

2018

An analysis of the lightning jump algorithm using the GOES-16 geostationary lightning mapper

Nathan L. Curtis

Follow this and additional works at: <https://louis.uah.edu/uah-theses>

Recommended Citation

Curtis, Nathan L., "An analysis of the lightning jump algorithm using the GOES-16 geostationary lightning mapper" (2018). *Theses*. 233.
<https://louis.uah.edu/uah-theses/233>

**AN ANALYSIS OF THE LIGHTNING JUMP
ALGORITHM USING THE GOES-16 GEOSTATIONARY
LIGHTNING MAPPER**

by

NATHAN L. CURTIS

A THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Atmospheric Science
in
The Department of Atmospheric Science
to
The School of Graduate Studies
of
The University of Alabama in Huntsville

HUNTSVILLE, ALABAMA

2018

In presenting this thesis in partial fulfillment of the requirements for a master's degree from The University of Alabama in Huntsville, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by my advisor or, in his/her absence, by the Chair of the Department or the Dean of the School of Graduate Studies. It is also understood that due recognition shall be given to me and to The University of Alabama in Huntsville in any scholarly use which may be made of any material in this thesis.



Nathan L. Curtis

12-12-18

(date)

THESIS APPROVAL FORM

Submitted by Nathan L. Curtis in partial fulfillment of the requirements for the degree of Master of Science in Atmospheric Science in Atmospheric Science and accepted on behalf of the Faculty of the School of Graduate Studies by the thesis committee.

We, the undersigned members of the Graduate Faculty of The University of Alabama in Huntsville, certify that we have advised and/or supervised the candidate of the work described in this thesis. We further certify that we have reviewed the thesis manuscript and approve it in partial fulfillment of the requirements for the degree of Master of Science in Atmospheric Science in the Department of Atmospheric Science.

Lawrence Carey 12/5/2018 Committee Chair
Dr. Lawrence Carey (Date)

Christopher Schultz 12/4/2018
Dr. Christopher Schultz (Date)

 Dr. Phillip Bitzer 2018/12/04
(Date)

John R. Mecikalski _____ Dr. John Mecikalski (Date) Department Chair

Emmanuel G. W. Shello
Dr. Sundar Christopher 12/14/2018 College Dean
(Date)

 Dr. David Berkowitz 1/21/19 (Date) Graduate Dean

ABSTRACT

School of Graduate Studies
The University of Alabama in Huntsville

Degree Master of Science College/Dept. Science/Atmospheric Science
in Atmospheric Science

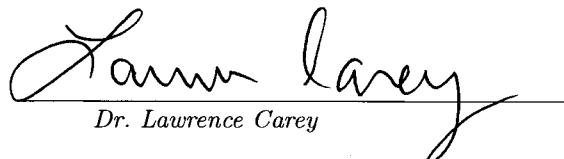
Name of Candidate Nathan L. Curtis

Title An Analysis of the Lightning Jump Algorithm
Using the GOES-16 Geostationary Lightning Mapper

The objective of this study is to implement the two-sigma lightning jump algorithm (LJA), initially developed using Lightning Mapping Arrays (LMAs), with GOES-16 Geostationary Lightning Mapper (GLM) flashes, evaluate its performance, and identify any needed adjustments to the algorithm to optimize operational skill. The GLM is projected to have lower detection efficiency (DE) (70-90 percent) than operational LMAs (95-99 percent). The reduced GLM DE coupled with the coarser spatial resolution of the GLM could have impacts on flash rates and trends that could affect the LJA in various ways. The LJA between the two systems are looked at in both a context of severe storm reports and the radar intensity metrics of MESH and VIL. Initial comparisons between the GLM and LMA on a subset of 5 severe and 2 non-severe storms show a number of differences in flash rates and trends between the two. These differences are maximized in extraordinarily intense storms with high flash rates in the LMA. It was found that LMA was better correlated with MESH (VIL) at 0.498 (0.484) than the GLM was at 0.225 (0.238). Flash rate differences affected the jumps using the two lightning measurement systems in various ways with

the most important being the increase in jump frequency in the GLM relative to the LMA during weak non-severe storms and decaying storms. This increase in jumps in the GLM for non-severe and decaying storms are an indication that the conceptual model of the LJA is failing on the GLM and is likely due to one or both of the following classes of error: GLM flash measurement error or GLM jump algorithm error. These differences and multiple sources of potential error suggest that a larger sample size study must be conducted in order to properly evaluate and optimize the LJA with the GLM. This larger study analyzed 930 storms and when verified against storm reports at its original configuration yielded a POD of 67.9%, FAR of 82.2%, and CSI of 0.16. While the POD was close, but slightly lower, to other studies using LMA and GLM proxy data, the FAR was 20% higher than any prior study. The high FAR is an important finding and is looked into more in-depth. It was found that false alarms in the GLM are found to be mainly associated with weaker MESH and VIL values than hits indicating that the high FAR is likely influenced by physical model, GLM measurement or jump algorithm errors.

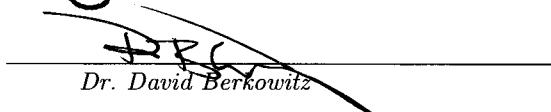
Abstract Approval: Committee Chair


Dr. Lawrence Carey

Department Chair


Dr. John Mecikalski

Graduate Dean


Dr. David Berkowitz

ACKNOWLEDGMENTS

I would first like to acknowledge my committee, Dr. Larry Carey, Dr. Chris Schultz, and Dr. Phil Bitzer, for all of their help and guidance throughout the last two years so that I was able to get to this point of writing this thesis. Dr. Carey you were an incredible advisor and motivator. Not only did you help guide me through the science aspects of becoming a researcher in this field, but also through just the overall day to day grind of graduate school. You also did an incredible job in helping keep my confidence up, especially during the times where I was trying to work a full time job as well and felt many times I was in over my head. Dr. Schultz, you got to be the lucky one that had their office closest to mine and thus got to be the first person I came to when I had any questions or comments regarding anything since you were a simple walk down the hall. Your door was always open, and you know I utilized that to the fullest extent, and I sincerely appreciate you always finding time to talk when needed even when you were busy with your own work. Also I can never thank you enough for pushing me to apply for the NASA Pathways position that I ended up receiving. I didn't personally feel yet that I had the ability to receive an offer of that magnitude and may not have even applied if it wasn't for your guidance. Dr. Bitzer, I thank you also for always being able to set time aside to meet and answer any questions I had. Those few meetings we had were full of great talks and information that has allowed me to reach this point.

Next, I would like to acknowledge my colleagues and educators from my undergraduate years at Indiana University. Starting with Robert Conrick and Dr. Paul Staten, you were the first to introduce me to the world of research. I always thought I wanted to be an operational meteorologist with the weather service and that I wouldn't be happy with anything else. You both let me into a research project you had both been working on and allowed me to go through the process of presenting at a conference and getting published. This led to a newfound interest in research that led me to where I am today. I would also like to thank Dr. Cody Kirkpatrick for being one of the most incredible mentors I have ever had. Dr. K not only were your classes incredibly well-done and informative to the point I still use your notes today, but you always just went the extra mile at every opportunity you could. From staying after class to chat with me more in depth about some things we went over, to creating extra classes (and thus a ton of extra work for yourself) to make sure a few of us got what we needed to further our education/careers, and finally just being someone I knew I could always talk to. I'm not sure I would have even made it through my undergraduate degree if it weren't for you.

I would like to also acknowledge all my colleagues and friends here at the University of Alabama Huntsville. There are way too many of you to list you all, but thanks for making graduate school a great experience. It is invaluable getting help and just being around other people going through the same stages we were all going through. There are a few of you I would like to specifically mention. First and foremost I'd like to thank Sarah Stough. Sarah, you were always open and welcoming to me coming by or sending you an e-mail and asking questions even though I know

you are super busy with your own work. You helped me a lot with a lot of the code that I used in this project, especially pertaining to WDSS-II. Also, you even did some extra analysis for me that I used in the LMA height and size properties. Thank you for just making this process easier for me in anyway you could. I'd also like to thank Austin Clark as well for being a fantastic friend and colleague. You were always someone I could vent to or just talk to about school that made things less stressful. You also were the one who had no problem calling me out when I was being over-dramatic or overly stressed about something I shouldn't be! It again just made the whole process more easy and more fun!

Lastly, I would like to thank my friends and family. Again there are too many people here to list everyone, but there are a few specific people I would like to call out. First to my immediate family, my Mom, Dad, and sister Jessica. You guys always supported my dream of becoming a meteorologist and were there for all of the ups and downs of the journey. More importantly you guys helped shape me into the person and professional I am today and absolutely would not be here if it wasn't for your love and support. Last, but surely not least, I have to thank my wonderful fianceé Tiffany. You have also continued to support me throughout this journey no matter what it took. I woke up one day deciding that I wanted to go to Alabama for graduate school and you had zero hesitation in supporting and joining me on this journey. You have been so patient with me as we put a lot of our life goals on hold so I could get through school. I'm not sure how you put up with me the last five years, but I'm so happy to be able to call you my best friend.

This project was supported by the NOAA GOES-R Series Risk Reduction Research (R3) and NASA MSFC-UAH Cooperative Agreement (NNM11AA01A).

TABLE OF CONTENTS

	Page
List of Figures	xiii
List of Tables	xvi
 Chapter	
1 Introduction	1
2 Background	4
2.1 Non-Inductive Charging	4
2.1.1 Overview	4
2.1.2 Charging of Particles	5
2.1.3 Charge Structure	7
2.1.4 Relation to Thunderstorm Intensity	8
2.2 Lightning's Relation to Severe Weather	9
2.3 Lightning Jump Algorithm	12
2.4 Lightning Detection	14
2.4.1 Lightning Mapping Arrays	15
2.4.2 Geostationary Lightning Mapper	16
2.4.3 Key Differences	17
2.5 Main Objectives	18

3 Data and Methodology	19
3.1 Lightning	19
3.2 Radar	20
3.3 Environmental	21
3.4 Domain	22
3.5 WDSS-II	23
3.5.1 Gridding	23
3.5.2 LMA Flash Clustering	23
3.5.3 Cell Tracking (VILFRD)	24
3.6 Verification Methods	27
3.6.1 Storm Reports	27
3.6.2 Radar Intensity Metrics	28
3.7 GLM Versus LMA Comparisons	30
3.8 Large Sample Study	31
4 Results	33
4.1 GLM vs LMA Comparisons	33
4.1.1 Specific Case Studies	33
4.1.1.1 North Alabama April 22, 2017 (Severe)	33
4.1.1.2 Central Oklahoma May 17, 2017 (Severe)	35
4.1.1.3 North Alabama April 22, 2017 (Non-Severe)	38
4.1.1.4 Central Oklahoma May 17, 2017 (Non-Severe)	40
4.1.2 Quantifying Differences in LMA versus GLM Trends and Radar Intensity Metrics	41

4.1.3	LMA Flash Size and Height Properties	43
4.1.4	Summary	48
4.2	Large Sample Size Study	50
4.2.1	Results Using Original LJA Configuration	50
4.2.2	Sensitivity Testing	52
4.2.2.1	Spatial Sensitivity	52
4.2.2.2	Threshold Sensitivity	54
4.2.3	Lightning Jump MESH and VIL Statistics	57
4.2.4	Timing Between First Jump and MESH Thresholds	62
4.2.5	Summary	65
5	Discussion	68
5.1	Additional Speculative Findings and Potential Errors	68
5.1.1	Differing GLM Data Sets	70
5.1.2	Storm Tracking	71
5.1.3	Large Flashes	74
5.1.4	Weak Jumps	75
6	Summary and Conclusions	78
6.1	Summary	78
6.2	Conclusion	80
6.3	Future Work	81
Bibliography		85

LIST OF FIGURES

Figure	Page
2.1 A figure from Saunders et al. (2006) showing the boundaries found between positive and negative polarities for rimers post-impact for various laboratory experiments. The solid line is Saunders et al. (2006), the solid line broken by dots is Pereyra et al. (2000), the dotted line is Takahashi (1978), and the dashed line is Saunders and Peck (1998).	6
2.2 A figure from Schultz et al. (2015) showing time series of total lightning with updraft volume, graupel volume, and updraft speeds for a storm in a storm in central Tennessee on 11 June 2012.	11
2.3 A figure from Stano et al. (2014) showing all of the currently active and proposed LMA's along with their approximate coverage areas.	15
3.1 Map of the continental United States with the large sample size study domain marked by the blue square. Only storms over land in the United States are considered.	21
3.2 Same as Figure 3.1 except showing all of the sub-domains used in this study.	22
3.3 This figure shows the VILFRD tracking process for a set of storms on April 22nd, 2017 at 22:41 UTC. Panel A: Radar Reflectivity (dBZ). Panel B: GLM Flashcounts (flashes per 5 minutes). Panel C: VILFRD Values. Panel D: K-Means Tracked Features.	25
3.4 A time series of GLM flash rates from a storm in N. Alabama on April 22, 2017 that produced a weak EF-1 tornado in Jones Chapel, AL. Stars mark when a lightning jump occurred, green triangles are hail reports, blue squares are wind reports, and red upside down triangles are tornado reports. A green (red) box surrounding a jump star indicates that jump was verified (a false alarm). A green (red) box surrounding storm reports indicates they are hits (misses).	26

4.1 Time series plots of lightning and radar data for a severe storm in North Alabama on April 22, 2017. Top: LMA flash rates (green line), LMA lightning jumps (green stars), GLM flash rates (black line), GLM lightning jumps (black stars). Bottom: VIL (blue line) and MESH (black line). The red triangle denotes the time of tornadogenesis at 2240 UTC.	34
4.2 The same as Figure 4.1 except for a severe MCS in Central Oklahoma on May 17, 2017.	36
4.3 The same as Figure 4.1 except for a non-severe single cell in North Alabama on April 22, 2017.	38
4.4 The same as Figure 4.1 except for a non-severe MCS in Central Oklahoma on May 17, 2017.	39
4.5 Filtered LMA flash and GLM flash differences (blue) compared with flash sizes calculated from LMA (red) for four storms. The four storm locations and their storm id numbers from Table 3.1 from top to bottom: Northern Alabama on 22 April 2017 (ID: 1), Northern Alabama on 22 April 2017 (ID: 2), Central Colorado on 8 May 2017 (ID: 4), and Texas Panhandle on 15 May 2018 (ID: 7). Green (black) boxes highlight areas where there was an decrease (increase) in flash size associated with a increase (decrease) in the LMA GLM difference.	45
4.6 The same as Figure 4.5 except now LMA calculated flash altitudes in red. Green (black) boxes highlight areas where there was an decrease (increase) in flash altitude associated with a increase (decrease) in the LMA GLM difference.	46
4.7 Results of the threshold sensitivity testing on the large sample size study showing POD (top), FAR (middle), and CSI (bottom). Darker blues indicated the higher performing values.	55
4.8 Normalized histograms of MESH values (top) and VIL values (bottom) for hits (blue) and false alarms (green) for the fifteen minutes surrounding a jump. Values range from 0 to 100 in bins of 5 and all zero values are thrown out.	59
4.9 Distributions of the difference in timing between given MESH thresholds and the first jump. A positive (negative) time means the first jump occurred before (after) the given thresholds. The thresholds are from top to bottom: 25.4 mm, 40.0 mm, and maximum.	63

LIST OF TABLES

Table	Page
3.1 Table showing the storms and their attributes used in the GLM vs LMA comparisons	31
4.1 Pearson correlations between the differing lightning systems and radar intensity metrics for each storm and the average Pearson correlations for all seven storms combined.	41
4.2 Pearson correlations calculated between differences in LMA versus GLM flashes and LMA calculated flash size and altitudes for four chosen storms using both the raw and filtered results. The storm locations and ID's are as in Table 3.1	47
4.3 Dates and Times (UTC) of the reprocessed GLM data used for the large sample study	51
4.4 Results of the spatial sensitivity testing, domain numbers are the same as shown in Figure 3.2. NaN's indicate not a number errors due to dividing by 0 (hits and misses both 0).	53
4.5 Mean and median statistics for all MESH and VIL values for hits and false alarms for the fifteen minutes surrounding a jump. Values of 0 are not considered.	58
4.6 Skill scores for the LJA on the GLM at a two-sigma 10 flash per minute minimum threshold for varying minimum MESH thresholds.	60

Chapter 1

INTRODUCTION

Lightning's relation to severe weather has been studied at least as far back as the 1960's (Kohl 1962; Stanford et al. 1971; Goodman et al. 1988; MacGorman et al. 1989). The non-inductive charging mechanism for electrification of thunderstorms has been widely accepted as the primary charging mechanism in linking lightning and severe weather due to its consistent ability to predict expected behavior in laboratory, observational, and modeling studies. In non-inductive charging, thunderstorm electrification is achieved due to collisions and subsequent charge separation of graupel and small hail particles with ice crystals in the presence of supercooled liquid water (Takahashi 1978; Saunders et al. 2006). A key result of Goodman et al. (1988) and MacGorman et al. (1989) was that measurements of total lightning are needed to asses the link between lightning and severe weather as there were no trends found in cloud-to-ground measurements only.

More recent advancements in technology have brought forth Lightning Mapping Arrays (LMAs) (Rison et al. 1999) that can detect 90% or greater of total lightning (intra-cloud and cloud-to-ground lightning flashes) using an array of Very High Frequency (VHF) sensors within a 125 km range of the center of the array

(Chmielewski and Bruning 2016). Total flash rate data is a key in linking severe weather to lightning production (Schultz et al. 2011). Using these total flash rate measurements from LMAs an automated algorithm, named the Lightning Jump Algorithm (LJA), was developed with the purpose of predicting severe weather by measuring rapid, two-sigma increases in total lightning (Schultz et al. 2009). The largest drawback of this algorithm is due to the high costs of building and maintaining LMAs, coupled with their limited field of views (FOVs) of only up to 125 km, that there are very few places the LJA can actually be implemented.

With the 2016 launch of the Geostationary Lightning Mapper (GLM) on board the GOES-16 satellite there is now access to hemispheric total flash rate data. Despite coarser spatial resolution versus LMA coupled with other potential challenges for the GLM such as; flashes with a bright background during the day, low altitude flashes, and accurately separating relatively small flashes in higher flash rate storms, the GLM is theorized to detect 70%-90% of total lightning (Goodman et al. 2013). Due to its relatively high detection efficiency (DE) given its wide FOV, the GLM is a good candidate to apply the LJA to in the near-future.

Remembering that the LJA was originally built and tested to run on LMA networks it is important to understand the differences between LMAs and the GLM and how they work with the LJA. With the lower DE of the GLM due to the aforementioned reasoning, the two instruments are also measuring completely different properties of lightning. LMA detects sources of very high frequency (VHF) electromagnetic radiation produced by lightning, while satellite based instruments, like the GLM, detect optical radiation produced by lightning (Nag et al. 2015). Due to these

differences it is important to compare and understand the differences in the GLM vs LMA flash rates and trends to identify any needed adjustments to the LJA to optimize operational skill of the LJA with the GLM.

Chapter 2

BACKGROUND

2.1 Non-Inductive Charging

2.1.1 Overview

The non-inductive charging mechanism is the leading charging process that explains the majority of observed and modeled electrical and lightning behavior in thunderstorm (Takahashi 1978; Baker and Dash 1994; Saunders et al. 2006). In non-inductive charging the separation of charged particles and consequent electrification of a storm is done in the absence of a background electric field. This separation of charge is caused by collisions of small ice crystals and small hail and/or graupel in the presence of supercooled water (Takahashi 1978). It is theorized that in the presence of supercooled water that there exists a quasi-liquid layer on ice particles, and during collisions there is a transfer of mass through this layer that leads to opposite polarities of the particles involved in the collision (Baker and Dash 1994; Saunders et al. 2006). Theories on how it is determined which particle gets which charge and how that leads to thunderstorm electrification will be explored further in this section.

2.1.2 Charging of Particles

Takahashi (1978) was one of the first attempts at measuring which particle will get which polarity in the non-inductive charging mechanism. Laboratory cold room experiments were conducted that simulated thunderstorm conditions with both ice and supercooled drops existing in the chamber with electric probes measuring the electric field. It was found during these experiments that to produce charge high enough for thunderstorm electrification that ice and supercooled water had to coexist in the chamber. This led to sensitivity studies to identify the polarity of charge transferred to graupel during collisions with ice crystals in the presence of supercooled water.

Takahashi postulated that cloud water content and temperature both played a role in determining particle polarity after collisions. This is later found to be due to the effect these two variables have on the physical structure of the particles and theoretical quasi-liquid layers (Baker and Dash 1994; Saunders et al. 2006). The colliding particles were split into two categories: the rimer (small hail and/or graupel) and the impactor (deposition ice). Due to the nature of the electric probes used, only the polarity of the rimers were captured, and it was assumed the impactors had opposite polarity. It was found that both cloud water content and temperature played a role in post-collision polarities. It was shown that in general rimers take a negative charge except outside of a region of modest to high cloud water content (roughly .2 to 2 gm^{-3} depending on the study) collocated with cold temperature from -10 °C to -30 °C.

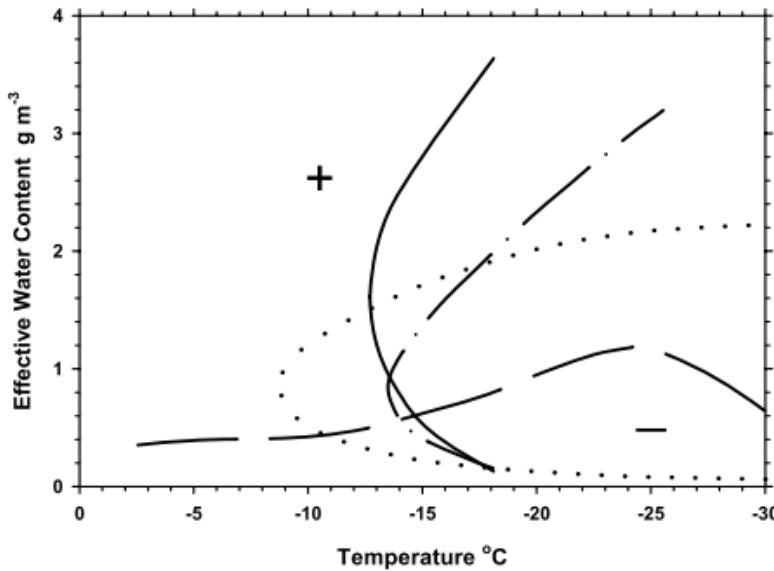


Figure 2.1: A figure from Saunders et al. (2006) showing the boundaries found between positive and negative polarities for rimers post-impact for various laboratory experiments. The solid line is Saunders et al. (2006), the solid line broken by dots is Pereyra et al. (2000), the dotted line is Takahashi (1978), and the dashed line is Saunders and Peck (1998).

Given the results of Takahashi (1978) it was important to try and understand physically why there was a change in polarity of the rimer based off of the background cloud water content and temperature during collisions. Baker et al. (1987) proposed a theory that relative diffusional growth rates, which change due to background thermodynamics such as temperature and water content, of impacting ice surfaces determined charge polarity. It was hypothesized that whichever ice particle was growing faster via diffusion during impact would become the positively charged particle. Williams et al. (1991) and Saunders et al. (2001) looked at accretion and supersaturation respectively during this process and came to conclusions that reinforced the relative growth rate hypothesis.

Since the original study of Takahashi (1978), there have been additional laboratory studies to try and reinforce and refine the initial findings of rimer charge polarity with temperature, cloud water content, and other properties related to characteristics of the ice and cloud water (e.g., Saunders and Peck 1998; Pereyra et al. 2000; Saunders et al. 2006). Figure 2.1 shows the result of all of these studies on one chart. The biggest takeaway from this figure is that there appears to be a reversal of polarity starting at temperatures ranging from -10 °C to -18 °C that could possibly be absent at higher cloud water contents.

2.1.3 Charge Structure

Knowing the physical basis of initial charge separation and that the polarity of the rimer versus impactor changes based on temperature and cloud water content, it is important to try and understand the way charge is structured throughout a thunderstorm. The leading hypothesis is that charge structure is determined by gravitational sedimentation caused by differential particle vertical motions which is due to a combination of vertical motions within the storm caused by the storm's updraft and differing particle fall speeds. Ice, small hail, and graupel have differing sizes and densities and thus the differing terminal fall speeds. Gravitational sedimentation aided by vertical motions then separates the charge in the storm scale into regions of net positive and negative charge (Williams and Lhermitte 1983). Conceptually there would be expected a region of positive charge at upper levels and negative charge at mid to low levels due to the lighter and normally positively charged ice crystals being lofted higher than heavier and normally negatively charged small hail and graupel.

This hypothesis was reinforced by in situ measurements from aircraft in storms in Montana shortly after (Gardiner et al. 1985). In later studies a tripole structure was hypothesized starting with an upper positively charged region, a middle negatively charged region, and lower positive charged region. The lower positive charged region was hypothesized to occur due to the positive charging of larger graupel and small hail particles, which usually occur at warmer temperatures lower in the storm (Jayaratne et al. 1983; Williams 1989).

While the tripole structure is still considered the idealized conceptual model of charge structure in a normal thunderstorm, many recent studies have shown that charge structure is a lot more complex. A study using electric field soundings through the convective regions of thunderstorms found the charge structures can be much more complex and consist of more than just three charge regions (Stolzenburg et al. 1998). A more recent numerical cloud modeling study shows just how complex charge can be in thunderstorms given non-inductive charging and gravitational sedimentation (Mansell et al. 2005). A study using measurements from LMAs and simulated measurements shows that there are many different combinations of dipole and tripole structures seen in storms and those that deviate from the conceptual positive-negative-positive model are dubbed to be of anomalous polarity (Bruning et al. 2014).

2.1.4 Relation to Thunderstorm Intensity

As a thunderstorm becomes more intense there are kinematic and microphysical properties of the storm that can lead to the enhancement of the non-inductive charging process within the storm. As an updraft increases in strength there should

be an increase in the size of the mixed phased updraft due to an increase in super-cooled water being lofted by the updraft. Additionally, there is an increase in ice, graupel, and small hail with increasing updraft intensity (Carey and Rutledge 1996). There is also an increase in turbulence with increasing updraft strength due to larger horizontal gradients of vertical motion (Holton 2004). With increased turbulence and supercooled and ice particles, conceptually it should increase the amount of collisions within the storm and thus increase the overall charge (Bruning and MacGorman 2013). The next section will dive deeper into the link between non-inductive charging and thunderstorm intensity.

2.2 Lightning's Relation to Severe Weather

Early studies in relating lightning to severe weather showed little utility in using cloud-to-ground lightning measurements, but hinted at maybe a stronger correlation between total lightning and severe weather (Kohl 1962; Stanford et al. 1971; Goodman et al. 1988; MacGorman et al. 1989; Carey and Rutledge 1998). As newer technology evolved there were improvements in measuring the total lightning of the storm. Williams et al. (1999) was one of the first studies that quantified total lightning flash counts to the occurrence of severe weather over a larger sample. It was found in this study that there were peaks in flash rates between 5 and 20 minutes prior to severe weather reports on the ground. This study also coined the “lightning jump” term and defined it as “abrupt increases in flash rate in advance of the maximum flash rate for the storm”.

Following this landmark study there was an increase in the quality and amount of LMAs that map total lightning using VHF sensors (Rison et al. 1999). This advancement in technology has allowed for more in depth studies linking lightning to thunderstorm intensity. Deierling et al. (2008) looked at total lightning measurements and thunderstorm updraft characteristics with multi-Doppler radar. They found there was a high correlation ($r = 0.93$) between total updraft volume and total lightning, also there were positive, but not as high, correlations between maximum updraft velocity and total flash rate ($r = 0.69$).

Dual-polarization radar studies also showed high correlations between total flash rates and graupel mass and updrafts greater than or equal to 10 ms^{-1} in the $-10 \text{ }^{\circ}\text{C}$ to $-40 \text{ }^{\circ}\text{C}$ range (Carey and Rutledge 1996, 2000; Schultz et al. 2015, 2017). Figure 2.2 shows a total flash rate time series with various microphysical and kinematic properties of the updraft from Schultz et al. (2015). It is important to note from this figure that increases in updraft characteristics tended to precede increases in total lightning on the order of 5-10 minutes.

Lastly, other studies have also linked total flash rates to radar intensity metrics as it is easier to validate without needing to be in a close baseline multi-Doppler lobe for updraft microphysics and kinematics or using ground reports. Chronis et al. (2015) looked at the relationship between total flash rate and the Maximum Expected Size of Hail (MESH) over a large sample of over 700 storms across three LMAs in Washington D.C., Alabama, and Oklahoma. It was found that thunderstorms with at least one lightning jump had a range of average MESH values from 11.0 to 18.0 mm while storms with no lightning jumps had average MESH values from 6.5 to 10.0

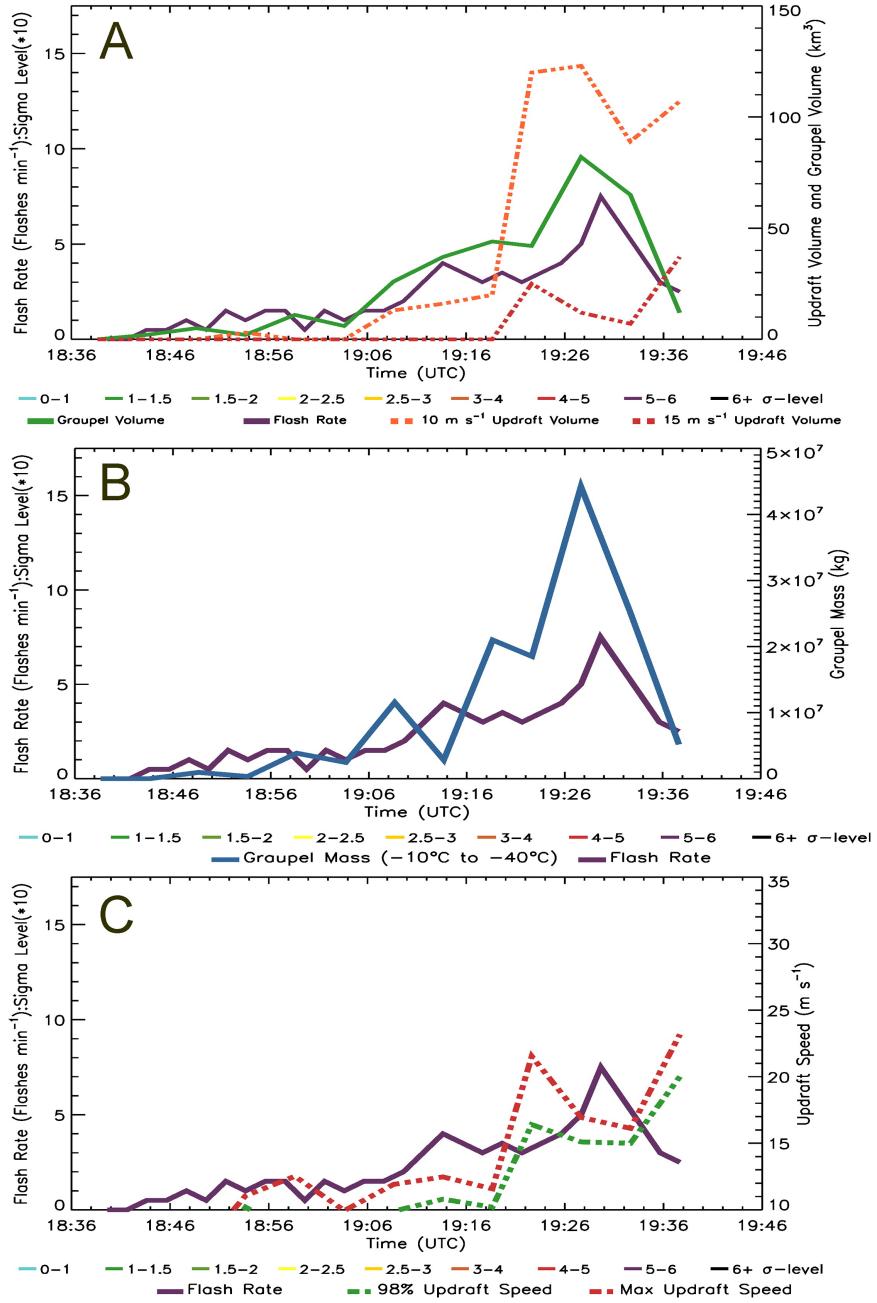


Figure 2.2: A figure from Schultz et al. (2015) showing time series of total lightning with updraft volume, graupel volume, and updraft speeds for a storm in central Tennessee on 11 June 2012.

mm. Thus it was found that there was again a positive correlation between total lightning and intensity of the thunderstorm.

These studies suggest a physically and statistically significant link between thunderstorm updraft intensity and total flash rates. With this link solidified the next step was to see how this relationship could be used to assist in the operational forecasting of severe weather.

2.3 Lightning Jump Algorithm

With both a conceptual and physical link found through many laboratory and field studies between total lightning and thunderstorm intensity the natural next step was to determine if there was any way to implement this relationship in operational severe weather forecasting. Studies in Florida showed that there was a jump, or localized peak, in lightning preceding severe weather reports on the ground (Williams et al. 1999). Steps were taken to see if there was a way to automatically detect these jumps in lightning that lead to severe weather. Initially three different algorithms were tested and verified using severe storm reports (Schultz et al. 2009). The three algorithms were mixes of identifying increases in total flash rates over selected standard deviation thresholds given different temporal and minimum flash rate thresholds. The algorithm that had the best verification skill scores was called the sigma algorithm.

From Schultz et al. (2009) the sigma algorithm looked for increases in the total lightning flash rates greater than or equal to set standard deviation thresholds over twelve minute moving windows for given minimum flash rate thresholds. Through sensitivity testing it was found that two sigma (two standard deviation) increases

in lightning with a minimum flash per minute threshold of 10 flashes per minute performed the best. An overview of the algorithm is as follows:

- Total flash rates are binned into 2-minute time periods.
- Consecutive two minute bins are subtracted from each other to find a time rate of change in total flash rate (DFRDT).
- A standard deviation of the last five DFRDT bins prior to the current time is taken.
- A ratio between the current DFRDT bin and the standard deviation of the last five DFRDT bins is taken, and that is the sigma level.
- If more than 14 minutes of data has been recorded, the flash rates are at least 10 flashes per minute, and the sigma level is 2 or greater, then that time period is considered a lightning jump.

This jump algorithm was validated in four separate ways. Schultz et al. (2009) put a 45 minute warning out after each jump, during each warning each severe storm report was counted as a hit, and every jump without a severe weather report was a false alarm. A single storm report could only verify a single jump and would verify the warning that started the earliest in the case of overlapping warnings. Storm reports were binned into 6-minute increments. A miss was a storm report that fell outside a jump warning window. Using this method and 107 manually tracked thunderstorms the jump had a probability of detection (POD) of 87% and a false alarm ratio (FAR) of 33%. Schultz et al. (2011) used the same verification methods on a larger sample

of 711 manually tracked storms and found a POD of 79% and a FAR of 36%. Schultz et al. (2016b) mirrored its verification more like the National Weather Service and used the previous methodology except one event could verify multiple jumps and reports were not binned in 6 minute increments. This study also used an automated tracking algorithm of the same days and regions of Schultz et al. (2011), but analyzed all of the storms during the periods and locations in the prior studies, along with GLM proxy data and found a POD of 69% and FAR of 63%.

One of the most important findings was that when using an automated algorithm, GLM proxy data, and a larger and more objectively selected sample that the POD decreased and FAR increased significantly. Exact reasons why were not verified, but could be due to the differences in the lightning data used.

2.4 Lightning Detection

Up to this point a majority of the lightning jump studies have been conducted using LMAs as they have had the widest domain of total lightning measurements. This study aims to explore the possibility of running the LJA on the newly launched GLM. In order to move the LJA to a new instrument it is important to further understand the intricacies of the instruments and the differences between them to determine any potential challenges or differences in how the LJA will perform on the new instrument.

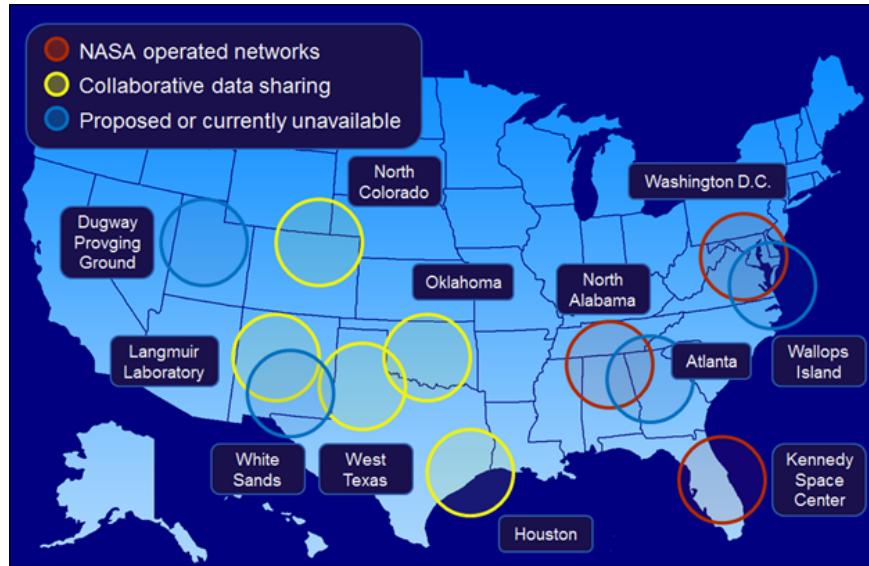


Figure 2.3: A figure from Stano et al. (2014) showing all of the currently active and proposed LMA's along with their approximate coverage areas.

2.4.1 Lightning Mapping Arrays

LMAs were first introduced in 1999 in Central New Mexico (Rison et al. 1999).

LMAs consist of arrays of sensors that detect very high frequency (VHF) electromagnetic radiation. These sensors are able to detect intra-cloud and cloud-to-ground lightning at a 90% or greater detection efficiency within a 125 km range of the center (Chmielewski and Bruning 2016). These sensors are in contrast to long baseline very low frequency and low frequency sensors (e.g. NLDN, ENTLN, etc.) that have an overall high DE with cloud-to-ground strikes and relatively low, when compared to LMA, DE with intra-cloud lightning, but at a much wider domain (Nag et al. 2015). Spatial and temporal resolutions in LMAs are expressed by their location and timing uncertainties. At a 125 km range LMAs have horizontal uncertainties of 400-800 m, vertical uncertainties of 500-1000 m, and temporal uncertainties of 80-100 μ s (Koshak

et al. 2004; Thomas et al. 2004). The largest drawbacks of LMAs are their relative high costs to operate and maintain coupled with their relatively small field of views. Figure 2.3 shows that as a result of the previous point there is relatively sparse coverage of total lightning measurements needed to implement the LJA across the United States.

2.4.2 Geostationary Lightning Mapper

The first lightning instrument in geostationary orbit was the GLM which was launched on board the GOES-16 satellite in November of 2016 (Goodman et al. 2013). The GLM detects the optical radiance of lightning in roughly 8 km by 8 km grid boxes at nadir to 14 km by 14 km at the corners over a domain covering most of North and South America. The GLM utilizes a wide FOV lens combined with a Charge Coupled Device (CCD). The CCD has a size of 1372 x 1300 pixels and detects intensities of light at the 777.4 nm near-IR band. Thus the GLM detects lightning by measuring changes in light intensity within the CCD array. The GLM clusters these changes in light into three categories: events, groups, and flashes. Events are any one pixel exceeding a set background light threshold over a 2 ms integration window. It is important to note that while events are an attempt to measure a single optical pulse due to lightning, it is theoretically possible that multiple pulses can occur in the 2 ms threshold. Lightning is often large enough that it will illuminate multiple GLM pixels within the same 2 ms window. Therefore events are also clustered into groups. A group is a series of event pixels that are adjacent to each other during the integration period. In the event of a small optical pulse that only illuminates a single GLM

pixel, the group will consist of just the single event pixel. Finally spatially associated groups from temporally contiguous integration periods are clustered into lightning flashes. More specifically, flashes are clusters of groups within 16 km and 330 ms of each other. These flash clustering thresholds are chosen to try and produce similar results to the definition of a flash on other instruments. It was theorized that the DE of the GLM would be in the 70%-90% range. Many studies since launch have shown the DE of the GLM to vary based on storm type and intensity, location, flash heights, and flash sizes (Bitzer 2018). At the 2018 GLM Science Team Meetings there were presentations that showed DEs ranging from 20% to 100% when compared to LMA across differing storms (Thomas et al. 2018; Trostel et al. 2018; Rutledge et al. 2018; Bateman 2018).

2.4.3 Key Differences

There are a handful of key differences between LMAs and the GLM. First and foremost is they are detecting different properties of lightning. The LMA's VHF electromagnetic radiation is associated with the breakdown of air via channel formation and leader processes, while GLM's optical radiation is associated with electric currents with intense radiation in optical frequencies due to sudden intense heating (Nag et al. 2015). GLM also is much coarser spatially ranging from 8 km by 8 km to 14 km by 14 km grid boxes versus LMAs having location uncertainties from 400 - 1000 m. Although GLM validation is still ongoing, preliminary analysis described earlier has demonstrated that the GLM can have varying DE on a storm by storm basis and even sometimes within the same storm. However, the biggest advantage GLM has

is its large domain allowing for a much larger spatial area with constant total flash rate measurements versus the current LMA technology. Despite these differences, the large domain of the GLM makes it a productive candidate to try and apply the LJA to for operational applications over a wide domain.

2.5 Main Objectives

Given what we understand about how the LJA operates and the differences in how LMAs and the GLM detect lightning the objectives of this thesis are to answer the following important questions regarding the potential use of the LJA with the GLM:

- What are the differences in flash rates and trends between LMA (within its 125 km operational range) and the GLM for the same subset of severe and non-severe storms?
- If large differences exist, how do those differences affect the LJA and its performance?
- How does the LJA perform when run independently on the GLM versus previous verification studies using LMA on a large sample?
- Are there any different combinations of sigma and minimum flash per minute thresholds that improve the performance of the LJA on the GLM?
- Are there any recommendations on next steps for a LJA or a LJA-like algorithm that will perform well on the GLM?

Chapter 3

DATA AND METHODOLOGY

3.1 Lightning

The lightning data used in this project comes from multiple sources. Most prominently flash rate data from the GOES-16 GLM are utilized. The GLM data are flash centroids from the Lockheed Martin reprocessed data set that corrects some of the navigational and timing issues that originally existed in the operational datasets from 2017 (Mach 2017a). This reprocessed data set includes eleven days from the GOES-16 GLM Post-Launch Test Field Campaign, and ten of these days in April and May 2017 are used. In addition, this study uses undecimated VHF source data from three LMAs, the North Alabama LMA (NALMA) (Koshak et al. 2004), the West Texas LMA (WTLMA) (Bruning 2011), and the Oklahoma LMA (OKLMA) (DiGangi et al. 2016), as well as decimated data from the Colorado LMA (COLMA) (Lang et al. 2014) alongside the GLM data. Only data within the 125 km operational range of the LMA are considered due to detection efficiency concerns noted in Chmielewski and Bruning (2016).

3.2 Radar

Level II NEXRAD WSR-88D Radar data will be used for both cell tracking and comparisons to lightning data from 103 different radar sites (Crum and Alberty 1993). The data are obtained through the National Centers for Environmental Information (NCEI) radar archive. Radar data are combined into a multi-radar composite and are used for both tracking and analysis. The two main products used are vertically integrated liquid (VIL) (Greene and Clark 1972) and maximum expected size of hail (MESH) (Witt et al. 1998). VIL is an integral of liquid water content derived from radar reflectivity values at varying elevation angles. MESH is the maximum hail expected from a storm based on several steps using radar reflectivity and temperature profiles. First flux values of hail kinetic energy are found by a reflectivity weighting function used to define a transition zone between rain and hail. The flux is then weighted by a temperature based on the 0 °C and -20 °C height levels. Finally these values are vertically integrated between the height of the 0 °C line and the height of the top of the storm. VIL is mainly used for tracking storms as well as a radar intensity metric. MESH is used exclusively as a radar intensity verification of the jump. MESH has been used as a severe weather indicator both in jump studies (Chronis et al. 2015) and in other studies involving severe weather (Cintineo et al. 2012; Rudlosky and Fuelberg 2013).

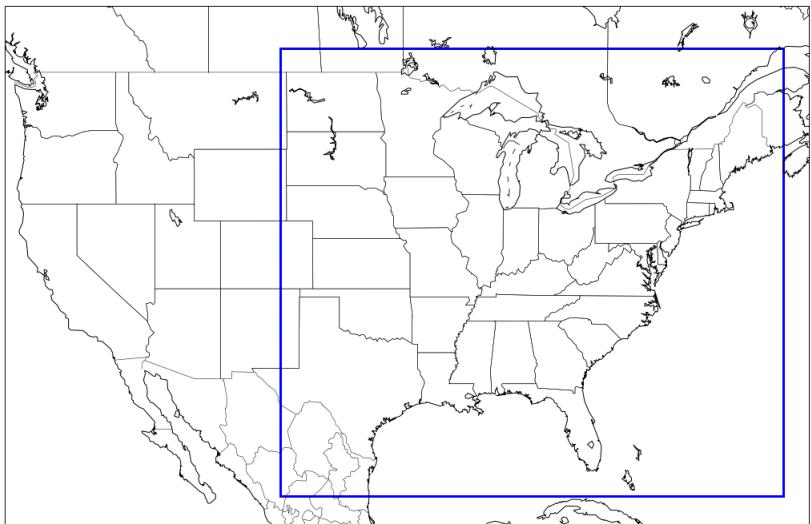


Figure 3.1: Map of the continental United States with the large sample size study domain marked by the blue square. Only storms over land in the United States are considered.

3.3 Environmental

The Rapid Refresh (RAP) model is used to define the environment in each case. Environmental data are updated every hour on the hour for the entire case. These data are mainly used for the calculation of the radar-based MESH. For the calculation of MESH both the 0 °C and -20 °C temperature levels are required. These temperature levels will be taken from simulated soundings at each grid point from this data.

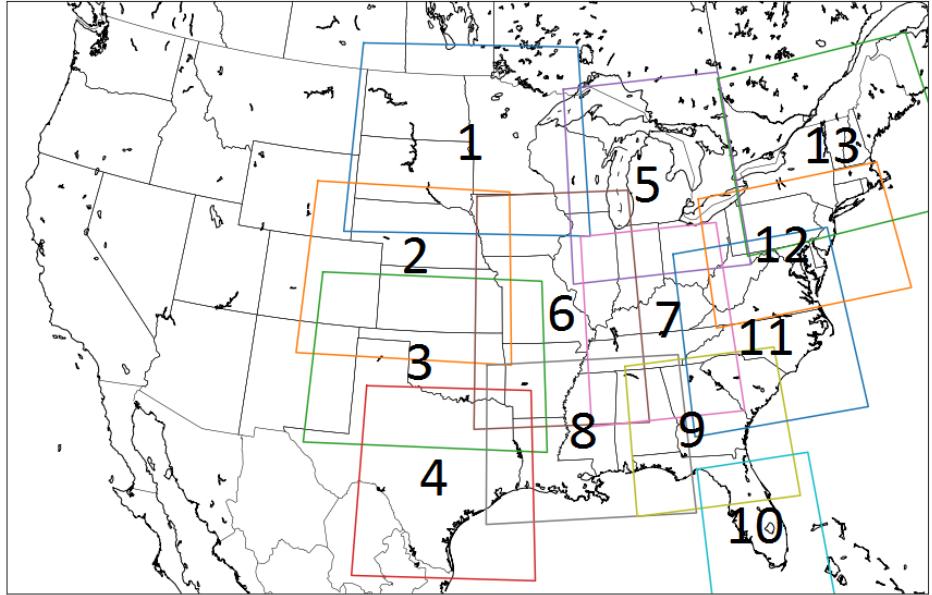


Figure 3.2: Same as Figure 3.1 except showing all of the sub-domains used in this study.

3.4 Domain

The domain for the large sample sized study includes lightning data and radar data for all sites (103 radars) east of the Rocky Mountains and the entire domain can be seen in Figure 3.1. The domain is broken up into 13 sub-domains due to computational restrictions as shown in Figure 3.2. While there is overlap of domains due to the domain requirements of the system used, no storm is counted more than once by using an objective storm merging technique to be explained in a later section. This domain is chosen to include the exact locations, and the locations in between, where previous LJA studies done using LMA were located for the most meaningful comparisons (Schultz et al. 2009, 2011, 2016b; Chronis et al. 2015). Only storms over land in the United States are considered.

3.5 WDSS-II

All gridding, radar calculating, flash clustering, and cell tracking are done within the Warning Decision Support System - Integrated Information (WDSS-II) framework (Lakshmanan et al. 2007). This program allows for the merging of up to nine radars for multi-radar multi-sensor analysis per domain and has a wide array of built-in algorithms that are used in this study ranging from LMA flash clustering to k-means tracking (Lakshmanan et al. 2003).

3.5.1 Gridding

The LMA flash data are gridded in $0.01^\circ \times 0.01^\circ$ grid boxes while GLM data are in $0.08^\circ \times 0.08^\circ$ grid boxes. All radar data are gridded into $0.01^\circ \times 0.01^\circ \times 1$ kilometer grid boxes. All environmental data are gridded into $0.01^\circ \times 0.01^\circ \times 1$ kilometer grid boxes.

3.5.2 LMA Flash Clustering

Unless otherwise mentioned, the LMA flash clustering are accomplished by using the default w2lmaflash algorithm. In this algorithm six stations must detect a source for it to be considered real and a minimum of 10 sources are required for a flash. Events that are collocated are only considered to be a part of the same flash if they occurred within 250 ms from each other. Simultaneous occurring events are only part of the same flash if they are within $5 * (1 + .00001 * (r - 100)^2)$ km of each other where r is the range from the center of the LMA network.

3.5.3 Cell Tracking (VILFRD)

For cell tracking an automated algorithm named VILFRD developed by Schultz et al. (2016b) is used along with WDSS-II's w2segmotionll. VILFRD uses a combination of vertically integrated liquid (VIL) and five minute GLM Flash Rate Densities (FLCT5) to assign values to and track storms. Figure 3.3 shows this process for one of the cases used in this study. Reflectivity-based VIL is calculated and combined with 5-minute GLM flash counts as per Equation 3.1.

$$VILFRD = 100 * \left[\left(\frac{VIL}{45} \leq 1 \right) + \left(\sqrt{\frac{FLCT5}{45}} \leq 1 \right) \right] \quad (3.1)$$

According to Schultz et al. (2016b), this algorithm tracks more based on radar data when flash counts are low and vice versa. For a storm to be considered a trackable feature it must have VILFRD values greater than 20 covering a specified spatial area. There are different scales in which the specified area needed to track ranges from 32 km² to 243 km². Schultz et al. (2016b) found different scale sizes had a direct impact on skill scores in the original study using LMA and GLM proxy data, and that an area of 162 km² (scale 5) yielded the best skill scores. Therefore scale 5 is the spatial scale utilized in this study. WDSS-II outputs storm features of minimum area 162 km² for areas that there is VILFRD greater than or equal than 20. Statistics of flash counts, max MESH, and max VIL, are calculated within each feature over the length of the feature's lifespan. These statistics are what is used in the LJA and verification methods.

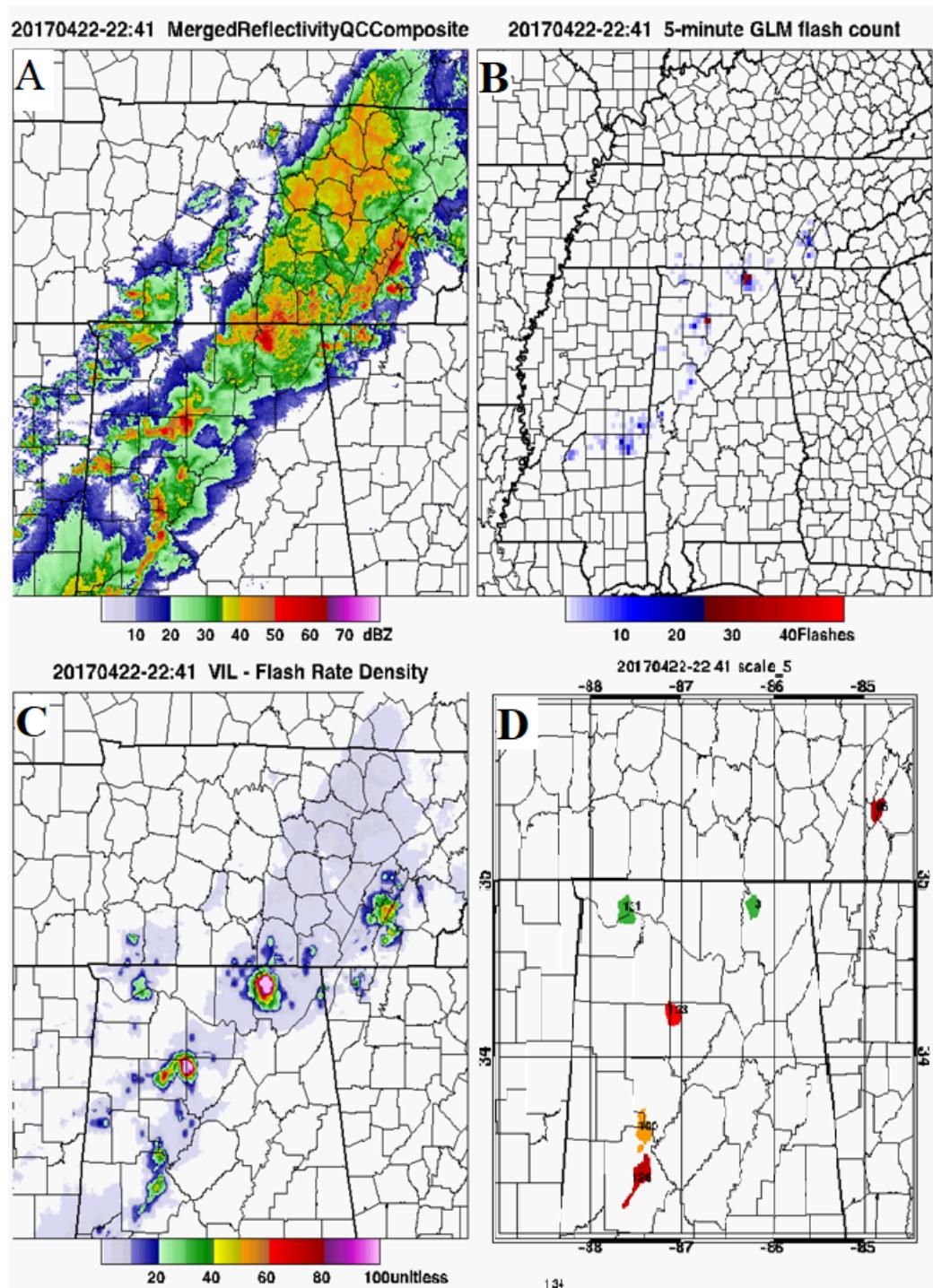


Figure 3.3: This figure shows the VILFRD tracking process for a set of storms on April 22nd, 2017 at 22:41 UTC. Panel A: Radar Reflectivity (dBZ). Panel B: GLM Flashcounts (flashes per 5 minutes). Panel C: VILFRD Values. Panel D: K-Means Tracked Features.

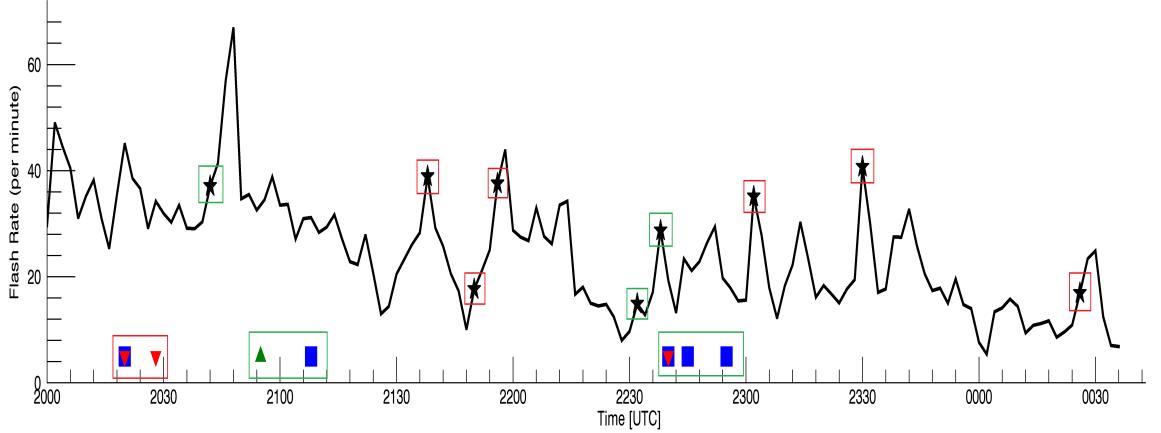


Figure 3.4: A time series of GLM flash rates from a storm in N. Alabama on April 22, 2017 that produced a weak EF-1 tornado in Jones Chapel, AL. Stars mark when a lightning jump occurred, green triangles are hail reports, blue squares are wind reports, and red upside down triangles are tornado reports. A green (red) box surrounding a jump star indicates that jump was verified (a false alarm). A green (red) box surrounding storm reports indicates they are hits (misses).

Due to the overlapping domains and spatial resolution of VILFRD it is necessary to merge tracks to ensure that there is an accurate representation of the storm and that the same storm is not counted twice. Near storm tracks are objectively merged if a new cell begins within 2 minutes of a previous track ending and is within 15 km of the end of that track. If two tracks are within 5 km of each other at simultaneous time steps it is assumed that they are the same storm just across multiple domains and the tracks are merged into one taking the maximum statistics which are assumed to be in the larger feature.

3.6 Verification Methods

3.6.1 Storm Reports

Storm reports are taken from the NCEI Storm Data database. After every jump there is a 45 minute warning placed on the storm starting at the time of the jump. A circle of area equal to the feature area is centered on the feature to find storm reports associated with the particular storm. A 5 km buffer is added to the radius of the circle to account for location errors in the storm reports. Only storm reports that fall within a tracked feature are considered. Storm reports are binned into six minute bins starting at the time of the first report as in Schultz et al. (2011). If two jumps occur within six minutes of each other, only the first jump is used for verification purposes. Every storm report within a 45 minute warning window is considered a hit, and every report outside of a warning window is considered a miss. False alarms are jumps that had no reports within its 45 minute warning. A single storm report can verify multiple jumps as in Schultz et al. (2016b). A mixture of previous storm report verification techniques are used here to follow close to what is actually used in operations while still trying to account for errors in report times and locations. Figure 3.4 shows an example of this storm report verification methodology on a single storm.

Once all of the hits, misses, and false alarms are stored they are then plugged into various statistics to see how well it performed. The three statistics calculated are probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI), and their calculations can be seen in Equation 3.2 (Wilks 2011). Where 100%

represents a perfect POD, 0% represents a perfect FAR, and 1 represents a perfect CSI.

$$POD = \frac{hits}{hits + misses}; FAR = \frac{false}{false + hits}; CSI = \frac{hits}{hits + misses + false} \quad (3.2)$$

There are certain caveats to using storm reports as a verification metric that are important to consider. The first being, just because there is a lack of a storm report does not mean severe weather did not occur. If the storm tracked through a sparsely populated area there may not be anyone there to report that anything happened (Davis and LaDue 2004; Dobur 2005). Another caveat is the timing of reports. Storm reports are either called in by the public and storm spotters or are taken from damage surveys / pictures after the fact. In both scenarios the times are usually estimated to the best ability of the witness or the radar. With that there can be deviations in time from when the severe weather actually occurred to when it was reported to have occurred. These caveats are why this study also looks at radar intensity metrics as a form of verification.

3.6.2 Radar Intensity Metrics

There are a few different ways this study utilizes the MESH and VIL measurements to verify or give more context to how the LJA is performing with the GLM. First for all the cases shown in Table 3.1 average Pearson correlations are taken between GLM flash rates, LMA flash rates, VIL, and MESH to see the relationship

between lightning and radar. This is done because prior studies noted in the Background section have shown positive correlations between various radar statistics and lightning flash rates and trends (Schultz et al. 2017; Chronis et al. 2015; Schultz et al. 2015; Carey and Rutledge 1996, 2000; Deierling et al. 2008). Another is measuring the time between the first jump in the storm and the first appearance of 25.4 mm MESH, 40 mm MESH, and maximum MESH. The MESH thresholds represent the severe hail and significant severe hail thresholds as defined by the National Weather Service. These are taken to determine if jumps are occurring at times that fit the conceptual model of the jump. The first jump of the storm should be occurring before the larger and maximum MESH thresholds as it should have responded to the prior intensification leading up to these values. If the first jump is occurring after these thresholds it may indicate an issue with the LJA as it is not jumping during peak intensification. Lastly, MESH and VIL values along the fifteen minutes centered on the jump time are taken to see if there is any relationship between radar intensity and if a jump was verified or a false alarm. As noted in previous sections it has been found that storms with jumps have higher average MESH values than those without. Thus if a majority of the false alarms are occurring in storms with weaker MESH values it points to a physical model issues with the LJA on the GLM, likely driven by GLM measurement or jump algorithm system errors, rather than a problem in storm report times and locations.

There is a caveat to using MESH as well. MESH tends to overestimate the hail size that falls to the ground (Witt et al. 1998), but it can still give a good indication of how strong the storm is.

3.7 GLM Versus LMA Comparisons

One of the main objectives is to see how the GLM and LMA flash counts and trends compare and contrast to one another. The overall purpose of this section is to see if there are major differences in the LJA between LMA and the GLM. For this seven different storms are subjectively chosen consisting of different types, locations, and severity. These storms are analyzed with both GLM and LMA flash rate statistics as well as MESH and VIL values. These storms can be seen in further depth in Table 3.1. The LJA is ran for each storm using both LMA and GLM data to see how the number of jumps and the timing of jumps between the two compare.

For all storms in this section averaged Pearson correlations between GLM flash rates, LMA flash rates, MESH, and VIL values are taken. For a select few of these storms there are additional statistics on LMA flash size and height taken using LMATools (Fuchs et al. 2016). In this portion flashes are clustered using the following parameters: at least 7 stations have to see a source, distance threshold of 3 km, a timing threshold of 3 s, and a maximum source altitude of 20 km. Flash statistics are then determined using a convex hull technique. These statistics are used to attempt to diagnose what the physical attributes of the flashes are between when there is good agreement and bad agreement between GLM and LMA flash rates. If there exists a clear correlation between flash size and/or height and GLM versus LMA DE it could have direct impacts on the performance of the LJA on the GLM as it will cause fluctuations in total flash rate counts that may not be physical.

Table 3.1: Table showing the storms and their attributes used in the GLM vs LMA comparisons

Storm #	Location	Date	Time (UTC)	Type	Severe?	LMA Flash Size/Height?
1	N. Alabama	April 22, 2017	1936-0032	Semi-Isolated Cell	Yes	Yes
2	N. Alabama	April 22, 2017	2100-0036	Semi-Isolated Cell	Yes	Yes
3	N. Alabama	April 22, 2017	2103-2141	Isolated Cell	No	No
4	Central Colorado	May 8, 2017	2024-2242	Isolated Cell	Yes	Yes
5	Central Oklahoma	May 17, 2017	0552-0734	Linear	Yes	No
6	Central Oklahoma	May 17, 2017	0433-0519	Linear	No	No
7	Texas Panhandle	May 15, 2018	2200-0312	Isolated Cell	Yes	Yes

3.8 Large Sample Study

A large sample study is also conducted with the overall purpose to independently verify how the LJA performs on GLM flash rates. The large sample consists of 930 storms and 402 storm reports across the ten GOES-R Post Launch Field Campaign days in April and May 2017 in the domains and sub-domains shown in Figure 3.1 and Figure 3.2. Several verification methods are ran on the large sample as described in detail in Section 3.6 and compared to prior LJA verification studies (Schultz et al. 2009, 2011, 2016b; Chronis et al. 2015).

The reasoning for a large sample size study is it helps to reduce errors in smaller scale anomalies and errors in the verification methods. However, an error is added on the topic of radar quality. In such a large study it is hard to know if the radars are operating properly, however this error should be minimal due to WSR-88D's having a nominal up time of 95%+ (Horvat et al. 2011).

Sensitivity testing is conducted on the storm report verification method to find what changes to the LJA, if any, yield the best results. Parameters tweaked in the

sensitivity testing are the 10 flash per minute threshold to initiate the algorithm and the 2-sigma value that identifies a jump. The minimum flash per minute threshold is varied from 0 to 25 flashes per minute in increments of 1 and the sigma value is varied from 1.0 to 4.0 in increments of 0.1. For each pair of values there are new FAR, POD, and CSI values calculated.

Chapter 4

RESULTS

4.1 GLM vs LMA Comparisons

4.1.1 Specific Case Studies

4.1.1.1 North Alabama April 22, 2017 (Severe)

April 22, 2017 had multiple supercells track across southern Tennessee and Northern Alabama. Figure 4.1 shows lightning and radar statistics of one of these storms that produced a large swath of damaging wind and hail reports as well as an EF-0 tornado in Skyline, Alabama at 2240 UTC. The first thing to note is the large differences in the magnitudes of GLM flashes versus LMA flashes. For most of the period LMA saw 2 to 3 times as many flashes as the GLM. At its peak, LMA saw 118 flashes per minute while the GLM peaked at 42 flashes per minute. Along with these differences in magnitudes there are periodic opposites in trends as well. Most notably, in the period between 2100 UTC and 2200 UTC there is a local minimum in GLM flashes while the storm is its most intense according to MESH and VIL. During this same period, LMA fluctuates a lot as well and has a slight decreasing trend, but it is not nearly as large as a drop as seen in the GLM.

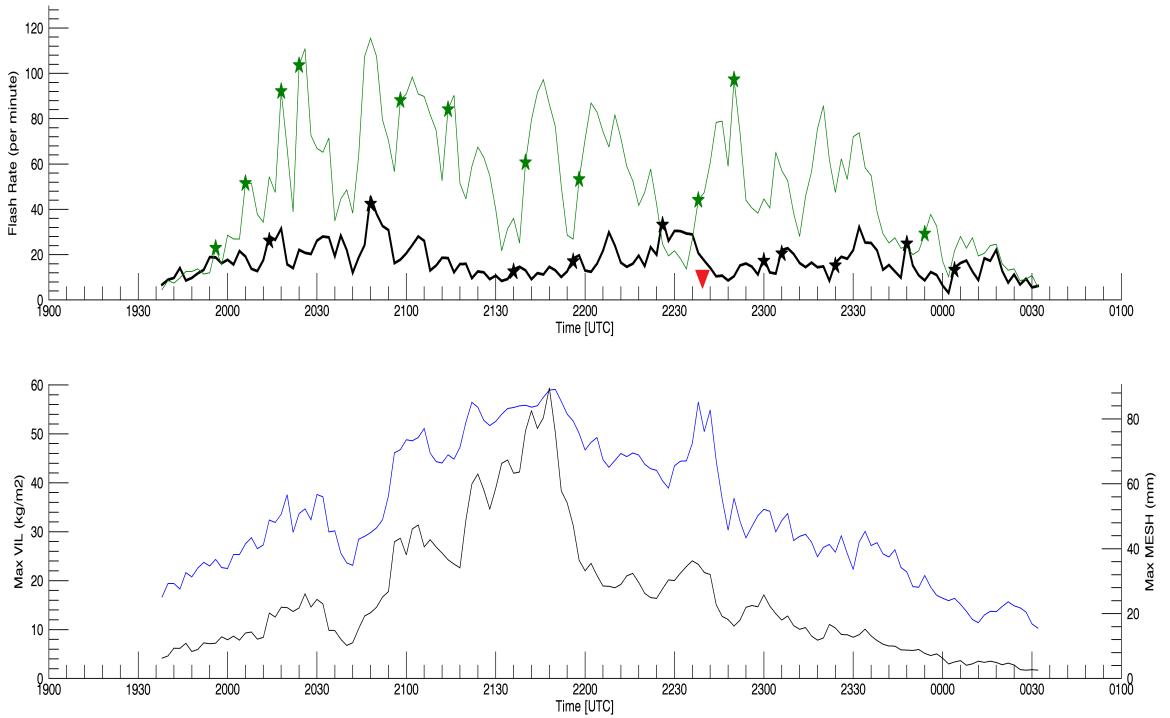


Figure 4.1: Time series plots of lightning and radar data for a severe storm in North Alabama on April 22, 2017. Top: LMA flash rates (green line), LMA lightning jumps (green stars), GLM flash rates (black line), GLM lightning jumps (black stars). Bottom: VIL (blue line) and MESH (black line). The red triangle denotes the time of tornadogenesis at 2240 UTC.

Looking at tornadogenesis in particular, both GLM and LMA saw jumps and then large drop offs in lightning within 15 minutes of the tornado. The GLM saw a jump at 2225 UTC, approximately 15 minutes prior to tornadogenesis, and LMA saw a jump at 2236 UTC, approximately 4 minutes prior to tornadogenesis. This example just prior to tornadogenesis, while extreme, is a part of a trend seen throughout the life cycle of the storm that when there were GLM and LMA jumps temporally close to each other GLM flashes tended to “jump” slightly earlier than LMA flashes.

Looking at the jumps between GLM and LMA flashes overall there does not seem to be much of a correlation. There were four jumps near the middle of the storm's life cycle that lined up within 15 minutes of each other between both the GLM and LMA, but the beginning and the end of the storm seem to have opposite trends in GLM and LMA jumps. At the beginning of the storm there were many more jumps in the LMA (5) than the GLM (2). Most of these LMA jumps coincided well with increases in storm intensity measured by MESH and VIL. At the end of the storm's life cycle there were many more jumps in the GLM (5) versus the LMA (1). All of these GLM jumps occurred while the storm was decreasing in intensity via MESH and VIL which suggests that the GLM is jumping at a relatively high rate when the storm is weakening. This suggests there is some error in the LJA on the GLM here driven by either flash measurements inconsistencies or algorithmic errors. For example, the conceptual model may be wrong with the LJA on the GLM as the conceptual model states that lightning jumps should be controlled by increases in updraft intensity not decreases. This failing of the conceptual model may be due to inaccuracies in the GLM flash data as there are still large differences in LMA vs GLM flash rates and trends at this time. Or there could be jump algorithm error such as how jumps are calculated in the GLM, how cells are being tracked, or how lightning is associated with tracked features.

4.1.1.2 Central Oklahoma May 17, 2017 (Severe)

May 17, 2017 started with multiple tornadic supercells forming over the Panhandle of Texas and Western Oklahoma near sunset. These storms quickly merged

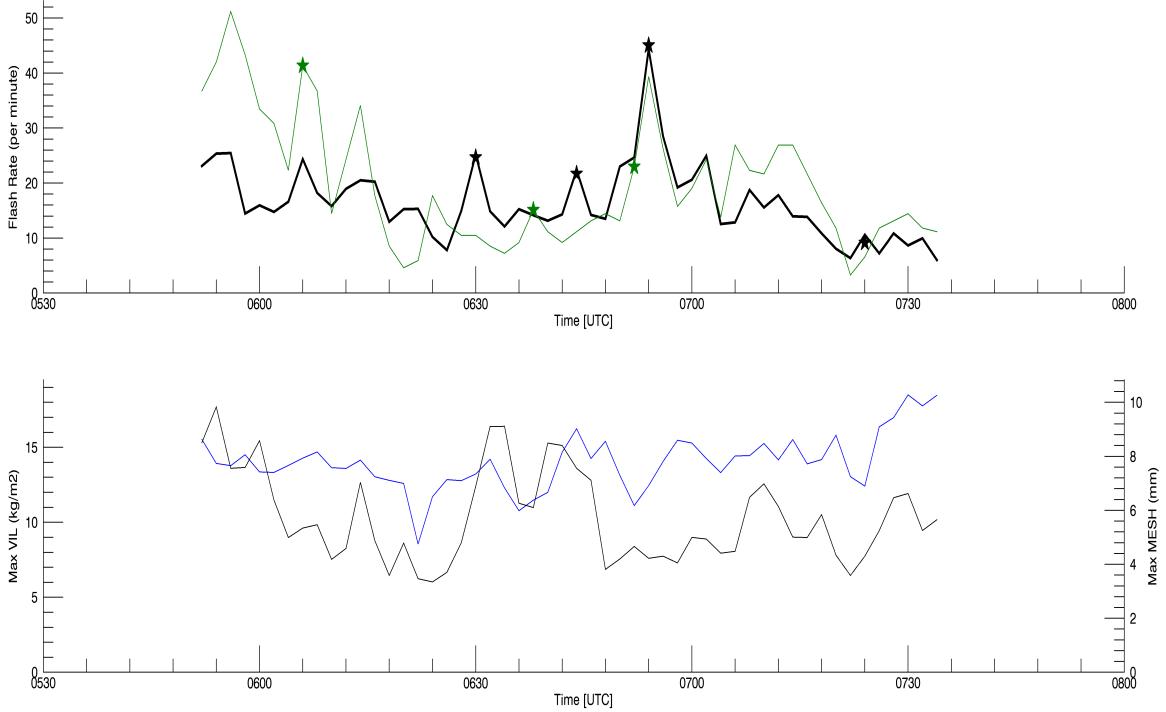


Figure 4.2: The same as Figure 4.1 except for a severe MCS in Central Oklahoma on May 17, 2017.

in multiple severe and non-severe MCS's that tracked across the state of Oklahoma. The size of this tracked feature was roughly two times larger than the aforementioned North Alabama supercell on average (500 km^2 versus 250 km^2). Figure 4.2 shows lightning and radar statistics of one of the severe MCS's that produced multiple severe wind reports through Central Oklahoma. In this case we see a notable difference between this storm and the severe storm in section 4.1.1.1 in that there is much better agreement in the magnitudes of GLM and LMA flashes. Compared to the aforementioned storm, this storm has lower flash rates in both the GLM and the LMA, and weaker MESH (peak of 10 mm versus 80 mm) and VIL (peak of 18 kg/m^2 versus

60 kg/m^2) values indicating this is a weaker storm overall. Despite this much better agreement in the magnitudes, there are still fairly different trends between the two.

As seen in the previous section, the jumps in the middle of the storm's life cycle line up much better between the GLM and the LMA than at the beginning or end of the storm. In the middle of the storm there are 3 GLM jumps compared to 2 in the LMA. All of these jumps coincide with increasing MESH and VIL values. At the beginning of the storm there is an LMA jump that occurs just before an increase in MESH with no real change in VIL and no coinciding GLM jump. At the end of this storm there is one final jump in GLM that occurs just prior to increases in both VIL and MESH with no coinciding LMA jump. This final jump is subjectively fairly weak as there does not appear to be a relatively large increase in flash rate, it is just a small peak during an otherwise decaying trend. This likely points towards issues with the GLM jump system itself in that the two-sigma algorithm is detecting a jump that subjectively does not appear to be one.

Overall this storm had much better agreement in magnitudes of flashes from the differing lightning detection systems, but there are still large differences in the timing of jumps. It is worth noting that this storm did not show a time period of increased jumps in the GLM coinciding with the weakening of the storm, but the tracking of this storm ends well before the storm itself dissipated due to a constraint of usable GLM data for the day, so that behavior may have just not been recorded.

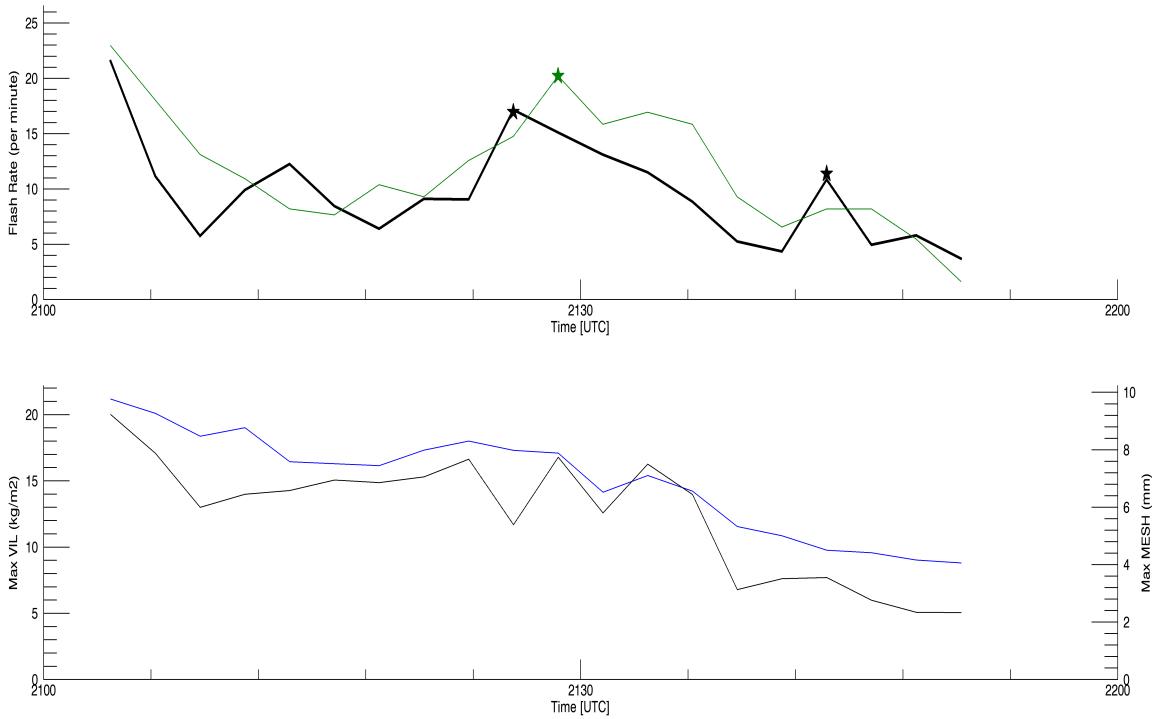


Figure 4.3: The same as Figure 4.1 except for a non-severe single cell in North Alabama on April 22, 2017.

4.1.1.3 North Alabama April 22, 2017 (Non-Severe)

Figure 4.3 shows another storm that occurred in North Alabama on April 22, 2017. This storm was a relatively short-lived non-severe single cell storm. In this storm there is fairly good agreement in the magnitudes of LMA and GLM flash rates. The trends also line up fairly well between the two except for a few minor differences here and there. Despite this storm producing no severe weather and looking relatively weak compared to the severe supercells on this day (Storm ID 1 and 2) given its VIL (22 kg/m^2) and MESH (9 mm) values there were two jumps using GLM flashes and one jump using LMA flashes. The one LMA jump was within 10 minutes of one of

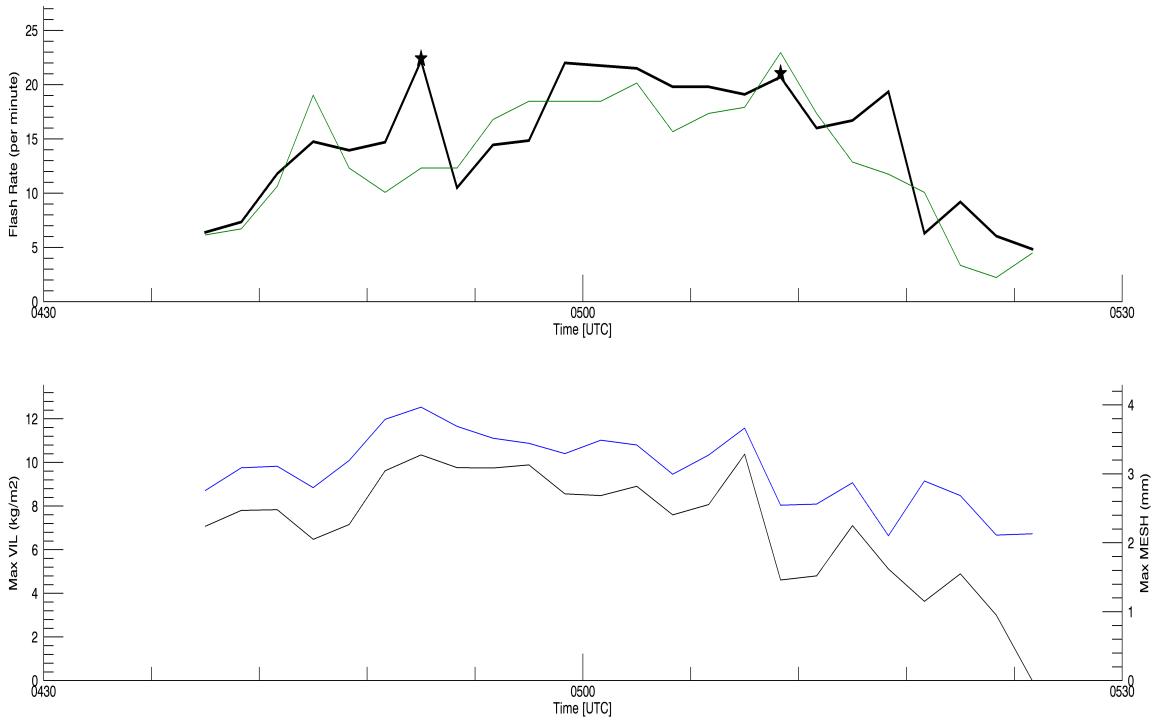


Figure 4.4: The same as Figure 4.1 except for a non-severe MCS in Central Oklahoma on May 17, 2017.

the GLM jumps and did see a slight increase in MESH of 3 mm associated with it despite the lack of any reported severe weather.

Again, an important finding is the other GLM jump occurs as the storm is decreasing in strength. Not only is the storm decreasing in strength, it also has very low VIL (9 kg/m^2) and MESH (2 mm) values that show the storm is fairly weak at this time. Up to this point all three of these selected case studies have shown GLM jumps that have no corresponding increase in storm intensity.

4.1.1.4 Central Oklahoma May 17, 2017 (Non-Severe)

Figure 4.4 shows another one of the MCS's that tracked across Oklahoma on May 17, 2017. This particular MCS did not produce any severe weather reports during the entirety of its tracked lifetime. There does seem to be fairly good agreement between the magnitudes of GLM and LMA flashes. This agreement was noted in most of the relatively weaker storms looked at in Table 3.1 while the stronger storms had much larger differences as in the Skyline, AL storm show in Section 4.1.1.1. Despite showing good agreement in flash magnitudes the two are fairly different trend-wise.

This storm had zero jumps using LMA flashes while there were two jumps using GLM flashes. One of these jumps corresponded with no change in MESH and VIL values while the other occurred while there was decreasing MESH (3 mm to 1 mm) and VIL (12 kg/m^2 to 8 kg/m^2). There were no reports of severe weather with this storm and the MESH and VIL values indicated this was a relatively weak storm compared to the severe storms sampled in this portion of the study. This storm adds to the evidence of there being jumps in GLM flashes that are not seen in LMA flashes that correspond with no change or a weakening of storm intensity via radar intensity metrics. Also, as in a jump seen in Section 4.1.1.2, this again may be a problem of the GLM jump system and two-sigma algorithm as while this is objectively considered a jump, subjectively it is hard to justify. A speculative discussion of these jumps that do not visually appear to be jumps will be included in the discussion Section 5.1.4 of this thesis.

Table 4.1: Pearson correlations between the differing lightning systems and radar intensity metrics for each storm and the average Pearson correlations for all seven storms combined.

Storm:	GLM vs. LMA	GLM vs VIL	GLM vs MESH	LMA vs VIL	LMA vs MESH
N. Alabama (Skyline)	0.351	0.127	-0.019	0.507	0.421
N. Alabama (Jones Chapel)	0.569	0.514	0.440	0.643	0.642
N. Alabama (Non-Severe)	0.773	0.617	0.650	0.701	0.760
Central Colorado (Denver)	0.360	0.343	0.178	0.565	0.591
Central Oklahoma (Severe)	0.563	-0.211	0.093	0.090	0.266
Central Oklahoma (Non-Severe)	0.801	0.412	0.517	0.384	0.457
Texas Panhandle	0.439	-0.136	-0.286	0.501	0.351
Averaged Correlations	0.551	0.238	0.225	0.484	0.498

4.1.2 Quantifying Differences in LMA versus GLM Trends and Radar Intensity Metrics

From the specific case studies a few things are evident in the differences between LMA and GLM flashes and their jumps. One being that even when magnitudes of flashes were similar between the two systems there were still obvious inconsistencies in the trends between the two. The second was that there tended to be a lot more jumps in GLM flashes than in LMA flashes and that the GLM jumps had no increases or even decreases in MESH and VIL values. In this section there is an attempt to objectively quantify these differences.

To quantify differences in trends Pearson correlations are taken for various metrics for each of the seven storms and an averaged Pearson correlation for all the storms is also taken. The metrics correlated are: GLM flashes versus LMA

flashes, GLM flashes versus VIL, GLM flashes versus MESH, LMA flashes versus VIL, and LMA flashes versus MESH. This is done in order to determine how the trends line up between the two systems together, and how each system lines up with independently with MESH and VIL values.

Table 4.1 shows the correlations for all of the aforementioned metrics for each storm and an average for all storms. There was a relatively large dynamic range of correlations when comparing the two lightning systems to each other ranging from 0.351 to 0.801 with an average correlation of 0.551. For two systems said to measure total lightning, these correlations are low, especially when considering GLM was theorized to have a DE of roughly 90% when compared to LMA. The storm that had the least correlation between GLM and LMA flashes was the Skyline, Alabama storm on April 22, 2017 that was highlighted in Section 4.1.1.1. Interestingly this storm had one of the higher correlations between LMA and radar metrics and one of the lower correlations between GLM and radar metrics. This observation goes against the conceptual model of the LJA which assumes that there is a high correlation between lightning production and radar-inferred storm intensity. This could be caused by a combination of aforementioned errors such as GLM's measurement of flash rates or GLM lightning jump algorithm error.

Looking at each system versus the radar intensity metrics we see a similar story. In all but one storm, the non-severe MCS in Oklahoma, LMA flashes were correlated better with radar intensity metrics than GLM flashes. In most of the cases these differences were quite large. In three different storms there was actually a negative correlation between GLM and either one or both of the radar intensity metrics with

the worst correlations being the Texas Panhandle case. Overall, the weaker storms as noted by radar-inferred storm intensity metrics had better correlations between lightning and the radar metrics than the stronger storms. Looking at the averages for all the storms GLM flashes had correlation values of 0.238 with VIL and 0.225 with MESH compared to LMA flashes having a correlation of 0.484 with VIL and 0.498 with MESH.

These results suggest that LMA flashes fit better with the conceptual model of the LJA than GLM flashes in that there is an increased relation between LMA flash rate and storm intensity over GLM flash rate and storm intensity. The failing of the LJA conceptual model in the GLM could be due to one or both of the following: GLM flash measurement errors and GLM LJA system errors. These results are consistent with what was shown in the case studies presented in Section 4.1.1 such that there tended to be more jumps using GLM flashes that corresponded with little increase or actual decreases in storm intensity than using LMA jumps.

4.1.3 LMA Flash Size and Height Properties

Given previous results that there can exist large differences in LMA and GLM flash rate trends and magnitudes it is important to try and understand why some of these differences are occurring. Especially considering the supplementary result that there are relatively low correlations found between GLM flashes and storm intensity based on radar MESH and VIL metrics. It was theorized from the beginning that the GLM would have lower DE's with small flashes due to its coarse spatial resolution and low altitude flashes due to optical extinction. This section will look at flash size

and height differences calculated via LMA sources compared to differences in GLM versus LMA flashes to attempt to diagnose how much of a role these physical flash properties play in the relatively large GLM and LMA differences found in the previous two subsections. For four of the cases with large GLM and LMA flash rate differences average flash size and height at each time step in the storm will be calculated and compared to the flash rate differences. The data is then put through a filter to reduce the high frequency noise to better assess overall patterns both quantitatively and qualitatively. The filter used is a 3rd order Butterworth low-pass filter with a sampling rate of $0.5 \text{ (min}^{-1}\text{)}$ and a cutoff frequency of $0.05 \text{ (min}^{-1}\text{)}$ (Butterworth 1930). Pearson correlations will be taken as an attempt to quantify how much of a factor flash size and height may play in the flash rate differences.

Figure 4.5 shows the plots of the LMA and GLM flash differences along with storm averaged flash size calculated via LMA. Considering flash size versus the flash rate differences there is definitely a noticeable trend in the two Alabama storms and the Texas storms. It appears that the largest differences in GLM and LMA flash rates occur when the storm is undergoing its smallest sized flashes. As flash size increases in this storm the flash rate differences tend to decrease. For the Central Colorado storm, the trend is still evident, but the changes are more subtle and not as easy to see as in the other three storms. It is again worth mentioning that the Colorado storm is the only one that uses decimated LMA data and that may be playing a role in the results found in this analysis. Throughout most of this storm's life cycle the flash sizes were small and the difference between the GLM and LMA flash rates were large.

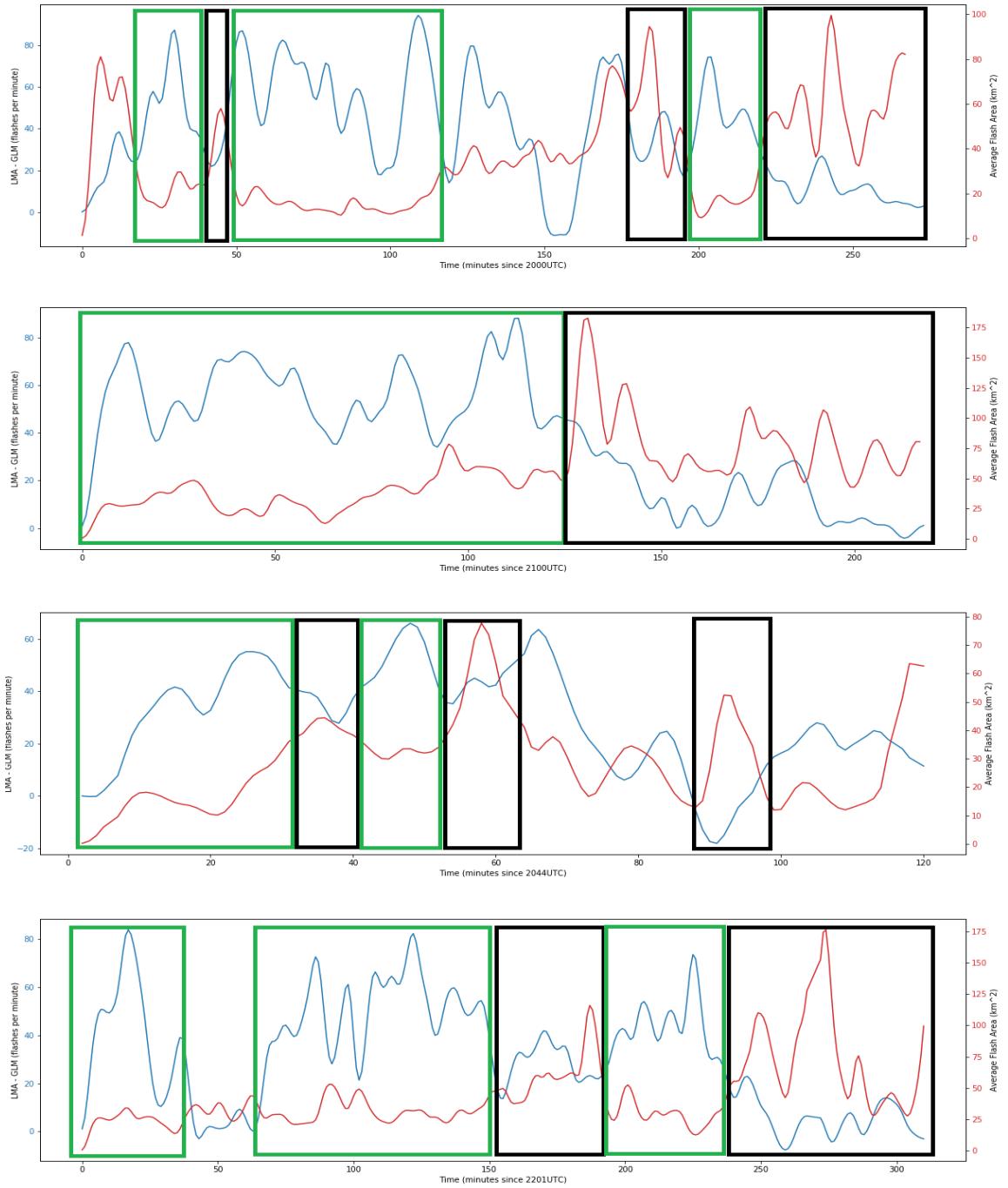


Figure 4.5: Filtered LMA flash and GLM flash differences (blue) compared with flash sizes calculated from LMA (red) for four storms. The four storm locations and their storm id numbers from Table 3.1 from top to bottom: Northern Alabama on 22 April 2017 (ID: 1), Northern Alabama on 22 April 2017 (ID: 2), Central Colorado on 8 May 2017 (ID: 4), and Texas Panhandle on 15 May 2018 (ID: 7). Green (black) boxes highlight areas where there was a decrease (increase) in flash size associated with a increase (decrease) in the LMA GLM difference.

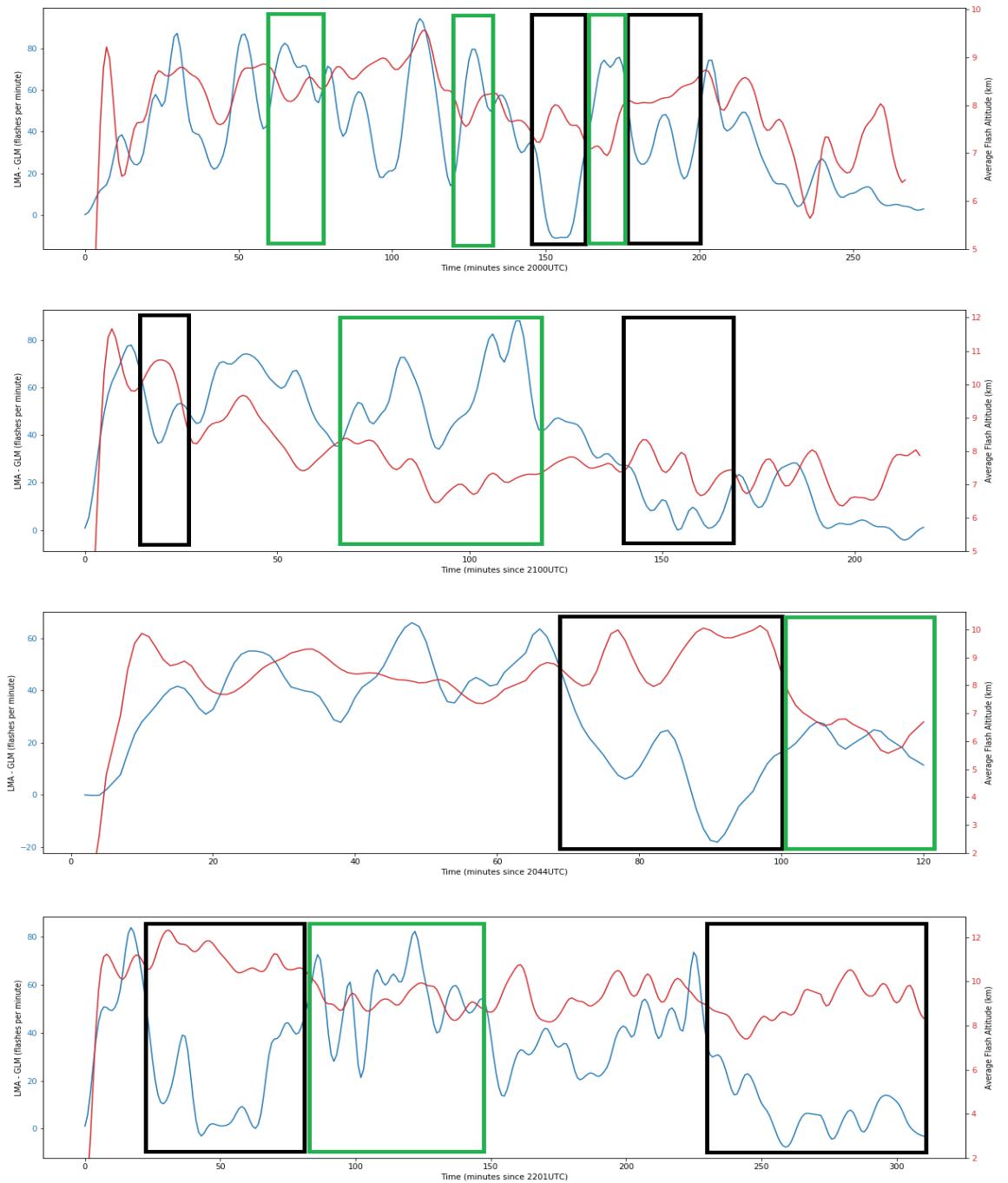


Figure 4.6: The same as Figure 4.5 except now LMA calculated flash altitudes in red. Green (black) boxes highlight areas where there was an decrease (increase) in flash altitude associated with a increase (decrease) in the LMA GLM difference.

Table 4.2: Pearson correlations calculated between differences in LMA versus GLM flashes and LMA calculated flash size and altitudes for four chosen storms using both the raw and filtered results. The storm locations and ID's are as in Table 3.1

Storm	Filtered Lightning Difference vs Flash Size	Filtered Lightning Difference vs Flash Altitude	Raw Lightning Difference vs Flash Size	Raw Lightning Difference vs Flash Altitude
North Alabama (ID: 1)	-0.43	0.47	-0.39	0.38
North Alabama (ID: 2)	-0.42	0.42	-0.31	0.35
Central Colorado (ID: 4)	-0.20	0.02	0.01	-0.15
Texas Panhandle (ID: 7)	-0.44	0.04	-0.37	-0.01
Averages	-0.37	0.24	-0.27	0.14

Figure 4.6 shows the same plots of flash rate differences except now compared to storm averaged flash altitudes calculated via LMA. There is not as clear as of a trend in comparing flash altitude to flash rate differences as there was with flash size. However, there are a few time frames in each storm that show a trend of decreased difference in LMA and GLM flash rates with increased flash height and vice versa.

Given the mixed results from the figures, Pearson correlations were also taken to objectively quantify these trends and are shown in Table 4.2. There are negative raw correlations between flash size and flash rate differences ranging from -0.31 to -0.39 in the Alabama and Texas storms, while there was a near-zero correlation in the Colorado storm. The near-zero correlation in the Colorado case was due to small flashes and high flash rate differences for a majority of the storm. This is a weakness in using Pearson correlations for the raw results in this analysis as a consistent trend in both samples as in the Colorado storm is not a physically null result although it can lead to near-zero correlation coefficients. For the filtered data there were more negative correlations between flash size and flash rate differences ranging between

-0.20 and -0.43 for all of the storms. For flash height there are positive correlations in the raw data for the Alabama storms at 0.38 and 0.35 while there is a slight negative correlation in the Colorado storm at -0.15 and a near-zero correlation for the Texas storm. For the filtered data and flash altitudes the Alabama storms were more highly positive still at 0.47 and 0.42 while the Colorado and Texas storms were near-zero at 0.02 and 0.04 respectively. Averaging all of these there is a negative correlation between flash size and flash rate differences in the filtered (raw) data at -0.37 (-0.27) and a slight positive correlation between the flash altitude and flash rate differences in the filtered (raw) data at 0.24 (0.14).

These results suggest that the GLM is detecting less of the smaller flashes that can at least partially explain some of the large differences in the magnitudes and trends of GLM flashes versus LMA flashes, especially in these more intense and severe storms. Flash altitude does not seem to have much of a contribution to flash rate differences. The slight positive average Pearson correlation value in flash altitudes was mainly dominated by the two Alabama storms. In these two storms there were other complex flash rate changes and flash size changes going on at the same time which could explain why they were more of an outlier.

4.1.4 Summary

In comparing the GLM and LMA flash rates and trends for a few select storms there were some similarities, but a lot of differences between the two systems which led to differences in the LJA between the two lightning detection systems. One of the similarities was there were a total of 13 jumps that lined up fairly well between the

GLM and LMAs. Another similarity was that for weaker storms via radar intensity metrics with lower flash rates there was good overall agreement in flash rates and even in gross trends as noted by the higher Pearson correlation coefficients between the GLM and LMA. Despite the better agreement in gross trends, there were still periodic differences that led to jump timing differences.

Outside of the few similarities there were quite a few differences discovered as well. First was for fairly intense storms there were not only disagreements in flash trends between the two systems, but large differences in flash rates as well. These differences made it such that there were 25 jumps that were not within 10 minutes of each other versus 13 jumps that were. Another difference was the lack of jumps in the GLM as the storm initially intensifies. There were a few different cases seen that the LMA began jumping earlier and more often at the beginning of the intensification of the storm that captured the initial severe weather that the GLM did not. Another was the GLM continued to jump, sometimes at a high rate, as the storm was weakening. There were 11 jumps in the GLM during periods of storm decay as noted by radar intensity metrics versus only 2 in the LMA. There were also 4 jumps in the GLM for weak non-severe convection versus only 1 in the LMA. Both of these results point towards a potential issue of false alarms with the LJA on the GLM. Lastly, it appears that the GLM is fairly sensitive to flash size as there was a correlation, average filtered of -0.37, between a decreasing average flash size and an increasing difference in GLM versus LMA flash rates. However, GLM was shown to not be very sensitive to average flash altitude, with an average filtered correlation of 0.14, in this small sample of storms.

With varying similarities and differences found in this small scale inter-comparison study it is important to understand how the LJA works on the GLM on a larger scale. There are times where at least the trends and jumps between the two systems line up well, but there are more times where there is little to no agreement. An important finding was the GLM jumps seen as the storms weakened and seen in non-severe weak convection. A larger sample study alongside this small scale inter-comparison study will help in diagnosing the biggest problems in the LJA with the GLM by yielding better comparisons to previous studies while also seeing which of the aforementioned similarities and differences stand out. This large sample study will be discussed in the next section.

4.2 Large Sample Size Study

4.2.1 Results Using Original LJA Configuration

The large sample sized study consists of all storms within the domain shown in Figure 3.1 using reprocessed GLM data from ten of the eleven days and times during the GOES-R Post Launch Test Field Campaign which is shown in Table 4.3. The first analysis in the large sample study is running the storm report verification as described in Section 3.6.1 on all the cases at the two-sigma and 10 flash per minute minimum thresholds to see how they compare to previous studies.

The results of this first run through of the large sample are as follows: 930 unique storms, 273 hits, 129 misses, 1265 false alarms, 67.9% POD, 82.2% FAR, and a 0.16 CSI score.

Table 4.3: Dates and Times (UTC) of the reprocessed GLM data used for the large sample study

Date	Times (UTC)
16 April 2017	0631 - 1417
18 April 2017	1657 - 2308
20 April 2017	2139 - 0441
22 April 2017	1928 - 0120
27 April 2017	0627 - 1143
29 April 2017	0300 - 1054
08 May 2017	1857 - 0358
12 May 2017	1308 - 2050
14 May 2017	1119 - 1909
17 May 2017	0200 - 0951

Compared to the Schultz et al. (2009) and Schultz et al. (2011) studies the POD is 10%-20% lower while the FAR is roughly 50% higher. This indicates that either the LJA while run on the GLM tends to be a lot more “jumpy” compared to the initial studies or that there is an issue of under-reporting given the initial studies used a subjectively selected set of storms that were probably better observed with reports. Compared to the Schultz et al. (2016b) study that also employed the automated tracking method along with a more objective sample by using all the suitable storms in the chosen domain, the results are closer and are a better comparison due to closer methodologies. Between these two studies there is only a 2% decrease in POD but still a modest 20% increase in FAR.

The biggest takeaway from the initial results is that there are many more false alarms than seen in any other of the prior LJA studies. There were roughly six times as many false alarms than there were hits. Comparing this with the relatively low

CSI scores it seems the jump was more random than an accurate forecaster (e.g., CSI of 0.16 is only slightly better than no skill at all or CSI of 0.00). This high false alarm rate is consistent with what was seen in the small sample study where there were many jumps with no obvious increases in storm intensity via radar metrics. Given the previous, it is obvious that this significantly increased false alarm rate is an important initial finding and will be looked at in-depth over the next few sections. There are a few different ways this study tries to diagnose the cause behind the high false alarm rate: spatial sensitivity testing to see if any particular area has drastically higher/lower false alarm rates, threshold sensitivity testing to see if there is a different combination of sigma and minimum flash per minute thresholds that works better with the GLM, and radar statistics around jump times between hits and false alarms to diagnose if there are any obvious storm intensity differences between false alarms and hits.

4.2.2 Sensitivity Testing

4.2.2.1 Spatial Sensitivity

The first batch of sensitivity testing is running the original two-sigma 10 flash per minute thresholds for all of the sub-domains shown in Figure 3.2. The purpose of this spatial sensitivity testing is to determine if there are any areas where the LJA performs better than others, especially in regards to false alarms. The spatial variability of the domains have multiple controls that could effect the LJA in storm morphology, population densities, average storm intensities, overall meteorology, and

Table 4.4: Results of the spatial sensitivity testing, domain numbers are the same as shown in Figure 3.2. NaN's indicate not a number errors due to dividing by 0 (hits and misses both 0).

Domain	Storms	Hits	Misses	False Alarms	POD	FAR	CSI
1	54	12	13	61	48.0%	83.6%	0.14
2	158	48	19	213	71.6%	81.6%	0.17
3	196	43	18	292	70.5%	87.2%	0.12
4	34	0	0	57	NaN	100%	0.00
5	52	1	1	62	50.0%	98.4%	0.02
6	149	48	31	201	60.1%	80.7%	0.17
7	85	41	14	136	74.5%	76.8%	0.21
8	78	56	15	90	78.9%	61.6%	0.35
9	27	21	2	34	91.3%	61.8%	0.37
10	12	0	0	17	NaN	100%	0.00
11	33	4	11	39	26.7%	90.7%	0.07
12	36	0	4	44	0.00%	100%	0.00
13	16	0	1	22	0.00%	100%	0.00

etc. If it is found that the results are spatially sensitive it could give a starting point for future studies in trying to diagnose the specific roles of aforementioned potential errors and the high false alarm rates.

The spatial sensitivity shows a dynamic range of false alarm rates from 61.6% to 100% and CSI scores from 0.00 to 0.37 as seen in Table 4.4. Breaking down the initial results spatially, the false alarm rates are still much higher than the Schultz et al. (2009) and Schultz et al. (2011) studies, however there were a couple locations, domains 8 and 9, that were more in line with the Schultz et al. (2016b) study. Unfortunately the lowered false alarm rate in domain 9 could be explained by there being a small sample and its sample being dominated by fairly intense storms on 22 April 2017. Domain 8, which encompasses portions of Texas, Oklahoma, and Arkansas as

well as all of Louisiana, Mississippi, and Alabama, does have a sizable sample size and the results may be physically plausible. It is also interesting to note the Domain 8 includes, but is not limited to, the domain used in Schultz et al. (2016b).

Outside of the two aforementioned domains, false alarm rates are fairly consistent in the 80% to 100% range as seen in the original testing for all domains. Despite one or two domains performing better and more like previous results, the other eleven show that the high false alarms are more of a systematic issue that is not controlled by spatial bounds. This result points toward the high false alarm rates not being affected much by spatial controls which may indicate that the conceptual model of the LJA on the GLM is failing due to errors in GLM flash rate measurements and/or GLM LJA system errors.

4.2.2.2 Threshold Sensitivity

Sensitivity tests are also run by varying the sigma and minimum flash per minute thresholds. The original two-sigma 10 flash per minute thresholds were chosen after sensitivity testing was done on them using LMA's. Since the GLM is a different instrument giving us potentially different magnitudes and trends of lightning at times as shown in Section 4.1, it is important to rerun this sensitivity testing to see if there is a better combination that fits the GLM better.

In this testing sigma levels are varying between 1.0 and 4.0 by increments of 0.1, and the minimum flash per minute thresholds are varied from 0 to 25 by increments of 1. At each pair of thresholds POD, FAR, and CSI are all calculated to

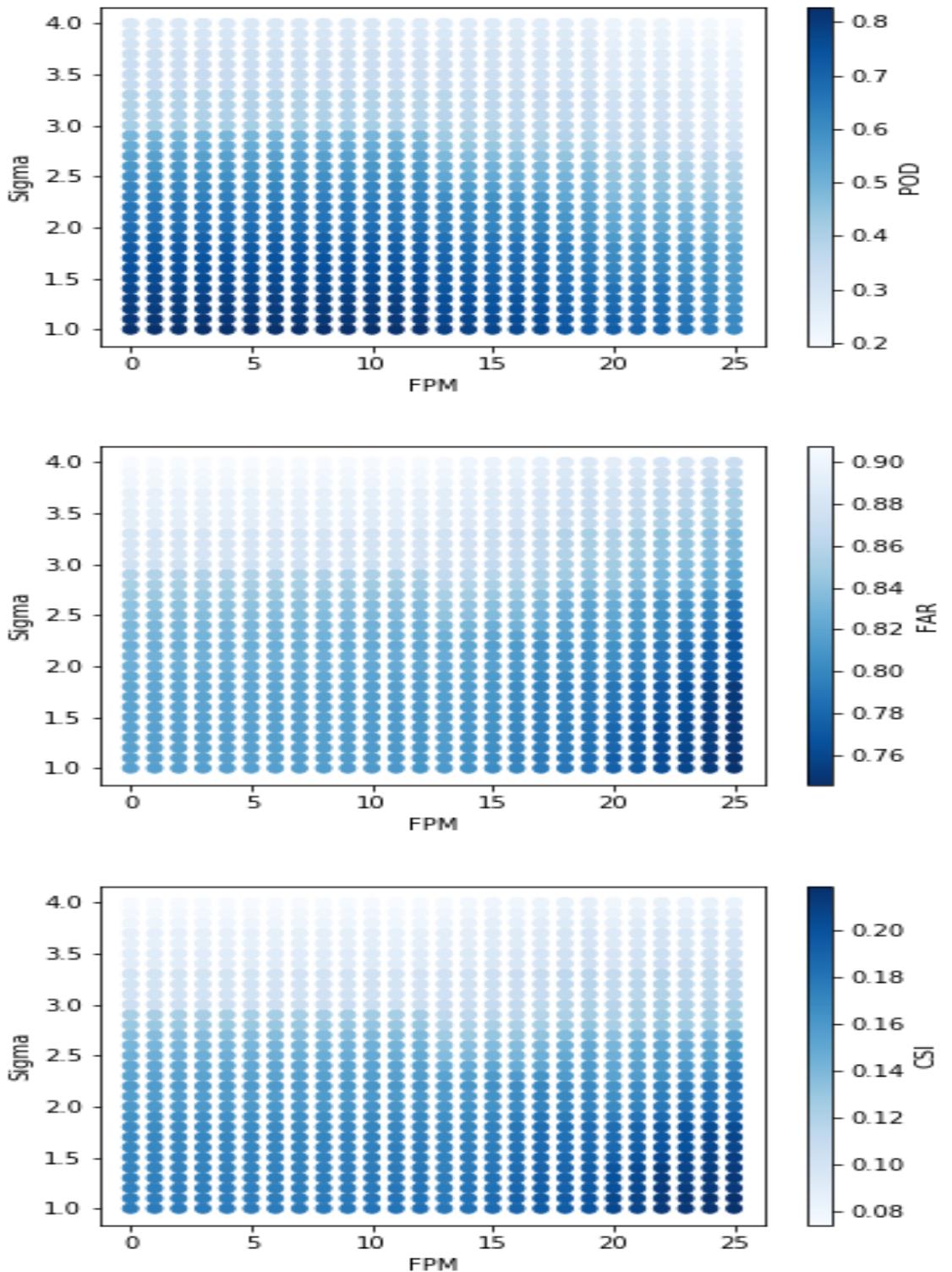


Figure 4.7: Results of the threshold sensitivity testing on the large sample size study showing POD (top), FAR (middle), and CSI (bottom). Darker blues indicated the higher performing values.

try and determine what combination works the best. Figure 4.7 shows the results of this sensitivity testing.

The POD has the highest dynamic range and varies between 19.4% and 82.8%. The best PODs occur when the thresholds on the LJA are the most relaxed with the lowest sigma and flash per minute thresholds. The worst PODs occur with the most constraining LJA thresholds at the higher FPM and sigma levels. This makes intuitive sense as a more relaxed LJA will jump much more leading to a higher probability of a jump occurring around a severe weather report.

The FAR has a much more restrictive dynamic range and varies between 74.6% and 90.8%. It is worth noting here that the smaller dynamic range of the FARs compared to the large dynamic range of the PODs is consistent with previous sensitivity testing done in Schultz et al. (2016b). The FARs varies more in the minimum flash per minute thresholds than in the sigma thresholds. The FARs are minimized at low sigmas and high flash per minute thresholds and maximized at high sigmas and low flash per minute thresholds. The lower FARs at higher flash per minute thresholds make intuitive sense as a more restrictive algorithm should lead to less jumps and only jumps in relatively intense storms are considered. However, the decreasing FAR with decreasing sigma seems counter-intuitive given the prior logic seen in this sensitivity testing thus far. Remembering that FAR is weighted by hits (c.f. Eqn. 3.2) and that there were roughly four times the amount of false alarms than hits at the original configuration, this result is likely caused by there being a higher increase in the percentage of hits over the increase in the percentage of false alarms caused by

a less restrictive LJA. This hypothesis is backed up by the increased POD at lower sigmas.

The CSI scores trended much like the FARs where it was maximized at higher FPM thresholds and lower sigma thresholds while being minimized at the opposite end. The CSI scores ranged from 0.07 to 0.22. The best CSI score of 0.22 occurred at a sigma level of 1.1 and a minimum flash per minute threshold of 25. At these thresholds there was a POD of 61.2% and a FAR of 74.5%. Even at these optimized values, the FAR is still roughly 10% higher than any of the prior studies. Overall these results suggest that there are better combinations of sigma and minimum flash per minute thresholds for GLM that increase its skill scores, but the LJA is still under-performing on the GLM versus LMAs, especially when it comes to the amount of false alarms.

4.2.3 Lightning Jump MESH and VIL Statistics

To understand the high FARs using the LJA on the GLM, statistics are derived from the radar intensity metrics because of the previously documented relationship between lightning and radar. For every jump that occurs within the large sample dataset, values for MESH and VIL are taken for the fifteen minutes centered on the jump. These statistics are split up into hits and false alarms. Normalized histograms, means, and medians are calculated for each of the two groups for MESH and VIL. All zero values of MESH and VIL are assumed to be due to radar dropouts and are not considered in these statistics. This is done as an effort to see if there are any notable differences in radar derived storm intensities between hits and false alarms.

Table 4.5: Mean and median statistics for all MESH and VIL values for hits and false alarms for the fifteen minutes surrounding a jump. Values of 0 are not considered.

	Mean	Median
MESH Hit	24.6 mm	21.9 mm
MESH False	17.1 mm	14.7 mm
VIL Hit	33.4 kg/m ²	32.9 kg/m ²
VIL False	27.7 kg/m ²	25.1 kg/m ²

This could help in the understanding of whether the false alarms are more of a result of under-reporting of severe events or a failing of the conceptual model of the LJA on the GLM, which would be due to GLM flash rate and/or GLM LJA system errors.

Figure 4.8 shows the normalized histograms for VIL and MESH for hits and false alarms while Table 4.5 shows mean and median values for each category. Starting with MESH it is clear to see that while there is some overlap, there is a clear difference between hits and false alarms. The distribution for false alarms peaks at much smaller values of MESH than the hits and falls off much quicker at larger values. The result from the histogram is mirrored in the statistics where the mean and median MESH values for hits are about 7 mm higher than for false alarms. For 78% of the time sampled with false alarms the values were below a MESH value of 25.4 mm which is the threshold for severe hail by the National Weather Service. A T-test yields a p-value of much less than 0.05 ($p << 0.01$) indicating that the means between false alarms and hits are statistically significantly different. The T-test was chosen because despite there being large differences in sample size the variances were similar and the

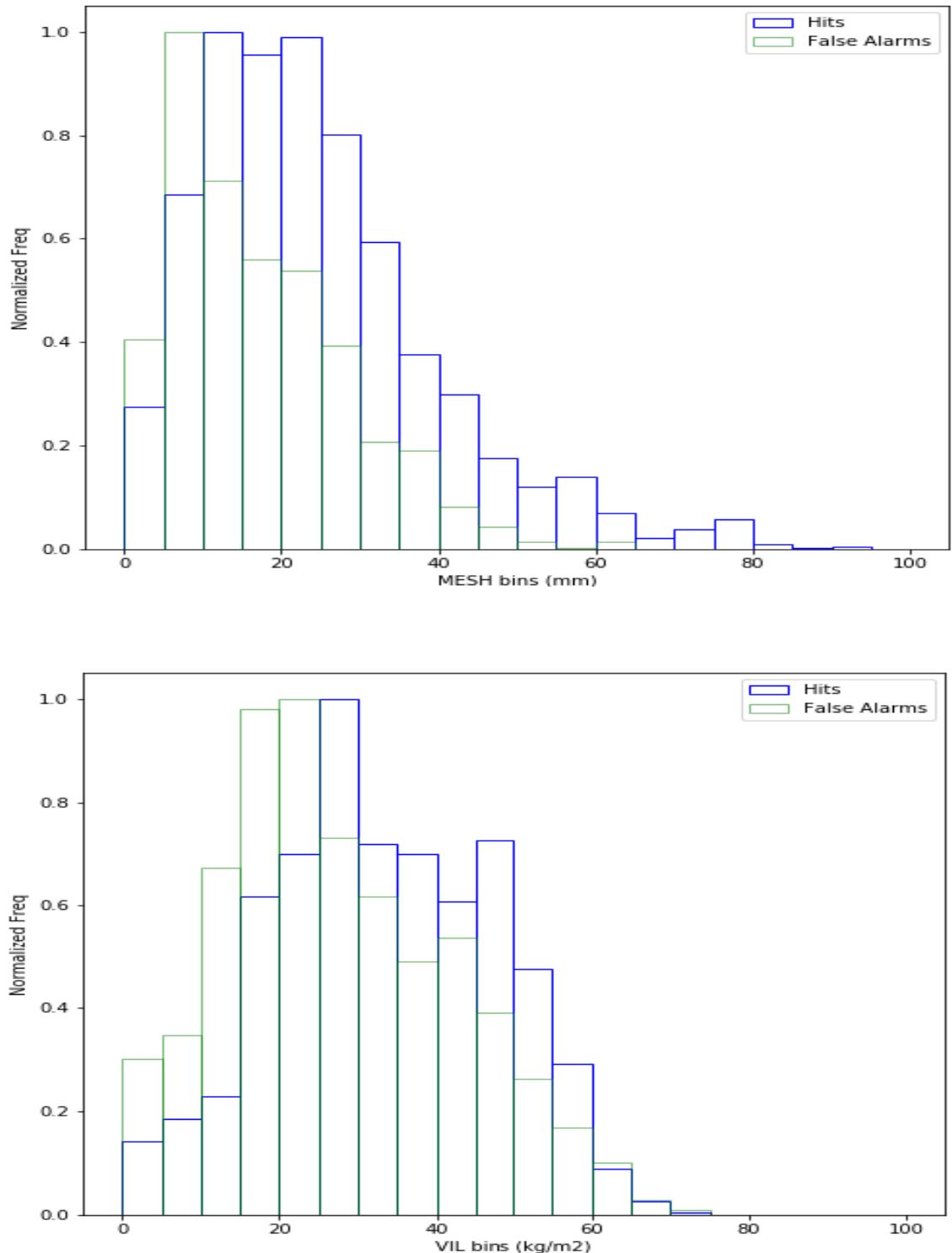


Figure 4.8: Normalized histograms of MESH values (top) and VIL values (bottom) for hits (blue) and false alarms (green) for the fifteen minutes surrounding a jump. Values range from 0 to 100 in bins of 5 and all zero values are thrown out.

Table 4.6: Skill scores for the LJA on the GLM at a two-sigma 10 flash per minute minimum threshold for varying minimum MESH thresholds.

MESH Thresholds	Hits	Misses	False Alarms	POD	FAR	CSI
None	273	129	1265	68%	82%	0.16
10 mm	248	154	655	61%	72%	0.23
15 mm	215	187	402	53%	65%	0.27
20 mm	186	216	235	46%	55%	0.29

distributions were close to normal making the T-test a valid method of measuring significance.

Looking at VIL there are similar results but there is more overlap between the two distributions. False alarms for VIL are still peaking at lower values than for hits but the fall off at higher values is not as dramatic as seen in MESH. There is a difference of 5.7 kg/m² in the mean and 7.8 kg/m² in the median. A T-test is ran on the VIL distributions as well and again yields a p-value of much less than 0.05 ($p << 0.01$) indicating again that the means between false alarms and hits are statistically significantly different.

Given that the peaks of the false alarm distributions are on the weaker side of the MESH and VIL statistics, it suggests that most of the false alarms are physically plausible and less of a result of other non-physical parameters such as under-reporting of severe events. That being said there is still likely error due to under-reporting as noted by the overlap of the distributions at higher MESH and VIL values. It is likely that this issue of under-reporting was present in prior studies and there is not any obvious evidence found here to suggest this study has an increased susceptibility to it.

These results suggest that adding a radar intensity threshold may be beneficial to the operational skill of the LJA on the GLM (Rudlosky and Fuelberg 2013, e.g.,). A drawback of adding a radar threshold however would be needing to be in a radar domain in order to run the jump which cuts down on the amount of spatial area the LJA could be ran. Table 4.6 shows how the original two-sigma 10 flash per minute algorithm runs on the large sample on the GLM with objectively chosen minimum MESH thresholds taken from the normalized histograms. If a traditional jump occurs while the current MESH value is below the chosen minimum threshold it is not counted as a jump. This analysis was only done on MESH because there was a bigger separation of distributions than in VIL which indicates that there would be better success adding a MESH threshold versus a VIL threshold. However, a prior study by Metzger and Nuss (2013) found that out of 73 jumps using LMA that 49 of them had a rise in VIL, 17 of them had a decrease, and 7 had no net change. Therefore there may be some utility in using VIL thresholds as well, but given the results of the distributions in this particular study MESH should perform better.

Going from no threshold to a 10 mm threshold shows the biggest decrease in FAR (10% decrease) while not having a relatively large effect on POD (7% decrease). The raw false alarm count was cut in roughly half, but the FAR did not decrease as much due to the concurrent decrease in hits. The CSI score was also raised by 0.07 up to a value of 0.23. As we increase the MESH value needed for a jump, the false alarms continue to fall, but the hits do as well. Despite a rising CSI through these results the higher thresholds would likely be rejected for most operational uses due to dropping the POD to near 50% and below.

4.2.4 Timing Between First Jump and MESH Thresholds

Given the high false alarm rates and lower skill scores compared to previous studies, this section describes another way to look at the jump to reinforce or dispute the hypothesis that there appears to be physical errors in the LJA on the GLM consisting of one or more of: physical conceptual model errors, GLM flash measurement errors, and/or GLM LJA system error. This analysis will look at the timing of the first jump of the storm compared to three different values of MESH: 25.4 mm (severe hail threshold considered by the NWS), 40.0 mm (significant severe hail threshold considered by the NWS), and the maximum of each storm. Theoretically if the LJA is working properly on the GLM there should be a jump around the time the lowest (25.4 mm) threshold is crossed as that is the first sign of a strong intense storm, a jump prior to the 40.0 mm threshold as the storm has intensified enough to possibly produce significant severe hail and thus should have had at least one jump by this time, and finally there should be a jump prior to the maximum MESH of the storm as it indicates the strongest point in the storm's life cycle. A prior study using LMAs found that for the 25.4 mm threshold the median first jump time was 11 minutes after and for the maximum threshold the median first jump time was 16 minutes before, which fits with the conceptual model laid out here and prior (Schultz et al. 2016a). It is worth noting that since MESH thresholds of severe and significant severe hail sizes are used that it has been found that MESH tends to slightly overestimate the reported hail size on the ground (Brimelow et al. 2002).

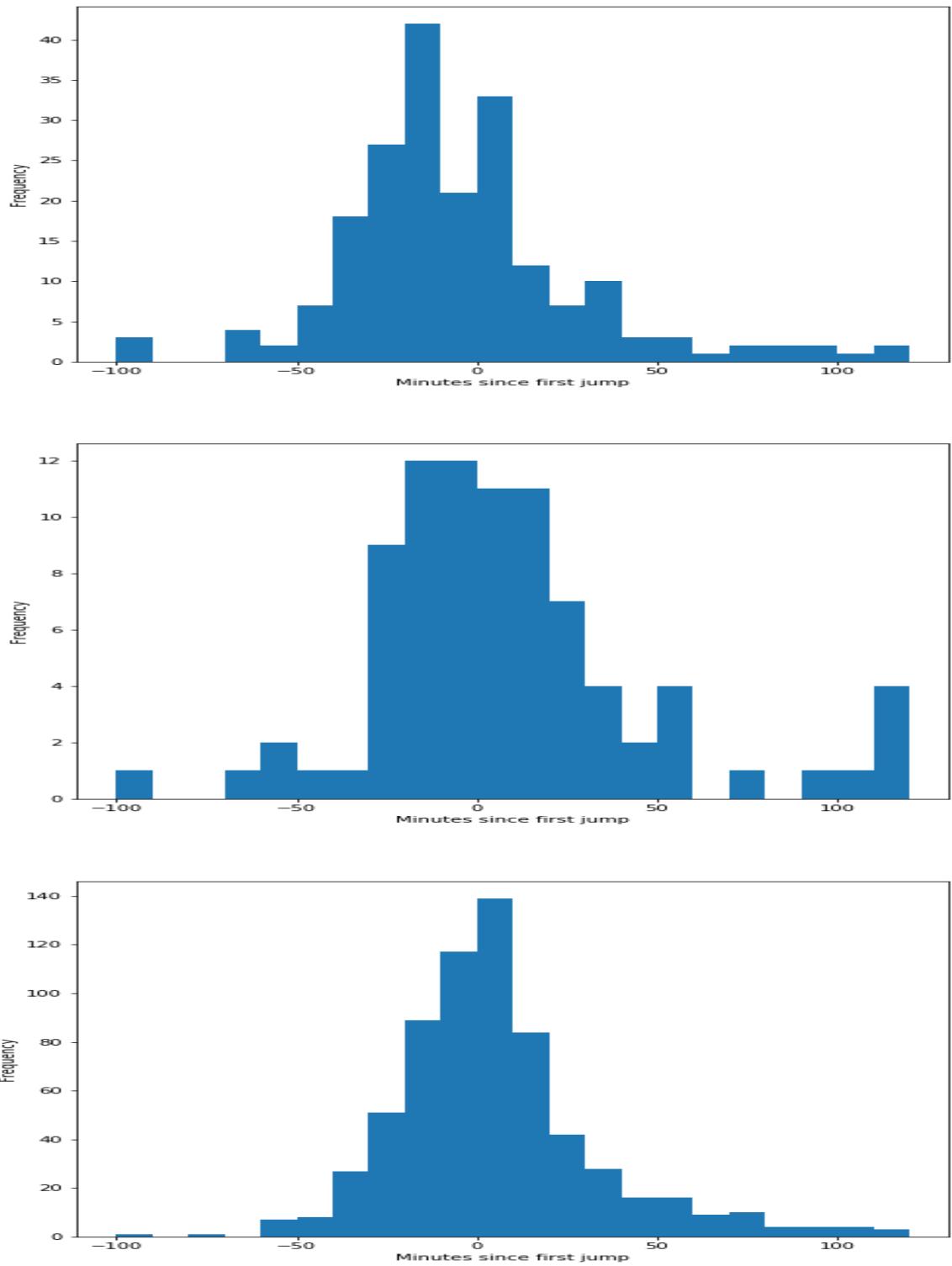


Figure 4.9: Distributions of the difference in timing between given MESH thresholds and the first jump. A positive (negative) time means the first jump occurred before (after) the given thresholds. The thresholds are from top to bottom: 25.4 mm, 40.0 mm, and maximum.

The results of this analysis are given in Figure 4.9. The 25.4 mm MESH value tends to occur prior to the first jump in the GLM, however there is a long tail of jumps that occur prior to this threshold. The mean jump time was 3 minutes after this threshold and the median was 10 minutes after. This lines up fairly well with Schultz et al. (2016a) findings with LMA where it was found the jump median was 11 minutes after this threshold for LMA. This also agrees with the conceptual model that the first jump should be occurring right around this time as it indicates the initial intensification of the storm.

The 40.0 mm MESH value has a large peak of values right around the 0 minute time frame meaning that the 40.0 mm MESH value occurs nearly simultaneously with the first jump of the storm. There are also large tails in both directions with this threshold. The mean jump time was 3 minutes before while the median jump time was 2 minutes before for this threshold. Given the conceptual model of the LJA the first jump would have been expected to occur well prior to this. A 40.0 mm MESH values indicates a fairly significant storm that should have had a jump associated with it by this point more often than not. There are a relatively large amount of first jumps that occur after this value as well. This lines up well with what was shown in the small sample study where there were not as many jumps in the GLM as in the LMA at the beginning of the storm when it was intensifying the most.

Moving onto the maximum MESH of a storm the distribution is a near normal distribution with its peak again near the 0 minute line. There is a longer tail in the positive direction than in the negative for this threshold. This threshold had a mean first jump time of 6 minutes before and a median first jump time of 2 minutes before.

This median jump time of 2 minutes prior is significantly less than the LMA study from Schultz et al. (2016a) that had a median time of 16 minutes at this threshold. This is again reinforcing a prior result where it was found the GLM was not jumping as much during the storm's initial intensification periods when severe weather might be expected after the first pulse in the updraft.

Overall, outside of the 25.4 mm threshold, there was a decrease in mean and median jump times prior to the other two thresholds being met. This result is consistent with prior findings in this study that have shown the lack of jumps in the GLM compared to the LMAs during initial intensification periods. The signs of the means and medians are consistent with prior studies suggesting that there is not a complete failure of the LJA on the GLM, however the vastly different magnitudes of these statistics still suggest that the aforementioned potential errors are leading to a degraded performance when compared to LMAs.

4.2.5 Summary

The large sample study confirmed the issues found in the smaller scale study. With one of the most important findings being related to the number of false alarms. In the small sample study it was found there were an abnormally high amount of jumps in the GLM in non-severe convection and during the decaying portion of more intense storms compared to the LMAs. This larger sample confirmed the jumps that appear to go against the conceptual model to be a larger scale problem as well as there were 1265 false alarms and 82.2% FAR given the original configuration of the LJA. Conversely, the POD was fairly consistent but slightly lower than the

previous studies at 68%, which indicated that the high number of false alarms is the biggest issue with the LJA on the GLM.

Spatial sensitivity testing found that there was not much difference in the results based on where you were within the domain with a small exception of a single sub-domain performing better than the rest. The single sub-domain, Domain 8, had a POD of 78.9%, a FAR of 61.6%, and a CSI of 0.35 with 78 storms sampled. Threshold sensitivity testing found that slightly higher skill scores, up to 0.22, were yielded at lower sigma levels and high minimum flash per minute thresholds. The highest skill score of 0.22 was at a configuration of 1.1 sigma and a minimum of 25 flashes per minute and had a POD of 61.2% and a FAR of 74.5%.

Statistics of MESH and VIL were taken for the fifteen minutes centered on each jump split into hits and false alarms to better understand what was physically going on in the storm when there were false alarm jumps. It was found for both MESH and VIL that false alarms were occurring in weaker storms as compared to hits. Roughly 78% of the times sampled with false alarms were associated with MESH values less than 25.4 mm. This indicates that there were many jumps not associated with intense/severe storms. With this result the large sample was run again putting a minimum threshold on a concurrent MESH value for a jump to occur. Increasing the MESH threshold led to higher skill scores as FAR's were cut down much quicker than POD's but there were still large decreases in POD's that would be less useful in an operational setting.

Finally the timing between the first jump of a storm and objectively chosen MESH values were calculated to reinforce or refute the hypothesis that there appears

to be physical errors in the LJA on the GLM. This analysis found that jumps were occurring before the higher MESH thresholds and after the lowest MESH threshold which does line up with what would be expected conceptually. However, the timing of the jumps before the higher thresholds was a lot closer than what was found in a prior study using LMAs and there were a relatively higher distribution of jumps where the first jump actually occurred after these values. This lined up well with a result from the small sample study where there were less jumps in the GLM during initial intensification.

Chapter 5

DISCUSSION

5.1 Additional Speculative Findings and Potential Errors

During the course of this research the high false alarm rate was the most important finding and became the main focus. Throughout this thesis it was mentioned that overall performance degradation of the GLM LJA versus the LMA LJA was due to two potential sources of error: GLM flash measurement error or GLM LJA system error. Speculative discussion on what exact errors are occurring within those two broader sources of error follows.

The GLM flash measurement error is errors in the GLM LJA caused due to the GLM producing inaccurate flash rates. In the stronger storms errors in GLM flash rate magnitudes were apparent as the GLM was seeing up to 3 times less flashes than the LMA at times. During these strong storms there were also mismatches in trends such that there were periods of decreasing GLM flash rates during increasing LMA flash rates and increasing MESH and VIL. In the weaker storms there was better agreement between the GLM and LMA flash rates and gross trends, but there were still times where there would be spikes, and sometimes jumps, in GLM flash rates not seen in the LMA and without an increase of MESH or VIL. There are also

other GLM data sets that currently exist during the current calibration and validation research that utilize differing filtering and clustering methods that change the flash rate and trend data given similar tracked features (Ringhausen 2018). Finally, there is an issue where the GLM can split relatively large flashes into many small flashes (Bruning 2017; Mach 2017b) which could potentially explain some of the jumps the GLM sees in weak or decaying convection. Flash size and altitudes, as measured by the LMA, also seemed to have an effect on GLM DEs, especially in flash sizes. This may relate to potential errors caused by optical radiance extinction as there will be less light escaping the top of the cloud due to increased scattering (flash altitude) or less overall light originally being emitted (flash size).

The errors in the GLM LJA system have a wide range from the algorithm itself to tracking of features. During the small scale GLM vs LMA comparisons there were a couple of jumps (noted in Sections 4.1.1.2 and 4.1.1.4) that technically met the criteria of a jump, but not the conceptual model of a jump as noted by Williams et al. (1999). These jumps were associated with very minor flash rate increases (as low as a 2 flash per minute increase), but also with little variability in the 12 minutes leading up to it, thus triggering the sigma algorithm. It was also found that there were some false alarms that occurred at times of high MESH and VIL values indicating that under-reporting of severe weather events may have some effect on the performance, but this was also noted to be a potential issue in the prior LJA studies (Schultz et al. 2009, 2011, 2016b). There could also be some issues in how features are identified and the tracking algorithm as a whole. It was noted throughout the small scale comparisons that the tracking algorithm tended to focus more on VIL even in the

higher flash rates. This could compound potential errors in lightning association to features as well as there was lightning not associated nor counted due to falling just outside of the tracked feature. There also existed issues where the tracking algorithm merged storms based on their radar quantities briefly which greatly increased feature sizes and thus increased the flash rate statistics. Finally, there were also issues in setting an arbitrary GLM grid to track and obtain lightning statistics to the point where changing where the grid started actually changed the number of jumps and thus the POD, FAR, and CSI of a single storm.

Overall, there are a lot of potential sources of errors in this study and perhaps there are more than what are mentioned here. It is likely that the degraded performance of the LJA when used on the GLM versus the LMA is a combination of many of these errors rather than a single source of error itself. Throughout this research a few of these aforementioned errors were looked into a bit more in depth and will be discussed further.

5.1.1 Differing GLM Data Sets

During the process of this research it was found through personal contacts and the annual GLM science team meetings that there are currently multiple GLM data sets currently in use. These differing data sets have been used in other calibration and validation efforts to try and maximize the utility and accuracy of the GLM. These data sets feature differing flash clustering methods, noise reduction techniques, and other processes before it reaches the final level two stage. Thus it is important to once again reiterate that this study, outside of two small sample storms that utilizes

the current GLM live feed, utilizes ten of the eleven days from the Lockheed Martin reprocessed data set for the GLM calibration and validation field campaign.

One of these other data sets produced by a research group at the University of Alabama Huntsville (UAH) was briefly analyzed for the Skyline, Alabama case on 22 April 2017 (Storm ID: 1) (Ringhausen 2018). For this case it was found using the same tracked feature that there were roughly 1.5 times as many GLM flashes on average. While it was still lower than then LMA flash rates, they were much closer in magnitude. It was also noted that there were at least some differences in trends between the two data sets as well. Both of these differences together could create differences in how the LJA would run for the same storm.

Given this preliminary finding just looking at one other data set it is likely that any changes in how the GLM data is processed would cause differences in how the LJA runs on the GLM. It is not certain at this time to what extent these differences would be. Therefore it should be noted that this study should only be considered valid for the data set used in it and that any future changes to GLM flash processing would require continued research into the performance of the GLM based LJA.

5.1.2 Storm Tracking

This section focuses on the VILFRD and WDSS-II methods used for the tracking of storms. During the process of creating the small sample storms the tracking was analyzed to make sure it was performing as expected. During this process it was noted that the VILFRD method tended to trend more with radar data than the GLM data even at higher flash rates. A sudden drop in GLM data had a much smaller

impact on the values of VILFRD in the feature than a sudden drop in VIL. This led to times where there was a lot of GLM flashes falling outside the tracked feature that was likely actually associated to that storm because the tracked feature was more focused on the core of the storm via VIL than the extent of the lightning production of the storm. To try and combat this a buffer was added to a few of the small sample storms, 10 km to the East and West and 20 km to the North and South. For supercells this buffer worked well in identifying more lightning that was likely associated to the storm and changing the magnitudes and trends of GLM flash rates. Thus this buffer could also have a potential effect on the skill of the LJA. However, this buffer did not perform as well for linear and multi-cell storms as it led to more lightning from surrounding storms to be associated with the original storm. This would complicate any issues of trying to automate this buffer in a larger sample size study to see the magnitude of the effects it has on the operational skill of the LJA. A potential way of mitigating this issue would be using GLM flash extent density rather than single point flash centroids. These issues of the tracking being more sensitive to radar and all of the lightning likely produced by a storm not being associated with it suggests that further exploration into improved methods of storm feature identification and tracking is needed.

Another finding related to storm tracking is that it was found that where and how the domains were chosen mattered. The first way it mattered has to do with how WDSS-II and Multi-Radar Multi-Sensor (MRMS) works. Within the research build of WDSS-II only nine radars can be used within a single domain. MRMS radar values, including VIL, are then calculated using those nine radars. The same storm

sampled by different combinations of radars based on how the domain was chosen would produce differing exact results of values such as VIL and slightly different spatial properties as well. This in turn effects values of VILFRD which will effect tracked feature sizes and locations. Differing tracked feature sizes and locations would cause differences in both the radar intensity metrics and lightning metrics used in this study which could have an impact on the LJA. The other nuance of domain placement is where exactly the GLM grid is started. As mentioned in the methodology section, for each domain GLM flash centroids are gridded into 8 km by 8 km boxes. If the domain is shifted slightly it will change the values of the individual GLM grid boxes which again will effect both VILFRD and the flash rate magnitudes and trends for a storm which could again have impacts on the LJA. A specific example of this was looked at for the Skyline, Alabama storm on 22 April 2017 (Storm ID: 1). The storm was sampled in differing domains from the large sample and small sample studies. Between the two domains there were slight differences in all of the verification metrics: GLM jump times, number of GLM jumps, hits, misses and false alarms.

During the small sample study it was also noted that from time to time there were relatively high fluctuations in tracked feature areas in a relatively short period of time. Some of these fluctuations were large enough that they were very likely non-physical. An example from one of the North Alabama supercells (Storm ID 2) was that there was a brief (5 minute) near-doubling of feature size due to a merging of two storms based on their VIL values. Rapidly increasing cell areas could cause artificial jumps in lightning purely due to more lightning being encapsulated within the tracked feature rather than actual physical increases in lightning production. This could lead

to an increase in false alarms purely due to issues in tracking. While definitely an issue to investigate, it did appear that this was mainly an isolated issue as only a couple cases of this were noted in the small sample study.

5.1.3 Large Flashes

GLM splitting relatively large flashes into many small flashes has been discussed as a potential problem in the current GLM flash clustering algorithm at the previous GLM science team meetings (Bruning 2017; Mach 2017b) . This has potential implications for the LJA in that a single or multiple large flashes could artificially cause a jump due to flash clustering issues rather than a physical increase in lightning production. Looking at Figure 4.5 from the LMA flash size and heights section, there are plenty of examples of sudden increases in average flash size at a particular time-step which likely points to one or more large flashes occurring at that time. During many of these large increases there is a subsequent drop in the difference between LMA and GLM flash rates. Sometimes this drop in difference is significant enough that it shows that the GLM was seeing more flashes than the LMA at that time. Looking through a handful of these LMA flash size increases with subsequent flash rate difference drops, the flash rate drops are mainly associated with increases in GLM flash rates rather than decreases in LMA flash rates. This is likely related to the previous issue found of the GLM incorrectly splitting a single or multiple large flashes.

A portion of these GLM flash rate increases associated with LMA flash size spikes also had GLM jumps associated with them, especially during the decaying

time period of the storm. A finding of Mecikalski et al. (2015) was that there were increases in flash size during weaker or decaying periods of storms. Therefore this flash splitting issue may explain at least a portion of the false jumps in the GLM found during the decaying period of severe storms and during non-severe storms as well. This particular issue may require more attention as it may be the most directly related to the high false alarm rate and may be something that gets fixed in later versions of the GLM processing methods.

5.1.4 Weak Jumps

There were two GLM jumps noted in the small scale comparisons (Sections 4.1.1.2 and 4.1.1.4) that did not subjectively appear to be jumps. These did not subjectively appear to be jumps because there was only a small increase in flash rates (2 flash increase in one case) associated with the jump. Both of these jumps occurred within times that there was little variability in GLM flash rates the prior 10 minutes of the storm. These near constant flash rates over long enough periods of time to effect the jump is not often seen in the LMA which appears to have a lot more variability. As mentioned in the background, sigma is the current time rate of change of two consecutive two minute flash rate bins (DFRDT) divided by the standard deviation of the prior 5 DFRDT bins (10 minutes). Therefore, it is possible that the GLM is more susceptible to weak jumps caused by small flash rate increases coupled with little prior 10 minute variability not often seen in the LMA.

Figure 5.1 shows cumulative distribution frequencies of DFRDT values at the time of a jump and the standard deviation of DFRDT values in the 10 minutes prior.

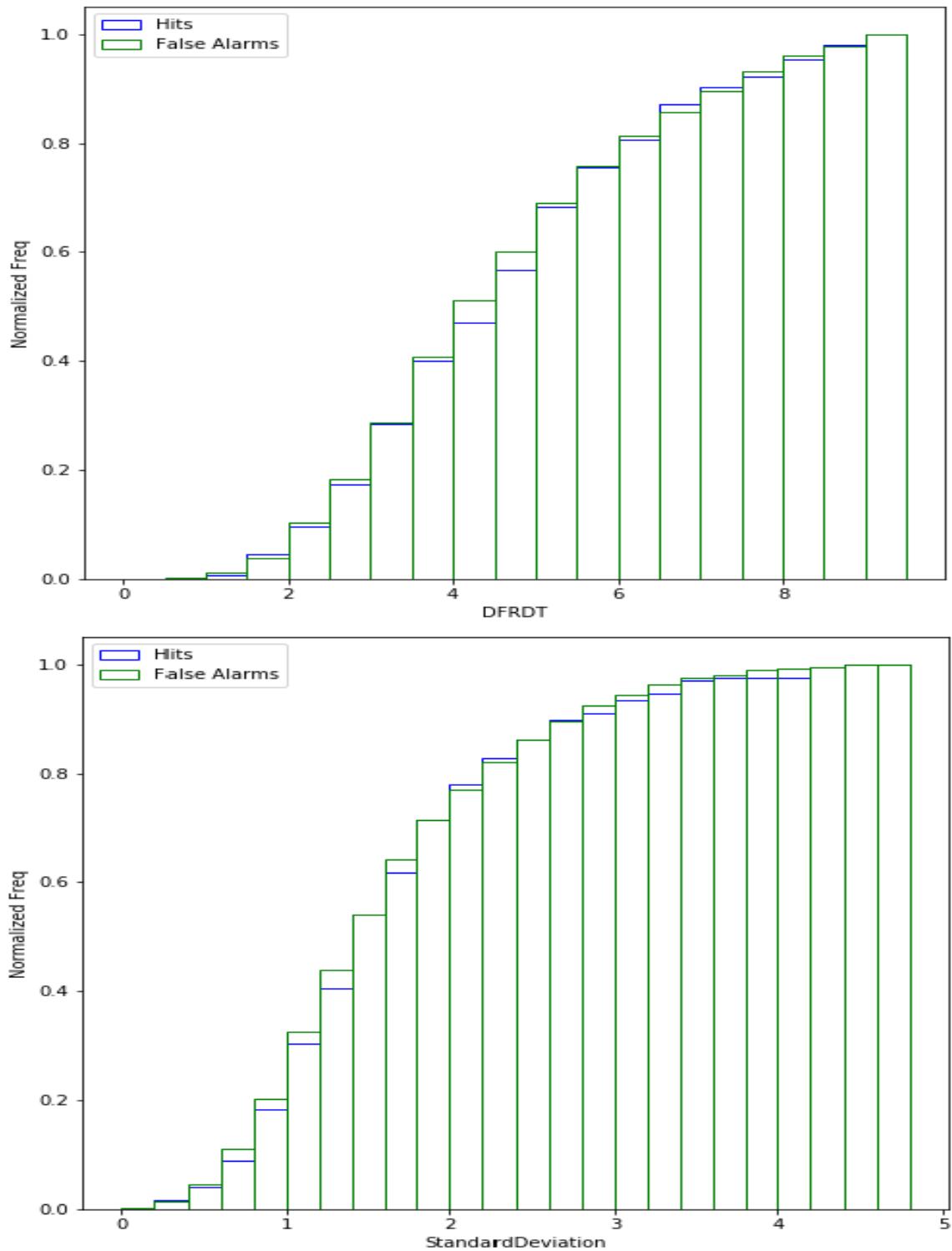


Figure 5.1: Cumulative distribution frequencies of DFRDT values at the time of jumps (top) and the standard deviation of the DFRDT values in the 10 minutes prior to the jump (bottom) for both hits (blue) and false alarms (green).

These distributions were made to determine if there were more false alarms occurring due to weaker standard deviation values due to decreased variability than hits. It does appear that there are slightly more false alarms at weaker standard deviations than hits. However, there is significant overlap, thus adding any sort of minimum standard deviation threshold as well would also decrease the amount of hits. This again indicates that weak jumps may explain some small portion of the false alarms, but there are some weak jumps also associated with hits.

Chapter 6

SUMMARY AND CONCLUSIONS

6.1 Summary

During the small sample comparisons of the GLM and the LMA for the same subset of storms it was found that there were differences, sometimes significant, in flash rates and trends between the two systems. These disagreements caused for there to be differences in how the LJA acted between the two systems. There were a few items of particular note that were also looked at during the large sample study conducted. The biggest item was that initial results running the original algorithm on the GLM showed there were more false alarm jumps during non-severe convection and during the decaying portion of stronger storms as well the lightning being less correlated with radar metrics, which indicates there are errors in the LJA on the GLM. These errors are thought to be one or all of the following: a breakdown of the physical conceptual model of the LJA, GLM flash measurement error, or GLM LJA system error including tracking and lightning association. In the large sample study it was found that there was a relatively high false alarm rate at 82% which is roughly 20% higher than any other study had seen prior, while the POD of 68% lined up well with the Schultz et al. (2016b) study. Distributions of MESH and VIL were taken for

the fifteen minutes centered on every jump in the large sample study. It was found many of the false alarms came from weaker storms via MESH and VIL as seen in the small sample study.

Storms in the LMA vs GLM comparison were taken from various regions to try and diagnose if there was any bias in how the two systems compared based on region. It was found that the aforementioned issues occurred in most of the storms regardless of region. The large sample study also did spatial sensitivity testing through 13 sub-domains and found similar results for all the domains except two, with one of those two likely being biased by a small sample. Therefore it was concluded that the LJA is not likely spatially sensitive.

In the large sample study there was also sensitivity testing done on the sigma and minimum flash per minute thresholds. Probability of detection was its greatest when these values were the most relaxed (low sigma and low flash per minute minimums). This was likely due to artificially increasing the number of jumps by changing its definition leading to more hits due to randomness. The false alarm rate was lowest at lower sigma and high flash per minute minimums. The lower values at lower sigma were associated with a large increase in the raw number of false alarms, but the percentage increase in false alarms was less than the increase in hits causing a decrease in the FAR. The best skill scores was at a configuration of 1.1 sigma and 25 flash minimum that had a POD of 61% and FAR of 74%.

In the GLM vs LMA comparisons it was also noted that the LMA started jumping earlier and more often at the beginning of a storm's intensification when compared to the GLM. These results were mirrored in the large sample study where

it was found that there were many cases where the first jump of the storm did not occur until after very high values (greater than 40 mm) of MESH and after the max MESH values of a storm.

Overall, the LJA while ran on the GLM performed worse than on the GLM than on LMAs in terms of false alarms and skill scores, and similar to slightly worse in terms of probability of detection.

6.2 Conclusion

The key results from this study are summarized in the following list:

- The most significant differences in flash rates and trends were found in extraordinary intense storms with high flash rates as seen by the LMA. During these storms LMA saw as many as three times the flashes as the GLM.
- Given the large differences in the subset of 7 storms there were subsequent differences in how the LJA ran as well. The LMA tended to jump earlier and more often during the initial intensification of the storms. During the mature phase of the storm the jumps between the two systems lined up the best with a total of 13 jumps being within 10 minutes of each other. However, overall there was little agreement in jump timing with 25 jumps not lining up well. Finally, as the storm began to decay the GLM began to jump a lot with 11 jumps with only 2 jumps seen in the LMA. During non-severe storms the GLM was seen to jump at a relatively high rate with 4 jumps when compared to the LMA at 1 jump. The latter results suggest there are errors in the LJA on the GLM

thought to be pertaining to one or more of the following: a breakdown of the physical conceptual model of the LJA, GLM flash measurement error, or GLM LJA system error including tracking and lightning association.

- When the LJA was run independently on the GLM over a large sample of 930 storms at its original 2-sigma minimum 10 flash per minute threshold there were 273 hits, 129 misses, 1265 false alarms, 67.9% POD, 82.2% FAR, and a 0.16 CSI score. Spatial sensitivity testing yielded a result of the LJA not being sensitive to various spatial controls on the GLM. Threshold sensitivity testing found that lower sigma values with higher flash per minute thresholds had the highest skill scores due to higher increases in the percentage of hits despite large increases in the raw number of false alarms. The configuration with the best skill score was 1.1 sigma and 25 flashes per minute and yielded a POD of 61.2%, FAR of 74.5%, and CSI of 0.22.
- The high false alarm rate was the most important finding of the LJA on the GLM. MESH and VIL distributions of jumps that were hits and false alarms yielded a result that false alarms were on average weaker than hits. For MESH false alarms were 7.5 mm weaker on mean and 7.2 mm weaker on median. For VIL false alarms were 5.4 kg/m² on mean and 7.8 kg/m² on median.

6.3 Future Work

There are many additional lines of potential future work found in this study. A lot of the work has to do with diagnosing and potentially fixing the issues of the

high number of false alarms. The hypothesized set of potential errors in this study was discussed in depth in Chapter 5. Recommendations on the work moving forward from that speculative discussion and a few other notes are discussed in this section.

The first recommendation is to do further and deeper analysis on the tracking method as discussed in Section 5.1.2. A lot of potential issues were found with the tracking throughout this study. These potential issues ranged from choosing where to start your grid matters to lightning being missed due to the tracking being more biased towards the radar values. One recommendation is moving towards using flash extent densities rather than flash centroids to better lightning association. It is also recommended that the tracking is moved to a GLM primary or GLM only tracking method using the fixed GLM grid. This would remove the need to arbitrarily choose the grid for each individual case and may alleviate some of the problems with lightning not being associated with features due to the tracking being biased by radar. In a GLM only tracking algorithm other instruments on the GOES series satellites could be used to track storms prior to initial lightning production.

Another recommendation is studying the potential issue of the GLM splitting large flashes into many small flashes as discussed in Section 5.1.3. Larger flashes tend to occur in weaker storms and during the decay of stronger storms. These are two periods where it was noted in the small sample study that there were an increased number of lightning jumps in the GLM where they normally would not be expected and not seen in the LMA. This issue could also have large impacts on the false alarm rate. A flash by flash study during the aforementioned times could be done

to determine the exact effects of this issue, and bettering flash clustering algorithms and filtering methods could also help to alleviate this issue.

A further recommendation would be to analyze the sensitivity of these results to varying GLM data sets as mentioned in Section 5.1.1. There exists other GLM data sets that go through different processing techniques that given different flash rates and trends. It would be important to see if differing data sets given the same sample of storms noted here would result in notably different outcomes, which could help narrow down what potential errors are causing the most problems in the LJA on the GLM.

In section 5.1.4 the potential error of weak jumps due to decreased variability of GLM flash rates is discussed. It was found that there are slightly more false alarms at weaker standard deviation values than hits, but there is large overlap between the two distributions. It is recommended that similar distributions are made using the LMA data sets from prior studies such as Schultz et al. (2016b) to compare the two systems. If there are overall more jumps occurring due to decreased variability of GLM flash rates when compared to the LMA it may be worth looking into different algorithms to diagnose jumps on the GLM than the sigma algorithm as defined in Schultz et al. (2009).

Since the beginning of optical lightning detecting systems there have been studies looking into the decreased DEs of these systems due to optical extinction. Optical extinction is caused when the light from a lightning flash is scattered so much by cloud ice and water that it reduces or completely suppresses the amount of light escaping the cloud to the point that the optical instruments will not detect it (Thomson and

Krider 1982). The GLM has been noted to also have reduced DEs when there is more cloud ice and water (Rutledge et al. 2018). Therefore another recommendation is to look at dual-polarization radar products and hydrometer identifications to determine the potential role of increase light scattering and optical extinction on GLM flash rate variability.

Another potential area that is recommend to be looked into further is flash optical energy. The GLM is able to detect the magnitude of the light escaping the top of a cloud from a lightning flash (Goodman et al. 2013). Prior studies have shown that while individual flashes emit more optical energy in weaker storms, that overall stronger storms emit more total optical energy (Peterson and Liu 2013). Future studies will try and further explore this link between optical energy and storm strength in the context of the GLM.

It may also be worth investigating in more depth than was done in this study whether or not adding radar thresholds, such as a minimum MESH value, benefits the LJA on the GLM in any way. Lastly, an analysis on the lead times of GLM jumps to severe weather reports will also be conducted on the large sample storms.

BIBLIOGRAPHY

- Baker, B., M. B. Baker, E. R. Jayaratne, J. Latham, and C. P. R. Saunders, 1987: The influence of diffusional growth rates on the charge transfer accompanying rebounding collisions between ice crystals and soft hailstones. *Quarterly Journal of the Royal Meteorological Society*, **113** (478), 1193–1215, doi:10.1002/qj.49711347807.
- Baker, M. B. and J. G. Dash, 1994: Mechanism of charge transfer between colliding ice particles in thunderstorms. *Journal of Geophysical Research: Atmospheres*, **99** (D5), 10 621–10 626, doi:10.1029/93JD01633.
- Bateman, M., 2018: Preliminary glm detection efficiency and climatology. GLM Science Team Meetings.
- Bitzer, P., 2018: Glm flash size and detection efficiency. Personal Communication.
- Brimelow, J. C., G. W. Reuter, and E. R. Poolman, 2002: Modeling maximum hail size in alberta thunderstorms. *Weather and Forecasting*, **17** (5), 1048–1062, doi: 10.1175/1520-0434(2002)017<1048:MMHSIA>2.0.CO;2.
- Bruning, E., 2011: A lightning mapping array for west texas: Deployment and research plans. 5th Conference on Meteorological Applications of Lightning Data.
- Bruning, E., 2017: Glm and wtlma flash rate and energy analyses. GLM Science Team Meetings.
- Bruning, E. C. and D. R. MacGorman, 2013: Theory and observations of controls on lightning flash size spectra. *Journal of the Atmospheric Sciences*, **70** (12), 4012–4029, doi:10.1175/JAS-D-12-0289.1.
- Bruning, E. C., S. A. Weiss, and K. M. Calhoun, 2014: Continuous variability in thunderstorm primary electrification and an evaluation of inverted-polarity terminology. *Atmospheric Research*, **135-136**, 274 – 284, doi:<https://doi.org/10.1016/j.atmosres.2012.10.009>.
- Butterworth, S., 1930: On the theory of filter amplifiers. *Experimental Wireless & Wireless Engineering*, **7**, 536–541.
- Carey, L. D. and S. A. Rutledge, 1996: A multiparameter radar case study of the microphysical and kinematic evolution of a lightning producing storm. *Meteorology and Atmospheric Physics*, **59** (1), 33–64, doi:10.1007/BF01032000.

- Carey, L. D. and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe hailstorm. *Journal of Geophysical Research: Atmospheres*, **103** (D12), 13 979–14 000, doi:10.1029/97JD02626.
- 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, doi:10.1175/1520-0493(2000)128<2687:TRBPAL>2.0.CO;2.
- 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, doi:10.1002/2016JD025159.
- Chronis, T., L. D. Carey, C. J. Schultz, E. V. Schultz, K. M. Calhoun, and S. J. Goodman, 2015: Exploring lightning jump characteristics. *Weather and Forecasting*, **30** (1), 23–37, doi:10.1175/WAF-D-14-00064.1.
- Cintineo, J. L., T. M. Smith, V. Lakshmanan, H. E. Brooks, and K. L. Ortega, 2012: An objective high-resolution hail climatology of the contiguous united states. *Weather and Forecasting*, **27** (5), 1235–1248, doi:10.1175/WAF-D-11-00151.1.
- Crum, T. D. and R. L. Alberty, 1993: The wsr-88d and the wsr-88d operational support facility. *Bulletin of the American Meteorological Society*, **74** (9), 1669–1688, doi:10.1175/1520-0477(1993)074<1669:TWATWO>2.0.CO;2.
- Davis, S. M. and J. G. LaDue, 2004: Nonmeteorological factors in warning verification. 22nd Conference on Severe Local Storms.
- Deierling, W., W. A. Petersen, J. Latham, S. Ellis, and H. J. Christian, 2008: The relationship between lightning activity and ice fluxes in thunderstorms. *Journal of Geophysical Research: Atmospheres*, **113** (D15), doi:10.1029/2007JD009700.
- DiGangi, E. A., D. R. MacGorman, C. L. Ziegler, D. Betten, M. Biggerstaff, M. Bowlan, and C. K. Potvin, 2016: An overview of the 29 may 2012 kingfisher supercell during dc3. *Journal of Geophysical Research: Atmospheres*, **121** (24), 14,316–14,343, doi:10.1002/2016JD025690.
- Dobur, J. C., 2005: A comparison of severe thunderstorm warning verification statistics and population density within the nws atlanta county warning area. National Weather Service.
- 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, doi:10.1002/2015JD024663.
- Gardiner, B., D. Lamb, R. L. Pitter, J. Hallett, and C. P. R. Saunders, 1985: Measurements of initial potential gradient and particle charges in a montana summer

- thunderstorm. *Journal of Geophysical Research: Atmospheres*, **90** (D4), 6079–6086, doi:10.1029/JD090iD04p06079.
- Goodman, S. J., D. E. Buechler, P. D. Wright, and W. D. Rust, 1988: Lightning and precipitation history of a microburst-producing storm. *Geophysical Research Letters*, **15** (11), 1185–1188, doi:10.1029/gl015i011p01185.
- Goodman, S. J., et al., 2013: The goes-r geostationary lightning mapper (glm). *Atmospheric Research*, **125-126** (1), 34 – 49, doi:<https://doi.org/10.1016/j.atmosres.2013.01.006>.
- Greene, D. R. and R. A. Clark, 1972: Vertically integrated liquid water—a new analysis tool. *Monthly Weather Review*, **100** (7), 548–552, doi:10.1175/1520-0493(1972)100<0548:VILWNA>2.3.CO;2.
- Holton, J. R., 2004: *An introduction to dynamic meteorology*. 4th ed., International Geophysics Series, Elsevier Academic Press,, 535 pp.
- Horvat, D. J., C. A. Horvat, C. Calvert, and T. Crum, 2011: The refreshed wsr-88d level ii data collection and distribution network. WSR-88D Radar Operations Center.
- Jayaratne, E. R., C. P. R. Saunders, and J. Hallett, 1983: Laboratory studies of the charging of soft-hail during ice crystal interactions. *Quarterly Journal of the Royal Meteorological Society*, **109** (461), 609–630, doi:10.1002/qj.49710946111.
- Kohl, D. A., 1962: Sferics amplitude distribution jump identification of a tornado event. *Monthly Weather Review*, **90** (11), 451–456, doi:10.1175/1520-0493(1962)090<0451:SADJIO>2.0.CO;2.
- Koshak, W. J., 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, doi:10.1175/1520-0426(2004)021<0543:NALMAL>2.0.CO;2.
- Lakshmanan, V., R. Rabin, and V. DeBrunner, 2003: Multiscale storm identification and forecast. *Atmospheric Research*, **67-68**, 367 – 380, doi:[https://doi.org/10.1016/S0169-8095\(03\)00068-1](https://doi.org/10.1016/S0169-8095(03)00068-1).
- Lakshmanan, V., T. Smith, G. Stumpf, and K. Hondl, 2007: The warning decision support system—integrated information. *Weather and Forecasting*, **22** (3), 596–612, doi:10.1175/WAF1009.1.
- 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, doi:10.1175/MWR-D-13-00184.1.

- MacGorman, D. R., D. W. Burgess, V. Mazur, W. D. Rust, W. L. Taylor, and B. C. Johnson, 1989: Lightning rates relative to tornadic storm evolution on 22 may 1981. *Journal of the Atmospheric Sciences*, **46** (2), 221–251, doi:10.1175/1520-0469(1989)046<0221:LRRTTS>2.0.CO;2.
- Mach, D., 2017a: Reprocessed glm data. Personal Communication.
- Mach, D., 2017b: Road to glm provisional validation: Deep dive on selected challenges to provisional, part 12+. GLM Science Team Meetings.
- 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*, **110** (D12), doi:10.1029/2004JD005287.
- 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, doi:10.1175/MWR-D-14-00250.1.
- Metzger, E. and W. A. Nuss, 2013: The relationship between total cloud lightning behavior and radar-derived thunderstorm structure. *Weather and Forecasting*, **28** (1), 237–253, doi:10.1175/WAF-D-11-00157.1.
- 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, doi:10.1002/2014EA000051.
- Pereyra, R. G., E. E. Avila, N. E. Castellano, and C. P. R. Saunders, 2000: A laboratory study of graupel charging. *Journal of Geophysical Research: Atmospheres*, **105** (D16), 20 803–20 812, doi:10.1029/2000JD900244.
- Peterson, M. and C. Liu, 2013: Characteristics of lightning flashes with exceptional illuminated areas, durations, and optical powers and surrounding storm properties in the tropics and inner subtropics. *Journal of Geophysical Research: Atmospheres*, **118** (20), 11,727–11,740, doi:10.1002/jgrd.50715.
- Ringhausen, J., 2018: Differing glm datasets. Personal Communication.
- Rison, W., R. J. Thomas, P. R. Krehbiel, T. Hamlin, and J. Harlin, 1999: A gps-based three-dimensional lightning mapping system: Initial observations in central new mexico. *Geophysical Research Letters*, **26** (23), 3573–3576, doi:10.1029/1999GL010856.
- Rudlosky, S. D. and H. E. Fuelberg, 2013: Documenting storm severity in the mid-atlantic region using lightning and radar information. *Monthly Weather Review*, **141** (9), 3186–3202, doi:10.1175/MWR-D-12-00287.1.
- Rutledge, S., K. Hilburn, B. Fuchs, W. Xu, and A. Clayton, 2018: Impact of ice water path and flash area on glm de. GLM Science Team Meetings.

- Saunders, C., S. Peck, G. A. Varela, E. Avila, and N. Castellano, 2001: A laboratory study of the influence of water vapour and mixing on the charge transfer process during collisions between ice crystals and graupel. *Atmospheric Research*, **58** (3), 187 – 203, doi:[https://doi.org/10.1016/S0169-8095\(01\)00090-4](https://doi.org/10.1016/S0169-8095(01)00090-4).
- Saunders, C. P. R., H. Bax-norman, C. Emersic, E. E. Avila, and N. E. Castellano, 2006: Laboratory studies of the effect of cloud conditions on graupel/crystal charge transfer in thunderstorm electrification. *Quarterly Journal of the Royal Meteorological Society*, **132** (621), 2653–2673, doi:10.1256/qj.05.218.
- Saunders, C. P. R. and S. L. Peck, 1998: Laboratory studies of the influence of the rime accretion rate on charge transfer during crystal/graupel collisions. *Journal of Geophysical Research: Atmospheres*, **103** (D12), 13 949–13 956, doi:10.1029/97JD02644.
- 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*, **30** (6), 1591–1621, doi:10.1175/WAF-D-14-00147.1.
- Schultz, C. J., L. D. Carey, E. V. Schultz, and R. J. Blakeslee, 2017: Kinematic and microphysical significance of lightning jumps versus nonjump increases in total flash rate. *Weather and Forecasting*, **32** (1), 275–288, doi:10.1175/WAF-D-15-0175.1.
- Schultz, C. J., T. Chronis, S. M. Stough, L. D. Carey, and D. J. Cecil, 2016a: The temporal and probabilistic relationship between lightning jump occurrence and radar-derived thunderstorm intensification. 28th Conference on Severe Local Storms.
- Schultz, C. J., W. A. Petersen, and L. D. Carey, 2009: Preliminary development and evaluation of lightning jump algorithms for the real-time detection of severe weather. *Journal of Applied Meteorology and Climatology*, **48** (12), 2543–2563, doi:10.1175/2009JAMC2237.1.
- Schultz, C. J., W. A. Petersen, and L. D. Carey, 2011: Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. *Weather and Forecasting*, **26** (5), 744–755, doi:10.1175/WAF-D-10-05026.1.
- Schultz, E. V., C. J. Schultz, L. D. Carey, D. J. Cecil, and M. Bateman, 2016b: Automated storm tracking and the lightning jump algorithm using goes-r geostationary lightning mapper (glm) proxy data. *Journal of Operational Meteorology*, **4** (7), 92–107, doi:10.15191/nwajom.2016.0407.
- Stanford, S. L., M. A. Lind, and G. S. Takle, 1971: Electromagnetic noise studies of severe convective storms in iowa: The 1970 storm season. *Journal of the Atmospheric Sciences*, **28** (3), 436–448, doi:10.1175/1520-0469(1971)028<0436:ENSOSC>2.0.CO;2.
- Stano, G. T., C. J. Schultz, L. D. Carey, D. R. MacGorman, and K. M. Calhoun, 2014: Total lightning observations and tools for the 20 may 2013 moore, oklahoma,

- tornadic supercell. *Journal of Operational Meteorology*, **2** (7), 71–88, doi:10.15191/nwajom.2014.0207.
- Stolzenburg, M., W. D. Rust, B. F. Smull, and T. C. Marshall, 1998: Electrical structure in thunderstorm convective regions: 1. mesoscale convective systems. *Journal of Geophysical Research: Atmospheres*, **103** (D12), 14 059–14 078, doi:10.1029/97JD03546.
- Takahashi, T., 1978: Riming electrification as a charge generation mechanism in thunderstorms. *Journal of the Atmospheric Sciences*, **35** (8), 1536–1548, doi:10.1175/1520-0469(1978)035<1536:REAACG>2.0.CO;2.
- Thomas, R., P. Krehbiel, W. Rison, M. Stanley, and A. Attanasio, 2018: Comparison of glm with lma lightning detection. GLM Science Team Meetings.
- 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. *Journal of Geophysical Research: Atmospheres*, **109** (D14), doi:10.1029/2004JD004549.
- Thomson, L. W. and E. P. Krider, 1982: The effects of clouds on the light produced by lightning. *Journal of the Atmospheric Sciences*, **39** (9), 2051–2065, doi:10.1175/1520-0469(1982)039<2051:TEOCOT>2.0.CO;2.
- Trostel, J., J. Losego, and M. Frank, 2018: Intercomparison of lma and glm over north georgia. GLM Science Team Meetings.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3d ed., Oxford, 676 pp.
- Williams, E., et al., 1999: The behavior of total lightning activity in severe florida thunderstorms. *Atmospheric Research*, **51** (3), 245 – 265, doi:[https://doi.org/10.1016/S0169-8095\(99\)00011-3](https://doi.org/10.1016/S0169-8095(99)00011-3).
- Williams, E. R., 1989: The tripole structure of thunderstorms. *Journal of Geophysical Research: Atmospheres*, **94** (D11), 13 151–13 167, doi:10.1029/JD094iD11p13151.
- Williams, E. R. and R. M. Lhermitte, 1983: Radar tests of the precipitation hypothesis for thunderstorm electrification. *Journal of Geophysical Research: Oceans*, **88** (C15), 10 984–10 992, doi:10.1029/JC088iC15p10984.
- Williams, E. R., R. Zhang, and J. Rydock, 1991: Mixed-phase microphysics and cloud electrification. *Journal of the Atmospheric Sciences*, **48** (19), 2195–2203, doi:10.1175/1520-0469(1991)048<2195:MPMACE>2.0.CO;2.
- Witt, A., M. D. Eilts, G. J. Stumpf, J. T. Johnson, E. D. W. Mitchell, and K. W. Thomas, 1998: An enhanced hail detection algorithm for the wsr-88d. *Weather and Forecasting*, **13** (2), 286–303, doi:10.1175/1520-0434(1998)013<0286:AEHDAF>2.0.CO;2.