Transient Astronomy with the Gaia Satellite

Simon Hodgkin, Lukasz Wyrzykowski, Sergey Koposov, Nadejda Blagorodnova, Guy Rixon, Floor van Leeuwen, Vasily Belokurov, Nic Walton

Institute of Astronomy, Cambridge

Acknowledgements

- Giuseppe Altavilla
- Vasily Belokurov
- Josh Bloom
- Ross Burgon
- Nadejda Blagorodnova
- Gisella Clementini
- Andrew Drake
- Gerry Gilmore
- Sergey Koposov
- Floor van Leeuwen
- Ashish Mahabal
- Francois Mignard
- Timo Prusti
- Guy Rixon
- Iain Steele
- Rachel Street
- Yiannis Tsapras
- Massimo Turatto
- Nic Walton
- Roy Williams
- Lukasz Wyrzykowski

Outline

- Motivation: Transient Astronomy
- The Gaia mission
- Detecting Science Alerts with Gaia
- Supernovae and Microlensing
- Details: Classification from Photometry
- Details: Classification from Spectroscopy
- Verification and Follow-up

(The Simpsons, Bart's Comet, S06E14)

Why do transient Astronomy ?

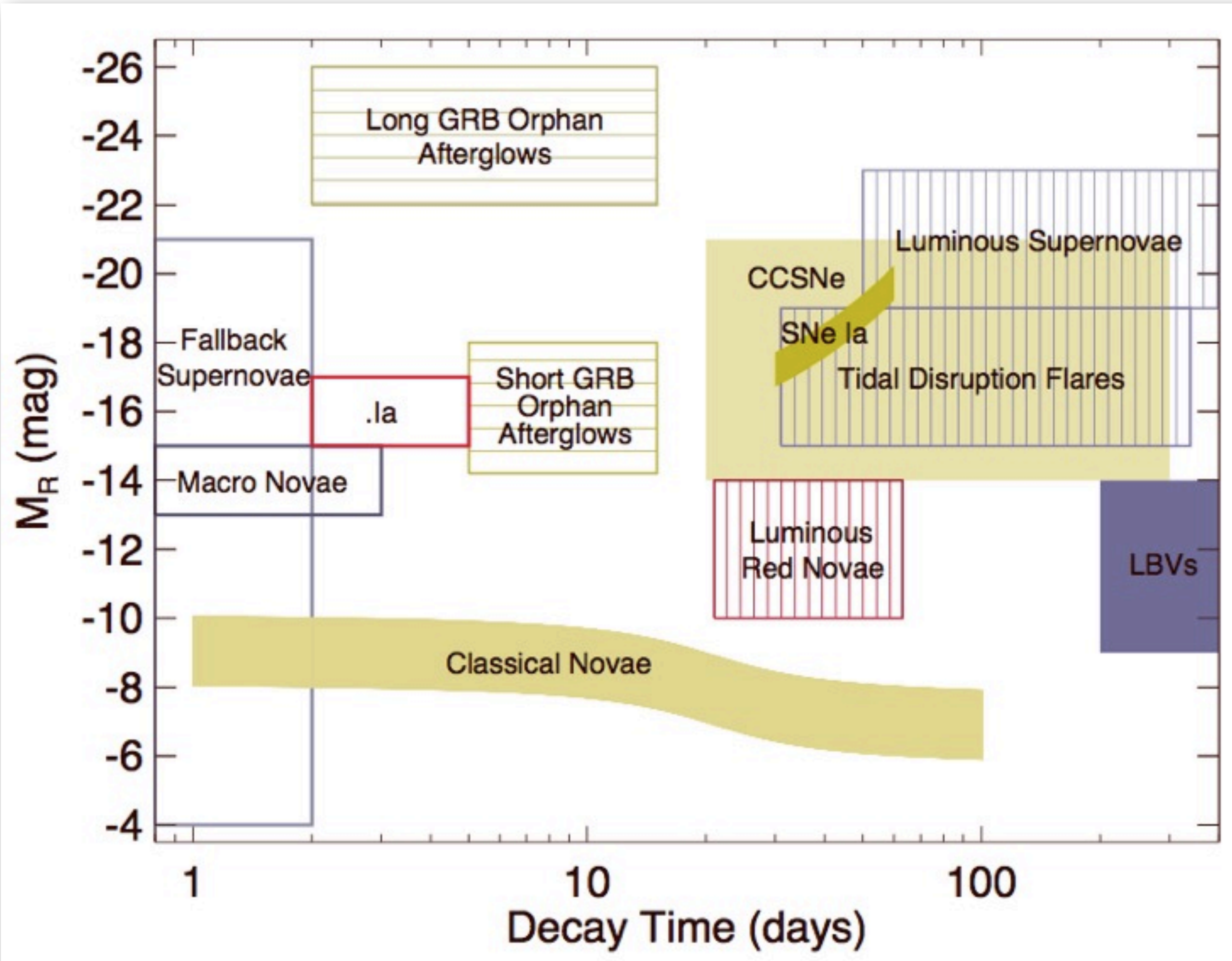
A (childlike) desire to see what's out there.

Variability is everywhere, and a useful diagnostic in Astrophysics.

Studying variable/transient behaviour leads to improved/new physics

Let's go burn down the observatory so this will never happen again.

the transient zoo: from fast to slow



well studied

few cases

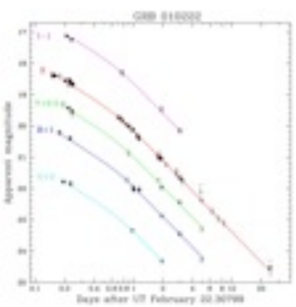
not seen yet

made up

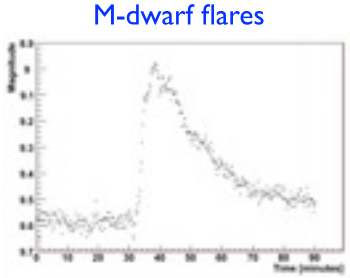
EXPLORING THE OPTICAL TRANSIENT SKY WITH THE PALOMAR TRANSIENT FACTORY, Rau et al. 2009

Potential Triggers

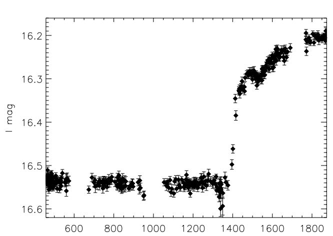
GRBs optical counterparts

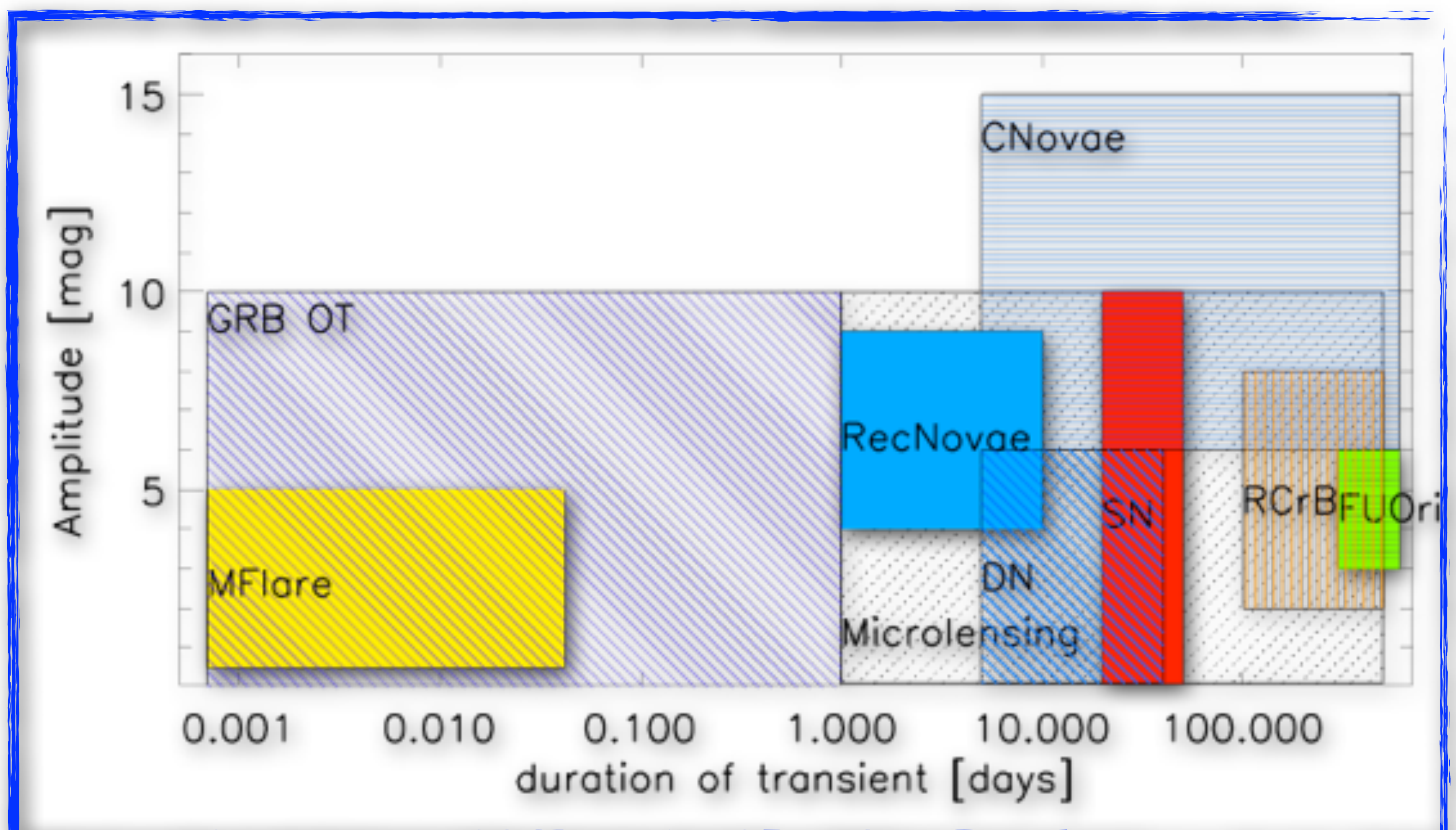


M-dwarf flares

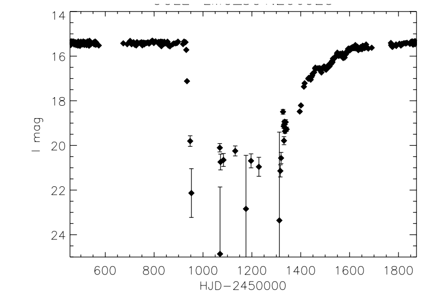


Be stars

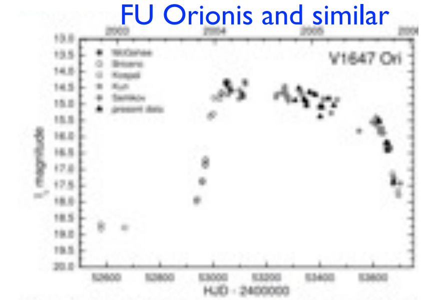




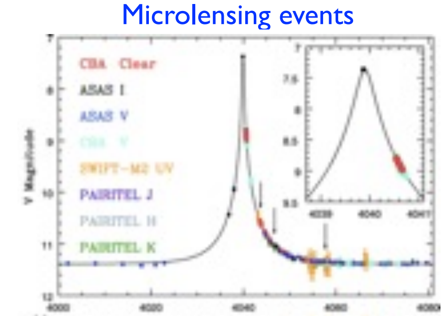
R Coronae Borealis

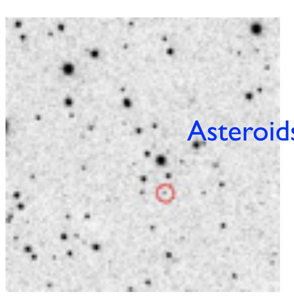


FU Orionis and similar



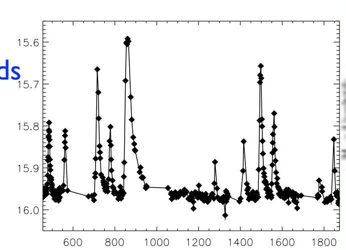
Microlensing events



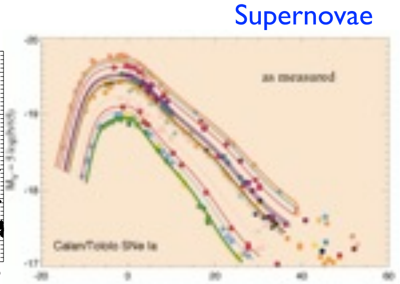


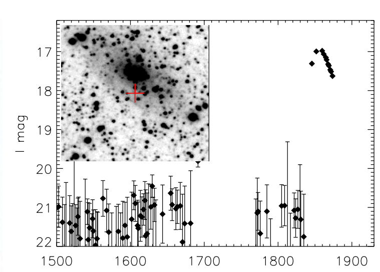
Asteroids

Dwarf novae

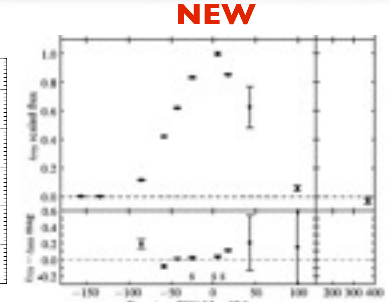


Supernovae

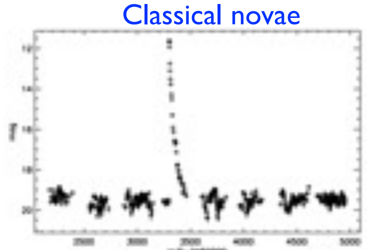




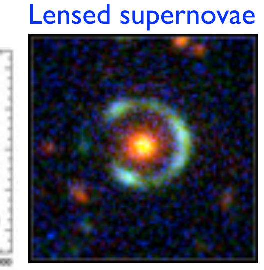
NEW



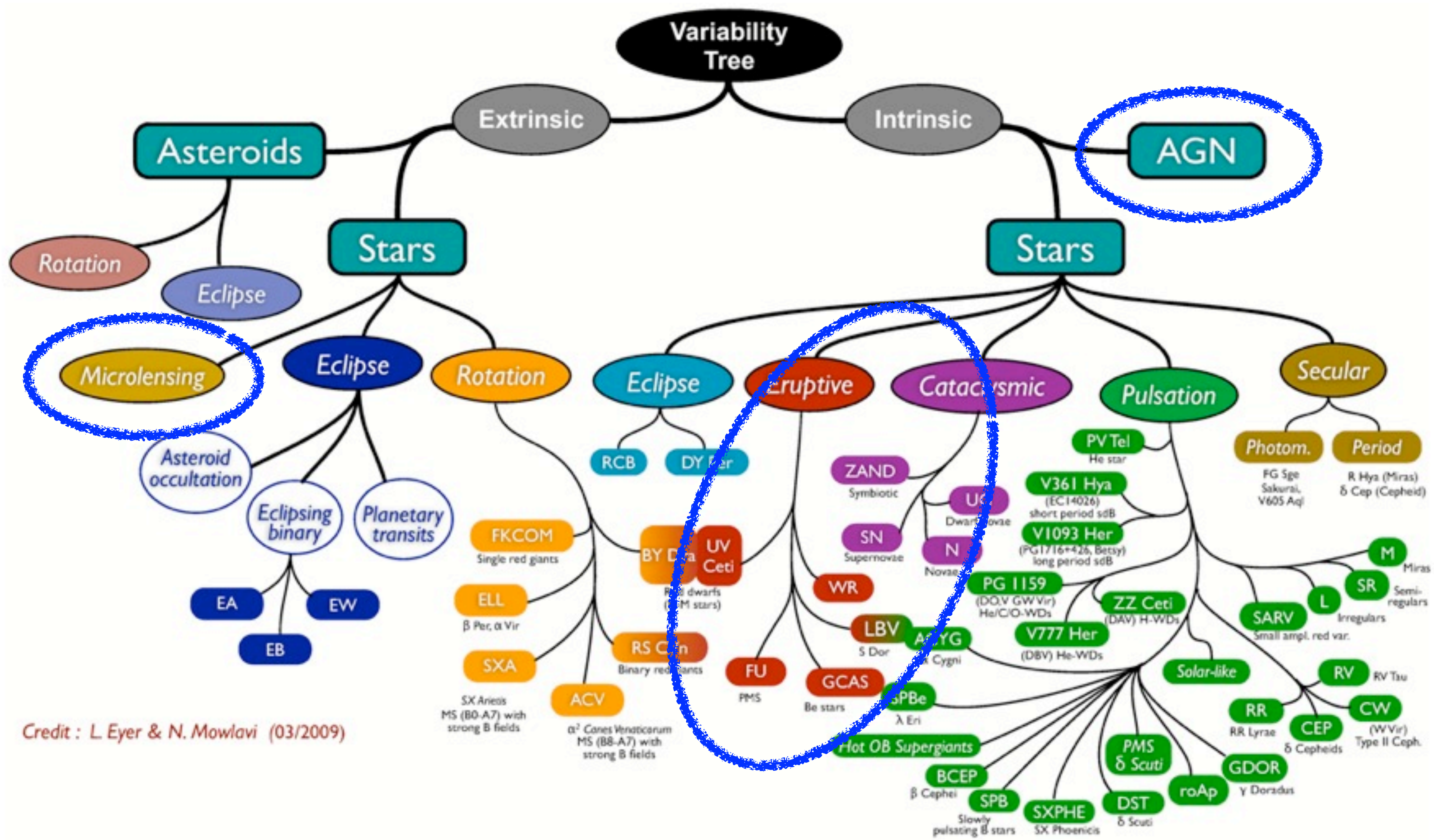
Classical novae



Lensed supernovae



Variable stars vs Science Alerts



Science Alerts: science data that would have little or no value without quick ground-based follow-up

Ongoing Transient/ Variability Experiments

Outline

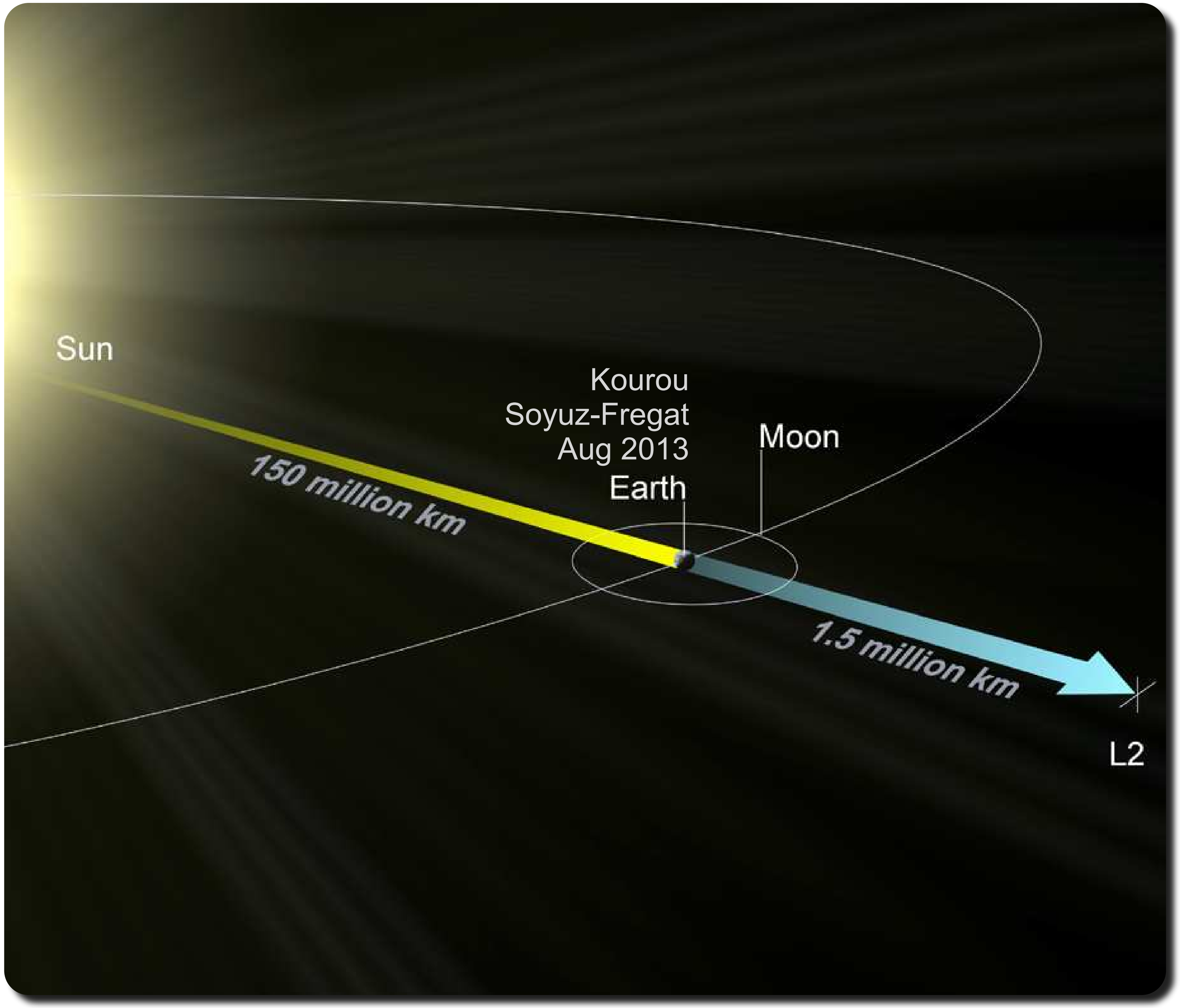
- Transient astronomy
- The Gaia mission
- Detecting transients with Gaia
- Supernovae and microlensing
- Details: Classification from photometry
- Details: Classification from spectroscopy
- Verification and follow-up

what is Gaia?

the fades, bbc3, episode 3

The ESA Gaia mission

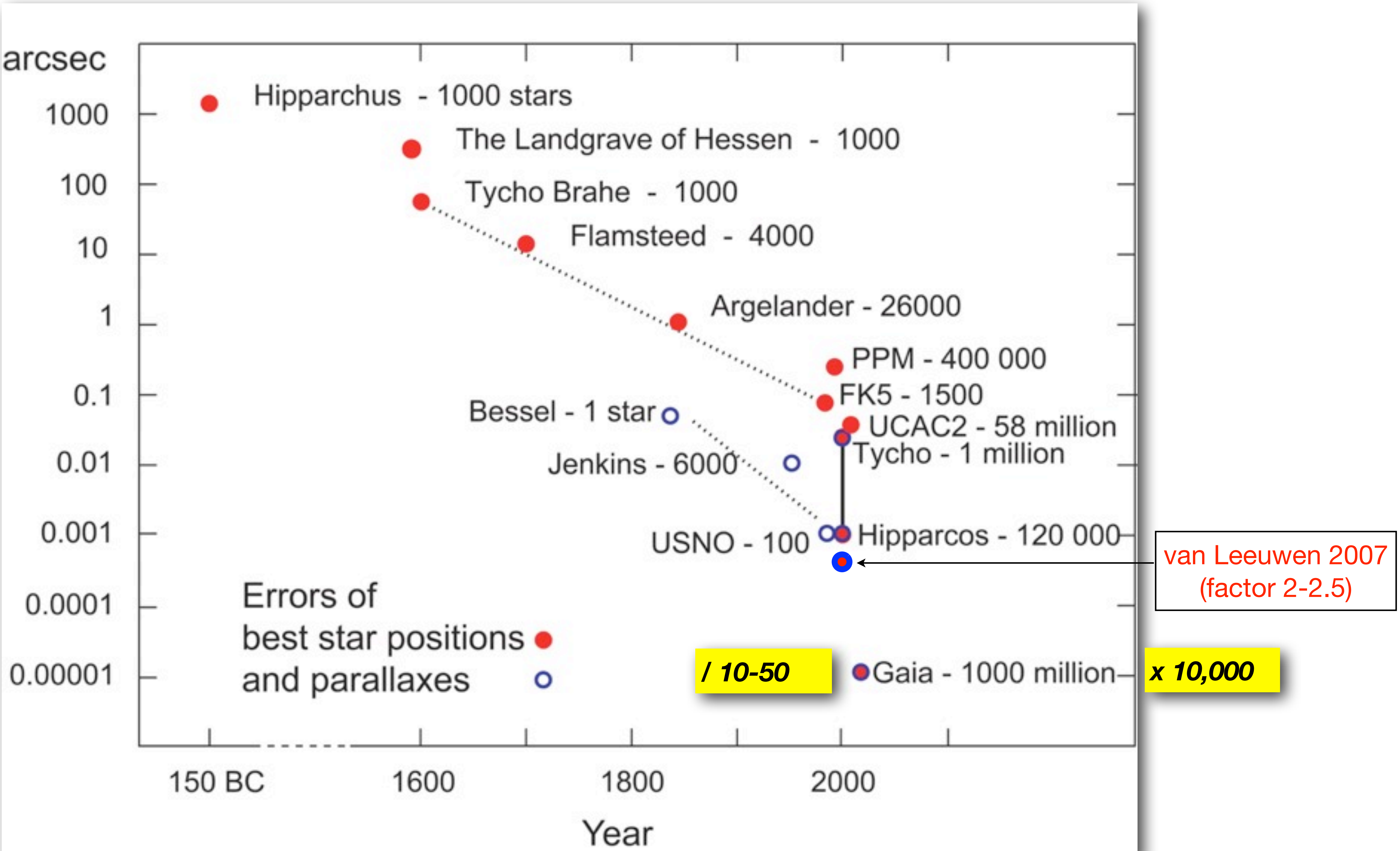
- 1 billion objects V=5-20 (~1% of the Galaxy's stars)
- Astrometry, Photometry, Spectrophotometry, Spectroscopy (radial velocities)
- 5 (+1) years (70x all sky): final results 2020-2021



http://www.rssd.esa.int/Gaia

10 μas is very, very, very, very, very, small

astrometry

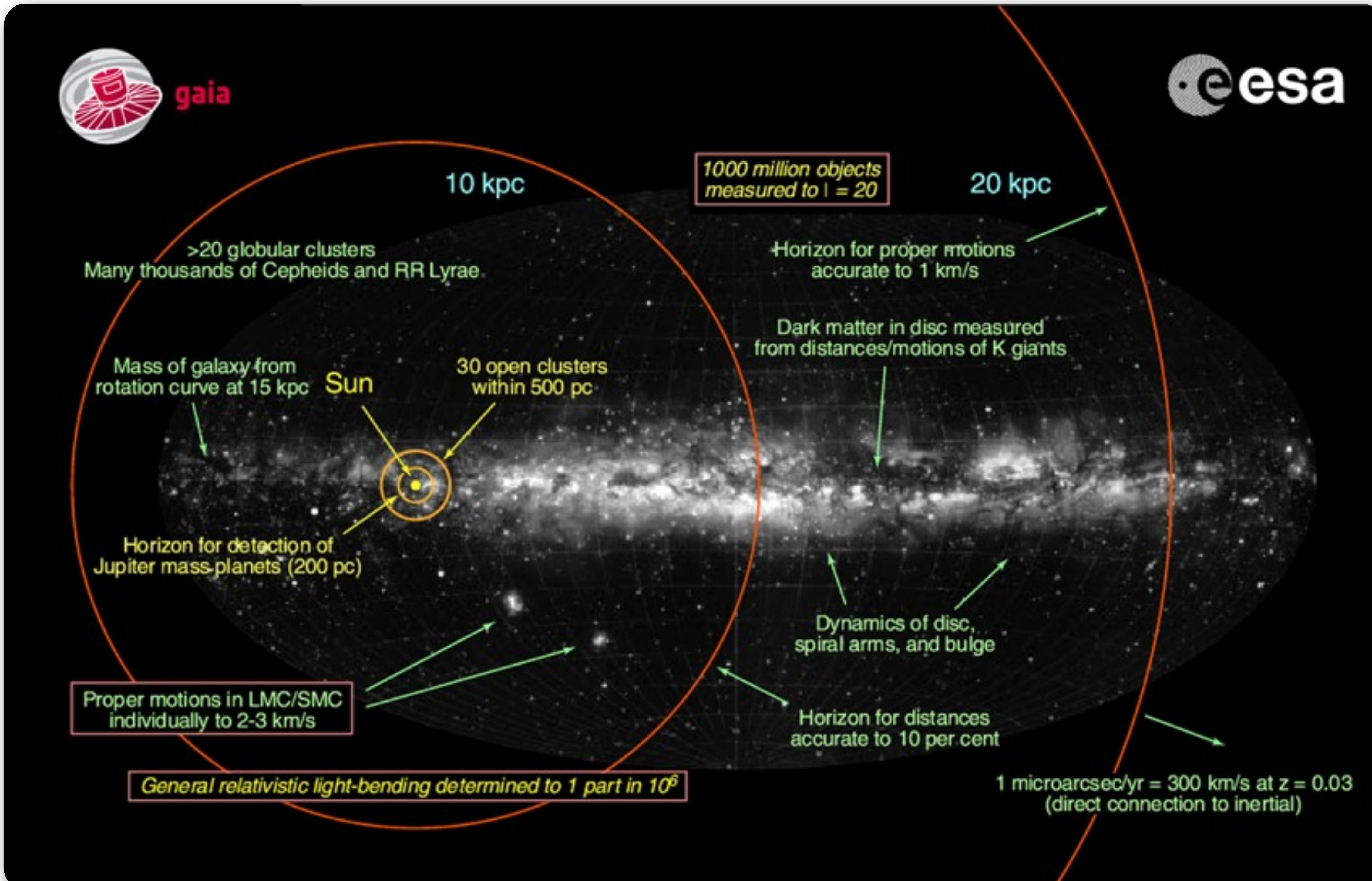


arcsec
1000
100
10
1
0.1
0.01
0.001
0.0001
0.00001
Hipparchus - 1000 stars
The Landgrave of Hessen - 1000
Tycho Brahe - 1000
Flamsteed - 4000
Argelander - 26000
PPM - 400 000
FK5 - 1500
Bessel - 1 star
UCAC2 - 58 million
Tycho - 1 million
Jenkins - 6000
USNO - 100
Hipparcos - 120 000
van Leeuwen 2007
(factor 2-2.5)
Errors of
best star positions
and parallaxes
/ 10-50
Gaia - 1000 million
x 10,000
150 BC
1600
1800
2000
Year

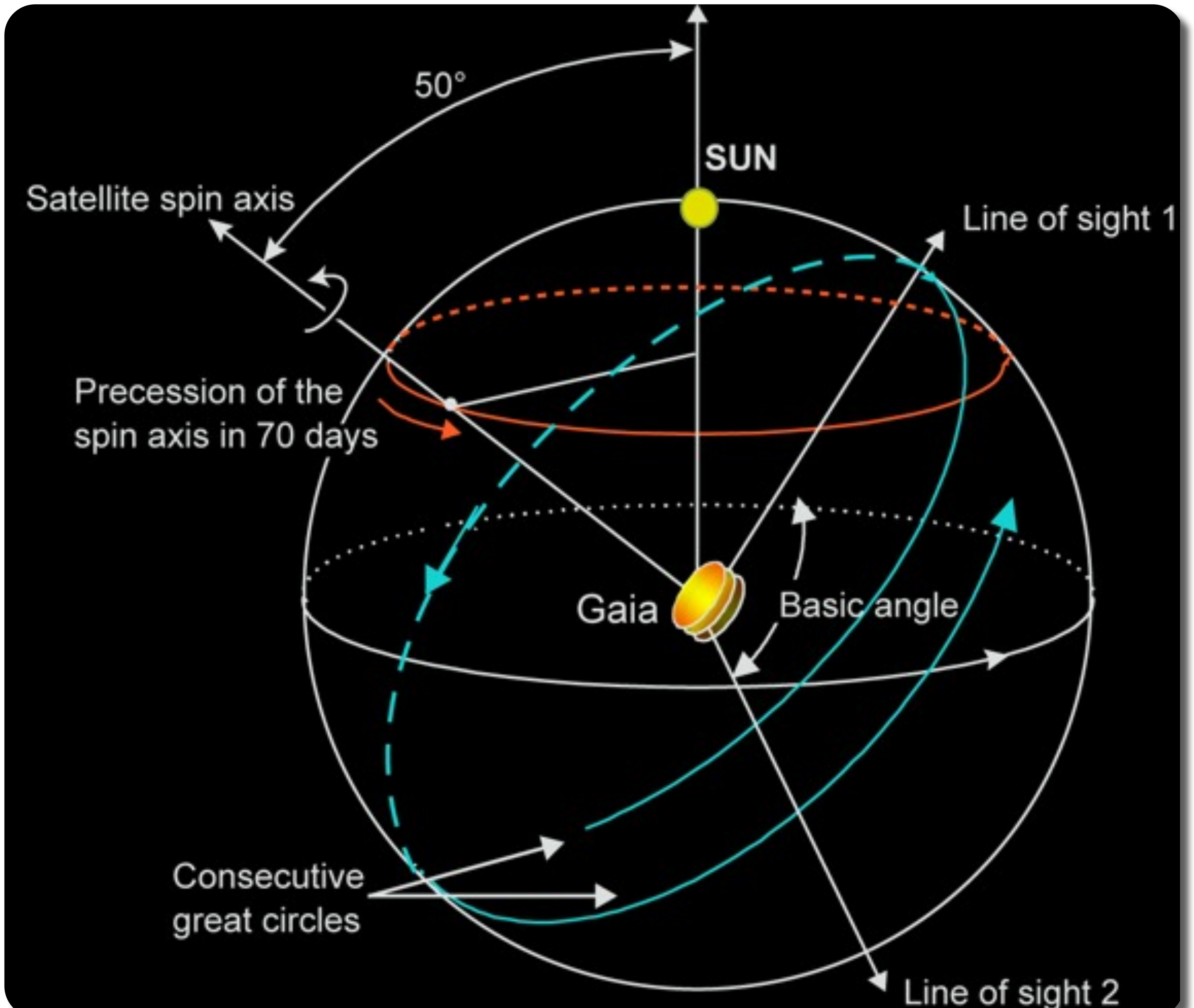
Gaia end of mission parallax errors

	B1V	G2V	M6V
$V-I_C$ [mag]	-0.22	0.75	3.85
Bright stars	5-14 µas (6 mag < V < 12 mag)	5-14 µas (6 mag < V < 12 mag)	5-14 µas (8 mag < V < 14 mag)
V = 15 mag	26 µas	24 µas	9 µas
V = 20 mag	330 µas	290 µas	100 µas

Science Goals

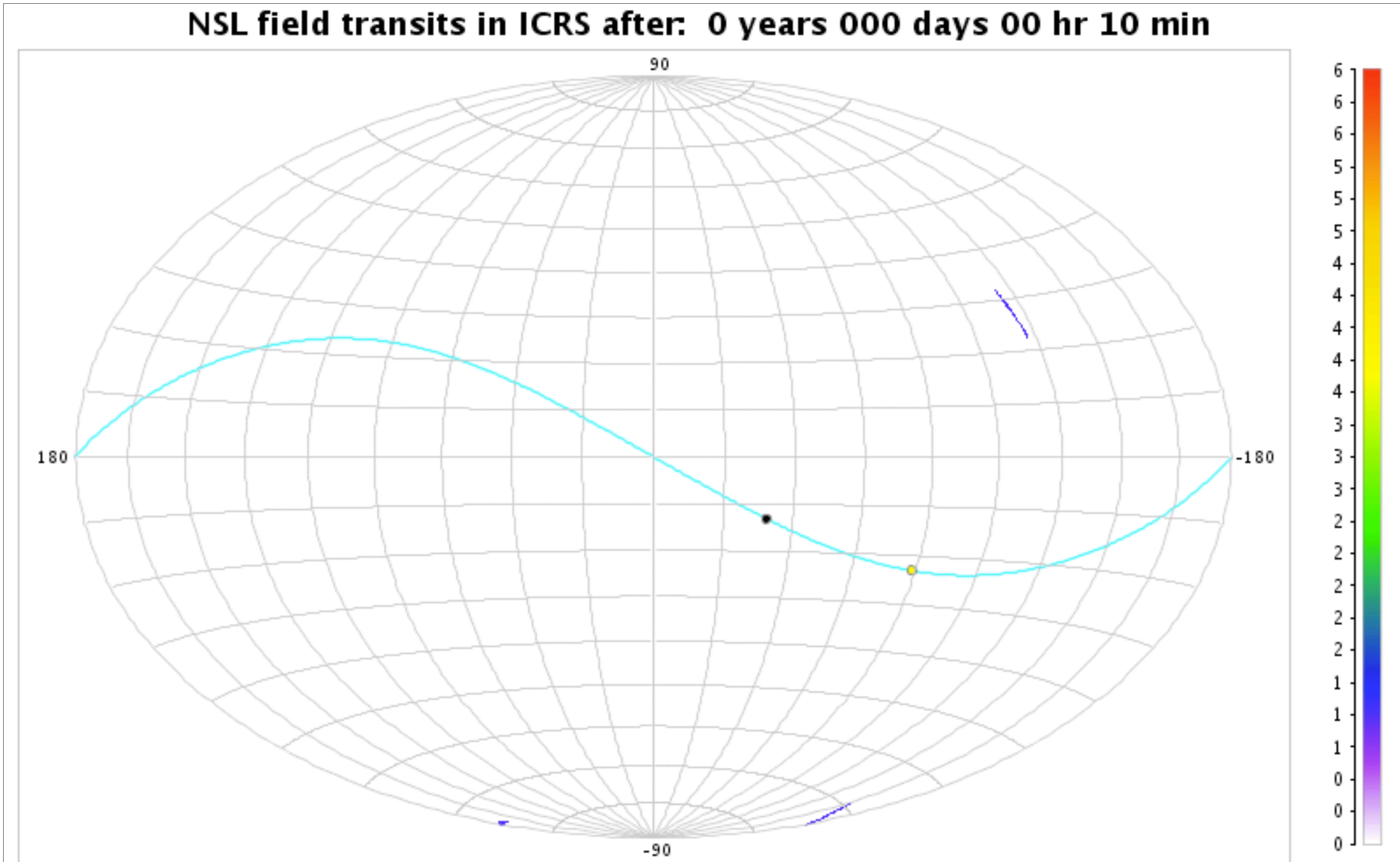


Scanning Law

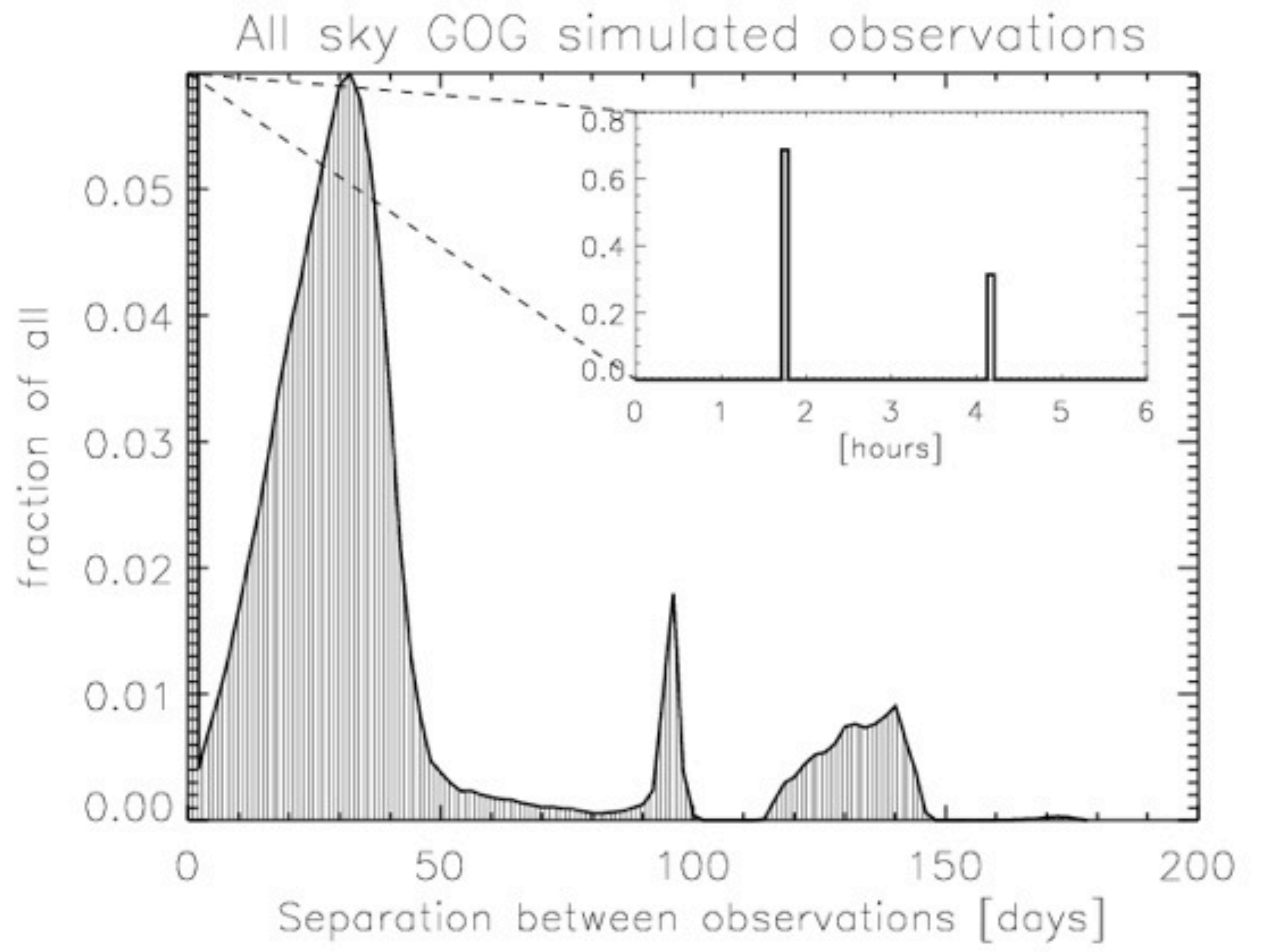


- Two telescopes, one focal plane
- Time between FOVs: 106.5m
- Time between successive scans: 6h
- Field revisited every ~30 days
- Each object measured ~70 times
- Densest coverage ~200 epochs

Scanning Law



Scanning Law

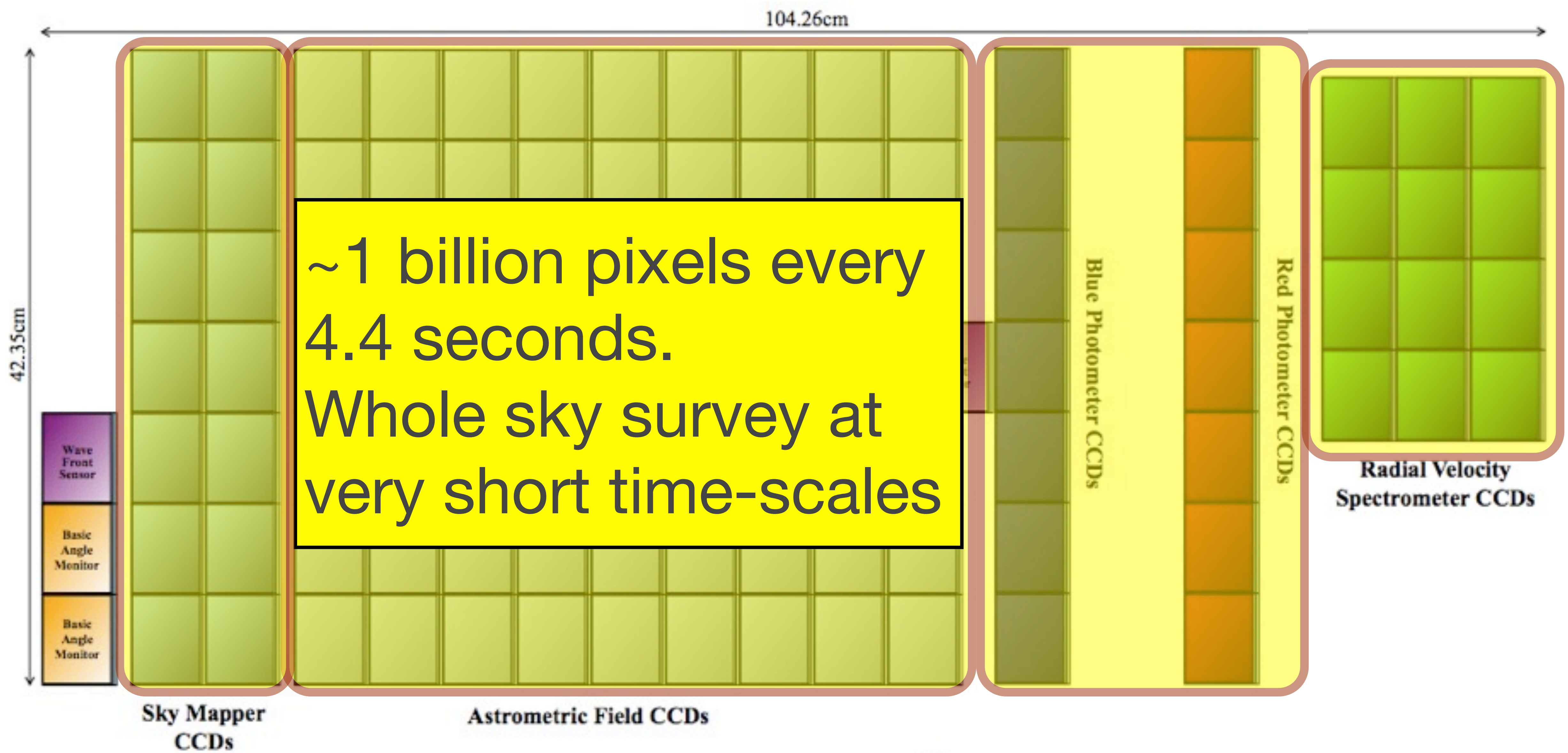


- Two telescopes, one focal plane
- Time between FOVs: 106.5m
- Time between successive scans: 6h
- Field revisited every ~30 days
- Each object measured ~70 times
- Densest coverage ~200 epochs

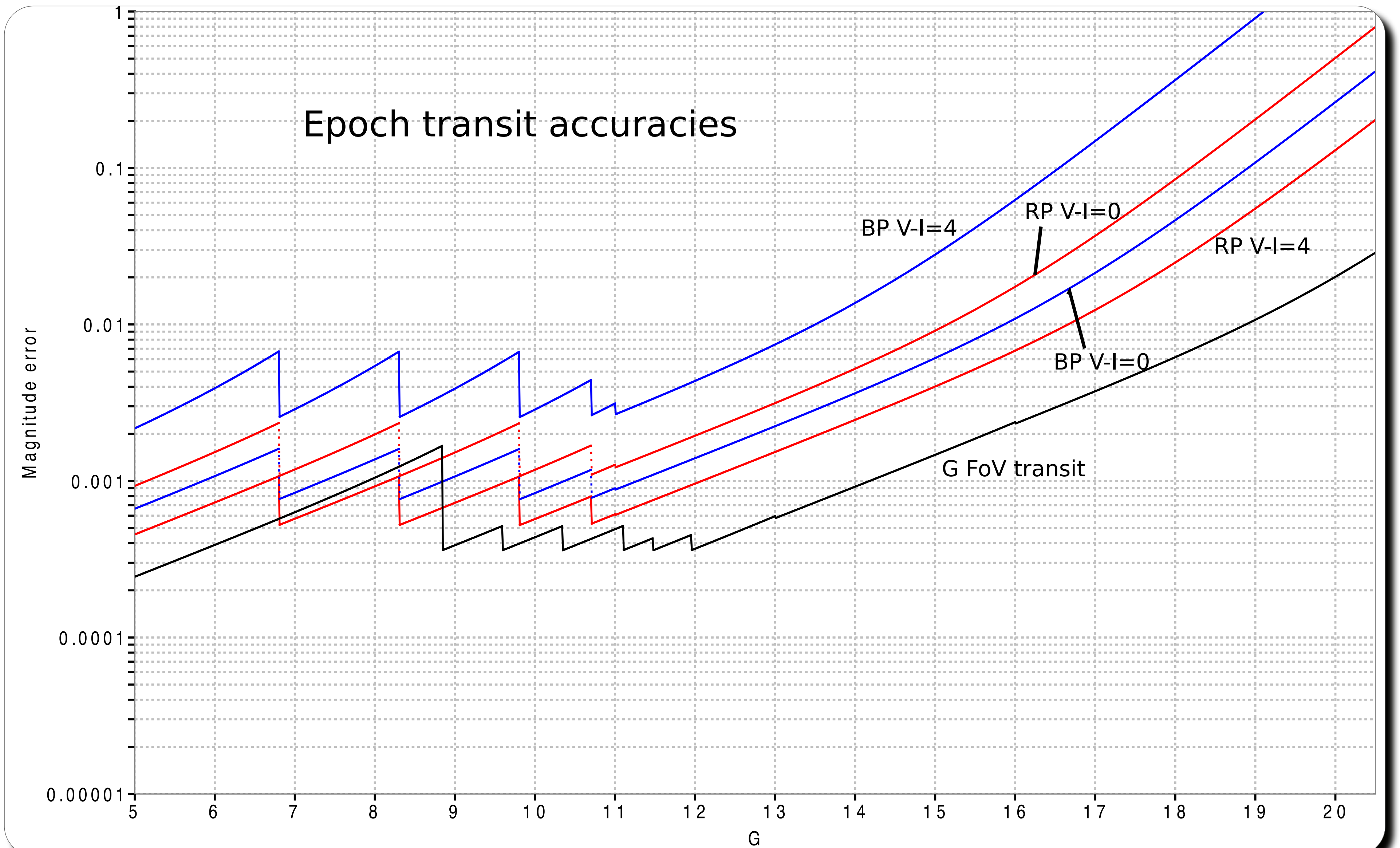
Focal Plane

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

106 CCDs ~ 938 million pixels ~ 2800 cm²

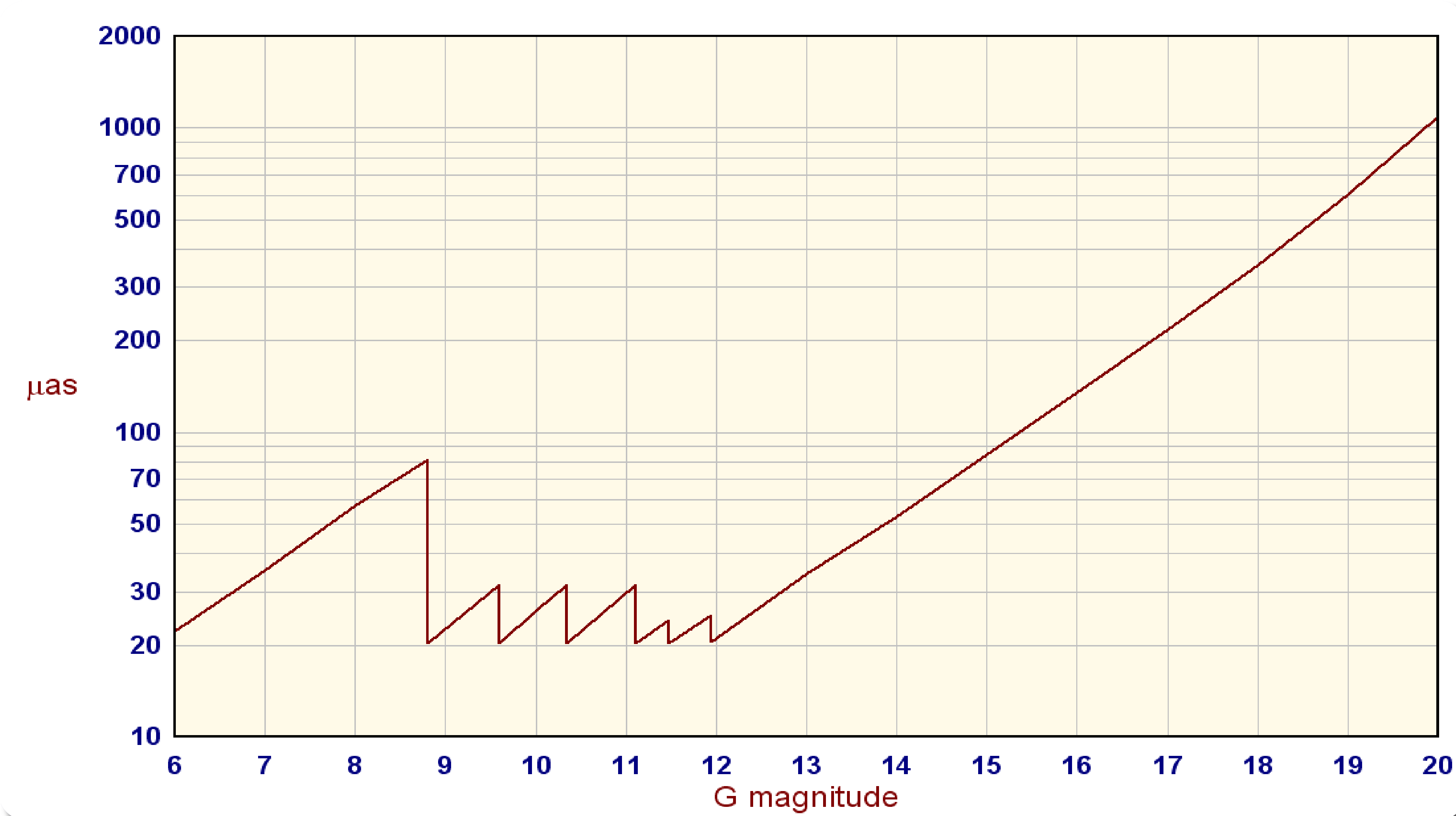


Photometry per transit



- 1% at G=19 (colours to ~10%)
- <2 mmag precision for G<12
- CCD TDI gates avoid pixel saturation for bright stars

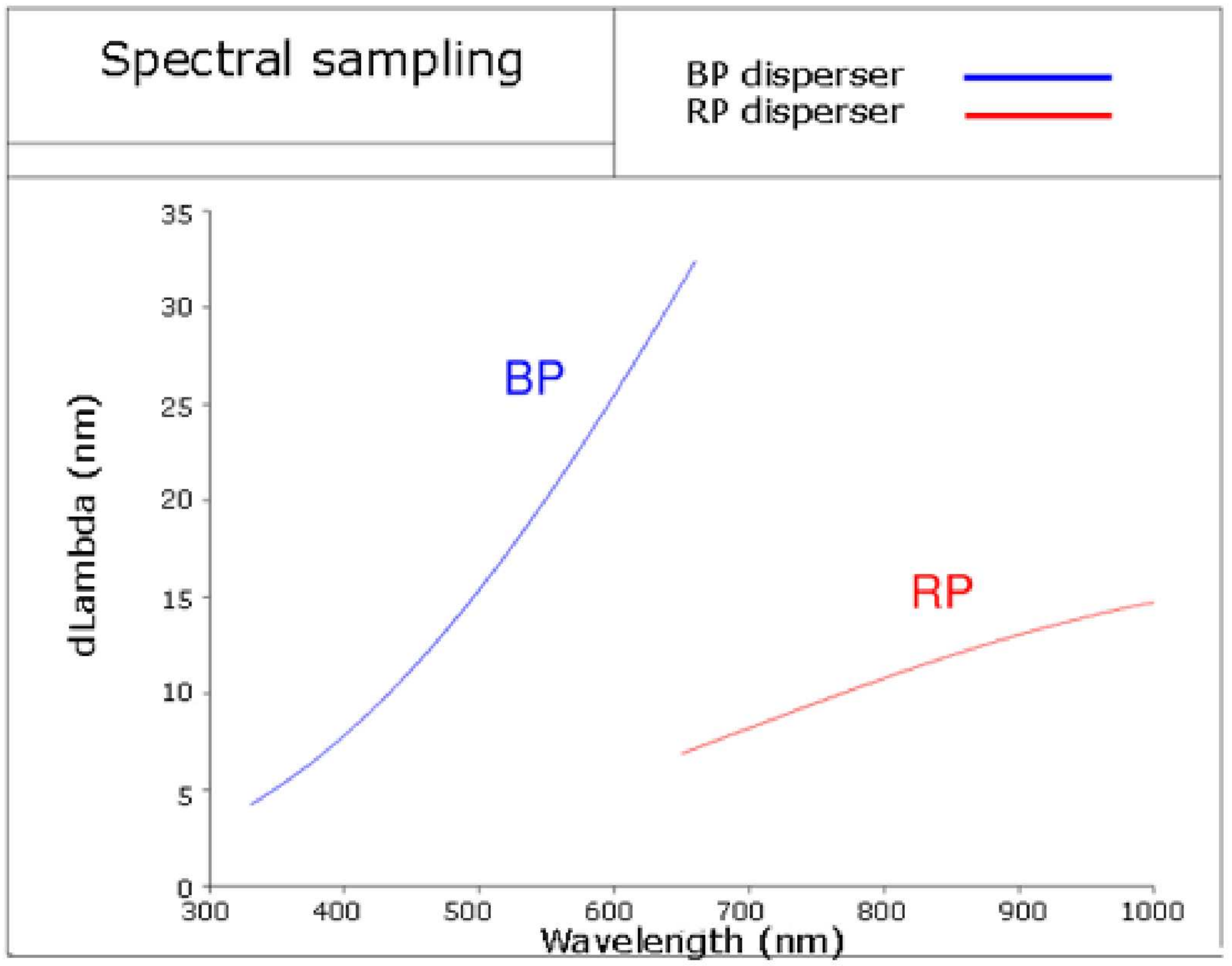
Astrometry per transit



- OGA1: 50 milli arcsec (IDT)
- OGA2: 100 micro arcsec (24hr later)

BP/RP spectra: classification

- two low-res fused-silica prisms
- BP 330-680nm @ 4-32 nm/pixel
- RP 640-1000nm @ 7-15 nm/pixel



50-150 million variables

- 0.5 or 2-3 or 7 million Eclipsing Binaries (Söderhjelm 2004, Eyer & Cuypers 2000, Zwitter 2002)
- 5,000-30,000 Planetary transit systems (Robichon 2002) ?
- 60,000-240,000 δ Scuti stars (Eyer&Cuypers 2000)
- 70,000 RR Lyrae stars (Eyer&Cuypers 2000)
- 2,000-8,000 Cepheids (Eyer&Cuypers 2000).. 9,000 (Windmark et al. 2010)

Outline

- Transient Astronomy
- The Gaia mission
- Detecting Science Alerts with Gaia
- Supernovae and Microlensing
- Details: Classification from Photometry
- Details: Classification from Spectroscopy
- Verification and Follow-up

Timeline for Data Flow: new

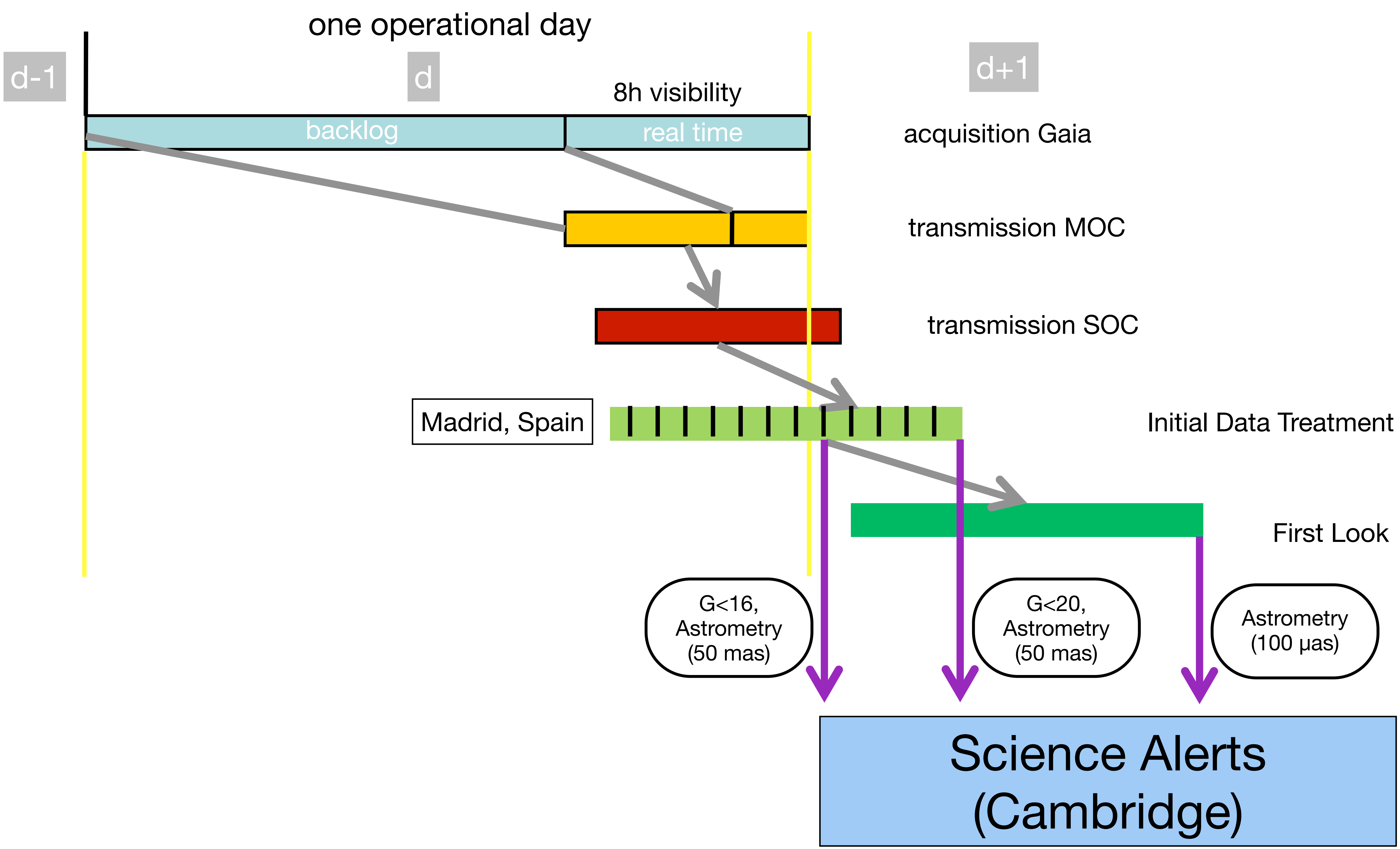
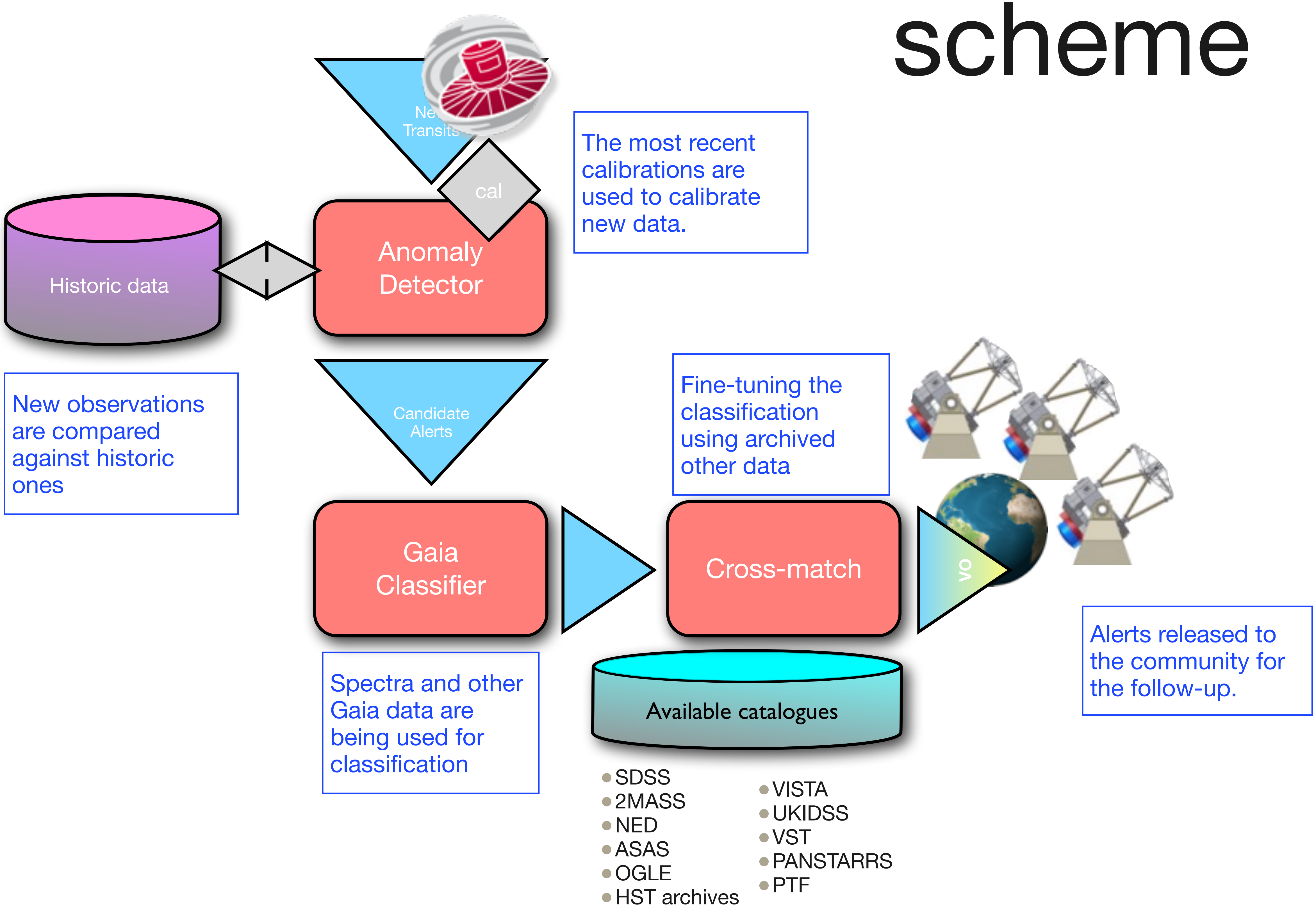


Figure courtesy Francois Mignard, updated by LW+STH

Operational scheme

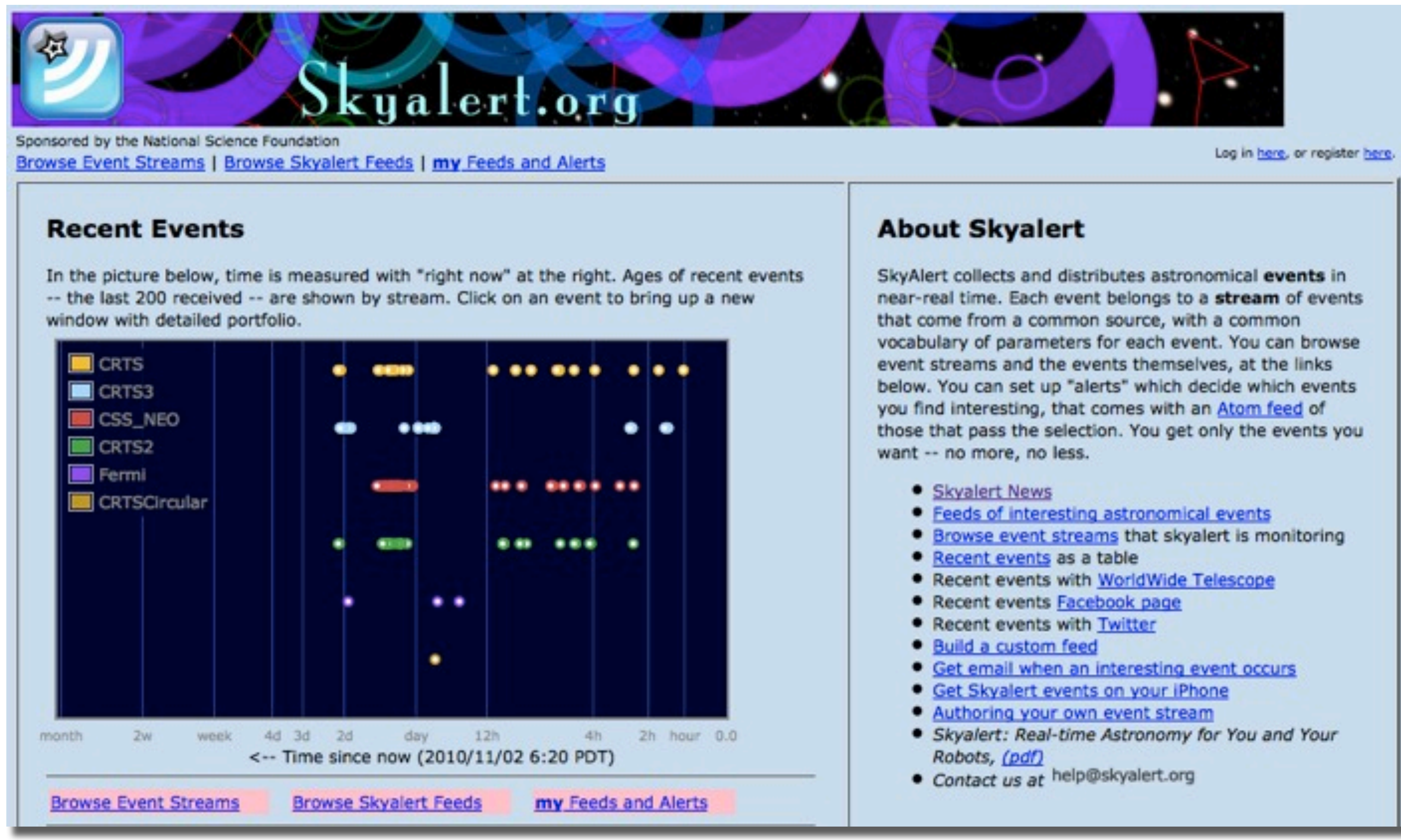


Alert dissemination

- Publication of Alerts to the entire community: no proprietary data.

- VOEvent - machine-readable format, can be displayed in e.g. Google Sky

- Skyalert.org - will host both alerts and follow-up data



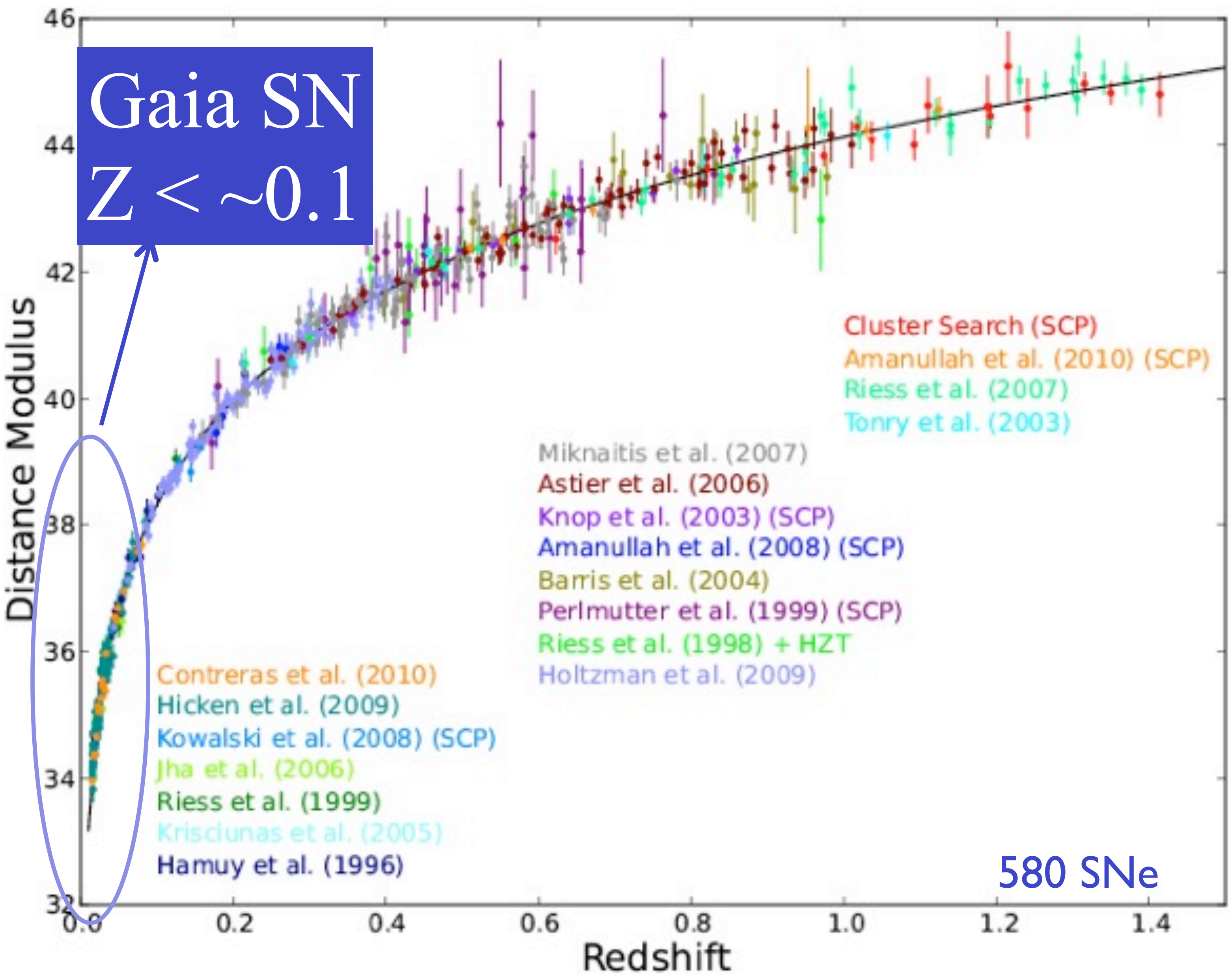
Watch List

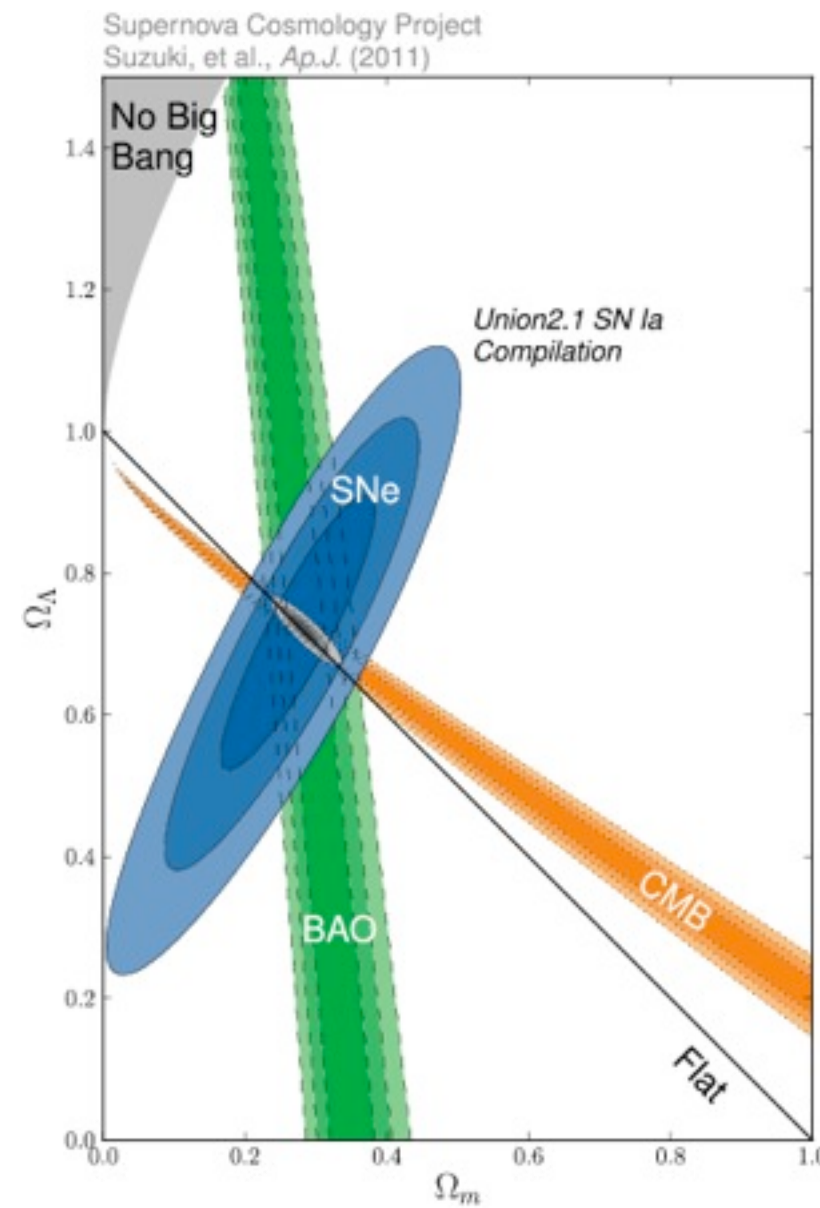
- Known variables will typically be excluded from a transient survey.
- So we will be monitoring a pre-decided set of known interesting objects.
- Flexible - add an object to the list of alerts during the mission.
- Normally detected alerts will end-up in the Watch List.
- Real power will come from comparing with other surveys

Outline

- Transient astronomy
- The Gaia mission
- Detecting transients with Gaia
- Supernovae and microlensing
- Details: Classification from photometry
- Details: Classification from spectroscopy
- Verification and follow-up

Supernova Cosmology





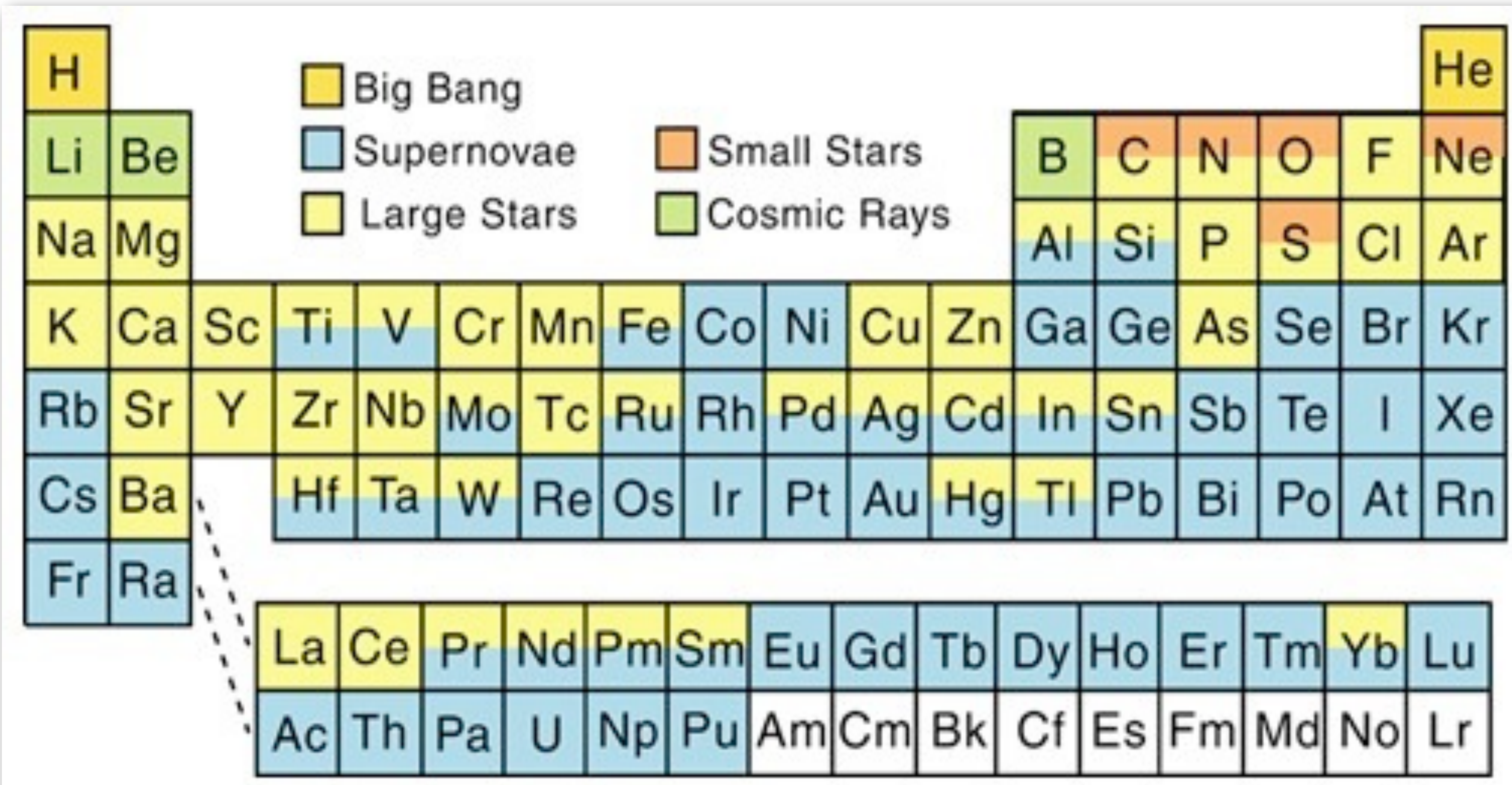
SCP 'Union 2.1' (Suzuki et al 2011)

Why are local Ia's interesting ?

- standard model: thermonuclear explosion of an accreting CO white dwarf, but detailed physics is not understood.
- Progenitor/companion spectral indicators at early times?
- Super-Chandrasekhar (and faint and fast) explosions imply some diversity in progenitors
- Existing follow-up programs don't yet provide adequate rapid spectroscopic/LC monitoring.
- C/O ratios, opacities and temperatures of the ejecta, role of environment (low metallicity dwarf galaxies).. any evidence for evolutionary effects.
- PESSTO survey (Public ESO Spectroscopic Survey for Transient Objects, PI: Smartt)

Core collapse SNe

- Important engines of nucleo-synthesis



Core-collapse SNe

Smartt, Annu. Rev. Astron. Astrophys. 2009. 47:63–106

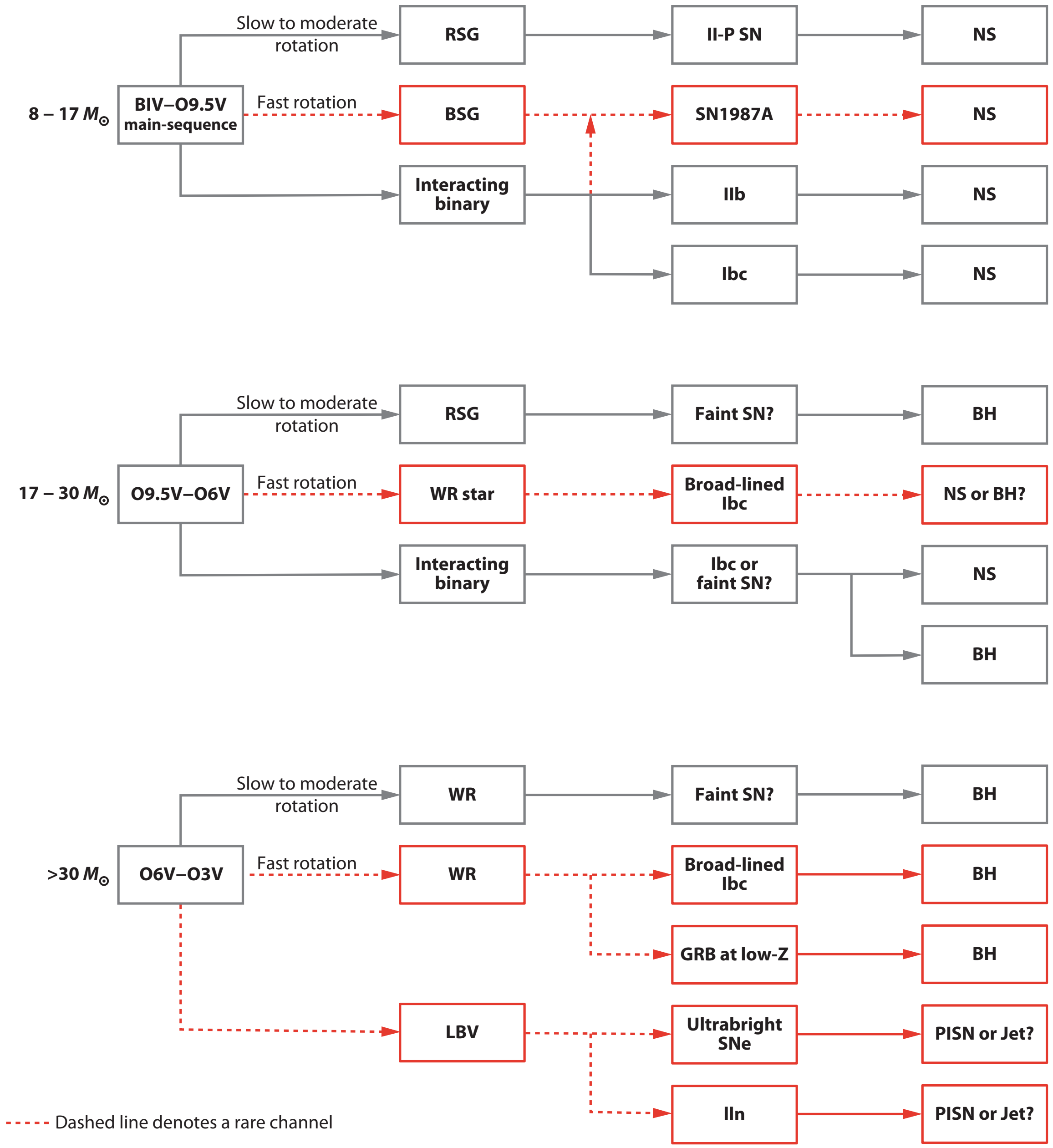
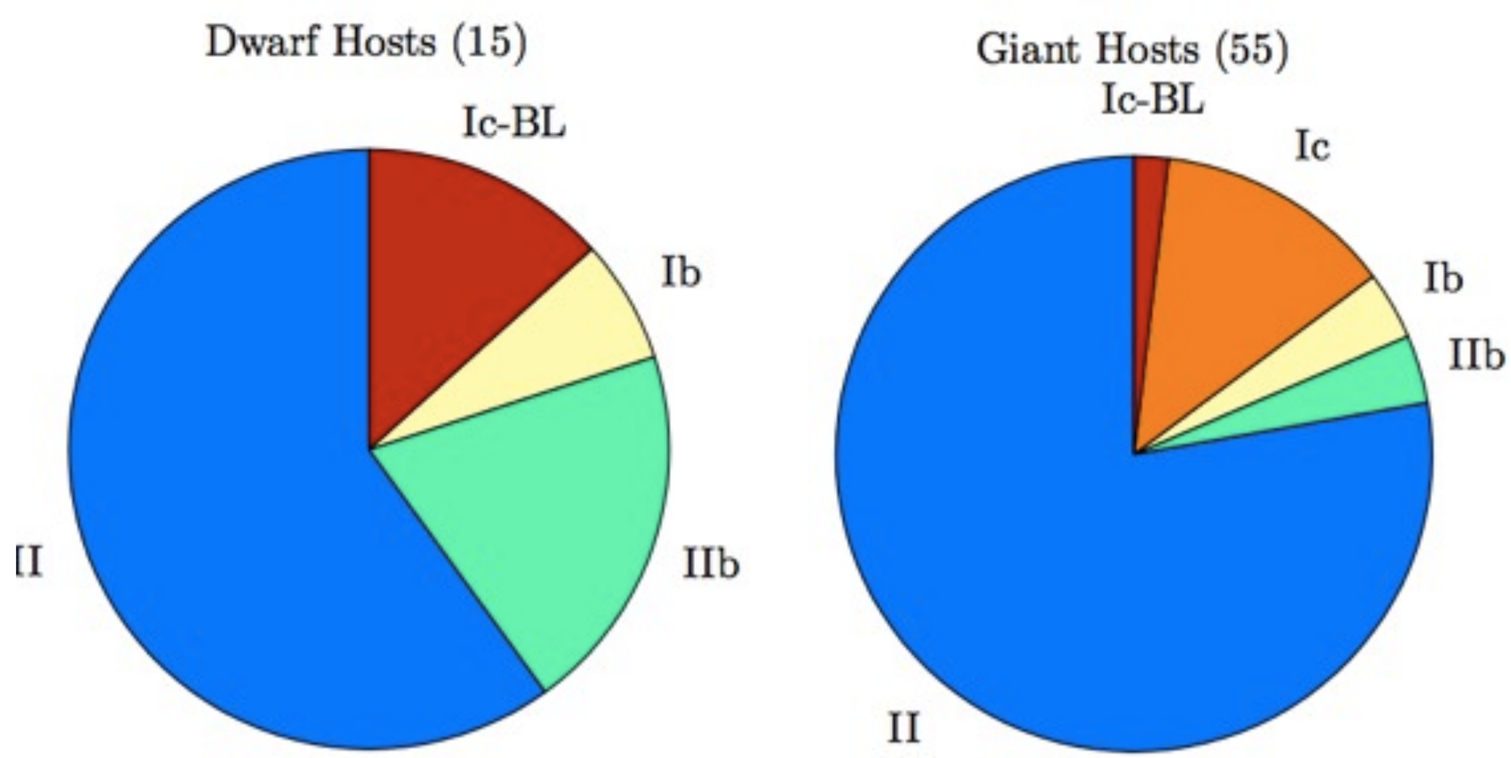


Figure 12
A summary diagram of possible evolutionary scenarios and end states of massive stars. These channels combine both the observational and theoretical work discussed in this review, and the diagram is meant to illustrate the probable diversity in evolution and explosion. It is likely that metallicity, binarity, and rotation play important roles in determining the end states. The acronyms are neutron star (NS), black hole (BH), and pair-instability supernova (PISN). The probable rare channels of evolution are shown in red. The faint supernovae are proposed and have not yet been detected.

- Significant interest in rates of various sub types of SN. Large followup campaign (PESSTO – PI: Smartt) aimed at characterisation and analysis of large sample (ESO:NTT 2012-2017)
- Require as input high quality triggers – Gaia potential source.
- Early detection and flagging of SN in the alert stream is important.

Need larger samples

- Large “unbiased” samples of core-collapse supernovae: the role of environment. (*SNe with no host galaxies, in dwarfs, in metal poor hosts*)
- Connecting progenitors to events.
- Local Ia SNe for the calibration of the SNe cosmology project

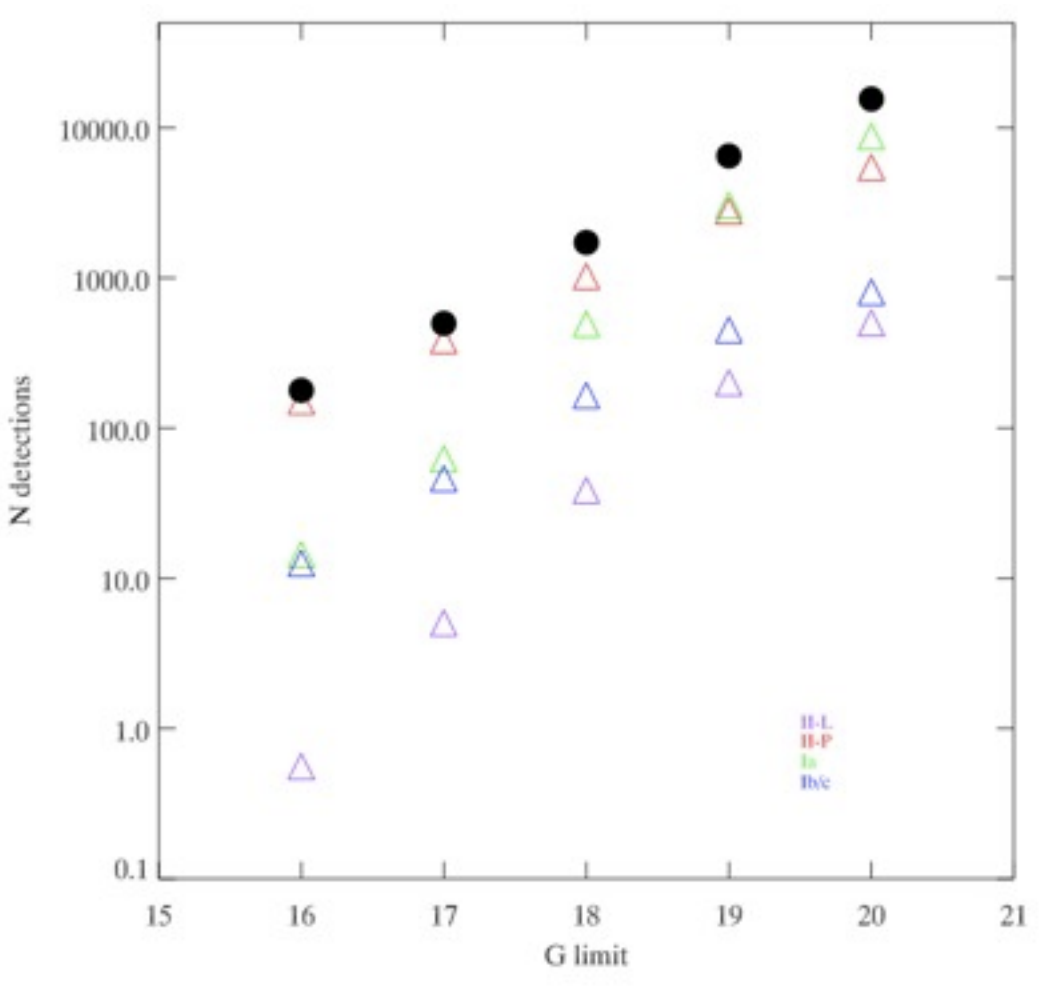


Core-Collapse Supernovae from the Palomar Transient Factory: Indications for a Different Population in Dwarf Galaxies, Arcavi et al., 2010

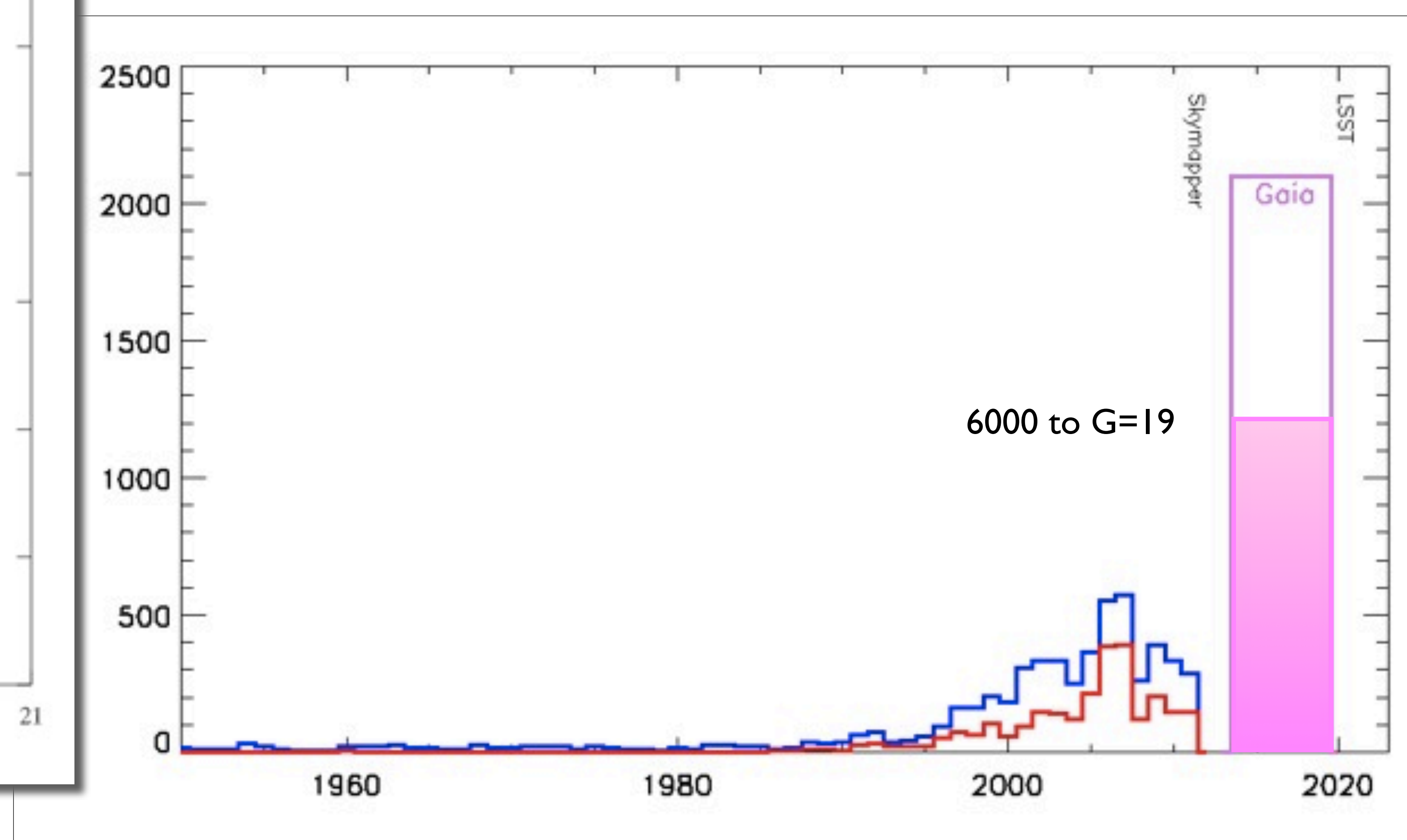
SNa discovery rates

GAIA-C5-TN-IOA-SHO-001-00

In this note, we revisit the simulations of Belokurov and Evans (2003), and make adjustments to account for the current mission parameters (the biggest change from the published numbers arising from the scanning law which roughly halves the number of observations of each part of the sky).

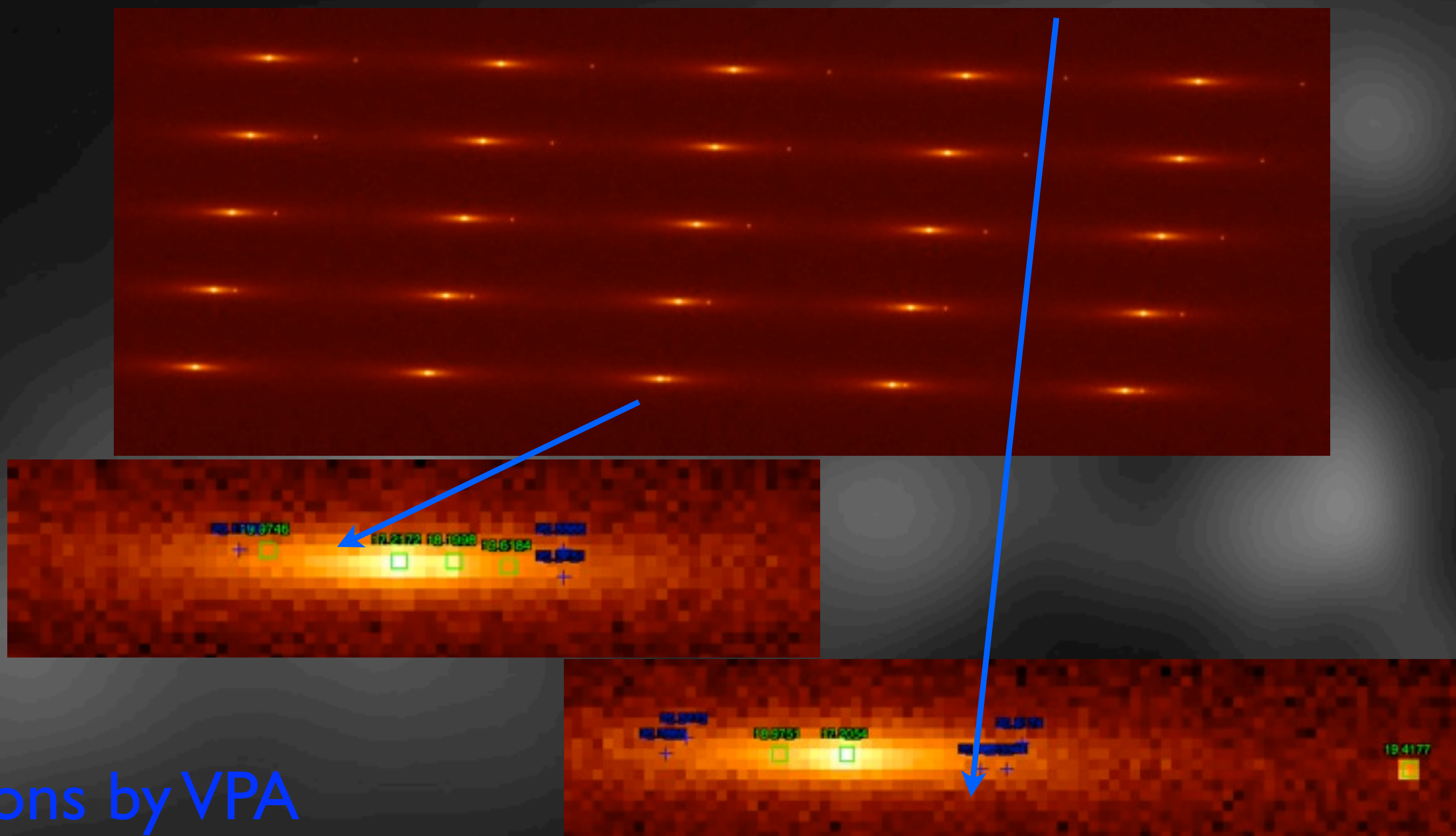


http://www.cbat.eps.harvard.edu/lists/Supernovae.html



See also: *Supernovae and Gaia*, Altavilla et al. 2011, 2012Ap&SS.tmp...66A

Supernovae simulated at increasing distance from a galaxy



Detections by VPA

brighter than 20mag

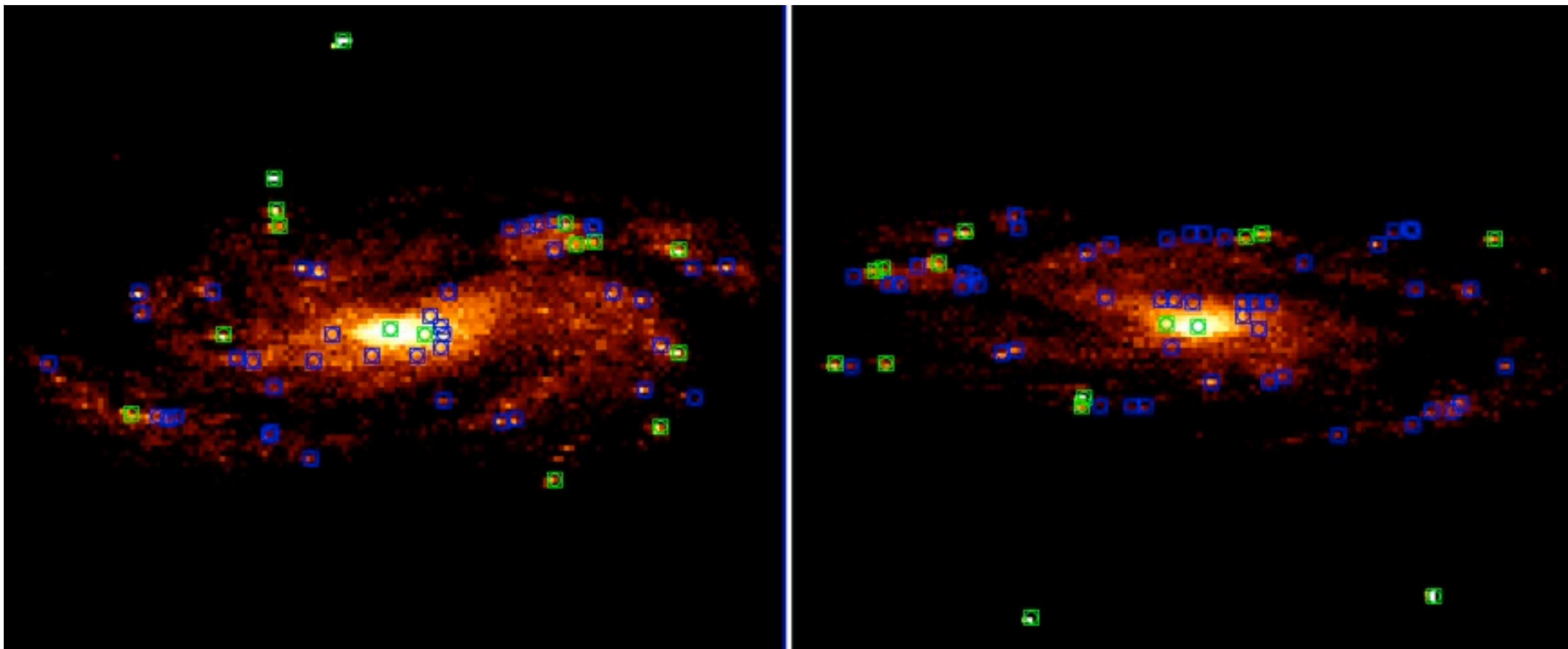
No windowing information

No priorities for transmission

resolved into ~0.1 arcsecond

What will Gaia see?

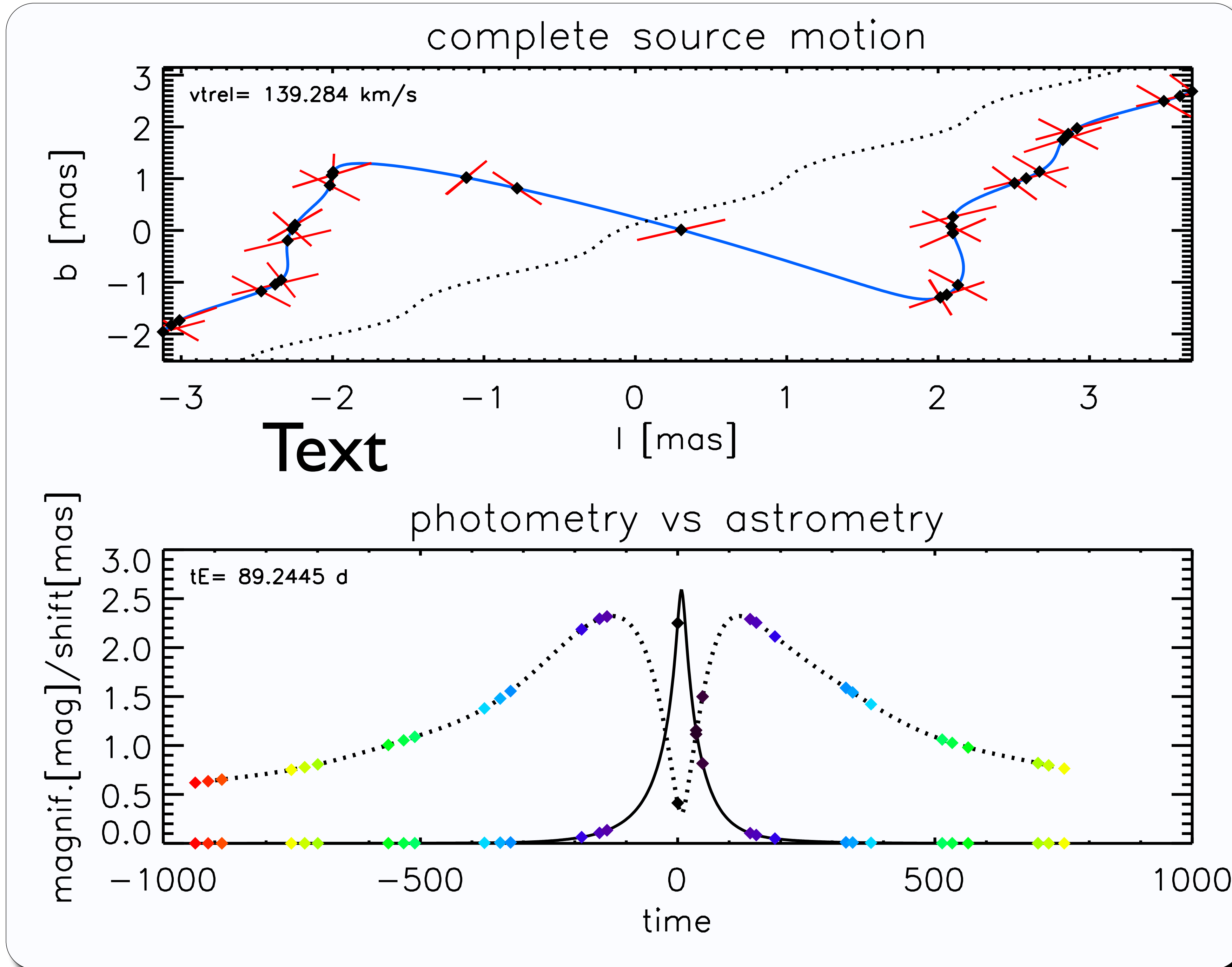
Simulated data for M100, based on the HST image. Each transit gives a different scanning angle of the galaxy, hence might reveal different knots.



blue: all detections from VPA   green: detections brighter than 20 mag

Astrometric microlensing

- Combine astrometric and photometric events to solve for the lens mass.
- discovery of astrometric **Black Hole** and **Brown Dwarf** lensing events would be a significant first for Gaia.
- **BH** events will cause astrometric signals around 2 milliarcsec (distance dependent). **BD** events are an order of magnitude smaller.



simulation from Lukasz Wyrzykowski: ds=8.6kpc, dl=1kpc, ml=6msun

Outline

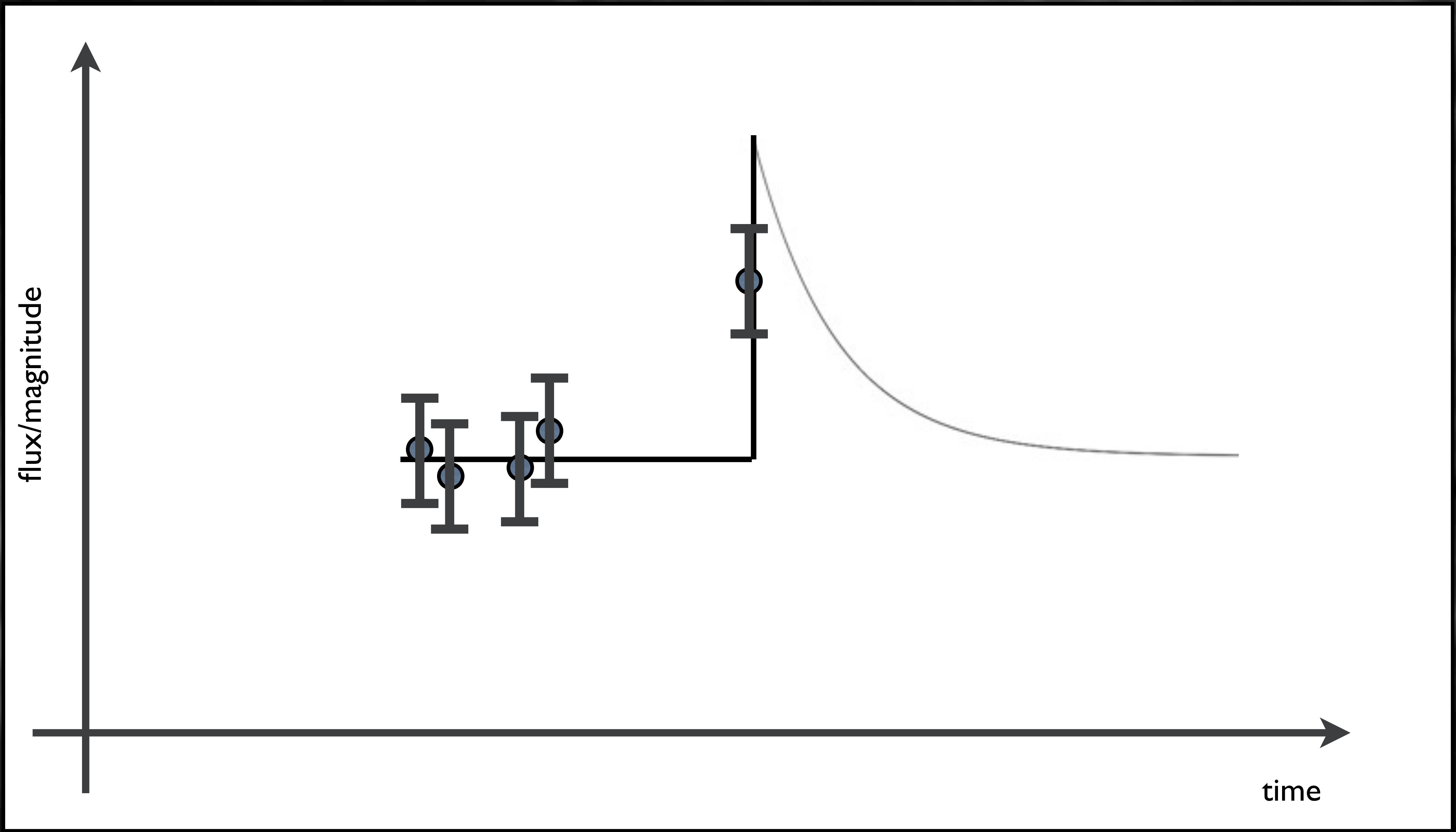
- Transient astronomy
- The Gaia mission
- Detecting transients with Gaia
- Supernovae and microlensing
- Details: Classification from photometry
- Details: Classification from spectroscopy
- Verification and follow-up

Automated Classification

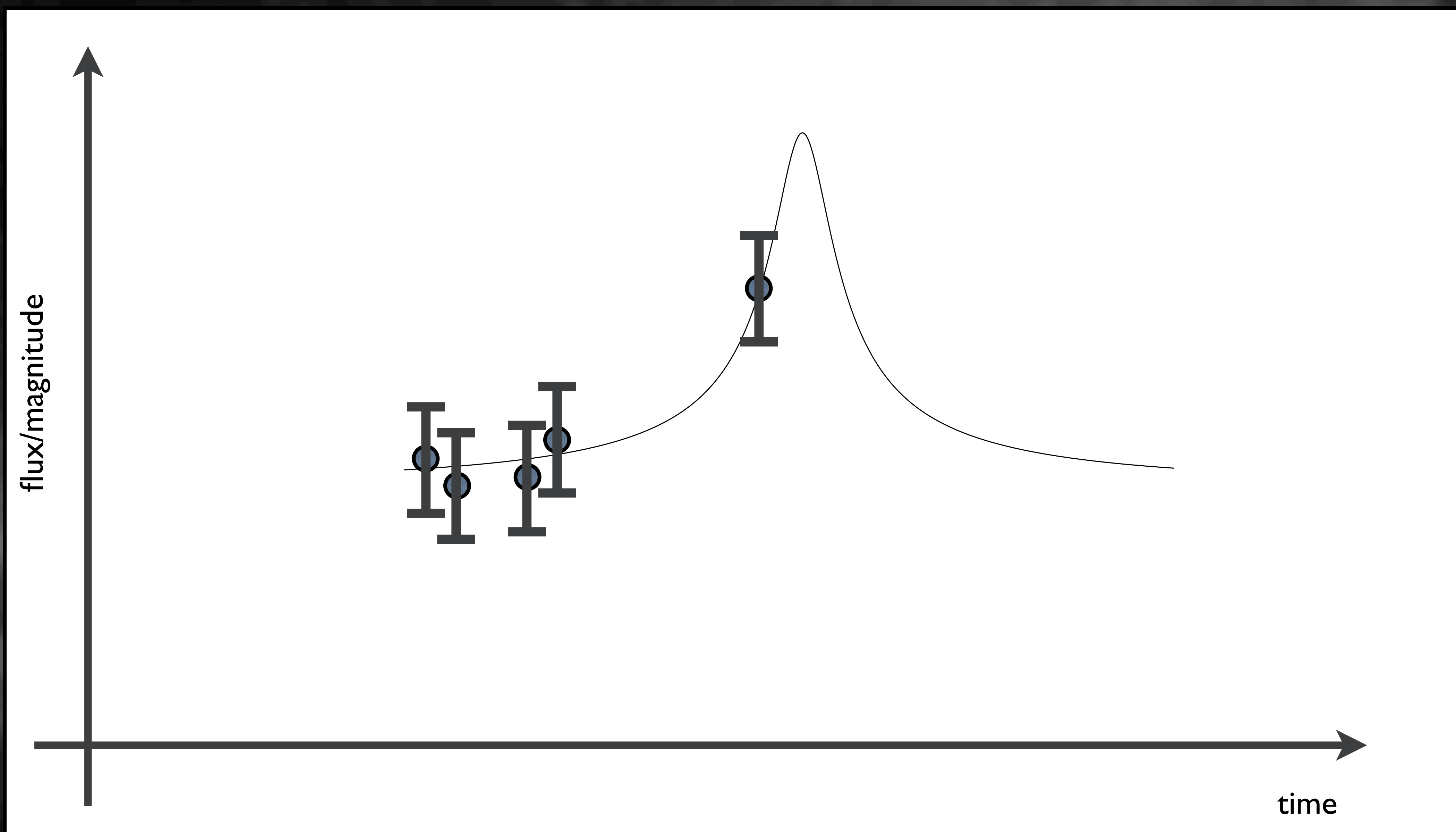
- 44 million transits per day
- 150->800 GByte/day
- we expect 100s-1000s of potential astrophysical triggers per day (real variables/moving objects).
- additional contaminants from noise (dominated by systematics)
- this precludes visual classification of a rich data stream
- based on streaming data (i.e. not waiting until the lightcurve is complete)
- automated methods are fast, repeatable and tuneable

get them out the door   assess completeness/errors   adjust strategy

Transient light curve classification challenges: ... this ?



Transient light curve classification challenges: ... or this ?

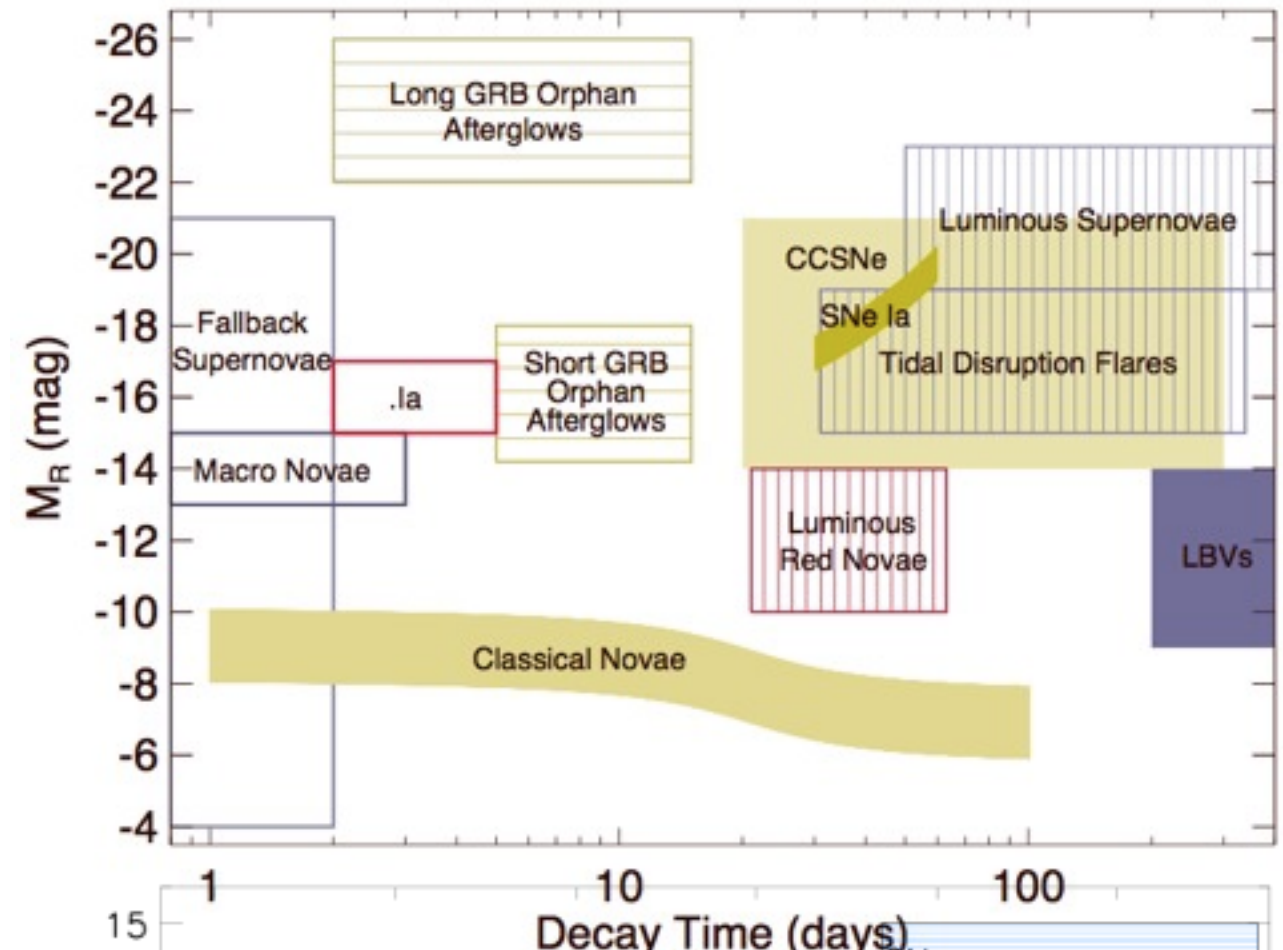


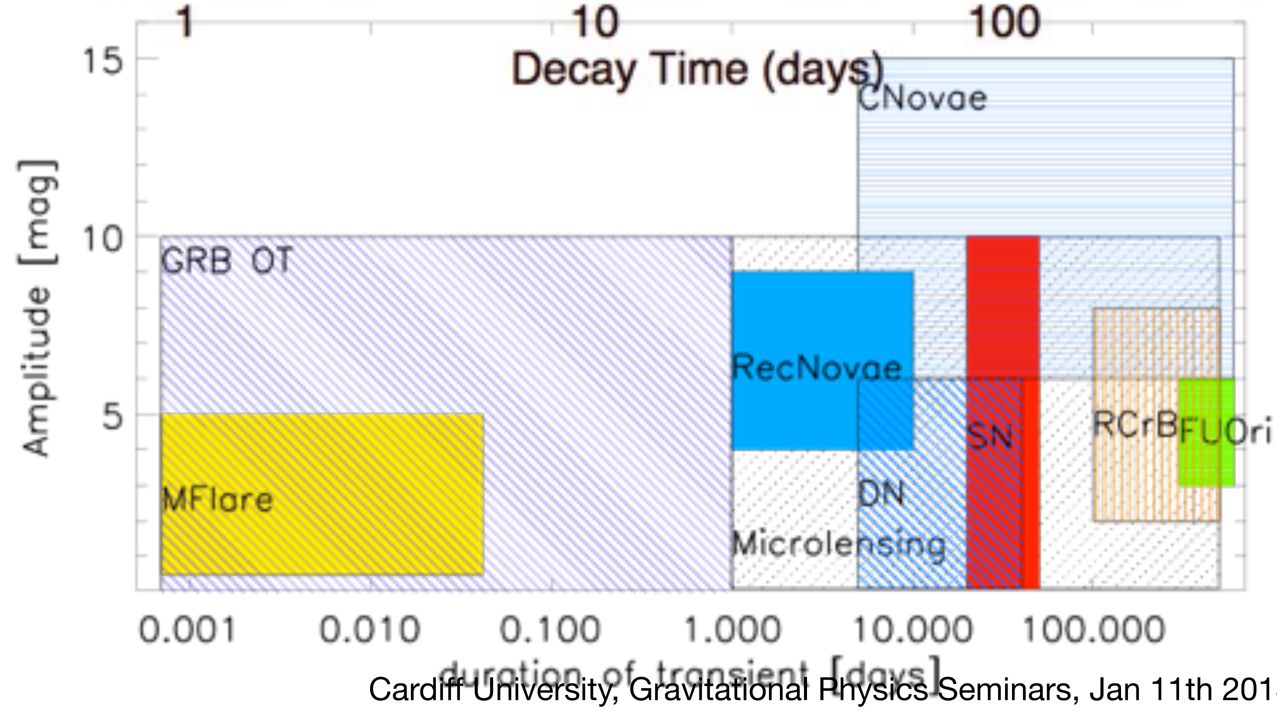
Features

- signal vs background events distinguished on the basis of features
- e.g. magnitude, signal-to-noise, fwhm: star/galaxy/cosmics
- the problem is how to maximise separation between classes of event

Strategies for classification

- multi-parameter thresholding (e.g. changes in amplitudes)
  tends to treat parameters as independent



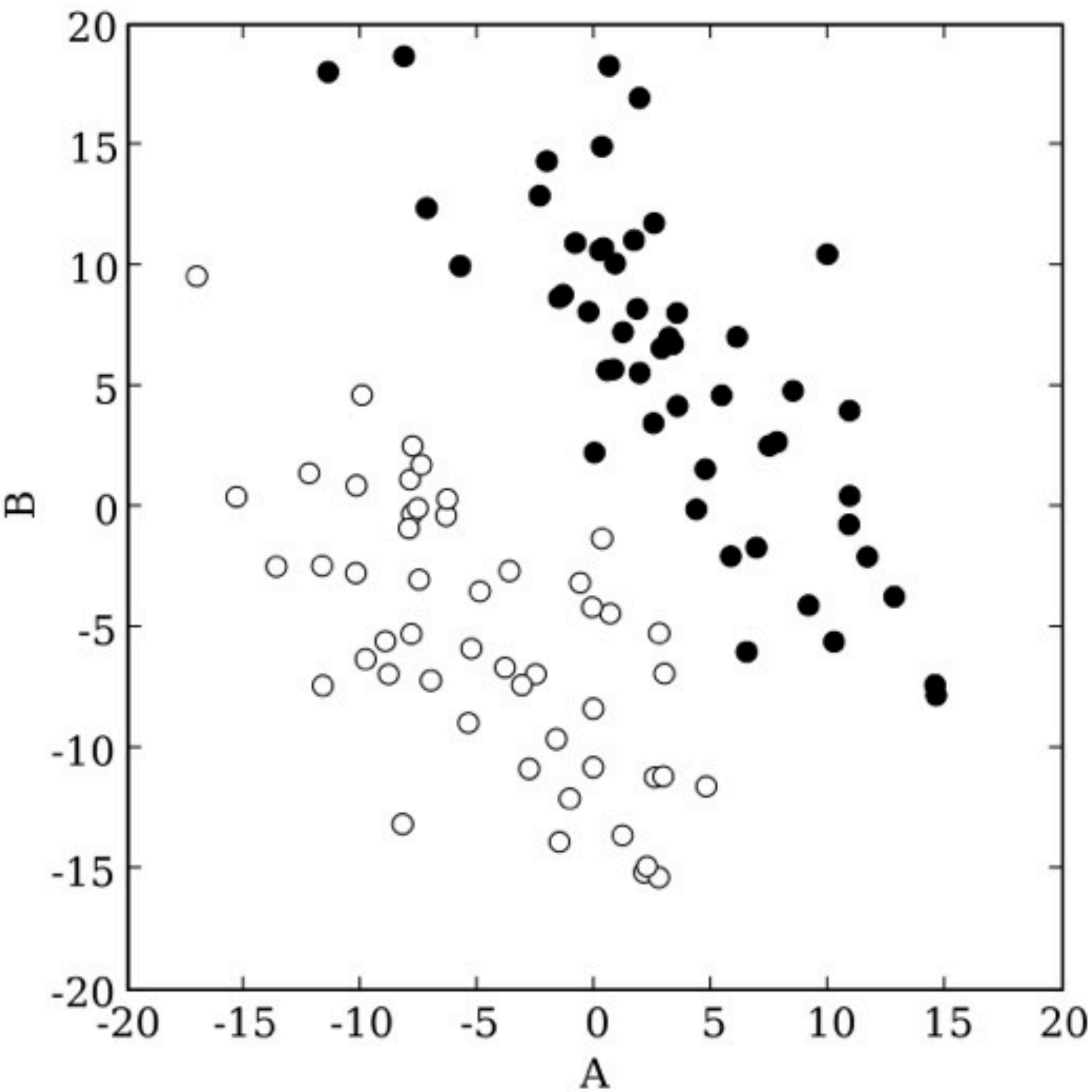


Strategies for classification

- multi-parameter thresholding (e.g. changes in amplitudes)

  tends to treat parameters as independent

  parameters are often correlated

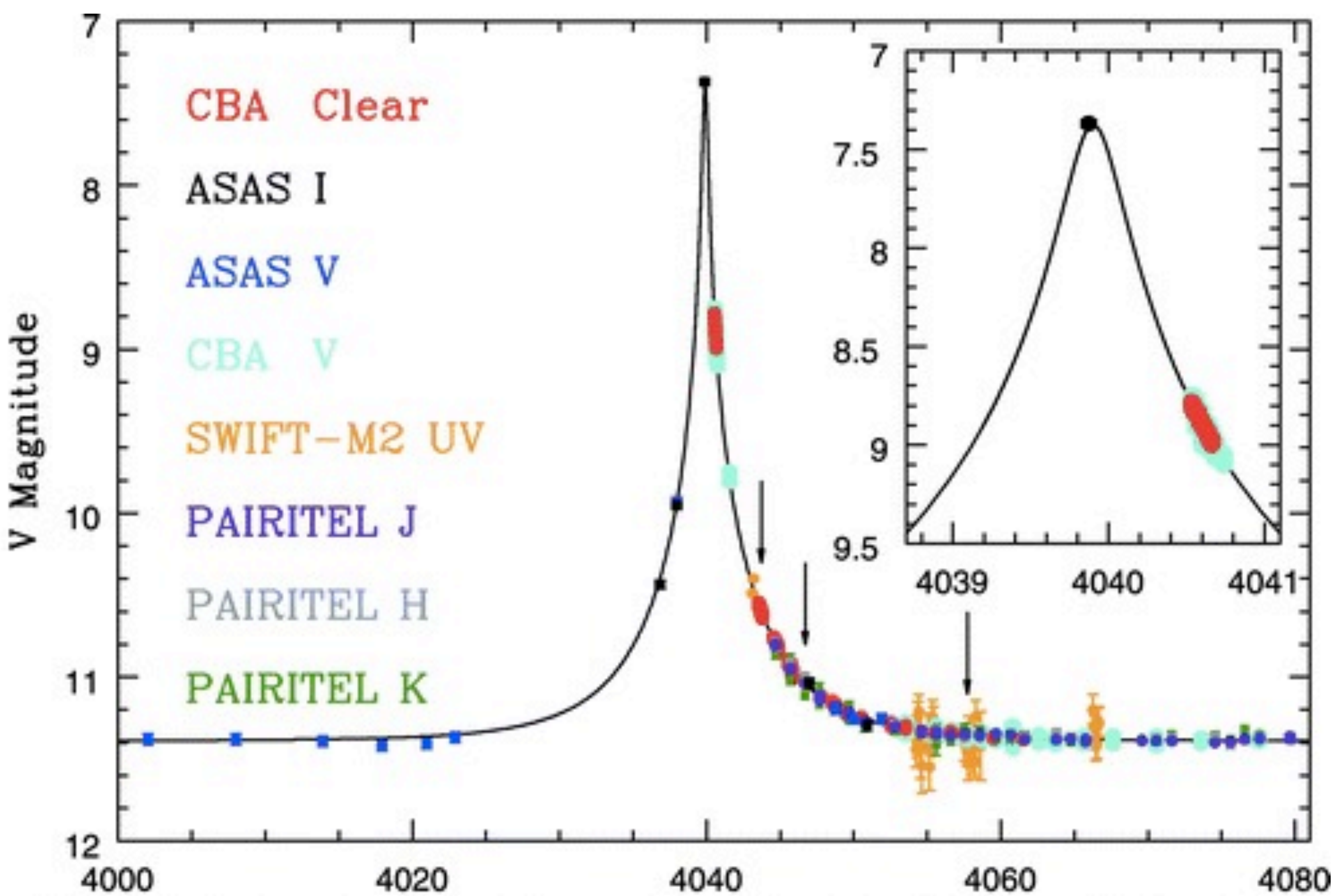


Strategies for classification

- multi-parameter thresholding (e.g. changes in amplitudes)
- matched filters (compare LC to model/template)

ignores all the other classes but pretty much necessary for some cases

Microlensing events



Strategies

- class-specific machine-learning methods (e.g. decision trees, random forests, support vector machines, automated neural networks)

significant improvement in efficiency for SN factory, Bailey et al. 2007

the values of A,B,C,D are arrived at by training

this decision allows for a distinction between high and low SNR detections

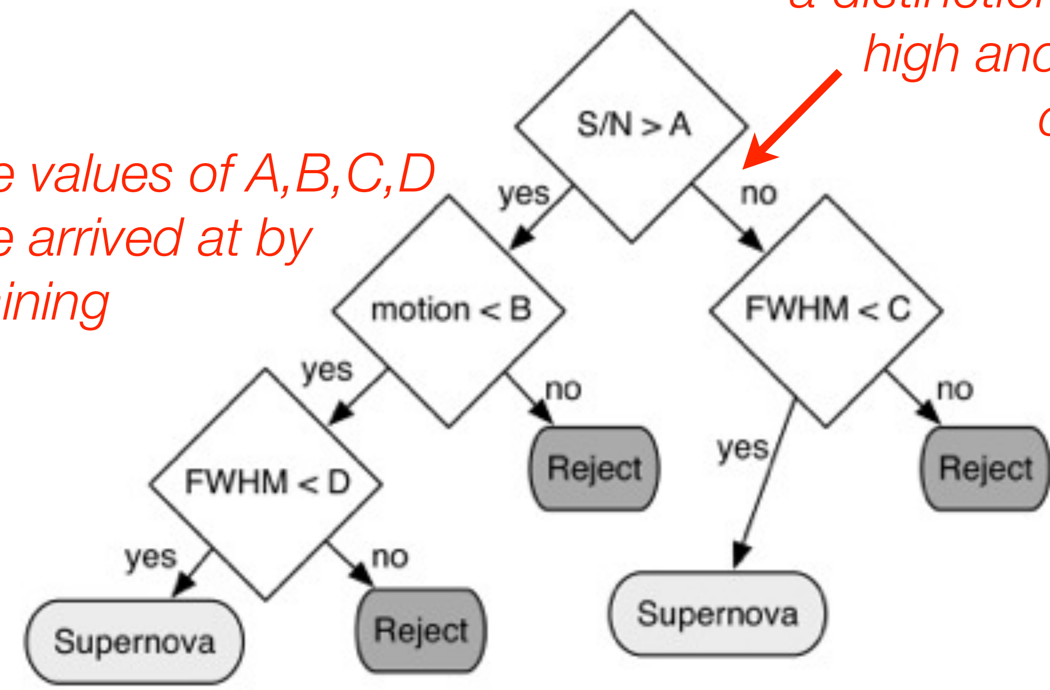


FIG. 2.—Example decision tree that would treat high signal-to-noise ratio objects differently from low signal-to-noise ratio objects. In practice, a real decision tree has many more branches and the same variable can be used to branch at many different locations with different cut values.

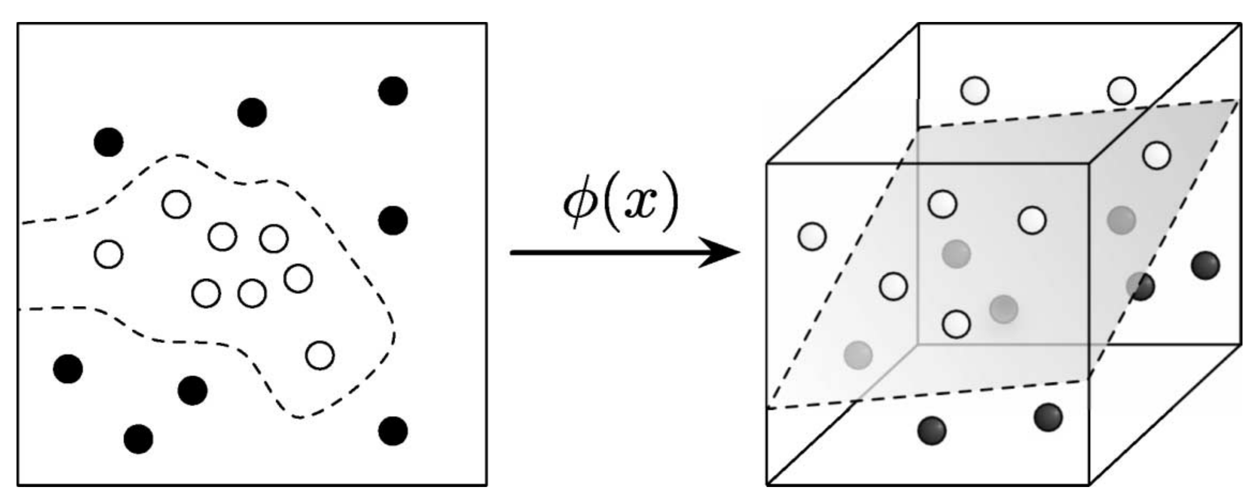


FIG. 3.—Illustration of a support vector machine. An input space of features is mapped into a higher dimensional space where the separation of classes becomes easier. The separation boundary in the original space may be quite complex, even disjoint. In the higher dimensional space, the separation surface is a hyperplane whose parameters are entirely determined by the subset of events (the support vectors) nearest to the boundary.

Decision Trees

- Decision trees (Breiman et al. 1984) separate signal from background events by making a cascading set of event splits (generalization of threshold cuts)
- A training procedure automatically selects the features and cut values to generate a tree with maximal separation of signal and background events.
- However, a small change in the training set can produce a considerably different tree.
- Boosting algorithms improve the performance of a classifier by giving greater weight to events that are hardest to classify.

Bailey et al. 2007, *HOW TO FIND MORE SUPERNOVAE WITH LESS WORK: OBJECT CLASSIFICATION TECHNIQUES FOR DIFFERENCE IMAGING*

Random Forests

- Random forests (Breiman 2001) also generate multiple decision trees for a given training set and use a weighted average of the trees as the final decision metric.
- When training a tree, at each branch the training cycle only considers a random subset of the possible features available to use.
- This has the effect of washing out the typical training instabilities of decision trees and produces a classifier that is fast to train and robust against outliers.

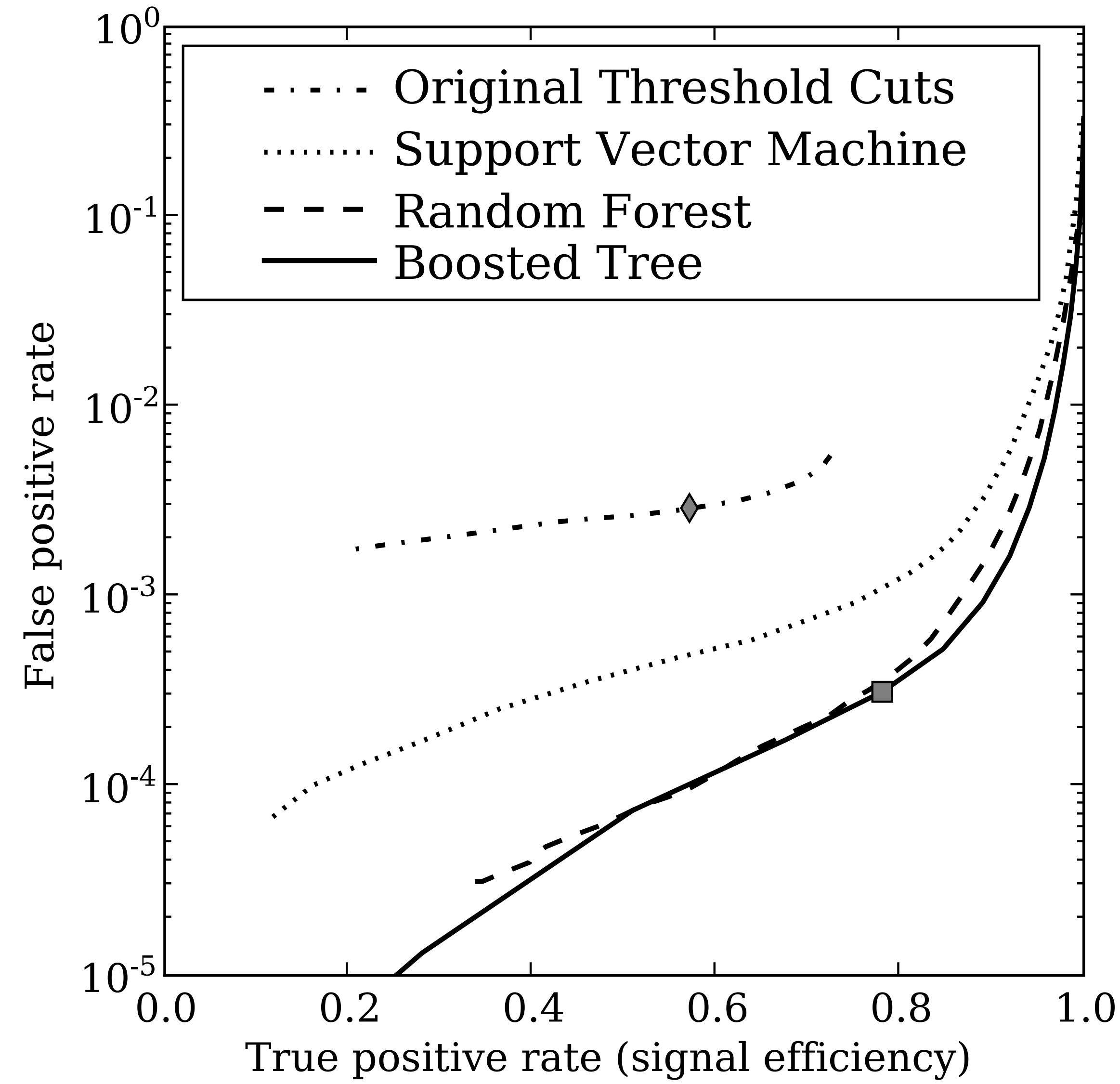


FIG. 4.—Comparison of boosted trees (*solid line*), random forests (*dashed line*), SVM (*dotted line*), and threshold cuts (*dash-dotted line*) for false-positive identification fraction vs. true-positive identification fraction. For the threshold cuts, the signal-to-noise ratio, motion, and shape cuts were varied to adjust signal and background rates. The gray diamond shows the performance of the threshold cuts used during the SNfactory summer 2006 search; the gray square shows the performance achieved with boosted trees, which were used for the fall 2006 SNfactory search. The lower right corner of the plot represents ideal performance. [*See the electronic edition of the Journal for a color version of this figure.*]

Bailey et al. 2007, *HOW TO FIND MORE SUPERNOVAE WITH LESS WORK: OBJECT CLASSIFICATION TECHNIQUES FOR DIFFERENCE IMAGING*

PTF classification

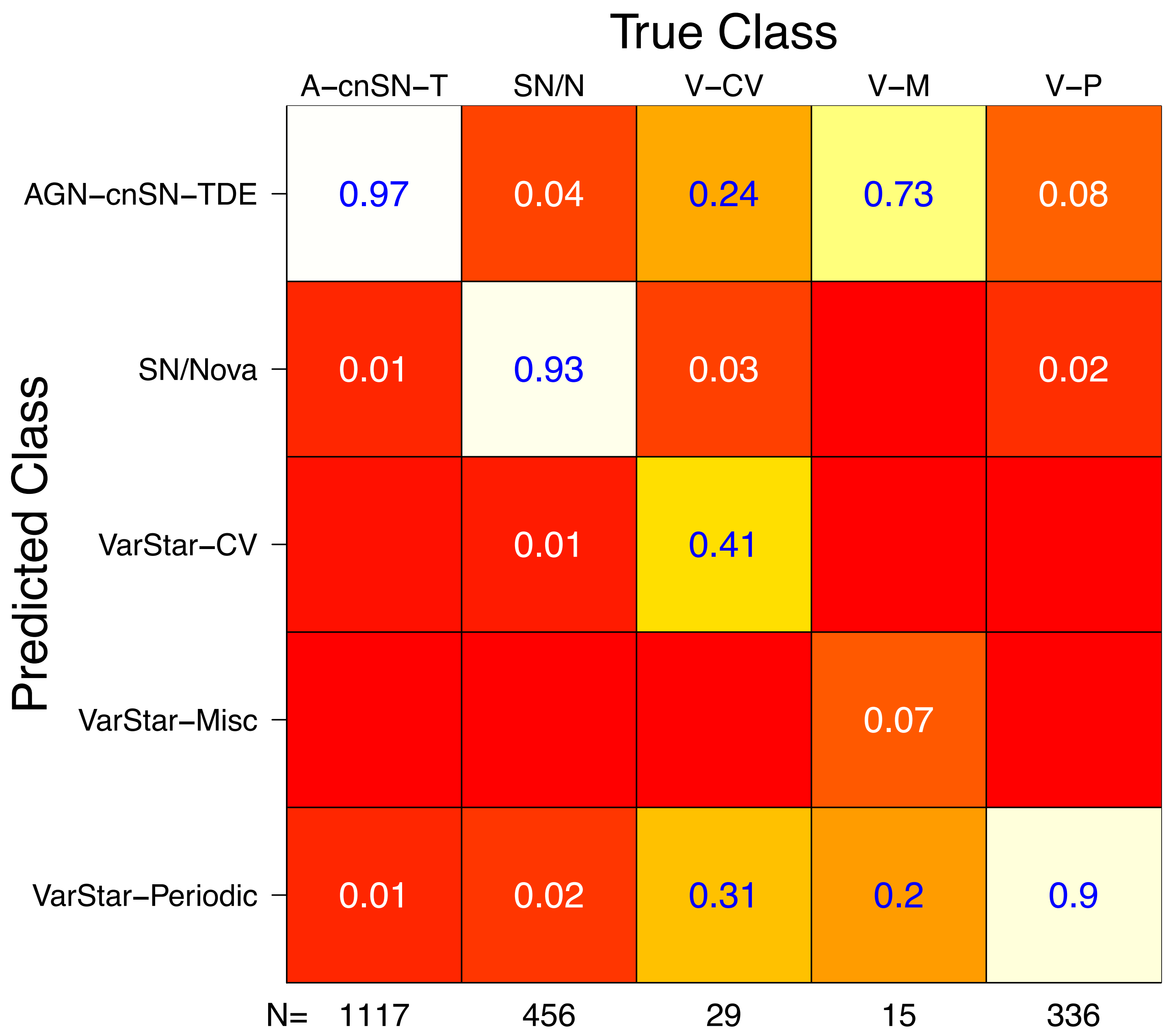
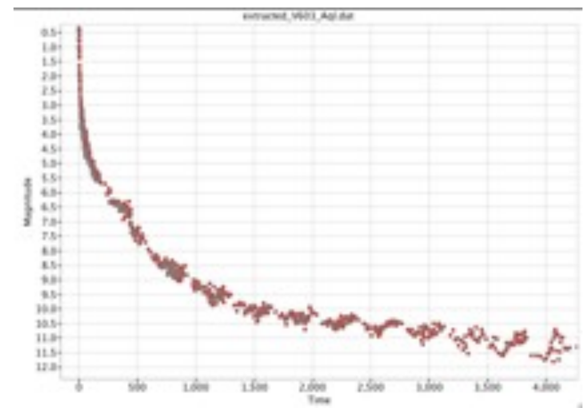


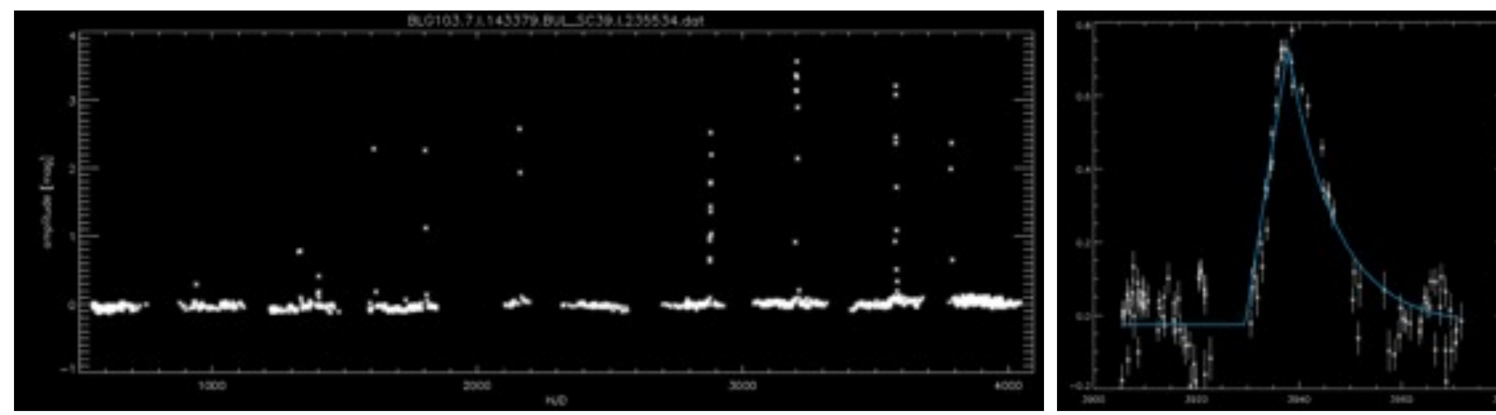
Fig. 12.— Confusion matrix for robotclass random forest classification. Classes are aligned so that entries along the diagonal corresponds to correct classification. Probabilities are normalized to sum to unity for each column. Recovery rates are ≥90%, with very high purity, for the three dominant classes. Classification accuracy suffers for the two classes with small amounts of data (note: class size is written along the bottom of the figure).

Preparation of the Templates

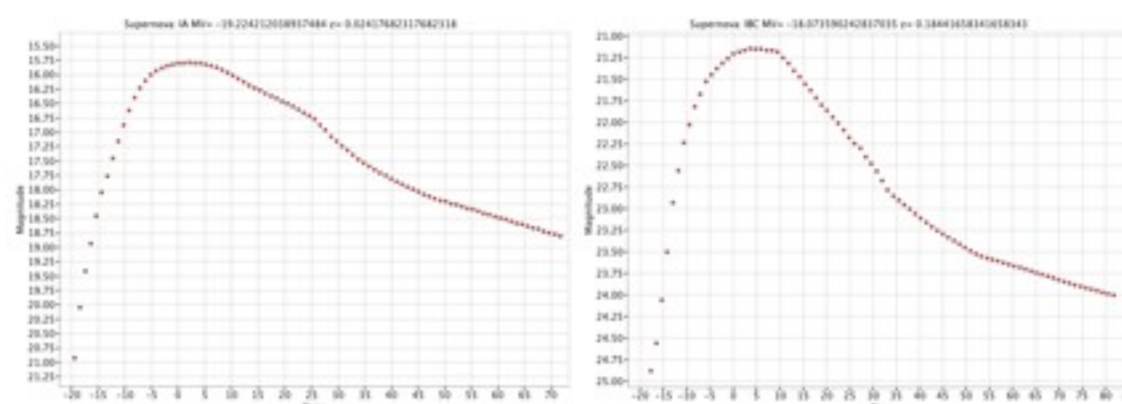
- Novae - from AAVSO



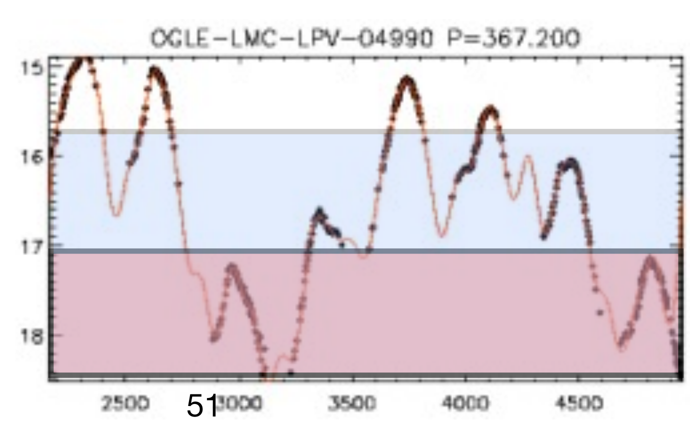
- DNe - model from OGLE



- SNe - all types from Nugent



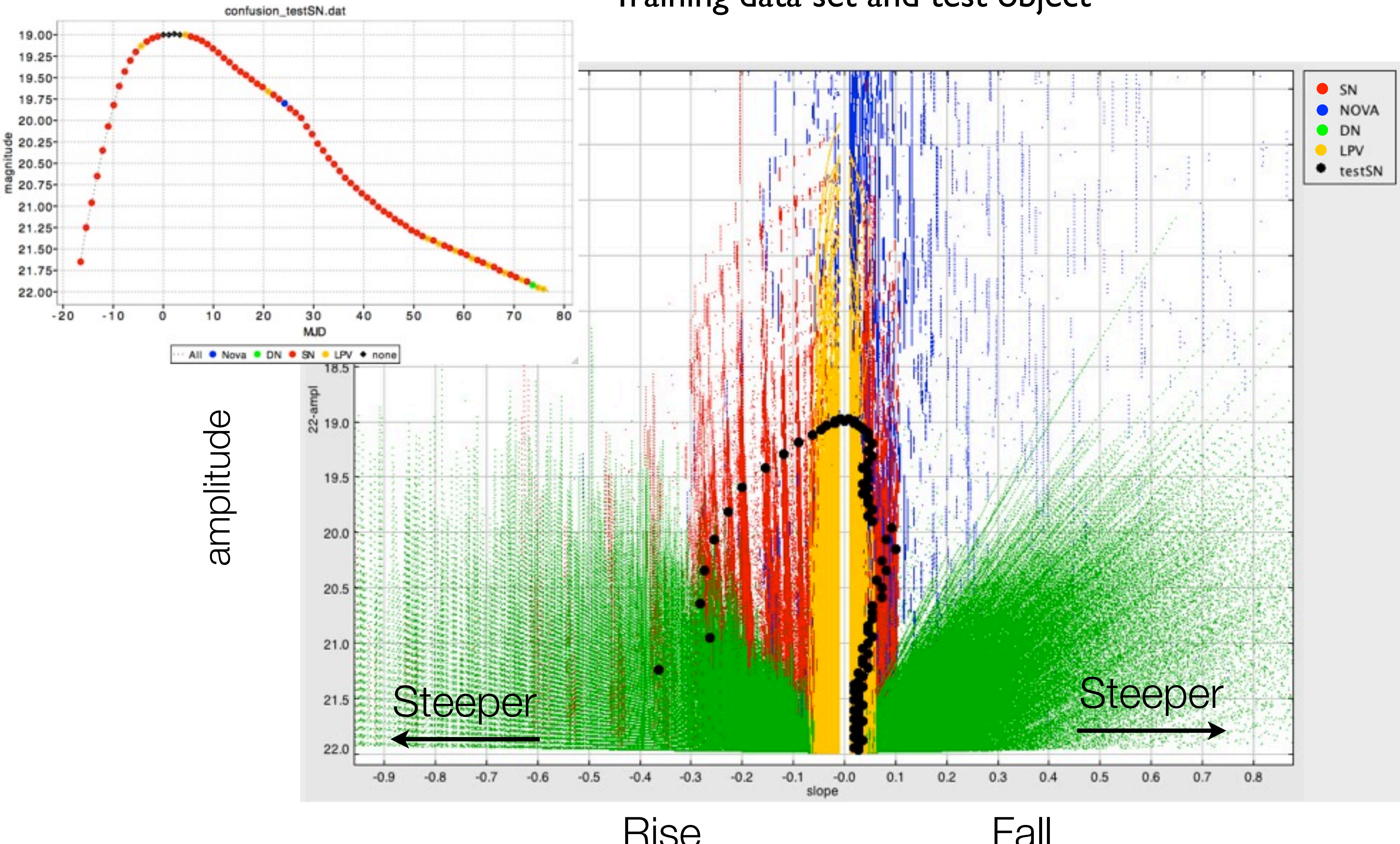
- LPVs - from OGLE



Method: training sets

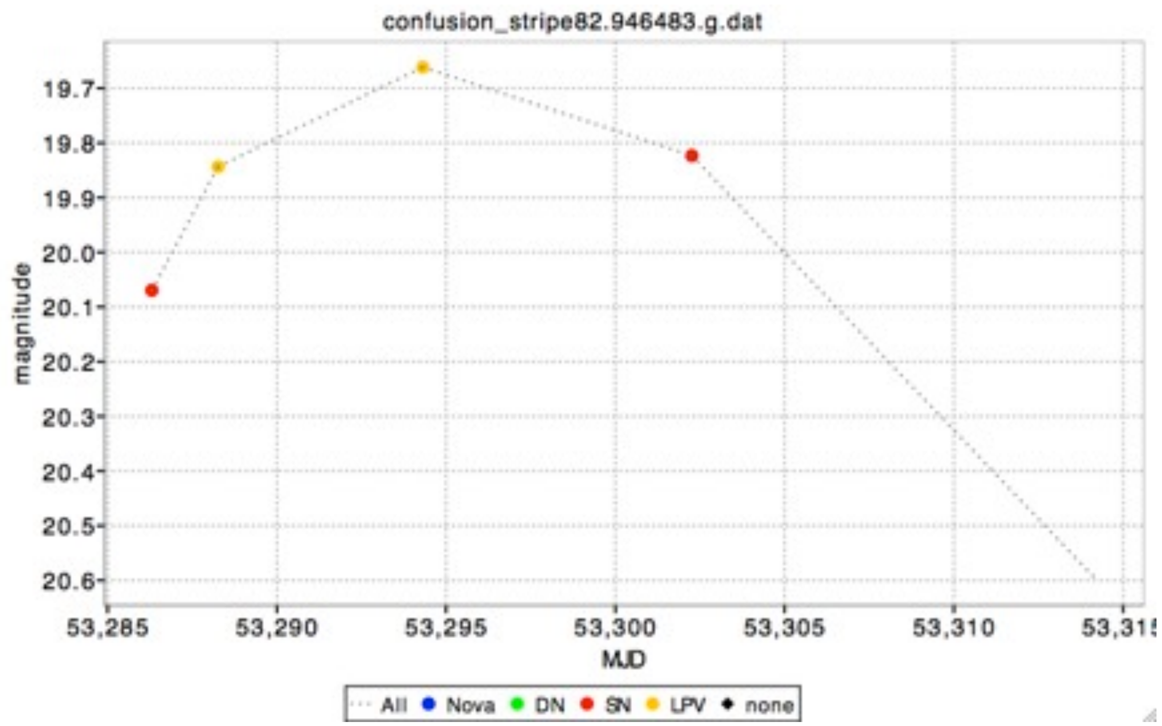
Template light curve of SN Ia

Training data set and test object

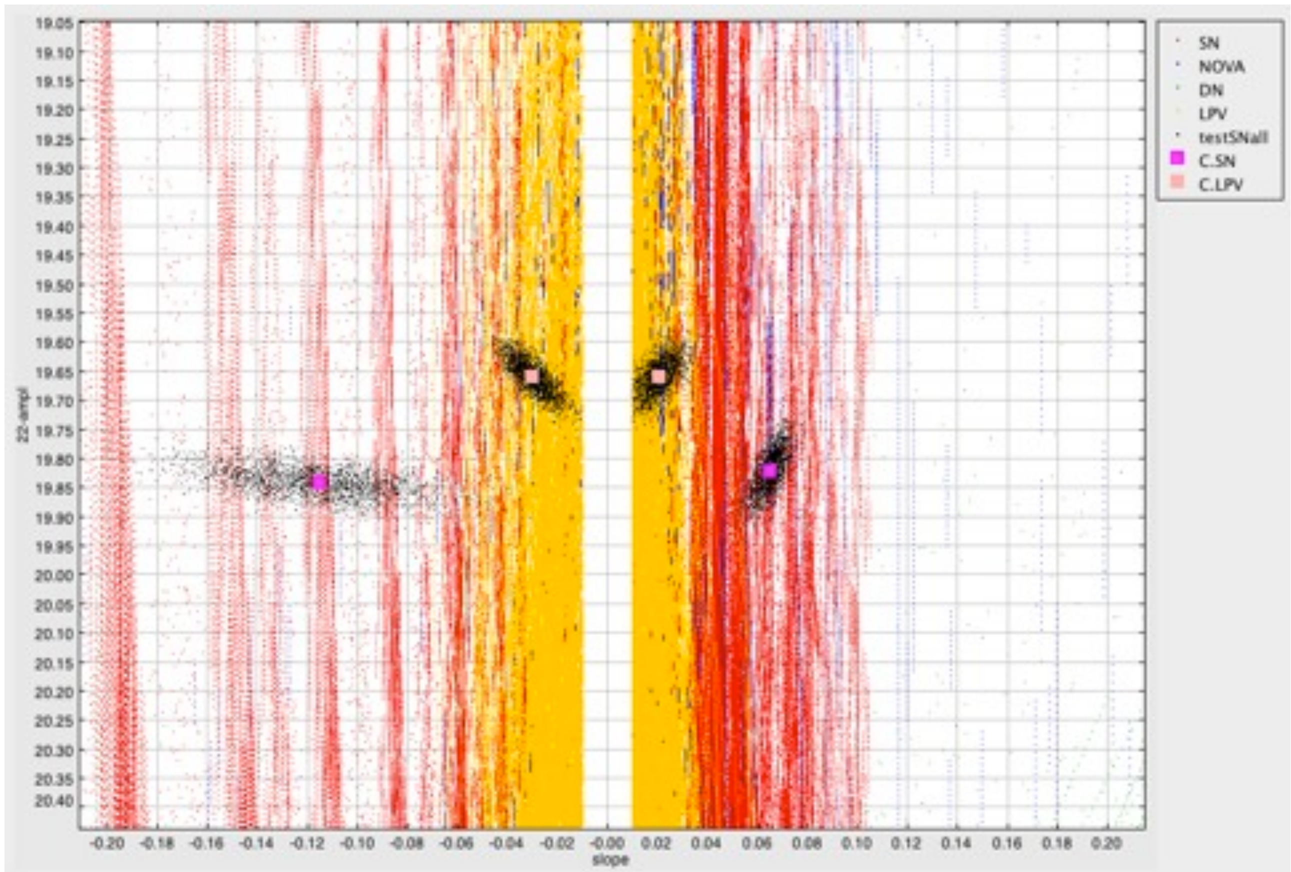


Method: a test object

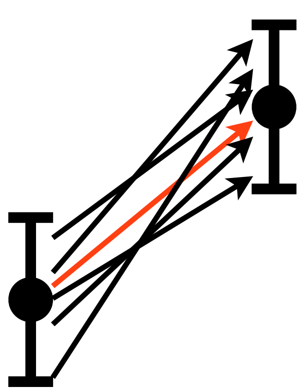
known supernova from Stripe82

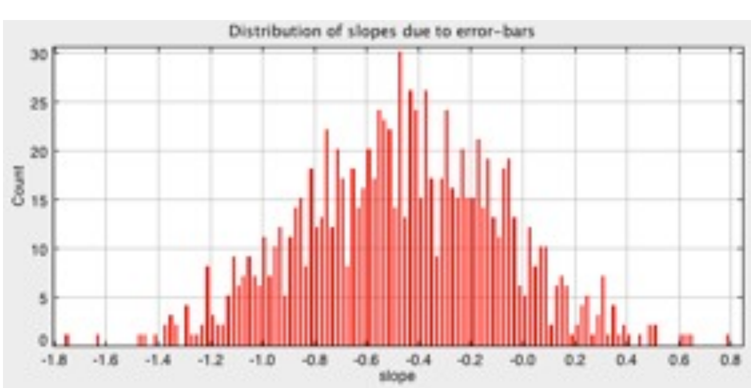


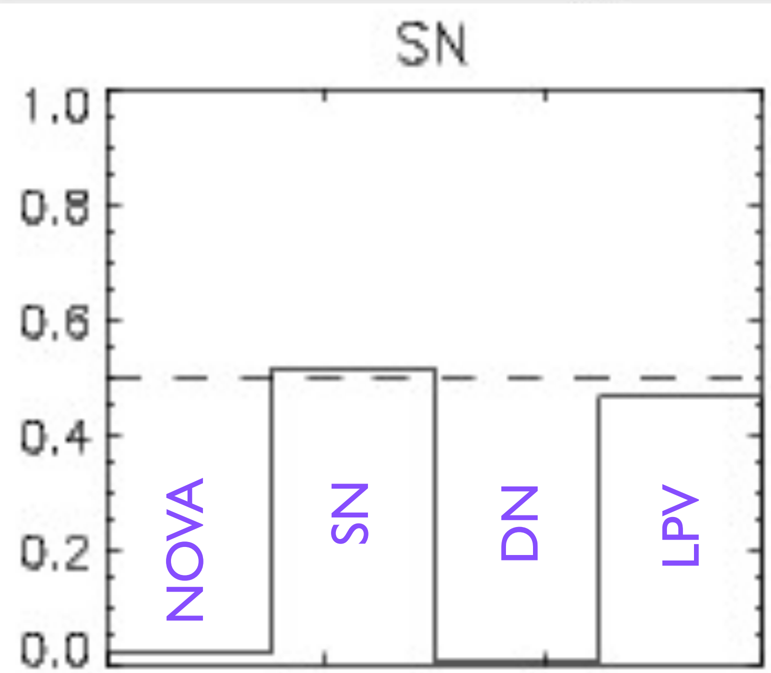
Training data set and test object



Each pair has a distribution of parameters due to error-bars







For more than one pair the final resulting PDF is sum of PDFs

AlertPipe Features

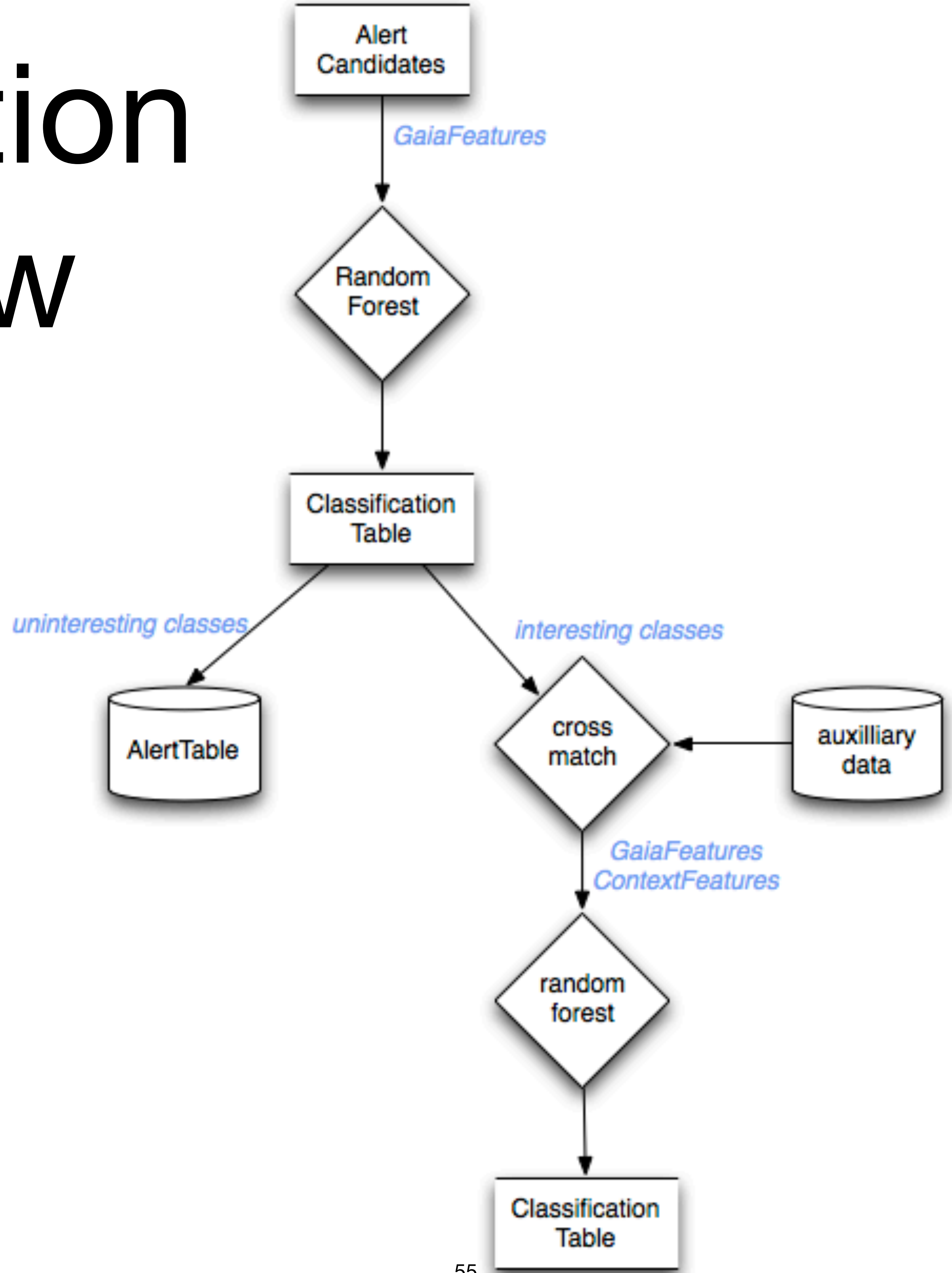
Lightcurve
gradient
amplitude
historic rms
magnitude
SNR
transit rms

Auxillary
neighbour star
shape pars
motion pars
coords
crowding
calibration offset
correlations
QC pars

Spectrum
flux v lambda
colours
SSCs
SpTy

Xid Environment
near known star mags
near known star cols
near known variable class
near galaxy
near galaxy Z
circumnuclear

Curation Flow

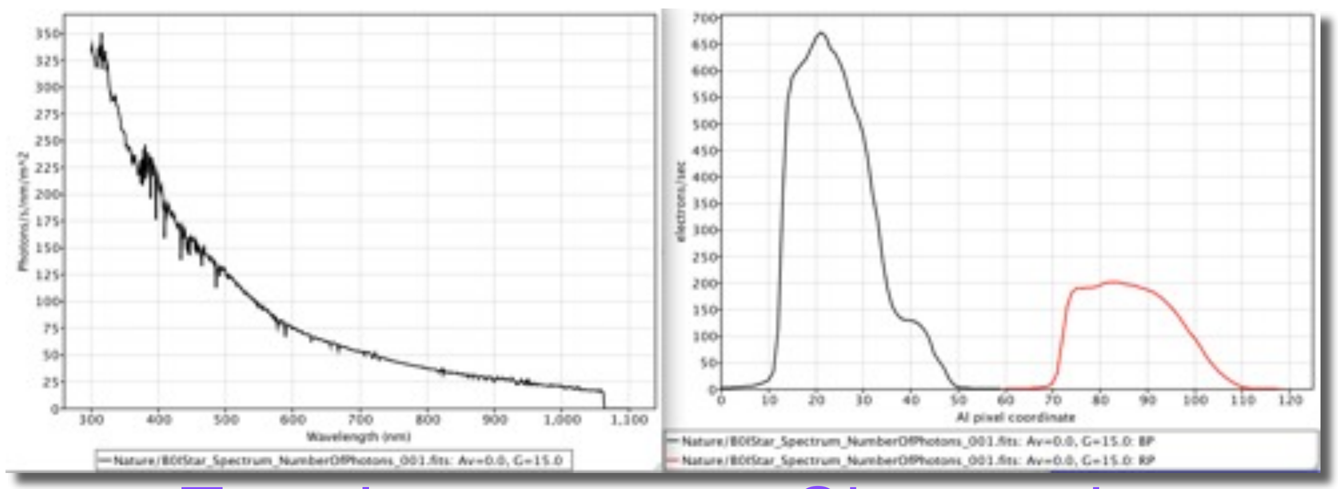


Outline

- Transient astronomy
- The Gaia mission
- Detecting transients with Gaia
- Supernovae and microlensing
- Details: Classification from photometry
- Details: Classification from spectroscopy
- Verification and follow-up

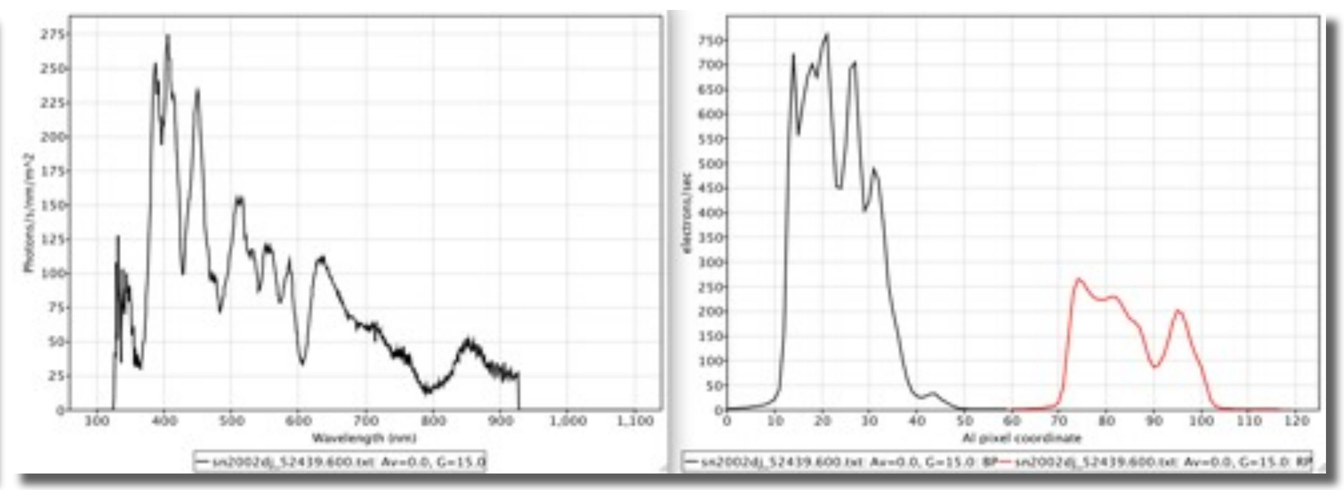
Spectral classification

star: supergiant B0

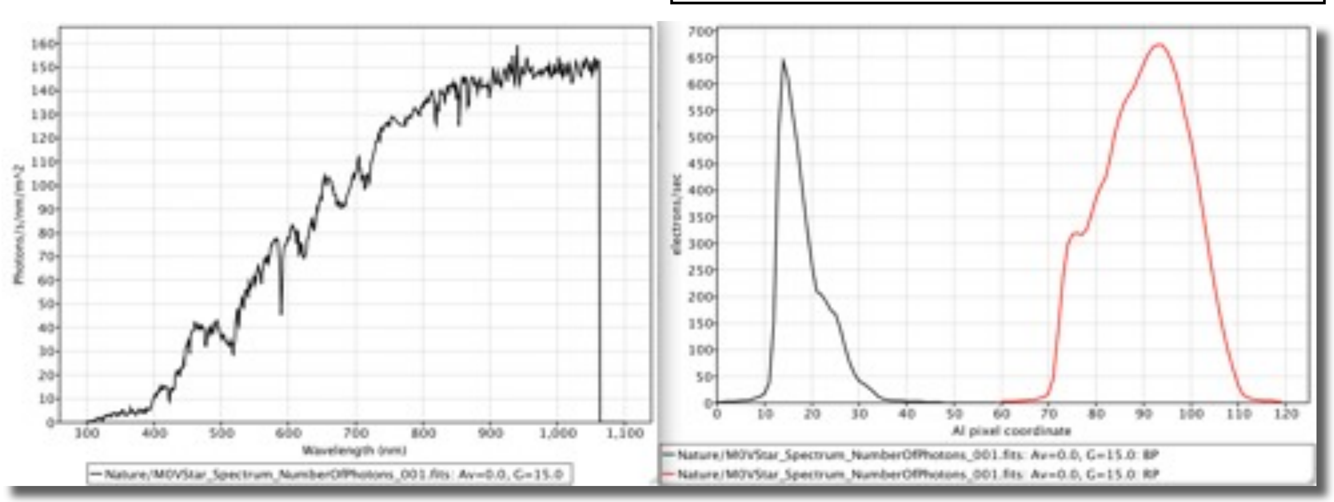


Template Observed

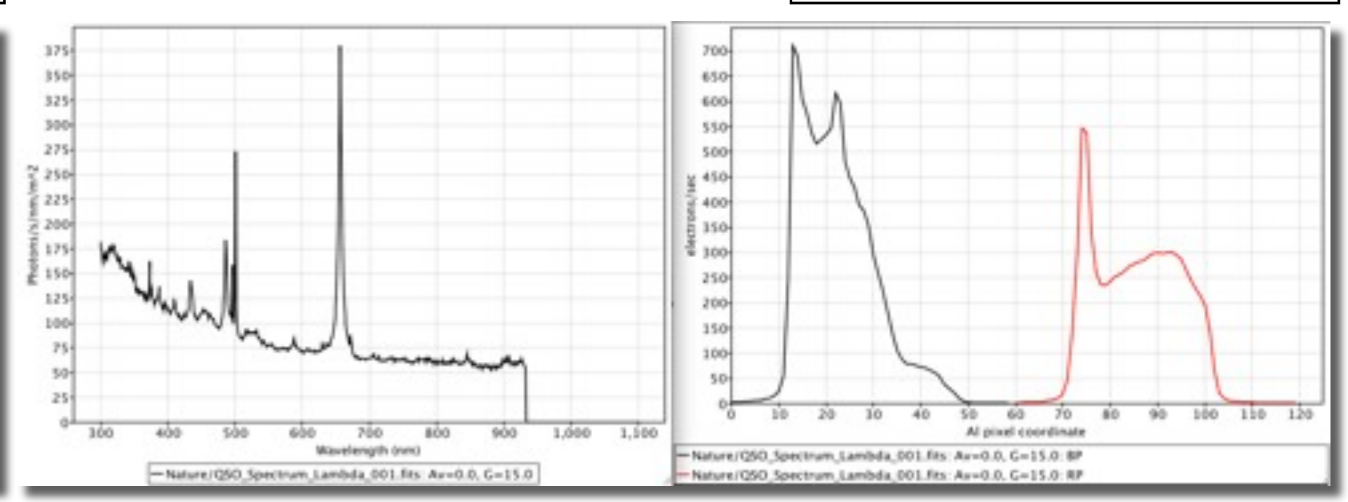
supernova 2002dj at -11d



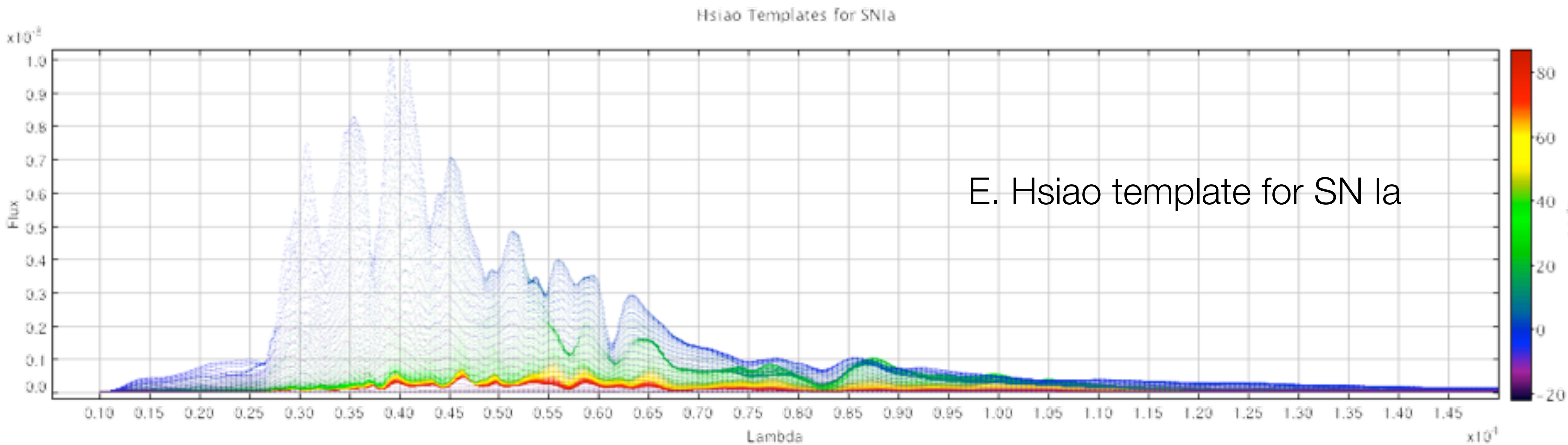
star: main sequence M0

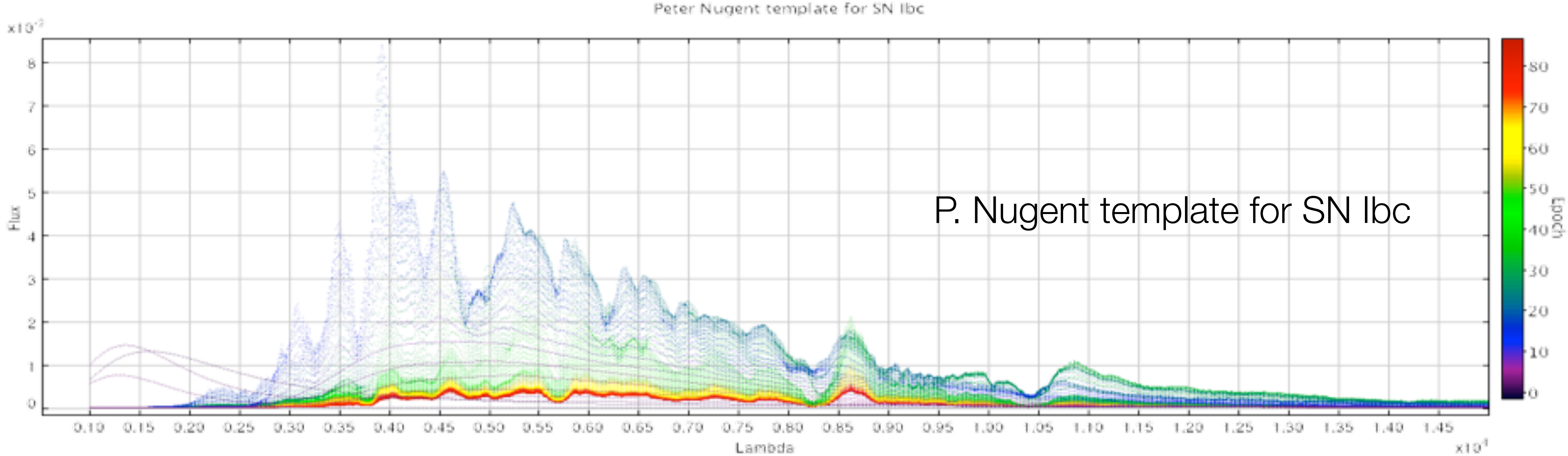


zero-redshift quasar



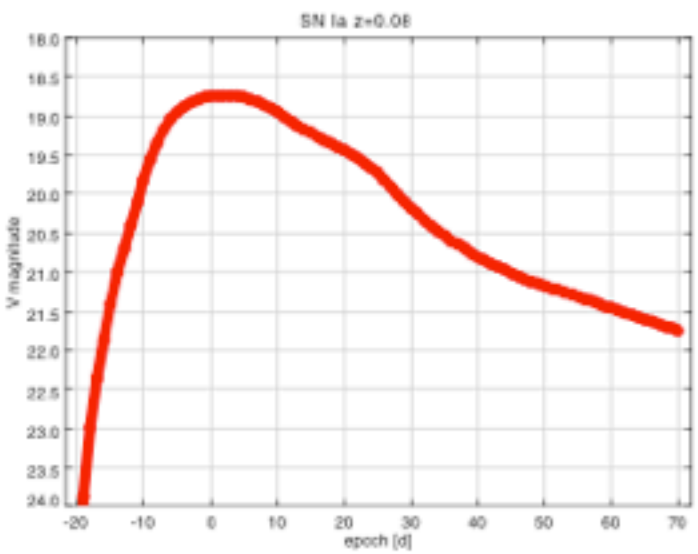
SuperNova library creation





BP/RP SN Spectra: Parameter Estimation

Supernova Ia
z=0.08



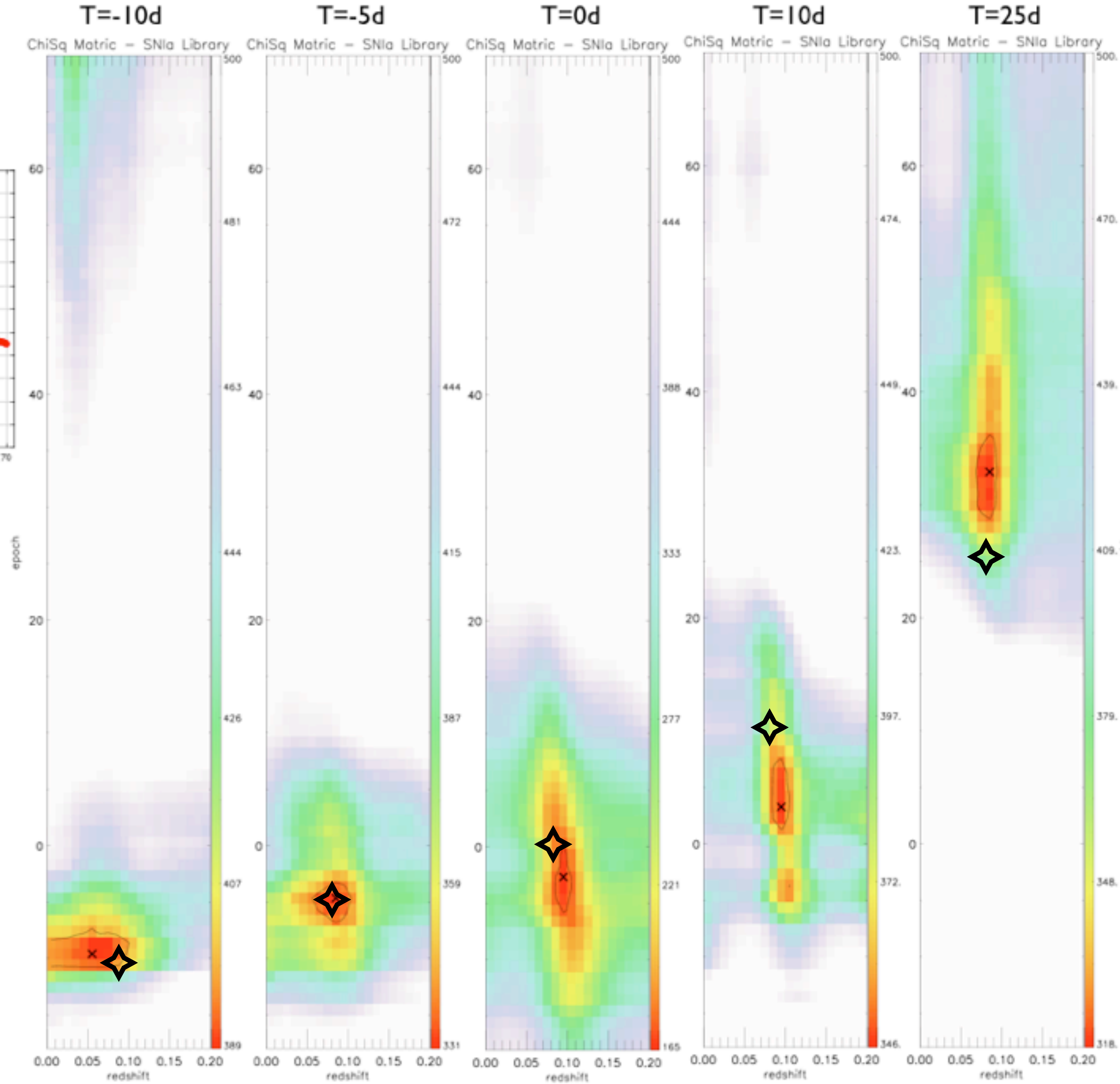
SNIa at z= 0.08 at peak will have mag= 18.7
Result for mag 19.8
found z= 0.05 -0.01 +0.03 (real= 0.08)
found epoch= -10.0 -1.0 +2.0 (real= -10.0)
Chi_min= 388.93

SNIa at z= 0.08 at peak will have mag= 18.7
Result for mag 19.0
found z= 0.08 -0.02 +0.01 (real= 0.08)
found epoch= -5.0 -1.0 +1.0 (real= -5.0)
Chi_min= 330.55

SNIa at z= 0.08 at peak will have mag= 18.7
Result for mag 18.7
found z= 0.09 -0.01 +0.01 (real= 0.08)
found epoch= -3.0 -2.0 +1.0 (real= 0.0)
Chi_min= 165.06

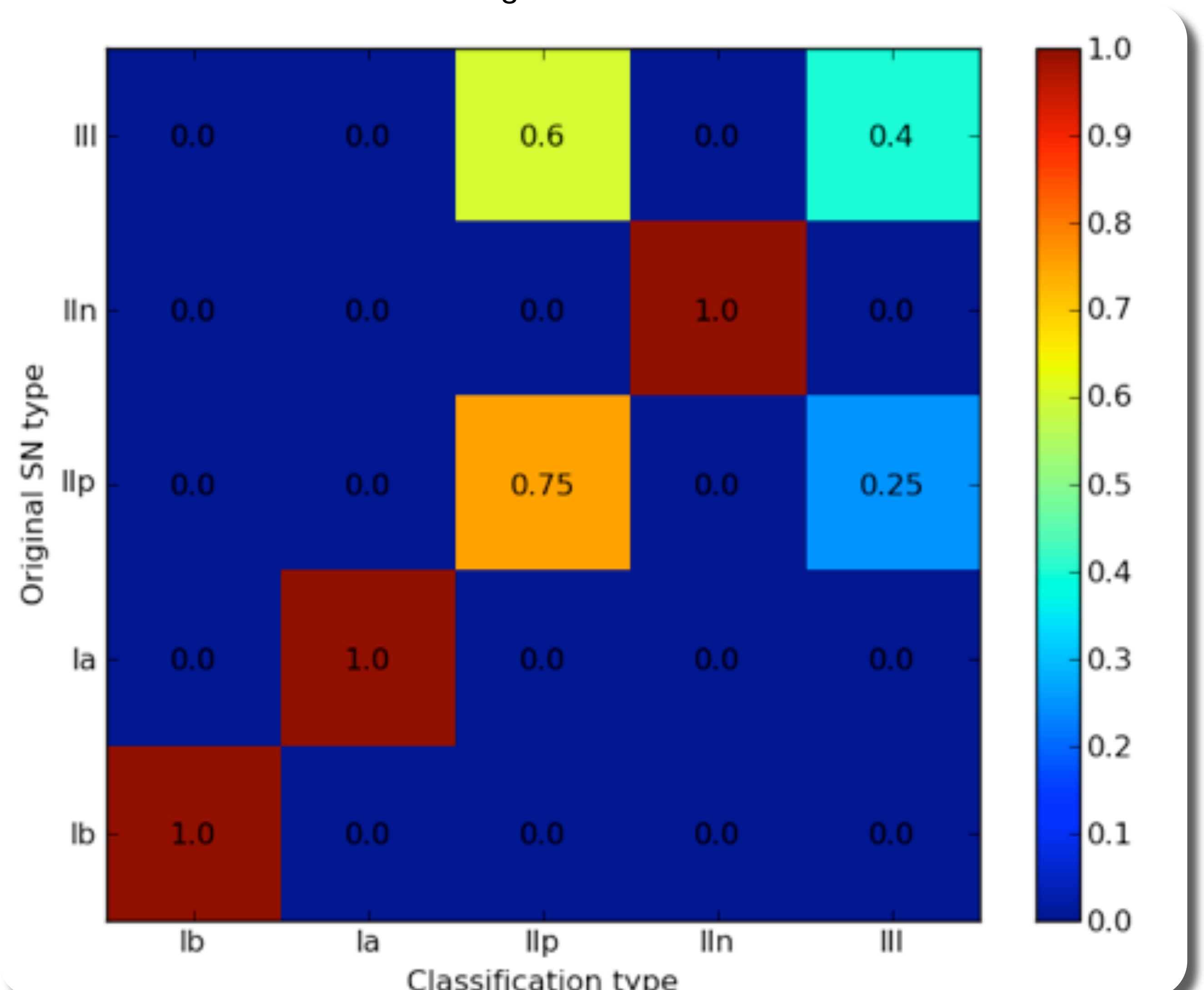
SNIa at z= 0.08 at peak will have mag= 18.7
Result for mag 19.0
found z= 0.09 -0.01 +0.01 (real= 0.08)
found epoch= 3.0 -2.0 +3.0 (real= 10.0)
Chi_min= 345.89

SNIa at z= 0.08 at peak will have mag= 18.7
Result for mag 19.7
found z= 0.08 -0.02 +0.01 (real= 0.08)
found epoch= 33.0 -1.0 +2.0 (real= 25.0)
Chi_min= 317.76

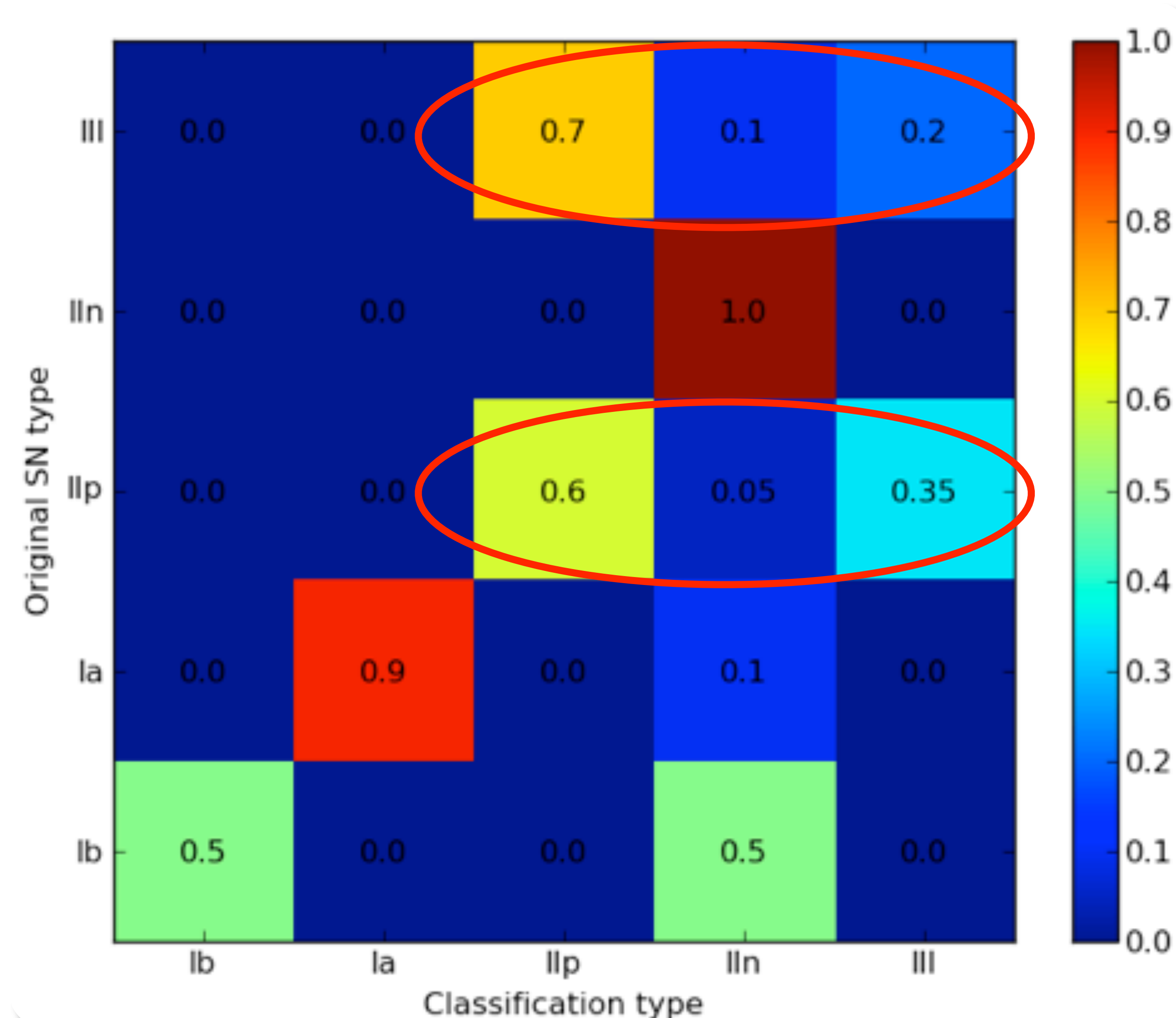


BP/RP SN Spectra: Classification of Type

m_g = 17 mag



m_g = 18 mag



- Range of model templates (Nugent, Hsiao)
- Perturb spectra (magnitude, redshift), add noise.
- Classify

Nadejda Blagorodnova, PhD @ IoA

Confusion among SNIIp and IIL.
→Very similar spectra.
→Main differences in lightcurve.

Outline

- Transient astronomy
- The Gaia mission
- Detecting transients with Gaia
- Supernovae and microlensing
- Details: Classification from photometry
- Details: Classification from spectroscopy
- Verification and follow-up

Purposes of Verification

- Demonstrate transient detection works
- Demonstrate transient classification works
- Test thresholds
- Validate associated classification probabilities
- Investigate Gaia Science Alert population : completeness and contamination
- **To build a training data set**

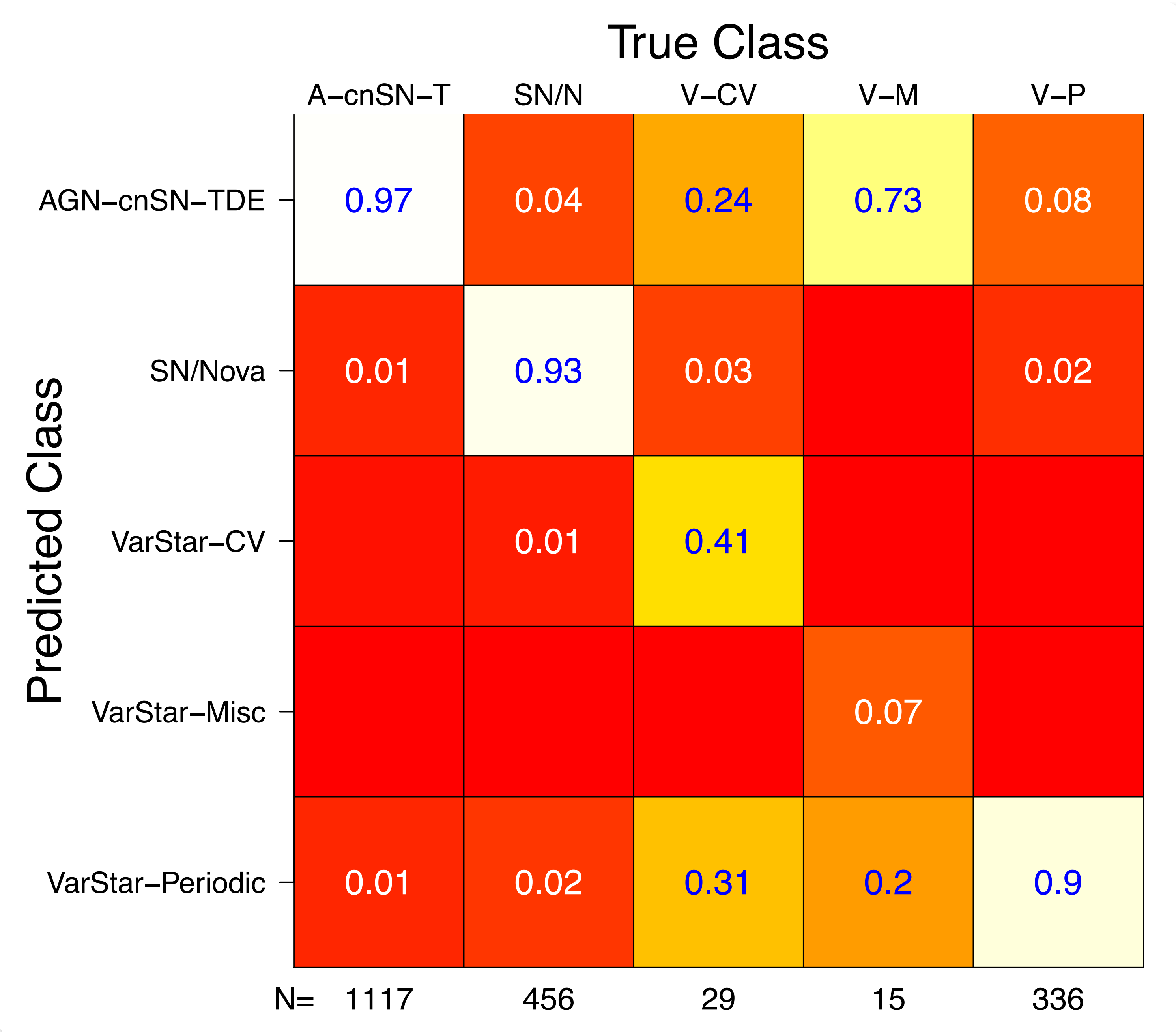
How do we verify?

- Using the Gaia BP/RP spectroscopy
- Cross-comparison with other transient/ variability studies : CRTS, PTF, ASAS, Skymapper, i.e. large sample of known variables (cheap).
- But not great for Gaia transients.. by definition. Therefore we will need our own follow-up programmes.

Size of Verification Plan?

- Depends on the goals
- For a training set: 100s per broad class seems to be a reasonable estimate
- This is of order 500 ‘follow-ups’

http://arxiv.org/abs/1106.5491v1



Automating Discovery and Classification of Transients and Variable Stars in the Synoptic Survey Era
J. S. Bloom[1], J. W. Richards[1,2], P. E. Nugent[3,1], R. M. Quimby[4], M. M. Kasliwal[4], D. L. Starr[1], D. Poznanski[1,3], E. O. Ofek[4], S. B. Cenko[1], N. R. Butler[1], S. R. Kulkarni[4], A. Gal-Yam[5], N. Law[6]

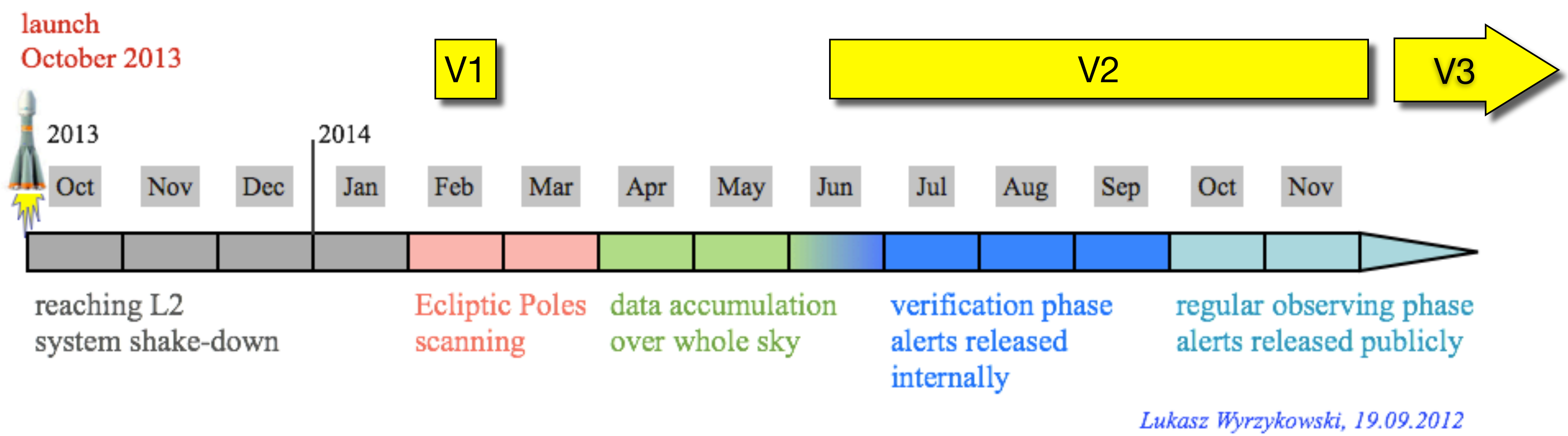
Gaia Operations: Year 1

- Ecliptic Pole Scanning Law at start of operations
- Nominal Scanning Law data accumulation starts early 2014
- I envisage two main verification efforts

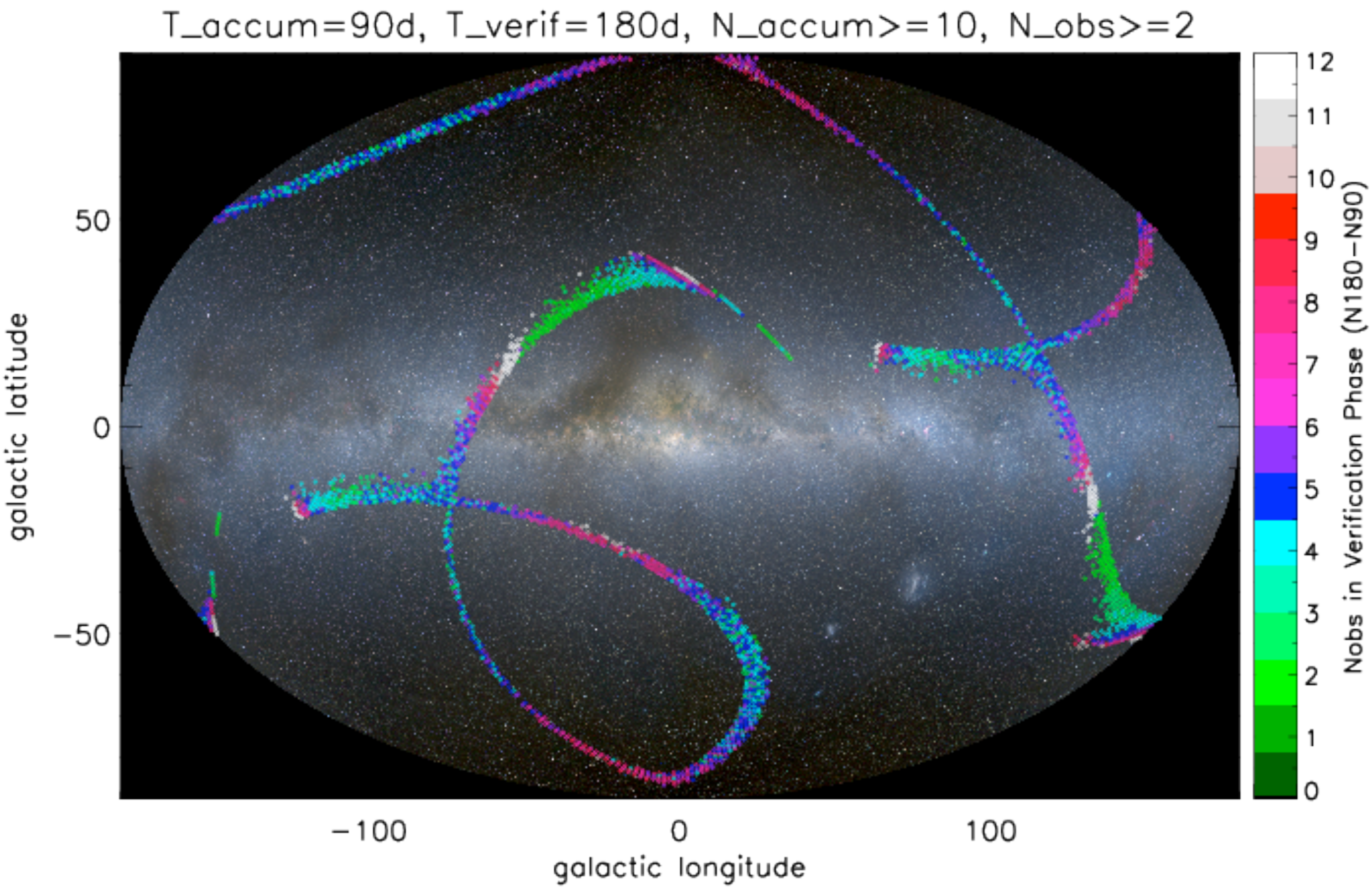
  V1 : towards the end of EPSL mode

  V2 : commences as soon as sufficent sky has been observed enough times to define the baseline catalogue

  V3: extended verification for rare classes and monitoring



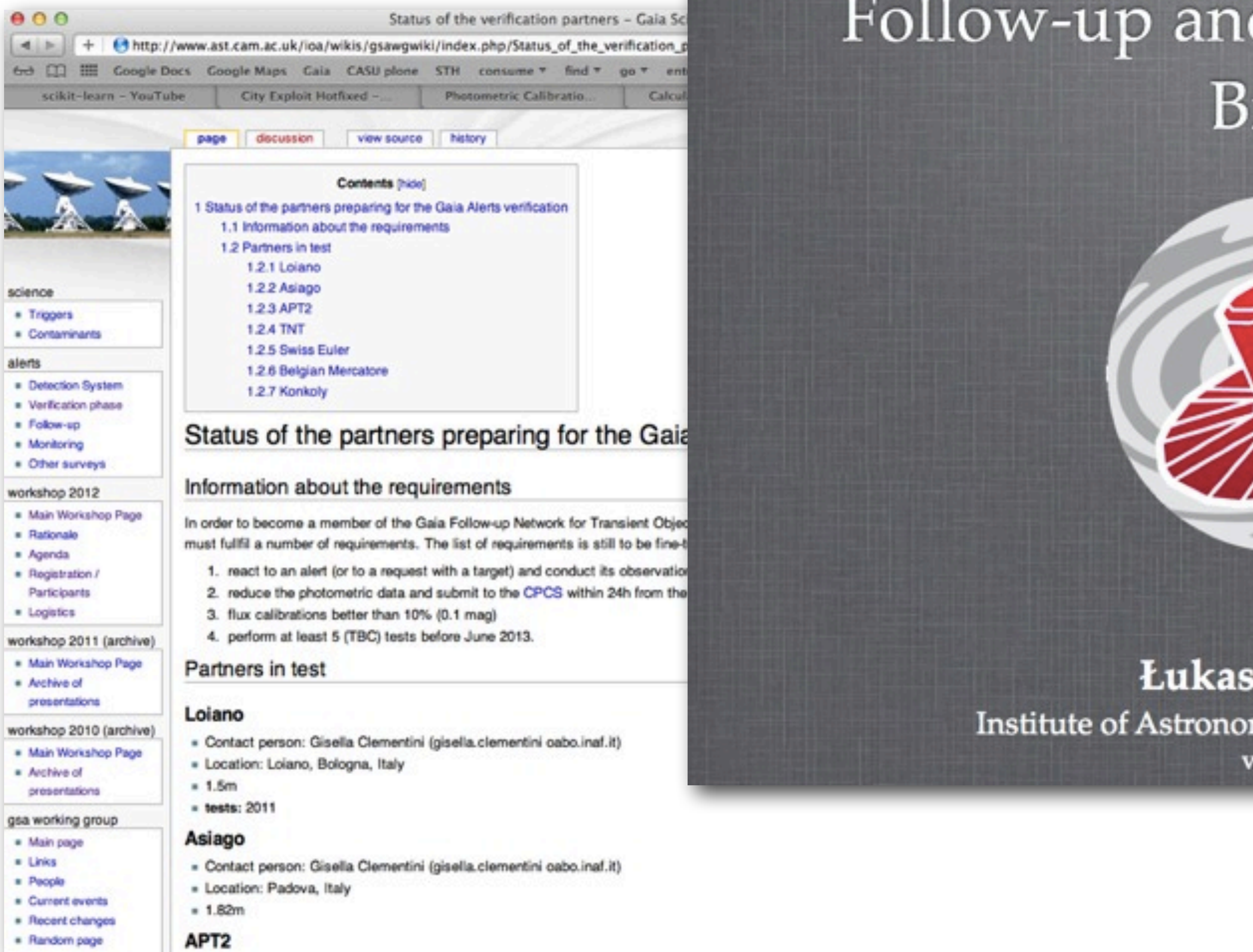
Fraction of the sky to be observed during the Verification



Requiring at least **10** points collected by day **90**

and at least **2** points before day **180**

GAIA-FUN-TO



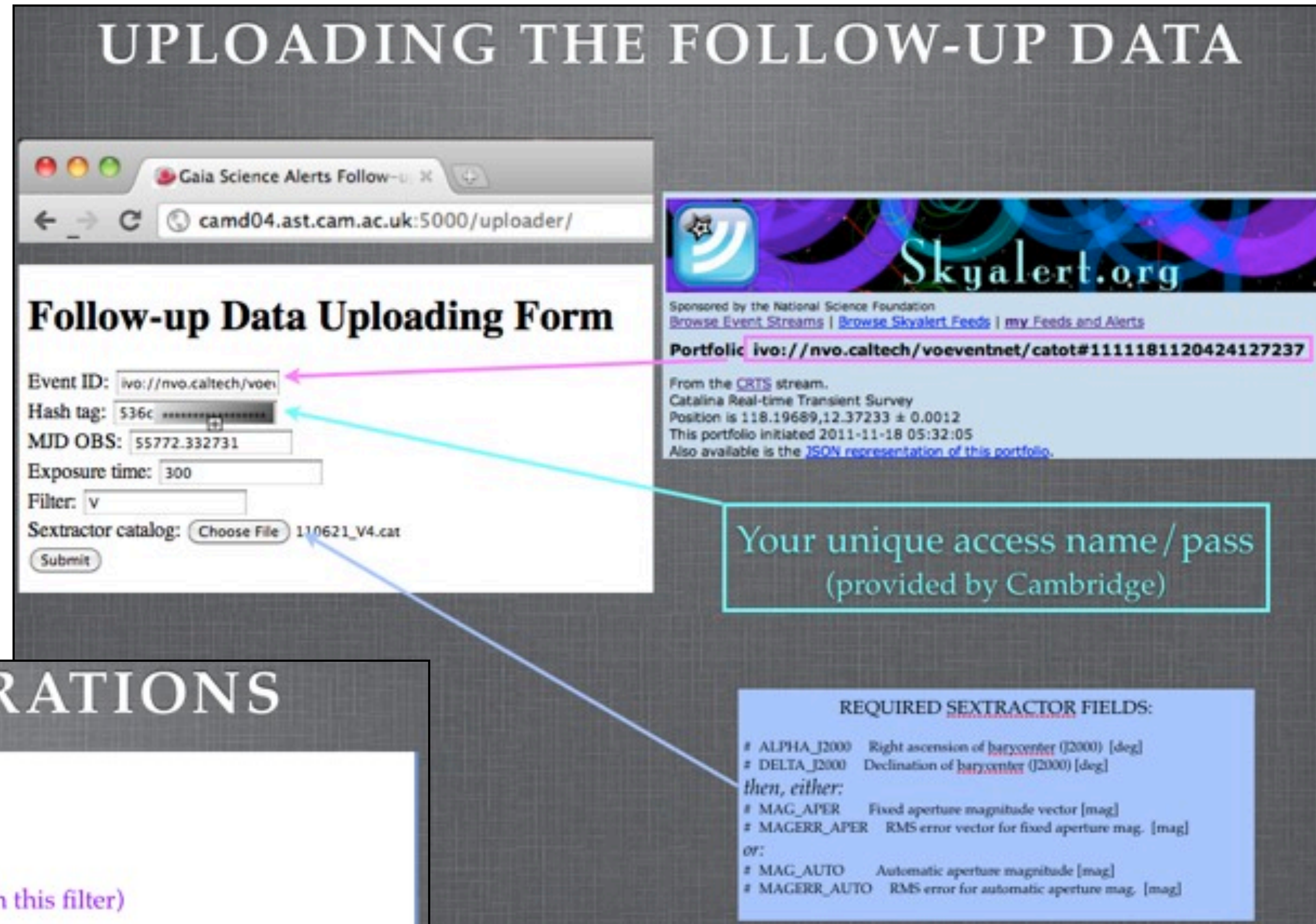
http://www.ast.cam.ac.uk/ioa/wikis/gsawgwiki

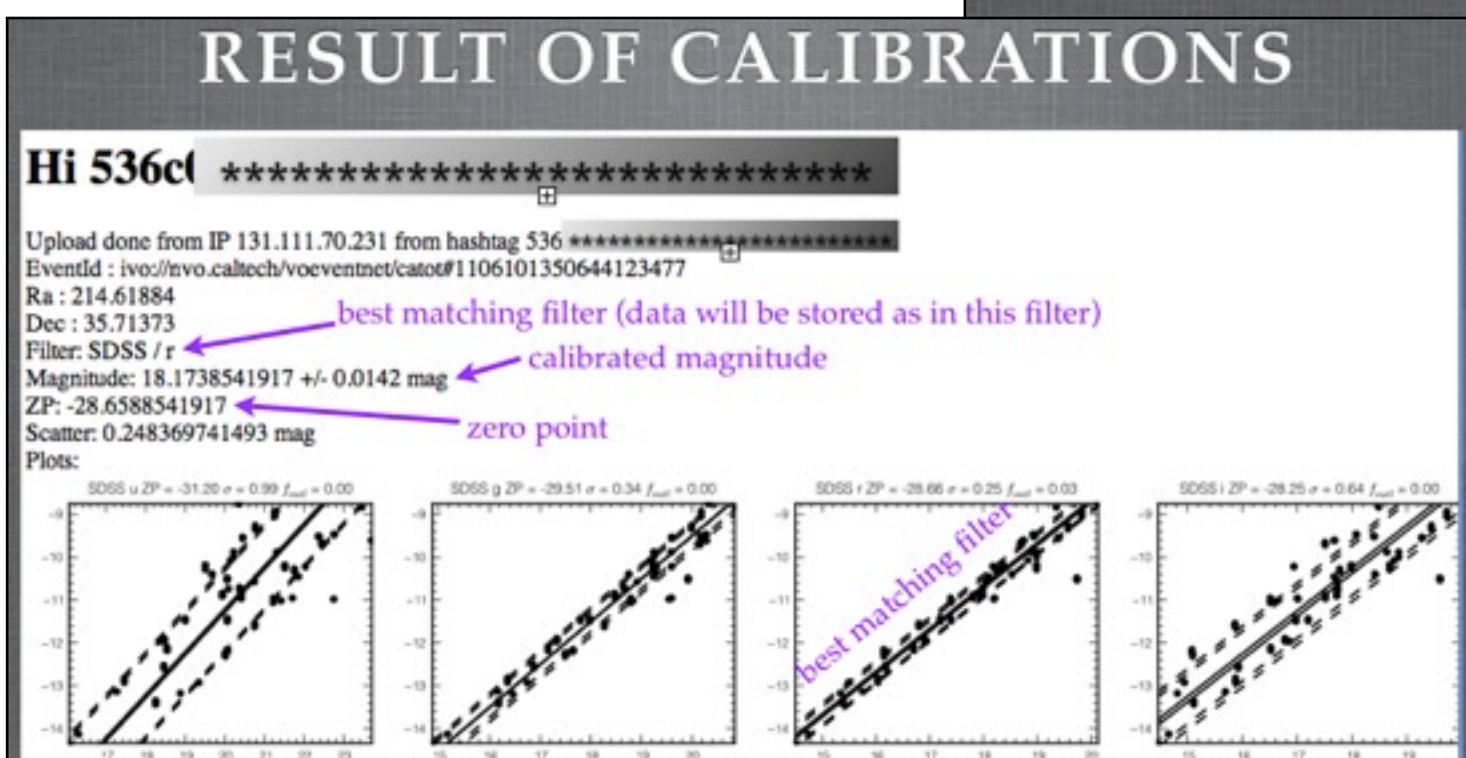
Additional Photometry

- Tells us that a source is real !
- Can recover morphological/ environmental information (limited SM/AF info for faint targets).
- Monitor the lightcurve with better coverage - important for classification.

Follow-up calibration server

for CRTS follow-up phase





Server has been tested by Italian and American colleagues

software developed by Sergey Koposov, IoA

How much additional spectroscopy?

- If we deem the Gaia spectroscopy insufficient.
- If we assume that a single ground-based spectrum is sufficient for broad classification.
- If we cannot rely on contemporaneous overlapping experiments
- 500 objects will each require ~30 minutes (generous).
- We would need ~250 hours or ~1 month total

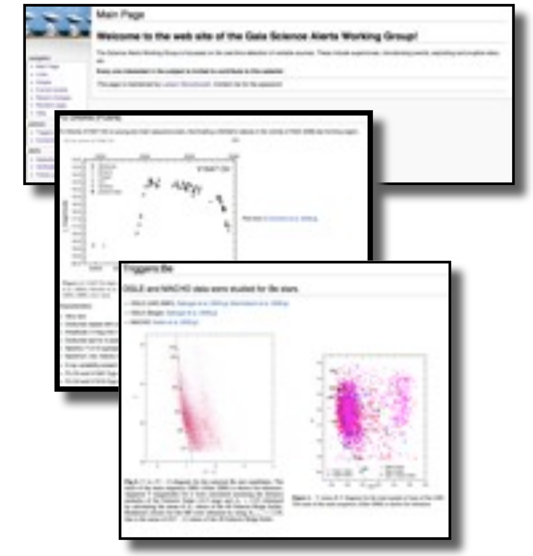
Verification issues

- Homogeneity of the spectra (wavelength range, dispersion, SNR)
- Suggests (to me) dedicated programmes on relatively few instruments.
- Less of a problem for the photometry
- Data management
- People power

What makes Gaia transients precious?

- All sky
- Huge dynamic range: 5-20 mag
- Spectro-photometry for each object means classification on the go
- Astrometry: can mean a lot for local SNe
- High temporal resolution

Summary



- Gaia: A rich source of transient phenomena from late-2014
- The alert stream is non-proprietary and the will form the first data from Gaia
- Significant verification program in mid-2014
- Watch List: call for targets shortly after launch
- Huge scope for multi-messenger science
- The science alerts software framework is in place (detection, curation, reporting)
- Plenty of scope for involvement, especially with verification