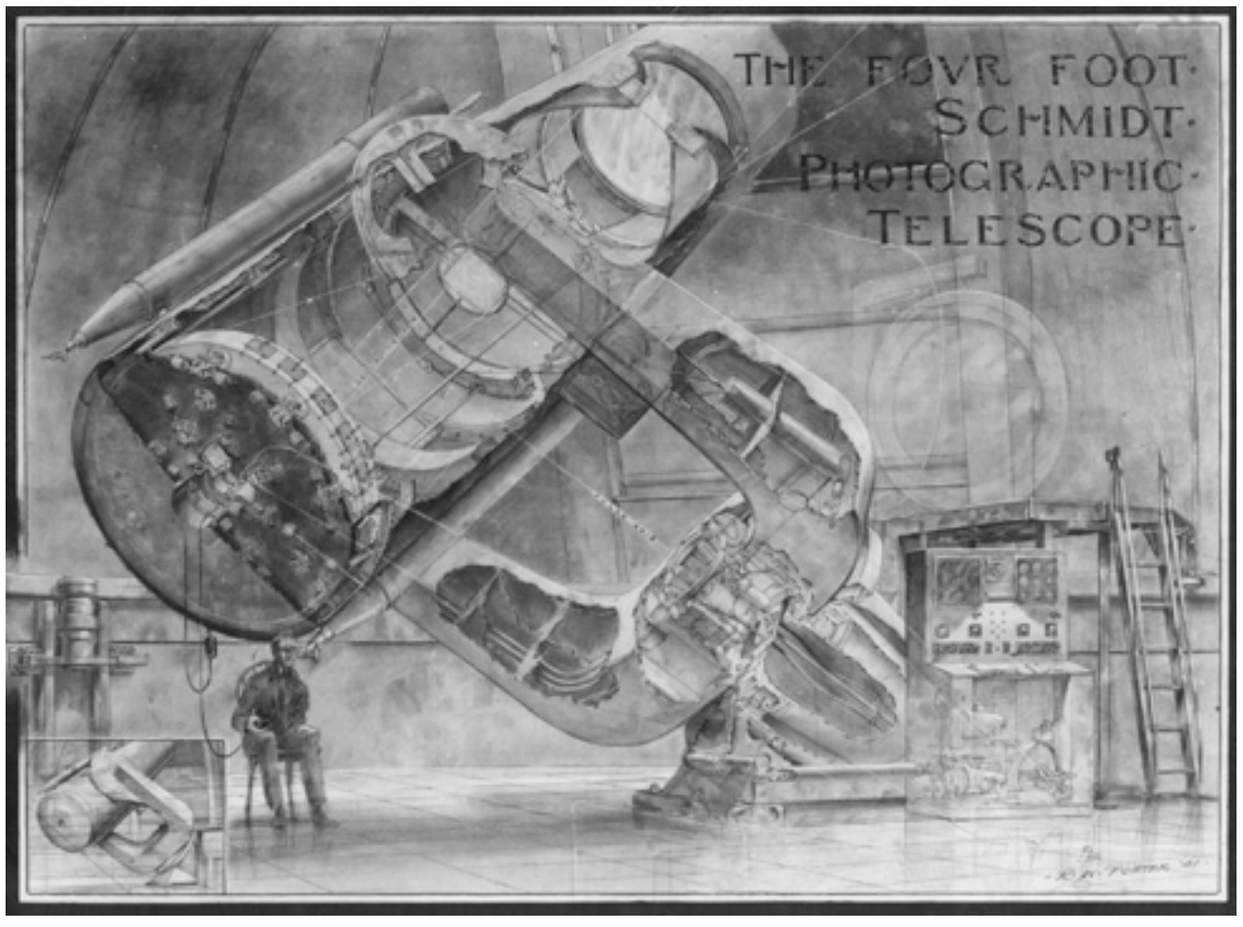
Data Mining the Optically Variable Sky since 1950



astro.caltech.edu

Greg Madsen (IoA, Cambridge)
B. Gaensler (Univ. of Sydney)

Motivation

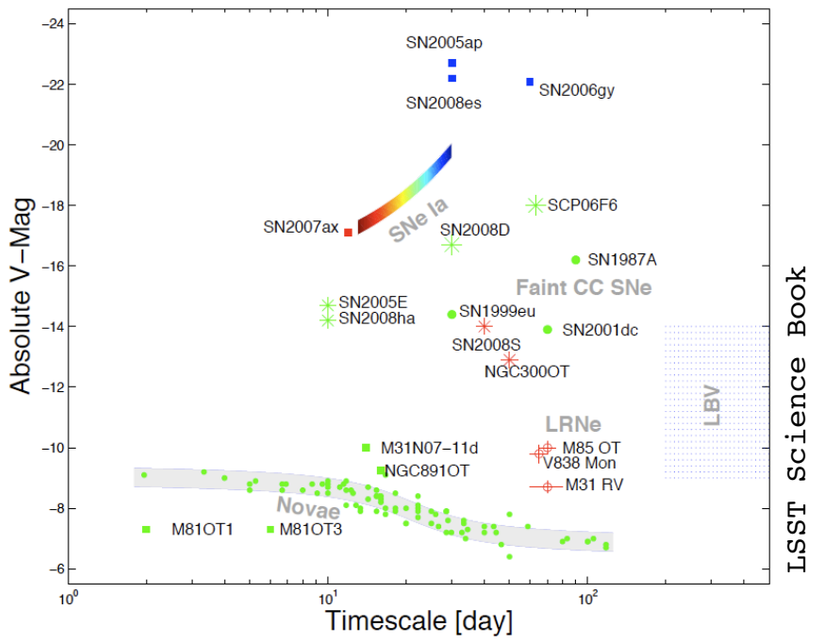
- Current and next generation transient surveys limited to a few year time-scales

- Take advantage of long temporal & huge sky coverage of historical optical survey data

astro.caltech.edu

48" Schmidt at Palomar

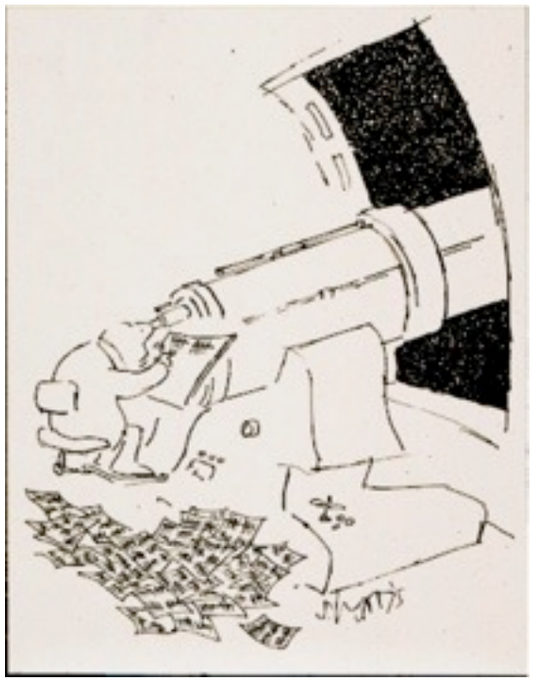
- Statistical analysis of large classes of variables
  - QSOs: structure of central engine
  - RR Lyr: structure of Galactic halo
  - Period changes in Miras, eclipsing binaries
- Discovery & understanding of outbursts from rare objects
  - dwarf novae, FU Ori, R Cor Bor, LBVs
- Explore new parameters in luminosity-duration space
- Improve historical photometry; legacy value



Approach: Use Existing Catalogues

- Photographic plate surveys go back ~100 years; digitised to create modern catalogues
  (e.g. USNO-B, GSC, SuperCosmos, DASCH)
- Plates very good for astrometry, OK for photometry
- USNO-B:
  three-band, five-epoch photometry of a billion objects

- Sloan Digital Sky Survey (SDSS) Data Release 9:
  accurate, five-band photometry of 260M *point sources*
- Compare SDSS & USNO-B catalogs to conduct blind search for variability
  - sensitivity (≈ 20th mag)
  - areal coverage (≈ 30% of sky),
  - temporal coverage (~ 60 years)



Monet et al 2003; Ahn et al 2012

Method

- determine common photometric system
- select well-measured point sources
- positional cross-match
  (be careful with proper motions)
- characterise blending effects
- identify artefacts
- compare magnitudes
  → recalibration of USNO-B required

$$O_{SDSS} = g + 0.452(g - r) + 0.08$$

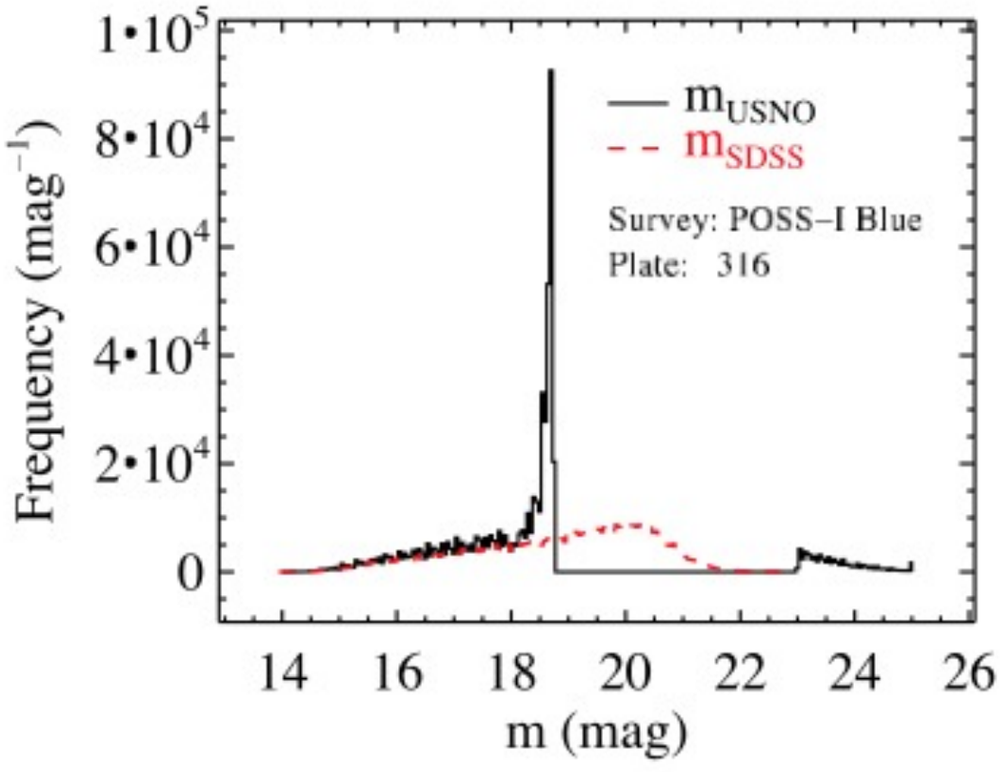
$$E_{SDSS} = r - 0.086(g - r) - 0.20$$

$$J_{SDSS} = g + 0.079(g - r) + 0.06$$

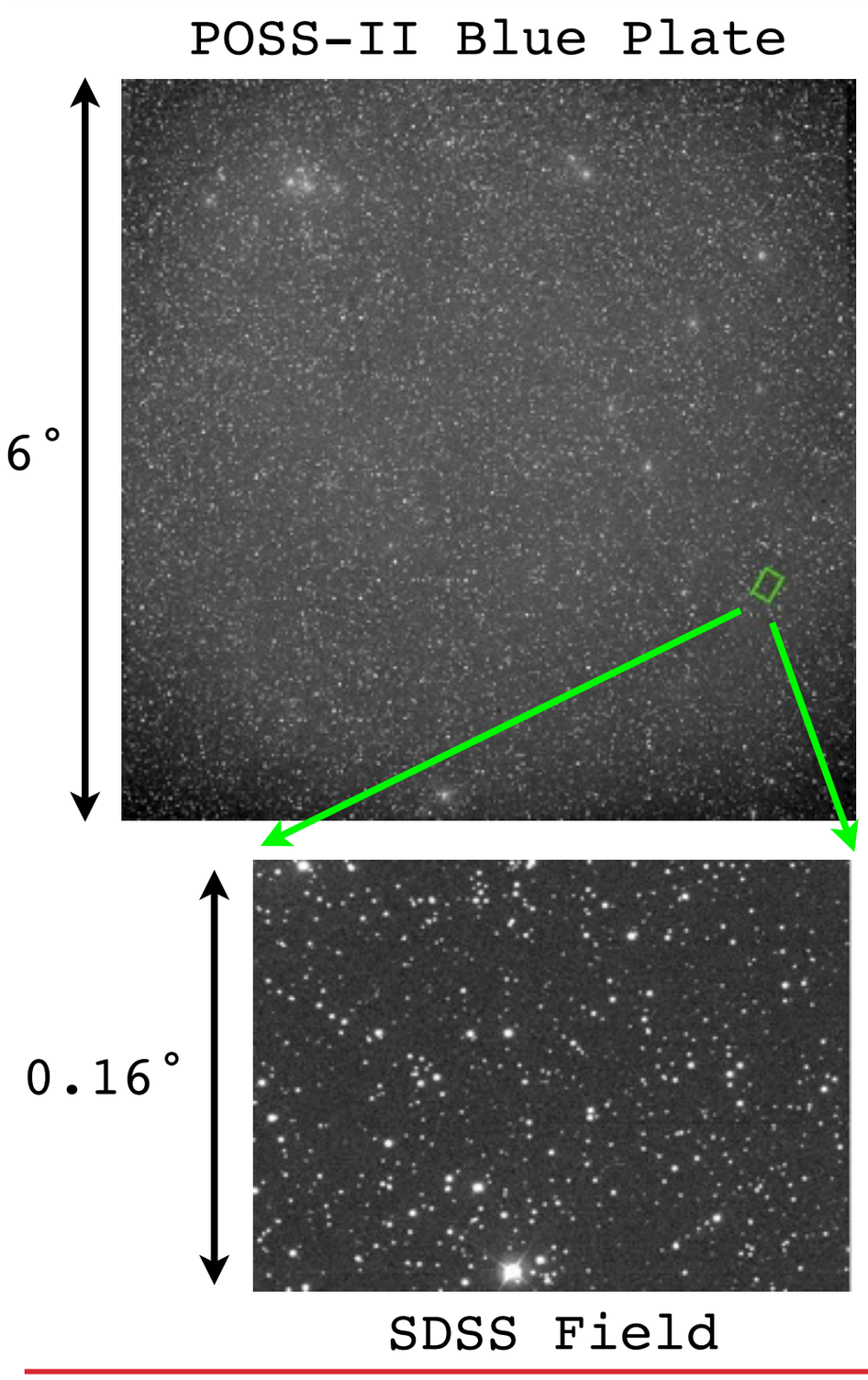
$$F_{SDSS} = r - 0.109(g - r) - 0.09$$

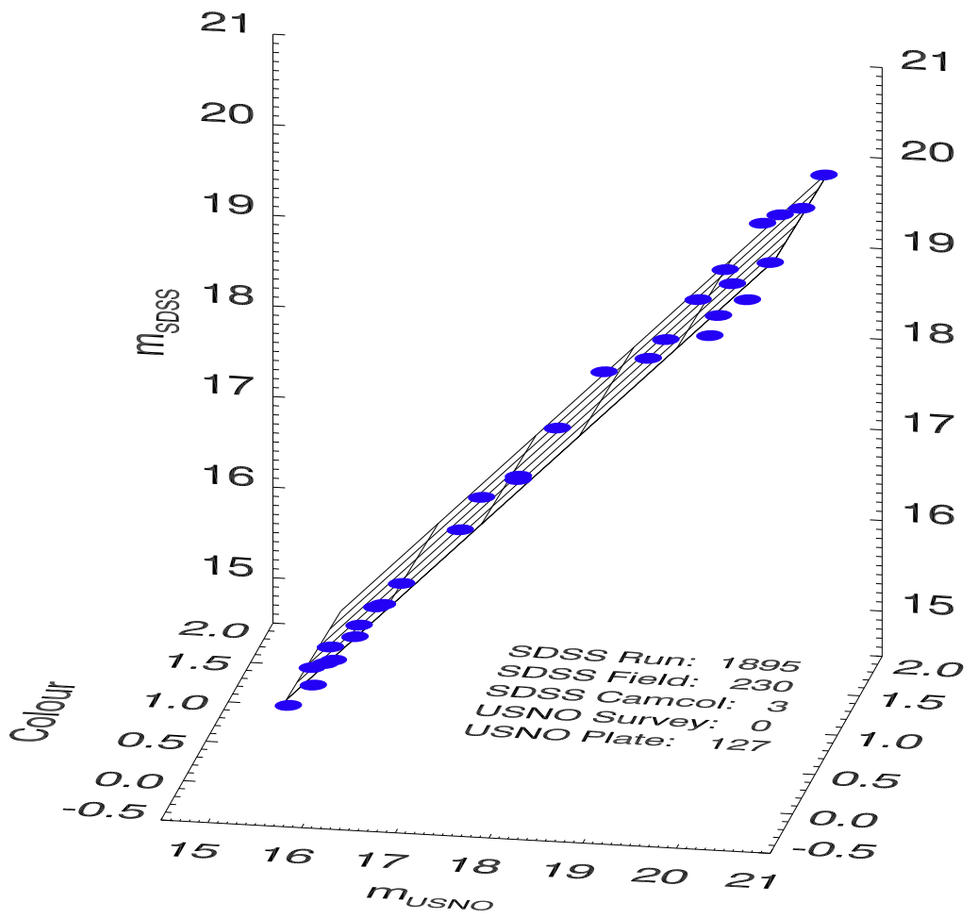
$$N_{SDSS} = i - 0.164(r - i) - 0.44,$$

SDSS→USNO system



Recalibration: Stage 1



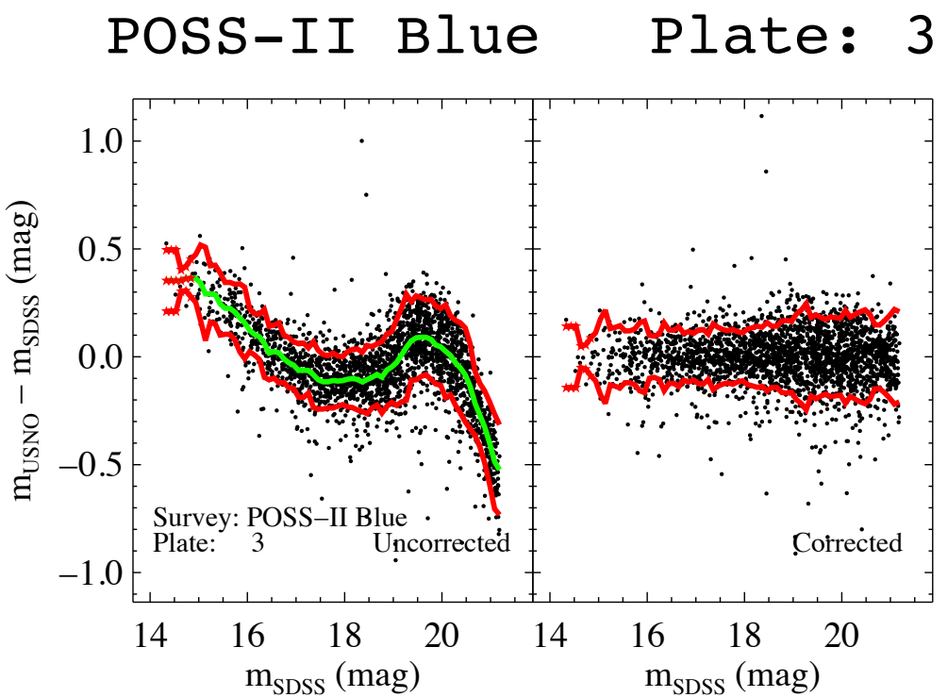


- Identify calibration objects
- Calculate coefficients:

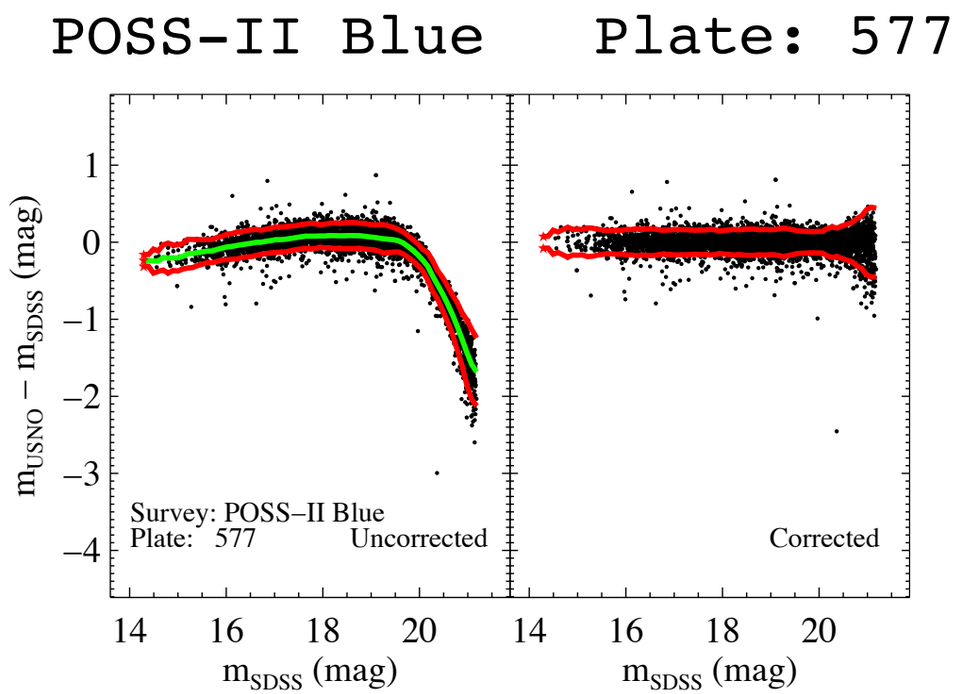
  $A * m_{USNO} + B * (colour) + C$

  that minimise $m_{SDSS} - m_{USNO}$
- Adjust USNO magnitudes

Recalibration: Stage 2

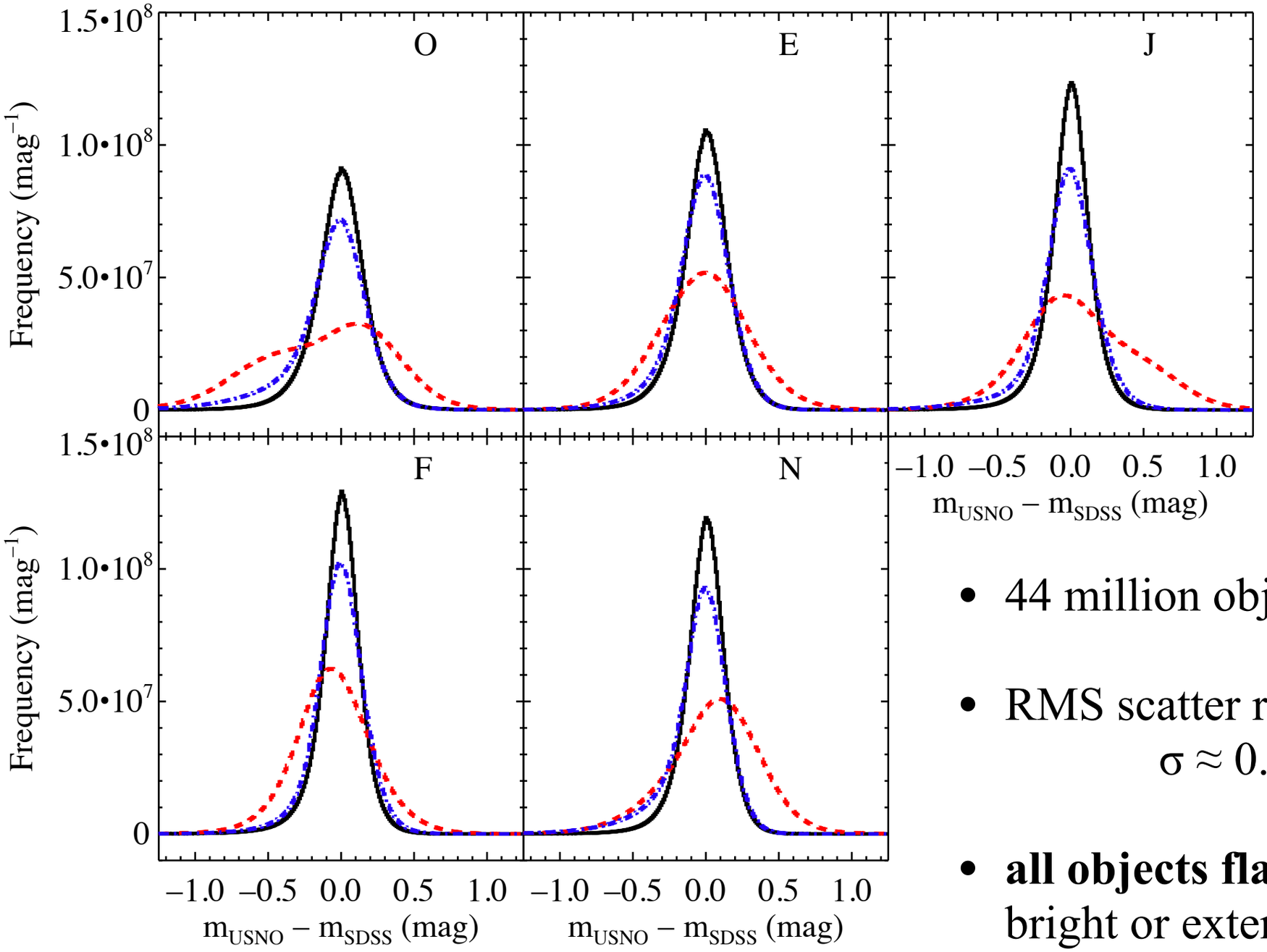


- Plate-wide (large angular scale) calibration problems remain



- Common pattern in residual magnitude vs. magnitude
- USNO magnitude adjusted to remove pattern
- Residual scatter used as proxy for photometric accuracy

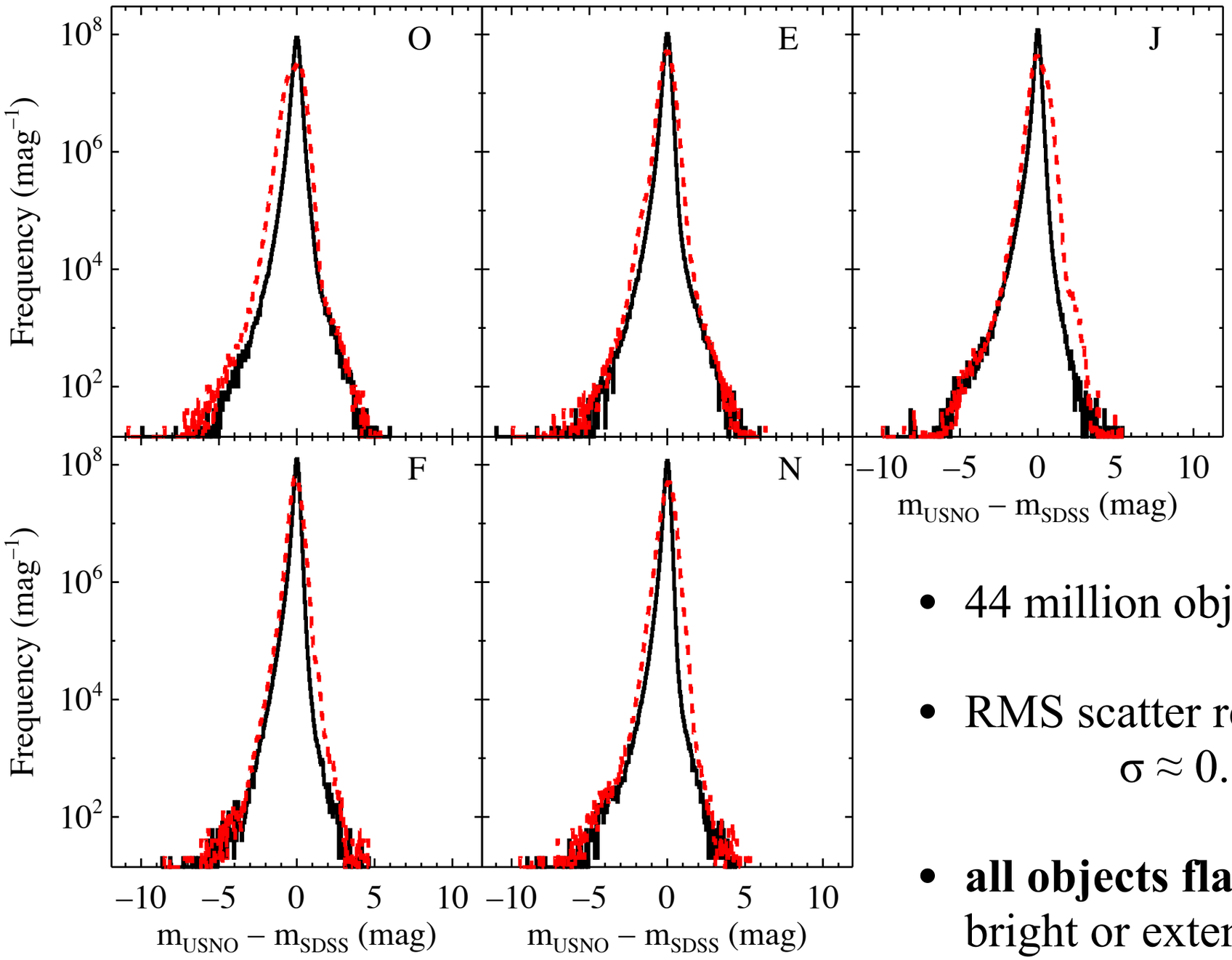
Recalibration: Results



Original
Stage 1
Stage 2

- 44 million objects recalibrated
- RMS scatter reduced by factor of ≈ 4
  $\sigma \approx 0.1$ mag
- **all objects flagged** for proximity to bright or extended objects, blended sources, consistent magnitudes

Recalibration: Results



Original
Stage 1
Stage 2

- 44 million objects recalibrated
- RMS scatter reduced by factor of ≈ 4
  $\sigma \approx 0.1$ mag
- **all objects flagged** for proximity to bright or extended objects, blended sources, consistent magnitudes

Lessons & Limitations

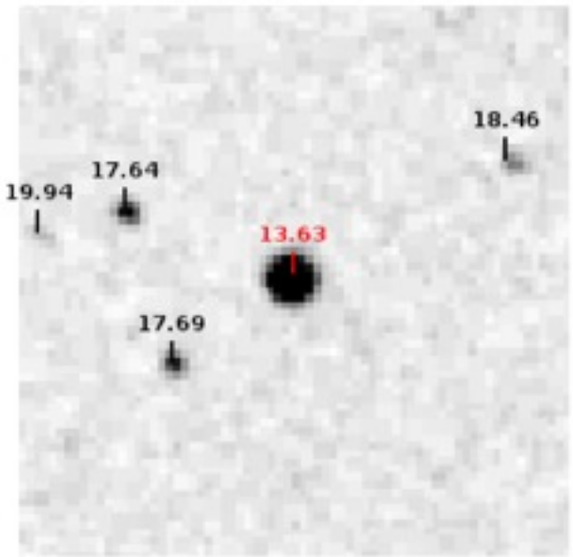
- Careful, self-consistent treatment of proper motions and cross-matching

- Robust identification of objects blended with stars, galaxies, & artefacts

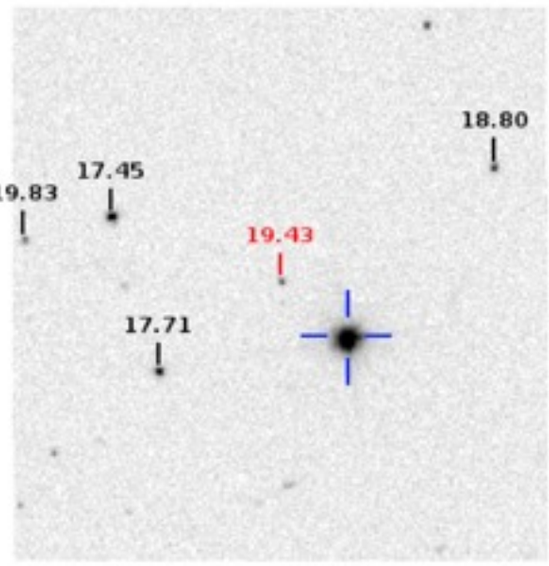
- Check consistency of USNO, GSC, and SuperCosmos magnitudes of the *same* objects

- Visual inspection of large samples used to assess catalog reliability

- Remaining issues
  - blends with unusual artefacts
  - ‘pathological’ proper motions
  - inaccurate Sloan magnitudes (!)
  - object is not point source

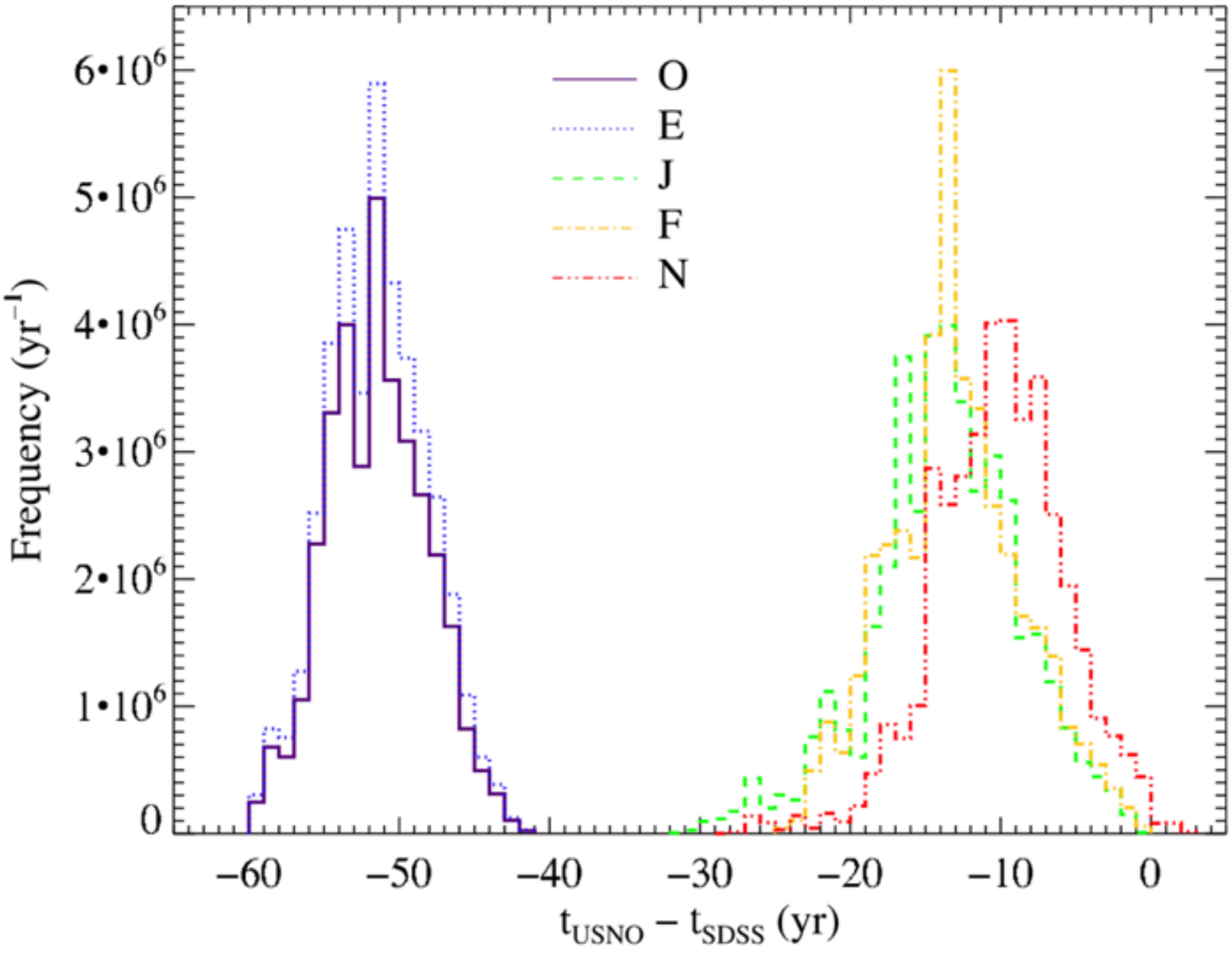


E''_{USNO}



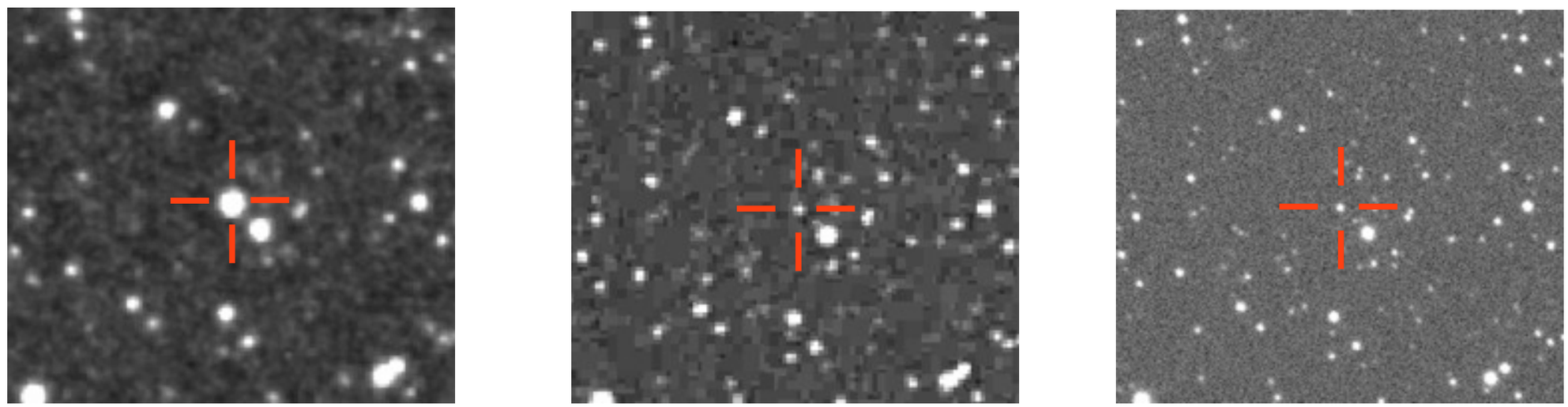
E_{SDSS}

Distribution of Epochs



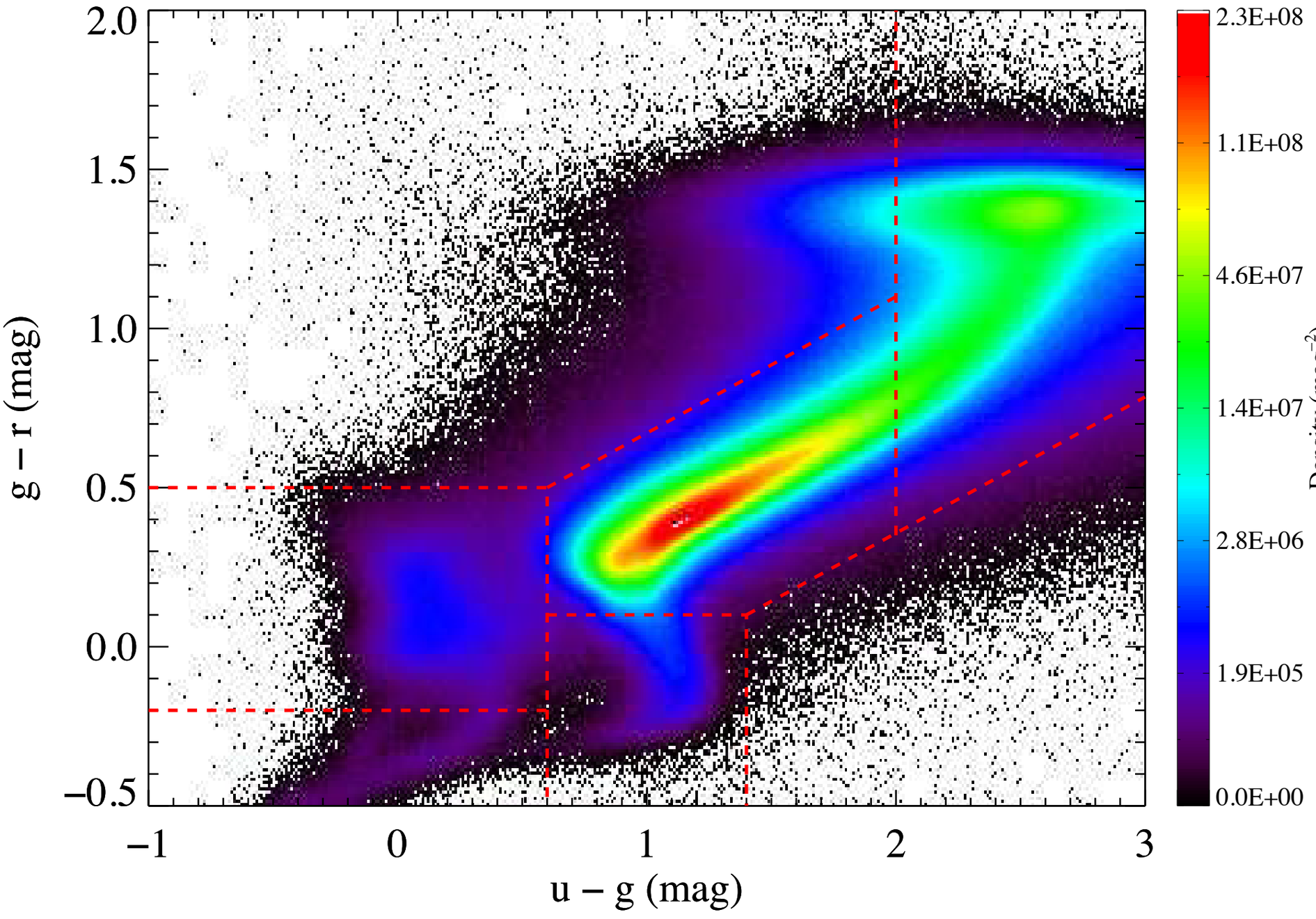
Large Amplitude Variables

- ≈ 1.5M objects exhibit $\Delta m > 1.0$ mag
  - *98% are spurious*; removed by flagging
- ≈ 250,000 candidate large amplitude (4σ) variables
  - visual inspection suggests ≈ 70% are real
  - only 4% appear in major variable catalogues (VSX, CRTS)



Classifying Variables

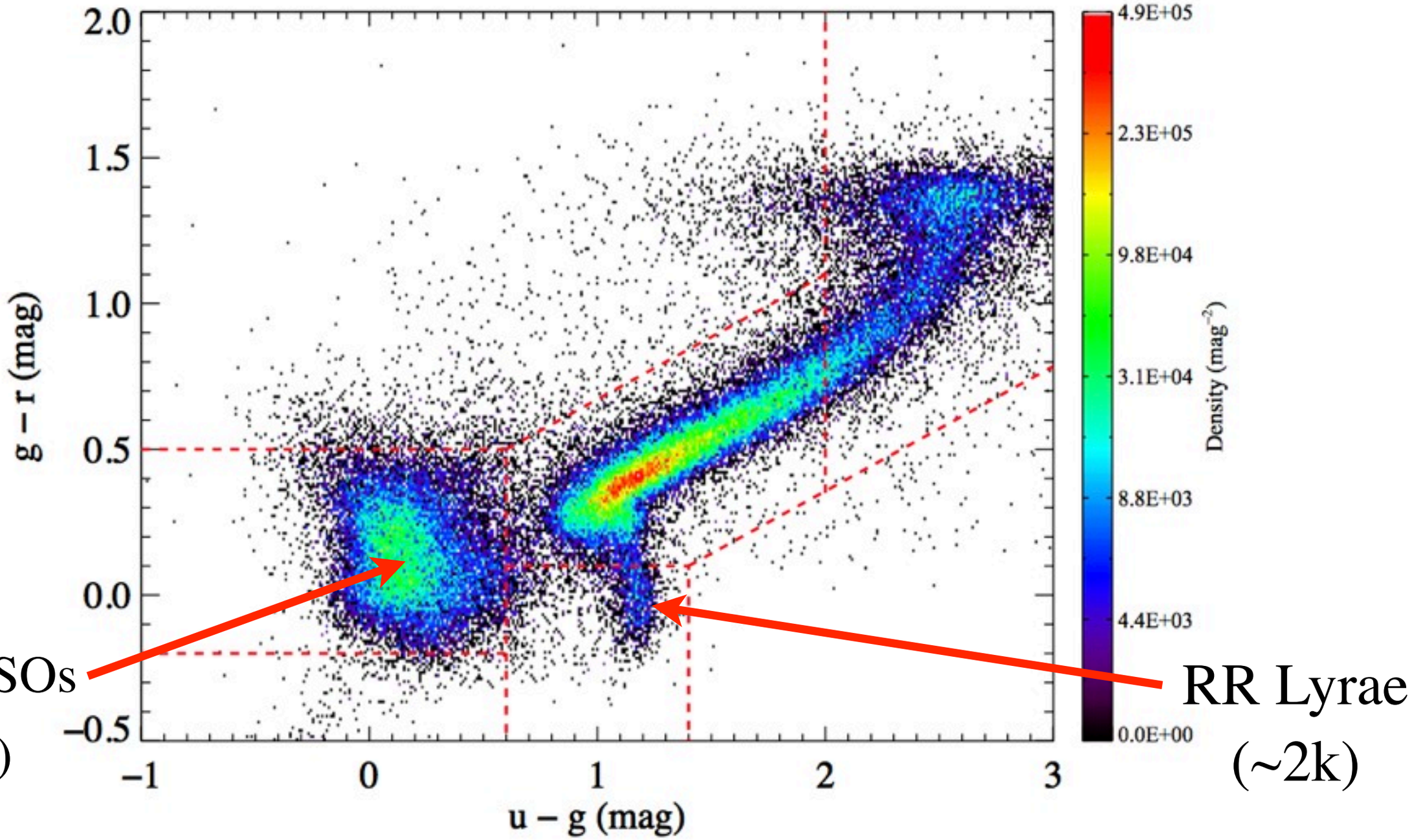
All cross-matched SDSS point sources



Richards et al 2002

Classifying Variables

4σ variables ("*J*"-band; 75k objects)



Low-*z* QSOs (~20k)

RR Lyrae (~2k)

Richards et al 2002

Summary

- Improved historical photometric catalogue
  - six-epoch photometry of ≈ 44M stars & quasars (1950-2010)
  - 95% complete to $g \approx 20.0$ mag; accurate to ≈ 0.1 mag
  - very low contamination rate
  - > 250,000 large amplitude variable candidates; mostly uncatalogued
  - legacy value for future surveys (e.g., LSST)
  - publicly available in convenient formats (arXiv:1309.6322; ApJS)
- Future extensions
  - ingest ongoing wide-field optical surveys to extend sky coverage; extend cross-matching to UV, IR, radio, etc.;
  - compare to large, higher-cadence variability surveys (OGLE, CRTS, PTF, TDSS, etc.)
  - identify “transients”