Universal Lightning Data Format

Ron Thomas
New Mexico Tech

- Goals
- Format evolution
- Examples
- Requirements
- Header format
- Data format
- Sumary

Goals

- Make Lightning mapping data easy to share.
- Make applications and tools easy to share.
- Data files should be self documenting
  - Header
  - Data
- Data and Header should be both people and machine readable
- Format must be flexible and dynamic

Format Evolution

- In 1998 we used a binary format.
  - No information in file on date.
  - With several version it was confusing for us.
  - Sharing data required our software.
- In 2000 at STEPS began an ASCII format.
  - Header with necessary information.
  - Compressed for archived files similar in size to binary data.
  - Standard data in .dat – other data in .aux files
- Format still evolving

Universal Format Requirements

- Dynamic and Self-Documenting
- ASCII
- Header
  - Should contain all information needed to use the data
  - Machine and human readable
  - Must be correct and complete
- Data
  - One line per LMA located event
  - White space delimited

Example of STEPS Data

New Mexico Tech's Lightning Mapping System -- Analyzed Data
Data start time: 06/04/00 00:20:00
Number of seconds analyzed: 600
Location: STEPS
Analysis program: /usr/local/nfs/bin/analysis_v5
Analysis started : Fri Feb 23 06:09:21 2001
Analysis finished: Fri Feb 23 06:11:51 2001
Number of active stations: 11
Active stations: B C D E G H I J K L N
Data: time (UT sec of day), lat, lon, alt(m), reduced chi^2, # of stations contributed
Data format: 15.9f 10.6f 10.6f 7.1f 5.2f 2d
Number of events: 18282
*** data ***

1200.317573156	39.864453	-101.482293	9899.3	0.22	7
1200.318110481	39.865432	-101.481430	9987.6	0.13	9
1200.319533869	39.865844	-101.480565	10392.5	0.11	7
1200.321055870	39.865133	-101.481050	10140.6	0.15	8
1200.322498469	39.866089	-101.481565	9999.0	0.05	8
1200.323015223	39.865779	-101.482934	9886.8	0.25	7
1200.323313556	39.866065	-101.483332	9795.7	0.44	6
1200.324001501	39.866260	-101.480819	9777.6	0.18	10

Example of STEPS Data

Key words end with a :

```
New Mexico Tech's Lightning Mapping System -- Analyzed Data
Data start time: 06/04/00 00:20:00
Number of seconds analyzed:  600
Location: STEPS
Analysis program: /usr/local/nfs/bin/analysis_v5
Analysis started : Fri Feb 23 06:09:21 2001
Analysis finished: Fri Feb 23 06:11:51 2001
Number of active stations: 11
Active stations: B C D E G H I J K L N
Data: time (UT sec of day), lat, lon, alt(m), reduced chi^2, # of stations contributed
Data format: 15.9f 10.6f 10.6f 7.1f 5.2f 2d
Number of events:       18282
*** data ***
  1200.317573156  39.864453 -101.482293  9899.3  0.22  7
  1200.318110481  39.865432 -101.481430  9987.6  0.13  9
  1200.319533869  39.865844 -101.480565 10392.5  0.11  7
  1200.321055870  39.865133 -101.481050 10140.6  0.15  8
  1200.322498469  39.866089 -101.481565  9999.0  0.05  8
```

Date and time of data

List of data and units

One line per event

*** Separates header and data

Time, Lat, Long,Alt,Chi-square, number of stations

Example of North Alabama Data

New Mexico Tech's Lightning Mapping System -- Ananlyzed Data
Data start time: 10/26/03 09:00:00
Number of seconds analyzed: 3600
Location: North Alabama
Analysis program: analysis_v6 -n1
Analysis started : Wed Mar 10 14:55:25 2004
Analysis finished: Wed Mar 10 15:46:00 2004
Number of active stations: 10
Active stations: A B C D E F G H I J
Data: time (UT sec of day), lat, lon, alt(m), reduced chi^2, # of stations contributed
Data format: 15.9f 10.6f 10.6f 7.1f 5.2f 2d
Number of events: 37807
*** data ***

32402.675876139	31.462247	-88.643320	16194.1	1.44	01ea	20
32402.688570792	31.576022	-88.571686	27600.5	0.37	01de	22
32407.302014723	34.889736	-86.560429	10794.2	1.96	0276	-3
32409.872435733	34.653678	-86.813101	12154.4	0.16	00ed	-13
32416.521024918	33.226298	-88.156660	7448.2	1.05	01ea	17
32416.527998168	33.235596	-88.152357	8416.0	0.29	0176	17
32416.560868966	33.213007	-88.134395	4049.5	0.67	01ee	17
32416.562985492	33.249044	-88.112230	3824.7	0.94	01fe	18

Improved data by adding two new fields.
What are they?

Data: and format: information are not correct and don't agree with data

Edited North Alabama Data

Created new key word

New Mexico Tech's Lightning Mapping System -- Ananlyzed Data
Data start time: 10/26/03 09:00:00
Number of seconds analyzed: 3600
Location: North Alabama
Analysis program: analysis_v6 -n1
Analysis started : Wed Mar 10 14:55:25 2004
Analysis finished: Wed Mar 10 15:46:00 2004
Number of active stations: 10
Active stations: A B C D E F G H I J
Station mask: first station in active station list is least significant digit
Data: time (UT sec of day), lat, lon, alt(m), reduced chi^2, station mask(hex), power(dBW)
Data format: 15.8f 10.4f 10.4f 7i 5.1f 4z 4i
Number of events: 37807
*** data ***

Define last two fields with units and format

```
32402.67587614  31.46225 -88.64332 16194  1.4 1ea  20
32402.68857079  31.57602 -88.57169 27601  0.4 1de  22
```

Can reduce number of digits to match precision 10 ns and 1 m. Reduces file size by 12%.

```
32402.675876139  31.462247 -88.643320 16194.1  1.44 01ea  20
32402.688570792  31.576022 -88.571686 27600.5  0.37 01de  22
```

Data before

Information to be included

- Data - required
  - Time, Latitude, Longitude, Altitude, Chi-square
- Data – highly useful
  - Station mask, source power (STEPS has this information in an auxiliary file)
- Data – optional from normal processing (for STEPS this was in an auxillary file)
  - Raw powers, above thresholds (median or raw)
  - Raw times used
  - Error ellipse
- Data – optional from further processing (XLMA output)
  - Flash number, charge identification, etc.

Header Information

- Documentation of data
- Useful Additional Information
  - Raw data version, data window length
  - Station Location (needed for finding the error ellipse)
  - Station Hardware Information
  - Station: “A”, Bob's place, 39.71528N, -101.4537E, 1076 m elevation, 137 ns delay, NMT-rev3, data version 7, 80 us time window.

Draft specifications: Lightning data format

- **Header**
  - **First line is a title**
  - **All other lines begin with a keyword**
  - **Keywords**
    - **end with a colon (:)**
    - **Must include these keywords**
      - Data start time: Location: Data: Data format: Number of events:
    - **New keywords can be added as needed**
    - **Use standard keywords if possible**
- **Separate data and header by a line beginning with ***
- **Data**
  - **One line per event -- as many fields as necessary – all line in same format**
  - **Required – time, lat, long, altitude, chi-squared**
  - **Fields are white space delimited**
  - **Number of lines equals the number of events indicated in the header**
  - **Other data can follow events – information on flashes, etc.**

Summary

- Easy to read and share for humans
  - ASCII file
  - Header with complete and accurate information
- Machine readable
  - Consistent key words
- Efficient storage
  - Files can be compressed
  - Limit number of digits to needed precision