Title: The North Alabama Lightning Mapping Array (NALMA) dataset

Author(s):

PI Contact:
Dr. Larry Carey, UAH
320 Sparkman Dr. Huntsville, AL, 35805
larry.carey@nsstc.uah.edu

Data questions contact:
Dr. Rich Blakeslee, NASA Marshall Space Flight Center
320 Sparkman Dr. Huntsville, AL, 35805
rich.blakeslee@nasa.gov

Dr. Eric Bruning, Texas Tech University (lmatools data set)
Box 41053 Lubbock, TX 79409-2101
Eric.bruning@ttu.edu

Document Authors:
Christopher Schultz, NASA MSFC/UAH and Mariana Scott, ESSC UAH
320 Sparkman Dr. Huntsville, AL, 35805
christopher.j.schultz@nasa.gov, schultz@nsstc.uah.edu

1.0 Data Set Overview:

This dataset is total lightning measurements from the North Alabama Lightning Mapping Array. Included are: raw VHF source information from the LMA that has been reprocessed and quality controlled, lightning flash output from a flash clustering algorithm that combines VHF source information into flashes using space and time criteria, and a gridded flash product using lmatools.

Time period covered by the data: May 15, 2012-June 15, 2012
Center of LMA: 34.7246100 -86.6453300
Data source: NASA MSFC
LMA website: http://weather.msfc.nasa.gov/sport/lma/ or
http://branch.nsstc.nasa.gov/PUBLIC/NALMA/

2.0 Instrument Description:

Instrument description: http://weather.msfc.nasa.gov/sport/lma/
Table of specifications and limitations: see Koshak et al. 2004, JTECH, Vol .21, 543-558

3.0 Data Collection and Processing:

Data are collected autonomously through a series of 11 sensors across north Central Alabama and 2 stations near Atlanta, Georgia. Each LMA sensor samples in 80 µs intervals to observe VHF radiation as electrical breakdown occurs. VHF source locations from at least 6 sensors are used to obtain the VHF source location using a time of arrival technique.

VHF source locations are then clustered into flashes using temporal and spatial criteria according to McCaul et al. (2005). Although no minimum source number per flash was prescribed in the flash data output, we recommend defining 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 package and provided to UAHuntsville researchers by Dr. Eric Bruning from Texas Tech University. VHF source locations are clustered into flashes using a space and time criteria using the McCaul et al. (2009) algorithm. For this data delivery, a 3 km spatial and 0.15 second temporal thresholds were used and a gridded product was created. Four product types were created: source density, flash initiation density, flash extent density, and average flash area. A thorough description of the product created by Dr. Bruning's lmatools package can be found at https://bitbucket.org/deeplycloudy/lmatools.

More information on lightning mapping array technology can be found in the following papers:

Rison et al. 1999, GRL, VOL. 26, NO. 23, PAGES 3573-3576, DECEMBER 1, 1999
Krehbiel et al. 2000, EOS, Vol 81, No. 3, 21-32
Thomas et al. 2001, GRL, VOL. 28, NO. 1, PAGES 143-146, JANUARY 1, 2001
Thomas et al. 2004, JGR, VOL. 109, D14207, doi:10.1029/2004JD004549, 2004

Please contact Jeff Bailey (email: Jeffrey.C.Bailey@nasa.gov, collection of data) or John Hall (email: john.hall@nasa.gov; post processing of data), for specific details on the collection or reprocessing procedure.

4.0 Data Format:

4.0.1 NALMA VHF Sources

Data file structure is ASCII, with naming convention for the files being LYLOUT_yymmdd_hhmmss_duration (seconds).dat

Data Format: Header information is 65 lines long and is in each hourly file. Header information includes the following:
LMA center location

Station locations
number of active stations
chi-squared used in reprocessing
minimum number of stations used in processing

List of parameters with units, sampling intervals, frequency, range:
Sampling rate: 80 μs
Range: < 50 m 3D error within 150 km of the center of the lightning mapping array. Thus data most accurate within 150 km of the LMA center. See Koshak et al. 2004, JTECH for more information.

ASCII Columns:
time (UTC sec of day), latitude (decimal degrees), longitude (decimal degrees), altitude (m), reduced chi^2 (unit less), Power (dBW), mask (unit less)

format(15.9f 10.6f 11.6f 7.1f 5.2f 5.1f 4x)

4.0.2 NALMA Flashes

Data file structure of hourly (UTC) flash file is ASCII, with naming convention for the files being: LMA_NA_6.2_080_yyyy-mm-dd_hh-mm-ss.dat

ASCII Columns:
VHF source number (unit less), time (UTC second of day), latitude (decimal degrees), longitude (decimal degrees), altitude (m), reduced chi^2 (unit less), Power (dBW), range to center of network (km), number of stations (unit less), flash number (unit less)

format(1x,i8,f13.6,f9.5,f11.5,f10.3,3f10.4,i3,i6)

VHF sources with the same "flash number" belong to the same flash according to the clustering criteria of McCaul et al. (2005). As noted earlier, no minimum source number per flash has been applied to the hourly flash files. We recommend defining valid flashes as containing ≥ 10 sources per flash to mitigate the presence of noise. The more experienced NALMA data user may wish to use a different source number per flash threshold to define a valid lightning flash, which is why all potential flashes are provided regardless of number of sources.

4.0.3 Gridded Flash Products using lmatools

Data file structure of four hourly (UTC) flash files, one for each product. The naming convention is NALMA_20120603_190000_3600_10src_0.0328deg-dx_flash_extent.nc

Where NALMA indicates the data type, 20120603 is the data collection date, 190000 is the data collection beginning time, 3600 is the duration in seconds of the data in the file, 10src is the minimum point per flash and 0.0328deg-dx is the grid size in the x direction, flash_extent is the product type, and .nc is the file extent 9in this case, netCDF). More information about these file types are available at https://bitbucket.org/deeplycloudy/lmatools.

Data product types:
There are four data products available in the gridded flash file type. Flash density is a count of the VHF sources on the grid for a valid flash. Flash initiation density is the first VHF source point for each flash, added to each cell in the grid. Flash extent density is a count of the number of flashes that passed through each grid cell, highlighting regions of local maximum electric potential. Finally, average flash area is the average area of all flashes that passes through each cell grid. This paragraph is only a short summary of each product type. For more information about these product types one should visit Dr. Bruning's documentation website at https://bitbucket.org/deeplycloudy/lmatools.

Data header:
All gridded LMA data has the following dimensions: longitude and latitude are 134 with 12 time steps "ntimes". The variables are dependent to each flash product file and are as follows:

LMA Source Density	**LMA Flash Footprint**	**Flash Initiation**	**Flash Extent**
int lma_source(ntimes, lon,lat)	float flash_footprint(ntimes, lon, lat)	int flash_initiation(ntimes, lon, lat)	int flash_extent(ntimes, lon, lat)
units = "Count per grid pixel per 5.0 min"	units = "km^2 per flash"	units = "Count per grid pixel per 5.0 min"	units = "Count per grid pixel per 5.0 min"
long_name = "LMA source density"	long_name = "LMA local mean flash area"	long_name = "LMA flash initiation density"	long_name = "LMA flash extent density"
grid_mapping = "crs"	grid_mapping = "crs"	grid_mapping = "crs"	grid_mapping = "crs"
missing_value = -9999	missing_value = -9999	missing_value = -9999	missing_value = -9999

The variables common to all gridded flash products are as follows:

int crs	float time(ntimes)	float longitude(lon)	float latitude(lat)
semi_major_axis = 6378137.f	units = "seconds since 2012-05-03 00:00:00"	units = "degrees_east"	units = "degrees_north"
longitude_of_prime_meridian = 0.f	long_name = "time"	long_name = "longitude"	long_name = "latitude"
grid_mapping_name = "latitude_longitude"		standard_name = "longitude"	standard_name = "latitude"
inverse_flattening = 298.2572f			

5.0 Data Remarks:

PI's assessment of the data (i.e., disclaimers, instrument problems, quality issues, etc.)

Missing data periods: None. NALMA stations that are not active are denoted in the header section of the source files starting at line 31. If the number of source collected column is 0, then the station was down.

During the over-whelming majority of the DC3 field campaign, the full 11 stations were operating in Northern Alabama with a few exceptions noted in the DC3 Field Catalog (FC). However, a post-processing error resulted in the loss of 2 sensor stations on the western side of the network (Firetower and Hospital sensors as seen in Fig. 1) from contributing to lightning source and flash location data.

It was determined after the DC3 field campaign that read errors associated with the NALMA sensor data header information during processing had caused the inadvertent loss of undecimated (i.e., not reduced for communication) data from two of the eleven VHF sensors in the NALMA (Firetower and Hospital; Fig. 1). Unfortunately, this error was not identified until after the hard drives associated with the two VHF sensors in question had been recycled (ie., reformatted and returned to the field), thus losing the sensor data and effectively reducing the sensor network coverage to 9 sensors in Alabama and 2 in Georgia during DC3. The remaining 9 sensors in northern Alabama still provide an LMA network of sufficient quality for DC3 research and analysis but the effective detection efficiency and location accuracy will be reduced somewhat from typical values (Koshak et al. 2004), especially in the western portions of the DC3 domain where the sensors were located (Fig. 1). Fortunately, the Doppler radar coverage and most of the DC3 cases were located toward the eastern side of the NALMA network so the impact of the two lost sensors should be somewhat mitigated. UAHuntsville and NASA MSFC will continue to assess the revised detection efficiency and location accuracy of the modified network and provide this information to DC3 participants so as to minimize the impact to DC3 science.

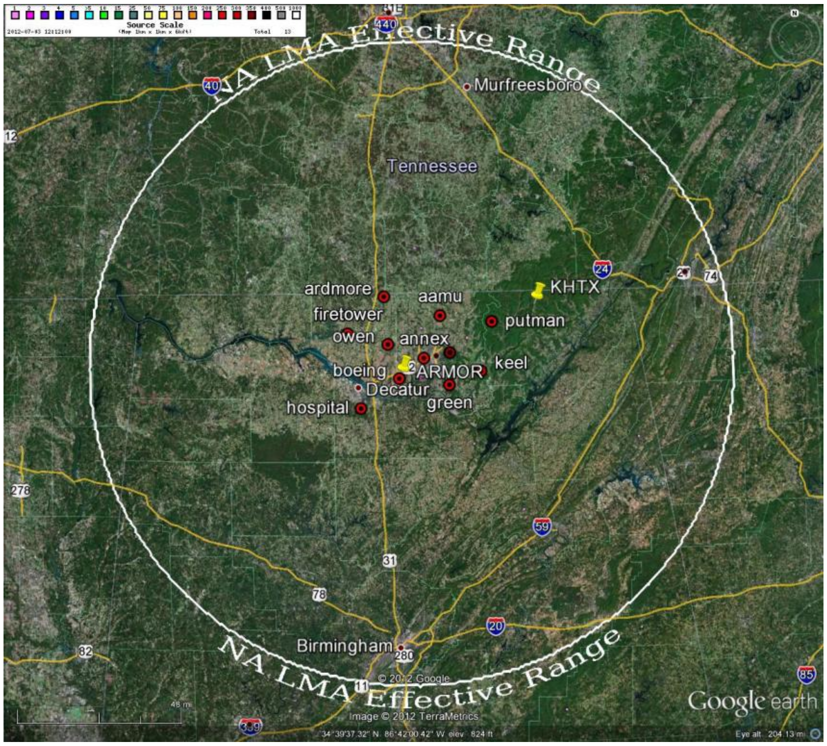

Figure 1. NALMA network sensor locations and effective range in Northern Alabama. Radar locations (ARMOR and KHTX) are shown for comparison.

Software compatibility (i.e., list of existing software to view/manipulate the data)

XLMA software package. Contact Dr. Ron Thomas, New Mexico Tech, to obtain the software.
Email: thomas@nmt.edu

6.0 References:

Bruning, Eric 2012: Gridded Flash Products using lmatools. https://bitbucket.org/deeplycloudy/lmatools

Koshak et al. 2004, JTECH, Vol .21, 543-558

Krehbiel et al. 2000, EOS, Vol 81, No. 3, 21-32

McCaul, E. W., J. Bailey, J. Hall, S. J. Goodman, R. Blakeslee, and D. E. Buechler, 2005: A flash clustering algorithm for North Alabama Lightning Mapping Array data. Preprints, Conf. on Meteorological Applications of Lightning Data, San Diego, CA, Amer. Meteor. Soc., 5.2. [Available online at http://ams.confex.com/ams/Annual2005/techprogram/paper_84373.htm.]

——, S. J. Goodman, K. M. LaCasse, and D. J. Cecil, 2009: Forecasting lightning threat using cloud-resolving model simulations. Wea. Forecasting, 24, 709–729.

Rison et al. 1999, GRL, VOL. 26, NO. 23, PAGES 3573-3576, DECEMBER 1, 1999

Thomas et al. 2001, GRL, VOL. 28, NO. 1, PAGES 143-146, JANUARY 1, 2001

Thomas et al. 2004, JGR, VOL. 109, D14207, doi:10.1029/2004JD004549, 2004