Deep convection lifecycle characteristics: a database from GoAmazon experiment

Camila da Cunha Lopes[1], Rachel Ifanger Albrecht[1], Douglas Messias Uba[2], Thiago Souza Biscaro[2], and Ivan Saraiva[3]

[1]Departamento de Ciências Atmosféricas, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Rua do Matão, 1226, 05508-090, São Paulo, SP, Brasil

[2]Divisão de Satélites e Sensores Meteorológicos, Coordenação-Geral de Ciências da Terra, Instituto Nacional de Pesquisas Espaciais, Rodovia Presidente Dutra, Km 40, SP-RJ, 12630-000, Cachoeira Paulista, SP, Brasil

[3]Centro Gestor e Operacional do Sistema de Proteção da Amazônia, Av. do Turismo, 1350, 69049-630, Manaus, AM, Brasil

Correspondence: Camila da Cunha Lopes (camilalopes.ccl@gmail.com)

Abstract. The Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5) Experiment provided a comprehensive suite of cloud-aerosol-precipitation observations with both *in situ* and remote sensing instruments. In this study, we apply a tracking methodology to volumetric radar data, creating a refined database focused on deep convective systems with full lifecycle, incorporating lightning data. This refined deep convection database is shown to be a robust sample of the complete dataset in terms of convective systems morphology. The analysis reveals significant seasonal and diurnal variations in convective morphology and intensity, with most intense systems occurring during the dry-to-wet season transition. The filtered dataset offers a robust sample for future studies on Amazonian convection.

1 Introduction

The Amazon tropical rainforest serves as a natural test bed for several studies on cloud-aerosol-precipitation and land-atmosphere interactions due to its large territorial extent which includes pristine forest, agricultural expansion and a large urban zone with an industrial ~~pole~~center. This complex ecosystem is one of the main centers of convection regulating the climate (Nobre et al., 2009; Artaxo et al., 2022) and the South American Monsoon System (SAMS) (Zhou and Lau, 1998; Jones and Carvalho, 2002). Several convection patterns are present in the region, mainly modulated by the Hadley circulation and the corresponding position of the Intertropical Convergence Zone (ITCZ) which determines the wet (austral summer) and dry (austral winter) seasons.

Several field experiments were conducted in the region in order to study different aspects of the cloud-aerosol-precipitation interactions: the Amazon Boundary Layer Experiment ABLE 2A (Harriss et al., 1988) and 2B (Harriss et al., 1990)) focused on the chemistry and dynamics of lower atmosphere in the dry and wet seasons, respectively; the Large-Scale Biosphere-Atmosphere Experiment in Amazonia LBA Program (Silva Dias et al., 2002) were responsible for numerous field experiments during late 1990s and early 2000s, including the first major mesoscale atmospheric campaign as part of the Tropical Rainfall Measuring Mission (TRMM) validation campaigns and the CHUVA Project (Cloud Processes of the Main Precipitation Systems in Brazil: A Contribution to Cloud-Resolving Modeling and to the Global Precipitation Measurement (GPM) (Machado et al., 2014a). The Green Ocean Amazon Experiment (GoAmazon 2014/5) (Martin et al., 2017) was the first long-term experiment to analyze the effects of the Manaus pollution plume in different experimental sites around Manaus, and included two intensive operation periods (IOPs) in the wet and dry seasons. Unlike previous experiments, an operational weather radar was available during GoAmazon, operated by *Sistema de Proteção da Amazônia* (SIPAM, System for the Protection of the Amazon), which covers all experimental sites with cloud remote sensing data in great temporal and spatial resolutions.

A few studies provide insights about convection characteristics during GoAmazon, such as Giangrande et al. (2017), Machado et al. (2018), Giangrande et al. (2020), and Biscaro et al. (2021). Each of them uses different definitions of cloud features with remote sensing data and even diverges in the definition of wet and dry seasons, which makes it difficult for other studies to follow a homogeneous methodology of convection measurements, specially the ones that analyze cloud-aerosol-precipitation interactions. For this reason, this study aims to create a comprehensive database of convective systems that can be used in future studies regarding GoAmazon data.

2 Materials and Methods

2.1 Data

Data sources for this study are the field experiments GoAmazon (Martin et al., 2017) e CHUVA-Manaus (Machado et al., 2014b) that occurred between January 2014 and December 2015 around Manaus, Amazonas. The main goal of the GoAmazon experiment was to analyze cloud-aerosol-precipitation interactions between the forest and the Manaus metropolitan region, specially the transformation of air plumes from the pristine forest to the Manaus pollution plumes and its eastern propagation. The goal of the CHUVA-Manaus experiment was to characterize the convection regimes in the region with remote sensing, with the installation of a X-Band radar during the experiment and partnerships with CENSIPAM (Centro Gestor e Operacional do Sistema de Proteção da Amazônia) for surface radar data and NASA-JAXA (National Aeronautics and Space Administration — Japan Aerospace Exploration Agency) for satellite radar data from TRMM (Tropical Rainfall Measuring Mission) and GPM (Global Precipitation Measurement Mission).

The main site of these experiments, nominated T3, was located in Manacapuru, Amazonas, Brazil (3.213°S, 60.598°W), about 70 km west from Manaus. A wide range of cloud, precipitation, aerosols and atmospheric instruments were installed at the site, such as part of the ARM (Atmospheric Radiation Measurement) mobile facility AMF1 that took measurements during the most part of the experiment. Some additional instrument and site (resumed in Fig. 1 of Martin et al. (2017)) took measurements during two intensive operation periods (IOPs), with IOP1 being between February 1st and March 31st 2014 (wet season) and IOP2 between August 15th and October 15th 2014 (dry season). More details about the campaigns can be found in Martin et al. (2016, 2017).

In order to create the convective systems database, radar volumes from Sistema de Proteção da Amazônia (SIPAM) radar located in Manaus (3.149°S, 59.991°W, 102.4 m altitude, Fig. 1) were selected in the period of GoAmazon experiment (2014-01-01 to 2015-12-31). These volumes consist of CAPPIs (Constant Altitude Plan Position Indicators) processed by Dr. Thiago

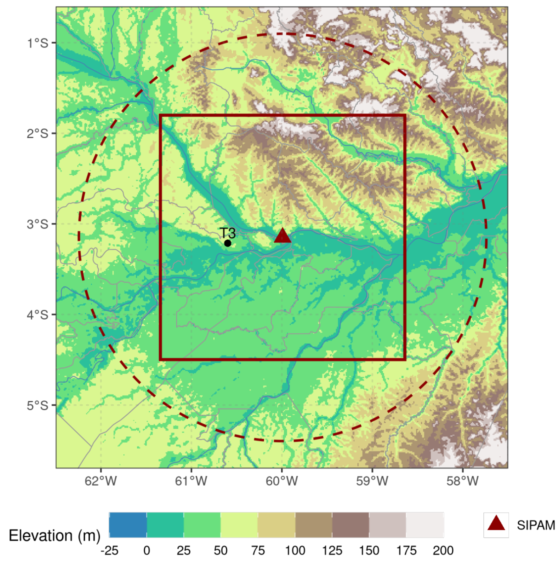

Figure 1. Amazonian region used in this study, with SIPAM radar location (red triangle), 250-km originally calculated coverage (dashed red lines), 150 km x 150 km bounding box (red square) used in the study and T3 site location in Manacapuru where surface data were collected (not shown in this study)

Biscaro (INPE) (Biscaro, 2019) considering limitations in the radar operation such as inconstant sweep numbers and antenna oscillation between volumes. Additionally, a bias correction calculated by Schumacher and Funk (2018) was applied. Table 1 shows the data settings, including the bias values applied throughout the period.

A second data source was employed to calculate parameters related to lightning activity in the convective systems. The Vaisala GLD360 lightning network Demetriades et al. (2010) measures cloud-to-ground return strokes with about 70% efficiency in the VLF (Very Low Frequency) range using two location techniques, magnetic direction finding (MDF) and time-of-arrival (TOA), as well as lightning recognition algorithms. Strokes data were accumulated over 12 minutes in the same timestamps as the radar data and selected within the clusters polygons delimited by TATHU.

2.2 Tracking methodology

The convective systems database was created with the TATHU (Tracking and Analysis of Thunderstorms) software package (Uba et al., 2022) applied to the radar data described in the previous section. The software is a free, open-source python package available at , which addresses convective system tracking as a multi-target tracking problem Makris and Prieur (2014). The main modules are observation, detection, description, tracking, and forecast (not used in this study). The algorithm detects agglomerates – called herein **clusters** – ~~of~~ in an CAPPI field according to the (one or more) threshold(s) and extracts its statistics such as size (in pixels), mean and maximum reflectivity, among others. From subsequent CAPPIs, it tracks and names (via a universal unique identifier — *uuid*) the **convective systems** that occurred in the period and its status during the described

Table 1. SIPAM radar data settings processed by CPTEC-INPE

Type	CAPPIs
Format	Binary, 15 x 500 x 500 elements
Resolution (vertical, horizontal, temporal)	1 x 1 x 1 km, 12 min
Min, max height	2 km, 16 km
CAPPI processing software	RSL (Radar Software Library)
Bias correction (Schumacher and Funk, 2018)	2014-01-03 to 2014-02-05: + 1.0 dB
	2014-02-06 to 2014-08-19: + 3.0 dB
	2014-08-20 to 2014-10-16: - 2.5 dB
	2014-10-17 to 2015-03-06: - 5.5 dB
	2015-03-07 to 2015-07-05: - 4.0 dB
	2015-07-06 to 2015-10-28: - 1.5 dB
	2015-10-29 to 2015-12-31: + 1.0 dB

life cycle, with status being “spontaneous generation” (new cluster), “continuity” (growing or decaying cluster), “split” (when a single cluster separates into two or more clusters after a time step) or “merge” (when two or more clusters merge into a single cluster after a time step). Output data was stored in a PostGIS database, a Data Base Management System (DBMS) with geospatial support that allows storage of large volumes of data in tabular form, including geolocated geometries. The database was converted to GeoJSON datasets in order to become easily available at Lopes (2024). Fig. 2 illustrates the definition of a convective system and its several clusters observed in subsequent timestamps, as adopted in this study.

Table 2 shows the TATHU settings used in this study. The CAPPI height chosen (3 km) is above cloud base, near maximum reflectivity usually found in convective cores. The ~~reflectivity and~~ decision to use CAPPI at this level considered the radar’s scan strategy and the characteristics of convection in the study region. The SIPAM radar used in this study has a 1.8° beam width, which significantly influences the vertical resolution of the scans, particularly at greater distances from the radar. Lower CAPPI heights are “filled” by a limited number of elevation angles at distances exceeding 50 km, reducing the spatial consistency of reflectivity data at these levels. Given that the mean cloud base height in Central Amazonia during the afternoon is approximately 1.5 km (with a range from 1 to 2 km) (Fisch et al., 2004; Souza et al., 2023), a CAPPI at 3 km ensures the detection and tracking of convective systems while minimizing biases due to beam filling effects. Moreover, lower elevations closer to the radar are significantly affected by ground clutter and beam blockage (Giangrande et al., 2016; Schumacher and Funk, 2018), further limiting the reliability of CAPPI data at lower heights. The reflectivity and area thresholds represent light and significant precipitation cores. The relative overlap area strategy considers two subsequent clusters in time as the same convective system when there’s at least a 10% overlap between their areas (i.e., polygons). The maximum interval between images con-

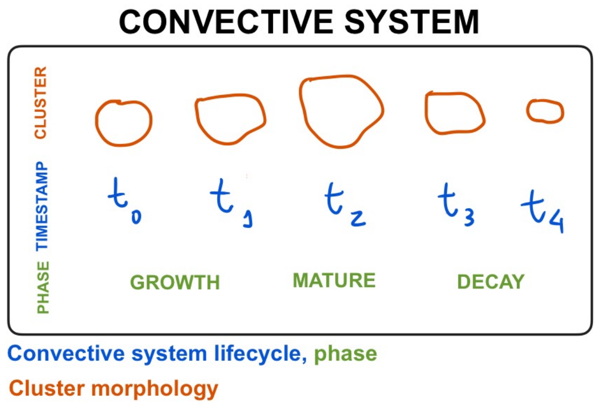

Figure 2. Illustrative convective system as observed by TATHU: clusters morphology (orange) in 5 subsequent timestamps that represent the growth, mature and decay phases of a full lifecycle.

siders a data gap sufficient to ensure continuity of the convective systems, but can imply in different convective systems being tracked as the same if they are close enough, considering that the average lifecycle of tropical convection is smaller than 60 minutes. Within the main statistics, the number of layers~~(~~, being either "0" for only having the ~~first~~ 20-dBZ reflectivity threshold or "1" for having both ~~reflectivity thresholds) is important to filter out convective systems that were not deep enough for this study~~20 and 40-dBZ reflectivity thresholds, can be used to separate systems with or without deep convective cores.

Two datasets were defined based on the TATHU tracking. The original - called herein **raw** - dataset contains all the convective systems observed in the period, regardless of duration, size and status during the life cycle. The **filtered** dataset contains only the convective systems that met the filtering criteria described in Table 2. The second threshold criteria (40 dBZ in a 40 km^2 minimum area) is used here as a definition of deep convective cores present during the lifecycle. Clusters that were identified in the borders of the grid had their associated convective systems discarded to avoid systems without full lifecycle within the study area. Systems with only one timestep were also discarded. We created this filtered dataset to provide a subset of deep convective systems with full life cycle, important criteria for several convection studies.

The original *systems* ~~table, with no filters applied,~~ dataset contains 91609 convective systems and 322896 clusters, while the filtered *systems_filtered* ~~table~~ dataset contains 5976 convective systems (6.5% of the original ~~table~~dataset) and 40394 clusters (12.5% of the original ~~table~~dataset). Using the filtered ~~table~~ dataset with the additional lightning data source, several other parameters were calculated for each cluster (Table 3) related to storm morphology. The following equations were applied for VIWL (Vertically Integrated Warm Liquid), VII (Vertically Integrated Ice) and VIL (Vertically Integrated Liquid), respectively:

$$VIWL = \sum_{i=2km}^{5km} 3.44 \times 10^{-6} \left(\frac{Z_i + Z_{i+1}}{2} \right)^{\frac{4}{7}} \qquad (1)$$

Table 2. TATHU parameters (original names in parentheses) and values chosen for the generation of *systems* (raw) and *systems_filtered* (filtereddata) tablesdatasets.

Input data	3-km CAPPI in a 150 x 150 km box (300 x 300 elements) centered at the radar location, between January 2014 and December 2015
Reflectivity thresholds (*threshold value*)	values greater than 20 dBZ, 40 dBZ
Minimum cluster sizes (*minarea*)	100 km^2, 40 km^2
Tracking technique (*trackers class*)	Relative overlap area (*RelativeOverlapAreaStrategy*)
Minimum cluster overlap area	10%
Maximum interval between images	60 min
Statistics (*stats*)	Maximum, mean and standard deviation of reflectivity, size (amount of pixels), number of layers (corresponding to having one or two reflectivity and minimum size thresholds)
Output	PostGIS database (*systems* and *systems_filtered* tables)
Filters applied	- Have at least one 40 dBZ/40 km^2(corresponding to deep convection) pixel in any timestamp - Do not "touch" the borders bounding box of the grid (corresponding to probably have part of the cluster outside the tracking region) - Last longer than 12 minutes (one timestamp) or have relation with other convective system (by split or merge)

$$VII = \sum_{i=7km}^{16km} \pi \rho_i N_o^{\frac{3}{7}} \left[\frac{5.28 \times 10^{-18}}{720} \left(\frac{Z_i + Z_{i+1}}{2} \right) \right]^{\frac{4}{7}} \tag{2}$$

$$VIL = \sum_{i=2km}^{16km} 3.44 \times 10^{-6} \left(\frac{Z_i + Z_{i+1}}{2} \right)^{\frac{4}{7}} \tag{3}$$

Table 3. Clusters additional parameters calculated on the filtered ~~data table~~dataset

Variable	Description	Reference
gld	GLD strokes within cluster area detected within 12 min (interval between scans)	
echotop_0, *echotop_20*, *echotop_40*	0, 20 and 40 dBZ echo top heights	
z_freq	Reflectivity frequencies per height; 15 x 16 matrices with reflectivities between -10 and 70 dBZ (every 5 dBZ) and heights between 2 and 16 km	Yuter and Houze (1995)
viwl_kgm2	VIWL (Vertically Integrated Warm Liquid) of cluster in kg/m^2, Equation 1	
vii_kgm2	VII (Vertically Integrated Ice) of clusters in kg/m^2, Equation 2	Petersen and Rutledge (2001)
vil_kgm2	VIL (Vertically Integrated Liquid) of cluster in kg/m^2, Equation 3	Greene and Clark (1972)
nae_s_1	Normalized area expansion in s^{-1} between clusters of a same convective system	Machado and Laurent (2004)
gld_strmin	Strokes rate between clusters of a same convective system in strokes/min	
echotop0_kmmin, *echotop20_kmmin*, *echotop40_kmmin*	0, 20 and 40 dBZ echo top rate between clusters of a same convective system in km/min	

3 Results

3.1 Raw and filtered ~~tables~~ datasets characteristics

Fig. ~~2~~ 3 shows the size — represented by area in km^2 — distribution of clusters and maximum size of convective systems. Both distributions are exponential, with higher exponential in the filtered than the raw ~~data table~~dataset, i.e., the maximum area value is 6 times larger (60000 km^2 vs 10000 km^2) and the distribution drops faster in the filtered data. These characteristics show the filtering effect around the border of the 150 km bounding box (Fig. 1), which excluded very large clusters: a 200 ×

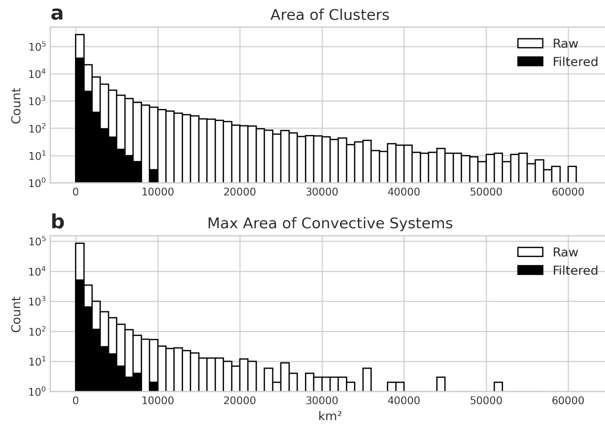

Figure 3. Distribution of clusters area (a) and convective systems' max area (b) of raw (white) and filtered (black) ~~data tables~~datasets.

200 points clusters with a 40000 km^2 area, for example, can be considered too large because it occupies 2/3 of the grid and probably touches the border of the bounding box in a given time stamp.

Fig. ~~3~~ 4 shows the mean and maximum clusters reflectivity distribution of raw (a) and filtered (b) ~~tables~~datasets. On raw data, the distributions show no significant frequency peaks, with mean reflectivity distributed mainly (frequency above 20%) between 20 and 40 dBZ and maximum reflectivity between 35 and 50 dBZ (frequency above 15%). Oppositely, on the filtered data, peaks (above 35%) can be found between 25 and 40 dBZ and between 50 and 55 dBZ of mean and maximum reflectivity, respectively. These differences between the distributions represent the filtering effect on the type of the selected clusters: the filters ended up selecting more intense clusters (larger mean and maximum reflectivity) and excluded mainly the ones that did not exceed 40 dBZ — observe that the clusters with maximum reflectivity below 40 dBZ are significantly less frequent (below 5%) compared to the raw data ones.

Table 4 shows some characteristics of the convective systems of raw and filtered ~~tables~~ datasets using the clusters classifications on each time step. On both datasets, the percentage of spontaneously generated convective systems was similar (above 70%); on the raw data, not all these systems had their full lifecycle covered, while on the filtered data this is true because one of the filtering criteria is to exclude convective systems that leave the radar coverage area. 53% of the filtered convective systems had split or merge during their lifecycle, compared to 37% of the raw systems; an important point here is that 31% of the raw systems and only 2% of the filtered systems lasted only one time step (12 min) (not shown), meaning that more raw convective systems did not last long enough to go through split or merge. The majority (80%) of the filtered and half (49%) of the raw convective systems were considered with full lifecycle (last time step was "continuity"), which also is explained by the high percentage of raw systems with only one time step (i.e., only "spontaneous generation" status).

Table 5 shows the distribution of convective systems durations of raw and filtered data ~~tables~~datasets. The majority of raw (40%) and filtered (47%) systems lasted up to 1 hour, indicating the predominance of isolated convective systems. Only 10% of raw systems lasted between 1 and 3h, compared to 36% of filtered systems. ~~Another significant difference occurred in the long~~

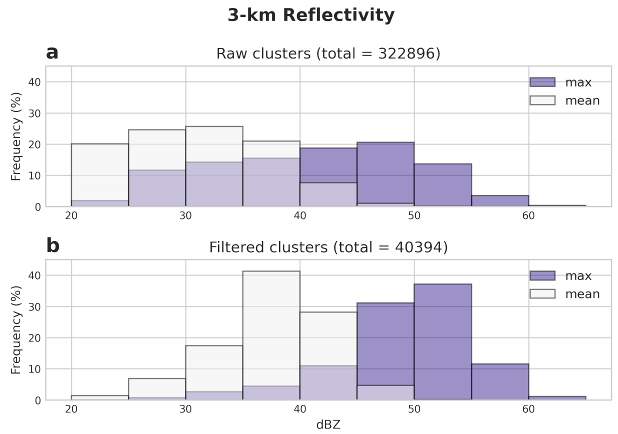

Figure 4. Distribution of clusters max (purple) and mean (white) reflectivity of raw (a) and filtered (b) ~~data tables~~datasets.

Table 4. Frequency of spontaneously generated (did not generate from splits or first steps of algorithm rounds), with split and/or merge and full lifecycle (last time step in continuity) convective systems (CS) of raw and filtered ~~data tables~~datasets.

	Raw	Filtered
Convective Systems (total)	91609	5976
CS spontaneously generated (%)	72	73
CS with split/merge (%)	37	53
CS with full lifecycle (%)	49	80

~~duration systems : almost two times~~ The majority of convective systems were short-lived, with 40% of raw and 47% of filtered systems lasting up to 1 hour, highlighting the prevalence of isolated convective events. A notable difference is observed in the 1–3h range, where only 10% of raw systems persisted, compared to 36% in the filtered dataset. For long-duration systems, a significant disparity emerges: nearly twice as many raw systems (631~~vs. 318) raw systems~~) lasted longer than 6h compared to those in the 5~~-~~–6h ~~, opposed to less (~~range (318). In contrast, the filtered dataset shows fewer systems exceeding 6h (18~~vs. 27)filtered systems. The filtering within the 150 km bounding box affects mainly the long-lasting systems (>~~) than those lasting between 5–6h ~~) because these systems probably were generated and/or dissipated outside the bounding box, explaining the smaller proportion of long-lasting systemsin the filtered data table~~(27), indicating a stronger filtering effect on prolonged convective systems.

Fig. ~~4~~ 5 shows the monthly distribution of raw and filtered clusters and convective systems. Seasons and intensive operation periods (IOPs) were defined according to Machado et al. (2018): dry season between August and October, dry-to-wet season

Table 5. Frequency of spontaneously generated (did not generate from splits or first steps of algorithm rounds), with split and/or merge and full lifecycle (last time step in continuity) convective systems (CS) of raw and filtered data tables.

	Raw	Filtered
CS duration $\leq 1h$ $\leq 1h$	36254 79301 (86,6%)	2805 3448 (57,7%)
$1h$ < CS duration $\leq 2h$ $\leq 2h$	6950 (7,6%)	1535 (25,7%)
$2h$ < CS duration $\leq 3h$ $\leq 3h$	2597 (2,8%)	635 (10,6%)
$3h$ < CS duration $\leq 4h$ $\leq 4h$	1170 (1,3%)	239 (4%)
$4h$ < CS duration $\leq 5h$ $\leq 5h$	641 (0,7%)	74 (1,2%)
$5h$ < CS duration $\leq 6h$ $\leq 6h$	318 (0,3%)	27 (0,5%)
CS duration > $6h$	631 (0,6%)	18 (0,3%)
Total	91609 (100%)	5976 (100%)

between November and December, wet season between January and March, IOP1 between February 1st and March 31st 2014 and IOP2 between August 15th and October 15th 2014. In general, a larger frequency of clusters and convective systems occurred in the wet and transition (wet-to-dry and dry-to-wet) seasons, with a peak of filtered clusters/systems in November 2015 (almost two times more than the raw clusters/systems). The proportion between raw and filtered systems changes over the months, with a larger frequency of raw clusters and convective systems on the wet seasons and the opposite on the dry seasons. This difference can be explained by the filter of very large clusters cited previously, which are more common in the wet season. Considering the climatological characteristics of each season, it is expected that more clusters/convective systems occur during the wet season than of the dry season, which was the case for both raw and filtered data.

Table 6 shows presents the duration of raw and filtered convective systems by seasons. Similar proportions occur among seasonsand IOP. The distribution of system durations is generally consistent across seasons, with similar proportions observed between raw and filtered systems. Similar datasets. In contrast to Table 5, where less than half of the raw systems lasted up to 1h, seasonal distributions show that more than 80% (of raw systems and 50%) of the raw (filtered) lasted up to 1 hour, compatible with isolated systems, and of filtered systems persisted for no more than 1h, indicating the predominance of isolated convective systems. Additionally, 97% (of raw and 94%) lasted up to 3 hours. The same characteristics can be found during the IOPs (Table 7), which are samples of the wet (IOP1) and dry (IOP2) seasonsof filtered systems had durations of 3h or less.

Distribution of convective systems (CS) durations of raw and filtered data tables separated by dry, dry-to-wet and wet seasons. Raw Filtered Raw Filtered Raw Filtered CS duration $\leq 1h$ 14431 716 10820 750 24174 643 $1h$ < CS duration $\leq 2h$ 1033 311 937 301 2254 297 $2h$ < CS duration $\leq 3h$ 365 138 348 131 818 113 $3h$ < CS duration $\leq 4h$ 167 52 142 47 364 53

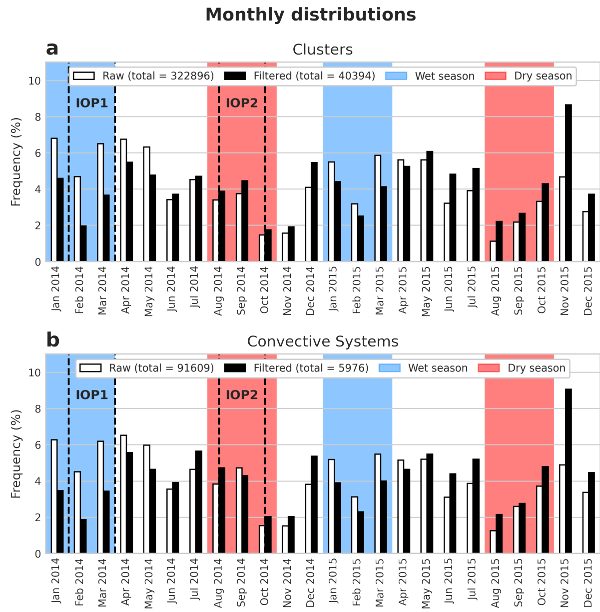

Figure 5. Monthly relative frequency distribution of clusters (a) and convective systems (b) of raw (white) and filtered (black) ~~data tables~~datasets. The blue and red areas delimit wet and dry seasons, respectively, and dashed lines delimit the intensive operation periods IOP1 and IOP2.

~~$4h <$ CS duration $\leq 5h$ 87 14 73 10 209 19 $5h <$ CS duration $\leq 6h$ 37 4 46 8 103 1 CS duration $> 6h$ 62 3 78 2 255 6 Total 16182 1238 12444 1249 28178 1132~~

Fig. ~~5~~ 6 shows the hourly distribution of raw and filtered clusters divided by ~~seasons~~IOP. Comparing the raw and filtered clusters, in all seasons there are a larger frequency of raw clusters during late night/dawn and a larger frequency of filtered clusters during late morning/afternoon. In the dry season (IOP2), this difference is more pronounced: the filtered clusters are more frequent (above 15%) between 1400 and 1500 local time compared to the raw clusters (below 10%). ~~In the dry-to-wet season, the difference between raw and filtered frequencies is much smaller, specially during the day. Comparing the IOPs (Fig. 6), similar characteristics are found, specially on IOP2 (Fig. 6b) during the dry season, where the difference between raw and filtered clusters are even larger during the afternoon.~~ These differences between raw and filtered clusters indicate a more diurnal characteristic of the filtered clusters, while the raw clusters are more nocturnal, probably represented by the very large and long-lasting (above 6h) clusters described previously.

~~Hourly distribution of clusters of raw (white) and filtered (black) data tables separated by dry (a), dry-to-wet (b) and wet (c) seasons. The blue and yellow areas delimit diurnal and nocturnal periods, respectively.~~

Fig. 7 shows the hourly distribution of raw and filtered convective systems initiation divided by ~~season~~IOP. Comparing the raw and filtered systems, in ~~all~~ both seasons there are a larger frequency of raw systems initiating during dawn, while a larger frequency of filtered systems initiate during morning/afternoon~~ — this~~. This difference is even greater (almost 10%) in the dry

Table 6. Distribution of convective systems (CS) durations of raw and filtered ~~data tables~~ datasets separated by intensive operation periods IOP1 and IOP2.

	IOP1		IOP2	
	Raw	Filtered	Raw	Filtered
CS duration ~~$\leq 1h$~~ $\leq 1h$	8427 (86,1%)	176 (55,7%)	6565 (90%)	272 (55,5%)
$1h$ < CS duration ~~$\leq 2h$~~ $\leq 2h$	772 (7,9%)	88 (27,8%)	432 (5,9%)	130 (26,5%)
$2h$ < CS duration ~~$\leq 3h$~~ $\leq 3h$	277 (2,8%)	34 (10,8%)	146 (2%)	53 (10,8%)
$3h$ < CS duration ~~$\leq 4h$~~ $\leq 4h$	118 (1,2%)	10 (3,2%)	60 (0,8%)	24 (4,9%)
$4h$ < CS duration ~~$\leq 5h$~~ $\leq 5h$	66 (0,7%)	6 (1,9%)	41 (0,6%)	7 (1,4%)
$5h$ < CS duration ~~$\leq 6h$~~ $\leq 6h$	40 (0,4%)	0 (0%)	19 (0,3%)	2 (0,4%)
CS duration >~~$6h$~~ $\leq 6h$	91 (0,9%)	2 (0,6%)	31 (0,4%)	2 (0,4%)
Total	9792 (100%)	316 (100%)	7294 (100%)	490 (100%)

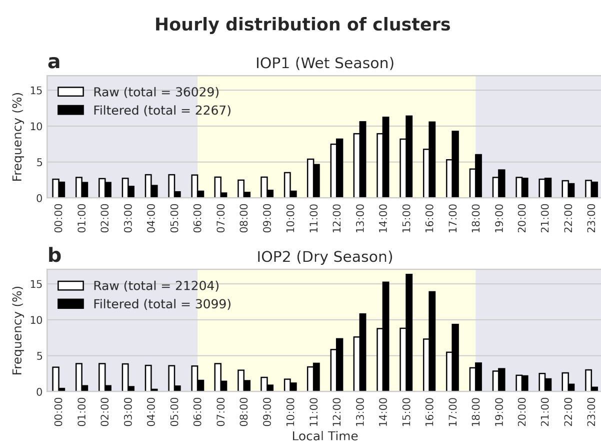

Figure 6. Hourly relative distribution of clusters of raw (white) and filtered (black) ~~data tables~~ datasets separated by intensive operation periods IOP1 (a) and IOP2 (b). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

season. Comparing the seasons, ~~there is a peak (frequency above 10%) in the initiation of the filtered systems at 1100~~ both in the dry and ~~1400 local time in the dry-to-wet season (Fig. 7b), while in the dry (Fig. 7a) and wet (Fig. 7c)~~ wet seasons the initiation peak occurs only at 1400 local time. ~~This indicates that a significant fraction of the filtered convective systems tend to initiate earlier in the dry-to-wet season, consistent with a different forcing (possibly different from the diurnal cycle) acting~~

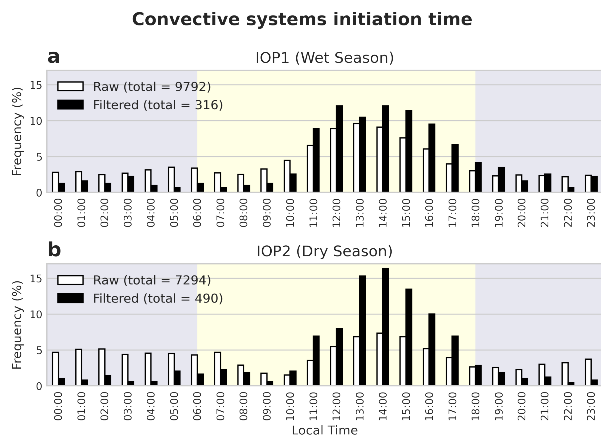

Figure 7. Hourly relative distribution of clusters convective systems initiation time of raw (white) and filtered (black) data tables datasets separated by dry intensive operation periods IOP1 (a) , dry-to-wet and IOP2 (b)and wet (c) seasons. The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

in the initiation of these systems. During the IOPs (Fig. 8), there are similar characteristics to the seasons associated with each IOP, such as the difference between the raw and filtered systems and the nocturnal and diurnal initiation, respectively. Specifically about the filtered systems, the 1400 local time peak during the wet season is not highlighted in IOP1(Fig. 8a), with another peak at 1200 local time.

Hourly distribution of clusters initiation time of raw (white) and filtered (black) data tables separated by intensive operation periods IOP1 (a) and IOP2 (b). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

3.2 Characteristics specific to the filtered systems

Starting with the monthly distribution, Fig. 9 8 shows the clusters and convective systems separated by lightning activity, as well as the distribution of GLD strokes. More than double the clusters (27214 vs. 13180) had no electrical activity, while more convective systems (3758 vs. 2218) showed lightning, which means that, in general, the convective systems with lightning consists of only a few clusters with lightning. There are significant differences between clusters and systems frequency: a larger frequency of clusters and systems without lightning occurs between wet and dry seasons, while peaks of clusters and systems frequency occurs in the dry-to-wet season, comparable to the peak strokes frequency. These findings are similar to what is found in previous works such as Albrecht et al. (2011, 2016).

Separating the analysis in systems with and without lightning, Fig. 10 9 shows the hourly distribution of clusters and initiation of convective systems with lightning and GLD strokes separated by dry, dry-to-wet and wet seasonsIOP. All variables are more frequent during late morning/afternoon, with significant differences between seasons: on the dry seasonIOPs: on IOP2, the clusters and strokes peaks occurs at 1500 local time, which is also the preferential time of convective systems initiation; on the dry-to-wet seasonIOP1, the clusters and strokes peaks occur earlier, on 1400 local time, while the systems initiation occurs

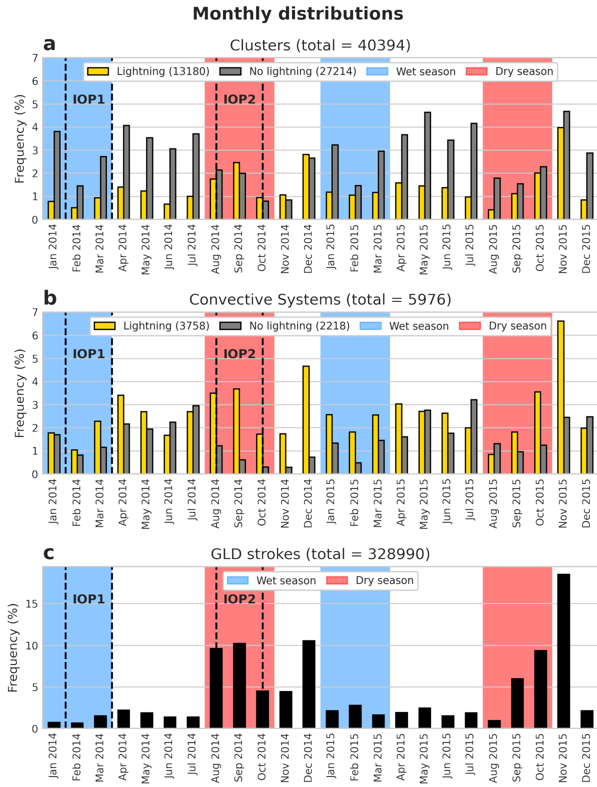

Figure 8. Monthly relative distribution of clusters (a), convective systems (b) and GLD strokes (c) of filtered data table datasets separated by lightning (yellow bars) or no lightning (gray bars) occurrence. The blue and red areas delimit wet and dry seasons, respectively, and dashed lines delimit the intensive operation periods IOP1 and IOP2.

preferentially even earlier, at 1100 local time, as observed previously (Fig. 5); on the wet season, the clusters and strokes peaks occur later, at 1600 local time, while the systems initiation occurs preferentially earlier, at 1400 local time. The correspondence between clusters and strokes peaks means that most lightning was distributed among more clusters, not necessarily that were clusters with lightning peaks; an example of clusters with possible lightning peaks is during the dry season, at 1900 local time, where the lightning frequency is almost double that the frequency of clusters. Another interesting point is about the frequency peaks among seasons: the dry and wet season peaks were always larger than the dry-to-wet peaks, except for the lightning activity, which indicates that the clusters and the convective systems initiation were more distributed throughout the day, but the electrical activity was higher, which is expected for this time of the year, as cited previously. During the IOPs (Fig. 11), there are similar characteristics in the wet and dry seasons relating to clusters and lightning, but the It's important to note that the low amount of systems influenced the frequency distribution, for this reason it is not possible to observe any meaningful tendency.

~~Hourly distribution of clusters with lightning (a), initiation time of clusters with lightning (b) and GLD strokes (c) of filtered data table during dry (red bars), dry-to-wet (green bars) and wet (blue bars) seasons. The blue and yellow areas delimit diurnal and nocturnal periods, respectively.~~

~~Hourly distribution of clusters with lightning (a), initiation time of clusters with lightning (b) and GLD strokes (c) of filtered data table during intensive operation periods IOP1 (blue) and IOP2 (red). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.~~

~~Fig. 12 shows the hourly distribution of~~ . For the clusters and initiation of convective systems without lightning ~~separated by dry, dry-to-wet and wet seasons. As with the clusters and systems with lightning, the~~ (not shown), the largest frequencies occur during late morning/afternoon~~; an interesting exception is on the initiation (Fig. 12b) during the dry-to-wet season, where a secondary peak can be observed during dawn (0400 and 0500 local time), showing another (probably synoptic) forcing influencing the formation of these nocturnal systems.~~ . The initiation frequency peaks on the dry and wet seasons occur at 1500 local time, while on the dry-to-wet season the distribution is more dispersed, with approximately equal peaks at 1200, 1300 e 1600 local time. ~~Similar characteristics are observed during the IOPs (Fig. 13), also with a low amount of convective systems and consequently no meaningful tendencies.~~

~~Hourly distribution of clusters without lightning (a) and initiation time of clusters without lightning (b) of filtered data table during dry (red bars), dry-to-wet (green bars) and wet (blue bars) seasons. The blue and yellow areas delimit diurnal and nocturnal periods, respectively.~~

~~The initial location and propagation direction of the convective systems separated by month are shown in Fig. 14 and Fig. 15 for the years 2014 and 2015, respectively. There is no preferential region for the initiation of the convective systems and, in general, at least half of them barely propagates during its lifecycle, compatible with the duration up to 1h of most of them (Table 5). The direction of propagation differs significantly over the months: in most months, the longer systems propagate in the same direction, like in January 2014 (southwest) and July 2014 and June 2015 (west), but some months have systems without a preferred direction, such as November 2014, December 2014 and March 2015.~~

~~Initiation (red dots) and propagation direction (black arrows with size normalized by distance between start and end) of convective systems in the filtered data table that occurred in 2014 separated by month. The gray circle represents the 150-km radius of the radar location.~~

~~Initiation (red dots) and propagation direction (black arrows with size normalized by distance between start and end) of convective systems in the filtered data table that occurred in 2015 separated by month. The gray circle represents the 150-km radius of the radar location.~~

In order to analyze ~~only~~ the propagation direction of the convective systems, Fig. ~~16~~ 10 shows the frequency distribution of the ~~complete time series. The main propagation direction (about 30% of total systems) is west, followed by west-northwest and west-southwest. This predominant direction from east to west is~~ propagation direction separated by IOP. The predominant direction of the convective systems is from the east, consistent with the main dynamic forcing in the Amazon: moisture flux from the tropical Atlantic by the easterlies influenced by the position of the Intertropical Convergence Zone (ITCZ) (Silva Dias and Carvalho, 2016).

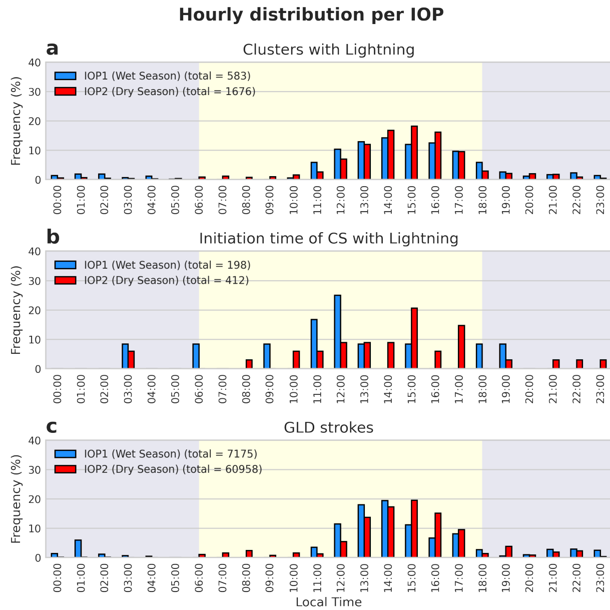

Figure 9. Hourly relative distribution of clusters ~~without~~ with lightning (a)~~and~~ , initiation time of clusters ~~without~~ with lightning (b) and GLD strokes (c) of filtered ~~data table~~ dataset during intensive operation periods IOP1 (blue) and IOP2 (red). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

~~Convective systems' propagation direction distribution in the filtered data table. The direction was defined by the distance between the first and last centroids of the convective system.~~

~~Fig. 17 shows the frequency distribution of the propagation direction separated by dry (a), dry-to-wet (b) and wet (c) seasons. The predominant direction of the convective systems is also west in all seasons, with some deviations:~~ Comparing the IOPs, in the wet season, about 25% propagated from the east and east-northeast, while in the dry season ~~,~~ more than 30% of the systems propagated ~~west and west-northwest; in the dry-to-wet season, about 30% propagated west and 25% propagated west-southwest; in the wet season, about 25% propagated west and west-southwest. The propagation direction during the IOPs (Fig. 18) was similar to the corresponding seasons.~~ from the east and east-southeast. This direction shift from ~~west-northwest to west-southwest~~ east-northeast to east-southeast is related to the shift in the position of the ITCZ and the cold fronts propagation in the South American continent, which affect the zonal winds regime in the Amazonian region (Rickenbach et al., 2002). These results are also consistent with ~~Gupta2024-pi~~Gupta et al. (2024), which focused on isolated convective systems near Manacapuru (T3) site.

~~Convective systems' propagation direction distribution in the filtered data table separated by dry (a), dry-to-wet (b) and wet (c) seasons. The direction was defined by the distance between the first and last centroids of the convective system.~~

In order to analyze the intensity of the convective systems, Fig. ~~19~~ 11 shows the frequency distribution of maximum heights and variation rates (considering the 12-min interval between radar scans) of the 0, 20 and 40 dBZ echoes, associated with the

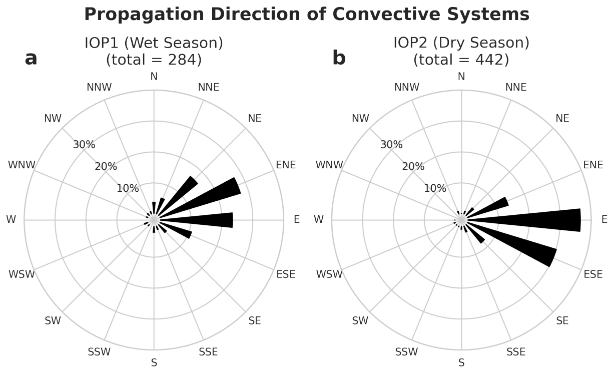

Figure 10. Convective systems' relative propagation direction distribution in the filtered ~~data table~~ dataset separated by intensive operation periods IOP1 (a) and IOP2 (b). The direction was defined by the distance between the first and last centroids of the convective system.

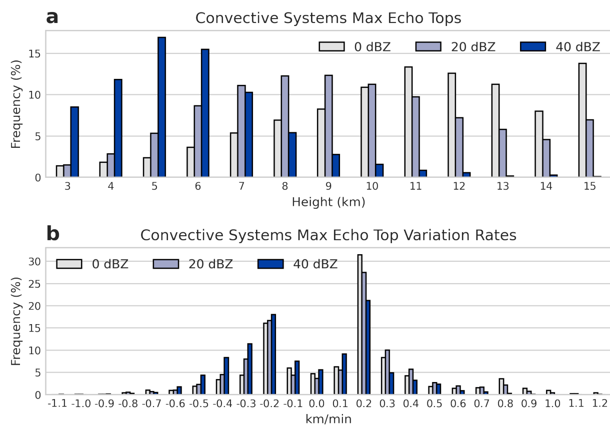

Figure 11. Convective systems' 0 (gray bars), 20 (light blue bars) and 40 dBZ (dark blue bars) max echo tops (a) and variation rates (b).

cloud top height, precipitable hydrometeors height and intense precipitation heights, respectively. The maximum ~~cloud top height (Fig. 19a) was frequently~~ top heights are more frequent above 10 km, with ~~high frequencies in the maximum height available (~~peaks in 11 km (dry season) and 15 km ~~).~~ (wet season). The maximum precipitable height was more frequent between 7 and 11 km. The maximum intense precipitation height was also more frequent between 4 and 7 km, with peaks in 5 km (both dry and wet seasons). These high tops, complemented by maximum precipitation height more frequently between 7 and 11 km and maximum intense precipitation height more frequently between 4 and 7 km, show how these systems were predominantly deep in its most intense moment. The variation rates (Fig. ~~19a~~11b) were similar between the echoes, with frequency peaks in ~~-200 and 200 m~~-0,2 and 0,2 km/min, indicating significant fluctuations of the echo tops throughout its life cycle.

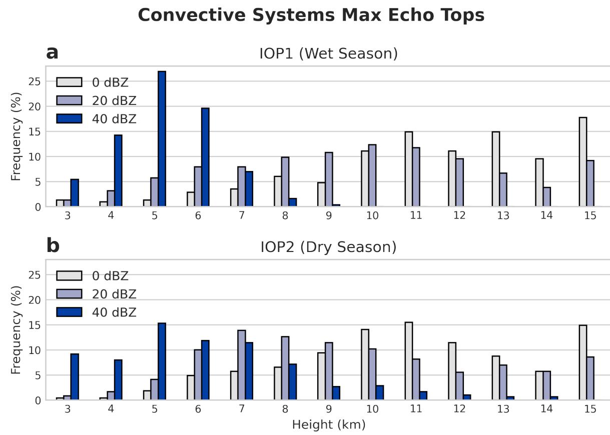

Figure 12. Convective systems' 0 (gray bars), 20 (light blue bars) and 40 dBZ (dark blue bars) max echo tops separated by intensive operation periods IOP1 (a) and IOP2 (b).

Fig. ~~20~~ 12 shows the frequency distribution of the maximum echo tops of 0, 20 and 40 dBZ separated by ~~dry (a), dry-to-wet (b) and wet (c) seasons~~IOPs. As in the complete time series, the maximum top heights are more frequent above 10 km, with peaks in 11 km (~~dry season~~IOP2) and 15 km (~~dry-to-wet and wet seasons~~IOP1). The maximum precipitable height was also more frequent between 7 and 11 km~~, but with a secondary (10%) peak in 15 km in the dry-to-wet season~~. The maximum intense precipitation height was also more frequent between 4 and 7 km, with peaks in 5 km~~(dry and wet seasons) and 6 km (dry-to-wet season). The small differences among the seasons indicate that the dry-to-wet season systems are slightly deeper (0 dBZ peak in 15 km height) and intense (higher peaks in 20 and 40 dBZ). The distribution during the IOPs (Fig. 21) are similar to the corresponding seasons.~~ .

~~Convective systems' 0 (gray bars), 20 (light blue bars) and 40 dBZ (dark blue bars) max echo tops separated by dry (a), dry-to-wet (b) and wet (c) seasons.~~

In order to analyze the clusters vertical profile, Fig. ~~22~~ 13 shows the frequency diagram by altitude of the clusters reflectivity of the complete time series. The most frequent profile is of a small reflectivity variation with height, 25 dBZ on the surface to 10 dBZ on 15 km height (i.e., a -1 dBZ per km variation rate). Less frequent (between 5 and 10%) profiles have very low (5 to 10 dBZ) or high (40 dBZ) reflectivity on the surface, and up to 25 dBZ on 15 km height, with a minimum of 0 dBZ on 12 km height.

Fig. ~~23~~ 14 shows the clusters frequency by altitude diagram separated by ~~dry (a), dry-to-wet (b) and wet (c) seasons~~ IOP as well the different between them~~(d, e and f)~~. Considering the largest frequencies, the profiles are similar between them (and with the complete time series profile), but ~~with differences throughout the seasons: the dry season profile is slightly more intense (up to 10%) than the dry-to-wet season (Fig. 23d), in a sense that higher reflectivities (above 15 dBZ) are more frequent in the full profile; the dry season profile is significantly more intense (up to 20 dBZ) than the wet season (Fig. 23e), specially near the surface, where the reflectivity in the dry season is up to 45 dBZ; the dry-to-wet season profile is more intense (up to 10%) than~~

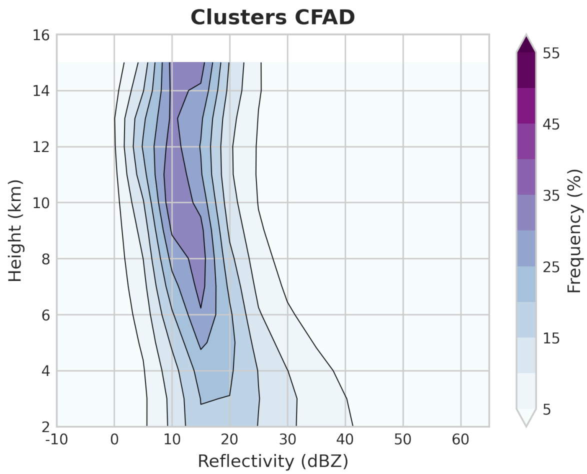

Figure 13. Mean Contoured Frequency by Altitude Diagram (CFAD) of clusters' reflectivity with 5-dBZ bins.

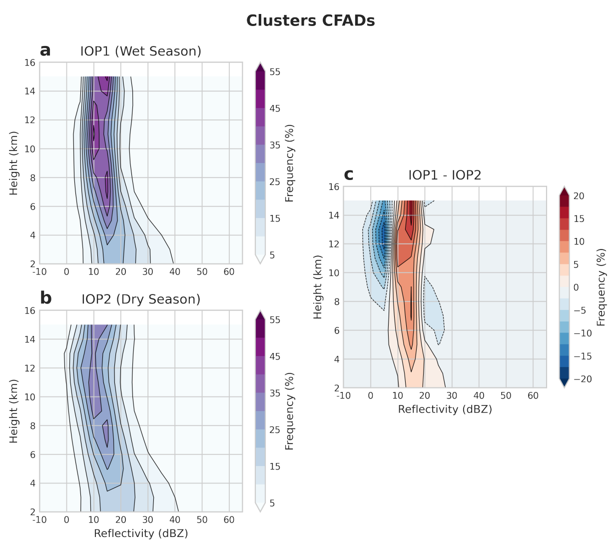

Figure 14. Mean Contoured Frequency by Altitude Diagram (CFAD) of clusters' reflectivity with 5-dBZ bins separated by intensive operation periods IOP1 (a) and IOP2 (b) as well as IOP1 - IOP2 anomaly (c).

~~the wet season (Fig. 23f), specially above 20 dBZ. Comparing the IOPs (Fig. 24), the profiles are similar to the corresponding seasons, but the difference between them is inverse: the~~ the IOP1 profile (wet season) is more intense (up to 20%) than the IOP2 (dry season), specially between 10 and 20 dBZ.

~~Mean Contoured Frequency by Altitude Diagram (CFAD) of clusters' reflectivity with 5-dBZ bins separated by dry (a), dry-to-wet (b) and wet (c) seasons, as well as dry-to-wet - dry (d), dry - wet (e) and dry-to-wet - wet (f) anomalies.~~

4 Conclusions

A database of convective systems that occurred during GoAmazon experiment was created to provide comprehensive convection data for future GoAmazon studies. The *systems* and *systems_filtered* datasets cover the main convective characteristics, including morphology and intensity, as well as electrical activity. The filtered dataset is shown to be an acceptable sample of the complete dataset, selecting deep convective systems with full lifecycle within the research area. Convection seasonality is also well represented, with more intense convective systems between dry and dry-to-wet seasons and less intense in the wet season, typically occurring during late morning/early afternoon. The preeminent propagation direction of these systems are associated with easterlies with a transition from slightly north to slightly south associated with the ITCZ position.

It is important to consider the limitations in the convection description when using these data for future research. Since the SIPAM radar main role is operational, its settings are not optimal for convection research: low spatial (1 km) and temporal (12 min) resolution (considering it is a weather radar), beam blockage during the experiment (Giangrande et al., 2016; Tian et al., 2021), radar software settings change during the experiment. Another limitation is in the tracking itself, specially when dealing with system split/merge (which occurs in a significant portion of the convective systems database) that can be more complex in multicellular systems. Even with these limitations, the database is an important source of convection characteristics for cloud-aerosol-precipitation research.

5 Code availability

The TATHU software package is available at https://github.com/uba/tathu (Uba et al., 2022). The code developed to create the datasets with TATHU is available at https://github.com/cclopes/tathu/tree/sipam-tracking/sipam-tracking.

6 Data availability

The *systems* and *systems_filtered* datasets are available at https://doi.org/10.5281/zenodo.13732692 (Lopes, 2024). SIPAM radar data are available at https://ftp.cptec.inpe.br/chuva/goamazon/experimental/level_0/eq_radar/esp_band_s/st_sipam/.

Author contributions. CL and RA designed the study. CL processed the data and wrote the manuscript advised by RA. DU developed the tracking software, helped during data processing, and reviewed the manuscript. TB processed the radar data and reviewed the manuscript. IS helped in providing the radar data and reviewed the manuscript.

Competing interests. No competing interests are present.

Acknowledgements. This study was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - grant 88887.464412/2019-00), Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp - grants 2009/15235-8, 2017/17047-0, 2022/13257-9, 2023/04358-9) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - grants 438638/2018-2, 313355/2021-5, 440171/2022-9). We would like to thank Sistema de Proteção da Amazônia (SIPAM) for providing the radar data and the ARM GoAmazon 2014/5 operations and science team for their efforts during the experiment. We also thank Vaisala Inc. for providing the GLD360 lightning dataset for this study.

References

Albrecht, R. I., Morales, C. A., and Dias, M. A. F. S.: Electrification of precipitating systems over the Amazon: Physical processes of thunderstorm development, Journal of Geophysical Research, D: Atmospheres, 116, D08 209, https://doi.org/10.1029/2010jd014756, 2011.

Albrecht, R. I., Goodman, S. J., Buechler, D. E., Blakeslee, R. J., Christian, H. J., Albrecht, R. I., Goodman, S. J., Buechler, D. E., Blakeslee, R. J., and Christian, H. J.: Where are the lightning hotspots on Earth?, Bulletin of the American Meteorological Society, 97, BAMS–D–14–00 193.1, https://doi.org/10.1175/bams-d-14-00193.1, 2016.

Artaxo, P., Hansson, H. C., Machado, L. A. T., and Rizzo, L. V.: Tropical forests are crucial in regulating the climate on Earth, PLOS Climate, 1, e0000 054, https://doi.org/10.1371/journal.pclm.0000054, 2022.

Biscaro, T. S.: PROCESSOS FÍSICOS QUE CONTROLAM A FORMAÇÃO DE NUVENS RASAS E SUA EVOLUÇÃO A CONVECÇÃO PROFUNDA NA REGIÃO CENTRAL AMAZÔNICA, Ph.D. thesis, INPE, 2019.

Biscaro, T. S., Machado, L. A. T., Giangrande, S. E., and Jensen, M. P.: What drives daily precipitation over the central Amazon? Differences observed between wet and dry seasons, Atmospheric Chemistry and Physics, 21, 6735–6754, https://doi.org/10.5194/acp-21-6735-2021, 2021.

Demetriades, N. W. S., Murphy, M. J., and Cramer, J. A.: Validation of Vaisala's global lightning dataset (GLD360) over the continental United States, in: Preprints, 29th Conf. on Hurricanes and Tropical Meteorology, Tucson, AZ, Amer. Meteor. Soc. D, vol. 16, vaisala.com, https://www.vaisala.com/sites/default/files/documents/6.Demetriades,%20Murphy,%20Cramer.pdf, 2010.

Fisch, G., Tota, J., Machado, L. A. T., Silva Dias, M. A. F., da F. Lyra, R. F., Nobre, C. A., Dolman, A. J., and Gash, J. H. C.: The convective boundary layer over pasture and forest in Amazonia, Theoretical and applied climatology, 78, 47–59, https://doi.org/10.1007/s00704-004-0043-x, 2004.

Giangrande, S. E., Toto, T., Jensen, M. P., Bartholomew, M. J., Feng, Z., Protat, A., Williams, C. R., Schumacher, C., and Machado, L.: Convective cloud vertical velocity and mass-flux characteristics from radar wind profiler observations during GoAmazon2014/5, Journal of Geophysical Research, D: Atmospheres, 121, 12,891–12,913, https://doi.org/10.1002/2016JD025303, 2016.

Giangrande, S. E., Feng, Z., Jensen, M. P., Comstock, J. M., Johnson, K. L., Toto, T., Wang, M., Burleyson, C., Bharadwaj, N., Mei, F., Machado, L. A. T., Manzi, A. O., Xie, S., Tang, S., Silva Dias, M. A. F., de Souza, R. A. F., Schumacher, C., and Martin, S. T.: Cloud characteristics, thermodynamic controls and radiative impacts during the Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5) experiment, Atmospheric Chemistry and Physics, 17, 14 519–14 541, https://doi.org/10.5194/acp-17-14519-2017, 2017.

Giangrande, S. E., Wang, D., and Mechem, D. B.: Cloud regimes over the Amazon Basin: perspectives from the GoAmazon2014/5 campaign, Atmospheric Chemistry and Physics, 20, 7489–7507, https://doi.org/10.5194/acp-20-7489-2020, 2020.

Greene, D. R. and Clark, R. A.: Vertically Integrated Liquid Water—A New Analysis Tool, Monthly Weather Review, 100, 548–552, https://doi.org/10.1175/1520-0493(1972)100<0548:VILWNA>2.3.CO;2, 1972.

Gupta, S., Wang, D., Giangrande, S., Biscaro, T., and Jensen, M. P.: Lifecycle of updrafts and mass flux in isolated deep convection over the Amazon rainforest: insights from cell tracking, Atmospheric chemistry and physics, 24, 4487–4510, https://doi.org/10.5194/acp-24-4487-2024, 2024.

Harriss, R. C., Wofsy, S. C., Garstang, M., Browell, E. V., Molion, L. C. B., McNeal, R. J., Hoell, Jr, J. M., Bendura, R. J., Beck, S. M., Navarro, R. L., Riley, J. T., and Snell, R. L.: The Amazon Boundary Layer Experiment (ABLE 2A): dry season 1985, Journal of geophysical research, 93, 1351, https://doi.org/10.1029/jd093id02p01351, 1988.

Harriss, R. C., Garstang, M., Wofsy, S. C., Beck, S. M., Bendura, R. J., Coelho, J. R. B., Drewry, J. W., Hoell, Jr, J. M., Matson, P. A., McNeal, R. J., Molion, L. C. B., Navarro, R. L., Rabine, V., and Snell, R. L.: The Amazon Boundary Layer Experiment: Wet season 1987, Journal of geophysical research, 95, 16 721–16 736, https://doi.org/10.1029/jd095id10p16721, 1990.

Jones, C. and Carvalho, L.: Active and break phases in the South American monsoon system, Journal of Climate, 15, 905–914, https://doi.org/10.1175/1520-0442(2002)015<0905:AABPIT>2.0.CO;2, 2002.

Lopes, C.: GoAmazon convective systems datasets (systems and systems_filtered), https://doi.org/10.5281/ZENODO.13732692, 2024.

Machado, L. A. T. and Laurent, H.: The Convective System Area Expansion over Amazonia and Its Relationships with Convective System Life Duration and High-Level Wind Divergence, Monthly Weather Review, 132, 714–725, https://doi.org/10.1175/1520-0493(2004)132<0714:TCSAEO>2.0.CO;2, 2004.

Machado, L. A. T., Dias, M. A. F. S., Morales, C., Fisch, G., Vila, D., Albrecht, R., Goodman, S. J., Calheiros, A. J. P., Biscaro, T., Kummerow, C., Cohen, J., Fitzjarrald, D., Nascimento, E. L., Sakamoto, M. S., Cunningham, C., Chaboureau, J.-P., Petersen, W. A., Adams, D. K., Baldini, L., Angelis, C. F., Sapucci, L. F., Salio, P., Barbosa, H. M. J., Landulfo, E., Souza, R. A. F., Blakeslee, R. J., Bailey, J., Freitas, S., Lima, W. F. A., and Tokay, A.: The Chuva Project: How Does Convection Vary across Brazil?, Bulletin of the American Meteorological Society, 95, 1365–1380, https://doi.org/10.1175/bams-d-13-00084.1, 2014a.

Machado, L. A. T., Silva Dias, M. A. F., Morales, C., Fisch, G., Vila, D., Albrecht, R., Goodman, S. J., Calheiros, A. J. P., Biscaro, T., Kummerow, C., Cohen, J., Fitzjarrald, D., Nascimento, E. L., Sakamoto, M. S., Cunningham, C., Chaboureau, J.-P., Petersen, W. A., Adams, D. K., Baldini, L., Angelis, C. F., Sapucci, L. F., Salio, P., Barbosa, H. M. J., Landulfo, E., Souza, R. A. F., Blakeslee, R. J., Bailey, J., Freitas, S., Lima, W. F. A., and Tokay, A.: The Chuva Project: How Does Convection Vary across Brazil?, Bulletin of the American Meteorological Society, 95, 1365–1380, https://doi.org/10.1175/BAMS-D-13-00084.1, 2014b.

Machado, L. A. T. T., Calheiros, A. J. P. P., Biscaro, T., Giangrande, S., Dias, M. A. F. S., Cecchini, M. A., Albrecht, R., Andreae, M. O., Araujo, W. F., Artaxo, P., Borrmann, S., Braga, R., Burleyson, C., Eichholz, C. W., Fan, J., Feng, Z., Fisch, G. F., Jensen, M. P., Martin, S. T., Poschl, U., Pohlker, C., Pohlker, M. L., Ribaud, J.-F. F., Rosenfeld, D., Saraiva, J. M. B. B., Schumacher, C., Thalman, R., Walter, D., Wendisch, M., Dias, M. A. F. S., Cecchini, M. A., Albrecht, R., Andreae, M. O., Araujo, W. F., Artaxo, P., Borrmann, S., Braga, R., Burleyson, C., Eichholz, C. W., Fan, J., Feng, Z., Fisch, G. F., Jensen, M. P., Martin, S. T., Poschl, U., Pohlker, C., Pohlker, M. L., Ribaud, J.-F. F., Rosenfeld, D., Saraiva, J. M. B. B., Schumacher, C., Thalman, R., Walter, D., Wendisch, M., Dias, M. A. F. S., Cecchini, M. A., Albrecht, R., Andreae, M. O., Araujo, W. F., Artaxo, P., Borrmann, S., Braga, R., Burleyson, C., Eichholz, C. W., Fan, J., Feng, Z., Fisch, G. F., Jensen, M. P., Martin, S. T., Poschl, U., Pohlker, C., Pohlker, M. L., Ribaud, J.-F. F., Rosenfeld, D., Saraiva, J. M. B. B., Schumacher, C., Thalman, R., Walter, D., and Wendisch, M.: Overview: Precipitation characteristics and sensitivities to environmental conditions during GoAmazon2014/5 and ACRIDICON-CHUVA, Atmospheric Chemistry and Physics, 18, 6461–6482, https://doi.org/10.5194/acp-18-6461-2018, 2018.

Makris, A. and Prieur, C.: Bayesian multiple-hypothesis tracking of merging and splitting targets, IEEE transactions on geoscience and remote sensing: a publication of the IEEE Geoscience and Remote Sensing Society, 52, 7684–7694, https://doi.org/10.1109/tgrs.2014.2316600, 2014.

Martin, S. T., Artaxo, P., Machado, L. A. T., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Andreae, M. O., Barbosa, H. M. J., Fan, J., Fisch, G., Goldstein, A. H., Guenther, A., Jimenez, J. L., Pöschl, U., Silva Dias, M. A., Smith, J. N., and Wendisch, M.: Introduction: Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5), Atmospheric Chemistry and Physics Discussions, 16, 30 175–30 210, https://doi.org/10.5194/acpd-15-30175-2015, 2016.

Martin, S. T., Artaxo, P., Machado, L., Manzi, A. O., Souza, R. A. F., Schumacher, C., Wang, J., Biscaro, T., Brito, J., Calheiros, A., Jardine, K., Medeiros, A., Portela, B., de Sá, S. S., Adachi, K., Aiken, A. C., Albrecht, R., Alexander, L., Andreae, M. O., Barbosa, H. M. J., Buseck, P., Chand, D., Comstock, J. M., Day, D. A., Dubey, M., Fan, J., Fast, J., Fisch, G., Fortner, E., Giangrande, S., Gilles, M., Goldstein, A. H., Guenther, A., Hubbe, J., Jensen, M., Jimenez, J. L., Keutsch, F. N., Kim, S., Kuang, C., Laskin, A., McKinney, K., Mei, F., Miller, M., Nascimento, R., Pauliquevis, T., Pekour, M., Peres, J., Petäjä, T., Pöhlker, C., Pöschl, U., Rizzo, L., Schmid, B., Shilling, J. E., Silva Dias, M. A., Smith, J. N., Tomlinson, J. M., Tóta, J., and Wendisch, M.: The Green Ocean Amazon Experiment (GoAmazon2014/5) Observes Pollution Affecting Gases, Aerosols, Clouds, and Rainfall over the Rain Forest, Bulletin of the American Meteorological Society, 98, 981–997, https://doi.org/10.1175/BAMS-D-15-00221.1, 2017.

Nobre, C. A., Obregón, G. O., Marengo, J. A., Fu, R., and Poveda, G.: Characteristics of Amazonian climate: Main features, in: Amazonia and Global Change, Geophysical monograph, pp. 149–162, American Geophysical Union, Washington, D. C., ISBN 9780875904764, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GM000903, 2009.

Petersen, W. A. and Rutledge, S. A.: Regional Variability in Tropical Convection: Observations from TRMM, Journal of climate, 14, 3566–3586, https://doi.org/10.1175/1520-0442(2001)014<3566:RVITCO>2.0.CO;2, 2001.

Rickenbach, T. M., Ferreira, R. N., Halverson, J. B., Herdies, D. L., and Silva Dias, M. A. F.: Modulation of convection in the southwestern Amazon basin by extratropical stationary fronts, Journal of Geophysical Research, D: Atmospheres, 107, LBA 7–1–LBA 7–13, https://doi.org/10.1029/2000JD000263, 2002.

Schumacher, C. and Funk, A.: GoAmazon2014/5 Three-dimensional Gridded S-band Reflectivity and Radial Velocity from the SIPAM Manaus S-band Radar, Tech. rep., Oak Ridge National Lab.(ORNL), Oak Ridge, TN (United States). Atmospheric . . . , https://www.osti.gov/biblio/1459573, 2018.

Silva Dias, M. A. F. and Carvalho, L. M. V.: The South American Monsoon System, in: The Global Monsoon System, vol. 9 of *World Scientific Series on Asia-Pacific Weather and Climate*, pp. 25–33, World Scientific, ISBN 9789813200906, https://doi.org/10.1142/9789813200913_0003, 2016.

Silva Dias, M. A. F., Rutledge, S., Kabat, P., Silva Dias, P. L., Nobre, C., Fisch, G., Dolman, A. J., Zipser, E., Garstang, M., Manzi, A. O., Fuentes, J. D., Rocha, H. R., Marengo, J., Plana-Fattori, A., Sá, L. D. A., Alvalá, R. C. S., Andreae, M. O., Artaxo, P., Gielow, R., and Gatti, L.: Cloud and rain processes in a biosphere-atmosphere interaction context in the Amazon Region, Journal of geophysical research, 107, S1, https://doi.org/10.1029/2001jd000335, 2002.

Souza, C. M. A., Dias-Júnior, C. Q., D'Oliveira, F. A. F., Martins, H. S., Carneiro, R. G., Portela, B. T. T., and Fisch, G.: Long-term measurements of the Atmospheric Boundary Layer height in Central Amazonia using remote sensing instruments, Remote sensing, 15, 3261, https://doi.org/10.3390/rs15133261, 2023.

Tian, Y., Zhang, Y., Klein, S. A., and Schumacher, C.: Interpreting the diurnal cycle of clouds and precipitation in the ARM GoAmazon observations: Shallow to deep convection transition, Journal of geophysical research, 126, https://doi.org/10.1029/2020jd033766, 2021.

Uba, D. M., Negri, R. G., Enoré, D. P., Costa, I. C. d., and Jorge, A. A. S.: TATHU - Software para rastreio e análise do ciclo de vida de sistemas convectivos, http://urlib.net/ibi/8JMKD3MGP3W34T/47AF772, 2022.

Yuter, S. E. and Houze, R. A.: Three-Dimensional Kinematic and Microphysical Evolution of Florida Cumulonimbus. Part II: Frequency Distributions of Vertical Velocity, Reflectivity, and Differential Reflectivity, Monthly Weather Review, 123, 1941–1963, https://doi.org/10.1175/1520-0493(1995)123<1941:tdkame>2.0.co;2, 1995.

Zhou, J. and Lau, K.-M.: Does a monsoon climate exist over South America?, Journal of climate, 11, 1020–1040, https://doi.org/10.1175/1520-0442(1998)011<1020:damceo>2.0.co;2, 1998.