Daily to Sub-daily precipitation downscaling based on multiple datasets using artificial neural networks in Brazil

Rogério Batista[1] (rogerio.batista@inpe.br) and Alan Calcheiros[1] (alan.calcheiros@inpe.br)
[1] National Institute of Space Research (INPE), Brazil

Abstract

The goal of this work is to approach the efficiency of using Artificial Neural Networks (ANN) in the processing of meteorological data. The main focus of this is to evaluate the results of an supervisioned ANN applied to downscaling daily rainfall data. In order to better represent the characteristics of the diurnal cycle and the physical processes of the different regions of Brazil we applied two different types of ANN, the Deep Neural Network (DNN) and the Recurrent Neural Network (RNN).

Data and Methods

The main information from daily retrievals comes from a satellites-based technique corrected by rain gauges, called MERGE which was developed by INPE in Brazil. The target is a sub-daily rainfall with temporal resolution of 3 hours. Meteorological variables with physical relationship with the rain in previous studies were selected as input, like infrared brightness temperature from GOES satellite, hourly precipitation estimates from microwave sensors (IMERG), and environmental data (e.g. humidity, wind, etc) from ERA5 reanalysis. Each of the chosen variables was pre-processed, producing averages (or accumulated) values and other 3-hour temporal resolution measurements. Table 1 shows all of the inputs and its characteristics.

Input variable	Dataset	Source	Spatial size	Values
Daily precipitation calibrated (mm)	MERGE	INPE	0.10°	1
10 years mean precipitation calibrated (mm)	MERGE	INPE	0.10°	8
Sub-daily precipitation estimated (mm)	IMERG	NASA	0.10°	8
Average infrared brightness temperature (K)	MERGIR	NASA	0.036°	8
Variance infrared brightness temperature (K)	MERGIR	NASA	0.036°	8
Average relative humidity 1000-500 mb (%)	ERA5	ECMWF	0.25°	8
Wind direction (degress)	ERA5	ECMWF	0.25°	8
Wing magnitude (m/s)	ERA5	ECMWF	0.25°	8
Total column water (kg/m²)	ERA5	ECMWF	0.25°	8

Table 1: Details of input variables

Observed data from 610 rain gauges were used, of which approximately 10% were reserved for ANN validation in places where data were not used in the ANN training, as well as all data from the year 2020 were also reserved for validation of results in a period that was not used in the learning process. Only the geographic points where daily rain was observed were used in the ANN training.

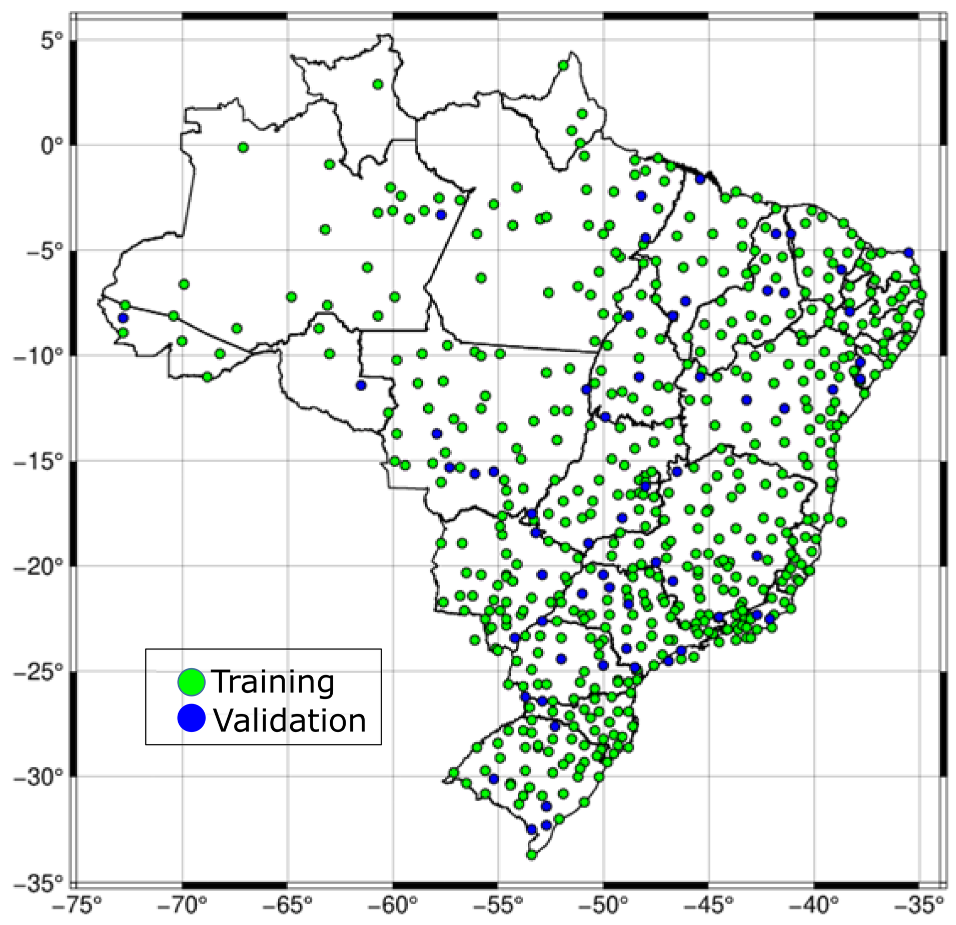

Figure 1: Distribution of observed data

The values were organized in daily series of eight values accumulated every 3 hours, totaling 691,753 daily series in the set used for ANN training and testing (90.6%) and 71,474 in the validation set (9.4%) for the period from 2000 to 2019. The data set for the year 2020, reserved for temporal validation, totaled another 50,576 daily records.

Hyperparameters and training

A deep learning was used for both DNN and RNN, the learning rate was 0.001 and the size of minibatches was 512-elements. For performance metric the mean square error (MSE) was used and to optimation ADAM was the gradient of the cost function. The Sigmoid function was applied between the intermediate layers and the Relu function was applied in the output layer.

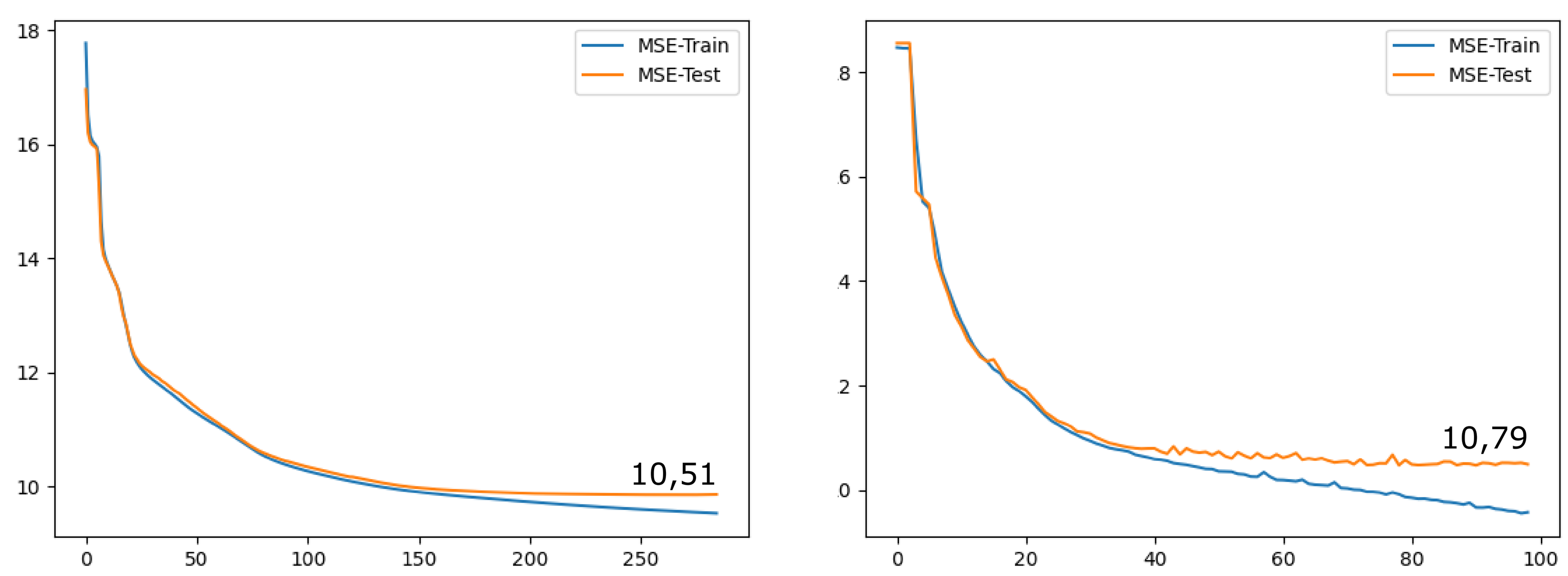

Figure 2: DNN (left) and RNN (rigth) training metrics

Results

Results obtained by the both ANN are in a better agreement when compared to IMERG product (the reference). Over all validation set the DNN obtained an MSE of 11.09 mm and RNN shows a value of 11.88 mm. Also, DNN shows better results for all the different regions of Brazil as well as for the different seasons, Figure 3 and 4.

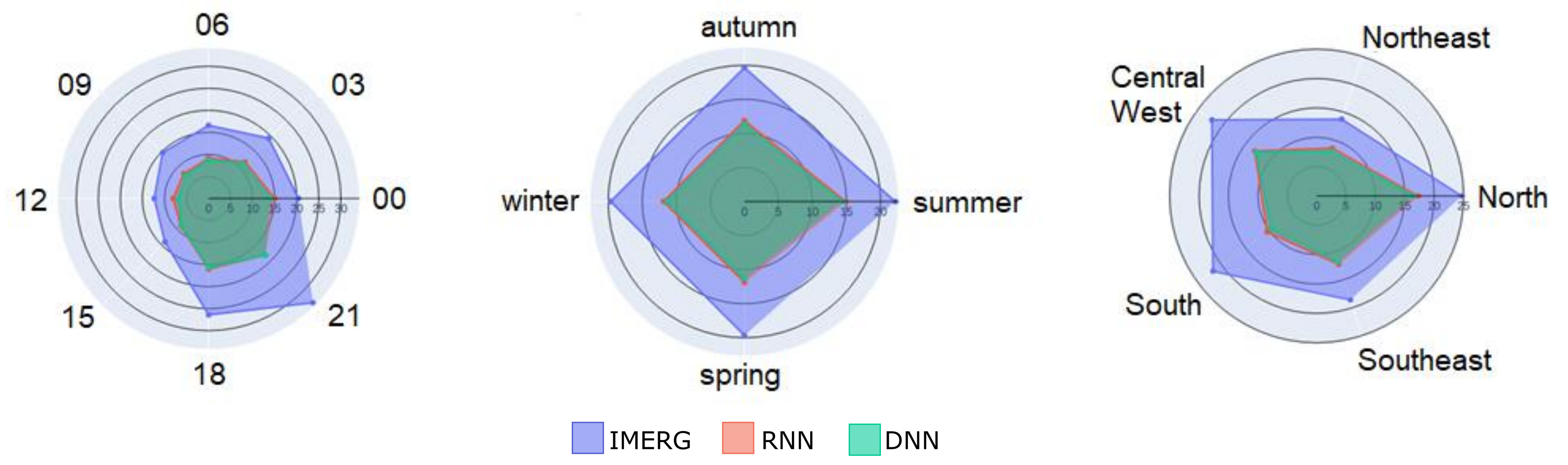

Figure 3: MSE over validation set by time, season and region

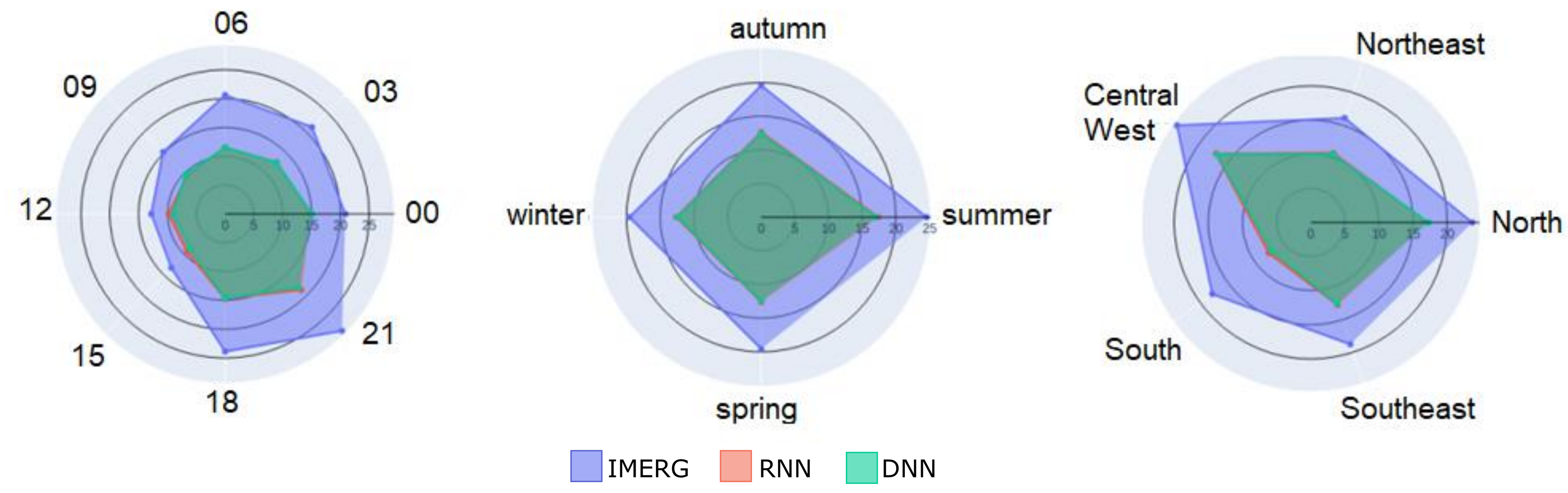

Figure 4: MSE over 2020 set by time, season and region

However, the rain screening (areas with rainfall) is slightly better for IMERG, but with a superestimation of the precipitation.

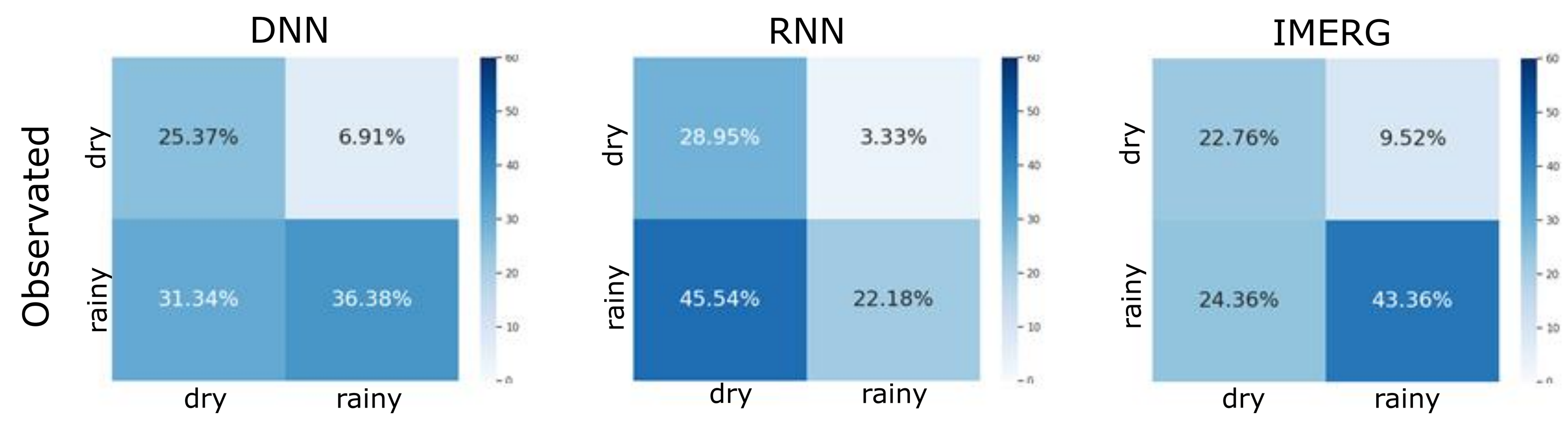

Figure 5: Cross validation of the technics in validation set

BIAS for RNN is better for hours with low precipitation, while DNN and IMERG are better for rainy periods (18 and 21 GMT). However, BIAS differences between DNN and RNN are very small and MSE shows a slightly better values to DNN for all times.

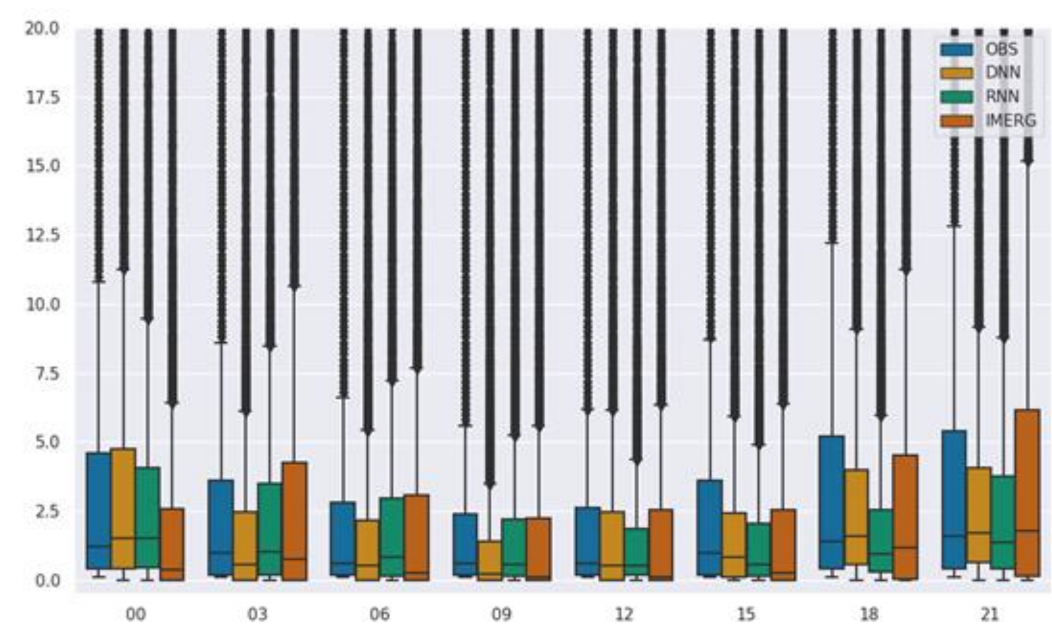

Figure 6: Statistical comparison of the differents technics

Final considerations

DNN was chosen better than RNN for the proposed. In Figure 7 we can see the moving average of the network on the validation set and compared to the IMERG we can observe a better performance. ANN have proved to be a promising alternative to the downscaling of daily precipitation which is computationally less complex than the execution of dynamic models, however more studies and analyses are needed.

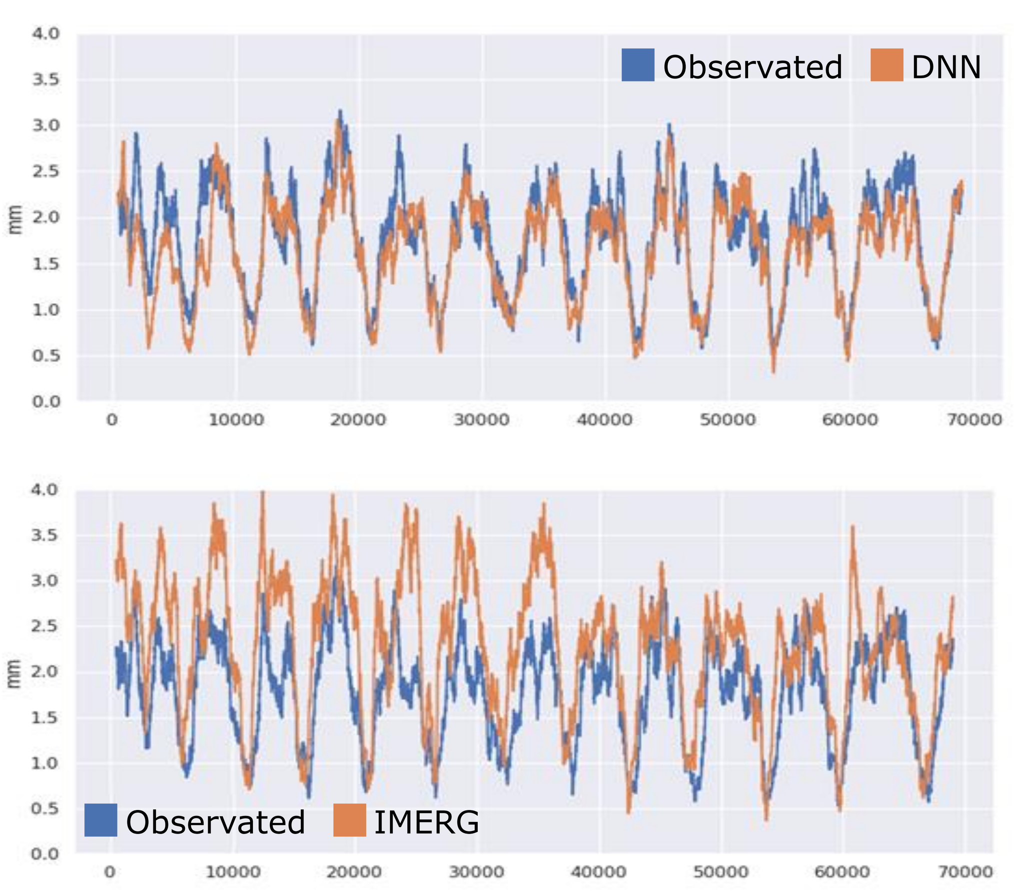

Figure 7: Moving average of technics (500 values)

References

KUMAR, J.; BROOKS, B.-G. J.; THORNTON, P. E.; DIETZE, M. C. Sub-daily Statistical Downscaling of Meteorological Variables Using Neural Networks. Procedia computer science, v. 9, p. 887–896, 2012.

KHAN, M. S.; COULIBALY, P.; DIBIKE, Y. Uncertainty analysis of statistical downscaling methods. Journal of Hydrology, 2006.

ROZANTE, J. R.; MOREIRA, D. S.; DE GONCALVES, L. G. G.; VILA, D. A. Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America. Weather and Forecasting, v. 25, n. 3, p. 885–894, 2010. American Meteorological Society.

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan (2019), GPM IMERG Early Precipitation L3 1 day 0.1 degree x 0.1 degree V06, Edited by Andrey Savtchenko, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC)