DISSERTATION

MICROPHYSICAL, DYNAMICAL, AND LIGHTNING PROCESSES ASSOCIATED WITH ANOMALOUS CHARGE STRUCTURES IN ISOLATED CONVECTION

Submitted by

Brody Fuchs

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2017

Doctoral Committee:

Advisor: Steven Rutledge

Brenda Dolan
Susan van den Heever
Jeffrey Pierce
Richard Eykholt

Copyright by Brody Fuchs 2017

All Rights Reserved

ABSTRACT

MICROPHYSICAL, DYNAMICAL, AND LIGHTNING PROCESSES ASSOCIATED WITH ANOMALOUS CHARGE STRUCTURES IN ISOLATED CONVECTION

Internal storm charge structures are linked to storm microphysics and dynamics. This study leverages available radar-based microphysical and dynamical information from recent field campaigns to investigate the processes that influence storm-scale charge structures. Nine normal polarity (mid-level negative charge) cases that occurred in northern Alabama, and six anomalous polarity (mid-level positive charge) cases that occurred in northeastern Colorado are studied in detail. The results suggest the presence of positively charged mid-level graupel in anomalous polarity storms, which is consistent with large amounts of supercooled liquid water (SCLW). Even though the normal polarity storms have more thermodynamic instability, the anomalous polarity storms have broader and stronger updrafts in addition to more robust mixed-phase microphysics. We expect the broader and stronger updrafts in anomalous Colorado storms are more resistant to dilution by entrainment. Using representative updraft speeds and warm cloud depths, the amount of time a parcel spends in the warm phase of a cloud was estimated for each storm observation. This metric is found to be the key discriminator between the two storm populations as the stronger updrafts and shallower warm cloud depths in Colorado lead to much shorter warm cloud residence time in those storms. We hypothesize this parameter strongly influences the amount of SCLW in the mid-levels because it impacts the loss of liquid water in the warm phase of the cloud via autoconversion and coalescence.

Using a recently developed automated flash clustering algorithm on multiple years of ground-based lightning mapping array (LMA) data, approximately 63 million lightning flashes were identified and analyzed from Washington DC, northern Alabama, and northeast Colorado. While LMA-based

average annual flash density values in Washington DC (~ 20 flashes $km^{-2} yr^{-1}$) and Alabama (~ 35 flashes $km^{-2} yr^{-1}$) are within 50% of corresponding satellite estimates, LMA-based estimates are approximately a factor of 3 larger (~ 50 flashes $km^{-2} yr^{-1}$) than satellite estimates in northeast Colorado. By estimating the initiation and propagation of lightning channels with LMA data, we find that flashes were produced at lower altitudes in Colorado, compared to Alabama or Washington DC. This is a result of the storm charge structures in these regions as normal polarity storms (common in Alabama and Washington DC) produce systematically higher altitude flashes and anomalous storms (common in Colorado) produce systematically lower altitude flashes.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Steve Rutledge, for all the support over these past 6 years. I really appreciate the balance that he is able to strike between leading us and allowing us to figure things out on our own. We have had a great working relationship throughout my time at CSU. Even though Timothy Lang left CSU before my Masters, he have still been a guiding force in my growth as a scientist and a good friend as well. I would also like to thank my committee members Drs. Brenda Dolan, Sue van den Heever, Jeffrey Pierce, and Richard Eykholt for the valuable contributions to this research, especially forcing me to consider this research from different perspectives.

Special thanks to Paul Hein for his assistance with various computer-related issues, in addition to the patience he has shown during the countless encounters we have had. The Rutledge research group, past and present: Brett Basarab, Angela Rowe, Tiffany Meyer, Doug Stolz, Nick Guy, Weixin Xu, Elizabeth Thompson, Brenda Dolan, Trent Davis, Karly Reimer, and Julie Barnum has been helpful in assisting with various aspects of this research. Assistance with LaTeX code, which was used to compile this dissertation, was provided by Chris Slocum and Brandon Wolding.

Numerous people were necessary to make this research possible because of the large amount of data in addition to the various data types. Chris Schultz and Larry Carey provided polarimetric and dual-Doppler data from the Alabama region. The efforts of Brett Basarab, Brenda Dolan and Alyssa Mathews made the synthesis of Colorado data in a reasonable amount of time possible. Lightning Mapping Array data for various regions was provided by Drs. Rich Blakeslee, Paul Krehbiel, Bill Rison, Ron Thomas, and Don MacGorman. Installation and maintenance of Lightning Mapping Arrays was provided by the folks at New Mexico Tech. Eric Burning was instrumental in the development of the original flash clustering algorithm and assisted with integration to the CSU software framework. Finally, friends and family deserve many thanks for their support throughout my entire time at CSU and indeed throughout my entire academic career. Funding for this research was provided by NSF grants AGS1010657, AGS1429925, AGS1460585 and NOAA grant NA14OAR4320125.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. STUDY MOTIVATION

It has long been established that lightning is closely linked to storm microphysics and dynamics via strong electric fields caused by charge-separating collisions of ice particles in the presence of supercooled liquid (SCLW; e.g. Workman and Reynolds 1950; Reynolds et al. 1957; Takahashi 1978; Latham 1981; Saunders and Peck 1998; Saunders 2008; Takahashi et al. 2017). Storm electrification and lightning have thus been useful tools to study microphysical and dynamical processes within thunderstorms (e.g. Carey and Buffalo 2007). Specifically, this study will focus on the processes that determine macroscale charge structures of storms, and how lightning responds to variability in storm-scale charge structure. Because lightning initiates in strong electric fields (Marshall et al. 1995; Dwyer 2005; Maggio et al. 2005), charge structures are thought to influence lightning characteristics. A large majority of thunderstorms possess a main mid-level (temperatures -10°C to -30 °C) negative charge between layers of upper- and lower-level positive charge (Krehbiel et al. 1979; Williams 1989; Mansell et al. 2005). These storms tend to produce mostly negative-polarity cloud-to-ground (-CG) flashes, meaning they transfer negative charge from the mid-levels of the cloud to the ground (Krehbiel 1986). Additionally, they produce intra-cloud (IC) flashes between mid-level negative charge and upper-level positive charge. Indeed, nearly 90% of all CG flashes in the continental United States (CONUS) are negative polarity (Orville 1994; Zajac and Rutledge 2001; Rudlosky and Fuelberg 2010). However, significant spatial variability in the percentage of positive CG (+CG%) has prompted interest in the factors that influence the production of +CGs. Specifically, +CG% is substantially higher in the High Plains region of the United States (Boccippio et al. 2001). It is hypothesized that many of these +CG flashes are produced by "anomalous" or "inverted" storms (hereafter anomalous), which possess a main mid-level positive charge region (e.g. Williams 1989; Lang et al. 2004; Wiens et al. 2005; Rust et al. 2005; Bruning et al. 2014;

Fuchs et al. 2015). Because graupel is believed to be the dominant charge carrier in the mid-levels of a thunderstorm, normal polarity storms are hypothesized to contain negatively charged mid-level graupel and anomalous storms are hypothesized to contain positively charged mid-level graupel.

Numerous laboratory studies of graupel charging have established that the charge polarity acquired by graupel upon collisions with ice crystals depends on the ambient temperature and SCLW content. At temperatures near 0 °C, graupel charges positively regardless of SCLW content. At temperatures near -20 °C, however, graupel charges negatively for modest SCLW amounts, and charges positively for high SCLW amounts. Accordingly, upper level ice crystals acquire opposite charge to the mid-level graupel. Since the mixed-phase charging zone is fixed (0 °C to -40 °C), the storm-scale charge structure depends on the mixed-phase SCLW amount (Williams et al. 2005). Logically it follows that anomalous storms should contain high SCLW amounts in the mid-levels, while normal polarity storms should contain more modest SCLW amounts. Therefore, charge structures may be used to infer SCLW amounts, and glean information about processes that may impact SCLW amounts (Lang and Rutledge 2011; Fuchs et al. 2015). In other words, charge structures can be used to investigate internal storm processes.

The approach of numerous previous studies was to relate environmental characteristics to lightning characteristics, both on a storm-by-storm basis (e.g. Smith et al. 2000; Carey and Buffalo 2007) and a climatological basis (Carey and Rutledge 2003; Williams et al. 2005). The physical underpinnings of this method are as follows: the environment influences microphysical and dynamical processes within a storm, which influence charging processes and macroscale charge structures, which in turn determine the lightning characteristics. Essentially, these studies were just working backwards. With lightning characteristics (such as +CG%), the macroscale charge structure can be inferred, and therefore the SCLW can be inferred. This process is illustrated schematically in Figure 1.1. Despite the research that has been conducted on the variability of storm-scale charge structures, our understanding remains incomplete. Many of the hypotheses about the connections between different processes are physically based, but many of them have not been tested to this point because of inadequate observations and datasets, including our inability to measure microphysical and charge structures in-situ, especially in strong convection.

The availability of polarimetric and dual-Doppler observations during recent field experiments provides an opportunity to investigate storm microphysics and dynamics in concert with LMA-based lightning information. However, merging multiple data types for numerous storms quickly becomes overwhelming. A method must be developed to synthesize the wealth of microphysical and dynamical information from numerous storms into meaningful insights. The approach of this study is to more directly evaluate hypotheses about the links between environment, microphysics, dynamics, charge structures, and lightning (Figure 1.1) by using newly developed automated tools and data processing. These tools allow us to both increase the detail with which we investigate storms, in addition to increasing sample size, leading to more robust insights. The tools used in this study are detailed in the next section.

1.2. Analysis Tools

1.2.1. *Flash clustering algorithm*

Total lightning flash rates (IC+CG) are a critical piece of information that is closely related to storm dynamics (Deierling and Petersen 2008). The first tool developed during this study is the flash-clustering algorithm of lightning mapping array (LMA) data. LMA networks are comprised of multiple stations (6 or more) and use time-of-arrival (TOA) techniques to locate bursts of very high frequency (VHF; 60-66 MHz) radiation produced by the discontinuous breakdown of lightning (Rison et al. 1999; Thomas et al. 2001, 2004). LMAs detect numerous VHF bursts (also called sources) in each flash, and can therefore accurately map flashes that are within approximately 100-150 km of the network center (Fuchs et al. 2016). Note that these networks have limited spatial coverage, but have nearly 100% detection efficiency, a feature unique to LMA networks. Because the number of sources from LMA data can be on the order of millions, most LMA-based analysis has been on a single flash or a single storm. With our development of an automated algorithm that clusters VHF sources in space and time into flashes, we are able to do analysis on large numbers of flashes on a storm scale as well as on a climatological scale (Chapter 2).

After VHF sources are clustered into flashes, total flash rates in addition to information about each flash such as flash size and location become available. We are then able to construct flash rate climatologies to understand long term patterns in lightning activity. Furthermore, LMA-derived flash rates can be compared with flash rates from other methods of lightning detection (e.g., satellite, Chapter 2). Interesting results were found from comparisons of annual flash densities between LMAs and satellites. Differences were found to be region and charge-structure dependent. We also develop a novel approach to mapping the three-dimensional structure of lightning channel initiation and propagation in Chapter 3 to address some unresolved questions from Chapter 2.

1.2.2. *Charge structure inference*

After a lightning flash is initiated, leaders of opposite polarity propagate into regions of charged hydrometeors (Coleman et al. 2003). Interestingly, the negative breakdown branch into positive charge tends to be more discontinuous, and therefore produce more VHF radiation than positive breakdown in regions of negative charge (Rison et al. 1999). We use this to our advantage by integrating VHF radiation from all flashes and inspecting the vertical distribution of VHF source density. We infer that the dominant positive charge resides near the mode of the vertical VHF source distribution (Wiens et al. 2005; Lang and Rutledge 2011). Since normal polarity storms typically contain an upper-level dominant positive charge region, LMA mode temperatures are usually around -40 °C, while anomalous storms with dominant mid-level positive charge typically have LMA mode temperatures around -20 °C (Wiens et al. 2005; Fuchs et al. 2015). This charge inference method is preferred to the +CG fraction method because many storms of interest in this study do not produce appreciable amounts of CGs, which makes this method of inference rather noisy and unreliable.

1.2.3. *CLEAR*

Many previous studies on charge structure variability have relied on a few case studies to understand the processes associated with anomalously electrified storms. In this study, we use the CSU Lightning, Environmental, Aerosol, Radar (CLEAR) framework (Lang and Rutledge 2011; Fuchs et al. 2015). CLEAR was designed as an objective and automated case study framework that identifies cells based on contiguous regions of a certain quantity (reflectivity in this case) and then attributes generic data types to cells based on spatial and temporal matching. This results in a statistical database of storms that can be queried for certain characteristics (such as inferred charge structure; Chapter 4). Even though LMA networks have limited coverage, we can accumulate a large number of cases as long as a continuous record of radar data is available.

1.3. Dissertation Outline

The goal of this dissertation is to use LMA (lightning) data along with available polarimetric and dual-Doppler data to understand how lightning characteristics respond to charge structure variability in addition to the mechanisms that are associated with anomalously electrified storms. The dissertation is comprised of three journal articles and will be laid out as follows. Chapter 2 establishes the flash clustering algorithm is titled “Climatological analyses of LMA data with an open-source lightning flash-clustering algorithm” and is published in the Journal of Geophysical Research: Atmospheres. Chapter 3 addresses some hypotheses put forth in Chapter 2 and is titled “Estimating lightning flash locations in isolated convective thunderstorms using LMA observations”. It is currently in review in the Journal of Geophysical Research: Atmospheres. Chapter 4 is a deep-dive into the microphysical and dynamical processes that influence storm-scale charge structure and is titled “Microphysical and dynamical processes associated with anomalous charge structures in isolated convection” and is currently in review in the Journal of Geophysical Research: Atmospheres. Finally, Chapter 5 will summarize some of the key findings of each chapter and offer some ideas to further the research.

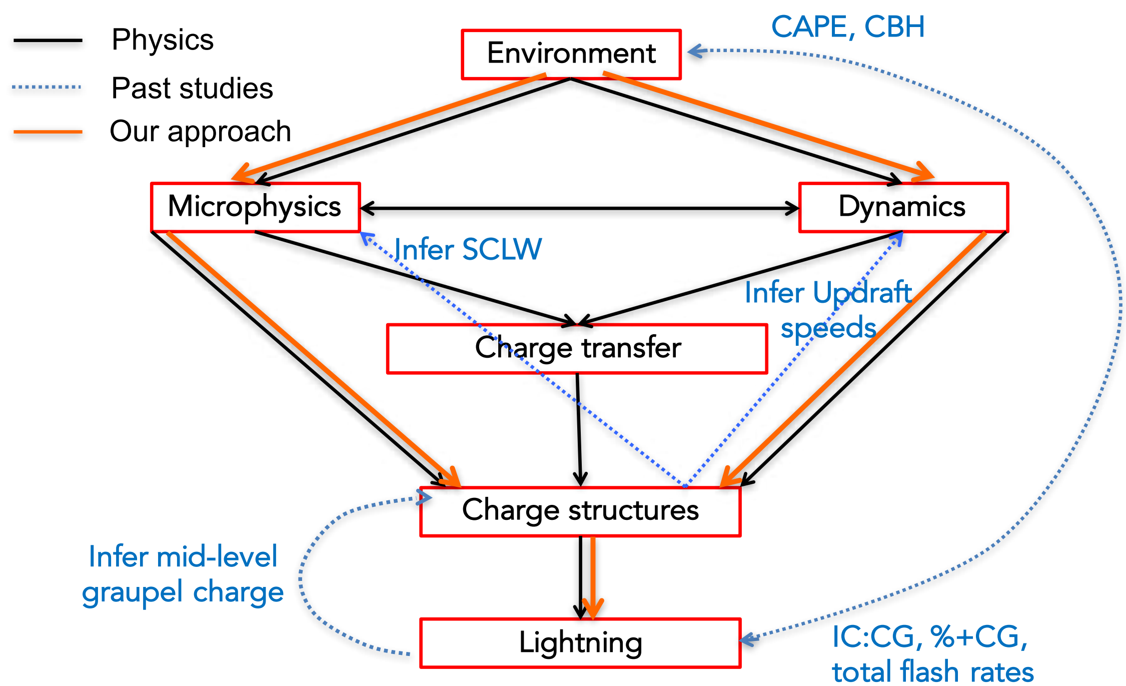

FIG. 1.1. Schematic of the physical interactions between storm environment, microphysics, dynamics, and electrification. Blue dashed lines indicate the methods of past studies. Orange line indicate relationships tested in this study.

CHAPTER 2

Climatological analyses of LMA data with an open-source lightning flash-clustering algorithm

2.1. Introduction

2.1.1. *Lightning processes*

A lightning flash is a continuous plasma channel that develops bi-directionally away from a region containing a strong electric field (Maggio et al. 2005) and propagates into potential energy wells (Coleman et al. 2003). Along this plasma channel, breakdown processes and currents produce electromagnetic radiation across a wide spectrum of frequencies. In focusing on the plasma channel and the processes along it, we are consistent with the flash definition that is used by ground- and space-based instruments (Christian et al. 1999, 2003; Cummins and Murphy 2009). Regardless of the detection method, the flash is the highest-level grouping of detected processes to emanate from a connected set of local breakdown processes. The large electric fields present in thunderstorms are hypothesized to be produced by charge-separating collisions between precipitation ice (graupel and hail) and ice crystals in the presence of supercooled liquid water (e.g., Takahashi 1978; Saunders et al. 1991). Accordingly, lightning is tightly coupled to the microphysical and dynamical processes within a storm (e.g., Williams 1985; Carey and Buffalo 2007; Fuchs et al. 2015; Schultz et al. 2015). Indeed, considerable evidence has been shown that lightning flash rates are strongly tied to updraft speeds, for example, (Deierling and Petersen 2008) and ice water contents (Petersen et al. 2005) in the mixed-phase region ($\sim$ 0 to -40 °C). The connection between lightning flash characteristics and storm physics can be used to study processes not observable by other methods. The LMA-based flash rate algorithm described in this study provides characteristics of flashes, such as size and location in addition to total flash counts to facilitate investigation of more storm processes.

This chapter is published in the Journal of Geophysical Research: Atmospheres as Fuchs et al. (2016).

2.1.2. *Lightning mapping array (LMA) detection of lightning*

The advent of LMA networks has provided unprecedented detail of lightning channels on the sub-storm scale. LMAs use time-of-arrival (TOA) techniques from multiple stations to detect the time and location of very high frequency (VHF) radiation bursts (or "sources") produced by the discontinuous propagation during breakdown that forms lightning channels (Rison et al. 1999). Radiation is detected in an unused local television channel frequency (usually $\sim$ 60-66 MHz or 5 m wavelength) to avoid contamination from local noise. There may be a few to a few thousand VHF sources reported by the LMA for a single lightning flash, depending on its spatial extent, its proximity to the network and network detection efficiency (Fuchs et al. 2015). Detection of VHF sources along lightning channels results in highly accurate three-dimensional (3D) mapping of the in-cloud portions of intra-cloud (IC) and cloud-to-ground (CG) flashes. See Figure 1 from Thomas et al. (2004) for an example of the detail that can be detected by an LMA. Sub-flash processes, such as the bidirectional nature of lightning breakdown can also be observed with an LMA (Behnke et al. 2005; Maggio et al. 2005; Velde and Montanyà 2013). Such detail has facilitated investigations of storm processes, such as cloud microphysics and turbulence (Bruning et al. 2007; Bruning and MacGorman 2013) in addition to macroscale electrical characteristics of storms. Examples of such investigations include the vertical charge structure and total lightning activity in thunderstorms occurring in geographic diverse regions in the United States (e.g. Lang and Rutledge 2011; Fuchs et al. 2015).

2.1.3. *LMA source clustering*

Since LMA networks only detect VHF sources produced by flash propagation, processing of those sources is required to get higher-level information about the overall flash. These algorithms are known as flash counting or flash clustering algorithms. Multiple flash clustering algorithms currently exist and all perform the same basic function of clustering VHF sources by space and time into their parent flashes (e.g. Thomas et al. 2003; MacGorman et al. 2008; McCaul et al. 2009). The grouping of sources is possible because the spatial and temporal separation between successive sources within a single flash is much smaller than the separation of sources in different flashes in all but the most extreme flash rate storms (see Figure 3 in Fuchs et al. 2015). This paper utilizes the open-source flash-clustering algorithm first discussed in Fuchs et al. (2015). The algorithm produces not only total flash rates, but also calculates characteristics about each flash such as location (initiation and centroid) and the plan-position area of each flash. The plan-position area of a flash can be thought of as the area enclosed by a rubber band wrapped around the flash viewed from above (Bruning and MacGorman 2013).

This algorithm currently performs the similar basic functionality as other flash clustering algorithms. In fact, flash rate differences between the XLMA algorithm (Thomas et al. 2003) and this algorithm have been shown to be consistently within 10-15% of each other (Fuchs et al. 2015). However, this algorithm is able to efficiently process large amounts of LMA data to produce climatological-scale results, such as those shown in this paper and is conducive to real-time use. The results in the paper elucidate the strengths and weaknesses of this algorithm, provide some insight about the ways the algorithm may be improved and raise some points about lightning physics. Since the algorithm is open source, anyone may contribute to the improvement of the algorithm.

2.1.4. *LMA-detected lightning in the context of other datasets*

Lightning flashes produce radiation across a large range of frequencies, each with distinct propagation characteristics (Shumpert et al. 1982; Budden 1988; Cummins and Murphy 2009; Nag et al. 2015). Since various lightning detection systems detect radiation from different portions of the spectrum, they are subject to their own strengths and weaknesses. Lower frequencies (from the very low frequency (VLF) to high frequency (HF) bands), subject to relatively little attenuation, are utilized by long-range detection systems such as Global Lightning Detection 360 network (GLD360; Said et al. 2013) and Earth Networks Total Lightning Network (ENTLN; Rudlosky 2015). These systems are able to continuously monitor lightning activity across the globe but are only able to detect strong flashes (typically CGs), since radiation from those processes is at a lower frequency and typically travels long distances before reaching detectors. The National Lightning Detection Network (NLDN; Cummins and Murphy 2009) is able to continuously detect weaker lightning flashes at the cost of spatial coverage and can discriminate between IC and CG flashes, although the uncertainty in the classification depends on several factors (Nag et al. 2013). Optical lightning detectors aboard satellites, such as the Lightning Imaging Sensor (LIS; Christian et al. 1999 and the Optical Transient Detector (OTD; Christian et al. 2003) are able to provide information about global lightning distributions via snapshots of storms, but are unable to provide temporal evolution of individual storms and are best suited for climatological studies. It is important to note that this limitation will be remedied with the launch of the Geostationary Lightning Mapper (GLM) aboard the Geostationary Earth Observing System R series satellite (GOES-R). GLM will continuously detect optical emission from lightning on a hemispheric scale from approximately 55°N to 55°S (Goodman et al. 2013).

2.1.5. *Differences between LMA and satellite flash data*

Lightning flash climatologies from the LMA provide some valuable information when compared to corresponding climatologies constructed from satellite-based LIS and OTD optical sensors. Some differences are inevitable because of the differences between detection methods and data collection characteristics of the two systems. These differences must be kept in mind when comparing these datasets, and may actually provide useful information about the flashes themselves when assessing any differences between flash climatologies.

LMA networks detect lightning fundamentally different from satellites. Satellite sensors use a charge-coupled device (CCD) to detect near-optical (777.4 nm) wavelength emissions from channel heating associated with oxygen excitation and relaxation (Christian et al. 1999, 2003). Conversely, the LMA detects VHF (5 m wavelength) radiation produced by discontinuous lightning channel propagation (Rison et al. 1999). Multiple sensors detect each source of radiation, and a TOA method is used to estimate the location and time of each radiation source. Optical radiation is subject to larger rates of attenuation when propagating through a thunderstorm, which may make it more difficult to detect flashes with certain characteristics such as those that do not produce much light.

In addition to the detection frequencies, the data collection characteristics are also quite different between the two systems. LMAs are stationary and always on, so they continuously map the channels that comprise each lightning flash, permitting temporal analysis of storm evolution and climatological analysis. Since LMAs continuously detect lightning, they sample a distribution of storm intensities and lifecycle phases as each storm initiates, matures and then eventually decays. Given a long enough record, LMAs should sample the entire spectrum of storms types and lifecycle phases to produce a representative measure of lightning characteristics near the network. Conversely, satellites capture short snapshots of lightning in a storm, and may capture multiple storms during an overpass of a particular region. All of the satellite overpasses from multiple years are aggregated to produce an estimate of total flash density in a grid box. Presumably all the storms sampled during an overpass will not have identical intensities and will be in different phases of their lifecycle. Therefore, we argue that it is reasonable to assume the long record of overpasses should sample the full spectrum of storm intensities, from weak to the very intense. The end result is similar to the continuous observations of storms by an LMA network.

Another factor that warrants consideration is that the operational period of the LMA networks is not concurrent with the operational period of the satellites. As a result, the climatologies constructed by both instruments were not sampling the same population of storms. This is of particular concern in the Washington DC and Colorado regions, which were observed by the OTD (north of $\sim 38°$ latitude), which flew from 1995 to 2000. It is possible that a small sample of storms could bias the results in these regions. The Alabama region is at lower latitude than the other regions and was observed by the LIS instrument (south of $\sim 38°$ latitude), which was operational for much longer than the OTD (1997- 2014).

All of these caveats must be taken into account when attempting to compare climatologies constructed from LMA and satellite systems.

2.1.6. *Overview of paper*

This study describes the development of the open-source flash-clustering algorithm and how the algorithm groups LMA sources into flashes. We will then demonstrate use of the algorithm in the form of lightning flash climatologies for the Washington DC, northern Alabama, and northeast Colorado LMA networks. These networks are located in environmentally distinct locations (Williams et al. 2005; Fuchs et al. 2015), which provide a wide range of storm intensities. In addition to flash densities, we will examine flash characteristics such as flash area, initiation height, duration and average power, and the sensitivity of these quantities to network detection characteristics. This will allow us to fully investigate the flash clustering algorithm and the detection performance of the LMA networks included in this study. Flash density climatologies will be compared to climatologies from satellite-based optical detectors (LIS and OTD) and implications of the similarities and differences will be explored.

2.2. Algorithm Description

2.2.1. *Clustering method*

This study uses the flash algorithm introduced (but only briefly described) by Fuchs et al. (2015). This algorithm is similar to other published flash algorithms (Thomas et al. 2003; Wiens et al. 2005; MacGorman et al. 2008; McCaul et al. 2009) in that it uses spatial and temporal thresholds to cluster individual VHF sources into flashes. The major differences with this algorithm are the ability to process large amounts of data and that it is open source. Here we further describe how the algorithm groups sources into flashes and some of the implementation details.

LMA data are reported as tuples of time, latitude, longitude, and altitude coordinates. Each source has an associated chi-squared metric, which describes the goodness-of-fit of the overdetermined solution for source location and time in addition to the number of stations that contributed to the location retrieval (Thomas et al. 2004). Initial filtering is performed to discard solutions (LMA sources) detected by fewer than six stations (by default) or a chi-squared value greater than 1.0. This is in an attempt to filter out noise (i.e. VHF sources not produced by lightning). Sensitivity studies (not shown) for a few select storms indicate that the flash counts and characteristics are not very sensitive to the choice of the chi-squared threshold.

At its core, the clustering algorithm uses the Density-Based Spatial Clustering of Applications with Noise (DBSCAN; Ester et al. 1996) algorithm, which is a general machine-learning clustering algorithm that assumes no a priori cluster shape. DBSCAN emulates the space and time thresholding behavior of other flash sorting algorithms, but is distinct because it searches for clusters of high density LMA sources in a four-dimensional space-time matrix (Pedregosa et al. 2011), rather than searching for the mere presence of LMA sources that are in close proximity to each other (Fuchs et al. 2015). In order to use the DBSCAN algorithm implemented in the Python scikit-learn package (Pedregosa et al. 2011), it is necessary to form a unified space-time coordinate vector that is searched for high-density clusters.

DBSCAN begins by picking random sources to start searching for clusters (or flashes). Each flash begins with a core LMA source surrounded by at least a specified number of points, each at a distance less than or equal to a specified maximum distance threshold. In the limit of the lowest possible density, a flash is set of points at a distance exactly equal to the specified distance from the core source (which is also the flash centroid) and distributed uniformly in all directions. A linear segment of points belonging to a flash would have an average spacing somewhat less than the specified distance. Connecting sets of points that meet the above criteria forms the flash. Therefore, the DBSCAN algorithm operates much like a simple space-time thresholding algorithm for randomly distributed points, but places a more stringent requirement on peripheral points. Compared to a simple dot-to-dot algorithm, which can connect a long string of points together as long as one other point is under the threshold distance, DBSCAN requires an outer point to be near another point which itself must be within the threshold distance of at least the minimum specified number of other sources.

Since the separation in both distance and time between sources in a flash must be considered simultaneously, a transformation must be made to make the units compatible for use in the DBSCAN algorithm. The space-time coordinate data is normalized with the Weighted Euclidean Distance used by Mach et al. (2007). After the spatial coordinates are converted to an earth-centered Cartesian coordinate system, the location of the LMA coordinate center is subtracted. Dividing the spatial dimensions by a prescribed distance value and the time dimension by a specified time value normalizes the coordinates. The maximum distance value must be large enough to include the nominal distance between VHF sources produced in a flash, while not being too large to group nearby flashes together. In Fuchs et al. (2015), the normalization distance value was chosen to be 3 km for the Colorado network and 6 km for the less sensitive Alabama and DC networks, in accordance with values from other algorithms (MacGorman et al. 2008; McCaul et al. 2009). The time coordinate is normalized by 150 milliseconds. Unlike McCaul et al. (2009) the algorithm does not use a distance-dependent normalization value that loosens the clustering criteria for sources far from the network, where source location errors are a significant concern (Thomas et al. 2004). Doing so attempts to compensate for missed sources with low powers that may result in missed flashes. We decided not to employ this adaptive thresholding for this paper because we want to provide a base performance for the algorithm before introducing any uncertainties associated with adaptive thresholding. Furthermore, investigating the impacts of detection efficiency on flash characteristics also allows us to characterize the performance of the LMAs included in this study.

LMA data rates may be on the order of 10^5 sources min^{-1}. The DBSCAN implementation in scikit-learn uses a pairwise distance matrix to determine clustering, which is an O(N^2) operation, making it prohibitive to process an entire thunderstorm at once. To work around this limitation, we take advantage of the fact that individual lightning flashes are time-limited and impose (by default) an arbitrary 3-second maximum duration on the lightning flashes. Impacts of this threshold will be explored in the results and discussion sections. Processing is then accomplished in a streaming mode, where a VHF source buffer whose time span is twice the maximum possible flash duration ensures that flashes less than or equal to the maximum duration are not split (Figure 2.1). After DBSCAN identifies clusters within the buffer, those clusters with a source in the first half of the buffer are pruned out and further processed as flashes. Some sources will remain in the second half of the buffer, and these are retained for processing with a new half-buffer of sources. This streaming process results in significant speed increase. As an example, the algorithm is fast enough to run in real time in the Colorado region, where LMA source rates from ~ 1000 flashes can be on the order of 1 million sources in a matter of 5 minutes.

2.2.2. *Clustering parameters*

DBSCAN cluster formation is controlled by two parameters: a Euclidean distance threshold ε and the minimum number of points N_{min} required for a cluster to not be considered noise. Because the coordinate vector has been normalized to the specified distance and time values, we set ε equal to 1.0, since ε is essentially a multiplicative constant for the space and time thresholds. In previous studies, N_{min} has ranged from 2 to 10, depending on the sensitivity of the particular LMA network (McCaul et al. 2009; Fuchs et al. 2015). For this study, N_{min} is set to 2 for Alabama and DC, and set to 10 for Colorado. While singleton points and small clusters of less than N_{min} points (noise in DBSCAN's terminology) are not valid DBSCAN clusters, they are still removed from the first half of the processing buffer and retained in the final output data as small flashes. Later, when calculating flash rates, it is customary to filter these small flashes out by once again thresholding valid flashes on their number of points.

Similar to other existing algorithms, this algorithm is sensitive to the choices of clustering parameters. Larger space and time thresholds used in normalization may result in more sources being included in flashes or incorrectly grouping two nearby flashes together into one larger flash. Conversely, smaller space and time thresholds may incorrectly result in breaking a large flash into multiple smaller flashes. N_{min} may affect the number of flashes identified in addition to the characteristics of flashes, such as size and duration. Smaller values of N_{min} will permit smaller clusters that may not be physical flashes, while a larger N_{min} will be more restrictive but may miss smaller flashes that have fewer sources, particularly for flashes that are far from the network. Differences in clustering parameters are mandated by differences in network sensitivities. The Alabama and DC networks are less sensitive than the Colorado network and therefore detect fewer low-power sources, resulting in fewer total detected sources and potentially missed flashes. To compensate for the lower detection efficiency of the Alabama and DC networks, a larger spatial threshold is employed. This attempts to compensate for sources produced by lightning channels that are missed and may result in a breakup of a flash because source distances are too far apart.

2.2.3. *Differences with existing algorithms*

The algorithm and instructions for use are currently available on github. Since it is open-source, anyone may download the package and contribute to it. The package is modular and permits other flash clustering algorithms to be added and used. Additionally, it has been incorporated into an automated framework by Fuchs et al. (2015) and can process many LMA files at a time to produce large-scale results, such as those shown in this study. However, one of the goals of this paper is to show the strengths and weaknesses of this algorithm and outline potential ways that the algorithm may be improved.

By adopting an open-source trans-disciplinary machine-learning package, we gain the benefit of rapid dissemination of bug fixes identified by a much larger community. This reduces the software maintenance burden on the lightning community, whose fundamental software task is reduced to lightning-specific preparation of the data for use in a standardized algorithm.

2.3. Results

2.3.1. *Washington D.C. region*

2.3.1.1. *Spatial flash variability*

The algorithm was applied to eight years (2007-2014) of LMA VHF source data for the Washington DC region. The network was comprised of 8 stations before 2009 and 10 stations after 2009. The addition of stations likely increased the detection of efficiency of the LMA. However, quantifying this impact is difficult because the storms were not identical before and after the addition. Unlike the other LMA networks included in this study, the DC LMA is configured to detect radiation of slightly higher frequency (local Ch. 10 vs. Ch. 4). This is important to note because lightning produces lower intensity radiation at higher frequencies, which are closer to the inherent noise floors of the stations, making source detection slightly more difficult by this network. Approximately 10 million flashes were gridded on a 0.15° latitude x 0.2° longitude grid (to make them approximately square). Conclusions reached in this study were not dependent on grid box size. We are assuming that the LMA detection efficiency is 100% over the network for all flashes, including CGs. This is because there is no currently accepted detection efficiency behavior for LMA networks. The implications of this assumption should be kept in mind when analyzing the results.

We not only wanted to investigate raw flash counts (and densities), which can be compared to other measurements such as satellite observations, but also characteristics of the flashes themselves. Figure 2.2 shows maps of VHF LMA source density, total lightning flash density, plan-position flash area and vertical distribution of flash initiation heights along with an azimuthally integrated relationship as a function of distance from the center of the network. The VHF source density in Figure 2.2a has a maximum near the center of the network with generally decreasing values farther from the network center. LMA source detection is line-of-sight, so sources that are close to the network are less likely to be blocked by the curvature of the Earth. Additionally, VHF sources are assumed to radiate isotropically

(Rison et al. 1999). Consequently, low-power sources are less likely to be detected if they are far from the network because the power falls off according to the inverse-square law.

Figure 2.2b shows a map of flash density processed by the algorithm. A pattern similar to the LMA source densities consisting of a maximum near the center of the LMA with decreasing values with increasing distance from the network is evident. The average flash density values within 50 km of the network are $\sim$ 18 flashes $km^{-2} yr^{-1}$ (Figures 2.2b, 2.2e). The decrease in flash density with increasing distance from the network is slower than the LMA source density, as shown in Figs. 2.2b, 2.2e. The LMA source density falls off to half of its maximum value at approximately 50 km from the center of the network while the flash density falls off to half its maximum value at approximately 125 km.

To investigate the characteristics of flashes themselves, Figure 2.2c shows a map of the median plan flash size for all flashes in a spatial bin. Similar to the VHF source density map, median flash size decreases with range from the maximum value. However, the location of maximum flash area is offset northeast of the network center towards the Baltimore area.

The vertical distribution of flash initiation height locations is shown in Figure 2.2d with respect to the distance from the network center. The flash initiation density maximum around 10 km MSL is consistent with IC flashes that initiate between mid-level negative charge and upper-level positive charge in normal polarity storms common in this region (Zajac and Rutledge 2001; Fuchs et al. 2015). The average flash initiation heights of these flashes, shown in Figure 2.2e, increase from 9 km over the network to 13.5 km at 300 km from the center of the network. This follows from the expected R^2 source height location errors outside of an LMA network, where R is the distance from the network center to the source (Thomas et al. 2004). The relative minimum around 7 km MSL is consistent with the main mid-level negative charge region, which is not a likely initiation point for lightning flashes (MacGorman et al. 2001), because it is a potential well, regardless of charge polarity (Maggio et al. 2005). The relative maximum around 5 km is consistent with the large electric field between mid-level negative charge and lower-level positive charge in normal polarity storms (Figure 2.2d). The lack of flashes at low altitudes and longer range is due in part to Earth's curvature blocking line of sight for those particular sources.

2.3.1.2. *Range sensitivities of flashes*

Decreasing source detection efficiency impacts the algorithm-derived flash characteristics. If sources produced by a flash are missed, an incomplete picture of the flash results. Implications of this are explored in Figure 2.3 for the DC region. First, Figure 2.3a shows distributions of flash durations, partitioned by distance from the LMA center. Flash duration is defined here as the time difference between the first and last source attributed to a flash. A consistent decrease in flash duration is apparent as distance increases from the LMA center. This is likely due to undetected sources that are occurring at the beginning or the end of the flash, thus shortening the duration of the flash. It is important to note that for flashes within 50 km of the network where values are most likely to be representative of the true flashes, the distribution of durations is nearly lognormal (note the log scale on the vertical axis) with a median around 200 milliseconds. Additionally, less than 3% of flashes had more than a 1 second duration, suggesting the 3 second cutoff in flash processes is more than sufficient.

Distributions of plan-position flash areas calculated by the algorithm are shown in Figure 2.3b. Similar to the behavior of flash durations, the algorithm-calculated flash areas decrease with range from the network. Just as undetected sources near the beginning or end of the flash can shorten the duration, undetected sources near the spatial boundary of the flash will result in a calculated area that is smaller than the physical flash. It should be noted that the algorithm assigns an area of zero if there are only 2 sources in a detected flash.

The distributions of the average source power in a flash shown in Figure 2.3c exhibit the opposite trend of previous panels. Average flash power increases monotonically with distance from the network. This is consistent with the inverse-square law, as weaker sources are more likely to go undetected if they are farther from the network, resulting in detection of only the strongest source powers in flashes far from the network. Conversely, the network is able to detect weaker sources inside of the network, which is why the median of the average source powers is approximately 0 dBW.

Figure 2.3d shows the distributions of the number of detected source points per flash. The median values decrease rapidly with increasing distance from the network, from approximately 20 sources per flash within the network to 2 at 250 km from the network. The lines on Figure 2.3d indicate the two thresholds at 2 and 10 points. It is clear that the flash count values are highly sensitive to the choice of N_{min}, particularly for flashes far from the network. There appears to be no best choice for this threshold. The choice of 2 minimum points may erroneously classify flashes simply because 2 VHF sources may be coincidentally adjacent, however a strict threshold of 10 would result in missing nearly all of the flashes at 250 km from the center of the network. Because of this, flash rates and quantities should be met with a great deal of skepticism at these long ranges from the network center. It should be noted, however, that the strict chi-squared source detection parameter should filter out much of the noise sources and mitigate identification of erroneous flashes.

Figure 2.3e shows a map of the number of lightning hours for the DC region. If one flash occurs in a grid box within a certain hour, it is counted as one lightning hour. This can be used as a measure of how common thunderstorms and lightning are in the DC region, and can be compared to other regions in the study. Over the center of the network, there are approximately 80 lightning hours, with decreasing values with distance from the network. The spatial variation in the lightning hour field is relatively smooth, much more so than the flash density map.

The combination of flash density and lightning hour maps provides another useful piece of information. Figure 2.3f shows a map of the average number of lightning flashes that occur in a grid box during an hour that at least one flash occurs. This metric can be thought of as an average grid box total flash rate, as the units are flashes hour^{-1}, however is not strictly a cell flash rate because multiple storms may occur in a grid box during a given hour. Regardless, this metric gives an indication of the intensity of the lightning-producing storms in this region. The values near the network are relatively uniform, ranging from 50-70 flashes hour^{-1}. The slower trend of decreasing values with range is not surprising when compared to flash densities, considering that only one flash in a grid box is needed to trigger a lightning hour.

2.3.1.3. *NLDN CG lightning*

NLDN CG flash densities are subtracted from the total flash densities provided by the algorithm to estimate IC flash densities in each grid box. Bulk total and CG flash density quantities are subtracted in each grid box, following the methodology of Boccippio et al. (2001). Recall that we are making the assumption of 100% flash detection efficiency with the LMA. Only NLDN CG flashes, subject to peak current filtering (Cummins and Murphy 2009) that occurred during the time period of 2007-2014 are included in this study to ensure the same storms are being included in this analysis. The detection efficiency of the NLDN network is above 95% throughout the DC domain. The detection characteristics of the NLDN may have been altered after its upgrade in 2013 (Nag et al. 2013), however we assumed 100% detection efficiency for the NLDN for the entire observation period, since it will not change any of the conclusions in this paper.

First, Figure 2.4a shows the NLDN CG density in units of flashes $km^{-2} yr^{-1}$. The significant spatial variation in CG activity is evident. Specifically, the highest values lie along the Atlantic coast and near Baltimore. Values decrease both inland towards the higher terrain, and offshore to the east. Figure 2.4b shows a map of the ratio of the estimated IC flash rate (based on LMA total flash rate and NLDN CG flash rate) to the NLDN CG flash rate. The maximum values are near 5-6 in the Baltimore area, with generally decreasing values farther from the network. This decrease is not surprising, since flashes are being missed by the LMA and clustering algorithm at long range, but are still being detected by the NLDN, which has little variability in detection efficiency over the range of the LMA network (Cummins and Murphy 2009). The IC:CG values are most trustworthy near and inside the network (approximately 50 km), where LMA detection efficiency is the highest. Accordingly, the differences between IC:CG values shown here are 50-100% higher than in Boccippio et al. (2001) and more recently Medici et al. (2017), which used satellite-based optical sensors for total flash rate estimates. Implications for these differences are explored in the discussion section.

2.3.2. *Northern Alabama*

2.3.2.1. *Spatial flash variability*

The northern Alabama LMA network covers portions of Alabama, Mississippi, Georgia and Tennessee. The network consists of 12 stations in northern Alabama centered near Huntsville, and 2 stations near Atlanta, Georgia. Approximately 43 million flashes from 7 years of LMA data (2008-2014) are shown here. The LMA source density map in Figure 2.5a shows that the maximum is located near the center of the network with a rapid decrease outside of the network, similar to the DC LMA. However, the magnitudes of the maximum LMA source density values are approximately a factor of 3 larger than in the DC region. The preference for sources on the south and west side of the network is apparent. Detection artifacts due to asymmetries in station locations are likely not the cause (Thomas et al. 2004; Koshak et al. 2004), since station locations are approximately uniform.

The maximum in source density near Nashville, Tennessee is not physical, however. Further analysis (not shown) indicates that a VHF emitter in the LMA frequency window is responsible for the additional sources there. An attempt to remove these sources is difficult, if not impossible, without a large radar dataset (to determine the presence of storms), and is beyond the scope of the present study. Additionally, the characteristics of the sources (such as altitude and power) were not substantially different from sources produced by lightning (not shown).

A map of total flash density is shown in Figure 2.5b. The average value within the LMA network is around 35 flashes $km^{-2} yr^{-1}$, much higher than the values in the DC region. The bias of lightning activity to the south and west is also evident in the total flash density map in Figure 2.5b, providing some evidence that the source density variability is indeed a physical signal, rather than a detection artifact of the LMA. Lower flash densities over the elevated terrain in the eastern portion of the network suggest that the terrain may have an impact on storms in the area. Since the terrain only rises to approximately 300 m above the altitude of Huntsville, it is unlikely that any line-of-sight blockage preventing LMA detection of VHF radiation is occurring. The decrease in flash density with range in the Alabama network is much more gradual than the DC network, as the azimuthally integrated flash density falls to half of its maximum value around 225 km, much longer than the 125 km in the DC region. This is likely due in part to the greater number of stations in the Alabama network (12) compared to DC in addition to the greater sensitivity as a result of lower noise floors. Note that the same point with high source densities near Nashville results in high flash densities in that same area. This implies that the production of sources by the anomalous emitter is close enough in space ($\sim$ 1 km) and time ($\sim$ 0.1 s) to be erroneously clustered into a flash by the algorithm.

The median plan flash areas map in Figure 2.5c and the azimuthally integrated values in Figure 2.5e shows an unexpected result. The median flash area peaks around 50 km from the center of the network, which is located just outside of the network radius. Median flash areas in that region are approximately 50% larger than over the center of the network. Unlike flash density, no significant flash area variability with azimuth is observed. Possible explanations for this will be explored in the discussion section.

The vertical distribution of flash initiation altitudes in the Alabama region shows a similar trend to the DC region. Most flashes initiate around 9-10 km MSL, likely between the upper-level positive charge region and the mid-level negative charge region of a normal polarity storm. Indeed, Fuchs et al. (2015) found that the overwhelming majority of storms in this region had inferred positive charge at temperatures colder than -40 °C, consistent with normal polarity storms containing positively charged ice crystals in the upper portions of the cloud (Williams 1985). The detected initiation altitude increases with distance from the network, as the mode height increases from 10 km over the network to 14 km at 300 km from the network due to source height location errors (Thomas et al. 2004). Detected flashes at lower altitudes (< 5 km MSL) are relatively rare, particularly at ranges greater than 200 km because those are below the horizon due to Earth's curvature.

2.3.2.2. *Range sensitivities of flashes*

Similar to the DC region, flashes are composited and analyzed by their distance from the center of the network to investigate the impacts of detection characteristics on the flashes themselves. The flash durations in Figure 2.6a show that algorithm's representation of flashes tends to become shorter for flashes far from the LMA, similar to the DC region. Note that the distribution of flash durations inside of 50 km is quite similar to corresponding flashes in the DC region. Undetected sources near the beginning or the end of the physical flash will effectively shorten the detected flash. Similar reasoning is consistent with the general decrease in flash area with increasing distance from the network in Figure 2.6b. The plan-position area estimate is sensitive to sources near the periphery of a flash. If those sources are not detected, the area estimate is erroneously low. This effect is particularly apparent when comparing the 95th percentiles of flash area for each distribution. It is important to note that the largest median area corresponds to flashes 50 - 75 km from the center of the network. Inspecting the 75th and 95th percentiles of flash area around 50-75 km show that fewer flashes have very large area, suggesting that some detail of the flashes may be lost at that range, assuming no substantial difference in physical flash properties. However, inspecting the 25th percentile of flash sizes at that range shows there are fewer small flashes, suggesting the smallest flashes are undetected at that range, thereby increasing the median flash area. The decrease in the 25th percentile of flash sizes at longer ranges is likely owed to undetected sources in a flash, artificially decreasing the calculated flash size.

The distribution of average flash powers (Figure 2.6c) follows a similar trend to the DC region. Lower average flash powers inside and near the network result from the ability of the LMA to detect weaker sources at close range. The average flash powers near the network in Alabama are much higher

($\sim$ 10 dBW) than in the DC region ($\sim$ 0 dBW). Given the lower noise floors and greater number of stations, the lower source powers in DC seem counterintuitive, but the difference can likely be explained by the different frequencies used by the networks. Recall that lightning produces lower power radiation at higher frequencies. The number of points per flash in Figure 2.6d supports the notion of the more sensitive Alabama network. While the points per flash follows the same trend as the DC region, the median values are much higher than in the DC region. For flashes between 225 and 250 km from the Alabama network, approximately 20% of the flashes are made up of at least 10 points, compared to approximately 1% of similar flashes in the DC region. Similar to the DC network, flash rates and quantities at this long range should be met with skepticism and should be analyzed to ensure their quality.

The number of lightning hours in each grid box is shown in Figure 2.6e, and indicates a different thunderstorm environment than in the DC region. With values above 200 lightning hours per year inside 50 km, this region is an active region for thunderstorm development and persistence. Figure 2.6f shows that the average grid box flash rates during a lightning hour in Alabama are not substantially higher than the DC region. This suggests that there are more numerous lightning-producing storms in Alabama than in DC, rather than more intense thunderstorms producing more lightning per storm. These results agree with the cell-based flash rates from Fuchs et al. (2015), which showed similar cell flash rates between the Alabama and DC regions. The maximum in lightning hours (Figure 2.6e) and relative minimum in grid total average flash rate (Figure 2.6f) over the network is a surprising feature which suggests that spurious clusters of points are being considered a flash by the algorithm. However, this effect appears to be limited to the interior of the network, a small portion of the whole domain, where network sensitivity is the highest.

2.3.2.3. *NLDN CG lightning*

Figure 2.7a shows the distribution of peak-current filtered NLDN CG flash density for the Alabama region. The westward preference for lightning production is also evident here, as northeastern Mississippi and northwestern Alabama have substantially higher CG flash densities than the eastern side of the network where the high terrain is located. Figure 2.7b shows that the average value of IC:CG within the network is around 5, nearly twice as high as previous estimates from Boccippio et al. (2001). The IC:CG should be most trustworthy inside of the network (50 km), as the highest flash detection efficiency is located there. Note the highest IC:CG values to the southeast of Huntsville over a region of high terrain known as Sand Mountain. This location corresponds to a relative minimum in NLDN CG density, while no comparable drop is observed in total flash density.

2.3.3. *Northeast Colorado*

2.3.3.1. *Spatial climatologies*

The northeast Colorado LMA network was installed during the winter of 2011-2012 for the DC3 field campaign (Barth et al. 2015). The network consists of 15 stations ranging from locations just northeast of Denver up to the Wyoming and Nebraska borders. Currently, it is the newest of the LMA networks in the United States and has the lowest noise floor of any network, due to the updated electronics and remote locations of most stations. The diameter of the network is approximately 90 km, also the largest of any network to date. This is expected to further increase the detection efficiency (Rison et al. 1999; Thomas et al. 2004). However, data during most of 2013 were not used since the network only had 5-6 active stations operating then due to various technical difficulties. For this reason, results in this section include approximately 10 million flashes from 2012 and 2014.

The LMA source density map in Figure 2.8a shows much larger magnitudes and spatial variability than either of the other networks. Complex topography in the region and a small sample size of 2 years are likely significant contributors to the variability of the LMA source density map. Unlike the other networks, the maximum VHF source density is not located within the network, but rather northeast of the network near Fort Morgan, about 70 miles northeast of Denver. A sharp gradient in LMA source density is present in the foothills of the Rocky Mountains as well. Rapid decrease in LMA source density is observed toward the east around 150 km from the center of the network, farther than the other networks.

The flash density map in Figure 2.8b exhibits some different behaviors than the source density map. The magnitudes of flash density are striking, with values as high as 80 flashes $km^{-2}yr^{-1}$ in some areas. The average flash density inside the network (50-55 flashes $km^{-2}yr^{-1}$) is approximately 3 times larger than in the DC network and almost twice as large as the Alabama network.

Concerning the spatial distribution of flash densities, the values are much more homogenous east of the Rocky Mountains, while lightning activity falls off rapidly towards the west in the Rocky Mountains. This is also shown in Figure 2.8e, as the flashes in the eastern portion of the domain are indicated separately. The large gradient is likely due to weaker storm intensities in the foothills where thermodynamic environments are not conducive to intense convection, relative to the adjacent foothills where instability and cloud base heights are higher (Williams et al. 2005; Jirak and Cotton 2006; Fuchs et al. 2015). A relative maximum in lightning density is present east of Denver on the north side of the Palmer Divide. This region is hypothesized to be a favorable location for storm initiation due to the Denver Convergence and Vorticity Zone (DCVZ; Crook et al. 1990). More years of data should help verify the spatial patterns in flash density, as spatial noise from a low number of storm days will be averaged out. Note that with the strict clustering thresholds employed in this region, we suspect that these flash density quantities may actually be a lower bound on the estimate of flash densities, particularly far from the network.

A drastically different signal in median plan flash area is observed in the Colorado region compared to the other regions, as the larger flash areas over the Rocky Mountains dominate the signal. The sharp longitudinal gradient in flash area is coincident with both the flash density gradient and the elevation gradient. Put another way, there are fewer but larger flashes over the Rocky Mountains compared to the adjacent plains of Eastern Colorado. This signal is consistent with the claims made by Bruning and MacGorman (2013) that weaker storms with less turbulence produce fewer but spatially larger flashes. They hypothesize that the flash size spectrum is strongly related to the amount of turbulence in a storm, because larger charge reservoirs build up in weaker storms that result in large flashes when breakdown is initiated. This is in contrast to strong storms with intense vertical motions that produce numerous pockets of charge that result in frequent but smaller flashes (Bruning and MacGorman 2013; Basarab et al. 2015).

Perhaps a less obvious, but nonetheless surprising signal is apparent in the data if only flashes that occurred in the eastern half of the domain are considered (dotted lines in Figure 2.8e). The median flash size increases from the center of the network to 125 km, after which the median flashes decreases monotonically, a trend similar to the Alabama network. The flashes in the eastern half of the domain are treated separately in this case to remove the terrain and meteorological effects of the Rocky Mountains to the west.

The vertical distribution of flash initiation points is also substantially different than the other regions. The maximum in flash initiation density is located around 7 km, coincident with the location of the relative minimum in the other regions. This implies that the vertical distribution of electric fields is drastically different in Colorado storms compared to Alabama or DC storms. Indeed, Fuchs et al. (2015) found that storms in Colorado exhibited different charge structures than storms in Alabama or DC. Storms in Colorado were much more likely to possess an anomalous charge structure, characterized by strong low- or mid-level positive charge when compared to a normal charge structure, which is characterized by strong mid-level negative and upper-level positive charge. It is then possible that the locations of strongest electric fields are located at different heights, as Figure 2.8d suggests. The rapid decrease in flash densities past 100 km is due in large part to the presence of the Rocky Mountains and the lack of flashes there. Inspecting Figure 2.8e shows that flash densities do not drop off until roughly

200 km from the center of the network for flashes in the eastern portion of the domain. The effect of height location errors can be observed in the increasing flash heights between 200 and 300 km from the network (Figure 2.8d).

2.3.3.2. *Range sensitivity of flashes*

In an effort to remove the effects of terrain on the lightning characteristics, only flashes within the eastern half of the Colorado LMA domain were grouped based on their distance from the LMA center and analyzed to investigate the impacts that detection variations have on calculated flash characteristics. In Figure 2.9a, it is immediately apparent that the sensitivity of flash characteristics to network proximity is markedly different from the other networks. The median duration holds steady around 200 milliseconds, regardless of distance from the LMA, out to 250 km. The consistency of flash duration with range is evidence of the network's high detection efficiency. It should be noted, however, that the distributions of flash durations within the network are remarkably similar to corresponding flashes in DC and Alabama, respectively.

A similar signature is present in Figure 2.9b, which shows the plan-position flash areas with respect to range from the LMA. The median flash areas increase from 5 km^2 inside of 25 km up to 12 km^2 at 150 km and decreases to 9 km^2 at 250 km. It is important to note here that the median flash area is actually larger at 250 km than inside of 25 km. Since this analysis is restricted to the flashes in the eastern portion of the domain for reasons discussed earlier, this provides strong evidence for the high sensitivity of the Colorado network. There is a maximum in median flash area at moderate ranges from the LMA center, similar to Alabama. However, looking at various percentiles of the distribution reveals a different behavior than the Alabama network. All major percentiles of flash area increase from the center of the network out to approximately 125 km, suggesting that detected flashes may be larger coincident with smaller flashes going undetected. Given the complex terrain in this region, it is reasonable to suspect that physical flash characteristics may vary with terrain and that may be influencing these relationships. However, it is extremely difficult to decouple those potential impacts from detection artifacts.

The distribution of average flash powers in Figure 2.9c reveals a consistent behavior for all ranges from the LMA center. The consistency provides some evidence of the high detection efficiency of the CO LMA. The 5th percentile of all distributions is above 5 dBW, indicating that few flashes have low average power.

The number of points per flash in the Colorado region (Figure 2.9d) is much larger than either the Alabama or DC regions. For flashes within the network, the median number of points per flash is over 100, or approximately a factor of 3 more than in DC and a factor of 2 than in Alabama. Nonetheless, the decrease in sensitivity at long range is evident as the median number of points monotonically decreases to approximately 30 points per flash at 250 km, still greater than any distribution of flashes in Alabama or DC. It is possible that physical flashes are being missed at distances around 250 km due to the strict 10-point threshold in this region, implying that these flash density values may be a lower bound on the real flash density value.

The number of thunderstorm hours in the Colorado region, shown in Figure 2.9e, reveals some undeniable spatial patterns. Immediately evident is the maximum in the sharp elevation gradient in foothills of the Rocky Mountains, with an eastward extension on the Palmer Lake Divide, south of Denver. Somewhat surprising are the values both over the network, and the location of maximum flash density north and east of Denver. The number of lightning hours in these areas is approximately 90-100 per year, similar values to those in the DC region and about a factor of 2 lower than Alabama. Since the values of flash density are much larger in Colorado than the other regions, it follows that more lightning is produced per storm. This is borne out in Figure 2.9f, which shows average grid total flash rates as high as 180, roughly a factor of 3 larger than either Alabama or DC. This is in accordance with the Fuchs et al. (2015) hypothesis that the unique thermodynamic ingredients in this region conspire to produce intense, highly electrified storms in Colorado. Note that the foothills of the Rocky Mountains, location of the lightning hours maximum, have a very low average grid total flash rate.

2.3.3.3. *NLDN CG lightning*

The local terrain undoubtedly influences the production of CG flashes in the Colorado region. Studies with longer datasets have shown a longitudinal gradient of CG flash density along the foothills of the Rocky Mountains in addition to a latitudinal gradient in the adjacent plains (Zajac and Rutledge 2001; Orville and Huffines 2001; Vogt and Hodanish 2014). Specifically, CG density maxima are present on the relatively higher terrain of the Palmer Divide and the Cheyenne Ridge (1.8 km above sea level) and a minimum in the Platte River valley (1.5 km above sea level). The limited temporal extent of this dataset shows similar features, with a maximum of CG activity to the east of Denver on the north side of the Palmer Divide on Figure 2.10a.

The large total flash density values in this region result in much higher IC:CG values than previous studies (e.g. Boccippio et al. 2001). Maximum IC:CG values on Figure 2.10b approach 25 near the network and in the Fort Morgan area. These values are striking, considering they are approximately a factor of 3 higher than Boccippio et al. (2001). Even though CG production is substantially lower in the Rocky Mountains compared to the adjacent plains, IC:CG values are also much lower. This is in accordance with the much lower total flash rates and larger flash sizes, hypothesized to result from different storm environments and weaker vertical motions in storms over the Rocky Mountains.

The numbers here likely represent a lower bound on the actual IC:CG values, since we are assuming that all CG flashes are detected by the LMA. CG flashes typically produce very little VHF radiation near the ground, however there is usually plenty of in-cloud activity associated with any CG flash. For this study, we have made no adjustments to the total flash counts and assume that the detection efficiency of the LMA is 100% because no published detection efficiencies currently exist. This is obviously not correct, but should be close for flashes within the network (50 km from the network center).

2.3.4. *Flash merging*

In an effort to test and improve the performance of the algorithm, a simple flash merging model was implemented. The goal of this was to mitigate the errors produced by the algorithm separating branches of one larger physical flash into multiple smaller flashes. Separation of branches may result in erroneously high values of flash rates and densities, while producing erroneous flash characteristics as well. The simple model implemented here tests for nearby flashes that are sufficiently close in time. More specifically, if the initial point of flash is within 150 milliseconds and 3 km from any point in any other flash, that flash is considered as a branch to the earlier nearby flash by the model. This model was tested on isolated cells with maximum reflectivity above 55 dBZ (to select the strongest storms with the most lightning) from the storm database produced by Fuchs et al. (2015). Figures 2.11a,c,e show the total number of flashes in each cell and the number of merged branches in Alabama, DC and Colorado, respectively. The merged branches can be thought of as the decreased total flash number for each cell. A linear least-squares best-fit slope is also indicated in each panel. The slope represents the percentage decrease of total flashes as a result of the leader merging. The average decrease in flash number is 4% for Alabama storms, 1% for DC storms and 8% for Colorado storms. That the effect is largest in Colorado is somewhat surprising, since it is the most sensitive network in this study. A possible explanation is that the greater network sensitivity in Colorado resolves lightning channels more fully, making them easier to meet the proximity criteria to other parts of the flash. Conversely, it would be more difficult to meet the proximity criteria if networks with lower sensitivity are not resolving parts of the lightning channels.

To assess the dependence of the simple flash-merging model on the detection efficiency, storms were partitioned by distance from the network center and a similar least-squares fit was performed for each population of cells. Figures 2.11b,d,f show the slope of the linear fit as a function of distance from the network center storms in Alabama, DC and Colorado, respectively. There is a general decreasing trend with increasing distance from the center of the network, particularly in DC and Colorado, meaning that fewer leaders are merged to parent flashes at longer ranges. This may be the result of decreasing detection sensitivity resulting in branches not meeting proximity criteria at longer ranges, similar to the larger percentage of leaders merged in Colorado, compared to the other regions. Slope sensitivities to particular cells are indicated by the thickness of the lines using the jackknife resampling method (Efron and Efron 1982). There is more sensitivity to individual storms at short range because the cell sample size is smaller, in large part because area of an annulus depends on its radii. Assuming a constant density of storms, there will be fewer storms near the network center, compared to farther from the network.

2.4. Summary and discussion

2.4.1. *Flash algorithm*

This paper has described a new open-source flash-clustering algorithm designed to perform the same basic function of spatial and temporal VHF source grouping as other flash clustering algorithms. One of the key differences with this open-source algorithm is that it facilitates collaboration and continuous improvements by the larger lightning community. By virtue of the algorithm's ability to automatically process large amounts of LMA data, this study is the first of its kind to provide climatological scale analyses on millions of LMA-detected lightning flashes in multiple regions of the United States with distinct environmental characteristics. These results highlight some of the differences between the LMA networks in Alabama, DC and Colorado. Some of the strengths and weaknesses of the algorithm were elucidated, which may steer future improvements to the algorithm.

2.4.2. *Flash and detection characteristics*

Comparisons between flash distributions partitioned by distance from the LMA showed some differences between the different networks. Flash characteristics in DC exhibited strong sensitivity to proximity to the network center, suggesting that an appreciable number of sources go undetected. The proximal dependence of flash characteristics is weaker in Alabama, but sensitivities still exist and may bias estimates of flash area integrated over a storm. Flash characteristics in the Colorado region were shown to be much less sensitive to network proximity. This is expected because the Colorado network has the most stations and lowest noise floors of any permanent network currently in operation.

It is these differences in network proximity sensitivity between networks that highlight some of the shortcomings of the algorithm and differences in network detection efficiencies. There are a couple of factors that contribute to network sensitivity. First, since VHF radiation sources radiate isotropically, the power flux decreases rapidly with distance, resulting in undetected sources (particularly weak ones) at long distances. This is an unavoidable problem determined by the laws of physics and the noise floor of a particular network.

Errors of source location estimation increase rapidly with distance outside of the network due to the geometry of intersecting hyperbolas (Thomas et al. 2004). This may result in VHF sources being scattered in space and time, to the point where sources are no longer close enough to meet the clustering thresholds. A potential remedy to this problem is to loosen the clustering thresholds with increasing distance, known as "adaptive thresholding" in McCaul et al. (2009). This would likely result in more detected flashes by the algorithm. The effect on flash characteristics is not immediately obvious, and would require a comparison of flash characteristics before and after the adaptive thresholding is applied. Additionally, there are actually competing effects on the flash area calculation: the source errors and the undetected sources, the former increasing the calculated area and the latter decreasing the calculated area.

Incorrectly classifying a branch of a lightning flash as a separate flash has been a known problem in flash clustering algorithms (Thomas et al. 2003; MacGorman et al. 2008). If a branch is delayed by a time interval after the start of the parent flash, it may be identified as a separate flash, thereby incorrectly inflating the number of flashes in storm. This phenomenon needs to be understood in a quantitative manner to ensure that the flash counts and characteristics are representative of the physical flashes that are being detected. A simple model for merging branches to their parent flashes was attempted in this study. Figure 2.11 shows that the effects separating leaders was minimal in most cases. Storms in Colorado were most sensitive to the branch-merging model, however flash counts were decreased by more than 10% in an extremely small number of cases.

For all the observed differences in network sensitivities and consequently, flash quantities between the regions, some similarities also arise. The distributions of flash durations are remarkably similar between regions, particularly if only considering flashes within each network. In DC, the 5th, 50th and 95th percentiles of flash duration are approximately 25, 200 and 500 milliseconds, respectively. In Alabama, the 5th, 50th and 95th percentiles of flash duration are approximately 25, 200 and 700 milliseconds, respectively. In the plains of Colorado (the eastern half of the network), the 5th, 50th and 95th percentiles of flash duration are approximately 50, 200 and 500 milliseconds, respectively. Recall that we are only considering the eastern half of the Colorado domain to mitigate terrain effects. The distributions of plan-position flash area are likewise similar in all three regions. The similarities of calculated lightning flash quantities provide some confidence that the algorithm is performing as expected. The differences in flash power and number of flash points rely on the detection of LMA sources, and are consequently going to have differences based on the ability of each network to detect VHF radiation produced by lightning. It is encouraging to see that variations in flash characteristics far from the network are smooth which suggests that those flashes with few points are indeed flashes and not random noise being clustered into a flash.

2.4.3. *Comparisons with satellite values*

One must bear in mind the detection and data collection differences between LMA and satellite when making comparisons of corresponding climatologies. Recall that satellites detect optical radiation produced by lightning, while LMA networks detect VHF radiation produced by discontinuous lightning breakdown. LMA networks continuously detect the same storms while satellites get a snapshot of lightning in a storm.

Average flash densities within the DC network (where LMA detection efficiency is highest) were 18 flashes $km^{-2} yr^{-1}$ compared to 10-15 flashes $km^{-2} yr^{-1}$ for the combined LIS/OTD climatology (HRFC_COM product from http://thunder.nsstc.nasa.gov; Cecil et al. 2014). LMA flash densities over the northern Alabama network are roughly 33% larger than comparable satellite observations (40 vs. 30 flashes $km^{-2} yr^{-1}$) by Christian et al. (1999, 2003) and Cecil et al. (2014). It should be noted that these values did not significantly change if gridded to the 0.5° x 0.5° resolution of Cecil et al. (2014). The comparisons between LMA and satellite constructed climatologies in DC and Alabama are relatively close to each other. In both of these regions, the large majority of storms are generally weak with low flash rates (Fuchs et al. 2015). Weaker storms are hypothesized to have less turbulence and correspondingly larger flashes (Bruning and MacGorman 2013).

By a wide margin, the largest difference between LMA and satellite flash densities is in the northeast Colorado region. Average LMA flash densities within the network are $\sim$ 50-55 flashes $km^{-2} yr^{-1}$, compared to $\sim$ 15-20 flashes $km^{-2} yr^{-1}$ for satellite detectors (OTD at latitudes greater than $\sim$ 38 °). This striking difference is roughly a factor 3 times larger for the LMA climatology than the OTD climatology. The flash sensitivity studies in this paper indicate the highest confidence in the flash density values in the Colorado region. Flash characteristics were least sensitive to network proximity and the strictest clustering thresholds were imposed in this region (section 2). It is difficult to estimate error bars with the LMA flash density given the numerous unknowns that cannot be tested with millions of flashes. Given the comparisons with XLMA in Fuchs et al. (2015) and subjective analysis of several storms (similar to Figure 3 from Fuchs et al. 2015), we estimate the errors to be approximately 10-20%. The differences between LMA and satellite climatologies are far outside of these bounds. It is worth noting that the short duration of both the LMA observation period of 2 years and the satellite observation of 5 years in Colorado may play a role in the large differences between the corresponding flash density estimates, especially because the observation periods are not concurrent. Additionally, the satellite climatologies are gridded more coarsely which will smooth out small-scale variations in flash density, which may have otherwise had higher values. However, the coarse grid should not result in any missed flashes by the satellites and therefore should remain a valid comparison with azimuthally averaged values of flash density, as in Figures 2.2e, 2.5e and 2.8e.

Perhaps not coincidentally, the flash density differences between satellite and LMA datasets are correlated with storm intensity. Fuchs et al. (2015) showed that isolated storms in Colorado most frequently produced large flash rates greater than 10 flashes min^{-1}. This was of particular importance for the bandwidth limited OTD instrument, as storms with large flash rates saturated sometimes the instrument and resulted in an erroneously low flash rate. It is important to note that the bandwidth limit will not be a problem on the GLM instrument to be launched later this year. Fuchs et al. (2015) showed that storms in Colorado most frequently produced high flash rates and had the highest IC:CG values of any region in the study. It is reasonable to suppose that the flash area distribution in high flash rate storms is skewed towards smaller flashes, following Bruning and MacGorman (2013). Those small flashes, more common in Colorado storms, may go undetected by satellites. Furthermore, comparisons of Figures 2.2d, 2.5d and 2.8d reveal that a much larger fraction of flashes initiate at lower altitudes in Colorado storms compared to Alabama or DC storms. The optical path for these flashes to reach satellite detectors is much longer than flashes in the upper portion of the cloud, which may result in undetected flashes by satellite instruments. This is corroborated by Thomas et al. (2000), which found that the LIS instrument detected a relatively small fraction of flashes that occurred in the lower portions of an Oklahoma storm. Fuchs et al. (2015) showed that a substantial fraction of storms in Colorado contained inferred mid-level positive charge, in contrast with storms in the Alabama and DC regions, suggesting that charge configurations in a storm may impact the number of lightning flashes detected by satellites during an overpass. By extension, regions with similar environments conducive to strong storms with a preference for low-altitude IC flashes (e.g. high CBH and instability) may have similarly undetected flashes.

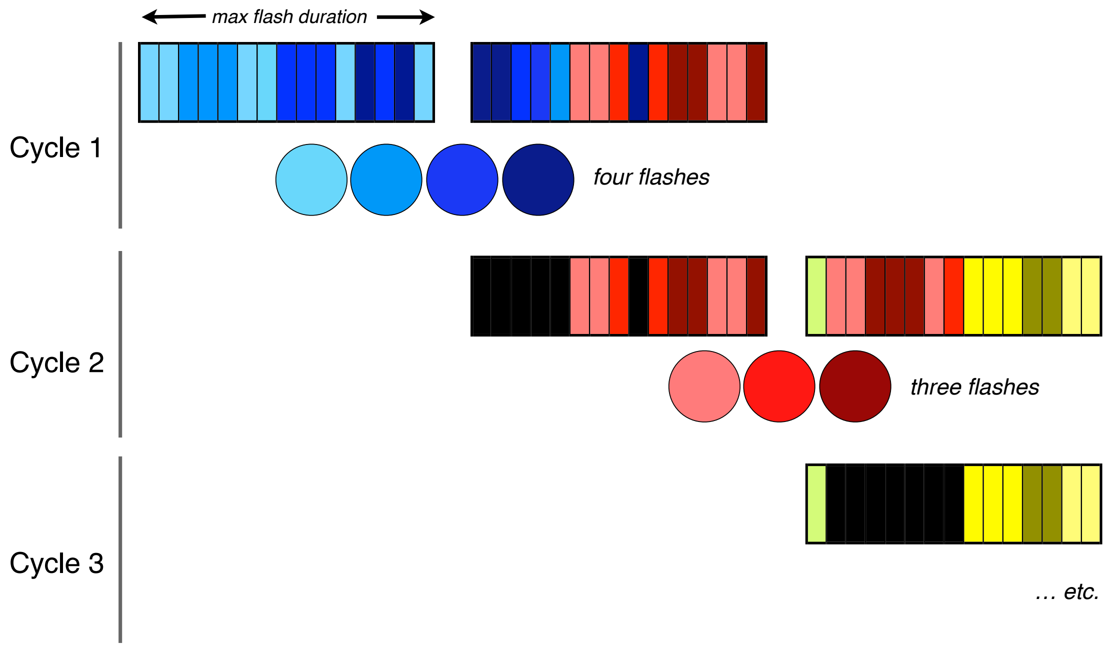

FIG. 2.1. Graphical illustration of the streamed clustering approach that utilizes a rolling buffer of VHF sources. In each cycle, the two horizontal bars represent the first and second halves of the source buffer. The vertical subdivisions of each large horizontal bar represent VHF sources. Colors encode clusters of sources identified as flashes. At the end of the first cycle, all sources (in blue) in the first half of the buffer are assigned to flashes and cleared. Some sources (in blue) from the second half of the buffer are removed because they were also clustered as part of flashes during the first cycle. The remaining sources in the second half of the buffer (in red) are joined by a new half-buffer of sources (in yellow and red), and the process repeats.

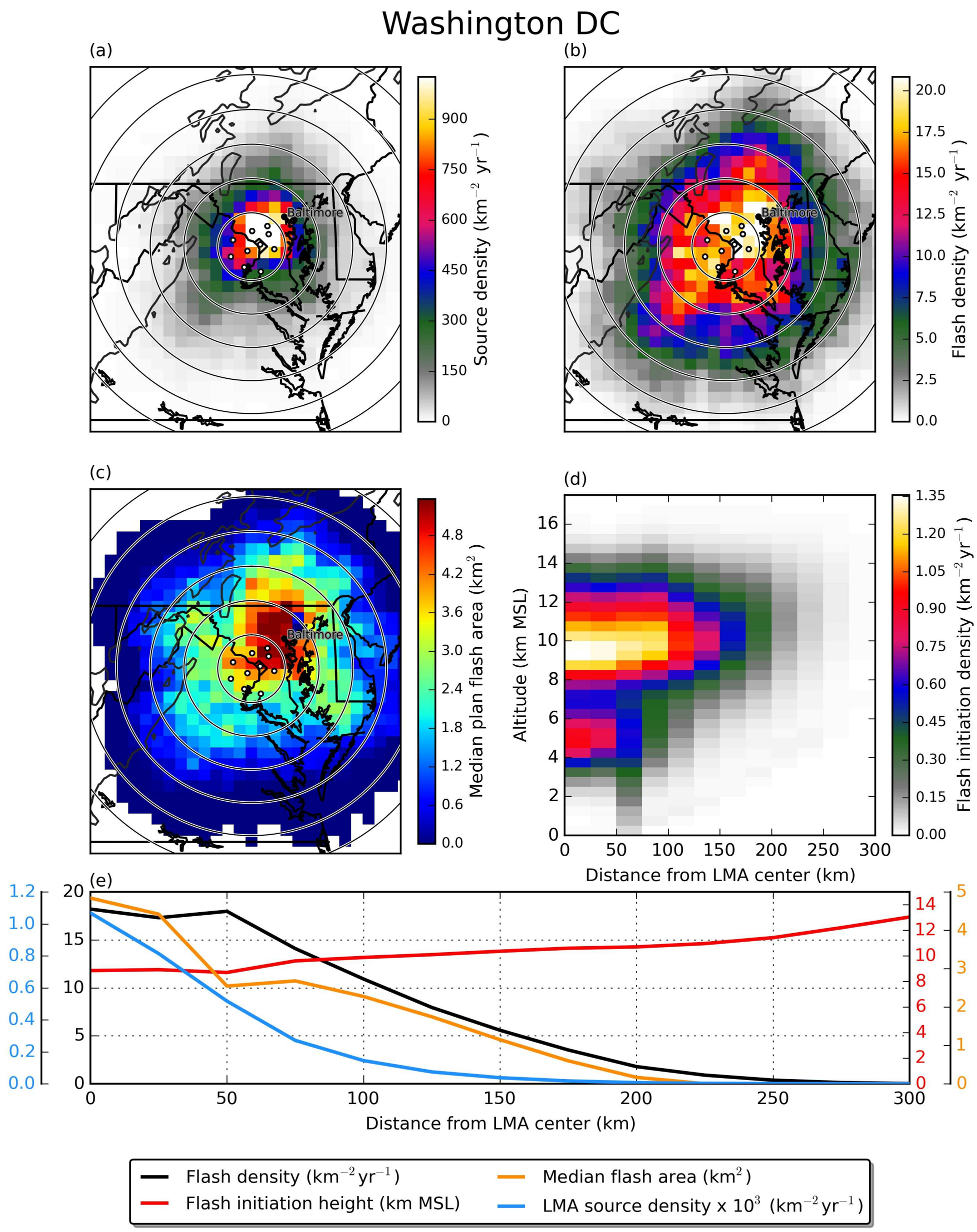

FIG. 2.2. (a) LMA source density for the Washington, DC network on a 0.15° x 0.2° grid. (b) Total lightning flash density from the clustering algorithm on the same grid. (c) Median plan flash area for all flashes within a grid box. Locations with less than 1 flash $km^{-2}yr^{-1}$ are excluded to minimize outlier effects. Dark gray contours on (a)-(c) are 300, 1000, and 2000 m MSL ground elevation. (d) Flash initiation height density as function of distance from the center of the network. (e) Flash quantities azimuthally integrated to give behavior as function of distance from the center of the network. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

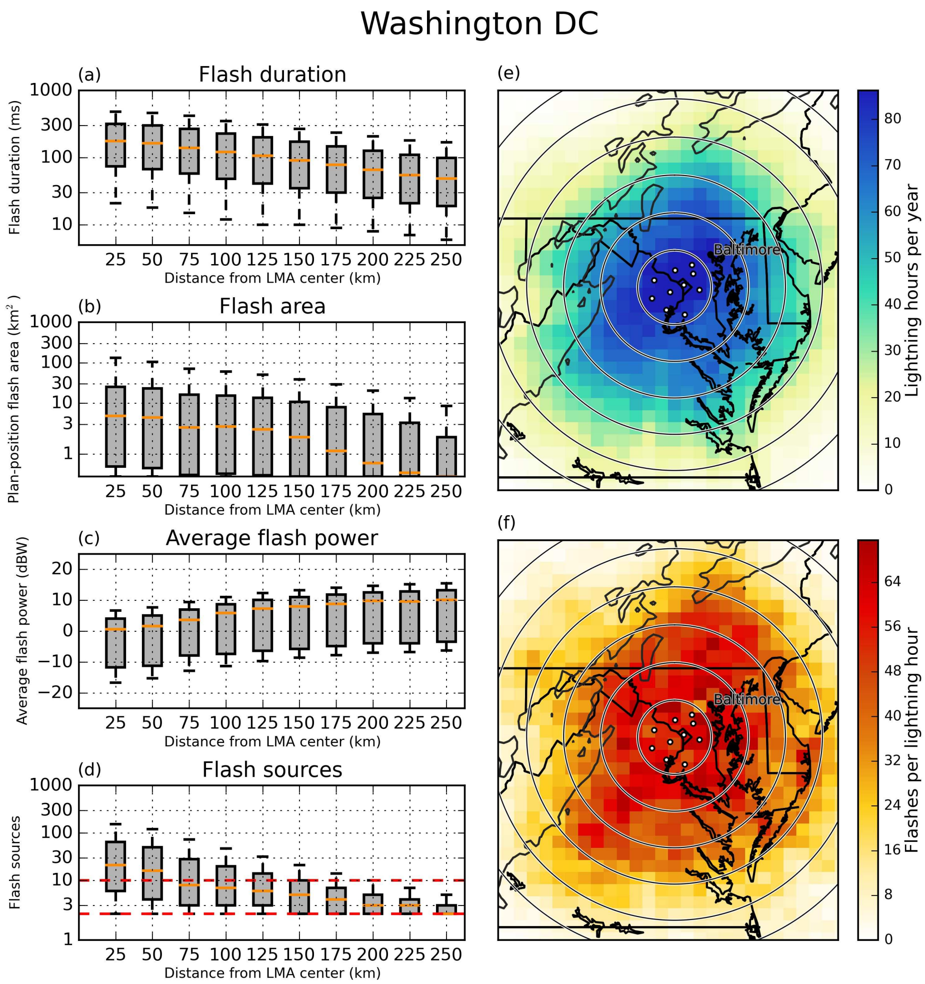

FIG. 2.3. (a) Distributions of flash duration for flashes in each range bin from the DC network center. Bars indicate the median value. Top and bottom of the box indicate the 25th and 75th percentiles, respectively. Whiskers indicate the 5th and 95th percentiles, respectively. (b) Similar to (a) for plan-position flash area. (c) Similar to (a) for the average source power within a flash. (d) Similar to (a) for the total number of points associated with each flash. The red dashed lines indicate the typical thresholds of 2 and 10 points for reference. (e) Map of lightning-hours for each grid box. Indicative of how many hours during a year at least one detected flash occurred in a particular grid box. (f) Map of the average number of flashes that occurred during an hour when at least one lightning flash occurred. This can be thought of as an average grid total flash rate.

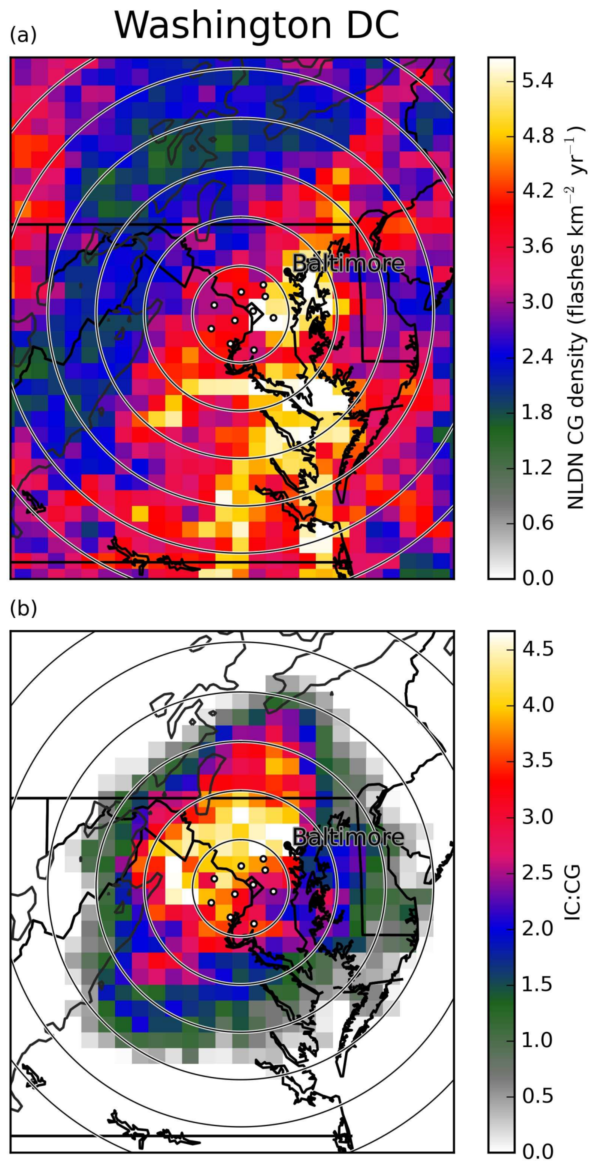

FIG. 2.4. (a) Peak current-filtered CG flash density map from the NLDN from the same timeframe as the available LMA data in DC. (b) Derived IC:CG values from the NLDN CG rate and the LMA calculated total flash rate assuming 100% detection efficiency. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

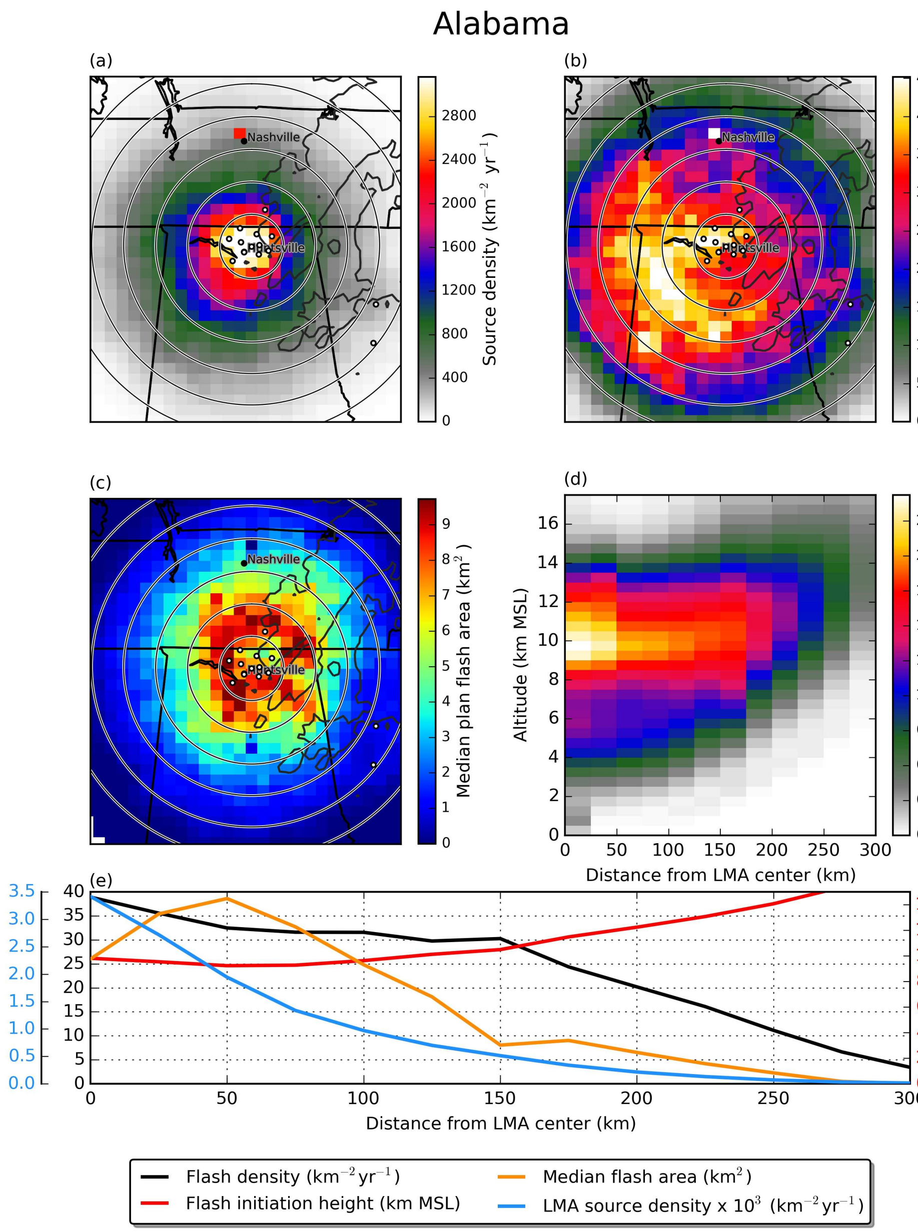

FIG. 2.5. (a) LMA source density for the northern Alabama network on a 0.15° x 0.2° grid. (b) Total lightning flash density from the clustering algorithm on the same grid. (c) Median plan flash area for all flashes within a grid box. Locations with less than 1 flash $km^{-2}yr^{-1}$ are excluded to minimize outlier effects. Dark gray contours on (a)-(c) are 300, 1000, and 2000 m MSL ground elevation. (d) Flash initiation height density as function of distance from the center of the network. (e) Flash quantities azimuthally integrated to give behavior as function of distance from the center of the network. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

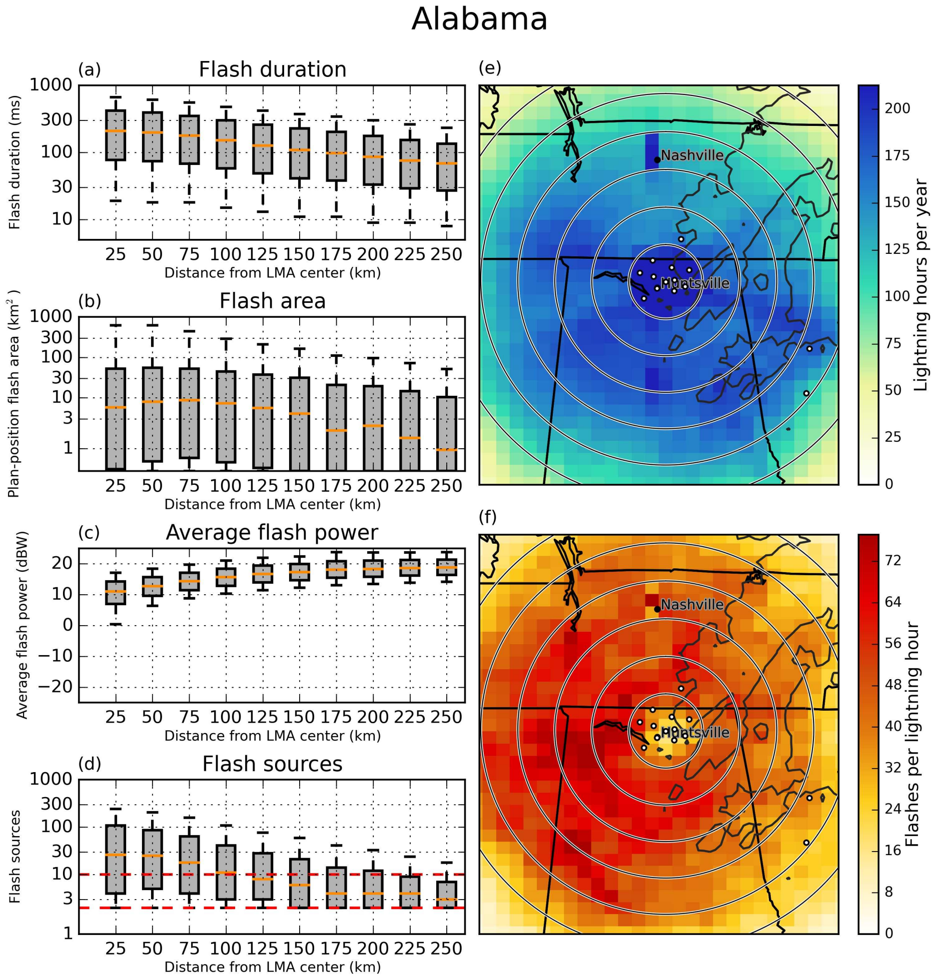

FIG. 2.6. (a) Distributions of flash duration for flashes in each range bin from the northern Alabama network center. Bars indicate the median value. Top and bottom of the box indicate the 25th and 75th percentiles, respectively. Whiskers indicate the 5th and 95th percentiles, respectively. (b) Similar to (a) for plan-position flash area. (c) Similar to (a) for the average source power within a flash. (d) Similar to (a) for the total number of points associated with each flash. The red dashed lines indicate the typical thresholds of 2 and 10 points for reference. (e) Map of lightning-hours for each grid box. Indicative of how many hours during a year at least one detected flash occurred in a particular grid box. (f) Map of the average number of flashes that occurred during an hour when at least one lightning flash occurred. This can be thought of as an average grid total flash rate.

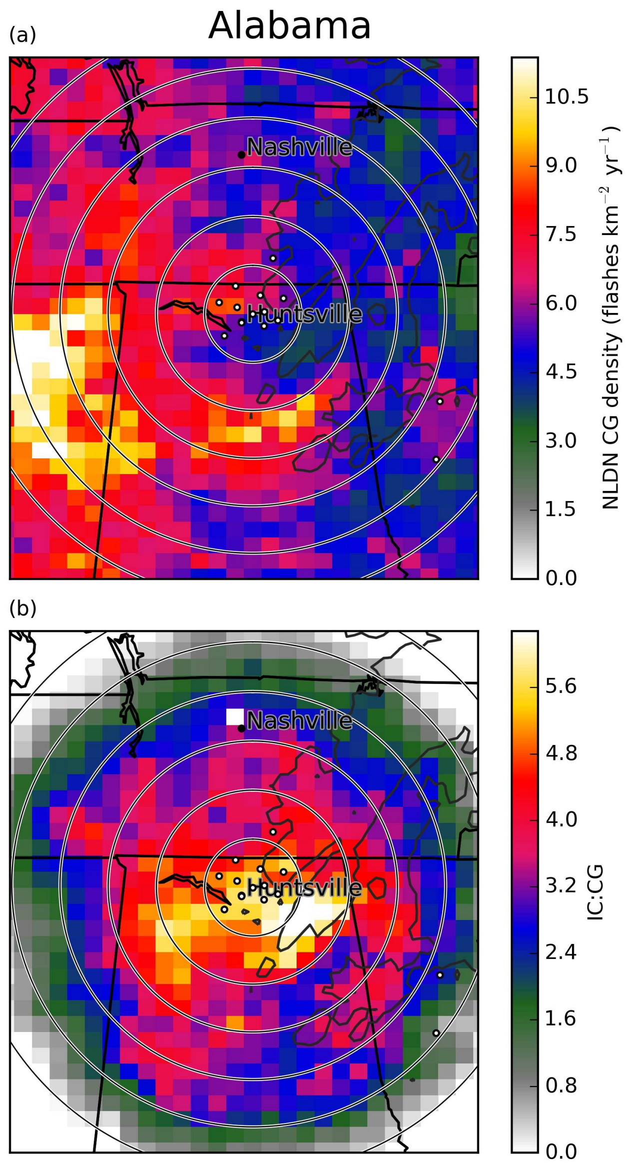

FIG. 2.7. (a) Peak current-filtered CG flash density map from the NLDN from the same timeframe as the available LMA data in northern Alabama. (b) Derived IC:CG values from the NLDN CG rate and the LMA calculated total flash rate assuming 100% detection efficiency. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

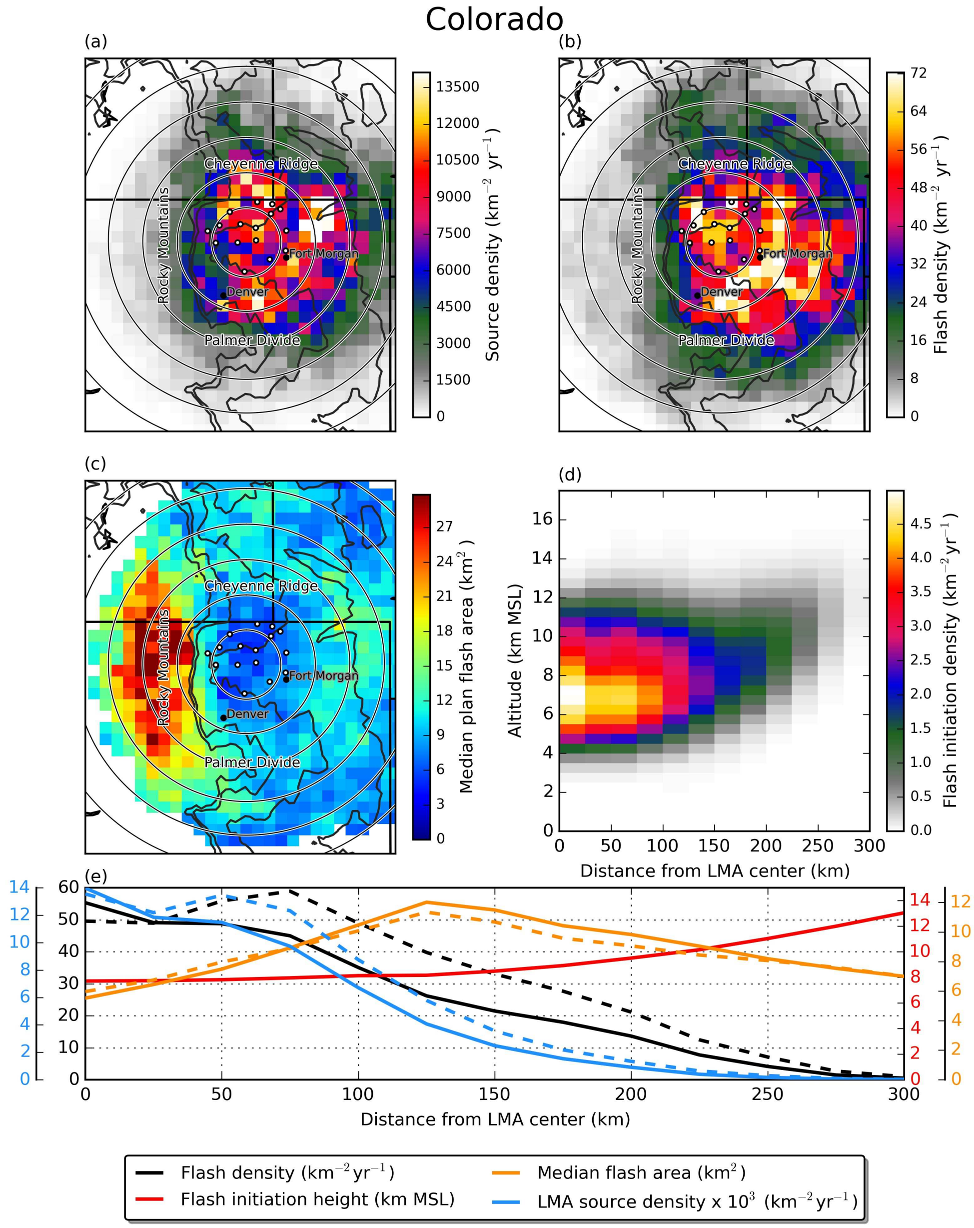

FIG. 2.8. (a) LMA source density for the Colorado network on a 0.15° x 0.2° grid. (b) Total lightning flash density from the clustering algorithm on the same grid. (c) Median plan flash area for all flashes within a grid box. Locations with less than 1 flash $km^{-2}yr^{-1}$ are excluded to minimize outlier effects. Dark gray contours on (a)-(c) are 1200, 1500, 1800, 2000, 4000 m MSL ground elevation. (d) Flash initiation height density as function of distance from the center of the network. (e) Flash quantities azimuthally integrated to give behavior as function of distance from the center of the network. Dotted lines indicate values when only considering sources and flashes in the eastern portion of the network. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

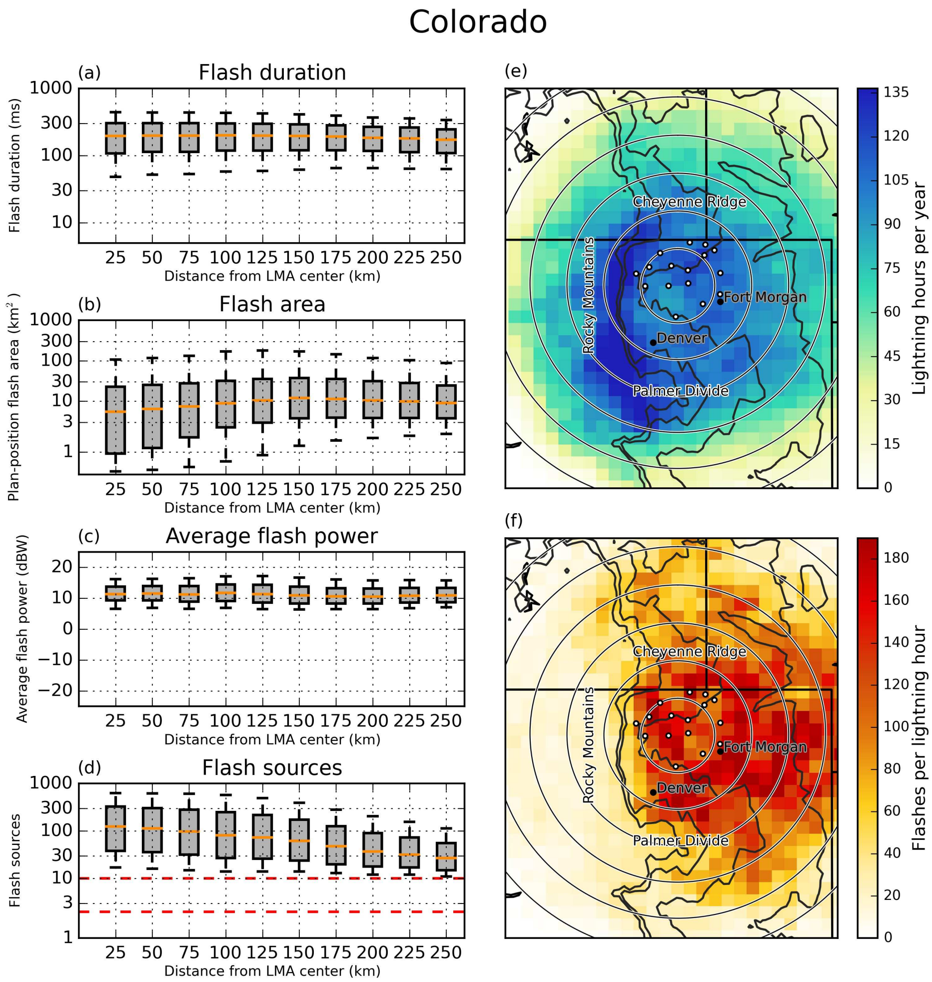

FIG. 2.9. (a) Distributions of flash duration for flashes in each range bin from the Colorado network center. Bars indicate the median value. Top and bottom of the box indicate the 25th and 75th percentiles, respectively. Whiskers indicate the 5th and 95th percentiles, respectively. (b) Similar to (a) for plan-position flash area. (c) Similar to (a) for the average source power within a flash. (d) Similar to (a) for the total number of points associated with each flash. The red dashed lines indicate the typical thresholds of 2 and 10 points for reference. (e) Map of lightning-hours for each grid box. Indicative of how many hours during a year at least one detected flash occurred in a particular grid box. (f) Map of the average number of flashes that occurred during an hour when at least one lightning flash occurred. This can be thought of as an average grid total flash rate.

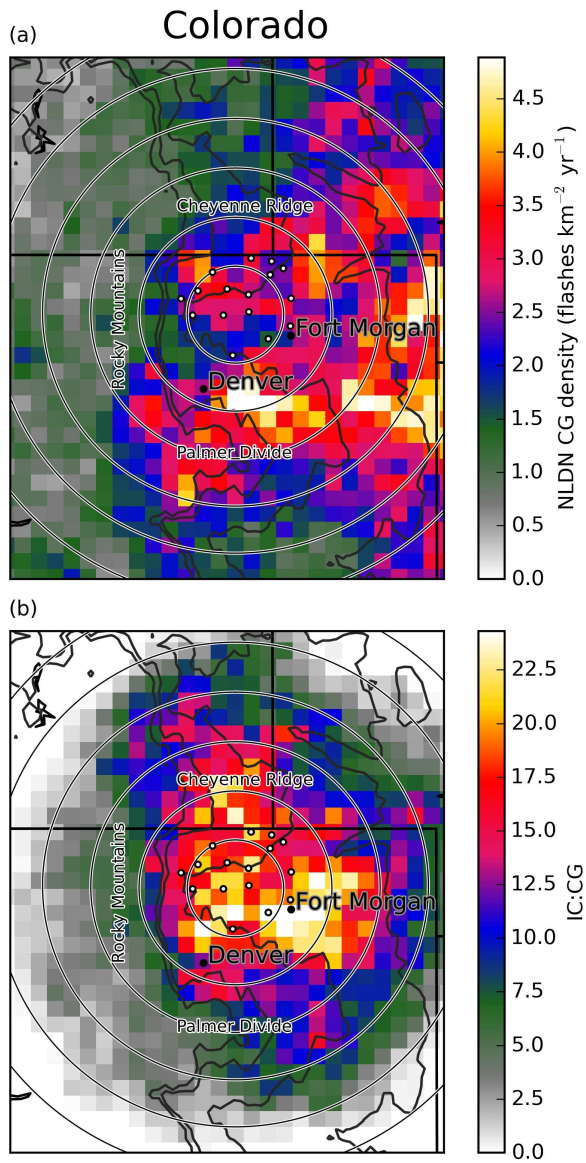

FIG. 2.10. (a) Peak current-filtered CG flash density map from the NLDN from the same timeframe as the available LMA data in Colorado. (b) Derived IC:CG values from the NLDN CG rate and the LMA calculated total flash rate assuming 100% detection efficiency. White dots indicate the location of LMA stations present during the analysis period. Range rings are increments of 50 km.

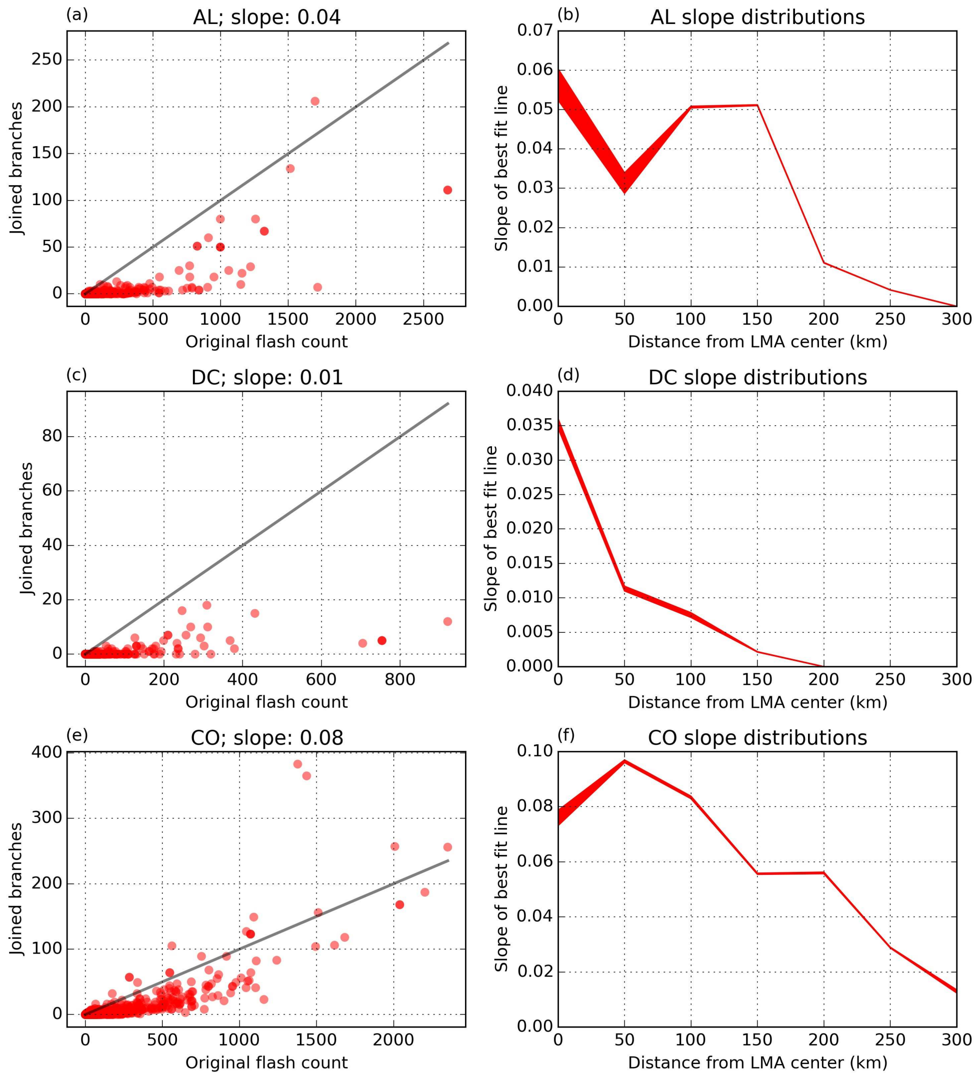

FIG. 2.11. (a) Scatter plot of original number of flashes in a cell and the number of flashes joined together in Alabama. The black line indicated 10% of flashes have been joined together. The slope of the least-squares best-fit line is indicated in the title. (b) The slope of the best-fit line if only considering cells in a binned distance from the Alabama network center. The width of the line indicates the 5th and 95th percentile of slopes using the jackknife method. (c) Same as (a) for the DC region. (d) Same as (b) for the DC region. (e) Same as (a) for the Colorado region. (f) Same as (b) for the Colorado region.

CHAPTER 3

Estimating lightning flash locations in isolated convective thunderstorms using LMA observations

3.1. Introduction

Lightning is generated by microscale processes but has impacts on much larger scales. Collisions between large precipitation ice particles (graupel) and small ice crystals in the presence of supercooled cloud water result in significant charge separation on the particle scale (e.g., Reynolds et al. 1957; Takahashi 1978; Jayaratne et al. 1983). Strong electric fields result when gravity and convective updrafts vertically separate large and small particles, which carry electrical charge of opposite polarity (Williams 1985). The sign of charge on graupel depends on the local temperature and amount of supercooled liquid water (SCLW; e.g., Saunders and Peck 1998). The general consensus of numerous laboratory studies is that graupel particles (ice crystals) tend to charge positively (negatively) in portions of the cloud with warm temperatures and high SCLW (Saunders et al. 1991; Takahashi et al. 2017). Accordingly, graupel particles (ice crystals) tend to charge negatively (positively) in portions of the cloud with colder temperatures and lower SCLW. Because the mixed-phase charging zone in thunderstorms spans a fixed temperature zone (0 °C to -40°C), it is hypothesized that the charge structure within a storm depends on SCLW in the mixed-phase (e.g., Bruning et al. 2014). Most thunderstorms possess a normal charge structure consisting of mid-level (∼ -20 °C) negative charge situated between regions of upper (∼ -40 °C) and lower-level (∼ 0°C) positive charge (e.g., Williams et al. 1989). The mid-level negative charge is thought to be comprised of negatively charged graupel particles, which suggests that SCLW amounts are modest, based on laboratory studies. Accordingly, ice crystals acquire positive charge from collisions with mid-level graupel and are carried to the upper level. The dipole comprised of the mid-level

This chapter is in review in the Journal of Geophysical Research: Atmospheres negative and upper-level charge regions is usually responsible for most of the lightning flashes in a normal polarity storm (Krehbiel et al. 1979; Krehbiel 1986). In some cases, the “anomalous” or “inverted” charge structures are present. These storms possess dominant mid- or low-level positive charge, and have been termed anomalous or inverted (e.g., Stolzenburg 1994; Lang et al. 2004). These storms are thought to have relatively large SCLW amounts that lead to positive graupel charging in the mid-levels. Accordingly, negatively charged ice crystals are carried to the upper-levels of these storms to form the upper dipole.

Regardless of the polarity of nearby regions of charge, lightning initiates in regions where the local electric field exceeds the breakdown threshold of air (Maggio et al. 2005). After initiation, lightning channels propagate into potential energy wells created by charged hydrometeors (MacGorman et al. 2001; Coleman et al. 2003). Therefore, the location of lightning flash initiation and flash propagation are largely governed by the storm-scale charge structure.

Fuchs et al. (2015) recently detailed significant variability in macroscale charge structures in isolated convective storms for one warm season near four Lightning Mapping Arrays (LMAs) in four environmentally distinct regions. In northern Alabama, Washington DC, and central Oklahoma, the overwhelming majority of isolated convective storms possessed a normal polarity charge structure, marked by LMA-inferred dominant positive charge in the upper levels near -40 °C (Section 3.2.3). Conversely, the vast majority of isolated convective storms in northeastern Colorado possessed an anomalous charge structure, marked by LMA-inferred dominant positive charge at temperatures warmer than -25 °C (Section 3.2.3). In an effort to understand the mechanisms that lead to anomalous storms in the Colorado region, Fuchs et al. (2015) found that Colorado storms had substantially higher cloud base heights and shallower warm cloud depths. It was hypothesized that these quantities influence the mixed-phase SCLW amounts, which in turn dictate the storm-scale charge structure. Given the regional differences in storm-scale charge structures, and the theoretical controls of storm-scale charge structures on flash locations, it is reasonable to expect variability in the vertical lightning locations between the normal polarity storms in Alabama, DC, and Oklahoma and the anomalous storms in Colorado.

In a comparison between multiple years of satellite and LMA lightning observations, Fuchs et al. (2016) found corresponding annual flash density estimates in northern Alabama and DC were within 50%. These differences were considered reasonable because the platforms use different detection techniques and observe different physical processes. Specifically, LMAs detect sub-flash processes in the very high frequency (VHF) portion of the electromagnetic spectrum (Rison et al. 1999) while satellite detectors use optical sensors viewing the storm from above (Christian et al. 1999, 2003). In northeastern Colorado, LMA-based flash density estimates were approximately 300% higher than corresponding satellite-based estimates, which was considered outside the uncertainty limits and therefore physically significant. Furthermore, the ratio of intra-cloud to ground flashes (IC:CG) in northeastern Colorado were approximately 300% higher than previous studies based on satellite estimates (e.g., Boccippio et al. 2001). One of the possible reasons Fuchs et al. (2016) offered for the annual flash density discrepancies was flash initiation altitude. They showed that the majority of flashes in the Alabama and DC regions initiated near 9-10 km MSL, while the large majority of flashes in northeastern Colorado initiated near 6-7 km MSL. However, that study only had flash initiation altitude information so they could only hypothesize that the lower altitude flashes may be more difficult to detect by optical sensors viewing storms from above.

Lightning impacts its environment, including the production of LNO_x (e.g. Schumann and Huntrieser 2007). Heated lightning channels dissociate diatomic nitrogen and oxygen, creating NO and NO_2 molecules, collectively called NO_x (e.g. Goldenbaum and Dickerson 1993; Huntrieser et al. 1998). NO_x is important in the atmosphere, especially in the upper troposphere, because it is a precursor for tropospheric ozone, an important greenhouse gas (e.g., Price et al. 1997; DeCaria et al. 2005). In fact, Schumann and Huntrieser (2007) claim that lightning is the dominant source of NO_x in the upper troposphere. Importance is placed on the vertical variability of flash channels, since NO_x ozone production efficiency depends on the ambient temperature, hydrocarbon concentration, NO_x concentration and NO_x lifetime, all of which depend on altitude (Lin et al. 1988; Finney et al. 2016). Therefore, any variability in vertical lightning locations may impact the amount of ozone produced by a particular storm. Furthermore, many chemistry models place LNO_x uniformly in the horizontal within the 20 dBZ isopleth (Pickering et al. 1998; Ott et al. 2010), so it is important to understand where lightning channels are produced with respect to reflectivity.

The substantial differences in flash initiation heights in thermodynamically and electrically distinct regions in Fuchs et al. (2015, 2016) provide significant motivation to investigate vertical locations of lightning channels. Furthermore, the location of lightning influences the subsequent transport of LNO_x and conversion to ozone (Pickering et al. 1998; DeCaria et al. 2000) and possibly the ability to optically detect lightning from spaceborne platforms (Fuchs et al. 2016). This study estimates lightning flash channel locations using highly accurate LMA observations of lightning flash initiation and propagation from the statistically significant dataset of isolated convective storms originally crafted by Fuchs et al. (2015). Calculated lightning channels are investigated with respect to height as well as collocated reflectivity. Importance will be placed on regional differences of these quantities and their relationship with LMA-inferred charge structure.

3.2. Data and methodology

3.2.1. *Radar*

The radar data in this study is derived from the National Mosaic and Multi-Sensor Quantitative Precipitation Estimates (NMQ) mosaic 3D radar dataset (Zhang et al. 2011). NMQ data is comprised of gridded Next Generation Radar Weather Surveillance Radars (NEXRAD-WSR88D) reflectivity data and covers the entire continental United States (CONUS) with a temporal resolution of 5 minutes during the observation period (2011-2012). NMQ Cartesian mosaic data provides approximately 1km x 1km horizontal resolution with a stretched vertical grid from 500 meters to 18 km above mean sea level (MSL). The vertical resolution varies from 500 meters near the surface to 2 km near the top of the domain.

3.2.2. *CLEAR*

The dataset for this study was originally created by Fuchs et al. (2015) using the CSU Lightning, Environmental, Aerosol and Radar (CLEAR) automated case study framework (Lang and Rutledge 2011). CLEAR uses objective cell identification and tracking on composite reflectivity fields, then attributes various types of information, such as lightning flashes from LMA data. Fuchs et al. (2015) imposed strict criteria for cell identification: contiguous 35 dBZ regions of area $\sim$ 20 km^2 containing a contiguous 45 dBZ region of area $\sim$ 13 km^2. This resulted in the removal of most small and non-lightning producing cells while preserving storms near the mature phase of their lifecycle. CLEAR was applied to NMQ mosaic data centered near LMAs in four environmentally distinct regions (Washington DC, northern Alabama, central Oklahoma and northeast Colorado) for one warm season to produce a database of over 4000 observations of isolated convective storms containing over 500000 flashes. The data in Alabama, DC, and Oklahoma come from the 2011 warm season and the data in Colorado comes from the 2012 warm season. The CO LMA was installed in the spring of 2012 for the Deep Clouds and Convective Chemistry field experiment (DC3; Barth et al. 2015).

3.2.3. *LMA*

LMA networks use time-of-arrival differences from multiple antennae to detect very high frequency (VHF; $\sim$ 60-66 MHZ or 5 m wavelength) radiation sources produced by the discontinuous breakdown of lightning channels (Rison et al. 1999). LMAs can accurately detect the in-cloud portion of flash channels (Krehbiel et al. 2000). LMAs may detect a few VHF sources to a few thousand VHF sources from a single flash, depending on the spatial extent of the physical flash, the location of the flash relative to the network, and the detection capabilities of the network (Thomas et al. 2004; Fuchs et al. 2016). VHF source location errors are on the order of tens of meters within the physical boundary of the network and increase with distance from the center of the network (Thomas et al. 2004). Additionally, detection efficiency decreases with increasing range from the network center (Thomas et al. 2004; Koshak et al. 2004; Chmielewski and Bruning 2016; Fuchs et al. 2016). Accordingly, only identified cells within 125 km of the LMA center are included in the analysis to mitigate LMA detection efficiency and location artifacts.

LMA detection of leader propagation depends on channel polarity (Rison et al. 1999). Once a lightning flash is initiated, leaders of opposite polarity propagate into potential energy wells to neutralize charge buildup. Negative leader propagation into positive charge is more discontinuous than propagation of positive leaders into negative space charge (Rison et al. 1999). Since VHF burst radiation is produced during discontinuities in leader propagation, negative leaders propagating in positive charge produce more VHF radiation than positive leaders propagating in negative charge (Rison et al. 1999; Williams 2006). This has both positive and negative implications. We can exploit this asymmetry to infer the location of positive charge, since positive charge is noisier in the VHF portion of the spectrum. The integrated result over a single flash or a whole storm is that more VHF radiation is produced in positive charge regions, which allows us to infer the dominant positive charge to be near the mode of vertical VHF source distribution (Wiens et al. 2005). Normal polarity storms are typically defined as having an LMA mode temperature near -40 °C and anomalous polarity storms are typically defined as having an LMA mode temperature around -20 °C (Wiens et al. 2005; Lang and Rutledge 2011; Fuchs et al. 2015). There are downsides to differences in leader propagation discontinuity. LMA depictions of flashes can be incomplete because the positive breakdown in negative charge can go undetected. Implications of the asymmetry and its impacts on the calculation of flash channels are discussed further in Secs. 3.2.4 and 3.2.5.

3.2.4. *Flash channel calculation*

Because LMA networks detect sub-flash features (Rison et al. 1999; Thomas et al. 2001; Bruning et al. 2007) further processing of VHF source data is required to retrieve information about individual flashes. Bruning (2013) and (Fuchs et al. 2015, 2016) described a flash-clustering algorithm that groups sources by space and time for flash identification. Once the VHF sources are grouped into flashes, useful information about each flash, such as location and spatial extent can be extracted. Of particular interest to this study are the three-dimensional channels of flash propagation. To estimate the spatial channel distribution of channels produced by each flash, a grid is imposed on the VHF sources comprising a flash. Any grid box containing at least one VHF source is considered to have contained a segment (or channel) of that flash. This process is essentially a three-dimensional extension of the flash extent density product (Mansell 2014).

Figure 3.1 shows an illustration of the process on a longitude-altitude cross-section of a sample flash with a 1 km x 1 km x 1 km grid. This flash occurred in Colorado, approximately 20 km from the center of the Colorado LMA network, and was chosen because the high detection efficiency of the Colorado LMA (Fuchs et al. 2016) yields a near-complete representation of the physical flash. In Figure 3.1a, grid boxes containing at least 1 LMA VHF source from the flash are highlighted to indicate that a segment of that flash passed through those boxes. The corresponding vertical distribution of calculated flash channels is shown in Figure 3.1e, which shows most of the flash channels residing between 5-8 km MSL. The average power of the VHF sources in the original flash is 14.9 dBW, which is typical of Colorado flashes (Fuchs et al. 2016).

To simulate decreasing detection efficiency, Figure 3.1b shows the same flash as Figure 3.1a, but with the lowest 25% of VHF source powers removed. Because the power from an emitted source follows the inverse square law, weaker sources may not be detected at longer ranges. Therefore, removing low VHF source powers is analogous to moving the flash farther away from the network center. The distribution of VHF source powers for this flash is shown in Figure 3.1f, with the 25th, 50th and 75th percentiles of those VHF source powers indicated by the vertical lines. The 25th percentile of VHF source powers for this flash is approximately 8 dBW (6.3 W). The distribution of calculated flash channels in Figure 3.1b largely resemble the flash channels from the original flash in Figure 3.1a, except a couple of grid boxes are missing. This is to be expected if the LMA is not detecting the entire flash. It is important to note that the vertical distribution of calculated flash channels is nearly identical to the original flash (Figure 3.1e). Further simulated degradation of LMA source detection is shown in Figure 3.1c (Figure 3.1d), which shows only the sources that have powers greater than the 50th (75th) percentiles within the flash. As sources are removed from the flash, the depiction of the flash becomes less complete. A key point is that the vertical distribution of calculated flash channels for each case Figures 3.1b-3.1d (each line in Figure 3.1e) closely resembles that of the original flash. Since the vertical distribution of flash channels is of principal interest to this study, perhaps variable detection efficiency will not lead to a substantial bias in calculated flash channels. However, Thomas et al. (2001) showed for a couple of flashes that negative breakdown into regions of positive charge produced higher VHF source powers than positive breakdown into regions of negative charge. If that is indeed the case in a majority of flashes, we would expect decreasing detection efficiency to lead to undetected positive breakdown in negative charge regions, resulting in biases in calculated flash channels toward positive charge. We removed low-power VHF sources in numerous random flashes in a manner similar to Figure 3.1 and found no systematic differences between vertical distributions of calculated flash channels. The difference in average power from the original flash (Figure 3.1a) to the flash subset of the 25% highest powers (Figure 3.1d) is approximately 4 dBW. Results from Fuchs et al. (2016) suggest that a $\sim$ 4-5 dBW increase in average power is equivalent to moving the flash from inside the network to $\sim$ 100 km farther away from the network.

3.2.5. *Grid size sensitivity*

The calculation of flash channels requires an imposed grid upon which the flash channels are mapped out. Therefore, it is reasonable to expect that the calculated flash channels will depend on the imposed grid size. Figure 3.2 illustrates how the grid size may affect the resulting flash channels on the same flash shown in Figure 3.1. Figure 3.2a shows a time-height cross-section of the flash. The first sources occur around 10 km MSL. Shortly after initiation, the heights of the sources descend to approximately 6 km MSL during the first 100 ms. Figure 3.2b shows the vertical distribution of LMA VHF sources (without any flash processing), where a maximum near 7 km MSL is evident. Based on the altitude of the LMA source maximum (Wiens et al. 2005) and the initial downward propagation (Shao and Krehbiel 1996; Bruning et al. 2007), we would expect the positive charge to reside around 6-8 km MSL. The other lines in Figure 3.2b show the vertical distribution of flash channels (similar to Figure 3.1e), gridded to 0.5 km, 1.0 km, and 2.0 km. The flash channels gridded to 0.5 km show a similar behavior to the LMA source distribution, with a prominent maximum near 6.5 km MSL. As the imposed grid increases in size (decreasing resolution), the resulting flash channels change as well. The flash channels calculated from a 1.0 km grid have a maximum at approximately 7.0 km MSL. Unsurprisingly, the vertical distribution is smoother than the 0.5 km grid flash channels. Finally, the flash channels calculated from the 2.0 km are even smoother and have a maximum around 8 km MSL.

Strictly speaking, as the grid cell size increases, the vertical distribution of the calculated flash channels becomes smoother because adjacent data points become more spread out. However, since only one VHF source is needed inside a box to assign a channel segment to that box, the asymmetry between the positive and negative leaders of the flash impacts the calculated flash channels. As the grids become larger, the likelihood of a box containing a VHF source produced by positive breakdown in negative charge increases. Conversely, the calculated flash channels from the negative breakdown into positive charge portion of the flash will remain largely unchanged because it is more likely that the smaller boxes already contain a VHF source since the LMA detects more of the negative leaders. These effects result in decreasing the weight of the positive charge in the flash channel calculation, which shifts the distribution towards the negative charge. This effect is observed in Figure 3.2, and was observed in numerous other flashes (not shown). It quickly proved unfeasible to visually verify this shift on the ~ 500000 flashes in this dataset, but we do expect this shift in most flashes that have an appreciable vertical extent.

3.3. Results

3.3.1. *Flash channels with respect to height and reflectivity*

This section will focus on the height of calculated flash channels and their collocated reflectivity values (from NMQ mosaic radar data). Due to the mismatch between radar data and the imposed grids, calculated flash channels are matched with the reflectivity from the nearest radar grid point. Sensitivity studies (not shown) indicate little dependence on the reflectivity collocation method. For every flash in each valid isolated convective cell, the channels are calculated based on the imposed grid, then the height and corresponding reflectivity values are recorded and saved. All the heights and reflectivities corresponding to each channel of every flash are summed and normalized by the number of valid cells for each region.

The resulting distributions from the flash channel calculations are shown in Figures 3.3-3.5, which show the climatological distributions of collocated reflectivity values for different grid sizes. The analysis is conducted on different grid sizes for reasons discussed earlier, namely the different weights implicitly assigned to positive charge. Figure 3.3 shows the distributions with flash channels gridded to 0.5 km for each of the 4 regions. In Alabama (AL), the majority of the calculated flash channels occurred between 8-10 km MSL. The average height of all calculated flash channels for AL was 8.9 km MSL and the average collocated reflectivity was 29.7 dBZ. Interestingly, 37% of all flash channels occurred in reflectivity regions of less than 20 dBZ. The flash channel distribution in the DC region is quite similar to the AL region, as most of the flash channels reside between 8-10 km MSL. The average flash channel height is the same as in AL, while the collocated reflectivity is slightly higher at 32.4 dBZ. Roughly one-quarter of the calculated flash channels were situated in reflectivities < 20 dBZ. The flash channel distribution in Oklahoma (OK) is shifted downward relative to the AL and DC regions, with most of the flash channels residing between 7-9 km MSL. The average flash channel height is 8.2 km MSL, 0.7 km, slightly lower than the AL and DC regions. The average collocated reflectivity is 37.9 dBZ, which is more than 5 dBZ higher than either AL or DC. Only a small fraction (14%) of the flash channels in Oklahoma occurred in regions with reflectivities < 20 dBZ. These results are consistent with the Oklahoma cases being generally more intense compared to the storms observed in AL and DC. OK storms had higher reflectivity values in the 7-10 km MSL layer compared to AL and DC storms (Figures 3.5b, 3.5c in Fuchs et al. 2015). The flash channel distribution in Colorado is the most bottom-heavy, as the average flash channel height was 7.1 km MSL, a full kilometer lower than OK and nearly 2 km lower than AL or DC. The average collocated reflectivity value is 41.2 dBZ, the highest of any region. Only 7% of calculated flash channels occurred in regions of less than 20 dBZ.

Figure 3.4 was constructed in the same manner as Figure 3.3, but the flash channels were calculated with a 1.0 km grid. Figure 3.4 shows similar results to Figure 3.3, with some small but important systematic differences. In AL, the average flash channel height is now 8.6 km MSL, 0.3 km lower than the flash channel distribution calculated with a 0.5 km grid. Perhaps coincidentally, the same downward shift in average flash channel height is evident in the DC and OK regions as well (0.3 km). The CO region has a downward shift of 0.2 km. The average collocated reflectivity values all change, but the trend remains the same: CO has the highest reflectivity, followed by OL and DC, with AL having the lowest.

Figure 3.5 shows flash channel distributions calculated with a 2.0 km grid. The average flash channel height is again shifted downward (relative to the 1.0 km grid) to 8.3 km MSL in AL. Similarly, in DC, the average flash channel height is shifted downward to 8.1 km MSL. In OK, the average flash channel height was shifted downward 0.5 km to 7.4 km MSL. Finally, the average flash channel height in CO storms is 6.6 km MSL. Note that for each distribution for the different grid sizes, the difference in average calculated flash channel height ($\sim$ 2 km) is approximately constant between CO and AL. The rank in average collocated reflectivity remained the same, regardless of grid size. This is expected, given that the storms in the CO region were the most intense of any region in this dataset (Fuchs et al. 2015).

As described above, the calculated flash channel distributions in each region shifted downward as the imposed grid size increased. This downward shift may be consistent with a normal polarity charge structures in AL and DC. Smaller grids give more weight to the upper-level positive charge, resulting in calculated flash channels at higher altitudes. Conversely, larger grids give less weight to positive charge, and therefore shift the distributions downward toward the mid-level negative charge. Indeed, Fuchs et al. (2015) found that a large majority of the isolated convective storms in AL and DC have LMA mode temperatures (proxy for dominant positive charge) near -40 °C, a strong indicator of normal polarity charge structure. The mechanism for the downward shift in CO with increasing grid size is more nebulous. Fuchs et al. (2015) showed that a large majority of isolated convective storms in Colorado had LMA mode temperatures near -20 °C, indicative of dominant mid-level positive charge and anomalous polarity. Following the logic from the AL and DC regions, the downward shift of average flash channel height with coarse grids suggests positive charge overlying negative charge. Perhaps this is indicative of mid-level positive charge and lower-level negative charge in anomalous polarity storms in Colorado.

3.3.2. *Regional vertical flash extent variability*

The flash channel calculation and co-located reflectivity has yielded some interesting results. However, the flash channel calculation method has shortcomings that relate to the inability of the LMA to adequately detect positive leader propagation in negative space charge. We address the validity of the flash channel calculation method by using a different metric to estimate the vertical extent of flashes. Figure 3.6 shows a joint histogram of the bottom (measured by the lowest altitude source in a flash) and top (measured by the highest altitude source in a flash) of every flash in our exhaustive dataset. Figure 3.7 shows the distributions of the minimum, mean, and maximum VHF source altitude for all flashes in each region. This method is output directly by the flash clustering algorithm, and is largely independent from the flash channel calculation. The average minimum flash altitude for all flashes in AL is 8 km MSL, the average mean flash altitude is approximately 10 km MSL and the average maximum flash altitude is 11.5 km MSL (Figures 3.6,3.7). The distributions of flash metrics in DC resembles AL, with a systematic downward shift of approximately 0.5 km. Each flash metric in OK is shifted lower than DC by approximately 0.3 km. The flashes in CO have some interesting behaviors. Two separate maxima are apparent in the joint distribution in Figure 3.6, which can be seen in Figure 3.7 as well. The mode of minimum flash altitudes in CO is around 4.5 km MSL, while a shoulder is present at approximately 9 km MSL. A similar signal is observed from the mean and maximum flash altitudes. The means of each flash height parameter distribution in Figure 3.7 are all approximately 2.0 km lower than corresponding values in AL. This number is very similar to the differences between average calculated flash channel heights as well.

3.3.3. *Charge structure dependence*

The propensity for anomalous charge structures (Fuchs et al. 2015), lower flash initiation altitudes (Fuchs et al. 2016) and flash channels (this study) in Colorado provide some evidence that the storm-integrated flash channels are controlled to first order by its macroscale charge structure. To investigate this claim more directly, Figure 3.8 shows a joint distribution of flash bottom and flash top (measured by the same method as Figure 3.6), partitioned by the LMA mode temperature of the storm in which each flash occurred, with all regions combined. Normal polarity charge structures are usually characterized by LMA mode temperatures (proxy for dominant positive charge) < -30 °C while anomalous charge structures are usually characterized by LMA mode temperatures > -30 °C (Wiens et al. 2005; Lang and Rutledge 2011). Figure 3.8a shows the flash top and bottom for all flashes that occurred in storms with an LMA mode between -60 °C and -40 °C. Most flashes in these storms have bottom heights of $\sim$ 9 km MSL or higher and most top heights are $\sim$ 11 km MSL or higher. Fuchs et al. (2015) showed that storms with very cold LMA modes are typically quite intense, and are argued to have positively charged ice crystals at very high heights as a result of strong updrafts. For lightning flashes that occurred in normal polarity storms with LMA mode temperatures between -40 °C and -30 °C, the average altitude of flash bottoms was 8 km MSL and the average altitude of flash tops was $\sim$ 10 km MSL. Notice that both heights are lower than flashes from Figure 3.8a, possibly due to the lower altitude (higher temperature) of the strong positive charge, as inferred by the LMA. Fuchs et al. (2015) showed that these storms had lower flash rates than storms with LMA mode temperatures < -40 °C, and therefore were not as intense. Figure 3.8c shows the top and bottom of all flashes in storms with LMA mode temperatures between -30 °C and -15°C, which are typically thought of as cases with anomalous polarity (strong mid-level positive charge). The downward shift in both flash bottom and flash top is evident, as the average value of the bottom of flashes is 7.1 km MSL. Average top height is 9.6 km MSL. This downward shift is likely due to the presence of a region of strong positive charge at warmer temperatures than is typically the case in normal polarity storms. Finally, the joint distribution of all flashes that occurred in storms with LMA mode temperatures between -15 °C and 0 °C is shown in Figure 3.8d. These flashes are the lowest of all storm categories, with an average value of 6.0 km MSL for the bottom of flashes and an average top altitude of 8.7 km MSL.

There is substantial regional disparity in lightning production as a function of charge structure. For example, 84% of the flashes in the AL region were produced by normal polarity storms (as defined earlier). In contrast, only 29% of the flashes in the CO were produced by normal polarity storms while 69% of the flashes were produced by anomalous polarity storms. It appears that storm charge structure exerts a significant control in the vertical location and extent of the flashes and is the reason that the

Colorado region stands out in comparison to the other regions in this study. Note that the percentages of flashes in each region do not add to exactly 100%, this is because a small percentage of flashes occurred in storms with LMA mode temperatures colder than -60 °C or warmer than 0 °C.

3.4. Summary and discussion

VHF-based mapping of lightning flashes by LMAs from over 4000 storms in diverse environments demonstrated the variability of calculated flash channel distributions with respect to collocated radar reflectivity and height, which produced some consequential insights. The flash channel calculation method presented herein undoubtedly inherited some of the shortcomings of LMAs and their inability to detect positive breakdown in negative charge. We have attempted to address these shortcomings by using a (mostly) independent method to estimate the vertical extent of flashes and are confident that the altitude differences observed between Colorado and the other regions not merely artifacts of the LMA.

In the Alabama and DC regions, storms produced the highest calculated flash channel heights, while the lowest calculated flash channel heights were found in Colorado storms. The lower altitude flash channels in the Colorado region, where the bulk of the storms have high cloud base heights and shallow warm cloud depths, corroborates the claims made by Fuchs et al. (2016). Indeed, the flashes that initiate at lower altitudes in Colorado translate to flash channel distributions centered lower in the storm, implying that flashes that start low tend to stay low in these storms. These flashes were also located in regions of relatively high radar reflectivity.

The low altitude flashes may also contribute to discrepancies between LMA and satellite climatologies in the Colorado region where annual flash densities from LMA observations are approximately 300% larger than satellite estimates (Cecil et al. 2014; Fuchs et al. 2016). Much smaller differences were observed in Alabama and DC, where flash altitudes were considerably higher, and in lower radar reflectivity. Note that the satellite-based estimates of annual flash density for Alabama were based on the Lightning Imaging sensor (LIS; Christian et al., 1999), while the estimates in DC and Colorado were only based on the Optical Transient Detector (OTD; Christian et al. 2003). However, Cecil et al. (2014) and Fuchs et al. (2016) note that the different detectors are not expected to significantly change the flash density estimates. The population of storms with shallow warm cloud depths was coincident with the highest fraction of intense anomalous storms by our LMA mode metric, and a low flash channel centroid. The flash channel dependence on warm cloud depth and charge structure implies that other regions of the globe with similar conditions to Colorado may have a portion of their flashes undetected by satellites. Apparently, the combination of low flash channels and high radar reflectivity reduces the ability of photons to escape through the top of the cloud and be detected by optical detectors on satellites. It would then be useful to understand other regions of the globe that may have the propensity for producing anomalous storms, as those regions may have similar undetected flashes by satellites.

The lower altitude flashes in Colorado storms in this dataset may also have implications for LNO_x transport and conversion to ozone. The Colorado storms in this dataset produced the highest flash rates, and therefore produced the highest amount of LNO_x, assuming for the sake of argument that each flash produces an equal amount of NO_x. However, since the anomalous storms in Colorado preferentially produce flashes at lower altitudes, the ozone production efficiency may be different in those storms. Furthermore, the vertical distribution of LNO_x may result in significant variability in LNO_x advection, and subsequent ozone production efficiency (Lin et al. 1988). NO_x produced at lower altitudes could be more subject to removal by storm downdrafts, transporting air towards the surface, LNO_x produced in the upper portion of the storm would be more likely to end up in the upper troposphere, assuming a typical profile of vertical motion (Battan and Theiss 1970; Cifelli et al. 2002). Additionally, Colorado storms had the highest percentage of calculated flash channels in reflectivities greater than 20 dBZ, the metric used by several chemical models. Conversely, more than a third of the calculated flash channels in Alabama storms occurred in reflectivities less than 20 dBZ. It is unclear what (if any) impact the parameterization placing NO_x within the 20 dBZ isopleth, regardless of storm characteristics, may have on the resulting NO_x and ozone distributions in a model. Regardless, all these factors need to be understood in order to fully understand the production of NO_x and ozone by lightning.

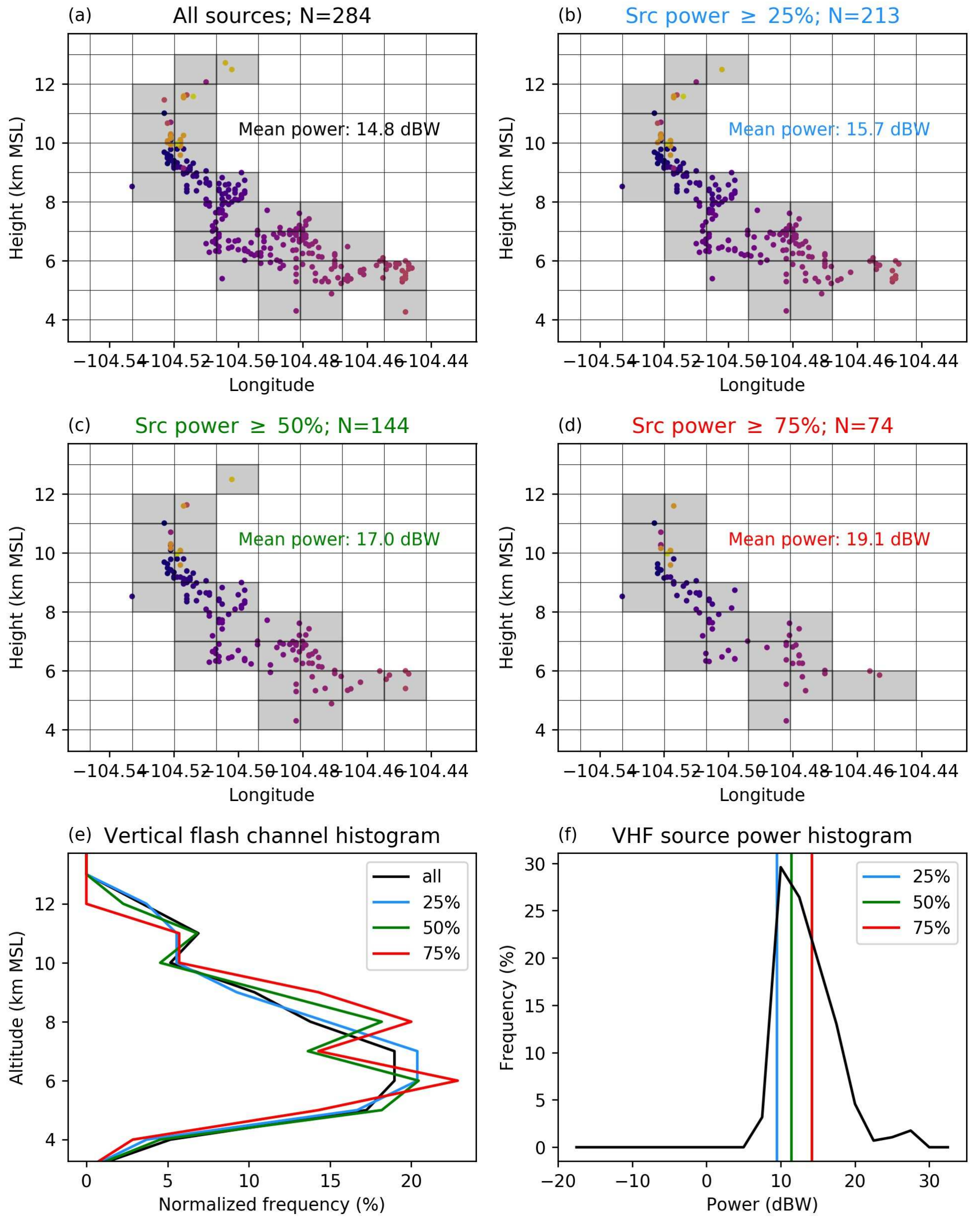

FIG. 3.1. (a) Illustration of the flash channel calculation. Points are LMA VHF sources colored by time (dark purple to light yellow). Grid boxes highlighted in gray are considered to contain a segment of the flash. (b) Same as (a) but with the lowest 25% of VHF source powers removed to simulate a decrease in LMA detection sensitivity. (c) Same as (a) but with the lowest 50% of VHF source powers removed. (d) Same as (a) but with the lowest 75% of VHF source powers removed. (e) Vertical distribution of flash channels from each representation of the flash in (a)-(d). (f) Distribution of VHF source powers (black) with the threshold powers shown in the vertical lines.

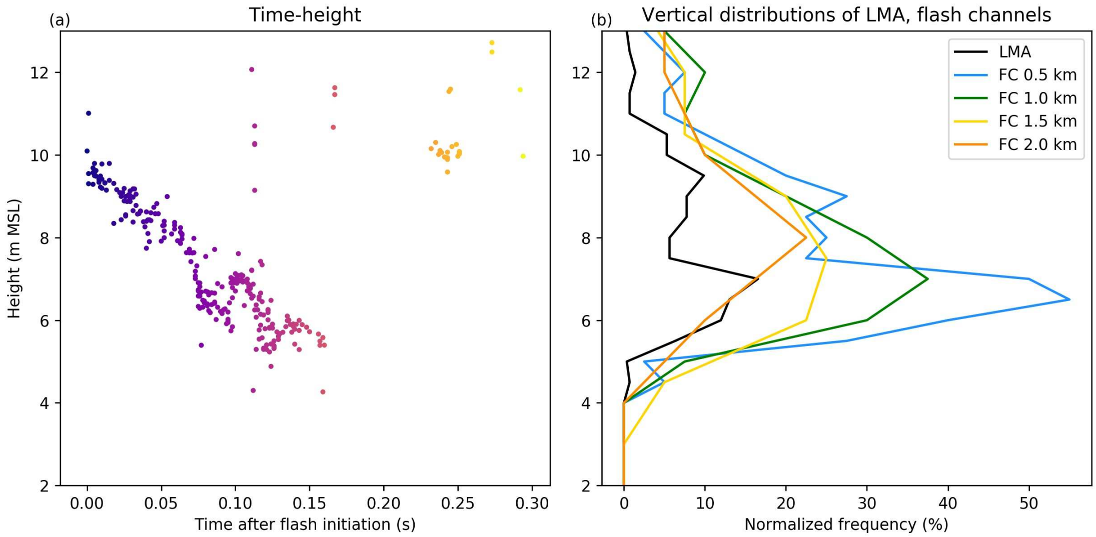

FIG. 3.2. (a) Time-height cross section of a sample flash. Points are LMA VHF sources colored by time (dark purple to light yellow). (b) Vertical distribution of LMA VHF sources (black), flash channels (FC) calculated on a 0.5 km (blue), 1.0 km (green), 2.0 km (orange) grid.

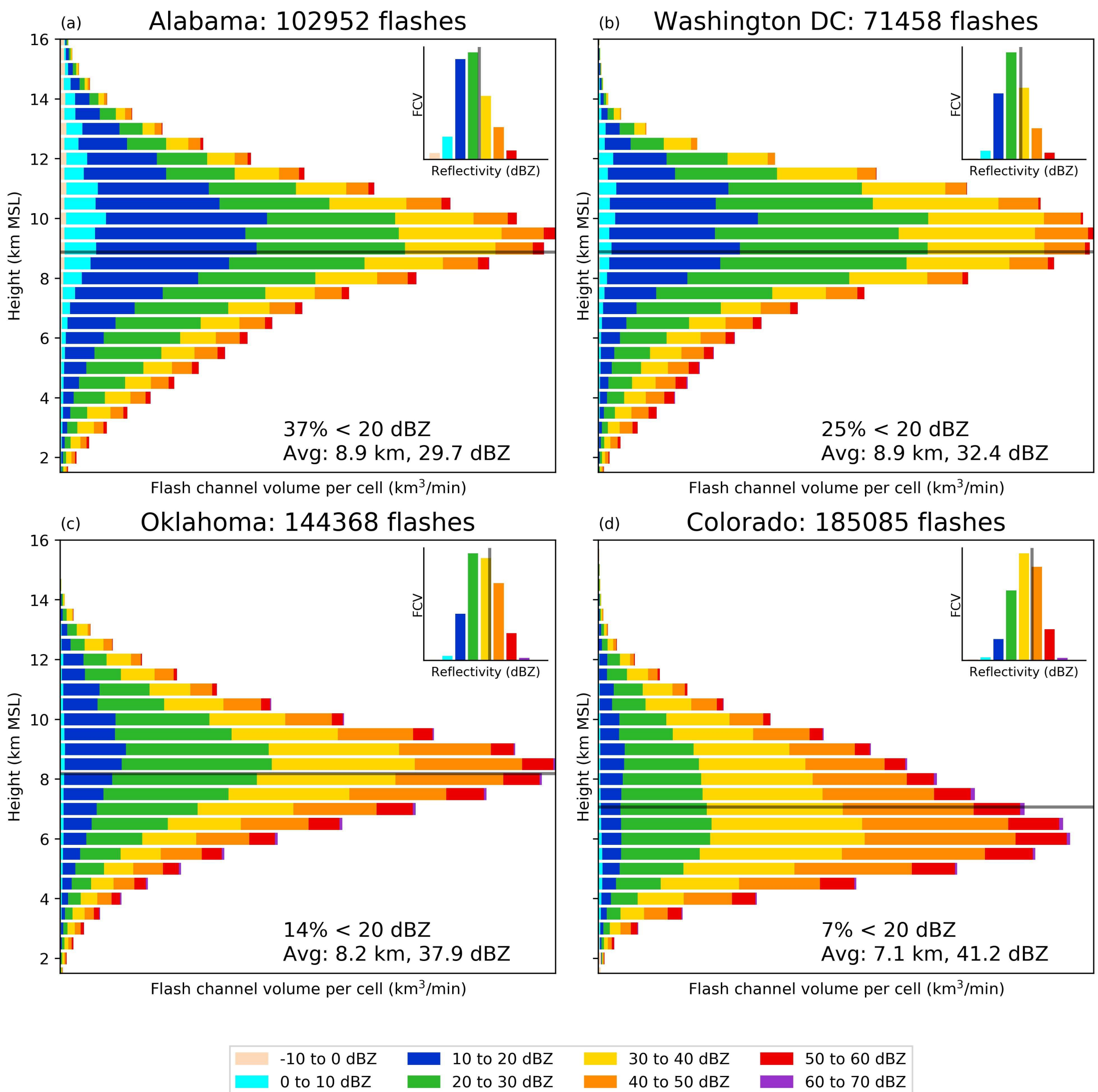

FIG. 3.3. Collocated flash channels calculated on a 0.5 km grid with respect to height and reflectivity for (a) in Alabama, (b) Washington, DC, (c) Oklahoma, and (d) Colorado. The inset panels show flash channel volume (FCV) with respect to reflectivity only. The average collocated flash channel height and reflectivity are shown in each panel as well as the fraction of calculated flash channels in regions of less than 20 dBZ.

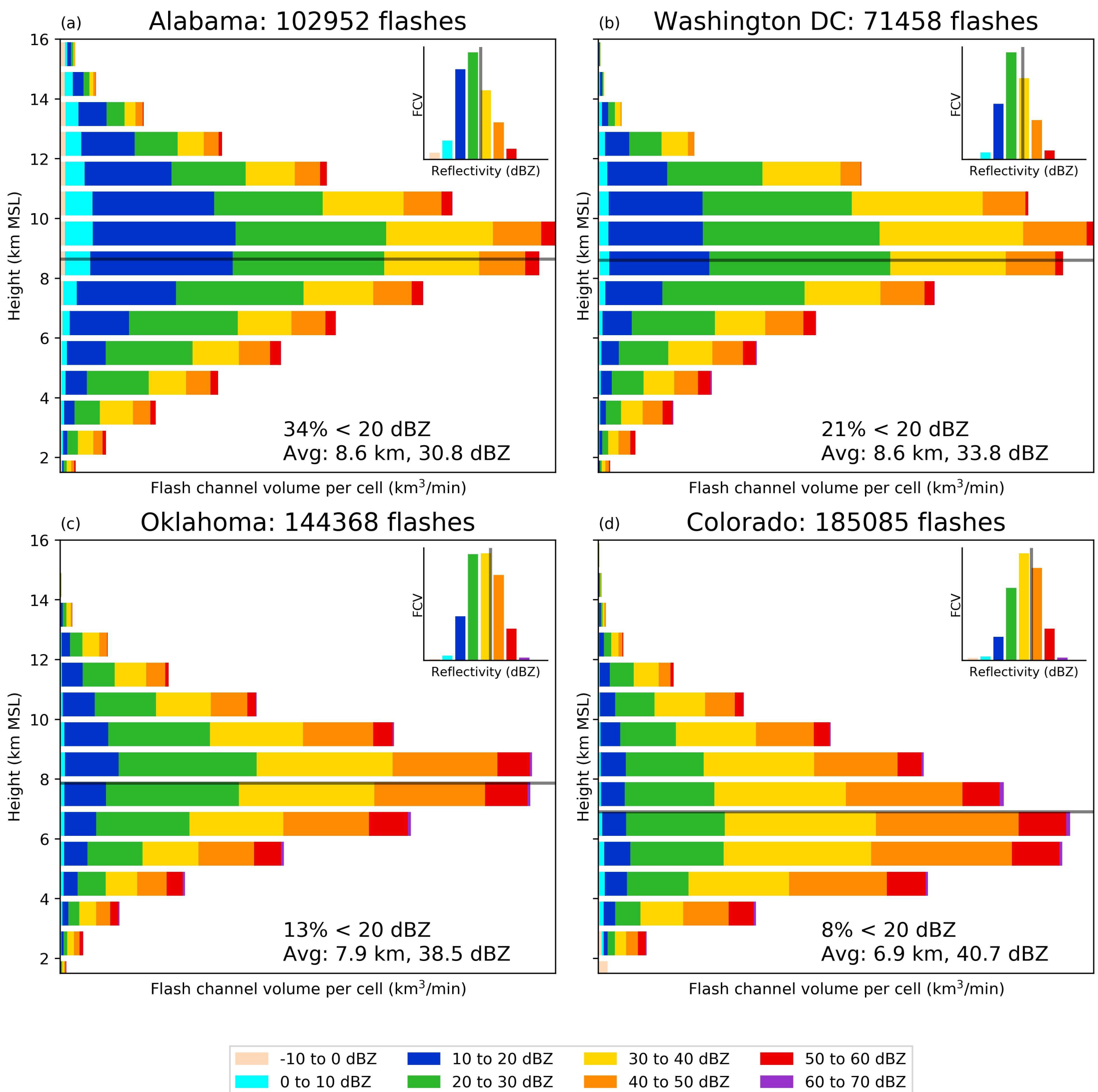

FIG. 3.4. Same as Figure 3, but the flash channels are calculated on a 1 km grid.

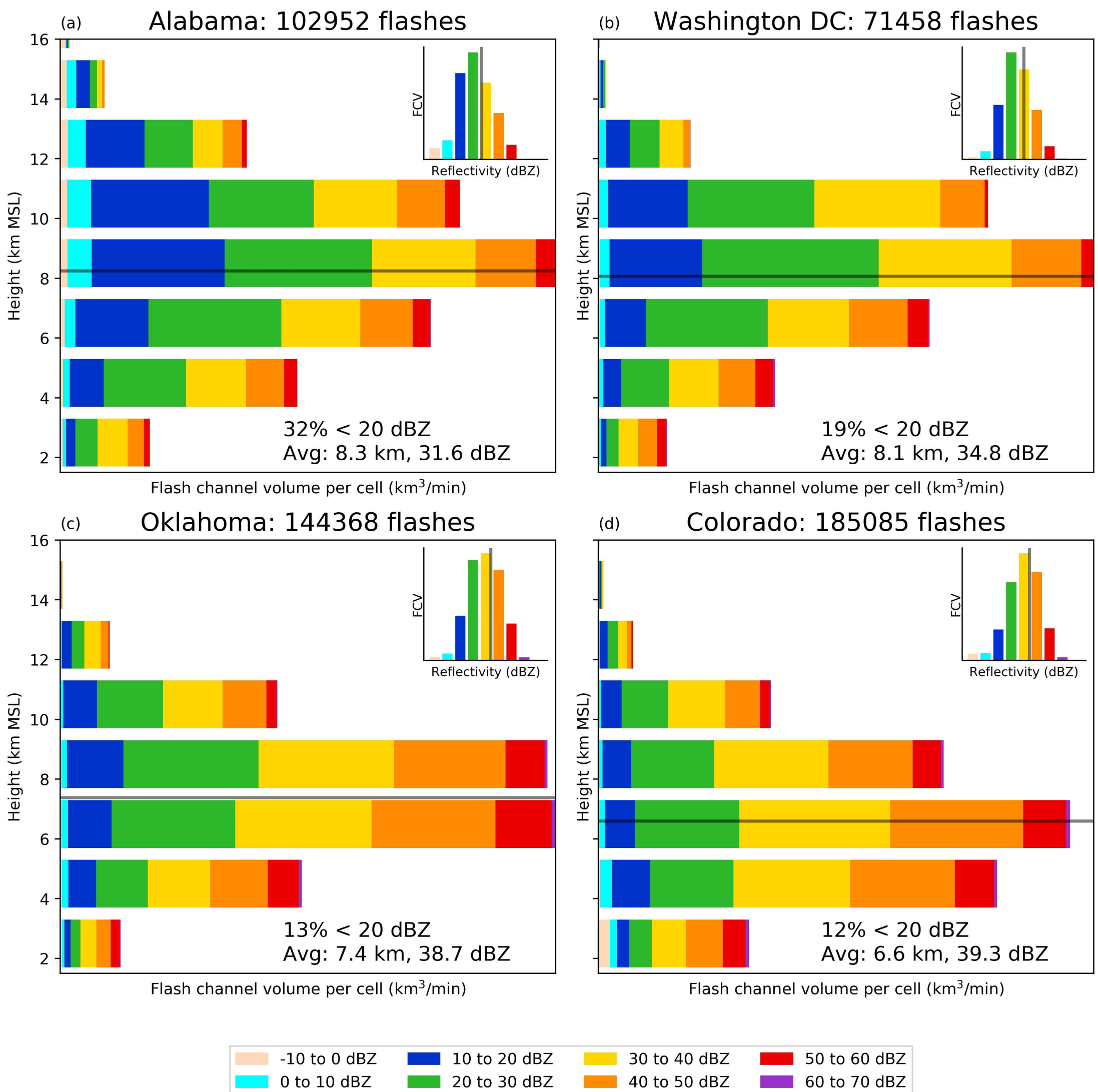

FIG. 3.5. Same as Figure 3, but the flash channels are calculated on a 2 km grid.

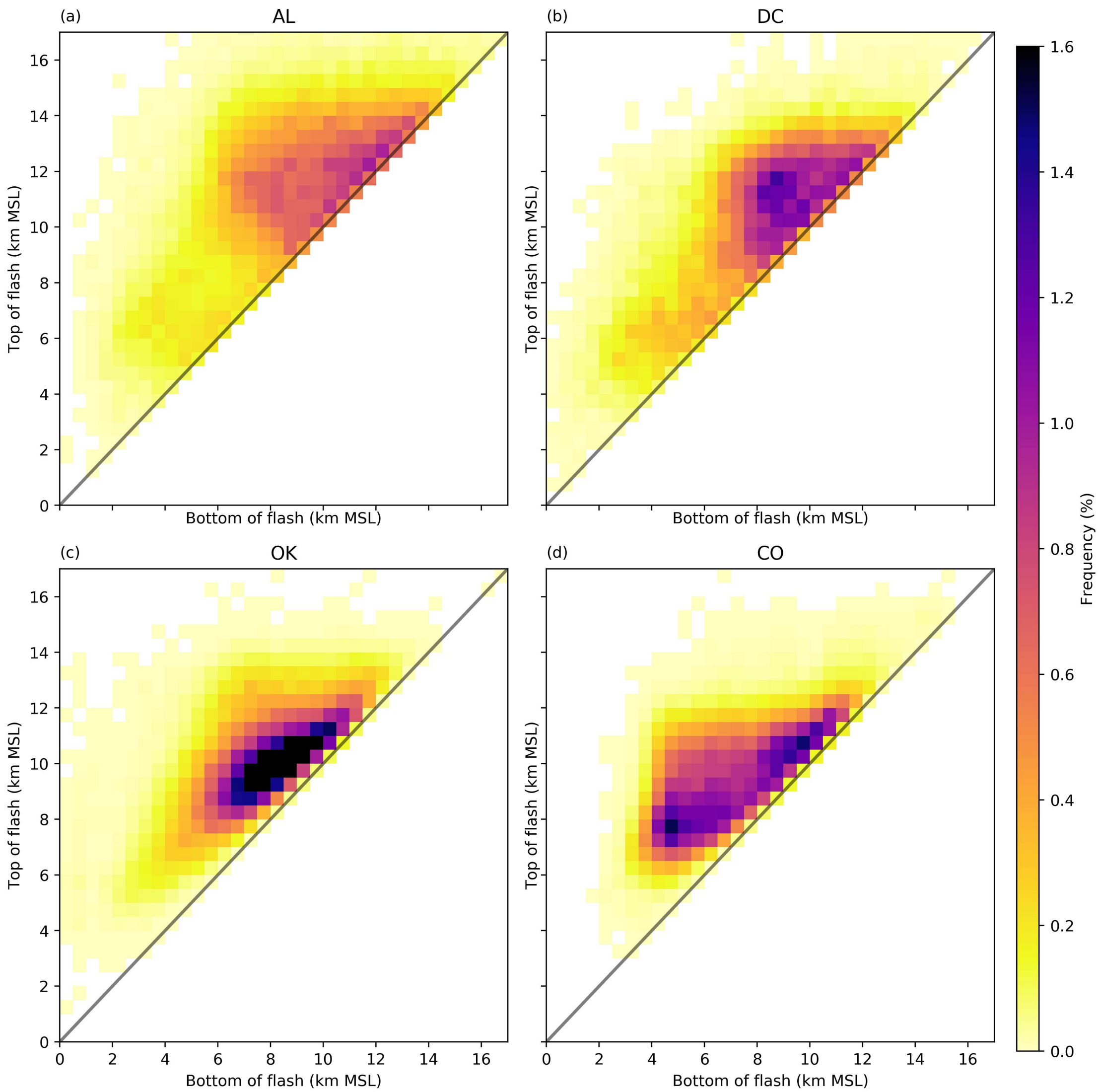

FIG. 3.6. Joint histogram of the bottom (as measured by the lowest altitude LMA VHF source) and the top (as measured by the highest altitude LMA VHF source) of each flash in (a) Alabama, (b) Washington DC, (c) Oklahoma, and (d) Colorado.

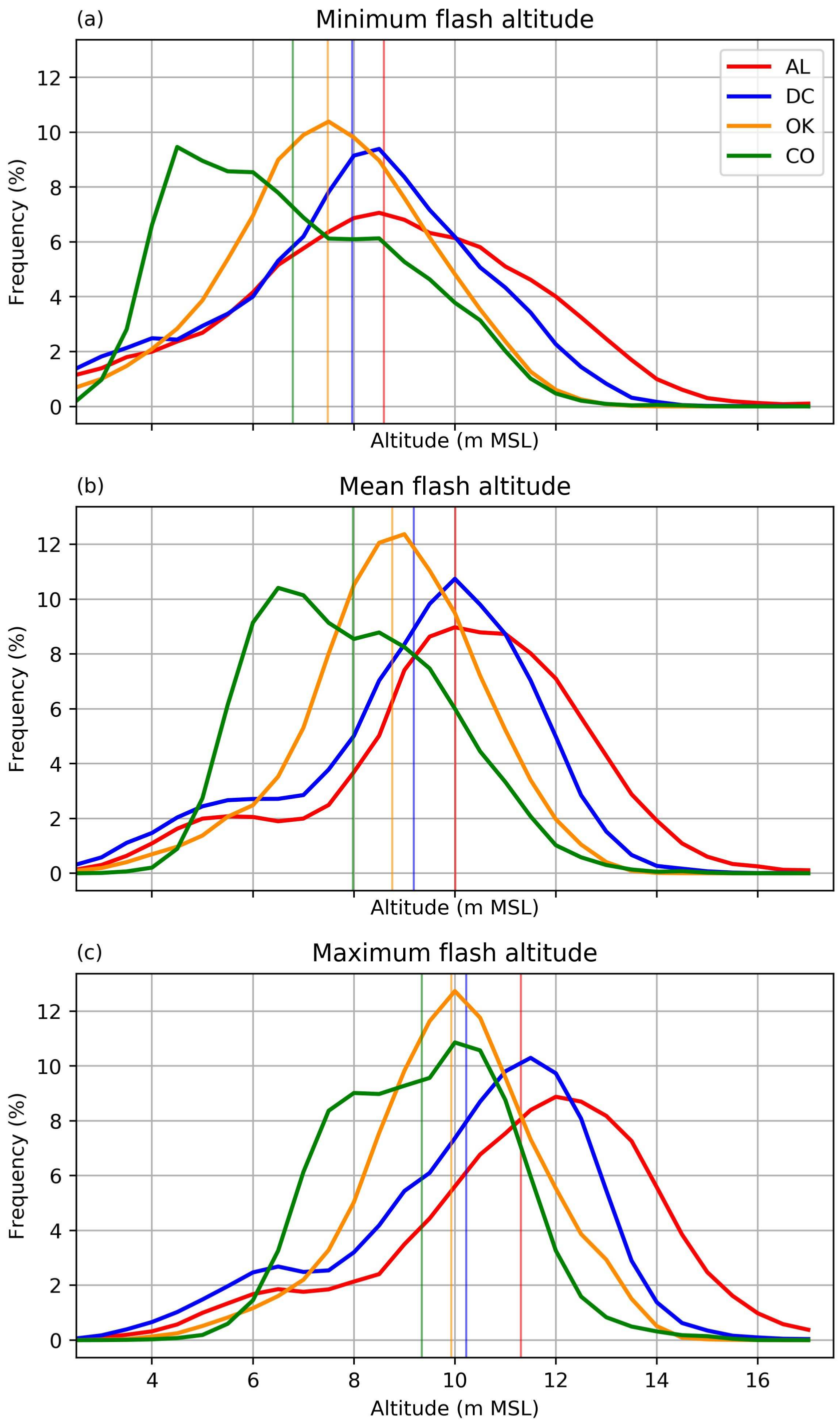

FIG. 3.7. Histograms of (a) minimum flash altitude, (b) mean flash altitude, (c) maximum flash altitude for each region (colors follow the legend). The vertical lines show the average values for each region using the same colors.

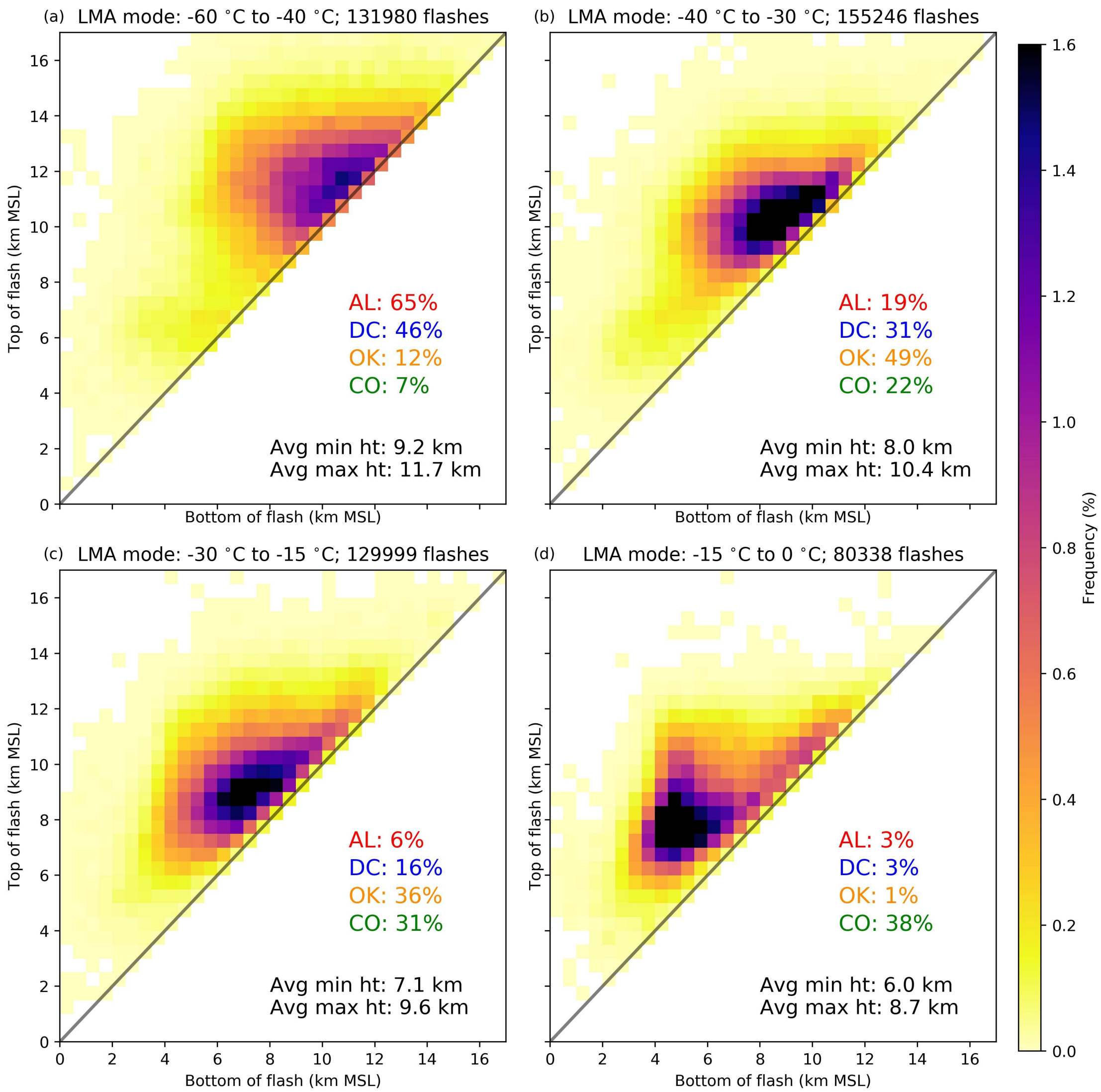

FIG. 3.8. Joint histograms of the bottom and top of each flash for (a) LMA mode temperatures between -60 °C and -40 °C, (b) LMA mode temperatures between -40 °C and -30 °C, (c) LMA mode temperatures between -30 °C and -15 °C, and (d) LMA mode temperatures between -15 °C and 0 °C. Average flash minimum and maximum heights and percentage of total flashes from each region indicated in each panel in addition to the percentage of the total flashes in each region that were produced by storms in a particular LMA mode temperature range. Note that for each region, the numbers do not add up to exactly 100% because a small fraction of flashes were produced by storms with LMA mode temperatures less than -60 °C or greater than 0 °C.

CHAPTER 4

Microphysical and dynamical processes associated with anomalous charge structures in isolated convection

4.1. Introduction

The non-inductive mechanism for cloud electrification is fairly well understood (e.g. Reynolds et al. 1957; Williams 1985; Saunders et al. 2006; Takahashi et al. 2017). The basic premise is that charge separation occurs when small (ice crystals) and large (graupel) ice particles undergo rebounding collisions in the presence of supercooled liquid water (SCLW), followed by separation of these particles by convective air motions (Williams 1985). Accordingly, charge structures and resultant lightning are directly linked to microphysical and dynamical processes (e.g. Smith et al. 2000; Williams et al. 2005; Carey and Buffalo 2007). The variability of storm-scale charge structures have been of interest for many years. Most convective storms possess a "normal" polarity charge structure with mid-level (approximately -10 °C to -30 °C) negative charge between regions of upper- and lower-level positive charge (Wilson 1920; Krehbiel 1986; Williams 1989). However, a class of storms, termed "inverted" or "anomalous", have been observed and are thought to possess mid-level positive charge between layers of lower- and upper-level negative charge (e.g. Lang et al. 2004; Rust et al. 2005; Wiens et al. 2005; MacGorman et al. 2008).

Early lab studies conducted by Reynolds et al. (1957) and Takahashi (1978) determined that the polarity and magnitude of charge acquired by graupel depended on temperature and the amount of SCLW (e.g. Jayaratne et al. 1983; Saunders and Peck 1998). Many subsequent laboratory studies have been conducted and the common thread is that cloud environments with warmer temperatures and larger SCLW contents promote positive charging of graupel (e.g. Saunders et al. 2006). Additionally, the magnitude of charge transfer per collision is enhanced as SCLW increases. Baker and Dash (1989)

This chapter is in review in the Journal of Geophysical Research: Atmospheres

hypothesized that the particle growing faster by deposition will transfer mass (accompanied by negative charge) to the particle it collides with, thereby acquiring net positive charge. Additionally, Avila and Pereyra (2000) found that graupel was more likely to charge positively when the distribution of supercooled cloud droplets was shifted to smaller sizes.

Because the mixed-phase charging zone spans from $\sim$ 0 °C to -40 °C for many thunderstorms, this implies that variability of SCLW may to first order determine the macroscale charge structure within a storm (Bruning et al. 2014). In normal polarity storms, graupel in the lower portion of the thunderstorm (temperatures between $\sim$ 0 °C and -15 °C) acquires positive charge regardless of SCLW content, based on laboratory studies (e.g. Takahashi 1978). In the storm mid-levels, graupel acquires negative charge at most liquid water contents ($\sim$ 1-4 g m^{-3}, depending on the particular laboratory study). The ice crystals that acquire positive charge from collisions with graupel are lofted to the top of the storm comprising the upper-level positive charge. Conversely, if a storm has large SCLW contents ($> \sim$ 1-4 g m^{-3}) in the mid-levels, graupel would become positively charged in the mid-levels, while ice crystals would acquire net negative charge, leading to upper-level negative charge region. Therefore, it follows that if the charge structure of a storm, particularly the charge of the mid-level graupel, can be determined or inferred, relative SCLW amounts can also be inferred.

The dependence of charge structures on SCLW content has sparked interest in the dependence of supercooled water on more fundamental processes. In this manner, charge structures can be used as a lens to study fundamental microphysical and dynamical processes in storms. Negatively charged mid-level graupel in normal polarity storms is likely due to weaker vertical velocities or rapid liquid water depletion rates (Bruning et al. 2014). Fallout of warm-phase precipitation from robust collision-coalescence processes and higher rates of entrainment have been observed for tropical convection that commonly have relatively narrow updraft widths (e.g. LeMone and Zipser 1980; Bringi et al. 1997; Williams and Stanfill 2002) and deep warm cloud depths (WCD; Atlas and Ulbrich 2000; Stolz et al. 2015), resulting in rapid depletion of available liquid water. Smaller liquid water depletion rates at warmer temperatures permit larger SCLW contents and increased likelihood of positive charge at mid-levels. Williams et al. (2005) claimed instability promotes strong updrafts that in turn lead to large SCLW in the mixed-phase region and positive charging of graupel in anomalous storms. Additionally, slow depletion rates are expected in storms with high cloud base heights (CBH) or shallow WCDs, where parcels traverse through the warm cloud region in a short time, resulting in decreased warm-phase precipitation growth and fallout (Carey and Buffalo 2007; Fuchs et al. 2015). Additionally, stronger and broader updrafts are less prone to entrainment, which reduces SCLW contents (Williams and Stanfill 2002; Williams et al. 2005; Bruning et al. 2014).

Determining charge structures and supercooled water contents in nature is difficult. Therefore, researchers have been forced to rely on proxy data to infer charge structures within storms. Most studies have used the fraction of cloud-to-ground (CG) flashes that are positive polarity (+CG; Wiens et al. 2005). This is because most CG flashes originate in the mid-levels (Krehbiel 1986), so the polarity of CG flashes provides some information about the charge residing in the mid-levels of a storm. In normal polarity storms with mid-level negative charge, nearly all the CG flashes are negative polarity (e.g. Krehbiel et al. 1979). In intense anomalous polarity storms, the fraction of +CG flashes is elevated, near 100% in some cases (e.g. Reap and MacGorman 1989; Stolzenburg 1994; Lang et al. 2004). However, there are other scenarios that can lead to increased production of +CG fractions such as precipitation unshielding (Carey and Rutledge 1998; Pawar and Kamra 2007) and the tilted dipole mechanism (Brook et al. 1982). Furthermore, intense anomalous polarity storms often produce relatively few CG flashes (e.g. Tessendorf et al. 2007), which can make it difficult to use CG polarity to diagnose storm polarity in some situations.

Despite the extensive research into the variability of storm charge structures, our understanding is far from complete. Furthermore, the lack of understanding of anomalous storms highlights the gaps in our knowledge of microphysical and dynamical processes in storms. The goal of the present study is to improve our understanding of the processes leading to anomalous charge structures in thunderstorms by more directly investigating the links between the environment, microphysics, dynamics and charge structure. This investigation will be carried out by comparing relevant microphysical and dynamical quantities between two storm populations: normal polarity storms (in northern Alabama) and anomalous polarity storms (in eastern Colorado). The study is organized as follows: section 4.2 will lay out the array of data types and the methodology of the study, section 4.3 enumerates the environmental, microphysical and dynamical differences between the normal and anomalous polarity storms in the dataset, and section 4.4 discusses the insights that can be gained from this paper and the robustness of those insights.

4.2. Data and Methods

4.2.1. *Radar*

A majority of cases in this study come from the Deep Clouds and Convective Chemistry Experiment (DC3; Barth et al. 2015), which took place during the summer of 2012 in northern Alabama, eastern Colorado, central and western Oklahoma. One of the main objectives of DC3 was to understand the convective transport of various trace gasses from the boundary layer to the upper troposphere. Therefore, there are numerous cases with microphysical and three-dimensional dynamics available in the dataset. Polarimetric radar data in the Colorado region comes from the CSU-Chicago Illinois (CSU-CHILL) polarimetric S-band radar (Bringi et al. 2011). Polarimetric data from the Alabama region comes from UAH ARMOR polarimetric C-band radar (Petersen et al. 2007). Both radars provided the standard polarimetric quantities: horizontal reflectivity (Z_H), differential reflectivity (Z_{DR}), cross-correlation coefficient at zero lag (ρ_{HV}) and differential phase (ϕ_{dp}). See Doviak and Zrnić (1993) and Bringi and Chandrasekar (2001) for more details about polarimetric radar. During DC3, CHILL operated in an alternating H/V mode allowing for the collection of linear depolarization ratio (LDR) in addition to the standard polarimetric quantities. Specific differential phase (K_{DP}) was calculated using the DROPS algorithm developed by Wang and Chandrasekar (2009). See Mecikalski et al. (2015) for more information regarding post-processing on ARMOR data in Alabama. Before analysis, all CHILL (ARMOR) radar quantities were interpolated to a 0.5 km x 0.5 km x 0.5 km (1.0 km x 1.0 km x 1.0 km) Cartesian grid using the NCAR Sorted Position Radar INTerpolator (SPRINT; Mohr and Vaughan 1979; Miller et al. 1986). Prior to interpolation, the velocity data were unfolded with National Center for Atmospheric Research (NCAR) solo3 software (Oye and Case 1995).

Combinations of various polarimetric quantities can provide information about hydrometeor shape, size, and phase, which can be used to estimate microphysical quantities of interest. Recently, a number of microphysical retrieval algorithms developed at CSU have been compiled in a Python-based package called CSU-Radartools, which is publicly available on Github (Lang et al. 2016; Mroz et al. 2017). The retrievals used in this study include fuzzy-logic inferred hydrometeor identification (HID; Dolan et al. 2013), water and graupel ice mass mixing ratios (Carey and Rutledge 2000; Cifelli et al. 2002), and blended rain rates (Bringi and Chandrasekar 2001; Cifelli et al. 2011). These calculations were carried out on the Cartesian grid after the data was quality controlled.

Dual-Doppler scanning for the CO cases was performed in coordination with the S-band CSU-PAWNEE Doppler radar (Lang et al. 2014, 2016; Basarab et al. 2015). For the AL cases, dual-Doppler scanning was performed in coordination with the S-band KHTX NEXRAD radar (Barth et al. 2015). For the DC3 field project, the majority of the scanning was done with a 5 to 6 minute dual-Doppler update time. The 3D wind synthesis from the Doppler radial velocities was performed with the NCAR Custom Editing and Display of Reduced formation in Cartesian Space (CEDRIC; Mohr et al. 1986). It should be noted that the errors in the vertical velocity (W) obtained by this methodology are typically on the order of 1-3 m s^{-1}, depending on the magnitude of vertical velocities and methods used to estimate the three-dimensional wind fields (e.g. Potvin et al. 2012; Calhoun et al. 2013). For further details of the DC3 dual-Doppler analysis and radar configurations, see Barth et al. (2015) and Basarab et al. (2015).

4.2.2. *CSU Lightning, Environmental, Aerosol, and Radar (CLEAR) Framework*

To synthesize the large amount of data from an array of data types and number of cases, we employ the CSU Lightning, Environmental, Aerosol, and Radar (CLEAR) automated case study framework (Lang and Rutledge 2011). CLEAR begins by objectively identifying storm cells using contiguous regions of reflectivity based on multiple, tunable reflectivity and size thresholds. In this study, the composite reflectivity field was queried for regions of 30 dBZ that were larger than 20 km^2, that also had an area of 40 dBZ that was larger than 10 km^2, following Fuchs et al. (2015). Each identified cell for each radar volume scan time is treated as an individual and independent storm observation. Restricting the analysis to isolated cells avoids complicating factors such as mergers and influences from nearby convection. Additionally, storm metrics, such as flash rates and updraft volumes, are easier to interpret if storms are isolated.

Ellipses are fit to each identified cell, then cells are tracked by spatially matching ellipses in consecutive radar volumes. Once cells are identified and tracked, other types of data (such as lightning flash rates and environmental soundings) can be attributed to cells by spatial and temporal matching. Cell identification and tracking was performed on 6 days in Alabama and 5 days in Colorado. It should be noted that all cells included in this study were located in dual-Doppler lobes and were within 125 km of their respective LMA network. Once the other data types were attributed to the cells, a database was constructed based on inferred storm-scale charge structure (section 4.2.5). This database was queried for the purposes of single case studies as well as bulk statistical analysis.

4.2.3. *Lightning Mapping Array (LMA)*

This study utilizes the Lightning Mapping Array (LMA) networks in northern Alabama and northeastern Colorado. LMA networks are comprised of 6 or more radio antennas that detect and locate very high frequency (VHF) radiation produced by discontinuous breakdown of lightning channels (Rison et al. 1999; Thomas et al. 2004). During DC3, the northern Alabama network was comprised of

11 stations in an approximate circle centered near Huntsville, Alabama and 2 additional stations near Atlanta, Georgia. The northeastern Colorado network is comprised of 15 stations centered near the Greeley, Colorado airport (KGXY). Both networks have reliable detection of LMA sources and flashes past 100 km from the respective network centers (Koshak et al. 2004; Chmielewski and Bruning 2016; Fuchs et al. 2016). To remove any potential noise from the LMA data, only sources that had at least 7 solution-contributing stations and a chi-squared fit value less than 2.0 were attributed to cells. LMA sources that satisfied filtering criteria were attributed to the appropriate cell if they were located within the spatial extent of the cell, defined by the composite reflectivity. LMA sources outside of any identified cells were attributed to the nearest cell if they were less than 10 km from a cell otherwise they were not considered in the analysis.

4.2.4. *Flash counting*

LMA networks do not detect lightning flashes. Instead, LMAs detect sources of VHF radiation produced by lightning propagation (Rison et al. 1999). Therefore, further processing on VHF source data is required to obtain information about the physical lightning flashes detected by LMAs. This study uses the flash algorithm developed by Bruning (2013) and Fuchs et al. (2015, 2016), which clusters VHF sources that are close together in space and time. Once VHF sources are initially clustered into flashes, some quality-control filtering is conducted. In the Colorado (Alabama) region, any analyzed flashes with less than 10 (2) sources are discarded from the analysis, following Fuchs et al. (2015, 2016). Furthermore, any flashes that are more than 10 km from any identified cells are discarded from the analysis, similar to the LMA source attribution. Once the flashes are filtered and attributed, information such as location, duration, and spatial extent of each flash can be estimated. Furthermore, statistics such as cell flash rate can be calculated. For more information on the flash clustering algorithm used in this study, see Fuchs et al. (2016).

4.2.5. *Charge structure inference*

In addition to flash-level information, LMA data can also be used to infer the charge structure within a thunderstorm, because LMA detection of leader propagation depends on polarity. Flashes typically initiate in regions of strong electric fields between positive and negative charge regions. Leaders of both polarities propagate away from the initiation point (MacGorman et al. 2001; Coleman et al. 2003; Maggio et al. 2005). Propagation of negative leaders into positive space charge is more discontinuous than propagation of positive leaders into negative space charge (Rison et al. 1999). This usually results in more LMA-detected VHF sources from positive charge regions. Therefore, the height of dominant positive charge is inferred to be near the altitude of the LMA source density maximum within a storm (Wiens et al. 2005; Lang and Rutledge 2011). By this method, the mode of LMA source density in a normal (anomalous) polarity storm is typically near -40 °C (-20 °C). Figure 2b in Fuchs et al. (2015) shows a schematic depiction of the relationship between charge structures and LMA source densities. It should be noted that since the positive leader propagation into negative charge is not well-resolved by the LMA, it is very difficult to know the precise location of the negative charge by this method (Wiens et al. 2005).

4.2.6. *Environment*

Hourly model reanalysis data was attributed to each identified cell for relevant environmental data and some polarimetric retrievals. Model data was used to attempt to capture spatial and temporal variability of the environment, particularly since many of the cases in Alabama occurred between 18Z and 21Z, several hours before the 00Z National Weather Service daily radiosonde launch. The Rapid Update Cycle (RUC; Benjamin et al. 2004) 13-km reanalysis was used for storms that occurred before 1 May 2012, thereafter the Rapid Refresh (RAP; Benjamin et al. 2006) 13-km reanalysis was used. A representative inflow point for each storm was calculated by selecting the nearest grid point approximately 40 km from the storm along the low-level upwind vector, similar to Fuchs et al. (2015). All relevant quantities at that grid point were temporally interpolated to the radar volume scan time before being attributed to the cell (Lang and Rutledge 2011). The reanalysis data included fundamental atmospheric quantities such as temperature, wind, and pressure as well as standard calculated quantities such as freezing height and convective available potential energy (CAPE). Additional variables were calculated based on the fundamental quantities, such as normalized CAPE (NCAPE; Blanchard 1998), CBH (Bradbury 2000) and WCD (Carey and Buffalo 2007).

4.3. Results

4.3.1. *Overview*

The analysis method in this study is to compare the environmental, electrical, microphysical, and dynamical characteristics of 73 individual radar volumes from 9 normal polarity cases from northern Alabama and 47 radar volumes from 6 anomalous polarity storms over northeast Colorado. The focus of this study is on isolated convective storms, since those are the easiest to interpret, free from interactions with adjacent cells. Recall that every storm at each radar volume time is treated as a separate individual cell. The macroscale charge structure was defined by the temperature of the LMA mode (Sec. 4.2). Normal polarity storms were defined to have an LMA mode temperature of < -30 °C. Storms with an LMA mode temperature > -25 °C were classified as anomalous.

4.3.2. *Case studies*

4.3.2.1. *Normal Alabama case study*

The archetypal normal polarity case from Alabama occurred on 21 May 2012 during the DC3 field project. The storm developed around 2000Z, intensified between 2000-2030Z before dissipating around 2100Z. The storm was located about 70 km northeast of the ARMOR radar site located near Huntsville, AL and was in the northern dual-Doppler lobe. Figure 4.1 shows a synopsis of the storm: both a snapshot during the mature phase of the storm at 2023Z (Figure 4.1a) and timeseries of the electrical, dynamical and microphysical characteristics of the storm (Figures 4.1 b-d). Figure 4.1a shows an x-z cross section of radar reflectivity along with updraft speeds and LMA source density. Maximum updrafts in this storm were approximately 12 m s^{-1} and were situated around 6 km MSL (-20 °C). The updraft remained strong to approximately 9 km MSL (equilibrium level was $\sim$ 12 km MSL). Radar reflectivities in the main updraft ranged from 35-55 dBZ. Additionally, the updraft primarily contained large raindrops, hail, and high-density graupel as inferred by polarimetric radar observations (not shown). The LMA source density maximum, used as a proxy for positive charge location, is situated between -30 °C and -40 °C near the top of the updraft.

During the period from 2000 to 2016Z, the flash rate increased from 0 to 3 min^{-1} (Figure 4.1b). The charge structure, as inferred from the vertical LMA source density, is noisy but suggests an active lower positive charge around 4-5 km MSL (Figure 4.1b). Mecikalski et al. (2015) and Carey et al. (2016) showed substantial âĂŞCG activity during this time, suggesting normal charge polarity. Updraft speeds are modest during this time, increasing from $\sim$ 6 m s^{-1} to 12 m s^{-1} (Figure 4.1d). At 2008Z, the strongest updraft speeds were located below $\sim$ 6 km, which is situated near the storm maximum graupel ice mass content (Figure 4.1c). The collocation of graupel and positive charge at lower levels is shown by the Pearson correlation coefficient relating the vertical distributions of LMA source density and average graupel ice mass (Figure 4.1c). The correlation is between 0.4 and 0.6 during the early stages of the storm.

An abrupt shift in charge structure occurs around 2022Z, when the dominant positive charge appears around 8-10 km MSL or approximately -35 °C. After the shift in storm charge structure to upper positive charge, the maximum graupel ice mass remains below approximately 8 km MSL, which leads to a graupel ice mass/LMA source correlation near 0 and suggests that graupel and positive charge are not collocated which implies graupel does not carry positive charge, rather negative charge. Maximum updrafts increase from ~ 12 m s^{-1} at 2010Z to ~ 18 m s^{-1} at 2025Z in concert with an increase in the altitude of maximum updraft strength. This updraft burst is coupled with an increase in the altitude of the maximum LMA source density, perhaps lofting the positively charged ice crystals to even greater heights. After 2033Z, the flash rates and average ice mass decrease as updraft speeds weaken. The storm dissipated shortly thereafter. Additional details regarding the radar and lightning properties of this storm can be found in Mecikalski et al. (2015) and Carey et al. (2016).

4.3.2.2. *Archetypal anomalous CO case*

The presentation of the information about the archetypal anomalous CO case in Figure 4.2 follows the same format as Figure 4.1, facilitating direct comparisons between the cases. The storm was located approximately 60 km to the east-northeast of CHILL and was in the eastern dual-Doppler lobe formed by CHILL and the CSU-Pawnee S-band radar. Figure 4.2a indicates the updraft is much stronger in this case, with peak values greater than 20 m s^{-1}. The updraft extends to about 13 km MSL, just above the sounding-derived equilibrium level (12.5 km MSL). The updraft contains largely rain below the melting level and low-density graupel above the freezing level (not shown), and reflectivities range from 20-40 dBZ, although much higher reflectivities exist just to the east of the main updraft. The maximum in LMA source density is located near 6 km MSL ($\sim$ -10 °C). The LMA source density maximum is located in a region of weak vertical motion, adjacent to the strong updraft.

This case was first identified by the objective tracking algorithm at 2155Z and quickly produced updrafts > 20 m s^{-1}. Rapidly the storm became electrically active, as flash rates increased from ~ 2 flashes min^{-1} to 20 flashes min^{-1} at 2212Z. During this time, the dominant positive charge region was located in the mid-levels of the storm between 5-8 km MSL, collocated with a maximum in average graupel ice mass. This resulted in a high graupel ice-mass/LMA source correlation that ranges from 0.6 to 0.9 in the early and middle stages of the storm lifecycle. The high correlation indicates that graupel is collocated with positive charge, from which we infer that the mid-level graupel is the dominant carrier of positive charge. The storm remained intense with strong updrafts, high flash rates and mid-level positive charge until 2243Z, when abrupt changes were observed.

The abrupt shift in the storm charge structure around 2243Z is peculiar, because the dissipation of the mid-level positive charge is coincident with a dissipation in mid-level graupel ice mass. Furthermore, as the upper-level positive charge becomes dominant, it is accompanied by the development of an upper-level graupel ice mass maximum, which is indicated by the graupel ice mass / LMA source correlation staying above 0.5. Perhaps this indicates that the positive charge above -30 °C is carried by graupel, which acquired its charge in situ or was transported upward from lower levels in the storm. This abrupt charge structure shift is accompanied by an increase in updraft strength at high altitudes, and a flash rate increase of $\sim$ 100%. After 2243Z, the updrafts weakened considerably, coincident with descent of the upper positive charge, descent of the graupel ice mass and decrease in storm flash rates.

4.3.3. *Statistical analysis*

4.3.3.1. *Electrical characteristics*

To gain more insights regarding differences between normal and inverted storms, we merge all observations (multiple storms, multiple sample times) for each charge structure classification and analyze them statistically. It is important to note that the observations are comprised of storms from the developing, mature and dissipating stages. However, the strict threshold imposed when identifying cells with the objective algorithm tend to favor cells near their mature phase (Fuchs et al. 2015). The vertical distribution of LMA source densities from each storm observation of (AL) normal cases are shown in Figure 4.3a. By design, the peak of each distribution is < -30 °C. Note that few distributions have modes colder than -40 °C. Note that some profiles have a smaller, secondary maximum at warmer temperatures, consistent with a lower positive charge (Wiens et al. 2005) and the presence of a normal polarity tripole. Conversely, the vertical distributions of LMA source densities from the anomalous (CO) cases are shown in Figure 4.3b. The spread in LMA mode temperatures is larger in these cases, with some distributions peaking around -25 °C and some around -10 °C. In fact, the distributions bear resemblance to both schematics in the middle and right panels of Figure 2b from Fuchs et al. (2015). Bruning et al. (2014) argue that the charge structures exist in a continuum, rather than in the binary classifications of normal and inverted. Indeed, the storms in the middle and right panels of Figure 2b from Fuchs et al. (2015) may be the same, but the charge centers may be vertically shifted.

The cell flash rate distributions are shown in Figure 4.3c. Median cell flash rate (total lightning) for the normal AL samples is approximately 5 flashes min^{-1}, whereas the median cell flash rate from the anomalous CO samples is roughly 10 flashes min^{-1}. For both storm populations, approximately 80% of samples have cell flash rates < 20 min^{-1}. The maximum cell flash rate for the normal AL population is 45 min^{-1}; 75 min^{-1} for the anomalous CO population. A comparison with the large statistical sample of isolated convective cells in Figure 7 of Fuchs et al. (2015) shows that an overwhelming majority of isolated convective cells in the vicinity of the northern Alabama LMA have LMA mode temperatures < -35 °C and an overwhelming majority of the CO events have LMA mode temperatures > -25 °C. Furthermore, the flash statistics in Figure 5a from Fuchs et al. (2015) shows that a cell flash rate of 5 flashes min^{-1} is approximately the 70th percentile of cell flash rates in the larger sample in northern Alabama. Cell flash rates of 10 min^{-1} in northeast Colorado are about the 60th percentile of flash rates from the larger Fuchs et al. (2015) sample. Hence, we argue the cells in this study are largely representative of their respective local climatologies based on the combination of cell flash rates and LMA mode temperatures compared with the larger Fuchs et al. (2015) dataset.

4.3.3.2. *Environmental characteristics*

The environment influences storm microphysics and dynamics, which in turn, impact charge structures and lightning. Therefore, it is important to understand the environmental context of these storm populations. Figure 4.4 shows the distributions of relevant thermodynamic quantities from the present storm populations. CAPE and normalized CAPE (NCAPE) are both higher in the normal AL storms than the anomalous CO storms by about 40%. The CO storms have much higher cloud base heights and much shallower warm cloud depths in comparison to the AL data. The precipitable water values are about 50% higher in AL storms than in the CO storms. It is perhaps somewhat surprising that the adiabatic water content is similar in both storm populations. This apparent inconsistency can be explained by the higher surface temperatures in anomalous CO cases in combination with similar surface dew points between the populations. Higher surface temperatures with the same surface dew points will result in larger surface dew point depressions, which directly contributes to cloud base height (Bradbury 2000). However, adiabatic water content depends only on surface dew point and pressure (Pruppacher et al. 1998). It is also surprising that mid-level relative humidity (as measured by average humidity between 600 mb and 500 mb) is higher in the CO samples. This may have implications for parcel dilution by entrainment. Surface to 6 km shear values are slightly larger in AL samples, but the values in both populations are modest, with respect to severe weather production (Weisman and Klemp 1982). The convective inhibition (CIN) is substantially larger in the CO samples, typically considered an ingredient for more intense convection (Riemann-Campe et al. 2009; Rasmussen and Houze Jr 2016). Finally, the equilibrium level heights are not significantly different between the two populations.

4.3.3.3. *Dynamics*

With the electrical and environmental characteristics laid out for context, we can examine the impacts of the environment on convective cell dynamics and the impacts of the overall cell dynamics on charge structures. Figures 4.5a,b show vertical motion composite contoured frequency by altitude diagrams (CFADs; Yuter and Houze Jr 1995) by averaging the CFAD for each observation in both storm populations. Probabilities corresponding to each W bin are color filled for each gridded altitude. Immediately apparent for each population is that most grid points have vertical velocities near zero at all altitudes, indicating that the majority of grid points within these storms are not associated with rising or sinking air, consistent with previous studies (e.g. Yuter and Houze Jr 1995). One of the largest differences between the vertical velocities in the composite CFADs are that updrafts greater than 20 m s^{-1} exist in the anomalous CO population, but are not present in the normal AL population. This indicates that the peak updrafts are stronger in the anomalous CO cases than they are in the normal AL cases, a signature evident in the maximum W cumulative distributions as well (Figure 4.5c). The average updraft and downdraft profiles are shown in the line plots in Figures 4.5a,b. Average updrafts in the anomalous CO population are stronger than the normal AL population by approximately 100%. Furthermore, average updrafts and downdraft speeds in AL normal cases are nearly constant with height, while the average updrafts and downdrafts in anomalous CO cases show significant variations with height. The average updraft speed in anomalous CO storms peaks around 8 km MSL, which is also where peak updraft speeds are maximized.

Updraft volumes > 5 m s^{-1} (UV5) and > 10 m s^{-1} (UV10) are also much larger in the anomalous CO samples. The median value of UV5 is approximately a factor of three larger in the CO storms compared to the AL storms. The message of these comparison statistics is clear: the updrafts in the anomalous CO storms are stronger and larger in volume than the normal AL storms. Given that the instability, as measured by CAPE and NCAPE in this study, is larger in the normal AL storms, one would expect stronger updraft speeds in storms with larger instability based on simple parcel and buoyancy theory. These results (Figure 4.5) indicate that parcel theory alone is insufficient to explain the larger and stronger updrafts in anomalous CO storms. We suggest the higher CBHs and shallower WCDs in the anomalous CO cases reduce rainout loss of liquid water, which would allow for more robust mixed-phase processes and additional latent heat release. Since the updrafts in the anomalous CO storms are larger in volume, we would expect them to be wider as well, which may act to reduce entrainment and dilution of instability and SCLW.

To investigate the issue of updraft width and potential impacts on entrainment, Figure 4.6 breaks down the areas of updrafts, downdrafts and neutral air with respect to temperature. Figure 4.6a shows the fraction of the storm area containing updrafts (defined as W > 5 m s^{-1}) for each storm in both populations. Note that the results were not particularly sensitive to the W threshold choice. The thick line indicates the median value of updraft fraction for all storms in the population as a function of temperature, with the interquartile range filled. Stark differences in updraft area fraction exist between the populations at nearly every temperature level. Most notably, the anomalous CO storms have much wider updrafts than the normal AL storms. The areas of the updrafts in the anomalous CO storms increase from 0 °C to approximately -25 °C. This is perhaps some evidence of the thermal broadening hypothesis (Morton et al. 1956; Williams and Stanfill 2002), since the updraft area increases with height. The fact that the updraft area is largest near -25 °C is interesting because that is typically in the mid-levels of the storm where graupel charge polarity is dependent on SCLW amounts, and we therefore expect this level to be consequential in determining the storm-scale charge structure. The -25 °C level also happens to be close to 8 km MSL in most storms in our sample set, which is where the strongest average and peak updrafts are observed in the anomalous CO storms (Figure 4.5). In contrast, the median updraft fraction distribution in normal AL storms is nearly constant with height. Recall that the vertical mean updraft speed profile is nearly constant with height as well, suggesting updraft strength and updraft width may be related.

Figure 4.6b was compiled in an identical manner as Figure 4.6a, but shows the fraction of grid points with W < -3 m s^{-1}. Differences between the storm populations are evident, especially at lower altitudes. Downdraft fractions in normal AL samples are maximized near the base of the storm. In contrast, the largest downdraft areas in anomalous CO samples are found between -10 °C and -30 °C, overlapping with the largest updraft areas. The mechanisms for this mid-level downdraft maximum are not clear. Downdraft fractions near the base of the storm are a small fraction of the total area in anomalous CO cases. It is important to note that three-dimensional wind fields can only be calculated for regions where radar scatterers are present, so motions outside of the storm radar echo cannot be retrieved. Figure 4.6c shows the remainder of the points at each level, which we classify as neutral because vertical motions are not particularly strong in either direction at these points. Note that most points are not located in significant updraft or downdraft regions as identified by this analysis, consistent with the CFADs in Figure 4.5.

It is essential to place the LMA sources within the content of the three-dimensional wind fields to understand how the dynamics may influence the inferred charge structures of these storms. Figure 4.7a shows the fraction of the storm total LMA sources in vertical motion bins for each radar grid height level averaged over all of the normal AL storms. Most of the LMA sources (rough proxy for positive charge) are located around 10 km MSL in regions of weak vertical motion. Figure 4.7b was constructed in a similar manner, but the sources are collocated with the horizontal gradient of vertical motion. Large values of this gradient indicate LMA sources may be located near an updraft, downdraft or an updraft/downdraft interface (Dye et al. 1986). Figure 4.7b shows that the LMA sources (and positive charge) are not necessarily located near a strong updraft. When paired with the results in Figure 4.7a, this suggests that the positive charge is located at the top of the updraft, as was the case in the archetypal case shown in Figure 4.1. In contrast, Figure 4.7c shows that most of the LMA sources in the CO anomalous storms are located between 5-8 km MSL, but are located in regions of relatively weak vertical motion. However, many of the LMA sources occur in regions of moderate-to-strong horizontal gradients of vertical motion, indicating that the positive charge in the anomalous CO storms is likely located near, but not in, the strongest updrafts. This phenomenon was observed in the archetypal case shown in Figure 4.2, but is confirmed on a larger scale here.

4.3.3.4. *Microphysics*

After elucidating some of the dynamical differences between the normal AL and anomalous CO storm populations, we will now contrast their microphysical characteristics. Figure 4.8 shows some relevant storm total metrics. Figure 4.8a shows the distributions of mixed-phase graupel volume (derived from the number of radar grid points between -5 °C and -40 °C with graupel or hail as the inferred dominant hydrometeor type) for both populations are very similar. Figure 4.8b shows that the mixed-phase 30 dBZ volumes (from grid points between -5 °C and -40 °C with reflectivity greater than 30 dBZ) are slightly larger for the normal AL cases compared to the anomalous CO cases. Figure 4.8c shows the distributions of maximum graupel height in both populations is quite similar, with median values differing by less than 1 km. Evidently the normal AL cases are just as tall and deep as the anomalous CO cases, at least for storms included in this dataset. Finally, the average mixed-phase graupel ice mass is much larger in anomalous CO cases. Perhaps this is a sign of more robust mixed-phase microphysics in anomalous CO cases.

To further understand the differences in vertical structure between the two storm populations, Figure 4.9 shows the vertical profile of radar reflectivity (VPRR; Zipser and Lutz 1994) for each population. Figures 4.9a-b shows the mean VPRR at each radar grid height for each normal AL storm (Figure 4.9a) and anomalous CO storm (Figure 4.9b), as well as an averaged composite for each population. Throughout a large depth ($\sim$ 2-12 km MSL), the average reflectivity is higher in the anomalous CO composite than in the normal AL composite. The magnitude of the difference varies from 0.5 dBZ near 12 km MSL to 5 dB at 7 km MSL, and is 3 dB or greater at most altitudes, which corresponds to a factor of roughly 25 larger ice mass, based on simple scaling arguments (Rutledge et al. 1992). Note that we are assuming perfect attenuation correction when making reflectivity and mass comparisons between radars with different wavelengths. Furthermore, the reflectivity lapse rate between 5 km MSL and 8 km MSL (approximately 0 °C to -20 °C) is drastically different between the two populations. The average reflectivity difference in the anomalous CO composite is only 1.5 dB (0.5 dB km^{-1}), compared to $\sim$6 dB (2 dB km^{-1}) in the normal AL composite. This difference is suggestive of more robust mixed-phase microphysics in the anomalous CO storms, which is consistent with stronger updrafts, precipitation ice growth, and potentially larger supercooled water contents. The maximum vertical profiles of radar reflectivity are shown in Figures 4.9c-d. The composite average of each population is shown in the bold lines. Contrary to the mean profiles, the maximum profiles are more similar to each other. The maximum reflectivities are higher in normal AL cases above 10 km MSL. It is important to note, however, that the 5-8 km lapse rate difference is still intact. In the composite profile for the anomalous CO storms, the difference is 2.5 dB (0.8 dB km^{-1}), whereas the reflectivity difference in the normal AL storms is approximately 10 dB (3.3 dB km^{-1}). This difference is again indicative of stronger mixed-phase hydrometeor growth, particularly where graupel would be located in anomalous CO storms. The composite maximum reflectivity in normal AL storms is higher than anomalous CO cases below 5 km MSL, suggestive of more robust warm-phase precipitation processes in those storms.

It is important to place the LMA sources within the context of the storm microphysics to better understand how microphysical processes may influence the charge structures of these storms (as we did for the dynamical discussion earlier). The archetypal case from the normal AL population suggested that the graupel was not vertically collocated with the positive charge, as measured by the correlation between the vertical distributions of LMA sources and average graupel ice mass. Therefore, it was suggested that this was evidence for ice crystals being the positive charge carriers in the upper-levels. Conversely, in the anomalous CO case, the maximum in LMA sources in the mid-levels was collocated with the maximum in graupel mass. This was used as evidence that graupel is the carrier of mid-level positive charge in the anomalous CO cases. These correlations are further explored in Figure 4.10, which shows the distributions of the Pearson linear correlation coefficient, r, between the vertical distribution of LMA sources and the vertical distribution of average graupel mass. Figure 4.10a shows the correlation coefficient is less than 0 for nearly all of the normal AL cases, indicating that graupel is not collocated with positive charge and therefore likely carries negative charge in the mid-levels, according to the standard normal polarity charge model. However, in the anomalous CO cases, the majority of the samples have correlation coefficients greater than 0.75, indicating that graupel is collocated with positive charge, and may be the positive charge carrier in the mid-levels. If indeed this is the case, then the supercooled water contents in anomalous CO cases would have to be relatively large in accordance with laboratory charging studies.

While the convective samples included in this study are limited to isolated convection, we acknowledge that a 1D correlation may be insufficient to claim whether graupel and positive charge are collocated. To address this potential concern, Figure 4.10b shows the full 3D correlation coefficient between LMA source density and average graupel ice mass. Since LMA source densities can be a noisy quantity, especially in lower flash rate storms, we would expect the 3D correlation coefficient to be closer to 0 regardless of the 1D correlation. Indeed, we do observe this in Figure 4.10b, with nearly all normal AL cases having a vanishingly small correlation coefficient. However, nearly all of the anomalous CO samples have a positive, nonzero correlation coefficient, suggesting graupel and positive charge are collocated. Hence graupel is likely the dominant charge carrier in the anomalous storms. Note that we have not discussed the possible role of hail as a charge carrier. This is because radar-inferred dominant hydrometeor type (not shown) indicates the volume containing hail was a small fraction of overall storm sizes.

4.4. Summary and Discussion

This study used an automated and objective case study framework to compile a number of case studies from Colorado and Alabama in an effort to investigate the microphysical and dynamical processes that influence storm-scale charge structures. In particular, we wanted to understand the processes that lead to anomalous charge structures in thunderstorms. LMA mode temperature was used to infer the storm charge structures, where storms with LMA mode temperatures < -30 °C (> -25 °C) were classified as normal (anomalous) polarity. Detailed analysis of case studies, similar to Figures 4.1-4.2, informed the relevant quantities to investigate in the statistical framework.

The statistical analysis revealed that the inferred positive charge in normal polarity AL storms was located near the top of the updraft in regions of weak horizontal gradients of vertical motion. Furthermore, the low (and often negative) correlation coefficients between the vertical distributions of LMA sources (proxy for positive charge) and graupel ice mass in normal AL storms indicate that positive charge and graupel are not collocated in those storms. This suggests, with reasonable confidence, that smaller ice crystals are the positive charge carriers in the upper levels of normal AL storms, conforming with the normal polarity charging model (e.g. Williams 1985; Mansell et al. 2005). It is more difficult to conclude with certainty that graupel in the mid-levels of those storms carries negative charge based on the LMA mode methodology, but it would be a logical corollary to upper-level positively charged ice crystals. This seems reasonable, given that updrafts are weaker and narrower, coupled with the lower CBHs and deeper WCDs, all of which are expected to lead to less active mixed-phase microphysics and lower SCLW contents, despite having higher instability. Finally, note that the 6-km shear is stronger in the normal AL cases, but it is difficult to know what effect that may have had on the internal processes in the storms in this study.

In contrast, the positive charge in the anomalous CO storms was located at lower altitudes (and warmer temperatures) near regions of relatively strong updrafts. In addition, the high (often near 1) correlation coefficients between the vertical distributions of LMA sources and graupel mass in anomalous CO storms indicate that positive charge and graupel are collocated in those storms. It is impossible, however, to conclude with certainty that the mid-level graupel is indeed the positive charge carrier because it is possible that undetectable ice crystals (by the radars in the study) may be comingling with the graupel. If, however, graupel is the positive charge carrier in the mid-levels of anomalous CO storms, it would be logical to conclude that SCLW amounts are relatively high, based on laboratory studies. Recall that graupel charges positively (negatively) at relatively high (low) SCLW amounts at temperatures near -20 °C (e.g., Takahashi 1978; Saunders and Peck 1998). This seems reasonable when considering that updrafts in anomalous cases are broader and stronger compared to normal polarity storms. The larger graupel ice mass and smaller mixed-phase reflectivity lapse rate suggest that mixed-phase microphysical processes are active and robust, which is consistent with high SCLW contents.

Fuchs et al. (2015) advanced the hypothesis that the majority of warm-season isolated convective storms in CO are anomalous polarity because of the short amount of time that parcels spend in the warm phase of the cloud, which we refer to here as the warm cloud residence time (WCT). With information about storm environment and dynamics in this study, we can test this hypothesis. Figure 4.11a shows the distributions of WCT for both normal AL storms and anomalous CO storms. The WCT was calculated by simply dividing the WCD (distance) by a representative particle speed (difference between updraft and particle fall speed). Each box and whisker plot shows WCT calculations based on different representative W (speed) using different percentiles of vertical motion, below the melting level. Each panel in Figure 4.11 shows the WCT calculations with a terminal fall velocity range of 0-6 m s^{-1}, which spans reasonable drop diameters (e.g. Gunn and Kinzer 1949). If the terminal fall speed is larger than the representative updraft speed, the WCT is set to be 3600 s. This value is somewhat arbitrary and is employed to avoid negative values of WCT in cases where the particle fall speed exceeds the representative updraft speed. In these cases, particles would never reach the mixed-phase region of the cloud.

Regardless of the W percentile choice or raindrop terminal fall speed, the WCTs are much longer for normal AL storms than for anomalous CO storms. At the 50th percentile of W with no terminal fall speed (to represent cloud droplets), the median WCT for the normal AL population is approximately 9 minutes, while it is about 3 minutes for the anomalous CO population. To get a sense for the sensitivity of the WCT to particle fall speed, Figure 4.11b shows the same WCT calculation but includes a constant terminal fall speed of 2 m s^{-1}. The median of the distribution of the WCTs for the same 50th percentile of W has increased from about 9 minutes to 20 minutes, well within the timeframe for warm rain processes to produce precipitation sized drops (e.g. Berry and Reinhardt 1974; Lau and Wu 2003). In contrast, the median value for the anomalous CO population has only increased from about 3 minutes to 5 minutes using the 50th percentile of W. For larger raindrop terminal fall speeds, WCTs increase to the imposed limit of 60 minutes in normal AL cases for multiple W percentiles. WCTs only reach the imposed limit for anomalous CO cases for a terminal fall speed of 6 m s^{-1}.

Given that the normal cases in this study all came from Alabama and all the anomalous cases came from Colorado, we attempted to find and analyze an anomalous AL case, but were unable to find a storm with sufficient flash rates and a sustained region of dominant positive charge at warm temperatures. However, we did identify a case study in the Colorado region that resembled a normal polarity charge structure. Figure 4.12 shows a detailed analysis of the storm, similar to Figures 4.1-4.2. The important features of this storm are: 30-55 dBZ reflectivities in the updraft region, relatively low total flash rates of less than 5 $\min^{-1}$ for its entire lifetime, an upper-level dominant positive charge near -40 °C for the first half hour of its lifetime when it was producing lightning, and maximum updraft speeds were approximately 10-15 m s^{-1} for most of its lifetime. Roughly speaking, this storm resembles a normal polarity AL case. However, an inspection of the vertical distribution of average graupel ice mass reveals that in the first 5-10 minutes of storm development, the LMA source maximum near 10 km MSL is accompanied by relatively high average ice mass at that same level, which leads to a correlation coefficient of 0.95. This suggests that the upper positive charge may be comprised of positively charged graupel. As the storm progresses, the correlation coefficient continuously decreases. During this time, the maximum in vertical LMA source density remains around 10 km MSL, while the majority of the average ice mass is located below 8 km, much like the normal AL cases.

The local environmental conditions (not shown) for this case do not resemble a normal AL storm. CBH values are high, WCD values are low, CAPE values are ∼500 J kg^{-1}, and the NCAPE is ∼0.06 m s^{-2}. Interestingly, the estimated WCTs (using the 75th W percentile) for this case range from approximately 7 minutes during the development phase and increase to approximately 12 minutes after 2240Z, mainly due to the weaker updraft speeds. The WCTs in this case more closely resemble the normal AL storms than the anomalous CO storms.

With all of the different possible permutations of environmental quantities, it seems that WCT is perhaps the best discriminator between the normal AL storms and the anomalous CO storms in this study. Therefore, it is logical to argue that WCT influences SCLW contents (and charge structures) through the amount of liquid water loss via rainout in addition to average supercooled cloud droplet size. We argue these differences in WCTs between the populations control the loss of potential SCLW via warm phase precipitation processes. Simply speaking, the amount of warm-phase precipitation fallout is dependent on the ratio between WCT and the time required for large droplets to form. If this ratio is much less than 1, droplets that form in the warm phase of the updraft will not grow to sufficient sizes to permit fallout. On the other hand, if this ratio is at least 1, droplets may be able to grow large enough to fall out, depleting the parcel of liquid water and thus robbing the mixed phase region of potential supercooled water. We expect the drop formation ratio in anomalous storms to be much less than 1, placing the depletion of SCLW at a minimum. In addition, the broader updrafts observed in Figure 4.6 provides some observational evidence of a relationship between higher CBHs and broader updrafts. These broader updrafts would likely be less prone to entrainment and dilution of available liquid water, especially in the updraft core.

In normal polarity storms, longer warm cloud residence times result in the drop-formation ratio approaching or exceeding 1, resulting in robust warm phase precipitation processes that reduce SCLW in the mixed phase region. The narrower updraft widths, which would make the updrafts more prone to entrainment and dilution, contribute to large liquid water depletion (Williams and Stanfill 2002; Williams et al. 2005). Furthermore, since the cloud droplets spend more time in the warm phase of the cloud, we would expect the cloud droplets to be larger in normal polarity storms, which tends to result in graupel particles charging negatively in the mid-levels (Avila and Pereyra 2000). Note that no physical mechanism for drop size impacting charge polarity was given in the paper. Conversely, shorter WCTs in anomalous storms should lead to smaller supercooled cloud droplet sizes, which would increase the likelihood of positive graupel charging in the mid-levels.

In this study, we have directly investigated the relationships between environmental quantities and storm processes in addition to the relationships between storm processes and charge structures. The hypotheses relating to mixed-phase microphysics (SCLW in particular) put forward based on the analysis in this study would help us better understand fundamental storm processes. However, without in situ observations of SCLW amounts, in addition to information about particle types and concentrations, it is impossible to test the hypotheses put forward in this study.

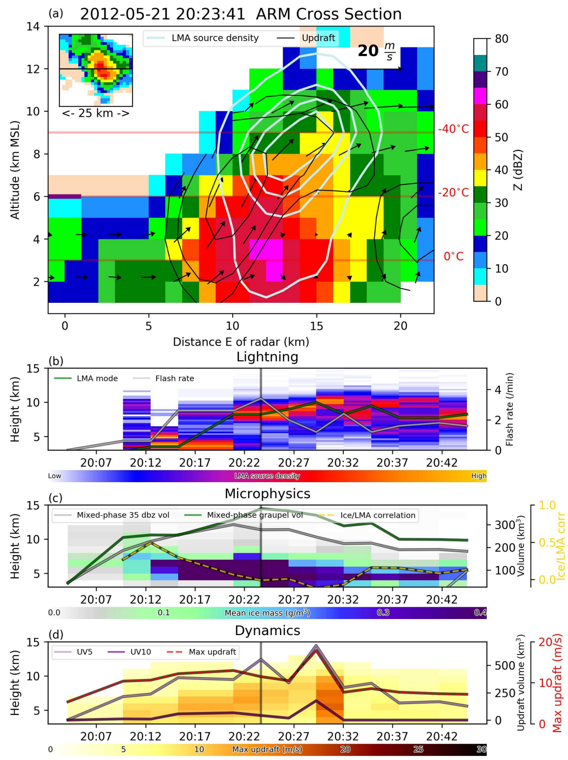

FIG. 4.1. Archetypal normal AL case study from 21 May 2012. (a) X-Z cross section of case at 202341Z, during the mature phase of the storm. Reflectivity values shown in the color fill, following the colorbar to the right of the panel. Vectors show the U and W components of the wind, following the legend in the upper right portion of the panel. The two-dimensional LMA source density (only for sources within 1 km of the cross section) is shown in light contours and is normalized to the maximum absolute value for this case. Updraft speeds are shown in dark contours of 3, 5, 10, 20, and 30 m s^{-1}. A plan view of composite reflectivity is shown in the inset panel.The location of the X-Z cross section is illustrated by the black line. (b) Timeseries of relevant electrical characteristics. Color fill shows vertical distribution of LMA source density, following the colorbar just below the panel, at each radar scan volume time. The LMA mode height (green) follows the left axis, while the cell total flash rate (gray) following the right axis. (c) Timeseries of relevant microphysical characteristics. The color fill shows the average ice mass at each vertical level, following the colorbar just below the panel, at each radar scan volume time. The mixed-phase 35 dBZ volume (gray) and the mixed-phase graupel volume (green) follow the right axis. The correlation coefficient between the vertical distributions of LMA source density and graupel ice mass (yellow dashed) follows the rightmost axis. (d) Timeseries of relevant dynamical characteristics. 95th percentile of updraft speed at each vertical level is shown in the color fill. UV5 and UV10 shown in the light and dark purple colors, respectively, and follow the right axis.

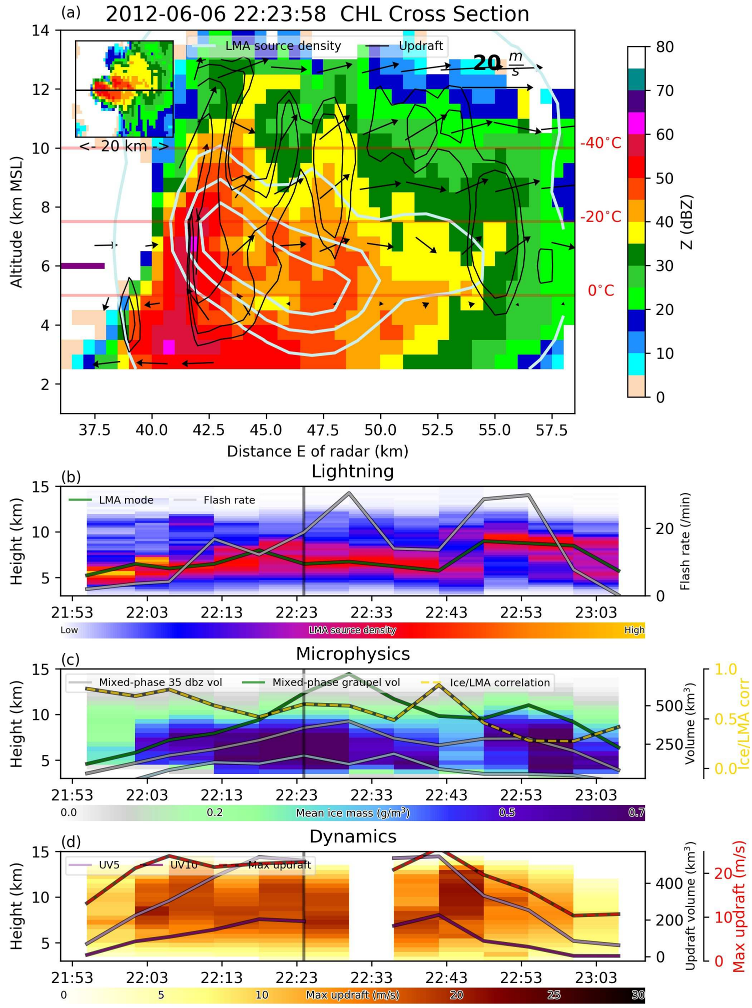

FIG. 4.2. Similar to Figure 1 but for the archetypal anomalous CO case study from 06 June 2012. Note that dual-Doppler data is not available between 2230Z and 2236Z due to a temporary issue with PAWNEE.

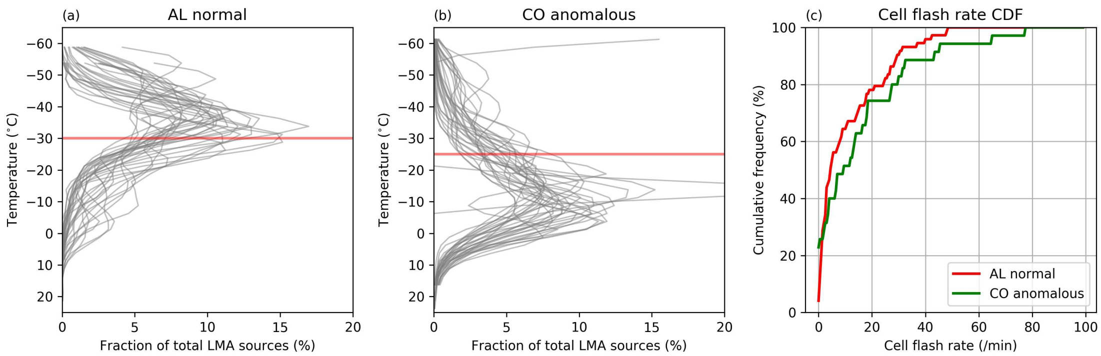

FIG. 4.3. (a) Vertical distribution of normalized LMA source densities for each normal AL storm, with temperature as the vertical coordinate. The red line indicates the -30 °C threshold bifurcating cells into normal or anomalous polarity. (b) Same as (a) but for each anomalous CO storm sample, red line indicates the -25 °C threshold. (c) Cumulative distribution of cell total flash rates for each storm sample.

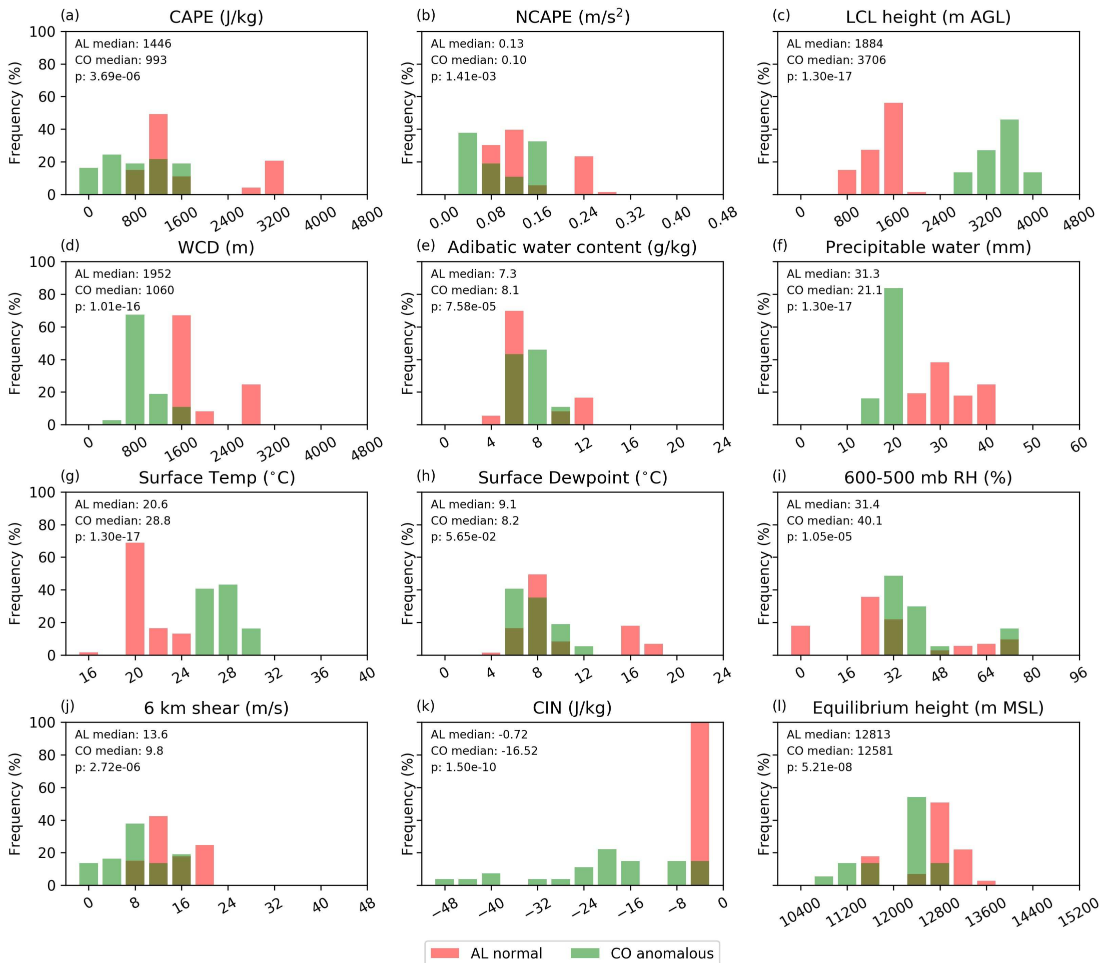

FIG. 4.4. Distributions of (a) CAPE (J kg^{-1}), (b) NCAPE (m s^{-2}), (c) LCL height (m AGL), (d) WCD (m), (e) adiabatic water content (g kg^{-1}), (f) precipitable water (mm), (g) surface temperature (°C), (h) surface dew point (°C), (i) average relative humidity (%) between 600-500 mb, (j) surface to 6 km shear (m s^{-1}), (k) CIN (J kg^{-1}), and (l) equilibrium height (m MSL) from attributed inflow soundings. The bars indicate the quantities associated with storm samples in this study. Median values of each quantity for both regions region and each study are indicated in each panel, in addition to the Spearman ranksum p value, which tests the null hypothesis by calculating the probability that both distributions are subsets of the same distribution.

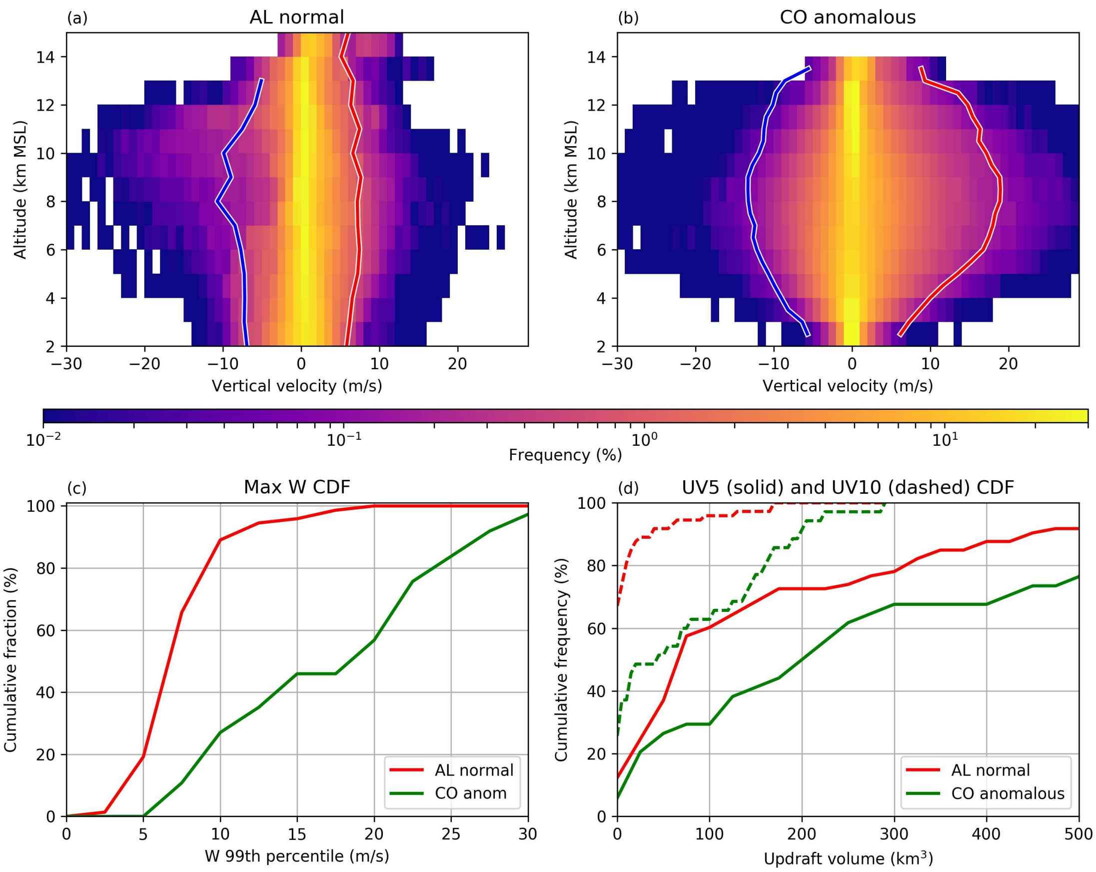

FIG. 4.5. (a) Composite contoured frequency by altitude diagram (CFAD) for all normal AL storm samples, following the log-scale color bar below the panel. The lines indicate the average updraft (red) and downdraft (blue) for all normal AL cases. (b) Same as (a) but for the anomalous CO storm samples. (c) Cumulative distribution of the maximum W (measured by the 99th percentile) for both storm populations. (d) Cumulative distributions of updraft volume greater than 5 m s^{-1} (UV5; solid line) and updraft volume greater than 10 m s^{-1} (UV10; dashed line) for both storm populations.

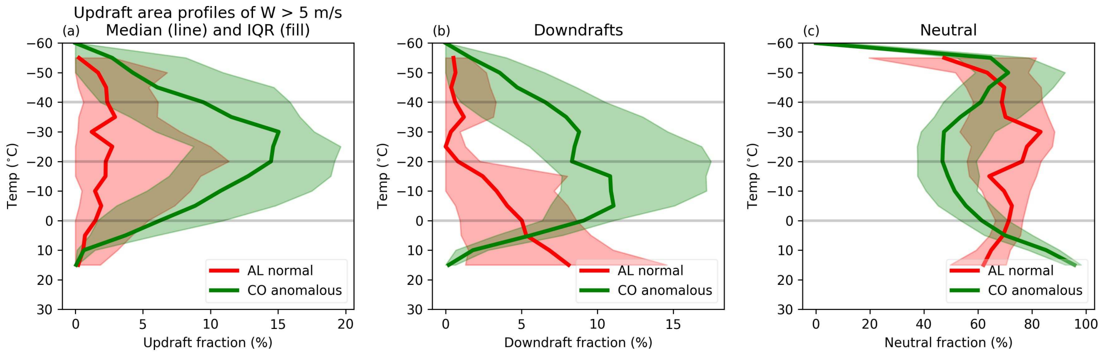

FIG. 4.6. (a) Median (line) and interquartile range (color fill) updraft area (W $\geq$ 5 m s^{-1}) for all storm samples at a particular temperature for each storm population. Converting radar heights to temperatures is done by interpolation using the attributed inflow sounding. (b) Same as (a) but for downdraft area (W < -3 m s^{-1}). (c) Same as (a) but for W between -3 and 5 m s^{-1}.

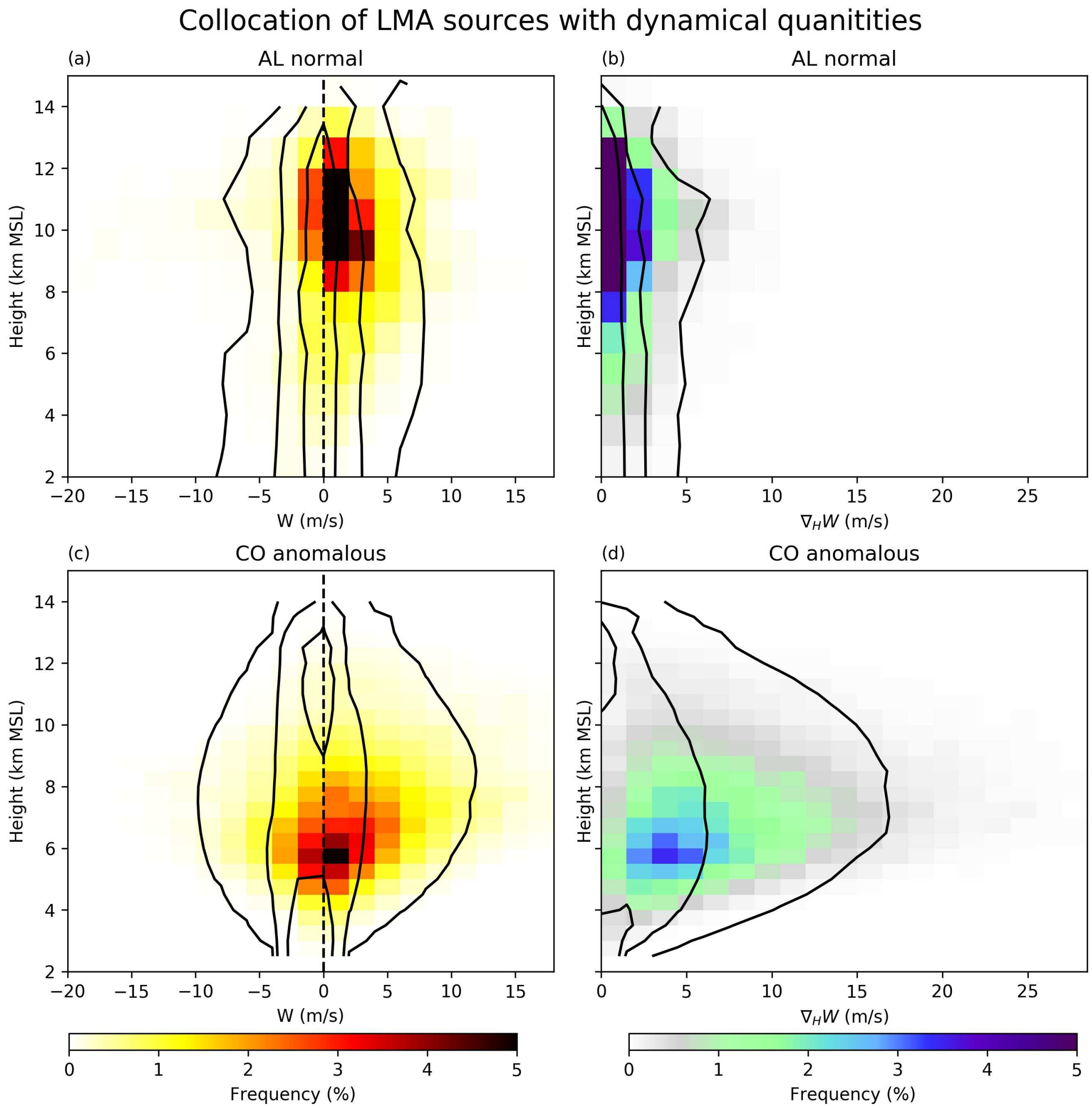

FIG. 4.7. (a) Fraction of total LMA sources collocated with a particular W bin at a particular height, averaged together for all normal AL storm samples. Contours show CFAD values of 1%, 10% and 30% of W. (b) Percentage of total LMA sources collocated with a particular bin of the horizontal gradient of W at a particular height for all normal AL storm samples. (c) Same as (a) but for the anomalous CO storm samples. (d) Same as (b) for the anomalous CO storm samples.

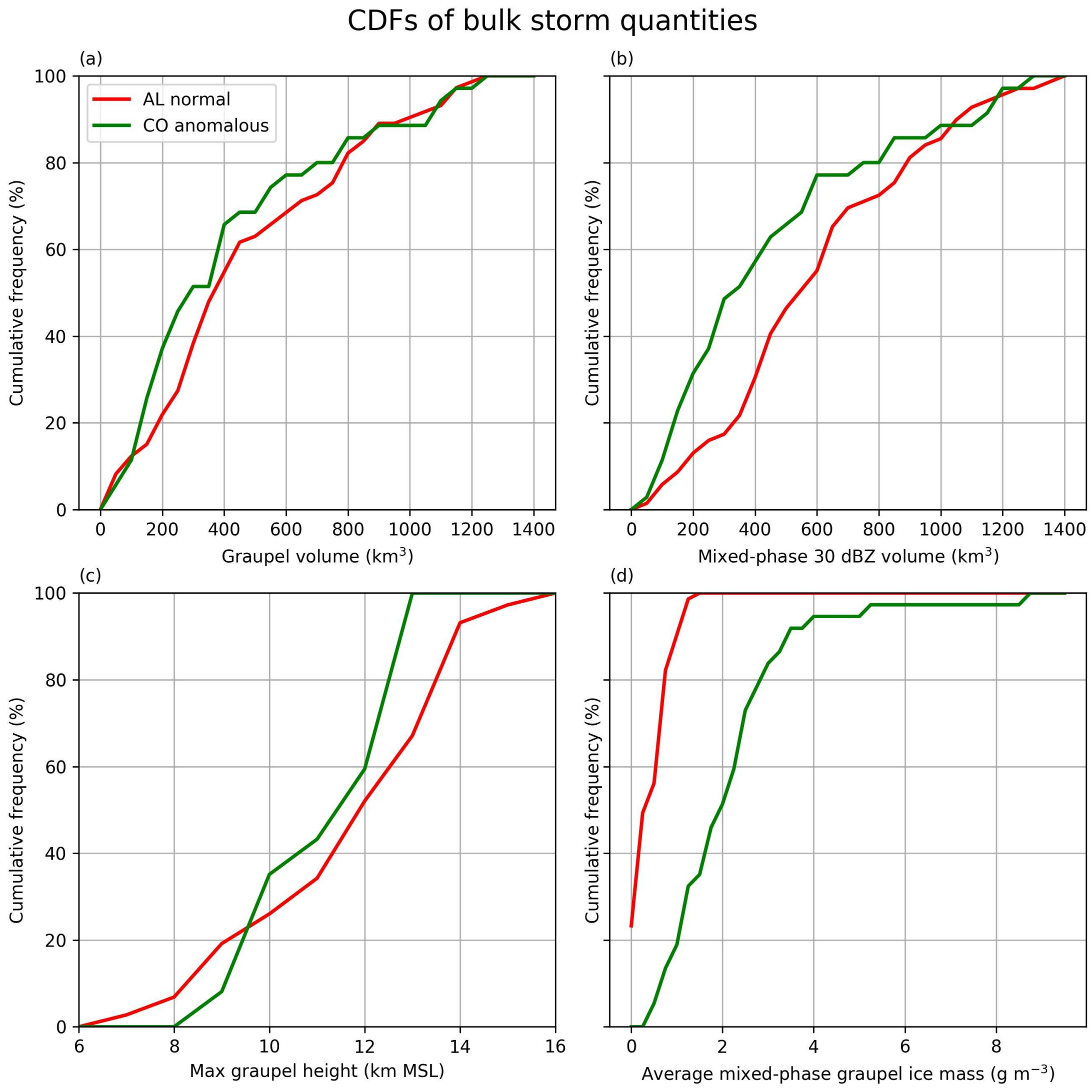

FIG. 4.8. Cumulative distributions of (a) mixed-phase graupel volume, (b) mixed-phase 30 dBZ volume, (c) maximum graupel height (from inferred dominant hydrometeor type), (d) average mixed-phase graupel mass mixing ratio.

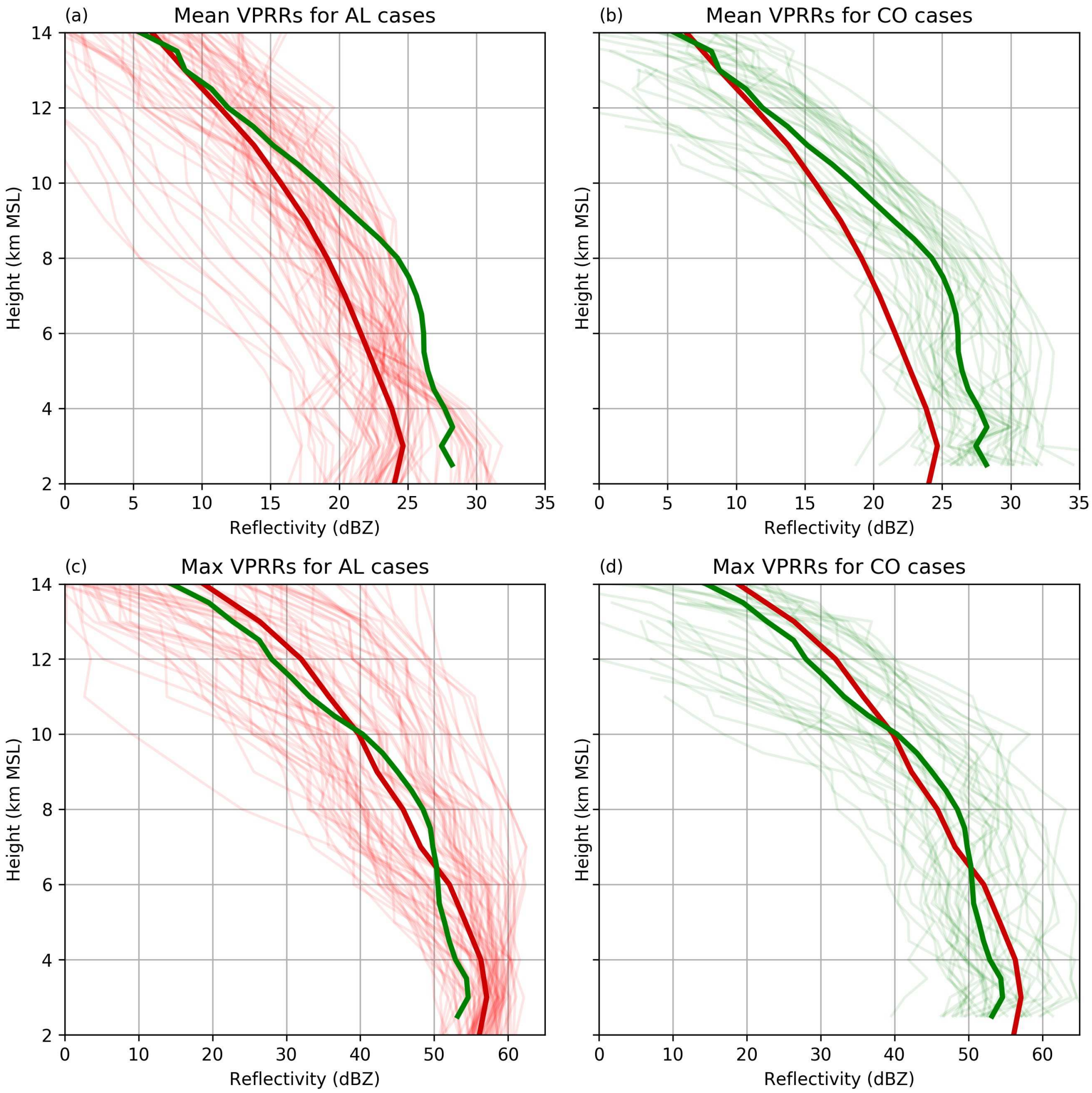

FIG. 4.9. (a) Mean vertical profile of radar reflectivity (VPRR) for each normal AL storm sample (light red lines) and the composite VPRR for the normal AL storm population (dark red line) and the composite VPRR for all anomalous CO storm population (dark green line). (b) Same as (a) but each thin green line is the mean VPRR for each individual anomalous CO storm sample. (c) Same as (a) but for max values instead of mean values in normal AL storm samples. (d) Same as (b) but for max values instead of mean values in anomalous CO storm samples.

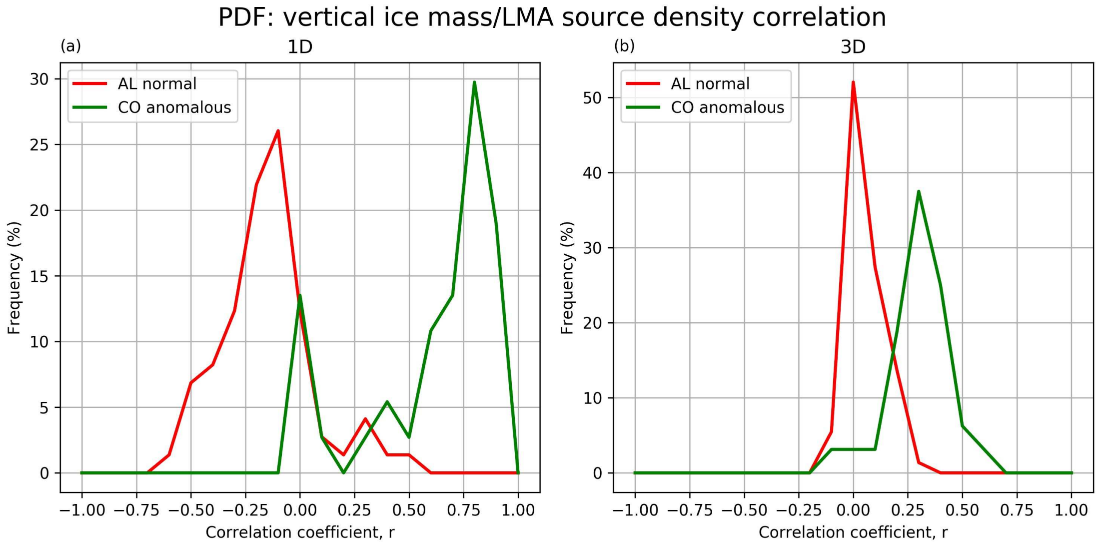

FIG. 4.10. (a) Distributions of correlation coefficient between the vertical distribution of average graupel ice mass and the vertical distribution of LMA source density for every storm in both populations. (b) Same as (a) but correlations are computed between three-dimensional values of graupel ice mass and LMA source density.

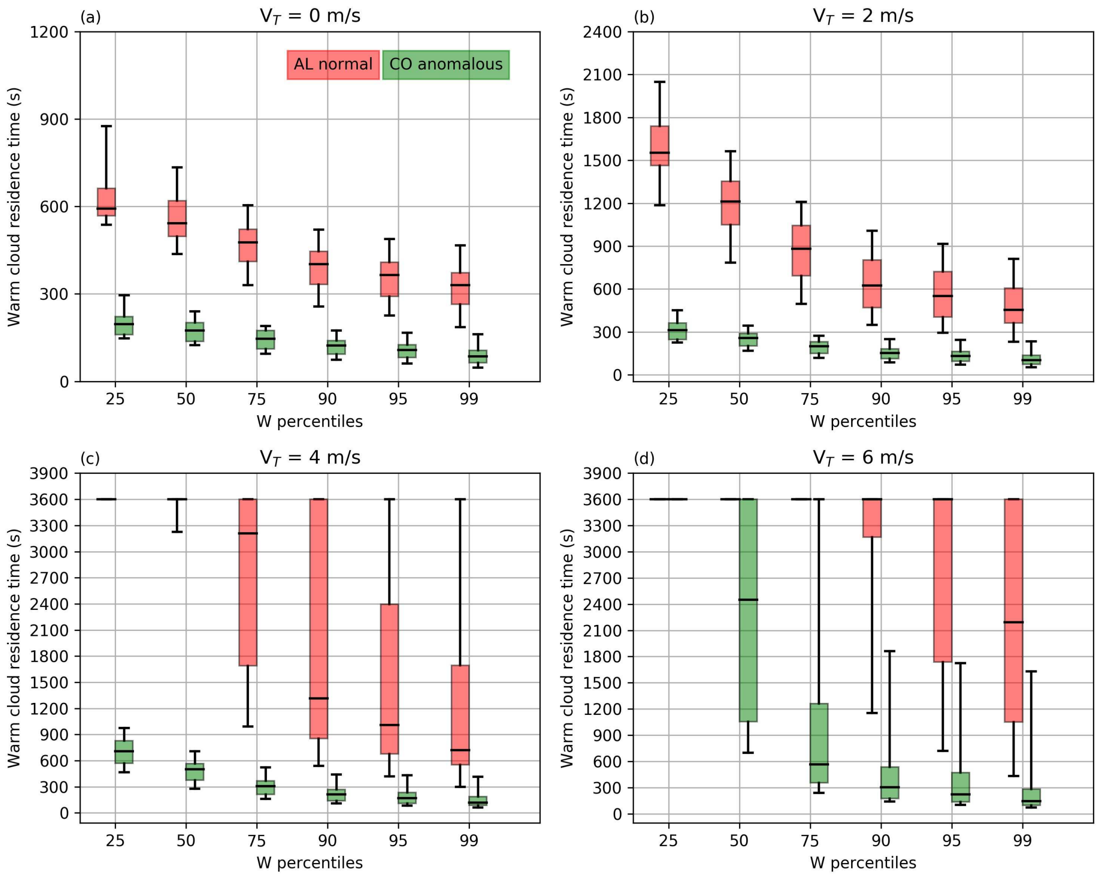

FIG. 4.11. (a) Estimates of warm cloud residence time for each population using different values of W. No particle fall speeds are included in the calculation of warm cloud residence times. (b) Same as (a) but assuming a constant 2 m s^{-1} particle fall speed. (c) Same as (a) assuming a constant 4 m s^{-1} particle fall speed. (d) Same as (a) assuming a constant 6 m s^{-1} particle fall speed. If particle fall speed is greater than representative updraft speed, warm cloud residence time is set to 3600 s.

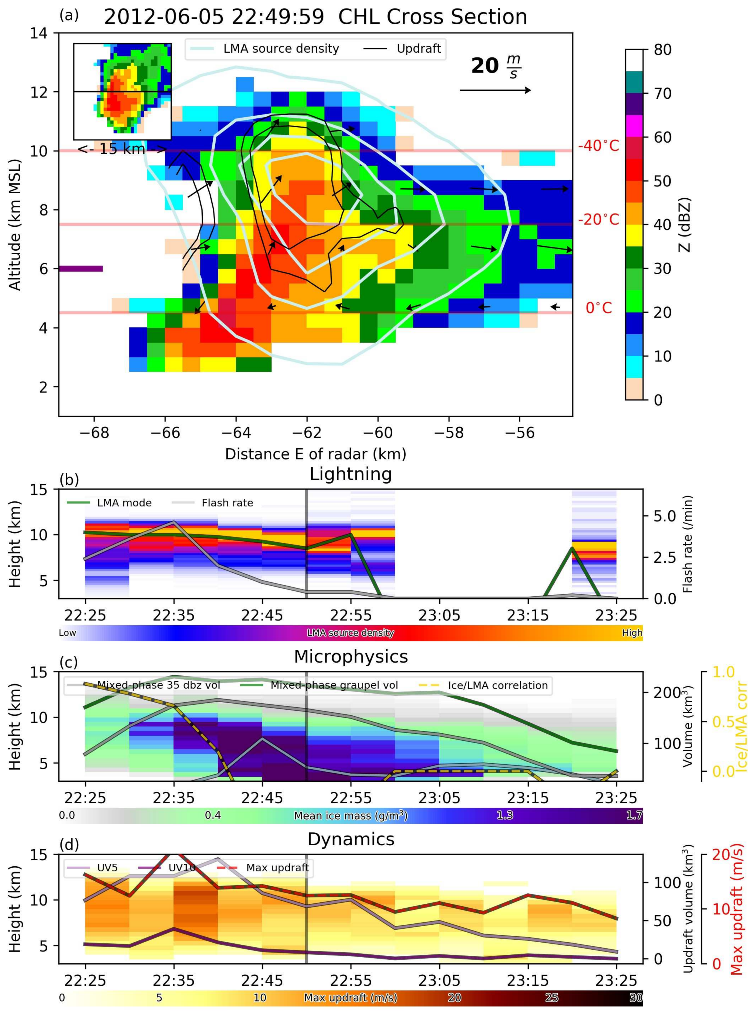

FIG. 4.12. Same as Figure 1, but for a normal polarity case in CO from 05 June 2012. VHF source distributions are not plotted from 2300Z to 2320Z because no lightning occurred within the cell during that time.

CHAPTER 5

Conclusions

5.1. Study summary

The purpose of this study was to improve our understanding of the processes that drive storm-scale charge structure variability, and to understand how storm-scale charge structure impacts lightning locations. Using tools developed in this study, we were able to leverage a wealth of observations of lightning, storm microphysics, and dynamics. These tools provided cell-based information, such as total lightning flash rate and mixed-phase graupel mass, and allowed us to distill a plethora of data for dozens of storms into revealing physical insights. These insights about the processes associated with anomalous storms in Colorado, as well as the distinct lightning behavior associated with those storms, are a significant contribution to the field of lightning studies as well as storm microphysics and dynamics. Furthermore, the tools developed in this study are currently being used in other areas of research, including the relationship between lightning and severe weather and lightning-generated nitrogen oxides. Hopefully, the insights and tools from this study serve as a springboard for new ideas and research in the future.

5.2. Physical mechanisms in anomalous storms

Much of the previous work on anomalous storms was based on the relationship between lightning characteristics (usually +CG%) and the local storm environment, and relied on multiple inferences to arrive at the conclusion that they contained high SCLW contents in the mixed-phase (Carey et al. 2003). In this study, we more directly investigated the processes associated with anomalous charge structures (Figure 1.1). Prominently, we showed for the first time that LMA-inferred dominant positive charge was collocated with mid-level graupel in anomalous polarity storms. We used this information to hypothesize that mid-level graupel was the positive charge carrier in anomalous storms, which leads us to infer the presence of high SCLW contents based on numerous laboratory studies. The links we were able to directly explore are consistent with high SCLW amounts in anomalous storms, namely:

(1) Anomalous storms had stronger and broader updrafts, even though they had somewhat lower instability

(2) Updraft area increased with height in anomalous storms, corroborating the thermal broadening hypothesis

(3) Anomalous storms had significantly shorter warm cloud residence times than normal polarity storms.

The thermal broadening hypothesis (Morton et al. 1956; Williams and Stanfill 2002) has existed for many years and has been invoked to explain charge structure variability through +CG% variability, albeit without much observational evidence. This study provides some evidence for the thermal broadening as the updraft area in anomalous CO storms increased with height. The higher CBH in anomalous CO storms allow the thermal to expand more as it rises to its higher condensation level. Based on Williams and Stanfill (2002), it is then reasonable to expect less entrainment in anomalous CO storms, which would effectively increase the SCLW content in the mixed-phase of anomalous storms. This may relate to the reason that updrafts are stronger and broader in anomalous CO storms, even though they had less instability. Obviously parcel theory is insufficient to explain the stronger updraft metric in CO storms.

Regardless of the underlying mechanism, the stronger updrafts and shallower WCDs in result in much shorter WCTs anomalous CO storms. We hypothesize that WCT is a key discriminator between normal and anomalous polarity storms, as storms with short WCTs have less time for warm-phase processes to grow droplets to sizes large enough to fall as precipitation. This leads to more cloud water in the mixed-phase region and high SCLW contents, which leads to positive graupel charging and anomalous charge structures. Indeed, the results from Chapter 4 showed more robust mixed-phase processes (e.g. higher average graupel mass and mixed-phase reflectivity) in anomalous storms than in normal storms.

5.3. Lightning response to charge structures

Beyond understanding the processes that lead to anomalous charge structures, we also used the automated flash clustering algorithm to produce lightning flash climatologies (Chapter 2). Variability in flash metrics (such as flash size and number of sources) were used to investigate the behavior of the clustering algorithm in addition to the detection ability of each LMA in the study. Spatial maps of annual flash density estimates were constructed within 250 km of LMAs in northern Alabama, Washington DC, and eastern Colorado. Annual flash density estimates ranged from 30-45 flashes km^{-2} yr^{-1} in Alabama, 15-20 flashes km^{-2} yr^{-1} in Washington DC, and up to 75 flashes km^{-2} yr^{-1} in eastern Colorado. That the largest annual flash densities lie in Colorado are consistent with Fuchs et al. (2015) which showed Colorado had the highest per-cell flash rates in the regions considered. On the other hand, it was surprising that Colorado had the highest LMA-based flash density estimates, which is at odds with satellite-based estimates (Cecil et al. 2014).

In a more direct comparison between satellite- and LMA-based annual flash density estimates, corresponding estimates were within 50% in Alabama and DC, but LMA-based estimates were approximately 300% higher than satellite-based estimates in Colorado. Average flash initiation heights in Colorado were approximately 2.5 km lower than in Alabama or DC. It was hypothesized that the anomalous storms, which are common in Colorado, were producing flashes with lower initiation heights. This was a substantial motivating factor in Chapter 3, which attempted to calculate not only the flash initiation location, but also how flashes propagated in three dimensions.

Since LMA networks can three-dimensionally map flashes (assuming they are close enough to the network), we were able to estimate the initiation and channel propagation of each flash produced by over 4000 isolated convective storms. The flash channel calculation method in Chapter 3 has shortcomings related to the inability of LMAs to detect positive breakdown in negative charge. However, a mostly independent method was used to corroborate regional differences in flash channel heights. Indeed, it was shown that the flashes occurred at systematically lower altitudes in Colorado storms, regardless of method. Furthermore, it was directly shown that lower flashes were produced in storms with relatively warm LMA modes (i.e. anomalous polarity storms), which were most common in Colorado. Since lightning initiates in regions of strong electric fields and propagates into potential wells comprised of charged hydrometeors, it makes sense that a significant portion of flashes channels were produced near the inferred dominant positive charge within a storm.

Connecting flash channel heights with storm-scale charge structures means that we may be able to bootstrap these results from the Alabama, DC, and Colorado up to the global scale. There are only a handful of LMAs in operation, each of which has a spatial coverage of approximately 250 km. By understanding the relationship between storm-scale charge structures and flash altitudes, we surmise that low altitude flashes may be missed by satellites in other regions of the globe where anomalous storms occur. We have hypothesized that photons produced by low-altitude flashes are more difficult to detect above the cloud with an optical detector because of the longer optical path for those photons. Given that some storm intensity metrics are based on lightning flash rates (Nesbitt et al. 2000; Lang and Rutledge 2002), incorrect satellite estimates may result in an incomplete understanding of the global distribution of storm intensity.

Lightning is the largest source of NO_x in the upper troposphere and produces ozone, which has implications for human health and climate (Fishman et al. 1979; Ebi and McGregor 2008; Ainsworth et al. 2012). Further, the ozone production efficiency of NO_x is dependent on altitude, with some estimates claiming that ozone production efficiency of NO_x in the upper troposphere up to 6 times higher than NO_x produced in the boundary layer Finney et al. (2016). This implies the ozone production efficiency of NO_x produced by the lower altitude flashes may be relatively low, compared to high altitude flashes. It logically follows that anomalous storms (with low altitude flashes) may have lower overall ozone production efficiency than normal polarity storms (with high altitude flashes). Additionally, we would expect advection of LNO_x to depend on lightning location as well as storm dynamics (Ott et al. 2010). Perhaps these factors will guide future research in lightning studies and atmospheric chemistry.

5.4. Future Research

The analysis conducted and insights gained in this study have advanced our understanding of the processes that influence storm-scale charge structures and provided some clues about why anomalous storms tend to occur in certain regions. Furthermore, this study has revealed some systematic behaviors in lightning altitude in response to storm-scale charge structure variability. The results in this study raise some issues that should be investigated in the future:

(1) Further testing of hypotheses about the processes associated with anomalous storms

(2) Reasons for flash count differences in Colorado, and perhaps all anomalous storms

(3) Possible existence of anomalous storms in other regions than eastern Colorado.

There are a couple of avenues that could be followed to address issue (1): in situ observations of SCLW contents in addition to particle types and concentrations, perhaps with a storm penetrating aircraft. This would be the only way to truly assess the veracity of the claims made in this study. Unfortunately, in situ observations would be very difficult given the strong vertical motions and mixed-phase hydrometeors. An alternative to in situ observations would be to use a model with a charge-separation parameterization, such as the model described in Mansell et al. (2005). Such a model would allow the tracking of SCLW in the mixed-phase region, indeed in all regions, of a storm. One could assess sensitivities to charge structures, SCLW amounts, and WCTs to changing environmental parameters.

In addition to the findings in Chapter 2 pertaining to the differences between LMA- and satellite-based flash density estimates, Thomas et al. (2001) found in a single Oklahoma storm that low-altitudes went undetected by satellites, but were detected by the LMA. They hypothesized (as we have) that the low altitude flashes are more difficult to detect by downward-looking satellites. However, it is possible that other factors may contribute to the mismatch in annual flash density estimates. Flashes that are larger in spatial extent may give off more photons, and therefore would be more likely to be detected. It is also possible that CG flashes and IC flashes have different brightness values, which would impact satellite detection as well. These are factors that can now be investigated with a large dataset of LMA-detected flashes (Chapter 2).

In order to address issue (3), two possibilities are apparent: expand the analysis framework presented in Chapter 4 to other regions of the globe with a similar suite of observations. The problem with this approach is that polarimetric radars would be needed for microphysical information and multiple Doppler radars would be needed for dynamical information. As stated in Chapter 4, this suite of observations is only available from coordinated field campaigns. Furthermore, any field campaign would have to take place in a region where anomalous storms occur. In order to gain an intuition about other locations where anomalous storms occur, we would likely need to look for other regions that have similar environmental characteristics to eastern Colorado, such as high CBH and shallow WCD with sufficient instability. One such possibility is the Remote sensing of Electrification, Lightning, And Mesoscale/Microscale Processes with Adaptive Ground Observations (RELAMPAGO; projectorelampago.org) project planned for the Fall of 2018 in the Mendoza region of Argentina in the foothills of the Andes Mountains. Indeed, this region is in the lee of a tall mountain range and experiences high flash rates and severe weather (Rasmussen et al. 2014), similar to the eastern Plains of Colorado. Therefore, it may be reasonable to suspect that anomalous storms may occur in this region. Beyond RELAMPAGO, it would be useful to estimate the likelihood of anomalous storms, perhaps through an environmental-based proxy that has physical connections to WCT, since we expect that to be the key discriminator between normal and anomalous polarity storms. Expanding the storm database to other regions of the world would allow for more comparisons of regions with different combinations of environmental parameters. Furthermore, comparing anomalous storms in Colorado with anomalous storms for other regions of the globe would undoubtedly lead to more robust insights.

REFERENCES

Ainsworth, E. A., C. R. Yendrek, S. Sitch, W. J. Collins, and L. D. Emberson, 2012: The effects of tropospheric ozone on net primary productivity and implications for climate change. *Annual review of plant biology*, **63**, 637–661.

Atlas, D. and C. W. Ulbrich, 2000: An observationally based conceptual model of warm oceanic convective rain in the tropics. *Journal of Applied Meteorology*, **39 (12)**, 2165–2181.

Avila, E. E. and R. G. Pereyra, 2000: Charge transfer during crystal-graupel collisions for two different cloud droplet size distributions. *Geophys. Res. Lett*, **27 (23)**, 3837–3840.

Baker, M. B. and J. G. Dash, 1989: Charge transfer in thunderstorms and the surface melting of ice. *J. Cry. Grow.*, **97**, 770–776.

Barth, M. C., et al., 2015: The deep convective clouds and chemistry (DC3) field campaign. *Bulletin of the American Meteorological Society*, **96 (8)**, 1281–1309.

Basarab, B., S. Rutledge, and B. Fuchs, 2015: An improved lightning flash rate parameterization developed from Colorado DC3 thunderstorm data for use in cloud-resolving chemical transport models. *Journal of Geophysical Research: Atmospheres*, **120 (18)**, 9481–9499.

Battan, L. J. and J. B. Theiss, 1970: Measurement of vertical velocities in convective clouds by means of pulsed-Doppler radar. *Journal of the Atmospheric Sciences*, **27 (2)**, 293–298.

Behnke, S. A., R. J. Thomas, P. R. Krehbiel, and W. Rison, 2005: Initial leader velocities during intracloud lightning: Possible evidence for a runaway breakdown effect. *Journal of Geophysical Research: Atmospheres*, **110 (D10)**.

Benjamin, S., D. Devenyi, and S. S. Weygandt, 2004: An hourly assimilation forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.

Benjamin, S. G., et al., 2006: From the 13-km RUC to the Rapid Refresh.

Berry, E. X. and R. L. Reinhardt, 1974: An analysis of cloud drop growth by collection: Part I. Double distributions. *Journal of the Atmospheric Sciences*, **31 (7)**, 1814–1824.

Blanchard, D. O., 1998: Assessing the vertical distribution of convective available potential energy. *Weather and Forecasting*, **13 (3)**, 870–877.

Boccippio, D., K. L. Cummins, H. J. Christian, and S. J. Goodman, 2001: Combined satellite and surface-based estimation of the intracloud-cloud-to-ground lightning ratio over the continental United States. *Mon. Wea. Rev.*, **129**, 108–122.

Bradbury, T., 2000: *Meteorology and Flight: Pilot's Guide to Weather*. AC Black London.

Bringi, V. and V. Chandrasekar, 2001: *Polarimetric Doppler weather radar: Principles and applications*. Cambridge University Press.

Bringi, V., R. Hoferer, D. Brunkow, R. Schwerdtfeger, V. Chandrasekar, S. Rutledge, J. George, and P. Kennedy, 2011: Design and performance characteristics of the new 8.5-m dual-offset Gregorian antenna for the CSU–CHILL radar. *Journal of Atmospheric and Oceanic Technology*, **28 (7)**, 907–920.

Bringi, V., K. Knupp, A. Detwiler, L. Liu, I. Caylor, and R. Black, 1997: Evolution of a Florida thunderstorm during the Convection and Precipitation/Electrification Experiment: The case of 9 August 1991. *Monthly weather review*, **125 (9)**, 2131–2160.

Brook, M., M. Nakano, P. Krehbiel, and T. Takeuti, 1982: The electrical structure of the Hokuriku winter thunderstorms. *J. Geophys. Res.*, **87**, 1207–1215.

Bruning, E., S. A. Weiss, and K. M. Calhoun, 2014: Continuous variability in thunderstorm primary electrification and an evaluation of inverted-polarity terminology. *Atmos. Res.*, **135**, 274–284.

Bruning, E. C., 2013: Streamed clustering of lightning mapping data in Python using sklearn. *Scientific Computing With Python*, Vol. 2.

Bruning, E. C. and D. R. MacGorman, 2013: Theory and observations of controls on lightning flash size spectra. *J. Atmos. Sci.*, **70**, 4012–4029.

Bruning, E. C., W. D. Rust, T. J. Schuur, D. R. MacGorman, P. R. Krehbiel, and W. Rison, 2007: Electrical and polarimetric radar observations of a multicell storm in TELEX. *Monthly weather review*, **135 (7)**, 2525–2544.

Budden, K. G., 1988: *The propagation of radio waves: The theory of radio waves of low power in the ionosphere and magnetosphere*. Cambridge University Press.

Calhoun, K. M., D. R. MacGorman, C. L. Ziegler, and M. I. Biggerstaff, 2013: Evolution of lightning activity and storm charge relative to dual-Doppler analysis of a high-precipitation supercell storm. *Monthly Weather Review*, **141 (7)**, 2199–2223.

Carey, L. D. and K. M. Buffalo, 2007: Environmental control of cloud-to-ground lightning polarity in severe storms. *Mon. Wea. Rev.*, **135**, 1327–1353.

Carey, L. D., W. Koshak, H. Peterson, and R. M. Mecikalski, 2016: The kinematic and microphysical control of lightning rate, extent, and NO_X production. *Journal of Geophysical Research: Atmospheres*, **121 (13)**, 7975–7989.

Carey, L. D. and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *J. Geophys. Res.*, **103**, 13 979–14 000.

Carey, L. D. and S. A. Rutledge, 2000: The relationship between precipitation and lightning in tropical island convection: A C-band polarimetric radar study. *Monthly weather review*, **128 (8)**, 2687–2710.

Carey, L. D. and S. A. Rutledge, 2003: Characteristics of cloud to ground lightning in severe and nonsevere storms over the central United States from 1989–1998. *J. Geophys. Res.*, **108**, D15.

Carey, L. D., S. A. Rutledge, and W. A. Petersen, 2003: The relationship between severe storm reports and cloud-to-ground lightning polarity in the contiguous United States from 1989 to 1998. *Mon. Wea. Rev.*, **131**, 1211–1228.

Cecil, D. J., D. E. Buechler, and R. J. Blakeslee, 2014: Gridded lightning climatology from TRMM-LIS and OTD: Dataset description. *Atmospheric Research*, **135**, 404–414.

Chmielewski, V. C. and E. C. Bruning, 2016: Lightning Mapping Array flash detection performance with variable receiver thresholds. *Journal of Geophysical Research: Atmospheres*, **121 (14)**, 8600–8614.

Christian, H. J., R. J. Blakeslee, S. J. Goodman, D. Boccippio, and coauthors, 2003: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector. *J. Geophys. Res.*, **108**, 4–15.

Christian, H. J., R. J. Blakeslee, S. J. Goodman, D. A. Mach, and coauthors, 1999: The Lightning Imaging Sensor. *NASA conf. pub.*, **1**, 746–749.

Cifelli, R., V. Chandrasekar, S. Lim, P. Kennedy, Y. Wang, and S. Rutledge, 2011: A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *Journal of Atmospheric and Oceanic Technology*, **28 (3)**, 352–364.

Cifelli, R., W. A. Petersen, L. D. Carey, S. A. Rutledge, and M. A. da Silva Dias, 2002: Radar observations of the kinematic, microphysical, and precipitation characteristics of two MCSs in TRMM LBA. *Journal of Geophysical Research: Atmospheres*, **107 (D20)**.

Coleman, L., T. Marshall, M. Stolzenburg, T. Hamlin, P. Krehbiel, W. Rison, and R. Thomas, 2003: Effects of charge and electrostatic potential on lightning propagation. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **108 (D9)**.

Crook, N. A., T. L. Clark, and M. W. Moncrieff, 1990: The Denver Cyclone. Part I: Generation in low Froude number flow. *J. Atmos. Sci.*, **47**, 2725–2742.

Cummins, K. and M. Murphy, 2009: An overview of lightning locating systems: History, techniques, and data uses, with an in-depth look at the US NLDN. *Electromagnetic Compatibility*, **51**, 499–518.

DeCaria, A. J., K. E. Pickering, G. L. Stenchikov, and L. E. Ott, 2005: Lightning-generated NOx and its impact on tropospheric ozone production: A three-dimensional modeling study of a Stratosphere-Troposphere Experiment: Radiation, Aerosols and Ozone (STERAO-A) thunderstorm. *Journal of Geophysical Research: Atmospheres*, **110 (D14)**.

DeCaria, A. J., K. E. Pickering, G. L. Stenchikov, J. R. Scala, J. L. Stith, J. E. Dye, B. A. Ridley, and P. Laroche, 2000: A cloud-scale model study of lightning-generated NOx in an individual thunderstorm during STERAO-A. *Journal of geophysical research*, **105**, 11.

Deierling, W. and W. A. Petersen, 2008: Total lightning activity as an indicator of updraft characteristics. *J. Geophys. Res.*, **106**, 20 395–20 402.

Dolan, B., S. A. Rutledge, S. Lim, V. Chandrasekar, and M. Thurai, 2013: A robust C-band hydrometeor identification algorithm and application to a long-term polarimetric radar dataset. *Journal of Applied Meteorology and Climatology*, **52 (9)**, 2162–2186.

Doviak, R. J. and D. S. Zrnić, 1993: *Doppler radar and weather observations*. Courier Dover Publications.

Dwyer, J. R., 2005: The initiation of lightning by runaway air breakdown. *Geophysical research letters*, **32 (20)**.

Dye, J. E., et al., 1986: Early electrification and precipitation development in a small, isolated Montana cumulonimbus. *J. Geophys. Res.*, **91**, 1231–1247.

Ebi, K. L. and G. McGregor, 2008: Climate change, tropospheric ozone and particulate matter, and health impacts. *Environmental health perspectives*, **116 (11)**, 1449.

Efron, B. and B. Efron, 1982: *The jackknife, the bootstrap and other resampling plans*, Vol. 38. SIAM.

Ester, Martin, Kriegel, and coauthors, 1996: A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*, Vol. 96, 226–231.

Finney, D. L., R. M. Doherty, O. Wild, P. J. Young, and A. Butler, 2016: Response of lightning NOx emissions and ozone production to climate change: Insights from the Atmospheric Chemistry and Climate Model Intercomparison Project. *Geophysical Research Letters*, **43 (10)**, 5492–5500, doi: 10.1002/2016GL068825, URL `http://dx.doi.org/10.1002/2016GL068825`, 2016GL068825.

Fishman, J., V. Ramanathan, P. Crutzen, and S. Liu, 1979: Tropospheric ozone and climate. *and S. C. Liu(NOAA, Boulder, Colo.). Nature*, **282**, 818–820.

Fuchs, B. R., E. C. Bruning, S. A. Rutledge, L. D. Carey, P. R. Krehbiel, and W. Rison, 2016: Climatological analyses of LMA data with an open-source lightning flash-clustering algorithm. *Journal of Geophysical Research: Atmospheres*, **121 (14)**, 8625–8648.

Fuchs, B. R., et al., 2015: Environmental controls on storm intensity and charge structure in multiple regions of the continental United States. *Journal of Geophysical Research: Atmospheres*, **120 (13)**, 6575–6596.

Goldenbaum, G. and R. Dickerson, 1993: Nitric oxide production by lightning discharges. *Journal of Geophysical Research: Atmospheres*, **98 (D10)**, 18 333–18 338.

Goodman, S. J., et al., 2013: The GOES-R geostationary lightning mapper (GLM). *Atmospheric research*, **125**, 34–49.

Gunn, R. and G. D. Kinzer, 1949: The terminal velocity of fall for water droplets in stagnant air. *Journal of Meteorology*, **6 (4)**, 243–248.

Huntrieser, H., H. Schlager, C. Feigl, and H. Höller, 1998: Transport and production of NO x in electrified thunderstorms: Survey of previous studies and new observations at midlatitudes. *Journal of Geophysical Research: Atmospheres*, **103 (D21)**, 28 247–28 264.

Jayaratne, E. R., C. P. R. Saunders, and J. Hallett, 1983: Laboratory studies of the charging of soft-hail during ice crystal interactions. *Q. Jour. of the Roy. Met. Soc.*, **109**, 609–630.

Jirak, I. L. and W. R. Cotton, 2006: Effect of air pollution on precipitation along the Front Range of the Rocky Mountains. *Journal of Applied Meteorology and Climatology*, **45 (1)**, 236–245.

Koshak, W., et al., 2004: North Alabama Lightning Mapping Array (LMA): VHF source retrieval algorithm and error analyses. *Journal of Atmospheric and Oceanic Technology*, **21 (4)**, 543–558.

Krehbiel, P., 1986: *The Electrical Structure of Thunderstorms*. National Academies Press.

Krehbiel, P., R. J. Thomas, W. Rison, T. Hamlin, M. Davis, and J. Harlin, 2000: Lightning mapping observations in central Oklahoma. *Eos*, **81 (3)**, 21–25.

Krehbiel, P. R., M. Brook, and R. A. McCrory, 1979: An analysis of the charge structure of lightning discharges to ground. *Journal of Geophysical Research: Oceans*, **84 (C5)**, 2432–2456.

Lang, T. J. and S. A. Rutledge, 2002: Relationships between convective storm kinematics, precipitation and lightning. *Mon. Wea. Rev.*, **130**, 2492–2506.

Lang, T. J. and S. A. Rutledge, 2011: A framework for the statistical analysis of large radar and lightning datasets: Results from STEPS 2000. *Mon. Wea. Rev.*, **139**, 2536–2551.

Lang, T. J., S. A. Rutledge, B. Dolan, P. Krehbiel, W. Rison, and D. T. Lindsey, 2014: Lightning in wildfire smoke plumes observed in Colorado during summer 2012. *Monthly Weather Review*, **142 (2)**, 489–507.

Lang, T. J., et al., 2004: The Severe Thunderstorm Electrification and Precipitation Study. *Bull. Amer. Meteor. Soc.*, **85**, 1107–1125.

Lang, T. J., et al., 2016: Observations of two sprite-producing storms in Colorado. *Journal of Geophysical Research: Atmospheres*, **121 (16)**, 9675–9695.

Latham, J., 1981: The electrification of thunderstorms. *Quarterly Journal of the Royal Meteorological Society*, **107 (452)**, 277–298.

Lau, K. and H. Wu, 2003: Warm rain processes over tropical oceans and climate implications. *Geophysical Research Letters*, **30 (24)**.

LeMone, M. A. and E. J. Zipser, 1980: Cumulonimbus vertical velocity events in GATE. Part I: Diameter, intensity and mass flux. *J. Atmos. Sci.*, **37**, 2444–2457.

Lin, X., M. Trainer, and S. Liu, 1988: On the nonlinearity of the tropospheric ozone production. *Journal of Geophysical Research: Atmospheres*, **93 (D12)**, 15 879–15 888.

MacGorman, D. R., W. D. Rust, T. J. Schuur, M. J. Biggerstaff, J. M. Straka, and coauthors, 2008: TELEX the thunderstorm electrification and lightning experiment. *Bull. Amer. Met. Soc.*, **89**, 997–1013.

MacGorman, D. R., J. M. Straka, and C. L. Ziegler, 2001: A lightning parameterization for numerical cloud models. *J. Appl. Meteor.*, **40**, 459–478.

Mach, D. M., H. J. Christian, R. J. Blakeslee, D. J. Boccippio, S. J. Goodman, and W. L. Boeck, 2007: Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **112 (D9)**.

Maggio, C., et al., 2005: Lightning-initiation locations as a remote sensing tool of large thunderstorm electric field vectors. *Journal of Atmospheric and Oceanic Technology*, **22 (7)**, 1059–1068.

Mansell, E. R., 2014: Storm-scale ensemble Kalman filter assimilation of total lightning flash-extent data. *Monthly Weather Review*, **142 (10)**, 3683–3695.

Mansell, E. R., D. R. MacGorman, C. L. Ziegler, and J. M. Straka, 2005: Charge structure and lightning sensitivity in a simulated multicell thunderstorm. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **110 (D12)**.

Marshall, T. C., M. P. McCarthy, and W. D. Rust, 1995: Electric field magnitudes and lightning initiation in thunderstorms. *Journal of Geophysical Research: Atmospheres*, **100 (D4)**, 7097–7103.

McCaul, E. W., S. J. Goodman, K. M. LaCasse, and D. J. Cecil, 2009: Forecasting lightning threat using cloud-resolving model simulations. *Wea. and Fore.*, **24**, 709–729.

Mecikalski, R. M., A. L. Bain, and L. D. Carey, 2015: Radar and lightning observations of deep moist convection across northern Alabama during DC3: 21 May 2012. *Monthly Weather Review*, **143 (7)**, 2774–2794.

Medici, G., K. L. Cummins, D. J. Cecil, W. J. Koshak, and S. D. Rudlosky, 2017: The intra-cloud lightning fraction in the contiguous United States. *Monthly Weather Review*, **(2017)**.

Miller, J. L., C. G. Mohr, and A. J. Weinheimer, 1986: The simple rectification to Cartesian space of folded radial velocities from Doppler radar sampling. *Journal of Atmospheric and Oceanic Technology*, **3 (1)**, 162–174.

Mohr, C. and R. Vaughan, 1979: An economical procedure for cartesian interpolation and display of reflectivity factor data in three-dimensional space. *Journal of Applied Meteorology*, **18 (5)**, 661–670.

Mohr, C. G., L. Jay Miller, R. L. Vaughan, and H. W. Frank, 1986: The merger of mesoscale datasets into a common Cartesian format for efficient and systematic analyses. *Journal of Atmospheric and Oceanic Technology*, **3 (1)**, 143–161.

Morton, B. R., G. Taylor, and J. S. Turner, 1956: Turbulent gravitational convection from maintained and instantaneous sources. *Proc. R. Soc. Lond.*, **234**, 1–23.

Mroz, K., A. Battaglia, T. J. Lang, D. J. Cecil, S. Tanelli, and F. Tridon, 2017: Hail-Detection Algorithm for the GPM Core Satellite Sensors. *Journal of Applied Meteorology and Climatology*, **(2017)**.

Nag, A., M. Murphy, A. Pifer, K. Cummins, and J. Cramer, 2013: Upgrade of the US National Lightning Detection Network in 2013. *AGU Fall Meeting Abstracts*, Vol. 1, 0337.

Nag, A., M. J. Murphy, W. Schulz, and K. L. Cummins, 2015: Lightning locating systems: Insights on characteristics and validation techniques. *Earth and Space Science*, **2 (4)**, 65–93.

Nesbitt, S. W., E. J. Zipser, and D. J. Cecil, 2000: A census of precipitation features in the tropics using trmm: Radar, ice scattering, and lightning observations. *Journal of Climate*, **13 (23)**, 4087–4106.

Orville, R. E., 1994: Cloud-to-ground lightning flash characteristics in the contiguous United States: 1989–1991. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **99 (D5)**, 10 833–10 841.

Orville, R. E. and G. R. Huffines, 2001: Cloud-to-ground lightning in the United States: NLDN results in the first decade. *Mon. Wea. Rev.*, **125**, 631–638.

Ott, L. E., et al., 2010: Production of lightning NOx and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations. *Journal of Geophysical Research: Atmospheres*, **115 (D4)**.

Oye, D. and M. Case, 1995: REORDER: A program for gridding radar data. Installation and use manual for the UNIX version. *NCAR/ATD*.

Pawar, S. D. and A. K. Kamra, 2007: End-of-storm oscillation in tropical air mass thunderstorms. *Jour. Geophys. Res.*, **112**.

Pedregosa, F., et al., 2011: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Petersen, W. A., H. J. Christian, and S. A. Rutledge, 2005: TRMM observations of the global relationship between ice water content and lightning. *Geophys. Res. Lett.*, **32 (14)**.

Petersen, W. A., K. R. Knupp, D. J. Cecil, J. R. Mecikalski, C. Darden, and J. Burks, 2007: The University of Alabama Huntsville THOR Center instrumentation: research and operational collaboration. *33rd International Conference on Radar Meteorology*.

Pickering, K. E., Y. Wang, W.-K. Tao, C. Price, and J.-F. Müller, 1998: Vertical distributions of lightning NOx for use in regional and global chemical transport models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **103 (D23)**, 31 203–31 216.

Potvin, C. K., L. J. Wicker, and A. Shapiro, 2012: Assessing errors in variational dual-Doppler wind syntheses of supercell thunderstorms observed by storm-scale mobile radars. *Journal of Atmospheric and Oceanic Technology*, **29 (8)**, 1009–1025.

Price, C., J. Penner, and M. Prather, 1997: NOx from lightning: 1. Global distribution based on lightning physics. *Journal of Geophysical Research: Atmospheres*, **102 (D5)**, 5929–5941.

Pruppacher, H. R., J. D. Klett, and P. K. Wang, 1998: Microphysics of clouds and precipitation. Taylor & Francis.

Rasmussen, K. and R. Houze Jr, 2016: Convective initiation near the Andes in subtropical South America. *Monthly Weather Review*, **144 (6)**, 2351–2374.

Rasmussen, K. L., M. D. Zuluaga, and R. A. Houze, 2014: Severe convection and lightning in subtropical South America. *Geophysical Research Letters*, **41 (20)**, 7359–7366.

Reap, R. M. and D. R. MacGorman, 1989: Cloud-to-ground lightning: Climatological characteristics and relationships to model fields, radar observations and severe local storms. *Mon. Wea. Rev.*, **117**, 518–535.

Reynolds, S. E., M. Brook, and M. F. Gourley, 1957: Thunderstorm charge separation. *J. Meteor.*, **14**, 426–436.

Riemann-Campe, K., K. Fraedrich, and F. Lunkeit, 2009: Global climatology of convective available potential energy (CAPE) and convective inhibition (CIN) in ERA-40 reanalysis. *Atmospheric Research*, **93 (1)**, 534–545.

Rison, W., R. Thomas, P. Krehbiel, T. Hamlin, and J. Harlin, 1999: A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico. *Geophys. Res. Lett.*, **26**, 3573–3576.

Rudlosky, S. D., 2015: Evaluating ENTLN Performance Relative to TRMM/LIS. *Journal of Operational Meteorology*, **3 (2)**.

Rudlosky, S. D. and H. E. Fuelberg, 2010: Pre-and postupgrade distributions of NLDN reported cloud-to-ground lightning characteristics in the contiguous United States. *Monthly Weather Review*, **138 (9)**, 3623–3633.

Rust, W. D., et al., 2005: Inverted-polarity electrical strucutres in thunderstorms in the Severe Thunderstorm Electrification and Precipiation Study (STEPS). *Atmos. Res.*, **76**, 247–271.

Rutledge, S. A., E. R. Williams, and T. D. Keenan, 1992: The Down Under Doppler and Electricity Experiment (DUNDEE): Overview and preliminary results. *Bull. Amer. Met. Soc.*, **73**, 3–16.

Said, R., M. Cohen, and U. Inan, 2013: Highly intense lightning over the oceans: Estimated peak currents from global GLD360 observations. *Journal of Geophysical Research: Atmospheres*, **118 (13)**, 6905–6915.

Saunders, C. P. R., 2008: Charge separation mechanisms in clouds. *Space Sci. Rev.*, **137**, 335–353.

Saunders, C. P. R., H. Bax-Norman, C. Emersic, E. E. Avila, and N. E. Castellano, 2006: Laboratory studies on the effect of cloud conditions on the graupel/crystal charge transfer in thunderstorm electrification. *Quar. J. of the Royal Met. Soc.*, **132**, 2653–2673.

Saunders, C. P. R., W. D. Keith, and R. P. Mitzeva, 1991: The effect of liquid water on thunderstorm charging. *J. Geophys. Res.*, **96**, 11 007–11 017.

Saunders, C. P. R. and S. L. Peck, 1998: Laboratory studies of the influence of the rime accretion rate on charge transfer during graupel/crystal collisions. *J. Geophys. Res.*, **103**, 13 949–13 956.

Schultz, C. J., L. D. Carey, E. V. Schultz, and R. J. Blakeslee, 2015: Insight into the kinematic and microphysical processes that control lightning jumps. *Weather and Forecasting*, **(2015)**.

Schumann, U. and H. Huntrieser, 2007: The global lightning-induced nitrogen oxides source. *Atmospheric Chemistry and Physics*, **7 (14)**, 3823–3907.

Shao, X. and P. Krehbiel, 1996: The spatial and temporal development of intracloud lightning. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **101 (D21)**, 26 641–26 668.

Shumpert, T., M. Honnell, and G. Lott Jr, 1982: Measured spectral amplitude of lightning sferics in the HF, VHF, and UHF bands. *Electromagnetic Compatibility, IEEE Transactions on*, **(3)**, 368–369.

Smith, S. B., J. G. LaDue, and D. R. MacGorman, 2000: The relationship between cloud-to-ground lightning polarity and surface equivalent potential temperature during three tornadic outbreaks. *Mon. Wea. Rev.*, **128**, 3320–3328.

Stolz, D., S. A. Rutledge, and J. R. Pierce, 2015: Simultaneous influences of thermodynamics and aerosols on deep convection and lightning in the tropics. *J. Geophys. Res.*

Stolzenburg, M., 1994: Observations of high ground flash densities of positive lightning in summertime thunderstorms. *Mon. Wea. Rev.*, **122**, 1740–1750.

Takahashi, H., K. Suzuki, and G. Stephens, 2017: Land–ocean differences in the warm-rain formation process in satellite and ground-based observations and model simulations. *Quarterly Journal of the Royal Meteorological Society*, **143 (705)**, 1804–1815.

Takahashi, T., 1978: Riming electrification as a charge generation mechanism in thunderstorms. *Atmos. Sci.*, **35**, 1536–1548.

Tessendorf, S. A., S. A. Rutledge, and K. C. Wiens, 2007: Radar and lightning observations of normal and inverted multicellular storms from STEPS. *Mon. Wea. Rev.*, **135**, 3682–3706.

Thomas, R., P. Krehbiel, W. Rison, J. Harlin, T. Hamlin, and N. Campbell, 2003: The LMA flash algorithm. *Proc. 12th Int. Conf. on Atmospheric Electricity*, 655–656.

Thomas, R. J., P. R. Krehbiel, W. Rison, T. Hamlin, D. J. Boccippio, S. J. Goodman, and H. J. Christian, 2000: Comparison of ground-based 3-dimensional lightning mapping observations with satellite-based LIS observations in Oklahoma. *Geophysical research letters*, **27 (12)**, 1703–1706.

Thomas, R. J., P. R. Krehbiel, W. Rison, T. Hamlin, J. Harlin, and D. Shown, 2001: Observations of VHF source powers radiated by lightning. *Geophysical research letters*, **28 (1)**, 143–146.

Thomas, R. J., P. R. Krehbiel, W. Rison, S. J. Hunyady, W. P. Winn, T. Hamlin, and J. Harlin, 2004: Accuracy of the lightning mapping array. *Jour. Geophys. Res.*, **109**, 14 207–14 216.

Velde, O. A. and J. Montanyà, 2013: Asymmetries in bidirectional leader development of lightning flashes. *Journal of Geophysical Research: Atmospheres*, **118 (24)**.

Vogt, B. J. and S. J. Hodanish, 2014: A high-resolution lightning map of the state of Colorado. *Monthly Weather Review*, **142 (7)**, 2353–2360.

Wang, Y. and V. Chandrasekar, 2009: Algorithm for estimation of the specific differential phase. *Journal of Atmospheric and Oceanic Technology*, **26 (12)**, 2565–2578.

Weisman, M. L. and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520.

Wiens, K. C., S. A. Rutledge, and S. A. Tessendorf, 2005: The 29 June 2000 supercell observed during STEPS. Part II: Lightning and charge structure. *J. Atmos. Res.*, **62**, 4151–4177.

Williams, E. R., 1985: Large-scale separation in thunderclouds. *J. Geophys. Res.*, **90**, 6013–6025.

Williams, E. R., 1989: The tripole structure of thunderstorms. *J. Geophys. Res.*, **94**, 13 151–13 167.

Williams, E. R., 2006: Problems in lightning physics, the role of polarity asymmetry. *Plasma Sources Science and Technology*, **15 (2)**, S91.

Williams, E. R., V. Mushtak, D. Rosenfeld, S. Goodman, and D. Boccippio, 2005: Thermodynamic conditions that lead to superlative updrafts and mixed-phase microphysics. *Atmos. Res.*, **76**, 288–306.

Williams, E. R. and S. Stanfill, 2002: The physical origin of the land-ocean contrast in lightning activity. *Comp. Rend. Phys.*, **3**, 1277–1292.

Williams, E. R., M. E. Weber, and R. E. Orville, 1989: The relationship between lightning type and convective state of thunderclouds. *J. Geophys. Res.*, **94**, 13 213–13 220.

Wilson, C. T. R., 1920: Investigations on lightning discharges and on the electric field of thunderstorms. *Phil. Trans. Roy. Soc. London*, 73–115.

Workman, E. J. and S. E. Reynolds, 1950: Electrical phenomena occurring during the freezing of dilute aqueous solutions and their possible relationship to thunderstorm electricity. *Phys. Rev.*, **78**, 254–259.

Yuter, S. E. and R. A. Houze Jr, 1995: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Monthly weather review*, **123 (7)**, 1941–1963.

Zajac, B. and S. A. Rutledge, 2001: Cloud-to-ground lightning activity in the contiguous United States from 1995 to 1999. *Mon. Wea. Rev.*, **129**, 999–1019.

Zhang, J., K. Howard, C. Langston, and coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) System: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338.

Zipser, E. J. and K. R. Lutz, 1994: The Vertical Profile of Radar Reflectivity of Convective Cells: A Strong Indicator of Storm Intensity and Lightning Probability? *Mon. Wea. Rev.*, **122**, 1751–1759.