Severe Weather Prediction Using Lightning Data

Iván Venzor-Cárdenas
Universidad Autónoma de Nuevo León
ivenzor@gmail.com

Nadia Ahmed
University of California, Irvine
ahmedn@uci.edu

Maria J. Molina
National Center for Atmospheric Research
molina@ucar.edu

Marek Slipski
Jet Propulsion Laboratory
marek.j.slipski@jpl.nasa.gov

Mark Cheung
Lockheed Martin

Clemens Tillier
Lockheed Martin

Samantha Edgington
Lockheed Martin

Gregory Renard
Augustus Intelligence

Abstract

Increases in flash rates detected in ground-based lightning data can be a precursor to severe weather hazards [12, 6, 9]. Lightning data from the Geostationary Lightning Mapper (GLM) aboard the GOES-16 satellite is not currently a systematic part of an operational model used by forecasters and is underutilized in severe storm research. We harness the spatial and temporal advantages of geostationary satellite data to create a machine learning model that augments the current meteorological practices and capabilities of forecasters during extreme weather. Our results suggest that false alarms for warned thunderstorms could be decreased by half, and that tornadoes and severe hail could be correctly identified 8 out of 10 times using GLM data.

1 Motivation and Research Problem

Severe weather, defined in this work as thunderstorms that produce tornadoes and large hail (1-inch or greater in diameter), can claim lives and cause substantial economic damage [7, 13]. In recent years, societal exposure and costs associated with severe thunderstorm events has been increasing [13, 1]. The average warning lead time, which refers to the amount of time (in minutes) before the onset of severe weather, is approximately 15 minutes [2]. While the forecast skill of National Weather Service (NWS) meteorologists is high [4], decreasing the number of false alarms remains a priority and can further increase public trust in severe weather warnings.

Various studies have been recently conducted that use machine learning to improve the lead time and accuracy of severe hazard predictions. Convolutional neural networks (CNNs) have been used to create skillful one-hour lead time forecasts of large hail from convection-permitting model output [10] and of tornadoes using radar [11]. Arguably one of the most comprehensive machine learning models for the prediction of severe hazards is the empirical ProbSevere model [4], but it does not incorporate GLM information. The large geographic scope and high temporal frequency of underutilized GLM data presented an opportunity to improve upon forecasts of severe weather.

2 Technical Contribution

Using GLM data we derived and gridded 8 higher-level quantities (flash extent density, flash centroid density, average flash area, total energy, group extent density, group centroid density average group

34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

area and minimum flash area) onto 2 km x 2 km grids at a 1 minute cadence using the GLM Tools package [3]. The nearly 4 million GLM L2 files (1.5 TB) spanned 2018 - 2020.

A supervised learning approach was pursued to train machine learning models using NOAA severe weather reports as ground truth. Graph Neural Networks as well as some variants of Convolutional Neural Networks (CNN) architectures, such as U-NET for image segmentation, were explored. We also invested a considerable amount of time working in temporal-spatial architectures such as 3D-CNN and recurrent neural network architectures such as GRU-CNN. While these approaches revealed some promising initial results these architectures require further tuning, data augmentation, and training time to achieve desired results.

GLM data can also be represented as time series information. Using the 8 derived quantities from GLM gridded data, we built time series in a 60 km x 60 km area centered on the coordinates of the tornado and hail reports by summing the quantities in each pixel at each 1 minute interval. These aggregated time series events comprise our positive class. For our null class (negative class), we used severe weather warnings that were previously issued by NWS meteorologists and had no associated severe weather reports within a 24-hour window. The time series generation for the null class is identical to the positive class but centered in the centroid coordinates of the warning. These considerations framed the challenge as a time series classification problem with a specific lead time to allow nowcasting. Effectively, our model used 45 minutes of data in a certain region to assess the probability of a severe weather event occurring within the next 15 minutes.

We explored InceptionTime neural network architecture [8] for time series classification, but decided to use a recently developed method called ROCKET (RandOm Convolutional KErnel Transform). According to the authors, ROCKET [5] is orders of magnitudes faster to train than InceptionTime and offers similar skill compared to state of the art methods. ROCKET transforms the multivariate time series using a large number of random convolutional kernels (random length, dilation, padding, weights and bias) to create features, which can be used to train a classifier. A substantial difference between ROCKET and a neural network is that the weights and biases are not learned but rather randomly generated. This characteristic makes ROCKET extremely fast and simple to use, yet powerful due to the large number of convolutional kernels used, these convolutions effectively translate into features. Once the features are created, a classifier can be trained. We tested a Ridge classifier and a Gradient Boosted Decision Trees (XGBoost) classifier, but as there were little difference in the results, the Ridge classifier was chosen due to its simplicity. The skill of the model is assessed using two widely used metrics in atmospheric science, the Critical Success Index (CSI) and the False Alarm Ratio (FAR). The definition of these metrics are based on the 2x2 contingency table (confusion matrix). The CSI metric is defined as the ratio of hits to the sum of hits, misses, and false alarms (higher values are better) and the FAR metric is defined as the ratio of false alarms to the sum of false alarms and hits (lower values are better).

When compared to the state of the art [4], our results for the CSI and FAR metrics are better. However, caution is needed when comparing these metrics, as the state of the art provides greater lead time, larger geographic coverage, and a longer temporal range.

Table 1: State of the art comparison

	Our results	State of the art
Lead time	15 min	> 15 min
Coverage	Central US (1,000 km x 800 km)	CONUS
Period	Mar-Jun 2019	May-Jul 2014, Mar-Dec 2016
CSI	0.49	0.35
FAR	0.41	0.66

There are several limitations that are important to acknowledge in this study. Our machine learning labels consisted of NOAA severe thunderstorm reports, which contain issues related to population biases and reporting practices [14]. The strong cross-correlations among the input GLM variables also present challenges in interpreting the reasons for high model skill and discovery of important physical features that may be related to severe thunderstorm activity. Moving forward, these limitations may be addressed and the work may be enhanced in the following ways. Extending the model into other geographic regions and seasons and a longer lead time could increase its value to a wider population. The addition of radiance information in the time series model could also increase our model's skill.

Broader Impact

The goal of this research is intended to help forecasters. Ultimately, we aim to have our models tested at the NOAA Hazardous Weather Testbed in Norman, Oklahoma first-hand by forecasters. However there is much future work and testing needed in order to achieve this goal.

Acknowledgments and Disclosure of Funding

This research was conducted during the 2020 NASA Frontier Development Lab (FDL) program, a public/private partnership between NASA, SETI and industry partners including Lockheed Martin, IBM, Google Cloud, Intel, and NVIDIA Corporation. The authors wish to thank in particular Google Cloud for generously providing computing resources. We gratefully thank all our mentors for guidance and useful discussion.

References

[1] Ashley, W. S., & Strader, S. M. (2016) Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bulletin of the American Meteorological Society* **97**(5), 767-786.

[2] Brooks, H. E., & Correia Jr, J. (2018) Long-term performance metrics for National Weather Service tornado warnings. *Weather and Forecasting* **33**(6), 1501-1511.

[3] Bruning, E. C., Tillier, C. E., Edgington, S. F., Rudlosky, S. D., Zajic, J., Gravelle, C., ... & Schultz, C. J. (2019) Meteorological Imagery for the Geostationary Lightning Mapper. *Journal of Geophysical Research: Atmospheres* **124**(24), 14285-14309.

[4] Cintineo, J. L., Pavolonis, M. J., Sieglaff, J. M., Lindsey, D. T., Cronce, L., Gerth, J., ... & Gravelle, C. (2018) The NOAA/CIMSS ProbSevere model: Incorporation of total lightning and validation. *Weather and Forecasting* **33**(1), 331-345.

[5] Dempster, A., Petitjean, F., & Webb, G. I. (2020) ROCKET: Exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery* 1-42.

[6] Darden, C. B., Nadler, D. J., Carcione, B. C., Blakeslee, R. J., Stano, G. T., & Buechler, D. E. (2010) Utilizing total lightning information to diagnose convective trends. *Bulletin of the American Meteorological Society* **91**(2), 167-176.

[7] Doswell, C. A. (2001) Severe convective storms—An overview. *In Severe convective storms*, pp. 1-26. American Meteorological Society, Boston, MA.

[8] Fawaz, H. I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., ... & Petitjean, F. (2020) Inceptiontime: Finding alexnet for time series classification. *Data Mining and Knowledge Discovery* 1-27.

[9] Gatlin, P. N., & Goodman, S. J. (2010) A total lightning trending algorithm to identify severe thunderstorms. *Journal of atmospheric and oceanic technology* **27**(1), 3-22.

[10] Gagne II, D. J., Haupt, S. E., Nychka, D. W., & Thompson, G. (2019) Interpretable deep learning for spatial analysis of severe hailstorms. *Monthly Weather Review* **147**(8), 2827-2845.

[11] Lagerquist, R., McGovern, A., Homeyer, C. R., Gagne, D. J., & Smith, T. (2020) Deep Learning on Three-dimensional Multiscale Data for Next-hour Tornado Prediction. *Monthly Weather Review.*

[12] Schultz, C. J., Petersen, W. A., & Carey, L. D. (2011) Lightning and severe weather: A comparison between total and cloud-to-ground lightning trends. *Weather and forecasting* **26**(5), 744-755.

[13] Smith, A. B., & Matthews, J. L. (2015) Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Natural Hazards* **77**(3), 1829-1851.

[14] Verbout, S. M., Brooks, H. E., Leslie, L. M., & Schultz, D. M. (2006) Evolution of the US tornado database: 1954–2003. *Weather and Forecasting* **21**(1), 86-93.