LA-UR-20-25866 (Accepted Manuscript)

The Hazards Posed by Mesoscale Lightning Megaflashes

Peterson, Michael Jay
Stano, Geoffrey

To be published in: Earth Interactions

DOI to publisher's version: 10.1175/EI-D-20-0016.1

Permalink to record: http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-20-25866

The Hazards Posed by Mesoscale Lightning Megaflashes

Michael Peterson[1], Geoffrey Stano[2]

[1] ISR-2, Los Alamos National Laboratory, Los Alamos, New Mexico
[2]The University of Alabama in Huntsville, Huntsville, Alabama

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

LA-UR-20-25866

Abstract

Lighting megaflashes extending over >100 km distances have been observed by the Geostationary Lightning Mappers (GLMs) on NOAA's 16-series Geostationary Operational Environmental Satellites (GOES). The hazards posed by megaflashes are unclear, however, due to limitations in the GLM data. We address these by reprocessing GOES-16 GLM measurements from 1/1/2018 to 1/15/2020 and integrating them with Earth Networks Global Lightning Network (ENGLN) observations. 194,880 GLM megaflashes are verified as natural lightning by ENGLN. Of these, 127,479 flashes occurred following the October 2018 GLM software update that standardized GLM timing. Reprocessed GLM/ENGLN lightning maps from these post-update cases provide a comprehensive view of how individual megaflashes evolve.

This megaflash dataset is used to generate statistics that describe their hazards. The average megaflash produces 5-7 CG strokes that are spread across 40-50% of the flash extent. As flash extent increases beyond 100 km, megaflashes become concentrated in key hotspot regions in North and South America while the number of CG and IC events per flash and the overall peak current increase. CGs in the larger megaflashes occur over 80% of the flash extent measured by GLM, while the majority contain regions where the megaflash is the only lightning activity in the preceding hour. These statistics demonstrate that there is no safe location below an electrified cloud that is producing megaflashes and current lightning safety guidance is not always sufficient to mitigate megaflash hazards.

1 Introduction

While lightning occurs most frequently in intense convection, the overall lightning hazard encompasses all surrounding regions where an individual or an operation might be adversely affected by lightning. The lightning hazard differs according to which part of the thunderstorm is being considered. Lightning is common in the convective core (Peterson and Liu, 2011) where other hazards such as hail and strong winds exist that motivate individuals to seek shelter. Lightning flash rate trends are symptomatic of updraft characteristics (Deierling and Petersen,2008) and ice fluxes (Deierling et al., 2008), and sudden increases (i.e, "jumps") in lightning activity have been used to predict the onset of severe weather (Williams et al., 1999; Schultz et al., 2009).

The 30-30 rule for lightning safety that was proposed by a Lightning Safety Group (LSG) at the Annual Meeting of the American Meteorological Society (AMS) in 1998 (Holle et al., 1999) works best with this convective-type lightning. By this rule, lightning is considered dangerous if the time difference between the flash of light and the clap of thunder is less than 30 s. This delay is due to the difference between the speed of light and the speed of sound in air, and works out to describe a lightning strike within ~10 km of the observer. Holle et al. (1999) noted the 30-s "flash-to-bang" part of the rule was insufficient for certain types of lightning. Lopez and Holle (1999) suggested that greater distances should be considered for large, organized convective systems. This creates a problem for the perception of danger, however. They note that lightning is not perceived to be close to the observer when longer flash-to-bang times are used. This can lead the observer to not appreciate the risk until the next strike occurs at their location. Moreover, the perception of low risk is amplified when the apparent flash rate is low - with minutes between visible strokes. Due to these limitations, some organizations do not recommend

using the 30-30 rule. For example, the guidance provided by NOAA recommends seeking shelter on any detection of thunder (i.e., "when thunder roars, go indoors"; NOAA, 2018). If the lightning is close enough that an observer can hear the audible shockwave it generates, then it is potentially close enough to strike them.

Lightning flashes outside of the convective core pose a unique hazard compared to convective lightning. This is because there exists a natural opposition between flash frequency and flash size (Bruning and MacGorman, 2013). While the heterogeneous electric field in the convective core produces high rates of relatively small flashes, homogeneous non-convective electrified clouds are infrequently discharged by lightning flashes that develop laterally over long horizontal distances. The overall maximum size of the flash is only limited by the extent of the charge reservoir that it can access in the electrified cloud. Large flashes are particularly common in Mesoscale Convective Systems (MCSs), while the largest cases occur exclusively in these organized convective systems. MCSs are favorable for large lightning because they produce electrified stratiform regions that can extend over hundreds of kilometers (Marshall and Rust, 1993; Stolzenburg et al., 1994; Lang et al., 2004) through charge advection from the convective line (Carey et al., 2005) and in-situ generation (Rutledge and MacGorman, 1988; Ely et al., 2008; Lang and Rutledge, 2008). These long horizontal lightning flashes have been termed "megaflashes" (Lyons et al., 2020) and are defined as a mesoscale lightning flash that is at least 100 km long.

The factors that describe non-convective lightning hazards – long horizontal flashes occurring in low flash rate regions of larger organized storm systems - are each, individually, conducive to an underappreciated lightning threat. Their combination is an ideal mix for a "bolt from the blue" if under clear skies or a "bolt from the grey" (Lyons, 2020) if under cloudy skies.

Perhaps the storm passed long ago with only low stratiform clouds remaining. Then, suddenly, a lightning flash comes from over the horizon and streaks across the sky putting down multiple Cloud-to-Ground (CG) strokes along its path. Those stratiform clouds overhead were electrified and, even though they were not actively flashing on their own, they still serve as a charge reservoir for lightning initiated elsewhere (Marshall and Rust, 1993; Lang et al., 2004; Carey et al., 2005). Further adding to the hazard, this type of lightning often produces positive CG (+CG) strokes with high peak currents and continuing current (CC) that lead to large charge moment changes. The physical attributes of these strokes are favorable for initiating forest fires (Latham and Williams, 2001) and generating exotic upper-atmosphere electrical discharges such as sprites (Franz et al., 1990; Williams, 1998; Lyons et al., 2009; Williams et al., 2010).

Scenarios as described above have been documented for individual cases of megaflashes that were mapped from space by NOAA's Geostationary Lightning Mapper (GLM: Goodman et al., 2013; Ruslosky et al., 2019). GLM is the first operational lightning detector that can map individual flash extent over broad (hemispheric scale) geospatial domains. Ground-based radio-frequency (RF) lightning networks including the National Lightning Detection Network (NLDN) resolve the locations of strokes and some cloud pulses, but these sparse detections are not sufficient to resolve megaflash structure. Lyons et al. (2020) showed an impressive megaflash case where the most distant NLDN events associated with the GLM flash were 500 km apart – starting on the Oklahoma-Texas border and ending in central Kansas. The strongest +CG strokes had peak currents exceeding 300 kA and charge moment changes > 3100 C km (well within the range for sprite production).

Individual case studies are instructive for demonstrating what megaflashes are capable of but documenting the lightning hazard posed by megaflashes requires taking a statistical

105 approach. The geostationary orbit of the GOES satellites allows GLM to record a staggering
106 amount of lightning data. The GOES-16 GLM detects on the order of a million lightning flashes
107 per day. Each year of GOES-16 GLM observations includes around 365 million flashes, which is
108 nearly 10x more lighting than the 38 million flashes that NASA's Optical Transient Detector
109 (OTD) and Lightning Imaging Sensor (LIS) instruments could have observed (i.e., after
110 correcting for instrument Detection Efficiency) during their combined 25 years in Low Earth
111 Orbit.

112 Unfortunately, the operational GLM data does not permit megaflashes to be identified
113 routinely. Strict latency requirements have resulted in hard limits being imposed by the
114 operational GLM ground system software (Goodman et al., 2010) for the maximum complexity
115 and duration of a single lightning flash. When a flash exceeds either 101 "groups" (an
116 approximation for individual optical pulses) or 3 s in duration, it will be terminated by the
117 ground system software and any additional detections will be assigned to a new flash. This
118 results in megaflashes being artificially split into dozens of smaller flashes.

119 To identify these megaflash cases, the operational GLM lightning data needs to be
120 reprocessed to resolve each complete and distinct lightning flash. We employ a "reclustering"
121 approach (Peterson, 2019) that evaluates the clusters in the operational GLM data produced by
122 NOAA, identifies any flashes that contain groups that should be clustered into the same flash,
123 and then merges the split flashes into a single flash cluster. The largest case of natural lightning
124 in the reclustered GLM dataset was a 709-km megaflash that recently has been recognized by the
125 World Meteorological Organization (WMO) as the global lightning extreme for flash extent
126 (Peterson et al., 2020a). Another 16.73 s flash in this dataset was also recognized by the WMO
127 as the global lightning extreme for flash duration.

In the present study, we integrate ground-based RF lightning measurements with our reclustered GLM dataset to document the lightning hazard posed by megaflashes across the Americas. As with Lyons et al. (2020), the RF measurements provide information on the ground strike locations and peak currents that are not measured by GLM. We use this combined dataset to produce statistics on the number of strokes per megaflash, the peak current and polarity of megaflash strokes, and the fraction of the megaflash horizontal extent where ground strikes occur. These statistics reinforce the unpredictable nature of the megaflash lightning hazard. Ground strikes can occur anywhere within the megaflash extent and frequently have high peak currents that are capable of causing damage, injury, or igniting fires. As suggested by Lopez and Holle (1999), greater care should be taken with organized convective systems – especially when near electrified anvil and stratiform clouds that are capable of producing a megaflash. Lightning in these regions may be infrequent, but it only takes one unexpected lightning flash to spark a tragedy.

2 Data and Methodology

2.1 Geostationary Lightning Mapper (GLM) Data

Megaflashes are identified in the reclustered GOES-16 GLM science data described at length in Peterson (2019) and more recently in Peterson et al. (2020a). This reprocessed dataset extends from 1/1/2018 until 1/15/2020 and includes the whole GOES-16 GLM domain that covers North and South America from 54° S to 54° N.

GLM detects lightning as transient increases in cloud illumination at the 777.4 nm Oxygen emission triplet. The GLM domain is imaged at 500 frames per second on a 1372x1300

pixel Charge-Coupled Device (CCD) imaging array. The GLM imaging array features variable pitch pixels that maintain a relatively consistent horizontal resolution projected to ground ranging from 8 km at nadir to 14 km at the limb. The steady-state radiant energy of the background scene at each pixel is subtracted from the instantaneous pixel energy, and then an "event" is registered if this difference exceeds the threshold for detection (Rudlosky et al., 2019).

The GLM data is organized into a hierarchy of cluster features that describe lightning over a range of temporal and spatial scales. Individual events during a single integration frame are the basic unit of GLM detection. Events do not represent complete physical processes, but rather describe locations on the CCD array that light up during lightning phenomena. Events are clustered into "groups" that describe contiguous regions on the CCD array that light up simultaneously. Thus, groups approximate cloud illumination from individual optical pulses generated by lightning. This is only an approximation because the 2-ms duration of GLM integration frames is considerably larger than the duration of individual optical pulses (usually on the order of 100 microseconds). Thus, the possibility exists that a single GLM group might capture multiple pulses. On the other hand, CC generates sustained optical emission that would last for multiple 2-ms GLM groups.

Groups that are close in space and time are then clustered into higher-level features that describe distinct lightning flashes. The process for constructing flashes is based on the clustering technique employed with LIS and validated over its 17-year mission on the Tropical Rainfall Measuring Mission (TRMM) satellite (Mach et al., 2007). For LIS, group centroids were evaluated for flash assignment by a three-term Weighted Euclidean Distance (WED) model applied in geolocated space. The three terms were the zonal difference in position (DX) between groups, the meridional difference in position (DY), and the time difference (DT). The spatial

terms were weighted by a threshold of 5.5 km while the temporal term was weighted by 330 ms. If two groups fell within the sphere defined by WED = 1, then they were determined to belong to the same flash.

The GLM clustering algorithm described in Goodman et al. (2013) differs from this LIS algorithm in two key ways. First, rather than using the group centroid locations as the basis for clustering, the GLM algorithm considers the positions of all events that constitute the group feature. If any of these events satisfy the WED model with an event in another group, they will be clustered into the same flash. The second key difference is the spatial threshold chosen. GLM uses the same 16.5 km threshold that was employed with the OTD instrument instead of the 5.5 km LIS threshold to accommodate the larger 8-14 km GLM pixels. Mach (2020) evaluated the clustering scheme used for GLM and found that variations in algorithm thresholds did not lead to large changes in the resulting flash rates for all but the most active thunderstorms (>40 flashes per minute).

The reclustered GLM data aims to extend the standard operational GLM data while preserving its structure and conventions. This post-processing evaluates the flash clusters generated by the GLM ground system as described above, identifies cases where flashes are artificially split by the hard limits in flash group count (101) and flash duration (3 s) coded into the ground system software, and then merges the split flash features together into a single distinct and complete flash feature. This processing also adds two feature levels to the GLM hierarchy that are not implemented in the ground system processing. “Area” features that approximate thunderstorm snapshots in the former LIS / OTD data are added that combine flashes in close spatial and temporal proximity into a single feature. “Series” features (Peterson et al., 2017) are also added that describe distinct periods of sustained optical emission from a single flash.

196 Finally, the post-processing adds flash metrics including flash extent (Peterson et al., 2018) and
197 optical multiplicity (Peterson and Rudlosky, 2019), and constructs gridded products – such as
198 Flash Extent Density (FED: Lojou and Cummins, 2005) and convective probability (Peterson et
199 al., 2020b) - that are packaged alongside the lightning cluster feature data.

200 In this study, the maximum flash extent, defined as the maximum great circle distance
201 between any two group centroids in a single flash, will be used to identify megaflashes. Any
202 flash that exceeds 100 km in extent will be designated a megaflash. Due to the meandering
203 nature of long horizontal lighting channels, the actual flash length would likely be greater.
204 However, space-based instruments like GLM are limited in the level of detail that they can
205 resolve with their kilometer-scale pixels, and methods that attempt to quantify the unique flash
206 length (i.e., not counting re-illumination) are computationally expensive (Peterson et al., 2018).
207 Identifying megaflashes using a 100 km extent threshold is a computationally-inexpensive way
208 to ensure that smaller flashes are not included in the sample, but smaller megaflashes with total
209 lengths > 100 km and extents < 100 km will be missed.

210 2.2 Earth Networks Global Lightning Network (ENGLN) Data

211 Beyond the flash length versus flash extent issue, there are two key caveats in using the
212 GLM data to identify megaflashes and the hazard that they pose. First, GLM does not report the
213 locations of ground strikes. GLM is a total lightning detector that cannot reliably differentiate
214 individual intracloud (IC) discharges from Cloud-to-Ground (CG) strokes. Ground networks
215 excel at identifying the locations and times of strokes. Combining GLM and ground network
216 observations mitigates the lack of GLM stroke information and informs the origins of the optical
217 pulses recorded by GLM. Second, the GLM data contains artifacts from solar contamination that
218 can masquerade as megaflash activity (Peterson, 2020a). These flashes can be additionally

219 screened by looking for a lack of coincidence with ground network observations. Thus, both key
220 caveats are mitigated through data fusion with a ground network.

221 In this study, Earth Networks Global Lightning Network (ENGLN) data are acquired
222 from Earth Networks and integrated into the GLM clustering hierarchy for the megaflash cases.
223 ENGLN is a distributed heterogeneous global network of long-range ground-based RF lightning
224 sensors. ENGLN integrates observations from two networks: the Earth Networks Total Lightning
225 Network (ENTLN: Zhu et al., 2017) and the World-Wide Lightning Location Network
226 (WWLLN: Jacobson and Holzworth, 2006; Hutchins et al., 2012). ENGLN data includes the
227 position and time of lightning events, their type (CG or IC), and also their peak current and
228 polarity. However, it should be noted that distant +CG strokes can be reported as -CGs if the
229 ground wave becomes attenuated. Thus, -CGs reported from megaflashes might, in fact, be mis-
230 classified +CGs.

231 2.3 Adding ENGLN Events to GLM Megaflashes

232 Our approach for clustering the ENGLN data into the GLM data tree assumes that (1) all
233 ENGLN reports (CG or IC) that are co-located with a GLM group contribute optical energy to
234 that group , and (2) not all ENGLN reports will lead to GLM groups (for example, if the cloud is
235 too optically thick to allow transmission to space that are bright enough for GLM to detect).
236 Thus, ENGLN reports should be close to the GLM events that comprise groups in space and
237 time, but some leeway should be granted to limit the number of missed reports.

238 We treat ENGLN events as “groups” (approximating complete lightning pulses) for
239 clustering purposes and assign them to GLM flashes if they occur within 16.5 km and 500 ms of
240 any GLM event within one of the groups from that flash. We use the box-distance clustering
241 algorithm from OTD rather than the WED method used by LIS and GLM to reduce

computational expense. While this clustering approach is applied to all ENGLN events that share coincidence with GLM megaflashes, it is important to note that the rates of matched events are not uniform in space and time. The GLM operating software was updated multiple times during our two-year period in the reprocessed GLM record (2018-2020), some of which improved the geolocation and timing accuracy in the later portion of the data record. These changes have minimal impact on whether a GLM megaflash contained an ENGLN event but will affect the number and locations of matched ENGLN events in a given GLM flash. Thus, we focus our assessment of matched GLM megaflash characteristics on the 10/31/2018 – 1/15/2020 period with the best timing information.

The other major factor impacting clustering uniformity is the fact that ENGLN does not have a uniform sensor density. Dense observations permit more events (especially weaker events) to be resolved. The sensor density is greatest in the United States, and the ENTLN domain in the surrounding regions contains drastically more events per square kilometer than the remainder of the GOES-16 GLM Field of View (FOV). As with GLM timing, this is not expected to impact whether a GLM megaflash will have ENGLN coincidence, but it will influence the number of coincident ENGLN events (especially IC events) per flash and their relative extent compared to the GLM flash extent.

Figure 1 shows an example GLM meagaflash with ENLGN events added. This particular megaflash over Louisiana was identified in Peterson (2019) as having the greatest unique footprint area reported by GLM, a 634-km overall extent, and a duration of nearly 10.5 s. The groups in this flash (connected by line segments in the central panel) developed westward from the flash start position at the rear of the convective line and then spread in many directions throughout the stratiform region of the MCS. The latitude extent of each time-ordered group in

265 the flash is shown to the right of the map while the longitude extent of each group is shown
266 above the map. A timeseries of group area (above the time axis) and group energy (below the
267 time axis) is shown along the bottom of the figure. ENGLN CGs are added as asterisk symbols in
268 both the map and the top timeseries, while ENGLN ICs are depicted as box symbols. In total,
269 126 ENGLN events were reported during this flash including 36 -CG strokes and 17 +CG
270 strokes. The first of these strokes occurred 1.126 s into the flash while the last occurred 0.601 s
271 before the end of the GLM flash. The strokes were not clustered in a single location, but rather
272 scattered throughout the 114,000-km^2 footprint of the GLM flash. The strongest -CG stroke from
273 this flash had a peak current of -118 kA while the strongest +CG stroke had a peak current of
274 +133 kA.

275 This information about the strokes that occurred during this flash was not available in the
276 previous analysis in Peterson (2019) because it only considered GLM and did not add ground
277 network observations. On the other hand, the ground networks do not map lightning flashes with
278 a sufficient level of detail to identify flash structure – information that is readily available with
279 GLM. Data fusion between these optical and RF measurements from the same flash enable
280 comprehensive assessments of individual megaflashes that are not possible with either
281 phenomenology in isolation.

282 Our merged GLM / ENGLN data contains 194,880 GLM megaflashes like the example in
283 Figure 1 that were observed between 1/1/2018 and 1/15/2020 across the GOES-16 GLM Full
284 Disk domain. This megaflash data is hosted at Peterson (2020b). These flashes are associated
285 with a total of 4.5 million ENGLN events. 1 million of these events (22%) were from CGs while
286 the remaining 3.5 million events (88%) were ICs. We will focus, however, on the period with
287 improved GLM timing accuracy (10/31/2018 onward), reducing the size of the megaflash sample

288 considered to 127,479 flashes (65% of all GLM megaflashes) across the GOES-16 GLM full
289 disk.

290

291 **3 Results**

292 The following sections assess the megaflash lightning hazard. Section 3.1 maps the
293 locations and peak extents of ENGLN-matched GLM megaflashes, and then summarizes their
294 overall attributes that define the lighting hazard. Section 3.2 elaborates on the statistics of
295 ENGLN matches by examining their frequencies and peak currents as a function of GLM
296 megaflash extent. Finally, Section 3.3 addresses the questions of whether megaflashes pose a risk
297 of CG strikes over their full spatial extent as mapped by GLM, and whether megaflashes are
298 sufficiently isolated in time that the public might have resumed outdoor activities when these
299 flashes occur if following the 30-30 rule.

300

301 *3.1 Overall Statistics on GLM Megaflash Location and Composition*

302 Megaflashes may be relatively uncommon in the GLM record compared to convective
303 lightning, but there are certain regions in the Americas that produce, on average, one-or-more
304 megaflashes per day. Figure 2a shows the locations of these "hotspot" regions: the Great Plains,
305 Gulf Coast, and Eastern Seaboard of the United States, coastal Central America from Mexico to
306 Colombia, and portions of southern Brazil, Uruguay, Paraguay, Bolivia, and Argentina in South

America. Note that Figure 2 is the only analysis in this study that uses all 194,880 ENGLN-matched GLM megaflashes.

While the term "megaflash" is applied to each case of 100+ km lightning, some flashes far exceed this threshold and extend for multiple hundreds of kilometers. These longer megaflashes exhibit notably different behavior than their 100-km counterparts. The first example of this is in Figure 2b, which shows the peak megaflash extent across the Americas. While 100-km megaflashes can occur anywhere, the largest flashes observed at most locations across the GLM FOV are 100-200 km across. The largest megaflashes that have been observed by GLM thus far are 500-700 km in extent (Lyons et al., 2020; Peterson et al., 2020), and these have only been detected in the Great Plains in North America and the La Plata basin in South America. Megaflashes and their associated hazards might be common in certain coastal and oceanic regions - for example, along the Central American coast - but only these continental basins have been shown to produce MCS thunderstorms capable of generating megaflashes that cover the equivalent land areas of entire states at a time. These large megaflashes have the potential to be particularly dangerous because of their exceptional distance from the convective core of the parent thunderstorm. Locations far removed from the lightning maxima in the storm core may be interpreted as having a low risk for lightning impacting outdoor activities. However, as long as these outlying clouds remain electrified, they can provide a conduit for megaflashes to strike "out of the grey."

The lightning hazard posed by megaflashes, in general, is quantified in Table 1. To improve the likelihood of matching GLM and ENGLN events, only the 10/31/218 – 1/15/2020 data (described in section 2.3) is used from this point forward. The average megaflash across the GOES-16 GLM Full Disk domain contains 5.5 ENGLN events that include 4.5 -CGs, 1 +CG,

330 and 17.7 IC pulses. When CG strokes are detected, their average maximum separation is 51.8
331 km, or 37.1% of the overall GLM flash extent. When IC strokes are detected, their average
332 maximum separation is 75.6 km, or 56.8% of the GLM flash extent.

333 Despite using only the most recent GLM data to make these assessments, these numbers
334 still underrepresent the megaflash hazard due to the inclusion of sparse ENGLN observations far
335 from the dense portion of the network. The ENGLN-Only region outside of the United States has
336 fewer CGs and ICs per GLM flash that are spread over notably smaller fractions of the GLM
337 extent. However, it is possible this is due to the physical differences between land-based and
338 oceanic or tropical and subtropical megaflashes rather than just ENGLN detection efficiency.
339 Thus, Table 1 specifically compares the continental hotspot regions in North and South America
340 that both contain large and complex megaflash cases. Megaflashes in the North America hotspot
341 (within the ENTLN domain) contain 2.1x the number of CGs and 2.3x the number of IC pulses
342 than their South American counterparts. The CGs in these North American hotspot flashes
343 extend over half the GLM flash extent, while the ICs extend over 77% of the GLM flash extent.

344 The statistics in Table 1 show that megaflashes are not only able to generate multiple
345 ground strikes along their path, but that these CGs are also separated by a significant portion of
346 the flash extent measured by GLM. This demonstrates that the lightning hazard is not limited to
347 the regions surrounding the convective core of the thunderstorm. However, GLM flashes within
348 the ENTLN domain are resolved in greater detail by ENGLN than the flashes outside of this
349 domain. Thus, the lightning hazard posed by megaflashes outside of the ENTLN domain may be
350 underrepresented in some cases For this reason, the analyses of how the lightning hazard changes
351 with megaflash extent that will be presented in Section 3.2 and 3.3 will only use the data from
352 North America. This includes 41,616 megaflashes of the 127,479 total cases from 10/31/2018 or

353 later. The same analyses for the full disk are still performed, and these will be included as
354 Supplemental Information (SI), for reference. However, these full disk analyses will not be
355 discussed in detail in the following sections.

356

357 *3.2 Megaflash Lightning Hazards Posed by ENGLN Event Count and Peak Current*

358 It was shown in Section 3.1 that the average megaflash produces multiple CG strokes
359 over its 100+ km extent. However, do longer flashes generate more CGs? With access to a larger
360 charge reservoir, do these longer megaflashes generate greater peak currents that can be
361 particularly hazardous? To answer these questions, we produce two-dimensional histograms that
362 catalog megaflashes according to their GLM extent and either their ENGLN event count (Figure
363 3) or their ENGLN peak current (Figure 4). The number of GLM megaflashes in each bin is
364 depicted as a color contour plot. Cumulative Distribution Functions (CDFs) are also computed
365 for flashes with similar sizes, and line plots are overlaid showing the median (thick solid), 25th
366 and 75th percentile (thin solid), 10th and 90th percentile (dashed), and 5th and 95th percentile
367 (dotted) values.

368 The two-dimensional histograms in Figure 3 show that the ubiquitous 100 km
369 megaflashes in the sample can contain a wide range of ENGLN event counts (from 1 to >100).
370 The median number of ENGLN events are 19 ICs (Figure 3a) and 4 CGs (Figure 3b, 3 of which
371 are -CGs (Figure 3c). More than half of the 100-km megaflashes do not produce a +CG (Figure
372 3d). As we move up to larger GLM megaflashes, however, the percentile curves shift towards

373 increased numbers of ENGLN events per megaflash. The largest GLM megaflashes have median
374 IC and CG counts of 34 and 45, respectively, while 95% have at least 11 -CGs and 8 +CGs.

375 As the statistics for megaflashes with intermediate extents fall between these two
376 extremes, the risk of multiple megaflash ground strikes only increases with flash extent. At 100
377 km, there is still a sizable number of megaflash cases with a single CG (Figure 3b). However,
378 95% of megaflashes that are >140 km in extent contain multiple CGs and 95% of megaflashes
379 >290 km contain multiple +CGs. Figure 4 shows distributions of ENGLN CG peak current in
380 GLM megaflashes. For all flash extents, megaflash +CG peak currents are greater than -CG peak
381 currents. For 100 km megaflashes, 90% of -CG peak currents and 70% of +CG peak currents are
382 < 75 kA. However, by 430 km, over half of -CG peak currents and 90% of +CG peak currents
383 exceed 75 kA. This is an exceptional peak current threshold, especially for land-based lightning
384 (i.e., Said et al., 2013). Furthermore, 95% of the largest flashes have +CGs in excess of 93 kA.
385 Large flash extents lead to both an increased number of CGs as well as CGs with high peak
386 currents (especially +CGs).

387

388 *3.3 Megaflash Lightning Hazards Posed by ENGLN Event Extents and GLM Flash Rates*

389 A megaflash generating multiple CGs does not guarantee that strikes can happen
390 throughout its enormous extent. Moreover, the rarity of megaflashes does not, necessarily, mean
391 that they occur in isolation from other types of lightning. Over what fraction of the megaflash

392 extent does the risk of a ground strike exist? How often do megaflashes exist in regions where
393 someone observing the 30-30 rule would be caught off-guard?

394 Figure 5 shows two-dimensional histograms that compare GLM flash extent with the
395 maximum separation of ENGLN CG strokes (left) and ENGLN IC events (right) following the
396 conventions of Figures 3 and 4. These comparisons are made in terms of absolute great circle
397 distance (top) and as a fraction of the megaflash extent resolved by GLM (bottom). The ENGLN
398 maximum event separations increase nearly linearly with GLM events for both CGs (Figure 5a)
399 and ICs (Figure 5b). While the detected ENGLN events can be concentrated in a small portion of
400 the megaflash (especially in the smaller 100-km megaflashes), ENGLN CGs and ICs usually
401 exist throughout the megaflash extent measured by GLM.

402 Table 1 showed that the peak separation of ENGLN CGs is only 50% of the GLM extent,
403 overall. However, half of 330+ km megaflashes have ENGLN CGs spread across 80% of their
404 GLM extent and nearly 95% of the largest GLM megaflashes have ENGLN CGs covering three-
405 quarters of their extent. Median IC separations (Figure 5d) are near this 80% fraction of the
406 GLM flash extent over the full range of flash sizes. If a GLM flash is observed to be 700 km in
407 extent, one can reasonably expect ENGLN CG and IC sources to occur over a 400-600 km
408 distance. Therefore, it should not be assumed that a safe region exists below an electrified cloud
409 that is producing megaflashes. The first portion of the 30-30 rule (which results in only a 10 km
410 standoff distance) is not appropriate for megaflash cases.

411 But what about the second portion of the 30-30 rule? Is 30 minutes from the last lightning
412 flash a sufficient period of time to wait before resuming outdoor activities? To answer this
413 question, we use GLM Flash Extent Density (FED) to quantify the flash rates at all locations

414 within the megaflash footprint. FED is a gridded product that increments a given gridpoint once
415 for every flash that extends into that gridpoint. If even one of the events that comprises a flash
416 occurs over a particular location, then that flash is counted in the local flash rate. Otherwise, the
417 flash is not counted. Also note that because GLM is a total lightning sensor, the FED grids
418 describe both Cloud-to-Ground (CG) and Intracloud (IC) flashes. These FED analyses represent
419 a worst-case scenario for testing the 30 minute cessation time because it assumes that all ICs will
420 be audible.

421 Figure 6 shows two-dimensional histograms for the minimum flash rate (Figure 6a) and
422 the mean flash rate (Figure 6b) within the boundaries of each GLM megaflash. The minimum
423 value in each plot is one flash per hour (the maximum time period considered in this analysis) .
424 A sizable portion of the 100 km megaflashes occur in relatively active thunderstorm regions with
425 minimum (Figure 6a) and average (Figure 6b) FED values reaching 10 flashes per minute. More
426 than 95% of all flashes of each size occur in regions where the average FED flash rate exceeds 1
427 flash in 10 minutes. In these cases, the second half of the 30-30 rule would be appropriate.
428 However, Figure 6a also shows that a sizable fraction of megaflashes extend into regions where
429 they are the only lightning during a 20-30 minute period. In fact, half of the 100-km megaflashes
430 that occur in relatively active storm regions also span inactive regions where the 30-minute part
431 of the rule is tested. Meanwhile, the majority of larger megaflashes that extend over multiple
432 hundreds of kilometers violate the 30-minute part of the 30-30 rule somewhere within their
433 extent. We thus conclude that neither portion of the 30-30 rule is entirely sufficient to ensure
434 safety from megaflash cases.

435 Recent analyses of a new operational GLM “stoplight” lightning safety product by Stano
436 et al. (2019) came to the same conclusion. The GLM stoplight product was developed in

response to requests from emergency managers to have a real-time lightning product that quickly showed the spatial extent of lightning and how long ago the lightning occurred. Specifically, the stoplight product breaks the visualization into three bins of 0-10, 11-20, and 21-30 minutes. This binning was purposely selected to match with the operational rule of thumb for waiting 30 minutes after the last observation of lightning (either visually or audibly). In the course of the product development, a simple grid cell analysis was conducted for the 79 hours of available data. The available interflash times were analyzed in Stano et al. (2019) to identify how often the 30 minute time period was violated (i.e., a flash after 30 minutes in the same location). In total, 7.4 million "flash pairs" were analyzed and in 1.4% the time between two flashes exceeded 30 minutes. Another 0.4% exceeded 40 minutes. This basic analysis showed that, in a bulk sense, the 30-minute wait time is valid, but the risk is non-zero.

4 Conclusion and Discussion

In this study, we quantify the lightning hazard that is posed by megaflashes that propagate horizontally over distances of at least 100 km. These flashes are distinct from normal convective lightning that generally extend only a few kilometers horizontally. The 30-30 rule for lightning safety recommends taking shelter if lightning is followed by thunder within 30 seconds, and to remain indoors for 30 minutes after the last lightning flash has occurred. However, it is important to remember this is based on convective lightning flashes. The 30 s flash-to-bang interval equates to lightning within ~10 km from the observer – only 10% of the minimum distance covered by megaflash events.

458 By analyzing the extent of megaflash cases from GLM with the locations of ENGLN
459 strokes, we are able to demonstrate that megaflashes typically strike multiple locations along
460 their 100+ km paths. Larger flashes also have CG activity over more of their overall extents.
461 While 100-km flashes may only produce CGs over a 50 km distance, the top megaflashes
462 typically produce CGs over ~80% of the overall flash extent measured by GLM, and these CGs
463 also have higher peak currents. Since CG strikes are spread throughout a significant portion of
464 the megaflash footprint, it should not be assumed that there is a safe area below an electrified
465 cloud that is producing megaflashes – regardless of the flash-to-bang times that are measured.

466 The megaflashes observed by GLM occur primarily outside of the convective core of the
467 parent thunderstorm. Low flash rates in these regions and large distances from intense
468 convection (reaching 100 km or more) can cause the lightning hazard to be underappreciated.
469 Most megaflashes contain regions where the megaflash is the only lighting activity noted by
470 GLM in the previous hour. The second half of the 30-30 rule may be adequate for megaflashes
471 over much of their extent (especially close to the convective core), but not all locations impacted
472 by megaflash activity.

473 These results lead us to conclude that additional caution must be taken with regard to the
474 large and organized convective systems that are known to produce megaflashes – including
475 below electrified anvil and stratiform clouds. There is no true safe distance when dealing with
476 megaflashes that span hundreds of kilometers. If a thunderstorm produces one megaflash, it can
477 probably generate another that might impact different locations where lightning was previously
478 not observed. An operational meteorologist responsible for decision support services, for
479 example, could recommend the cessation of outdoor activities over a larger area than the 30-30
480 rule suggests behind a thunderstorm that has demonstrated that it is capable of generating

481 megaflashes, and that outdoor activities remain paused until the electrified stratiform cloud has
482 moved out of the area completely or the meteorological conditions are no longer favorable for
483 continued megaflash activity.

484 In light of the new megaflash measurements enabled by GLM, the time has come for the
485 community to revisit lightning safety guidance by convening a new Lightning Safety Group
486 (LSG) as described in Holle et al. (1999). This community review needs to evaluate common
487 guidance standards against emerging lightning research, new lightning detection capabilities, and
488 also updated risk analyses that have become available since 1998.

489

490 **Acknowledgments**

491 This work was supported by the US Department of Energy through the Los Alamos
492 National Laboratory (LANL) Laboratory Directed Research and Development (LDRD) program
493 under project number 20200529ECR. Los Alamos National Laboratory is operated by Triad
494 National Security, LLC, for the National Nuclear Security Administration of U.S. Department of
495 Energy (Contract No. 89233218CNA000001). The GLM LCFA and ABI data sets may be
496 obtained from NOAA via their CLASS service, which is located at https://www.class.noaa.gov/
497 (GLM: https://www.avl.class.noaa.gov/saa/products/search?datatype_family=GRGLMPROD,
498 ABI: https://www.avl.class.noaa.gov/saa/products/search?datatype_family=GRABIPRD). GLM
499 megaflash data is collected in the database in Peterson (2020b). The ENGLN data used in this
500 study were provided by Earth Networks, Inc. (https://www.earthnetworks.com/), and may be
501 ordered from them.

502

References

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.

Carey, L. D., Murphy, M. J., McCormick, T. L., & Demetriades, N. W. (2005). Lightning location relative to storm structure in a leading-line, trailing-stratiform mesoscale convective system. *Journal of Geophysical Research: Atmospheres*, *110*(D3).

Deierling, W., & Petersen, W. A. (2008). Total lightning activity as an indicator of updraft characteristics. *Journal of Geophysical Research: Atmospheres*, *113*(D16).

Deierling, W., Petersen, W. A., Latham, J., Ellis, S., & Christian, H. J. (2008). The relationship between lightning activity and ice fluxes in thunderstorms. *Journal of Geophysical Research: Atmospheres*, *113*(D15).

Ely, B. L., Orville, R. E., Lawrence, D. C., & Hodapp, C. L. (2008). Evolution of the total lightning structure in a leading-line, trailing-stratiform mesoscale convective system over Houston, Texas. *J. Gephys. Res*., **113,** doi:10.1029/2007JD008445

Franz, R. C., Nemzek, R. J., & Winckler, J. R. (1990). Television image of a large upward electrical discharge above a thunderstorm system. *Science*, *249*(4964), 48-51.

Goodman, S. J., Blakeslee, R. J., Koshak, W. J., Mach, D., Bailey, J., Buechler, D., ... & Stano, G. (2013). The GOES-R geostationary lightning mapper (GLM). *Atmospheric research*, *125*, 34-49.

Goodman, S. J., Mach, D., Koshak, W. J., & Blakeslee, R. J. (2010). GLM Lightning Cluster-Filter Algorithm (LCFA) Algorithm Theoretical Basis Document (ATBD). *NOAA NESDIS Center for Satellite Applications and Research.(Available as https://www. goes-r. gov/products/ATBDs/baseline/Lightning_v2. 0_no_color. pdf, posted 24 Sept. 2010)*.

Holle, R. L., López, R. E., & Zimmermann, C. (1999). Updated recommendations for lightning safety–1998. *Bulletin of the American Meteorological Society*, *80*(10), 2035-2042.

Lang, T. J., Rutledge, S. A., & Wiens, K. C. (2004). Origins of positive cloud-to-ground lightning flashes in the stratiform region of a mesoscale convective system. *Geophysical research letters*, *31*(10).

Lang, T., & Rutledge, S. A. (2008). Kinematic, microphysical, and electrical aspects of an asymmetric bow echo mesoscale convective system observed during STEPS. *J. Geophys. Res*., **113,** doi: 10.1029/2006JD007709

Latham, D., & Williams, E. (2001). Lightning and forest fires. In *Forest Fires* (pp. 375-418). Academic press.

548 Lopez, R. E., and R. L. Holle, 1999: The distance between suc-cessive lightning flashes. NOAA
549 Tech. Memo. ERL NSSL-1XX, National Severe Storms Laboratory, Norman, OK, 28 pp.
550 [Available online at:
551 https://repository.library.noaa.gov/view/noaa/13393/noaa_13393_DS1.pdf]
552
553 Lyons, W. A., Bruning, E. C., Warner, T. A., MacGorman, D. R., Edgington, S., Tillier, C., &
554 Mlynarczyk, J. (2020). Megaflashes: Just How Long Can a Lightning Discharge Get?.
555 *Bulletin of the American Meteorological Society*, *101*(1), E73-E86.
556
557 Lojou, J.-Y., & Cummins, K. (2005). *On the representation of two-and three dimensional total*
558 *lightning information.* Paper presented at the Proceedings, Conf. on Meteorological
559 Applications of Lightning Data, Amer. Meteorol. Soc., San Diego, CA, paper.
560
561 Lyons, W. A., Stanley, M. A., Meyer, J. D., Nelson, T. E., Rutledge, S. A., Lang, T. L., &
562 Cummer, S. A. (2009). The meteorological and electrical structure of TLE-producing
563 convective storms. In *Lightning: Principles, instruments and applications* (pp. 387-415).
564 Springer, Dordrecht.
565
566 Mach, D. M. (2020). Geostationary lightning mapper clustering algorithm stability. *Journal of*
567 *Geophysical Research: Atmospheres*, *125*(5), e2019JD031900.
568
569 Mach, D. M., Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Goodman, S. J., & Boeck, W. L.
570 (2007). Performance assessment of the optical transient detector and lightning imaging
571 sensor. *Journal of Geophysical Research: Atmospheres*, *112*(D9).
572
573 Marshall, T. C., & Rust, W. D. (1993). Two types of vertical electrical structures in stratiform
574 precipitation regions of mesoscale convective systems. *Bulletin of the American*
575 *Meteorological Society*, *74*(11), 2159-2170.
576
577 NOAA. *"When Thunder Roars, Go Indoors!" - Lightning Safety Awareness*, NOAA's National
578 Weather Service, 18 June 2018, www.weather.gov/ind/LightningSafetyAwareness.
579 Accessed 7/20/2020.
580
581 Peterson, M., & Liu, C. (2011). Global statistics of lightning in anvil and stratiform regions over
582 the tropics and subtropics observed by the Tropical Rainfall Measuring Mission. *Journal*
583 *of Geophysical Research: Atmospheres*, *116*(D23).
584
585 Peterson, M. (2019). Research applications for the Geostationary Lightning Mapper operational
586 lightning flash data product. *Journal of Geophysical Research: Atmospheres*, *124*(17-18),
587 10205-10231.
588
589 Peterson, M. (2020a). Removing solar artifacts from Geostationary Lightning Mapper data to
590 document lightning extremes. *Journal of Applied Remote Sensing*, *14*(3), 032402.
591
592 Peterson, M. (2020b), Lightning Megaflash Data. https://doi.org/10.7910/DVN/YSDLWJ,
593 Harvard Dataverse, V1

Peterson, M., & Rudlosky, S. (2019). The time evolution of optical lightning flashes. *Journal of Geophysical Research: Atmospheres*, *124*(1), 333-349.

Peterson, M., Rudlosky, S., & Deierling, W. (2017). The evolution and structure of extreme optical lightning flashes. *Journal of Geophysical Research: Atmospheres*, *122*(24), 13-370.

Peterson, M., Rudlosky, S., & Deierling, W. (2018). Mapping the lateral development of lightning flashes from orbit. *Journal of Geophysical Research: Atmospheres*, *123*(17), 9674-9687.

Peterson, M. J., Lang, T. J., Bruning, E. C., Albrecht, R., Blakeslee, R. J., Lyons, W. A., et al.. (2020a). New WMO Certified Megaflash Lightning Extremes for Flash Distance (709 km) and Duration (16.73 seconds) recorded from Space. *Geophysical*

Peterson, M., S. Rudlosky, and D. Zhang, 2020b: Thunderstorm Cloud-Type Classification from Space-Based Lightning Imagers. *Mon. Wea. Rev.*, **148**, 1891–1898, https://doi.org/10.1175/MWR-D-19-0365.1.

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.

Rutledge, S. A., & MacGorman, D. R. (1988). Cloud-to-ground lightning activity in the 10-11 June 1985 mesoscale convective system observed during the Oklahoma-Kansas PRESTORM project. *Mon. Wea. Rev.*, **116**, 1393–1408

Said, R. K., Cohen, M. B., and Inan, U. S. (2013), Highly intense lightning over the oceans: Estimated peak currents from global GLD360 observations, *J. Geophys. Res. Atmos.*, 118, 6905– 6915, doi:10.1002/jgrd.50508.

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.

Stano, G. T., Smith, M. R., & Schultz, C. J. (2019). Development and Evaluation of the GLM Stoplight Product for Lightning Safety. *Journal of Operational Meteorology*, *7*(7).

Stolzenburg, M., Marshall, T.C., Rust, W. D., & Smull, B.F. (1994). Horizontal distribution of electrical and meteorological conditions across the stratiform region of a mesoscale convective system. *Mon. Wea. Rev.*, **122,** 1777–1797

Williams, E. R. (1998). The positive charge reservoir for sprite-producing lightning. *Journal of atmospheric and solar-terrestrial physics*, *60*(7-9), 689-692.

Williams, E., Boldi, B., Matlin, A., Weber, M., Hodanish, S., Sharp, D., ... & Buechler, D. (1999). The behavior of total lightning activity in severe Florida thunderstorms. *Atmospheric Research*, *51*(3-4), 245-265.

Williams, E. R., Lyons, W. A., Hobara, Y., Mushtak, V. C., Asencio, N., Boldi, R., ... & Holzworth, R. H. (2010). Ground-based detection of sprites and their parent lightning flashes over Africa during the 2006 AMMA campaign. *Quarterly Journal of the Royal Meteorological Society*, *136*(S1), 257-271.

650 **Table 1.** General statistics describing the average number of ENGLN events in GLM
651 megaflashes and their lateral separations.
652

	GLM Flash Count	Average ENGLN Events per GLM Flash				Average CG Max Separation		Average IC Max Separation	
		All	+CGs	-CGs	ICs	Distance [km]	Percent of Flash Extent	Distance [km]	Percent of Flash Extent
GOES-16 GLM Field of View									
Full Disk	127479	5.5	1.0	4.5	17.7	51.8	37.1	75.6	56.8
ENGLN-Only Domain	80890	4.0	0.6	3.4	9.8	41.5	29.6	60.5	45.4
Subtropical Large Megaflash (300+ km) Hotspot Regions									
North America	46576	8.1	1.6	6.5	31.5	69.8	50.2	101.8	76.6
South America	50402	5.1	0.9	4.2	13.2	54.1	37.2	76.8	55.9

653
654

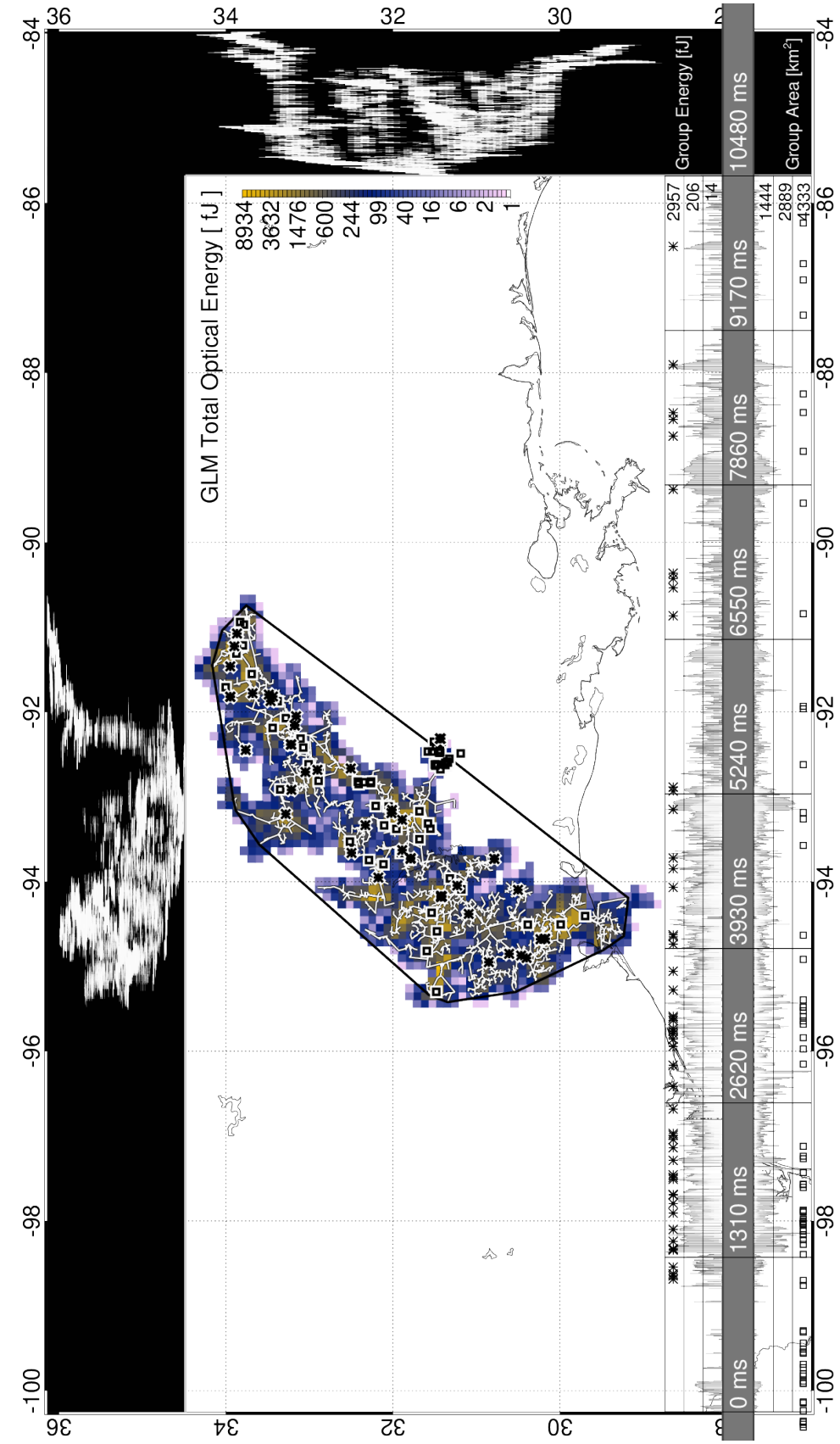

Figure 1. An example GLM megaflash with coincident ENGLN events added. GLM Total Optical Energy is mapped as a color contour. The group-level structure is overlaid with white line segments on the map, while the latitude and longitude extents of the time-ordered groups are depicted to the right of and above the map. Timeseries of group energy (above the time axis) and area (below) are shown below the map. ENGLN CG strokes are depicted as asterisks in both the map and timeseries, while IC events are drawn as box symbols.

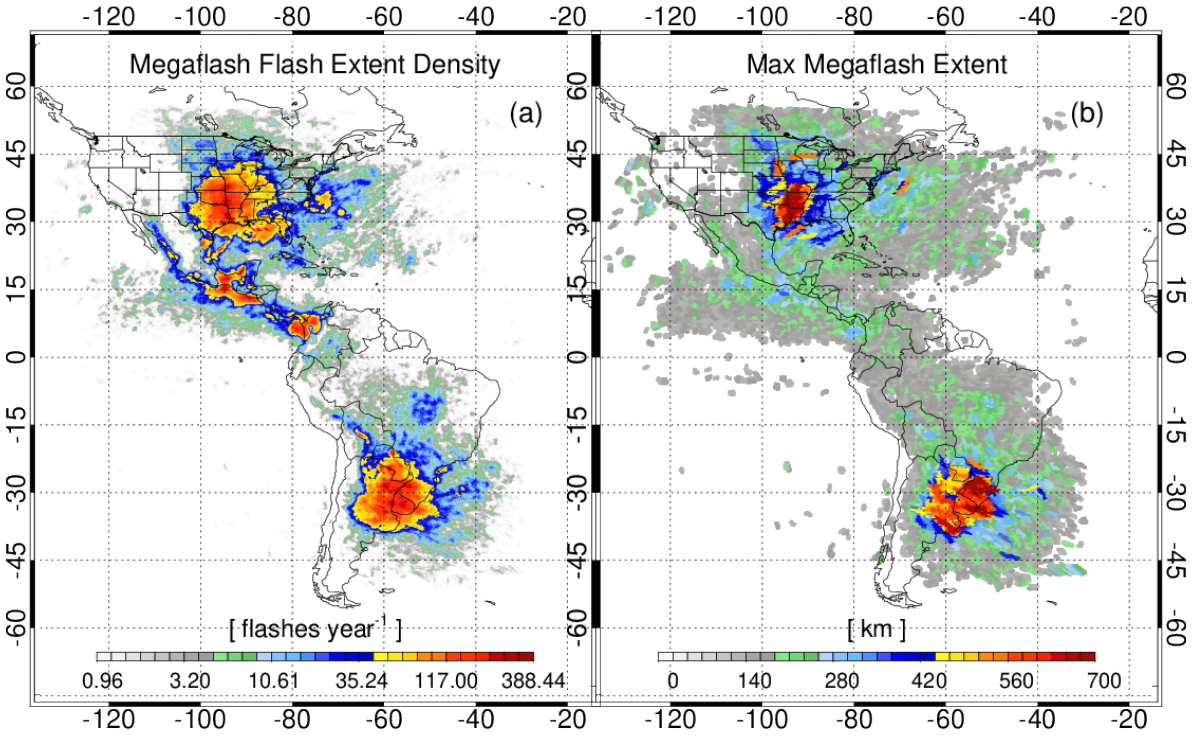

Figure 2. Hemispheric distributions of GLM megaflash frequency depicted as a Flash Extent Density (a), and peak megaflash extent (b) in the 1/1/2018-1/15/2020 record.

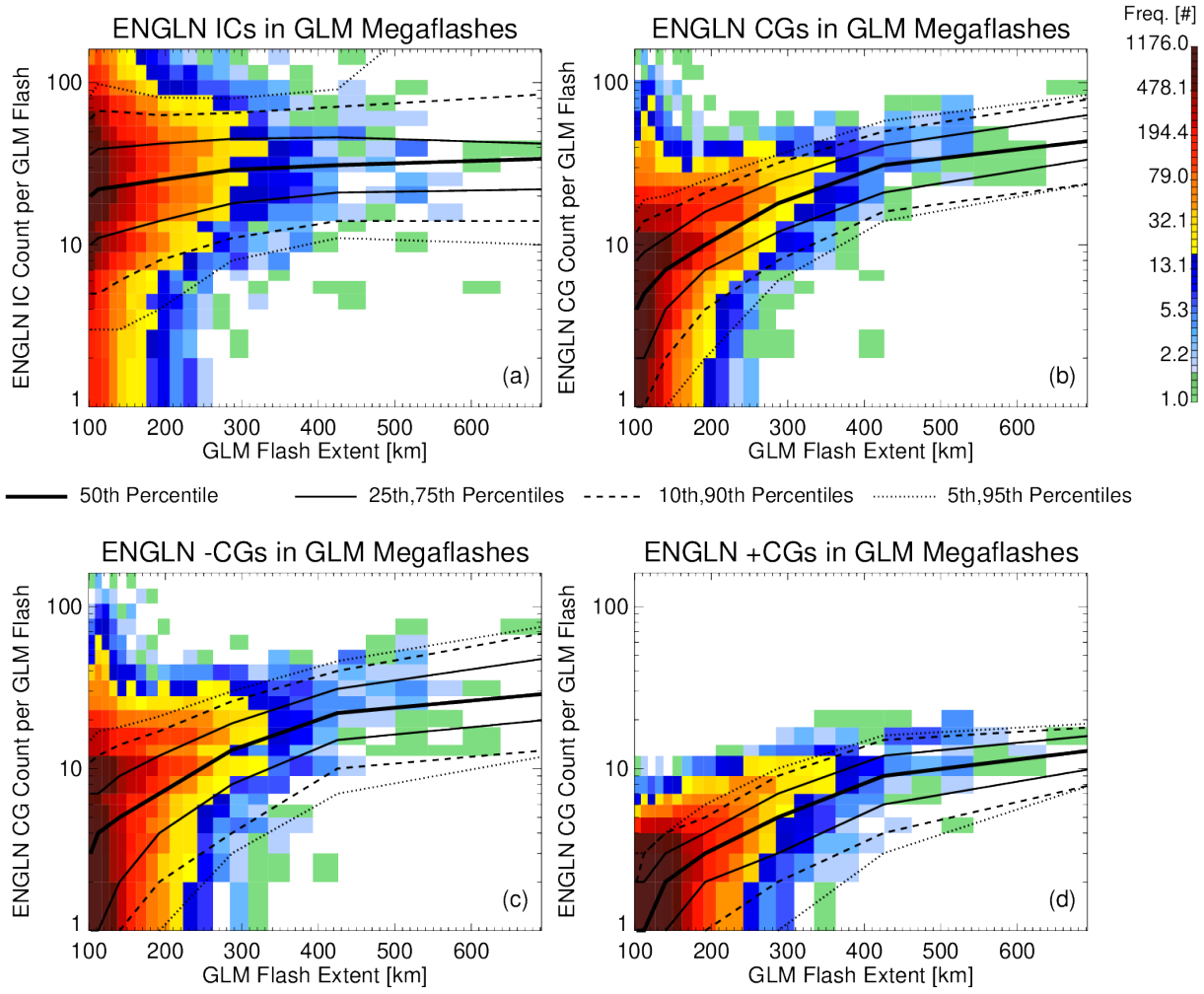

Figure 3. Two-dimensional histograms (color contours) of GLM flash extent and ENGLN (a) IC count, (b) CG count, (c) -CG count, and (d) +CG count per megaflash. CDFs are produced for GLM megaflashes of similar sizes, and the median (thick solid lines), 25th and 75th percentiles (thin solid lines), 10th and 90th percentiles (dashed lines), and 5th and 95th percentiles (dotted lines) are overlaid.

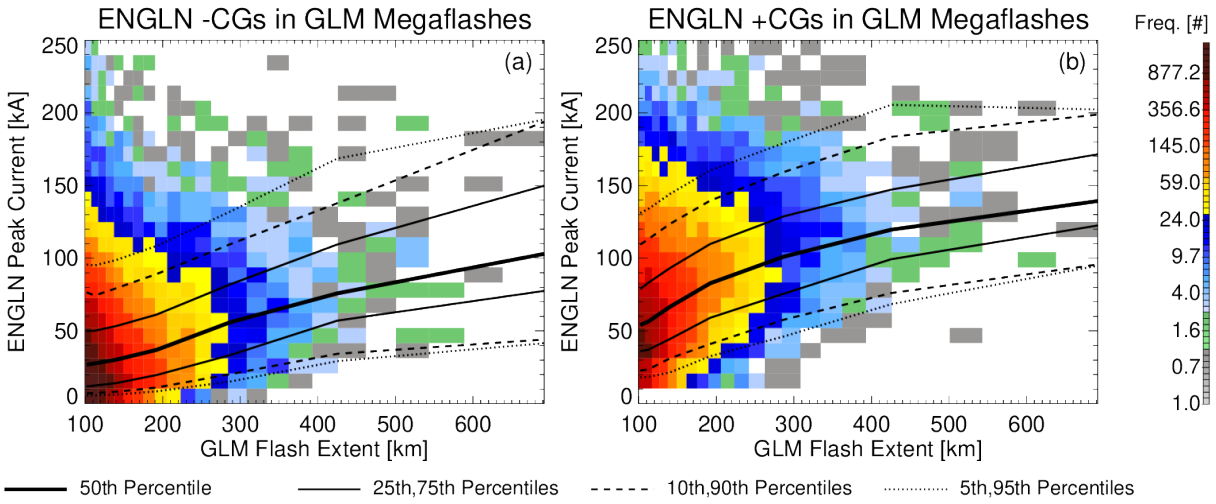

Figure 4. Two-dimensional histograms in the style of Figure 3, but between GLM flash extent and ENGLN (a) -CG peak current, and (b) +CG peak current.

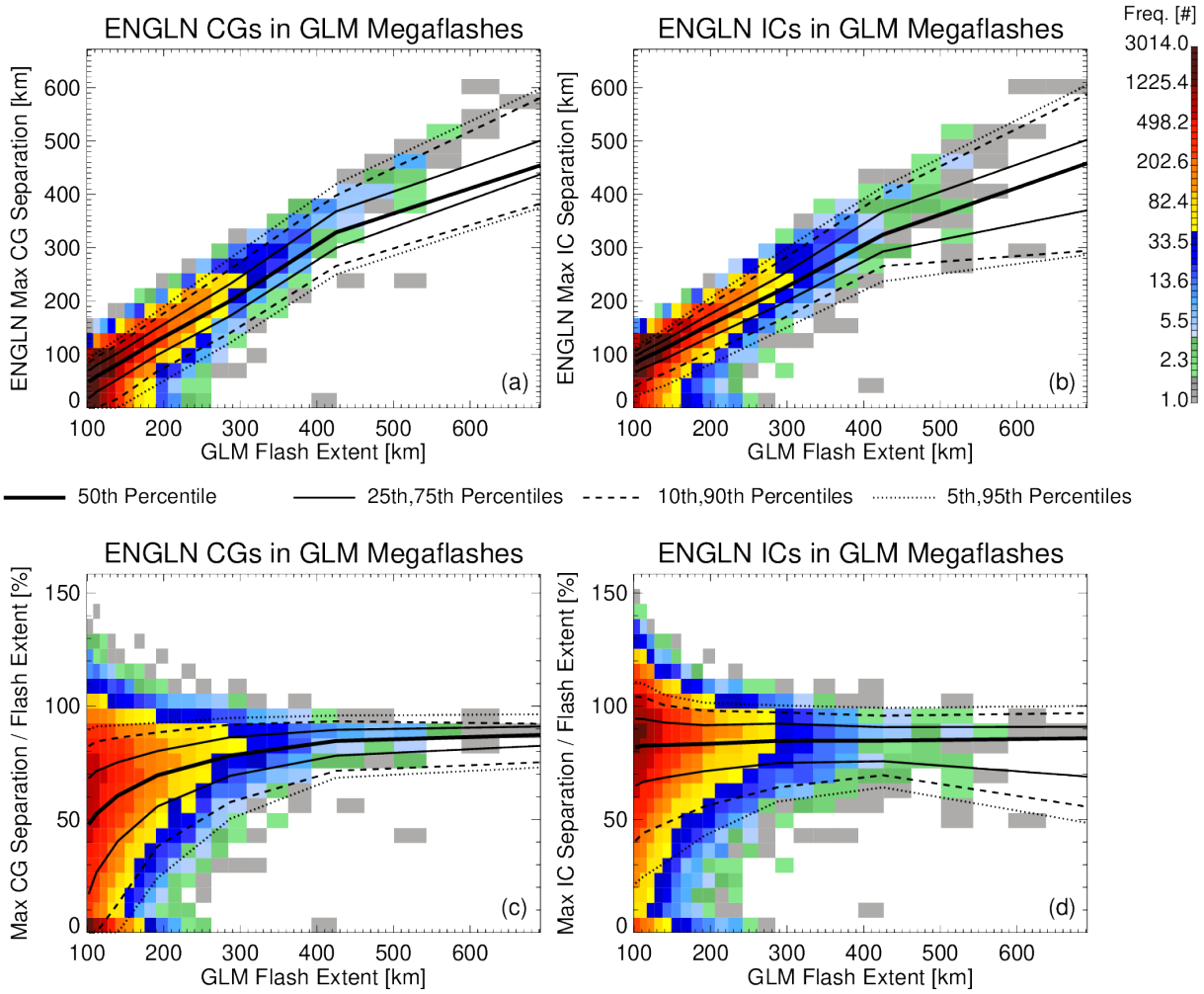

Figure 5. Two-dimensional histograms in the style of Figure 3, but between GLM flash extent and ENGLN (a) maximum CG separation distance, (b) maximum IC separation distance, (c) the maximum CG separation distance fraction of the GLM extent, and (d) maximum IC separation distance fraction of the GLM extent

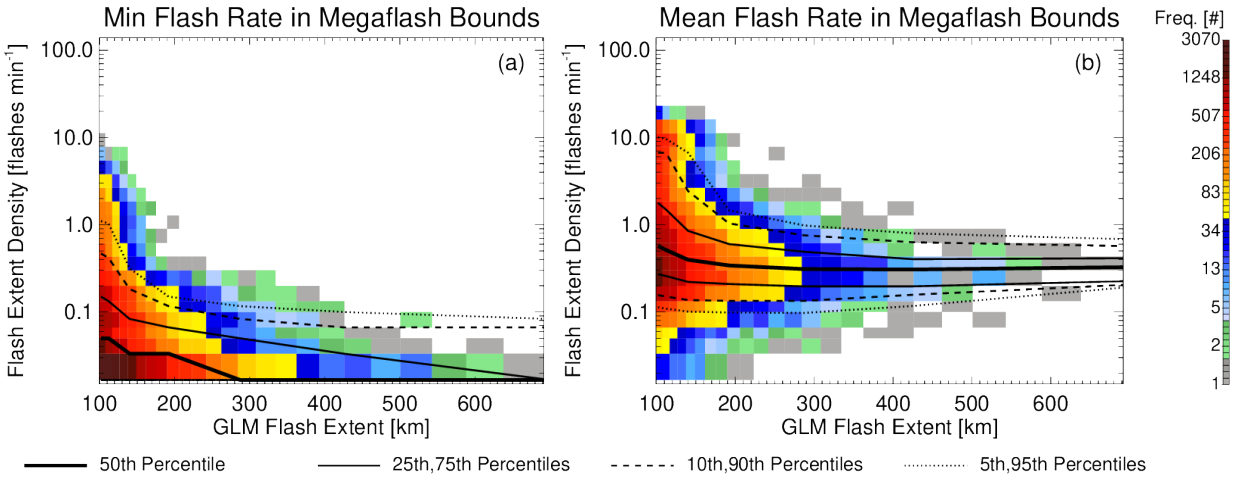

Figure 6. Two-dimensional histograms in the style of Figure 3, but between GLM flash extent and GLM (a) minimum flash rate, and (b) mean flash rate within the boundaries of each megaflash. Flash rate is quantified as a Flash Extent Density.