Title: DC3 flash product data for select IOPs 23 May 2014

Author(s)

PI: Dr. Eric Bruning
Assistant Professor, Atmospheric Science Group
Department of Geosciences
Texas Tech University, Box 41053, Lubbock, TX 79409-1053 USA
Ofc: 806 834-3120 / Fax: (806)-742-1738
eric.bruning@ttu.edu

1.0 Data Set Overview:

This dataset is comprised of lightning flash statistics derived from the original VHF measurements provided by the North Alabama, Colorado, Oklahoma, and West Texas Lightning Mapping Arrays.

These include 2-minute time series of flash count and size statistics derived from the flash data, as well statistics of the pixels in the 5-minute gridded flash product data

Time period covered by the data: Specific IOPs from May 14, 2012-June 30, 2012

2.0 Instrument Description:

See the description of each LMA in its respective source dataset.

3.0 Data Collection and Processing:

The process to sort the ASCII VHF source data into flashes (stored along with their constituent VHF source data in HDF5 format) and to create the basic product grids (in NetCDF) is described in the documentation for the WTLMA lightning flash products, with minor differences to adjust for each network center. That text is repeated in section 3.0.1 for convenience.

This dataset contains flash statistics and plots derived from the flash-level HDF5 files and NetCDF product grids. Both the flash-level and gridded data were subsetted using a temporally varying spatial mask. That mask was the intersection the following two rectangular domains:

1. IOP-specific bounding boxes in IOP*-boundingbox.txt
2. The domain of the flash product grids

The IOP bounding boxes were provided by Kristin Cummings, and cover storm complexes, not individual storm cells. The flash product grids were originally created so that the perimeter of the grids represented an approximate limit to the good 2D coverage domain for each LMA network. Most bounding boxes lasted longer than the extent of the

2D LMA coverage. Therefore, this dataset represents a high-level overview of flash statistics for each IOP while the storm was in good LMA coverage.

All of the analysis data files and figures can be reconstructed from the flash data and derived grids and by running the included Python scripts as derived in environment_setup.README.

3.0.1 Description of LMA flash data (*repeated from the WTLMA flash documentation*)

Data were collected through a combination of automated and manual collection to ensure data from as many stations as possible were incorporated. Each LMA sensor recorded the peak pulse of all all pulses to break threshold in an 80 μs interval. VHF source locations from at least 5 sensors are used to obtain the VHF source location using a time of arrival technique. A two-pass processing technique was used, with the first pass accepting all retrieved sources with chi-squared less than 5.0, and a second pass that permitted additional points with large chi-squared values.

VHF source locations are then clustered into flashes using temporal and spatial criteria according to McCaul et al. (2005, 2009). All sources with chi-squared less than 5.0, altitude less than 20 km, and detected by at least six stations were candidates for clustering. A 3 km spatial and 0.15 second temporal threshold was used for clustering, and no flash was allowed to exceed 3 s in length. Although no minimum source number per flash was prescribed in the flash data output, we recommend defining as valid flashes as containing ≥ 10 sources per flash to mitigate the presence of noise. The more experienced user may wish to use a different source number per flash threshold to define a valid flash, which is why all potential flashes are provided regardless of source number.

The gridded flash products were created using the lmatools Python software package (Bruning 2013). A minimum of ten sources was required for a flash to appear on the grid. Four product types were created: VHF source density, flash initiation density, flash extent density, and average flash area. These products were created in 5-min windows on a regular lat-lon grid with nominal (at the grid center) spacing of 3 km. For each of these products, except for perhaps flash initiation density, a logarithmic color mapping usually works best. A brief description of these product types follows.

Source density
This product is a simple count of the VHF sources on the grid. It is the most basic form of gridded data that can be derived from the LMA data. It only counts those sources that contributed to a valid flash.

Flash initiation density
For each flash, the first VHF source point in time is taken as the flash initiation location. Therefore, each flash is represented by a single grid cell location A sum of all values on this grid gives the total number of flashes in the domain. The electric field rebuilds

relatively rapidly in regions with larger flash initiation rates, since flashes initiate in regions of large electric field.

Flash extent density
Flash extent density gives a column-local flash rate; it is a count of how many flashes passed through that grid cell. This product highlights a key advantage of the LMA: it shows regions where it was energetically favorable (local maximum in electric potential) for flashes to frequently revisit.

Average flash area (footprint)
The average flash area product calculates the average area of all flashes that passed through a grid cell. It is the sum of the areas of all flashes that passed through a grid cell divided by the flash extent density. This product is good at highlighting regions with infrequent but very extensive flashes, such as MCS stratiform regions and supercell anvils. It can also highlight convective regions that have unusually small, frequent flashes.

4.0 Data Format:

4.0.1 Analysis derived from the HDF5 flash-level data
This analysis was primarily produced by code in the energy_stats.py module.

Plot of all flash centroids.
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/flashes.png
One dot per flash colored by time. Color scale is from blue to green to yellow to red corresponds to the range of times listed in the title of the plot
Flash count, size and energy statistics
./figures/IOP*.output.flash_stats.csv
ASCII, comma separated. Every two minues. One header line. Columns are:
start_isoformat, end_isoformat: UTC start and end times in ISO 8601 format
number : count of flashes
mean, variance, skewness, kurtosis: raw moments of the flash size distribution
energy: number * (mean*mean + variance)
energy/number: the energy column divided by the number column
Plot of data from flash_stats.csv
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/moment-energy-timeseries.pdf
Plot of flash energy spectra for each LMA data file
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/LYLOUT*-energy.pdf
Raw output from the processing script for each run
./figures/IOP*.output.txt

4.0.2 Analysis derived from the flash product grids
For each IOP there are four subfolders, one for each flash product grid type:
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/grids_flash_extent/

./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/grids_flash_footprint/
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/grids_flash_initiation/
./figures/IOP*-boundingbox-thresh-0.15_dist-3000.0_pts-10/grids_lma_source/

Within each of the above folders are the following:
Statistics of pixels (grid boxes) that comprise each 5 minute grid interval
e.g., flash_extent_20120603_015501.png
Columns are:
time (ISO): UTC start time of the interval in ISO 8601 format
max count per grid box: Maximum value of all pixels
sum of all grid boxes: Sum of values of all pixels
5th percentile, 50th percentile, 95th percentile: Value at that percentile of the distribution of values of all pixels
Plots of each 5 minute grid interval
e.g., flash_extent_20120603_015501.png
A movie assembled from the series of the above images
e.g. flash_extent_20120602.mp4 — H.264 encoded MPEG4 video file

4.0 References:

Bruning, E. 2013-2014: lmatools source repository, https://github.com/deeplycloudy/lmatools, retrieved 30 April 2014.