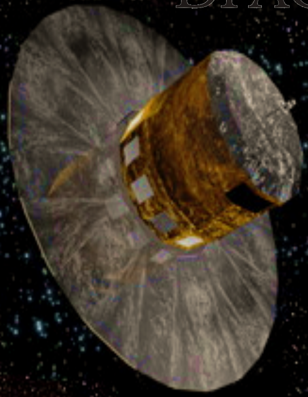
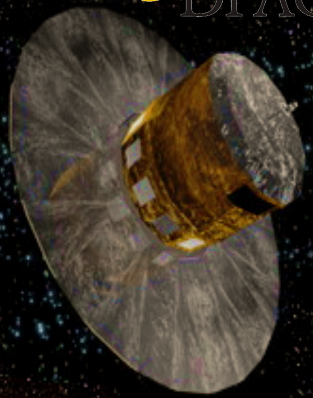
# Gaia, an overview



F. Arenou, Observatoire de Paris/CNRS
With many slides contributed from DPAC members from CU1 to CU9:
C. Bailer-Jones, A. Brown, D. Evans, L. Eyer, J. Hernandez, D. Katz,
F. Mignard, U. Lammers, W. O'Mullane, D. Pourbaix, T. Prusti, P. Tanga

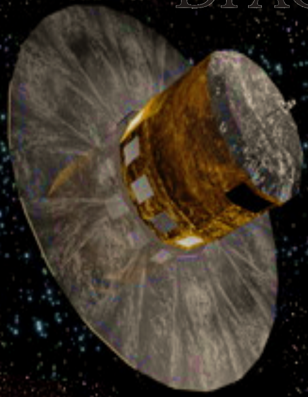
# Outline



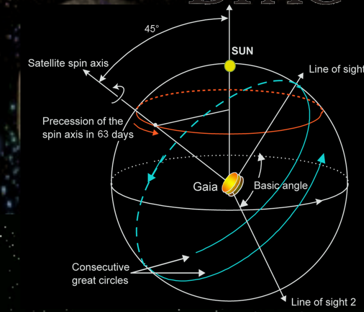
1. The Gaia satellite
   - ❑ Instrument
   - ❑ Scanning law
2. The DPAC processing
   - ❑ Organisation
3. Data processing
   - ❑ AGIS
   - ❑ CUs
4. DR1
   - ❑ Content
   - ❑ Performances
   - ❑ Caveats
5. DR2 and beyond
   - ❑ DR2
   - ❑ Expected improvements
   - ❑ Mission extension

A big challenge !

# The Gaia mission



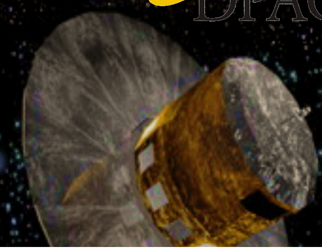
# Gaia in operation

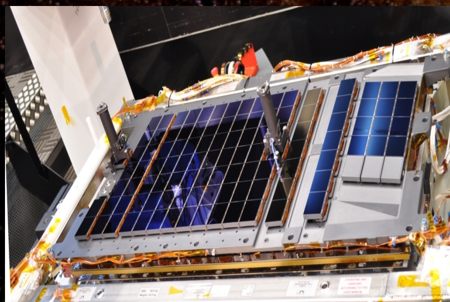


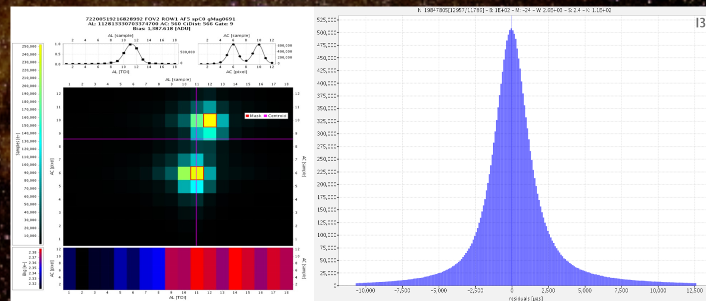


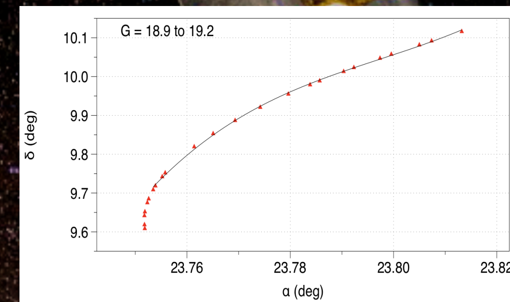
- In routine operations
  - Since July 2014
  - Nominal mission ends mid-2019
- All sources 3 < G < 20.7 (and beyond)
  - Bright stars covered with special observations
  - Crowded regions observed with sky mappers

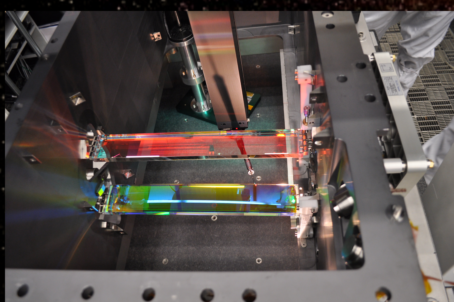
# Gaia: an observatory in orbit

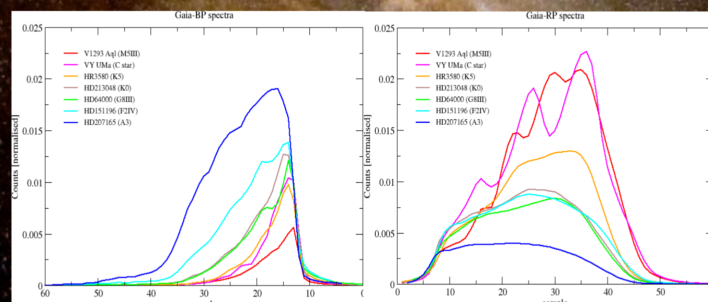


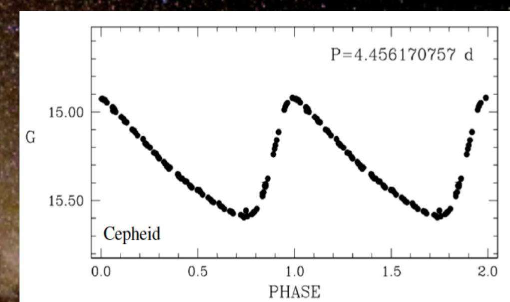


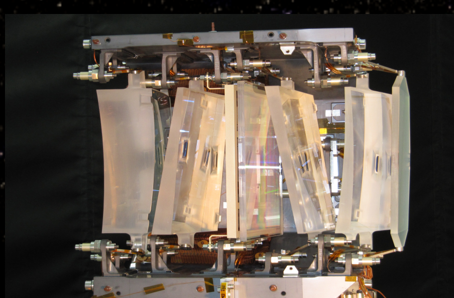


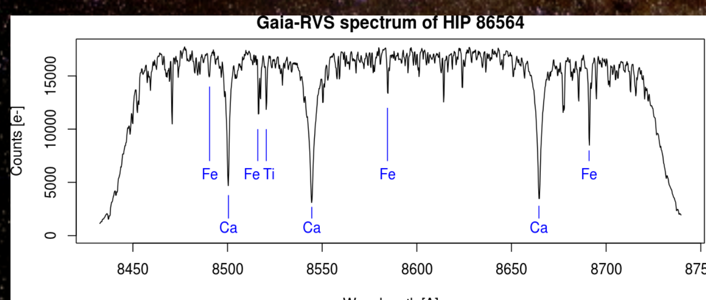



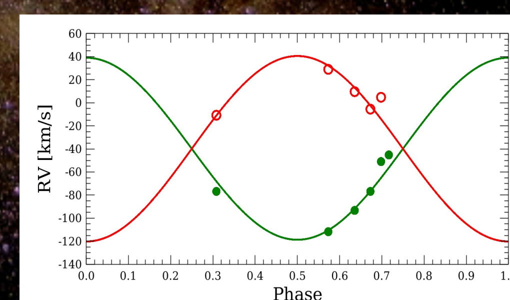





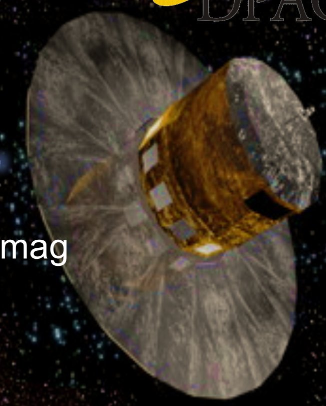






Credit: Airbus, CU3, CU4, CU5, CU6, CU7, CU8

# Focal plane transit

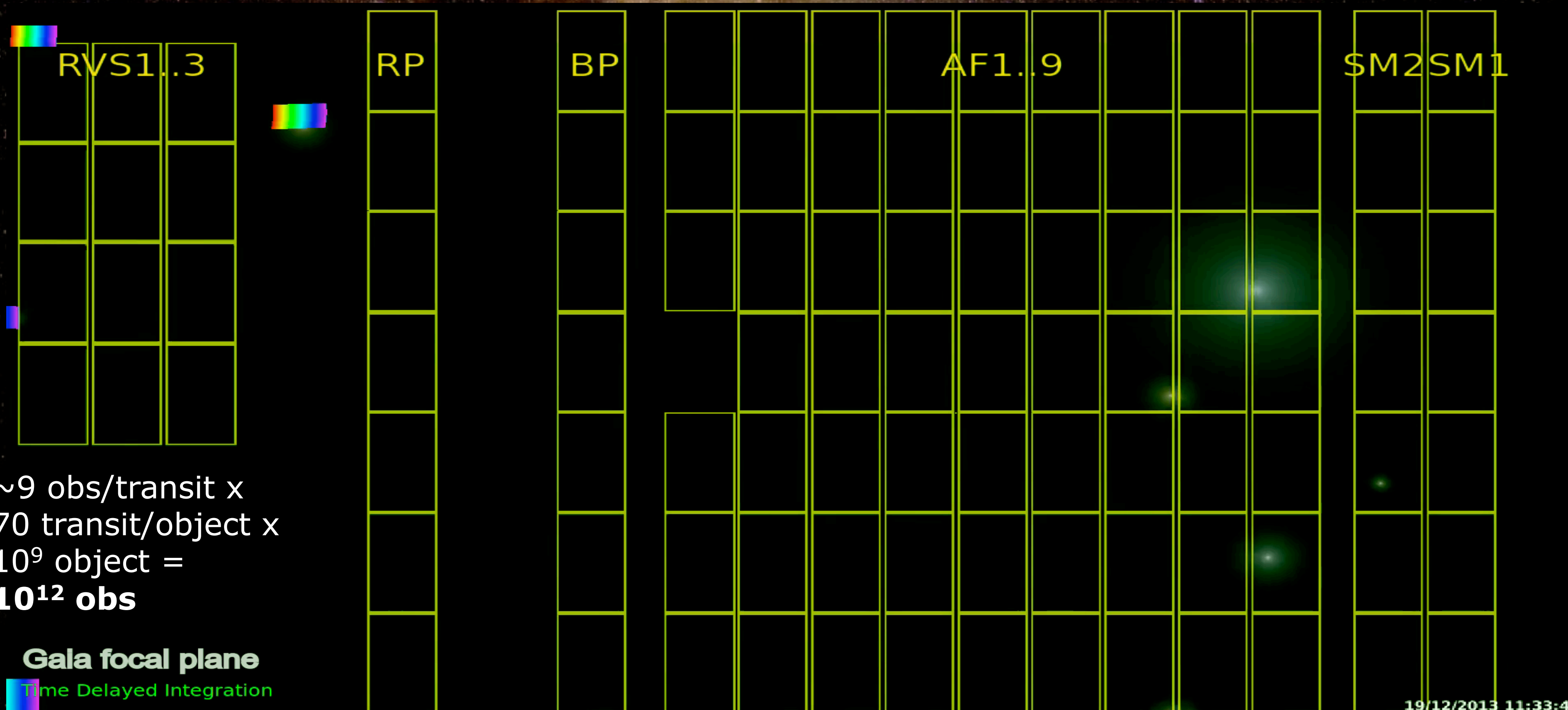


**Photometry:**
- spectro-photometer
- blue and red CCDs

**Spectroscopy:**
- high-resolution spectra
- red CCDs

**Total field:**
- active area: 0.75 $deg^2$
- CCDs: 14 + 62 + 14 + 12
- 4500 x 1966 pixels (TDI)
- pixel size = 10 µm x 30 µm
  = 59 mas x 177 mas

**Sky mappers:**
- detects all objects to 20.7 mag
- rejects cosmic-ray events
- FoV discrimination



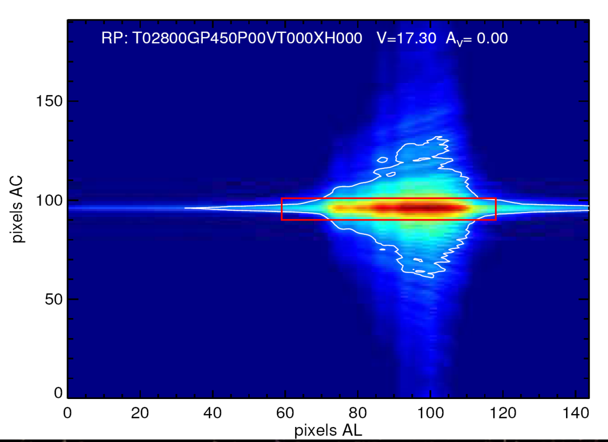

~9 obs/transit x
70 transit/object x
$10^9$ object =
**$10^{12}$ obs**

Credit: U. Lammers/AGIS

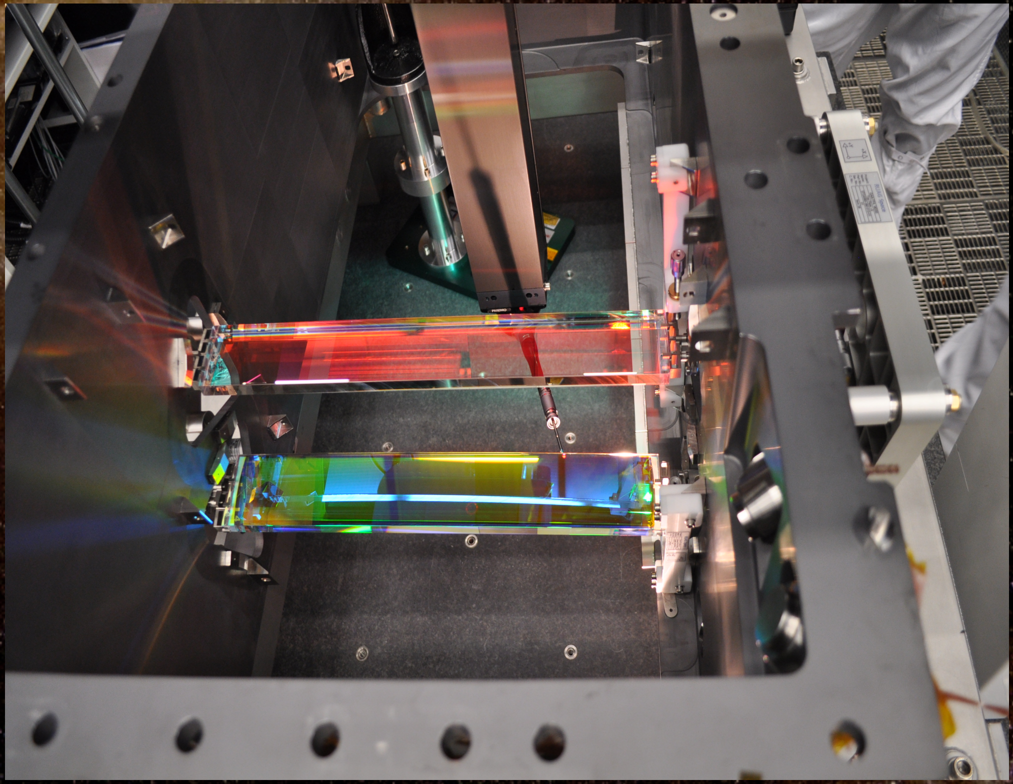
# Photometry Measurement

Blue photometer:
330–680 nm

Red photometer:
640–1050 nm

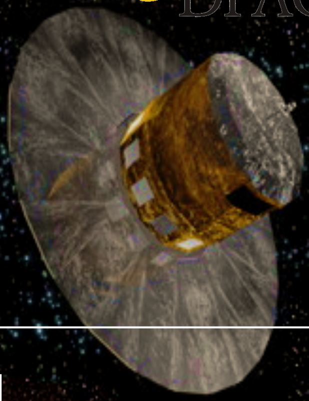



RP spectrum of M dwarf (V=17.3)
Red box: data sent to ground
White contour: sky-background level
Colour coding: signal intensity

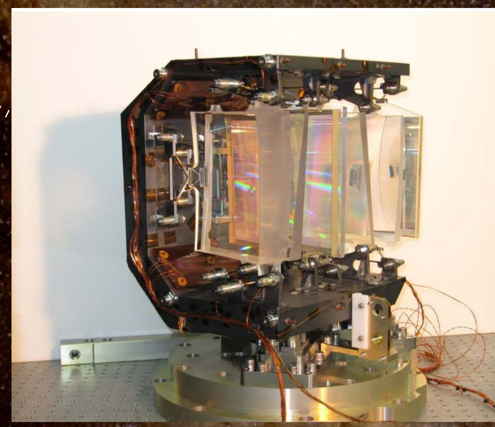


Figures courtesy EADS-Astrium

# Radial Velocity Measurement

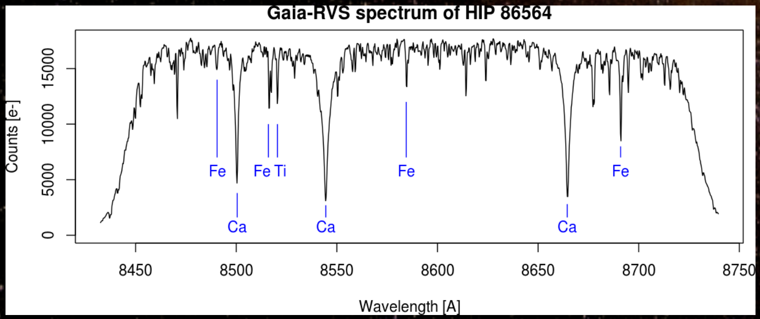


Field of view



RVS spectrograph

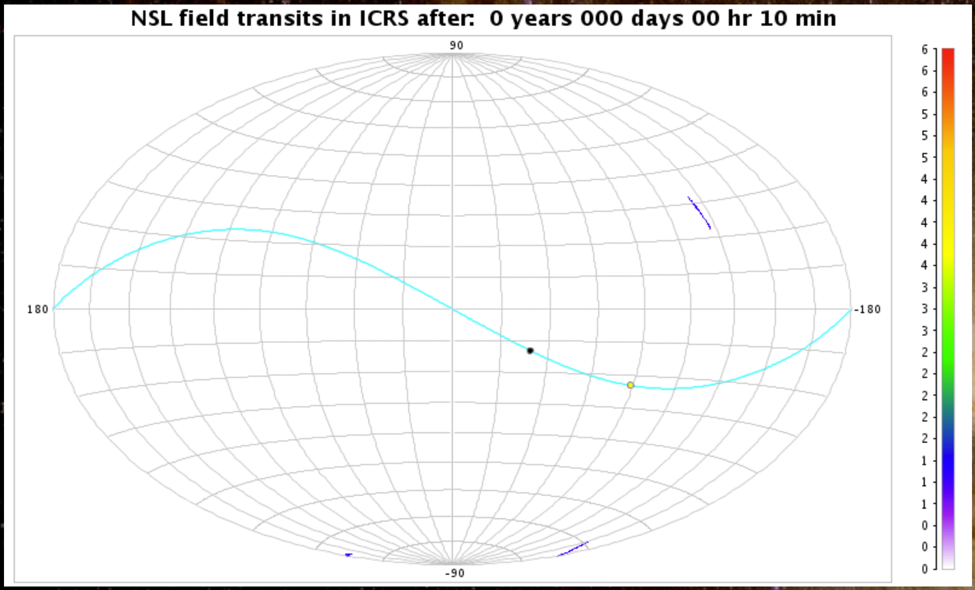
CCD detectors

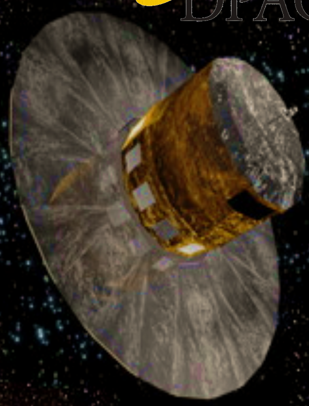


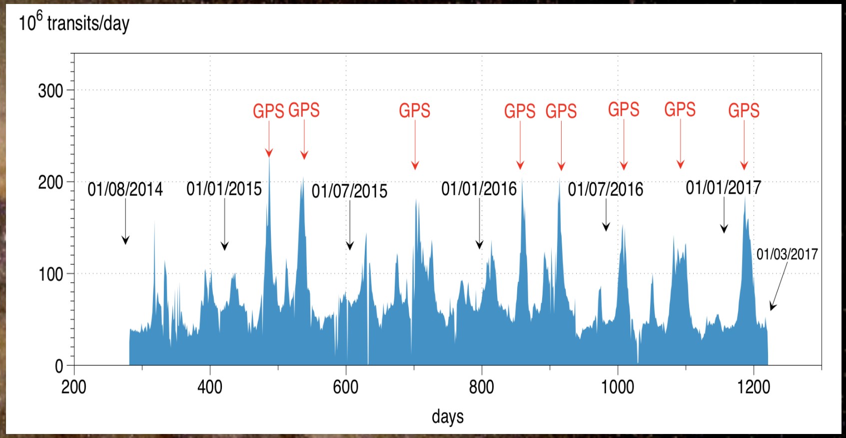

- Slitless spectroscopy
- using Ca triplet (848-872 nm)
- R=10800
- $G_{RVS} < 16.2$

Figures courtesy D. Katz, O. Marchal

# A typical Gaia day






Credit: F. Mignard

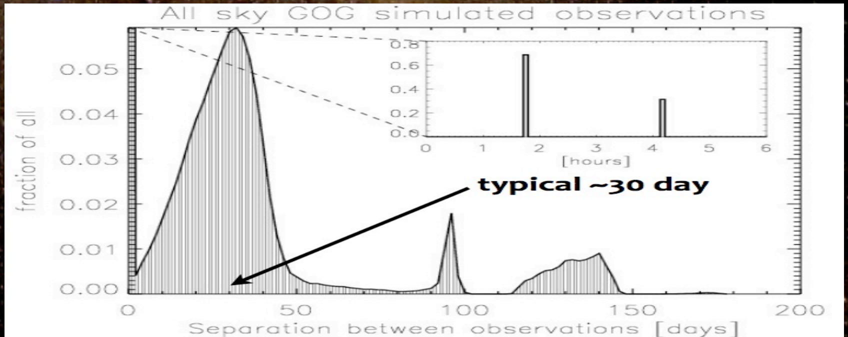
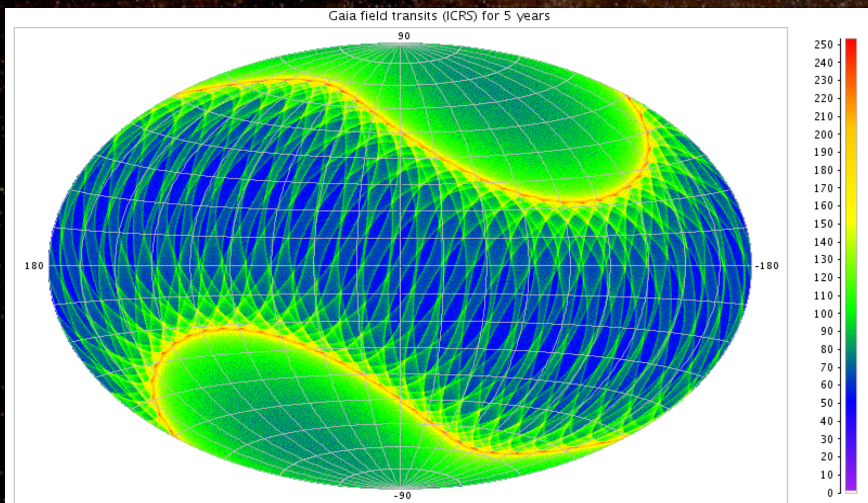



Figure courtesy: Lukasz Wyrzykowski

- 70 million sources
  - 600 million astrometric measurements
  - 155 million BP/RP measurements
  - 13 million spectra
- 40 GB telemetry to the ground

# End of mission performances




| EOM | B1V | | | G2V | | | M6V | | |
|---|---|---|---|---|---|---|---|---|---|
| **G [mag]** | **G** | **BP** | **RP** | **G** | **BP** | **RP** | **G** | **BP** | **RP** |
| **3 - 13** | 0.2 | 1 | 1 | 0.2 | 1 | 1 | 0.2 | 1 | 1 |
| **14** | 0.2 | 1 | 1 | 0.2 | 1 | 1 | 0.2 | 2 | 1 |
| **15** | 0.2 | 1 | 2 | 0.2 | 1 | 1 | 0.2 | 5 | 1 |
| **16** | 0.4 | 2 | 3 | 0.4 | 2 | 2 | 0.4 | 10 | 1 |
| **17** | 0.6 | 3 | 6 | 0.6 | 4 | 4 | 0.6 | 25 | 2 |
| **18** | 0.9 | 6 | 15 | 0.9 | 9 | 8 | 0.9 | 63 | 4 |
| **19** | 1.8 | 15 | 37 | 1.8 | 23 | 20 | 1.8 | 157 | 8 |
| **20** | 3.7 | 37 | 91 | 3.7 | 56 | 48 | 3.7 | 395 | 20 |

| | **B1V** | **G2V** | **M6V** |
|---|---|---|---|
| **$V-I_C$ [mag]** | -0.22 | 0.75 | 3.85 |
| **Bright stars** | 5-16 µas (3<V<12) | 5-16 µas (3<V<12) | 5-16 µas (5<V<14) |
| **V = 15 mag** | 26 µas | 24 µas | 9 µas |
| **V = 20 mag** | 600 µas | 540 µas | 130 µas |

| **Spectral type** | **V [mag]** | **Radial-vel. uncertainty [km $s^{-1}$]** |
|---|---|---|
| **B1V** | <7.5 | <1 |
| | 11.3 | 15 |
| **G2V** | <12.3 | <1 |
| | 15.2 | 15 |
| **K1III-MP (metal-poor)** | <12.8 | <1 |
| | 15.7 | 15 |

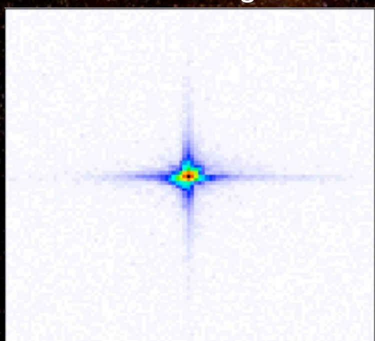
Gaia DPAC

# Bad surprises

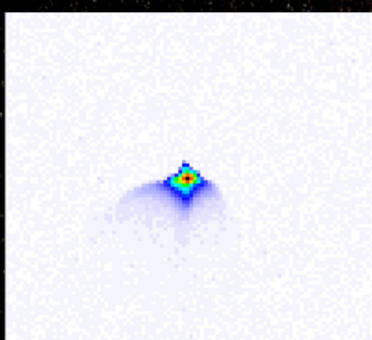
- Most is going fine (CCD, communication, attitude, etc.), but…
- Basic Angle Variations larger than expected
  - Corrected with Basic Angle Monitoring data for Gaia DR1 and DR2
- Contamination, transmission loss
  - Last decontamination August 2016; no sign of transmission loss since
- Attitude variations: micrometeoroids and micro-clanks
  - Will be better taken into account for Gaia DR2
- Stray light
  - Main impact on faint sources and spectroscopy
  - On-board software modified from read-out to background dominated

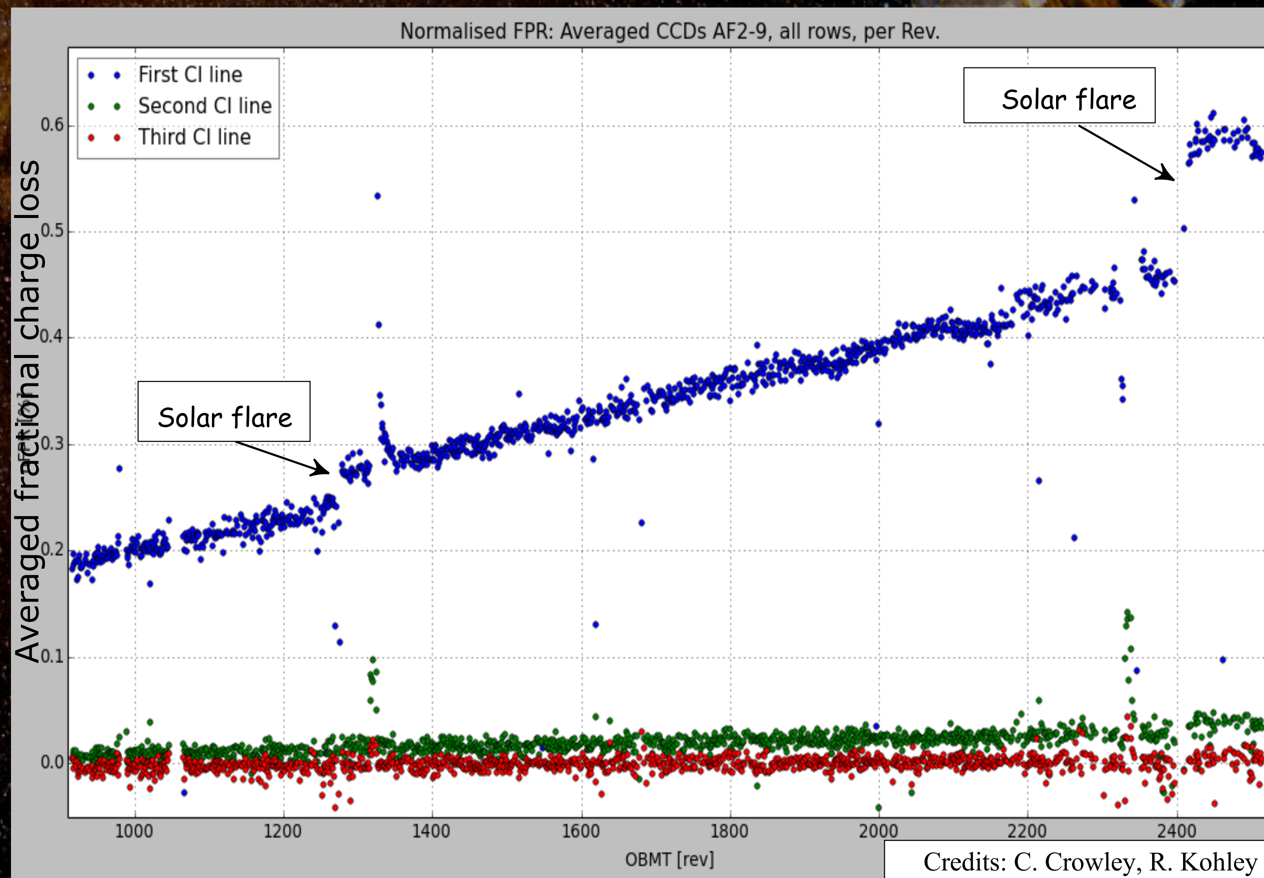
# Some good surprises: radiation

Nominal signal

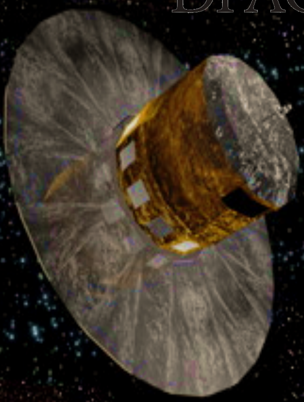


Signal after radiation damage





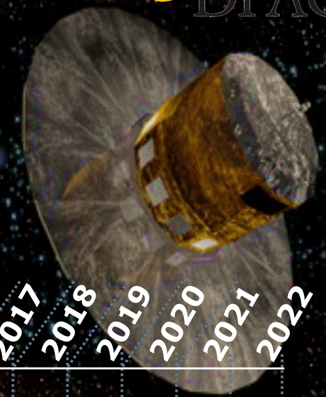

- ❑ CTI effect function of amplitude of the signal + radiation history
- ❑ Radiation damage+CTI evolution monitored through charge injections
- ❑ Projection to end of mission is a factor ~10 smaller than expected

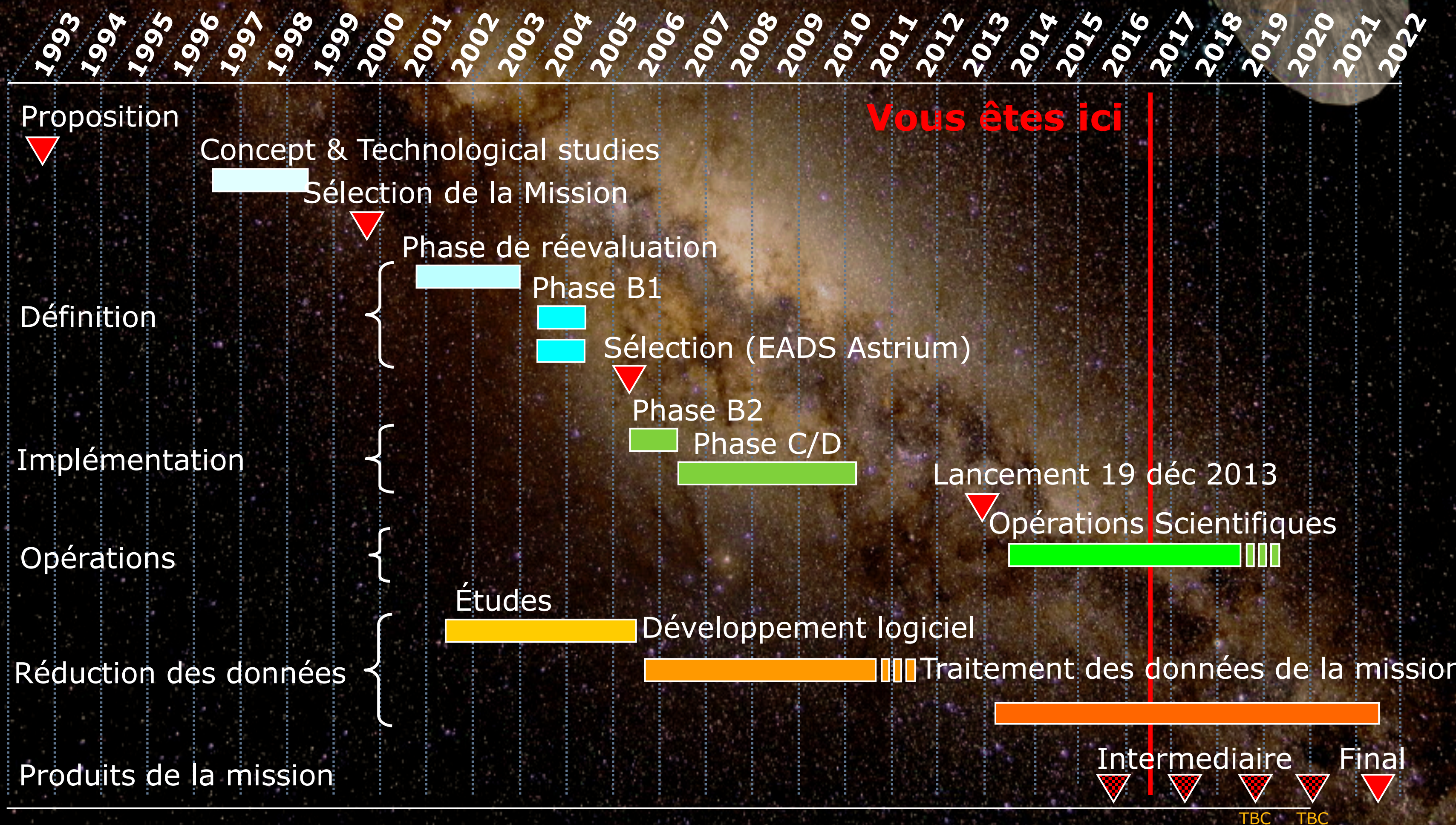


DPAC

# Data Processing & Analysis Consortium

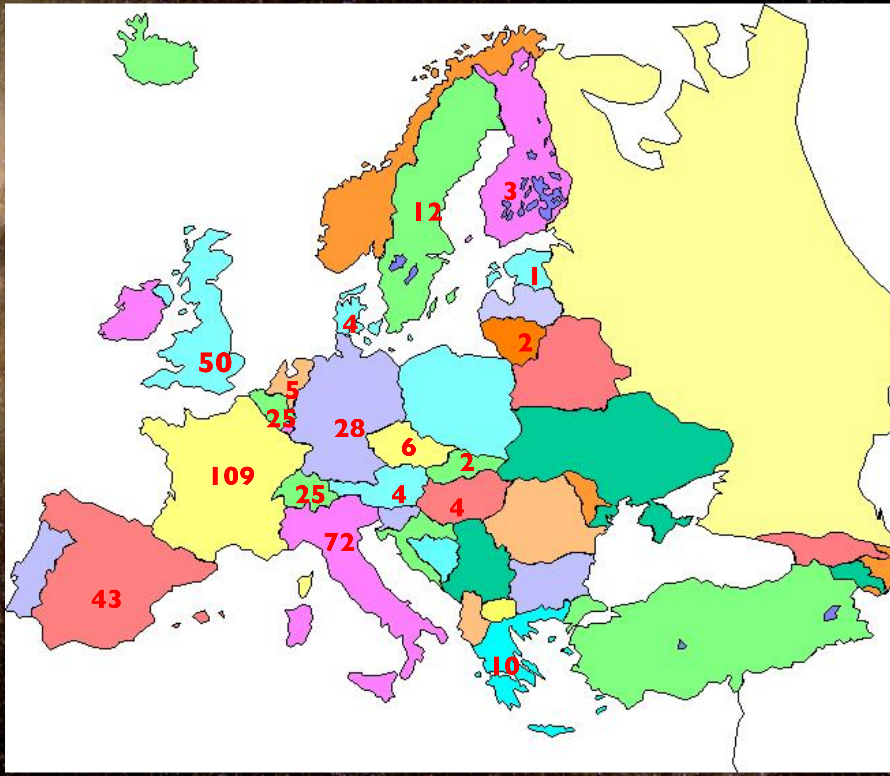
# Gaia: calendrier





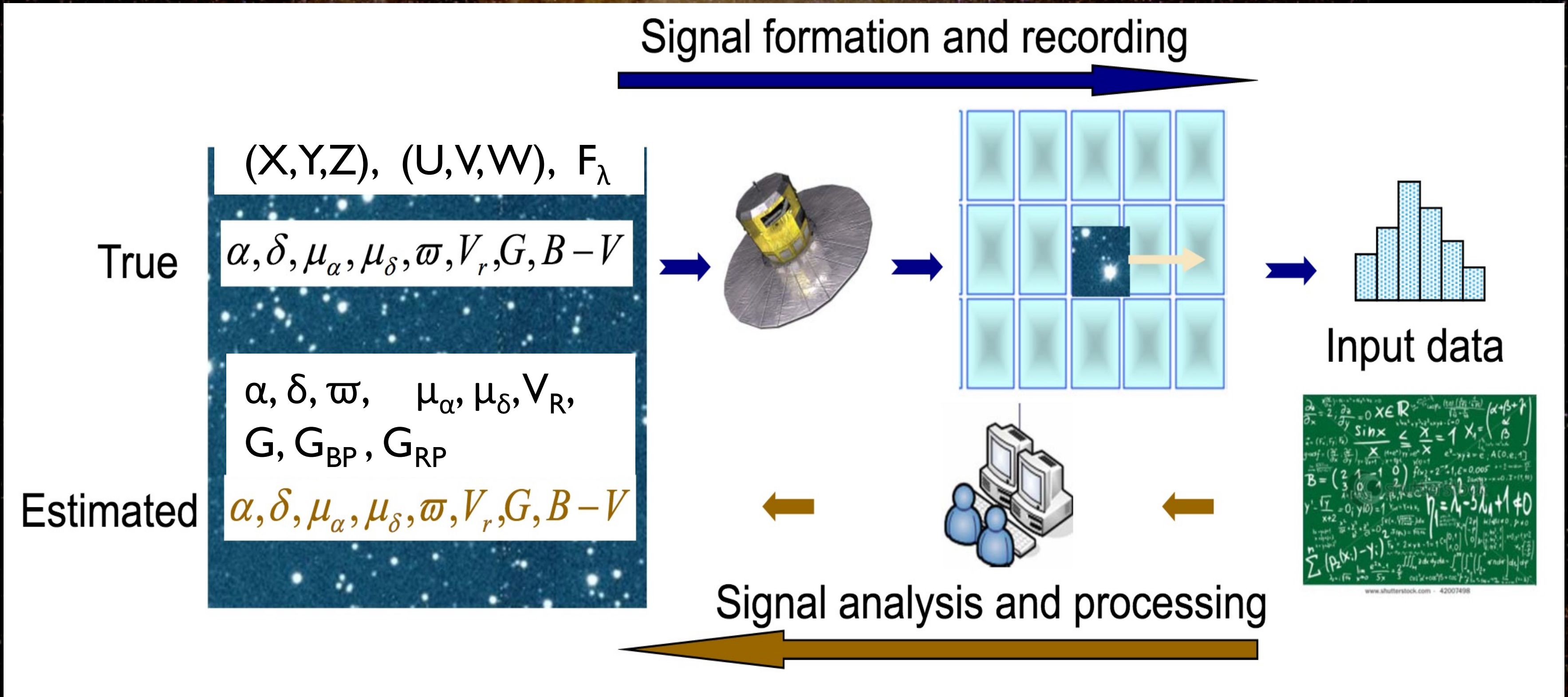
# Ground segment

- DPAC: European Consortium
  - For Member States (25 agencies)
  - ~450 people
  - 2006-2016: > 1000 P-Y
- A large ESA involvement
  - ESTEC - Mission Manager
  - ESTEC - Scientific Support (3p.)
  - ESAC – Sci. Ground Seg. (20p.)
  - ESOC - Operations (23p.)
  - ESTEC - Project Team (18p.)

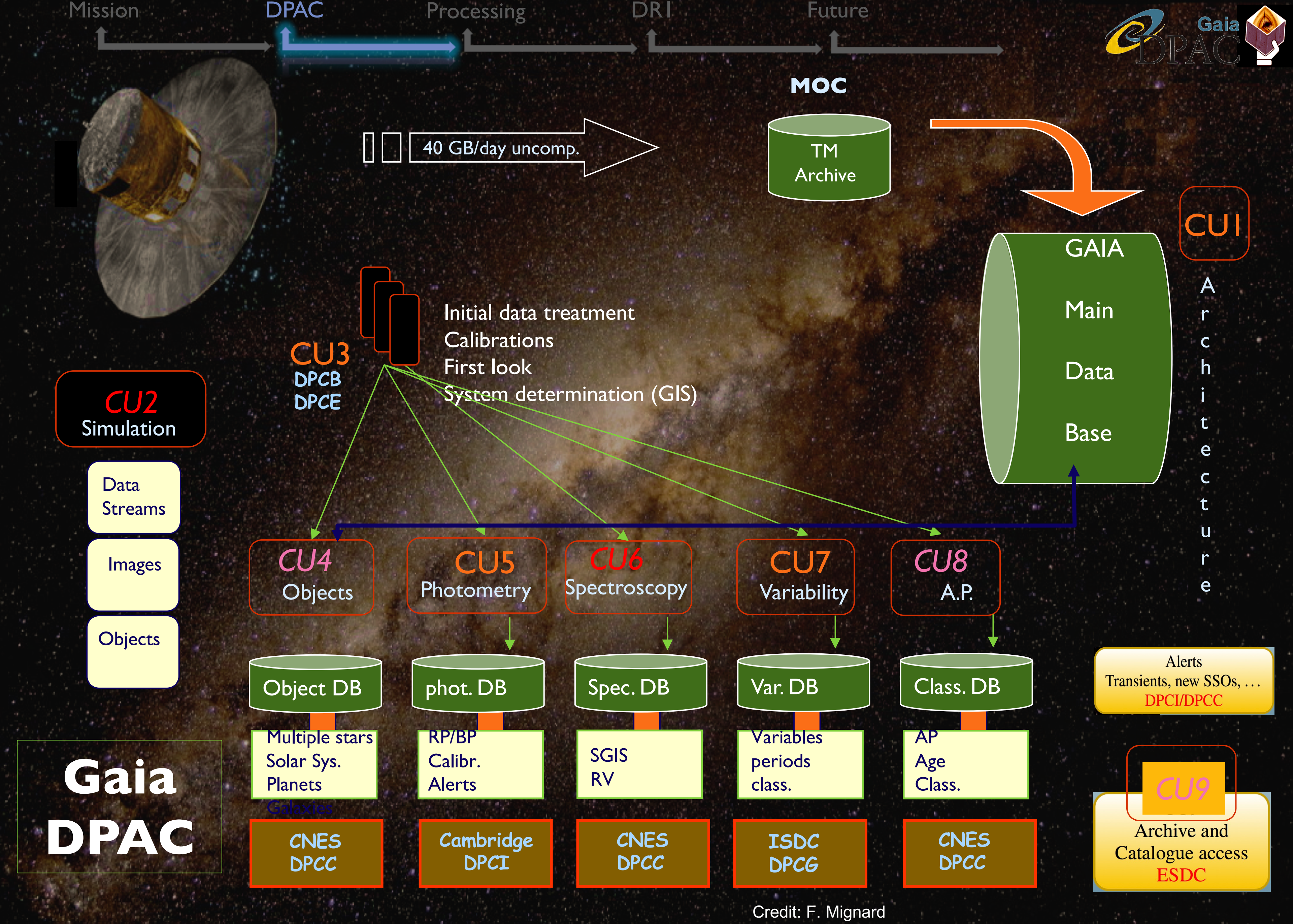


Credit: F. Mignard

# From time to space

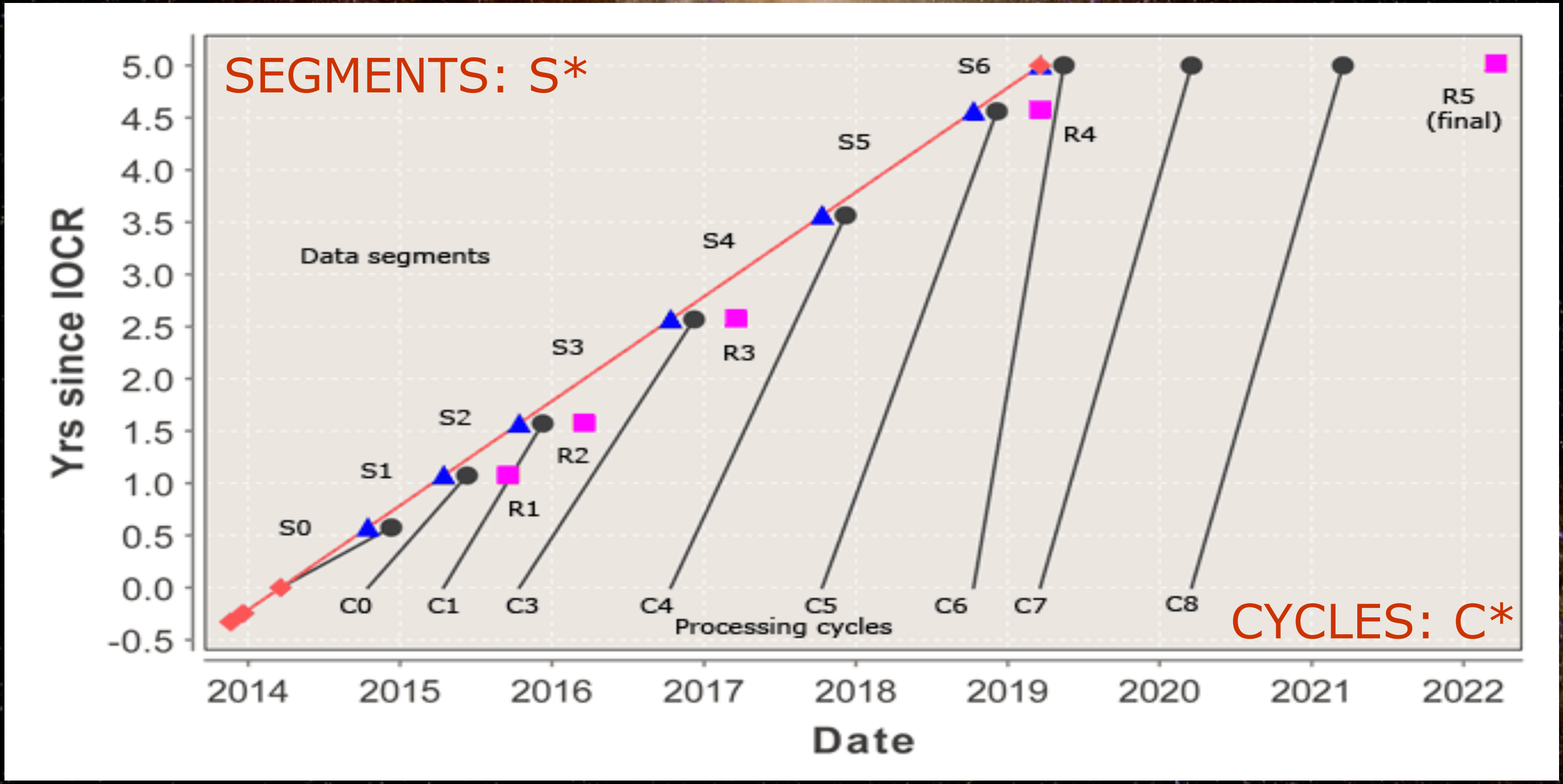


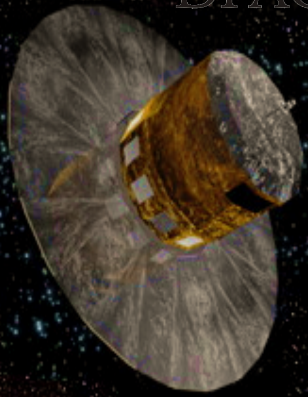

Measuring time → 1D positions → 2D positions → parallaxes → Measuring space


Mission
DPAC
Processing
DR1
Future
Gaia DPAC
MOC
40 GB/day uncomp.
TM Archive
CU1
GAIA Main Data Base
Architecture
Initial data treatment
Calibrations
First look
System determination (GIS)
CU3
DPCB
DPCE
CU2
Simulation
Data Streams
Images
Objects
CU4 Objects
CU5 Photometry
CU6 Spectroscopy
CU7 Variability
CU8 A.P.
Object DB
phot. DB
Spec. DB
Var. DB
Class. DB
Multiple stars
Solar Sys.
Planets
Galaxies
RP/BP
Calibr.
Alerts
SGIS
RV
Variables
periods
class.
AP
Age
Class.
CNES DPCC
Cambridge DPCI
CNES DPCC
ISDC DPCG
CNES DPCC
Alerts
Transients, new SSOs, …
DPCI/DPCC
CU9
Archive and
Catalogue access
ESDC
Gaia DPAC
Credit: F. Mignard

# Data flow: segments, cycles

- Data is artificially split in time segments (6 to 12 months)
- Data processing is iterative (e.g. CU4 needs CU3), thus cyclic
  - So the dates of Data Releases are driven by total duration of cycles
  - While the release quality is driven by # of segments used + # cycles

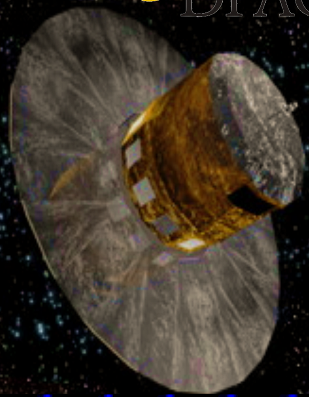


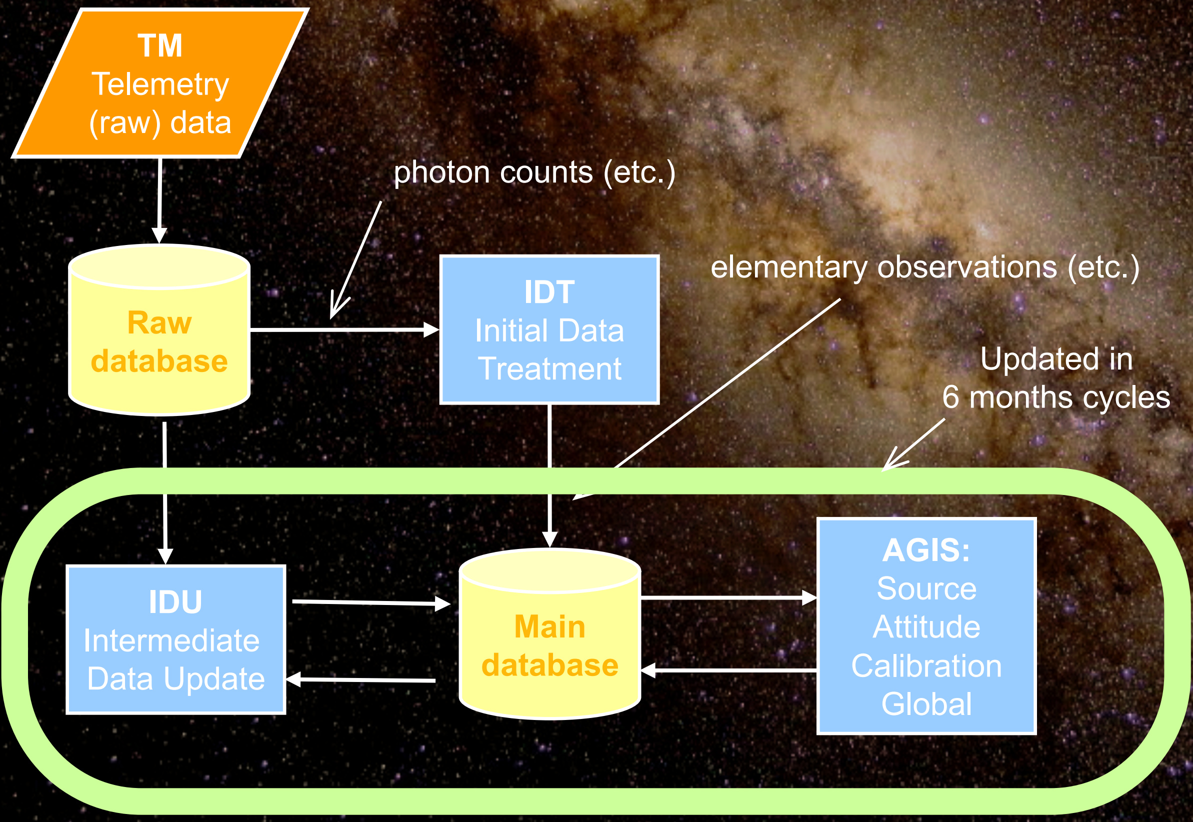


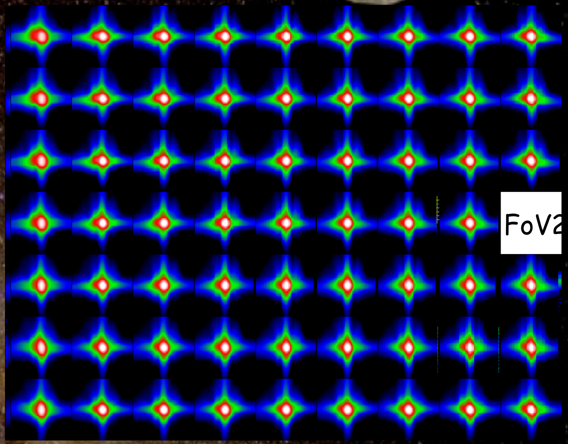
CU3, CU4, CU5, CU6, CU7, CU8, CU9

# Data processing

# CU3: Overall processing

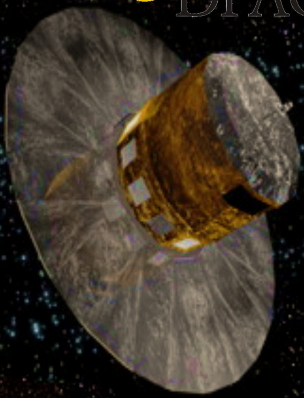





Credit: U. Lammers / AGIS/CU3

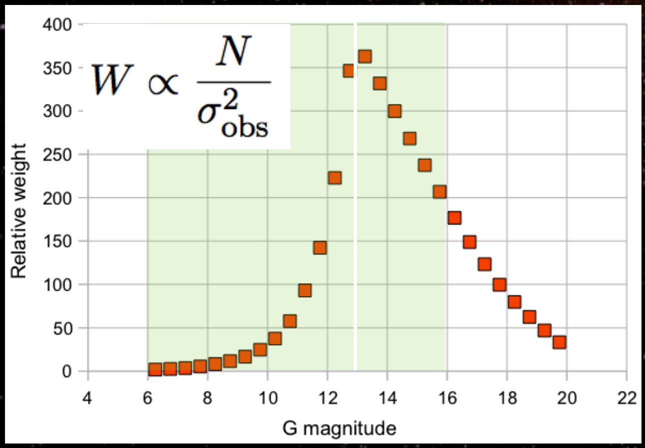
# CU3: AGIS



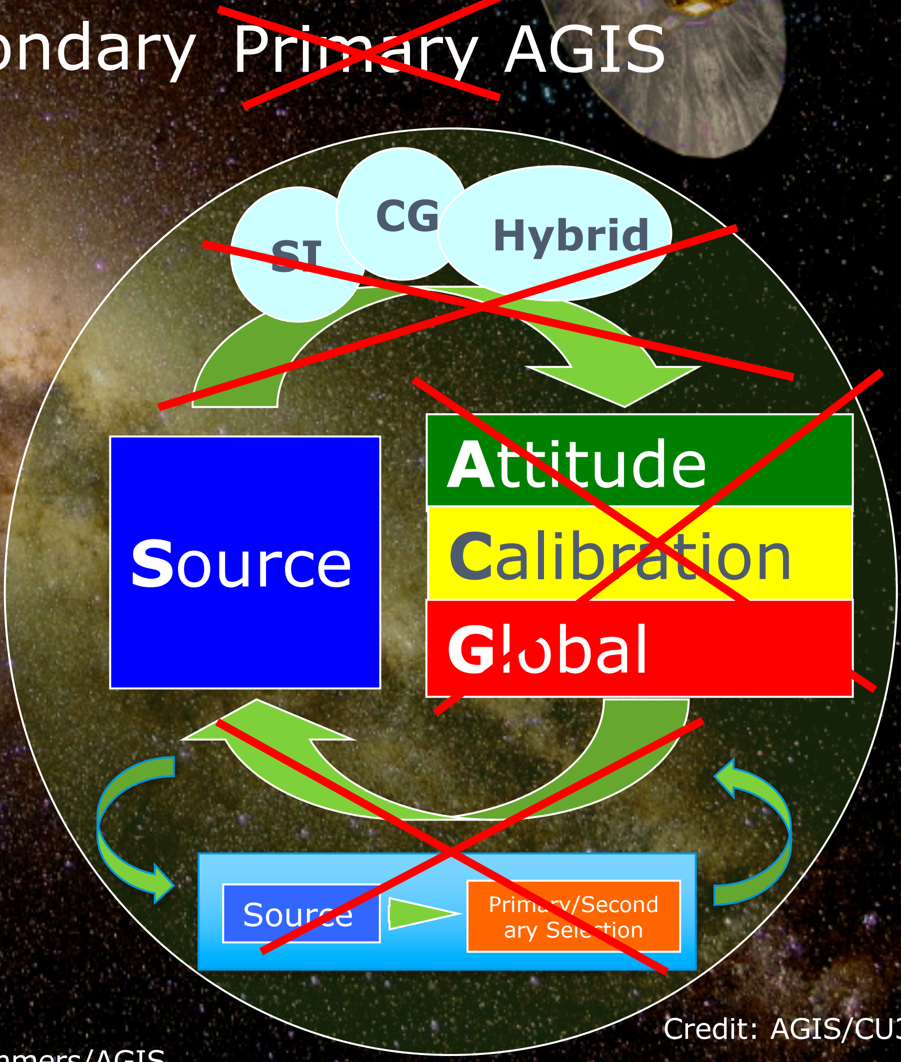
- AGIS
  - Astrometric, Global, Iterative, Solution
- Measurements:
  - Transit time, when the centroid of the LSF moves over the center of the light-sensitive area of the CCD + mean acros-scan column
  - From the transit time and satellite attitude, determine the proper direction of the object
- Direct solution is not possible
  1. S: Adjust 5 billion source param. using estimated attitude + calibration
  2. A: Adjust attitude using improved source parameters (unknowns)
  3. C: Adjust calibrations using improved source parameters (unknowns)
  - Goto 1, until updates are small enough

# CU3: Astrometric data flow

- Selection of primary stars
  - Well behaved, enough observations for attitude determination and calibration (G<13 for the saturation gates)
  - Maximising the total astrometric weight N/σ² (typically 13<G<16)

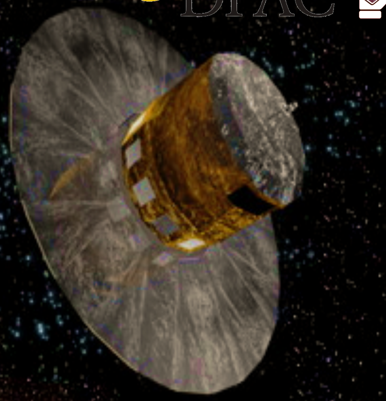


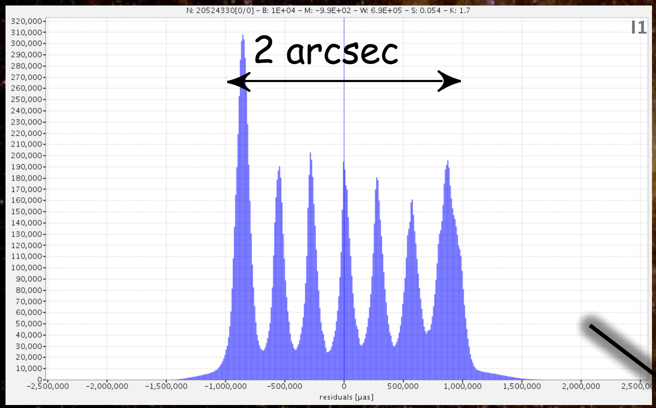

Credit: U. Lammers/AGIS

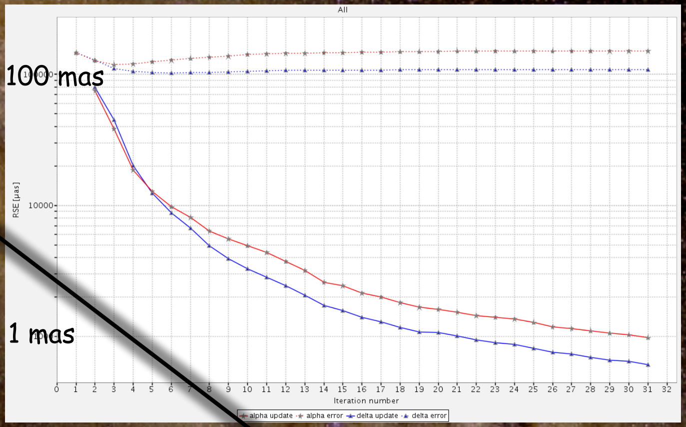


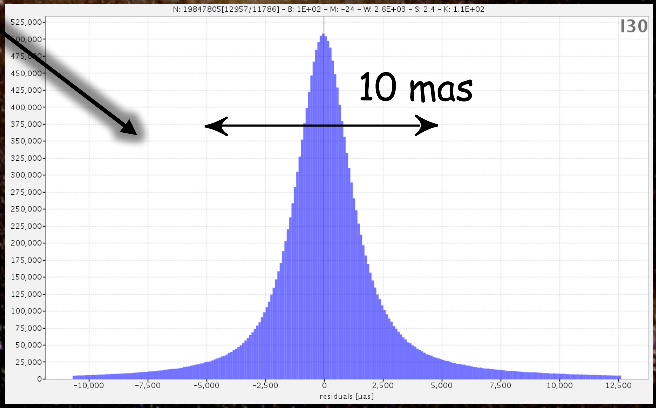

Credit: AGIS/CU3

# CU3: Convergence



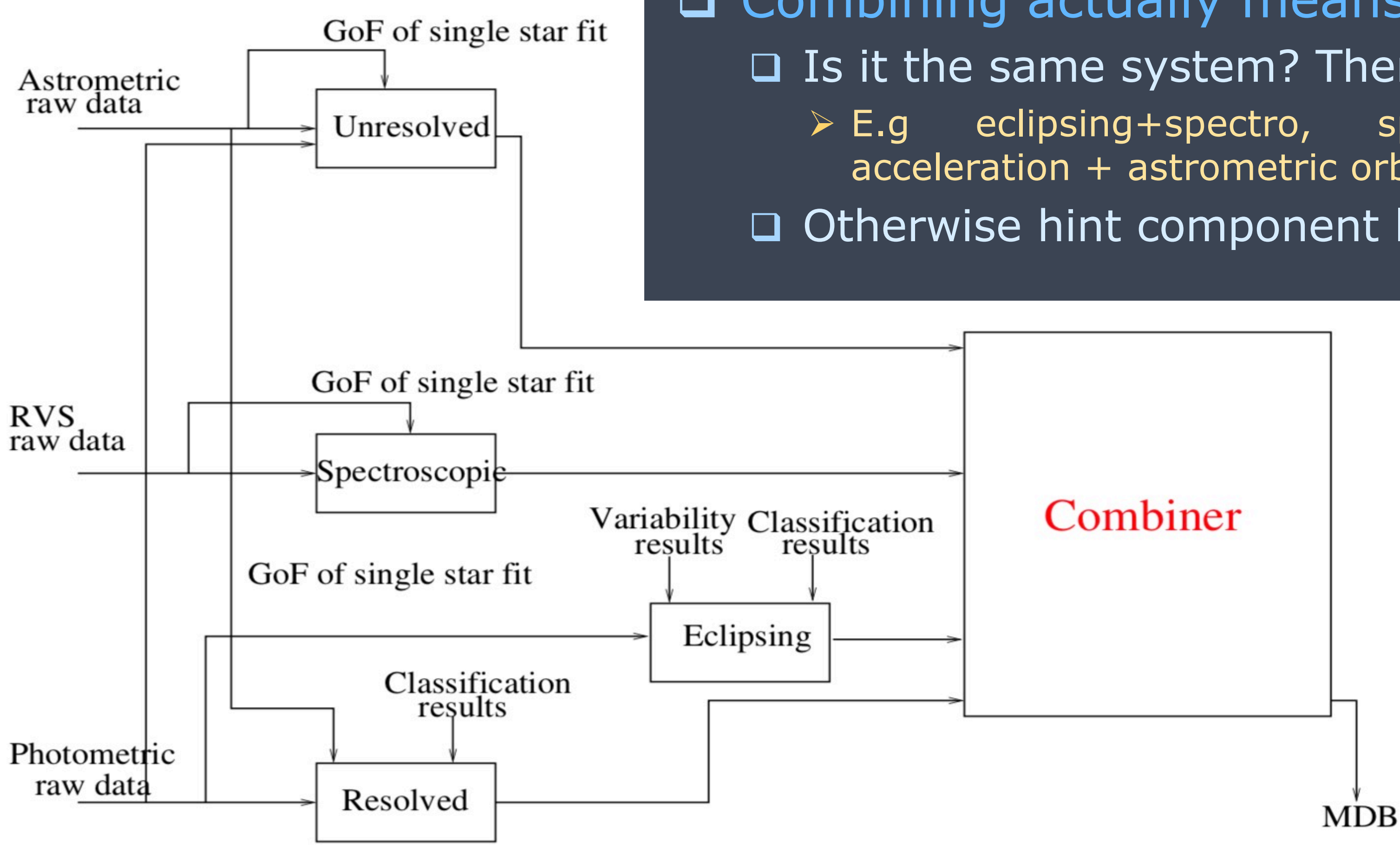



- From ~1000 billion measurements
- 5 billion unknown source parameters (***S***)
  - Plus Global or nuisance parameters:
  - ***A***ttitude: cubic splines between knot intervals ~15s for quaternions: 40 million parameters
  - ***C***alibration
  - Unfortunately with a complex interconnection

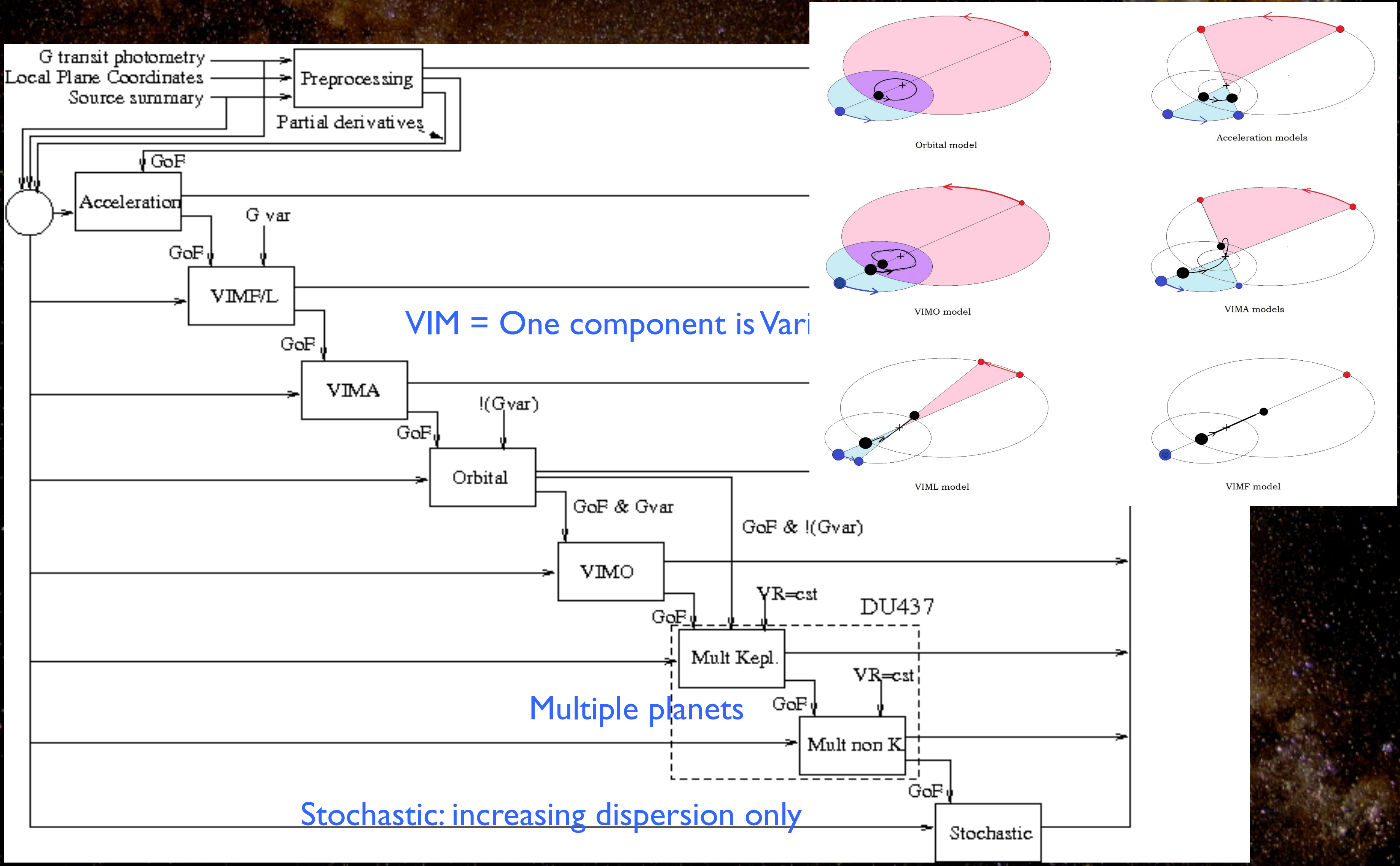
# CU4: Multiple stars + multiple instruments




- Combining actually means
  - Is it the same system? Then try.
    - E.g eclipsing+spectro, spectro acceleration + astrometric orbital
  - Otherwise hint component level

D. Pourbaix FNRS @ ULB – Belgium NSS in Gaia DPAC

# CU4: astrometric zoo of non-singles



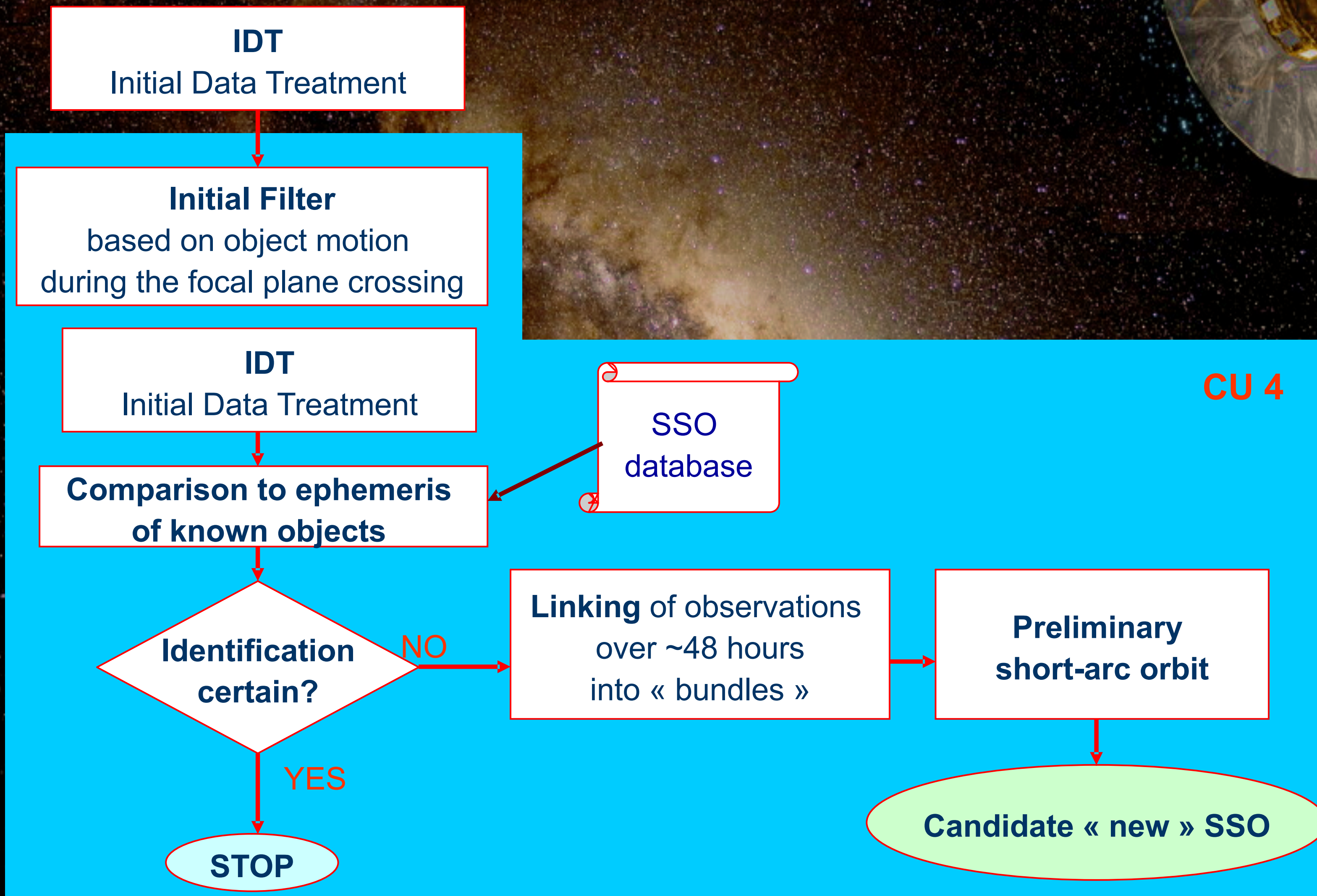

VIM = One component is Vari

Multiple planets

Stochastic: increasing dispersion only

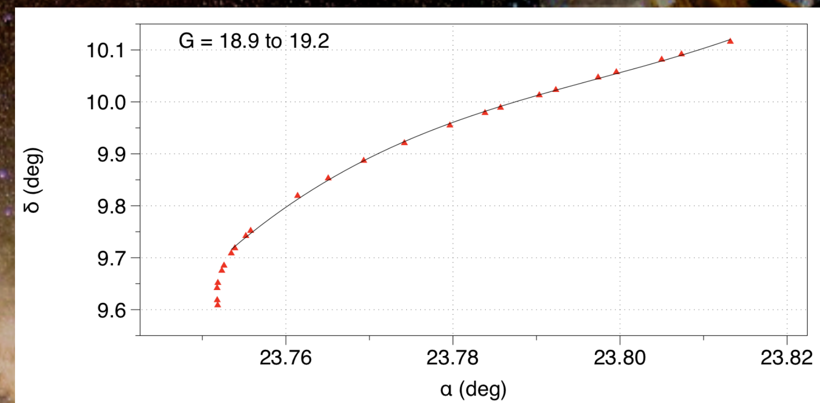
Gaia DPAC

# CU4: daily pipeline Solar System Obs.




Credit: P. Tanga

# CU4: Long term SSO processing




Credit: P. Mignard

**IDT**
Initial Data Treatment

**Comparison to ephemeris of known objects**

**Linking** of all available observations for each object into « threads »

SSO database

**Accurate orbits**

**Physical properties**

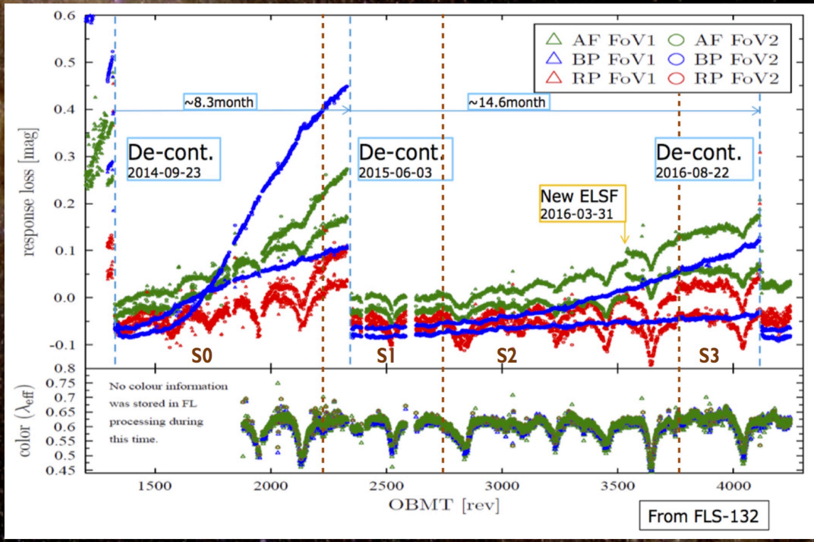
**Global parameters**

Mission outcome

Credit: P. Tanga

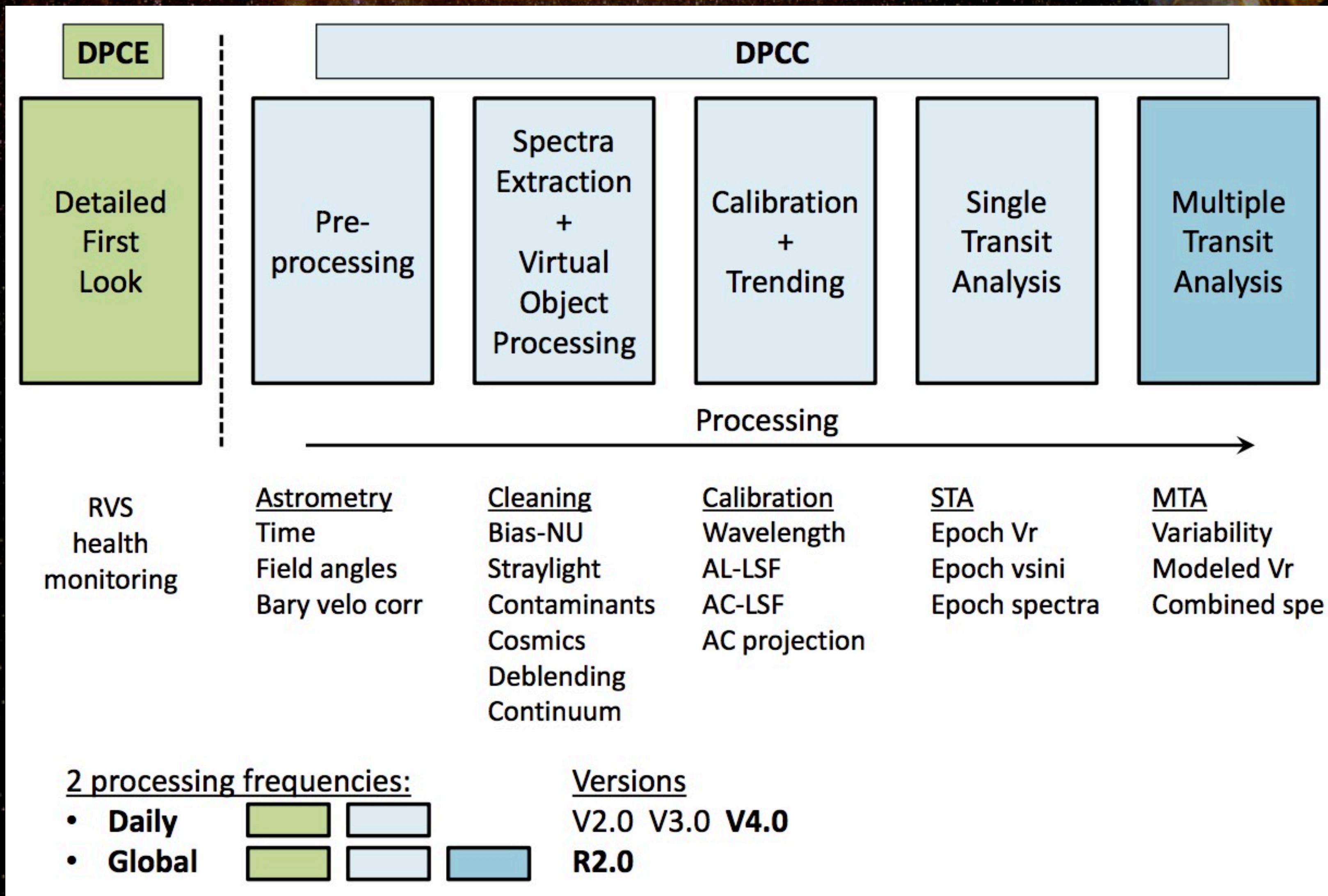
# CU5: Photometry

- Basic calibrations, photometric parameters, calibration model
  - Large Scale calibrations: variations in telescope & detectors on timescales of about one day
  - Small Scale calibrations: local response variations on the CCDs
  - Large complexity due to saturation + gates
  - Study of residuals: magnitude terms, colour effects, AC dependency
- Accumulation of mean flux information & epoch photometry
- Pass-band characterisations and absolute flux calibrations




Credit: J. Martin-Fleitas

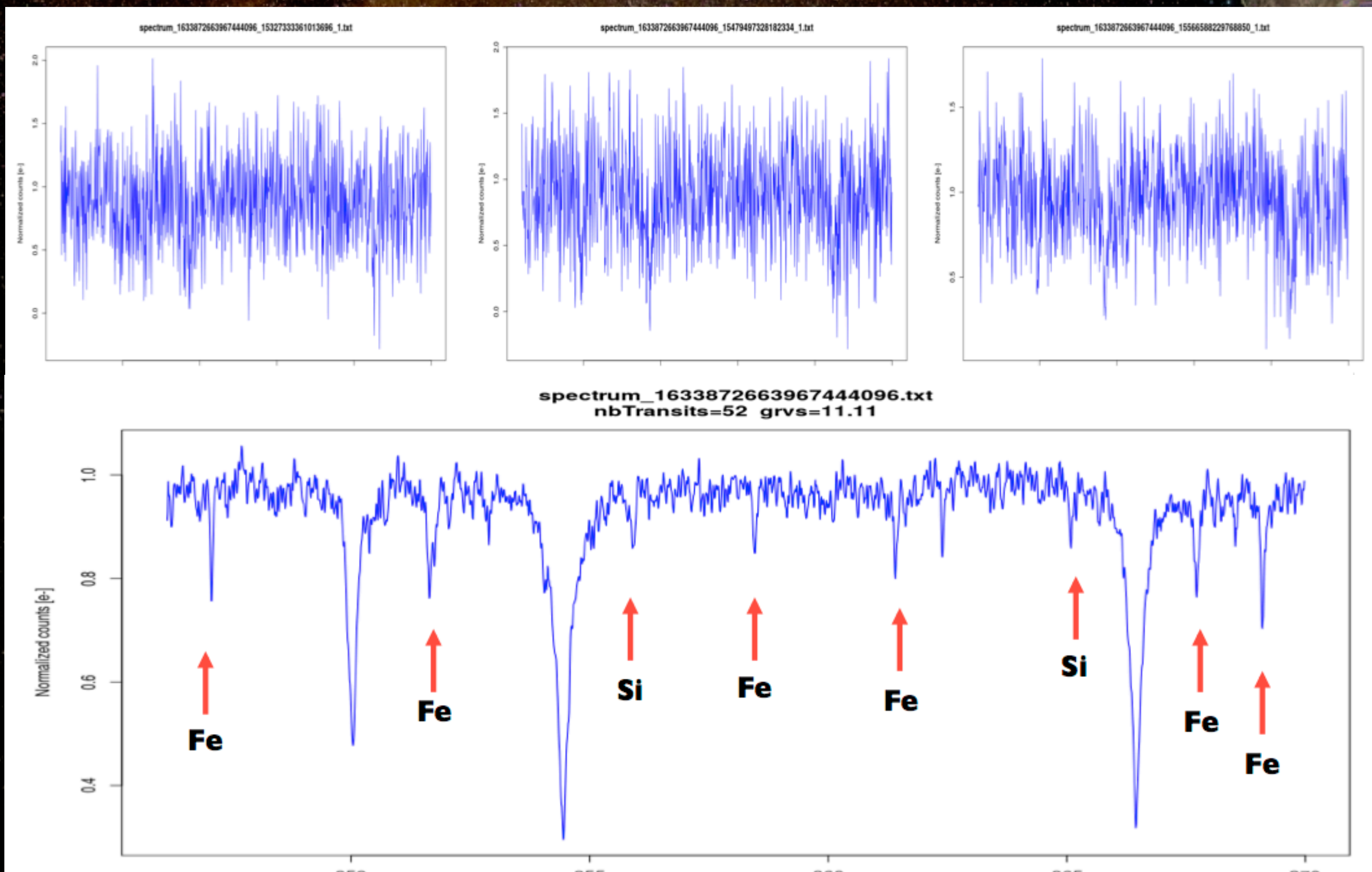
# CU6: Radial Velocity Spectrometer




Credit: D. Katz/CU6

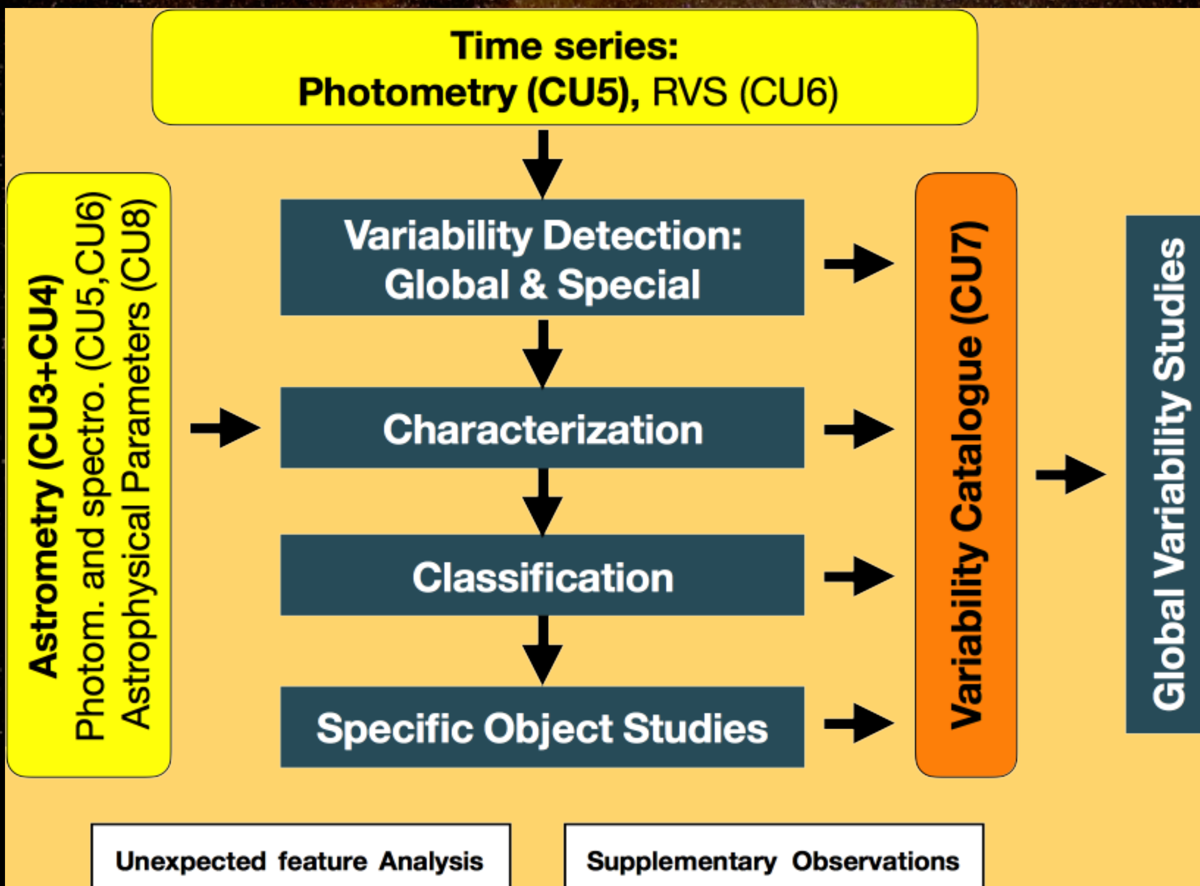
# CU6: Radial Velocity Spectrometer

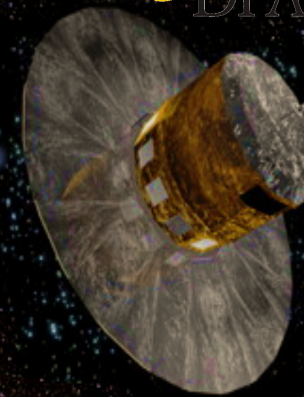
- Exemple: G ~ 11 and 52 spectra stacked

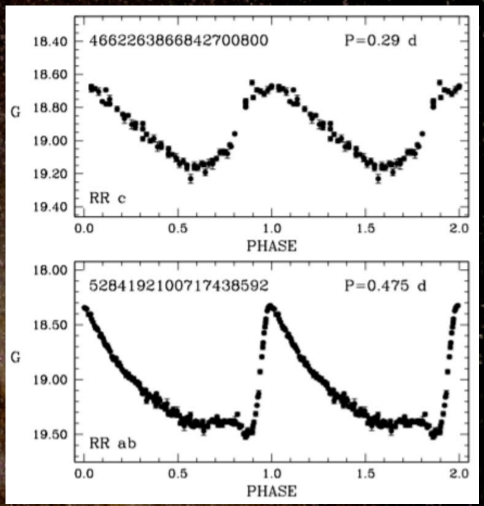



Credit : Marchal, Katz,Benson, Seabroke /CU6

# CU7: Variability

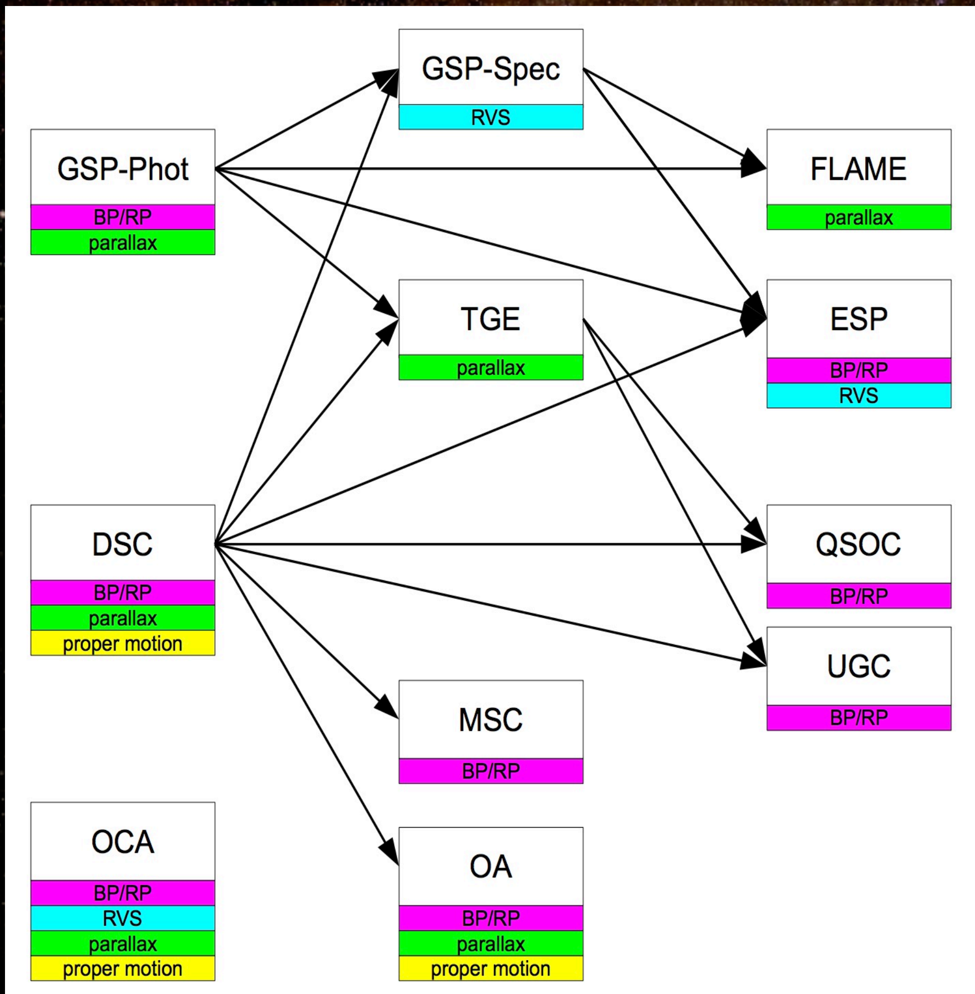


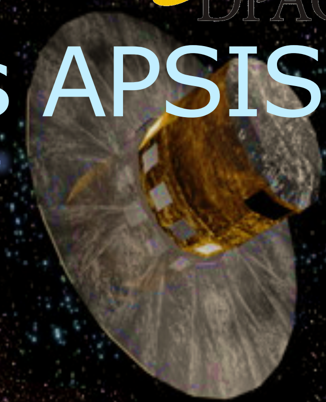






Credit: L. Eyer / CU7

# CU8: astrophysical parameters APSIS

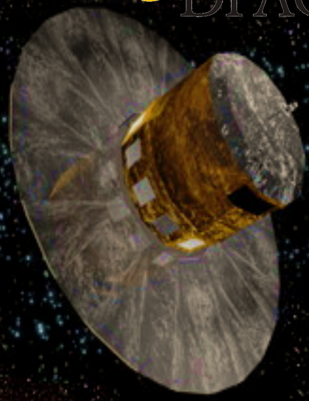


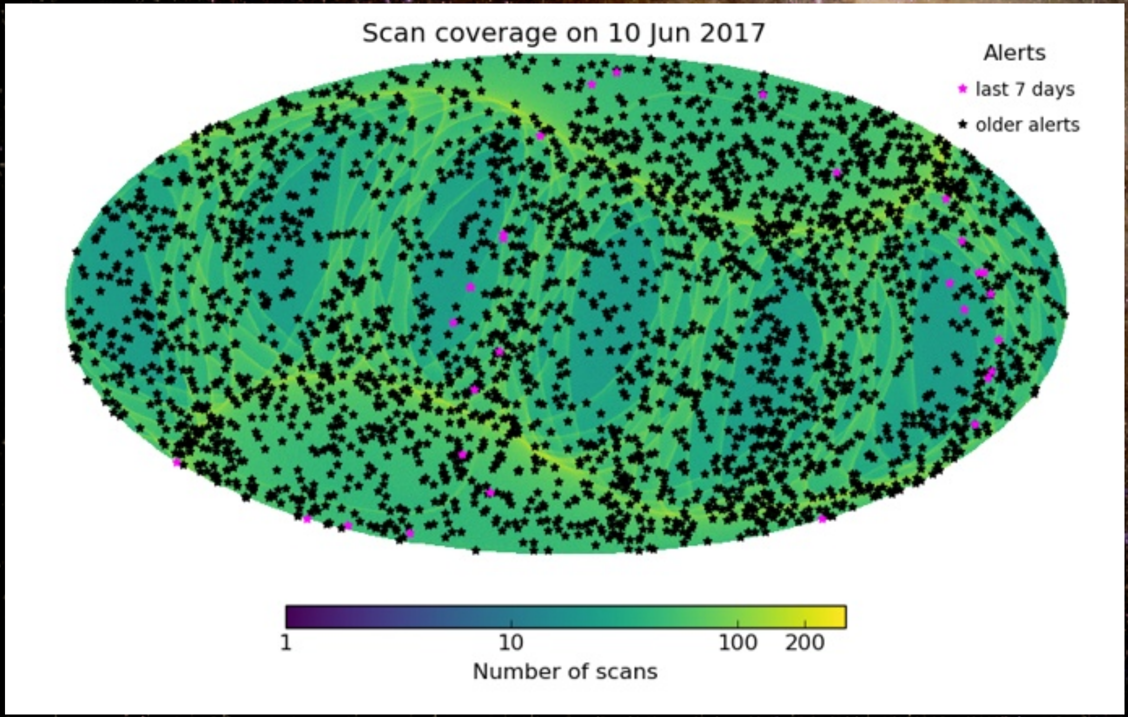
| Acronym | Name |
|---|---|
| DSC | Discrete Source Classifier |
| ESP | Extended Stellar Parametrizer: |
| -CS | ESP – Cool Stars |
| -ELS | ESP – Emission Line Stars |
| -HS | ESP – Hot Stars |
| -UCD | ESP – Ultra Cool Dwarfs |
| FLAME | Final Luminosity Age and Mass Estimato |
| GSP-Phot | Generalized Stellar Parametrizer – Photo |
| GSP-Spec | Generalized Stellar Parametrizer – Spectr |
| MSC | Multiple Star Classifier |
| OA | Outlier Analysis |
| OCA | Object Clustering Algorithm |
| QSOC | Quasar Classifier |
| TGE | Total Galactic Extinction |
| UGC | Unresolved Galaxy Classifier |

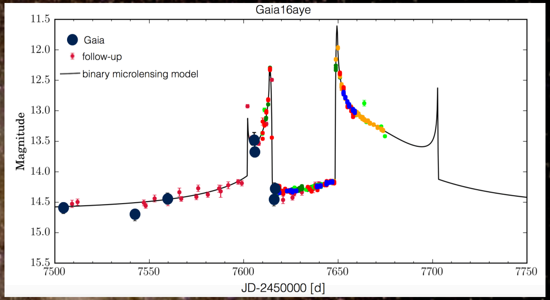
Credit: C. Bailer-Jones / CU8

# Science alerts: photometry



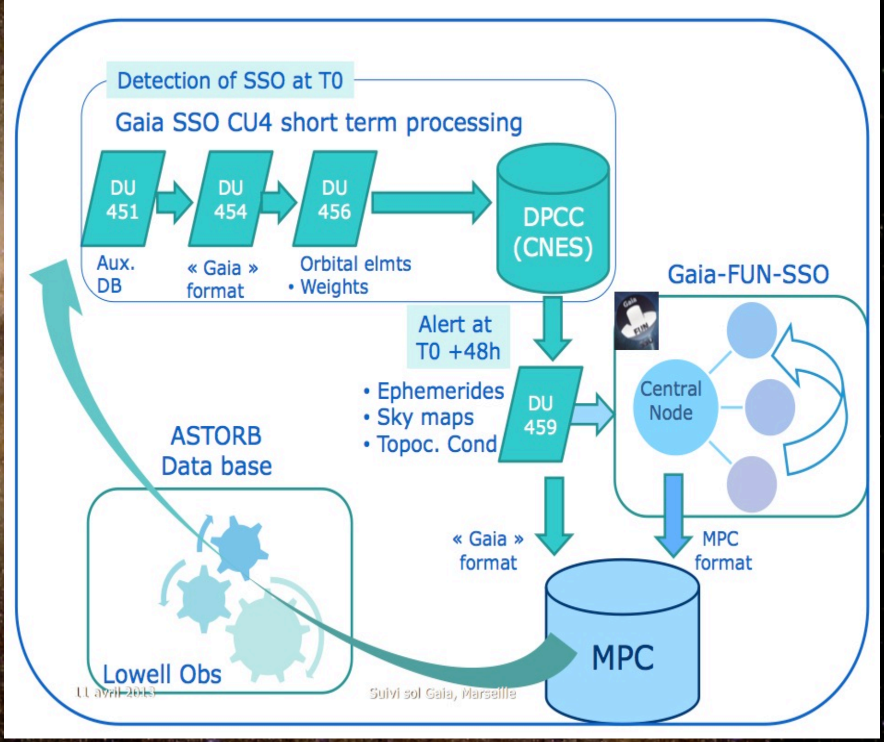
- As of June 10, 2,789 alerts
  - Among which 1,439 SN candidates
- Alert within 48h of observations
  - About 6 alerts/day last month
  - http://gsaweb.ast.cam.ac.uk/alerts/home



# Science alerts: Solar System Obj.

- ❑ Complementary Ground Based observations required
  - ❑ to confirm new detections
  - ❑ to avoid loss of moving objects
  - ❑ to improve orbits
- ❑ Gaia-FUN-SSO
  - ❑ 27 observing Sites
  - ❑ 80 observers
  - ❑ https://gaiafunsso.imcce.fr/index.php




Court: W.Thuillot, ObsParis

# CU9: DPAC Publication Unit

- CU9 is the one responsible for making all Gaia data available
  - except the Science Alerts
- ESA helps for the infrastructure and support
- CU9 tasks
  - Documentation
  - Architecture and technical development
  - Scientific validation
  - Operations
  - Education and outreach
  - Science enabling applications
  - Visualization
- For the DPAC workers: as defined in the SMP, the Gaia mission follows a `no special data rights' policy
  - Meaning: restricted access to the data, and not any science with it

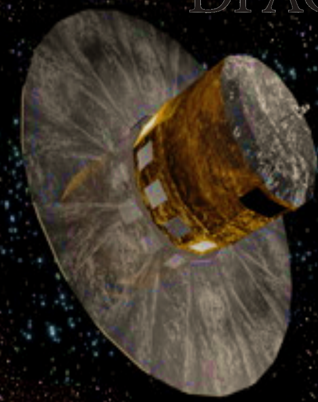
# CU9: scientific validation

- Transversal analysis of the catalogue content before delivery
  - Beside the verifications + validations done in each Coordination Unit
- Several rehearsals done with preliminary data
  - Leading to data improvement (and filtering at the end)
- About 55 tests implemented for DR1
  - Plus many interactive analysis

| DR1 validation \ W.P. | Internal tests | Models | External comparisons | Statistical analysis | Variability tests | Clusters |
|---|---|---|---|---|---|---|
| Completeness, large scale | ✗ | ✗ | ✗ | | | |
| Completeness, small scale | ✗ | | ✗ | | | ✗ |
| Astrometric accuracy | | ✗ | ✗ | ✗ | | ✗ |
| Astrometric precision | ✗ | | ✗ | ✗ | ✗ | ✗ |
| Photometric data | ✗ | | ✗ | ✗ | ✗ | ✗ |
| Variability data | | | | | ✗ | |

DR1

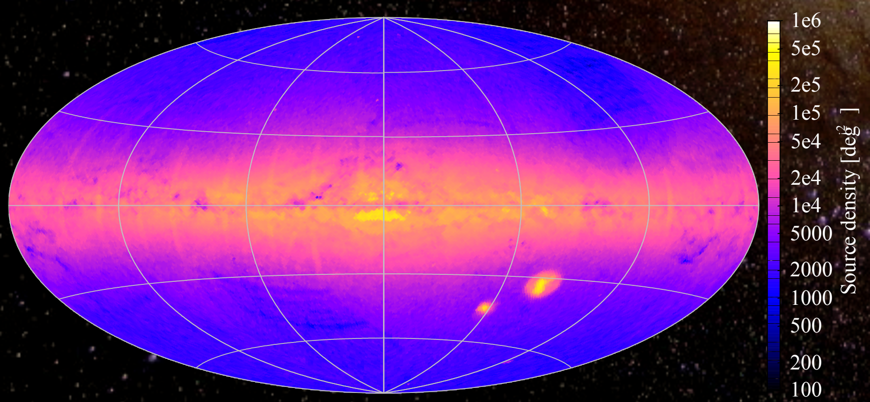
# Data Release 1



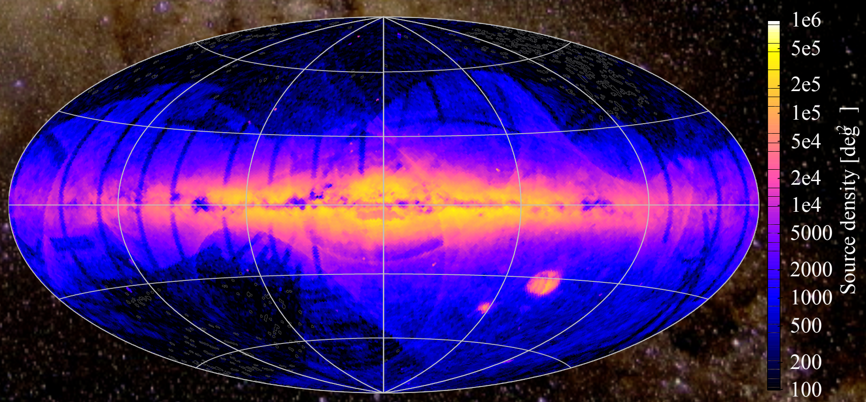
# 14 September 2016: Data Release 1

- Exactly 1000 days after launch
  - Processing of 14 months of data
  - Large access to the archive
    - More than 10.000 users within 24h, 60 TB downloaded over the first two weeks
- Articles
  - 12 DPAC DR1 Papers
  - 1st science publication 9h after DR1
  - June: > 200 science publications

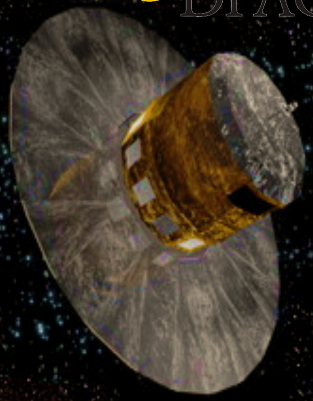
685 million sources matched to IGSL



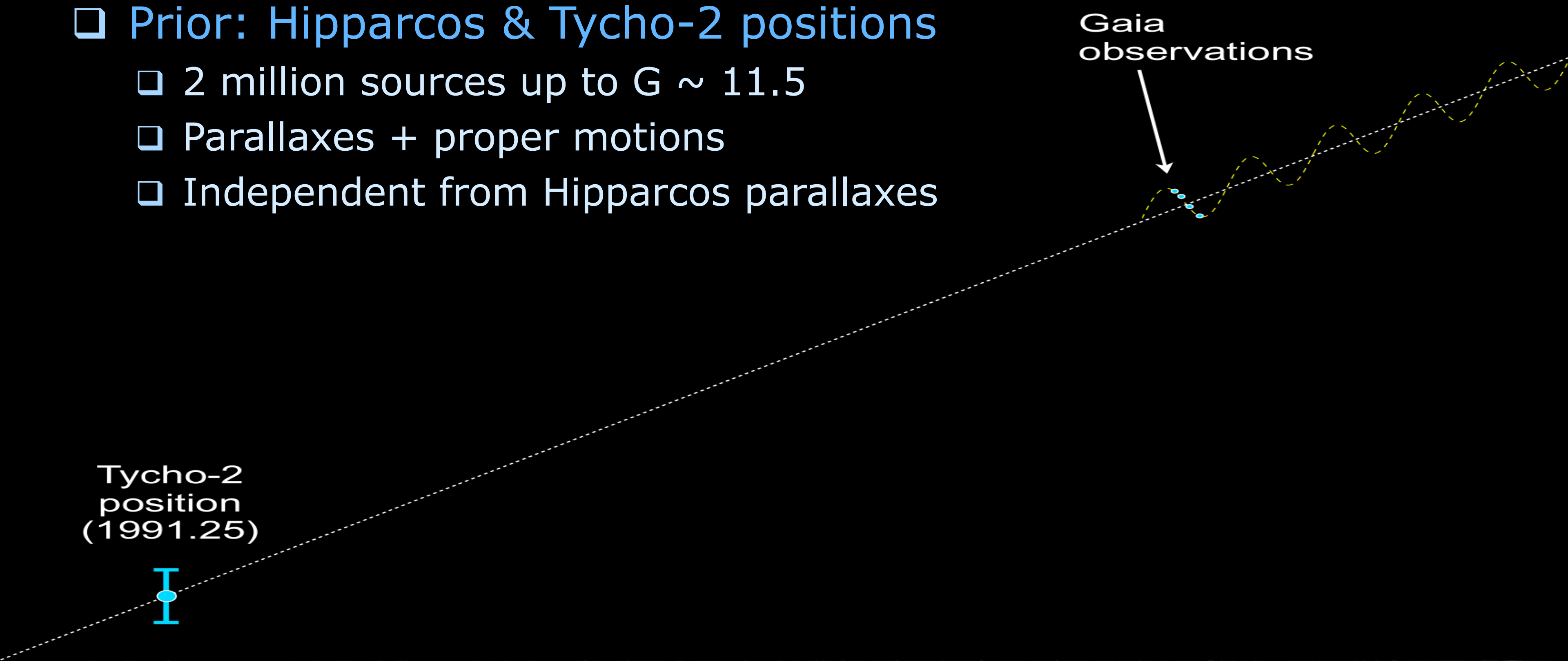

456 million new sources in Gaia DR1



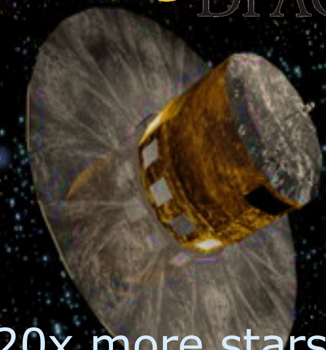
# The TGAS solution



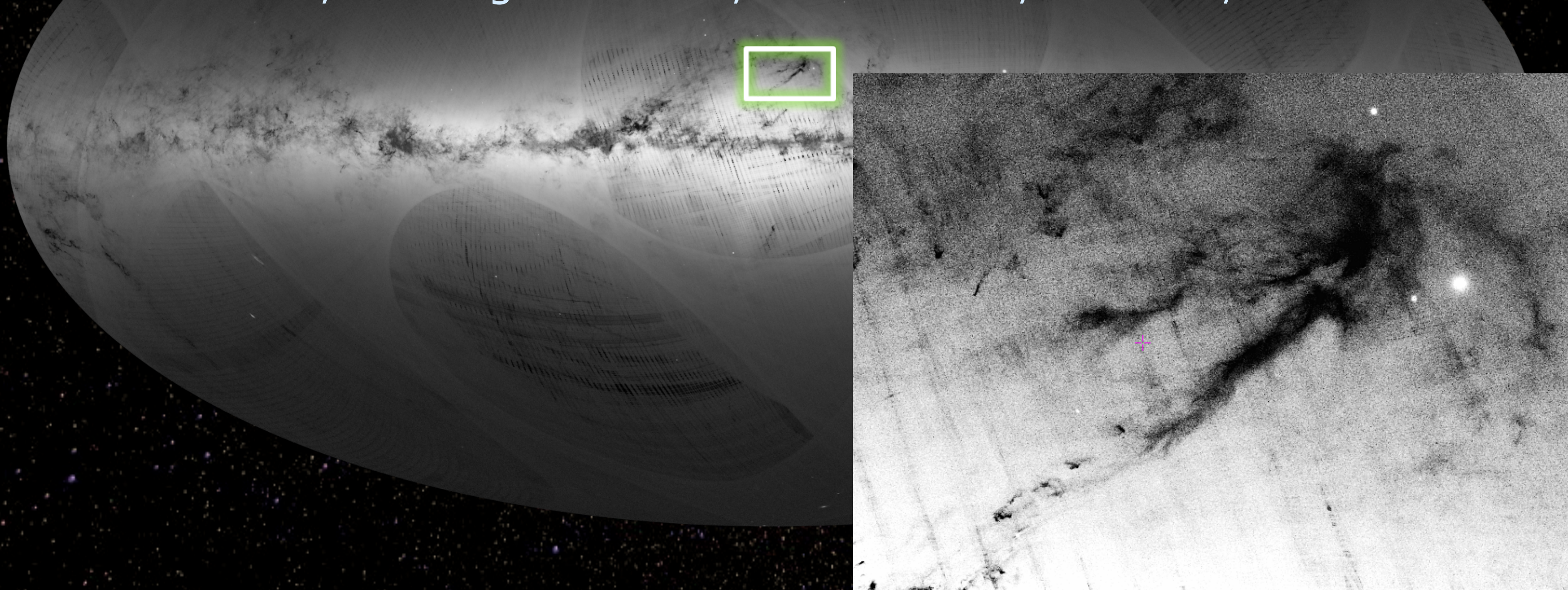
- 14 months data only
  - How to decouple parallaxes from proper motion
- Prior: Hipparcos & Tycho-2 positions
  - 2 million sources up to G ~ 11.5
  - Parallaxes + proper motions
  - Independent from Hipparcos parallaxes



# Catalogue content



- An atlas of 1.1 billion stars, up to G ~ 20.7, σ ~10 mas
- TGAS: 2 million parallaxes and proper motion
  - Parallaxes at least twice more precise than Hipparcos for 20x more stars
  - P.m. precision ~1 mas/yr for Tycho stars, 0.05 mas/yr for Hipparcos stars
- 3000 light curves of cepheids and RR Lyrae, σ ~ 0.003 -0.01 mag
- Pos. + mag. for 2000 ICRF QSOs, ICRF<0.1 mas, rotation<0.03mas $yr^{-1}$
- Documentation, interrogation tools, visualisation, X-match, etc.



Gaia DPAC

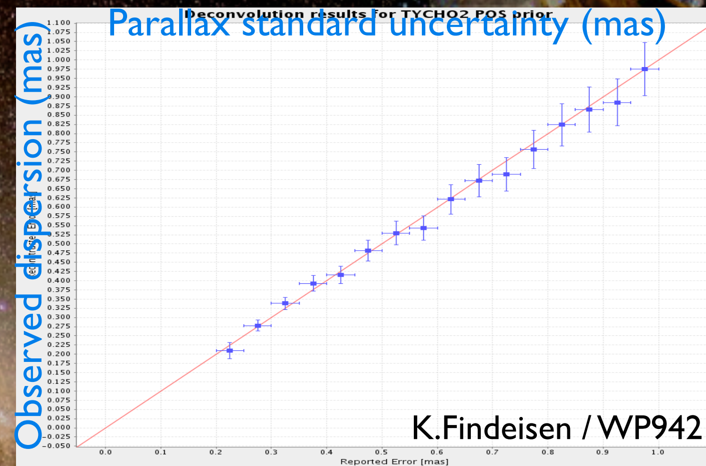
# Validation: parallax std. uncertainty

- Uncertainty slightly pessimistic

C. Babusiaux / WP944

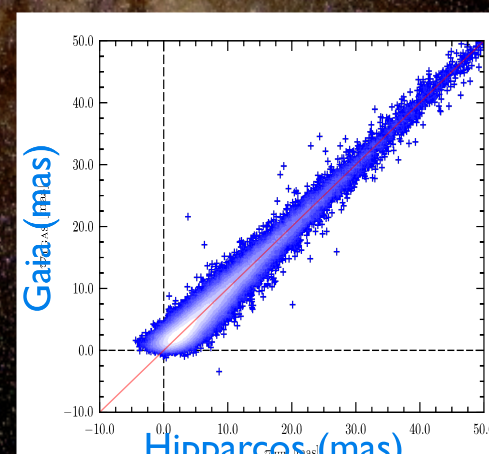
| Catalogue | $\varpi$ extra standard dev |
|---|---|
| Hipparcos | 0.58 ± 0.005 |
| VLBI | - |
| HST | 0.6 ± 0.2 |
| RECONS | -0.9 ± 0.5 |
| VLBI & HST & RECONS | 0.42 ± 0.13 |
| Cepheids | -0.18 ± 0.01 |
| RRLyrae | -0.16 ± 0.02 |
| Cepheids & RRLyrae | -0.17 ± 0.01 |
| RAVE | -0.06 ± 0.02 |
| APOGEE | -0.12 ± 0.01 |
| LAMOST | -0.17 ± 0.02 |
| PASTEL (J-K>0.3) | 0.1 ± 0.05 |
| APOKASC | -0.15 ± 0.01 |
| LMC | -0.14 ± 0.03 |
| SMC | -0.09 ± 0.09 |
| ICRF2 QSO auxiliary solution | -0.17 ± 0.01 |

Extra std. dev on parallax uncertainties (mas)

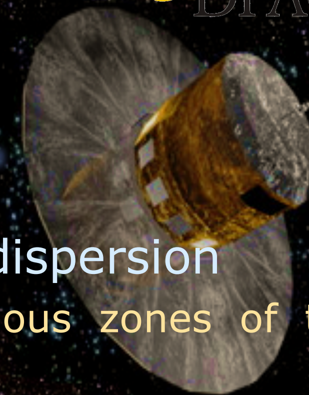



Deconvolution of the negative parallax tail for Tycho stars
extra dispersion: −0.11 ± 0.01 mas

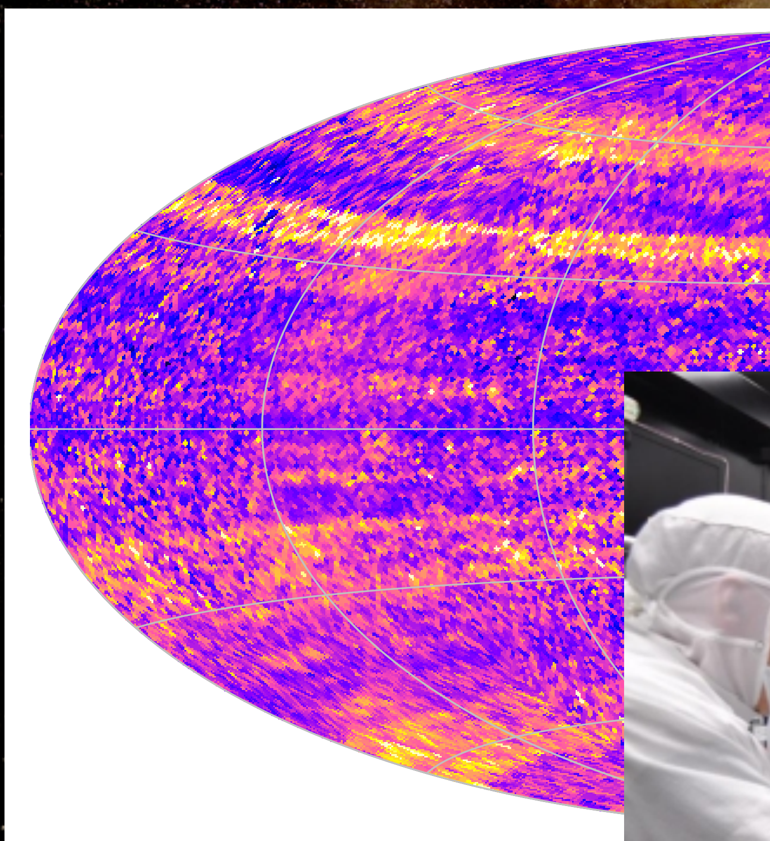
Width of Gaia parallaxes smaller than Hipparcos



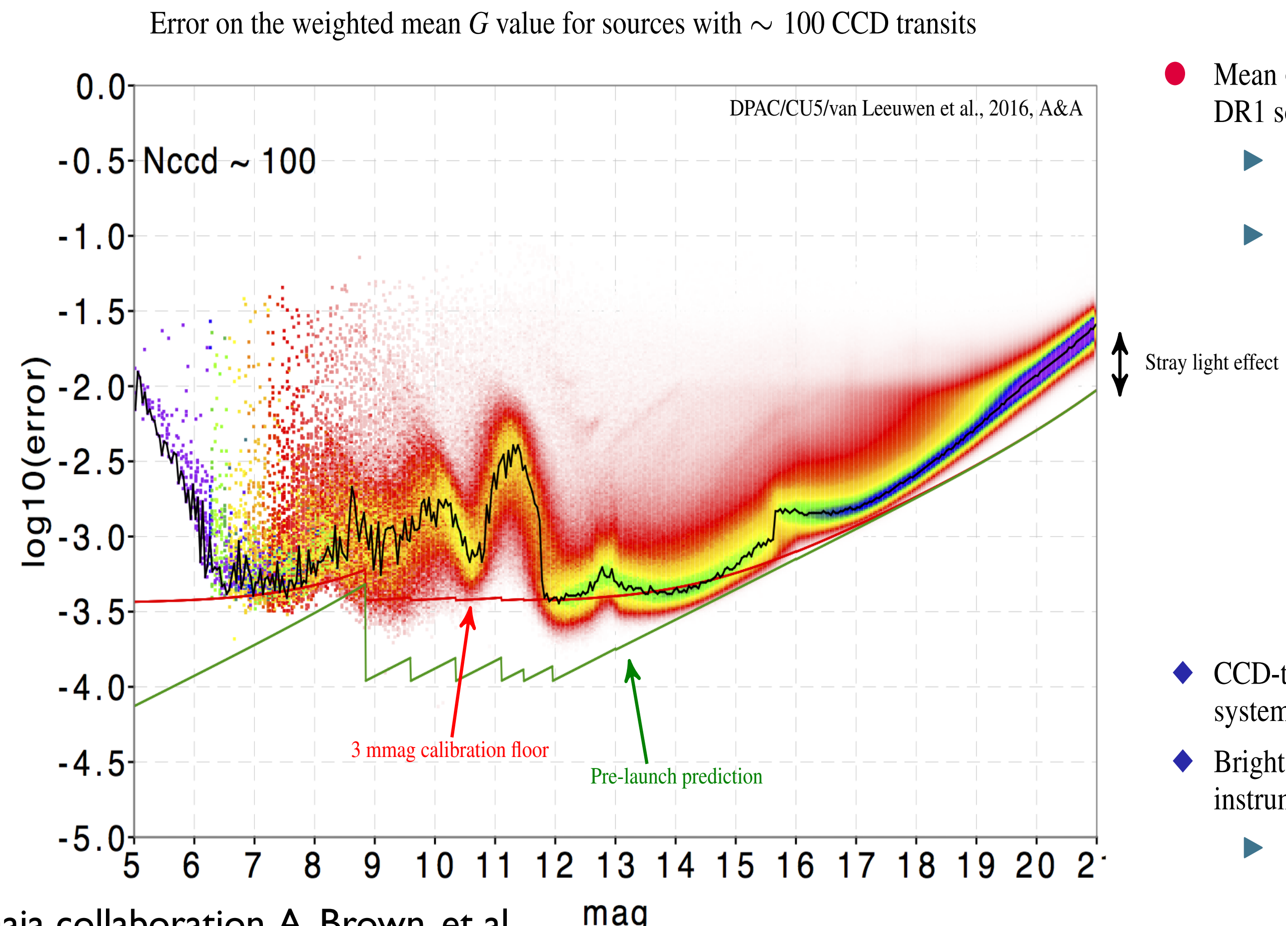
# Compared to Tycho-2 p.m.



- Differences Gaia-Tycho-2 contain a 1.8 mas yr $^{-1}$ extra dispersion
  - Problems in Tycho-2 p.m. probably originate from the various zones of the Astrographic Catalogue (e.g. +55° to +64° for Vatican)



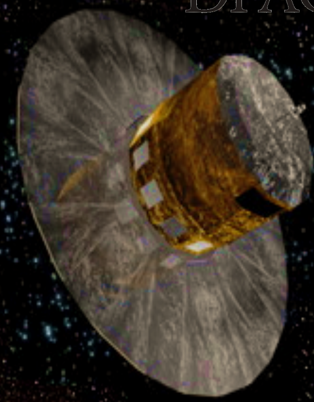
# Photometric precision




Gaia collaboration, A. Brown, et al.

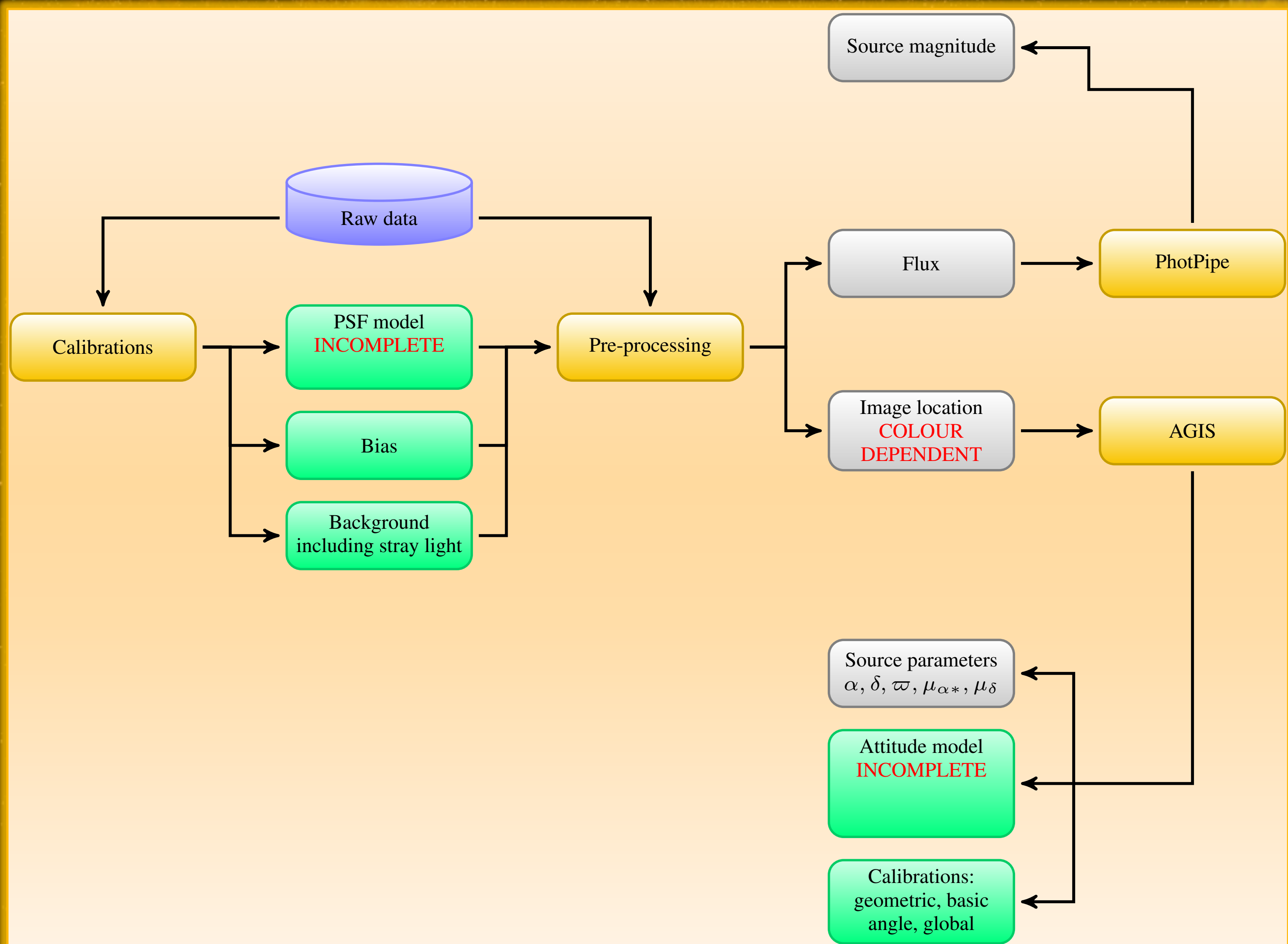
- Mean *G*-band fluxes and errors for all Gaia DR1 sources
  - *G* magnitudes in VEGAMAG, zeropoints for AB
  - No pass-band calibrations, transformations to other systems to be provided

- CCD-transit *G*-band calibration systematics at the ∼ 3 mmag level
- Bright end features related to on-board instrument configuration changes
  - will be calibrated out in future releases



# Known limits of first data release

# Trade-offs for DR1



# Trade-offs for DR1

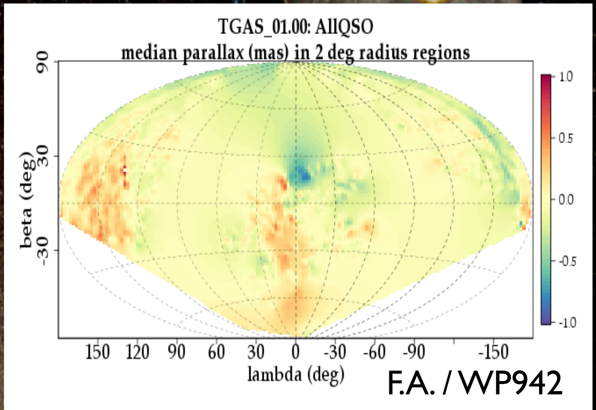
- Lessons learned from DR1 validation
  - DR1: a nice `proof of concept' for Gaia and the work done by DPAC
- Early publishing asked however for several trade-offs
  - Small timespan (14 months)
    - Bright stars from Hipparcos/Tycho used as priors
  - Difficult calibration model
    - Bright stars are the worst cases (saturated)
  - Attitude perturbed model and other unmodelled effects (μclanks)
    - Increasing the level of systematics in astrometry
  - Cross-matching inadequacies
    - Coming from ground-based catalogues, spurious sources
  - Source model: all stars assumed single
    - Bad goodness of fit, incompleteness effects
  - All this leads to spatially correlated systematics
- Given all this, DR1 is really
  - The best that could be achieved during the time available
  - Could be done only thanks to the strong DPAC competence and motivation

# Validation: parallax accuracy

C. Babusiaux / WP944

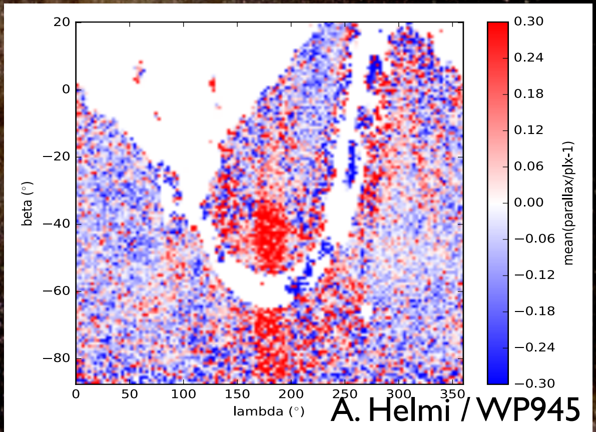
| Catalogue | Outliers | ϖ bias |
|---|---|---|
| Hipparcos | 0.09% | -0.094 ± 0.004 |
| VLBI | 0 / 9 | 0.083 ± 0.12 |
| HST | 2 / 19 | -0.11 ± 0.19 |
| RECONS | 0 / 13 | -1.04 ± 0.58 |
| VLBI & HST & RECONS | 2 / 41 | -0.08 ± 0.12 |
| Cepheids | 0 / 207 | -0.014 ± 0.014 |
| RRLyrae | 0 / 130 | -0.07 ± 0.02 |
| Cepheids & RRLyrae | 0 / 337 | -0.034 ± 0.012 |
| RAVE | 47 / 5144 | 0.07 ± 0.005 |
| APOGEE | 0 / 2505 | -0.06 ± 0.006 |
| LAMOST | 6 / 317 | -0.01 ± 0.02 |
| PASTEL (J-K>0.3) | 1 / 218 | 0.05 ± 0.02 |
| APOKASC | 0 / 969 | -0.07 ± 0.009 |
| LMC | 2 / 142 | 0.11 ± 0.02 |
| SMC | 0 / 58 | -0.12 ± 0.05 |
| ICRF2 QSO auxiliary solution | 1 / 2060 | -0.046 ± 0.01 |

Fully consistent with QSO results: $-0.036 \pm 0.002$ mas

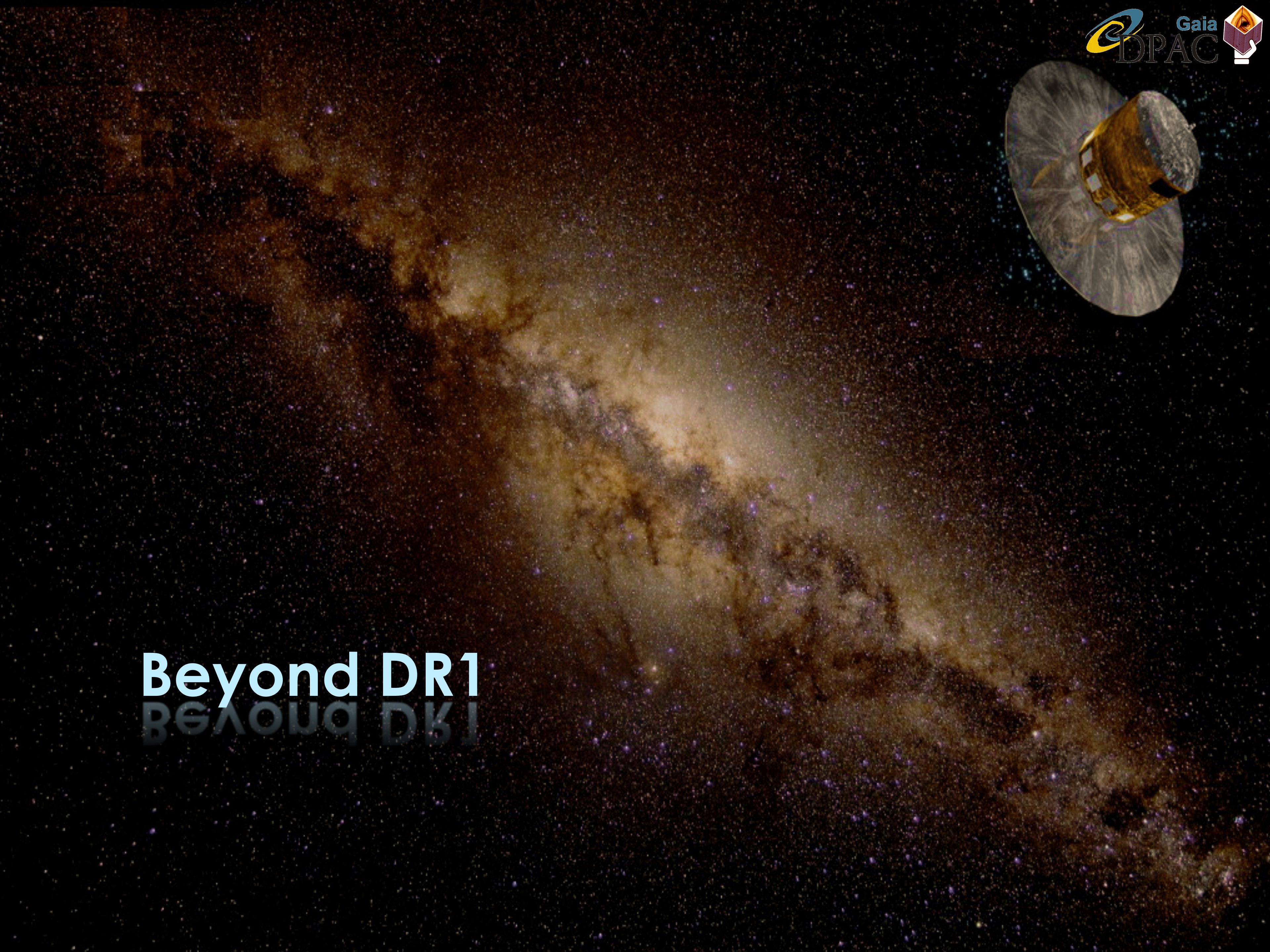


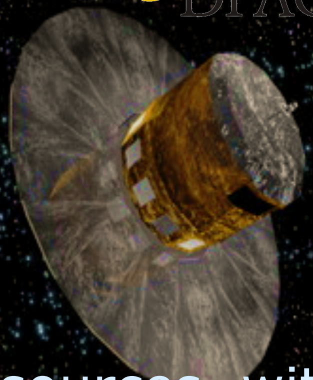

QSO parallaxes

RAVE (ϖ ratio -1)



# Beyond DR1





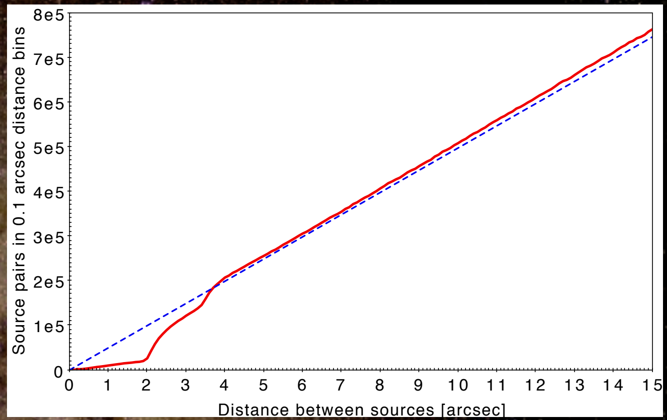
# DR2 schedule

- April 2018 (firm)
- Content
  - Five-parameter astrometric solutions for all sources with acceptable formal standard errors > 1 billion
    - Positions only for the other
  - *G* and integrated $G_{BP}$ and $G_{RP}$ and photometric fluxes and magnitudes for all sources
  - Median radial velocities for $G_{RVS}$ < 12 sources
  - Estimates of $T_{eff}$ and where possible A (0, G, BP, RP) based on integrated photometry for G < 17 sources
  - Photometric data for a sample of variables stars
  - Epoch photometry for a preselected list of 10 000 asteroids

# Many improvements expected for DR2

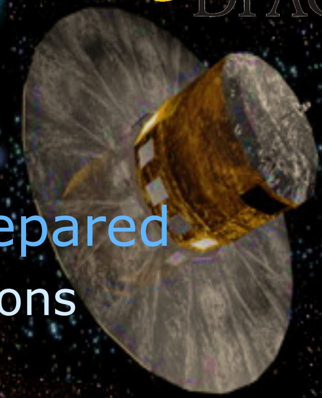
- ❑ Improvements
  - ❑ 5p astrometry for most stars
  - ❑ Better astrometric precision
  - ❑ Systematics strongly reduced
  - ❑ Small-scale incompleteness reduced
  - ❑ Source list has been cleaned up

- ❑ Better than yesterday but not as good than tomorrow!
  - ❑ Precision, accuracy and data content will improve at each Data Release




Number of sources vs distance between sources (arcsec) in high density field for DR1 (credit C. Fabricius). Will much improve at later releases

# A few remarks

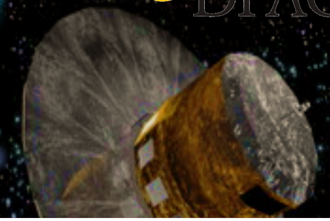


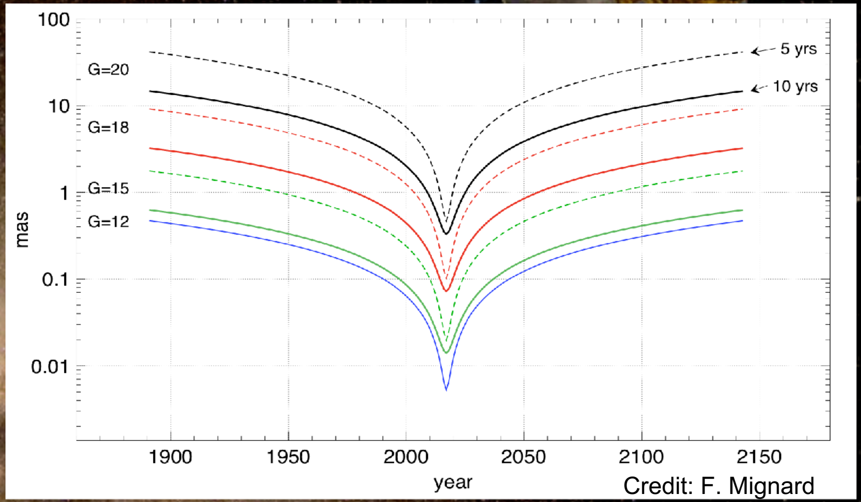
- Organisation: most of the work has been well prepared
  - Simulating data is really needed, years before operations
  - Many Operations Rehearsals done (fortunately)
  - Usefulness of a Project Office with competent people (not managers)
  - Usefulness of a Coord. Unit devoted in advance to architecture dev.
- Despite all this, reality is always more complex
  - Disordered, corrupted telemetry, unavoidable gaps, S/W changes
    - One million lines managed at SOC (mostly java, could have been worse).
  - Development phase does not stop with production phase… manpower
  - Complexity, interdependency between data delays the process
  - Difficulty of X-match: merging/splitting sources: effects downstream
  - Many more challenges to come with multiple stars… ☺
- Importance of the interchange file format
  - Should be resilient to small data model changes & have analysis tools
  - Data models are very important

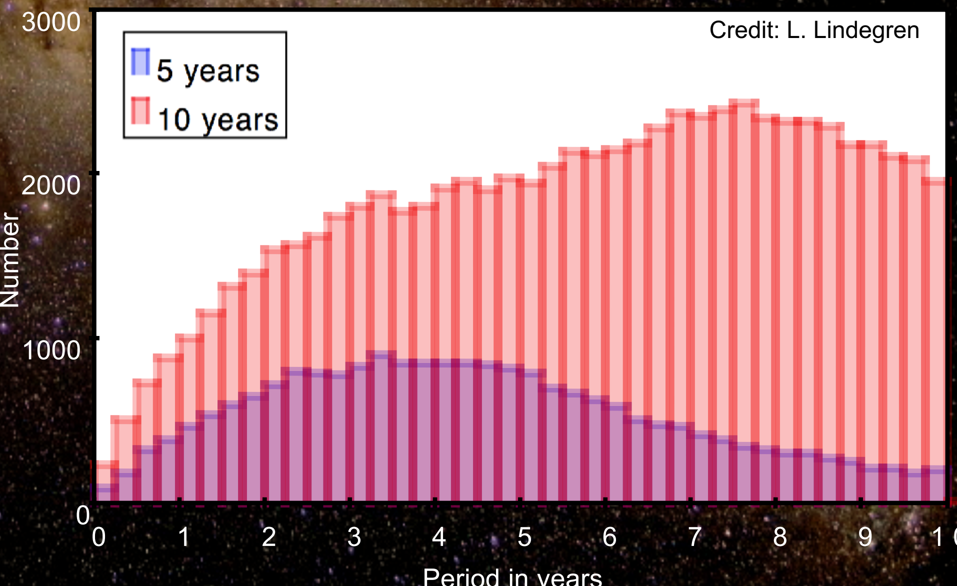
# Mission extension

- The scientific case
  - Will slow down the degradation of the reference frame
  - Precision improvement
    - Proper motion ∝ $t^{-3/2}$
    - Positions, parallaxes ∝ $t^{-1/2}$
    - Also for photometry, radial velocities
  - Longer orbital periods reachable
    - Binaries & exoplanets
  - Variability studies improved

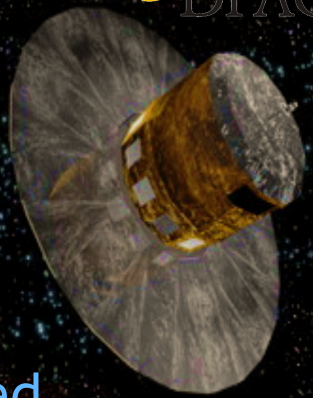
- In practice
  - Nominal end mid-2019
  - Gas, sunshield: end-2023±1yr
  - Science case prepared for 5yr extension, but ESA cycle is 2+2
  - Ministerial level decision on ESA science budget next November





# Conclusion



- Gaia is performing globally well, mostly as planned
  - More stars than initially planned: bright~G=3, faint end G > 20.7
  - Small problems with reduced scientific impacts
  - Complexity larger than expected, slightly delaying the whole process
  - However DR1 saw the light and DR2 will be present in 10 months
- If Gaia data is useful to you, then please acknowledge the DPAC+ESA work in your articles !
  - This will help ensuring national funding + secure mission extension

*ʿAbd al-Raḥmân ibn ʿOmar al-Ṣoûfî (903-986), كتاب الكواكب الثابتة“ book of fixed stars”.*