GOES-R AWG Lightning Team: Lightning Cluster Filter Algorithm (LCFA)
June 8, 2010

Presented By: Douglas Mach[1]
[1] University of Alabama in Huntsville

The Lightning Team Members

- **Co-Chairs:**
  - » Steven Goodman
  - » William Koshak (NASA/MSFC-AWG)
  - » Richard Blakeslee (NASA/MSFC-R3)
- **Application Team:**
  - » Douglas Mach (UAH-AWG)
  - » Walter Petersen (NASA/MSFC-R3)
  - » Dennis Buechler (UAH-R3)
  - » Larry Carey (UAH-R3)
  - » Monte Bateman (USRA-AWG)
  - » Eugene "Bill" McCaul (USRA-R3)
- **AWG Application Team (cont):**
  - » Donald MacGorman (OAR/NSSL-R3)
  - » Henry Fuelberg (GRA FSU-R3)
  - » Eric Bruning (CICS-AWG,R3)*
  - » Rachel Albrecht (CICS-R3)*

    (* Left project in early 2010)
- **Product Development Team:**
  - » STAR:
    - – Shanna Sampson
    - – Hua Xie
    - – Lihang Zhou
    - – Walter Wolf

Outline

- Executive Summary
- Algorithm Description
- ADEB and IV&V Response Summary
- Requirements
- Validation Strategy
- Validation Results
- Future Algorithm Improvements
- Summary

Executive Summary

- The GLM Lightning Cluster Filter (LCFA) generates the Option 1 products of Events, Groups, and Flashes along with QA bits and the indexing needed to reconstruct the flash and its group and event components.
- Version 5 will be delivered July 16, 2010. ATBD (100%) is on track for a June 16, 2010 delivery.
- An algorithm, based on Lightning Imaging Sensor (LIS) and Optical Transient Detector (OTD) heritage, has been developed that can process the peak expected Level 1b data rates within specified latency of 4 seconds.
- Validation proxy data sets have been developed and used to test the GLM LCFA.
- Our analyses indicate spec compliance for all products.

Requirements

Lightning Detection

Name	User & Priority	Product Geographic Coverage	Vertical Resolution	Horizontal Resolution	Mapping Accuracy	Measurement Range	Measurement Accuracy	Product Refresh Rate /Coverage Time	VAGL	Measurement Precision	Temporal Coverage Qualifiers	Product Extent Qualifier	Cloud Cover Conditions Qualifier	Product Statistics Qualifier
Lightning Detection -Events -Groups -Flashes	GOES -R	Full Disk	Sfc to Cloud Top.	10 km	5 km	Real Time	70% minimum Flash Detection Efficiency (FDE)	Continuous	20 sec	5% (Std. Dev. of FDE)	Day and night	Quantitative out to at least 65 degrees LZA and qualitative beyond	Cloud cover conditions permitting obs. of lightning associated with threshold accuracy	Over lightning cases and mesoscale-sized surrounding regions

Algorithm Description

Algorithm Summary

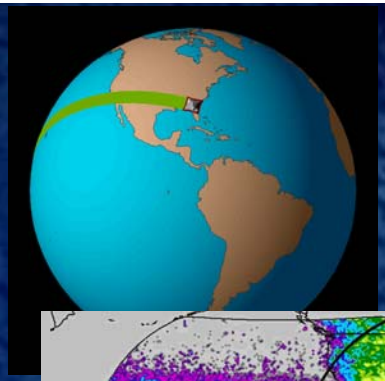
- The GLM Lightning Cluster Filter (LCFA) generates the Option 1 products of Events, Groups, and Flashes along with the QA bits and sindexing needed to reconstruct the flash and its group and event components.
- It uses enhanced LIS/OTD heritage code to cluster the input Level 1b data into events, groups, and flashes within the latency requirements.
- We use a unique “cheat” to speed clustering of new into multiple “active” flash regions spread across the GLM FOV to increase code efficiency and decrease latency (horseshoe_handgrenade).
- We use a “full fit” clustering system where all events, groups, and flashes that CAN be clustered ARE clustered.
- New events can result in the combination of previously separate groups and new groups can result in the combination of previously separate flashes.
- The LCFA is designed to run in a continuous mode as the GLM data will be continuous (the GLM is NOT an imager)
- The LCFA code has “hooks” for adding “filter” code to identifying any events, groups, or flashes (or parts thereof) that may be from non-lightning sources that were not filtered by the L0 to L1b algorithm.
- The LCFA code is capable of being parallelized if needed to meet required latency (currently not implemented).

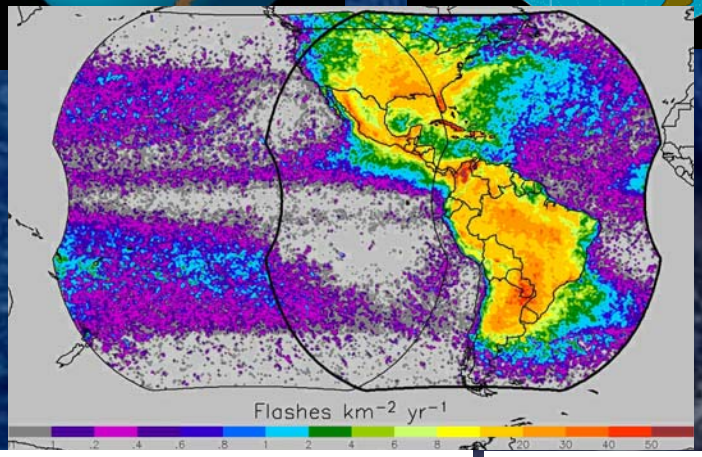
Motivation for Algorithm Selection

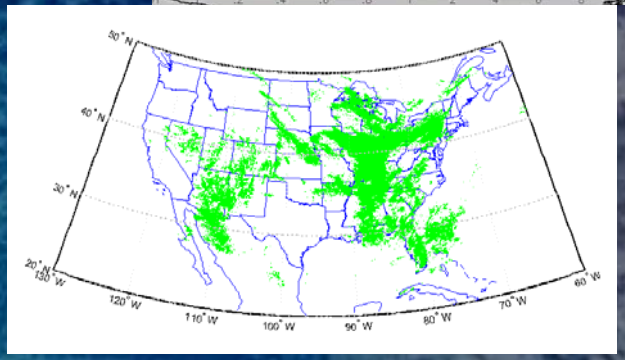
- The LCFA represents an enhancement of the proven heritage algorithm used in the LIS and OTD orbital lightning detectors.
  - » Basic algorithm has been used successfully for over a decade to process orbital lightning data into events, groups, and flashes.
  - » “Hooks” added to code anticipating post Level 1b “filtering” are based on our experience with prior instruments.
  - » Enhancements are based on our experiences with prior implementations of the heritage algorithm to “real world” instruments.
  - » No other algorithms were considered because there are no other algorithms

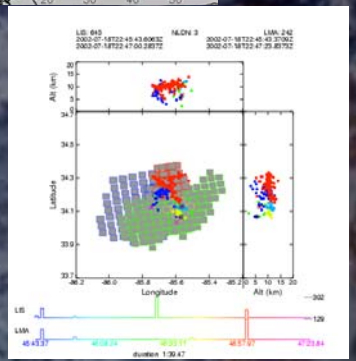
Expected GLM Performance Relative to Other Lightning Sensors

- Orbital Sensors
  - » GLM vs. OTD – greater sensitivity, greater FOV, better SNR, similar spectral band, similar pixel size, both detect cloud and ground lightning, similar time resolution
  - » GLM vs. LIS – greater sensitivity, greater FOV, better SNR, similar spectral band, larger pixel size, both detect cloud and ground lightning, similar time resolution
- Ground Sensors
  - » GLM vs. LDAR – much larger FOV, optical vs. VHF lightning detection, LDAR has greater time & spatial resolution, 2D vs. 3D
  - » GLM vs. NLDN – greater sensitivity, larger FOV, optical vs. VLF lightning detection, NLDN mostly detects ground flashes, similar time resolution









LCFA Processing Flowchart Clustering Details (1)

A. Grab Next Event In Buffer (LCFA processes one event at a time, possibly parallel processing opportunity)

B. Update active Groups & Flashes (using event time, determine which groups & flashes are now "DEAD")

C. Find Parent Group (look through list of active groups for one or more than this event can be clustered with)

   a. If more than 1 – goto F

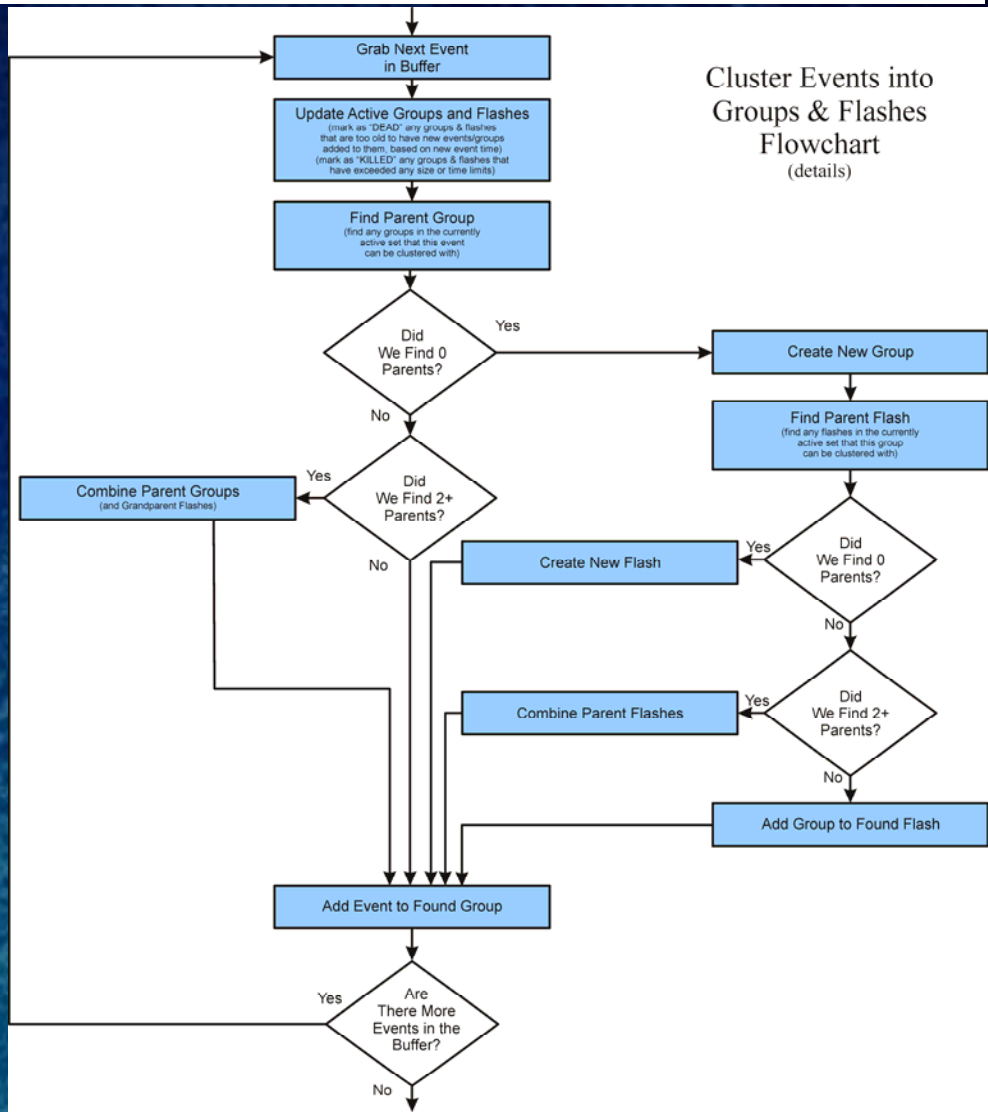
   b. If less than 1 – goto H

D. Add Event to the Group (update group & parent flash statistics)

E. Go To Top of Loop (A) (if there are more events in the buffer, otherwise exit section)

F. Combine All Groups That Can Be A Parent To The Event (and combine grandparent flashes if the groups have different parents)

G. Go To D (above)



LCFA Processing Flowchart Clustering Details (2)

H. Create New Group Based On Event

I. Find Parent Flash of New Group (same concept as C prior slide)

   a. If more than 1 – goto L

   b. If less than 1 – goto N

J. Add Group to Parent Flash

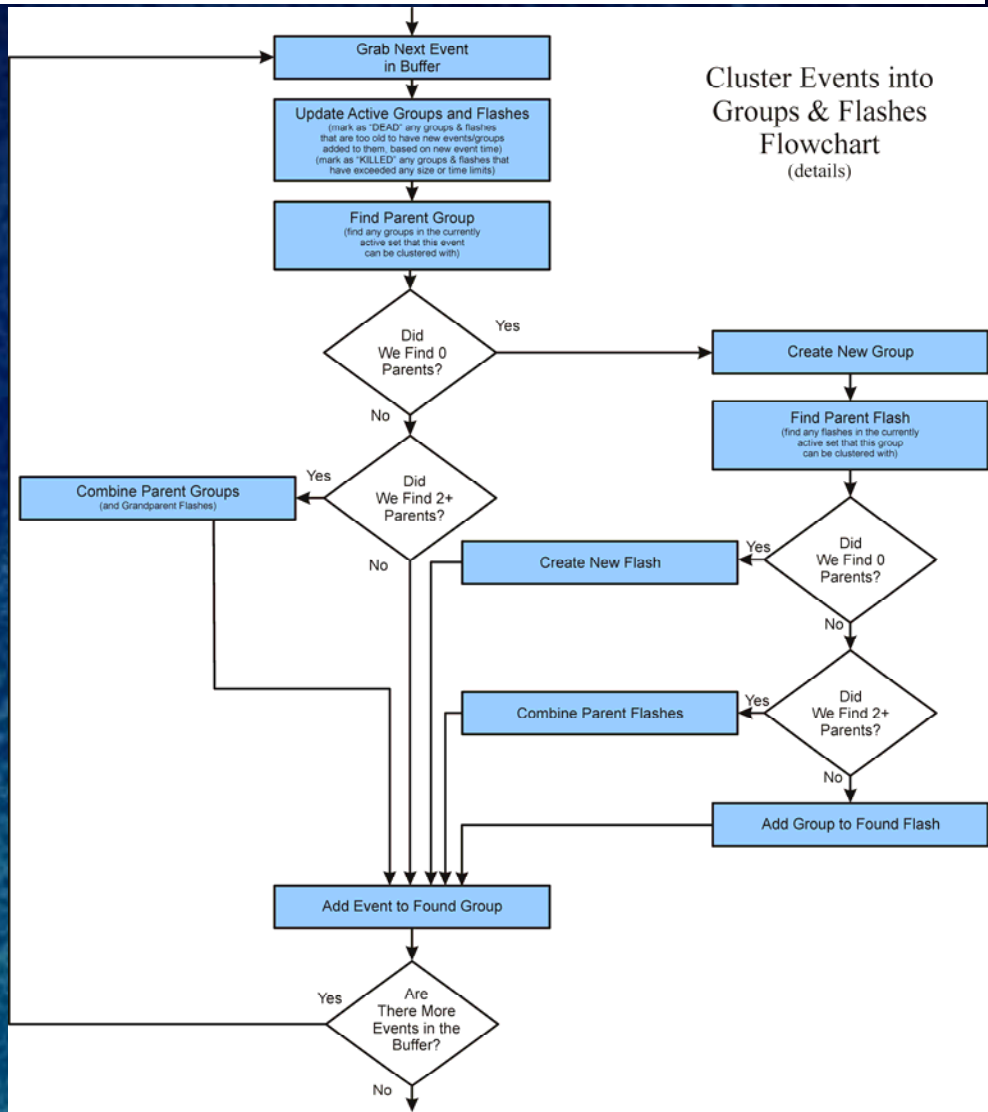
K. Go To D (prior slide)

L. Combine All Flashes That Can Be A Parent To The Group (same concept as F prior slide)

M. Go To D (prior slide)

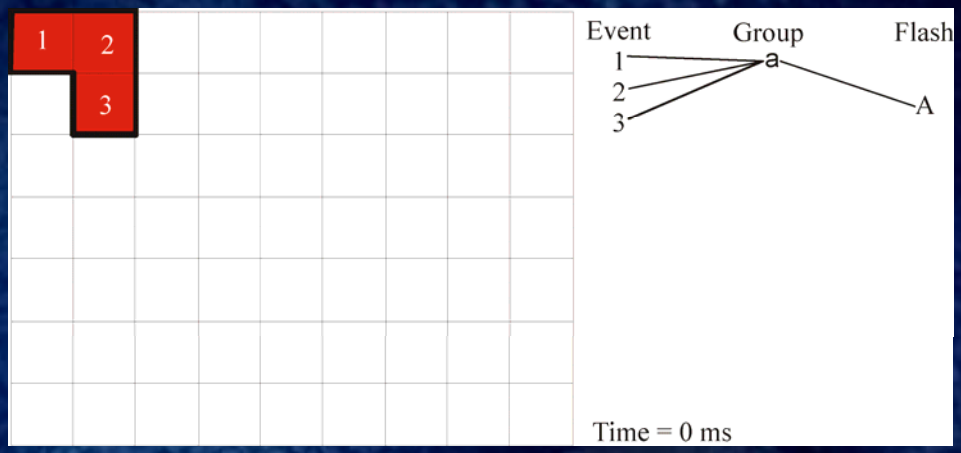
N. Create New Flash Based On Group

O. Go To D (prior slide)

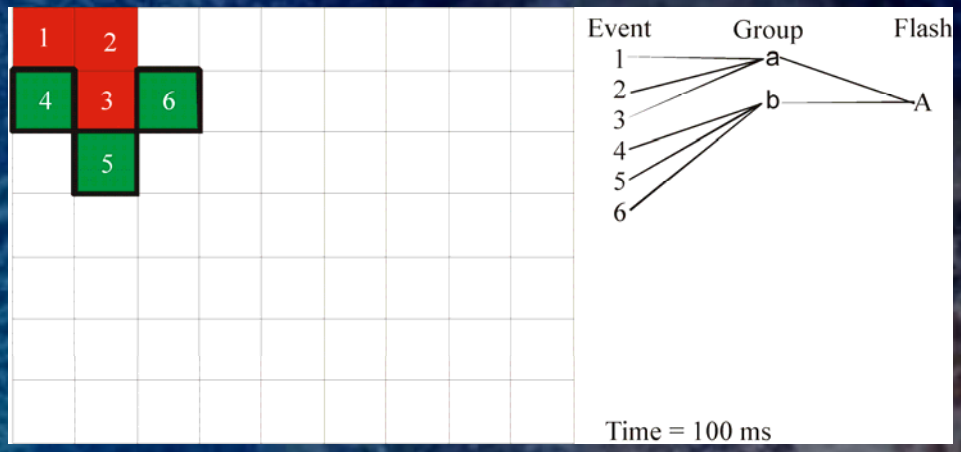


Example LCFA Output (1)

- **Time = 0 ms**: The first time integration is shown to the right. Three (**1, 2, 3**) events occur at this time integration. Since the events are simultaneous and register in adjacent (i.e., neighboring or diagonal) pixels, they are collected into a single group (**a**). The group is assigned a new parent flash (**A**). Note that the flash information is not moved to the output stream/files until the flash is completed (DEAD).

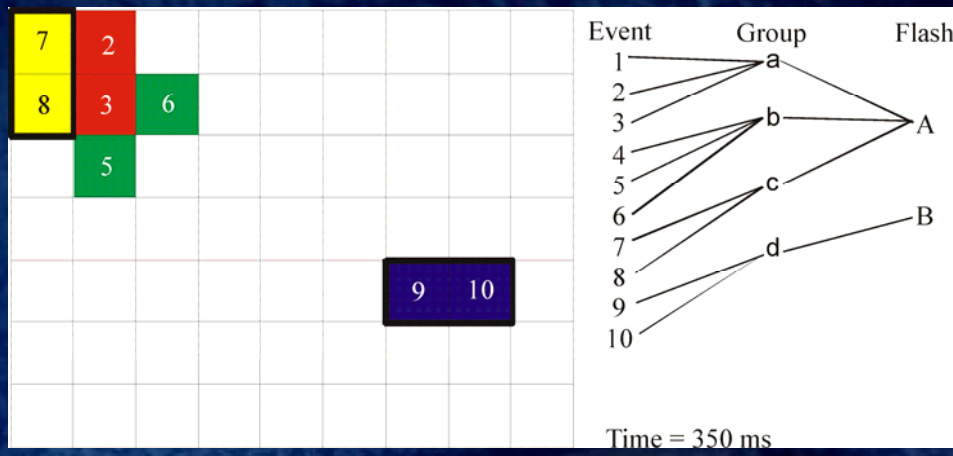


- **Time = 100 ms**: The next time integration with data is shown to the right. At this time (100 ms after the first one), there are three more events (**4, 5, 6**). As in the previous case, these three new events are all assigned to a new group (**b**). These events are not assigned to group **a** since they occur at a different time. Since group b is within 16.5 km of group a (actually, they touch), and the groups occur within 330 ms of each other, group **b** is assigned to flash **A**.

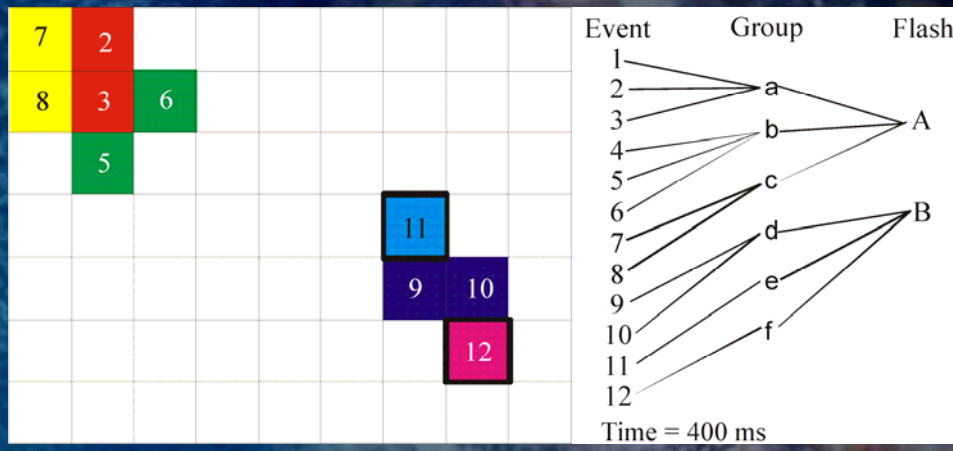


Example LCFA Output (2)

- **Time = 350 ms:** The next integration time with data is shown. The time is 350 ms after the time of the first events, but only 250 ms after the time of the last events. At this time there are four (**7, 8, 9, 10**) more events. Events **7** and **8** are adjacent to each other and are assigned to a new group (**c**). Events **9** and **10** are not adjacent to events **7** and **8**, but are adjacent to each other. They are assigned to another new group (**d**). Since group **c** is within 330 ms of the last group of flash **A** (250 ms) and is also within 16.5 km of the parts of flash **A**, group **c** is assigned to flash **A**. Although group **d** also occurred within 330 ms of the last group of flash **A**, it is greater than 16.5 km away from any part of flash A so it is assigned to a new flash (**B**).

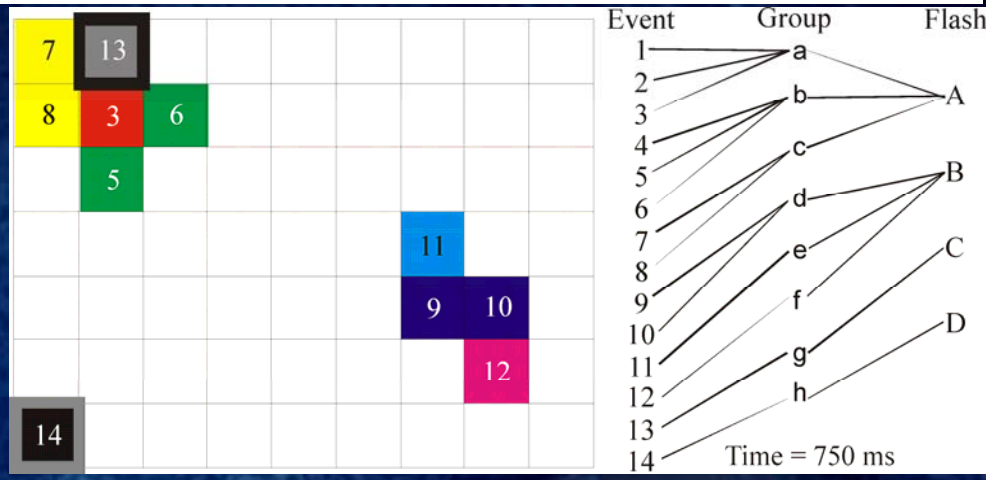


- **Time = 400 ms:** The time is 400 ms after the first events and 50 ms after the latest events. Two more events occur (**11, 12**) at this time. These two events are at the same time, but they are not adjacent to each other. They are assigned to two new groups (**e** for **11** and **f** for **12**). The two new groups are less than 330 ms (50 ms) from the time of the last group of flash **B** and are within 16.5 km (adjacent) of the parts of flash **B** so the two groups are assigned to flash **B**.



Example LCFA Output (3)

- **Time = 750 ms:** The last time with events (for this example) is shown. At this time integration, 750 ms after the first events and 350 ms after the last events, there are two new events (**13, 14**). The events are not adjacent, so they are assigned to two new groups (**g** for **13** and **h** for **14**). Group **g** overlaps the parts of flash **A**, however, it has now been 400 ms (greater than 330 ms) since the last group associated with flash **A**. Therefore, group **g** is assigned to a new flash (**C**). Group **h** is not within 16.5 km of any current flash, so it is assigned another new flash (**D**). Since flashes **A** and **B** have now 'expired', their summary statistics are sent to the output streams and files as bandwidth and processing power permits. 330 ms after the last events for flashes **C** and **D**, they are also sent to the output streams/files.



Algorithm Changes from 80% to 100%

- Changing output format to reflect bits/widths of GRB output.
- Metadata output will be added.
- Quality flags will be finalized.
- “Fallback” code will be added.

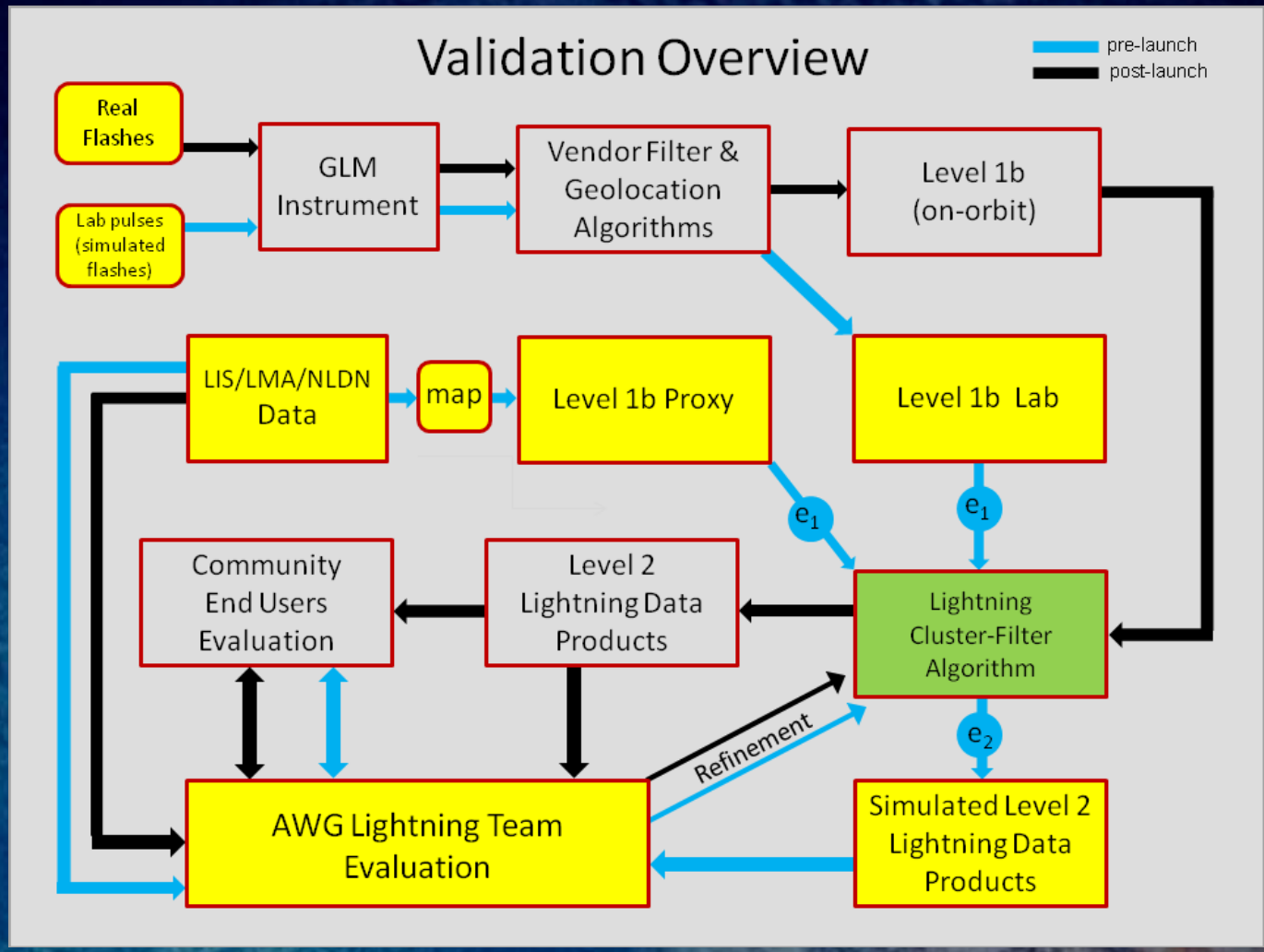
ADEB and IV&V Response Summary

- All ATBD errors and clarification requests have been addressed.
- No feedback required modifications to the approach.

Validation Approach

Validation Approach (1)

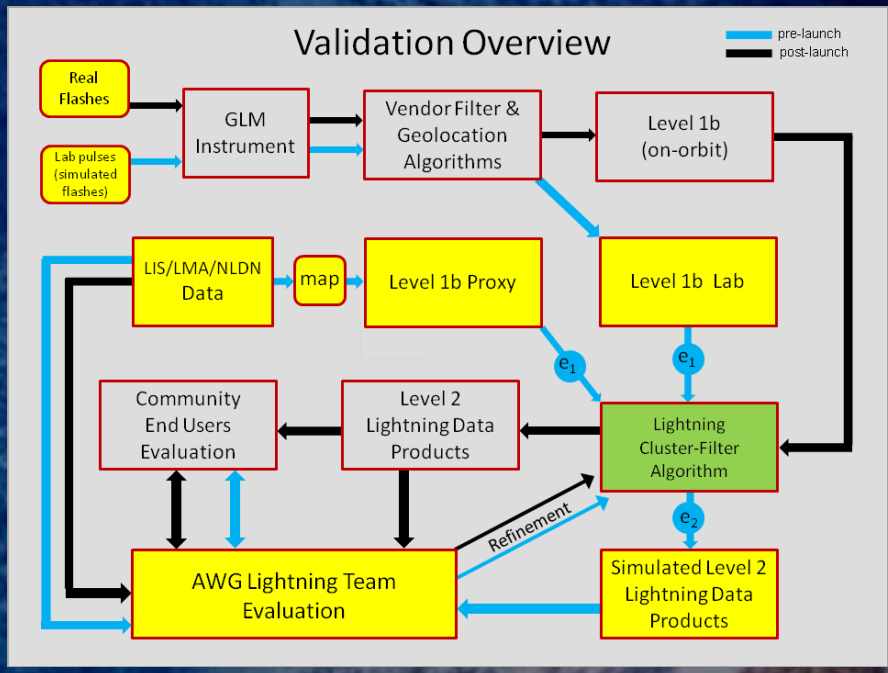
- The approach consists of two different phases:
  - » pre-launch evaluation phase given by the blue arrows
  - » post-launch evaluation phase given by the black arrows.



Validation Approach (2)

Pre-launch Algorithm Evaluation Phase

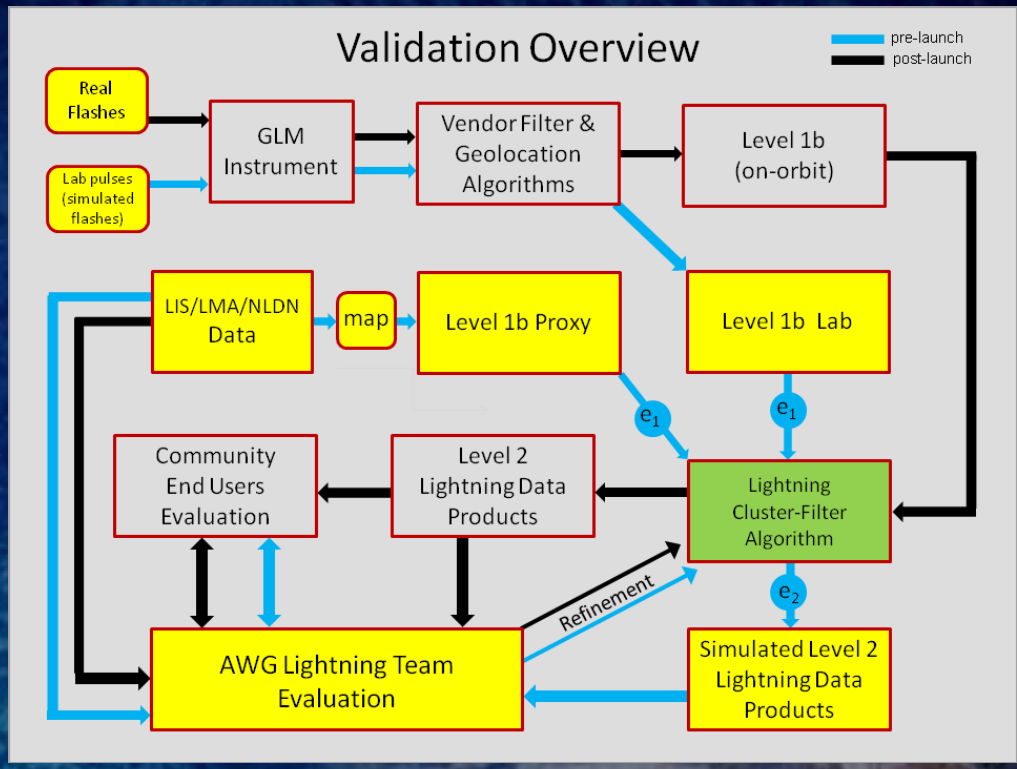
- Stage 1: Proxy Data Development
  - » Use LIS, LMA, OTD, and NLDN data to develop empirical mapping to GLM Level 1b proxy
  - » Use above results to develop “fake but realistic” GLM L1b proxy data to test code limits
  - » Add errors to proxy to test code exception handling
  - » Create “fake and unrealistic” GLM L1b proxy data to test code limits
- Stage 2: Instrument Calibration Testing
  - » GLM illuminated with optical pulses in a controlled laboratory environment
  - » Optical pulse characteristics are carefully chosen to simulate as closely as feasible the expected cloud-top optical emissions produced from real lightning flashes
  - » GLM instrument output is a simulated Level 0 output
  - » Vendor filtration & geolocation algorithms are used on the output to produce the “Level 1b Lab” data
  - » Level 1b Lab data is then tested like the Level 1b Proxy data (see Stage 1)



Validation Approach (3)

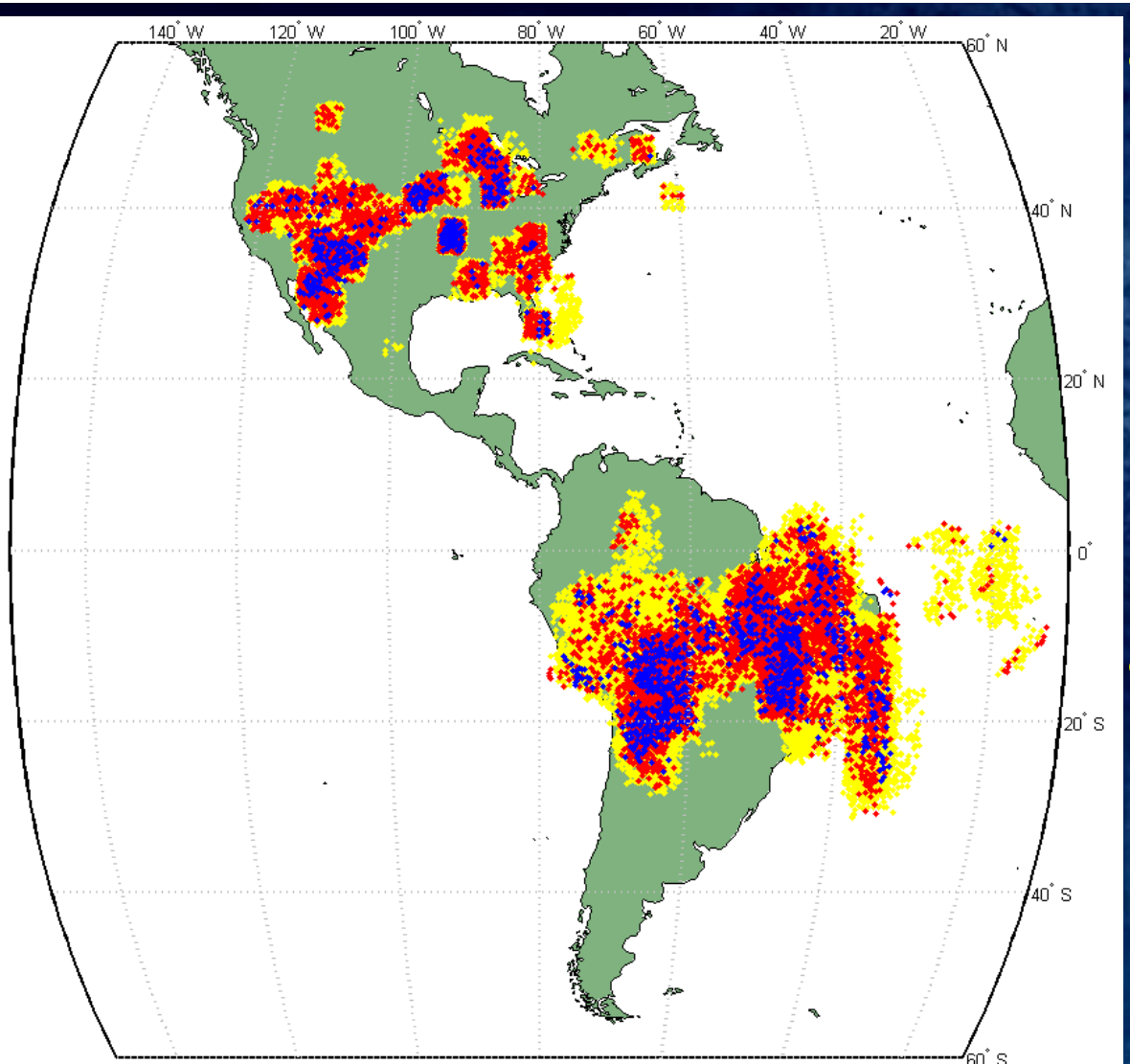
Post-Launch Algorithm Evaluation Phase

- Actual GLM data will be evaluated by other sensors
  - » LIS (historical)
  - » LMA
  - » NLDN
  - » Radar
  - » ΔE
  - » etc.
- This time the evaluation will be based on comparisons with actual Level 2 Data rather than simulated Level 2 Data



Validation Results

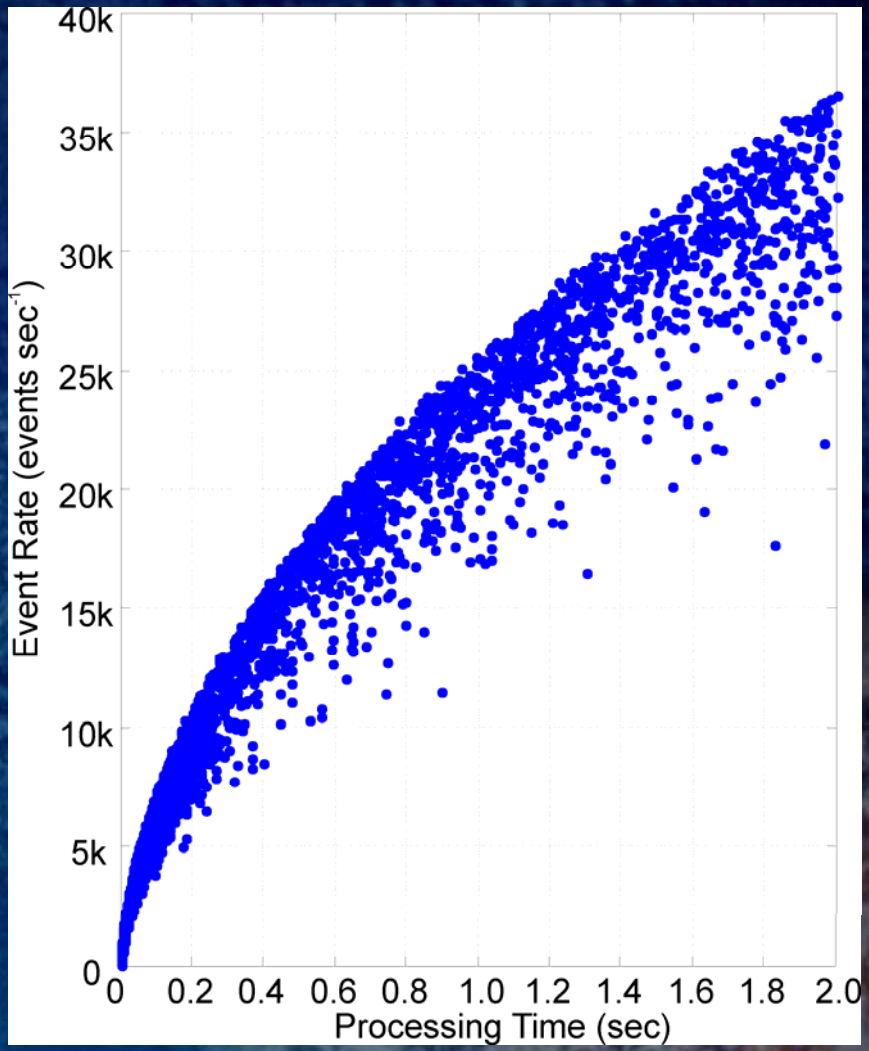
GLM LCFA Validation Results (1)



- NLDN to Seed Artificial Flashes to GLM proxy to GLM Algorithm Output
  - » yellow = all flashes
  - » red = 1/10th of all flashes
  - » blue = 1/100th of all flashes
- Totals:
  - » 1.0×10^6 events
  - » 4.4×10^5 groups
  - » 1.3×10^4 flashes

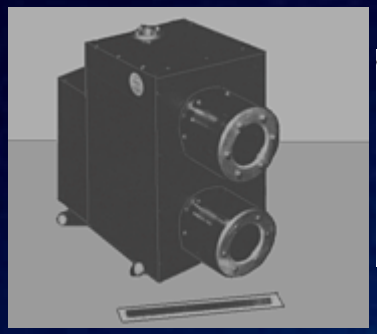
GLM LCFA Validation Results (2)

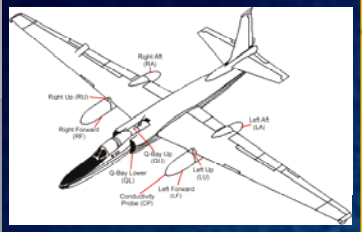
- Proxy data tests on single thread GLM code on target system indicate that the code can process 20,000 to 25,000 events in less than 1 second of computer time
- Current estimates of maximum GLM data rates are around 15,000 events/sec (mean rates nearer 150 events/sec)
- We can process a nominal second of GLM event data in less than one second of computer time

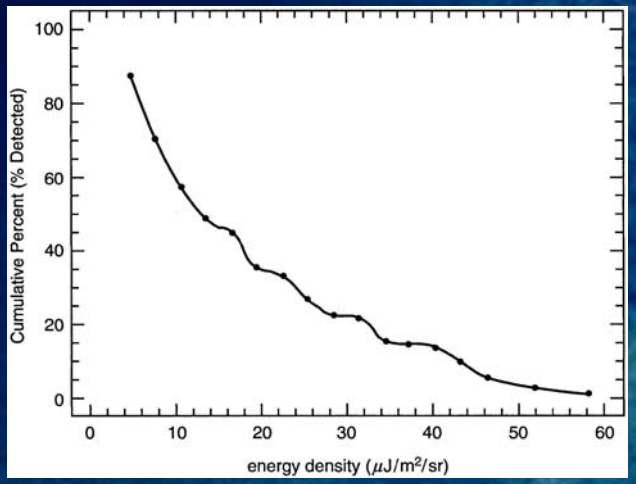


GLM LCFA Validation Results (3)

U2 Aircraft Optical Pulse Data



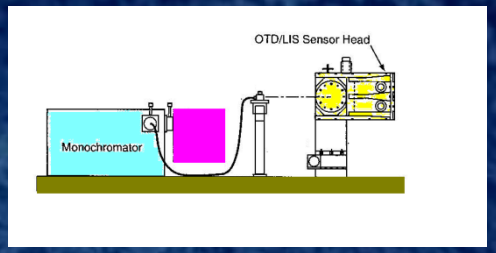




(Christian & Goodman , 1987)
(Goodman et al., 1988)

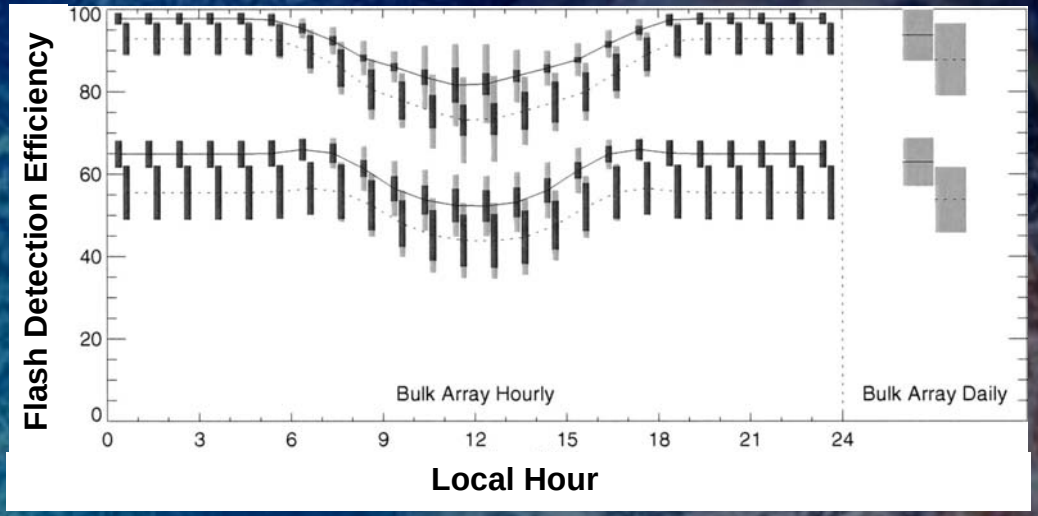
Synthesize

OTD/LIS Cal Data & Instrument Specs



(Koshak et al., 2000)

Flash Detection Efficiency (Boccippio et al., 2002)

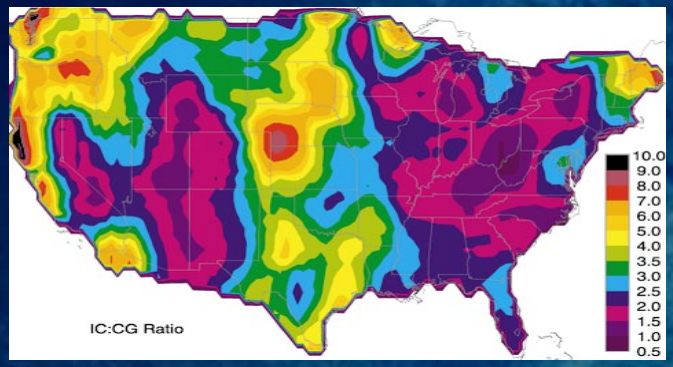


GLM LCFA Validation Results (4)

OTD

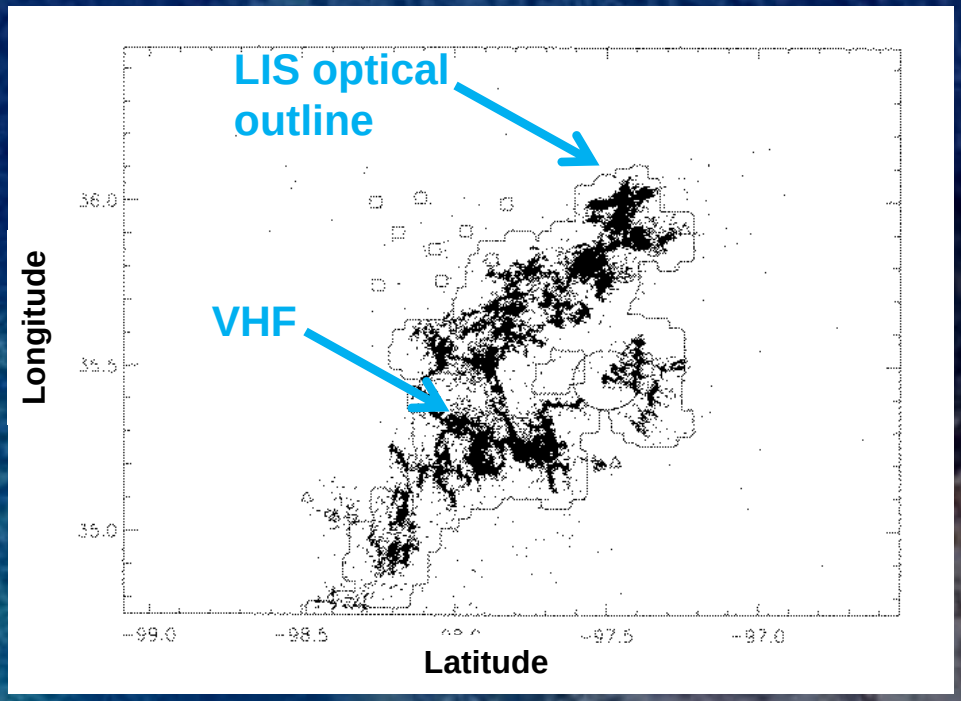
NLDN

combine



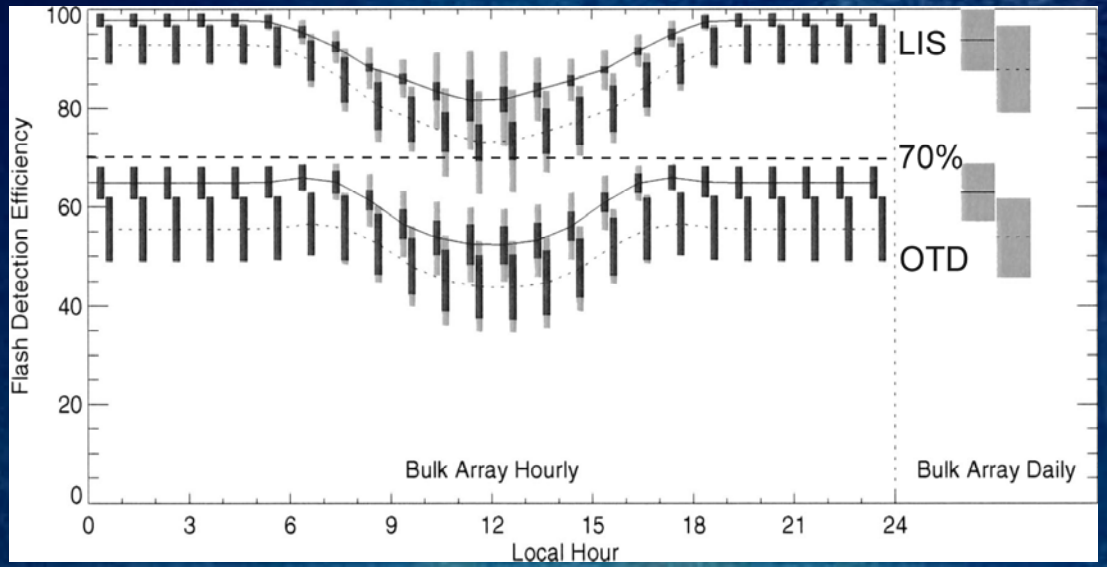
IC:CG Ratio (Boccippio et al., 2001)

VHF (OKLMA) Compared to LIS Optical (Thomas et al., 2000)



GLM LCFA Validation Results Summary

- Latency for L2 Events, Groups & Flashes
  - » Input buffer (1 sec latency)
  - » 1 second to process (0-1 sec latency)
  - » Data output by flash end (0-1 sec latency)
  - » Output buffer (1 sec latency)
  - » Total of 2-4 seconds of latency for GLM data processing
- Flash detection efficiency based on prior instruments study
  - » GLM efficiency will be greater than LIS/OTD due to system improvements



Flash Detection Efficiency (Boccippio et al., 2002)

Future Algorithm Improvements

- None anticipated at this time

- Possible ideas:
  - » Use cloud height to refine parallax correction
  - » Use cloud mask as a “sanity check” on lightning sources
  - » Extend groups in time (so “touching” will include subsequent time frames)
  - » Work up parallel processing version (separate out ingest, filter, cluster, and output into separate threads)

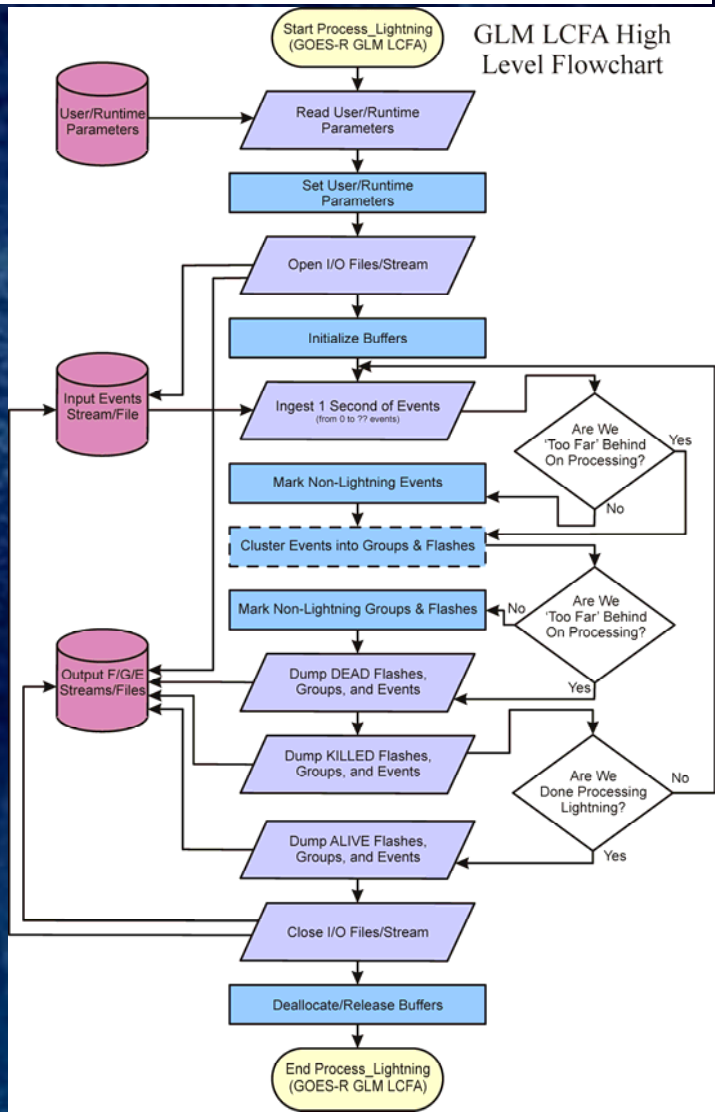
Summary

- The GLM Lightning Cluster Filter Algorithm provides a heritage algorithm based solution to the clustering of the Level 1b GLM data into Level 2+ products.
- Version 5 (7/16/2110) and the 100% ATBD (6/16/2010) will be delivered.
- The products meet specifications.

Extra Slides
(more details)

LCFA Processing Flowchart Overview

- Read User Parameters (limits, buffer sizes, etc.)
- Set User Parameters
- Open I/O Files/Streams (data to/from files or streams)
- Initialize Buffers (active flashes, groups, events, I/O buffers, etc.)
- Ingest 1 Second of L1b Data (for I/O efficiency)
- Mark Non-Lightning Events (events that got past the L0-L1b filters), (will still be clustered)
- Cluster Events into Groups & Flashes (the meat of the algorithm, will get its separate slide)
- Mark Non-Lightning Groups & Flashes (with extra knowledge that clustering provides, mark other items as "non-lightning)
- Dump DEAD Flashes/Groups/Events (flashes and their children & grandchildren that have naturally expired)
- Dump KILLED Flashes/Groups/Events (flashes and their children & grandchildren that were terminated for cause)
- Close I/O Files/Streams
- Deallocate/Release Buffers (clean up after itself)



Example LCFA Output

- **Summary Data:** In the example data processing sequence just described, there were fourteen events, eight groups, and four flashes (see Tables below). This example shows how the GLM algorithm will convert events into groups and flashes. The full list of what will be output is listed in the Table to the right. During the GLM mission, the start-time of the flashes (the time of the first event in the flash) will be a relative time (seconds from mission epoch, currently 1/1/2000). The group (first event time of the group) and event times will be in GLM frames (2 ms) from the start of the parent flash to minimize data transmission bandwidth.

Resultant Flash Data

flash_id	start_time	delta	event_	latlon_	child_	Child_ids
A	0	350	8	6	3	a, b, c
B	350	50	4	4	3	d, e, f
C	750	0	1	1	1	G
D	750	0	1	1	1	H

Resultant Group Data

group_ id	parent_id	group_	event_	latlon_	child_	Child_ids
a	A	0	3	3	3	1, 2, 3
b	A	100	3	3	3	4, 5, 6
c	A	350	2	2	2	7, 8
d	B	350	2	2	2	9, 10
e	B	400	1	1	1	11
f	B	400	1	1	1	12
g	C	750	1	1	1	13
h	D	750	1	1	1	14

Flash Data L2

bits	Description
56	uint - time, 24 bit Day, 32 bit μs/day, epoch 1/1/2000 (time of first event in flash)
31	int - lat/lon x 100 per ea 16 bit long (signed), 15 bit lat (signed)
13	uint - # of GLM frames (~2 ms) in flash (max 16.4 s)
20	uint – footprint, scaled log value
15	uint - flash optical energy density ($\mu Jm^{-2}sr^{-1}\mu m^{-1}$)
10	uint - group count-1 (max 1025)
10	uint - first group index
13	QA bits
168	21 Bytes

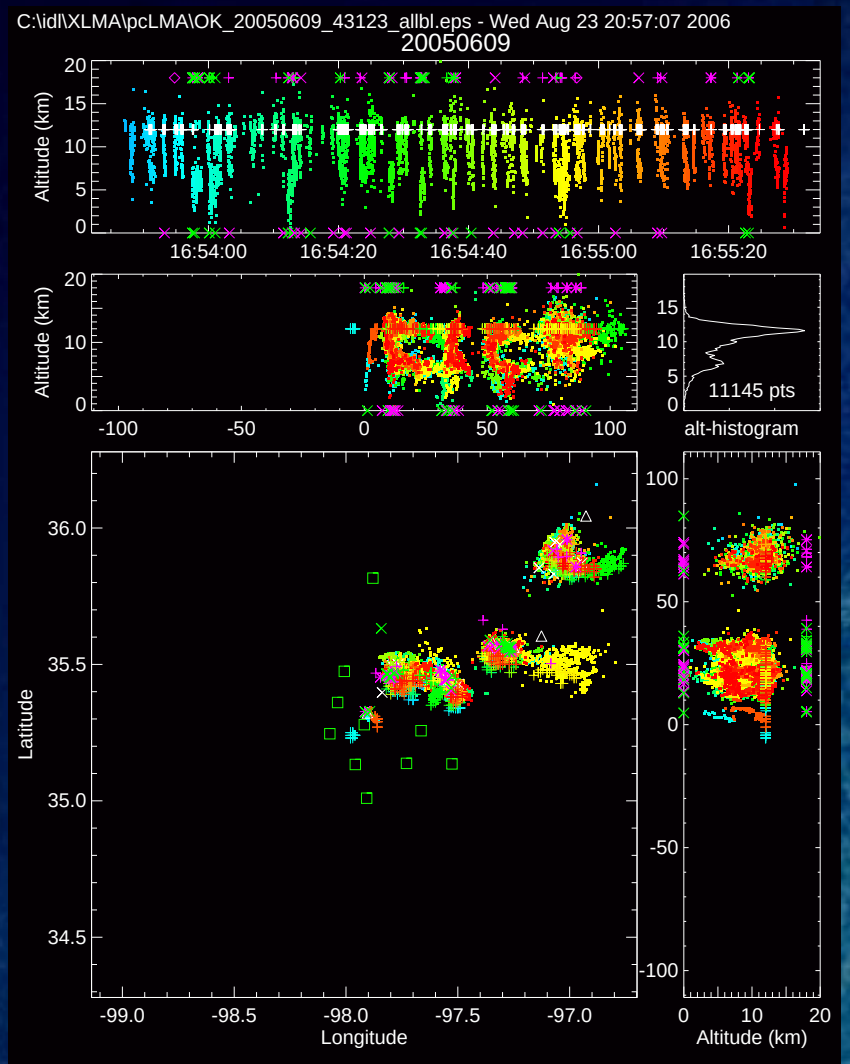
Group Data L2

bits	Description
13	uint - delta GLM frame (~2 ms) count from the start of parent flash to this group
31	int - lat/lon x 100 per ea 16 bit long (signed), 15 bit lat (signed)
11	uint - # of GLM frames (~2 ms) in group (max 4.1 s)
19	uint - footprint, scaled log value
14	uint - group optical energy density ($\mu Jm^{-2}sr^{-1}\mu m^{-1}$)
9	uint – parent flash index
11	uint - event count-1 (max 2049)
11	uint - first event index
9	QA bits
128	16 Bytes

Event Data L1b

bits	Description
13	uint - delta GLM frames (~2 ms) from the start of the grandparent flash to this event
31	int - lat/lon x 100 per ea 16 bit long (signed), 15 bit lat (signed)
12	uint - event optical energy density ($\mu Jm^{-2}sr^{-1}\mu m^{-1}$)
10	uint - parent group index
6	QA bits
72	9 Bytes

GLM LCFA Validation Results (extra)



LIS Data Orbit 43123 6/09/2005

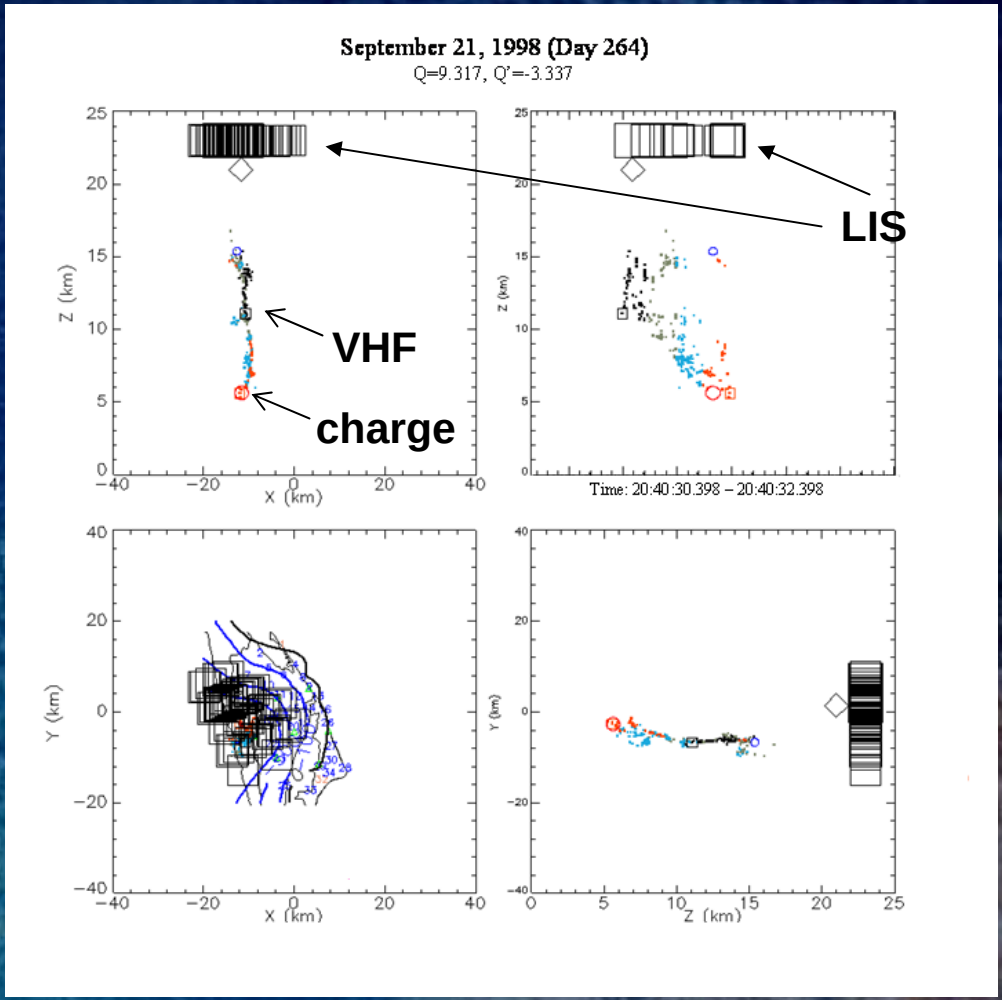
- LIS 73 flashes
- 1215 groups
- 3843 events
- LMA 843 flashes
- 11145 LMA events
- 162 LASA
- 40 NLDN

451 co-incident LIS-LMA

GLM LCFA Validation Results (extra)

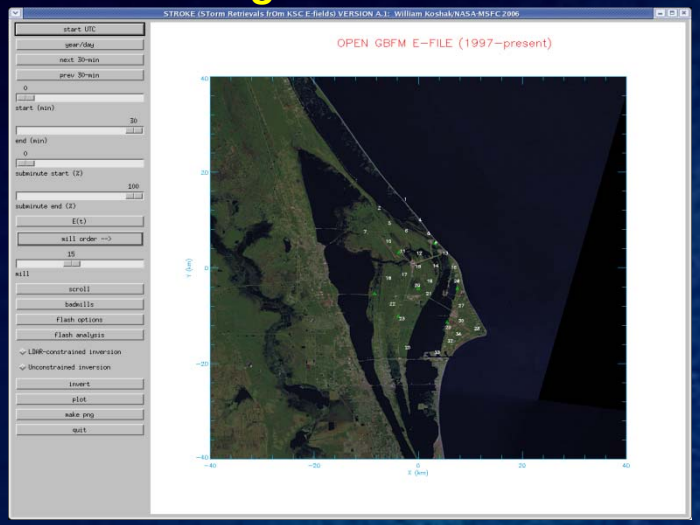
- Data in plots:
  - » LIS events (squares)
  - » LIS flash centroid (diamonds)
  - » Charge (circles)
  - » VHF (color dots)
- Able to co-locate LIS, charge data, and VHF mapping data to validate LIS
- Same validation data/technique for GLM data

(Koshak et al., 2007)

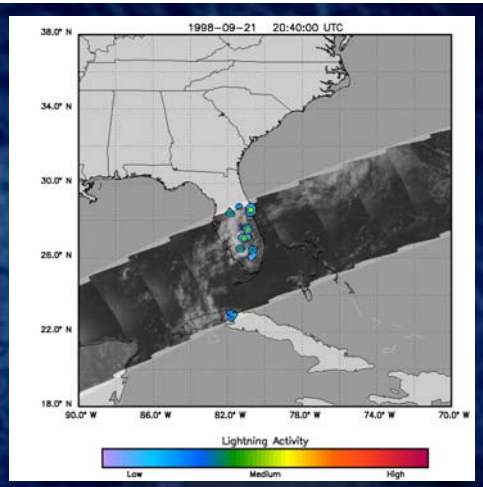


GLM LCFA Validation Results (extra)

MSFC Charge Retrieval Software



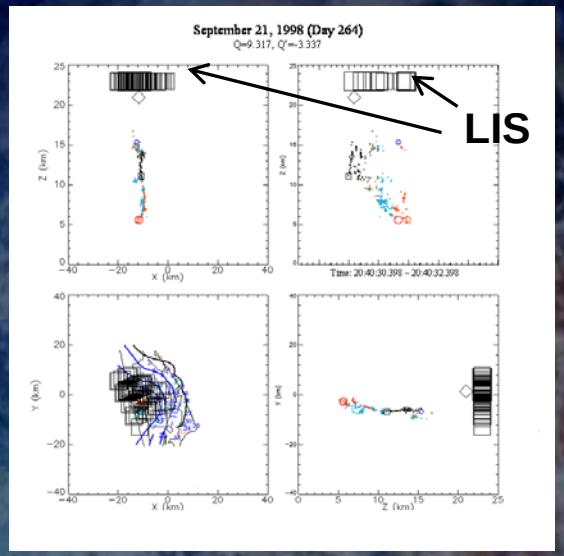
LIS Overpasses



(Koshak et al., 2007)

Field Mill

LDAR

Over-plots:

- LIS
- Charge
- VHF