Overview of Catalina Sky Survey PDS Archive

R. Seaman[1], C. Neese[2], J. Stone, E. Christensen

15 September 2024

1 Introduction

The history, telescopes, and operations of the Catalina Sky Survey are discussed in separate documents. Here we will discuss the details of the archival holdings that have resulted from about two decades of nightly survey and follow-up observations. Readers interested in the broader history, data, and planning of the near-Earth asteroid community may find it at NASA's Planetary Defense Coordination Office[3] or the Minor Planet Center[4] and Center for NEO Studies[5] that PDCO funds.

2 Catalina Sky Survey Archive Bundle Structure

The basic structure of the Catalina Sky Survey PDS bundle is outlined in figure 1. The subdirectories listed are for CSS telescopes (giving their MPC site codes) currently operating. Availability of current and legacy data sets from each telescope depends on many best faith factors. The legacy holdings for the Siding Spring Survey, MPC code E12, will be ingested into the PDS Small Bodies Node at a later phase of the CSS archive reprocessing project. Data from future telescopes such as the Schulman telescope, MPC code G84, of the Mount Lemmon SkyCenter will be considered for inclusion in the SBN holdings of the Catalina Sky Survey as those projects evolve and are implemented. Precise details are yet to be settled.

Bundle	Collection	Subdirectory
Catalina	Data_Raw	703
		G96
		I52
		V00
		V06
	Data_Partially_Processed	703
		G96
		I52
		V00
		V06

[1] rseaman@arizona.edu, Lunar and Planetary Laboratory, 1629 E University Drive, Tucson, AZ 85721
[2] Planetary Science Institute, 1700 Fort Lowell Road, Suite 106, Tucson, AZ 85719
[3] https://www.nasa.gov/planetarydefense/overview
[4] https://minorplanetcenter.net
[5] https://cneos.jpl.nasa.gov

	Data_Calibrated	703
		G96
		I52
		V00
		V06
	Data_Derived	703
		G96
		I52
		V00
		V06
	Miscellaneous	
	Calibration	
	Document	

Figure 1 – PDS bundle design for Catalina Sky Survey data holdings.

2.1 Reprocessing of CSS nightly data sets

Images taken by Catalina Sky Survey telescopes are processed by data reduction pipelines that run on computers in the telescope domes. For remotely operated telescopes, these computers may reside at remote locations, but the nightly pipelines are all triggered with short latency immediately after the camera shutter closes. At the end of each night, data from each telescope take a convoluted but fully automated path from the telescopes to the University of Arizona campus and then on to the Small Bodies Node of the Planetary Data System.

CSS data (images and derived data products as described here) can also be archived verbatim into the SBN after the fact. This is also automated, generally in large batches running in parallel. A third option is to reprocess the original images to create new derived data products. Since the CSS pipeline has changed over the years, reprocessing is appropriate for data sets older than when some major feature was added. For CSS, this date has been selected to be January 1, 2020. All data sets older than this date have been reprocessed through the current pipeline.

There are no inherent implications for the user in using reprocessed versus non-reprocessed data. However, since astrometry submitted to the Minor Planet Center originated from the original pipeline data products, such reprocessed nights cannot easily be matched against the MPC catalog. (Re)submitting astrometry to the Minor Planet Center is not currently part of this project, though that may change. That being the case, some file types may be omitted from the SBN holdings and may have been missing from data sets that passed through earlier versions of the CSS pipeline in any event. These can include files related to generating astrometry with extensions *mpcd*, *neos*, *fail*, *rjct*, *detf*, *detl*, *detb*, *mrpt*, and *ades*. But also a variety of files that may not have been created or saved at the telescope originally, such as more recent JSON parameter files, logs, pointing files, etc. And in general, any telescope night may be missing optional files such as *pass1*, *avgs*, *avgr* that depend on different paths through the pipeline. For example, most of the “miscellaneous” files are optional.

3 Catalina Bundle Contents

In the tables below the subdirectory nomenclature is "tel" for the MPC code of the telescope and "yyyy" and "yyMmmdd" are the UTC of the next morning, eg, "G96/2020/20Nov13". See section 4 below for the file naming convention. The "#" column is the number of each data product to expect, though various pipeline exceptions can result in files being missing, i.e., most image classes are optional. The "#" code letters are:

- E = number of exposures in the night, typically 0-800
- F = number of fields observed, typically 0-200
- D = number of NEO discoveries, varies night to night
- G = variable numbers depending on weather and other factors

3.1 Catalina Bundle

Bundle	LID
Catalina	urn:nasa:pds:gbo.ast.catalina.survey

Product	Format	Description	#	Pipeline?
bundle_gbo.ast.catalina.survey.xml	XML	Bundle XML Label	1	no

3.2 Raw Data

Raw data for the Catalina Sky Survey are large numbers of short exposure FITS images from CCD (or perhaps in the future, CMOS) cameras. CSS survey cameras are widefield from about 1 square degree up to 19 square degrees. CSS follow-up cameras are typically a fraction of a square degree, though all cameras and telescopes can serve both functions, eg, a widefield survey telescope may be used for a targeted follow-up of a short-arc object whose location is highly uncertain. Survey exposures are taken at sidereal rates with trailing asteroids, while follow-up exposures are often at asteroid rates with trailing stars. Multiple exposures separated by seconds or minutes are used to detect moving objects.

Collection	LID
Data_Raw	urn:nasa:pds:gbo.ast.catalina.survey:data_raw

Product	Format	Description	#	Pipeline?
collection_data_raw.xml	XML	Collection XML Label	1	yes
collection_inventory_data_raw.csv	CSV	Collection inventory	1	yes

Subdirectory: **tel/yyyy/yyMmmdd**

Product *(filenames)*	Format	Description	#	Pipeline?
[filename].fits	FITS	Raw camera exposures	E	yes

Most cameras used by CSS have been single monolithic CCD detectors and the resulting raw FITS files have been single image format (SIF). By contrast, the 90Prime camera used at the Bok telescope (MPC code V00) is a 2x2 mosaic of CCDs and creates multi-extension format (MEF) FITS. Note that raw camera images have not always been archived by Catalina Sky Survey.

3.3 Partially Processed Data

The CSS pipeline produces a variety of partially processed data products that are input to all later algorithms. These may be directly useful for diverse purposes and also support later reprocessing. While reviewing these data products, it may be useful to consult the CSS Data Flow Diagram as discussed in the operations document. Raw images are first calibrated using the usual CCD techniques to produce the *calb* files. The CSS pipeline is iterative and sometimes files are persisted from the first pass (*pass1*) though these are usually omitted, and the CSS pipeline follows dual paths using both catalog and image differencing techniques (*csub*) which are later merged. CSS follow-up observations often rely on the NEO community's so-called "track and stack" image coaddition at both sidereal and object rates and in such cases there may be *avgs* and *avgr* files.

Collection	LID
Data_Partially_Processed	urn:nasa:pds:gbo.ast.catalina.survey:data_partially_processed

Product	Format	Description	#	Pipeline?
collection_data_partially_processed.xml	XML	Collection XML Label	1	yes
collection_inventory_data_ partially_processed.csv	CSV	Collection inventory	1	yes

Subdirectory: **tel/yyyy/yyMmmdd**

Product *(filenames)*	Format	Description	#	Pipeline?
[filename].calb	FITS	CCD calibrated	E	yes
[filename].pass1	FITS	Pass1 images	E	yes
[filename].csub	FITS	Difference images	E	yes
[filename].avgs	FITS	Track & Stack coadd, sidereal	E	yes
[filename].avgr	FITS	Track & Stack coadd, asteroid rate	E	yes

As with raw FITS files, CCD-calibrated (calb) images from the Bok (V00) 90Prime camera are multi-extension format (MEF) FITS. Other partially processed imaging data from CSS, including from V00, are single image format FITS.

3.4 Pipeline Calibrated Data

The pipeline-calibrated images supply the pixels for source detection and all that follows. Legacy data sets may begin with the calibrated *arch* FITS files because raw data were not retained in the early years of the survey. (Why calibrated CSS FITS data are called "arch" is an obscure fact lost to history.) Most archive users will find that the *arch* files are the best place to start, including their most complete set of image header metadata. CSS relies on the standard FITS FPACK compression format[6] and users may need to install this software to use CSS data. Partially processed data with FITS world coordinates for each image are astrometrically calibrated while photometric calibrations rely on the full pipeline to have completed in the *arch* files.

Collection	LID
Data_Calibrated	urn:nasa:pds:gbo.ast.catalina.survey:data_calibrated

Product	Format	Description	#	Pipeline?
collection_data_calibrated.xml	XML	Collection XML Label	1	yes
collection_inventory_data_calibrated.csv	CSV	Collection inventory	1	yes

Subdirectory:	**tel/yyyy/yyMmmdd**			
Product ***(filenames)***	**Format**	**Description**	**#**	**Pipeline?**
[filename].arch	FITS	Calibrated exposures	E	yes

Fully calibrated imaging data from CSS, whether originating from a single CCD detector or a mosaic camera like 90Prime at the Bok, are composited into a single two dimensional image to support further pipeline processing through source detection, astrometric solutions, moving object detection, and etc.

3.5 Derived Data Products

Derived data products are very diverse in the Catalina Sky Survey workflow. The basic strategy of the CSS pipeline is founded on building a catalog of point sources for each image using the third party Sextractor software. See the document CSS_operations.pdf for details. For FITS format files it may be helpful to consult the FITS standard[7] as well as derived standards such as the Leiden Data Analysis Center (LDAC)[8] format used by the binary Source Extractor catalogs. Various techniques correlate and classify the point source catalogs to identify candidate moving objects and match them against known asteroid ephemerides. Human validation and machine heuristics identify candidate Near-Earth Asteroids.

Collection	LID
Data_Derived	urn:nasa:pds:gbo.ast.catalina.survey:data_derived

Product	Format	Description	#	Pipeline?
collection_data_derived.xml	XML	Collection XML Label	1	yes
collection_inventory_data_derived.csv	CSV	Collection inventory	1	yes

[6] https://heasarc.gsfc.nasa.gov/fitsio/fpack/
[7] https://fits.gsfc.nasa.gov/fits_standard.html
[8] https://marvinweb.astro.uni-bonn.de/data_products/THELIWWW/LDAC/LDAC_concepts.html

Subdirectory:	**tel/yyyy/yyMmmdd**			
Product *(filenames)*	**Format**	**Description**	**#**	**Pipeline?**
[filename].sext	ASCII	Bright sources	E	yes
[filename].sexb	FITS (bintable)	Deep sources	E	yes
[filename].sexs	ASCII	rate coadd sext output	G	yes
[filename].mtds	ASCII	Moving Target Detections	F	yes
[filename].mtdf	ASCII	Annotated MTD objects	F	yes
[filename].dets	ASCII	Asteroid detections	F	yes
[filename].tran	ASCII	Transient sources	E	yes
[filename].tran_ai	ASCII	AI-scored transient sources	E	yes
[filename].rjct	ASCII	Rejects	F	yes
[filename].mpcd	ASCII	MPC batch	F	yes
[filename].neos	ASCII	NEOs and NEO candidates	D	yes
[filename].fail	ASCII	Failed astrometry	D	yes
[filename].tssexb	FITS (bintable)	Track & Stack sextractor output	G	yes
[filename].avgrsexb	FITS (bintable)	Rate coadd sexb output	G	yes
[filename].detf	FITS (bintable)	merge detection metadata	F	yes

As with other automated astronomical pipelines, Catalina Sky Survey's nightly workflow converts raw input images into a diversity of derived data products. These are described in more detail in the CSS Operations and Processing document. To enable the user to understand their meaning and use, CSS files fall into a smaller number of general classes:

- Metadata about point source detections (extensions **sexb**, **sext**, **sexs**, **tssexb**, **avgrsexb**). In the current CSS pipeline these are produced by the third-party SExtractor[9] software from input data at either sidereal or asteroid rates and either single snapshots of the sky or coadded stacks. The software originally creates a FITS binary table format, but this is later converted to tabular ASCII and directly annotated by CSS software to add columns.

- Data are also filtered, merged, and extended into per-image files of transient sources (extension **tran**) that are available for subsequent linking with similar format files from other observatories. These sources are scored (extension **tran_ai**) for artifacts (diffraction spikes, etc.) using a PyTorch transient detection AI classification model. These scores are used by CSS for tuning our pipeline for NEO candidates presented to the observers for validation, and are also available to third parties for linking multi-telescope tracklets across multiple nights, for instance.

- Metadata about moving target detections (files with extensions **mtds** and **mtdf**). The heart of the beast that converts individual point source detections into candidate moving object detections, or *tracklets*. Point sources correspond to individual images of the sky (or coadded equivalents), while tracklets represent multiple exposures designed to detect and measure motion.

[9] https://www.astromatic.net/software/sextractor/

- Metadata about machine and human validated tracklets (files with extensions **dets**, **rjct**, **fail**). To find each dot or streak of light that has reflected off a small body in the solar system, every survey and follow-up project must sort through many thousands of background stars and galaxies, imaging artifacts, and foreground artificial satellites. Many plausible false tracklets are processed for each actual moving object detection. Software limits which of these are presented to the human observers for validation.

- Metadata about new NEO candidates and NEO follow-up observations (**neos** extension) versus metadata about incidental known objects and unknown "low digest2" detections (**mpcd** extension). Many thousands of previously discovered main belt asteroids are redetected for each new near-Earth asteroid.

- Metadata merged (**detf** – and **detl** and **detb** in the miscellaneous collection) from various of the above files, as well as from the processed images, to enable generation of astrometric submissions to the Minor Planet Center, described below.

3.6 Miscellaneous

A variety of documents round out the bundle to capture input parameters, pointing information, survey coverage, MD5 checksums, etc. The collection of miscellaneous derived data products includes a variety of parameter file inputs and log files as well as intermediate file types that are later merged into tabular structures in the section above. This class also includes logistically useful files such as text FITS headers for the arch files.

Collection	LID
Miscellaneous	urn:nasa:pds:gbo.ast.catalina.survey:miscellaneous

Product	Format	Description	#	Pipeline?
collection_miscellaneous.xml	XML	Collection XML Label	1	yes
collection_inventory_miscellaneous.csv	CSV	Collection inventory	1	yes

Subdirectory: **tel/yyyy/yyMmmdd**

Product *(filenames)*	Format	Description	#	Pipeline?
[filename].detl	ASCII	input detection list	D	yes
[filename].detb	FITS (bintable)	Per-batch or per-tracklet metadata	D	yes
[filename].iext	ASCII	Sources from difference imaging	E	yes
[filename].strp	ASCII	Field catalog	E	yes
[filename].strm	ASCII	Field catalog	E	yes
[filename].scmp	ASCII	SCAMP output	E	yes
[filename].ephm	ASCII	Ephemeris	F	yes
[filename].achk	ASCII	Ephemeris from astcheck	F	yes
[filename].hits	ASCII	Hits	F	yes
[filename].arch_h	ASCII	FITS header card images (no newlines)	E	yes
[filename].ast	ASCII	Astrometry	D	yes

[filename].mrpt	ASCII	IAU 80-column astrometry submission	D	yes
[filename].ades	XML	ADES astrometry submission	D	yes
[filename].focheck	ASCII/JSON	Astrometry orbital fit check	D	yes
[filename].param	ASCII	Parameters	G	no
[filename].params	ASCII	Input parameters	G	no
[filename].outparams	ASCII	Output parameters	G	no
[filename].json	ASCII	JSON	D	no
[filename].log	ASCII	Logs	D	no
[filename].point	ASCII	Nightly telescope pointing	1	no
[filename].pt	JSON	Per-image pointing (MPC format)	E	yes
[filename].xmls	ASCII	SCAMP params	E	no
[filename].cov	ASCII	Survey coverage	1	no
model.*.csv	ASCII/CSV	AI training set for PyTorch transient detection classification model	1	no
model.*.pth	Binary	Serialized AI PyTorch state dictionary, including the model's weights, biases, and other parameters	1	no
ai_config.json	JSON	AI configuration parameters for PyTorch transient classification model	1	no
MPCORB.DAT	ASCII	Orbit catalog from Minor Planet Center, updated daily	1	no
mpcorb.sof	ASCII	Orbit catalog in Standard Orbit Format, updated daily	1	no
ELEMENTS.COMET	ASCII	Orbital elements for comets from NASA's Jet Propulsion Lab	1	no
ObsCodes.html	ASCII	List of observatory codes from the Minor Planet Center	1	no
[filename].txt	ASCII	Various utility files (eg, software versions, survey and follow-up planning)	G	no
signature.md5	ASCII	MD5 file hashes	1	no

One can infer by the nature of the word "miscellaneous" that these data products are very diverse. First, to finish the list from section 3.5, the primary goal of a NEO survey is to convert pixel data into astrometric data for submission to the Minor Planet Center:

- Astrometric data represented in the two IAU-standard formats (**ades** and **mrpt** extensions). ADES[10] is the newer format with additional metadata in XML format. MRPT files are in the old 80-column[11] IAU format.

[10] https://github.com/IAU-ADES/ADES-Master/blob/master/ADES_Description.pdf
[11] https://www.minorplanetcenter.net/iau/info/OpticalObs.html

- Files with **focheck** extensions are a simple check that the astrometry links with reasonably small residuals to the target object's orbit using the community find_orb[12] orbit fitting software. The format of this file is the IAU 80-column tracklet concatenated with JSON-format output of detailed parameters of the find_orb fit.

- Metadata about point sources detected through the difference imaging branch of the CSS pipeline (**iext** extension). These are formatted similarly, but not identically, to the sext files.

- Per-image stellar lists and photometry (extensions **strm** and **strp**).

- Input and output for SCAMP[13] software (extension **xmls** and **scmp**) which converts point source detections into astrometric and photometric measurements. The CSS pipeline then further processes these values.

- Files related to telescope and survey operations (extensions **point**, **pt**, **cov**, **followup.txt**, **userfields.txt**, **survey*.txt**).

- Files related to an AI PyTorch[14] model trained to classify transient source detections (**ai_config.json**, **model*.csv**, **model*.pth**). These are used to tune the CSS pipeline to exclude artifacts like diffraction spikes, etc., when presenting NEO candidates to the observers for validation.

- Community resource files current as of the individual night at the telescope (or for legacy data, of the date of pipeline reprocessing the night's images). These include: **MPCORB.DAT**[15] and **ObsCodes.html**[16] from the Minor Planet Center and **ELEMENTS.COMET**[17] from the NASA Jet Propulsion Laboratory, as well as orbital elements in Project Pluto's Standard Orbit Format in the **mpcorb.sof**[18] file.

- FITS headers in 80-column card image format without newlines (**arch_h**).

- Per-image asteroid ephemerides (**ephm**) and per-field identifications against these known objects (**hits**). The **achk** files, if present, are extended ephemerides generated using Project Pluto's astcheck[19] software.

- Unidentified incidental (low digest2) IAU 80-column tracklets (extension **ast**).

---

[12] https://www.projectpluto.com/find_orb.htm
[13] https://www.astromatic.net/software/scamp/
[14] https://pytorch.org
[15] https://minorplanetcenter.net/iau/MPCORB.html
[16] https://minorplanetcenter.net/data
[17] https://ssd.jpl.nasa.gov/dat/ELEMENTS.COMET
[18] https://www.projectpluto.com/orb_form.htm
[19] https://www.projectpluto.com/astcheck.htm

- Input and output parameters of various software (file extensions **param**, **params**, **outparams**, **json**).

- Log files generated by various software processes (extension **log**).

3.7 Calibrations

Catalina Sky Survey CCD data is reduced in a similar fashion to other CCD imaging data, but NEO survey data in general are generally distinguished by large pixels. CSS relies on pragmatic and efficient logistics above almost everything else. CSS relies on pragmatic focusing, queue-based data-taking, flat fields selected exposure-by-exposure from a flat field library, and anything that streamlines operations. Flat fields that were used on a given night are copied from the library into the nightly data directory.

Collection	LID
CALIBRATION	urn:nasa:pds:gbo.ast.catalina.survey:calibration

Product	Format	Description	#	Pipeline?
collection_calibration.xml	XML	Collection XML Label	1	yes
collection_inventory_calibration.csv	CSV	Collection inventory	1	yes

Subdirectory: **tel/yyyy/yyMmmdd**

Product *(filenames)*	Format	Description	#	Pipeline?
flat.fits	FITS	Flat field images	varies	yes

Catalina's flat field images have the same dimensionality as the raw FITS files. For most CCD cameras these are single image format FITS, while for mosaic detectors such as 90Prime at the Bok telescope these are multi-extension format (MEF) FITS.

3.8 Documents

Collection	LID
DOCUMENT	urn:nasa:pds:gbo.ast.catalina.survey:document

Product	Format	Description	#	Pipeline?
collection_document.xml	XML	Collection XML Label	1	no
collection_inventory_document.csv	CSV	Collection inventory	1	no

Product *(filenames)*	Format	Description	#	Pipeline?
CSS_bundle_overview.pdf	PDF/A	This document	1	No
CSS_history.pdf	PDF/A	History of the Catalina Sky Survey	1	No
CSS_telescopes.pdf	PDF/A	CSS Telescopes and Instrumentation	1	No
CSS_operations.pdf	PDF/A	CSS Operations and Processing	1	No

4 Catalina Sky Survey Documents

- *Overview of Catalina Sky Survey PDS Archive*, Seaman, Neese, Stone, and Christensen, 13 Nov 2020 (this document)
- *History of the Catalina Sky Survey*, CSS team, 12 Nov 2020
- *Catalina Sky Survey Telescopes and Instrumentation,* CSS team, 12 Nov 2020
- *Catalina Sky Survey Operations and Processing*, CSS team, 12 Nov 2020
- *Catalina Sky Survey Data Flow Diagram,* CSS team, 15 Aug 2021

5 CSS Data Holdings and File Naming

The Catalina Sky Survey (CSS) acquires optical imaging data of the night sky of both standardized survey fields and targeted follow-up to support the discovery of Near-Earth Objects (NEOs) as well as incidental astrometric observations of asteroids, comets, and other solar system objects. Fields are imaged multiple times (almost always four times) in support of moving object detection against the background stars. All subsequent data products, and their names, derive from four initial raw CCD images (see below for explanation of sub-fields):

```
G96_20200927_2B_N20001_01_0001.fits.fz[20]
G96_20200927_2B_N20001_01_0002.fits.fz
G96_20200927_2B_N20001_01_0003.fits.fz
G96_20200927_2B_N20001_01_0004.fits.fz
```

In survey mode these are typically 30 second exposures, acquired about 7 minutes apart, interleaved with observations of other adjoining fields. Follow-up observations vary in exposure time depending on the expected brightness of the targets and are often taken in rapid sequence.

An alternate observing mode for follow-up is to “track and stack” in which each of the four frames is composed of separate frames that are coadded at the expected rate of motion of the targeted object. In that case there will be additional raw images – two dozen or more are not unusual. Several of the file types discussed below will only appear for track and stack observations.

Archival holdings at CSS’s University of Arizona headquarters are universally in a directory structure: /archive/TEL/YYYY/NIGHT[21], where TEL is one of the MPC codes listed below, YYYY is the calendar year, and NIGHT is of the form: YYMmmDD, for example:

```
/archive/G96/2020/20Sep27
```

[20] FITS images are FPACK compressed at the telescope during the night. This leaves the headers readable.
[21] In practice, annual datasets are hosted in ZFS pools in a growing mass datastore that is replicated between the CSS offices at the Lunar and Planetary Laboratory and the campus data center, UITS.

CSS image names and derived data products within these directories are comprised of several segments separated by underscores with an appended extension:

- Telescope – three letter MPC[22] site code
- Night – UT date at the end of the night, YYYYMMDD
- Binning – CCD pixel binning, eg, 1B, 2B, or 4B
- Field – field ID, Xnnnnn where X is N, S, F, or U
- Repetition – 1 for the first time a field is imaged per night, incremented if repeated
- Sequence – sequence number of exposures, 1-4 (or more for track-and-stack)
- Extension – "fits" for the raw data or many other options (see below)
- Compression – "fz" for FITS tile compression, "gz" for gzip, or absent (optional)

The enumerated CSS MPC codes are our four survey telescopes:

- 703 – the original Catalina Sky Survey (CSS) 0.7m Schmidt on Mount Bigelow
- G96 – the Mount Lemmon Survey (MLS) 60" telescope
- E12 – the Siding Spring Survey (SSS) using the 0.5m Uppsala Schmidt (retired in 2013)
- V00 – the Steward Observatory 90" Bok telescope on Kitt Peak, when used for NEO surveying in partnership with Spacewatch and the University of Minnesota

And three follow-up[23] telescopes:

- I52 – 40" follow-up telescope next to G96 on Mount Lemmon
- V06 – the 61" Kuiper telescope on Mount Bigelow, when used by CSS for follow-up
- G84 – the 32" Schulman telescope on Mount Lemmon, when used by CSS for follow-up

CSS survey field nomenclature is N for fields north of the celestial equator and S for southern fields. The first two digits are rounded to a grid in degrees of declination. The final three digits are a grid in degrees of right ascension. Each telescope has a different set of standard fields depending on their field of view. For example, the Schmidt has a larger field of view so there are fewer standard field centers.

Follow-up nomenclature has evolved and for most purposes can be regarded as "F" (for follow-up) and "U" (for user) followed by 5 random but unique characters. The distinction between follow-up and user fields has largely vanished, and currently the 5 characters are assigned as a subset of the temporary object designation as an aid to the observers. Temporary designations have also evolved, sometimes rapidly, and can be best regarded as random unique strings, generally 7 characters, assigned by the different NEO surveys using various schemes.[24]

There is a lengthy list of CSS file extensions, generated for a variety of purposes leading ultimately to submitting asteroid and comet astrometry to the IAU Minor Planet Center. These

---

[22] The IAU Minor Planet Center, https://minorplanetcenter.net

[23] All CSS telescopes can operate in both survey and follow-up mode, the division of labor depends on several factors, most critically the field of view of the camera. NEO survey cameras are very wide field.

[24] CSS designations currently begin with "C" followed by a 5-character high-order base running number and a trailing digit indicating the telescope: 1 for 703, 2 for G96, and so forth. These are arbitrary 7-character strings for the purposes of this document.

may or may not be encountered in gzip compressed form (with an appended ".gz" extension) or FITS tile compressed (FPACK) format (with a ".fz").

Ext	***Format***	***Collection***	***Description***
fits	FITS	*Data_Raw*	Raw CCD exposures (not for legacy data)
calb	FITS	*Data_Partially_Processed*	Basic CCD calibrated
pass1	FITS	*Data_Partially_Processed*	Pipeline pass 1 images (rare)
csub	FITS	*Data_Partially_Processed*	Difference images
avgs	FITS	*Data_Partially_Processed*	Sidereal track-and-stack coadds
avgr	FITS	*Data_Partially_Processed*	Asteroid rate T-and-S coadds
arch	FITS	*Data_Calibrated*	Fully calibrated images
pt	JSON	*Miscellaneous*	MPC JSON format pointing metadata
sext	ASCII	*Data_Derived*	Bright sextractor[25] source lists
sexb	bintable[26]	*Data_Derived*	Deep sextractor source lists (binary)
iext	ASCII	*Miscellaneous*	Sources from difference imaging
strp	ASCII	*Miscellaneous*	Field catalog list
strm	ASCII	*Miscellaneous*	Field matches
ephm	ASCII	*Miscellaneous*	Ephemeris predictions for field
achk	ASCII	*Miscellaneous*	Ephemeris predictions using astcheck
mtds	ASCII	*Data_Derived*	Moving Target Detection sources
mtdf	ASCII	*Data_Derived*	Filtered and annotated MTD list
dets	ASCII	*Data_Derived*	Candidate asteroid detections
hits	ASCII	*Miscellaneous*	Hits of detections against ephemeris
rjct	ASCII	*Data_Derived*	Objects rejected during validation
mpcd	ASCII	*Miscellaneous*	Non-NEO asteroid astrometry
neos	ASCII	*Miscellaneous*	NEO asteroid astrometry
fail	ASCII	*Data_Derived*	Failed astrometry (various reasons)
arch_h	ASCII	*Miscellaneous*	FITS header card images (no newlines)
tssexb	bintable	*Miscellaneous*	T-and-S sextractor output
avgrsexb	bintable	*Miscellaneous*	Rate coadd sextractor output (binary)
sexs	ASCII	*Miscellaneous*	Rate coadd sextractor output (text)
detf	bintable	*Data_Derived*	Merged per-detection metadata
tran	ASCII	*Data_Derived*	Transient source detections
tran_ai	ASCII	*Data_Derived*	AI-scored transient sources

Files with the above extensions will have names meeting the criteria in the first paragraph, but not all file extensions are created for all four (or more) raw FITS images.[27] Per-image processing includes image calibration and source extraction, while per-field processing applies to steps after

[25] https://www.astromatic.net/software/sextractor
[26] FITS binary table format. These will not be compressed.
[27] See bundle spreadsheet.

the initial moving target detection, for example the merged per-detection metadata results in one FITS binary table per four raw input images:

```
G96_20200927_2B_N20001_01_0001.detf
```

In addition, five of the file types from the Miscellaneous collection have slight name variations:

Ext	*Format*	*Collection*	*Description*
scmp	ASCII	*Miscellaneous*	SCAMP output
detl	ASCII	*Miscellaneous*	Input detection list for ADES generation
detb	bintable	*Miscellaneous*	Batched per-detection metadata
mrpt	ASCII	*Miscellaneous*	Old (“80-column”) astrometry submission
ades	XML	*Miscellaneous*	ADES astrometry submissions

The CSS pipeline relies on SCAMP (Software for Calibrating AstroMetry and Photometry)[28] to compute FITS world coordinates to each image individually, and as needed to coadded images. SCAMP also is the first step in CSS photometric calibration. SCAMP appends an “_1” to input files names since CSS images have a single FITS extension.

```
G96_20200927_2B_N20001_01_0001_1.scmp.gz
```

The final four files are only created by the pipeline if NEOs are discovered (or recovered) in a field. In this case the detection number (out of many candidates) is appended along with “neos” since this detection comes from the human-validated NEO astrometry in the neos file:

```
G96_20200927_2B_N20001_01_0001.29.neos.detl
G96_20200927_2B_N20001_01_0001.29.neos.detb
G96_20200927_2B_N20001_01_0001.29.neos.mrpt
G96_20200927_2B_N20001_01_0001.29.neos.ades
```

At the very end of the night up to two batches[29] (sometimes very large) are created of “incidental” astrometry (automated discovery of asteroids or comets). In this case the naming scheme is truncated because the discoveries / recoveries pertain to the night as a whole:

```
G96_20200927_2B.mpcd.detl     G96_20200927_2B.neos.detl
G96_20200927_2B.mpcd.detb     G96_20200927_2B.neos.detb
G96_20200927_2B.mpcd.mrpt     G96_20200927_2B.neos.mrpt
G96_20200927_2B.mpcd.ades     G96_20200927_2B.neos.ades
```

In addition to the per-image, per-field, and per-detection files described above, there are a variety of mostly text documents that describe the nightly observing and data processing as a whole. The number of documents has varied over the years. On the other hand, CSS image calibrations are

[28] https://www.astromatic.net/software/scamp

[29] The incidental neos and/or mpcd files are only generated if each such kind of detections were made. There are various reasons for both survey and follow-up mode why this may not happen.

simplified to be as efficient as possible to carry out and flat fields will generally be the only calibration files archived.

Ext	***Format***	***Collection***	***Description***
fits	FITS	*Data_Calibration*	Flat fields
param	ASCII	*Miscellaneous*	Pipeline and other parameters
params	ASCII	*Miscellaneous*	Input parameters
outparams	ASCII	*Miscellaneous*	Output parameters
json	ASCII	*Miscellaneous*	JSON parameters
log	ASCII	*Miscellaneous*	Processing logs
point	ASCII	*Miscellaneous*	Nightly telescope pointings
xmls	ASCII	*Miscellaneous*	SCAMP parameters
cov	ASCII	*Miscellaneous*	Survey coverage
md5	ASCII	*Miscellaneous*	MD5 file hashes