TECHNIQUES FOR FORECASTING THE CESSATION OF LIGHTNING AT CAPE CANAVERAL AIR STATION AND THE KENNEDY SPACE CENTER

THESIS

Michael W. Holmes, First Lieutenant, USAF

AFIT/GM/ENP/00M-08

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

AFIT/GM/ENP/00M-08

TECHNIQUES FOR FORECASTING THE CESSATION OF LIGHTNING AT CAPE CANAVERAL AIR STATION AND THE KENNEDY SPACE CENTER

THESIS

Presented to the Faculty

Department 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 Meteorology

Michael W. Holmes, B.S.

First Lieutenant, USAF

March 2000

AFIT/GM/ENP/00M-08

TECHNIQUES FOR FORECASTING THE CESSATION OF LIGHTNING AT CAPE CANAVERAL AIR STATION AND THE KENNEDY SPACE CENTER

Michael W. Holmes, B.S.
First Lieutenant, USAF

Approved:

Gary R. Huffines (Chairman)	date
Cecilia A. Miner (Member)	date
William F. Bailey (Member)	date

Acknowledgements

There are a number of people who directly contributed to the successful completion of this research project. Most importantly is the support and encouragement I have received from my family and especially my loving wife Shannon. As she's so often done in the past, she has patiently endured my long absences from home with a tolerance that I could only dream about and when we did cross paths, she only had kind words of encouragement to offer. She is beautiful inside and out and I wouldn't be where I am today without her. My son Noah, while only 3 years old, was also instrumental in the completion of this project as a fantastic stress reliever. He never placed much priority on the books and papers I would be forever reading, and as a result, didn't hesitate to remove them from my hands and insist that I play with him.

On a more professional note, I would like to thank my thesis committee members, Lieutenant Colonel Cecilia Miner, and Dr. William Bailey and especially my thesis chairman, Major Gary Huffines who always provided guidance when asked, but more importantly, let me conduct the research as I saw fit. This large degree of freedom made this research effort the single most valuable learning experience I have had here at AFIT. I would also like to thank the rest of the AFIT faculty, Pete Rahe, and of course my fellow students who despite their best intentions didn't always provide intelligent answers to my questions, but were always the model of consistency when it came to comic relief.

Michael W. Holmes

Table of Contents

List of Figures

Figure Page

List of Tables

Table Page

Abstract

The focus of this research effort is directed toward identifying new methods of forecasting the cessation of lightning along the Central Atlantic Coast of Florida. Cloud-to-ground lightning flashes place Air Force (AF) personnel and assets at risk almost daily at this location. Providing a more accurate method of forecasting the cessation of lightning would allow for safer and more efficient execution of AF operations.

A data set consisting of 40 thunderstorm cases was identified within a 90 nautical miles (nmi) region surrounding the Melbourne, Florida WSR-88D (KMLB) site. Each case falls between the months of May and September and the years of 1995 through 1997. Simple and multiple linear regression models are built using this dataset. Variables included max Vertically Integrated Liquid water (VIL), max reflectivity, max peak current, peak cumulative flash rate, peak negative flash rate, and peak positive flash rate. Results indicate that three of the simple linear regression models to some extent accurately represent the data. Additionally, when the data set is separated by thunderstorm cell type (multi or single) and cell specific regressions are built, results indicate that the regressions based on the single-cell data set produce a substantial increase in forecast skill compared to that of climatology. In fact, some regressions are shown to improve forecast accuracy by 90 % over that of climatology. Moreover, multiple linear regression models are shown to produce similar results and further reinforce the notion that each thunderstorm cell type (multi or single) behaves substantially different from the other with respect to forecasting the cessation of lightning.

TECHNIQUES FOR FORECASTING THE CESSATION OF LIGHTNING AT CAPE CANAVERAL AIR STATION AND THE KENNEDY SPACE CENTER

1. Introduction

1.1 Background

The National Aeronautics and Space Administration (NASA) Kennedy Space Center (KSC) located on the Central Atlantic Coast of Florida is recognized as an area having among the highest lightning flash densities in the country (Orville 1991; Orville and Silver 1997). The high frequency of lightning places over 25,000 employees and over 7 billion dollars in facilities at risk each time a thunderstorm passes over the region (Boyd et al. 1995). With these points in mind, NASA has placed a high priority on accurately forecasting the initiation and cessation of lightning since it is critical to the safe and timely execution of ground and launch operations.

A number of previous research efforts have focused on identifying when a convective cloud will produce its first lightning flash, but little attention has been given to identifying when an existing thunderstorm will cease to produce lightning. Forecasting the cessation of lightning is equally as important as forecasting the initiation of lightning since ground, air, and launch operations remain on hold throughout the lifecycle of a nearby thunderstorm.

Analyzing some of the techniques used in forecasting the initiation of lightning may provide some insight into forecasting the cessation of lightning. In addition, some

unique and previously untried methods may also prove to be useful in forecasting the cessation of lightning and are the focus of this research effort.

1.2 Problem Statement and Objective

Starting from the premise that a convective cloud has generated sufficient charge separation to initiate at least one lightning flash, the problem is then identifying when this thunderstorm will cease to produce lightning. The 45th Weather Squadron at the KSC and the United States Air Force (USAF) weather community in general would benefit if this problem were solved. Thus, the objective of this research effort is to provide an accurate method of forecasting the cessation of lightning.

1.3 Overall Approach

Forecasting the cessation of lightning, despite its importance to the aviation community has received very little direct attention from researchers. In fact, only one published research effort has focused directly on forecasting the cessation of lightning (Hinson 1997), and it produced mixed results. The major limiting factor to this lone research effort was the sample size. Hinson's analysis method only allowed him the opportunity to analyze three thunderstorm cases, which makes it difficult to draw any statistically significant results. Some of his findings however, were compelling and are discussed in the literature review chapter of this thesis along with a description of his methodology. Ultimately, the analysis method and time restraints limited the number of cases he could evaluate. With this in mind, a major focus of this research effort is to decrease the analysis time and increase the number of cases.

2. Literature Review

2.1 Convective Lifecycle of a Thunderstorm

The convective lifecycle of a thunderstorm in an environment of weak vertical wind shear can be broken up into three stages (Byers and Braham 1949). First is the cumulus stage (Fig. 1a). During this stage of development, updrafts on the order of 1 – 10 m s^{-1} are the dominant feature throughout the cell (Bluestein 1993). As the precipitation particles continue to grow within the convective cell, they eventually become so heavy that the updraft can no longer sustain their weight. Subsequently, the precipitation particles begin to fall, initiating a downdraft. This downdraft is then enhanced due to evaporative cooling as the precipitation particles pass through unsaturated air making the air more negatively buoyant. The point at which precipitation particles reach the surface marks the transition from the cumulus stage to the mature stage (Fig. 1b). The mature stage is characterized by the presence of both an updraft and a downdraft. In this idealized model, as precipitation continues to fall, the updraft is essentially completely overtaken by the downdraft. This point marks the transition from the mature stage to the dissipating stage (Fig. 1c) and can be characterized by a steady decrease in the rate of precipitation at the surface.

The lifecycle of a multi-cell thunderstorm is very similar to that of a single-cell thunderstorm. Each cell that develops in a multi-cell thunderstorm behalves individually as does a single-cell thunderstorm. The distinction between the two occurs in the dissipating stage of the lifecycle. As the downdraft collides with the surface of the earth, it disperses horizontally. In a multi-cell thunderstorm, this horizontal dispersion of the

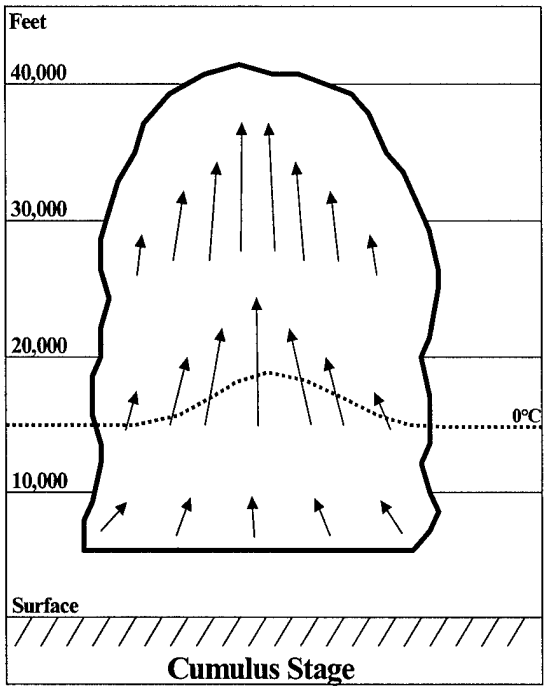

Figure 1a. Cumulus stage of development. (Adapted from Byers and Braham 1949)

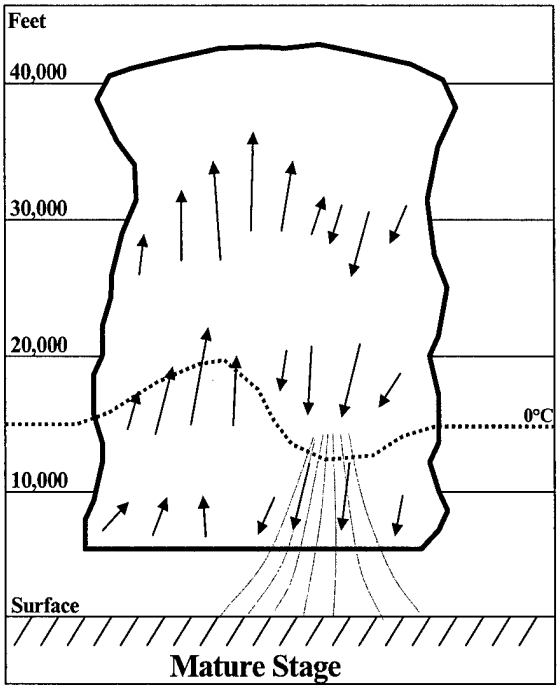

Figure 1b. Mature stage of development. (Adapted from Byers and Braham 1949)

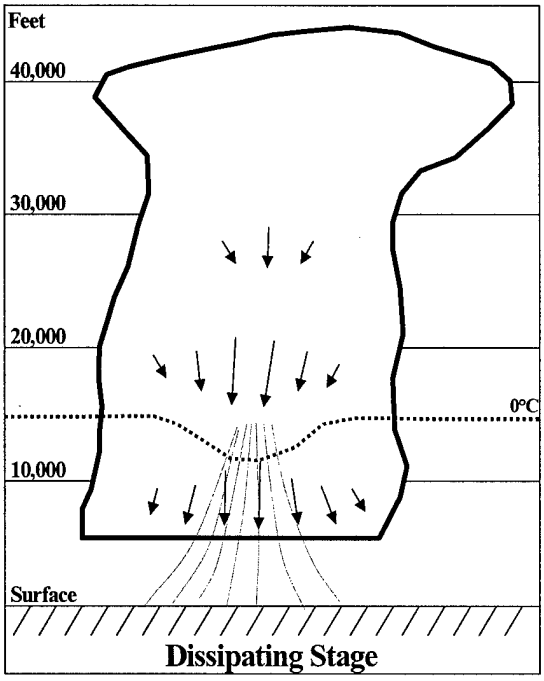

Figure 1c. Dissipating stage of development. (Adapted from Byers and Braham 1949)

downdraft converges with another boundary (sea breeze, outflow from another storm, etc). Since this convergence zone is bounded below by the surface of the earth, the only direction that the air can move is upward, spawning a new convective cell. This process of cell regeneration continues until thermodynamic conditions are no longer favorable, or convergence between airmasses ceasees. In the single-cell case, the downdraft collides and disperses horizontally just as in the multi-cell case. However, there is either no boundary present at the surface for the storm outflow to converge with or the convergence is not sufficient to initiate another convective cell, ultimately, indicating the death of the cell and the entire thunderstorm.

2.2 Lightning

2.2.1 Cloud Electrification

Fundamentally, lightning flashes are a direct result of the separation between strongly positive and negative charged regions that exist in some convective clouds. The processes by which the charges and subsequent separations are generated are not fully understood and are currently an area of active research. There are, however, two basic theories used to explain the dipole nature of thunderstorms (Uman 1987; Saunders 1993).

The first theory focuses on cloud electrification due to convective processes. Once convection is initiated, positively charged fair weather ions are picked up and carried aloft into the cloud where it is thought they then attach themselves to water droplets (Saunders 1993). These positively charged water droplets are then carried higher into the cloud by the ongoing convection. A negatively charged screening layer is generated along the outer boundary of the cloud and carried downward by the convective downdrafts. The result of this process is the separation of charged regions within the

cloud. If the positive and negative charged regions within the cloud are sufficiently large, a lightning flash may occur.

The second major theory focuses on particle charging mechanisms and can be further broken into inductive and noninductive charging theories. Both theories suggest that charged particles are generated as a result of collisions with one another inside the cloud. The processes differ in that the inductive charging mechanism requires the presence of a vertical electric field, while the noninductive mechanism does not. There are a number of possible noninductive charging mechanisms, but only one (grauple-ice mechanism) has been duplicated in laboratory work (MacGorman and Rust 1998) and has been shown to be consistent with observations (Rutledge and Carey 1997).

2.2.2 The Lightning Flash

2.2.2.1 Negative Cloud to Ground Lightning

The majority of lightning flashes in thunderstorms lower a net negative charge to the ground. In fact, Orville (1994) found in a study of cloud-to-ground (CG) flash characteristics conducted in the United States over a six-year period that more than 90% of all thunderstorms lowered a net negative charge to the surface. A generalized picture of the charge separation in a thunderstorm that produces a negative CG discharge is illustrated in Figure 2.

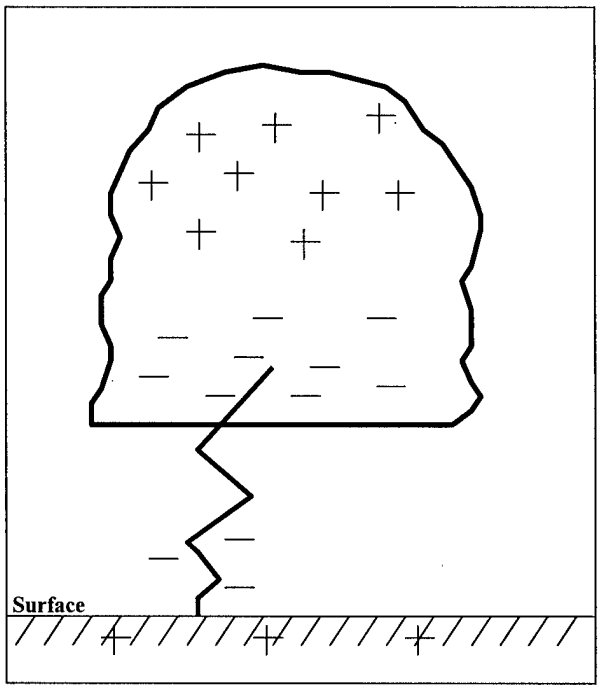

Figure 2. A generalization of the charge separation within a thunderstorm. (Adapted from Uman 1987)

Once sufficiently large positive and negative regions have been created within the cloud a coronal discharge may occur. A coronal discharge that initiates from the positive or negative charged region of the cloud is the beginning of a positive or negative lightning flash, respectively. The remainder of this section specifically focuses on the lifecycle of a single negative CG flash. An outline of this process is described below and is meant to provide a general overview of the process. For a detailed discussion of all processes involved in a single lightning CG flash, refer to Uman (1987) or MacGorman and Rust (1998).

When the breakdown potential of the free atmosphere is reached, a coronal or point discharge is initiated from the negatively charged region of the cloud into the free

atmosphere. Once the breakdown potential is reached, negative charge begins lowering to ground in steps. The sum of these steps is called the stepped leader. Each step is approximately 50 m in length and travels at approximately 10^5 m s^{-1} (Uman 1987). As the stepped leader approaches the ground, the electric field at the surface begins to build on tall or irregularly shaped objects until the breakdown potential is reached. An upward moving positive discharge is then initiated and is termed the attachment process or attachment leader. When the attachment leader meets with the stepped leader, the ionized path is completed between the cloud and the ground and a negative charge is transferred to the earth. At this point a return stroke is initiated. It travels from the ground back up the ionized path to the source region of charge within the cloud. Upon the completion of the return stroke, if sufficient charge is still available in the charged region of the cloud, then a dart leader will travel down the previously ionized path and initiate another return stroke. This dart leader and subsequent return stroke may repeat several times and is responsible for the flicker of lightning as seen by the human eye.

2.2.2.2 Positive Cloud to Ground Lightning

Less than 10 percent of all CG flashes lower a net positive charge to the surface of the earth (Orville 1994). One of the main differences between positive CG flashes and negative CG flashes is that positive flashes can have peak currents that are much larger than those of negative CG flashes and they generally are more frequent as you move to higher latitudes, higher altitudes or during the winter season (Uman 1987). However, the discharge process are similar.

2.2.3 National Lightning Detection Network

The National Lightning Detection Network (NLDN) was established in 1987 by combining three previously independent lightning detection networks into one all encompassing lightning detection network that covers the entire United States (Cummings et al. 1998). In 1995, an upgrade of the NLDN sensors was completed. Following the upgrade, the NLDN was comprised of two distinct lightning detection sensors that were to be used concurrently. The two sensors are the Time of Arrival (TOA) and the Improved Accuracy from Combined Technology (IMPACT) sensor that utilized both a TOA sensor and a Magnetic Detection Finder (MDF) sensor. There are 59 TOA sensors and 47 IMPACT sensors distributed over the continental United States (Fig. 3). Cummings et al. (1998) estimated that after the 1995 upgrade the location accuracy increased to 0.5 km and the flash detection efficiency increased to 80-90% over most of the United States. An evaluation of the NLDN conducted over the Northeastern United States during and following the upgrade indicated that there was a modest increase in detection efficiency when compared to the network prior to the upgrade (Idone et al 1998).

2.3 Weather Surveillance Radar

The Weather Surveillance Radar 1988 Doppler (WSR-88D) also know as the Next Generation Weather Radar (NEXRAD) was commissioned in 1988 and was in operation at the Melbourne, Florida, site in 1994. The WSR-88D is a horizontally polarized, S band radar that transmits at 3 GHz, and has a wavelength of 10.71 cm (Rinehart 1997). This radar offers a wide range of products that can be used to evaluate a

storm. However, only two are used in this research effort, specifically, the base reflectivity and VIL. Both are discussed in detail in later sections.

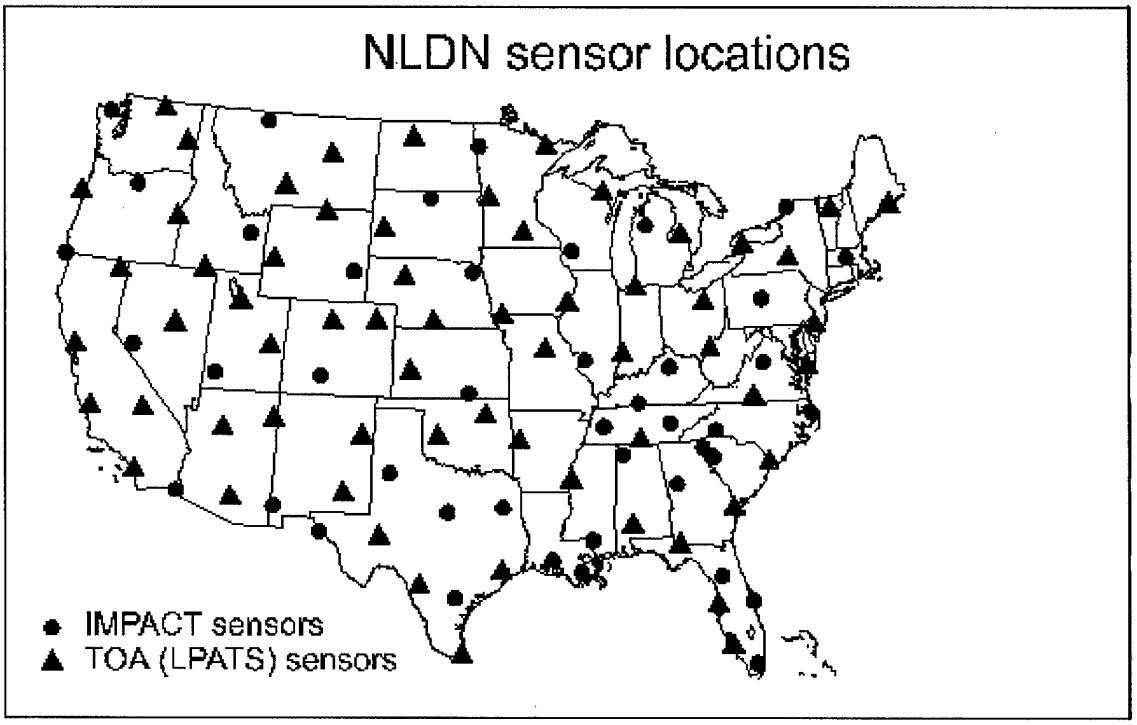

Figure 3. NLDN map. Depicts the geographic location of the IMPAC and TOA sensors over the continental United States (From Cummings et al. 1998).

2.3.1 Modes of Operation

The WSR-88D operates in two different modes, the most important of which is the precipitation mode, which utilizes the most effective scan strategies to identify potentially threatening meteorological events. All of the radar data gathered in this research effort was taken while the WSR-88D was in the precipitation mode. The second mode that the WSR-88D can operate in is the clear air mode. It is more sensitive than the precipitation mode and is mainly useful in identifying wind speeds on days when meteorological targets are not present.

2.3.2 Volume Coverage Patterns

Two volume coverage pattern (VCP) scan strategies are available when the WSR-88D is in precipitation mode. The first is VCP21, which collects data at each of the following elevation scans; 0.5, 1.45, 2.4, 3.35, 4.3, 6.0, 9.9, 14.6, 19.5 degrees (Fig. 4). This is the scan strategy that was utilized by the Melbourne radar site (KMLB) prior to 1997 (Gremillion and Orville 1999). This particular scan strategy samples the lower levels completely. However, there are large gaps between 4.3 and 19.5 degree scans, which could contain potentially important meteorological information. The second scan strategy available in precipitation mode is VCP11 (Fig. 5). In 1997 the Melbourne radar site switched to this scan strategy which fills many of the gaps between 4.4 and 19.5 degree elevation slices that occur in the VCP21 scan strategy.

2.3.3 Reflectivity Products

The WSR-88D generates a number of products that can be used to interpret various details within a storm. Of all the products, the one most used by researchers and operational forecasters is the base reflectivity product. The base reflectivity product displays the echo intensity of a particular target based on how much electromagnetic radiation is returned from that target back to the antenna. One base reflectivity product is returned for each elevation scan that is accomplished and can be requested in 0.54, 1.1, or 2.2 nmi resolution. It can be used to estimate rainfall and intensity, storm structure and identify the storms boundaries.

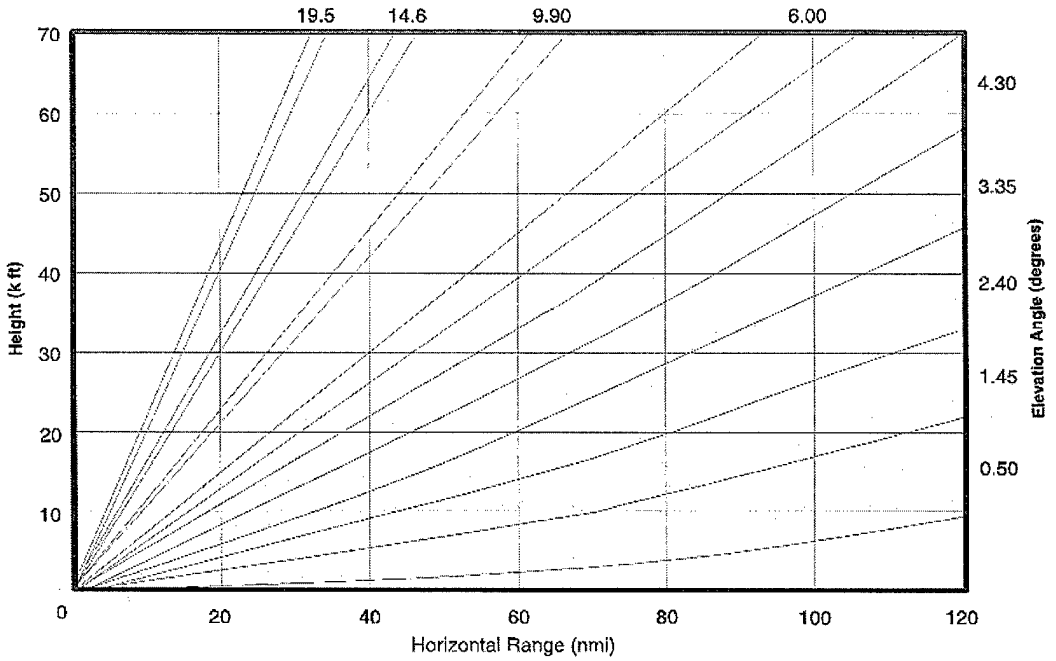

Figure 4. Scan strategy of Volume Coverage Pattern (VCP) 21 (From WSR-88D Handbook 1998).

The composite reflectivity product is also commonly used. It is a volumetric product, which means that a complete volume scan must be available for it to be displayed. In general, it is a composite summary of the maximum observed base reflectivity value measured above each geographic grid point. This product is available every 5 to 6 minutes depending of the scan strategy being used at the time and allows the radar user to quickly identify and track the maximum reflectivity within the storm.

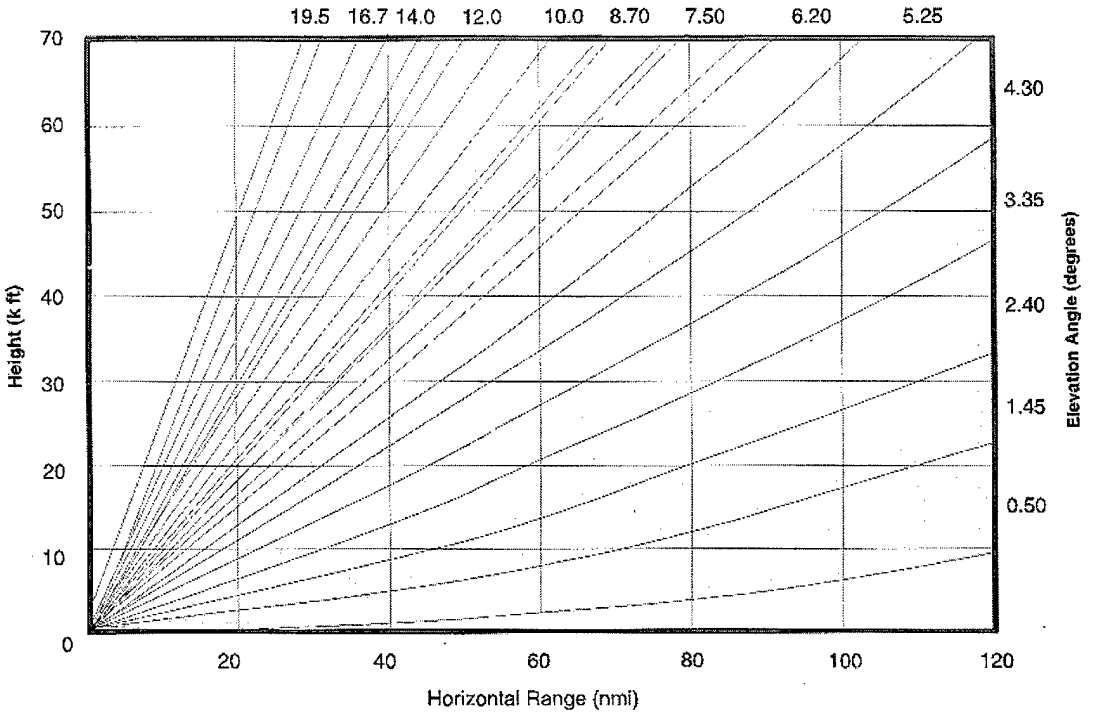

Figure 5. Scan strategy of Volume Coverage Pattern (VCP) 11 (From WSR-88D Handbook 1998).

2.3.4 Vertically Integrated Liquid Water Product

The vertically integrated liquid water (VIL) product provides the radar user with a graphical representation of the liquid water content (kg m^{-2}) that is contained within a 2.2 x 2.2 nmi column of air above a particular geographic point. This product, like composite reflectivity is a volumetric product and is only available after the completion of a volume scan.

2.4 WATADS

The WSR-88D Algorithm Testing and Display System (WATADS) was built by the National Severe Storms Laboratory (NSSL) to facilitate after-the-fact analysis of WSR-88D data. It was also designed to evaluate a particular algorithm's performance

over a specific region after adjusting the adaptable parameters. WATADS is a Unix based software package that ingests Level II archive radar data from an 8 mm tape. After reading in the archive Level II data, WATADS then processes it through the necessary algorithms. When the data have been completely processed, nearly all the products that are available to a user at a Principal User Processor (PUP) are also available through WATADS via the Radar Analysis and Display System (RADS).

2.5 Radar and Lightning Studies

Over the last two decades, weather radars operating at different power levels and wavelengths have been used to sample thunderstorms. One particular research use of these radars has been to sample thunderstorms and to correlate some radar parameter with the initiation of cloud electrification and a subsequent lightning flash. Florida, and more specifically, the Kennedy Space Center (KSC) located on the Atlantic Coast has been a favorite location for such lightning research given its high flash density (Orville 1991; Orville and Silver 1997).

A Lightning Detection and Ranging (LDAR) network located at the Kennedy Space Center has been used to identify correlations with radar parameters (Lhermitte and Krehbiel 1979). Their research took place prior to the installation of the Melbourne, FL WSR-88D. They used a network of three smaller Doppler radars to perform their research. After two summers, they found that LDAR data are usually confined between an altitude of 7 to 10 km, and there was a positive correlation between the maximum LDAR data rate and the maximum vertical velocity, and radar reflectivity. No correlations were made between Doppler radar parameters and the electrification or first flash of lightning.

Other research efforts have looked at different types of precipitation particles to identify a possible correlation between the type of precipitation particles and the occurrence of lightning (Jameson et al. 1996; Lopez and Aubagnac 1997; Carey and Rutledge 1998). In an evaluation of three Florida thunderstorms, the onset of electrification was shown to begin just prior to the rapid growth of the reflectivity volume above the –7 degree C level (Jameson et al. 1996). In another study, temporal variations in the mass of small hail, grauple, and supercooled liquid water were shown to be coincident with variations in the flash rate (Lopez and Aubagnac 1997).

A study of 1,257 thunderstorm cells in the Southern Plains compared the maximum CG flash rates with maximum values of reflectivity, VIL, and thickness of the reflectivity $\geq$ 30 dBZ above the 0 degree C height (MacGorman and Filiaggi 1997). The results of the study indicated that 70% of the cells that produced CG flashes had a peak in the flash rate at, or slightly after, the same time as the radar inferred parameters peaked. The 30 dBZ thickness values by themselves correlated well with increased CG flash rates. In fact, 92% of the cells that had no CG flashes also had a 30 dBZ thickness less than 7 km, and 85% of the cells having a high CG flash rate greater than 4 min^{-1} also had a 30 dBZ thickness greater than 7 km.

Radar reflectivity values at particular heights have also been evaluated for their use in identifying regions in a developing convective cell that may produce a lightning flash. Larsen and Stansbury (1974) suggested that the 43 dBZ precipitation cores at a level of 7 km plotted on a Constant Altitude Plan Position Indicator (CAPPI) were most likely the source regions of observed lightning.

A study by Buechler and Goodman (1990) proposed a new storm identification algorithm for the NEXRAD (Next Generation Weather Radar). Their proposed algorithm would recognize a storm by the appearance of a 40 dBZ echo at the –10 degree C level. After evaluating 15 thunderstorms that occurred over Alabama, New Mexico, and Florida, their results indicated a 100% probability of detection of a lightning-producing storm with lead-times ranging from 4 to 33 minutes.

As mentioned earlier, very few published research efforts have looked at forecasting the cessation of lightning. However, in the summer of 1996 three thunderstorms that occurred near the Kennedy Space Center were analyzed specifically to look for correlations between the cessation of lightning in convective cells and radar reflectivity values (Hinson 1997). The 40 and 45 dBZ reflectivity values of each of the three convective cells were analyzed at both the -10 degree C (6000 m) and -20 degree C (7500 m) levels. The last indication of a reflectivity value (40 or 45 dBZ) at a particular level (-10 or -20 degree C) was to be an indication of the impending cessation of lightning within the convective cell. The time at which the last reflectivity value at a particular level was observed was recorded then compared to the time at which the last CG flash occurred as recorded by the NLDN. A time lag was then computed from each comparison. Results indicated that in general, the time lag between the last observed reflectivity echoes at the different temperature heights and the last CG flashes were reasonably consistent. However, the time lags observed using the 45-dBZ reflectivity echoes at the -10 degree C height were the most consistent (30 min for each case).

In addition to the above comparisons, Hinson (1997) also looked for correlations between electric field variations as measured by field mill and NLDN data for the same

three convective cells. Unlike the radar reflectivity comparisons, the electric field variations did not prove to be useful in forecasting the cessation of lightning.

3 Methodology

3.1 Introduction

Forecasting the cessation of lightning has been a research topic that has received very little attention by the research community. The results of the lone research effort that primarily focused on forecasting the cessation of lightning by Hinson (1997) were minimized because only three thunderstorms were evaluated. With this in mind, the initial focus of this thesis was to identify a larger sample size of thunderstorms and thereby allow for a more statistically significant result. This was an ambitious goal that absorbed a large proportion of the allotted research time. The remainder of this chapter focuses on the methods, techniques, and criteria used to identify the thunderstorms in this thesis.

3.2 Geographic Region of Interest

The 45th Weather Squadron (45WS) located at Kennedy Space Center sponsored this research effort. With increasing sample size in mind, a large region that included all of the 45 WS areas of forecast responsibility was considered in this study (Fig. 6). This region is centered on the Melbourne (KMLB) WSR-88D site (28.1128° N, 80.6589° W) and extends out 90 nmi around KMLB. Each of the valid thunderstorm cited later in this section occurred in this region.

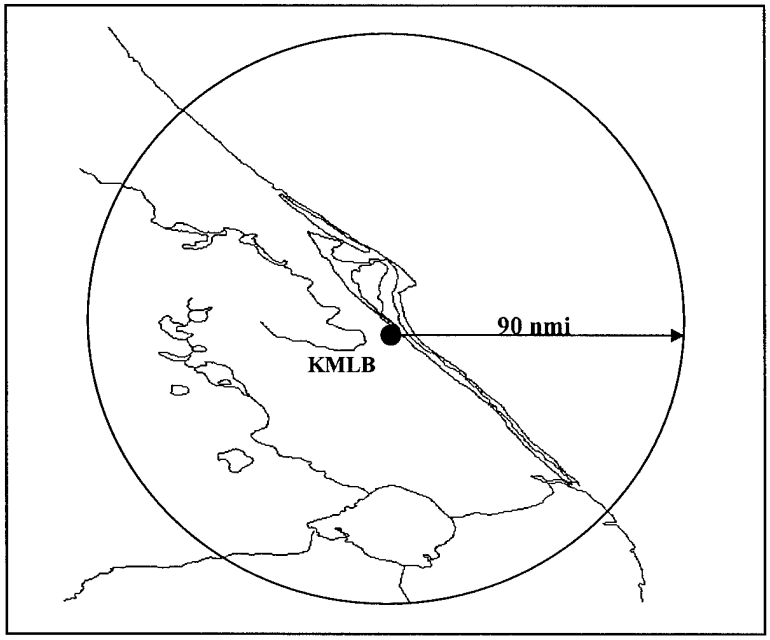

Figure 6. Geographic region of interest.

3.3 Identification of Cases

The thunderstorm cases in this thesis were identified using a combination of climatology, Weather Surveillance Radar 88D (WSR-88D), and NLDN data. The specifics of each are described in the following sections.

3.3.1 Applying Climatology

Airmass thunderstorms that occur over Florida and develop in low shear environments have been shown to have the highest frequency of cloud-to-ground flashes in the United States (Orville 1991). This type of thunderstorm can be initiated either by localized heating or by low level convergence (Byers and Rodebush 1948). Typically, airmass thunderstorms initiated by low level convergence that develop over Florida

during the summer months are due to interactions between sea and land breezes and the low level wind or convective outflow from a nearby thunderstorm. These conditions, along with the absence of synoptic scale forcing, are most commonly observed between the months of May and September (Neumann 1971), and, consequently, only thunderstorms that occurred between May and September were considered in this research project.

3.3.2 Identification Using NLDN Data

In order to maintain consistency of the data set, only thunderstorms that occurred after the NLDN upgrade in 1994 were included in this research project. An additional restriction was that all valid thunderstorms must not have progressed beyond the mature stage of development (Fig. 1b) in their lifecycle as they entered the 90 nmi umbrella surrounding KMLB. This restriction was meant to ensure that, as each thunderstorm moved under the 90 nmi umbrella it was in a phase where collisions between water and ice particles were actively generating and/or maintaining sufficient charge separation within the cell to produce a lightning flash. The last restriction was that each thunderstorm must be reasonably isolated, meaning that the lightning flashes from a nearby thunderstorm were not significantly contaminating the thunderstorm of interest. Some contamination was allowed in the early stages of the thunderstorms lifecycle as long as the contaminant flashes were not recorded as the maximum peak current, and did not significantly affect the flash rate.

With these criteria in place, NLDN data archived at the Air Force Institute of Technology (AFIT) were then used to identify specific thunderstorm cases. This was accomplished by identifying days in which a thunderstorm occurred within the 90 nmi

radius of KMLB. An Interactive Data Language (IDL) program (Appendix A) was written to extract and write to a file the total number of flashes that were recorded beneath the umbrella surrounding KMLB for each day (0000 UTC – 2359 UTC) that fell between the months of May and September and between the years of 1995 and 1997, inclusively. After running this program a list of 566 valid thunderstorm days were produced.

With a list of valid thunderstorm days in hand, it was then possible to generate Graphics Interchange Format (GIF) animations of the CG lightning over the region of interest for each given day (Fig. 7). The region of interest was expanded to 120 nmi to allow tracking of a thunderstorm as it moved into or out of the 90 nmi umbrella surrounding KMLB. An IDL program was written to generate a GIF image every ten minutes for each of the 566 valid thunderstorm days (Appendix B). The ten minute time interval was selected as a compromise between storm detail and hard disk space availability. The ten minute interval was a small enough time step to capture the general details of an individual thunderstorm while at the same time conserving disk space usage. A more detailed analysis of the lightning data using a five minute time interval was conducted later, and the details are discussed in section 3.4. For the purpose of storm identification, the ten minute interval was sufficient. The NLDN data was then run through this IDL program for each of the 566 valid thunderstorm days and ultimately generated more than 80,000 GIF images that occupied more than 10 Gigabytes in hard disk space.

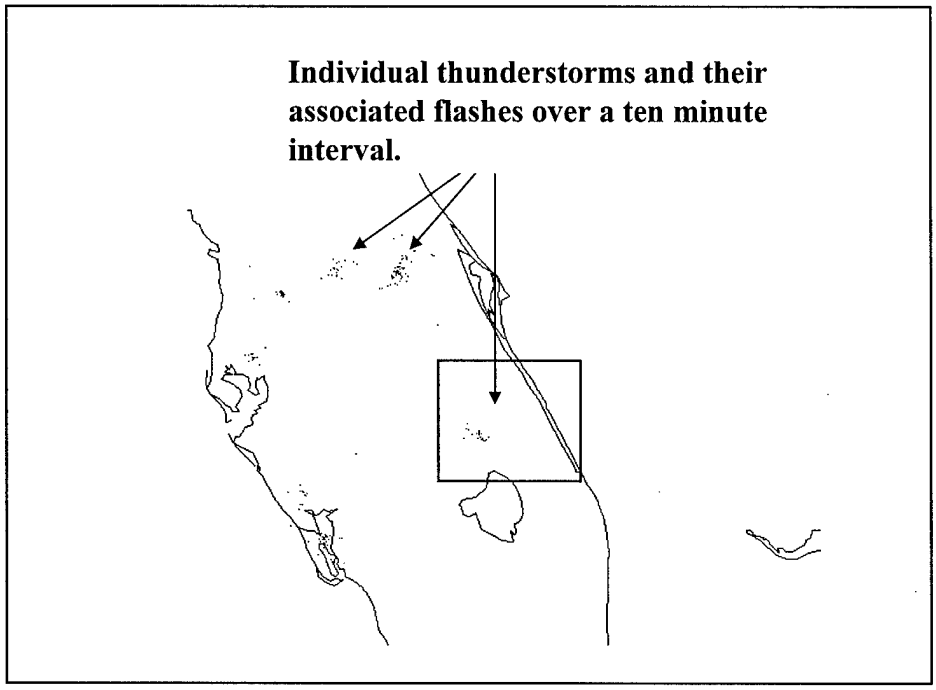

Figure 7. Example of large scale GIF image used to identify thunderstorm cases. Boxed region is magnified in Figure 8.

With the GIF images generated, it was then possible to begin identifying specific thunderstorm cases. The software program Xanimate® running on a SUN Sparc® Workstation was used to animate each series of thunderstorm GIF images. The most obvious dates to begin with were the dates in which previous research efforts had identified low shear thunderstorm events that occurred over Florida. Both Hinson (1997) and Gremillion (1999) identified these types of storms that occurred over the KSC complex. A total of eight of these previously identified storms met the limiting criteria described previously. By animating the lightning plots, an additional 38 previously uncited thunderstorms were identified that met the limiting criteria. This brought the total number of valid thunderstorm cases to 46.

3.3.3 Further Identification Using WSR-88D Data

Having identified the date and time of each of the 46 thunderstorm cases, the corresponding days of WSR-88D Archive Level II data were then requested through the Air Force Combat Climatology Center (AFCCC). Upon receipt of the Level II archive data stored on 8 mm tape, each of these 46 cases was then processed through WATADS, and then reevaluated using radar data.

The limiting criterion for the WSR-88D data is similar to that of the lightning data. Each thunderstorm must not be beyond the mature stage of development (Fig. 1b) as it enters the 90 nmi umbrella surrounding KMLB. This limitation ensures that when the radar is on its lowest elevation scan (0.5°) that the -10° C level (6000 m) of each thunderstorm can still be sampled by the radar. Additionally, this limitation also ensures that there is still an updraft present within the thunderstorm and subsequently, higher reflectivity particles (larger) are still suspended in the cell. The stage of development was determined objectively by evaluating vertical cross sections of each thunderstorm with the aid of WATADS. After applying this criterion, the previously identified data set it was reduced from 46 to 44 cases. The excluded 2 cases had incomplete or corrupted Level II archive data sets and were therefore discarded.

3.4 Extraction of Specific Lightning Parameters

With the 44 valid thunderstorm cases identified, a more detailed analysis of each case was then possible. An IDL program (Appendix C) was written to extract specific lightning parameters (described in Table 1) from the NLDN data archive.

Table 1. Description of lightning parameters.

Variable	Definition
Tot Flashes	Total number of flashes recorded throughout the life of the thunderstorm.
Last Flash	Time occurrence of last flash, recorded in minutes following 0000 UTC.
Max PK Curr	Maximum absolute peak current recorded during the life of the thunderstorm (kA).
t/o Max PK Curr	Time occurrence of peak current, recorded in minutes following 0000 UTC.
PK Flash Rt	Maximum flash rate recorded over a 5 minute interval during the life of the thunderstorm.
t/o PK Flash Rt	Time occurrence of peak pos flash rate, recorded in minutes following 0000 UTC.
PK Pos Flash Rt	Maximum pos flash rate recorded over a 5 minute interval during the life of the thunderstorm.
t/o PK Pos Flash Rt	Time occurrence of peak pos flash rate, recorded in minutes following 0000 UTC.
PK Neg Flash Rt	Maximum neg flash rate recorded over a 5 minute interval during the life of the thunderstorm.
t/o PK Neg Flash Rt	Time occurrence of peak neg flash rate, recorded in minutes following 0000 UTC.

This IDL program also produced two graphical outputs. The first type is a GIF image that depicts the location of each flash on a high resolution map (Fig. 8). These images are very similar to the ones initially used to identify potential thunderstorm cases except they were generated every 5 minutes and were focused in on one particular thunderstorm. An animation of a thunderstorm as depicted by lightning flashes can be built when the GIF images are looped in sequence. The second graphical output is a time series plot of the four lightning parameters used in this research effort (Fig. 9), those values being maximum absolute peak current (solid) measured along the left y-axis, peak flash rate (dash), peak negative flash rate (dot), and peak positive flash rate (dash dot). The latter three are measured along the right y-axis.

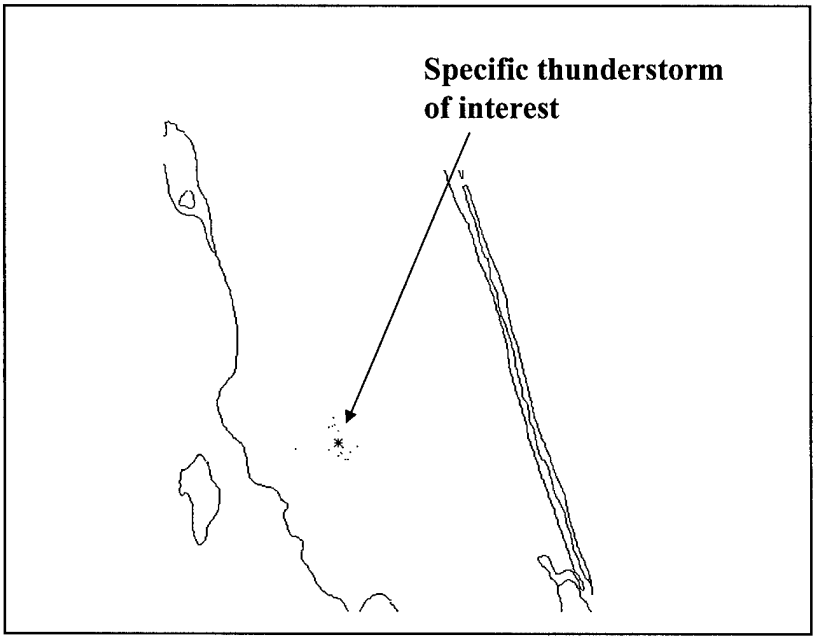

Figure 8. Example of small scale GIF image used to isolate each thunderstorm case.

3.5 Extraction of Specific WSR-88D Parameters

Extraction of specific WSR-88D parameters was a very important part of this research project, since a large majority of prior research efforts cited in Chapter 2 of this thesis had previously identified relationships between various radar and lightning parameters. The WSR-88D derived parameters of particular interest to this research project are listed in Table 2. Automating the extraction of these specific WSR-88D parameters was not possible, and they were instead extracted manually, volume scan by volume scan, using WATADS.

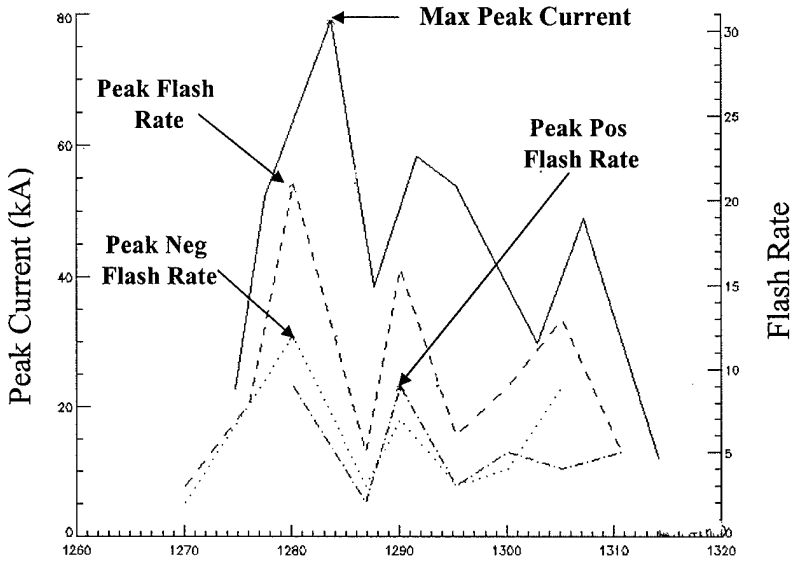

Figure 9. Example of a time series plot.

3.5.1 Data Extraction Problems

Using WATADS, the extraction of the base reflectivity values at unique temperature levels, proved very time consuming and more importantly, the results were inconsistent. WATADS allows a user to manually extract a reflectivity value and its respective height by merely placing the mouse over the bin of interest and clicking the left mouse button. When this sequence of events is accomplished, WATADS displays the reflectivity value and its height in the lower right hand portion of the RADS display window. This process was time consuming but initially appeared to accurately produce the desired values. However, after reevaluating a test case three separate times, it was

Table 2. Description of radar parameters.

Variable	Definition
Max VIL	Vertically integrated liquid water. Liquid water content (kg m^{-2}) that is contained within a 2.2 x 2.2 nmi column of air above a particular geographic point.
t/o Max VIL	Time occurrence of maximum VIL, recorded in minutes following 0000 UTC.
Max Ref	Maximum Reflectivity. Represents the echo intensity of a particular target based on how much electromagnetic radiation is returned from that target back to the antenna and has units of dBZ.
t/o Max Ref	Time occurrence of maximum Reflectivity, recorded in minutes following 0000 UTC.

discovered that this method produced three separate results. This was because, in WATADS, as you move the mouse within a specific bin and click the mouse button, you can get different reflectivity and height values literally based on the pixel you have selected. This fact made it very difficult to duplicate the same procedure consistently for one case and even more difficult to duplicate over 44 cases. As such, this method of analysis was abandoned since any results that were produced would be extremely subjective.

3.6 Complete Data Set

The culmination of applying the limiting criterion to the extraction of specific lightning and WSR-88D parameters for each thunderstorm case is displayed in Table 3. All variables listed in Table 3 have been described in the previous sections with the exception of case number, (which is self explanatory) and values followed by "dt", which represent the time difference (in minutes) between the peak occurrence of that variable and the last flash of the thunderstorm. Additionally, this table represents all of the data that will be used in the analysis section of this thesis.

Table 3. Complete data set. Represents all the data used in analysis section (M denotes missing data). Cases 4 and 13 have been removed for reasons previously discussed.

Case	1	2	3	5	6	7	8	9	10	11	12.1	14	15	16	17	18
Tot Flashes	145	117	9	33	15	10	89	130	118	308	544	181	228	1024	108	220
Max VIL	43	40	20	39	34	31	40	48	40	49	37	49	47	88	38	46
VIL dt	20	39	10	33	15	14	82	41	52	101	52	108	56	73	51	48
Max Ref	60	59	57	66	57	58.5	59	66	64.5	61	62	65.5	67	68.5	63.5	63
L/O Max Ref	5800	5300	1300	6300	2100	1700	2600	5400	4200	2300	4300	1300	4800	4100	1300	6200
Ref dt	26	33	10	57	0	14	35	51	47	38	48	103	46	66	91	55
Peak Curr	98	98	26	81	74	58	130	82	58	157	101	83	102	107	111	64
PK Curr dt	12	28	0	6	5	3	9	31	55	104	73	56	10	34	51	27
PK Flash Rt	25	34	5	8	4	5	14	27	23	25	51	24	34	145	14	24
PK Flash dt	34	34	20	24	12	10	53	40	53	129	83	92	64	62	45	46
PK Pos Flash Rt	1	2	1	1	1	0	1	4	4	7	11	5	2	40	3	3
PK Pos Flash dt	18	30	20	7	22	M	53	40	58	83	93	77	29	57	49	36
PK Neg Flash Rt	25	33	4	8	4	5	13	23	21	24	43	21	34	112	13	22
PK Neg Flash dt	34	34	20	24	12	10	53	40	53	129	83	92	64	62	40	52

Case	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Tot Flashes	56	96	11	136	46	123	16	26	20	29	3	14	207	12	35	800
Max VIL	42	38	19	29	35	36	41	34	31	32	24	39	57	39	31	43
VIL dt	24	40	15	24	58	31	17	14	8	18	34	4	39	13	26	129
Max Ref	64	61	55	67.5	58	57.5	63.5	54	58.5	57	57.5	63.5	63.5	63.5	62.5	67
L/O Max Ref	5700	2200	1500	2900	2500	3800	4200	3700	2400	1600	1000	2700	8200	6100	4400	5100
Ref dt	30	80	15	24	40	25	17	28	3	13	22	4	45	18	41	134
Peak Curr	81	68	26	121	113	56	41	75	46	128	45	48	55	45	100	115
PK Curr dt	11	52	7	7	23	42	9	35	4	13	4	1	57	12	0	75
PK Flash Rt	22	15	2	17	10	24	8	8	9	7	1	7	22	5	9	68
PK Flash dt	16	36	26	23	35	42	12	20	11	16	20	3	31	10	30	89
PK Pos Flash Rt	0	2	1	2	1	1	0	1	1	1	0	1	2	1	1	21
PK Pos Flash dt	M	36	16	63	10	42	M	20	11	16	M	8	41	16	24	64
PK Neg Flash Rt	22	13	2	17	10	24	8	7	8	6	1	7	22	5	9	62
PK Neg Flash dt	16	36	26	23	35	46	12	20	11	16	20	3	31	10	30	84

Case	35	36	37	38	39	40	41	42.1	43	44	45	46
Tot Flashes	93	55	52	7	86	20	11	148	26	28	573	24
Max VIL	37	29	36	25	42	33	25	48	25	32	42	22
VIL dt	68	64	52	23	38	17	8	43	29	26	52	2
Max Ref	64	55.5	55	56	58	56.5	55.5	60.5	60.5	55	63	55.5
L/O Max Ref	4800	1200	4500	1800	4700	5100	5100	8200	6400	3600	5100	2100
Ref dt	68	69	26	17	18	17	13	28	29	26	47	2
Peak Curr	123	178	215	28	79	46	24	79	30	49	88	36
PK Curr dt	17	26	13	17	31	13	6	19	11	14	33	20
PK Flash Rt	14	10	10	3	21	9	5	22	6	12	68	10
PK Flash dt	70	55	44	8	34	15	5	20	15	23	39	9
PK Pos Flash Rt	2	2	2	0	9	1	2	8	1	0	8	0
PK Pos Flash dt	24	19	24	M	25	4	5	35	15	M	54	M
PK Neg Flash Rt	13	9	9	3	12	9	4	19	5	12	67	10
PK Neg Flash dt	70	55	44	8	34	15	7	20	10	23	39	9

4. Analysis and Results

4.1 Introduction

With the complete data set generated, it was possible to begin evaluating techniques for forecasting the cessation of lightning based on the lightning and radar parameters listed in Table 3. It should be noted that the data set of 44 thunderstorm cases represent a substantial increase in sample size from Hinson's research effort conducted in 1997, and though results may not be conclusive, they will be more statistically significant.

This chapter focuses on detailed analysis of these parameters along with the results. Specifically, the analysis involves performing statistical evaluations on the data to identify if any significant correlations exist between the parameters. Where correlations are significant, regressions are generated and evaluated for their usefulness in forecasting the cessation of lightning.

4.2 Analysis Tools

4.2.1 x-y Plots

An x-y plot or scatter plot is an efficient way of displaying paired data sets and allows for quick identification of many common patterns (linear, nonlinear) that may exist. For this reason, these plots are used extensively throughout this chapter. They are generated by plotting one value versus the other on a Cartesian grid in a 2-dimensional plane (see figures 11a – 11f). Ideally, if a pattern does exist, then a mathematical model may be derived to describe the relationship and ultimately be used as a model to forecast the cessation of lightning.

4.2.2 Least Squares Line

If the plotted points of a particular comparison tend to cluster about a straight line, i.e., linear, then the x-y plot will also be accompanied by a least squares line associated with the x and y values. Equation 1 is the mathematical representation of a least squares line, where y is the dependent variable, x is the independent variable, β_0 is the y intercept (2), and β_1 is the slope of the line (3). It should be noted that the simple linear regression models discussed in this research effort will also be expressed in this form (1).

$$y_{est} = \beta_0 + \beta_1 \cdot x \tag{1}$$

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \cdot \sum_{i=1}^{n} x_i}{n} \tag{2}$$

$$\beta_1 = \frac{n \cdot \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i \right)^2} \tag{3}$$

This linear equation (1) will serve as a single variable linear regression model for the plotted x-y values. The accuracy of each comparison for the regression line with respect to the actual data points will be evaluated using other statistical calculations such as standard deviation (s), coefficient of determination (R^2), and standardized residuals (e_s), all of which are described later in this section.

4.2.3 Standardized Residuals

Residuals represent the difference between the predicted value of the dependent variable (y_{est}) and the observed value (y) for the same point (4), and essentially

$$e = y - y_{est} \tag{4}$$

indicates the degree of error in the vertical direction at each given data point that is associated with the linear regression model (Wilks 1995).

Standardized residuals (5) can be calculated by dividing the residual (4) by the standard error of estimate (6) and are especially useful in determining whether the regression model is valid (Devore 1995). If a plot of standardized residual versus the independent variable does not display any distinct pattern (i.e., it's randomly distributed), and the majority of the plots lie between 2 and –2, then the regression model in question may accurately represent the data.

$$e_s = \frac{e}{se\_e(e)} \tag{5}$$

$$se\_e(e) = \sqrt{\frac{\left(\sum_{i=1}^{n} e\right)}{n-2}} \tag{6}$$

It is important to emphasize that standardized residual plots by themselves do not definitively show that a regression model is accurate. They're just one statistical tool that may be used in conjunction with others to help make this determination (Devore 1995). Specifically, the statistical tools that accompany the standardized residuals in this chapter

are the coefficient of determination (R^2), and plots of y_{est} (fitted values) versus y (observed values).

4.2.4 Coefficient of Determination

Assuming a normal distribution and an approximate linear relationship exists between the independent and dependant variable, then the coeffecient of determination (R^2) may be used to help identify how a particular regression model may perform. Devore (1995) defines R^2 as the proportion of observed dependant variable (y) variation that can be explained by the simple linear regression model.

R^2 can take on values that range from 0 to 1. Higher values of R^2 indicate that the regression model explains the variation in y very well. Conversely, lower values indicate that the regression model does not effectively explain the variation in y, and an alternative model should be used.

$$R^2 = 1 - \frac{SSE}{SST} \tag{7}$$

$$SSE = \sum_{i=1}^{n} \left(y_i - y_{est}\right)^2 \tag{8}$$

$$SST = \sum_{i=1}^{n} \left(y_i - y_{mean}\right)^2 \tag{9}$$

The calculation of R^2 (7) involves subtracting the product of the sum of squared errors (SSE) and the inverse of the total sum of squares (SST^{-1}) from 1. SSE and SST are defined by Devore (1995) as the sum of squared deviations about the least squares line (8) and the sum of squared deviations about their mean (9) respectively.

4.2.5 Standard Deviation

An estimate of the variance (σ^2) is obtained by dividing SSE by the degrees of freedom (n-2) of the estimate (10) and represents the average of all squared deviations from the mean (Devore 1995). The estimated standard deviation (s) is merely the positive square root of σ^2 (11) and is a convenient measure of dispersion. Each of these

$$\sigma^2 = \frac{SSE}{n-2} \tag{10}$$

$$s = \sqrt{\sigma^2} \tag{11}$$

estimated values will be used as a measure of the validity of the regression models. Larger values will be an indication that the regression model may not be a useful forecast tool.

4.2.6 Evaluation of Forecast Skill

Forecast skill is a useful measure of how effectively a given regression model performs and refers to the relative accuracy of a regression model compared to the accuracy of climatology (Wilks 1995). In this research effort, each regression model is compared directly to the climatological life span of a single-cell thunderstorm. This introduces some error since the regression each model is built using time occurrence of specific lightning and radar parameters as indicators of the impending cessation of lightning. The time occurrences of these parameters may or may not coincide with the time occurrence of the onset of electrification of the cell. Hence, the skill values presented in the analysis section should not be taken literally, but instead as a general indicator of the performance of the model.

The accuracy of the regression models and climatological averages can be described in terms of the mean absolute error (MAE) or the mean squared error (MSE), each of which are described below. In Equation 12, skill score (SS) is calculated in terms

$$SS_{MAE} = 1 - \frac{MAE}{MAE_{climo}} \tag{12}$$

of MAE (SS_{MAE}) and can be easily adjusted to be calculated in terms of MSE by merely exchanging MAE for MSE.

The two most common measures of forecast accuracy are MAE and MSE (Wilks 1995). Mathematically speaking, MAE is the average of the sum of absolute values of the difference between the forecast (y_{est}) generated by the regression model and the observed value (y) (13). MAE increases as the differences between forecasts and observations increase and can be interpreted as a typical magnitude for the forecast error (Wilks 1995). MAE_{climo} is analogous to MAE except y_{est} is replaced with climatological average (climo) (14).

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| y_{est_i} - y_i \right| \tag{13}$$

$$MAE_{climo} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| climo - y_i \right| \tag{14}$$

The MSE is the average of the sum of the squared differences between the forecast (y_{est}) generated by the regression model and the observed value (y) (15). The

squaring of the differences makes the MSE more sensitive to large errors than the MAE, but in order to express the typical magnitude of forecast error the square root of MSE (RMSE) must be calculated (16). The MSE_{climo} is analogous to MSE except y_{est} is replaced with climatological average (climo) (17).

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(y_{est_i} - y_i \right)^2 \tag{15}$$

$$RMSE = \sqrt{MSE} \tag{16}$$

$$MSE_{climo} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(climo - y_n \right)^2 \tag{17}$$

4.3 Simple Linear Regression Analysis

After viewing the initial x-y plots, it was discovered there were four cases that had anomalous numbers of flashes. Creating a box and whiskers plot (Fig.10) of the flashes identified the anomalous cases as outliers (Devore 1995) and these cases were subsequently eliminated from the study. The removal of these cases reduced the data set to 40 thunderstorms.

The remainder of this chapter focuses on evaluating the peak occurrence of a particular lightning or radar parameter and determining its usefulness, if any, in forecasting the cessation of lightning. Specifically, the radar parameters are peak VIL and reflectivity, and the lightning parameters are peak cumulative, positive, and negative flash rates and absolute maximum peak current.

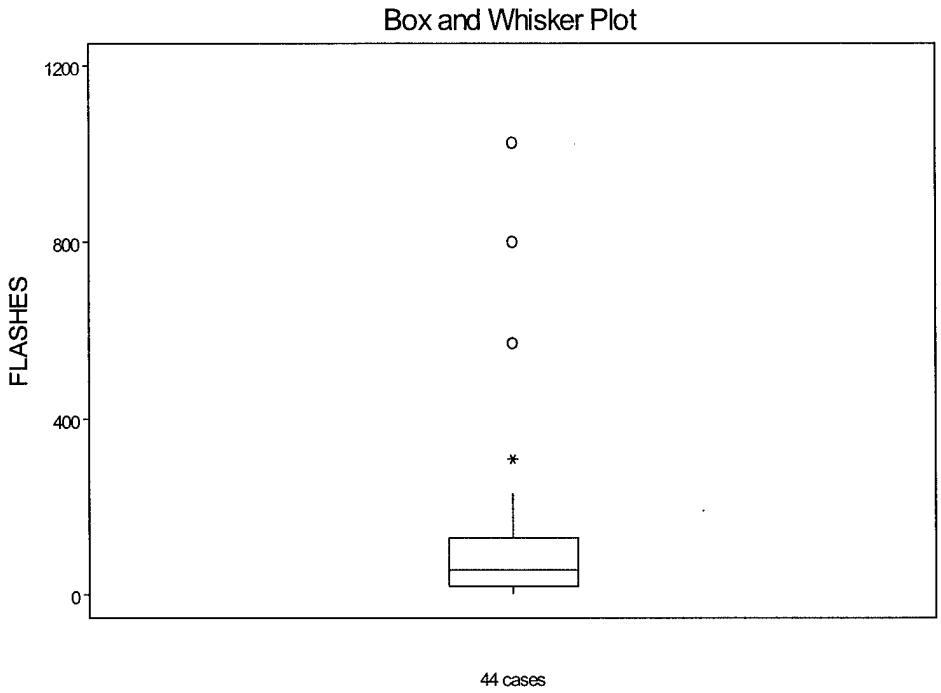

Figure 10. Box and whiskers plot. Depicts the total number of flashes for each thunderstorm case. * denotes outliers, and O denotes extreme outliers.

4.3.1 Analysis of Cases

An initial analysis of each case was conducted using x-y plots to look for any graphical evidence of a simple linear relationship that may exist between the respective paired data sets (Fig. 11a - 11f). Each of these figures was constructed by plotting the peak value in question (VIL, reflectivity, maximum peak current, etc) recorded for each thunderstorm (x axis) versus the time difference between the occurrence of that particular peak value and the last flash of the thunderstorm (y axis). Overlaid on each of these plots is a least squares line (discussed earlier), which represents for each case the simplest linear regression model that can be constructed from the respective data sets. Note, β_0 (y intercept) and β_1 (slope) of the regression models are given in Table 4.

Table 4. Simple linear regression variables. Specific slope (β_1), y intercept (β_0), standard deviation (s) and coefficient of determination (R^2) for each case.

Case	β_0	β_1	R^2	s
Peak VIL vs. Peak VIL dt	-7.68	1.17	0.26	20
Peak Reflectivity vs. Peak Reflectivity dt	-20.7	0.91	0.14	22
Peak Current vs. Peak Current dt	15	0.07	0.03	19
Peak Flash Rate vs. Peak Flash Rate dt	10.94	1.37	0.46	16
Peak Pos Flash Rate vs. Peak Pos Flash Rate dt	16.28	3.92	0.21	18
Peak Neg Flash Rate vs. Peak Neg Flash Rate dt	11.05	1.48	0.44	17

Recall that one method of measuring the validity of a regression model is to plot the standardized residuals. Figures 12a through 12f are the standardized residual plots for each comparison, and in each case, they are randomly distributed, and only a few values exceed 2. Thus, each of the regressions is considered statistically valid, but as mentioned earlier, that does not necessarily indicate that it's accurate.

By visually inspecting each of the x-y plots and noting the modest values for the coefficients of determination (R^2) and the high standard deviations (s) in Table 4, it can be concluded that the regression models do not completely represent the data accurately however, some represent the data better than others. Specifically, the comparisons involving the peak cumulative and peak negative flash rates produced least squares lines indicated in each case an increase in flash rate corresponded to an increase in the time interval between the occurrence of that peak and the last flash. Additionally, the R^2 values for these two cases were greater than 0.4, but their standard deviations were

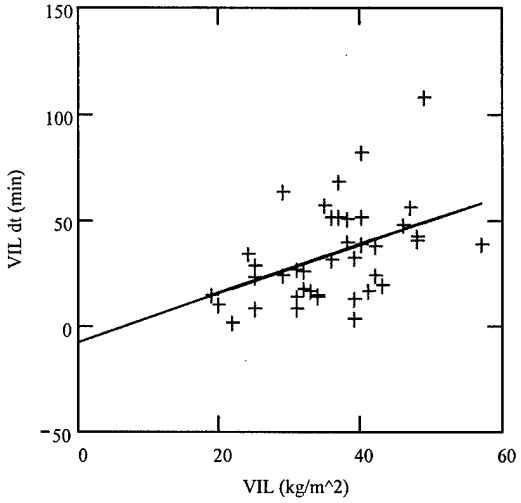

Figure 11a. x-y Plot of Peak VIL vs. VIL dt.

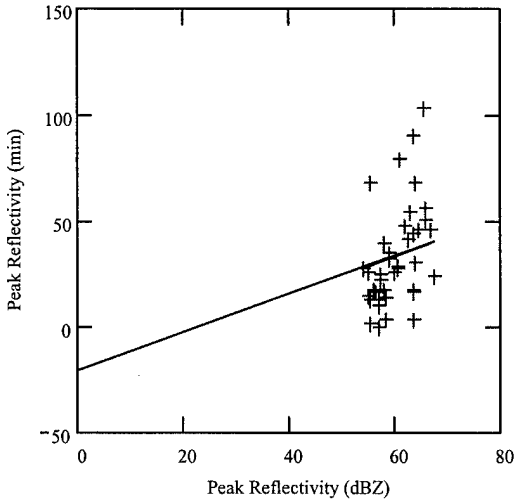

Figure 11b. x-y Plot of Peak Reflectivity vs. Reflectivity dt.

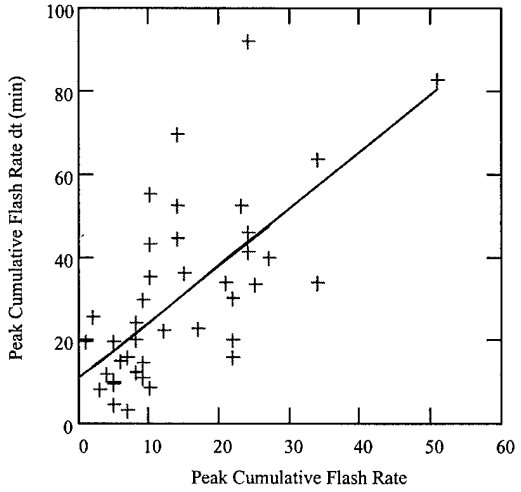

Figure 11c. x-y Plot of Peak Flash Rate vs. Peak Flash Rate dt.

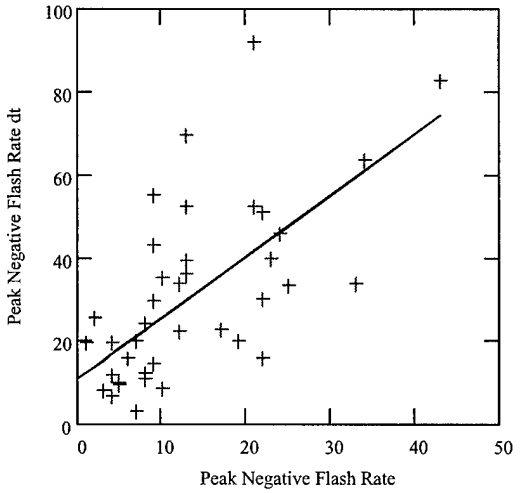

Figure 11d. x-y Plot of Peak Negative Flash Rate vs. Peak Negative Flash Rate dt.

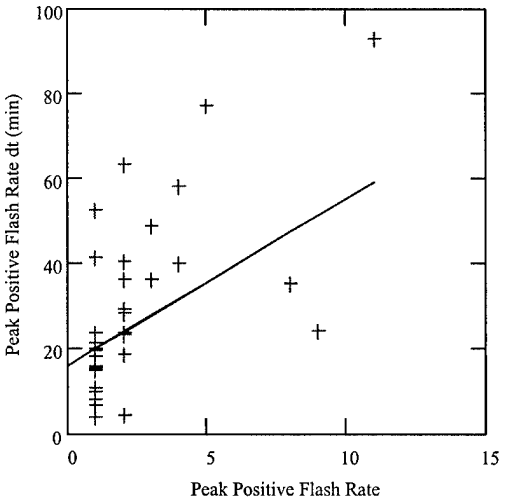

Figure 11e. x-y Plot of Peak Positive Flash Rate vs. Peak Positive Flash Rate dt.

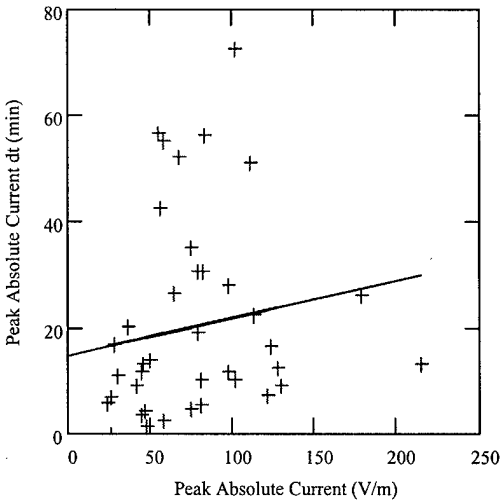

Figure 11f. x-y Plot of Maximum Peak Current vs. Maximum Peak Current dt.

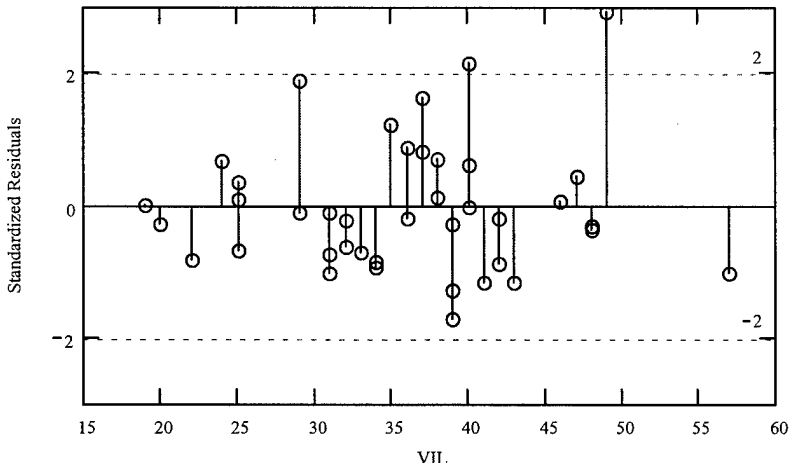

Figure 12a. Standardized Residual Plot for VIL Case.

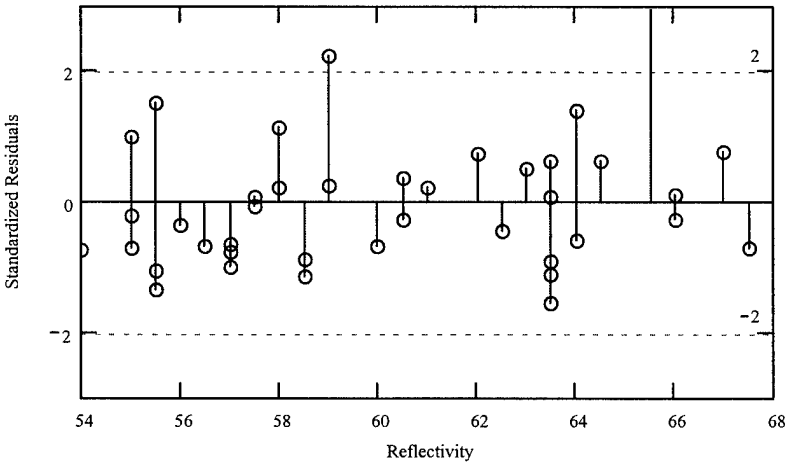

Figure 12b. Standardized Residual Plot for Reflectivity Case.

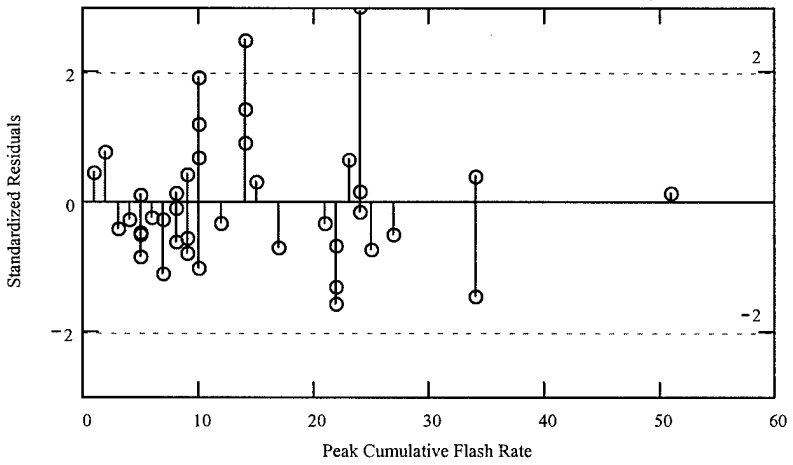

Figure 12c. Standardized Residual Plot for Peak Flash Rate Case.

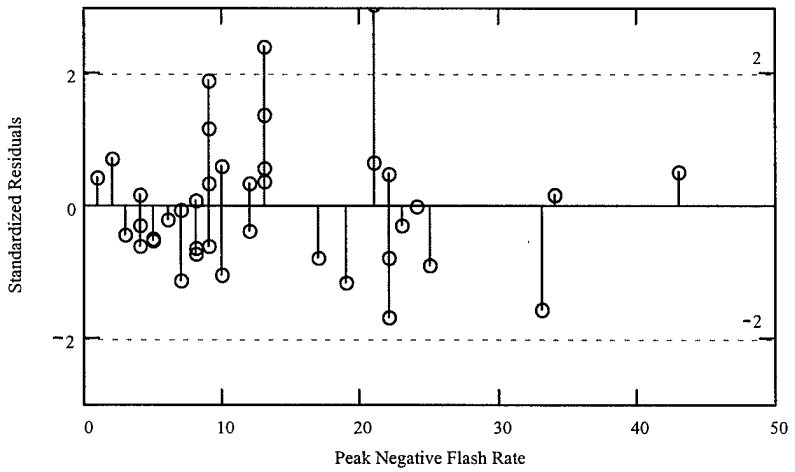

Figure 12d. Standardized Residual Plot for Peak Negative Flash Rate Case.

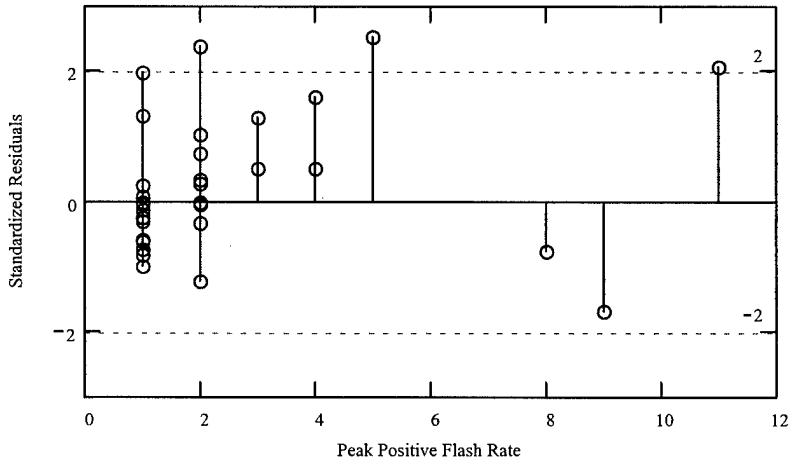

Figure 12e. Standardized Residual Plot for Peak Positive Flash Rate Case.

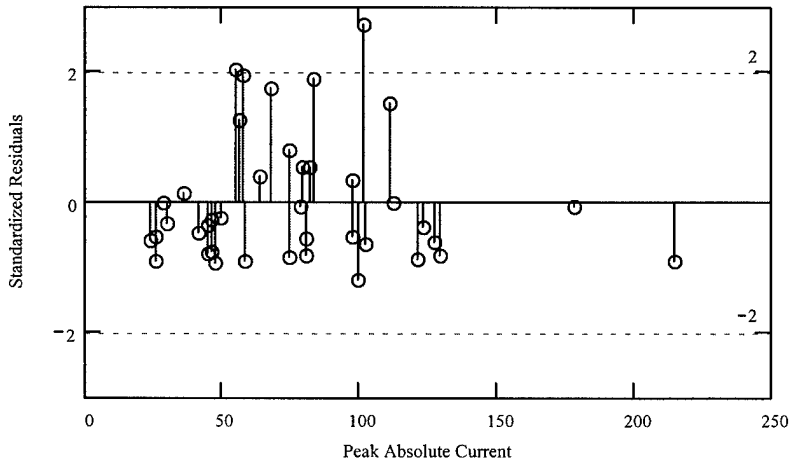

Figure 12f. Standardized Residual Plot for Max Peak Current Case.

still on the order of 15 minutes. To a lesser degree, the comparison involving peak VIL also produced a least squares line that described the general trend of the data, but this case only produced a modest R^2 value of 0.26 and a relatively high standard deviation of 20 minutes. The remaining comparisons produced x-y plots, R^2 and standard deviations that were not consistent with linear behavior.

Additional x-y plots were generated using the exact same data as previously noted. However, these new sets of plots (Fig. 13a –13f) were slightly different from the initial set (Fig. 12a –12f) because they distinguished between single and multi-cell thunderstorms (plus and circle symbols respectively) and were plotted on logarithmic scales. Each plot (excluding the positive flash rate comparison) consists of 18 single-cell cases and 22 multi-cell cases. Additionally, they graphically illustrate the distinction that can be drawn between multi-cell and single-cell thunderstorms. Clearly, these figures suggest that evaluating each thunderstorm cell type independently may, but not necessarily, produce a more accurate regression model. Such an evaluation was accomplished for each comparison, and some of the results are displayed in Table 5.

When the values in Table 5 are compared to the original values in Table 4 it is evident that there are substantial differences between the evaluation methods. Of particular interest are the comparisons displaying the highest degree of linearity. Those comparisons were peak cumulative flash rate, peak negative flash rate, and peak VIL. For each of these cases, the standard deviations of the regressions generated using

Table 5. Simple linear regression variables distinguished by thunderstorm type. Specific slope (β_1), y intercept (β_0), standard deviation (s) and coefficient of determination (R^2) for each case and cell type.

Comparison	Cell Type	β_0	β_1	R^2	s
Peak VIL vs. Peak VIL dt	Single	6	0.35	0.11	10
	Multi	11	0.89	0.18	21
Peak Reflectivity vs. Peak Reflectivity dt	Single	1.1	0.25	0.12	10
	Multi	-4	0.85	0.21	22
Peak Current vs. Peak Current dt	Single	3	0.14	0.19	8
	Multi	47	-0.19	0.03	21
Peak Flash Rate vs. Peak Flash Rate dt	Single	9	0.68	0.34	7
	Multi	21	1.08	0.28	18
Peak Pos Flash Rate vs. Peak Pos Flash Rate dt	Single	15	0.35	0	8
	Multi	26	2.53	0.13	22
Peak Neg Flash Rate vs. Peak Neg Flash Rate dt	Single	9.2	0.69	0.34	7
	Multi	27	0.99	0.27	18

the single-cell data set were smaller than those that were generated using the original data set which made no distinction between cells. Conversely, the standard deviations for the regressions built using the multi-cell data set were all greater that the original set of regressions. This result is not surprising since a single-cell thunderstorm typically has a life span (in terms of minutes) with a lower bound near 30 and an upper bound near 50. This is not necessarily the case for multi-cell thunderstorms. They may be nearly as short lived as a single-cell thunderstorm or may persist for hours. Therefore, the lifespan of a multi-cell thunderstorm can take on a larger range of values than that of a single-cell thunderstorm and subsequently is more likely to produce a larger standard deviation.

Another interesting difference between the two tables is that for the cases of interest each of the β_0 (y intercept) values in Table 5 intersect the positive y axis (positive

time). This is important since these regression models will be used as forecast tools and a physically realistic result of positive time is desired.

The last notable difference between the tables is that in each case, the R^2 values for both the multi and single-cell cases in Table 5 are smaller than the original values in Table 4. This indicates a decrease in the linear correlation between the two variables in each case and subsequently lowers the confidence in the regression models.

A true measure of the accuracy of the regression model is to plot the predicted values versus the observed values. The more accurate the model is the more closely the plot will mirror a line with a slope of one (i.e., a 45° line). Figures 14a through 14f are plots of this type broken up by cell. A brief inspection of these figures reveals for each case that neither the single nor the multi-cell regressions bear any resemblance to a 45° line. In fact, in all cases, the slopes of the plots approach zero indicating an inaccurate regression model. Additionally, in each case, the regression models are an over estimate of the smaller observations and an under estimate of the larger. This behavior is consistent through each case and doesn't appear to be any less a factor in the previously cited cases where the degree of correlation was noted as among the highest (i.e., peak flash rate, peak negative flash rate, and VIL).

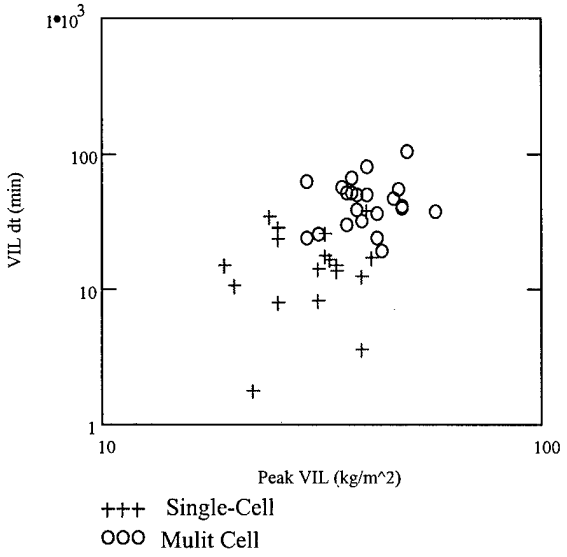

Figure 13a. Logarithmic Plot of Peak VIL vs. VIL dt.

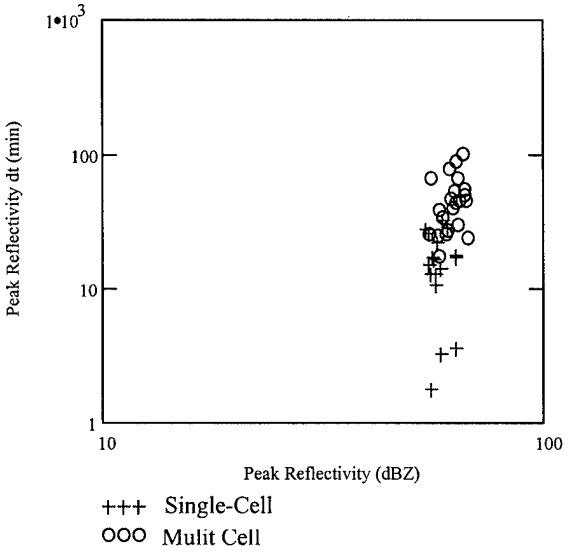

Figure 13b. Logarithmic Plot of Peak Reflectivity vs. Reflectivity dt.

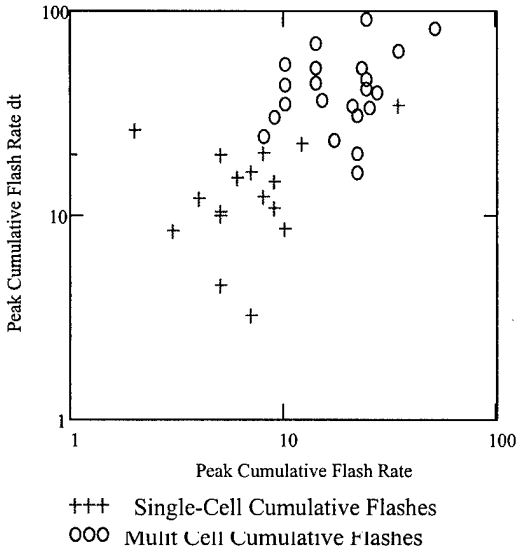

Figure 13c. Logarithmic Plot of Peak Flash Rate vs. Peak Flash Rate dt.

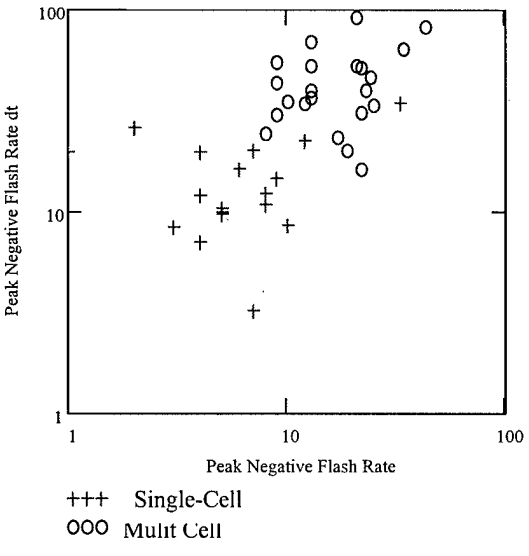

Figure 13d. Logarithmic Plot of Peak Neg Flash Rate vs. Peak Neg Flash Rate dt.

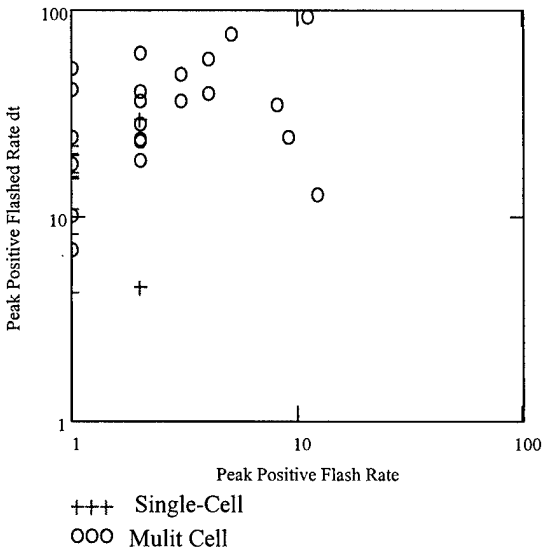

Figure 13e. Logarithmic Plot of Peak Pos Flash Rate vs. Peak Pos Flash Rate dt.

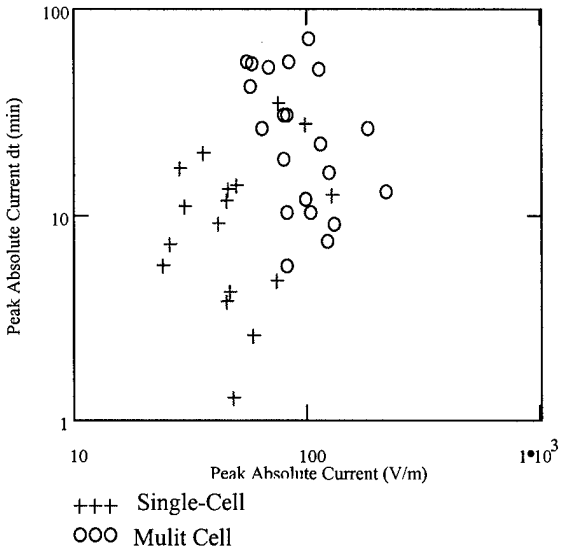

Figure 13f. Logarithmic Plot of Max Peak Current vs. Max Peak Current dt.

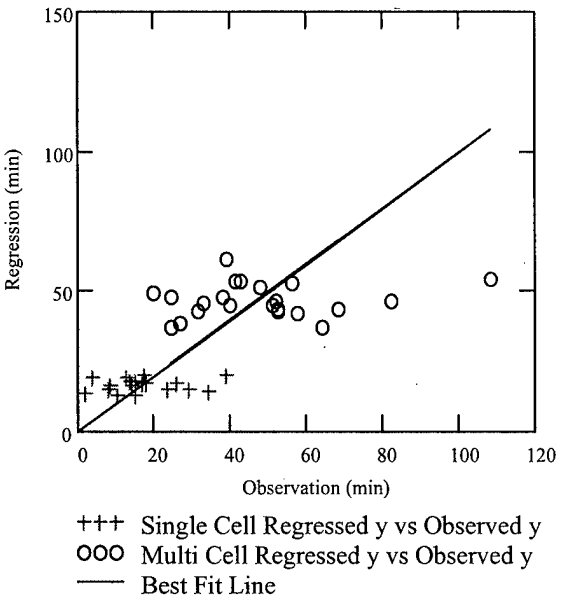

Figure 14a. Comparison of Forecast vs. Observation for VIL Case.

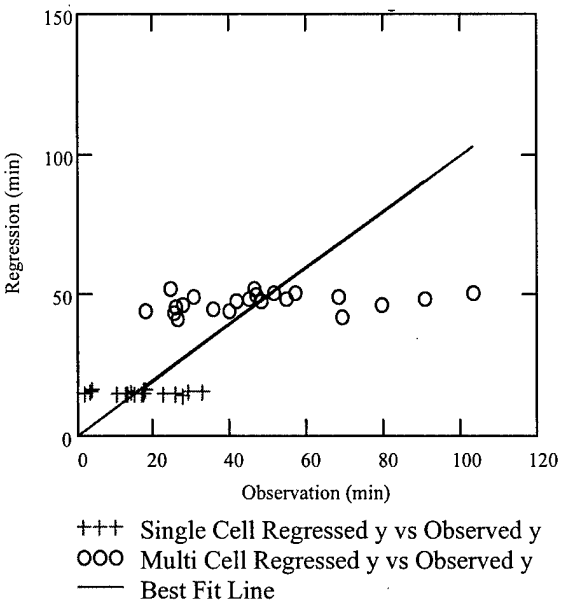

Figure 14b. Comparison of Forecast vs. Observation for Reflectivity Case.

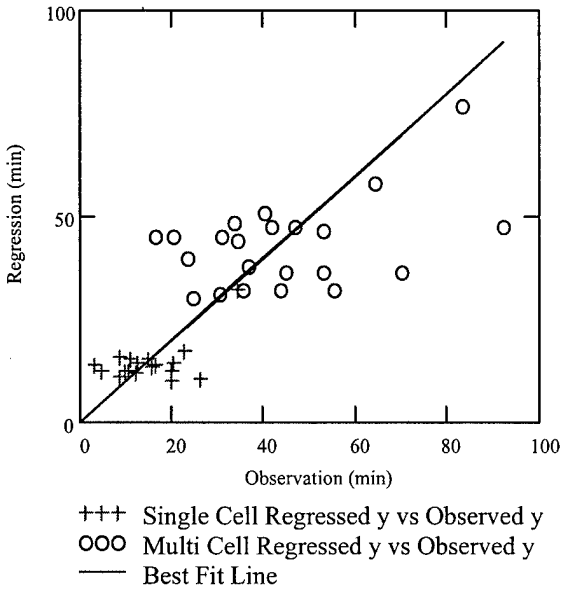

Figure 14c. Comparison of Forecast vs. Observation for Peak Flash Rate Case.

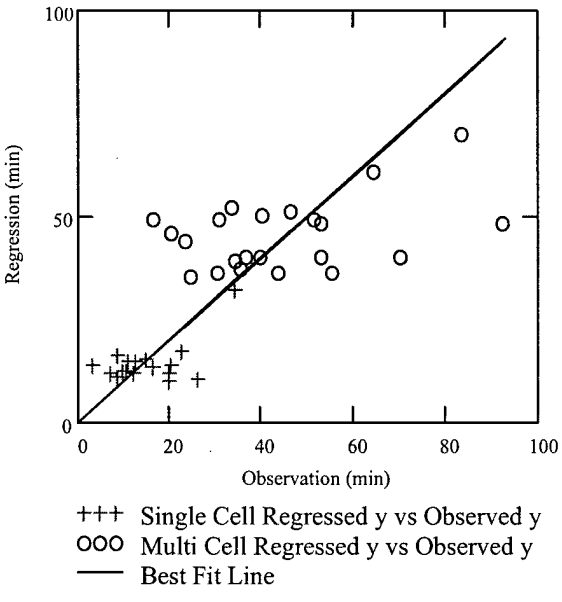

Figure 14d. Comparison of Forecast vs. Observation for Peak Neg Flash Rate Case.

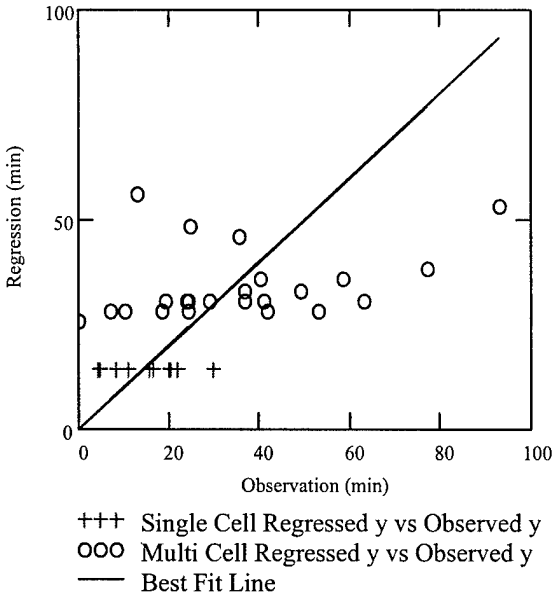

Figure 14e. Comparison of Forecast vs. Observation for Peak Pos Flash Rate Case.

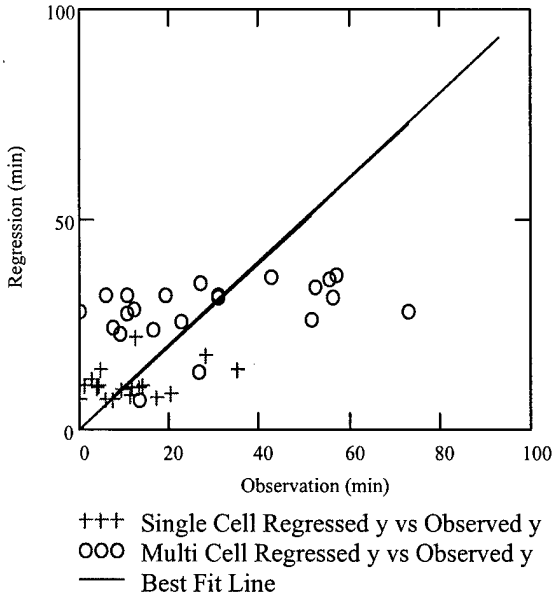

Figure 14f. Comparison of Forecast vs. Observation for Max Peak Current Case.

4.3.2 Skill Evaluation of Simple Linear Regressions

The previous section focused on determining the validity of each regression model along with an initial estimate of the accuracy based on R^2, standard deviations, and finally x-y plots of regressed values versus observed values. This section has a slightly different focus. It will explicitly concentrate on the ability of a particular regression model to forecast the cessation of lightning by calculating the regression's skill and comparing it to that of the climatological skill.

The initial findings discussed in the previous section suggest that the regression models would, at best, produce marginal results. However, after calculating each regression's skill as measured by comparing it to that of the climatological accuracy, it was evident that the regressions proved in each case to be a better forecasting tool than climatology. It should be noted that for reasons previously discussed, only single-cell thunderstorms have clear climatological life spans. For this reason, only the regression models built using the single-cell thunderstorm data set (18 cases) along with the complete data set (40 cases) are used in the skill evaluation. The regressions based solely on the multi-cell thunderstorms data set (22 cases) were excluded since no representative climatological average exists. Figures 15a through 15f are plots of each case and graphically illustrate the increase in forecast skill of each regression. The y-axis of these plots represents the percent increase or decrease in forecast accuracy of the regression model versus climatology and the x-axis is merely the theoretical climatology in

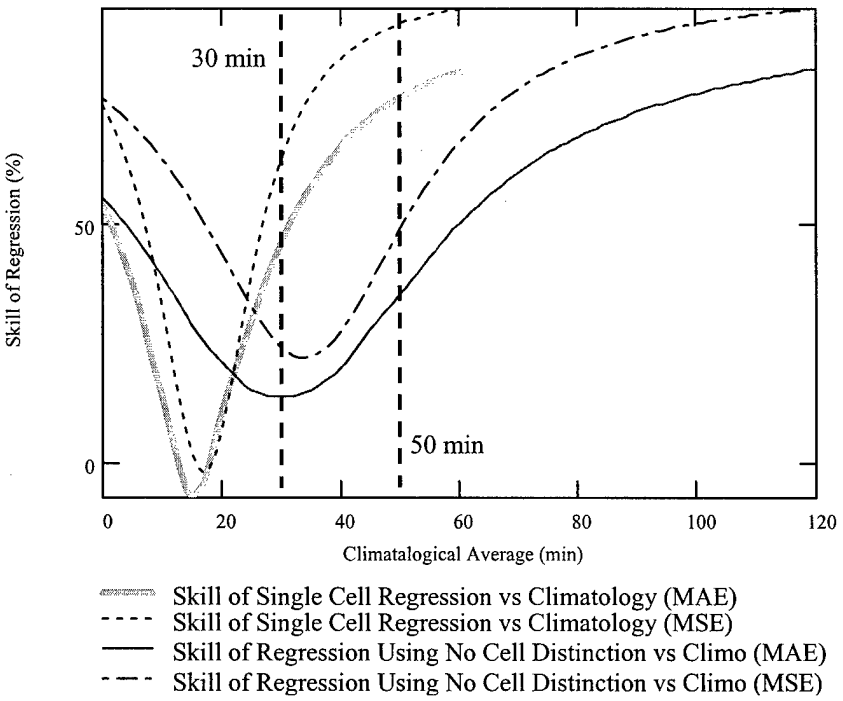

Figure 15a. Skill of VIL Based Regression vs. Climatology.

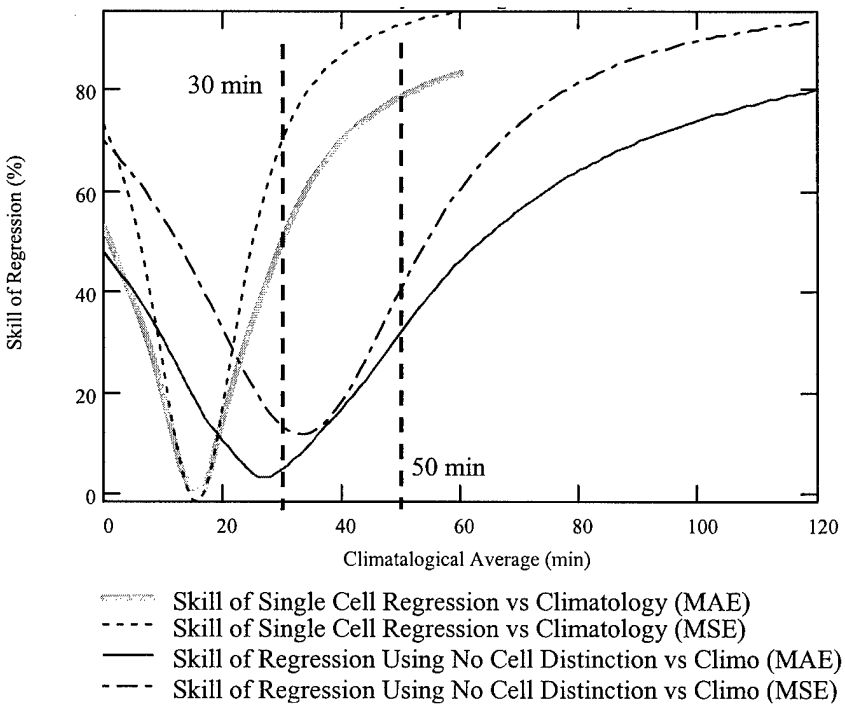

Figure 15b. Skill of Reflectivity Based Regression vs. Climatology.

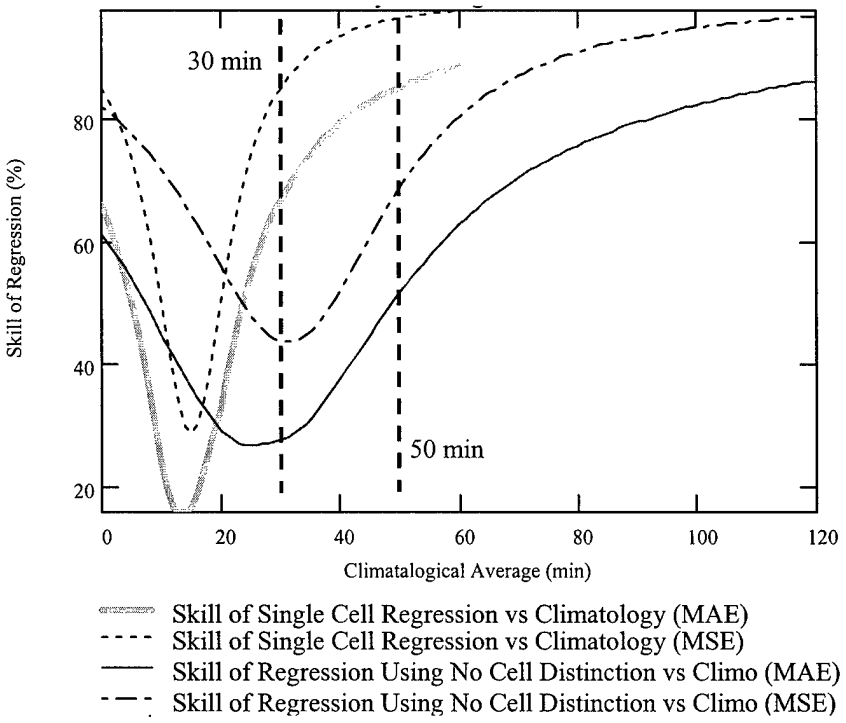

Figure 15c. Skill of Peak Flash Rate Based Regression vs. Climatology.

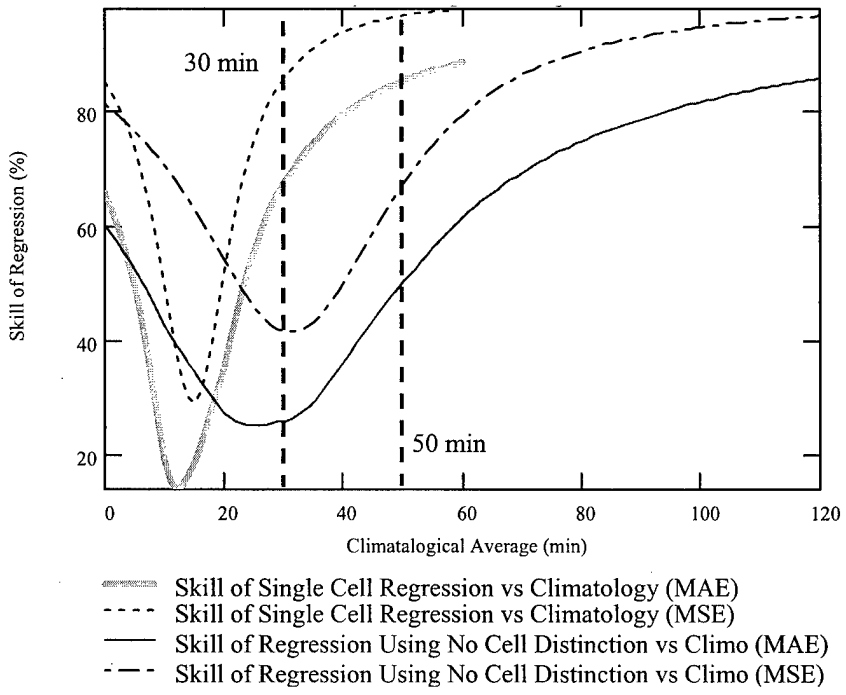

Figure 15d. Skill of Peak Neg Flash Rate Based Regression vs. Climatology.

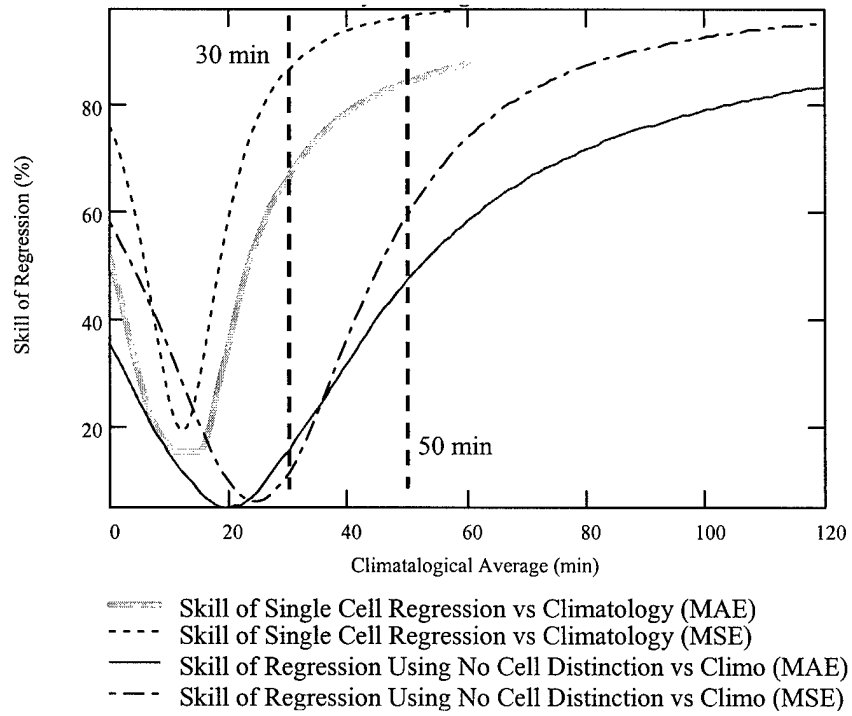

Figure 15e. Skill of Peak Pos Flash Rate Based Regression vs. Climatology.

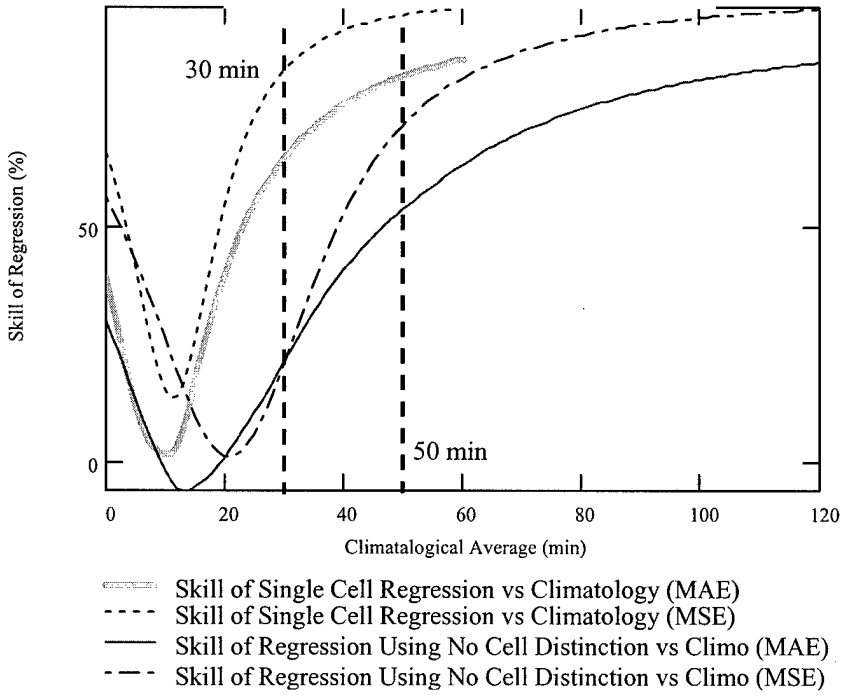

Figure 15f. Skill of Max Peak Current Based Regression vs. Climatology.

units of minutes with the actual climatological average of a single-cell thunderstorm identified as the area that falls between the two dashed vertical lines. The remainder of this section will focus strictly on this small area between the vertical lines. Table 6 summarizes some of the more important results.

Table 6. Summary of skill values for each simple linear regression.

Comparison	Skill @ 30 min (%)				Skill @ 50 min (%)				RMSE	MAE
	Single		No distinction		Single		No distinction			
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE		
VIL	46	64	36	49	77	92	14	25	9.5	7.3
Reflectivity	51	70	32	41	79	93	5	14	9.1	6.9
Current	64	83	21	21	82	95	54	72	8.2	6.5
Flash Rate	67	86	52	44	86	97	28	69	6.2	4.8
Pos Flash Rate	67	87	16	11	84	96	48	60	6.9	5.6
Neg Flash Rate	68	86	26	42	86	97	51	68	6.1	4.8

The skill values presented in Table 6 represent the percent increase in forecast skill over climatology calculated at eight discrete points on each of the figures above. Skill values are calculated four times at the climatological minimum life span of a single-cell thunderstorm (i.e., where the 30 minute vertically dashed line intersects the mean absolute error (MAE) and mean squared error (MSE) curves of each regression). Skill values are calculated four more times at the climatological maximum of a single-cell thunderstorm (i.e., where the 50 minute vertically dashed line intersects the MAE and MSE curves of each regression).

By carefully evaluating Figures 15a through 15f, one can identify three common features. First, all of the regression models produce an increase in forecast skill over climatology. Second, for each case the regressions built using the single-cell thunderstorm data set not only out performed climatology, but they also consistently out

performed the regressions built using the complete data set. Lastly, the method of calculating forecast skill using MSE consistently out performed the MAE method.

An evaluation of the data presented in Table 6 also produced some expected and unexpected results. For instance, the peak flash rate and peak negative flash rate cases produced the lowest MAE and RMSE, and that's not surprising since they displayed the highest linear correlation. But, what is surprising is that these two cases did not always produce the greatest increase in forecast skill although they did produce among the greatest. Another unexpected result was that the peak positive flash rate and maximum peak current cases consistently produced among the highest increases in forecast skill. In fact, these cases out performed the VIL case in almost every comparison.

4.4 Multiple Linear Regressions

A description of multiple linear regressions is very similar to that of the simple linear regression previously discussed. Essentially, the difference is that multiple regressions relate two or more independent variables to a single dependant variable. They typically are written in the form of Equation 18 where x_1 and x_2 are the independent variables and y_{est} is the dependent variable. It should be noted that it's possible to evaluate more than two independent variables, but with time being the limiting factor in this research effort, only regressions involving two independent variables were considered.

$$y_{est} = \beta_0+\beta_1 \cdot x_1+\beta_2 \cdot x_2 \tag{18}$$

$$\text{Regression A} = 10.04+1.915 \cdot x_1-1.346 \cdot x_2 \tag{19}$$

$$\text{Regression B} = 12.22+2.478 \cdot x_1-1.279 \cdot x_2 \tag{20}$$

$$\text{Regression C} = 9.9+1.866 \cdot x_1-.009 \cdot x_2 \tag{21}$$

The analysis portion of this section involved evaluating twelve different combinations of the lightning and radar parameters. These different combinations were then used to generate twelve independent models from each of the single-cell, multi-cell and no distinction data sets. In all, 36 regression models were created. R^2 values for each of the 36 regressions were then calculated. The model from each of the three data sets (single-cell, multi-cell, and no distinction) with the highest R^2 value was singled out to be evaluated for its usefulness in forecasting the cessation of lightning. Specifically, regressions A, B, and C (19, 20 and 21 respectively) were generated using single-cell, no distinction, and multi-cell data sets respectively. A summary of the results is listed in Table 7 and a plot of the skill of each regression versus climatology is shown in Figure 16. It should be noted that the skill values presented in Table 7 and Figure 16 were calculated using the MSE method since it produced the greatest increase in forecast skill over climatology. This analysis technique is slightly different than in the previous comparison where both MSE and MAE were used to calculate the skill values. Referring back to the Table 6, recall that the skill values calculated using MSE were consistently higher than those calculated using the MEA method. The same is true for these regressions so, to reduce clutter on the plots, the skill evaluation using MAE was excluded.

By comparing the data in Table 7 for regressions A and B to the data in Table 6 generated using a simple linear regressions it's possible to evaluate which technique of linear regression (simple or multiple) produces the highest increase in forecast skill.

Table 7. Summary of skill values for each multiple linear regression. Dependant (y_{est}) and independent (x_1, x_2) variables used to generate each multiple regression and their associated skill scores at the location that each intersects the vertically dashed lines in Figure 16.

Comparison	Skill @ 30 min (%)	Skill @ 50 min (%)	RMSE	x_1	x_2	y_{est}
Regression A (single-cell)	86	97	7.55	Peak Flash Rate	Peak Neg Flash Rate	Peak Flash Rate dt
Regression B (multi-cell)	45	70	15.35	Peak Flash Rate	Peak Neg Flash Rate	Peak Flash Rate dt
Regression C (no distinction)	71	49	14.11	Peak VIL	Height of Max Ref	Peak VIL dt

When this was accomplished, it was evident that regressions A and B offer no substantial improvement over the best performing simple linear regression models. In fact, compared to the peak flash rate case (Table 6), the skill values are almost identical and the RMSE is actually slightly higher for regressions A and B.

While a comparison of the data in Table 7 to that of Table 6 does not indicate that multiple linear regression models are any more accurate that simple linear regression models, the plotted data (Fig. 16) does illustrate than there are substantial differences between cell types. Each of the skill curves presented in Fig. 16 has a similar shape, but all have different minimums (with respect to time). These minimums are unique to the data set used to generate them and clearly reinforce an earlier conclusion that each thunderstorm type (single or multi-cell) must be treated individually. This is an important result since it explicitly shows that knowing the type of thunderstorm can substantially improve the forecast skill. For example, if a forecaster was trying to predict the cessation of lightning within a particular thunderstorm and he/she knew it was of the single-cell variety, then a prudent forecaster would maximize his/her forecast skill by applying a single-cell regression model (regression A). An improvement in forecast skill

could also be gained if regressions B or C were used, but clearly, they don't offer the same improvement as regression A that was built using the single-cell data set.

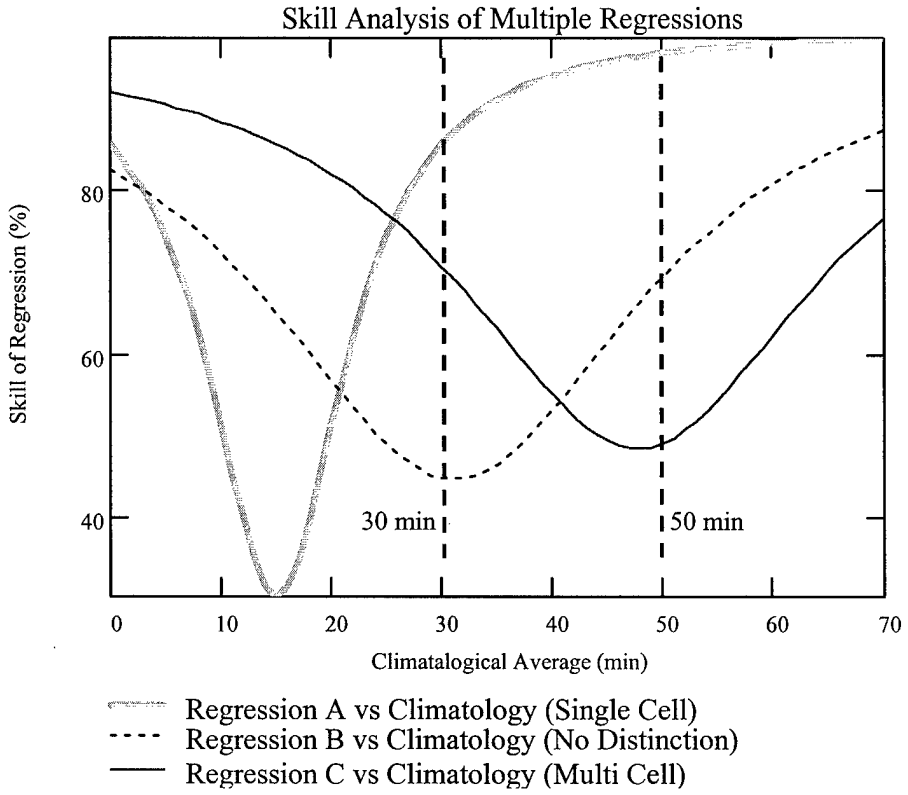

Figure 16. Skill of Each Multiple Linear Regression vs. Climatology.

5 Summary and Conclusions

Thunderstorms and more specifically the CG lightning flashes that they produce, place Air Force (AF) personnel and assets at risk every time they occur near or over an AF installation. To counter this risk, personnel and assets are justifiably removed from the area or delayed in entering the area. In either case, the effect is a decrease in operational efficiency, which ultimately results in a loss of tax dollars. The ability to predict the lightning producing life span of an individual thunderstorm would allow for a more timely execution of AF operations and in the process save large sums of US tax dollars.

Forecasting the life span of the lightning producing portion of a thunderstorm is a very challenging topic. The focus of this particular research effort was directed toward identifying methods of forecasting the cessation of lightning by observing specific lightning and radar parameters. Additionally, a special emphasis was placed on generating a large enough sample size of thunderstorm cases to make the results as statistically significant as possible. Subsequently, a large portion of this research effort was devoted to generating the thunderstorm data set of 46 cases that was later reduced to 40.

The analyses of the data set involved using x-y plots to identify any correlations and linear or nonlinear patterns that exist. Three cases evaluated using simple linear regression techniques exhibited some linear behavior. Specifically the cases involving peak flash rate (Fig. 11c), peak negative flash rate (Fig. 11d), and peak VIL (Fig. 11a)

each produced least squares lines that were reasonably consistent with the plotted data and had R^2 values that were among the highest calculated (Table 4).

A further evaluation of the usefulness of simple linear regression models was conducted by differentiating between single and multi-cell thunderstorms in the data set. An additional set of regression models built using the newly separated single-cell data set was generated (Table 6) and compared to the initial regression models which were built without distinction between cells. Results were mixed. The R^2 values of the single-cell regressions actually decreased in comparison to the initial regressions in every instance. This may in part be due to the small sample size of 18 thunderstorm cases, but an exact cause was not determined. Conversely, forecast skill as compared to climatology for the regressions generated using the single-cell data set increased substantially compared to that of the regressions generated using no distinction between cells (Fig. 15a – 15f). This result illustrated that the different thunderstorm cell types exhibited substantially different behavior with respect to forecasting the cessation of lightning.

Multiple linear regressions were also generated using various combinations of the lightning and radar parameters. Skill values of the models generated using the combinations that produced the highest R^2 values for each data set (Table 7) were compared to the best performing simple linear regression models. Results indicated that while multiple linear regression models equaled the skill performance of the simple linear regression models, they didn't offer any substantial improvement over them. Additionally, a plot of forecast skill of the multiple regression models (Fig. 16) further reinforced the notion that each specific thunderstorm type is unique from the other with respect to its relationship to forecasting the cessation of lightning. Thus, if one is trying to

forecast the cessation of lightning, and the specific type of thunderstorm cell is known, then one could substantially increase his or her forecast skill by utilizing a cell specific regression model.

In summary, providing an accurate method to forecast the cessation of lightning is a very difficult task and will most likely take several more attempts before the problem is solved. This research effort has evaluated some previously untried methods for forecasting the cessation of lightning, and while the results have in no way indicated that the problem has been solved, they did, however, produce substantial increases in forecast skill over that of the skill of climatology. Beyond showing that regression models built using observable parameters (lightning and radar) consistently outperform climatology, this research effort also contributes a reasonably large sample size of that can be used in following research efforts.

5.1 Recommendations for Further Research

Applying the techniques identified by Hinson (1997) to the thunderstorm data set generated in this research project may prove to be the best approach since it would increase the sample size of his method and thereby yield a more statistically significant comparison to the results described in this research effort. This would be time consuming and require an efficient method of evaluating certain radar reflectivity values at discrete heights.

Another method would be to evaluate the usefulness of LDAR data. This lightning detection system is unique to the KSC area and provides a three-dimensional view of lightning data. Evaluating LDAR data rates at specific heights (corresponding to peak reflectivity and VIL values) may provide some insight into solving this problem.

There are some variations of this research effort that could also prove to be worthwhile, for example, determining if weighting the regression models improves the forecast. Another variation would be to generate multiple linear regression models that have more than two independent variables.

Bibliography

Bluestein, H. B., 1993: *Synoptic-Dynamic Meteorology in the Midlatitudes*, vol. II, *Observations and Theory of Weather Systems*, Oxford Univ. Press, 594 pp.

Boyd, B.F., W.P. Roeder, J.B. Lorens, D.S. Hazen, and J.W. Weems, 1995: Weather support to pre-launch operations at the Eastern Range and Kennedy Space Center. *Sixth Conf. on Aviation Wea. Sys.*, Dallas, TX, 135-140

Buechler, D. E., and S. J. Goodman, 1990: Echo size and asymmetry: Impact on NEXRAD storm identification. *J. Appl. Meteor.*, **29**, 962-969.

Byers, H. R., and R. R. Braham, 1949: *The Thunderstorm*. Supt. of Documents, U. S. Government Printing Office, Washington D.C., 287 pp.

Byers, H. R., and H. R. Rodebush, 1948: Causes of thunderstorms on the Florida peninsula. *J. Meteor.*, **5**, 275-280.

Carey, L. D., and S. A. Rutledge, 1998: Electrical and multiparameter radar observations of a severe thunderstorm. *J. Geophys. Res.*, **103,** 13979-14000.

Cummings, K.L., M.J. Murphy, E.A. Bardo, W.L. Hiscox, R.B. Pyle, and A.E. Pifer, 1998: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network. *J. Geophys. Res.*, **103,** 9035-9044.

Devore, J. L., 1995: *Probability and Statistics for Engineering and the Sciences,* Fourth Edition., Duxbury Press, 743 pp.

Gremillion, M.S., and R.E. Orville, In Press: Thunderstorm characteristics of cloud-to-ground lightning at the Kennedy Space Center, Florida: A study of lightning initiation signatures as indicated by the WSR-88D. In press, *Wea. Forecasting.*

Hinson, M. S., 1997: A study of the characteristics of thunderstorm cessation at the NASA Kennedy Space Center. *M.S. Thesis, Texas A&M University*, 90 pp.

Idone, V. P., and D. A. Davis, P.K. Moore, Y. Wang, R.W. Henderson, M. Reis, and P.F. Jamason, 1998: Performance evaluation of the U.S. National Lightning Detection Network in eastern New York. *J. Geophys. Res.*, **103,** 9045-9055.

Jameson, A. R., M. J. Murphy, and E. P. Krider, 1996: Multi-parameter radar observations of isolated Florida thunderstorms during the onset of electrification. *J. Appl. Meteor.*, **35**, 343-354.

Larsen, H. R., and E. J. Stansbury, 1974: Association of lightning flashes with precipitation cores extending to a height of 7 km. *J. Atmos. Terr. Phys.*, **36**, 1547-1553.

Lhermitte, R., P. R. Krehbiel, 1979: Doppler radar and radio observations of thunderstorms. *IEEE Trans. Geo. Elec.*, **17**, 162-171.

Lopez, R. Z., J. P. Aubagnac, 1997: The lightning activity of a hailstorm as a function of changes in its microphysical characteristics inferred from polarimetric radar observations. *J. Geophys. Res.*, **102,** 16799-16813.

MacGorman, D., and T. Filiaggi, 1997**:** Lightning ground flash rates relative to radar-inferred storm properties. *Preprints, 28th Conf on Radar Meteor.*, 7 – 12 Sept 1997, Austin, TX, 143-144.

MacGorman, D. R., and D. W. Rust, 1998: *The Electrical Nature of Storms*. Oxford Univ. Press, 422 pp.

Marshal, J. S., and S. Radhakant, 1978: Radar precipitation maps as lightning indicators. *J. Appl. Meteor.*, **17**, 206-212.

Neumann, C.J., 1971:The thunderstorm forecasting system at the Kennedy Space Center. *J. Appl. Meteor.*, **10**, 921-936.

Orville, R. E., 1991: Lightning ground flash density in the contiguous United States - 1989. *Mon. Wea. Rev.*, **119,** 573-577.

_____, 1994: Cloud-to-Ground lightning flash characteristics in the contiguous United States: 1989-1991. *J. Geophys. Res.*, **99**, 573-577.

Orville, R. E., and A. C. Silver, 1997: Lightning ground flash density in the contiguous United States. 1992-1995. *Mon. Wea. Rev.*, **125,** 631-638.

Rinehart, R. E., 1997: *Radar for Meteorologists*, Third edition., Rinehart Publications, 428 pp.

Rutledge S. A., and L. D. Carey, 1997: Lightning ground flash rates relative to radar-inferred storm properties. *Preprints, 28th Conf on Radar Meteor.*, 7 – 12 Sept 1997, Austin, TX, 147-148.

Saunders, C. P. R., 1993: A review of thunderstorm electrification processes. *J. Appl. Meteor.*, **32**, 642-653.

Uman, M. A., 1987: *The Lightning Discharge*. International Geophysical Series, No. 39, W. L. Donn, Ed., Academic Press, 377 pp.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. International Geophysical Series, No. 59, R. Dmowska and J.R. Holton, Ed., Academic Press, 467 pp.

APPENDIX A

IDL PROGRAM TO CALCULATE THE NUMBER OF FLASHES PER DAY

```
pro flash_eval

        ; This program uses NLDN data to calculate the number of flashes that occur
        ; over a given geographic region for each day of each year specified.
        ; Lastly, it writes the number of flashes and date to a file.

        ; Written By: Micheal Holmes
        ; Last updated: 10 June 1999

years = ['95', '96', '97', '98']
months = ['jan', 'feb', 'mar', 'apr', 'may', 'jun', 'jul', 'aug', 'sep', 'oct', 'nov', 'dec']
nyears = n_elements(years)
nmonths = n_elements(months)
inpath = '/home/fujita12/flash/lgh19'
files = strarr(nyears*nmonths)
for y = 0, nyears - 1 do                $
        for m = 0, nmonths - 1 do       $
                files[y*nmonths + m]=inpath + years[y] $
                + '/' + months[m] + years[y] + '.lgh'
        Range = ['01/01/95', '12/31/98']
d1 = julday(1,1,1995)
d2 = julday(12,31,1998)
ndays = d2-d1+1
counts = lonarr(ndays)
num = strcompess(sindgen(100),/remove_all)
num[0:9] = '0'+num[0:9]
        for day = 0, ndays-1 do begin
                caldat, d1+day, m, d, y
                caldat, d1+day+1, m_2, d_2,y_2
                dates=[num[m]+'/'+num[d]+'/'+num[y mod 100] +' '+'00:00:00',
num[m_2]+'/'+num[d_2]+'/'+num[y_2 mod 100] +' '+'00:00:00']
        findtime, dates, startind, startpos, lastind, lastpos, files, 11L
        currentind = startind
        currentpos = startpos
        done = ((currentind GE lastind) AND (currentpos GE lastpos))
        region = [26.12, 30.11, -82.92, -78.4]
```

```
        nflashes=0
        while not(done) do begin
        f=getchunk(files,startind,stoppos,lastind,lastpos,region,currentind,currentpos,11L,
        50000)


                if (n_elements(f) GT 1) then begin
                nflashes = nflashes + n_elements(f)/11
                f = exp_lgh(f)
                endif
        done = ((currentind GE lastind) AND (currentpos GE lastpos))
        endwhile
        start_date=[num[m]+'/'+num[d]+'/'+num[y mod 100] +' '+'00:00:00']
        print, nflashes, start_date
        openu, outfile, 'filename.txt', /get_lun, /append
        printf, outfile, nflashes, start_date
        close, outfile
        free_lun, outfile
endfor
end
```

APPENDIX B

IDL PROGRAM THAT GENERATES GIF IMAGES OF LIGHTNING FLAHSES

```
pro master_gif_pass
        ; This program will plot the flashes in ten minute intervals
        ; over a specified geographic region then write the image as
        ; a gif file to a specified directory.

        ; Written By: Michael Holmes
        ; Last updated: 13 July 1999

;  Declaration of variables:
yy = ' '
mmm = ' '
dd =' '
mmddyy1 =' '
mmddyy2 =' '
hh1 =' '
mm1 =' '
ss1 =' '
hh2 =' '
mm2 =' '
ss2 =' '
stop_hr =' '
stop_min =' '

openr, 8, '96_cont.txt'
while not(eof(8)) do begin
readf, 8, mm, dd, yy, hh1, mm1, hh2, mm2, mmm, $
        format = '(a2, 1x, a2, 1x, a2, 1x, a2, 1x, a2, 4x, a2, 1x, a2, 4x, a3)'

; These commands insure that hh and mm are always two digets
; i.e., hh=07, mm=10, ....
num = strcompress(sindgen(100), /remove_all)
num(0:9) = '0'+num(0:9)
months = mmm
years = yy
days = dd

; This portion fixes the number of digits to certian variables.
dd = fix(dd)
```

```
mm = fix(mm)
h1 = fix(hh1)
m1 = fix(mm1)
stop_hr = fix(hh2)
stop_min = fix(mm2)
h2 = fix(hh1)
min_int= 10
print, mm, dd, yy, hh1, mm1, hh2, mm2, mmm
m2 = m1+min_int
        if (m1 GE 50) then begin
                h2 = h2+1
                m2 = 0
        endif
count = mm1/10
print, 'count =',count
done = ((h1 EQ stop_hr) and (m1 EQ stop_min))
while not(done)  do begin
dates=[num(mm) +'/'+ num(dd) +'/'+ yy + ' ' + num(h1) +':'+ num(m1) +':'+ '00',
num(mm) +'/'+ num(dd) +'/'+ yy + ' ' + num(h2) +':'+ num(m2) +':'+ '00']
        inpath = '/home/fujita12/flash/lgh19'
        files = inpath + years + '/' + months + years + '.lgh'
        print, dates
        findtime, dates, startind, startpos, lastind, lastpos, files, 11L
        currentind = startind
        currentpos = startpos
        done = ((currentind GE lastind) AND (currentpos GE lastpos))
        region = [25.0, 50.0, -125.0, -67.0]
        nflashes=0
        f=getchunk(files,startind,stoppos,lastind,lastpos,region,currentind,currentpos,11L,
        50000)
        print, n_elements(f)
        nflashes = n_elements(f)/11
        print, nflashes
        if (nflashes EQ 0) then goto, fwd
                dfile = mmm + num(dd) + num(yy) + '_' + num(h1) + num(m1) + '_' +
                num(h2) + num(m2)
                f = exp_lgh(f)
                help, f, /structure
                print, n_elements(f)
                map_set, 0, -100, 0, limit = [26, -83, 30, -78], /usa, /lambert, color=255,
                /noborder, $
                title = dfile
                plots, f.lon, f.lat, psym = 3, color = 255
                write_gif, '/home/fujita8/gif/96_plots/' + dfile+'.gif', TVRD()
        fwd: m1 = m1 + min_int
```

```
        m2 = m1 + min_int
        if (m1 EQ 60) then begin
                m2 = m1 - 50
        endif

        if ((m1 GE 50) AND (m1 LT 60)) then begin
                h2 = h2 +1
                m2 = 0
        endif
        count = count + 1
        if (count EQ 6) then begin
                h1=h1+1
                count = 0
                m1 = 0
        endif
      print, h1, stop_hr
      print, m1, stop_min
        done = ((h1 EQ stop_hr) and (m1 EQ stop_min))
 endwhile
 wdelete
 wait, 1
endwhile
close, 8
end
```

APPENDIX C

IDL PROGRAM THAT GENERATES A TIME SERIES PLOT OF A THUNDERSTORM

```
pro master_xy_gif_plot

        ; This program was written to generate an x, y time series plot
        ; of peak absolute current, peak cumulative flash rate, peak negative
        ; flash rate, and peak positive flash rate. It also generates a plot
        ; of lightning flashes and writes them as gif files to a specified directory.
        ; You will be prompted to select the 25 nmi range circle, then a rectangular
        ; region of interest to plot NLDN data over.

        ; Written By: Micheal Holmes
        ; Last updated: 23 Oct 1999

;  Declaration of variables:

yy = ' '
mmm = ' '
dd =' '
mmddyy1 =' '
mmddyy2 =' '
hh1 =' '
mm1 =' '
ss1 =' '
hh2 =' '
mm2 =' '
ss2 =' '
stop_hr =' '
stop_min =' '

; These commands insure that hh and mm are always two digets
; i.e., hh=07, mm=10, ....
num = strcompress(sindgen(100), /remove_all)
num(0:9) = '0'+num(0:9)

; This portion of the program prompts the user for the dates
; and times for which to generate gif images.
print, 'please enter the year (yy)'
        read, PROMPT='yy:', yy, format='(a2)'
print, 'please enter the month (mmm)'
```

```
        read, PROMPT='mmm:', mmm, format='(a3)'

print, 'please enter the day (dd)'
        read, PROMPT='dd:', dd, format='(a2)'
years = yy
months = mmm
days = dd
print, 'Please enter initial  date (mm/dd/yy)'
        read, Prompt='mmddyy:', mmddyy1, format='(a9)'
print, 'Please enter initial hour (hh)'
        read, Prompt='hh:',hh1, format='(a2)'
print, 'Please enter initial min (mm)'
        read, Prompt='mm:',mm1, format='(a2)'
print, 'Please enter final  date (mm/dd/yy)'
        read, Prompt='mmddyy:', mmddyy2, format='(a9)'
print, 'Please enter final hour (hh)'
        read, Prompt='hh:',hh2, format='(a2)'
print, 'Please enter final min (mm)'
        read, Prompt='mm:',mm2, format='(a2)'

test = increment(hh1, hh2, mm1, mm2, nx)          ; This portion calculates the number
of 5 min increments.
print , nx
h1 = fix(hh1)                                     ; This portion fixes the number of
digits to certian variables.
m1 = fix(mm1)
stop_hr = fix(hh2)
stop_min = fix(mm2)
h2 = fix(hh1)
Date_legend = mmm + num(dd) + num(yy) + '_' + num(h1) + num(m1) + '_' +
num(stop_hr) + num(stop_min)
read_gif, 'kxmr3.gif', img, r,g,b                 ;  This portion picks the region in
which to grab data.
tvlct, r,g,b,0
tvlct, 0, 0, 0,    19      ;black
tvlct, 180,180,255, 15
tvlct, 255,200,255, 16  ;light blue
tvlct, 210,210,210, 17
tvlct, 255,255,255, 18  ;white
a =  size(img)
window, xsize=a(1), ysize=a(2)
tv, img
print, 'Select 25 nm North radial'
cursor, x,y, /up, /device
        y_adj = (y-(a(2)/2))/(25*1.852)
```

```
i = 0
done = (i EQ 2)
while not(done) do begin                    ; Select rectangular region of
interest
                                            ; beginning with NW, then
SE corner.
        cursor, x,y, /up, /device
        xnew = x-(a(1)/2)
        ynew = y-(a(2)/2)
        Lon_lat0 = [-1.4078, .4907]
        Az = atan(xnew,ynew)
        Arc_dist = ((ynew/y_adj)/(cos(Az))/6371)
        coordinates = LL_arc_distance(Lon_lat0, Arc_dist, Az)
        deg_coord = coordinates *(180/!pi)
        if (i EQ 0) then begin
                N_lat = coordinates(1)*180/!pi
                W_lon = coordinates(0)*180/!pi
                x_min = x
                y_max = y
                print,'NLAT,Wlon=', N_lat, W_lon, y_max, x_min
        endif

        if (i EQ 1) then begin
                S_lat = coordinates(1)*180/!pi
                E_lon = coordinates(0)*180/!pi
                x_max = x
                y_min = y
                print, 'SLAT/Elon=', S_lat, E_lon, y_min, x_max
        endif
        i= i+1
        done = (i EQ 2)
endwhile

min_int= 5
m2 = m1+min_int
        if (m1 GE 55) then begin        ; Start rime loop.
                h2 = h2+1
                m2 = 0
        endif
count = mm1/5
flash_total = 0
maxflash=0
```

```
maxpcount=0
max_peak_tot=0
maxncount=0

max_time = fltarr(nx)
max_peak = fltarr(nx)
pk_flash_rate = fltarr(nx)
pk_pos_flash_rate = fltarr(nx)
pk_neg_flash_rate = fltarr(nx)
max_time_flash = fltarr(nx)
max_time_flash_tot = fltarr(nx)
max_time_flash_pos = fltarr(nx)
max_time_flash_neg = fltarr(nx)
last_flash = fltarr(nx)
max_time_peak = fltarr(nx)
i=0
j=0
done = ((h1 EQ stop_hr) and (m1 EQ stop_min))
while not(done)  do begin
        dates=[mmddyy1 + ' ' + num(h1) +':'+ num(m1) +':'+ '00', mmddyy2 + ' ' +
num(h2) +':'+ num(m2) +':'+ '00']
        dfile =  mmm + num(dd) + num(yy) + '_' + 'storm2' + '_' + num(h1) + num(m1) +
'_' + num(h2) + num(m2)
;       dfile =  mmm + num(dd) + num(yy) + '_' + num(h1) + num(m1) + '_' + num(h2) +
num(m2)
        print, dates
        inpath = '/home/fujita12/flash/lgh19'
        files = inpath + years + '/' + months + years + '.lgh'
        findtime, dates, startind, startpos, lastind, lastpos, files, 11L
        currentind = startind
        currentpos = startpos
        done = ((currentind GE lastind) AND (currentpos GE lastpos))
        region = [S_lat, N_lat, W_lon, E_lon]
        nflashes=0
        f=getchunk(files,startind,stoppos,lastind,lastpos,region,currentind,currentpos,11L,
        50000)
        nflashes = n_elements(f)/11
        flash_total = flash_total + nflashes
        pk_flash_rate(j)=nflashes
        if (nflashes EQ 0) then goto, fwd
                f = exp_lgh(f)
                map_set, 0,-100 ,0, limit=[S_lat, W_lon, N_lat, E_lon], $
                 /usa, /lambert, color=19, /noborder
                map_continents, hires=1, /coasts, color=255
                map_continents, /rivers, hires=4, color=255
```

```
pos = where(f.peak GT 0.0, pcount)
neg = where(f.peak LT 0.0, ncount)
max_pk = max(abs(f.peak))
max_hr = f(nflashes-1).hour
max_minute = f(nflashes-1).minute
max_second = f(nflashes-1).second
time = (float(f.hour)*60) + float(f.minute) + (float(f.second)/60)
max_time(i) = time(where(abs(f.peak) EQ max_pk))
max_peak(i) = max_pk
last_flash(i) = max_hr*60.0 + max_minute*1.0 + max_second/60.0

if (max_pk GE max_peak_tot) then begin
        max_peak_tot = max_pk
        max_time_peak(i) = max_time(i)
endif
i=i+1
if (nflashes GT maxflash) then begin
        maxflash = nflashes
        max_time_flash_tot(j) = time(where(nflashes EQ maxflash))
endif

if (pcount GT maxpcount) then begin
        maxpcount = pcount
        max_time_flash_pos(j)=time(where(pcount EQ maxpcount))
endif

if (ncount GT maxncount) then begin
        maxncount = ncount
        max_time_flash_neg(j)=time(where(ncount EQ maxncount))
endif

max_time_flash(j) = time(where(nflashes EQ nflashes))
pk_pos_flash_rate(j)=pcount
pk_neg_flash_rate(j)=ncount
print, nflashes, pk_flash_rate(j), pk_pos_flash_rate(j),
pk_neg_flash_rate(j)
print, max_time_flash_tot(j), max_time_flash_pos(j),
max_time_flash_neg(j)
j=j+1
if (ncount EQ 0) then begin
        ratio = 'INF'
endif
if (ncount NE 0) then begin
        ratio = pcount*1.0/ncount*1.0
endif
```

```
        print, ' Number of total flashes = ',flash_total
        print, ' Number of posative flashes = ',pcount
        print, ' Number of negative flashes = ',ncount
        print, ' Peak Current = ',max(abs(f.peak))
        pk_curr =(max_peak)*1.0
        if (pcount EQ 0) then begin
                print, ' Max Pos Curr = no pos flashes'
                Max_pos_curr = 'no pos flashes'
        Endif

        if (pcount NE 0) then begin
                print, ' Max Pos Curr = ',max(f(pos).peak)
                Max_pos_curr = max(f(pos).peak)
        endif

        print, ' Last Flash = ', max_hr, max_minute, max_second
        print, ' Ratio = ', ratio
        plots, f.lon, f.lat, psym = 3, color = 255
        plots, f(nflashes-1).lon, f(nflashes-1).lat, psym = 2, color = 255,
        symsize=1

        XYOUTS,0.5, .97, 'Date'+string(dfile)+'     ', CHARSIZE=1,/Normal, $
        alignment=1, color=255
        XYOUTS,0.5, .97, 'Flashes'+string(nflashes), CHARSIZE=1,/Normal, $
        alignment=0, color=255
        XYOUTS,0.5, .94, 'Storm Tot Flashes'+string(flash_total), $
        CHARSIZE=1,/Normal, alignment=0, color=255
        XYOUTS,0.5, .91, 'Peak Curr'+string(max_peak_tot), $
        CHARSIZE=1,/Normal, alignment=0, color=255
        XYOUTS,0.5, .88, ' Max Pos Curr = '+string(Max_pos_curr), $
        CHARSIZE=1,/Normal, alignment=0, color=255 $
        XYOUTS,0.5, .85, 'Last
        Flash'+string(max_hr)+string(max_minute)+string(max_second), $
        CHARSIZE=1,/Normal, alignment=0, color=255
        XYOUTS,0.5, .82, ' Ratio = '+ string(ratio), CHARSIZE=1,/Normal, $
        alignment=0, color=255
        write_gif, '/home/fujita8/gif/cases/last_flash/' + dfile+'.gif', TVRD()

fwd:    m1 = m1 + min_int
        m2 = m1 + min_int
        if (m1 EQ 60) then begin
                m2 = m1 - 55
        endif

        if ((m1 GE 55) AND (m1 LT 60)) then begin
```

```
            h2 = h2 +1
            m2 = 0
        endif
        count = count + 1



        if (count EQ 12) then begin
            h1=h1+1
            count = 0
            m1 = 0
        endif
        done = ((h1 EQ stop_hr) and (m1 EQ stop_min))
openu, outfile1, '/home/kramer1/users/mholmes/thesis/text_output/'+Date_legend+'.txt', $
/get_lun, /append

printf, outfile1, dfile, max(last_flash), flash_total, max(max_peak),  $
max(max_time_peak), $
        max(pk_flash_rate), max(max_time_flash_tot), max(pk_pos_flash_rate), $
        max(max_time_flash_pos), max(pk_neg_flash_rate), max(max_time_flash_neg)

close, outfile1
free_lun, outfile1
endwhile
wdelete
wait, 2
plot, max_time,  max_peak, /ynozero,  color=255, $
        xrange=[last_flash(0)-5, max(last_flash)+5], $
        subtitle='Date      '+string(Date_legend), background=19, $
        ytitle = 'Peak Currents V/m', min_value=1, $
        xstyle=8, ystyle=8,  $
        xmargin=[8,8], ymargin=[4,4]
        plots, max(max_time_peak), max(max_peak), color=255, psym=2, symsize=1
        xyouts, max(max_time_peak), max(max_peak), max(max_time_peak), $
            color=255, alignment=.5, charsize=1.5
        axis, yaxis=1, yrange=[0, max(pk_flash_rate)+10], $
        ystyle=1, ytitle='Peak Flash Rate per 5min', color=255, /save
oplot, max_time_flash, pk_flash_rate, linestyle=2, color=255, min_value=1
oplot, max_time_flash, pk_pos_flash_rate, linestyle=3, color=255, min_value=1
oplot, max_time_flash, pk_neg_flash_rate, linestyle=1, color=255, min_value=1
xyouts, max(max_time_flash_tot), max(pk_flash_rate), max(max_time_flash_tot), $
        color=255, alignment=.5, charsize=1.5
plots, max(max_time_flash_tot), max(pk_flash_rate), psym=2, color=255, symsize=1
xyouts, max(max_time_flash_pos), max(pk_pos_flash_rate), max(max_time_flash_pos),
$ color=255, alignment=.5, charsize=1.5
```

```
plots, max(max_time_flash_pos), max(pk_pos_flash_rate), psym=2, color=255, $
symsize=1
xyouts, max(max_time_flash_neg), max(pk_neg_flash_rate), max(max_time_flash_neg),
$ color=255, alignment=.5, charsize=1.5
plots, max(max_time_flash_neg), max(pk_neg_flash_rate), psym=2, color=255, $
symsize=1
xyouts, max(last_flash),  0, max(last_flash), color=255, alignment=.5, charsize=1.5 $
plots, max(last_flash), 0, psym=2, color=255, symsize=1

        write_gif, '/home/fujita8/gif/cases/xy_plots/' + Date_legend +'_xy'+'.gif', $
TVRD()
wait, 2
openu, outfile2, '/home/kramer1/users/mholmes/thesis/text_output/master_flash.txt', $
/get_lun, /append

print,  'Date:              ' +' '+ 'Last Flash:' +' '+ 'Tot Flashes:'  +' '+'Pk Cur:' +' '+ 'T Pk  $
Cur:'+' '+ 'Pk Flash Rt:'  +' '+'T Pk Flash Rt:' +' '+  'Pk + Flash Rt:' +' '+  'T Pk + Flash $
Rt:'+' '+  'Pk - Flash Rt:' +' '+ 'T Pk - Flash Rt:'

print,  Date_legend, max(last_flash), flash_total, max(max_peak), max(max_time_peak),
$ max(pk_flash_rate), max(max_time_flash_tot/60), max(pk_pos_flash_rate), $
max(max_time_flash_pos), max(pk_neg_flash_rate), max(max_time_flash_neg)


printf, outfile2, Date_legend, max(last_flash), flash_total, max(max_peak), $
max(max_time_peak), max(pk_flash_rate), max(max_time_flash_tot), $
max(pk_pos_flash_rate), max(max_time_flash_pos), max(pk_neg_flash_rate), $
max(max_time_flash_neg)

close, outfile2
free_lun, outfile2
end
```

Vita

Michael W. Holmes was born in December 1967 in Salt Lake City, Utah. He graduated from Ogden High School in 1986 and enlisted in the Air Force in the same year, where he served as an Aircraft Maintenance Specialist (Crew Chief) on F-15's. During his enlistment, he was stationed at the 318 FIS at McChord AFB, Washington and the 555 TFTS at Luke AFB, Arizona until he separated from the Air Force in 1992 and immediately enrolled in the Meteorology program at the University of Utah.

In 1995 he accepted a student internship at the National Weather Service (NWS), Western Region Scientific Services Division (WRSSD) where he performed a comparison of WSR-88D precipitation estimates with gauge data. He later published his results in a technical attachment (TA 95-30) and presented them at the Second Annual Workshop on Weather Prediction in the Intermountain West. In this same year, he was also awarded the Shia-Kung Kao Scholarship and an AFROTC Scholarship. In 1996 he graduated with his B.S. in meteorology and was shortly thereafter commissioned in the Air Force as a Weather Officer where he was stationed at the Tanker Airlift Control Center (TACC) Scott AFB, Illinois. In 1998 he was assigned to the Air Force Institute of Technology (AFIT) at Wright-Patterson AFB, OH. He will graduate with his M.S. in meteorology from AFIT in March 2000 and move to his next assignment at the 11th Operational Weather Squadron (OWS) Elmendorf, AFB AK.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2000	Master's Thesis

4. TITLE AND SUBTITLE
TECHNIQUES FOR FORECASTING THE CESSATION OF LIGHTNING AT CAPE CANAVERAL AIR STATION AND THE KENNEDY SPACE CENTER

5. FUNDING NUMBERS

6. AUTHOR(S)
Michael W. Holmes, First Leutenant, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 P Street, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GM/ENP/00M-08

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
45th Weather Squadron
Attn: Mr. William P. Roeder
1201 Edward H. White II St., MS 7302
Patrick AFB, FL 32925-3238 DSN 467-8410

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Chairman: Major Gary R. Huffines, ENP, DSN: 785-3636, ext. 4511
Member: Lt Col Cecilia A. Miner, ENP, DSN: 785-3636, ext. 4645
Member: Assoc. Prof. William F. Bailey, ENP, DSN: 785-3636, ext. 4501

12a. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The focus of this research effort is directed toward identifying new methods of forecasting the cessation of lightning along the Central Atlantic Coast of Florida. Cloud-to-ground lightning flashes place Air Force (AF) personnel and assets at risk almost daily at this location. Providing a more accurate method of forecasting the cessation of lightning would allow for safer and more efficient execution of AF operations. A data set consisting of 40 thunderstorm cases was identified within a 90 nautical miles (nmi) region surrounding the Melbourne, Florida WSR-88D (KMLB) site. Each case falls between the months of May and September and the years of 1995 through 1997. Simple and multiple linear regression models are built using this dataset. Variables included max Vertically Integrated Liquid water (VIL), max reflectivity, max peak current, peak cumulative flash rate, peak negative flash rate, and peak positive flash rate. Results indicate that three of the simple linear regression models to some extent accurately represent the data. Additionally, when the data set is separated by thunderstorm cell type (multi or single) and cell specific regressions are built, results indicate that the regressions based on the single-cell data set produce a substantial increase in forecast skill compared to that of climatology. In fact, some regressions are shown to improve forecast accuracy by 90 % over that of climatology. Moreover, multiple linear regression models are shown to produce similar results and further reinforce the notion that each thunderstorm cell type (multi or single) behaves substantially different from the other with respect to forecasting the cessation of lightning.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Lightning, Cessation of Lightning.	87
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL