2018 25th International Lightning Detection Conference &
7th International Lightning Meteorology Conference
March 12 - 15 | Ft. Lauderdale, Florida, USA

Evaluating the Efficacy of Electric Field Mills to Predict Lightning Events near Cape Canaveral Using Convolutional Recurrent Neural Networks

Daniel E. Hill, Ryan G. Hefron, and Richard S. Seymour
Air Force Institute of Technology, Wright-Patterson AFB, United States
Daniel.Hill@afit.edu, Ryan.Hefron@afit.edu, Richard.Seymour@afit.edu

***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 can improve lightning warnings has had mixed results. This paper used a convolutional recurrent neural network to predict lightning events near Cape Canaveral. Our method summarized the EFM data across a 60 second time window and then used 30 minutes of this summarized data to predict lightning after a 15-minute warning period. The best dataset achieved a total accuracy of 90.3% on a test dataset with a true positive rate of 77.6%, probability of false detection rate of 8.3%, and an Operational Utility Index (OUI) of 53.9%. This EFM-only performance is comparable to past methods which used additional predictors and shows potential for lightning prediction using the EFM sensor array in future research.**

Keywords—Artificial neural network; Convolutional recurrent neural network; Electric field mill; Cape Canaveral; Lightning; Lightning prediction; Machine learning

I. Introduction

The area surrounding the 45th Weather Squadron (45 WS) and Cape Canaveral Air Force Station experiences one of the highest rates of lightning events in the United States between May and September. This area also houses many of the space flight operations for US military, government, and industry. In a work by Finn et al. [2010], it was estimated that $20+ billion in facilities, multi-billion dollar boosters/payloads, and 25,000+ personnel are threatened by lightning in this region. Incidents such as the destruction of the Atlas-Centaur rocket in 1987 caused by a lightning strike illustrate this threat [Christian et al., 1989]. After this event, the Lightning Advisory Panel was established and launch criteria was implemented to reduce the risk of weather related accidents. Additionally, a sensor array of 31 Electric Field Mills (EFMs) was built around the Cape Canaveral area to improve understanding of the electrification of the atmosphere near launch sites [Canright, 2001].

This paper sought to improve upon 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 30 EFMs in the Cape Canaveral region the mean was calculated using a statistics windows (SW) of 60 seconds. A measurement window (MW) of 30 minutes was used to create running windows of the means. 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 an 8.04 km (5 mile) radius around Cape Canaveral, known as the area of concern (AOC). A Convolutional Recurrent Neural Network (CRNN) was trained and validated on a portion of the dataset and the results were presented in the context of previous research. Our results show that models built using only EFM data can produce excellent short-term forecasts for the AOC when a model incorporates spatial and temporal context of the EFM data.

II. Background

Many studies have investigated if using EFMs can improve lightning prediction, but results have been inconclusive. To discuss past research, evaluation metrics must be clearly defined. False alarm rate and false alarm ratio are often used interchangeably which has lead to a great deal of confusion [Barnes et al., 2009]. For this reason, this paper will use the terms Probability of False Detection (POFD) for the false positive rate and Probability of False Alarm (POFA) for the false positive ratio. The calculations of these metrics are given in (2) and (3).

Additionally, this paper will attempt to objectively compare past research into lightning prediction with the understanding that each of the past methods differ significantly in their methodologies. A confusion matrix shown in Table 1 is used

TABLE I. CONFUSION MATRIX EXAMPLE

		Actual	
		Lightning	*No Lightning*
Predicted	*Lightning*	*a*	*b*
	No Lightning	*c*	*d*

where *a* is the number of true positives, *b* is the number of false positives, *c* is the number of false negatives, and *d* is the number of true negatives.

The True Positive Rate (TPR), called in many papers the probability of detection, is the ratio of lightning events correctly predicted to total lightning events (1); the POFD, also false positive rate or false alarm rate, is the probability of non-lightning events being predicted as lightning events (2). The POFA, also false positive ratio or false alarm ratio, is the ratio of falsely predicted lightning events to the total number of lighting predictions made (3).

$$TPR = \frac{a}{a+c} \tag{1}$$

$$POFD = \frac{b}{b+d} \tag{2}$$

$$POFA = \frac{b}{a+b} \tag{3}$$

The Operational Utility Index (OUI) (6) [D'Arcangelo, 2000; Kehrer et al., 2006] is a nonstandard statistic used by the 45 WS to weight the importance of TPR, the True Skill Statistic (KSS) (7) [Hanssen & Kuipers, 1965; Woodcock, 1976], and the POFA to their purposes. It is a metric that has a maximum of 0.83 that most heavily weights the detection of lightning, which results in the fewest number of false negatives. This reduces the likelihood of the 45 WS predicting a nonlightning event when one actually occurs, placing personnel and assets at the most risk.

It is important to note the difference between false alarm rate and false alarm ratio in the OUI calculation. In the work by D'Arcangelo [2000], the d-index, later called the OUI, is calculated using the POFA, but is referred to as the false alarm rate. The OUI calculation presented in the work by Kehrer et al. [2006] also uses the term false alarm rate, but again uses the calculation for POFA presented in (3). For this reason, this paper explicitly uses the terms POFA and POFD to clearly distinguish between the two calculations in (2) and (3).

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

$$TSS = TPR - POFD = \frac{a}{a+c} - \frac{b}{b+d} \tag{7}$$

A large corpus of techniques to forecast lightning strikes have been developed by researchers, and several models have included EFM data as a predictor variable. Here we examine recent efforts that incorporate EFM data into their predictive model. The evaluation metrics for these research methods are presented in Table 2 (the work by Maier and Huddleston [2017] is omitted due to missing metrics).

In the work by Mazany et al. [2002], a logistic regression model was created using four predictor variable: the 30-minute maximum value across all 31 EFMs (read every 5 minutes), Global Positioning System (GPS) Integrated Precipitable Water Vapor (IPWV), the change of GPS IPWV over a 9-hour period, and the *K* Index (KI). Data was collected May-September of 1999. The purpose of the research was to improve the forecasting skill and increase the lead time of lightning forecasting for the Kennedy Space Center (KSC). Their model calculated the probability of a lightning event and used a threshold of 0.7 as the determining value for a lightning prediction to be made. If the model produced a value below 0.7, then a lightning event was predicted to occur in a radius of 37.04 km (20 nautical miles) centered as Cape Canaveral. The method used a long-term prediction window (12.5h) that provided indication of a lightning event for the entire day. The long prediction window and single season of data resulted in a limited number of observed lightning events in this study.

This long prediction window differs significantly from the 15-minute prediction window used in this paper, thus it is important to not directly compare metrics without considering the differing intents. They evaluated their methods by counting the number of times a lightning event occurred/did not occur after 90 minutes of their model producing a value less than 0.7. They presented a confusion matrix where a true positive was a lightning event occurring more than 90 minutes after their model prediction fell below 0.7 and a false negative being when their model fell below 0.7 and no lightning event occurred during that day. The results are presented in Table 2. They reported that the EFM maximum value provided little benefit to the long-term (≥90 minute) predictability of lightning events; however, EFM maximum value did appear to change significantly immediately prior to a lightning event [Mazany et al., 2002]. This change partly motivated our examination of short-term lightning forecasting using only EFMs.

The work by Kehrer et al. [2006] optimized the work by Mazany et al. [2002] for 2 hour and 9 hour forecasting intervals. They used a logistic regression model with four predictor variables gathered May-September of 2000-2003: 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. EFM data was excluded due to its low p-value during a forward and backward feature selection. This indicates that EFM data was not significant as a predictor in a logistic regression model at the specified forecasting intervals.

For the 2-hour forecasting interval, their intent was to create a model that provided a 30-minute warning prior to lightning activity for a 9.26 km (5 nautical mile) radius around Cape Canaveral. The other 90 minutes of the 2-hour forecast interval were a processing and communication lag time to mimic real-world delays between data gathering and warning implementation. This closely resembles the intent of this paper and so the 2-hour forecast results are presented and the 9-hour

results are omitted. The comparison in Table 2 corresponds to these results. Kehrer et al. [2006] discuss the potential for improvements to lightning forecasting by using nonlinear techniques, such as neural networks, and by adding predictors, such as EFM data.

In the work by Murphy et al. [2008], the researchers sought to understand the contribution of EFMs to lightning warning systems. They analyzed the 10 and 60 second means of 2 EFMs from June-August of 2004-2005 to predict lightning within a square created by extending in the cardinal directions 10 and 20 km centered between the two EFMs. Three conditions were established that triggered a warning, two of which included a 1 kV/m or 2 kV/m threshold for the mean of the EFMs. Once a condition was met, a warning was issued.

If a lightning event occurred in the AOC after this point, it was considered a true positive and the time between the prediction and the event was recorded as the lead time. A false positive was defined as a warning criterion being met but no subsequent lightning event occurring (no timeframe was given in which the event must occur). The researchers expressed that the addition of the EFM data to the warning criteria resulted in worse performance than if they were excluded. They hypothesized that the effective range of the EFMs was ineffective at detecting the electrification of storms whose center is well above sea level. They suggest their work may be improved by changing the size of the AOC or changing the orientation of the AOC to the EFMs [Murphy et al., 2008].

The work by Da Silva Ferro et al. [2011] used a single EFM in Southeastern Brazil to determine criteria to provide lightning warnings. Circular AOCs of 5, 10, and 15 km and annular AOCs of 0 to 5 km, 5 to 10 km, and 10 to 15 km were centered around the EFM. Multiple EFM thresholds were chosen to use a criterion for warnings; that is, when the EFM exceeded the given threshold, a warning was issued. A true positive was recorded when a lighting event occurred within the 45-minute warning and a false positive if one was predicted but did not occur. The circular AOCs were most pertinent to this analysis so only the best metrics presented for the circular AOCs and for a threshold of 0.9 kV/m are presented in Table 2.

The wok by Da Silva Ferro et al. [2011] heavily emphasized the importance of altitude of the EFM in its ability to detect and predict lightning, which supports the conclusions mentioned above in the work of Murphy et al. [2008]. Due to the elevation of the EFM (800 meters above sea level) in their study, they saw improved TPR and POFA using one EFM with a 10 km AOC than other similar studies with small AOCs and more EFMs. With this in mind, they suggest using a network of EFMs and an adjustment of the AOC to account for the improved metrics attained by the altitude of their sensor [Da Silva Ferro et al., [2011].

A more recent work by Maier and Huddleston [2017] analyzed multiple EFMs in the Cape Canaveral region and looked at the range of the EFM readings over 3 minute running windows known as envelopes. Their intent was to detect initial cloud electrification and the onset of lightning. Six stormy days of data in the summer of 2014 were used to predict lightning within 9.26 km (5 nautical miles) of Cape Canaveral. A warning was issued when the range exceeded a given threshold of 40, 60, or 80 V/m. This resulted in a POFD of 4.7% (6/128), but no other metrics or numbers were reported.

III. DATA

Data from May-June of 2012-16 was collected from the Kennedy Space Center (KSC) website (https://kscwxarchive.ksc.nasa.gov/). Fig. 1 shows the location of the 31 EFMs in the Cape Canaveral region. The EFM readings led to a dataset consisting of a date, time, and volts per meter (V/m) measured at 50 hertz. The Lightning Detection and Ranging data was provided by the 45 WS and included only pertinent lightning events from 2013-2016. The LDAR data had features date, time, and x, y, and z location of the lightning events.

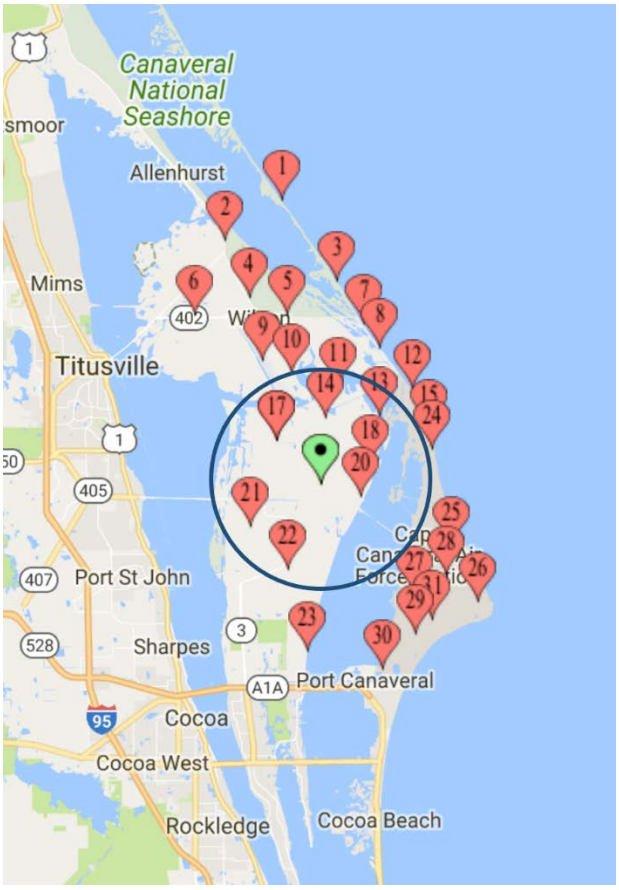

Fig. 1. Map depicting the location of the 31 EFMs supporting the 45 WS and the Cape Canaveral region. The Green identifier is the location of the center of the LDAR array and the center of the AOC selected for this study [Google Maps, 2018].

The data arrived as 30 minutes of readings for each EFM, but was consolidated into 30 minutes of data that contained all EFM readings. Data from only 30 EFMs were useable due to integrity problems associated with the excluded site. If an EFM did not have a reading for a given time, a 'NaN' value was recorded. This ensured no gaps in time existed in the data, since

temporal continuity is important when processing data using convolutional and recurrent layers later in the analysis.

A statistics window (SW) of 60 seconds was chosen. The data were separated into contiguous, non-overlapping 60 second chunks and the mean was calculated for each chunk. The mean values at each of the 30 EFMs were the features for a given observation. The data was then consolidated into days.

A min-max scaling was applied across all the data using (6) and (7). This scaled the mean calculations to values between -1 and 1 to improve network training. Equations (6) and (7) calculate the min-max scaled data $x_{i_{scaled}}$ where M and m are the maximum and minimum values the scaled data can obtain (1 and -1 were chosen), respectively. The min(x) and max(x) are the minimum and maximum values obtained across the entire timeframe by the EFM from which the x_i was drawn.

$$x_{i_s} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{6}$$

$$x_{i_{scaled}} = x_{i_s} * (M - m) + m \tag{7}$$

The missing data replaced with 'NaN' values in the original data were ignored to calculate the mean when possible. An additional variable was added called the missing data indicator if insufficient data was available to calculate the mean. When this occurred, then the value for the statistic was set to 0 and the missing data indicator was flipped for that EFM. Each EFM had a missing-data indicator variable, resulting in an additional 30 features per observation.

A measurement window (MW) of 30 minutes was chosen. The MW was used to expand the dataset by taking a rolling window of SW observations that covered a time period of MW. This resulted in a 3-dimensional dataset: number of MW sequences in a day, sequence of length MW (30), and 60 features (mean and missing-data indicator for each EFM location for a given SW). These sets were expanded into 4-dimensional datasets where the third dimension corresponded to each EFM location and the fourth dimension was the mean and the missing data indicator for a given location. This provided a way to consider spatial and temporal context to each observation processed by the convolutional recurrent neural network.

The LDAR dataset was then used to determine if lightning events occurred in the area of concern (AOC) during each SW. An AOC distance of 8.04 km (5 mile) radius around the center of the LDAR array near Cape Canaveral was selected. If the LDAR dataset indicated a lightning event during a SW, then a response of '1' was given; a '0' value otherwise. Based on this data, final target values were created by determining if a lightning event occurred within a 15-minute window, 15 minutes after the last SW in the MW. That is, a warning window (WW) and prediction window (PW) were both chosen to be 15 minutes. The neural network attempted to predict if lightning would occur within the PW. Fig. 2 depicts this setup. A '1' response was given if a lightning event occurred within the AOC during the PW; '0' otherwise.

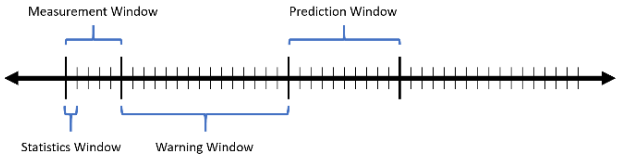

Fig. 2. Diagram showing the SW, MW, WW, and PW in a timeline. A MW is comprised of SWs and is used to predict an event within a PW after a WW.

IV. Methodology

A. Data Augmentation/Reduction

While lightning is extremely common in the central Florida region, the number of recorded lightning events compared to nonlightning events in the dataset is small, approximately 10%. For this reason, the training data were augmented to equalize the number of lightning and nonlightning events by randomly oversampling the lightning events. The oversampling method described in the work by Lemátre et al. [2017] was used to randomly oversample.

The oversampling required at least one lightning event in each dataset. Due to neural network processing requirements and the structure of the data, this meant each day in the analysis must contain at least one lightning event. Any day that did not contain a lightning event was removed from the analysis.

As mentioned earlier, if a statistic could not be calculated due to missing data the missing-data indicator was flipped. If a dataset contained more than one-hour total of SWs that had flagged missing-data indicators for all 30 EFMs, then that day was deemed unusable and removed from the analysis.

After these cleaning actions 138 days of data remained. Each day was randomly placed into a training, validation, or test dataset with a probability of 70% training, 15% validation, and 15% test. By splitting the data by day, we ensured complete separation of the training, validation, and test sets with no overlap of the rolling windows.

B. Model Selection and Hyperparameter Optimization

Application of a CRNN to this problem domain was modeled after work done by Hefron et al. [2018], who showed that CRNNs significantly outperformed other neural network approaches to analyze electroencephalographic data. While their application to workload estimation differed significantly from lightning prediction, there were many parallels in their approach. Their data consisted of a time-sequence of voltage readings from 64 electrodes located at known locations on the surface of a participant's scalp. Our data consists of time sequences of V/m readings from sensors with known locations. Since both datasets were time-series data from a sensor array, a convolutional recurrent network fit well.

While several models were examined during model building, our final model was comprised of 3 convolution layers and 3 recurrent layers. The first two convolution layers learned features that were independent of each EFM location across time. It was hypothesized that useful temporal features associated with atmospheric electrification should be similar at each EFM location, but that patterns of correlation across the EFM sensor array leading to lightning events at the AOC may

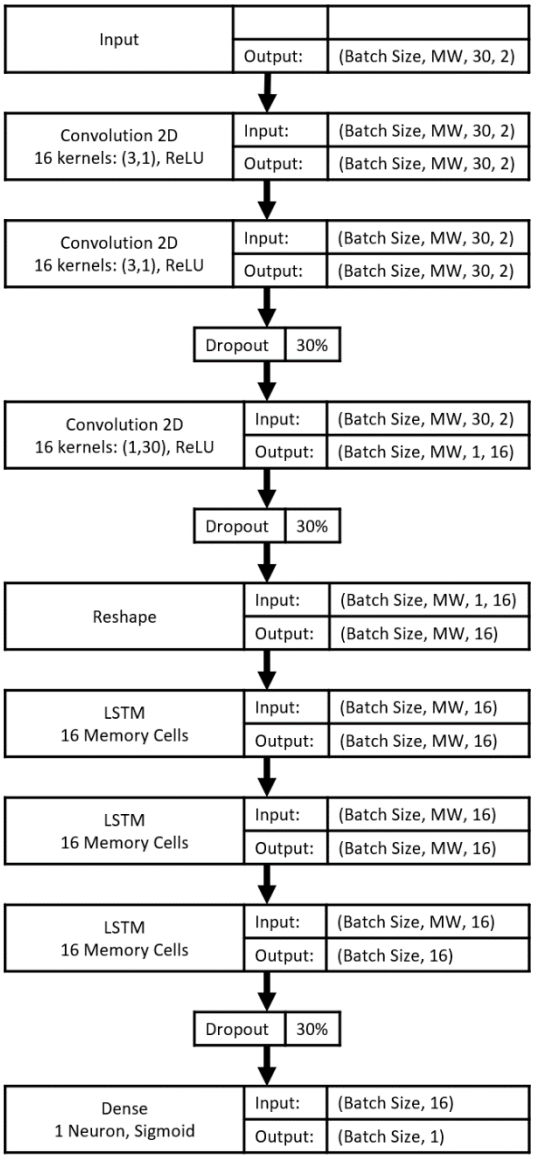

Fig. 3. This diagram depicts the structure of the neural network used in this analysis. It contains two convolutional layers capturing the features seen across time at each EFM, one convolutional layer capturing the correlation across all EFMs, and three LSTM layers capturing patterns across time of the feature maps from the convolution layers.

differ depending on weather patterns. The first two convolutional layers learn useful atmospheric electrification characteristics. Sixteen kernels of size (3,1) striding by (1,1) with same padding was used with a Rectified Linear Unit (ReLU) activation function. To decrease overfitting, a dropout layer with 30% dropout was added after these two convolution layers.

The third convolution layer captured the correlation across the entire region of the change seen at each EFM location. By using a kernel of (1,30), this layer learned the pattern of features seen across EFMs prior to a lightning event in the AOC and acted as a spatial filter. This was thought of as the 'state' of the regional change of the EFMs. A stride of 1 with valid padding

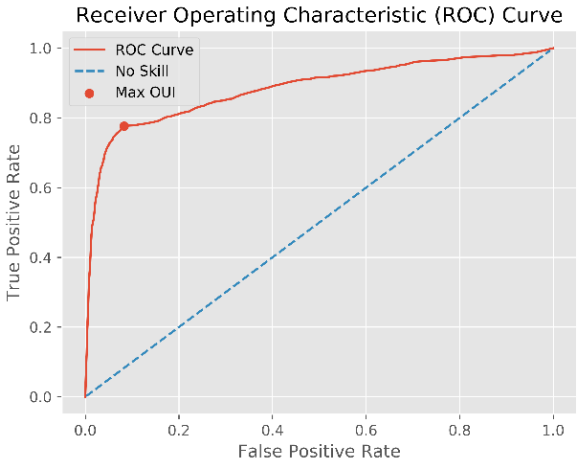

Fig. 4. A ROC curve shows how the TPR and the FPR change as the threshold used to determine classification changes. This figure shows the point along the curve at which the OUI is maximized.

was used with a ReLU activation. Another dropout layer with 30% dropout was added after this layer.

The next three recurrent layers captured the temporal patterns in the temporally and spatially filtered data. The input to these layers was a sequence of activation maps of the state of the region. These layers used 16 Long Short Term Memory (LSTM) units per layer to learn the change of the state of the region leading up to a lightning event in the AOC. An additional dropout layer was added after these layers with a 30% dropout rate.

The last layer was the output layer and consisted of a single densely connected neuron with a sigmoid activation function. A diagram of the model architecture is displayed in Fig. 3. The Adam optimizer [Kingma and Ba, 2014]with an initial learning rate of 0.01 was paired with a binary cross-entropy loss function. The learning rate was decayed by a factor of 0.5 when validation loss failed to improve over 5 epochs. Early stopping was implemented with a patience of 50 monitoring validation loss. The model was trained using the training dataset and the model that achieved the best validation loss was kept. This model was used to evaluate performance using the test dataset.

V. RESULTS AND DISCUSSION

Using a Receiver Operating Characteristic (ROC) curve shown in Fig. 4, the classification threshold that resulted in the maximum OUI was found. The ROC curve depicts the trade-off between TPR and POFD as the classification-determining threshold changes. Since OUI is the measure specified by the 45 WS as being the best indicator of a quality predictor, this paper focused on the threshold that optimized this statistic.

The optimal OUI was achieved at a threshold of 0.547. Using this threshold, a confusion matrix was calculated and is shown in Fig. 5. This threshold yielded a TPR of 77.6%, a POFD of 8.3%, a POFA of 48.1%, and an OUI of 53.9%. The results are compared to the previous research in Table 2. When comparing the results of this research with the research cited in the introduction, it should be noted that these metrics compare

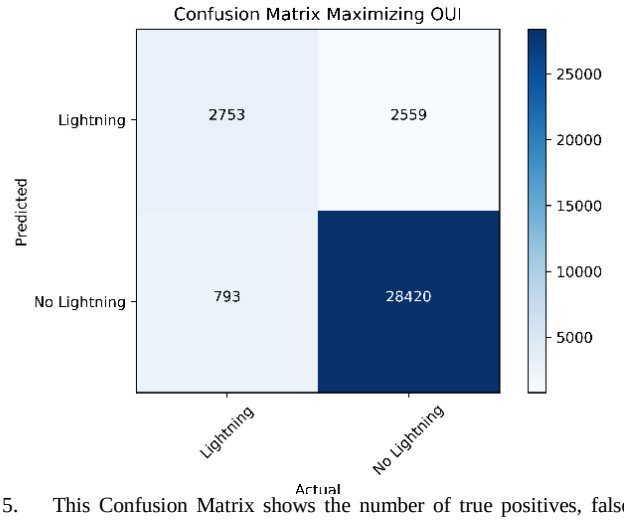

Fig. 5. This Confusion Matrix shows the number of true positives, false positives, false negatives, and true negatives (moving left to right, top to bottom) achieved on a test set. These values were achieved by finding the threshold that achieved the highest OUI.

research with different intents, methods, and parameters. A conclusion that one method is better than another cannot be made; rather Table 2 highlights a number of lightning prediction methods' evaluation metrics.

Our method achieved average performance in terms of OUI and TPR compared to other works, but achieved a much lower POFD than all other methods (excluding the work by Maier and Huddleston [2017]) indicating this method may be the best method to ensure false predictions are not made. Because our method was focused on short-term prediction of lightning events, many more prediction windows were available compared to other work. This allowed us to explore fitting a deep neural network to the data rather than relying on more traditional models such as logistic regression which tend to better handle a smaller number of observations. However, using more observations to forecast short-term warnings highlighted a challenge not present in the other works we compare our results to—our dataset was highly biased towards nonevent conditions. Because of the longer prediction windows used in other studies, their data sources were far more balanced than ours. This made model fitting more complex and challenging than using a balanced dataset.

Despite the challenges, our results may provide utility through augmentation of long-term forecasts by reducing the likelihood of false detections for short-term periods that overlap with the long-term prediction windows. In this way, if a long-term forecast predicts a lightning event with a lengthy prediction window, our method in combination with others, may improve fidelity of when an event will occur.

A shortcoming of our methodology which we intend to improve in the future is that our predictions are not in reference to the first lightning event associated with a storm. Rather, our method continues to make predications even after the onset of a storm. Adjusting our data processing pipeline for this factor and extending warning windows are left for future work.

TABLE 2. RESULTS COMPARISON

	TPR	POFD	POFA	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][a]	37.7%	N/A	71.0%	N/A
[Da Silva Ferro et al., 2011][a]	60%	N/A	41%	N/A
This Paper	77.6%	8.3%	48.1%	53.9%

[a.] The works by Murphy et al. [2008] and Da Silva Ferro et al. [2011] do not use a method that results in the calculation of true negative events. This makes it impossible to calculate POFD and OUI.

VI. Conclusion

With these results, it is concluded that a short-term predictive relationship exists between the EFM sensor array and lightning events in the Cape Canaveral area. This paper shows that EFMs alone are capable of predicting lightning for the region around Cape Canaveral at a success rate comparable to past methods as determined by OUI, yet with significantly differing warning and prediction windows. Future work should seek to improve these results by including data from the remainder of the stormy season (May-September). Additionally, determining the optimal time periods for MW and SW or including higher order statistics, time of day, and month of year as additional features may improve performance. Finally, existing methods can be fused with ours to improve the overall model.

Acknowledgment

The authors would like to thank the following individuals:

- Mr. William Roeder and the 45th Weather Squadron for the opportunity to work on this problem and the assistance given.
- Dr. Lisa Huddleston and her support team for their assistance in retrieving the data required to complete this study.
- Mr. Marcin Owczarczyk for writing a program to automate the downloading of data from the ksc.gov website; the study would not have been possible without this program.
- Dr. Brett Borghetti for his guidance and instruction in deep learning.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the United States Air Force or the U.S. Government.

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.

Canright, S. (2001). NASA - Lightning and Launches. Retrieved July 15, 2017, from https://www.nasa.gov/audience/foreducators/9-12/features/F_Lightning_and_Launches_9_12.html

Christian, H. J., Mazur, V., Fisher, B. D., Ruhnke, L. H., Crouch, K., & Perala, R. P. (1989). The Atlas/Centaur lightning strike incident. *Journal of Geophysical Research*, *94*(D11), 13169.

D’Arcangelo, D. L. (2000). Forecasting the onset of cloud-ground lightning using layered vertically integrated liquid water, *298*(704), 1–71.

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.

Finn, F. C., Roeder, W. P., Buchanan, M. D., McNamara, T. M., McAllenan, M., Winters, K. A., … Huddleston, L. L. (2010). *Lightning Reporting at 45th Weather Squadron: Recent Improvements*.

Google Maps. (2018). *Cape Canaveral Region with Electric Field Mill Site Locations.* Retrieved from https://www.mapcustomizer.com/map/EFMs%20and%20LDAR%20

Hanssen, A. W., & Kuipers, W. J. A. (1965). On the Relationship Between the Frequency of Rain and Various Meteorological Parameters. *Koninklijk Nederlands Meteorologisch Instituut Mededelingen En Verhandelingen*, (81).

Kehrer, K. C., Graf, B., & Roeder, W. (2006). *Global Positioning System (GPS) Precipitable Water in Forecasting Lightning at Spaceport Canaveral*. Cocoa Beach, FL.

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization.

Lemátre, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, *18*(17), 1–5.

Maier, L. M., & Huddleston, L. L. (2017). Automated Identification of Initial Storm Electrification and End-of-Storm Electrification Using Electric Field Mill Sensors. In *American Meteorological Society Annual Meeting*. Seattle, WA.

Mazany, R. A., Businger, S., Gutman, S. I., & Roeder, W. (2002a). A Lightning Prediction Index that Utilizes GPS Integrated Precipitable Water Vapor*. *Weather and Forecasting*, *17*(5), 1034–1047.

Mazany, R. A., Businger, S., Gutman, S. I., & Roeder, W. (2002b). A Lightning Prediction Index that Utilizes GPS Integrated Precipitable Water Vapor*. *Weather and Forecasting*, *17*(5), 1034–1047.

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.*

Woodcock, F. (1976). The Evaluation of Yes/No Forecasts for Scientific and Administrative Purposes. *Monthly Weather Review*.