AGU ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

RESEARCH ARTICLE
10.1029/2021EA001945

The Illumination of Thunderclouds by Lightning: 4. Volumetric Thunderstorm Imagery

Michael Peterson[1] **and Douglas Mach**[2]

[1]ISR-2, Los Alamos National Laboratory, Los Alamos, NM, USA, [2]Science and Technology Institute, Universities Space Research Association, Huntsville, AL, USA

This article is a companion to Peterson et al. (2021), https://doi.org/10.1029/2021EA001943; Peterson et al. (2021), https://doi.org/10.1029/2021EA001944; and Peterson et al. (2021), https://doi.org/10.1029/2021JD035579.

Abstract Optical instruments such as the Geostationary Lightning Mapper (GLM) detect lightning based on transient changes in cloud illumination. The horizontal location of lightning is determined from the coordinates of the pixels on the imaging array illuminated during the flash. However, the vertical position of the lightning pulses (approximated by GLM "groups") below the cloud top cannot be routinely measured from a single space-based instrument. In our prior work, we have developed a machine learning algorithm that can infer optical source altitude for a given pulse based on how the optical energy is distributed across the group footprint and the local Advanced Baseline Imager Cloud-Top Height (CTH). In this fourth part of our thundercloud illumination study, we leverage these source altitudes to generate volumetric GLM imagery of a Colombia thunderstorm. We find that 3D versions of the current GLM meteorological imagery products (that describe thunderstorm kinematics) and thundercloud imagery products (that depict how the flashes appear from space) provide additional insights into lightning activity in the thunderstorm that are lost in the vertical integration used to generate the current 2D GLM gridded products. This new volumetric imaging capability provides a more comprehensive picture of where lightning occurs in the storm, how its physical characteristics vary across three-dimensional space, and how its optical emissions interact with surrounding the cloud medium.

Plain Language Summary Optical lightning imagers including NOAA's Geostationary Lightning Mapper (GLM) detect lightning by viewing the Earth from space with a specialized high-speed camera that triggers whenever one of its pixels brightens in response to lightning illuminating the surrounding clouds. Lightning can be located and have its structure mapped in two-dimensions by recording which pixels light up during the flash and projecting their angular coordinates down to the Earth. GLM data and imagery products are generated from this 2D composite view of the lightning activity in the thunderstorm. However, this is not a complete picture, as the frequency and behavior of lightning differs between vertical levels. We previously developed a method for retrieving source altitude based on how the energy from the optical pulses is spread horizontally across the cloud. In this study, we use these altitude estimates to construct 3D GLM imagery products that describe lightning across the full volume of the parent thunderstorm. This volumetric imagery provides additional insights into lightning and thunderstorms that are lost in the vertical integration employed by the current 2D GLM imagery products.

Key Points:

- Source altitude information for each Geostationary Lightning Mapper group makes it possible to generate thunderstorm imagery as volumetric grids
- Example 3D imagery is generated for a Colombia thunderstorm—including Flash Extent Density, Min Flash Area, and optical energy grids
- These 3D grids provide additional information about storm structure that are masked by the vertical integration in the current 2D grids

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

M. Peterson,
mpeterson@lanl.gov

Citation:

Peterson, M., & Mach, D. (2022). The illumination of thunderclouds by lightning: 4. Volumetric thunderstorm imagery. *Earth and Space Science*, *9*, e2021EA001945. https://doi.org/10.1029/2021EA001945

Received 5 AUG 2021
Accepted 9 DEC 2021

Author Contributions:

Conceptualization: Michael Peterson
Data curation: Michael Peterson
Formal analysis: Michael Peterson
Funding acquisition: Michael Peterson
Investigation: Michael Peterson
Methodology: Michael Peterson

1. Introduction

In Peterson et al. (2021c), we demonstrated that it is possible to estimate the source altitude for individual groups detected by a lightning imager—particularly NOAA's Geostationary Lightning Mapper (GLM: Goodman et al., 2013; Rudlosky et al., 2019)—from measurements of the spatial energy distribution provided by the instrument. Groups generated by high-altitude sources take on a substantially different appearance than groups from low-altitude sources (Peterson, 2020) because increased scattering over a thicker cloud layer broadens the spatial and temporal energy distributions (Light, Suszcynsky, & Jacobson, 2001; Light, Suszcynsky, Kirkland, & Jacobson, 2001; Koshak et al., 1994; Suszcynsky et al., 2000). As a result, low-altitude groups have broad and textured spatial distributions of radiant energy, while high-altitude groups might have almost all their optical energy concentrated in a single pixel—as shown previously with GLM groups coincident with Gigantic Jets that leave the cloud top (Boggs et al., 2019). In previous work (Peterson et al., 2021c), we examined flashes in a thunderstorm with altitudes provided by a Lightning Mapping Array (LMA: Rison et al., 1999) in central Colombia. We then used machine learning methods to find which combination of group metrics describing the amplitude, breadth, and texture of the group spatial energy distributions provided the best balance between altitude prediction accuracy and computational expense. The resulting random forest model was able to not only

Project Administration: Michael Peterson
Resources: Michael Peterson
Software: Michael Peterson
Validation: Michael Peterson
Visualization: Michael Peterson
Writing – original draft: Michael Peterson, Douglas Mach
Writing – review & editing: Michael Peterson

reproduce the GLM-matched LMA altitude distributions throughout the time history of the thunderstorm with a median absolute error of 1.33 km, but also correctly map the vertical development of individual flashes (Peterson et al., 2021c).

While adding altitude information to Level-2 GLM cluster feature data will be useful for GLM analyses of lightning physics, its primary benefit is expected to be in the generation of gridded products. GLM-derived meteorological imagery (Bruning et al., 2019) is the preferred data product for forecasters, and it is currently produced as a collection of two-dimensional grids that aggregate flashes from all vertical levels in the storm. Gridded products can be divided into four general categories: lightning rate girds—including Flash Extent Density (FED: Lojou & Cummins, 2004) and Group Extent Density (GED); flash characteristic grids—including Average Flash Area (AFA), Mean Flash Extent (MFEx), and Minimum Flash Area (MFA); cloud illumination grids—including Total Optical Energy (TOE) and the measured/modeled energies in Peterson (2019a); and thunderstorm retrieval grids—including the cloud type product in Peterson, Rudlosky, and Zhang (2020). Note that while NOAA only routinely produces a subset of the GLM grids that have been developed by the lightning community (and the list above is also not comprehensive), any of the gridded products may be constructed from the Level-2 GLM cluster feature data using the glmtools Python package (Bruning et al., 2019) or other event-based or group-based gridding techniques.

The lack of altitude information is a key limitation for the current two-dimensional gridded products because lightning imagers including GLM preferentially detect flashes with current-carrying channels close to the cloud top (Thomas et al., 2000) and these high-altitude flashes have different characteristics than flashes that are confined to low altitudes. Both the frequent small flashes that extend above 10 km altitude and the infrequent large flashes that remain near the cloud base impact the FED, AFA and TOE values. Thus, trends in these gridded products result not only from changes in flash rates and flash structure in response to thunderstorm kinematics, but also from variations in flash altitude. For example, AFA tends to increase outward from the storm core. While this can result from thunderstorm structure—anvil or stratiform flashes occurring outside of the convective core following the natural opposition between flash rate and flash area (Bruning & MacGorman, 2013)—the more common cause of this behavior is energetic low-altitude sources illuminating surrounding cloud regions that do not produce lightning and are not otherwise illuminated by dimmer flashes (Peterson et al., 2017).

Adding altitude information to GLM groups will allow us to construct volumetric gridded products to compare the behavior of lightning at different vertical levels within the thunderstorm. This fourth part of our thundercloud illumination study uses the GLM-retrieved altitudes from the Colombia thunderstorm in Peterson et al. (2021c) to demonstrate the utility of these new volumetric GLM gridded products in documenting lightning behavior and measuring thundercloud illumination from different altitudes in the storm. While we primarily consider the GLM-retrieved altitudes in this study, other altitude measurements (for example, from LMAs) are also suitable for constructing volumetric lightning imagery. Moreover, many of these products—like volumetric Flash Extent Density—can be constructed without GLM, using only measurements from LMAs or similar instruments.

2. Data and Methodology

The Colombia thunderstorm case that we have been examining since Part 1 (Peterson et al., 2021a) occurred on 01 November 2019 in the vicinity of the Colombia LMA (COLLMA: López et al., 2016; Aranguren et al., 2018). The thunderstorm and its lightning activity are described in detail in Peterson et al. (2021a). The storm is an interesting case for GLM because (a) it occurred near the satellite subpoint where parallax is small and instrument thresholds are low, making it possible to resolve flashes accurately and with an exceptional level of detail, and (b) it contained a diverse collection of lightning (including long horizontal stratiform flashes) over a large vertical depth extending above 15 km altitude due to high tropopause heights in the inner tropics.

Section 2.1 will describe GLM measurements of this Colombia thunderstorm case. Section 2.2 will briefly discuss the machine learning model developed in Peterson et al. (2021c) for retrieving source altitudes for arbitrary GLM groups. Finally, Section 2.3 will address how volumetric grids are generated from the GLM data.

2.1. GLM Cluster Feature Data

GLM is a lightning imager that is based on the design of NASA's Optical Transient Detector (OTD: Christian et al., 2003) and Lightning Imaging Sensor (LIS: Christian et al., 2000; Blakeslee et al., 2020). These optical instruments record the cloud top brightness of the scene below the satellite in a narrow spectral band around the 777.4 nm Oxygen emission line triplet at a high frame rate (nominally 500 FPS). An "event" is generated whenever the radiant energy in one of its pixels during a single 2-ms frame exceeds a specified threshold above the slowly changing background illumination at that pixel. These events are the basic building-blocks of lightning detections and describe portions of the cloud top that are momentarily illuminated during a lightning process. Individual events do not describe complete lightning pulses. Optical sources may be larger than a GLM pixel or occur at pixel corners (Zhang & Cummins, 2020) in some cases. However, even the abundant small lightning sources can illuminate cloud areas much larger than a pixel via scattering through the cloud medium as long as they are sufficiently bright (Suszcynsky et al., 2001).

Lightning imagers employ clustering algorithms to translate event data into features that describe distinct lightning processes. Beginning with the LIS/OTD missions, features have been defined to approximate individual optical pulses (termed "groups") and complete lightning flashes (Mach, 2020; Mach et al., 2007). Groups are defined as clusters of events that fill a contiguous region on the instrument's imaging array. Flashes are defined as clusters of groups that occur in close spatial and temporal proximity based on either a box-distance model (OTD) or a Weighted Euclidian Distance (WED) model (LIS, GLM).

GLM clustering is performed in real time by the lightning Cluster Filter Algorithm (LCFA: Goodman et al., 2010; Mach, 2020). This algorithm is based on the LIS clustering algorithm, but its group-to-flash clustering uses the OTD distance threshold to account for the larger GLM pixels, and proximity is determined from the constituent event data for each group rather than the group centroids. In addition to these changes to the clustering algorithm from LIS to GLM, the strict GLM latency requirements have led to hard thresholds being coded into the LCFA that terminate flashes once they reach a certain level of complexity. These thresholds of 101 events per group, 101 groups per flash, and a 3 s flash duration are not based on lightning physics and are far stricter than the previous LIS thresholds of 2000 groups per flash and 30 s in flash duration (Peterson et al., 2017). When a flash is terminated for violating one of these thresholds, it will be marked with a "degraded" quality flag in the operational GLM data and any subsequent events/groups will define a new and independent flash feature. This causes cases of long horizontal lightning megaflashes (Lyons et al., 2020; Peterson, 2021a; Peterson, Lang, et al., 2020) to be split into multiple (often tens of) "flash" features in the GLM LCFA data with all but the final emissions along each branch being designated as degraded quality.

Subsequent data products—including the meteorological imagery described in Bruning et al. (2019)—assume that the LCFA clusters the event data correctly up to the flash level. However, these long-horizontal megaflashes that are artificially split by the LCFA are prominent outside of the convective core and are largely responsible for the spatial variations in the flash characteristic grids that reveal thunderstorm organization and structure. Grids like AFA and MFA differentiate between small convective flashes and large stratiform flashes. If the large stratiform flashes are split into 101-group pieces with smaller areas than the overall flash, the contrast between convective and stratiform lightning will be reduced by even up to 1–2 orders of magnitude, as in the case of the 114,000 km^2 GLM flash discussed in Peterson (2019b).

Fortunately, the events and groups that comprise these split flashes are preserved in the operational GLM data and can be repaired to encapsulate the complete and distinct lightning flashes intended by the clustering algorithm in the LCFA. Peterson (2019b) developed software to repair the GLM flash cluster data, which has since been improved to handle cases of flashes split between separate data files. This software also adds flash metrics that better describe the development of each flash than the standard parameters in the operational LCFA data, and adds a feature level between groups and flashes to represent lightning activity between their 2 and 330 ms time scales. These "series" features (Peterson & Rudlosky, 2019) capture periods of sustained optical emission from the flash and approximate sub-flash processes such as K-waves (Winn et al., 2011) and continuing current (Bitzer, 2017).

We will use this reprocessed GLM data set (known as GLM-CIERRA: Peterson, 2021b) to construct the volumetric gridded products, as it provides a scientifically accurate picture of GLM flashes at all scales—including new

world records for flash distance (709 km) and duration (16.73 s) that have recently been recognized by the World Meteorological Organization (WMO) (Peterson, Lang, et al., 2020).

2.2. Retrieving Source Altitudes for GLM Groups

In Peterson et al. (2021c), we used Python's scikit-learn module (Pedregosa et al., 2011) to construct a random forest machine learning model for predicting the average altitude of COLLMA sources matched with GLM groups from group-level metrics describing the spatial distribution of optical energy across the group footprint. COLLMA data were provided over a 1.7° longitude (74.5°W – 72.8°W) by 1° degree latitude (6.5°N – 7.5°N) box within the LMA domain and LMA sources within a 10-km/10-ms window around a group footprint were assigned as "events" to the group of interest. If RF sources could match multiple groups in a series, the brightest matching group was chosen for assignment. The average matched source altitude was assumed to correspond to the altitude of the lightning channel that generated the optical emissions responsible for the group. However, since LMA sources are defined from a bottom-up view of the storm and GLM measurements of thundercloud illumination depend on the thickness of the cloud layer between the source and the cloud top, we use LMA altitudes that have been normalized relative to the local Advanced Baseline Imager (ABI: Schmit et al., 2017) Cloud Top Height (CTH) product (Heidinger, 2011), rather than the absolute LMA altitudes. Because LMA measurements are generally accurate to spatial scales on the order of meters (Thomas et al., 2004), the accuracy of our GLM-retrieved altitudes is critically dependent on the accuracy of the measurements of the optically relevant cloud top height.

The input feature data is comprised of a subset of 16 metrics (Table 1 in Peterson et al., 2021c) that describe the amplitude, breadth, and texture of the spatial energy distributions for GLM groups, the local estimated GLM threshold that the flash was subjected to, and the spatial extent of the process (series) that produced the group. A detailed description of the parameter selection and model training process is provided in Peterson et al. (2021c). The resulting random forest model had an overall median absolute error for multi-event groups of 1.33 km in the testing data set. Multi-event groups accounted for 66% of all groups detected by GLM from the 01 November 2019 Colombia thunderstorm. We applied the model to all of these groups and found that the GLM-retrieved altitudes resolved changes in the LMA-matched vertical altitude distribution over the duration of the thunderstorm—including responses to convective invigoration and maturation—and was able to map the three-dimensional development of individual flashes.

2.3. Constructing Volumetric GLM Gridded Products

To facilitate comparisons with the grids generated by NOAA, we base the X and Y coordinate systems of our grids on the GLM event pixels at the time of the observations and add a Z dimension consisting of 1-km layers from 0 to 20 km. Gridded products are computed on this 3D grid based on the horizontal positions of their constituent events and the altitude retrieved for the group. The GED product, for example, loops through each event in the group and increments each voxel corresponding to the point (event longitude, event latitude, group altitude) by one. FED, meanwhile, sets all these voxels equal to one for a given flash and then computes the total for all flashes. The energy products are computed in the same manner but increment the 3D voxels by their local energy. Finally, the flash characteristic grids and thunderstorm retrieval grids compute the minimum/mean/maximum attributes of the flashes that extend into each voxel.

The construction of volumetric imagery is exemplified in Figure 1 for two different flashes: a convective flash (Figure 1a) and a long horizontal flash (Figure 1b). Note that the horizontal extent of the images are scaled to fill the plot, causing the long horizontal flash to appear to be closer in size to the convective flash than it should be. The greyscale points represent the 3D positions of the GLM group centroids and are also projected onto the rear panels to show their longitude-altitude and latitude-altitude distributions. The groups are colored by time with darker shades representing groups near the beginning of the flash and lighter shades representing newer groups.

We can then visualize the 3D volume occupied by the groups in the flash (red boxes) or make contour plots representing horizontal/vertical slices through the 3D grid. In this example, we integrate GED vertically through the domain in the color contour plot below the 3D plot to produce similar imagery to NOAA's 2D GED product, and horizontally to generate longitude-altitude and latitude-altitude contour plots along the back of the figure.

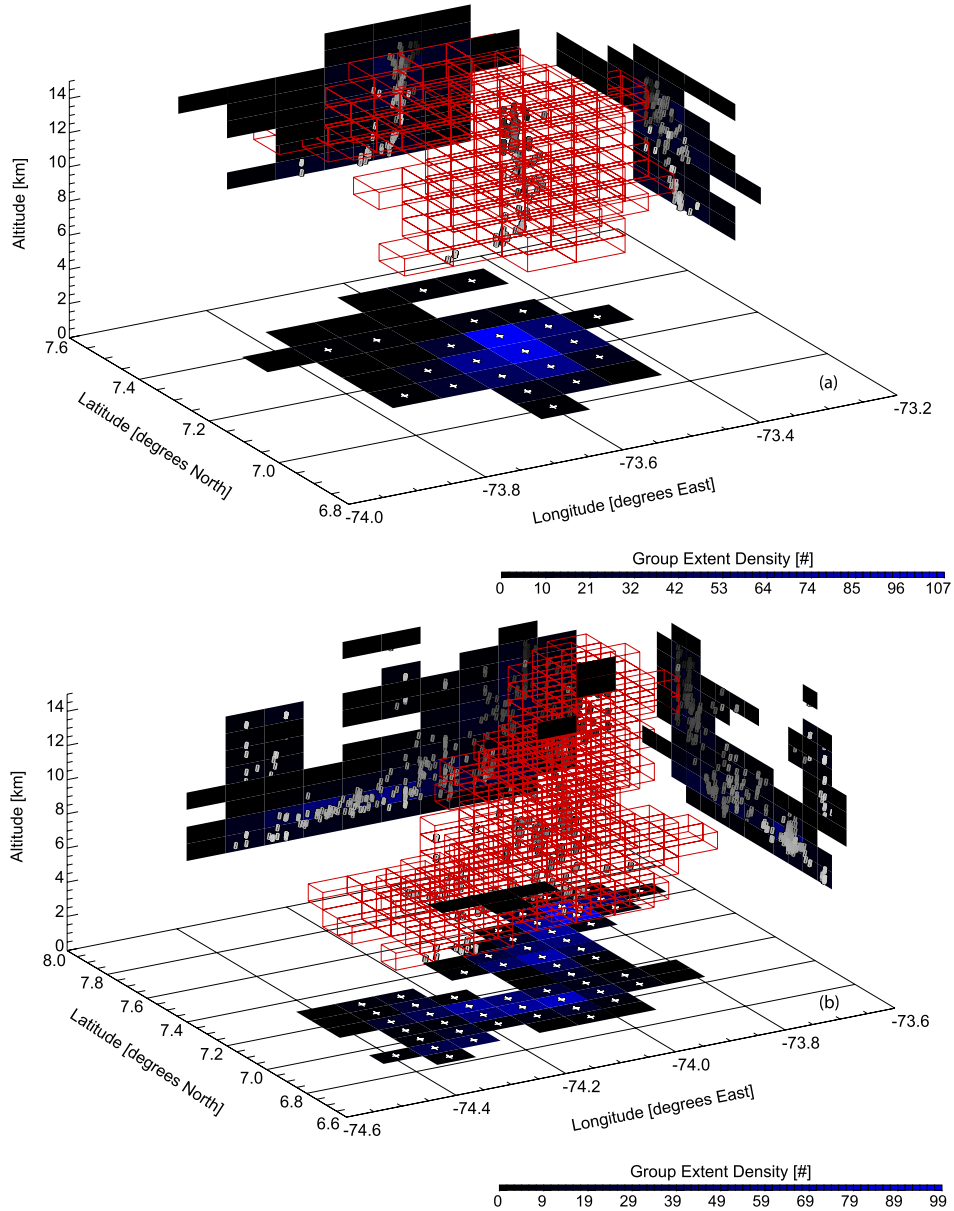

Figure 1. Volumetric Group Extent Density (GED) for an example (a) primarily vertical flash and (b) long horizontal flash. Positions of Geostationary Lightning Mapper groups (grayscale by time order from dark to light) and voxels illuminated by their constituent events are shown in 3D space. Vertical projections of GED are mapped at $z = 0$ km, while horizontal projections of GED and the group-level structure with longitude (back) and latitude (right) are plotted vertically. Note that each panel has a different latitude and longitude scale.

We can see from this imagery that both flashes begin (dark gray) at high altitudes in the cloud and descended over time (lighter gray). This is typical of lightning, which frequently develops between two or more vertical layers in the cloud. When we discuss high or low altitude flashes in the following sections, we are referring to flashes that extend to these altitudes, not, necessarily, flashes that occur entirely at a given altitude. The flash in Figure 1a also produced events (boxes) that occurred far from the group centroids (greyscale points), indicating that the flash was bright enough to illuminate distant cloud regions (particularly between 10 and 13 km altitude in this case). This poses a caveat for our methodology: we are assuming that all illumination across these larger groups occurs at the same altitude. This may not be true near the edge of the storm where the cloud depth can vary substantially across the group footprint and we can have cases of reflections off other nearby clouds. This does not appear to be

Table 1
List of Geostationary Lightning Mapper (GLM) Volumetric Thunderstorm Imagery Products Generated in This Study

Parameter name	Units	Description
Flash Extent Density	#	Total flash count per voxel
Group Extent Density	#	Total group count per voxel
Mean Groups per Flash	#	Mean number of groups per flash in each voxel
Average Flash Area	km^2	Average illuminated area of all flashes in each voxel
Minimum Flash Area	km^2	Minimum illuminated area of all flashes in each voxel
Mean Flash Extent	km	Mean group-level extent of all flashes in each voxel
Mean Flash Duration	ms	Mean duration of all flashes in each voxel
Total Optical Energy	fJ	Total optical energy from all events in each voxel
Mean Flash Energy	fJ	Average optical energy in a given voxel from all flashes in that voxel
Mean Group Energy	fJ	Average optical energy in a given voxel from all groups in that voxel
Measured – Modeled Energy	fJ	Energy difference between GLM-measured TOE in a given voxel and the expected energy from modeling the radial energy distribution for each group

an issue with the long horizontal stratiform flash in Figure 1b, whose bright groups better reflect the group-level structure of the flash.

All of our storm-level volumetric imagery is generated by aggregating the GLM data over all flashes during 15-min GLM-CIERRA (Peterson, 2021b) data files aligned to the UTC hour. A 15-min aggregation is chosen because we find that it provides a desirable level of spatial filling across the grid—even in low flash rate thunderstorms. Other aggregation periods including 1 min or 5 min would also be useful. In total, we generate 12 volumetric gridded products, which are listed in Table 1. While not all products will be discussed for brevity, example imagery for each product will be included as Supporting Information S1.

3. Results

The following sections will analyze the volumetric GLM gridded products generated for the Colombia thunderstorm using our GLM source altitude estimates. For a more thorough discussion of the evolution of the Colombia case, refer back to Part 1 (Peterson et al., 2021a). For an examination of how the instrument threshold affects the GLM data, refer to Part 2 (Peterson et al., 2021b). Finally, for a detailed description of the altitude model and its performance, refer back to Part 3 (Peterson et al., 2021c). Our analyses here will start with a general discussion of time-altitude trends in the gridded products over the duration of the thunderstorm in Section 3.1. Then, we will examine snapshots from various points in the storm that illustrate the origins of these trends and demonstrate the value of having vertical information available when interpreting GLM imagery. This analysis will be divided into two parts: Section 3.2 will discuss the meteorological imagery (encompassing the lightning rates and flash characteristics categories from before), while Section 3.3 will discuss the thundercloud imagery grids (encompassing the GLM energy products) that approximate what a high-altitude aircraft or space-based sensor would capture on film while recording the storm from above (i.e., Figure 1 in Peterson, 2019a).

3.1. Time-Altitude Trends in GLM Gridded Products

As we discussed in Part 2 (Peterson et al., 2021b), trending gridded products over time can be problematic if the instrument threshold changes substantially over the analysis period. In the Colombia thunderstorm case, however, the approximate threshold remained below 1 fJ over most of the storm duration (see Figure 2 in Peterson et al., 2021b). Dramatic increases in threshold were only noted before 02:30 UTC when lightning activity was first entering the LMA data domain and after 12:30 UTC as the storm was dissipating—and these increases were only up to a maximum of 2 fJ. Thus, threshold changes are not expected to be a significant source of bias in this case.

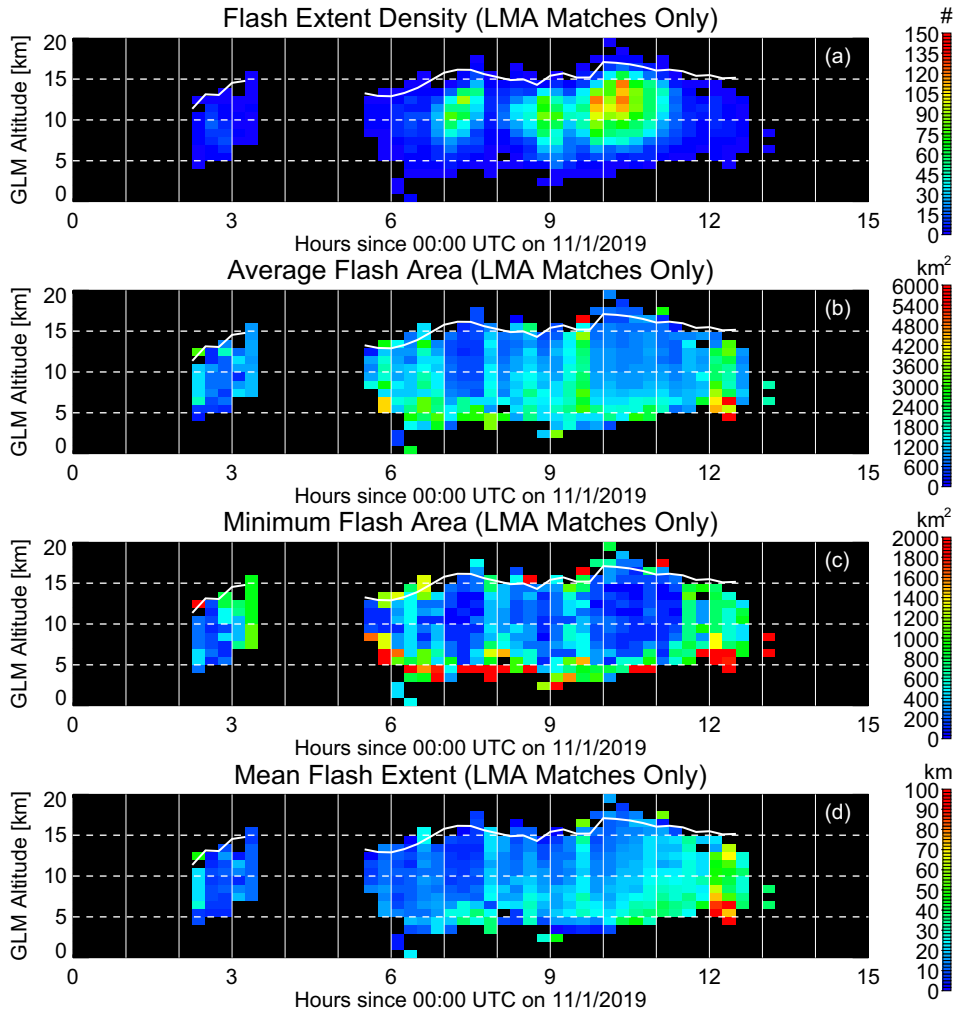

Figure 2. Timeseries of (a) Flash Extent Density, (b) Average Flash Area, (c) Minimum Flash Area, and (d) Mean Flash Extent from Lightning Mapping Array-matched Geostationary Lightning Mapper (GLM) flashes in each 1-km altitude layer over the duration of the Colombia thunderstorm. The white lines in each panel signify the maximum Advanced Baseline Imager Cloud-Top Height coincident with the GLM groups.

Time-altitude grids are shown in Figure 2 for FED, AFA, MFA, and MFEx and Figure 3 for TOE, Mean Flash Energy (MFEn), Mean Group Energy (MGE), and Mean Groups per Flash (MGPF). The white lines in each panel signify the maximum ABI CTH coincident with GLM groups. Note that these timeseries are generated using only flashes that are matched to LMA sources to enable comparisons with the altitude timeseries figures from Part 1 (Peterson et al., 2021a). As in Part 1, flashes that occur outside of the LMA data domain (or that straddle its edges) are not considered here. We also provide the same plots constructed from the measured LMA source altitudes for each group rather than the GLM predictions as Figures S1 and S2 in Supporting Information S1 to demonstrate that similar products can be derived from other types of lightning altitude data.

Two distinct periods of lightning activity on 01 November 2019 can be noted in Figures 2 and 3: a short earlier period around 03:00 UTC where thunderstorm activity grazed the southern boundary of the LMA data domain (Figure 1f in Peterson et al., 2021a), and a longer period from 05:30 UTC until 13:00 UTC that contained most of the lightning activity mapped by both GLM and the LMA. This later period included three episodes of convective invigoration (around 07:00 UTC, 09:00 UTC, and 10:00 UTC) where peaks in the GLM group rates and LMA source rates could be noted in Figure 2 of Peterson et al., 2021a. The volumetric FED plot in Figure 2a here not only has local maxima corresponding to the peaks in these periods, but also increases in flash activity

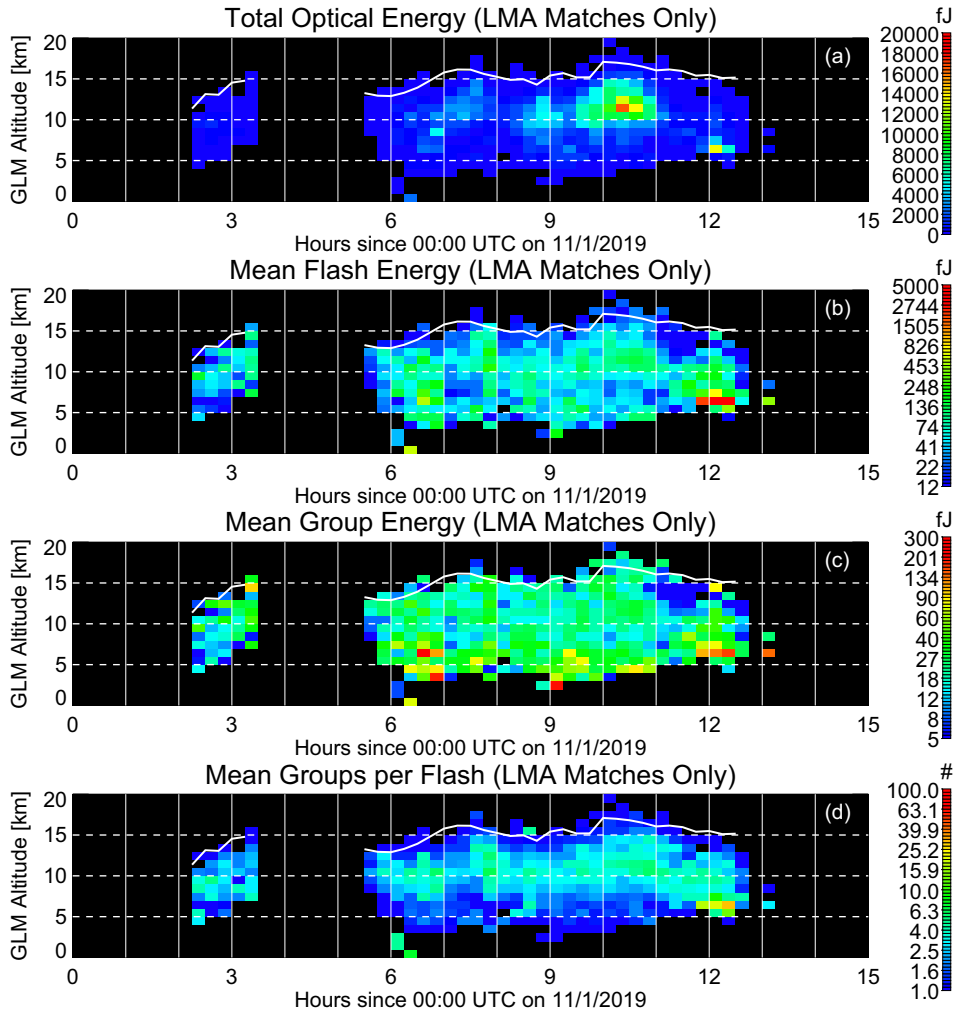

Figure 3. Timeseries of (a) Total Optical Energy, (b) Mean Flash Energy, (c) Mean Group Energy, and (d) Mean Groups per Flash from Lightning Mapping Array-matched Geostationary Lightning Mapper (GLM) flashes in each 1-km altitude layer over the duration of the Colombia thunderstorm. The white lines in each panel signify the maximum Advanced Baseline Imager Cloud-Top Height coincident with the GLM groups.

specifically at high altitudes in response to the strengthening updraft (particularly during the 08:00 UTC and 10:00 UTC peaks).

These periods of intensification are accompanied by a reduction in average flash area over most of the vertical column in Figure 2b. The key exception to this is with flashes that are located near the cloud base, whose AFA values generally remain above 1,500 km^2 throughout the storm duration. The MFA plot in Figure 2c amplifies this contrast between lightning at the cloud base and lightning higher in the thunderstorm. The large values come from both convective flashes that illuminate large areas of nearby clouds and long horizontal stratiform flashes, as illustrated in Section 3.2. While these two types of large flashes can be difficult to separate in the standard grids produced by NOAA, the Mean Flash Extent product (Figure 2d) captures long horizontal flashes while not being particularly sensitive to flashes that happen to illuminate a large area of nearby cloud. This product, when viewed in conjunction with the flash-area products, shows that most of the large-area flashes during the earlier periods (05:00–10:00 UTC) illuminated cloud regions far beyond the extent of the flash structure mapped by its groups, while the later large flashes were cases of long horizontal lightning.

The TOE grid in Figure 3a resembles the FED grid from Figure 2a but is more concentrated in the 10–12 km layer than FED and has an unmatched low-altitude peak in the 12:00 UTC hr. These differences result from variations in the amount of energy provided by each flash (Figure 3b) and group (Figure 3c) as well as the number of

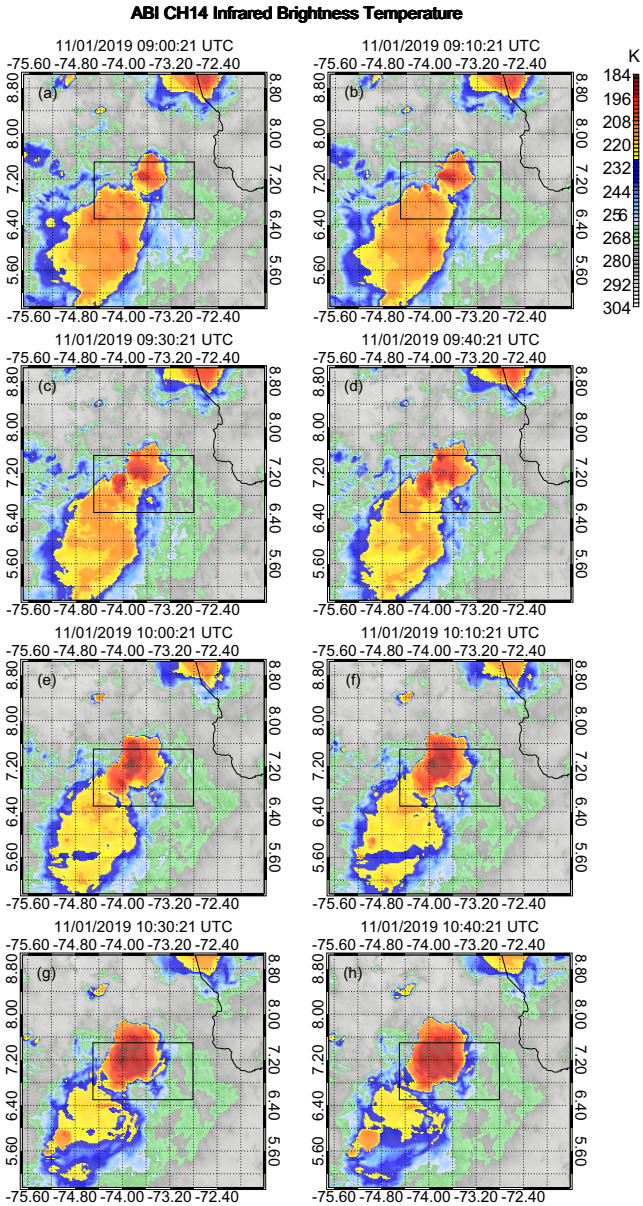

Figure 4. Maps of Advanced Baseline Imager (ABI) Channel 14 (11.2 μm) infrared brightness temperature across central Colombia during the nearest ABI scans to the 15-min GLM-CIERRA data files at (a) 09:00 UTC, (b) 09:15 UTC, (c) 09:30 UTC, (d) 09:45 UTC, (e) 10:00 UTC, (f) 10:15 UTC, (g) 10:30 UTC, and (h) 10:45 UTC. The ABI scan start times are listed in the titles for each panel. The boxes region represents the COLLMA data domain considered in this study.

optical pulses that are resolved from each flash (Figure 3d) at each altitude. Before these later periods that start around 11:00 UTC, flashes in the vertical layers around the LMA source peak (~10 km altitude from Figure 2 in Peterson et al., 2021a) consist of more groups than the flashes that extend to near the cloud base or the cloud top (Figure 3d). The low-altitude groups can be more energetic than groups above 7 km altitude (Figure 3c), but the few bright groups that can be resolved from such low altitudes are not sufficient to overcome the total energy contributed by the frequent dim groups that are detected higher up. Thus, the Mean Flash Energy (Figure 3b) near the cloud base is not remarkable in this plot, and the TOE values from these lower layers are small compared to layers around the 10-km LMA peak.

The exception to these generally low MFE values at low altitudes is when the storm begins to dissipate and long horizontal stratiform flashes become prominent. Flash altitudes decrease and flash rates fall after the 10:30 UTC peak, while the gridded products begin to change to reflect the organization and structure of the dissipating storm. These changes are most prominent around 12:00 UTC where low-altitude peaks form in the AFA, MFA, and Mean Flash Extent grids in Figure 2. The larger flashes responsible for these peaks have more groups per flash (Figure 3d) and produce more energy per group (Figure 3c) or flash (Figure 3b), leading to the second-highest TOE peak in Figure 3a.

This behavior at low altitudes in the storm would not be as notable in the current 2D gridded products because they integrate the lightning activity from all vertical levels. As GLM has a detection advantage for sources closer to the cloud top, the 2D grids for this storm would be biased towards flash activity near the 10-km LMA source maximum where GLM tends to resolve smaller/fainter pulses that have not been spatially broadened to the same extent as low-altitude optical pulses. Constructing volumetric grids allows us to resolve nuances in the meteorological and thundercloud imagery that are not apparent in the standard 2D gridded products generated by NOAA.

3.2. Meteorological Imagery From Volumetric GLM Grids

The time-altitude grids in Section 2.1 provide an overview of lightning activity within the LMA data domain that still might integrate the behavior of lightning in multiple distinct thunderstorm features. In this section, we will examine spatial variations in lightning rates and flash characteristics between 09:30 UTC and 10:30 UTC, surrounding the third and strongest period of intensification. To gain a broader perspective on the lightning activity in the region, we will no longer exclusively consider GLM groups that are matched to LMA sources, or even groups that occur within the LMA data domain. Instead, all GLM data from the Colombia thunderstorm will be analyzed within 2° latitude or longitude from the center of the LMA data domain.

ABI Channel 14 (11.2 μm) infrared brightness temperatures for the 09:00 UTC and 10:00 UTC hr are shown across central Colombia in Figure 4. The box outlined in the center of each panel corresponds to the LMA data domain used in the analyses in Figures 1 and 2. This period describes the merger of two cold cloud features (contiguous regions of ABI CH4 IR $T_b < 215$ K). The larger feature in the southwest corner of Figure 4a depicts a large Mesoscale Convective System (MCS) with an expansive stratiform region extending to the southwest. This feature was first noted around 02:00 UTC and grew over the next four hours before encroaching upon the LMA data domain by 06:00 UTC. The second northern feature depicts a convective cell that developed in isolation around 06:00 UTC and grew into a larger storm over the next three hours. As the features began to merge by 09:00 UTC (Figure 4a), warming could already be noted in the ABI CH14 IR T_b values within the southern feature, indicating that the feature was

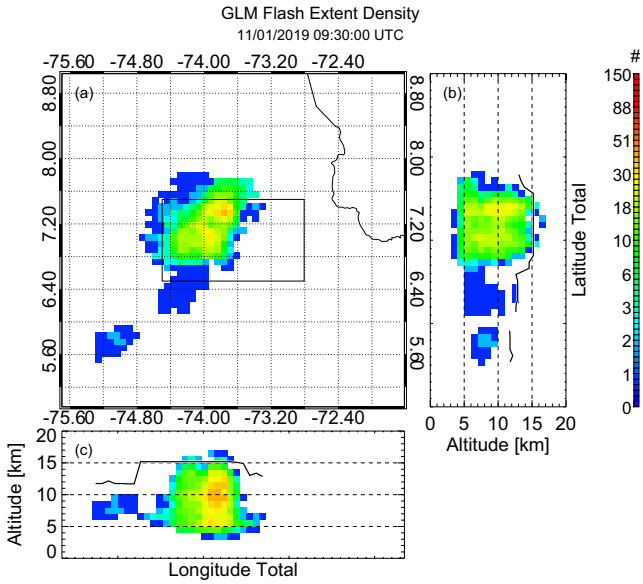

Figure 5. Volumetric Flash Extent Density valid from 09:30 to 09:45 UTC expressed (a) as a vertical integration, and horizontal integrations resulting in (b) a latitude-altitude distribution, and (c) a longitude-altitude distribution. The boxed region in (a) represents the COLLMA data domain, while the solid lines in (b) and (c) show the maximum Advanced Baseline Imager Cloud-Top Height coincident with Geostationary Lightning Mapper groups at each latitude or longitude.

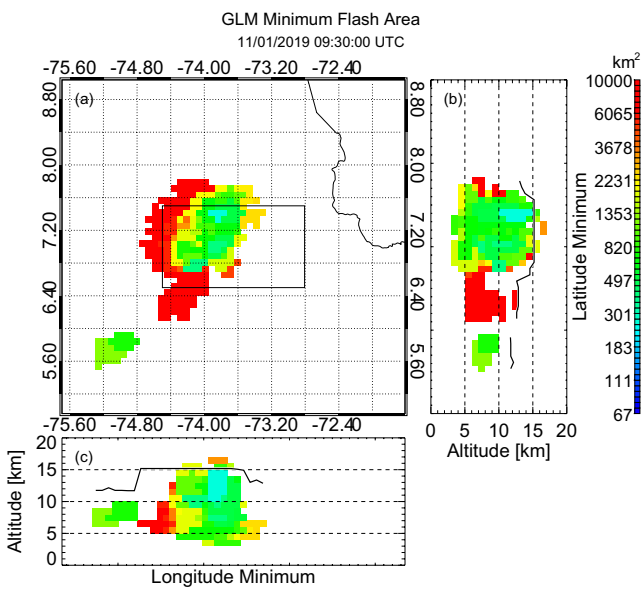

Figure 6. Volumetric Minimum Flash Area valid from 09:30 to 09:45 UTC, plotted as in Figure 5.

beginning to dissipate. Meanwhile, the newer northern feature still had cloud top temperatures <196 K. The merger of these features over the two-hour period shown in Figure 4 was accompanied by intensification, resulting in decreasing infrared brightness temperatures over a large fraction of the LMA data domain.

To examine how the lightning responded to these developments, Figures 5–11 summarize the 3D gridded products by plotting horizontal and vertical integrations through the domain mapped in Figure 4. Figure 5 shows GLM FED at 09:30 UTC—at the beginning of the period in Figure 2a where FED values began to rise throughout the column and specifically near the cloud top. The top-down integration in Figure 5a is consistent with the current 2D grids (though, not downscaled to the finer ABI fixed grid), while the horizontal integrations in the panels next to the map show the latitude-altitude (Figure 5b) and longitude-altitude (Figure 5c) distributions of FED across the mapped region. As in Figure 4, the LMA data domain is indicated with a box in Figure 5a. Additionally, the maximum ABI CTH values coincident with GLM groups are indicated by solid line overlays in Figures 5b and 5c.

The larger regions of ABI infrared brightness temperatures <215 K contain multiple convective-scale features in the FED plots in Figure 5 that are aligned with the coldest cloud-tops in Figure 4. At 09:30 UTC, two such features can be noted: a northern FED maximum within the northern cold cloud feature that includes the overall FED maximum for the storm and extends down to 5 km altitude (Figure 5b), and a southern convective feature in the southern cold cloud feature that corresponds to the dissipating MCS. Later, a convective feature further north within the northern cold cloud feature will form, which will we refer to as the "northernmost" convective feature to distinguish it from the northern convective feature shown here. Moreover, we use the term "convective core" to colloquially describe the region surrounding both (or all) convective features where at least 3 flashes occurred in the 15-min window shown (i.e., 1 flash every 5 min).

The convective core of the northern thunderstorm feature is surrounded by lower FED values (1–2 flashes in the 15 min window). These boundary regions are primarily populated by flashes at lower altitudes between 5 and 10 km. Figure 6 plots the 3D MFA grid to better explain what these boundary flashes represent. The small flashes around the FED maxima (mostly in the northern feature) and larger flashes along the western flank of the storm are typical for the MFA product, but—as mentioned previously—could result from either bright optical pulses interacting with nearby clouds or long horizontal stratiform flashes. Gradients in the ABI infrared imagery in Figure 4 suggests that the large MFA values along the northwest flank of the storm could result from edge illumination (as CH14 IR T_bs are much warmer than the in nearby convective core), while we would expect to find long horizontal flashes in the stratiform region extending from the southwestern flank of the convective core of the mature southern feature.

One method to verify this interpretation of the GLM data is to examine additional aspects of flash behavior in these regions. The MFEx and the MGPF gridded products differentiate between single radiant groups illuminating a cloud region and propagating flashes extending through that cloud region. Both products are included as SI, but we will show MGPF in Figure 7. The MGPF values in Figure 7a along the northwestern flank of the storm are low (1–2 along the periphery, increasing towards the convective core), confirming that these regions were illuminated occasionally by only the largest/brightest groups produced during large flashes. The MGPF values in the southwestern

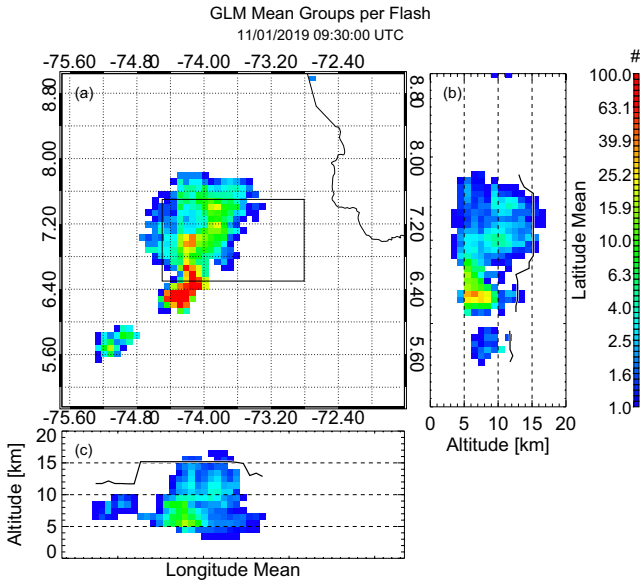

Figure 7. Volumetric Mean Groups per Flash valid from 09:30 to 09:45 UTC, plotted as in Figure 5.

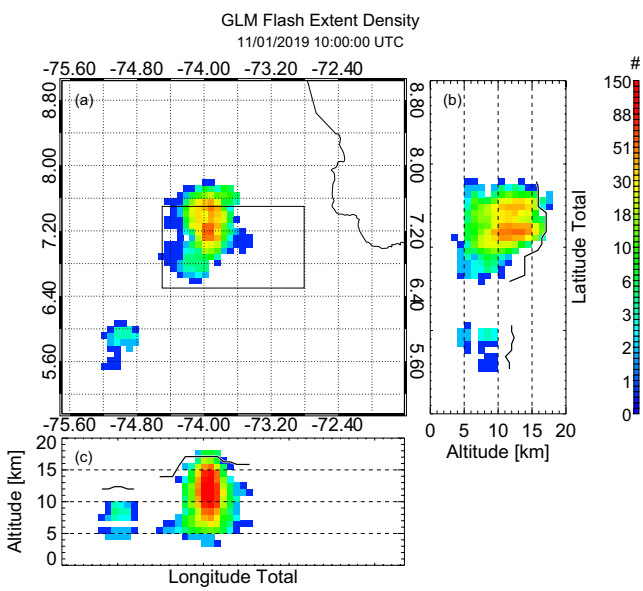

Figure 8. Volumetric Flash Extent Density valid from 09:45 to 10:00 UTC, plotted as in Figure 5.

flank of the primary contiguous thunderstorm feature, meanwhile, are exceptionally high—indicating that the average flash in this region produced 20+ groups whose events extended into each of these pixels. The latitude-altitude distribution in Figure 7b shows that these large MGPF values occur exclusively at low altitudes in the cloud (<10 km) extending southward from the southern convective feature.

Between 09:30 UTC and 10:00 UTC (Figures 4c–4e), the original southern convective feature continued to weaken while a new cell developed along the northern boundary of the LMA data domain. The original northern feature also intensified around 10:00 UTC, leading to lower ABI infrared brightness temperatures corresponding to higher cloud top heights. Figure 8 shows the volumetric FED grid at 10:00 UTC. The ABI CTH traces in Figures 8b and 8c are ∼2 km higher than in Figure 5 with significant lightning activity (FED > 30 flashes over the 15-min period) extending beyond 15 km altitude. The MFA plot in Figure 9 shows that this increase in convective lightning is accompanied by a substantial decrease in MFA at all vertical levels within the storm core due to the presence of flashes that illuminate the equivalent area of just 1–3 GLM pixels.

This suggests that the increase in column-integrated FED in Figure 5a is driven by frequent small convective flashes at high levels in the storm in response to the strengthening updraft that likewise caused flash heights to increase between 09:30 UTC and 10:00 UTC. However, the 3D grids also illustrate that despite the single smallest flashes decreasing in area, the AFA at low altitudes did not drastically change, and there were still some cases of large flashes along the western flank of the storm in Figure 9a. This increase in high-altitude FED was also accompanied by a decrease in low-altitude FED, though lightning continued to be detected at low altitudes over this period. Thus, these changes in the vertical source altitude distribution describe an upward migration in flash activity rather than a cessation of lightning activity at low levels. This can also be noted in the FED timeseries in Figure 2 (for GLM altitudes) and Figure S1 in Supporting Information S1 (for LMA altitudes). While not as prominent as in Figures 8 and 9 due to integrating all lightning activity over the LMA data domain into a single profile, the upward shift in source altitude is evident in these timeseries and reaches its maximum vertical displacement in the 10:15 UTC bin.

The storm began to dissipate after 10:15 UTC. By 10:30 UTC, all convective features within the storm had started to weaken. The FED values in Figure 10 are notably lower than in the previous plots, though the MFA values in Figure 11 remained small following the convective burst. MFEx in Figure 2d was beginning to increase and, within the next hour, AFA (Figure 2b) and MFA (Figure 2c) would follow suit as the vertical peak in MGPF (Figure 3d) fell below 10 km altitude—marking the transition from a convective flash dominance to prominent non-convective lightning activity.

3.3. Thundercloud Imagery From Volumetric GLM Grids

The GLM energy grids listed in Table 1 show what GLM flashes look like embedded in the clouds. TOE, MFEn, and MGE all represent what an observer in space would record with a camera—with TOE representing a typical long (15-min) exposure, MFEn approximating the average energy on flash time scales (<1 s) and MGE reporting the average energy on group time scales (2 ms). This conceptual approximation is not perfect for the latter two quantities, as they are computed by taking the TOE grid and dividing by either FED or GED. Thus, their actual definitions are the average amount of energy in a given pixel or voxel provided by each flash or group that illuminated that pixel or voxel. This reduces their dependence on lightning frequency to highlight variations in thundercloud illumination across the scene.

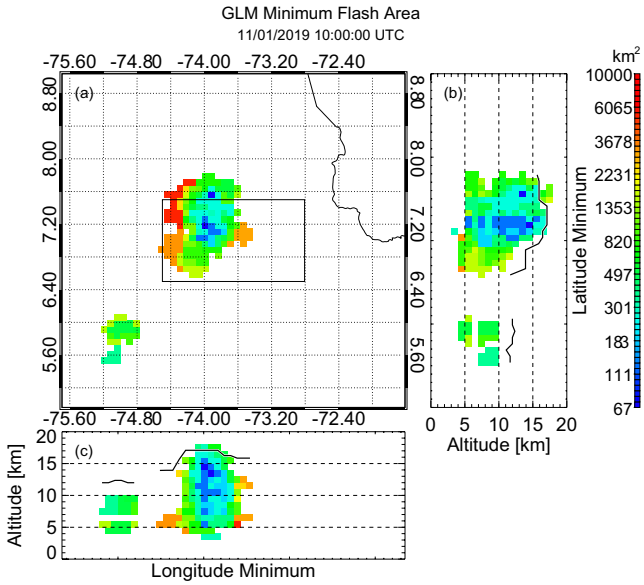

Figure 9. Volumetric Minimum Flash Area valid from 09:45 to 10:00 UTC, plotted as in Figure 5.

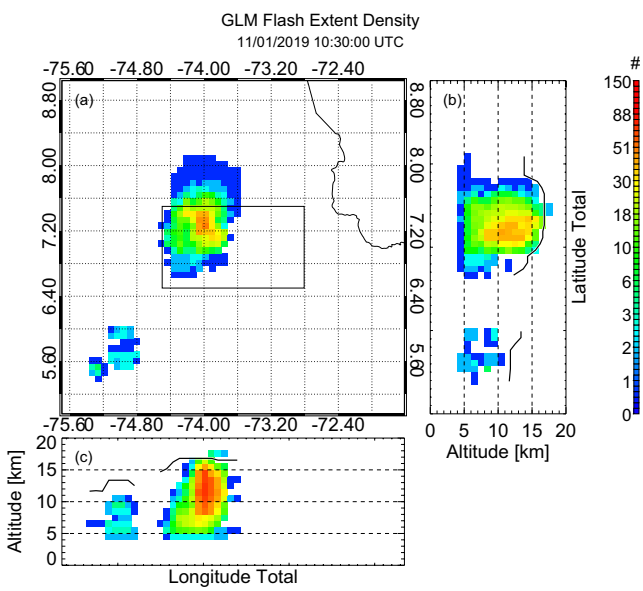

Figure 10. Volumetric Flash Extent Density valid from 10:30 to 10:45 UTC, plotted as in Figure 5.

Thundercloud imagery can be more difficult to interpret than the meteorological imagery in the previous section because the amount of energy that is measured from orbit depends not only on the frequency and characteristics of lightning, but also the optical properties of the surrounding clouds. Where the lightning is located relative to the cloud structure (particularly with respect to altitude and proximity to cloud edges) will affect how the illuminated clouds appear from space. Therefore, features will be evident in the thundercloud imagery that are not apparent in the metrological imagery. The unique perspective provided by these products can be useful for tracking thunderstorm development and explaining trends in the meteorological imagery.

A key example is the case of radiance anomalies around convective clouds. We previously inferred that bright groups illuminating neighboring cloud regions resulted in the large flashes along the northern flank of the convective core. However, with MGE and MME, we can directly see that there is increased illumination in this region. Figure 12 shows MGE imagery for the Colombia thunderstorm at 09:30 UTC. A greyscale color palette is used in this product to emphasize contrasts in the imagery. The vertical integration in Figure 12a shows a considerable amount of texture across the scene surrounding the convective features from Figure 5a, including a low-energy region adjacent to the northern FED hotspot and two particularly bright regions located on the northern and southern flanks of the storm core. The volumetric imagery (Figures 12b and 12c) shows that these bright regions coincide with the large low-altitude flashes shown in Figures 6 and 7.

The northern MGE peak is part of a broader region of enhanced MGE that extends across the northwestern flank of the storm and wraps around to the western flank in Figure 12a (red arrows). It also extends from 5 km altitude up 15 km altitude—thus, multiple pulses from all altitudes at the edge of the cloud are contributing to the northern MGE maximum in Figure 12a. In these cases, optical emissions from sources at the edge of the cloud can travel along a relatively clear path to the imager with reduced attenuation compared to transmitting through the full vertical cloud volume, and this leads to biases towards large/energetic flashes in the meteorological imagery. By comparison, the southern MGE maximum is part of a linear feature of enhanced MGE that extends southward from the dissipating MCS. This peak is located in a region where the bright MGE values are in a vertically thin layer that extends from 5 to 7 km altitude - far below the local cloud top. Optical emissions depend on both the electrical current traversing the channel and the channel length. Flashes with long horizontal channels are thus able to generate more energy for a given current (particularly during strokes or K-changes) than small convective flashes, and this could result in the increased MGE values here.

The dark region near the northern FED maximum requires a more detailed analysis to explain. This four-pixel MGE feature extends throughout the vertical column, leading to the decreased MGE values south of the northern bright spot in Figure 12b. Our recent work with radiance anomalies (Peterson, 2021d) offers two possibilities on how radiant events can be adjacent to some of the dimmest events: either there is a poorly transmissive cloud present, or there is no cloud at all to scatter photons towards the satellite. The low altitudes of these groups and cold ABI IR T_b values in Figure 4c rule out the latter explanation. Thus, it is likely a poorly transmissive cloud that is preventing certain pixels from being illuminated in these GLM groups.

Figure 13 shows a representative large (>1,000 km^2) group from this dark region. These large groups account for the top few percent of all groups from the Colombia thunderstorm (2% overall, but sometimes exceeding 6%) because most of the emissions detected by lightning imagers including GLM are faint in-cloud pulses. Only a few groups in the average flash are particularly

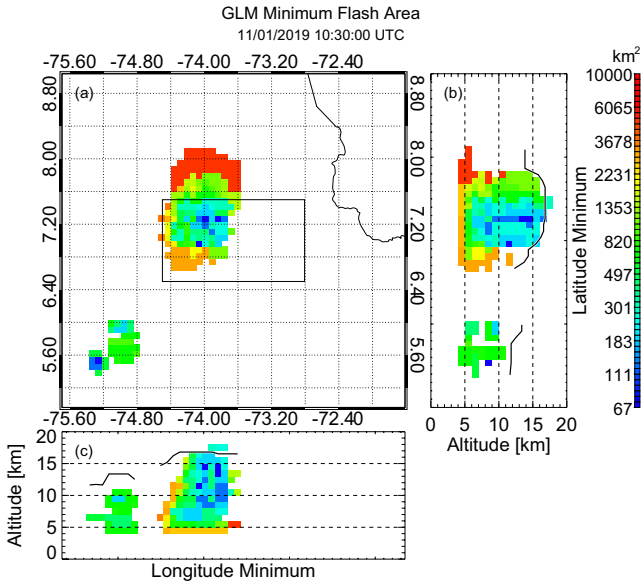

Figure 11. Volumetric Minimum Flash Area valid from 10:30 to 10:45 UTC, plotted as in Figure 5.

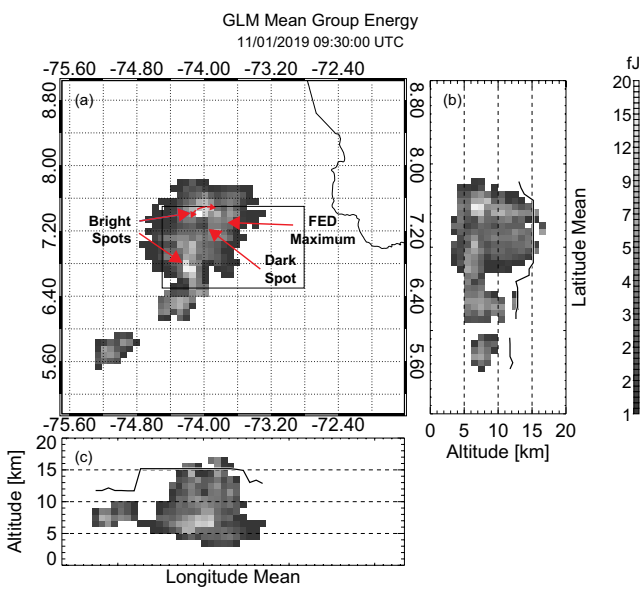

Figure 12. Volumetric Mean Group Energy valid from 09:30 to 09:45 UTC, plotted as in Figure 5.

energetic (i.e., Figure 4a in Peterson & Rudlosky, 2019) and illuminate a large area. We perform the steps necessary to identify radiance anomalies (as in Peterson, 2020) and construct the Measured – Modeled Energy product from Table 1. This approach attempts to normalize the spatial energy measurements from individual GLM groups to compensate for the natural radiance fall-off with distance from the optical source, thereby amplifying local contrasts in the radiance patterns from inhomogeneities in the cloud medium. The spatial energy distribution for the group is shown in Figure 13a. The group consisted of a single very bright event (356 fJ) and a footprint comprised of 96 other dimmer events that extended to the southwest, northwest, and northeast of the brightest event. Note that this footprint did not extend due south, and the GLM pixel adjacent to the 356 fJ event did not trigger. Instead, the footprint seems to wrap around the western edge of a feature in the cloud medium that is blocking the light in this particularly bright pulse from reaching orbit. This radiance anomaly was only a few events away from forming a hole in the group, as we saw previously with LIS, and other groups in the area (not shown for brevity) had similar difficulties penetrating this cloud region.

We plot the radial energy profile for this group in Figure 13b and then fit the data points to a 4-term Gaussian model to describe the radiance fall-off with distance from the brightest event in the group (dashed line). Unlike Peterson (2019a), we now clamp the radiance profile to the origin at the brightest event and the most distant event radius by giving these points additional weight in the numerical model (equal to the total number of events in the group). We construct and record these Gaussian models for every GLM group that exceeds 1,000 km^2 in illuminated area along with optical pulse metrics such as the Half Width of Half Maximum (dotted line in Figure 13b), but discard cases where the model fit is either unrealistic (constant or increasing energy with distance) or too far from the energies of the clamp points (>5% in the normalized energy from Figure 13b).

As these issues are not a concern for the group in question, we can use its Gaussian model to compute an idealized spatial energy distribution in Figure 13c. Unlike the measured energy distribution in Figure 13a, there are no radiance anomalies present and GLM energy depends only on the radial distance from the brightest measured event. Note that there is no maximum distance or minimum energy in the modeled energy distribution. To mitigate bias from extrapolation below the local GLM threshold, we ignore all pixels that are either further from the origin than the most distant observed event or that are less energetic than the dimmest event in the group.

Finally, the measured energy distribution in Figure 13a is compared with the modeled energy distribution in Figure 13c to generate the imagery in Figure 13d. This energy comparison forms the basis for the thundercloud imagery in Peterson (2019a) and the poorly transmissive cloud identification algorithm in Peterson (2020). However, we are presenting it here as a difference between measured and modeled energy in units of fJ rather than as a radiance ratio. The Measured/Modeled Energy product gives each group an equal weight, regardless of its brightness, and highlights texture in the group spatial energy distribution at larger distances from the origin. The Measured – Modeled Energy shown in Figure 13d emphasizes contrasts near the origin and weights each group according to its maximum event energy. There are advantages to each approach, but we chose to introduce the energy difference product here because it reduces the impact of elongated optical sources (which produce higher-than-expected energies at large distances) while highlighting cases where optical emissions can escape the cloud along relatively clear paths. For the group in question, this product clearly shows a measured

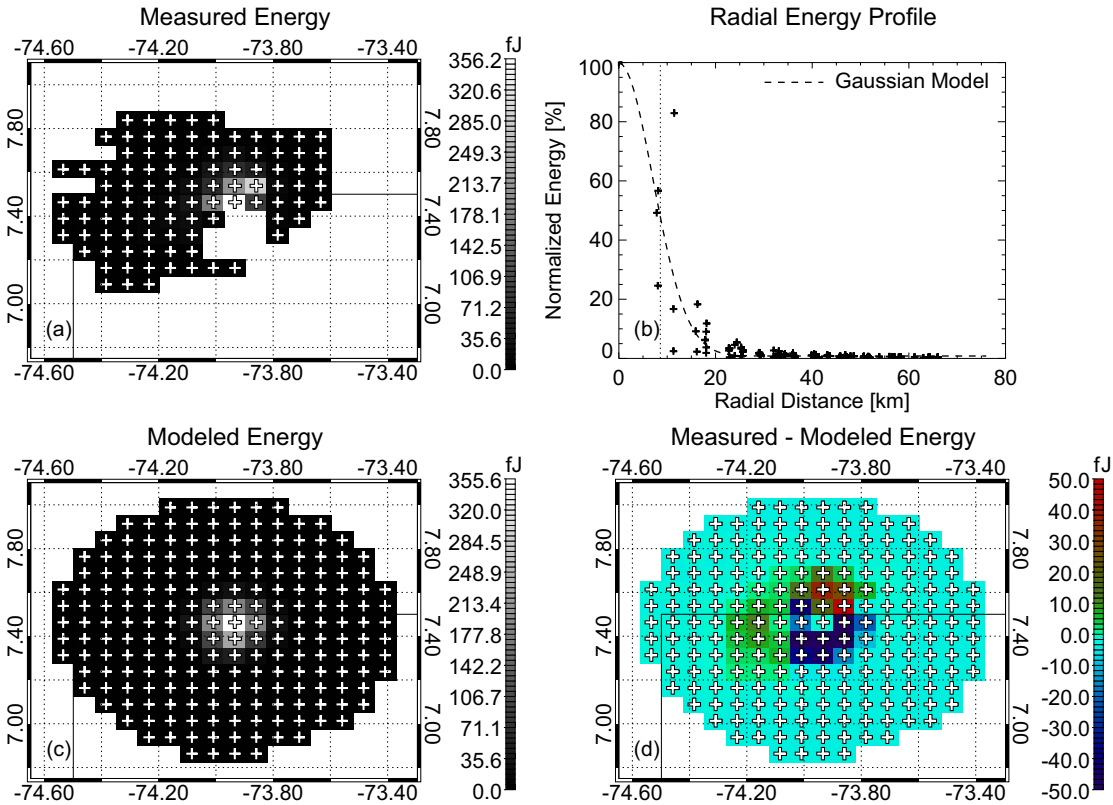

Figure 13. Spatial energy analysis for an example large Geostationary Lightning Mapper (GLM) group located along the northern flank of the thunderstorm core. (a) Measured event energy across the group. (b) Radial energy profile relative to the brightest event in the group (plus symbols) and Gaussian model fit (dashed). The Half Width of Half Maximum (HWHM) distance is indicated with a vertical dotted line. (c) Idealized spatial energy distribution from the Gaussian model. (d) Energy difference between the GLM measurements and the Gaussian model.

energy deficit to the south and east of the brightest event in the group and a measured energy surplus in the northern and western regions where the group footprint wraps around the poorly transmissive cloud.

The Measured – Modeled Energy (MME) grids (i.e., Figure 13d) are accumulated from all GLM groups >1,000 km^2 and the resulting volumetric MME imagery for 09:30 UTC is mapped in Figure 14 along with ABI CH14 infrared brightness temperatures (Figure 14a) as an overall vertical integration (Figure 14b) and at 6 different vertical levels. As before, the LMA data domain is indicated with a solid outline at the center of each panel. A prominent feature of this new imagery is that there is a pronounced low energy bias across the feature (cool colors). This results from when GLM does not record events at a certain pixel which the model suggests should have triggered due to its proximity to the brightest event. As in Figure 14d, this can occur as small differences along the edge of the group (<1 fJ per group) or large differences near the brightest event in the group (on the order of 1–10 fJ per group). The first category of small differences can be problematic in high group rate environments because they are almost always negative and—despite individually meeting the 5% threshold mentioned above—their aggregate sum introduces a non-trivial negative bias into the accumulated grids. To reduce this bias, we do not allow contributions from the smallest energy differences in the sample. The minimum energy difference that we consider 0.1 fJ in this case.

The panels in Figure 14 indicate regions where more energy than expected escapes the cloud in red and regions that are darker than expected in blue. The vertical integration in Figure 14b masks the contributions from individual layers that are shown in the subsequent panels. Many of the prominent features in Figure 14b originate below 9 km altitude (Figures 14d and 14e). The high-altitude layers primarily describe distinct convective features that have an energy deficit at their centers surrounded by an energy surplus where light can escape the side of the cloud. Two of these features can be noted in Figure 14 (most visible in Figure 14f) that are aligned with the older MCS feature in the south and the newer storm feature in the north from Figure 14a. However, while the southern

cold centers in each altitude layer are aligned vertically in Figures 14e–14g, the northern cold feature has a southward tilt such that it is only aligned with the highest cloud-tops from Figure 14a in the top layer that corresponds to 12–15 km (Figure 14g).

The reason for this apparent tilt is that the GLM MME grid is sensing the development of the northernmost cold cloud feature between 09:30 UTC and 10:00 UTC. The cold centers in the lower layers are displaced further north than the coldest ABI cloud tops because they describe the illumination of this new feature while the higher cold centers are still sensing the older northern feature that is aligned with the minimum ABI IR T_b. In the next ABI scan, the northernmost convective feature would strengthen, generating a new FED maximum to the northwest of the 09:30 UTC FED maximum (i.e., Figure 8a) with increasing ABI CTH values and frequent lightning at high altitudes in the storm.

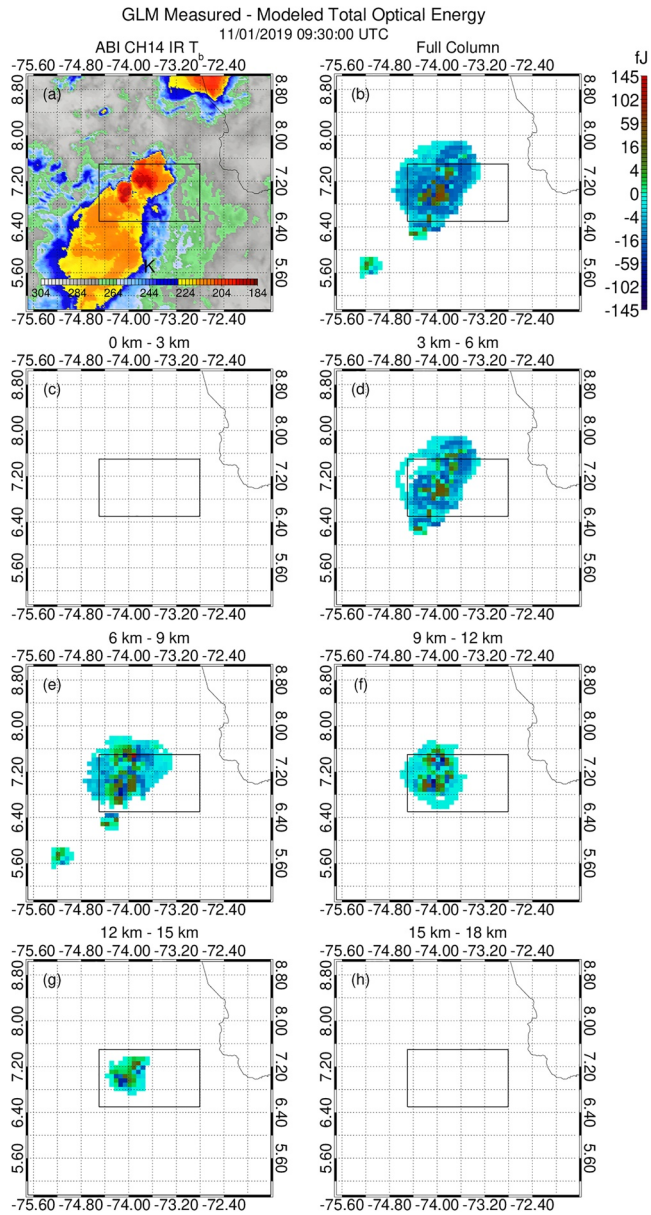

Figure 14. Maps of (a) Advanced Baseline Imager Channel 14 (11.2 μm) infrared brightness temperature and the Geostationary Lightning Mapper MME product as (b) a vertical integration through the full column and in layers between (c) 0–3 km, (d) 3–6 km, (e) 6–9 km, (f) 9–12 km, (g) 12–15 km, and (h) 15–18 km altitude valid at 09:30 UTC. Layers are greater than or equal to the lower limit and less than the upper limit.

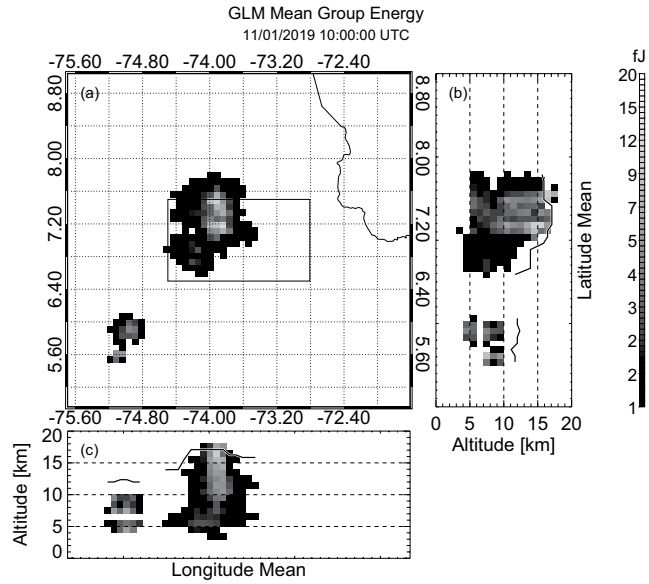

Figure 15. Volumetric Mean Group Energy valid from 10:00 to 10:15 UTC, plotted as in Figure 5.

By the end of this period at 10:00 UTC, the MGE (Figure 15) and MME (Figure 16) features corresponding to the older MCS had largely eroded, leading to low MGE values mostly below 10 km altitude and only slight MME gradients. By contrast, the strengthening northern features each produced their own distinct vertically aligned MGE maximum and MME minimum that extended from the cloud top down to at least 9 km altitude. The brightest pixels do not result from stratiform flashes or cases of edge illumination at this time step, but instead are collocated with the FED maxima. Unlike at 09:30 UTC, the highest MGE values come from the small and frequent flashes extending to high altitudes that dominate the FED at this time and the increasing MGE with altitude results from the decreasing optical depth of cloud separating these flashes from the imager. Still, there are cases of groups with radiance anomalies that preferentially illuminate the gap between the two features rather than either feature. These groups have a nearly linear appearance and are responsible for the positive MME anomalies in Figures 16f and 16g.

While the location of the lightning sources relative to the cloud features is important for determining how the illuminated clouds appear from orbit, there is a general correlation between strengthening/weakening convection and the peak amplitude of the MME product. As with the northern convective feature, strong convection tends to have an energy deficit in the tens of femtojoules at individual vertical levels, while weakening convection might have an energy deficit around 1 fJ or smaller, or even a slight surplus. While the goal of this work is to demonstrate the 3D imaging capability, these results show that additional work is merited to examine how well these GLM thundercloud imagery products (including TOE, MFEn, and MGE) track changes in cloud products derived from radar, passive microwave, and infrared/visible thunderstorm measurements.

4. Conclusion

In this fourth part of our thundercloud illumination study, we use the GLM group-level source altitude retrieval developed in Peterson et al. (2021c) to construct volumetric meteorological and thundercloud imagery from the GLM data collected during a Colombia thunderstorm. Analyses of these 3D gridded products demonstrate that de-coupling the trends from lightning at different altitudes reveals aspects of thunderstorm development that are masked by the vertical integration in the current NOAA 2D GLM grids. Not only do flashes at low altitudes behave differently than lightning closer to the cloud top, but flash rates also vary with time and altitude.

Tracking these changes can be useful for diagnosing changes in convection (including invigoration and dissipation) that determine the risk of severe weather. Sudden increases or, "jumps," in the thunderstorm flash rate have been shown to be symptomatic of the convective invigoration that leads to severe weather events and Lightning

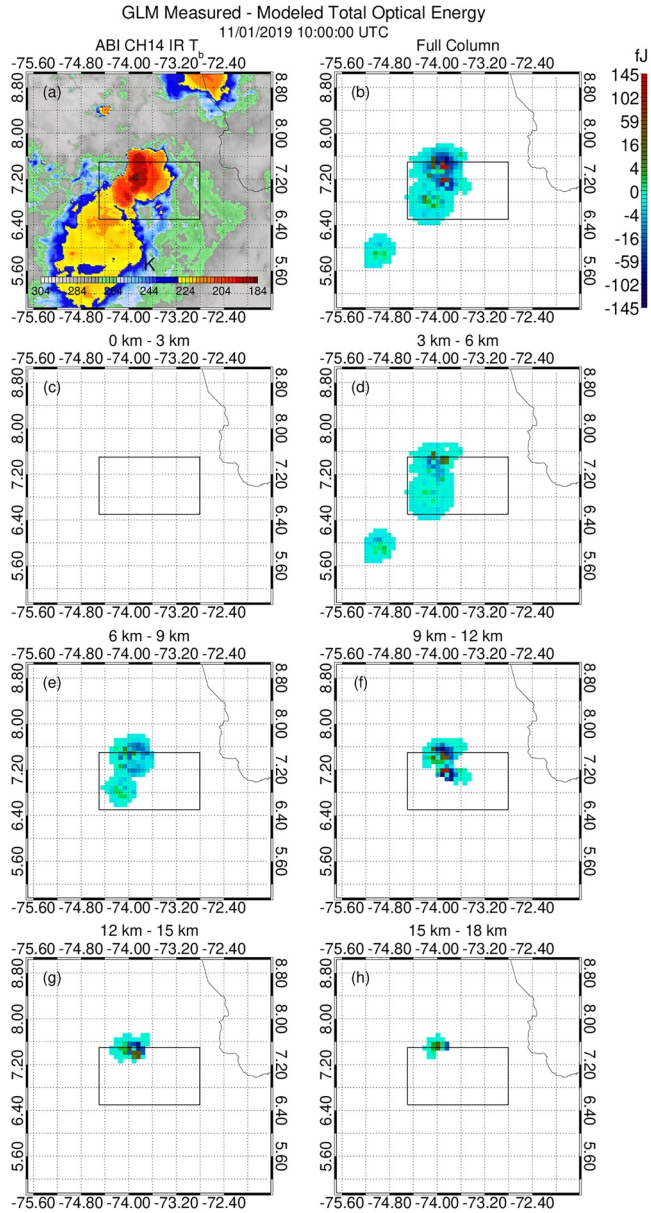

Figure 16. As in Figure 14, but valid at 10:00 UTC.

Jump Algorithms (LJAs) have been developed for LMA data (Schultz et al., 2009) to predict imminent severe weather events. However, poor GLM detection of compact flashes at low altitudes (whose optical energy must transmit through a dense cloud medium to trigger GLM) poses a problem for identifying jumps in the flash rate data. GLM-derived flash altitudes could provide an alternate method for identifying these strengthening updrafts that is not adversely affected by poor GLM performance near the cloud base. Moreover, early periods of vertical development might be detected at low altitudes in the 3D meteorological or thundercloud imagery products (as we saw with MME in our discussion of Figure 16) before signals at high altitudes (i.e., the 2D FED and ABI products like CTH).

These 3D grids, if constructed for the full disk, would provide a more comprehensive picture of lightning activity across the GLM domain than the current 2D gridded products. Moreover, the 20 s update cycle of GLM data would enable rapid update volumetric imagery to be generated to complement imagery from the ABI mesoscale sectors. If this can be done in a time-efficient manner, then it would provide a new perspective on GLM data

to operational users. However, the key challenge will be to develop a universal altitude retrieval that works for every storm type within the GLM FOV. Future work will improve the machine learning approach developed in Peterson et al. (2021c) for use with multiple types of thunderstorms across the GLM FOV and with the global data generated by LIS and OTD since 1995.

Data Availability Statement

The GLM/LMA matched data used in the study are available at the Harvard Dataverse and may be accessed via the DOI listed in Peterson (2021b). The machine learning models developed in this study are also available at the Harvard Dataverse and may be accessed via the DOI listed in Peterson (2021c).

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 NASA 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 Colombia LMA at the Technical University of Catalonia and Dr. Jesús López for sharing their processed LMA data for the presented case. Dr. Ken Cummins for providing valuable feedback on this work across its various parts.

References

Aranguren, D., Lopez, J., Montanya, J., & Torres, H. (2018). Natural observatories for lightning research in Colombia. In *2018 International Conference on Electromagnetics in Advanced Applications (ICEAA)* (pp. 279–283). IEEE. https://doi.org/10.1109/iceaa.2018.8520371

Bitzer, P. M. (2017). Global distribution and properties of continuing current in lightning. *Journal of Geophysical Research: Atmospheres*, *122*(2), 1033–1041. https://doi.org/10.1002/2016jd025532

Blakeslee, R. J., Lang, T. J., Koshak, W. J., Buechler, D., Gatlin, P., Mach, D. M., et al. (2020). Three years of the lightning imaging sensor onboard the international space station: Expanded global coverage and enhanced applications. *Journal of Geophysical Research: Atmospheres*, *125*, e2020JD032918. https://doi.org/10.1029/2020JD032918

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., 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*, *124*, 14285–14309. https://doi.org/10.1029/2019JD030874

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

Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Boeck, W. L., Buechler, D. E., Driscoll, K. T., & Stewart, M. F. (2003). Global frequency and distribution of lightning as observed from space by the Optical Transient Detector. *Journal of Geophysical Research*, *108*(D1). ACL 4-1–ACL 4-15. https://doi.org/10.1029/2002jd002347

Christian, H. J., Blakeslee, R. J., Goodman, S. J., & Mach, D. M. (Eds.). (2000). *Algorithm Theoretical Basis Document (ATBD) for the Lightning Imaging Sensor (LIS)*. NASA/Marshall Space Flight Center. Retrieved from http://eospso.gsfc.nasa.gov/atbd/listables.html

Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., et al. (2013). The GOES-R geostationary lightning mapper (GLM). *Atmospheric Research*, *125–126*, 34–49. https://doi.org/10.1016/j.atmosres.2013.01.006

Goodman, S. J., Mach, D., Koshak, W. J., & Blakeslee, R. J. (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

Heidinger, A. (2011). *Algorithm theoretical basis document ABI cloud height*. NOAA NESDIS Center for Satellite Applications and Research. Retrieved from http://www.goes-r.gov/products/ATBDs/baseline/Cloud_CldHeight_v2.0_no_color.pdf

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

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

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

Lojou, J.-Y., & Cummins, K. L. (2004). On the representation of two- and three-dimensional total lightning information. In *Preprints, Conference on Meteorological Applications of Lightning Data*. (pp. Paper 2.4, AMS Annual Meeting, San Diego, CA, USA).

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

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

Mach, D. M. (2020). Geostationary Lightning Mapper clustering algorithm stability. *Journal of Geophysical Research: Atmospheres*, *125*(5), e2019JD031900. https://doi.org/10.1029/2019jd031900

Mach, D. M., Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Goodman, S. J., & Boeck, W. L. (2007). Performance assessment of the optical transient detector and lightning imaging sensor. *Journal of Geophysical Research*, *112*(D9), D09210. https://doi.org/10.1029/2006jd007787

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peterson, M. (2019a). Using lightning flashes to image thunderclouds. *Journal of Geophysical Research: Atmospheres*, *124*, 10175–10185. https://doi.org/10.1029/2019jd031055

Peterson, M. (2019b). 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. (2020). 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. (2021a). Where are the most extraordinary lightning megaflashes in the Americas? *Bulletin of the American Meteorological Society*, *102*(3), E660–E671. https://doi.org/10.1175/bams-d-20-0178.1

Peterson, M. (2021b). *GLM-CIERRA*. https://doi.org/10.5067/GLM/CIERRA/DATA101

Peterson, M. (2021c). *Coincident optical and RF lightning detections from a Colombia thunderstorm*. Harvard Dataverse, V1. https://doi.org/10.7910/DVN/5FR6JB

Peterson, M. (2021d). Holes in optical lightning flashes: Identifying poorly transmissive clouds in lightning imager data. *Earth and Space Science*, *8*, e2020EA001294. https://doi.org/10.1029/2020EA001294

Peterson, M., Light, T., & Mach, D. (2021a). The illumination of thunderclouds by lightning: Part 1: The extent and altitude of optical lightning sources. *Journal of Geophysical Research: Atmospheres*.

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*. https://doi.org/10.1029/2021ea001944

Peterson, M., & Rudlosky, S. (2019). The time evolution of optical lightning flashes. *Journal of Geophysical Research: Atmospheres*, *124*, 333–349. https://doi.org/10.1029/2018jd028741

Peterson, M., Rudlosky, S., & Deierling, W. (2017). The evolution and structure of extreme optical lightning flashes. *Journal of Geophysical Research: Atmospheres*, *122*, 13370–13386. https://doi.org/10.1002/2017jd026855

Peterson, M., Rudlosky, S., & Zhang, D. (2020). 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. (2020). 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

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

Rudlosky, S. D., Goodman, S. J., Virts, K. S., & Bruning, E. C. (2019). Initial geostationary lightning mapper observations. *Geophysical Research Letters*, *46*(2), 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

Schultz, C. J., Petersen, W. A., & Carey, L. D. (2009). Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *Journal of Applied Meteorology and Climatology*, *48*(12), 2543–2563. https://doi.org/10.1175/2009jamc2237.1

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

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

Thomas, R., Krehbiel, P. R., Rison, W., Hamlin, T., Boccippio, D. J., Goodman, S. J., & Christian, H. J. (2000). Comparison of ground-based 3-dimensional lightning mapping observations with satellite-based LIS observations in Oklahoma. *Geophysical Research Letters*, *27*(12), 1703–1706. https://doi.org/10.1029/1999gl010845

Thomas, R. J., Krehbiel, P. R., Rison, W., Hunyady, S. J., Winn, W. P., Hamlin, T., & Harlin, J. (2004). Accuracy of the lightning mapping array. *Journal of Geophysical Research*, *109*, D14207. https://doi.org/10.1029/2004JD004549

Winn, W. P., Aulich, G. D., Hunyady, S. J., Eack, K. B., Edens, H. E., Krehbiel, P. R., et al. (2011). Lightning leader stepping, K changes, and other observations near an intracloud flash. *Journal of Geophysical Research*, *116*, D23115. https://doi.org/10.1029/2011JD015998

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), e2019JD032024. https://doi.org/10.1029/2019jd032024