ZTF Data System: Phase-II Plans & Schedule

Frank Masci & the IPAC/Caltech ZTF Team

October 23, 2020

The ZTF Science Data System (ZSDS)

- The ZSDS is housed at IPAC, Caltech
- IPAC is a multi-mission science center (IRAS, 2MASS, *Spitzer*, WISE/NEOWISE, LSST, Euclid, WFIRST..)
- Responsibilities for ZTF:
  - data processing: from instrumental calibration to alert packets and lightcurves.
  - quality assurance; product archiving; user-interfaces to retrieve and analyze data.
  - management of data releases and user-support.
  - on-demand forced-photometry lightcurve service.

Cahill Center for Astronomy & Astrophysics

IPAC

Mt Palomar ~ 150 km

ZTF Computing & Storage at IPAC

- We ingest ~ 300 GB of raw camera data *per night* to generate ~ 3.8 TB in products.
- Processing is in "real-time" and alert packets leave IPAC ~10 to 20 minutes since observation.
- Expect ~ 5.7 PB in data products at end of Phase-II.
- Compute cluster consists of 66 compute nodes (1192 processor cores @2.5GHz each).

Racks containing 66 compute nodes

Archive fileservers/disk arrays

ZTF Public Release Data Products

- Like Phase-I, Phase-II public data releases will include:
  - Refinements to products and data quality flags in all previous releases (back to start of phase-I).
  - Products derived from additional observations:

    - Raw camera & calibration image data.
    - Calibrated single-epoch images, difference images, masks, PSFs, and source catalogs.
    - Lightcurves derived from positional re-matching across all epochs (back to start of Phase-I).
    - New reference images (co-adds), depth-of-coverage maps, and source catalogs.
    - Source database drawn from reference image catalogs to facilitate lightcurve retrieval.

- Public Alerts from the public surveys are not tied to any bulk data release
  - These continue to be distributed close to real-time for consumption by Alert Brokers.

Phase-II Public Data Release plan

- Move from a 6-month to 2-month release cycle for all data.
- **For public survey data:** move from a 6-month to 2-month embargo following DR4.
- **For private data** (partnership/Caltech science programs): continue with 18-month embargoing.

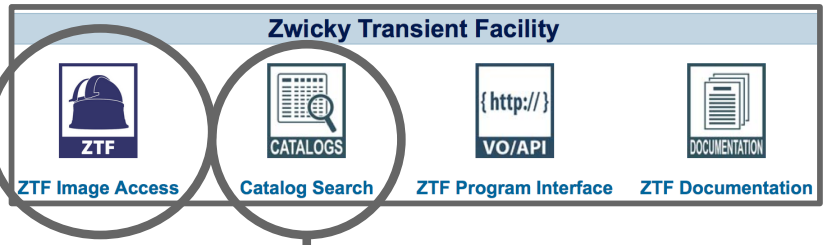
Release	Release Date	Public survey observation span	Private surveys observation span
DR4	12/09/20	03/17/18 – 06/30/20	03/17/18 – 06/30/19
DR5	03/31/21	03/17/18 – 01/31/21	03/17/18 – 09/30/19
DR6*	06/30/21	03/17/18 – 04/30/21	03/17/18 – 12/31/19
DR7*	08/31/21	03/17/18 – 06/30/21	03/17/18 – 02/29/20
DR8*	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
…	…		

* Bimonthly release cycle

Data Access and Visualization Tools

https://irsa.ipac.caltech.edu/Missions/ztf.html

- Access is through IRSA at IPAC.
- Can search for images and source catalog files by position or object name (including SSOs), sources extracted from co-adds & their lightcurves; overlays, time series viewer w/ interactive manipulation.
- Accompanying APIs (command-line driven retrieval) are available for all services.



Lightcurve retrieval / Time Series Tool

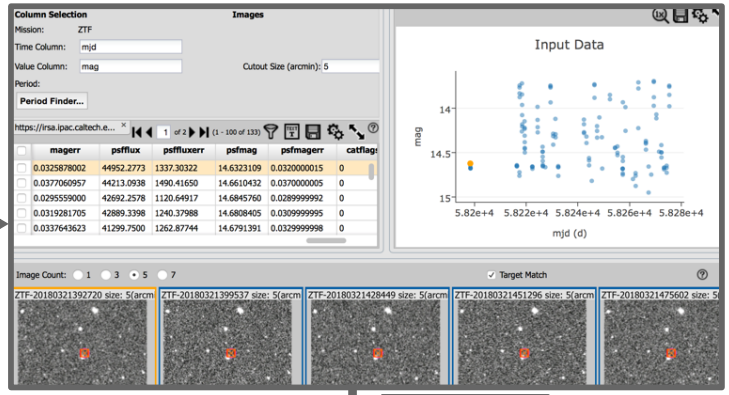
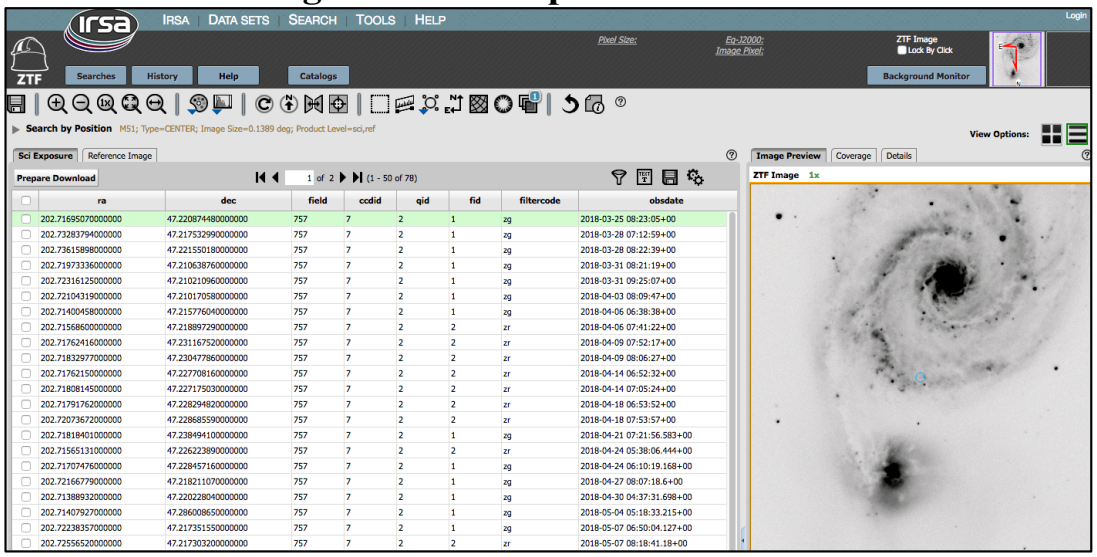
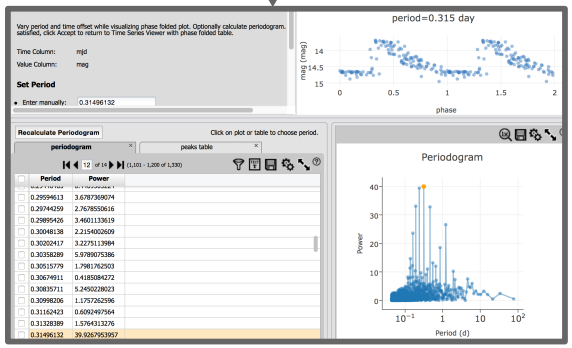


Image viewer and product retrieval



Period finding



Moving Object Search Tool (MOST)
asteroids pre-covered by ZTF imaging

NASA/IPAC INFRARED SCIENCE ARCHIVE

IRSA | DATA SETS | SEARCH | TOOLS | HELP | Login

MOST - Moving Object Search Tool

The Moving Object Search Tool (MOST) can determine the orbit for a given solar system object then find images t
object's predicted positions in select image datasets housed at IRSA (see Instructions). It can serve as a "precovery"
newly discovered objects were previously observed.

Submit Reset

Image Dataset ztf

For ZTF: Time Range = 2017-10-15 to present
For complete range, leave limits blank (but this may take a long time)
Example: Antonia 2017-11-01 to 2017-12-31

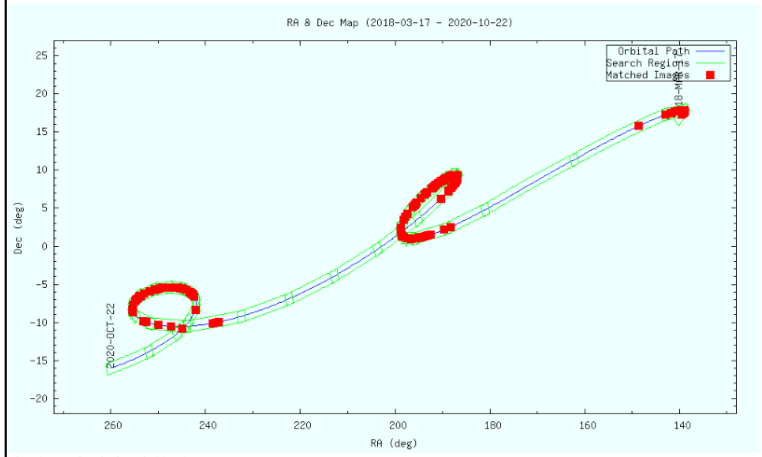
Observation Begin (UTC) 2018-03-17 **Observation End (UTC)** 2020-10-22

Ephemeris Step Size (day) 1 **Output Mode** Regular

Create Fits and DS9 Region Files Tarballs

- **Solar System Object Name Input:** Masci
- **Solar System Object NAIF ID Input:**
- **MPC One-Line Element Input:**
  - **Object Type:** Asteroid
- **Orbital Elements Manual Input:**

Orbital Path Plot for Asteroid: 438973 Masci



View/Download Plot: Orbit Plot

Images with a Matched Object Position

NOTE: Image ID may be only an identifier, not an actual file name.

Download/ View	DS9 File	Image ID	date_obs	time_obs	mjd_obs	ra_obj	dec_obj	r	Δ
			(UTC)	(UTC)	(day)	(deg)	(deg)	(AU)	(AU)
	region	ztf_20180325264456_000569_zg_c06_o_q4_sciimg.fits	2018-03-25	06:20:50	58202.26446760	139.423371	17.288453	3.755	3.012
	region	ztf_20180329190660_000569_zg_c06_o_q4_sciimg.fits	2018-03-29	04:34:34	58206.19067130	139.188657	17.458317	3.759	3.061
	region	ztf_20180329254352_000569_zg_c06_o_q4_sciimg.fits	2018-03-29	06:06:16	58206.25435190	139.185162	17.460838	3.759	3.062

Phase-II: major data-system upgrades

Capability / functionality	Release date
Public forced photometry service	2020-12-01
Forced photometry histories in alert packets (T – 30 days)	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 format and database	2021-10-01
More frequent release of lightcurves (tied to new datastore)	2021-11-01 (commence DR8)
P60 archive & data access service	2021-10-01

Phase-II: other improvements & features

- Update astrometric calibration framework to use *Gaia Early DR3*.
- Update automated PSF-estimation software to further improve photometric accuracy:
  goal is <~ 1% *absolute* precision (relative to PanSTARRS1), consistently on photometric nights.
- 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.
- Refine Star/Galaxy classification scores of nearest PS1 sources in alert packets using latest ML.
- Reprocess subsets of improperly calibrated Phase-I data following upgrades.

Data Access & Documentation

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

- **Access to Images, Catalogs, lightcurves, and analysis tools:**
  https://irsa.ipac.caltech.edu/Missions/ztf.html

- **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

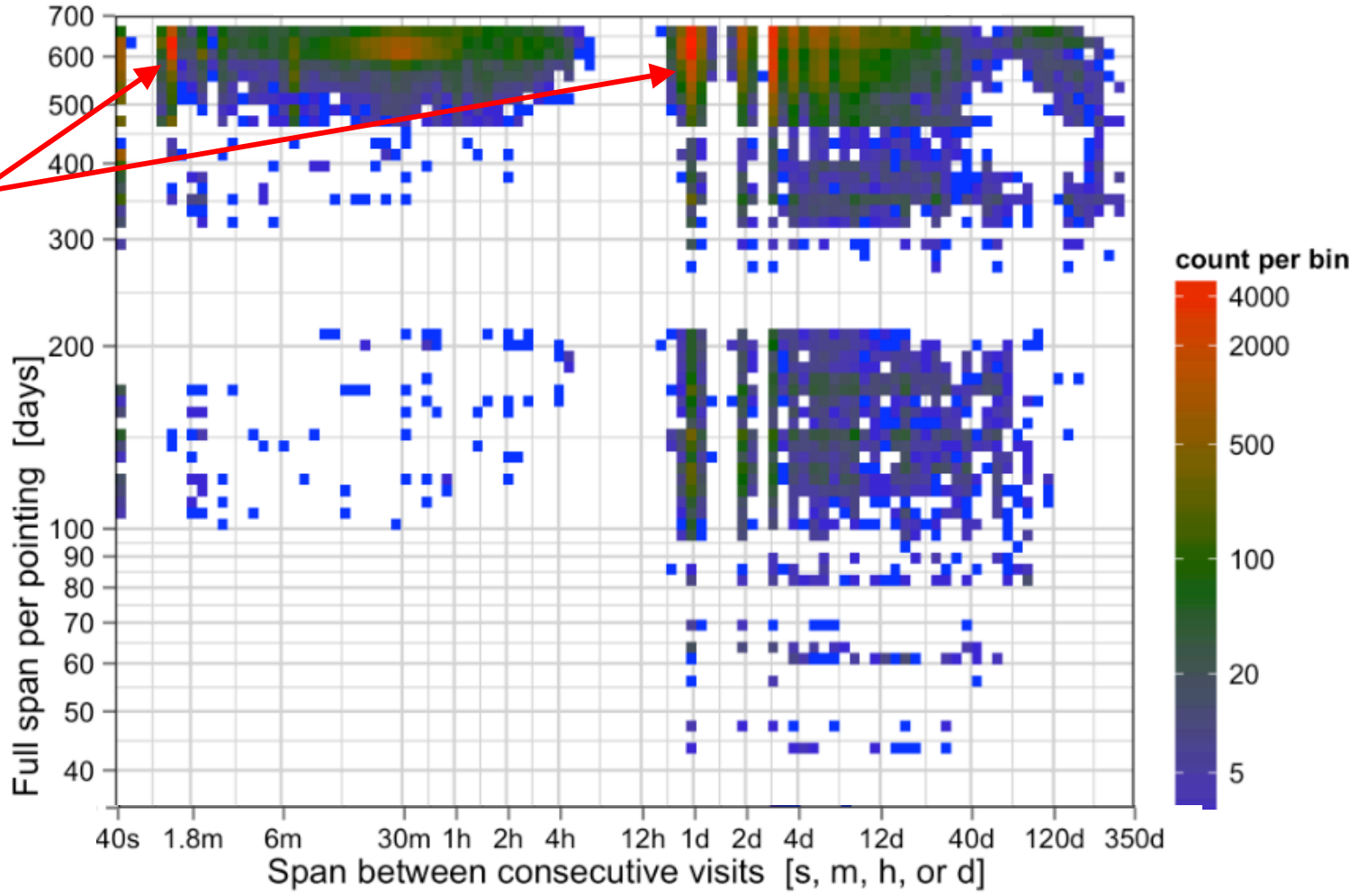
- **Growing archive of raw public alerts and usage documentation:**
  https://ztf.uw.edu/alerts/public/

Back up slides

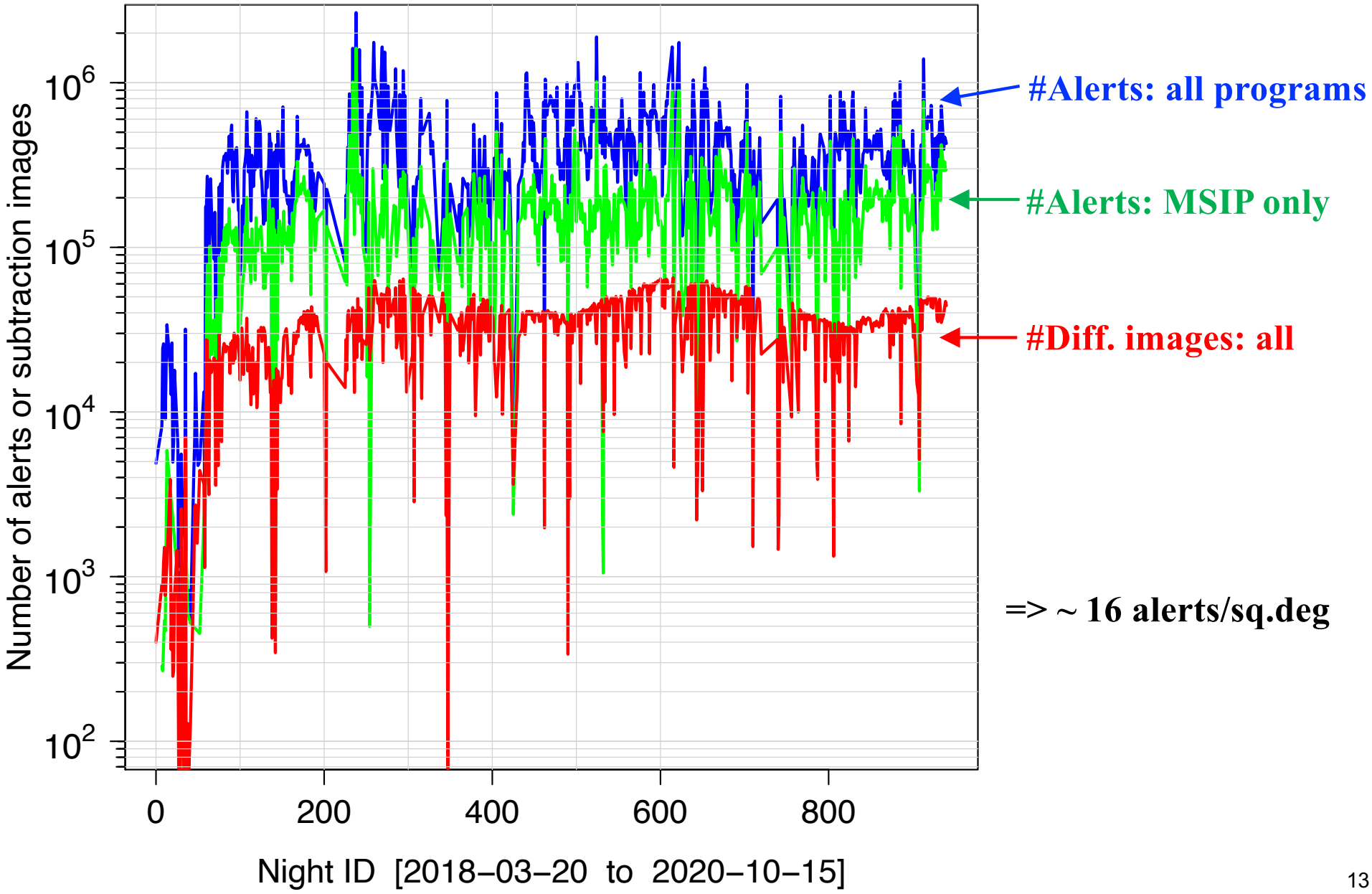
What can DR3 provide?
Lightcurve Timespans and Cadences

- Counts per bin are #visits *per 47 square degree field*.
- Sampling is non-uniform and irregular due to seasonal variation in field visibility from P48.
- Constrains the variability-related archival science projects possible with DR3.

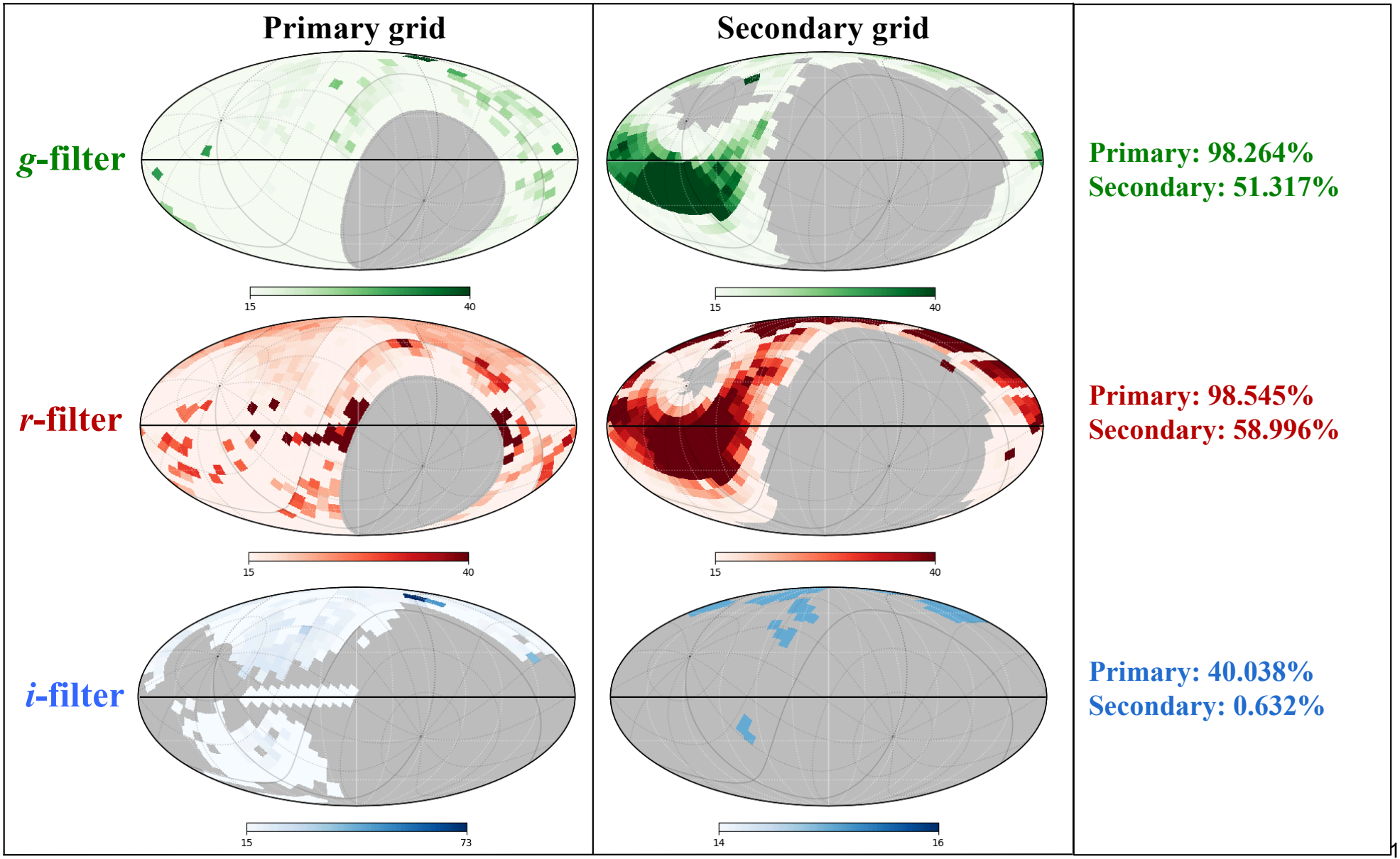
includes galactic plane!



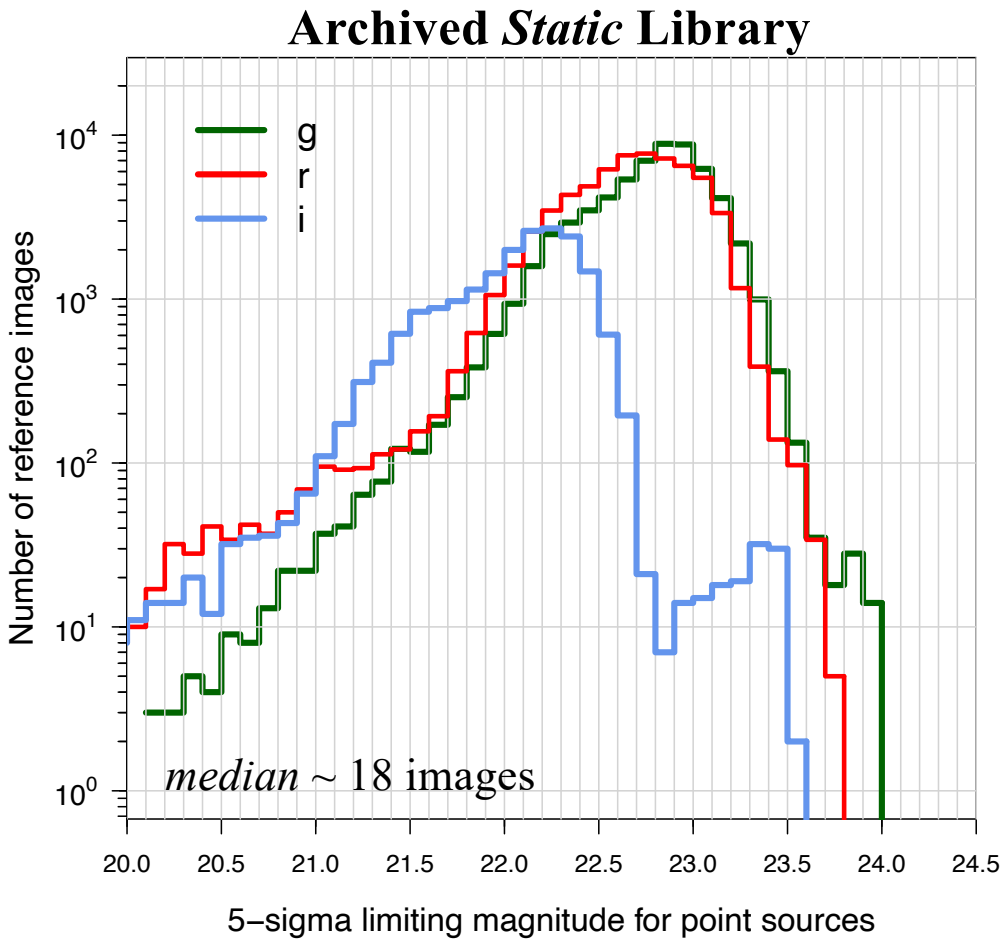
Alert Statistics (Mar 20, 2018 – Oct 15, 2020)

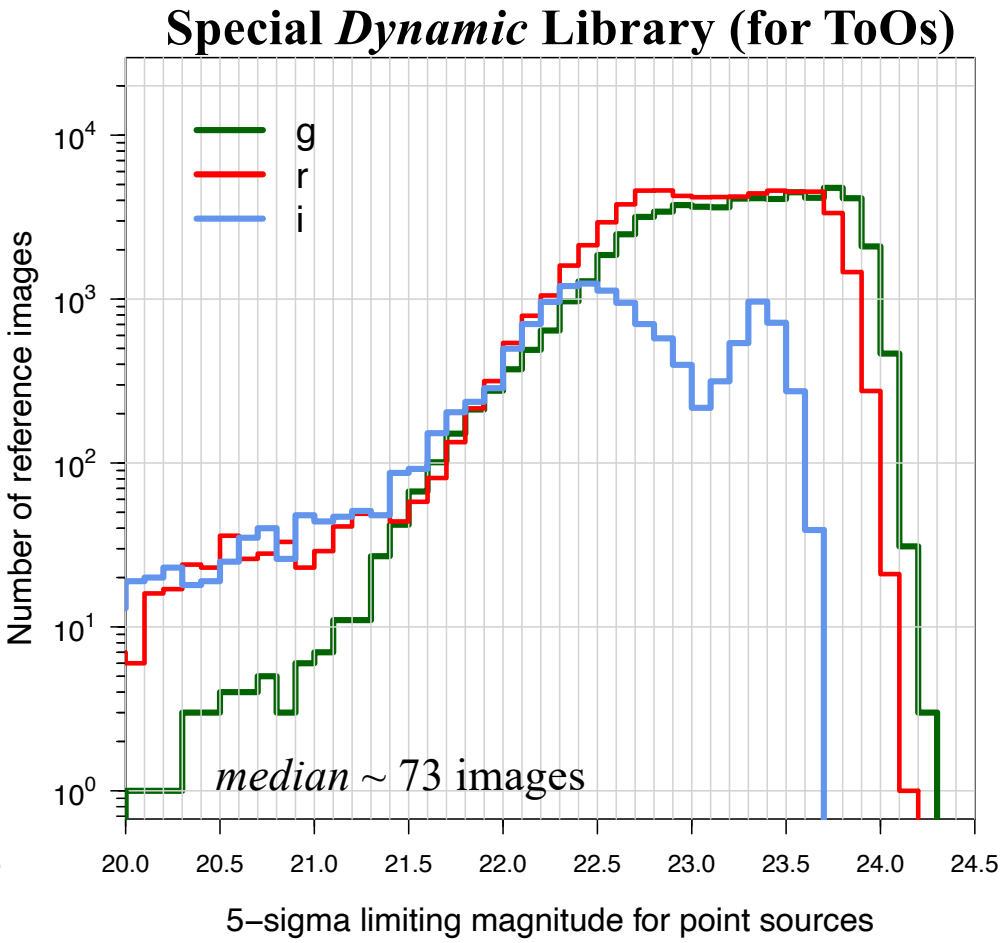


Reference Image Coverage as of Oct 15, 2020
(l, b = 0, 0 centered)



Reference Image Depths
Archived versus Special (*internal*)





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