LI L2 Product User Guide [LIL2PUG]

Doc.No. : EUM/GEO/TEN/15/828715
Issue : v2D e-signed
Date : 6 May 2022
WBS/DBS : MTG-831420

EUMETSAT
Eumetsat-Allee 1, D-64295 Darmstadt, Germany
Tel: +49 6151 807-7
Fax: +49 6151 807 555
http://www.eumetsat.int

© EUMETSAT
The copyright of this document is the property of EUMETSAT.

Document Signature Table

	Name	Function	Signature	Date
Book Captain:	Katja Hungershoefer Bartolomeo Viticchie	IFCT Engineer LI Scientist		
Reviewed by:	Fabian Müller	MTG LI Instrument Engineer		
Reviewed by:	Claude Ledez	IFCE Team Leader		
Reviewed by:	Maria Jose Marquez	Quality & Product Assurance Manager		
Approved by:	Susanne Dieterle	MTG System Manager		

Co-authors/Contributors

Name	Organisation
Stephan Bojinski, Corvin-Petrut Cobarzan, Sven-Erik Enno, Jochen Grandell, Rory Hutson, Fabian Müller, Heikki Pohjola, Timothy Patterson, Bartolomeo Viticchie	EUMETSAT

Distribution List

Distribution list	
Name	***No. of Copies***
EUMETSAT MTG Team	Electronic Distribution

Document Change Record

Issue / Revision	*Date*	*DCN. No*	*Changed Pages / Paragraphs*
1 Draft	13/10/2015	-	First draft
1 Draft	25/02/2016	-	Updates to Chapter 3 and Chapter 4
1A Draft	04/04/2016	-	Updated to new template Added information including high level information from LIL2PS and LIL2FS Added CDL format descriptions. General updates. Made read-only for review
1B Draft	25/04/2016	-	Added Section 6.7 with information about reference grid Added Section 7: Product Usage Made read-only for review
2A Draft	30/07/2021	-	Updated version for review by LI MAG Members
2B Draft	22/03/2022	-	Updated version following the review of the LI MAG Members
2C	21/04/2022	-	• Update document structure in Section 1.4 • Added Section 7.6 on how time is stored in general and additional information about the accumulation_start_times in the accumulated products • Updated grid parameters in Table 11 • Information on the FCI L1c reference grid added in Section 8.1 • Completed the section LI Level 2 Product Usage • Fixed broken bookmarks
2D	06/05/2022		Version to be published and shared with the users

Table of Contents

Table of Figures

Table of Tables

1 INTRODUCTION

1.1 Scope

This document is the Product User Guide for LI Level 2 products. This release is a draft version to accompany the release of an LI Level 2 test data package. The document is based on our current best knowledge of the LI instrument functionality and characteristics, data processing, and output format. Any additional evolutions in this knowledge up to the launch of the first MTG Imaging platform and during the commissioning phase, will lead to updates in future releases of this document. In addition, some areas of the document are currently incomplete. These will be detailed and expanded in subsequent releases.

1.2 Acronyms and Definitions

Abbreviation/Term	Meaning
AF	Absolute Factor
CC	Cloud to Cloud
CG	Cloud to Ground
DE	Detection Efficiency
EURD	End Users Requirements Document
ENVISAT	Environmental Satellite
FDE	Flash Detection Efficiency
FOV	Field of View
FT	False Transient
GLM	Geostationary Lightning Mapper
IC	Intra-Cloud
L1b	Level 1b
L2	Level 2
LI	Lightning Imager
LOH	LI Optical Head
MERIS	Medium Resolution Imaging Specrometer

MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
MTG-I	Meteosat Third Generation Imager Satellite
MTG-S	Meteosat Third Generation Sounder Satellite
NWC	Nowcasting
NWP	Numerical Weather Prediction
OC	Optical Channel
PICs	Pseudo-Invariant Calibration Sites
SRD	System Requirement Document
SSD	Spatial Sample Distance
TT	True Transient
WED	Weighted Euclidian Distance

1.3 Applicable and Reference Documents

1.3.1 Reference Documents

Acronym	Reference Number	Title
[WMO-386]	Available from https://public.wmo.int/en/resources/library	WMO Manual on the Global Telecommunication System-Volume I. 2009 Edition.
[CF]	http://cfconventions.org/	CF Conventions Document
[NACDD]	https://geo-ide.noaa.gov/wiki/index.php?title=NetCDF_Attribute_Convention_for_Dataset_Discovery	NetCDF Attribute Convention for Dataset Discovery
[MACH07]	Mach, D., Christian, H., Blakeslee, R., Boccippio, D., Goodman, S., Boeck, W., J. Geophys. Res., Vol. 112, 2007	Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor

Acronym	Reference Number	Title
[LIL2ATBD]	EUM/MTG/DOC/11/0155 v6 available online.	Algorithm Theoretical Basis Document (ATBD) for Level 2 processing of the MTG Lightning Imager data
[FEGS19]	Quick, M. G., 2019. Sub-flash Comparison of FEGS and GLM Observation from GOES-R Flight Campaign. Presentation at the 2019 GLM Annual Science Team Meeting, September 10-12, 2019, Huntsville, AL. Available online.	Sub-flash Comparison of FEGS and GLM Observation from GOES-R Flight Campaign
[KOS10]	Koshak, W. J., 2010. Journal of Atmospheric and Oceanic Technology, 27 (11), 1822–1838. https://doi.org/10.1175/2010JTECHA1405.1	Optical Characteristics of OTD Flashes and the Implications for Flash-Type Discrimination

1.4 Document Structure

The sections of this document present the following information

Section 1 An overview of the document

Section 2 A brief introduction to the MTG programme, the MTG platform, and the on-board instruments

Section 3 The Lightning Imager instrument, detection principle, and useful references to the documents describing the Level 0 and Level 1b processing

Section 4 The description of the LI Level 2 processing algorithms

Section 5 The description of the LI Level 2 products disseminated to the users

Section 6 Information about the file naming convention

Section 7 Characteristics of the LI Level 2 product

Section 8 Useful information about the product usage

Section 9 Information about the quicklooks available in the archive

Appendix A A detailed look at the NetCDF formats including complete format descriptions

2 METEOSAT THIRD GENERATION (MTG)

2.1 The MTG Programme

The Meteosat Third Generation (MTG) programme provides meteorological data over Europe and Africa and maintains continuity of the Meteosat programme, continuing and expanding the service provided by Meteosat Second Generation (MSG).

2.2 The MTG Platform

MTG is a twin satellite concept based on 3-axis stabilised platforms. The twin satellites comprise an imaging satellite, MTG-I, and a sounding satellite, MTG-S. Four imaging and two sounding satellites are planned.

The MTG-I payload comprises:
- The Flexible Combined Imager (FCI)
- The Lightning Imager (LI)
- The Data Collection System (DCS)
- Search and Rescue (GEOSAR)

The MTG-S payload comprises:
- The Infrared Sounder (IRS)
- The Sentinel-4 Ultra-violet, Visible and Near-infrared Sounder (UVN)

Twice a year, around each equinox, a yaw flip manoeuvre is performed in order to keep the same face of the satellite protected from direct Sun illumination. During the manoeuvre, the spacecraft is rotated 180 degrees around the z-axis (so inverting the spacecraft centred +/- X axes and +/- Y axes).

3 LIGHTNING IMAGER (LI)

3.1 The LI Mission

One of the new instruments selected for MTG is the Lightning Imager (LI) payload. The LI mission is intended to provide real time lightning detections of the total lighting activity, namely, cloud-to-cloud and intra-cloud (CC+IC) and cloud-to-ground (CG) lightning from a geostationary orbit. Such measurements will support Nowcasting (NWC) and Very Short Range Forecasting of severe storm hazards and lightning strike warning. As lightning is strongly correlated with storm-related phenomena like precipitation, hail and gust, the LI measurements can indicate intensive convection related to ice formation, updraft strength and convective rainfall. Lightning can also serve as a proxy for adiabatic and latent heating to be assimilated in global/mesoscale Numerical Weather Prediction (NWP) models. Finally, in the field of atmospheric chemistry, lightning plays a significant role in generating nitrogen oxides. The natural nitrogen oxide global budget is a matter of great uncertainty at this time, and long-term observations of one of its sources will prove valuable as the subject develops.
When a lightning discharge takes place, a significant portion of the emitted optical radiation is in the form of emission lines (a triplet) of the neutral atomic oxygen that are excited by the lightning channel. Those emission lines have wavelengths around 777.4 nm. Hence, observing the Earth around the 777.4 nm favours the detection of photons generated by lightning discharges, also when such photons emerge from bright clouds. One of the key elements that affects lightning detection from space is the scattering through the clouds of photons emitted along the lightning channel, which eventually reach the top of the clouds. A space-born, geostationary lightning imager observes continuously cloud tops and captures the optical pulses arising from lightning activity. In fact, the result of photon scattering in clouds is a broadening and delay of the optical pulse in time. In detail, pulses typically present a fast rise, followed by a slightly slower decay. Often, during their evolution, they have multiple peaks of emission. The most advanced and detailed description of the variation of optical emissions due to lightning has been derived with the Fly Eye GLM Simulator (FEGS; see [FEGS19]). FEGS measurements show that optical pulses have a duration of about 1 ms, with peak emissions which last about 0.5 ms. In addition, the scattering through clouds creates optical emissions at the top of clouds with footprints that extend over several kilometres and being almost always larger than 10 km (see [FEGS19]).
The main objective of the LI mission is to detect lightning flashes with an average Flash Detection Efficiency (FDE) over the LI FOV larger than 70% (see the definition of flash in Section 4.2). This will enable the identification of areas at risk of heavy convection which requires primarily information on flashes and their density in space and time.

3.2 Instrument Characteristics

3.2.1 Instrument Design

The Lightning Imager Instrument is an imaging filter spectrometer with on-board data processing for detecting lightning by measuring the lightning optical pulse signal from the top of the atmosphere. The so-called LI Optical Head (LOH) includes four identical Optical Channels (OC). Each OC is composed of a baffle for stray light, an optical system, a focal

plane assembly with a 1000×1170 pixels CMOS detector and Front End Electronics including the real-time pixel processors for the on-board processing.
Figure 1 shows how each detector (optical channel) observes a different portion of the Earth disc: the required FOV coverage for the mission is 84% of the Earth disc observable from geostationary position, and 100% of the territories of all the EUMETSAT member states. The spatial resolution of LI varies within its FOV: at sub-satellite point the sampling distance is 4.5 km, while over central Europe it is about 7 km (see Figure 1).

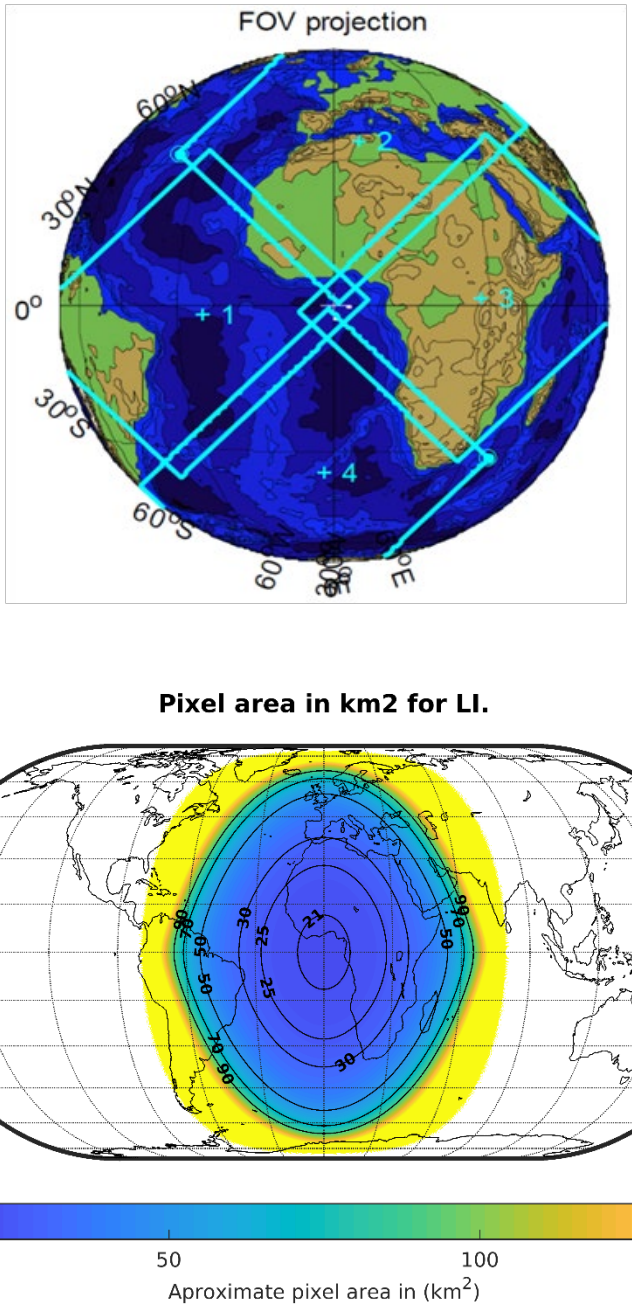


Figure 1: LI spatial coverage of the four Optical Channels (top panel) and LI pixel size (bottom panel). See also Section 7.3 for changes of the detector orientation related to the platform yaw flip.

The key design features of LI are driven by the key properties of optical pulses generated by lightning (see Section 3.1). In detail:

- **Spectral band**: narrowband filter centred over the neutral oxygen triplet at 777.4 nm with 1.9 nm bandwidth; the filter is designed to enable *i*) the selective sensing of photons from the emission of the oxygen triplet, *ii*) deal with the chromatic aberration which can impact the detection at the edges of the LOHs, and *iii*) measurements with high signal-to-noise ratio.
- **Spatial sampling**: the sampling at sub-satellite point is 4.5 km, i.e., smaller than typical dimensions of optical pulses expected from lightning activity (of the order of 10 km of diameter); it is worth stressing that over southern/central Europe the spatial sampling will be smaller than the pulse size (see Figure 1).
- **Temporal sampling**: the integration time of 1 ms period is comparable to the typical duration of pulses and twice the duration of the "peak phase" (i.e., about 0.5 ms).

The instrument can be operated in different modes including OPERATIONAL and CALIBRATION mode. In operational mode, the LI instrument is fully operating and performing all functionalities to accomplish the mission objectives (see Section 3.2.2). The instrument will be switched to CALIBRATION mode about once every six months in night-time conditions to perform routine calibration. No lightning detection is available during these calibration periods. A description of the acquisitions performed in calibration mode is available in Section 3.2.3.

3.2.2 Detection Principle and on-board Processing

The events/detections measured by the LI are identified as follows (see Figure 2) for each pixel of the four detectors:

1. the measurement of the radiance of each LI pixel is considered to be the sum of the background and the lightning signal;
2. the background of each pixel is evaluated by a configurable running average that allows the suppression of the radiometric noise (present in each single acquisition) while preserving the properties of the background scene;
3. the background evaluated in point 2 is removed from each measured radiance with the aim of retaining only the lightning signal;
4. the detection principle consists in comparing the signal derived after background removal against a detection threshold which is designed to have the LI sensitive to optical emissions that are as faint as the expected radiometric noise fluctuations for a given background;
5. those signals that are found to be above the detection threshold are identified as candidate lightning events (or detected transients).

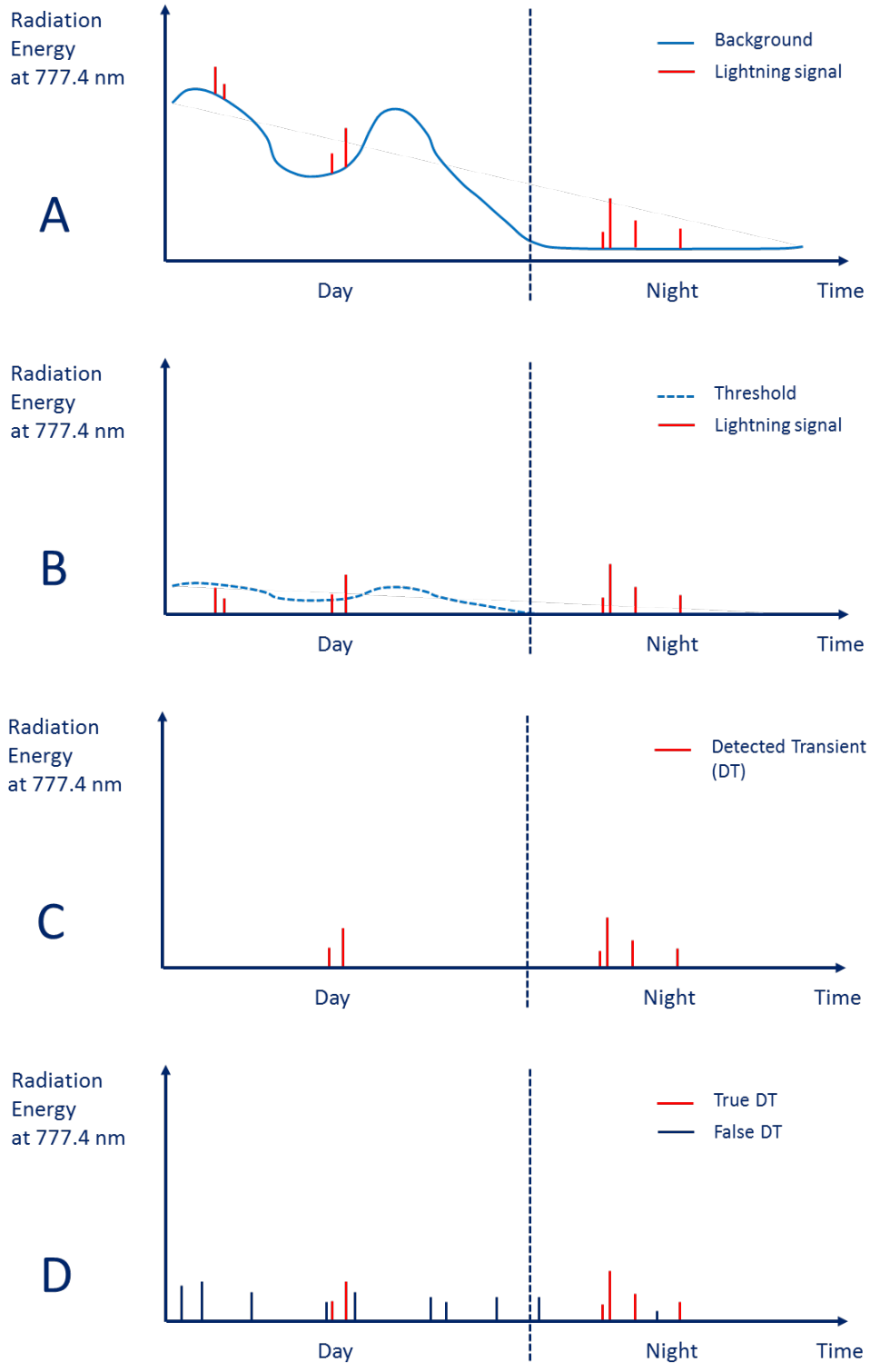


Figure 2: Schematic representation of the detection principle for one LI pixel. A) representation of the lightning signal on top of the background (variable from day to night); B) threshold computation and comparison of the threshold against the net lightning signal; C) DTs selected from the time sequence; D) Identification of true and false DTs after the on-board processing.

Figure 3 presents a schematic 2-D representation of the outcome of the steps A-C of Figure 2.

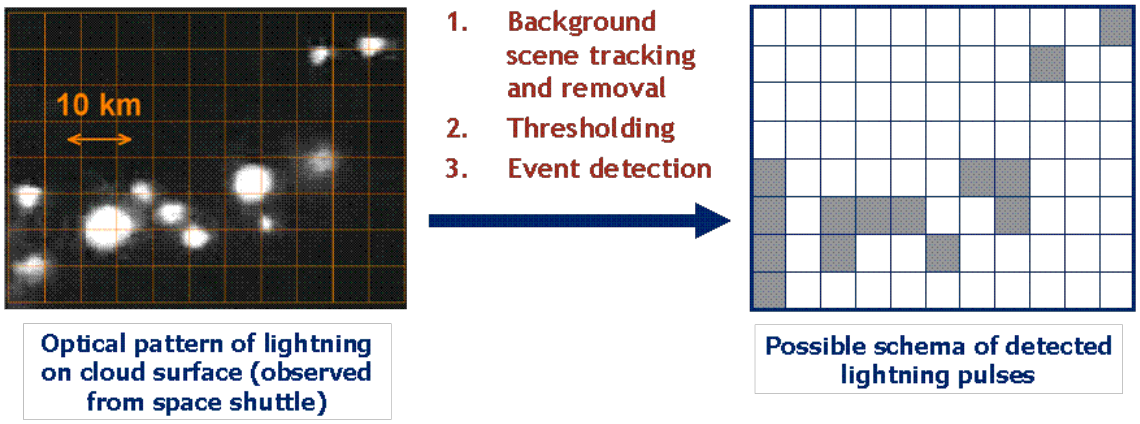


Figure 3: Detection principle of LI. A real observation from the space shuttle at night time (left panel) is used as an example image sampled with a schematic representation of the LI pixel grid; after steps 1-5, the candidate lightning events are located on the LI grid (right panel).

Candidate lightning events can originate from lightning or from different noise sources (see Figure 2 panel D) such as:

a. radiometric noise[1],
b. micro-vibration of the platform,
c. particle impacts on the focal plane,
d. Sun glint,
e. Sun stray light,
f. ghost effect,
g. etc.

For this reason, the sole detection process is not sufficient for the identification of lightning from space. As a consequence, a sequence of processing steps aimed at checking the spatial, temporal, and physical properties of the measurements is necessary after the detection. The main challenge of the LI processing is to differentiate between the events triggered by lightning (true DTs) and false events (false DTs) generated by noise.

The processing of the LI events is done by employing the following quantities that are available for each event:

a. location,
b. detection time,
c. radiance measurement, and
d. background estimation (as evaluated through the running average).

The LI instrument can be configured to provide the radiance measurements (c) and the background measurements (d) at the location/pixel of the detection only (1×1 pixel window),

[1] The LI detection principle favours the generation of a large number of false events due to radiometric noise. In fact, using a detection threshold that is derived from the theoretical level of radiometric noise (at 1σ; see bullet 4 in Section 3.2.2), implies detecting as events all those local fluctuations of noise that go above 1σ.

or at the pixel of the detection plus the measurements of its 8 neighbouring pixels, i.e. for a 3×3 pixel window (default configuration during routine operations). This enables a more refined interpretation of the radiance measurements at the location of the detection for the discrimination between true and false events.

The volume of data transmitted to ground depends on the number of lightning events occurring over one second (data rate). The current bandwidth budget of 30 Mbps will enable sending up to 80000 events per second (including their 3×3 pixel windows), and a full image of the background with a configurable repeat cycle. Namely, 15 or 30 or 45 or 60 seconds with 60 seconds being the baseline during routine operations.

In total, 4.7 million pixels are processed on-board every millisecond. In order to fit within the data rate limitations of the downlink bandwidth (30 Mbps), the number of false events needs to be considerably reduced already on-board by a number of on-board processing steps.
The rejection of false events is achieved through a sequence of filtering steps that starts on-board (Level 0, prior to the downlink) and continues on ground through the Level 1 and Level 2 processing. For the high-level description of the Level 0 and Level 1b filtering, the reader can refer to Sections 3.2 and 3.3 of [LIL2ATBD], respectively. In this document, we present a high-level description of the Level 2 filtering (see Section 4). Additional, more detailed information on the Level 2 processing is given in the [LIL2ATBD].

3.2.3 Calibration

The LI does not have an on-board unit to perform radiometric calibration while in-flight. For this reason, the in-flight radiometric characterization, calibration, and monitoring is done through analyses of dark scenes acquisitions during night-time and vicarious calibration techniques, respectively[2].
Dark scene measurements have to be performed when the Earth disk is totally dark. The best conditions to perform dark acquisitions are found when the Sun is eclipsed by the Earth. This happens every six months. For this reason, the dark acquisitions for the in-flight characterization of LI will take place twice a year around midnight. When set in CALIBRATION mode, LI acquires images with the four OCs (see Figure 1) simultaneously, and with different, configurable exposure times.
In order to characterize the detector Offset and Readout Noise, multiple acquisitions with the shortest possible exposure time are processed. For each pixel, one can derive these two quantities by computing the pixel-based average and standard deviation of the acquisitions, respectively. The outcome of this analysis (per pixel) is used to calibrate the measurements once the instrument is back in the operational mode. To characterize the Dark Current, two different exposure times are employed. Also in this case, multiple acquisitions for the two exposure times are performed. After Offset removal, a linear fit is done to derive the variation of the measurement as a function of the exposure time. The result of the linear fit (i.e., the slope) is the Dark Current.
In order to monitor the radiometric performances of the LI instrument, a database of Top of Atmosphere (TOA) reflectance 4-D histograms is built. Each histogram allows to sample the geographical space (2-D), the time during the day (1-D), and the reflectance space (1-D). The reflectance values are computed from the Level 1b background measurements. A new

[2] It is important to stress that the acquisition of a flat field is not possible; in fact, the Earth disk, characterized by a very strong pixel-by-pixel signal variability, is not a suitable target for this analysis.

histogram is created every 15 days, this will allow to sample the seasonal variation during the year. In addition, a new database is created every time the yaw-flip status changes.
The vicarious calibration makes use of those histograms: From each histogram in the database, the reflectance of different Pseudo-Invariant Calibration Sites (PICS) observed by the LI are extracted. Such sites, with typical areas of about 100×100 km^2, have a radiometric stability of about 2-3% and are located in desert regions, mainly in the Sahara. The reference reflectance values, to be compared against the ones stored in the database for the calibration, have been derived by EUMETSAT by using a combined climatology of ENVISAT MERIS and MSG SEVIRI data. The reference values also include daily and seasonal variations. For a set of available PICS (PIC_i), the observed reflectance (ρ) is compared to the reference one (ρ_{ref}) with the ratio $\left[\frac{\rho}{\rho_{ref}}\right]_{PIC_i}$. The average of the ratios is employed to derive the correction to the last Absolute calibration Factor (AF_{T-1}) with the aim of deriving an updated value (AF_T):

$$AF_T = \frac{AF_{T-1}}{\left[\frac{\rho}{\rho_{ref}}\right]_{AVG}}.$$

The new absolute calibration factor is uploaded to the instrument and is used to convert the signal in counts to radiance units.

4 LI LEVEL 2 PROCESSING ALGORITHMS

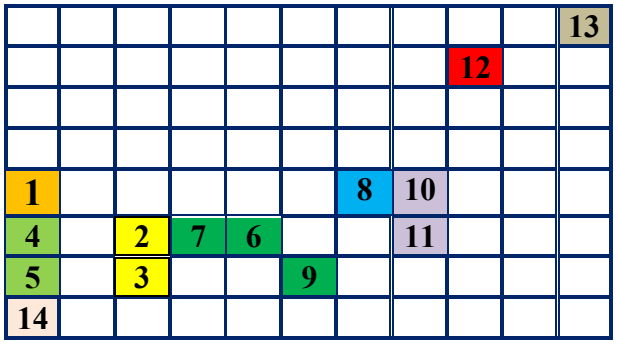
The statistical properties of LI events are as much depending on the instrument design (e.g., number of detector elements, integration time, etc.) as on the underlying geophysical characteristics of lightning itself. Therefore, the Level 2 algorithms make use of the properties of lightning to overcome the design related features. In addition to identifying lightning events, the Level 2 processing attempts in removing as many false events as possible.

4.1 Acceptance of Level 1b events and group processing

The processing of events at Level 2 starts from the acceptance of Level 1b events obtained from the operational Level 1 processing. These are still strongly dominated by false events due to noise (see Section 3.2.2). The acceptance step takes into account the outcome of the analyses at Level 1b to select those events that will be processed at Level 2. At this point of the processing, about 35,000 events per second are selected to be processed at Level 2. At the acceptance step, a quality indicator is associated to each event. After the acceptance of the events, the Level 2 groups are computed. Each group is a collection of connected events on the same acquisition frame.

4.1.1 Example group clustering sequence

In order to illustrate a simple case of group clustering, a conceptual example is presented in the following (see Figure 4). The example is based on the definition of group: a collection of connected events on the LI grid acquired within the same acquisition frame.



Event ID	Temporal frame Number	Group ID
1	1	1
2	2	2
3	2	2
4	5	3
5	5	3
6	5	4
7	5	4
8	5	5
9	5	4
10	12	6
11	12	6
12	15	7
13	15	8
14	18	9

Figure 4: Conceptual example of group clustering. Left panel: location of the events in a portion of the LI grid. The colour coding allows to relate the left panel to the right panel. Right panel: compact description of the outcome of the group clustering based on the location of the events in the left panel.

The group clustering sequence that results in the identification of a total of nine groups is the following one:

Group #1: The first event is found in the first frame and is assigned to the 1st group. The first group is a single-event group.

Group #2: Events #2 and #3 are connected events that are in the same acquisition frame (i.e., the second one), for this reason they belong to the same group.

Group #3, #4, and #5: a total of six events are available in frame number five; based on their location in the LI grid, three different groups can be defined. In this example, a 8-neighbour schema is employed for the group definition.

From fourteen events, the group clustering algorithm has created nine groups. Once the groups have been computed, the processing focusses on checking the properties of groups in five steps. It is important to stress that the group clustering and analysis is done per OC.

4.1.2 Check on the particle signature

Particle impacting the focal plane are known to generate clusters of events within a single acquisition frame (i.e., groups, see Section 4.1.1) that are highly elongated. Accordingly, false events due to particle impacts are identified through the group elongation. If the elongation is above a configurable threshold, the group is classified as false.

4.1.3 Check on the saturation of the radiance with the group

In the event of either Sun intrusion[3], Sun glint, or Sun stray-light, events with very high radiances (either saturated or close to the saturation) are expected to be detected. Those groups containing a fraction larger than a configurable threshold of the events that have radiances above a configurable radiance threshold (set very high in the dynamic range) are classified false.

4.1.4 Check on the Sobel gradient of the background

By analysing the average background gradient at the scene of the events in the group, one can spot groups that are composed by false events generated by the platform micro-vibrations at the location of a steep background gradient. In addition, the Level 2 processor allows the user to complement this analysis step with the monitoring of the total number of events for each frame. In fact, if micro-vibrations are so dominant to introduce regular peaks in the total number of events as a function of time, one can use this information to corroborate the classification of false groups with high average background gradient. If a group with high average background gradient (above a configurable threshold) is detected over a frame corresponding to an increase of total number of events, it is classified as false. If the processing is configured not to use the monitoring of the peaks of events, the classification is done only through the comparison between the average background gradient in the group and the configurable threshold.

[3] Sun within the instrument FOV.

4.1.5 Check on the radiance of the group

By analysing the radiance of the events composing a group, one can spot groups that are composed by false events generated by noise. This is done by exploiting the knowledge of the expected minimum detectable energy as a function of the background radiance. By employing a configurable threshold on the radiance of the events in the group, one can compute the fraction of events in the group that with radiances above the radiance threshold. If the fraction is larger than a configurable threshold, the group is considered true. Conversely, if a group does not have enough events above the configurable radiance threshold, it is classified as false.

4.1.6 Check on the size of the group

By analysing either the number of events in the group, or the physical size (in km^2) of the group, one can spot groups that are composed by false events. This stems from the fact that the radiometric noise should hardly create collections of transients over the same frame that are coherent in space. A group is classified as true if its size is larger than a configurable threshold.

4.1.7 Rejection of Groups prior to the Computation of Flashes

In order to perform a definitive rejection of a group, all the analysis steps presented in Sections 4.1.2, 4.1.3, 4.1.4, 4.1.5, and 4.1.6 are considered. In detail, two of the five steps can cause a rejection of a group, independently of the outcome of the others: the particle signature and the saturation (presented in Sections 4.1.2 and 4.1.3, respectively). If a group passes both these steps, it is classified based on the remaining three analysis steps. In detail, in order to be rejected, the group must be classified as false by the check on the relative Sobel gradient, the radiance, and the size. If the group is classified as true by one of these analysis steps, it goes to the flash computation. The use of each single analysis step can be turned on or off through the configuration of the Level 2 processing. For example, if after launch it will be assessed that the micro-vibrations of the platform are smaller than expected, one could decide to ignore the outcome of the analysis of the background gradient.
Finally, at the group rejection step, a quality indicator is associated to each group. Such indicator is computed by means of the outcome of all the group analysis steps and the average quality of the events in the group (see Section 4.1). Users will be provided with such quality indicator in the Level 2 group product (see Section 5).

4.2 Flash processing

The Level 2 flash computation consists in the aggregation of groups in space and time using spatial distance and temporal interval to define the limits within which such aggregation can take place.
This processing step can be performed in two ways for LI:

1. Flash clustering based on the Weighted Euclidian Distance (WED) algorithm: according to the WED flash clustering algorithm, a group has to be spatially and temporally near one other group in a flash to be part of a multi-group flash, such that the following relationship is fulfilled [MACH07]:

$$WED = \sqrt{\left(\frac{D}{D_{diff}}\right)^2 + \left(\frac{T}{T_{diff}}\right)^2} \leq 1.0$$

Where:

a. T is the time difference between the two groups (in ms);

b. D is the distance between the two groups (in km);

c. D_{diff} is the configurable threshold on the distance between groups;

d. T_{diff} is the configurable threshold on the time difference between groups.

2. Flash clustering based on separate time/distance checks: in this approach the distance and time checks are considered separately so that there are no trade off effects, e.g., with spatially close groups that are separated by a longer time. In this method two groups are considered as part of the same flash only when both of the following conditions are met:

   $D \leq D_{diff}$ and $T \leq T_{diff}$ with the same meaning of the quantities in bullet 1.

The default values for thresholds are D_{diff} = 16.5 km and T_{diff} = 330 ms. In addition, the flash computation is done for the four OCs independently. This allows one to speed-up the flash clustering (i.e., one of the most time consuming steps in the LI Level 2 processing).

4.2.1 Example flash clustering sequence

Flashes are defined by clustering groups. For the LI, a "full fit" flash clustering approach is adopted (in line with GLM). This means that, if a group can be clustered to multiple flashes, those flashes will be merged in a single one. The example that follows uses group centroids for the demonstration of the WED principle. The Level 2 processing of LI allows the user to configure the group clustering to use the distance between groups computed from the events composing them.
As an example demonstrating the clustering, we employ the following case study: 11th orbit on 4 Jan 2007 of TRMM/LIS. We focus on an area in which the LI flash clustering algorithm produces one flash from a total of 27 groups (see Figure 5).

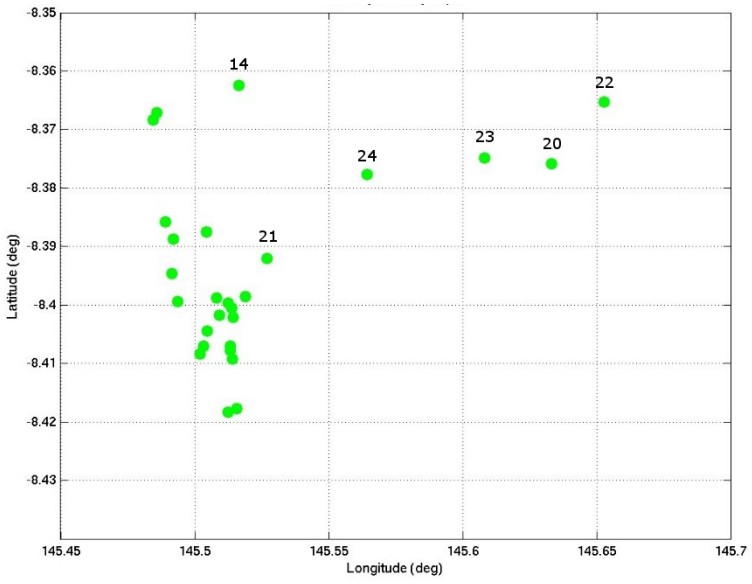


Figure 5. Snapshot of the area of interest with 27 groups. The numbers highlight the key groups for the definition of the flash. The 21 groups that are not explicitly numbered belongs to the cluster in the left side of the plot. Their locations and time is available in Table 1.

The useful information for the groups within the area in Figure 5 is available in Table 1. Groups #1 to #19 are clustered in the same flash (#1). Group #20 is the one that breaks the chain of clustering, with a minimum WED of 2.38 from the cluster of groups #1-#19: group #20 belongs to a new flash (#2). Group #21 belongs to flash #1. Groups #22 and #23 are added to flash #2. Group #24 is the bridge between the two flashes so far created. In fact, it has values for the WED distance that are below 1 with respect to group #21 and #23 (0.8003 and 0.8787, respectively); for example, the WED with respect to group #14 is 1.0771, right above the threshold at one. Flashes #1 and #2 can be merged at the time group #24 is analysed. The following three groups #25-#27 are added to the flash (these are located within the main cluster of the flash).

Table 1: Groups included in the snapshot area in Figure 5.

Group #	Time [sec]	Lat [deg]	Lon [deg]	Group time difference to previous group [ms]
1	441824316.9069	-8.3887	145.492	0
2	441824316.9155	-8.3947	145.492	8.59999657
3	441824316.9210	-8.3994	145.494	5.5000186
4	441824316.9301	-8.3875	145.504	9.10001993
5	441824316.9336	-8.4083	145.502	3.49998474
6	441824316.9391	-8.3988	145.508	5.5000186
7	441824316.9406	-8.4071	145.503	1.49995089
8	441824316.9497	-8.3986	145.519	9.10001993
9	441824316.9517	-8.4017	145.509	1.99997425
10	441824316.9642	-8.4183	145.513	12.5000477
11	441824317.0698	-8.3857	145.489	105.599999
12	441824317.0713	-8.3683	145.485	1.50001049
13	441824317.0878	-8.3671	145.486	16.4999962
14	441824317.2436	-8.3625	145.517	155.799985
15	441824317.2487	-8.4177	145.516	5.10001183
16	441824317.2522	-8.4071	145.513	3.49998474
17	441824317.2542	-8.4093	145.514	1.99997425
18	441824317.3381	-8.4078	145.513	83.9000344
19	441824317.3577	-8.3996	145.512	19.5999742
20	441824317.3667	-8.3758	145.633	9.00000334
21	441824317.3687	-8.3921	145.527	2.00003386
22	441824317.3687	-8.3653	145.653	0
23	441824317.3708	-8.3749	145.608	2.09999084
24	441824317.3723	-8.3777	145.564	1.50001049
25	441824317.3743	-8.4044	145.505	1.99997425
26	441824317.4260	-8.4006	145.514	51.699996
27	441824317.4959	-8.4021	145.514	69.8999763

The flash clustering and analysis (see Sections 4.2.2-4.2.10) are applied to the groups created by the four LI OCs separately. The analysis of the flashes consists in nine steps which are described in the following subsections.

4.2.2 Check on the single-group flashes

Single-group flashes (flashes composed by one group) are a specific case of flashes that are usually related to false events. However, in order to avoid any a-priori decision that may cause the rejection of true events, one can exploit the outcome of the analyses described in Section 4.1.7 to classify this kind of flash. In fact, single-group flashes can be described as

very specific groups: these have passed the group filtering, but after the last processing step they could be considered to be composed by false events. For this reason, in order to keep single-group flashes a second threshold-based filtering is performed on the group quality (defined in Section 4.1.7).

4.2.3 Check on the number of groups within the flash

By analysing the number of groups in a flash, one can spot flashes composed by false events (i.e., false flashes). This stems from the fact that the radiometric noise should hardly create sequences of groups that are coherent in space and time to form flashes with large number of groups. A flash is classified as true if it has a number of groups larger than a configurable threshold.

4.2.4 Check on the "footprint" of a flash

By analysing the size of the flash footprint, one can spot flashes composed by false events. This stems from the fact that the noise should hardly create sequences of false groups that are coherent in space and time to form flashes with a large footprint. The method allows one to handle two different cases: *i)* flash footprint composed by a single patch, and *ii)* flash footprint composed by multiple patches (i.e., scattered flash). Similarly to the group size step (see Section 4.1.6), the flash footprint can be computed and analysed in number of pixels or the physical size (in km^2).

4.2.5 Check on the time correlation between groups within a flash

By analysing the temporal correlation between subsequent groups within a flash, one can compute the probability of the flash being composed by false events. This stems from the fact that the noise should hardly create sequences of false groups that are closely correlated in time. The analysis checks the average time between subsequent groups. In the case in which this quantity is above a configurable threshold, the flash is considered false.

4.2.6 Check on the spatial correlation between groups in a flash

By analysing the spatial correlation between subsequent groups within a flash, one can compute the probability of the flash being composed by false events. This stems from the fact that the noise should hardly create sequences of false groups that are closely correlated in space. The analysis checks the average spatial distance between subsequent groups. In the case in which this quantity is above a configurable threshold, the flash is considered false.

4.2.7 Check on the Sobel gradient of the background

By analysing the average background gradient at the scene of the events in the flash, one can spot groups that are composed by false events generated by the platform micro-vibrations at the location of a steep background gradient. Differently from what was done for the groups (see Section 4.1.4), this analysis step cannot be complemented by the information on the number of events as a function of time. This is due to the fact that flashes are defined over multiple frames.

4.2.8 Rejection of flashes

In order to perform a definitive rejection of a flash, all the analysis steps presented in Sections 4.2.2, 4.2.3, 4.2.4, 4.2.5, 4.2.6, 4.2.7 are considered. In detail, a flash is rejected if a single analysis step classifies it as false. It is important to stress that (as for the group filtering step in Section 4.1.7) the use of each single analysis step can be turned on or off through the configuration of the Level 2 processing.
Finally, at the flash rejection step, a quality indicator is associated to each flash. Such indicator is computed by means of the outcome of all the flash analysis steps and the average quality of the groups in the flash (see Section 4.1.7). Users will be provided with such quality indicator in the Level 2 flash product (see Section 5.1).

4.2.9 A posteriori re-introduction of flashes

This analysis step has the aim of using the outcome of the Level 2 processing to re-introduce flashes that have been rejected by the Level 2 filtering (see Section 4.2.8). The analysis performs a check of the distance between the rejected flashes and the flashes that passed Level 2 filtering. When a rejected flash is found to have at least one event at the same location of events of flashes considered true by the Level 2 processing: the rejected flash will be reintroduced in the Level 2 product. In order to meet the timeliness, the check is done between flashes within the same Level 2 chunk of 10 sec.

4.2.10 Removal of duplicate flashes in overlap regions

In order to prepare the Level 2 flashes for the definition of the Level 2 products, the removal of duplicate flashes in the overlap between OCs must be performed. This prevents an over-representation of the flash activity in overlap regions that would have an impact on all the Level 2 products.
Given a pair of flashes from different OCs within pre-defined overlap regions, a check is done on the time intervals covered by the two flashes. If the two temporal windows of the flashes overlap within the threshold used in the flash clustering (see Section 4.2), a check is done on the physical distance between those groups of the two flashes that are acquired within the temporal window of overlap. If the smallest distance so computed is below the distance employed for the flash clustering, the flash with the largest total radiance is kept for further processing, the other one is removed.

5 PRODUCT TYPES

5.1 Group and Flash products

Two products providing point-like information for the continuous monitoring of the lightning activity are delivered to the users. Namely:
1. Level 2 groups (LI-2-LGR-x-FD-x), and
2. Level 2 flashes (LI-2-LFL-x-FD-x).

In the following, these are also referred to as "Initial Products". Level 2 events (LI-2-LEF-x-FD-x), i.e., the finest scale over which LI can monitor lightning activity, are not disseminated to users. These data are archived at EUMETSAT, and can be ordered by users via dedicated queries to the EUMETSAT archive.
Both group and flash products are delivered with a repeat cycle of 10 seconds. The details of their format are provided in Sections A.4 and A.6, respectively. In Table 2 and Table 3, the list of the most important variables in these two products are presented and described.

Table 2: Key variables in the LI-2-LGR-x-FD-x product.

Variable	Description
`group_time`	UTC time of the group.
`latitude`	Average latitude of the events in the group weighted by the event radiance.
`longitude`	Average longitude of the events in the group weighted by the event radiance.
`radiance`	Total radiance of the group[4].
`number_of_events`	Number of events in the group.
`flash_id`	ID of the flash that allows one to relate the groups in the LI-2-LGR-x-FD-x to the LI-2-LFL-x-FD-x (see Table 3).
`group_filter_qa`	Quality of the group computed at the group-rejection step; the step at which groups are analysed singularly (see Section 4.1.7).

Table 3: Key variables in the LI-2-LFL-x-FD-x product.

Variable	Description
`flash_time`	UTC time of the first group in the flash.
`flash_duration`	Duration of the flash, i.e., time difference between the first a the last group of the flash.
`latitude`	Average latitude of the events in the flash weighted by the event radiance.
`longitude`	Average longitude of the events in the flash weighted by the event radiance.
`radiance`	Total radiance of the flash[4].

[4] See the Appendix in [KOS10] for the interpretation of the group/flash radiance from LI.

number_of_events	Number of events in the flash.
number_of_groups	Number of groups in the flash.
flash_footprint	Size of the flash footprint in number of pixels.
flash_id	ID of the flash.
flash_filter_confidence	Quality of the filter computed at the group-rejection step; the step at which groups are analysed singularly (see Section 4.2.8).

The variables group_time, flash_time, latitude, longitude, flash_duration, are the "key variables" of the LI-2-LGR-x-FD-x and LI-2-LFL-x-FD-x products. In fact, these provide the minimum required information for the continuous monitoring of the lightning activity through LI. In order to understand why a set of "key variables" has been defined, it must be stressed that during the initial phase of the dissemination of the products (i.e., right after the LI Commissioning), the "key variables" will be the ones that should be considered the most reliable by users.
The variables number_of_events, number_of_groups, flash_footprint, and flash_id are directly related to the basic elements of the LI detection and processing. For this reason, they should be considered reliable form the beginning of the dissemination phase. These variables provide users with information about the number of LI events that compose groups and flashes, i.e., how frequent in time and spatially extended are the optical emissions associated to flashes. The latter being an attribute that is derived from the imaging capability of LI.
The reliability of the radiance variable in both LI-2-LGR-x-FD-x and LI-2-LFL-x-FD-x will strongly depend on the completeness of the assessment of the LI radiometric performances (both absolute and relative) at the time of the dissemination. Such an assessment may require a period longer than the duration of the Commissioning.
Finally, group_filter_qa and flash_filter_confidence shall be regarded to be reliable after the LI System will have reached a stable configuration. In fact, these two variables will provide users with information about the margin with which each group and flash has passed the Level 2 filtering. The LI System will undergo tuning campaigns that will be followed by thorough/long-term assessments of the properties of groups and flashes composing the Level 2 products (including the quality indicators). Only after EUMETSAT will be confident about the reliability of the information provided by these two variables, users will be advised to exploit them.

5.2 Accumulated (gridded) products

Differently from the point-like products presented in Section 5.1, the LI accumulated products provide the users with data that make use of the imaging information of LI. In fact, they contain accumulated information computed from the events that compose the flashes that passed the Level 2 filtering (see Section 4). The accumulation is done over 30 seconds, which is also the repeat cycle of the LI Level 2 accumulated products. As a consequence, the accumulated products do not enable a continuous monitoring of the lightning activity.
For a given accumulated product, each of the events in the Level 2 flashes is employed for the processing. The integration time over which the products are defined is meant to be a compromise between delivering event-based information and satisfying users who require rapid access to such information. Those users who want to collect data over longer time

periods can simply stack the 30-second-long products. The MTG-FCI IR Level 1c grid (see Section 7.7) has been selected as the accumulation grid in order to enable an easy comparison/overlay of lightning data with MTG FCI Level 1c images. This facilitates the synergetic use of FCI and LI products in the users' applications.
In the accumulation process, the events belonging to those flashes within the accumulation time window are used as input. Being these events defined on the LI grid (1000×1170 pixels per OC; see Section 3.2.1), they are employed to compute the following quantities:

a. **Level 2 Accumulated Flash Area** (see Figure 6): the events are employed to derive the flash footprint on the LI grid and each flash contributes with a value equal to one to the accumulation.

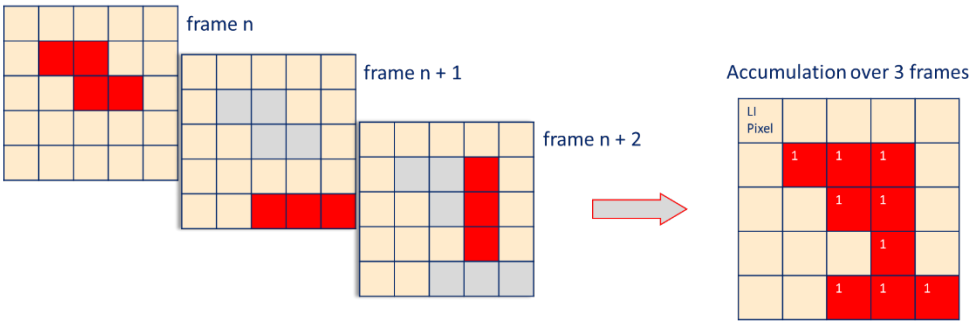


Figure 6: Graphic representation of the accumulation of events belonging to a flash and contributing to the AFA. Left: the different groups composing the flash and their respective acquisition frame; in grey, the accumulated events up to the accumulation frame while in red, the events being added to the accumulation at each frame. Right: The final result of the accumulation.

b. **Level 2 Accumulated Flash** (see Figure 7): the events are employed to derive the accumulated number of events per LI grid element belonging to the flash footprint; after this accumulation, the accumulated value for each element is normalized to the total number of events within the flash.

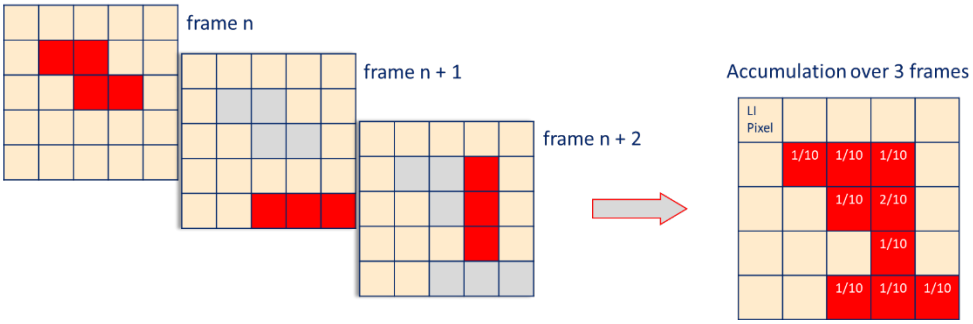


Figure 7: Graphic representation of the accumulation of events belonging to a flash and contributing to the AF. Left: the different groups composing the flash and their respective acquisition frame; in grey, the accumulated events up to the accumulation frame while in red, the events being added to the accumulation at each frame. Right: The final result of the accumulation.

c. **Level 2 Accumulated Flash Radiance** (see Figure 8): the radiances of the events (R# in Figure 8) are employed to derive the accumulated radiance per LI grid element belonging to the flash footprint.

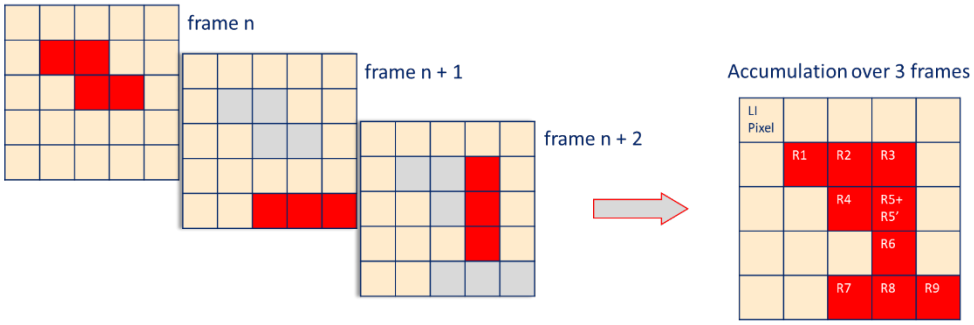


Figure 8: Graphic representation of the accumulation of events belonging to a flash and contributing to the AFR. Left: the different groups composing the flash and their respective acquisition frame; in grey, the accumulated events up to the accumulation frame while in red, the events being added to the accumulation at each frame. Right: The final result of the accumulation.

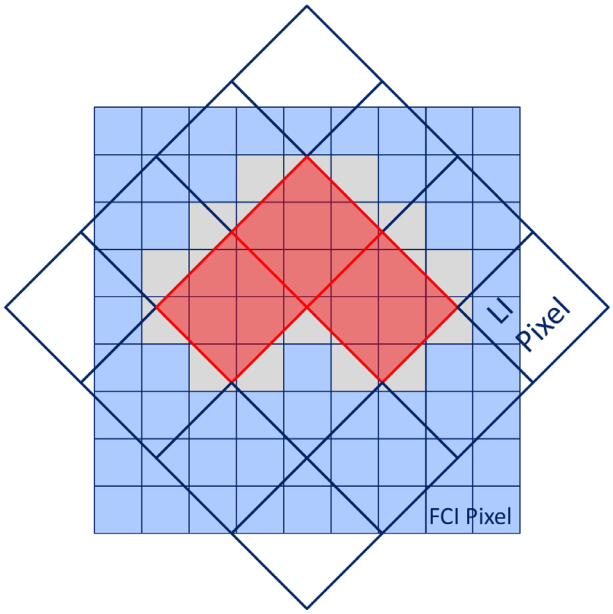


Figure 9: Diagram depicting the re-gridding of 3 pixels of a LI accumulated product (red pixels within the 45 deg tilted grid; pixel size 4.5 km) onto the FCI Level 1c grid (grey pixels within the sky-blue grid; pixel size 2 km).

The information computed in a, b, and c, available on the LI grid, is then transferred onto the FCI Level 1c 2 km grid through a re-gridding step as depicted in Figure 9. The outcome of the accumulation defines the content of the three LI accumulated products on the FCI 2 km grid:

1. Accumulated Flash Area (LI-2-AFA-x-FD-x)
2. Accumulated Flash (LI-2-AF-x-FD-x)
3. Accumulated Flash Radiance (LI-2-AFR-x-FD-x)

In Table 4, Table 5, and Table 6, the lists of the most important variables in the accumulated products are presented.

Table 4: Key variables in the LI-2-AFA-x-FD-x product.

Variable	Description
x	x-location in the FCI IR 2 km grid (see Section 7.7).
y	y-location in the FCI IR 2 km grid (see Section 7.7).
accumulated_flash_area	Value of AFA at the locations {x, y}.
average_flash_qa	Average flash quality of the set of flashes employed for the computation of the product (see Section 4.2.8 and Table 2). 0=higher probability true; 1=higher false probability.

Table 5: Key variables in the LI-2-AF-x-FD-x product.

Variable	Description
x	x-location in the FCI IR 2 km grid (see Section 7.7).
y	y-location in the FCI IR 2 km grid (see Section 7.7).
flash_accumulation	Value of AF at the locations {x, y}.
average_flash_qa	Average flash quality of the set of flashes employed for the computation of the product (see Section 4.2.8 and Table 2). 0=higher probability true; 1=higher false probability.

Table 6: Key variables in the LI-2-AFR-x-FD-x product.

Variable	Description
x	x-location in the FCI IR 2 km grid (see Section 7.7).
y	y-location in the FCI IR 2 km grid (see Section 7.7)..
flash_radiance	Value of AFR at the locations {x, y}.
average_flash_qa	Average flash quality of the set of flashes employed for the computation of the product (see Section 4.2.8 and Table 2). 0=higher probability true; 1=higher false probability.

The variables `accumulated_flash_area` and `flash_accumulation`, are the "key variables" of the LI-2-AFA-x-FD-x and LI-2-AF-x-FD-x products, respectively. In fact, these are derived from the events composing the flashes and are directly related to the basic elements of the LI detection. For this reason, they should be considered reliable from the beginning of the dissemination phase. These variables provide users with information about the flash occurrences (or flash rate, if normalized by the time) and density of LI events within flashes, respectively. As pointed out in Section 5.1:

- the reliability of the variable `flash_radiance` will strongly depend on the completeness of the assessment of the LI radiometric performances (both absolute and relative) at the time of the dissemination, and
- the `average_flash_qa` (directly related to the `flash_filter_confidence`) shall be regarded to be reliable after the LI System will have reached a stable configuration.

Also for these products, EUMETSAT will communicate to users when a variable shall be considered reliable.

6 PRODUCT NAMING CONVENTION

All MTG products have a WMO-compatible name, following the WMO file naming convention [WMO-386] (cf Attachment II-15 p25 2009 edition). The filename will consist of the dataset (or product) name with a file_type and a compression field:

(dataset_name).(file_type)(compression)

Where, `dataset_name` is composed of the following fields, separated by underscore symbols:

(pflag)_(productidentifier)_(oflag)_(originator)_(yyyyMMddhhmmss)_(freeformat)

`productidentifier` is composed of the following fields, separated by commas:

(locationindicator),(datadesignator),(freedescription)

`freedescription` is composed of the following fields with plus symbol or dash symbol separators:

(spacecraftid)-(data_source)-(processing_level)-(type)-(subtype)-(coverage)-(subsetting)-(component1)-(component2)-(component3)-(purpose)-(format)

`freeformat` is composed of the following fields, separated by underscore symbols:

(facility_or_tool)_(environment)_(start_time)_(end_time)_(processing_mode)_(special_compression)_(disposition_mode)_(repeat_cycle_in_day)_(count_in_repeat_cycle)

The order of the fields is mandatory. It is worth noting that when there is no relevant value within the `freedescription` or `freeformat` sections, the field is left out (indicated by an empty string). This can lead to the repetition of dashes or underscores.
Table 7 shows the fully expanded set of name fields in the correct order, with values described for LI Level 2 datasets. Following the main table, subsequent subsections describe the allowed values for the selected fields in greater detail.

Table 7: Breakdown of the fields in the LI Level 2 dataset naming convention.

Name Field	Description	LI Level 2 Product Values
`pflag`	WMO mandated	"W"
`locationindicator`	WMO mandated	"XX-EUMETSAT-Darmstadt"
`datadesignator`	The type of data with respect to the categories and subcategories defined in [|WMO-386],	"IMG+SAT"
`spacecraftid`	Spacecraft indicator	"MTIn" for MTG Imager n where n = 1,2,3 or 4
`data_source`	Instrument, platform or SAF	"LI"
`processing_level`	Processing Level	"2"

Name Field	Description	LI Level 2 Product Values
type	Indentifies the type of data	"LEF" - Filtered lightning events "LGR" - Lightning groups "LFL" - Lightning flashes "AF" - Accumulated flashes "AFA" - Accumulated flash area "AFR" - Accumulated flash radiance
subtype	Identifies a sub-type for the type.	""
coverage	Coverage of the full accumulation interval	"FD" for full disk
subsetting	Identification of the type of subsetting performed	""
component1	Identifies a first level component of the product	"CHK" for chunk "QCK" for a quick-look file
component2	Identifies a second level component of the product	"BODY" for a body chunk "TRAIL" for a trailer chunk "IMAGE" for an image quick-look
component3	Identifies a third level component of the product	""
purpose	The intended purpose of the dataset. This normally refers to the intended final recipient.	"ARC" for an archival dataset "DIS" for a dissemination dataset
format	The intended encoding format of the dataset.	"NC4E" for NetCDF-4 enhanced model "PNG" for a quick-look PNG image
oflag	WMO mandated	"C"
originator	WMO mandated	"EUMT"
yyyyMMddhhmmss	Is the UTC time of the processing, defined as the time of the formatting of the dataset/product by the processor [TBC-EUMETSAT], formatted in Abbreviated Generalised Time format e.g. yyyy = year MM = month dd = day of month hh = hour of day mm = minute of hour ss = second of minute	
facility_or_tool	Facility or tool producing the dataset	"L2PF" = Level 2Processing Facility "GTT" = Generic Test Tool
environment	Ground Segment Environment producing the dataset	"OPE" - Operational
start_time	UTC Time of start of Sensing Data formatted in Abbreviated Generalised Time format (see above).	For the body chunk, this will be the time of the first measurement in the chunk. For a trailer chunk or a quick-look, this is the start time of the first body chunk in the repeat cycle.
end_time	UTC Time of end of Sensing Data formatted in Abbreviated Generalised Time format (see above).	For the body chunk, this will be the time of the last measurement in the chunk. For a trailer chunk or a quick-look, this is the end time of the last body chunk in the repeat cycle.
processing_mode	Identification of the mode of processing	"N" = nominal
special_compression	This field provides identification of a special compression technique that has been applied to one or more variables in the dataset. Special compression does not include the standard NetCDF data compression or "deflation" using in-built zlib support which is transparent to the user.	No special compression is applied for LI L2 products

Name Field	Description	LI Level 2 Product Values
disposition_mode	Shows disposition of the dataset from the perspective of an end-user"s needs.	"C" = commissioning "O" = operational "T" = testing
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the current group repeat cycle in the day for this particular dataset. For LI, which does not have a natural repeat cycle, this is expected to correspond to the number of accumulation interval for the archived product (e.g. every 10 minutes [TBC]). The counter starts at 0001 for the first group accumulation interval at or after midnight and resets for the next group accumulation interval at or after the following midnight. This value can be used to associate all chunks of a repeat cycle dataset.	Variable
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuities when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicable (e.g. datasets which are not chunk-able).	Variable
file_type	Indicator of the encoding format of the data, according to WMO conventions.	".nc" – NetCDF ".png" – PNG image
compression	Indicator of compression applied to the dataset as a whole according to WMO conventions (as opposed to the internal compression of variable indicated by the "special_compression" name field).	None

7 CHARACTERISTICS OF THE LEVEL 2 PRODUCTS

7.1 Overview

Six LI Level 2 products are produced operationally as indicated in Table 8. The product ID defined as:

`<DATA_SOURCE>-<PROCESSING_LEVEL>-<TYPE>-<SUBTYPE>-<COVERAGE>(-<SUBSETTING>)`,

is used to identify each product. The individual fields are the same as defined in Table 8. When a field has no value, this is represented in the Product ID as a lower case "x".

Table 8: Summary of LI Level 2 Products

Product ID	Data	Archived	Disseminated
LI-2-LEF-x-FD-x	Filtered lightning events	Y	-
LI-2-LGR-x-FD-x	Lightning groups	Y	Y
LI-2-LFL-x-FD-x	Lightning flashes	Y	Y
LI-2-AF-x-FD-x	Accumulated flashes	Y	Y
LI-2-AFA-x-FD-x	Accumulated flash area	Y	Y
LI-2-AFR-x-FD-x	Accumulated flash radiance	Y	Y

As most MTG Level 2 products, the LI Level 2 products are formatted as Network Common Data Format (NetCDF-4). The NetCDF data format supports dimensions, variables and attributes. Variables have a name, a shape defined by their dimensions, and a data type. Attributes store information about variables. Every attribute has an associated variable or group. Global attributes are considered attached to the root group. Every attribute has a name, a value, a data type, and a length. Example attributes are long_name, units, scale_factor and add_offset.

The term data chunk is used to denote a portion of a dataset in view of ground segment processing. The length of the data chunk is sized to meet timeliness and processing criteria. A dataset may be composed of differing types of data chunks. The two possible chunk types are body and trailer chunks. One or more body data chunks per repeat cycle or group accumulation interval are possible. They contain the Level 2 data and share a common format (NetCDF-4).

The trailer chunk always contains at least the list of preceding body chunks that have been produced for the dataset. The trailer chunk is also a NetCDF-4 file. The information if a dataset is chunked and the chunk type can be obtained from the product file name from the fields 'component1' and 'component2' (see Table 7).

For the near real-time dissemination (see the last column in Table 8), short duration chunks are generated in order to meet timeliness requirements. The inital products (LI-2-LGR-x-FD-x and LI-2-LFL-x-FD-x) have a configurable chunk size between 7 and 11 seconds[5], while the accumulated products have a typical chunk size of 30 seconds.

[5] The current assumption is 10 seconds.

All products listed in Table 8 are archived and retrievable by users, and only the LI-2-LEF product is not disseminated operationally. The archived version of the products will be generated over longer time scales to make storage and retrieval easier. The format of a given product is the same regardless of the temporal coverage of the product. The detailed format for each product is given in Appendix A.

7.2 NetCDF Structures

The LI Level 2 products have a flat structure without any internal NetCDF groups. The only exception is the LI-2-LEF-x-FD-x product which uses groups to separate the data from the four optical heads (Figure 10 and Section 7.3).

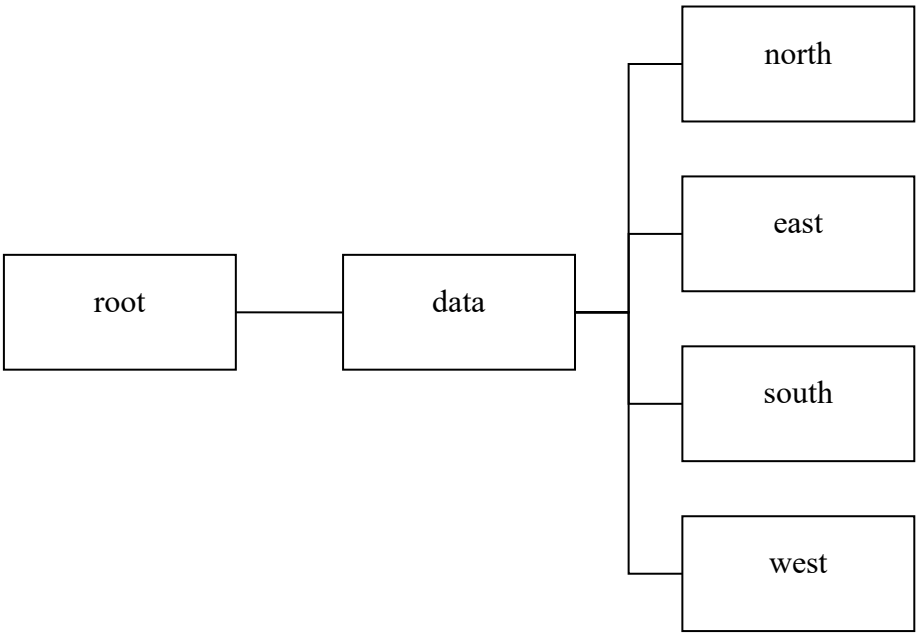


Figure 10: Figure showing the group structure of the LI-2-LEF-x-FD-x product groups.

7.3 Detector Orientation

The four optical heads are identified within LI products as north, south, east, and west, irrespective of the yaw flip status. This means that the correspondence between north, south, east, west and detector numbers changes with yaw flip status. The correspondence is described in Table 9:

Table 9: Relationship between detector group names, detector identifiers and yaw flip mode

Detector Group Name	Detector Identifier	
	Yaw Flip Summer	Yaw Flip Winter
west	detector 1	detector 3
north	detector 2	detector 4
east	detector 3	detector 1
south	detector 4	detector 2

7.4 Global Attributes

All MTG NetCDF products, whether flat or grouped, share a common set of global attributes containing metadata about the product. Some general examples are given in Table 7 and product-specific examples may be found in the format descriptions in the Appendix A. Many of these global attributes have analogues in the product name (see Table 7).

Table 10: Common set of global attributes for MTG NetCDF products.

Name	Description
Conventions	Conventions that the product conforms to, e.g. "CF-1.6, Unidata Dataset Discovery v1.0"
title	Dataset/product name formatted as set out in Section 6
summary	As defined in the relevant dataset/product format specification
keywords	As defined in the relevant dataset/product format specification
keywords_vocabulary	As defined in the relevant dataset/product format specification
history	"original generated file"
institution	"EUMETSAT"; this field may be extended with other values should datasets/products be generated in other locations
location_indicator	"XX-EUMETSAT-Darmstadt"; this field may contain other values should products be generated in other locations.
data_designator	As per the dataset name field "data_designator" in Table 7
platform	As per the dataset name field "spacecraft" in Table 7
data_source	As per the dataset name field "data_source" in Table 7
processing_level	As per the dataset name field "processing_level" in Table 7
coverage	As per the dataset name field "coverage" in Table 7
type	As per the dataset name field "type" in Table 7
subtype	As per the dataset name field "subtype" in Table 7
component1	As per the dataset name field "component1" in Table 7
component2	As per the dataset name field "component2" in Table 7
component3	As per the dataset name field "component3" in Table 7
product_id	The identifying product_id as used in the SIP
baseline_version	Allows products with a contiguous baseline to be identified
release_version	Used to tag a reprocessing version. Near-real-time data will be tagged as "Original"
processor_version	Processor version
algorithm_version	Algorithm version
format_version	Format version of the dataset/product
time_coverage_start	As per the dataset name field "start_time" in Table 7
time coverage_end	As per the dataset name field "end_time" in Table 7
processing_mode	As per the dataset name field "processing_mode" in Table 7
special_compression	As per the dataset name field "special_compression" in Table 7

Name	Description
subsetting	Empty for LI L2 products
disposition_mode	As per the dataset/product name field "disposition_mode" in Table 7
source	Characterisation of the type of data as per [CF]
runtime_data	Audit trail information listing auxiliary input data
parent_data	Audit trail information listing parent product data
linked_data	Catalogue information listing data to be linked to in the archive
facility_or_tool	As per the dataset name field "facility_or_tool" in Table 7
environment	As per the dataset name field "environment" in Table 7
references	"www.eumetsat.int"; it is intended that users of the dataset/product can access published, web-based references describing the data and the methods used to produce it at this address
comment	Useful comments. Default value: "None"
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> Where: < > indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY is the year value of the "date-time_position" root attribute DDD is the day in year value derived from the "date-time_position" root attribute, left padded with zeroes: 001 = Jan 1st, etc. NNNN is a copy of the "repeat_cycle_in_day" field
repeat_cycle_in_day	As per the dataset/product name field "repeat_cycle_in_day" in Table 7
count_in_repeat_cycle	For chunked products, this is the cumulative count of the dataset chunk in the repeat cycle or group accumulation interval starting from 0001. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data. It allows the receiver of the data to check that no data was lost during dissemination. Products that are not chunked will have a value of 0000.
instrument_configuation_id	Value of the "instrument configuration identifier" from the Level 0 data ICU-I auxiliary data
instrument_configuation_id_version	Value of the "instrument configuration identifier version" from the level 0 data ICU-I auxiliary data
mtg_name	String field containing the MTG WMO-convention name for the file
alternative_name	String field containing a possible alternative name for the file (e.g., Sentinel-4 naming convention)
purpose	As per the dataset/product name field "purpose" in Table 7
format	As per the dataset/product name field "format" in Table 7
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise left empty.
naming_authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file). Otherwise left empty.

Name	Description
geospatial_lat_min	Specifies the southernmost latitude covered by the dataset. This attribute is not present if the dataset has no relevant geolocation information available.
geospatial_lat_max	Specifies the northernmost latitude covered by the dataset. This attribute is not present if the dataset has no relevant geolocation information available.
geospatial_lon_min	Specifies the westernmost longitude covered by the dataset. This attribute is not present if the dataset has no relevant geolocation information available.
geospatial_lon_max	Specifies the easternmost longitude covered by the dataset. This attribute is not present if the dataset has no relevant geolocation information available.

7.5 Encoded Float Values

In order to compact the data, some float values are encoded as integers with a scaling factor and offset provided by associated netCDF attributes as per the CF conventions (see Appendix B). The attributes are called `scale_factor` and `add_offset`, respectively. CF-aware tools are able to automatically decode the original float value from the stored integer, but it can also be retrieved as:

```
float_value = (integer_value × scale_factor) + add_offset
```

7.6 Time as Real Data Type

The time is by default the UTC time encoded as an NC_DOUBLE data type following the CF conventions for a time coordinate. The units attribute will have a default value of “seconds since 2000-1-1 00:00:00.0” unless a different unit or epoch time are required for reasons of enhanced precision or precision over a longer time period.
This default encoding is shown in the CDL example below:

```
dimensions:
        time = 100;

variables:
        double time (time);
                time:standard_name = "time";
                time:long_name=”Time of observation”
                time:units = "seconds since 2000-01-01 00:00:00.0";
```

This example shows an array of 100 times, stored in an array of NC_DOUBLE data types. The times are stored as seconds since an epoch of midnight on January 1st 2000.

7.6.1 Time in LI Level 2 Accumulated Products

The accumulation start time indicates the time of the first flash employed in the accumulation. The values of the accumulation start time in the accumulated products are given in seconds since 2000-01-01 00:00:00.
The format of the LI Level 2 accumulated products (Appendices A.8 to A.13) is defined in such a way that it can be used for the disseminated and the archived accumulated products.

The main difference between the two is the time interval covered. Whereas the disseminated accumulated products cover typically 30 seconds, the accumulated products stored in the archive cover 1 to 10 minutes of data, i.e. an aggregation of 2 to 20 30-second chunks.
For each 30-second interval, the accumulation start time, given by the variable *accumulation_start_times,* is the same for all pixel-based parameters, e.g. the flash_accumulation in the LI-2-AF-x-FD-x. That is the reason why the dimension *accumulations* has a value of 1 for the disseminated products. In the archived accumulated products the dimension *accumulations* takes values between 2 and 20 covering 1 to 10 minutes.
For a given accumulation in the dataset, the associated parameters in variables with pixel dimension can be retrieved using the *accumulation_offsets* index. For accumulation 'i' the parameters can be accessed as per following examples:

```
start_pixel = accumulation_offset[i]
end_pixel = accumulation_offset[i+1]-1 // if 'i' is the last accumulation then use the last array index
num_accumulation_pixels = end_pixel – start_pixel +1
single_flash_accumulation[j] = flash_accumulation[start_pixel+j] // for j in range [0…num_accumulation_pixels]
```

7.7 Cross-Referencing in Initial Products

The identifier values in the fields `group_id` and `flash_id` are consistent across the three initial products produced for the same time period and can be used for cross-referencing of information.

7.8 Reference Grid

The geo-location information for the LI accumulation products is based on the standard idealized geostationary projection used by MTG FCI Level 1c products for the infrared channels.

7.8.1 Row and Column Numbering

A row is defined as a line accumulation bins running West to East. The rows are numbered from the south to north starting from 1 (Figure 11).

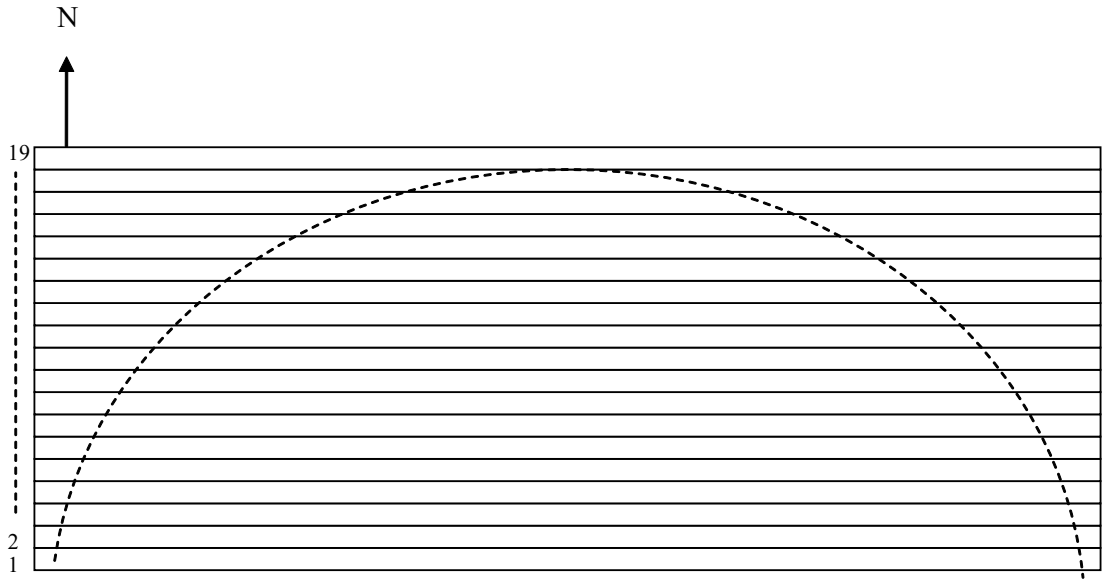


Figure 11: Illustration of row numbering within a rectifictaion grid

A column is defined a line of accumulation bins running in South to North. The columns are numbered from the West to East starting from 1 (Figure 12).

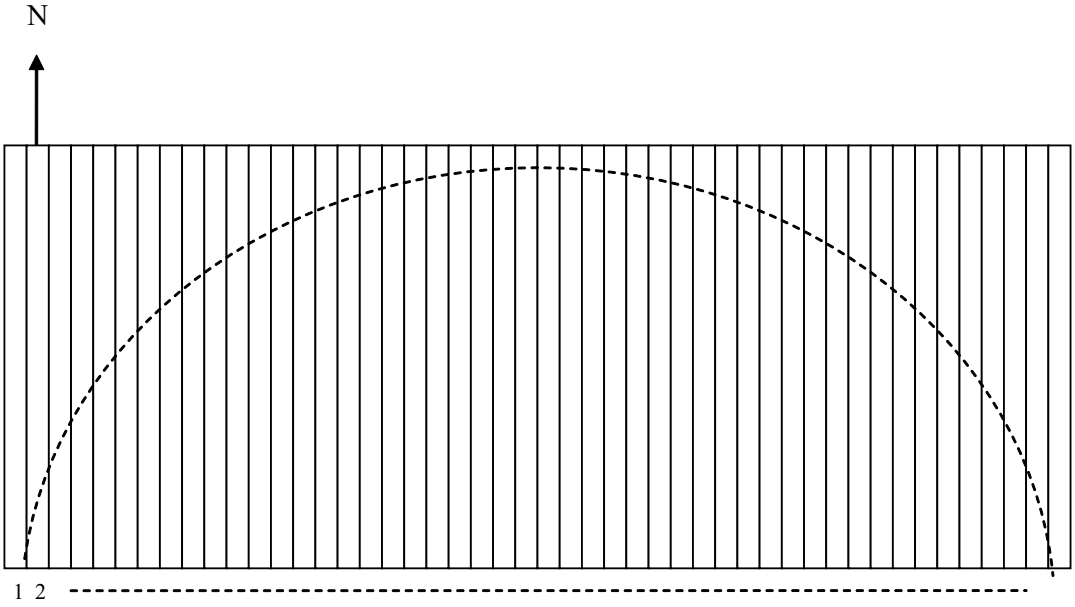


***Figure 12: Illustration of column numbering within a rectified* image**

The origin pixel at position [1,1] is located in the SW corner.

7.8.2 Reference Grid Definition

The reference grid defines the geo-referenced position of the pixels or accumulation bins. The grid steps are equiangular both in satellite azimuth and elevation and equal to the spatial sampling angle of an FCI Level 1c channel that has a projected distance at the sub-satellite point (the spatial sampling distance or SSD), of 2 km.

The normalized geostationary projection describes the view from a virtual satellite to an idealized Earth. The virtual satellite is in a geostationary orbit, perfectly located in the Equator plane at the given longitude, λ_D (normally 0 deg). This point on the equator is the origin of the projection. The distance between virtual satellite and centre of Earth (the geostationary radius) is given by the geostationary altitude above the surface and the equatorial radius of the Earth's reference ellipsoid. This normalized geostationary projection defines the line of sight of each pixel centre P as a vector representing the view from a virtual satellite in geostationary orbit, perfectly located in the Equator plane at the given longitude λ_D. This vector is expressed as a function of two angles called elevation (ϕ_S) and azimuth (λ_S) and is defined as follows:

$$\lambda_S = tan^{-1}\left(\frac{r_2}{r_1}\right)$$

$$\phi_S = sin^{-1}\left(\frac{r_3}{\sqrt{r_1^2 + r_2^2 + r_3^2}}\right)$$

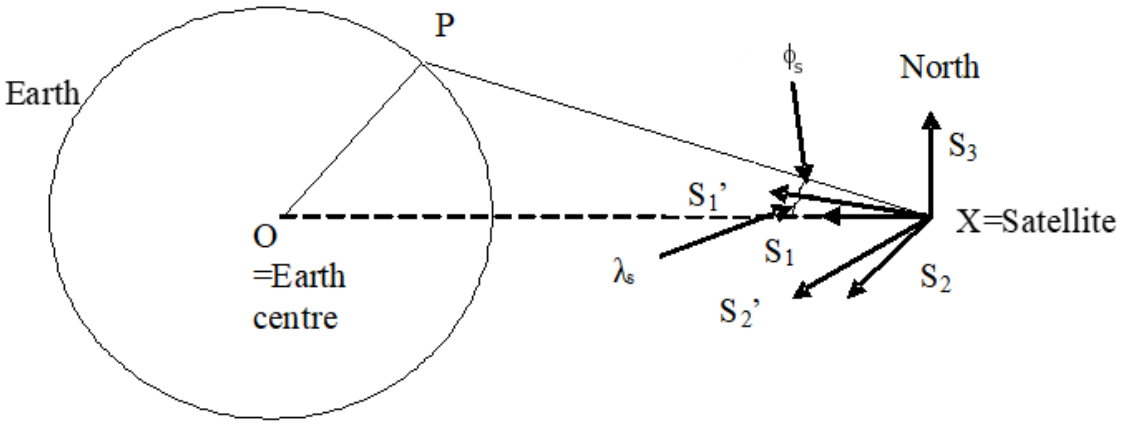


Figure 13: Angular Definition of the Reference Grid

Figure 13 shows the angular definition of the reference grid where:

- the frame (S_1, S_2, S_3) has its origin at the satellite position, (S_3) points northwards, and (S_1) directs to the centre of the Earth
- the vector r of coordinates (r_1, r_2, r_3) in the frame (S_1, S_2, S_3) is a pixel line of sight vector with r = XP/norm(XP)

In order to geolocate the radiances, the user must first calculate the corresponding azimuth, elevation coordinate for each row and column pixel, and then calculate the corresponding latitude, longitude coordinate from the azimuth, elevation information. This is described in the following.
Let (r, c) be the coordinates (row and column) of any pixel of the product. Row and columns are counted increasingly when going from bottom to up (south to north) and left to right (west to east) and beginning at 1. Therefore, the South West corner of the grid has coordinates (1, 1). The correspondence between the row and column position (r, c) and the azimuth and elevation position (λ, ϕ) is written:

$$\lambda_s = \lambda_0 - (c\text{-}1) \cdot Azimuth\_Grid\_Sampling$$
$$\phi_s = \phi_0 + (r\text{-}1) \cdot Elevation\_Grid\_Sampling$$

where:

1. *Azimuth_Grid_Sampling* and *Elevation_Grid_Sampling* are the reference grid spatial sampling angles, representing viewing angle increments between pixels in the W-E and S-N directions, respectively. The corresponding values are given in Table 11.
2. λ_0 and ϕ_0 are the angles from the centre of the projection to the centre of the *pixel* in the first *row* and first *column* of the *reference grid, respectively*. Note that the first row, column of the reference grid is indexed (1, 1). The values correspond to *Azimuth_Grid_Sampling* * (columns –1)/2 or *Elevation_Grid_Sampling* * (rows – 1)/2, respectively, and are given in Table 11, too.

Note that the E-W viewing angle (λ_0) does not correspond to the standard definition of azimuth, for an observation from the instrument perspective, which runs from negative to positive from West to East. Instead, it runs from negative to positive from East to West.

The N-S viewing angle corresponds to the standard definition of elevation, for an observation from the instrument perspective.

SSD (km)	λ_0		ϕ_0		Grid Sampling		Columns in Full Disc	Rows in Full Disc
	degrees	radians	degrees	radians	degrees	radians		
2	8.91303970830389	1.55561889270898E-01	-8.91303970830389E	-1.55561889270898E-01	3.20209797316468E-03	5.58871526031607E-05	5568	5568

Table 11 Values for the LI Level 2 accumulation grid

With these values, the coordinates of the Earth centre (origin of the projection) are (2784.5, 2784.5).

7.8.3 Normalized Geostationary Projection

The virtual *satellite* is in a geostationary orbit, perfectly located in the Equator plane at the given longitude, λ_D (normally 0 deg). Assuming all trignometric values are in degrees, the transformation from satellite scanning angles (λ_s, ϕ_s) to geographical coordinates (lon, lat) is given by the inverse projection function:

$$\begin{pmatrix} lon \\ lat \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{s_2}{s_1}\right) + \lambda_D \\ \arctan\left(s_4 \cdot \frac{s_3}{s_{xy}}\right) \end{pmatrix}$$

where:

$$s_1 = h - s_n \cdot \cos(\lambda_s) \cdot \cos(\phi_s)$$

$$s_2 = -s_n \cdot \sin(\lambda_s) \cdot \cos(\phi_s)$$

$$s_3 = s_n \cdot \sin(\phi_s)$$

$$s_4 = \frac{r_{eq}^2}{r_{pol}^2}$$

$$s_5 = \left(h^2 - r_{eq}^2\right)$$

$$s_{xy} = \sqrt{s_1^2 + s_2^2}$$

$$s_n = \frac{h \cdot \cos(\lambda_s) \cdot \cos(\phi_s) - s_d}{\cos^2(\phi_s) + s_4 \cdot \sin^2(\phi_s)}$$

$$s_d = \sqrt{(h \cdot \cos(\lambda_s) \cdot \cos(\phi_s))^2 - (\cos^2(\phi_s) + s_4 \cdot \sin^2(\phi_s)) \cdot s_5}$$

The shape of the Earth is described by an oblate ellipsoid with a single flattening parameter f

$$f = \frac{r_{eq} - r_{pol}}{r_{eq}}$$

where r_{eq} and r_{pol} denote the equatorial and polar radius of the Earth, respectively. The appropriate values for the Earth are f = 1/298.257223563 and r_{eq} = 6378.137 km. The parameter h in the equations above refers to the geostationary radius. The geostationary radius is the distance from the Earth's centre to the satellite in geostationary orbit and can be calculated from the sum of the geostationary altitude (35786.4 km) and the equatorial Earth radius r_{eq}.

8 LI LEVEL 2 PRODUCT USAGE

8.1 Reference Grid of the Accumulated Products

Accumulated products do not have associated geolocation coordinate variables in order to reduce the size of the product (see Section 5.2). Pixel positions are provided as x and y coordinates. Their raw counts (packed values) correspond to column and row numbers, respectively. After applying the scale_factor and add_offset as described in Section 7.5, the azimuth (x) and elevation (y) are obtained. Knowing the azimuth and elevation angles, geographical (lat,lon) coordinates can be calculated with the formula provided in Section 7.8.3. In addition, the NetCDF Climate and Forecast (CF) convention grid_mapping variables for the geostationary projection are included in the products to allow CF-Convention-aware tools to geolocate the grid_mapping associated variables (cf. variable mtg_geos_projection). To allow users to check their readers and implementations, lat/lon arrays of the FCI Level 1c reference grids are available from summer 2022 on the EUMETSAT webpage for download (cf. https://www.eumetsat.int/mtg-test-data).

8.2 Pixel size and LI Level 2 products

LI-2-LGR-x-FD-x and LI-2-LFL-x-FD-x products contain two variables that provide the size of each group and flash in units of LI pixels (see Section 3.2.1 for the definition of the LI pixel grid), i.e., `number_of_events` (see Table 2) and `flash_footprint` (see Table 3), respectively. In order to compute the exact physical size of a group/flash, users should use the information available in the LI-2-LEF-x-FD-x product. In fact, one can derive such descriptor only by knowing which events compose a group/flash and employ the physical size of the projection of each event on the Earth surface for the computation. Alternatively, one can use an approximation based on the assumption that, over the typical extention of a group/flash, the size of LI pixels does not change significantly. The user shall just use the size of the Earth-projection of the LI pixel at the location of the group/flash (i.e., `latitude` and `longitude`, see Table 2 and Table 3), and use this pixel size to derive the size of the group/flash as this value times `number_of_events` for groups and `flash_footprint` for flashes, respectively. The auxiliary file to derive the LI pixel size for a specific geolocation will be delivered to users through the EUMETSAT public page.
Accumulated flash products provide users with three different measurements per pixel in the FCI 2 km geostationary grid (see Section 5.2). The size of the Earth-projected pixel varies with the geolocation. The auxiliary files needed to describe the latitude and longitude values of the grid and the physical size of the pixels as a function of the geolocation will be delivered to users through the EUMETSAT public page.

8.3 Comparing/combining LI data to external lightning location systems

MTG LI data shall be regarded as complementary measurements to both space-based and ground-based lightning products/measurements produced by other systems. For this reason, it is worth exaplaining what LI will bring in the field of lightning diagnostic and how to relate/compare the “key variables” in the products (see Section 5) to other data.
The LI-2-LGR-x-FD-x product contains LI groups (see Section 5.1). These are closely related to GLM and ISS-LIS groups. In fact, the definition of a group is shared by the three systems: collections of pixel-based lightning events that are acquired within the same acquisition frame and are spatially clustered. LI groups provide users with information about the time-

slicing imaging (over the LI acquisition time, i.e., one millisecond) of lightning optical emissions. When comparing LI groups with either GLM or ISS-LIS groups, users must consider the differences in design between instruments, such as, integration time, spatial sampling/resolution. Both GLM and ISS-LIS acquire over two millisecond. When observing the same storm, this difference in design can potentially create considerable differences in the total number of groups, as well as differences between the acquisition times of the groups. In addition, differences will be found also for the geolocation of groups. In general, the discrepancies mentioned above are expected to be of the order of few millisecond for the group time and of the order of few kilometres for the group geolocation.
Ground-based systems are composed by radio antennas and provide users with strokes (CG lightning) and pulses (CC and IC lightning). CG strokes present times and locations of CG lightning ground contact (strike) points and are easily detected by modern ground-based systems. In contrast, only a minority of CC and IC lightning pulses are normally detected as their radio emissions are weaker and harder to process. Ground-based systems tend to be more sensitive to powerful and (vertically) extensive CC and IC pulses. In order to understand how to relate LI groups to ground-based strokes, one must can consider the following example. A stroke, located by a ground-based system, produces an optical pulse that is obsevable at the top of the cloud after multiple scattering through the cloud itself. The scattering causes the optical pulse to have a certain spatial extension, duration, and radiance. Space-based lightning imagers sense such optical emission at the top of the cloud; depending on the properties of such emission, several groups can be observed (due to time-slicing and spatial sampling). In addition to this, space-based lightning imager are more sensitive to IC+CC lighting events. Therefore, when observing the same storm, the number of groups provided by space-based systems exceeds the number of strokes and pulses from ground-based systems (see Figure 14). For example, from a rough evaluation, the ratio between the number of GLM groups and GLD-360 strokes is 7:1.

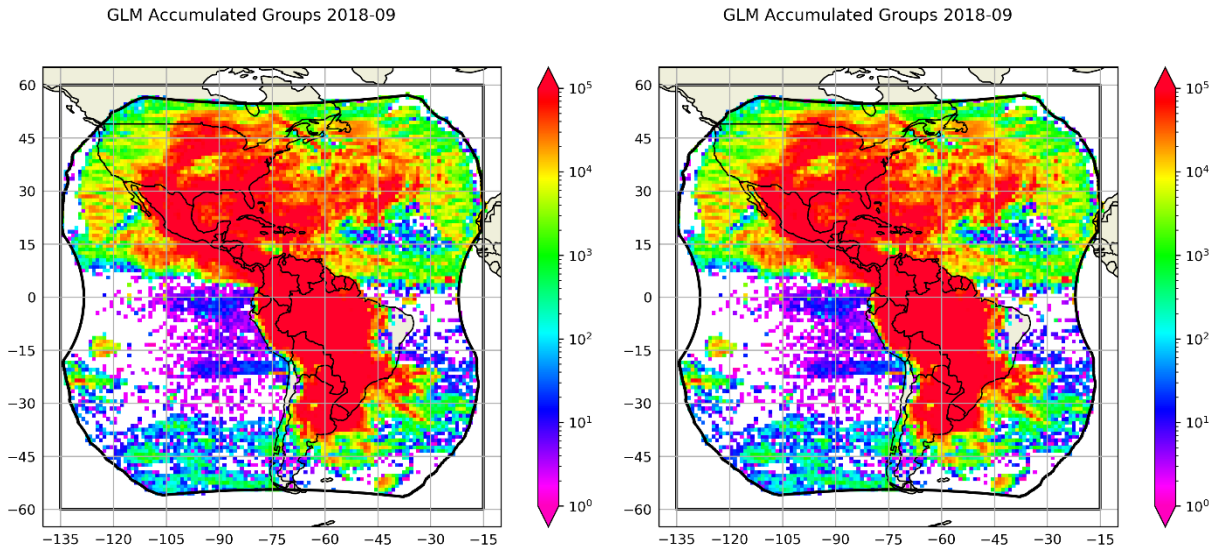


Figure 14: maps of accumulated GLM groups (left) and GLD360 strokes (right) for the month of September 2018.

The LI-2-LFL-x-FD-x product contains LI flashes (see Section 5.1). The definition of a flash is shared by LI and GLM: collections of groups that are correlated in space and time within

the two "windows" of 330 millisecond (temporal window) and 16.5 km (space window), respectively. Even if the definition of a flash is not uniform among all lightining location systems, the simple fact that a flash is a collection of groups/strokes correlated in space and time somewhat mitigates the differences in the way different types of lightning sensors interpret different lightning processes. This makes flash datasets of different lightning location systems more comparable than group/stroke datasets. In order to understand why this is, the reader can refer to the diagram in Figure 15. At the top of the diagram, one finds the black time axis on which lightning strokes and pulses are marked (these can be of any type; i.e., CG, CC, or IC). Below this, three additional time axes are represented, these are meant to show the temporal occurrence/detection of:

1. Ground-based strokes (yellow);
2. LI groups (green);
3. GLM groups (purple).

The ground-based system detects three strokes, while the two space-based instruments detect a much larger number of groups (this is meant to be representative of the difference presented in Figure 14). LI and GLM detect groups that are anyway different in number and temporal distribution. Each system reports different times for the beginning of the flash (B) and end of the flash (E), as well as average flash location. Despite these differences, the three systems detect one flash each. That is why, when comparing GLD360 flashes and GLM flashes as done in Figure 16, one finds a much better agreement compared to the group to stroke comparison over the same time (Figure 14).

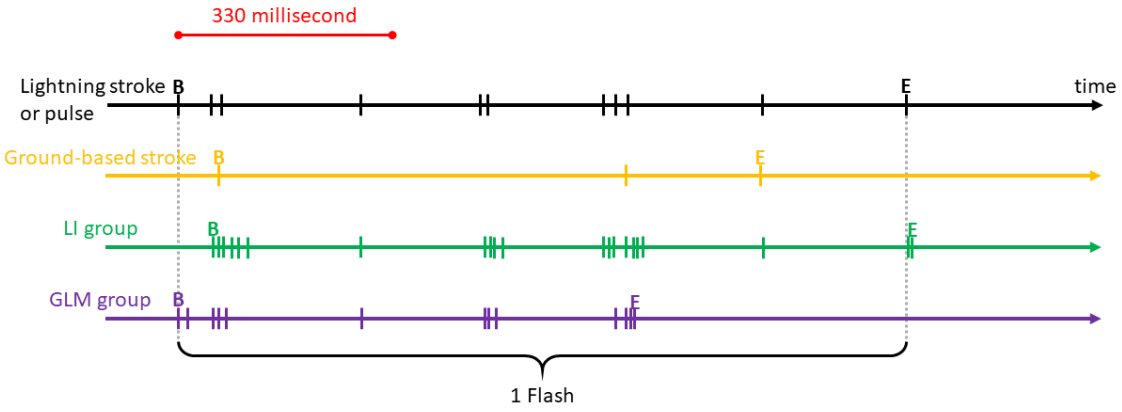


Figure 15: Simplified representation of the collection of groups/strokes into flashes from different types of lightning location systems (i.e., ground-based and space-based). The diagram represents only the group/stroke time with the assumption that groups/strokes are also correlated in space. The letters B and E mark the Begin and End of the flas, respectively.

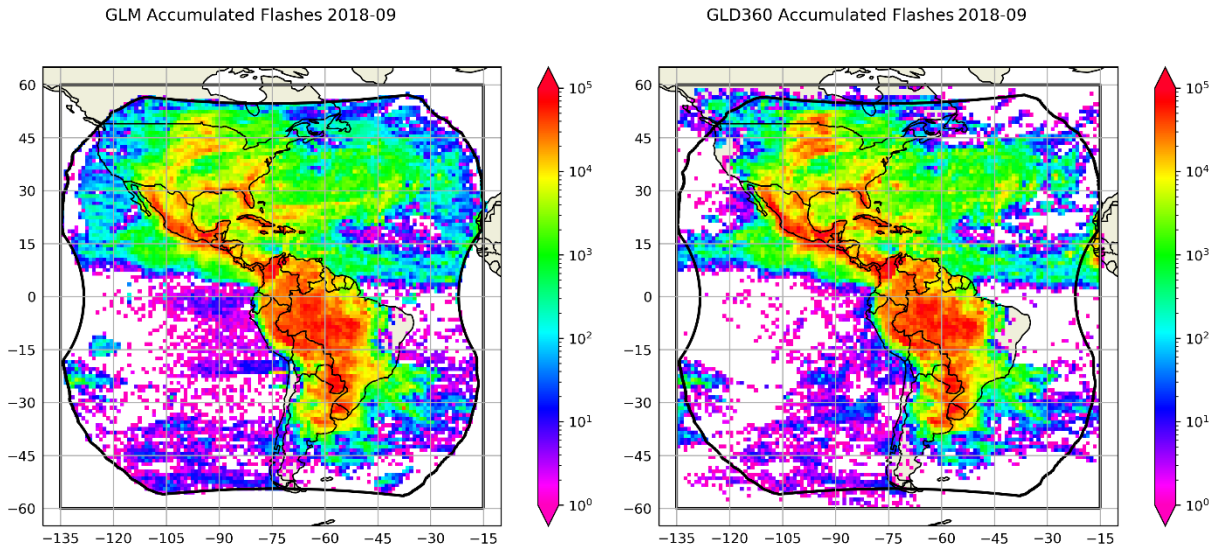


Figure 16: maps of accumulated GLM flashes (left) and GLD360 flashes (right) for the month of September 2018.

From this stems that:

1. LI-2-LGR-x-FD-x is a product that provides users with data that represent the LI capability of detecting optical pulses due to lightning activity down to its integration time (1 millisecond). Since the content of the group product is strongly influenced by the LI design, its use in combination with data from other systems can be challenging.
2. LI-2-LFL-x-FD-x is a product that provides users with data that are less reprentative of the full LI optical pulse diagnostic capabilities, but comparable/relatable to data from other lighitning location systems.

It is very important to stress that when combining LI groups/flashes with data from other lightning location systems, users need to use two types of corrections:

1. Parallax correction, to be applied to the location of groups/flashes (see Section 8.4.1 details). This correction is needed to mitigate the impact of the projection error on the Earth surface, due to the height of the optical pulse emission (of the order of 10 km), from the point of view of a GEO observer. In fact, when observing lightning away from the sub-satellite point, the parallax can introduce errors of the order to 10-20 km in locating the groups/flashes on the Earth surface. Correcting for the parallax is paramount to combine LI data with other lightning detection data. The error due to parallax can be comparable to the scale over which multiple storms can take place, especially in correspondence small cloud systems.
2. Time correction, to be applied to the time of groups/flashes (see Section 8.6 for details). This correction is needed to take into account the travel-time of photons emitted by the optical pulse from the top of the cloud to the LI instrument located in GEO orbit. In fact, the time information available in LI products stems from the on-board/acquisition time. Correcting for the travel-time of photons is paramount to combine LI data with other lightning location data. The correction is of the order 120 millisecond, comparable to the time scale over which multiple strokes/pulses can happen and a lot of groups can be detected, especially in active storms.

These two corrections allow one to achieve the best available time and space alignment between the LI dataset and actual lightning before using the data or performing any type of

additional processing; for example, in EUMETSAT such corrections will be applied prior to the "matching exercise" between LI data and external lightning location data for monitoring the LI detection performances.

LI-2-AFA-x-FD-x, LI-2-AF-x-FD-x, and LI-2-AF-x-FD-x are the three products that are providing users with imaging information from LI. In fact, these products contain spatially extended lightning measurements derived form the basic element of the LI lightning detection, i.e., LI (pixel-based) detections (see Sections 3.2.2 and 5.2). Accumulated products are one of the most important features of the LI mission and complement the standard group/flash products. In detail:

1. LI-2-AFA-x-FD-x provides user with data about the flash mapping by using the area covered by the optical emission of each flash in the LI-2-LFL-x-FD-x product. It is important to keep in mind that each flash is treated as a "flat" (uniform) optical emission in this product. LI-2-AFA-x-FD-x allows one to monitor the regions within a cloud top from which lightning-related optical emissions over 30 sec are emerging and accumulating and to know the number of flashes that where observed within the FCI grid pixels composing those regions. For example, from the LI-2-AFA-x-FD-x product one can derive the flash rate for each pixel of the FCI 2 km grid. This is a considerable improvement compared to the simple description of the flash by means of the variable `flash_footprint` available in the LI-2-LFL-x-FD-x product.
2. LI-2-AF-x-FD-x complements the LI-2-AFA-x-FD-x by providing one with the variation of the number of events within those regions reported to have lightning flashes in the LI-2-AFA-x-FD-x product. LI-2-AF-x-FD-x provides users with data about the mapping of number of LI events/detections rather than mapping of flashes. One shall keep in mind that the absolute value within each pixel of the LI-2-AF-x-FD-x has no real physical meaning; it is rather a proxy for the pixel-by-pixel variation of the number of events. It is worth noting that one can derive the flash rate over a region encompassing a complete lightning feature (not withing a FCI grid pixel) in the LI-2-AF-x-FD-x product; this stems from the definition in Section 5.2.
3. LI-2-AFR-x-FD-x is meant to describe the pixel-by-pixel variation of the optical emission accumulated over 30 sec within the FCI 2 km grid. This stems from the events contributing to the LI-2-AF-x-FD-x product (each one contributing with its own radiance) and it can be thought as the "appearance" of the accumulated optical emission over 30 sec as seen by LI.

GLM is also providing users in the US with "gridded products" (originally not in the product baseline). According to GLM experts, these are possibly the most used products from GLM. GLM accumulated products are computed over 5-10 min with a refreshing rate of about 1 min. LI accumulated products will be provided every 30 sec. These can be easily combined (accumulated) over a generic time, e.g., typical FCI repeat cycle of the order of 10 min. In Table 12 one finds the comparison between the LI Accumulated Products and the GLM Gridded Products. The table has been compiled with the help of GLM experts. It is worth noting that the most used GLM Gridded Products is the FED, directly related to the LI-2-AFA-x-FD-x.

Table 12: comparison between LI Accumulated Products and GLM Gridded Products

MTG LI Accumulated Products	**LI-2-AFA**	**LI-2-AFR**	**LI-2-AF**	**–**
GLM Gridded	**FED**	**TOE**	**–**	**MFA**

Products				
GLM Descriptive Name	Flash extent density	Total optical energy	TBD	Minimum flash area
Notes	Primary product used in US operations	GLM uses energy; LI uses radiance. Relative values in an image will be locally the same	Placeholder name: Event flash fraction or *illuminated flash fraction*	Preferable to average flash area – better signal to noise ratio
Product Construction	Count of flashes passing through each pixel	Total light emitted	Each flash adds +1 to the whole grid, weighted in each pixel by the number of events within the flash at that pixel	Area of the smallest flash passing through each pixel
Physical Basis	Visualizes and quantifies variation in local flash rates. Best-studied measure of mixed-phase region updraft variability.	Brighter flashes result from greater heating of the air via time-integrated current. What the instrument actually observes: total light above background.	Highlights regions of more frequent flickering within each flash.	Regions with active convective updrafts make small flashes, supporting convective/stratiform discrimination.

8.3.1 Parallax correction

The parallax effect causes a bias in the projection of lightning position to ground coordinates due to the fact that the optical pulses observed by LI are located at the top of clouds, not on the ground. This bias increases with the angular distance from nadir (or sub-satellite point). For instance, an optical pulse at 0 deg longitude and 46 deg latitude which is located at 12 km height above ground, is observed at 46.14 deg latitude. This difference in latitude corresponds to a bias of 15.5 km on the ground. Depending on the way a user wants to employ LI-2-LGR-x-FD-x and LI-2-LFL-x-FD-x products, it may be necessary to apply a parallax correction to the data. In detail, if the user wants to exploit the products in any application that requires information about lightning location on the Earth surface, the parallax correction will be necessary to make the best possible use of the products.
The best way to evaluate the height of the cloud top from which optical emissions observed by LI are emerging would be employing the latest available information from FCI Level 2 cloud products. However, this implies having access to the continuous stream of FCI products. As an alternative, EUMETSAT will provide users with a parallax correction derived from a 15-year-long Cloud Top Height (CTH) climatology derived from the SEVIRI products. In detail, the correction will be formatted as a portable/static Look-Up Table (LUT) from which the parallax correction for both latitude and longitude will be available per month and geo-location in the FOV. A Python script aimed at extracting the information from the LUT for a specific location and time will be shared with the users. Finally, additional information will be included in the LUT file, such as two types of CTH baed on different cloud types and the uncertainties on the CTH assessment. EUMETSAT considers this solution to be a good balance between simplicity-of-use and refinement for the description of the CTH. In the future, the parallax correction derived from FCI Level 2 products will be considered. Such information, delivered to users with a repeat cycle of 10 min, would provide one with near-real-time data for parallax correction.

8.3.2 Photon travel time correction

As stated in Section 8.4.1 for the parallax correction, if the user wants to exploit LI products in any application that requires information about lightning occurrence time on Earth, the correction for the time needed by photons emitted by optical pulses to reach the LI instrument (at about 36000 km above the equator) is needed. In spherical coordinates, the computation of the pixel-to-satellite distance uses standard spherical trigonometry. Non spherical geometry would give one corrections of the order of 10 km, which translate in a time-correction well below the required precision of 1 ms (i.e., the LI acquisition time; see 3.2.1). Finally, the computation uses the distance from the Earth surface rather than the cloud top given the small correction this would imply in terms of photon travel time. The auxiliary file needed to describe photon travel time correction as a function of latitude and longitude will be delivered to users through the EUMETSAT public page.

9 ARCHIVE QUICK-LOOKS

Archive Quick-Looks are images related to a product that enable fast viewing for the purpose of selection by archive users. These images are included when products are sent to the archive. Table 13 provides and overview of the archive image for each product defined in this format specification. If no archive quicklook is to be generated, then that is stated here. The naming of the quicklooks will follow the conventions defined in Section 6.

Product ID	Archive Quicklook Description
LI-2-LEF-x-FD-x	Bar plot of filtering stats with totals and from each algorithm.
LI-2-LGR-x-FD-x	Binned histogram of group sizes with overlaid stats.
LI-2-LFL-x-FD-x	Binned flash duration with overlaid stats.
LI-2-AFA-x-FD-x	500x500 pixel PNG formatted image showing accumulated product values overlaid on coastline with colour scale. The accumulation values should be binned into the smaller QL size and contain accumulation values covering the whole product accumulation interval.
LI-2-AF-x-FD-x	no – specific QL generated for this product. Link to AFA QL.
LI-2-AFR-x-FD-x	No specific QL generated for this product. Link to AFA QL.

Table 13 Archive quicklook description

APPENDIX A FORMAT DESCRIPTIONS

A.1 Common Definitions

A.1.1 Enumerated Types

Name	Type	Description	Value	String
boolean	ubyte	boolean type - two states		
			0	false
			1	true
trilean	ubyte	trilean type - three states		
			0	false
			1	true
			2	undefined
auxiliary_dataset_status_type	ubyte	Possible states for an auxiliary dataset used in processing		
		OK	0	OK
		dataset was used but was out of its stated validity time	1	out_of_validity_time
		auxiliary dataset was not available	2	not_available
detector_type	ubyte	Indicates which detector the data in the group corresponds to		
			1	detector_1
			2	detector_2
			3	detector_3
			4	detector_4

A.1.2 Bit Masks

Name	Type	Bit	String	Description	Meaning	Range
None defined						

A.2 LI-2-LEF-BODY

A.2.1 Group:root (/)

A.2.1.1 Dimensions

Name	Description	Type	Values	Shape
None defined				

A.2.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.2.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		

processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the	string		

	"special_compression" field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words "NO URL".			
disposition_mode	As per the dataset/product name field "disposition_mode" in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field "facility_or_tool" in dataset name	string		
environment	As per the dataset name field "environment" in dataset name	string		
references	"www.eumetsat.int"	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, "None."	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at	string		

	or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty	string		

	string.			
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.2.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
None defined					

A.2.2 Group:/state

A.2.2.1 Dimensions

Name	Description	Type	Values	Shape
None defined				

A.2.2.2 User Types

Name	Description	Type	Values	Shape
None defined				

A.2.2.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.2.2.4 Variables

Name	Attribute	Description	Type	Values	Shape
None defined					

A.2.3 Group:/state/processor

A.2.3.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved inm processing the dataset		<runtime_value>	

A.2.3.2 User Types

Name	Description	Type	Values	Shape

auxiliary_dataset_status_type	See Enum types	enum ubyte		

A.2.3.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.2.3.4 Variables

Name	Attribute	Description	Type	Values	Shape
auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.	string		auxiliary_dataset
auxiliary_dataset_status		See Enum types	auxiliary_dataset_status_type		auxiliary_dataset

A.2.4 Group:/data

A.2.4.1 Dimensions

Name	Description	Type	Values	Shape
detector	Fixed value of 4		4	

A.2.4.2 User Types

Name	Description	Type	Values	Shape
detector_type	See Enums worksheet	enum ubyte		

A.2.4.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.2.4.4 Variables

Name	Attribute	Description	Type	Values	Shape
l1b_missing_warning		Flag is set if there were gaps in the expected L1b	byte	runtime	1

		inputs.			
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if a configurable proportion of the input L1b events have any of the following l1b quality flags set: *pixel_scene_quality::geolocation_error *pixel_scene_quality::geolocation_warning	byte	runtime	1
	long_name		*string*	*"L1b event geolocation warning"*	
l1b_radiometric_warning		Set to 1 if a configurable proportion of the input L1b events have the l1b quality flag, pixel_detector_quality::radiometric_correction_warning, set.	byte	runtime	1
	long_name		*string*	*"L1b event radiometric warning"*	

A.2.5 Group:/data/<detector>

A.2.5.1 Dimensions

Name	Description	Type	Values	Shape
events	Number of L2 events in product (determined at runtime - <= #L1B events)		runtime	

A.2.5.2 User Types

Name	Description	Type	Values	Shape
None defined				

A.2.5.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.2.5.4 Variables

Name	Attribute	Description	Type	Values	Shape
Detector		See Enums worksheet	detector_type		1
	long_name		*string*	*"ID of detector for this group"*	
epoch_time		Time (UTC) of the first frame from the first full L1b chunk that contributed to this dataset.	double		1
	long_name		*string*	*"Start time of integration frame"*	

	standard_name		*string*	*"time"*	
	units		*string*	*"seconds since 2000-01-01 00:00:00.0"*	
	precision		*string*	*"1 millisecond"*	
	time_standard		*string*	*"UTC"*	
time_offset		Time offset of each event from epoch_time in seconds. This means events carried over from the previous processing cycle will have a negative offset.	float		events
	long_name		*string*	*"Time offset from epoch time"*	
	_FillValue		*float*	*NC_FILL_FLOAT*	
	units		*string*	*"seconds"*	
event_id		ID of L2 lightning event	uint		events
	long_name		*string*	*"ID of LI L2 Event"*	
detector_row		Detector row index of the central pixel of the event window	ushort		events
	long_name		*string*	*"Detector row position of event pixel"*	
	_FillValue		*ushort*	*NC_FILL_USHORT*	
	units		*string*	*"1"*	
detector_column		Detector column index of the central pixel of the event window	ushort		events
	long_name		*string*	*"Detector column position of event"*	
	_FillValue		*ushort*	*NC_FILL_USHORT*	
	units		*string*	*"1"*	
latitude		Latitude position of the central pixel of the event window. Expected to come from previous background image.	short		events
	long_name		*string*	*"Latitude of event"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_north"*	
	standard_name		*string*	*"latitude"*	
	scale_factor		*float*	*0.0027*	
	add_offset		*float*	*0*	
longitude		Longitude position of the central pixel of the event window. Expected to come from previous background image.	short		events
	long_name		*string*	*"Longitude of event"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_east"*	
	standard_name		*string*	*"longitude"*	
	scale_factor		*float*	*0.0027*	
	add_offset		*float*	*0*	
radiance		Radiance measurement for the event	ushort		events
	long_name		*string*	*"Radiance of event pixel"*	
	_FillValue		*ushort*	*NC_FILL_USHORT*	
	units		*string*	*"mW.m-2.sr-1"*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.0103*	
group_id		ID of L2 lightning group object in the associated LI-2-GR product	uint		events
	long_name		*string*	*"ID of associated LI L2 Group object"*	
flash_id		ID of L2 lightning flash object in the associated LI-2-LFL product	uint		events
	long_name		*string*	*"ID of associated LI L2 Flash object"*	

event_filter_qa		Quality Assurance value derived from the L1b event pre-filtering process, in the range 0 to 1 where 0 means low confidence, and 1 means high confidence.	ubyte		events
	long_name		*string*	*"L2 event pre-filtering quality assurance value"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.3 LI-2-LEF-TRAIL

A.3.1 Group:root (/)

A.3.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.3.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.3.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		

data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed	string		

	(2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group	string		

	accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		

creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.3.1.4 Variables

Name	Attribute	Description	Type	Values	Shape

available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of body chunk files produced for this product"*	

A.4 LI-2-LGR-BODY

A.4.1 Group:root (/)

A.4.1.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	
groups	Number of L2 groups in product (determined at runtime)		runtime	

A.4.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
auxiliary_dataset_status_type	See Enum types	enum byte		

A.4.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in	string		

	dataset name			
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	String		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	String		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	String		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact	string		

	internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	String		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating	string		

	the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following:	string		

	'person', 'group', 'institution', or 'position'.			
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.4.1.4 Variables

Name	Attribute	Description	Type	Values	Shape

auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.	string		auxiliary_dataset
	long_name		*string*	*"Auxiliary dataset identifier"*	
	title		*string*	*"Identifier of auxiliary dataset or type of auxiliary dataset used to produce this product"*	
auxiliary_dataset_status		Possible states for an auxiliary dataset 0 = OK 1 = dataset was used but was out of its stated validity time 2 = auxiliary dataset was not available	auxiliary_dataset_status_type		auxiliary_dataset
	long_name		*string*	*"Status of auxiliary dataset"*	
l1b_missing_warning		Flag is set if there were gaps in the expected L1b inputs.	byte	runtime	1
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if a configurable proportion of the input L1b events have any of the following l1b quality flags set: * pixel_scene_quality::geolocation_error * pixel_scene_quality::geolocation_warning	byte	runtime	1
	long_name		*string*	*"L1b event geolocation warning"*	
l1b_radiometric_warning		Set to 1 if a configurable proportion of the input L1b events have the l1b quality flag, pixel_detector_quality::radiometric_correction_warning, set.	byte	runtime	1
	long_name		*string*	*"L1b event radiometric warning"*	
group_time		Group time based on integration frame of contributing events	double	runtime	groups
	long_name		*string*	*"Start time of integration frame"*	
	standard_name		*string*	*"time"*	
	units		*string*	*"seconds since*	

				2000-01-01 00:00:00.0"	
	precision		*string*	*"1 millisecond"*	
	time_standard		*string*	*"UTC"*	
latitude		Average latitude position of the group	short	runtime	groups
	long_name		*string*	*"Latitude of group"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_north"*	
	standard_name		*string*	*"latitude"*	
	scale_factor		*float*	*0.00275*	
	add_offset		*float*	*0*	
longitude		Average longitude position of the group	short	runtime	groups
	long_name		*string*	*"Longitude of group"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_east"*	
	standard_name		*string*	*"longitude"*	
	scale_factor		*float*	*0.00275*	
	add_offset		*float*	*0*	
radiance		Radiance value for the group	ushort	runtime	groups
	long_name		*string*	*"Radiance of event pixel"*	
	_FillValue		*ushort*	*NC_FILL_USHORT*	
	units		*string*	*"mW.m-2.sr-1"*	
	scale_factor		*float*	*<configured_value>*	
	add_offset		*float*	*<configured value>*	
group_id		ID of L2 lightning group object in this product	uint	runtime	groups
	long_name		*string*	*"LI L2 Group IDs"*	
flash_id		ID (per group in this product) of L2 lightning flash object linked to the associated LI-2-LFL product	uint	runtime	groups
	long_name		*string*	*"ID of associated LI L2 Flash object with each group"*	
number_of_events		Number of LI L2 events linked to each group.	ushort	runtime	groups
	long_name		*string*	*"Number of events in each group"*	
group_filter_qa		Quality assurance value derived from the L2 group filtering process, in range 0-1 where 0=low confidence and 1=highest confidence.	ubyte	runtime	groups
	long_name		*string*	*"L2 filtered group quality assurance value"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.5 LI-2-LGR-TRAIL

A.5.1 Group:root (/)

A.5.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.5.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.5.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMET SAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		

subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.	string		
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_ value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this	string		

	would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)			
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_ YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		

subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		

geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.5.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of body chunk files produced for this product"*	

A.6 LI-2-LFL-BODY

A.6.1 Group:root (/)

A.6.1.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	
truncated_flash	Number of flashes forced to complete.		runtime	
flashes	Number of L2 flashes in product (determined at runtime)		runtime	

A.6.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
auxiliary_dataset_status_type	See Enum types	enum byte		

A.6.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		

data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	String		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	String		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed	string		

	(2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group	string		

	accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		

creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.6.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset.	string		auxiliary_dataset

		If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.			
	long_name		*string*	*"Auxiliary dataset identifier"*	
	title		*string*	*"Identifier of auxiliary dataset or type of auxiliary dataset used to produce this product"*	
auxiliary_dataset_status		Possible states for an auxiliary dataset 0 = OK 1 = dataset was used but was out of its stated validity time 2 = auxiliary dataset was not available	auxiliary_dataset_status_type		auxiliary_dataset
	long_name		*string*	*"Status of auxiliary dataset"*	
l1b_missing_warning		Flag is set if there were gaps in the expected L1b inputs.	byte	runtime	1
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if a configurable proportion of the input L1b events have any of the following l1b quality flags set: * pixel_scene_quality::geolocation_error * pixel_scene_quality::geolocation_warning	byte	runtime	1
	long_name		*string*	*"L1b geolocation warning"*	
l1b_radiometric_warning		Set to 1 if a configurable proportion of the input L1b events have the l1b quality flag, pixel_detector_quality::radiometric_correction_warning, set.	byte	runtime	1
	long_name		*string*	*"L1b radiometric warning"*	
flash_time		Flash time based on first contributing event	double		flashes
	long_name		*string*	*"Nominal flash time"*	
	standard_name		*string*	*"time"*	
	units		*string*	*"seconds*	

				since 2000-01-01 00:00:00.0"	
	precision		*string*	*"1 millisecond"*	
	time_standard		*string*	*"UTC"*	
latitude		Latitude position of the flash	short		flashes
	long_name		*string*	*"Latitude of flash"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_north"*	
	standard_name		*string*	*"latitude"*	
	scale_factor		*float*	*0.00275*	
	add_offset		*float*	*0*	
longitude		Longitude position of the flash	short		flashes
	long_name		*string*	*"Longitude of flash"*	
	_FillValue		*short*	*NC_FILL_SHORT*	
	units		*string*	*"degrees_east"*	
	standard_name		*string*	*"longitude"*	
	scale_factor		*float*	*0.00275*	
	add_offset		*float*	*0*	
radiance		Radiance value for the flash	ushort		flashes
	long_name		*string*	*"Radiance of flash"*	
	_FillValue		*ushort*	*NC_FILL_USHORT*	
	units		*string*	*"mW.m-2.sr-1"*	
	scale_factor		*float*	*<configured_value>*	
	add_offset		*float*	*<configured_value>*	
flash_id		ID of L2 lightning flash objects in this product	uint		flashes
	long_name		*string*	*"LI L2 Flash IDs"*	
number_of_groups		Number of LI L2 groups linked to each flash	uint		flashes
	long_name		*string*	*"Number of groups in each flash"*	
number_of_events		Number of LI L2 events linked to each flash	ushort		flashes
	long_name		*string*	*"Number of events in each flash"*	
flash_duration		Duration between earliest and latest events in each flash (ms)	ushort		flashes
	long_name		*string*	*"Flash duration"*	
	units		*string*	*"ms"*	
flash_footprint		Number of LI pixels composing the flash footprint.	ushort		flashes
	long_name		*string*	*"Flash*	

				footprint size"	
truncated_flashes		indices of flashes that were forced to complete because their duration exceeded a maximum threhsold.	ushort		truncated_flash
	long_name		*string*	*"Truncated flash indices"*	
flash_filter_confidence		Confidence value derived from the L2 flash filtering process, in the range 0-1 where 0 means high confidence of a true flash and 1 means low confidence of a true flash	byte	runtime	flashes
	long_name		*string*	*"L2 filtered flash confidence"*	
	_FillValue		*byte*	*NC_FILL_BYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.7 LI-2-LFL-TRAIL

A.7.1 Group:root (/)

A.7.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	
time_bins	Number of 'per second' bins for relevant hsitograms. Depends on pseudo_RC duration + time slots from previous RC carried over.	byte	601	
filters	Number of group and flash analysis filters	byte	17	
event_energy_bins	Number of 'event energy' bins for historgram	byte	runtime	
event_bkg_radiance_bins	Number of 'event background radiance' bins for historgram	byte	runtime	
group_total_energy_bins	Number of 'group total energy' bins for histogram	byte	runtime	
group_area_bins	Number of 'group area' bins for histogram	byte	runtime	
flash_total_energy_bins	Number of 'flash total energy' bins for histogram	byte	runtime	
flash_area_bins	Number of 'flash area' bins for histogram	byte	runtime	
flash_duration_bins	Number of 'flash duration' bins for histogram	byte	runtime	
groups_per_flash_bins	Number of 'group-per-flash' bins for histogram	byte	runtime	
events_per_flash_bins	Number of 'Events-per-flash' bins for histogram	byte	runtime	
events_per_group_bins	Number of 'Events-per-group' bins for histogram	byte	runtime	
group_sobel_bins	Number of 'Group Average Sobel' bins for histogram	byte	runtime	
flash_sobel_bins	Number of 'Flash Average Sobel' bins for histogram	byte	runtime	

A.7.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.7.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is	string	e.g."CF-1.7"	

	applicable to netCDF4.			
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	"original generated file"	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	"EUMETSAT"	
location_indicator	As per the dataset name field" location_indicator" in dataset name	string		
data_designator	As per the dataset name field "data_designator" in dataset name	string		
platform	As per the dataset name field "spacecraft" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field "data_source" in dataset name	string		
processing_level	As per the dataset name field "level" in dataset name	string		
coverage	As per the dataset name field "coverage" in dataset name	string		
type	As per the dataset name field "type" in dataset name	String		
subtype	As per the dataset name field "subtype" in dataset name	string		
component1	As per the dataset name field "component1" in dataset name	string		
component2	As per the dataset name field "component2" in dataset name	string		
component3	As per the dataset name field "component3" in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		

format_version	Format version of the dataset/product.	String		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.	string		
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute	string		

	Convention for Dataset Discovery			
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the "instrument configuration identifier" from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two "100" entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the "instrument configuration identifier version" from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention	string		

	name for the file			
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in	string		

	date_time_position.			
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.7.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of body chunk files produced for this product"*	
Level 2_flashes_per_second		Total number of 'true' flashes per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of flashes per second"*	
Level 2_groups_per_second		Total number of 'true' groups per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of groups per second"*	
Level 2_events_per_second		Total number of 'true' events per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of events per second"*	
Level 2_single_group_flash_per_second		Total number of 'true' one-group-flashes per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of single group flashes per second"*	
Level 2_fragmented_flash_per_second		Total number of 'true' flashes-with-multiple-subfootprints per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of flashes with fragmented footprint per second"*	

Level 2_small_flash_per_second		Total number of 'true' flashes-wth-small-footprint per second	ushort	runtime	time_bins
	long_name		*string*	*"Number of flashes with small footprint per second"*	
filter_rejection_rates		Fraction of 'pre-filtered' L1b events that are rejected by each of the group and flash analysis steps relative to the number of groups or flashes in the reporting period	byte	runtime	filters
	long_name		*string*	*"Rejection rates for each L2 filter"*	
	filter_names		*string*	*"G_L2Part G_L2SaturRad G_L2RelSobel G_L2DTPeaks G_L2Rad G_L2Size G_L2QA G_L2EarlyRej F_L2SingleGroups F_L2Groups F_L2Foot F_L2TimeCorr F_L2DistCorr F_L2AvgRelSobel F_L2QA F_Reintroduction F_DuplicateFlash "*	
group_rejection_rate		Fraction of rejected groups wrt the total number of identified groups in the reporting period	byte	runtime	1
	long_name		*string*	*"Overall group rejection rate"*	
flash_rejection_rate		Fraction of rejected flashes wrt the total number of identified flashes in the reporting period	byte	runtime	1
	long_name		*string*	*"Overall flash rejection rate"*	
average_group_size		Average number of events per group	byte	runtime	1
	long_name		*string*	*"Average number of events per group"*	
average_bkg_gradient_groups		Average background gradient in group	byte	runtime	1
	long_name		*string*	*"Average background gradient in group"*	
average_bkg_gradient_flashes		Average background	byte	runtime	1

		gradient in flashes			
	long_name		*string*	*"Average background gradient in flashes"*	
event_energy_ranges		Event energy bin boundaries. The bin 'n' is defined from (>=start[n] : < start[n-1]). The last bin is defined as (>= start[n])	float	runtime	event_energy_bins
	long_name		*string*	*"Event Energy bin boundaries"*	
	units		*string*	*"J/(m2.sr)"*	
event_energy_histogram		Frequency of events with energy values in each defined bin range.	ushort	runtime	event_energy_bins
	long_name		*string*	*"Frequency of events with energy in defined ranges"*	
event_bkg_radiance_ranges		Event background radiance bin boundaries. The bin 'n' is defined from (>=start[n] : < start[n-1]). The last bin is defined as (>= start[n])	float	runtime	event_bkg_radiance_bins
	long_name		*string*	*"Event bacground radiance bin boundaries"*	
	units		*string*	*"mW.m-2.sr-1"*	
event_bkg_radiance_histogram		Binned event background radiance values per defined bin range of radiances.	ushort	runtime	event_bkg_radiance_bins
	long_name		*string*	*"Frequency of event background radiances in defined ranges"*	
	bin_boundaries		*string*	*"0 50 100 150 …. "*	
	bin_units		*string*	*"mW.m-2.sr-1"*	
group_total_energy_ranges		Group energy bin start values.	ushort	runtime	group_total_energy_bins
group_total_energy_histogram		Binned group total energy values per defined bin range of energy values.	ushort	runtime	group_total_energy_bins
group_area_ranges		Group area bin start values	ushort	runtime	group_area_bins
group_area_histogram		Binned group area (km^2) values per defined bin range.	ushort	runtime	group_area_bins
flash_total_energy_ranges		Flash energy bin start values.	ushort	runtime	flash_total_energy_bins
flash_total_energy_histogram		Binned flash total	ushort	runtime	flash_total_energ

		energy values per defined bin range of energy values.			y_bins
flash_area_ranges		Flash area bin start values	ushort	runtime	flash_area_bins
flash_area_histogram		Binned flash area (km^2) values per defined bin range.	ushort	runtime	flash_area_bins
flash_duration_ranges		Flash duration bin start values	ushort	runtime	flash_duration_bins
flash_duration_histogram		Flash duration (ms) values per defined bin range.	ushort	runtime	flash_duration_bins
groups_per_flash_ranges		Group-per-flash count bin start values	ushort	runtime	groups_per_flash_bins
groups_per_flash_histogram		group-per-flash values per defined bin range.	ushort	runtime	groups_per_flash_bins
events_per_flash_ranges		Events-per-flash count bin start values	ushort	runtime	events_per_flash_bins
events_per_flash_histogram		Events-per-flash values per defined bin range.	ushort	runtime	events_per_flash_bins
events_per_group_ranges		Events-per-group count bin start values	ushort	runtime	events_per_group_bins
events_per_group_histogram		Events-per-group values per defined bin range.	ushort	runtime	events_per_group_bins
group_sobel_ranges		Group-average-sobel count bin start values	ushort	runtime	group_sobel_bins
group_sobel_histogram		Group-average-sobel values per defined bin range.	ushort	runtime	group_sobel_bins
flash_sobel_ranges		Flash-average-sobel count bin start values	ushort	runtime	flash_sobel_bins
flash_sobel_histogram		Flash-average-sobel values per defined bin range.	ushort	runtime	flash_sobel_bins

A.8 LI-2-AF-BODY

A.8.1 Group:root (/)

A.8.1.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	
pixels	Number of reference grid points containing lightning data		runtime	
accumulations	Number of accumulated lightning datasets stored in this product. For dissemination this is generally 1, and for archive this is generally between 2 and 20 (covering 1-10 minutes).		runtime	

A.8.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
auxiliary_dataset_status_type	See Enum types	enum byte		

A.8.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMET SAT”	

location_indicator	As per the dataset name field" location_indicator" in dataset name	string		
data_designator	As per the dataset name field "data_designator" in dataset name	string		
platform	As per the dataset name field "spacecraft" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field "data_source" in dataset name	string		
processing_level	As per the dataset name field "level" in dataset name	string		
coverage	As per the dataset name field "coverage" in dataset name	string		
type	As per the dataset name field "type" in dataset name	string		
subtype	As per the dataset name field "subtype" in dataset name	string		
component1	As per the dataset name field "component1" in dataset name	string		
component2	As per the dataset name field "component2" in dataset name	string		
component3	As per the dataset name field "component3" in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	String		
time_coverage_start	As per the dataset name field "start_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field "end_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field "processing_mode" in dataset name	string		
special_compression	As per the dataset name field "special_compression" in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the "special_compression" field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal	string		

	compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_ value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	String		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_ <subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain	string		

	DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type	string		

	specified by the creator_type attribute) principally responsible for creating this data.			
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.8.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of	string		auxiliary_dataset

		lower case x characters.			
	long_name		*string*	*"Auxiliary dataset identifier"*	
	title		*string*	*"Identifier of auxiliary dataset or type of auxiliary dataset used to produce this product"*	
auxiliary_dataset_status		Possible states for an auxiliary dataset 0 = OK 1 = dataset was used but was out of its stated validity time 2 = auxiliary dataset was not available	auxiliary_dataset_status_type		auxiliary_dataset
	long_name		*string*	*"Status of auxiliary dataset"*	
mtg_geos_projection		Grid mapping variable for the MTG GEOS Projection	int	runtime	1
	long_name		*string*	*"MTG geostationary projection"*	
	grid_mapping_name		*string*	*"geostationary"*	
	perspective_point_height		*double*	*35786400*	
	semi_major_axis		*double*	*6378137*	
	semi_minor_axis		*double*	*6356752.31424518*	
	inverse_flattening		*double*	*298.257223563*	
	latitude_of_projection_origin		*double*	*0*	
	longitude_of_projection_origin		*double*	*0*	
	sweep_angle_axis		*string*	*"y"*	
accumulation_start_times		Start time for each accumulation dataset.	double	runtime	accumulations
	long_name		*string*	*"Accumulation start time"*	
	units		*string*	*"seconds since 2000-01-01 00:00:00.0"*	
	precision		*string*	*"1 millisecond"*	
accumulation_offsets		One offset per accumulation dataset. These provide index into the arrays of dimension 'pixel'	uint	runtime	accumulations
x		X coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"azimuth angle encoded as column"*	
	axis		*string*	*"X"*	
	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_x_coordinate"*	
	scale_factor		*double*	-	

				5.588715260 31607e-05	
	add_offset		*double*	*0.155617776 423501*	
	valid_range		*string*	*1, 5568*	
y		Y coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"zenith angle encoded as row"*	
	axis		*string*	*"Y"*	
	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_y_coordinate"*	
	scale_factor		*double*	*5.588715260 31607e-05*	
	add_offset		*double*	*-0.155617776 423501*	
	valid_range		*string*	*1, 5568*	
flash_accumulation		Flash accumulation value per contributing pixel on the reference grid.	ushort	runtime	pixels
	long_name		*string*	*"Per area accumulation of flashes"*	
	grid_mapping		*string*	*"mtg_geos_projection"*	
	units		*string*	*"flashes/pixel"*	
	coordinate		*string*	*"sparse: x y"*	
l1b_missing_warning		Flag is set if there were gaps in the expected L1b inputs for each of the contribnuting accumulation periods.	byte	runtime	accumulations
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag	byte	runtime	accumulations
	long_name		*string*	*"L1b event geolocation warning"*	
l1b_radiometric_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag in the accumulation period	byte	runtime	accumulations
	long_name		*string*	*"L1b event radiometric warning"*	
average_flash_qa		Average QA value from contributing flashed in the accumulation period. 0=higher probability true; 1=higher false probability.	ubyte	runtime	accumulations
	long_name		*string*	*"average flash confidence value"*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.9 LI-2-AF-TRAIL

A.9.1 Group:root (/)

A.9.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.9.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.9.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMET SAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		

data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	String		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	String		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	String		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of:	string		

	(1) an internal compression code as listed in the "special_compression" field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words "NO URL".			
disposition_mode	As per the dataset/product name field "disposition_mode" in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_ value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field "facility_or_tool" in dataset name	string		
environment	As per the dataset name field "environment" in dataset name	string		
references	"www.eumetsat.int"	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, "None."	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type >_<subtype>_YYYY_DDD_NNNN_<release_versio n> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no	string		

	repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		

creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.9.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of*	

				body chunk files produced for this product"	

A.10 LI-2-AFA-BODY

A.10.1 Group:root (/)

A.10.1.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	
pixels	Number of reference grid points containing lightning data		runtime	
accumulations	Number of accumulated lightning datasets stored in this product. For dissemination this is generally 1, and for archive this is generally between 2 and 20 (covering 1-10 minutes).		runtime	

A.10.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
auxiliary_dataset_status_type	See Enum types	enum byte		

A.10.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	

location_indicator	As per the dataset name field" location_indicator" in dataset name	string		
data_designator	As per the dataset name field "data_designator" in dataset name	string		
platform	As per the dataset name field "spacecraft" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field "data_source" in dataset name	string		
processing_level	As per the dataset name field "level" in dataset name	string		
coverage	As per the dataset name field "coverage" in dataset name	string		
type	As per the dataset name field "type" in dataset name	string		
subtype	As per the dataset name field "subtype" in dataset name	string		
component1	As per the dataset name field "component1" in dataset name	string		
component2	As per the dataset name field "component2" in dataset name	string		
component3	As per the dataset name field "component3" in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	String		
time_coverage_start	As per the dataset name field "start_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field "end_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field "processing_mode" in dataset name	string		
special_compression	As per the dataset name field "special_compression" in dataset name	string		

subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the "special_compression" field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words "NO URL". If the value is "MULTI", then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the "special_compression" field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words "NO URL".	string		
disposition_mode	As per the dataset/product name field "disposition_mode" in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field "facility_or_tool" in dataset name	string		
environment	As per the dataset name field "environment" in dataset name	string		
references	"www.eumetsat.int"	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, "None."	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version>	string		

	where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field			
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the "instrument configuration identifier" from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two "100" entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the "instrument configuration identifier version" from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		

mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual	string		

	observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.			
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.10.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.	string		auxiliary_dataset
	long_name		*string*	*"Auxiliary dataset identifer"*	
	title		*string*	*"Identifier of axuliary dataset or type of auxiliary dataset used to produce this product"*	
auxiliary_dataset_status		Possible states for an auxiliary dataset 0 = OK 1 = dataset was used but was out of its stated validity time 2 = auxiliary dataset was not available	auxiliary_dataset_status_type		auxiliary_dataset
	long_name		*string*	*"Status of auxiliary dataset"*	
mtg_geos_projection		Grid mapping variable for the MTG GEOS Projection.	int	runtime	1
	long_name		*string*	*"MTG geostationary projection"*	
	grid_mapping_name		*string*	*"geostationary"*	
	perspective_point_height		*double*	*35786400*	

	semi_major_axis		*double*	*6378137*	
	semi_minor_axis		*double*	*6356752.31424518*	
	inverse_flattening		*double*	*298.257223563*	
	latitude_of_projection_origin		*double*	*0*	
	longitude_of_projection_origin		*double*	*0*	
	sweep_angle_axis		*string*	*"y"*	
accumulation_start_times		Start time for each accumulation dataset.	double	runtime	accumulations
	long_name		*string*	*"Accumulation start time"*	
	units		*string*	*"seconds since 2000-01-01 00:00:00.0"*	
	precision		*string*	*"1 millisecond"*	
accumulation_offsets		One offset per accumulation dataset. These provide index into the arrays of dimension 'pixel'	uint	runtime	accumulations
x		X coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"azimuth angle encoded as column"*	
	axis		*string*	*"X"*	
	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_x_coordinate"*	
	scale_factor		*double*	*-5.58871526031607e-05*	
	add_offset		*double*	*0.155617776423501*	
	valid_range		*string*	*1, 5568*	
y		Y coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"zenith angle encoded as row"*	
	axis		*string*	*"Y"*	
	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_y_coordinate"*	
	scale_factor		*double*	*5.58871526031607e-05*	
	add_offset		*double*	*-0.155617776423501*	
	valid_range		*string*	*1, 5568*	
accumulated_flash_area		Flash Area accumulation values per contributing pixel on the reference grid.	uint	runtime	pixels
	long_name		*string*	*"Number of contributing unique flashes to each pixel"*	
	grid_mapping		*string*	*"mtg_geos_projection"*	
	coordinate		*string*	*"sparse: x y"*	

l1b_missing_warning		Flag is set if there were gaps in the expected L1b inputs.	byte	runtime	accumulations
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag	byte	runtime	accumulations
	long_name		*string*	*"L1b event geolocation warning"*	
l1b_radiometric_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag in the accumulation period	byte	runtime	accumulations
	long_name		*string*	*"L1b event radiometric warning"*	
average_flash_qa		Average QA value from contributing flashed in the accumulation period. 0=higher probability true; 1=higher false probability.	ubyte	runtime	accumulations
	long_name		*string*	*"average flash confidence value"*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.11 LI-2-AFA-TRAIL

A.11.1 Group:root (/)

A.11.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.11.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.11.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMET SAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		

data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	String		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	String		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal	string		

	compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_ value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		

count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		

standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.11.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of body chunk files produced for this product"*	

A.12 LI-2-AFR-BODY

A.12.1 Group:root (/)

A.12.1.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	
pixels	Number of reference grid points containing lightning data		runtime	
accumulations	Number of accumulated lightning datasets stored in this product. For dissemination this is generally 1, and for archive this is generally between 2 and 20 (covering 1-10 minutes).		runtime	

A.12.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
auxiliary_dataset_status_type	See Enum types	enum byte		

A.12.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	

location_indicator	As per the dataset name field" location_indicator" in dataset name	string		
data_designator	As per the dataset name field "data_designator" in dataset name	string		
platform	As per the dataset name field "spacecraft" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field "data_source" in dataset name	string		
processing_level	As per the dataset name field "level" in dataset name	string		
coverage	As per the dataset name field "coverage" in dataset name	string		
type	As per the dataset name field "type" in dataset name	string		
subtype	As per the dataset name field "subtype" in dataset name	string		
component1	As per the dataset name field "component1" in dataset name	string		
component2	As per the dataset name field "component2" in dataset name	string		
component3	As per the dataset name field "component3" in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field "start_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field "end_time" in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field "processing_mode" in dataset name	string		
special_compression	As per the dataset name field "special_compression" in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the "special_compression" field in dataset name then it is followed by two strings:	string		

	(1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_ value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		

repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		

creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.12.1.4 Variables

Name	Attribute	Description	Type	Values	Shape

auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.	string		auxiliary_dataset
	long_name		*string*	*"Auxiliary dataset identifier"*	
	title		*string*	*"Identifier of auxiliary dataset or type of auxiliary dataset used to produce this product"*	
auxiliary_dataset_status		Possible states for an auxiliary dataset 0 = OK 1 = dataset was used but was out of its stated validity time 2 = auxiliary dataset was not available	auxiliary_dataset_status_type		auxiliary_dataset
	long_name		*string*	*"Status of auxiliary dataset"*	
mtg_geos_projection		Grid mapping variable for the MTG GEOS Projection.	int	runtime	1
	long_name		*string*	*"MTG geostationary projection"*	
	grid_mapping_name		*string*	*"geostationary"*	
	perspective_point_height		*double*	*35786400*	
	semi_major_axis		*double*	*6378137*	
	semi_minor_axis		*double*	*6356752.314245518*	
	inverse_flattening		*double*	*298.257223563*	
	latitude_of_projection_origin		*double*	*0*	
	longitude_of_projection_origin		*double*	*0*	
	sweep_angle_axis		*string*	*"y"*	
accumulation_start_times		Start time for each accumulation dataset.	double	runtime	accumulations
	long_name		*string*	*"Accumulation start time"*	
	units		*string*	*"seconds since 2000-01-01 00:00:00.0"*	
	precision		*string*	*"1 millisecond"*	
accumulation_offsets		One offset per accumulation dataset. These provide index into the arrays of dimension 'pixel'	uint	runtime	accumulations
x		X coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"azimuth angle encoded as column"*	
	axis		*string*	*"X"*	

	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_x_coordinate"*	
	scale_factor		*double*	*-5.58871526031607e-05*	
	add_offset		*double*	*0.155617776423501*	
	valid_range		*string*	*1, 5568*	
y		Y coordinate in reference grid of each contributing lightning data point.	short	runtime	pixels
	long_name		*string*	*"zenith angle encoded as row"*	
	axis		*string*	*"Y"*	
	units		*string*	*"radian"*	
	standard_name		*string*	*"projection_y_coordinate"*	
	scale_factor		*double*	*5.58871526031607e-05*	
	add_offset		*double*	*-0.155617776423501*	
	valid_range		*string*	*1, 5568*	
flash_radiance		Flash radiance values (from resampled flash footprints)	ushort	runtime	pixels
	long_name		*string*	*"Area averaged flash radiance accumulation"*	
	grid_mapping		*string*	*"mtg_geos_projection"*	
	coordinate		*string*	*"sparse: x y"*	
	units		*string*	*"mW.m-2.sr-1"*	
	_FillValue		*string*	*"NC_FILL_SHORT"*	
	scale_factor		*float*	*<configured_value>*	*runtime*
	add_offset		*float*	*<configured_value>*	*runtime*
l1b_missing_warning		Flag is set if there were gaps in the expected L1b inputs.	byte	runtime	accumulations
	long_name		*string*	*"Expected L1b inputs missing"*	
l1b_geolocation_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag	byte	runtime	accumulations
	long_name		*string*	*"L1b event geolocation warning"*	
l1b_radiometric_warning		Set to 1 if any of the contributing L2-Flash products raised the same flag in the accumulation period	byte	runtime	accumulations
	long_name		*string*	*"L1b event radiometric warning"*	
average_flash_qa		Average QA value from contributing flashed in the accumulation period. 0=higher probability true; 1=higher false probability.	ubyte	runtime	accumulations
	long_name		*string*	*"average flash confidence value"*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.13 LI-2-AFR-TRAIL

A.13.1 Group:root (/)

A.13.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.13.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.13.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		

data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO	string		

	URL". If the value is "MULTI", then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the "special_compression" field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words "NO URL".			
disposition_mode	As per the dataset/product name field "disposition_mode" in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field "facility_or_tool" in dataset name	string		
environment	As per the dataset name field "environment" in dataset name	string		
references	"www.eumetsat.int"	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, "None."	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected	string		

	repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely	string		

	identify the creator's institution.			
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	String		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.13.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body	String		body_chunk

		chunk files that were produced for the current repeat cycle or equivalent time period.			
	long_name		*String*	*"Names of body chunk files produced for this product"*	

A.14 LI-2-LE-BODY

A.14.1 Group:root (/)

A.14.1.1 Dimensions

Name	Description	Type	Values	Shape
None defined				

A.14.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type definition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.14.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
data_source	As per the dataset name field “data_source” in dataset name	string		

processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”. If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.	string		
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		

source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime _value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field "facility_or_tool" in dataset name	string		
environment	As per the dataset name field "environment" in dataset name	string		
references	"www.eumetsat.int"	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, "None."	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the "repeat_cycle_time_position" field DDD = day in year value derived from the "repeat_cycle_time_position " field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the "repeat_cycle_in_day" field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.	string		
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan	string		

	pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).			
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any	string		

	standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.			
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.14.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
None defined					

A.14.2 Group:/state

A.14.2.1 Dimensions

Name	Description	Type	Values	Shape
None defined				

A.14.2.2 User Types

Name	Description	Type	Values	Shape

None defined

A.14.2.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.14.2.4 Variables

Name	Attribute	Description	Type	Values	Shape
None defined					

A.14.3 Group:/state/processor

A.14.3.1 Dimensions

Name	Description	Type	Values	Shape
auxiliary_dataset	Number of auxiliary datasets involved in processing the dataset		<runtime_value>	

A.14.3.2 User Types

Name	Description	Type	Values	Shape
auxiliary_dataset_status_type	See Enum types	enum ubyte		

A.14.3.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.14.3.4 Variables

Name	Attribute	Description	Type	Values	Shape
auxiliary_dataset_identifier		Unique identifier for the auxiliary dataset. If available, the filename should be used. If the auxiliary file was not available, the file name template should be stated, with unknown values such as times set to the correct length of lower case x characters.	string		auxiliary_dataset
auxiliary_dataset_status		See Enum types	auxiliary_dataset_status_type		auxiliary_dataset

A.14.4 Group:/data

A.14.4.1 Dimensions

Name	Description	Type	Values	Shape
detector	Fixed value of 4		4	

A.14.4.2 User Types

Name	Description	Type	Values	Shape
detector_type	See Enums worksheet	enum ubyte		

A.14.4.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.14.4.4 Variables

Name	Attribute	Description	Type	Values	Shape
None defined					

A.14.5 Group:/data/<detector>

A.14.5.1 Dimensions

Name	Description	Type	Values	Shape
unfiltered_events	Number of L1b events that were considered for L2 processing		runtime	
l1b_chunks	Number of L1b-LE-chunks that contributed events to the current product.		runtime	
filters	Number of L2 group and flash filters		17	

A.14.5.2 User Types

Name	Description	Type	Values	Shape
None defined				

A.14.5.3 Group Attributes

Name	Description	Type	Values	Shape
None defined				

A.14.5.4 Variables

Name	Attribute	Description	Type	Values	Shape
detector		See Enums worksheet	detector_type		1
	long_name		*string*	*"ID of detector for this group"*	
epoch_time		Start UTC time of first frame in the first contributing L1b chunk for this dataset.	double		1
	long_name		*string*	*"Start time of integration frame"*	
	standard_name		*string*	*"time"*	
	units		*string*	*"seconds since 2000-01-01 00:00:00.0"*	
	precision		*string*	*"1 millisecond"*	
	time_standard		*string*	*"UTC"*	
time_offset		Time offset of each event from epoch_time. Events carried over from previous processing cycle will have a negative offset.	float		unfiltered_events
	long_name		*string*	*"Time offset from epoch time"*	
	_FillValue		*float*	*NC_FILL_FLOAT*	
	units		*string*	*"seconds"*	
event_id		ID of L2 lightning event as per LI-L2-LEF	uint		unfiltered_events
	long_name		*string*	*"ID of LI L2 Event"*	
group_id		ID of L2 lightning group object in the associated LI-2-GR product	uint		unfiltered_events
	long_name		*string*	*"ID of associated LI L2 Group object"*	
flash_id		ID of L2 lightning flash object in the associated LI-2-LFL product	uint		unfiltered_events
	long_name		*string*	*"ID of associated LI L2 Flash object"*	
l1b_chunk_ids		List of all L1b event dataset chunks identified by: concatenation of the L1b chunk filename components as: 10000 * repeat_cycle_in_day + count_in_repeat_cycle	uint		l1b_chunks
	long_name		*string*	*"Array of L1b event chunk IDs"*	
l1b_chunk_offsets		Idenfifies the	uint		l1b_chunks

		'unfiltered_events' indexing boundaries between events coming from different L1b chunks.			
	long_name		*string*	*"Array offsets for L1b event chunk boundaries"*	
l1b_window		window' index for original event from L1b chunk.	uint		unfiltered_events
	long_name		*string*	*"window index of associated L1b event"*	
filter_values		Result of each L2 group and flash filter per event.	ubyte		unfiltered_events, filters
	long_name		*string*	*"L2 filter results"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	filter_names		*string*	*"G_L2Part G_L2SaturRad G_L2RelSobel G_L2DTPeaks G_L2Rad G_L2Size G_L2QA G_L2EarlyRej F_L2SingleGroups F_L2Groups F_L2Foot F_L2TimeCorr F_L2DistCorr F_L2AvgRelSobel F_L2QA F_ReIntroduction F_DuplicateFlash"*	
	scale_factor		*float*	*0.004*	
	add_offset		*float*	*0*	
l1b_filter_qa		Confidence result for the L1b pre-filtering of true events: 0=low confidence; 1=high confidence	ubyte		unfiltered_events
	long_name		*string*	*"L1b event confidence"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	
Level 2_group_filter_qa		Confidence result for the L2 groups: 0=low confidence; 1=high confidence	ubyte		unfiltered_events
	long_name		*string*	*"L2 group confidence"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	
Level 2_flash_filter_qa		Confidence result for the L2 flashes: 0=low confidence; 1=high confidence	ubyte		unfiltered_events
	long_name		*string*	*"L2 flash confidence"*	
	_FillValue		*ubyte*	*NC_FILL_UBYTE*	
	add_offset		*float*	*0*	
	scale_factor		*float*	*0.004*	

A.15 LI-2-LE-TRAIL

A.15.1 Group:root (/)

A.15.1.1 Dimensions

Name	Description	Type	Values	Shape
body_chunk	Number of body chunks that were produced for the current repeat cycle or equivalent time period.		<runtime_value>	

A.15.1.2 User Types

Name	Description	Type	Values	Shape
boolean	See Enums spreadsheet. There is no boolean type in netCDF. This enumerated type at root level can be used by all datasets/products. This user type defininition only needs to be present when it is used within the dataset.	enum byte	See Enums table	
trilean	See Enums spreadsheet. For situations where an undefined state is also required. This user type defininition only needs to be present when it is used within the dataset.	enum byte	See Enums table	

A.15.1.3 Global Attributes

Name	Description	Type	Values	Shape
Conventions	Conventions that the product conforms to. This could be a future version of the CF Conventions that is applicable to netCDF4.	string	e.g."CF-1.7"	
title	Dataset/product name	string		
summary	As defined in the relevant dataset/product format specification.	string		
keywords	As defined in the relevant dataset/product format specification.	string		
keywords_vocabulary	As defined in the relevant dataset/product format specification.	string		
history	As per [CF]	string	“original generated file”	
institution	This field may be extended with other values should datasets/products be generated in other locations.	string	“EUMETSAT”	
location_indicator	As per the dataset name field” location_indicator” in dataset name	string		
data_designator	As per the dataset name field “data_designator” in dataset name	string		
platform	As per the dataset name field “spacecraft” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		

data_source	As per the dataset name field “data_source” in dataset name	string		
processing_level	As per the dataset name field “level” in dataset name	string		
coverage	As per the dataset name field “coverage” in dataset name	string		
type	As per the dataset name field “type” in dataset name	string		
subtype	As per the dataset name field “subtype” in dataset name	string		
component1	As per the dataset name field “component1” in dataset name	string		
component2	As per the dataset name field “component2” in dataset name	string		
component3	As per the dataset name field “component3” in dataset name	string		
product_id	The identifying product_id as used in the SIP	string		
baseline_version	Baseline version. The baseline version will reference of all other version numbers. Assumes processor_version is not sufficient for this.	string		
release_version	Release version. Used to tag datasets that can be considered to have a contiguous consistency sufficient for example, for consideration as a climate set.	string		
processor_version	Processor version. Currently assumes a single processor version number suffices for the relevant IDPF or L2PP. Currently undefined if processor version also includes configuration of static auxiliary data and processor switch configuration, etc.	string		
algorithm_version	Algorithm version. Currently unclear how this would be used and it may be redundant with processor_version.	string		
format_version	Format version of the dataset/product.	string		
time_coverage_start	As per the dataset name field “start_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
time_coverage_end	As per the dataset name field “end_time” in dataset name. Renamed in line with Attribute Convention for Dataset Discovery	string		
processing_mode	As per the dataset name field “processing_mode” in dataset name	string		
special_compression	As per the dataset name field “special_compression” in dataset name	string		
subsetting	If this field is empty then no further strings follow. If this value is a single specified internal compression method as listed in the “special_compression” field in dataset name then it is followed by two strings: (1) human-readable parameters describing the exact internal compression performed (2) either a URL providing a description of the internal compression method or the words “NO URL”.	string		

	If the value is “MULTI”, then this is followed by sets of triplets of strings (one per internal compression applied) A triplet consists of: (1) an internal compression code as listed in the “special_compression” field dataset name ; (2) human-readable parameters describing the exact internal compression performed; (3) either a URL providing a description of the internal compression method or the words “NO URL”.			
disposition_mode	As per the dataset/product name field “disposition_mode” in dataset name	string		
source	Characterisation of the type of data as per [CF].	string		
runtime_data	Space-separated string array of the SIP names of all nonproduct input datasets used in the creation of the dataset (auxiliary data, configuration file, DPP files, etc.) (Was part of <source> field)	string	<runtime_value>	
parent_data	Space-separated string array of the SIP names of all parent products/datasets used in the creation of the dataset (Was part of <source> field)	string		
linked_data	Space-separated string array of the SIP names of all datasets to be linked with this dataset in the archive (e.g. for a Level 0 dataset this would be all additional datasets required to create the virtual L0+ dataset in the archive). (Was part of <source> field)	string		
facility_or_tool	As per the dataset name field “facility_or_tool” in dataset name	string		
environment	As per the dataset name field “environment” in dataset name	string		
references	“www.eumetsat.int”	string		
comment	Miscellaneous information about the data, not captured elsewhere. (See [CF]). Unless otherwise specified in the relevant dataset/product format specification, “None.”	string		
date_created	UTC time of processing formatted in Abbreviated Generalised Time format and defined as the time of the formatting of the dataset/product by the processor. Renamed in line with Attribute Convention for Dataset Discovery	string		
group_tag	String that represents a grouping of datasets that allows chunks and quick-looks to be linked together. The string has the format: <platform>_<datsource>_<processing_level>_<type>_<subtype>_YYYY_DDD_NNNN_<release_version> where: <> indicates the same value as the named global metadata field in the brackets (as described in this table) YYYY = the year value of the “repeat_cycle_time_position” field DDD = day in year value derived from the “repeat_cycle_time_position ” field, left padded with zeroes: 001 = Jan 1st, etc. NNNN = copy of the “repeat_cycle_in_day” field	string		
repeat_cycle_in_day	4-digit number (right-justified, zero-filled) indicating the expected current repeat cycle or group accumulation interval in the day for this particular dataset. For details on how to determine	string		

	the expected repeat cycle see [EXPRC]. The counter starts at 0001 for the first repeat cycle at or after midnight (based on the time_position value) and resets for the next repeat cycle at or after the following midnight. Datasets/products that have no repeat cycle or group accumulation interval (e.g. certain DPP files) should use a fixed value of 0000 to indicate the field is not applicable.			
processed_count_in_repeat_cycle	Cumulative count of the dataset chunk in the repeat cycle or group accumulation interval. Resets when the repeat_cycle_in_day value changes. The counter increments for each created chunk in a repeat cycle or accumulation interval. It does not increment when a chunk is not created due to missing parent data.	string		
count_in_repeat_cycle	4-digit number (right-justified, zero-filled) indicating the expected count value of the dataset chunk in the repeat cycle or group accumulation interval based on the scan pattern or equivalent information. The counter will have discontinuties when chunks are not produced. The counter starts from 1 and resets when the repeat_cycle_in_day value changes. The counter increments for each chunk in a repeat cycle or accumulation interval (whether header, body or trailer). A value of 0 is used for datasets for which the counter is not applicble (e.g. datasets which are not chunk-able).	string		
instrument_configuration_id	List of space-separated values of the “instrument configuration identifier” from the level 0 data ICU-I auxiliary data. Each unique ICID/ICID Version combination produces an entry in the list e.g. an ICID 100 that exists in the product with ICID Versions 1 and 2 will produce two “100” entries in the list.	string		
instrument_configuration_id_version	List of space-separated values of the “instrument configuration identifier version” from the level 0 data ICU-I auxiliary data. Each ICID in the instrument_configuration_id field should have a matching ICID Version entry in the same position in this list.	string		
subsettable_groups	Space separated list of paths to groups that have the subsettable="yes" group attribute.	string		
subsettable_groups_present	Space separated list of paths to groups that are present in the product. Will be the same as subsettable_groups for unsubsetted products.	string		
mtg_name	String field containing the MTG WMO-convention name for the file	string		
alternative_name	String field containing a possible alternative name for the file (e.g. Sentinel-4 naming convention)	string		
purpose	As per the dataset/product name field “purpose” in dataset name	string		
format	As per the dataset/product name field “format” in dataset name	string		
id	Can contain a DOI for reprocessed climate datasets (configuration file). Otherwise set to an empty string.	string		
naming authority	Will contain the DOI issuing authority for reprocessed climate datasets (configuration file) if id attribute is used. Otherwise set to an empty string.	string		

creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'.	string		
creator_institution	The institution of the creator; should uniquely identify the creator's institution.	string		
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.	string		
license	URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.	string		
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.	string		
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas	string		
time_coverage_duration	Describes the duration of the data set.	string		
time_coverage_resolution	Describes the targeted time period between each value in the data set	string		
cdm_datatype	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS [THREDDS]	string		
date_time_position	This is the start time of the repeat cycle (accumulation interval) shifted forwards or backwards to the nearest 30 seconds bin counting from 00:00:00. This removes minor varinations and offsets in the actual observation start time of the repeat cycle. Repeat cycle Observations starting at 11:59:58, 12:00:00 and 12:00:05 would all have a date_time_position value with a time of 12:00:00. An observation starting at 23:59:45 would have value of 00:00:00 and be the first repeat cycle of the next day.	string		
time_position	This is the time string taken from date/time string in date_time_position.	string		
geospatial_lat_min	Geospatial_lat_min specifies the southernmost latitude covered by the dataset.	double		
geospatial_lat_max	Geospatial_lat_max specifies the northernmost latitude covered by the dataset.	double		
geospatial_lon_min	Geospatial_lon_min specifies the westernmost longitude covered by the dataset.	double		
geospatial_lon_max	Geospatial_lon_max specifies the easternmost longitude covered by the dataset.	double		

A.15.1.4 Variables

Name	Attribute	Description	Type	Values	Shape
available_body_chunks		Names of all the body chunk files that were produced for the current repeat cycle or equivalent time period.	String		body_chunk
	long_name		*String*	*"Names of body chunk files produced for this product"*	

APPENDIX B NETCDF AND APPLICABLE STANDARDS AND CONVENTIONS

B.1 NetCDF

The LI Level 2 products are NetCDF-4 files and use the enhanced data model. In addition, they utilise the Hierarchical Data Format version 5 (HDF5) as the storage layer and so can also be read as HDF-5 files. Use of the enhanced NetCDF-4 data model allows groups to be created to aid with the natural collection of various data and the subsetting of channels. In addition, enumerated variable types allow flags to be defined once and used throughout the dataset.

B.2 CF Conventions

The current Climate and Forecast Conventions (CF 1.6) are applicable to version 3 of the NetCDF data model. As such, the LI Level 2 products cannot conform terms of the conventions although they do try to follow the spirit of the conventions as far as possible. The draft CF 1.7 document is also NetCDF-3 specific, but there are plans to create a CF-2 document to cover the enhanced NetCDF-4 model.

B.3 NetCDF Attribute Convention for Dataset Discovery

The NetCDF Attribute Convention for Dataset Discovery [NACDD] document defines a recommended set of metadata for netCDF-based meteorological products. The MTG products are conformant with all the Highly Recommended attributes and the majority of those Recommended attributes that are applicable to the products.

The full list of NACDD metadata present in MTG netCDF products is given in the table below.

Attribute	Description
Highly Recommended	
title	A short phrase or sentence describing the dataset. In many discovery systems, the title will be displayed in the results list from a search, and therefore should be human readable and reasonable to display in a list of such names. This attribute is also recommended by the NetCDF Users Guide and the CF conventions. Nominally contains the product name
summary	A paragraph describing the dataset, analogous to an abstract for a paper.
keywords	A comma-separated list of key words and/or phrases. Keywords may be common words or phrases, terms from a controlled vocabulary (GCMD is often used), or URIs for terms from a controlled vocabulary (see also "keywords_vocabulary" attribute).
Conventions	A comma-separated list of the conventions that are followed by the dataset. For files that follow this version of ACDD, include the string 'ACDD-1.3'. (This attribute is described in the NetCDF Users Guide.)
Recommended	
id	An identifier for the data set, provided by and unique within its naming authority. The combination of the "naming authority" and the "id" should be globally unique, but the id can be globally unique by itself also. IDs can be URLs, URNs, DOIs, meaningful text strings, a local key, or any other unique string of characters. The id should not include white space characters.

naming_authority	The organization that provides the initial id (see above) for the dataset. The naming authority should be uniquely specified by this attribute. We recommend using reverse-DNS naming for the naming authority; URIs are also acceptable. Example: 'edu.ucar.unidata'.
history	Provides an audit trail for modifications to the original data. This attribute is also in the NetCDF Users Guide: 'This is a character array with a line for each invocation of a program that has modified the dataset. Well-behaved generic NetCDF applications should append a line containing: date, time of day, user name, program name and command arguments.' To include a more complete description you can append a reference to an ISO Lineage entity; see NOAA EDM ISO Lineage guidance.
source	The method of production of the original data. If it was model-generated, source should name the model and its version. If it is observational, source should characterize it. This attribute is defined in the CF Conventions. Examples: 'temperature from CTD #1234'; 'world model v.0.1'.
processing_level	A textual description of the processing (or quality control) level of the data.
comment	Miscellaneous information about the data, not captured elsewhere. This attribute is defined in the CF Conventions.
license	Provide the URL to a standard or specific license, enter "Freely Distributed" or "None", or describe any restrictions to data access and distribution in free text.
standard_name_vocabulary	The name and version of the controlled vocabulary from which variable standard names are taken. (Values for any standard_name attribute must come from the CF Standard Names vocabulary for the data file or product to comply with CF.) Example: 'CF Standard Name Table v27'.
date_created	The date on which this version of the data was created. (Modification of values implies a new version, hence this would be assigned the date of the most recent values modification.) Metadata changes are not considered when assigning the date_created. The ISO 8601:2004 extended date format is recommended, as described in the Attribute Content Guidance section.
creator_name	The name of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.
creator_email	The email address of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.
creator_url	The URL of the person (or other creator type specified by the creator_type attribute) principally responsible for creating this data.
institution	The name of the institution principally responsible for originating this data. This attribute is recommended by the CF convention.
project	The name of the project(s) principally responsible for originating this data. Multiple projects can be separated by commas, as described under Attribute Content Guidelines. Examples: 'PATMOS-X', 'Extended Continental Shelf Project'.
geospatial_lat_min	Describes a simple lower latitude limit; may be part of a 2- or 3-dimensional bounding region. Geospatial_lat_min specifies the southernmost latitude covered by the dataset.
geospatial_lat_max	Describes a simple upper latitude limit; may be part of a 2- or 3-dimensional bounding region. Geospatial_lat_max specifies the northernmost latitude covered by the dataset.
geospatial_lon_min	Describes a simple longitude limit; may be part of a 2- or 3-dimensional bounding region. geospatial_lon_min specifies the westernmost longitude covered by the dataset. See also geospatial_lon_max.
geospatial_lon_max	Describes a simple longitude limit; may be part of a 2- or 3-dimensional bounding region. geospatial_lon_max specifies the easternmost longitude covered by the dataset. Cases where geospatial_lon_min is greater than geospatial_lon_max indicate the bounding box extends from geospatial_lon_max, through the longitude range discontinuity meridian (either the antimeridian for -180:180 values, or Prime Meridian for 0:360 values), to geospatial_lon_min; for example, geospatial_lon_min=170 and geospatial_lon_max=-175 incorporates 15 degrees of longitude (ranges 170 to 180 and -180 to -175).
time_coverage_start	Describes the time of the first data point in the data set. Use the ISO 8601:2004 date format, preferably the extended format as recommended in the Attribute Content Guidance section.
time_coverage_end	Describes the time of the last data point in the data set. Use ISO 8601:2004 date format, preferably the extended format as recommended in the Attribute Content Guidance section.
time_coverage_duration	Describes the duration of the data set. Use ISO 8601:2004 duration format, preferably the extended format as recommended in the Attribute Content Guidance section.

time_coverage_resolution	Describes the targeted time period between each value in the data set. Use ISO 8601:2004 duration format, preferably the extended format as recommended in the Attribute Content Guidance section.
Suggested	
creator_type	Specifies type of creator with one of the following: 'person', 'group', 'institution', or 'position'. If this attribute is not specified, the creator is assumed to be a person.
creator_institution	The institution of the creator; should uniquely identify the creator's institution. This attribute's value should be specified even if it matches the value of publisher_institution, or if creator_type is institution.
keywords_vocabulary	If you are using a controlled vocabulary for the words/phrases in your "keywords" attribute, this is the unique name or identifier of the vocabulary from which keywords are taken. If more than one keyword vocabulary is used, each may be presented with a prefix and a following comma, so that keywords may optionally be prefixed with the controlled vocabulary key. Example: 'GCMD:GCMD Keywords, CF:NetCDF COARDS Climate and Forecast Standard Names'.
platform	Name of the platform(s) that supported the sensor data used to create this data set or product. Platforms can be of any type, including satellite, ship, station, aircraft or other. Indicate controlled vocabulary used in platform_vocabulary.
instrument	Name of the contributing instrument(s) or sensor(s) used to create this data set or product. NOTE: Metadata with this name is not present but the root attribute data_source contains equivalent metadata
cdm_data_type	The data type, as derived from Unidata's Common Data Model Scientific Data types and understood by THREDDS. (This is a THREDDS "dataType", and is different from the CF NetCDF attribute 'featureType', which indicates a Discrete Sampling Geometry file in CF.)
references	Published or web-based references that describe the data or methods used to produce it. Recommend URIs (such as a URL or DOI) for papers or other references. This attribute is defined in the CF conventions.
Highly Recommended Variable Attributes	
long_name	A long descriptive name for the variable (not necessarily from a controlled vocabulary). This attribute is recommended by the NetCDF Users Guide, the COARDS convention, and the CF convention.
standard_name	A long descriptive name for the variable taken from a controlled vocabulary of variable names. We recommend using the CF convention and the variable names from the CF standard name table. This attribute is recommended by the CF convention.
units	The units of the variable's data values. This attribute value should be a valid udunits string. The "units" attribute is recommended by the NetCDF Users Guide, the COARDS convention, and the CF convention.

APPENDIX C NETCDF TOOLS

C.1 Overview

The MTG NetCDF datasets make use of a number of features of the enhanced NetCDF-4 data model, including groups, unsigned integer data types and enumerated data types. Not all NetCDF tools are capable of utilizing enhanced NetCDF-4 datasets. However, the NetCDF-4 files also use HDF-5 as the data layer, and so the datasets may also be examined with HDF-5 tools. This Appendix lists freely available tools that are known to be compatible with the MTG NetCDF-4 datasets.
This is not an exhaustive list as other tools and libraries may also be compatible with the enhanced NetCDF-4 model, or may be updated to be so in future.

C.2 NetCDF Libraries and Tools

The latest version of the NetCDF (Network Common Data Form) libraries should be installed. At a minimum, NetCDF 4.3.3.1 is required for writing datasets in the MTG format. NetCDF libraries are being developed by Unidata, a member of the UCAR Comunity Programs. Libraries can be downloaded from their webpage: http://www.unidata.ucar.edu. The NetCDF distribution provides a number of command line tools for looking at the structure and contents of NetCDF datasets. HDF-5 and gzip need to be installed before NetCDF.

C.2.1 gzip

Gzip is used as the default internal compression tool for the MTG NetCDF-4 datasets. The gzip libraries need to be installed before installing HDF-5.

C.2.2 HDF-5

HDF-5 (Hierarchical Data Format, version 5) is used as the storage layer for the MTG NetCDF-4 datasets. The HDF-5 libraries need to be compiled before installing NetCDF-4. HDF 5 is being developed by The HDF Group, The latest libraries can be downloaded from their webpage: https://www.hdfgroup.org.

C.3 Panoply

Panoply is a freely available, cross-platform java application that provides as GUI for browsing and plotting geo-gridded and other arrays from NetCDF datasets. It can also handle other formats such as GRIB, HDF, etc. It is supported by NASA and is available from: http://www.giss.nasa.gov/tools/panoply/. As it is implemented in Java, it provides the same GUI in different operating systems and does not require administrative or root privileges to install. It can display the CDL description as well as images, and makes use of many of the CF conventions. For instance, it can automatically convert integer values to real values using the scale_factor and add_offset variable attributes.

C.4 HDFView

HDFView is a freely available, cross-platform java application with a GUI for browsing and editing HDF4 and HDF5 files. It is available from: http://www.hdfgroup.org/products/java/hdfview/

APPENDIX D EXAMPLES USAGE OF THE PYTROLL SOFTWARE

Pytroll is an easy to use, modular, free and open source python framework for the processing of earth observation satellite data. The provided python packages are designed to be used both in R&D environments and in 24/7 operational production. An overview description of all packages and their mutual dependencies, maturity and known operational usage and as well as a complete list of pytroll packages is available on github.com: https://github.com/pytroll.

Satpy is a Python library for reading, manipulating, and writing data from remote-sensing earth-observing meteorological satellite instruments. Satpy provides users with readers that convert geophysical parameters from various file formats to the common Xarray DataArray and Dataset classes for easier interoperability with other scientific python libraries. Satpy also provides interfaces for creating RGB (Red/Green/Blue) images and other composite types by combining data from multiple instrument bands or products. Various atmospheric corrections and visual enhancements are provided for improving the usefulness and quality of output images. The Pyresample package is used to resample data to different uniform areas or grids. The documentation is available at http://satpy.readthedocs.org/.

Satpy will also include readers of the LI Level 2 data. The corresponding links and examples will be included in this product user guide, once they are available.

APPENDIX E COMPARABLE GLM PRODUCTS

This part will be added in the next version of this document.