



Kepler: A Search for Terrestrial Planets

Kepler Archive Manual

KDMC-10008-006

Written by

Susan E. Thompson, Dorothy Fraquelli, Jeffrey E. van Cleve and Douglas A. Caldwell

May 2, 2016



**Space Telescope Science Institute
3700 San Martin Drive
Baltimore, MD 21218**

Prepared by: Susan E. Thompson Date 4 May 2016
Susan E. Thompson, Kepler Science Office

Prepared by: Dorothy Fraquelli Date 4 May 2016
Dorothy Fraquelli, MAST Archive Scientist

Approved by: Faith Abney Date May 4, 2016
Faith Abney, DMC and Data Processing and Archive Services Branch Manager

Approved by: Dwight Sanderfer Date 5/9/16
Dwight Sanderfer, Kepler Science Operation Center Manager

Approved by: Michael R. Haas Date 5/8/16
Michael R. Haas, Kepler Science Office Director

Approved by: Steve B. Howell Date 4 May '16
Steve B. Howell, Kepler Project Scientist

Document Change Log

Change Number	Change Date	Pages/Sections	Changes/Notes
001	Oct 17, 2011	Section 2.3.3	Description of FFI WCS coordinates
002	Oct 17, 2011	Section 2.3.4	Description of the Cotrending Basis Vectors
003	Oct 17, 2011	Section 2.3.5	Description of Pixel Response Functions
004	Oct 17, 2011	Section 2.2.1	How to Search for Custom Aperture Targets
005	Oct 17, 2011	Section 2.3.1.1	PDC-MAP and how it affects the PDC light curve, including keywords in the headers.
006	Oct 17, 2011	Table 2-3	Quality Flag updates
007	Oct 17, 2011	Section 1.10	Acknowledging Kepler in Papers
008	Oct 17, 2011	Section 2.3.1	Aperture Extension changes for light curve files.
009	Oct 17, 2011	Section 2.4	New keywords added to the keyword dictionary.
010	Oct 17, 2011	Section 2.3.3	Description of Uncertainties FFI
011	Oct 17, 2011	Table 3-2, Section 3.2.1	Data Tables now include Crowding, Flux Fraction and CDPP values.
012	Oct 17, 2011	p. 3	Added Change Log
013	Oct 21, 2011	Section D.2	Updated Acronyms
014	Oct 21, 2011	Appendix C2	Added FITS keywords for PRFs and CBVs
015	Mar 13, 2012	Section 2.3.6	Added Background FITS file description
016	Mar 13, 2012	Section 2.3.7	Added Collateral FITS file description
017	Mar 13, 2012	Section 2.3.8	Added ARP FITS file description
018	Mar 13, 2012	Table 2-1	Added Background, Collateral and ARP files
019	Apr 9, 2012	Section 3.3.5	Added CDPP retrieval section
020	Apr 9, 2012	Section 2.3.4	Removed details about CBVs
021	Apr 11, 2012	Section 3.3.3	Added Section on Enhanced Kepler Searches
022	Apr 9, 2012	Section 3.3.4	Updated CasJobs Section
023	Apr 9, 2012	Table 1-1	Revise data release date
024	Apr 16, 2012	Section 1.8	Update data release text
025	Apr 16, 2012	p. 4	File Version Table

for KDMC-10008-005

026	Apr 16, 2013	Table 2-2	File Version Table
027	Apr 16, 2013	Section 2.4	Added PDC related Keywords
028	Apr 16, 2013	Section 2.3.1.1	Updated description of PDC
029	Apr 16, 2013	Table 2-3	Added Quality flags
030	Apr 17, 2013	Section A.1b	Updated FITS keyword list
031	Apr 17, 2013	Section D.2	Updated acronym list
032	May 28, 2013	Section 2.3.9	Reverse Clock Added
033	May 28, 2013	Section 3	Updates to MAST Tools
034	May 28, 2013	Section— Changing File Formats	Removed
035	Apr 23, 2014	Section – Registering as User	Removed
036	Apr 23, 2014	Section – Data Rel.	Removed

for KDMC-10008-006

Change Number	Change Date	Pages/Sections	Changes/Notes
037	Jan 8, 2016	Signatures Page	Added/Changed Signatures
038	Jan 8, 2016	Versions Table	Updated Version Numbers
039	Jan 8, 2016	Introduction	Information about K2 and reaction wheel failures
040	Jan 11, 2016	Section 2.3.1.1	Added a Quality Flags section
041	Jan 11, 2016	Section 2.3.1.2	More information on the crowding correction as relates to PDC
042	Jan 11, 2016	Section 2.3.2	Added column RB_LEVEL description
043	Jan 12, 2016	Section 2.3.7	Added rolling band Information
044	Jan 20, 2016	Table 2-2	Added focal plane characterization version number
045	Feb 22, 2016	Section 2.3.5	FC models added to this section
046	Feb 22, 2016	Chapter 3	Chapter Updated to include Portal
047	March 25, 2016	Section 2.3.3	Added information about Golden FFIs

File Version Numbers Described by this version of the Archive Manual

File Type	Version *	Archive Manual Section
Light Curves Files	6.1	2.3.1
Target Pixel Files	6.1	2.3.2
Full Frame Images	3.1	2.3.3
Background	3.0	2.3.6
Collateral/ARP	2.0	2.3.7 - 2.3.8
CBVs	4.0	2.3.4
Focal Plane Characterization	1.0	2.3.5

*See the keyword FILEEVER in the header of each file.

*Earlier versions of the data files are described by previous Kepler Archive Manuals, however we strongly suggest using the most recent version of the *Kepler* data products when possible.

Table of Contents

Document Change Log	3
Table of Contents	6
Chapter 1 Introduction.....	8
1.1 Overview of Kepler	8
1.2 Overview of Data Flow	9
1.3 Related Documents.....	10
1.4 Overview of MAST.....	10
1.5 User Support Services.....	11
1.6 Getting Your Data	11
1.7 Publication Acknowledgement.....	12
Chapter 2 Kepler Data Products.....	13
2.1 Introduction	13
2.1.1 Overview of Data in the Archive.....	13
2.1.2 Kepler Time System.....	13
2.1.3 FITS headers.....	14
2.2 File Name Syntax	14
2.2.1 Custom Aperture Targets	15
2.3 Data in the Archive.....	16
2.3.1 Light Curve Files.....	16
2.3.2 Target Pixel Data	22
2.3.3 Full Frame Image.....	24
2.3.4 Co-trending Basis Vectors	25
2.3.5 Focal Plane Characterization Models	26
2.3.6 Background Data Files	47
2.3.7 Collateral Data Files	48
2.3.8 Artifact Removal Pixel Data Files	51
2.3.9 Reverse Clock Data Files	52
2.3.10 Ancillary Engineering Data	52
2.4 Keyword Definitions	52
Chapter 3 Searching the Archive for Kepler Data	55
3.1 The MAST Kepler Catalogs	55
3.2 The MAST Home Page.....	57
3.3 MAST Discovery Portal (aka Portal).....	59
3.4 MAST Classic Interface	62
3.5 Stellar Parameter Tables.....	71
3.6 FFI and Engineering Data Search and Retrieval or Download	71
3.7 The CasJobs Implementation.....	74
3.8 Other useful tools at MAST.....	75
3.8.1 Kepler Object of Interest (KOI) Search Form.....	75
3.8.2 Confirmed Planets Search Forms.....	75
3.8.3 Eclipsing Binaries	76
3.8.4 The Kepler-GALEX Crossmatch.....	76
3.8.5 Retrieving CDPP values.....	77
3.9 Alternate Methods for Downloading Data.....	78
References	80

Appendices	82
A.1: Kepler Light Curve File Headers.....	83
A.2: Target Pixel File Headers.....	90
B.1 Calibrated Full Frame Image (FFI) Primary Header	101
B.2 Calibrated FFI Channel Header	102
B.3 FFI Uncertainties Primary Header.....	104
B.4 Uncertainties FFI Channel Header	105
B.5 Ancillary Data Description	108
C.1 Cotrending Basis Vector Headers	113
C.2 Pixel Response Function Headers	113
D.1 Glossary of Terms	115
D.2 Common Acronyms.....	118

Chapter 1 Introduction

Data from the *Kepler* mission are housed in the Archive at the Space Telescope Science Institute (STScI) and accessed through MAST (Mikulski Archive for Space Telescope). General information about *Kepler* may be found at the *Kepler* Mission <http://kepler.nasa.gov/> and the *Kepler* for astrophysicists' <http://keplerscience.arc.nasa.gov/> web sites.

1.1 Overview of Kepler

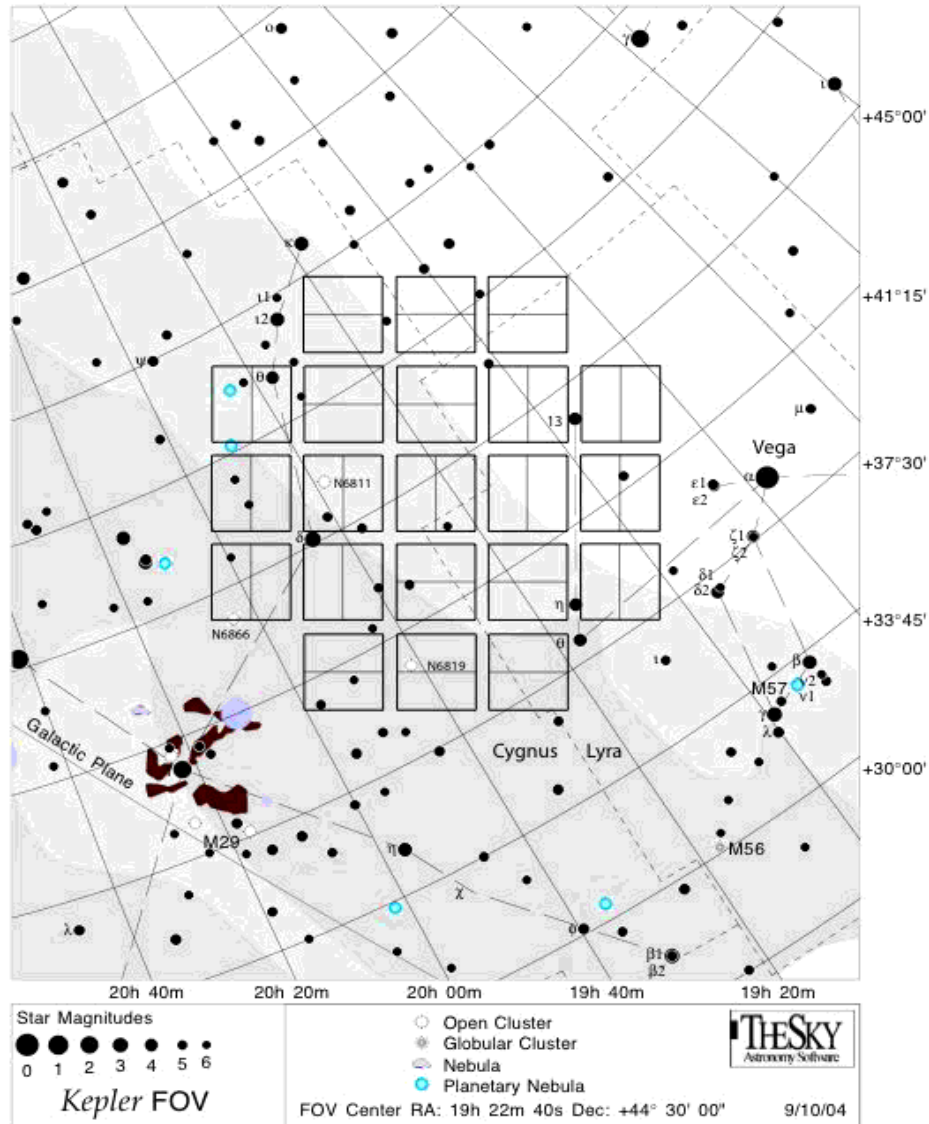
The *Kepler* mission was designed to survey a region of the Milky Way galaxy to detect and characterize Earth-size and smaller planets in or near the habitable zone by using the transit method of planetary detection. This was accomplished by observing changes in the brightness of stars in the same patch of sky for 4 years between May 2009 and May 2013. After *Kepler* lost two reaction wheels, the spacecraft was repurposed to observe the ecliptic plane, and is called the *K2* mission (Howell et al. 2014). The *K2* archive products are very similar to *Kepler*'s but are not specifically addressed in this document. For more information on *K2* see <http://keplerscience.arc.nasa.gov/>.

The *Kepler* telescope has a 0.95-meter aperture and a 115.6 deg² field-of-view (FOV). It was pointed at and recorded data from the same region of the sky for the duration of the mission. The single instrument on board, a photometer, is an array of 42 CCDs arranged in 21 modules. Each CCD has 2 outputs. The half-maximum bandpass is 435 to 845 nm, with >1% relative spectral response as short as 420 nm and as long as 905 nm. Each 50x25 mm CCD has 2200x1024 pixels. The interval between reads of a given pixel of a CCD is composed of an exposure time set to 6.019803 seconds and a fixed readout time of 0.51895 seconds. For each integration, all pixels were read-out and temporally summed in the Science Data Accumulator (SDA). Target lists determined which of those SDA summed pixels were read-out of the SDA and transmitted to the Solid State Recorder for later downlink. On average, 32 pixels are read out of the SDA per target. The *long cadence* data (LC) are summed into 29.424-minute bins (270 integrations), while the *short cadence* data (SC) are one-minute bins (9 integrations).

Kepler is in an Earth-trailing heliocentric orbit with a period of approximately 372.5 days. In this orbit the spacecraft slowly drifts away from the Earth and was at a distance of over 0.4 A.U. after 4 years. The orbit permits continuous pointing on a single region of the sky. Additional advantages are the very stable pointing attitude and the avoidance of the high radiation dosages associated with an Earth orbit.

The observed star field is near the galactic plane, centered on galactic coordinates $l = 76.32^\circ$, $b = +13.5^\circ$ (RA=19h 22m 40s, Dec=+44° 30' 00"). Figure 1-1 shows the field with the *Kepler* FOV superimposed. The squares show the 5 square degree FOV of each of the 21 modules. The gaps between the modules are aligned so that about half of the 15 stars in the FOV brighter than $m_v=6$ fall in these gaps. The 42 CCDs cover a four-way symmetrical pattern on the sky such that most of the same stars stay visible during the mission, even after a quarterly 90° roll. In addition, the orientation of the rows and columns of each module's location on the sky is preserved for all roll orientations except for the center module, which is only 180 degrees symmetric. The roll is necessary to keep the solar arrays oriented towards the Sun and the radiator pointed towards deep space.

Figure 1-1 Kepler Field of View



1.2 Overview of Data Flow

Data were downloaded from the spacecraft (S/C) through the Deep Space Network (DSN). The Mission Operations Center (MOC) at LASP received the data and telemetry packets, which were then binned into files by data type. The data were then sent to the Data Management Center (DMC) at STScI, where they are archived. The data were then decompressed, sorted by cadence (long or short) and pixel type (target, background or collateral), and converted to the FITS (Flexible Image Transport System) format. At this point in processing, the data were historically called “original”, but are more commonly called “raw” or “uncalibrated”. The data were then sent to the Science Operations Center (SOC) at NASA Ames, where detailed calibration was performed and light curve and target pixel files were produced.

The data were then returned to the DMC for archiving. For a given target, archive users can access the raw and calibrated pixel values in the target pixel files, or the integrated flux values in the light curve files. Because of improvements to the SOC data-processing pipeline, the cadence and FFI data were processed and archived many times. Each processing is uniquely identified with a data release number, and described by an accompanying Data Release Note. MAST only makes accessible the last version of light curve and target pixel files.

1.3 Related Documents

Documentation is available on-line for all archive holdings. The main archive page, <http://archive.stsci.edu>, provides links to a general introduction to MAST and a "getting started" page. Each mission page has links to mission specific information, and a mission specific "getting started" page. The MAST Kepler page is located at <http://archive.stsci.edu/kepler/>. Users are encouraged to explore the MAST Discovery Portal (The Portal; see §3.3) at <https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>. Another useful link for scientific papers about Kepler is <http://keplerscience.arc.nasa.gov/kpub-kepler.html>.

A description of *Kepler*, its design, performance and operational constraints may be found in the Kepler Instrument Handbook (KIH, Van Cleve & Caldwell 2016). A description of Kepler calibration and data processing is described in the Kepler Data Processing Handbook (KDPH, Jenkins et al. 2016 (in prep); Fanelli et al. 2011). Science users should also consult the special ApJ Letters devoted to early Kepler results and mission design (April 2010, ApJL, Vol. 713 L79-L207). Additional technical details regarding the data processing and data qualities can be found in the Kepler Data Characteristics Handbook (KDCH, Christiansen et al. 2013) and the Data Release Notes (DRN). This archive manual specifically documents the file formats, as they exist for the last data release of Kepler, Data Release 25 (KSCI-19065-002). The earlier versions of the archive manual and data release notes act as documentation for the earlier versions of the data files. The *Kepler* manuals and release notes may all be downloaded from the Kepler Documentation page at MAST (<http://archive.stsci.edu/kepler/documents.html>).

1.4 Overview of MAST

The Mikulski Archive for Space Telescopes supports a variety of astronomical data archives, with the primary focus on scientifically related data sets in the optical, ultraviolet, and near-infrared parts of the spectrum. See <http://archive.stsci.edu/missions.html> for a full list of the missions hosted by MAST and <http://archive.stsci.edu/hlsp/index.html> for the high-level science products, surveys, and catalog data distributed by MAST. MAST is a component of NASA's distributed Space Science Data Services (SSDS).

The staff of the Archive Sciences Branch and the Mikulski Archive for Space Telescopes (MAST) provide:

- world-wide technical and scientific leadership in archive system design
- secure storage and reliable retrieval services for data from all MAST-supported missions
- user-friendly and scientifically useful search and cross-correlation tools
- development and support for inter-archive communication and data transfer standards
- accurate and useful mission archive documentation
- helpful user support services with a typical response time of 1 business day

MAST provides a large suite of searches, including customized searches for each mission. It also provides several cross-mission search tools. MAST also archives sets of community contributed High-Level Science Products (HLSPs) solicited from the astronomical community including *Kepler* and *K2*.

The MAST Users Group provides essential user perspectives on archive operations and development, including suggesting priorities for short and long term operational and scientific enhancements to the archive. User feedback is obtained via an annual survey. Send e-mail to archive@stsci.edu to participate in the next survey.

The archive website, <http://archive.stsci.edu>, is the best place to start learning about MAST and what it can do to enable your research. The web site should always be consulted for the most current information.

1.5 User Support Services

Archive Hotseat

Help or answers to any questions about archive issues may be obtained by sending e-mail to archive@stsci.edu. The helpdesk staff will respond to questions concerning the archive and archive databases, and DVDs and hard disks provided by STScI. Helpdesk personnel also authorize accounts so that PIs and GOs can access their data. They will also provide advice concerning basic search strategies, and will investigate and document all problem reports. The archive helpdesk staff may not always know how to solve a problem, but they are responsible for finding out who does know the answer and for continuing to work with you until the problem is resolved. All initial communication from the user community to the archive should be directed to the archive helpdesk.

Questions and Comments

Communication regarding all aspects of the archive should normally be directed to the archive helpdesk (e-mail: <mailto:archive@stsci.edu>). This will allow Archive staff to respond to your requests even when individual members of the group are away.

1.6 Getting Your Data

Everyone may retrieve Kepler data via the MAST. A list of all data products and how to obtain them are tabulated at http://archive.stsci.edu/kepler/data_products.html. Use the Data Search and Retrieval form, http://archive.stsci.edu/kepler/data_search/search.php, to obtain *Kepler* light curves and target pixel files. Users may search for data sets by various target-based parameters such as Investigation ID and Kepler ID. All light curves, target pixel files and FFIs are also available for direct download from the public ftp site, <http://archive.stsci.edu/pub/kepler>. Also, MAST has added Kepler data to its Discovery Portal (<https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html>). Users are encouraged to explore the capabilities of the Portal. See Chapter 3 for more information about different ways to obtain *Kepler* data.

1.7 Publication Acknowledgement

Publications based on *Kepler* data should carry the following acknowledgement.

“This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission Directorate.”

Those publications based on data retrieved from MAST should carry the following acknowledgment.

"Some/all of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts.”

See the MAST Data Use Policy http://archive.stsci.edu/data_use.html for the current MAST grant number.

This Kepler Archive Manual should be referenced as
S. E. Thompson, D. Fraquelli, J. van Cleve and D. Caldwell, 2016, “Kepler Archive Manual”, KDMC-10008-006, <http://archive.stsci.edu/kepler/documents.html>.

Chapter 2 *Kepler* Data Products

2.1 Introduction

A variety of data products are archived for the *Kepler* space mission. These include science data of astronomical interest (light curves, target pixel data, and Full Frame Images (FFI)) and auxiliary data used to calibrate the images and determine the status of the spacecraft (focal plane characterization files, engineering data and telemetry data). The Science Operations Center (SOC) processes the data prior to it being archived.

Many *Kepler* specific terms are used to concisely and accurately describe the data and the processing. We have defined many of these terms in the glossary and list of acronyms in Appendix D.

2.1.1 Overview of Data in the Archive

Kepler's primary mission was to obtain flux time series of individual targets. As such, the majority of the *Kepler* science data is organized by target. Those interested in studying time variable phenomenon will find the light curve files and/or the target pixel files of interest. Additionally, *Kepler* downloaded and calibrated a single cadence of the entire *Kepler* field each month. These Full Frame Images (FFIs) are not target specific and can be downloaded from their own search and download page at MAST.

All science targets collected by *Kepler* are archived at the MAST. In some rare cases, targets are not processed through the entire SOC pipeline. Primarily, this happens when the optimal aperture, recalculated after the data are collected, contains no valid pixels for aperture photometry. This may also happen if the target is declared a catalog artifact. For each of these cases, no light curve file is available. Only pixel level data are available (i.e., the target pixel file).

Besides the astronomical data, *Kepler* archives much of the auxiliary data that either directly or indirectly affects the astronomical data products. Engineering data describes the state of the instrument and the spacecraft during data collection. Focal plane characterization models, artifact removal pixels and collateral CCD data are used to calibrate and process the data.

2.1.2 *Kepler* Time System

The readout time for each cadence is recorded as a Vehicle Time Code (VTC). This timestamp is produced within 4 ms of the readout of the last pixel of the last frame of the last time slice (see glossary). When the data is downloaded to Earth, the Mission Operations Center converts VTC to Coordinated Universal Time (UTC), correcting for leap seconds and any drift in the spacecraft clock, as measured from telemetry. UTC times are converted to Barycentric Dynamical Time (TDB) then to Barycentric Julian Date (BJD) to correct for the motion of the spacecraft around the center of mass of the solar system. TDB is a time system that does not include the leap seconds that bedevil calculations of periods in the UTC system. TDB agrees with the time systems TDT and TT to better than 2 ms. See Eastman et al. (2010) for a recent discussion of the various time systems common in astronomy.

Time stamps expressed in Modified Julian Date (MJD) or UTC are geocentric and not corrected to the TDB time system. Time is specified in the data files with an offset from BJD. The offset has a value of 2454833.0 and is specified in each file header (see BJDREF). The units are clearly defined in the headers of the *Kepler* data files. The SOC uses SPICE kernels, which are calculated from the spacecraft

telemetry, to calculate barycentric corrections. For more information on the SPICE kernels and SPICE tools visit <http://naif.jpl.nasa.gov/naif/>.

The quoted times for any cadence are believed to be accurate to within ± 50 ms. This requirement was developed so that knowledge of astrophysical event times is limited by the characteristics of the event, rather than the characteristics of the flight system, even for high SNR events. Users who require temporal accuracy of better than 1 minute should read the KDCH (§6) and KIH (§7.3) carefully. These documents contain additional details regarding the times, including accuracy, corrections for readout time slice offsets, change in units, etc.

2.1.3 FITS headers

Kepler science data use the FITS (Flexible Image Transport System) format to comply with astronomical data standards. Data headers use standard FITS keywords to formulate the data definition and comply with all FITS recommended keyword usage current at the time of header design. The primary headers contain keywords inherited by all subsequent extensions. Primary header keywords specify data processing inputs, data quality, observational modes (long cadence, short cadence, FFI), target information, etc. The primary header/data units (HDU) do not contain a data array. FITS headers for cadence science data are specified in Appendix A.

2.2 File Name Syntax

Many file types are archived for *Kepler*. To prevent confusion, a standard syntax is used for *Kepler* filenames.

Kepler filenames have 3 components:

- **Rootname:** Usually a timestamp, for some file types the rootname contains other identifiers such as the Kepler Identification number (KepID) or the module/output (mod/out) number. The rootname begins with the character string ‘kplr’ followed by a time stamp of the form yyyydddhhmmss, where ddd is the day of year. If KepID or mod/out is present in the rootname, it precedes the timestamp (e.g., kplr<kepler_id>_<stop_time>). See Table 2-1 for a list of rootnames.
- **Suffix:** The suffix indicates the type of data in the file within the data set (short cadence light curve, target pixel, background, etc.) See Table 2-1 for a list of suffixes.
- **Extension:** The extension indicates the format of data contained in the file (e.g. fits, csv, or xls).

These three components are concatenated as shown to form the file name.

kplr<rootname>_<suffix>.<extension>

In the archive, the data set name is the rootname.

Table 2-1 Filenames

Type of Data	Rootname*	Suffix	Extension
Calibrated Light Curves:			
LC calibrated light curves	<kepler_id>-<stop_time>†	llc	fits
SC calibrated light curves	<kepler_id>-<stop_time>†	slc	fits
Target Pixel Data:			
LC Target pixel data	<kepler_id>-<stop_time>†	lpd-targ	fits
SC Target pixel data	<kepler_id>-<stop_time>†	spd-targ	fits
Full Frame Image:			
FFI original data	<stop_time>	ffi-orig	fits
FFI calibrated data	<stop_time>	ffi-cal	fits
FFI uncertainties	<stop_time>	ffi-uncert	fits
Auxiliary Data:			
Background data	<modout>-<stop_time>†	bkg	fits
LC Collateral data	<modout>-<stop_time>†	coll	fits
SC Collateral data	<modout>-<stop_time>†	cols	fits
Artifact Removal Pixel data	<modout>-<stop_time>†	arp	fits

* All rootnames begin with kplr.

† For historical reasons, the <stop_time> for time series files is given in local Pacific time for the last processed cadence of that quarter or month. All other stop times are given in UTC.

2.2.1 Custom Aperture Targets

Certain targets observed by Kepler, known as custom targets, have special apertures either because the target is not in the Kepler Input Catalog (KIC, see Chapter 3) or requires a special aperture to collect the appropriate pixels. The custom aperture targets are given a Kepler ID number between 100,000,000 and 100,004,300 to distinguish them from the typical Kepler targets. These special ID numbers are used in the data file names, as the Kepler ID in the header, and for the Kepler ID field on the data search page at MAST. Custom targets are given a different number each quarter and for some targets, more than one custom aperture is defined each quarter. Data searches by the traditional KIC identification number will not return the data files for the custom apertures of the specified target. The MAST has, however, linked the custom targets to the correct RA and Dec positions. To find all custom apertures for a particular target, perform a cone search around the RA and Dec position of the desired target.

There are two types of custom apertures of particular scientific interest to the general community. The first aperture type contains data from the open clusters NGC 6819 and NGC 6791. Each cluster is tiled with contiguous custom apertures that are placed side-by-side to create a 200 by 200 pixel block. The best way to find these targets is to search for an investigation ID of “STC” and a custom aperture number greater than 100 million. You can distinguish the clusters by their RA and Dec values. The second aperture type of interest is the background super apertures, which cover relatively dark regions of sky at a common location on all CCD channels. These data were collected continuously between quarters 5 and 17. To find these data, do a data search for an investigation ID of “EXBA”.

2.3 Data in the Archive

Most users of *Kepler* data will be interested in the light curves and the target pixel data. The full frame images (FFI) are useful for examining contamination near targets, but are not as useful for time series analysis of individual targets. A person interested in identifying targets on *Kepler's* field of view will be interested in the Kepler Input Catalog (KIC) and Characteristics Table (CT). The specific file versions discussed in this archive manual are listed in Table 2-2. See the FILEVER keyword in the primary header to determine the version number of a specific file. Also the Data Release number is specified in the primary header of most archive data files to indicate the version of the data contained in the files. The last *Kepler* data release for the FFIs, target pixel, light curve, collateral, ARP, focal plane characterization, and background files is data release 25.

Table 2-2 File Version Numbers described in this Archive Manual

File Type	Version *	Archive Manual Section
Light Curves Files	6.1	2.3.1
Target Pixel Files	6.1	2.3.2
Full Frame Images	3.1	2.3.3
Background	3.0	2.3.6
Collateral/ARP	2.0	2.3.7 - 2.3.8
CBVs	4.0	2.3.4
Focal Plane Characterization	1.0	2.3.5

*See the keyword FILEVER in the header of each file.

2.3.1 Light Curve Files

Light curve files are produced for each target using simple aperture photometry. For any particular quarter, there are more than 160,000 long cadence targets and up to 512 short cadence targets being archived. Short cadence targets always have a corresponding long cadence light curve. Long cadence targets will be observed for at least a quarter and short cadence targets will be observed for at least a month (except for Q4 where targets on module 3 were lost due to a hardware failure). In the case where a target is observed at both long and short cadence, there will be one long cadence light curve each quarter and up to three short cadence light curves, one for each month.

A light curve file contains time series data. Each data point corresponds to a measurement from a long or short cadence. Long and short cadence data are not mixed in a given light curve file. For each data point there are multiple flux and centroid values along with uncertainties. The value NaN is specified for any missing data values.

The light curves are packaged as FITS binary table files with a primary header, a light curve extension and an aperture extension. The long cadence FITS headers are listed in Appendix A.1.

Primary Header

The primary header contains information pertaining to the entire file, such as target information and version processing information. This header contains information about which CCD was used to collect these data, which quarter the data was collected, and which Data Release Notes apply to this processing of the data. The header contains keywords for the properties of the target star as reported in the KIC.

Light Curve Binary Extension

The binary table contains the following data columns:

TIME [64-bit floating point] – The time at the mid-point of the cadence in BKJD. *Kepler* Barycentric Julian Day (BKJD) is Julian day minus 2454833.0 (UTC=January 1, 2009 12:00:00) and corrected to be the arrival times at the barycenter of the Solar System. The pipeline uses the right ascension and declination of the object (found in the primary header), along with the location of the spacecraft at the time of the cadence to perform this calculation. This column can be converted to BJD using the following formula for each member of the series [i]:

$$\text{BJD}[i] = \text{TIME}[i] + \text{BJDREFI} + \text{BJDREFF},$$

where BJDREFI and BJDREFF are given as keywords in the header.

TIMECORR [32-bit floating point] – The barycenter correction calculated by the pipeline plus the time slice correction. This column allows users to revert back to non-barycentric times, if required. To convert the barycentered times in the TIME column to the Julian Day of the observation, use the following formula:

$$\begin{aligned} \text{JD}[i] &= \text{BJD}[i] - \text{TIMECORR}[i] + \text{time_slice_correction} \\ &= \text{BJD}[i] - \text{TIMECORR}[i] + (0.25 + 0.62(5 - \text{TIMSLICE}))/86400 \end{aligned}$$

where TIMSLICE is given in the header. To obtain the times in Julian Day reported by the spacecraft, simply subtract the TIMECORR column from the TIME column, and then add the integer offset, 2454833, stored in BJDREFI and BJDREFF.

CADENCENO [32-bit integer] – The cadence number is a unique integer that is incremented by one with each cadence.

SAP_FLUX [32-bit floating point] – The flux in units of electrons per second contained in the optimal aperture pixels collected by the spacecraft. This light curve is the output of the PA module in the SOC pipeline.

SAP_FLUX_ERR [32-bit floating point] – The error in the simple aperture photometry as determined by PA in electrons per second. The reported errors for each cadence are a sum of the minimal error calculation (shot noise plus read noise) and an offset term to account for the extra error from the full propagation of errors.

SAP_BKG [32-bit floating point] – The total background flux summed over the optimal aperture in electrons per second. The background flux for each pixel is calculated by fitting a surface to the background pixels on each mod/out.

SAP_BKG_ERR [32-bit floating point] – The 1-sigma error in the simple aperture photometry background flux.

PDCSAP_FLUX [32-bit floating point] – The flux contained in the optimal aperture in electrons per second after the PDC module has applied its cotrending algorithm to the PA light curve. To better understand how PDC manipulated the light curve, read Section 2.3.1.2 and see the PDCSAPFL keyword in the header.

PDCSAP_FLUX_ERR [32-bit floating point] – The 1-sigma error in PDC flux values.

SAP_QUALITY [32-bit integer] – Flags containing information about the quality of the data. Table 2-3 explains the meaning of each active bit. See the Data Characteristics Handbook and Data Release Notes for more details on safe modes, coarse point, argabrightenings, attitude tweaks, etc. Unused bits are reserved for future use.

PSF_CENTR1 [64-bit floating point] – The column centroid calculated by fitting the point-spread function (PSF) combined with the modulation from pointing jitter and other systematic effects (also referred to as pixel response function). This value is only calculated for selected stars, see discussion of PPA targets in the Kepler Data Processing Handbook.

PSF_CENTR1_ERR [32-bit floating point] – The 1-sigma error in PSF-fitted column centroid.

PSF_CENTR2 [64-bit floating point] – The row centroid calculated using the PSF fitting described above. This value is only calculated for selected stars.

PSF_CENTR2_ERR [32-bit floating point] – The 1-sigma error in PSF-fitted row centroids.

MOM_CENTR1 [64-bit floating point] – The column value for the flux-weighted centroid (first moment) position of the target at this cadence.

MOM_CENTR1_ERR [32-bit floating point] – The 1-sigma error in the column value for the first moment centroid position.

MOM_CENTR2 [64-bit floating point] – The row value for the flux-weighted centroid (first moment) position of the target at each cadence.

MOM_CENTR2_ERR [32-bit floating point] – The 1-sigma error in the row value for the first moment centroid position.

POS_CORR1: [32-bit floating point] – The column component of the local image motion calculated from the motion polynomials. We report the motion in pixels relative to the mid-cadence of the quarter.

POS_CORR2: [32-bit floating point] – The row component of the local image motion calculated from the motion polynomials. We report the motion in pixels relative to the quarter's mid-cadence.

2.3.1.1 Quality Flag Definitions

The target pixel, light curve and DV time series (see KSCI-19079-001) files contain a data column called QUALITY or SAP_QUALITY to indicate when certain spacecraft events occurred, or when the pipeline flagged interesting phenomenon. Table 2-3 gives the summary of these flags. Most flags provide information and do not indicate that the data is unusable. For instance, the cosmic ray flag, bit 8, indicates that a cosmic ray was corrected in the optimal aperture on that cadence. The SPSD flag (bit 11, Single Pixel Sensitivity Drop-out, see KDCH), set by PDC, flags the last good cadence before the SPSD is detected. The manual exclude (bit 9) and coarse point (bit 16) flags are set manually based on information collected during pre-processing. The rolling band flags (bits 18 and 19) indicate that rolling band artifact, see KIH, was detected at some level in the optimal aperture and the entire mask, respectively. Further investigation of the size of the rolling band is warranted, depending on the science case, to decide if the cadence should be considered invalid.

Table 2-3 – Bits for the QUALITY and SAP_QUALITY data column.

Bit	Value	Explanation
1*	1	Attitude Tweak
2*	2	Safe Mode
3*	4	Spacecraft is in coarse point. It is set manually to pad not-in-fine point data.
4*	8	Spacecraft is in Earth point. The first real cadence after Earth point is marked.
5	16	Reaction wheel zero crossing
6*	32	Reaction wheel desaturation event
7*	64	Argabrightening detected across multiple channels on this cadence
8	128	Cosmic Ray was found and corrected in optimal aperture pixel
9*	256	Manual Exclude. The cadence was excluded because of an anomaly.
10	512	This bit is unused by <i>Kepler</i> .
11	1024	SPSD detected. This bit is flagged on the last non-gapped cadence before the maximum positive change due to the detected SPSP.
12	2048	Impulsive outlier removed before cotrending
13*	4096	Argabrightening event on specified CCD mod/out detected
14	8192	Cosmic Ray detected on collateral pixel row or column in optimal aperture.
15*	16384	Detector anomaly flag was raised.
16*	32768	Spacecraft is not in fine point.
17*	65536	No data collected.
18	131072	Rolling Band detected in optimal aperture.
19	262144	Rolling Band detected in full mask.
20	545288	Possible thruster firing. Not set in <i>Kepler</i> data
21	1048576	Thruster firing. Not set in <i>Kepler</i> data.

* indicates that these cadences are gapped by the pipeline (either in CAL, PA or PDC). The original pixel level data is available in most cases.

Aperture Extension

The aperture extension contains a single image that describes which pixels were collected by the spacecraft, which pixels are contained in the optimal aperture and which pixels were used to calculate the centroids. The pixels in the optimal aperture are those used to create the SAP_FLUX light curve and are likely different for each quarter. The FITS standard requires a rectangular bounding box even though many target apertures are not rectangles. Therefore the image contains null pixels that were never collected (i.e., the image includes the extra pixels necessary to create a rectangular image). See Table 2-4 for a description for the reason each bit is set.

Table 2-4 – Aperture image bit descriptions for light curve files

Bit	Value	Meaning
1	1	pixel was collected by spacecraft
2	2	pixel was part of the optimal aperture
3	4	pixel was used to calculate the flux weighted centroid
4	8	pixel was used to calculate the PRF centroid

Note: in Data Release 25, bits 3 and 4 were not set in long cadence light curves. One halo of pixels around the optimal aperture is used to calculate the flux-weighted centroid and the entire set of collected pixels is used to calculate the PRF centroid.

2.3.1.2 The PDC light curves (PDCSAP_FLUX)

The primary purpose of the Presearch Data Conditioning (PDC) module of the Kepler data processing pipeline is the removal of signatures in the light curves that are correlated with systematic error sources from the telescope and spacecraft, such as pointing drift, focus changes, and thermal transients. In most cases PDC achieves its goal of removing these errors while preserving planet transits and other astrophysically interesting signals. To do this, PDC uses a Bayesian Maximum A Posteriori (MAP) approach to establish a range of “reasonable” robust fit parameters. These robust fit parameters are then used to generate a “Bayesian Prior” and a “Bayesian Posterior” probability distribution function which, when maximized, finds the best fit that simultaneously removes systematic effects while reducing the signal distortion and noise injection that commonly afflicts simple Least Squares (LS) fitting.

For data formatted as v5.0 and greater, Long Cadence PDC light curves use an msMAP (multiscale MAP) approach (Stumpe et al. 2014). msMAP is a wavelet-based band-splitting framework for removing systematics from the light curves. In each band, a subset of highly correlated and quiet stars is used to generate a cotrending basis vector (CBV) set. Each of the bands is handled differently where the options are: 1) perform a simple robust fit to the CBVs, 2) perform a MAP fit based on a Bayesian Prior (Smith et al. 2012) or 3) perform no fit. The type of fitting performed in each band is specified in the FITS headers (see FITTYPE_j). msMAP does not always perform better than regular, single-band MAP. PDC calculates both and provides the better of the two reductions, determined by examining the total PDC goodness metric. The keyword PDCMETHD in the FITS header indicates whether the ‘multiscale MAP’ or ‘regular MAP’ light curve was chosen. See §2.4 for more information on the keywords.

The Short Cadence PDC processing uses a single-band MAP fit. However, because of the limited number of short cadence targets, the long cadence basis vectors are interpolated to create the short cadence CBVs. Also, the Bayesian priors used to do the MAP fit are drawn from the long cadence priors.

PDC performs several other critical services. One is the identification and removal of Sudden Pixel Sensitivity Dropouts (SPSDs, or Discontinuities) that result in abrupt drops in pixel flux with short recovery periods. These corrections are described using the keywords NSPSDDDET and NSPSDCOR for the number of SPSPDs detected and corrected, respectively. Another service is the adjustment of light curve fluxes to account for excess flux in the optimal apertures due to star field crowding. This is done using the following equation:

$$\text{crowdingCorrectedFlux} = \text{cotrendedFlux} - (1 - \text{crowdingMetric}) * \text{median}(\text{cotrendedFlux})$$

The crowding metric can be found in the FITS header keyword CROWDSAP. PDC also corrects for the fraction of flux collected by the optimal aperture; the flux fraction, is stored in the header keyword FLFRCSAP.

PDC “protects” known transits from falsely being detected as Sudden Pixel Sensitivity Dropouts (SPSPDs) or other types of outliers. Cadences containing known transits and eclipses are computed using the known epoch, period and duration of the events, and assuming a linear ephemeris. No SPSPDs or outliers are flagged during the known transits. This helps preserve transit depths and shapes from corruption by the SPSPD and outlier correction algorithms. Note: this only affects transits known at the time the data were processed. There is still the risk of transit corruption for undetected transits.

Goodness metrics are also provided in the FITS headers for long-cadence PDC MAP data. These goodness metrics characterize the “goodness” of the PDC cotrending with regard to four characteristics. A total goodness is also given as the geometric mean of the four components. The values range between 0 and 1, where 0 is poor and 1 is perfect goodness. Percentile values are also given for each component and the sum. While a percentile value is given for all targets, the statistical range is only generated for non-custom targets; this is to prevent custom targets from skewing the statistics for standard targets. Large values for all four components does not necessarily mean the resultant PDC light curve has been reliably corrected, only that it is good with respect to those metrics. For more information on the PDC goodness metrics and how they are defined see Stumpe et al. (2012).

The four goodness metrics and their percentiles found in the header of the light curve extension are as follows:

- 1) Correlation: Cotrending attempts to remove correlated systematics between the targets. Any residual correlation is characterized by a lower value. See keywords PDC_COR and PDC_CORP.
- 2) Variability: Using an estimation of the stellar variability (Smith et al. 2012), any change in the intrinsic variability due to the cotrending will result in a lower value for this metric. A value of one means the variability was perfectly preserved. See the keywords PDC_VAR and PDC_VARP.
- 3) Noise: Indicates if high frequency noise, just below the Nyquist Frequency, was added to the cotrended light curves. See keywords PDC_NOI and PDC_NOIP.
- 4) Earth Point: A value close to one means that an earth point thermal recovery signature was removed well. See keywords PDC_EPT and PDC_EPTP.

2.3.2 Target Pixel Data

For each target, *Kepler* only acquires the pixels contained within a predefined mask. These pixels are used to create the data found in the light curve files. Each target pixel file packages these pixels as a time series of images in a binary FITS table. The intent of these files is to provide the data necessary to perform photometry on the raw or calibrated data when needed (or desired) to understand (or improve) the automated results of the *Kepler* pipeline.

In the binary table, the pixel values are encoded as images. Each element in the binary table contains the pixels from a single cadence. Missing integer values are filled with the value -1, missing floating-point values are filled with the value NaN as described by the FITS standard, and keywords with missing values are either left as blank or filled with an empty string.

Each target pixel file has a primary header and two extensions: target table and aperture. The primary header describes the target and the processing software. The target table extension contains the flux time series for both the raw and calibrated pixels. The aperture extension describes the target pixel mask and optimal aperture. The headers for each extension are listed in Appendix A.2.

Primary Header

The primary header contains information pertaining to the entire file, such as target information and version processing information. This header contains information about which CCD was used to collect these data, which quarter the data was collected, and which Data Release Notes apply to this processing of the data. The header contains keywords for the properties of the target star, obtained from the KIC. Primary header keywords are supplied to better understand the data processing and the target properties; they are not intended for publication without first understanding their source. Appendix A.2a contains a description of all the keywords in this header.

Target Table Extension

The target table extension contains pixel time series. It has 13 columns containing a series of either scalar values or images. The header defines these columns according to the FITS standard (Pence et al. 2010). The keywords are explained in Appendix A.2b.

The image dimensions vary from target to target, however all images contained in a single target pixel file have the same dimensions. The location of the pixels on the specified CCD is provided in the header. The celestial World Coordinate System solutions for each image column is also specified, however not all FITS viewers support these keywords.

The definition of each column in the Target Table Extension is as follows:

TIME [64-bit floating point] – The time at the center of the cadence in BKJD. This column is identical to the light curve file TIME column. *Kepler* Barycentric Julian Day is a Julian day minus 2454833.0 (UTC=January 1, 2009 12:00:00) and corrected to be the arrival times at the barycenter of the Solar System. The pipeline uses the right ascension and declination of the object (found in the primary header), along with the location of the spacecraft at the time of the cadence to perform this calculation. (For objects with no KIC number, this value is calculated for the RA and Dec at the center of the mask.) This column can be converted to BJD using the following formula for each member of the series [i]:

$$\text{BJD}[i] = \text{TIME}[i] + \text{BJDREFI} + \text{BJDREFF},$$

where BJDREFI and BJDREFF are given as keywords in the header.

TIMECORR [32-bit floating point] – The barycenter correction calculated by the pipeline plus the time slice correction (see KIH). This column is identical to the light curve file TIMECORR column. This column allows users to revert back to non-barycentric times, if required. To convert the times in the first column to the Julian Day for the channel use the following formula:

$$\begin{aligned} \text{JD}[i] &= \text{BJD}[i] - \text{TIMECORR}[i] + \text{time_slice_correction} \\ &= \text{BJD}[i] - \text{TIMECORR}[i] + (0.25 + 0.62(5 - \text{TIMSLICE})) / (86400) \end{aligned}$$

where TIMSLICE is given in the header. To obtain the times in Julian Day reported by the spacecraft, subtract the TIMECORR column from the TIME column.

CADENCENO [32-bit integer] – The cadence number is a unique integer that is incremented with each cadence. This column is identical to the light curve file CADENCENO column.

RAW_CNTS [2D array of signed 32-bit integers] – One image per cadence of the raw counts measured in each pixel downloaded from *Kepler*. To restore the values contained in this column to the Analog to Digital Units read off the photometer, subtract the appropriate “fixed offset”, LC or SC (keywords LCFXDOFF and SCFXDOFF), and add the mean black level (keyword MEANBLCK) times the number of readouts (keyword NREADOUT). Then, each count represents one Analog to Digital Unit.

FLUX [2D array of 32-bit floating point] – One image per cadence of the measured flux in each pixel after processing by the pipeline module CAL, the removal of the interpolated background, and the removal of cosmic rays. The units are electrons per second. See Quintana et al. (2010) on pixel level calibrations or the KDPH for more details on the processing. This column may be used, along with the optimal aperture, to reproduce the SAP_FLUX values found in the light curve file for this target by summing the values in the optimal aperture pixels for each cadence.

FLUX_ERR [2D array of 32-bit floating point] – An image of the 1-sigma error in the measured flux as calculated by CAL in electrons per second. This error includes the error from the background subtraction. This per pixel error does not, and cannot, include the pixel-to-pixel correlated background errors that are included in the errors calculated for the light curve files.

FLUX_BKG [2D array of 32-bit floating point] – An image of the background flux subtracted from the data in electrons per second. The background flux is calculated by interpolating a 2 dimensional surface from approximately 4500 dedicated background pixels on each channel.

FLUX_BKG_ERR [2D array of 32-bit floating point] – An image of the 1-sigma error in the background flux. This per pixel error does not include the pixel-to-pixel correlated background errors that are included in the error columns of the light curve files.

COSMIC_RAYS [2D array of 32-bit floating point] – An image of the cosmic ray flux identified by the module PA. The units are electrons per second. For most cadences, this image will be an array of NaNs to indicate that there were no cosmic rays. The quality flag, bit 8, indicates when a cosmic ray falls in the optimal aperture.

QUALITY [32-bit integer] – Flags containing information about the quality of the data. Table 2-3 explains what activates each bit. See the Data Characteristics Handbook and Data Release notes for more details on safe modes, coarse point, argabrightenings, attitude tweaks etc. Unused bits are reserved for future use.

POS_CORR1: [32-bit floating point] – The column component of the local image motion calculated from the motion polynomials. We report the motion in pixels relative to the mid-cadence of the quarter.

POS_CORR2: [32-bit floating point] – The row component of the local image motion calculated from the motion polynomials. We report the motion in pixels relative to the quarter’s mid-cadence.

RB_LEVEL: [2D array of 32-bit floats] – A 2D array of the rolling band level measured for each CCD row and each pulse duration by the Dynablack module. The rows of the 2D array are the same as the CCD rows in the FLUX and RAW_CNTS columns. The array’s columns are in order of increasing pulse duration; the pulse durations in hours are given in the header under RBTDUR*i*. See Section 2.3.7 for more information on rolling bands and dynablack.

Aperture Extension

The aperture extension contains a single image that describes which pixels were collected by the spacecraft and which pixels are contained in the optimal aperture. Those pixels in the optimal aperture are used to create the SAP_FLUX light curve described in Section 2.3.1. The FITS standard requires a rectangular bounding box even though many target apertures are not rectangles. Therefore the image contains null pixels that were never collected (i.e., the image includes the extra pixels necessary to create a rectangular image). The aperture mask is a 32-bit integer with a value between zero and three. Other bits are reserved for future use. See Table 2-5.

Table 2-5 – Aperture image bit descriptions

Bit	Value	Meaning
1	1	pixel was collected by S/C
2	2	pixel is in the optimal aperture

2.3.3 Full Frame Image

A Full Frame Image (FFI) contains values for every pixel in each of the 84 channels. Nominally three FFIs are taken each quarter and are used to confirm the proper orientation and placement of the detector on the sky and to assess photometer health. The DMC process generates an FFI data file that contains the uncalibrated pixels (ffi-orig). The SOC produces the calibrated and uncertainty images (ffi-cal and ffi-uncert). The uncalibrated, calibrated, and uncertainty FFIs contain one extension per mod/out. A complete FFI dataset consists of 3 files, each with a filename like kplr<stop_time>_suffix. Tables 2-6 and 2-1 give details on the file names and data set.

Table 2-6 – FFI Data set and files

Data type	File suffix	Format
FFI – uncalibrated	ffi-orig	84 FITS image extensions
FFI – calibrated	ffi-cal	84 FITS image extensions
FFI – uncertainties	ffi-uncert	84 FITS image extensions

The original, or uncalibrated, FFI created by the DMC contains the raw counts collected from the spacecraft as well as the collateral (black and smear) data in units of electrons per cadence. The header keywords do not match the keywords of the calibrated FFI files created by the SOC, and are self-documenting. The calibrated FFI was processed through the CAL portion of the Kepler pipeline and has units of electrons per second. Standard calibrations, such as flat fields, blacks, and smears have been applied. For more details see the KDPH. The uncertainty file, also in units of electrons per second, contains the propagated uncertainties on the flux for each pixel in the calibrated FFI.

The FFIs contain a World Coordinate System (WCS) solution calculated from the motion polynomials. Prior to data release 24, the pipeline did not calculate a motion polynomial directly from the FFI. So the FFI WCS was based on the motion polynomial from the nearest long cadence. In DR24 and DR25, the WCS is measured using motion polynomials created directly from the FFI. The WCS is stored in each extension header using the keywords specified by Greisen & Calbretta (2002) and Calbretta & Greisen (2002). The distortion coefficients that describe non-linearities in the shape of the focal plane are accounted for using Simple Imaging Polynomials (SIP) outlined by Schupe (2005). The WCS solution typically achieves an accuracy of 0.1 pixels (0.4 arcseconds).

Eight “Golden” FFIs are considered part of quarter 0 and were taken during commissioning after the telescope had been focused, thermal stability obtained and pointing established. Along with the FFIs from Q2 and Q17, these were last processed as part of DR21 and include a WCS that is not calculated directly from the image.

2.3.4 Co-trending Basis Vectors

The co-trending basis vectors (CBVs) represent the set of systematic trends present in the ensemble flux data for each CCD channel. The CBVs are computed when creating the PDC time series and are provided as part of the quarterly delivery of data. The user may fit the CBVs to light curves to remove the common instrumental effects from the data when the PDC time series is insufficient for their science goals. More details on the method used to generate these basis vectors are given in the KDCH.

CBV Format

The Kepler mission creates one cotrending basis vector (CBV) file each quarter. The cotrending basis vectors are named according to the following format: `kplr <yyyydddhhmss>-q<##>-d<##>_lcbv.fits`. Where the time represents the stop time of the data set in UTC, the `q<##>` represents the quarter number and the `d<##>` represents the data release number. A new CBV file is created each time the data is reprocessed. The `DATA_REL` keyword in the header of the CBV file should match the same keyword in the data being cotrended.

The CBV consists of a primary header and 84 data extensions, one for each mod/out. Each data extension contains the following columns:

TIME [32 bit floating point] – The time at mid-cadence in modified Julian date.

CADENCENO [integer] – The cadence number is a unique integer that is incremented with each cadence.

GAPFLAG [integer] – This flag is set to a value of one when the PDC portion of the pipeline has gapped the data. This occurs because the spacecraft is in coarse point, safe mode, at earth point, during a momentum dump, a detector anomaly occurs, an argabrightening event was detected on that mod/out, a manual exclude, or an attitude tweak occurs. The light curves are linearly interpolated across these gaps with a piecewise cubic Hermite polynomial prior to creating the basis vectors. PA data do exist for many of these events; however, the use of CBVs on these gapped data points is not recommended.

VECTOR_1 [32 bit floating point] – The first co-trending basis vector. The file includes the first 16 basis vectors as the subsequent columns of this binary data table.

Warning: Currently, *Kepler* provides 16 basis vectors. However, using the vectors above eight generally does not provide much improvement to the co-trending and has been known to introduce non-physical signals into the light curve. Also, the multi-scale band split basis vectors used by PDC (see §2.3.1.2) are not provided to the end user.

2.3.5 Focal Plane Characterization Models

Before launch and during commissioning, *Kepler* obtained data to model the characteristics of the focal plane. These focal plane characterization (FC) models are used for various calibration and diagnostic tasks in the *Kepler* Data Processing Pipeline, with retrieval and usage details in Allen, Klaus, & Jenkins (2010). The FC models, subsection numbers, and supporting technical documents (available from MAST) are shown in Table 2-7. The order of the sections is roughly that of the calibration steps described in detail in Quintana et al. (2010) and Kolodziejczak et al. (2010). The values archived at MAST are those used for the final *Kepler* data processing and release (i.e., Data Release 25).

Several models are related to each other and need to be used consistently. For example, one can choose to use the 1D black or dynamic black models to represent the spatial, temporal, and temperature dependence of *Kepler*'s black image, while the 2D black model gives the static part of the black image. The 1D black model in turn must be used with the scene-dependent row model to give best results. Similarly, one must use the pointing, roll-time, and focal-plane geometry models together or not at all to get intelligible results for mapping RA and DEC on the sky to focal plane pixel coordinates.

Several models were not actually used in *Kepler* data processing for any data release. These models either did not improve the aggregate noise performance (invalid pixels, spectral response) or the algorithms required were not developed well enough to incorporate in the pipeline (video crosstalk). They are documented here for those who wish to improve on the *Kepler* pipeline in the future.

The focal plane characterization models are available at MAST at <https://archive.stsci.edu/kepler/fpc.html>.

Table 2-7 – List of FC Models. Column 4 indicates whether the file was used by the pipeline in the final *Kepler* data release.

FC Model	§	File Name	P/L?	Supporting Documents
clock-state-mask	1	clock_state_mask_KADN-26205_20081218.fits	Y	KADN-26205
2-D black	2	kplr2015238_2dblack.fits, kplr2015238_2dblack-uncert.fits	Y	
dynamic black	3	black-retrieval-tool directory, kplr<YYYYDOYHHMSS>-q<##>-<mmo>- dr<##>_dynablack.mat ("blob file")	Y	
1-D black	4	extract_one_d_black_from_blob.m, kplr<YYYYDOYHHMSS>-q<##>-<mmo>- dr<##>_ldblack.mat	Y	
scene-dependent rows	5	is_scene_dependent.m kplr2015279231411- dr25_blackrowmask.txt	Y	
bad smear columns	6	is_saturated_smear_column.m	Y	
read noise	7	kplr2008102416_read-noise.txt	Y	
gain	8	kplr2008072318_gain.txt	Y	KADN-26196
linearity	9	kplr2009060215_linearity.txt	Y	KADN-26196
undershoot	10	kplr2008102809_undershoot.txt	Y	KADN-26197
large-scale flat	11	kplr2015236_largeflat.fits	Y	
small-scale flat	12	kplr2015236_smallflat.fits, kplr2015236_smallflat-uncert.fits	Y	
invalid pixels	13	n/a	N	
pointing	14	kplr2013050211_pointing.txt	Y	KADN-26285
roll-time	15	kplr2013040800_rolltime.txt	Y	
focal-plane geometry	16	kplr2009060418_geometry.txt	Y	KADN-26285
simplified (RA, DEC) to pixel mapping	14-16	morc_2_ra_dec_4_seasons.xls	N	
pixel response function	17	kplr<module>.<output>_<yyyyddd>.fits	Y	
video crosstalk	18	kplr_2015111701_video_crosstalk_rec_s rc.csv	N	
spectral response	19	kplr_2015120120_spectral_response.csv	N	

2.3.5.1 Clock State Mask

The Fine Guidance Sensor (FGS) frame transfer (FT) and parallel transfer (PT) clocks crosstalk into the science data as described in KIH §6.2. The clock state mask describes the state of the FGS clocks during the digitization of a science pixel for users who wish to model the time and temperature dependence of FGS crosstalk. To first order, all science pixels having the same FGS clock state can be expected to have the same crosstalk offset and thermal behavior. The clock state mask includes both a 3-pixel delay due to the analog-to-digital conversion process, and a 3-pixel advance to provide a baseline

just before the crosstalk, with the net result that the first FGS-PT pixel identified in the clock state mask for each PT clock sequence occurs just as the FGS is clocked, but 3 pixels before the effect of this clocking appears in the digitized science pixel output. The clocking crosstalk pattern repeats 5 times during a science frame read-out (due to the 10 Hz FGS and the 0.5 sec science CCD readout times).

The file format is a FITS image of unsigned integers, 1132 columns by 1070 rows. The values 16-31 indicate the 16 FT pixel states, values 32-63 the 32 PT pixel states, and the value 96 indicates the serial transfer state. Values 0-15 are not used. The clock state mask is the same for every channel. The file name is `clock_state_mask_KADN-26205_20081218.fits`. See KADN-26205 (Van Cleve, 2008) for a more detailed discussion.

2.3.5.2 Static 2D Black

Black Level Overview: The black level (or bias) is the output of a CCD when it is not illuminated, and dark current is negligible at low temperatures or short integration times. Due to the nature of *Kepler*'s readout electronics, this black level has a two-dimensional spatial structure that also varies with time and temperature (KIH §6). Pipeline pixel processing before DR 24 used only the time and temperature-independent part of the two-dimensional black, described in this section, along with the time varying 1D black (§2.3.5.4; Quintana et al., 2010). For DR 24 and later releases, the time and temperature dependence (see Kolodziejczak et al. 2010) was modeled using the 2D dynamic black model “Dynablack” (§2.3.5.3) or the 1D approximation (§2.3.5.4) to the 2D dynamic black, depending on which model had a lower variance for a given quarter and channel. In cases where the Dynablack model was computed, but not used by the pipeline, there will be both Dynablack and 1D black models from which the user may choose. Users can determine which black algorithm was used for a particular data set from the FITS keyword `BLKALGO` (with values: “exponentialOneDBlack,” or “dynablack”). Note that the static 2D black and the 1D black are designed to be used together in that the 1D black is a fit to the residual signal after subtracting the static 2D black. The Dynamic 2D black model includes the mean black signal along with the spatial and temporal dependence of the black. The short-duration quarters Q0, Q1, and Q17 use the 1D black for all channels.

The static 2D black includes a constant electronic offset that has been added to the CCD voltage to ensure that only positive signals are input to the analog-to-digital converter (ADC). The 2D black model consists of a full-frame image for each of *Kepler*'s 84 detector channels and is stored in the file `kplr2015238_2dblack.fits`. The FITS file format follows the same format as *Kepler* full-frame star images: an empty primary data HDU followed by 84 image extensions. The 2D black model is in units of counts (or DN) per read and is scaled to be appropriate for a single CCD frame. Therefore, multiply the model by 9 to apply it to short-cadence data consisting of 9 frames per cadence (FITS keywords `NUM_FRM`, `NREADOUT`), and multiply the model by 270 for use with long-cadence data. To correct an image for the 2D black signal, first shift the raw data as described under `RAW_CNTS` in §2.3.2, then subtract the scaled 2D black model from the shifted raw data pixel-by-pixel. The 2D black uncertainty (`kplr2015238_2dblack-uncert.fits`) contains the single-frame 1σ uncertainty for each pixel's black value and follows the same file format.

The 2D black correction is the first CCD pixel calibration after the pixel values are corrected for the fixed offset and mean black levels as part of the onboard requantization and compression. The 2D black signal has some dependence on the local detector electronics (LDE) temperature. The static 2D black model was generated from 25 unilluminated full-frame images collected at a temperature within the *Kepler* flight temperature range. The 2D black model is appropriate for use with all *Kepler* science data. Low-level thermal effects are corrected for in later calibration stages (see §2.3.5.4 1D Black and

§2.3.5.3 Dynablack for details). Quintana et al. (2010) describe all of the steps involved in *Kepler* pipeline pixel calibrations.

Because of the synchronized readout of the *Kepler* science CCDs and the simultaneous readout of the FGS CCDs sharing the same local detector electronics, there are spatially dependent signals in the black level that vary from channel-to-channel. The 2D black FC model contains four primary components:

1. a scalar component corresponding to the mean black level of the channel,
2. a smoothly varying row-wise spatial component,
3. a smoothly varying column-wise ringing at the start of each row,
4. and a high-spatial frequency FGS clocking crosstalk pattern (see KIH §5 & §6 for details).

2.3.5.3 Dynamic 2D Black

The dynamic 2D black represents the time and temperature variation of the black level (Kolodziejczak et al. 2010), and is used in DR 24 and later when it gives a better fit to the data for a given quarter and channel than the 1D approximation (§2.3.5.4). This dynamic 2D black is calculated in the pipeline’s “Dynablack” module. Because of the complexity of the model and the data needed to compute it, we have provided MATLAB code and data files (“Dynablack blobs”) that allow the user to compute the dynamic 2D black level for all pixels for all valid long cadences in Q2-Q16, even in cases when the pipeline selected the 1D black (§2.3.5.4) as a better fit. For further details on the Dynablack model and the fitting process see Kolodziejczak et al. (2010).

The Dynablack blobs are packaged per quarter and channel. The dynamic 2D black model units, raw data shift, and scaling with number of reads are the same as for the static 2D black (§2.3.5.2). The naming convention for the blob files is:

```
kplr<YYYYDOYHHMMSS>-q<##>-<mno>-dr<##>_dynablack.mat
```

where YYYYDOYHHMMSS is the time stamp associated with the data, q## is the quarter number, mno is the module and output number, and dr## is the data release number (e.g., kplr2012004120508-q11-241-dr25_dynablack.mat, for quarter 11, module 24, output 1, data release 25).

The Dynablack model fits for the thermal dependence of the 2D black structure and it includes the thermal dependence of the fine-guidance sensor clocking cross-talk pixels (see §2.3.5.1). The Dynablack blob file incorporates information from the long-cadence data (§2.3.2), reverse-clocked flight data (§2.3.9), full-frame images (§2.3.3), artifact removal pixels (§2.3.8), and ancillary engineering data (§2.3.10). The black retrieval tool uses default configuration parameters to reconstruct the Dynablack models, matching the nominal configuration for the *Kepler* pipeline. The fit for each valid long cadence is assigned the mid-cadence modified Julian date (MJD). Black estimates for any other MJD are interpolated from the fitted ones, which allows the dynamic black to be used in correcting short cadence data. There are no Dynablack fits for any module/output channel in the short Quarters 0, 1, and 17, or for any output on module 3 in the quarter in which it failed (Q4).

The MATLAB function call to recover the dynamic black level from a give blob file is:

```
>> [ black, deltaBlack, initializedModels, relativeIdx ] =  
extract_blacks_from_dynablack_blob( blobFile, rows, cols, mjds );
```

Inputs:

- `blobFile` = [string]; full path to .mat file containing the Dynablock blob
- `rows` = [int]; nPixels x 1 array containing the one-based CCD row coordinates of the pixel blacks to retrieve
- `cols` = [int]; nPixels x 1 array containing the one-based CCD column coordinates of the pixel black values to retrieve
- `mjds` = [double]; nCadences x 1 array of MJD timestamps on which to return the pixel black values

Outputs:

- `black` = [double]; nPixels x nCadences array containing the pixel black values in DN/read. Typical values are ~720 DN/read.
- `deltaBlack` = [double]; (optional) nPixels x nCadences array containing the uncertainties on the pixel black values in DN/read
- `initializedModels` = [struct]; (optional) structure containing models read from the Dynablock blob and initialized for use in `dynablock_reader.m`, which can be used for faster subsequent calculations with the same model.
- `relativeIdx` = [double]; (optional) vector of MJDs converted to relative cadence indices, for use in `dynablock_reader.m`

The black levels and uncertainties returned by `extract_blacks_from_dynablock_blob.m` are in counts (or DN) per read. In order to apply them to LC or SC data, they should be multiplied by 270 or 9 respectively.

Example 1: Return the black level and uncertainty from module 24, output 1 in quarter 13 for the pixels in the rectangular patch from rows 101 to 104, columns 501 to 503 and MJD time stamps: 56015.2379, 56015.25830, and 56016.2787 using the Dynablock blob file `kplr2012179063303-q13-241-dr25_dynablock.mat`. First define arrays for the pixel row and column values and timestamps, then call the function:

```
>> rows = [101;102;103;104;101;102;103;104;101;102;103;104];
>> cols = [501;501;501;501;502;502;502;502;503;503;503;503];
>> mjds = [56015.2379; 56015.25830; 56015.2787];
>> [black, deltaBlack, initializedModels, relativeIdx] =
extract_blacks_from_dynablock_blob
('kplr2012179063303-q13-241-dr25_dynablock.mat', rows, cols, mjds);
```

The call returns 12 x 3 arrays of `black` and `deltaBlack` values for each of the 12 pixels and three cadences. The black values in DN/read for the pixel at $(row,col) = (101, 501)$ are

```
>> black(1,:) = 716.2078 716.2033 716.2064
```

Note that if you read the blob file directly into MATLAB, without using `extract_blacks_from_dynablock_blob`, you will get these expected warnings:

```
Warning: Element(s) of class 'ModelComponent' do not match the current
constructor definition. The element(s) have been converted to
structures.
```

```
Warning: Element(s) of class 'DataSubset' do not match the current
constructor definition. The element(s) have been converted to
structures.
```

`extract_blacks_from_dynablock_blob` will nonetheless run correctly as shown above.

Example 2: Return an image of all pixels in the dynamic 2D black for a single cadence, using `dynablack_reader`, the `relativeIdx` and `initializedModels` from a previous call to `extract_blacks_from_dynablack_blob`, and the MATLAB `meshgrid` and `reshape` functions:

```
>>[colGrid, rowGrid] = meshgrid(1:1132, 1:1070);
>>rowV = reshape(rowGrid,1070*1132,1);
>>colV = reshape(colGrid,1070*1132,1);
>>[black, deltaBlack] = dynablack_reader(rowV, colV, relativeIdx(100),
    initializedModels);
>>blackImage = reshape(black,1070,1132);
>>imagesc(blackImage)
```

While most of the details in the optional output “initializedModels” are beyond the scope of this document, it can be used to obtain the list of long cadence MJD timestamps for the entire quarter for which the Dynablack fits were performed; as given by:

```
>> lcMjd = initializedModels.midTimestamps(initializedModels.FCLC_list)
```

The Dynamic Black Retrieval Tool software package includes the following MATLAB functions and classes. Users should either run this tool from the top-level directory, or place this directory first in the MATLAB path. Proper operation of this tool depends on the three classes listed. Keeping the directory structure intact with the class directories as sub-directories of the top-level directory ensures these classes are available to any function run in the top-level directory.

Listing of Retrieval Tool functions and required classes:

black-retrieval-tool (top level directory):

```
    dynablack_reader.m
    extract_blacks_from_dynablack_blob.m
    initialize_dynablack_reader.m
```

@modelClass:

```
    FGSFrame_components.m
    FGSParallel_components.m
    get.m
    Horizontal_errors.m
    set.m
    Vertical_errors.m
    FGSFrame_errors.m
    FGSParallel_errors.m
    Horizontal_components.m
    modelClass.m
    Vertical_components.m
```

@modelComponentClass:

```
    assemble.m
    get.m
    modelComponentClass.m
    set.m
```

@modelParametersClass:

```
    estimates.m
    get.m
    initialize.m
    modelParametersClass.m
    set.m
```

2.3.5.4 1D Black Correction

The row-dependent 1D black model is a correction added to the static 2D black (§2.3.5.2) as an approximation to the 2D dynamic black (§2.3.5.3) when the 2D dynamic black model is noisy, fails to converge, or is unavailable. Because of the complexity of the model and the data needed to compute it, we have provided MATLAB code and data files that allow the user to compute the dynamic 1D black level for all pixels for cadences and channels for which the 1D black correction was selected over the 2D dynamic black model. The MATLAB function `extract_one_d_black_from_blob.m` uses a binary file from the final pipeline processing (DR25) for the specified detector channel (module/output) and quarter number to return the one-dimensional black level and its uncertainty for a set of pixel rows. Below we describe the 1D black data format (“1D black blob”) and the use of the black retrieval tool to retrieve requested values.

The 1D correction is determined for each long cadence by fitting a six-component row-dependent model to the trailing black collateral pixels after the static 2D black correction has been applied. The 1D black model includes scalar and linear components for the masked smear rows (1-20), and a scalar, linear, and two exponential components for the image portion of the array (rows 21-1070). The decay constants for the exponential terms for each detector channel were determined from flight data and are fixed in the 1D black model as shown in the function `exponential_constants()` in the file `extract_one_d_black_from_blob.m`. The 1D black fit coefficients, their covariance, and associated cadence timestamp information are packaged in a MATLAB data structure and stored as MATLAB `.mat` files (“1D black blobs”). Each 1D black blob contains fit information for a single channel over a single quarter. Such blobs exist for each channel/quarter pair where the default black correction algorithm (BLKALGO = “exponentialOneDBlack”) was selected for Data Release 25 processing. The naming convention for the blob files is:

```
kplr<YYYYDOYHHMMSS>-q<##>-<mmo>-dr<##>_1dblack.mat
```

where YYYYDOYHHMMSS is the time stamp associated with the data, q## is the quarter number, mmo is the module and output number, and dr## is the data release number (e.g., `kplr2012088054726-q12-092-dr25_1dblack.mat`, for quarter 12, module 9, output 2, data release 25).

The MATLAB function call to recover the 1D black from a give blob file is:

```
>> [black, deltaBlack, oneDModel] =  
    extract_one_d_black_from_blob(blobFile, rows, mjds, interpMethod)
```

Inputs:

- `blobFile` = [string]; full path to `.mat` file containing the 1D black blob
- `rows` = [int]; nRows x 1 array containing the one-based CCD row coordinates of the 1D black corrections to retrieve (1-1070)
- `mjds` = [double]; nCadences x 1 array of modified Julian day timestamps on which to return the 1D black corrections
- `interpMethod` = (optional) [string]; interpolation method to be used when interpolating 1D black corrections over MJDs ('linear' or 'nearest'); default = 'linear'

Outputs:

- `black` = [double]; nRows x nMJDs array containing the 1D black correction in DN/read. Typical values are a few DN with a spatial and temporal variation of a few tenths of a DN per read.

- `deltaBlack` = (optional) [double]; nRows x nMJDs array containing the propagated uncertainty in the 1D black correction in DN/read
- `oneDModel` = (optional) [double]; nCCDRows (1070) x nCoeffs (6) array containing the 1D black model for all CCD rows

This tool returns the fitted 1D black values and optionally their uncertainties at the requested CCD rows and MJDs from the 1D black blob for the selected channel. An optional input argument determines the interpolation method applied. The applied model design matrix (`oneDModel`) is also optionally returned. For the masked smear rows 1-20, the first four columns of `oneDModel` are zero and the 5th and 6th columns are the constant and linear functions. For the rest of the image (rows 20-1070), the first four columns are the constant, linear, long exponential, and short exponential functions, while the 5th and 6th columns are zero. The black levels and uncertainties returned by `extract_one_d_black_from_blob.m` are in counts (or DN) per read. In order to apply them to long- or short-cadence data, they should be multiplied by 270 or 9, respectively.

The *Kepler* pipeline pixel calibration module (CAL) fills data gaps for the 1D black correction using nearest-neighbor interpolation of the fitted coefficients from ungapped timestamps. The user should note that using linear interpolation for MJDs greater than 15 minutes past the endpoints of the fitted cadence time stamps may result in unrepresentative 1D black corrections. It is best to use a “nearest neighbor” interpolation for such points.

The 1D black correction for CAL short cadence processing is a linear interpolation of the 1D black correction for long cadence from ungapped long cadence timestamps plus an additional stochastic bias correction per cadence. This correction is equal to the robust mean of the trailing black collateral data after subtraction of the model without this stochastic term, thus forcing model errors to a mean of zero in the trailing black collateral data for each cadence. These bias terms are not available in the 1D black blob and should be determined by the user for short cadence data using a similar method.

2.3.5.5 Scene-dependent Rows

Bright stars at high CCD column numbers complicate the 2D and 1D dynamic black model fits, as undershoot from these bright stars corrupts the black level measured in the trailing black region. In order to avoid corrupting the fitted model, the pipeline excludes the rows containing bright stars in columns > 1000 from the fit (Kolodziejczak et al., 2010). Since *Kepler* viewed the same scene for the entire mission, we developed a fixed listing of excluded rows for each channel and season combination. The “scene-dependent” row table is included as part of the MATLAB code used to extract the 1D black model, but is not necessary to use the 1D black model provided. We provide the MATLAB function `is_scene_dependent.m` to map the scene dependent CCD rows for each channel/season pair. In addition, the scene-dependent table is delivered as a flat text file: `kp1r2015279231411-dr25_blackrowmask.txt`. The black-row-mask file indicates rows that are potentially corrupted and should be excluded from the 1D black model fit if users wish to do their own 1D black model fits. The format is a comma-separated file with 336 rows and 1072 columns. There is one row for each channel and season, with the channel number (1-84) and season (0-3) indicated by the first two entries in each row, respectively. The remaining entries of each row consist of 1070 flags indicating whether the corresponding row was used in the model fit (0) or excluded due to the presence of a bright star (1).

The MATLAB function call to recover scene dependent rows is:

```
>> [isSceneDep] = is_scene_dependent(channel, season, oneBasedRows)
```

Inputs:

- channel = [int]; *Kepler* channel number (1-84)
- season = [int]; *Kepler* season (0-3)
- oneBasedRows = [int array]; nRowsx1 array of one-based CCD row numbers (1-1070)

Outputs:

- isSceneDep = [logical array]; nRowsx1 array corresponding to OneBasedRows. true = yes, this row is scene dependent, false = otherwise

This function determines whether or not CCD rows contain scene dependent signals corrupting the trailing black collateral data given the channel and season for the original *Kepler* FOV.

While the MATLAB function `extract_one_d_black_from_blob.m` along with the blob files are sufficient to extract the fitted 1D black for any *Kepler* long cadence data, we describe the contents of the 1D black blobs for the ambitious user who may wish to perform their own analysis of the fitted coefficients using the scene-dependent row model. Each 1D black blob is a MATLAB .mat file containing a single variable, `inputStruct`. The field contents in the following example are for module/output 17/2 for a quarter containing 4044 long cadences. There are six components to the row dependent 1D black model: constant, linear, long exponential, short exponential, masked smear constant, and masked smear linear.

```
inputStruct =
    cadences: [4044x1 double]           % long cadence number
    startTimeStamps: [4044x1 double]    % cadence start time (mjd)
    midTimeStamps: [4044x1 double]      % cadence midpoint time
    endTimeStamps: [4044x1 double]      % cadence end time
    module: 17                          % CCD module
    output: 2                            % CCD output
    blackCorrectionStructLC: [1x1 struct] % fit coefficient structure

inputStruct.blackCorrectionStructLC =
    timestamp: [4044x1 double]          % fit cadence time stamp (same as
midTimeStamps)
    gapIndicators: [4044x1 logical]     % cadence gap indicators
    original: [4044x6 double]           % fitted ond-d-black model
coefficients; nCadences x nModelComponents
    originalCovariance: [4044x6x6 double] % model coefficient covariance
    smoothed: [4044x6 double]          % smoothed version of fitted
coefficients
    smoothedCovariance: [4044x6x6 double] % smoothed version of coefficient
covariance
```

2.3.5.6 Saturated Smear Columns

The *Kepler* photometer operates without a shutter, so all stars leave a “smear” signal during CCD readout. The CCDs are operated in such a way that this smear signal is measured in two locations for each column: the “masked” region under an aluminum mask covering the first twenty physical CCD rows and the “virtual” region composed of 26 over-clocked rows at the end of each readout (KIH, §4.5). Quintana et al. (2010) give further details on the smear correction implemented in the pipeline.

The pipeline normally averages the masked- and virtual-smear measurements to determine the smear signal for each column. However, in some cases, the presence of a bright saturating star corrupts the associated measurement, and causes charge to spill across CCD rows. Typically tens of rows are affected, but for very bright targets the spilled charge can affect hundreds of rows. If these targets are sufficiently close to the masked- or virtual-smear regions, then the proximate smear pixels are affected.

Because the downlinked smear pixel data for each column are coadded on-board over 12 smear rows, saturated smear pixels can lead to rollover in the 23-bit accumulation register so that downlinked values are not representative of the total accumulated charge.

In order to avoid using these corrupted smear values in the smear correction, the FFI and long-cadence data are used to identify columns where saturated targets are bleeding into the masked- or virtual-smear regions. This is accomplished by searching the long cadence collateral smear data for spatial outliers by thresholding the smear difference (masked – virtual). The corrupted smear may be due to a bright saturating star on the target channel, or via video crosstalk, on any of the other three output channels on the same target module (see §2.3.5.18). The undershoot signal from the saturating star (see §2.3.5.10) generally corrupts some additional smear columns, so the bad smear values tend to come in clusters of both columns and output channels within a module. The full-frame images (FFI) are used to unambiguously locate the responsible star and flag either the masked or virtual smear, as appropriate. By searching all of the long cadence collateral data and the FFIs taken in a given quarter, variable stars that cause intermittent smear corruption are identified. The tool below provides the mapping of flagged bad smear columns for each CCD channel and observing season as used in the final data release (DR 25).

The saturated smear columns may be retrieved for any *Kepler* observing season/channel using the MATLAB function `is_saturated_smear.m`.

```
[isSaturatedMasked, isSaturatedVirtual] = is_saturated_smear_column(channel, season, columns)
```

Inputs:

- `channel` = [int]; *Kepler* channel number (1-84)
- `season` = [int]; *Kepler* season (0-3)
- `columns` = [int array]; nCols x 1 array of one-based CCD column numbers (1-1132)

Outputs:

- `isSaturatedMasked` = [logical array]; nCols x 1 array corresponding to columns. true == yes, this column is saturated in the masked smear region, false == otherwise
- `isSaturatedVirtual` = [logical array]; nCols x 1 array corresponding to columns. true == yes, this column is saturated in the virtual smear region, false == otherwise

Example-1: Retrieve the saturated smear columns for channel 5 (module 3, output 1) season 0 (Summer). First define a column vector of column numbers, then call the function.

```
>> columns = [1:1132]';  
>> [isSaturatedMasked, isSaturatedVirtual] = is_saturated_smear_column(5, 0, columns);
```

The call returns two 1132x1 logical arrays with the following indices set to “true”

```
>> masked = find(isSaturatedMasked)  
>> masked =  
    338  
    339  
    340  
    341  
    342  
>> virtual = find(isSaturatedVirtual)  
>> virtual =  
    504  
    505
```

2.3.5.7 Read Noise

Electronic noise sources intrinsic to the CCD on-chip amplifiers and off-chip analog signal chain are referred to collectively as read noise, which is well-represented by a white Gaussian noise that is included in the signal every time the CCD is read out. Read noise differs for each module/output, but is relatively constant with time. The delivered read noise model is derived from a ground test of the array in Aug-Sep 2008. The measured FFI noise values were scaled for the number of coadds, and quantization noise was removed, to give the read noise per read in units of DN. The values are the same as those shown in Table 13 of the KIH.

The model consists of one “|”-separated text file for the entire focal plane, with one line per module/output, with the following columns:

1. MJD -- Modified Julian Day, start time for model use
2. Module
3. Output
4. Read Noise

MJD is repeated on each line, but in practice these start times are the same for all lines. The file name is `kplr2008102416_read-noise.txt`, with the first line reading:

```
54503.936910|2|1|0.790
```

2.3.5.8 Gain

The system gain is the number of electrons per CCD digital number (DN), which is assumed to be constant with electron count. For small signals, the conversion of electrons to photons is also approximately constant, with deviations captured by the linearity model (§2.3.5.9).

The gains were calculated using the "photon transfer" method in the same test as the linearity model (§2.3.5.9). Flat field images were taken at varying levels of illumination and the dark-subtracted mean level compared with the variance in 160x160 pixel non-overlapping regions. The variance is linearly proportional to the photoelectron count, so the slope of the variance as a function of the mean level, in units of DN, gives the gain in electrons/DN. The values are the same as those shown in Table 13 of the KIH. As indicated in Table 2-7, more detail can be found in §2 of KADN-26196 (Van Cleve, 2016a).

The model consists of one “|”-separated text file for the entire focal plane, with one line for each of the 84 CCD module/output, and the following columns:

1. MJD -- Modified Julian Day, start time for model use
2. Module
3. Output
4. Gain

The file name is `kplr2008072318.txt`, with the first line reading

```
54671.556|02|1|111.230
```

The gain model should be used in combination with the linearity model (§2.3.5.9) in order to convert DN to photons in *Kepler* pixel data.

2.3.5.9 Linearity

Linearity is the ratio of the small-signal linear approximation of the detector photon transfer curve (DN output vs. photons input) to that observed. The linearity is represented by a polynomial fit to the ratio as a function of input DN scaled to a single read. That is, the input DN values range from 1 to 16383 ($2^{14} - 1$). In practice, the *Kepler* CCDs saturate before the maximum analog-to-digital converter (ADC) value, so the largest DN in the data is $\sim 10,000$.

Linearity measurements of the science channels were made during ground testing of the Detector Array Assembly at various LDE operating temperatures; the current results were collected at 23 C. A 660 nm pulsed LED light source flat-field flood-illuminated all modules simultaneously, and the LED pulse width was used to precisely vary the photon count. This photon transfer data set was inverted and differentiated with respect to the observed signal. The resulting reverse differential transfer function was fit to a quadratic with the constraint of zero derivative at the maximum observed flood illumination signal C_{obs_max} . The fit parameters were then transformed and integrated to give the standard polynomial fit format for delivery, and covariances calculated using a Monte Carlo method. As indicated in Table 2-7, more detail can be found in KADN-26196 (Van Cleve, 2016a).

To correct an image for gain and non-linearity, divide the dark-subtracted image by the number of reads NUM_FRM in order to get the observed signal C_{obs} in units of DN/read. The correction is then:

$$C_{lin} = \left(\sum_{n=1}^N A_n (hC_{obs})^{n-1} \right) C_{obs}$$

where C_{lin} is the corrected signal in DN/read, h is the scale factor (`scalex` in the table below), and the A_i are the polynomial coefficients. Offset and origin shifts are not shown, since they are identically zero in this delivery.

At this point the corrected DN values can be multiplied by the scalar gain (§2.3.5.8) for the appropriate module/output to get electrons, then multiplied by the number of coadds to get the number of electrons in the original image. The polynomials are valid over the range from 0 to C_{max_domain} , which is the median of pixel values observed in the *Kepler* FOV between the well depth (as defined in KADN-26196, Van Cleve, 2016a) and C_{obs_max} . While the behavior of individual pixels becomes complex for observed signals between C_{max_domain} and C_{obs_max} , as charge begins to bleed into adjacent pixels, the sum of pixels in a photometric aperture is much simpler. As stated in Gilliland *et al.* (2011), “Saturation occurs at about $K_p = 11.5$, but conservation of charge is preserved beyond saturation and excellent photometry is easily recovered from targets up to 7 magnitudes beyond saturation by including all the pixels into which charge bleeds.”

The linearity model consists of a “|”-separated text file with one line for each CCD module/output. The following 84 lines – one per mod/out— contain linearity polynomial coefficients and covariance matrix values for the coefficients as shown in Table 2-8:

Table 2-8: Contents of Linearity Model

Column	Contents	Description	Data Type
1	MJD	Modified Julian Day, start time for model use	Float
2	module	The module for this gain value	Integer
3	output	The output for this gain value	Integer
4	order	The order of the polynomial fit (= 2 for all)	Integer
5	type	The type of polynomial used (= ‘standard’ for all)	String
6	xindex	Not used	Integer
7	offsetx	Not used	Float
8	scalex (<i>h</i>)	Scaling parameter for polynomial fit (= 0.01 for all)	Float
9	originx	Not used	Float
10	max_domain	Maximum input at which polynomial fit is valid	Integer
11-13	coefficients A_1, A_2, A_3	Polynomial coefficients in ascending polynomial order	Float
14-22	covariances	9 columns of covariance terms for the polynomial coefficients, ordered cov11 cov12 ... cov13 cov21 ... cov33	Float

The parameters MJD, order, type, xindex, offsetx, scalex, and originx are identical for each of the 84 channels. The file name is `kplr2009060215_linearity.txt`, with the first line reading:

```
54896.500|02|1|2|standard|-1|0.0|1.0e-02|0.0|10116|1.0184|-4.9214E-04|1.4412E-
06|5.6182E-07|-1.4324E-08|4.1947E-11|-1.4324E-08|4.4952E-10|-1.3164E-12|4.1947E-
11|-1.3164E-12|3.8552E-15
```

2.3.5.10 Undershoot

Hardware limitations of the Local Detector Electronics (LDE) analog signal chain introduce an undershoot/overshoot at the output, which is digitized. Each pixel read out is then the weighted sum of the original pixel value in the CCD and previous pixels in the same read out row. Undershoot is particularly noticeable at sharp steps in the signal, such as saturated columns. Undershoot can be modeled as a linear digital filter, and thus, can be corrected by applying the corresponding inverse digital filter. See §6.6 of the KIH for further description and KADN-26197 (Van Cleve, 2016b) for a detailed discussion of the model.

The forward model, available from the archive, transforms the “true” signal **X** coming from the CCDs into the observed signal **Y**, which contains the undershoot. The reverse model transforms the observed signal into the “true” signal without undershoot. The forward model file consists of a single “|”-separated text file for the entire focal plane, with one line for each module/output with the columns shown in Table 2-9:

Table 2-9: Contents of Undershoot Model

Column	Contents	Description
1	MJD	Modified Julian Day, start time for model use
2	module	The module for these undershoot
3	output	The output for these undershoot coefficients
4	N_b	The number of the feedforward (numerator) coefficients (equal to the feedforward filter order +1) = 20 for all
5 to 24	B_i	numerator (B) filter coefficients
25 to 44	σ_i	B filter coefficient uncertainties

The modified Julian day is used to denote the time at which the model becomes applicable. $N_b = 20$ for the table in MAST. The file name is `kplr2008102809_undershoot.txt`.

The MATLAB usage of the forward model to show the effect of undershoot with coefficients \mathbf{B} on the input or “true” signal \mathbf{X} is

```
>Y = filter(B,1,X),
```

where the MATLAB filter function in this simplified case implements

$$y_n = \sum_{i=1}^{\min(n,N_b)} B_i x_{n-i+1}$$

Then “true” pixel values from columns $n-19$ to n inclusive are used to calculate the undershoot-affected pixel value in the n th column, for $N_b = 20$. To reverse the undershoot and correct data row-by-row one should use:

```
>X = filter(1,B,Y).
```

See the MATLAB documentation for more about how the filter function works; note that the MATLAB documentation refers to a coefficient vector \mathbf{B} which has $nb + 1$ elements, where “nb” is the feed-forward filter order. In our notation the number of feed-forward coefficients $N_B = nb + 1$. Users not using MATLAB may correct the undershoot by implementing this recursive formula:

$$x_n = \frac{1}{B_1} \left[y_n - \sum_{i=2}^{\min(n,N_b)} B_i x_{n-i+1} \right],$$

which is numerically faster than representing the forward model as a lower-triangular matrix $\mathbf{Y}=\mathbf{B}\mathbf{X}$ and inverting it to do the correction: $\mathbf{X}=\mathbf{B}^{-1}\mathbf{Y}$. The inverse exists and is numerically robust.

Users of the undershoot file to correct raw cadence data in the target pixel files will have to make sensible assumptions about pixels upstream of the collected pixels. For example, the CAL module of the Pipeline uses FFIs for approximate upstream pixel content. If the filter is applied without these assumptions, it is equivalent to assuming that the preceding (unseen) pixel stream is all zeros, which may not be desirable.

2.3.5.11 Large-Scale Flat

Because of the difficulties in uniformly illuminating *Kepler*'s large field of view over the 0.95-meter aperture, suitable flat field data that include the full photometer and optics were not available from ground testing. Therefore, the flat field correction for *Kepler* is divided into two parts: 1) a large-scale flat representing the telescope vignetting function (this section and KIH §3.6) and 2) a small-scale flat representing the local pixel response non-uniformity (PRNU; §2.3.5.12 and KIH §4.12, §4.14). Both the large- and small-scale flat field values represent the relative sensitivity of each pixel to uniform illumination. The large-scale flat is used as part of the response variation correction for *Kepler*'s CCDs during pixel-level calibration (Quintana et al. 2010). The large-scale flat consists of a full-frame image for each of *Kepler*'s 84 detector channels. The FITS file format follows that for *Kepler* full-frame images (§2.3.3): an empty primary data HDU followed by 84 image extensions. The file name is `kplr2015236_largeflat.fits`. In order to correct an image for the large-scale flat field response, the user should divide the data image by the large-scale flat image pixel-by-pixel. Because the large-scale flat file is based on an analytic model, there is no associated uncertainty file.

The large-scale flat is a 2D polynomial fit to a vignetting model that accounts for the 0.95 meter aperture of the Schmidt telescope and the central obscuration from the *Kepler* focal plane and local detector electronics. Details of the vignetting model are given in KIH §3.6. The large-scale flat gives the relative response of the pixels on each CCD module/output. Pixels near the center of the focal plane where there is no vignetting have a large-scale flat response of 1. Pixels from the non-imaging areas of the CCD have large-scale flat values set to 1. The 5th order 2D polynomials used to represent the large-scale flat only capture low spatial frequency effects. Pixel-to-pixel response variations are captured by the small-scale flat. For a complete flat field correction, *Kepler* data should be divided by both the small-scale flat and the large-scale flat.

2.3.5.12 Small-Scale Flat

The small-scale flat is used as part of the response variation correction for *Kepler*'s CCDs during pixel-level calibrations (Quintana et al. 2010). The small-scale flat consists of a full-frame image for each of *Kepler*'s 84 detector channels. The FITS file format follows that for *Kepler* full-frame images (§2.3.3): an empty primary data HDU followed by 84 image extensions. The file name is `kplr2015236_smallflat.fits`. In order to correct an image for the small-scale flat field response, the user should divide the data image by the small-scale flat image pixel-by-pixel. The small-scale flat uncertainty contains the one-sigma relative uncertainty for each pixel's flat value in the same format. The file name is `kplr2015236_smallflat-uncert.fits`.

Small-scale flat values are relative to the mean pixel sensitivity; that is, the mean of all the flat pixel values is defined to be 1. Pixels from the non-imaging areas of the CCD have small-scale flat values set to 1 and their uncertainties set to NaN. For a complete flat field correction, *Kepler* data should be divided by both the small- and large-scale flats.

The small-scale flat was computed by comparing the response of each pixel with the robust average of the response of its neighbors from a 9x9 box centered on the pixel of interest. The flat is derived from a series of ground-based test images with quasi-uniform illumination collected prior to assembly of the photometer. The small-scale flat images are constructed from two-exposure coadds (`NUM_FRM = 2`) with an exposure time of 2.08 seconds (`INT_TIME`) and the standard *Kepler* readout time of 0.519 sec (`READTIME`). The number of useful cadences per flat image varies from five to seven depending on the channel and is recorded in the FITS keyword `NCADPFLT` in the image extension header for each channel. The small-scale flat does not include any of the response variations caused by *Kepler*'s flight

optics (Schmidt corrector, primary mirror, field flattener lenses), as they comprise the large-scale flat (§2.3.5.11).

Because of the illumination pattern in the ground test used to generate the small-scale flat, a small number of pixels will have anomalously low or high values. In order to prevent computational problems, the minimum pixel value has been clipped to 0.333 in the small-scale flat. Uncertainties for these clipped pixels should be regarded with caution. The computed mean over pixels on a channel differs slightly from 1.0 due to these clipped pixel values and due to outlier pixel rejection during computation of the small-scale flat. The maximum deviation of the mean from 1.0 is -4×10^{-4} , resulting from three bad columns on channel 43. The *Kepler* pipeline did not use an explicit bad-pixel model for pipeline processing (see §2.3.5.13). However, the small-scale flat can be used to identify out-of-family pixels with unusually low or high photoresponse. For instance, defining a response threshold of $\pm 10\%$ (i.e., pixels with values less than 0.9 or greater than 1.1) results in 15195 flagged pixels, an average of about 180 per channel.

2.3.5.13 Invalid Pixel List

Invalid pixels are defined as those having various anomalous properties, which either reduce signal or add noise to images and light curves. In typical astronomical applications, such pixels are either excluded from processing, or have their values replaced by those of neighboring good pixels. For high precision photometry, however, removing or replacing invalid pixels introduces its own systematic errors as image motion changes the amount of flux falling on the removed or replaced pixel. In the *Kepler* pipeline, these systematic errors were considered worse than the effects of including anomalous pixels, and an invalid pixel list was never used.

Users wishing to contemplate the removal of invalid pixels may refer to the small-scale flat (KAM 2.3.5.12) to identify pixels with anomalously high or low photoresponse, and exclude them from their calculations. While pixels with anomalously high dark current were identified before launch, they were rare (no more than 0.1% on any channel), and no measurements were done after dust cover ejection to identify pixels that developed high dark current while on-orbit. Measuring dark current in a shutterless system viewing a complex scene is a research project in itself, and was never attempted for *Kepler*.

2.3.5.14 Pointing Model

The pointing model is calculated by the pipeline (Middour et al. 2010, Jenkins et al. 2016) in the routine `compute_pointing_model.m` to emulate the action of the spacecraft's fine guidance sensors (FGS) in the presence of velocity aberration. The routine adjusts the vehicle's attitude to minimize the centroid errors between the apparent (aberrated) positions of guide stars and their assigned fiducial positions on the FGS as the velocity aberration changes with orbital position. In other words, the pointing model approximates the transformation between body-centered spacecraft coordinates and the center of the aberrated FOV such that the focal-plane assembly achieves the desired optical pointing as a function of time. In the mechanical frame of the spacecraft, these adjustments have an amplitude of $\sim 40''$ peak-to-peak.

The pipeline uses this pointing model, along with the focal plane geometry (§2.3.5.16) and roll-time (§2.3.5.15) models, to convert between sky and CCD coordinates. This coordinate transformation is combined with the PRF model (§2.3.5.17) in the target aperture definition (TAD) process used to determine which pixels to collect and downlink for each target (Bryson et al. 2010). In the sky-to-CCD calculations, the positions of targets are first aberrated, so that all subsequent calculations are done in this aberrated frame with respect to the origin given by the pointing model. This scheme permits calculation of the “differential velocity aberration (DVA)”, which is the change in the angular distance

between two points in the focal plane due to aberration. Note that DVA is much more important across *Kepler*'s large FOV, than in narrow-field imagers like HST. DVA varies between targets at the center and edge of the FOV by up to 0.6 pixels over an observing quarter even though the FPA center is fixed in RA and Dec by the FGS (Bryson et al. 2010). While these corrections are important for data collection and initial data processing, these model-derived apertures were replaced by data-driven optimal apertures for the final data release (DR25). Since the model does not include the ad-hoc attitude adjustments made at the start of some observing periods, which can be of the same order of magnitude as the maximum DVA, users are advised to use similar data-driven methods when updating aperture definitions during their own data reanalysis. Hence, this pointing model is provided primarily as historical documentation of *Kepler* methods, along with the pipeline code (KDPH, Jenkins et al. 2016).

Users may note that target pixel and FFI files contain World Coordinate System (Greisen and Calabretta, 2002) headers (§2.3.2, §2.3.3) calculated from motion polynomials using an algorithm described in detail in KADN-26285 (Van Cleve, 2010). In addition, zero-based locations of channel centers and corners that are accurate to within 1.0 pixels are provided in the simplified (RA, DEC) to pixel mapping file `morc_2_ra_dec_4_seasons.xls` (excerpt shown in Table 12 of KIH). Where these two sets of values provide sufficient accuracy, the pointing reconstruction outlined in this section can be ignored.

There is one pointing model for the entire FOV. The file name is `kplr2013050211_pointing.txt`. Each line in the model file has four single space-separated columns:

1. MJD (sampled every 0.5 days)
2. RA (decimal degrees)
3. Dec (decimal degrees)
4. Roll (decimal degrees) with respect to nominal (unaberrated) for that season, which is one of the four orientations of the spacecraft (see Figure 24 in KIH) that places the focal-plane assembly on the celestial FOV (Figure 1 in KIH). The unaberrated roll orientations for the *Kepler* telescope are 123°, 213°, 303°, and 33° for seasons 0, 1, 2, 3, respectively, which include the spacecraft season-0 orientation (110°), the focal-plane assembly-to-spacecraft offset (13°), and the seasonal quarterly rolls of 90°. See Table 3-1 for the dates of each quarter/season.

All values in all columns are double-precision floating-point values, presented with the maximum precision attainable (i.e., 15 significant digits). The pointing model values and formats changed substantially after the second of *Kepler*'s four reaction wheel failed, ending the primary mission on MJD 56423 = May 11, 2013; the delivered file is the last pointing model computed before that date, and contains predictions up to MJD 57660.

The nominal RA, Dec, and first spacecraft roll angle are 290.6667, 44.5000, and 110 degrees, respectively, as shown in the `rollTimeClass` function `rollTimeClass.m` and the seasonal variation is shown in the `raDec2PixClass` function `get_direction_matrix.m`.

2.3.5.15 Roll Time Model

The roll-time model gives the actual dates in MJD on which the *Kepler* spacecraft was rolled 90° to achieve a new *Kepler* season (or roll orientation). The file name is `kplr2013040800_rolltime.txt`. Collection of valid cadences started between 12 h and 3 d after the roll, depending on mission events. The starting calendar dates for data collection in a season are shown in Table 3-1, and the exact times in the KDCH. The roll-time is used with the pointing model (§2.3.5.14) and the focal plane geometry model (§2.3.5.16) to convert between sky coordinates and CCD coordinates. The roll time file is a comma-separated text file with 17 lines. The first column is the

MJD and the second column is the season starting on that date. The first line of this file is for 18-Mar-2009, after launch but before dust cover ejection, and applies to Q0 and Q1 as Spring seasons.

The roll time model values and the file format changed substantially after the second reaction wheel failure ended the *Kepler* primary mission on MJD 56423 = May 11, 2013. The file delivered contains actual roll times up to the last *Kepler* primary mission roll on April 8, 2013.

2.3.5.16 Focal Plane Geometry

The *Kepler* focal plane geometry model (file name `kplr2009060418_geometry.txt`) consists of a set of coefficients that specify the location, orientation, and optical distortion of each channel with respect to the center of the focal plane array (FPA), which the attitude control system nominally holds at the center of the celestial FOV. The focal plane geometry model is used with the pointing model (§2.3.5.14) and roll-time model (§2.3.5.15) in calculations (including aberration) that transform sky coordinates (right ascension and declination) to/from pixel coordinates (module, output, row, column) in the *Kepler* data processing pipeline. The FPG coefficients were determined in-flight by fitting the positions of a number of stars on each CCD module/output (Tenenbaum and Jenkins, 2010).

The FPG file is a text file with 253 lines. The first line is the MJD at which the model became applicable. Lines 2-43 are space-separated columns containing the Euler 3-2-1 rotation angles (Goldstein, 1980) of the 42 CCD chips (see Figure 23 in KIH) with respect to the FPA axes (see Figure 24 in KIH). These angles are sorted by module, by chip A, and then by B (see Figure 23 in KIH). These Euler angles are applied in the order Z'-Y'-X' in the FOV coordinate system (see Figure 24 in KIH) to convert the RA, Dec, and roll angles given by the pointing and roll-time models to the origin of the chip coordinate system, which is defined as the last column (1100) and the last physical row (1044) plus half the inter-CCD gap width (see below).

Lines 44-85 are three space-separated columns representing minor adjustments to the chip coordinate system representing the as-built focal plane module positions:

1. Rotation in degrees which must be added to the last column in lines 2-43 (in degrees) as a tweak.
2. Half-gap width from the module center in the CCD row direction (in pixels), which varies by a few pixels around its 39-pixel nominal design value. The full physical gap is shown schematically for a single module in Figure 20 of KIH, while the position and orientation of gaps over the whole FPA are shown on the left side of Figure 4 in Tenenbaum and Jenkins (2010).
3. CCD column offset (in pixels), with magnitude of 3 pixels or less.

Lines 86-169 have a single column that defines the platescale in arcsec/pixel at the center of the channel for each of the 84 channels. Lines 170-253 specify the “pin cushion” optical effect (in inverse square arcseconds) for each of the 84 outputs, which causes the plate scale to decrease with radial distance from the center of the module (Tenenbaum and Jenkins, 2010). Note that the formulae in Tenenbaum & Jenkins (2010) §4.2 are written in-line as one would write MATLAB code. Without the pin cushion factor, the conversion of sky angle to pixel position can be off by a full pixel in both row and column.

2.3.5.17 Pixel Response Function

The *Kepler* Pixel Response Function (PRF) is determined from a combination of the *Kepler* optical point spread function (PSF) and various pointing and electronic systematics of the *Kepler* spacecraft during 14.7-minute long exposures. The PRFs were measured only once using data collected during spacecraft commissioning. For more information on the PRF, how it was measured, and how it is used in the *Kepler* pipeline, refer to Bryson et al. (2011).

The PRF varies in size and shape across the *Kepler* focal plane. To describe this variation, the PRF model is comprised of five images per channel as shown in Table 2-10. These images describe the PRF in the four corners and the center of the channel. Each PRF was determined by using 10 stars near the corners (or at the center) of the CCD. For computational simplicity, each PRF may be assumed to be located exactly at the active CCD corner (or center). This approximation has negligible impact on the interpolated values across a channel.

The discrete PRF model is over-sampled by a factor of 50 to allow for sub-pixel interpolation. The model is comprised of a 550 x 550 (or 750 x 750) grid that covers an 11 x 11 (or 15 x 15) pixel array. The size of the PRF depends on the channel. The provided array is large enough to encompass all the light from an unsaturated point source collected by *Kepler*.

There are three primary uses of the PRF:

- 1) To predict (or model) pixel values due to a star at a specified pixel location with a specified magnitude.
- 2) To estimate a star's location based on its pixel values via non-linear PRF fitting. The PRF location and amplitude that produces the smallest chi-squared value between the pixel values estimated by the PRF and the observed pixel values provides the estimated location of the star.
- 3) To estimate the stellar magnitude resulting from the PRF fit.

Users may interpolate between the PRFs at a discrete set of points to capture the spatial variation of the PRF within a channel.

The *Kepler* Project provides the five PRFs for each channel as FITS (Flexible Image Transport System) formatted files with five image extensions. Each channel is stored in a separate file with names formatted as `kplr<module>.<output>_<creation_date>.fits`. The creation date is formatted as `<yyyyddd>`. Note all the PRF files are bundled in a single tar file.

Each image extension within the file contains one of the five PRFs calculated for that channel. The first four image extensions contain the PRF for the four corners of the CCD and the last extension contains the PRF for the center of the array. The LOCATION keyword in conjunction with the table below may be used to determine the location of the center of each PRF. (These detector locations are zero-based and include the collateral row and columns.)

Table 2-10: Row and column values for the five locations in the PRF files.

Location	Column	Row
1	12	20
2	12	1043
3	1111	1043
4	1111	20
5	549.5	511.5

The FITS headers also contain physical WCS coordinates that can be used to determine these locations. The PRF grid spacing, which is the inverse of the over-sampling factor, is specified in the physical WCS keywords in each image header (CDELTA1P and CDELTA2P).

2.3.5.18 Video Crosstalk

Intra-module video crosstalk occurs when the output signal from one channel is electronically coupled to the output signals of the other three channels on the same module as a result of them sharing a wire bundle or circuit board (KIH §6.3; Caldwell et al. 2010; Coughlin et al. 2015). Inter-module video crosstalk, in which signals from different modules are coupled, was reduced to undetectable levels by the staggered module readout within board pairs and the shielding between board pairs in the LDE operations and design. Since the pixels on the four channels of a module are read out synchronously, the intra-module crosstalk signal appears on the pixel with the same row and column coordinates as its source on another channel. The coefficient amplitude is large enough that a legitimate Jupiter-sized transit signal on one channel can induce an artificial Earth-sized transit signal on an associated channel (i.e., same module, different output). Such events are often identified as KOIs and must be culled from the catalog as false positives. The file name for the model is `kplr_2015111701_video_crosstalk_rec_src.csv`.

The assumptions of the video crosstalk model (after Caldwell et al. 2010) are:

1. The coefficient C_{ij} between source channel j and recipient channel i is the same for all pixels in a given source-recipient channel pair,
2. Crosstalk is linear and additive (i.e., the total crosstalk signal in a recipient channel is the product of the source intensity and the crosstalk coefficient summed over the three possible source channels on that module),
3. Crosstalk can be asymmetric, so in general, C_{ij} does not equal C_{ji} ,
4. Crosstalk can be either positive or negative, and
5. All signals (e.g., stellar flux, charge bleeding, undershoot) have the same coefficient.

The analysis method is described in Caldwell et al. (2010). Briefly, FFIs were examined to identify pixels that contained a bright source in one (and only one) source channel and only dim pixels in the recipient channels. For every such pixel, the coefficient C_{ij} was defined as the ratio of the background-removed flux in the recipient channel (i) to the background-removed flux in the source channel (j). The median over all pixels meeting this bright-dim selection criterion constitutes the reported value. This procedure neglects the contribution of recipient channel stellar flux to the recipient channel pixel values, and thereby tends to overestimate the coefficient amplitudes.

The video crosstalk model is not used by the *Kepler* data processing pipeline, but has been used to examine KOIs with the same pixel coordinates as a bright, variable source on an associated channel. A general ephemeris-matching method that rejects false KOIs produced by a variety of contamination processes (Coughlin et al. 2015) has shown that very few KOIs are impacted by video crosstalk. Nonetheless, ephemeris matching does not address the residual risk from stars that are not observed by Kepler or which cause false positives on only one Kepler target. Thus, video crosstalk is an indispensable model for assessing risk due to stars located on adjacent readout channels that could be the source of a transit-like signal. The video crosstalk model may also be helpful to users performing their own pixel-level corrections for other scientific purposes, either for FFIs or for long-cadence data where the requisite pixels were collected on the relevant channels. Users should be aware that, in addition to possibly overestimating the magnitude of the coefficients, the video crosstalk model was derived from data collected during the first year of the mission, and no attempt has been made to monitor the coefficients as the LDE ages. Visual examination of video crosstalk-corrected images suggests that the accuracy of the coefficients is ~25%, but no formal error estimates have been developed.

The format of the video crosstalk model is a comma-separated value (csv) file listing recipient channel, source channel, and double-precision coefficient (i, j, C_{ij} in the notation of Caldwell et al. 2010). There are 12 asymmetric coefficients per module x 21 modules = 252 crosstalk coefficients. For example, the first line of the file reads:

```
1, 2, 0.000286392931119
```

to indicate that a star on channel 2 (mod.out 2.2) will create a crosstalk signal on channel 1 (mod.out 2.1) that is 0.000286 times the pixel value on channel 2.

The method of use is to calculate the crosstalk image in DN from the raw images (in DN) of the other three channels of the module, multiply by the gain of the recipient image, and then subtract the result from the calibrated image. For example, the total crosstalk signal in DN for a pixel on channel 1 is $S_{X1} = C_{12}S_2 + C_{13}S_3 + C_{14}S_4$, where C_{ij} is the crosstalk coefficient and S_i is the uncalibrated intensity in DN from the corresponding pixel on source channel i . The crosstalk signal in e-/s for this pixel is then $S_{X1}G_1/t_{exp}$, where G_1 is the gain (e-/DN) for channel 1 and t_{exp} is the exposure time of 1625.4 seconds for both long-cadence and FFI data. In the rare case that one is calculating the per-cadence crosstalk into an LC target from a complete set of source channel LC pixels, the raw LC data must be shifted as described for RAW_CNTS in §2.3.2 before applying the crosstalk coefficients. For the more typical case of estimating crosstalk from FFI data, there is no raw data shift.

2.3.5.19 Spectral Response

The spectral response model gives the number of electrons generated in a CCD per photon entering the telescope aperture as a function of wavelength, λ . The signal generated by a star is then proportional to the product of the stellar spectrum and the spectral response model. The spectral response model includes the transmission of the Schmidt corrector assembly (SCA), the reflectance of the primary mirror assembly (PMA), the transmission of the field flattener lens (FFL) on each module including its bandpass filter (BPF) coating, as well as the quantum efficiency (QE) of each CCD chip. The model does not include vignetting or the 23% areal central obscuration of the telescope by the local detector electronics and focal-plane assembly, and does not include radiation degradation of components since launch. The SCA, PMA, and FFL + BPF were measured at moderate spectral resolution ($\lambda/\Delta\lambda > 100$) by their vendors, while the CCDs were measured in six discrete bandpasses (i.e., 400, 500, 600, 700, 800, and 850 nm) upon receipt at Ball. The CCD response was interpolated between these points before multiplying by the other factors to give the per-chip total spectral response. Since all factors are the same for the two output channels on a CCD, the spectral response model is identical for each channel pair. Figure 13 of the KIH shows the spectral response model for module 11, output 2 (i.e., channel 34), and shows the factors common to all channels (SCA and PMA) as well as the factors particular to channel 34 (the FFL + BPF transmission and CCD QE). Spectral response varies considerably over a narrow wavelength range at the edges of the band defined by the BPF, and in a narrow band around 534 nm where several of the BPFs have a fringe. Otherwise, the variation is <5% between 450 and 830 nm.

The spectral response file is a csv file with 10 header lines starting with #, followed by 85 columns of data: the wavelength in nm, followed by spectral response of the 84 channels. The numerical format is floating point with 3 decimal places. The file name is `kplr_2015120120_spectral_response.csv`.

2.3.6 Background Data Files

Kepler observes a grid of 4464 background pixels on each channel at long cadence in order to remove the zodiacal light and the unresolved background stars from the data. The Kepler pipeline fits these pixels to determine the background applicable for each observed aperture. The results of this fit are available in the light curve and target pixel files for the specified mask. The pixel time series of these background data are contained in the background data files. Users interested in modeling the background for specific targets or in the flux variations of these small patches of sky should refer to these background data files.

The background data files contain the raw and calibrated background flux for all the background pixels on each mod/out. One file exists for each mod/out each quarter. The mod/out is found in the name of the file (see Table 2-1) as well as the primary header.

Primary Header

The primary header contains information pertaining to the entire file, such as quarter, season, channel, and pipeline version numbers.

Background Binary Extension

This data extension contains the background pixels as well as the results of fitting the background for the entire channel. The background pixels are not a continuous portion of the CCD, therefore the data are stored as arrays and not images. The “Pixels” binary extension, see below, describes the CCD location of each element in these arrays. The binary table in this extension contains the following columns:

TIME_MJD: [64-bit floating point] – The UTC time at mid-point of the cadence in modified Julian date.

BKJD_TIME: [array of 64-bit floating point] – An array of barycenter corrected times in BKJD is given for each cadence. The barycenter correction was performed for the location of each background pixel. The right ascension and declination of each pixel are determined using the motion polynomial of each cadence and are specified in the header of the pixels binary extension.

CADENCENO: [32-bit integer] – The cadence number is a unique integer that is incremented by one with each cadence.

RAW_CNTS: [array of 32-bit signed integers] – The raw flux for each background pixel in units of counts.

FLUX: [array of 32-bit floating point] – The calibrated flux for each background pixel in units of electrons per second.

FLUX_ERR: [array of 32-bit floating point] – The error in the calibrated flux for each background pixel in units of electrons per second.

BKG_CO: [array of 64-bit floating point] – The background polynomial coefficients determined by fitting the calibrated background pixels on the specified channel. The number of background polynomial coefficients for each cadence is given by $(K+1)*(K+2)/2$, where K is the order of the fit, see the KDPH.

BKG_CO_ERR: [image of 64-bit floating point] – The covariance matrix that results from the background polynomial fit for the specified channel.

COSMIC_RAYS: [array of 32-bit floating point] – The cosmic rays identified and removed from each background pixel in units of electrons per second.

QUALITY: [32-bit integer] – Bit flag containing information about the data quality. See Table 2-3.

POS_CORR1: [array of 32-bit integers] – An array containing the column component of the local image motion calculated from the motion polynomial. Each element of the array represents a different background pixel. The locations are given in the pixels binary extension. We report the motion in pixels relative to the mid-cadence of the quarter.

POS_CORR2: [array of 32-bit integer] – An array containing the row component of the local image motion calculated from the motion polynomials. Each element of the array represents a different background pixel. The locations are given in the pixels binary extension. We report the motion in pixels relative to the quarter's mid-cadence.

Pixels Binary Extension

This extension describes the CCD row and column values for each background pixel found in the data array of the first binary extension (RAW_CNTS, FLUX, FLUX_ERR, POS_CORR1, POS_CORR2). The data column RAWX contains the column coordinate and the data column RAWY contains the row coordinate. The header of this extension contains the World Coordinate solution according to the FITS standard and is used to determine the values in BKJD_TIME, POS_CORR1 and POS_CORR2.

2.3.7 Collateral Data Files

The collateral data for each channel consists of masked and virtual smear columns along with trailing black rows. The short cadence collateral data also includes the overlap summed-pixels (black/masked and black/smear). See Figure 7 and §2.6.3 of the KIH for a description of the collateral data. Users interested in calibrating raw pixel data will be interested in these data files. A description of how the Kepler pipeline uses the collateral data to calibrate the raw pixels can be found in the KDPH.

In addition to the raw and calibrated collateral data, the BLACK extension contains information about the rolling bands detected by the Dynablack module. The KIH gives a detailed description of the source of the rolling bands observed on the *Kepler* CCDs. The rolling bands are 10 — 100 pixels across and drift with temperature variations. To detect rolling band artifacts the residual black data is convolved with a square wave transit kernel (at 14 different test pulse durations) to produce a time series of transit depths. These depths are used to produce a severity level for each CCD row at each cadence. If a bright star exists near the trailing black columns then the row is considered to be Scene Dependent. The column cutoff is given in the keyword DBCOLCO in the binary table header. Because of undershoot that occurs after reading out a bright star, the effective black level is reduced for the pixels read-out after the bright star. As a result, Dynablack, does not use these Scene Dependent rows when fitting the black model, and as a result, cannot detect rolling band in the center of this region. Note, some interpolation is done resulting in rolling band being detected at the edge of a scene dependent region. Those rows impacted by rolling band, as well as the scene dependence, is marked in the BLACK extension of the LC collateral data files. For more information on Dynablack and how the pipeline detects rolling band, see the KDPH.

The number of row (or column) pixels that are summed on the spacecraft to create the smear and black data can be found in the respective headers (see keywords NROWVSMR, NROWMSMR, and NCOLBLK). The file contains a separate extension for each of the black, virtual smear, masked smear, and SC overlap pixels. Each of these extensions contains a pixel list that gives the row (or column) number of each element in the array.

Primary Header

The primary header contains information pertaining to the entire file, such as quarter, channel, and pipeline version numbers.

Binary Data Extensions

The following columns are found in the binary data extensions (BLACK, VIRTUALSMEAR, MASKEDSMEAR and SC2DCOLLATERAL):

All Binary Data Extensions:

TIME_MJD: [64-bit floating point] – The UTC time at mid-point of the cadence in modified Julian date.

CADENCENO: [32-bit integer] – The cadence number is a unique integer that is incremented by one with each cadence.

BLACK Binary Extension:

BLACK_RAW: [array of 32-bit signed integers] – The raw black co-added pixels used to estimate the black level correction for the specified channel in units of counts.

BLACK_RES: [array of 32-bit floating point] – The residuals of the co-added black pixels after removing a fit to the black and the cosmic rays. The units are counts per second.

BLACK_RES_ERR: [array of 32-bit floating point] – The error in the black residuals in units of counts per second.

BLACK_CR: [array of 32-bit floating point] – The cosmic rays identified and removed from the co-added black pixels in units of counts per second.

RB_LEVEL: [2D array of 32-bit floating point] – The rolling band level for every CCD row and test pulse duration. The test pulse durations are given in hours in the keywords RBT DUR_{*i*}.

RB_FLAG: [2D array of 32-bit floating point] – The rolling band flags for every CCD row and test pulse duration. The value is a 2-bit flag where the bits are as follows: 1) Scene Dependence and 2) Rolling Band Detected.

VIRTUAL SMEAR Binary Extension:

VSMEAR_RAW: [array of 32-bit signed integers] – The raw flux of the co-added, virtual smear pixels in units of counts.

VSMEAR_FLUX: [array of 32-bit floating point] – The calibrated flux of the co-added, virtual smear pixels in units of electrons per second.

VSMEAR_FLUX_ERR: [array of 32-bit floating point] – The error in the calibrated co-added, virtual smear pixels in units of electrons per second.

VSMEAR_CR: [array of 32-bit floating point] – The cosmic rays identified and removed from the co-added virtual smear pixels in units of electrons per second.

MASKED SMEAR Binary Extension:

SMEAR_RAW: [array of 32-bit signed integers] – The raw flux of the co-added, masked smear pixels in units of counts.

SMEAR_FLUX: [array of 32-bit floating point] – The calibrated flux of the co-added, masked smear pixels in units of electrons per second.

SMEAR_FLUX_ERR: [array of 32-bit floating point] – The error in the calibrated co-added, masked smear pixels in units of electrons per second.

SMEAR_CR: [array of 32-bit floating point] – The cosmic rays identified and removed from the co-added, masked smear pixels in units of electrons per second.

SC2DCOLLATERAL Binary Extension (only in SC files):

BMASKED_RAW: [signed 32-bit integer] – The raw counts for the sum of the pixels in the cross-sections of the trailing black columns and the masked smear rows in units of counts.

BMASKED_RES: [32-bit floating point] – The residuals of the masked black after removing the fit to the black, correcting for the number of co-added pixels, and correcting for the identified cosmic rays. The units are counts per second per pixel.

BMASKED_RES_ERR: [32-bit floating point] – The error in the residuals of the masked black in units of counts per second per pixel.

BMASKED_CR: [32-bit floating point] – The cosmic rays identified and removed from the masked black co-added pixels in units of counts per second per pixel.

BVIRTUAL_RAW: [signed 32-bit integer] – The raw counts for the sum of the pixels in the overlapping virtual smear rows and trailing black columns in units of counts.

BVIRTUAL_RES: [32-bit floating point] – The residuals of the virtual black values after removing the fit to the black, correcting for the number of co-added pixels and correcting for the identified cosmic rays. The units are in counts per second per pixel.

BVIRTUAL_RES_ERR: [32-bit floating point] – Error in the residuals of the virtual black in units of counts per second per pixel.

BVIRTUAL_CR: [32-bit floating point] – The identified and removed cosmic rays found in the virtual black region in units of counts per second per pixel.

Binary Pixel List Extensions

The pixel list extensions (BLACKPIXELLIST, VIRTUALSMEARPIXELLIST, and MASKEDSMEARPIXELLIST) describe the location (row or column) of the values in the arrays in the

black and smear extensions. For short cadence data, these pixel lists are not contiguous because the collateral data is only read-out for those rows and columns that contain a collected pixel.

2.3.8 Artifact Removal Pixel Data Files

The Artifact Removal Pixels (ARPs) are individual pixels in the collateral data that are read-out at the long cadence rate. They are a set of pixels defined to sample the states of the fine guidance sensor (FGS) crosstalk with the science CCD pixels. They are used to monitor the crosstalk levels in both the pre-clocked bias region ("leading black") and the over-scanned bias region ("trailing black"). The pixels in the ARP files are not contiguous; therefore the data is stored in arrays instead of images for each cadence. ARP data are used by the pipeline for dynamic 2-D black correction. See the KIH (§6.8) and KDPH for further details.

Primary Header

The primary header contains information pertaining to the entire file, such as quarter, season, channel, and pipeline version numbers.

Artifact Removal Binary Extension

The following columns are found in this data extension:

TIME_MJD: [64-bit floating point] – The UTC time at mid-point of the cadence in modified Julian date.

CADENCENO: [32-bit integer] – The cadence number is a unique integer that is incremented by one with each cadence.

RAW_CNTR: [array of 32-bit integers] – The raw counts collected for each ARP pixel in units of counts. See the pixels binary extension for the row and column of each element of the array.

FLUX: [array of 32-bit floating point] – The array of calibrated ARP pixels in units of electrons per second.

FLUX_ERR: [array of 32-bit floating point] – The error in the calibrated ARP pixels in units of electrons per second.

COSMIC_RAYS: [array of 32-bit floating point] – The array of cosmic rays identified and removed from the ARP pixels in units of electrons per second. See the pixels binary extension for the row and column of each element of the array.

QUALITY: [32-bit integer] – Bit flag containing information about the data quality. See Table 2-3.

Pixels Binary Extension

This extension describes the CCD row and column values for each ARP pixel found in the data array of the first binary extension. RAWX contains the column coordinate and RAWY contains the row coordinate.

2.3.9 Reverse Clock Data Files

The Kepler CCD is reverse clocked several times every month to obtain a measure of the bias (see KIH §5.1). These reverse clock data files are created by the DMC and are available at the MAST through the FFI and Engineering data download page. The reverse clock data set is comprised of the target pixels (lcs-targ), the background pixels (lcs-bkg) and the collateral pixels (lcs-col). Each file contains one extension per mod/out. Each extension contains one floating point array of all the reverse clock measurements in that mod/out. To determine the physical CCD location of each element of the array, refer to the pixel mapping file specified in the header. The primary header keywords LCTPMTAB, BKGPMTAB and LCCPMTAB give the name of the pixel mapping file associated with the target, background and collateral reverse clock data, respectively. The pixel mapping files are also available from the MAST.

2.3.10 Ancillary Engineering Data

Ancillary engineering data is a subset of the engineering telemetry (KIH §2.5.5) collected by the spacecraft that may prove useful for science data analysis and flight system characterization. Generally speaking, items associated with attitude determination and control (e.g. pointing error, wheel speeds) and various temperatures are the most useful for systematic error removal. This assessment is based on studies and results for early versions of the pipeline (version 7.x and before), before the pipeline moved to light curve ensemble-based systematic error removal (Stumpe et al., 2012; Smith et al., 2012).

There is one ancillary engineering data file per LC data collection interval, and in some cases ancillary engineering data exists for time intervals where no valid science data was collected. The ancillary engineering data files are in FITS format. Values for each mnemonic are contained in an individual FITS binary table extension containing 2 columns: one for the sample time (in MJD) and the other for the parameter value. The sampling rates are not the same for all parameters. The keywords in each extension give a rough description of the parameter and its units. These files have roughly 200 extensions. The extension index, mnemonic, description, and units for each item are tabulated in Appendix B.5. While maps of the exact locations of the sensors are not available in the archive, approximate locations can be inferred from the telemetry item descriptions and the flight system drawings in KIH §2.5.

Not all engineering data found to be relevant are included in the archive. For example, the signal levels of variable guide stars and the reaction wheel housing temperatures (TH1RW3T and TH1RW4T) are correlated with photometric errors, but are not in the delivered ancillary engineering data. See the KDCH §4.5 and §5.2 for further discussion.

2.4 Keyword Definitions

Several keywords in the headers of archive products cannot be explained very well in the 40 characters allotted for the comment field. Here we provide a short dictionary to clarify some of these keywords.

BACKAPP: This keyword is set to true if the background has been subtracted from the FLUX column of the target pixel file or the SAP_FLUX column of the light curve file. The background flux has not been subtracted for the FFIs causing BACKAPP to be set to false.

CDPP3_0: The root mean square CDPP (combined differential photometric precision) value calculated for 3-hour intervals by the Kepler pipeline. CDPP6_0 contains the 6-hour rms CDPP value

and CDPP12_0 contains the 12-hour rms CDPP value for the same target. CDPP corresponds to the depth of a box-car transit signal with a duration of N-hrs that results in a detection Signal-to-Noise of unity in the Transiting Planet Search (TPS) module. CDPP is calculated on a per cadence basis, thus the rmsCDPP only captures the first moment of the CDPP time series, and not its time evolution or distribution of CDPP values. See the KDPH for more details on how rms CDPP is calculated by the pipeline. These CDPP values, while accurate to assess the quarterly noise levels, are unsuitable for occurrence rate calculations. Algorithmic and parametric differences exist between these values and those used by TPS to find exoplanets. The final RMS CDPP values, which were explicitly designed for this purpose (Burke et al. 2015), can be found in the DR25 DV Time Series Files (KSCI-19079-001) or the DR25 Robust RMS CDPP column of the Stellar table (see NExSci).

CROWSAP: The ratio of the flux from the target to the total flux in the optimal aperture as estimated from the sources available in the KIC.

DATA_REL: The version of the data release notes that corresponds to the data set. Each time data is processed it is associated with a new set of data release notes.

EQUINOX: The equinox of the celestial coordinate system used to describe the Right Ascension and Declination of the target.

FILEVER: File format version. This is incremented each time the FITS format of a file type changes. This is specified as a string of the form “<major_update>.<minor_update>”. Increments in the minor_update number should maintain compatibility with code written for that file type and major_update number.

FITTYPEj: The type of fit used by PDC for the band indicated by the integer j. Values include ‘robustFit’, ‘priorFit’ or ‘none’. The number of bands is indicated by the keyword NUMBAND.

FLFRCSAP: Fraction of the target flux contained in the optimal aperture calculated by the target aperture definition portion of the pipeline using the sources available in the KIC.

GAIN: The value of the gain measured for the specified channel.

INT_TIME: The integration time for a single frame. Kepler sums 270 of these integrations to create one long cadence value and 9 of these for one short cadence value; see NUM_FRM.

LC_START: The time at the middle of the first cadence in the file in modified Julian date.

LC_END: The time at the middle of the last cadence in the file in modified Julian date.

LCFXDOFF: The value of the flight software black level added to the long cadence data by the spacecraft as part of the requantization processes prior to downloading the data. See the Instrument Handbook for more details (van Cleve & Caldwell 2009, KSCI-19033).

NPIXSAP: Number of pixels in the optimal aperture.

NPIXMISS: Number of pixels that should be in the optimal aperture, but were not collected by the spacecraft.

NUMBAND: Number of wavelet length-scale bands used by PDC to produce the PDCSAP_FLUX time series. The bands are listed in order from the longest length-scale (i.e., period) to the shortest.

NUM_FRM: The number of frames summed to create the data contained in the file.

PDCVAR: Measure of the variability as measured by PDC (see equation 11 of Smith et al. 2012). This value is used to weight the prior relative to the probability distribution function.

PDCMETHD: The method used to calculate the PDC light curve provided in the file. The options include ‘multiscaleMAP’ or ‘regularMAP’.

PR_GOOD j : The goodness of the PDC prior fit associated with the band indicated by the value j . The number of bands is indicated by the keyword NUMBAND (see equation 19 of Smith et al. 2012).

PR_WGHT j : The prior weight used by PDC with the band indicated by the value j . The number of bands is indicated by the keyword NUMBAND.

RADESYS: The reference frame for all celestial coordinates reported by Kepler is in ICRS (International Celestial Reference System).

READNOIS: The value of the readnoise in electrons measured for the specified channel.

SCFXDOFF: The value of the flight software black level added to the short cadence data by the spacecraft as part of the requantization processes prior to downloading the data. See the KIH.

TIERRELA: The relative timing error. All times are accurate relative to each other within 50 ms.

TIERABSO: The absolute timing error. This value is blank in the header, however a limit on the absolute timing error is discussed in the KDCH.

TSTART: The start time of the observations contained in the file measured at the beginning of the first cadence. The units are in BJD-BJDREF, where the keywords BJDREFI and BJDREFF make up the value of BJDREF.

TSTOP: The stop time of the observations contained in the file measured at the end of the last cadence. The units are in BJD-BJDREF, where BJDREF is a keyword listed in the file.

Chapter 3 Searching the Archive for Kepler Data

This chapter describes the MAST Kepler search forms and interfaces as they exist at the time of writing. Users are encouraged to consult the Kepler main page at MAST for new features, interfaces, and documentation that may be added after the official close-out of the mission. Layout of the MAST pages is also subject to change.

The first section describes the various catalogs that form a basis for these interfaces and includes information on “target” searches, which are now used to search either the Kepler Input Catalog (KIC Search) as delivered by the Kepler Project or the KIC as augmented by MAST (Target Search). Subsequent sections describe the MAST Home Page, the MAST Data Discovery Portal, which is a new search interface that will eventually replace the Classic MAST interfaces, the Kepler Classic MAST interfaces, and various other data access tools and interfaces.

Kepler data and products may be searched for and/or retrieved via the MAST Discovery Portal (aka Portal), the Classic MAST search forms, via CasJobs (search only) or directly through wget/curl/ftp or your browser.

Common searches are based on position, time of observation, target name or kepler_id (Kepler Identification Number), but additional catalog fields are searchable. MAST allows the upload of a file containing a list of up to 10,000 kepler_id's, coordinates and/or target names for desired searches. Cross correlation with catalogs is possible using CasJobs or the Portal. Please note kepler_id and target name are different identifiers. MAST makes this distinction so that users may enter, say, NGC 6791 or TrES-2, as the target name and the Resolver will return the coordinates. The resolver does accept planet names (e.g., Kepler-68c), KIC numbers and KOI numbers as target names.

The archive at MAST contains Kepler light curves and target pixel files, which are the basic data products produced by Kepler, several catalogs used by proposers to select targets, several catalogs that are used by the pipelines as part of the processing, information on the KOIs (Kepler Object of Interest) and a number ancillary files and tables. Pointers to these files may be found on the MAST Kepler home page <https://archive.stsci.edu/kepler/>. Users are encouraged to explore the content listed in the left gutter menus and in the tabs at the top of the page.

3.1 The MAST Kepler Catalogs

This section describes the primary catalogs underlying the MAST Data Search & Retrieval Interface (§3.4), the Kepler Target interface and the KIC interface.

The Target Search and the KIC Search interfaces were designed as a way to locate targets on-silicon for future observation by Kepler. As implied by the name, the KIC Search is used to search the Kepler Input Catalog (KIC), which is the catalog as received from the Kepler Project. As an aid to proposers, MAST combined the KIC, the Characteristics Table, additional targets in the 2MASS catalogs (the KIC was incomplete in this respect), the Lucas (UK-IRT) survey in J (KIS), fluxes and colors from GALEX, Sloan (SDSS/DR9, to be replaced with future SDSS releases over time), Everett-Howell (UBV), and the Kepler Isaac Newton Telescope Survey (IRT) and created the Target Search interface. The KIS and UBV surveys are described in Griess et al. (2012) and Everett, Howell, & Kinemuchi (2012). See https://archive.stsci.edu/kepler/kepler_fov/explanations.html, the Explanations and Caveats page, for

descriptions of the surveys and the rules used to form the color names. Cross correlation was done to match the KIC targets to each catalog or data set. Users are warned to combine KIC and KIS colors and magnitudes with care, as they are based on different zero point magnitude systems (AB and Vega/Johnson).

The KIC contains information on approximately 13 million sources, most of which are visible in or nearby the *Kepler* field of view. Creation of the KIC is discussed in Brown et al. (2011). Each source has an identification number, called the `kepler_id`, and a position (RA and Dec). Additional fields may or may not have values for each source. The KIC search interface is <https://archive.stsci.edu/kepler/kic10/search.php>.

The Characteristics Table (CT) contains parameters indicating if a given `kepler_id` is on-silicon by Kepler for each of the 4 seasons, as well as other characteristics. Not all sources in the KIC have values for all parameters. In fact, only about one third of the KIC entries are ever on-silicon. MAST does not provide a direct search capability for the CT. Instead, the CT fields have been included in the Kepler Target table. The Kepler Target search form is used to search the Kepler Target Table. Some information about the CT parameters is given in the MAST Field Descriptions files found at the following link: https://archive.stsci.edu/search_fields.php?mission=kepler_fov. CT derived fields are prefixed in the “Column Name” by “`kct_`”.

For a given object, the `kepler_id` is the same in the KIC and the KTC. The one exception is for KIC targets observed with custom apertures (see §2.2.1): in this case, the `kepler_id` listed in the KTC is **not** in the KIC or the CT.

Facts about the KIC, CT and Kepler Target tables:

- The full KIC contains 13,161,029 rows (objects).
- The CT contains 6,569,466 rows (about half of the KIC total).
- The total number of targets in the Kepler Target table is > 15 million.
- The number of KIC targets on CCD every season is 4,353,971.

The Target search Interface, http://archive.stsci.edu/kepler/kepler_fov/search.php, also contains a “Fields Description” link to a page with a short description, a range of valid values and the data type for available fields.

When doing a target search, several catalog values are listed by season, necessary because Kepler was rotated quarterly. The planned start date for each season is given in Table 3-1. The seasons are numbered from 0 to 3, with 0 corresponding to Summer. Some targets are not on the CCD every quarter.

Table 3-1 Start Dates by Season

Year	Spring	Summer	Fall	Winter
Quarters	0,1,5,9,13,17	2,6,10,14	3,7,11,15	4,8,12,16
Season	3	0	1	2
2009	May2, May 13	Jun 20	Sep 18	Dec 19
2010	Mar 20	Jun 24	Sep 23	Jan 06,2011
2011	Mar 21	Jun 27	Sep 29	Jan 05,2012
2012	Mar 29	Jun 28	Oct 05	Jan 12,2013
2013	Apr 9			

Note on Contamination/Crowding values: Contamination is a floating-point number between 0 and 1 representing the fraction of light in the aperture due to the target star. MAST calculates and provides the value as a rough estimate of the contamination for each KIC target and season in the MAST *target* search interface. **However, this contamination value is NOT to be used to correct *Kepler* flux light curves.** These values are intended only to facilitate the selection of isolated stars as targets. The contamination value relevant to a specific flux light curve can be substantially different from the predictive value. It varies because of changes in photometric aperture size, the pixel response function, optical throughput, and other position-dependent characteristics of the focal plane. The aperture for *Kepler* photometry is approximately defined before data collection, but is not finalized until post-processing is complete.

The values in the Data Search form for crowding (from which the contamination value is derived) and flux fraction are the calculated values used by the *Kepler* pipeline to create the PDC light curves, and may be used for contamination correction. These are the values found in the headers of the individual files for light curve files of v2.1 and later.

The search for data (actual observations) uses the archive table containing the exposure metadata, the KIC and the Kepler Target Catalog (KTC). The KTC contains all objects observed or scheduled for observation by *Kepler*. At launch, there were some 150,000 targets in the KTC. The KTC was updated on a quarterly basis and holds observation start and stop times for each target. The times are given in both Modified Julian Date (MJD) and standard date format (i.e., YYYY-MM-DD HH:MM:SS). See §3.4.

A list of the fields in the Data Search and Retrieval Interface is available via the "Fields Description" link, https://archive.stsci.edu/search_fields.php?mission=kepler, on the page.

3.2 The MAST Home Page

The MAST Home Page is shown in Figure 3-1. The link to MAST is <https://archive.stsci.edu>. Note the following items. At the top of the page is the "Mission_Search" menu. Use this menu to reach the Portal, and the classic MAST Kepler Data and Kepler Target pages. The lower portion of the page contains the Portal Search Box. Enter a target name or coordinates in the "one box" and click search or simply click on search to go to the Portal.

Barbara A. MIKULSKI ARCHIVE OF SPACE TELESCOPES

MAST STScI Tools Mission_Search Search Website Follow Us

About MAST Getting Started

FAQ

High-Level Science Products

Software

FITS

Archive Manual

Related Sites

NASA Datacenters

MAST Services

MAST and the VO

Newsletters & Reports

Data Use Policy

Dataset Identifiers

Acknowledgments

As of April 13, the archive is now using the STScI Single Sign-On (SSO) identity manager. To check on your account click [here](#). For more information about how accounts were transitioned click [here](#).

The Mikulski Archive for Space Telescopes (MAST) is a NASA funded project to support and provide to the astronomical community a [variety of astronomical data archives](#), with the primary focus on scientifically related data sets in the optical, ultraviolet, and near-infrared parts of the spectrum. MAST is located at the Space Telescope Science Institute (STScI).

NEW Search Using the MAST Data Discovery Portal

- MAST Cross-Mission Search
- Integrated 3D All-Sky Viewer
- Access to Data from other Archives
- [User's Guide](#)

Enter [Target Name \(or Coordinates\)](#):

Search

Enter value or just click on search for the Portal

One box entry

Google Google Search

WWW MAST

Note: STScI may provide links to Web pages that are not part of the STScI, AURA, NASA, or ESA domain. These sites are managed by organizations, companies or individuals not under our control, and neither STScI, AURA, NASA, nor ESA are responsible for the information or links you may find there. We provide these links as a convenience and the presence of these links is not an endorsement of the site.

News

January 22, 2016:
New HLSP: Hubble Tarantula Treasury Project (HTTP)

December 17, 2015:
GLASS team releases data for two more clusters!

December 14, 2015:
Hubble Legacy Archive Release 9

December 07, 2015:
Big Portal Updates Over the Past Year

December 03, 2015:
2 Guest Observer Cycle Program Announced

Missions

Hubble

Hubble Legacy Archive

HSTonline

DSS

JWST

K2

KEPLER

SwiftUVOT

XMM-OM

BEFS (ORFEUS)

Copernicus

EPOCH

EUVE

FUSE

GALEX

GSC

Figure 3-1 MAST Home Page (archive.stsci.edu)

3.3 MAST Discovery Portal (aka Portal)

The Portal is MAST’s general-purpose data discovery and retrieval interface. First time users are encouraged to consult the help, the User Manual and view the tutorials. Access to all of these is available from the Portal home page, as shown in Figure 3-2. The Portal has many features that enable new search potential, both within a given mission and especially between missions, both MAST and non-MAST.

The Portal home page contains a “one box”. Enter a target name or a set of coordinates; click “Search” and results will be displayed, as shown in Figure 3-3 for KIC1026957. Note: By default all MAST missions are searched.

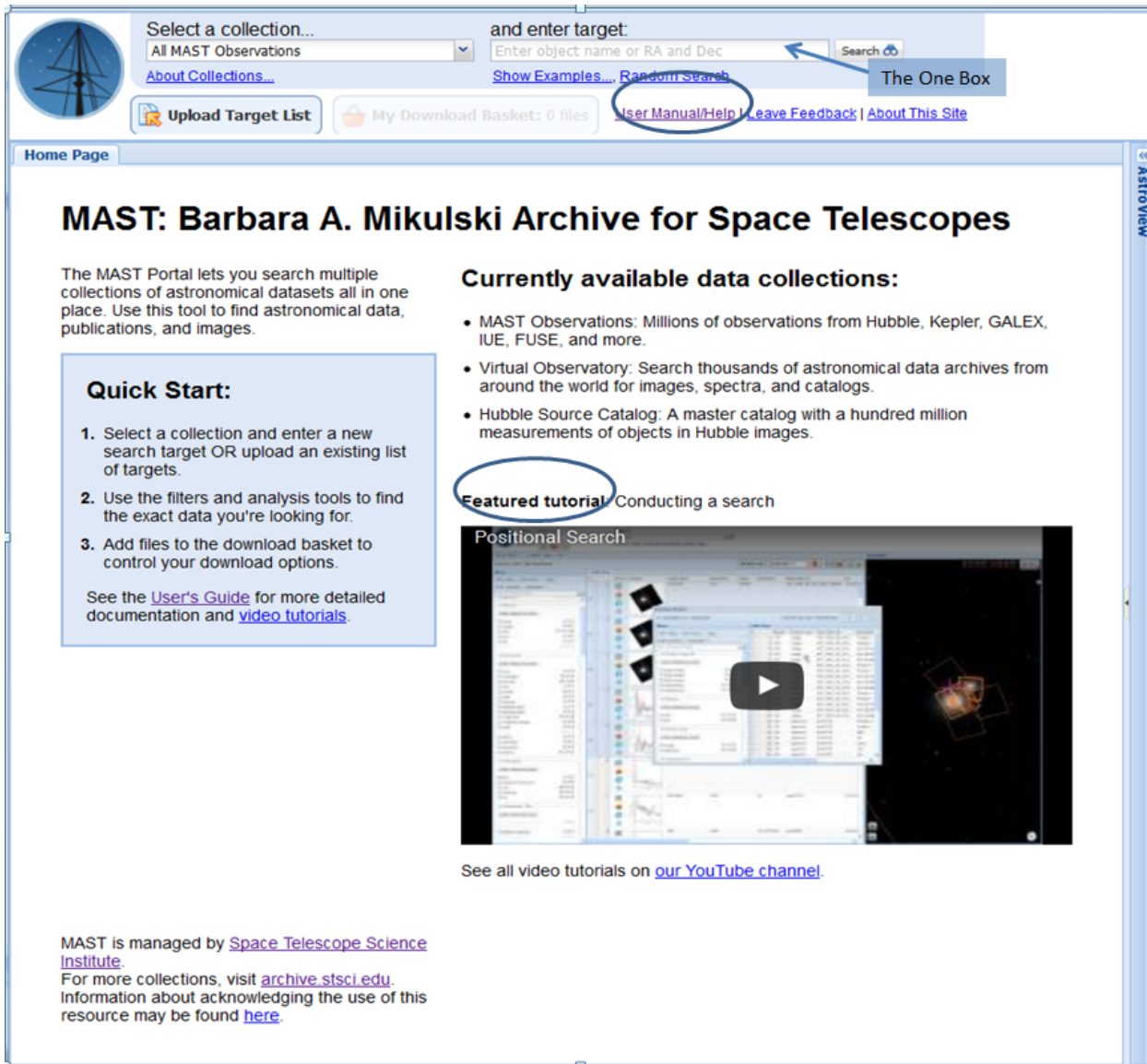


Figure 3-2 Portal Home Page.

Notice the top portion of the page remains the same, and shows the target name. The number of found results, 211, is given on the left side of the page. The main portion of the page is divided into three panels: Filters, List of Observations, and AstroView. The Filters and AstroView panes may be minimized to better view the metadata for the observations.

The 211 observations for a single KIC object should strike the reader as high. This is because the cone search currently has a default radius of 0.2 degrees. Users have the option of entering a radius in the one box (i.e., KIC1026957 r=0.005) or using the filters to reduce the number of displayed results. The radius may be specified in different units, e.g., r=0.2m for arc minutes, r=10s for arc seconds.

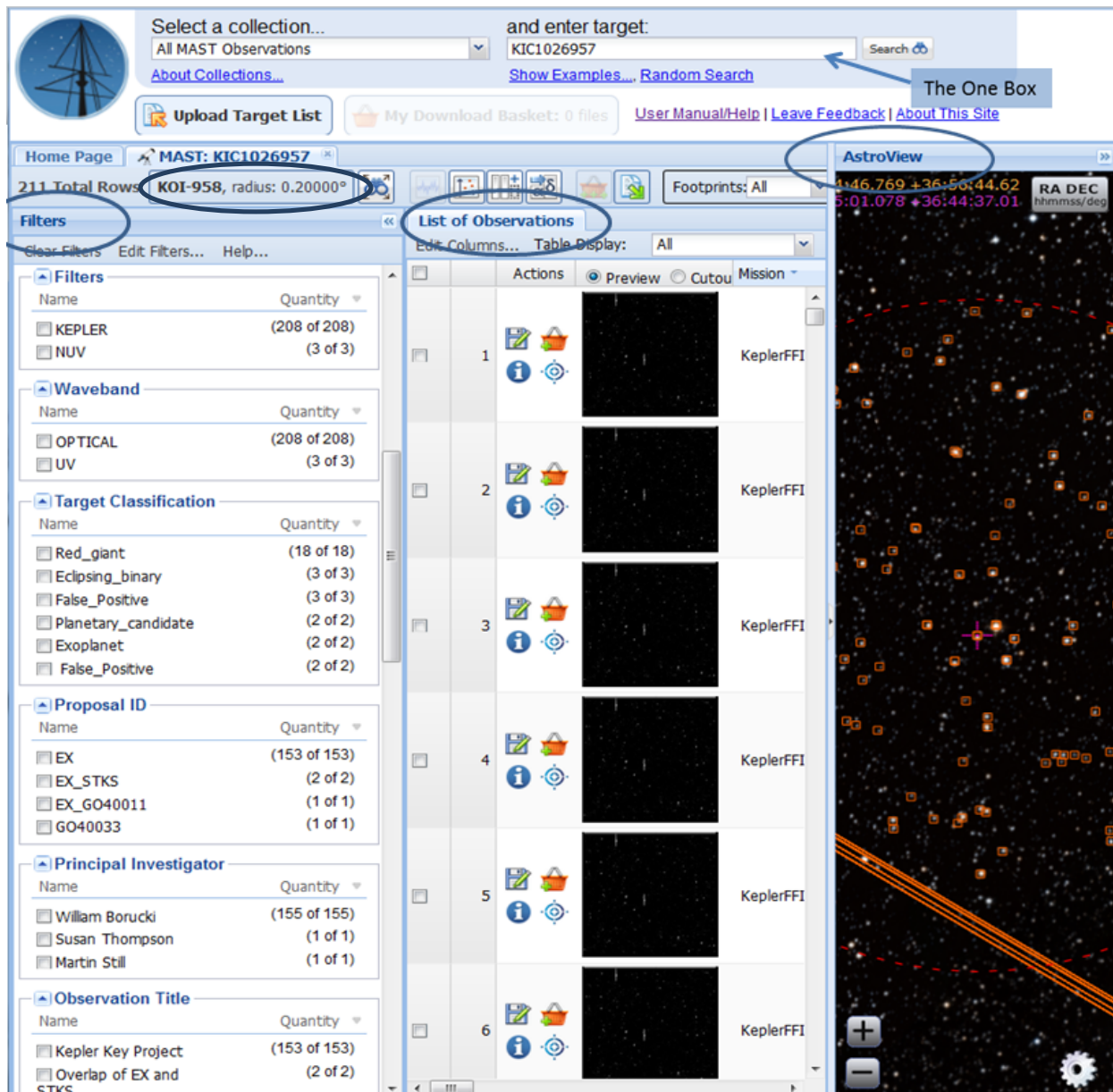


Figure 3-3 Portal Results

The search results are stored in tabs, in this case "KOI-958, radius: 0.20000," so the results of multiple searches are retained. Each search is stored in a separate tab, which can be closed using the 'X' button on each tab. Columns in the metadata section can be hidden or re-ordered and the filters may be used to limit the number of rows displayed. Figure 3-4 shows such a filtered page, where only the light curves for KIC 1026957 are given.

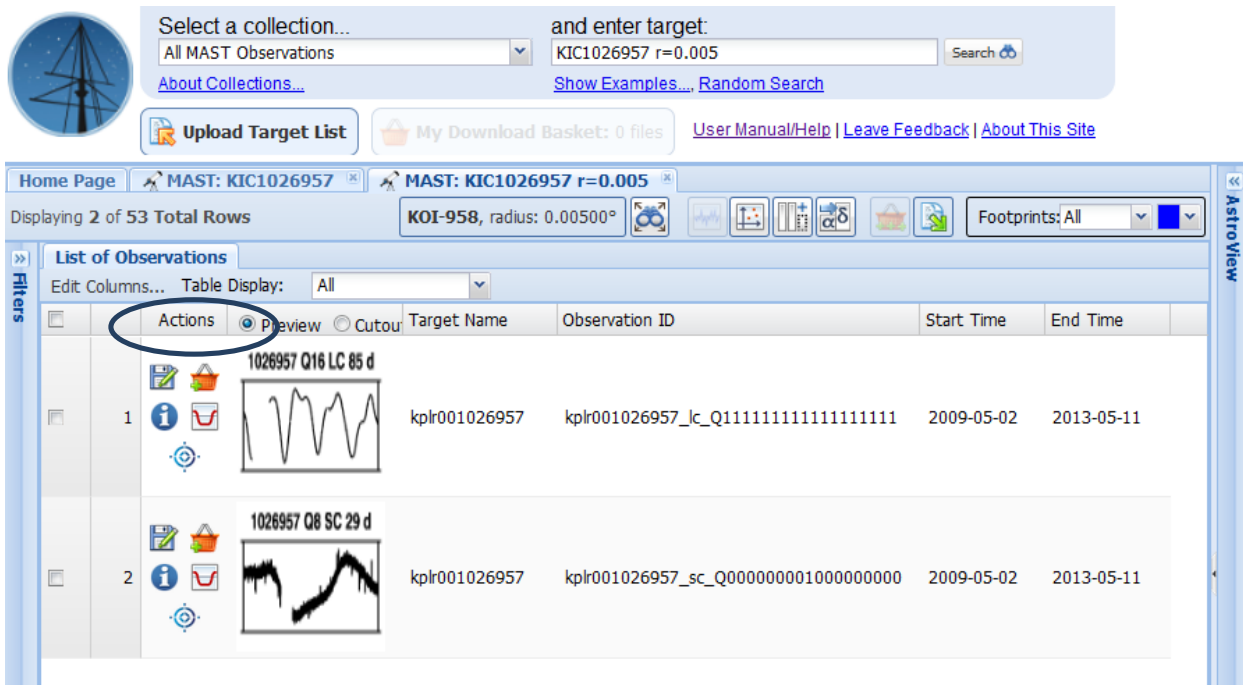


Figure 3-4 Filtered Portal results for KIC1026957.

Note the 5 icons in the “Actions” column. Hovering the cursor over each icon reveals its purpose. The floppy disk icon is an immediate data download. The basket is a shopping cart, to hold data for later download. The graph is a timeseries viewer. The "information" icon displays extra metadata, access to a larger preview image, and other MAST pages for this target. The "target" icon centers AstroView on this entry. See the help for more details.

Clicking on the graph brings up the Timeseries Viewer. One of the available light curves for the target will be displayed by default. Additional light curves may be added to the plot by selecting them. The Viewer contains a Phase Folding feature. Figure 3-5 shows the Q0 and Q1 SAP light curves phase folded by 21.76 days. Note the “Increment by” box, which allows the user to easily increase or decrease the entered period.

The Portal contains many other features that users will find helpful. Consult the help, the User Manual and the video tutorials for details and examples.

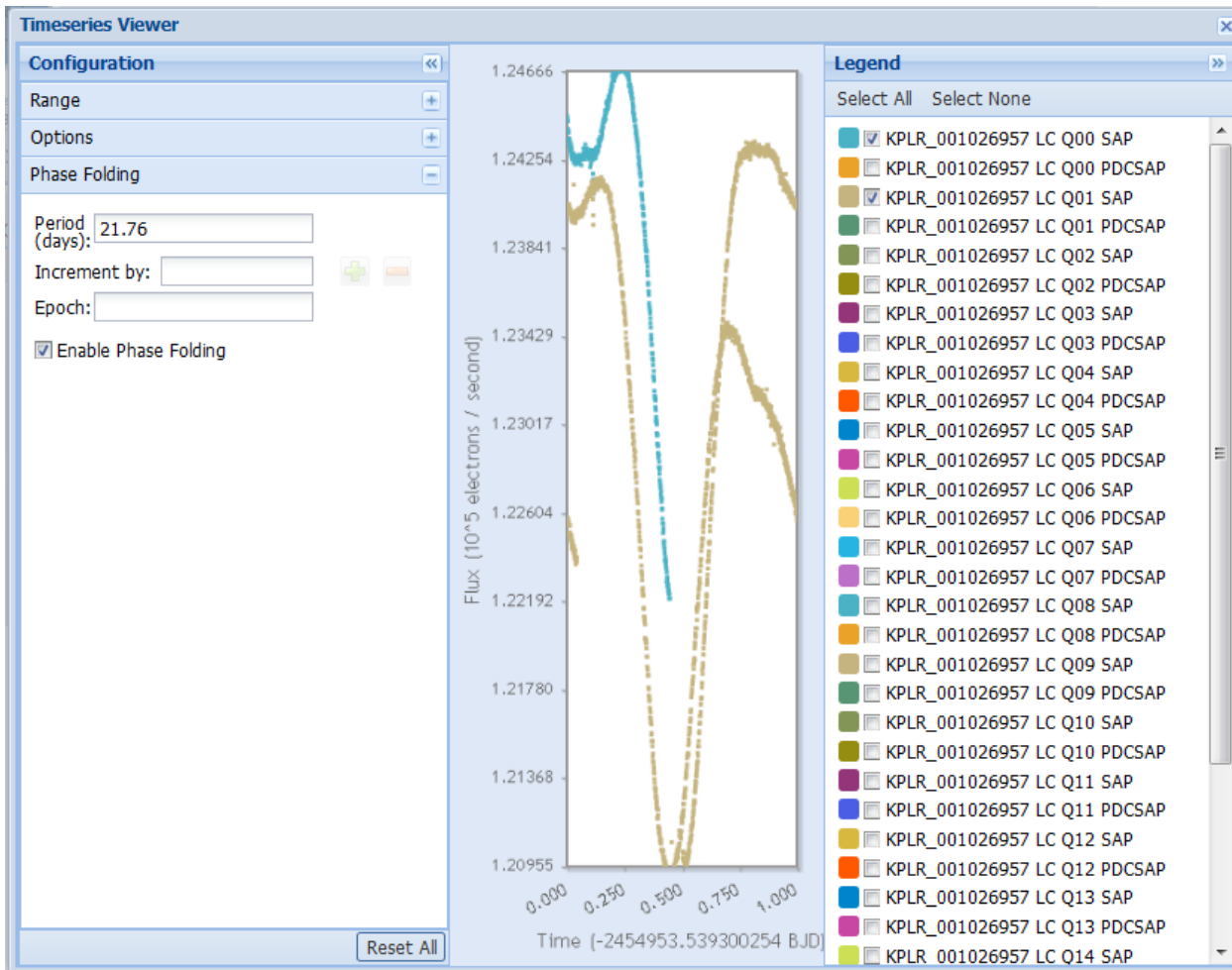


Figure 3-5 Phase folded Q0 and Q1 SAP light curves for KIC1026957. The period used is 21.76 days.

3.4 MAST Classic Interface

Users wishing to search for and/or retrieve Kepler data may use the Data Search & Retrieval Form (i.e., the Data Search option in the Mission Search menu), which is available at https://archive.stsci.edu/kepler/data_search/search.php. On-line help is available. The search for data uses information from the exposure metadata in the archive table, the KIC, and the KTC.

A note about target pixel files: Where a light curve has a matching target pixel file, users will only see the light curve file listed in the search results. However, the user may request the target pixel file be delivered with the light curve. If no light curve file is available, but a target pixel file exists, the target pixel file record will be included in the search results. The “Datasets Marked for Retrieval” page will show records for both light curve and target pixel files. The “Class” field on this page indicates if the record is for a light curve (archive class CSC or CLC) or a target pixel file (archive class TPL or TPS), where the ‘L’ and ‘S’ stand for long or short cadence in both cases.

The Kepler data tables contain several values pulled directly from the archived light curve and target pixel files. These include the actual start time, the actual stop time, crowding, flux fraction, 3-hr CDPP, 6-hr CDPP, and 12-hr CDPP.

Figure 3-7 shows the Data Search & Retrieval Form, hereafter called simply Data Search, set up for a very simple search, with only the Kepler ID as a qualifier. Search results are shown in Figure 3-8.

In the Search Results Form, use the “Mark” column to indicate the data to retrieve. To send a retrieval request for these data, click the “Mark” boxes or click on “Mark all”, then click on “Submit marked data for retrieval from STDAS.” The list of datasets to retrieve will be displayed. Unselect those not wanted, then click on “Submit marked data for retrieval from STDAS.” This will open up the Kepler Retrieval Options page, as shown in Figure 3-9.

The Retrieval Options Form is used for retrieval of public and proprietary data. Since all Kepler data are public, users do not need to use STScI SSO account credentials (i.e., your registered e-mail address and password with STScI) as anonymous retrieval may be used instead (Retrieve anonymously). Users who prefer to use an account rather than be anonymous **must** use an SSO account. For anonymous retrieval, enter your e-mail address. For all requests, even anonymous, a valid e-mail account is necessary in order for the archive to send status information regarding the retrieval request.

Under Delivery Options, indicate how the data should be delivered. If requesting “sftp” or “FTP” delivery, fill out the boxes on the right side of the page, giving the name of the receiving computer, the location for the data, and a valid account name and password. Click on “Send retrieval request to ST-DADS.” Some error checking is done for data that are to be sent by ftp or sftp, to ensure access to the computer and location for the data. A confirming e-mail is sent if the checks pass. If there is a problem, an e-mail reporting the problem is sent to the user. For a successful request, the page shown in Figure 3-10 is displayed.

If the “Archive staging area” is selected, the retrieved data are written to an Archive staging disk. The data will remain in this location for ~2 days. The user is responsible for ftp’ing the data from the staging disk. After clicking on “Send retrieval request to ST-DADS.”, the page shown in Figure 3-10 is displayed.

Standard ftp or sftp may be used to retrieve data from the staging disk. Users are advised to issue bin (for binary) and, if using mget, prompt (to turn off prompting). Attempting to ftp the data from the staging disk before the retrieval is complete will result in incomplete datasets. The ftp should not be started until after the completion e-mail has been received.

For all retrievals, e-mail is sent confirming the retrieval request was received, where the data will be placed and what datasets will be retrieved. After the retrieval is complete, a second e-mail is sent listing the files that were delivered. If the “Archive staging” option was used, the location of the data is sent.

[Archive Status](#) **Kepler Data Search & Retrieval** [Help](#)
[Field Descriptions](#)

[Standard Form](#) [File Upload Form](#)

Target Name Resolver Resolve
Right Ascension Declination Radius (arcmin) 0.02
Equinox J2000

Kepler ID 1571152 Investigation ID 2Mass ID
KFP Mag Target Type Long Cadence Short Cadence Release Date
Teff Log G Quarter
Condition Flag All Targets

User-specified field 1 Kepler ID Field Descriptions
User-specified field 2 Kepler ID Field Descriptions
User-specified field 3 Kepler ID Field Descriptions
User-specified field 4 Kepler ID Field Descriptions

Output Columns
Mark
Kepler ID
Dataset Name
Target Type
Archive Class
Actual Start Time
Actual End Time
Release Date
 Mark

Sort By: ang_sep (') Reverse
Kepler ID Reverse
null Reverse

Output Coords: Sexagesimal Degrees Hours

Output Format HTML_Table

Remove Null Columns Make Rows Distinct Skip formatting

Maximum Records: 5001
Records per Page: 500

Figure 3-7 Kepler Data Search & Retrieval form

[Display numeric columns graphically using VOPlot](#) ([Help](#))

number of rows returned = 19

note: reload page if no results are shown

[Table Info](#)

Mark	Kepler ID	Dataset Name	Archive Class	Actual Start Time	Actual End Time
<input type="checkbox"/>	1571152	KPLR001571152-2009131105131	CLC	2009-05-02 00:54:56	2009-05-11 17:51:31
<input type="checkbox"/>	1571152	KPLR001571152-2009166043257	CLC	2009-05-13 00:15:49	2009-06-15 11:32:57
<input type="checkbox"/>	1571152	KPLR001571152-2009201121230	CSC	2009-06-20 00:10:56	2009-07-20 19:12:30
<input type="checkbox"/>	1571152	KPLR001571152-2009259160929	CLC	2009-06-20 00:25:09	2009-09-16 23:09:29
<input type="checkbox"/>	1571152	KPLR001571152-2009350155506	CLC	2009-09-18 17:19:58	2009-12-16 23:55:06
<input type="checkbox"/>	1571152	KPLR001571152-2010078095331	CLC	2009-12-19 21:03:56	2010-03-19 16:53:31
<input type="checkbox"/>	1571152	KPLR001571152-2010174085026	CLC	2010-03-20 23:47:15	2010-06-23 15:50:26
<input type="checkbox"/>	1571152	KPLR001571152-2010265121752	CLC	2010-06-24 22:44:09	2010-09-22 19:03:09
<input type="checkbox"/>	1571152	KPLR001571152-2010355172524	CLC	2010-09-23 16:08:24	2010-12-22 01:10:41
<input type="checkbox"/>	1571152	KPLR001571152-2011073133259	CLC	2011-01-06 20:45:08	2011-03-14 20:18:16
<input type="checkbox"/>	1571152	KPLR001571152-2011177032512	CLC	2011-03-21 00:24:25	2011-06-26 10:10:29
<input type="checkbox"/>	1571152	KPLR001571152-2011271113734	CLC	2011-06-27 08:14:32	2011-09-28 18:22:51
<input type="checkbox"/>	1571152	KPLR001571152-2012004120508	CLC	2011-09-29 16:56:19	2012-01-04 19:50:25
<input type="checkbox"/>	1571152	KPLR001571152-2012088054726	CLC	2012-01-05 21:49:54	2012-03-28 12:32:43
<input type="checkbox"/>	1571152	KPLR001571152-2012179063303	CLC	2012-03-29 05:42:32	2012-06-27 13:18:20
<input type="checkbox"/>	1571152	KPLR001571152-2012277125453	CLC	2012-06-28 15:17:47	2012-10-03 19:40:10
<input type="checkbox"/>	1571152	KPLR001571152-2013011073258	CLC	2012-10-05 23:39:07	2013-01-11 15:18:15
<input type="checkbox"/>	1571152	KPLR001571152-2013098041711	CLC	2013-01-12 14:21:10	2013-04-08 11:02:28
<input type="checkbox"/>	1571152	KPLR001571152-2013131215648	CLC	2013-04-09 17:26:44	2013-05-11 12:01:39

Figure 3-8 Kepler Data Search Results

Kepler Retrieval Options

[Archive Status](#)

[Important Downtime Message](#)

38 datasets (38 public, 0 proprietary) marked.

User Information

Retrieve anonymously

Email (required) [?](#)

Use your STScI ID

STScI ID (your email; required) [?](#)

Password (required) [?](#)

Alternate Email (optional) [?](#)

Delivery Options

Deliver the data to the Archive staging area [?](#)

Send the data to me on DVD. [?](#)

Compress the files using gzip [?](#)

sftp (OpenSSH v2) or

FTP the data to:

Hostname (required) [?](#)

directory (required) [?](#)

username (required) [?](#)

password (required) [?](#)

Select the type of files you want...

Light Curves [?](#)

Target Pixel Files [?](#)

Focal Plane Characterization Files [?](#)

Full-Frame Images [?](#)

Cadence Data [?](#)

Ancillary Data [?](#)

Pixel Mapping Reference Files [?](#)

Telemetry Files [?](#)

... or select or enter one or more extensions:

LPD-TARG [?](#)

LLC

SLC

SPD-TARG [?](#)

[?](#)

Figure 3-9 Retrieval Options Page

Kepler Retrieval Options

[Archive Status](#)

Your request was successfully submitted. The following datasets were requested:

kp1r001571152-2010265121752, kp1r001571152-2013131215648, kp1r001571152-2011271113734, kp1r001571152-2009131105131, kp1r0015

The request number is **anonymous27812**.

Wed Jan 27 13:49:10 2016
archive@stsci.edu

Copyright © 2016 [AURA, Inc.](#)
All Rights Reserved.

Figure 3-10 Confirmation of successful retrieval.

Example Data Searches

All searches for Kepler data using the Data Search form are, by definition, searches of the KTC, https://archive.stsci.edu/kepler/data_search/search.php. The ultimate purpose of a data search is retrieval of data from the archive. When searching, check the fields listed in the “Output Columns” box. Also check the format in the “Output Format” box. By default, search results are displayed as an html table, but several other options are available.

Example Data Searches: Use Case 1

I want to download all the data for a given program.

This is a simple procedure. Starting at the main MAST web site, <https://archive.stsci.edu/>, under “Mission Search” in the bar across the upper portion on the page, select “Kepler Data.” This will display the Kepler Data Search form, https://archive.stsci.edu/kepler/data_search/search.php.

Put the investigation id (program id) in the box labeled “Investigation_ID.” Note wildcards are automatically used on all program id searches. Figure 3-11 shows the top portion of the results page. Follow the instructions in §3.4 (above) to retrieve the data.

number of rows returned = 137

note: reload page if no results are shown

[Table Info](#)

Mark	Kepler ID	Investigation ID	Dataset Name	Quarter
<input checked="" type="checkbox"/>	3439031	EX_GO20025	KPLR003439031-2010265121752	6
<input checked="" type="checkbox"/>	3439031	EX_GO20025	KPLR003439031-2010355172524	7
<input checked="" type="checkbox"/>	3439031	EX_GO20025	KPLR003439031-2011073133259	8
<input checked="" type="checkbox"/>	3439031	EX_GO20025	KPLR003439031-2011177032512	9
<input checked="" type="checkbox"/>	3440230	EX_GO20023_GO20025	KPLR003440230-2010265121752	6
<input checked="" type="checkbox"/>	3440230	EX_GO20023_GO20025	KPLR003440230-2010355172524	7
<input checked="" type="checkbox"/>	3440230	EX_GO20023_GO20025	KPLR003440230-2011073133259	8
<input checked="" type="checkbox"/>	3440230	EX_GO20023_GO20025	KPLR003440230-2011177032512	9
<input checked="" type="checkbox"/>	3542573	EX_GO20025	KPLR003542573-2010265121752	6
<input checked="" type="checkbox"/>	3542573	EX_GO20025	KPLR003542573-2010355172524	7
<input checked="" type="checkbox"/>	3542573	EX_GO20025	KPLR003542573-2011073133259	8
<input checked="" type="checkbox"/>	3542573	EX_GO20025	KPLR003542573-2011177032512	9
<input checked="" type="checkbox"/>	4740676	EX_GO20025	KPLR004740676-2010265121752	6
<input checked="" type="checkbox"/>	4740676	EX_GO20025	KPLR004740676-2010355172524	7
<input checked="" type="checkbox"/>	4740676	EX_GO20025	KPLR004740676-2011073133259	8
<input checked="" type="checkbox"/>	4740676	EX_GO20025	KPLR004740676-2011177032512	9
<input checked="" type="checkbox"/>	4839180	EX_GO20025	KPLR004839180-2010355172524	7
<input checked="" type="checkbox"/>	4839180	EX_GO20025	KPLR004839180-2011073133259	8
<input checked="" type="checkbox"/>	4839180	EX_GO20025	KPLR004839180-2011177032512	9
<input checked="" type="checkbox"/>	4848423	EX_GO20023_GO20025	KPLR004848423-2010355172524	7
<input checked="" type="checkbox"/>	4848423	EX_GO20023_GO20025	KPLR004848423-2011073133259	8
<input checked="" type="checkbox"/>	4851217	EX_STKL_GO20023_GO20025	KPLR004851217-2010355172524	7
<input checked="" type="checkbox"/>	4851217	EX_STKL_GO20023_GO20025	KPLR004851217-2011073133259	8
<input checked="" type="checkbox"/>	4851217	EX_STKL_STKS_GO20023_GO20025	KPLR004851217-2011177032512	9
<input checked="" type="checkbox"/>	4912991	EX_GO20025	KPLR004912991-2010265121752	6

Figure 3-11 Use case 1 search results

Example Data Searches: Use Case 2

I want to know if there are Kepler data in the archive for cool giant stars.

The Kepler Team has provided a list of targets they identify as Red Giants. To find observations of these targets, use the “Condition Flag” on the Kepler Data Search form as shown in Figure 3-12.

The screenshot shows the 'Kepler Data Search & Retrieval' interface. At the top left is a link for 'Archive Status' and at the top right is a 'Help' link with a sub-link for 'Field Descriptions'. Below these are two tabs: 'Standard Form' (selected) and 'File Upload Form'. The search area contains three buttons: 'Search', 'Reset', and 'Clear Form'. The search criteria are organized into several rows of fields:

- Target Name, Resolver (set to 'Resolve'), Radius (arcmin) (set to 0.02)
- Right Ascension, Declination, Equinox (set to J2000)
- Kepler ID, Investigation ID, 2Mass ID
- KEP Mag, Target Type (with checked options for 'Long Cadence' and 'Short Cadence'), Release Date
- Teff, Log_G, Quarter
- Condition Flag (set to 'Red Giant')

At the bottom, there are four 'User-specified field' sections, each with a dropdown menu (all set to 'Kepler ID') and a 'Field Descriptions' input box.

Figure 3-12 Red Giant search

There may be other targets in the archive that are Red Giants. The KIC values for temperature and surface gravity may be used to perform a search for these data. Set up the Data Search & Retrieval form by putting the temperature and log g ranges in the “Temperature” and “Log G” boxes. Adjust the “Output Columns” and check the “Output Format.” Click “Search” to initiate the search. See Figure 3-13 for the qualified form.

Mark	Kepler ID	Investigation ID	Dataset Name	Quarter
<input type="checkbox"/>	1873513	GO40063	KPLR001873513-2012277125453	14
<input type="checkbox"/>	1873513	GO40063	KPLR001873513-2013011073258	15
<input type="checkbox"/>	1873513	GO40063	KPLR001873513-2013098041711	16
<input type="checkbox"/>	1873513	GO40063	KPLR001873513-2013131215648	17
<input type="checkbox"/>	2860885	EX	KPLR002860885-2009166043257	1
<input type="checkbox"/>	2860885	EX	KPLR002860885-2009259160929	2
<input type="checkbox"/>	2860885	EX	KPLR002860885-2009350155506	3
<input type="checkbox"/>	2860885	EX	KPLR002860885-2010265121752	6
<input type="checkbox"/>	2860885	EX	KPLR002860885-2011271113734	10
<input type="checkbox"/>	2860885	EX	KPLR002860885-2012277125453	14
<input type="checkbox"/>	3430205	EX	KPLR003430205-2009131105131	0
<input type="checkbox"/>	3430205	EX	KPLR003430205-2009166043257	1
<input type="checkbox"/>	3430205	EX	KPLR003430205-2009259160929	2
<input type="checkbox"/>	3430205	EX	KPLR003430205-2009350155506	3
<input type="checkbox"/>	3430205	EX	KPLR003430205-2010078095331	4
<input type="checkbox"/>	3430205	EX	KPLR003430205-2010265121752	6
<input type="checkbox"/>	3430205	EX	KPLR003430205-2010355172524	7
<input type="checkbox"/>	3430205	EX	KPLR003430205-2011271113734	10
<input type="checkbox"/>	3430205	EX	KPLR003430205-2012004120508	11
<input type="checkbox"/>	3430205	EX	KPLR003430205-2012277125453	14
<input type="checkbox"/>	3750264	EX	KPLR003750264-2010174085026	5
<input type="checkbox"/>	3750264	EX	KPLR003750264-2010265121752	6
<input type="checkbox"/>	3750264	EX	KPLR003750264-2010355172524	7
<input type="checkbox"/>	3750264	EX	KPLR003750264-2011073133259	8
<input type="checkbox"/>	3750264	EX	KPLR003750264-2011177032512	9
<input type="checkbox"/>	3750264	EX	KPLR003750264-2011271113734	10
<input type="checkbox"/>	3750264	EX	KPLR003750264-2012004120508	11
<input type="checkbox"/>	3750264	EX	KPLR003750264-2012088054726	12
<input type="checkbox"/>	3750264	EX	KPLR003750264-2012179063303	13
<input type="checkbox"/>	3750264	EX	KPLR003750264-2012277125453	14
<input type="checkbox"/>	3750264	EX	KPLR003750264-2013011073258	15

Figure 3-13 Partial results for a Red Giant search using astrophysical parameters

3.5 Stellar Parameter Tables

The stellar parameters provided by the Kepler Stellar Working Group, which are used in the Kepler pipeline for the DR25, Q1-Q17, processing and search for transiting planets, are accessible via a classic MAST search form. The page is at <http://archive.stsci.edu/kepler/stellar17/search.php>. It is also available from the main menu tab and under the “Stellar 17” menu.

3.6 FFI and Engineering Data Search and Retrieval or Download

Kepler Full Frame Image (FFI) and Engineering data may be located and retrieved via the Kepler Full Frame Image (FFI) & Engineering Data Search page at <https://archive.stsci.edu/kepler/ffi/search.php>. The amount of metadata for FFIs is limited, with “start time”, “end time” and “quarter” being the main fields, and that for Engineering data is even more limited. An example of the FFI Search form is given below in Figure 3-14.

To simply download the existing FFIs, go to <https://archive.stsci.edu/pub/kepler/ffi/>, either directly or from the MAST/Kepler home page via the “Search & Retrieval” tab in the left hand gutter. Click on the filename to download the file.

To display FFIs, without downloading them, go to the FFI display tool, either from the MAST/Kepler home page or directly at https://archive.stsci.edu/kepler/ffi_display.php. The opening page of this tool is shown in Figure 3-15, below.

The FFI & Engineering Data Search Page, linked in the first paragraph of this section, is also used to search for and retrieve engineering data, ancillary engineering data and a variety of ancillary data files. The list of retrievable files is: Long and Short Cadence Collateral (CLL & CLS) files, Artifact Removal Pixel (ARP) files, Background Pixel (BKG) files, and Reverse Clock (RVC) files. See §2.3.5 to §2.3.10. To search for, say, Artifact Removal Pixel files, select Artifact Removal Pixel (ARP) from the pull down menu in the “Archive Class” form element. The page will change to include form elements appropriate for ARP data. For example, form elements for “quarter,” “module” and others will appear. These elements may be qualified to see all ARPs for a given quarter, for a given mod, or both.

Kepler Full Frame Image (FFI) & Engineering Data Search

[Standard Form](#)

[File Upload Form](#)

<p>Archive Class Full Frame Images (FFI) ▾</p> <p>Start Time <input type="text"/></p>	<p>File Name <input type="text"/></p> <p>End Time <input type="text"/></p>	<p>Quarter <input type="text"/></p>
<p>User-specified field 1 Field Descriptions File Name ▾ <input type="text"/></p>	<p>User-specified field 2 Field Descriptions File Name ▾ <input type="text"/></p>	

<p>Output Columns</p> <div style="border: 1px solid #ccc; padding: 5px; min-height: 100px;"><p>Mark</p><p>File Name</p><p>Archive Class</p><p>Start Time</p><p>End Time</p><p>Archive Date</p><p>Quarter</p></div> <p><input type="button" value="up"/> <input type="button" value="down"/> <input type="button" value="remove"/> <input type="button" value="reset"/></p> <p><input type="button" value="add"/> Mark ▾</p> <p><input type="button" value="add all"/> <input type="button" value="remove all"/></p>	<p>Sort By:</p> <p>File Name ▾ <input type="checkbox"/> Reverse</p> <p>null ▾ <input type="checkbox"/> Reverse</p> <p>null ▾ <input type="checkbox"/> Reverse</p> <p>Output Format HTML_Table ▾</p> <p><input type="checkbox"/> Remove Null Columns <input type="checkbox"/> Make Rows Distinct <input type="checkbox"/> Skip formatting</p> <p>Maximum Records: 5001 ▾</p> <p>Records per Page: 500 ▾</p>
--	--

Figure 3-14 Standard FFI and Engineering Data Search

Kepler Data Search

FAQ

Kepler Science Center

Search & Retrieval

Download Options

Data Products

Documentation

Data Release Notes

Data Reduction & Analysis

Related Sites


Publications/News

MAST Services

Images

Data Use Policy

Acknowledgments



Full Frame Image (FFI) Display

This tool enables you to browse the Kepler field of view using the currently available FFI and FFI Uncertainty images, and to identify specific objects observed by GALEX, SDSS, and 2MASS in the Kepler field, as well as Kepler objects for which data are now public, "dropped," or published. The timestamp in the FFI file names are in the format yyyydddhhmmss, where ddd refers to Day of Year (for example, 114 = April 24 UTC), and times refer to the end of the exposure.

See the [README.txt](#) file in the FFI ftp directory for more information on the WCS corrections. Because the spacecraft rotates 90 degrees each quarter (clockwise in the image below), objects will appear in four different CCD channels. For example, in the image [below](#) representing the detector position in quarter 1, a target in channel 1 will appear sequentially in channels 53, 81, and 29, before returning to channel 1 in quarter 5.

For FFIs delivered before Oct. 2011, we recommend using SAOImage ds9 to correctly display the celestial coordinate system. Outdated WCS keywords exist in the FFI headers that cause some FITS viewers, eg. Gaia and Skycat, to display the wrong coordinates. The problem has been remedied in all FFIs delivered after Sep. 31, 2011. More recently, all the FFIs were reprocessed to include a one minute barycentric time correction. The corrected files were made available online at MAST in June, 2013. FFIs were once again reprocessed (with improved WCS values) as part of Data Release 24. The new files replaced the previous online versions on June 2, 2015.

Display channel:

for FFI image:

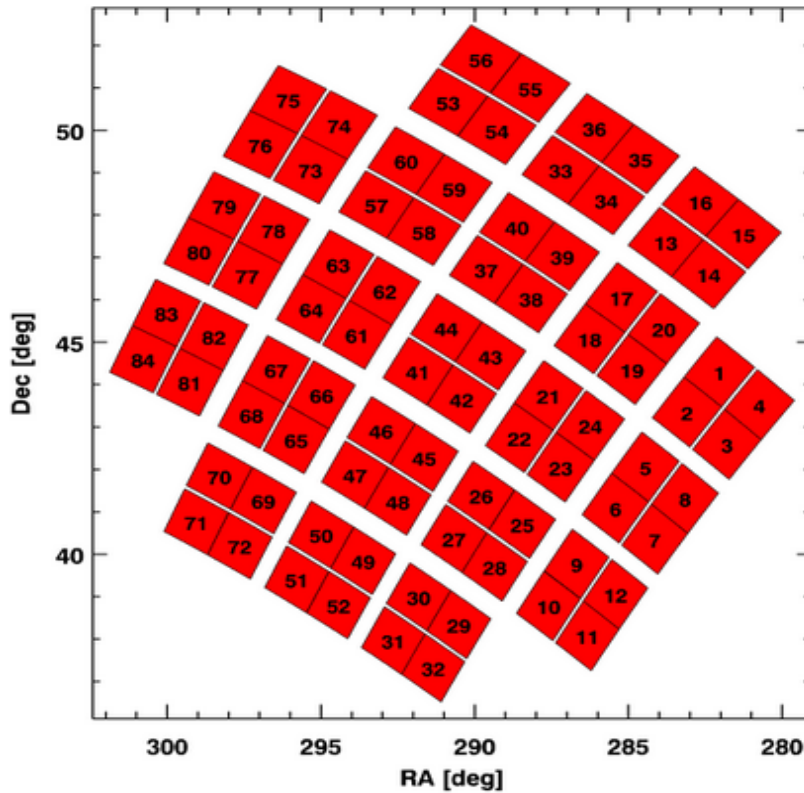


Figure 3-15: FFI Display Tool opening page

3.7 The CasJobs Implementation

MAST has adapted the “CasJobs” tool constructed at Johns Hopkins University as an alternate method to download large data sets. Use of this tool, accessible at <http://mastweb.stsci.edu/kplrcasjobs/>, requires a one-time registration unless users have registered for the MAST/CasJobs tool already, (e.g., for bulk queries of GALEX data).

When should one use CasJobs? Here are several cases:

- **The query’s target list is long.** The interface Form has a target limit of 50,001. Casjobs can be used in its “instant” or “query” (come back later) modes. In the later mode, there is no limit to the number of returns.
- **The use of customized filtering conditions.** The Form necessarily provides a limited number of computed columns, e.g. filter colors. Want to filter on objects with (FUV – i)? Can’t find it on the form! Use CasJobs.
- **The query offers access to other (non-Kepler provided or related) database tables.** Queries can be constructed by imposing conditions or variables from other MAST database tables as well as the primary table used.
- **CasJobs provide the facility of “persistence.”** This allows users to return to the MyDB tables and use them for two or more step queries. This can be helpful when you want to create a list of complicated sets of conditions and do not want to search a multi-million entry database – such queries can be extremely time-intensive and often fail in the first attempt. Moreover, the results from the first step may give the user ideas for additional queries to arrive at the final target list. Finally, results (and their queries) are stored so that you can revisit or modify past searches at any time, something the Classic web forms can not do.

Submit queries using SQL. Searches using the structured query language (SQL) provide more flexibility than a search form interface.

Because many users may not be familiar with SQL, this tool includes its own general help page, GO help page (<http://mastweb.stsci.edu/kplrcasjobs/GOHelpKC.aspx#Part1>), and SQL tutorial. It is essential to read the help pages to avoid common pitfalls, e.g. forgetting to change the ‘context’ tab (which points to a database table) from *kepler* to the user’s own database area (“**MyDB**”). Users may also contact MAST personnel, via archive@stsci.edu, for help with CasJobs and in formulating their queries. In addition, the GO help page is updated whenever new features, such as a new catalog, is added to the CasJobs interface.

Figure 3-16 exhibits the “Query” page of the CasJobs/Kepler. After logging in, users can make queries by first consulting the relevant database tables. This is done by selecting the MyDB tab in the upper menu and the *kepler* context selection in the tables menu tab in the upper left (directly under “Home” and “Help” in the upper menu). This action will create an array of tables in the left panel. Most of the Kepler related tables can now be queried. Clicking “MyDB” will show a list of tables and documentation.

Once you are ready to write and submit a query, click on Query in the upper pane and an open query box will be created. A few SQL sample queries are shown in the indicated tab. These examples can be

customized to return more refined lists without knowing too much about SQL syntax. Use of this query requires consultation of the column names of interest. These can be found in MyDB.

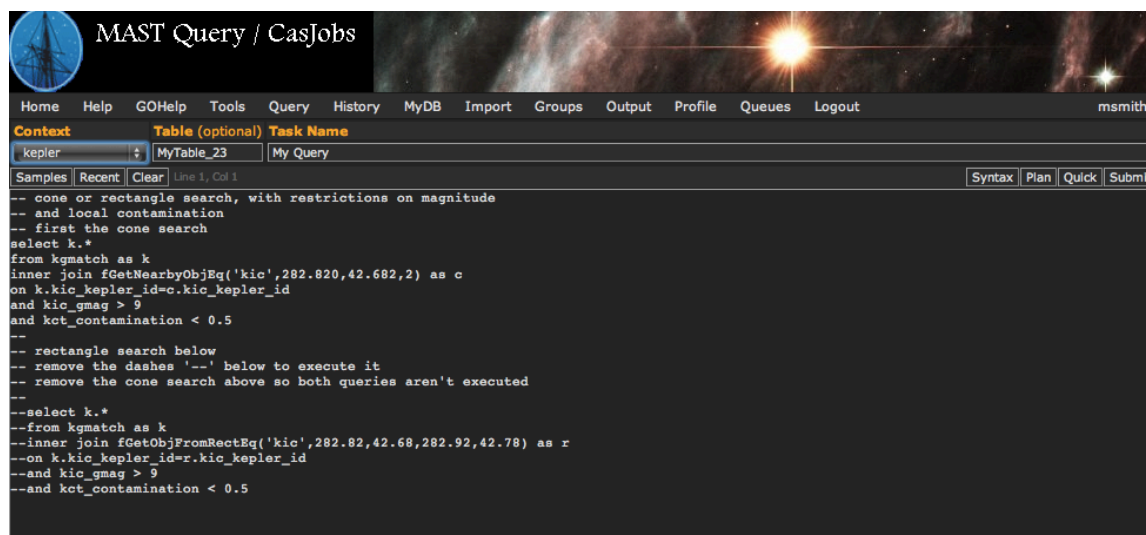


Figure 3-16 The Kepler/GALEX CasJob form. This page can be used to formulate sophisticated queries to constrain a class of objects. This example shows a comment-annotated query for selection of matching objects around a designated area in the sky from the Complete table.

3.8 Other useful tools at MAST

MAST has provided several other ways to retrieve useful data and tables. Here we describe a few that may be of use to the Kepler user. To see a full list, consult the left hand gutter on the Kepler home page at <https://archive.stsci.edu/kepler/> and look under the “Search & Retrieval” menu in this gutter.

3.8.1 Kepler Object of Interest (KOI) Search Form

The Kepler Objects of Interest (KOI) table is produced by the Kepler project and provided to MAST by the NASA Exoplanet Science Institute (NExSci). The table lists planetary candidates, false positives, and confirmed exoplanets. Transit information, stellar parameters, orbital information on possible exoplanets, and disposition fields from the Kepler Project and NExSci are included. The search form can be found at <https://archive.stsci.edu/kepler/koi/search.php>.

Users will also find access to the KOI False Positive Table, and the KOI Astrophysical Positional Probabilities (KAPP) under the Search and Retrieval Tab.

3.8.2 Confirmed Planets Search Forms

Two separate search forms are available to search for Kepler confirmed Exoplanets. The first, Kepler Confirmed Planets KOI Information, at https://archive.stsci.edu/kepler/confirmed_planets/search.php, displays information obtained from the Kepler pipeline processing software. This information, which is also stored in the KOI table, describes the transit estimates and derived orbital parameters. Links are also

provided to display available archived data and to query Simbad for the selected exoplanet. For non-KOI targets, these columns are blank.

Users should take care when entering values in the various KOI search forms. For example, Kepler-68c will return results if entered as the “Target Name.” Other valid forms for the “Target Name” element are kepler-68c, kepler-68 c, and kepler 68 c. Note the spaces in these last two examples. However, for the “Planet Name” element, available in a “User-specified field,” only the form kepler-68 c (**note the space**) yields any returns. Users are encouraged to wild card these values, for example, %68%, to reduce the error rate.

A second interface, Kepler Confirmed Planets Published Information, at https://archive.stsci.edu/kepler/published_planets/search.php lists the same set of confirmed planets but contains information extracted from the published papers by the staff at NExSci. The information available depends on the authors of each paper, so many entries are left undefined, however this is the only table with information on non-KOI exoplanets. As above, the syntax for “Planet Name” is very specific, Kepler-68 c (**note the upper case K and the space**). Use of wildcards (%68 c or %68%) is encouraged.

Both forms should return the same list of exoplanets. Both search results include a link to see the results for the same entry from the other table.

3.8.3 Eclipsing Binaries

The current version of the Kepler Eclipsing Binary catalog, produced by the team at Villanova University, is available at https://archive.stsci.edu/kepler/eclipsing_binaries.html. Columns can be sorted by clicking the column headings, and links are provided for displaying archived data, displaying light curve plots from Villanova, or for seeing an overview of the target as produced by Villanova. Work on this catalog is still continuing and columns with missing data will be populated as the Eclipsing Binary team finishes their analysis. Note a few targets with multiple periods have multiple entries. More information on the catalog can be found at <http://keplerebs.villanova.edu/>.

3.8.4 The Kepler-GALEX Crossmatch

The Kepler-GALEX cross match (KGxmatch) was created to mitigate the problem of selecting blue objects from the KIC/Kepler Target tables by providing GALEX magnitudes as a substitute for u-band photometry, a value that does not exist in the KIC. Although GALEX observations do not cover the entire Kepler field of view, there is sufficient coverage to warrant performing a cross match.

MAST cross-matched the GALEX Release #6 “mcat” catalog, with the Kepler Target table version of the KIC. Our positional results are contained in two Kepler/Cross matched tables described below. We expect to update the crossmatch using GALEX Release #7 in 2016.

No extensive list of cross-matched objects observed by two missions is likely to be perfect. Our tables use as the criterion for matching the angular separation between positions of objects in the KIC and GALEX catalogs, and this criterion alone is not always sufficient. For example, the best apparent match to a KIC entry may be a GALEX catalog entry for which the closest match is some other KIC entry. To accommodate such ambiguities, MAST has generated two catalogs: a Complete (alternatively, “KGMATCH”) and an Accurate (“GoldStandard”) table.

The Complete table gives all possible GALEX matches to each KIC entry within a search radius of 5". This table in general gives multiple matches (and reverse matches), ranked by increasing separation. The Complete table returns possible GALEX matches of Kepler entries out to 5", and vice versa; all potential matches are ranked by distance from the Kepler entries.

The Accurate table gives all unique matches for search radii out to 2.5", both in the KIC-to-GALEX and GALEX-to-KIC match directions. Although this table is incomplete - because it misses those rare correct matches to GALEX entries with coordinates just beyond the 2.5" search radius - the matches are unambiguous. Note that while this table is designated "Accurate", it can occasionally generate a false match if the correct match is to a Kepler entry that has a GALEX entry (and vice versa) that does not have the very closest coordinates. Such errors are common in crowded fields and for data collated from different bandpasses and at the edge of detectors where field distortions are greatest.

Note also that both tables are subject to the greatest errors for GALEX AIS (all sky) survey tiles, for which the exposures are short. There are 81 GALEX sky tiles (each a circle of radius ~0.6 degrees) that overlap the Kepler FOV, of which 79 were observed in both FUV and NUV bands. The user can expect that most GALEX objects will have both magnitudes represented. Users should note that the number of matched objects is limited by both the incomplete GALEX areal coverage and the brighter faint magnitude limit relative to the KIC's.

Kepler-GALEX Crossmatch interface form

MAST provides two interfaces to the KGxmatch results. One interface is a classic MAST web form that allows searches of either the Complete or the Accurate table, accessible at <https://archive.stsci.edu/kepler/kgmatch/search.php>. The other is CASJobs (see the next section) The classic interface will be familiar to most MAST users, and with the aid of a field description page, usage should be almost self-explanatory. The table allows uploading of target coordinates, and requires checking of either the Complete or Accurate table box. The default is Accurate. This form is recommended for relatively simple queries that do not require more filter conditions than those fields exhibited on the form. The returned fields include fluxes and colors from GALEX, SCP-Sloan, and 2MASS missions. The output is available in several formats.

3.8.5 Retrieving CDPP values

The *Kepler* Mission provides the root mean square Combined Differential Photometric Precision (CDPP) values for every target observed. The values are calculated for various time scales, chosen to reflect typical transit durations. The CDPP values for the 3-, 6- and 12-hour time scales are provided each quarter. For more information on how this noise metric is calculated see Jenkins et al. (2010). These values are available in the light curve data headers and in the archive catalog. See the definition in §2.4.

For a small number of objects, users may obtain the CDPP values via the standard MAST/Kepler search form by adding `cdpp3`, `cdpp6` and/or `cdpp12` to the Output Columns. For an intermediate number of objects, users should use an *http get* request (or, similarly using the GNU Wget unix command). Examples are available on line at the following link: https://archive.stsci.edu/vo/mast_services.html. In addition to the examples, the page contains links to field names for each mission. The field names are needed as part of the *http get* request. For the CDPP values, and any other data related search, use the "Kepler data" link. Any of the fields may be used to qualify the results of an *http get* request.

For example, what are the CDP3 values for investigation GO20010? Note the use of “max_records” and “outputformat” in the request. A partial listing of the returns is given below the example. NB: if you cut and paste this get request, remove the blanks added by the end of lines.

https://archive.stsci.edu/kepler/data_search/search.php?kvc_investigation_id=GO20010&max_records=2000&selectedColumnsCsv=kvc_kepler_id,sci_data_quarter,sci_Cdpp3_0,sci_Cdpp6_0,sci_Cdpp12_0&outputformat=CSV&action=Search

```
Kepler ID,Quarter,,,
integer,integer
2968820,6,276.9526,251.3615,228.124
2968820,7,260.9334,240.7988,226.9803
2968820,8,188.7593,184.3977,177.6128
2968820,9,181.8454437255859,176.8077697753906,174.5309295654297
2984406,6,224.5499,205.8647,193.4872
2984406,7,223.7221,205.881,191.9032
2984406,8,153.0623,148.5383,146.9009
2984406,9,138.8018188476563,135.3896789550781,133.1756439208984
3096721,6,200.063,200.4038,190.1847
```

Magnitude and color cuts can also be specified. Here the search is for data where $J > 16$ and $(J-K) < 2$. A partial list of results is shown below the example. NB: if you cut and paste this get request, remove the blanks added by the end of lines.

https://archive.stsci.edu/kepler/data_search/search.php?twoMass_jmag>16&twoMass_jkcolor<2&max_records=3000&selectedColumnsCsv=kvc_kepler_id,sci_data_quarter,sci_Cdpp3_0,sci_Cdpp6_0,sci_Cdpp12_0&outputformat=CSV&action=Search

```
Kepler ID,Quarter,,,
integer,integer
757076,0,107.1907272338867,87.06771087646484,74.18241882324219
757076,1,98.90160369873048,74.26993560791016,55.08943557739258
757076,2,107.1909103393555,84.28044128417969,66.48330688476562
757076,3,107.0135803222656,82.63387298583984,62.61486434936523
757076,4,105.0075378417969,80.99000549316406,63.43953704833984
757076,5,106.4529,76.9312,53.2947
757076,6,109.9427,80.9282,60.1158
757076,7,104.9784,77.6593,55.3902
757076,8,104.5721,79.6476,57.4747
```

3.9 Alternate Methods for Downloading Data

For archive users who do not wish to search for Kepler data, but rather simply download, say, all the public data, or their GO data, MAST provides some alternative means for downloading data. The “Download Options” menu the left hand gutter of the MAST Kepler home page goes to a page that contains details on the light curve bundles, directory syntax, scripting examples (wget, Curl), a direct link to downloadable catalogs and other useful information. Below we describe how to access and download the quarterly-based tar files of Kepler light curves.

Public Data Download

Most of Kepler's data are staged in a directory area that is available through anonymous ftp or through a browser. These data are found at <https://archive.stsci.edu/pub/kepler/lightcurves/tarfiles/>. Each quarterly set of data has been broken into smaller files per quarter, each no larger than 5 GB. This was done because large tarfiles were presenting a significant problem for some users. The syntax for these smaller files is EX_QX_I.tgz, where X is the quarter and I is an integer starting at 1. The data are also tarred together by investigation ID (e.g. EX, STKS, GO) in the “teamfiles” directory for those looking for only a subset of the data.

To access the light curves do the following (X represents the quarter number):

```
ftp archive.stsci.edu
login (as anonymous or with your archive user name and password)
cd /pub/kepler/lightcurves/tarfiles/QXteamfiles
bin
mget *QX*
```

A set of wget scripts is also provided that, when executed, will download the light curves. These scripts are located in the same directory as the tarfiles, <https://archive.stsci.edu/pub/kepler/lightcurves/tarfiles/>. Consult the README file in their directory for details of the scripts.

There are subdirectories and tarfiles containing lightcurves for objects identified as Red Giants, Eclipsing Binaries and KOIs. Consult the README file for important information on the KOI version, as creation of the KOI tarfiles lags the Kepler Mission announcement of a new KOI release. The subdirectories are at <https://archive.stsci.edu/pub/kepler/lightcurves/tarfiles/>.

References

All project documents (KSCI, KADN) are available on the MAST Kepler web site:
<http://archive.stsci.edu/kepler/documents.html>

Allen, C., Klaus, T., and Jenkins, J. M., Proc. SPIE, Vol. 7740, 77401E

Borucki, W. J. et al., 2010, ApJL, L126

Brown et al., 2011, AJ, 142, 112

Bryson, S. T., Jenkins, J. M., Klaus, T. C. et al., 2010, Proc. SPIE, Vol. 7740, 77401D

Bryson, S. T., Tenenbaum, P.; Jenkins, J. M. et al., 2011, ApJ, 713, L97

Burke, C. et al. 2015,

Calabretta, M. R., and Greisen, E. W., 2002, A&A, 395, 1077

Caldwell, D. A., Van Cleve, J. E., Jenkins, J. M., et al., 2010, Proc. SPIE, Vol. 7731, 773117

Christansen, J. L., Jenkins, J. M., Caldwell, D. A., Barclay, T. et al., 2013, Kepler Data Characteristics Handbook, KSCI-19040-004

Coughlin, J. L., Mullally, F., Thompson, S. E., et al., 2016, ApJS Accepted,
<http://arxiv.org/abs/1512.06149>

Dunham, T., et al, 2010, ApJL, L136

[Eastman, J. et al. 2010, PASP, 122, 935](#)

[Everett, M., Howell, S. B. & Kinemuchi, K., 2012, PASP, 124, 316E](#)

Fanelli, M. N., Jenkins, J., Bryson, S. T. et al., 2011, “Kepler Data Processing Handbook”, KSCI-19081-001

Gilliland, R. L., Chaplin, W. J., Dunham, E. W., et al., 2011, ApJS, 197, 6

Goldstein, H., 1980, *Classical Mechanics*, pp. 143-166 and 608-610

Greisen, E. W., and Calabretta, M. R., 2002, A&A, 395, 1061

[Greiss et al. 2012, AJ, 144, 24G](#)

Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398

Jenkins, J. M., Chandrasekaran, C., McCauliff, S. D., et al., 2010, Proc. SPIE, Vol. 7740, 77400D

Jenkins, J. M., et al. 2010, ApJL, L87

Jenkins, J. M., et al., 2016, “Kepler Data Processing Handbook”, KSCI-19081-002, in prep

Koch, D. et al. 2010, ApJL, L79

Kolodziejczak, J. J., Caldwell, D. A., Van Cleve, J. E., et al., 2010, Proc. SPIE, Vol. 7742, 77421G

Pence et al. 2010 A&A, 524, A42

Quintana, E. V., Jenkins, J. M., Clarke, B. D., et al., 2010, Proc. SPIE, Vol. 7740, 77401X

Schupe et al., 2005, ASPC, 347, 491

Smith J. C. et al., 2012, PASP, 124, 1000

Stumpe, M. et al., 2012, PASP, 124, 985

Stumpe, M. et al. 2014, PASP, 126, 100

Tenenbaum, P., and Jenkins, J. M, 2010, Proc. SPIE, Vol. 7740, 77401C

Van Cleve, J. E., 2008, KADN-26205

Van Cleve, J. E., 2010, KADN-26285

Van Cleve, J. E., 2016a, KADN-26196

Van Cleve, J. E., 2016b, KADN-26197

Van Cleve, J. E., and Caldwell, D. A., 2009, “Kepler Instrument Handbook”, KSCI-19033-001

Appendices

Appendix A. Calibrated Data Headers – Light Curve and Target Pixel Files

A.1: Kepler Light Curve File Headers

A.1a: Light Curve File Primary Header

This header describes how the data was taken and processed along with information about the target contained in the file. These keywords describe the instrument and season the data was collected. CREATOR describes the code and version of the code that created the file. PROCVER describes the version of the pipeline that processed the data. DATA_REL relates which version of the data release notes describes these data. FILEVER contains the version of the file format.

Example Primary Header

```
SIMPLE = T / conforms to FITS standards
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T / file contains extensions
NEXTEND = 2 / number of standard extensions
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1 / extension version number (not format version)
ORIGIN = 'NASA/Ames' / institution responsible for creating this file
DATE = '2013-10-18' / file creation date.
CREATOR = '558083 FluxExporter2PipelineModule' / pipeline job and program used t
PROCVER = 'svn+ssh://murzim/repo/soc/tags/release/9.1.4 r53267' / SW version
FILEVER = '6.1 ' / file format version
TIMVERSN= 'OGIP/93-003' / OGIP memo number for file format
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 7206837' / string version of KEPLERID
KEPLERID= 7206837 / unique Kepler target identifier
CHANNEL = 46 / CCD channel
SKYGROUP= 62 / roll-independent location of channel
MODULE = 14 / CCD module
OUTPUT = 2 / CCD output
QUARTER = 17 / Observing quarter
SEASON = 3 / mission season during which data was collected
DATA_REL= 25 / version of data release notes for this file
OBSMODE = 'long cadence' / observing mode
RADESYS = 'ICRS ' / reference frame of celestial coordinates
RA_OBJ = 293.765570 / [deg] right ascension
DEC_OBJ = 42.737940 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
PMRA = 0.0000 / [arcsec/yr] RA proper motion
PMDEC = 0.0000 / [arcsec/yr] Dec proper motion

PMTOTAL = 0.0000 / [arcsec/yr] total proper motion

PARALLAX= / [arcsec] parallax
GLON = 75.711056 / [deg] galactic longitude
GLAT = 10.706512 / [deg] galactic latitude
GMAG = 10.021 / [mag] SDSS g band magnitude
RMAG = 9.691 / [mag] SDSS r band magnitude
IMAG = 9.661 / [mag] SDSS i band magnitude
ZMAG = 9.648 / [mag] SDSS z band magnitude
```

```

D51MAG =          9.865 / [mag] D51 magnitude,
JMAG   =          8.827 / [mag] J band magnitude from 2MASS
HMAG   =          8.628 / [mag] H band magnitude from 2MASS
KMAG   =          8.574 / [mag] K band magnitude from 2MASS
KEPMAG =          9.769 / [mag] Kepler magnitude (Kp)
GRCOLOR =         0.330 / [mag] (g-r) color, SDSS bands
JKCOLOR =         0.253 / [mag] (J-K) color, 2MASS bands
GKCOLOR =         1.447 / [mag] (g-K) color, SDSS g - 2MASS K
TEFF   =          6304 / [K] Effective temperature
LOGG   =          4.169 / [cm/s2] log10 surface gravity
FEH    =          0.140 / [log10([Fe/H])] metallicity
EBMINUSV=         0.037 / [mag] E(B-V) reddening
AV     =          0.115 / [mag] A_v extinction
RADIUS =          1.590 / [solar radii] stellar radius
TMINDEX =        1304921800 / unique 2MASS catalog ID
SCPID  =          / unique SCP processing ID
CHECKSUM= '9fiGBciD9ciDAciD' / HDU checksum updated 2013-10-18T19:39:21Z
END

```

A.1b: Light Curve File LIGHTCURVE Extension Header

This header describes the detector and the time period over which the data was collected.

Example LIGHTCURVE header

XTENSION= 'BINTABLE' / marks the beginning of a new HDU
BITPIX = 8 / array data type
NAXIS = 2 / number of array dimensions
NAXIS1 = 100 / length of first array dimension
NAXIS2 = 4573 / length of second array dimension
PCOUNT = 0 / group parameter count (not used)
GCOUNT = 1 / group count (not used)
TFIELDS = 20 / number of table fields
TTYPE1 = 'TIME ' / column title: data time stamps
TFORM1 = 'D ' / column format: 64-bit floating point
TUNIT1 = 'BJD - 2454833' / column units: barycenter corrected JD
TDISP1 = 'D14.7 ' / column display format
TTYPE2 = 'TIMECORR' / column title: barycenter - timeslice correction
TFORM2 = 'E ' / column format: 32-bit floating point
TUNIT2 = 'd ' / column units: day
TDISP2 = 'E13.6 ' / column display format
TTYPE3 = 'CADENCENO' / column title: unique cadence number
TFORM3 = 'J ' / column format: signed 32-bit integer
TDISP3 = 'I10 ' / column display format
TTYPE4 = 'SAP_FLUX' / column title: aperture photometry flux
TFORM4 = 'E ' / column format: 32-bit floating point
TUNIT4 = 'e-/s ' / column units: electrons per second
TDISP4 = 'E14.7 ' / column display format
TTYPE5 = 'SAP_FLUX_ERR' / column title: aperture phot. flux error
TFORM5 = 'E ' / column format: 32-bit floating point
TUNIT5 = 'e-/s ' / column units: electrons per second (1-sigma)
TDISP5 = 'E14.7 ' / column display format
TTYPE6 = 'SAP_BKG ' / column title: aperture phot. background flux
TFORM6 = 'E ' / column format: 32-bit floating point
TUNIT6 = 'e-/s ' / column units: electrons per second
TDISP6 = 'E14.7 ' / column display format
TTYPE7 = 'SAP_BKG_ERR' / column title: ap. phot. background flux error
TFORM7 = 'E ' / column format: 32-bit floating point
TUNIT7 = 'e-/s ' / column units: electrons per second (1-sigma)
TDISP7 = 'E14.7 ' / column display format
TTYPE8 = 'PDCSAP_FLUX' / column title: aperture phot. PDC flux
TFORM8 = 'E ' / column format: 32-bit floating point
TUNIT8 = 'e-/s ' / column units: electrons per second

TDISP8 = 'E14.7 ' / column display format
TTYPER9 = 'PDCSAP_FLUX_ERR' / column title: ap. phot. PDC flux error
TFORM9 = 'E ' / column format: 32-bit floating point
TUNIT9 = 'e-/s ' / column units: electrons per second (1-sigma)
TDISP9 = 'E14.7 ' / column display format
TTYPER10 = 'SAP_QUALITY' / column title: aperture photometry quality flag
TFORM10 = 'J ' / column format: signed 32-bit integer
TDISP10 = 'B16.16 ' / column display format
TTYPER11 = 'PSF_CENTR1' / column title: PSF-fitted column centroid
TFORM11 = 'D ' / column format: 64-bit floating point
TUNIT11 = 'pixel ' / column units: pixel
TDISP11 = 'F10.5 ' / column display format
TTYPER12 = 'PSF_CENTR1_ERR' / column title: PSF-fitted column error
TFORM12 = 'E ' / column format: 32-bit floating point
TUNIT12 = 'pixel ' / column units: pixel (1-sigma)
TDISP12 = 'E14.7 ' / column display format
TTYPER13 = 'PSF_CENTR2' / column title: PSF-fitted row centroid
TFORM13 = 'D ' / column format: 64-bit floating point
TUNIT13 = 'pixel ' / column units: pixel
TDISP13 = 'F10.5 ' / column display format
TTYPER14 = 'PSF_CENTR2_ERR' / column title: PSF-fitted row error
TFORM14 = 'E ' / column format: 32-bit floating point
TUNIT14 = 'pixel ' / column units: pixel (1-sigma)
TDISP14 = 'E14.7 ' / column display format
TTYPER15 = 'MOM_CENTR1' / column title: moment-derived column centroid
TFORM15 = 'D ' / column format: 64-bit floating point
TUNIT15 = 'pixel ' / column units: pixel
TDISP15 = 'F10.5 ' / column display format
TTYPER16 = 'MOM_CENTR1_ERR' / column title: moment-derived column error
TFORM16 = 'E ' / column format: 32-bit floating point
TUNIT16 = 'pixel ' / column units: pixel (1-sigma)
TDISP16 = 'E14.7 ' / column display format
TTYPER17 = 'MOM_CENTR2' / column title: moment-derived row centroid
TFORM17 = 'D ' / column format: 64-bit floating point
TUNIT17 = 'pixel ' / column units: pixel
TDISP17 = 'F10.5 ' / column display format
TTYPER18 = 'MOM_CENTR2_ERR' / column title: moment-derived row error
TFORM18 = 'E ' / column format: 32-bit floating point
TUNIT18 = 'pixel ' / column units: pixel (1-sigma)
TDISP18 = 'E14.7 ' / column display format
TTYPER19 = 'POS_CORR1' / column title: column position correction
TFORM19 = 'E ' / column format: 32-bit floating point

TUNIT19 = 'pixels ' / column units: pixel
TDISP19 = 'E14.7 ' / column display format
TTYPER20 = 'POS_CORR2' / column title: row position correction
TFORM20 = 'E ' / column format: 32-bit floating point
TUNIT20 = 'pixels ' / column units: pixel
TDISP20 = 'E14.7 ' / column display format
INHERIT = T / inherit the primary header
EXTNAME = 'LIGHTCURVE' / name of extension
EXTVER = 1 / extension version number (not format version)
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 8191672' / string version of target id
KEPLERID= 8191672 / unique Kepler target identifier
RADESYS = 'ICRS ' / reference frame of celestial coordinates
RA_OBJ = 299.407028 / [deg] right ascension
DEC_OBJ = 44.035052 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
EXPOSURE= 86.02766139 / [d] time on source
TIMEREFF = 'SOLARSYSTEM' / barycentric correction applied to times
TASSIGN = 'SPACECRAFT' / where time is assigned
TIMESYS = 'TDB ' / time system is barycentric JD
BJDREFI = 2454833 / integer part of BJD reference date
BJDREFF = 0.00000000 / fraction of the day in BJD reference date
TIMEUNIT= 'd ' / time unit for TIME, TSTART and TSTOP
TELAPSE = 93.44383910 / [d] TSTOP - TSTART
LIVETIME= 86.02766139 / [d] TELAPSE multiplied by DEADC
TSTART = 906.83540556 / observation start time in BJD-BJDREF
TSTOP = 1000.27924466 / observation stop time in BJD-BJDREF
LC_START= 55739.34343282 / mid point of first cadence in MJD
LC_END = 55832.76587042 / mid point of last cadence in MJD
DEADC = 0.92063492 / deadtime correction
TIMEPIXR= 0.5 / bin time beginning=0 middle=0.5 end=1
TIERRELA= 5.78E-07 / [d] relative time error
TIERABSO= / [d] absolute time error
INT_TIME= 6.019802903270 / [s] photon accumulation time per frame
READTIME= 0.518948526144 / [s] readout time per frame
FRAMETIM= 6.538751429414 / [s] frame time (INT_TIME + READTIME)
NUM_FRM = 270 / number of frames per time stamp
TIMEDEL = 0.02043359821692 / [d] time resolution of data
DATE-OBS= '2011-06-27T07:59:49.865Z' / TSTART as UTC calendar date
DATE-END= '2011-09-28T18:37:33.934Z' / TSTOP as UTC calendar date
BACKAPP = T / background is subtracted

DEADAPP = T / deadtime applied
VIGNAPP = T / vignetting or collimator correction applied
GAIN = 112.32 / [electrons/count] channel gain
READNOIS= 80.690688 / [electrons] read noise
NREADOUT= 270 / number of read per cadence
TIMSLICE= 1 / time-slice readout sequence section
MEANBLCK= 742 / [count] FSW mean black level
LCFXDOFF= 419400 / long cadence fixed offset
SCFXDOFF= 219400 / short cadence fixed offset
CDPP3_0 = 54.561180114746094 / [ppm] RMS CDPP on 3.0-hr time scales
CDPP6_0 = 40.61515808105469 / [ppm] RMS CDPP on 6.0-hr time scales
CDPP12_0= 31.42128562927246 / [ppm] RMS CDPP on 12.0-hr time scales
CROWDSAP= 0.9967 / Ratio of target flux to total flux in op. ap.
FLFRCSAP= 0.8995 / Frac. of target flux w/in the op. aperture
NSPSDDDET= 0 / Number of SPSDs detected
NSPSDCOR= 0 / Number of SPSDs corrected
PDCVAR = 1.274828553199768 / Target variability
PDCMETHD= 'multiScaleMap' / PDC algorithm used for target
NUMBAND = 3 / Number of scale bands
FITTYPE1= 'robust ' / Fit type used for band 1
PR_GOOD1= 0.8318247199058533 / Prior goodness for band 1
PR_WGHT1= 0.0 / Prior weight for band 1
FITTYPE2= 'prior ' / Fit type used for band 2
PR_GOOD2= 0.6127007007598877 / Prior goodness for band 2
PR_WGHT2= 63.41262435913086 / Prior weight for band 2
FITTYPE3= 'prior ' / Fit type used for band 3
PR_GOOD3= 0.7650205492973328 / Prior goodness for band 3
PR_WGHT3= 79.17725372314453 / Prior weight for band 3
PDC_TOT = 0.5447844862937927 / PDC total goodness metric for target
PDC_TOTP= 0.346148818731308 / PDC_TOT percentile compared to mod/out
PDC_COR = 0.9994268417358398 / PDC correlation goodness metric for target
PDC_CORP= 85.61112976074219 / PDC_COR percentile compared to mod/out
PDC_VAR = 0.9866594076156616 / PDC variability goodness metric for target
PDC_VARP= 54.49952697753906 / PDC_VAR percentile compared to mod/out
PDC_NOI = 0.8432802557945251 / PDC noise goodness metric for target
PDC_NOIP= 46.72538375854492 / PDC_NOI percentile compared to mod/out
PDC_EPT = 1.0 / PDC earth point goodness metric for target
PDC_EPTP= 79.75627136230469 / PDC_EPT percentile compared to mod/out
CHECKSUM= 'ZqDfgnDdZnDdfnDd' / HDU checksum updated 2015-10-02T20:42:26Z
END

A.1c Light Curve File Aperture Extension Header

The aperture extension describes the mask for the designated target in the file.

Example Aperture Header

```
XTENSION= 'IMAGE ' / marks the beginning of a new HDU
BITPIX = 32 / array data type
NAXIS = 2 / number of array dimensions
NAXIS1 = 11 / length of first array dimension
NAXIS2 = 20 / length of second array dimension
PCOUNT = 0 / group parameter count (not used)
GCOUNT = 1 / group count (not used)
INHERIT = T / inherit the primary header
EXTNAME = 'APERTURE' / name of extension
EXTVER = 1 / extension version number (not format version)
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 7206837' / string version of KEPLERID
KEPLERID= 7206837 / unique Kepler target identifier
RADESYS = 'ICRS ' / reference frame of celestial coordinates
RA_OBJ = 293.765570 / [deg] right ascension
DEC_OBJ = 42.737940 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
WCSAXES = 2 / number of WCS axes
CTYPE1 = 'RA---TAN' / right ascension coordinate type
CTYPE2 = 'DEC--TAN' / declination coordinate type
CRPIX1 = 5.734425136711934 / [pixel] reference pixel along image axis 1
CRPIX2 = 9.020901042644141 / [pixel] reference pixel along image axis 2
CRVAL1 = 293.7655695 / [deg] right ascension at reference pixel
CRVAL2 = 42.73794 / [deg] declination at reference pixel
CUNIT1 = 'deg ' / physical unit in column dimension
CUNIT2 = 'deg ' / physical unit in row dimension
CDEL1 = -0.001103289223486 / [deg] pixel scale in RA dimension
CDEL2 = 0.001103289223486305 / [deg] pixel scale in Dec dimension
PC1_1 = 0.8574946837902978 / linear transformation element cos(th)
PC1_2 = -0.5119469045520156 / linear transformation element -sin(th)
PC2_1 = -0.5127083574373209 / linear transformation element sin(th)
PC2_2 = -0.8600875986916059 / linear transformation element cos(th)
WCSNAMEP= 'PHYSICAL' / name of world coordinate system alternate P
WCSAXESP= 2 / number of WCS physical axes
CTYPE1P = 'RAWX ' / physical WCS axis 1 type CCD col
CUNIT1P = 'PIXEL ' / physical WCS axis 1 unit
CRPIX1P = 1 / reference CCD column
CRVAL1P = 229 / value at reference CCD column
CDEL1P = 1.0 / physical WCS axis 1 step
CTYPE2P = 'RAWY ' / physical WCS axis 2 type CCD row
CUNIT2P = 'PIXEL ' / physical WCS axis 2 units
CRPIX2P = 1 / reference CCD row
CRVAL2P = 915 / value at reference CCD row
CDEL2P = 1.0 / physical WCS axis 2 step
NPIXSAP = 29 / Number of pixels in optimal aperture
NPIXMISS= 0 / Number of op. aperture pixels not collected
CHECKSUM= '1VohmUoZlUoflUoZ' / HDU checksum updated 2013-10-18T19:39:21Z
```

A.2: Target Pixel File Headers

A.2a Target Pixel File Primary Header

This header describes how the data was taken and processed along with information about the target contained in the file. The data is processed in the Science Operations Center at NASA/Ames. CREATOR describes the code and version of the code that created the file. PROCVER describes the version of the pipeline that processed the data. DATA_REL relates which version of the data release notes describes these data.

Example Primary Header

```
SIMPLE = T / conforms to FITS standards
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T / file contains extensions
NEXTEND = 2 / number of standard extensions
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1 / extension version number (not format version)
ORIGIN = 'NASA/Ames' / institution responsible for creating this file
DATE = '2013-10-21' / file creation date.
CREATOR = '562203 TargetPixelExporterPipelineModule' / pipeline job and program
PROCVER = 'svn+ssh://murzim/repo/soc/tags/release/9.3.29 r60484' / SW version
FILEVER = '6.1 ' / file format version
TIMVERSN= 'OGIP/93-003' / OGIP memo number for file format
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 7206837' / string version of KEPLERID
KEPLERID= 7206837 / unique Kepler target identifier
CHANNEL = 46 / CCD channel
SKYGROUP= 62 / roll-independent location of channel
MODULE = 14 / CCD module
OUTPUT = 2 / CCD output
QUARTER = 17 / Observing quarter
SEASON = 3 / mission season during which data was collected
DATA_REL= 25 / version of data release notes for this file
OBSMODE = 'long cadence' / observing mode
RADESYS = 'ICRS ' / reference frame of celestial coordinates
RA_OBJ = 293.765570 / [deg] right ascension
DEC_OBJ = 42.737940 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
PMRA = 0.0000 / [arcsec/yr] RA proper motion
PMDEC = 0.0000 / [arcsec/yr] Dec proper motion
PMTOTAL = 0.0000 / [arcsec/yr] total proper motion
PARALLAX= / [arcsec] parallax
GLON = 75.711056 / [deg] galactic longitude
GLAT = 10.706512 / [deg] galactic latitude
GMAG = 10.021 / [mag] SDSS g band magnitude
RMAG = 9.691 / [mag] SDSS r band magnitude
IMAG = 9.661 / [mag] SDSS i band magnitude
ZMAG = 9.648 / [mag] SDSS z band magnitude
D51MAG = 9.865 / [mag] D51 magnitude,
JMAG = 8.827 / [mag] J band magnitude from 2MASS
HMAG = 8.628 / [mag] H band magnitude from 2MASS
KMAG = 8.574 / [mag] K band magnitude from 2MASS
```

```

KEPMAG =          9.769 / [mag] Kepler magnitude (Kp)
GRCOLOR =         0.330 / [mag] (g-r) color, SDSS bands
JKCOLOR =         0.253 / [mag] (J-K) color, 2MASS bands
GKCOLOR =         1.447 / [mag] (g-K) color, SDSS g - 2MASS K
TEFF =           6304 / [K] Effective temperature
LOGG =           4.169 / [cm/s2] log10 surface gravity
FEH =            0.140 / [log10([Fe/H])] metallicity
EBMINUSV=        0.037 / [mag] E(B-V) reddening
AV =             0.115 / [mag] A_v extinction
RADIUS =          1.590 / [solar radii] stellar radius
TMINDEX =        1304921800 / unique 2MASS catalog ID
SCPID =          / unique SCP processing ID
CHECKSUM= 'Nc8ZOc7ZNc7ZNc7Z' / HDU checksum updated 2013-10-21T15:05:21Z
END

```

A.2b Target Pixel File TARGETTABLES Header

Format of binary table. Each column of the binary table is described with the FITS standard keywords TTYPE, TFORM and TUNIT. Columns are made of images and have the keyword TDIM to specify the dimensions of the image. Each image also contains keywords to specify the row and column values of the pixels on the specified channel.

Example TARGETTABLES Header

XTENSION= 'BINTABLE' / marks the beginning of a new HDU
BITPIX = 8 / array data type
NAXIS = 2 / number of array dimensions
NAXIS1 = 588 / length of first array dimension
NAXIS2 = 4573 / length of second array dimension
PCOUNT = 0 / group parameter count (not used)
GCOUNT = 1 / group count (not used)
TFIELDS = 13 / number of table fields
TTYPE1 = 'TIME ' / column title: data time stamps
TFORM1 = 'D ' / column format: 64-bit floating point
TUNIT1 = 'BJD - 2454833' / column units: barycenter corrected JD
TDISP1 = 'D14.7 ' / column display format
TTYPE2 = 'TIMECORR' / column title: barycenter - timeslice correction
TFORM2 = 'E ' / column format: 32-bit floating point
TUNIT2 = 'd ' / column units: day
TDISP2 = 'E14.7 ' / column display format
TTYPE3 = 'CADENCENO' / column title: unique cadence number
TFORM3 = 'J ' / column format: signed 32-bit integer
TDISP3 = 'I10 ' / column display format
TTYPE4 = 'RAW_CNTS' / column title: raw pixel counts
TFORM4 = '20J ' / column format: image of signed 32-bit integers
TUNIT4 = 'count ' / column units: count

TDISP4 = 'I8 ' / column display format
TDIM4 = '(5,4) ' / column dimensions: pixel aperture array
TNULL4 = -1 / column null value indicator
WCSN4P = 'PHYSICAL' / table column WCS name
WCAX4P = 2 / table column physical WCS dimensions
1CTY4P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY4P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN4P = 'PIXEL ' / table column physical WCS axis 1 unit
2CUN4P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV4P = 120 / table column physical WCS ax 1 ref value
2CRV4P = 298 / table column physical WCS ax 2 ref value
1CDL4P = 1.0 / table column physical WCS a1 step
2CDL4P = 1.0 / table column physical WCS a2 step
1CRP4P = 1 / table column physical WCS a1 reference
2CRP4P = 1 / table column physical WCS a2 reference
WCAX4 = 2 / number of WCS axes
1CTYP4 = 'RA---TAN' / right ascension coordinate type
2CTYP4 = 'DEC--TAN' / declination coordinate type
1CRPX4 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX4 = 2.772054486225443 / [pixel] reference pixel along image axis 2
1CRVL4 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL4 = 47.13156 / [deg] declination at reference pixel
1CUNI4 = 'deg ' / physical unit in column dimension
2CUNI4 = 'deg ' / physical unit in row dimension
1CDLT4 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT4 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC4 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC4 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC4 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC4 = 0.8459233094122672 / linear transformation matrix element cos(th)
TTYPE5 = 'FLUX ' / column title: calibrated pixel flux
TFORM5 = '20E ' / column format: image of 32-bit floating point
TUNIT5 = 'e-/s ' / column units: electrons per second
TDISP5 = 'E14.7 ' / column display format
TDIM5 = '(5,4) ' / column dimensions: pixel aperture array
WCSN5P = 'PHYSICAL' / table column WCS name
WCAX5P = 2 / table column physical WCS dimensions
1CTY5P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY5P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN5P = 'PIXEL ' / table column physical WCS axis 1 unit
2CUN5P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV5P = 120 / table column physical WCS ax 1 ref value

2CRV5P = 298 / table column physical WCS ax 2 ref value
1CDL5P = 1.0 / table column physical WCS a1 step
2CDL5P = 1.0 / table column physical WCS a2 step
1CRP5P = 1 / table column physical WCS a1 reference
2CRP5P = 1 / table column physical WCS a2 reference
WCAX5 = 2 / number of WCS axes
1CTYP5 = 'RA---TAN' / right ascension coordinate type
2CTYP5 = 'DEC--TAN' / declination coordinate type
1CRPX5 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX5 = 2.772054486225443 / [pixel] reference pixel along image axis 2
1CRVL5 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL5 = 47.13156 / [deg] declination at reference pixel
1CUNI5 = 'deg ' / physical unit in column dimension
2CUNI5 = 'deg ' / physical unit in row dimension
1CDLT5 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT5 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC5 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC5 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC5 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC5 = 0.8459233094122672 / linear transformation matrix element cos(th)
TTYPER6 = 'FLUX_ERR' / column title: 1-sigma calibrated uncertainty
TFORM6 = '20E ' / column format: image of 32-bit floating point
TUNIT6 = 'e-/s ' / column units: electrons per second (1-sigma)
TDISP6 = 'E14.7 ' / column display format
TDIM6 = '(5,4) ' / column dimensions: pixel aperture array
WCASN6P = 'PHYSICAL' / table column WCS name
WCAX6P = 2 / table column physical WCS dimensions
1CTY6P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY6P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN6P = 'PIXEL ' / table column physical WCS axis 1 unit
2CUN6P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV6P = 120 / table column physical WCS ax 1 ref value
2CRV6P = 298 / table column physical WCS ax 2 ref value
1CDL6P = 1.0 / table column physical WCS a1 step
2CDL6P = 1.0 / table column physical WCS a2 step
1CRP6P = 1 / table column physical WCS a1 reference
2CRP6P = 1 / table column physical WCS a2 reference
WCAX6 = 2 / number of WCS axes
1CTYP6 = 'RA---TAN' / right ascension coordinate type
2CTYP6 = 'DEC--TAN' / declination coordinate type
1CRPX6 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX6 = 2.772054486225443 / [pixel] reference pixel along image axis 2

1CRVL6 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL6 = 47.13156 / [deg] declination at reference pixel
1CUNI6 = 'deg ' / physical unit in column dimension
2CUNI6 = 'deg ' / physical unit in row dimension
1CDLT6 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT6 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC6 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC6 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC6 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC6 = 0.8459233094122672 / linear transformation matrix element cos(th)
TTYPER7 = 'FLUX_BKG' / column title: calibrated background flux
TFORM7 = '20E ' / column format: image of 32-bit floating point
TUNIT7 = 'e-/s ' / column units: electrons per second
TDISP7 = 'E14.7 ' / column display format
TDIM7 = '(5,4) ' / column dimensions: pixel aperture array
WCSN7P = 'PHYSICAL' / table column WCS name
WCAX7P = 2 / table column physical WCS dimensions
1CTY7P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY7P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN7P = 'PIXEL ' / table column physical WCS axis 1 unit
2CUN7P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV7P = 120 / table column physical WCS ax 1 ref value
2CRV7P = 298 / table column physical WCS ax 2 ref value
1CDL7P = 1.0 / table column physical WCS a1 step
2CDL7P = 1.0 / table column physical WCS a2 step
1CRP7P = 1 / table column physical WCS a1 reference
2CRP7P = 1 / table column physical WCS a2 reference
WCAX7 = 2 / number of WCS axes
1CTYP7 = 'RA---TAN' / right ascension coordinate type
2CTYP7 = 'DEC---TAN' / declination coordinate type
1CRPX7 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX7 = 2.772054486225443 / [pixel] reference pixel along image axis 2
1CRVL7 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL7 = 47.13156 / [deg] declination at reference pixel
1CUNI7 = 'deg ' / physical unit in column dimension
2CUNI7 = 'deg ' / physical unit in row dimension
1CDLT7 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT7 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC7 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC7 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC7 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC7 = 0.8459233094122672 / linear transformation matrix element cos(th)

TTYPE8 = 'FLUX_BKG_ERR' / column title: 1-sigma cal. background uncertain
TFORM8 = '20E ' / column format: image of 32-bit floating point
TUNIT8 = 'e-/s ' / column units: electrons per second (1-sigma)
TDISP8 = 'E14.7 ' / column display format
TDIM8 = '(5,4) ' / column dimensions: pixel aperture array
WCSN8P = 'PHYSICAL' / table column WCS name
WCAX8P = 2 / table column physical WCS dimensions
1CTY8P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY8P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN8P = 'PIXEL ' / table column physical WCS axis 1 unit
2CUN8P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV8P = 120 / table column physical WCS ax 1 ref value
2CRV8P = 298 / table column physical WCS ax 2 ref value
1CDL8P = 1.0 / table column physical WCS a1 step
2CDL8P = 1.0 / table column physical WCS a2 step
1CRP8P = 1 / table column physical WCS a1 reference
2CRP8P = 1 / table column physical WCS a2 reference
WCAX8 = 2 / number of WCS axes
1CTYP8 = 'RA---TAN' / right ascension coordinate type
2CTYP8 = 'DEC--TAN' / declination coordinate type
1CRPX8 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX8 = 2.772054486225443 / [pixel] reference pixel along image axis 2
1CRVL8 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL8 = 47.13156 / [deg] declination at reference pixel
1CUNI8 = 'deg ' / physical unit in column dimension
2CUNI8 = 'deg ' / physical unit in row dimension
1CDLT8 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT8 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC8 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC8 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC8 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC8 = 0.8459233094122672 / linear transformation matrix element cos(th)
TTYPE9 = 'COSMIC_RAYS' / column title: cosmic ray detections
TFORM9 = '20E ' / column format: image of 32-bit floating point
TUNIT9 = 'e-/s ' / column units: electrons per second
TDISP9 = 'E14.7 ' / column display format
TDIM9 = '(5,4) ' / column dimensions: pixel aperture array
WCSN9P = 'PHYSICAL' / table column WCS name
WCAX9P = 2 / table column physical WCS dimensions
1CTY9P = 'RAWX ' / table column physical WCS axis 1 type, CCD col
2CTY9P = 'RAWY ' / table column physical WCS axis 2 type, CCD row
1CUN9P = 'PIXEL ' / table column physical WCS axis 1 unit

2CUN9P = 'PIXEL ' / table column physical WCS axis 2 unit
1CRV9P = 120 / table column physical WCS ax 1 ref value
2CRV9P = 298 / table column physical WCS ax 2 ref value
1CDL9P = 1.0 / table column physical WCS a1 step
2CDL9P = 1.0 / table column physical WCS a2 step
1CRP9P = 1 / table column physical WCS a1 reference
2CRP9P = 1 / table column physical WCS a2 reference
WCAX9 = 2 / number of WCS axes
1CTYP9 = 'RA---TAN' / right ascension coordinate type
2CTYP9 = 'DEC---TAN' / declination coordinate type
1CRPX9 = 3.8390115151408395 / [pixel] reference pixel along image axis 1
2CRPX9 = 2.772054486225443 / [pixel] reference pixel along image axis 2
1CRVL9 = 291.2423505 / [deg] right ascension at reference pixel
2CRVL9 = 47.13156 / [deg] declination at reference pixel
1CUNI9 = 'deg ' / physical unit in column dimension
2CUNI9 = 'deg ' / physical unit in row dimension
1CDLT9 = -0.001099218131143 / [deg] pixel scale in RA dimension
2CDLT9 = 0.001099218131143 / [deg] pixel scale in DEC dimension
11PC9 = -0.8391590040673064 / linear transformation matrix element cos(th)
12PC9 = 0.5368907516025715 / linear transformation matrix element -sin(th)
21PC9 = 0.5404001415376445 / linear transformation matrix element sin(th)
22PC9 = 0.8459233094122672 / linear transformation matrix element cos(th)
TTYPER10 = 'QUALITY ' / column title: pixel quality flags
TFORM10 = 'J ' / column format: signed 32-bit integer
TDISP10 = 'B24.24 ' / column display format
TTYPER11 = 'POS_CORR1' / column title: column position correction
TFORM11 = 'E ' / column format: 32-bit floating point
TUNIT11 = 'pixel ' / column units: pixel
TDISP11 = 'E14.7 ' / column display format
TTYPER12 = 'POS_CORR2' / column title: row position correction
TFORM12 = 'E ' / column format: 32-bit floating point
TUNIT12 = 'pixel ' / column units: pixel
TDISP12 = 'E14.7 ' / column display format
TTYPER13 = 'RB_LEVEL' / column title: rolling band level
TFORM13 = '20E ' / column format: image of 32-bit floating point
TUNIT13 = 'sigma ' / column units: detection significance
TDISP13 = 'E14.7 ' / column display format
TDIM13 = '(5,4) ' / column dimensions: pixel aperture array
INHERIT = T / inherit the primary header
EXTNAME = 'TARGETTABLES' / name of extension
EXTVER = 1 / extension version number (not format version)
TELESCOP= 'Kepler ' / telescope

INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 10139564' / string version of target id
KEPLERID= 10139564 / unique Kepler target identifier
RADESYS = 'ICRS ' / reference frame of celestial coordinates
RA_OBJ = 291.242351 / [deg] right ascension
DEC_OBJ = 47.131560 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
EXPOSURE= 84.44683458 / [d] time on source
TIMEREFSYS = 'SOLARSYSTEM' / barycentric correction applied to times
TASSIGN = 'SPACECRAFT' / where time is assigned
TIMESYS = 'TDB ' / time system is barycentric JD
BJDREFI = 2454833 / integer part of BJD reference date
BJDREFF = 0.00000000 / fraction of the day in BJD reference date
TIMEUNIT= 'd ' / time unit for TIME, TSTART and TSTOP
TELAPSE = 93.44315636 / [d] TSTOP - TSTART
LIVETIME= 86.02703284 / [d] TELAPSE multiplied by DEADC
TSTART = 906.83547052 / observation start time in BJD-BJDREF
TSTOP = 1000.27862688 / observation stop time in BJD-BJDREF
LC_START= 55739.34343282 / mid point of first cadence in MJD
LC_END = 55832.76587042 / mid point of last cadence in MJD
DEADC = 0.92063492 / deadtime correction
TIMEPIXR= 0.5 / bin time beginning=0 middle=0.5 end=1
TIERRELA= 5.78E-07 / [d] relative time error
TIERABSO= / [d] absolute time error
INT_TIME= 6.019802903270 / [s] photon accumulation time per frame
READTIME= 0.518948526144 / [s] readout time per frame
FRAMETIM= 6.538751429414 / [s] frame time (INT_TIME + READTIME)
NUM_FRM = 270 / number of frames per time stamp
TIMEDEL = 0.02043359821692 / [d] time resolution of data
DATE-OBS= '2011-06-27T07:59:49.865Z' / TSTART as UTC calendar date
DATE-END= '2011-09-28T18:37:33.934Z' / TSTOP as UTC calendar date
BACKAPP = T / background is subtracted
DEADAPP = T / deadtime applied
VIGNAPP = T / vignetting or collimator correction applied
GAIN = 103.04 / [electrons/count] channel gain
READNOIS= 89.572672 / [electrons] read noise
NREADOUT= 270 / number of read per cadence
TIMSLICE= 1 / time-slice readout sequence section
MEANBLCK= 739 / [count] FSW mean black level
LCFXDOFF= 419400 / long cadence fixed offset
SCFXDOFF= 219400 / short cadence fixed offset
CDPP3_0 = / [ppm] RMS CDPP on 3.0-hr time scales

CDPP6_0 = / [ppm] RMS CDPP on 6.0-hr time scales
CDPP12_0= / [ppm] RMS CDPP on 12.0-hr time scales
CROWDSAP= 0.7705 / Ratio of target flux to total flux in op. ap.
FLFRCSAP= 0.8273 / Frac. of target flux w/in the op. aperture
DBCOLCO = 1000 / Column cutoff used by Dynablack
DBTHRES = 2.8321184431653808 / [e-/s] Flux threshold used by Dynablack
RBTDUR1 = 3 / rootname of residual baseline image
RBTDUR2 = 6 / rootname of residual baseline image
RBTDUR3 = 12 / rootname of residual baseline image
RBTDUR4 = 24 / rootname of residual baseline image
RBTDUR5 = 31 / rootname of residual baseline image
BLKALGO = 'exponentialOneDBlack' / black algorithm used
CHECKSUM= 'gfhIjZf9gffGgZf9' / HDU checksum updated 2015-10-02T22:25:31Z
END

A.2c Target Pixel File Aperture Header

The aperture extension describes the mask for the designated target in the file. The physical coordinates of the bottom left pixel of the mask is given by CRVAL1P and CRVAL2P and the RA and Dec are provided as WCS keywords according to the FITS standard.

Example Aperture Header

```
XTENSION= 'IMAGE' / marks the beginning of a new HDU
BITPIX = 32 / array data type
NAXIS = 2 / number of array dimensions
NAXIS1 = 11 / length of first array dimension
NAXIS2 = 20 / length of second array dimension
PCOUNT = 0 / group parameter count (not used)
GCOUNT = 1 / group count (not used)
INHERIT = T / inherit the primary header
EXTNAME = 'APERTURE' / name of extension
EXTVER = 1 / extension version number (not format version)
TELESCOP= 'Kepler' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBJECT = 'KIC 7206837' / string version of KEPLERID
KEPLERID= 7206837 / unique Kepler target identifier
RADESYS = 'ICRS' / reference frame of celestial coordinates
RA_OBJ = 293.765570 / [deg] right ascension
DEC_OBJ = 42.737940 / [deg] declination
EQUINOX = 2000.0 / equinox of celestial coordinate system
WCSAXES = 2 / number of WCS axes
CTYPE1 = 'RA---TAN' / right ascension coordinate type
CTYPE2 = 'DEC--TAN' / declination coordinate type
CRPIX1 = 5.734425136711934 / [pixel] reference pixel along image axis 1
CRPIX2 = 9.020901042644141 / [pixel] reference pixel along image axis 2
CRVAL1 = 293.7655695 / [deg] right ascension at reference pixel
CRVAL2 = 42.73794 / [deg] declination at reference pixel
CUNIT1 = 'deg' / physical unit in column dimension
CUNIT2 = 'deg' / physical unit in row dimension
CDEL1 = -0.001103289223486 / [deg] pixel scale in RA dimension
CDEL2 = 0.001103289223486305 / [deg] pixel scale in Dec dimension
PC1_1 = 0.8574946837902978 / linear transformation element cos(th)
PC1_2 = -0.5119469045520156 / linear transformation element -sin(th)
PC2_1 = -0.5127083574373209 / linear transformation element sin(th)
PC2_2 = -0.8600875986916059 / linear transformation element cos(th)
WCSNAMEP= 'PHYSICAL' / name of world coordinate system alternate P
WCSAXESP= 2 / number of WCS physical axes
CTYPE1P = 'RAWX' / physical WCS axis 1 type CCD col
CUNIT1P = 'PIXEL' / physical WCS axis 1 unit
CRPIX1P = 1 / reference CCD column
CRVAL1P = 229 / value at reference CCD column
CDEL1P = 1.0 / physical WCS axis 1 step
CTYPE2P = 'RAWY' / physical WCS axis 2 type CCD row
CUNIT2P = 'PIXEL' / physical WCS axis 2 units
```

```
CRPIX2P =          1 / reference CCD row
CRVAL2P =         915 / value at reference CCD row
CDELTA2P =         1.0 / physical WCS axis 2 step
NPIXSAP =          29 / Number of pixels in optimal aperture
NPIXMISS=          0 / Number of op. aperture pixels not collected
CHECKSUM= '3W0d3V0a3V0a3V0a' / HDU checksum updated 2013-10-21T15:05:21Z
END
```

Appendix B. Full Frame Image Headers

B.1 Calibrated Full Frame Image (FFI) Primary Header

The FFIs contain a primary header and 84 extension headers, one for each mod/out. The FFI are in units of electrons per second.

```
SIMPLE = T / conforms to FITS standards
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T / file contains extensions
NEXTEND = 84 / number of standard extensions
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1 / extension version number (not format version)
ORIGIN = 'NASA/Ames' / institution responsible for creating this file
DATE = '2011-08-26' / file creation date.
CREATOR = '198129 ffi assembler' / pipeline job and program used to produce this
PROCV = 'svn+ssh://murzim/repo/soc/branches/integ/8.1-il r44395' / SW version
FILEVER = '3.1 ' / file format version
TIMVERSN= 'OGIP/93-003' / OGIP memo number for file format
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBSMODE = 'full frame image' / observing mode
DATSETNM= 'kplr2009292020429' / data set name
DCT_TIME= '2009292020429' / data collection time: yyyydddhhss
DCT_TYPE= 'FFI ' / data type
DCT_PURP= 'Monthly FFI' / purpose of data
IMAGTYPE= 'SocCal ' / FFI image type: raw, SocCal, SocUnc
QUARTER = 3 / Observing quarter
DATA_REL= / SOC Data release version
SEASON = 1 / mission season during which data was collected
FINE_PNT= T / fine point pointing status during accumulation
MMNTMDMP= F / momentum dump occurred during accumulation
SCCONFIG= 57 / commanded S/C configuration ID
PIXELTYP= 'all ' / pixel type: target, background, coll., all
REV_CLCK= F / reverse clocking in effect?

VSMRSROW= 1046 / collateral virtual smear region start row
VSMREROW= 1057 / collateral virtual smear region row end
NROWVSMR= 12 / number of rows binned in virtual smear
VSMRCOL = 12 / collateral virtual smear region start column
VSMRECOL= 1111 / collateral virtual smear region end column
NCOLVSMR= 1100 / number of columns in virtual smear region
MASKSROW= 6 / science collateral masked region start row
MASKEROW= 17 / science collateral masked region end row
NROWMASK= 12 / number of rows binned in masked region
MASKSCOL= 12 / science collateral masked region start
```

```

MASKECOL=          1111 / science collateral masked region end
NCOLMASK=          1100 / number of columns in masked region
BLCKSROW=           0 / science collateral black region start row
BLCKEROW=         1069 / science collateral black region end column
NROWBLCK=          1070 / number of rows in black region
BLCKSCOL=          1118 / science collateral black region start
BLCKECOL=          1131 / science collateral black region end column
NCOLBLK =           1070 / number of columns binned in black region
OPERTEMP=          198129.0 / [C] commanded FPA temperature set point
FOCPOS1 =           -27.168 / [microns] mechanism 1 focus position
FOCPOS2 =            78.7336 / [microns] mechanism 2 focus position
FOCPOS3 =          -67.9754 / [microns] mechanism 3 focus position
RADESYS = 'ICRS      ' / reference frame of celestial coordinates
EQUINOX =           2000.0 / equinox of celestial coordinate system
RA_NOM  =           290.667 / [deg] RA of spacecraft boresight
DEC_NOM =            44.5 / [deg] declination of spacecraft boresight
ROLL_NOM=           200.0 / [deg] roll angle of spacecraft
CHECKSUM= '6F7m6D616D616D61' / HDU checksum updated 2011-08-26T00:06:47Z
END

```

B.2 Calibrated FFI Channel Header

There are 84 extensions in total, one for each mod/out.

```

XTENSION= 'IMAGE      ' / marks the beginning of a new HDU
BITPIX  =           -32 / array data type
NAXIS   =             2 / NAXIS
NAXIS1  =           1132 / length of first array dimension
NAXIS2  =           1070 / length of second array dimension
PCOUNT  =             0 / group parameter count (not used)
GCOUNT  =             1 / group count (not used)
INHERIT =             T / inherit the primary header
EXTNAME = 'MOD.OUT 2.1' / name of extension
EXTVER  =             1 / extension version number (not format version)
TELESCOP= 'Kepler     ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
CHANNEL =             1 / CCD channel
SKYGROUP=           53 / roll-independent location of channel
MODULE  =             2 / CCD module
OUTPUT  =             1 / CCD output
TIMEREf = 'SOLARSYSTEM' / barycentric correction applied to times
TASSIGN = 'SPACECRAFT' / where time is assigned
TIMESYS = 'TDB        ' / time system is barycentric JD
MJDSTART=          55123.06602474 / [d] start of observation in spacecraft MJD
MJDEND  =          55123.0864583 / [d] end of observation in spacecraft MJD
BJDREFI =          2454833 / integer part of BJD reference date

```

```

BJDREFF =          0.00000000 / fraction of the day in BJD reference date
TIMEUNIT= 'd      '          / time unit for TIME, TSTART and TSTOP
TSTART  =          290.56733021 / observation start time in BJD-BJDREF
TSTOP   =          290.58776377 / observation stop time in BJD-BJDREF
TELAPSE =          0.02043356 / [d] TSTOP - TSTART
EXPOSURE=          0.01881185 / [d] time on source
LIVETIME=          0.01881185 / [d] TELAPSE multiplied by DEADC
DEADC   =          0.92063492 / deadtime correction
TIMEPIXR=          0.5 / bin time beginning=0 middle=0.5 end=1
TIERRELA=          5.78E-07 / [d] relative time error
INT_TIME=          6.019802903270 / [s] photon accumulation time per frame
READTIME=          0.518948526144 / [s] readout time per frame
FRAMETIM=          6.5387514294144005 / [s] frame time (INT_TIME + READTIME)
NUM_FRM =          270 / number of frames per time stamp
FGSFRPER=          103.7897052288 / [ms] FGS frame period
NUMFGSFP=          58 / number of FGS frame periods per exposure
TIMEDEL =          0.02043356 / [d] time resolution of data
DATE-OBS= '2009-10-19T01:35:04.537Z' / TSTART as UTC calendar date
DATE-END= '2009-10-19T02:04:29.997Z' / TSTOP as UTC calendar date
BTC_PIX1=          536.0 / reference col for barycentric time correction
BTC_PIX2=          567.0 / reference row for barycentric time correction
BUNIT   = 'electrons/s'      / physical units of image data
BARYCORR=          1.3054676E-03 / [d] barycentric time correction
BACKAPP =          F / background is subtracted
DEADAPP =          T / deadtime applied
VIGNAPP =          T / vignetting or collimator correction applied
READNOIS=          87.604748 / [electrons] read noise
NREADOUT=          270 / number of read per cadence
TIMSLICE=          4 / time-slice readout sequence section
MEANBLCK=          727.0 / [count] FSW mean black level
RADESYS = 'ICRS      '      / reference frame of celestial coordinates
EQUINOX =          2000.0 / equinox of celestial coordinate system
WCSNAMEP= 'PHYSICAL'      / name of world coordinate system alternate P
WCSAXESP=          2 / number of WCS physical axes
CTYPE1P = 'RAWX      '      / physical WCS axis 1 type CCD col
CUNIT1P = 'PIXEL     '      / physical WCS axis 1 unit
CRPIX1P =          1 / reference CCD column
CRVAL1P =          0 / value at reference CCD column
CDEL1P =          1.0 / physical WCS axis 1 step
CTYPE2P = 'RAWY      '      / physical WCS axis 2 type CCD row
CUNIT2P = 'PIXEL     '      / physical WCS axis 2 units
CRPIX2P =          1 / reference CCD row
CRVAL2P =          0 / value at reference CCD row
CDEL2P =          1.0 / physical WCS axis 2 step
CTYPE1  = 'RA---TAN-SIP'    / Gnomonic projection + SIP distortions
CTYPE2  = 'DEC--TAN-SIP'    / Gnomonic projection + SIP distortions
CRVAL1  =          299.2617001144113 / RA at CRPIX1, CRPIX2

```

```

CRVAL2 = 44.01502810934495 / DEC at CRPIX1, CRPIX2
CRPIX1 = 533.0 / X reference pixel
CRPIX2 = 521.0 / Y reference pixel
CD1_1 = -0.0005004780578635 / Transformation matrix
CD1_2 = 9.850651507779759E-4 / Transformation matrix
CD2_1 = 9.847518004452874E-4 / Transformation matrix
CD2_2 = 4.998471812902435E-4 / Transformation matrix
A_ORDER = 2 / Polynomial order, axis 1
B_ORDER = 2 / Polynomial order, axis 2
A_2_0 = 6.340797800196409E-7 / distortion coefficient
A_0_2 = 1.749990982371E-07 / distortion coefficient
A_1_1 = 5.382453881118889E-7 / distortion coefficient
B_2_0 = 1.175242752788E-07 / distortion coefficient
B_0_2 = 7.164025680025224E-7 / distortion coefficient
B_1_1 = 5.496396742485739E-7 / distortion coefficient
AP_ORDER= 2 / Inv polynomial order, axis 1
BP_ORDER= 2 / Inv polynomial order, axis 2
AP_1_0 = 3.76588935122466E-6 / inv distortion coefficient
AP_0_1 = -4.851187332073E-06 / inv distortion coefficient
AP_2_0 = -6.40464415381175E-7 / inv distortion coefficient
AP_0_2 = -1.767332402001E-07 / inv distortion coefficient
AP_1_1 = -5.343953631364E-07 / inv distortion coefficient
BP_1_0 = -2.637543835256E-06 / inv distortion coefficient
BP_0_1 = -1.547946592422E-06 / inv distortion coefficient
BP_2_0 = -1.064163345654E-07 / inv distortion coefficient
BP_0_2 = -7.220223503958E-07 / inv distortion coefficient
BP_1_1 = -5.404568839297E-07 / inv distortion coefficient
A_DMAX = 0.3818408549009291 / maximum distortion, axis 1
B_DMAX = 0.40086247834881306 / maximum distortion, axis 2
CHECKSUM= 'ZARXd8PXZAPXd5PX' / HDU checksum updated 2011-08-26T00:05:45Z
END

```

B.3 FFI Uncertainties Primary Header

The FFI uncertainties contain a primary header and 84 extensions.

```

SIMPLE = T / conforms to FITS standards
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T / file contains extensions
NEXTEND = 84 / number of standard extensions
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1 / extension version number (not format version)
ORIGIN = 'NASA/Ames' / institution responsible for creating this file
DATE = '2011-08-26' / file creation date.
CREATOR = '198129 ffi assembler' / pipeline job and program used to produce this
PROCVR = 'svn+ssh://murzim/repo/soc/branches/integ/8.1-i1 r44395' / SW version
FILEVER = '2.0 ' / file format version
TIMVERSN= 'OGIP/93-003' / OGIP memo number for file format
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
OBSMODE = 'full frame image' / observing mode

```



```

DATSETNM= 'kplr2009292020429' / data set name
DCT_TIME= '2009292020429' / data collection time: yyyyddhhss
DCT_TYPE= 'FFI' / data type
DCT_PURP= 'Monthly FFI' / purpose of data
IMAGTYPE= 'SocCal' / FFI image type: raw, SocCal, SocUnc
QUARTER = 3 / Observing quarter
SEASON = 1 / mission season during which data was collected
FINE_PNT= T / fine point pointing status during accumulation
MMNTMDMP= F / momentum dump occurred during accumulation
SCCONFIG= 57 / commanded S/C configuration ID
PIXELTYP= 'all' / pixel type: target, background, collateral, all
REV_CLCK= F / reverse clocking in effect?
VSMRSROW= 1046 / collateral virtual smear region start row
VSMREROW= 1057 / collateral virtual smear region row end
NROWVSMR= 12 / number of rows binned in virtual smear
VSMRCOL = 12 / collateral virtual smear region start column
VSMRECOL= 1111 / collateral virtual smear region end column
NCOLVSMR= 1100 / number of columns in virtual smear region
MASKSROW= 6 / science collateral masked region start row
MASKEROW= 17 / science collateral masked region end row
NROWMASK= 12 / number of rows binned in masked region
MASKSCOL= 12 / science collateral masked region start
MASKECOL= 1111 / science collateral masked region end
NCOLMASK= 1100 / number of columns in masked region
BLCKSROW= 0 / science collateral black region start row
BLCKEROW= 1069 / science collateral black region end column
NROWBLCK= 1070 / number of rows in black region
BLCKSCOL= 1118 / science collateral black region start
BLCKECOL= 1131 / science collateral black region end column
NCOLBLK = 1070 / number of columns binned in black region
OPERTEMP= 198129.0 / [C] commanded FPA temperature set point
FOCPOS1 = -27.168 / [microns] mechanism 1 focus position
FOCPOS2 = 78.7336 / [microns] mechanism 2 focus position
FOCPOS3 = -67.9754 / [microns] mechanism 3 focus position
RADESYS = 'ICRS' / reference frame of celestial coordinates
EQUINOX = 2000.0 / equinox of celestial coordinate system
RA_NOM = 290.667 / [deg] RA of spacecraft boresight
DEC_NOM = 44.5 / [deg] declination of spacecraft boresight
ROLL_NOM= 200.0 / [deg] roll angle of spacecraft
CHECKSUM= '6F7m6D6l6D6l6D6l' / HDU checksum updated 2011-08-26T00:06:47Z
END

```

B.4 Uncertainties FFI Channel Header

```

XTENSION= 'IMAGE' / marks the beginning of a new HDU
BITPIX = -32 / array data type
NAXIS = 2 / NAXIS
NAXIS1 = 1132 / length of first array dimension
NAXIS2 = 1070 / length of second array dimension

```

```

PCOUNT = 0 / group parameter count (not used)
GCOUNT = 1 / group count (not used)
INHERIT = T / inherit the primary header
EXTNAME = 'MOD.OUT 2.1' / name of extension
EXTVER = 1 / extension version number (not format version)
TELESCOP= 'Kepler ' / telescope
INSTRUME= 'Kepler Photometer' / detector type
CHANNEL = 1 / CCD channel
SKYGROUP= 53 / roll-independent location of channel
MODULE = 2 / CCD module
OUTPUT = 1 / CCD output
TIMEREF = 'SOLARSYSTEM' / barycentric correction applied to times
TASSIGN = 'SPACECRAFT' / where time is assigned
TIMESYS = 'TDB ' / time system is barycentric JD
MJDSTART= 55123.06602474 / [d] start of observation in spacecraft MJD
MJDEND = 55123.0864583 / [d] end of observation in spacecraft MJD
BJDREFI = 2454833 / integer part of BJD reference date
BJDREFF = 0.00000000 / fraction of the day in BJD reference date
TIMEUNIT= 'd ' / time unit for TIME, TSTART and TSTOP
TSTART = 290.56733021 / observation start time in BJD-BJDREF
TSTOP = 290.58776377 / observation stop time in BJD-BJDREF
TELAPSE = 0.02043356 / [d] TSTOP - TSTART
EXPOSURE= 0.01881185 / [d] time on source
LIVETIME= 0.01881185 / [d] TELAPSE multiplied by DEADC
DEADC = 0.92063492 / deadtime correction
TIMEPIXR= 0.5 / bin time beginning=0 middle=0.5 end=1
TIERRELA= 5.78E-07 / [d] relative time error
INT_TIME= 6.019802903270 / [s] photon accumulation time per frame
READTIME= 0.518948526144 / [s] readout time per frame
FRAMETIM= 6.5387514294144005 / [s] frame time (INT_TIME + READTIME)
NUM_FRM = 270 / number of frames per time stamp
FGSFRPER= 103.7897052288 / [ms] FGS frame period
NUMFGSFP= 58 / number of FGS frame periods per exposure
TIMEDEL = 0.02043356 / [d] time resolution of data
DATE-OBS= '2009-10-19T01:35:04.537Z' / TSTART as UTC calendar date
DATE-END= '2009-10-19T02:04:29.997Z' / TSTOP as UTC calendar date
BTC_PIX1= 536.0 / reference col for barycentric time correction
BTC_PIX2= 567.0 / reference row for barycentric time correction
BUNIT = 'electrons/s' / physical units of image data
BARYCORR= 1.3054676E-03 / [d] barycentric time correction
BACKAPP = F / background is subtracted
DEADAPP = T / deadtime applied
VIGNAPP = T / vignetting or collimator correction applied
READNOIS= 87.604748 / [electrons] read noise
NREADOUT= 270 / number of read per cadence
TIMSLICE= 4 / time-slice readout sequence section
MEANBLCK= 727.0 / [count] FSW mean black level
RADESYS = 'ICRS ' / reference frame of celestial coordinates
EQUINOX = 2000.0 / equinox of celestial coordinate system

```

```

WCSNAMEP= 'PHYSICAL'           / name of world coordinate system alternate P
WCSAXESP=                2 / number of WCS physical axes
CTYPE1P = 'RAWX   '           / physical WCS axis 1 type CCD col
CUNIT1P = 'PIXEL  '           / physical WCS axis 1 unit
CRPIX1P =                  1 / reference CCD column
CRVAL1P =                  0 / value at reference CCD column
CDELTP1 =                  1.0 / physical WCS axis 1 step
CTYPE2P = 'RAWY   '           / physical WCS axis 2 type CCD row
CUNIT2P = 'PIXEL  '           / physical WCS axis 2 units
CRPIX2P =                  1 / reference CCD row
CRVAL2P =                  0 / value at reference CCD row
CDELTP2 =                  1.0 / physical WCS axis 2 step
CTYPE1  = 'RA---TAN-SIP'      / Gnomonic projection + SIP distortions
CTYPE2  = 'DEC--TAN-SIP'      / Gnomonic projection + SIP distortions
CRVAL1  =    299.2617001144113 / RA at CRPIX1, CRPIX2
CRVAL2  =    44.01502810934495 / DEC at CRPIX1, CRPIX2
CRPIX1  =                533.0 / X reference pixel
CRPIX2  =                521.0 / Y reference pixel
CD1_1   =   -0.0005004780578635 / Transformation matrix
CD1_2   =   9.850651507779759E-4 / Transformation matrix
CD2_1   =   9.847518004452874E-4 / Transformation matrix
CD2_2   =   4.998471812902435E-4 / Transformation matrix
A_ORDER =                  2 / Polynomial order, axis 1
B_ORDER =                  2 / Polynomial order, axis 2
A_2_0   =   6.340797800196409E-7 / distortion coefficient
A_0_2   =   1.749990982371E-07 / distortion coefficient
A_1_1   =   5.382453881118889E-7 / distortion coefficient
B_2_0   =   1.175242752788E-07 / distortion coefficient
B_0_2   =   7.164025680025224E-7 / distortion coefficient
B_1_1   =   5.496396742485739E-7 / distortion coefficient
AP_ORDER=                  2 / Inv polynomial order, axis 1
BP_ORDER=                  2 / Inv polynomial order, axis 2
AP_1_0  =   3.76588935122466E-6 / inv distortion coefficient
AP_0_1  =  -4.851187332073E-06 / inv distortion coefficient
AP_2_0  =  -6.40464415381175E-7 / inv distortion coefficient
AP_0_2  =  -1.767332402001E-07 / inv distortion coefficient
AP_1_1  =  -5.343953631364E-07 / inv distortion coefficient
BP_1_0  =  -2.637543835256E-06 / inv distortion coefficient
BP_0_1  =  -1.547946592422E-06 / inv distortion coefficient
BP_2_0  =  -1.064163345654E-07 / inv distortion coefficient
BP_0_2  =  -7.220223503958E-07 / inv distortion coefficient
BP_1_1  =  -5.404568839297E-07 / inv distortion coefficient
A_DMAX  =    0.3818408549009291 / maximum distortion, axis 1
B_DMAX  =    0.40086247834881306 / maximum distortion, axis 2
CHECKSUM= 'EeGfGZ9dEbEdEZ9d' / HDU checksum updated 2011-08-26T00:05:45Z
END

```

B.5 Ancillary Data Description

The following lists the names and short descriptions of the engineering data that is available in the ancillary engineering data files, listed in binary table extension order.

Mnemonic	Description	Units
ADACTCFGSET	Current Actuator Config Set	dn
ADACTCFGID	Current Actuator Config ID	dn
ADATTERDX	Fine pointing error standard deviation	rad
ADATTERDY	Fine pointing error standard deviation	rad
ADATTERDZ	Fine pointing error standard deviation	rad
ADATTERMX	Fine pointing error mean	rad
ADATTERMY	Fine pointing error mean	rad
ADATTERMZ	Fine pointing error mean	rad
ADATTERMS	The amount of time over which the last mean and deviation da	sec
ADATTERMT	Time Stamp associated with the mean and deviation data.	sec
ADRW1SPD_	wheel 1 speed	rpm
ADRW2SPD_	wheel 2 speed	rpm
ADRW3SPD_	wheel 3 speed	rpm
ADRW4SPD_	wheel 4 speed	rpm
TH1SCMNTT	Spacecraft-Telescope Mount Temp (Sensor 022)	C
TH2SCMNTT	Spacecraft-Telescope Mount Temp (Sensor 022)	C
TH1SPIDT	Spider Temp (Sensor 181)	C
TH2SPIDT	Spider Temp (Sensor 181)	C
TH1TELET	Telescope Structure Temp (Sensor 182)	C
TH2TELET	Telescope Structure Temp (Sensor 182)	C
TH1FPAT	Focal Plane Array Spreader Temp (Sensor 170)	C
TH2FPAT	Focal Plane Array Spreader Temp (Sensor 170)	C
TH1LDET1	LDE Temp 1 (Sensor 050)	C
TH2LDET1	LDE Temp 1 (Sensor 050)	C
TH1LDET2	LDE Temp 2 (Sensor 051)	C
TH2LDET2	LDE Temp 2 (Sensor 051)	C
PW1ESBV	Essential Bus Voltage	V
PW1ESBI	Essential Bus Current	a
PW1SMBI	Safe Mode Bus Current	a
PW1NEBI	Non-Essential Bus Current	a
PW1PLBHZBI	Payload + Hazardous Bus Current	a
PW2ESBV	Essential Bus Voltage	V
PW2ESBI	Essential Bus Current	a
PW2SMBI	Safe Mode Bus Current	a
PW2NEBI	Non-Essential Bus Current	a
PW2PLBHZBI	Payload + Hazardous Bus Current	a
TCSCCFGID	Mimic of commanded Spacecraft Configuration Id.	dn
FDMLDEFFINUM	the commanded ffi num	dn
FDMINTPER	the commanded integration time	dn
FDMLCPER	the LDE commanded long cadence period	dn
FDMSCPER	the LDE commanded short cadence period	dn
PEDSHDT1	SunshadeStructure #1 Temperature	C
PEDSHDT2	Sunshade Structure #2 Temperature	C
PEDLDET3	LDE Temperature #3	C

PEDRADHPT	Radiator Heatpipe Mount Temperature	C
PEDCRRT1	Schmidt Corrector #1 Temperature	C
PEDCRRT2	Schmidt Corrector #2 Temperature	C
PEDCRRT3	Schmidt Corrector #3 Temperature	C
PEDCRRT4	Schmidt Corrector #4 Temperature	C
PEDSPIDT1	Spider #1 Temperature	C
PEDSPIDT2	Spider #2 Temperature	C
PEDRADTELT1	Telescope Radiator Mount Temperature #1	C
PEDUTELT2	Telescope Upper Structure #2 Temperature	C
PEDRADT1	FPA Radiator #1 Temperature	C
PEDRADT2	FPA Radiator #2 Temperature	C
PEDLTELT1	Telescope Lower Structure #1 Temperature	C
PEDLTELT2	Telescope Lower Structure #2 Temperature	C
PEDBLKHT1	Telescope Bulkhead #1 Temperature	C
PEDTELMNTT1	Telescope Spacecraft Mount #1 Temperature	C
PEDPMAT1	Primary Mirror #1 Temperature	C
PEDPMAT2	Primary Mirror #2 Temperature	C
PEDPMAT3	Primary Mirror #3 Temperature	C
PEDPMAT4	Primary Mirror #4 Temperature	C
PEDFOC1T	Focus Mechanism Motor #1 Temperature	C
PEDFOC2T	Focus Mechanism Motor #2 Temperature	C
PEDFOC3T	Focus Mechanism Motor #3 Temperature	C
PEDESBBRDT	Essential Bus Power Supply Temperature	C
PEDSMBBRDT	Safe Mode Bus Power Supply Temperature	C
PEDLPSBRDT2	LDE Power Supply #2 Temperature	C
PEDLPSBRDT1	LDE Power Supply #1 Temperature	C
PEDLDEDCSP5V	LDE_D_CS_+5VDC	V
PEDFPAHI	Fpa Heater Current	a
PEDLDEAPP7_5V	LDE_AP_+7.5VDC	V
PEDLDEANM7_5V	LDE_AN_-7.5VDC	V
PEDLDECPP12V	LDE_CP_+12VDC	V
PEDLDECNM12V	LDE_CN_-12VDC	V
PEDLDEBPP32V	LDE_BP_+32VDC	V
PEDLDEBNM15V	LDE_BN_-15VDC	V
PEDLDEBPP15V	LDE_BP_+15VDC	V
PEDFPAHCT1	Operational Heater Temperature readings most recently read f	C
PEDFPAHCT2	Operational Heater Temperature readings most recently read f	C
PEDFPAHCT3	Operational Heater Temperature readings most recently read f	C
PEDFPAHCT4	Operational Heater Temperature readings most recently read f	C
PEDFPAHCT5	Operational Heater Temperature readings most recently read f	C
PEDFPAHCSETPT	Commanded set point to control FPA temperature.	C
PEDFPAHCERR	The most recent difference value written to the EDHC.	dn
PEDFPAMOD25T	CCD Module 25 Temperature.	C
PEDFPAMOD20T	CCD Module 20 Temperature.	C
PEDFPAMOD15T	CCD Module 15 Temperature.	C

PEDFPAMOD10T	CCD Module 10 Temperature.	C
PEDFPAMOD05T	CCD Module 5 Temperature.	C
PEDDRV1T	Keltic 1 Driver Board Temperature.	C
PEDACQ1T	Keltic 1 Acquisition Board temperature.	C
PEDFPAMOD24T	CCD Module 24 Temperature.	C
PEDFPAMOD19T	CCD Module 19 Temperature.	C
PEDFPAMOD14T	CCD Module 14 Temperature.	C
PEDFPAMOD09T	CCD Module 9 Temperature.	C
PEDFPAMOD04T	CCD Module 4 Temperature.	C
PEDDRV2T	Keltic 2 Driver Board temperature.	C
PEDACQ2T	Keltic 2 Acquisition Board temperature.	C
PEDFPAMOD03T	CCD Module 3 Temperature.	C
PEDFPAMOD08T	CCD Module 8 Temperature.	C
PEDFPAMOD13T	CCD Module 13 Temperature.	C
PEDFPAMOD18T	CCD Module 18 Temperature.	C
PEDFPAMOD23T	CCD Module 23 Temperature.	C
PEDDRV3T	Keltic 3 Driver Board temperature.	C
PEDACQ3T	Keltic 3 Acquisition Board temperature.	C
PEDFPAMOD02T	CCD Module 2 Temperature.	C
PEDFPAMOD07T	CCD Module 7 Temperature.	C
PEDFPAMOD12T	CCD Module 12 Temperature.	C
PEDFPAMOD17T	CCD Module 17 Temperature.	C
PEDFPAMOD22T	CCD Module 22 Temperature.	C
PEDDRV4T	Keltic 4 Driver Board temperature.	C
PEDACQ4T	Keltic 4 Acquisition Board temperature.	C
PEDFPAMOD01T	CCD Module 1 Temperature.	C
PEDFPAMOD06T	CCD Module 6 Temperature.	C
PEDFPAMOD11T	CCD Module 11 Temperature.	C
PEDFPAMOD16T	CCD Module 16 Temperature.	C
PEDFPAMOD21T	CCD Module 21 Temperature.	C
PEDDRV5T	Keltic 5 Driver Board temperature.	C
PEDACQ5T	Keltic 5 Acquisition Board temperature.	C
PEDFOC1POS	Current position of focus mechanism 1.	um
PEDFOC2POS	Current position of focus mechanism 2.	um
PEDFOC3POS	Current position of focus mechanism 3.	um
PEDLDEDP5V	LDE_D_+5VDC	V
ADSTATE	current adcs state	dn
PW1XEPC2SW	PDB1 switch 5, X-Band EPC 2 switch status	dn
PW1KEPC2SW	PDB1 switch 4, Ka-Band EPC 2 switch status	dn
PW1KEPC1SW	PDB1 switch 3, Ka-Band EPC 1 switch status	dn
PW1XEPC1SW	PDB1 switch 2, X-Band EPC 1 switch status	dn
PW2XEPC2SW	PDB2 switch 5, X-Band EPC 2 switch status	dn
PW2KEPC2SW	PDB2 switch 4, Ka-Band EPC 2 switch status	dn
PW2KEPC1SW	PDB2 switch 3, Ka-Band EPC 1 switch status	dn
PW2XEPC1SW	PDB2 switch 2, X-Band EPC 1 switch status	dn
TC1KAEXCST	Word A1: Transponder state/ID - State of Ka-band exciter	dn
TC1XBEXCST	Word A1: Transponder state/ID - State of X-band exciter	dn
TC2KAEXCST	Word A1: Transponder state/ID - State of Ka-band exciter	dn

TC2XBEXCST	Word A1: Transponder state/ID - State of X-band exciter	dn
NN1ANNLH2ST	BCA Sw12 - Anneal2 (NIC BLM#34)	dn
NN1ANNLH1ST	BCA Sw11 - Anneal1 (NIC BLM#33)	dn
NN2ANNLH2ST	BCA Sw12 - Anneal2 (NIC BLM#34)	dn
NN2ANNLH1ST	BCA Sw11 - Anneal1 (NIC BLM#33)	dn
FDMFINEPTLCK	If true then ADCS is in fine point state	dn
FDMMOMDUMP	If true then the thrusters are firing to reduce momentum	dn
PEDFPAHMDST	Current mode for PEDM Thermal.	dn
PEDFPAHCT1ST	Status of reading Operational Heater Temperature read from	dn
PEDFPAHCT2ST	Status of reading Operational Heater Temperature read from	dn
PEDFPAHCT3ST	Status of reading Operational Heater Temperature read from	dn
PEDFPAHCT4ST	Status of reading Operational Heater Temperature read from	dn
PEDFPAHCT5ST	Status of reading Operational Heater Temperature read from	dn
PEDFPAHCTUSED1	Status of Operational Heater Temperature used in FPA temp ca	dn
PEDFPAHCTUSED2	Status of Operational Heater Temperature used in FPA temp ca	dn
PEDFPAHCTUSED3	Status of Operational Heater Temperature used in FPA temp ca	dn
PEDFPAHCTUSED4	Status of Operational Heater Temperature used in FPA temp ca	dn
PEDFPAHCTUSED5	Status of Operational Heater Temperature used in FPA temp ca	dn
PEDFPAMOD25ST	CCD Module 25 Temperature status.	dn
PEDFPAMOD20ST	CCD Module 20 Temperature status.	dn
PEDFPAMOD15ST	CCD Module 15 Temperature status.	dn
PEDFPAMOD10ST	CCD Module 10 Temperature status.	dn
PEDFPAMOD05ST	CCD Module 5 Temperature status.	dn
PEDDRV1ST	Keltic 1 Driver Board temperature status.	dn
PEDACQ1ST	Keltic 1 Acquisition Board temperature status.	dn
PEDFPAMOD24ST	CCD Module 24 Temperature status.	dn
PEDFPAMOD19ST	CCD Module 19 Temperature status.	dn
PEDFPAMOD14ST	CCD Module 14 Temperature status.	dn
PEDFPAMOD09ST	CCD Module 9 Temperature status.	dn
PEDFPAMOD04ST	CCD Module 4 Temperature status.	dn
PEDDRV2ST	Keltic 2 Driver Board temperature status.	dn
PEDACQ2ST	Keltic 2 Acquisition Board temperature status.	dn
PEDFPAMOD03ST	CCD Module 3 Temperature status.	dn
PEDFPAMOD08ST	CCD Module 8 Temperature status.	dn
PEDFPAMOD13ST	CCD Module 13 Temperature status.	dn
PEDFPAMOD18ST	CCD Module 18 Temperature status.	dn
PEDFPAMOD23ST	CCD Module 23 Temperature status.	dn
PEDDRV3ST	Keltic 3 Driver Board temperature status.	dn
PEDACQ3ST	Keltic 3 Acquisition Board temperature status.	dn
PEDFPAMOD02ST	CCD Module 2 Temperature status.	dn

PEDFPAMOD07ST	CCD Module 7 Temperature status.	dn
PEDFPAMOD12ST	CCD Module 12 Temperature status.	dn
PEDFPAMOD17ST	CCD Module 17 Temperature status.	dn
PEDFPAMOD22ST	CCD Module 22 Temperature status.	dn
PEDDRV4ST	Keltic 4 Driver Board temperature status.	dn
PEDACQ4ST	Keltic 4 Acquisition Board temperature status.	dn
PEDFPAMOD01ST	CCD Module 1 Temperature status.	dn
PEDFPAMOD06ST	CCD Module 6 Temperature status.	dn
PEDFPAMOD11ST	CCD Module 11 Temperature status.	dn
PEDFPAMOD16ST	CCD Module 16 Temperature status.	dn
PEDFPAMOD21ST	CCD Module 21 Temperature status.	dn
PEDDRV5ST	Keltic 5 Driver Board temperature status.	dn
PEDACQ5ST	Keltic 5 Acquisition Board temperature status.	dn
NN1KTWT1PWST	KTWT1 Beam on/off status (NIC BLM#15)	dn
NN1XTWT1PWST	XTWT1 Beam on/off status (NIC BLM#13)	dn
NN1KTWT2PWST	KTWT2 Beam on/off status (NIC BLM#31)	dn
NN1XTWT2PWST	XTWT2 Beam on/off status (NIC BLM#29)	dn
NN2KTWT1PWST	KTWT1 Beam on/off status (NIC BLM#15)	dn
NN2XTWT1PWST	XTWT1 Beam on/off status (NIC BLM#13)	dn
NN2KTWT2PWST	KTWT2 Beam on/off status (NIC BLM#31)	dn
NN2XTWT2PWST	XTWT2 Beam on/off status (NIC BLM#29)	dn

Appendix C. Auxillary File Headers

C.1 Cotrending Basis Vector Headers

CBV Primary Header

```
SIMPLE = T / conforms to FITS standard
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1.0 / extension version number
ORIGIN = 'NASA/Ames' / organization that generated this file
TELESCOP= 'Kepler ' / telescope
DATE = '2011-07-22' / file creation date
CREATOR = '/Develop/designFITS/code/pdc-cbv/make_pdcfits.py r43962' / FITS SW
BVVER = ' Prototype Uhat r43545' / basis vector software revision
DATAVER = 'q7_archive_ksop752/lc/mpe_true/pdc-matlab-4158-163239' / input data d
QUARTER = 7 / quarter pertaining to this file
SEASON = 1 / mission season
DATA_REL= 10 / version of data release notes
CHECKSUM= 'frDEho9CfoCCfo9C' / HDU checksum updated 2011-07-22T10:44:30
DATASUM = '0 ' / data unit checksum updated 2011-07-22T10:44:30
END
```

Example Binary Extension

```
SIMPLE = T / conforms to FITS standard
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1.0 / extension version number
ORIGIN = 'NASA/Ames' / organization that generated this file
TELESCOP= 'Kepler ' / telescope
DATE = '2011-07-22' / file creation date
CREATOR = '/Develop/designFITS/code/pdc-cbv/make_pdcfits.py r43962' / FITS SW
BVVER = ' Prototype Uhat r43545' / basis vector software revision
DATAVER = 'q7_archive_ksop752/lc/mpe_true/pdc-matlab-4158-163239' / input data d
QUARTER = 7 / quarter pertaining to this file
SEASON = 1 / mission season
DATA_REL= 10 / version of data release notes
CHECKSUM= 'frDEho9CfoCCfo9C' / HDU checksum updated 2011-07-22T10:44:30
DATASUM = '0 ' / data unit checksum updated 2011-07-22T10:44:30
END
```

C.2 Pixel Response Function Headers

PRF Primary Header

```
SIMPLE = T / conforms to FITS standard
BITPIX = 8 / array data type
NAXIS = 0 / number of array dimensions
EXTEND = T
EXTNAME = 'PRIMARY ' / name of extension
EXTVER = 1.0 / extension version number
ORIGIN = 'NASA/Ames' / organization that generated this file
```

```

DATE      = '2011-09-22'          / file creation date
CREATOR   = 'MAKECALPRF'         / SW version used to create this file
PROCVVER  = 'svn+ssh://murzim/repo/so/trunk/Develop/designFITS/code/prf_r44769'
FILEVER   =                      1.0 / file format version
TELESCOP= 'Kepler  '            / telescope
INSTRUME= 'Kepler photometer'    / detector type
CHANNEL   =                      3 / CCD channel
MODULE    =                      2 / CCD module
OUTPUT    =                      3 / CCD output
CHECKSUM= 'fAjdf3gdf9gdf9gd'    / HDU checksum updated 2011-09-22T01:43:48
DATASUM   = '0'                  / data unit checksum updated 2011-09-22T01:43:48
END

```

Example PRF Image Extension Header

```

XTENSION= 'IMAGE  '            / Image extension
BITPIX   =                    -32 / array data type
NAXIS    =                      2 / number of array dimensions
NAXIS1   =                    750
NAXIS2   =                    750
PCOUNT   =                      0 / number of parameters
GCOUNT   =                      1 / number of groups
INHERIT  =                      T / inherit primary keywords
EXTNAME  = 'PRF_M2_O3_P1'      / extension name
EXTVER   =                    1.0 / extension version number
TELESCOP= 'Kepler  '            / telescope
INSTRUME= 'Kepler photometer'    / detector type
LOCATION   =                      1 / code representing position on CCD channel
WCSNAMEP= 'PHYSICAL'          / name of world coordinate system alternate P
WCSAXESP=                      2 / number of WCS physical axes
CTYPE1P  = 'RAWX  '            / physical WCS axis 1 type CCD col
CUNIT1P  = 'PIXEL '            / physical WCS axis 1 unit
CRPIX1P  =                    375 / pixel image reference column
CRVAL1P  =                   11.99 / value at pixel image reference column
CDELT1P  =                    0.02 / physical WCS axis 1 step
CTYPE2P  = 'RAWY  '            / physical WCS axis 2 type CCD row
CUNIT2P  = 'PIXEL '            / physical WCS axis 2 units
CRPIX2P  =                    375 / pixel image reference row
CRVAL2P  =                   19.99 / value at pixel image reference row
CDELT2P  =                    0.02 / physical WCS axis 2 step
CHECKSUM= '3SS2APS16PS1APS1'    / HDU checksum updated 2011-09-22T01:43:48
DATASUM  = '2365760863'        / data unit checksum updated 2011-09-22T01:43:48
END

```

Appendix D: Acronym and Definition List

D.1 Glossary of Terms

Cadence: A cadence is the frequency with which summed data are read out of the SDA. Short cadence is a 1-minute sum while long cadence is a 30-minute sum.

Channel: The sequential numbering, from 1 through 84, of the mod/outs. See Figure H.1 for the channel numbering.

Column: Each channel has 1132 columns and 1070 rows. There are 1100 science columns enumerated as columns 12 through 1111. Collateral data is enumerated as columns 0 through 11. Columns 1112 through 1131 are virtual columns used to measure electronic bias levels. These are not reported in the CT since an astronomical object located at that position would fall on the adjacent channel. A particular column is, therefore, an integer ranging from -205 through 1111. Negative values are enumerated in order to provide information on the distance of a target from a CCD channel.

Data Availability flag: A flag that indicates if data has been or will be observed by the Kepler photometer. A value of zero indicates that the target has not been observed. A value of one indicates that the target is either planned to be observed or has been observed but the data have not yet been archived. A value of two indicates that data for that target has been archived. Data availability flags are updated quarterly.

Data set: A group of one or more files that are related to each other. For example, a data set may consist of the calibrated cadence data, the corresponding collateral data and the processing history file. The files in a data set are archived as a group.

Data set name: The archive name for a group of one or more files that are archived as a group. Retrieval of a data set will result in retrieval of all files in the group.

Full Frame Image: Called FFI, this is a full readout of every CCD pixel. An FFI is typically taken at the end of each month. The FFI is formatted as a FITS image.

Light Curve: For each target, the SOC provides two calibrated light curves for each quarter. Both are the result of simple aperture photometry. One has been co-trended (“PDCSAP_FLUX”) and the other has not (“SAP_FLUX”). Both light curves are contained within a single FITS formatted file.

Mod/out: Short for module/output, mod/out indicates which CCD recorded the data and which read amplifier was used to read out the detector. The values for modules start at 2 and run through 24, omitting 5 and 21. (Modules 1, 5, 21 and 25 are the fine guidance sensors.) The values of output range from 1 to 4. Although not used in this document, the syntax for mod/out is m.o, where m is the module number and o is the output number. Mod/out ranges from 2.1, 2.2, 2.3, 2.4, 3.1.. 24.4, omitting modules 5 and 21. For a given target, the mod/out will change on a quarterly basis. See Figure D-1 for module and channel numbers.

Motion Polynomial: Two-dimensional polynomials are separately fit to the collection of row and column, flux-weighted, centroids of specified targets as a function of target right ascension and declination. The motion polynomials are computed cadence by cadence. In addition to providing seeds for PSF fitted centroids, motion polynomials are utilized elsewhere in the SOC Pipeline for focal plane geometry fitting, systematic error correction, attitude determination, and computation of instrument performance metrics. See the KDPH for more details.

Output: See Mod/Out

Pixel Mapping Reference Files: Called PMRF files, these files provide the key to identifying which data values in a given cadence data set belong to which targets. These files are produced at the DMC from the quarter-specific target and aperture definitions provided by the SOC. The PMRF files are non-proprietary.

Row: Each channel has 1132 columns and 1070 rows. There are 1024 science rows enumerated as rows 20 through 1043. Collateral data is enumerated as rows 0 through 19 and 1044 through 1069. ROW is an integer ranging from -232 through 1098. Values outside 0 through 1069 are enumerated in order to provide the location of a target in relation to a CCD channel.

Sky Group: As the spacecraft rolls from quarter to quarter (season to season), the stars fall on different CCD channels. The sky group (specified by Skygroup_ID at MAST) is an integer that groups stars together on the sky and, consequently, is time invariant. It is primarily used to specify custom apertures. The Skygroup_ID is equivalent to the channel number an object falls on (or near) during season 2. Stars very close to the boundary separating pixels on different channels can jump from one sky group to another from season to season. Only one sky group value is reported in the KIC. Users requiring sky groups for custom apertures should be especially careful for targets close to the channel boundary. The Skygroup_ID for a target near column 1111 can be misleading. One should consult the on-line pixel calculator for verification.

Target Pixel Data: For each target, the SOC creates a file that contains the individual data values for each pixel for each cadence. These target-based data are archived.

Time Slice: The readout of different modules is staggered in time as described in §5.1 of the KIH. Most modules have a readout time that is a 0.25--3.35 seconds **before** the recorded timestamp for the cadence. The magnitude of this difference, known as the time slice offset, is given by

$$t_{ts} = 0.25 + 0.62(5 - n_{\text{slice}}) \text{ seconds,}$$

where n_{slice} is the module's time slice index. The (module dependent) value for n_{slice} is given in Figure 34 of the KIH. This value is included in BJD times seen by the end user. Because of the quarterly rotation of the spacecraft, a target will lie on a different module each quarter, and may therefore have a different time slice offset from quarter to quarter.

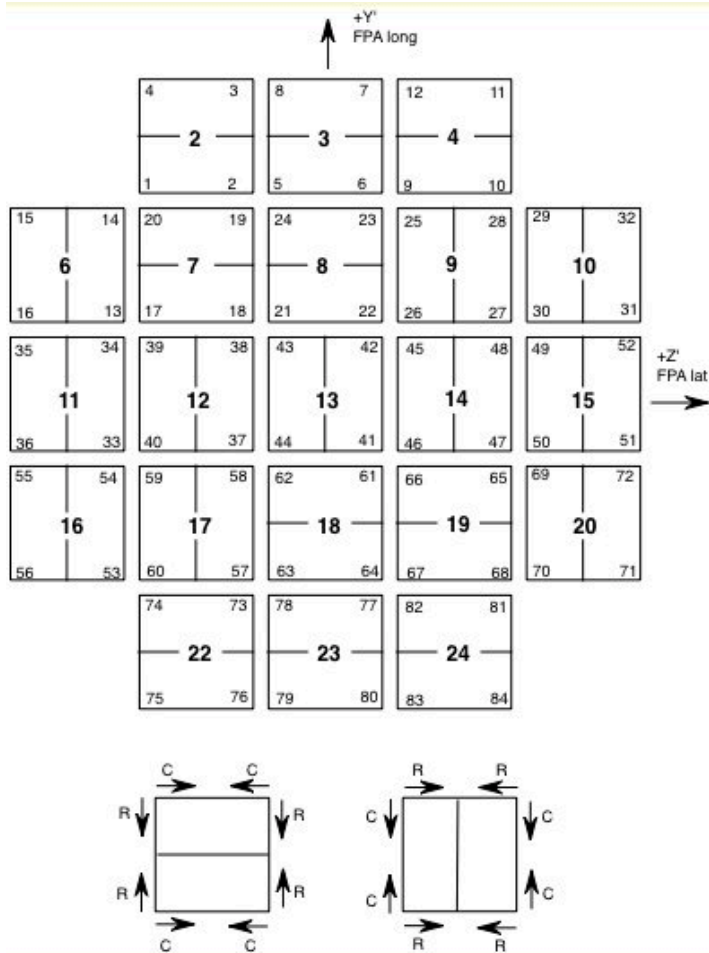


Figure D-1 Modules and channels with column and row directions. Each square shows 2 CCDs. The bold number in the each square is the module number. The smaller numbers in the corners of each square are the channel numbers. In the lower part of the figure, the column (C) and row (R) directions are indicated.

D.2 Common Acronyms

ARC	NASA Ames Research Center
ARP	Artifact Removal Pixels
BKJD	Barycentric Kepler Julian Date -- Equivalent to BJD minus 2454833.0
CAL	Pixel Calibration module
CBV	Co-trending Basis Vector
CCD	Charged Couple Device
CDPP	Combined Differential Photometric Precision
CR	Cosmic Ray
CRCT	Cosmic Ray Correction Table
CT	Characteristics Table
DR	Data Release (usually followed by a data release number)
DIA	Differential Image Analysis
DMC	Data Management Center (for Kepler)
DSN	Deep Space Network (NASA)
DV	Data Validation, a <i>Kepler</i> pipeline module
FC	Focal Plane Characterization
FFI	Full Frame Image
FITS	Flexible Image Transport System
FOV	Field of View
GO	Guest Observer
HDU	Header/Data Units in the FITS format
HLSP	High Level Science Product
HST	Hubble Space Telescope
ICRS	International Coordinate Reference System
KDCH	<i>Kepler</i> Data Characterization Handbook (KSCI-19040)
KDPH	<i>Kepler</i> Data Processing Handbook (KSCI-19081-001)
KeplID	<i>Kepler</i> Identification Number
KIC	<i>Kepler</i> Input Catalog
KIH	<i>Kepler</i> Instrument Handbook (KSCI-19033-001)
KOI	<i>Kepler</i> Object of Interest
KRC	<i>Kepler</i> Results Catalog
KTC	<i>Kepler</i> Target Catalog
LASP	Laboratory for Atmospheric and Space Physics
LC	Long Cadence
MAP	<i>Maximum a Posteriori</i> (msMAP is multi-scale MAP)
MAST	Mikulski Archive at Space Telescope
MJD	Modified Julian Date MJD = JD - 2400000.5 days
MOC	Mission Operations Center
NAIF	Navigation and Ancillary Information Facility (part of NASA's Planetary Sciences Division)

NaN	Not a Number
NED	NASA/IPAC Extragalactic Database
NASA	National Aeronautics and Space Administration
PA	Photometric Analysis module
PDC	Pre-search Data Conditioning Module
PDC-LS	Least Squares Pre-search Data Conditioning Module
PDF	Probability Distribution Function
PI	Principal Investigator
PMRF	Pixel Mapping Reference File
PSF	Point-Spread Function
PRF	Pixel Response Function
SAP	Simple Aperture Photometry
S/C	Spacecraft (i.e., Kepler)
SC	Short Cadence
SCP	Stellar Classification Program
SDA	Science Data Accumulator
SIMBAD	Set of Identifications, Measurements and Bibliography for Astronomical Data.
SOC	Science Operations Center
SPICE	Part of NAIF, SPICE is an information system designed for space instruments
SPSD	Sudden Pixel Sensitivity Dropout
SQL	Standard Query Language
SSDS	Space Science Data Services
STScI	Space Telescope Science Institute
TBD	To Be Determined
TDB	Barycentric Dynamical Time
TDT	Terrestrial Dynamical Time
TT	Terrestrial Time
UCD	Uniform Content Descriptor
VTC	Vehicle Time Code
WCS	World Coordinate System