

GOES-17 GLM Level 2 (Events, Groups, Flashes) Release  
Provisional Data Quality  
December 20, 2018  
Read-Me for Data Users

The GOES-R Peer/Stakeholder Product Validation Review (PS-PVR) for GLM L2 Provisional Maturity was held on December 20, 2018. As a result of this review, the PS-PVR panel chair declared the data was of Provisional validation maturity.

The GLM L2 product consists of geo-located and time-stamped *events, groups, and flashes*, with associated calibrated optical amplitudes (in units of Joules).

Provisional maturity, by definition, means that:

- Validation activities are ongoing, and the general research community is now encouraged to participate;
- Severe algorithm anomalies are identified and under analysis. Solutions to anomalies are in development and testing.
- Incremental product improvements may still be occurring;
- Product performance has been demonstrated through analysis of a small number of independent measurements from select locations, periods, and associated ground-truth/field program efforts;
- Product analysis are sufficient to communicate product performance to users relative to expectations;
- Documentation of product performance exists that includes recommended remediation strategies for all anomalies and weaknesses. Any algorithm changes associated with severe anomalies have been documented, implemented, tested, and shared with the user community;
- Testing has been fully documented;
- Products ready for operational use & for use in comprehensive cal/val activities and product optimization.

Provisional users bear all responsibility for inspecting the data prior to use and for the manner in which the data are utilized. Persons desiring to use the GOES-17 GLM Provisional-maturity L2 products for any reason, including but not limited to scientific and technical investigations, are encouraged to consult the GLM Calibration/Validation science team members for feasibility of the planned applications.

Commentary on important GLM data characteristics and known issues being resolved include:

**1. Flash Detection:** GOES-17 GLM flash detection efficiency (FDE) over the field-of-view and over a 24 hour period averaged 71% at the West park position (137.2 W longitude), and depends on the ground-network reference data employed; i.e., comparisons with just GLD360 ground reference showed an average FDE of 88%. Adjustments of the Ground Segment RTEP settings on 3 Dec 2018 (Algorithm Discrepancy Report, ADR 855) have led to these improved FDE values. Insufficient lightning sample size was available (i.e. due to Northern Hemisphere Winter analysis period) to evaluate extent of potential FDE depletion in the NW CONUS region. In addition, there is evidence that the FDE is substantially smaller in anomalous (i.e.

inverted polarity) storms, and in severe (e.g., hail-producing) storms, or storms with deep liquid water path. In general, because the FDE associated with reference data is itself variable and typically below 100%, it is not always possible to exactly/unambiguously determine the GLM FDE in all cases.

2. **False Events:** False events due to blooming, created from glint and/or solar intrusion, are still present but a blooming filter, and improvements to glint processing, have been developed, tested, and are planned for implementation next year. Adjustments to the 2<sup>nd</sup> level threshold filter on 3 Dec 2018 (ADR 855) have muted a hot pixel region over the South Pacific. Owing to the improvements to ground processing software in preparation for GOES-16 GLM Provisional Maturity, false events due to high energy radiation particles, aka "radiation dots" have been reduced by removing Single Group Flashes (SGFs), but future adjustments to this filter will be tested so as to mitigate removal of legitimate flashes; e.g. a change to instead remove Single Event Flashes (SEFs) is planned. Moreover, a formal "put back" filter is being tested to improve retention of legitimate flashes.

3. **Position Errors:** GOES-17 GLM flash location errors are on par with latest GOES-16 GLM location error analyses. Parallax errors near the edges of the field-of-view are a persistent feature, but improvements on 3 Dec 2018 (ADR 855) have already been achieved by optimizing the equator and pole parameters of the cloud-top "lightning ellipsoid" model parameters. Long-term plans are to replace the static "lightning ellipsoid" with monthly 3 degree resolution cloud-top surfaces that have, in detailed tests, proven to further mitigate parallax. Mitigation of INR inaccuracies due to diurnal variations has not yet been achieved, but is in analysis.

4. **Timestamp:** There are still time-order issues with L2 data event times (as well as group time, flash start time, and flash end time), but these issues normally do not pose any major problems for most analyses; a fix is planned for next year.

5. **Family Links:** Family linkages refer to the correspondence between events, groups and flashes. For example, a "childless" group is a group with no events, and an "orphan" event is an event with no parent group. All downward family links have been fixed (effective when the DO.07.00.00 software version went operational on October 15, 2018). Upward family links will not be done, due to bandwidth constraints.

6. **Unsigned Integer Read:** In order to save storage space, some floating-point variables (such as times, latitude, and longitude) are stored in the GLM NetCDF file as a lower resolution internal format with a "scale\_factor" and "add\_offset" attribute. Some of the GLM data is stored in a non-standard format (as unsigned integers). This is an issue that affects multiple instruments on GOES-16/17, and a pilot fix was worked via ADR 844 with implementation on November 5, 2018. The classic model for NetCDF does not support internal storage of unsigned integers larger than 8 bits. Despite this, many of the variables in the GOES-16/17 data are stored internally in the NetCDF files as either 16-bits or 32-bits unsigned integers. A future version of NetCDF (CDF5) will have options for internal storage of unsigned integers, but the GOES program does not use that version (CDF5). To get around this lack of unsigned integer support, the GOES program added a non-standard attribute "\_Unsigned", to designate which variables are stored internally as unsigned. Unfortunately, with the DO.07 version of the code, some variables that are stored as

unsigned integers are not marked as such with the “\_Unsigned” attribute (this is a bug in the code). So, until a fix is achieved (for both the non-standard “\_Unsigned” attribute and the totally unmarked unsigned integers), we recommend using the low level NetCDF readers (the ones that read the NATIVE file formats) and use the following process to convert from the native format to the actual data format:

Retrieve the variable data (using low level routines): For MOST of the variables that should be read as unsigned integers, there is an attribute “\_Unsigned” for that variable. The internal variable should then be read and cast to an unsigned integer of either 16 or 32 bits (depending on the size of the internal variable). This step must be completed before applying scale\_factor and add\_offset values to convert from scaled unsigned integer to science units.

Exceptions to the Rule: As it turns out, not all unsigned internal variables are actually marked with the attribute “\_Unsigned”. We have designated these variables as “stealth unsigned variables” (SUVs). There is no physical way to determine which of the remaining variables are SUVs. The list below indicated the current set (the ones we have found) that are SUVs in DO.07 and will need to be converted like the variables that have the “\_Unsigned” attribute (note, each of these):

Variable	Meaning	When Added
event_time_offset	GLM L2+ Lightning Detection: event’s time of occurrence	DO.01.00.00
group_time_offset	GLM L2+ Lightning Detection: mean time of group occurrence actually marked with the attribute	DO.01.00.00
flash_time_offset_of_first_event	GLM L2+ Lightning Detection: time of occurrence of first constituent event in flash	DO.01.00.00
flash_time_offset_of_last_event	GLM L2+ Lightning Detection: time of occurrence of last constituent event in flash	DO.01.00.00
group_frame_time_offset	GLM L2+ Lightning Detection: mean frame time of group’s constituent events’ times of occurrence	DO.07.00.00
flash_frame_time_offset_of_first_event	GLM L2+ Lightning Detection: frame time of occurrence of first constituent event in flash	DO.07.00.00

flash_frame_time_offset_of_last_event	GLM L2+ Lightning Detection: frame time of occurrence of last constituent event in flash	DO.07.00.00
---------------------------------------	---	-------------

As an aside, the times in the prior versions of the output (prior to DO.07.00.00) have been changed from milliseconds to seconds. Although this change is actually documented, it is still a major change (and might require changes to the reading routines).

*How to Determine If Data is DO.07.00.00 or a Prior Version:* There are no external indications in the file that the format has changed. One way to determine which version of the file you are working with is to count the number of variables in the file. If it is 45, it is the OLD format, will not have the SUVs (but will still have the “\_Unsigned” designation for the non-standard unsigned integers). If the variable count is 48, it is the DO.07.00.00 version of the file and will have the list of SUVs (above) along with the “\_Unsigned” designation for the non-standard unsigned integers.

**7. Gridded Data & Data Quality Products:** These products are not yet available but are being developed, with plans for eventual submission to the Ground Segment via a formal Algorithm Discrepancy Report (ADR) along with fully tested meta-code.

**8. Data Recommended:** The Calibration Working Group (CWG) recommends that data on or after 4 December 2018 is best to use because this follows the software fix for ADR 855 (mentioned several times above) that was implemented on 3 December 2018.

Contact for further information: OSPO User Services at [SPSD.UserServices@noaa.gov](mailto:SPSD.UserServices@noaa.gov)

GLM contacts for specific information on the GLM L2 data:

Bill Koshak [william.koshak@noaa.gov](mailto:william.koshak@noaa.gov)

Scott Rudlosky [scott.rudlosky@noaa.gov](mailto:scott.rudlosky@noaa.gov)