are smaller than OBS90 in some but also higher in other days.

The quantitative evaluation of the two datasets using the other statistical indexes are shown in Figure 7 respectively: ETS(a), POD(b), BIAS(c) and FAR (d) for the Summer and Winter trimesters. MERGE presents slightly better ETS mainly in the range between light and moderate precipitation (0.254 to 25.4 mm). Moreover, MERGE POD shows better results for all ranges except for intense precipitation (above 38.1 mm) where its performance is similar to OBS90. In other hand, the BIAS score suggests that MERGE has a slight tendency to overestimate precipitation while OBS90 shows amounts closer to the observed in particular above 0.254mm. The FAR results show that MERGE and OBS90 have the same performance when indicating precipitation that did not occur.

We conclude that drawing 10 % of the stations randomly over the entire domain (South America) does not produce significant impacts of using the MERGE technique over the interpolated observation fields (OBS90). Nevertheless, as shown in Figure 1, the irregular spatial distribution of observations is biased towards the continental edges, in

particular over Northeastern Brazil. Despite 10% of the stations drawn were chosen randomly, about 92% of them were located in the high density regions. Consequently combined TRMM estimates with observations were of very little impact when compared to interpolated surface observations.

b. Cross-validation sub-sampling over selected areas

To show how the satellite information improves precipitation analysis over areas with little observations we proposed to remove randomly 10% of the stations over the regions with lower density (longitudes west of 56W).

The temporal evolution of RMSE pentads for the Summer (January, February and March) and Winter (June, July and August) of 2007 are shown in Figures 8a and 8b respectively. During the Summer MERGE presents a significant reduction in the errors when compared to OBS90, in particular for January and March. Furthermore, during the Winter the MERGE combination reduced the errors for the entire period except for a few days in August.

Figure 9 shows the quantitative precipitation indexes for the same periods (Summer and Winter) and domain: ETS(a), POD(b), BIAS(c) and FAR(d). MERGE presents more satisfactory ETS results when compared to OBS90 (Fig. 9a) for all precipitation thresholds. MERGE also has a better performance on detecting precipitation that really occurred as show by the POD (Fig 9b) for all precipitation thresholds. However MERGE slightly overestimates precipitation for all thresholds as shown by the BIAS index (Fig. 9c) while OBS90 has a tendency to underestimation. There is a slightly higher number of false alarms (FAR) when comparing the sub-sampling over a region

with lower density with the entire domain for both MERGE and OBS90 as shown in Figure 9d.

The average precipitation for the Summer trimester over the entire continent is shown in Figure 10 for OBS90 (a) and MERGE (b). It is noticeable that for the regions with high observations density the results are very similar unlike the regions with sparse observational network. MERGE presents higher precipitation amounts over regions of low density. For instance, in central Amazon MERGE precipitation amounts to 10 to 14 mm while OBS90 are less than 4mm. This is in part due to the lack of information for the OBS90 product over that region which causes the interpolated precipitation amounts to be low. Figure 11 shows the average precipitation over the entire continent of the Winter trimester for OBS90 (a) and MERGE (b). As verified during the Summer, MERGE presents higher (and more consistent with expected climatological values) precipitation over regions with sparse observation network.

Another important feature when comparing MERGE with OBS90 is that with TRMM estimates it is possible to extend the precipitation analysis to the adjacent oceans bringing some influence from remote observations over land as shown in Figures 10 and 11.

4. Summary and Conclusions

A combination of TRMM satellite precipitation estimates with surface observations over the South American continent was performed for the Summer and Winter trimesters of 2007 hereafter called MERGE. Two different approaches for the evaluation of the performance of this product against observations were tested: a cross-

validation sub-sampling the entire continent and other sub-sampling only areas with sparse observations. Over regions with high density of surface stations, we found no significant improvements in the MERGE product (where in fact there is little contribution from TRMM) over simply interpolating the existing observations (namely OBS90). Nevertheless, the results analyses over low density regions (west of 56W) show substantial (positive) impact of the MERGE product when compared against OBS90. Despite the MERGE expected lower false alarm ratio (FAR) over the regions with lower observations density (also found in the OBS90), MERGE has been proved to be a valuable analysis in a regular grid for model results evaluation.

In summary, this study emphasizes two important aspects that need attention on precipitation evaluation of atmospheric models, in particular at operational centers such as CPTEC/INPE: first, quick and efficient techniques to blend satellite and surface observations are needed for model evaluation to become practical at the daily routine of operational centers as the MERGE methodology presented in this paper; secondly, model results evaluation over large areas have to be performed with care, in particular, over regions with irregular and sparse observational network such as South America. Regions with different observation density should be evaluated differently.

Nonetheless, this study reinforces the significance of remote sensed precipitation estimates, in particular TRMM products, over regions with sparse ground information, showing its superiority to surface observations alone.

REFERENCES

Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2003: Sensitivity of Precipitation Forecast Skill Scores to Bilinear Interpolation and a Simple Nearest-Neighbor Average Method on High-Resolution Verification Grids. *Wea. Forecasting*, **18**, 918–932.

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations, NOAA Tech. Memo. ERL NSSL-62 *National Severe Storms Laboratory*, Norman, OK 73069, 60 pp. [NTIS COM-73-10781.], 1973. 2359

C. Caruso and F. Quarta, Interpolation methods comparison, *Computers and Mathematics with Applications*, **35** (12) (1998), pp. 109–126.

de Goncalves, L. G. G., W. J. Shuttleworth, B. Nijssen, E. J. Burke, J. A. Marengo, S. C. Chou, P. Houser, and D. L. Toll, 2006: Evaluation of model-derived and remotely sensed precipitation products for continental South America, *J. Geophys. Res.*, **111**, D16113, doi:10.1029/2005JD006276

Franchito, S. H., V. B. Rao, A. C. Vasques, C. M. E. Santo, and J. C. Conforte (2009), Validation of TRMM precipitation radar monthly rainfall estimates over Brazil, *J. Geophys. Res.*, **114**, D02105, doi:10.1029/2007JD009580.

Huffman, G.J., R.F. Adler, E.F. Stocker, D.T. Bolvin, and E.J. Nelkin, 2003: Analysis of TRMM 3-Hourly Multi-Satellite Precipitation Estimates Computed in Both Real and Post-Real Time. Combined Preprints CD-ROM, 83rd *AMS Annual Meeting*, Poster

P4.11 in: 12th Conf. on Sat. Meteor. and Oceanog., 9-13 February 2003, Long Beach, CA, 6 pp.

Huffman G. J., Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55

Koch S. E., M. desJardins, and P. J. Kocin, 1983: An interactive Barnes objective map analysis scheme for use with satellite and conventional data. *J. Climate Appl. Meteor.*, **22**, 1487–1503.

Kummerow, C., J. Simpson, O. Thiele, W. Barnes, A.T.C. Chang, E. Stocker, R.F. Adler, A. Hou, R. Kakar, F. Wentz, P. Ashcroft, T. Kozu, Y. Hong, K. Okamoto, T. Iguchi, H. Kuroiwa, E. Im, Z. Haddad, G. Huffman, B. Ferrier, W.S. Olson, E. Zipser, E.A. Smith, T.T. Wilheit, G. North, T. Krishnamurti, and K. Nakamura, 2000: The Status of the Tropical Rainfall Measuring Mission (TRMM) after Two Years in Orbit. *J. Appl. Meteor.*, **39**, 1965–1982.

Mesinger, F.; K. Brill., 2004: Bias normalized precipitation scores. Preprints, *17th Conf. on Probability and Statistics*, Seattle, WA, Amer. Meteor. Soc., CD-ROM, J12.6, 2004.

Mills, G. A., Weymouth, G., Jones, D., Ebert, E. E., Manton, M., Lorkin, J. and Kelly, J., 1997: A National objective daily rainfall analysis system. *BMRC Techniques Development Report*, **1**, 30 pp

- Rozante, J. R., and I. F. A. Cavalcanti, 2008: Regional Eta model experiments: SALLJEX and MCS development, *J. Geophys. Res.*, **113**, D17106, doi:10.1029/2007JD009566.
- Sinha, S. K., Narkhedkar, S. G. and Mitra, A. K., 2006: Barnes objective analysis scheme of daily rainfall over Maharashtra (India) on a mesoscale grid, *Atmosfera*, **19**, 59–76.
- Valverde Ramírez, M. C., 2003: Use of Neural Network for Analysis and Forecast over the South and Southeast Brazil. PhD dissertation, Instituto Nacional de Pesquisas Espaciais (INPE), 238 pp. [Available from INPE-13950-TDI/1061 <<http://urlib.net/sid.inpe.br/jeferson/2004/02.16.14.01>>].
- Vila, D.A., L.G.G. de Goncalves, D.L. Toll, and J.R. Rozante, 2009: Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. *J. Hydrometeor.*, **10**, 533–543.

List of Tables

Table 1: 24h accumulated TRMM precipitation combined with observations. Table shows a section of the list of stations ID created after the precipitation merge as example.

List of Figures

Figure 1: Spatial distribution of raingauges from different sources (i.e. GTS, regional and local agencies and automated stations) over South America. Source data: INMET, CPTEC/INPE and regional centers of meteorology.

Figure 2: Example of TRMM 24h accumulated precipitation estimation (shaded) March, 8th, 2003. The white dots on the left panel (a) represent the reporting raingauges on that particular day. White squares on the right panel (b) represent the 0.25° TRMM grid boxes near the raingauges shown in the left panel.

Figure 3: Section of the domain where the Barnes Objective Analysis is applied illustrating how the proposed method works.

Figure 4: Root mean square error calculated at different values of the gamma parameter.

Figure 5: Schematic diagram representing the evaluation methodology applied in this study.

Figure 6: Time-series of RMSE pentads calculated for the following periods of 2007: (a) summer trimester and (b) winter trimester. Calculations were performed considering 10% random sub-setting over the entire domain.

Figure 7: Statistical indexes (ETS(a), POD(b), BIAS(c) e FAR(d)) computed for the entire period of study (summer and winter of 2007) after randomly removing of 10% of the observations over the whole domain.

Figure 8: Temporal evolution of the RMSE pentads for: summer (a) and winter (b) trimesters of 2007 after randomly removing 10% of the observations over the lower density areas of the domain (west of 56°W).

Figure 9: Statistical indexes (ETS(a), POD(b), BIAS(c) e FAR(d)) computed for the entire period of study (summer and winter of 2007) after randomly removing of 10% of the observations over the lower density areas of the domain (west of 56°W).

Figure 10: Precipitation averaged over the summer trimester computed using OBS90 (a) and MERGE (b).

Figure 11: Precipitation averaged over the winter trimester computed using OBS90 (a) and MERGE (b).

Table 1: 24h accumulated TRMM precipitation combined with observations. Table shows a section of the list of stations ID created after the precipitation merge as example.

	ID	latitude	longitude	prec.	
T R M M	33333	-49.875	-82.875	3.3	M E R G E
	33333	-49.625	-82.875	22.1	
	
	
O B S	83849	-25.513	-49.171	32.0	
	83780	-23.615	-46.667	14.0	
	
	

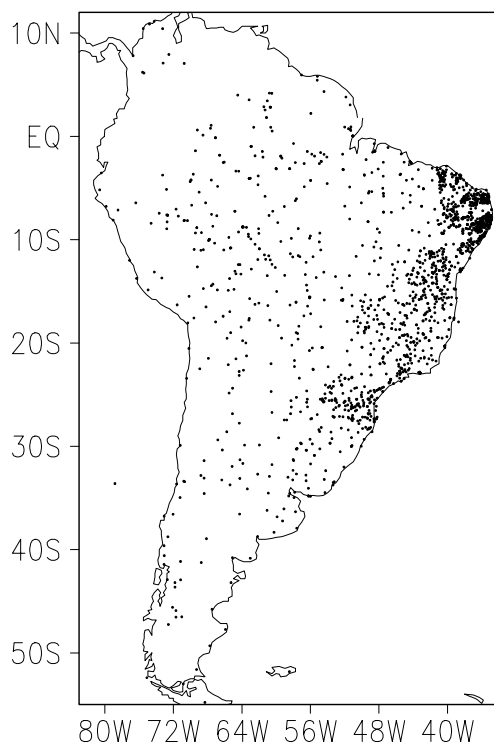


Figure 1: Spatial distribution of raingauges from different sources (i.e. GTS, regional and local agencies and automated stations) over South America. Source data: INMET,CPTEC/INPE and regional centers of meteorology.

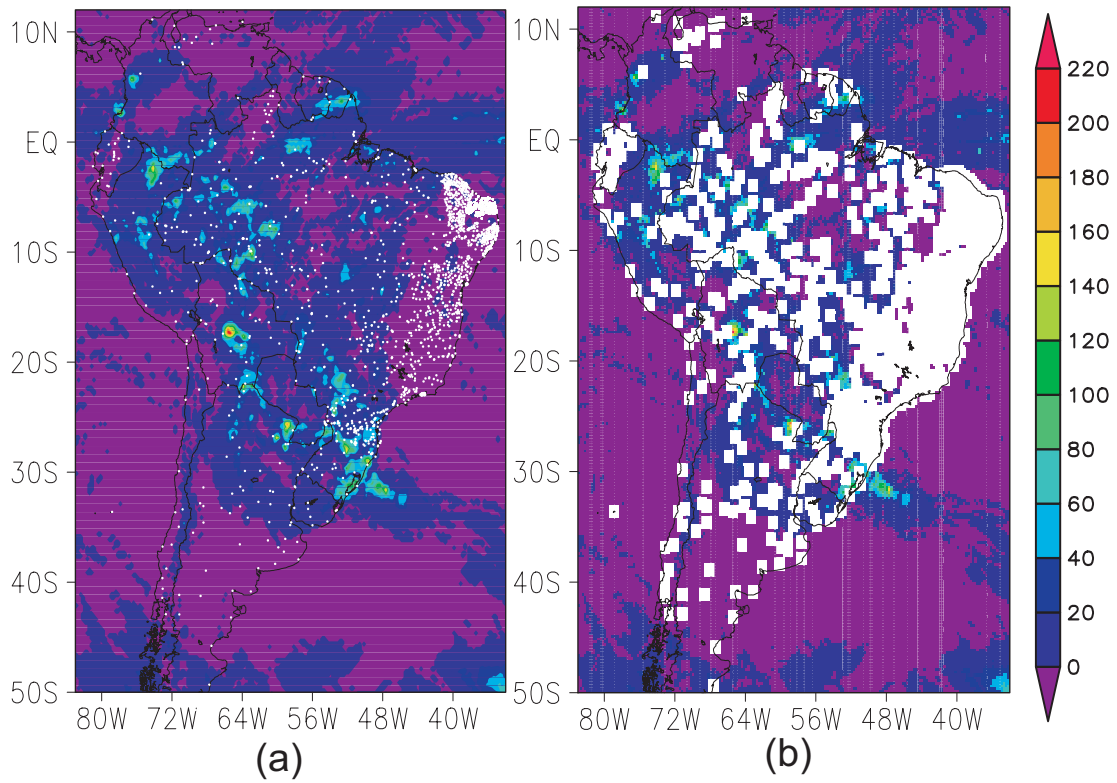


Figure 2: Example of TRMM 24h accumulated precipitation estimation (shaded) March, 8th, 2003. The white dots on the left panel (a) represent the reporting raingauges on that particular day. White squares on the right panel (b) represent the 0.25° TRMM grid boxes near the raingauges shown in the left panel.

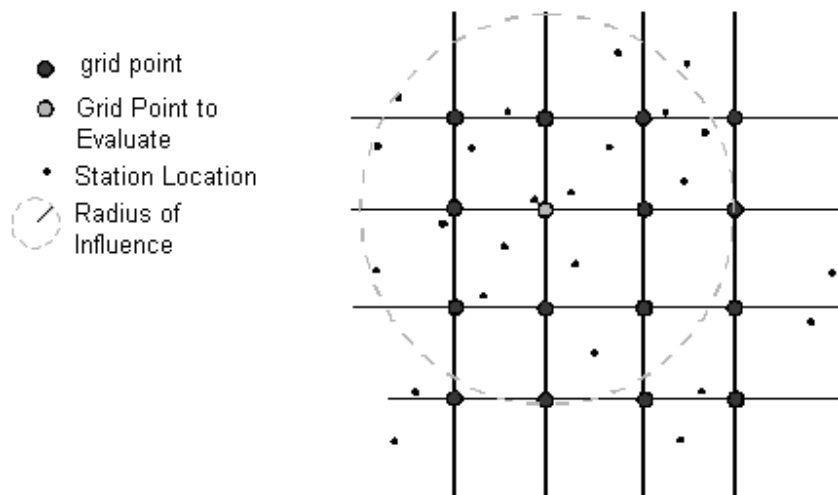


Figure 3: Section of the domain where the Barnes Objective Analysis is applied illustrating the different elements used in the method (GEMPAK Online Tutorial - <http://www.unidata.ucar.edu/software/gempak/tutorial/barnes.html>).

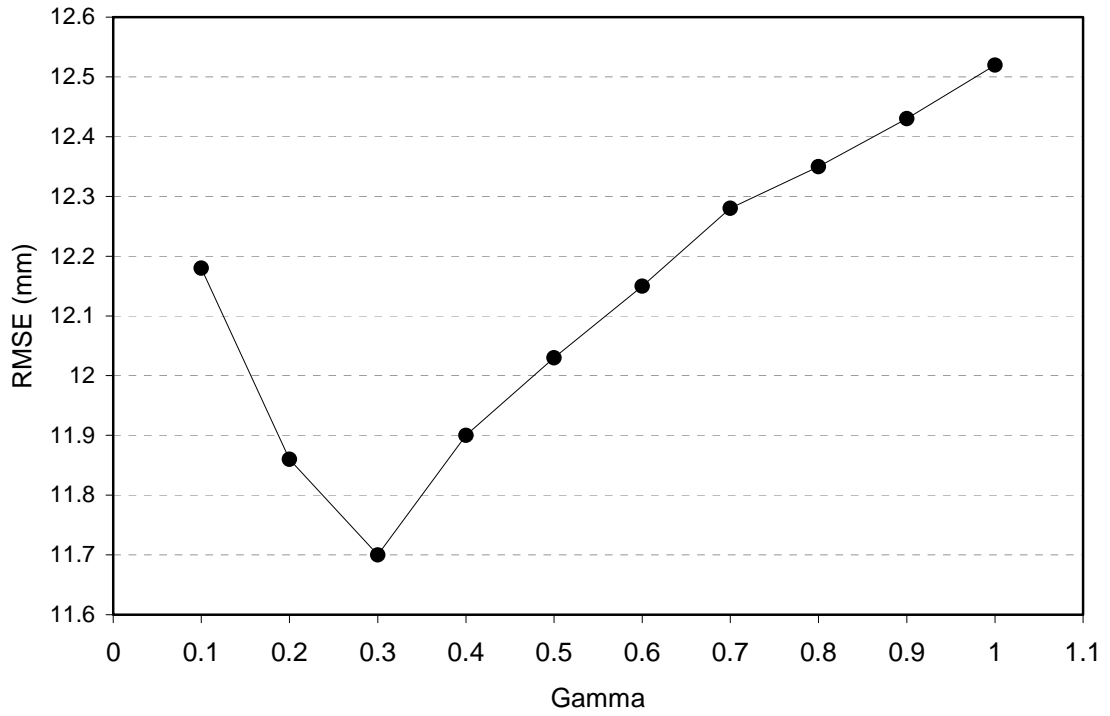


Figure 4: Root mean square error calculated at different values of the gamma parameter.

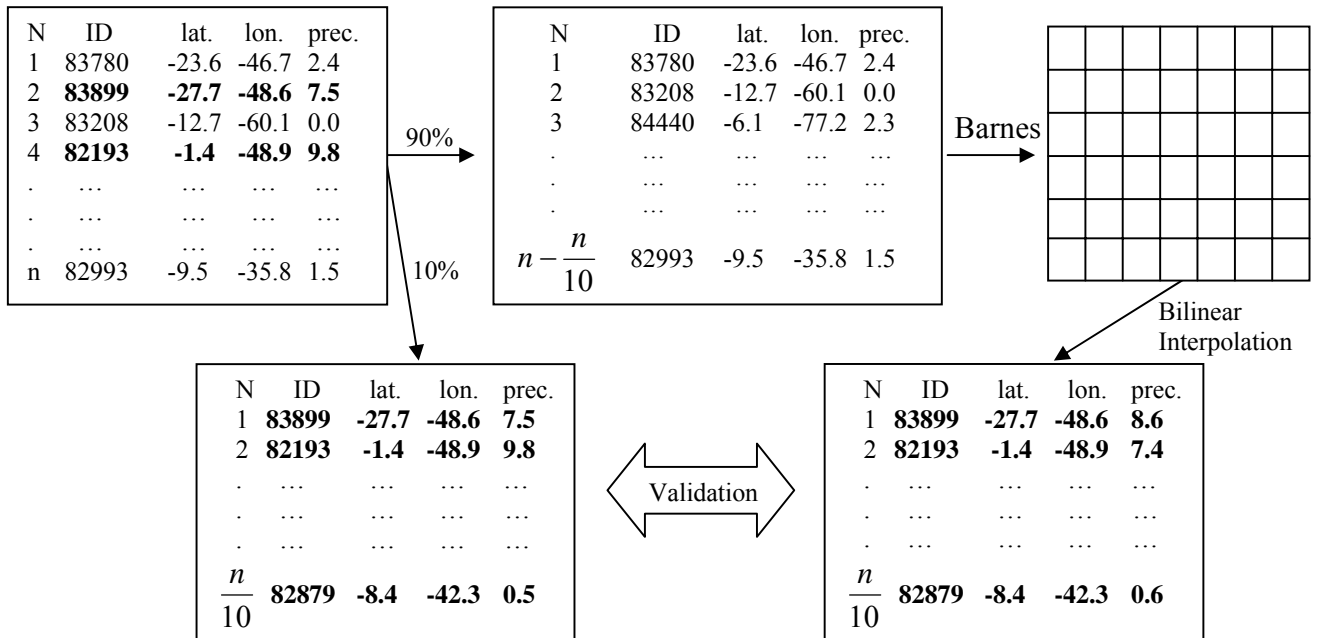


Figure 5: Schematic diagram representing the evaluation methodology applied in this study.

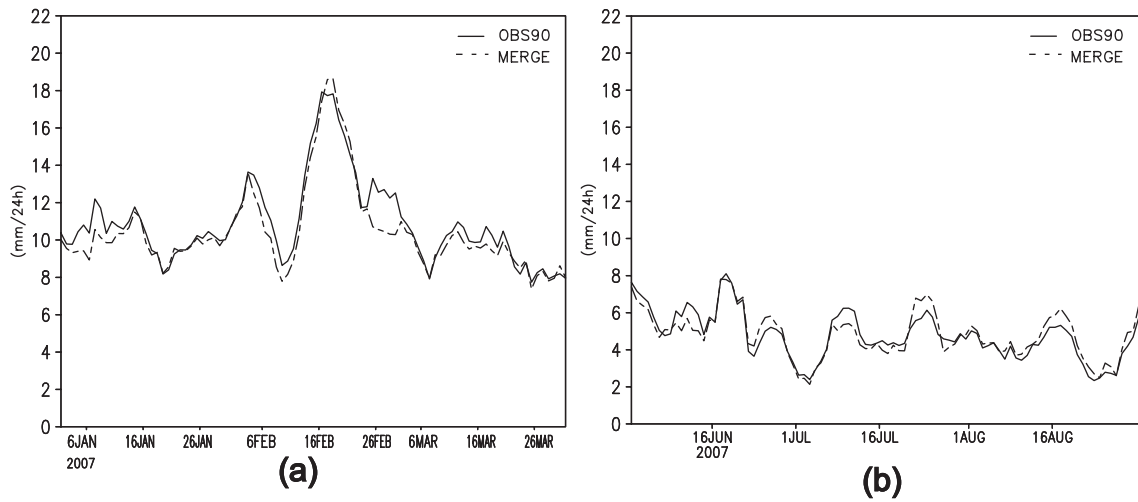


Figure 6: Time-series of RMSE pentads calculated for the following periods of 2007: (a) summer trimester and (b) winter trimester. Calculations were performed considering 10% random sub-setting over the entire domain.

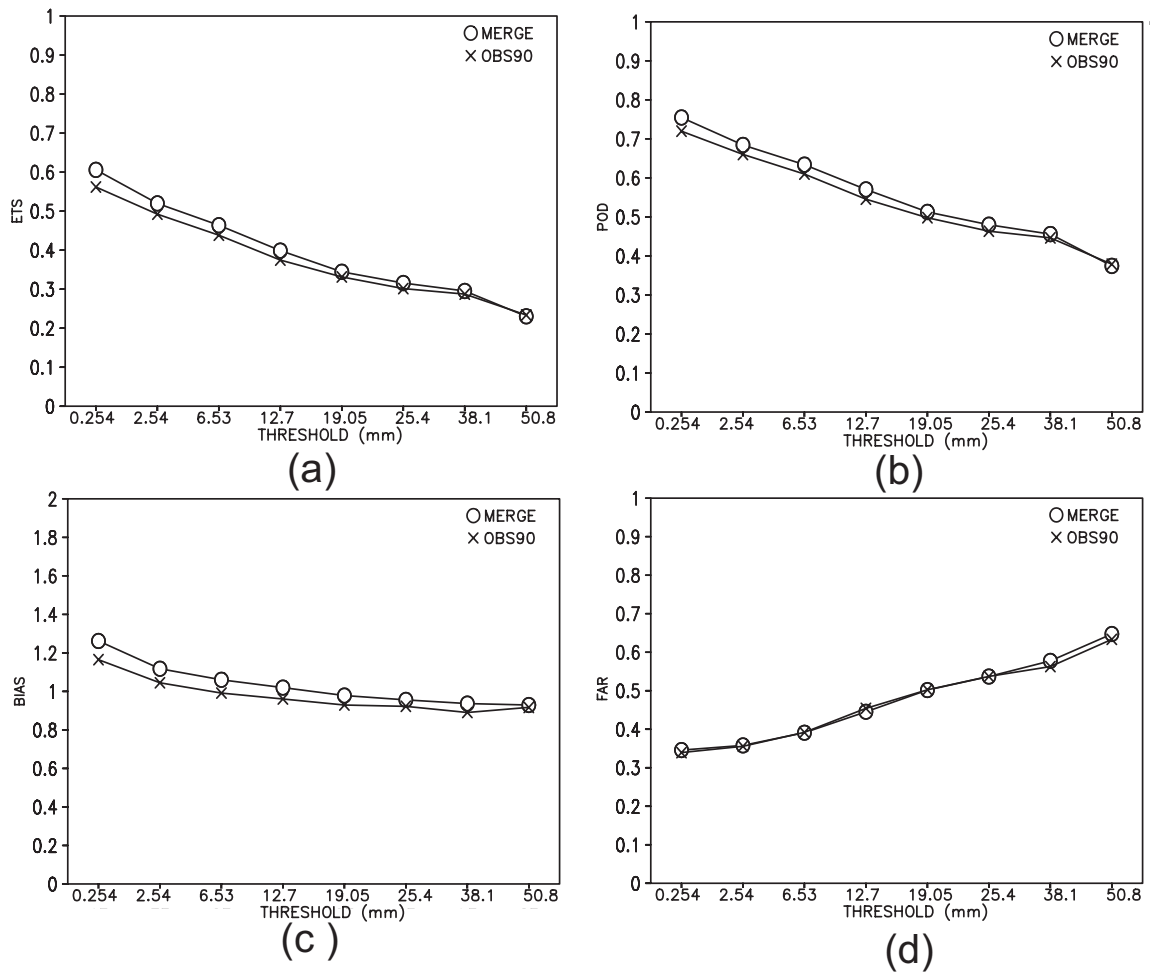


Figure 7: Statistical indexes (ETS(a), POD(b), BIAS(c) e FAR(d)) computed for the entire period of study (summer and winter of 2007) after randomly removing of 10% of the observations over the whole domain.

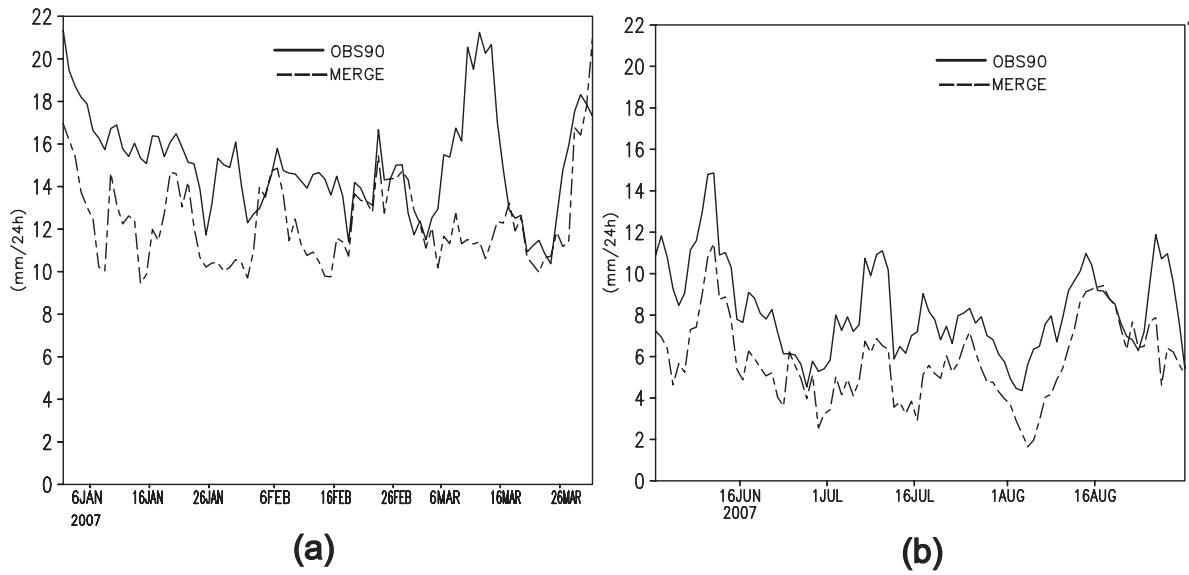


Figure 8: Temporal evolution of the RMSE pentads for: summer (a) and winter (b) trimesters of 2007 after randomly removing 10% of the observations over the lower density areas of the domain (west of 56°W).

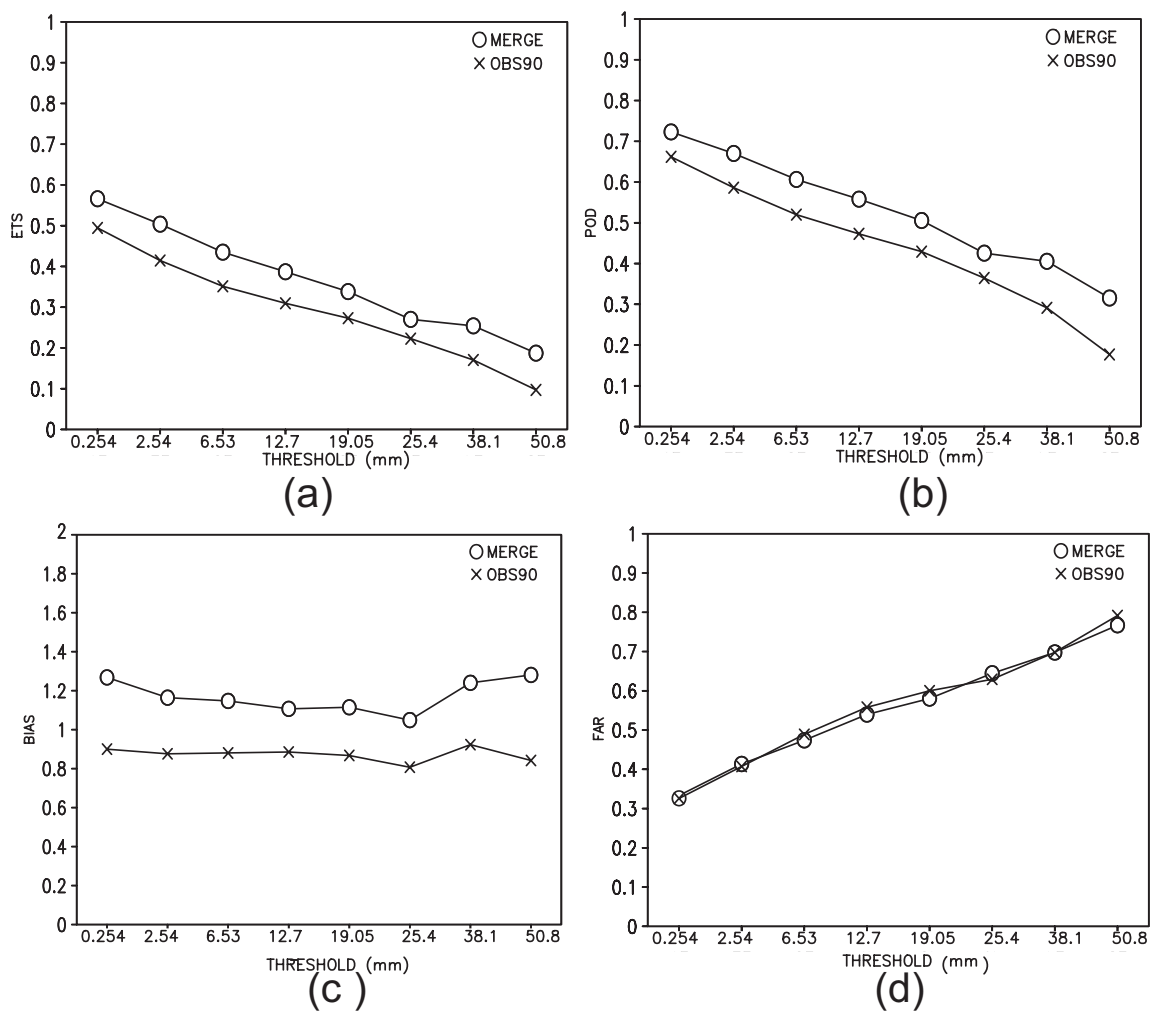


Figure 9: Statistical indexes (ETS(a), POD(b), BIAS(c) e FAR(d)) computed for the entire period of study (summer and winter of 2007) after randomly removing of 10% of the observations over the lower density areas of the domain (west of 56°W).

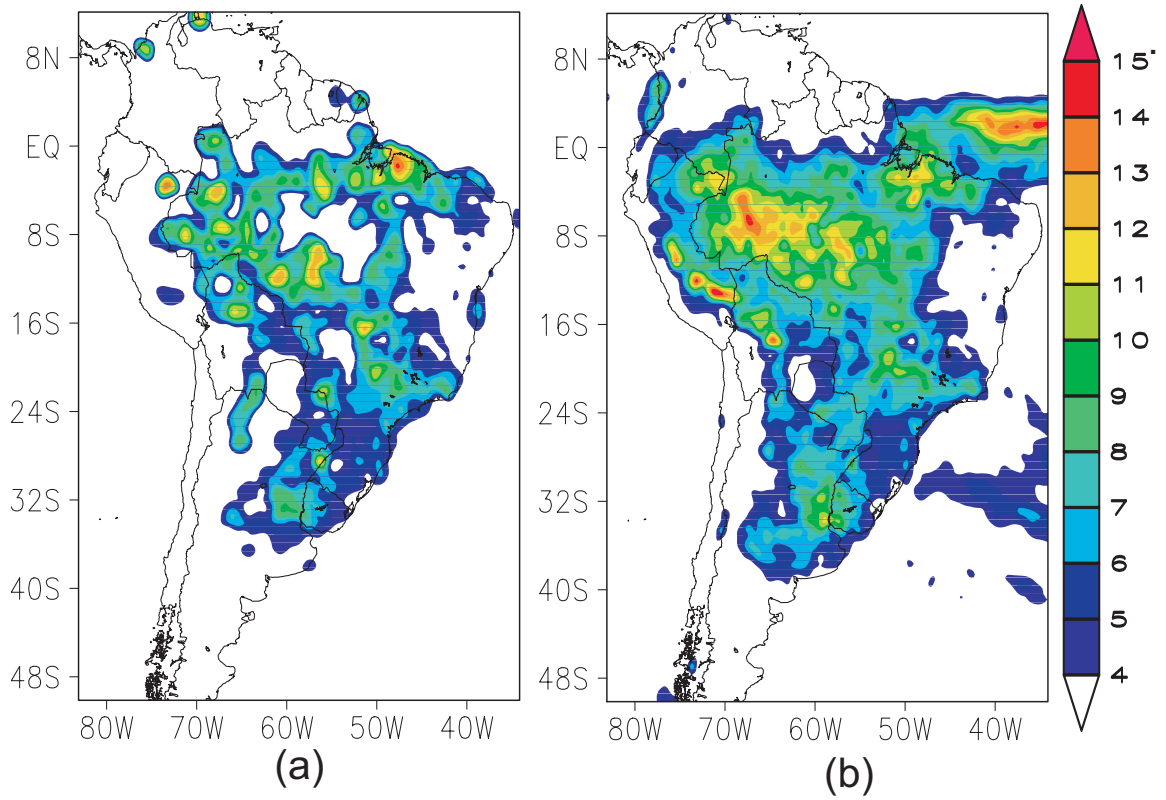


Figure 10: Precipitation averaged over the summer trimester computed using OBS90 (a) and MERGE (b).

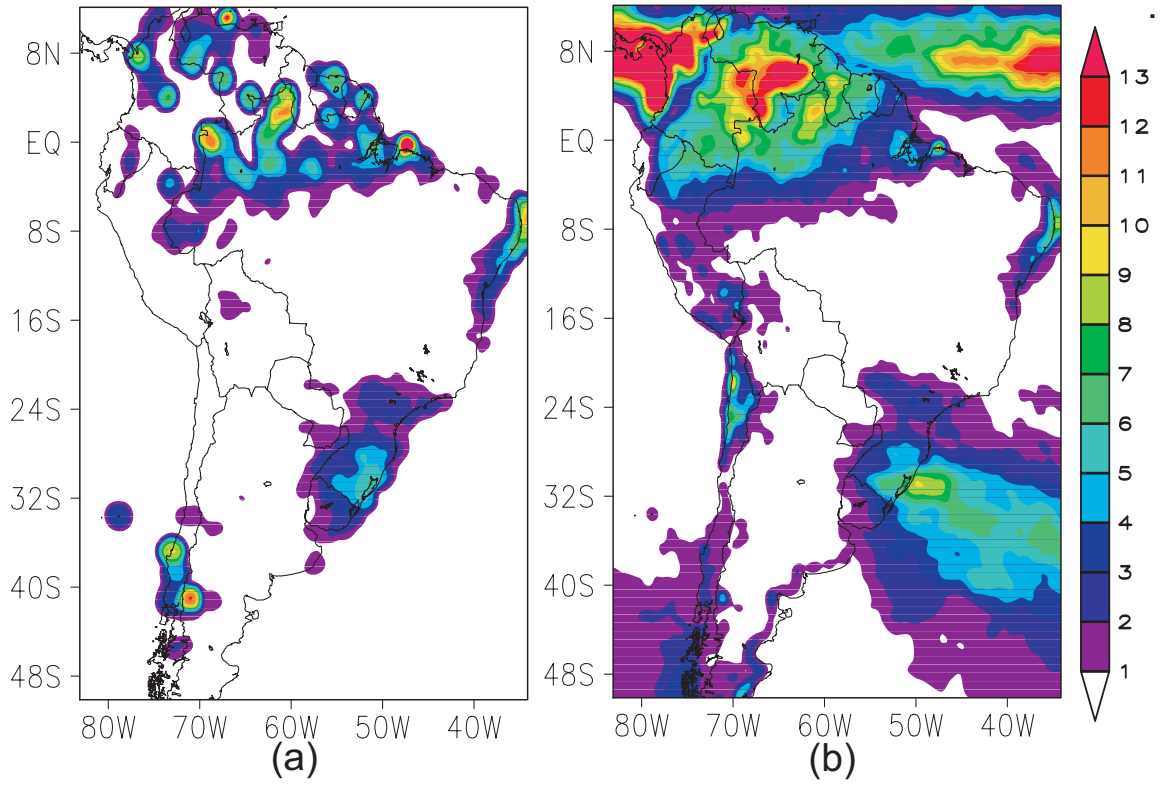


Figure 11: Precipitation averaged over the winter trimester computed using OBS90 (a) and MERGE (b).