URSI GASS 2023, Sapporo, Japan, 19 – 26 August 2023

Updates to the Earth Networks Total Lightning Network

Elizabeth A. DiGangi*(1), Yanan Zhu(1), Jeff Lapierre(1), and Michael Stock(2)
(1) AEM, Germantown, MD, USA; email: elizabeth.digangi@aem.eco
(2) Cooperative Institute for Severe and High-Impact Weather Research and Operations/NOAA/OAR/National Severe Storms Laboratory, Norman, OK USA

Abstract

The Earth Networks Total Lightning Network (ENTLN) incorporates advanced lightning location technology to deliver competitive lightning detection efficiency, location accuracy, and classification. It consists of over 1800 wideband sensors deployed worldwide to detect lightning and generate real-time localized storm alerts. ENTLN is constantly evolving through network expansion, as well as hardware and software development. In this paper, we will describe recent advances to the ENTLN hardware and processor, and we will show verification statistics for location accuracy, detection efficiency, and classification accuracy of the new processor.

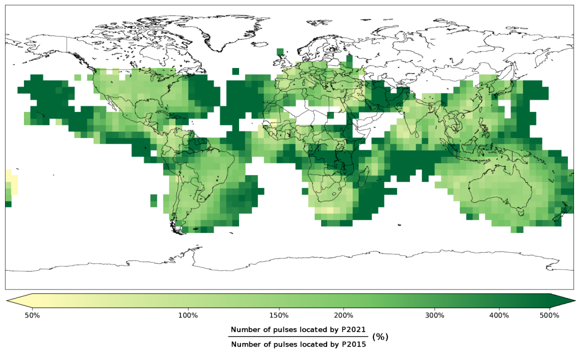

Figure 1. Global distribution of pulse detection gains for P2021 over P2015.

1 Introduction

The Earth Networks Total Lightning Network (ENTLN) consists of over 1800 sensors deployed globally that detect wideband electric field signals emitted by both intracloud (IC) and cloud-to-ground (CG) lightning. The electric field signals recorded by the sensors are continuously sent back, in compressed data packets, to the central processor, where lightning events are geolocated in real-time using the time-of-arrival (TOA) technique. The ENTLN data processor has been upgraded several times since its inception. The performance characteristics of the processor that was operational as of August 2015 (P2015) were evaluated using both natural and rocket-triggered lightning recorded in Florida [1]. In December 2021, the ENLTN launched a new processor (P2021). The upgrades made in P2021 that led to improved performance characteristics, such as a new classification algorithm, a new propagation model, and new regional data processing architecture, are described in this study. We also demonstrate the improved performance of P2021. Most of the content in this study is formally published in [2].

2 Upgrades

The new ENTLN sensor has been redesigned to use a dipole sensing element to help reduce the requirements of a strong ground. These new sensors are currently operational and produce comparable waveforms to the previous monopole antenna. New upgrades to the lightning location algorithm have increased the detection efficiency (DE), location accuracy (LA), and classification accuracy (CA) of the network. ENTLN continues to use data from the World Wide Lightning Location Network (WWLLN), enhanced via raw signals from approximately 200 ENTLN sensors, to locate lightning in extremely remote regions like the deep oceans. The rest of this section will review the upgrades made to the ENTLN processor.

2.1 Propagation Model

A lightning-produced field pulse has a longer rise time after propagating a longer distance over finitely conducting ground [3, 4]. For a lightning locating system that uses peak-finding and TOA techniques, the elongation of the rise time of field pulses causes delayed arrival times at sensors, which contributes to location errors [5]. An empirical arrival time correction has been implemented in P2021 using the rocket-triggered lightning data to improve LA [2]. The distance calculation was also updated to model a complete oblate spheroid, further improving the LA.

2.2 Detection

The processing architecture of P2021 involves splitting the world into a number of regions. Each region uses waveform data from sensors all over the world as inputs, but only looks for lightning within the bounds of the region. A grid-based location approach is implemented in an iterative manner for every second of data arriving at the processor. First, the region is split into relatively large cells, and each one of those cells searches for lightning within its bounds through an approximate (and thus quick) TOA location calculation. If a cell contains lightning, that cell is split up into smaller cells, and the search iterates with tightening constraints until initial lightning location guesses in the region are accurate and achieve a high enough spatial resolution. A better initial guess reduces the likelihood that the optimizer stage fails to converge to a solution. The new propagation model contributes to improved detection as well, especially when detecting lightning that is distant from the network. Improving the distance calculation results in better arrival time agreement of the waveform peaks at detecting sensors, thereby reducing the chance that real lightning gets rejected due to inadequate quality criteria. Regional processing also allows each region to have its own processing configuration, which provides the ENTLN the flexibility to have customized parameters in the processing algorithms for each different region. P2021 includes a sophisticated probabilistic model involving a set of quality metrics to calculate the confidence of a solution, which better discerns if a solution is real lightning or noise. The confidence and score threshold for discriminating between real and false detections are trained for each region individually, which allows maximal performance in all regions.

Table 1. Performance characteristics of P2015 and P2021 evaluated using natural lightning

Processor	**Old (P2015)**	**New (P2021)**
# flashes	219	219
# strokes	608	608
Flash DE	99%	98%
Flash CA	97%	98%
Stroke DE	96%	97%
Stroke CA	91%	94%

2.3 Classification

ENTLN classifies lightning pulses as either IC or CG. CG lightning has a channel which connects to the ground, and IC lightning does not. When ENTLN pulses are grouped into flashes, a flash will be called a CG if at least one pulse in the flash is classified as a CG; CG flashes will contain IC pulses as well, since CG lightning always has in-cloud components to it. As part of P2021, it was critical to develop a new classification algorithm using machine learning, so that the system would be less likely to include human error. The P2015 algorithm was a deterministic decision tree which classified pulses as IC or CG based on hard thresholds for various waveform parameters. Those thresholds were derived empirically from a limited ground-truth dataset from Florida (see sections 3.1-3.2). The P2021 classifier is a probabilistic algorithm which classifies segments of waveforms by calculating a classification score from the same waveform parameters. Each waveform parameter is linearized and weighted before they're all summed to get the classification score. The weighting coefficients were derived by applying a logistic regression machine learning algorithm to a verification dataset of manually classified waveforms (see section 3.3). The resulting classification score is converted to a "probability of CG" metric and retained in the final data feed for future classification evaluation efforts.

Table 2. Performance characteristics of P2015 and P2021 evaluated using rocket-triggered lightning

Processor	**Old (P2015)**	**New (P2021)**
# flashes	36	36
# strokes	175	175
# detected flashes	36	36
# correctly classified flashes	35	34
# detected strokes	169	170
# correctly classified strokes	145	150
Flash DE	100%	100%
Flash CA	97%	94%
Stroke DE	97%	97%
Stroke CA	86%	88%
Median location error (m)	215	92

3 Performance Evaluation Methods

There are three separate verification datasets used for validating different aspects of P2015 and P2021. One consists of rocket-triggered lightning information, one consists of natural lightning observed with a high-speed camera, and one is the manual classification dataset described in section 2.3. These datasets are described in sections 3.1-3.3. The overall performance characteristics of P2021 evaluated in the North Central Florida region were derived from the two ground-truth datasets described in sections 3.1 and 3.2. Because the raw waveforms from ENTLN sensors are archived indefinitely, we generated the P2021 ENTLN ground-truth evaluation dataset by running the same waveforms through the new processor. The criteria for searching for the ENTLN pulse that corresponds to a stroke in the ground-truth datasets is as follows: For every return stroke in the ground-truth datasets, we identified all the ENTLN-reported pulses (if any) in a 2 ms time window (±1 ms relative to the GPS time of the ground truth return stroke) and within 40 km of the Lightning Observatory in Gainesville (LOG) or Camp Blanding. A stroke is considered as a missed event if no ENTLN pulses were found meeting the criteria. If the detected ENTLN pulse was reported as a CG, we regard it as a correctly classified event. Otherwise, it is considered as misclassified.

3.1 Rocket-triggered lightning

The rocket-triggered lightning dataset was acquired at Camp Blanding, Florida in 2014 and 2015. Since the location of the ground termination point is known, location error can be evaluated in addition to DE and CA. All return strokes in rocket-triggered lightning are physically similar to subsequent return strokes in natural lightning [6, 7]. There are 36 flashes (175 strokes) in this dataset.

3.2 Natural lightning

All the natural CG lightning flashes in this ground-truth dataset were simultaneously recorded by electric field change meters and high-speed video cameras at the LOG in 2014 and 2015. There are 219 flashes (608 strokes) in this dataset. The time of occurrence and polarity of return strokes in natural lightning were determined by the electric field measurements. All return strokes were confirmed with cloud-to-ground channels in video recordings. However, their ground-truth locations and peak currents are unknown. The natural lightning dataset can be used to evaluate DE and CA.

3.3 Manually classified ENTLN waveforms

To produce a large enough training and testing dataset for the new classification algorithm, we developed a time-delay evaluation technique for manually evaluating raw waveform data and identifying them as CGs or ICs. This technique does not rely on looking for patterns in the shapes of the waveforms in order to avoid introducing human bias into the final dataset. The primary limitation to this technique is that a pulse can only be manually classified if it was detected by an ENTLN sensor within 15 km of the pulse. With this technique, we were able to accumulate a dataset of almost 10,000 pulses to use for training and testing the new algorithm. To verify that our technique works as expected, we did a double-blind test of it by mixing known CG waveforms from the natural lightning ground-truth data in with other waveforms. The technique proved 99% accurate by this measure. Therefore, we felt confident using it to develop a larger dataset.

4 Results

Globally, P2021 is locating approximately 50% more pulses than P2015. In moderately remote regions of the world, performance gains can be higher (see Figure 1). In this section, we will review the general performance characteristics of P2021 compared with P2015 which can be evaluated by the ground-truth datasets from Florida, including LA, DE, and CA. We will further expand on the CA improvements as evaluated from a portion of the manual verification dataset.

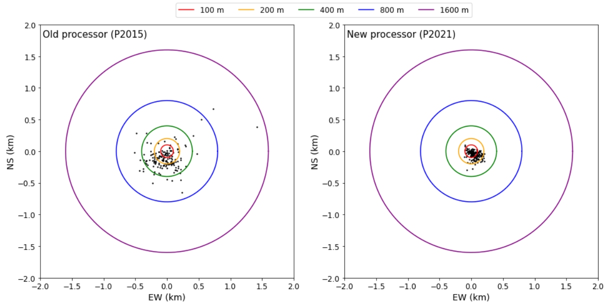

Figure 2. Location accuracy for P2015 (left) and P2021 (right) relative to rocket-triggered lightning.

4.1 Location Accuracy and Detection Efficiency

The performance characteristics of P2021 evaluated using the natural and rocket-triggered lightning datasets are summarized in Tables 1 and 2, respectively. The performance characteristics of the P2015, which were evaluated using the same ground-truth datasets, are also given in these tables for comparison. Compared to P2015, both the flash and stroke DEs of P2021 are similarly good. The LA improved significantly, with the median location error decreasing from 215 m to 92 m. The median peak current estimation error slightly improved from 15% to 13%. The locations of ENTLN pulses corresponding to return strokes in rocket-triggered lightning that are detected by P2015 and P2021 are shown in Figure 2.

4.2 Classification Accuracy

Tables 1 and 2 show the classification algorithm accuracy for P2021 and P2015 when applied to the natural and rocket-triggered lightning datasets in FL. The stroke CA increased from 91% to 94% for natural lightning and from 86% to 88% for rocket-triggered lightning. This increase may not look substantial, but considering that P2015's classification scheme was trained on the FL data and P2021's was explicitly not, that increase is significant. Figures 3 and 4 show the legacy and new classification algorithm skills, respectively, as tested on a subset of the manual verification dataset, which includes both a larger number of pulses to test against and more geographic variation in pulse location compared with the FL dataset. The CG probability of detection (POD) increased by 10 percentage points, from 77% in P2015 to 87% in P2021, while the CG false alarm rate (FAR) only increased by less than half a percentage point. An increase in FAR is expected with an increase in POD, but we were able to achieve a significant POD increase with a relatively tiny FAR increase.

5 Summary

The upgraded processor of the ENTLN was deployed in December 2021. The performance characteristics of the new processor (P2021) were evaluated using the same

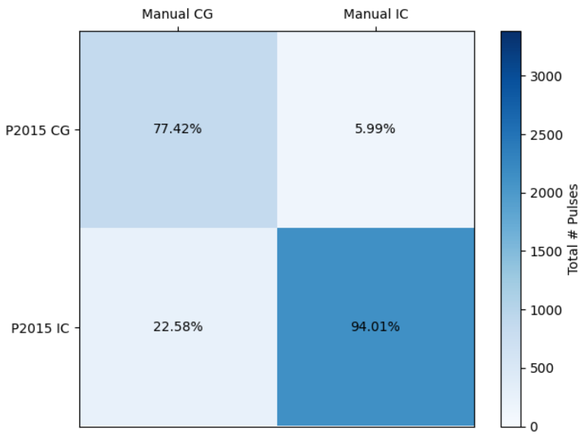

Figure 3. Confusion matrix depicting the classification skill (in terms of POD and FAR) of P2015.

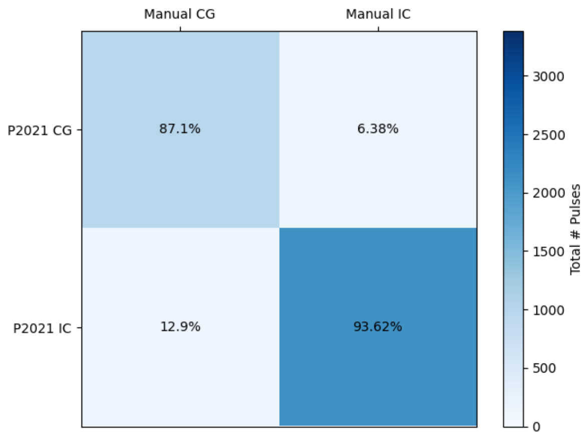

Figure 4. Confusion matrix depicting the classification skill (in terms of POD and FAR) of P2021.

ground-truth natural and rocket-triggered data acquired in Florida that was used to evaluate P2015 previously [1]. The stroke DEs of P2021 and P2015 are similarly good (97%). The LA of P2021 (92 m) improved significantly compared with that of P2015 (215 m). The improved LA is due to the new propagation and distance model, which corrects arrival time delays and calculates a more accurate ground wave propagation path. The decision tree classification algorithm in P2015 was replaced by a probabilistic classification algorithm in P2021. This algorithm was developed using a verification dataset of manually-classified waveforms. Testing both algorithms on a subset of the manual verification dataset showed that P2021 performed significantly better than P2015 on correctly classifying CGs, increasing that POD from 77% to 87% while the corresponding CG FAR remained very low (6%). The number of ENTLN pulses reported by P2021 all over the world is about 50% higher than that of P2015. The detection gains are modest in regions of the world with a fairly dense network of sensors, while the gains increase significantly in regions where sensor density is lower. The new regional processing architecture, new propagation model, and a sophisticated probabilistic source/noise discrimination model of P2021 contribute to these detection gains.

Acknowledgements

The authors thank faculty and previous student members of the Lightning Lab of University of Florida, Vladimir Rakov, Manh Tran, Douglas Jordan, Martin Uman, Jaime Caicedo, Daniel Kotovsky, Robert Wilkes, Felipe Carvalho, Terry Ngin, William Gamerota, John Pilkey, and Brian Hare, for their contributions to the collection of ground-truth data used in this study.

References

[1] Y. Zhu et al., “Evaluation of ENTLN Performance Characteristics Based on the Ground Truth Natural and Rocket-Triggered Lightning Data Acquired in Florida,” *J. Geophys. Res. Atmos.*, **122**, 8, Sep. 2017, pp. 9858–9866, doi: https://doi.org/10.1002/2017JD027270.

[2] Y. Zhu, M. Stock, J. Lapierre, and E. DiGangi, “Upgrades of the Earth Networks Total Lightning Network in 2021,” *Remote Sensing*, **14**, 9, 2022, doi: 10.3390/rs14092209.

[3] V. Cooray and S. Lundquist, “Effects of propagation on the rise times and the initial peaks of radiation fields from return strokes,” *Radio Sci.*, **18**, 3, May 1983, pp. 409–415, doi: 10.1029/RS018i003p00409.

[4] M. A. Uman, C. E. Swanberg, J. A. Tiller, Y. T. Lin, and E. P. Krider, “Effects of 200 km propagation on Florida lightning return stroke electric fields,” *Radio Sci.*, **11**, 12, Dec. 1976, pp. 985–990, doi: 10.1029/RS011i012p00985.

[5] Y. Zhu, W. Lyu, J. Cramer, V. Rakov, P. Bitzer, and Z. Ding, “Analysis of Location Errors of the U.S. National Lightning Detection Network Using Lightning Strikes to Towers,” *J. Geophys. Res. Atmos.*, **125**, 9, May 2020, doi: https://doi.org/10.1029/2020JD032530.

[6] Z. Ding, S. Chen, V. A. Rakov, I. Kereszy, and Y. Zhu, “Comparison of Far Electric Field Waveforms Produced by Rocket-Triggered Lightning Strokes and Subsequent Strokes in Natural Lightning,” in *Proceedings of the 35th International Conference on Lightning Protection (ICLP)*, 2021, **1**, pp. 1–5, doi: 10.1109/ICLPandSIPDA54065.2021.9627431.

[7] J. Wang et al., “Multiple-station measurements of a return-stroke electric field from rocket-triggered lightning at distances of 68-126 km,” *IEEE Trans. Electromagn. Compat.*, **61**, 2, 2019, pp. 440–448, doi: 10.1109/TEMC.2018.2821193.