Public Data Release 4

Contents

1. Summary of the ZTF Surveys
   a. Acknowledging ZTF and Referencing Data Product usage
2. Scope of this Release
3. Summary of Data Products
4. File-based Product Statistics
5. Source and Lightcurve Statistics
6. DR4 Sky Coverage
7. Lightcurve products: timespans, cadence, and epoch distributions
8. File-based Products in Archive
   a. Single-Exposure Image Products
   b. Reference Image Products
   c. Other File Products
9. Epoch-based Data Quality Flags and Filtering
   a. Quality Filtering of Single-Exposure File Products
   b. Quality Filtering of Epochs in Lightcurves
10. What you should know about the ZTF Release Products
11. Extended (detailed) Advisories and Cautionary Notes
12. How to Access ZTF Data Products
   a. Product-specific Access GUIs and Recipes
   b. Example Queries using the APIs
      i. Querying Single-Exposure Science Image Products
      ii. Querying Reference Image Products
      iii. Querying Lightcurves
      iv. Solar System Object Precovery
   c. Bulk Download of all Lightcurves
13. Help Desk
14. Further Documentation & Relevant Publications

An overview of ZTF and the science objectives is given in Bellm et al. 2019 and Graham et al. 2019. The ZTF Science Data System (ZSDS) is housed at IPAC/Caltech and is described in Masci et al. 2019. This page provides a summary and quick guide to the ZTF Public DR4 products.
1. Summary of the ZTF Surveys

During Phase-I of the survey (March 2018 - September 2020), ZTF observing time was split three ways:

(i) The Public Survey (40%);
(ii) Private Collaboration and Partnership Surveys (40%);
(iii) Programs granted by the Caltech Time Allocation Committee (20%).

Funding is provided by both the ZTF collaboration and the U.S. National Science Foundation through the Mid-Scale Innovations Program (MSIP).

The Public Survey in particular enabled a wide range of community science and constituted the following:

- A Northern-equatorial sky survey: a three-night cadence of the Northern sky visible from Palomar Observatory ($\approx 2.5$ to $3\pi$ steradians), comparable to the LSST wide-fast-deep cadence, irrespective of Galactic latitude. This component surveys $\approx 15,000$ to $20,000$ square degrees every three nights.
- A Galactic plane-targeted survey: a one-night cadence survey of the Galactic plane at $|b| \leq 7^\circ$ and Dec. $>-25^\circ$, whenever a survey-field falls in this footprint. The area covering this footprint is seasonal and varies from $\approx 1000$ to $2000$ square degrees.
- For each of the above surveys, there is one $g$-filter and one $r$-filter exposure of the same pointing separated by at least 40 minutes.

1.a. Acknowledging ZTF and Referencing Data Product usage

For publications that use ZTF Science Data Products from Phase-I of the survey, please include the following text in your acknowledgments:

*Based on observations obtained with the Samuel Oschin 48-inch Telescope at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under Grant No. AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW.*

Furthermore, we appreciate you citing the following publication:


2. Scope of this Release

The Fourth ZTF Public Data Release (DR4) builds upon the Third Data Release (DR3) to include products from (i) an additional six months of data acquired under the Public Survey (+ January to Juner 2020) and (ii) data acquired under Private Survey time (from both partnership and Caltech programs) from approximately the first 15.4 months of science operations (~ March to June 2019). The latter is included due to the 18 month proprietary period up to this release date.

More specifically, DR4 includes the following:

1. **Single Exposure Products** associated with individual science observation epochs (lightcurve epochs) in $g$ and $r$ filters only satisfying:
   - Program ID = 1: implying products from the Public Survey, FOR:
   - Observation MJD $\approx 58194.0$ (= 2018-03-17 00hrs UTC), AND
2. **Single Exposure Products** associated with individual science observation epochs (lightcurve epochs) in g, r, and i filters satisfying:
   - Observation MJD = 59030.0 (= 2020-06-30 00hrs UTC), AND
   - Observation MJD = 58194.0 (= 2018-03-17 00hrs UTC), AND
   - Observation MJD = 58664.0 (= 2019-06-30 00hrs UTC), AND
   - Program IDs = 2 and 3: implying products from the Private Surveys, FOR:
   - Observation MJD ≥ 58194.0 (= 2018-03-17 00hrs UTC), AND
   - Observation MJD ≤ 58664.0 (= 2019-06-30 00hrs UTC), AND

3. All **Reference (co-add) Image Products** available in the ZTF Archive as of November 15, 2020 (when products were frozen). Only good quality science exposures from all survey programs, irrespective of the epoch spans above, were included in the co-adds.

**NOTE:** The public distribution of transient alerts extracted from difference images constructed from products (1) and (3) above commenced on June 4, 2018. These alerts continue to be generated and distributed using new (unreleased) image data from Phase-II of the Public Survey.

### 3. Summary of Data Products

In accord with the criteria for data-inclusion in Section 2, the following products are available in DR4:

- **Raw CCD-based image data files and metadata in FITS format.**
- **CCD-quadrant-based calibration image data and metadata in FITS format.**
  - Stacked bias images.
  - Relative pixel-responsivity maps (flatfield corrections).
  - Bad-pixel masks derived from these calibrations.
- **Instrumentally calibrated CCD-quadrant-based, single-exposure science images, source-catalogs, and metadata; specifically:**
  - Science images with astrometric and photometric calibration solutions, and QA information.
  - Corresponding bit-mask images.
  - PSF templates (image and look-up-table) for each science image.
  - Source-catalog FITS-table files containing exclusively PSF-fit photometry measurements.
  - Source-catalog FITS-table files containing exclusively concentric aperture photometry & source-shape metrics.
  - Difference images constructed in the sense: "science (single-exposure) image minus reference image". For DR4, these are only available for science images that had a preexisting reference image when the science image was first processed.
  - Effective PSF template for difference image.
  - Tar file of already publicly-released alert packet files (in Avro format) corresponding to the CCD-quadrant.
  - Processing log files.
- **Reference Images (co-added/stacked single-exposures) per CCD-quadrant, source-catalogs, and metadata; specifically:**
  - Co-add images with astrometric and photometric calibration solutions, and QA information.
  - Corresponding pixel-uncertainty images.
  - Corresponding pixel depth-of-coverage images (recording effective number of input exposures per pixel).
  - Source-catalog FITS-table files containing exclusively PSF-fit photometry measurements.
  - Source-catalog FITS-table files containing exclusively concentric aperture photometry & source-shape metrics.
  - Processing log files.
- **Objects (database) Table storing sources from all Reference Image PSF-fit-based catalog files.**
  - Includes "collapsed-lightcurve" metrics (from prior position-matching of detected sources across all overlapping, CCD-quadrant-based exposures in the public survey).
  - This table facilitates initial object-position searches and filtering on any of the lightcurve-collapsed metrics for subsequent extraction of lightcurve data.
- **A datastore of lightcurves containing additional metrics at the single-epoch (exposure) level.** This is searchable by interfacing with queries on the Objects Table. Lightcurve data can be visualized and downloaded via a GUI or API (Sections 12a and 12b).
- **Entire lightcurve holdings packaged into gzipped tar files (Section 12c)**
For details on data and file formats, associated ancillary products, product construction and usage, see the ZSDS Explanatory Supplement. See Section 10 below for a list of MUST READ ADVISORIES about these products. See Section 11 below for a more detailed list of advisories and known caveats.

4. File-based Product Statistics

The ZTF camera consists of 16 CCDs, each of which is partitioned into 4 readout quadrants. Therefore, 64 CCD-quadrant images are generated per ZTF exposure. A CCD-quadrant is the basic image-unit for pipeline processing and from which all science data products are derived. A CCD-quadrant covers $0.854° \times 0.854°$ on the sky. Table 1 summarizes the CCD-quadrant-based (single exposure) image counts. The second column shows the number of full camera exposures from which these were generated. The CCD-quadrant image counts are separated into "bad" (probably unusable due to non-photometric conditions) and "good" images. The criteria used to identify "bad" vs "good" images are described in Section 2.4 of the ZSDS Advisories & Cautionary Notes. The relative fractions of bad and good images in this table provide an indicator of the overall fraction of "bad" (likely suspect) single-epoch photometry measurements within individual lightcurves.

<table>
<thead>
<tr>
<th>Filter(s)</th>
<th>#Exposures</th>
<th>#Good (usable)</th>
<th>#Bad (suspect)</th>
<th>Total number</th>
<th>Bad / Total [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td>125,835</td>
<td>7,199,211</td>
<td>484,361</td>
<td>7,683,572</td>
<td>6.30</td>
</tr>
<tr>
<td>$r$</td>
<td>169,921</td>
<td>9,575,649</td>
<td>769,871</td>
<td>10,345,520</td>
<td>7.44</td>
</tr>
<tr>
<td>$i$</td>
<td>8,052</td>
<td>432,369</td>
<td>33,011</td>
<td>465,380</td>
<td>7.09</td>
</tr>
<tr>
<td>$g + r + i$</td>
<td>303,808</td>
<td>17,207,229</td>
<td>1,287,243</td>
<td>18,494,472</td>
<td>6.96</td>
</tr>
</tbody>
</table>

Table 1: Number of exposures and "good / bad" CCD-quadrant-based science images in DR4

The number of CCD-quadrant-based reference images (co-adds) per filter in DR4 are shown in Table 2. Note: unless brought to our attention, all of these are expected to have good photometric and astrometric quality in general. This is because they were constructed from good-quality single-exposure images satisfying a range of criteria, which are more stringent than those used to tag the "good" images in Table 1. The criteria used to select inputs for generating reference images are described in Section 6.7 of the ZSDS Explanatory Supplement.

The existence of a reference image for a specific survey field, CCD-quadrant, and filter is important since it means that for this region of sky: (i) image differencing and alerts can be triggered, and (ii) objects can be extracted and then used as seeds to generate lightcurves. The third column in Table 2 shows the percentage of ZTF survey fields/CCD-quadrants with at least one exposure (regardless of quality) that have a reference image. Note: at least 15 good-quality images are required to generate a reference image. There are fields/CCD-quadrants not visited as often during the survey and therefore lack reference images. The fourth column in Table 2 shows the approximate percentage of P48-accessible sky covered in reference images, as constrained by the area covered by any exposures on the primary survey grid.

<table>
<thead>
<tr>
<th>Filter(s)</th>
<th>#Reference Images</th>
<th>Visited-sky coverage [%]</th>
<th>Accessible-sky coverage [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td>61,910</td>
<td>81.24</td>
<td>98.16</td>
</tr>
<tr>
<td>$r$</td>
<td>65,142</td>
<td>81.31</td>
<td>98.31</td>
</tr>
<tr>
<td>$i$</td>
<td>17,368</td>
<td>30.35</td>
<td>41.75</td>
</tr>
<tr>
<td>$g + r + i$</td>
<td>144,420</td>
<td>67.63</td>
<td>79.41</td>
</tr>
</tbody>
</table>
5. Source and Lightcurve Statistics

Table 3 reports the number of sources in all PSF-fit-based and aperture-based catalog file products (summarized in Section 3), separately for single-exposure (sci) images and reference (ref or co-add) images. The PSF-fit and aperture catalogs are treated independently (with no merging or cross-linking of records) in the Data System. This means a large fraction of the same sources are present in both flavors of catalog. The PSF-fit-based catalogs contain more sources since first, PSF-fitting generally yields more accurate photometry at faint fluxes and hence is more sensitive, and second, it includes deblending of closely-separated sources. Deblending is more important in high source-density regions such as the Galactic plane.

Only sources from the PSF-fit-based sci catalogs are positionally-matched across observation epochs to construct lightcurves (see Table 4 for numbers). As mentioned above, sources in the deeper PSF-fit-based ref catalogs are initially used as seeds to trigger source-matching and generate lightcurves.

<table>
<thead>
<tr>
<th>Filter(s)</th>
<th>#PSFcat-sci sources</th>
<th>#Aperturecat-sci sources</th>
<th>#PSFcat-ref sources</th>
<th>#Aperturecat-ref sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td>74,126,685,884</td>
<td>46,764,103,340</td>
<td>2,205,550,406</td>
<td>689,728,722</td>
</tr>
<tr>
<td>$r$</td>
<td>217,553,318,884</td>
<td>134,048,046,583</td>
<td>3,051,221,036</td>
<td>1,040,799,684</td>
</tr>
<tr>
<td>$i$</td>
<td>4,241,251,453</td>
<td>2,529,992,274</td>
<td>842,675,797</td>
<td>257,078,955</td>
</tr>
<tr>
<td>$g + r + i$</td>
<td>295,921,256,221</td>
<td>183,342,142,197</td>
<td>6,099,447,239</td>
<td>1,987,607,361</td>
</tr>
</tbody>
</table>

Table 3: Number of sources in CCD-quadrant-based catalog files in DR4, according to extraction and image type

Table 4 reports the number of lightcurves from positional-matching of sources in the single-exposure PSF-fit-based science catalogs with length (number of independent epochs) greater than some limit. The distribution in the number of observation epochs and corresponding timespans is shown in Figure 3. All observation epochs, regardless of data quality are included in the counts in Table 4. For an estimate of the average fraction of "bad" epochs per lightcurve (with likely suspect and unusable photometry), see Table 1. The criteria used to identify "bad" vs "good" epochal data are described in Section 2.4 of the ZDS Advisories & Cautionary Notes.

<table>
<thead>
<tr>
<th>Filter(s)</th>
<th>#lightcurves with $N_{obs} \geq 1$</th>
<th>#lightcurves with $N_{obs} \geq 2$</th>
<th>#lightcurves with $N_{obs} \geq 5$</th>
<th>#lightcurves with $N_{obs} \geq 10$</th>
<th>#lightcurves with $N_{obs} \geq 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$</td>
<td>1,130,129,239</td>
<td>935,173,342</td>
<td>777,837,837</td>
<td>652,970,013</td>
<td>492,664,548</td>
</tr>
<tr>
<td>$r$</td>
<td>1,869,471,288</td>
<td>1,602,146,112</td>
<td>1,367,437,169</td>
<td>1,200,623,178</td>
<td>998,686,568</td>
</tr>
<tr>
<td>$i$</td>
<td>231,760,229</td>
<td>168,246,778</td>
<td>115,370,894</td>
<td>91,981,800</td>
<td>57,062,399</td>
</tr>
<tr>
<td>$g + r + i$</td>
<td>3,231,360,756</td>
<td>2,705,566,232</td>
<td>2,260,645,900</td>
<td>1,945,574,991</td>
<td>1,548,413,515</td>
</tr>
</tbody>
</table>
6. DR4 Sky Coverage

Figure 1 shows the approximate spatial distribution in the number of single-exposure epochs in each filter included in DR4, as Aitoff-Hammer projections in equatorial coordinates. These are approximate since the spatial resolution (bin size) of each map is 1° while a CCD-quadrant subtends \( \approx 0.854° \times 0.854° \) at the center of a map. The small-scale dark regions are not holes in coverage, but due to aliasing from the warped reprojection and resampling of CCD-quadrant centers onto larger scales. The \( \sim 3 \) arcmin gaps between the camera's CCDs also contribute to these holes.

FITS-formatted image representations of Figure 1 with Word Coordinate System metadata in their headers can be downloaded for each filter:

- \( g \)-filter single-exposure coverage map: sciimage_galcovmap_dr04_g.fits.gz
- \( r \)-filter single-exposure coverage map: sciimage_galcovmap_dr04_r.fits.gz
- \( i \)-filter single-exposure coverage map: sciimage_galcovmap_dr04_i.fits.gz

These maps can be visualized using a FITS-image viewer to determine if a specific sky-position is covered by a survey-field/CCD-quadrant. If so, the approximate number of overlapping epochs on/near this position (to within \( \approx 1° \)) can be determined from examining the pixel values.

**Figure 1** - Sky coverage and number of observation epochs per approximate CCD-quadrant footprint represented in DR4, in equatorial coordinates centered at \( ra, dec = 0, 0 \). **Left:** \( g \)-filter; **Center:** \( r \)-filter; **Right:** \( i \)-filter. Color bar represents the number of observation epochs. Dark regions on small scales are not holes in coverage, but due to the coarse resampling of CCD-quadrant centers on 1° scales. This resampling also distorts the true number of epochs per pointing (see Figure 3a for the true epoch distribution). Click on a panel to enlarge.

Figure 2 shows the approximate spatial distribution in CCD-quadrants per filter from which lightcurves were generated for inclusion in DR4. These maps closely track the Reference Image coverage since (PSF-catalog-based) sources extracted therefrom were used as seeds to trigger the lightcurve generation. As in Figure 1, these maps are approximate due to their finite spatial resolution (bin size) of 1° compared to the remapped CCD-quadrants that subtend \( \approx 0.854° \times 0.854° \) at the center of a map. The small-scale dark regions are not holes in coverage, but due to aliasing from the warped reprojection and resampling of CCD-quadrant centers onto larger scales. The \( \sim 3 \) arcmin gaps between the camera's CCDs also contribute to these holes.

FITS-formatted image representations of Figure 2 with Word Coordinate System metadata in their headers can be downloaded for each filter:

- \( g \)-filter lightcurve (and \( \approx \) Reference Image) coverage map: matchfile_galcovmap_dr04_g.fits.gz
- \( r \)-filter lightcurve (and \( \approx \) Reference Image) coverage map: matchfile_galcovmap_dr04_r.fits.gz
- \( i \)-filter lightcurve (and \( \approx \) Reference Image) coverage map: matchfile_galcovmap_dr04_i.fits.gz

These maps can be visualized using a FITS-image viewer to determine if a specific position is covered by a survey field/CCD-quadrant. If so, one can determine if lightcurves (and/or Reference Images) exist on/near this position, to within \( \approx 1° \).
7. Lightcurve products: timespans, cadence, and epoch distributions

To complement the epoch-coverage maps in Figure 1 and lightcurve numbers in Table 4, Figure 3 shows the distribution in overall timespan per survey-field (latest minus earliest exposure in DR4) versus the number of observation epochs in each.

Figure 3 - Timespan of observations per survey-field as a function of the number of epochs in each. These timespans are proxies for the overall spans of lightcurves in DR4.

Figure 4 encapsulates the available timespans and cadences (separation between consecutive epochs) in DR4, for each filter. These plots can be used as follows: for a given observation or effective lightcurve timespan per survey field (or pointing), one would take a horizontal slice to infer the distribution of visit separations (cadences) therein. The number of times a lightcurve is sampled at that cadence is indicated by the color bar.
Figure 4 - Effective timespan per pointing (per \( \approx 7^\circ \times 7^\circ \) survey field) versus visit cadence. See text above for details. **Left:** g-filter; **Center:** r-filter; **Right:** i-filter. **Click on a panel to enlarge.**

## 8. File-based Products in Archive

### 8.a. Single-Exposure Image Products

The calibrated CCD-quadrant-based science images have the following generic root paths and filenames in the archive:

```
YYYY/MMDD/fracday/ztf_filefracday_field_filtercode_cccdid_imgtypecode_qqid_sciimg.fits
```

where:
- **YYYY** = year
- **MMDD** = two-digit month and two-digit day
- **fracday** = fractional day since UT 0hrs
- **filefracday** = YYYYMMDDfracday
- **field** = six-digit survey field ID
- **filtercode** = zg, zr, or zi for g, r, i filters respectively
- **ccdid** = two-digit CCD identifier: 01..16
- **imgtypecode** = "o" for on-sky data; "f" for dome flat; "b" for bias
- **qid** = one-digit quadrant ID within specific ccdid: 1..4

An archived "sci" image can be prepended with the following root URL to access and download the specific `sciimg.fits` file:

https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/

For example:

https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/2018/0325/251921/ztf_20180325251921_000314_zg_c08_o_q2_sciimg.fits

The path and filename identifiers for science images covering a sky position and/or time range can be obtained using an API query (see Section 12b.i).

Alternatively, the `sciimg.fits` suffix can be replaced with any of the following ancillary file suffixes, with products summarized in Section 3:
mskimg.fits - bit-mask image corresponding to sciimg.fits
psfcat.fits - PSF-fit source catalog for sciimg.fits
sexcat.fits - aperture-based (SExtractor) photometry catalog for sciimg.fits
sciimgdao.psf - PSF for sciimg.fits in DAOPhot look-up-table format
sciimgdaopsfcent.fits - PSF stamp at center of sciimg.fits
scimrefdiffimg.fits.fz - difference image in fpack'd FITS format
diffimgpsf.fits - effective PSF stamp for difference-image
alerts.tar.gz - gzipped tar-directory of all alert packets
sciimlog.txt - processing log for sciimg.fits (instrumental calibration)
diffimlog.txt - processing log for difference image and event extraction
log.txt - overall processing log for realtime pipeline run

8.b. Reference Image Products

Reference images (co-added science images) have the following generic root paths and filenames in the archive:

```
prefield/field<field>/filtercode/<ccdid>/q<qid>/ztf_field_filtercode_<ccdid>_q<qid>_refimg.fits
```

where:
- `prefield` = first three digits of survey field ID
- `field` = six-digit survey field ID
- `filtercode` = zg, zr, or zi for g, r, i filters respectively
- `ccdid` = two-digit CCD identifier: 01..16:
- `qid` = one-digit quadrant ID within specific ccdid: 1..4

An archived "ref" image can be prepended with the following root URL to access and download the specific `refimg.fits` file:

https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/

For example:

https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/000/field000245/zg/ccd02/q1/ztf_000245_zg_c02_q1_refimg.fits

The path and filename identifiers for reference images covering a sky position can be obtained using an API query (see Section 12b.ii).

Alternatively, the `refimg.fits` suffix can be replaced with any of the following ancillary file suffixes, with products summarized in Section 3:

- `refcov.fits` - pixel depth-of-coverage image
- `refunc.fits` - pixel (1σ) uncertainty image
- `repsfcat.fits` - PSF-fit source catalog for refimg.fits
- `refsexcat.fits` - aperture-based (SExtractor) photometry catalog for refimg.fits
- `refimlog.txt` - processing log for sciimg.fits (instrumental calibration)
- `log.txt` - overall processing log for realtime pipeline run

8.c. Other File Products

Other relevant root URLs to access file-based products in DR4 are as follows.

The root URL to access raw CCD-based image files is: https://irsa.ipac.caltech.edu/ibe/data/ztf/products/raw/
The root URL to access calibration (cal) related files is: https://irsa.ipac.caltech.edu/ibe/data/ztf/products/cal/

For a description of all path and filename identifiers, see Section 7 of the ZSDS Explanatory Supplement.

9. Epoch-based Data Quality Flags and Filtering

The ZTF Observing System is fully robotic. One limitation is that it cannot determine in advance when conditions are non-photometric, for example, when intermittent clouds partially (or fully) cover the field-of-view during an exposure. Scattered moonlight can also wreak havoc. This severely affects the accuracy of the derived photometric calibration solutions. There is one
10. What you should know about the ZTF Release Products

The following are high-level advisories. Visit the link in Section 11 for a more detailed list.

1. DR4 contains products from both the Public and Private Surveys, derived from individual exposures covering the observation spans summarized in Section 2. All i-filter data was acquired under private survey time. All exposures in the public survey have an integration time of 30 seconds. Exposures in the private surveys have a mixture of integration times: 30, 60, 90, 120, 240, and 300 seconds, with the most common being 30 seconds.

2. The archive contains products with basic automated data-quality assurance applied during processing. The primary criterion for archiving a science image (along with its derived products) is that it must have passed astrometric and photometric calibration using thresholds on metrics tuned to yield potentially usable science products. "Bad" (unusable) data in the archive is still possible. This is data affected by clouds, low atmospheric transparency and spatial variations therein, moon contamination, bright source scattering artifacts, unmasked aircraft and artificial satellites, transient detector behavior including cross-talk, bad seeing or focus, and/or inaccurate telescope tracking. The archive metadata contains flags at both the image-level and epoch-level within individual lightcurves to identify and omit likely unusable data from queries (see Section 9).

3. Calibrated ZTF photometry can show biases of up to 0.025 mag for predominately bright sources with g, r, i magnitudes <~ 15.5 relative to source photometry in the Pan-STARRS1 (PS1) MeanObject DR1 Catalog. However, see note 17 below. These biases are present in both the PSF-fit-derived and aperture-derived catalogs. These biases are field, CCD-quadrant, and magnitude dependent. The lightcurve photometry (when retrieved through either the GUI, API, or tar balls) includes corrections for various biases, resulting in final photometric accuracies (with respect to PS1) of typically <~ 0.01 mag.

4. There is no separate searchable database storing source measurements from the single-exposure catalog files. These measurements can only be accessed by downloading the catalog files, which are in FITS-binary.
5. Single-exposure source photometry can be accessed at the source-level indirectly by retrieving entire lightcurves through either the GUI, API, or tar balls. These products also have various photometric corrections applied (note 3 above), while the single-exposure catalog files do not include these refinements.

6. Large spatial queries using the GUI-driven lightcurve service may take longer than expected depending on user-load. An alternative for retrieving lightcurves is to submit source positions to the lightcurve API. We recommend first querying the Objects Table to ensure your positions return data (however see notes 9, 10, and 11 below).

7. Lightcurves were constructed using exclusively measurements from the calibrated single-exposure PSF-fit-derived catalogs, not the aperture-based photometry catalogs.

8. The photometry catalogs, including lightcurve products, contain measurements of both point-like and extended objects. The PSF-fit photometry measurements are optimal for point-sources only. Concentric aperture photometry measurements are more optimal for extended objects.

9. Lightcurves were constructed using seed detections from the deeper (co-add) Reference Image PSF-fit catalogs. These "seed sources" were loaded into the searchable Objects Table where their photometry is ~ 2 to 2.5 magnitudes deeper than the single-exposure catalogs. Therefore, a majority of faint sources in the Objects Table (fainter than the single-exposure sensitivity limits) will not have lightcurves. Futhermore, objects with photometry hovering near the single-exposure detection limits will have incomplete temporal-coverage and therefore more sparsely sampled lightcurves.

10. Reference Images and their catalogs (loaded into the Objects Table) were generated by stacking single-exposures acquired from all science programs (public, private, and Caltech-time observations). These programs covered two overlapping tiling grids on the sky, a primary grid and a secondary grid, to mitigate gaps between CCDs on the focal plane. The single-exposures included in DR4 however (from which lightcurves were constructed), do not cover all survey fields from both grids. Therefore, expect to see missing lightcurves for many sources in the Object Table.

11. When querying the Objects Table to retrieve lightcurves, the nobsrel (and ngoodobsrel) metrics report the total (and total unflagged) number of lightcurve epochs included in DR4. For DR4, one caveat is that nobsrel = ngoodobsrel for all objects. This was due to a coding oversight. The total number of epochs, including unreleased epochs from all science programs (public, private, and Caltech-time observations) are indicated by the nobs and ngoodobs metrics. The latter will usually exceed nobsrel and ngoodobsrel respectively. All lightcurve-collapsed metrics in the Objects Table are computed using all (including unreleased) epochal data from all science programs.

12. The lightcurve photometry measurements (mag values) already have their epoch-dependent photometric zeropoints applied (magzp column in returned metadata). All photometry is in the native ZTF photometric system. No color corrections are applied (or rather, zero color in the AB system is assumed). To transform photometry onto the AB (PS1) system, you will require knowledge of the source color in that system. The exact color used during initial photometric calibration against PS1 is defined by column pcolor and the color coefficient to transform to the PS1 system is given by column clrcoeff. The procedure for applying color corrections is described in Section 10.1.1 of the ZSDS Explanatory Supplement.

13. In DR4, a difference image (file suffix *scimrefdiffimg.fits.fz; see Section 8a) will only accompany a single-exposure science image in the archive if a Reference Image corresponding to the same field/CCD-quadrant/filter existed on/before the single-exposure image was processed. Reference Image generation continued throughout the survey and image differencing could only be triggered if a Reference Image was present. Therefore, missing difference image products are more prevalent during the early phase of the survey.

14. The units of the pixel values in all image products are in Digital Numbers (DN). Methods on how to apply the photometric calibration solutions to your own measurements using these products are outlined in Section 10 of the ZSDS Explanatory Supplement.

15. When querying lightcurves using the GUI, the mjdtimestamps attached to the lightcurve measurements pertain to the start of each exposure while hjd and hmjd pertain to the middle, i.e., with 0.5*EXPTIME added to the start observation times, where EXPTIME represents the integration time (FITS header keyword) of the corresponding exposure. The date/timestamps attached to all file-based products however (i.e., values for the FITS header keywords: OBSJD, OBMJD, SHUTOPEN, DATE-OBS) pertain to the start of the exposure, not the middle.

16. Varying atmospheric transparency on the scale of an individual CCD-quadrant can have a significant impact on the quality of the i-filter fringe corrections. Fringe corrections were calibrated from high throughput images acquired under photometric conditions. Therefore, the presence of intermittent cloud cover, depending on its spatial variation, can lead to erroneous corrections and unusable images in the i-Filter.

17. Be warned that calibrated photometry extracted from image data acquired with CCD #06 after October 23, 2019
may show systematic residuals with respect to PS1 of up to 10%. This was traced to a camera tuning update. Future reprocessing will attempt to correct the biased photometry.

18. The lightcurve photometry measurements (mag values) are sometimes (but seldomly) accompanied by negative error estimates (magerr). These negative errors are meaningless and imply that PSF-fit solutions did not converge. We advise treating all measurements with negative errors as bogus in analyses.

11. Extended (detailed) Advisories and Cautionary Notes

A more detailed compilation of advisories and known caveats specific to each data product are given in: ZDS Advisories & Cautionary Notes

These notes are progressively updated as issues are resolved, new ones become known, or as we learn of new tips or software that could be of use to the community.

12. How to Access ZTF Data Products

Access to all ZTF data products is through the online (GUI-based) web-tools and API services of the NASA/IPAC Infrared Science Archive (IRSA) linked from: https://irsa.ipac.caltech.edu/Missions/ztf.html.

Below we provide links to specific GUIs and documentation describing how to access the archive products. Examples using the API services are also given.

12.a. Product-specific Access GUIs and Recipes

- To visualize, analyze, and retrieve any of the file-based products (single-exposure science or reference images, and/or their catalogs or other files), use the following GUI: https://irsa.ipac.caltech.edu/applications/ztf/. You can supply a sky position, a list of positions, or names of objects that can be resolved, including Solar System objects or an orbital ephemeris. The query returns the CCD-quadrant images touching these positions and metadata for additional filtering. An important metadata column you will want to filter is the INFOBITS value, primarily for "bad" quality data (see Section 9a above). The GUI allows selected-image previews and interactive analysis. Entire file products or cutouts on specific regions can be packaged and downloaded.

- Lightcurves can be retrieved, visualized, and analyzed using the GUI services through three steps. The third step below uses the Time Series Tool to visualize and analyze an individual lightcurve.
  1. First query the DR4 Objects Table containing sources extracted from Reference Images (co-added single-exposures): ZTF DR4 Objects. You can perform a spatial cone-search, box or polygon search, supply a list of search positions, as well as include filtering on any of the pre-computed lightcurve-collapsed metrics shown in the form. This query returns a list of the objects for which lightcurves potentially exist (subject to advisories 9, 10, 11 in Section 10). This list can be saved to a file and fed to the lightcuve API, or the results can be sent directly to the Time Series Tool (see next step).
  2. On the results page following the query in step 1, click on "To Time Series Tool" to retrieve the lightcurve(s) of your object(s). At this stage, you can click "Download Light Curve Table" to save the lightcurves for all objects to a single ASCII table, concatenated and distinguished by object ID (oid). Metadata is also written, which includes the catflags column for identifying and filtering bad observation epochs (see Section 9b above).
  3. Alternatively, given the list of object IDs that have lightcurve data in step 2, you can click on an individual object and send its lightcurve to the Time Series Tool. This tool allows interactive filtering on the lightcurve metadata (for example, omitting bad observation epochs by filtering the catflags value). This tool provides a visualization of the single-exposure images for neighboring lightcurve epochs, allows you to compute a periodogram and a period therefrom, and then plot a phase-folded lightcurve. There is also an option to supply your own period for the phase-folding.

Figure 5 shows a schematic of this workflow.
A custom tool for recovering and characterizing known Solar System Objects (SSOs) in ZTF data is the Moving Object Search Tool (MOST): [https://irsa.ipac.caltech.edu/applications/MOST/](https://irsa.ipac.caltech.edu/applications/MOST/). This is a precovery tool where you can specify a known SSO name, NAIF ID, ephemeris, or orbital elements, and the observation time range. The service returns the single-exposure science image filenames containing the SSO, associated archive metadata, orbit plots, refined ephemeris, and more orbital parameters. Images or cutouts (and/or associated photometry catalogs) can then be retrieved using the image-download API. For more information, see the full usage instructions for the MOST GUI. An API for the MOST tool is also available.

**Figure 5** - Schematic of the workflow for retrieving and analyzing lightcurves using the GUI services (see Section 12a). Click to enlarge.

### 12.b. Example Queries using the APIs

Application Programming Interfaces (APIs) accompany most of the GUI data-retrieval services. These APIs can be executed from within your own software to enable repetitive and/or bulk data downloads.

An overview for retrieving file-based products using APIs is given on the Image/Catalog API Page. API queries can include thresholding on any of the available archive image metadata parameters.

Examples of using APIs to retrieve file-based products and lightcurves are given below.

#### i. Querying Single-Exposure Science Image Products using the API

To retrieve file-based single-exposure science-image products that touch a fixed R.A.,Dec position, fall within an observation JD range, and are likely to be of "good" (usable) quality with INFOBITS < 33554432 (see Section 9a), first query their metadata using the `wget` utility, e.g.,

```
wget "https://irsa.ipac.caltech.edu/ibe/search/ztf/products/sci?
P0S=255.9302,11.8654&WHERE=obsjd>2458219.9678+AND+obsjd<2458228.8155+
AND+infobits<33554432" -O out.tbl
```

where all inputs reside on one line. Then use the output metadata table (out.tbl) contents to construct the image URL-paths/filenames. Using the metadata column names, a science-image will have the generic URL-path/filename:

```
https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/YYYY/MMDD/fracday/ztf_filefracday_000field_filtercode_ccdid_imgtypecode_qqid_sciimg.fits
```

where identifiers in **bold green** are the values of actual columns in the out.tbl table, and `YYYY` (year), `MM` (month), `DD` (day), and `fracday` (fractional time of day) can be extracted from `filefracday`. These images can then be retrieved using the `wget` utility. For the example above, the first metadata record in out.tbl would be retrieved using:

```
wget https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/2018/0411/467847/ztf_20180411467847_000535_zr_c11_o_q3_sciimg.fits
```

Any of the accompanying products can be retrieved by replacing `sciimg.fits` with another filename suffix (see Section 8a). Instead of retrieving the entire image, you can download a square cutout centered at a specific R.A.,Dec and specific size, and save the output to a file (optionally gzipped). For the example above:

```
wget "https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/2018/0411/467847/ztf_20180411467847_000535_zr_c11_o_q3_sciimg.fits?
center=255.8535,12.0503&size=60arcsec&gzip=false" -O cutout.fits
```
If the API fails to return a file (or its metadata), you can try navigating the archive tree directly to search for the specific product: https://irsa.ipac.caltech.edu/ibe/data/ztf/products/sci/. For guidance, listings of the DR4 single-exposure Science products are given in Section 8a. A description of the path and filename identifiers is given in Section 7 of the ZSDS Explanatory Supplement.

ii. Querying Reference Image Products using the API

To retrieve file-based Reference Image products that touch a fixed R.A.,Dec position in only the \( r \)-filter (\( fid=2 \)), first query their metadata using the `wget` utility, e.g.,

```bash
wget "https://irsa.ipac.caltech.edu/ibe/search/ztf/products/ref?
  POS=358,25.6&WHERE=fid=2" -O out.tbl
```

where all inputs reside on one line. Then use the output metadata table (out.tbl) contents to construct the image URL-paths/filenames. Using the metadata column names, a Reference Image will have the generic URL-path/filename:

```plaintext
https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/
  prefield/field/filtercode/ccd/aid/qid/
  ztf_field_filtercode_caid_qid_refimg.fits
```

where identifiers in **bold green** are the values of actual columns in the out.tbl table, and `prefield` is the first three (left-zero-padded) digits of the six-digit `field` value. These images can then be retrieved using the `wget` utility. For the example above, the first metadata record in out.tbl would be retrieved using:

```bash
wget https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/
  001/field001596/zr/ccd16/q2/ztf_001596_zr_c16_q2_refimg.fits
```

Any of the accompanying products can be retrieved by replacing `refimg.fits` with another filename suffix (see Section 8b). Instead of retrieving the entire image, you can download a square cutout centered at a specific R.A.,Dec and specific size, and save the output to a file (optionally gzipped). For the example above:

```bash
wget "https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/
  001/field001596/zr/ccd16/q2/ztf_001596_zr_c16_q2_refimg.fits?
  center=357.46871,26.00549&size=100arcsec&gzip=false" -O cutout.fits
```

If the API fails to return a file (or its metadata), you can try navigating the archive tree directly to search for the specific product: https://irsa.ipac.caltech.edu/ibe/data/ztf/products/ref/. For guidance, listings of the DR4 Reference Image products are given in Section 8b. A description of the path and filename identifiers is given in Section 7 of the ZSDS Explanatory Supplement.

iii. Querying Lightcurves using the API

The API functionality corresponding to the GUI-driven recipe for retrieving lightcurves is described on the Lightcurve Programming Interface Page. Numerous examples are given on this page. Below we provide a more generic example that includes most parameters of interest. You will always want to exclude observation epochs associated with bad/unusable data by masking `catflags` values equal to 32768, -- the "cloud-affected and/or moon-contamination" #flag (decimal bit 15); see Section 9b. Alternatively, you can exclude epochs with any non-zero bit by using the `catflags` mask value 65535 in the query below.

The following query returns a concatenated list of lightcurves for objects falling within a 10 arcsec (~0.0028°) radius from position R.A.,Dec = 255.9302°, 11.8654° in the \( r \)-filter, where each has \( \geq 3 \) epochs that fall within the MJD range 58194.0...58483.0 and all consist of only "good" (likely usable) measurements. Lightcurves are stored in the output ASCII table: out.tbl. The concatenated lightcurves are distinguished by their ObjectIDs therein (column `oid`).

```bash
wget "https://irsa.ipac.caltech.edu/cgi-bin/ZTF/nph_light_curves?
  POS=CIRCLE+255.9302+11.8654+0.0028&BANDNAME=r&NOBS_MIN=3&
  TIME=58194.0+58483.0&BAD_CATFLAGS_MASK=32768&FORMAT=ipac_table" -O out.tbl
```

All inputs reside on one line. A lightcurve (with common `oid` values) is represented by columns `mag ± magerr` versus `mjd` or `hjd`. The out.tbl also contains metadata to reconstruct the single-exposure science image path/file names that you can later download using the `wget` utility, or alternatively, cutouts on your position of interest (see Section 12b.i).
You can also upload the out.tbl file from above to the Time Series Tool to enable visualization, analysis, and period finding, analogous to the GUI-driven method described in Section 12a. Ensure you select "ztf" from the menu before uploading. Following your upload, ensure only one ObjectID (oid) is selected for visualizing, using the filter (📍) icon in the Time Series Tool.

### iv. Solar System Object Precovery using the API

The API equivalent of the Moving Object Search Tool (GUI) is described on the MOST Programming Interface Page. For example, to retrieve all single-exposure science images containing Asteroid 438973 Masci, first query their metadata using the `wget` utility:

```bash
wget "https://irsa.ipac.caltech.edu/cgi-bin/MOST/nph-most?catalog=ztf&input_type=name_input&obj_name=Masci&obs_begin=2018+03+17&obs_end=2018+12+31&output_mode=Brief" -O out.tbl
```

where all inputs reside on one line. Then use the output metadata table (out.tbl) contents to construct the science image URL-paths/filenames (see Section 12b.i above). These images can then be downloaded using the `wget` utility. Alternatively, cutouts centered on the object's R.A.,Dec in each image, as listed in out.tbl, can be downloaded.

### 12.c. Bulk Download of all Lightcurves

The lightcurves included in this release have been packaged into gzipped tar files. A selection of basic (most necessary) metadata accompany each lightcurve, primarily IDs to enable the retrieval of associated metadata from the archive, for example, images or Object Table metrics using the access methods above.

The following URL contains a listing of the tar-gzipped filenames and their (compressed) sizes:

https://irsa.ipac.caltech.edu/data/ZTF/lc_dr4/

When uncompressed, their total volume is ~ 9.7 TB. You can click on individual files to download them, or use the `wget` utility to automate the downloads. If your operating system does not automatically unzip these, you can use `gunzip file.tar.gz` followed by `tar -xvf file.tar` to unpack. A list of corresponding MD5 checksums is also provided for those who want to verify file contents following a download. On a Unix/Linux OS, the files can be verified using `md5sum -c checksums.md5`.

The files are labelled by survey-Field ID where each Field spans ≈ 7° x 7°. Each filename also includes the minimum to maximum RA and Dec values spanned by all the objects contained therein. These RA, Dec ranges can be used to identify the sky region(s) of interest in case you are interested in downloading a subset. The DR4 lightcurves are spread across 1029 fields.

Each file unpacks into an ASCII table with the following structure:

```plaintext
# ObjectID (oid), nDR4epochs, FilterID, FieldID, RcID, ObjRA, ObjDec
   HMJD, mag, magerr, clrcoeff, catflags
...
Click for example (truncated)
```

The lightcurves are delimited by lines: "# Object / lightcurve metadata ...". Following this line is the actual lightcurve for that object. The columns are as follows:
**ObjectID:** Unique Object Identifier. This is the *oid* that is returned by queries on the Objects Database Table. This ID can be used to retrieve additional metadata for an Object or lightcurve.

**nDR4epochs:** The number of observation epochs in the lightcurve for this object, included in this data release only.

**FilterID:** Filter for the lightcurve photometry: 1 => *g*; 2 => *r*; 3 => *i*.

**FieldID:** The survey Field Identifier. Lightcurves in DR4 are spread across 1020 Fields. This identifier can be used to retrieve additional archive metadata.

**RcID:** The readout-channel Identifier (or CCD-Quadrant index) that contains the object. The RcID runs from 0 to 63 over the ZTF focal plane. In combination with FieldID, this can be used to retrieve additional archive metadata.

**ObjRA:** Object Right Ascension from the input Reference Image (seed) catalog corresponding to the FieldID, RcID, and FilterID [decimal degrees].

**ObjDec:** Object Declination from the input Reference Image (seed) catalog corresponding to the FieldID, RcID, and FilterID [decimal degrees].

----------------

**HMJD:** Heliocentric-based Modified Julian Date [days] corresponding to the middle of each exposure and approximate position of the object. The HMJD is derived from "HJD - 2400000.5" where HJD is the Heliocentric Julian Date.

**mag:** Calibrated magnitude for a source with color *g* - *r* = 0 in the AB photometric system. See Cautionary Note #12 in Section 10.

**magerr:** Corresponding 1-σ uncertainty in *mag* estimate; excludes the (usually small) uncertainty associated with the calibration zeropoint and possible systematics relative to external photometry catalogs (for example PS1).

**clrcoeff:** Linear color coefficient term from photometric calibration that can be applied to the *mag* estimate if you know the AB *g* - *r* color of your source. See Cautionary Note #12 in Section 10.

**catflags:** Photometric/image quality flags encoded as bits (Section 9b).

In particular, you will always want to exclude observation epochs affected by clouds and/or the moon. These epochs have catflags = 32768 (decimal bit 15).

A utility to extract and parse the lightcurves in the unpacked ASCII tables is called *zort*. This tool enables custom filtering, the identification of lightcurves belonging to the same object across different filterIDs, and plotting.

### 13. Help Desk

For any issues related to data-access, GUI tools, APIs, data-quality, formats, processing algorithms or product usage, please email:

irsasupport@ipac.caltech.edu

### 14. Further Documentation & Relevant Publications

A detailed description of the processing pipelines, data products (including alert packets), instructions for product retrieval, filename identifiers, data formats, and product usage can be found in:

**The ZSDS Explanatory Supplement:** Pipelines, Definitions, Data Products & Access

Two relevant publications from the special ZTF Focus Issue of PASP:
An overview of the ZTF Science Data System:
*The Zwicky Transient Facility: Data Processing, Products, and Archive, 2019, vol 131 (issue 995), 18003*

An overview of the technical specifications of the camera and observing system, its performance, and ZTF survey design:
*The Zwicky Transient Facility: System Overview, Performance, and First Results, 2019, PASP, vol 131 (issue 995), 18002*

Last update: 2020 November 10