Variability Analysis and Processing of Gaia EPSL data

Laurent Eyer on behalf of Coordination Unit 7

Liverpool, UK

Tuesday November 11, 2015

The Gaia multi-epoch survey

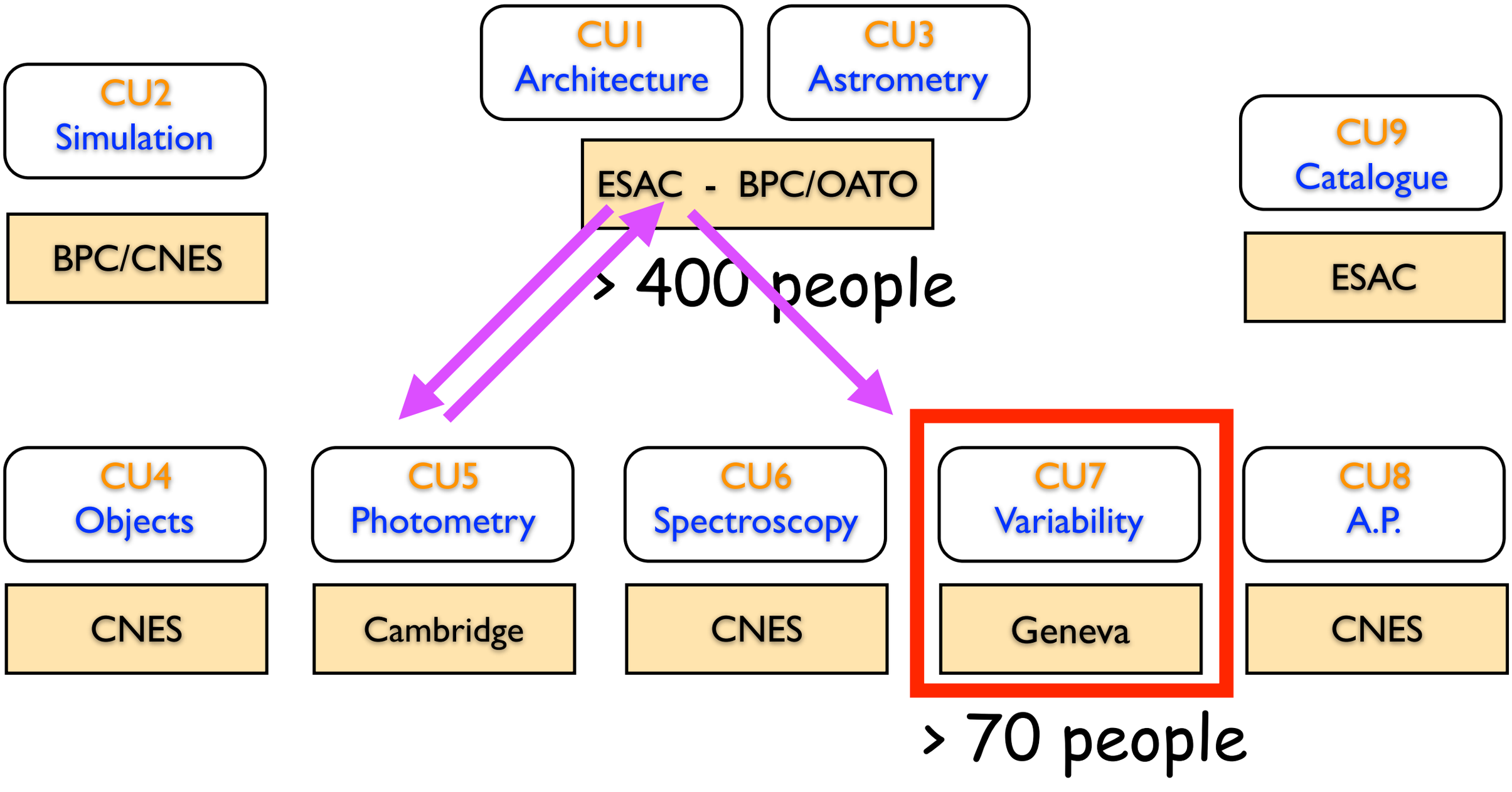
Gaia is an exceptional survey, because:

- Gaia furnishes exquisite parallaxes
- Gaia surveys the entire sky with one set of instruments (**G** photometry, **BP**, **RP** spectrophotometry, **RVS** spectrometry)
- Gaia performs nearly simultaneous measurements in these different instruments
- Gaia measures the “bright sky”
- Gaia has a peculiar sampling

This talk is “only” about G magnitude

Gaia Data Processing and Analysis Consortium DPAC

Two concepts: 1. Coordination Units (CU)
2. Data Processing Centers (DPC)



Operation Rehearsals

Satellite has been in operation (since July 2014)

Variability Processing and Analysis is **not** "in operation" **(yet)**

We have been training…. on

— "Gaia" Simulated data
— some real data (Hipparcos, OGLE, EROS)

And in the beginning of 2015 on

— **Real Gaia Ecliptic Pole Scanning Law data**

This exercise/training was called **Operation Rehearsal (OR5 stage 2)**

Official goal of the operation rehearsal: not to get science results
Goal is to turn the crank for the software

Variability Processing and Analysis: about 590,000 lines of Java-R-SQL-XML-XSD code

The Operation Rehearsal data

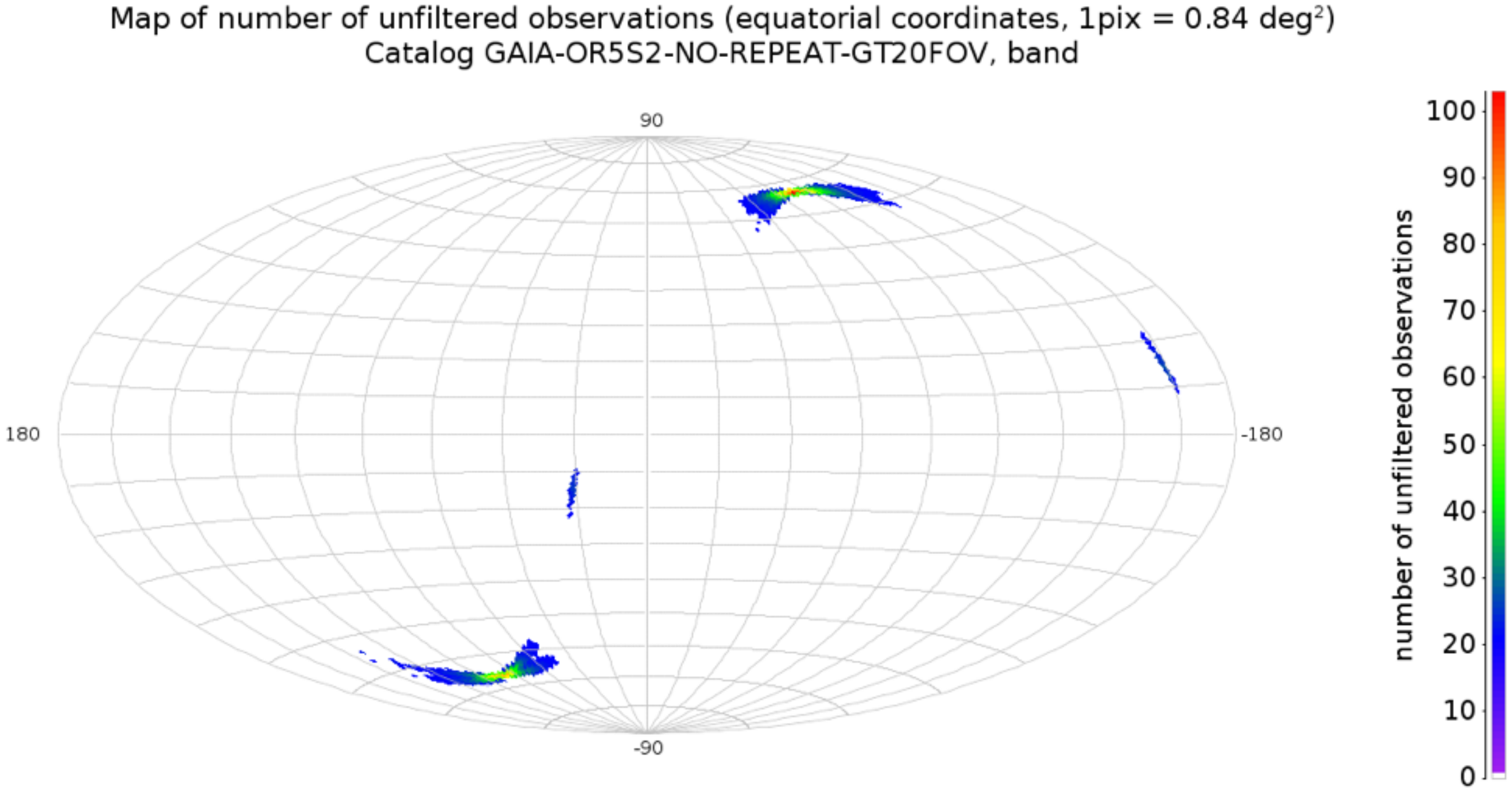
Data set {
- 28 days of Ecliptic Scanning law
- 3 days of Nominal Scanning law

69 million sources
received from Photometric Processing (Cambridge University)

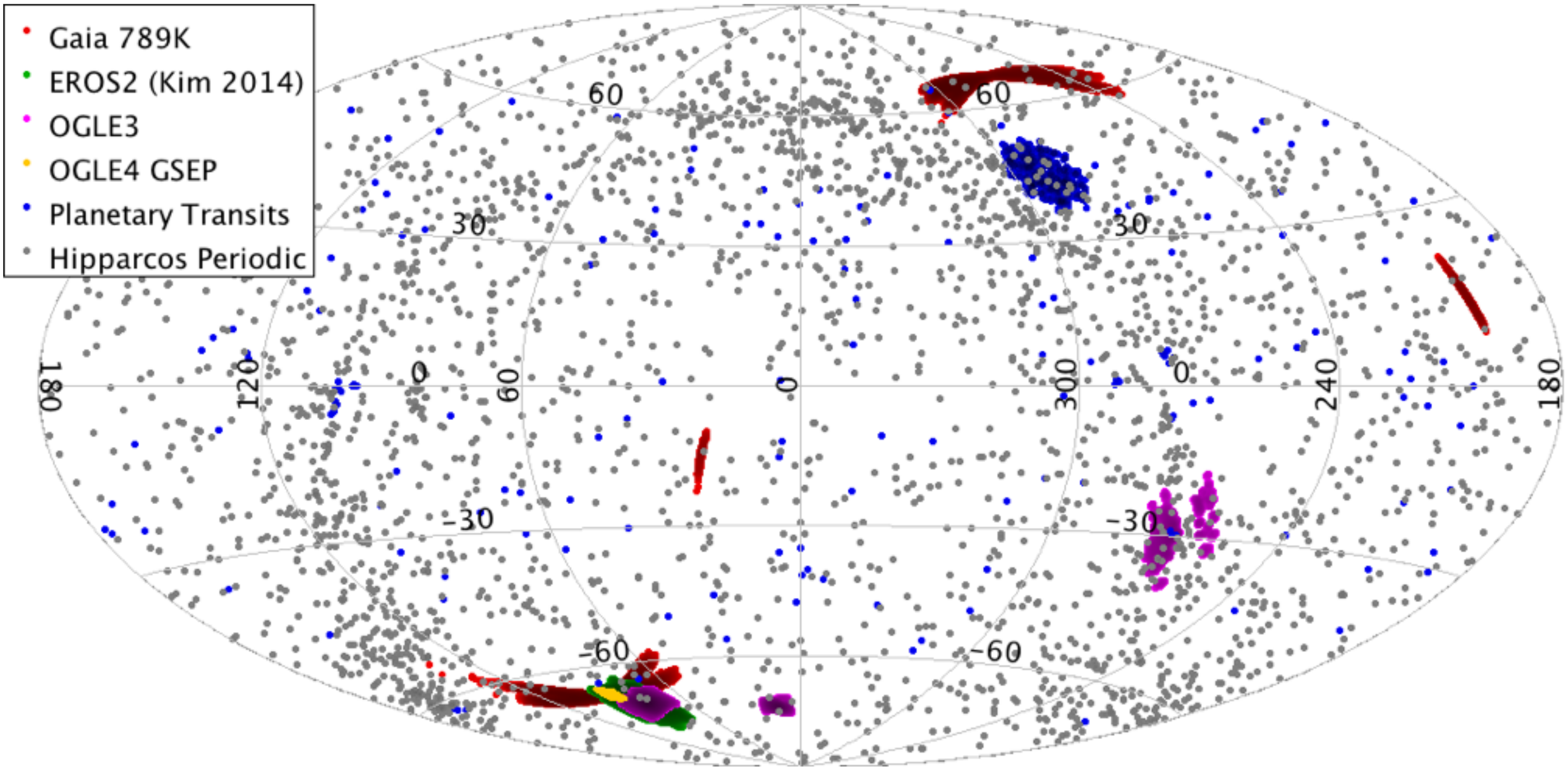
Selection:
20 measurements in either G, BP, RP
No repeated observations within 100 minutes

790,000 sources

Mean number of obs/source (789K)



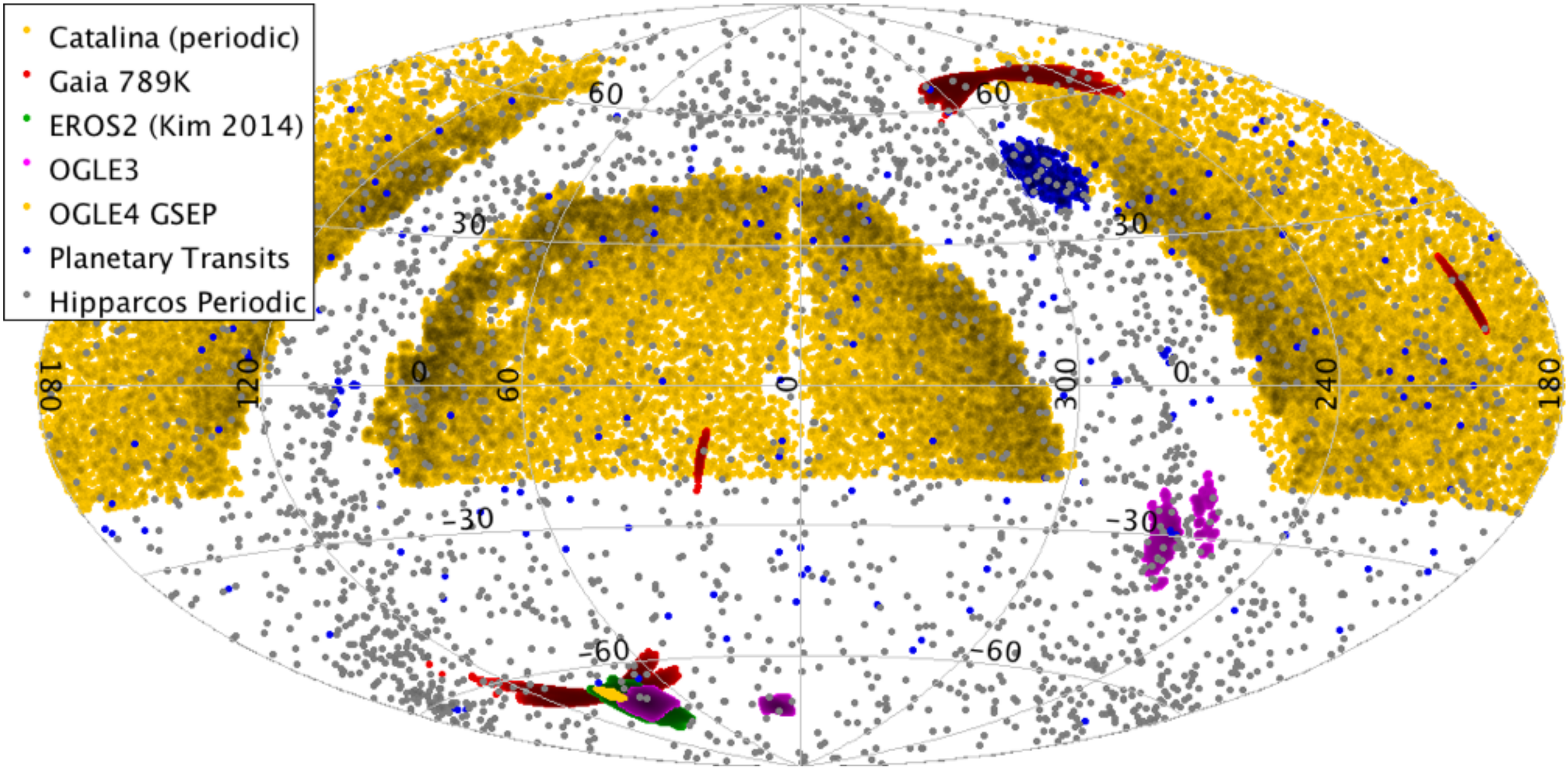
Gaia 789K + EROS2, OGLE, Hipparcos, Planetary Transits



(Equatorial coordinates, deg)

Courtesy of L.Rimoldini

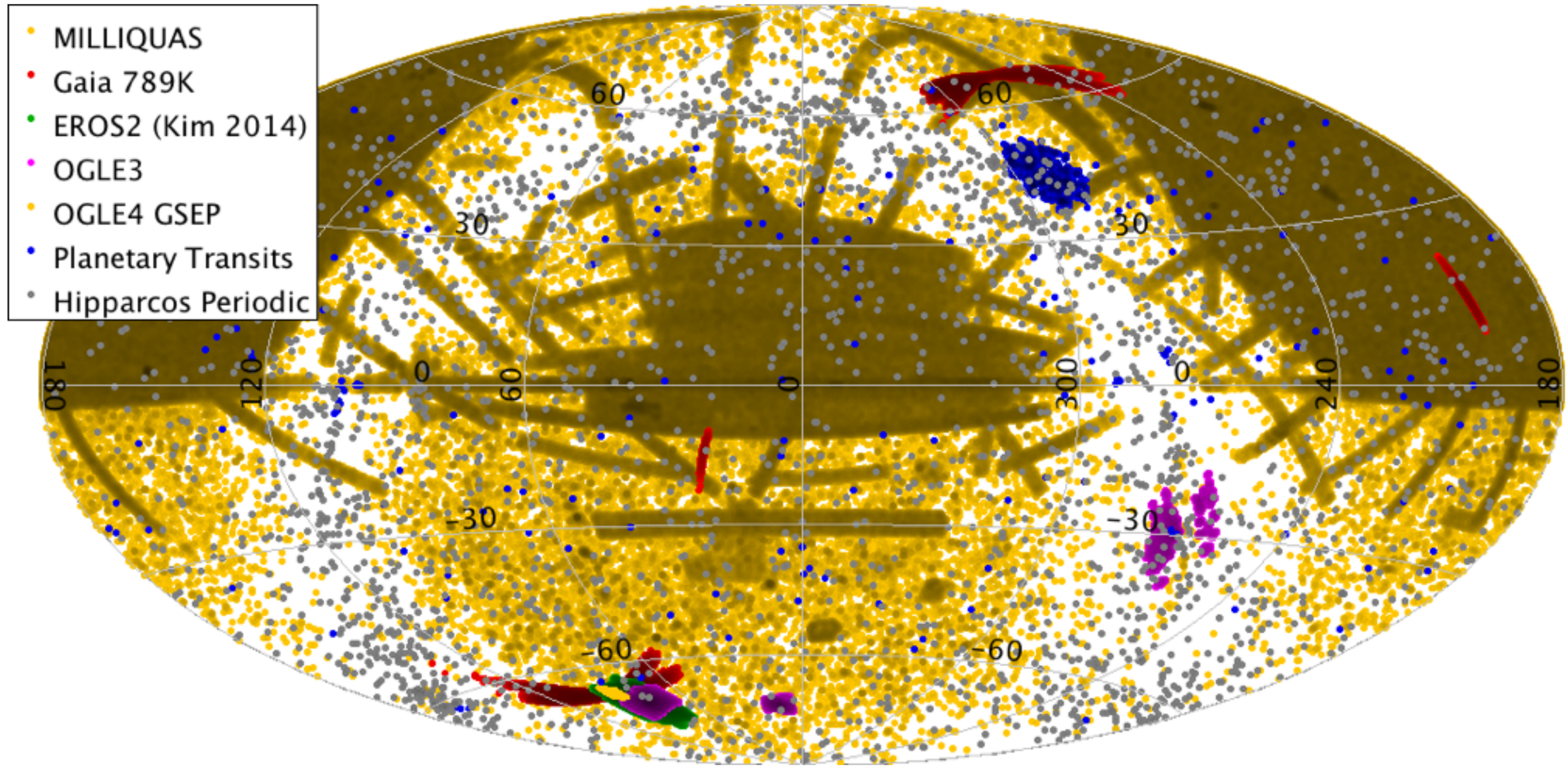
+ Catalina (periodic)



(Equatorial coordinates, deg)

Courtesy of L.Rimoldini

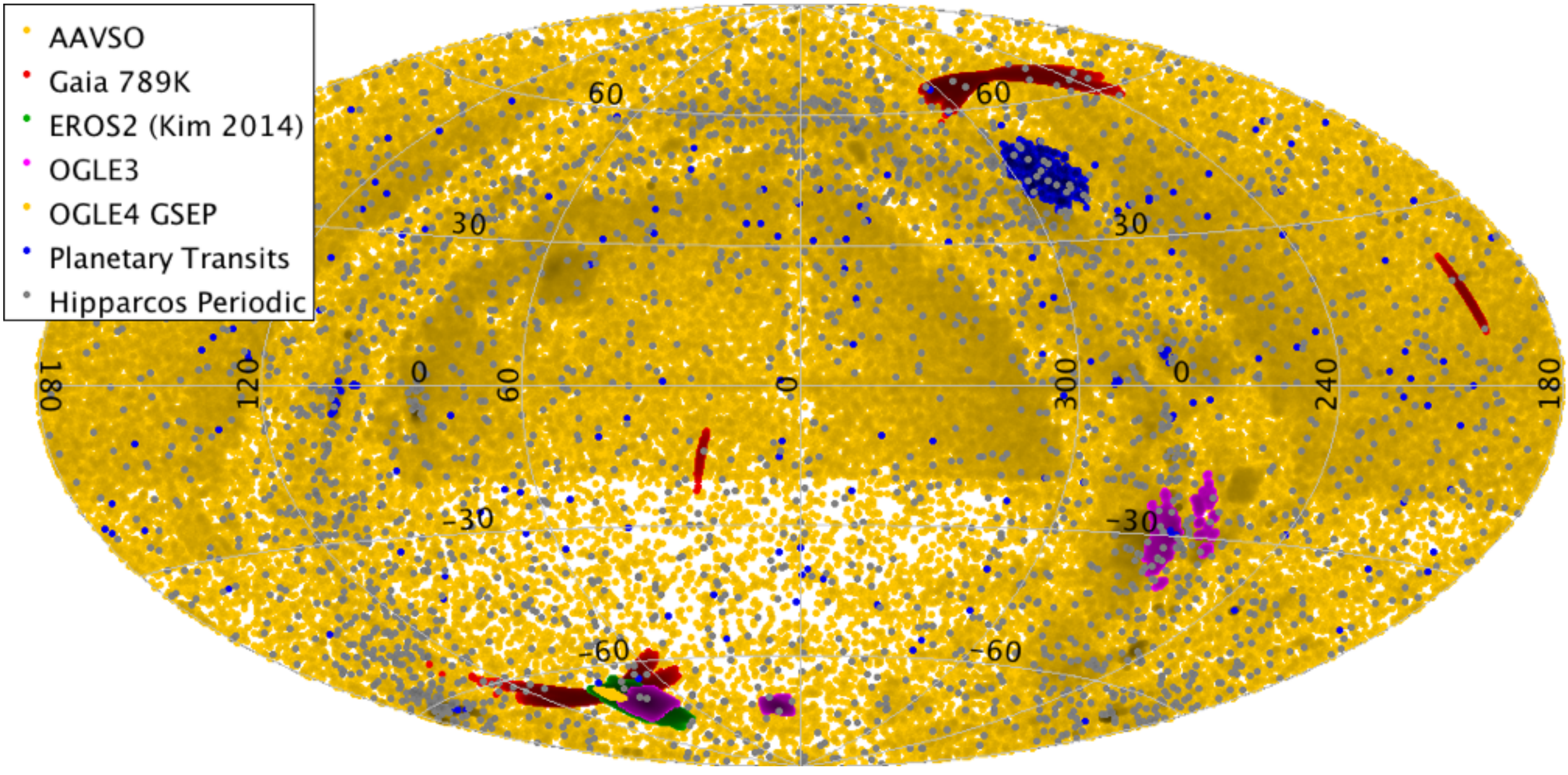
+ MILLIQUAS



(Equatorial coordinates, deg)

Courtesy of L.Rimoldini

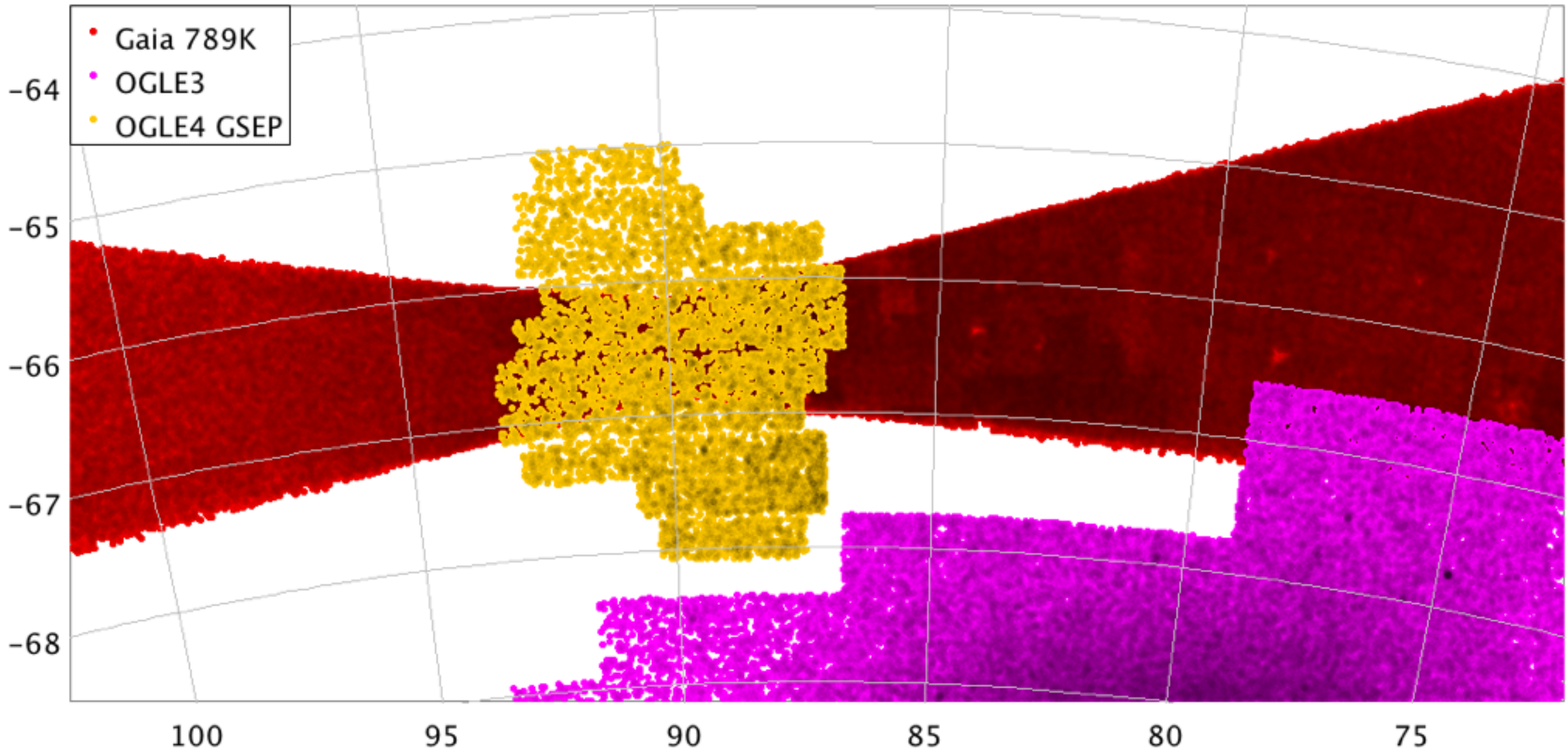
+ AAVSO



(Equatorial coordinates, deg)

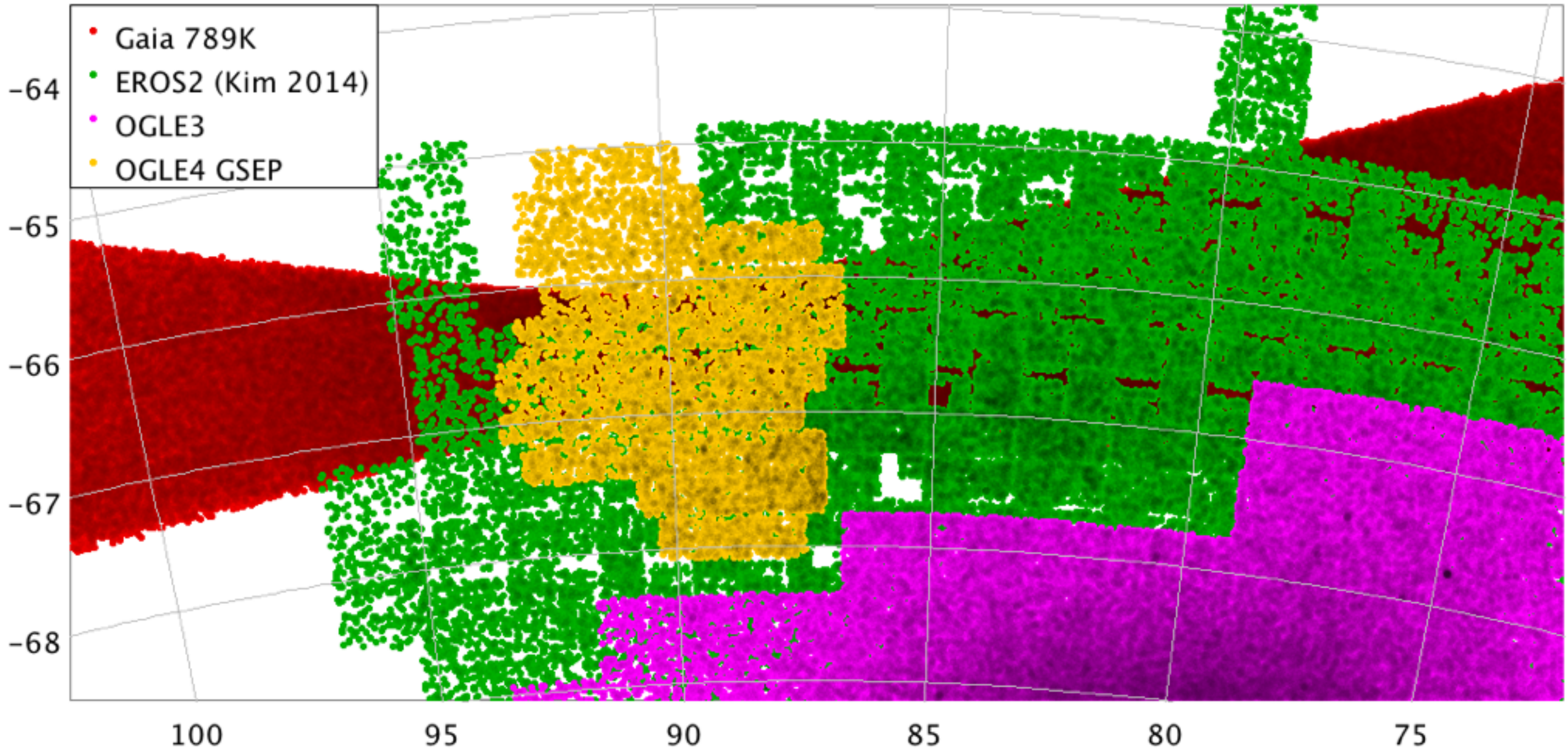
Courtesy of L.Rimoldini

SEP: Gaia + OGLE



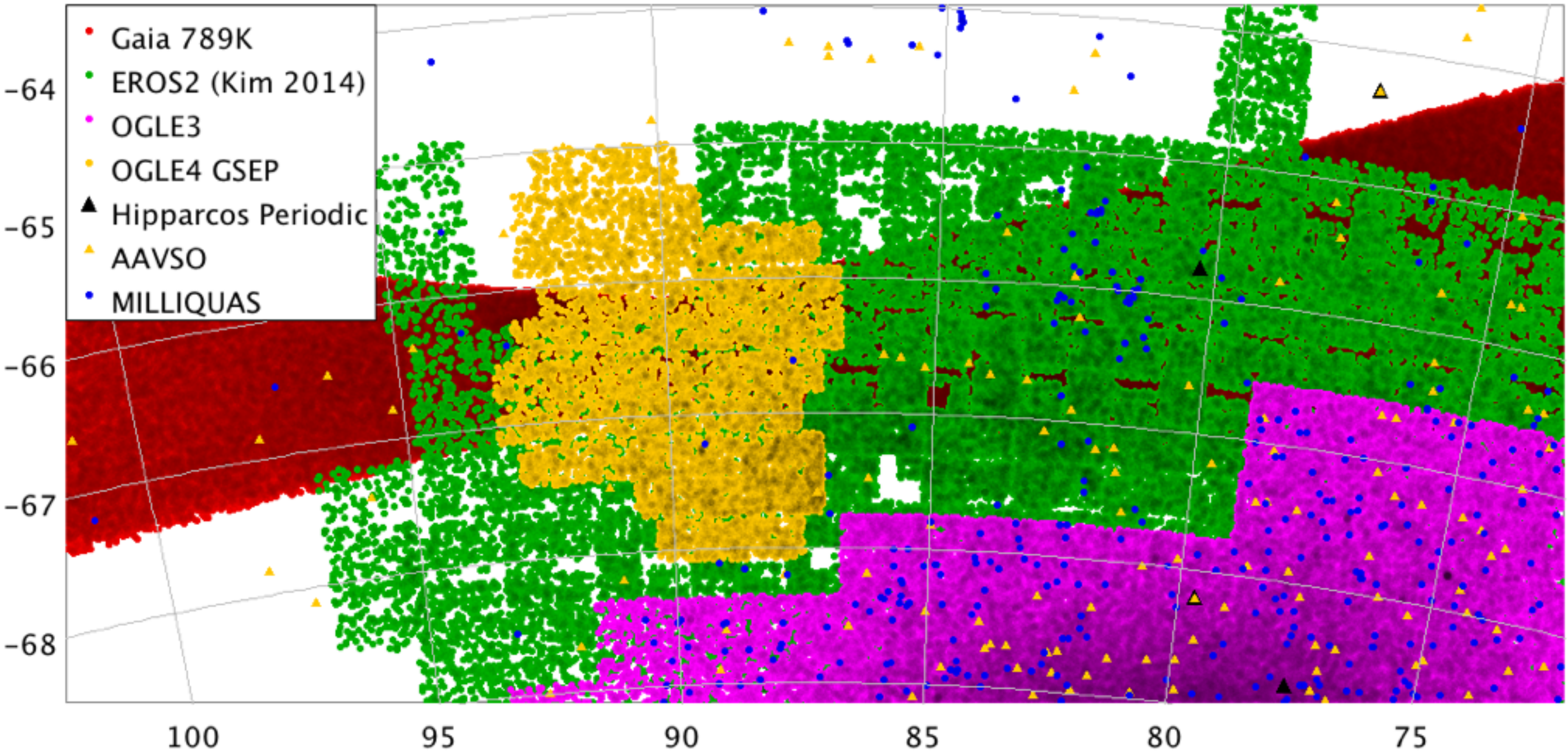
Courtesy of L.Rimoldini

SEP: Gaia + OGLE, EROS2



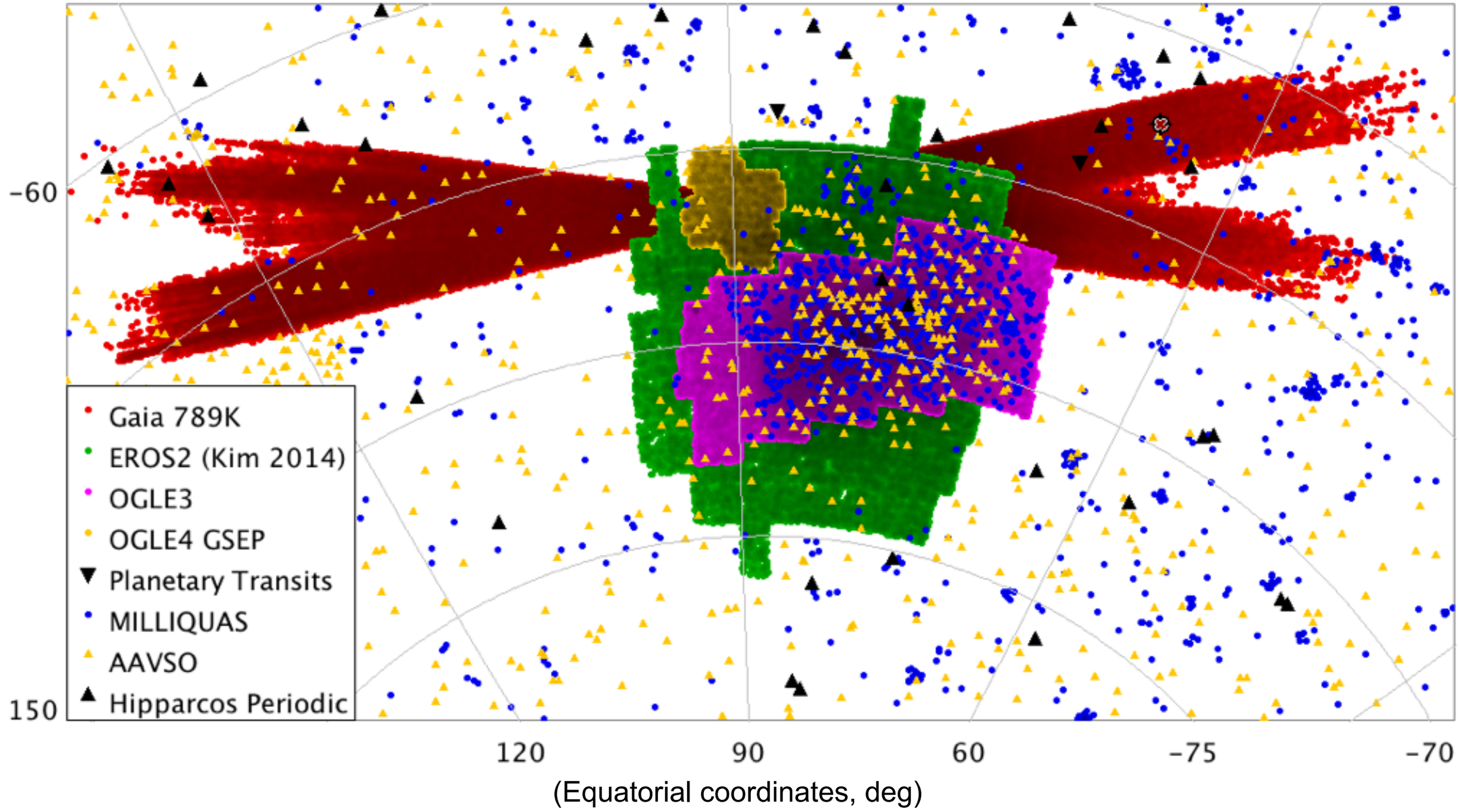
Courtesy of L.Rimoldini

SEP: Gaia + OGLE, EROS2, Hipparcos, AAVSO, MILLIQUAS



Courtesy of L.Rimoldini

SEP: Gaia + OGLE, EROS2, Hipparcos, AAVSO, MILLIQUAS, Planetary Transits

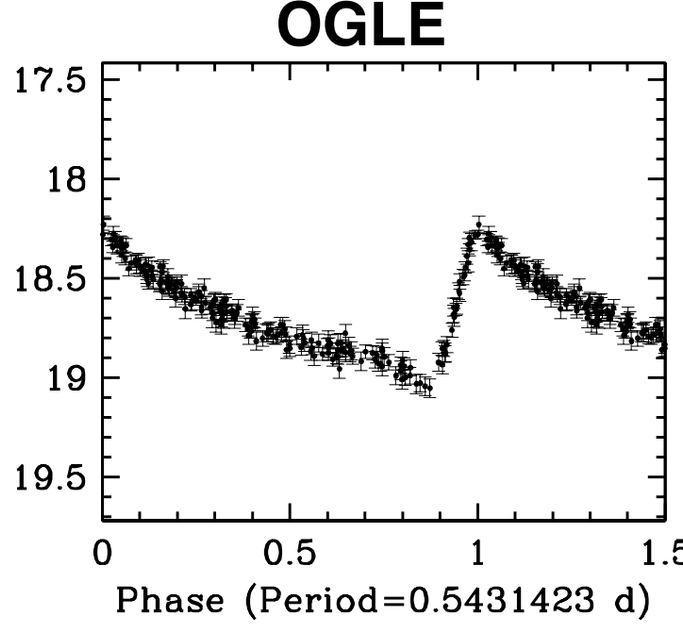


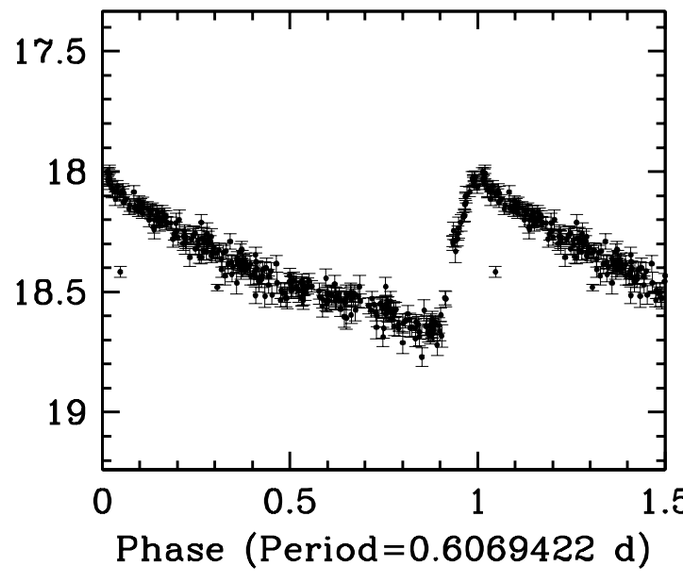
Courtesy of L.Rimoldini

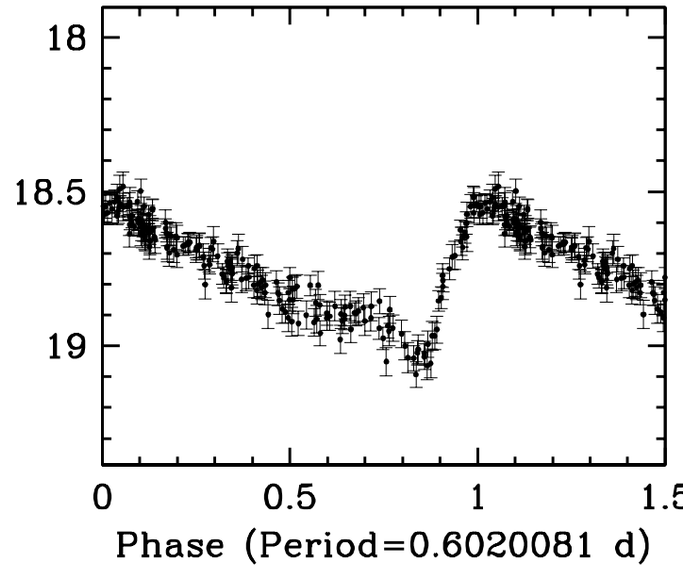
To get the data flavour
Comparison with OGLE

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

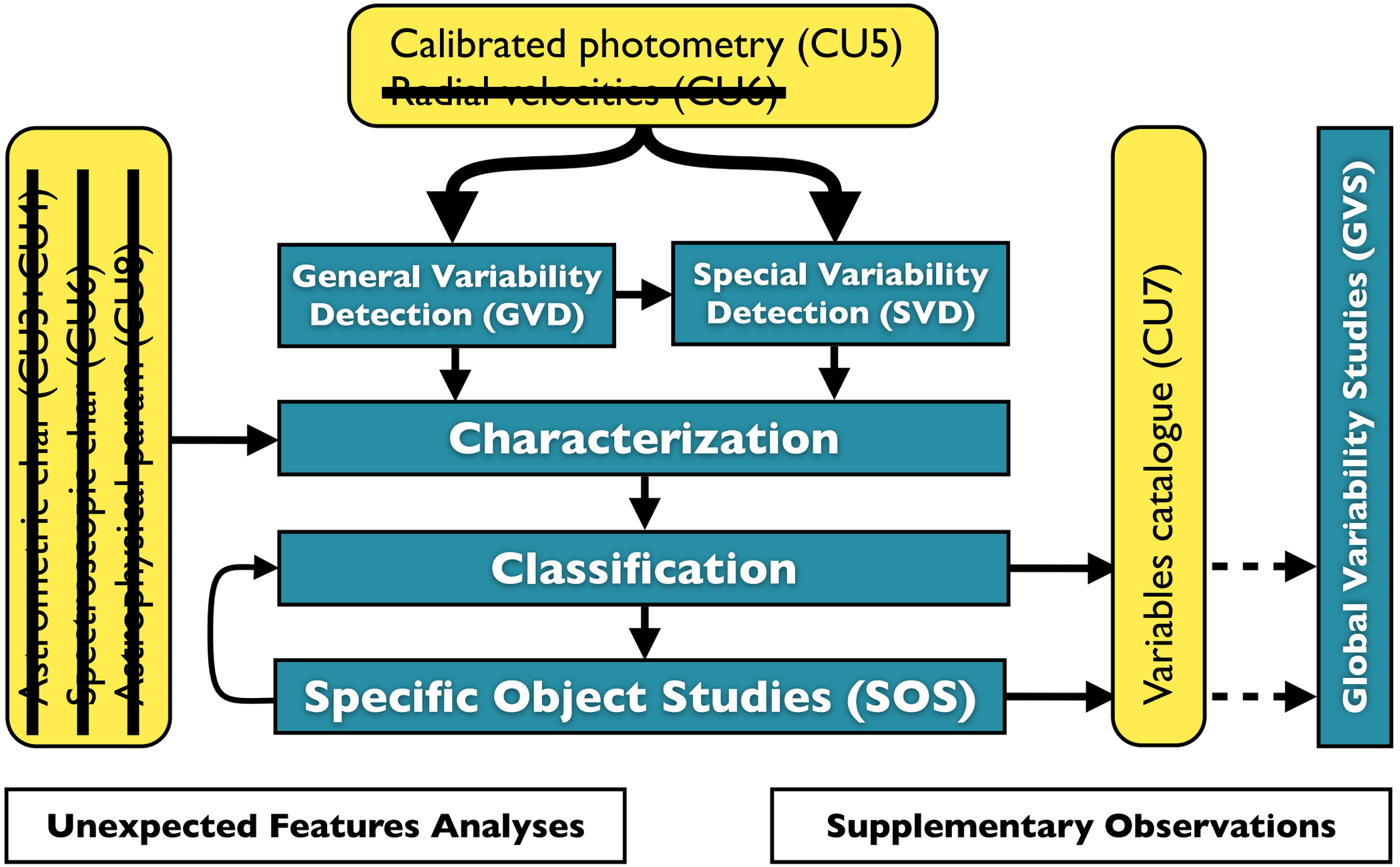
Credits: ESA/Gaia/DPAC/CU5/CU7/INAF-OABo, Gisella Clementini, Dafydd Evans, Laurent Eyer, Krzysztof Nienartowicz, Lorenzo Rimoldini and the Geneva CU7/DPCG and CU7/INAF-OACN teams.







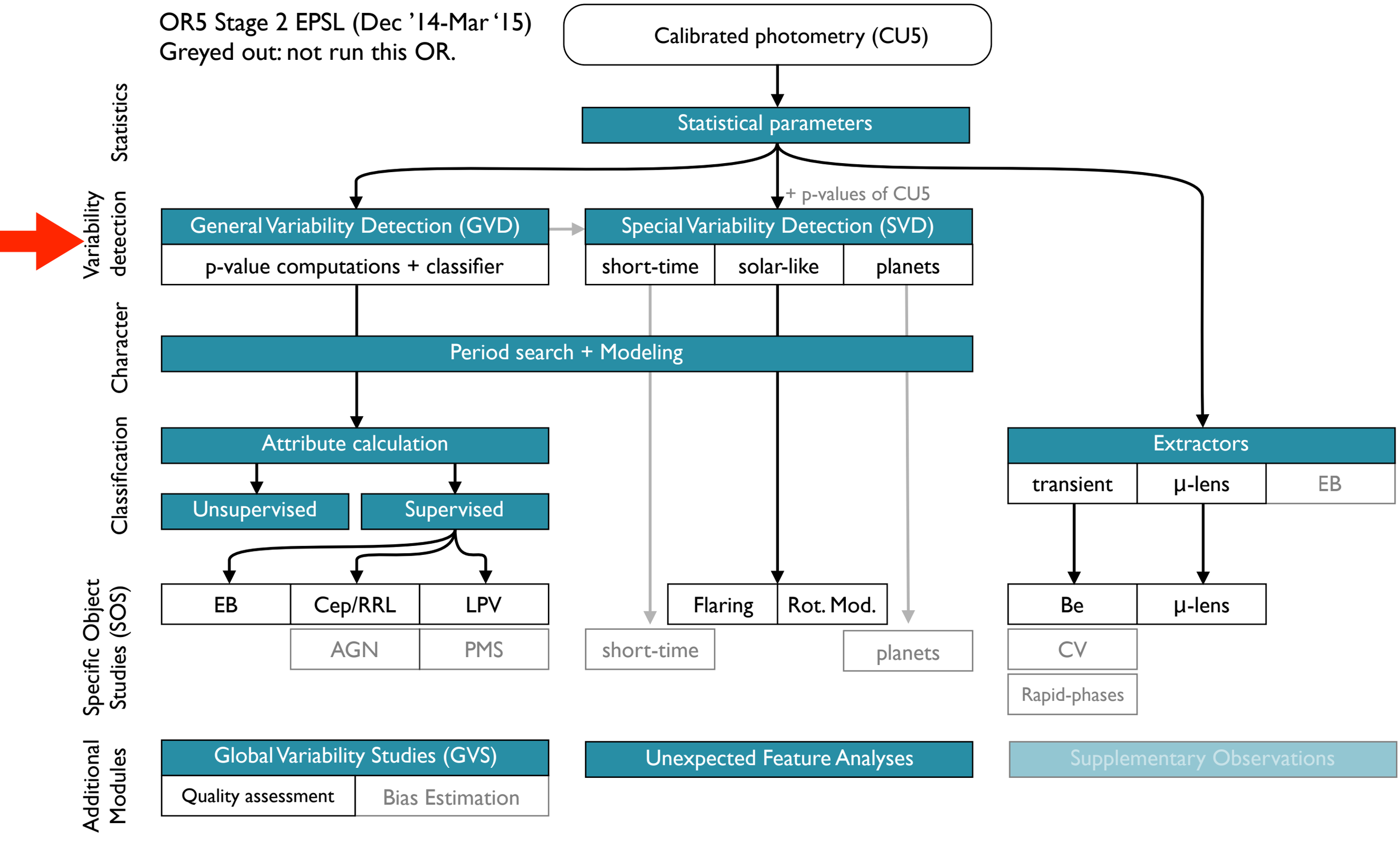
CU7 / DPCG Variability Analysis



Courtesy of B.Holl

CU7 processing chain

OR5 Stage 2 EPSL (Dec '14-Mar '15)
Greyed out: not run this OR.



Courtesy of B.Holl

General Variability Detection

Isabelle Lecoeur, Lorenzo Rimoldini, Diego Ordonez, Laurent Eyer

Two fundamental quantities to estimate:

- **Completeness**
- **Contamination**

Detection was done with a classifier (Random Forest)

attributes were computed

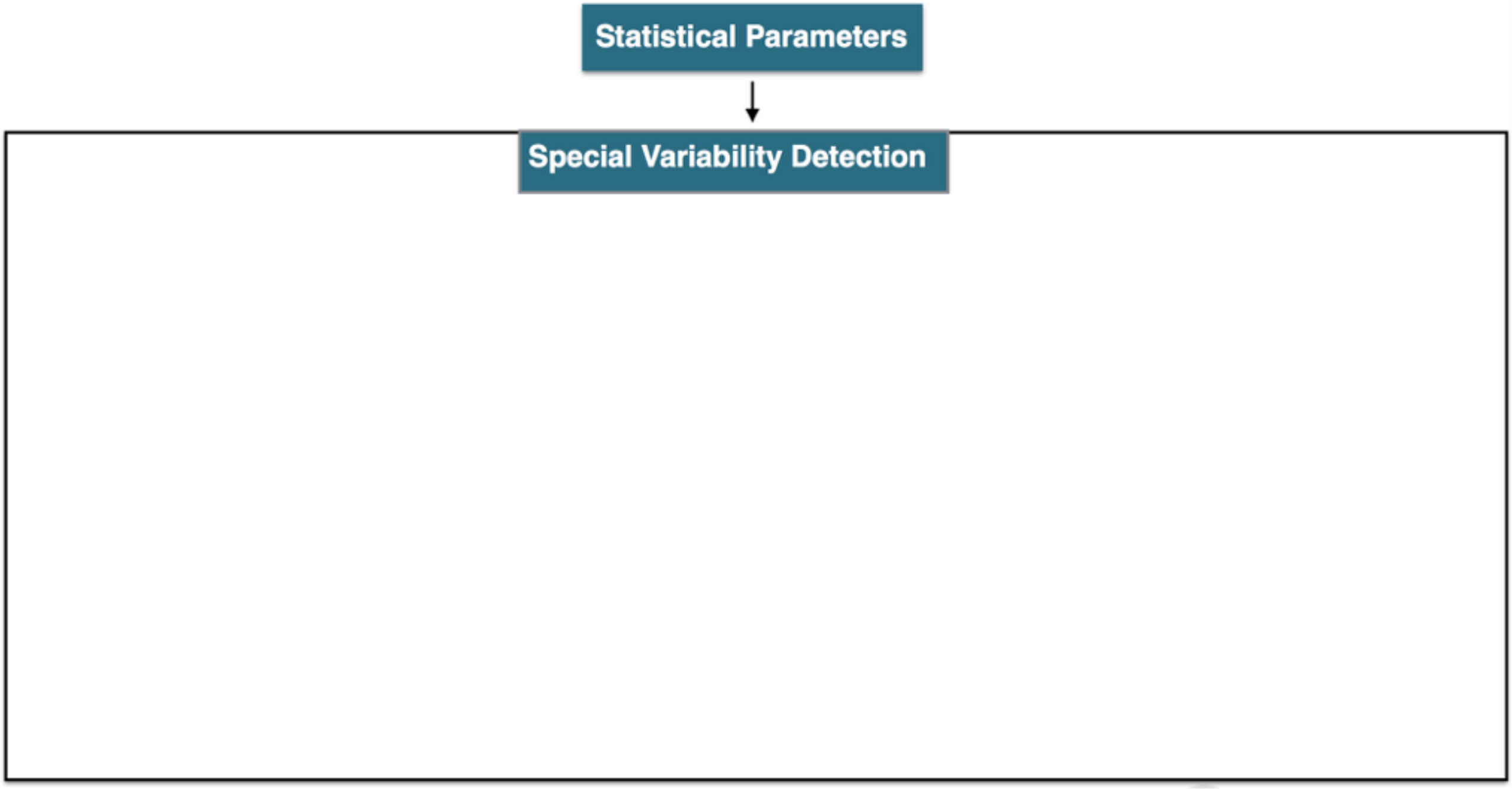
a training set was defined (based on OGLE)

Classifier result: The confusion matrix

		VARIABLE	CONSTANT
376	VARIABLE	80	20
546	CONSTANT	5	95
	Contamination	8	13

Special Variability Detection:

Elisa Distefano, Shay Zucker, Brandon Tingley, Laurent Eyer, Isabelle Lecoeur, Maroussia Roelens, Leanne Guy, Alessandro Lanzafame



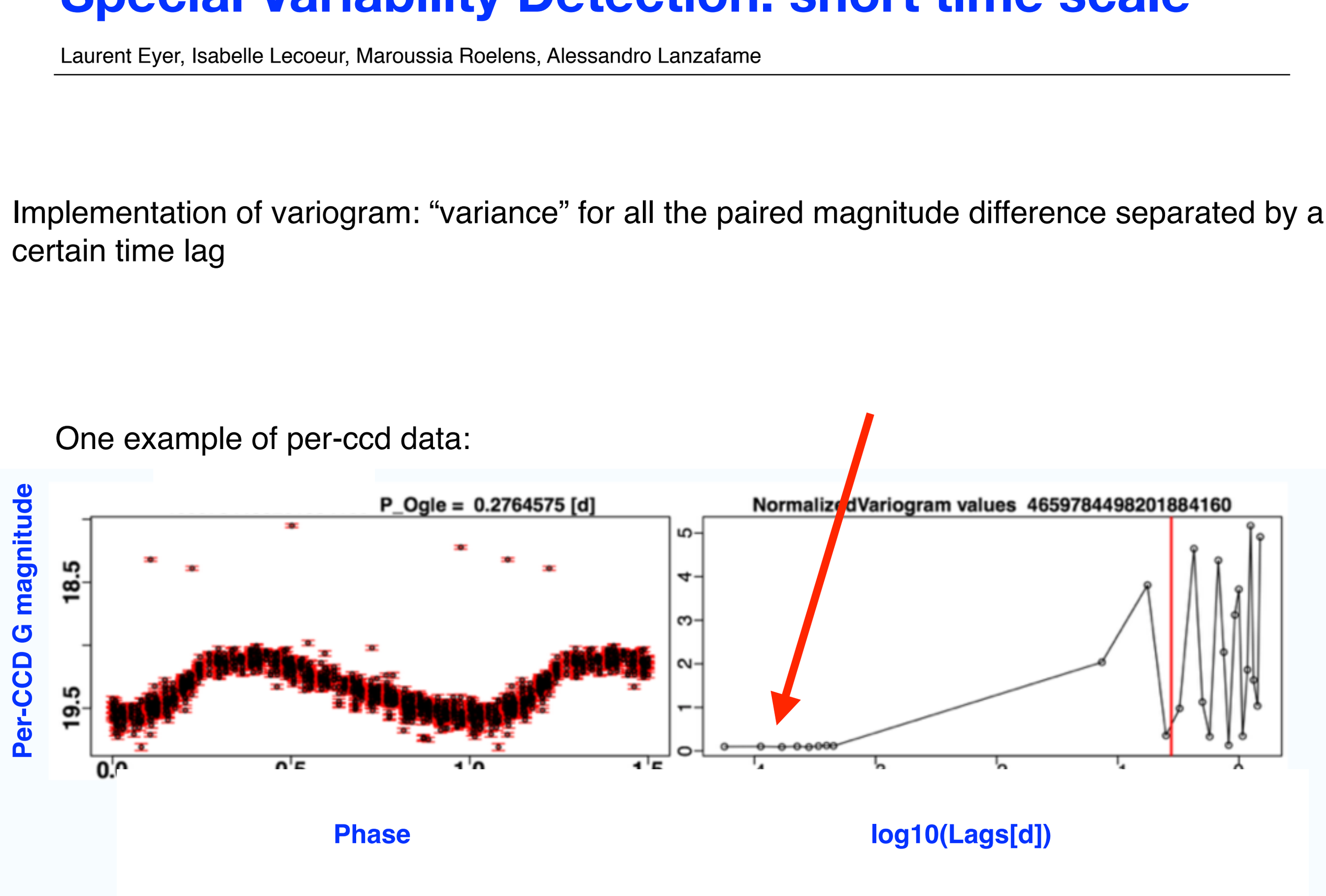
Courtesy of L.Guy

Special Variability Detection: short time scale

Laurent Eyer, Isabelle Lecoeur, Maroussia Roelens, Alessandro Lanzafame

Implementation of variogram: “variance” for all the paired magnitude difference separated by a certain time lag

One example of per-ccd data:



Courtesy of M.Roelens/I.Lecoeur

Special Variability Detection: exo-planet transits

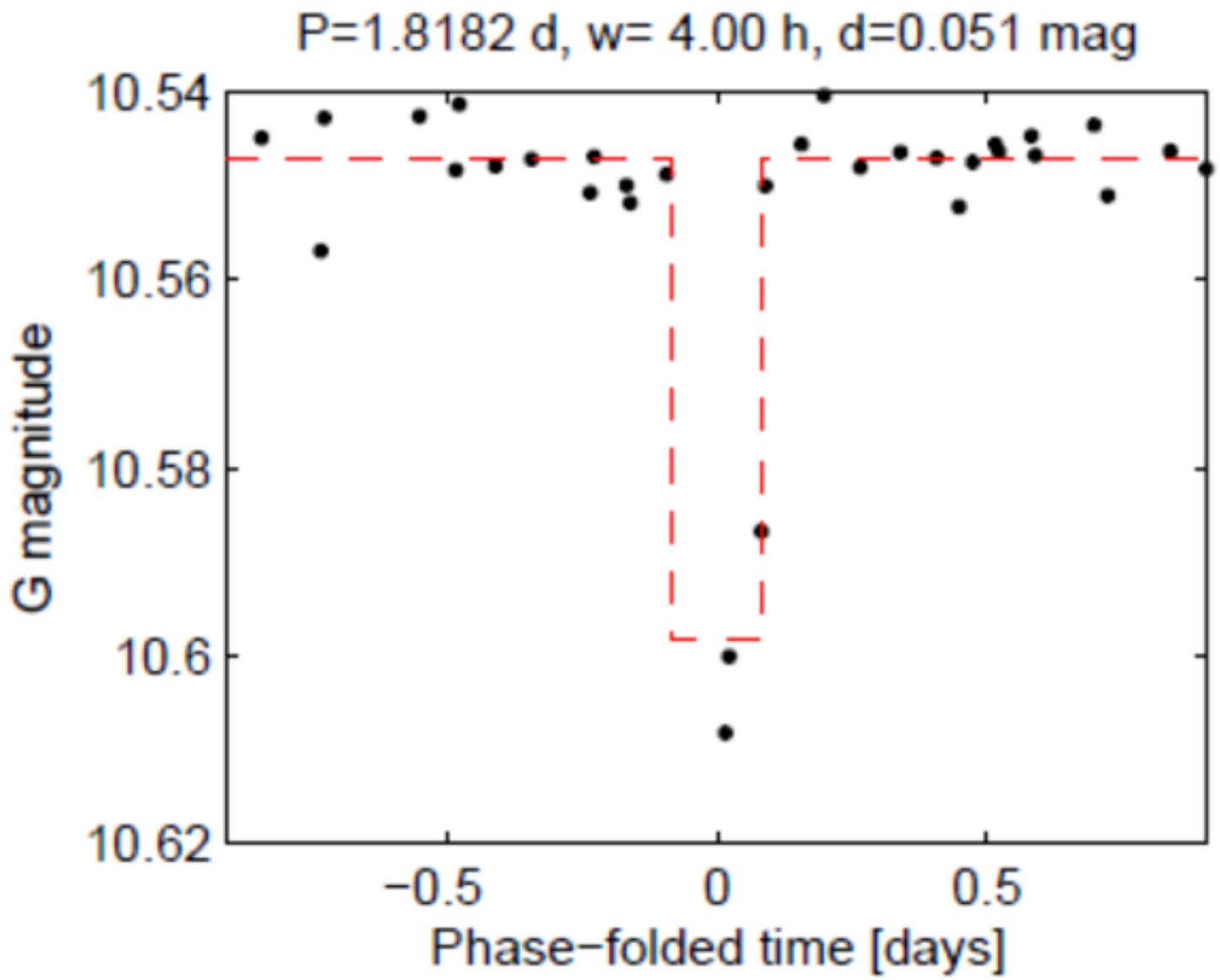
Shay Zucker, Brandon Tingley, Leanne Guy, Alessandro Lanzafame

Two algorithms:

—Box-Least Square

—Outlier Probability, Tingley (A&A 2011)

The Box Least Square algorithm gives about 200 candidates

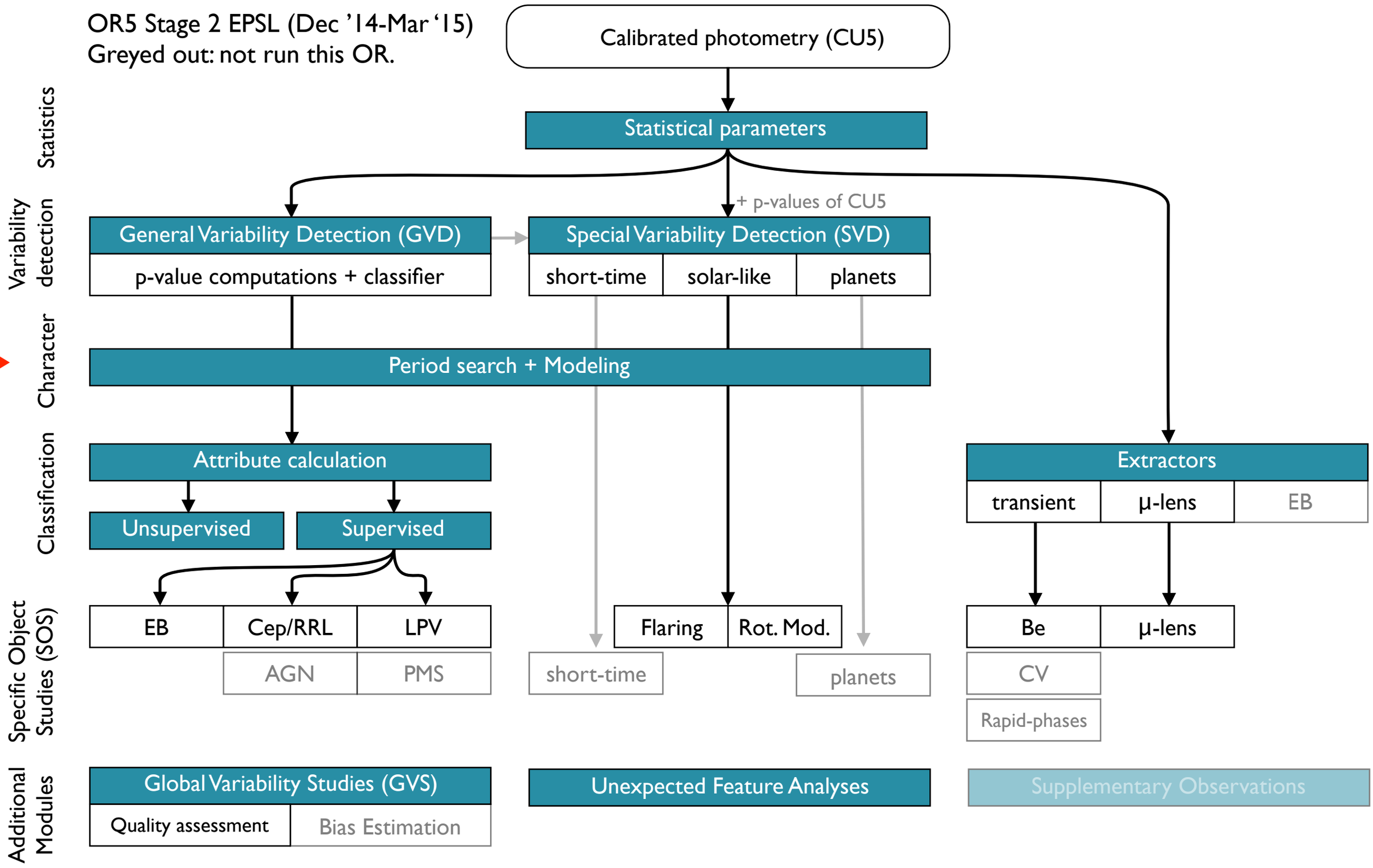


Courtesy of S. Zucker

We do not claim any detection yet !

Conclusion Box-Least Square is functioning well

CU7 processing chain



Courtesy of B.Holl

Characterisation

Jan Cuypers, Leanne Guy, Lorenzo Rimoldini, Joris De Ridder

Time series per object:

Time(i), **G**-, **BP**-, **RP**- mag(i) [or **RV**/radial velocity(i)] i=1,…, number of measurements

Goal: To define attributes

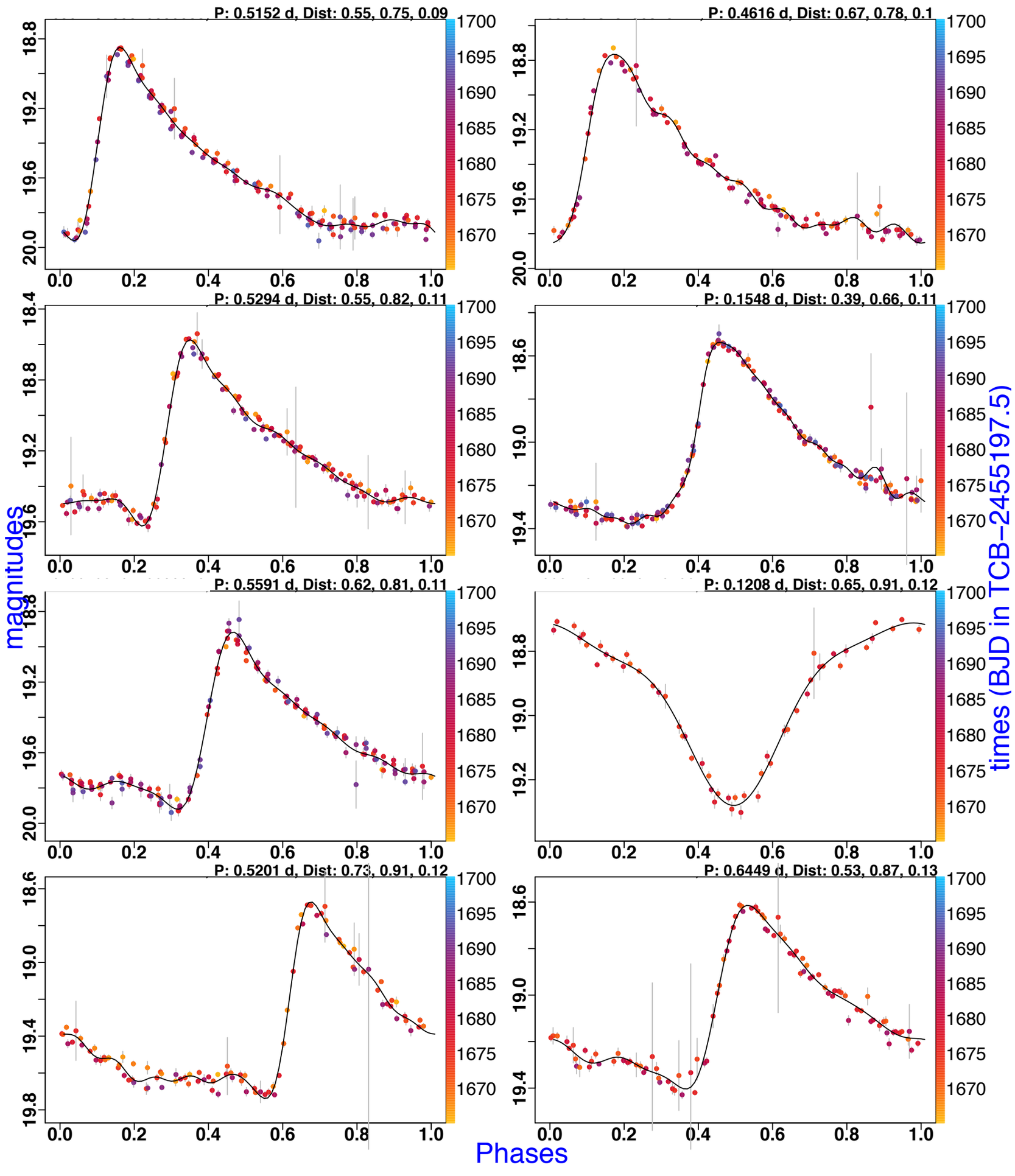
- statistical parameters

- Modelling

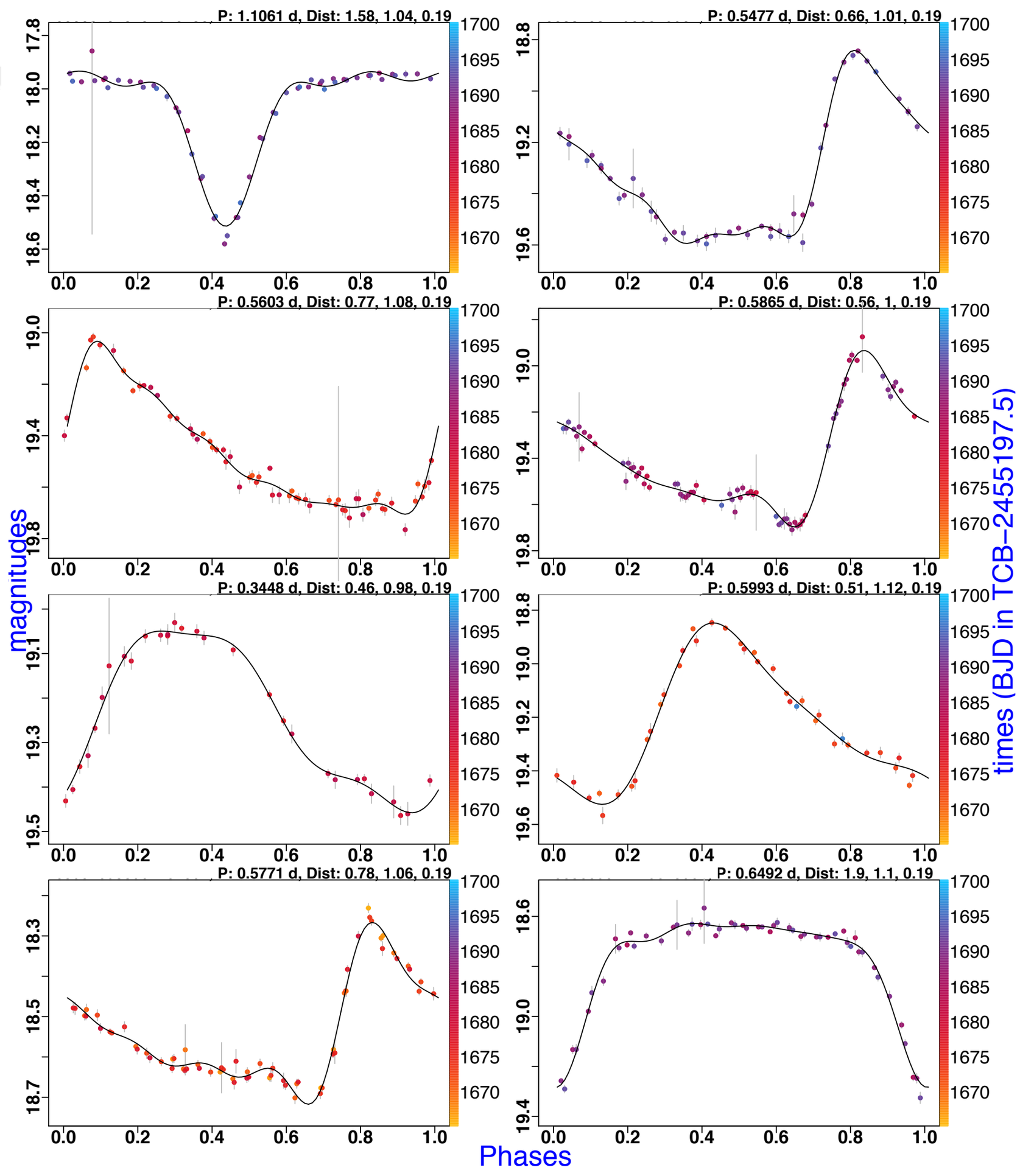
  —Period search

  —Fourier Series and polynomial fit

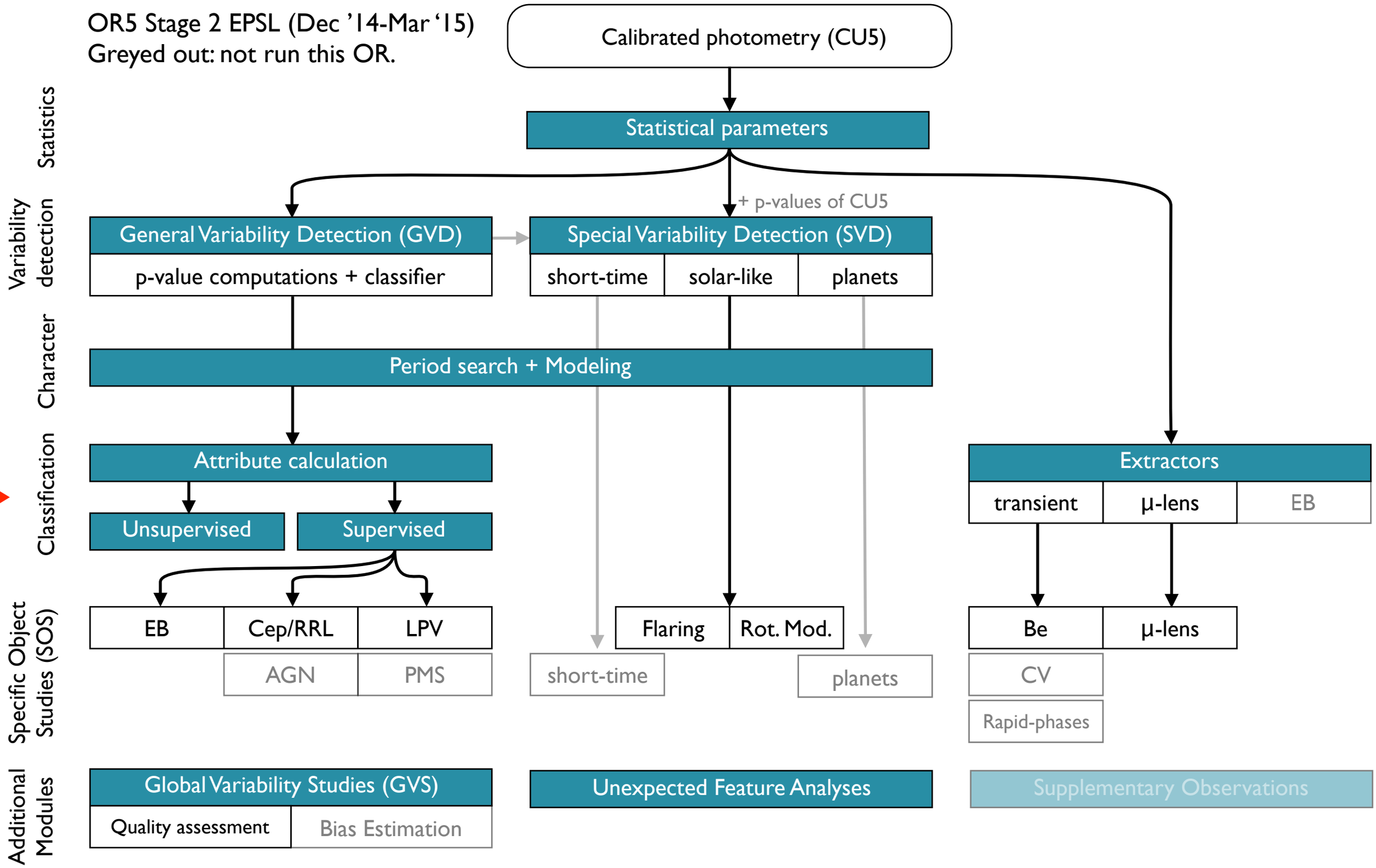
Few examples of modelling



Few examples of modelling



CU7 processing chain



Courtesy of B.Holl

Classification

Joris de Ridder, Berry Holl, Lorenzo Rimoldini, Luis Sarro, Sara Regibo, Mauro Lopez, Jonas Debosscher, Maria Sueveges

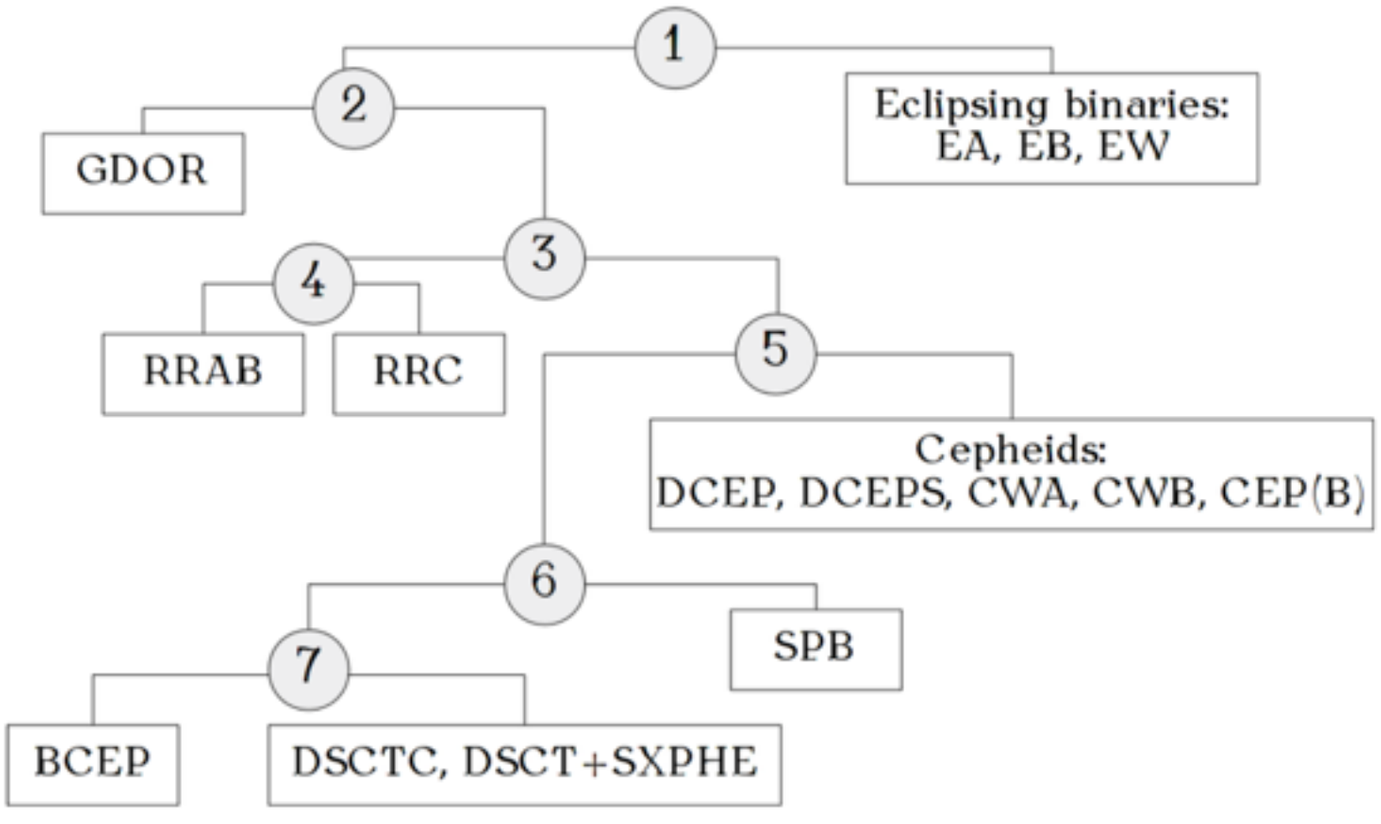
Supervised classification (several methods):

Multistage tree:
Bayesian networds

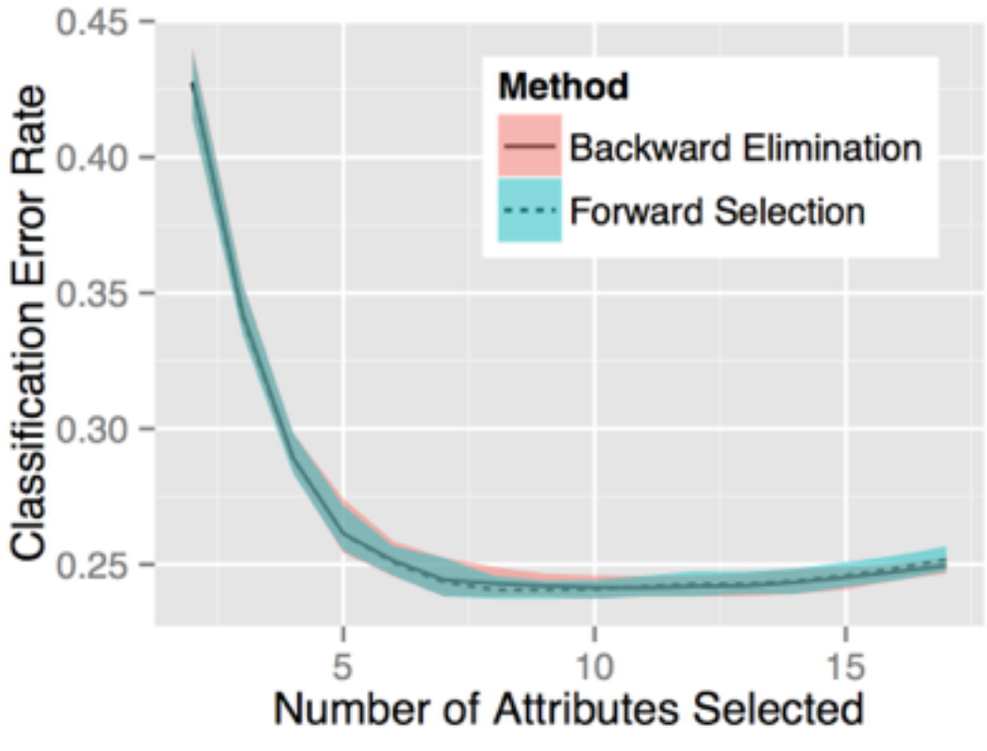
Multistage tree:
Gaussian mixture

Random Forest

Tree for Gaussian Mixture:



Furnish training set
built from Crossmatched data

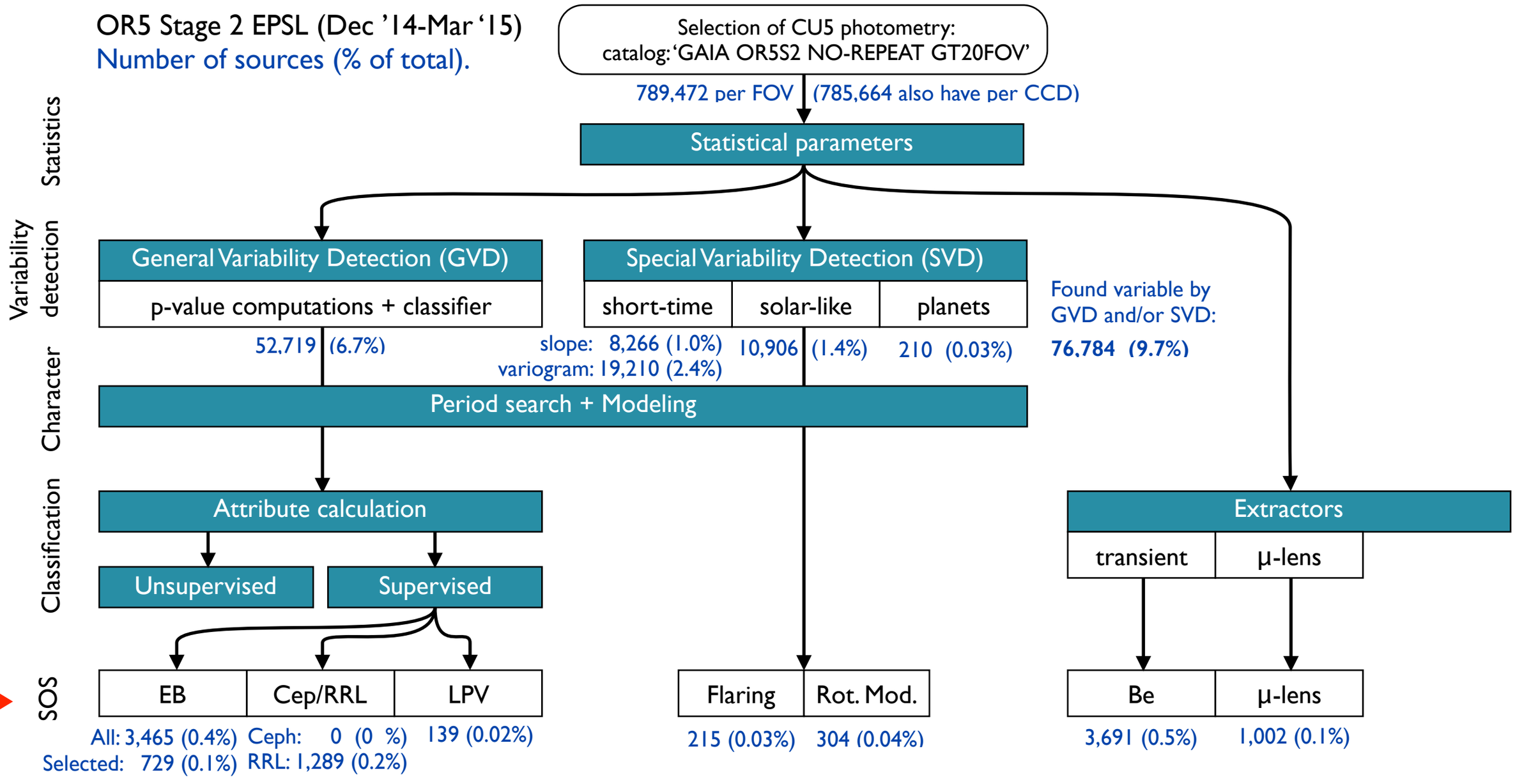


Classification

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	_	_	_

CU7 processing chain



Courtesy of B.Holl

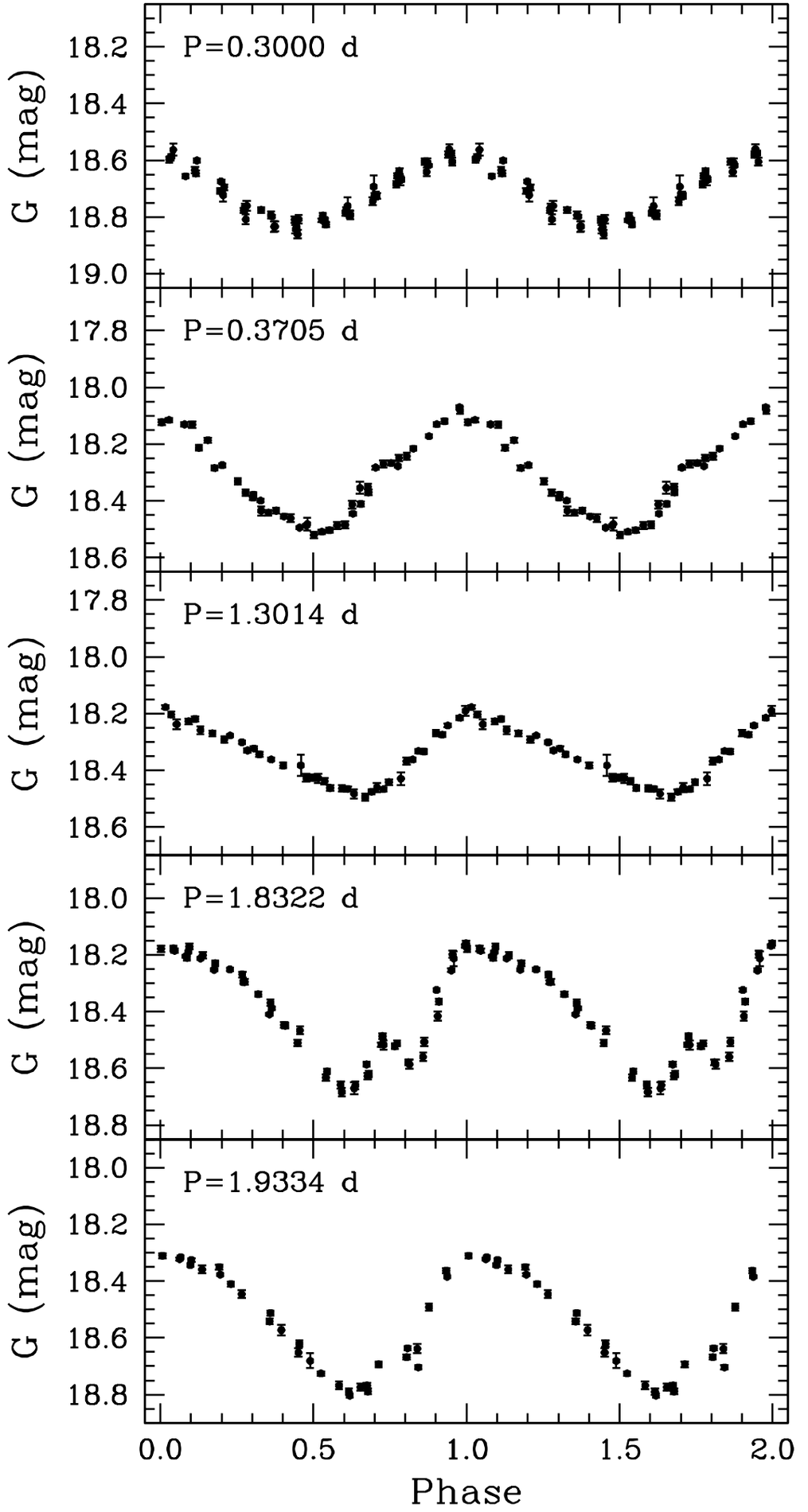
Specific Object Studies
RR Lyrae and Cepheid stars

Gisella Clementini, Silvio Leccia, Vincenzo Ripepi, Nami Mowlavi, Isabelle Lecoeur

Image of the Week (May 28):

Classical overtone Cepheid
3 candidate anomalous Cepheids
Type 2 Cepheid

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.

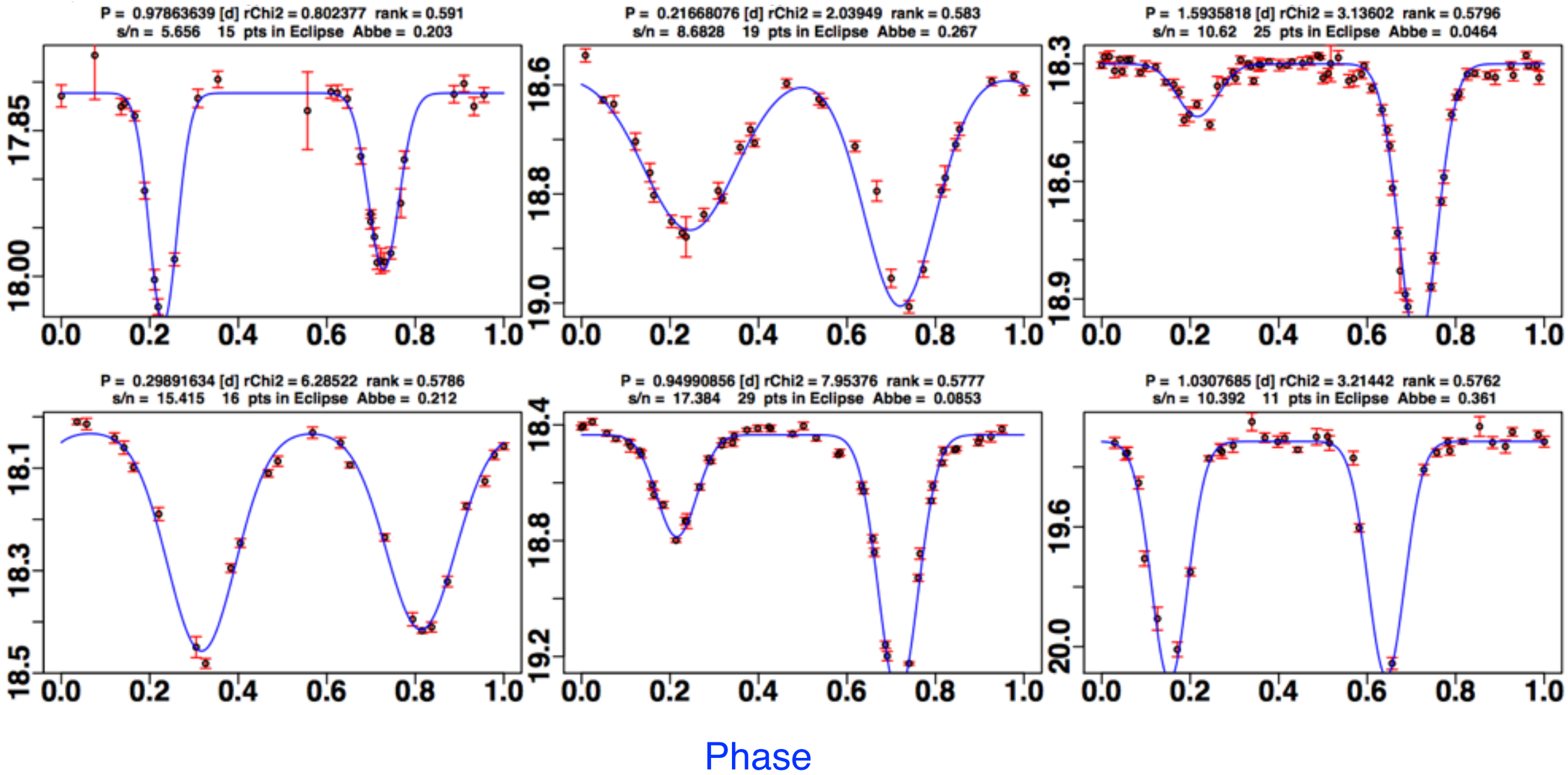


Specific Object Studies: Eclipsing binaries

Nami Mowlavi, Berry Holl, Isabelle Lecoeur, Fabio Barblan, Lorenzo Rimoldini

Eclipsing binaries go to CU4 for a full modelling
In CU7, some simple modelling are made
Solutions allow a ranking

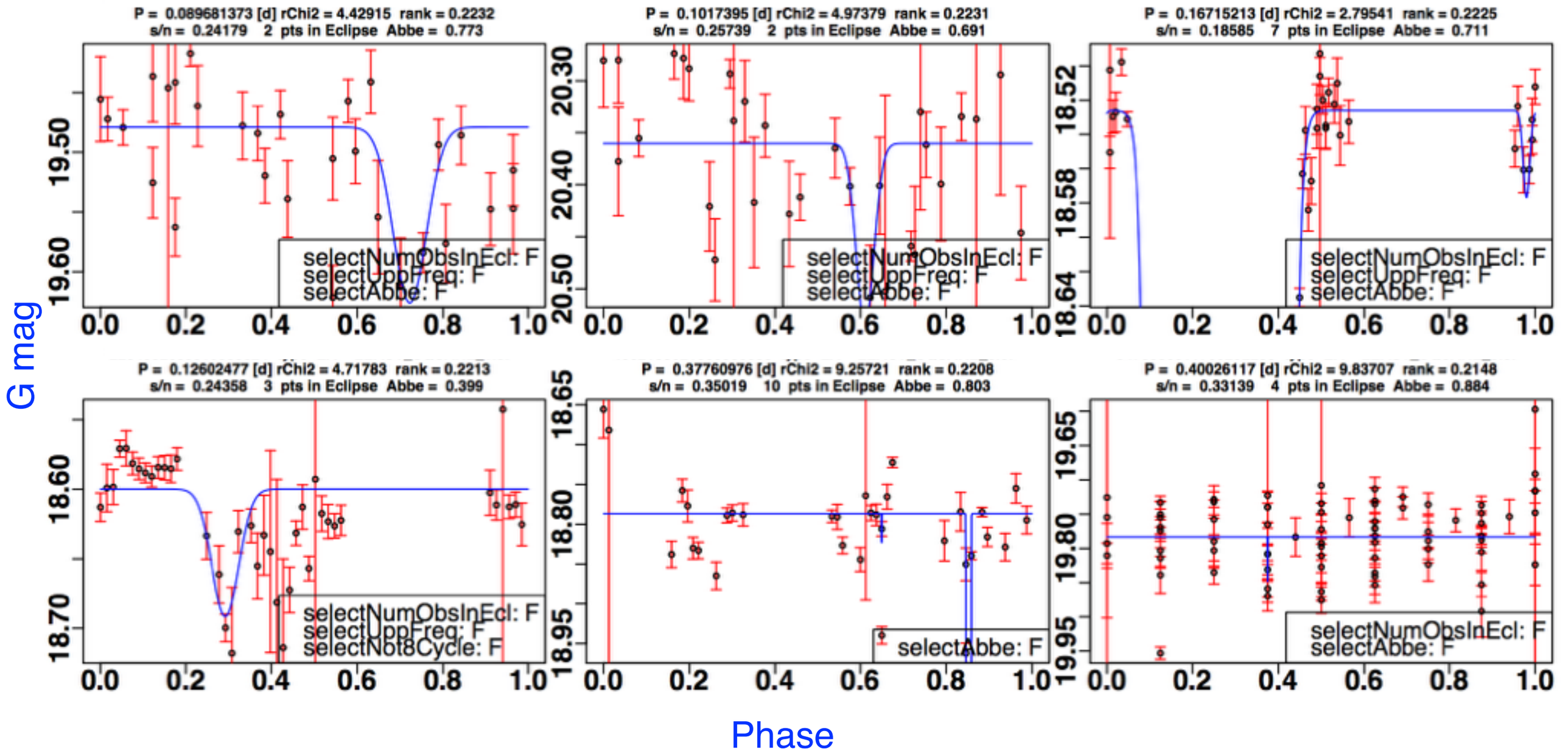
Highest rank



Specific Object Studies: Eclipsing binaries

Nami Mowlavi, Berry Holl, Isabelle Lecoeur, Fabio Barblan, Lorenzo Rimoldini

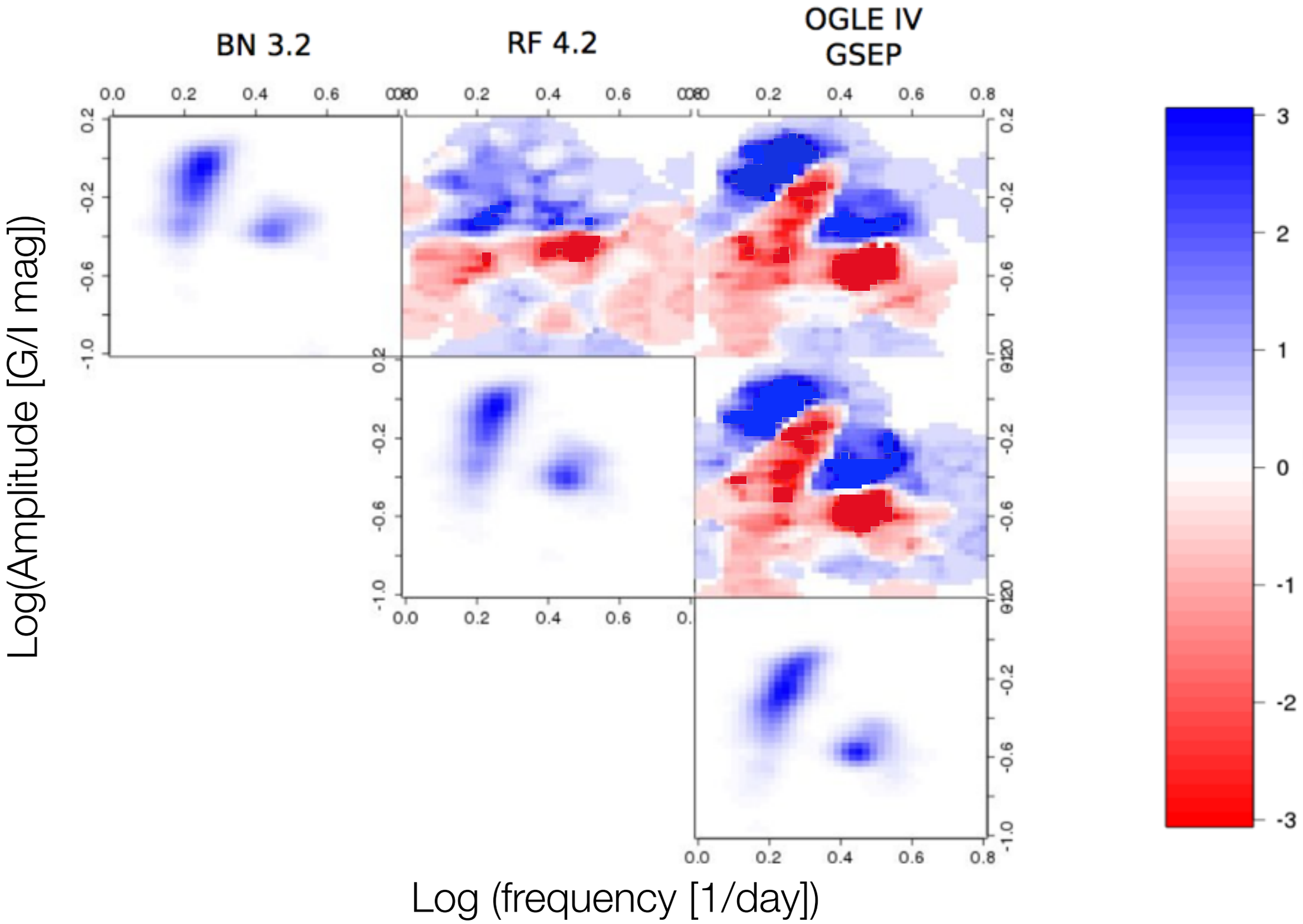
Lowest rank



Global Variability studies

Luis Sarrro, Manuel, André Moitinho, Leanne Guy

One example: Comparison of distribution functions of RR Lyrae stars



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

Release scenario

spacecraft operations start

nominal mission end

extended mission end?

2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022

Release 1:

- **α and δ, mean G-magnitude**
- **Commissioning data**
- **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}
- 5-par astrometry
- Object classifications and Astrophysical Parameters
- **Orbital solution of binaries**
- **Mean RVS spectra**

Release 4:

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

Final release:

- everything !

Courtesy of B.Holl

Conclusions

- We can remark that for a first reduction, the quality of G band is remarkable

- The photometry has improved already (for the current Operation Rehearsal)

- The Variability Processing and Analysis seems to be on the right track !

- We (CU5/CU7 joint effort) may release some EPSL data variability analysis (Data Release 1)