Application of the DBSCAN algorithm for identifying morphological features of atmospheric systems over the amazon basin

LEAL NETO, H. B.[1]; CALHEIROS, A. J.[1]; Applied Computing – National Institute for Space Research (INPE)

Email: {helvecio.neto, alan.calheiros}@inpe.br

INTRODUCTION

In the Amazon region precipitation-generating and atmospheric process are modulated by different scales (micro, meso, and synoptic). Studying the morphological characteristics (size) of the atmospheric events is fundamental to understand the dynamics of meteorological systems. The thermodynamic processes that lead to the development of convective systems (CS) over the Amazon depend on several local factors. The life cycle patterns of CSs have a strong relationship with the diurnal cycle through the radiative heating that occurs in the region [1]. Studying the morphological characteristics of CSs are important because they allow estimating the structures and intensity of atmospheric phenomena occurring in the Amazon basin region.

DATA AND STUDY AREA

The data used in this work are available in public repositories. In the case of satellite, the data is available on the FTP servers of INPE. For radar, the data is available on the ARM (Atmospheric Radiation Measurement) platform. The radar data are part of a compilation related to the 2014 campaign of the GoAmazon project, which among its experiments, used several instruments to investigate the interaction between forest and atmosphere in the Amazon basin region.

From the satellite, the infrared channel 4 (10.2 µm) will be used with an overlay over the whole Amazon basin region (Figure 1a). For the Radar the CAPPI (Constant Altitute Plan Position Indicator) of 3km was used (Figure 1b).

METHODOLOGY

DBSCAN is an unsupervised methods for identifying groups present in multivariate data [2]. This clustering technique does not require predefined classes. The main purpose of DBSCAN is find groups based on distance metrics between individuals points present in an information density. The applied thresholds were: 20, 30 and 40 dBZ (decibels at Z-level) for RADAR and 235, 220 and 210 K (*Kelvin*) for GOES-13 Satellite. These values represent different intensity features in the reflectivity and brightness temperature measurements. The coordinate locations of the points based on the intensity thresholds was the input for the DBSCAN training process. Such information is passed on to the algorithm, and with that, clusters are classified based on the distribution of coordinate points.

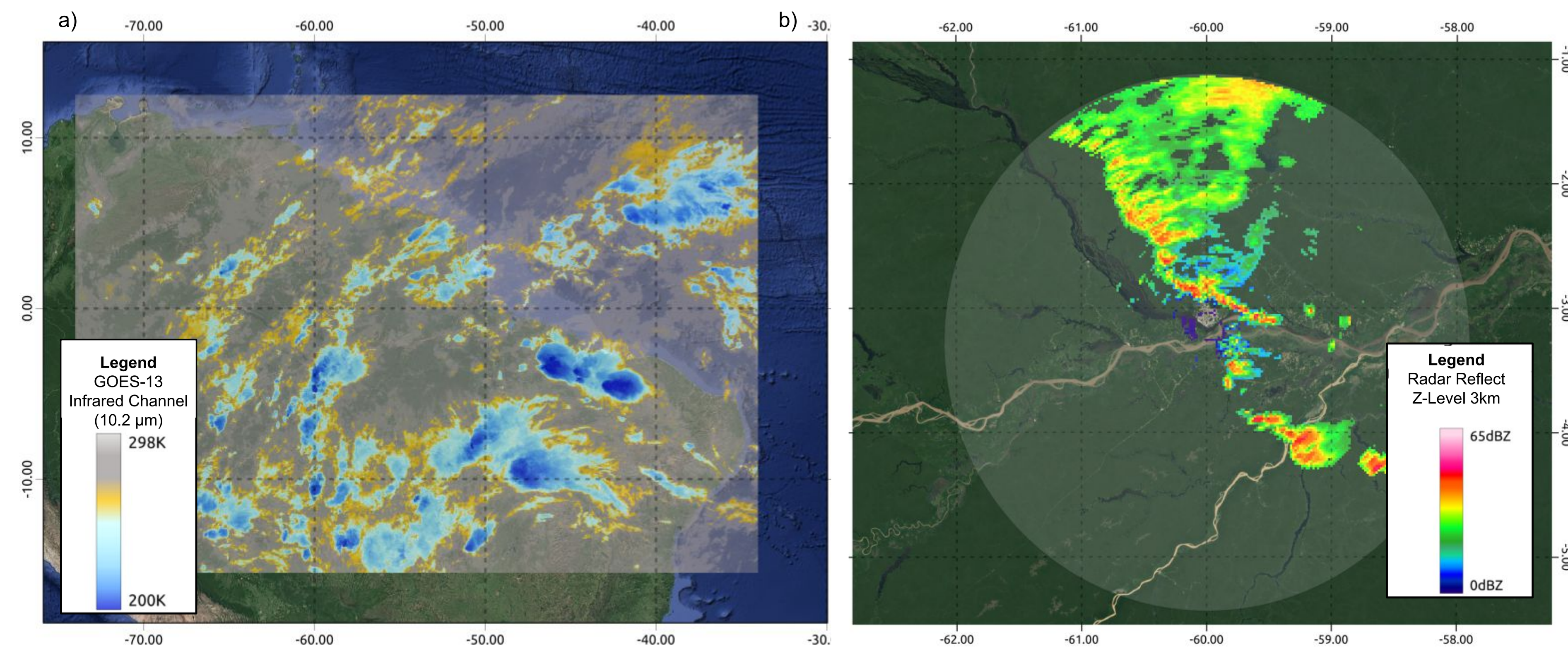

Figure 1. Study area at the Amazon basin. (a) GOES-13 satellite data with the infrared channel and (b) Radar volumetric scanning at the CAPPI level of 3km.

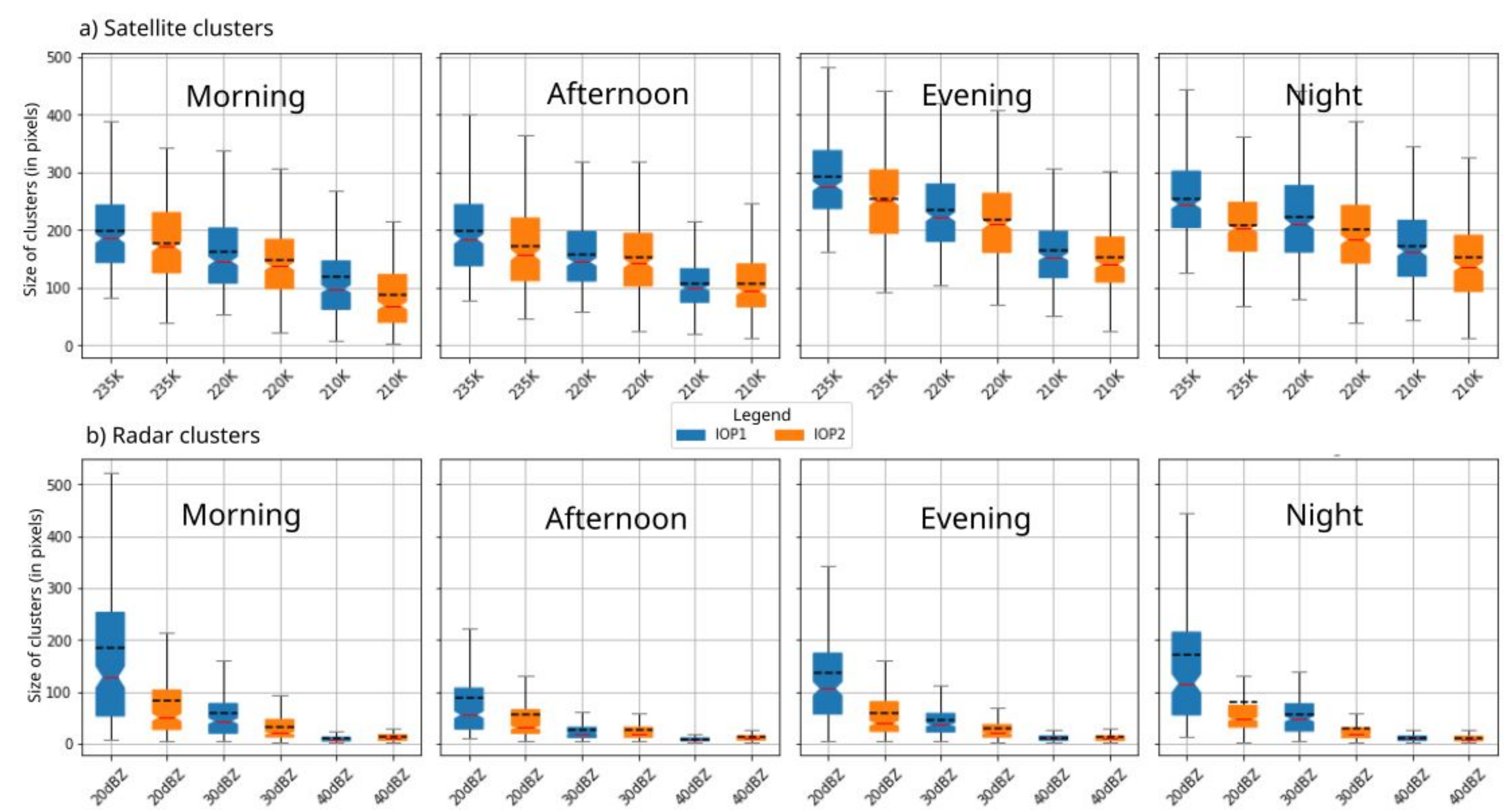

Figure 2. Distribution of the cluster size in the daytime and nighttime periods during the IOPs. (a) Satellite clusters size and (b) Radar clusters both, for different thresholds.

RESULTS

To characterize the morphology of the clusters found by DBSCAN, we subdivided into groups (Morning, Afternoon, Evening and Night) the clusters referring to the thresholds previously mentioned. Figure 2 shows the distribution for the size of the clusters in the different IOPs for the satellite and radar data. It is emphasized that the size of the clusters is based only on the number of pixels counted, the realistic size of the clusters should be considered based on the spatial resolution of the sensors (Satellite Resolution: 4km; Radar Resolution: 2km). The difference between the cluster sizes for the satellite (Figure 2a.) was not so significant if one looks at the means and medians in the boxplots, unlike the radar cluster sizes (Figure 2b.), when evaluating the difference between the IOPs it becomes clear the difference between the cluster sizes, especially at lower thresholds (20 dBZ).

It is also possible to notice that the number of clusters of lower intensity was more prominent in IOP1, this characteristic is related to the rainy period in the Amazon region. This period is also characterized by presenting rain cells with large stratified regions [3], which corroborates to the higher number of clusters with lower intensity thresholds. Another point to be highlighted in Figure 2, is the higher occurrence of clusters during the afternoon period, because in this period the systems are more intense and present greater dynamics of vertical development [3,4].

CONCLUSION AND FUTURE WORK

The results obtained using DBSCAN clustering algorithm demonstrate the different characteristics of atmospheric systems over the study region. With this, it's possible evaluate during the morning period that systems have larger areas, especially for thresholds with less intense systems. Evaluating the size of the clusters it's noted that especially for radar, the difference between clusters of lower intensity (20 dBZ) were greater among the IOPs, this is mainly due to rainy season in the Amazon, where, systems with large stratified areas form clouds with large area extensions. In contrast, during IOP2 events tend to be more isolated and intense.

References:

1] LAURENT, H. et al. Characteristics of the amazonian mesoscale convective systems observed from satellite and radar during the wetamc/lba experiment. Journal of Geophysical Research: Atmospheres, Wiley Online Library, v. 107, n. D20, p. LBA–21, 2002.
[2] ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: kdd. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231.
[3] MACHADO, L. A. et al. Overview: Precipitation characteristics and sensitivities to environmental conditions during goamazon2014/5 and acridicon-chuva. Atmospheric Chemistry and Physics, Copernicus GmbH, v. 18, n. 9, p. 6461–6482, 2018.
[4] NUNES, A. M. et al. Severe convection features in the amazon basin: A trmm-based 15-year evaluation. Frontiers in Earth Science, Frontiers Media SA, v. 4, p. 37, 2016.

Acknowledgements:
The authors wish to thank the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - Brasil (CAPES) for their financial support.

Author affiliations:
Helvecio Neto (INPE, BR); Alan Calheiros (INPE, BR);

Data Science and Machine Learning applied to the fields of ecology, environment and socio-economics (supported by PARCEC) , 04 October 2022, São Paulo, Brazil