The Illumination of Thunderclouds by Lightning: Part 1: The Extent and Altitude of Optical Lightning Sources

Michael Jay Peterson[1], Tracy Ellen Lavezzi Light[2], and Douglas Michael Mach[3]

[1]ISR-2,Los Alamos National Laboratory
[2]Los Alamos National Laboratory (DOE)
[3]Universities Space Research Association

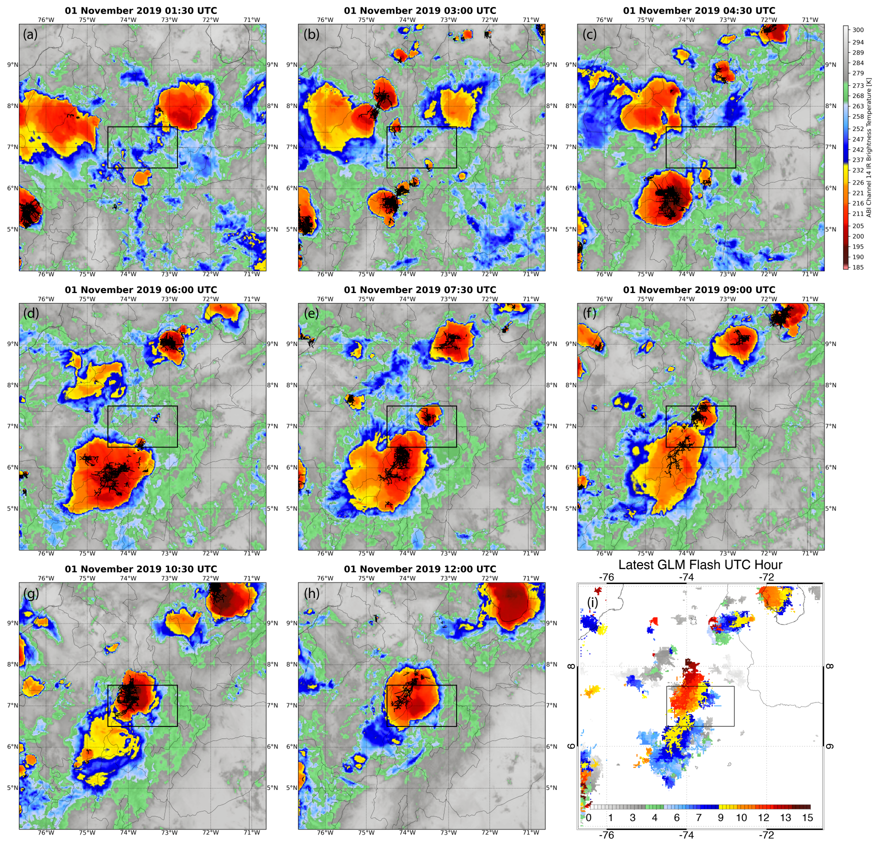
November 26, 2022

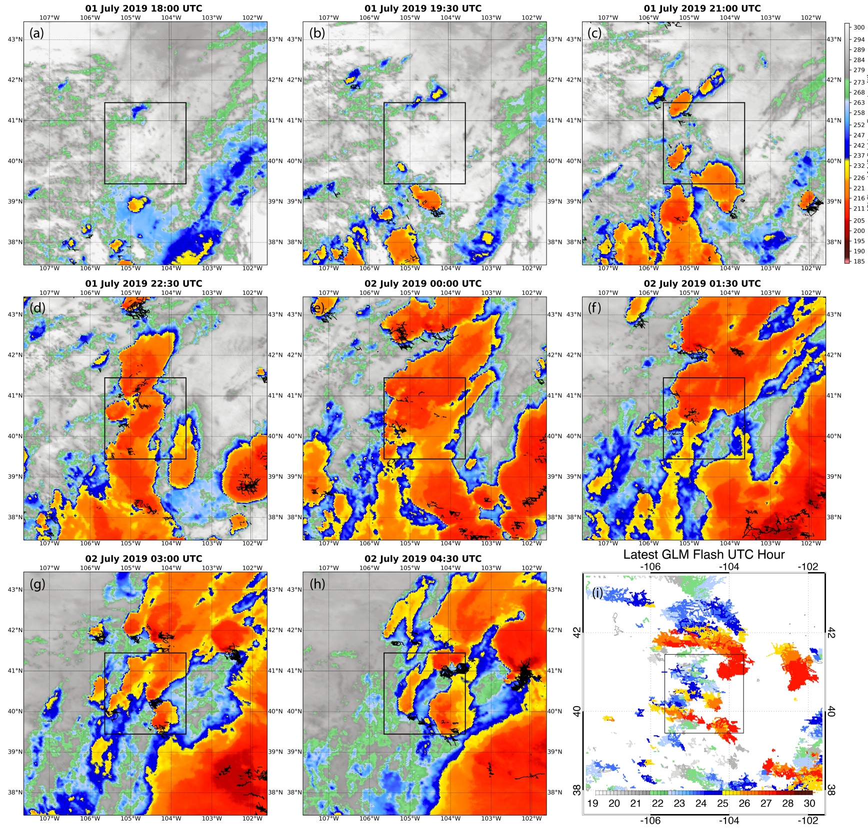
Abstract

Optical space-based lightning sensors including NOAA's Geostationary Lightning Mapper (GLM) detect lightning though its transient illumination of the surrounding clouds. What space-based optical lightning sensors measure is influenced by the physical attributes of the light source, the location of the source within the cloud scene, and the spatial variations in cloud composition. We focus on the lightning channels that serve as optical sources for GLM groups and flashes in this first part of our thundercloud illumination study. We match Lightning Mapping Array (LMA) sources with GLM groups and flashes during two thunderstorms to examine channel segments that are active during optical emission. We find that in each storm, the LMA sources matched with LMA groups are small (median: 2-3 km) compared to GLM pixels (nominal: 8 km), and preferentially come from high altitudes in the cloud (>8-10 km). The detection advantage for high-altitude sources permits GLM to resolve faint optical pulses near the cloud top that might be missed from lower altitudes. However, the most energetic groups can be detected from all altitudes, and the largest groups largely originate at low altitudes. The relationship between group brightness and illuminated area depends on flash development within the cloud medium, and flash development into different cloud regions can be identified by tracking GLM metrics of cloud illumination over time.

Hosted file

peterson-028-2021_fulminogram3d_part1_si.doc available at https://authorea.com/users/530648/articles/606426-the-illumination-of-thunderclouds-by-lightning-part-1-the-extent-and-altitude-of-optical-lightning-sources


01 November 2019 01:30 UTC
(a)
01 November 2019 03:00 UTC
(b)
01 November 2019 04:30 UTC
(c)
01 November 2019 06:00 UTC
(d)
01 November 2019 07:30 UTC
(e)
01 November 2019 09:00 UTC
(f)
01 November 2019 10:30 UTC
(g)
01 November 2019 12:00 UTC
(h)
Latest GLM Flash UTC Hour
(i)



01 July 2019 18:00 UTC
01 July 2019 19:30 UTC
01 July 2019 21:00 UTC
01 July 2019 22:30 UTC
02 July 2019 00:00 UTC
02 July 2019 01:30 UTC
02 July 2019 03:00 UTC
02 July 2019 04:30 UTC
Latest GLM Flash UTC Hour
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
ABI Channel 14 IR Brightness Temperature [K]

The Illumination of Thunderclouds by Lightning:

Part 1: The Extent and Altitude of Optical Lightning Sources

Michael Peterson[1], Tracy E. L. Light[1], Douglas Mach[2]

[1] ISR-2, Los Alamos National Laboratory, Los Alamos, New Mexico

[2]Science and Technology Institute, Universities Space Research Association, Huntsville, AL, USA

Corresponding author: Michael Peterson (mpeterson@lanl.gov), B241, P.O. Box 1663 Los Alamos, NM, 87545

Key Points:

- GLM measurements of thundercloud illumination are compared with LMA measurements of flash structure and ENGLN stroke detections
- The GLM detection advantage for high-altitude sources is quantified, and shown to vary with group area and energy
- Group maximum separation is a better approximation of LMA flash extent than event-based size metrics, but it is limited by GLM sensitivity

LA-UR-21-22307

Abstract

Optical space-based lightning sensors including NOAA's Geostationary Lightning Mapper (GLM) detect lightning though its transient illumination of the surrounding clouds. What space-based optical lightning sensors measure is influenced by the physical attributes of the light source, the location of the source within the cloud scene, and the spatial variations in cloud composition. We focus on the lightning channels that serve as optical sources for GLM groups and flashes in this first part of our thundercloud illumination study. We match Lightning Mapping Array (LMA) sources with GLM groups and flashes during two thunderstorms to examine channel segments that are active during optical emission. We find that in each storm, the LMA sources matched with LMA groups are small (median: 2-3 km) compared to GLM pixels (nominal: 8 km), and preferentially come from high altitudes in the cloud (>8-10 km). The detection advantage for high-altitude sources permits GLM to resolve faint optical pulses near the cloud top that might be missed from lower altitudes. However, the most energetic groups can be detected from all altitudes, and the largest groups largely originate at low altitudes. The relationship between group brightness and illuminated area depends on flash development within the cloud medium, and flash development into different cloud regions can be identified by tracking GLM metrics of cloud illumination over time.

Plain Language Summary

Lightning flashes are detected from space by monitoring cloud-top brightness for rapid changes due to illumination from lightning. The amount of lightning that instruments like the Geostationary Lightning Mapper (GLM) can detect depends on how the clouds are illuminated by lightning. Small, dim flashes are difficult to detect because they only faintly illuminate the surrounding clouds. However, even bright sources below particularly-thick clouds might not cause enough illumination to trigger the sensor. This study begins a comprehensive analysis of the thundercloud illumination that is measured by GLM, impacts on instrument performance, and the opportunities it presents for charactering flashes and their environments in new and unique ways.

1 Introduction

Lightning flashes are comprised of vast networks of hot ionized plasma channels (da Silva et al., 2019) that extend over tens or even hundreds of kilometers (Lang et al., 2017; Peterson et al., 2017a; Lyons et al., 2020; Peterson, 2019). Electrical currents traversing the various branches of the lightning ”tree” cause intense heating along the channels, leading to the atmospheric constituent gasses undergoing dissociation, excitation, and recombination (as summarized in Christian et al., 2000). This process results in particularly-strong emissions at the atomic lines for the atmospheric gasses, which space-based optical lightning detectors including the Lightning Imaging Sensor (LIS: Christian et al., 2000; Blakeslee et al., 2020) and Geostationary Lightning mapper (GLM: Goodman et al., 2013; Rudlosky et al,. 2019) leverage to measure total lightning (Cloud-to-Ground lightning and Intracloud lightning) during all hours of the day and night.

Instruments based on the LIS / GLM design measure the brightness of the scene below the spacecraft within a narrow spectral band surrounding the Oxygen emission line triplet at 777.4 nm at a high frame rate (nominally 500 Frames per Second). The instruments detect rapid increases in brightness at any point across their Charge-Coupled Device (CCD) imaging arrays caused by lightning illuminating the surrounding clouds. This approach yields high overall detection efficiencies at the flash level (69-88% for LIS: Cecil et al., 2014 derived from Boccippio et al., 2002; up to 90% for GLM: Bateman et al., 2020) relative to detailed ground-based measurements, while the pixelated imaging array enables coarse (kilometer-scale resolution) two-dimensional mapping of the development of the lightning tree (Peterson et al., 2018) for flashes within the instrument Field of View (FOV).

For sensors in geostationary orbit such as GLM, lightning can be mapped over most of the near-facing hemisphere (Peterson, 2019). This is important for documenting flash development in remote regions where other lightning measurements are sparse (i.e., over the open ocean, or deep within the Amazon rainforest), and for extending regional observations beyond their traditional ranges (i.e., mapping distant portions of flashes observed by Lightning Mapping Arrays: LMAs, Rison et al,. 1999). As flash structure is intimately linked to the organization and kinematics of the parent thunderstorm (Bruning and MacGorman, 2013), observing how flashes evolve can provide key insights into convective processes and their associated hazards (for example, Fierro et al., 2016; Peterson et al., 2020a,b; Thiel et al., 2020).

Space-based lightning imagers detect the illumination of the thunderclouds rather than the lightning channels, directly. This indirect measurement of the illuminated lightning channels raises some serious issues for detection. What lightning can be detected and at what level of detail are both determined by the optical characteristics of the clouds and how they are illuminated by the lightning in question. Optical lightning emissions interact with the surrounding cloud medium through absorption and scattering (Thomson and Krider, 1984), which disperse and attenuate the optical signals. For the simplest case of an optical point source embedded in a homogeneous slab cloud, the total optical energy from the event will be spread radially (Light et al., 2001a; Peterson, 2020a) and the optical waveform will be broadened temporally (Koshak et al., 1994; Suszcynsky et al., 2000; Light et al., 2001b) according to the paths taken by the scattered photons. The far edges of the spatial and temporal energy distributions will also be eroded by increased absorption at longer path lengths from additional particle interactions.

Under idealized circumstances, the effect of radiative transfer within the cloud on instrument DE is straight-forward. Increasing the optical thickness of the cloud amplifies these effects until the cloud reaches a point where the optical signals that escape the cloud-top fall below the instrument threshold for detection. Thus, the DE is reduced. The lightning and cloud scenes found in nature, however, are often far more complicated:

(1) The ionized lightning channels that generate the optical lightning emissions have variable geometries. The horizontal extent of the illuminated lightning channels and their vertical altitudes are not consistent between flashes – or even at different times within the same flash.

(2) Spatial variations in cloud composition cause the optical emissions to preferentially transmit through certain cloud regions compared to others. An extreme case of this is when "holes" occur in LIS or GLM groups where the clouds surrounding a particularly opaque cloud region are simultaneously illuminated while the central region remains dark (Peterson, 2020b). This occurs in both tall convection and with overhanging anvil clouds that are presumably illuminated from below.

(3) If the optical emissions encounter a cloud boundary, they can access neighboring clouds and take a "shortcut" path to the satellite compared to transmitting through the full optical depth of cloud above the source (Peterson, 2020a). This causes particularly-radiant pulses to simultaneously illuminate exceptional cloud areas (up to 10,000 km^2) that extend far beyond the electrically-active thunderstorm core (Suszcynsky et al., 2001; Peterson et al., 2017a). All cases of "warm lightning" that have been found in the LIS dataset so far (Peterson et al., 2017a) are from a combination of (2) and (3). In these cases, LIS only detects the illumination of nearby warm clouds and not illumination within the optically thick storm core.

All these factors affect not just the DE of instruments like LIS and GLM, but also degrade their Location Accuracy (LA) and introduce substantial biases into the gridded products generated from the flash cluster data that describe flash and group characteristics across the storm (Bruning et al., 2019).

Considering how thunderclouds are illuminated by lightning is necessary to ensure proper interpretations of space-based observations from lightning imagers (for example, recognizing when their limitations are hampering detection). These indirect lightning measurements can also be leveraged for novel applications that provide additional information about lightning and the surrounding storm clouds that are not possible with direct lightning measurements - including those from Radio-Frequency (RF) sensors. In this study, we will analyze the cloud illumination measured by GLM, examine the factors that determine what GLM can detect, and explore how this information can be used in a new application: estimating the altitudes of optical sources within the cloud. This study is organized into four parts, each with a specific focus, that will all use the same set of combined lightning observations from the GOES-16 GLM, an LMA, and the Earth Network Global Lightning Network (ENGLN). These data were collected from two different thunderstorms. The first was a Colombia thunderstorm near the GOES-16 satellite subpoint that had a normal charge structure and low GLM detection threshold. The second was an inverted-polarity thunderstorm over Colorado where most of the lightning activity occurred at low altitudes and signal loss from radiative transfer effects was further amplified by a high GLM detection threshold. GOES-16 Advanced Baseline Imager (ABI: Schmit et al., 2017) observations from these thunderstorm cases will also be considered.

The focus here in Part 1 is on the altitudes and geometries of the lightning channels that serve as optical sources for cloud illumination. We will use combined optical and RF lightning measurements to infer the sizes and altitudes of the optical sources responsible for GLM groups, and examine how GLM measurements respond to changes in source position and structure. Future work in Part 2 (Peterson et al., 2021b) will focus on the GLM instrument and examine how the GLM data products change under different detection thresholds. Part 3 (Peterson et al., 2021c) will then use GLM measurements of cloud illumination to develop a methodology for retrieving source altitude. Finally, Part 4 (Peterson et al., 2021d) will construct and evaluate volumetric meteorological and thundercloud imagery from the GLM data.

2 Data and Methodology

2.1 The GOES-16 Geostationary Lightning Mapper

GLM is the first space-based lightning sensor operated on NOAA spacecraft, and the first lightning sensor to be placed in geostationary orbit. We use GLM data from the GOES-16 satellite, which was launched in November 2016 and has been providing lightning data to the public from the GOES-East position since December 2017. The GOES-16 GLM FOV extends from the Pacific Ocean in the west to the coast of west Africa in the east, and between 54 degrees north and south latitude (Rudlosky et al., 2019). This includes the full width of the Americas landmass between Argentina and southern Canada.

Cloud illumination is measured using pixel-level GLM event data that is captured during a 2-ms GLM integration frame. Event detection is not consistent across the GLM FOV, however, due mostly to the curvature of the Earth. While GLM pixels around the satellite subpoint (75.2° W, 0° N) are measured from nadir, the pixels at the edge of the GLM FOV approach a side view of the thunderstorm. This causes a few issues for GLM performance. The first is that the area of the Earth's surface (or, more accurately, the surface of the ellipsoid chosen to correspond to cloud-top altitude) contained within each pixel increases with slant angle. GLM partially mitigates this effect by employing a variable-pitch focal plane that preserves a ~8 km pixel resolution over most of the CCD array (only increasing up to ~14 km at the edge of the FOV), but there are still local variations in pixel size that impact how source energy density translates to total pixel energy. These variations are minimized by examining thunderstorms near the satellite subpoint.

The second issue is that the instrument threshold varies across the instrument FOV. Thresholds are generally lowest near the satellite subpoint and increase radially from there – but, as with pixel size, there are also local variations imposed by the instrument hardware. These variations are caused by the Real Time Event Processors (RTEPs) rather than the focal plane, and thus are aligned with the sub-arrays handled by each RTEP. Selecting cases near the satellite subpoint also provides the best thresholds to examine faint cloud illumination.

The event data recorded by GLM is then processed by the Lightning Cluster Filter Algorithm (LCFA: Goodman et al., 2010) in the GLM ground system, which introduces additional issues that make it into the operational GLM data product that is distributed by NOAA. The primary role of the LCFA is to cluster contiguous simultaneous events on the GLM imaging array into "group" features that approximate optical pulses and then cluster groups into "flash" features that nominally describe complete and distinct lightning flashes. Filtering is also applied based on the event and clustered data to remove obvious artifacts. The LCFA is subject to strict latency requirements, however, that limit how much lightning can actually be clustered. To prevent latency issues, the LCFA introduces hard thresholds on how many events may comprise a group, how many groups may comprise a flash, and the maximum duration of a flash.

Once a group exceeds 101 events or a flash exceeds 101 groups or 3 s, it is terminated by the LCFA, and any subsequent activity will be clustered into a new and independent group or flash feature. Of course, lightning has no hard limits and the thresholds chosen by the LCFA are quite low – even compared the for previous LIS instrument (Peterson et al., 2017b). Therefore, a non-negligible fraction of lightning becomes split into multiple pieces by the LCFA – including the largest and most exceptional flashes on Earth (Peterson et al., 2020c). In Peterson (2019), we document an approach to correct these LCFA issues and produce science-quality GLM data. We will use that dataset here, which is available at Peterson (2021a).

In this study, we will compare cloud illumination in an ideal thunderstorm case with a problematic thunderstorm case. The selection criterion for an ideal case is simply proximity to the satellite subpoint. However, a problematic case should have as many unfavorable factors for GLM detection as possible including: (1) a high GLM threshold, (2) most of the lightning activity occurring near the cloud base, and (3) occurring in a region of the CCD array where there are substantial local variations in threshold and pixel size. Additional limitations on both cases are that they should occur close to the center of an LMA network where accurate VHF source information is available, and they should occur after the late 2018 GLM software updates (Koshak et al., 2018) that improved timing and geolocation accuracy. Two such cases are found: one within the domain of the Colombia LMA, and another within the domain of the Colorado LMA.

2.2 The Colombia and Colorado Lightning Mapping Arrays

2.2.1 The Colombia LMA

The closest LMA to the GLM satellite subpoint is the Colombia LMA (COLLMA: Lopez et al., 2016; Aranguren et al., 2018). Note that the Colombia LMA has been abbreviated as COLLMA as well as COLMA in the literature, but the later acronym conflicts with the Colorado LMA that is universally abbreviated COLMA – so we will exclusively use the COLLMA term to describe the Colombia LMA here. COLLMA was deployed to Colombia as ground support for the Atmospheric Space Interaction Monitor (ASIM: Neubert et al., 2019) on the International Space Station (ISS) and became the first LMA system to be installed in the inner tropics. The equatorial location of the system has allowed charge structures in the particularly-tall convective clouds that occur in Colombia to be resolved (Lopez et al., 2019), which are thought to be favorable for Terrestrial Gamma-ray Flashes (TGFs: Split et al., 2010; Fabró et al., 2015) and Gigantic Jets (GJs: Chen et al., 2008; Boggs et al., 2019).

The COLLMA was initially deployed in northern Colombia in 2015 surrounding the city of Santa Marta on the Caribbean coast as a 6-sensor network configured to have a 5-20 km baseline. Lightning data were collected in Santa Marta until 2018, when the network was redeployed Barrancabermeja in central Colombia, which sees greater overall lightning activity (Albrecht et al. 2016; Peterson et al., 2021a).

We will use COLLMA data collected during this second deployment because it occurred after the late 2018 GLM software updates (Koshak et al., 2018). LMA data over a 1.7º longitude (74.5º W – 72.8º W) by 1º degree latitude (6.5 º N – 7.5º N) box within the LMA domain from 01 November 2019 were provided by Lopez (2020, personal communication) for comparison with GLM. The LMA sources were clustered and quality controlled by Lopez (2020, personal communication) using the algorithms developed by van der Velde and Montanyà (2013). Noise sources are identified and removed according to source density in three-dimensional (3D) space-time boxes whose sides describe a horizontal distance (XY), a vertical distance (Z), and a time difference (T). The XY, Z, and T thresholds are derived empirically to represent "low," "medium," or "high" levels of noise suppression. The data provided were subject to the medium setting where two-or-fewer sources in boxes with sides of XY=5 km, Z=1.5 km, and T=0.5 s are eliminated.

2.2.2 The Colorado LMA

The Colorado LMA (COLMA) is a nominal 15-station LMA network that has been operational in northern Colorado since 2012 (Rison et al., 2012). COLMA is a large LMA with stations spread across a 100 km distance and a nominal range of around 350 km. Each station is designed to be autonomous with power provided by solar panels and communications provided by cellular modems. Previously-analyzed COLMA data from multiple 2019 thunderstorms were provided by Cummins (2020, personal communication). Flashes were clustered using the XLMA software and quality control was performed subjectively using on an empirically-derived reduced chi-squared threshold.

We only consider lightning sources near the center of the COLMA network in this study. Sources are selected from a latitude / longitude box that is 2º longitude (105.6º W – 103.6º W) by 2º degree latitude (39.4 º N –41.4º N). This larger box than the COLLMA data provided from the Colombia thunderstorm case accommodates the larger thunderstorm features in the Colorado case while still capturing only the lightning activity that occurred near the center of the array.

2.3 The Earth Networks Global Lightning Network

The Earth Networks Global Lightning Network (ENGLN) combines lightning observations from Earth Networks Total Lightning Network (ENTLN: Zhu et al., 2017) sensors with the World-Wide Lightning Location Network (WWLLN: Lay et al., 2004; Rodger et al., 2006; Jacobson et al., 2006; Hutchins et al., 2012) to detect and geolocate Cloud-to-Ground (CG) strokes and intracloud discharges. ENGLN data from across the GLM field of view was provided by Earth Networks for the entirety of 2019. We only consider the CG data within the ENGLN dataset in this study since we have LMA observations available that map the in-cloud portions of each lightning flash.

2.4 Matching LMA Sources and ENGLN Strokes with GLM Groups and Flashes

The matching scheme in this study is based on the GLM/ENGLN approach used in Peterson and Lay (2020). We assume that all RF events that occur within the footprint of a GLM group are part of the active lightning channels that contributed to the optical energy recorded during the group. ENGLN strokes and LMA sources are interpreted as an RF analog to the optical GLM events, and we ingest them into the GLM clustering hierarchy accordingly. A GLM group might be assigned multiple RF events within its footprint, but RF events cannot have multiple parent groups.

RF events are not perfect analogs to optical GLM events, and this leads to two important caveats with our matching scheme. The first is that GLM is not able to detect every active portion of the flash, and in some cases, this will cause RF events to occur beyond the GLM group footprint. In Peterson and Lay (2020), we accounted for this possibility by allowing ENGLN events to match GLM groups if they occurred within a specified distance threshold from the GLM group footprint. A few thresholds were tried, and 10 km was ultimately selected. We will use the same 10-km threshold in this study.

The second caveat is that the RF events might not occur at the same time as the optical illumination. This is expected to be a greater issue with LMA events compared to ENGLN strokes because VHF emissions largely cease once the active channel becomes conductive, while optical emissions are sustained as long as current continues to flow in the channel. Thus, the LMA data coincident with a GLM group might not describe the full extent of the illuminated channel that generated the group. To address this possibility, we also impose a generous time threshold on the GLM/RF matches. RF events are assigned to the overall most-energetic GLM group that occurs within 10 ms of the RF event – not the group that is closest in time. This ensures that the RF events capture the peak of the light curve from whatever process (stroke, K-change, etc.) is causing the channel illumination.

Only the GLM groups and flashes that are entirely within the LMA data domain boundaries are considered for matching. Flashes that straddle the boundaries or occur outside of the LMA domain will be shown in Section 3.1 to describe the broader thunderstorm, but are otherwise not included in the results comparing the GLM and LMA aspects of the lightning detected during these storms.

3 Results

The following sections describe the joint GLM / LMA behavior of lightning during thunderstorms in Colombia and Colorado. The overall history of these storms will be summarized in Section 3.1. Section 3.2 compares the LMA and GLM extents of matched flashes. Section 3.3 analyzes the altitude distributions of LMA sources in matched cases. Finally, Section

3.4 examines how the illumination of the surrounding clouds changes as flashes propagate through the cloud medium.

3.1 Lightning Measurements from Thunderstorms in Colombia and Colorado

3.1.1 The Colombia Thunderstorm Case

The Colombia thunderstorm is in an advantageous location for GLM detection near the GOES-16 satellite subpoint, but it is also an ideal case to examine because all stages of the convective life cycle are sampled, resulting in a diverse collection of flash types within the combined LMA and GLM domain. Figure 1 shows the history of the storm. Figure 1a-h show the ABI Channel 14 (11.2 μm) infrared brightness temperatures of the clouds across the mapped region from 01:30 UTC to 12:00 UTC and the GLM-derived horizontal structure of each flash (black line segments). Figure 1i sorts the GLM data by time and then overlays the GLM measurements from all flashes produced by the Colombia thunderstorm to show the latest time when lightning activity was detected at each point on the map.

The thunderstorm moved over the LMA domain from south to north, and the boxed region captures the full longitudinal width of lightning activity from the storm as it passed though. The first lightning within the LMA box occurred between 02:00 UTC and 04:00 UTC when two small convective features crossed into the box (Figure 1b). Timeseries are shown in Figure 2 of GLM, LMA, and ENGLN lightning rates (Figure 2a), LMA altitude distributions (Figure 2b), and LMA (Figure 2c) and GLM (Figure 2d) extent distributions. The GLM flash rate during this period approached 2 flashes per minute with 1 ENGLN -CG every 2.5 minutes and 1 ENGLN +CG every 10 minutes during this peak. These flashes also generated a maximum of 30 GLM groups per minute and 70 LMA sources per minute. All quality-controlled LMA sources were between 5 and 15 km altitude, and both the LMA flashes and GLM flashes were small (mostly < 20 km in extent) during this period.

The most active period of lightning within the LMA domain extended from 05:30 UTC to 13:15 UTC. Peak GLM flash rates and ENGLN -CG rates exceeded 10 per minute, with an additional 2 +CGs per minute and hundreds of GLM groups and thousands of LMA sources per minute (Figure 2a). This period actually describes the passage of two distinct convective features in Figure 1. The first of these features started off as disorganized convection to the south of the LMA domain at 01:30 UTC (Figure 1a), which first started to produce lightning by 03:00 UTC (Figure 1b). It then organized into a large convective feature by 04:30 UTC (Figure 1c), and started to encroach upon the LMA domain by 06:00 UTC (Figure 1d). This feature then started to mature and eventually dissipate by the end of the period, resulting in the long-horizontal flashes that we first see at 08:00 UTC, but become prevalent within the LMA domain after 9:00 UTC (Figure 1f, Figure 2c-d).

The second thunderstorm feature initiated within the LMA domain starting in the 06:00 UTC hour (Figure 1d). This feature grew and developed while the first feature was maturing between 07:30 UTC (Figure 1e) and 10:30 UTC (Figure 1g). By the end of the period, this second feature was the primary source of lightning within the LMA domain (Figure 1h). Due to the staggering of the two thunderstorm features in time, GLM and the LMA were sensing both the compact flashes associated with new convection and the long horizontal flashes associated with maturation from 09:00 UTC onward. This time period provides a robust variety of flash extents, altitudes, and optical energies that allow us to examine what GLM can detect relative to the LMA.

3.1.2 The Colorado Thunderstorm Case

The Colorado thunderstorm is mapped in Figure 3, while the same timeseries of lightning rates and flash characteristics as Figure 2 are shown in Figure 4. Note that this storm occurred around UTC midnight and the listed hours are relative to 00:00 UTC on the first day of the storm (01 July 2019). The Colorado case started off as disorganized convection that grew between 21:00 UTC on 01 July and 00:00 UTC on 02 July, and then continued to produce lightning over the region into the night.

LMA data were only available between 20:00 UTC on 01 July and 03:00 UTC on 02 July (hour 27 in Figure 4a), but they showed that the flashes produced by this storm were close to the cloud base (Figure 4b) and fairly compact – only occasionally exceeding 40 km (Figure 4d). GLM flashes (Figure 4c) were typically smaller than the LMA flashes, and mostly < 20 km across. While GLM flash rates were higher in the Colorado thunderstorm (Figure 4a), there are indications of poor detection efficiency in the data. The group rates were within an order of magnitude of the LMA source rates during the Colombia case, but they are separated by a full two orders of magnitude in Figure 4a. For every GLM group, there were approximately 100 LMA sources detected. Also, the GLM group-level structure (plotted with black line segments) frequently occurs outside of the cold cloud region rather than within the convective storm core in the GLM / ABI maps in Figure 3 – particularly at 19:30 UTC (Figure 3b) and 21:00 UTC (Figure 3c) on 01 July, and 03:00 UTC (Figure 3g) and 04:30 UTC (Figure 3h) on 02 July. This can happen with optically-thick clouds that attenuate the optical signals to the point of preventing detection, entirely. In these cases, only the optical emissions that escape the side of the cloud and illuminate nearby lower cloud decks are detected from space. We've previously noted this behavior with LIS as the source of apparent cases of "warm lightning" (Peterson et al., 2017a).

3.1.3 Relative Detection Rates between GLM, the LMAs, and ENGLN

We can use our GLM/RF matching scheme to quantify the fraction of the lightning in the Colombia and Colorado thunderstorms detected by each instrument. Table 1 computes the amount and percentage of GLM flashes and groups that also contain ENGLN strokes or LMA sources. There were a total of 2154 GLM flashes and 56,399 GLM groups within the LMA box during the Colombia thunderstorm case. 21.9% of the flashes contained at least one ENGLN stroke, while 90.1% were matched with LMA sources. At the group level, ENGLN strokes accounted for just 1.1% of groups, while LMA sources were linked to 40.2% of groups. Note that these percentages are low estimates for the relative trigger rates because processes like strokes and K-changes might generate multiple groups with only one being counted here.

The Colorado thunderstorm case, meanwhile, produced 5278 flashes. Of these flashes, 14.5% contained ENGLN strokes, while almost all flashes (99.9%) were linked to LMA sources. This may be due to a greater LMA sensitivity, but a lower GLM detection efficiency could also play a role if the flashes that are resolved by GLM are also favorable to LMA matching. Relative event rates are also higher at the group level – with 2.6% of groups matching an ENGLN stroke and 70% of groups containing LMA sources.

Table 2 inverts Table 1 and lists the quantities and percentages of RF events that are successfully matched to GLM groups and flashes. The Colombia thunderstorm generated 1246 ENGLN -CGs. Of these, 1013 (81.3%) were matched to GLM flashes, while 49.8% were matched to GLM groups. As for +CGs, ENGLN reports 71 total with 49 (69%) matched to GLM flashes and 13 (18.3%) matched to GLM groups. The LMA reports a total of 376,482 sources, with 96.9% matching GLM flashes and 48.1% matching GLM groups. The remaining RF events were not close enough to a GLM flash or group to constitute a match, and these might be considered missed events.

It is important to note that the percentages listed in Table 2 do not correspond to GLM Detection Efficiency (DE) values, as there are additional nuances that need to be considered with DE to make a fair comparison. Still, we can use these match rates to comment on differences in GLM detection between the two cases. The Colorado case generated 3123 -CGs and 104 +CGs that were detected by ENGLN. 35.9% of these -CGs and 51% of the +CGs were matched to GLM flashes, while 23% of the -CGs and 39.4% of the +CGs were matched to GLM groups. The LMA resolved 5,658,247 VHF sources and only 22.7% matched GLM flashes and 2.8% matched GLM groups. Generally, GLM had more difficulty detecting the optical emissions associated with RF events during the Colorado case. An exception could be +CGs, which had matching GLM groups more often in the Colorado case than in the Colombia case, but this could be an artifact of the low sample size of +CG strokes.

3.2 LMA and GLM Measurements of Lightning Extent

Differences in GLM performance are expected to impact the flash characteristics reported by GLM. If GLM has difficulty measuring illumination in certain cloud regions, then extent, duration, optical energy, etc. may be reduced when flashes propagate into these clouds. Figure 5 compares the overall extent of the LMA sources matched to GLM flashes with the GLM flash extent measured using either group centroid (left) or event pixel (right) locations. The Colombia case is considered in Figure 5a,b while the Colorado case is examined in Figure 5c,d. To highlight the relative scale of the GLM and LMA flashes, the vertical axis shows the ratio of the GLM flash extent to the LMA flash extent, with unity corresponding to flashes of the same size. Note that these LMA-derived extents are not the same as the LMA flash extents from Figures 2d and 4d, as multiple LMA flashes might occur within the footprint of a GLM flash. In such cases, all of these LMA flashes contribute to the cloud illumination detected by GLM, and the GLM flash extent should capture the combined extent of all matched LMA flashes. To account for this, we record the LMA flash indices of each LMA source matched to the constituent groups in the GLM flash, and then compute LMA flash extent as the maximum Great Circle distance between all LMA sources with those flash indices. This results in LMA extents that are larger than the flash extents noted previously – including some cases that appear to reach 100 km. Moreover, GLM flashes that are comprised of a single group in the left panels of Figure 5, or whose events only illuminate one pixel in the right panels will have reported extents of 0 km. These flashes are shown along the bottom of the plots. Slanted lines are also drawn to indicate constant distances representing the GLM pixel size. Finally, the solid thick line tracks the average GLM : LMA extent ratio for each LMA extent.

In the past, we have used the separation of groups rather than the separation of events to document flash size with GLM-like instruments because groups are less sensitive to radiative transfer effects in the cloud than events and it is possible to resolve flash extents smaller than a GLM pixel from the radiance-weighted group centroid data. Figure 5a,b shows why this approach is more appropriate than measuring flash size using event data. Under the ideal conditions of the Colombia case, the average GLM flash extent (solid black line) is close to the LMA measured source extent (near the horizontal line at 1.0) for flashes larger than ~5 km. For smaller LMA flashes, the GLM group extent overestimates the LMA extent because sources located at pixel boundaries can effectively double the extent of the GLM flash (Zhang et al., 2020). GLM can still over-estimate the flash size in larger cases, but it is far more likely that the LMA will detect lightning activity that GLM does not resolve. By contrast, the GLM event extent (Figure 5b) overestimates the LMA extent for all but some of the largest flashes detected in the Colombia thunderstorm. Differences between GLM event separation and LMA source extent can be small for cases of propagating flashes that approach the megaflash scale (100+ km in total length), but for most convective-scale flashes, GLM group separation provides the more accurate measure of flash extent. The flash areas reported by GLM are also subject to these high biases because they are computed using event data rather than group data. This will impact gridded products including AFA and Minimum Flash Area that are, likewise, derived from event data.

GLM can produce reasonable measurements of flash extent for larger flashes under ideal viewing conditions, but thunderstorms that are subject to poor GLM performance will not resolve flashes to the same extent as an LMA. Group separations in the Colorado case are almost always smaller than their matched LMA source extents, with mean GLM : LMA ratios decreasing from 0.6 for 2-km LMA extents to 0.02 for 100-km LMA extents. The coarse GLM pixel size partly explains the decline in mean GLM : LMA ratios with distance – which can be seen as a local maximum above and following the 1-pixel slanted contour line. However, the primary reason for GLM underestimating flash extent (even up to the megaflash-scale at 100 km) is that GLM simply does not detect optical emissions from most of the lightning channels in the flash that are mapped by the LMA. Even when illumination does occur at levels that GLM can detect, the GLM flash extent is most likely 0 km (i.e., along the bottom of the plot in Figure 5c). This indicates one of two possibilities: that GLM only detected one group during these LMA flashes that span tens of kilometers, or that all subsequent groups are comprised of single events that all occur in the same GLM pixel. The first scenario has event separations >0 km in Figure 5d, while the latter case also has 0-km event separations. In either case, very little of the flash is being resolved by GLM.

While GLM sensitivity can severely impact which lightning channels in the flash can be mapped by GLM, we can also explore how GLM sensitivity impacts which optical sources give rise to GLM groups. Figure 6 repeats the GLM event / LMA source extent analyses from Figure 5b and d at the group level (Figure 6a,c) while additionally showing overall histograms for the maximum extent of LMA sources along the active lightning channels during the GLM group. As before, the top panels correspond to the Colombia case while the bottom panels correspond to the Colorado case. In both thunderstorms, GLM groups are most frequently comprised of 1 or 2 events, corresponding to separations of 0 km (bottom row of the figure) or 1 pixel (~8 km, first slanted line). Larger groups that are comprised of 5-10 events, meanwhile, occur over a range of LMA extents and are not strongly correlated with LMA source extent. This supports the idea that optical energy is a stronger factor for determining group size than source geometry (Suszcynsky et al., 2001) due to the broadening effects of scattering by the cloud medium on the optical signals emitted by the sources.

At the same time, the extent of LMA sources along the active lightning channels during GLM groups in Figure 6b and d are usually quite small. The median LMA source extent is between 2 and 3 km in both thunderstorms, while the LMA sources during 83% of the matched groups from the Colombia case and 90% of the groups from the Colorado case are smaller than one GLM pixel. Particularly-long optical sources that span multiple GLM pixels do occur in cases of long horizontal flashes, but they are rare with only ~1% of LMA extents exceeding 3 GLM pixels. These long optical sources are a special case representing just one type of illumination that we see in horizontally-propagating flashes, albeit one that can last for tens of milliseconds while producing many consecutive groups (an example will be shown later in

Section 3.4). Other modes, including the frequent "flickering" at the ends of developing branches as the lightning channels extend through the cloud are far more localized, typically with only 1-2 GLM events per group.

3.3 The Altitudes of LMA Sources during GLM Groups

Thunderstorm charge structure plays an important role in shaping what GLM detects from a given storm by determining the altitudes at which lightning activity occurs. GLM has a detection advantage for resolving lightning near the cloud-top, as the optical thickness of the layer between the source and satellite is small compared to the full cloud depth. Thus, the signals will be less attenuated by scattering and absorption within the cloud medium. This is expected to be an important factor behind the difference in GLM performance between the Colombia case and the Colorado case.

To examine the effect of source altitude on GLM detection, Figure 7 computes the overall altitude distributions for all LMA sources during the Colombia and Colorado thunderstorms (Figure 7a,d), and then compares the source altitude distributions from the LMA sources matched with GLM flashes (Figure 7b,e) and GLM groups (Figure 7c,f) by subtracting the normalized matched distributions from the overall distributions from Figure 7a,d. The Colombia thunderstorm was a normal polarity thunderstorm with an upper positive layer above ~10 km altitude and a lower negative layer around 5 km altitude. Most of the LMA sources resulted from development through the upper layer. The Colorado case, meanwhile, was an inverted-polarity thunderstorm with most of the LMA source occurring in the positive charge layer around 5 km altitude.

Both thunderstorms show that GLM is predisposed towards detecting high-altitude sources while missing low-altitude sources at the flash level (Figure 7b,e) and at the group level (Figure 7c,f). However, the amplitudes of these detection differences are only a few percent in either direction. While still a noticeable departure from the overall LMA source distribution, it is far from the notion that GLM detects *only* high-altitude sources. In fact, GLM can preferentially detect certain types of low-altitude sources. There is a slight positive bias in the lowest altitude bins for the Colorado case in Figure 7e,f. This positive bias is only present in the 01 July case (two other inverted polarity Colorado cases were examined, but not shown), and it only occurs at certain hours during the storm (most notably in the 00 UTC hour on 02 July). It appears to be related to low-altitude flashes near the edge of the convective core that GLM can easily detect. Both the intensity of the discharge and the cloud scene surrounding the optical emitter factor in to GLM detection. Light escaping the side of an opaque cloud can lead to apparent cases of "warm lightning" (Peterson et al., 2017a) where only the surrounding clouds are illuminated brightly enough to trigger a lightning imager, while even particularly opaque clouds can still be illuminated by sufficiently-bright optical pulses (Peterson, 2020b).

Figure 8 demonstrates how the source altitude profiles from Figure 7 vary with group energy and illuminated area in the Colombia case. Rather than tallying all matched LMA sources, Figure 8 shows separate two-dimensional histograms for the maximum (top tow), mean (middle row), and minimum (bottom row) LMA source altitudes associated with a given group. These 2D histograms are normalized such that the total frequency per unique group energy (left column) or area (right column) on the horizontal axis sums to 100%. The middle column shows the overall source altitude histogram for the matched LMA sources for reference.

For the small and dim groups detected by GLM, the corresponding LMA sources come primarily from the upper layer at 10 km. This is a reflection of the overall source altitude distribution seen in Figure 7a. However, as we move towards larger and more radiant groups, the lower charge layer becomes increasingly important. Eventually, a secondary maximum forms in the group energy-altitude distributions in Figure 8a,d, and g (around 100 fJ). The largest groups (~1000 km^2 in Figure 8c,f, and i) are more likely to be associated with sources in the lower layer than the upper layer. Some of these large and bright groups come from strokes (Koshak, 2010), but Table 1 and Table 2 suggest that stroke coincidence is too rare to explain all of them. Furthermore, when Figure 8 is generated using only GLM groups matched with ENGLN strokes (included as Supporting Information in Figure S1), stroke detections occurred across the full range of GLM group energies and areas shown here. Strokes generally produce larger and brighter groups than in-cloud pulses, but this is not always the case. Low peak currents or thick clouds between the optical source and the satellite can cause strokes to generate GLM groups that are not exceptional.

The trends in Figure 8 can be explained by three simultaneous factors: (1) small dim groups originating in the lower charge layer being attenuated to the point where GLM does not easily resolve them, (2) the additional optical depth available for scattering broadening the optical signals from the lower charge layer – leading to larger groups (as we saw in Figure 7f of Peterson, 2020a), and (3) large, energetic groups arising from lightning at the edge of the thunderstorm where the optical signals can transmit through / reflect off of thinner cloud layers to reach the satellite.

These factors become more important for the Colorado case where the sources are concentrated in the lower charge layer (Figure 7d). Figure 9 repeats the GLM energy / area and LMA source altitude analyses from Figure 8 for the Colorado case. In this case, GLM groups at all energies primarily originate from optical emissions near the cloud base. There is no secondary maximum in the vertical profiles as we saw with the Colombia case in Figure 8. Instead, group frequency tapers off with increasing altitude (despite a similar detection advantage from these higher sources in Figure 7f). Few groups originating from the 8 km (flashes) to 9 km (groups) changeover point in Figure 7e,f and higher GLM thresholds over Colorado make it difficult for GLM to detect flashes in this storm, let alone resolve their detailed development over time.

3.4 Variations in Cloud Illumination with Flash Propagation

How thunderclouds are illuminated by lightning depends on the optical characteristics of the cloud scene and the position and geometry of the source. We've shown that optical sources are typically much smaller than GLM pixels (Figure 6 indicates 2-3 km extents). However, flashes frequently develop beyond these scales and may even extend between clouds regions with different optical characteristics –notable examples being flashes that develop horizontally between convective and stratiform clouds and flashes that develop vertically between two charge layers. We might expect a flash to infrequently generate larger groups while it develops through the lower layer and then transition to frequent small / dim groups after it reaches the upper layer. Moreover, if optical pulses truly are localized processes in most cases, then flashes that remain in one of the two charger layers should illuminate the same clouds in the same way with each optical pulse. This should lead to cases of "repeater flashes" where the group illuminated area is a strong function of only group brightness.

To search for these repeater flashes, we examine the groups that comprise each flash during the Colombia thunderstorm and compare the maximum event energy per group with the group footprint area. For flashes that consist of at least 10 groups, we then fit to a polynomial model to these group metrics and compute its reduced chi^2 statistic. Repeater flashes that illuminate the cloud in the same way should have a strong correlation between the brightest pixel within the group footprint and the group illuminated area, as pulses of equal energy at different points in the flash evolution would still generate the same group footprint.

Three flash cases are plotted in Figures 10-12 and animated in S2-S4. These figures resemble XLMA-style plots with a central plan view (d) of GLM group energy (the largest group in Figures 10-12, each group in the animations) with LMA source locations overlaid. Above and to the right of the plan view panel are longitude-altitude (c) and latitude-altitude (e) plots of the LMA sources. Further outward are longitude (a) and latitude (f) cross sections of GLM group energy, where each event on the map is depicted with a square symbol, and the total energy along the cross section is shown as a bar graph. The bottom panels, then show timeseries of LMA source altitude (g) and GLM group energy (i), as well as an overall LMA source altitude distribution for the whole thunderstorm in the 15-minute interval encompassing the flash (h). Finally, the top-right panel (b) shows a scatterplot of GLM group area and group maximum event energy with the polynomial fit (dashed line) overlaid. The group data in (i) and (b) is colored by time with darker groups occurring earlier in the GLM flash and lighter groups occurring near the end. The group shown in (d) is indicated with a red symbol in (b).

Figure 10 shows an example of a repeater-type GLM flash. The polynomial fit in (b) captures the group data with a reduced-chi^2 of 0.16. The flash is comprised of entirely low-altitude LMA sources (< 10 km), and the radiance pattern of the largest group mapped in (d) shows a high level of complexity with a dark center surrounded by a ring of illumination – indicating a dense cloud region above the group center. Local variations in the spatial radiance distribution are not detrimental to generating these repeater flashes. The important factor is to similarly illuminate the surrounding cloud.

Cases of horizontally extensive sources can also be repeater flashes or contain long-lasting repeater series within the larger flash. Figure 11 shows an example of a repeater flash in the top percentile of group-level LMA source extent from Figure 6b. GLM only detected two dim (1-3 fJ) groups in the first 600 ms of flash development, followed by two long-lasting series – each encompassing an ENGLN -CG stroke - where substantial portions of the lightning channels mapped by the LMA were simultaneously illuminated over a ~50 ms period. This case is clearly not from a localized optical source, as the GLM group footprint and its brightest events follow the curved path of the LMA sources from the convective core of the storm in the northeast of the plot to the -CGs in the northwest. Yet, despite the long extent of the source, group area remains a strong function of maximum group energy between each group in the flash. We can still have the optical source similarly illuminate the clouds if the geometry of the source remains constant over time (the animation in S3 shows that this is at least true for the first series) and if the cloud mass surrounding the flash is sufficiently homogeneous (which is expected for stratiform clouds).

While these repeater flashes only represent a small subset of all GLM lightning, generalizing this type of analysis to consider how the area / max. energy distributions change over the flash duration can reveal when flashes develop between clouds regions with different optical characteristics. Figure 12 shows an example flash like the hypothetical case described at the beginning of the section that began in the lower (5 km) charge layer before later developing into the upper (~10 km) charger layer. As predicted, the early groups in the flash described infrequent yet brighter (20-50 fJ max energy per series) illumination of the surrounding clouds with two of the three early series coming from ENGLN -CG strokes. Later activity in the flash (after it developed into the upper layer) produced frequent GLM activity from dim pulses on the order of a few femtojoules. Examining the group area / max. energy distribution in Figure 12b shows that the early groups (dark grey) from the lower layer had relatively low peak optical energies given their reported areas, while later groups (light grey) were particularly energetic for their sizes. This difference in how the clouds are illuminated from sources in each layer causes the distribution in Figure 12b to resemble the two distinct curves that are joined at the top right from the bright (~100 fJ) groups generated while the flash developed vertically between the layers. The separation of these two curves suggests that the illumination from different charge layers is sufficiently distinct to enable classification or even to retrieve the altitudes of optical sources below the cloud top. We will address this possibility later in Part 3 of this study.

5 Discussion and Conclusions

This study combines GLM data with LMA and ENGLN observations to examine how the inferred geometry of the active lightning channel at the time of GLM groups affects how the clouds are illuminated. Two thunderstorms are considered: a thunderstorm in Colombia near the satellite subpoint with favorable conditions for GLM detection, and an inverted-polarity thunderstorm in Colorado with unfavorable conditions for GLM detection.

These cases demonstrate the limits of GLM's ability to measure flash horizontal extent. Under ideal conditions for GLM detection (as in the Colombia case), the GLM maximum distance between group centroids generally provides a reasonable measurement of flash size compared to the LMA flash extent as long as the LMA flash is larger than around one-half of a GLM pixel. By contrast, measuring flash extent as the maximum distance between GLM events (or approximating flash size as the footprint area of the illuminated cloud) generally overestimates the sizes of LMA flashes due to light being scattered across the cloud scene. Under unfavorable conditions for GLM detection (as in the Colorado case), however, both group-based and event-based measurements of flash size underestimate the LMA flash extent because portions of the lightning channel are not resolved by GLM. GLM extents in the Colorado case were frequently reported as 0 km – indicating that only single groups (Figure 5c) or events (Figure 5d) were detected by GLM from the LMA flashes.

The LMA sources matched to GLM groups are used to approximate the portions of the lightning channel that are active during individual optical pulses. The LMA source extent at the group level is usually smaller than the cloud regions illuminated during GLM groups in both the Colombia and Colorado thunderstorms. The median extents of LMA sources matched to a GLM group are 2 km (Colorado) to 3 km (Colombia), which only span a portion of a GLM pixel (nominally 8 km) in either case. Larger optical sources that are one or more GLM pixels across account for the top 10% (Colorado) and 17% (Colombia) of LMA source extents in these two thunderstorms, while the top 1% of LMA source extents span 3-or-more GLM pixels. The most extensive sources come from the large-scale horizontal rearrangement of charge during long-horizontal lightning flashes where the GLM group footprints trace out the paths of the illuminated lightning channels through the cloud (Figure 11 shows an example of this type of illumination associated with a -CG stroke). Therefore, the approximation of optical emitters as localized sources (which might be approximated as point sources on the scale of GLM pixels) is generally reasonable (especially for convective flashes), but this assumption does not always hold for flashes outside of the convective core that propagate horizontally over considerable distances.

Both thunderstorm cases show a bias in GLM detection towards high-altitude sources. Compared to the overall LMA source altitude distribution, the GLM-matched LMA source distribution was notably amplified at high altitudes and suppressed at low altitudes (except in cases where light escapes the sides of the cloud). The changeover altitude between amplification and suppression depends on the storm in question, and was between 7 and 10 km over th durations of the two storms examined. However, this GLM detection advantage for high-altitude sources does not mean that GLM detects *only* high-altitude sources. Low-altitude pulses generate GLM groups as well, and these detections largely depend on the source intensity. Even in the Colombia case where most of the lightning activity occurred in the upper 10-km layer, the largest and most radiant GLM groups were at least equally likely to originate from the lower layer at 5 km. While increased scattering and absorption in the cloud medium can attenuate weaker signals from the lower layer to the point where they are not detected by GLM, the most intense pulses – including but not limited to strokes – are still detected by GLM.

These results support the concept that the brightness of the optical source and the nature of the cloud medium between the source and satellite have a greater impact on how the resulting groups appear from orbit than the geometry of the optical source in most cases. This is why low-altitude groups often have considerable spatial variations in their spatial energy distributions from the optical emissions interacting with thick cloud depths. It also explains how we can find "repeater" flashes where group illuminated area is a strong function of group maximum event energy – even as the flash develops horizontally through the cloud over time. The spatial structure of GLM groups has been infrequently studied, but it is key to understanding how clouds are illuminated by lightning. Understanding this aspect of GLM measurements will potentially lead to new GLM applications for describing lightning and its surrounding cloud medium.

Acknowledgments

This work was supported by the US Department of Energy through the Los Alamos National Laboratory (LANL) Laboratory Directed Research and Development (LDRD) program under project number 20200529ECR. Los Alamos National Laboratory is operated by Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001). The work by co-author Douglas Mach was supported by ASA 80MFSC17M0022 “Cooperative Agreement with Universities Space Research Association” and NASA Research Opportunities in Space and Earth Science grant NNX17AJ10G “U.S. and European Geostationary Lightning Sensor Cross-Validation Study.” We would like to thank the operators of the Colorado LMA at Colorado State University and the Colombia LMA at the Technical University of Catalonia, and Drs. Ken Cummins and Jesús López for sharing their processed LMA data for the presented cases. The data used in this study is available at Peterson (2021b).

References

Albrecht, R. I., Goodman, S. J., Buechler, D. E., Blakeslee, R. J., & Christian, H. J. (2016). Where are the lightning hotspots on Earth?. *Bulletin of the American Meteorological Society*, *97*(11), 2051-2068.

Aranguren, D., Lopez, J., Montanya, J., & Torres, H. (2018, September). Natural observatories for lightning research in Colombia. In *2018 International Conference on Electromagnetics in Advanced Applications (ICEAA)* (pp. 279-283). IEEE.

Blakeslee, R.J., Lang, T.J., Koshak, W.J., Buechler, D., Gatlin, P., Mach, D.M., Stano, G.T., Virts, K.S., Walker, T.D., Cecil, D.J., Ellett, W., Goodman, S.J., Harrison, S., Hawkins, D.L., Heumesser, M., Lin, H., Maskey, M., Schultz, C.J., Stewart, M., Bateman, M., Chanrion, O. and Christian, H. (2020), Three Years of the Lightning Imaging Sensor Onboard the International Space Station: Expanded Global Coverage and Enhanced Applications. *J. Geophys. Res. Atmos*., **125**: e2020JD032918. https://doi.org/10.1029/2020JD032918

Boccippio, D. J., Koshak, W. J., & Blakeslee, R. J. (2002). Performance Assessment of the Optical Transient Detector and Lightning Imaging Sensor. Part I: Predicted Diurnal Variability, *Journal of Atmospheric and Oceanic Technology*, *19*(9), 1318-1332. https://doi.org/10.1175/1520-0426(2002)019<1318:PAOTOT>2.0.CO;2

Boggs, L. D., Liu, N., Peterson, M., Lazarus, S., Splitt, M., Lucena, F., ... & Rassoul, H. K. (2019). First observations of gigantic jets from geostationary orbit. *Geophysical Research Letters*, *46*(7), 3999-4006. https://doi.org/10.1029/2019GL082278.

Bruning, E. C., & MacGorman, D. R. (2013). Theory and Observations of Controls on Lightning Flash Size Spectra, *Journal of the Atmospheric Sciences*, *70*(12), 4012-4029. https://doi.org/10.1175/JAS-D-12-0289.1.

Bruning, E., Tillier, C. E., Edgington, S. F., Rudlosky, S. D., Zajic, J., Gravelle, C., et al. (2019). Meteorological imagery for the geostationary lightning mapper. *Journal of Geophysical Research: Atmospheres*, 2019; 124: 14285– 14309. https://doi.org/10.1029/2019JD030874.

Cecil, D. J., Buechler, D. E., & Blakeslee, R. J. (2014). Gridded lightning climatology from TRMM-LIS and OTD: Dataset description. *Atmospheric Research*, *135*, 404-414. https://doi.org/10.1016/j.atmosres.2012.06.028

Chen, A. B., et al. (2008), Global distributions and occurrence rates of transient luminous events, *J. Geophys. Res.*, 113, A08306, https://doi.org/10.1029/2008JA013101.

Christian, H. J., R. J. Blakeslee, S. J. Goodman, and D. M. Mach (Eds.), 2000: Algorithm Theoretical Basis Document (ATBD) for the Lightning Imaging Sensor (LIS), NASA/Marshall Space Flight Center, Alabama. (Available as http://eospso.gsfc.nasa.gov/atbd/listables.html, posted 1 Feb. 2000)

da Silva, C. L., Sonnenfeld, R. G., Edens, H. E., Krehbiel, P. R., Quick, M. G., & Koshak, W. J. (2019). The plasma nature of lightning channels and the resulting nonlinear resistance. *Journal of Geophysical Research: Atmospheres*, 124, 9442– 9463. https://doi.org/10.1029/2019JD030693

Fabró, F., Montanyà, J., Marisaldi, M., van der Velde, O. A., & Fuschino, F. (2015). Analysis of global Terrestrial Gamma Ray Flashes distribution and special focus on AGILE detections over South America. *Journal of Atmospheric and Solar-Terrestrial Physics*, *124*, 10-20. https://doi.org/10.1016/j.jastp.2015.01.009.

Fierro, A. O., Gao, J., Ziegler, C. L., Calhoun, K. M., Mansell, E. R., & MacGorman, D. R. (2016). Assimilation of Flash Extent Data in the Variational Framework at Convection-Allowing Scales: Proof-of-Concept and Evaluation for the Short-Term Forecast of the 24 May 2011 Tornado Outbreak, *Monthly Weather Review*, **144**(11), 4373-4393. https://doi.org/10.1175/MWR-D-16-0053.1.

Goodman, S. J., D. Mach, W. J. Koshak, and R. J. Blakeslee. (2010). *GLM Lightning Cluster-Filter Algorithm (LCFA) Algorithm Theoretical Basis Document (ATBD)*. Retrieved from https://www.goes-r.gov/products/ATBDs/baseline/Lightning_v2.0_no_color.pdf, posted 24 Sept. 2010

Goodman, S. J., D. Mach, W. J. Koshak, and R. J. Blakeslee. (2010). *GLM Lightning Cluster-Filter Algorithm (LCFA) Algorithm Theoretical Basis Document (ATBD)*. Retrieved from https://www.goes-r.gov/products/ATBDs/baseline/Lightning_v2.0_no_color.pdf, posted 24 Sept. 2010

Goodman, S. J., R. J. Blakeslee, W. J. Koshak, D. Mach, J. Bailey, D. Buechler, L. Carey, C. Schultz, M. Bateman, E. McCaul Jr., and G. Stano, 2013: The GOES-R geostationary lightning mapper (GLM). *J. Atmos. Res.,* **125-126**, 34-49. https://doi.org/10.1016/j.atmosres.2013.01.006.

Hutchins, M. L., Holzworth, R. H., Brundell, J. B., and Rodger, C. J. (2012), Relative detection efficiency of the World Wide Lightning Location Network, *Radio Sci.*, 47, RS6005, https://doi.org/10.1029/2012RS005049.

Jacobson, A.R., R. Holzworth, J. Harlin, R. Dowden, and E. Lay, 2006: Performance Assessment of the World Wide Lightning Location Network (WWLLN), Using the Los Alamos Sferic Array (LASA) as Ground Truth. *J. Atmos. Oceanic Technol.,* **23**, 1082–1092, https://doi.org/10.1175/JTECH1902.1.

Koshak, W. J., 2010: Optical characteristics of OTD flashes and the implications for flash-type discrimination. *J. Atmos. Oceanic. Technol.,* **27,** 1,822 – 1,838. https://doi.org/10.1175/2010JTECHA1405.1.

Koshak, W. J., Solakiewicz, R. J., Phanord, D. D., and Blakeslee, R. J. (1994), Diffusion model for lightning radiative transfer, *J. Geophys. Res.*, 99(D7), 14361– 14371, https://doi.org/10.1029/94JD00022.

Koshak, W. J., D. Mach, M. Bateman, P. Armstrong, and K. Virts (2018). GOES-16 GLM Level 2 DataFull Validation Data Quality: Product Performanc Guide For Data Users. https://www.ncdc.noaa.gov/sites/default/files/attachments/GOES16_GLM_FullValidation_ProductPerformanceGuide.pdf

Lang, T. J., Pédeboy, S., Rison, W., Cerveny, R. S., Montanyà, J., Chauzy, S., ... & Krahenbuhl, D. S. (2017). WMO world record lightning extremes: Longest reported flash distance and longest reported flash duration. *Bulletin of the American Meteorological Society*, *98*(6), 1153-1168. https://doi.org/10.1175/BAMS-D-16-0061.1.

Lay, E. H., Holzworth, R. H., Rodger, C. J., Thomas, J. N., Pinto, O., and Dowden, R. L. (2004), WWLL global lightning detection system: Regional validation study in Brazil, *Geophys. Res. Lett.*, 31, L03102, doi:10.1029/2003GL018882.

Light, T. E., Suszcynsky, D. M., and Jacobson, A. R. (2001a), Coincident radio frequency and optical emissions from lightning, observed with the FORTE satellite, *J. Geophys. Res.*, 106(D22), 28223– 28231, https://doi.org/10.1029/2001JD000727.

Light, T. E., Suszcynsky, D. M., Kirkland, M. W., and Jacobson, A. R. (2001b), Simulations of lightning optical waveforms as seen through clouds by satellites, *J. Geophys. Res.*, 106(D15), 17103– 17114, https://doi.org/10.1029/2001JD900051.

López, J. A., Montanyà, J., van der Velde, O., Romero, D., Aranguren, D., Torres, H., ... & Martinez, J. (2016, September). First data of the Colombia lightning mapping array—COLMA. In *2016 33rd International Conference on Lightning Protection (ICLP)* (pp. 1-5). IEEE.

López, J. A., Montanyà, J., van der Velde, O. A., Pineda, N., Salvador, A., Romero, D., et al. (2019). Charge structure of two tropical thunderstorms in Colombia. *Journal of Geophysical Research: Atmospheres*, 124, 5503– 5515. https://doi.org/10.1029/2018JD029188

Lyons, W. A., Bruning, E. C., Warner, T. A., MacGorman, D. R., Edgington, S., Tillier, C., & Mlynarczyk, J. (2020). Megaflashes: Just how long can a lightning discharge get?. *Bulletin of the American Meteorological Society*, *101*(1), E73-E86. https://doi.org/10.1175/BAMS-D-19-0033.1.

Neubert, T., Østgaard, N., Reglero, V., Blanc, E., Chanrion, O., Oxborrow, C. A., ... & Bhanderi, D. D. (2019). The ASIM mission on the international space station. *Space Science Reviews*, *215*(2), 26. https://doi.org/10.1007/s11214-019-0592-z.

Peterson, M. (2019). Research applications for the Geostationary Lightning Mapper operational lightning flash data product. *Journal of Geophysical Research: Atmospheres*, 124, 10205– 10231. https://doi.org/10.1029/2019JD031054

Peterson, M. (2020a). Modeling the transmission of optical lightning signals through complex 3-D cloud scenes. *Journal of Geophysical Research: Atmospheres*, 125, e2020JD033231. https://doi.org/10.1029/2020JD033231

Peterson, M. (2020b). Holes in Optical Lightning Flashes: Identifying Poorly-Transmissive Clouds in Lightning Imager Data. *Earth and Space Science*, 7, e2020EA001294. https://doi.org/10.1029/2020EA001294.

Peterson, M. (2021a). GLM-CIERRA http://dx.doi.org/10.5067/GLM/CIERRA/DATA101

Peterson, M. (2021b). Coincident Optical and RF Lightning Detections from a Colombia Thunderstorm. https://doi.org/10.7910/DVN/5FR6JB, Harvard Dataverse, V1

Peterson, M. J., & Lay, E. H. (2020). Geostationary Lightning Mapper (GLM) Observations of the Brightest Lightning in the Americas. Journal of Geophysical Research: Atmospheres, 125, e2020JD033378. https://doi.org/10.1029/2020JD033378

Peterson, M., Deierling, W., Liu, C., Mach, D., and Kalb, C. (2017a), The properties of optical lightning flashes and the clouds they illuminate, *J. Geophys. Res. Atmos.*, 122, 423– 442, https://doi.org/10.1002/2016JD025312.

Peterson, M., Rudlosky, S., & Deierling, W. (2017b). The evolution and structure of extreme optical lightning flashes. *Journal of Geophysical Research: Atmospheres*, 122, 13,370–13,386. https://doi.org/10.1002/2017JD026855

Peterson, M., Rudlosky, S., & Deierling, W. (2018). Mapping the lateral development of lightning flashes from orbit. *Journal of Geophysical Research: Atmospheres*, 123, 9674–9687. https://doi.org/10.1029/2018JD028583

Peterson, M., Rudlosky, S., & Zhang, D. (2020a). Changes to the appearance of optical lightning flashes observed from space according to thunderstorm organization and structure. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD031087. https://doi.org/10.1029/2019JD031087

Peterson, M., Rudlosky, S., & Zhang, D. (2020b). Thunderstorm Cloud-Type Classification from Space-Based Lightning Imagers, *Monthly Weather Review*, *148*(5), 1891-1898. https://doi.org/10.1175/MWR-D-19-0365.1.

Peterson, M. J., Lang, T. J., Bruning, E. C., Albrecht, R., Blakeslee, R. J., Lyons, W. A., et al. (2020c). New World Meteorological Organization certified megaflash lightning extremes for flash distance (709 km) and duration (16.73 s) recorded from space. *Geophysical Research Letters*, 47, e2020GL088888. https://doi.org/10.1029/2020GL088888

Peterson, M., Mach, D., & Buechler, D. (2021a). A Global LIS/OTD Climatology of Lightning Flash Extent Density. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033885. https://doi.org/10.1029/2020JD033885

Peterson, M., Light, T., & Mach, D. (2021b). The Illumination of Thunderclouds by Lightning: Part 2: The Effect of GLM Instrument Threshold on Detection and Clustering. *Journal of Geophysical Research: Atmospheres*.

Peterson, M., Light, T., & Mach, D. (2021c). The Illumination of Thunderclouds by Lightning: Part 3: Retrieving Optical Source Altitude. *Journal of Geophysical Research: Atmospheres*.

Peterson, M., & Mach, D. (2021d). The Illumination of Thunderclouds by Lightning: Part 4: Volumetric Thunderstorm Imagery. *Journal of Geophysical Research: Atmospheres*.

Rison, W., R. J. Thomas, P. R. Krehbiel, T. Hamlin, and J. Harlin, 1999: A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico. *Geophys. Res. Lett., **26***(23), 3573-3576. https://doi.org/10.1029/1999gl010856.

Rison, W., Krehbiel, P. R., Thomas, R. J., Rodeheffer, D., & Fuchs, B. (2012). The Colorado lightning mapping array. *AGUFM*, *2012*, AE23B-0319. https://doi.org/10.1175/WAF-D-15-0037.1.

Rodger, C. J., Werner, S., Brundell, J. B., Lay, E. H., Thomson, N. R., Holzworth, R. H., & Dowden, R. L. (2006, December). Detection efficiency of the VLF World-Wide Lightning Location Network (WWLLN): initial case study. In *Annales Geophysicae* (Vol. 24, No. 12, pp. 3197-3214). Copernicus GmbH.

Rudlosky, S. D., S. J. Goodman, K. S. Virts, and E. C. Bruning, 2019: Initial geostationary lightning mapper observations. *Geophys. Res. Lett.*, **46**, 1097– 1104. https://doi.org/10.1029/2018GL081052

Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., & Lebair, W. J. (2017). A closer look at the ABI on the GOES-R series. *Bulletin of the American Meteorological Society*, *98*(4), 681-698. https://doi.org/10.1175/BAMS-D-15-00230.1.

Splitt, M. E., Lazarus, S. M., Barnes, D., Dwyer, J. R., Rassoul, H. K., Smith, D. M., Hazelton, B., and Grefenstette, B. (2010), Thunderstorm characteristics associated with RHESSI identified terrestrial gamma ray flashes, *J. Geophys. Res.*, 115, A00E38, https://doi.org/10.1029/2009JA014622.

Suszcynsky, D. M., Light, T. E., Davis, S., Green, J. L., Guillen, J. L. L., and Myre, W. (2001), Coordinated observations of optical lightning from space using the FORTE photodiode detector and CCD imager, *J. Geophys. Res.*, 106(D16), 17897– 17906, https://doi.org/10.1029/2001JD900199.

Suszcynsky, D. M., Kirkland, M. W., Jacobson, A. R., Franz, R. C., Knox, S. O., Guillen, J. L. L., and Green, J. L. (2000), FORTE observations of simultaneous VHF and optical emissions from lightning: Basic phenomenology, *J. Geophys. Res.*, 105(D2), 2191–2201, https://doi.org/10.1029/1999JD900993.

Thiel, K. C., Calhoun, K. M., Reinhart, A. E., & MacGorman, D. R. (2020). GLM and ABI characteristics of severe and convective storms. *Journal of Geophysical Research: Atmospheres*, **125**, e2020JD032858. https://doi.org/10.1029/2020JD032858

Thomson, L.W. and E.P. Krider, 1982: The Effects of Clouds on the Light Produced by Lightning. *J. Atmos. Sci.,***39**, 2051–2065, https://doi.org/10.1175/1520-0469(1982)039<2051:TEOCOT>2.0.CO;2

van der Velde, O. A., & Montanyà, J. (2013). Asymmetries in bidirectional leader development of lightning flashes. *Journal of Geophysical Research: Atmospheres*, *118*(24), 13-504. https://doi.org/ 10.1002/2013JD020257.

Zhang, D., & Cummins, K. L. (2020). Time evolution of satellite-based optical properties in lightning flashes, and its impact on GLM flash detection. *Journal of Geophysical Research: Atmospheres*, *125*(6). https://doi.org/10.1029/2019JD032024.

Zhu, Y., Rakov, V. A., Tran, M. D., Stock, M. G., Heckman, S., Liu, C., … Hare, B. M. (2017). Evaluation of ENTLN performance characteristics based on the ground truth natural and rocket-triggered lightning data acquired in Florida. *Journal of Geophysical Research: Atmospheres*, 122. 9858– 9866, https://doi.org/10.1002/2017JD027270.

Table 1. Frequencies of GLM groups and flashes matching with ENGLN strokes and LMA sources during the Colombia and Colorado thunderstorm cases.

	All GLM Features	GLM Features Matched with ENGLN Strokes		GLM Features Matched with LMA Sources	
	Total	Total	Percent	Total	Percent
		Colombia Case			
GLM Flashes	2154	471	21.9	1942	90.1
GLM Groups	56399	631	1.1	22681	40.2
		Colorado Case			
GLM Flashes	5278	767	14.5	5275	99.9
GLM Groups	28335	744	2.6	19831	70.0

Table 2. Frequencies of ENGLN CGs and LMA sources matching with GLM flashes and groups during the Colombia and Colorado thunderstorm cases.

	All RF Events	RF Events Matched with GLM Flashes		RF Events Matched with GLM Groups	
	Total	Total	Percent	Total	Percent
		Colombia Case			
ENGLN -CGs	1246	1013	81.3	621	49.8
ENGLN +CGs	71	49	69.0	13	18.3
LMA Sources	376482	364851	96.9	181049	48.1
		Colorado Case			
ENGLN -CGs	3123	1123	35.9	720	23.0
ENGLN +CGs	104	53	51.0	41	39.4
LMA Sources	5658247	1287623	22.7	161204	2.8

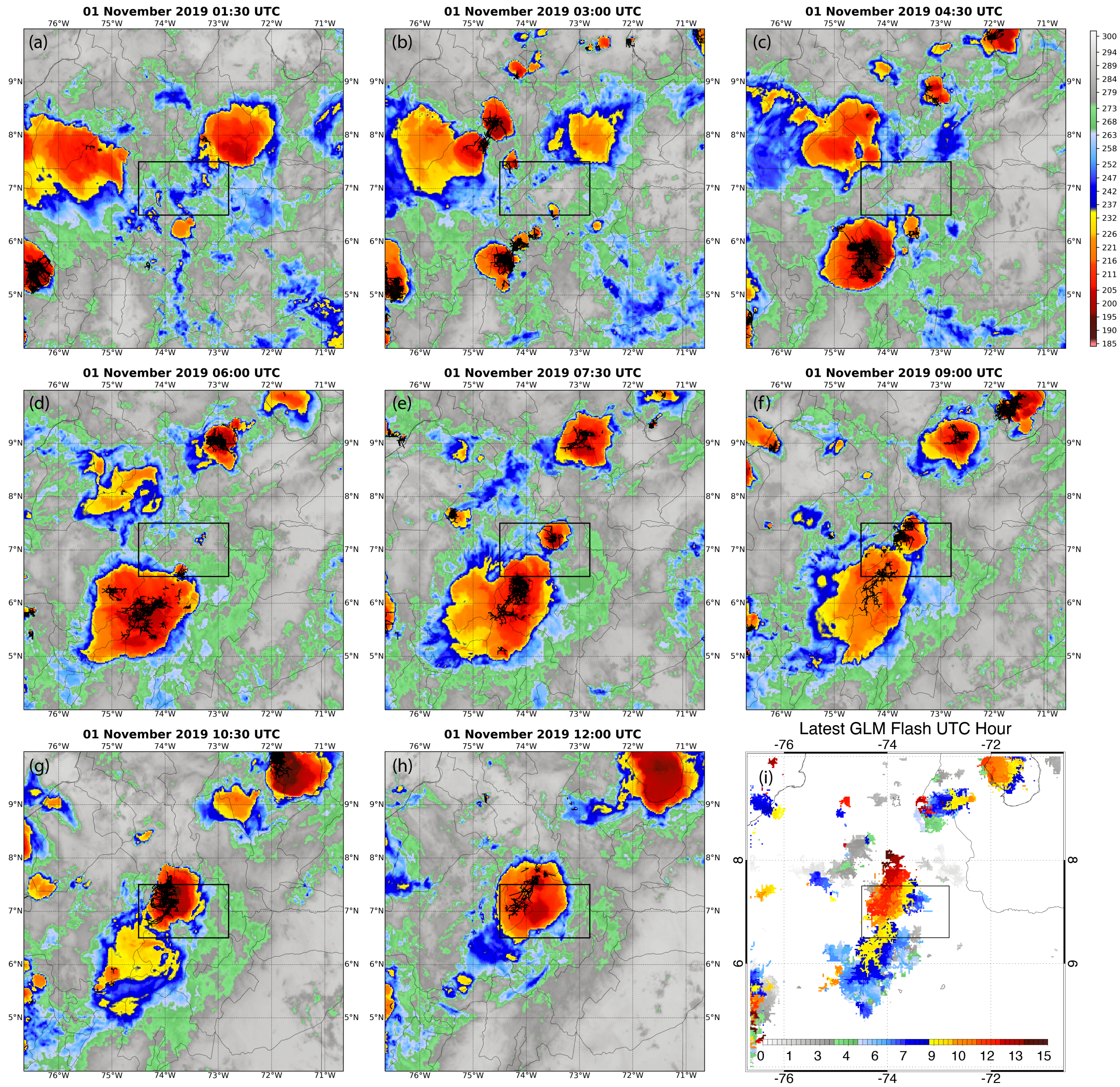


Figure 1. GLM lightning activity (black line segments showing group extent) and ABI Channel 14 (11.2 µm) infrared brightness temperatures (color contours) over the history of the Colombia thunderstorm on 01 November 2019 (a-h) and the time of latest lightning over the mapped region (i). The black boxed region shows the domain where LMA data are available.

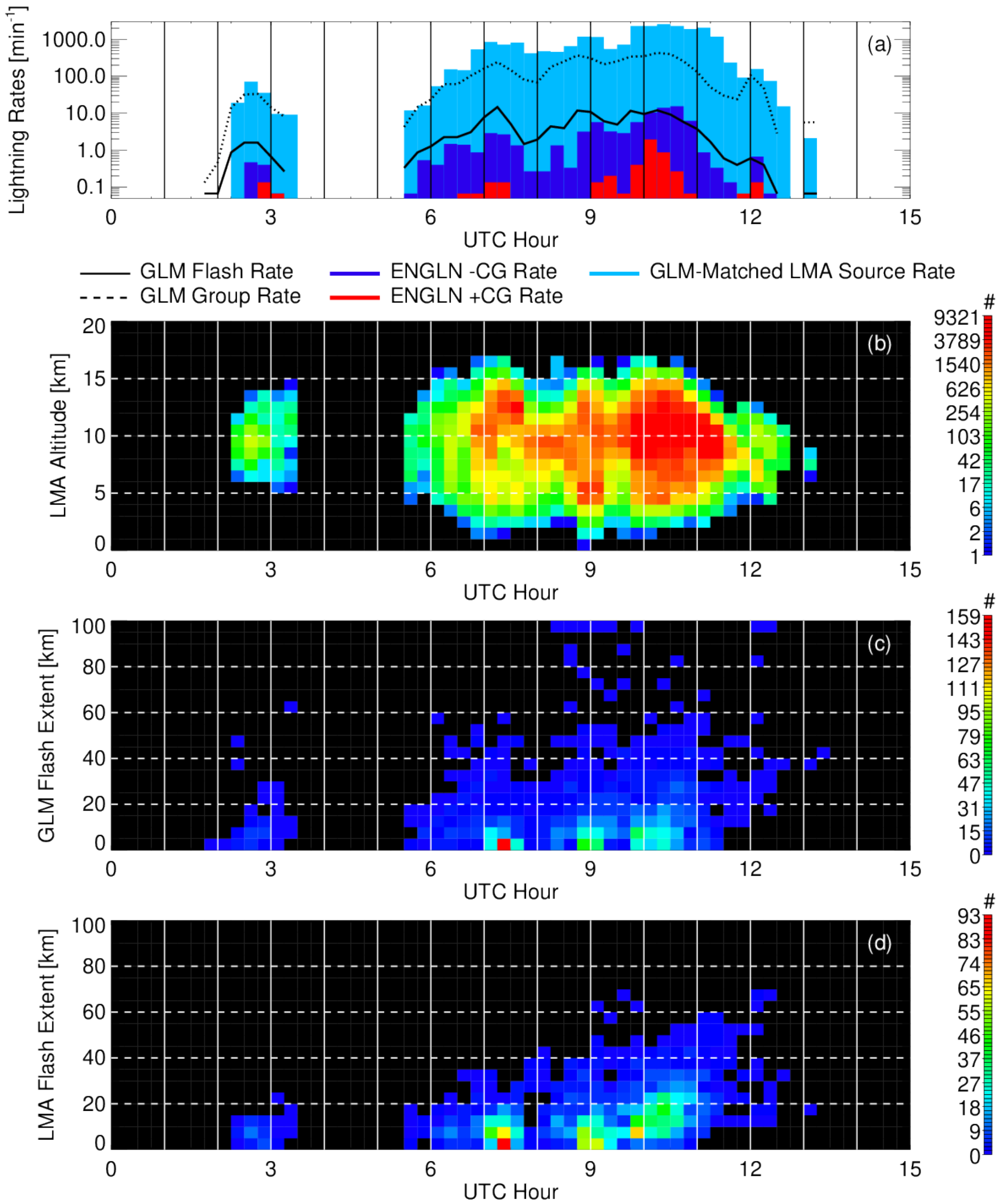


Figure 2. Timeseries of (a) GLM, ENGLN, and LMA lightning rates, (b) LMA source altitude distributions, (c) GLM flash extent distributions, and (d) LMA flash extent distributions for the Colombia thunderstorm case.

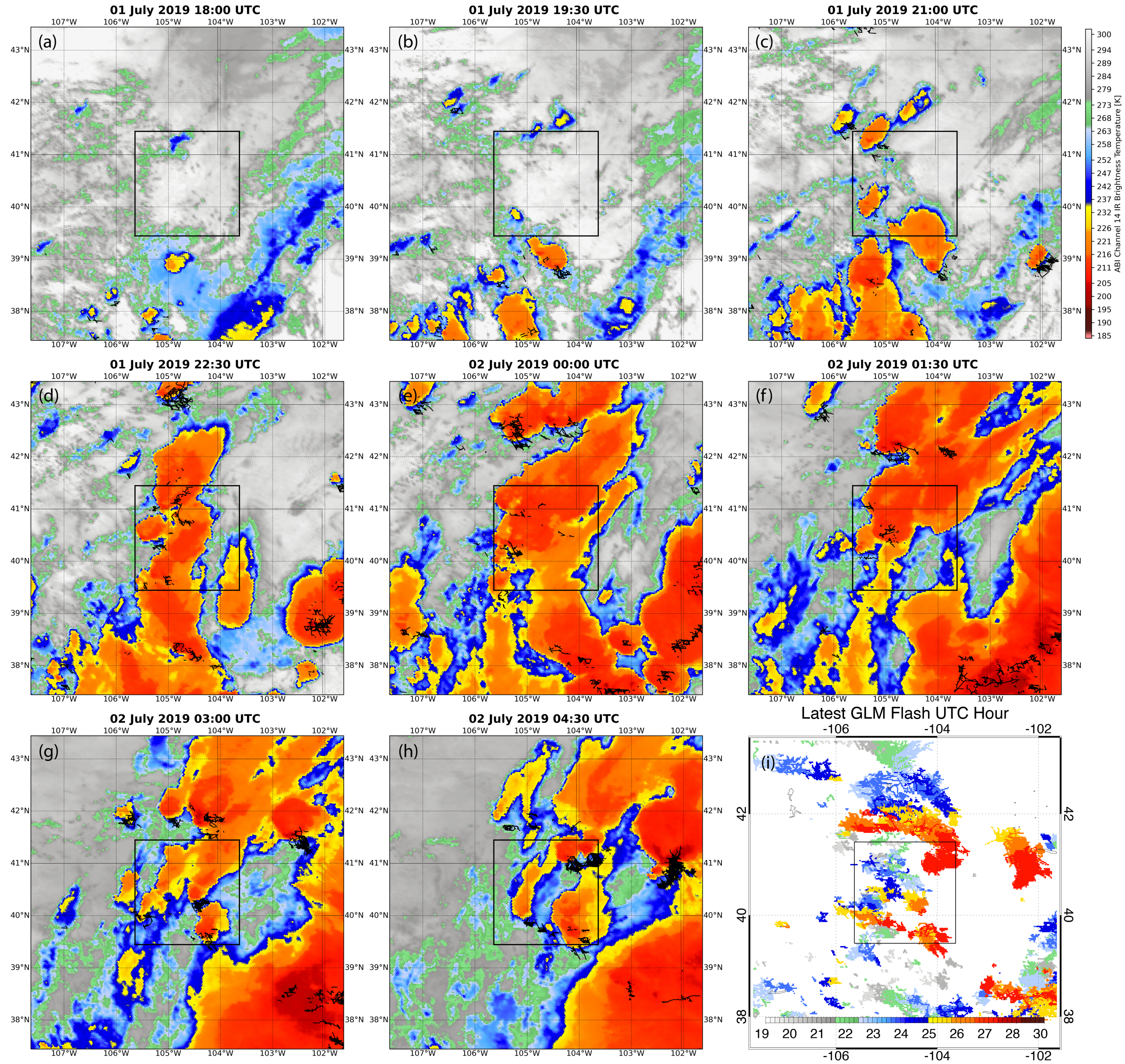


Figure 3. As in Figure 1, but showing the history of the Colorado case on 01 July 2019 – 02 July 2019.

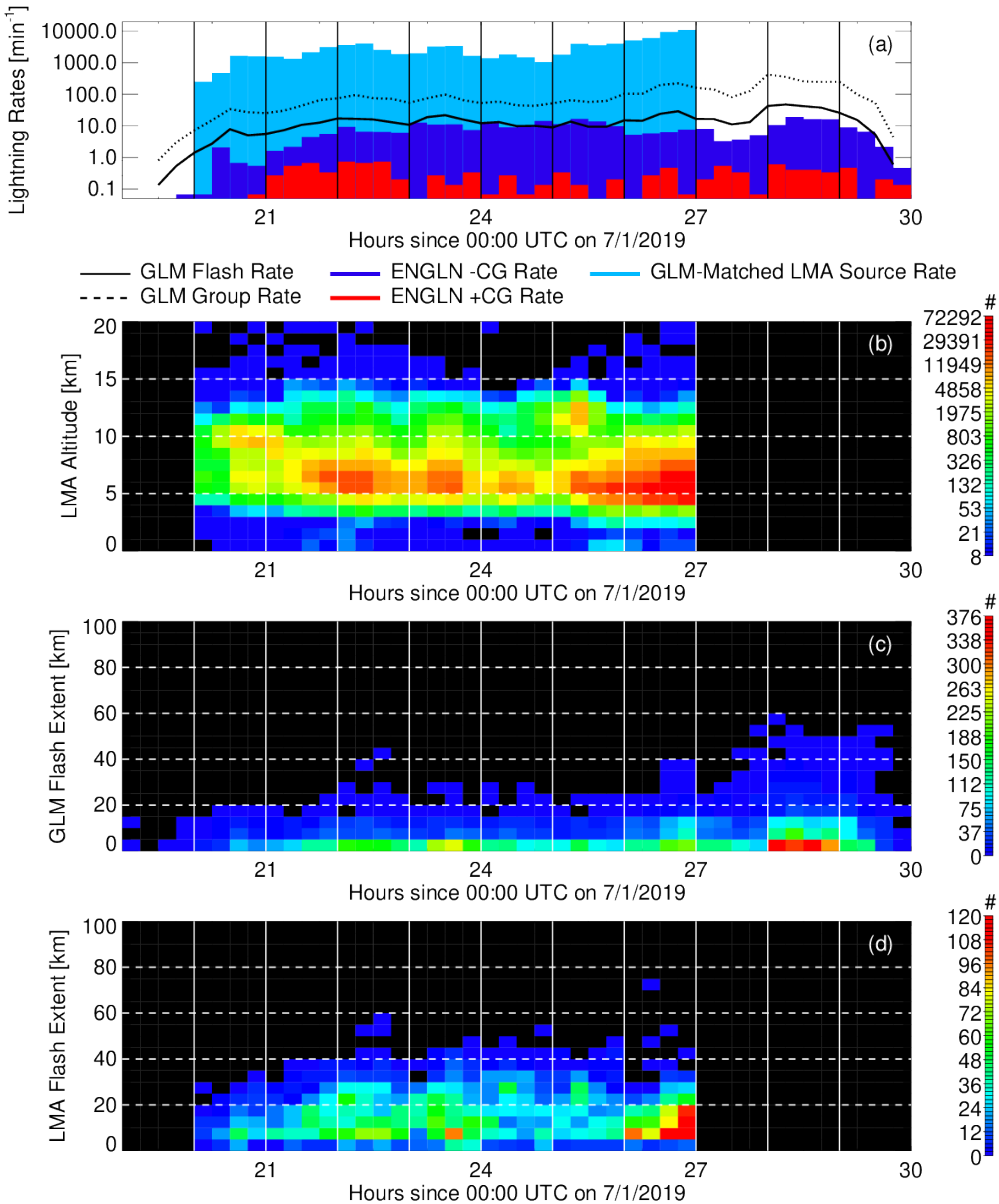


Figure 4. Timeseries of lightning frequency and flash characteristics for the Colorado case following the format of Figure 2.

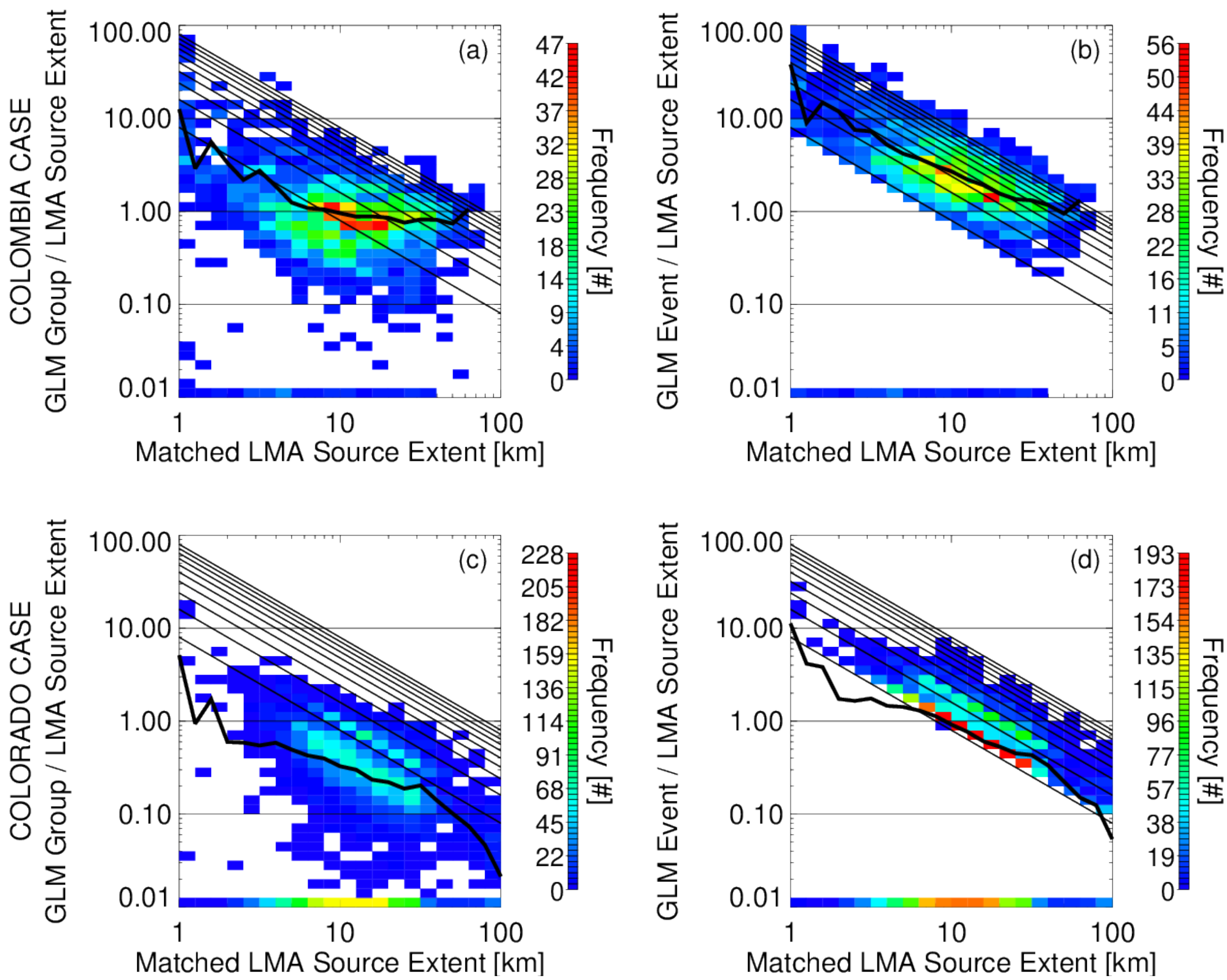


Figure 5. Two-dimensional histograms of the overall extent of LMA flashes matched with GLM flashes and the ratio of LMA : GLM flash extent measured using group centroid locations (a,c) and event locations (b,d). The Colombia case is shown in (a,b) while the Colorado case is shown in (c,d). GLM flashes with extents of 0 km are shown along the bottom of the histograms, while the average ratio for each flash size is depicted with a solid line overlay. Slanted lines indicate constant distances corresponding to the sizes of 1-10 GLM pixels.

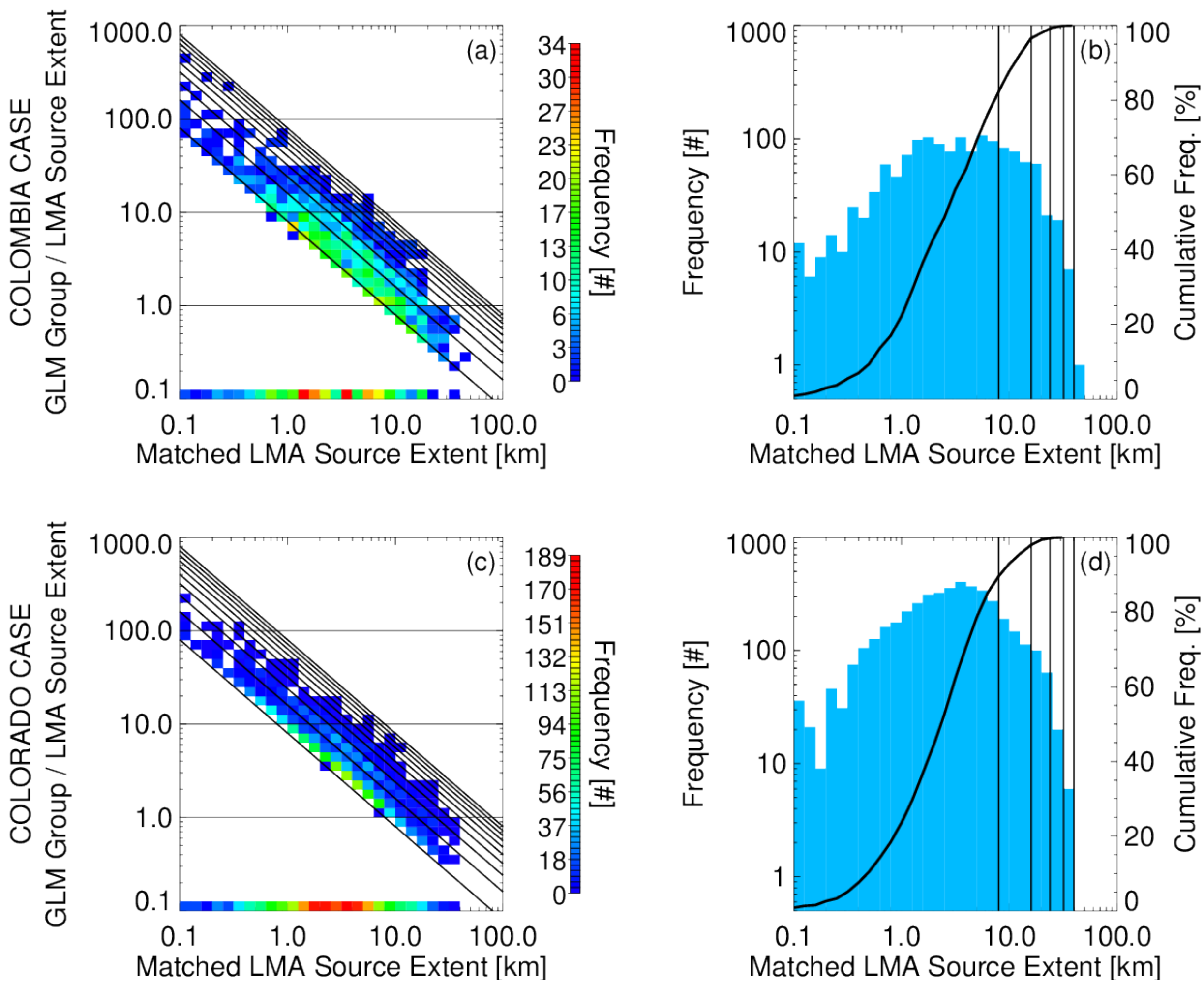


Figure 6. Histograms of GLM group extent and the extent of coincident LMA sources. Two-dimensional histograms in the style of Figure 5b and d are shown in (a) and (c) for groups rather than flashes. Histograms and Cumulative Density Functions (CDFs) of matched LMA source extent are shown in (b) and (d). As in Figure 5, the Colombia case is shown in (a,b) while the Colorado case is shown in (c,d).

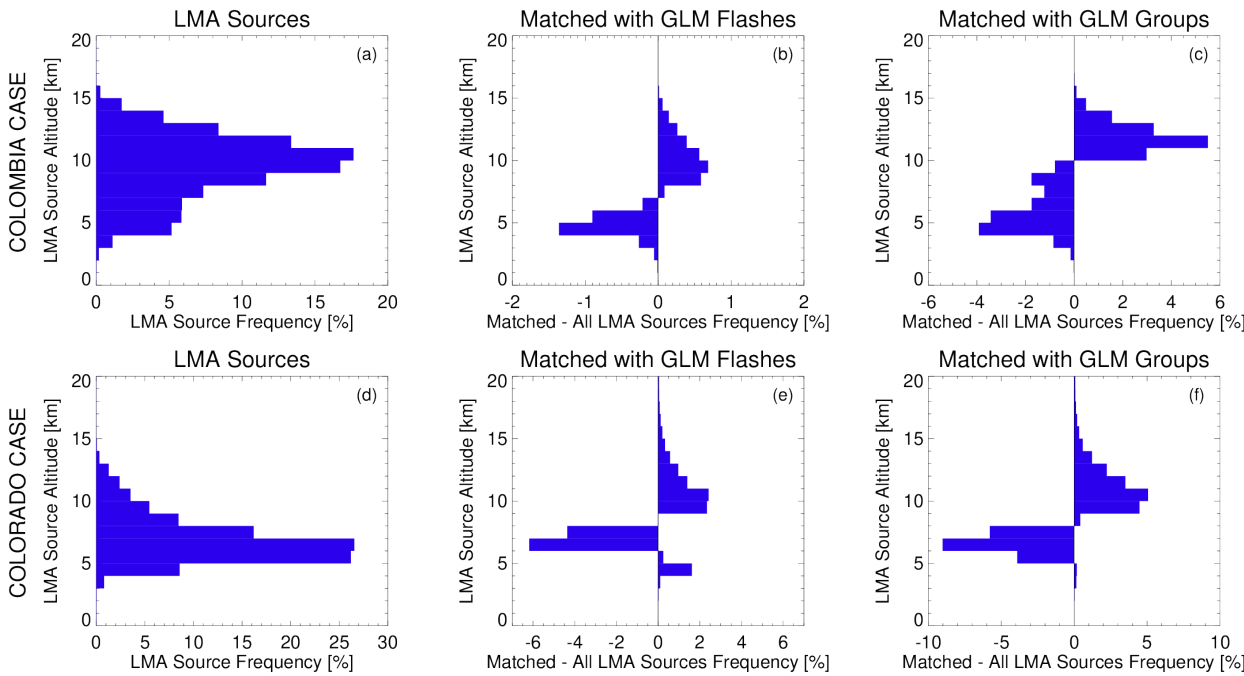


Figure 7. LMA source altitude distributions (a,d) and departures from the overall altitude distribution for LMA sources matched to GLM flashes (b,e) and groups (c,f).

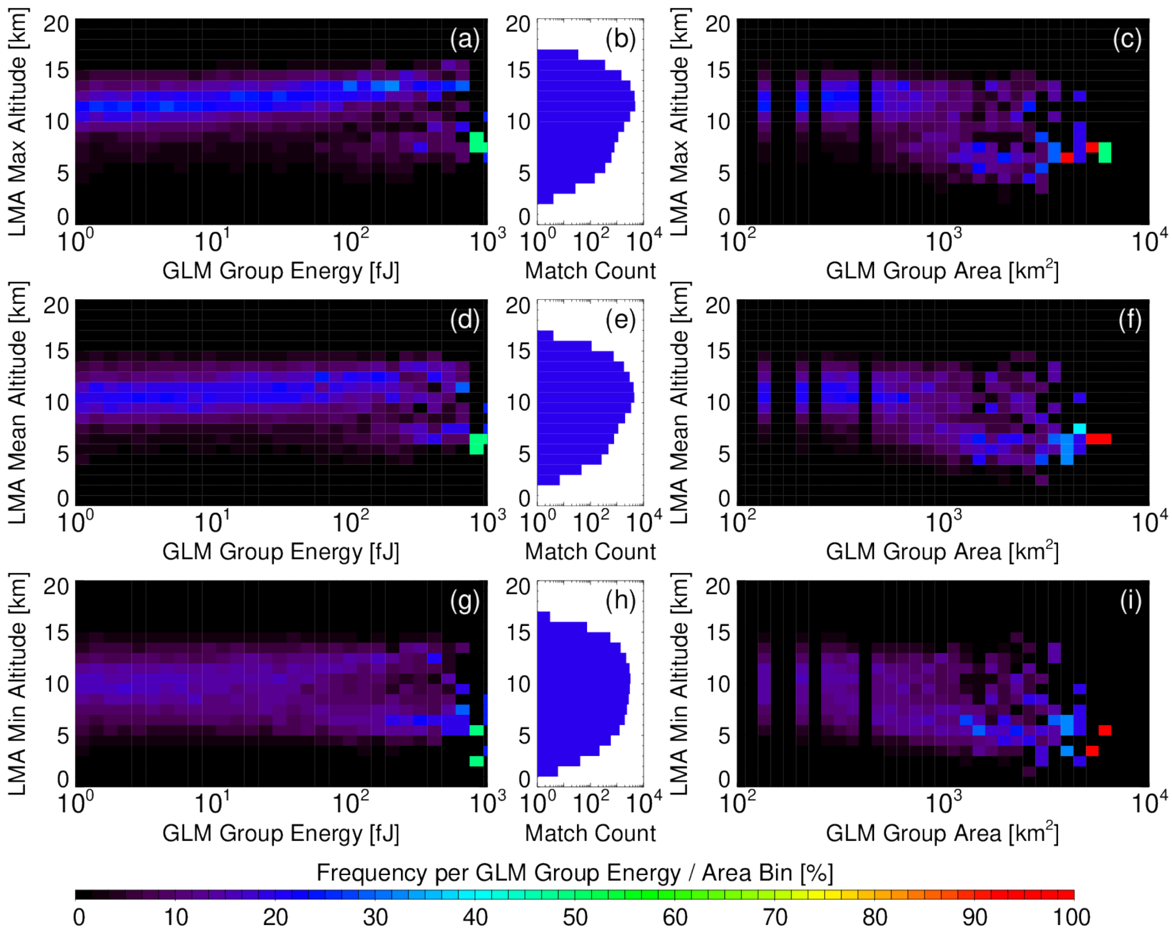


Figure 8. LMA maximum (a-c), mean (d-f), and minimum (g-i) source altitude distributions for matched GLM groups from the Colombia case at various energies (a,d,g) and areas (c,f,i). Vertical frequencies in the contour plots sum to 100% for each energy or area value shown. The central panels (b,e,h) show the overall matched LMA source altitude distributions for all GLM groups.

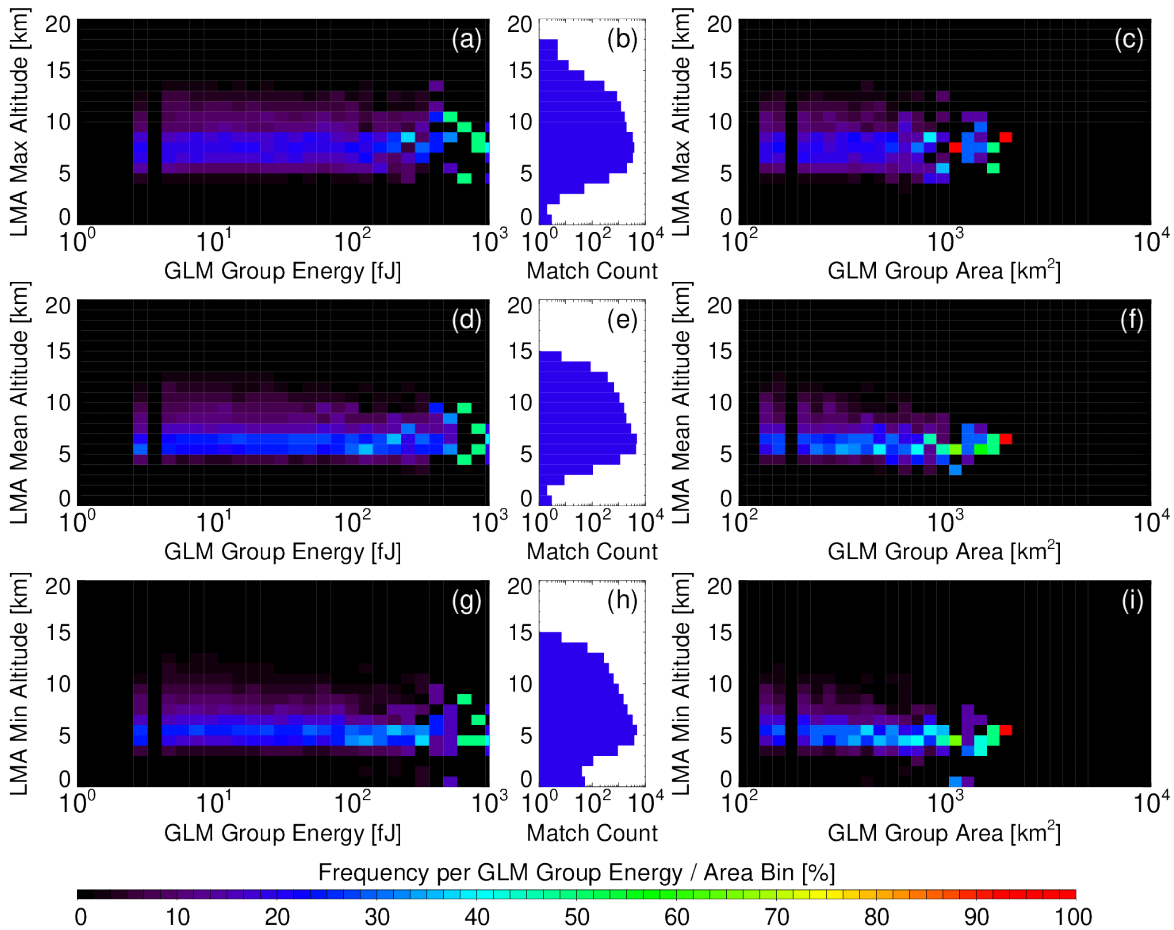


Figure 9. As with Figure 8, but for the Colorado case.

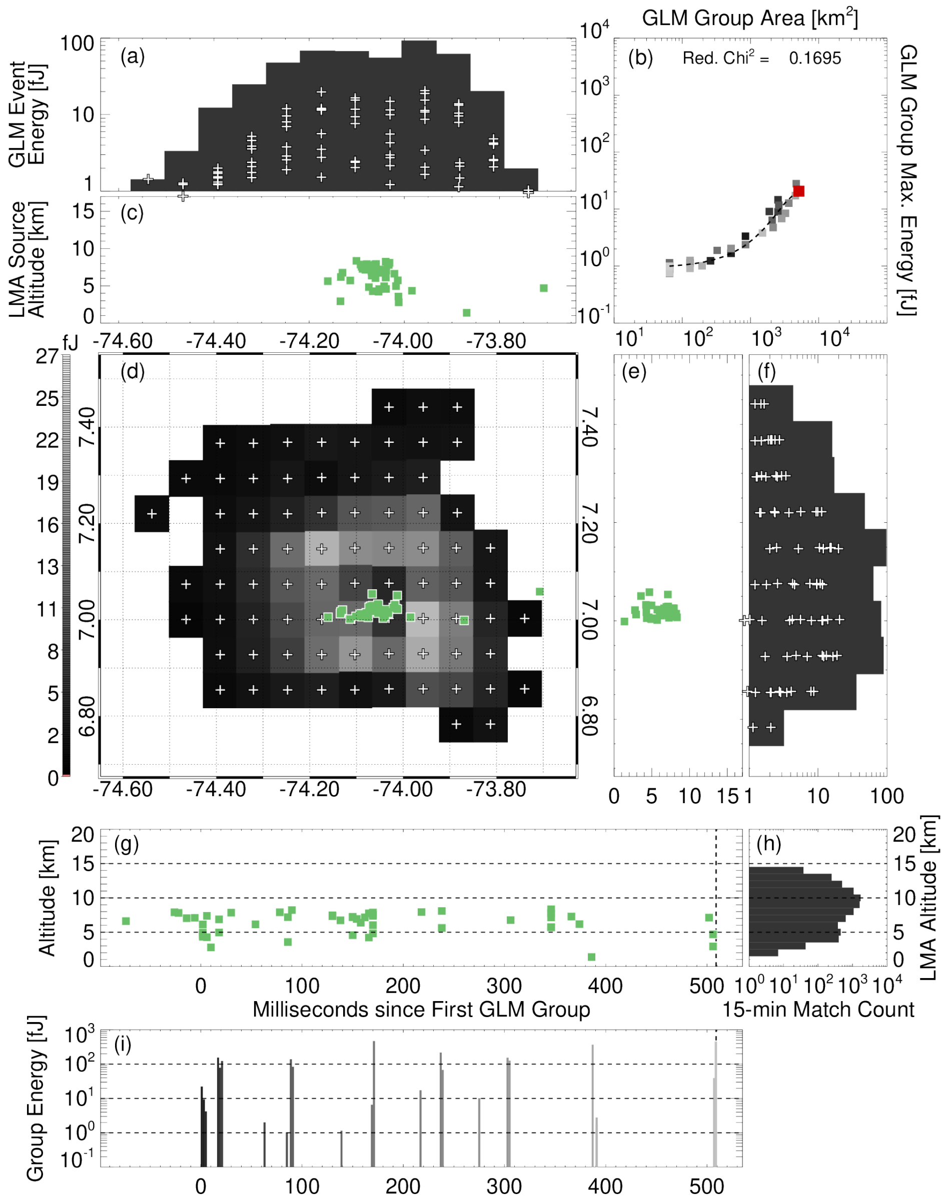


Figure 10. Combined GLM and LMA flash evolution plot for a case during the Colombia thunderstorm where subsequent groups illuminated the cloud in a consistent manner. The central panel (d) maps the spatial GLM event energy distribution during the largest group in the flash (greyscale pixels) and LMA flash structure (small green boxes). The panels above (d) show longitude-altitude LMA source distributions (c) and GLM event energy distributions (a) that include individual events (plus symbols) and total energy (bars). (e) and (f) do the same for latitude. The panels below (d) show timeseries of LMA source altitude (g) and GLM group energy (i) as well as the overall LMA source altitude distribution during the 15-minute period containing the flash (h). Finally, a scatterplot of GLM group area and maximum event energy is shown in (b) with a polynomial fit overlaid as a dashed line. GLM groups are colored according to time from dark to light. The current group is marked in (b) with a red box. ENGLN strokes, if present, are indicated with blue (red) asterisk symbols for -CGs (+CGs).

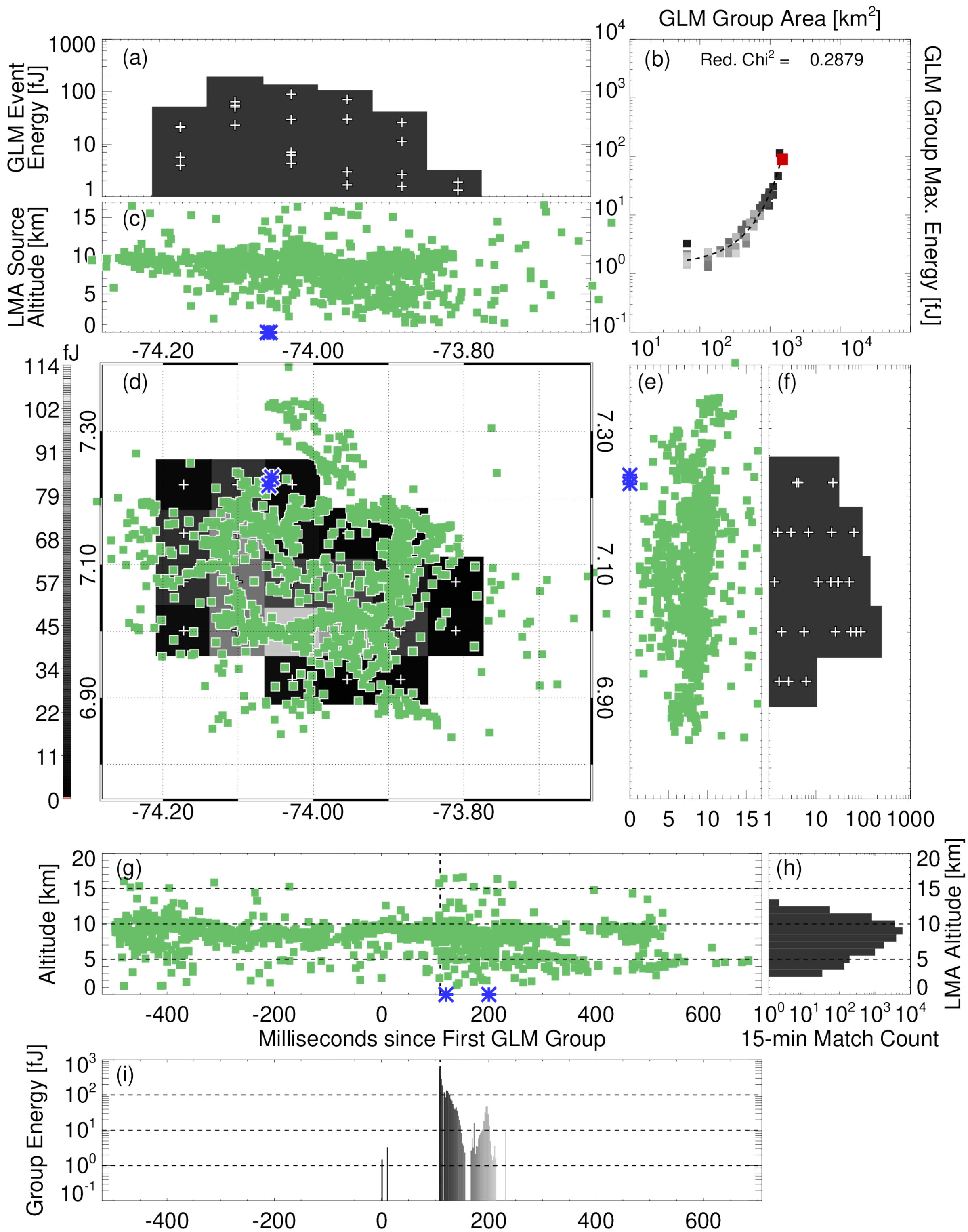
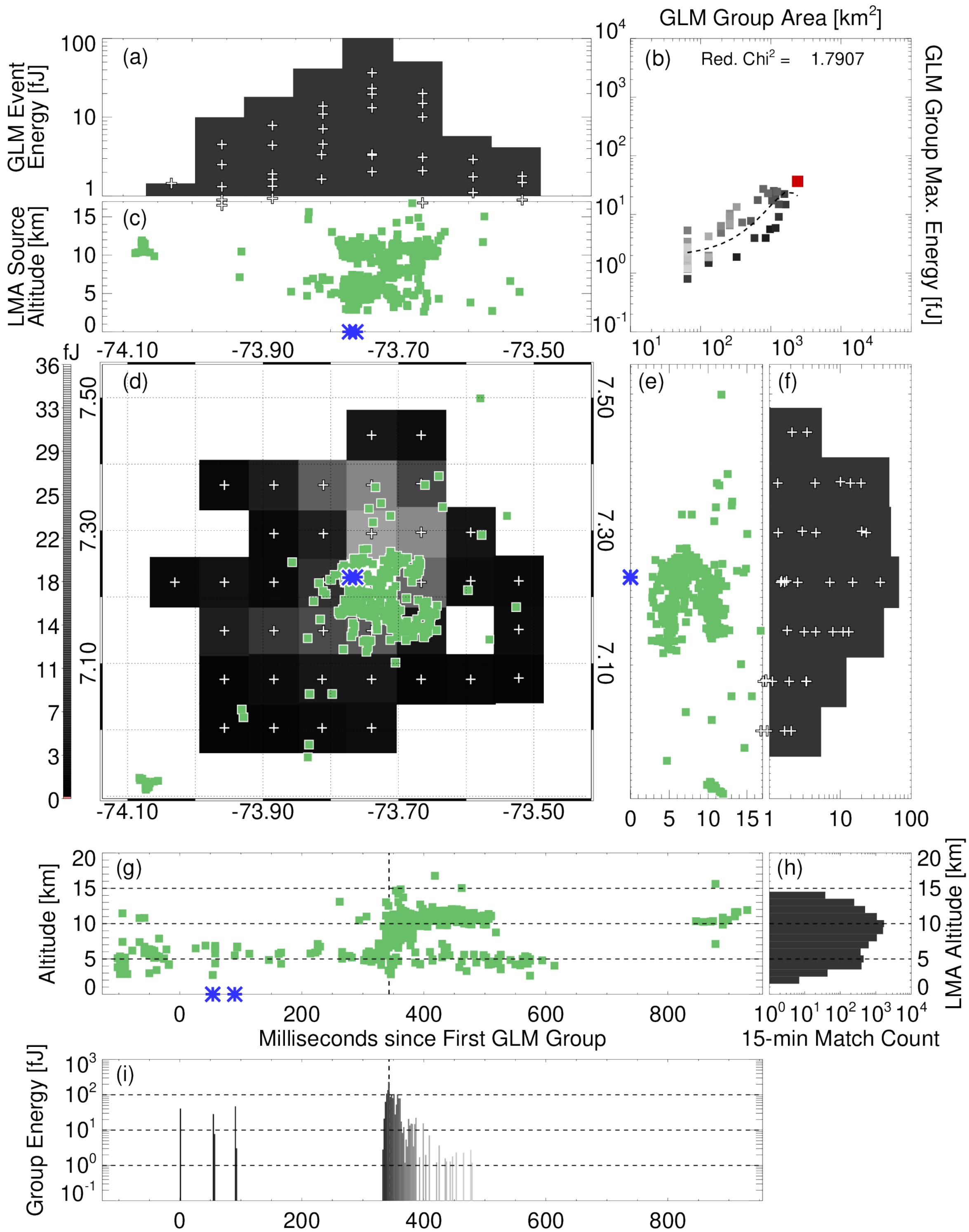


Figure 11. As with figure 10, but for a long horizontal lightning flash.



990

991 **Figure 12.** As with Figure 10, but for a flash whose groups illuminated the cloud in different

992 ways before and after development into the upper charge layer.

993