NOAA ROSES Semi-Annual Report

Reporting Period: March 2021 – August 2021 (2nd report)

PI: Stephanie Stevenson
Co-PI(s): Kate Musgrave (CIRA/CSU Institutional PI)

Project Title: Enhancing forecast applications of the GOES-R GLM in tropical cyclones using multi-platform data fusion and AI to assess environment and storm structure

Executive Summary (1 paragraph max)

This project aims to utilize machine learning to aid in the development of an automated real-time predictive tool that can assess links between the Geostationary Lightning Mapper (GLM), tropical cyclone (TC) structure, and TC intensity change to improve intensity forecasts. The Year 1 milestones focus on data collection and data processing in anticipation for the machine-learning training and tool development in Years 2 and 3. Steady progress continued during this reporting period, with the main accomplishments focused on proper quality control of the GLM data and continued assembly of the machine-learning dataset. The core project team continues to hold bi-weekly meetings to maintain communication and keep the project on track.

Progress toward FY20 Milestones and Relevant Findings (with any Figs)

Year 1 milestones were focused on collection, assembly, and intercomparisons of the machine-learning data inputs. The bulk of this reporting period was focused on quality control of GLM data and assembling the collected and processed data for machine-learning. The team met with GLM instrument experts and other researchers utilizing the data for TC applications in May 2021 to discuss strategies for quality control. Subsequent analysis of the GLM lightning dataset revealed several issues that may impact the climatological evaluation and machine-learning if not addressed, including spurious lightning reported by blooming, sun glint, false lightning on the subarray boundaries (e.g., “Bahamas bar”). Based on identification of these false lightning events, several quality control approaches were explored:

- Discarding GLM lightning observed outside of convective regions identified by the Advanced Baseline Imager (ABI) infrared channel using a brightness temperature threshold within 0.5 degree of the observed lightning
- Identifying thresholds of GLM group energy and area typically associated with erroneous GLM lightning data (**Fig. 1**)
- Variability in the lightning data quality (including diurnal characteristics) as a function of GLM build number
- Using machine-learning techniques such as convolutional neural networks (**Fig. 2**) to train a model on good lightning data (including ground-based lightning networks like WWLLN) to filter erroneous GLM lightning

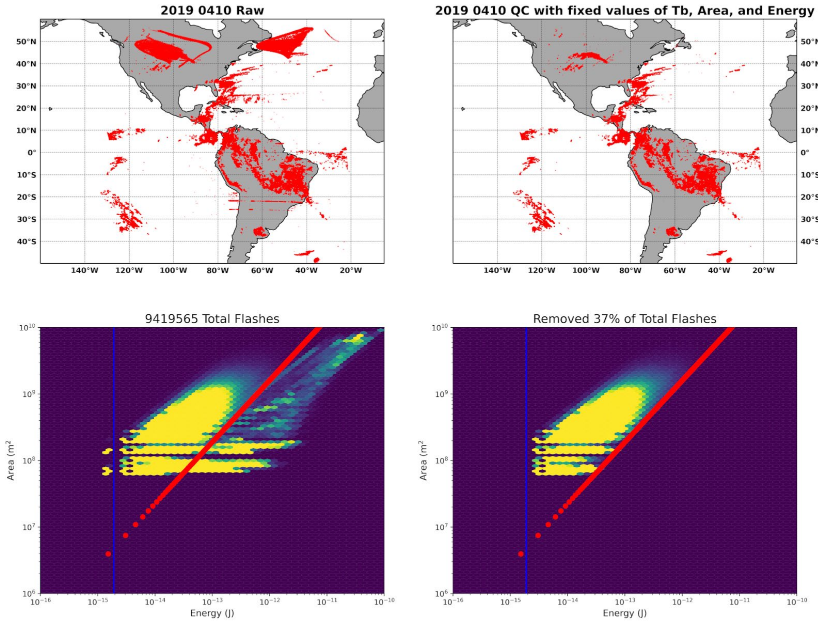

Figure 1. Sample during April 10, 2019 of raw GLM flashes (top left) and the distribution of flash area and energy for the flashes (bottom left). The right panels show quality-controlled data using an ABI Ch 13 brightness temperature threshold of -10°C, an energy threshold of $2x10^{-24}$ J (blue line, bottom panels), and an area threshold equivalent to the function plotted that varies for a given energy (red line, bottom panels).

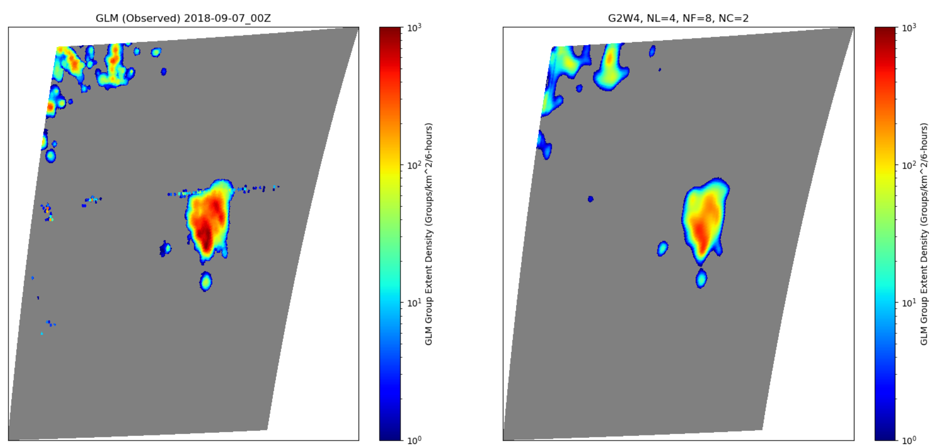

Figure 2. Sample during TC Florence (2018) of the (left) GLM observed group extent density [groups / km^2 / 6 h] and (right) auto-encoder machine-learning output of GLM group extent density trained to bias-corrected WWLLN.

Progress towards Year 2 milestones also continued during this reporting period. A large machine-learning dataset was constructed using 74+ input variables and associated metadata. A majority of the input variables are centered around lightning group extent density, group energy, and group area from GLM, WWLLN, or ISS-LIS accumulated over multiple time periods (1-, 3-, and 6-hour). Additional variables include ABI imagery, TC geometry, solar geometry, environmental variables of SST, OHC, and AOD, and measures of vertical wind shear. Given that TC intensity and intensity change prediction is the goal of the project, the output variables for the dataset include measures of maximum tangential wind speed and minimum central pressure, and their change with time.

Plans for Next Reporting Period

The project team plans to wrap up the climatological comparisons between the lightning databases and AOD, SST, and OHC during the next reporting period. While these were part of the Year 1 objectives, one of the key project team members moved on to another position during this reporting period. We do not anticipate significant delays in the project given that several of the Year 2 milestones were started in Year 1. Initial progress on the Year 1 objectives will be presented at the American Geophysical Union (AGU) and American Meteorological Society (AMS) annual meetings in December 2021 and January 2022, respectively. Progress will continue on Year 2 milestones with machine learning testing with the compiled input datasets aimed at understanding spatial lightning features and their association with TC intensity and intensity change.