



Deep convection life cycle characteristics: a database from the GoAmazon experiment

Camila da Cunha Lopes¹, Rachel Ifanger Albrecht¹, Douglas Messias Uba², Thiago Souza Biscaro², and
Ivan Saraiva³

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

²Divisão 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, Brazil

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

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

Received: 26 September 2024 – Discussion started: 30 October 2024

Revised: 17 July 2025 – Accepted: 31 July 2025 – Published: 17 October 2025

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 a full life cycle, incorporating lightning data. This refined deep convection database is shown to be a robust sample of the complete dataset in terms of convective system morphology. The analysis reveals significant seasonal and diurnal variations in convective morphology and intensity, with most intense systems occurring during the dry-season–wet-season transition. The filtered dataset offers a robust sample for future studies on Amazonian convection. The database is available at <https://doi.org/10.5281/zenodo.13732692> (Lopes, 2024).

1 Introduction

Due to its large territorial extent that includes pristine forest, agricultural regions, and a large urban zone with an industrial center, the Amazon tropical rainforest serves as a natural test bed for studies on cloud–aerosol–precipitation and land–atmosphere interactions. 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 Amazonian wet (austral summer) and dry (austral winter) seasons.

Several field experiments have been conducted in the region in order to study different aspects of cloud–aerosol–precipitation interactions in the Amazon. The Amazon

Boundary Layer Experiment campaigns ABLE 2A (Harriss et al., 1988) and 2B (Harriss et al., 1990) focused on the chemistry and dynamics of the lower atmosphere in the dry and wet seasons, respectively. The Large-Scale Biosphere–Atmosphere Experiment in Amazonia (LBA) program (Silva Dias et al., 2002) was 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 (Cloud Processes of the Main Precipitation Systems in Brazil) project, a contribution to cloud-resolving modeling and to the Global Precipitation Measurement (GPM) mission (Machado et al., 2014a). The Observations and Modeling of the Green Ocean Amazon (GoAmazon2014/5; Martin et al., 2017) campaign was the first long-term experiment to analyze the effects of the Manaus pollution plume at different experimental sites

around Manaus, and it included two intensive operation periods (IOPs) during the wet and dry season. Unlike previous experiments, an operational weather radar, operated by Sistema de Proteção da Amazônia (SIPAM – System for the Protection of the Amazon), was available during GoAmazon, and it enabled coverage of all experimental sites with cloud remote-sensing data at great temporal and spatial resolutions. The main goal of the GoAmazon experiment was to analyze cloud–aerosol–precipitation interactions between the forest and the Manaus metropolitan region, specifically the transformation of air plumes from the pristine forest to the Manaus pollution plumes and their subsequent eastern propagation. The goal of the CHUVA-Manaus experiment was to characterize the convection regimes in the region using remote sensing, with the installation of an X-band radar during the experiment and the establishment of partnerships with CENSIPAM (Centro Gestor e Operacional do Sistema de Proteção da Amazônia – Manager and Operational Center of the System for the Protection of the Amazon) for surface radar data and NASA–JAXA (National Aeronautics and Space Administration–Japan Aerospace Exploration Agency) for satellite radar data from TRMM and GPM.

A few studies, such as Giangrande et al. (2017), Machado et al. (2018), Giangrande et al. (2020), and Biscaro et al. (2021), have provided insight into the convection characteristics during GoAmazon. Each of them use different definitions of cloud features employing different sources of remote-sensing data, and they even diverge with respect to their definition of the wet and dry seasons, thereby making it difficult for other studies to follow a homogeneous methodology of convection measurements, especially those that analyze cloud–aerosol–precipitation interactions. For this reason, this study aims to create a comprehensive database and methodology for convective systems based on radar data that can be used in future studies utilizing GoAmazon data.

2 Materials and methods

2.1 Data

Data sources for this study are the GoAmazon (Martin et al., 2017) and CHUVA-Manaus (Machado et al., 2014b) field experiments that occurred between January 2014 and December 2015 around Manaus, Amazonas. The main site of these experiments, named T3, was located in Manacapuru, Amazonas, Brazil (3.213° S, 60.598° W), about 70 km west of Manaus. A wide range of cloud, precipitation, aerosol, and atmospheric instruments were installed at the site, as part of the ARM (Atmospheric Radiation Measurement) mobile facility AMF1 that took measurements during most of the experiment. Some additional instruments and other sites (shown in Fig. 1 of Martin et al., 2017) took measurements during two intensive operation periods (IOPs), with IOP1 being between 1 February and 31 March 2014 (wet season) and IOP2 being between 15 August and 15 October 2014 (dry

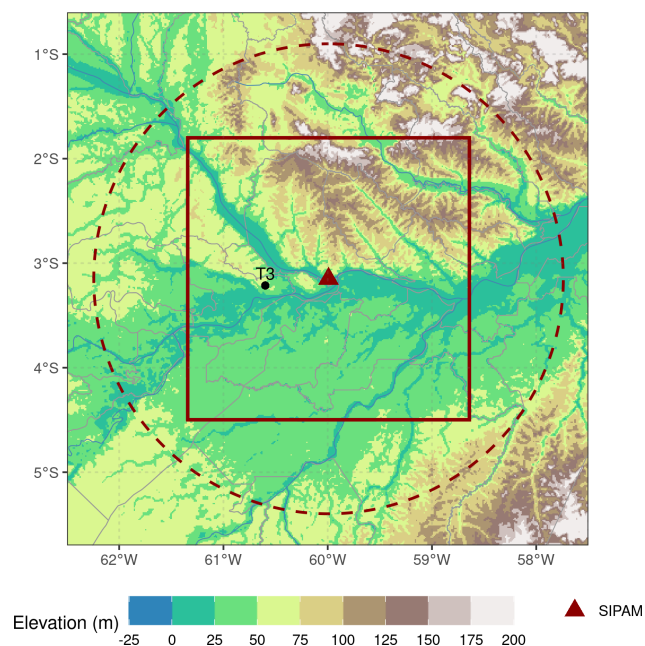


Figure 1. Amazonian region used in this study, showing the SIPAM radar location (red triangle), the 250 km originally calculated coverage (dashed red lines), the 150 km $\times$ 150 km bounding box (red square) used in the study, and the T3 site location in Manacapuru where surface data (not shown in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

season). More details about the campaigns can be found in Martin et al. (2016, 2017).

In order to create the convective system database, radar volumes from the SIPAM (Sistema de Proteção da Amazônia – System for the Protection of the Amazon) single-polarization S-band radar located in Manaus (3.149° S, 59.991° W, 102.4 m altitude, Fig. 1) were selected during the period of the GoAmazon experiment (1 January 2014 to 31 December 2015). These volumes consist of CAPPis (constant-altitude plan position indicators) with a bias correction calculated by Schumacher and Funk (2018); Biscaro (2016). Table 1 shows the data settings, including the bias values applied throughout the period.

Given the coarse characteristics of this radar and its scan strategy, such as varying numbers of elevation sweeps and antenna vibrations between volumes, the radar data used in this study consist of quality-controlled, bias-corrected CAPPis, gridded onto a 1 $\times$ 1 km horizontal resolution, following previous studies using the same source of data (e.g., Gupta et al., 2024). Following the approach of Saraiva et al. (2016), we restricted our analysis to a 150 $\times$ 150 km box centered on the radar location. This domain allows for consistent and reliable detection of convective structures while also minimizing the effects of beam broadening, ground clutter, and incomplete vertical sampling at farther ranges.

A second data source was employed to calculate parameters related to lightning activity in the convective systems.

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

Type	CAPPIs
Format	Binary, 15 × 500 × 500 elements
Resolution (vertical, horizontal, temporal)	1 × 1 × 1 km, 12 min
Min, max height	2 km, 16 km
CAPPI processing software	RSL (Radar Software Library)
Bias correction (Schumacher and Funk, 2018)	3 Jan to 5 Feb 2014: +1.0 dB 6 Feb to 19 Aug 2014: +3.0 dB 20 Aug to 16 Nov 2014: −2.5 dB 17 Oct 2014 to 6 Mar 2015: −5.5 dB 7 Mar to 5 Jul 2015: −4.0 dB 6 Jul to 28 Oct 2015: −1.5 dB 29 Oct to 31 Dec 2015: +1.0 dB

The Vaisala Global Lightning Detection Network (GLD360; Demetriades et al., 2010) measures cloud-to-ground return strokes in the VLF (very low frequency) range using triangulation techniques, magnetic direction finding (MDF), and time of arrival (TOA), as well as lightning recognition algorithms. Limitations of this network include the detection efficiency in Brazil, which is 70 % in average (Naccarato et al., 2010) but can be significantly lower in areas with deficient coverage such as northern Brazil, including the Amazon region, as well as the definition of a measured stroke as a single or multiple real return strokes (Murphy and Nag, 2015). Stroke data were accumulated over 12 min using the same time stamps as the radar data and were selected within the cluster polygons delimited by the tracking algorithm.

2.2 Tracking methodology

The convective system 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 Uba (2025a), and it 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 of pixels – hereafter referred to as *clusters* of storm cells – in an input field at a single time step according to the (one or more) threshold(s) and extracts their statistics, such as size (in pixels) and mean and maximum values of the input field. From subsequent time steps, based on spatial overlap, it tracks and names (via a universal unique identifier – *uuid*) the *convective systems* that occurred in the period and their status during the described 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). The identification of a split or a merge depends mainly on the cluster’s propagation between subsequent time steps, where discrepancies can occur in situations such as clusters generating within the same relative overlap area of other propagating clusters. This split/merge problem is common to several tracking algorithms that use area overlap strategies, such as SCIT (Storm Cell Identification and Tracking Algorithm; Johnson et al., 1998) and TITAN (Thunderstorm Identification, Tracking, Analysis, and Nowcasting; Dixon and Wiener, 1993).

Table 2 shows the TATHU settings used in this study. Based on several considerations, we selected SIPAM radar reflectivity greater than 20 dBZ at the 3 km CAPPI field as the input field to the TATHU tracking algorithm. First, this altitude lies above the typical cloud base in the region (Fisch et al., 2004; Souza et al., 2023) and coincides with the level at which maximum reflectivity is frequently observed in convective precipitation cores below freezing layer. Second, lower elevations near the radar are significantly affected by ground clutter and beam blockage (Giangrande et al., 2016; Schumacher and Funk, 2018), compromising the reliability of reflectivity data closer to the surface. Additionally, the SIPAM radar’s relatively coarse resolution limits the ability to resolve vertical structures accurately, particularly storm tilts. However, in the tropical environment of the Amazon, vertical wind shear is generally weak, reducing the likelihood of significant storm tilt or vertical displacement of convective cores. These factors together support the use of a horizontal 3 km CAPPI as a consistent and suitable input for identifying and tracking convective systems in this study. Other tracking algorithms consider composite reflectivity (e.g., Heikenfeld et al., 2019; Sokolowsky et al., 2024) instead of 2D reflectivity fields to offer a more comprehensive depiction of the 3D structure of convective systems. However, considering the context of our study and the nature of this dataset (exposed above), the 3D structure of the convective systems should be captured. A comprehensive study of the implica-

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

Input data	3 km CAPPI in a 150 × 150 km box (300 × 300 elements) centered at the radar location, between January 2014 and December 2015
Reflectivity thresholds (<i>threshold value</i>)	Values greater than 20 and 40 dBZ
Minimum cluster sizes (<i>minarea</i>)	100 and 40 km ²
Tracking technique (<i>trackers class</i>)	Relative overlap area (<i>RelativeOverlapAreaStrategy</i>)
Minimum cluster overlap area	10 %
Maximum interval between images	60 min
Statistics (<i>stats</i>)	Maximum, mean, and standard deviation of reflectivity; size (number of pixels); and number of layers (corresponding to having one or two reflectivity and minimum size thresholds)
Output	PostGIS database (<i>systems</i> and <i>systems_filtered</i> tables)
Filters applied	– There is at least one 40 dBZ/40 km² (corresponding to deep convection) pixel in any time stamp – Does not intersect the bounding box of the grid (corresponding to part of the cluster probably being outside the tracking region) – Lasts longer than 12 min (one time stamp) or has a relation with another convective system (by splitting or merging)

tions with respect to the reflectivity thresholds for tracking convective systems in CAPPI fields can be found in Leal et al. (2022) but is not within the scope of this study. Moreover, our approach is methodologically consistent with previous studies that have successfully employed single-level CAPPI fields to characterize convective systems in the Amazon (e.g., Laurent and Laurent, 2002; Albrecht et al., 2011; Leal et al., 2022; Gupta et al., 2024). These studies have provided valuable insights into storm morphology, evolution, and vertical structure using similar techniques. Adopting a comparable methodology ensures continuity and comparability with the existing body of literature on Amazonian convection.

The relative overlap area strategy considers two subsequent clusters in time as the same convective system when there is at least a 10 % overlap between their areas (i.e., polygons). The maximum interval between images considers a data gap sufficient to ensure the continuity of the convective systems, but it can result in different convective systems being tracked as the same system if they are in the overlap area, considering that the average life cycle of tropical convection is less than 60 min. Within the main statistics, clusters with a number of threshold layers equal to 0 (with only the 20 dBZ reflectivity threshold) or 1 (with both the 20 and 40 dBZ reflectivity thresholds) can be used to separate systems with or without deep convective cores.

For better illustration of the terminology of the TATHU tracking system applied to SIPAM 3 km CAPPI fields, Fig. 2 shows a schematic of these definitions: a *cluster* is a contiguous region (polygon) of pixels within the 3 km CAPPI re-

flectivity field that exceeds the 20 dBZ reflectivity threshold *identified at a single time step*; a *convective system* is a time-continuous sequence of clusters that are linked across radar volumes based on spatial overlap. A *cluster* can have one or several cores that exceed the 40 dBZ reflectivity threshold, an estimation of deep convection, but these cores are considered to be part of a single cluster during the description and tracking of the convective systems.

Output data were stored in a PostGIS database, a database management system (DBMS) with geospatial support that allows for the storage of large volumes of data in a tabular format, including geolocated geometries. The database was converted to GeoJSON datasets in order to make it easily available in Lopes (2024).

Two datasets were defined based on the TATHU tracking. The original – hereafter referred to as the *raw* dataset – contains all of the convective systems observed in the period, regardless of duration, size, or status during the life cycle. In contrast, 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 km² minimum area) is used here as a definition of deep convective cores present during the life cycle. Clusters intersecting the borders of the grid had their associated convective systems discarded to exclude systems without a full life cycle within the study area. Systems with only one time step were also discarded. We created this filtered dataset to provide a subset of deep convective systems with a full life cycle, an important criteria for several convection studies.

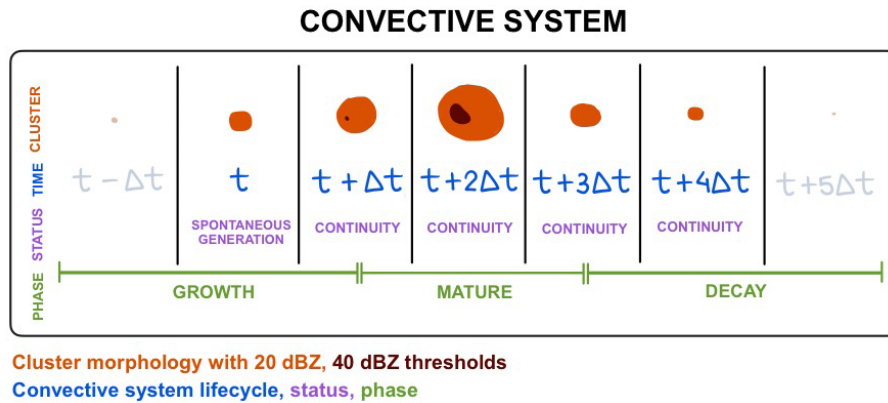


Figure 2. Illustrative deep convective system as observed by TATHU: cluster morphology with 20 dBZ (orange) and 40 dBZ (dark red) thresholds in five valid subsequent time stamps with the TATHU status (spontaneous generation or continuity) and different phases (growth, mature, and decay) of a full life cycle.

The original *systems* dataset contains 91 609 convective systems and 322 896 clusters, while the filtered *systems_filtered* dataset contains 5976 convective systems (6.5 % of the original dataset) and 40 394 clusters (12.5 % of the original dataset). Using the filtered dataset with the additional lightning data, 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:

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

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

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

3 Results

3.1 Raw and filtered dataset characteristics

This subsection compares the physical characteristics between the raw and filtered datasets in order to define what type of convective system each dataset represents. Figure 3 shows the size – represented by area (in km²) – distribution of clusters and the maximum size of convective systems. Both distributions have a maximum frequency in the smallest (1000 km²) area. The maximum area of the raw systems is 6 times larger (60 000 km² vs. 10 000 km²) and the distribution drops faster in the filtered systems. These characteristics show the filtering effect around the border of the 150 km bounding box (Fig. 1), which excluded very large clusters. A

200 × 200 point cluster with a 40 000 km² area, for example, can be considered too large because it occupies two-thirds of the grid and probably intercepts the border of the bounding box in a given time stamp.

Figure 4 shows the mean and maximum cluster reflectivity distributions for the raw (Fig. 4a) and filtered (Fig. 4b) datasets. On raw data, the distributions show no significant frequency peaks, with the mean reflectivity distributed mainly (frequency above 20 %) between 20 and 40 dBZ and the maximum reflectivity between 35 and 50 dBZ (frequency above 15 %). Conversely, 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 those 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 the raw and filtered datasets using the cluster 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 of these systems had their full life cycle covered, whereas on the filtered data, all of the systems had their full life cycle covered (as one of the filtering criteria is to exclude convective systems that leave the radar coverage area). A total of 53 % of the filtered convective systems split or merged during their life cycle, 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 raw convective systems did not last long enough to go through a split or merge. The majority (80 %) of the filtered and half (49 %) of the raw convective systems were

Table 3. Clusters’ additional parameters calculated on the filtered dataset.

Variable	Description	Reference
gld	GLD strokes within a 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; 15×16 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 the cluster (in kg m^{-2}), Eq. (1)	
vii_kgm2	VII (vertically integrated ice) of the cluster (in kg m^{-2}), Eq. (2)	Petersen and Rutledge (2001)
vil_kgm2	VIL (vertically integrated liquid) of the cluster (in kg m^{-2}), Eq. (3)	Greene and Clark (1972)
nae_s_1	Normalized area expansion (in s^{-1}) between clusters of the same convective system	Machado and Laurent (2004)
gld_strmin	Stroke rate between clusters of the same convective system (in strokes per minute)	
echotop0_kmmin, echotop20_kmmin, echotop40_kmmin	0, 20, and 40 dBZ echo top rate between clusters of the same convective system (in km min^{-1})	

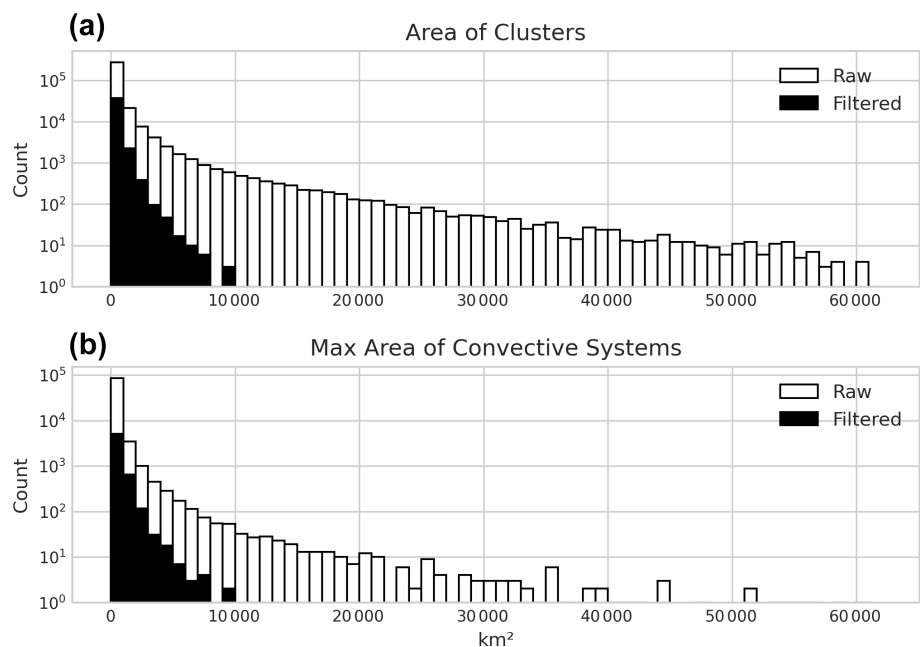


Figure 3. Distribution of clusters’ area (a) and convective systems’ max area (b) for the raw (white) and filtered (black) datasets.

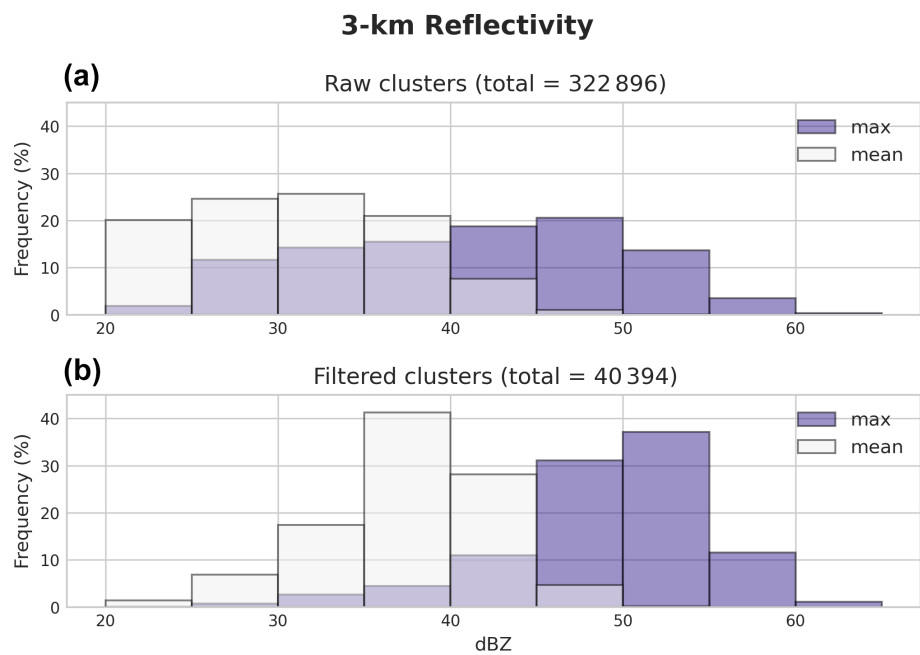


Figure 4. Distribution of clusters’ max (purple) and mean (white) reflectivity for raw (a) and filtered (b) datasets.

Table 4. Frequency of spontaneously generated (did not generate from splits or the first steps of algorithm rounds) convective systems (CSs), split and/or merged CSs, and full-life-cycle (last time step in continuity) CSs for the raw and filtered datasets.

	Raw	Filtered
CSs (total)	91 609	5976
CSs spontaneously generated (%)	72	73
CSs with split/merge (%)	37	53
CSs with full life cycle (%)	49	80

considered with their full life cycle (last time step was “continuity”), which is also explained by the high percentage of raw systems with only one time step (i.e., only “spontaneous generation” status). The 20 % of filtered convective systems without a full life cycle are within the not spontaneously generated systems – products of a split or a merge with other systems.

Table 5 shows the distribution of convective system durations for the raw and filtered datasets. The majority of the raw (40 %) and filtered (47 %) systems lasted up to 1 h, indicating the predominance of isolated convective systems (Giangrande et al., 2023; Viscardi et al., 2024; Gupta et al., 2024). Only 10 % of the raw systems lasted between 1 and 3 h, compared to 36 % of filtered systems. The majority of convective systems were short-lived, with 40 % of the raw and 47 % of the filtered systems lasting up to 1 h, highlighting the prevalence of isolated convective systems. A notable difference is observed in the 1–3 h range, where only 10 % of the raw systems persisted, compared to 36 % in the filtered dataset.

Table 5. Frequency of spontaneously generated (did not generate from splits or the first steps of algorithm rounds) CSs, split and/or merged CSs, and full-life-cycle (last time step in continuity) CSs for the raw and filtered data tables.

	Raw	Filtered
CS duration ≤ 1 h	79 301 (86.6 %)	3448 (57.7 %)
1 h < CS duration ≤ 2 h	6950 (7.6 %)	1535 (25.7 %)
2 h < CS duration ≤ 3 h	2597 (2.8 %)	635 (10.6 %)
3 h < CS duration ≤ 4 h	1170 (1.3 %)	239 (4 %)
4 h < CS duration ≤ 5 h	641 (0.7 %)	74 (1.2 %)
5 h < CS duration ≤ 6 h	318 (0.3 %)	27 (0.5 %)
CS duration > 6 h	631 (0.6 %)	18 (0.3 %)
Total	91 609 (100 %)	5976 (100 %)

For long-duration systems, nearly twice as many raw systems (631) lasted longer than 6 h compared to those in the 5–6 h range (318). In contrast, the filtered dataset shows fewer systems exceeding 6 h (18) than those lasting between 5 and 6 h (27), indicating a stronger effect of the area filtering on prolonged convective systems.

Figure 5 shows the monthly distribution of the 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-season–wet-season transition between November and December, wet season between January and March, IOP1 between February 1 and March 31 2014, and IOP2 between 5 August and 15 October 2014. In general, a larger frequency of clusters and convective systems occurred in the wet and

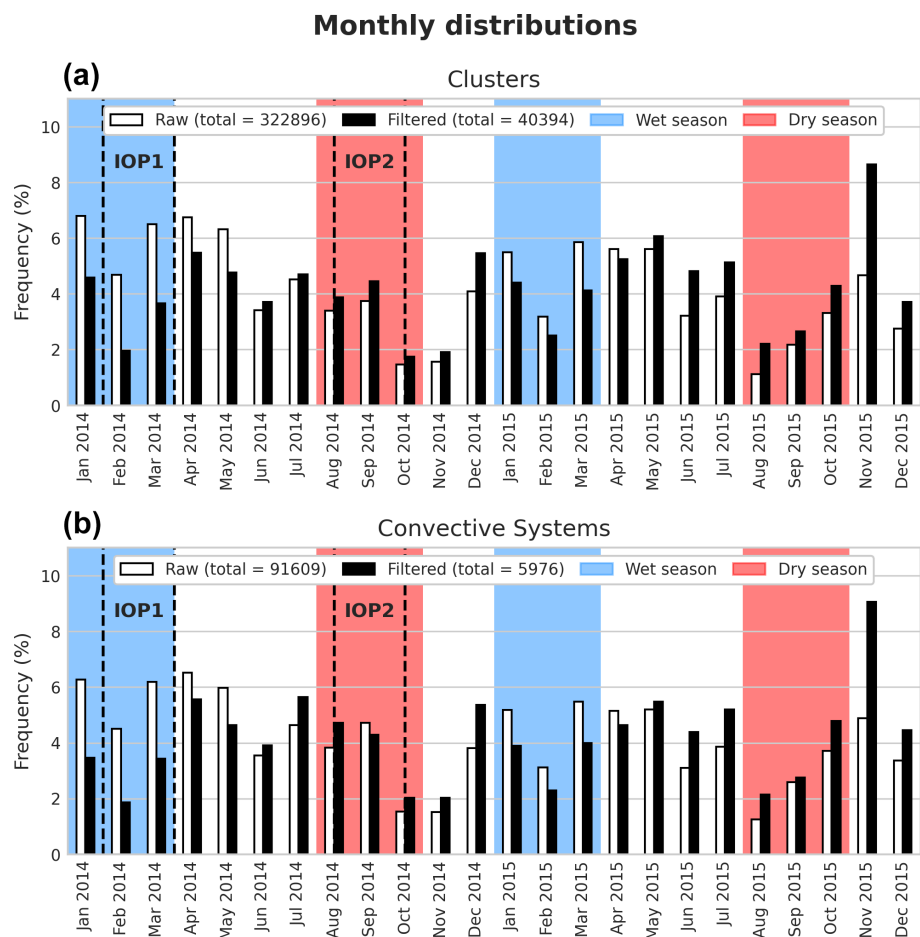


Figure 5. Monthly relative frequency distribution of clusters **(a)** and convective systems **(b)** for the raw (white) and filtered (black) datasets. The blue and red areas delimit wet and dry seasons, respectively, and dashed lines delimit IOP1 and IOP2.

Table 6. Distribution of convective system (CS) durations for the raw and filtered datasets separated by IOP1 and IOP2.

	IOP1		IOP2	
	Raw	Filtered	Raw	Filtered
CS duration ≤ 1 h	8427 (86.1 %)	176 (55.7 %)	6565 (90 %)	272 (55.5 %)
1 h < CS duration ≤ 2 h	772 (7.9 %)	88 (27.8 %)	432 (5.9 %)	130 (26.5 %)
2 h < CS duration ≤ 3 h	277 (2.8 %)	34 (10.8 %)	146 (2 %)	53 (10.8 %)
3 h < CS duration ≤ 4 h	118 (1.2 %)	10 (3.2 %)	60 (0.8 %)	24 (4.9 %)
4 h < CS duration ≤ 5 h	66 (0.7 %)	6 (1.9 %)	41 (0.6 %)	7 (1.4 %)
5 h < CS duration ≤ 6 h	40 (0.4 %)	0 (0 %)	19 (0.3 %)	2 (0.4 %)
CS duration > 6 h	91 (0.9 %)	2 (0.6 %)	31 (0.4 %)	2 (0.4 %)
Total	9792 (100 %)	316 (100 %)	7294 (100 %)	490 (100 %)

transition (wet-to-dry and dry-to-wet) seasons, with a peak in the filtered clusters/systems in November 2015 (almost 2 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 in the wet seasons and the opposite in the dry seasons. This difference can be explained by the filtering 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 would occur during the wet season than during the dry season, which was the case for both the raw and filtered data.

Table 6 presents the duration of the raw and filtered convective systems by IOP. The distribution of system durations

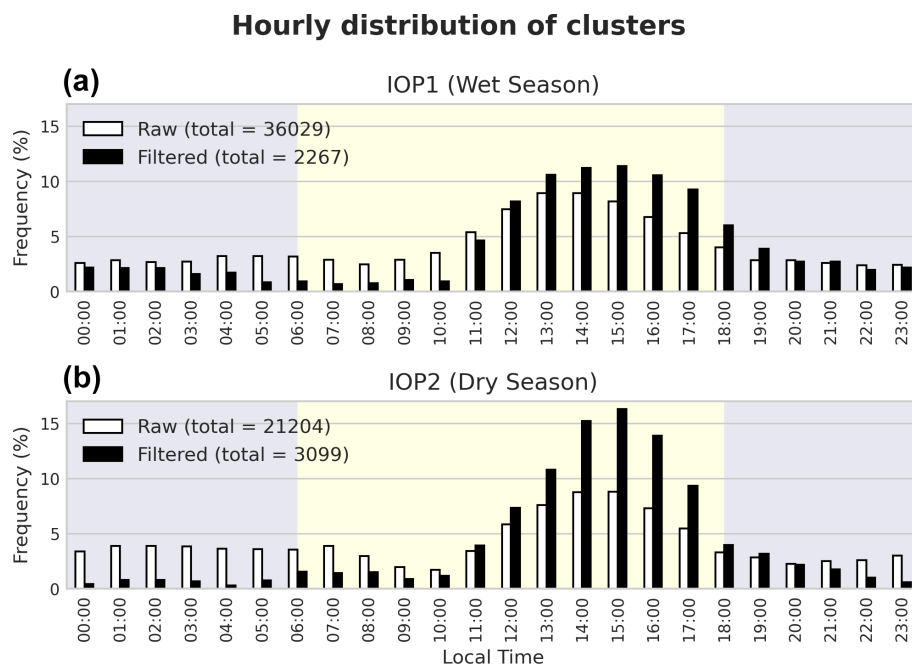


Figure 6. Hourly relative distribution of clusters for the raw (white) and filtered (black) datasets separated by IOP1 (a) and IOP2 (b). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

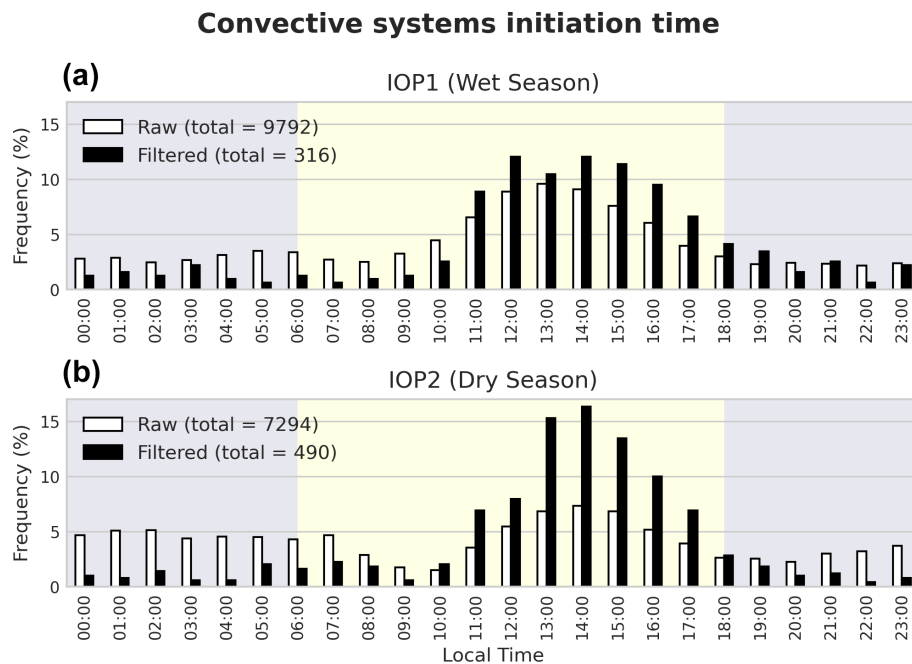


Figure 7. Hourly relative distribution of convective system initiation time for the raw (white) and filtered (black) datasets separated by IOP1 (a) and IOP2 (b). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

is generally consistent across seasons, with similar proportions observed between the raw and filtered datasets. In contrast to Table 5, where less than half of the raw systems lasted up to 1 h, seasonal distributions show that more than 80 % of the raw systems and 50 % of filtered systems persisted for

no more than 1 h, indicating the predominance of isolated convective systems. Additionally, 97 % of raw and 94 % of filtered systems had durations of 3 h or less.

Figure 6 shows the hourly distribution of the raw and filtered clusters by IOP. Comparing the raw and filtered clus-

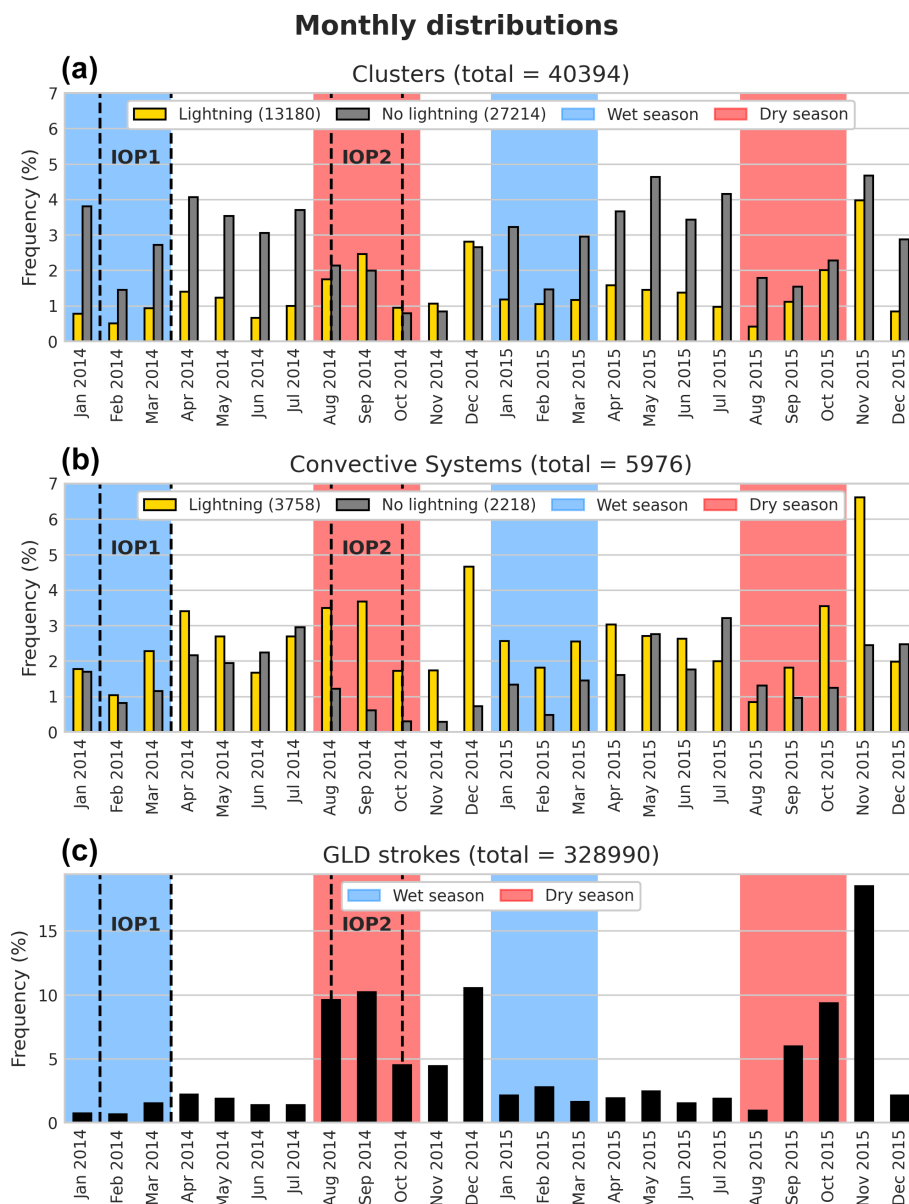


Figure 8. Monthly relative distribution of clusters (a), convective systems (b), and GLD strokes (c) for filtered datasets separated by lightning (yellow bars) or no lightning (gray bars) occurrence. The blue and red areas delimit wet and dry seasons, respectively, and dashed lines delimit IOP1 and IOP2.

ters, in all seasons, there is a larger frequency of raw clusters during the late-night/dawn period and a larger frequency of filtered clusters during late-morning/afternoon period. In the dry season (IOP2), filtered clusters are more frequent (above 15 %) between 14:00 and 15:00 LT (local time) compared to the raw clusters (below 10 %), while the raw clusters are 5 % more frequent than the filtered clusters between 23:00 and 07:00 LT. These differences between the 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 6 h) clusters described previously.

Figure 7 shows the hourly distribution of the raw and filtered convective system initiation (defined here as the first time stamp) by IOP. Comparing the raw and filtered systems, in both seasons, there is a larger frequency of raw systems initiating at dawn, whereas a larger frequency of filtered systems initiate during the morning/afternoon. This difference is even greater (almost 10 %) in the dry season. Comparing the seasons, in both the dry and wet seasons, the initiation peak occurs only at 14:00 LT. Specifically regarding the filtered systems, the 14:00 LT peak during the wet season is not highlighted in IOP1, with another peak at 12:00 LT.

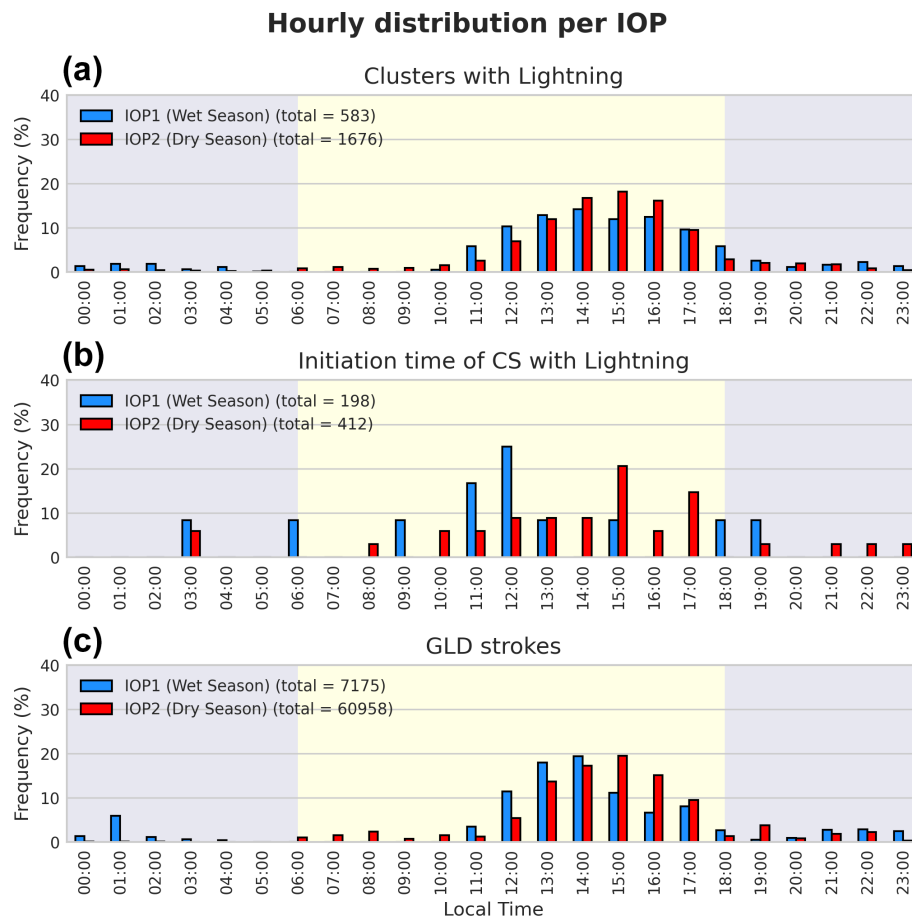


Figure 9. Hourly relative distribution of clusters with lightning (a), initiation time of clusters with lightning (b), and GLD strokes (c) for filtered dataset during IOP1 (blue) and IOP2 (red). The blue and yellow areas delimit diurnal and nocturnal periods, respectively.

When comparing the physical characteristics of the raw and filtered datasets, the filters significantly decrease the sampling size but define a specific subset of convective systems: in general, they have areas of up to 10 000 km², a maximum reflectivity mainly above 50 dBZ, a full life cycle (growth, mature, and decay phases), and last up to 3 h. The raw dataset includes these same systems as well as diverse convective types, such as mesoscale (considering the areas above 50 000 km²), both short-lived (12 min duration) and long-lived (duration above 6 h), and stratiform (reflectivity below 30 dBZ) systems. Both datasets are useful for further convection studies, but they should be chosen carefully depending on the study objectives. If the convection phase is important, for example, the filtered dataset would be more appropriate, whereas some studies might benefit from convection patterns in the Amazon region that can be extracted from the raw dataset.

3.2 Characteristics specific to the filtered systems

Starting with the monthly distribution, Fig. 8 shows the clusters and convective systems separated by lightning activity as

well as the distribution of GLD strokes. More than double the clusters (27 214 vs. 13 180) had no electrical activity, while more convective systems (3758 vs. 2218) showed lightning, which means that, in general, the convective systems with lightning consist of only a few clusters with lightning. A larger frequency of clusters and systems without lightning occurs between wet and dry seasons, while peaks in cluster and system frequency occur in the dry-season–wet-season transition, comparable with the peak stroke frequency. These findings are similar to what has been found in previous work such as Albrecht et al. (2011, 2016).

Separating the analysis into systems with and without lightning, Fig. 9 shows the hourly distribution of clusters and the initiation (first time stamp) of convective systems with lightning and GLD strokes by IOP. All variables are more frequent during the late-morning/afternoon period, with significant differences between IOPs. In IOP2, the cluster and stroke peaks occur at 15:00 LT, which is also the preferential time for convective system initiation. In IOP1, the cluster and stroke peaks occur around the same time, but system initiation occurs preferentially earlier, at 12:00 LT. The

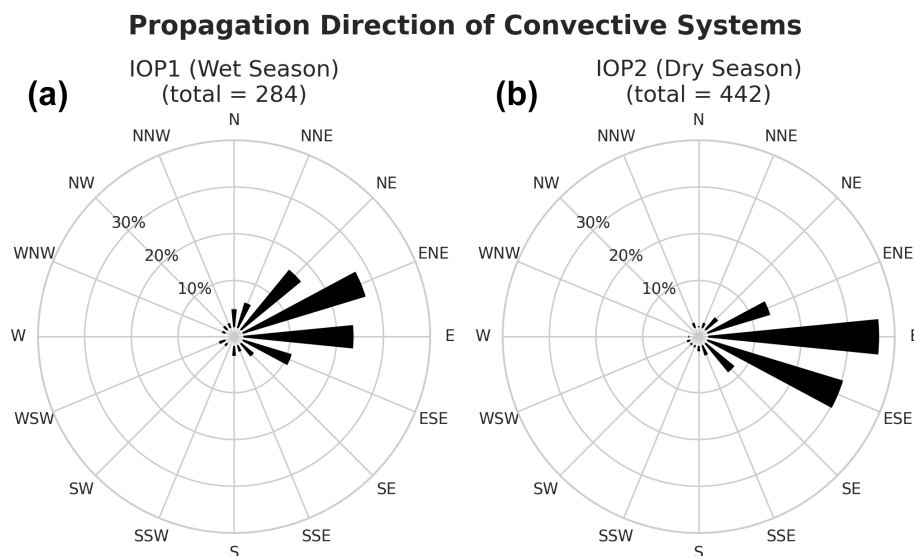


Figure 10. Convective systems' relative propagation direction distribution in the filtered dataset separated by IOP1 (a) and IOP2 (b). The direction was defined by the distance between the first and last centroids of the convective system.

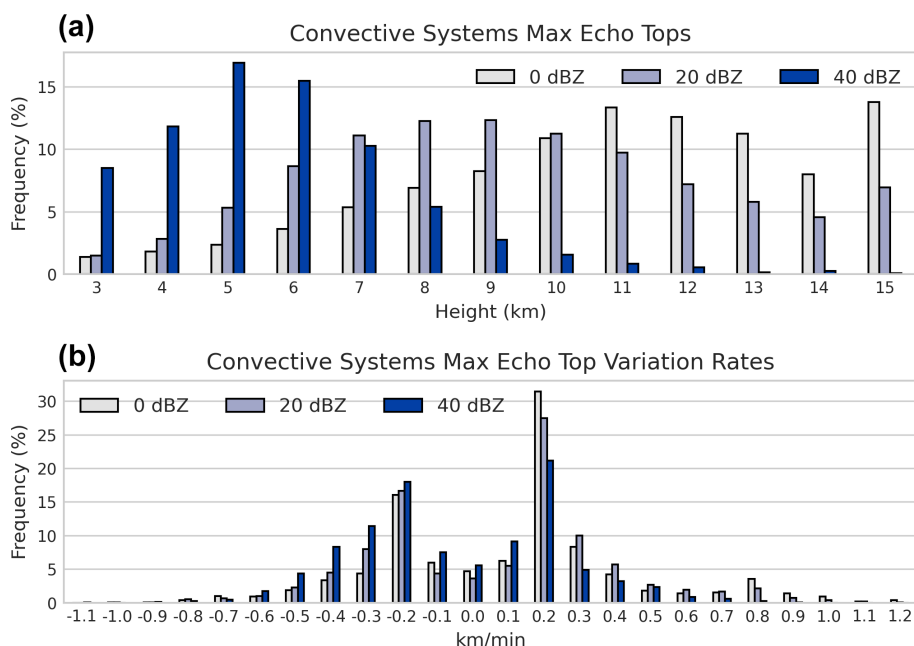


Figure 11. Convective systems' 0 (gray bars), 20 (light-blue bars), and 40 dBZ (dark-blue bars) max echo tops (a) and max echo top variation rates (b).

low number of systems (198 in IOP1 and 412 in IOP2) influenced the frequency distribution, but the results were similar to their corresponding seasons (not shown). Convective systems without lightning are more frequent during the late-morning/afternoon period (not shown). The initiation frequency peaks in the dry and wet seasons occur at 15:00 LT, whereas during the dry-season–wet-season transition, the distribution is more dispersed, with approximately equal peaks at 12:00, 13:00, and 16:00 LT.

In order to analyze the propagation direction of the convective systems, Fig. 10 shows the frequency distribution of the propagation direction by IOP. The predominant direction of the convective systems is from the east, consistent with the main dynamic forcing in the Amazon, which is a moisture flux from the tropical Atlantic due to the easterlies that is influenced by the position of the Intertropical Convergence Zone (ITCZ) (Silva Dias and Carvalho, 2016). Comparing the IOPs, in the wet season, about 25 % of systems propa-

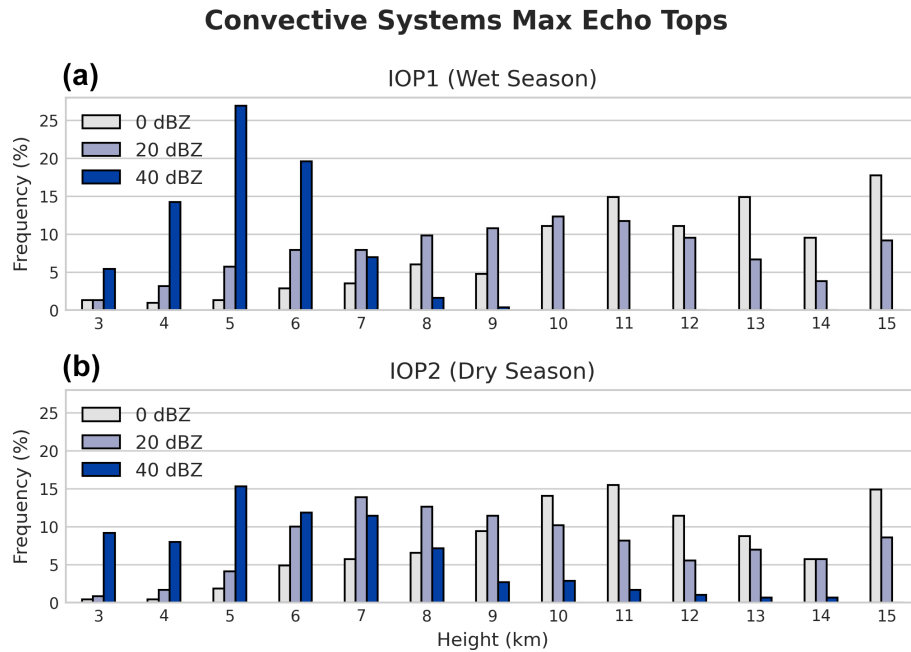


Figure 12. Convective systems' 0 (gray bars), 20 (light-blue bars), and 40 dBZ (dark-blue bars) max echo tops separated by IOP1 (a) and IOP2 (b).

gated from the east and east-northeast, whereas in the dry season, more than 30 % of systems propagated from the east and east-southeast. This direction shift from east-northeast to east-southeast is related to the shift in the position of the ITCZ and cold-front propagation over the South American continent, which affect the zonal wind regime in the Amazonian region (Rickenbach et al., 2002). These results are also consistent with Gupta et al. (2024), who focused on isolated convective systems near the Manacapuru (T3) site.

In order to analyze the intensity of the convective systems, Fig. 11 shows the frequency distribution of the maximum heights and variation rates (considering the 12 min interval between radar scans) of the 0, 20, and 40 dBZ echoes, associated with cloud-top height, precipitable-hydrometeor height, and intense-precipitation height, respectively. The maximum cloud-top heights are more frequent above 10 km, with peaks at 11 km (dry season) and 15 km (wet season). The maximum precipitable-hydrometeor 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 at 5 km (in both the dry and wet seasons). These high tops, complemented by a maximum precipitation height more frequently between 7 and 11 km and a maximum intense-precipitation height more frequently between 4 and 7 km, show how these systems were predominantly deep at their most intense moments. The variation rates (Fig. 11b) were similar between the echoes, with frequency peaks at -0.2 and 0.2 km min^{-1} , indicating significant fluctuations in the echo tops throughout their life cycle.

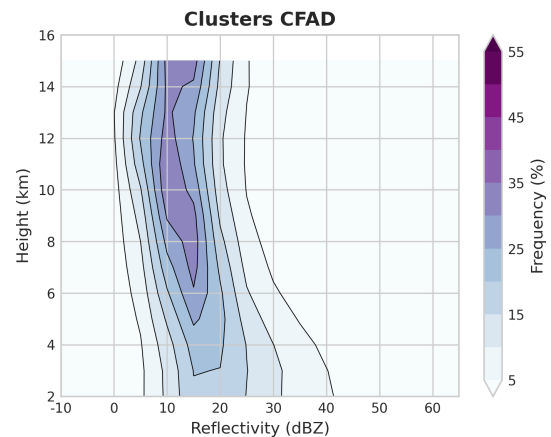


Figure 13. Mean contoured frequency by altitude diagram (CFAD) of clusters' reflectivity with 5 dBZ bins.

Figure 12 shows the frequency distribution of the maximum echo tops of 0, 20, and 40 dBZ by IOP. As in the complete time series, the maximum top heights are more frequent above 10 km, with peaks at 11 km (IOP2) and 15 km (IOP1). The maximum precipitable height was also more frequent between 7 and 11 km. The maximum intense precipitation height was also more frequently located between 4 and 7 km, with peaks at 5 km.

In order to analyze the clusters' vertical profiles, Fig. 13 shows the frequency diagram by altitude of the cluster reflectivity of the complete time series. The most frequent profile is of a small reflectivity variation with height: 25 dBZ at the

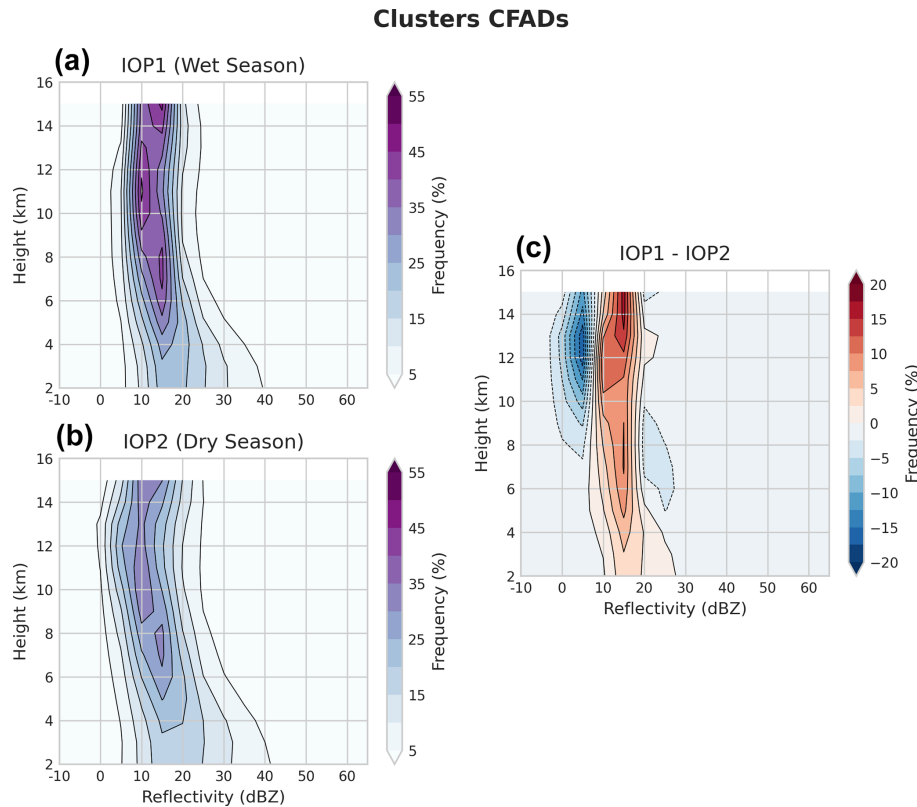


Figure 14. Mean contoured frequency by altitude diagram (CFAD) of clusters' reflectivity with 5 dBZ bins separated by IOP1 (a) and IOP2 (b); the IOP1 – IOP2 anomaly (c) is also shown.

surface to 10 dBZ at 15 km height (i.e., a -1 dBZ km^{-1} variation). Less frequent (between 5 % and 10 %) profiles have very low (5 to 10 dBZ) or high (40 dBZ) reflectivity at the surface and up to 25 dBZ at 15 km height, with a minimum of 0 dBZ at 12 km height.

Figure 14 shows the clusters' frequency by altitude diagrams by IOP as well the difference between them. Considering the largest frequencies, the profiles are similar between the IOPs (and with the complete time series profile), but the IOP1 profile (wet season) is more intense (up to 20 %) than IOP2 profile (dry season), especially between 10 and 20 dBZ.

4 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/ (Biscaro, 2016).

5 Code availability

The TATHU software package is available at <https://doi.org/10.5281/zenodo.17298147> (Uba, 2025a). The code developed to create the datasets with TATHU is

available at <https://doi.org/10.5281/zenodo.17297271> (Uba, 2025b).

6 Conclusions

A database of convective systems that occurred during the GoAmazon experiment was created to provide a comprehensive convection data collection for future GoAmazon studies. The *systems* and *systems_filtered* datasets cover the main convective characteristics (including morphology and intensity) and electrical activity. The filtered dataset is shown to be an acceptable sample of the complete dataset, selecting deep convective systems with a full life cycle within the research area. Convection seasonality is also well represented, with more intense convective systems between the dry season and the dry-season–wet-season transition and less intense systems in the wet season, typically occurring during the late-morning/early-afternoon period. The preeminent propagation direction of these systems is connected with the 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. As the SIPAM radar's main role is to serve as an operational weather radar instrument, 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), and radar software setting changes during the experiment. Another limitation is in the tracking itself, especially when dealing with system splitting/merging (which occurs in a significant portion of the convective system database), which can be more complex in large, mesoscale systems. Even with these limitations, the database is an important source of convection characteristics for cloud–aerosol–precipitation research.

Author contributions. CL and RA designed the study. CL processed the data and wrote the manuscript, with advice from RA. DU developed the tracking software, helped with data processing, and reviewed the manuscript. TB processed the radar data and reviewed the manuscript. IS provided the radar data and reviewed the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. 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.

Financial support. This research has been supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (grant no. 88887.464412/2019-00) (C. Lopes), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – grant nos. 438638/2018-2, 313355/2021-5, and 440171/2022-9) (R. Albrecht), and Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp – grant nos. 2009/15235-8, 2017/17047-0, 2022/13257-9, and 2023/04358-9) (field experiment operations).

Review statement. This paper was edited by Montserrat Costa Surós and reviewed by two anonymous referees.

References

Albrecht, R. I., Morales, C. A., and Dias, M. A. F. S.: Electrification of precipitating systems over the Amazon: Physical pro-

- cesses of thunderstorm development, *J. Geophys. Res.-Atmos.*, 116, D08209, <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?, *B. Am. Meteorol. Soc.*, 97, BAMS-D-14-00193.1, <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, e0000054, <https://doi.org/10.1371/journal.pclm.0000054>, 2022.
- Biscaro, T. S.: SIPAM radar data during GoAmazon [data set], https://ftp.cptec.inpe.br/chuva/goamazon/experimental/level_0/eq_radar/esp_band_s/st_sipam/ (last access: 17 March 2025), 2016.
- 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, *Atmos. Chem. Phys.*, 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, vol. 16, [vaisala.com](https://www.vaisala.com/sites/default/files/documents/6.Demetriades,Murphy,Cramer.pdf), Amer. Meteor. Soc. D, <https://www.vaisala.com/sites/default/files/documents/6.Demetriades,Murphy,Cramer.pdf> (last access: 17 July 2025), 2010.
- Dixon, M. and Wiener, G.: TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting – A Radar-based Methodology, *J. Atmos. Ocean. Tech.*, 10, 785–797, [https://doi.org/10.1175/1520-0426\(1993\)010<0785:TTITAA>2.0.CO;2](https://doi.org/10.1175/1520-0426(1993)010<0785:TTITAA>2.0.CO;2), 1993.
- 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, *Theor. Appl. Climatol.*, 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, *J. Geophys. Res.-Atmos.*, 121, 12891–12913, <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, *Atmos. Chem. Phys.*, 17, 14519–14541, <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, *Atmos. Chem. Phys.*, 20, 7489–7507, <https://doi.org/10.5194/acp-20-7489-2020>, 2020.
- Giangrande, S. E., Biscaro, T. S., and Peters, J. M.: Seasonal controls on isolated convective storm drafts, precipitation intensity, and life cycle as observed during GoAmazon2014/5, *At-*

- mos. Chem. Phys., 23, 5297–5316, <https://doi.org/10.5194/acp-23-5297-2023>, 2023.
- Greene, D. R. and Clark, R. A.: Vertically Integrated Liquid Water – A New Analysis Tool, *Mon. Weather Rev.*, 100, 548–552, [https://doi.org/10.1175/1520-0493\(1972\)100<0548:VILWNA>2.3.CO;2](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, *Atmos. Chem. Phys.*, 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, *J. Geophys. Res.*, 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, *J. Geophys. Res.*, 95, 16721–16736, <https://doi.org/10.1029/jd095id10p16721>, 1990.
- Heikenfeld, M., Marinescu, P. J., Christensen, M., Watson-Parris, D., Senf, F., van den Heever, S. C., and Stier, P.: tobac 1.2: towards a flexible framework for tracking and analysis of clouds in diverse datasets, *Geosci. Model Dev.*, 12, 4551–4570, <https://doi.org/10.5194/gmd-12-4551-2019>, 2019.
- Johnson, J. T., MacKeen, P. L., Witt, A., Mitchell, E. D. W., Stumpf, G. J., Eilts, M. D., and Thomas, K. W.: The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm, *Weather Forecast.*, 13, 263–276, [https://doi.org/10.1175/1520-0434\(1998\)013<0263:tsciat>2.0.co;2](https://doi.org/10.1175/1520-0434(1998)013<0263:tsciat>2.0.co;2), 1998.
- Jones, C. and Carvalho, L.: Active and break phases in the South American monsoon system, *J. Climate*, 15, 905–914, [https://doi.org/10.1175/1520-0442\(2002\)015<0905:AABPIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0905:AABPIT>2.0.CO;2), 2002.
- Laurent, H. and Laurent, D.: Characteristics of the Amazonian mesoscale convective systems observed from satellite and radar during the WETAMC/LBA experiment, *J. Geophys. Res.*, <https://doi.org/10.1029/2001JD000337>, 2002.
- Leal, H. B., Calheiros, A. J. P., Barbosa, H. M. J., Almeida, A. P., Sanchez, A., Vila, D. A., Garcia, S. R., and Macau, E. E. N.: Impact of Multi-Thresholds and Vector Correction for Tracking Precipitating Systems over the Amazon Basin, *Remote Sens.*, 14, 5408, <https://doi.org/10.3390/rs14215408>, 2022.
- Lopes, C.: GoAmazon convective systems datasets (systems and systems_filtered), Zenodo [data set], <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, *Mon. Weather Rev.*, 132, 714–725, [https://doi.org/10.1175/1520-0493\(2004\)132<0714:TCSAEO>2.0.CO;2](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., Chaboureaud, 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?, *B. Am. Meteorol. Soc.*, 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., Chaboureaud, 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?, *B. Am. Meteorol. Soc.*, 95, 1365–1380, <https://doi.org/10.1175/BAMS-D-13-00084.1>, 2014b.
- Machado, L. A. T., Calheiros, A. J. P., Biscaro, T., Giangrande, S., Silva Dias, M. A. F., 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., Pöschl, U., Pöhlker, C., Pöhlker, M. L., Ribaud, J.-F., Rosenfeld, D., Saraiva, J. M. B., Schumacher, C., Thalman, R., Walter, D., and Wendisch, M.: Overview: Precipitation characteristics and sensitivities to environmental conditions during GoAmazon2014/5 and ACRIDICON-CHUVA, *Atmos. Chem. Phys.*, 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 T. Geosci. Remote*, 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), *Atmos. Chem. Phys.*, 16, 4785–4797, <https://doi.org/10.5194/acp-16-4785-2016>, 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, *B. Am. Meteorol. Soc.*, 98, 981–997, <https://doi.org/10.1175/BAMS-D-15-00221.1>, 2017.
- Murphy, M. J. and Nag, A.: Cloud lightning performance and climatology of the U.S. based on the upgraded U.S. National Lightning Detection Network, https://ams.confex.com/ams/95Annual/webprogram/Manuscript/Paper262391/AMS2015_MALD_8.2_murphy.pdf (last access: 17 July 2025), 2015.

- Naccarato, K. P., Pinto, Jr, O., Garcia, S. A. M., Murphy, M., Demetriades, N., and Cramer, J.: Validation of the new GLD360 dataset in Brazil: First results, in: International Lightning Detection Conference, <https://www.vaisala.com/sites/default/files/documents/7.Naccarato,%20Pinto,%20Garcia.pdf> (last access: 17 July 2025), 2010.
- 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, American Geophysical Union, Washington, D.C., 149–162, ISBN 9780875904764, <https://doi.org/10.1029/2009GM000903>, 2009.
- Petersen, W. A. and Rutledge, S. A.: Regional Variability in Tropical Convection: Observations from TRMM, *J. Climate*, 14, 3566–3586, [https://doi.org/10.1175/1520-0442\(2001\)014<3566:RVITCO>2.0.CO;2](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, *J. Geophys. Res.-Atmos.*, 107, LBA 7-1–LBA 7-13, <https://doi.org/10.1029/2000JD000263>, 2002.
- Saraiva, I., Silva Dias, M. A. F., Morales, C. A. R., and Saraiva, J. M. B.: Regional Variability of Rain Clouds in the Amazon Basin as Seen by a Network of Weather Radars, *J. Appl. Meteorol. Clim.*, 55, 2657–2675, <https://doi.org/10.1175/JAMC-D-15-0183.1>, 2016.
- 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., ORNL – Oak Ridge National Lab., Oak Ridge, TN, USA, <https://www.osti.gov/biblio/1459573> (last access: 19 March 2025), 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, World Scientific, 25–33, 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, *J. Geophys. Res.*, 107, S1, <https://doi.org/10.1029/2001jd000335>, 2002.
- Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., Heikenfeld, M., Brunner, K. N., Bruning, E. C., Collis, S. M., Jackson, R. C., Leung, G. R., Pfeifer, N., Raut, B. A., Saleeby, S. M., Stier, P., and van den Heever, S. C.: *tobacv1.5*: introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena, *Geosci. Model Dev.*, 17, 5309–5330, <https://doi.org/10.5194/gmd-17-5309-2024>, 2024.
- 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 Sens.*, 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, *J. Geophys. Res.*, 126, <https://doi.org/10.1029/2020jd033766>, 2021.
- Uba, D.: uba/tathu: v1.0.0, Zenodo [code], <https://doi.org/10.5281/zenodo.17298147>, 2025a.
- Uba, D.: Ccloses/Tathu: SIPAM Tracking Module on TATHU, Zenodo [code], <https://doi.org/10.5281/zenodo.17297271>, 2025b.
- Uba, D. M., Negri, R. G., Enoré, D. P., Costa, I. C. D., and Jorge, A. A. S.: TATHU – Software para rastreo e análise do ciclo de vida de sistemas convectivos, <http://urlib.net/ibi/8JMKD3MGP3W34T/47AF772> (last access: 19 March 2025), 2022.
- Viscardi, L. A. M., Torri, G., Adams, D. K., and Barbosa, H. D. M. J.: Environmental controls on isolated convection during the Amazonian wet season, *Atmos. Chem. Phys.*, 24, 8529–8548, <https://doi.org/10.5194/acp-24-8529-2024>, 2024.
- 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, *Mo. Weather Rev.*, 123, 1941–1963, [https://doi.org/10.1175/1520-0493\(1995\)123<1941:tdkame>2.0.co;2](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?, *J. Climate*, 11, 1020–1040, [https://doi.org/10.1175/1520-0442\(1998\)011<1020:damceo>2.0.co;2](https://doi.org/10.1175/1520-0442(1998)011<1020:damceo>2.0.co;2), 1998.