Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data
## Document Change Record

<table>
<thead>
<tr>
<th>Issue / Revision</th>
<th>Date</th>
<th>DCN. No</th>
<th>Changed Pages / Paragraphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>13 July 2011</td>
<td>1</td>
<td>• System PDR RIDs #108 to #118 (for RID #109 also a new figure replacing the old one)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #119. Added new figure 3 with explanation in text.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #122</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RIDs #124 to #126</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #128</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RIDs #130 to #139</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RIDs #141 to #153</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #154, Figure 24 edited.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #155 (modified table).</td>
</tr>
<tr>
<td>1A</td>
<td>14 July 2011</td>
<td>2</td>
<td>• System PDR RIDs #15 to #16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #17 (modified list in section 2.2 instead of Table 8 as in the RID).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #45 (Table 3 modified).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #49 (note added to definitions).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #50 (note added to section 2.3).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #51 (note on ASPKE added in section 2.3).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RIDS #55 &amp; #58 (timeliness requirements added in a new section 2.3.4).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RIDS #56, #61, #82, #84, #86 (added product content description with an example illustration in section 4.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• System PDR RID #60, #99 (note added on flash definition and CHUVA campaign in</td>
</tr>
</tbody>
</table>
section 5.2.3)

- System PDR RID #65, related to #154. The figure in 8.2.5.2 has been edited with accompanying text.

- System PDR RID #67 (note added in ).

- System PDR RID #71

- System PDR RIDs #78 and #123: a misplaced bullet in section 3.1 deleted.

- System PDR RID #79: bullet list in section 3.1 modified.

- System PDR RID #85: cut-off rationale explained, taking into account the RID comment.

- System PDR RID #94: L1b/L2 validation discussion in the text separated.

- System PDR RID #96: Note added on prototype processor performance testing limitations.

- System PDR RID #98: Error! Reference source not found. “zoomed” to allow a better view of details.

- System PDR RID #104: modified to explain isolated strokes.

- System PDR RID #74: Word “spectrometer” omitted.

- System PDR RID #91: Clustering distance discussion enhanced in Section 5.2.3.

- System PDR RID #73: Section 2.2 (list of benefits) edited, so it is no longer a summary session.

- System PDR RID #80: Table 2 modified to include jitter.

- System PDR RID #106: Section 4.3 modified to include description on how group/flash locations are computed.

- System PDR RID #87: Section 5.2.1 modified to describe baseline vs. future enhancement.

- System PDR RID #107: Section 4.3 and 5.2.4
to better indicate that baseline processing is independent of auxiliary data sources.

<table>
<thead>
<tr>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
</table>
| 1A 14 July 2011 3 | • System PDR RIDs #1: availability estimate of operational processor related parts of the document (section 7) added.  
• System PDR RIDs #2, #75, #76, #120: full rewrite of section 2.3 parts dealing with system requirements, including explaining notes on issues brought up by review. 
• System PDR RIDs #72: modified section 2.2 |
| 1A 12 August 2011 4 | • Added the Product requirements table for lightning detection as Annex I as requested by PDR panel |
| 1E 12 November 2012 5 | • Modified section 5.2.4 on Parallax correction to highlight that it is done on L2, not on L1b, and by removing references to ATDnet data usage. |
| 1E 12 November 2012 6 | **Density product related changes:**  
• Removed the density product from section 9, on other potential MTG LI data products.  
• Changed wordings in the document from “baseline L2 product” to “baseline L2 lightning data product”.  
• Density product added to Table 5.  
• Added section 4.4 on “L2 Lightning density products”  
• Added main section 6 on “Algorithm description for lightning density generation”, detailing the density products.  
• Removed obsolete prototype processor example images, which showed flash/group/event “densities” which have not been computed as defined for L2. |
| 1E 31 January 2013 7 | • Reworded the description on redundancy filtering and removed a note suggesting that it would not be baseline. It is baseline but the contents are still not defined as long as L1 processing has not been defined. |
## Density product modifications:

- Products no longer called density products as no division with area is done in the computation, and is left for the users for any given area, with the help of detailed grid information.

- In line with the above, three newly defined products are defined: *accumulated flashes*, *accumulated flash index*, and *accumulated flash radiance*.

- Initially two periodicities were described (30 sec and 2.5 minutes). The latter was intended for direct comparison with FCI imagery. However, as the user can easily combine successive 30 sec products in order to reach any multiple of 30 seconds, the 2.5 independent product was regarded as unnecessary.

- Describing the resampling necessary for the transition from the LI grid to the target IR grid (2 km) in a new sub-section.

### Change Log

<table>
<thead>
<tr>
<th>Draft</th>
<th>Date</th>
<th>Page</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11 February 2013</td>
<td>8</td>
<td>Density product modifications:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Products no longer called density products as no division with area is done</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>in the computation, and is left for the users for any given area, with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>the help of detailed grid information.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• In line with the above, three newly defined products are defined:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>accumulated flashes</em>, <em>accumulated flash index</em>, and *accumulated flash</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>radiance*.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Initially two periodicities were described (30 sec and 2.5 minutes).</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The latter was intended for direct comparison with FCI imagery. However,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>as the user can easily combine successive 30 sec products in order to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>reach any multiple of 30 seconds, the 2.5 independent product was</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>regarded as unnecessary.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Describing the resampling necessary for the transition from the LI grid</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to the target IR grid (2 km) in a new sub-section.</td>
</tr>
<tr>
<td>2</td>
<td>12 February 2013</td>
<td>9</td>
<td>New version for the system PDR (delta)</td>
</tr>
<tr>
<td>3</td>
<td>18 April 2013</td>
<td>10</td>
<td>Updated based on LIST comments (editorial)</td>
</tr>
<tr>
<td>3</td>
<td>22 April 2013</td>
<td>11</td>
<td>New Section 5.2.4.1 explaining the selected approach for parallax correction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• New Sections 5.2.2.1 and 5.2.3.1 explaining the methodology for computing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>the group/flash locations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• New Section 5.2.3.1 describing the methodology for computing the X/Y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>distances for the WED-algorithm.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Editing parts of text mentioning LI integration time or spatial resolution</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>to reflect the current state of design (1 ms and 4.5 km SSP).</td>
</tr>
<tr>
<td>3</td>
<td>25 April 2013</td>
<td>12</td>
<td>Changing the WED algorithm from computing the direct distance between</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>groups instead of (X, Y) distances.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Modified the description of filtering to be done at L2, consists now of:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>redundancy filtering (event level), 1st and 2nd level flash filtering (after L2 flash clustering and</td>
</tr>
</tbody>
</table>
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

<table>
<thead>
<tr>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 October 2013</td>
<td>Updates based on the Consistency Checkpoint Review (CCR) actions:</td>
</tr>
<tr>
<td></td>
<td>• OBT added to list of acronyms (RID #135)</td>
</tr>
<tr>
<td></td>
<td>• Definition of orbital radius (RID #172)</td>
</tr>
<tr>
<td></td>
<td>• Typos corrected (RID #124, #132, #161, #164)</td>
</tr>
<tr>
<td></td>
<td>• Added footnotes on references to other parts in the text (RID #133)</td>
</tr>
<tr>
<td></td>
<td>• LI uses for other research purposes mentioned (RID #175)</td>
</tr>
<tr>
<td></td>
<td>• Updated the MSG operational services timeline (RID #189)</td>
</tr>
<tr>
<td></td>
<td>• Updated wording in the accumulated product definitions (RID #179)</td>
</tr>
<tr>
<td></td>
<td>• Document chronology added (RID #209)</td>
</tr>
<tr>
<td></td>
<td>• Terminology in Section 1.1 updated (RID #187)</td>
</tr>
<tr>
<td></td>
<td>• Added a bullet on complementary use of ground-based observations (RID #191)</td>
</tr>
<tr>
<td></td>
<td>• Timeliness definition added (RID #196)</td>
</tr>
<tr>
<td></td>
<td>• Removing statement on using LI data alone for nowcasting of convection, as it would in reality always be used in conjunction with modelling or additional data sources (RID #213)</td>
</tr>
<tr>
<td></td>
<td>• Footnote added on definition of the clustering distances (left open currently) (RID #185)</td>
</tr>
<tr>
<td></td>
<td>• Footnote added describing the limited purpose of the prototype processor (RID #169)</td>
</tr>
<tr>
<td></td>
<td>• Added a note on quality of products and the respective flags (RID #123)</td>
</tr>
<tr>
<td></td>
<td>• Footnote added explaining the adjacency principle (RID #149)</td>
</tr>
</tbody>
</table>
- “first fit” and “full fit” explanation expanded (RID #159)
- Footnote added to explain the “N%” of the accumulated product processing of LI grids into the FCI grid (RID #162)
- Spatial sampling definition and difference to LIS clarified (RUD #177)
- Removed the text “...with the size of the grid box in km², and in addition divided...” from the accumulated flash product description. The accumulated flashes are not density products (divided by km²) as the text would have indicated (RID #178)
- Added a note that FCI grid information needs to be available to the accumulated product processing (RID #180)
- Accumulated product conceptual examples (in Figure 26 to Figure 28) clarified to refer to the accumulated flash index product, with a note added how the corresponding accumulated flash product examples would be in comparison (RID #186)
- Added a full Section on an up-to-date description of the LI instrument. Removing TBC and TBD throughout the document to reflect this major update (RID #130).

Other changes:

- Remove main section “Assumptions and limitations”, which is now just a placeholder. It can be added later if needed.
- Remove the current contents on the main Section “Synergy with ground-based observations”, and add a note that these issues will be dealt with in the LI Cal/Val document in due time.
- Updated Figure 10 on “Overview of MTG LI processing chain up to L2” to reflect the current status of the processing scheme, especially regarding the false event/group/flash filtering which will be encompassing both L1 and L2 processing. In relation, also modified the
text introducing the image, to reflect the change in how false event filtering is now understood.

- Added a reference to the follow-on study on the LIProxy simulator
- Added an up-to-date picture of the LIProxy GUI
- Removed Section 11 on “Other potential MTG LI data products” which is not considered relevant for the purpose of the document.
- Significant changes in the LINET-based proxy data description on include the CHUVA-related new results and modifications to the method
- Changed a note on whether IC or CG lightning is more easily observed by its optical pulse above clouds to reflect literature and LIST comments (Section 2.3.1).

<table>
<thead>
<tr>
<th>4</th>
<th>6 March 2014</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clarified statement in Section 6.2.2 on group clustering, in that integration time is not “fixed” at 1 ms but can be changed and shall be tested also during commissioning. This is leading to the need to have also the possibility to include temporally neighbouring time frames for group clustering (and not only events taking place in the same frame, which is the LIS approach).</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Added a subsection 6.2.3.2 on “Flash clustering based on separate distance and time threshold criteria”. This means that two separate methods shall be available in the L2PF for flash clustering (WED and the separate distance/time criteria). The change is based on the L2 algorithm study outcome run as a LI MAG study.</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table of Contents

List of Figures .......................................................................................................................... 12
List of Tables .............................................................................................................................. 15
List of Acronyms ......................................................................................................................... 16

1 Introduction .......................................................................................................................... 18
  1.1 Terminology ..................................................................................................................... 18
  1.2 Document evolution ......................................................................................................... 20

2 Overview ................................................................................................................................ 22
  2.1 The Meteosat Third Generation (MTG) system ............................................................... 22
  2.2 The need for lightning measurements from a geostationary orbit .................................. 24
  2.3 MTG LI System requirements .......................................................................................... 25
    2.3.1 Special notes on system requirements ...................................................................... 31
    2.3.2 Design drivers .......................................................................................................... 32
    2.3.3 Restricted mission operations ................................................................................... 33
    2.3.4 Timeliness requirements of MTG LI L2 products .................................................... 33

3 MTG Lightning Imager ............................................................................................................ 34
  3.1 Main characteristics .......................................................................................................... 34
  3.2 Lightning pulse detection concept and functional description ........................................ 36
  3.3 Real Time Pixel Processor (RTPP) description, background determination and subtraction .......................................................... 37
  3.4 Window Manager ............................................................................................................ 38
  3.5 Single Pixel Filtering ...................................................................................................... 39
  3.6 Science Data Packetiser ................................................................................................. 40

4 Measurement approach ........................................................................................................ 41
  4.1 Transition from LEO to GEO observations .................................................................... 41
  4.2 Lightning optical signal ................................................................................................... 43
  4.3 Measurement principle .................................................................................................... 43

5 Product generation from L1 to L2 ...................................................................................... 45
  5.1 Summary of non-lightning event/group/flash filtering ...................................................... 45
  5.2 Outcome of L1 processing: L1b data ............................................................................... 47
  5.3 Baseline L2 Lightning data product .................................................................................. 47
  5.4 L2 accumulated products ............................................................................................... 49
  5.5 Data requirements .......................................................................................................... 50

6 Algorithm description for flash generation ........................................................................ 53
6.1 Algorithm overview

6.2 Significant algorithm details
   6.2.1 Redundancy filtering at event level and flash filtering
   6.2.2 Group clustering
   6.2.3 Flash clustering
   6.2.4 Parallax correction

7 Algorithm description for lightning accumulated product generation

7.1 Algorithm overview

7.2 Resampling of event/flash data into the accumulated product grid

7.3 Significant algorithm details

7.4 Example of the accumulated product usage

8 Algorithm calibration and validation

8.1 Validation

8.2 Quality control

9 Product processor development and realisation

9.1 Prototype processor

9.2 Operational processor

10 Proxy data sets and tools

10.1 Proxy data based on statistical and conceptual models (LIProxy tool)
   10.1.1 Concepts
   10.1.2 Realisation
   10.1.3 Software

10.2 Proxy data based on LINET/LIS/LMA comparisons
   10.2.1 Background
   10.2.2 Results of the first data comparison of LIS and LINET observations
   10.2.3 CHUVA field campaign outcome and results of the second data comparison
   10.2.4 Steps for generating proxy data from LINET observations based on the first data comparison
   10.2.5 Further enhancement of the LINET proxy data method based on the CHUVA campaign outcome
   10.2.6 Application to Europe
   10.2.7 Examples

11 Potential for synergy with ground-based observations

12 References

Annex I: Product requirements for Lightning Detection
LIST OF FIGURES

Figure 1. Histogram of the flash duration as derived from global observations of LIS optical signals between 2002 and 2010. .......................................................... 19

Figure 2: European Territories of EUMETSAT Member States .......................................................... 26

Figure 3. Lightning Pulse Temporal and Spatial Templates. ................................................................. 29

Figure 4. Event radiant energy of LIS observations (2002-2010) with a day/night separation .......... 32

Figure 5. Schematic picture of the LI optical bench. ......................................................................... 35

Figure 6. LI geographical coverage area. .......................................................................................... 36

Figure 7. Annual flash density derived from NASA OTD and LIS observations (1995-2006). ........... 41

Figure 8. Example optical pulses measured with the FORTE system [Light et al., 2003]. ............... 43

Figure 9. Detection principle of the MTG Lightning Imager, where a real observation from the space shuttle is used as an example image of how a storm with lightning could look like from space (in night-time conditions). ......................................................... 44

Figure 11. Flash tree example describing the contents of the baseline L2 product. .......................... 48

Figure 12. Flow chart of MTG LI L2 clustering algorithm overview ............................................... 54

Figure 13. Flow chart of MTG LI group clustering. The lower left box updates the “flash statistics” (and not “group statistics”) as these statistics are on the root level where all flashes and the related groups and events are traced. .......................................................... 55

Figure 14. Flow chart of MTG LI flash clustering. ........................................................................... 56

Figure 15. Detected events shown in a detector X/Y matrix format covering several integration frames (of t ms length) .................................................................................. 59

Figure 16. Illustration of how to compute the distances D in Option 2, where each event in a group are compared to events in the other group(s). The shortest distance measured is then the baseline for D. The red crosses indicate the location of the groups, which are based on the radiance weighted average locations of events. ........................................................................ 64

Figure 17. Snapshot to area of interest with two flashes. ................................................................. 66

Figure 18. Snapshot to area of interest with 17 groups. The numbers refer to the group numbering used in Table 6. ......................................................................................... 66

Figure 19. Snapshot to area of interest with two flashes with the “first fit” LIS algorithm. ............ 69

Figure 20. Snapshot to area of interest with one flash with the “full fit” MTG LI algorithm. Note that the location of the flash is determined by a radiance weighted average of the group locations, not a geometric average. ................................................. 69
Figure 21. Snapshot to area of interest with 27 groups. The numbers refer to the group numbering used in Table 8 ................................................................. 70

Figure 22. LI grid data to be resampled into the accumulated product grid (equal to the FCI IR grid). Only one LI grid element is shown positioned on a sample of IR grid elements. Those target grid pixels which are “selected” based on the LI grid element coverage (>50% requirement) are shaded ........................................................................................................... 75

Figure 23. Flow chart of MTG LI L2 accumulated flashes algorithm overview .............................................................. 76

Figure 24. Flow chart of MTG LI L2 accumulated flash index algorithm overview .............................................................. 77

Figure 25. Flow chart of MTG LI L2 accumulated flash radiance algorithm overview .............................................................. 78

Figure 26. Example of the accumulated flash index principle. Left, flash #1 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid ......................................................................................................................... 79

Figure 27. Example of the accumulated flash index principle. Left, flashes #1 and #2 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid ......................................................................................................................... 79

Figure 28. Example of the accumulated flash index principle. Left, flashes #1, #2, and #3 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid ......................................................................................................................... 80

Figure 29. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is T₀ = 30 seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil) ........................................................................................................................................ 82

Figure 30. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is T₀ = 60 seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil) ........................................................................................................................................ 83

Figure 31. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is T₀ = 90 seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil) ........................................................................................................................................ 84

Figure 32. Screenshot of the Matlab-based prototype L2 processor ................................................................................................................................. 89

Figure 33. Schematic picture of the proxy data structure. Photons compose optical pulses of certain size and energy. The pulses with spatial position and time are organized in storms. ........................................................................................................................................ 92

Figure 34. LIProxy GUI example ........................................................................................................................................................................................................ 96

Figure 35. Pulses (top), photons (middle) and overlaid detected events (bottom) for a simulated storm. The number of displayed photons was reduced for sake of clarity of the figure ........................................................................................................................................ 98

Figure 36. Schematics diagram of stroke related group modelling. One stroke coincident group is followed by normal distributed (in time) ........................................................................................................................................ 103

Figure 37. Schematic of LINET stroke detection and smallest detectable signal ........................................................................................................................................ 105
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Figure 39. Array of the smallest detectable LINET stroke (absolute peak current) over the Sao Paulo area derived from data of 10 Feb and 27 March 2012. Array resolution is 0.1 degree in each direction .................................................. 106

Figure 40. LIS groups per LINET stroke (GPS) from coincident flash observations for LIS overpasses in different areas .......................................................................................................................... 107

Figure 41. The LINET configuration of sites in Europe in 2008 ................................................................................................................................. 108

Figure 42. Distribution of the smallest detected peak current during 2010. (a) All strokes observed with smallest plotted on top, (b) Detection matrix with a 0.1 x 0.1 degree resolution .............................................. 109

Figure 43. Application of LINET proxy data on 28 July 2006, showing 5 minutes of data overlaid on MSG SEVIRI 10.8 µm channel image. The full European area is shown in (a) and the white square indicates the area highlighted in (b) ................................................................................................................................. 111
LIST OF TABLES

Table 1. L2 ATBD significant milestones. ................................................................. 20
Table 2. Summary of some basic differences of LEO and GEO lightning observations. .......... 42
Table 3. Typical NLEs affecting MTG LI lightning optical pulse detection.............................. 45
Table 4. L2 accumulated product characteristics. ......................................................... 50
Table 5. Data requirements for L2 processing. ............................................................ 51
Table 6. Groups included in the snapshot area of interest of Case #1. ............................... 67
Table 7. WED between group #16 and each candidate group #9 to #15. ....................... 68
Table 8. Groups included in the snapshot area of interest of case #2. There are two “first fit” LIS algorithm flashes, where the groups for flash #2 are shaded. ......................................................... 71
Table 9. WED between group #24 and all preceding groups (#1-#23). ............................ 72
Algorithm Theoretical Basis Document (ATBD) for L2 processing of
the MTG Lightning Imager data

LIST OF ACRONYMS

ADC Analog-to-Digital Converter
APS Active pixel sensor
ATDnet Arrival Time Difference Network (MetOffice ground based LLS)
CAF Central application facility
CC Cloud-to-cloud
CDF Common Data Format
CMOS Complementary Metal Oxide Semiconductor
CG Cloud-to-ground
CL Control logic
DCP Data collection platform
DCS Data collection system
DE Detection efficiency
DLR Deutsches Zentrum für Luft- und Raumfahrt
DT Detected transient
FAR False alarm rate
FCI Flexible combined imager
FEE Front End Electronics
FOC Full operational capability
FORTE Fast On-Orbit Recording of Transient Events
FOV Field of view
FPA Focal plane assembly
FPGA Field programmable gate array
FPN Fixed Pattern Noise
GLM Geostationary lightning mapper
GEO Geostationary orbit
GPS (LIS) Groups per (LINET) Stroke
GUI Graphical user interface
HDF Hierarchical Data Format
IC Intracloud
INPE Instituto Nacional de Pesquisas Espaciais
INR Image Navigation and Registration
IQT Instrument Quality Tool
IRS Infrared sounder
L2PF Level 2 Processing Facility
LCFA Lightning cluster-filter algorithm (for GLM)
LE Lightning event (caused by a real lightning optical pulse)
LEO Low earth orbit
LERMA Laboratoire d'Etude du Rayonnement et de la Matière en Astrophysique (Observatoire de Paris)
LI Lightning imager
LIS Lightning imaging sensor
LIST MTG Lightning Imager Science Team
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

LLS  Lightning Locationing System
LMA  Lightning Mapping Array
LME  Lightning Imager Main Electronics
LOH  Lightning Imager Optical Head
LWC  Liquid water content
MET  Meteorological Support Division
MSG  Meteosat second generation
MTG  Meteosat third generation
NASA National Aeronautics and Space Administration
NBF  Narrow bandpass filter
netCDF Network Common Data Format
NLE  Non-lightning event (caused by something else than a lightning optical pulse)
NMS  National Meteorological Service
NOAA National oceanic and atmospheric administration
OBT  On-board time
OTD  Optical transient detector
QE   Quantum efficiency
RF   Radio frequency
RMDCN Regional Meteorological Data Communication Network
RTPP Real time pixel processor
SAF  Satellite application facility
SAR  Search and rescue
S/C  Spacecraft
SRF  Solar rejection filter
SSP  Sub-satellite point
SW   Software
TBC  To be confirmed
TBD  To be decided
TBW  To be written
TC   Telecommand
TCE  Technical Computing Environment
TRMM Tropical Rainfall Measurement Mission
USP  University of Sao Paulo
UVN  Ultraviolet, Visible and Near-infrared sounder
VHF  Very high frequency
WED  Weighted Euclidian distance
XPOL X-polarisation
1 INTRODUCTION

This document describes the scientific baseline for the primary L2 products for the MTG LI, i.e. the “flashes” along with the intermediate product “groups”. The processing at L2 starts from quality controlled L1b “events”, which can be described as MTG LI triggered events caused by lightning optical pulses. The statistical properties of LI events are as much depending on the instrument design (number of detector elements, integration time) as on the underlying geophysical characteristics of lightning itself. Therefore, the L2 algorithms make use of the temporal and spatial coherency of lightning to overcome the design related features.

A brief description of the important terminology is first given in Section 1.1.

Background information on the need for lightning measurements from space, the Meteosat Third Generation (MTG) system, and the MTG Lightning Imager are given in Sections 2 and 3. The measurement approach and the basic principles of L2 product generation for the LI are described in Sections 4 and 5, respectively. The L2 algorithm details (baseline clustering and accumulated product algorithms) are discussed in Section 6, with a brief description on product calibration and validation in Section 8. The product processor is described in Section 9 (currently only the prototype processor). Two sets of proxy data created for the LI product development and validation tasks are described in Sections 10.

1.1 Terminology

Some basic MTG LI and data processing related terminology is given below. The reference is the MTG Conventions and Terms document (Ref: EUM/MTG/DEF/08/0034), where the full description of all MTG related terms and conventions are given. It should be noted that since the main heritage of the MTG LI mission and its baseline products comes from the Tropical Rainfall Measurement Mission – Lightning Imaging Sensor (TRMM-LIS) mission, also the definitions and terminology used in the MTG LI related documentation share the same heritage.

**Lightning Optical Pulse**

A lightning optical pulse is produced by an electric discharge within or below a cloud, where the optical radiation is emitted from the hot lightning channel. The lightning pulse duration is on the order of 50 µs and the released photons are transported to the cloud surfaces by scattering. The resulting lightning optical signal to be observed at the cloud top has a pulse duration delayed and widened in time to about 600 µs, distributed over an enlarged area of a minimum of about 100 km² up to a maximum area of about 10,000 km² depending on the number of scattering processes involved and the complexity of the flash “skeleton”.

**LI Groups**
Groups can be regarded as representing lightning strokes\(^1\). They result from combining all the L1b events (events with geolocation, UTC time stamp and calibrated radiance) adjacent pixels, detected during the same and/or adjacent integration frame. Groups have attributes based on their characteristics, such as: shape, duration, quality flags, duration of a group. Groups are an intermediate product to obtain flashes.

**Lightning Flash**

Lightning flashes are composed of multiples of CG return strokes or CC lightning, represented by LI groups, causing lightning optical pulses grouped by proximity in location over a flash duration. An example distribution of typical total flash duration observed with the LIS instrument is shown in Figure 1.

![Figure 1. Histogram of the flash duration as derived from global observations of LIS optical signals between 2002 and 2010.](image)

\(^1\) A return stroke is the massive electrical discharge following the bridging of the conductive channel of ionized air between the negative charges in the cloud and the positive surface charges below. This return stroke is the most luminous and noticeable part of the lightning discharge, and therefore also the part of the lightning process most easily detected by the optical pulse detection system such as the Lightning Imager.


**Triggered Event**  
A triggered event occurs when the energy registered by a detector element exceeds the LI trigger threshold.

**Trigger Threshold**  
The trigger threshold is used, at detector element level to discriminate a lightning optical pulse from the background radiance.

**Background radiance images**  
The Background Radiance for each LI detector element in the LI detector array averaged over a given time interval.

**Lightning Event**  
A lightning event is defined as a LI triggered event caused by a lightning optical pulse.

**Detection Efficiency / Detection probability**  
The LI detection efficiency is the probability that a LI triggered event is generated for a radiance sample originating from a lightning optical pulse.

**False Alarm**  
A false alarm is defined as a LI triggered event occurring in the absence of a lightning optical pulse after having applied any filtering (on-ground or on-board the satellite).

**False Alarm Rate**  
The false alarm rate (FAR) is the number of LI false alarms per second.

### 1.2 Document evolution

The milestones of the L2 ATBD document evolution are described in Table 1.

<table>
<thead>
<tr>
<th>Date</th>
<th>Version</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 2011</td>
<td>1</td>
<td>First release entering the system Preliminary Design Review</td>
</tr>
<tr>
<td>February 2013</td>
<td>2</td>
<td>As a result of discussions with the Lightning Imager Science Team (LIST), added the accumulated products as LI L2 baseline products.</td>
</tr>
<tr>
<td>Month</td>
<td>Version</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>May 2013</td>
<td>3</td>
<td>Final LIST reviewed update before the Consistency Checkpoint Review (CCR).</td>
</tr>
<tr>
<td>November 2013</td>
<td>4</td>
<td>Updates based on the CCR outcome.</td>
</tr>
</tbody>
</table>
2 OVERVIEW

2.1 The Meteosat Third Generation (MTG) system

The mission of the Meteosat Third Generation (MTG) System is to provide continuous high spatial and time resolution observations and geophysical parameters of the Earth System derived from direct measurements of its emitted and reflected radiation using satellite-based sensors from the geo-stationary orbit. To fulfill its mission it is required to deploy sustained capabilities to acquire, process and distribute to downstream application users and second tier processing centres environmental data on a broad spectral range (from UV to LWIR), covering extensive areas (global and regional), and within a variety of different time scales to continue and enhance the services offered by the Second Generation of the Meteosat System (MSG).

The MTG mission encompasses the following observation missions:

- Flexible Combined Imager (FCI) mission, allowing to scan either the full disc in 16 channels every 10 minutes with a resolution in the range 1-2 km (Full Disc High Spectral resolution Imagery (FDHSI) in support of the Full Disc Scanning Service (FCI-FDSS)) or a quarter of the earth in 4 channels every 2.5 minutes with a resolution twice better (High spatial Resolution Fast Imagery (HRFI) in support of the Rapid Scanning Service (FCI-RSS)).

- InfraRed Sounding (IRS) mission, covering the full disc in 60 minutes, providing hyperspectral sounding information in two bands, a Long Wave InfraRed (LWIR: 700 - 1210 cm\(^{-1}\)) and Mid Wave InfraRed (MWIR: 1600 - 2175 cm\(^{-1}\)) band with a resolution around 4 km.

- Lightning Imager (LI) mission, detecting continuously over almost the full disc, the lightning discharges taking place in clouds or between cloud and ground with a resolution around 10 km.

- Ultraviolet, Visible & Near-infrared (UVN) sounding mission, covering Europe every hour taking measurements in three spectral bands (UV: 290 - 400 nm; VIS: 400 - 500 nm, NIR: 755 - 775 nm) with a resolution around 10 km. The UVN mission is implemented with the GMES Sentinel-4 instrument accommodated in the MTG-S satellites.

The Space Segment of the MTG System consists of a satellite constellation. Three in-orbit satellites are needed to support the complete and total set of missions and functions listed above, the full operational capability (FOC). To span the operational life time of the programme over 20 years, there will be in whole 4 satellites dedicated to support the Imagery missions (MTG-I), and 2 satellites to support the sounding missions (MTG-S).
Each satellite is specified for a nominal lifetime (including commissioning) of ca. 8.5 years, carrying the payload complements or meteorological sensors according to the following split which have been confirmed in the definition and feasibility phase -Phase A- for its detailed design and implementation:

- MTG-I: FCI, LI, DCS and SAR
- MTG-S: IRS + UVN (GMES S-4)

This distribution of the payload complement and redundancies is to a significant part due to the novel nature of the sounding missions (IRS and UVN) and their development risks, including also their respective downstream applications using data from geostationary systems. The twin satellite concept is also balancing the payload mass distribution, power, consumables such as fuel, and data rates making effective use of the same platform.

Complementary to the direct observation missions summarised above and yet essential to satisfy key user needs, the following objectives have also to be fulfilled by MTG:

- Level 2 product extraction;
- Data Collection System (DCS), for collecting and transmitting observations and data from surface, buoy, ship, balloon or airborne Data Collection Platforms (DCP);
- Long term archiving in the EUMETSAT Data Centre including reprocessing;
- Near Real Time Data Dissemination & Relay services to users, including Foreign Satellite Data (FSD) collection and distribution (data from other EUMETSAT and Third Party satellite systems for calibration and global applications):
  - EUMETCast & High Rate dissemination services (including relay of Foreign Satellite Data (FSD));
  - RMDCN dissemination service;
  - Search And Rescue (SAR) relay service. Similarly to MSG, the MTG System has the capability to accommodate a GEOSAR transponder, enabling the operations of the mission under the aegis of the COSPAS-SARSAT System.
  - Internet dissemination services;
- Archived dataset retrieval services continue to be provided as part of the multi-mission EUMETSAT Data Centre services.
- User support services are enhanced to address MTG as well.

The MSG satellites are expected to deliver operational services in full nominal configuration at least until 2018 - hopefully longer - at which time a Meteosat Third Generation (MTG) System needs to be available to provide the imaging mission; noting however that the launch
date of MTG-I1 may be postponed e.g. to optimise the use of in-orbit MSG assets and the overall MSG/MTG lifetime.

The full nominal operational configuration includes a prime MTG-I satellite, a second MTG-I satellite (acting as in-orbit hot backup for the prime MTG-I satellite and supporting the RSS services) and an MTG-S satellite.

2.2 The need for lightning measurements from a geostationary orbit

The MTG Lightning Imager (LI) will complement existing ground based capabilities for the detection and location of lightning events with information on cloud-to-ground (CG) and cloud-to-cloud or intracloud (CC/IC) discharges, without a discrimination of the types with MTG LI data alone. Based on currently available total lightning data from LEO satellites and ground-based systems, it has been estimated that CG strokes count only roughly 15% of the total lightning activity, although the fraction of CG flashes (positive and/or negative) varies during a storm lifecycle. The highest benefit will be provided by the full disk observation of total lightning (intra-cloud and cloud to ground), with uniform detection efficiency over land and ocean and non-industrialized regions, e.g. in areas not covered by ground observations such as Africa and Atlantic Ocean. The targeted detection efficiency (see Section 1) will allow offering a consistent level of service providing thus a “space truth” reference for different ground based lightning observation systems over the MTG LI coverage area.

The LI will serve as an essential point of comparison for all lightning systems (ground- and space-based). However, the sensed phenomenon (optical pulses) is not directly comparable to any existing documented dataset, presenting challenges for dataset intercomparison. Comparisons of different detection systems should thus take an integrated approach, by looking to explain any differences that are found between the systems.

The main benefits of the MTG LI mission can be described as:

- The MTG LI measurements of total lightning (IC+CG) are complementing the global/regional measurements of CG lightning as provided by ground based systems and will improve the quality of information which is essential for air traffic routing and safety.
- Error characterized (i.e. after validation) IC+CG information can be assimilated to improve very short range forecasts of severe convective events or used to verify/validate other satellite data based NWC algorithms to forecast time and location of initiation of lightning in a new storm cell.
- Information on lightning can also serve as proxy for adiabatic and latent heating to be assimilated in global/mesoscale NWP models.
- The information on IC+CG will allow to assess the impact of climate change on thunderstorm activity by monitoring and long-term analysing lightning characteristics. In cooperation with the two NOAA GLMs on
GOES-R and GOES-S a major part of the globe is covered by a long term committed GEO lightning (IC+GC) observing system.

- Providing IC+CG information on a global scale will be a prerequisite for studying and monitoring the physical and chemical processes in the atmosphere regarding NOx, which is playing a key role in the ozone conversion process and acid rain generation.

- Use of total lightning information as a convective/stratiform separator for rain classification and rain retrieval.

- In high latitude boreal forests lightning is a major cause of forest fires. LI data can be used to issue warnings of high risk areas in affected regions.

- Lightning observations can be used to help diagnose the intensification of tropical cyclones over oceans.

- Lightning observations can be used to identify active convection for over-ocean air traffic.

- Providing a linkage to TRMM LIS science and climatological datasets for the tropics that have been developed since 1998. LIS climatology is based on very long term observations due to the short viewtime available from the instrument. Verifying and developing the climatologies obtained with LIS/OTD from GEO observations will be an important task in the future.

- Some ground based system operating in the LF/VLF and VHF regions are more suitable for monitoring utilities, airports and such, which require very high location accuracy down to hundreds of meters. However, observations from space offer a complementary data source by identifying, tracking and extrapolating electrically active areas with a uniform observation quality.

The product requirements for L2 lightning detection are presented in Annex I.

2.3 MTG LI System requirements

The main requirements from the MTG System Requirements document (SRD) are reproduced here for convenience [MTG, 2011b].
Coverage area

The Lightning Imager (LI) coverage shall contain at least the European territories of all the EUMETSAT member states and at least 84% of the visible earth disc (taken as a circle of 17.54° in diameter centred at SSP) when the satellite is within the nominal longitude range.

![Figure 2: European Territories of EUMETSAT Member States](image)

Note: The European territories of the EUMETSAT member states are shown in blue dots on the map. The whole of Turkey is to be taken to lie within Europe for this definition.

Measurement bandwidth
The LI triggered events shall consist in the measurements of the strongest lightning emission features within the cloud top optical spectra produced by the neutral oxygen lines in the near infrared.

Note: The OI(1) line at 777.4 nm is made of three lines of nearly equal intensity with a total separation of 0.34 nm.

Background images

The LI shall generate LI triggered events and LI background radiance images over the area of coverage.

The repeat cycle for averaging and transmission of the LI background radiance images shall be less than 60s.

Level 1 Radiometric Requirements

Unless otherwise stated the requirements in this section apply:

- for all illumination conditions,
- over each MTG-I satellite nominal operational lifetime,
- for a 50% cloud cover of the earth,
- for an average cloud albedo of 80%.

For any lightning optical pulse whose radiance has a spatial distribution and a temporal evolution enclosed within the templates of, the LI shall ensure that:

a) The LI average detection probability\(^2\) is greater than
   - 0.90 at 45°N latitude, SSP longitude

\(^2\) Definition of Detection probability is given in Section 1.1.
- 0.70 in average over the whole instrument coverage area
- 0.40 as a minimum anywhere over all EUMETSAT member states as a goal

b) The LI detection efficiency for any radiance sample, whose LI effective radiance from the given lightning optical pulse is raised to a value higher than the minimum specified radiance \( L_{LP_{min}} \), is greater than
- 0.90 at 45°N latitude, SSP longitude
- 0.70 in average over the whole instrument coverage area
- 0.40 as a minimum anywhere over all EUMETSAT member states as a goal

The value of \( L_{LP_{min}} \) is a function of the local background \( L_{bkg} \) (in \( \text{Wm}^{-2}\text{sr}^{-1}\mu m^{-1} \)) in accordance with the following formula:

\[
L_{LP_{min}} = 6.7 \sqrt{1 + 0.02 L_{bkg} \text{mW} / (\text{m}^2\text{sr})}
\]

Note 1: The background radiance generated by the clouds is to be estimated according to the formula delivered in [MTG,2011a]. Where \( E_s \) is the sun extraterrestrial spectral irradiance = 1164.3 \( \text{Wm}^{-2}\mu m^{-1} \) at 777.4 nm

Note 2: With the above formula, \( L_{LP_{min}} = 6.7 \text{ mWm}^{-2}\text{sr}^{-1} \) during night conditions, \( L_{LP_{min}} = 16.7 \text{ mWm}^{-2}\text{sr}^{-1} \) in day conditions, Summer solstice, 45°N latitude, SSP longitude, cloudy conditions with albedo \( \rho = 0.8 \) (these radiance values are equivalent, for a 0.6 ms rectangular pulse, to an energy density of 4\( \mu \text{Jm}^{-2}\text{sr}^{-1} \) and 10\( \mu \text{Jm}^{-2}\text{sr}^{-1} \), respectively)

Note 3: The specified DE values are aimed at optimising the instrument performance over the EUMETSAT member states.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Radiance vs. cross section

670 mW.m\(^{-2}\).sr\(^{-1}\)

45 km

45 km

10 km

LDmin

Distance on ground

Top view

10 km

100 km

Distance on ground

LTsma x(t) maximum radiance spatial template

LTsmin(t) minimum radiance spatial template

LTs(t) sample pulse

(b) Template for the spatial distribution of lightning pulse radiance (not to scale).

Note 2: For any t,x,y within the templates range, \(L_{T_{\text{min}}}(x,y) \leq L_{LP}(x,y,t) \leq L_{T_{\text{max}}}(x,y)\)

Figure 3. Lightning Pulse Temporal and Spatial Templates.

[SRD] LI-15510 iss: 3B | Level 1b | [EURD] LI-06100

The LI shall provide the lightning event radiance measured in the spectral interval centred at 777.4 nm and having a width of 0.34 nm, over the full range from 6.7 to 670 mWm\(^{-2}\).sr\(^{-1}\), with an error (at 1\(\sigma\)) less than:

a) 10% relative accuracy for radiances higher than 70 mWm\(^{-2}\).sr\(^{-1}\)

b) 7 mWm\(^{-2}\).sr\(^{-1}\) absolute accuracy for radiances lower than 70 mWm\(^{-2}\).sr\(^{-1}\)

[SRD] LI-15540 iss: 3B | Level 1b

The LI shall provide the LI background spectral radiance with an accuracy (at 1\(\sigma\)) better than:

a) 10% relative accuracy for spectral radiances higher than 1 Wm\(^{-2}\).sr\(^{-1}\).\(\mu\)m\(^{-1}\)

b) 0.1 Wm\(^{-2}\).sr\(^{-1}\).\(\mu\)m\(^{-1}\) absolute accuracy for spectral radiances lower than 1 Wm\(^{-2}\).sr\(^{-1}\).\(\mu\)m\(^{-1}\).
The LI shall measure and transmit to ground all the lightning events, as derived from LI-15390, assuming a maximum lightning pulse rate of 800 per second, an average lightning pulse duration of 1.0 ms (above L_{pmin}) and an average lightning pulse footprint area of 400km².

Note: The requirement is aimed at sizing the detector readout and downlink capacities in terms of lightning events per second.

**False Alarm and False Flash Rate (FAR/FFR)**

The LI False Alarm Rate (LI-FAR) shall be lower than 350 per second after ground processing (level 1b), based on the conditions given below:

a) the false event filtering uses only LI instrument data and LI instrument characterization and calibration information.

b) if spatio-temporal coherence is to be used for FAR reduction then a LI triggered event can be considered a false event if no other LI triggered event occurs within 0.5 seconds and within a distance of 50 km, assuming that false events are uncorrelated.

The LI False Flash Rate (FFR) shall be lower than 2.5 [TBC] per second after enhanced ground processing with no restrictions applied to the filtering applied, i.e. exploiting the optimum algorithms.

**Level 1 Spatial and Temporal Requirements**

The LI shall provide a spatial sampling less than or equal to 10km at 45°N for the sub-satellite longitude.

It shall be possible to determine the LI triggered event time with an accuracy less than or equal to 5ms (1 ms goal) with respect to OBT.

**Level 1 Geometric Requirements**
The absolute value of the LI absolute sample position knowledge error (ASPKE) evaluated over the complete Full Disc Coverage (FDC) shall be less than 4 km (112 µrad) at SSP, at a 99.73% confidence level.

Note: This refers to geo-location error of each detector element in the detector array after ground processing.

2.3.1 Special notes on system requirements

Both detection efficiency and false alarm/false flash rate requirements are instantaneous, i.e. with no temporal averaging.

Regarding the FFR values, it can be noted that the global true flash rate from the LIS/OTD statistics has been used to set feasible requirements. For the MTG LI FOV, an average (annual) flash rate of ~25 flashes/s is estimated based on these statistics (global rate 44+/−5 fl/s).

The Detection Efficiency is linked to the lightning optical pulses as defined in Section 1, i.e. the requirement on DE is at the top of a cloud. The issue whether the optical pulse detection from space would be more sensitive to IC or CG lightning is not so straightforward. Some evidence suggests that CG optical pulses have a tendency to actually be brighter than IC flashes. On the other hand, CG flashes with lightning channel elements contributing to the optical pulse only lower in cloud will appear dimmer to optical detection at top of cloud [Koshak, 2010].

The Lightning Imager cannot distinguish between IC and CG lightning on a flash-to-flash basis, or between lightning occurring at different altitudes.

Based on a long-term (2002-2010) data record of LIS statistics, it is shown that the radiant energies measured are as follows:

- Daytime event energies, mean 15.1 uJ/ster/m2, median 8.3 uJ/ster/m2
- Nighttime event energies, mean 10.8 uJ/ster/m2, median 5.2 uJ/ster/m2

It should be noted that these values related to events and not flashes, meaning a triggered LIS pixel due to an optical lightning pulse.
2.3.2 Design drivers

In order to enable lightning detection both at night as well as during the day, when the lightning signals are superimposed on high background signals from clouds, various filtering techniques have to be implemented in the instrument design:

- Spatial filtering.
- Temporal filtering, by using a short enough detector exposure time to distinguish between lightning signals and background signals from clouds.
- Spectral filtering, by measuring only the enhanced lightning spectral radiance at one specific wavelength, in this case at 777.4 nm. For this purpose a narrow bandpass filter is used.

Figure 4. Event radiant energy of LIS observations (2002-2010) with a day/night separation.
In order to suppress the background and the background noise from the clouds it is important that both the detector exposure time and the filter spectral bandwidth are small.

In order to identify lightning signals the detectors will be read out continuously with a high repetition rate. The signals will be compared to a threshold value that has been predefined or calculated by a certain algorithm. The thresholds will have to be set for each individual detector pixel in order to account for the background signal variations as a function of geolocation and time of day. The threshold will be carefully balanced with respect to the expected lightning signals to achieve a combination of high lightning detection efficiency and a sufficiently low False Alarm Rate (FAR). Even so, the FAR will be considerable, and the L0=>L1b=>L2 data processing software will have a task to reduce the FAR as much as possible using a variety of possible techniques, using both instrument and satellite internal data. Optionally, such supporting data for the filtering tasks could also come from the imager on the same MTG-I platform, i.e. the Flexible Combined Imager (FCI).

2.3.3 Restricted mission operations

Because of strong expected thermal distortions on the satellite structure, for the time periods from 5 minutes before the start of the eclipse until 120 minutes after its maximum depth, the absolute sample position knowledge error is relaxed from 4 km to 8 km.

In addition, and due to the straylight impact on the instrument radiometry, during periods when the sun is between 8° and 16° from the instrument boresight, the FAR is increased by a factor of 2.

2.3.4 Timeliness requirements of MTG L1 L2 products

The MTG System Requirements Document (SRD) specifies the timeliness requirements as follows:

- L2 products (groups/flashes): 30 seconds (target), 2 minutes (threshold).

---

3 Timeliness follows the MTG definition, as specified in the MTG Conventions and Terms document (EUM/MTG/DEF/08/0034). In essence, it states that timeliness is the time difference between the foreseen end of acquisition of the last contributing data by EUMETSAT and the end of reception of the corresponding data (processed) by the users.
3 MTG LIGHTNING IMAGER

3.1 Main characteristics

The LI instrument consists of the following main elements:

- A Lightning Imager Optical Head (LOH) consisting of four identical cameras, each camera covering a different part of the visible earth, as depicted in Figure 5. The covered percentage of the total visible earth disc amounts to about 86% and includes the European territories of all EUMETSAT Member States as shown in Figure 6. Each camera is equipped with:
  - A protective cover on the baffle aperture to prevent baffle and optics contamination during launch and pre-launch activities.
  - Baffle for stray light suppression.
  - A telescope with a Field Of View (FOV) to cover about ±5.1 degrees. The entrance aperture is 110 mm in diameter.
  - A solar rejection filter (SRF) that is designed to block as much as possible all sun light outside the wavelength range of interest 770-785 nm. Within the wavelength range of interest the filter transmission is as high as possible: >0.95. The performance of this filter is critical, because all unwanted light entering the system after the filter may cause thermal problems, stray light problems and optical degradation problems.
  - A spectral narrow bandpass filter (NBF) with an equivalent band width of 1.9 nm centred on the main atomic oxygen triplet lightning emission line at 777.4 nm. The performance of this filter is critical, because it determined to a large extent the signal-to-noise behaviour of the LI instrument for lightning occurring over bright clouds. If selected to broad, the filter transmits the white light from the cloud, whilst the lightning signal itself is not increased. The cloud background signal will introduce significantly more shot-noise and the overall instrument signal-to-noise is reduced, which in turn reduces the LI lightning detection efficiency. If the filter band width is selected too narrow, the lightning spectral signal may be cut off by the filter response (also as function of incidence angle of the filter), which also reduces the signal-to-noise and the LI lightning detection efficiency. The NBF filter band width has to be carefully optimised. The transmission of the NBF at 777.4 nm is >0.90.
  - Imaging optics to image the earth on the detector(s).
  - Backside illuminated CMOS detector(s) of 1170x1000 pixels with on-chip ADCs. The detector performance characteristics are critical for the overall LI performance. The quantum efficiency (QE) is required to be >0.70.
ADCs have 11 bit resolution (goal 12 bits) in order to meet the radiometric accuracy requirements. The 24 µm x 24 µm detector pixel pitch corresponds to a 4.5 km x 4.5 km ground sample at the subsatellite point. The detector pixel full well capacity is required to be >450000 electrons in order to cope with the radiometric dynamic range.

- The proximity electronics, also referenced as Front End Electronics (FEE), consisting of electronics boards and relevant frames and covers, are supported by the Focal Plane Assembly (FPA) structures.

- A Lightning Imager Main Electronics (LME) box located 1-2 meters away from the LOH. The LME box takes care of processing all triggered events (true and false) and packaging the data for downlinking to fit within the allocated 30 Mbps bandwidth.

- LOH to LME interconnection harness.

The LI mass is about 93 kg (including 15% contingency margin; required ≤93 kg), the average power consumption is about 194 W (required ≤320 W) and the data rate is 16.9 Mbps (required ≤30 Mbps).
3.2 Lightning pulse detection concept and functional description

The concept adopted for the lightning detection drives the overall architecture of the instrument, including the selection of the CMOS detector architecture, of the hardware and software needed to manage the detector and to process the measured images. For these reasons, the lightning pulse detection concept is initially introduced, so that a better understanding of the resulting LI architecture can be obtained. The detection philosophy implemented in the proposed LI configuration is based on the following functional and performance aspects:

1. Image acquisition for continuous monitoring of the presence of lightning in the FOV.

2. Calculation of pixel by pixel adaptive background to cope with non-uniformities of the image (oceans, clouds, area in night conditions and areas with daylight conditions).

3. Use of an adaptive threshold, to fully exploit the detection capability over portions of the image presenting different illuminating conditions.

Figure 6. LI geographical coverage area.
4. Achievement of the maximum flexibility in the detection method, to allow even in flight, some tuning of the key parameters to fit lightning pulse characteristics as experienced during the operational life.

5. The system must allow on ground testing and characterisation of the key parameters needed to optimise all the key instruments performances.

The request to perform continuous monitoring of lightning presence (bullet 1), combined with the lightning pulse duration of typically 0.6 ms, leads to set the exposure time as close as possible to the pulse duration, and continuous readout of the detector. An exposure time of 1 ms and image rate of 1 kHz have been selected as trade-off between the lightning duration and the time needed by the electronics and software to perform the functionalities reported in bullets 2 to 4 in real time. In addition, the high frame rate, combined with the large amount of data to be processed, requires a high computational throughput that could not be achieved via the software architecture only. For this reason the detector processing is implemented in a hardwired logic.

The necessity to perform adaptive background computation (bullet 2) requires a dedicated logic that computes for every detector pixel the background level. This background calculation is performed with a dedicated filter that rejects noise effects and spurious events, but allows for low term variations to take into account that local background can change during the day or with the clouds movements. The removal of the background level to the overall pixel signal will therefore allow obtaining the net illumination level.

To fully exploit detection capability (bullet 3), the detection logic uses different thresholds for different illumination levels. In this way lower thresholds can be used in dark areas of the scene (having lower shot noise), using higher thresholds only in highly illuminated areas (areas with higher shot noise). For this purpose the LI instrument implements a function that determines for each pixel the detection threshold on the basis of the estimated background level.

In addition, to allow the maximum flexibility of the system, the transfer function associating a threshold level to the estimated background is stored in the software and commanded to the hardwired logic (bullet 4). Therefore adaptation of the thresholds could be performed even in flight, on the basis of the measured operational conditions.

3.3 Real Time Pixel Processor (RTPP) description, background determination and subtraction

The full frame digital video signal is processed by a Real Time Pixel Processor (RTPP) that performs the following tasks:

- Reference background pixel estimation for each detector pixel.
- Definition of the detection threshold for each detector pixel. This is accomplished with a lookup table stored in the RTPP mapping the threshold
versus the background level. This look up table can be modified by the software running in the LI CPU.

- Pixel thresholding: each pixel signal is compared against the defined detection threshold.

- Detection Transient (DT) address identification: the address of each detected pixel is transmitted to the Control Logic (CL).

- Storage of pixel signal of each detected pixel and of the surrounding pixels.

- Storage of the reference background for each detected pixel and of the surrounding pixels.

The reference background estimation is performed in real time, individually for each pixel, in order to provide an accurate reference for pixel data thresholding. The estimating function is optimised considering two opposite needs:

- Long filtering (integration) interval is needed to reduce the reference background noise and to increase the signal-to-noise of the background-corrected detector pixel signals.

- Short filtering (integration) interval is needed to promptly follow the variation of the background radiance caused by natural phenomena (as changing illumination condition and cloud movement) and by spacecraft artefacts (e.g. pointing drift and microvibrations).

### 3.4 Window Manager

The Window Manager processes the events from the 60 RTPPs (4 ASICs, each containing 15 RTPPs). The Window Manager performs the following tasks:

- Read the coordinates from the event address memory in the RTPP;

- Perform clustering of adjacent events (i.e. falling on adjacent pixels of the detector), considered as belonging to the same lightning.

- Define the area of interest (window coordinates), containing the lightning image according the following requirements:
  
  - The window shall be rectangular in shape, with a maximum dimension of 8x8 pixels, corresponding at SSP to 37x37 km².

  - Optical pulses exceeding this dimension (this should be a negligible percentage) will be split over multiple windows and transmitted to ground as separate windows. The windows will include the cluster of Detected Transients plus an additional border of one pixel on left, right, bottom and upper sides of the envelopes of the cluster.
Acquire the signal data and the background data stored in the RTPP memory, for the pixels belonging to a window as above determined.

The Control Logic will manage the composition of the windows. A limit will be represented by the maximum number of windows that the FPGA can manage. In case this limit is exceeded no more windows are opened. In the case of an overflow, the Control Logic will forward to the LME the information that there are too many events generated by an ASIC, and command the ASIC to add the delta-thresholds to the lookup table output.

In addition to the new windows, identified on the basis of the DT distribution occurring in the actual frame, the FPGA Control Logic shall also be able to acquire video data for the whole set of windows identified in the previous frame (persistent windows), i.e. temporal windowing. This function is introduced to provide the higher quantity and quality of data to ground processing as it allows tracking the tail of the lightning pulse when the radiance could be lower than the detection threshold. A single frame of persistence (providing an overall acquisition interval of at least 2 ms) is considered sufficient to track the whole pulse duration considering that, as reported in literature, almost all lightning pulses have duration shorter than 1.5 ms. In case that a persistent window overlaps with a new window, the persistent window is not opened.

Acquisition of a persistent window can be enabled or disabled via telecommand (TC). The “Neighbour function” can also enabled or disabled via TC. If disabled, only the detected pixels are transmitted.

The data relevant to persistent windows will be sent from each optical channel to the LME, but the on-board software calculates the sum of the signals of the pixels belonging to the persistent window and will transmit only this data (together with the pixel coordinate) to the ground.

Thresholds applied for detection of transients will not be included in the DT data packet as both the content of the lookup tables and the reference background value used to link to the threshold in the lookup table are known to the ground processing. In addition, any autonomous software intervention on thresholds (i.e. coarse threshold regulation), will be reported to the ground as part of the LI instrument auxiliary data.

### 3.5 Single Pixel Filtering

The “Single Pixel Filter” is intended to keep single triggered events closer to the edge of the earth, when the lightning optical pulses become more and more elliptically shaped (rather than circularly shaped). The filter is a second-stage on-board filtering (after the first-stage on-board filtering that only keeps adjacent triggered events) and acts on isolated triggered pixels only. It compares the sum of the signals of the 3x3 pixels placed around the detected pixel against a threshold value. The comparison is done after subtracting the estimated backgrounds from the candidate triggered events and its surrounding pixels. If the sum of surrounding pixels is higher than the used threshold, the event is called a lightning triggered event and it is forwarded to the LME. Otherwise the event is flagged as a false event and the window is discarded. The “Single Pixel Filter” can be disabled for certain detector areas,
defined by a dedicated configuration memory, in order to avoid the lightning event filtering for geolocations where the minimal optical pulse of 10 km diameter is as large or larger as a single detector pixel.

3.6 Science Data Packetiser

The Science Data Packetiser prepares the science data to be sent to the LME. It collects the window starting address and dimensions, the pixels and the background data of a “valid” window and it provides this data package to the SpaceWire Link Control.

The data packets sent to the LME for each frame include:

- Housekeeping data relevant for the event data characterisation, including a frame counter used for identify the frame data.

- Events data, including windows pixels signal, related windows background, window coordinates and size (in case the “Neighbour function” is disabled, only pixel signal, background and coordinates of over threshold pixel are transmitted).

- Background data reference image. The RTPP computes continuously the background image. Each frame (1 ms) a portion of the complete detector background image (8 windows of 8x8 pixels each) is sent from each LOH channel to the LME. The full image is transmitted in about 2.4 seconds.
4 MEASUREMENT APPROACH

4.1 Transition from LEO to GEO observations

Feasibility of lightning detection from space by optical sensors has been successfully proven by the NASA instruments OTD (1995-2000) and LIS (1997-present) on low earth orbits [e.g. Mach et al., 2007]. One outcome of such data has been the first instrumentally uniform collection of longer term flash statistics within the coverage areas, see Figure 7.

![Figure 7. Annual flash density derived from NASA OTD and LIS observations (1995-2006).](image)

However, the transition from LEO to GEO observations is not straightforward, and certain elements need to be considered in order to understand the differences in e.g. processing of the observed data. Some of the factors of this transition can be regarded as positive, such as:

- Temporally, spatially and instrumentally uniform coverage of the visible disk over land and ocean,
- High temporal resolution,
- Excessive radiation noise hampering LEO missions (SAA – Southern Atlantic Anomaly) is likely less problematic for GEO observations.

In addition to the positive factors, there are also elements which cause a challenge to GEO observations:

- The larger distance affects radiometric sensitivity and resolution capabilities, and thus the instrument requirements and complexity,
- Fixed observation geometry with location dependent distortions and observation angles (northernmost member states),

- The spatial resolution does vary across the full disk coverage area of the LI more than in the case of LIS instrument, which varies from 4-8 km depending on viewing angle across the detector matrix. For LI, the requirement is a 10 km spatial sampling at 45 deg latitude SSP longitude. Considering the viewing geometry considerations, this is well achieved with the 4.5 km spatial sampling at SSP.

In addition, noise conditions (e.g. S/C external noise, solar glint, S/C attitude effects, cloud vs. S/C movement) are different to LEO, although both type of observations are subject to such effects and should be properly taken into account in processing.

The basic differences between LEO and GEO observations are also summarised in Table 2.

**Table 2. Summary of some basic differences of LEO and GEO lightning observations.**

<table>
<thead>
<tr>
<th></th>
<th>LEO (here: TRMM LIS)</th>
<th>GEO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geometry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altitude</td>
<td>400 km</td>
<td>36000 km</td>
</tr>
<tr>
<td>FOV</td>
<td>550 km x 550 km</td>
<td>Full disk</td>
</tr>
<tr>
<td><strong>Duration of observation</strong></td>
<td>90 seconds</td>
<td>continuous</td>
</tr>
<tr>
<td>Max. satellite zenith angle</td>
<td>42 deg</td>
<td>86 deg</td>
</tr>
<tr>
<td>Spatial footprint variation</td>
<td>Fixed to detection matrix</td>
<td>Depends on nadir angle</td>
</tr>
<tr>
<td><strong>Noise conditions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiation noise</td>
<td>Particles in van Allen belt</td>
<td>Solar flares</td>
</tr>
<tr>
<td>Solar glint</td>
<td>Changes with satellite motion</td>
<td>Always present at calculable positions</td>
</tr>
<tr>
<td><strong>Other effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satellite motion effects</td>
<td></td>
<td>Day and night segments always in FOV</td>
</tr>
<tr>
<td>Spacecraft microvibrations (“jitter”) at cloud edges</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.2 Lightning optical signal

A lightning optical pulse is defined in Section 1.1, when basic terminology used throughout the document is being introduced. The basic element of a lightning optical pulse is the scattering it encounters in the cloud, which eventually allows the signal to be observed at the cloud top. The results of scattering in the cloud is a broadening and delay of the optical pulse in time to about 600 µs, with a typical fast rise within 250 µs, followed by a fast (but slower) decay, with a total width between 300-800 µs. In addition, there will be a spatial widening up to several km and the partitioning of energy between upper and lower cloud surfaces [Light et al., 2003]. The size of the spatial widening depends on the scattering processes involved. The temporal characteristics of optical pulses measured with the FORTE system are illustrated in Figure 8.

![Figure 8. Example optical pulses measured with the FORTE system [Light et al., 2003].](image)

4.3 Measurement principle

The MTG LI measures an optical signal, which is the sum of the background signal and the lightning optical signal at a detector position. The detection principle consists of comparing this measured signal to a given threshold, after removal of the background for each detector element. The optical amplitude above the detection threshold is then identified as a lightning event. The background level depends largely on the local albedo (in particular the presence of clouds), and this albedo is largely changing with time, in particular because of the cloud motion and the evolving solar zenith angle. This basic measurement principle is illustrated in Figure 9.

The LI sensor will send the information to the ground relative to each detected lightning pulse including navigation, energy, and time. The maximum number of pulses per second is estimated to be in the order of 800 for an average lightning pulse footprint area of 400 km².

The lightning events will be sampled at a spatial resolution of 4.5 km (SSP) giving altogether 4.7 million pixels divided into 4 detectors. The integration time is 1.0 ms.
The volume of data transmitted to ground will then depend on the number of lightning events occurring at any instant, i.e. the data rate will be variable. The current budget is less than 30 Mbps including a full image of the background level downloaded every 60 seconds. It should also be noted that within the estimated peak rates (800 pulses/s) there should also be no difference in timeliness of lightning data to end-users.

**Figure 9. Detection principle of the MTG Lightning Imager, where a real observation from the space shuttle is used as an example image of how a storm with lightning could look like from space (in night-time conditions).**
5 PRODUCT GENERATION FROM L1 TO L2

In addition to on-orbit event detection, the data needs to be filtered to minimise the number of non-lightning events (NLE) and thus control the FAR. Ground processing of raw lightning data consists of the NLE filtering, geolocation, UTC time stamping of events, and a radiance conversion, after which L1b data is obtained. This is followed by L2 processing to do the clustering of the events to L2 groups and L2 flashes. Furthermore, the L2 accumulated products are generated with a 30 second duty cycle, and is based on the flash product with accompanying events. As filtering will benefit from the clustering progress, the filtering is an overall task encompassing both L1b and L2 processing and cannot thus be contained in only one processing step. A general overview of the MTG L1 processing chain is shown in Figure 10.

The L1b and L2 processing contain a number of possible NLE filtering steps, which will be based on experiences from LIS processing with some GEO-specific modifications.

The treatment of non-lightning events in this document is still limited, but shall be expanded when details are available.

5.1 Summary of non-lightning event/group/flash filtering

The NLE are caused by elements as summarised in Table 3. The provided list is only a general representation of the issue.

<table>
<thead>
<tr>
<th>Cause of NLEs</th>
<th>Description</th>
<th>Specific filter for processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>High energy particle collisions</td>
<td>Energetic particles impacting the detector matrix will cause NLEs of varying signatures, depending on the angle of impact.</td>
<td>Radiation filter</td>
</tr>
<tr>
<td>Noise (instrument, spacecraft etc)</td>
<td>Random noise without typical features of real lightning (temporal and spatial coherency)</td>
<td>Shot noise / coherency filter</td>
</tr>
<tr>
<td>Solar glint</td>
<td>Specular reflections from water surfaces will cause a very bright optical signal. After geolocation of data, the potential areas can be subjected to the specific filter.</td>
<td>Solar glint filter</td>
</tr>
<tr>
<td>Spacecraft motion (&quot;jitter&quot;)</td>
<td>Microvibrations may affect the detection process by affecting the background signal e.g. around cloud edges. The significance of jitter depends on the vibration frequency – a low enough frequency is not expected to be a problem for detection as a fast enough</td>
<td>Contrast filter</td>
</tr>
</tbody>
</table>
background signal updating should be able to compensate for the effect.
Figure 10. Overview of MTG LI processing chain up to L2.

5.2 Outcome of L1 processing: L1b data

The L1b data set consists of three data products:

- L1b Lightning data product. This consists of all triggered events with a flag to distinguish between false/true events. In addition, also other quality control flags are to be included (e.g. reason for filtering, reliability of a true event...).

- Background radiance product (including geolocation for all ground pixels). This is based on triggering all detector elements in a given time interval.

- Calibration data product.

The three L1b data products are implemented in the netCDF format.

The L1b background radiance product is made available every 60 seconds. Since the background is provided for each detector pixel, the granule size is always the same. Background data collection should not interfere with lightning observation.

The L1b geolocation grid product provides the geolocation grid attached to each ground pixel. This data product is made available every 60 seconds and its granule size is always the same. This can be part of the background data product.

The L1b calibration data products can have variable durations and sizes, depending on which calibration measurements have been performed. More details will be provided once the details of the LI instrument design and its operational scenario are known.

5.3 Baseline L2 Lightning data product

The baseline L2 lightning data product consists of a full “flash tree”, with the related groups and events related to a given flash. This is illustrated in Figure 11. Please refer to Section 1.1 for a definition of flash/group/events. As a distinction to L1b, only true events are included in the L2 “flash tree”, since the false events are not linked to any flash.

The groups and flash locations are computed based on radiance weighted centroids of events.

Lightning flashes are composed of multiples of groups caused by lightning, which are in turn grouped by proximity in location over a flash duration. It is expected that the flashes will be the primary L2 product for most users. Since the MTG LI is observing the optical emissions at cloud tops, no separation between CG and CC flashes can be done with using LI data alone. As a synergetic product (see Section 11 for further discussion) with ground-based networks, a separation of flash types might be envisaged. It should be noted that since the flash clustering algorithm is based on tracking the temporal and spatial distance between groups, no pre-
determined time limit has been placed on flashes. As long as there are additional groups within the algorithm defined temporal and spatial “windows”, a flash will in principle not be terminated. For practicality reasons a cut-off will however be implemented in the SW so that the processing buffer does not need to store a very large amount of groups in the very limited number of extra-long flash cases to be expected. Flashes that have been cut-off due to such a fixed limit can be flagged and combined in post-processing. The cut-off time is [TBC] and depends on feasibility of processing, buffer implications and the foreseen number of cut-off flashes. A long enough time (~10s) would minimise the number of cut-offs and the need for later recombining, while still keeping the flash buffer at a reasonable size.

As such, flashes (with the addition of the related groups and events) are the basic L2 products for MTG LI. Detailed description of any other possible products are beyond the scope of this ATBD, but a brief outlook into various possibilities with or without synergy with other measurements (ground-based or satellite) is provided in Section 1.

---

**Figure 11. Flash tree example describing the contents of the baseline L2 product.**
5.4 L2 accumulated products

The full flash tree (flash/group/event) with this information is provided with a target timeliness of 30 sec. Based on TRMM LIS, where the data ratio is roughly 1:12:56 (flashes:groups:events), the MTG LI data is potentially also very heavily event-dominated. However, EUMETSAT users are mainly interested in:

- Flashes, and an
- Understanding of “what kind of a flash” it is they are getting, i.e. “strength”, duration, and extent of the flash. In the case of optical measurements from space, the strength of a flash is expressed as radiances/radiant energies, where a direct link to peak currents as observed in ground-based systems is not considered feasible.

The assumption is that most users would be well served with the flashes (groups) and a supporting accumulated product describing the extent of the flashes, coming from the full flash tree product. The periodicity of the product shall be such that it fits well with the L2 timeliness or any internal buffering (30 sec) and/or the FCI product cycle (2.5/10 min). In practice, by producing such accumulated products with a 30 sec periodicity, any multiples of these can easily be constructed by the user, including the FCI-matched 2.5/10 minutes.

In order to use the full amount of information available from the L2 Lightning data product, the extent of the accumulated products will be defined by events, but units and values are defined by the flashes.

The products are not “density” products as such, since a division with pixel area (in km$^2$) is not done in the computation, and is left for the users for any given area, with the help of detailed grid information provided to the users. It is assumed that cumulative flash counts will be useful also as they are. Another complication for creating a full disk density product is the varying pixel size along the LI FOV. Since the density value would result from dividing the flash count by the varying pixel size, this means that storms of similar “strength” would appear different in the density plots depending on the LI pixel size, if not normalised in some way.

There are two possibilities for computing the index value for the accumulated product, both of which have their potential applications, as shown below. In addition, also the distribution of flash radiances can be presented as an accumulated product. The three different accumulated products are introduced in Table 4.

A more detailed description of the accumulated product computation is given in Section 7.
### Table 4. L2 accumulated product characteristics.

<table>
<thead>
<tr>
<th>Product Name</th>
<th>Periodicity</th>
<th>Unit</th>
<th>Grid</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 accumulated flashes</td>
<td>30 sec</td>
<td>flashes/30sec</td>
<td>2 km</td>
<td>Grid box values are the accumulated flashes (in all those grid boxes affected by the LI events responsible for the flash) during 30 seconds. Each flash during a 30 second interval contributes to the grid box value if that flash has an event inside the grid box. The grid box value increases by $1/N_i$ for each flash comprised of $N_i$ events in separate grids. <strong>This allows the user to integrate over a sub-area of the full FOV to get the cumulative flash count for that sub-area.</strong></td>
</tr>
<tr>
<td>L2 accumulated flash index</td>
<td>30 sec</td>
<td>flashes/30sec</td>
<td>2 km</td>
<td>Grid box values are the accumulated flashes (in all those grid boxes affected by the LI events responsible for the flash) during 30 seconds. Each flash during a 30 second interval contributes to the grid box value if that flash has an event inside the grid box. The grid box value increases by $I$ for each flash having an event inside the grid box. This allows the user to observe how many flashes have impacted that particular grid element in the given time. <strong>With this method, the accumulated flash amount is only correct for each grid element alone in the accumulated product grid.</strong></td>
</tr>
<tr>
<td>L2 accumulated flash radiance</td>
<td>30 sec</td>
<td>$\mu J/m^2/\mu m/30sec$ (TBC)</td>
<td>2 km</td>
<td>Grid box values are counted by dividing the radiances affecting that grid box by the number of grid boxes subject to each flash. This allows the user to integrate over a sub-area of the full FOV to get the cumulative flash radiances for that sub-area.</td>
</tr>
</tbody>
</table>

#### 5.5 Data requirements

A summary of the basic data requirements for L2 processing are shown in Table 5.
### Table 5. Data requirements for L2 processing.

#### Input data

<table>
<thead>
<tr>
<th>Data</th>
<th>Note</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1b Background radiances</td>
<td>Needed for geolocating the clustered groups and flashes in a lat/lon grid</td>
<td>Required</td>
</tr>
<tr>
<td>L1b Events</td>
<td>Lightning events, with a flag indicating true/false events (after filtering).</td>
<td>Required</td>
</tr>
<tr>
<td>MTG FCI grid information</td>
<td>The L2 accumulated products are processed in the FCI IR grid (2 km). Therefore, relevant grid information needs to be available to the accumulated product processing.</td>
<td>Required</td>
</tr>
<tr>
<td>ATDnet data</td>
<td>Possible use for redundancy filtering (storm identification)</td>
<td>Optional</td>
</tr>
<tr>
<td>MTG FCI L2 data</td>
<td>Possible use for false event filtering and general quality control by identifying cold cloud tops (high probability lightning intensive areas).</td>
<td>Optional</td>
</tr>
</tbody>
</table>

#### Output data

<table>
<thead>
<tr>
<th>Data</th>
<th>Note</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 Events</td>
<td>Lightning events, containing only true lightning events. In addition, events have possibly been removed by the redundancy filtering during L2 processing.</td>
<td>Required for Flash clustering</td>
</tr>
<tr>
<td>L2 Groups</td>
<td>Intermediate product to reach L2 Flashes</td>
<td>Required for Flash clustering</td>
</tr>
<tr>
<td>L2 Flashes</td>
<td>Final lighting data product of the baseline L2 processor</td>
<td>Final outcome disseminated to users.</td>
</tr>
<tr>
<td>L2 accumulated flashes</td>
<td>Accumulated flashes (normalised by grid affected by each flash) in a 30 sec temporal scale and a 2 km grid.</td>
<td>Final outcome disseminated together with the L2 Flash product.</td>
</tr>
<tr>
<td>L2 accumulated flash index</td>
<td>Accumulated flashes (not normalised) in a 30 sec temporal scale and a 2 km</td>
<td>Final outcome disseminated together with the L2 Flash</td>
</tr>
</tbody>
</table>
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

| L2 accumulated flash radiance | Accumulated flashes radiance in a 30 sec temporal scale and a 2 km grid. | Final outcome disseminated together with the L2 Flash product. | grid. | product. |
6 ALGORITHM DESCRIPTION FOR FLASH GENERATION

6.1 Algorithm overview

A general (high level) overview of the MTG LI clustering algorithm is presented in Figure 12, whereas more detailed flow charts of the group and flash clustering processes are shown in Figure 13 and Figure 14, respectively.

The flow charts shown are preliminary, with more details added when the processor maturity is developing.

The MTG LI clustering algorithm is based on a strong heritage from the OTD/LIS clustering algorithms, and also including modifications applied to the GLM LCFA algorithm. The differences (TBD) are mainly due to a differing time integration period between the MTG LI and LIS/GLM.

Although the baseline algorithm is based on a LIS heritage with only limited changes, there are significant improvements in the data usage and applications. These are stemming from the general difference of GEO vs. LEO observations, including continuous monitoring and homogeneous observations over the full disk. Please refer to Section 4.1 for a more detailed discussion on these aspects.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Processing of raw data up to **Level 1b** (NLE filtering, UTC time stamp, geolocation, radiance conversion)

---

**L1b**

---

**L2**

- Event buffer of N sec received for L2 processing
- Events in buffer?
  - no
  - yes
    - Redundancy filtering (for quality control of L1b output)
    - Clustering of groups and flashes
    - Write out all flash data (flashes, groups, events)
    - Storage

---

- Dissemination to users
- Application display

*Figure 12. Flow chart of MTG LI L2 clustering algorithm overview.*
Figure 13. Flow chart of MTG LI group clustering. The lower left box updates the “flash statistics” (and not “group statistics”) as these statistics are on the root level where all flashes and the related groups and events are traced.
Figure 14. Flow chart of MTG LI flash clustering.
6.2 Significant algorithm details

6.2.1 Redundancy filtering at event level and flash filtering

As a preliminary step before the clustering of events into groups and flashes, the data that has been filtered for NLEs in the L1 processing will be subject to an additional filtering step to verify the L1b data. There are various possibilities regarding what will be included in this processing step. However, at the current stage of development further study is still required to analyse their advantages and disadvantages. These include for example the use of FCI L2 products such as Optical Cloud Analysis, or ground-based ATDnet data for higher-level storm identification. In the overview shown in Figure 10, a redundancy filtering in separate stages of L2 processing is envisaged, as follows:

- The first part, called **redundancy filtering at event level**, is taking place at the beginning of L2 processing (quality control). This aims at removing NLEs from coming from the L1b processing.

- The second part is immediately following the L2 flash clustering, and can be called a **1st level flash filtering**. Flash filtering allows removing flashes (with all related events) from the final L2 flash products.

- The third part (**2nd level flash filtering**) is taking place after the accumulation product creation, i.e. after the 30 second collection period. This allows the removing of entire non-lightning related flash structures, if present. Similarly to the 1st level flash filtering, this filtering will remove also all related groups and events.

Uncertainty of the L1b processing specifications will potentially impact the amount of extra NLE filtering that is required in the Level 2 processing. Further studies by the LIST and with IFCT are expected to resolve this in order to provide more information for the specifying the filtering steps needed at L2.

6.2.2 Group clustering

As a reference to the LIS/GLM approach, the group clustering will be done in the pixel space (each pixel with X/Y-coordinates), where a lat/lon-conversion is not necessary before clustering. LIS/GLM have an integration time of 2 ms, which is limited by the technology chosen (CCD), whereas the typical optical pulse characteristics indicates that a somewhat smaller integration time would be more optimal in order to maximise signal and minimise noise. For the LI, a baseline integration time of 1 ms has been chosen by industry. This integration time is parameterised and one of the tasks during commissioning is to evaluate the impact of various integration times in different conditions (e.g. day/night time). Due to the most likely shorter integration time, the group clustering algorithm shall contain an option of including the adjacent integration frames in the group clustering.
The algorithm shall also retain flexibility in defining how many temporally adjacent frames can be included in a group \([1\ldots N]\), or how far in the spatial dimension the adjacency principle\(^4\) can be applied. This should allow a testing of the stability and robustness of the group clustering with reference to the adjacency definitions.

6.2.2.1 Group location

The groups locations are computed based on radiance weighted centroids of events. Although the clustering itself is done in the X/Y space (pixel coordinated) the group locations are expressed in geographical coordinated (Lat/Lon). Therefore, event geographical coordinated are needed for this computation. Computing the group locations is done as follows:

For Latitude \((\varphi_{\text{group}})\) in degrees:

\[
\varphi_{\text{group}}(\text{deg}) = \frac{\varphi_{\text{event} \#1} \cdot L_{\text{event} \#1} + \varphi_{\text{event} \#2} \cdot L_{\text{event} \#2} + \ldots + \varphi_{\text{event} \#N} \cdot L_{\text{event} \#N}}{L_{\text{event} \#1} + L_{\text{event} \#2} + \ldots + L_{\text{event} \#N}}
\]  

(1)

For Longitude \((\lambda_{\text{group}})\) in degrees:

\[
\lambda_{\text{group}}(\text{deg}) = \frac{\lambda_{\text{event} \#1} \cdot L_{\text{event} \#1} + \lambda_{\text{event} \#2} \cdot L_{\text{event} \#2} + \ldots + \lambda_{\text{event} \#N} \cdot L_{\text{event} \#N}}{L_{\text{event} \#1} + L_{\text{event} \#2} + \ldots + L_{\text{event} \#N}}
\]  

(2)

where

\(\varphi_{\text{event} \#1,2,\ldots,N}\) = latitude of events \#1,2,\ldots,N,

\(\lambda_{\text{event} \#1,2,\ldots,N}\) = longitude of events \#1,2,\ldots,N,

\(L_{\text{event} \#1,2,\ldots,N}\) = radiance of events \#1,2,\ldots,N.

6.2.2.2 Example group clustering sequence

In order to illustrate a simple case of group clustering, the example provided in Figure 9 is used as a starting point, with some assumptions on event timings. An example group clustering sequence is shown in Figure 15. The timing of the events is for this example

\(^4\) The core of the clustering algorithm is the combining ("clustering") of events, and later in the processing groups, based on how close they are to each other both spatially and temporally. The term “adjacency principle” refers to this methodology, where “adjacent” events/groups are clustered into flashes. It is a “principle”, since it does not mean that the elements need to be literally adjacent in the detector grid. This relates to how flashes are defined, which then again is transformed into how the clustering algorithm is defined.
arbitrary, being however at the same time realistic. In order to simplify the example, only the same time integration frame will be considered for the group clustering.

![Diagram of detected events in a detector X/Y matrix format covering several integration frames (of 1 ms length).](image)

**Figure 15.** Detected events shown in a detector X/Y matrix format covering several integration frames (of 1 ms length).

The group clustering sequence for this example is described below:

**Group #1:** The first event in the buffer is assigned to the 1st group. Since only the same time frame is considered for group clustering in this example, for simplicity reasons, the first group remains a single-event group.

**Group #2:** Event #2 is assigned group #2. Also event #3 is in the same timeframe, and gets assigned to the same group.

**Group #3:** Events #4 and #5 are adjacent and in the same time frame, hence are assigned to group #3.

**Group #4:** Events #6, #7, and #9 are in the same time frame, and are only adjacent to each other and are assigned to group #4.
Group #5: Event #8 is in the same timeframe as the previous events but not adjacent to any of them, and is hence assigned a new group #5.

Group #6: Events #10 and #11 are adjacent and in the same timeframe, and are assigned a new group #6.

Group #7: Event #12 becomes a single-event group #7.

Group #8: Event #13 is in the same time frame as event #12, but is not adjacent. Hence, it is assigned a new group #8.

Group #9: Event #14 is adjacent to event #5, but in a much later time frame. It is also assigned a single-event group #9.

Therefore, from 14 events we have clustered 9 groups. In reality, and based on LIS experiences, this event-to-group ratio would be much higher (in the order of 4:1 or 5:1) but for editorial reasons the example is simplified.

6.2.3 Flash clustering

The flash clustering of MTG LI should follow as closely as possible the approach of LIS/GLM in order to minimise processing related differences in data output. For OTD (in operation from 1995 to 2000), a method of box clustering was applied, where any group which was closer than 16.5 km from a flash or was separated in time with no more than 330 ms was considered to belong to the given flash. For LIS, a more elaborate comparative method has been developed, which is based on Weighted Euclidian Distances (WED) in spatial and temporal spaces. This method will also be applied to the GLM flash clustering. For the LI, both methods shall be available in the operational processor (L2PF), to be evaluated during commissioning and other supporting studies.

6.2.3.1 Flash clustering based on the 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 multigroup flash, such that the following relationship is fulfilled [Mach et al., 2007]:

\[
WED_{LIS/GLM} = \sqrt{\left(\frac{X}{X_{\text{diff}}}\right)^2 + \left(\frac{Y}{Y_{\text{diff}}}\right)^2 + \left(\frac{T}{T_{\text{diff}}}\right)^2} \leq 1.0, \tag{3a}
\]

where \(X\) and \(Y\) are the spatial distances in the X and Y dimensions, respectively, and \(T\) is the time difference between the two groups (in ms). The values for \(X_{\text{diff}}\) and \(Y_{\text{diff}}\) are also design dependent as they are related in addition to flash properties in general, also to the achievable pixel size. For LIS and GLM, the values used in processing are 5.5 km and 16.5 km, due to the ~5km and ~10km pixel sizes, respectively. On the other hand, \(T_{\text{diff}}\) is based on LIS
group/flash statistics, and a value of 330 ms is the default value for LIS (and presumed for GLM).

For MTG LI, as slight modification of the WED-algorithm is used, which simply combines the \((X, Y)\) differences into a direct distance difference:

\[
WED_{MTG LI} = \left( \frac{D}{D_{diff}} \right)^2 + \left( \frac{T}{T_{diff}} \right)^2 \leq 1.0, \hspace{1cm} (3b)
\]

where \(D_{diff}\) and \(T_{diff}\) are the distances of two groups in the spatial and temporal space, respectively. The exact values are still TBD, but default values close or equal to those for GLM are regarded as default.

A group-by-group comparison is conducted, such that a group only has to be spatially or temporally near one other group in a flash to be a part of a multigroup flash.

### 6.2.3.2 Flash clustering based on separate distance and time threshold criteria

With the second approach the distance and time elements are kept separate so that there are no trade off effects e.g. with spatially close groups that are separated by a longer time such as with the WED algorithm. In this method two groups, A and B, are considered as part of the same flash only when both of the following conditions are met:

\[
\begin{align*}
D &\leq D_{diff} \\
T &\leq T_{diff}
\end{align*}
\hspace{1cm} (3c)
\]

where \(D_{diff}\) and \(T_{diff}\) are defined similarly as in the WED case.

### 6.2.3.3 Computing the distances \(D\) for flash clustering

Depending on the final MTG LI design, simulations will be needed to analyse the impact of clustering distance as it is depending on the pixel size. In principle it would also be possible to have a changing clustering distance value over the FOV, since the footprint on ground increases towards the edges of the FOV. Such considerations are still TBD. The current understanding is that GLM processing uses only one value over the full FOV.

The values of \(D\) are either defined as:

**Option 1:** As in LIS approach: group centroid distances (in km) in east-west and north-south directions, respectively, or
Option 2: As in GLM approach: event footprint distances (in km) of individual groups in east-west and north-south directions, respectively.

Note: Choosing between these two methods shall be parameterised in the flash clustering.

Computing values for $D$ in the WED algorithm, Option 1:

The LIS approach is a rather straightforward computation of a distance between group centroids in east-west and north-south directions, respectively. The individual events contained in the group do not influence the computation.

The distances are computed using the so called great-circle distances, which are the shortest distance between any two points on the surface of a sphere measured along a path on the surface of the sphere (as opposed to going through the sphere's interior).

Because the Earth is nearly spherical, and the separation of relevant groups in the flash clustering are very small in comparison to the Earth’s radius, the equations for great-circle distance can be used to calculate the shortest distance between points on the surface of the Earth.

The most accurate methods for computing such distances (up to millimetres) are based on the assumption that the figure of the Earth is an oblate spheroid, instead of the great-circle distance method which assumes a spherical Earth.

The haversine formula \cite{Sinnott, 1984} can be used which is a numerically better-conditioned approximation for smaller distances, which are typical in the case of the computations needed for flash clustering:

$$\Delta \sigma = 2 \arcsin \left( \sqrt{\sin^2 \left( \frac{\Delta \varphi}{2} \right) + \cos \varphi_1 \cos \varphi_2 \sin^2 \left( \frac{\Delta \lambda}{2} \right)} \right),$$  \hspace{1cm} (5)

where

$\varphi_1$ = geographical latitude of group #1 (of the compared pair),

$\varphi_2$ = geographical latitude of group #2 (of the compared pair),

---

\cite{Sinnott, 1984} The definition of the thresholds for the $X/Y$ clustering distances and/or how they are determined is still open, but the various options are presented. A study on the clustering strategies will be conducted in early 2014, resulting in recommendations on the various options in the clustering algorithms. However, due to the novelty of the instrument, it is seen as vital that the processor retains flexibility in the parameterisation and some of the inner functioning of the algorithms, to be investigated during commissioning.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

\[ \Delta \varphi = \text{absolute difference of latitude of groups #1 and #2 (of the compared pair)}, \]

\[ \Delta \lambda = \text{absolute difference of longitude of groups #1 and #2 (of the compared pair)}. \]

After computing the central angle \( \Delta \sigma \) between the two points, the distance \( d \) can be computed by:

\[ d = R \cdot \Delta \sigma, \]

where \( R \) is the Earth’s radius (a mean value of 6371.009 km can be assumed).

**Computing values for \( D \) in the WED algorithm, Option 2:**

Instead of using only group centroids for computing the distanced X/Y for the WED clustering algorithm, all events contained in the groups can be used to compute the distances. This means that many more distance pairs are computed, since each event in the group is compared to events in the other groups. Figure 16 illustrates the basic principle of Option 2. Two groups containing a number of events (blue circles) are compared to decide if they spatially could be a part of the same flash. However, the illustration does not visualise the temporal aspect of the WED algorithm.

The same computational methods (formulas) as for Option 1 shall be used for Option 2.
Figure 16. Illustration of how to compute the distances D in Option 2, where each event in a group are compared to events in the other group(s). The shortest distance measured is then the baseline for D. The red crosses indicate the location of the groups, which are based on the radiance weighted average locations of events.

6.2.3.4 Deviations from LIS algorithm

The GLM algorithm contains some modifications to the LIS algorithm, which are to be taken into account when finalising the MTG LI algorithm [TBC]:

- Replacing of “first fit” with “full fit” in flash clustering. The LIS/OTD algorithm uses a “first fit” to add new groups to a flash, where a group is assigned to the first overlapping flash in the buffer. In the GLM “full fit” approach a group is assigned to all overlapping flashes in the buffer, and if the group can indeed be assigned to more than one flash, all flashes it can be assigned to will be merged. This has been considered to be more effective, since the method in use in the LIS era (“first fit”) means that in the case of flashes with groups (= strokes) linking them would result in two separate flashes, whereas in reality it would require a single merged flash as a result.

- The current LIS algorithm uses group centroids (radiance weighted average locations) to cluster flashes. The GLM algorithm will use event footprints of individual groups to cluster the groups into flashes, which will prevent spatially large flashes from splitting into separate flashes. In practice, when implemented, this would mean that the WED algorithm (Eq. 1) would compare footprint (edge) distances instead of centroids.

- Replacing the \((X, Y)\) distances with a direct distance \(D\).

It should be noted that this baseline clustering algorithm is based on current LIS heritage, and modifications should be analysed as science develops. A further definition of a lightning flash in terms of VHF, VLF/LF and optical measurements can be reached through e.g. measurements campaigns such as CHUVA (which is part of Brazilian contribution to GPM) taking place in Brazil in the 2011-2012 timeframe.

6.2.3.5 Flash location

The flash locations are computed based on radiance weighted centroids of groups, as follows:

For Latitude \((\varphi_{\text{group}})\) in degrees:

\[
\varphi_{\text{flash}}(\text{deg}) = \frac{\varphi_{\text{group}#1} \cdot L_{\text{group}#1} + \varphi_{\text{group}#2} \cdot L_{\text{group}#2} + \ldots + \varphi_{\text{group}#N} \cdot L_{\text{group}#N}}{L_{\text{group}#1} + L_{\text{group}#2} + \ldots + L_{\text{group}#N}} \quad (7)
\]
For Longitude ($\lambda_{\text{group}}$) in degrees:

$$\lambda_{\text{flash}} \text{ (deg)} = \frac{\lambda_{\text{group} \#1} \cdot L_{\text{group} \#1} + \lambda_{\text{group} \#2} \cdot L_{\text{group} \#2} + \ldots + \lambda_{\text{group} \#N} \cdot L_{\text{group} \#N}}{L_{\text{group} \#1} + L_{\text{group} \#2} + \ldots + L_{\text{group} \#N}}, \quad (8)$$

where

\begin{align*}
\varphi_{\text{group} \#1,2,...N} &= \text{latitude of groups \#1,2,...N}, \\
\lambda_{\text{group} \#1,2,...N} &= \text{longitude of groups \#1,2,...N}, \\
L_{\text{group} \#1,2,...N} &= \text{radiance of groups \#1,2,...N}.
\end{align*}

### 6.2.3.6 Example flash clustering sequence

Due to the basic principle of applying the WED algorithm, the flash clustering cannot be shown as a similar simple example as in the case of group clustering, where only temporal and spatial adjacency rules were applied.

Two cases of example clusterings are shown, which are this time based on real LIS data (groups) to be clustered according to the WED algorithm principle (with the LIS/GLM version). For the example clustering shown, the LIS data from 4 Jan 2007 have been used, i.e. the LIS science data file `TRMM_LIS_SC.04.1_2007`.

The first case is an example of a “normal” clustering approach, while the second case shown demonstrates the effect of applying the “full fit” (GLM/MTG LI approach) instead of the “first fit” (LIS approach), as described above.

All of the examples shown use group centroids instead of event footprints for the demonstration of the WED principle, due to editorial reasons.

**Case 1: Typical flash clustering example**

The following file for the first example case study is used:

- 1st orbit on 4 Jan 2007: `TRMM_LIS_SC.04.1_2007.001.52013`

This orbit file only has 12 LIS flashes, which makes it easier to manage.

We concentrate in an area where two flashes are produced by the MTG LI algorithm from the related groups. The plots of the orbit snapshot are shown in Figure 17 (flashes) and Figure 18 (groups).
Figure 17. Snapshot to area of interest with two flashes.

Figure 18. Snapshot to area of interest with 17 groups. The numbers refer to the group numbering used in Table 6.
As mentioned, we note that there are two flashes for the groups plotted in Figure 18 as processed by the MTG LI algorithm. The basic data for the groups within this area of interest are shown in Table 6. Looking at the time difference between the groups, which are listed in a chronological order, we note that group #16 strikes out as it is 279 ms apart from the previous group.

According to the Weighted Euclidean Distance (WED) algorithm, a group has to be spatially and temporally near one other group in a flash to be part of a multigroup flash, as defined by Eq 1.

Since the WED is above 1.0 already if T > 330 ms, we can conclude the group #16 can only be a part of the original flash started by group #1, if it is spatially very close to any of the groups #9-#15, where the maximum delay will be 326.1 ms (to group #9). To investigate this, we compute the WED to each of the candidate groups #9-#15, as shown in Table 7.

According to the calculation of WEDs, none of the groups have a WED less than 1.0, meaning that a new flash would need to be generated with group #16 as its first element (as is done by the MTG LI algorithm).

Table 6. Groups included in the snapshot area of interest of Case #1.

<table>
<thead>
<tr>
<th>Group #</th>
<th>Group time</th>
<th>Latitude (deg)</th>
<th>Longitude (deg)</th>
<th>Group time difference to previous group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>441769047.7740</td>
<td>-0.7785</td>
<td>33.5265</td>
<td>0.0000</td>
</tr>
<tr>
<td>2</td>
<td>441769047.8117</td>
<td>-0.7585</td>
<td>33.5382</td>
<td>37.6892</td>
</tr>
<tr>
<td>3</td>
<td>441769047.8167</td>
<td>-0.7760</td>
<td>33.5291</td>
<td>5.0201</td>
</tr>
<tr>
<td>4</td>
<td>441769047.8243</td>
<td>-0.7673</td>
<td>33.5420</td>
<td>7.5531</td>
</tr>
<tr>
<td>5</td>
<td>441769047.8871</td>
<td>-0.7385</td>
<td>33.5090</td>
<td>62.8052</td>
</tr>
<tr>
<td>6</td>
<td>441769047.8941</td>
<td>-0.7385</td>
<td>33.5090</td>
<td>7.0343</td>
</tr>
<tr>
<td>7</td>
<td>441769047.8991</td>
<td>-0.7735</td>
<td>33.5340</td>
<td>5.0201</td>
</tr>
<tr>
<td>8</td>
<td>441769047.9172</td>
<td>-0.7474</td>
<td>33.4928</td>
<td>18.0969</td>
</tr>
<tr>
<td>9</td>
<td>441769047.9243</td>
<td>-0.7552</td>
<td>33.5013</td>
<td>7.0343</td>
</tr>
<tr>
<td>10</td>
<td>441769047.9444</td>
<td>-0.7647</td>
<td>33.4678</td>
<td>20.0958</td>
</tr>
<tr>
<td>11</td>
<td>441769047.9494</td>
<td>-0.7381</td>
<td>33.4752</td>
<td>5.0354</td>
</tr>
<tr>
<td>12</td>
<td>441769047.9529</td>
<td>-0.7635</td>
<td>33.4690</td>
<td>3.5095</td>
</tr>
<tr>
<td>13</td>
<td>441769047.9549</td>
<td>-0.7415</td>
<td>33.4890</td>
<td>2.0142</td>
</tr>
<tr>
<td>14</td>
<td>441769047.9564</td>
<td>-0.7459</td>
<td>33.4811</td>
<td>1.5106</td>
</tr>
<tr>
<td>15</td>
<td>441769047.9710</td>
<td>-0.6985</td>
<td>33.4865</td>
<td>14.5721</td>
</tr>
<tr>
<td>16</td>
<td>441769048.2504</td>
<td>-0.7679</td>
<td>33.5211</td>
<td>279.3884</td>
</tr>
<tr>
<td>17</td>
<td>441769048.2519</td>
<td>-0.7290</td>
<td>33.5163</td>
<td>1.5106</td>
</tr>
</tbody>
</table>

6 This is due to the basic principle of the WED algorithm. If groups are located “on top” of each other, the maximum allowed time difference is 330 ms which means that WED is below 1. However, if the groups are moving further apart, the allowed time difference needs to decrease in order to keep the WED at below 1.
Table 7. WED between group #16 and each candidate groups #9 to #15.

<table>
<thead>
<tr>
<th>Group #</th>
<th>( \Delta X ) (km)</th>
<th>( \Delta Y ) (km)</th>
<th>( \Delta T ) (ms)</th>
<th>WED</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>2.1995</td>
<td>1.4083</td>
<td>326.1</td>
<td>1.0964</td>
</tr>
<tr>
<td>10</td>
<td>5.9327</td>
<td>0.3547</td>
<td>306.0</td>
<td>1.4240</td>
</tr>
<tr>
<td>11</td>
<td>5.1018</td>
<td>3.3002</td>
<td>301.0</td>
<td>1.4326</td>
</tr>
<tr>
<td>12</td>
<td>5.7913</td>
<td>0.4929</td>
<td>297.5</td>
<td>1.3890</td>
</tr>
<tr>
<td>13</td>
<td>3.5676</td>
<td>2.9181</td>
<td>295.5</td>
<td>1.2263</td>
</tr>
<tr>
<td>14</td>
<td>4.4517</td>
<td>2.4391</td>
<td>294.0</td>
<td>1.2827</td>
</tr>
<tr>
<td>15</td>
<td>3.8445</td>
<td>7.6803</td>
<td>279.4</td>
<td>1.7763</td>
</tr>
</tbody>
</table>

Case 2: “full fit” (GLM/MTG LI approach) versus “first fit” (LIS approach) example

As an example demonstrating the “full fit” vs “first fit”, we select the following file for the second case study:

- 11th orbit on 4 Jan 2007: TRMM_LIS_SC.04.1_2007.001.52015

This orbit file only has 5 LIS flashes, which makes it also easy to manage.

We concentrate in an area where the “first fit” LIS algorithm produces two flashes while the “full fit” MTG LI algorithm produces only one. The plots of the orbit snapshot are shown in Figure 19 (“first fit” LIS flashes), Figure 20 (“full fit” MTG LI flashes) and Figure 21 (groups).
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Figure 19. Snapshot to area of interest with two flashes with the “first fit” LIS algorithm.

Figure 20. Snapshot to area of interest with one flash with the “full fit” MTG LI algorithm. Note that the location of the flash is determined by a radiance weighted average of the group locations, not a geometric average.
As mentioned, we note that there is only one flash for the groups plotted in Figure 20, as processed by the MTG LI “full fit” algorithm, while the LIS “first fit” algorithm has produced two flashes. The basic data for the groups within this area of interest are shown in Table 8.

When comparing the “full fit” and “first fit” approaches, the main questions are:

- Is group #20 spatially and temporally so separated from the groups before that it should form a flash of its own?

- Are any of the other groups in the second LIS “first fit” flash (groups #22-#24) spatially and temporally close to flash #1 (with a WED < 1.0 to any of the groups belonging to flash #1) so that flash #2 could be combined through such a group pair to flash #1 and hence remove flash #2 (i.e. follow the “full fit” principle)?

First we test the 1st bullet point, and compute the WED of group #20 to groups #1-#19. Due to the spatial and temporal distance to the main “group of groups”, the WEDs computed (not shown here) range from 2.38 to 3.17. **It is therefore indeed appropriate to designate a new flash for group #20.**

In order to see if the two flashes can be recombined, one of the groups that would be assigned to the new flash (groups #22-#24) would need to have a WED < 1.0 to any of the groups belonging to flash #1. Let’s assume that group #24 is our most potential suspect, as it is located closest to the main group cluster forming flash #1. We compute the WED for that group to all preceding groups. These results are shown in Table 9.
From the results we see that groups #1 to #10 are in any case not possible candidates, since the $\Delta T$ is $> 330$ ms which alone lead to WED values $> 1.0$.

However, there are two groups which have a WED $< 1.0$, i.e. group #21 and #23. Of these groups, group #21 has already been assigned to flash #1, therefore we can combine the flashes and assign only one flash for all groups. This is what is done by the “full fit” MTG LI algorithm.

Looking at the distribution of groups, this makes also logical sense, since the remaining groups #25-#27 are located clearly within the “main group cluster” forming flash #1.

It should be noted that the examples provided are special cases intended to demonstrate the LIS vs. GLM/LI algorithm features. According to Mach et al (2007) the $T$ threshold is negligible for flash clustering after $\sim 300$ ms. For values $< 300$ ms the threshold flash/group counts are increasingly affected (with a flash rate dependency).

Table 8. Groups included in the snapshot area of interest of case #2. There are two “first fit” LIS algorithm flashes, where the groups for flash #2 are shaded.

<table>
<thead>
<tr>
<th>Group #</th>
<th>Group time (UTC)</th>
<th>Latitude (deg)</th>
<th>Longitude (deg)</th>
<th>Group parent address (flash)</th>
<th>Group time difference to previous group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>441824316.9069</td>
<td>-8.3887</td>
<td>145.492</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>441824316.9155</td>
<td>-8.3947</td>
<td>145.492</td>
<td>1</td>
<td>8.59999657</td>
</tr>
<tr>
<td>3</td>
<td>441824316.9210</td>
<td>-8.3994</td>
<td>145.494</td>
<td>1</td>
<td>5.5000186</td>
</tr>
<tr>
<td>4</td>
<td>441824316.9301</td>
<td>-8.3875</td>
<td>145.504</td>
<td>1</td>
<td>9.10001993</td>
</tr>
<tr>
<td>5</td>
<td>441824316.9336</td>
<td>-8.4083</td>
<td>145.502</td>
<td>1</td>
<td>3.49998474</td>
</tr>
<tr>
<td>6</td>
<td>441824316.9391</td>
<td>-8.3988</td>
<td>145.508</td>
<td>1</td>
<td>5.5000186</td>
</tr>
<tr>
<td>7</td>
<td>441824316.9406</td>
<td>-8.4071</td>
<td>145.503</td>
<td>1</td>
<td>1.49995089</td>
</tr>
<tr>
<td>8</td>
<td>441824316.9487</td>
<td>-8.3986</td>
<td>145.519</td>
<td>1</td>
<td>9.10001993</td>
</tr>
<tr>
<td>9</td>
<td>441824316.9517</td>
<td>-8.4017</td>
<td>145.509</td>
<td>1</td>
<td>1.99997425</td>
</tr>
<tr>
<td>10</td>
<td>441824316.9642</td>
<td>-8.4183</td>
<td>145.513</td>
<td>1</td>
<td>12.5000477</td>
</tr>
<tr>
<td>11</td>
<td>441824317.0698</td>
<td>-8.3857</td>
<td>145.489</td>
<td>1</td>
<td>105.599999</td>
</tr>
<tr>
<td>12</td>
<td>441824317.0713</td>
<td>-8.3683</td>
<td>145.485</td>
<td>1</td>
<td>1.50001049</td>
</tr>
<tr>
<td>13</td>
<td>441824317.0878</td>
<td>-8.3671</td>
<td>145.486</td>
<td>1</td>
<td>16.4999962</td>
</tr>
<tr>
<td>14</td>
<td>441824317.2436</td>
<td>-8.3625</td>
<td>145.517</td>
<td>1</td>
<td>155.799985</td>
</tr>
<tr>
<td>15</td>
<td>441824317.2487</td>
<td>-8.4177</td>
<td>145.516</td>
<td>1</td>
<td>5.10001183</td>
</tr>
<tr>
<td>16</td>
<td>441824317.2522</td>
<td>-8.4071</td>
<td>145.513</td>
<td>1</td>
<td>3.49998474</td>
</tr>
<tr>
<td>17</td>
<td>441824317.2542</td>
<td>-8.4093</td>
<td>145.514</td>
<td>1</td>
<td>1.99997425</td>
</tr>
<tr>
<td>18</td>
<td>441824317.3381</td>
<td>-8.4078</td>
<td>145.513</td>
<td>1</td>
<td>83.9000344</td>
</tr>
<tr>
<td>19</td>
<td>441824317.3577</td>
<td>-8.3996</td>
<td>145.512</td>
<td>1</td>
<td>19.5999742</td>
</tr>
</tbody>
</table>
21 441824317.3687 -8.3921 145.527 1 2.00003386
22 441824317.3687 -8.3653 145.653 2 0
23 441824317.3708 -8.3777 145.564 2 1.50001049
24 441824317.3743 -8.4044 145.505 1 1.99997425
25 441824317.4260 -8.4006 145.514 1 51.699996
26 441824317.4959 -8.4021 145.514 1 69.8999763

Table 9. \( WED \) between group \#24 and all preceding groups (\#1-\#23).

<table>
<thead>
<tr>
<th>Group #</th>
<th>( \Delta X ) (km)</th>
<th>( \Delta Y ) (km)</th>
<th>( \Delta T ) (ms)</th>
<th>WED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.9641</td>
<td>1.2215</td>
<td>465.3</td>
<td>2.0333</td>
</tr>
<tr>
<td>2</td>
<td>8.0161</td>
<td>1.8801</td>
<td>456.8</td>
<td>2.0389</td>
</tr>
<tr>
<td>3</td>
<td>7.7824</td>
<td>2.4051</td>
<td>451.2</td>
<td>2.0157</td>
</tr>
<tr>
<td>4</td>
<td>6.6079</td>
<td>1.0910</td>
<td>442.2</td>
<td>1.8107</td>
</tr>
<tr>
<td>5</td>
<td>6.8529</td>
<td>3.3916</td>
<td>438.7</td>
<td>1.9235</td>
</tr>
<tr>
<td>6</td>
<td>6.1725</td>
<td>2.3320</td>
<td>433.2</td>
<td>1.7782</td>
</tr>
<tr>
<td>7</td>
<td>6.7185</td>
<td>3.2534</td>
<td>431.7</td>
<td>1.8850</td>
</tr>
<tr>
<td>8</td>
<td>4.9995</td>
<td>2.3177</td>
<td>422.6</td>
<td>1.6260</td>
</tr>
<tr>
<td>9</td>
<td>6.0514</td>
<td>2.6600</td>
<td>420.6</td>
<td>1.7518</td>
</tr>
<tr>
<td>10</td>
<td>5.7101</td>
<td>4.4918</td>
<td>408.0</td>
<td>1.8093</td>
</tr>
<tr>
<td>11</td>
<td>8.2986</td>
<td>0.8922</td>
<td>302.5</td>
<td>1.7729</td>
</tr>
<tr>
<td>12</td>
<td>8.7763</td>
<td>1.0313</td>
<td>301.0</td>
<td>1.8475</td>
</tr>
<tr>
<td>13</td>
<td>8.6402</td>
<td>1.1698</td>
<td>284.4</td>
<td>1.8044</td>
</tr>
<tr>
<td>14</td>
<td>5.2605</td>
<td>1.6807</td>
<td>128.6</td>
<td>1.0771</td>
</tr>
<tr>
<td>15</td>
<td>5.3572</td>
<td>4.4308</td>
<td>123.6</td>
<td>1.3183</td>
</tr>
<tr>
<td>16</td>
<td>5.6077</td>
<td>3.2530</td>
<td>120.1</td>
<td>1.2336</td>
</tr>
<tr>
<td>17</td>
<td>5.5253</td>
<td>3.4973</td>
<td>118.1</td>
<td>1.2416</td>
</tr>
<tr>
<td>18</td>
<td>5.6296</td>
<td>3.3339</td>
<td>34.2</td>
<td>1.1941</td>
</tr>
<tr>
<td>19</td>
<td>5.7221</td>
<td>2.4269</td>
<td>14.6</td>
<td>1.1309</td>
</tr>
<tr>
<td>20</td>
<td>7.5963</td>
<td>0.2035</td>
<td>5.5</td>
<td>1.3817</td>
</tr>
<tr>
<td>21</td>
<td>4.1022</td>
<td>1.5950</td>
<td>3.5</td>
<td><strong>0.8003</strong></td>
</tr>
<tr>
<td>22</td>
<td>9.7545</td>
<td>1.3698</td>
<td>3.5</td>
<td>1.7910</td>
</tr>
<tr>
<td>23</td>
<td>4.8233</td>
<td>0.3059</td>
<td>1.5</td>
<td><strong>0.8787</strong></td>
</tr>
</tbody>
</table>

6.2.4 Parallax correction

Parallax causes bias in projection of lightning position to ground coordinates due to the elevated lightning optical source at an unknown cloud top or edge height. This bias increases with angular distances from nadir. For instance, a point at 0° longitude, 46° latitude and 12 km height above ground appears to be at 46.14° latitude, which corresponds to a bias of 15.5 km [Finke and Hauf, 2002].
There are some alternatives for parallax correction, which need to be investigated during development. For example, in order of complexity:

1. **(baseline) Assumption of standard cloud top height.** In this approach, a standard height is assumed for a cloud top, for which a geometric correction is applied. It should also be noted that the cloud top height is also dependent on latitude and season (i.e. tropopause height) and therefore a standard height assumption could also take such relationships into account, for example in the form of a look-up table.

2. Knowledge of the cloud top height provided by the Flexible Combined Imager (FCI) on board the MTG-I platform.

### 6.2.4.1 Parallax correction approach for MTG LI products

For consistency, and allowing maximum flexibility to users, a parallax correction method is harmonised with the approach selected for the relevant FCI L2 products. This includes computing every FCI cycle (10 minutes) a table of delta-\( \varphi \) and delta-\( \lambda \), which can then be applied to the given Lat/Lon values of products. However, implementing the parallax correction is left for the user. This combines the flexibility of users being able to decide on the parallax correction for a given application, while at the same time providing an easy-to-use method for the correction, if needed. The delta-\( \varphi \)/delta-\( \lambda \) table will be in the FCI grid, updated every 10 minutes, and based on either CTH or OCA cloud top heights. Thus, the user can search the closest (in location and time) delta-value pair and perform the correction for each LI product element.
7 ALGORITHM DESCRIPTION FOR LIGHTNING ACCUMULATED PRODUCT GENERATION

7.1 Algorithm overview

For the accumulated products, a grid with a spacing of 2 km is defined (equal to the FCI IR grid). In order to create the accumulated products, the L2 Flash data is buffered for the accumulated products’ data collection period. A period of 30 seconds is defined, which is short enough to satisfy those users’ needs who require a very rapid access to new lightning data, while those who want to collect data for a longer time can simply stack the 30-second product.

At the end of the buffer data collection, the accumulated products are computed such that the L2 Events define the extent of the lightning activity in the accumulated product grid, while the L2 Flashes define the values in the product grid.

Flow charts describing the high level overview of the accumulated product algorithms are shown in Figure 24 to Figure 25, for the accumulated flashes, accumulated flash index, and accumulated flash radiance, respectively.

7.2 Resampling of event/flash data into the accumulated product grid

The rectified FCI grid for IR channels (2 km SSP) has been selected as the accumulated product grid in order to enable an easy comparison of the accumulated products with MTG FCI data. Such synergetic use of IR imagery and LI product data is foreseen in the users’ applications.

How the resampling is to be done is illustrated in Figure 22. In essence, data is to be distributed to the new grid with a simple formula:

- If \( N\% \) of the target IR grid element (accumulated product grid) is filled by an LI grid event, then the target grid element is getting a value. This means that the target grid element event/flash counter is added by one (\( \text{counter} = \text{counter} + 1 \)).

- In the case of accumulated radiance, the radiance value in the target grid elements is to be added with the new radiance value such that the radiance from the LI event grid is equally distributed between the target IR grids to be filled. For example, if an LI grid radiance of value 100 is to be resampled into the target accumulated product grid, and 5 grid elements are to be filled, each element gets a value of 20.

\( N\% \) here refers to the fraction of the FCI grid that is covered by the LI grid data, as illustrated in Figure 22.
The $N\%$ which determined which target grid elements are to be filled shall be configurable (default value: 50%).

![Diagram showing LI grid data to be resampled into the accumulated product grid (equal to the FCI IR grid). Only one LI grid element is shown positioned on a sample of IR grid elements. Those target grid pixels which are “selected” based on the LI grid element coverage (>50% requirement) are shaded.]

### 7.3 Significant algorithm details

The basic principle of a L2 accumulated flash index product computation is described as a form of a simplified example in Figure 26 to Figure 28. The example shows three time steps of the accumulated product generation, i.e. when a 30 second product is generated, the three flashes in the example figures occur all during the 30 seconds product buffer. The example highlights how the extent of the accumulated product in the product grid is determined by the events, but the values are based on the flashes corresponding to the events. In case of the accumulated flash product, the example would be modified such that each of the grid box values increases by $1/N_i$ for each flash comprised of $N_i$ events in separate grids.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Accumulated flashes product

L2 Flash buffer of 30sec received for flash density processing

Flashes in buffer?

no

yes

Adding cumulative events in the accumulated product grid

Recording how many flashes are linked to the accumulated events in each accumulated product grid box

Dividing the index value (for each flash separately) by the number of grid elements attached to each flash.

Storage

Dissemination to users

Application display

Figure 23. Flow chart of MTG LI L2 accumulated flashes algorithm overview.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Figure 24. Flow chart of MTG LI L2 accumulated flash index algorithm overview.
Figure 25. Flow chart of MTG LI L2 accumulated flash radiance algorithm overview.
Figure 26. Example of the accumulated flash index principle. Left, flash #1 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid.

Figure 27. Example of the accumulated flash index principle. Left, flashes #1 and #2 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid.
Figure 28. Example of the accumulated flash index principle. Left, flashes #1, #2, and #3 of 3 during the 30 second buffer. Right, the cumulative event and flash counts in the buffer. Data on the right are still in the LI grid and have not been resampled to the IR grid.

7.4 Example of the accumulated product usage

A real-world example of the use of an accumulated product was created from TRMM LIS data. Due to being a LEO mission, each sub-satellite point is observed with LIS only for a short time and therefore only “snapshot” images can be created and trend analysis is not possible. Data from 20 Nov 2010 over South America were chosen for creating the example.

The examples shown in Figure 29 to Figure 31 have been processed following the “Accumulated flash index” algorithm, i.e. one cannot integrate over pixels for cumulative flash counts, like with the “Accumulated flashes” algorithm. There are three image pairs, at $T_0 = 30, 60$ and 90 seconds. The top image in the pairs shows simply the cumulative events in the 30 second buffer, and the bottom image goes one step further from the accumulated flash index product, by creating a grid-based flash density value by assuming a standard size for the grids. Therefore, the example demonstrated a possible application of the accumulated flash index product.

It must be noted that for this example, a resampling of the LIS data into an IR grid has not taken place as the idea has just been to demonstrate the product principle.

The examples highlight the usefulness of the accumulated products in general. Although the cumulative events also show the extent of storms or electrically active regions, they
might give misleading indications on where the most important areas to follow might be. The accumulated flash-based product (converted to densities in the examples), which uses events to define the extent and flashes to indicate strength, gives a more reliable indication on active areas. This is evident e.g. in Figure 29, where cumulative events would indicate strong activity in three different cells. However, in the derived density product the cells on the left are still very pronounced whereas the cell on the right is clearly attenuated when flash information has been taken into account.

(a) Cumulative events at $T_0 = 30$ seconds.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

(b) Accumulated flash index based density product at $T_0 = 30$ seconds.

Figure 29. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is $T_0 = 30$ seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil).

(a) Cumulative events at $T_0 = 60$ seconds.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Figure 30. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is $T_0 = 60$ seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil).

(a) Cumulative events at $T_0 = 90$ seconds.

(b) Accumulated flash index based density product at $T_0 = 60$ seconds.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

(b) Accumulated flash index based density product at $T_0 = 90$ seconds.

Figure 31. Example of accumulated product usage, showing first (a) the cumulative events, and then (b) the accumulated flash index based density product. The time step is $T_0 = 90$ seconds. The example is based on TRMM LIS Flash/Event data from 20 Nov 2010 (Brazil).
8  ALGORITHM CALIBRATION AND VALIDATION

As a result of a EUMETSAT funded study, LERMA has provided a report addressing a verification and validation strategy for the MTG LI [Defer, 2010]. The main outcome of the study final report is summarised here, but for a full description the reader is referred to the full report available from the EUMETSAT external web-pages.

The MTG LI will require calibration, verification and validation in various instances. These tasks are divided into several sub-categories, as follows:

- **Pre-launch:**
  - Instrument calibration of instrument output (not in the scope of this document),

- **Post-launch:**
  - Instrument calibration (depends on design, currently TBD),
  - L1b product validation,
  - L2 product validation,
  - L2 product quality control.

Of the items listed above, pre- and post-launch calibration is not in the scope of the ATBD, as well as L1b product validation. The study referred to [Defer, 2010], however, briefly touches these issues as well. It is, however, considered that knowledge of the preceding steps is necessary for a successful validation of L2 products.

8.1  Validation

**L1b validation** should assess:

- detection efficiency
- false alarm rate
- location accuracy
- radiance of the optical signal

**L2 validation** should assess:

- detection efficiency
- false alarm rate
• location accuracy
• performance of the algorithms used to associate optical pulses in flashes.

It should be noted that the system requirements state the DE in relation to events, i.e. L1b, whereas there are separate false alarm rate requirements (FAR for events/L1b; FFR for flashes/L2).

The L2 validation should determine the performances of the algorithms used to associate optical pulses in flashes.

The L2 validation activities should be performed based on simultaneous concurrent observations (of the same lightning activity) of MTG LI, ground-, air- and space-based lightning locating systems (GLLSs, ALLSs and SLLSs respectively). The validation scheme should encompass all type of lightning sensors, i.e. operational and scientific where instrumental specifications and clear description of recording methodologies, data processing and applied algorithms are available. Two types of validation campaigns should be envisaged: a validation campaign of comparisons performed with operational lightning detection networks (mainly ground-based); several field campaigns performed during scientific field experiments where research lightning detection networks are deployed either under direct funding from EUMETSAT or during multi-disciplinary scientific projects. Comparison with space borne lightning observations should also be envisaged with low orbit sensors, such as the optical sensor deployed on the Tool for the Analysis of RAdiation from lightNIng and Sprites (TARANIS) mission.

Observations from VHF receiver based LLSs should also be used as validation measurements.

Inter-comparison of MTG LI and GOES GLM should be performed. A cross-instrumentation validation with GOES GLM sensor is highly recommended in order to provide a consistent dataset to the operational users and to the scientific community for the region covered by these geostationary missions. A coordinated field experiment in the Northeastern part of Brazil, region covered by both MTG LI and GOES GLM, is also highly recommended in order to effectively tackle this objective.

The validation should be performed at different levels: MTG LI frame level (L1b), flash component level, flash level, storm level and regional level, i.e. levels that span over the entire range of applications of MTG LI observations. Consequently, according to the type of level, different ground-, air- and space-based instruments are required.

The results of MTG-LI validation should provide some inputs on the synergy of MTG LI and all-type-LLS observations, with and without additional multiple-instrument cloud characterization of the parent thunderstorms, for both research investigations and operational applications. One should take the opportunity of the validation campaigns to expand the investigations to scientific questions. For instance the preliminary results on the synergy of the different lightning and cloud instruments should provide some material for the
characterization of the convection (within MTG LI coverage area) and to develop some preliminary nowcasting tools and assimilation schemes in numerical weather prediction models. The observations recorded during MTG LI validation will also be of high interest for lightning physics.

8.2 Quality control

The validation will also require the monitoring of the behaviour of the instrument and its platform as well as of the aging of the instrument during the entire MTG LI lifetime. Specific parameters will have to be defined in order to provide the users some flags (e.g. status of the instrument, status of the platform, pointing accuracy) on the quality of the observations. It is suggested that MTG LI validation should be performed on a regular basis following a series of standardized procedures that EUMETSAT and LIST will have to define. In addition, standard quality control measures for L1b and L2 products such as existing CAF and SAF products are expected to be in place.

Regarding quality indicators for the algorithms themselves, there are a number of flags specified in the L2 format specification. These are giving a measure of the reliability of each Flash/Group, as well as a false event filtering status. However, since the observed parameter retrieved by the L2 algorithm is not measurement such as temperature or pressure, a direct error estimate is not provided.
9 PRODUCT PROCESSOR DEVELOPMENT AND REALISATION

9.1 Prototype processor

A prototype L2 processor\(^8\) has been developed in the EUMETSAT TCE environment using the Matlab language. The main objectives of the prototype processor are:

- Implementation, testing and verification of the L2 algorithms. This includes the ability to test parameterisation and the L2 algorithm response to various input files (proxy data events) and the processing parameters. This includes testing also the robustness of the algorithms to varying data streams.

- Verification and validation of the various proxy data used for L2 (and L0-L1) processor development. This includes for example monitoring of data statistics (events, groups, flashes).

- Although computational aspects can be investigated, the findings will only be indicative to the expected performance of the operational processor. The performance would only be relevant on the machine where tests are performed, and only in the language used (MATLAB). This would be much slower than an implementation in a system-level language like C. The prototype processor is intended strictly for algorithm testing and development.

The current prototype processor realisation is in Matlab. A screenshot of the GUI is shown in Figure 32.

The prototype processor SW is constructed with a modular approach where there are separate functions and routines for the following main elements:

- Input/output data management

- Group clustering
  - Group clustering parameter management

- Flash clustering
  - Flash clustering parameter management

- Data visualisation

---

\(^8\) The prototype processor is only intended for testing of L2 algorithm functionalities with various types of input data (proxy data). It is not intended to do testing of end-to-end processes, timeliness, or other similar issues. The input data for the prototype processor is also not a real representation of L1b data (in format).
• Output statistics
• Graphical User Interface (GUI)

Figure 32. Screenshot of the Matlab-based prototype L2 processor.

9.2 Operational processor

(Placeholder, expected availability of a dedicated description: end of 2015. In addition, a description on a reference processor will be added, and consequently the text on the prototype processor removed)
10 PROXY DATA SETS AND TOOLS

For instrument studies and the development of data processing algorithms it is necessary to generate data as proxy for the future instrument output from detected optical pulses. Thus, proxy data represent a stream of detected pixel events.

Data can be generated from existing and suitable data from other sources, such as LIS/OTD data or various ground-based networks. Independent of the data source, data has to be modified to take into account the characteristics of geostationary observations. Each of the input data sources come with their own characteristics and provide a challenge for their use as MTG LI proxy data. Another way for creating proxy data would be to simulate the end-to-end chain from lightning optical pulses to MTG LI events. This would create an “artificial”, or simulated, proxy data source. Steps in artificial proxy data creation are the (i) simulation of the generation of lightning pulses, (ii) the propagation of the optical radiation from source to detector and (iii) the detection of these pulses.

Within the MTG LI development, two independent methods for proxy data generation have so far been created, which will be described in the following. The first is based on the simulated concept, while the second is based on the use of a ground-based LINET data set as input data for the proxy data creation.

10.1 Proxy data based on statistical and conceptual models (LIProxy tool)

A tool has been developed in a study funded by EUMETSAT (“Generation of Artificial Proxy Data for the MTG Lightning Imager”, EUM/CO/09/4600000660) which enables the generation of artificial lightning data on pulse, storm and global scales. The development work and the data are briefly described in the following, but for a full description one should refer to the study Final Report and the SW user manual [Finke, 2010a; Finke, 2010b]. The methodology and the resulting Matlab-based tool has been extensively evaluated and further developed in a later study, which expanded the usability of the GUI and added several new components into the simulation software [Sist and Biron, 2012].

10.1.1 Concepts

The artificial proxy data generation is based on the following assumptions and idealisations regarding the optical signal from lightning and the detection process by the future instrument.

10.1.1.1 Optical lightning pulses and photons

A lightning discharge radiates optical energy. This optical radiation is transported from the source location inside the cloud to the cloud surfaces by multiple elastic scattering. At the cloud surface the optical signal appears as a blurred optical pulse with certain spatial extend and temporal duration. From the surfaces the optical energy is assumed to be radiated isotropically in all directions, i.e. according to the Lambert-Beer law.
In order to simulate this radiation an optical pulse from lightning arriving at the instrument from the cloud surface is represented as a large number of photons. The spatial pattern of the optical pulse and the temporal pulse shape are simulated by corresponding distributions of these photons in space and time.

The basic entity in the proxy data generation is the pulse consisting of photons. This pulse represents the optical signal from lightning at the cloud surface. The photons of the pulse are generated by random number generators. This concept allows for the artificial generation of optical pulses with any desired statistical properties.

Pulses are characterized by position parameters and by internal parameters. Internal parameters such as energy, spatial and temporal size, pulse shape and pattern are generated according to the established empirical probability distribution functions (pdf) or conceptual models. The internal parameters control the photon generation for each pulse. The position in time and space is taken from observed data, a certain storm model, or any other desired distribution (see Figure 33). For large scale distribution the global lightning and storm statistics is used.

The proxy data should obey the same statistical distribution as natural lightning and produce detection data as close as possible to the future instruments output. To achieve this goal, the pulse data are generated by random number generators according to the empirical distributions of optical lightning pulses. These empirical distribution functions were derived from the long time observation of lightning with the optical lightning detectors on low Earth orbits LIS and FORTE [Finke, 2006].
Figure 33. Schematic picture of the proxy data structure. Photons compose optical pulses of certain size and energy. The pulses with spatial position and time are organized in storms.

10.1.1.2 Pulse propagation and mapping

The pulses are emitted at the cloud surface. A part of the emitted energy is gathered by the instruments optics. This propagation process is simulated by a mapping of the generated pulses onto the geostationary field of view. Additional optical signals can be overlaid such as cloud reflection radiation or various types of noise.

10.1.1.3 Detection process

The generated photons are then detected by a simulated detection process, which consists of the same principal steps as in the presumed instrument. Since the artificial data are pulses consisting of a large number of photons the detection is simulated by summation over photons in certain time and space intervals. These intervals can be set to the assumed instrument parameters. Additionally, a detection threshold is set as a minimal number of photons. Other detection effects such as saturation can be introduced as well.
The lightning pulse is a short transient signal which is overlaid during daytime by the bright background of reflected solar radiation from clouds. This slowly changing background has to be subtracted from the signal. In the simulated detection the background processing is implemented in the following way: A moving average of the number of photon count for each pixel is calculated and subtracted from the actual detected signal. If this signal exceeds a certain threshold it is registered as a detected event.

The output of the detection process is the detected pixel events, or simply, events. Generally, a pulse can produce several events. Each event is characterized by discrete time and pixel position and energy, i.e. the number of detected photons.

10.1.1.4 Additional concepts for lightning and storm distribution

If the positions of pulses are not taken from observation but generated randomly the following distribution characteristics are realized in order to produce a realistic lightning distribution. The proxy data show autocorrelation in time and space on the following scales:

- Multiples strokes belonging to the same flash, typically close in space and separated by few tens of milliseconds.
- Lightning pulses are organised in storm clouds which follow a storm propagation and development cycle.
- Storms show a geographical distribution with the typical diurnal and seasonal variations.

This structure is generated by random number generators which base on the empirical pdfs and the global lightning data climatology.

10.1.2 Realisation

The generation of the artificial proxy data sets is realised in the following processing steps:

1. Scenario setup
2. Data generation
3. Reception simulation
4. Detection processing

Each of these steps is done by one or several software modules, which exchange data and parameter structures. The strict modular concept allows to modify every single step, and also to introduce additional effects, such as noise in the data stream.
10.1.2.1 Scenario Setup

The pulses are generated according to the desired spatial and temporal distribution on the scales defined above. The data on all these scales can be selectively randomized or fixed to user input values, e.g. real lightning and storm data.

**Lightning pulse positions and times.** These are imported from lightning observation or generated by conceptual model of storm propagation and evolution. For test purposes also unrealistic lightning distributions can be generated.

The global distribution of storms is simulated according to the climatology which was derived from the OTD and LIS observations.

A multiplicity of the pulses is simulated by the generation of a random number of subsequent pulses with random separation times and random location close to the first pulse.

**Lightning pulse parameters.** For each lightning pulse the optical parameters are generated randomly. These include the total energy, spatial size and temporal pulse width.

10.1.2.2 Data Generation

In this step for each of the pulses the photons are generated by the random number generators. The generated stream of photons is transformed to the satellite field of view.

**Photon data.** Each pulse consists of a large number of photons, which are distributed in space according to the spatial distribution function. The positions are generated firstly as distances (in kilometres) around zero, and then shifted to the desired pulse centre location on Earth (longitude, latitude). The time of the photons is simulated randomly according to the temporal pulse function shape. Output is a stream of single photons with longitude, latitude and time.

**Mapping to the GEO field of view.** The positions given in Earth longitude-latitude are mapped onto the GEO field of view according to the anticipated satellite position (0° longitude, 42164 km orbital radius i.e. Earth average radius plus the geostationary orbit altitude) and the non-spherical Earth model as assumed in the EUMETSAT technical documents. In this step the number of photons is reduced by the cosine of the satellite angle for the given location on Earth. This simulates the Lambertian radiation assumption.

10.1.2.3 Reception Simulation

This step models the data flow and transformation from the sensor’s entry to the detection matrix.

**Gridding of the data in time and space.** The data stream is detected in certain time intervals of typically 0.5-2.0 ms. The time grid is given by start time, end time and time window step. The detection matrix is represented as an equidistant grid consisting of e.g. 1024x1024 pixels covering the angular field of view of 16°x16°.
For each of the time intervals the photons falling in each of the pixels of the detector matrix are counted. The result is the number of photons for each of the pixels per time window interval.

**Adding of Noise.** Optical signals reaching the detector from other sources than lightning are considered as noise. This can represent internal detector noise, shot noise, the solar radiation reflected from the clouds, solar glint from water surfaces, etc. Also high energetic particles can interact with the detection matrix and produce photon counts. All these types of noise can be selectively added to the photon data on the detection grid.

**Amplification.** The incoming signal from lightning and noise can saturate the pixel detectors. This saturation effect is simulated by clipping the data. Other characteristics of the amplifier can be introduced as well.

### 10.1.2.4 Detection Processing

This step models the process of detection of lightning events in the gridded data stream.

**Background subtraction.** Detecting the faint lightning events on the bright background requires a certain background processing. The real time processing implementation in the future instrument is not known. It is assumed here, that a moving average is calculated for each pixel, which is subtracted from the actual data in order to discriminate the lightning signal.

**Thresholding.** Only pixels with energy (photon numbers) above a certain threshold are reported as events. Thus a data reduction and selection takes place.

**Pixel events.** Each pixel with a photon count above the threshold represents a detected event. The event data contain pixel position, time interval and number of photons. Additionally, the corresponding geolocation (longitude, latitude) and a radiance value corresponding to the photon number is calculated for each event.

Output is a collection of the events detected in the given time grid.

### 10.1.2.5 Further Processing

The event data stream is the basis for any subsequent data processing, for instance the association of the events into groups, which represent the detected lightning pulses. The constructed groups can be compared against the pulse data, which were used for the generation of the photon cloud. Other processing steps on this level include the removal of non lightning events, thus reducing the false alarm rate.

### 10.1.3 Software

The software for the generation of the artificial proxy data (LIProxy) was written in the Matlab programming language. The LIProxy software package consists of three parts:
1. Function Library - containing all the functions for data generation, detection and display

2. Script collection - demonstrating the typical use of the function library

3. Graphical User Interface (GUI) - serving for an easy configuration of proxy data scenarios, data display and data import and export.

It is dedicated for use in simulation of the detection process of optical pulses from lightning and in the development of algorithms for the detected lightning events.

An example of the GUI appearance for a typical simulation scenario is shown in Figure 34. In the example shown, the initial storm locations where optical pulses are to be generated are based on ATDnet strokes which is one of the inputs that can be selected to provide realistic storm distributions across the field of view.

Figure 34. LIPROxy GUI example.
Figure 35 shows an example for a generated proxy data set for a model storm. At first the pulse positions and times, and pulse parameters are generated (top). The next step is the generation of photons and mapping to the field of view (middle). Finally, the photons are detected and the detected events are identified (bottom).

For a simulation on the whole geostationary field of view the global storm climatology is used in order to generate the storm distribution for a certain month of the year and time of the day. The proxy data generation starts with a random initial storm distribution and applies a storm model to each storm. The storm model generates the random position and time for the pulses of the storm. Next step is the generation of photons and the detection of events, the resulting event density for an example simulation is shown in Figure 36.
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

Figure 35. Pulses (top), photons (middle) and overlaid detected events (bottom) for a simulated storm. The number of displayed photons was reduced for sake of clarity of the figure.

Figure 36. Density of detected events for a hemispheric simulation of lightning distribution for a day in June between 13 UTC and 20 UTC
10.2 Proxy data based on LINET/LIS/LMA comparisons

As one of the preparatory activities for the MTG LI mission, a study was launched to compare optical lightning observations from LIS and ground-based LINET measurements in the VLF/LF frequency band [Höller and Betz, 2010]. The main objective of the study was to get information on the relation of VLF/LF and optical measurements. Depending on the findings, a secondary objective was to enable the creation of proxy data for the MTG LI development work, which could be used as another proxy data source on the side of the artificial proxy data discussed in the previous Section. The full report of the study is available in [Höller and Betz, 2010].

A second study aimed at refining the earlier developed proxy data method by enabling a much more extensive data comparison. This became possible with the CHUVA campaign, which was organised in Brazil Sao Paulo area in 2011-2012, in support of a multiyear campaign in preparation for the Brazilian contribution to the GPM mission. This campaign allowed for the first time a direct comparison of all major lightning detection techniques, including LIS, LINET, LMA, ATDnet, as well as Vaisala and Earth Networks systems. The EUMETSAT contribution to the main campaign was funding the LINET mobile unit network measurements during the Intensive Observation Period (IOP), and the following data analysis with the aim to improve the LINET based proxy data method [Höller, 2013].

10.2.1 Background

LINET is a time-of-arrival system originally developed at the atmospheric research group at the physics department of the University of Munich, enabling the identification of IC events, including determination of the source heights of the VLF/LF emissions. It is claimed that the system allows detection of both IC events and CG strokes with a comparable efficiency. The detection range for 3D and 2D observations is within 100 km and 300 km, respectively. At present, the LINET network consists of more than 100 sensors sites in 20 European countries and is owned and operated by Nowcast GmbH. [Betz et al., 2004; Betz et al., 2007].

LINET lightning measurements were involved during a series of field campaigns focusing on lightning NOx production. For the EU-project TROCCINOX (campaign in 2005), DLR has installed a 6-sensor version of LINET during the field experiment in Brazil [Höller et al., 2009; Schumann and Huntrieser, 2007]. After the TROCCINOX campaign in Brazil, the system was set up in Australia (SCOUT-O3, 2005; TWP-ICE, 2006) and tropical West-Africa (AMMA, 2006). A comprehensive overview of all DLR-LINET deployments during the different field campaigns in the tropics is given by Höller et al. [2009]. The data analysis performed up to now includes all experimental regions but focused on the data from Benin (W-Africa) as this is the area most relevant for the MTG field of view, and were thus selected as the main source of LINET data for the LIS/LINET comparisons. The network in Benin operated from June to November 2006.
10.2.2 Results of the first data comparison of LIS and LINET observations

From the analysis of the LIS and LINET simultaneous observations of the first study [Höller and Betz, 2010] it was found that the overall flash numbers were about the same from both systems. On the other hand, this does not necessarily mean that the same flashes were observed by both systems. Only about 50% of all LINET (LIS) flashes had time- and space-coincident LIS (LINET) counterparts.

It was found that in many cases the LINET stroke was directly followed in time by a LIS group, which was recorded within 2 ms after the stroke. This delay can be attributed partly to the scattering processes of light within the thunderstorm for another part to the limited LIS time resolution using a 2 ms time frame. The inter-stroke optical activity is of a much more irregular nature. Thus, for generating proxy data for the MTG LI from LINET data, a statistical representation has been considered to be adequate. The coincident stroke/group events did not show strong correlations between LINET stroke peak current and LIS group radiance, thus here also a statistical treatment has been followed for the MTG LI optical signal simulation. As LINET can discriminate cloud-to-ground (CG) and intra-cloud (IC) strokes and also can indicate IC stroke height, a sub-division of the data along these categories was performed. The statements made above for the overall sample did not show distinctive differences in these categories. This indicates that also CG strokes produce enough optical radiation at cloud top, possibly due to sufficiently large in-cloud channel extension.

10.2.3 CHUVA field campaign outcome and results of the second data comparison

The goal of the second study comparing LIS and LINET with the objective of developing a proxy data method was to cover the main rainy season in Brazil October 2011 – April 2012 by LINET measurements in order to obtain a comprehensive data set for comparison with LMA and LIS. It was also for the first time that a large number of research and operational lightning detection networks was deployed or taking measurements for a field campaign. Thus within CHUVA a unique lightning data set could be established and was completed by radar observations of precipitation, electric fields or optical (high speed video) information from the ground.

A network of seven DLR LINET sensors was installed in the Sao Paulo metropolitan area during September-October 2011 with support of USP and INPE as the local partners. Operations were continued until end of April 2012 thus covering about 7 month of observations during the complete rainy season in the Sao Paulo area. The NASA LMA network of 11 VHF sensors covered the same time period and area and thus a comprehensive data set was obtained for analyzing coincident lightning observations with the optical LIS sensor from space (5 high priority cases out of 22 in total).

The main conclusions from the data comparisons conducted during the study are as follows:

- As found in previous studies, LINET strokes and LIS groups are well correlating,
- LINET strokes map the flash branches similar to LMA (with considerably less source points),
• An initial breakdown phase of vertically propagating sources can be often found in LINET and LMA data,
• Higher level LINET and LMA signals have higher probability to be optically detected,
• Lower level LINET and LMA signals are optically detected from above in case of missing high level precipitation (e.g. from radar),
• XPOL radar helps in interpretation of 3D cloud structure important for scattering of light,
• Horizontal location of optical emissions may be different from RF sources (reasons to be investigated further).

10.2.4 Steps for generating proxy data from LINET observations based on the first data comparison

A statistical analysis of the relations between the LIS and LINET data sets led to model distribution functions suitable for generating random proxy optical data from a given set of observed LINET strokes. These can be used for simulating the output of the future MTG-LI instrument. The model primarily aims to reproduce the number of optical events which is likely to arise during typical thunderstorm situations based on the data obtained from W-Africa and Europe. The focus is on the total number of events rather than on other details of the lightning characteristics like exact group timing or location. These features are highly variable and show strong statistical fluctuations as induced by the complexity of the lightning process.

The method is based on stroke – group relations. Each stroke is simulated to be accompanied by one directly time-coincident optical group. Additional groups are introduced following the stroke event within about 100 ms. Radiance and footprint sizes are simulated randomly from the statistical distribution functions. For each group a circular shape is assumed. Using projection methods for the geo-referenced lightning events a pixel matrix of the future MTG-LI instrument is simulated and proxy event data are produced as output.

Motivated by the results discussed above, the main model assumptions are as follows:

1. The total number of optical flashes equals the number of RF detected flashes recorded by LINET. This result does not imply that all flashes had a counterpart within the flash data of the respective comparative system. This was only the case for about half of the flashes. Although in some cases these differences seemed to be attributable to the LINET range-dependent detection efficiency, there could also be other reasons responsible for these differences, e.g. channel orientations or the vertical distribution of lightning within the clouds. All these factors of influence are not directly controllable from the information provided by the present data set and are beyond the modelling aim of the present study.
2. Each LINET stroke is directly associated with a coincident optical group. This property was discovered to be a consistent feature from the data set. The occurrence time of the optical group is given by the time of the LINET stroke. The location of the centre of the optical group is determined from the latitude/longitude stroke coordinates. The correlation of the peak current and group radiance is, however, rather poor. The data distribution is approximated by a bivariate log-normal distribution with parameters derived from a fitting procedure using an expectation maximization (EM) algorithm leading to maximum likelihood estimates of the parameters in a Gaussian (mixture) model.

3. Additional optical groups are simulated around each stroke. As shown by the analysis, there is a lot of optical activity around the time of occurrence of a LINET stroke. The average flash properties were characterized by about 7 optical groups per LINET stroke.

Whereas the random generation of the number of additional optical groups is quite straightforward, the simulation of the time of occurrence is more difficult. From the time series data analysed, it is obvious that there are many possible configurations for optical signals to occur during the course of a flash relative to the occurrence of the RF stroke signal. As there is anyway not a very clear rule for optical pulses to occur within the flash duration, it might be a practicable approach to connect the simulated additional optical pulses directly to each stroke. As it is difficult to infer any statistics on such configurations, a simplified approach is followed for the modelling of the stroke/group relationship. It is assumed that the additional strokes may be simulated by normally distributed occurrence times with a mean time delay of 50 ms after the LINET stroke. A schematic representation of this model is shown in Figure 37.
The radiance values associated with each of these additional groups is derived from the overall statistics of LIS group radiances taken from the complete sample of group radiances for the coincident flashes (LINET flash coincident with LIS flash, sub-sample of total LIS statistics).

Finally, one must simulate the location and the size of the additional groups. The following assumptions are made:

- The location of the centroid of the direct stroke coincident group and the additionally generated groups are assumed to be identical to the stroke lat/lon location.

- Like in the case of the radiance simulation the overall footprint statistics taken from all the coincident flashes is described by a log-normal distribution function. The group footprints have to be related to the group radiances as these parameters are not independent of each other. For the overall statistics from all LIS/LINET coincident flashes, a similar approach as for the coincident stroke/group statistics is followed. The results are not too different from the smaller data set of the stroke-related observations. The simulated radiance value from the time-radiance simulation is used to define a footprint function that uses a special radiance value as input.

- The final footprint size is determined from that radiance/footprint pair which has the closest radiance value. This procedure is due to the assumption of a bivariate distribution function of the time/radiance dependence, such that a maximum
number of strokes associated with maximum radiance occur around 50 ms after the LINET stroke.

10.2.5 Further enhancement of the LINET proxy data method based on the CHUVA campaign outcome

From the CHUVA field campaign a sufficient number of comparable data was available for detailed analysis. Especially the LMA offers a lot of discharge location data which were compared to the data obtained from the other observing systems like LIS and LINET. The aim of the network comparison was to arrive at an optimized proxy data generation method based on the elaborated lightning model arising from the integrated view provided by all the different detection systems used within CHUVA.

With the CHUVA data there was the opportunity to look at a high-sensitivity LINET configuration and thus extend the range of sensitivity towards the low peak current range, as compared to the previous LINET tropical measurements. For the first time the very small amplitude range of LINET strokes was covered by LIS observations. This was a good step forward towards the European LINET network with its variable sensitivity inferred from the inhomogeneous sensor spacing in the different regions.

10.2.5.1 Smallest detectable signal by LINET

It was expected that the number of LIS groups per LINET stroke was dependent on the LINET detection efficiency (or the smallest detectable stroke), and this dependence was investigated for the CHUVA case studies of intense convection over the network. The enhanced proxy data generator is based on these results. The detection of strokes by LINET depends on the stroke’s peak current and its position with respect to the sensors. For a stroke to be analyzed by LINET its peak current has to exceed a certain detection limit at least at four stations. Thus the detection limit is determined by the stroke’s distance (signal decreases with distance) and viewing angle (x- and y- rings of the field antenna) from the 4th closest station. Figure 38 illustrates the LINET stroke detection and the smallest detectable signal.
Figure 38. Schematic of LINET stroke detection and smallest detectable signal.

As a more general description for the minimum detectable peak current also valid for Europe with its larger scale network, the minimum detectable peak current can be expressed in a form of a lat/lon array determined experimentally. The result of such a mapping of the data is shown in Figure 39.

$$Amp_{kA} = R_{km} \times \sqrt{1 + \left(\frac{1}{\tan \Phi}\right)^2} \times 50 \times 0.0003024$$
Figure 39. Array of the smallest detectable LINET stroke (absolute peak current) over the Sao Paulo area derived from data of 10 Feb and 27 March 2012. Array resolution is 0.1 degree in each direction.

10.2.5.2 LIS groups per LINET stroke (GPS)

Based on the results from comparing LIS-LINET data from coincident cases during CHUVA, as well as the earlier African campaign with longer LINET baselines, it was possible to establish a relationship on the number of LIS groups observed per LINET stroke (GPS) within the coincident flashes. This relationship is the taken into account in the proxy data generator for MTG-LI.

Figure 40 shows the result of the analysis for all available data from the different tropical LINET campaigns. For deriving the ratio of LIS groups per LINET stroke the total number of strokes is used, and divided by the total number of groups within a coincident flash. These flashes have been arrived at by clustering together the LIS and LINET defined flashes.
From the results of the analysis it is noted that the CHUVA data add additional information in the low peak current regime. A linear fit is applied to the data and shows a generally increasing GPS with increasing minimum peak current or distance from the closest sensor. This can be explained by the decreasing LINET detection efficiency for small amplitude strokes with increasing flash distance from the sensors.

Another important result is that, especially for the high sensitivity CHUVA data, the GPS ratio can be well below 1. That means that LINET sees more strokes than LIS is able to detect the corresponding groups. This result is directly relevant to the proxy data generation as it implies that some of the LINET strokes from the generation procedure need to be eliminated and not converted to optical pulses. It is clear that this also applies to the European LINET network which has lower detection limits of 1-2 kA in its best covered areas.

10.2.6 Application to Europe

The new results obtained from the CHUVA study were intended to be applied to the European LINET network and the generation of LI proxy data over Europe. For this, the minimum detectable stroke peak currents for Central Europe needed to be established. These characteristics depend on the network configuration (sensor sites) which is shown in Figure
41. The number of sensors has increases from year to year and the coverage is still improving. The minimum of all detected peak currents over central Europe is shown in Figure 42. In accordance with the sensor distribution, we note the high detection efficiency over Germany and Austria.

![Figure 41. The LINET configuration of sites in Europe in 2008.](image-url)
Figure 42. Distribution of the smallest detected peak current during 2010. (a) All strokes observed with smallest plotted on top, (b) Detection matrix with a 0.1 x 0.1 degree resolution
10.2.7 Examples

A typical example of the LINET based proxy data is shown in Figure 43, where a case of significant thunderstorm activity has been chosen taking place on 28 July 2013. A cumulative set of 5 minutes of data is shown, overlaid on a MSG SEVIRI 10.8 µm channel image taken at 18:42 UTC. The second sub-image is a close up of the main European image, showing the structure of the lightning pulses within the area. As there is no pulse density shown in the image, there can be several pulses on top of each other in the same grid. The SEVIRI 3712x3712 grid is used for this example.

(a) LINET based LI proxy data over Europe, 28 July 2006 (5 minutes of data)
Algorithm Theoretical Basis Document (ATBD) for L2 processing of the MTG Lightning Imager data

(b) Close-up of (a) showing the LI proxy data structure over a storm (5 minutes of data)

Figure 43. Application of LINET proxy data on 28 July 2006, showing 5 minutes of data overlaid on MSG SEVIRI 10.8 μm channel image. The full European area is shown in (a) and the white square indicates the area highlighted in (b).
11 POTENTIAL FOR SYNERGY WITH GROUND-BASED OBSERVATIONS

Of the various ground-based networks available, the use of the MetOffice operated ATDnet data in support of MTG LI data processing and operations has been studied the most due to the good coverage of the data and the direct availability of the data in the MTG Ground Segment. LIST initiated studies have also been conducted by the MetOffice in support of the ATDnet data usage.

Based on LIST discussions and direct negotiations with the Met Office, a first study [Bennett et al., 2010] was launched in 2009 to get a first look into the feasibility of using global or at least LI-coverage ground-based lightning location systems and MTG LI in providing added value as synergetic products. The study team was led by Alec Bennett from the MetOffice. The study summarised the main requirements for lightning data users based on Met Office surveys, which formed a useful baseline from which to assess the suitability of systems and products used to convey information of lightning both over Europe and throughout the MTG footprint.

A second study [Collins et al., 2012] was launched in 2011, and was building on the outcome and assumptions on the previous study. The main tasks were to verify the feasibility of the proposed synergetic activities by comparing ATDnet and LIS (space-based optical) datasets both in selected case studies as well as in a more “global” statistical analysis. In addition, the study was tasked to analyse the expected evolution of the ATDnet system in the timeframe of preparing for MTG. The second study concluded that in addition to different detection efficiencies (DE), the two systems are sensitive to different parts of the lightning process. Therefore, lower-level synergy would be difficult, whereas at storm level there would be a potential for synergy, where a direct matching of individual strokes/flashes would not be necessary.

Due to the inherent nature of the two data sources, i.e. being sensitive to different parts of the lightning process, a complementary synergetic product can be envisaged. However, this is out of scope of the current L1b/L2 planning.

Regarding use of ATDnet data in Cal/Val, this would cover the long-term monitoring and verification in the SQDAR (Science Data Quality Analysis and Reporting). The LI instrument will be subject to in-orbit degradation, which would need to be monitored. In the case of the LI, direct calibration (radiometric, spectral, and spatial) is not possible, and instead various vicarious calibration techniques shall be applied. This includes e.g. the use of the LI background imagery in comparison with FCI data over deep convective clouds, sun glint, and deserts. The aim is to support long-term instrument and L1b product monitoring, calibration parameter (database) updates, and assists in resolution of anomalies. The use of well-defined external sources on ground, i.e. ADTnet would be highly beneficial for such long-term tasks in the operational phase.

For such a long-term operational use, ATDnet still needs a further verification. ATDnet network expansion is still open (schedule, outcome). In addition, due to the nature of ground-based observations (e.g. diurnal effects in the troposphere, network station status etc) the
stability of the ATDnet is not constant. Therefore, the long-term stability of the ATDnet system would also need to be monitored and assessed against the MTG LI performance.

The of ATDnet for long-term monitoring or other Cal/Val related activities, as well as use of other ground-based systems are discussed in the MTG LI Calibration and Validation plan, which is to be available during Phase C/D.
12 REFERENCES


## ANNEX I: PRODUCT REQUIREMENTS FOR LIGHTNING DETECTION

<table>
<thead>
<tr>
<th>Product</th>
<th>Product geographic coverage</th>
<th>Vertical resolution</th>
<th>Horizontal resolution</th>
<th>Temporal coverage</th>
<th>Timeliness</th>
<th>Detection Efficiency (DE)</th>
<th>False Flash Rate (FFR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 product for lightning detection (events, groups, flashes)</td>
<td>Full disk Surface to cloud top (no separation between CG and CC lightning)</td>
<td>10 km</td>
<td>24h (continuous day and night)</td>
<td>Goal: 30 sec Threshold: 2 min</td>
<td>&gt;90% for flashes consisting of events with radiant energies &gt; 10 µJm²sr⁻¹ at 45°</td>
<td>&lt; 2.5 flashes/s</td>
<td></td>
</tr>
</tbody>
</table>