AN AUTOMATED BOLIDE DETECTION PIPELINE FOR GOES GLM

Jeffrey C. Smith[1,2], Robert L. Morris[1,2], Clemens Rumpf[2], Randolph Longenbaugh[3], Nina McCurdy[2], Christopher Henze[2] and Jessie Dotson[2]
[1]The SETI Institute, [2]NASA Ames Research Center (jsmith@seti.org), [3]Sandia National Laboratories

The Geostationary Lightning Mapper (GLM) instrument onboard the GOES 16 and 17 satellites has been shown to be capable of detecting bolides (bright meteors) in Earth's atmosphere. Due to its large, continuous field of view and immediate public data availability, GLM provides a unique opportunity to detect a large variety of bolides, including those in the 0.1 to 3 m diameter range and complements current ground-based bolide detection systems, which are typically sensitive to smaller events. We present a machine learning based bolide detection and light curve generation pipeline being developed at NASA Ames Research Center as part of NASA's Asteroid Threat Assessment Project (ATAP). The goal is to generate a large catalog of calibrated bolide lightcurves to provide an unprecedented data set which will be used to inform meteor entry models on how incoming bodies interact with the Earth's atmosphere and to infer the pre-entry properties of the impacting bodies. The data set will also be useful for other asteroidal studies. The pipeline runs in an automated fashion and bolide light curves along with other measured properties are promptly published on a NASA hosted publicly accessible website, https://neo-bolide.ndc.nasa.gov.

Introduction NASA's Asteroid Threat Assessment Project (ATAP) is a NASA Ames Research Center activity in support of NASA's Planetary Defense Coordination Office (PDCO). A first element of the project involves estimating the characteristics of potentially hazardous objects prior to impact. To validate the reentry models ATAP is generating a database of bolide light curves. A prototype has been running for over a year now with over 1700 bolides posted at: https://neo-bolide.ndc.nasa.gov. Our goal is the detection and characterization of bolides in order to calibrate our reentry models, but the fused data set will be useful for numerous other science studies.

There exists no uniformly processed set of bolide light curves for use with tuning re-entry modeling software, which currently uses a very small number of well studied events, such as the Chelyabinsk meteor. In addition, most ground based camera systems have limited sight and easily saturate and thus do not produce accurate disintegration light curves. With the stereo detection nature of GOES 16 and 17 it is also possible to reconstruct the trajectory of the bolides, which is useful for both understanding the object's origin (I.e. Planetary Science) and final destination (for possible recovery).

The GOES Geostationary Lightning Mappers [1] positioned over the western hemisphere in geosynchronous orbit detect transient light events at 500 frames per second (2 ms). The GLM instrument is a staring CCD imager (1372 $\times$ 1300 pixels) which measure emissions in a narrow 1.1 nm pass-band centered at 777.4 nm, a principal wavelength for the neutral atomic oxygen emission (OI) line of the lightning spectrum. This passband is quite narrow, but GLM will still detect metal atom lines and continuum emission in the narrow pass band [2]. This means despite GLM's singular goal of detecting lightning, papers have already demonstrated the ability to detect bolide events in GLM data [2, 3].

The pipeline work has been carried out in a series of stages. Stage 1 was to develop a prototype detector using classic filter design [3] and demonstrate the feasibility of detecting bolides in GLM. Stage 2 was to wrap the prototype algorithms in a robust pipeline architecture for automatic and scheduled execution on the NASA Ames Supercomputer. Stage 3 was to collect detections from the prototype pipeline and then perform human vetting in order to generate a labeled data set. Stage 4 was to use the training data set to train a random forest machine learning classifier to improve the performance of the detector.

Pipeline Architecture The previous work [3] utilized classic filter design to create a set of detection filters. Much effort was placed on the design of these filters, or 'features' in Machine Learning parlance, with each feature identifying a distinct aspect of the data. Although an excellent first step, the classic filter design did have its limitations. Namely, limitations in the variance in the classification logic resulted in the model under-fitting. The algorithms developed in [3] were used as a basis for a first generation GLM bolide detection pipeline. The output of this pipeline was used to develop a training set to develop a second iteration random forest based pipeline, which is now running daily.

Figure 1 gives a flow chart of the current pipeline. The pipeline runs on each GLM 20-second Level 2 netCDF file. The GLM data groups are first clustered into potential bolide detections. The clusters are then classified into bolides and not bolides. The detection candidate information is then passed to a GOES Advanced Baseline Imager (ABI) cutout tool which identifies a section of a GOES ABI image data in numerous bands to examine evidence if clouds and weather activity in both day and night settings is present around the bolide candidate. A further figure is also generated which shows all GLM event data within a region around the candidate. All generated figures are then packaged up for manual

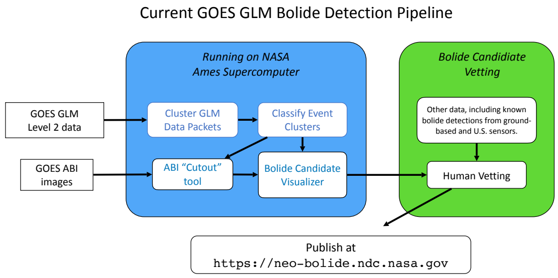

Figure 1: Flowchart of current GLM bolide pipeline.

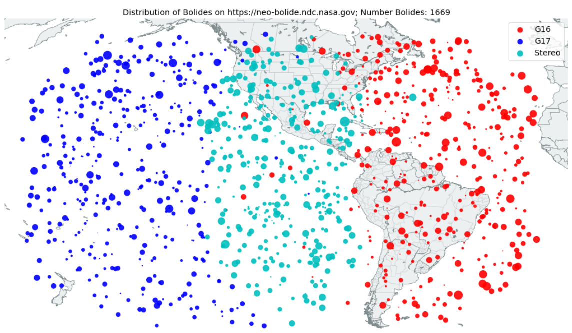

Figure 2: Distribution of 1669 bolides listed on the website as of 2020-11-28. Most are detected via the pipeline but a small number are found via other means.

perusal by the human vetters. Those candidates that pass vetting are deemed "discovered" bolides and posted on our website.

Pipeline Performance We show the distribution of bolides listed on our website in Figure 2. All bolides shown are in the GOES GLM data but not all were found in the pipeline. A small number were found by other means, but in those cases the GLM data was re-examined and if the data was present they were posted to the website, with appropriate comments attached to each event. Bolide total integrated energy is proportional to dot size. Those present in GOES-16 are in red, those in GOES-17 in blue and those present in both in cyan. A small number in the stereo region have observable data in only one of the satellites and show up as only blue or red within the cyan dots. The astute observer will also notice a single cyan dot well east of the North American coast. This bolide was indeed observed by both instruments when GOES-17 was in its staging location at -89.5 Degrees East before being parked in its final location at -137.2 Degrees East in November, 2018.

There is also a clear dependence in detected bolides with time, as shown in Figure 3. Peaks corresponding to known showers, [4], are highlighted. We are detecting many of the major showers and perhaps some of the

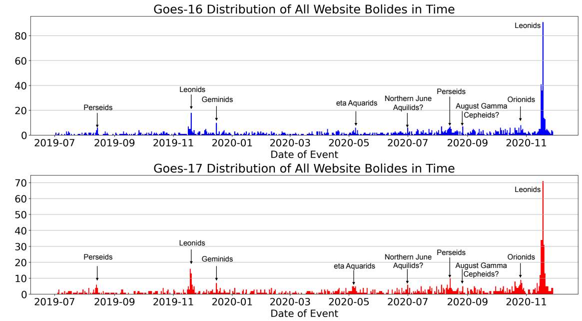

Figure 3: Distribution of 1669 bolides listed on the website as of 2020-11-28.

weaker showers. The Leonid meteor showers in November, 2019 and 2020 are very clear. It is not yet known why so many more Leonids were detected in 2020 than 2019. Further investigations are needed to determine if this is selection bias or a true increase in detectable meteoroids in 2020.

The current pipeline detection performance results in a detection precision of 45.9% and 41.2% for GOES-16 and 17 respectively. This precision results in about 5 detections per day per satellite and the burden on human vetters is sufficiently low. However for an automated pipeline with no human-in-the-loop inspection, higher precision is required and is being pursued.

The primary use of our data is to generate a set of calibrated bolide reentry light curves to tune reentry modelling software. But other uses are also apparent. We can clearly identify common meteor shower events in our data and a statistical analysis of this distribution could aid in understanding meteor shower bulk compositions. Within the stereo overlap region viewed by both GOES-16 and 17, it is possible to reconstruct the trajectory of the incident bolides. Such a reconstruction can aid in both investigations of the origins of the bolides and also increase the prospects of potential recovery.

References

[1] S. J. Goodman, *et al.*, *Atmospheric Research* **125**, 34 (2013).

[2] P. Jenniskens, *et al.*, *Meteoritics & Planetary Science* **53**, 2445 (2018).

[3] C. M. Rumpf, R. S. Longenbaugh, C. E. Henze, J. C. Chavez, D. L. Mathias, *Sensors (Basel, Switzerland)* **19**, 1008 (2019).

[4] A. meteor Society, *Meteor Shower Calendar 2020-2021* (2020). https://www.amsmeteors.org/meteor-showers/meteor-shower-calendar/.