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

Detecting Narrow Bipolar Events on a Global Scale with Machine Learning

Yanan Zhu* [(1)], Jeff Lapierre[(1)], Elizabeth DiGangi[(1)], Michael Stock[(2)], Si Chen[(3)], Vladimir Rakov[(3)], Ziqin Ding[(3)], Yunjiao Pu[(4)], and Steven Cummer[(4)]

(1) Advanced Environmental Monitoring/Earth Networks, Germantown, MD, USA; e-mail: yanan.zhu@aem.eco
(2) Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, OK, USA
(3) Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA
(4) Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA

Abstract

A machine-learning classification algorithm was developed to classify compact intracloud discharges (CIDs), also known as narrow bipolar events (NBEs), based on electric field waveforms recorded by Earth Networks Total Lightning Network (ENTLN) sensors. The classification accuracy of the model is about 95%, which is consistent with the accuracy additionally determined based on two independent validation datasets. From February to November 2022, a total of 10.4 million NBEs were detected by the ENTLN on a global scale, which account for 0.8% of all classified lightning events. The percent of lightning events that are NBEs is 1.1% in tropics, which is about double the 0.58% in mid-latitudes. We also found the strong latitude dependence of NBE percentage in the CONUS and Australia. The factors that cause the tropic/mid-latitude contrast as well as the latitude dependence remain to be investigated.

1 Introduction

Compact intracloud discharges known as narrow bipolar events (NBEs) are electrical discharges inside the cloud that are characterized by a relatively narrow (a few tens of microseconds) bipolar field pulse in low frequency (LF) band (e.g., [1]), powerful radiation in high frequency (HF) and very high frequency (VHF) bands (e.g., [2]), and inferred short (less than 1 km) channel length (e.g., [3]). Strong radio signals of NBEs have been routinely detected from a long distance by both ground and space-borne instruments (e.g., [4], [5]).

NBEs usually occur in isolation or at the beginning of a regular lightning flash [6], suggesting that some of them are closely tied to the lightning initiation processes (e.g., [7], [8]). Recently, based on the measurements of the Atmosphere-Space Interactions Monitor (ASIM), NBEs were found to be associated with blue corona discharges close to or above the top of thunderclouds (e.g., [9], [10]).

2 Earth Networks Total Lightning Network

The ENTLN consists of over 1,800 sensors deployed all over the world that detect wideband (1 Hz to 12 MHz) electric field signals emitted by lightning [11]. The electric field signals recorded by the sensors are non-linearly decimated, and then continuously sent back to the central processor, where the geolocation of lightning is implemented in real-time using the time-of-arrival (TOA) technique. On average, the ENTLN reports about 50 lightning events per second worldwide [12]. For each lightning event, time of occurrence, geolocation, event type, and peak current estimate are reported.

3 NBE Classification

The core of the NBE classification procedure is similar to the one used by Cordoba Marx Meter Array to classify cloud-to-ground and intra-cloud lightning [13]. Key aspects of the classification algorithm are described as follows.

3.1 Classification model

The classification model for the NBE classification is based on Support Vector Machines (SVMs). SVMs are a set of versatile and powerful supervised machine-learning models that can be used for linear or nonlinear classification. A nonlinear kernel was chosen for the NBE classification model. The core concept of SVMs is that its optimization objective is to maximize the distance between the decision boundary and data points that are closest to the decision boundary. The output of the model is binary (NBE vs. non-NBE).

3.2 Distance constraint

Only pulses from stations contributing to the geolocation of the lightning event and also within 600 km of the event were used for classification. The 600-km distance constraint was set because we found that distant NBE field pulses tend to be wider (due to the propagation over lossy ground) than the typical narrow bipolar pulses, and were harder to distinguish from other lightning pulses (e.g. initial breakdown pulses). In order to make the classification more reliable, pulses from stations beyond 600-km are not classified.

3.3 Preprocessing

We first truncated the raw decimated waveform using a 100-μs window (centered on the pulse peak). Then pre-check procedures were implemented to discard saturated pulses and also pulses without enough data points. Specifically, the initial half cycle and the overshoot of the pulse each need to have each more than five data points. Then raw decimated waveform in the window was linearly interpolated to 1 MHz. The pulse was normalized so that the peak amplitude is 1 (for positive pulses) or -1 (for negative pulses), where positive pulses correspond to negative charge moving up and negative pulses correspond to negative charge moving down. The 101 samples in the 100-us window are used as features to training the model. Examples of waveforms of NBEs of both polarities are shown in Figure 1. The raw decimated waveforms are shown in black dots while interpolated waveforms are shown in blue line.

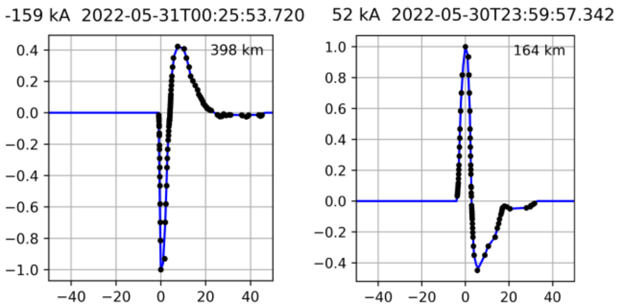

Figure 1.. Example waveforms of a –NBE (left) and a +NBE (right).

3.4 Dataset used for training and testing

The dataset used for testing and training consists of a total of 24,000 manually classified pulses evenly distributed in 4 categories (-NBE, -non-NBE, +NBE, +non-NBE). As mentioned before, only pulses from stations within 600-km of a lightning event were included. We trained a classification model individually for positive and negative pulses.

3.5 Model Performance

The classification model was trained and tested using a 4-fold cross-validation for tuning the hyper-parameters and evaluating the performance of the model. The average classification accuracies were 96% and 95% for negative and positive pulses, respectively.

3.6 Independent Validation

The classification accuracy was additional validated against two NBE datasets both recorded in Florida in 2022. The first dataset comprises 30 –NBEs recorded by the electric field sensors at the Lightning Observatory in Gainesville (LOG), Florida [14]. All of them were correctly classified as –NBEs by our model. The second dataset comprises a total of 494 +NBEs recorded by the low-frequency magnetic field sensor at Florida Institute of Technology, which were identified using an independent machine-learning model developed by Pu et al. [15]. It was found that 94% of 494 +NBEs were also classified as +NBEs by our model.

4 Initial results

From February to November 2022, a total of 10.4 million NBEs were identified, which account for 0.8% of total lightning events that were classified. Note that not all lightning events detected by the ENTLN were classified because of the distance constraint introduced in Section 3.2. The number of NBEs detected by the ENTLN is shown with a spatial resolution of 1° by 1° in Figure 2. Due to the limited coverage as well as the distance constraint, NBEs in some land regions (e.g., Amazon, Eastern Europe, and Northern Africa) and over deep oceans are not detected. One can see that the contiguous United States (CONUS), Argentina, Bangladesh, Southeastern Asia, and Northern Australia are regions with most NBEs. However, the number of NBEs is biased due to the detection efficiency of the ENTLN, which is not uniform across all regions that are covered.

To mitigate the influence of the detection efficiency, the percent of all classified events that are NBEs in 1° by 1° bounding box is calculated and shown in Figure 3a. We can see clear contrast in NBE percent in tropics (0°-23.5°) and mid-latitudes (30°-60°) and most of the NBE hotspots are in tropics. The overall NBE percent in tropics is 1.1%, which is about double the 0.58% in mid-latitudes. We also zoom in the CONUS and Australia, where the ENTLN has continuous coverage and very high detection efficiency, to determine the NBE dependence on latitude, and the results are shown in Figures 3b and 3c. In both regions, the NBE percentage decreases with increasing latitude. It is known that tropical regions have more lightning due to higher amount of moisture and heat present in the atmosphere. The warm, moist air in the tropics rises rapidly, leading to the formation of thunderstorms and lightning. Additionally, tropical regions are located near the equator, where the Earth's surface is heated most directly by the sun, leading to even more atmospheric instability and thunderstorm development. However, it is still not clear what meteorological factors cause the tropical thunderstorms to produce higher percent of narrow bipolar events.

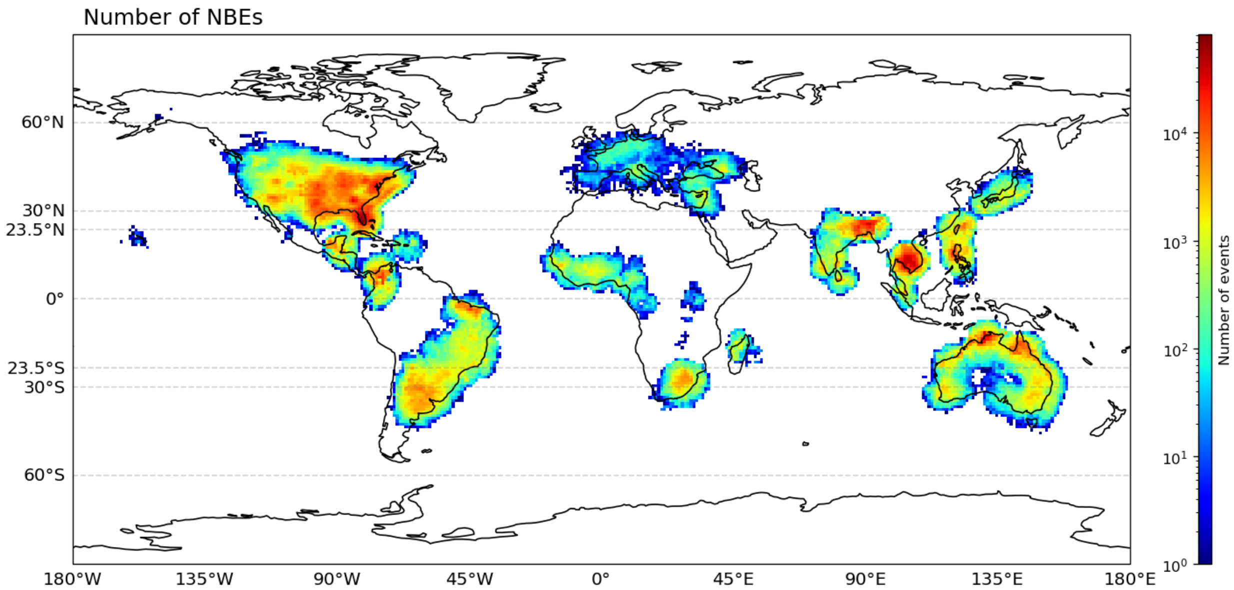

Figure 2. Number of NBEs (regardless of polarity) detected by the ENTLN with a resolution of 1° by 1°.

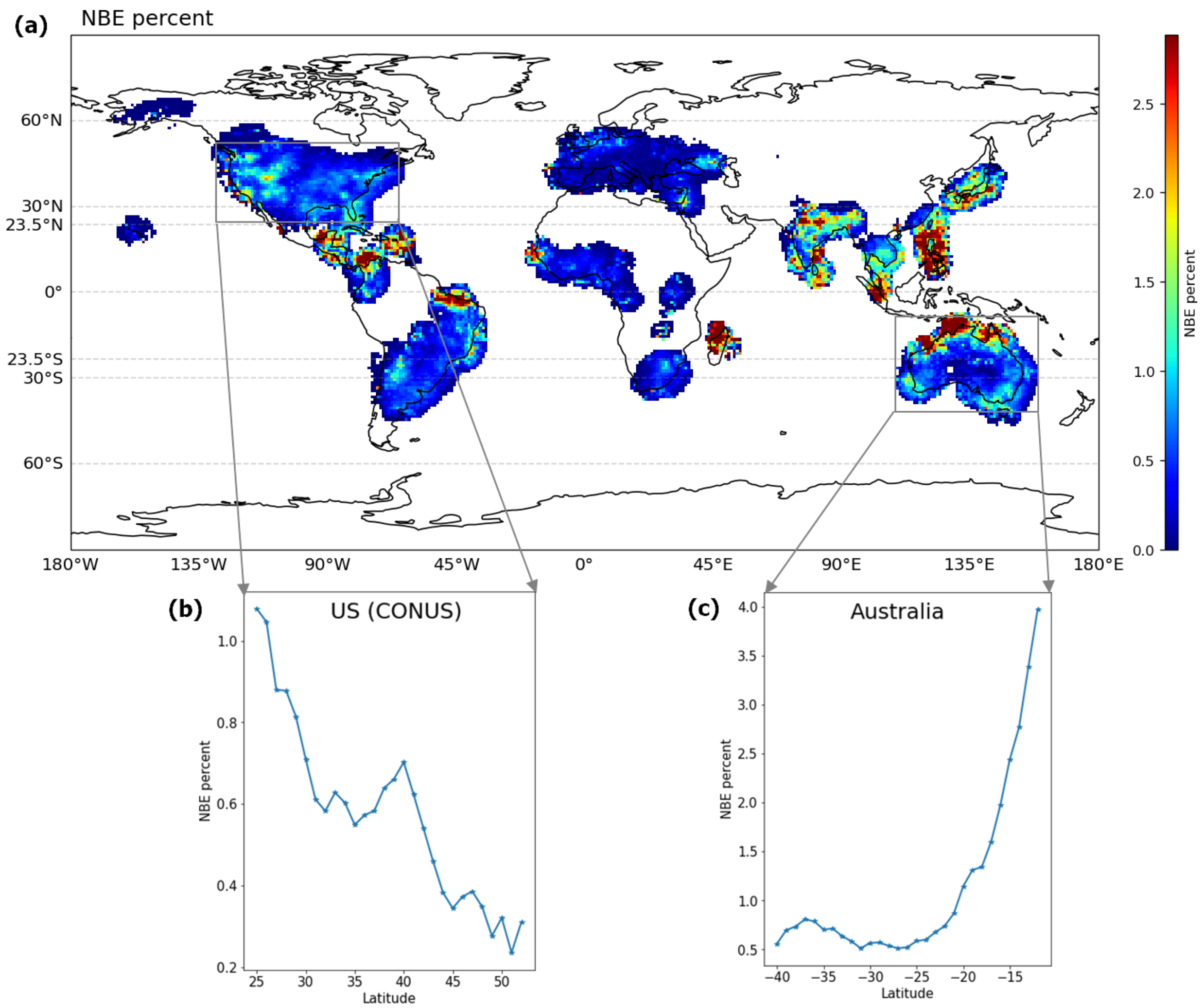

Figure 3. (a) Percent of lightning events that are NBEs with a resolution of 1° by 1°. (b) and (c) Variation of NBE percent versus latitude in the US (CONUS) and Australia, respectively.

Acknowledgements

S. Chen, Z. Ding, and V. A. Rakov were supported in part by U.S. NSF grants AGS-2114471 and AGS-2055178. Y. Pu and S. A. Cummer are supported by U.S. NSF grant AGS-2026304.

References

[1] D. A. Smith *et al.*, “A distinct class of isolated intracloud lightning discharges and their associated radio emissions,” *J. Geophys. Res. Atmos.*, vol. 104, no. D4, pp. 4189–4212, 1999.

[2] D. M. Le Vine, “Sources of the strongest RF radiation from lightning,” *J. Geophys. Res.*, vol. 85, no. C7, p. 4091, 1980.

[3] A. Nag and V. A. Rakov, “Compact intracloud lightning discharges: 2. Estimation of electrical parameters,” *J. Geophys. Res.*, vol. 115, no. D20, p. D20103, Oct. 2010.

[4] A. R. Jacobson, W. Boeck, and C. Jeffery, “Comparison of Narrow Bipolar Events with Ordinary Lightning as Proxies for Severe Convection,” *Mon. Weather Rev.*, vol. 135, no. 1, pp. 1144–1154, 2005.

[5] A. F. R. Leal and V. A. Rakov, “A study of the context in which compact intracloud discharges occur,” *Sci. Rep.*, vol. 9, no. 1, pp. 1–15, 2019.

[6] T. Wu, S. Yoshida, T. Ushio, Z. Kawasaki, and D. Wang, “Lightning-initiator type of narrow bipolar events and their subsequent pulse trains,” *J. Geophys. Res. Atmos.*, vol. 119, no. 12, pp. 7425–7438, Jun. 2014.

[7] W. Rison *et al.*, “Observations of narrow bipolar events reveal how lightning is initiated in thunderstorms,” *Nat. Commun.*, vol. 7, p. 10721, 2016.

[8] J. N. Tilles *et al.*, “Fast negative breakdown in thunderstorms,” *Nat. Commun.*, vol. 10, no. 1, p. 1648, Dec. 2019.

[9] F. Liu *et al.*, “Observations of Blue Discharges Associated With Negative Narrow Bipolar Events in Active Deep Convection,” *Geophys. Res. Lett.*, vol. 45, no. 6, pp. 2842–2851, Mar. 2018.

[10] S. Soler *et al.*, “Blue Optical Observations of Narrow Bipolar Events by ASIM Suggest Corona Streamer Activity in Thunderstorms,” *J. Geophys. Res. Atmos.*, vol. 125, no. 16, pp. 1–13, 2020.

[11] C. Liu and S. Heckman, “The application of total lightning detection and cell tracking for severe weather prediction,” in *91st AmericanMeteorological Society Annual Meeting*, 2011.

[12] Y. Zhu, M. Stock, J. Lapierre, and E. DiGangi, “Upgrades of the Earth Networks Total Lightning Network in 2021,” *Remote Sens.*, vol. 14, no. 9, p. 2209, May 2022.

[13] Y. Zhu, P. Bitzer, V. Rakov, and Z. Ding, “A Machine-Learning Approach to Classify Cloud-to-Ground and Intracloud Lightning,” *Geophys. Res. Lett.*, vol. 48, no. 1, pp. 1–8, 2021.

[14] V. A. Rakov *et al.*, “New insights into the lightning discharge processes,” *Plasma Sources Sci. Technol.*, vol. 31, no. 10, p. 104005, Oct. 2022.

[15] Y. Pu *et al.*, “Unsupervised Clustering and Supervised Machine Learning for Lightning Classification: Application to Identifying EIPs for Ground-based TGF Detection,” *Authorea*, 2022.