ZTF Data System: Status & Phase-II Plans

Frank Masci & the IPAC/Caltech ZTF Team

October 19, 2020














HEISING-SIMONS FOUNDATION

Some statistics for ZTF-I
Mar 17, 2018 – Sep 30, 2020

- Number of on-sky camera exposures: **178,865; 272,472; 27,947 = *g r i***
- Number of calibrated epochal quadrant-based *sci* images: **11,045,643; 16,764,041; 1,707,093**
- Percentage of un-calibratable (non-salvageable) quadrant images: **~ 3.7 %**

- Number of epochal image (non-difference) PSF-fit extractions: **~ 109B; 361B; 43B**
- Number of reference image PSF-fit extractions (= seeds for lightcurves): **~ 4.6B; 6.4B; 1.6B**

- Number of alerts from **positive** subtractions: **~ 47.48M; 105.6M; 7.05M**
- Number of alerts from **negative** subtractions: **~ 41.51M; 90.01M; 6.14M**
- Number of alerts associated with known SSOs (<= 3arcsec): **~ 6.5M**

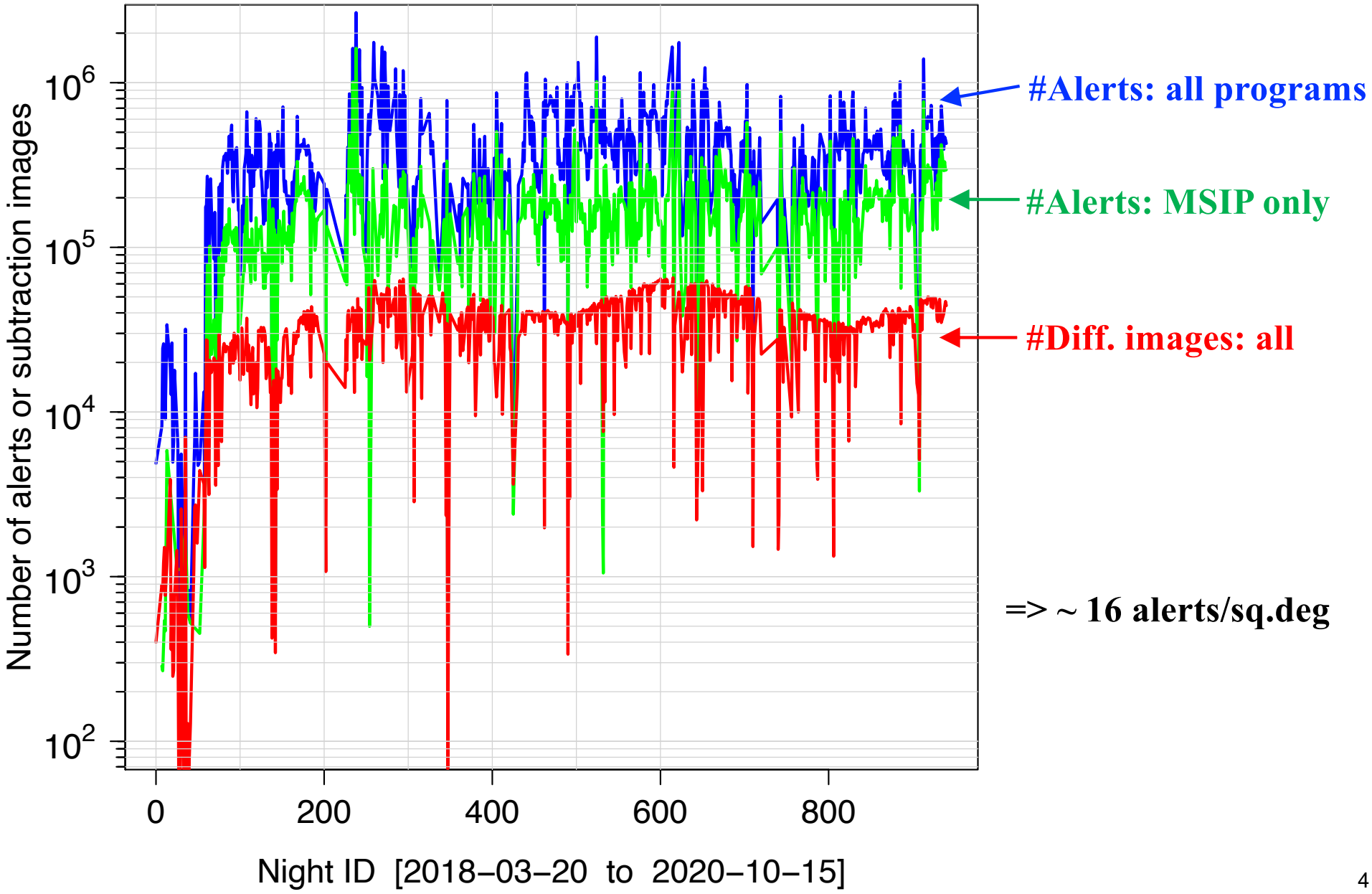
- Number of detected streaks associated with fast moving, known SSOs: **> 25,000**
- Number of linked point-sources (*tracklets*) associated with known SSOs: **2,934,868**

Forced Photometry usage for ZTF-I
Feb 6, 2019 – Sep 30, 2020

- Number of forced photometry requests (distinct sky-positions) since Feb 6, 2019: **8488**
- From **43** different users and broadly **8** institutions (according to email domain name):

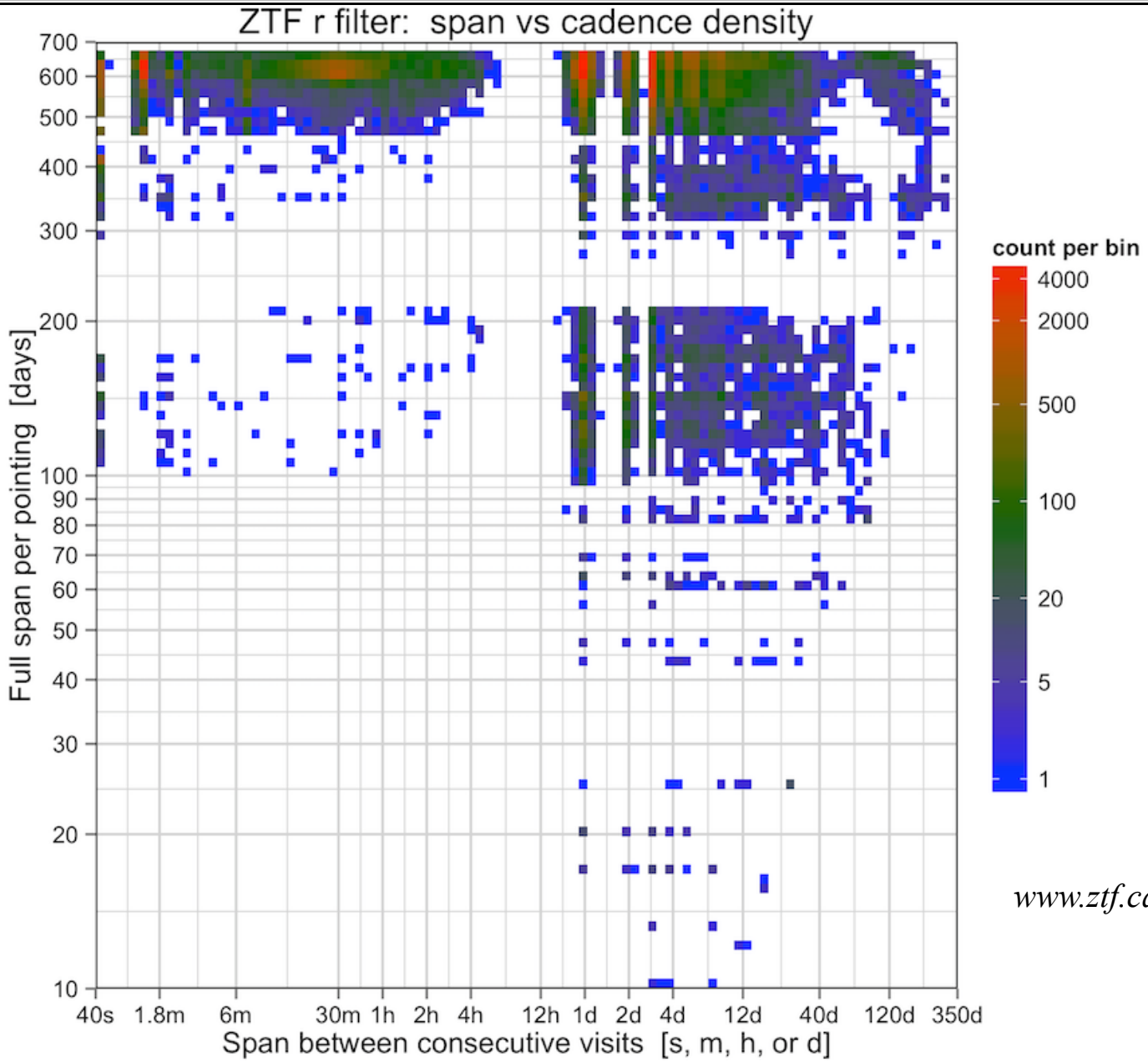
Domain	#Requests
fysik.su.se	4524
caltech.edu	1796
desy.de	1362
weizmann.ac.il	344
ipac.caltech.edu	232
astro.umd.edu	114
astro.su.se	103
uw.edu	13

Alert Statistics (Mar 20, 2018 – Oct 15, 2020)



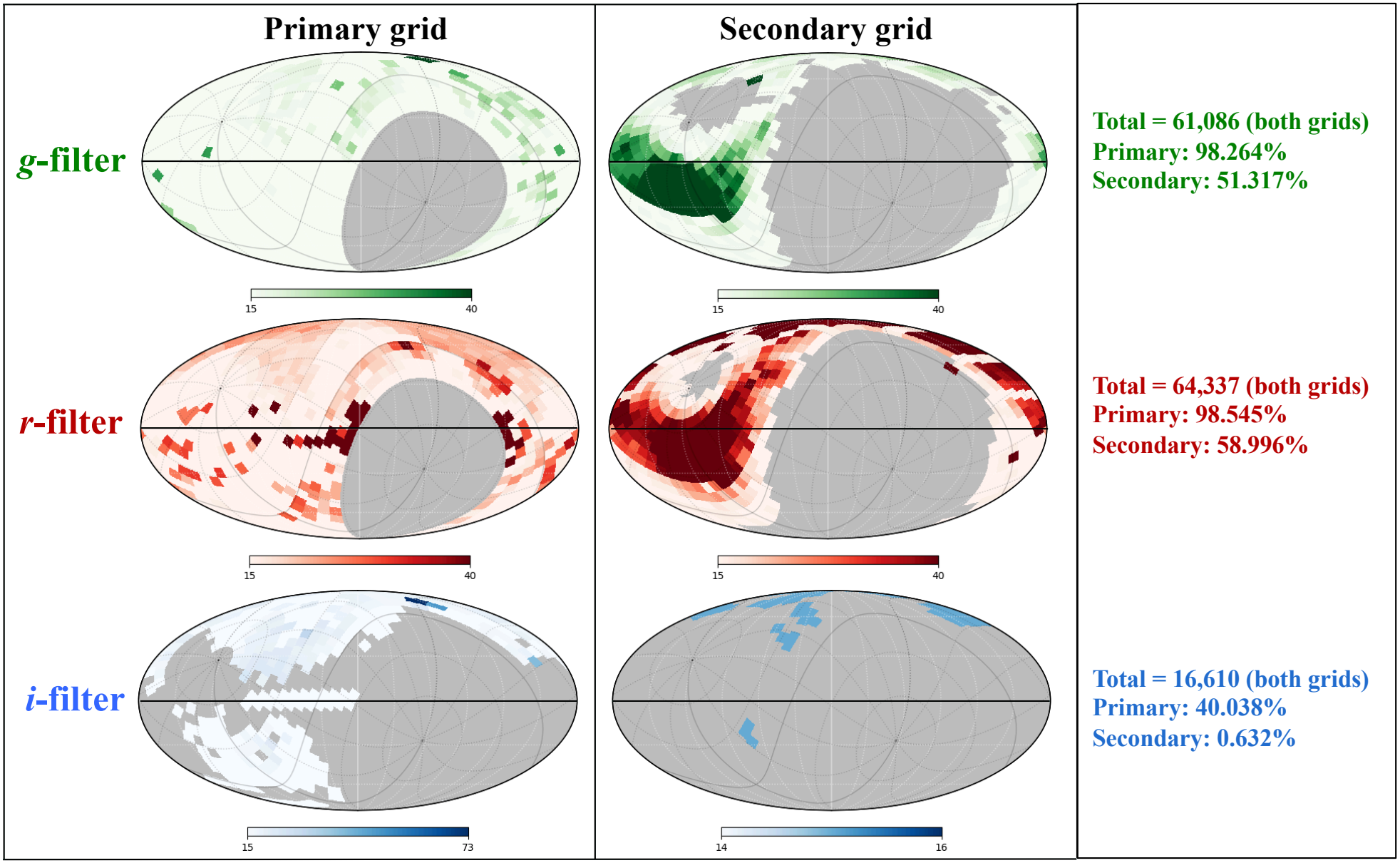
Third Public Data Release (June 2020)

lightcurve span vs sampling density

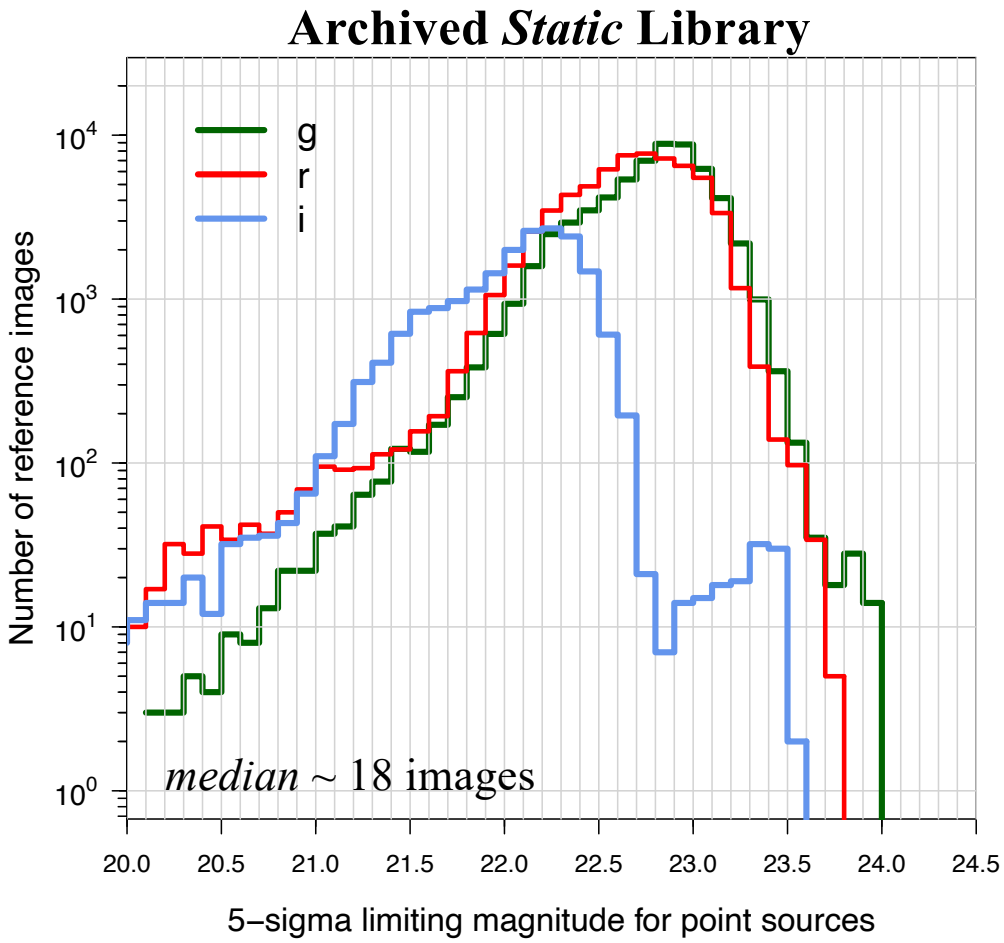


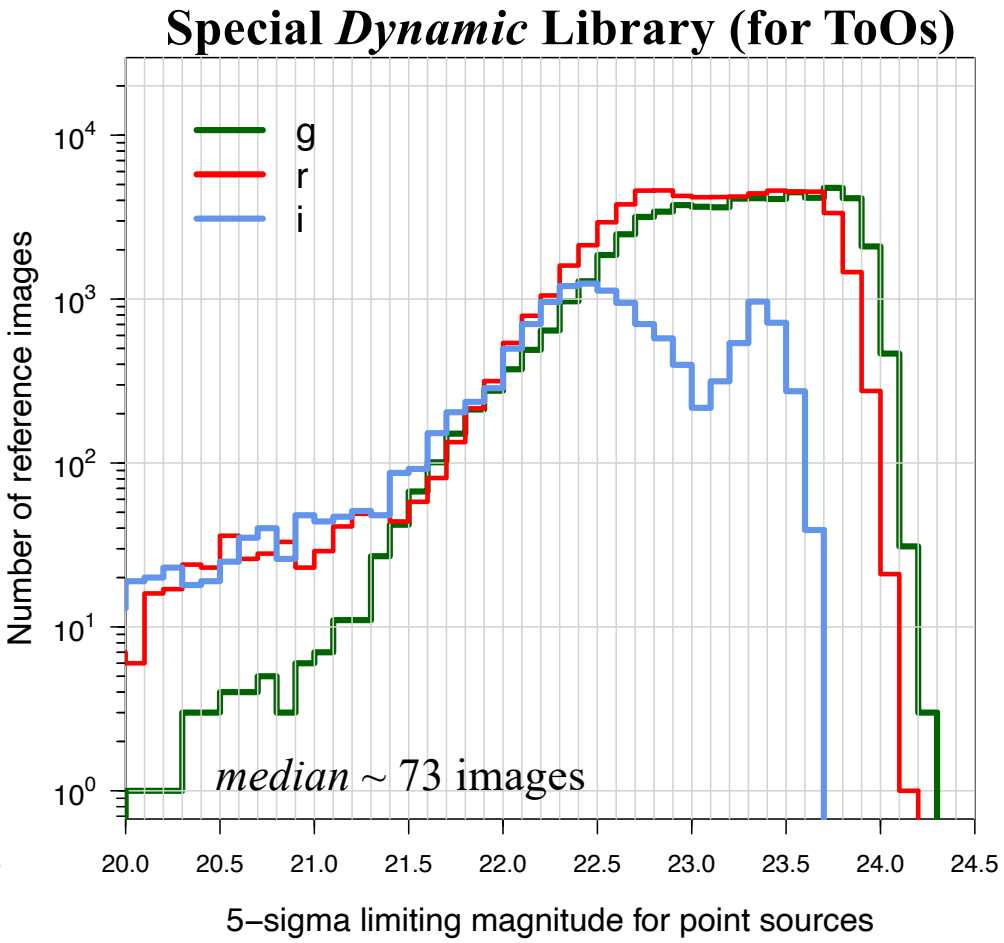
www.ztf.caltech.edu/page/dr3

Archived Reference Image Coverage: Oct 15, 2020
(l, b = 0, 0 centered)



Reference Image Depths
Archived versus Special (*internal*)





Fields with ultra-high source confusion:
noise & mag-limit estimators break down

Phase-II: Public Data Release plan

- **Each public data release contains:**
  - Raw CCD & calibration image data
  - Epochal calibrated science images and file-based source catalogs
  - Lightcurve photometry derived from positional re-matching across all epochs
  - New reference images and file-based source catalogs
  - Source database drawn from reference image catalogs to facilitate lightcurve retrieval
  - Updated quality flags and indicators for all the above

- **Move from a 6-month to 2-month embargo for public survey data after DR4**
- **Continue with 18-month embargo for private data: partnership and Caltech**

Release	Release Date	Public obs span	Private obs span
DR4	12/09/20	03/17/18 – 06/30/20	03/17/18 – 06/30/19
DR5	03/31/21	03/17/18 – 01/31/21	03/17/18 – 09/30/19
DR6*	06/30/21	03/17/18 – 04/30/21	03/17/18 – 12/31/19
DR7*	08/31/21	03/17/18 – 06/30/21	03/17/18 – 02/29/20
DR8*	11/03/21	03/17/18 – 08/31/21	03/17/18 – 04/30/20
DR9*	01/05/22	03/17/18 – 10/31/21	03/17/18 – 06/30/20

Etc …

* Bimonthly release cycle

Phase-II: prioritized major development & upgrades

Capability / functionality	Delivery date
Public forced photometry service	2020-12-01
Alert packed forced photometry histories	2020-12-20
Database, system, & infrastructure upgrades to support +3yr	2021-04-01
Bimonthly public release of file-based data products	2021-06-30 (commence DR6)
New lightcurve datastore	2021-10-01
More frequent release of lightcurves (tied to new datastore)	2021-11-01 (commence DR8)
Upgrade archive & access services for GRB Cam data	2021-02-01
P60 archive & access services	2021-10-01

Phase-II *ad hoc* improvements/features

- Mechanism for bulk data transfers to partner institutions.
- Update astrometric calibration framework to use *Gaia*-3.
- Propagate additional nearest *Gaia*-source metrics into alert packets; including proper motions.
- Include data in alert packets from the *PS1 Source Types & Redshifts with ML (PS1-STRM)* catalog.
- Update Star/Galaxy classification scores in alert packets using latest ML methods.
- Deploy Solar System Marshal at IPAC (vetting of streaks with *DeepStreaks* ML and scanning pages).
- Deployment of CCD-based "PIFF" automated PSF-estimation software from IN2P3 Team.
- Possible reprocessing of a subset of improperly calibrated Phase-I data following upgrades.
- Futuristic:
  - deploy *Tails* comet-finder ML framework at IPAC
  - use *Tails*-like architecture to find streaks

Reminder on documentation

- **Third Public Data Release: recipes for retrieving any data:**
  https://www.ztf.caltech.edu/page/dr3

- **Science Data System Explanatory Supplement**:
  https://irsa.ipac.caltech.edu/data/ZTF/docs/ztf_pipelines_deliverables.pdf

- **Science Data System paper**:
  https://iopscience.iop.org/article/10.1088/1538-3873/aae8ac

- **Archive access and services:**
  https://irsa.ipac.caltech.edu/Missions/ztf.html

- **Public alert archive and usage:**
  https://ztf.uw.edu/alerts/public/