Investigation of LI False Alarms Filtering Algorithms Concepts and Status

Joint MTG LI Mission Advisory Group
& GOES-R GLM Science Team workshop
27-29 May, 2015, Roma, Italy

V. Crombez

Topics of the presentation

Context: LI, events and false alarms

Objectives of the study

Noises and false alarms sources

The filtering algorithm

- A general filtering scheme
- The Bayesian approach

Calibration

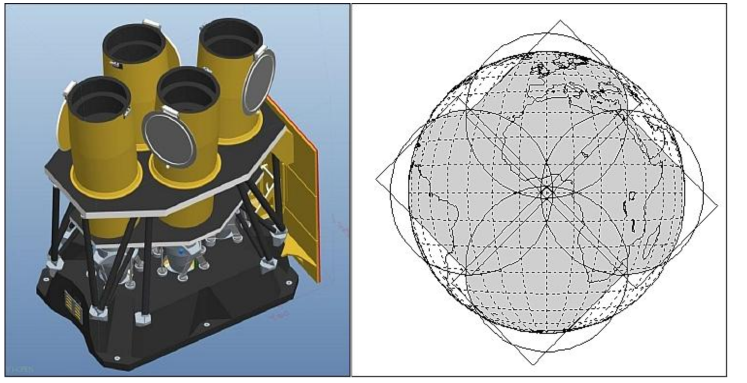
Performance estimation: the LIDEFAS simulator

Performances

Conclusion

Context: LI, events and false alarms

LI instrument (main features)



- Coverage: 80% Earth disk with 4 optical heads
- Sampling: 4.5Km at Sub-Satellite Point.
- Each optical head with protective cover, baffle, solar rejection filter
- Narrow band filter: 777.4 nm ± 0.17 nm
- Optical system with F/1.73 and D=110mm,
- CMOS detector size: 1000x1170pixels
- Read frequency: 1000 frames per second

Detection principle

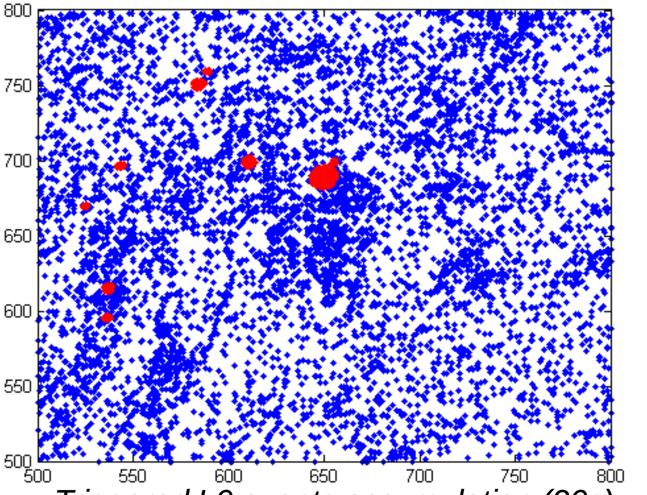
- Objective: to detect optical pulses of 0.6 ms duration, at any time and any location on the Earth, with energy down to 4µJ/m2/sr
- Background estimation per pixel: smoothing filter on few previous frames
- Substraction of background from current frame
- Result is thresholded => triggered events

On board filtering

- Events with 3x3 adjacent events are downloaded
- 7/8 of other events are rejected by a second threshold

L0 events: real events / false alarms

- Typical L0 FAR between 30 000 and 50 000 ev/s



Triggered L0 events accumulation (20s)
- *Blue: false alarms*
- *Red: real pulses*

Objectives of the study

To address the False Alarms filtering of LI data at Level 1/2 processing step

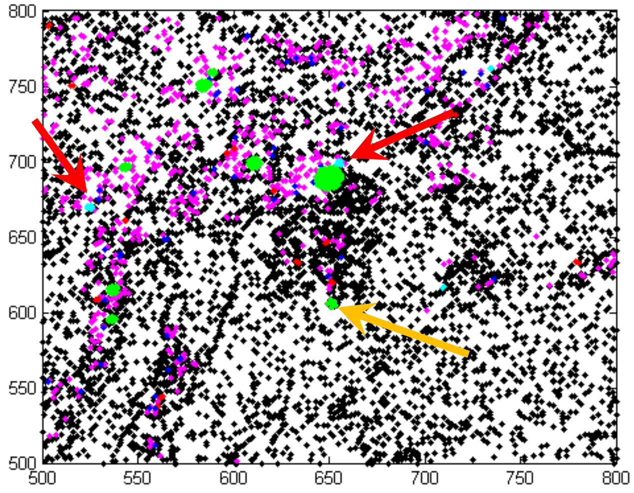
- We aim at reducing as much as possible the False Alarms (FA)
- while keeping Detection Efficiency (DE) at acceptable level

Study approach

- Analysis of the detection noises source
  - Evaluation of the noises impact through the whole chain,
  - From on-board detection and filtering up to level 2 output stream (flashes)
- Design and prototyping of the FA filtering algorithms
- Evaluation of the filtering efficiency in terms of performances

FAR and DE performances are always jointly analyzed

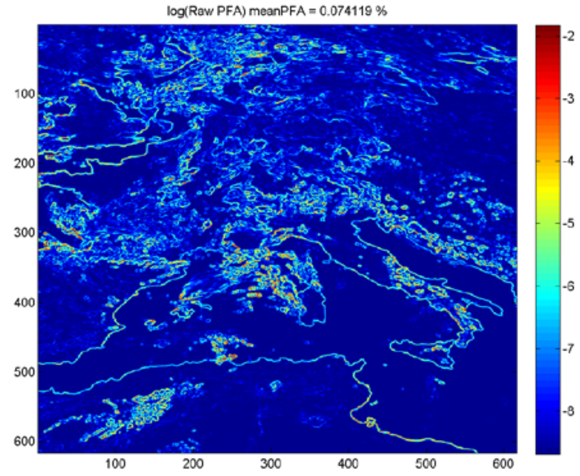
- At Level 0 events
  - False events rate, Pulses detection efficiency, downlink data rate
- At Level 2 output:
  - Events, groups and flashes



Classified L2 events accumulation (20s)

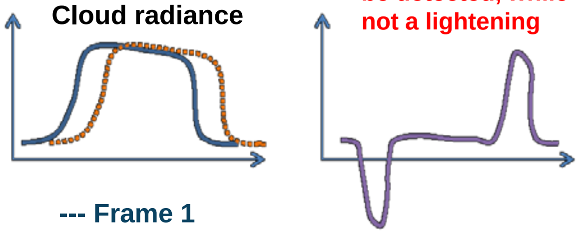
- *green: detected lightning pulses*
- *other colours: rejected events*
- *red arrows: non detections*
- *Yellow arrow: false alarms*

Noises and false alarms sources



Instrument noise

- CMOS detector pixel and Readout circuit designed to offer a 450ke- FWC.
- Readout circuit noise: 150e- 1-σ
- On chip 12bits ADC: 450ke-/4096 = 110e- sampling



Background noise

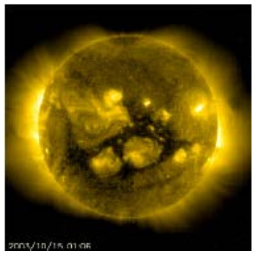
- The photonic noise up to 435 e- when considering very high earth background values (300W/m2/sr/µm)

Jitter noise

- Combination of 2 phenomenons:
  Background texture (borders) and Line Of Sight jitter

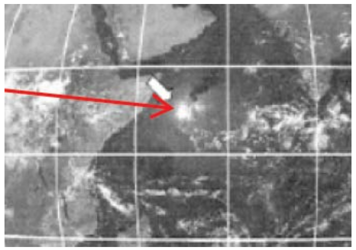
High energy particles on LI detector (Back Side Illuminated Thinned configuration)

- Galactic cosmic rays or sun radiation flux during solar flares
- Protons are the main contributors of single event phenomena (SEP) on detector
- Heavy ions can trigger few events with quasi circular shape

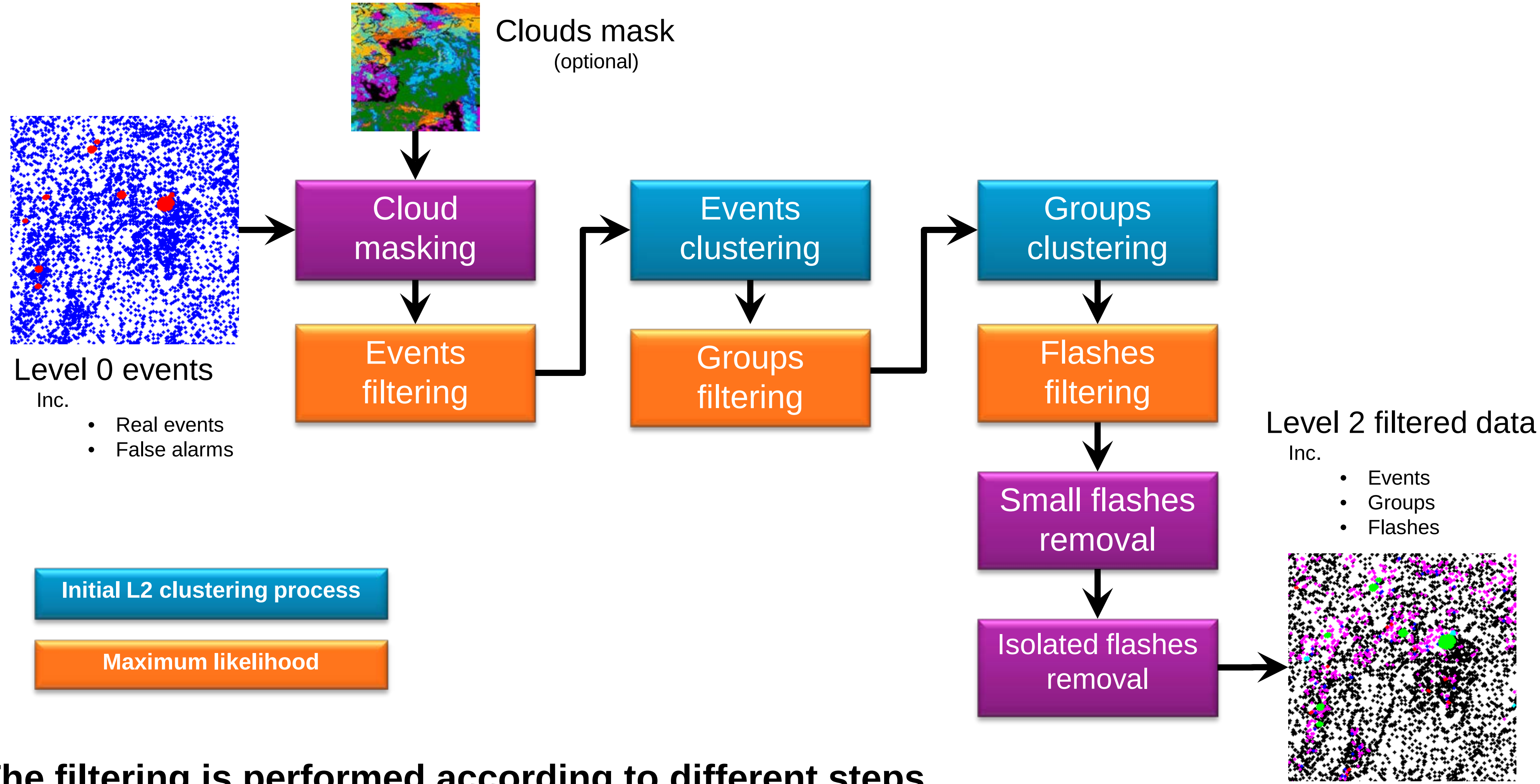


Sun glint

- Position accurately predictable
- Slow evolution compared to frame frequency
- Should no be different from bright clouds



The filtering algorithm: a general filtering scheme



The filtering is performed according to different steps

- Each step is fully configurable and optional

The filtering algorithm: the Bayesian approach

Set of elements (events, groups or flashes): Two classes: real (C=1) or false (C=-1)

Set of features: For each element, characteristics are computed: **F** = (F_1, F_2, …, F_n)

Classification:

- Bases on probabilities related to features: P(C=1 | F) and P(C=-1 | F)
- A binary classifier can be defined: $h \in \begin{cases} 1 & if\ \frac{P(C=1|\boldsymbol{F})}{P(C=-1|\boldsymbol{F})} > \mu \\ -1 & otherwise \end{cases}$

Decision h \ event c	C=1	C=-1
Confirmed (h=1)	ok	False alarm
Rejected (h=-1)	Non detection	ok

Each feature is characterized by its distribution:

- P(C=1 | Fi) is the probability of having a Real Event, provided that the feature #i has the value Fi
- P(C=-1 | Fi) is the probability of having a False alarm, provided that the feature #i has the value Fi

A "naïve" global criterion is defined, assuming independence between features

$$\prod_i \boldsymbol{P(C = 1/Fi)} > \boldsymbol{\mu}. \prod_i \boldsymbol{P(C = -1/Fi)}$$

- P(C=1 / Fi), P(C=-1/ Fi) and µ have to be calibrated

Calibration

Calibration consists in

- Estimating distributions curves P(C=1 / Fi), P(C=-1 / Fi)
- Tuning the thresholds, at each step of filtering

Features characteristics depends on observational conditions leading to define classes

- Time in the day (ex: daytime, night-time) based on local sun elevation
- Season in the year
- Geographical positions (ex: Europe, Africa, Oceans, …)

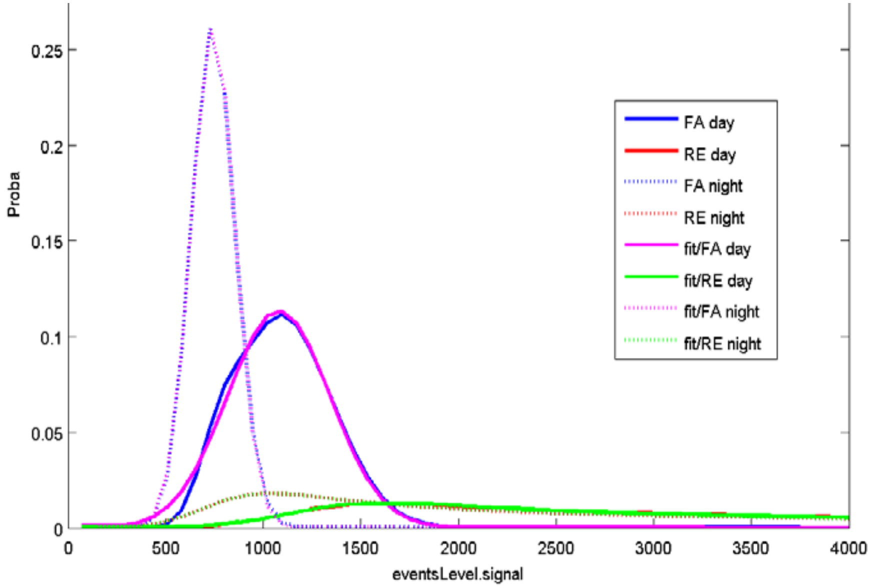
➢ It is necessary to define a set of parameters for each considered class

Calibration data set

- Acquired data set from LI
- Corresponding reference data issued from other means
- Sufficient data necessary for each considered class

Distributions estimation

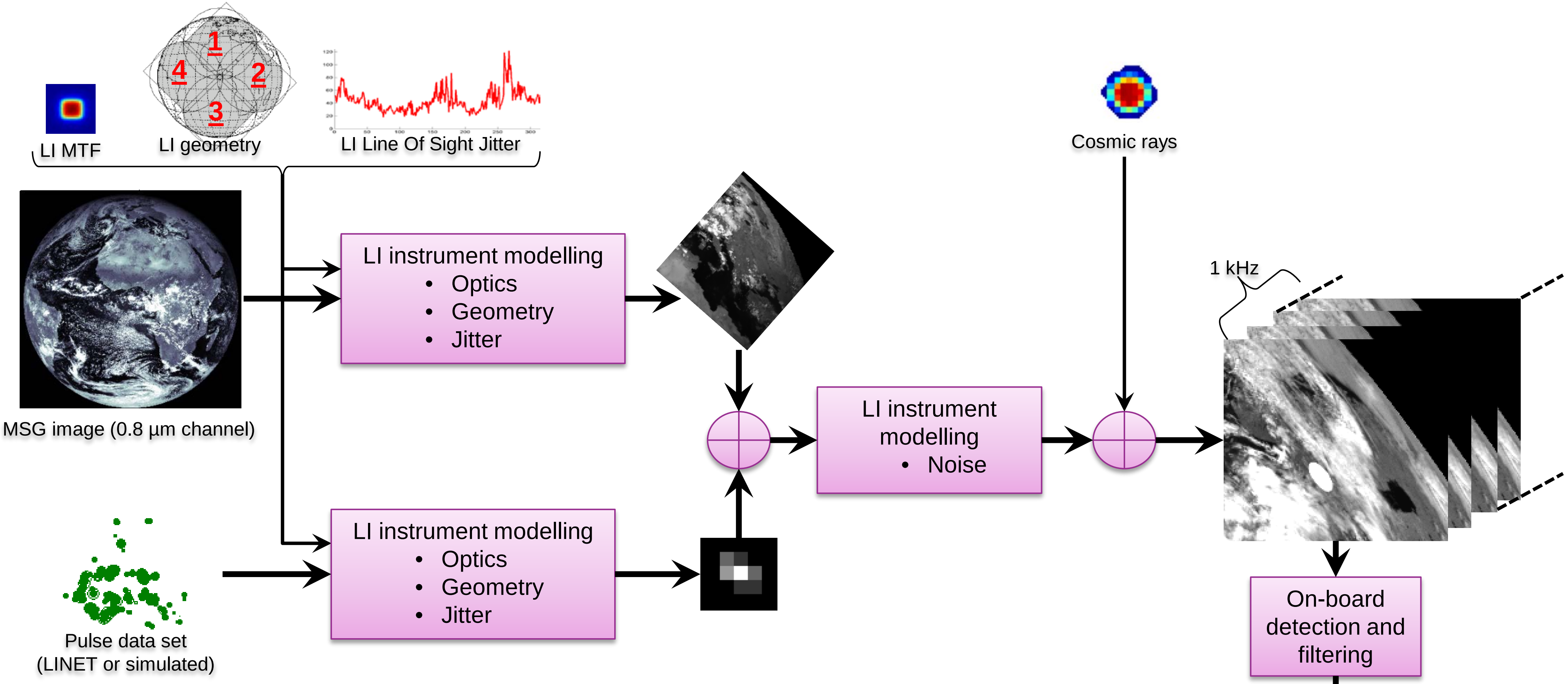
- Values Fi are computed for each feature #i according to data set
- Histograms are smoothed or fitted with theoretical curves

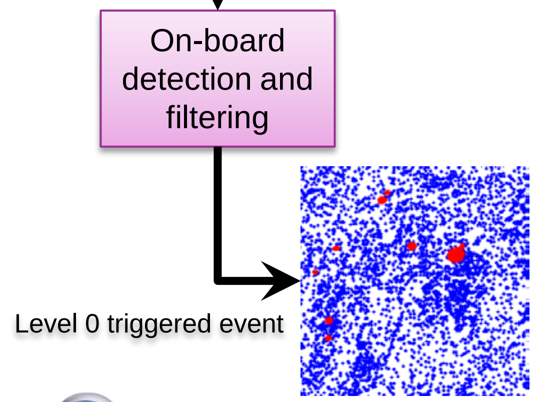


Thresholds are tuned

- To reach the “best compromise” between
  - false events rate (FAR) or false flashes rate (FFR)

Performance estimation: the LIDEFAS simulator





LI Detection Efficiency and False Alarms simulator

- An image simulator
- Is coupled with the filtering and clustering algorithms prototype

Performances

Comparison between the filtering prototype results and the simulator inputs allows a performance estimation

- Example of scenario
  - 24 sequences of 1 mn duration every hour over one day: 28/07/2006
  - Source image: MSG, channel 0.8 µm
  - Sources pulse: data set based on the ground-based LINET observations
- Obtained results summarized in the following table

cases	L0 event DE %	L0 Pulse DE %	L2 FAR ev/s	L2 FGR gr/s	L2 FFR fl/s	L2 Pulse DE %	L2 Flash DE %
Reference	94.1	97.6	19.1	5.9	1.7	93.2	91.9
Signal / SRD	78.5	85.4	183.2	93.6	12.9	72.8	69.0
SRD requirements			< 350		< 2.5 (TBC)	> 70	

- L0 columns provide Detection Efficiency (DE) for events and pulses related to the on-board detection and filtering only
- L2 columns provide performances related to the full acquisition chain (on-board detection and filtering, on-ground filtering and clustering)
- False alarms are related to events (FAR), groups (FGR) or flashes (FFR).
- Line “Signal/SRD” is obtained while replacing pulses radiance and size by minimal values defined in the System Requirement Document

- Performances are resulting from a compromise between DE and FAR that depends on thresholds tuning.

Conclusion

In the course of the study

- The noise sources have been identified and analysed all along the data acquisition and processing chain
- An algorithmic baseline for the LI false alarms filtering has been designed, prototyped and tested
- A complete LI Detection efficiency and False Performance simulator has been developed and is able to generate level 0 events from input reference data
- A calibration methodology and procedure has been proposed
- A performance estimation through simulation campaign has been conducted

The proposed filtering concept showed very good performances

- The algorithmic baseline remains relatively simple
- It is fully configurable