**RELAMPAGO LMA Level 2 File Documentation**

Level 2 data were created using the lmatools Python package (https://github.com/deeplycloudy/lmatools). A minimum of 6 detecting stations are required to consider a VHF source as part of a flash. Maximum allowable distance and maximum allowable time between successive sources in a flash is 3000 m and 0.15 s, respectively. Maximum allowable flash duration is 3 s. See references below for more information about LMA data, processing, and each variable.

Each RELAMPAGO LMA Level 2 file is an HDF5 architecture containing two overarching data tables: "events" and "flashes." Each of these houses a sub-table with the name "LMA_yymmdd_HHMMSS_0600," where yymmdd is two-digit year, month, and day, and HHMMSS is two-digit UTC time stamp of the 10-minute file. Note that RELAMPAGO LMA data always start (and end) 1 second after each 10-minute period (e.g., 000001, 001001, 002001, etc.). This is an artifact of how the Level 0 data were recorded.

Inside each of these sub-tables are structured arrays containing the actual data. For example, in the REL_190102_231001_0600.dat.flash.h5 file, the event data are all contained in a structured array within events/LMA_190102_231001_0600 and the flash data are all contained in a structured array within flashes/LMA_190102_231001_0600.

The *events* structured array reproduces much of the input Level 1 VHF source location file but also adds *flash_id*, a unique tag that matches each relevant event to a particular flash. The columns in order are as follows:

- 'alt' (float, altitude in m above the WGS84 ellipsoid)
- 'charge' (int, charge identification - not used, always 0)
- 'chi2' (float, chisqr value for the VHF source)
- 'flash_id' (int, matches each relevant source to its parent flash in the "flashes" table - see below)
- 'lat' (float, decimal latitude)
- 'lon' (float, decimal longitude)
- 'mask' (str, hexadecimal mask)
- 'power' (float, VHF source power)
- 'stations' (int, number of stations contributing to source's location solution)
- 'time' (double, second of day; note that source data from the 2350 UTC file from each day end 1 second into the following day)

The *flashes* structured array contains information about each flash identified in the dataset. The columns in order are as follows:

- 'area' (float, $km^2$)
- 'ctr_alt' (float, altitude of flash centroid in m above the WGS84 ellipsoid)
- 'ctr_lat' (float, decimal latitude of flash centroid)
- 'ctr_lon' (float, decimal longitude of flash centroid)

- ‘duration’ (float, duration of flash in seconds)
- ‘flash_id’ (int, unique identification number for each flash)
- ‘init_alt’ (float, altitude of flash initiation in m above the WGS84 ellipsoid)
- ‘init_lat’ (float, decimal latitude of flash initiation)
- ‘init_lon’ (float, decimal longitude of flash initiation)
- ‘init_pts’ (str, not used; always empty)
- ‘n_points’ (int, number of sources in flash)
- ‘specific_energy’ (float, J/kg)
- ‘start’ (double, second of day for first source in flash; note that flash data from the 2350 UTC file from each day end 1 second into the following day)
- ‘total_energy’ (float, J)
- ‘volume’ (float, $km^3$)

In Python, you can retrieve level 2 data into numpy structured arrays using the following example code snippet:

```
import numpy as np
import h5py
data = h5py.File('REL_190102_231001_0600.dat.flash.h5')
events = np.array(data['events']['LMA_190102_231001_0600']
flashes = np.array(data['flashes']['LMA_190102_231001_0600']
data.close()
```

**References**


Bruning, E.C. and D.R. MacGorman, 2013: Theory and Observations of Controls on Lightning Flash Size Spectra. J. Atmos. Sci., 70, 4012-4029, https://doi.org/10.1175/JAS-D-12-0289.1

Chmielewski, V. C., and Bruning, E. C. ( 2016), Lightning Mapping Array flash detection performance with variable receiver thresholds, J. Geophys. Res. Atmos., 121, 8600- 8614, doi:10.1002/2016JD025159.

Fuchs, B. R., Bruning, E. C., Rutledge, S. A., Carey, L. D., Krehbiel, P. R., & Rison, W. ( 2016). Climatological analyses of LMA data with an open-source lightning clustering algorithm. Journal of Geophysical Research: Atmospheres, 121, 8625-8648. https://doi.org/10.1002/2015JD024663

Rison, W., R. Thomas, P. Krehbiel, T. Hamlin, and J. Harlin (1999), A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico, Geophys. Res. Lett., 26, 3573-3576.

Thomas, R. J., Krehbiel, P. R., Rison, W., Hunyady, S. J., Winn, W. P., Hamlin, T., & Harlin, J. ( 2004). Accuracy of the lightning mapping array. Journal of Geophysical Research, 109, D14207. https://doi.org/10.1029/2004JD004549