3-22-2018

Forecasting Lightning Cessation Using Dual-Polarization Radar and Lightning Mapping Array near Washington, D.C.

Nancy M. Holden

Forecasting Lightning Cessation Using Dual-Polarization Radar and Lightning Mapping Array near Washington, D.C.

THESIS

Nancy Marie Holden, Capt, USAF

AFIT-ENP-MS-18-M-085

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-MS-18-M-085

FORECASTING LIGHTNING CESSATION USING DUAL-POLARIZATION RADAR AND LIGHTNING MAPPING ARRAY NEAR WASHINGTON, D.C.

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Atmospheric Science

Nancy Marie Holden, B.S.

Capt, USAF

22 Mar 2018

AFIT-ENP-MS-18-M-085

FORECASTING LIGHTNING CESSATION USING DUAL-POLARIZATION RADAR AND LIGHTNING MAPPING ARRAY NEAR WASHINGTON, D.C.

THESIS

Nancy Marie Holden, B.S.
Capt, USAF

Committee Membership:

Maj O. Nava, Ph.D.
Chair

Maj C. Lewis, Ph.D.
Member

William P. Roeder, M.S.
Member

Abstract

Accurate forecasts of thunderstorms are important to space launch, aviation, and public safety. While prior studies have primarily focused on atmospheric conditions leading to lightning onset, less research has been dedicated to the challenging problem of predicting lightning cessation. This study verifies the probabilistic lightning cessation model developed by Joseph Patton (2017) at Florida State University for use by the U.S. Air Force's 45th Weather Squadron at Cape Canaveral Air Force Station (CCAFS) and the National Aeronautics and Space Administration (NASA) Kennedy Space Center (KSC). The Washington, D.C. greater metropolitan area, which presents a climate different to that of central Florida, was chosen as the domain of study. Consistent results would build confidence for use of the method at CCAFS/KSC and lend credence for use at other locations and possible implementation as a product for the Next-Generation Radar network.

The lightning cessation algorithm employs the use of dual-polarization radar and New Mexico Tech Lightning Mapping Array in and around the Washington, D.C area. The algorithm incorporates the presence of graupel at -5°C, -10°C, -15°C, and -20°C levels, maximum reflectivity at 0°C, and composite (maximum) reflectivity in a generalized linear model. The model was tested for three probability thresholds: 95.0%, 97.5%, and 99.0%. A database of 47 isolated, warm season storms in the greater metropolitan Washington, D.C. area were tracked. Performance statistics show that the model revealed notable skill in the Washington, D.C. area, yet not to the desired level as indicated by the model's performance in central Florida.

Acknowledgements

First and foremost I would like to thank my loving mother and father who have been a constant source of motivation and support throughout my career and academic endeavors. I am also grateful for my boyfriend for providing continual love and support, as well as sharing his time and talent in graphic design.

I would like to thank my thesis advisor for the contributions he has made to this thesis by challenging me and guiding me throughout the entire process. I would also like to recognize the other committee members for sharing their expertise and and thoughtful feedback throughout the thesis process.

Finally, I would like to thank my classmates for the many hours spent collaborating and studying at AFIT and the motivation they provided. I would also like to give a special thanks to my lightning initiation counterpart, and close friend, for sharing long days and nights working with me in the weather lab, without whose constancy and support I would not have been able to complete this thesis. I wish you all the best in your career and future assignments.

Nancy Marie Holden

Table of Contents

List of Figures

List of Tables

List of Acronyms

45 WS 45th Weather Squadron

CA cloud-to-air

CC cloud-to-cloud

CCAFS Cape Canaveral Air Force Station

CG cloud-to-ground

CSI Critical Success Index

DCLMA Washington D.C. Lightning Mapping Array

EXPCOR Expected Correct

FAR False Alarm Ratio

GLM generalized linear model

GPS Global Positioning System

HCA Hydrometeor Classification Algorithm

HSS Heidke Skill Score

IC intra-cloud

KAKQ Norfolk/Richmond, VA

KCCX College State, PA

KDOX Dover AFB, DE

KIAD Sterling, VA

KLWX Sterling, VA

KSC Kennedy Space Center

LDAR Lightning Detection and Ranging

LDAR-II Second Generation Lightning Detection and Ranging

LLCC Lightning Launch Commit Criteria

LMA Lightning Mapping Array

MP mixed-phase

NASA National Aeronautics and Space Administration

NCEI National Centers for Environmental Information

NEXRAD Next-Generation Radar

NIC Non-inductive charging

NOAA National Oceanic and Atmospheric Administration

NSSL National Severe Storms Laboratory

NWS National Weather Service

PCHIP Piecewise Cubic Hermite Interpolating Polynomial

PM Percentile Method

POD Probability of Detection

PPI plan position indicator

RF radio frequency

SPC Storm Prediction Center

TOA time of arrival

TSS True Skill Statistic

USAF United States Air Force

UTC Coordinated Universal Time

VCP Volume Coverage Pattern

VHF very high frequency

WDSS-II Warning Decision Support System - Integrated Information

WSR-88D Weather Surveillance Radar, 1988, Doppler

I. Introduction

Over the past 30 years, lightning has ranked among the top five weather-related phenomena to induce the most fatalities in the United States (National Weather Service, 2017). As of 2016, the National Weather Service (NWS) reported a 30-year (1987-2016) average of 47 lightning-related deaths per year. Many of these fatalities are the result of the public's misjudgment on the storm's evolving capacity to generate lightning. In fact, the most dangerous periods for lightning-induced injury or fatality coincide with not just the first lightning flash, but also the last (Holle et al., 1999). As a storm appears to be dissipating, the perceived lightning threat diminishes. Often this occurs prematurely, leading to reduced safety precautions before the lightning potential has fully dissipated. Miscalculations of the timing of lightning cessation has resulted in avoidable casualties. In addition, industries with time sensitive outdoor activities, such as flight line operations at commercial airports, are interested in improved prediction of the timing of lightning cessation in order to assure personnel can return to work as quickly as possible while still maintaining their safety. Consequently, the necessity for further research and a better understanding of the driving mechanisms of lightning cessation is evident.

A recent study by Patton (2017), hereafter referred to as JP17, investigated a new probabilistic strategy for predicting lightning cessation. His model was designed for use by the 45th Weather Squadron (45 WS) at Cape Canaveral Air Force Station (CCAFS)/Kennedy Space Center (KSC). Florida leads the nation in lightning occurrence; it has one of the highest annual cloud-to-ground (CG) lightning density (flashes km^{-2}) in the nation as indicated by Figure 1 (Wolf, 2007). Consequently, there have been more than double the number of lightning-induced fatalities in Florida than in any other state from 2006 to 2015 (Holle, 2016).

The 45 WS utilizes lightning watches to forecast the potential for lightning with

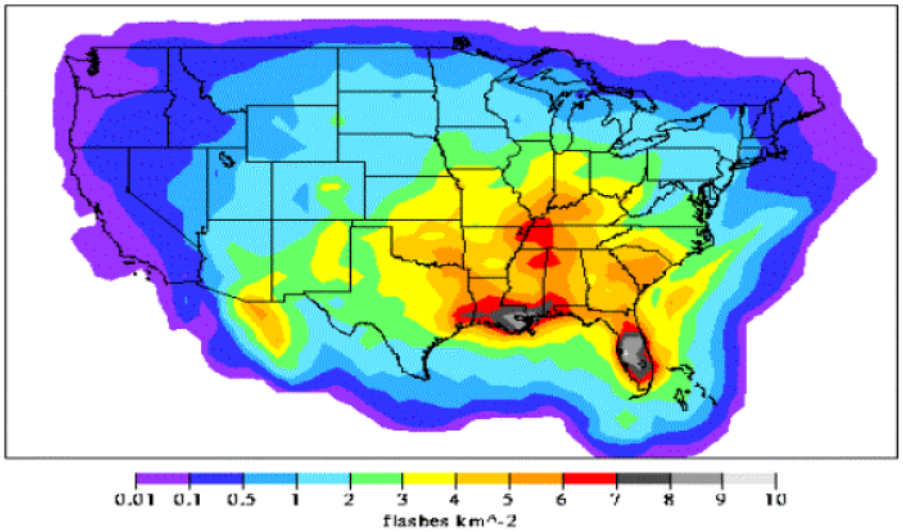

Figure 1. Annual CG lightning density (flashes km-2) across the contiguous United States (Wolf, 2007).

a desired lead-time of 30 min. These alerts prompt resource protection measures to ensure base safety and operational mission success. Once lightning is detected on station or if lightning is imminent, a lightning warning is issued, notifying base personnel and halting operations. Indeed, the prevalence of lightning and its impact on aviation and space launch operations drives the 45 WS's demand for improved lightning guidance. Because there is not as much skillful guidance on how to forecast the end of lightning, after-the-fact analysis has shown that the 45 WS lightning warnings are left issued too long. This is costly due to lost productivity of outside workers, which can delay preparation for space launch and eventually even the space launch schedule.

The 45 WS issues lightning watches and warnings for ten mostly overlapping lightning warning circles at CCAFS, KSC, Patrick AFB, and other locations as depicted in Figure 2 (Roeder et al., 2017). A lightning watch is issued when lightning is expected in the circle(s) with a desired lead-time of 30 min. A lightning warning is issued when lightning is imminent or occurring with the circle(s). The circles have a

Figure 2. The ten lightning warning circles used by 45 WS for CCAFS, KSC, Patrick AFB, and other locations (Roeder et al., 2017).

radius of 5 nm or 6 nm depending if a single small facility is being served, or if several close facilities or a single large facility is being served, respectively. The 45 WS issues the lightning watches and warnings based on total lightning (i.e. both CG lightning and lightning aloft). The top meteorological challenge to the 45 WS has been knowing when to cancel lightning warnings. A set of complex rules, namely the Lightning Launch Commit Criteria (LLCC), is employed by the 45 WS during space launches to protect space launch vehicles and their payloads from natural and rocket-triggered lightning (McNamara et al., 2010). However, these rules provide guidance primarily for the protection against rocket-triggered lightning. Accordingly, the LLCC is more stringent than the standard forecast rules for natural lightning, and the lightning cessation research in this thesis does not apply.

The lightning cessation model developed by JP17 is the latest effort by the 45 WS to improve their rules for lightning cessation and culminates several years of re-

search. A probabilistic framework for predicting lightning cessation has the potential to improve timing precision of the last flash occurrence. This improvement would allow lightning advisories to be canceled sooner and with more confidence. Therefore, operations would be resumed quicker, maximizing time available for outside operations the weeks to months of preparation before space launch.

Upon testing his lightning cessation algorithm in central Florida, JP17 observed positive results. The median wait time expressed by the model was significantly shorter than that of the 45 WS guidance (15-min wait time) currently used. A better understanding of how this model performs can be achieved by testing this method in a climate different to that of central Florida. This concept forms the basis of my study. JP17's lightning cessation model will be tested on a set of isolated thunderstorm cases in and around Washington, D.C. Evaluating and comparing the results of this model in the new environment will provide insight into the model's versatility and efficacy. Consistent results will lend credence to the model's effectiveness and build confidence in the 45 WS's use of the method operationally. With significant positive results, these findings could even contribute to an eventual implementation of a similar algorithm in the Next-Generation Radar (NEXRAD) network.

This chapter introduced the purpose of this study, and provided the scope of the problem at hand. Chapter II provides a review of the cloud electrification mechanisms and previous research on lightning cessation. Additionally, it provides a background of the instruments used, namely dual-polarization radar and the lightning mapping array. Chapter III details the methodology and data used for analysis. Chapter IV presents the results of the analysis. Lastly, Chapter V discusses the conclusions and the significance of the research findings as well as offers a recommendation for future work.

II. Background

2.1 Cloud Electrification

Lightning can be classified into four main categories: CG, cloud-to-cloud (CC), intra-cloud (IC), and cloud-to-air (CA). CG includes any lightning discharge between cloud and earth. Within the CG category there are four main types: downward negative lightning, upward negative lightning, downward positive lightning, and upward positive lightning. The polarity of the CG lightning is defined by the net transfer of charge to the ground: a negative CG flash transfers a net negative electric charge to the ground, and positive CG transfers a net positive charge to the ground. Each of the CG types can be classified as CG discharges, in which the associated electrical charge effectively reaches the surface. In this context, "effectively" refers to the resulting flow of electrons within a lightning channel due to the movement of electrons located elsewhere in the channel. In fact, a direct charge transfer from cloud to ground does not actually occur (Rakov, 2016). In contrast to CG lightning, CC, IC, and CA lightning take place in the air and are collectively referred to as cloud flashes.

There are many possible charging mechanisms that lead up to lightning discharge and flash initiation, both on the ground and in the atmosphere. The cloud electrification process begins with the electrification of individual hydrometeors, followed by the spatial separation of these charged hydrometeors due to polarity. A simplified example schematic of the separation of charges for two distinct locations is depicted in Figure 3.

The two primary electrification mechanism classifications are inductive and non-inductive. Inductive electrification describes a charge separation process involving an external electric field (Rakov, 2016; Saunders, 2008). In a non-inductive electrifica-

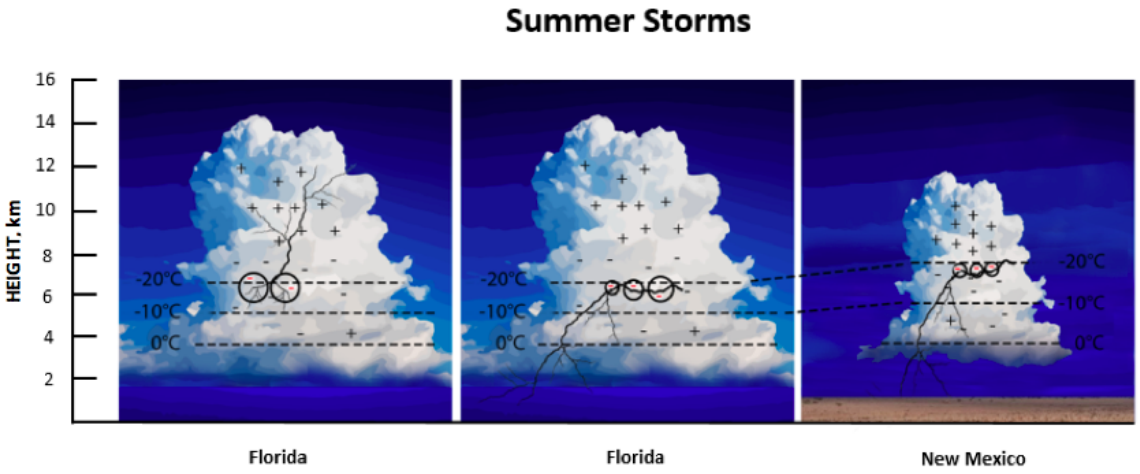

Figure 3. Conceptual model of the charge distribution and tripole structure of an isolated summer thunderstorm in Florida and New Mexico (adapted from Krehbiel (1986)).

tion process, the hydrometeors are polarized internally within the cloud, and do not require polarization by the ambient electric field (Rakov and Uman, 2003). Non-inductive charging (NIC) theory describes a process involving interactions among graupel, ice crystals, and supercooled water droplets (Takahashi, 1978). Both mechanisms play a significant and distinct role in the lightning electrification process. Inductive electrification enhances the electric field concentrations after the initial field is established, while the NIC mechanisms generate the initial charge distribution and polarity (Kuettner et al., 1981). Consequently, NIC drives the initial rapid development of the electric field and is deemed the most widely accepted electrification mechanism to date.

During the non-inductive electrification process, large scale separation of charged particles occurs due to gravity (Rakov, 2016). Ice particles and supercooled water droplets are also mobilized within the cloud. Following the schematic in Figure 4 , these particles collide with individual graupel particles and retain the charge of the opposite species (Zhang et al., 1991). Typically, ice crystals develop a positive

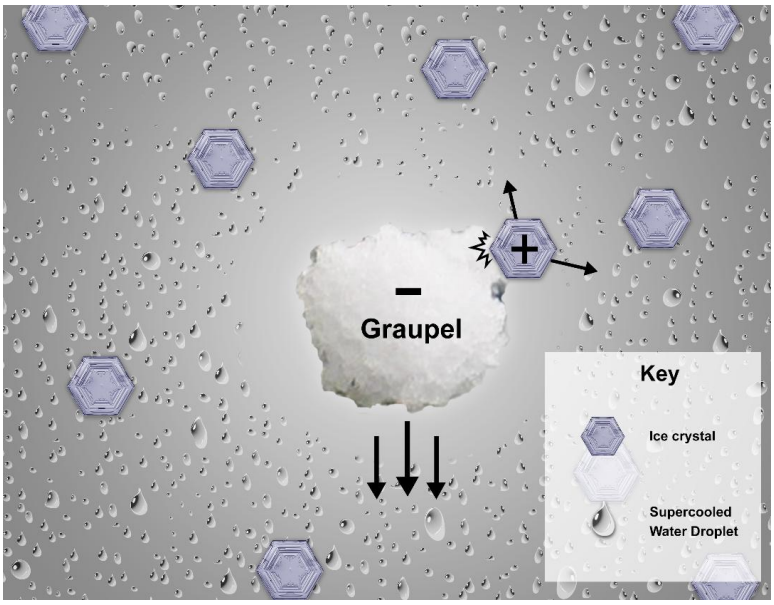

Figure 4. Conceptual schematic of the NIC (adapted from Saunders (2008)). Graupel particles interact with ice crystals in the presence of supercooled water droplets within the mixed-phase region. Collision of these particles prompts a sign change, contributing to the charge distribution necessary for cloud electrification.

charge, while the graupel exhibits a negative charge. However, according to laboratory studies (Jayaratne et al., 1983), a critical cloud temperature exists that dictates the charge of the graupel. Below the critical temperature, graupel particles acquire a negative sign. Once this temperature is exceeded, lower altitude graupel particles acquire a positive sign. This critical temperature is aptly referred to as the "reversal temperature" and ranges from -10°C to -20°C at a height of 6 km (Rakov, 2016).

The presence of graupel at key levels is a major indicator of cloud electrification. Graupel forms when supercooled water droplets, lofted from storm updrafts above the 0°C level, collect and freeze on ice crystals or snowflakes. This interaction typically occurs in the mixed-phase (MP) region of the storm. The MP region is defined as the region between the 0°C and approximately -40°C isothermal levels where supercooled water droplets, graupel, and ice crystals coexist Williams (1985).

As described by Williams (1989), thunderstorms typically exhibit a tripole

charge. The layer between -10°C and -25°C, where graupel formation is most prevalent, is predominantly negatively charged and enveloped by a large positively charged regions above and smaller positively charged region below the middle negative layer (Figure 3). The top positively charged layer is dominated by ice crystals lofted above by the updrafts within the cloud, and the lower positively charged layer is likely due to the positively charged graupel at warmer temperatures below the reversal temperature. According to NIC theory, this tripole structure is enhanced by the collisions of ice crystals and graupel particles within the cloud. It is theorized that when enough charge separation occurs, a preliminary breakdown is eventually able to take place (although the exact mechanism of this process is still unknown) which results in the initiation of the first stepped leader (Rakov, 2013). Therefore, the MP region plays a key role in the lightning electrification and subsequently lightning cessation processes.

Negative polarity CG lightning originates from the negative region of the cloud tripole structure and is prevalent in summertime thunderstorms (Krehbiel et al., 1979). In fact, at least 90% of the global CG lightning is accounted for by negative downward lightning (Rakov, 2013). IC lightning primarily exists between the upper positive and negative layers of the tripole and lightning aloft constitutes 75% of all global lightning (Rakov, 2013).

2.2 Previous Studies

Timing lightning cessation has proven to be particularly challenging due to its unpredictable nature and the lack of understanding with respect to the primary mechanisms that influence it. Anderson (2010) aimed to discern whether an empirical relationship established by Wolf (2007) for CG lightning onset could likewise be applied to lightning cessation: once the 40 dBZ horizontal reflectivity exceeds the -10°C level, CG lightning will occur. However, this hypothesis did not distinguish

between the mechanisms of charge generation versus charge dissipation. Thus, this criteria produced an insignificant correlation with the end of lightning occurrence.

Research motivated by the 45 WS over the past decade has utilized advanced meteorological tools such as Lightning Detection and Ranging (LDAR) and dual-polarization radar data to improve lightning cessation forecasts. A recent study driven by the 45 WS to develop empirical guidance for lightning cessation employed LDAR data, vertical sounding data, and select radar data from 116 storms (Stano et al., 2010). Five statistical and empirical schemes for forecasting lightning cessation were evaluated. Four of the five schemes failed by prompting the cancellation of the lightning warnings too early. The Percentile Method (PM) was the only scheme integrated into 45 WS operational lightning forecasts. Unlike the other four methods, the PM successfully captured the outlier maximum time intervals between strikes. However, it is important to note that this scheme had a propensity to overforecast the wait time after lightning cessation.

More recently, Preston and Fuelberg (2015) implemented the use of dual-polarization radar data in an endeavor to improve lightning cessation prediction. Their results indicated that the best lightning cessation forecasting algorithm identified the presence of graupel and a radar reflectivity of greater than or equal to 35 dBZ at the -10°C temperature altitude. 10 min after this criteria is no longer satisfied, lightning cessation is expected. The algorithm was an effective tool for forecasting lightning cessation in isolated thunderstorms. Thereafter, Davey and Fuelberg (2017) investigated the prediction of lightning cessation for specifically non-isolated cells employing the use of dual-polarization radar products at specific isothermal levels within the MP region. However, applying this method to non-isolated thunderstorms did not produce the same measure of success. In their study, a "non-isolated" thunderstorm was defined as a storm impacted by flashes originating from nearby cells. They concluded that

no combination of dual-polarization or conventional radar products had produced consistent results for safely predicting lightning cessation in non-isolated cells. As a result, more research is required to provide empirical guidance applicable to multi-cell and other non-isolated cell environments.

Founding the basis of my research, JP17 employed dual-polarization radar data to forecast lightning cessation using probabilistic guidance. JP17 utilized several techniques employed in the research by Preston and Fuelberg (2015) and Davey and Fuelberg (2017). First, dual-polarization radar data at specific isothermal heights in the MP region were extracted for input into his lightning cessation model. Specifically, JP17 used the dual-polarization derived Hydrometeor Classification Algorithm (HCA) product to determine graupel presence at the 0°C, -5°C, -10°C, -15°C, and -20°C levels. In addition, maximum reflectivity values at these key levels as well as the maximum composite reflectivity value were utilized.

Next, the radar data was incorporated into two separate predictive models which output a probability for lightning cessation. The idea was to designate probabilistic weights for parameters of statistical significance using a generalized linear model (GLM) to generate a best-fit logistic regression which delivered probabilistic guidance for total lightning occurrence. JP17 tested an independent version and a bootstrapped version of the GLM. While both methods proved to shorten the current 45 WS wait times for forecasting lightning cessation, the bootstrapped version of the GLM was the recommended method. JP17 found that the 97.5% probability threshold of the bootstrapped GLM offered both time savings and a low false alarm ratio. This study focuses on applying JP17's probabilistic GLM on an entirely new climate in the Washington, D.C. area. Its performance will be evaluated and compared to the model's results in central Florida.

2.3 Dual-Polarization Radar

The use of polarized radar sensing in meteorological applications can be traced back to the late 1940s associated with the detection of precipitation echoes (Doviak and Zrnic, 2014). The radar beams were horizontally polarized to increase the return signal scattered off hydrometeors. As rain drops fall, small drops remain spherical since the surface tension of the drop dominates over the aerodynamic forces. In contrast, for large drops the aerodynamic forces become stronger and the drops spread horizontally, such that the horizontal extent of the drop is greater than the vertical extent. Consequently, larger drops provide a greater intensity signal for a horizontally polarized radar beam. This conventional radar technology evolved tremendously over the following decades. When combined with the capabilities of the pulsed-Doppler radar, this technology enabled enhanced meteorological sensing via the penetration of clouds to detect tracers of wind and measure their radial velocities.

A collaborative effort by the National Severe Storms Laboratory (NSSL) and the NWS under the National Oceanic and Atmospheric Administration (NOAA) resulted in the most recent technological advancement of the traditional Doppler radar, the dual-polarization radar. The recent invention of dual-polarization radar has proven to be a useful meteorological forecasting tool in a variety of operational and research settings through the detection of hydrometeors in the atmosphere. Its unique capabilities set it apart from the conventional single-polarization Doppler radar developed in the 1990s. Both conventional radar and dual-polarization radar utilize an active sensing process in which short pulses of electromagnetic wave fields are initially transmitted from the source. The transmitted wave propagates outward until it reaches an object (e.g. rain, hail). The radar pulse scatters off the object and is received by the radar. The signal is then processed through algorithms which decipher characteristics regarding the sensed objects. The primary difference between the conventional

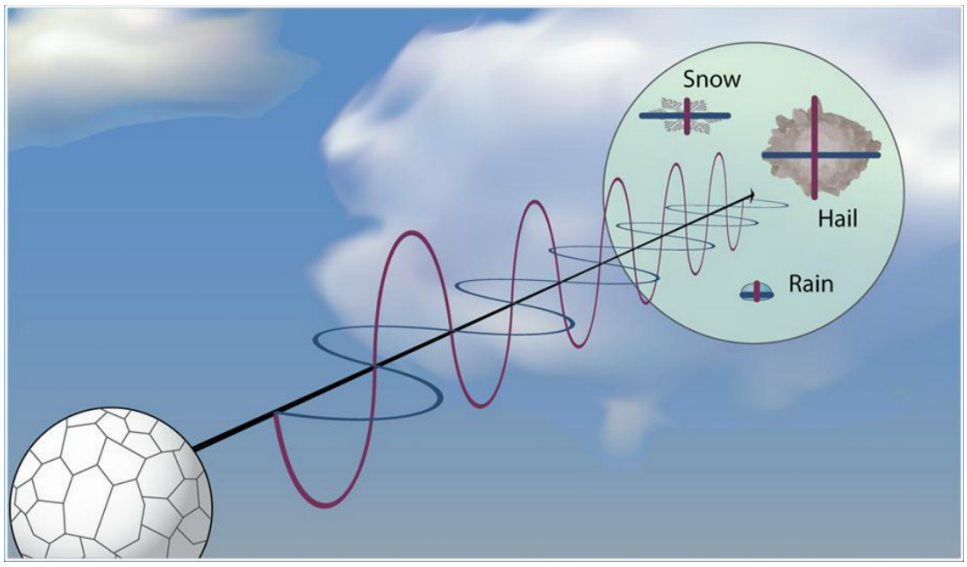

Figure 5. Utilization of horizontal and vertical polarizations allow dual-polarization radar to decipher the size and shape of hydrometeors (NSSL, 2017).

radar and dual-polarization radar is the orientation of polarization of the electromagnetic wave. While conventional radar systems employ horizontally polarized waves, dual-polarization radar employs both horizontally and vertically polarized waves as depicted in Figure 5. This advanced feature enables information regarding both the size and shape of the object to be recorded by the sensor (NSSL, 2017). Additionally, by combining the various dual-polarized radar returns, the most likely hydrometeor species can be identified.

The NWS operates 160 high-resolution 10-cm wavelength, S-band Doppler weather radars, namely the NEXRAD Weather Surveillance Radar, 1988, Doppler (WSR-88D). Beginning in 2011, WSR-88D NEXRAD sites across the nation upgraded to new dual-polarization radar capabilities. These dual-polarization radars provide 14 new NWS derived products in addition to the basic reflectivity and velocity products available through conventional radar systems. WSR-88D radars offer two primary data processing levels: Level-II and Level-III. These processing levels differ in reso-

lution, processing time, and products available. Level-II radar data has a resolution of 250 m x 0.5° in azimuth below 2.4° in elevation, while Level-III radar data has a resolution of 1 km x 1° in azimuth. Level-II base data consist of three meteorological data quantities: reflectivity, mean radial velocity, and spectrum width. There are over 75 Level-III products including both low-bandwidth base products in addition to derived, post-processed products such as HCA, precipitation estimates, and hail estimates.

The HCA can be used in conjunction with key temperature levels to infer possible cloud electrification. The algorithm categorizes each radar echo within the radar beam into the most likely hydrometeor type. The HCA distinguishes between 10 types of radar echoes: hail, graupel, big drops, heavy rain, rain, wet snow, dry snow, ice crystal, ground clutter, and biological. This classification process uses a fuzzy logic scheme described by Ryzhkov et al. (2005) and Schuur et al. (2003). The algorithm combines conventional base radar products as well as dual-polarization parameters with melting layer data (Kumjian, 2013). There are six radar parameters directly utilized in the algorithm: 1) horizontal radar reflectivity, 2) differential reflectivity, 3) cross-correlation coefficient, 4) specific differential phase, 5) a texture parameter of the reflectivity field, and 6) a texture parameter of the field of differential phase (Park et al., 2009). The identification of graupel presence within the MP region using the HCA product is a key component to JP17's lightning cessation model.

2.4 Lightning Mapping Array

Utilizing a network of stations to collectively establish a three-dimensional lightning display, the Lightning Mapping Array (LMA) locates the total lightning present, both aloft and on the ground. The New Mexico Institute of Mining and Technology

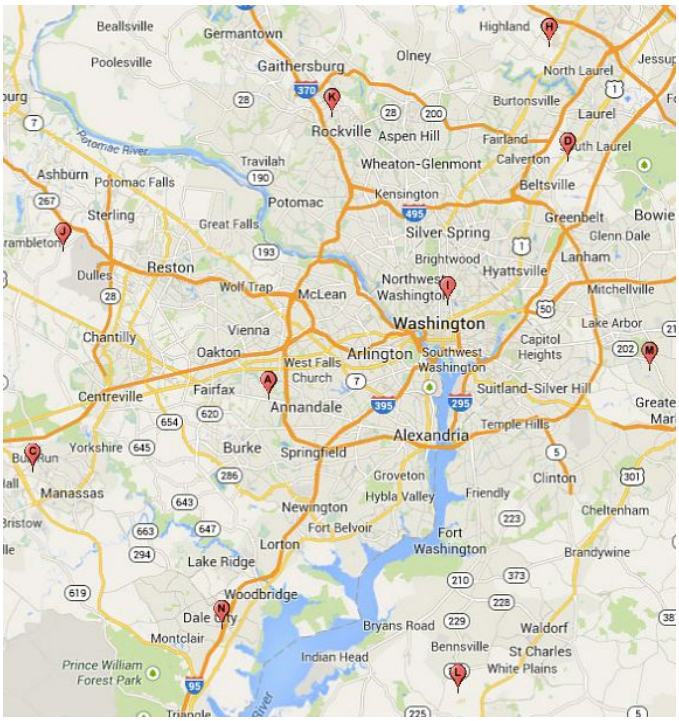

Figure 6. The DCLMA network comprised of 10 stations centered around the greater metropolitan Washington, D.C. area. The network encompasses an area of 70 x 100 km in extent.

LMA system was developed by Bill Rison, Paul Krehbiel, Ron Thomas and colleagues. The lightning detection network is patterned after the LDAR system developed for use at the National Aeronautics and Space Administration (NASA) KSC (Maier et al., 1995). The LMA network locates lightning discharges, including the small-scale, fast components of a lightning flash such as stepped leaders, when dielectric breakdown occurs. The LMA detects these lightning features by pinpointing sources of very high frequency (VHF) radiation using a Global Positioning System (GPS)-based time of arrival technique (Krehbiel et al., 2000; Rison et al., 1999). Pioneered by Proctor (1971), the time of arrival (TOA) method employs the use of six or more stations. Six or more stations are necessary to pinpoint the location of the lightning discharge in three-dimensional space, to eliminate false solutions, and to provide location accuracy

via statistical methods such as chi-square minimization (Thomas et al., 2004).

Arrival times of the VHF radiation in the 6 MHz bandwidth of an unused television channel (centered at approximately 60 MHz) are recorded independently for each of the stations. The TOA at a pair of stations are used to identify the time differences of arrival which constrain the source of the VHF pulse to a hyperbolic surface. A second pair of stations provides a second hyperbolic surface which intersects the first surface providing a three-dimensional curved line on which the discharge occurred. Three or more pairs of stations are necessary to identify the intersection of hyperbolic surfaces which pinpoints the source of the lightning discharge in three-dimensional space as shown in Figure 7 (Roeder, 2010). By connecting stepped leader locations in space and time from the originating flash, the lightning channels can be identified.

For each station, the peak radiation event is recorded in intervals of 80 to 100 microseconds when a predetermined radio frequency (RF) power noise threshold is exceeded (Rison et al., 1999). The LMA processing is done in approximately 1-second segments, whereby all the TOA values are placed into a single array that is sorted chronologically. The three-dimensional capabilities of the LMA system enable mapping of lightning channels with 95.0% flash detection efficiency and 76.9% source detection efficiency within a 100-km radius of all networks (Chmielewski and Bruning, 2016).

There are three established New Mexico Tech LMA networks across the nation in Oklahoma, Alabama, and Washington, D.C. The LMA network located in the Washington, D.C. greater metropolitan area, and used in this study, was developed in a collaborative effort between NASA, NOAA, and New Mexico Tech. The Washington D.C. Lightning Mapping Array (DCLMA) is comprised of 10 stations that span an area of 70 x 100 km in extent (Figure 6). Multiple stations are utilized to prevent the misidentification of noise spikes as sources of radiation at individual locations.

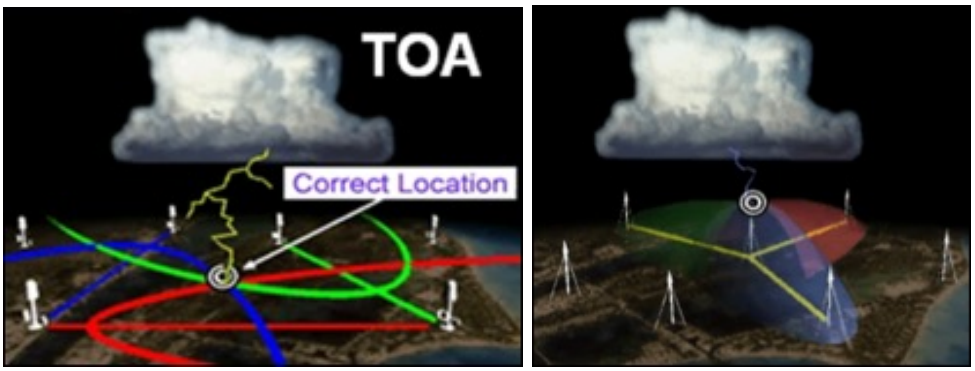

Figure 7. Stepped leader location via differences in time of arrival between pairs of stations (Roeder, 2010). In general, pinpointing the location of a stepped leader requires four intersecting hyperbolae. (a) Two-dimensional view of intersecting hyperbolae (left) and (b) three-dimensional view of intersecting hyperbolae (right)

The network maps total lightning activity over a 200-300 km diameter area around Washington, D.C.

III. Methodology

3.1 Sources of Meteorological Data

A final database of 47 thunderstorm cases were collected from the DCLMA archives for the Washington, D.C area. In order to most accurately test the lightning cessation model developed by JP17, the corresponding meteorological data sources and radar dataset characteristics were emulated as closely as possible. The thunderstorm cases were taken from warm season months (May to September) over a 6-yr span (2012 to 2017). Each storm was manually tracked using dual-polarization radar beginning 16 min prior to lightning cessation to 16 min following lightning cessation, for a total tracking timespan of 33 min. These time restrictions were chosen to emphasize the target time limits required to outperform the standard 15-min wait time presently used by the 45 WS.

3.1.1 Lightning Data

The New Mexico Tech DCLMA network located in Washington, D.C. was used to collect archived lightning flash data for each of the thunderstorm cases. The LMA network functions similarly to the Second Generation Lightning Detection and Ranging (LDAR-II) network (Poehler and Lennon, 1979; Roeder, 2010) located at KSC and used in JP17's study. Patterned after the LDAR-II system, the New Mexico Tech LMA pinpoints the location of VHF radiation to identify lightning channels in three dimensional space. This allows both IC flash channels and CG upper channels to be captured in addition to CG strikes. The LMA network measures the time of arrival of 60 MHz RF radiation from a lightning discharge and the data is displayed approximately once per second. Hourly post-processed data were used to identify lightning flashes. The time of lightning cessation was determined to be the minute

that the last flash was detected by the LMA network. The lightning flashes associated with each isolated storm were tracked using MATLAB code to pinpoint the last flash corresponding with the time of lightning cessation.

3.1.2 Radar Data

Four dual-polarization WSR-88D radars from the NWS in and around the Washington, D.C area were used to extract Level-II and Level-III radar data for each thunderstorm case. The primary radar used for the majority of the cases and within the closest proximity to the Washington, D.C. greater metropolitan area is located in Sterling, VA (KLWX).

A complicating factor with the Level-III data collection process was the lack of radar data for key dual polarization parameters above the 3.5° elevation angle. Specifically, the HCA values at higher elevation angles corresponding to the isothermal levels ranging from -10°C to -20°C could not be collected using KLWX as the sole radar source. Thus, three alternate radars displayed in Figure 8 were used to combat the dearth in radar data. The WSR-88D radar located at Dover AFB, DE (KDOX) was the second closest radar to the Washington, D.C. area at a distance of 144 km from the LMA network center. KDOX was used to collect the majority of the HCA values at higher altitudes. The remaining two alternate radars located at College State, PA (KCCX) and Norfolk/Richmond, VA (KAKQ) were used to collect HCA values at higher altitudes for storms located on the northern and southern fringe of the domain of study, respectively. Additionally, the use of multiple radars mitigated the impact of the "cone of silence", the inverted cone over the radar site where data is unavailable due to close proximity to the radar. These additional radars worked effectively to fill in gaps in data.

Raw Level-II and Level-III radar data downloaded from National Centers for

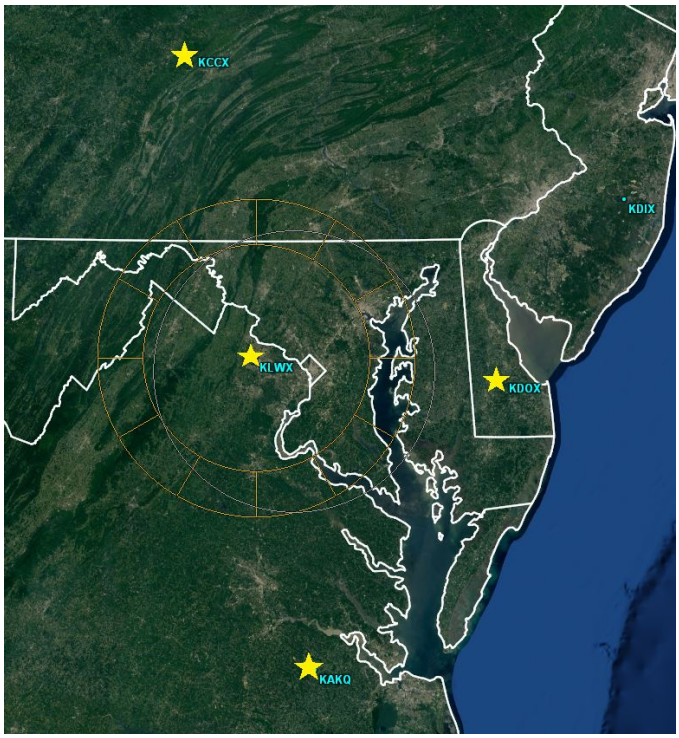

Figure 8. Locations of the four WSR-88D dual-polarization radars used in this study. The yellow stars denote the locations of KLWX, KDOX, KAKQ, and KCCX radars using the GR2Analyst interface.

Environmental Information (NCEI) was ingested and tracked manually via GRLevelX software. Specifically, GR2Analyst and GRLevel3 radar processing and display programs were used to analyze Level-II and Level-III radar data, respectively. GRLevelX was developed by Gibson Ridge Software, LLC, and features a high-speed visual interface for radar data displayed on a high resolution radial grid (1km x 1° x 230 km with 256 data levels).

The radar data was manually tracked at scanning intervals between 3 to 5 minutes, depending on the current Volume Coverage Pattern (VCP). At each interval, specific radar parameters were collected at five isothermal levels: 0°C, -5°C, -10°C, -15°C, and -20°C. These temperature levels were determined using Sterling, VA (KIAD) 00 Coordinated Universal Time (UTC) analyzed atmospheric sounding

Temperature level	Thermal layer (ft)
0°C	13,000-14,000
-5°C	14,000-18,000
-10°C	18,000-21,000
-15°C	21,000-23,000
-20°C	23,000-26,000

Table 1. Thermal layers calculated for each temperature level used for analysis.

text data from the University of Wyoming's Atmospheric Science Department. The height corresponding to each isothermal level was recorded for each of the 47 cases. Since the height variation for each level was minimal (a maximum height variation of 2 kft), values for each isothermal height were averaged for all 47 cases. Furthermore, isothermal heights were converted to isothermal layers for ease of manual analysis (Table 1). Finding the exact isothermal height for the parameters was indeed a challenge with manual analysis, especially with HCA due to limited elevation angles and only plan view capabilities. JP17 also noted vulnerabilities in the model regarding the large vertical gaps between isothermal temperature levels where key data could be overlooked. Thus, using layers instead of singular temperature levels provided a more comprehensive analysis that was also more operationally friendly.

3.2 Convective Cell Selection

An initial database of 135 thunderstorm cases from 2012-17 was collected from the LMA archives for the Washington, D.C. area. These thunderstorm cases were further refined by using an event ranking system that established limiting criteria. Seven limiting criteria were selected for case elimination: cells with non-isolated, severe, or linear convective characteristics, cells exceeding the "effective distance" from the radar and LMA network, and poor health of the radar and LMA network. The flow chart given by Figure 9 illustrates the stringent case selection process.

Each thunderstorm cell was scrutinized for key convective features using the

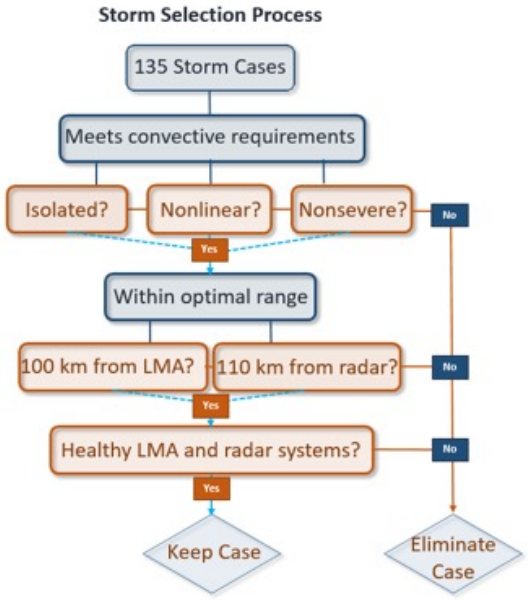

Figure 9. Flow chart outlining the thunderstorm case selection process.

Interactive Radar Map Tool provided by the NCEI under NOAA. Firstly, non-isolated cells were eliminated. Using guidance from Preston and Fuelberg (2015) and JP17, an isolated thunderstorm was defined as having no connecting reflectivity channels greater than 15 dBZ with surrounding cells (see Figure 15 for an observed example). This limitation led to the elimination of over half of the initial 135 cases. For a thunderstorm at the end of its life cycle, it is more common to find a non-isolated thunderstorm cell than an isolated thunderstorm cell. This is due to the increased number of thunderstorms in a given area and, consequently, increased number of thunderstorm cell interactions. Cells were also analyzed visually for linear and severe characteristics. Thunderstorm cells were verified severe if the Storm Prediction Center (SPC) listed any archived severe storm reports associated with the thunderstorm in question. These severe and non-linear thunderstorm cells were eliminated.

Secondly, range limits for cell selection were established from the center of both the LMA network and KLWX radar. According to Chmielewski and Bruning (2016),

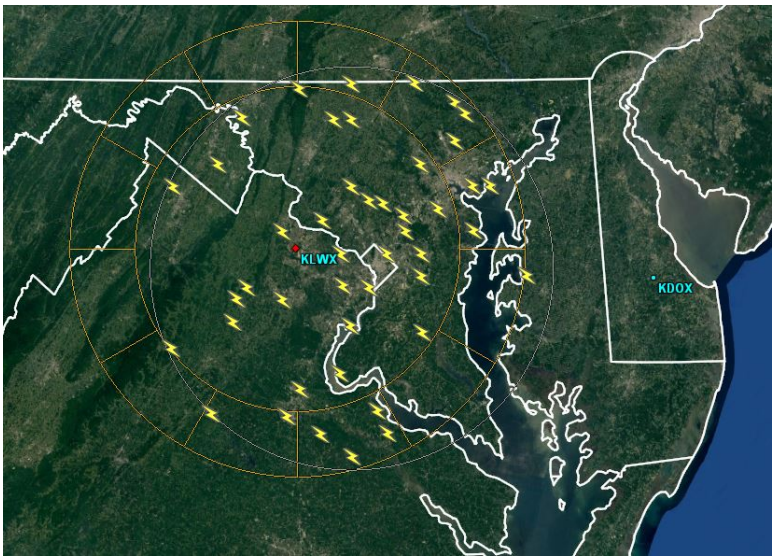

Figure 10. Thunderstorm locations for all 47 cases at the time of lightning cessation depicted by lightning icons. The outermost yellow range ring denotes the 110-km KLWX radar radius and the white range ring denotes the 100-km LMA radius.

within 100 km of the LMA network the flash detection efficiency exceeded 95.0%. Outside this range, the LMA lightning detection capabilities begin to deteriorate. Therefore, only centrally located thunderstorm cells within a 100-km radius from the DCLMA network were considered.

In addition, an effective radar range of 110 km from the primary KLWX radar was established. The Interactive Radar Map Tool provided three separate radar ranges based on the availability of beam coverage at specified altitudes from the ground: 4,000 ft (best coverage), 6,000 ft (better coverage), and 10,000 ft (fair coverage). These radar height designations were determined by NOAA's Radar Operations Center. The 6,000 ft coverage, corresponding to an approximate 110-km radar range, provided a large enough area to accrue the desired number of thunderstorm cases, while still maintaining adequate resolution. Since the maximum range for KLWX was 230 km, 110 km was deemed an acceptable range for KLWX. It is important to note that this radar range was not maintained for all radars, but only the primary

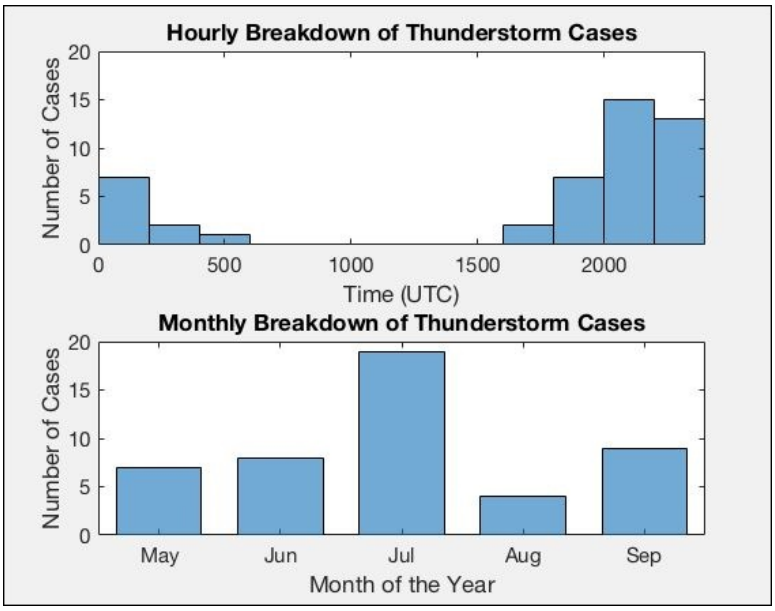

Figure 11. Hourly (top) and monthly (bottom) breakdown of all 47 thunderstorm cases.

radar KLWX. In fact, higher altitudes were desired for the alternate radars in order to identify HCA presence at the higher isothermal levels (-10°C to -20°C). Thus, the resolution was more coarse for these radars, and specifically the Level-III data.

Lastly, sufficient data and adequate data quality were key considerations during the case elimination process. Thunderstorm cases needed to have sufficient LMA data, meaning no noticeable gaps in hourly post-processed data, and at least six active LMA stations. Thunderstorm cases that didn't have any associated hourly post-processed data available were eliminated. With respect to radar data, cases with gaps in KLWX radar data due to radar outages were eliminated. For HCA data collection at higher altitudes, if sufficient data were not available through KDOX, KCCX, or KAKQ radars then the case was also eliminated.

The spatial distribution of the final 47 cases is depicted in Figure 10. There was a higher concentration of lightning cessation cases on the eastern half of the domain of study, since many cells to the west were avoided due to the impact of the

mountainous terrain on radar quality. The monthly and yearly breakdown of storms are presented in Figure 11. The majority of the lightning cessation cases occurred near the time of peak heating, or shortly thereafter. Monthly coverage was relatively evenly distributed amongst all months apart from July, which fostered the greatest number of storms.

3.3 Patton's Bootstrapped GLM

Popularized by McCullagh and Nelder (1989), GLM refers to a broad classification of models in which the response variable follows an exponential family distribution with the mean. The mean is assumed to be some, typically nonlinear, function of $x_i c_i$, where x_i comprises the known covariates and c_i comprises the coefficients to be estimated. Similarly, the bootstrapped GLM developed by JP17 utilizes a set of predictor and coefficient values in a best-fit logistic regression model for predicting lightning cessation.

The GLM was trained by extracting a 40% random sample of the 1-min interval observations from the total dataset. The 1-min observations were classified by their timing with respect to lightning cessation: before lightning cessation and after lightning cessation. The GLM was trained to associate future lightning with the radar-derived parameters consistent with the observations taken prior to lightning cessation. Additionally, the model was trained to associate no lightning with the radar parameters indicated by observations which occurred after lightning cessation. A predictive model was then calculated using the most statistically significant predictors. This process was repeated 1,000 times to pinpoint a median estimate of predictor coefficients. This use of sampling with replacement reveals the "bootstrapped" aspect of JP17's GLM.

For the observations taken prior to lightning cessation, graupel presence at one

or more isothermal levels as well as higher radar reflectivities are expected. Likewise, lack of graupel presence and lower radar reflectivities indicate observations after lightning cessation. The GLM incorporates these radar-derived parameters using a linear combination and then calculates the associated predictor coefficients that best reduce errors. This process is performed by using a least squares method for maximum likelihood estimation. The model is a form of binary logistic regression which utilizes a binomial distribution to estimate error distributions for the GLM. The "link" function, described in Agresti (2013), relates the coefficients and predictor values, listed in Table 3 and Table 2 respectively, to the lightning cessation probability given by Equation 1. Accordingly, radar parameters of statistical significance were assigned probabilistic weights, utilized to estimate the probability for lightning cessation:

$$Probability(Cessation) = \frac{exp(c_0 + c_1x_1 + c_2x_2 + c_3x_3 + ...)}{1 + exp(c_0 + c_1x_1 + c_2x_2 + c_3x_3 + ...)} \quad (1)$$

Following Equation 1, x_1, x_2, x_3,... represent radar-derived parameters deemed statistically significant by the GLM. The coefficients calculated by the GLM c_1, c_2, c_3,... are multiplied by these observed parameters.

Level	Parameter (x_i)
Composite (Maximum)	Maximum reflectivity
0°C	Maximum reflectivity
-5°C	Graupel presence (0 or 1)
-10°C	Graupel presence (0 or 1)
-15°C	Graupel presence (0 or 1)
-20°C	Graupel presence (0 or 1)

Table 2. Predictor values to be incorporated into JP17's bootstrapped GLM.

Parameter (x_i)	Coefficient Value (c_i)
Maximum composite reflectivity	-0.2472
Maximum reflectivity at 0°C	-0.0637
Graupel presence (0 or 1) at -5°C	-1.1189
Graupel presence (0 or 1) at -10°C	-0.8548
Graupel presence (0 or 1) at -15°C	-0.8072
Graupel presence (0 or 1) at -20°C	-0.9997
Intercept (c_0)	16.0826

Table 3. The six predictor coefficient values to be incorporated into JP17's bootstrapped GLM.

3.4 Lightning Cessation Algorithm Testing

Once a refined set of lightning cessation cases was collected, the time of lightning cessation, or the minute of the last recorded lightning flash, was determined for each of the 47 cases. To ensure lightning flashes detected via the LMA network corresponded to the appropriate thunderstorm cell, hourly LMA images were used as a first glance comparison tool. Figure 12 depicts a cluster of lightning flashes corresponding to an isolated cell. The lightning flash cluster location was then compared visually with radar reflectivity images to ensure the cluster aligned with the cell of interest.

Thereafter, hourly post-processed LMA data for each of the thunderstorms were ingested into MATLAB analysis software. The code utilized a set of predetermined latitude and longitude constrictions and time restraints based on the analyzed LMA images and radar reflectivity trends. Within these constrictions, the precise time and location of the last lightning flash was determined.

Each of the 47 lightning cessation cases were analyzed using the lightning cessation predictive model developed by JP17. JP17's recommended method for lightning cessation utilizes a bootstrapped GLM that incorporates six predictor values (Table 2) and their corresponding coefficient values (Table 3). Predictor values at a specific time are input into the GLM equation which generates the probability for lightning

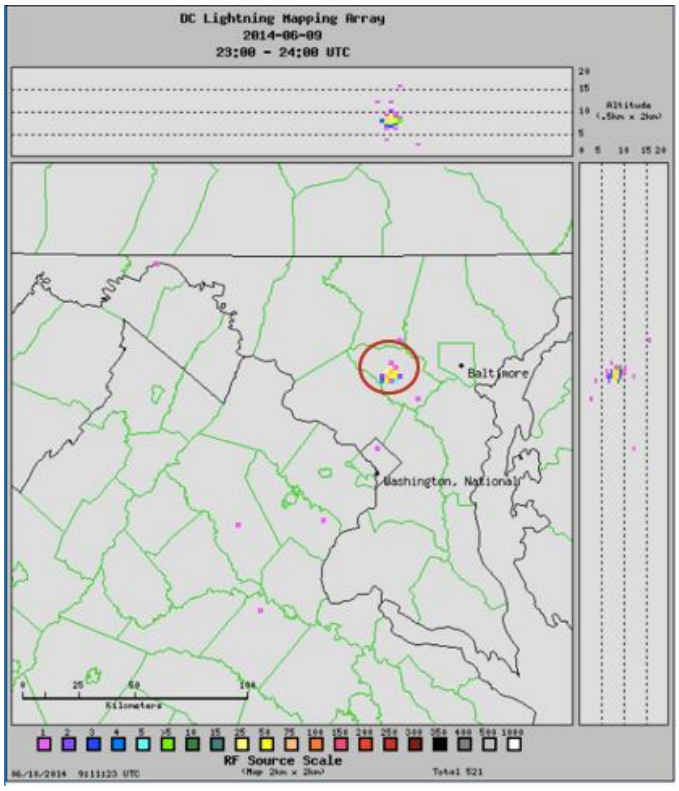

Figure 12. Cluster of lightning flashes circled in red that corresponds to a lightning cessation thunderstorm case. This is an hourly real-time 250 km image for June 9, 2014 at 2300 UTC.

cessation at that time. For this study, thunderstorm cells were analyzed from 16 min before the time of lightning cessation to 16 min following the time of lightning cessation, spanning a total of 33 min. Radar parameters were collected for each radar scan within the alloted time and grouped into 4-min bins. Tracking and analysis of radar data was done manually via GRLevelX software. Although manual analysis introduces the possibility of human error, storms were triple checked to ensure accuracy. Additionally, multiple radars were employed to ensure consistency and to fill in brief time gaps in radar data.

First, Level-III radar data were collected. GRLevel3 software was employed to analyze HCA and composite reflectivity values using a plan position indicator

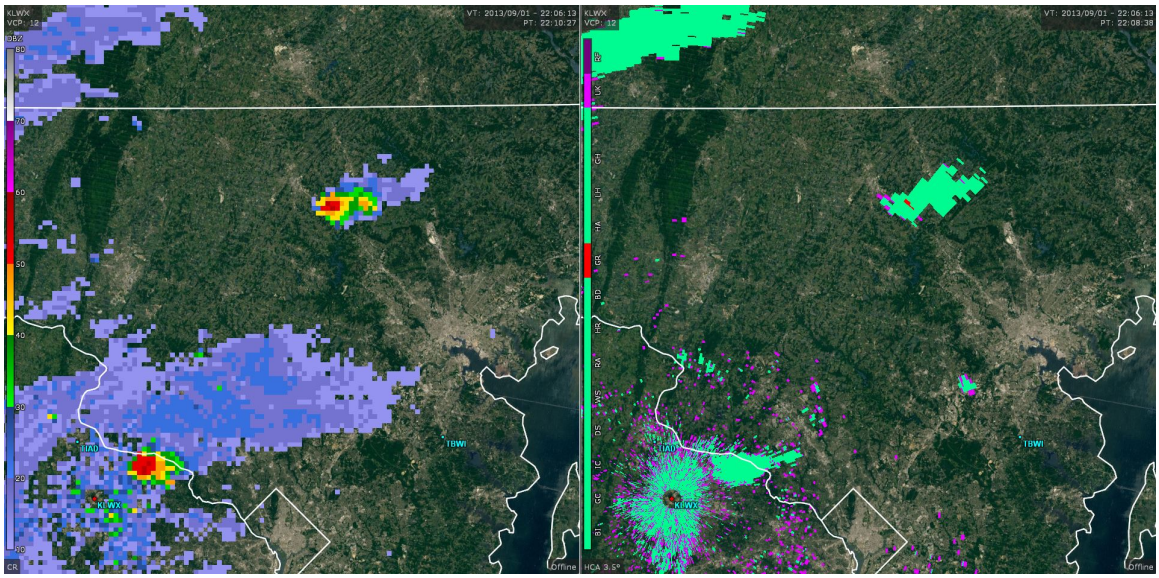

Figure 13. KLWX radar image taken at 2206 UTC, prior to lightning cessation. Composite reflectivity (left) and HCA (right) at the 3.5°elevation angle displayed using GrLevel3 2-panel view.

Figure 14. To facilitate the graupel identification process, the HCA color table was customized accordingly.

(PPI) display. GRLevel3 offers a multi-panel viewing platform that was utilized for analyzing Level-III data quickly and efficiently. Figure 13 displays an example case in which composite reflectivity and HCA at the 3.5° elevation angle were displayed at the first 4-min binned group, 16 min prior to lightning cessation. In this case, the maximum composite reflectivity for the storm was 57.5 dBZ. Next, HCA values were analyzed at each available elevation angle. The required elevation angles were determined based on the estimated thermal layers listed in Table 1. The elevation angle displayed on the panel to the right hand side of Figure 13 corresponds to the -5°C layer. Graupel presence is represented by the bright red color. Since the HCA

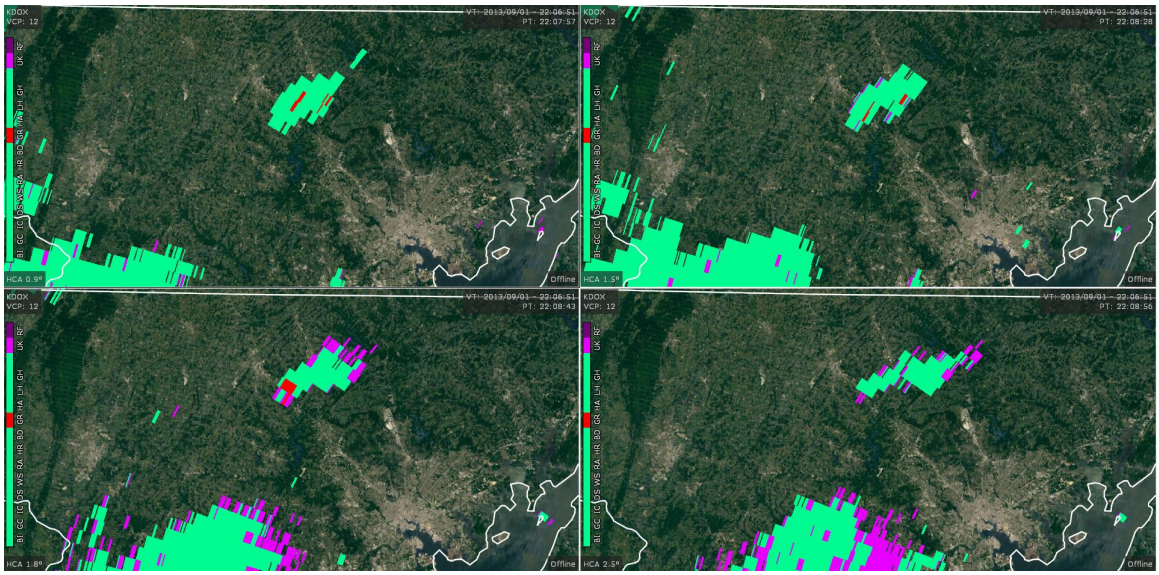

Figure 15. KDOX radar image taken at 2206 UTC, prior to lightning cessation. HCA values displayed using GrLevel3 4-panel view, each panel showing a different elevation angle: 0.9°(top left), 1.5°(top right), 1.8°(bottom left), 2.5°(bottom right).

color pallet in GRLevel3 encompasses 12 classes of hydrometeors, the HCA color table was customized according to Figure 14 to facilitate the graupel identification process. According to Figure 13, graupel presence was identified at the -5°C level. However, three remaining thermal layers must still be analyzed for graupel presence (-10°C, -15°C and -20°C). Since the elevation angles for HCA do not typically encompass the isothermal heights beyond -10°C with the archived Level-III KLWX radar data, an additional radar had to be employed to collect the remaining HCA data. In this case, KDOX was used. Figure 15 shows the remaining three elevation angles required for analysis (and the -5°C layer for double-checking purposes) using the KDOX radar.

After Level-III radar data were acquired, Level-II maximum reflectivity values at 0°C were collected. GR2Analyst software was used to visualize the data, and the cross-section feature was implemented to analyze the data. Figure 16 shows the layout of the cross-section feature with height and horizontal distance displayed on the vertical and horizontal axes, respectively. Toggle buttons in the right panel under

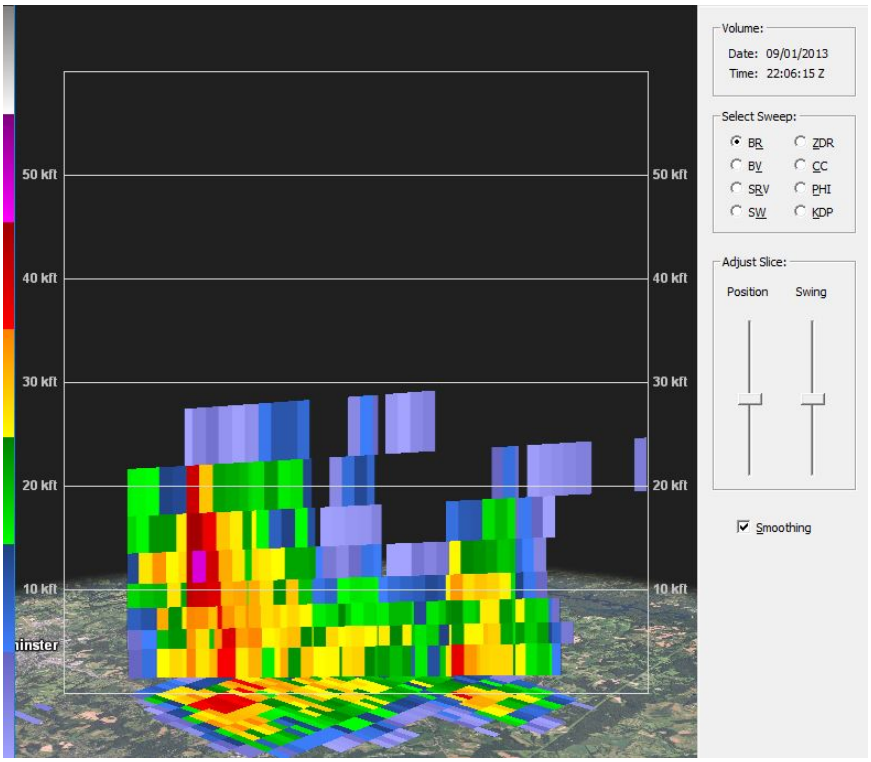

Figure 16. Example GR2Analyst radar cross-section of reflectivity of a thunderstorm cell.

"position" and "swing" were utilized as tools to scan through each thunderstorm cell and pinpoint the maximum reflectivity value at the 0°C level.

Following analysis, 4-min binned data were interpolated to 1-min intervals by adopting a cubic Hermite spline interpolation method. This method utilizes a piecewise, continuous function comprised of 3rd order polynomials. Due to it's piecewise construct, cubic spline interpolation prevents Runge's phenomenon, a manifestation of artificial oscillations in the function due to higher order polynomial interpolation. The interpolated value at a query point is based on a shape-preserving piecewise cubic interpolation of the values at neighboring grid points. Therefore, cubic spline interpolation allows for a smoother curve that is a better representation of what is observed in nature, while still maintaining its true shape. The cubic spline interpola-

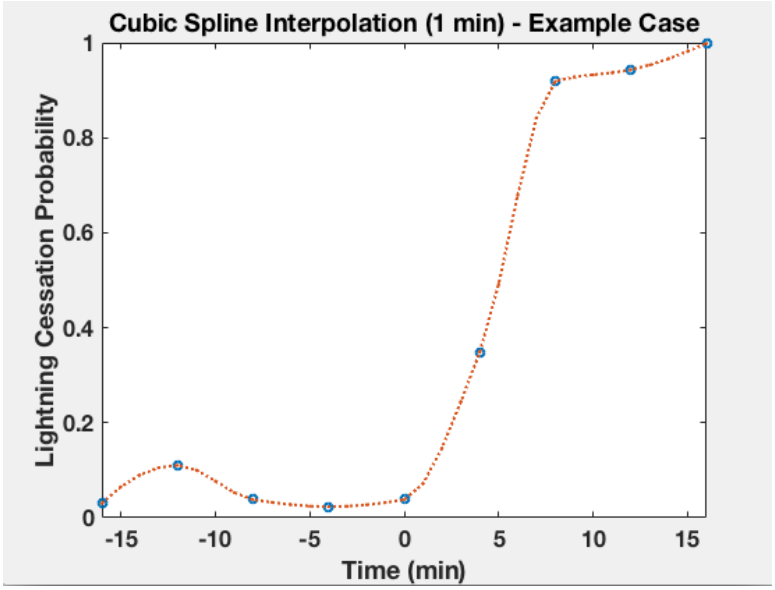

Figure 17. The probability of lightning cessation over time for an example thunderstorm case after cubic spline interpolation was applied.

tion was achieved using a MATLAB interpolation function, namely Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) (The Mathworks Inc., 2018). Figure 17 depicts the lightning cessation probability curve for an example thunderstorm case after spline interpolation.

3.5 Bootstrapping the Data

The resulting dataset was comprised of 47 cases, each with 33 independent 1-min observations; each observation corresponded to a percentage value indicating the likelihood of lightning cessation. In order to establish confidence intervals for the calculated performance statistics to effectively compare to JP17's statistics, the data were bootstrapped. Bootstrapping is a method of resampling introduced by Efron (1979) which enables one to calculate the distribution of a given statistic using a single original dataset. The resampling is performed by drawing n observations with replacement from the original dataset, which is then used to calculate the statistic

(Efron and Tibshirani, 1994; Hesterberg, 2015). This process is repeated many times, typically 10,000 times to create a sufficient sample size. These bootstrap statistics comprise the bootstrap distribution. This distribution can be utilized as an estimate to infer the characteristics of the original statistic, including sampling error, confidence intervals, and bias.

In this study, bootstrapping was performed on the lightning cessation probabilities determined for the 33 1-min interval data for all 47 cases. Specifically, this dataset was sampled with replacement 10,000 times, with corresponding performance statistics calculated for each sample. Next, the percentile method (Efron and Tibshirani, 1994) was applied to determine confidence regions for the calculated statistics. The percentile method was deemed an acceptable method for calculating confidence intervals for this study considering that the sample size after bootstrapping was sufficiently large (10,000 samples) (Wilks, 2011). The 95th percentile of the calculated statistics were determined, and the associated error bars were plotted with the statistics from the original dataset. These values were overlaid with JP17's central Florida statistics for comparison.

3.6 Forecast Metrics

The interpolated 1-min interval data from the 47 cases were used to calculate forecast metrics using a set of skill score statistics. The performance results were calculated using two approaches. First, each 1-min interval was treated as an independent observation, which corresponded to a probability value for lightning cessation. Second, each thunderstorm case was treated as a singular event for which the model was trying to correctly predict lightning cessation. Probability thresholds were set to indicate that above these threshold values, the model predicted lightning cessation. Below these probability thresholds, lightning was predicted to still be ongoing.

Three threshold probabilities were selected to match that of JP17: 95.0%, 97.5%, and 99.0%. These thresholds allow the model to be tested, and the performance metrics to be determined. Furthermore, these particularly high probability thresholds were selected to emphasize safety as a crucial priority when predicting lightning cessation.

For the first approach, every minute observation was categorized as either a hit, miss, false alarm, or correct null in accordance with Table 4. A false alarm corresponds to the prediction of lightning cessation when lightning is still ongoing. This is the most dangerous case in terms of safety. A hit refers to an observation/minute which indicates lightning cessation is correctly predicted. In contrast, a miss refers to the event in which lightning cessation is not forecast, but lightning has ended. Finally, the correct null refers to an observation that indicates lightning is ongoing, and the model correctly forbears the prediction of lightning cessation.

For the second approach, each storm was considered an individual event for which the model was attempting to correctly predict lightning cessation. The time that the model first predicted lightning cessation for each case was recorded. This represents the time a lightning advisory would be canceled if used real-time. A total of 47 lag-times after lightning cessation were determined, and the median lag-time was calculated. A storm with a lag-time after observed lightning cessation

		Was cessation observed at this minute?	
		Yes	No
Was cessation forecast at this minute?	Yes	Hit	False Alarm
	No	Miss	Correct Null

Table 4. Contingency table from JP17 that lists the four possible outcomes based on the connection between the model forecast and the observation ((Jolliffe and Stephenson, 2012; Wilks, 2011).

between 1 and 15 minutes was designated as a hit. A storm with a lag-time after observed lightning cessation beyond 15 minutes was designated as a miss. Lastly, a false alarm was characterized as a storm for which the model predicted lightning cessation prematurely. The correct null does not exist for this approach since all storms eventually experience lightning cessation. Thus, only verification statistics that do not incorporate the correct null were calculated for this approach.

Once the forecast metrics for the data were determined, a set of verification measures defined by Equations 2-7 were calculated. These measures were used to determine the skill of the model. Skill in this context refers to the ability of the model to correctly forecast the occurrence or non-occurrence of lightning cessation, more often than what would be expected by chance (Jolliffe and Stephenson, 2012). The Probability of Detection (POD), or hit rate, refers to the proportion of lightning cessation events that were correctly forecasted (Donaldson et al., 1975; Jolliffe and Stephenson, 2012) and is defined as:

$$Probability\ of\ Detection = POD = \frac{Hits}{Hits + Misses} \tag{2}$$

With regards to the POD, a value close to 1.0 is desired and conveys that there were minimal missed forecasts in relation to hit forecasts. This means the model is correctly detecting and predicting lightning cessation. It is important to note that the POD does not take into account the false alarms and should therefore not be used as the sole method for determining the skill of a forecast.

The False Alarm Ratio (FAR) refers to the probability of a false alarm given that an event was forecast (Donaldson et al., 1975; Jolliffe and Stephenson, 2012). In terms of lightning cessation, the FAR corresponds to the probability of forecasting

lightning cessation when lightning is still ongoing. The FAR is defined as:

$$False\ Alarm\ Ratio = FAR = \frac{False\ Alarms}{Hits + False\ Alarms} \tag{3}$$

FAR values range from 0.0 to 1.0. The ideal FAR is 0.0, signifying that the number of false alarms is limited. Similar to the POD, the FAR should not be used alone to measure the skill of a forecast due to the strong dependence on the number of hits. In this case, rare events will score higher than random forecasts of common events, and thus is deemed inequitable. In this study an equitable measure is one that gives all random forecasting systems the same score (Gandin and Murphy, 1992). This provides a no-skill baseline against which a forecaster can be examined in contrast to have skill. While the use of FAR alone is inequitable, using FAR and POD conjointly can provide a more accurate measure of the skill of the model; a perfect skill score would have a FAR of 0.0 and a POD of 1.0.

The Critical Success Index (CSI) provides a sample estimate of the probability of a hit occurring given that an event is either forecasted, observed, or both (Donaldson et al., 1975; Jolliffe and Stephenson, 2012). CSI is defined as:

$$Critical\ Success\ Index = CSI = \frac{Hits}{Hits + Misses + False\ Alarms} \tag{4}$$

Values range from 0.0 to 1.0 and a 1.0 CSI would indicate perfect skill, while a 0.0 would indicate no skill. Owing to the fact that this measure does not depend on the number of correct rejections, it presents an inequitable skill assessment. Rare events appear favored, since a non-occurrence of a rare event is easier to forecast, while common events are at a disadvantage. CSI will still be calculated, since JP17 calculated this measure, and the goal is to compare all statistics of the model's performance in Washington, D.C. to those of central Florida. However, the measure will be weighted

less than the following two, preferred verification measures: Heidke Skill Score (HSS) and True Skill Statistic (TSS).

HSS and TSS are verification measures that provide useful stand-alone performance statistics of the dataset. The HSS is a measure of the fractional improvement of the forecast over the standard forecast, chance (Jolliffe and Stephenson, 2012; Murphy and Daan, 1985). This measure utilizes a performance variable, namely the Expected Correct (EXPCOR), that incorporates all possible forecast metrics as well as the total number of events. EXPCOR is defined as:

$$\begin{aligned} Expected\ Correct = EXPCOR \\ &= \frac{(Hits + Misses) * (Hits + False\ Alarms)}{Total\ Events} \\ &+ \frac{(Correct\ Nulls + Misses) * (Correct\ Nulls + False\ Alarms)}{Total\ Events} \end{aligned} \tag{5}$$

EXPCOR represents the number of forecasts expected to verify based on chance. HSS is defined as:

$$Heidke\ Skill\ Score = HSS = \frac{(Hits + Correct\ Nulls) - EXPCOR}{Total\ Events - EXPCOR} \tag{6}$$

The sum of hits and correct nulls in the numerator represents the number of times that the forecast matches the actual observation. HSS values range from $-\infty$ to 1.0; a perfect forecast would obtain a score of 1.0. Alternatively, random forecasts would be awarded a score of 0.0 and negative values indicate that the random forecast is better.

The TSS, also known as the Peirce Skill Score, is a verification measure that takes all event outcomes outlined in Table 4 into consideration (Flueck, 1987; Jolliffe and Stephenson, 2012; Murphy and Daan, 1985; Peirce, 1884). The TSS is defined

as:

$$True\ Skill\ Statistic = TSS = \frac{(Hits * Correct\ Nulls) - (False\ Alarms * Misses)}{(Hits + Misses) * (False\ Alarms + Correct\ Nulls)} \tag{7}$$

TSS values range from -1.0 to 1.0 with a desired value of 1.0. A value of 1.0 indicates perfect skill, while a value of 0.0 indicates no skill. A value of -1.0 indicates perfect skill, but incorrect calibration. For rare events, the number of correct nulls is large and TSS is weighted accordingly. Thus, this score may be more useful for more frequent events. Nevertheless, both TSS and HSS are considered truly equitable, and will output an expected score of zero for both random and constant forecasts. As equitable, stand-alone measures these statistics are the most valuable. Thus, they will be emphasized in the following chapter.

IV. Results

Overview

A database consisting of all the minutes from the 47 thunderstorm cases were compiled. Two approaches were used to evaluate the performance results of JP17's bootstrapped GLM in the Washington, D.C. area. For the first approach, all 1,551 minutes of data were ingested, a forecast outcome was determined for each minute/observation, and then the corresponding performance statistics for all the minutes were calculated. The second approach managed the data on a storm by storm basis, whereby the model was trying to successfully predict lightning cessation for each case. For this approach, a forecast outcome was determined for each case, and then the corresponding performance statistics for all cases were calculated. For each approach, the model's performance results in the Washington, D.C. area were examined first, then these results were compared to JP17's model performance results in central Florida.

4.1 Minute by Minute Performance Results

A final dataset of 1,551 min was compiled, consisting of all the minute data from each of the 47 cases from the 16 min before lightning cessation through the 16 min following lightning cessation. Each minute/observation corresponds to a probability value for lightning cessation. Results test how the method performed by treating each 1-min interval as an independent observation. Similar to JP17, the 95.0%, 97.5%, and 99.0% probability thresholds were tested.

Hits, misses, false alarms, and correct nulls were determined in accordance with Table 4. It is important to note that with respect to lightning cessation these terms should not be interpreted in the same manner as traditional forecast metrics. This

	Hits	False Alarms	Misses	Correct Nulls
95.0%	247	6	505	793
97.5%	186	2	566	797
99.0%	126	1	626	798

Table 5. Minute by minute forecast verification metrics for the four possible outcomes using JP17's lightning cessation GLM in Washington, D.C.

is primarily due to the nature of the forecast; the goal is to correctly predict the end of a weather phenomenon, not its onset. This point is best demonstrated by the definition of a false alarm in terms of lightning cessation. A false alarm represents the prediction of lightning cessation when lightning is still ongoing. Indeed, this is the most dangerous outcome in terms of safety.

The forecast metrics for the Washington, D.C. area are listed in Table 5. By examining the distribution of the metrics, the greatest weight for all three probability thresholds was in the correct nulls. In Washington, D.C., the GLM performed well correctly delaying the prediction of lightning cessation before the event actually occurred. This also contributed to the low number of false alarms; the false alarms only comprise .40%, .13%, and .06% of the total metrics for the 95.0%, 97.5%, and 99.0% thresholds, respectively. This is encouraging since false alarms pose the greatest threat to public safety and resource protection.

The second most prevalent metric was miss. If the model did not forecast lightning cessation, yet lightning cessation had indeed occurred, it was recorded as a miss. Although this is not a dangerous result, it is not desired since it impedes operations by imposing unnecessary and costly delays. Furthermore, the number of misses is relatively high in relation to the number of hits. There are roughly twice as many misses than hits. Simply looking at the ratio of misses to hits gives the impression that the model forecasted lightning cessation incorrectly more often than correctly. However, the large number of correct nulls, which is a favorable forecast

outcome, should also be accounted for. Additionally, in the context of how this study was executed, the misses to hits ratio can be misleading.

During analysis, a time constraint was applied which limited the observation time to 33 min. This means that observations were not necessarily taken until the storm dissipated and lightning cessation was guaranteed. In consequence, the disproportion of misses to hits could be due to the fact that the model was not tested long enough for all the possible number of hits to be identified. This disproportion of metrics will also be self-evident in a few of the performance statistics for Washington, D.C.

Once the forecast metrics for the data were determined, statistical performance measures defined by Equations 2-7 and presented in Table 6 were calculated. These measures were used to determine the skill of the GLM in the Washington, D.C. area. The low FAR is congruent with the low number of false alarms in the dataset. The FARs for each of the three probability thresholds fall below 2.5%. A lower value is ideal since it indicates that a very small number of observations using the GLM forecasted lightning cessation too early. As discussed in section 3.6, the FAR and POD are best interpreted together. The POD values range from $\sim$ .10 to $\sim$ .30 (Table 6); these values are rather low since the desired POD score is a value near 1.0. Analyzing the POD and FAR conjointly gives conflicting results in terms of model performance. While the model correctly refrains from prematurely forecasting lightning cessation, it is also slow to identify lightning cessation when it is occurring.

DC Statistics	POD	FAR	HSS	TSS	CSI
95.0%	0.3285	0.0237	0.6422	0.3209	0.3259
97.5%	0.2473	0.0106	0.6109	0.2448	0.2467
99.0%	0.1676	0.0079	0.5790	0.1663	0.1673

Table 6. Minute by minute performance statistics for the Washington, D.C. area using JP17's lightning cessation GLM.

Thus, the number of hits in relation to misses is low.

The TSS and CSI reflect values similar to the POD, which has an approximate range from $\sim$ .10 to $\sim$ .30 (Table 6). A perfect score for both TSS and CSI would be 1.0. Thus, the calculated TSS and CSI values indicate that the model has relatively low skill. In contrast, the HSS indicates satisfactory performance with values ranging from $\sim$ 0.55 to $\sim$ .65. As mentioned in section 3.6, the HSS combines all forecast metrics in Table 5 to measure the fractional improvement of the forecast over the standard forecast. Thus, these HSS values mean that the model is performing much better than the standard forecast, or chance. However, it is important to note that these performance results do not necessarily imply that the model will perform in the same manner in real-time.

Next, the verification measures of the model's performance in the Washington, D.C. area were compared to the model's performance in central Florida. Table 7 displays the performance measures calculated in JP17 for central Florida. Figures 18 through 19 illustrate the similarities and differences in the GLM's performance statistics in the Washington, D.C. area versus central Florida. The error bars for each performance measure were calculated by bootstrapping the data using 10,000 resamples over a 95% confidence interval. These error bars reveal the variability, or range of values, of each performance statistic for the Washington, D.C. dataset. JP17's results for central Florida are overlaid, showing the relationship of the performance statistics for the two locations. If the performance statistics for central

FL Statistics	POD	FAR	HSS	TSS	CSI
95.0%	0.6989	0.0044	0.6650	0.6949	0.6968
97.5%	0.6318	0.0013	0.5976	0.6307	0.6312
99.0%	0.5451	0.0005	0.5098	0.5448	0.5448

Table 7. Minute by minute performance statistics from JP17 for central Florida using the lightning cessation GLM.

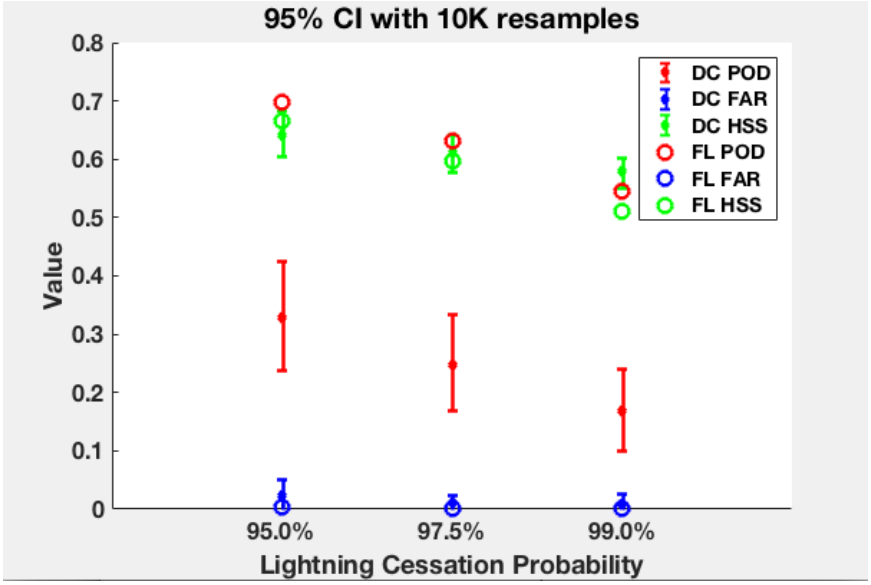

Figure 18. POD, FAR, and HSS performance statistics for JP17's GLM in both Washington, D.C. and central Florida. Washington, D.C. performance statistics and bootstrapped results utilizing the 95th percentile are depicted by the diamonds and error bars, respectively. Central Florida performance statistics are overlaid in circles.

Florida are contained within the Washington, D.C. error bars for the corresponding performance statistic, then it can be concluded with confidence that the performance statistics are statistically similar. Likewise, if the performance statistics for central Florida are not contained within the Washington, D.C. error bars then it can be concluded that the performance statistics are statistically different.

Figure 18 illustrates the comparison of POD, FAR, and HSS values for the three probability thresholds. The FARs, which are a measure of the reliability of the model, are relatively low for both locations and all probability thresholds. The central Florida FARs are contained within the Washington, D.C. FAR error bar ranges, thus they are statistically similar. In contrast, the POD values for both locations vary significantly. In fact, the Florida POD values double those of Washington, D.C. This contrast can possibly be construed by a key difference in methodologies. Theoretically, there should be a minimum of one hit for each probability threshold for each case (at

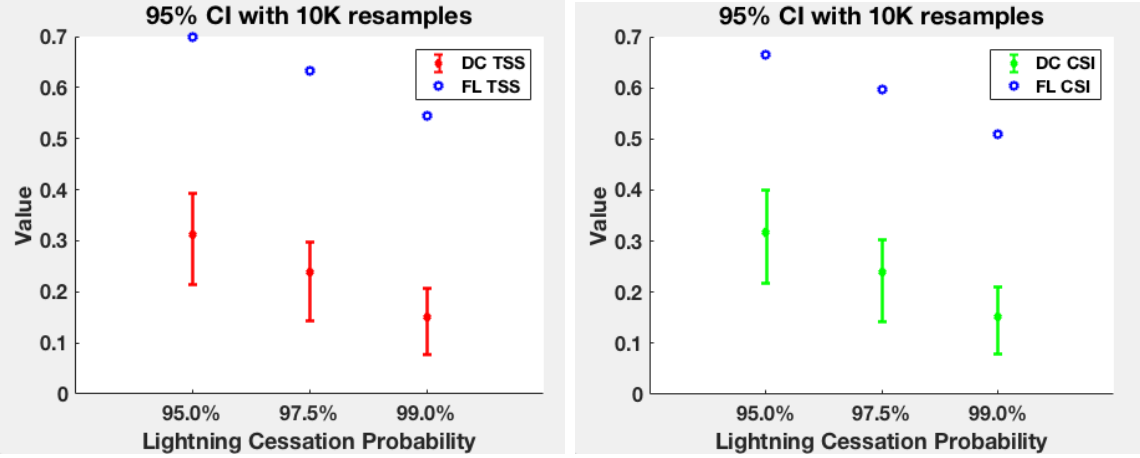

Figure 19. (a) TSS (left) and (b) CSI (right) performance statistics for JP17's GLM in Washington, D.C. and central Florida. Washington, D.C. performance statistics and bootstrapped results using the 95th percentile are depicted by the diamonds and error bars, respectively. Central Florida performance statistics are overlaid in blue circles.

least 47 hits per probability threshold). This is true because, for all lightning events, lightning cessation will eventually take place. However, in this study, time restrictions were set, limiting the analysis observation time to 33 min, 16 min before and after lightning cessation. JP17 did not utilize time restrictions in his study. He observed each thunderstorm case until the storm dissipated (or composite reflectivity dropped below 25 dBZ) and the model forecasted lightning cessation. His higher values for POD (Figure 18), TSS (Figure 19a), and CSI (Figure 19b) reflect this. This difference in methodology will be accounted for in section 4.2 with the alternative performance evaluation approach.

As discussed in section 3.6, the HSS and TSS are both truly equitable measures; therefore, they are more robust indicators of skill than CSI, POD, and FAR used as stand-alone measures. One distinction between the TSS and HSS is that for models with positive skill, TSS has a proclivity to treat overpredicting models more generously than HSS, and underpredicting models more harshly. In this study, this distinction is evidenced by the significant contrast between the TSS and HSS (Tables 6 and 7). Indeed, the model in the Washington, D.C. area tends to underpredict

lightning cessation as indicated by the disproportionate ratio between misses and hits in Table 5. This could explain why the HSS values are higher and more congruent with JP17's Florida results, while the TSS values are much lower in comparison.

4.2 Storm by Storm Performance Results

Similar to JP17's performance evaluation strategy, a second method was utilized to evaluate the model's performance in the Washington, D.C. area by treating each storm as an individual event. For this approach, the model was trying to correctly predict the time of lightning cessation for each case. The objective was to treat each thunderstorm case as if it were occurring in real-time. This approach identified the time when the model first exceeded the designated probability value and predicted lightning cessation, if at all.

Figure 20 presents a conceptual timeline of a thunderstorm case from storm initiation to storm dissipation. Viewing from left to right, the 0 minute highlighted by the first red line represents the time of observed lightning cessation. The second red dashed line indicates the pre-established 45 WS wait time after lightning cessation. This timeline helps to demonstrate the designation of possible forecast outcomes for each thunderstorm case, since they differ from that of the first approach (see section 3.6).

Each thunderstorm case was defined by a single forecast outcome: hit, miss, or false alarm. Hits, misses, and false alarms were all calculated based on strict time limits with respect to the observed time of lightning cessation. The objective was to correctly predict cessation within 1 to 15 min after lightning cessation occurrence. This method tested the model's performance against the 45 WS's presently used wait time of 15 min. The correct null does not exist for this approach since lightning cessation is observed for every case. Thus, there is no situation that yields a case

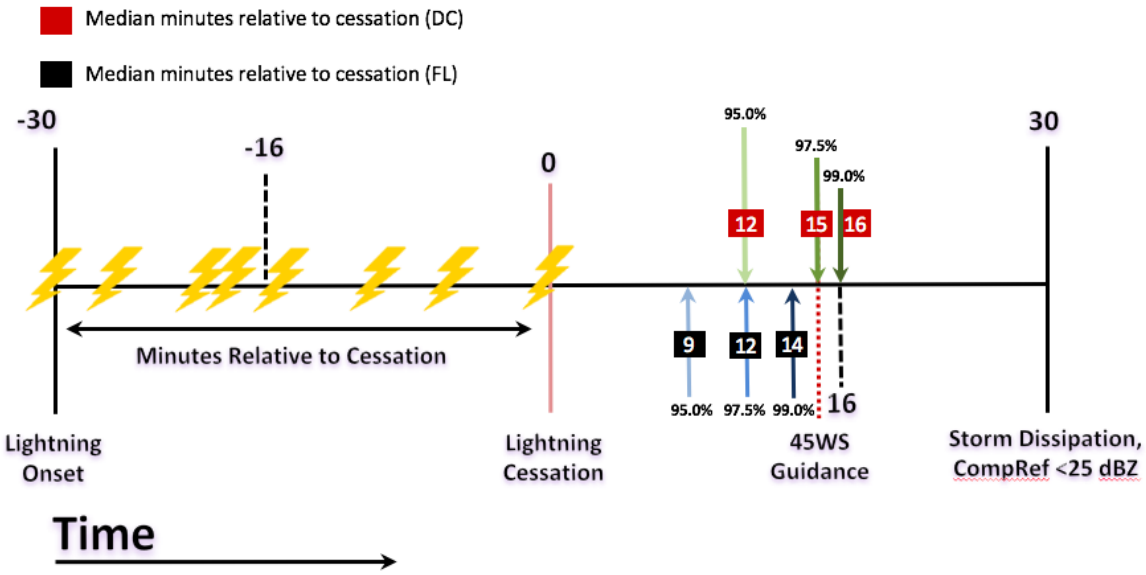

Figure 20. Timeline depicting the conceptual 60-minute lifetime of an isolated thunderstorm from lightning onset to storm dissipation (adapted from JP17). The solid red line at 0 minutes corresponds to the observed time of lightning cessation. The 16-minute markers prior to and following lightning cessation correspond to the time limits set for data analysis in this study. The green and blue arrows represent the three probability thresholds from Patton's bootstrapped GLM for Washington D.C. and central Florida, respectively. Similarly, the minutes in red and black blocks represent the median time relative to lightning cessation for Washington D.C. and central Florida, respectively.

outcome which correctly predicts the non-occurrence of lightning cessation. The Washington, D.C. metrics and performance measures for this approach are listed in Table 8 and Table 9, respectively.

By setting the 15-min time constraint for hit and miss designations, the storm by storm approach accounted for the discrepancy of POD and CSI values noted

	Hits	False Alarms	Misses
95.0%	25	3	19
97.5%	24	2	21
99.0%	17	1	24

Table 8. Storm by storm metrics for the three possible outcomes using JP17's GLM in Washington, D.C.

DC Statistics	POD	FAR	CSI
95.0%	0.5682	0.1071	0.5319
97.5%	0.5333	0.0769	0.5106
99.0%	0.4146	0.0556	0.4048

Table 9. Storm by storm performance statistics for the Washington, D.C. area using JP17's lightning cessation GLM.

in section 4.1. The Washington, D.C. performance statistics presented in Table 9 are significantly higher than those determined using the first evaluation approach (Table 6). In fact, POD and CSI values are almost double the performance statistics calculated by using the first approach. For this approach, the applied time constraint has no impact on the forecast metrics. Thus, the difference in CSI and POD values can be attributed in part to this characteristic.

With regard to the FAR (Table 9), the values are significantly higher than the FARs listed in Table 6 using the first approach. The FARs for the three probability thresholds range from $\sim 5.0\%$ to $\sim 11.0\%$ for the Washington, D.C. area. Although there is only a small percentage of total minutes in the dataset that are associated with false alarms, these minutes are distributed among several storms that performed worse as a whole. This has serious implications with respect to real-time use; even predicting lightning cessation a minute early can have dangerous safety repercussions.

Table 10 lists JP17's performance statistics for central Florida for comparison with the corresponding statistics for the Washington, D.C. area. Although there is an approximate .25 difference between the POD values for the 95.0% probability threshold for central Florida versus Washington, D.C., the values become more consistent

FL Statistics	POD	FAR	CSI
95.0%	0.8022	0.0135	0.7935
97.5%	0.7033	0.0154	0.6957
99.0%	0.5519	0.0098	0.5489

Table 10. Storm by storm performance statistics for JP17 in central Florida using the lightning cessation GLM.

with increasing probability thresholds (Table 9 and 10). There is a more significant drop-off of POD values from the 95.0% probability threshold to the 99.0% probability threshold for central Florida compared to Washington, D.C. Meanwhile, the FARs for central Florida remain relatively consistent for all three probability thresholds, while there is a significant drop-off in FAR from the 95.0% threshold to the 99.0% thresholds in Washington, D.C. Therefore, for these performance measures the Washington, D.C. results corresponding to the more conservative probability thresholds, 97.5% and 99.0%, are more consistent with those of central Florida. This observation is evidenced by Figure 21; both POD values and FAR values are closer to the 95% confidence interval for the 97.5% and 99.0% probability thresholds compared to the 95.0% threshold.

According to Figure 21, the FARs for central Florida do fall within the 95% confidence interval calculated for the Washington, D.C. area. However, the POD values do not fall within the Washington, D.C. 95% confidence interval. Thus, the FAR values are statistically similar, while the POD values are statistically different. Similar to the results from first approach, these are conflicting outcomes, which expose the model discrepancies in the Washington, D.C. area. Furthermore, it should be noted that the calculated error bars are much longer for this approach due to smaller sample size (47 outcomes versus 1,551 for the first method). These longer error bars indicate lower reliability.

Lastly, the model-predicted lag-time after lightning cessation was determined for the three probability thresholds for each of the 47 cases. These lag-times are depicted in Figure 22. The horizontal red line represents the time that lightning cessation was observed, and the horizontal dashed black line represents the presently used 45 WS wait time. It is important to note that the thunderstorm cases for which the model did not predict lightning cessation within the 33-min observation timespan are not

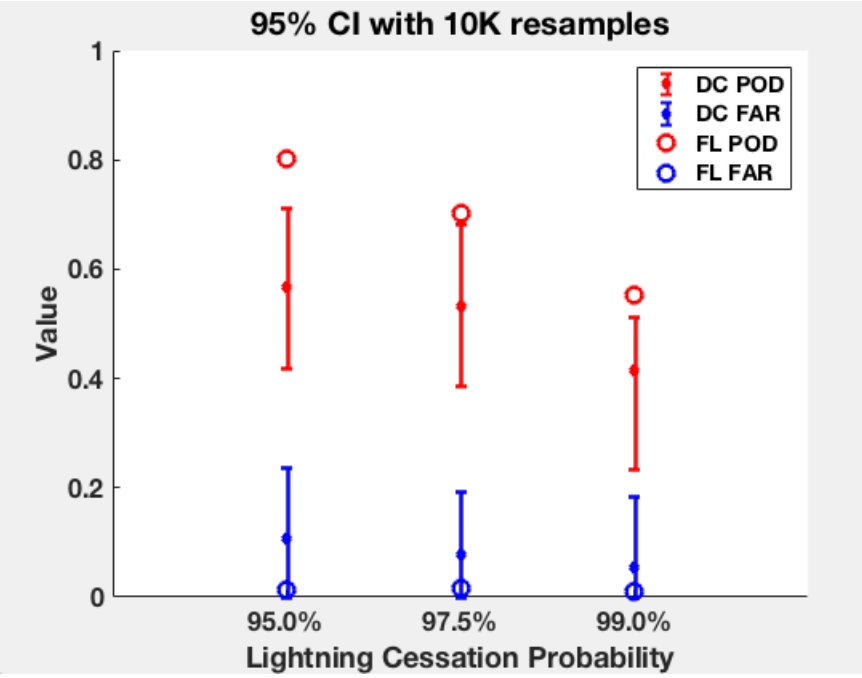

Figure 21. Storm by storm POD and FAR performance statistics for JP17's GLM in Washington, D.C. and central Florida. Washington, D.C. POD and FAR bootstrapped results utilizing the 95th percentile are depicted by the diamonds and error bars, respectively. Corresponding performance statistics for Central Florida are overlaid in circles.

displayed here. Thus, there are no results displayed above the 16 min mark. However, this does not skew the results. The objective was to test the model's performance with the desired result to have the model-predicted lightning cessation time fall between the 0 and 15-min timespan. Indeed, most of the predicted lightning cessation times exist between these two limits, with few events situated on or below the red line. The 95.0% probability threshold exhibits the most false alarms, yet these false alarms occurred just minutes shy of the observed time of lightning cessation.

The median lag-times were calculated to compare to the median lag-times determined in JP17's study for central Florida. Furthermore, 10,000 resamples of the dataset were achieved to determine the 95th percentile range of values. The bootstrapped results are illustrated by the error bars in Figure 23 and overlaid with JP17's

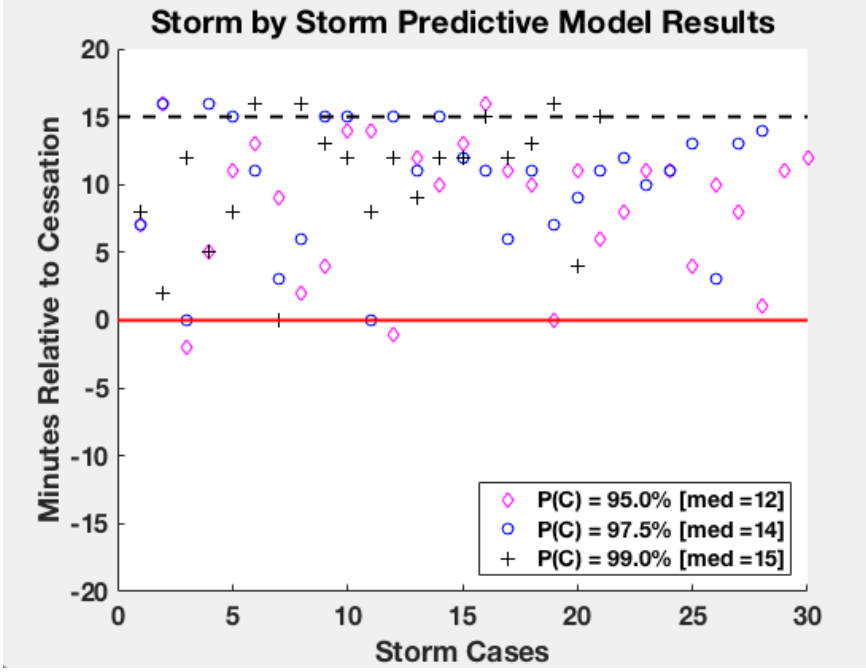

Figure 22. Storm by storm lag-times for thunderstorm cases in Washington, D.C. Results for all three probability thresholds of JP17's GLM are displayed. Markers beneath the solid red line indicate false alarms, and markers above the dashed black line indicate that the model waited too long to predict lightning cessation (beyond the 45 WS wait time of 15 min). The total 47 cases are not displayed because the observation time was constrained to 16 min before and after lightning cessation.

median values for central Florida. The median lag-times for Washington, D.C. are 12, 15, and 16 min for the 95.0%, 97.5%, and 99.0% probability thresholds, respectively. The cases for which the model did not predict lightning cessation within the 33-min observation window were still factored into the median lag-time calculation. These lag-time values were designated as an arbitrary number greater than 16 min. Overall, the resulting median lag-times for the lightning cessation model in the Washington, D.C. area are longer than the median lag-times in central Florida for all three probability thresholds. The error bars indicate that there is also significant variability in the median times based on the small sample size. Furthermore, both the 97.5% and 99.0% thresholds surpass the currently established 45 WS wait time. Only the 95.0% threshold outperforms the 15-min wait time, with a median lag-time of 12 min.

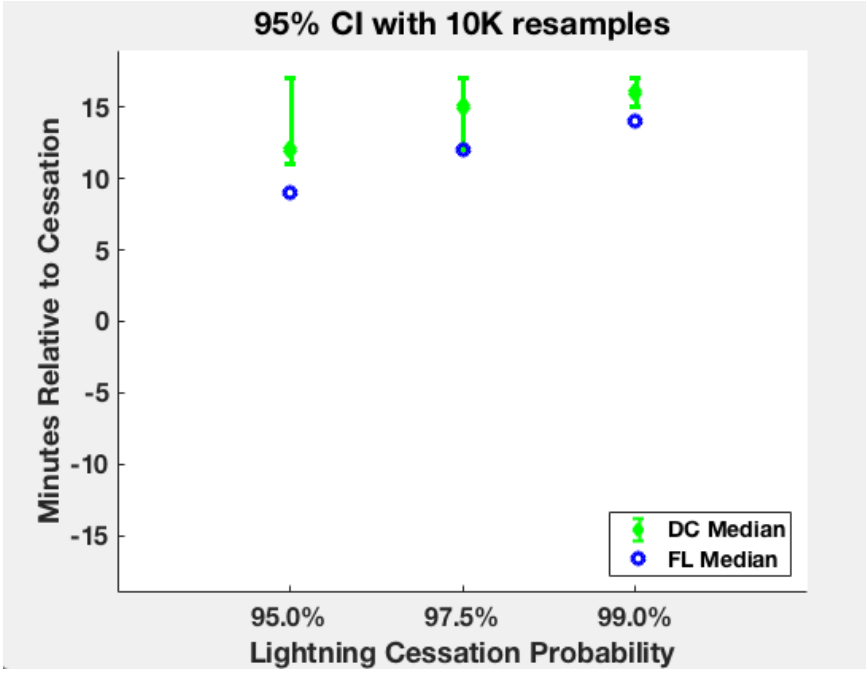

Figure 23. The median lag-time results for JP17's GLM in Washington, D.C. and central Florida. Washington, D.C. bootstrapped lag-times utilizing the 95th percentile are depicted by the green error bars. Central Florida median lag-times are overlaid in blue circles.

However, the increased FAR associated with the 95.0% probability threshold must also be taken into consideration.

V. Conclusions

5.1 Summary of Results

Probabilistic lightning cessation guidance using a bootstrapped GLM was recently developed by JP17 for the 45 WS's use at CCAFS/KSC. To test the model's overall consistency and versatility in other geographical locations, the GLM was tested in and around the greater metropolitan Washington, D.C. area. Similar to JP17's study, a three-dimensional lightning mapping system and dual-polarization radars were employed to test the model. A dataset of 47 isolated thunderstorms was collected following a similar selection process to that of JP17's study.

The bootstrapped GLM utilizes six key dual-polarization radar predictor values: graupel presence at the -5°C, -10°C, -15°C, and -20°C isothermal levels, maximum reflectivity at the 0°C level, and maximum composite reflectivity. These predictor values give insight into the mechanisms that sustain lightning generation and the physical properties indicative of lightning cessation. The model was tested by employing JP17's GLM for each thunderstorm case. The results were evaluated using a set of performance measures identical to that of JP17. Two approaches were used to verify the model's performance. The first approach considered every minute as an individual event with a singular outcome, while the second approach considered each storm as an individual event with a singular outcome. Median lag-times after lightning cessation were also determined for each probability threshold to compare with those calculated by JP17.

Comparison of the performance statistics for the GLM in the Washington, D.C. area and central Florida reveals that overall JP17's GLM did not perform as well in the Washington, D.C. area as it did in central Florida. The TSS values for all three probability thresholds for the Washington, D.C. area were approximately half

those of central Florida. Additionally, for both verification approaches, POD and CSI values for all probability thresholds were significantly lower for the Washington, D.C. area. Bootstrapped results established a 95th percentile range of values depicting the performance statistics' variability. Results show that POD, TSS, and CSI values were not contained within these value ranges, confirming the model discrepancies in the two geographically separated areas (Figures 18, 19, and 21).

Nevertheless, there are a few promising results that can be extracted from this study. The consistency between the HSS values in Washington, D.C. and central Florida (Tables 6 and 7) indicates that the model does convey skill, and is more accurate than the standard forecast. Furthermore, the low FARs indicated by the model's performance in Washington, D.C. reveal that the model tends to evade the most dangerous outcome by waiting until lightning ceases before predicting lightning cessation. The 95th percentile ranges from the bootstrapping validated these model consistencies for the FARs for both approaches and HSS (Figure 18).

These findings indicate that key radar parameters utilized by the model were statistically significant for predicting lightning cessation for both locations. Specifically, graupel presence within the MP region plays a vital role in the lightning electrification processes according to NIC theory. Also, the use of isothermal temperatures levels as opposed to height levels are excellent coordinates for identifying radar parameters, since they have consistent properties regardless of geographical location. This corroborates the results of JP17 and Preston and Fuelberg (2015) which identified graupel presence at specific isothermal level(s) as statistically significant. During analysis, graupel presence was almost always observed throughout the MP region during active lightning. Likewise, there was a noticeable trend in the rapid reduction of graupel presence promptly after observed lightning cessation. This decline in graupel presence started at the upper levels, and percolated to the lower levels with

the successive minutes after cessation. This pattern complemented JP17's lightning cessation model well. However, the model still tended to delay the forecast of lightning cessation for too long. This can be attributed to the reflectivity trends of the Washington, D.C. storms. The storm reflectivity values tended to be too high for too long, causing the model to unnecessarily delay the prediction of lightning cessation.

Median lag-times after cessation were also evaluated for the Washington, D.C. area. The analyzed median lag-times for the three probability thresholds in the Washington, D.C. area were longer than the lag-times in central Florida. The 95.0% threshold was the most promising result, with a 12-minute lag-time after cessation compared to the 9-minute lag-time for central Florida. Though this result is significant, the FAR for the 95.0% threshold must also be considered, since it displayed the highest FAR of all the probability thresholds: 2.4% and 10.7% using the first and second approach, respectively. However, the more conservative probability thresholds, 97.5% and 99.0%, convey no significant wait time improvement from the 45 WS presently used 15-minute wait time. Yet, these more conservative thresholds could provide some utility for the 45 WS by giving the forecaster greater confidence that the lightning has truly ceased.

Although the performance results for JP17's bootstrapped GLM were highly favorable for central Florida, the results for the Washington, D.C. area did not produce the same level of success. Discrepancies in the model's performance can likely be attributed to key climate differences for the two domains of study as well as methodology incongruities. With regard to climate, the difference in forcing mechanisms and the disproportionate concentration of aerosols in the environment convey key climate distinctions for the two locations. According to the 2010 Census (United States Census Bureau, 2010), Washington D.C. had a population density of 9,856 people per square mile and ranked among the top ten most populous metropolitan statistical

areas (Mackun and Wilson, 2011). This contrast in population density compared to central Florida suggests that there is a higher degree of anthropogenic aerosol emissions in Washington, D.C. compared to central Florida. This exposes the Washington, D.C. atmosphere to a different atmospheric particle composition capable of interacting differently with hydrometeors within the MP region. Consequently, more aerosol concentrations in the atmosphere could serve as cloud condensation nuclei for more ice crystal formation. This consequence could ultimately alter the electrification mechanisms within the cloud and even the frequency of lightning occurrence. However, these are currently conjectures that need to be analyzed and tested further for legitimacy.

Furthermore, the difference in forcing mechanisms prevalent in the Washington, D.C. area compared to those prevalent in central Florida has significant impacts on lightning generation and cessation. Although both locations are situated along the coast, the latitudinal differences create different atmospheric environments for storm development. While central Florida is characterized as a more barotropic environment, the Washington, D.C. environment exhibits more baroclinic tendencies. Furthermore, Washington, D.C. is under predominantly continental air with westerly flow while central Florida is predominantly under maritime air with easterly and westerly flow. This key difference exposes Washington, D.C. to more frontal systems than central Florida. Subsequently, central Florida is exposed to more airmass thunderstorms due to sea breezes originating from both east and west coastlines. Selecting solely warm season thunderstorms aimed to avoid the baroclinic systems prevalent in spring and fall which typically foster more linear, multicellular, and severe thunderstorms.

Nevertheless, trying to select a sufficient amount of thunderstorm cases that met the criteria discussed in section 3.2, yet were not influenced by baroclinic forcings,

was virtually impossible. In fact, collecting the 47 isolated thunderstorm cases near Washington, D.C. over a 6-year timespan was an arduous task. The majority of the observed storms were omitted due to linear or severe characteristics associated with frontal passage. Thus, the frequency of occurrence of isolated thunderstorms was much less frequent in Washington, D.C. than its southern counterpart. If identifying isolated storms was so difficult to achieve in the Washington, D.C. area, then JP17's lightning cessation model, which depends on isolated cell structure, is definitely not an ideal cessation model for that area. Perhaps lightning cessation guidance cannot be captured by a single comprehensive model, but rather relies on key climate distinctions.

Lastly, although this study endeavored to replicate JP17's methodology as precisely as possible, there were a few inevitable differences which could have contributed to performance inconsistencies. One of the most significant differences was the manner in which the radar analysis was performed. JP17 automated the process using Warning Decision Support System - Integrated Information (WDSS-II) software (Lakshmanan et al., 2007). However, in this thesis, the analysis and model testing was done manually in the Washington, D.C. area. This resulted in the use of layers instead of precise isothermal levels, vertical gaps in HCA data between available elevation angles, and decreased HCA resolution for the higher isothermal levels. Furthermore, JP17 tested the model using a robust database of 184 Florida thunderstorms, while this study consisted of only 47 Washington, D.C. thunderstorms. Nonetheless, error bar calculations for the performance statistics using the bootstrapping technique did account for these sample size differences.

5.2 Future Work

Future work is required to corroborate the aforementioned conclusions and clarify the discrepancies in the performance of JP17's lightning cessation predictive model in the Washington, D.C. area. This includes repeating the process by applying several adjustments to the methodology, increasing the number of thunderstorm cases by analyzing thunderstorms from other geographical locations, and testing different dual-polarization parameters.

Adjustments to the methodology to more closely mimic JP17's methods could help clarify the discrepancies in model performance. The primary adjustment would be to automate the analysis process by using WDSS-II or a similar program to ingest and analyze radar data. This would permit the automated merging of local WSR-88D radars and interpolation of essential radar parameters in three-dimensional space and time. Additionally, it would provide the highest resolution for all radar-derived parameters and avoid the opportunity for human error. Another adjustment to the methodology would be to analyze the storms until they fully dissipated (or the composite reflectivity dropped below 25 dBZ) rather than limiting analysis to 16 minutes before and after lightning cessation. This would help verify the performance statistics from the first verification approach.

The results of this study suggest that some aspects of the JP17 method may be applicable to climates different from central Florida. In addition to increasing the sample size in the Washington, D.C. area, the method should be tested in climates even more different from central Florida. Specifically, areas near an active dual-polarization radar and LMA network: Alabama and Oklahoma. The LMA in Alabama is more similar to the climate in central Florida, especially in summer, while the LMA in Oklahoma could provide a test in an entirely different climate. Moreover, LMAs supporting field research could provide spot checks in additional areas with

different climates.

Lastly, a similar probabilistic approach utilizing different dual-polarization parameters could be tested for the Washington, D.C. area. It was proven that although JP17's bootstrapped GLM did not perform as well in the Washington, D.C. area as it did in central Florida, the model did show promising skill. Graupel presence in the MP region was a key indicator of active lightning, while the reflectivity value threshold for lightning cessation in the model was too low. Perhaps a different combination of dual-polarization radar parameters or tweaking threshold values at select isothermal levels within the MP region would produce more optimal performance results. Ideally, this revised GLM would incorporate some of the same radar parameters used in JP17 in addition to other statistically significant dual-polarization parameters.

In conclusion, although JP17's bootstrapped GLM did not achieve the same level of success in the Washington D.C. area as it did in central Florida, the results were still significant. Performance statistics show that the model did achieve skill and the 95.0% probability thresholds did shorten the wait time by 3 minutes, while still achieving favorable HSS values. Additionally, the storm by storm verification approach revealed that while the model performed worse in the Washington D.C. area compared to central Florida, some of the performance statistics were still comparable and thus reassuring. Recommendation for action would be to retain the probabilistic guidance concept from JP17 and develop and test various GLMs with a new combination of radar parameters extracted from the MP region. Ideally, thunderstorms from select locations across the nation would be incorporated into the storm database. This could possibly achieve a more comprehensive lightning cessation predictive model that would deliver optimal lightning cessation guidance for locations with differing climates.

Bibliography

Agresti, A. (2013). Introduction to Generalized Linear Models. In *Categorical Data Analysis*, pages 113–162. John Wiley & Sons, 3rd edition.

Anderson, H. A. (2010). *Characteristics of decaying storms during lightning cessation at kennedy space center and Cape Canaveral Air Force station, MS Thesis*. Florida State University.

Chmielewski, V. C. and Bruning, E. C. (2016). Lightning Mapping Array flash detection performance with variable receiver thresholds. *J. Geophys. Res. Atmos.*, 121(14):8600–8614.

Davey, M. J. and Fuelberg, H. E. (2017). Using radar-derived parameters to forecast lightning cessation for nonisolated storms. *J. Geophys. Res. Atmos.*, 122:3435–3456.

Donaldson, R., Dryer, R., and Kraus, M. (1975). An objective evaluator of techniques for predicting severe weather events. In *Ninth Conference on Severe Local Storms*, pages 321–326, Norman, OK. Amer. Meteor. Soc.

Doviak, R. J. and Zrnic, D. (2014). *Doppler Radar and Weather Observations*. Academic press.

Efron, B. (1979). Bootstrap methods: Another look at the jacknife. *Ann. Stat.*, 7:1–26.

Efron, B. and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC Press, Boca Raton, FL.

Flueck, J. (1987). A study of some measures of forecast verification. In *Conference on Probability and Statistics in Atmospheric Science*, pages 69–73, Edmonton, Canada. Amer. Meteor. Soc.

Gandin, L. and Murphy, A. (1992). Equitable scores for categorical forecasts. *Mon. Weather Rev.*, 120:361–370.

Hesterberg, T. C. (2015). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *Am. Stat.*, 64(4):371–386.

Holle, R. L. (2016). Lightning Fatalities by State. Technical report, Vaisala Inc., Tuscon, AZ.

Holle, R. L., Lopez, R. E., and Zimmermann, C. (1999). Updated recommendations for lightning safety-1998. *Bull. Amer. Meteor. Soc.*, 80(10):2035–2041.

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

Jolliffe, I. T. and Stephenson, D. B. (2012). Deterministic forecasts of binary events. In *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, chapter 3, pages 31–59. John Wiley & Sons, Chichester, England, 2d edition.

Krehbiel, P., Thomas, R., Rison, W., Hamlin, T., Harlin, J., and Davis, M. (2000). GPS-based mapping system reveals lightning inside storms. *Earth and Space Sci. News*, 81(3):21–25.

Krehbiel, P. R. (1986). The electrical structure of thunderstorms. In *The Earth's Electrical Environment*, chapter 8, pages 90–113. The National Acadamy Press, Washington DC.

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

Kuettner, J. P., Levin, Z., and Sartor, J. D. (1981). Thunderstorm electrification-inductive or non-inductive? *J. Atmos. Sci.*, 38(11).

Kumjian, M. R. (2013). Principles and applications of dual-polarization weather radar. Part I: Description of the polarimetric radar variables. *J. Operational Meteor.*, 1(19):226–242.

Lakshmanan, V., Smith, T., Stumpf, G., and Hondl, K. (2007). The Warning Decision Support SystemIntegrated Information. *Wea. Forecasting*, 22(3):596–612.

Mackun, P. and Wilson, S. (2011). Population distribution and change: 2000 to 2010. Technical report.

Maier, L., Lennon, C., Britt, T., and Schaefer, S. (1995). LDAR system performance and analysis. In *Proceedings of the International Conference on Cloud Physics*.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman & Hall, 2d edition.

McNamara, T. M., Roeder, W. P., and Merceret, F. J. (2010). The 2009 update to the lightning launch commit criteria. In *14th Conference on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA. Amer. Meteor. Soc.

Murphy, A. and Daan, H. (1985). Forecast verification. In Murphy, A. and Katz, R., editors, *Probability, Statistics and Decision Making in the Atmospheric Sciences*, pages 379–437. Westview Press, Boulder, CO.

National Weather Service (2017). Weather Fatalities 2016.

NSSL (2017). Research Tools: Dual Polarized Radar.

Park, H., Ryzhkov, A., Zrnic, D., and Kim, K. (2009). The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, 24:730–748.

Patton, J. (2017). *Using radar-derived parameters to develop probabilistic guidance for lightning cessation within isolated convection near Cape Canaveral, Florida, MS Thesis*. Florida State University.

Peirce, C. S. (1884). The numerical measure of the success of predictions. *Science*, 4(93):453–454.

Poehler, H. A. and Lennon, C. L. (1979). Lightning Detection and Ranging (LDAR) system description and performance objectives. Technical report, NASA, Cocoa Beach, FL.

Preston, A. D. and Fuelberg, H. E. (2015). Improving lightning cessation guidance using polarimetric radar data. *Wea. Forecasting*, 30:308–328.

Proctor, D. E. (1971). A hyperbolic system for obtaining VHF radio pictures of lightning. *J. Geophys. Res.*, 76(6):1478–1489.

Rakov, V. A. (2013). The physics of lightning. *Surv. Geophys.*, 34(6):701–729.

Rakov, V. A. (2016). *Fundamentals of Lightning*. Cambridge University Press.

Rakov, V. A. and Uman, M. A. (2003). *Lightning: Physics and Effects*. Cambridge University Press.

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

Roeder, W. P. (2010). The four dimensional lightning surveillance system. In *Third Int. Lightning Meteorology Conf.*, Orlando, FL. Vaisala.

Roeder, W. P., McNamara, M., McAleenan, M., Winters, K. A., Maier, L. M., and Huddleston, L. L. (2017). The 2014 upgrade to the lightning warning areas used by the 45th Weather Squadron. In *18th Conference on Aviation, Range, and Aerospace Meteorology*, Seattle, WA. Amer. Meteor. Soc.

Ryzhkov, A., Schuur, T., Burgess, D., Heinselman, P., Griangrade, S., and Zrnic, D. (2005). The Joint Polarization Experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, 86(6):809–824.

Saunders, C. (2008). Charge separation mechanisms in clouds. In *Planetary Atmospheric Electricity*, pages 335–353. Springer, New York, NY.

Schuur, T., Ryzhkov, A., and Heinselman, P. (2003). Observations and classification of echoes with the polarimetric WSR-88D radar. Technical report, NOAA/National Severe Storms Laboratory, Norman, OK.

Stano, G. T., Fuelberg, H. E., and Roeder, W. P. (2010). Developing empirical lightning cessation forecast guidance for the Cape Canaveral Air Force Station and Kennedy Space Center. *J. Geophys. Res*, 115.

Takahashi, T. (1978). Riming electrification as a charge generation mechanism in thunderstorms. *J. Atmos. Sci.*, 35.

The Mathworks Inc. (2018). Piecewise Cubic Hermite Interpolating Polynomial (PCHIP).

Thomas, R. J., Krehbiel, P. R., Rison, W., Hunyady, S. J., Winn, W. P., Hamlin, T., and Harlin, J. (2004). Accuracy of the lightning mapping array. *J. Geophys. Res.*, 109(D14207).

United States Census Bureau (2010). Resident Population Data: Population Density.

Wilks, D. S. (2011). *Frequentist statistical inference*. Academic Press, 3rd edition.

Williams, E. R. (1985). Large-scale charge separation in thunderclouds. *J. Geophys. Res*, 90(D4):60136025.

Williams, E. R. (1989). The tripole structure of thunderstorms. *J. Geophys. Res.*, 94:13151–13167.

Wolf, P. (2007). Anticipating the initiation, cessation, and frequency of cloud-to-ground lightning, utilizing WSR-88D reflectivity data. *NWA Electron. J. Oper. Meteor.*

Zhang, R., Williams, E. R., and Rydock, J. (1991). Mixed-phase microphysics and cloud electrification. *J. Atmos. Sci.*, 48(19).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
02–26–2018	Master's Thesis	Oct 2016 — Mar 2018

4. TITLE AND SUBTITLE

FORECASTING LIGHTNING CESSATION
USING DUAL-POLARIZATION RADAR
AND LIGHTNING MAPPING ARRAY NEAR WASHINGTON, D.C.

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Holden, Nancy Marie, Capt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-MS-18-M-085

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

45th Weather Squadron
1201 Edward H. White Ste. C-129
Patrick AFB, FL 32925
COMM 321-853-8410
Email: William.Roeder@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

45 WS

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This study verifies the probabilistic lightning cessation model developed by Joseph Patton at Florida State University for use by the U.S. Air Forces 45th Weather Squadron at Cape Canaveral Air Force Station (CCAFS) and Kennedy Space Center (KSC). The Washington D.C. greater metropolitan area, which presents a climate different to that of central Florida, was chosen as the domain of study. Dual-polarization radar and Lightning Mapping Array data were used to track storms. Performance statistics show that the model revealed notable skill in the Washington D.C. area, yet not to the desired level as indicated by the models performance in central Florida.

15. SUBJECT TERMS

Lightning Cessation, Lightning Mapping Array, Dual-Polarization Radar

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Maj Omar Nava, AFIT/ENP
U	U	U	U	64	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636, x4518; Omar.Nava@afit.edu

Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18