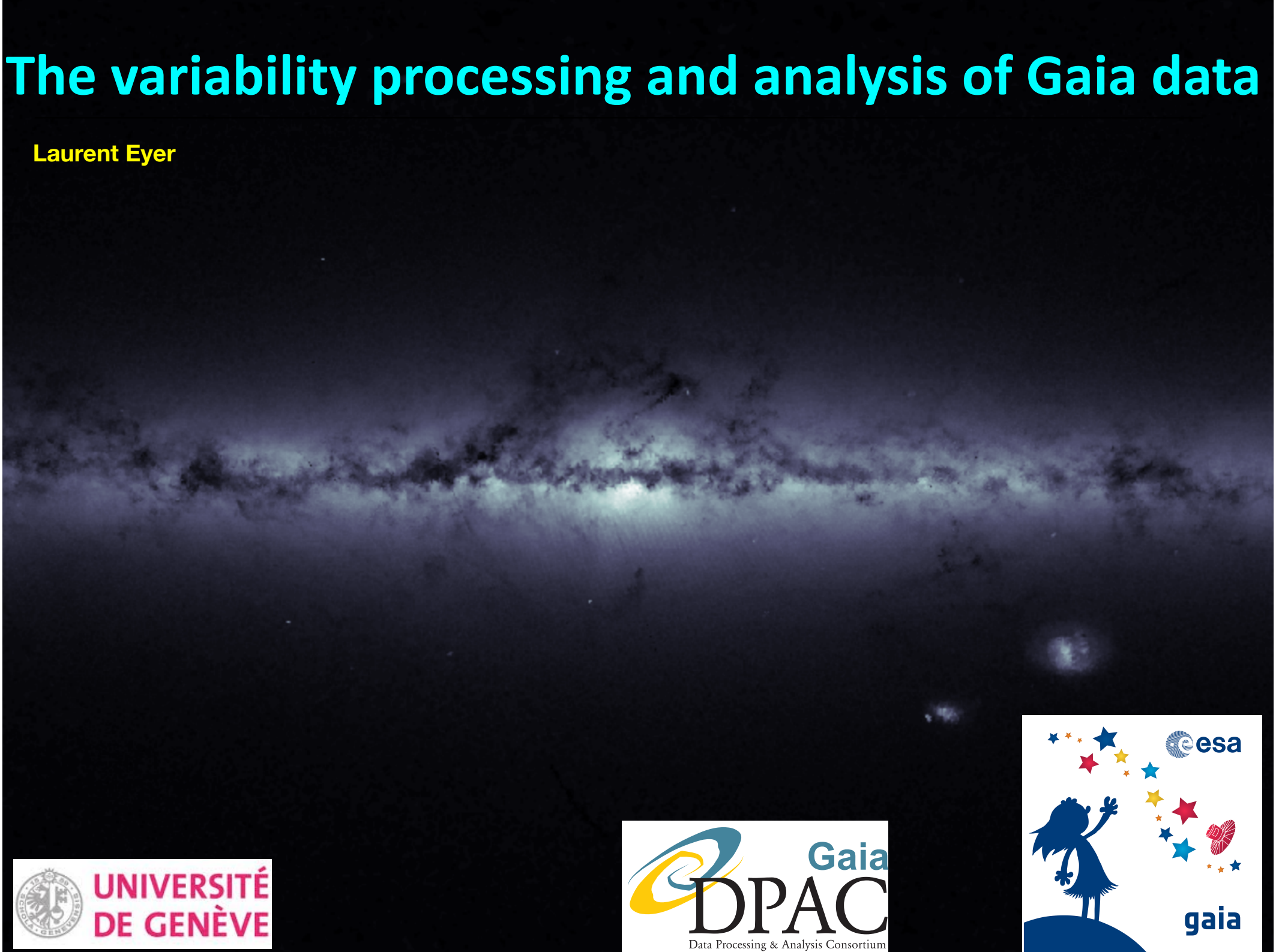
The variability processing and analysis of Gaia data

Laurent Eyer



UNIVERSITÉ DE GENÈVE

Gaia DPAC Data Processing & Analysis Consortium

esa gaia

The variability processing and analysis of Gaia data

Laurent Eyer[1]

Nami Mowlavi[1], Dafydd W. Evans[2], Berry Holl[1], Lorenzo Rimoldini[1], Alessandro Lanzafame[3], Leanne Guy[1], Shay Zucker[4], Brandon Tingley[5], Isabelle Lecoeur-Taïbi[1], Maroussia Roelens[1], Jan Cuypers[6], Joris De Ridder[7], Sara Regibo[7], Manuel López[8], Jonas Debosscher[7], Maria Süveges[1], Luis Sarro[9], Gisella Clementini[10], Silvio Leccia[11], Vincenzo Ripepi[11], Fabio Barblan[1], André Moitinho[12], Krzysztof Nienartowicz[1], Diego Ordoñez-Blanco[1], Jonathan Charnas[1], Grégory Jévardat de Fombelle[1], and many other distinguished colleagues

[1]Department of Astronomy, University of Geneva, Versoix, Switzerland
[2]Institute of Astronomy, University of Cambridge, Cambridge, United Kingdom
[3]Dipartimento di Fisica e Astronomia, Universita di Catania, Catania, Italy
[4]Department of Geosciences, Tel Aviv University, Tel Aviv, Israel
[5]Institut for Fysik og Astronomi, Aarhus Universitet, Aarhus, Denmark
[6]Royal Observatory of Belgium, Brussels, Belgium
[7]Instituut voor Sterrenkunde, KU Leuven, Leuven, Belgium
[8] Centro de Astrobiologia, Departamento de Astrofisica, Villanueva de la Canada, Spain
[9]Departamento de Inteligencia Artificial, UNED, Madrid, SpainI
[10]INAF Osservatorio Astronomico di Bologna, Bologna, Italy
[11]INAF-Osservatorio Astronomico di Capodimonte, Napoli, Italy
[12]Faculdade de Ciencias de Universidade de Lisboa, Lisboa, Portugal

LSST
Zira Hotel, Belgrade, Serbia
Friday June 24, 2016

UNIVERSITÉ DE GENÈVE

Gaia DPAC Data Processing & Analysis Consortium

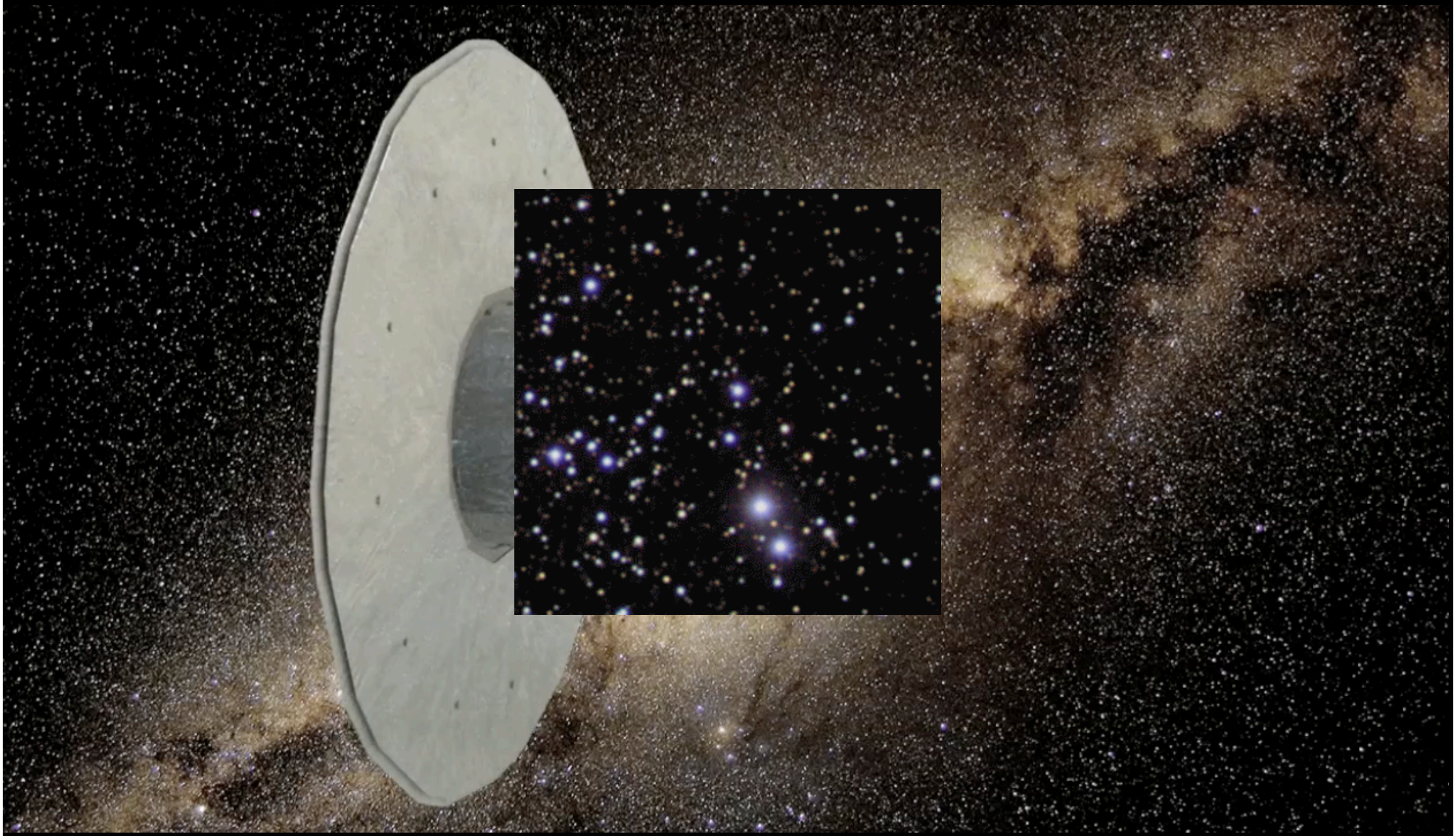
esa gaia

The Gaia mission

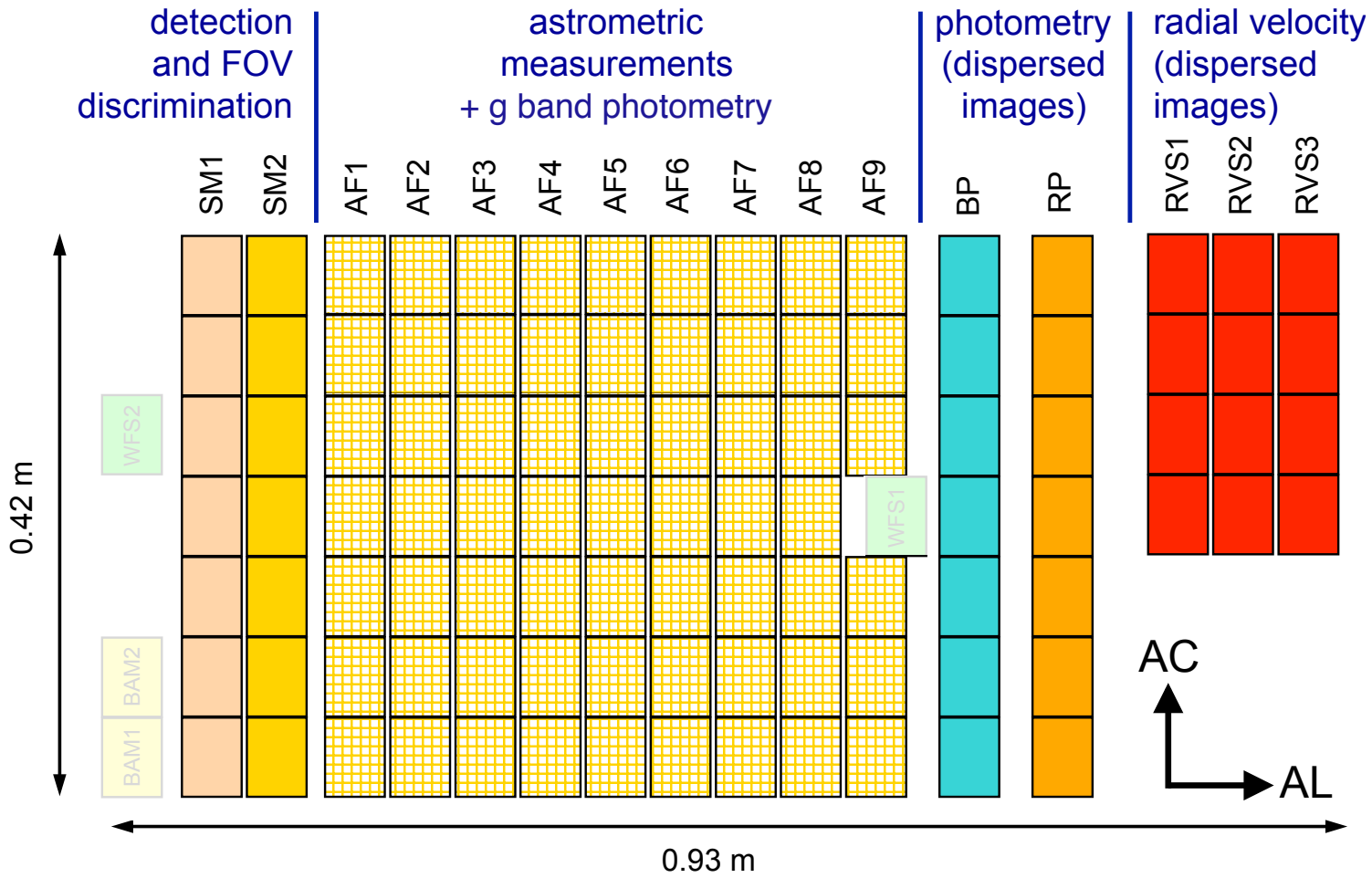
- Cornerstone mission of the European Space Agency
- Observations of all the objets brighter than a certain limit V ~ 20 (**>1 billion objects**)
- Measurement of:
  - positions (astrometry)
  - brightness, colors (spectro-photometry)
  - radial velocities (spectroscopy)
- Launch (Soyuz rocket): December 19 2013
- Length: 5 years (70 times all sky) (+5 year??)
- Final Catalogue: 2021-2022 (+ 5??)

Gaia: a spinning and scanning spacecraft

Gaia: a spinning and scanning spacecraft



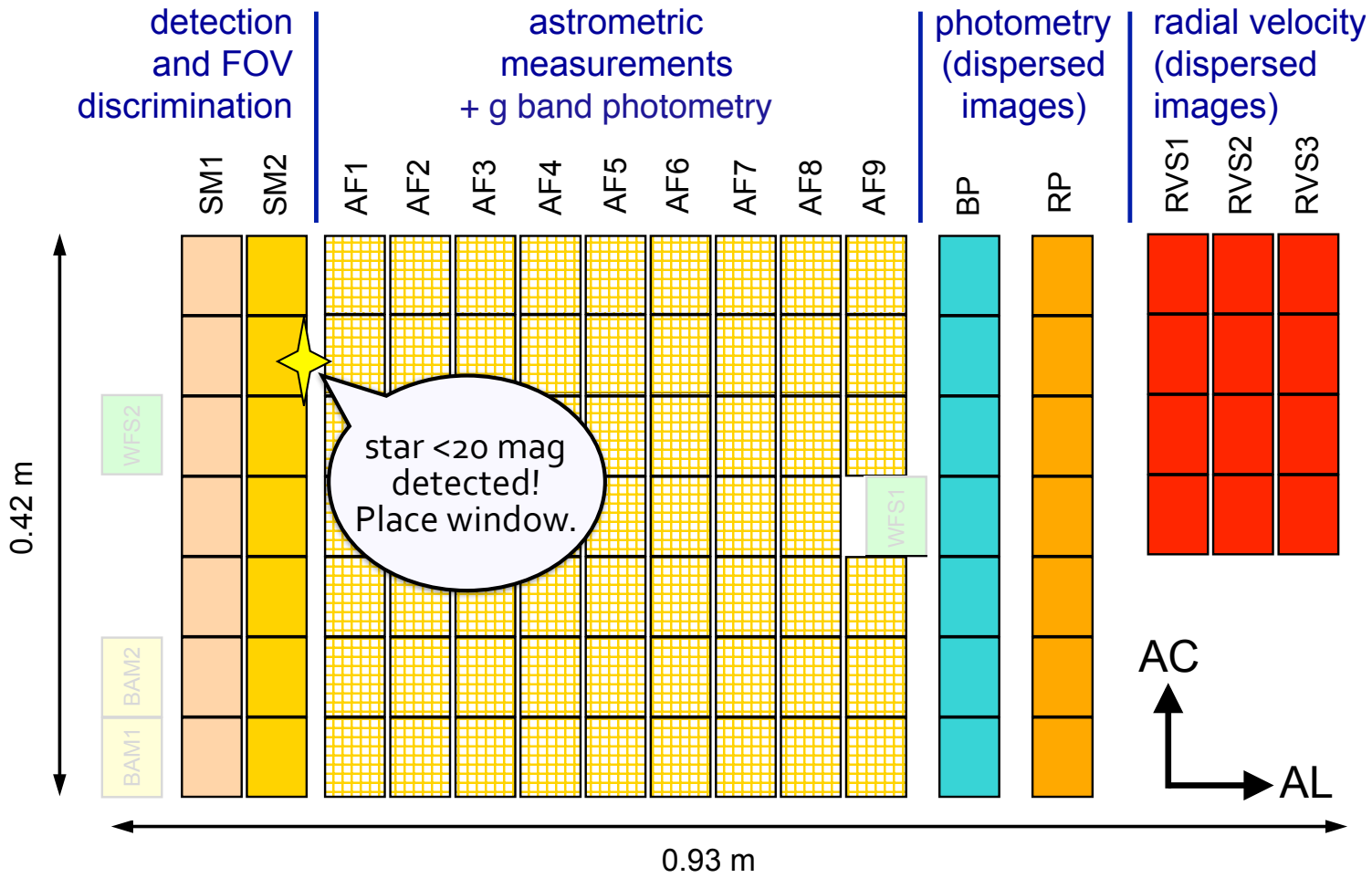
Focal plane



Gaia FoV: 0.7 deg x 0.7 deg
pixel: 0.059”(AL) x 0.177”(AC)

Courtesy of Berry Holl

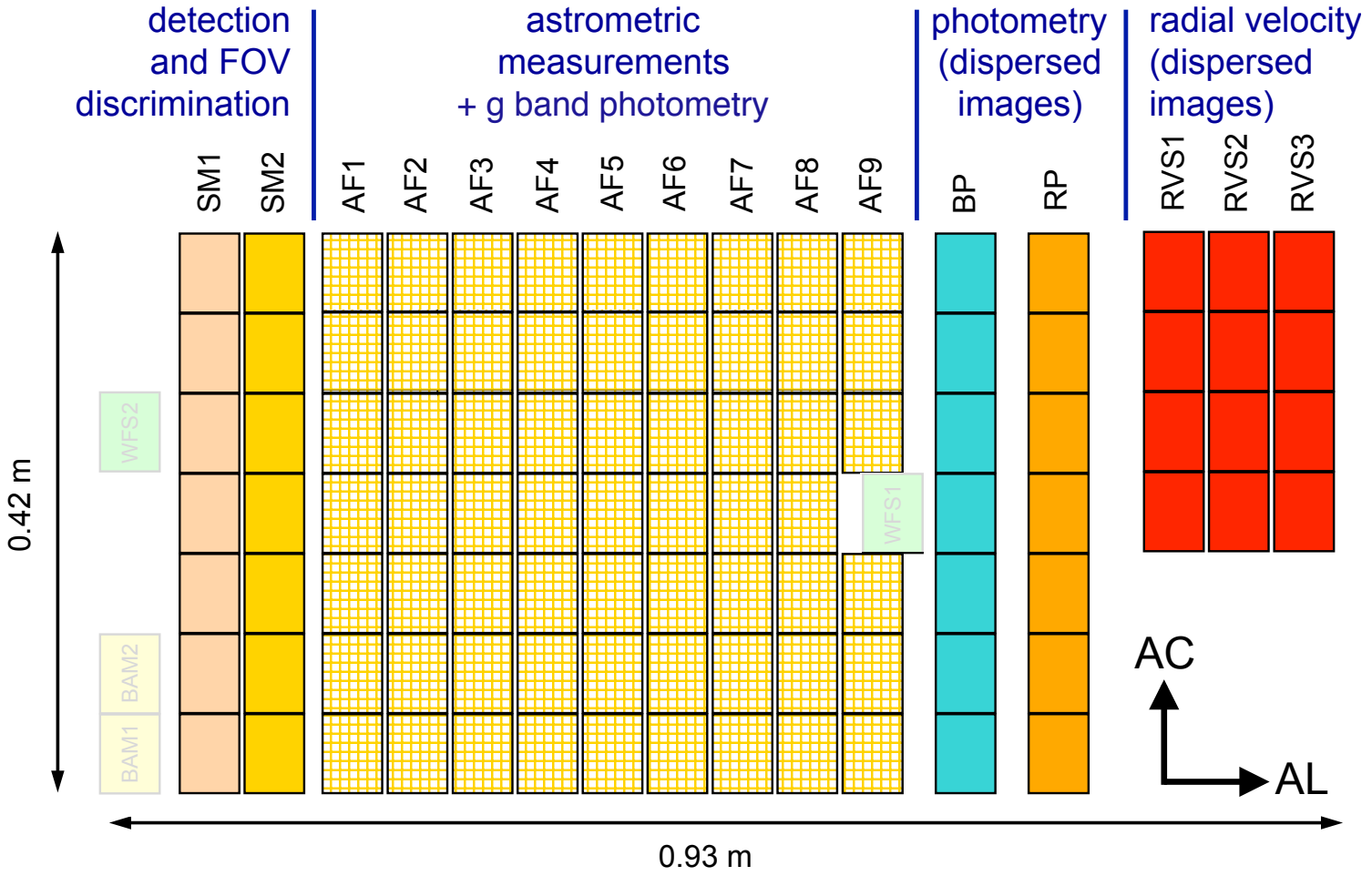
Focal plane



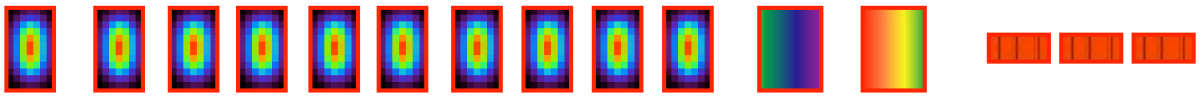
Gaia FoV: 0.7 deg x 0.7 deg
pixel: 0.059”(AL) x 0.177”(AC)

Courtesy of Berry Holl

Focal plane



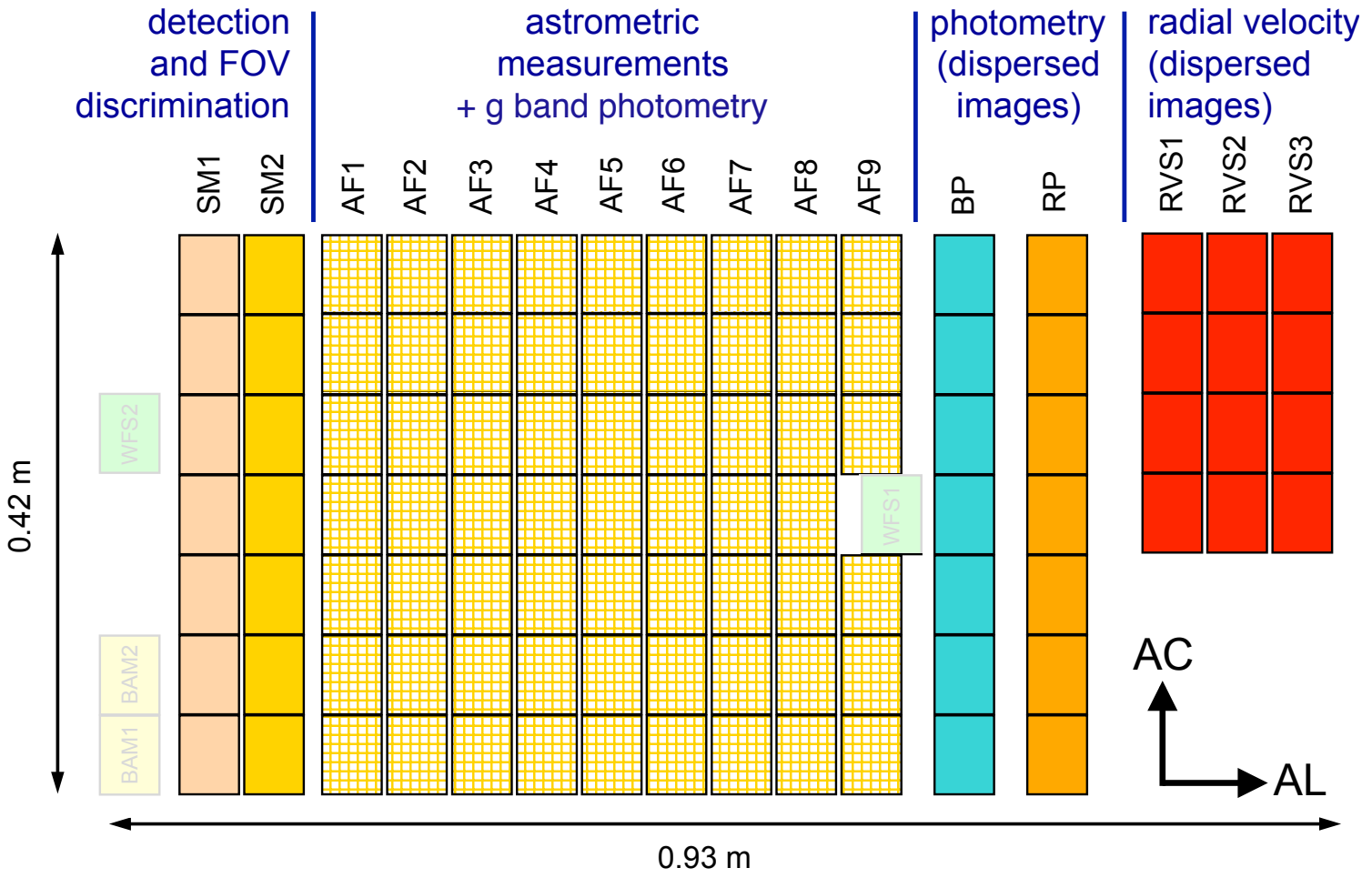
observed windows (not to scale):

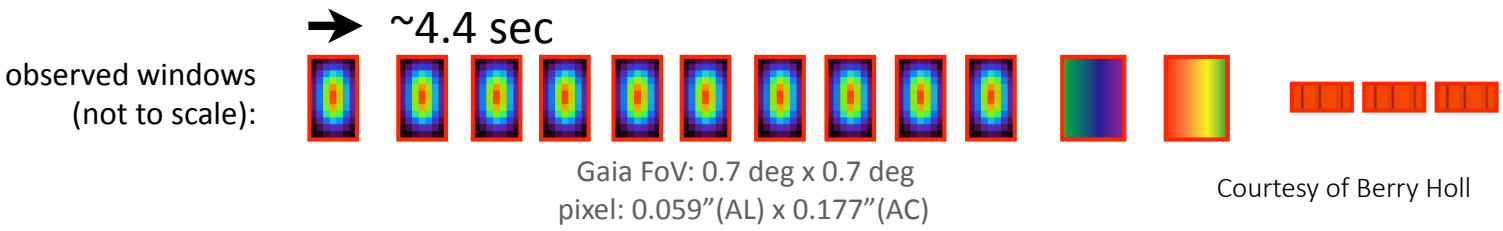


Gaia FoV: 0.7 deg x 0.7 deg
pixel: 0.059”(AL) x 0.177”(AC)

Courtesy of Berry Holl

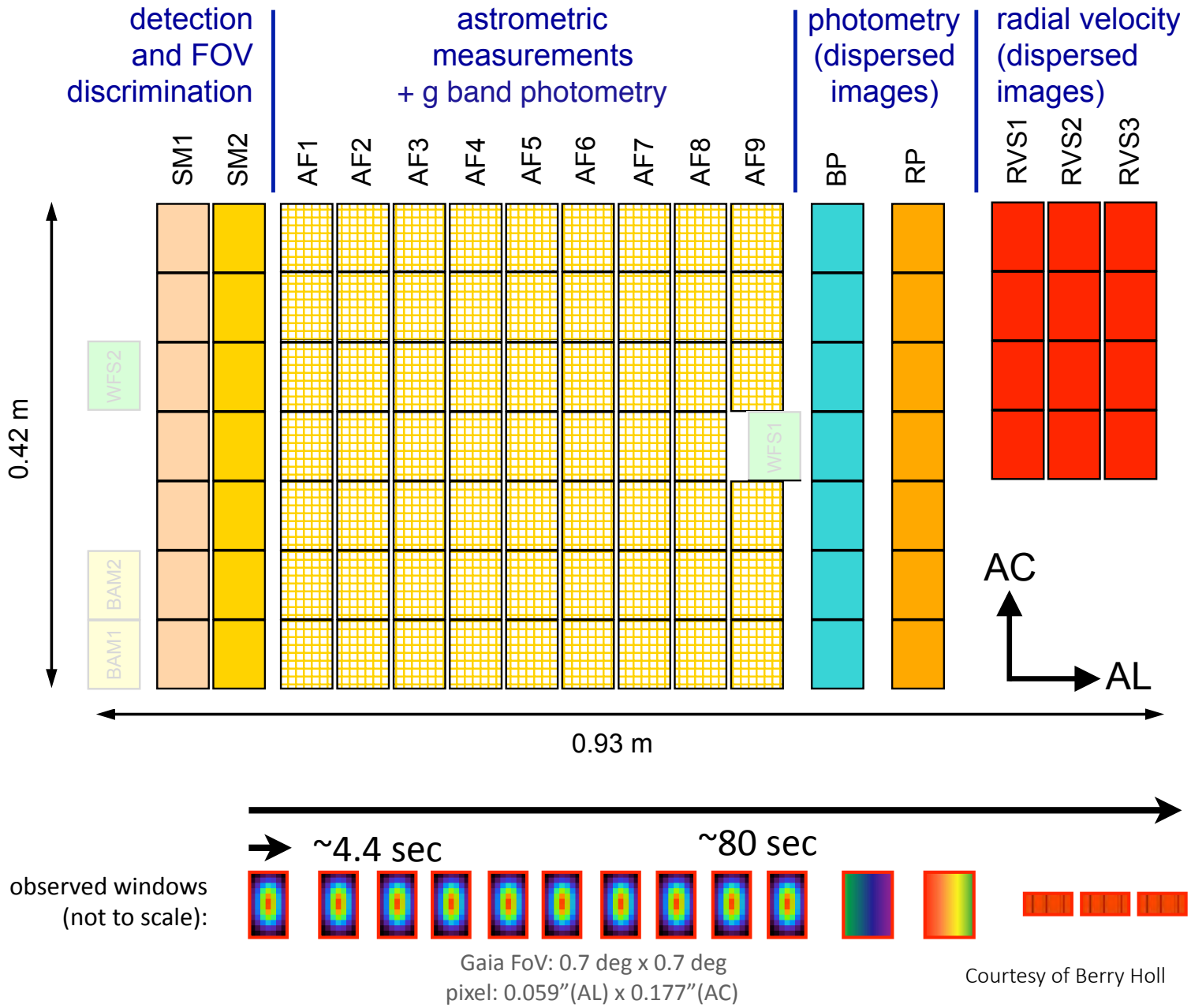
Focal plane



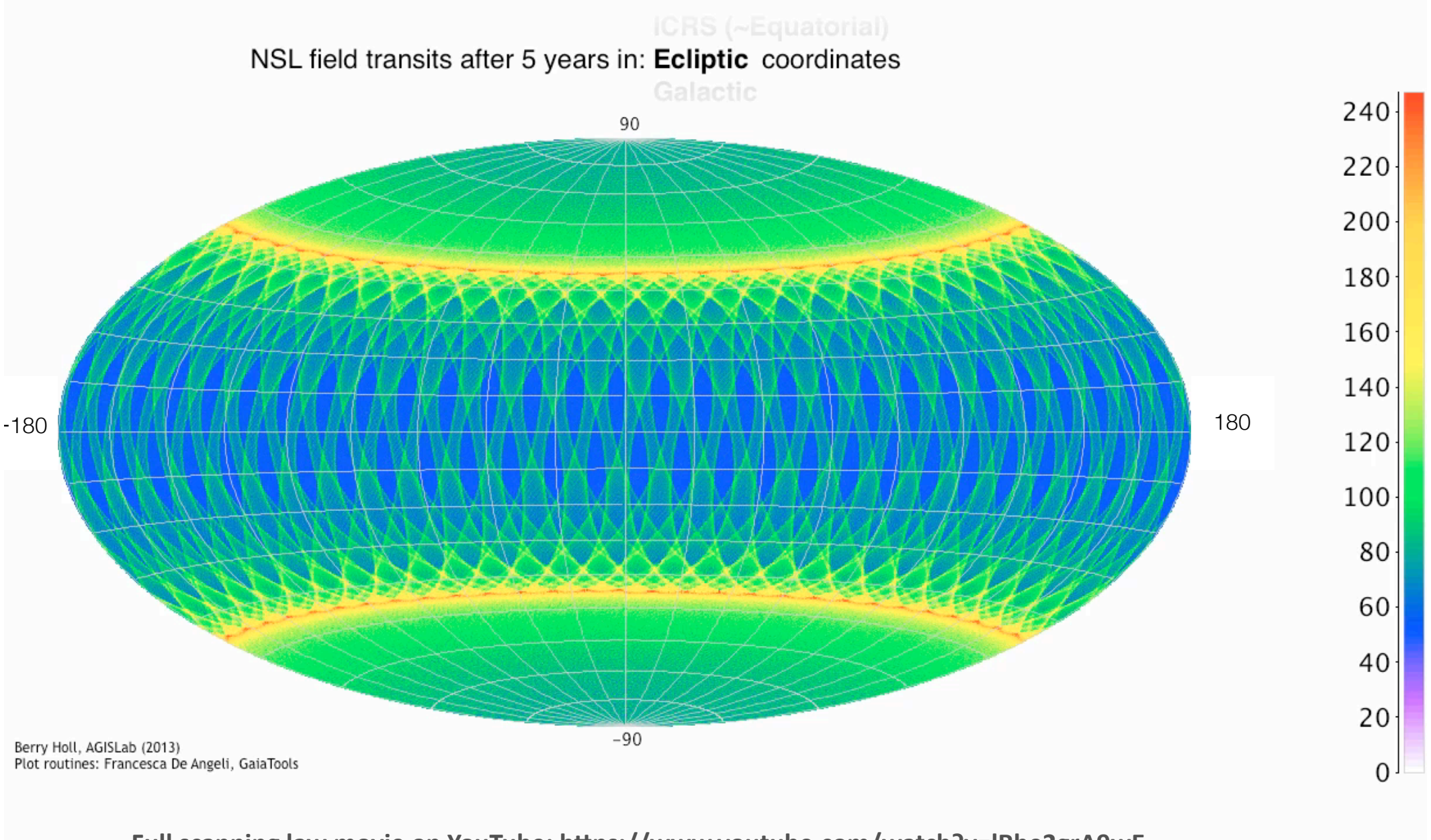


Courtesy of Berry Holl

Focal plane


detection and FOV discrimination
astrometric measurements + g band photometry
photometry (dispersed images)
radial velocity (dispersed images)
SM1
SM2
AF1
AF2
AF3
AF4
AF5
AF6
AF7
AF8
AF9
BP
RP
RVS1
RVS2
RVS3
WFS2
WFS1
BAM2
BAM1
0.42 m
AC
AL
0.93 m
~4.4 sec
~80 sec
observed windows (not to scale):
Gaia FoV: 0.7 deg x 0.7 deg
pixel: 0.059”(AL) x 0.177”(AC)
Courtesy of Berry Holl

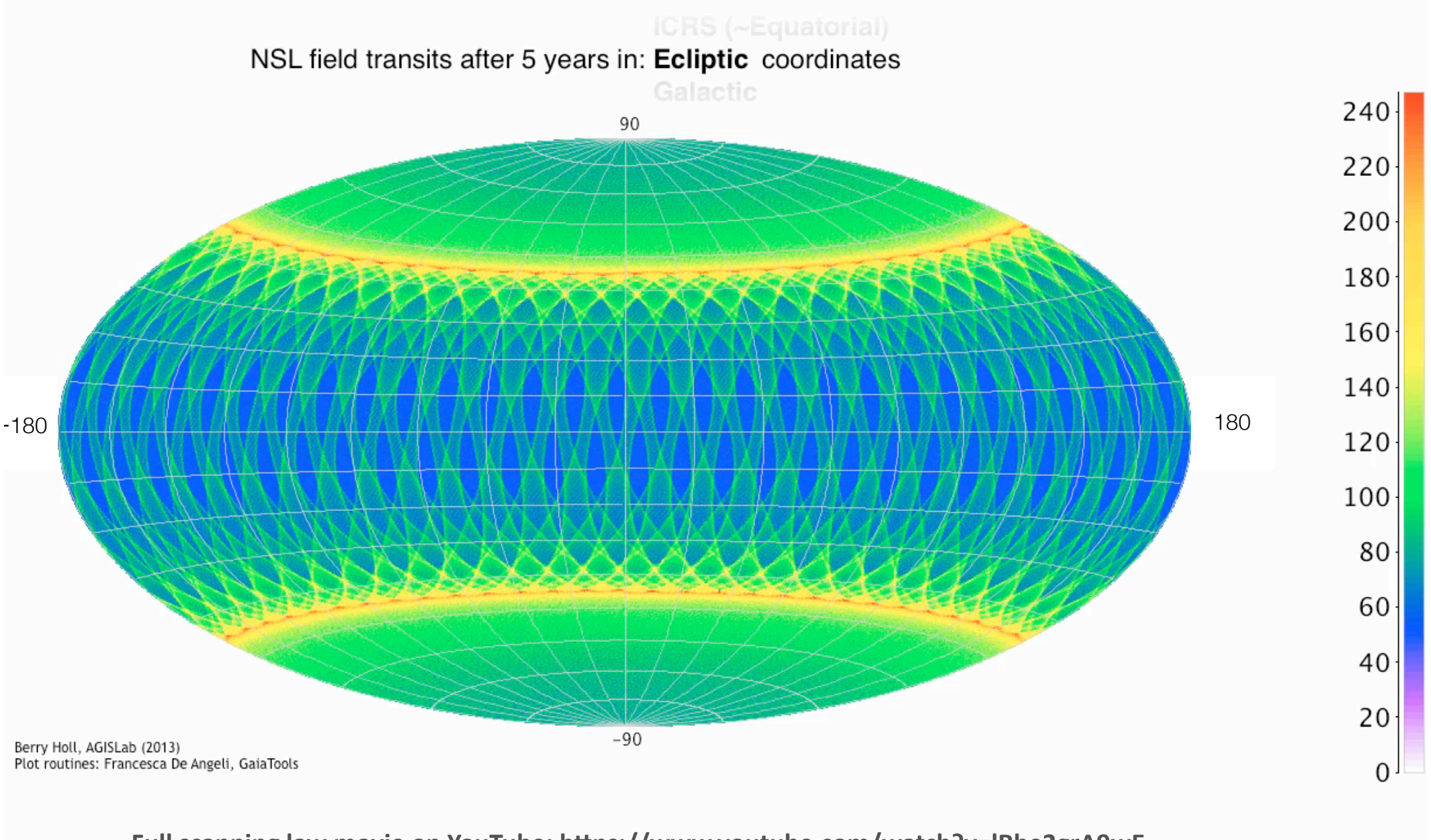
Five year nominal scanning law (NSL)



Full scanning law movie on YouTube: https://www.youtube.com/watch?v=lRhe2grA9wE

Courtesy of Berry Holl

Five year nominal scanning law (NSL)



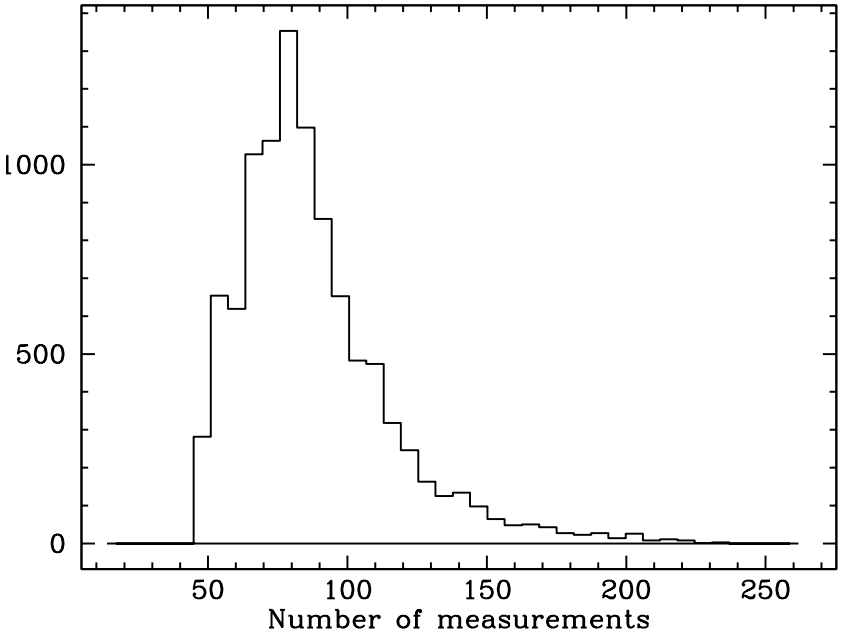
Full scanning law movie on YouTube: https://www.youtube.com/watch?v=lRhe2grA9wE

Courtesy of Berry Holl

Properties of the Gaia cadence

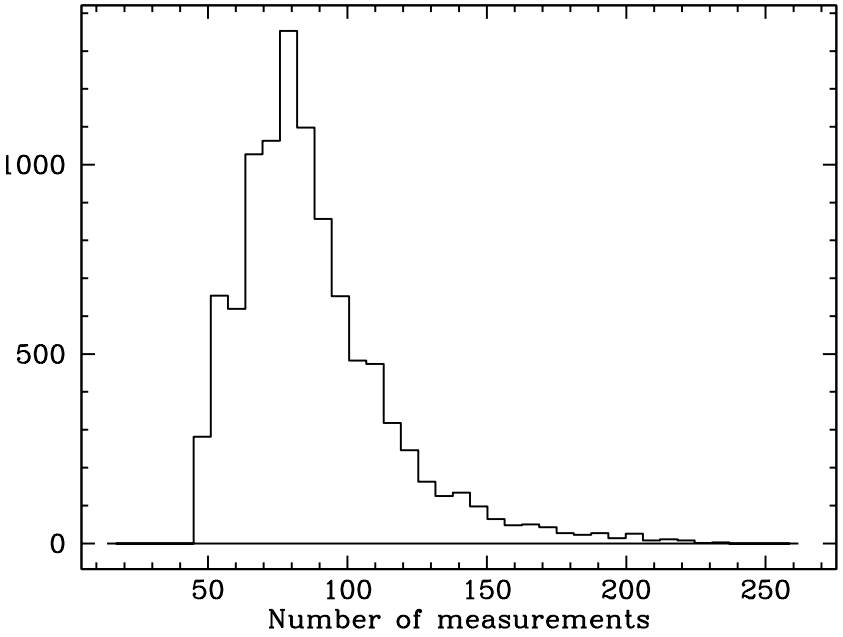
Properties of the Gaia cadence

1) Histogram of number of measurements
Mean number ~70 (dead time included)

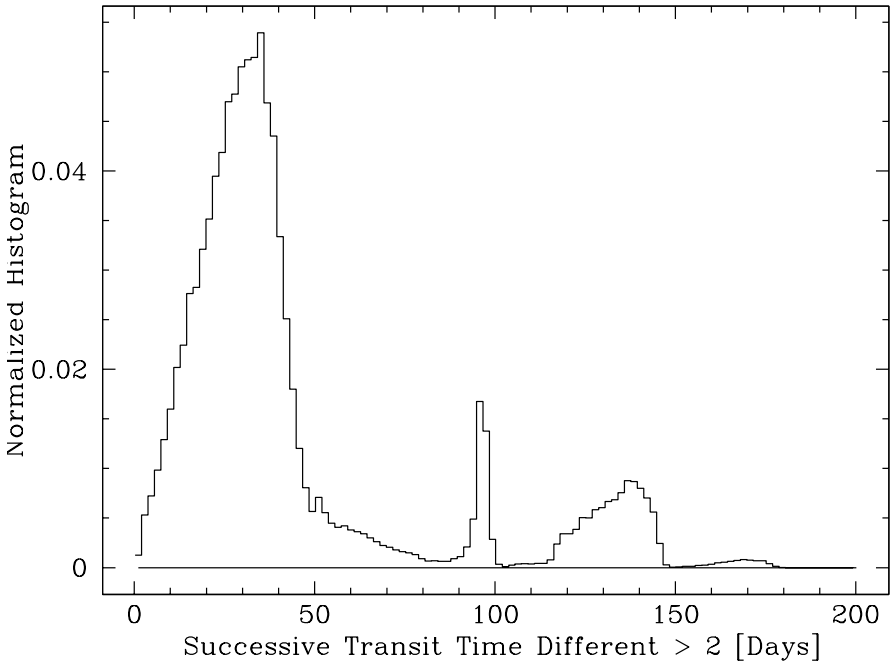


Properties of the Gaia cadence

1) Histogram of number of measurements
Mean number ~70 (dead time included)

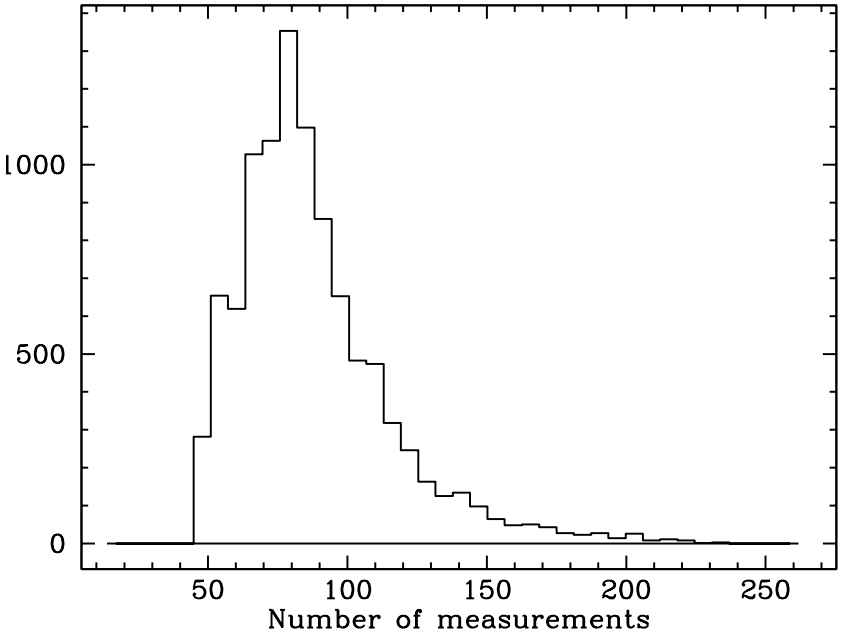


2) Histogram of successive time intervals

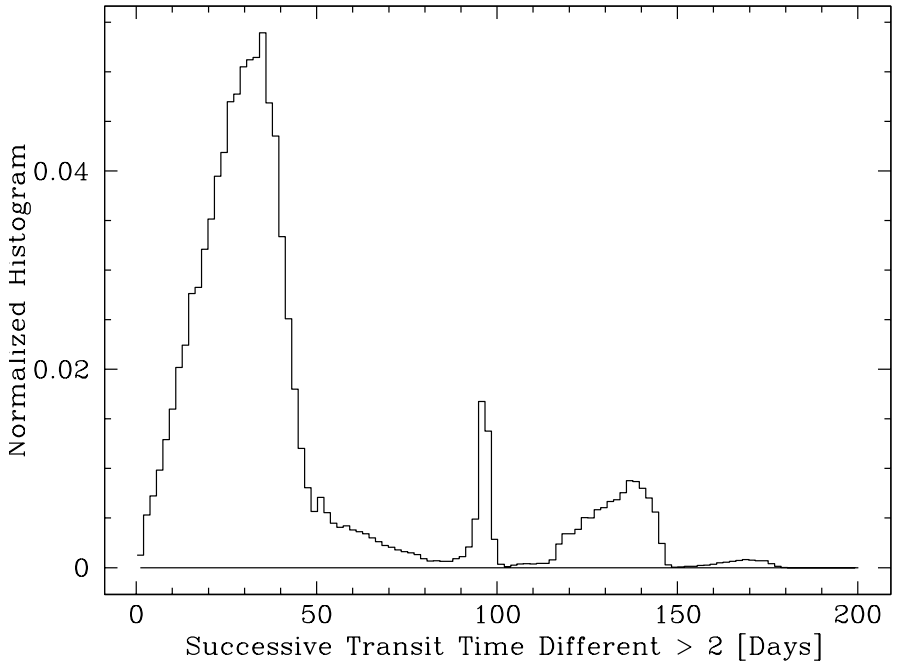


Properties of the Gaia cadence

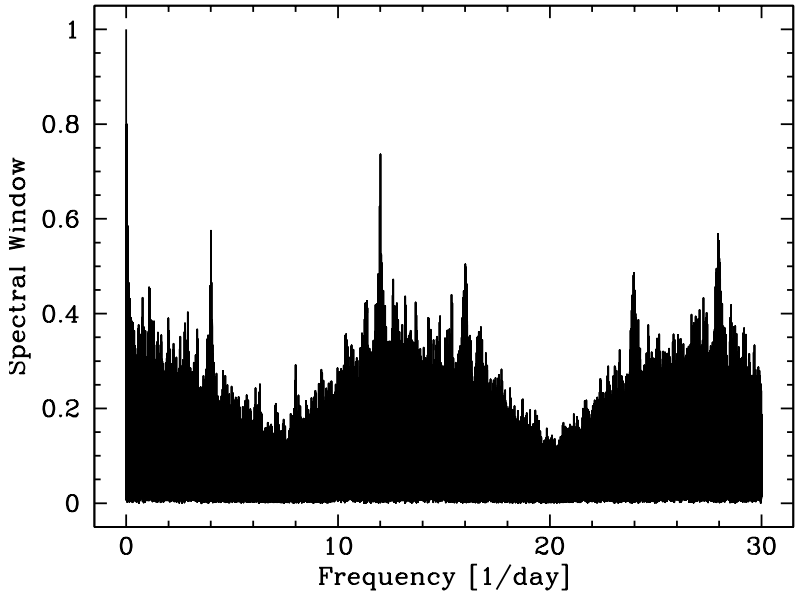
1) Histogram of number of measurements
Mean number ~70 (dead time included)



2) Histogram of successive time intervals



3) Spectral window



The Gaia astrometric performance: Mean parallax error (end-of-mission)

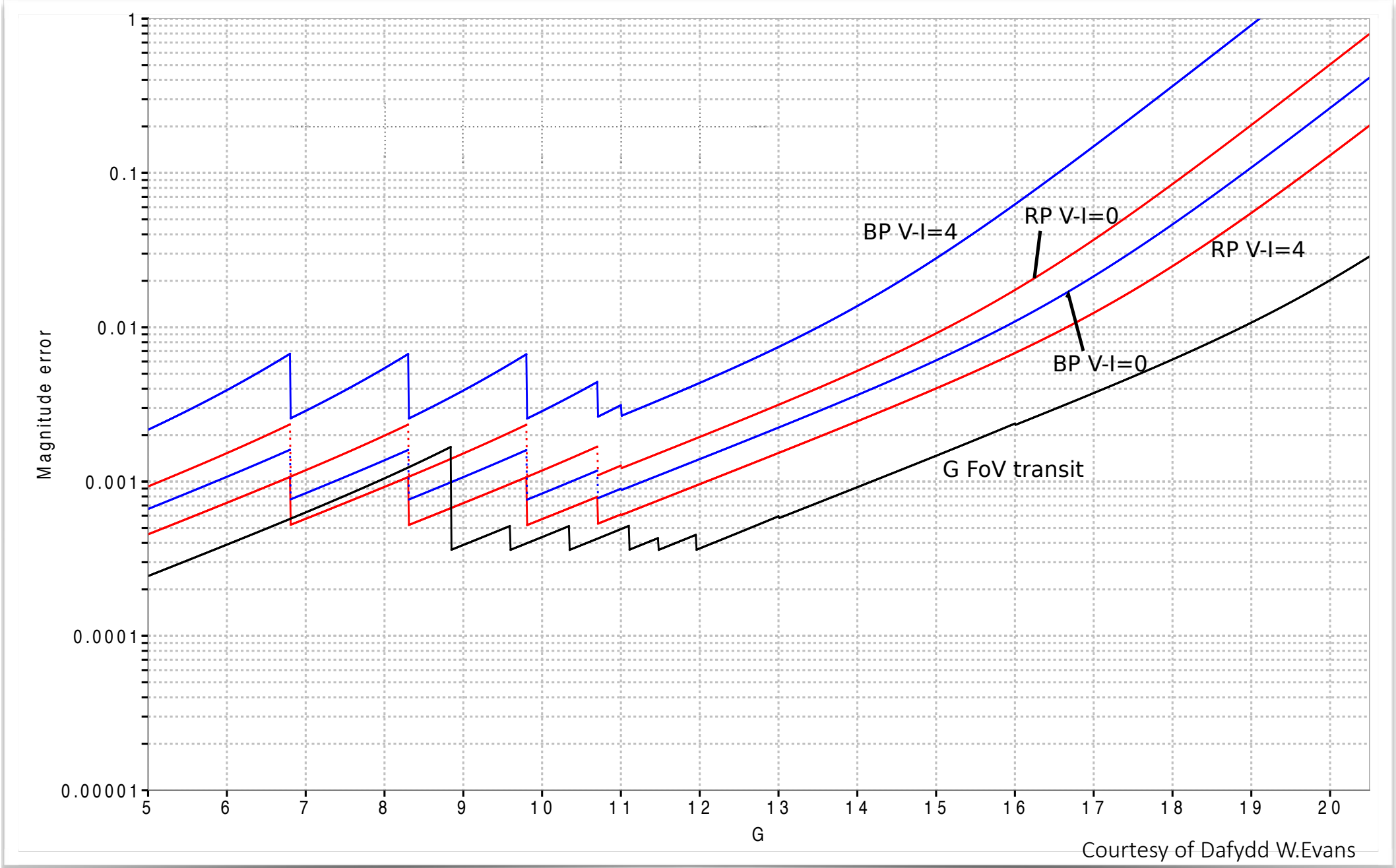
Mean parallax error end-of-mission (5 year)

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

Position error (μas): ... x ~0.7 ~ $1/(\text{mission length})^{0.5}$

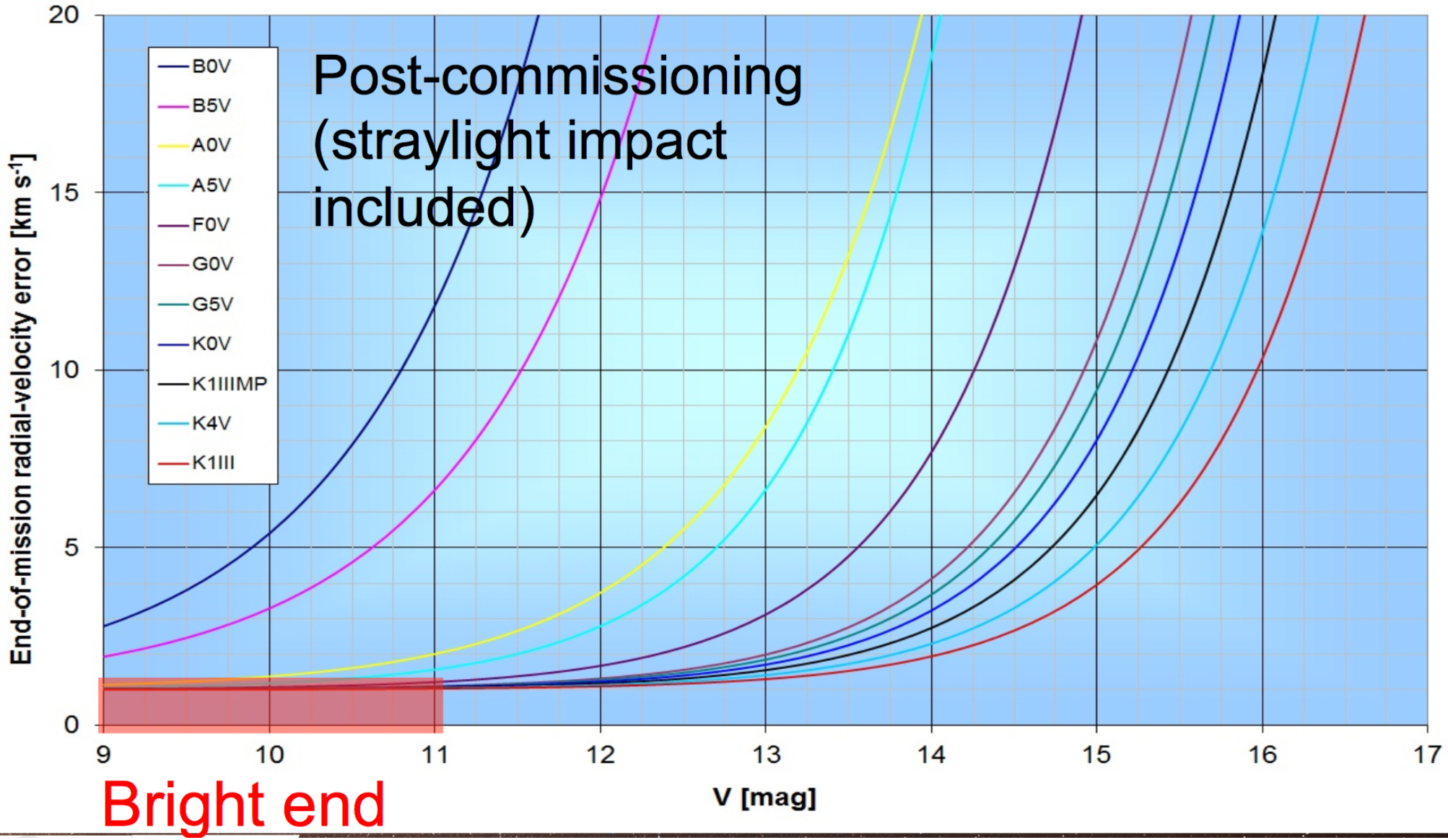
Proper motion error (μas/year): ... x ~0.5 ~ $1/(\text{mission length})^{1.5}$

Epoch Field-of-View photometric precision



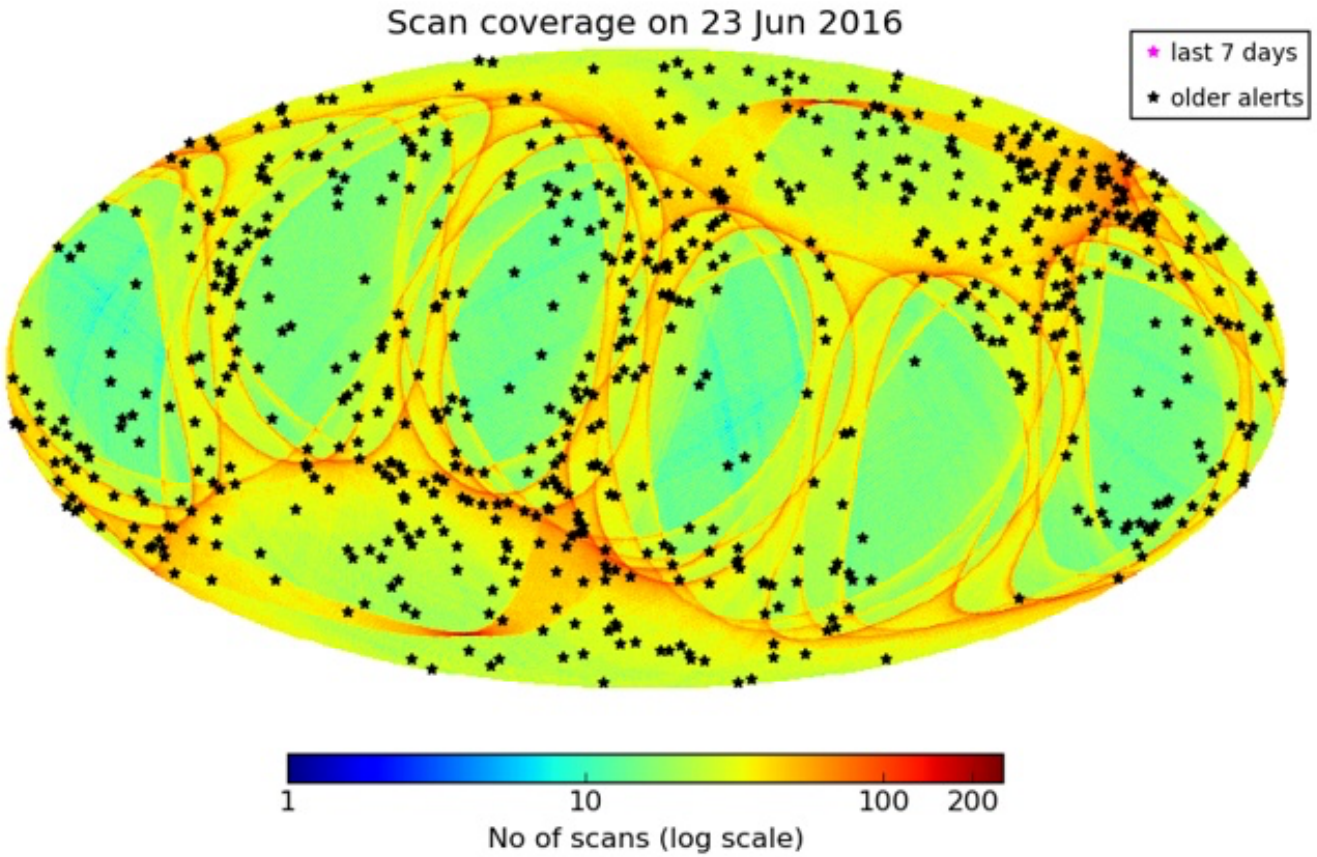
Courtesy of Dafydd W.Evans

Radial velocity end of mission precision



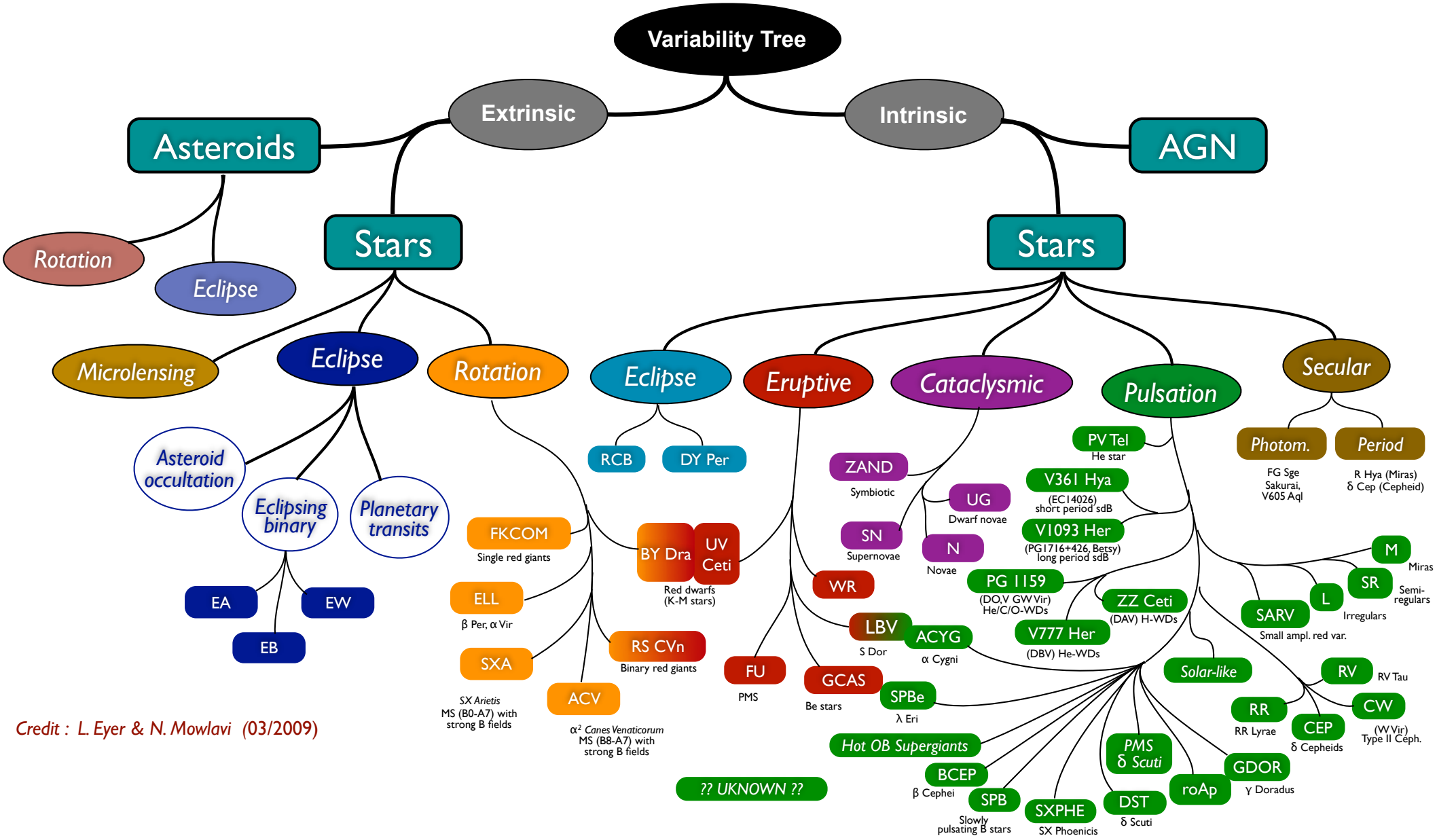
Courtesy of Georges Seabrook

Variability I: Gaia Science Alerts (Cambridge Univ.)



As of today: 696 Gaia alerts

Variability II

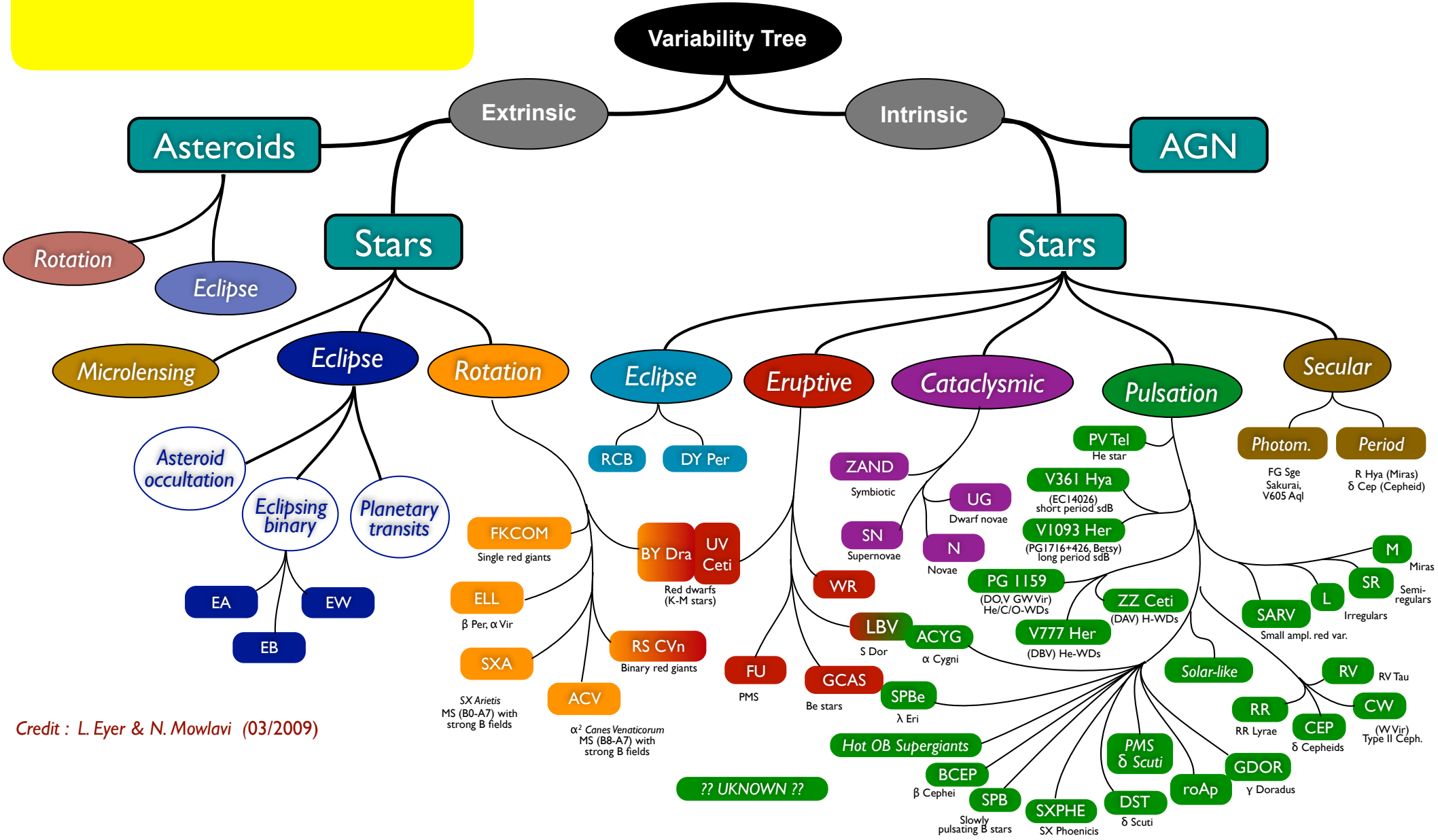


Credit : L. Eyer & N. Mowlavi (03/2009)

Variability II

For a source:

Gaia data

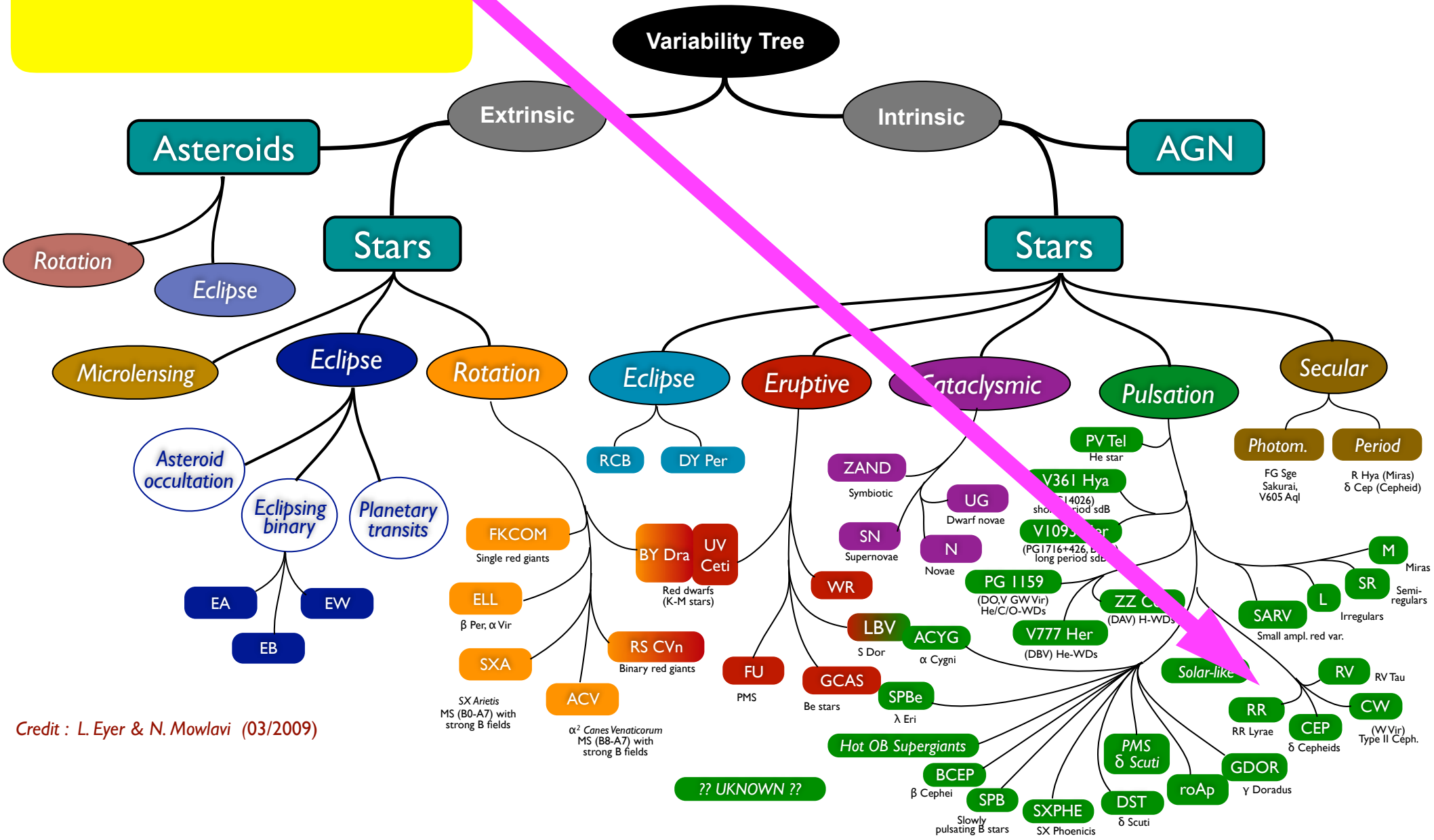


Credit : L. Eyer & N. Mowlavi (03/2009)

Variability II

For a source:

Gaia data



Credit : L. Eyer & N. Mowlavi (03/2009)

Photometry + radial velocities + astrometry

Example:
Hipparcos and CORAVEL

We searched for stars variable in Hipparcos
who had CORAVEL measurements

~ **160 Hipparcos variables** with more than 10 radial velocities measurements with CORAVEL (σ_{VR} ~ 1 km/s)

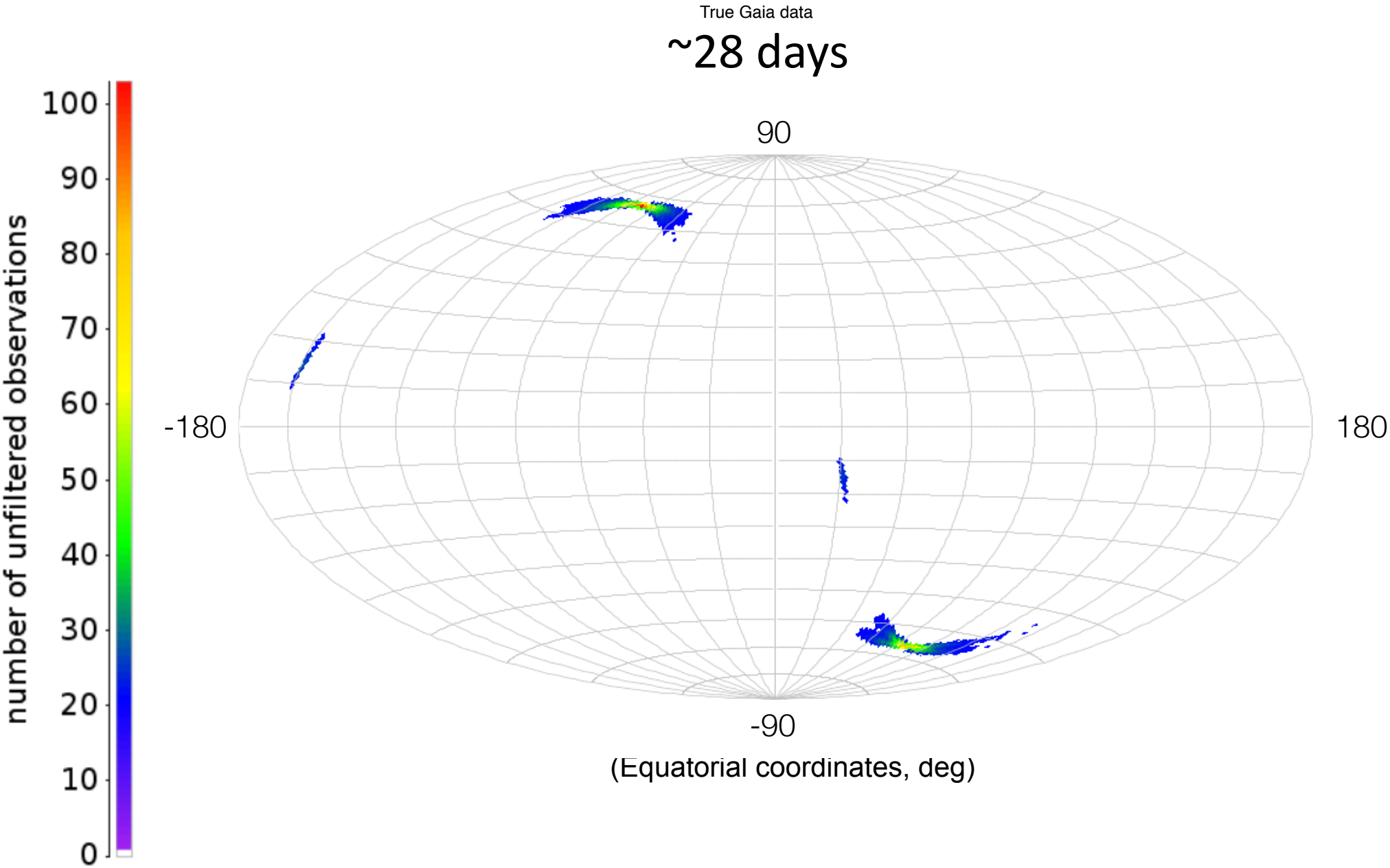
With parallaxes better than 10%

~30 stars

Variability Processing and Analysis: Pragmatic approach

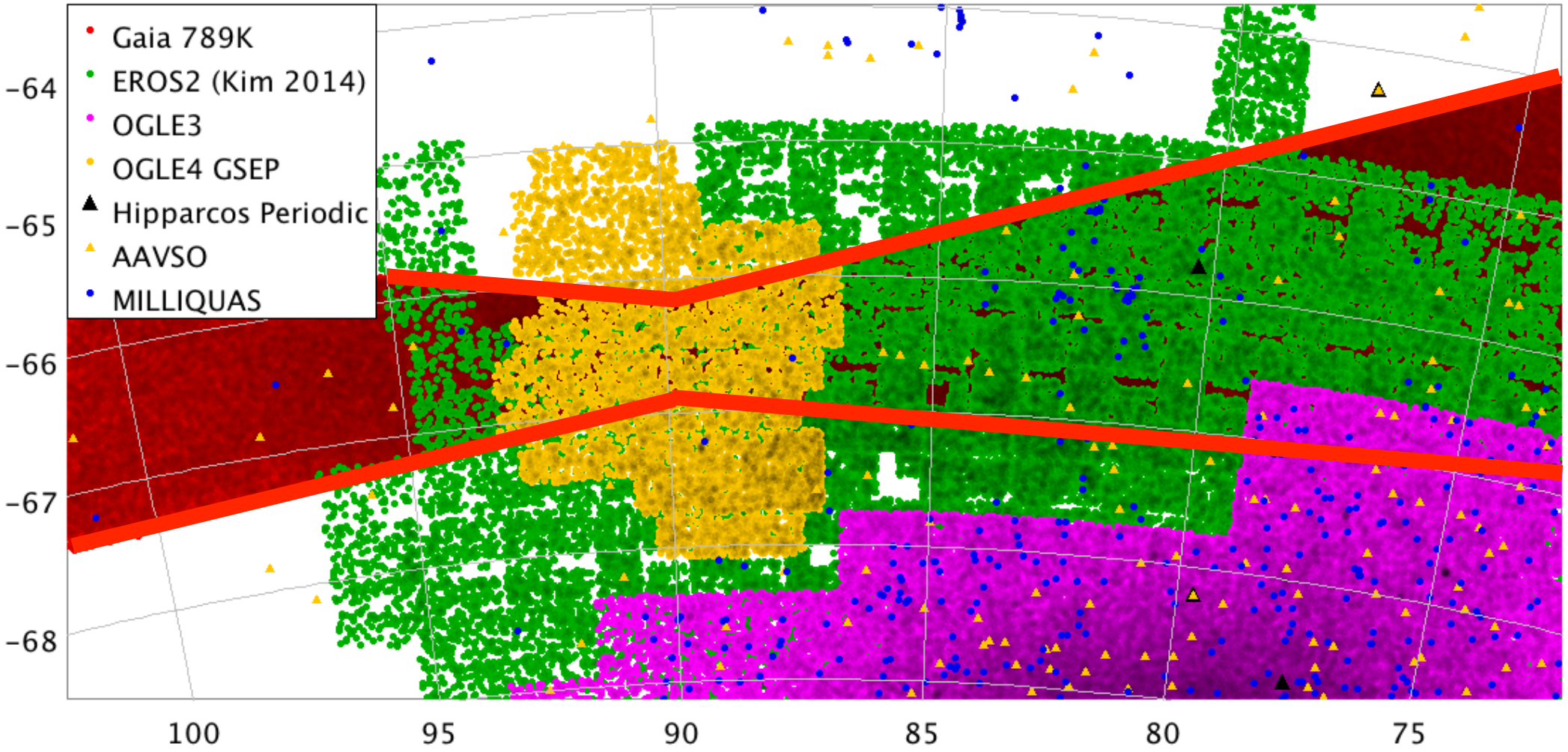
- **Develop a "method"**
  - Test on simulated data (often very simple modeling)
  - Test on real data
    - OGLE, EROS
    - Kepler
    - SDSS
    - Hipparcos-Coravel, etc.
- **Apply it to real Gaia data**
  - Start on "easy" signals (Cepheids/RR Lyrae stars), e.g. large signal to noise ratio, "regular" variables
  - Start with G-band, then include BP/RP, then astrometry, then RVS data (not really a choice)
  - Start with law numbers of objects
- **Interact with the data (subsamples) at all levels of the analysis**

Selection of sources observed by Gaia from Ecliptic Pole Scanning Law (790,000 sources) with 20+ measurements



Warning: **Sampling regular, different from Nominal Scanning law**

South Ecliptic Pole region (part of Large Magellanic Cloud): Gaia and other surveys

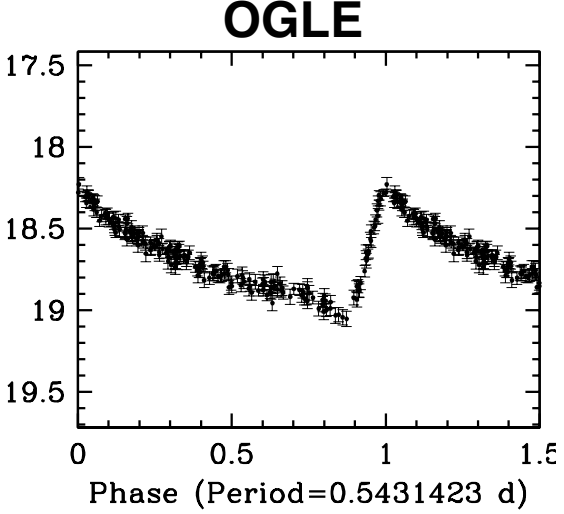


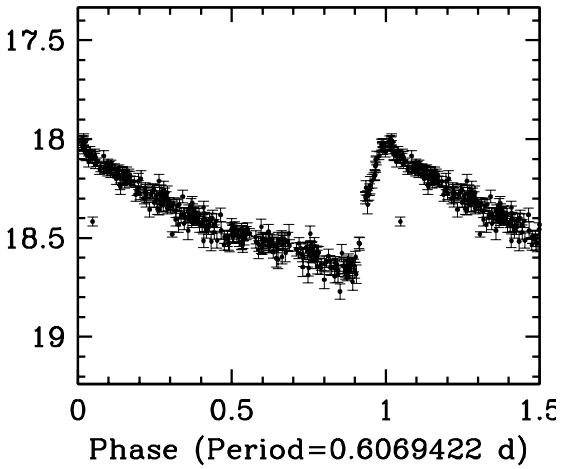
True Gaia data

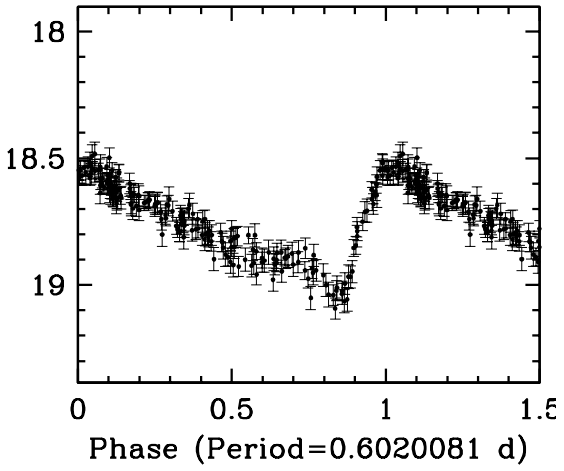
To get the data flavour
Comparison with OGLE

Image of the Week (March 05, 2015):
RR Lyrae stars

Credits: ESA/Gaia/DPAC/CU5/CU7/INAF-OABo, Gisella Clementini, Dafydd Evans, Laurent Eyer, Krzysztof Nienartowicz, Lorenzo Rimoldini and the Geneva CU7/DPCG and CU7/INAF-OACN teams.

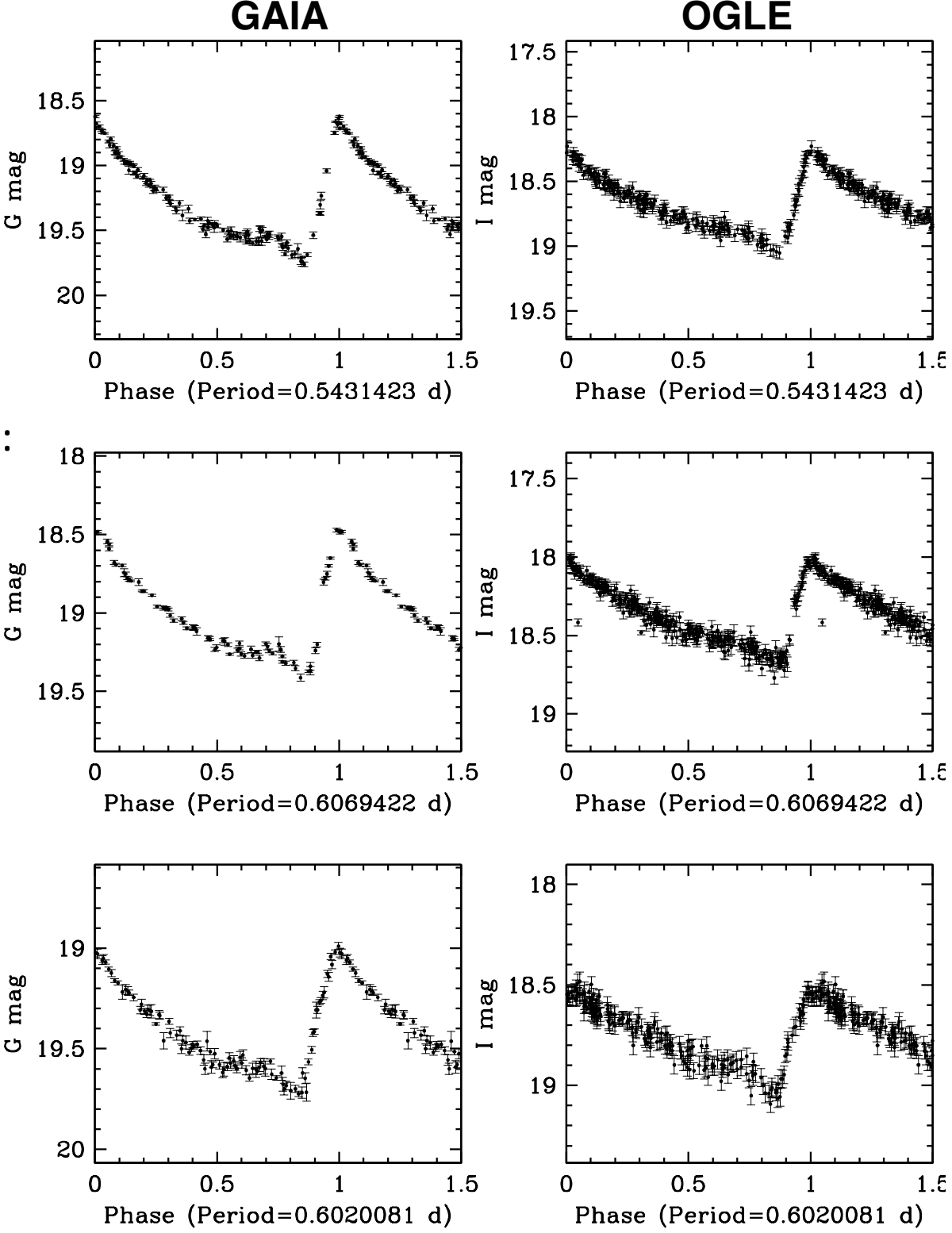






To get the data flavour
Comparison with OGLE

Image of the Week (March 05, 2015):
RR Lyrae stars



Credits: ESA/Gaia/DPAC/CU5/CU7/INAF-OABo, Gisella Clementini, Dafydd Evans, Laurent Eyer, Krzysztof Nienartowicz, Lorenzo Rimoldini and the Geneva CU7/DPCG and CU7/INAF-OACN teams.

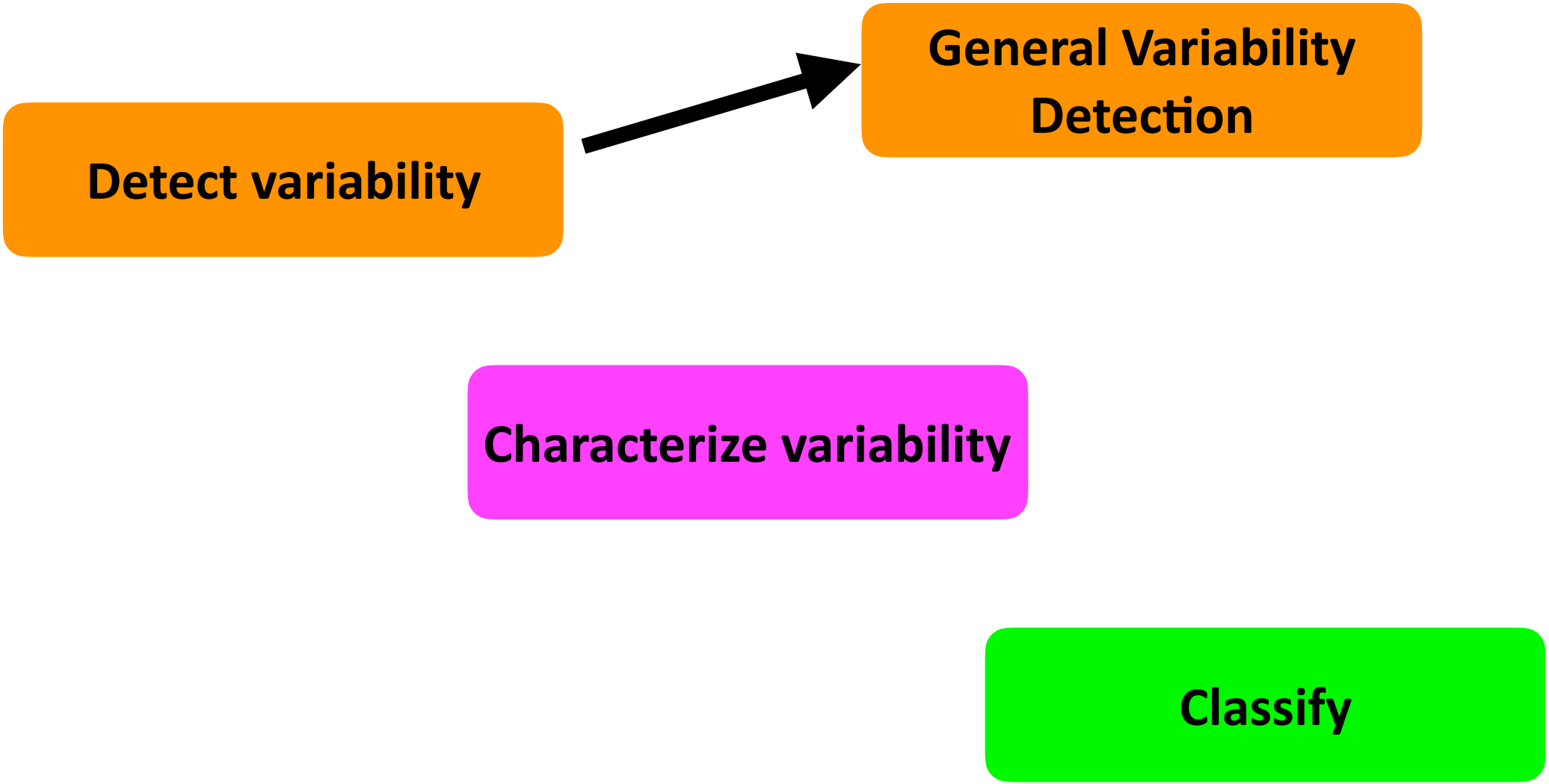
Gaia Variability Processing and Analysis

Detect variability

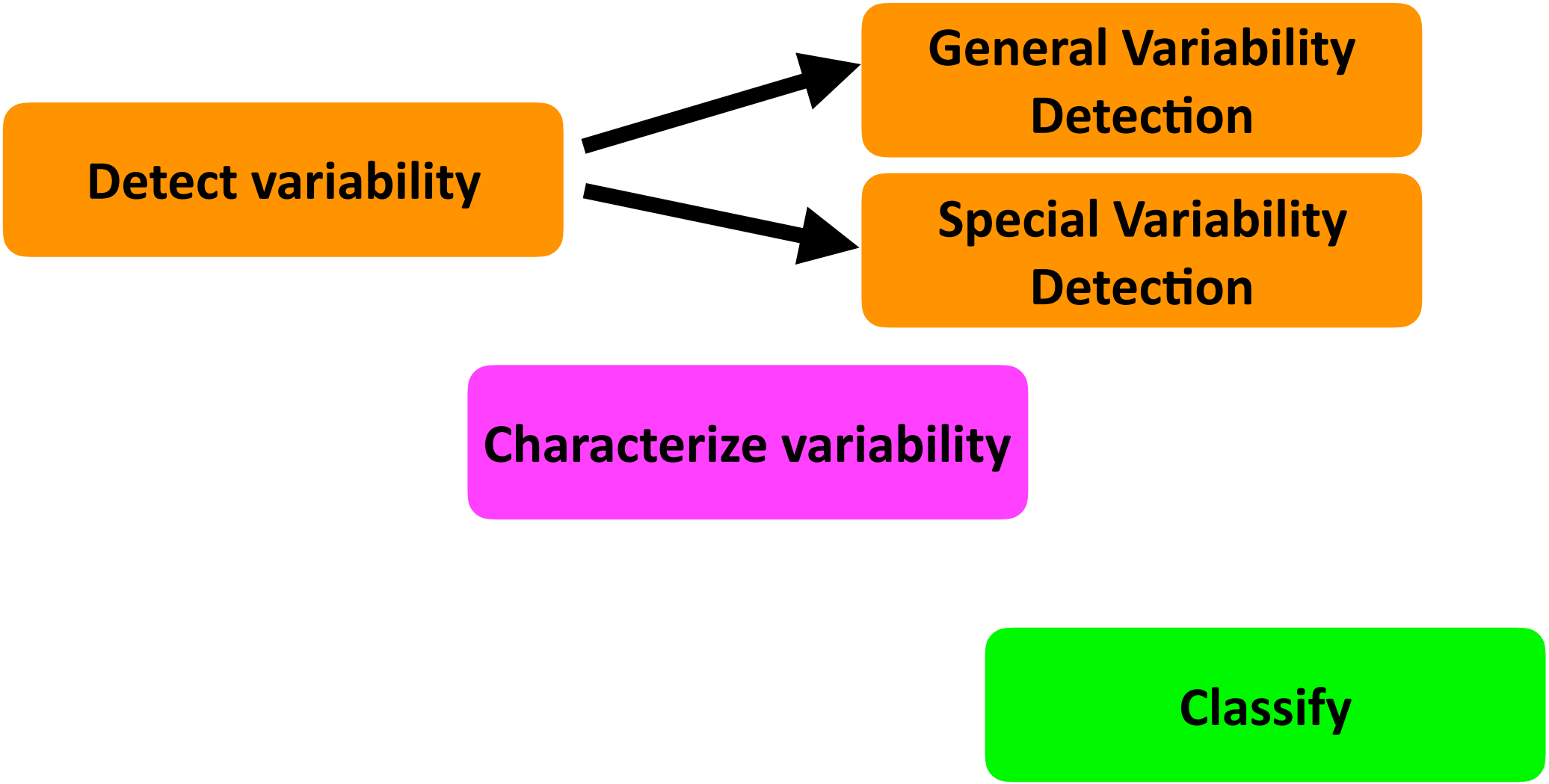
Characterize variability

Classify

Gaia Variability Processing and Analysis



Gaia Variability Processing and Analysis



General Variability Detection

- Classical hypothesis testing is difficult to applied because of underestimated uncertainties

General Variability Detection

- Classical hypothesis testing is difficult to applied because of underestimated uncertainties

- Detection was done with a classifier (Random Forest)
  - many attributes were computed
  - a training set was defined (based on OGLE)

General Variability Detection

- Classical hypothesis testing is difficult to applied because of underestimated uncertainties

- Detection was done with a classifier (Random Forest)
  - many attributes were computed
  - a training set was defined (based on OGLE)

two fundamental quantities to estimate:

—Completeness

—Contamination

General Variability Detection

- Classical hypothesis testing is difficult to applied because of underestimated uncertainties

- Detection was done with a classifier (Random Forest)
  - many attributes were computed
  - a training set was defined (based on OGLE)

two fundamental quantities to estimate:

—Completeness

—Contamination

General Variability Detection

- Classical hypothesis testing is difficult to applied because of underestimated uncertainties

- Detection was done with a classifier (Random Forest)
  - many attributes were computed
  - a training set was defined (based on OGLE)

two fundamental quantities to estimate:

—Completeness

—Contamination

For LSST: Importance to have other surveys (e.g. Kepler , Gaia, etc.) for preparation of simulations, calibration and comparison (K2)

Special Variability Detection

Solar like (magnetic) activity

Exo-planetary transits

Small amplitude

Short time scale (per FoV, per CCD data)

Variable type	Period range
β Cep	1.6h - 12h
δ Scuti	20min - 8h
roAp	5min - 25min
EC 14026	80s - 8min
Besty stars	30-150min
RR Lyrae	5h - 1d
ZZ Ceti	30s - 25min
V777 Her	2-16min
GW Vir	5-80min
AM CVn	5min - 65min

Special Variability Detection

Solar like (magnetic) activity

Exo-planetary transits

Small amplitude

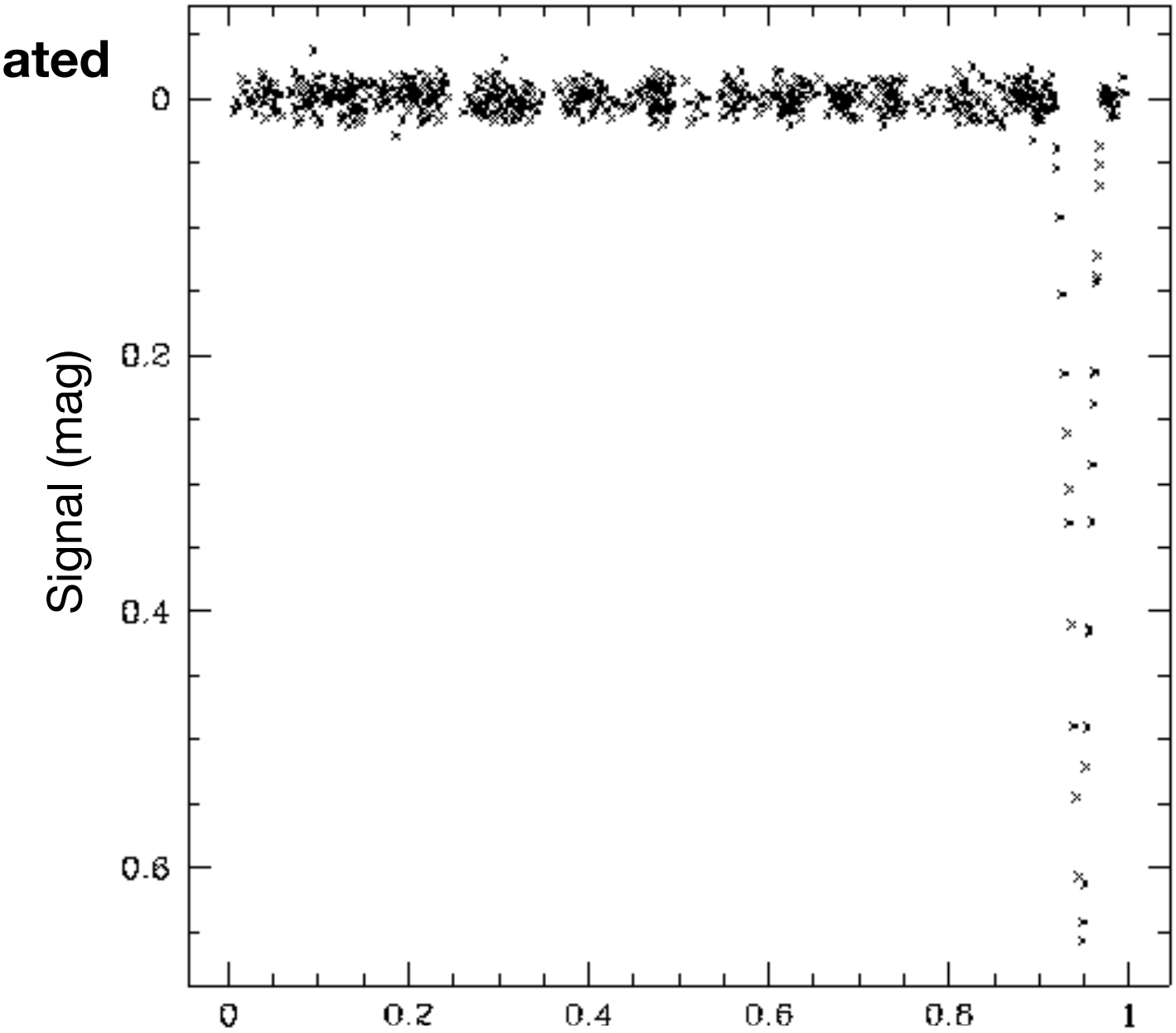
Short time scale (per FoV, per CCD data)

Variable type	Period range
β Cep	1.6h - 12h
δ Scuti	20min - 8h
roAp	5min - 25min
EC 14026	80s - 8min
Besty stars	30-150min
RR Lyrae	5h - 1d
ZZ Ceti	30s - 25min
V777 Her	2-16min
GW Vir	5-80min
AM CVn	5min - 65min

LSST: Many interesting objects for short cadence twilight survey, light curve should be simulated to study the performance of LSST

Special Variability Detection: short time scales

Simulated



Per CCD photometry

~30 are known, 200 AM CVn stars may be detected by Gaia (Nelemans 2013)

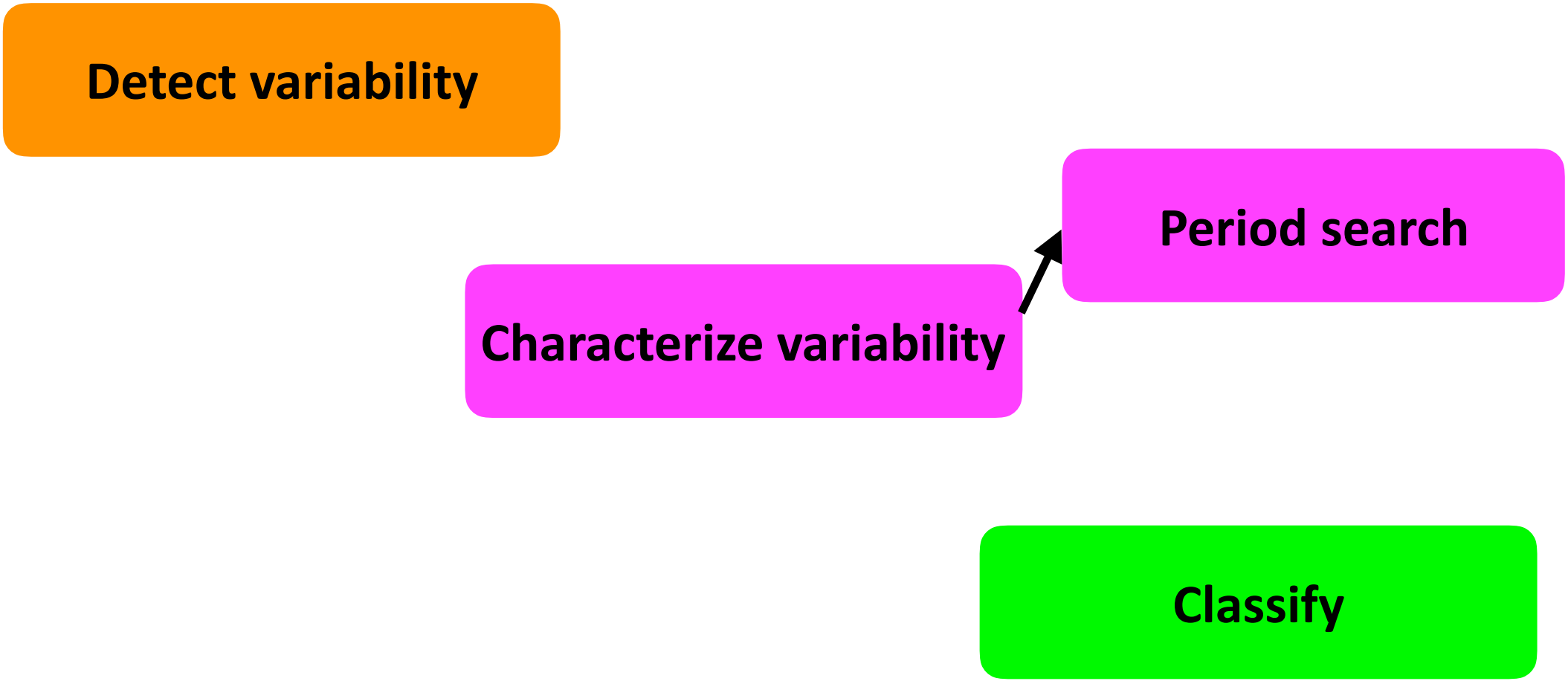
Gaia Variability Processing and Analysis

Detect variability

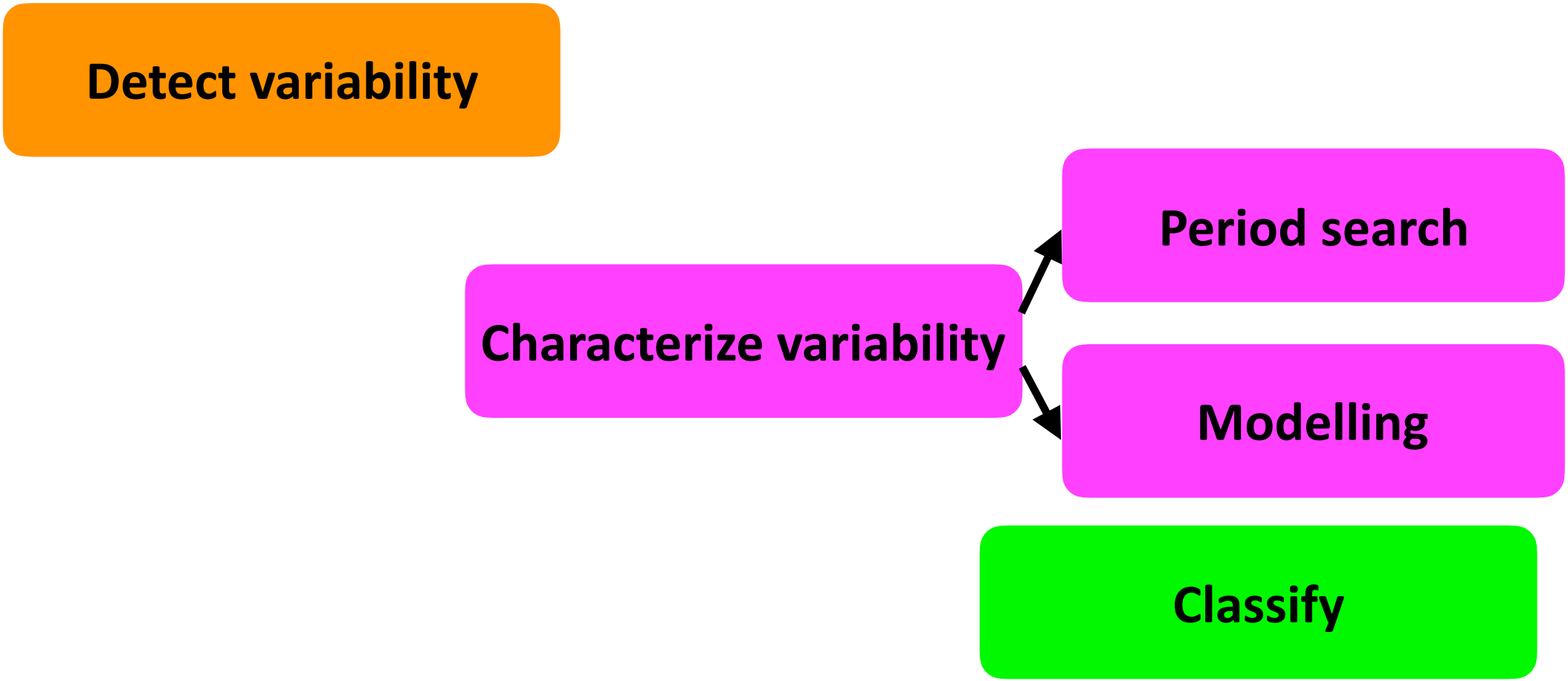
Characterize variability

Classify

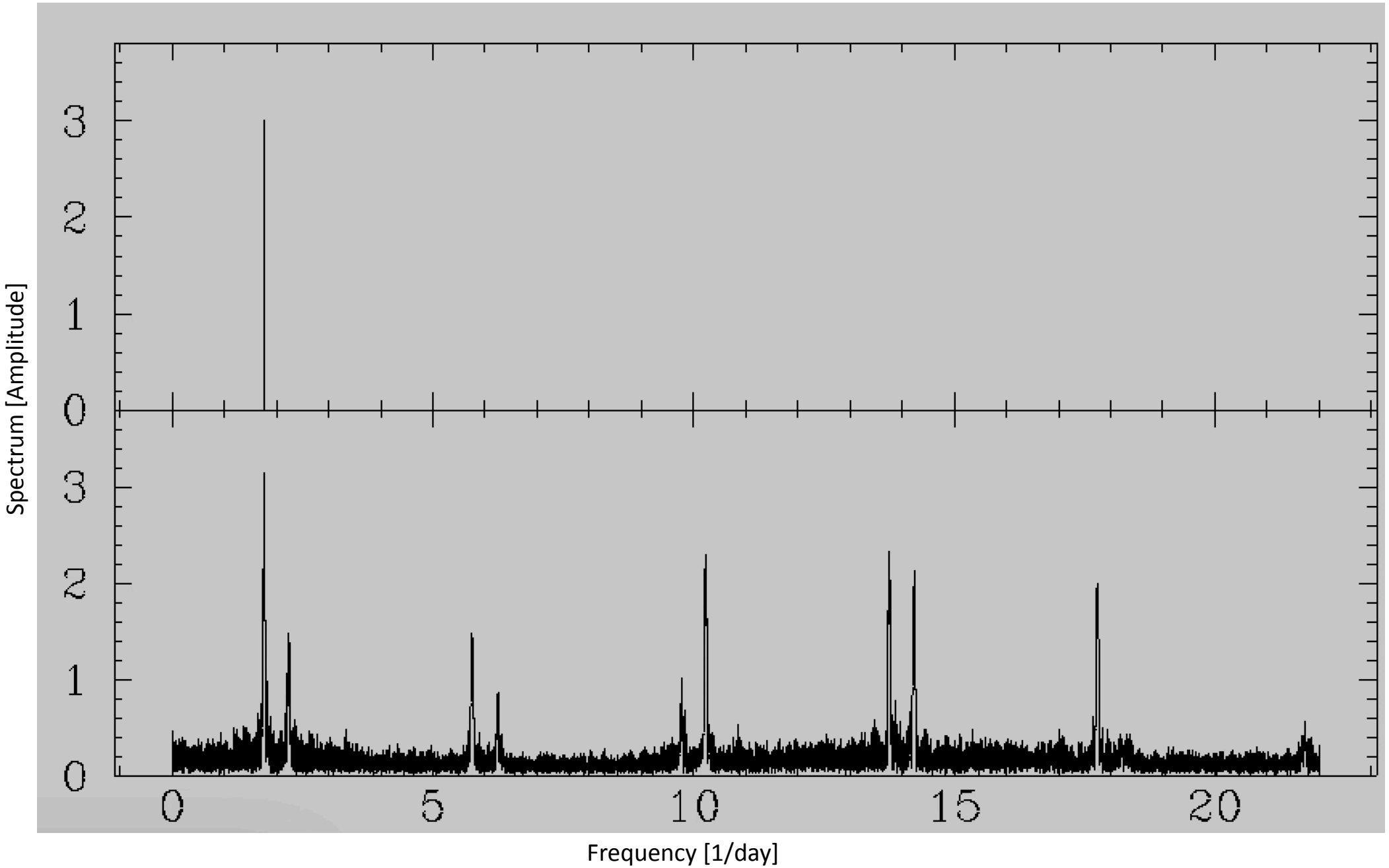
Gaia Variability Processing and Analysis



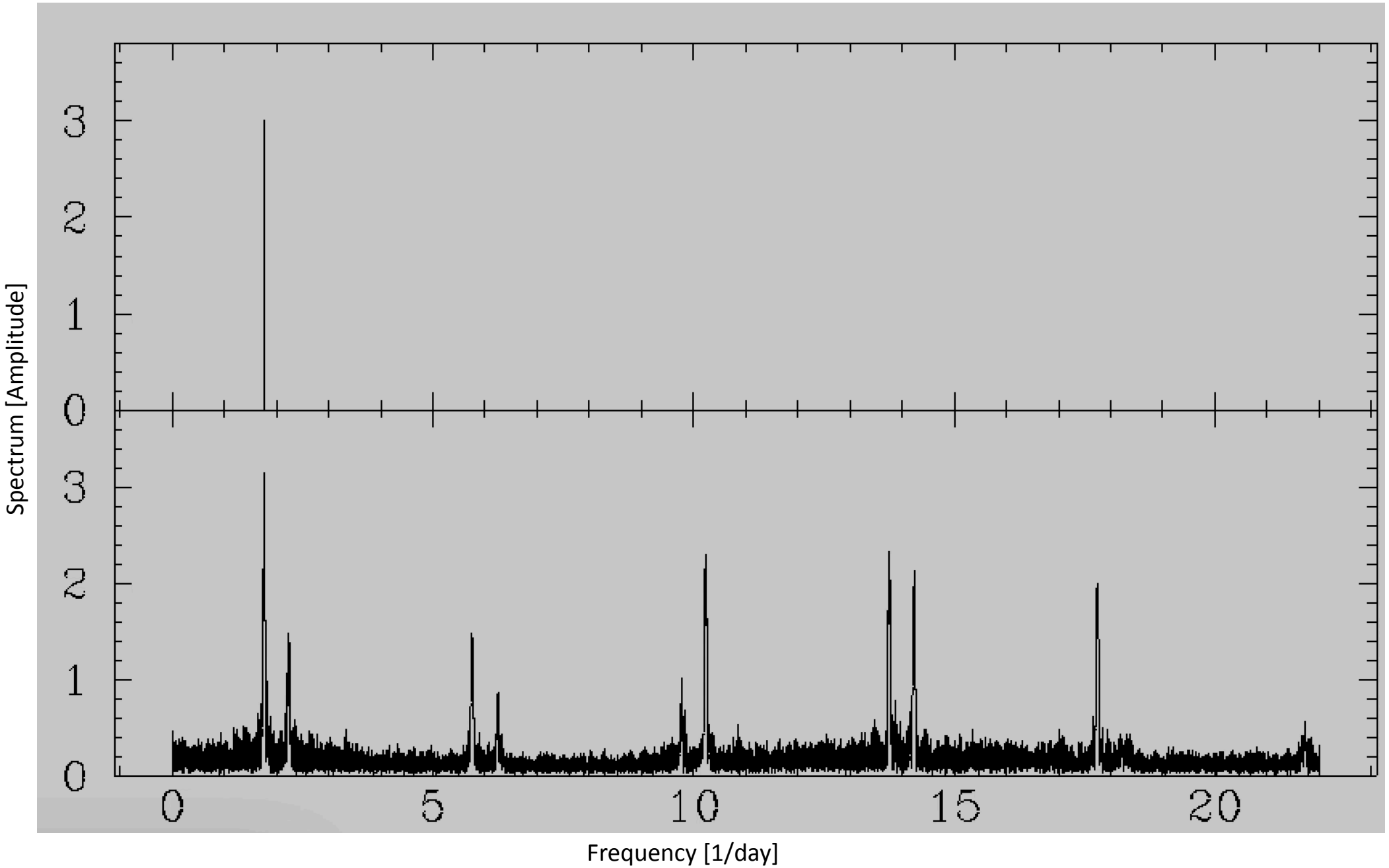
Gaia Variability Processing and Analysis



Few words about period search (Convolution)

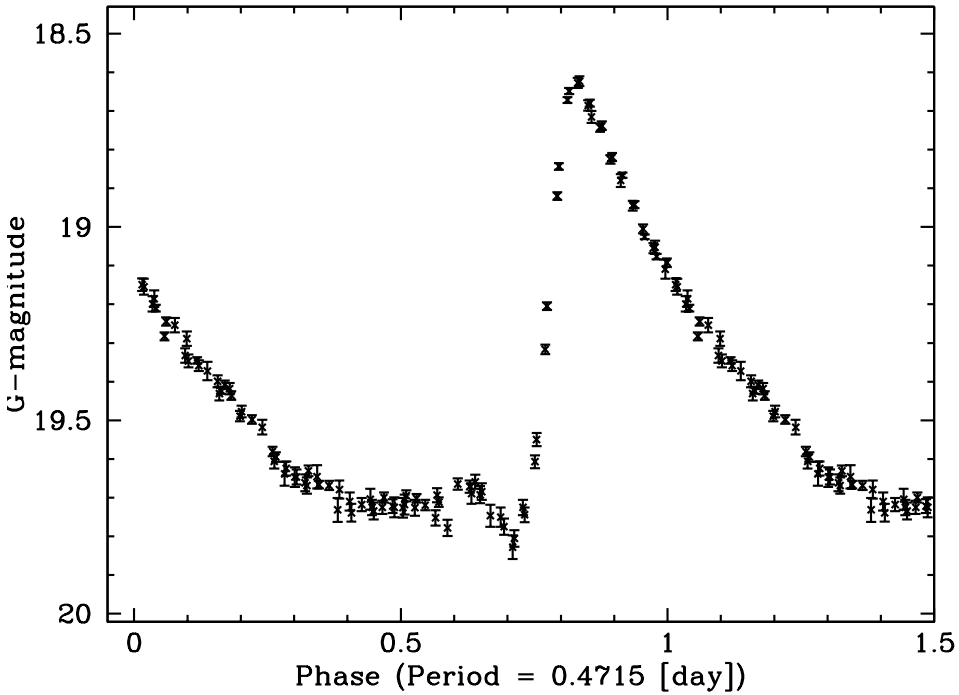


Few words about period search (Convolution)

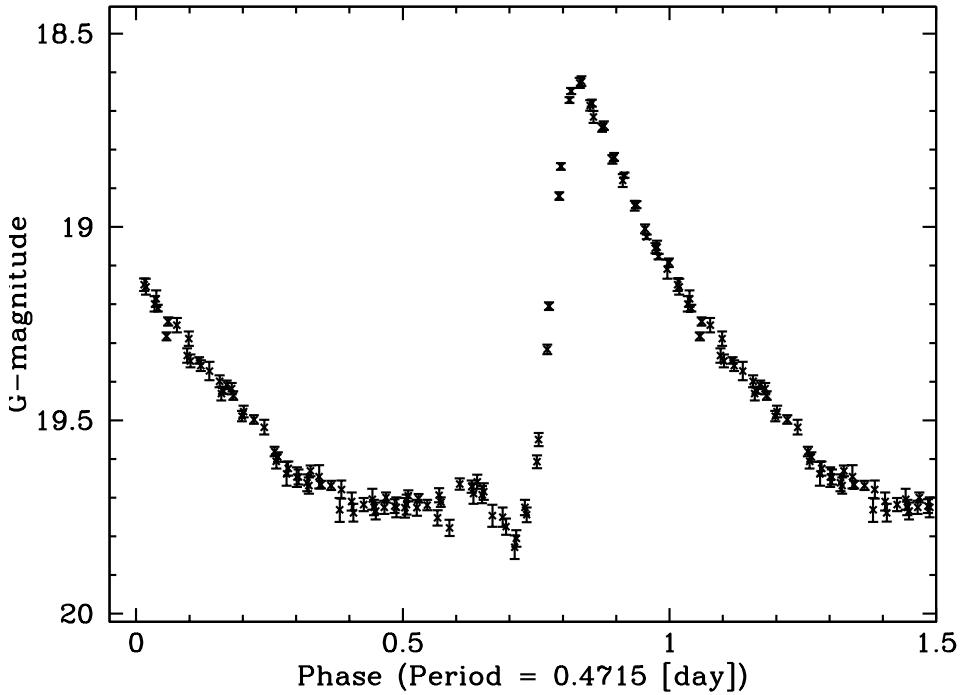


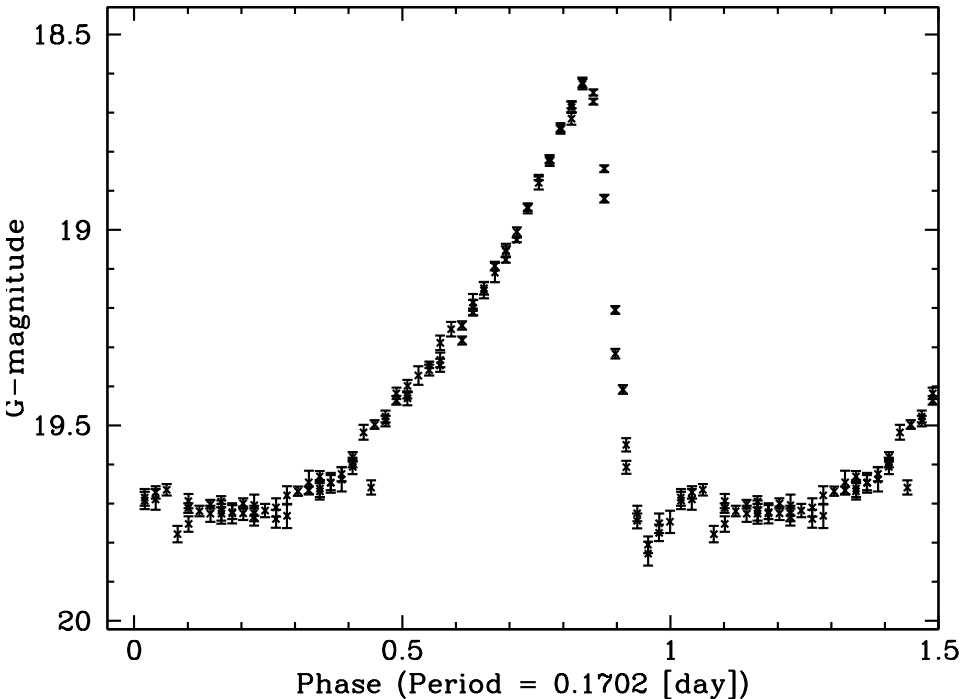
An example of aliasing in Gaia (same star with two different periods) in EPSL (too regular sampling)

An example of aliasing in Gaia (same star with two different periods) in EPSL (too regular sampling)

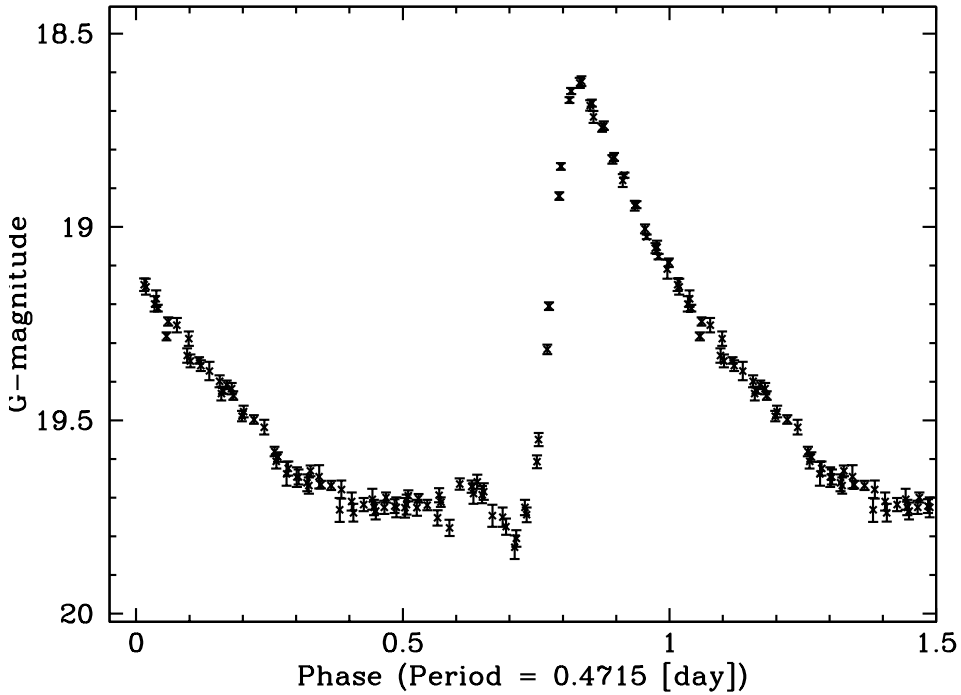


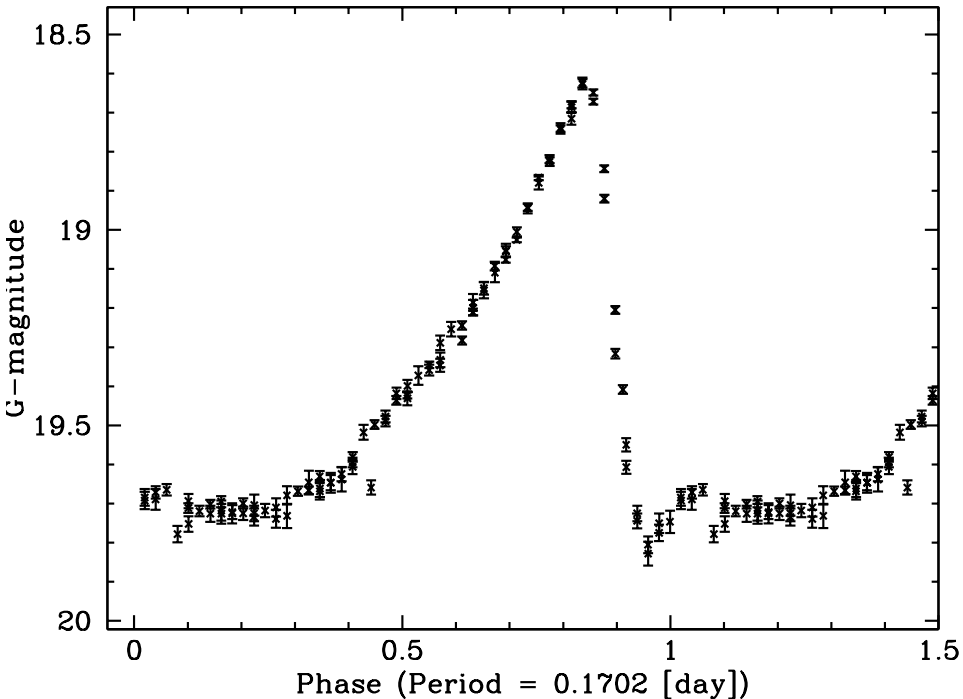
An example of aliasing in Gaia (same star with two different periods) in EPSL (too regular sampling)





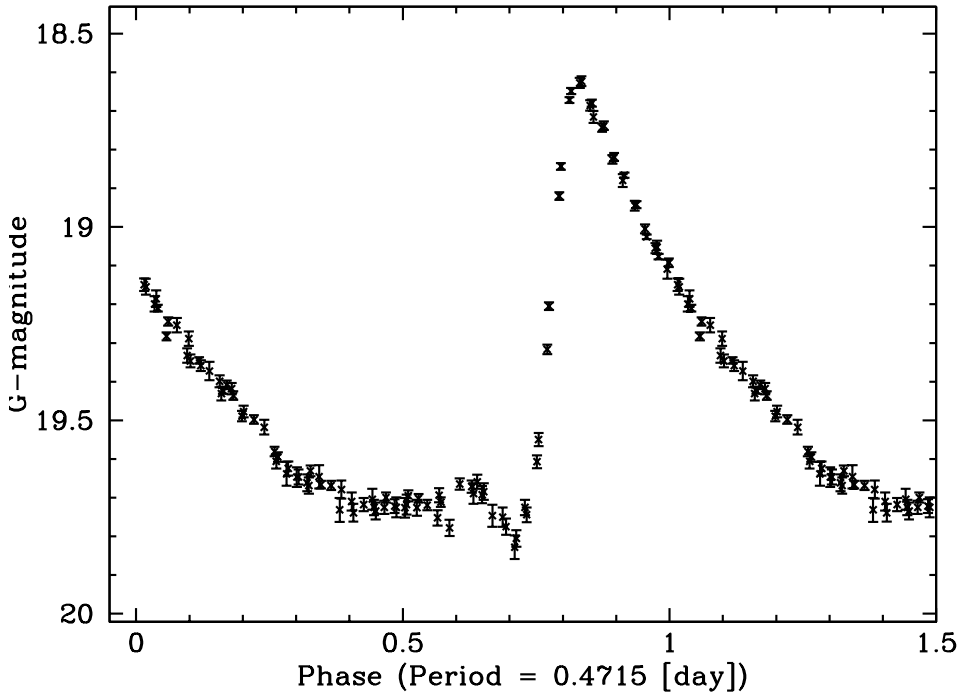
An example of aliasing in Gaia (same star with two different periods) in EPSL (too regular sampling)

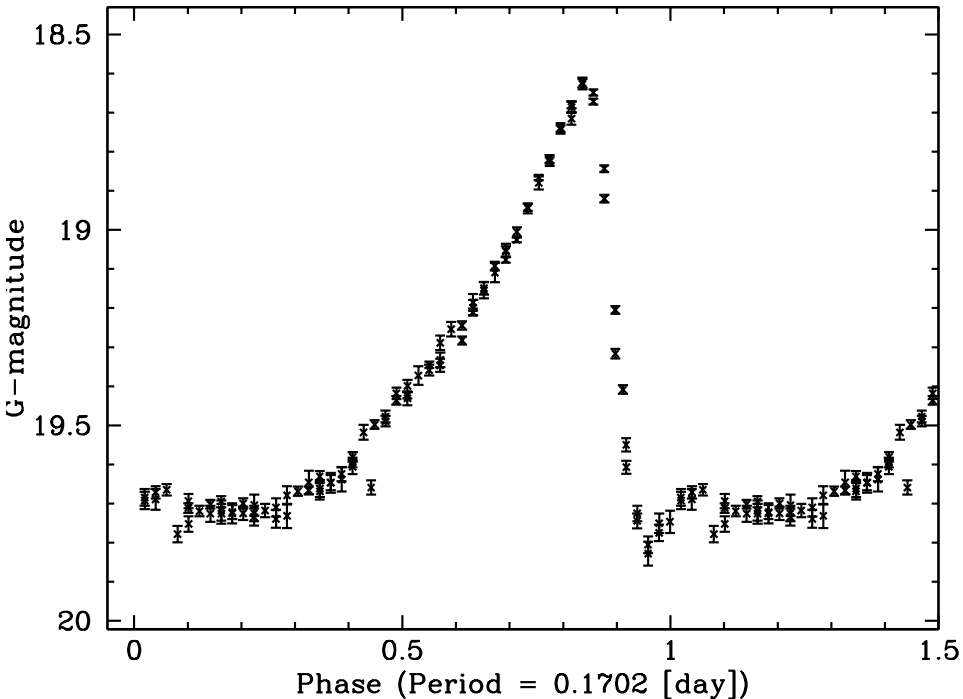




Stroboscopic effect/Wagon-wheel effect

An example of aliasing in Gaia (same star with two different periods) in EPSL (too regular sampling)





Stroboscopic effect/Wagon-wheel effect

For LSST:There are ways to deal with aliasing (it is not always a lost cause), see Dawson & Fabricki

Gaia Variability Processing and Analysis

Detect variability

Characterize variability

Classify

Classification

Supervised classification (several methods):

Classification

Supervised classification (several methods):

Multistage tree:
Bayesian networks

Multistage tree:
Gaussian mixture

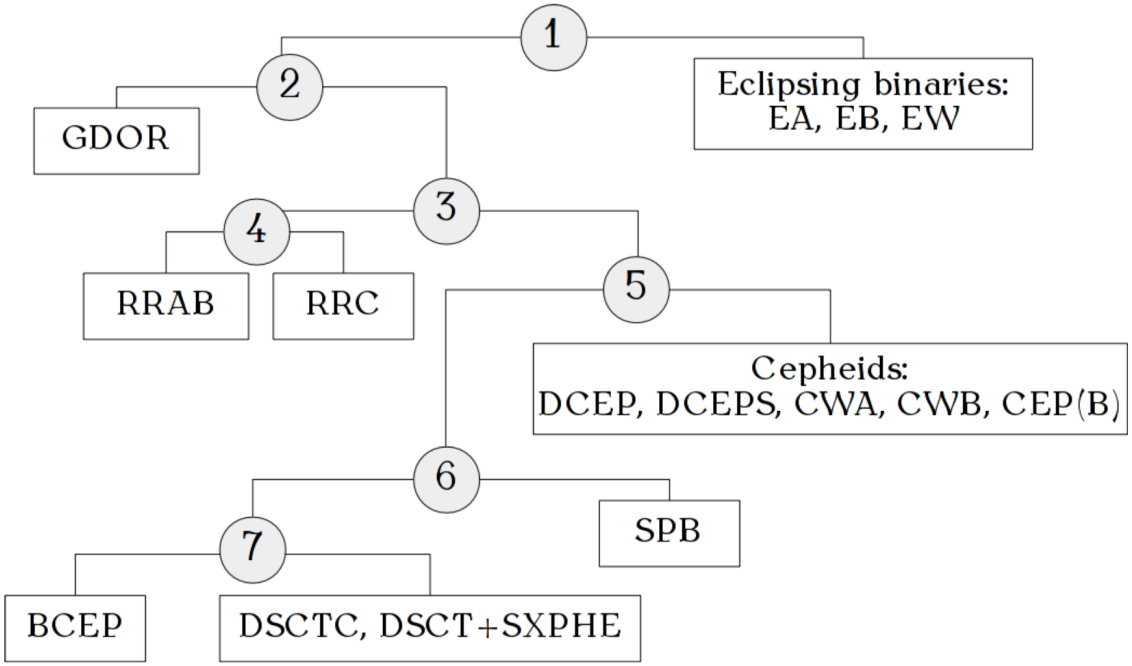
Random Forest

Classification

Supervised classification (several methods):

Multistage tree: Bayesian networks	Multistage tree: Gaussian mixture	Random Forest

Tree for Gaussian Mixture:



Classification

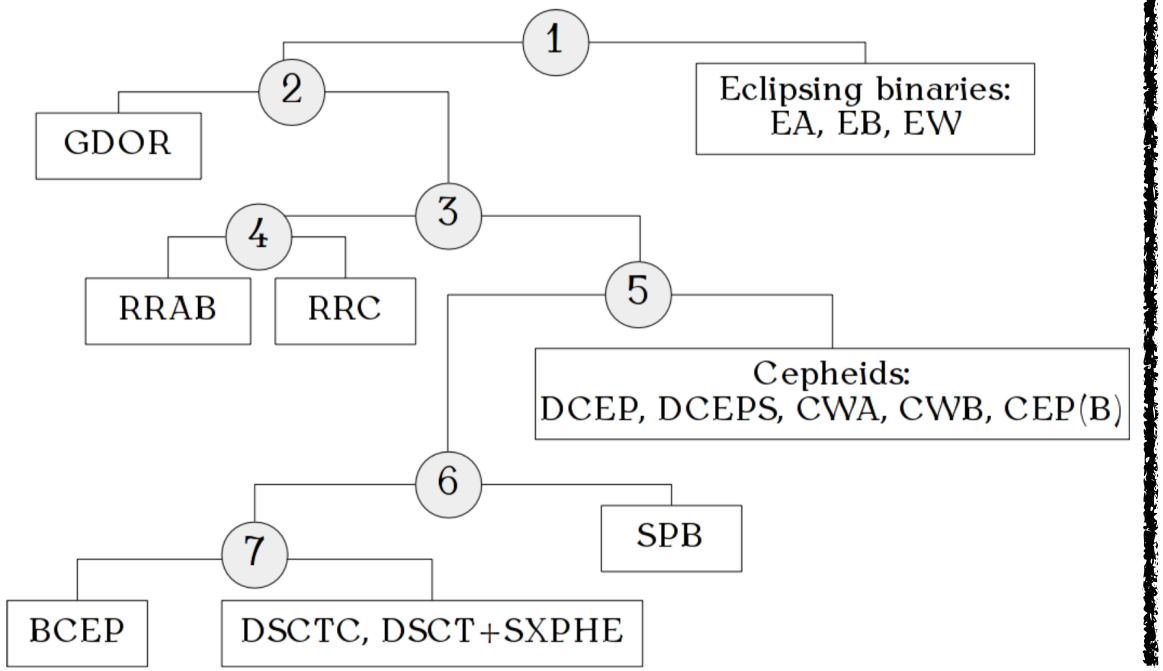
Supervised classification (several methods):

Multistage tree:
Bayesian networks

Multistage tree:
Gaussian mixture

Random Forest

Tree for Gaussian Mixture:



Furnish training set
built from Crossmatched data

Classification

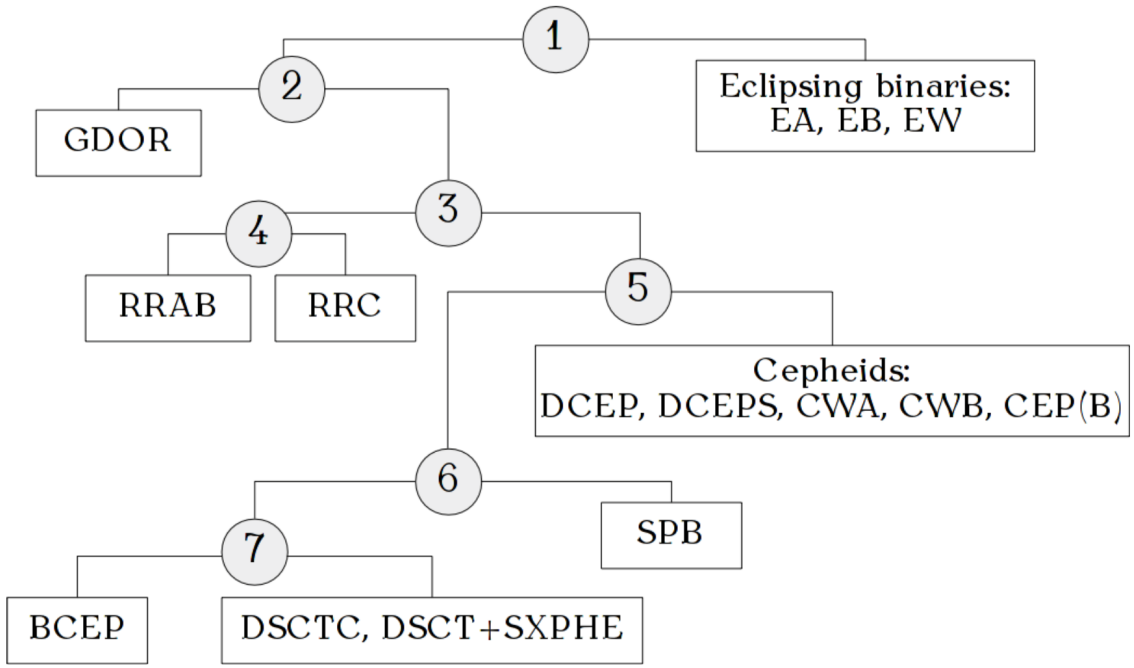
Supervised classification (several methods):

Multistage tree:
Bayesian networks

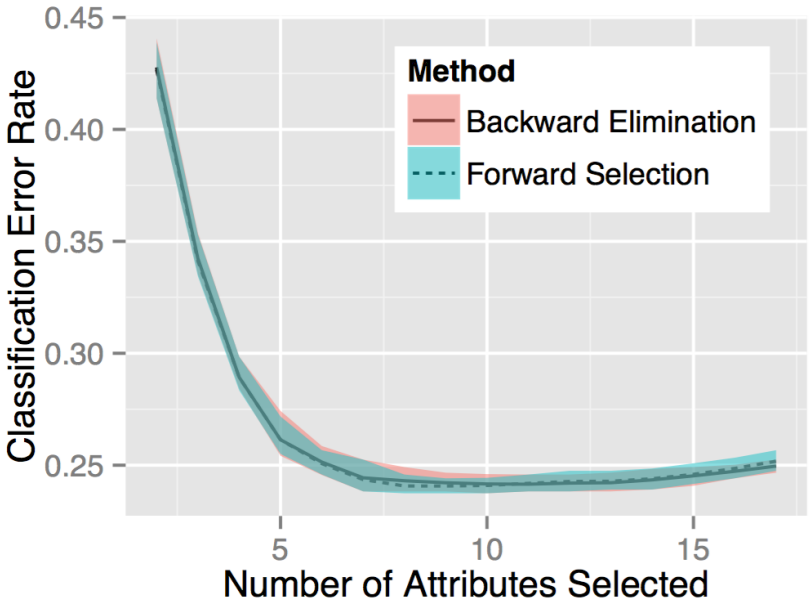
Multistage tree:
Gaussian mixture

Random Forest

Tree for Gaussian Mixture:



Furnish training set
built from Crossmatched data



Classification

Confusion matrix of Random Forest
using cross-matched data (OGLE, Hipparcos, AAVSO, Milliquas)

		CONSTANT	QSO	ECL	OTHER	RRLYR	LPV	ELL	DSCT	DCEP	UG
103	CONSTANT	79	8	5	9						
100	QSO	7	88	2	3						
135	ECL	4	3	84	7	1	1	1			
108	OTHER	13	4	12	69		2	1			
106	RRLYR	3		1	2	94					
27	LPV	30	19	7	30		15				
8	ELL	25	12		62						
7	DSCT	14		29	29	29					
2	DCEP	50		50							
2	UG					100					
Contamination		34	20	19	34	5	43	100	_	_	_

Gaia Variability Processing and Analysis

Detect variability

Characterize variability

Classify

Gaia Variability Processing and Analysis

Detect variability

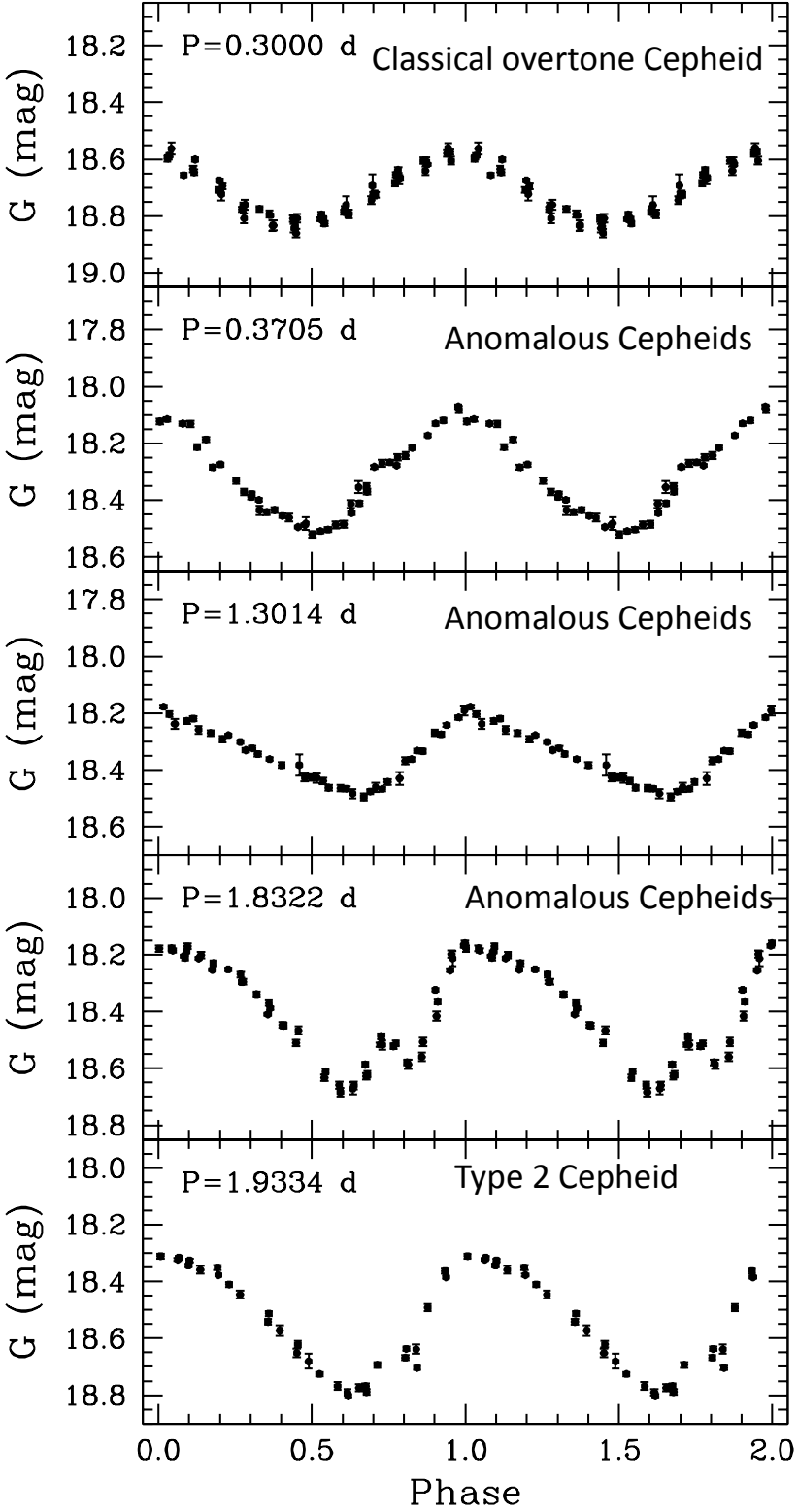
Characterize variability

Classify

Specific object Studies

Specific object studies
RR Lyrae and Cepheid stars

- Verify the classification
- Derive specific parameters
- Perform a sub-classification
- Give a feedback to the classification

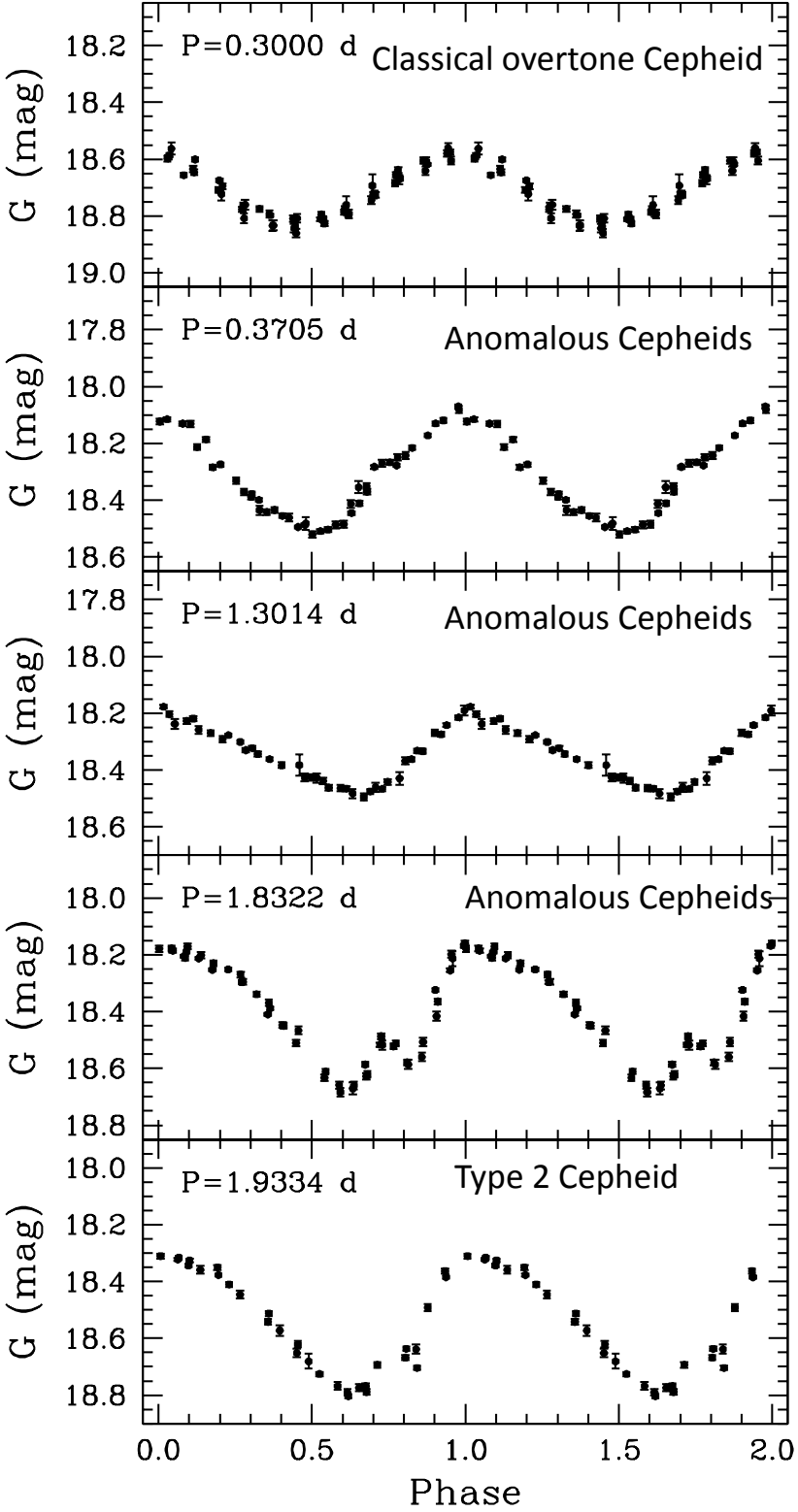


Credits: ESA/Gaia/DPAC/CU5/DPCI/CU7/INAF-OABo/INAF-OACn Gisella Clementini, Vincenzo Ripepi, Silvio Leccia, Laurent Eyer, Lorenzo Rimoldini, Isabelle Lecoeur-Taibi, Nami Mowlavi, Dafydd Evans, Geneva CU7/DPCG and the whole CU7 team. The photometric data reduction was done with the PhotPipe pipeline at DPCI; processing data were received from the IDT pipeline at DPCE.

Specific object studies
RR Lyrae and Cepheid stars

- Verify the classification
- Derive specific parameters
- Perform a sub-classification
- Give a feedback to the classification

"Active learning"



Credits: ESA/Gaia/DPAC/CU5/DPCI/CU7/INAF-OABo/INAF-OACn Gisella Clementini, Vincenzo Ripepi, Silvio Leccia, Laurent Eyer, Lorenzo Rimoldini, Isabelle Lecoeur-Taibi, Nami Mowlavi, Dafydd Evans, Geneva CU7/DPCG and the whole CU7 team. The photometric data reduction was done with the PhotPipe pipeline at DPCI; processing data were received from the IDT pipeline at DPCE.

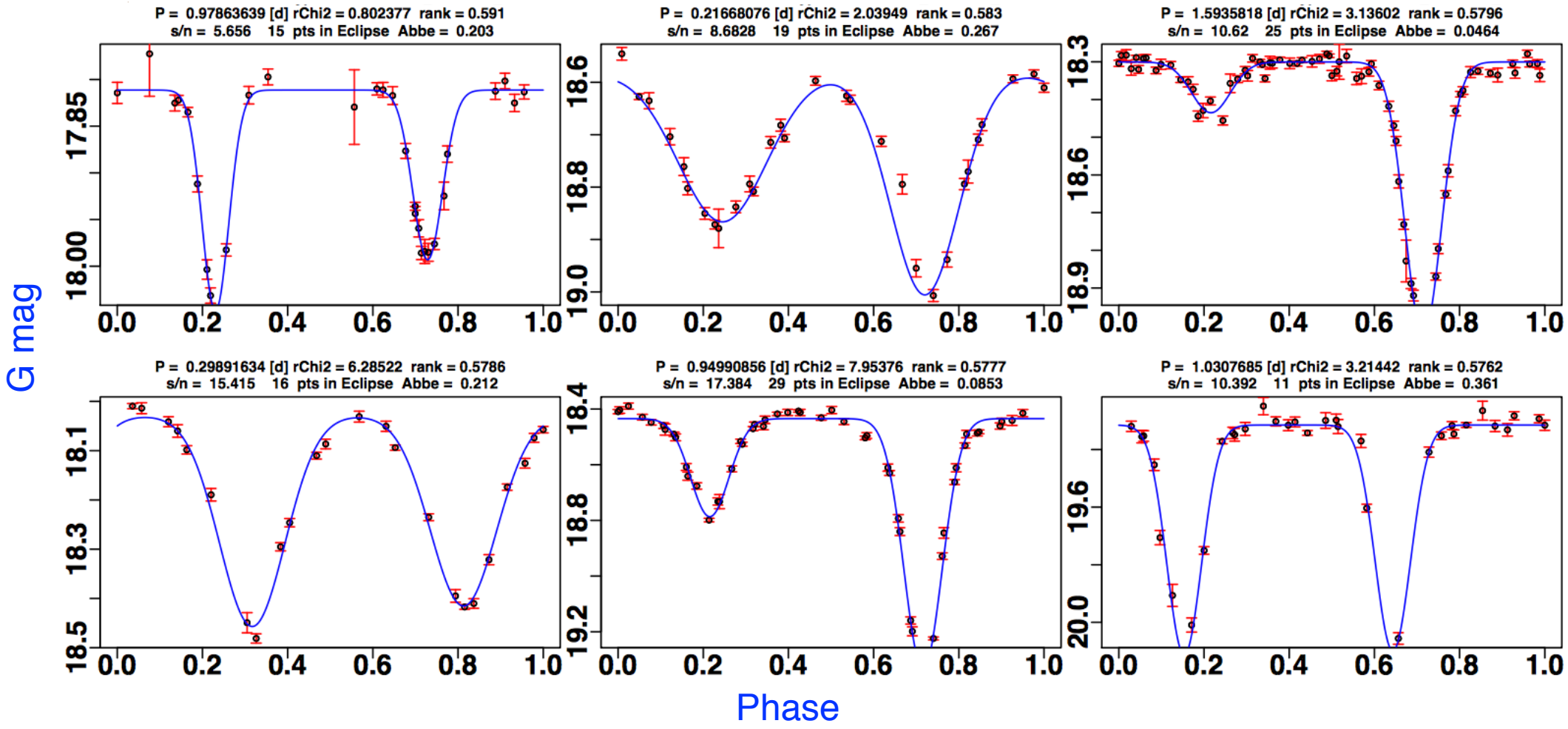
Specific Object Studies: An other example

Eclipsing binaries

Eclipsing binaries go to a dedicate treatment (Université Libre de Bruxelles) for a full modelling

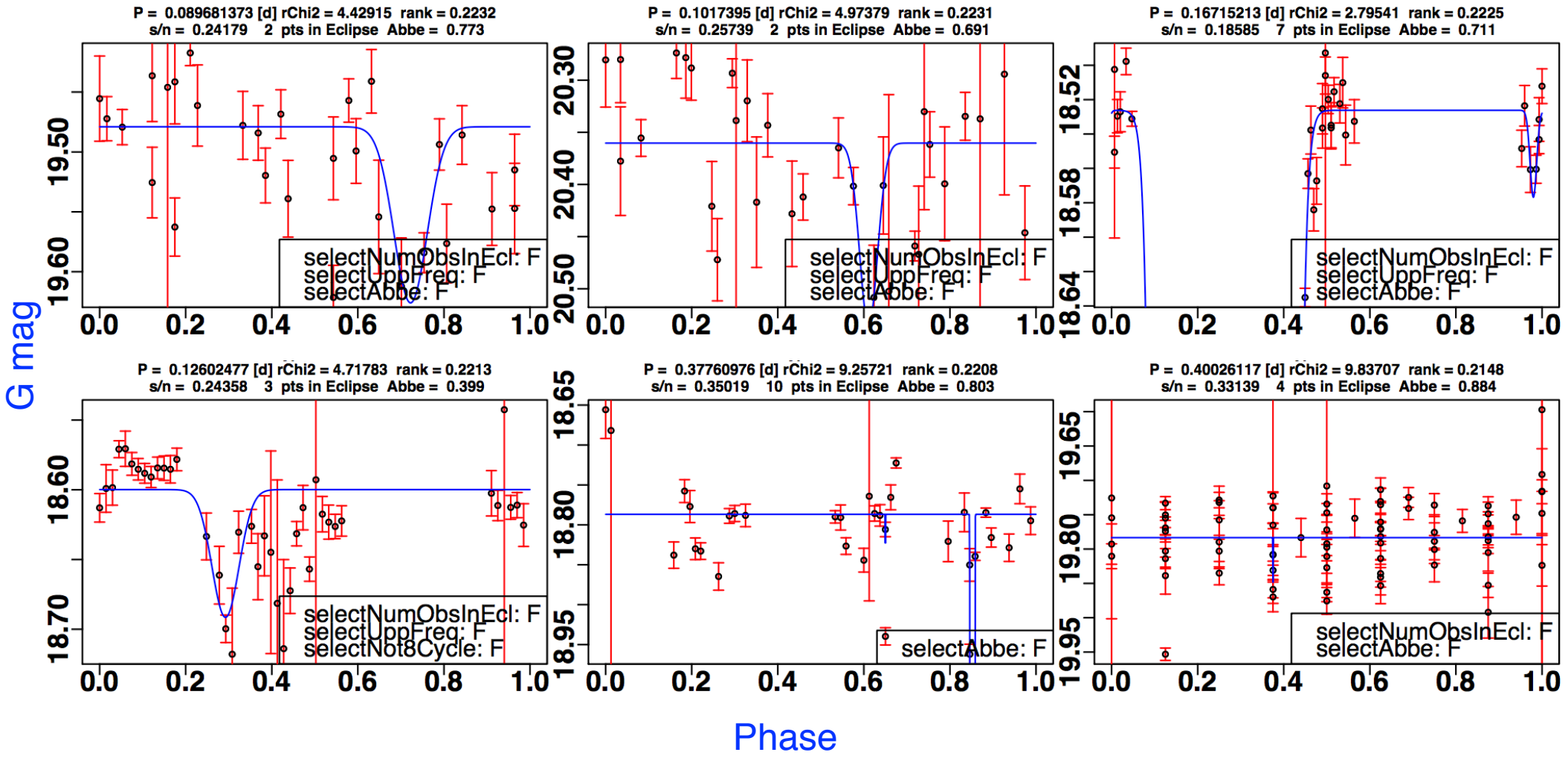
Here, simple modelling (fit with “two Gaussians”) are made
The solution enables a ranking

Highest rank



Specific Object Studies: Eclipsing binaries

Lowest rank



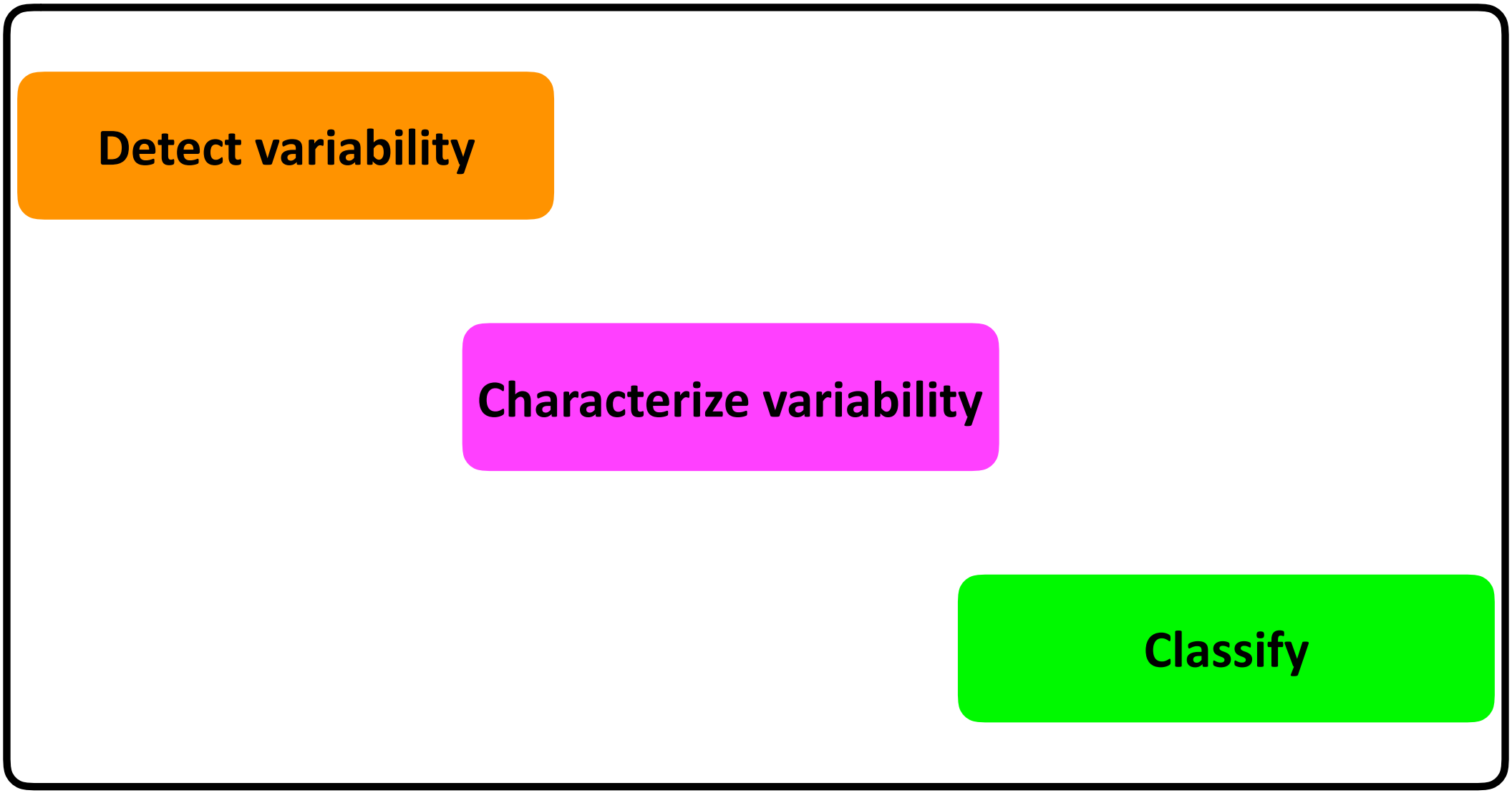
Gaia Variability Processing and Analysis

Detect variability

Characterize variability

Classify

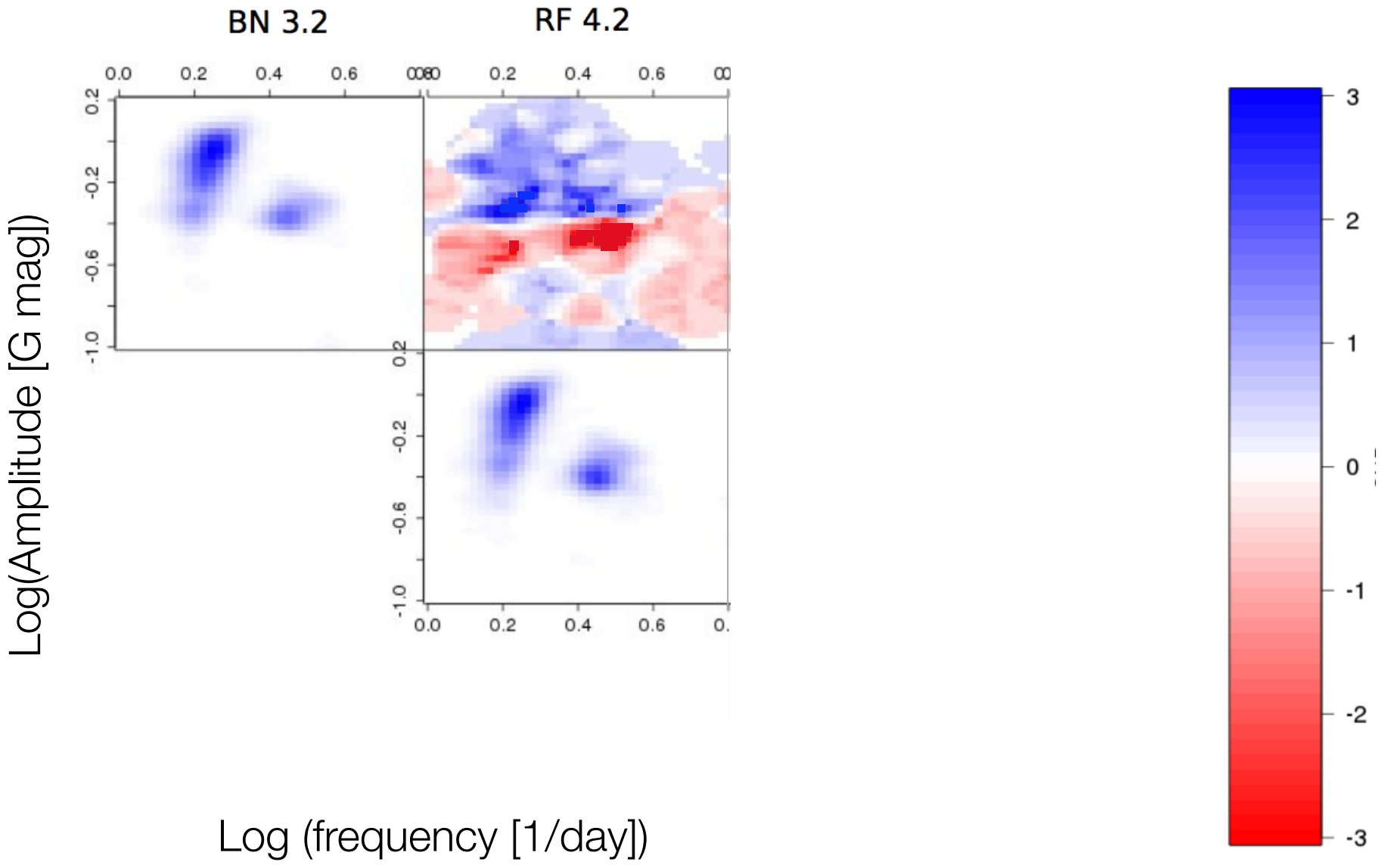
Gaia Variability Processing and Analysis



Global Variability Studies

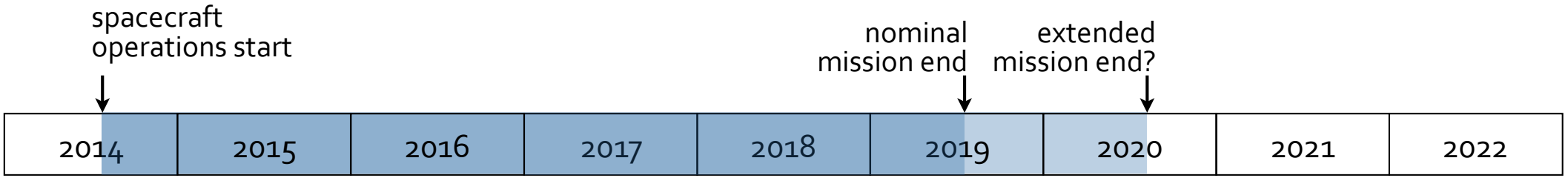
Global Variability Studies

Comparison of distribution functions of RR Lyrae stars



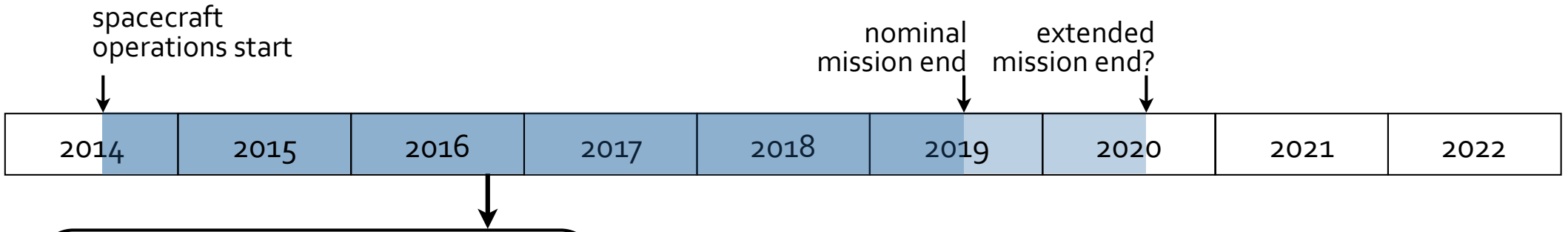
Courtesy of L. Sarro (modified by L.Eyer)

The data releases



Courtesy of B.Holl

The data releases

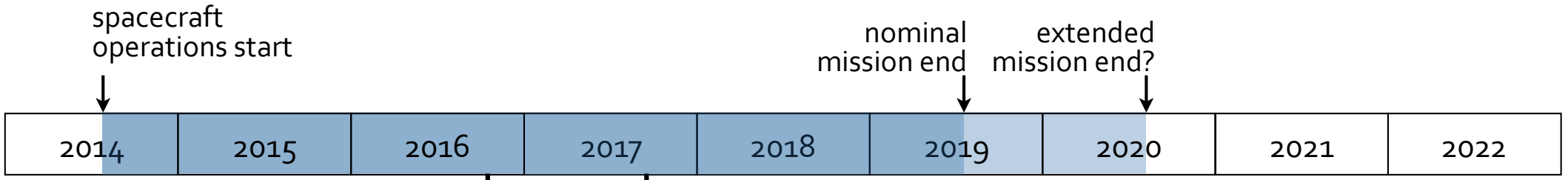


Release 1:

- **α and δ, mean G-magnitude for 1billion stars**
- **100K proper motion stars (Hipparcos+Gaia)**

Courtesy of B.Holl

The data releases



Release 1:

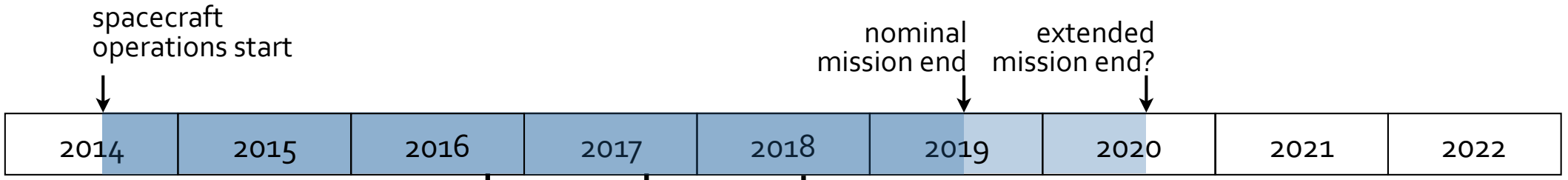
- **α and δ, mean G-magnitude for 1billion stars**
- **100K proper motion stars (Hipparcos+Gaia)**

Release 2:

- **5-parameter astrometric solutions for single star (parallax)**
- **Integrated BP/RP + Astrophysical parameters**
- **Mean V_{rad} (for non variable)**

Courtesy of B.Holl

The data releases



Release 1:

- **α and δ, mean G-magnitude for 1billion stars**
- **100K proper motion stars (Hipparcos+Gaia)**

Release 2:

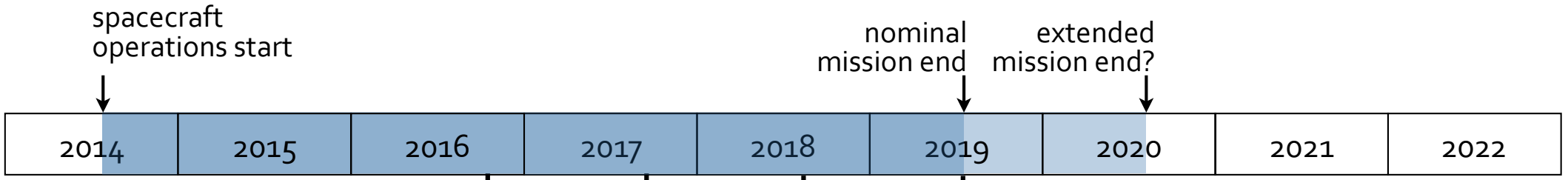
- **5-parameter astrometric solutions for single star (parallax)**
- **Integrated BP/RP + Astrophysical parameters**
- **Mean V_{rad} (for non variable)**

Release 3:

- Mean V_{rad}
- Object classifications and Astrophysical Parameters
- **Mean RVS, BP, RP spectra**
- **Orbital solution of binaries**

Courtesy of B.Holl

The data releases



Release 1:

- **α and δ, mean G-magnitude for 1billion stars**
- **100K proper motion stars (Hipparcos+Gaia)**

Release 4:

- **Variable stars classification**
- **Non-single star catalogue**
- **Solar system objects**

Release 2:

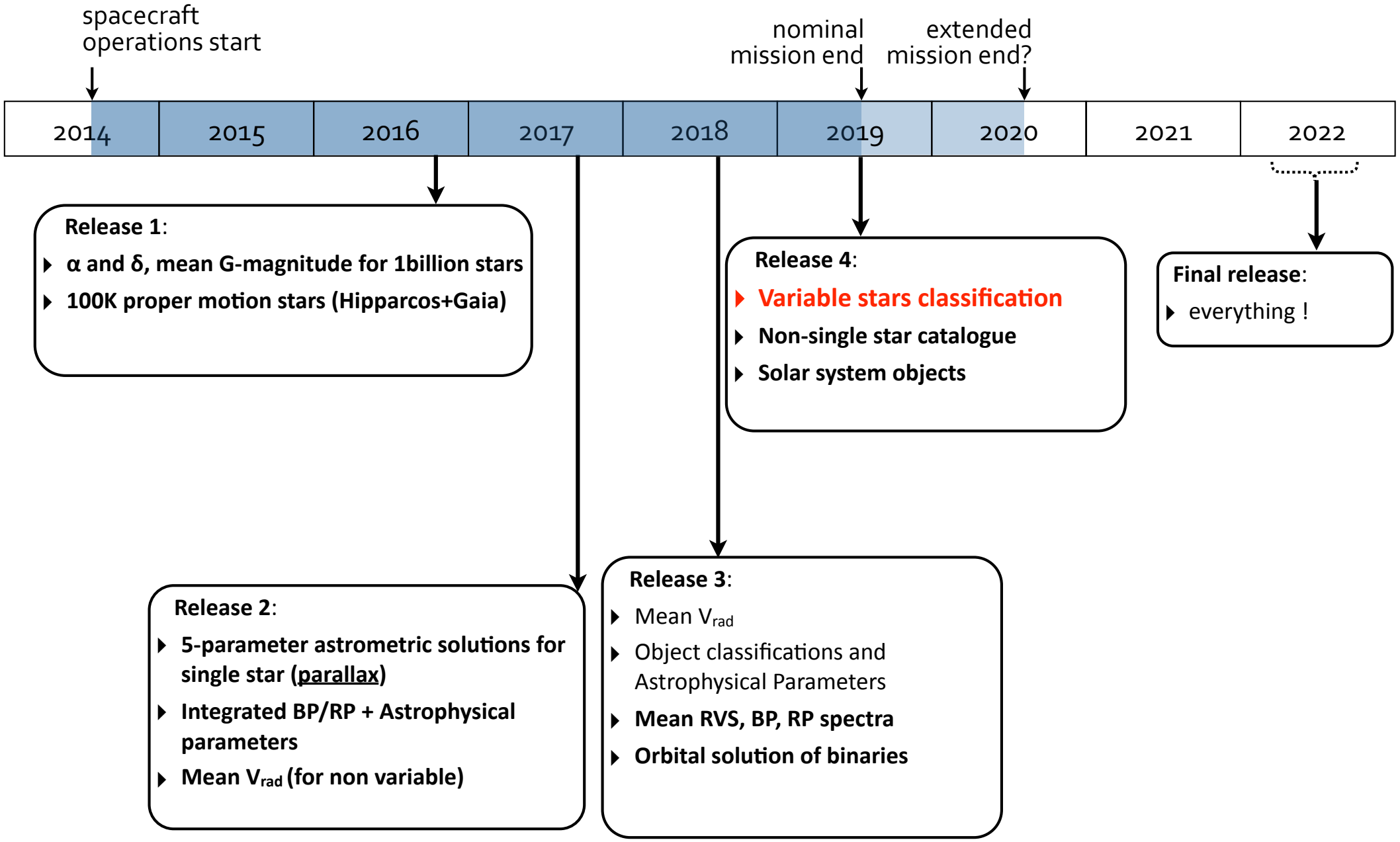
- **5-parameter astrometric solutions for single star (parallax)**
- **Integrated BP/RP + Astrophysical parameters**
- **Mean V_{rad} (for non variable)**

Release 3:

- Mean V_{rad}
- Object classifications and Astrophysical Parameters
- **Mean RVS, BP, RP spectra**
- **Orbital solution of binaries**

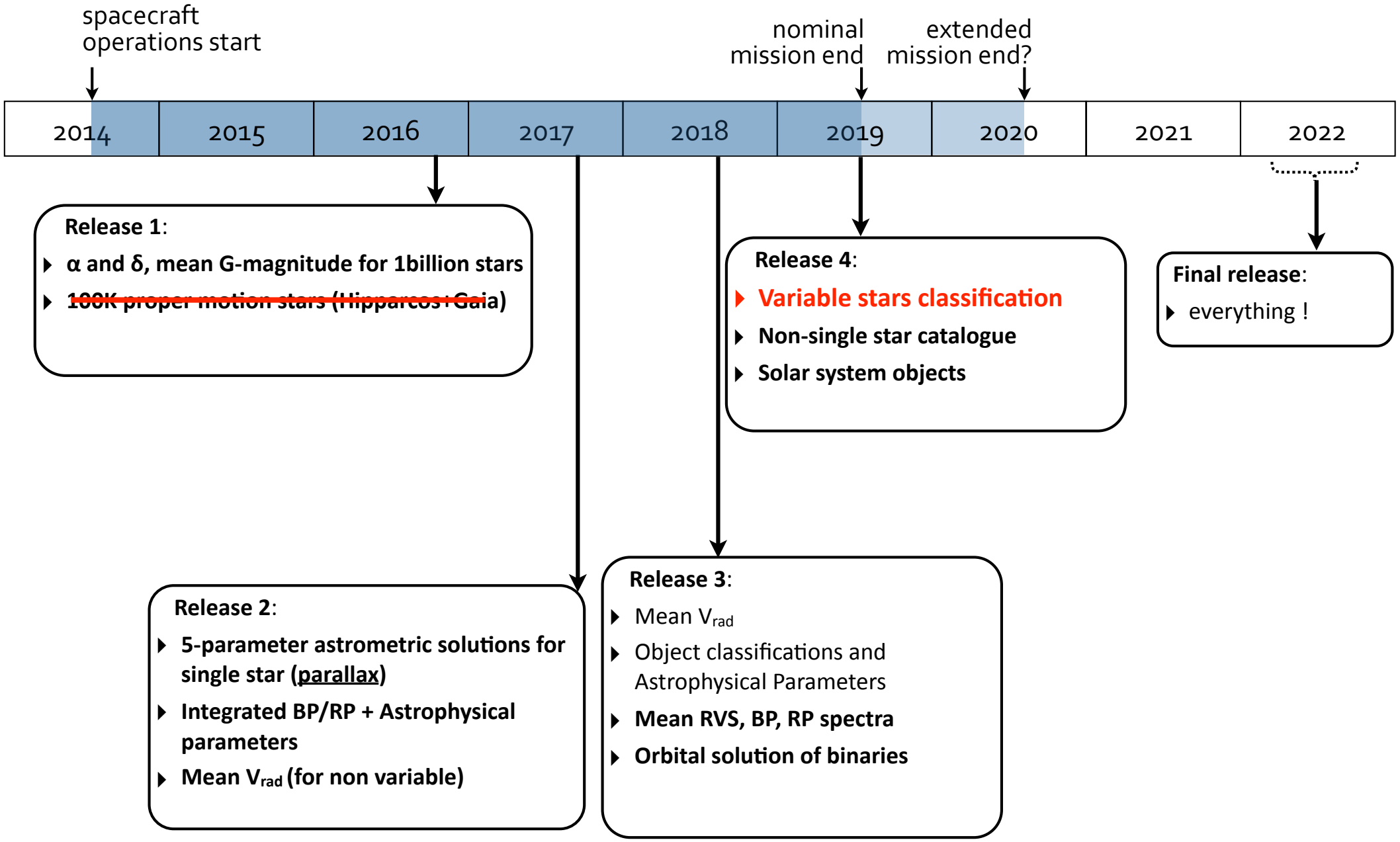
Courtesy of B.Holl

The data releases



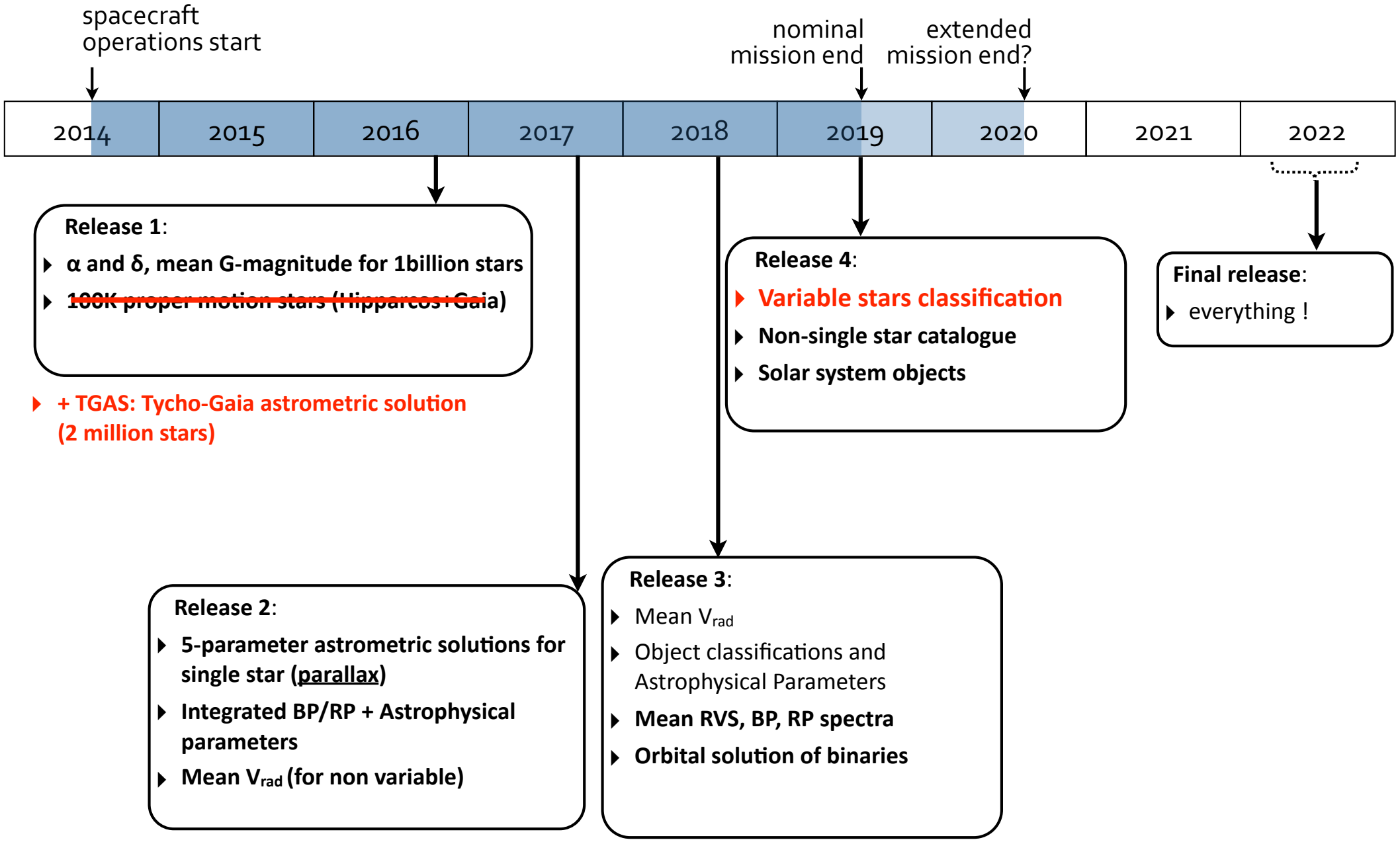
Courtesy of B.Holl

The data releases



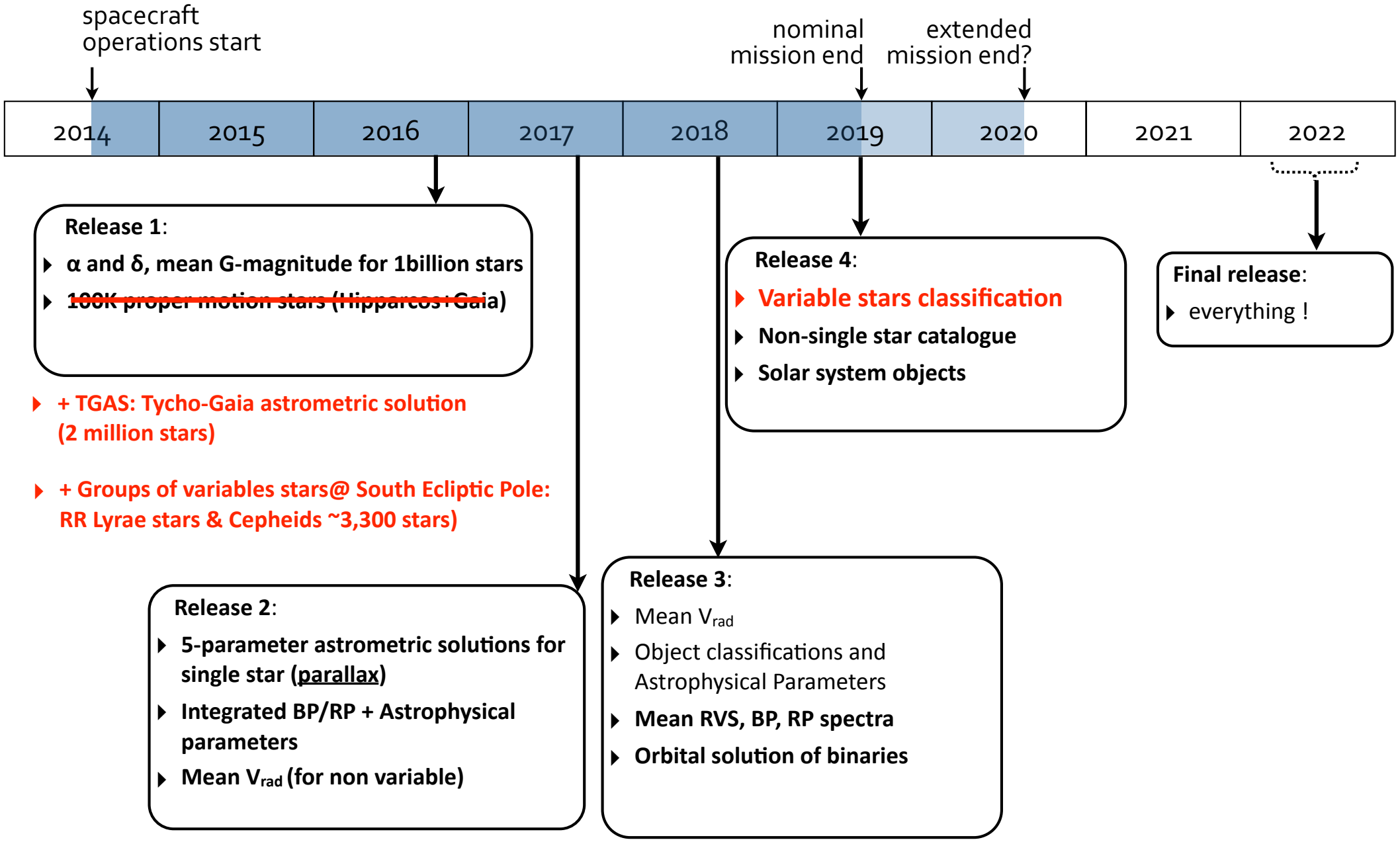
Courtesy of B.Holl

The data releases



Courtesy of B.Holl

The data releases



Courtesy of B.Holl

Thank you for
your attention