Kepler Data Release 8 Notes

KSCI-19048-001

Data Analysis Working Group (DAWG)
Pavel Machalek, Editor
Jessie Christiansen, Editor

Data Release 8 for Quarter Q5

Q.m		First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Num CINs	Start CIN	End CIN
5	LC	55275.9912	55370.6600	20-Mar-2010 23:47:20	23-Jun-2010 15:50:24	4633	16373	21006
5.1	SC	55275.9813	55307.5096	20-Mar-2010 23:33:04	21-Apr-2010 12:13:49	46290	479650	525939
5.2	SC	55308.7772	55336.4028	22-Apr-2010 18:39:10	20-May-2010 09:40:02	40560	527800	568359
5.3	SC	55337.0982	55370.6699	21-May-2010 02:21:24	23-Jun-2010 16:04:39	49290	569380	618669

Prepared by: ______________________ Date 10-25-10
Pavel Machalek, Kepler Science Office, for the DAWG (next page)

Prepared by: ______________________ Date 10/25/10
Jessie Christiansen, Kepler Science Office, for the DAWG (next page)

Approved by: ______________________ Date 10/25/10
Jeffrey Van Cleve, Editor-in-Chief

Approved by: ______________________ Date 10/25/10
Jon Jenkins, Co-I for Data Analysis & DAWG Lead

Approved by: ______________________ Date 10/25/10
Michael R. Haas, Science Office Director

These Notes are the collective effort of the Data Analysis Working Group (DAWG), composed of Science Office (SO), Science Operations Center (SOC), Guest Observer Office (GO) and Science Team (ST) members as listed below:

Jon Jenkins*, Chair

Doug Caldwell*, Co-Chair

Allen, Christopher L.
Bryson, Stephen T.
Christiansen, Jessie L.
Clarke, Bruce D.
Cote, Miles T.
Dotson, Jessie L.
Fanelli, Michael N.
Gilliland*, Ron (STSci)
Girouard, Forrest
Haas, Michael R.
Hall, Jennifer
Ibrahim, Khadeejah
Kinemuchi, Karen
Klaus, Todd
Kolodziejczak, Jeff (MSFC)
Li, Jie
Machalek, Pavel
McCauliff, Sean D.
Middour, Christopher K.
Mullally, Fergal
Quintana, Elisa V.
Seader, Shawn
Smith, Jeffrey Claiborne
Still, Martin
Tenenbaum, Peter G.
Thompson, Susan E.
Twicken, Joe
Uddin, Akm Kamal
Van Cleve, Jeffrey
Wohler, Bill
Wu, Hayley Y.

*Science Team

Document Control

Ownership

This document is part of the Kepler Project Documentation that is controlled by the Kepler Project Office, NASA/Ames Research Center, Moffett Field, California.

Control Level

This document will be controlled under KPO @ Ames Configuration Management system. Changes to this document **shall** be controlled.

Physical Location

The physical location of this document will be in the KPO @ Ames Data Center.

Distribution Requests

To be placed on the distribution list for additional revisions of this document, please address your request to the Kepler Science Office:

Michael R. Haas

Kepler Science Office Director

MS 244-30

NASA Ames Research Center

Moffett Field, CA 94035-1000

Michael.R.Haas@nasa.gov

Table of Contents

Prefatory Admonition to Users

The corrected light-curve product generated by Pre-search Data Conditioning (PDC) is designed to enable the Kepler planetary transit search. Although significant effort has been expended to preserve the natural variability of targets in the corrected light curves in order to enable astrophysical exploitation of the Kepler data, it is not possible to perfectly preserve general stellar variability on long timescales with amplitudes comparable to or smaller than the instrumental systematics, and PDC currently is known to remove or distort astrophysical features in a subset of the corrected light curves. In those cases where PDC fails, or where the requirements of an astrophysical investigation are in conflict with those for transit planet search, the investigator should use the uncorrected ('raw') light-curve product instead of the PDC ('corrected') light-curve product, and use the ancillary engineering data and image motion time series provided in the Supplement for systematic error correction. Investigators are strongly encouraged to study the Data Release Notes for any data sets they intend to use. The Science Office advises against publication of these Release 8 light curves without such careful consideration by the end user and dialog with the Science Office or Guest Observer Office as appropriate.

Users are encouraged to notice and document artifacts, either in the raw or processed data, and report them to the Science Office at kepler-scienceoffice@lists.nasa.gov.

Users who neglect this Admonition risk seeing their works crumble into ruin before their time.

1. Introduction

These notes have been prepared to give Kepler users of the Multimission Archive at STScI (MAST) a summary of flight system events that occurred during data collection that may impact quality, and a summary of the performance of the data processing pipeline used on this data set for this release. The Notes for each release of data to the public archive will be placed on MAST along with other Kepler documentation, at http://archive.stsci.edu/kepler/data_release.html.

These Notes are not meant to supplant the following documents, which are also needed for a complete understanding of the Kepler data:

1. **Kepler Instrument Handbook** (KIH, KSCI-19033) provides information about the design, performance, and operational constraints of the Kepler hardware, and an overview of the pixel data sets available. It was released on July 15, 2009, and is publicly available on MAST. Users will need to be familiar with the material in Sections 2 and 4.2-4.5 of the KIH to fully benefit from these Notes.
2. **Kepler Data Analysis Handbook** (KDAH) describes how these pixel data sets are transformed into photometric time series by the Kepler Science Pipeline, the theoretical basis of the algorithms used to reduce data, and a description of residual instrument artifacts after Pipeline processing. Until the KDAH is available, users seeking a discussion of pipeline processing at a deeper level of detail than that provided in these Notes are directed to the SPIE papers (Refs. 3-5), which are available from MAST at http://archive.stsci.edu/kepler/papers/ and from SPIE (http://spie.org/)
3. **Kepler Archive Manual** (KDMC-10008) describes file formats and the availability of data through MAST. The Archive Manual is available on MAST at http://archive.stsci.edu/kepler/manuals/K_archive_manual_v4_083009.htm.
4. **Kepler Mission Special Issue of Astrophysical Journal Letters** (Volume 713, Number 2, 2010 April 20) contained several papers providing background on mission definition (Ref. 11), target selection (Ref. 12), science operations (Ref. 13), the Kepler point spread function (Ref. 14), instrument performance (Ref. 15), and the data processing pipeline (Ref. 9). Two papers discuss the characteristics of the Long Cadence data (Ref. 7), and Short Cadence data (Ref. 8) respectively. Numerous additional papers also provide early science results in both planet detections and asteroseismology, placing the use of Kepler data in context.

Users unfamiliar with the data processing pipeline should read Section 4 first, then the ApJ papers, then the SPIE papers. A list of acronyms and abbreviations appears in Section 9. Questions remaining after a close reading of these Notes and the Instrument Handbook may be addressed to kepler-scienceoffice@lists.nasa.gov .

A sentence at the start of a Section flags those Sections as "recycled" from earlier Notes. If the Notes pertaining to a given data Release are revised, they will be reapproved for release and given an incremented document number KSCI-190XX-00n. n starts at 1 for the original version of the notes for a data Release. Reference to Release Notes will refer to the most recent version (highest n) unless otherwise stated.

Data that would be unwieldy to print in this document format are included in a tar file, the Data Release Notes Supplement, which has been released with this document. Supplement files are called out in the text, and a README file in the tar file also gives a brief description of the files contained. All supplement files are either ASCII or FITS format, though some are also provided as MATLAB *.mat files for the convenience of MATLAB users. The contents of the Supplement are described in Section 10.

Dates, Cadence numbers, and units: Each set of coadded and stored pixels is called a *Cadence,* while the total amount of time over which the data in a Cadence is coadded is the *Cadence period*, which in the case of the flight default operating parameters is 1766 s = 0.49 h, or 270 frame times for Long Cadence, and 58.85 s or 9 frame times for Short Cadence. Cadences are absolutely and uniquely enumerated with *Cadence interval numbers* (CIN), which increment even when no Cadences are being collected, such as during downlinks and safe modes. The *relative cadence index* (RCI) is the Cadence number counted from the beginning of a quarter (LC) or month (SC). RCIs are calculated from the first *valid* Cadence of a Quarter (LC) or Month (SC). For example, the first LC of Q1 would have an RCI = 1 and CIN = 1105 while the last LC of Q1 has RCI = 1639 and CIN = 2743.

Figures, tables, and supplement files will present results in CIN, RCI, or MJD, since MJD is the preferred time base of the Flight System and Pipeline, and can be mapped one-to-one onto CIN or RCI. On the other hand, the preferred time base for scientific results is Barycentric Julian Date (BJD); the correction to BJD is done on a target-by-target basis in the files users download from MAST, as described in detail in Section 7.4. Unless otherwise specified, the MJD of a Cadence refers to the time at the midpoint of the Cadence. Data shown will be for Q5, unless otherwise indicated in the caption. Flux time series units are always the number of detected electrons per Long or Short Cadence.

2. Release Description

A *data set* refers to the data type and observation interval during which the data were collected. The observation interval for Long Cadence data is usually a *quarter*, indicated by Q[n], though Q0 and Q1 are 10 days and one month, respectively, instead of 3 months as will be the case for the rest of the mission. Short Cadence targets can be changed every month, so SC observation intervals are indicated by Q[n]M[m], where m = 1 to 3 is the Month within that Quarter. The *data processing* descriptor is the internal Kepler Science Operations (KSOP) ticket used to track the data processing. The KSOP ticket contains a "Pipeline Instance Report," included in the Supplement, which describes the version of the software used to process the data, and a list of parameter values used. Released software has both a release label, typically of the form m.n, and a revision number (preceded by "r") precisely identifying which revision of the code corresponds to that label. For example, the code used to produce Data Release 8 has the release label "SOC Pipeline 6.2" and the revision number r38897. Unreleased software will, in general, have only a revision number for identification.

The same data set will, in general, be reprocessed as the software improves, and will hence be the subject of multiple releases. The combination of data set and data processing description defines a *data product*, and a set of data products simultaneously delivered to MAST for either public or proprietary (Science Team or GO) access is called a *data release*. The first release of data products for a given set of data is referred to as "new," while subsequent releases are referred to as "reprocessed."

Data products are made available to MAST users as FITS files, described in the Kepler Archive Manual and Section 7 of these Notes. While the Kepler Archive Manual refers to light curves which have not been corrected for systematic errors as 'raw', in these Notes they will be referred to as 'uncorrected' since the uncorrected light curves are formed from calibrated pixels, and 'raw' will refer only to the pixel values for which only decompression has been performed. The relationship of pipeline outputs to MAST files is shown in Figure 2. The keyword DATA_REL = 8 is in the FITS headers so users can unambiguously associate Release 8 FITS files with these Notes. In Release 8, all light curve files have FITS keyword QUARTER = 5.

Data Release 8 was produced with released code, with formal verification and validation of the pipeline and the resulting data products. While the Kepler data analysis pipeline continues to evolve to adapt to the performance of the flight system and our understanding of the data, the rate of evolution has decreased such that major upgrades can now be expected on a roughly annual basis.

2.1 Summary of Contents

Table 1: Contents of Release 8. CIN is the Cadence interval number described in Section 1. All Release 8 cadence data were processed under KSOP-568 with SOC Pipeline 6.2, revision number r38897, and are released for the first time. The Pipeline Instance ID (PID) for CAL, PA, and PDC is shown. All 4 channels of module 3 permanently failed at MJD 55205.745 during Q4 (Section 5.4) and no data are available after that date.

Q.m		CAL PID	PA PID	PDC PID	First Cadence MJD midTime	Last Cadence MJD midTime	Num CINs	Start CIN	End CIN
5	LC	2757	2817	2817	55275.9912	55370.6600	4633	16373	21006
5.1	SC	2837	2837	2837	55275.9813	55307.5096	46290	479650	525939
5.2	SC	2857	2857	2857	55308.7772	55336.4028	40560	527800	568359
5.3	SC	2858	2858	2858	55337.0982	55370.6699	49290	569380	618669

For Release 8, all science mission FFIs have been uniformly reprocessed using Pipeline 6.2. With the exception of cosmic ray cleaning, which is not available for FFIs, the same CAL processing (Section 4) has been applied to these images as to the Release 8 cadence files. FFIs are available from the MAST Kepler FFI search page: http://archive.stsci.edu/kepler/ffi/search.php.

2.2 Pipeline Changes Since Previous Release

This Section describes changes in the Pipeline code and parameters since the previous release of newly released data (the Q4 data in Data Release Notes 6). Changes are listed by Pipeline module outputs, and the corresponding data products on MAST. The software modules comprising the science data analysis Pipeline are described briefly in Section 4. Users unfamiliar with the Pipeline should read Section 4 before reading this Section. The PA and PDC versions and input parameters can be unambiguously referenced by their Pipeline Instance Identifier (PID), shown in Table 1.

2.2.1 CAL: calibrated pixels

There were significant changes to CAL were made between Release 6 and Release 8.

2.2.2 PA: uncorrected light curves and centroids

There were no significant changes to PA between Release 6 and Release 8.

2.2.3 PDC: corrected light curves

There were no significant changes to PDC between Release 6 and Release 8.

3. Current Evaluation of Performance

3.1 Overall

The Combined Differential Photometric Precision (CDPP) of a photometric time series is the effective white noise standard deviation over a specified time interval, typically the duration of a transit or other phenomenon that is searched for in the time series. In the case of a transit, CDPP can be used to calculate the S/N of a transit of specified duration and depth. For example, a 6.5 hr CDPP of 20 ppm for a star with a planet exhibiting 84 ppm transits lasting 6.5 hours leads to a single transit S/N of 4.2 σ.

The CDPP performance has been discussed by Borucki et al. (Ref. 2) and Jenkins et al. (Ref. 7). Jenkins et al. examine the 33.5-day long Quarter 1 (Q1) observations that ended 2009 June 15, and find that the lower envelope of the photometric precision on transit timescales is consistent with expected random noise sources. The Q5 data discussed in these Notes have the same properties, as shown in Figure **1**. Nonetheless, the following cautions apply for interpreting data at this point in our understanding of the Instrument's performance:

1. Many stars remain unclassified until Kepler and other data can be used to ascertain whether they are giants or otherwise peculiar. Since giant stars are intrinsically variable at the level of Kepler's precision, they must be excluded from calculations of CDPP performance. A simple, but not foolproof, way to do this is to include only stars with high surface gravity (log g > 4).
2. Given the instrument artifacts discussed in detail in the KIH and Ref. 15, it is not generally possible to extrapolate noise as 1/sqrt(time) for those channels afflicted by artifacts which are presently not corrected or flagged by the Pipeline.
3. Stellar variability and many instrumental effects are not, in general, white noise processes.
4. There is evidence from the noise statistics of Q0 and Q1 (see the Release 5 Notes) that the Pipeline is overfitting the data for shorter data sets (a month or less of LC data) and fainter stars, so users are urged to compare uncorrected and corrected light curves for evidence of signal distortion or attenuation. The problem is less evident in the Q2-Q5 data sets than in the Q0 and Q1 data of Release 5.

Example published data is shown in [2] and [10].

Further information may be gleaned from examining the $TMCDPP_k$ of subsets of the full target list, such as all targets with magnitude between 11.75 and 12.25 and log g > 4, loosely referred to as "12th magnitude dwarfs". Table 2 summarizes the median and percentile results for various target subsets in Q2. Note that the median CDPP over K = {all stars in a given magnitude bin} actually decreases as stars get fainter beyond 10th magnitude, since the proportion of all stars which are (quiet) dwarfs increases considerably as the stars get fainter.

The Jenkins et al. (Ref. 7) expression for the lower noise envelope was fitted to propagated uncertainties accounting for all known and quantified random errors such as shot noise from all sources in the aperture (including background sky flux), read noise, quantization noise, and processing noise from pixel-level calibrations (such as the 1-D black and smear corrections). The lower envelope of expected errors appear to be dominated by shot noise out to Kepler magnitude =14 or so. Extending this expression to the benchmark 6.5 hr transit time gives the results shown in Figure **1** and Table 2.

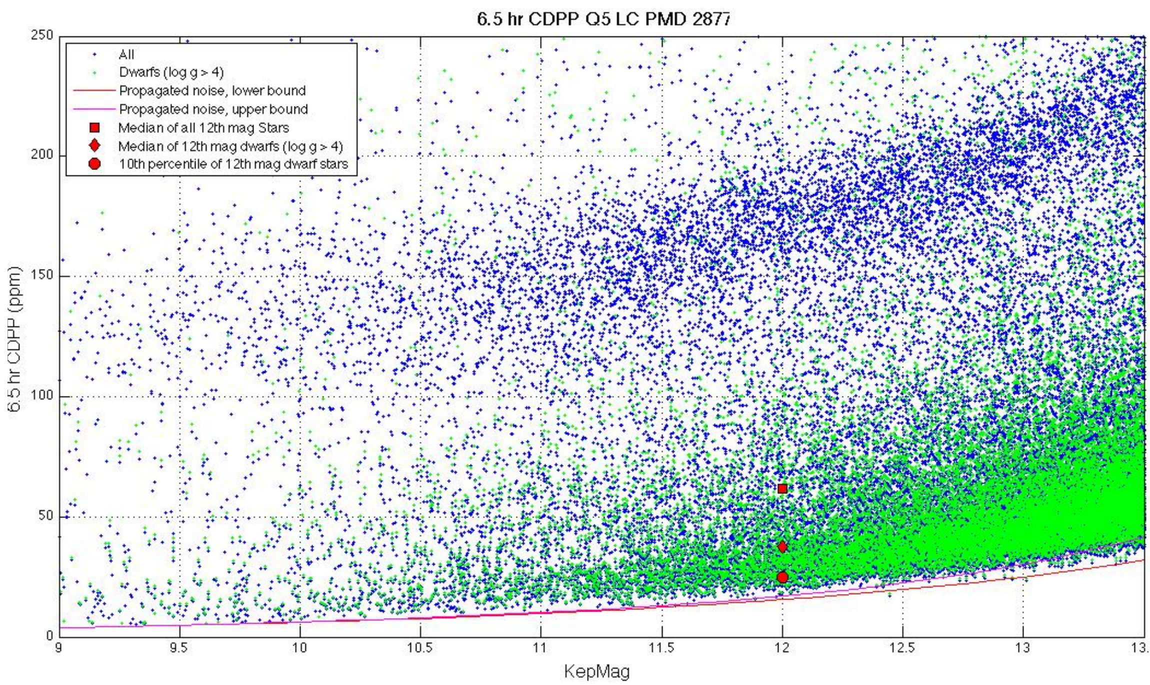

Figure 1**: 6.5 hr Temporal Median (TM) of the Quarter 5 CDPP time series calculated by the TPS Pipeline module for stars between 9th and 13.5th magnitude. The 6 hr TMCDPPs have been divided by sqrt(13/12) = 1.041 to approximate 6.5 hr TMCDPPs. Stars on the planetary target list with Kepler Magnitude < 13.5 and log g > 4, which are almost certainly dwarf stars, are shown as green +'s; other stars are marked with blue +'s. The red and magenta lines are the lower and upper envelopes, respectively, of CDPP.**

Table 2: Aggregate Statistics for the TMCDPPs plotted in Figure 1. Column Definitions: (1) Kepler Magnitude at center of bin. Bins are +/- 0.25 mag, for a bin of width 0.5 mag centered on this value. (2) Number of dwarfs (log g > 4) in bin. (3) 10th percentile TMCDPP for dwarfs in bin. (4) Median TMCDPP for dwarfs in bin. (5) Number of all stars in bin. (6) 10th percentile TMCDPP of all observed stars in bin. (7) Median TMCDPP for all stars in bin. (8) Lower envelope of simplified noise model CDPP, which does not include astrophysical noise. TMCDPP is in units of ppm.

center mag	number of dwarfs in bin	10th prc CDPP, dwarfs	median CDPP, dwarfs	Number of all stars in bin	10th prc CDPP, all stars	median CDPP, all stars	lower envelope of model CDPP
9	27	6.5	34.1	186	13.2	87.1	3.8
10	161	11.7	37.5	591	13.4	107.9	6
11	634	17.8	35.4	1775	20.7	96.8	9.5
12	2228	24.8	38.2	4310	26.7	63.1	15.2
13	7009	36.7	50.4	10114	37.9	58.2	24.4

3.2 Changes in Performance Since Previous Release

Performance is essentially unchanged since Release 6 (Q4). In Release 8 (Q5), the median 12th magnitude dwarf has a CDPP of 38 ppm, while in Release 6 (Q4) this benchmark was 40 ppm.

3.3 Known Calibration Issues

Topics under consideration by the DAWG which may change future calibration parameters or methods include:

1. Find a set of ancillary engineering data (AED) and Pipeline-generated metrics which more effectively remove systematic errors without overfitting the data (and hence distorting the astrophysical signal). Correlations between the corrected light curves of different targets suggest the existence of unrepresented systematic errors.
2. Improve the characterization of stellar variability to represent weaker and more complex waveforms, so cotrending can be more effective when the stellar variability is temporarily removed from the light curve.
3. Characterize and correct for in-orbit change of focus (Section 6.4).
4. Identify particular light curves that are poorly corrected, and understand why generally effective remedies do not work in these cases. Feedback from users to kepler-scienceoffice@lists.nasa.gov is essential for the SO and SOC to identify, flag, and fix all such "hard cases."
5. Mitigate or at least identify the Artifacts described in the KIH, Section 6.7.

6. Assess and improve the focal plane characterization models which are inputs to CAL.
7. Improve definition of photometric apertures, especially for saturated stars where failure to include pixels at the edge of charge bleeding columns can be particularly problematic (Section 5.5).

Calibration and data analysis issues related to the focal plane and its electronics are discussed in the Instrument Handbook.

4. Data Delivered – Processing History

4.1 Overview

This Section is unchanged since the Release 6 Notes.

The delivered FITS files were processed as shown in simplified form in Figure 2. What is referred to as "raw" flux time series in the Kepler Archive Manual is the result of calibrating pixels, estimating and removing sky background, and extracting a time series from a photometric aperture, and is referred to in these notes as "uncorrected" flux time series. The "corrected" flux time series has been decorrelated against known system state variables, such as pointing. In these Notes, we refer to "detrending" as an operation that removes low-frequency features of a light curve, using only the light curve data itself – such as subtracting the results of a median boxcar or centered polynomial (Savitzky-Golay) fit from the data. "Cotrending," on the other hand, removes features correlated between the light curve and ancillary data, with some loss of low-frequency information and consequent signal distortion. Cotrending is also referred to as "systematic error removal."

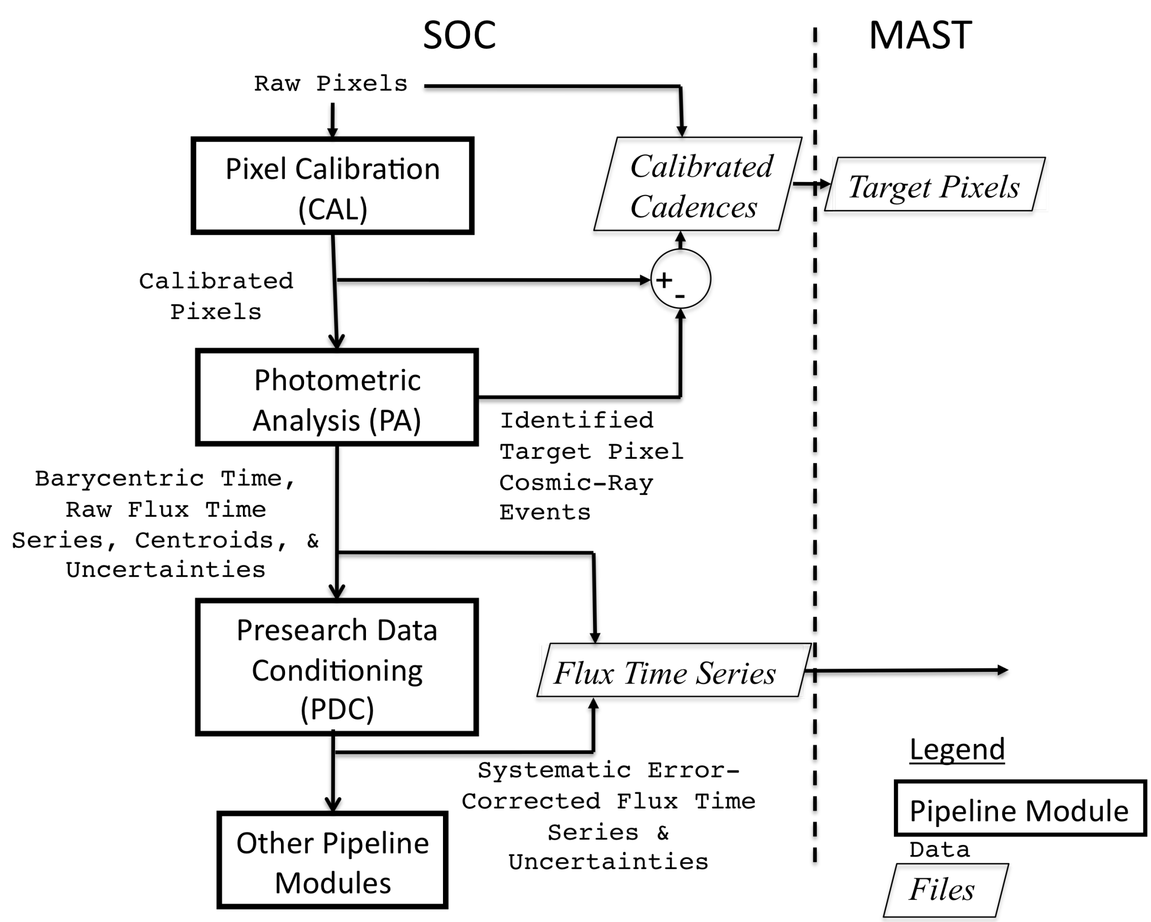

Figure 2: Processing of data from raw pixels to flux time series and target pixel files archived at MAST. The target pixel files generated at MAST from the calibrated Cadence

files delivered by the SOC have identified cosmic-ray events removed. The corrected flux time series delivered to MAST contain stellar variability and have -Infs for bad or missing data. See Section 7 and the MAST Kepler Archive Manual for details of MAST file contents.

4.2 Pixel-Level Calibration (CAL)

This section is unchanged since the Release 6 Notes.

The first step, pixel calibration (software module CAL), performs the pixel level calibrations shown in Figure 3. The SOC receives raw pixel data from each Kepler CCD, including collateral pixel data that is collected primarily for calibration. These collateral pixels include serial register elements used to estimate the black level (voltage bias), and masked and over-clocked rows used to measure the dark current and estimate the smear that results from the lack of a shutter on the spacecraft. Detailed models of each CCD have been developed from pre-flight hardware tests, along with full-frame images (FFIs) taken during commissioning prior to the dust cover ejection. These models are applied within CAL to correct for 2D bias structure, gain and nonlinearity of the ADU-to-photoelectron conversion, the electronic undershoot discussed in KIH Section 6.6, and flat field. CAL operates on long (30 min) and short (1 min) Cadence data, as well as FFIs [3, 9].

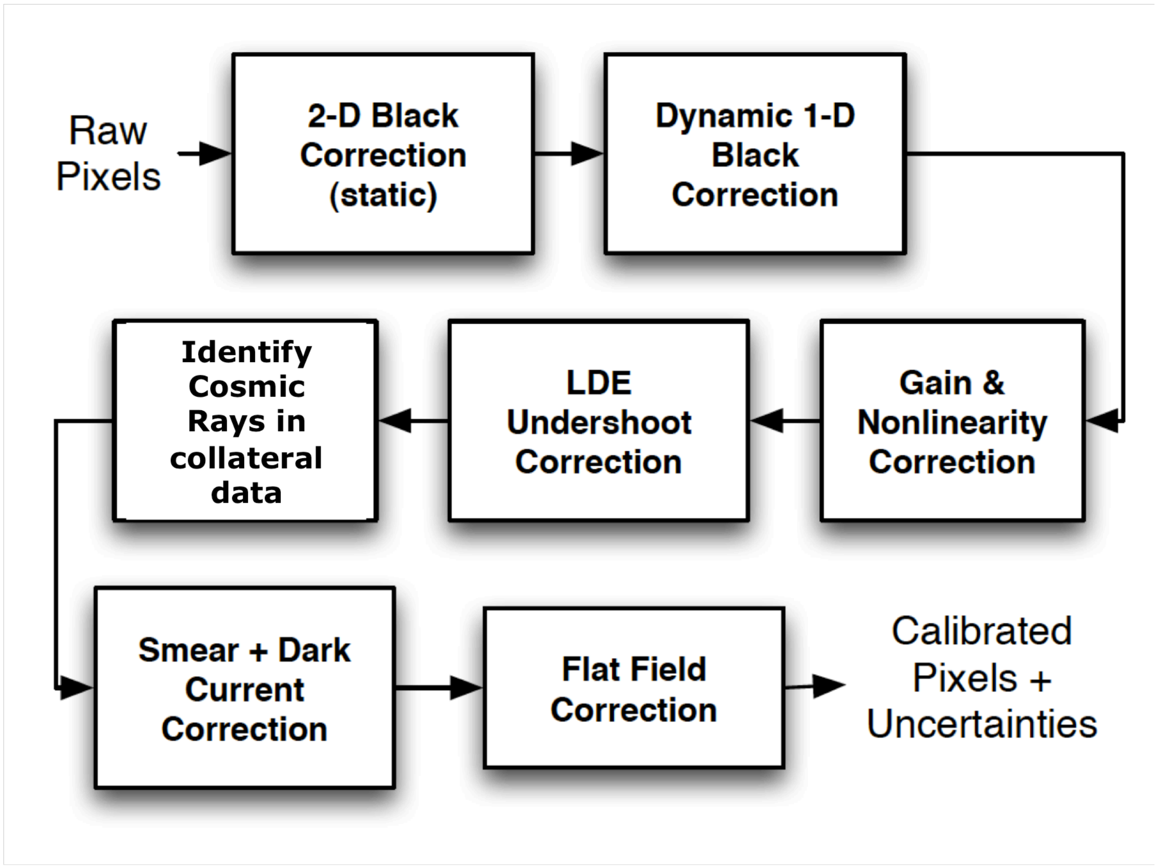

Figure 3: Pixel Level Calibrations Performed in CAL. See the Instrument Handbook for a discussion of signal features and image contents processed in CAL.

There may be rare cases where the bleeding smear value exceeds the 23 bits of the Science Data Accumulator. Users noticing repeated large step-like transitions between two discrete flux levels should consult the FFIs to see whether either the masked or virtual smear regions contain bleeding charge. If so, please contact the Science Office for further investigation.

4.3 Photometric Analysis (PA)

This Section is essentially unchanged from the Release 6 Notes, except to update Figure 4.

The primary tasks of this module are to compute the photometric flux and photocenters (centroids) for up to 170,000 Long Cadence (thirty minute) and 512 Short Cadence (one minute) targets across the focal plane array from the calibrated pixels in each target's aperture, and to compute barycentric corrected timestamps per target and cadence [4].

The tasks performed by Photometric Analysis (PA) are

1. Calculation of barycentric time correction, obviating the need for manual correction discussed in the Release 2 Notes (KSCI-19042).

2. Detection of Argabrightening events (Section 6.1). Argabrightening detection and associated Cadence gapping take place before CR detection; otherwise, many of the Argabrightening events would be cleaned as cosmic rays in the respective pixels, and not effectively detected and marked as data gaps.
3. Cosmic ray (CR) cleaning of background and target pixels, logging of detected CRs, and calculation of CR metrics such as hit rate and mean energy. The CR's are corrected by subtracting the residual differences after median filtering is performed on the detrended pixel time series. The same method and parameters are used for LC and SC. (Note 5)
4. Robust 2-D polynomial fitting to calibrated background pixels.
5. Background removal from calibrated target pixels.
6. Aperture photometry. In this Release, the flux is the sum of pixels in the optimal aperture after background removal (Simple Aperture Photometry, SAP).
7. Computation of flux-weighted (first moment) centroids (Note 1).
8. Fitting of 2-D motion polynomials to target row and column centroids, which smoothly maps (RA,DEC) to (row, column) for a given output channel (Note 4).
9. Setting gap indicators for Cadences with Argabrightening (Section 6.1). The gapped Cadences have all –Inf values in the FITS light curve files, except for the first three columns (see Section 7.2).

<u>*Notes*</u>

1. Flux-weighted (first moment) centroids are calculated for all targets. PRF centroids are also computed, but only for a small subset (PPA_STELLAR) of the Long Cadence targets due to the heavy computational requirements of the PRF centroiding algorithm. The PRF fitting does not necessarily converge for targets which are faint, or located in complex fields. Only flux-weighted centroids are exported to MAST in this Release; they are suitable for precision astrometry [6] in uncrowded apertures. Users wishing to improve on the flux-weighted centroids need to consider the distribution of flux from non-target sources in the optimal aperture pixels or use the PRFs provided in the KIH Supplement to do their own fits.

2. There is no identification of bad pixels in PA in this Release, nor is there any exclusion, gapping or other treatment of known bad pixels. Bad pixels may be identified in future releases. The treatment of bad pixels is TBD, and may depend on how the pixel is bad (high read noise, unstable photoresponse, low photoresponse, etc.) and its location in the target aperture. While the Pipeline flags bad data on a per mod.out, per cadence basis, bad pixels affect individual targets, and users are cautioned to carefully inspect the target pixels before believing peculiar light curves.

3. The output of PA is called ‘raw’ in the light curve FITS file, even though it is the sum of ‘calibrated’ pixels, because systematic errors have not been removed by PDC.

4. Motion polynomials are a means of estimating local image motion, and do not assume rigid body motion of the entire focal plane. They thus account for changes in plate scale, rotation, image distortion, and differential velocity aberration (DVA) on a channel-by-channel and Cadence-by-Cadence basis (Figure 4). There is no requirement for smoothness in time of motion polynomials for cotrending and other purposes, and there is no fitting or smoothing across time (see Section 10.2.1 for further discussion). The simplified mod.out center motion time series provided in the Supplement are the row and column of the nominal center (in RA and DEC) of the mod.out as calculated from these motion polynomials.

5. Data which are greater than 12 median absolute deviations (MAD), after the removal of a trend formed by a quadratic fit followed by a five Cadence wide median filter, are identified as CRs. The MAD is calculated over a sliding window 145 cadences wide after the trend is removed. The amplitude, cadence, and location of the removed CR will be made available to users in future Releases, either as cadence-to-cadence cosmic ray correction tables, or integrated into the target pixel files (Section 7.5), so that users may restore the CRs and use their own methods of CR detection and removal if desired.

Astrophysical phenomena of only a single cadence duration cannot be distinguished from CRs, as the Pipeline does not check for correlated outliers on adjacent pixels.

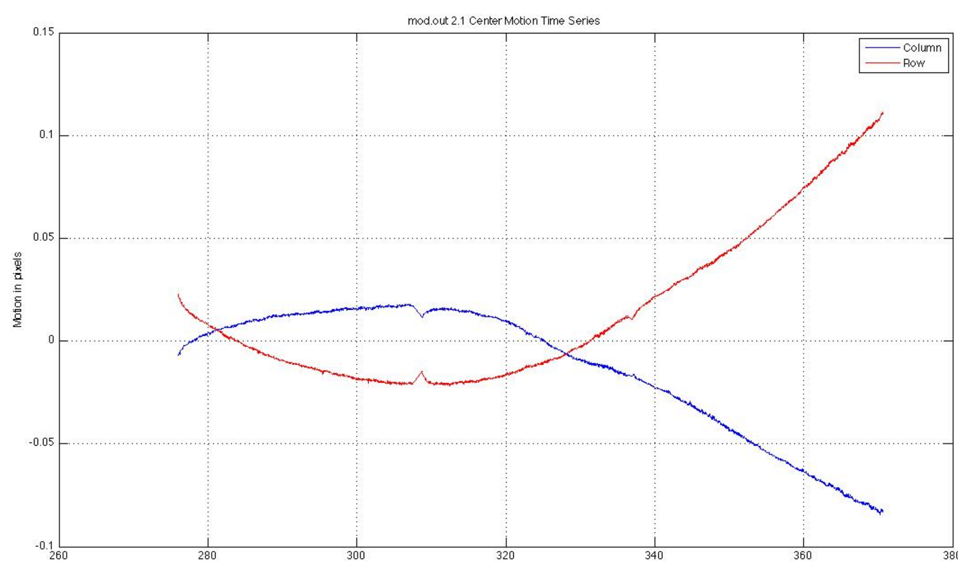

Figure 4: Mod.out 2.1 Center Motion Time Series calculated from motion polynomials for Q5. The median row and column values have been subtracted. Since this mod.out is at the edge of the field, it shows large differential velocity aberration (DVA) with respect to the center of the field, as well as a higher sensitivity to focus jitter and drift. The jumps at MJD ~ 55308 & 55337 are due to temperature changes during Earth downlinks. Except for these events, these curves are smoother than Q2 and earlier.

4.4 Pre-Search Data Conditioning (PDC)

This Section is unchanged from the Release 6 Notes.

The primary tasks of PDC for MAST users are to correct systematic errors and remove excess flux in target apertures due to crowding. PDC was designed to remove systematic errors that are correlated with ancillary engineering or Pipeline generated metrics (such as motion polynomials). Significant effort has been expended to preserve the natural variability of targets, though further effort is still required to strike the right balance between preserving stellar variability signals and removing signatures correlated with instrumental effects. Users will therefore need to be cautious when their phenomena of interest are much shorter (<1 h) or much longer (>5 d) than a transit, or have complex light curves with multiple extrema on transit time scales (such as eclipsing and contact binaries). Examples of astrophysical features removed or significantly distorted by PDC are shown in Section 4.4.3.

Tuning the parameters of PDC requires assessing the relative merits of removing instrumental artifacts, preserving transits and their shapes, and preserving other astrophysical phenomena, and it is not likely that any single choice can give satisfactory results for all observing conditions, targets, and phenomena of interest. Hence, PDC is discussed in greater detail in these Notes than is CAL or PA. Users concerned about the impact of PDC on their signals of interest are invited to use the ancillary engineering data (AED) and motion time series in the Supplement to perform their own systematic error correction.

4.4.1 Description

This Section is unchanged from the Release 6 Notes.

The tasks performed by PDC are:

1. Accept data anomaly flags for Cadences that are known to be lost or degraded (Section 4.4.2). These Cadences and their corresponding data anomalies are shown in Section 5.3.6
2. Resampling of AED to match the sampling rate of LC and SC data.
3. Identification and correction of unexplained discontinuities (i.e. unrelated to known anomalies), an iterative process.
4. Cotrend target flux time series against AED and motion polynomials derived by PA (Section 4.3 item 8) to remove correlated linear and nonlinear deterministic trends. Singular Value Decomposition (SVD) is used to orthogonalize the set of basis vectors and numerically stabilize the model fit.
5. Identify variable stars (>0.5% center-peak variability).
6. For variable stars only, perform coarse systematic error correction with the following steps:
    a. Correct discontinuities due to attitude tweaks.

    b. Compare phase-shifting harmonic fitting to simple polynomial fitting, and select the method that gives the smallest error for initial detrending.
    c. Correct thermal recovery transients with a polynomial fit for each target.
    d. Remove a low-order polynomial trend from the transient-corrected light curve
    e. Repeat the harmonic fit first done in Step 6.b, and save this improved harmonic fit for later restoration.
    f. Subtract the improved harmonic fit from the light curve resulting from Step 3, and cotrend the harmonic-removed light curve as in Step 4.
7. For stars initially identified as variable, if the standard cotrended result is not variable and cotrending has reduced the noise then the identification of the star as variable is considered mistaken. Then the result of the standard cotrending is retained. Otherwise, the result of the harmonic free cotrending is retained. The harmonic content is restored later in PDC.
8. Assess results of cotrending. If cotrending has increased the noise by >5%, restore the uncorrected light curve at this point.
9. Correction for the excess flux in the optimal aperture for each target due to crowding, as calculated over the optimal aperture.
10. Identification and removal of impulsive outliers after masking off astrophysical events such as giant transits, flares, and microlensing. A median filter is applied to the time series after the removal of obvious astrophysics, and the residual is determined by subtracting the median series from the target flux series. A robust running mean and standard deviation of the residual is calculated and points more than 12 σ (LC) or 8 σ (SC) from the mean are excluded. Not all astrophysical events are successfully masked, and hence may be falsely identified as outliers or may unnecessarily increase the noise threshold for outliers. The masked events are restored to the MAST corrected light curves.

<u>*Notes*</u>

1. The crowding metric is the fraction of starlight in an aperture that comes from the target star. For example, a crowding metric of 1 means that all the light in an aperture comes from the target, so the light curve needs no correction. A crowding metric of 0.5 means that half the light is from the target and half from other sources, so the flux must be decreased by half to obtain the correct light curve for the target. Note that the uncorrected flux time series are *not* corrected for crowding. The crowding metric is based on the Kepler Input Catalog (KIC) star locations and brightnesses, the local PRF of the target star and its neighbors, and the optimal aperture. It is averaged over a quarter, and neglects seasonal and secular changes in the PRF compared to the model established by observations during Commissioning. A given star will move to different parts of

the Kepler focal plane from quarter to quarter as Kepler rolls, so the PRF, aperture, and crowding metric will also vary from quarter to quarter.

2. Gaps are not filled in the MAST files, and are represented as -Infs. Intermediate data products generated by PDC and internal to the SOC do have gaps filled, before being passed to planetary search parts of the Pipeline.

3. Different frequencies of AED will physically couple to the photometric light curve with different strengths. PDC represents this by decomposing AED time series into low and high bandpass components, and allowing the coefficients of the components to vary independently in the fit.

4. The output of PDC is referred to as 'corrected' data in the delivered files. Users are cautioned that systematic errors remain, and their removal is the subject of ongoing effort as described in Section 3.3.

4.4.2 Performance

This Section is conceptually unchanged from the Release 6 and 7 Notes. It has been modified to show examples from Q5. Cases where the examples are not from Q5 are labeled as such in the captions.

PDC gives satisfactory results on most stars which are either intrinsically quiet (Figure 5), or have well-defined harmonic light curves above the detection threshold (Figure 6); in most of these cases, the standard deviation of the corrected flux is within a factor of 2 of the noise expected from read and shot noise in the calibrated pixels summed to form the uncorrected light curve. It also performs well in many cases where the star is variable, but without a dominant harmonic term (Figure 7). However, PDC will sometimes not identify a target-specific discontinuity (Figure 8), and will sometimes introduce noise into complex lightcurves (Figure 9). Conversely, PDC sometimes identifies eclipses as discontinuities and introduces a discontinuity in an attempt to correct the false discontinuity (Figure 10).

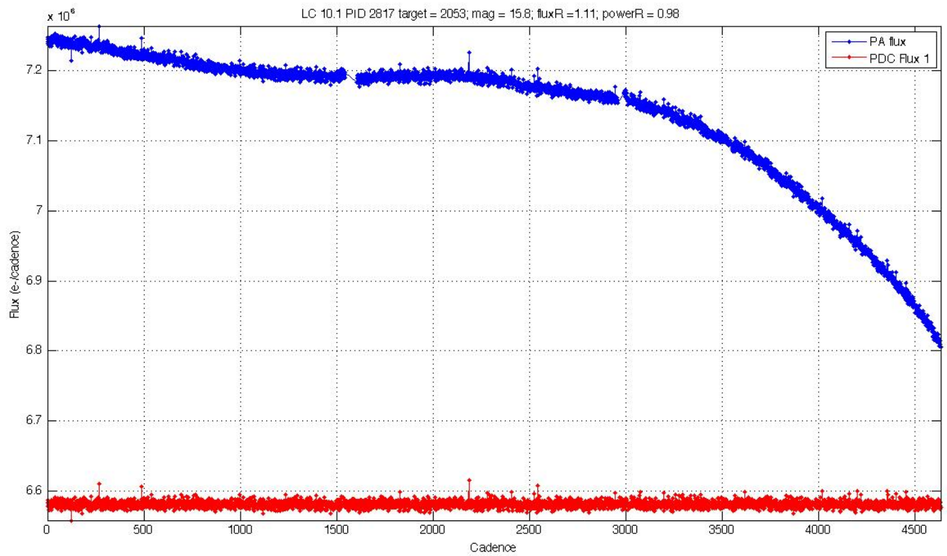

Figure 5: Q5 example of PDC removal of trends and discontinuities from the light curve of a quiet star of Kepler magnitude 15.8. The noise in the corrected light curve is only 11% greater than the noise expected from the calibrated pixels, a considerable improvement over the uncorrected light curve.

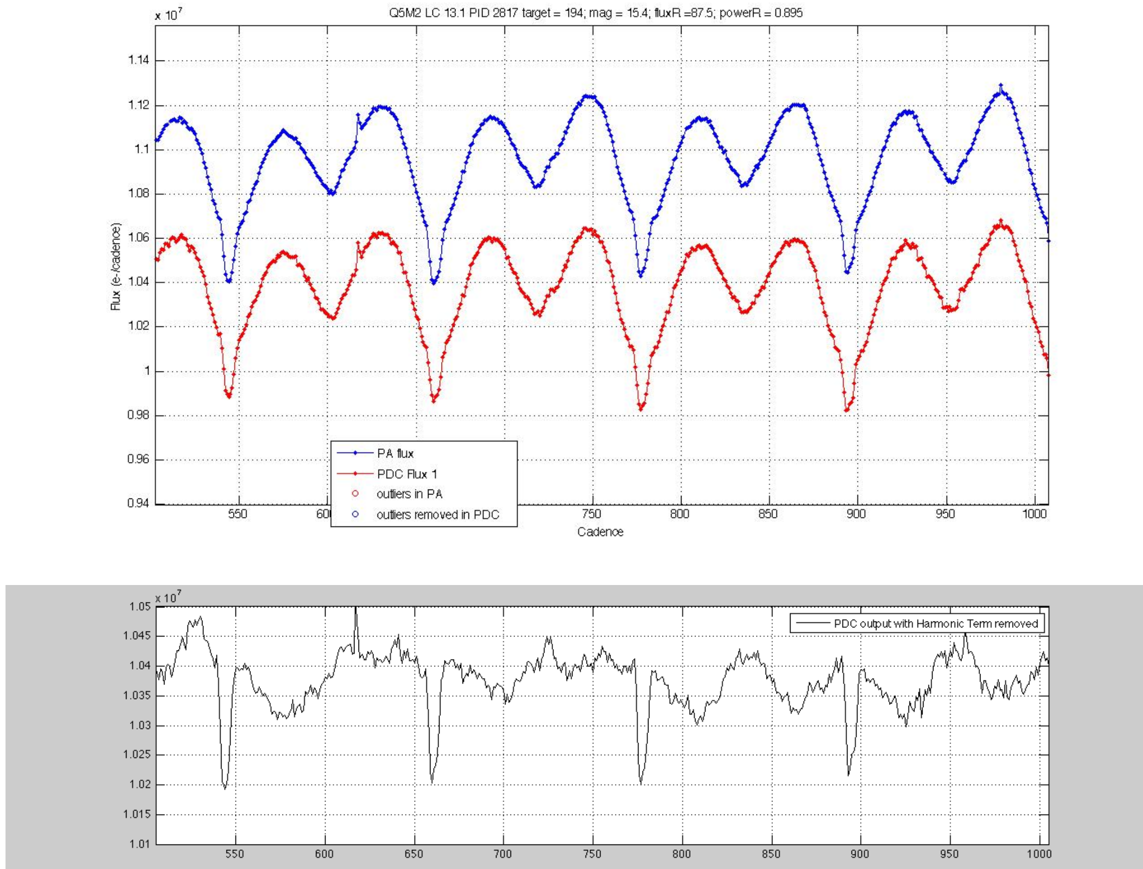

Figure 6: Q5 example of PDC correction of a harmonically variable star with eclipses. MAST users receive the light curve corrected for systematic errors, with the harmonic variability restored and gaps in the data represented by –Infs. The corrected light curve delivered to MAST is shown in red in the upper panel of this Figure. The lower panel shows the light curve with harmonics removed illuminating the efficacy of PDC harmonic removal, even with light curves with other features such as eclipses.

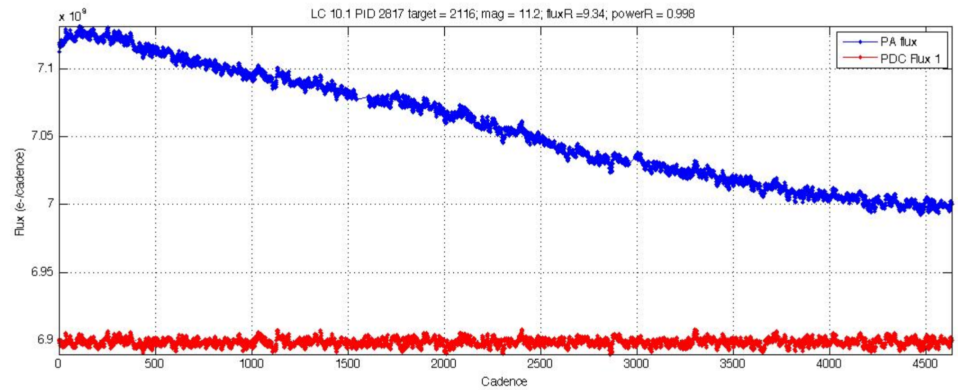

Figure 7: Q5 example of PDC correction of a non-harmonically variable star. The RMS of the corrected (red) curve is about 9x the noise calculated from the read and shot noise in the calibrated pixels, and is believed to be almost entirely due to intrinsic stellar variability. While this figure illustrates PDC's ability to fill data gaps the filled data is not delivered to MAST and users will have –Inf for the flux of those cadences.

Figure 8: Q5 example of an unidentified and hence uncorrected target-specific discontinuity at cadences 800 and 2500.

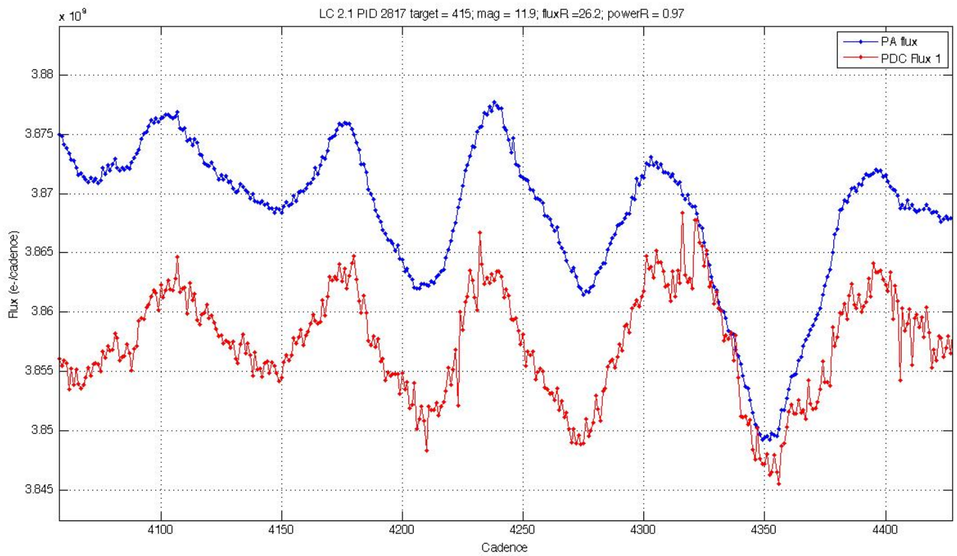

Figure 9: Q5 example of PDC adding short-period noise to a light curve, which it did not identify as a bad cotrend.

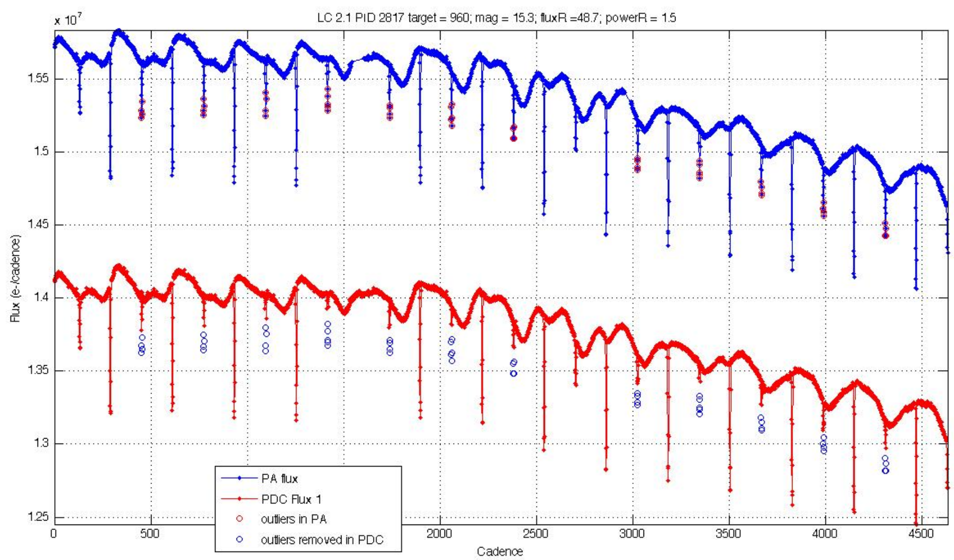

Figure 10: Q5 example of PDC misidentifying an eclipse as an outlier.

4.4.3 Removal of Astrophysical Signatures

This Section is conceptually unchanged from the Release 6 Notes. It has been modified to show examples from Q5. Cases where the examples are not from Q5 are labeled as such in the captions.

PDC can remove astrophysical signatures if they are:

1. Harmonic, but have periods > 5d and fall below PDC's detection threshold for stellar variability. In Release 8 , the center-peak threshold is 0.5%, which for otherwise quiet stars allows a harmonic with a peak-to-peak amplitude of 1.0% to go undetected.
2. Spikes a few Cadences wide (Figure 12), such as flares, or other steep gradients in flux.
3. More or less linear ramps over the processing interval.
4. Harmonic signals above the threshold, but the harmonic fit does not produce a good fit, and current algorithm fails to recognize that cotrending has performed badly.
5. Non-harmonic signals for which current algorithm fails to recognize that cotrending has performed badly.
6. Harmonic signals above the threshold for which the fit is good, but PDC incorrectly determines that target was cotrended well when treated as non-variable (Figure 11).

A thorough study of astrophysical signal distortion by PDC is underway, but has not been completed to date. Users may be helpful to this effort by reporting light curves, in which they suspect that a signal has been distorted or removed, to the Science Office at kepler-scienceoffice@lists.nasa.gov.

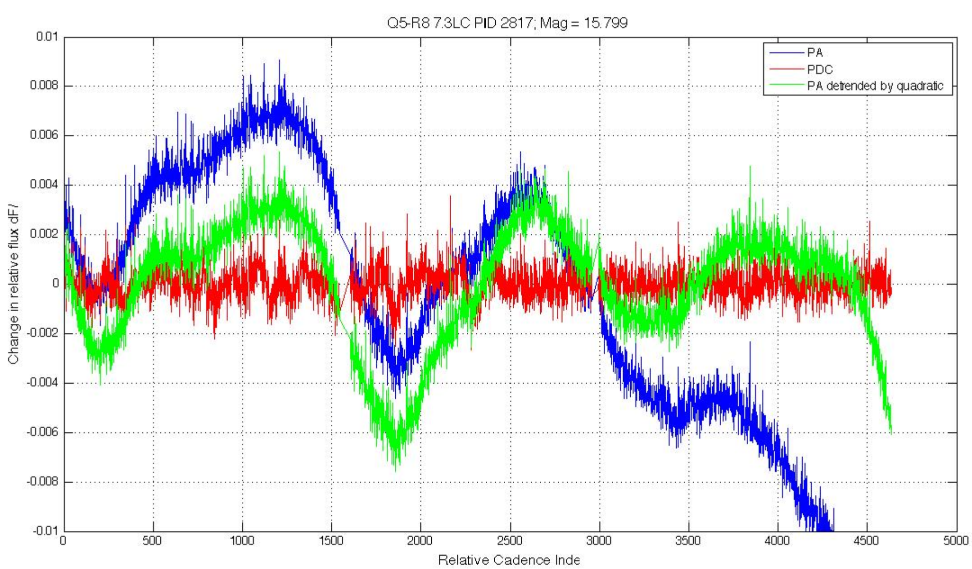

Figure 11: Q5 example of PDC removal of harmonic stellar variability. The amplitude of the variability with respect to a quadratic trend is ± 0.54%, just over the 0.50% threshold, before initial cotrending. Standard cotrending eliminated the variability but increased the noise, which PDC did not detect in this case. As described in

Section 4.4.1 Step 7, if PDC had detected the increase in noise, it would have retained the result of the harmonic-removed cotrending, and restored the harmonic content before export to MAST.

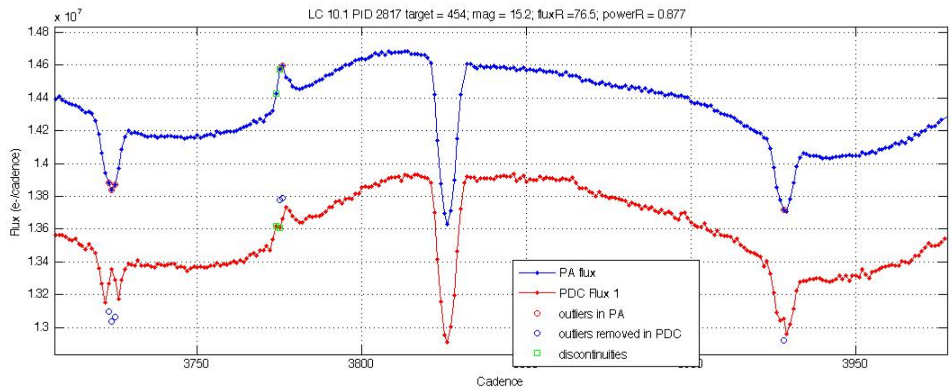

Figure 12: Q5 example of astrophysical events, possibly flares, identified by PDC and partially removed from the corrected light curve at cadence 3775. Open green squares show astrophysical events identified as discontinuity anomalies and "corrected." Unlike outliers, discontinuities are *not* restored to the light curves before delivery to MAST.

5. Lost or Degraded Data

In this Section, we discuss Cadences that are essentially lost to high-precision photometry due to planned or unplanned spacecraft events. Particularly important and unexpected phenomena are written up as Kepler Anomaly Reports (KARs), or mitigated by SOC change requests (KSOCs).

5.1 Momentum Desaturation

Solar radiation torque causes angular momentum to build up in the reaction wheels, which then must be desaturated by thruster firings when the wheels spin up to their maximum operating RPM. Desats occur every 3 days. The spacecraft (S/C) is not designed to maintain Fine Point control during these events, and enters Coarse Point mode. The subsequent image motion is sufficient to spoil the photometric precision of data collected during desats, and a few minutes after desats during which the spacecraft restores Fine Point control. One LC and several SCs are affected for each desaturation.

The momentum dump Cadences have -Infs in the delivered light curve files, but finite values in the uncalibrated and calibrated pixels. The dump Cadences are listed in Table 3 so that users of time series will know which -Infs are due to desats, and users of pixel data will know which Cadences to exclude from their own analyses. Tables of the more numerous SCs afflicted by desats is included in the Supplement, though they duplicate some of the information in the SC data anomaly tables (Section 5.3.6).

Table 3: Momentum dumps in Q5 and the corresponding Long Cadences. CIN = Cadence Interval Number, RCI = Relative Cadence Index.

```
List of Momentum Dump Cadences
CIN    RCI     Date(MJD)
 16444     72 55277.44194
 16591    219 55280.44568
 16736    364 55283.40855
 16882    510 55286.39186
 17029    657 55289.39560
 17174    802 55292.35847
 17320    948 55295.34177
 17467   1095 55298.34551
 17612   1240 55301.30839
 17758   1386 55304.29169
 17905   1533 55307.29543
 18054   1682 55310.34003
 18200   1828 55313.32334
 18346   1974 55316.30665
 18492   2120 55319.28995
 18638   2266 55322.27326
 18784   2412 55325.25656
 18930   2558 55328.23987
 19076   2704 55331.22318
 19222   2850 55334.20648
 19328   2956 55336.37244
 19473   3101 55339.33532
```

19619	3247	55342.31862
19765	3393	55345.30193
19911	3539	55348.28523
20057	3685	55351.26854
20203	3831	55354.25185
20349	3977	55357.23515
20495	4123	55360.21846
20641	4269	55363.20177
20787	4415	55366.18507
20933	4561	55369.16838
21005	4633	55370.63960

5.2 Reaction Wheel Zero Crossings

The descriptive part of this Section and the Figures are unchanged since Release 6 (Q4). Table 4 has been updated to show Q5 values.

Another aspect of spacecraft momentum management is that some of the reaction wheels cross zero angular velocity from time to time. The affected wheel may rumble and degrade the pointing on timescales of a few minutes. The primary consequence is an increased noise in the Short Cadence centroids, and pixel and flux time series. The severity of the impact to the SC flux time series seems to vary from target to target, with all SC centroid and pixel time series showing some impact. In some cases, we observe negative spikes of order 10^{-3} to 10^{-2} in SC relative flux time series (Figure 13), and these Cadences must be excluded from further analysis. The impact on Long Cadence data is much less severe in both amplitude and prevalence. Zero crossings are not gapped in this Release, and users will have to use Table 4 to identify possibly afflicted Cadences.

In Figure 13, the noise in centroids and loss of flux occurs on multiple stars during the zero crossing, so this noise is not the result of an uncorrected cosmic ray event or other local transient. Neither is it due to the momentum dumps (Section 5.1) labeled in the Figure, for which one or two Cadences right after the dump may have bad pointing, but are not flagged as data gaps by the Pipeline. The zero crossings occur at distinctly different times than the momentum dumps.

Since the Pipeline does not flag zero crossings as anomalous data, the zero crossing events are shown in Table 4. Events were identified in reaction wheel telemetry, which is not sampled synchronously with Cadences. For each zero crossing event, the last Cadence ending before the event and the first Cadence beginning after the event were identified. Overlap between events is due to this rounding of Cadence numbers at times when the slowest wheel had nonzero speed for a time interval shorter than 2 Cadence periods.

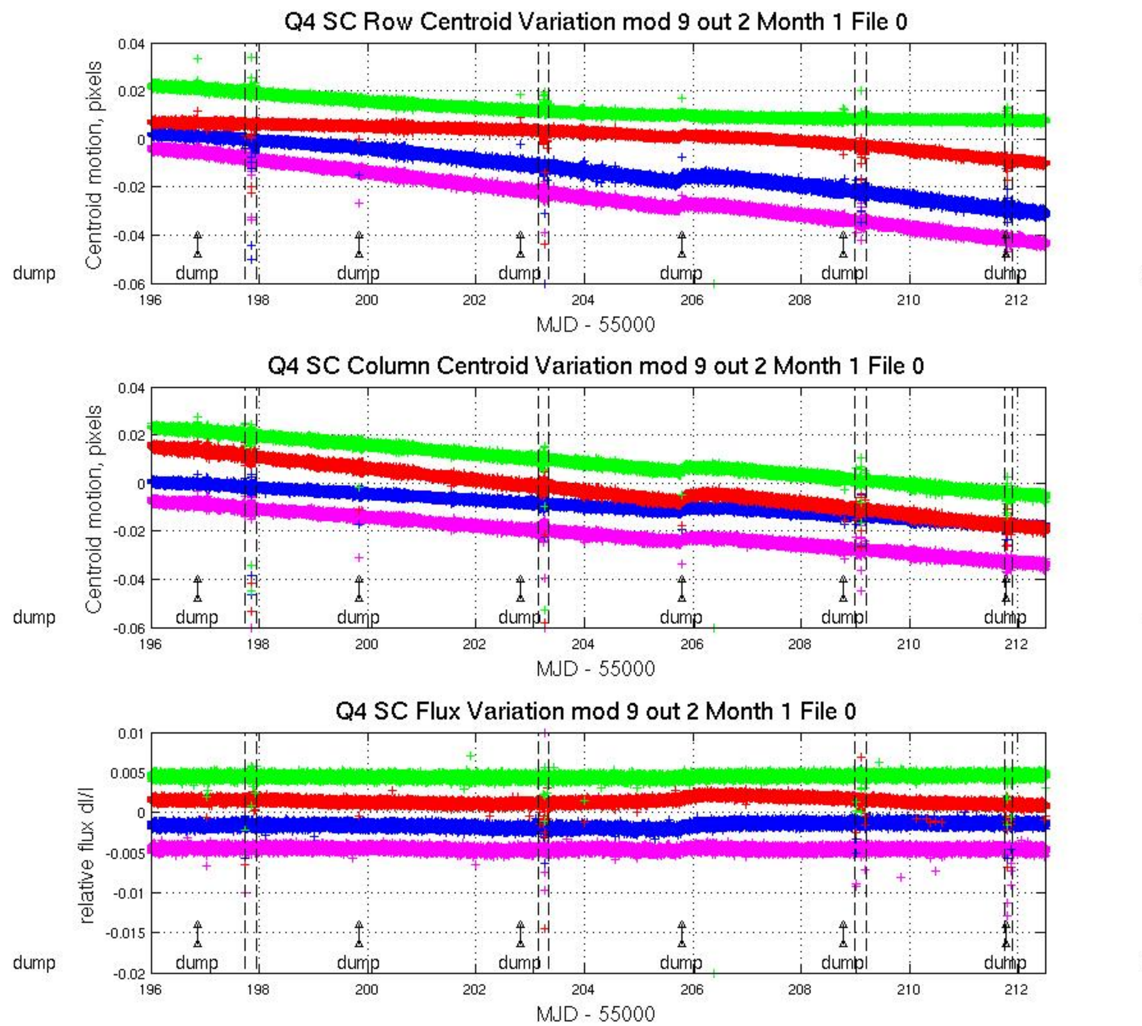

Figure 13: Example from Q4 (Release Notes 6) of the effect of reaction wheel speed zero crossing on SC flux and centroids. The plots show row and column centroid motion, and the relative flux change, in the neighborhood of zero crossings. The data on several stars are overplotted in different colors in each panel of the Figure. Vertical dashed black lines bracket the times during which at least one wheel had zero speed according to its telemetry. The curves are offset for clarity, and momentum dumps are labeled. The kink in the data at MJD = 55205.72 is the failure of mod 3.

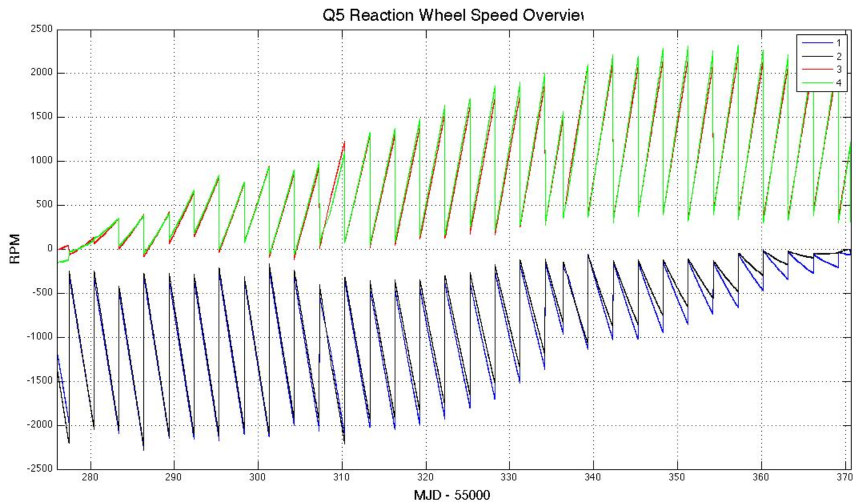

Figure 14: Overview of reaction wheel speeds in Q5 rounded to the nearest Short Cadence. Zero crossings occur in the first half of the quarter for wheels 3 & 4 and near the end of the quarter for wheels 1 & 2.

Table 4: Zero crossing events in Q5, defined as the time from first to last zero crossing in the event, rounded to the nearest Cadence. The corresponding cadence numbers for LC are in the supplement.

Column definitions:

MJDStart – MJD of start of zero crossing event

MJDEnd – MJD of end of zero crossing event

CINStart – Cadence Interval number at beginning of zero crossing event

CINEnd -– Cadence Interval number at end of zero crossing event

RCIStart – Relative cadence index at beginning of zero crossing event

RCIEnd - Relative cadence index at end of zero crossing event

NumTLMSamp – Number of engineering telemetry sample in the zero crossing event

Event#	MJDStart	MJDEnd	CINStart	CINEnd	RCIStart	RCIEnd	NumTLMSamp
1	55275.991	55276.327	479665	480157	16	508	242
2	55276.344	55276.367	480182	480216	533	567	17
3	55276.373	55276.400	480225	480265	576	616	20
4	55278.164	55278.181	482854	482880	3205	3231	13
5	55278.208	55278.580	482919	483466	3270	3817	269
6	55278.585	55278.602	483472	483497	3823	3848	13
7	55286.641	55286.923	495300	495714	15651	16065	204
8	55295.342	55295.480	508075	508278	28426	28629	100
9	55301.496	55301.595	517110	517256	37461	37607	72
10	55304.465	55304.544	521469	521585	41820	41936	57
11	55304.552	55304.607	521596	521678	41947	42029	41
12	55370.148	55370.164	617903	617926	135374	135397	12
13	55370.166	55370.659	617929	618653	135400	136124	356

5.3 Data Anomalies

5.3.1 Safe Mode

From time to time, the Kepler Spacecraft will go into Safe Mode, because of an unanticipated sensitivity to cosmic radiation, or unanticipated responses to command sequences.

There were no safe modes in Q5.

5.3.2 Loss of Fine Point

From time to time, the Kepler spacecraft will lose fine pointing control, rendering the Cadences collected useless for photometry of better than 1% precision. While the LOFPs are treated as lost data by the Pipeline, users with sources for which ~1% photometry is scientifically interesting may wish to look at the pixel data corresponding to those Cadences. On-orbit Flight Software upgrades have greatly reduced the frequency and duration of LOFPs.

There were no LOFPs in Q5.

5.3.3 Pointing Drift and Attitude Tweaks

Daily reference pixels are used by the SOC/SO to measure S/C attitude. The SOC PDQ software uses centroids of 3-5 stars per module/output to determine the measured boresight attitude compared with the pointing model (which accounts for differential velocity aberration). The Photometer Attitude

Determination (PAD) software performs a similar calculation to reconstruct the attitude using the Long Cadence science data when the data are processed after each downlink, and reprocessed on a Quarterly basis before delivery to MAST. The PAD attitude errors (RA, Dec, roll) for Q5 are shown in Figure 15. The maximum attitude residual (MAR) is the largest distance between the expected and actual location of a star in its aperture, for a given Cadence. The RSS sum of RA, Dec, and roll errors is an upper bound on the rigid body component of MAR and is also shown in the Figure.

Since continued attitude drift would invalidate target aperture definitions and lead to large photometric errors, small attitude adjustments ("tweaks") are performed if necessary to ensure that Maximum Attitude Residual (MAR) is always < 100 mpix.

There were no attitude tweaks in Q5.

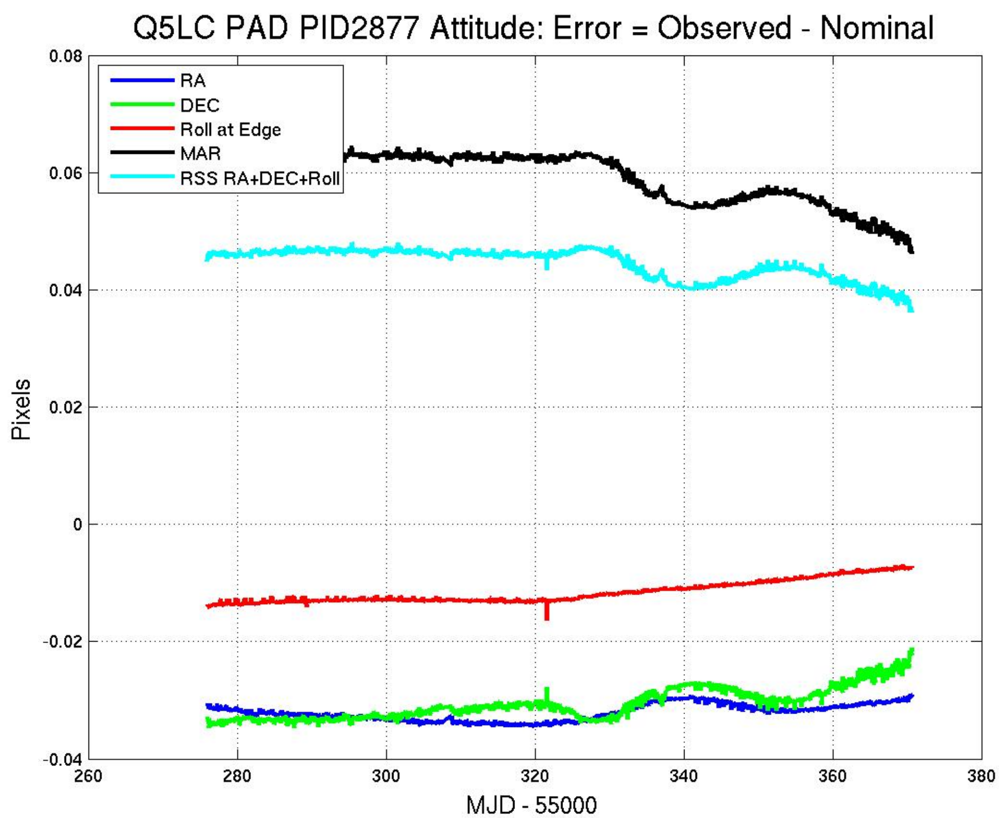

Figure 15: Attitude Error in Quarter 6, calculated using Long Cadence data.

5.3.4 Downlink Earth Point

Science data is downlinked once a month, and the spacecraft changes its attitude to point its fixed High Gain Antenna (HGA) at the Earth. Science data collection ceases, and the change in attitude induces a thermal transient in the

Photometer. In this Release, data collected after Earth Point are corrected in the same way as data after a Safe Mode.

5.3.5 Manually Excluded Cadences

Occasionally, a Cadence is manually excluded, usually near a gap or discontinuity in the data that makes it difficult to exclude them automatically. It was not necessary to do this in Q5 Release 8.

5.3.6 Anomaly Summary Table

Table 5 shows a summary of the Anomalies for both LC and SC data. COARSE_POINT in the SC table indicates normal attitude recovery after a momentum dump and not a Loss of Fine Point event (Section 5.3.2).

Table 5: Anomaly Summary Table for Long and Short Cadences.

```
Unique Anomaly Labels with bit codes:
  1 ARGABRIGHTENING
  2 EARTH_POINT

 Column Definitions
1. Cadence Interval Number
2. Relative Cadence Index
3. Cadence mid-Times, MJD
4. Number of Anomalies on Cadence
5. Anomaly List bit-code
 16373      1 55275.99115   1   2
 16430     58 55277.15587   1   1
 17916   1544 55307.52020   1   2
 17917   1545 55307.54064   1   2
 17918   1546 55307.56107   1   2
…
 19359   2987 55337.00589   1   2
 19360   2988 55337.02632   1   2
 19361   2989 55337.04676   1   2
 19362   2990 55337.06719   1   2
 19363   2991 55337.08762   1   2
(See Supplement for the entire table)
```

5.4 Module 3 failure

All 4 outputs of Module 3 failed at 17:52 UTC Jan 9, 2010, during LC CIN 12935 during Quarter 4. Reference pixels showed loss of stars and black levels decreased by 75 to 100 DN per frame. FFIs show no evidence of photons or electrically injected signals. The start of line ringing and FGS crosstalk are still present after the anomaly.

The loss of the module led to consistent temperature drops within the LDE, telescope structure, Schmidt corrector, primary mirror, FPA modules, and acquisition/driver boards.

After a review of probable causes, it was concluded that the probability of a subsequent failure was remote, a conclusion supported by continued operation of all the other Modules in the last 9 months.

The impact on science observations is that 20% of the FOV will suffer a one-quarter data outage every year as Kepler performs its quarterly rolls.

5.5 Incomplete Apertures Give Flux and Feature Discontinuities at Quarter Boundaries

Some users have reported larger than expected flux and flux slope discontinuities between Quarters. A degree of mismatch of flux at Quarter boundaries is expected, because the target has moved to a different mod.out with a possibly different aperture and crowding metric. Even worse, changes in relative feature depths between Quarters have also been seen. In each case to date, this problem has been due to the fact that the optimal aperture pixels (Ref. 16) have omitted bleeding charge from sources that saturate 3 or more pixels (Kepler magnitude 11 or brighter). These omissions are due to the low fidelity of the Kepler pre-flight saturation model and occasional errors in the KIC Kepler magnitude, particularly for variable stars. The depth variation problem at the Quarter boundary can often be substantially mitigated by summing all the calibrated pixels, not just those in the optimal aperture. However, if charge has bled outside the full target aperture (which includes a halo of pixels around the estimated optimal subset), then that information is irretrievably lost. Unfortunately, target pixel files are not yet available to users, so users concerned about large inter-quarter discontinuities in bright star flux time series need to contact the Science Office.

As the mission has progressed stars with poorly captured saturation have been identified and their apertures improved. Therefore there are more targets that have poorly captured saturation early in the mission. Users of bright targets are strongly advised to check the pixel files once they are available to be sure that the saturated flux is captured in the optimal aperture.

In the near future the problem of capturing saturation will be largely fixed by using the observed saturation for a star to set the optimal aperture for that star.

6. Systematic Errors

This Section discusses systematic errors arising in on-orbit operations, most of which get removed from flux time series by PA or PDC (Section 4). While the Release 8 data is cotrended against image motion (as represented by the Cadence-to-Cadence coefficients of the motion polynomials calculated by PA) as well as LDE board temperatures, other telemetry items which may be used for cotrending the data in future releases are included in the Supplement so that users can at least qualitatively assess whether features in the time series look suspiciously like features in the telemetry items. This telemetry has been filtered and gapped as described in the file headers, but the user may need to resample the data to match the LC or SC sampling. In addition, PDC corrects systematic effects only in the flux time series, and this Section and the associated Supplement files may be useful for users interested in centroids or pixel data (when available).

Most of the events described in this Section are either reported by the spacecraft or detected in the Pipeline, then either corrected or marked as gaps. This Section reports events at lower thresholds than the Pipeline, which affect the light curves and therefore may be of interest to some users.

6.1 Argabrightening

Argabrightening, named after its discoverer, V. Argabright of BATC, is a presently unexplained diffuse illumination of the focal plane, lasting on the order of a few minutes. It is known to be light rather than an electronic offset since it appears in calibrated pixel data from which the electronic black level has been removed using the collateral data. It is not a result of gain change, or of targets moving in their apertures, since the phenomenon appears with the same amplitude in background pixels (in LC) or pixels outside the optimal aperture (in SC) as well as stellar target pixels. Many channels are affected simultaneously, and the amplitude of the event on each channel is many standard deviations above the trend, as shown in Figure 16.

The method of detection is

1. Calculate the median, for each Cadence and mod.out, of the calibrated background (LC) or out-of-optimal-aperture (SC) pixels,
2. Detrend the data by fitting a parabola to the resulting time series and subtract the fit.
3. High-pass filter the detrended data by median filtering the detrended data using a 25 Cadence wide filter, and subtracting that median-filtered curve from the detrended data to form the residual background light curve.
4. Calculate the Median Absolute Deviation (MAD) of the residual. The Argabrightening statistic S_{Arg} is then the ratio of the residual to the MAD.

5. Find S_{Arg} which exceed the single-channel threshold T_{MAD} , and subsequently treat those Cadences as gaps for all pixels in that channel. In the current version of the Pipeline, T_{MAD} is the same for all channels.
6. A multichannel event is detected on a given Cadence if the number of mod.outs for which $S_{Arg} > T_{MAD}$ on that Cadence exceeds the multi-channel event threshold T_{MCE} . Then all mod.outs on that Cadence are marked as gaps, even those channels that did not individually exceed T_{MAD}. Multichannel event detection allows the use of lower T_{MAD} while still discriminating against spurious events on isolated channels.
7. For multichannel events, average S_{Arg} over all 84 outputs of the FPA to form $\langle S_{Arg} \rangle_{FPA}$

The Pipeline uses a rather high T_{MAD} = 100 for LC and 60 for SC, and a high T_{MCE} = 42 (half the mod.outs). While it appears that background subtraction has mostly removed this phenomenon from the delivered Long Cadence data, the residual effect has not been proven to be negligible in all cases, especially in Short Cadence data. There may also be significant Argabrightening events in both LC and SC, which do not exceed the thresholds. This Section gives a summary of events with lower thresholds T_{MAD} = 10 and T_{MCE} = 10 (Long Cadence in Table 6 and Short Cadence in Table 7), so that the user may consider whether some Cadences of interest might be afflicted by Argabrightening, but not identified as such by the Pipeline and gapped (i.e., -Inf in all columns of the light curve file, except those referring to time or CIN). The Supplement contains these detection summaries as ASCII files.

The Supplement also contains the channel-by-channel background time series so users can identify low-level or few-channel Argabrightenings using their own criteria. These time series may also be useful for correcting SC data collected during Argabrightening events, since the Pipeline background correction interpolates LC background data to calculate the background for SC data. Users may notice some “chatter” in the background time series. A preliminary study shows that the problem is present in the calibrated background pixels, but not in the raw pixels, and is present in about 25% of the channels, with an amplitude up to 3% of the background. The reasons are still under investigation.

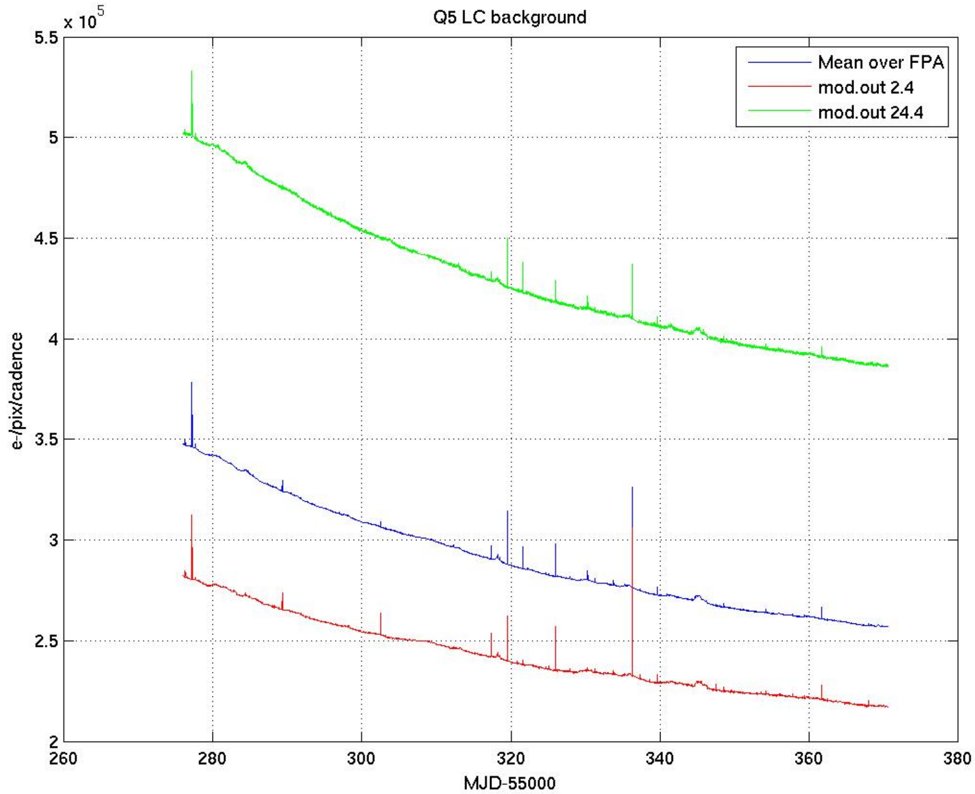

Figure 16: Background time series for Q5 showing the average over all mod.outs, and the modules furthest from (2.4) and nearest to (24.4) the Galactic plane. The narrow spikes common to all 3 curves are Argabrightening events.

Table 6: Q5 LC Argabrightening Events with amplitude $T_{MAD} > 10$, and occurring on a number of channels $T_{MCE} > 10$. The columns are (1) CIN = Cadence Interval Number for Argabrightening Cadences, (2) RCI = Relative Cadence Index for Argabrightening Cadences, (3) Date = Arg Cadence mid-Times, MJD, (4) Mean Argabrightening statistic over Channels of Arg Event $\langle S_{Arg} \rangle_{FPA}$ (5) N_chan = Channels exceeding threshold in Arg Cadence, (6) N_pipe = Channels exceeding default (Pipeline) threshold in ArgCadence. MAD is calculated on a channel-by-channel basis.

Q5

CIN	RCI	Date (MJD)	$\langle S_{arg} \rangle_{FPA}$	N_{chan}	N_{pipe}
16386	14	55276.25679	22.4	77	0
16430	58	55277.15587	265.6	80	78
16454	82	55277.64628	19.8	76	0
17024	652	55289.29343	45.7	69	3
17399	1027	55296.95603	6.9	17	0
17669	1297	55302.47310	18.0	22	6
18148	1776	55312.26079	7.0	17	0
18397	2025	55317.34876	55.8	80	3
18436	2064	55318.14567	9.2	29	0

18440	2068	55318.22740	14.4	56	0
18456	2084	55318.55434	12.4	50	0
18505	2133	55319.55559	220.8	80	78
18568	2196	55320.84291	6.9	14	0
18605	2233	55321.59895	93.2	68	29
18759	2387	55324.74572	8.6	28	0
18778	2406	55325.13396	6.7	23	0
18819	2447	55325.97174	138.8	80	54
18917	2545	55327.97423	9.6	34	0
19031	2659	55330.30366	35.7	74	0
19080	2708	55331.30491	16.2	72	0
19202	2830	55333.79781	20.6	66	0
19278	2906	55335.35076	6.1	13	0
19324	2952	55336.29071	426.5	80	79
19374	3002	55337.31239	7.1	28	0
19491	3119	55339.70312	33.3	79	0
19876	3504	55347.57006	6.4	25	0
19928	3556	55348.63261	14.4	53	0
20159	3787	55353.35277	6.2	22	0
20201	3829	55354.21098	14.7	60	0
20568	4196	55361.71011	50.0	79	0
20881	4509	55368.10583	7.1	16	0

Table 7: Same analysis as Table 6, for Q5 SC. Note consecutive detections of the largest events. A horizontal line separates the 3 Months of the Quarter. The relative cadence index (RCI) is reset at the start of each Month.

Q5

CIN	RCI	Date (MJD)	$<S_{arg}>_{FPA}$	N_{chan}	N_{pipe}
480052	403	55276.25509	24.2	78	0
481386	1737	55277.16370	266.9	80	80
481387	1738	55277.16438	23.0	79	0
482101	2452	55277.65070	16.9	64	0
499196	19547	55289.29445	8.0	27	0
499197	19548	55289.29513	36.7	70	12
518541	38892	55302.47072	13.4	17	9
532912	5113	55312.25909	7.3	12	0
540395	12596	55317.35591	13.5	66	0
540396	12597	55317.35659	18.0	55	1
540398	12599	55317.35795	25.8	60	1
542152	14353	55318.55264	8.6	27	0
543615	15816	55319.54912	133.0	80	78
543616	15817	55319.54980	109.6	80	69
545504	17705	55320.83575	7.7	21	0
546638	18839	55321.60814	101.1	66	50
551234	23435	55324.73857	9.4	32	0
551981	24182	55325.24737	39.2	79	6
553044	25245	55325.97140	53.3	80	28
553045	25246	55325.97208	28.0	80	0
553046	25247	55325.97276	18.2	69	0
553047	25248	55325.97344	26.9	74	2

553048	25249	55325.97412	15.8	58	0
555994	28195	55327.98070	9.9	36	0
559400	31601	55330.30060	41.5	75	14
560861	33062	55331.29571	17.9	71	0
564523	36724	55333.78998	9.9	41	0
564524	36725	55333.79066	9.9	36	0
568183	40384	55336.28288	440.5	80	80
568184	40385	55336.28356	24.6	60	10
568185	40386	55336.28424	6.6	14	2
569701	322	55337.31682	5.3	20	0
573192	3813	55339.69461	35.5	80	0
586320	16941	55348.63635	15.3	64	0
594508	25129	55354.21336	9.4	38	0
605521	36142	55361.71454	38.0	80	4
605522	36143	55361.71522	9.3	30	0
612091	42712	55366.18950	31.7	79	1

6.2 Variable FGS Guide Stars

Variable FGS Guide stars are not expected to be significant in Q5 photometry.

6.3 Pixel Sensitivity Dropouts

Section 6.3.1 is unchanged from Release 6.

6.3.1 Particle-induced

This Section is unchanged from Release 6.

Space-based focal planes respond to cosmic ray (CR) events in several ways:

1. A transient response is induced by the charge deposited by the CR, and is cleared by the next reset (destructive readout) of the pixel.
2. Medium-term alteration of detector properties, which recover to near or at their pre-event values after some time and resets without annealing.
3. Long-term alteration of detector properties, which are only restored by annealing the focal plane
4. Permanent damage

Typically, type 3 and 4 effects are caused by non-ionizing energy loss (NIEL), or "knock-on" damage, which can be caused by any baryonic particle.

Type 1 effects are removed by the Pipeline's CR detection algorithm. At this point in the mission, type 3 effects do not appear to be common enough to warrant the disruption of the observing schedule that would be caused by annealing, and both type 3 and type 4 effects will eventually be mitigated by updating the bad pixel map used for calibration. Type 2 effects are not corrected by the Pipeline at the pixel level (Figure 17). In this Release, the Pipeline corrects the aperture flux discontinuities (Figure **18**) resulting from these pixel discontinuities (Section 4.4), though users examining pixel data and uncorrected light curves need to remain aware of them.

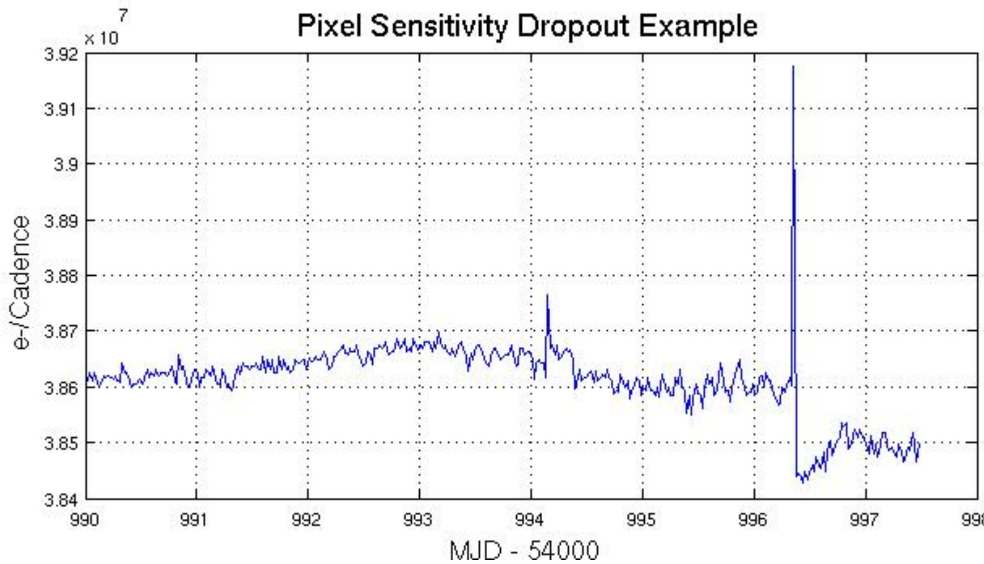

Figure 17: Pixel time series from Q1 (Release 2) showing discontinuity after large CR event. CRs have not been removed by the Pipeline at this stage of processing. Target: KeplerID = 7960363, KeplerMag = 13.3. Dropouts are not corrected on a pixel-by-pixel basis.

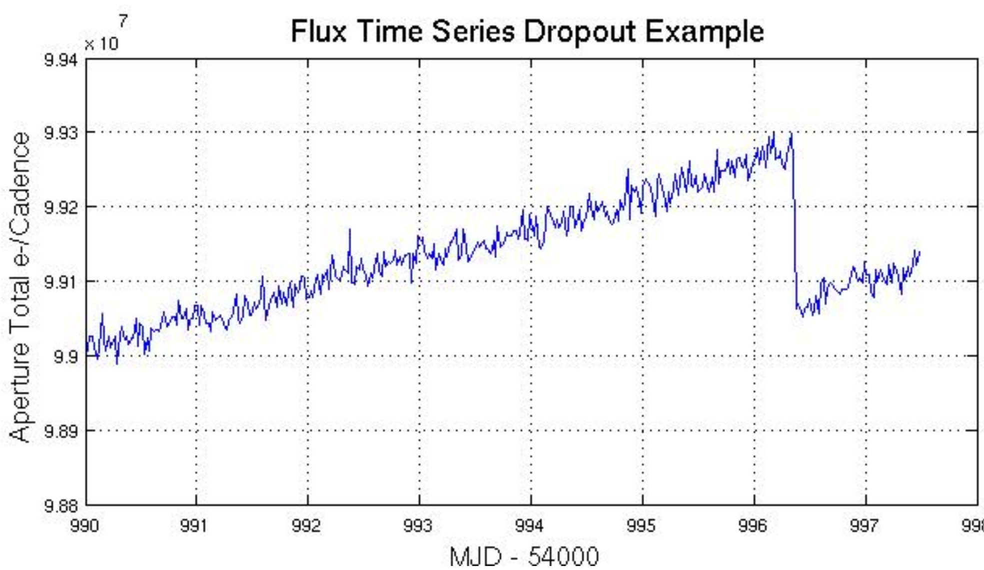

Figure 18: **Same event as for the previous Figure as seen in the uncorrected Simple Aperture Photometry (SAP) light curve produced by PA. CR hits have been removed by PA. This figure is presented for historical interest, as PDC identifies most of these discontinuities and removes them before producing the corrected light curves (Section 4.4).**

6.4 Focus Drift and Jitter.

Examination of Q1 data revealed that many of the science targets exhibit non-sinusoidal variations in their pixel time series with a period between 3 and 6 hours. The behavior was less frequent at the beginning of Q1 and becomes progressively worse with time. Initially, this phenomenon was associated with desaturation activities, but became nearly continuous about 15 days into the observations. The problem persisted through the end of Q3 (see Release 4 Notes, KSCI-19044). It should not be observable in Q5.

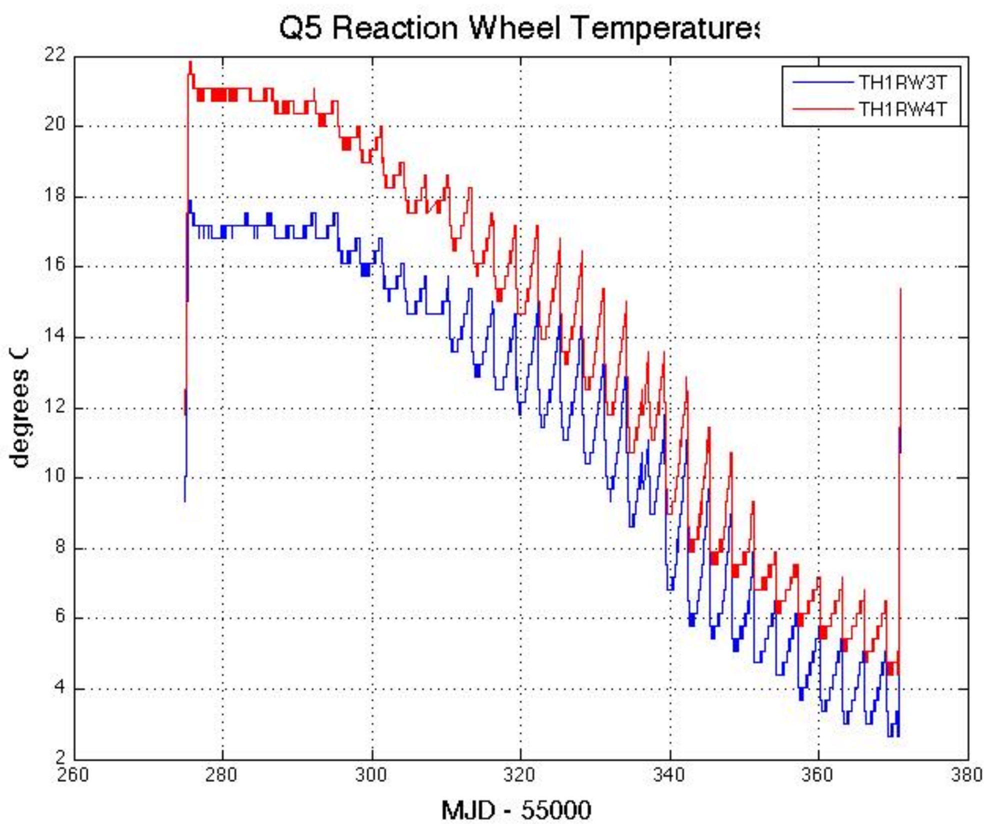

Figure 19: Reaction wheel adapter temperatures during Q5. Temperature variation in Q5 is dominated by a slow seasonal drift and the 3 day period of reaction wheel desaturations. The telemetry data in this Figure is not plotted for times when the spacecraft is not in Fine Point, and is smoothed with a 5 point median filter.

The DAWG has investigated whether there is a secular variation of the focus and PRF width. Preliminary results indicate that the seasonal cycle dominates, with a good correlation between PRF width variation and the temperature of the Launch Vehicle Adapter (TH2LVAT), as shown in Figure 20. The pattern has begun to repeat, now that a full year of science data collection has occured.

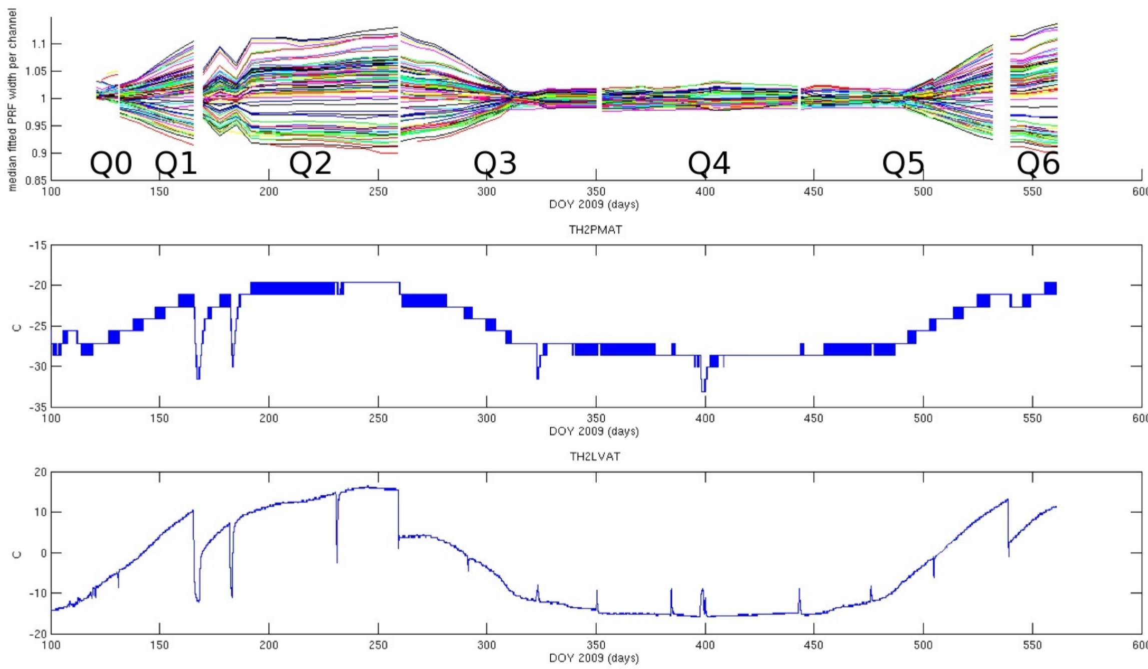

Figure 20: Correlation of variation in PRF width with various spacecraft temperatures, demonstrating the seasonal nature of focus and PRF changes.

For users of the PDC output, the focus changes are mostly captured by the motion polynomial coefficients used for cotrending. For users doing their own cotrending, the mod.out center motion time series provided in the Supplement will represent much of the image motion resulting from focus changes, for all targets on the corresponding mod.out. However, they do not represent local plate scale changes, which may contribute systematic errors to the light curves of individual targets on that mod.out. Thus the reaction wheel and Launch Vehicle Adapter temperature sensor telemetry for Q5 are also provided in the Supplement.

6.5 Short Cadence Requantization Gaps

This section is unchanged since Release 6, except to emphasize this is a Short Cadence phenomenon.

Short Cadence pixels at mean intensities >20,000 e- show banding as shown in Figure 21, with quantized values of number of electrons preferred. This is the result of the onboard requantization (KIH Section 7.4), and is considered benign since in the overall extraction the light curve is near the Poisson limit. These requantization gaps are expected, and a necessary cost associated with achieving the required compression rates on board Kepler. However, it is pointed out here so that users will not suspect a problem.

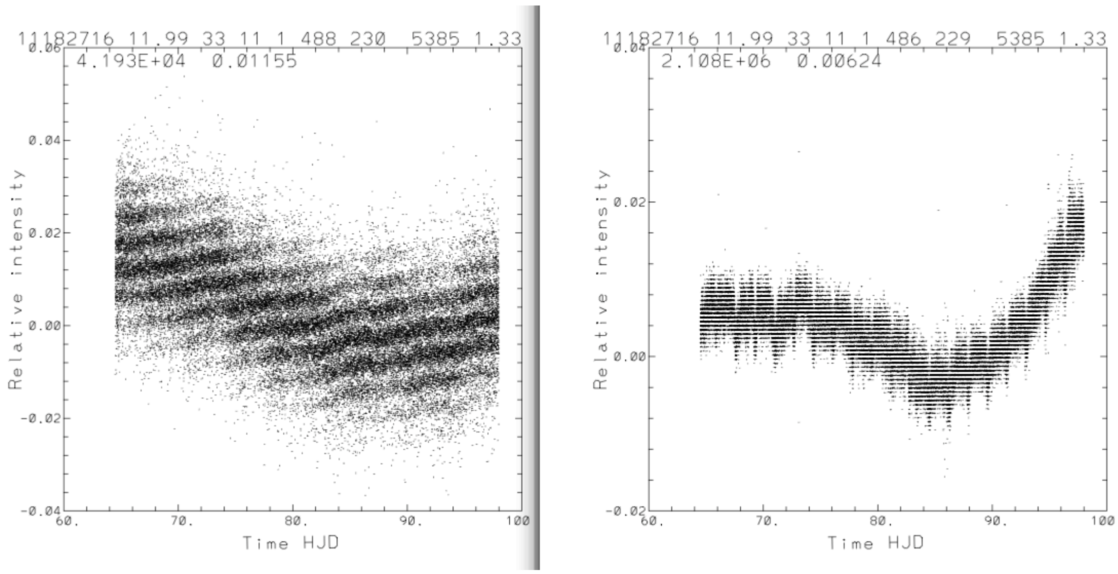

Figure 21: Requantization gap example in Q1 SC pixel time series. The 'band gaps' scale with mean intensity (42,000 e- left, 2.1e6 right). See KIH Section 7.4 for a discussion of quantization and the (insignificant) information loss it entails.

6.6 Spurious Frequencies in SC Data

Section 6.6.1 is unchanged from Releases 6.

6.6.1 Integer Multiples of Inverse LC Period

Spurious frequencies are seen in SC flux time series, and pixel data of all types – including trailing black collateral pixels. The frequencies have an exact spacing of 1/LC interval, as shown in Figure 22. As the SC data are analyzed in the frequency domain in order to measure the size and age of bright planetary host stars, the contamination of the data by these spurious frequencies will complicate these asteroseismology analyses, but will not compromise the core Kepler science. The physical cause of this problem is still under discussion, though the problem might be remedied with a simple comb notch filter in future releases even if no ancillary data can be found that exhibits these features.

This feature was first reported in Q1 data (Ref. 8). It has now been identified in pre-launch ground test data as well as later, and is therefore considered a normal feature of the as-built electronics. It is not an artifact introduced by the Pipeline, since it appears in raw trailing black collateral data.

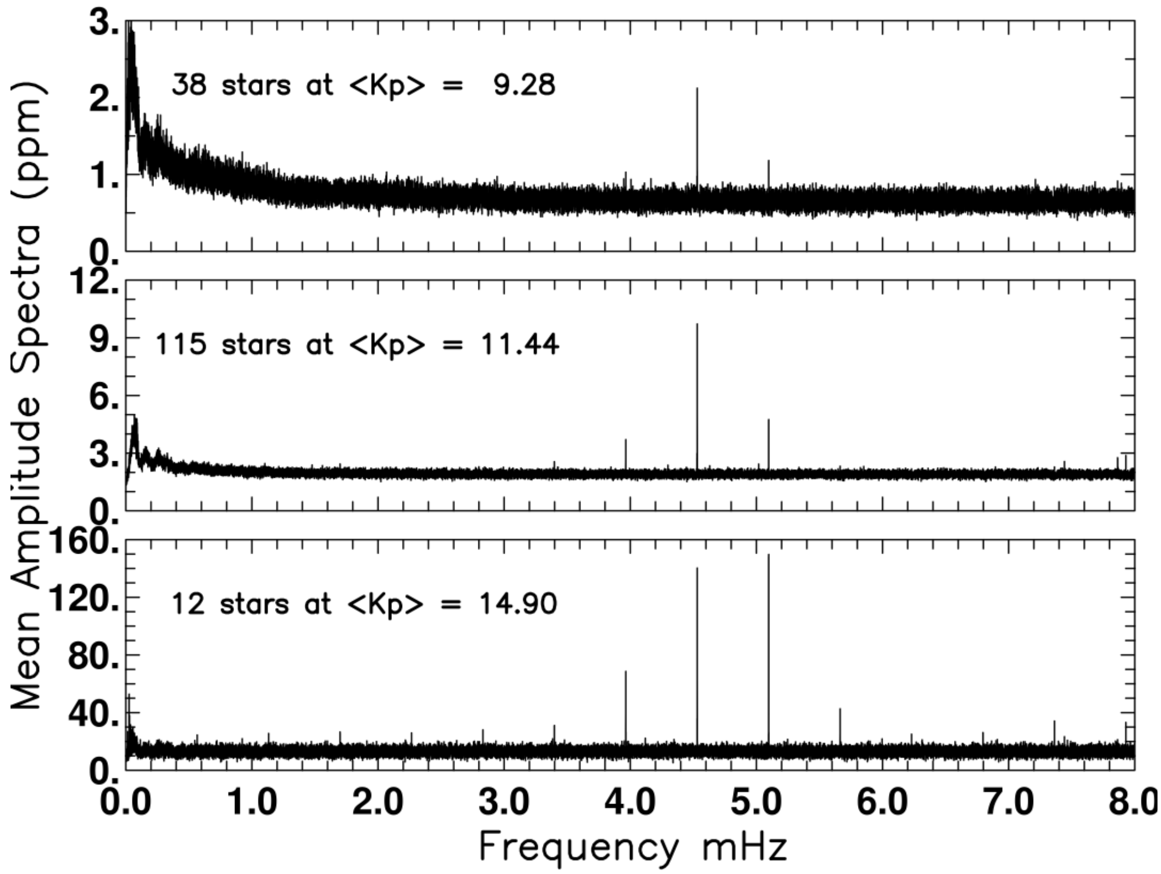

Figure 22: Mean amplitude spectra over samples of quiet stars from Q1, spanning more than a factor of 100 in brightness, showing spurious frequencies. The 1/LC-Cadence artifacts at the fundamental of 0.566391 mHz and all harmonics are visible for the faint star set in the bottom panel. Even at 9th magnitude in the upper panel this artifact remains a dominant spectral feature from the 7th and 8th harmonics. From Gilliland et al. (Ref. 8).

6.6.2 Other Frequencies

Further analysis of SC data in Q1 and subsequent Quarters showed several stars in which the SC data showed peak power at 7865 μHz (~127.16 seconds). This is not a harmonic of the 1/LC noise discussed in the previous Section. Across the Q2M1 safing event, the phase shifted for both the 1/LC harmonics and for the 7865 μHz feature. Across the Q2M3 safing event the phase remained fixed. Since stellar signals tend to stay at the same phase, the phase shift across Q2M1 is evidence that the n/LC and 7865 μHz features are instrumental. Peaks have also been reported at 7024, 7444, 7865, and 8286 μHz -- consistent with a splitting of 421 μHz = 2375.3 s, or 39.59 minutes.

In Q0-Q2, multiple groups reported the issues around 80-95 μHz which correspond to about 3.2 hours. The non-sinusoidal nature of these spurious signals leads to evenly spaced peaks, not unlike stellar oscillations. Since this is the same period as the temperature variation of the reaction wheel housing temperature, and that variation has been eliminated by reducing the corresponding temperature controller deadband (see DRN #6 on Q4). Users are, however, encouraged to examine the thermal telemetry shown in these Notes and provided in the Supplement to strengthen the case that detected spectral features are astrophysical, not instrumental.

In Q3, broadband features around 270 - 360 μHz occur in several stars, corresponding to periods of 0.75 - 1 hour. Q5 SC data has not yet been characterized for potential unique behavior.

A period of about 3 days has been reported multiple times, and is almost certainly associated with the momentum management cycle and associated temperatures (Figure 19).

Table 8: List of Possible Spurious Frequencies in SC data. Users are advised to check detections against this list, and report additional spurious frequencies to the Science Office. Labels: RW = reaction wheel passive thermal cycle associated with momentum cycle. RWTH = Reaction wheel housing temperature controller thermal cycling (believed not to be a problem from Q3 onward). U = unknown. Narrow lines are defined as $\nu/\Delta\nu$ > 50, broad lines as $\nu/\Delta\nu$ < 50.

SC spurious frequency summary

frequency uHz	frequency d^{-1}	Period S	Period Min	period hr	Period D	Label	width
8.9	0.33	112320.00	1872.000	72.0000	3.00000	RW	?
86.8	7.50	11520.00	192.000	3.2000	0.13333	RWTH	broad
290.0	25.06	3448.28	57.471	0.9579	0.03991	U1	broad
340.0	29.38	2941.18	49.020	0.8170	0.03404	U2	broad
360.0	31.10	2777.78	46.296	0.7716	0.03215	U3	narrow
370.4	32.00	2700.00	45.000	0.7500	0.03125	U4	narrow
421.0	36.37	2375.30	39.588	0.6598	0.02749	splittingU5-U8	narrow
566.4	48.94	1765.56	29.426	0.4904	0.02043	1/LC	narrow
1132.8	97.87	882.78	14.713	0.2452	0.01022	2/LC	narrow
1699.2	146.81	588.52	9.809	0.1635	0.00681	3/LC	narrow
2265.6	195.74	441.39	7.357	0.1226	0.00511	4/LC	narrow
2832.0	244.68	353.11	5.885	0.0981	0.00409	5/LC	narrow
3398.3	293.62	294.26	4.904	0.0817	0.00341	6/LC	narrow
3964.7	342.55	252.22	4.204	0.0701	0.00292	7/LC	narrow
4531.1	391.49	220.70	3.678	0.0613	0.00255	8/LC	narrow
5097.5	440.43	196.17	3.270	0.0545	0.00227	9/LC	narrow
5663.9	489.36	176.56	2.943	0.0490	0.00204	10/LC	narrow
6230.3	538.30	160.51	2.675	0.0446	0.00186	11/LC	narrow
6796.7	587.23	147.13	2.452	0.0409	0.00170	12/LC	narrow
7024.0	606.87	142.37	2.373	0.0395	0.00165	U5	narrow
7363.1	636.17	135.81	2.264	0.0377	0.00157	13/LC	narrow
7444.0	643.16	134.34	2.239	0.0373	0.00155	U6	narrow
7865.0	679.54	127.15	2.119	0.0353	0.00147	U7	narrow
7929.5	685.11	126.11	2.102	0.0350	0.00146	14/LC	narrow
8286.0	715.91	120.69	2.011	0.0335	0.00140	U8	narrow
8495.9	734.04	117.70	1.962	0.0327	0.00136	15/LC	narrow

7. Data Delivered – Format

7.1 FFI

The FFIs are one FITS file per image, with 84 extensions, one for each module/output. See the KIH to map the extension table number = channel number onto module and output.

A temporary procedure has been developed to populate FFIs with linear WCS information. Tests indicate that distortion and differential velocity aberration will cause systematic errors of ~< 1.5 pixels in the corners of the FFI.

In the future Releases, FFIs will use the SIP convention for representing distortion in FITS image headers (fits.gsfc.nasa.gov/registry/sip/SIP_distortion_v1_0.pdf).

7.2 Light Curves

This Section is unchanged since Release 6.

Light curves have file names like kplr<kepler_id>-<stop_time>, with a suffix of either llc (Long Cadence) or slc (Short Cadence), and a file name extension of fits.

A light curve is time series data, that is, a series of data points in time. Each data point corresponds to a measurement from a Cadence. For each data point, the flux value from simple aperture photometry (SAP) is given, along with the associated uncertainty. Only SAP light curves are available at this time. The centroid position for the target and time of the data point are also included.

The light curves are packaged as FITS binary table files. The fields of the binary table, all of which are scalar, are briefly described below and are listed in Table 9. There are 19 fields comprising 88 bytes per Cadence; however, fields 12-19 are not populated at this time. The FITS table header listed in the Appendix of the MAST manual is superseded by Table 9. The new keywords `DATA_REL` and `QUARTER` discussed in Section 2 are in the binary table header. The module and output are identified in the binary table extension header keywords `MODULE` and `OUTPUT`.

The following data values are given for each data point in a light curve:

- barycentric time and time correction for the midpoint of the Cadence
- for the simple aperture photometry (pixel sum) of optimal aperture pixels
    - first-moment centroid position of the target and uncertainty
    - uncorrected flux value and uncertainty. Gap Cadences are set to -Inf
    - corrected flux value and uncertainty. Gap Cadences are set to –Inf
    -

Table 9: Available light curve data table fields, modified after the MAST manual KDMC-10008 (August 30, 2009): SAP replaces OAP, and data in columns 12-19 is not available and are filled with -Inf. Time units are the same as in Releases 3-7.

Column Number	Field Name	Data Type	Bytes	Description	Units
1	Barytime	1D	8	barycentric time BJD – 2400000. See Section 7.4 for detailed discussion.	days
2	timcorr	1E	4	barycentric time correction. See Section 7.4 for detailed discussion	seconds
3	Cadence_number	1J	4	Cadence number (CIN)	N/A
4	ap_cent_row	1D	8	row pixel location	pixels
5	ap_cent_r_err	1E	4	error in row pixel location	pixels
6	ap_cent_col	1D	8	Column pixel location	Pixels
7	ap_cent_c_err	1E	4	error in column pixel location	pixels
8	ap_raw_flux	1E	4	SAP uncorrected flux	e- / Cadence
9	ap_raw_err	1E	4	SAP uncorrected flux error	e- / Cadence
10	ap_corr_flux	1E	4	SAP corrected un-filled flux	e- / Cadence
11	ap_corr_err	1E	4	SAP corrected un-filled flux error	e- / Cadence

Data Types:

1D – double precision floating point.

1E – single precision floating point. Note that, although all SOC calculations and internal data representation are double-precision, the SAP fluxes and errors are reported as single-precision floats, which will give roundoff errors of approximately 0.11 ppm (*Numerical Recipe*s Chapter 20 & confirmed by numerical experiments on MAST and internal SOC data).

1J – 32 bit integer

See Section 7.4 for a discussion of time and time stamps.

If you are an IDL user, the `tbget` program in the `astrolib` library extracts the data. If you are an IRAF user, tprint can be used to dump an ascii table of selected row and column values.

7.3 Pixels

This Section is unchanged since Release 6.

Target **pixel** data files contain all the pixels for a target from all cadences, while target **cadence** files contain pixels from all targets for a single cadence. Both raw counts and calibrated flux values are in the pixel data files. The raw counts is the integer value as recorded on the spacecraft. The calibrated pixel value is that provided by the SOC, and is equal to the output of CAL with background and cosmic rays removed.

Target Pixel Data Files are not currently available, but will be in the near future.

7.4 Time and Time Stamps

The primary time stamps available for each cadence in both LC and SC time series are intended to provide proper BJD times corrected to the solar system barycenter and are uniquely determined for each star individually.

Users are urged to read this Section if they have not previously read the Release 6 or Release 7 Notes, as a close reading may help them avoid attempting to do follow-up observations at the wrong time.

7.4.1 Overview

This Section is unchanged since Release 6.

The precision and accuracy of the time assigned to a cadence are limited by the intrinsic precision and accuracy of the hardware and the promptness and reproducibility of the flight software time-stamping process. The Flight System requirement, including both hardware and software contributions, is that the absolute time of the start and end of each cadence is known to ±50 ms. This requirement was developed so that knowledge of astrophysical event times would be limited by the characteristics of the event, rather than the characteristics of the flight system, even for high SNR events.

Several factors must be accounted for before approaching the 50 ms limit:

1. Relate readout time of a pixel to Vehicle Time Code (VTC) recorded for that pixel and cadence in the SSR. The VTC stamp of a Cadence is created within 4 ms after the last pixel of the last frame of the last time slice of that Cadence is read out from the LDE.
2. VTC to UTC of end of cadence, using information provided by the MOC to the DMC to convert between three time systems: 1) vehicle time code (VTC); 2) JPL Ephemeris Time (ET); and 3) Coordinated Universal Time (UTC). These conversions require leap second information and the spacecraft clock correlation.

3. Done by MOC, with precision and accuracy to be documented.
4. Convert UTC to Barycentric JD. This is done in PA (Section 4.3) on a target-by-target basis. The amplitude of the barycentric correction is approximately $(a_K/c)\cos\beta$, where $a_K \sim 1.02$ AU is the semi-major axis of Kepler's approximately circular ($e_K < 0.04$) orbit around the Sun, c the speed of light, and β is the ecliptic latitude of the target. In the case of the center of the Kepler FOV, with β = 65 degrees, the amplitude of the UTC to barycentric correction is approximately +/- 211 s. BJD is later than UTC when Kepler is on the half of its orbit closest to Cygnus (roughly May 1 – Nov 1) and earlier than UTC on the other half of the orbit. This correction is done on a target-by-target basis to support Kepler's 50 ms timing accuracy requirement.
5. Subtract readout time slice offsets (See KIH Section 5.1). This is done in PA (Section 4.3). The magnitude of the time slice offset is $t_{rts} = 0.25 + 0.62(5 - n_{slice})$ s, where n_{slice} is the time slice index (1-5) as described in the KIH. Note that this will in general be different from Quarter to Quarter for the same star, as the star will be on different mod.outs, so the relative timing of events across Quarter boundaries must take this into account.

7.4.2 Time Stamp Definitions

This Section is unchanged since Release 6.

Cadence files:

JD = Julian Date

MJD = Modified Julian Date

MJD = JD - 2400000.5

1. `STARTIME(i)` = MJD of start of i^{th} Cadence
2. `END_TIME` = MJD of end of i^{th} Cadence
3. `MID_TIME(i)` = MJD of middle of Cadence =`(STARTIME(i) + END_TIME(i))/2`
4. JD(i) `= MID_TIME(i) + 2400000.5`

Releases 4-6 light curves, with barycentric and time slice corrections:

1. `timcorr(i)` = dtB(i) - t_{rts}, where dtB(i) = barycentric correction generated by PA, a function of Cadence `MID_TIME(i)` and target position, and t_{rts} is the readout time slice offset described in Section 7.4.1. Units: seconds.
2. BJD(i) = barycentric Julian Date = `timcorr(i)/86400` + JD(i). Units: days

3. `barytime(i)` = Barycentric Reduced Julian Date = BJD(i) – 2400000 = `timcorr(i)/86400` + `MID_TIME(i)` + 0.5. Units: days
4. `LC_START` = MJD of beginning of first Cadence (uncorrected). Units: days
5. `LC_END` = MJD of end of last Cadence (uncorrected). Units: days

Or, as is summarized in the FITS table header:

```
COMMENT barytime(i)- timcorr(i)/86400 - 0.5 = utc mjd(i) for cadence_number(i)
```

Where `utc mjd(i) for cadence_number(i)` is the same as `MID_TIME(i)`

The difference between Release 3 and Releases 4-6

In Release 3, t_{rts} = 0 for all targets, while in Release 4 t_{rts} is calculated as described in Section 7.4.1. That is the only difference.

The vexing matter of the 0.5 days

Users should note that `barytime` follows the same conventions as Julian Date, and astronomers in general; that is, the day begins at noon. MJD, on the other hand, follows the convention of the civil world: that the day begins at midnight. If `timcorr` = 0, then MJD = `barytime` – 0.5 d and `barytime` = MJD + 0.5 d.

7.4.3 Caveats and Uncertainties

Section 7.4.3 is unchanged since Release 6.

Factors which users should consider before basing scientific conclusions on time stamps are:

1. The precise phasing of an individual pixel with respect to the Cadence time stamp (not understood to better than +/- 0.5 s) at this time.
2. General and special relativistic effects in the calculation of the barycentric correction. For example, time dilation at Kepler with respect to a clock at rest with respect to the solar system barycenter, but outside the Sun's gravity well, is 7.5×10^{-9} = 0.23 s/yr – so these effects cannot be dismissed out of hand at this level, and must be shown to be negligible at the level of Kepler's time accuracy requirement of 50 ms or corrected for.
3. The existing corrections have yet to be verified with flight data.
4. Light travel time and relativistic corrections to the user's target, if the target is a component of a binary system.
5. BJD as calculated in this Release is UTC based and high-precision users will want to use BJD in the Dynamical Time standard, which is the preferred absolute time reference for extra-terrestrial phenomena (See Ref. 17 for a thorough discussion).

The advice of the DAWG is not to consider as scientifically significant relative timing variations less than the read time (0.5 s) or absolute timing accuracy better than one frame time (6.5 s) until such time as the stability and accuracy of time stamps can be documented to near the theoretical limit.

7.5 Future Formats Under Discussion

This Section is unchanged since Release 6.

The Science Office recognizes that the MAST products are deficient in several ways, and is working on the following improvements:

1. Provision of an aperture extension, which will tell users which pixels were used to calculate uncorrected flux time series and centroids.
2. Data quality flags, encoding much of the information on lost or degraded data and systematic errors provided in these Notes in a way that will spare users the drudgery of fusing the data in the Supplement with the light curve data.
3. A local WCS coordinate system derived from linearized motion polynomials with Cadence-to-Cadence corrections from the mid-time of the data set. These corrections are a target-specific image motion time series for users to use in their own systematic error correction and are thus an improvement on the mod.out center motion time series provided in the Supplement.
4. Packaging target pixel files as columns of images rather than as pixel lists.

8. References

1. "Initial Assessment Of The Kepler Photometeric Precision," W.J. Borucki, NASA Ames Research Center, J. Jenkins, SETI Institute, and the Kepler Science Team (May 30, 2009)
2. "Kepler's Optical Phase Curve of the Exoplanet HAT-P-7," W. J. Borucki et al., Science Vol 325 7 August 2009 p. 709
3. "Pixel Level Calibration in the Kepler Science Operations Center Pipeline," E. V. Quintana *et al.*, SPIE Astronomical Instrumentation conference, June 2010.
4. "Photometric Analysis in the Kepler Science Operations Center Pipeline," J. D. Twicken *et al.*, SPIE Astronomical Instrumentation conference, June 2010.
5. "Presearch Data Conditioning in the Kepler Science Operations Center Pipeline," J. D. Twicken *et al.*, SPIE Astronomical Instrumentation conference, June 2010.
6. Dave Monet, private communication.
7. "Initial Characteristics of Kepler Long Cadence Data for Detecting Transiting Planets," J. M. Jenkins *et al.*, ApJ Letters **713**, L120-L125 (2010)
8. "Initial Characteristics of Kepler Short Cadence Data," R. L. Gilliland *et al.*, ApJ Letters **713,** L160-163 (2010)
9. "Overview of the Kepler Science Processing Pipeline," Jon M. Jenkins *et al.*, ApJ Letters **713**, L87-L91 (2010)
10. "Discovery and Rossiter-McLaughlin Effect of Exoplanet Kepler-8b," J. M. Jenkins *et al.*, submitted to ApJ http://arxiv.org/abs/1001.0416
11. "Kepler Mission Design, Realized Photometric Performance, and Early Science," D. Koch *et al.*, ApJ Letters **713**, L79-L86 (2010)
12. "Selection, Prioritization, and Characteristics of Kepler Target Stars," N. Batalha *et al.*, ApJ Letters **713**, L109-L114 (2010)
13. "Kepler Science Operations," M. Haas *et al.*, ApJ Letters **713**, L115-L119 (2010)
14. "The Kepler Pixel Response Function," S. Bryson *et al.*, ApJ Letters **713**, L97-L102 (2010)
15. "Instrument Performance in Kepler's First Months," D. Caldwell *et al.*, ApJ Letters **713**, L92-L96 (2010)
16. "Selecting Pixels for Kepler Downlink," S. Bryson *et al.*, SPIE Astronomical Instrumentation conference, June 2010.
17. "Achieving Better Than One-Minute Accuracy In The Heliocentric And Barycentric Julian Dates," Jason Eastman, Robert Siverd, B. Scott Gaudi, submitted to ApJ http://arxiv.org/abs/1005.4415v2

9. List of Acronyms and Abbreviations

ACS	Advanced Camera for Surveys
ADC	Analog to Digital Converter
ADCS	Attitude Determination and Control Subsystem
AED	Ancillary Engineering Data
ARP	Artifact Removal Pixel
BATC	Ball Aerospace & Technologies Corp.
BG	BackGround pixel of interest
BOL	Beginning Of Life
BPF	Band Pass Filter
CAL	Pixel Calibration module
CCD	Charge Coupled Device
CDPP	Combined Differential Photometric Precision
CDS	Correlated Double Sampling
CR	Cosmic Ray
CSCI	Computer Software Configuration Item
CTE	Charge Transfer Efficiency
CTI	Charge Transfer Inefficiency
DAA	Detector Array Assembly
DAP	Data Analysis Program
DAWG	Data Analysis Working Group
DCA	Detector Chip Assembly
DCE	Dust Cover Ejection
DIA	Differential Image Analysis
DMC	Data Management Center
DNL	Differential Non-Linearity of A/D converter
DSN	Deep Space Network
DV	Data Validation module
DVA	Differential Velocity Aberration
ECA	Electronic Component Assembly
EE	Encircled Energy
EOL	End of Life
ETEM	End-To-End Model of Kepler
FFI	Full Field Image
FFL	Field Flattener Lens
FGS	Fine guidance sensor
FOP	Follow-up Observation Program
FOV	Field of View
FPA	Focal Plane Assembly
FPAA	Focal Plane Array Assembly

FSW	Flight Software
GCR	Galactic Cosmic Ray
GO	Guest Observer
GUI	Graphical User Interface
HGA	high-gain antenna
HST	Hubble Space Telescope
HZ	Habitable Zone
I&T	Integration and Test
INL	Integral Non-Linearity of A/D converter
IRNU	Intra-pixel Response Nonuniformity
KACR	Kepler Activity Change Request (for additional data during Commissioning)
KAR	Kepler Anomaly Report
KCB	Kepler Control Box
KDAH	Kepler Data Analysis Handbook
KIC	Kepler Input Catalog
KSOP	Kepler Science OPerations
KTD	Kepler Tech Demo (simulated star field light source)
LC	Long Cadence
LCC	Long Cadence Collateral
LDE	Local Detector Electronics
LGA	low-gain antenna
LOS	Line of Sight
LPS	LDE Power Supply
LUT	look-up table
LV	Launch Vehicle
MAD	Median Absolute Deviation
MAST	Multi-mission Archive at STSci
MJD	Modified Julian Date = JD - 2400000.5
MOC	Mission Operation Center
MORC	Module, Output, Row, Column
NVM	Non-Volatile Memory
OFAD	Optical Field Angle Distortion
PA	Photometric Analysis module
PAD	Photometer Attitude Determination (Pipeline S/W)
PDC	Pre-Search Data Conditioning module
PID	Pipeline instance Identifier (unique number assigned to each run of the Pipeline)
PM	Primary Mirror
PMA	Primary Mirror Assembly
POI	Pixels of Interest

PPA	Photometer Performance Assessment (Pipeline S/W)
ppm	parts per million
PRF	Pixel Response Function
PRNU	Pixel Response Non-Uniformity
PSD	power spectral density
PSF	Point Spread Function
PSP	Participating Scientist Program
PWA	Printed Wiring Assembly
QE	Quantum Efficiency
RC	Reverse Clock
S/C	Spacecraft
S/W	Software
SAO	Smithsonian Astrophysical Observatory
SC	Short Cadence
SCo	Schmidt Corrector
SDA	Science Data Accumulator
SNR	Signal-to-Noise Ratio
SO	Science Office
SOC	Science Operations Center
SOL	Start-of-Line
SSR	Solid State Recorder
SSTVT	Single-String Transit Verification Test
STScI	Space Telescope Science Institute
SVD	Singular Value Decomposition
TAD	Target and Aperture Definition module
TDT	Target Definition Table
TPS	Transiting Planet Search module
TVAC	Thermal Vacuum testing

10. Contents of Supplement

This Section is conceptually unchanged since Release 6 (Q4). The files themselves describe Q5 data.

The Supplement is available as a full package (DataReleaseNotes_08_SupplementFull.tar).

10.1 Pipeline Instance Detail Reports

These files list the Pipeline version and parameters used to process the data, so that the Pipeline results in this Release can be reconstructed precisely at some future time. The file names are:

```
Q5M1_SC_r6.2_ksop568_mpe_asrun_Pipeline_Instance_Detail_Report_100929.txt
Q5M2_SC_r6.2_ksop568_mpe_asrun_Pipeline_Instance_Detail_Report_100929.txt
Q5M3_SC_r6.2_ksop568_mpe_asrun_Pipeline_Instance_Detail_Report_100929.txt
Q5_LC_r6.2_ksop568_CAL+PA_no_mpe_asrun_Pipeline_Instance_Detail_Report_100929.t
xt
Q5_LC_r6.2_ksop568_PA_+PDC_mpe_asrun_Pipeline_Instance_Detail_Report_100929.txt
```

10.2 Thermal and Image Motion Data for Systematic Error Correction

These files are provided so that users can perform their own systematic error correction, if they conclude that the methods used by PDC are not suitable for their targets and scientific goals. It is important to remember that inclusion of additional time series to the cotrending basis set may not improve the results if the cotrending time series are noisy, poorly sampled, or nearly degenerate. The thermal AED will, in general, have to be resampled to match the Cadence times, and on physical grounds it may be more effective to cotrend against bandpass-filtered AED as separate basis vectors. See the SPIE PDC paper (Ref. 5) for a brief discussion of synchronizing ancillary data to mid-Cadence timestamps, and the use of synchronized AED as a cotrending basis set.

10.2.1 Mod.out Central Motion

On rare occasions (<2% of the points), users may notice some “chatter” in the motion time series, which results from a known problem with the motion polynomial fitting algorithm and not actual jumps in telescope attitude or CCD position. A more robust, iterative algorithm has been identified and will be implemented in future Pipeline software to remedy this problem. Users will also clearly see DVA and the signatures of the variable FGS guides stars (Section 6.2) and the reaction wheel heaters (Section 0) in the motion time series.

Files:

`Q5LC_central_column_motion.txt`

`Q5LC_central_row_motion.txt`

Channel central column (row) motion from motion polynomials for all channels, sampled at the Long Cadence period.

Column Descriptions

1. Cadence Interval Number
2. Relative Cadence Index
3. Gap Indicator. 1 = Momentum Dump or Loss of Fine Point
4. Cadence mid-Times, MJD

5-88. Mod.out center column (row) for each channel. Units: pixels. mod.outs are shown

`Q5LC_central_motion.mat` – MATLAB file containing both row and column motion; this will spare MATLAB users the drudgery of parsing the text files.

10.2.2 Average LDE board Temperature

`Q5_LDC_BoardTemp.txt` – average of the ten LDE board temperatures.

Column descriptions:

1. MJD - 55000, units: d, sampling 6.92E-04 d = 59.75 s
2. Average temperature, units: C

10.2.3 Reaction Wheel Housing Temperature

`Q5_TH1RW34T_MJD_gap.txt`– Reaction wheel housing temperature. Data are gapped for desats and median-filtered with a box width = 5 samples.

Column definitions:

1. MJD, units: d, sampling (unfiltered) = 58.0 s
2. TH1RW3T – units: C
3. TH1RW4T – units: C

10.2.4 Launch Vehicle Adapter Temperature

`Q5_TH12LVAT_MJD_gap.txt`-- Launch Vehicle Adapter Temperature. Data are gapped for desats and median-filtered with a box width = 5 samples.

Column definitions:

1. MJD – 55000, units: d, sampling (unfiltered) = 58.0 s
2. TH1LVAT – units: C
3. TH2LVAT – units: C

10.3 Background Time Series

The background time series provide the median calibrated background pixel value on a given mod.out and Cadence. For LC, the background pixels are the dedicated background pixel set. For SC, the background pixels are the target pixels which are not in the optimal aperture. These values are calculated directly from the pixel sets, not from the Pipeline-derived background polynomials.

```
Short Cadence:
Q5M1_SC_background.txt
Q5M2_SC_background.txt
Q5M3_SC_background.txt

Long Cadence:
Q5_LC_background.txt
```

Column definitions

1. Cadence Interval Number
2. Relative Cadence Index for Argabrightening Cadences
3. Gap Indicator. 1 = No Data, Momentum Dump, or Loss of Fine Point
4. Cadence mid-Times, MJD
5. Median background current averaged over FPA, e-/Cadence. All zeros = no SC targets

6-89. Mod.out background in e-/Cadence for each channel. mod.outs are shown

Corresponding MATLAB files are provided to spare MATLAB users the drudgery of parsing the text files.

10.4 Flight System Events

<u>Argabrightening Detections</u>

```
ArgAgg_Q5_LC_PID2817_MADT010_MCT10_Summary.txt

ArgAgg_Q5M1_SC_PID2837_MADT010_MCT10_Summary.txt
ArgAgg_Q5M2_SC_PID2857_MADT010_MCT10_Summary.txt
ArgAgg_Q5M3_SC_PID2858_MADT010_MCT10_Summary.txt
```

Column Definitions:

1. Cadence Interval Number for Argabrightening Cadences
2. Relative Cadence Index for Argabrightening Cadences
3. Arg Cadence mid-Times, MJD
4. Mean SNR over Channels of Arg Event
5. Channels exceeding threshold in Arg Cadence
6. Channels exceeding default threshold in ArgCadence

<u>Out of Fine Point Cadence Lists</u>

`Q5M1_SC_isNotFinePoint.txt`

`Q5M2_SC_isNotFinePoint.txt`

`Q5M3_SC_isNotFinePoint.txt`

`Q5_LC_isNotFinePoint.txt`

10.5 Zero Crossing Events

`Q5_LC_ZeroCrossings.txt`

`Q5_SC_ZeroCrossings.txt`

`Column definitions:`

MJDStart – MJD of start of zero crossing event

MJDEnd – MJD of end of zero crossing event

CINStart – Cadence Interval number at beginning of zero crossing event

CINEnd -– Cadence Interval number at end of zero crossing event

RCIStart – Relative cadence index at beginning of zero crossing event

RCIEnd - Relative cadence index at end of zero crossing event

NumTLMSamp – Number of engineering telemetry sample in the zero crossing event

10.6 Calibration File READMEs

The calibration file names are not listed in the headers of the light curves and target pixel files. The calibration file names listed in the FITS headers of Cadence files and FFIs are not, in general, correct. The README files for the calibration files actually used for all releases to date are:

`kplr2008072318_gain.readme.txt`

`kplr2008102416_read-noise.readme.txt`

`kplr2008102809_undershoot.readme.txt`

`kplr2009060215_linearity_readme.txt`

`kplr2009060615-mmo_2d-black.readme.txt`

`kplr2009062300_lsflat.readme.txt`

`kplr2009062414-MMO_ssflat.readme.txt`

They are supplied with the Release 5 Supplement and not duplicated in the Release 8 Supplement.

10.7 Short Supplement Package

The Supplement also contains a short package suitable for emailing (DataReleaseNotes_08_SupplementSmall.tar). The small package does not contain the following files:

```
Q5M1_background.txt
Q5M2_background.txt
Q5M3_background.txt
Q5_LC_background.txt
Q5LC_central_column_motion.txt
Q5LC_central_row_motion.txt
Q5_TH12LVAT_MJD_gap.txt
Q5_TH1RW34T_MJD_gap.txt
```