LIGHTNING PREDICTION USING ARTIFICIAL NEURAL NETWORKS AND ELECTRIC FIELD MILL DATA

THESIS

Daniel E. Hill, Capt, USAF

AFIT-ENC-MS-18-M-002

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENC-MS-18-M-002

LIGHTNING PREDICTION USING ARTIFICIAL NEURAL NETWORKS AND ELECTRIC FIELD MILL DATA

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

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 Masters of Science in Applied Mathematics

Daniel E. Hill, B.S. Mathematics, B.A. Spanish

Capt, USAF

22 March 2018

AFIT-ENC-MS-18-M-002

LIGHTNING PREDICTION USING ARTIFICIAL NEURAL NETWORKS AND ELECTRIC FIELD MILL DATA

THESIS

Daniel E. Hill, B.S. Mathematics, B.A. Spanish
Capt, USAF

Committee Membership:

Lt Col R. Seymour, PhD
Chair

Dr. B. Borghetti
Member

Lt Col A. Geyer, PhD
Member

AFIT-ENC-MS-18-M-002

Abstract

Electric Field Mills (EFMs) located in the region surrounding Cape Canaveral record the electrification of the atmosphere near them. Research studying how these sensors could improve lightning warnings has had mixed results. This paper used a Convolutional Recurrent Neural Network (CRNN) and data from 30 EFMs from May-July of 2012-2016. The mean was calculated for every 60 second period and 30 minutes of this summarized data was used to create a lightning prediction with a warning period of 15 minutes. This method achieved a True Positive Rate (TPR) of 77.6%, a False Positive Rate (FPR) of 8.3%, a False Discovery Rate (FDR) of 48.1%, and an Operational Utility Index (OUI) of 53.9% (Kehrer et al., 2006). This suggests that the EFM sensor array, when used as a means to measure the electrification of the entire region, is capable of effectively predicting lightning for a 5-mile radius near Cape Canaveral. Moreover, achieving a 53.9% on the OUI rivals the best methods currently used implying that incorporating EFMs into lightning forecasting may reduced the FPR and save millions of dollars in delay and cancellation costs.

Table of Contents

LIGHTNING PREDICTION USING ARTIFICIAL NEURAL NETWORKS AND ELECTRIC FIELD MILL DATA

I. Introduction

1.1 Background

Research Motivation.

Cape Canaveral Air Force Station (CCAFS), Kennedy Space Center (KSC), and Patrick Air Force Base (AFB) experience the highest rate of thunderstorms in the world during the summer months due to their location in what is known as the thunderstorm capital of the United States (Florida Climate Center, 200). Lightning poses a great risk to the organizations' $17+ billion facilities, multi-billion dollar boosters/payloads, and 25,000+ personnel (Roeder, Hajek, Flinn, Maul, & Fitzpatrick, 2000). The 45th Weather Squadron (45 WS) is tasked with issuing weather forecasts and lightning warnings to ensure the risk to assets and personnel due to weather is minimized.

Lightning strikes caused $825.7 million in U.S. home insurance claims and 38 lightning related deaths, 9 of those deaths in Florida, in the United States in 2016 (Insurance Information Institute, 2016). A cloud-to-ground lightning event that strikes or nearly strikes a mission area or payload may result in injury or death of personnel and/or destruction of expensive equipment and facilities. A strike on a fuel-filled rocket may cause an explosion or lead to key guidance or navigational systems being destroyed. In 1963, Pan Am Flight 214 caught fire and crashed after being struck by lightning killing all 81 people passengers and crew (Erdman, 2014).

In 1987, lightning destroyed an Atlas-Centaur rocket during launch (Christian et al., 1989). Lightning warnings may have prevented such tragedies.

The Lightning Advisory Panel (LAP) was established in response to the Atlas-Centaur disaster and was charged with creating what is now known as the Lightning Launch Commit Criteria (LLCC) (Merceret et al., 2010). The LLCC is a set of parameters that must be met in order to conduct space flight launches. These criteria consist primarily of cloud thickness, reflectivity, temperatures, and electrification of the atmosphere as predictors of lightning events (Koons & Walterscheid, 1996). The two primary criteria that must be met for EFMs are:

1. Do not launch for 15 minutes after the absolute value of any electric field measurement at the surface within 5 NM of the flight path has been greater than 1500 V/m.

2. Do not launch for 15 minutes after the absolute value of any electric field measurement at the surface within 5 NM of the flight path has been greater than 1000 V/m (Koons & Walterscheid, 1996).

While it is generally accepted that voltage/meter of the atmosphere measured by EFMs is strongly related to the presence of lightning nearby the sensor, the relationship is not well understood. Several studies on EFMs and lightning prediction show inconclusive results on how to use EFM measurements to predict lightning, such as in Murphy, Holle, and Demetriades (2008). This thesis will attempt to more accurately predict lightning using Artificial Neural Networks (ANNs) and the measurements from the 31 EFMs surrounding the 45 WS.

1.2 Problem Statement

Research Questions.

Performing space flight operations in Florida will inevitably result in cancellations and delays due to weather, but better lightning prediction could lead to immense savings and increased safety. This research seeks to answer two questions:

1. Does a relationship exist between the electrification of the atmosphere measured by EFMs and lightning events near Cape Canaveral?

2. If this relationship exists, can it be used to predict lightning at or exceeding the success rate of past methods?

Hypothesis.

Past attempts at using EFMs to predict lightning, such as Murphy et al. (2008), focused on a small number of EFMs using the idea of a voltage/meter threshold as the indicator for a lightning event near the EFM that was measured. This thesis takes a more holistic view and looks at the voltage/meter of the region by considering all 31 EFMs simultaneously to predict a lightning event at a specific location. Additionally, instead of "smoothing" out the data by taking the mean over an interval, as Murphy et al. (2008) did, which can hide large spikes in the voltage/meter, this thesis calculates the mean, variance, kurtosis, and skew to capture the subtle changes across time more completely. It is hypothesized that by looking at the regional electrification by measuring all EFMs over a large period of time, the performance metrics evaluating lightning prediction will be improved.

1.3 Methodology

Datasets.

The EFM sensor array consists of 31 EFMs spread across central Florida. Each EFM collects the voltage per meter (V/m) of the surrounding atmosphere at 50 hertz creating data consisting of three features: date, time and voltage per meter. This information is available from 1996 to the present. This thesis combines data May-July from 2013-2016 with the LDAR data to create its datasets.

The LDAR array consists of eight sensors spread across central Florida that detect when a lightning event occurs. The system triangulates the lightning event's location using the readings from the eight sensors. The data consists of six features: Time, Date, X location, Y location, Z location, and Event type. Data is available from 1994 to the present. The data used in this analysis was received from the 45 WS and was preprocessed by them to only include the type of lightning events pertinent to this analysis.

Approach.

This paper differs from past methods by considering a much larger dataset than previous research (May-July, 2013-2016) coupled with the Lightning Detection and Ranging (LDAR) dataset from 2013-2016. For each EFM, summary statistics (mean, variance, skewness, and kurtosis) were calculated for every 20, 30, and 60 second period, known as the statistic window (SW). Two measurement windows (MWs) of 30 and 60 minutes were used to create running windows of data. Responses were created to identify when lightning occurred within a 15-minute prediction window (PW) after a 15-minute warning window (WW). Lastly, the LDAR dataset was reduced to only include lightning events that occurred within a 8.04 km (5 mile) and 16.09 km (10 mile) radius around Cape Canaveral, known as the area of concern (AOC).

A dataset was created for each of the possible SW, MW, PW, WW, and AOC combinations, resulting in 12 datasets. Considering the EFM data as hyperspectral images of the region across time, much like was done in the work by Tsironi, Barros, Weber, and Wermter (2017); and Wu and Prasad (2017), a convolutional recurrent neural network (CRNN) was chosen to capture both the spatial and temporal nature of the data. All 12 datasets were trained on the CRNN and the results were compared against past research and the current prediction methods used by the 45 WS.

1.4 Overview

Chapter 2 provides context to this paper by exploring some of the past research on lightning prediction and some examples of how ANNs have improved machine learning processes. Chapter 3 is the methodology and details this paper's research, to include the description of the neural networks, the selection of the parameters to create the datasets, and the metrics used to evaluate performance. Chapter 4 is the discussion of the research and the results. Chapter 5 presents the conclusions and recommendations for future research.

II. Literature Review

2.1 Introduction

This chapter defines a number of terms to allow for discussion and comparison of past research. Past research has been inconsistent with its definitions of statistics and performance metrics requiring a discussion on how this paper will evaluate success (Barnes, Schultz, Gruntfest, Hayden, & Benight, 2009). This chapter then discusses past research into lightning prediction and compares how each method performed using the terminology introduced. Lastly, artificial neural networks and their use in meteorology and time sequenced prediction is discussed.

2.2 Introducing Terminology

Model Parameter Terminology.

Reading past research into lightning prediction reveals that all research has several parameters in common. Each method of research requires the researcher to select the area of concern (AOC), the measurement windows (MW), and the warning window (WW). Some methods also use a prediction window (PW) and a statistics window (SW). While not all research explicitly states these values, the nature of the research and the context of the study requires these parameters. Mazany, Businger, Gutman, and Roeder (2002); Murphy et al. (2008); and Da Silva Ferro et al. (2011) had similar terminology, but this thesis seeks to more explicitly define these and describe past research in these common terms. For research which assumes but does not state these values, this thesis attempts to explicitly identify them.

To begin, AOC is defined as the area in which the researcher is attempting to predict lightning activity. The shape of this area may be circular, rectangular, or

elliptical. The total area covered by the AOC may range greatly depending on the research and the design decisions.

Da Silva Ferro et al. (2011) provides a graphic that depicts their chosen AOC, shown in Figure 1. Figure 1 shows they selected an AOC with a circular shape centered around an electric field mill (EFM). The graphic does not depict the radius of the AOC, but Da Silva Ferro et al. (2011) selected radii of 5, 10, and 15 kilometers (km) and performed an analysis with each radius.

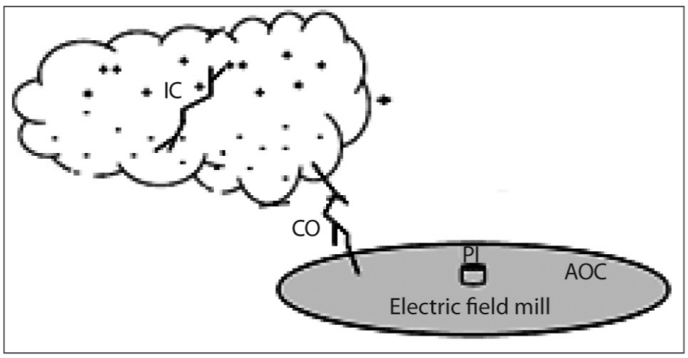


Figure 1. Graphical Representation of AOC

Once a researcher determines their AOC, the parameters for their data must be selected. Figure 2 gives a graphic representing the different parameters described here. First, the SW is a time period over which EFM readings may be summarized. In most research, this is a small amount of time between 20 seconds and 1 minute. The MW is the amount of time used by a model to predict a lightning event. For example, if a prediction is made based off a single EFM reading, then the MW is 0.02 seconds, since EFMs measure the V/m once every 0.02 seconds. If 30 minutes of EFM data is used to make a prediction, then the MW is 30 minutes.

The WW, referred to as warning time or lead time in some literature, is the amount of time between the PW and the MW. WW is so named because in a real-world lightning event, this is the amount of warning that one can expect to have before a lightning event occurs. It is important to note that some lightning forecasting studies

do not establish this prior to analysis. Instead, a criteria is established that triggers a lightning warning and the amount of time that passes between the instantiation of the warning and a lightning event is recorded. At the end of the study the mean of these times is calculated and established as the lead time. If it is assumed that the recorded lead times approach a normal distribution, then 50% of the warnings fail to meet the after-established lead time resulting in a constant 50% false negative rate. This is concerning and should be addressed in future studies.

The PW is the time window in which lightning is predicted (or not predicted) to occur. Given any given MW, a model will either predict lightning or not predict lightning and the PW is the time for which that prediction is made.

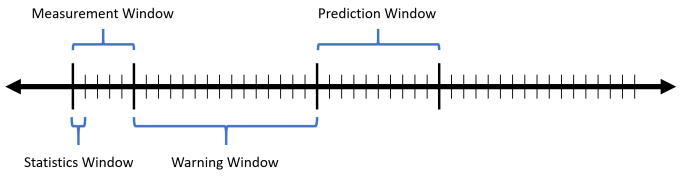


Figure 2. Graphical Representation of Parameters

In Figure 2, if each tick mark is one minute, then the SW is 1 minute, the MW is 5 minutes, the WW is 15 minutes and the PW is 10 minutes. This means that 5 minutes of data containing 5 one minute SW summarizations of the data are being used to make predictions for the time period 15 minutes (WW) after the end of the MW and 10 minutes (PW) after the end of the WW. Using a running window of MWs ensures that a prediction is made for all time periods within the timeline. Since the WW of one MW is part of the prediction window of a previous MW, a determination of the model's efficacy is not dependent on what happens in the WW. Only the events occurring within the PW are important. Figure 3 depicts two MWs in time. The lightning event occurs at the same time, but it happens in the first MW's PW and the second MW's WW. Since the first MW predicts lightning and the second does

not, then both are correct since each MW is only concerned with its own PW. The running window ensures the entire timeline has predictions.

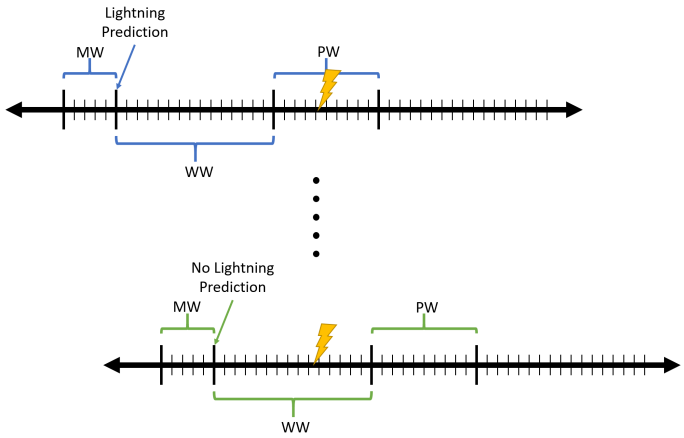


Figure 3. Running Window Covering all Time Periods

Once these parameters are selected, common terminology and definitions of the metrics used to evaluate a model. This section defines terminology and describes past research methods in these terms in order to make an objective comparison of their results. Chapters 4 and 5 will also use these terms to evaluate this thesis' methods and results.

Results Terminology.

To compare past methods of lightning prediction, the notion of accuracy must be more clearly defined. Using the above diagram depicting SW, MW, WW, and PW, this section explains the different situations that may arise and gives terms to each. Table 1 depicts a commonly used confusion matrix. Using this structure only allows for prediction within the PW, regardless of what lightning activity may be occurring elsewhere. Since a running window is used to predict lightning continuously, a correctly functioning predictor would warn of lightning prior to any event (as depicted

in Figure 3 above).

Table 1. Confusion Matrix Example

	Actual	
	Lightning	No Lightning
Predicted Lightning	True Positive	False Positive
Predicted No Lightning	False Negative	True Negative

A True Positive (TP), or hit, is when the model predicts a lightning event from a MW of readings and a lightning event occurs within that MW's PW. A TP is depicted in Figure 4.

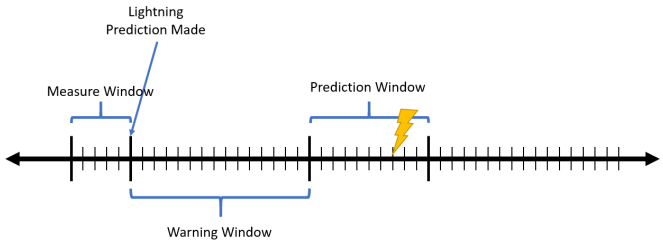


Figure 4. True Positive

A False Negative (FN) occurs when a model does not predict lightning for a PW, but a lightning event happens. This is the most dangerous real-world situation and should be minimized as much as realistically possible. Figure 5 depicts a FN.

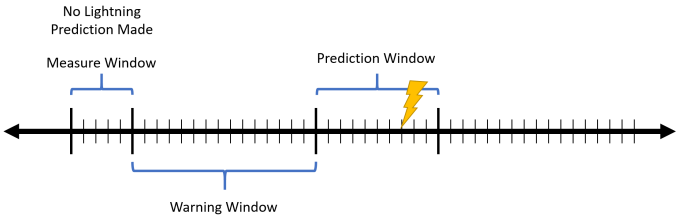


Figure 5. False Negative

The True Positive Rate (TPR), also known as probability of detection, hit rate, sensitivity, or recall; is a statistic describing the ability of a model to 'detect' lightning events. It describes how well a model does at accurately predicting true lightning

events. Equation 1 shows the calculation of this statistic. A model that has high TPR often trades this success for a high number of predictions of lightning when no lightning event occurs. While the risk of lightning would be absolutely minimized, this situation is not realistic. A model that has a high TPR and still predicts nonlightning events accurately is optimal.

$$\text{TPR} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{1}$$

A False Positive (FP), or false alarm, is when a lightning event is predicted, but one does not happen within the MW's PW. Figure 6 depicts this. This situation is realistically better than a FN because assets and personnel are not placed at risk. However, a model that predicts only lightning events results in a large number of FPs which would stop all space flight operations and not be realistic. Minimizing the FA rate as much as possible is also optimal.

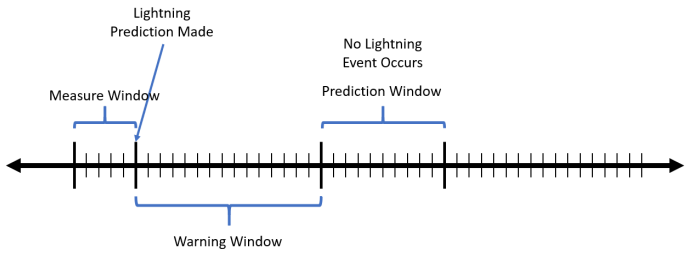


Figure 6. False Positive

A True Negative (TN) is the number of times a model did not predict lightning and lightning did not happen. The idea of 'correctly not warned' is applicable. Figure 7 depicts a TN.

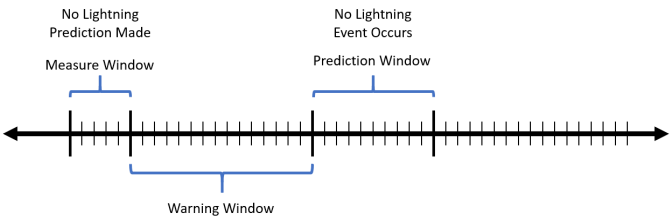


Figure 7. True Negative

Before proceeding, it is important to mention the confusion that has arisen, both within the meteorological community and outside it, regarding Equations 2 and 3. A brief discussion of the different terms used will be given and then a term will be selected and used for the remainder of this paper. Equation 2 is the ratio of FPs to all true nonlightning events. It can be thought of as the percentage of time a model will issue a false alarm and resources will be expended needlessly because a lighting event does not follow within a reasonable amount of time (for instance, the PW). Equation 3 is the ratio of the number of false positives to all events predicted as lightning. This can be thought of as the probability of a model to produce a false alarm.

Equation 2 has commonly been called the False Positive Rate or the False Alarm Rate. Barnes et al. (2009) suggest to use the term Probability of False Detection as a means of alleviating the confusion that has arisen in the meteorological community between False Alarm Rate and False Alarm Ratio. However, this creates further confusion as this term is also used by some to refer to Equation 3. In an attempt to remove ambiguity and unload these overloaded terms, this paper will use the terms False Positive Rate (FPR) to refer to Equation 2 and False Discovery Rate (FDR) to refer to Equation 3.

$$\text{FPR} = \frac{\text{False Positive}}{\text{False Positives} + \text{True Negatives}} \tag{2}$$

$$\text{FDR} = \frac{\text{False Positive}}{\text{False Positives + True Positives}} \tag{3}$$

The True Negative Rate (TNR) is the ratio of the number of true negatives to all true nonlightning events (see 4. It is the effectiveness of a model to correctly not warn of a lightning event.

$$\text{TNR} = \frac{\text{True Negatives}}{\text{True Negatives+ False Positives}} \tag{4}$$

The Hanssen and Kuipers discriminant, or True Skill Score (TSS), is a measure of how well the model discriminated between lightning and non lightning events (Hanssen & Kuipers, 1965; Woodcock, 1976). It is calculated by taking the difference between the TPR and the FPR (see Equation 5) and is a discriminant measure that has been shown to yield an unbiased measure of forecasting accuracy for yes/no forecasting (Woodcock, 1976). However, This statistic suffers heavily weighting the TPR resulting in it becoming equal to TPR when a model is attempting to predict rare events. This could be problematic for lightning events.

$$\text{TSS} = \text{TPR - FPR} = \frac{\text{(TP)(TN)-(FP)(FN)}}{\text{(TP+FN)(FP+TN)}} \tag{5}$$

In addition to these statistics, an Operational Utility Index (OUI) will be calculated. The OUI is a non-standard statistic created by the 45 WS to capture the importance of finding the optimal balance between TPR and FDR (D'Arcangelo, 2000; Kehrer, Graf, & Roeder, 2006). The OUI uses the TSS (see Equation 5) in addition to TPR and FDR and weights these three statistics to calculate a metric with a maximum value of 0.83 being the best possible model. Equation 6 shows the weights and calculation of the OUI.

$$\text{OUI} = \frac{3(\text{TPR}) + 2(\text{TSS}) - \text{FDR}}{6} \tag{6}$$

The metrics used to report the effectiveness of a lightning prediction varies widely. This paper attempts to compare methods according to these statistics to achieve as objective of a comparison as possible. It is important to keep in mind that the parameters and methods used in each research vary widely, so the notion of a 'best' method cannot be made. The comparison is done to show the success of the corpus of lightning prediction methods and how this paper's methods compare.

2.3 Research in Lightning Forecasting with Electric Field Mills

This section examines past research into lightning prediction and evaluates their methods using the terminology and statistics introduced in the last section. Terminology differs among researchers; when there is confusion, this paper attempts to clarify the intent of the authors as best as possible. A brief introduction of the past research is given followed by a comparison of the research decisions made in each. These are summarized in Tables 3 and 4. Lastly, the methods are evaluated and compared in Table 7.

The Lightning Launch Commit Criteria (LLCC) establishes criteria that must be met for space flight launches. These criteria include two explicit references to Electric Field Mills (EFMs):

1. Do not launch for 15 minutes after the absolute value of any electric field measurement at the surface within 5 NM of the flight path has been greater than 1500 V/m.

2. Do not launch for 15 minutes after the absolute value of any electric field measurement at the surface within 5 NM of the flight path has been greater than

1000 V/m (Koons & Walterscheid, 1996).

While the LLCC is used to mitigate the risk of triggered lightning at launch, the idea of a V/m threshold being exceeded triggering a warning is common. The 45 WS uses a similar threshold idea and other methods as criteria to issue Phase-1 lightning watches which indicate lightning is probable within a 5 nautical mile radius within 30 minutes (Roeder et al., 2000). The research presented here includes methods that use similar thresholds as seen in the LLCC and predict lightning for both long and short-term periods. Mixed results show a continued need to understand the contribution of EFM data to lightning prediction.

Lightning Prediction Techniques.

Mazany et al. (2002) created a lightning prediction index using Global Positioning System (GPS) Integrated Precipitable Water Vapor (IPWV), along with several other variables. Their research focused on improving the forecasting skill and increasing the lead time of lightning forecasting for the KSC. The researchers used a logistic regression model and performed statistical analysis to select four variables: the 30-minute maximum value across all 31 EFMs (read every 5 minutes), GPS IPWV, the change of GPS IPWV over a 9-hour period, and the K-Index (KI). The EFM data was measured every 5 minutes and a 30-minute maximum was used as the predictor variable (SW and MW both equal to 30 minutes). They fit a logistic regression model with these four variables from data collected between May and September of 1999. The output from this model was a probability of lightning which they called the GPS Lightning Index.

Looking at the success of different probability thresholds, they selected a threshold of 0.7; meaning that all GPS Lightning Indexes below 0.7 were predicted to be lightning and all those above were not. They set their WW at 90 minutes, but did

not specify a PW. The intent of their research was for long-term prediction ($\geq$90 minutes) for a 37.04 km (20 nautical mile) radius around Cape Canaveral (AOC). When evaluating their model, if a lightning event happened anytime 90 minutes after the GPS Lightning Index fell below 0.7, this was considered a TP. It is inferred based on the discussions in the paper that the prediction is intended for the entire day or at least for a 12 hour period after the 90 minute WW. Thus, this paper will set the PW at 13.5 hours.

Mazany et al. (2002) reported evaluation metrics by day, meaning that each day was seen as an observation. If a lightning event occurred during that day, then it was seen as a lightning event. As stated before, a TP was counted if the GPS Lightning Index fell below 0.7 and a lightning event happened more than 90 minutes after. A False Negative was when the GPS Lightning Index fell below 0.7 but lightning occurred less than 90 minutes after the prediction was made. Using these definitions, Mazany et al. (2002) presented the confusion matrix in Figure 2 on a test dataset.

Table 2. Mazany et al. Confusion Matrix of Test Dataset (2002)

	Actual	
	Lightning	No Lightning
Predicted Lightning	7	3
Predicted No Lightning	1	10

In summary, Mazany et al. (2002) chose a circular AOC with radius 20 nm, SW of 30 minutes, MW of 30 minutes, WW of 90 minutes, and a PW of 13.5 hours. Their evaluation metrics are shown in 7.

Lambert, Wheeler, and Roeder (2005) performed a similar logistic regression but used the Cape Canaveral Air Force Station (CCAFS) rawinsonde data (XMR) to predict lightning in the CCAFS area. The equation was designed to calculate the probability of lightning occurring during any given day between May and September. To attempt to optimize their technique, they created five equations, one for each

month May-September, which is the thunderstorm season for that region. To evaluate their equations, they used the Brier Skill Score (BSS) and a reliability diagram. Their methods and purpose differed significantly from this thesis', so their results were omitted.

Kehrer et al. (2006) optimized the work by Mazany et al. (2002) for forecast intervals of 2 hours and 9 hours. Their intent was to create 2-hour and 9-hour models to predict lightning within an AOC of 5 nautical miles around Cape Canaveral. Using data from May-September of 2000-2003, they determined four different variables to include in the logistic regression model: the 30-minute change in GPS Precipitable Water, the 7.5-hour change in GPS Precipitable Water, the current Precipitable Water, and the K-Index. The 2-hour forecast interval specifies a WW of 30 minutes with the remaining 90 minutes accounting for real-world expected delays between measurement and the warning being issued. Since this forecast interval more closely aligns with the intent of this paper, these results are reported and the 9-hour forecast interval results are omitted. The evaluation metrics are presented in Table 7. Kehrer et al. (2006) discussed the potential improvements by using a nonlinear method such as neural networks and by adding other predictors such as EFM data.

In summary, Kehrer et al. (2006) established an AOC of 5 nautical miles radius around Cape Canaveral, using a WW of 30 minutes. As with Mazany et al. (2002), an expiration was not explicitly given to the warning, so PW is difficult to determine; however, Kehrer et al. (2006) provided a confusion matrix and a full report of their evaluation metrics. The variables used in this study had MWs and SWs ranging from 7.5 hours to 0 (when measuring a current value of an instrument). Since the variables did not include EFM data, this is not as important to this paper.

In 2008, Murphy et al. (2008) attempted to integrate EFM data into the previous work by Murphy and Holle (2006) in order to better understand how EFMs could

contribute to lightning prediction methods. They used a threshold method that predicted lightning if the SW = 10 second and SW = 60 second means of two EFMs simultaneously read greater than a threshold. Since the mean of these time periods was the threshold measured, the MW is also set to the same 10 seconds and 60 seconds, respectively. Two thresholds of 1 kV/m and 2 kV/m were evaluated. The EFMs were located 3.8 kilometers (km) apart and the rectangular AOCs with sides measuring 20 and 40 km were centered between them. This research set WW = 0 since the prediction was made for the 15 minutes (PW) immediately following the MW.

As with several studies presented in this paper, the PW was not explicitly stated. Since the number of a FPs was presented, some criteria must have been used to determine when a warning should expire and a FA be counted. Murphy et al. (2008) does mention that a warning expires 15 minutes after a criteria is no longer met. This paper infers that this is the PW.

Using their criteria method and the two thresholds for the EFM means, they determined that the inclusion of the EFM data actually decreased the effectiveness of lighting prediction. They suggest that the effective range of an EFM does not allow it to detect the electrification of storm clouds which are usually high altitude given the size of the chosen AOC. By varying the size of the AOC and the orientation or distribution of EFMs around the AOC may improve results (Murphy et al., 2008). Their evaluation metrics are presented in Table 7.

Beasley, Williams, and Hyland (2008) examined if EFM readings exceeded a threshold a number of minutes before a lightning event within a certain radius of the lightning event. They asked three questions:

1. Given an amount of time, a distance, and a threshold, what fraction of cloud-to-ground lightning events do the EFM readings exceed the threshold within

the distance around the lightning event location?

2. In what fraction of cloud-to-ground lightning events do the EFM readings not exceed a threshold within a distance around the lightning event location?

3. Given a time window, a distance, and a threshold, in what fraction of the cloud-to-ground lightning events do the EMF readings exceed the threshold within the distance around the lightning event location?

Since Beasley et al. (2008) did not create a predictive model, this thesis' framework does not match exactly. However, Beasley et al. (2008) implied the intent of the work is to inform lightning warnings when they mentioned their work aligning with the Weather Launch Commit Criteria. For this reason, this thesis attempts to put their work into the same common terminology to allow for comparison of their findings with explicitly predictive models. Their methods thus established a SW of 0.

The first question they asked specified times of 10, 5, 2, and 1 minutes with distances of 10, 5, 2, and 1 km from the lightning event location. The EFMs within the specified radius had to surpass thresholds of +/- 1, 2, or 5 kV/m. Translating this to the common framework gives MWs of 10, 5, 2, and 1 minutes and PWs and WWs of 0. The AOC is circular with radii equal to the distances specified, but emphasizing that this is a radius around the lightning event, not around an area of interest. This differs significantly from other methods mentioned and the methods of this thesis. With these parameters specified, Beasley et al. (2008) reported that in 81.3% of the lightning events an EFM reading exceeded a +/- 1 kV/m threshold within a 10 km radius around the lightning event within 10 minutes prior to the event. This also results in 18.7% of cases where the threshold was not exceeded with the same parameters. While other results were given, this question fails to adequately address the purpose of this thesis' research since it does not capture the idea of prediction.

The other results are omitted and the first question is not included in the summary.

For the second question, Beasley et al. (2008) only specified a distance of 1 km (AOC), a threshold of +/- 1 kV/m, and a time period of 10 minutes (MW). This method resulted in a FNR in 18.7% of the cases.

For the third question, and the question most pertinent to this paper, Beasley et al. (2008) establishes thresholds of +/- 1, 2, and 5 kV/m, distances of 10, 5, 2, and 1 km (AOC), and time windows of 3-6, 6-9, 9-12, and 12-15 minutes. Using these time windows rather than a time period prior to the lightning event allowed for a notion of a WW. Thus, given the MWs of 3-6, 6-9, 9-12, and 12-15 minutes, the corresponding WWs were 3, 6, 9, and 12 minutes, respectively. PWs remained at zero. With these parameters, Beasley et al. (2008) saw EFMs exceed a 1 kV/m threshold at least 9 minutes prior to the strike within 10 km of the lightning event in 56% of cases, implying 44% of the cases this did not happen. Other results were given, but these parameters are highlighted and most represent the intent of this paper.

In 2011, Da Silva Ferro et al. (2011) a single EFM in São José dos Campos, Brazil and a threshold method to predict lightning events in order to issue lightning warnings. Their research was unique because of the altitude of the EFM. The AOC being collocated with an EFM sitting at 800 meters above sea level allowed for a more accurate reading of the electrification of a storm in the AOC, rather than the atmosphere of the ground below or near a storm, which is the case of studies with EFMs at lower altitudes.

Da Silva Ferro et al. (2011) used thresholds of: +/- 0.5; 0.8; 0.9; 1.0; 1.2; and 1.5 kV/m. They used two types of AOC: circular and annular around the EFM. Their annular method used regions ranging from 0 to 5 km; 5 to 10 km; and 10 to 15 km. Their circular method used radii of 5, 10, and 15 km. They established a PW of 45 minutes; that is, when a lightning warning was issued, a lightning event was expected

within 45 minutes of the issuing. If no lightning event occurred, then the warning expired after this 45 minutes period. If at any point the mean of the EFM values across a minute exceeded the threshold, then a warning was issued. This established their SW = 1 minute, MW = 1 minute, and WW = 0, meaning that the PW started immediately following a threshold-exceeding EFM mean reading. They did establish a notion of Lead Time (LT), which they defined as the time between a warning and a lightning event. This was gathered post-facto and not used in the same manner that WW is defined in this paper. However, for all true positive lightning warnings these LTs were recorded to give a notion of a warning metric. Da Silva Ferro et al. (2011) reported with a 1 kV/m threshold and a circular 10 km AOC around the EFM a TPR of 58% and a FPR of 41%.

Parameter Comparison.

This section uses the common terminology presented to compare the research methods and results of the aforementioned studies. Table 3 compares the AOCs and Table 4 compares the SW, MW, WW, and PW. Table 7 compares the TPR, FPR, FNR and OUI, when possible.

As discussed above, lightning prediction requires a notion of time and location. This paper calls the location the AOC and is defined in terms of the its center, the area it covers, or its shape. In most cases, researchers chose circular AOCs centered around a point of interest (POI), such as a launch site at Cape Canaveral (Mazany et al., 2002). In other cases, a technique was used to widen the AOC and create a larger area using a rectangle, as in Lambert et al. (2005) or Murphy et al. (2008). Another approach taken by Da Silva Ferro et al. (2011) created a circle around the EFM instead of a point of interest. Most researchers chose a distance from the center

of the AOC of 5 or 10 miles. Table 3 outlines the AOC parameters chosen for each of the studies.

Table 3. Area of Concern Parameters

Researcher	Location Relative to Sensors	Shape	Size
Mazany et al. (2002)	POI	Circular	20 nm
Kehrer et al. (2006)	POI	Circular	5 nm
Murphy et al. (2008)	POI	Square	10, 20 km
Beasley et al. (2008)	POI*	Circular	10, 5, 2, 1 km
Da Silva Ferro et al. (2011)	EFM	Circular	10, 5, 2, 1 km
Da Silva Ferro et al. (2011)	EFM	Circular	5, 10, 15 km

Point of Interest (POI) is a location chosen by the researcher for which the lightning warnings are issued

*Beasley et al. (2008) establishes the AOC around the lightning event and analyzes the sensor measurements around the event, thus making the AOC the event rather than a pre-established location

Once the location parameters are chosen, a researcher must scope the time. Many of the researchers looked at the EFM readings without summarizing them, resulting in a SW of 0. Others chose to smooth the reading and established a SW of 10 to 60 seconds. The WW also ranged widely, if used at all. The PW ranged from 0 to 13.5 hours. Table 4 shows the selections of each parameter.

Table 4. Time Window Parameters

Researcher	Statistics Window	Measure Window	Warning Window	Prediction Window
Mazany et al. (2002)	30m	30m	0	13.5h
Kehrer et al. (2006)*	Varied	Varied	30m	0
Murphy et al. (2008)	10s, 1m	10s, 1m	0	15m
Beasley et al. (2008)	20s	3m	10m, 5m, 2m, 1m	0
Da Silva Ferro et al. (2011)	1m	1m	0	45m
*Kehrer et al. (2006) the statistics window varied for each variable in their model s= seconds, m = minutes, h = hours				

With the space and time parameters clearly identified, the methods must be evaluated. The metrics comparing the ability of the models to predict lightning are shown in Table 7. When a statistic cannot be calculated, inferred, was not presented, or is not applicable, the metric is left blank. If a paper does not specify how a metric was calculated, it is assumed the method of calculation in Frei (2008) was used.

Table 5. Method Results Comparison

Researcher	TPR	FPR	FDR	OUI
Mazany et al. (2002)	87.5%	23.1%	30.0%	60.2%
Kehrer et al. (2006)*	95%	47%	45.3%	45%
Murphy et al. (2008)	37.7%	-	41%	-
Da Silva Ferro et al. (2011)	60%	-	41%	-

Mazany et al. (2002) statistics were calculated from the confusion matrix in Table 6.

Kehrer et al. (2006) statistics are from the validation set

Murphy et al. (2008) statistics calculated from the counts presented in Table 3 for a SW of 10 seconds and 1 kV/m threshold.

Da Silva Ferro et al. (2011) did not report values for the circular AOC method. Only the TPR and FPR are reported. Statistics for 10 km AOC and 1.0 kV/m threshold were used.

2.4 Artificial Neural Networks

General.

Artificial Neural Networks (ANNs) are machine learning methods loosely inspired by the neural synapses of the brain. ANNs, like other machine learning techniques, perform such tasks as regression, time series predictions, and classification. Their uniqueness lies in their ability to adapt to any arbitrary degree of complexity, making them capable of approximating a function to an arbitrary degree of precision (Hornik, 1991).

This capability makes ANNs particularly useful in nonlinear contexts, in which the nature of the phenomenon being studied is unknown. Their usefulness has been shown in the improvements in image classifications using Convolutional Neural Net-

works (CNNs) (Krizhevsky, Sutskever, & Hinton, 2012). Text completion, time series predictions (Sutskever, Vinyals, & Le, 2014), and speech recognition (Sak, Senior, & Beaufays, 2014) (Graves & Jaitly, 2014) have also seen vast improvements through the use of Recurrent Neural Networks (RNNs). Combining CNNs and RNNs, known as Convolutional Recurrent Neural Networks (CRNNs), or in some literature Recurrent Convolutional Neural Networks, has lead to even greater improvement.

Tsironi et al. (2017) experienced improvements to gesture recognition by combining CNNs with the Long-Short Term Memory RNN layers. Likewise, Wu and Prasad (2017) applied a CRNN to hyperspectral images for object recognition and saw improved performance over traditional CNNs and RNNs. This section of Chapter 2 explores the methods of CRNNs to ascertain why they experience performance improvements over CNNs and RNNs applied to the same problem.

Convolutional Recurrent Neural Networks.

CRNNs combine the feature extraction benefits of a CNN with the time series prediction strengths of a RNNs. CNNs convolve a matrix of weights over data, yielding a feature map indicating the presence or lack of a desired outcome. In image analysis, these high level features are generally lines, curves, shapes, or colors, and the outcomes are objects such as signs, animals, or people. RNNs are capable of such classification, but their strength truly lies in time series data.

RNNs identify long-term dependencies across datasets. This makes them useful for tasks where an outcome is dependent on both data immediately presented as well as data from past observations. In long-short term memory (LSTM) RNNs, the LSTM layers contain a memory unit that allows the net to retain information about all past observations. In this way, RNNs perform well on tasks such as sentence completion, translation, and speech recognition. By "remembering" how a sentence

was structured in past data, the RNN can predict well what word should come next given the most recent words given to it. These two tasks, feature extraction from images and sentence completion, can be related to many other tasks.

Consider weather data. In much the same way that sentence structures are similar throughout a story and that given a set of words limits the next word to follow, weather is dependent on what has happened in a region in the past and what has immediately happened leading up to a moment in time. In this way, weather data is similar to a problem optimally suited for a RNN. However, if the same information was presented as a large array of data consisting of many sensors placed throughout a region, the data may appear more similar to a hyperspectral image. This type of problem is more suited to a CNN. By combining these ideas, the CRNN is born.

Wu and Prasad (2017) applied this theory to hyperspectral images of an urban environment and a rural environment to classify the different types of terrain in each. The convolution layer in their CRNN captures the middle-level, abstract, and invariant spacial information of the images. The recurrent layers captured the contextual difference between the spectral bands of the images. In this way, the CRNN outperformed CNNs, RNNs, and several other ANNs in classifying the type of terrain in each image. CRNNs have also been applied successfully in classification of gestures.

Tsironi et al. (2017) analyzed a set of images of individuals performing 9 different gestures. By taking the differential of the images that comprised a gesture and training a CRNN (in this case, a LSTM) with them, Tsironi et al. (2017) achieved vastly improved performance over CNNs and RNNs. Tsironi et al. (2017) processed the images significantly and was careful to construct the dataset in such a way that the convolution layers and the recurrent layers had access to the features of the set of images required for it to learn meaningful information. For example, by calculating the differential images of the images comprising a gesture, the model could then extract

both the spatial features and perform a gesture sequence classification. By feeding each differential image to the CRNN and identifying the start and end of a gesture, the convolution layers extracted the change in position of the hands and body, which was then fed to the recurrent layers. These changes in position over time allowed for accurate classification of the combined gesture.

These examples of CRNNs being used to extract two distinct properties of a dataset are inspiration for the research performed in this thesis. The next chapter describes the methodology used and how, like Tsironi et al. (2017) and Wu and Prasad (2017), CRNNs can outperform other methods at classification in specific contexts.

2.5 Conclusion

Lightning is a complicated atmospheric event and is affected by many variables. The methods and data studied to predict lightning vary greatly. All of the presented works show the trade-off between FAR and TPR. This thesis analyzes the effectiveness of predicting lightning using CRNNs and the electric potential gradient measured by electric field mills.

III. Methodology

3.1 Introduction

This section describes the methodology chosen in this analysis. First, the equipment is described so duplication of the research is possible. Next, the datasets used are described in detail including the cleaning, augmentation, and reduction techniques. Lastly, the ANN used is detailed to include the reasoning for hyperparameter selection and network structure.

3.2 Materials and Equipment

Software.

ANNs require several pieces of software working together to perform the extensive calculations at the heart of neural networks. The recent resurgence of ANNs in deep learning has resulted in a number of open-source software packages, but these packages are sensitive to changes in other software versions. For this reason, this thesis states the specific software versions used to produce this research.

This research used Python 3.6.4 with PyCharm 2017.3.2 as an IDE running on a 64-bit Windows 10 Enterprise operating system. The machine used an Intel Xeon CPU E5-2680 and a NVIDIA Quadro K6000. Keras 2.0.9 with a TensorFlow 1.4.0 backend was chosen to produe the ANNs. To utilize the GPU maximally, CUDA 8.0 and cuDNN v5.1 were used. These versions are specified because they are not the most updated versions.

Sensors.

EFMs are instruments that measure the voltage per meter (V/m) of the nearby atmosphere. EFMs measure at a rate of 50 hertz with values ranging from approxi-

mately -15,000 V/m to 15,000 V/m (Livingston & Krider, 1978). There are 31 EFMs in the area surrounding the 45 WS (Finn et al., 2010). These sensors measure voltage per meter, but are not capable of explicitly identifying a lightning event.

The LDAR array records lightning flashes and is comprised of nine sensors surrounding the KSC. The LDAR array detects lightning flashes up to 100km from the KSC, providing the precise location in X, Y, Z distances from the center and the time of the lightning event using Chi-squared minimization of the estimate of all sensors (Global Hydrology Resource Center, n.d.). This thesis combines the LDAR lightning data with the EFM voltage per meter measurements to produce a dataset identifying the state of the atmosphere at the time of the lightning event.

3.3 Dataset Creation

The EFM data was pulled from the https://kscwxarchive.ksc.nasa.gov/ using a script written in Python 2.7. Due to server settings that restricted the amount of time a host can maintain connection to the server, the data was pulled in 5 day increments as a .zip file for each 5 day group. The data unzipped into a folder for each day with each folder containing .zip files for each 30 minute period during that day. Each of these .zip files unzipped into a folder that contained a .dat file. These .dat files were text files that had been compressed and required a .exe executable file from https://kscwxarchive.ksc.nasa.gov/ to extract the usable information. The .exe extracted a collection of .RAW data files, one file for each EFM. These .RAW data files required their extension to be changed to .txt to before being ready for further processing. The entire process from the initial unzipping of the five-day group file to the changing of the extension was called the preprocessing step.

The EFM data, once preprocessed, was organized into a folder for each day of data, meaning that folder contained 48 folders corresponding to every 30 minute time

period that made up a 24 hour day. Each of these folders contained 31 files of data, one file for each of the 31 EFMs. Thus, for a complete day of data, there were 31 x 48 = 1,488 data files, each with 90,000 (30 minutes worth of 50 hertz readings) observations. The data files contained three columns of data: date, time, and voltage per meter (V/m). This thesis analyzed data from May through July of 2013-2016, 301 out of the 364 days in this time period were available for download and were included in the analysis.

The first action taken to clean the data was to remove data from faulty sensors. An EFM consistently failed to provide meaningful data; the data from this sensor was removed entirely leaving 30 EFMs. The next step was to handle instances repeated data. Several files contained multiple EFM readings with the same date/time. Only the first occurring instances of these repeated data points were kept. Missing data was accounted for after the summary statistics were calculated.

Four summary statistics were chosen: mean, variance, skewness, and kurtosis. By using four summary statistics, a more holistic analysis of a time period was possible. For example, Figure 8 shows the graph of multiple probability density functions with mean = 0 and variance = 1. The nature of these distributions is very different, even though two of their summary statistics would indicate they are the same. This figure demonstrates the limited perspective given by only using one summary statistic to describe data. In the EFM data, if only mean and variance were measured, important information about the nature of the atmosphere may be missed. For this reason, all four statistics are calculated using several time periods to ensure the most accurate description of the electric gradient of the atmosphere is obtained.

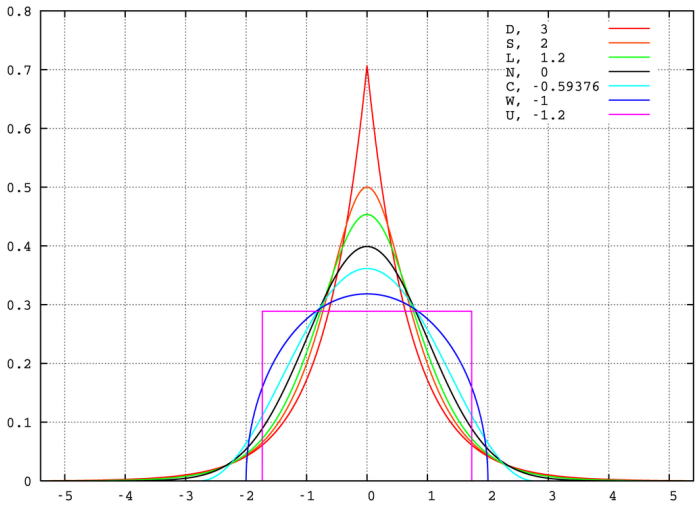


Figure 8. Effect of Kurtosis on Distributions with Mean=0 and Variance=1

These time periods were called statistics windows (SWs). The data was summarized over SWs 20, 30 and 60 seconds. This resulted in a dataset with timestamps of the start of the time period over which the data was summarized. For example, for time period 6/12/2013 12:00:00 to 6/12/2013 12:01:00 (3000 observations) and SW = 20, 3 observations were calculated with timestamps 12:00:00, 12:00:20, and 12:00:40. Additionally, this expanded the number of variables from 30 V/m readings to 120 statistics, 4 for each V/m reading. The 4 statistics were called channels, remaining consistent with the terminology of 4D tensors in CNNs. A fifth channel was added to account for any missing data.

The fifth channel was a binary value that indicated if insufficient data was available to calculate the summary statistics. If there were less than four values within a SW, then attempting to calculate the statistics would result in a NaN for at least one of the statistics. The corresponding fifth channel would be set to zero, indicating missing data. After the fifth channel indicator variables were set, any NaNs that resulted from attempting to calculate the statistics with insufficient data were set to 0. Using this fifth channel indicator variable allowed the ANN to learn which readings contained missing data and how to appropriately account for the resulting zero values. By

making this indicator variable a channel, this allowed for an indicator of missing data to be assigned to each EFM and each SW. To ensure the integrity of the data, the number of SWs for which all EFMs during that SW had indicator variables indicating missing data were counted. If the sum total of time of these SWs totaled more than 1 hour, then the entire day was removed from the analysis. Additionally, any day that did not contain a lightning event within the AOC was excluded from the analysis.

The 45 WS provided the preprocessed LDAR data for 2013-2016 . While additional data was available on the https://kscwxarchive.ksc.nasa.gov/ website, it required additional processing and interpreting to be meaningful to this analysis. For this reason, only the provided 2013-2016 data was used and the scope of the analysis was restricted to this time frame. The preprocessed LDAR data included date, time, and distances of lightning events from the center of the LDAR system. The LDAR data was processed to remove all lightning events not occurring within the AOC. Two circular AOCs were established with radii of 5 miles and 10 miles, both centered at the LDAR center at KSC. This created two LDAR datasets, one for each AOC. By comparing the time of the lightning event to the EFM statistic window time period, a response of 1 or 0 was given to the corresponding observation. The response was set to 1 if a lightning event occurred within the AOC in the SW and a 0 otherwise.

Analyzing these statistics revealed that the data was unfit for use in a neural network. Table 6 shows the maximum, minimum, mean, and variance of the means and variance calculations for 11 May 2013. These values are extremely large and could cause exploding gradients in an ANN. To prevent this and to maximize the likelihood of training, all statistics were min-max scaled.

Table 6. Descriptive Statistics Describing the Statistics Datasets

	Dataset of SW=20		Dataset of SW=30		Dataset of SW=60	
	Mean	Variance	Mean	Variance	Mean	Variance
Max	6546.12	4.5×10^7	6146.08	3.5×10^7	6114.21	3.5×10^7
Min	-9531.57	0	-8243.65	0	-7562.07	0
Mean	81.26	59440.98	81.19	69284.52	80.95	87480.06
Variance	1.2×10^9	1.2×10^{23}	1.2×10^9	1.2×10^{23}	1.1×10^9	2×10^{23}

To mitigate the risk of vanishing or exploding gradients, all the data was min-max scaled using Equations 7 and 8 where min and max are the minimum and maximum values the scaled data can achieve. The $\min(X)$ and $\max(X)$ are the minimum and maximum values attained by the Xth EFM to which the x value belongs; that is, x is a single value to be scaled from the Xth EFM that attained a maximum and minimum value of M and m, respectively, over all the observations.

$$x_{scaled} = x_{std} * (\mathrm{M} - \mathrm{m}) + (\mathrm{m}) = 2x_{std} - 1 \tag{7}$$

$$x_{std} = \frac{x - \min(X)}{\max(X) - \min(X)} \tag{8}$$

In summary, data from the EFM sensors near Cape Canaveral were collected from the https://kscwxarchive.ksc.nasa.gov/ website. The data was cleaned and consolidated into datasets containing 30 minutes worth of readings for 30 of the 31 EFMs. The data was then summarized across SWs 20, 30, and 60 seconds by taking the mean, variance, skewness, and kurtosis over each SW. A missing data indicator was added to identify statistics there were calculated without all SW of the readings available. This resulted in 3 separate datasets of 150 variables each. The LDAR data was

then modified by removing all lightning events that did not occur within AOCs of 5 and 10 miles radii centered at KSC. This resulted in 2 LDAR datasets. Combining the EFM and LDAR datasets allowed for the creation of the response variable by determining if a lightning event occurred during the time period over which the summary statistics were calculated: 1 if yes, 0 if no. This resulted in 6 distinct statistic datasets: $3\,\text{SWs} * 2\,\text{AOCs} = 6$ datasets.

The intent of this thesis was to predict lightning with sufficient warning time to seek shelter. To reflect this intent, the data was restructured. MWs of 30 and 60 minutes were chosen and a PW and WW were both set to 15 minutes. A running window of MW was taken over each statistics dataset resulting in 12 windowed datasets. For example, one day of the SW=20 dataset was originally 4320 observations, since there are 4320 non-overlapping, contiguous 20 second periods in a day. If the MW=30 minutes, then this dataset becomes a 3-dimensional dataset of (4140, 90, 150). There are 3 SWs per minute; thus, 90 SWs per MW, 45 SWs per PW, and 45 SWs per WW, which leads to 4140=4320-90-45-45. Some data was lost in this restructuring, but it was lost towards midnight where few lightning events occur. In general, this conversion resulted in 3-dimensional datasets consisting of (# of SW observations-MW-PW-WW, sequence length (MW), 30 EFMs x 5 channels = 150). To account for the prediction of lightning at a future time, a new response was created based on the selection of the WW and PW.

Previously each SW received a response. In this new construct, each MW time period of data receives a response. If at any point in the PW a lightning event occurred, then the entire MW received a 1 response. Figure 9 from Chapter 2 depicts the method used to create the new response. The notion of lightning onset prediction was attained by using a running window.

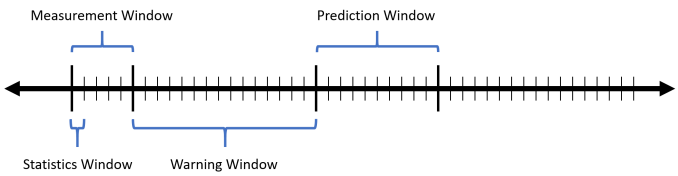


Figure 9. Graphical Representation of Windows

To appropriately evaluate the effectiveness of the model, the data was split into training, validation, and test datasets by randomly selecting 70% of the day datasets for training, 15% for validation, and 15% for testing. Since the number of lightning events versus nonlightning events was highly skewed toward nonlightning events, the training data was randomly oversampled to produce equal number of lightning and nonlightning events. The 3-dimensional arrays of (# of SW observations, sequence length (MW), 150) were then split into (# of SW observations, sequence length (MW), 30 EFMs, 5 channels) to reflect the notion of a hyperspectral image over time. This was thought of as a 5 channel picture of the regional electric gradient of the atmosphere.

3.4 Motivation and Creation of the ANN

A Convolutional Recurrent Network Approach.

CNNs are widely used in image classification due to their ability to create filters that identify distinguishing characteristics, i.e., vertical/horizontal lines, changes in colors, shapes, etc. It is hypothesized that EFM readings may also be able to capture some features of lightning events immediately prior to their happening. In Beasley et al. (2008), they discuss one such characteristic when they state that EFM readings often show a steep gradient several minutes prior to a flash in an area 5 km to 10 km from where the flash occurs. In the same way that CNNs identify distinguishing characters of objects in images, they also can identify patterns in EFM data that

indicate lightning events.

To express the motivation for using CNNs, consider a dataset consisting of 18000 EFM readings from 30 EFMs with the first reading being at time t_0. Since readings are taken at 50 hertz, these 18000 readings cover a time period of 6 minutes. Suppose this data is summarized with four statistics over a SW of 1 minute, reducing the dataset to 6 observations. This results in a dataset with index of $t_0, t_1, t_2, t_3, t_4, t_5$ where each index represents the starting time over which the data was summarized. The 120 columns of this dataset arise from taking the 4 summary statistics for each of the 30 EFMs. The dimension of the entire dataset is (6, 120): 6 observations and 120 statistic values. Alternatively, this could be seen as a dataset of size (6, 30, 4), but to remain consistent with how the actual data is organized and processed, this example will remain a dataset of size (6, 120).

Now set the MW to 3 minutes. The dataset is then transformed into a 3D tensor where each observation consists of 3-1 minute periods. This results in four observations: $[t_0, t_1, t_2]$, $[t_1, t_2, t_3]$, $[t_2, t_3, t_4]$, and $[t_3, t_4, t_5]$. The dimension of the dataset is now (4, 3, 120), since there are 4 sets of time periods each consisting of 3 statistic windows of 120 statistic values.

The data is reordered into a 4D tensor to place all four statistics for an EFM in the same dimension. That is, for t_0 of the first observation, the dimension is now (1, 1, 30, 4) where 30 represents the 30 EFMs and 4 represents the 4 summary statistics. By performing this on each time period, the dataset dimension is now (4,3,30,4): 4 observations, 3 time periods in each, 30 EFMs, and 4 statistics per EFM.

In this form, 5 kernels of size (1,30) are convolved across each observation, resulting in a new observation of dimension (1, 5) where each pass yields a sum of dot products of the kernel values with the values of the 4 statistics for each time period and each EFM. The dimension of the dataset is reduced to (4, 5) where each observation is

now a feature map of the entire region over the time periods $[t_0, t_1, t_2]$, $[t_1, t_2, t_3]$, $[t_2, t_3, t_4]$, and $[t_3, t_4, t_5]$. Each of these four observations are correlations of all the EFMs across the region over a SW window. This example illustrates the motivation for the selection of the model used in this analysis and how the data was visualized as a hyperspectral image. The methods and procedures in this example are similar to that used in this paper.

To translate this into the context of this analysis further, consider Figure 10. The correlation of the change at each of the 31 EFM locations should yield a more complete picture of the electrification of the region compared to readings from any subset of the EFMs. Figure 10 shows the location of all 31 EFMs in the Cape Canaveral region, as well as the center of the LDAR array which is co-located with the center of the 5 mile radius AOC chosen for this study.

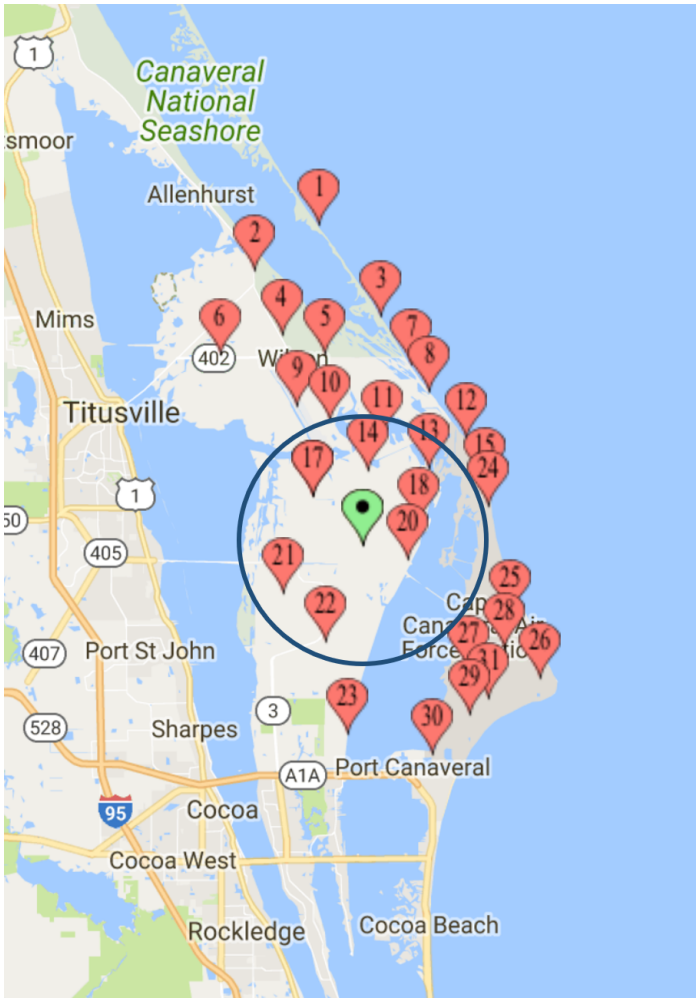


Figure 10. Location of EFM Sensors (red) and the center of the LDAR array (Green) and a five mile AOC around the LDAR center (Google Maps, 2018)

RNNs specialize in time series data. Their usefulness has been shown in speech and text recognition and in video analysis. Since the EFM datasets contain both a notion of time and space, combining a CNN and RNN into a CRNN allowed for their combined specializations to improve lightning prediction. By adding a long-short term memory recurrent layer after the (1,30) convolution, the CRNN would learn the patterns of change seen across the region in the minutes leading up to a lightning event in the AOC.

Model Selection and Hyperparameter Optimization.

Model selection began with a single LSTM layer with 4 neurons and a single neuron dense output layer with a sigmoid activation function. Adam was chosen as the optimizer and binary crossentropy as the loss function for all iterations of the model selection process and the final model (Kingma & Ba, 2014). This initial was able to achieve a 65% accuracy on the training dataset which indicated it did not have sufficient capacity to overfit to the training data. More neurons were added, and the most improvement was seen when 16 neurons were used.

Two additional LSTM layers were added each with 16 memory cells and the training accuracy jumped to 76%. A number of different initial learning rates were used, but the best performance was seen when using an initial rate of 0.1 decaying gradually. A convolution layer was added before the LSTM layers to attempt to capture patterns of change of each EFM across small windows of time. Eight kernels of size (3,1) were used with same padding and a Rectified Linear Unit (ReLU) activation function. This resulted in an improvement to 83%. Adding an additional convolution layer with the same parameters improved the accuracy to 86%.

Next, a convolution layer across the EFMs was used to capture the correlation of the EFMs across a time period. Eight kernels of size (1,30) were used with valid padding and a ReLU activation function. This saw an improvement to 93%. By increasing the number of kernels on all convolution layers to 16 and decreasing the initial learning rate to 0.01, the final configuration of the model was established. It achieved a 96% accuracy in the training data.

Once a model was found that overfit to the training data, regularization was added and the validation loss and accuracy were observed. Adding dropout layers with 30% dropout between the second and third convolution layers, between the third convolution and the first LSTM layer, and between the third LSTM layer and the output

layer resulted in a steadily decreasing validation loss over several epochs. This verified that the architecture had both the capacity necessary to learn the datasets, but could generalize to a validation dataset. A diagram of the model is shown in Figure 11. This model trained on all 12 datasets and then ran against the corresponding test datasets to calculate metrics to determine the models quality. These results are presented in the Chapter 4.

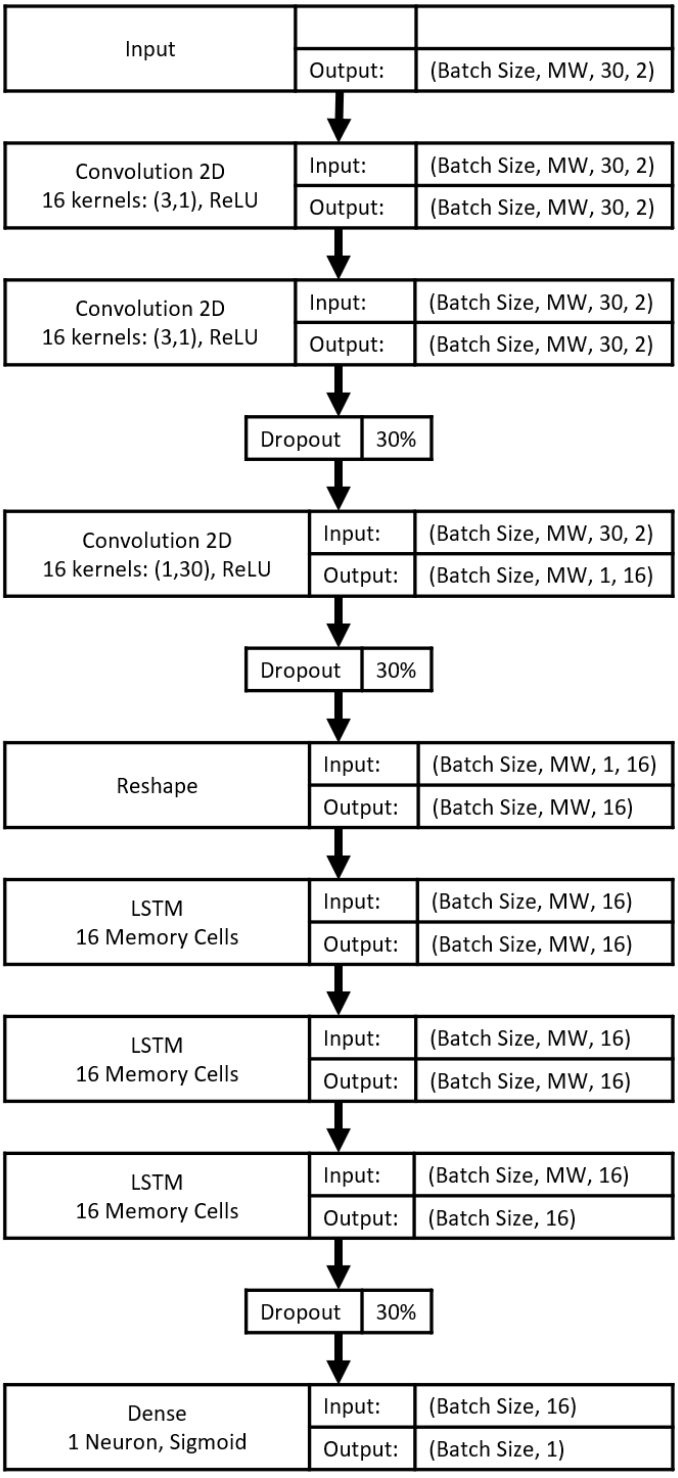


Figure 11. CRNN Architecture

IV. Discussion and Results

Due to time constraints and hardware limitations, only the dataset with SW = 60, MW = 30, and AOC = 5 miles was used to train the model. Additionally, the dataset was reduced in size to only include the mean of the EFMs and the missing data indicators. By analyzing such graphs as shown in Figure 12, it was concluded that mean may be sufficient to predict lightning. This graph shows the mean value of the last SW in 10000 sequences on the y-axis and what sequence the SW came from on the x-axis. The orange dots are the corresponding responses for the 10000 MW where a 0.1 reading is a lightning event and a 0 is a nonlightning event. The stark difference of the mean measurements during lightning events versus nonlightning events at this one EFM site indicates a clear pattern exists.

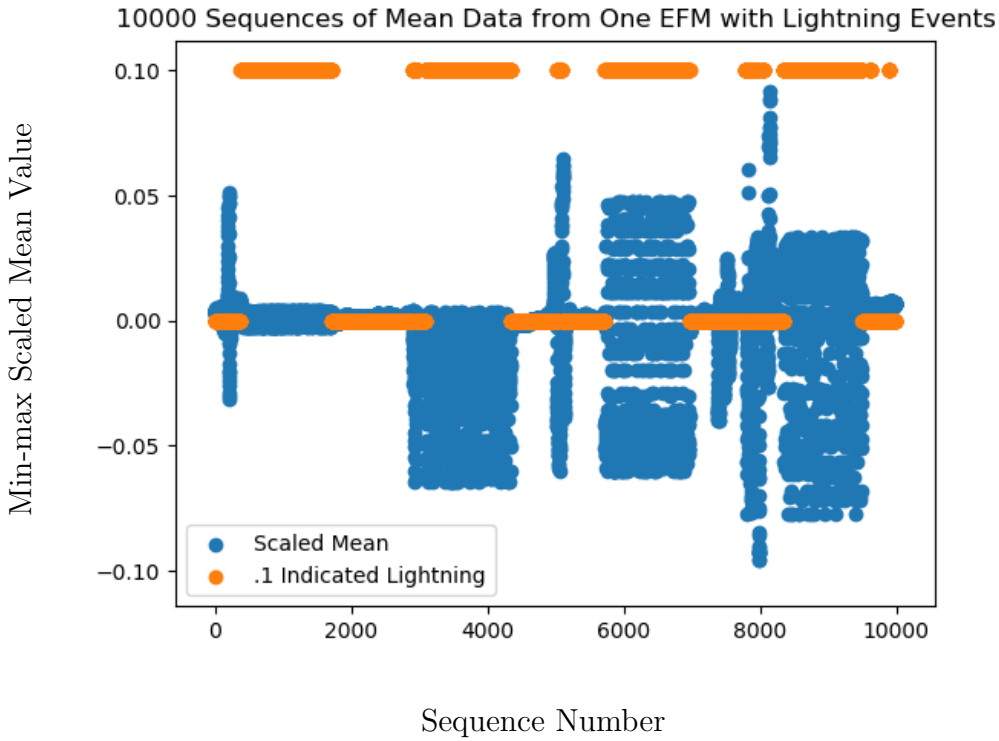


Figure 12. Mean of EFM with Lightning Event

Figures 13 and 14 show the training and validation accuracy and loss history for this dataset. The model trained for 27 epochs, but saw a plateauing loss on the validation set around epoch 15. The training loss continued to improve, but this was likely overfitting.

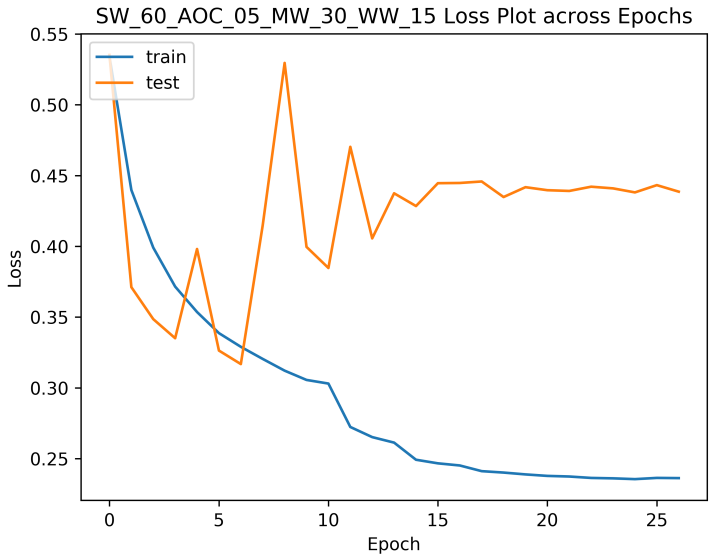


Figure 13. Training and Validation Loss History

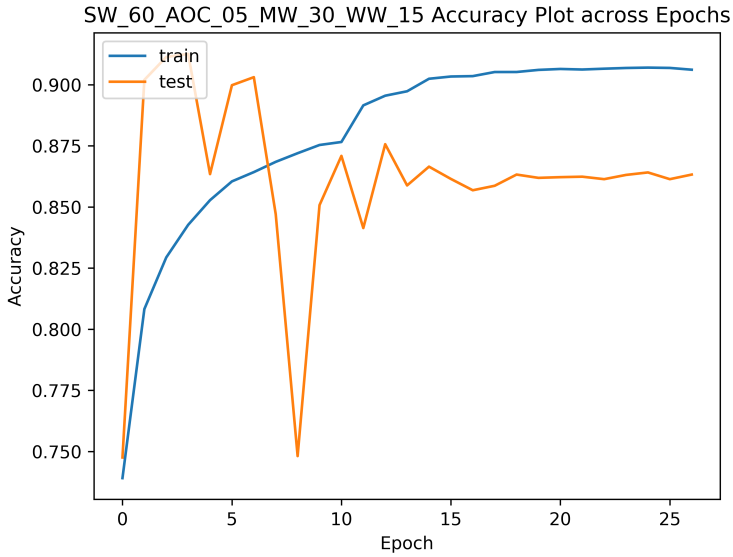


Figure 14. Training and Validation Accuracy History

Both the validation loss and accuracy were somewhat erratic towards the beginning of training. This is likely due to a high initial learning rate or to the sparse nature of the validation data. With additional optimization of the learning rate and kernel initialization weights, better validation loss may be achieved.

The model with the lowest validation loss was chosen and used on the test data. It achieved an accuracy of 90.3% on the test dataset. Since the default evaluation threshold for model predictions is .5, a Receiver Operating Characteristic (ROC) curve was used to identify a threshold that optimized the Operational Utility Index (OUI). Figure 15 shows this ROC curve with pointers identifying where the model achieved its OUI. The Area Under the Curve of a ROC is also an indication of a model's performance. The AUC for the presented ROC is 0.88 where 1 is a perfect model. Since the 45 WS established the OUI as the weighted metric that optimizes models for their intents, this paper focused on the it.

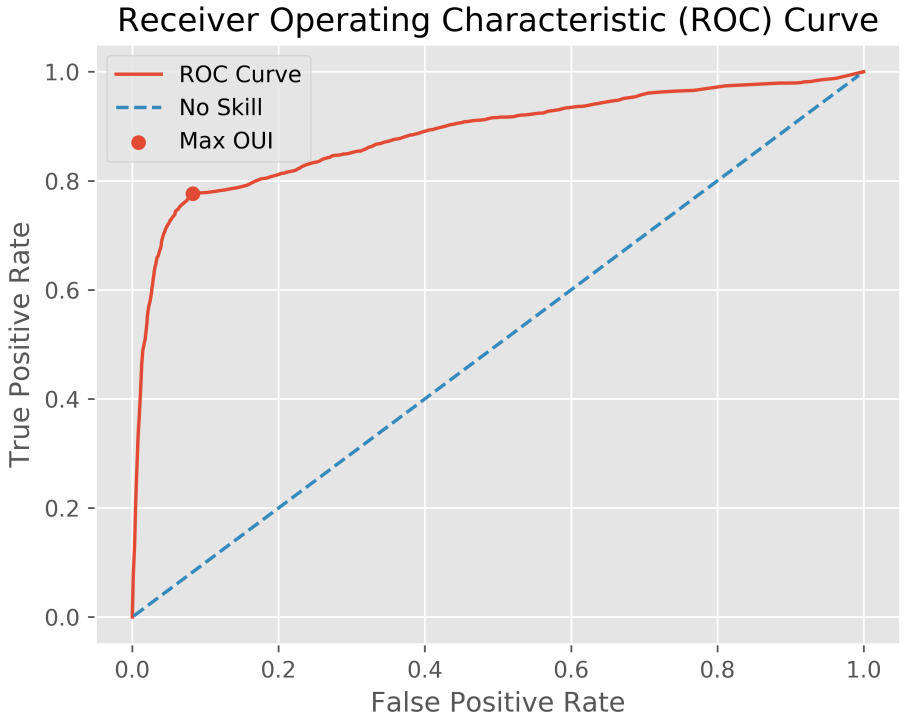


Figure 15. Receiver Operating Characteristic Curve

The maximum OUI was obtained at a threshold of 0.547. A confusion matrix was created at this threshold as is shown in Figure 16. The model achieved an overall accuracy of 90.3% with a True Positive Rate (TPR) of 77.6%, a Probability of False Detection (POFD) of 8.3%. and an OUI of 53.9%.

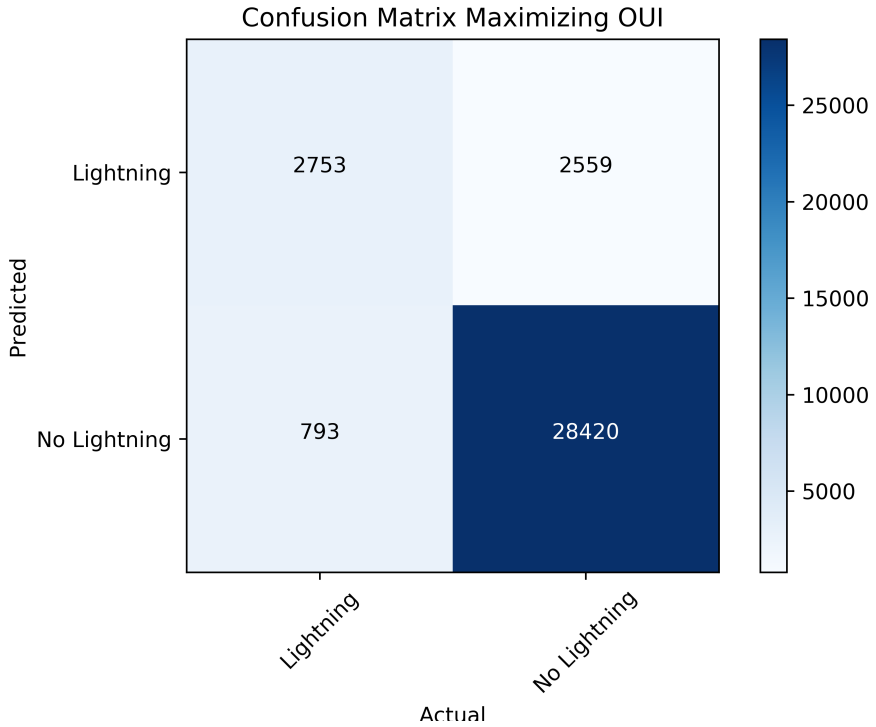


Figure 16. Confusion Matrices

These metrics are displayed in Table 7 alongside the past research discussed in Chapter. This method achieved comparable results to past methods in both the TPR and OUI, falling in the middle of all methods for which these metrics could be calculated. The FPR was significantly lower than all methods presented, but this may be the result of of the short WW and PW used in this method as opposed to Mazany et al. (2002) and Kehrer et al. (2006). In summary, these results indicate that if OUI is a metric that the 45 WS deems a good evaluation of methods, then this method performs comparably to past methods that used EFMs as predictors.

Table 7. Method Results Comparison

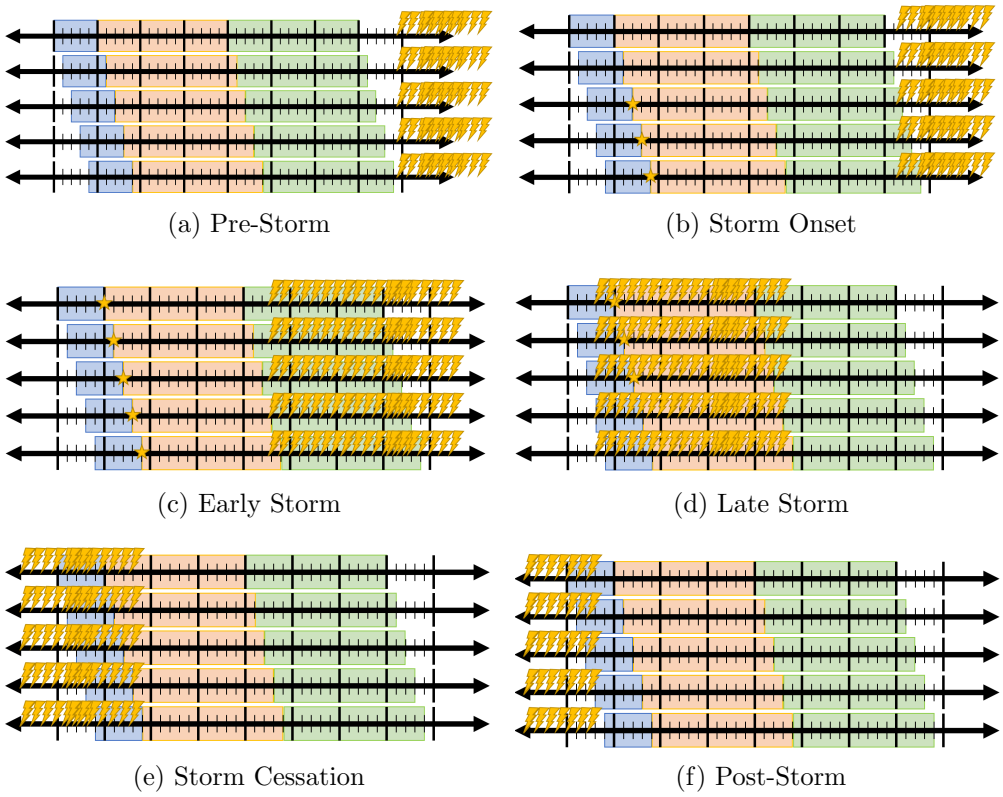
Researcher	TPR	FPR	FDR	OUI
Mazany et al. (2002)	87.5%	23.1%	30.0%	60.2%
Kehrer et al. (2006)	95%	47%	45.3%	45%
Murphy et al. (2008)	37.7%	-	41%	-
Da Silva Ferro et al. (2011)	60%	-	41%	-
This Paper	77.6%	8.3%	48.1%	53.9%

V. Conclusion and Future Work

This paper demonstrates that EFMs have a short-term predictive relationship with lightning events in the Cape Canaveral area. Using a method that takes into consideration 30 EFM sensors and accounts for both the temporal and geographic nature of lightning, the EFM array alone is capable of predicting lightning with a comparable success rate as other methods according to the 45 WS's OUI. The LLCC currently uses a threshold from a nearby EFM as a determining factor to permit space flight launches (Koons & Walterscheid, 1996)(McNamara, Roeder, & Merceret, 2010). Referencing works such as Murphy et al. (2008) who saw poor predictive results when using an EFM threshold as a predictor, this paper suggests that a potentially improved criteria may be the change of a running mean across a 30 minute time period.

Additionally, this method provides a running lightning prediction which should allow for a prediction of both lightning onset and lightning cessation. While the results presented in this paper focus on lightning prediction for a single set of time windows (MW, WW, PW), by looking at the prediction results of multiple sets of time windows this method can be used for prediction of lightning cessation and onset. Figure 17 displays an example of a storm moving across time windows of MW= 5, WW= 15, and PW = 15 (blue = MW, pink = WW, green = PW). These values are arbitrary and are used only to demonstrate how to use this method for lightning onset and cessation prediction.

While this paper uses a traditional confusion matrix from classical detection theory in which their are a total of four possible scenarios (TP, FP, FN, TN) (see Table 8), this traditional confusion matrix can be expanded into the modified confusion matrix presented in Table 9. Establishing multiple time windows allowed for the possibility of 12 outcomes. Table 8 and Table 9 are color coated to show how each of the four



(a) Pre-Storm (b) Storm Onset

(c) Early Storm (d) Late Storm

(e) Storm Cessation (f) Post-Storm

Figure 17. Example of Time Windows Moving Through a Storm

possible outcomes in Table 8 correspond to the possible outcomes in Table 9.

Using the hypothetical example of a storm moving across an AOC in Figure 17 and assuming a perfect predictor, Tables 10 a-f show the classical confusion matrices for the corresponding storm state from Figure 17. These tables look very similar and without labels Tables b and d cannot be distinguished. However, by looking at the same storm example and using a modified confusion matrix, a pattern emerges.

Looking at Table 11 shows a clear distinction between storm phases. The Storm

Table 8. Classical Confusion Matrix Example

	Actual	
	Lightning	No Lightning
Predicted Lightning	True Positive	False Positive
Predicted No Lightning	False Negative	True Negative

Table 9. Modified Confusion Matrix Example

	Actual							
	Lightning in PW				Lightning Not in PW			
	In WW		Not in WW		In WW		Not in WW	
	In MW	Not in MW	In MW	Not in MW	In MW	Not in MW	In MW	Not in MW
Predicted Lightning								
Predicted No Lightning								

Onset phase is denoted by predictions being made for lightning in the PW only and predictions being made for no lightning without any other lightning happening in the other windows. The Storm Cessation phase is distinguished by no lightning predictions being made but ongoing lightning happening within the MW and within and out of the WW. Examining this table suggests a follow-on study using a similar method as in this paper may be useful in predicting lightning cessation.

Additionally, future applications should apply this method to the multiple warning circles established by the 45 WS in Roeder et al. (2017) determine its effectiveness throughout the Cape Canaveral region. The proof-of-concept given in this paper at one location is expected to generalize across all locations near Cape Canaveral.

Future work should add the remainder of the stormy season (May-September) to this dataset and perform analysis to determine the optimal MW and SW. Expanding the results from this study to a 30 minute WW is expected to attain similar results as presented for the 15 minute WW. Adding additional features such as higher order statistics, time of day, and month of year may increase the effectiveness of this method. Lastly, the techniques from this paper should be augmented with other methods and features to improve the results of lightning forecasting.

Table 10. Classical Confusion Matrices for Hypothetical Storm Progression

(a) Pre-Storm

	Actual	
	Lightning	No Lightning
Predicted Lightning		
Predicted No Lightning		5

(b) Storm Onset

	Actual	
	Lightning	No Lightning
Predicted Lightning	3	
Predicted No Lightning		2

(c) Early Storm

	Actual	
	Lightning	No Lightning
Predicted Lightning	5	
Predicted No Lightning		

(d) Late Storm

	Actual	
	Lightning	No Lightning
Predicted Lightning	3	
Predicted No Lightning		2

(e) Storm Cessation

	Actual	
	Lightning	No Lightning
Predicted Lightning		
Predicted No Lightning		5

(f) Post-Storm

	Actual	
	Lightning	No Lightning
Predicted Lightning		
Predicted No Lightning		5

Table 11. Modified Confusion Matrices for Hypothetical Storm

		Actual							
		Lightning in PW				Lightning Not in PW			
		In WW		Not in WW		In WW		Not in WW	
		In MW	Not in MW	In MW	Not in MW	In MW	Not in MW	In MW	Not in MW
Pre-Storm	Predicted Lightning								
	Predicted No Lightning								5
Storm Onset	Predicted Lightning				3				
	Predicted No Lightning								2
Early Storm	Predicted Lightning		2		3				
	Predicted No Lightning								
Late Storm	Predicted Lightning	3							
	Predicted No Lightning					2			
Storm Cessation	Predicted Lightning								
	Predicted No Lightning					3		2	
Post-Storm	Predicted Lightning								
	Predicted No Lightning							3	2

References

Barnes, L. R., Schultz, D. M., Gruntfest, E. C., Hayden, M. H., & Benight, C. C. (2009). CORRIGENDUM: False Alarm Rate or False Alarm Ratio? *Weather and Forecasting*, *24*(5), 1452–1454. Retrieved from http://journals.ametsoc.org/doi/abs/10.1175/2009WAF2222300.1 doi: 10.1175/2009WAF2222300.1

Beasley, W. H., Williams, D. E., & Hyland, P. T. (2008). ANALYSIS OF SURFACE ELECTRIC-FIELD CONTOURS IN RELATION TO CLOUD-TO-GROUND LIGHTNING FLASHES IN AIR-MASS THUNDERSTORMS AT THE KENNEDY SPACE CENTER. *20th International Lightning Detection Conference and 2nd International Lightning Meteorology Conference*.

Christian, H. J., Mazur, V., Fisher, B. D., Ruhnke, L. H., Crouch, K., & Perala, R. P. (1989, 9). The Atlas/Centaur lightning strike incident. *Journal of Geophysical Research*, *94*(D11), 13169. Retrieved from http://doi.wiley.com/10.1029/JD094iD11p13169 doi: 10.1029/JD094iD11p13169

D'Arcangelo, D. L. (2000). Forecasting the onset of cloud-ground lightning using layered vertically integrated liquid water. , *298*(0704), 1–71. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA381063

Da Silva Ferro, M. A., Yamasaki, J., Roberto, D., Pimentel, M., Naccarato, K. P., Magalhães, M., & Saba, F. (2011). Lightning risk warnings based on atmospheric electric field measurements in Brazil. *São José dos Campos*, *3*(3), 301–310. Retrieved from http://radiometrics.com/data/uploads/2015/01/Ferro_JATM_2011.pdf doi: 10.5028/jatm.2011

Erdman, J. (2014). *Lightning's Scary and Damaging Power — The Weather Channel.* Retrieved from https://weather.com/storms/severe/news/lightning-damage-trees-homes-buildings-ground-20130719#/4

Finn, F. C., Roeder, W. P., Buchanan, M. D., McNamara, T. M., McAllenan, M., Winters, K. A., ... Huddleston, L. L. (2010, 4). *Lightning Reporting at 45th Weather Squadron: Recent Improvements* (Tech. Rep.). 45th Weather Squadron; National Aeronautics and Space Administration. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20100038320&qs=Nm%3D4294920545%7CdOrganization%7CWeather%2520Squadron%2520%2845th%29%7C%7C4294920912%7CSubject%2520Terms%7CCLOUD-TO-GROUND%2520DISCHARGES%7C%7C4294950535%7CSubject%2520Terms%7CLIGHTNING%7C%7C4294961781%7C

Florida Climate Center, F. S. U. (200). *Thunderstorms - Florida Climate Center.* Retrieved from http://climatecenter.fsu.edu/topics/thunderstorms

Frei, C. (2008). Section 5: Forecast Evaluation and Skill Scores. In *Analysis of climate and weather data* (chap. 5). Retrieved from ftp://ftp.pmodwrc.ch/pub/x_sections_ozone/anna/analisys%20of%20climate/Scores1[1].pdf

Global Hydrology Resource Center. (n.d.). *Lightning Detection and Ranging (LDAR) Dataset.* Retrieved from https://ghrc.nsstc.nasa.gov/uso/ds_docs/ldar/ldar_dataset.html

Google Maps. (2018). *Map of EFM Locations and LDAR Center.* Retrieved from https://www.mapcustomizer.com/map/EFMs%20and%20LDAR%20

Graves, A., & Jaitly, N. (2014). Towards End-to-End Speech Recognition with Recurrent Neural Networks. In N. Lawrence & M. Reid (Eds.), *The proceedings of machine learning research* (pp. 1764–1772). Beijing, China: PMLR. Retrieved from http://proceedings.mlr.press/v32/graves14.pdf

Hanssen, A., & Kuipers, W. (1965). On the Relationship Between the Frequency of Rain and Various Meteorological Parameters. *Koninklijk Nederlands Meteorologisch Instituut mededelingen en verhandelingen*(81).

Hornik, K. (1991). Approximation Capabilities of Muitilayer Feedforward Networks. *Neural Networks*, *4*, 251–257. Retrieved from http://zmjones.com/static/statistical-learning/hornik-nn-1991.pdf

Insurance Information Institute. (2016). *Lightning — III.* Retrieved from http://www.iii.org/fact-statistic/lightning

Kehrer, K. C., Graf, B., & Roeder, W. (2006, 1). *Global Positioning System (GPS) Precipitable Water in Forecasting Lightning at Spaceport Canaveral* (Tech. Rep.). Cocoa Beach, FL: NASA Kennedy Space Center. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20130011214&qs=Nm%3D4294920545%7CdOrganization%7CWeather%2520Squadron%2520%2845th%29%7C%7C4294920912%7CSubject%2520Terms%7CCLOUD-TO-GROUND%2520DISCHARGES%7C%7C4294950535%7CSubject%2520Terms%7CLIGHTNING%7C%7C4294961781%7C

Kingma, D. P., & Ba, J. (2014, 12). Adam: A Method for Stochastic Optimization. Retrieved from http://arxiv.org/abs/1412.6980

Koons, H. C., & Walterscheid, R. L. (1996). *Lightning Launch Commit Criteria.* Retrieved from http://www.dtic.mil/docs/citations/ADA305539

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In *Proceedings of the 25th international conference on neural information processing systems* (pp. 1097–1105). Lake Tahoe, Nevada: NIPS. Retrieved from http://www.cs.toronto.edu/~fritz/absps/imagenet.pdf

Lambert, W., Wheeler, M., & Roeder, W. (2005, 1). Objective Lightning Forecasting at Kennedy Space Center and Cape Canaveral Air Force Station using Cloud-to-Ground Lightning Surveillance System Data. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20130009921&qs=Nm%3D4294920545%7CdOrganization%7CWeather%2520Squadron%2520%2845th%29%7C%7C4294920912%7CSubject%2520Terms%7CCLOUD-TO-GROUND%2520DISCHARGES%7C%7C4294950535%7CSubject%2520Terms%7CLIGHTNING%7C%7C4294961781%7C

Livingston, J. M., & Krider, E. P. (1978). Electric Fields Produced by Florida

Thunderstorms. *Cl JOURNAL OF GEOPHYSICAL RESEARCH JANUARY*, *83*(20).

Mazany, R. A., Businger, S., Gutman, S. I., & Roeder, W. (2002, 10). A Lightning Prediction Index that Utilizes GPS Integrated Precipitable Water Vapor*. *Weather and Forecasting*, *17*(5), 1034–1047. Retrieved from http://journals.ametsoc.org/doi/abs/10.1175/1520-0434%282002%29017%3C1034%3AALPITU%3E2.0.CO%3B2 doi: 10.1175/1520-0434(2002)017⟨1034:ALPITU⟩2.0.CO;2

McNamara, T. M., Roeder, W., & Merceret, F. (2010). The 2009 update to the Lightning Launch Commit Criteria. In *14th conference on aviation, range, and aerospace meteorology.* Retrieved from https://ams.confex.com/ams/90annual/techprogram/paper_164180.htm

Merceret, F. J., Willett, J. C., Christian, H. J., Krider, E. P., O 'brien, T. P., Rust, W. D., & Walterscheid, R. L. (2010). *A History of the Lightning Launch Commit Criteria and the Lightning Advisory Panel for America's Space Program* (Tech. Rep.). Kennedy Space Center, FL: National Aeronautics and Space Administration. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110000675.pdf

Murphy, M. J., & Holle, R. L. (2006). Murphy -WARNINGS OF CLOUD-TO-GROUND LIGHTNING HAZARD BASED ON COMBINATIONS OF LIGHTNING DETECTION AND RADAR INFORMATION. In *19th international lightning detection conference.*

Murphy, M. J., Holle, R. L., & Demetriades, N. W. S. (2008). CLOUD-TO-GROUND LIGHTNING WARNINGS USING ELECTRIC FIELD MILL AND LIGHTNING OBSERVATIONS. *20th International Lightning Detection Conference*. Retrieved from http://www.vaisala.com/vaisaladocuments/scientificpapers/cloud-to-ground_lightning_warnings_using_electric_field_mill_and_lightning_observations.pdf

Roeder, W. P., Hajek, D. L., Flinn, F. C., Maul, G. A., & Fitzpatrick, M. E. (2000). METEOROLOGICAL AND OCEANIC INSTRUMENTATION AT SPACEPORT FLORIDA OPPORTUNITIES FOR COASTAL RESEARCH 45th Weather Squadron. Retrieved from https://www.researchgate.net/profile/William_Roeder/publication/229037605_73_METEOROLOGICAL_AND_OCEANIC_INSTRUMENTATION_AT_SPACEPORT_FLORIDA-OPPORTUNITIES_FOR_COASTAL_RESEARCH/links/00b7d536b5a1798fcf000000/73-METEOROLOGICAL-AND-OCEANIC-INSTRUMENTATION-AT

Roeder, W. P., Mcnamara, T. M., Mcaleenan, M., Winters, K. A., Maier, L. M., & Huddelston, L. L. (2017). The 2014 Upgrade to the Lightning Warning Circles Used by 45th Weather Squadron.

Sak, H., Senior, A., & Beaufays, F. (2014, 2). Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition. Retrieved from http://arxiv.org/abs/1402.1128

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neu-

ral networks. *Advances in Neural Information Processing Systems (NIPS)*, *27*, 3104–3112. Retrieved from http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural doi: 10.1007/s10107-014-0839-0

Tsironi, E., Barros, P., Weber, C., & Wermter, S. (2017, 12). An analysis of Convolutional Long Short-Term Memory Recurrent Neural Networks for gesture recognition. *Neurocomputing*, *268*, 76–86. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0925231217307555 doi: 10.1016/j.neucom.2016.12.088

Woodcock, F. (1976). *The Evaluation of Yes/No Forecasts for Scientific and Administrative Purposes* (Vol. 104) (No. 10). Retrieved from http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281976%29104%3C1209%3ATEOYFF%3E2.0.CO%3B2 doi: 10.1175/1520-0493(1976)104⟨1209:TEOYFF⟩2.0.CO;2

Wu, H., & Prasad, S. (2017, 3). Convolutional Recurrent Neural Networks forHyperspectral Data Classification. *Remote Sensing*, *9*(3), 1–20. Retrieved from http://www.mdpi.com/2072-4292/9/3/298 doi: 10.3390/rs9030298

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
02–03–2018	Master's Thesis	Sept 2016 — Mar 2018

4. TITLE AND SUBTITLE

LIGHTNING PREDICTION USING ARTIFICIAL NEURAL NETWORKS AND ELECTRIC FIELD MILL DATA

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Daniel E. Hill

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

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

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/ENC/MS/18-M-002

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

45th Weather Squadron (45 WS/WXT)
1201 Edward H. White II Street
Patrick AFB, FL 32925
DSN 467-8410, COMM 321-853-8410
Email: william.roeder@us.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFWA

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

12. DISTRIBUTION / AVAILABILITY STATEMENT

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Electric Field Mills (EFMs) located in the region surrounding Cape Canaveral record the electrification of the atmosphere near them. Research studying how these sensors could improve lightning warnings has had mixed results. This paper used a Convolutional Recurrent Neural Network (CRNN) and data from 30 EFMs from May-July of 2012-2016. The mean was calculated for every 60 second period and 30 minutes of this summarized data was used to create a lightning prediction with a warning period of 15 minutes. This method achieved a True Positive Rate (TPR) of 77.6%, a False Positive Rate (FPR) of 8.3%, a False Discovery Rate (FDR) of 48.1%, and an Operational Utility Index (OUI) of 53.9% (Kehrer et al., 2006). This suggests that the EFM sensor array, when used as a means to measure the electrification of the entire region, is capable of effectively predicting lightning for a 5-mile radius near Cape Canaveral. Moreover, achieving a 53.9% on the OUI rivals the best methods currently used implying that incorporating EFMs into lightning forecasting may reduced the FPR and save millions of dollars in delay and cancellation costs.

15. SUBJECT TERMS

Artificial Neural Networks, Electric Field Mills, Cape Canaveral, Lightning Forecasting, Convolutional Recurrent Neural Networks

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col R. Seymour, PhD, AFIT/ENC
U	U	U	U	50	19b. TELEPHONE NUMBER *(include area code)* 937-255-6565 , x4398; daniel.hill@afit.edu