Kepler Data Release 4 Notes

KSCI-19044-001

Data Analysis Working Group (DAWG)

Jeffrey Van Cleve, Editor

Data Release 4 for Quarter 3

Quarter .month	Data Type	First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Number of Cadence Intervals
3	LC	55092.7222	55181.9966	9/18/09 17:19	12/16/09 23:55	4370
3.1	SC	55092.7123	55123.0555	9/18/09 17:05	10/19/09 1:19	44550
3.2	SC	55123.9144	55153.9511	10/19/09 21:56	11/18/09 22:49	44100
3.3	SC	55156.0156	55182.0065	11/21/09 0:22	12/17/09 0:09	38160

Including FFIs and Dropped Targets for the Public

Prepared by: _____ Date 4/15/2010
Jeffrey Van Cleve, Kepler Science Office, for the DAWG (next page)

Approved by: _____ Date 4/15/2010
Jon Jenkins, Co-I for Data Analysis & DAWG Lead

Approved by: _____ Date 4/15/10
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), and Science Team (ST) members as listed below:

Jon Jenkins*, Chair

Doug Caldwell*, Co-Chair

Allen, Christopher L.

Batalha*, Natalie

Bryson, Stephen T.

Chandrasekaran, Hema

Clarke, Bruce D.

Cote, Miles T.

Dotson, Jessie L.

Gilliland*, Ron (STSci)

Girouard, Forrest

Haas, Michael R.

Hall, Jennifer

Ibrahim, Khadeejah

Klaus, Todd

Kolodziejczak, Jeff (MSFC)

Li, Jie

McCauliff, Sean D.

Middour, Christopher K.

Pletcher, David L.

Quintana, Elisa V.

Tenenbaum, Peter G.

Twicken, Joe

Uddin, Akm Kamal

Van Cleve, Jeffrey

Wohler, Bill

Wu, Hayley Y.

*Science Team

Affiliations are Kepler Science Office or Science Operations Center unless otherwise noted.

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

1. Introduction 7
2. Release Description 8
2.1 Summary of Contents 9
2.2 Pipeline Changes Since Previous Release 9
2.2.1 Change to CAL: calibrated pixels 9
2.2.2 Change to PA: raw light curves and centroids 10
2.2.3 Changes to PDC: corrected light curves 10
3. Current Evaluation of Performance 11
3.1 Overall 11
3.2 Changes in Performance Since Previous Release 13
3.3 Ongoing Calibration Issues 13
4. Data Delivered – Processing History 14
4.1 Overview 14
4.2 Pixel-Level Calibration (CAL) 15
4.3 Photometric Analysis (PA) 15
4.4 Pre-Search Data Conditioning (PDC) 17
4.4.1 Description 17
4.4.2 Performance 18
4.4.3 Removal of Astrophysical Signatures 22
5. Lost or Degraded Data 24
5.1 Momentum Desaturation 24
5.2 Reaction Wheel Zero Crossings 25
5.3 Data Anomalies 27
5.3.1 Safe Mode 27
5.3.2 Loss of Fine Point 27
5.3.3 Pointing Drift and Attitude Tweaks 28
5.3.4 Downlink Earth Point 29
5.3.5 Manually Excluded Cadences 29
5.3.6 Anomaly Summary Table 29
6. Systematic Errors 30
6.1 Argabrightening 30
6.2 Variable FGS Guide Stars 33
6.3 Pixel Sensitivity Dropouts 34
6.4 Focus Drift and Jitter 35

6.5 Requantization Gaps 38
6.6 Spurious Frequencies in SC Data with Spacing of 1/LC 39
6.7 Known Erroneous FITS header keywords 40
7. Data Delivered – Format 41
7.1 FFI 41
7.2 Light Curves 41
7.3 Pixels 42
7.4 Time and Time Stamps 42
7.4.1 Overview 42
7.4.2 Time Stamp Definitions for Release 4 43
7.4.3 Caveats and Uncertainties 44
8. References 45
9. List of Acronyms and Abbreviations 46
10. Contents of Supplement 49
11. Calibration Files and Parameters 50

1. Introduction

These notes have been prepared to give Kepler users of the Multimission Archive at STScI (MAST) a summary of how the data were collected and prepared, and how well the data processing pipeline is functioning on flight data. They will be updated for each release of data to the public archive and placed on MAST along with other Kepler documentation, at http://archive.stsci.edu/kepler/documents.html . They 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. The user should refer to the KIH for a glossary and acronym list.
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. The initial release of the KDAH, focused on the needs of MAST users and Guest Observers (GOs), is expected on 6/15/10. A significantly expanded version will be released in the fall of 2010, which will include topics which are specific to planetary transit detection and validation, or of general interest but still under development (such as Differential Image Analysis, or DIA).
3. **SOC Tools User's Guide** describes the data and model retrieval tools written by the SOC, which access intermediate data products which are not available through MAST. This User's Guide will only be available to and useful for Science Team members who have access to SOC workstations.
4. **Kepler Archive Manual** (KDMC-10008) describes file formats and the availability of data through MAST. The Archive Manual is available on MAST

Users unfamiliar with the data processing pipeline should read Section 4 first. 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 will indicate if it is "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 which would be unwieldy to print in this document format are included in a tar file, the Data Release Notes #4 Supplement, which will be 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. Supplement files are either ASCII or FITS format.

Note on dates and cadence numbers in these Notes: Cadences are absolutely 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). 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 Q3 unless otherwise noted as historical.

2. Release Description

A *data description* refers to the data type and observation interval during which the data were collected. The observation interval 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. The *data processing* descriptor is the internal Kepler Science Operations (KSOP) ticket used to request 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 4 has the release label "SOC Pipeline 6.1" and the revision number r36013. Unreleased software will, in general, have only a revision number for identification.

The same data will, in general, be reprocessed as the software improves, and will hence be the subject of multiple releases. The combination of data and data processing descriptions 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." See Data Release 2, KSCI-19042, for an example of reprocessed data.

The keywords `DATA_REL` = `4` has been added to the FITS headers so users can unambiguously associate Release 4 FITS files with these Notes. In addition, the keyword `QUARTER` has been added; in the case of Release 4, all light curve files have `QUARTER` = 3, though multiple quarters may be included in future releases.

Data Release 4 was produced with released code, with formal verification and validation of the pipeline and the resulting data products. The resulting data products are a substantial improvement over that available using the pipeline code that was developed before launch, and described in the Data Release 2 Notes (KSCI-19042). 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 is expected to be slower in the second year of the mission than in the first, with major upgrades on a roughly annual basis.

Data Release 4 gives users the opportunity to examine the data for possibly interesting science and to involve the users in improving the pipeline for future data releases. To perform the latter service, users are encouraged to notice and document artifacts, either in the raw or processed data, and report them to the Science Office.

Before users publish results based on Data Release 4, they must understand the data analysis performed and ascertain its impact on their results. The best resource to aid in this assessment is these Data Release Notes. The Science Office advises against publication of these Release 4 light curves without such careful consideration by the end user and dialog with the Science Office and Guest Observer Office, where deemed appropriate.

2.1 Summary of Contents

Table 1: Contents of Release 4. CIN is the cadence interval number described in Section 1. All Release 4 cadence data were processed under KSOP-400 with SOC Pipeline 6.1, revision number r36013, and are released for the first time. The Pipeline Instance ID (PID) for CAL, PA, and PDC is shown. The Table on the cover page is a subset of this Table. Refer to Section 7.4 for a discussion of time and time stamps.

Quarter .month		CAL PID	PA PID	PDC PID	First Cadence MJD midTime	Last Cadence MJD midTime	CIN Start	CIN end	Num CINs
3	LC	1109	1109	1109	55092.7222	55181.9966	7404	11773	4370
3.1	SC	1131	1131	1131	55092.7123	55123.0555	210580	255129	44550
3.2	SC	1131	1131	1131	55123.9144	55153.9511	256390	300489	44100
3.3	SC	1137	1137	1137	55156.0156	55182.0065	303520	341679	38160

Notes:

The FFIs use the same integration parameters as the Long Cadence science data. 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 cadence files in a given Pipeline revision. However, the FFIs are processed on a monthly basis as they are received, and have not been reprocessed uniformly using the same Pipeline revision as was used for the Release 4 Q3 cadence data. No formal validation of the operation of the Pipeline on FFIs has been performed. Thus, these images are good representations of what targets and their neighborhoods look like, but the photometric accuracy of the images has not yet been investigated by the DAWG.

The FFI files collected in Q3 are:

`kplr2009292020429_ffi-SocCal.fits`

`kplr2009322233047_ffi-SocCal.fits`

`kplr2009351005245_ffi-SocCal.fits`

These FFIs are available from the MAST public ftp site: http://archive.stsci.edu/pub/kepler/ffi/

2.2 Pipeline Changes Since Previous Release

Experience with real flight data has led to substantial improvements to the Pipeline for Release 4. In this Section, the 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 Change to CAL: calibrated pixels

Changes to CAL were relatively minor between Release 3 and Release 4. The only change of note was to improve the identification of bleeding columns in the virtual or masked smear collateral data, and compensate for undershoot in the subsequent columns. This was an important change, as badly corrected collateral data affect all targets which contain the afflicted column, and the errors are highly correlated. However, in cases where the bleeding smear value exceeds the 23 bits of the Science Data Accumulator, and the bleeding star is also variable, the difference between the masked and virtual smear will toggle between positive and negative, and

CAL will produce incorrect results. 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.

2.2.2 Change to PA: raw light curves and centroids

There were no significant changes to PA itself between Release 3 and Release 4. However, users will be pleased to learn that the MAST light curves now contain only flux-weighted (first moment) centroids. In Release 3, the Pipeline exported PRF centroids if available, or flux-weighted centroids if PRF centroids were not available. This caused confusion to archive users, because there was no outward indication of centroid type, which could vary from cadence to cadence.

2.2.3 Changes to PDC: corrected light curves

1. The EARTH_POINT anomaly flag was added after it became clear that there could be large thermal transients after monthly Ka-band downlinks (in addition to large transients after the safe modes).
2. PDC substitutes the raw flux for the cotrended flux when the cotrended flux has a higher noise than the raw flux. The difference between raw and corrected flux light curves in this case will be a constant estimate of excess flux due to crowding in the optimal aperture, as well as gaps identified in PDC which were not present in the PA output. In Release 3, the PDC substituted the coarsely detrended flux. At present, there is no light curve FITS header keyword indicating that this has been done, though users can difference the ‘raw’ and ‘corrected’ light curves to check.
3. PDC performs the systematic error correction twice for the targets that are initially identified as "variable". In Release 3, many of these stars were identified as variable because of the large flux level changes due to the spacecraft/data anomalies (Section 5.3), not because the stellar flux actually varied. If the corrected flux has lower noise than the raw flux but is not variable, then the target is treated as if it was not initially variable (i.e. no harmonics are separated out before systematic error correction).
4. In Release 4 there is an attempt to preserve transits in the vicinity of the known spacecraft/data anomalies, which may lead to insufficiently corrected discontinuities in some cases (Figure 7).
5. There were substantial changes to the coarse detrending function which determines the flux to which the harmonic content is actually fit.

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 gives 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 4.1 σ.

The CDPP performance has been discussed by Borucki et al. [2] and Jenkins et al. [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, indicating that Kepler has the capability to fulfill its mission objectives. The Q3 results 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.
2. Given the instrument artifacts discussed in detail in the KIH, 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 other instrumental effects are not, in general, white noise processes.

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

The Photometer Performance Assessment (PPA) tool formally calculates CDPP on 6 hr timescales as a function of cadence *c* for each target *k*. The temporal median of the CDPP for each target is $TMCDPP_k = M_{\{c\}}(CDPP_{ck})$, where $M_{\{c\}}$ denotes the median over the set of cadences {c}, which in this case is all the Cadences of a Quarter. $TMCDPP_k$ is then divided by sqrt(13/12) to approximate the results on the 6.5 hr benchmark time scale. The results (Figure 1) separate into two branches, mostly corresponding to giants with log g < 4 and dwarfs with log g > 4.

$TMCDPP_k$ can be further summarized by aggregate statistics, which are calculated over all targets which are members of a set *K* satisfying certain criteria, such as the aggregate median $CDPP_K = M_{\{k \in K\}}(TMCDPP_k)$, where $M_{\{k \in K\}}$ indicates the median calculated over all targets *k* satisfying the criteria for membership in *K*. For example, *K* can be {all targets with magnitude between 11.75 and 12.25 and log g > 4}, loosely referred to as "12th magnitude dwarfs." The aggregate median CDPP for 12th magnitude dwarfs is 39.1 ppm in Q3. Aggregate percentiles can be defined in the same way. The aggregate 10th percentile CDPP for 12th magnitude dwarfs is 24.6 ppm. It is believed that the 10th percentile value is a better representation of instrument and Pipeline performance than the median value, since even dwarf stars can show significant variability when observed with Kepler's unprecedented precision. Thus, the aggregate 10th percentile CDPP on 6.5 hr timescales for 12th magnitude dwarfs is ~23% higher than the requirement of 20 ppm for 12th magnitude stars. Table 2 summarizes the results at other magnitudes. Note that the aggregate 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. [7] expression for the lower noise envelope for stars brighter than 14th magnitude may be thought of as the RSS sum of shot noise and an effective read noise corresponding to a source extracted from 2.6 pixels (given a per-frame, per pixel read noise of 100 e-). Extending this expression to the benchmark 6.5 hr transit time gives the results shown in Figure 1 and Table 2.

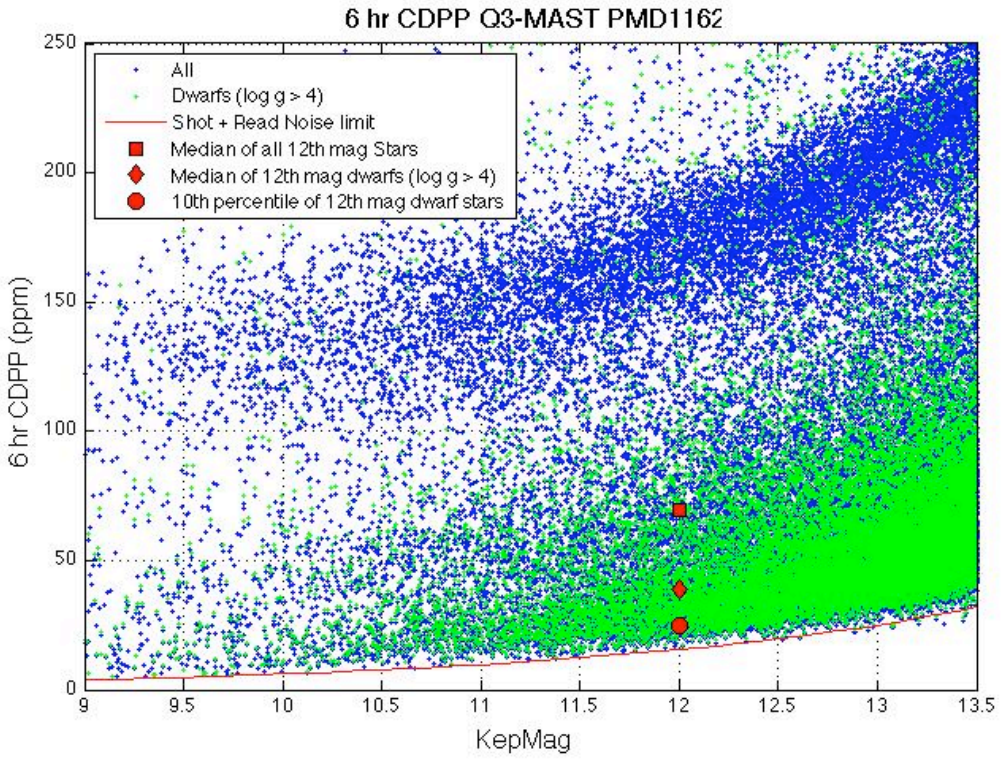

Figure 1: 6.5 hr Temporal Median (TM) of the per-Cadence CDPP calculated for Quarter 3 by PPA for stars between 9th and 13.5th magnitude. The 6 hr TMCDPPs have been divided by sqrt(13/12) = 1.041 to approximate 6.5 hr TMCDPPs. Stars on the planetary target list with Kepler Magnitude < 13.5 and log g > 4, which are almost certainly dwarf stars, are shown as green +'s; other stars are marked with blue +'s. The red line is the CDPP calculated from a simple shot and effective read noise model derived from Jenkins et al. [7].

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) Simplified noise model CDPP (9) Percentage of all observed stars with TMCDPP < noise model. 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	simple Model CDPP	% below model
9	29	9.9	30.5	224	11.9	76.3	3.8	0.0
10	169	11.2	36.0	711	15.2	109.8	6.0	0.0

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	simple Model CDPP	% below model
11	649	18.1	33.7	2004	21.6	103.9	9.5	0.1
12	2304	24.6	39.1	4812	27.2	69.5	15.2	0.0
13	7232	36.0	51.7	11500	38.1	64.5	24.4	0.1

3.2 Changes in Performance Since Previous Release

The Pipeline has not undergone qualitative changes in its algorithms since Data Release 3 for Q2 data. However, we expected considerable improvement in performance in the Q3 data for Data Release 4, since no pointing tweaks (Section 5.3.3) were required in Q3 and the focus jitter correlated with the reaction wheel heater (Section 6.4) was substantially diminished by operational changes. Both improved 12th magnitude dwarf CDPP and centroid stability were observed, though the cause of the improved CDPP has not been proven by analysis.

3.3 Ongoing Calibration Issues

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

1. Identify residual instrumental effects in fluxes and centroids in science data which are not yet correlated with ancillary data. The existence of uncorrected systematic noise is revealed by strong target to target correlations in the PDC-corrected flux.
2. Identify additional flight system state variables, which can be used to cotrend the data in PDC. Currently, only motion polynomials and the 10 LDE board temperatures are used. Other possible variables include the reaction wheel housing temperatures associated with plate scale variations (Section 6.4). The same algorithm would be used.
3. 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.
4. Characterize the in-orbit change of focus (Section 6.4).
5. Identify particular light curves that are poorly corrected, and understand why generally effective remedies do not work in these cases. Feedback from users is essential for the SO and SOC to identify, flag, and fix all such “hard cases.”
6. Mitigate or at least identify the Artifacts described in the KIH, Section 6.7.
7. Assess and improve the focal plane characterization models which are inputs to CAL.

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

The delivered FITS files were processed as shown in simplified form in Figure 2. What is referred to as "raw" flux time series is the result of calibrating pixels, estimating and removing sky background, and extracting a time series from a photometric aperture. 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."

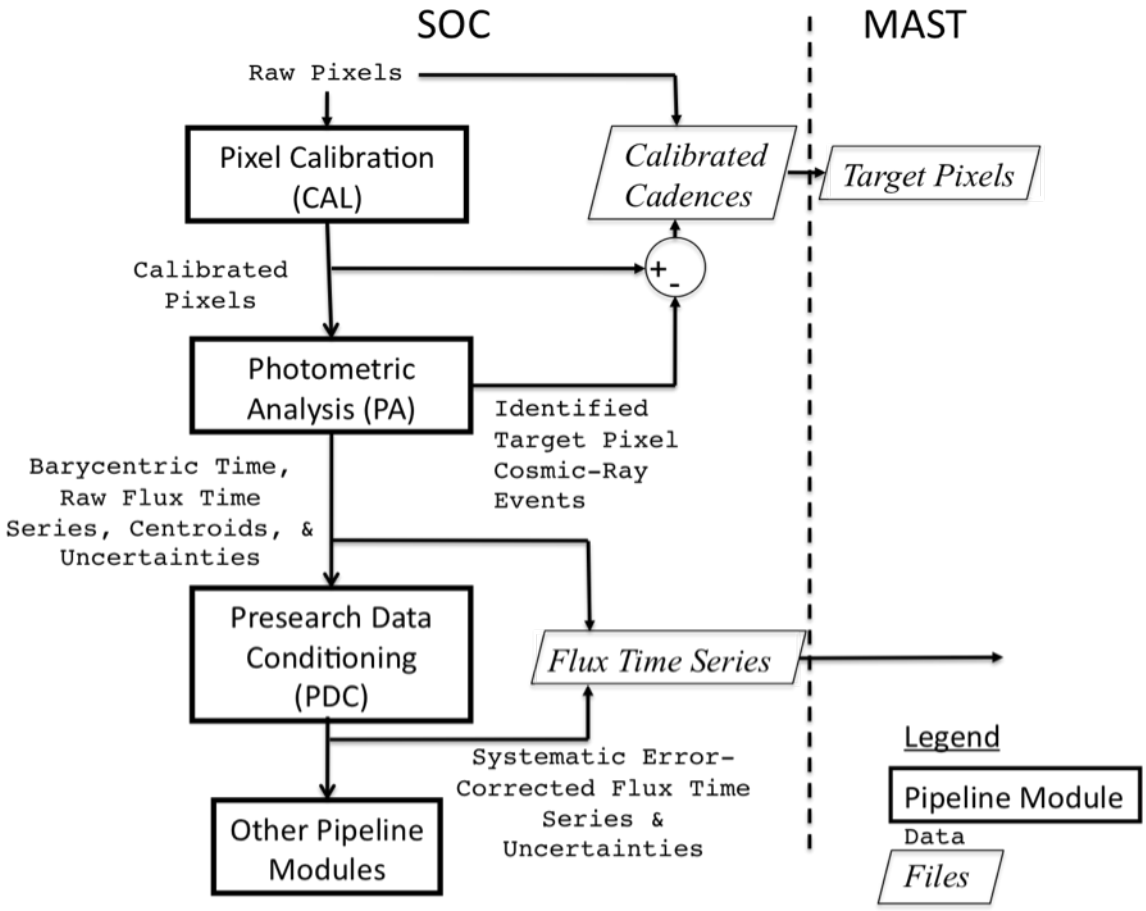

Figure 2: Processing of data from raw pixels to flux time series and target pixel files archived at MAST. The target pixel files generated at MAST from the calibrated cadence files delivered by the SOC have identified cosmic-ray events removed. The corrected flux time series delivered to MAST contain stellar variability and have -Infs for bad or missing data; conversely, the corrected light curves used internally in the SOC for detecting planets have stellar variability removed, and have bad or missing data filled by an autoregressive (AR) algorithm. Stellar variability identified by the Pipeline includes

impulsive outliers and harmonics. See Section 7 and the MAST Kepler Archive Manual for details of MAST file contents.

4.2 Pixel-Level Calibration (CAL)

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].

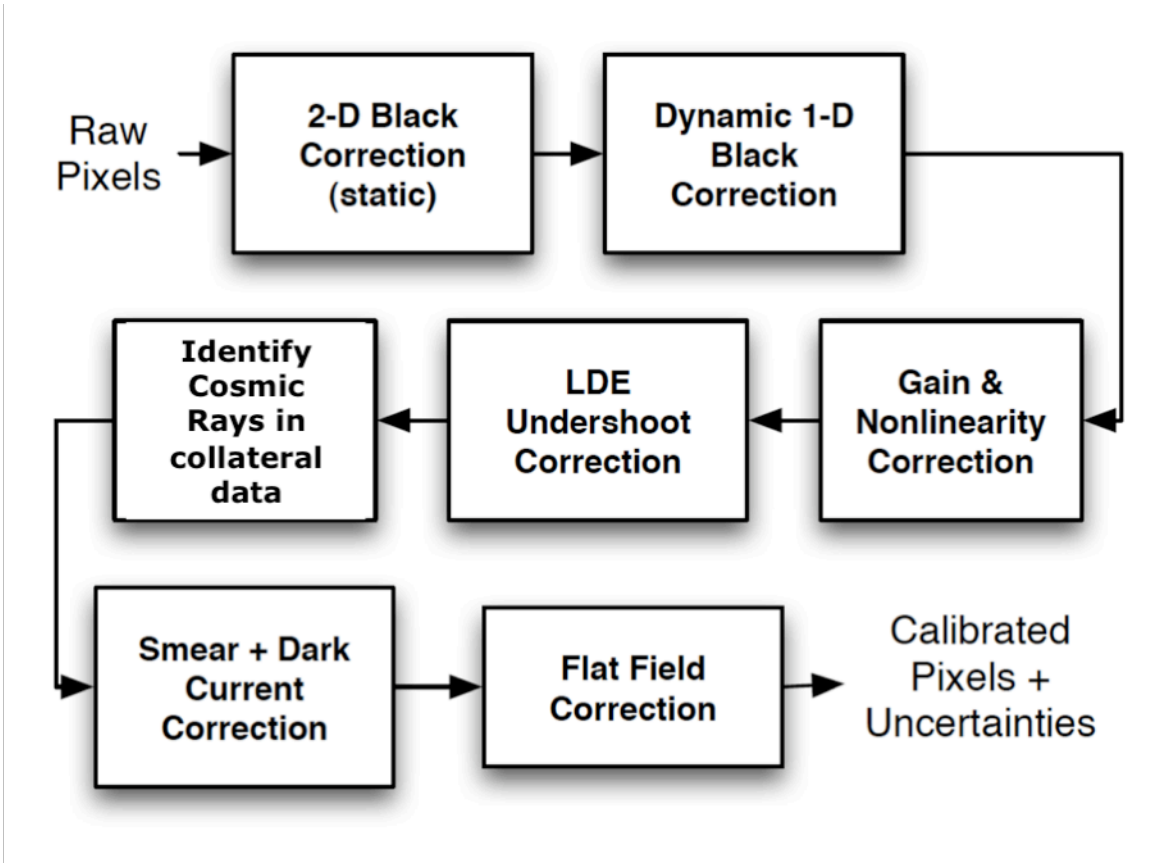

Figure 3: Pixel Level Calibrations Performed in CAL. See the Instrument Handbook for a discussion of signal features and image contents processed in CAL.

4.3 Photometric Analysis (PA)

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). Argabrightening detection and associated cadence gapping take place first; 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. Data which are greater than 12 median absolute deviations (MAD), after the removal of a trend formed by a quadratic fit followed by a 5 cadence wide median filter, are identified as CRs. The MAD is calculated over a sliding window 145 cadences wide. When fluctuations within the sliding window are larger, the CR threshold is higher. Otherwise, large random fluctuations will trigger the CR detector in periods when the flux is particularly noisy. As a result, the CR cleaning process is more sensitive when the flux is particularly quiet. 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.
4. Robust 2-D polynomial fitting to calibrated background pixels
5. Background removal from calibrated target pixels
6. Aperture photometry. In release 4, 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. See note below.
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. 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.
9. Setting gap indicators for cadences with Argabrightening (Section 6). The gapped cadences have all –Inf values in the FITS light curve files, except for the first two columns: time and cadence number.

Notes

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. While only flux-weighted centroids are exported to MAST in Release 4, 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.

There is no identification of bad pixels in PA in Release 4, 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.

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.

4.4 Pre-Search Data Conditioning (PDC)

The primary tasks of PDC are to correct systematic errors, remove excess flux in target apertures due to crowding, identify outliers, and fill gaps in raw flux light curves for long and short cadence targets. PDC was designed to remove systematic errors that are correlated with ancillary engineering or Pipeline generated data (such as motion polynomials), and also to condition long cadence light curves for the transiting planet search (TPS). 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 emphasizing transit signals. 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.

4.4.1 Description

The tasks performed by PDC are:

1. Accept data anomaly flags set by the Pipeline operator, for cadences which 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 ancillary spacecraft data 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. Identify variable stars (>0.5% center-peak variability).
5. For variable stars only, 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 which 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, and save this improved harmonic fit for later restoration.
    f. Remove the improved fit and pass the harmonic-removed light curve to step 6.
6. Cotrending target flux time series against ancillary data 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. The only ancillary data used for Release 4 are the 10 LDE board temperature. While the reaction wheel housing temperatures correlated with flux time series (Section 6.4) were not used for cotrending in Release 4, the motion polynomials contain similar information, and it remains to be seen whether explicit inclusion of the reaction wheel housing temperatures in the ancillary data set used for cotrending will actually improve matters.
7. Assess results of cotrending. If cotrending has increased the noise, restore the uncorrected light curve at this point.
8. For stars initially identified as variable, if cotrending has reduced the noise but the cotrended result is not variable, then the identification of the star as variable is considered mistaken, and the star is reprocessed as a quiet star (ie no harmonics are separated out before systematic error correction).
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 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 are restored to the MAST light curves.

Notes

The crowding metric is the fraction of starlight in an aperture which 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 before the correct light curve for the target is obtained. Note that the 'raw' flux time series are *not* corrected for crowding. The crowding metric is based in the Kepler Input Catalog (KIC) star locations and brightnesses, and 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.

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. This gap-filling will be described in the Data Analysis Handbook.

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

PDC gives satisfactory results on most stars which are either intrinsically quiet (Figure 4), or have well-defined harmonic light curves above the detection threshold (Figure 5); 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 6). However, sometimes discontinuities or other data anomalies mimic the astrophysical events that PDC is tuned to protect, and incomplete removal of the data anomaly results (Figure 7).

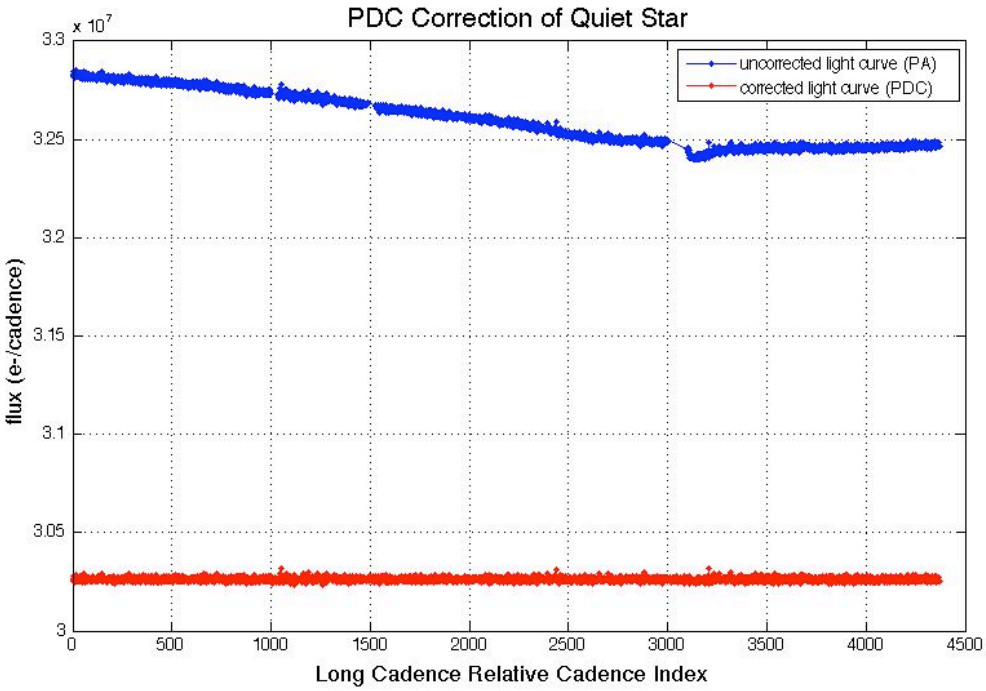

Figure 4: PDC removal of slopes and discontinuities from the light curve of a quiet star of Kepler magnitude 14.6. The noise in the corrected light curve is only 20% greater than the noise expected from the calibrated pixels, a considerable improvement over the uncorrected light curve. The gap between RCI = 3000 and 3100 is a Safe Mode event (Section 5.3.1) followed by a monthly download and subsequent thermal transient. Data from Q3.

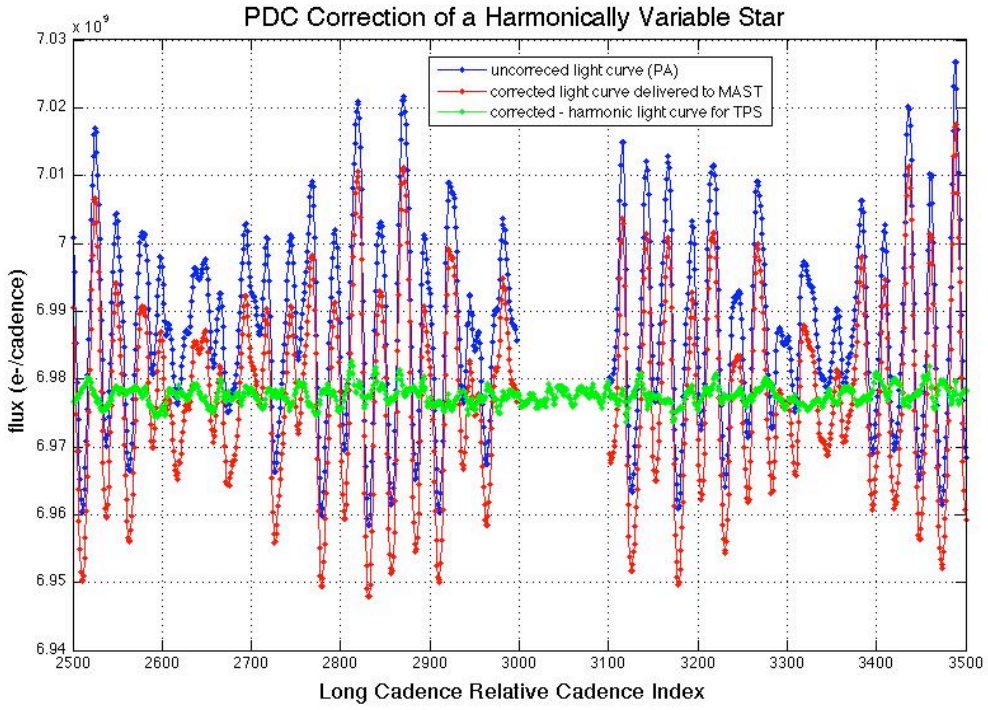

Figure 5: PDC correction of a harmonically variable star. MAST users receive the light curve corrected for systematic errors, with the harmonic variability restored and gaps in the data represented by -lnfs, as shown in red in this Figure. The green curve shows the light curve with the harmonic term is removed and gaps filled by an autoregressive (AR) algorithm. The standard deviation of the non-harmonic part of the corrected light curve is 240 ppm, much less than the ~2000 ppm standard deviation of uncorrected light curve. The gap between RCI = 3000 and 3100 in the uncorrected light curve is a Safe Mode event (Section 5.3.1) followed by a monthly download and subsequent thermal transient. Data from Q3.

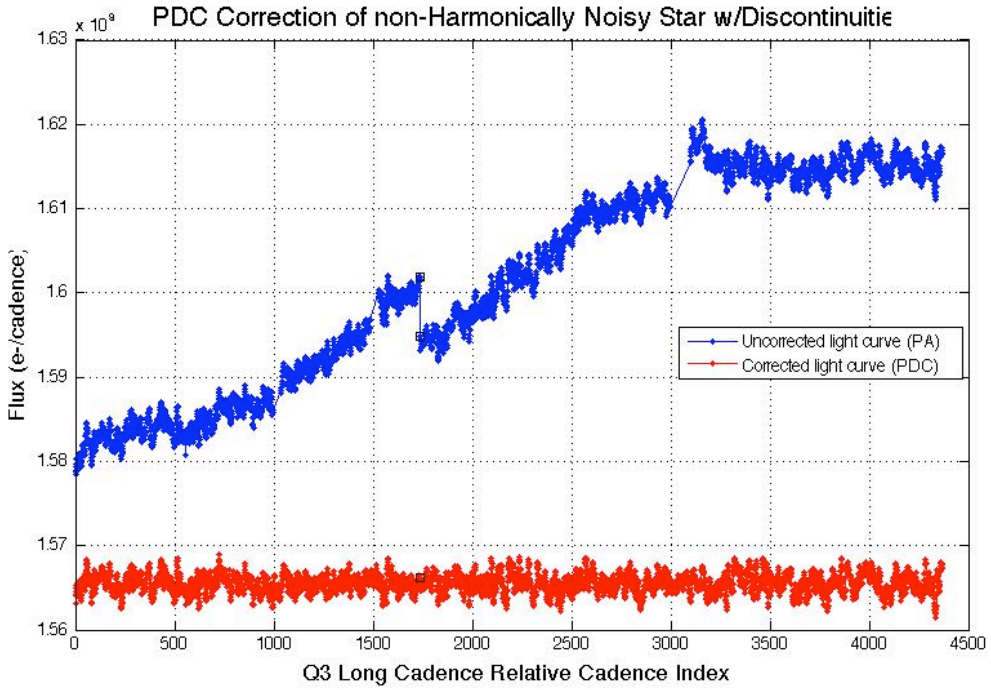

Figure 6: PDC correction of a non-harmonically variable star light curve with discontinuities. The RMS stellar variability in the corrected (red) curve is about 8x the noise calculated from the read and shot noise in the calibrated pixels. Data from Q3.

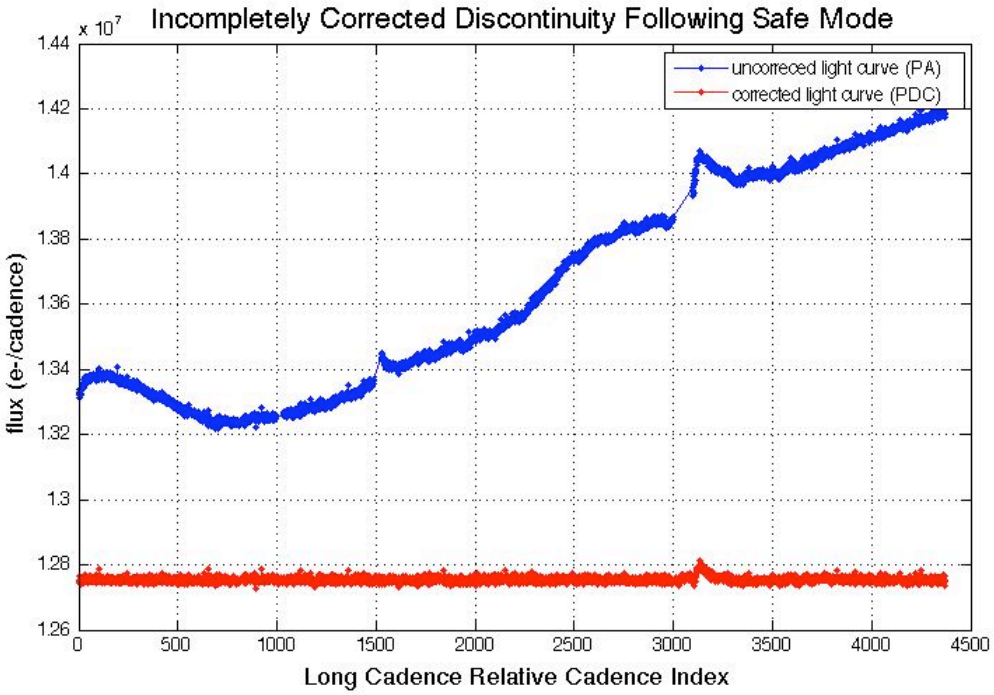

Figure 7: Incompletely corrected discontinuity following Safe Mode. PDC is attempting to preserve possible astrophysical events close to this discontinuity. While the Earth Point

at RCI = 1500 has been fully corrected, the signature after the Safe Mode between RCI 3000 and 3100 looks enough like an astrophysical event that it has been only incompletely corrected. It is expected that PDC parameters will be tuned in future Releases to strike a better balance between real event preservation and data anomaly removal. Data from Q3.

4.4.3 Removal of Astrophysical Signatures

PDC can remove astrophysical signatures if they are:

1. Harmonic, but have periods > 5d and fall below the detection threshold for stellar variability (Figure 8). In Release 4, 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 9)
3. More or less linear ramps over the processing interval.
4. Harmonic signals above the threshold, but the harmonic fit does not succeed or succeeds but is not a good fit, and current algorithm does not correctly recognize that cotrending has performed badly.
5. Non-harmonic signals for which current algorithm does not correctly 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.

A thorough study of astrophysical signal distortion by PDC, in which test patterns are injected into PA outputs and the PDC results compared to the inputs, has not been performed to date.

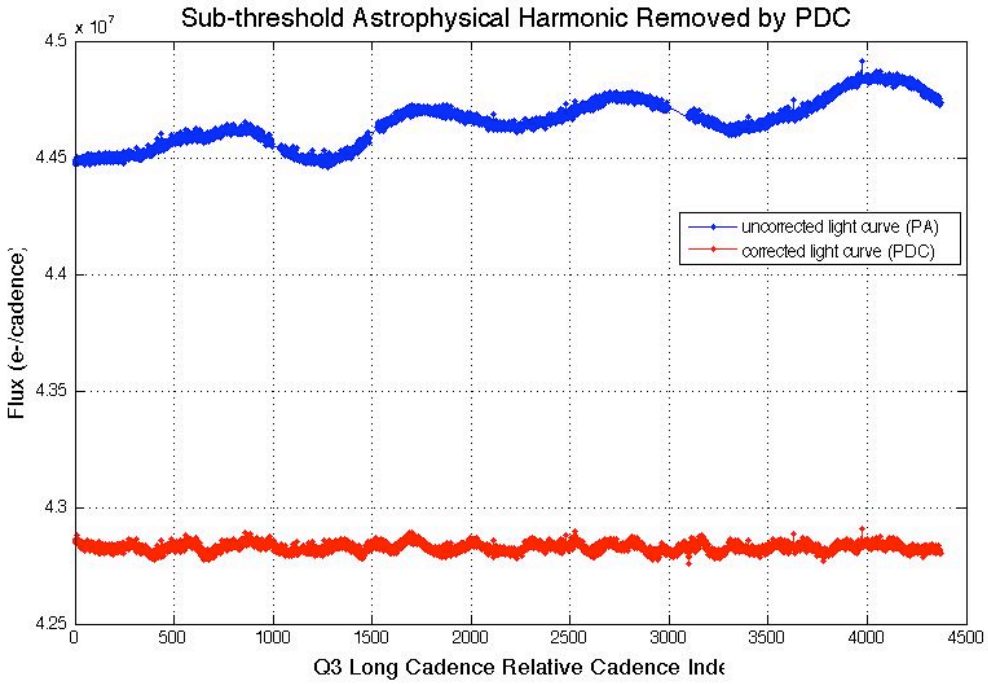

Figure 8: PDC removal of harmonic stellar variability below the harmonic variability detection threshold. If the harmonic variability exceeds the threshold, it is set aside before cotrending, then restored to the MAST corrected light curve.

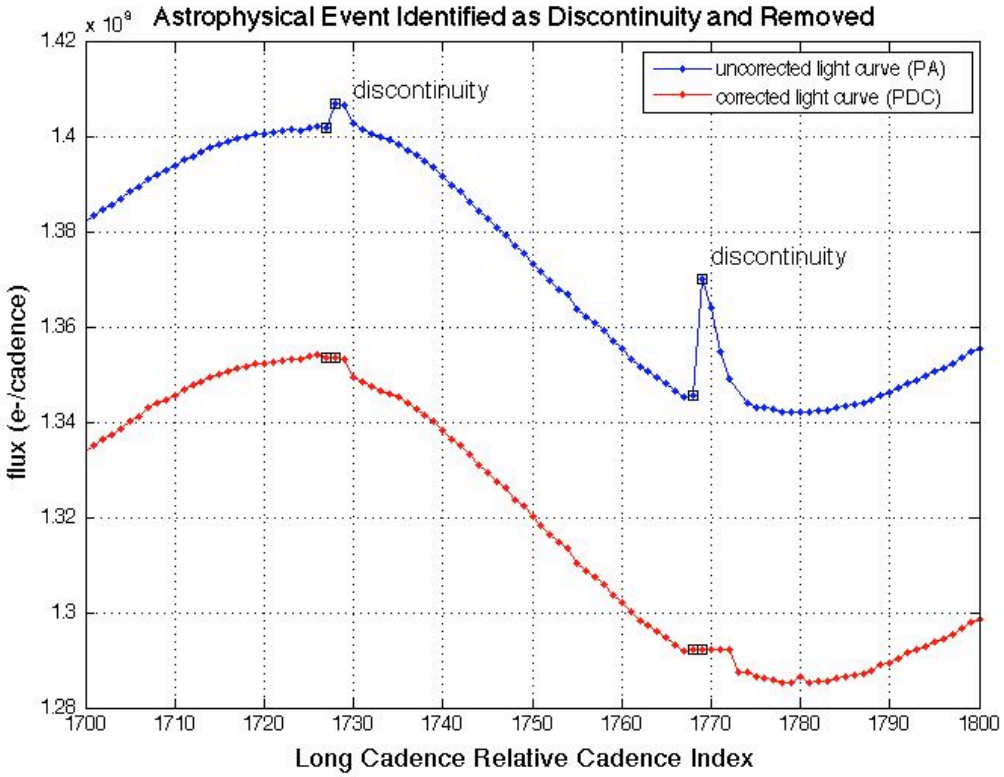

Figure 9: Astrophysical Event, possibly a flare, identified by PDC and removed from the corrected light curve. Open black squares show astrophysical events identified as discontinuity anomalies and "corrected." Data from Q3.

5. Lost or Degraded Data

In this Section, we discuss cadences which 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 SOC change requests (KSOCs). While KARs and KSOCs are not available to users, the KAR number is listed at the end of each Section for Science Office internal reference.

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 3.0 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 Q3 and the corresponding Long Cadences. CIN = Cadence Interval Number, RCI = Relative Cadence Index.

```
Kepler Data Analysis Handbook Supplement
MomentumDump Summary
Prepared by the DAWG
File Name: Q3_LC_MomentumDump.txt
This file created: 05-Feb-2010 17:30:05
MJD 55232.72922
List of Cadences during Momentum Dumps
CIN    RCI     Date(MJD)
  7465     62 55093.96866
  7611    208 55096.95196
  7756    353 55099.91484
  7902    499 55102.89814
  8048    645 55105.88145
  8194    791 55108.86475
  8340    937 55111.84806
  8486   1083 55114.83136
  8632   1229 55117.81467
  8778   1375 55120.79797
  8885   1482 55122.98437
  9030   1627 55125.94724
  9176   1773 55128.93054
  9322   1919 55131.91385
  9468   2065 55134.89715
  9614   2211 55137.88046
  9760   2357 55140.86376
  9906   2503 55143.84707
 10052   2649 55146.83037
 10198   2795 55149.81368
```

10344	2941	55152.79698
10547	3144	55156.94501
10693	3290	55159.92831
10840	3437	55162.93205
10985	3582	55165.89492
11131	3728	55168.87823
11276	3873	55171.84110
11423	4020	55174.84484
11569	4166	55177.82814
11714	4311	55180.79102
11770	4367	55181.93530

5.2 Reaction Wheel Zero Crossings

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 targets showing some impact to the centroid and pixel time series. In some cases, we observe negative spikes of order 10^{-3} to 10^{-2} in SC relative flux time series (Figure 10), and these cadences must be excluded from further analysis. The impact on long cadence data is much less severe in both amplitude and prevalence.

In Figure 10, 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), at MJD – 55000 = 102.90, 105.88, and 108.86, 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 in Release 4, the zero crossing events in Q3 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.

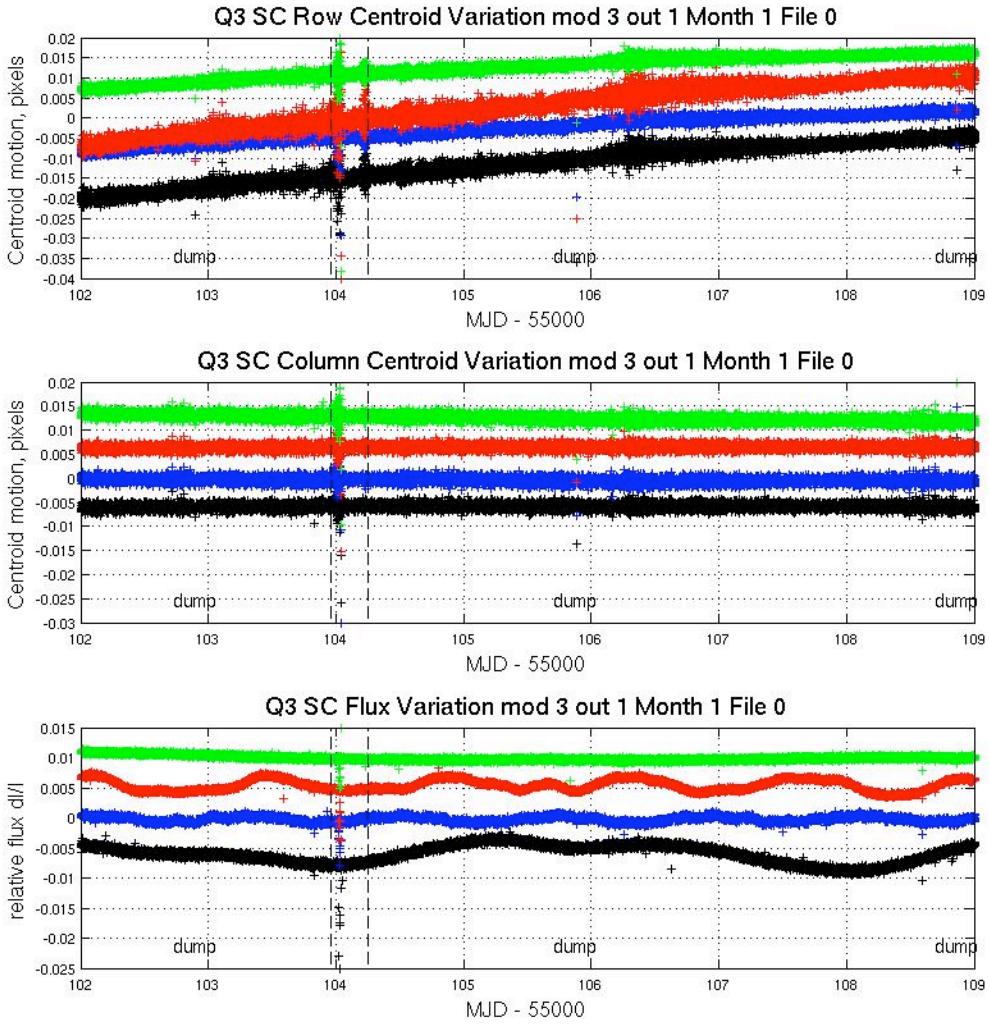

Figure 10: Example from Q3 of the effect of reaction wheel speed zero crossing on SC flux and centroids. The plots show row and column centroid motion, and the relative flux change, in the neighborhood of zero crossings. The data on several stars are overplotted in different colors in each panel of the Figure; vertical dashed black lines bracket the times during which at least one wheel had zero speed according to its telemery. The curves are offset for clarity, and momentum dumps are labeled.

Table 4: Zero crossing events in Q3, defined as the time from first to last zero crossing in the event, rounded to the nearest cadence. This Table is included in the Supplement as file `Q3_ZeroCrossings.txt`.

Kepler Data Release Notes 4 Supplement
Zero Crossing Listing
Prepared by the DAWG
This file created: 26-Jan-2010
MJD 55222.50642

Column Definitions			Cadence Numbers			
	LC	midTime MJD	Long Cadence		Short Cadence	
Event#	Start	End	Start	End	Start	End
1	55103.961	55104.247	7954	7968	227112	227504
2	55121.002	55121.166	8788	8796	252131	252346
3	55131.914	55131.996	9322	9326	268135	268228
4	55134.877	55134.918	9467	9469	272512	272540
5	55137.860	55137.901	9613	9615	276892	276924
6	55140.864	55140.884	9760	9761	281275	281302
7	55152.777	55152.838	10343	10346	298788	298841
8	55152.838	55152.920	10346	10350	298858	298947
9	55165.874	55165.956	10984	10988	318020	318111
10	55168.878	55168.960	11131	11135	322405	322498
11	55174.988	55175.029	11430	11432	331397	331414
12	55175.008	55175.376	11431	11449	331426	331920
13	55177.808	55177.889	11568	11572	335538	335631

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. While each individual event is unexpected, it is not unusual for newly-commissioned spacecraft to experience them, until the in-orbit idiosyncrasies of the flight system are understood. In Quarter 3, there was a Safe Mode event between MJD 55154.0 and 55156.0, as shown in Table 5. The LDE was turned off, but data previously collected remained in the Solid State Recorder (SSR) for retrieval after Safe Mode recovery. Data collected after resumption of science data collection shows a trend strongly correlated with the warmup of the LDE boards, which is for the most part mitigated by PDC (Section 4.4). Kepler's Flight Software has been modified to leave the LDE on during radiation-induced resets of the RAD750 processor, so that data degradation due to thermal transients after an LDE power cycle does not occur during this kind of Safe Modes.

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. The cause of these Loss of Fine Point (LOFP) events is presently under investigation. 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, shown in Table 5. In Quarter 3, one LOFP event between MJD = 55099.91 and 55100.04 was automatically self-corrected by Flight Software, with a loss of only 7 LCs, while the event between MJD 55113.05 and 55113.83 required ground commanding, and led to a loss of 39 LCs to high-precision photometry. Recent Flight Software upgrades are expected to reduce the frequency and duration of LOFPs.

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 3-4 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 residuals (RA, Dec, roll) for Q3 are shown in Figure 11. 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. 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 MAR is always < 100 mpix. In Q2, tweaks were necessary, and they introduced discontinuities into the data which were not fully compensated by the initial version of the Pipeline. A change to the FGS centroiding algorithm implemented at the start of Q3 has apparently greatly diminished the boresight drift. There were no tweaks in Q3 since the drift was an order of magnitude less than in Q2, and MAR remained between 23 and 38 mpix for the entire quarter.

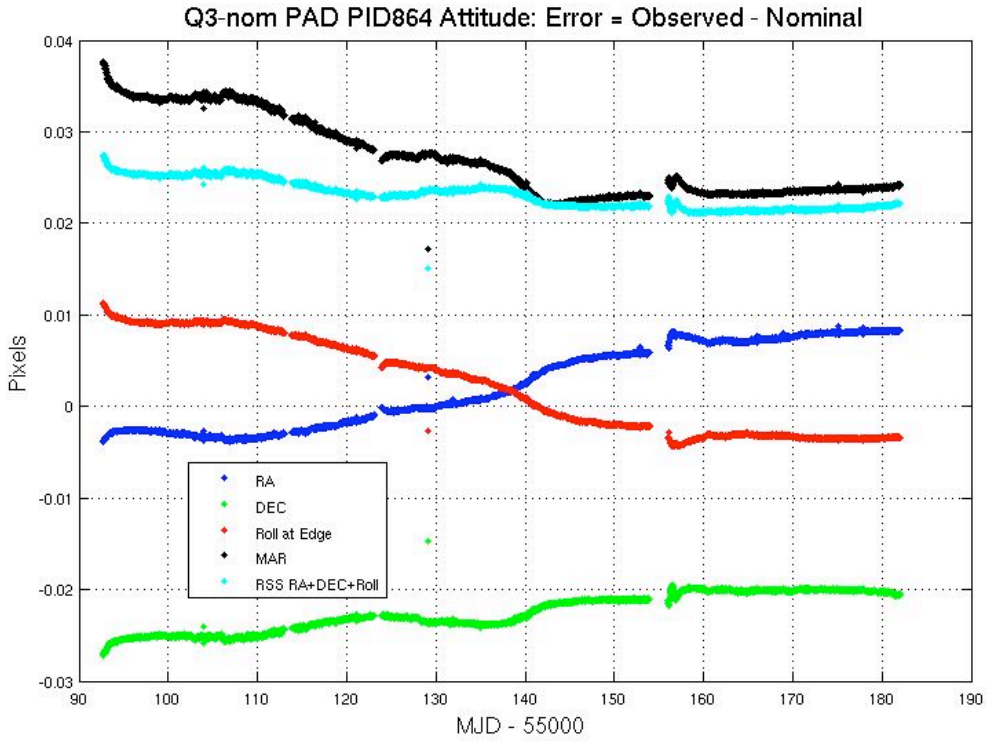

Figure 11: Attitude Error in Quarter 3, calculated by PAD using Long Cadence data. There was one Safe Mode from MJD = 55154.0 to 55156.0. The drift rate off nominal pointing was much less than that during Quarter 2. The stray point at MJD = 55129.07 is a Cadence afflicted by an Argabrightening (shown in Table 6) which was just below the Pipeline multi-channel event threshold, and hence was not gapped in the attitude solution time series. It does not indicate an actual pointing excursion of the telescope.

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 Release 4, 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 when edge effects make it difficult to exclude them automatically. It was not necessary to do this in Release 4 for Quarter 3.

5.3.6 Anomaly Summary Table

Table 5 shows a summary of the Anomalies for both LC and SC data.

Table 5: Anomaly Summary Table for Long and Short Cadences. COARSE_POINT = Loss of Fine Point (Section 5.3.2).

LC CIN			Cadence midTime MJD		
Start	End	Anomaly Type	Start	End	Note
7404	7404	EARTH_POINT	55092.7222	55092.7222	Earth Point at start of Quarter
7596	7596	ARGABRIGHTENING	55096.6455	55096.6455	
7756	7762	COARSE_POINT	55099.9148	55100.0374	
8399	8437	COARSE_POINT	55113.0536	55113.8301	
8889	8930	EARTH_POINT	55123.0661	55123.9039	Data downlink
9183	9183	ARGABRIGHTENING	55129.0736	55129.0736	
9368	9368	ARGABRIGHTENING	55132.8538	55132.8538	
10401	10501	EARTH_POINT	55153.9617	55156.0051	Safe Mode merged with subsequent Data downlink.
10520	10520	ARGABRIGHTENING	55156.3933	55156.3933	
11644	11644	ARGABRIGHTENING	55179.3607	55179.3607	

	SC CIN			Cadence midTime MJD		
M	Start	End	Anomaly Type	Start	End	Note
	210580	210580	EARTH_POINT	55092.7123	55092.7123	First cadence in Q3M1 for thermal transient
1	216350	216352	ARGABRIGHTENING	55096.6424	55096.6438	
1	221150	221321	COARSE_POINT	55099.9118	55100.0282	
1	240439	241599	COARSE_POINT	55113.0499	55113.8400	
n/a	255130	256390	EARTH_POINT	n/a	55123.9144	Data downlink between months 1 and 2.
2	263955	263956	ARGABRIGHTENING	55129.0671	55129.0678	
3	300490	303520	EARTH_POINT	n/a	55156.0156	Safe Mode merged with subsequent Data downlink.
3	304081	304086	ARGABRIGHTENING	55156.3977	55156.4011	

6. Systematic Errors

This Section discusses systematic errors arising in on-orbit operations, most of which will be removed from flux time series by PDC (Section 4). While the Release 4 data is cotrended against image motion (as represented by the motion polynomials calculated by PA) and the 10 board temperatures PEDACQ[1-5]T and PEDDRV[1-5]T, other telemetry items which may be used for cotrending the data in future releases are included 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 will have 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 may be useful for users interested in centroids or pixel data.

While most of the events described in this Section are either reported by the spacecraft or detected in the Pipeline and either corrected or gap-filled, this Section reports events at lower thresholds than the Pipeline, which 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. The method of detection is to fit a trend line to calibrated background (LC) or out-of-optimal-aperture (SC) pixels, fit a parabola, smooth the residual, and look for outliers in the difference between the residual and the smoothed residual. The Pipeline identifies Argabrightenings with this method, and subsequently treats those cadences as gaps for all pixels in that channel. While the Pipeline processes each channel in isolation, all channels are marked as gaps if Argabrightenings are detected on more than half of the channels.

The Pipeline uses a rather high threshold of 100x the median absolute deviation (MAD) for LC and 60x for SC. 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 which exceed a 10x MAD threshold on at least 11 channels (Table 6 and 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 the light curve file). The Supplement contains summaries (`ArgAgg_Q3*_Summary.txt`) and channel-by-channel (`ArgAgg_Q3*_ArgStatsAll.txt`) results.

Table 6: Q3 LC Argabrightening Events with amplitude greater than 10x the median absolute deviation (MAD), which occurred simultaneously on at least 11 of the 84 channels. 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 SNR over Channels of Arg Event, (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. Number of Arg Events Detected: 29. Events Per Month: 9.7.

CIN	RCI	Date (MJD)	MeanSNR	N_{chan}	N_{pipe}
7596	193	55096.64546	173.8	84	75
7857	454	55101.97863	7.5	20	0
7859	456	55102.01950	4.4	12	0
8157	754	55108.10871	7.3	28	0
8563	1160	55116.40475	14.1	67	0
8676	1273	55118.71374	12.0	49	0
8761	1358	55120.45060	7.6	14	0
8775	1372	55120.73667	7.4	13	0
8825	1422	55121.75835	12.9	52	0
9074	1671	55126.84632	6.2	12	0
9075	1672	55126.86675	6.2	12	0
9076	1673	55126.88718	5.5	11	0
9183	1780	55129.07358	104.3	82	39
9207	1804	55129.56398	2.7	12	0
9368	1965	55132.85379	169.5	84	72
9419	2016	55133.89591	14.8	65	0
9438	2035	55134.28415	13.0	64	0
9488	2085	55135.30583	7.6	21	0
9926	2523	55144.25574	10.6	37	0
9984	2581	55145.44089	11.1	44	0
10152	2749	55148.87373	7.7	30	0
10337	2934	55152.65395	46.0	83	3
10520	3117	55156.39330	509.6	84	84
10568	3165	55157.37411	6.6	24	0
10575	3172	55157.51715	5.7	12	0
10826	3423	55162.64598	18.2	71	0
10827	3424	55162.66642	8.0	33	0
11644	4241	55179.36066	121.9	84	59
11698	4295	55180.46408	7.2	19	0

Table 7: Same analysis as Table 6, for Q3 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.

CIN	RCI	Date (MJD)	MeanSNR	N_{chan}	N_{pipe}
216350	5771	55096.64240	81.2	83	57
216351	5772	55096.64308	60.6	83	34
216352	5773	55096.64376	31.1	68	6
224185	13606	55101.97897	7.6	20	0
240439	29860	55113.04989	206.9	83	83
240440	29861	55113.05057	190.4	83	82
240441	29862	55113.05125	122.1	83	81
245366	34787	55116.40577	11.1	48	0
248755	38176	55118.71408	7.4	18	0
251318	40739	55120.45980	7.2	15	0
251737	41158	55120.74518	7.5	15	0
253235	42656	55121.76550	13.3	50	0
263955	7566	55129.06711	75.4	79	52
263956	7567	55129.06779	21.5	33	10
264692	8303	55129.56909	2.4	11	0
269511	13122	55132.85141	14.6	66	0
269512	13123	55132.85209	27.5	80	0
269513	13124	55132.85277	23.8	78	0
269514	13125	55132.85345	21.3	56	2
269515	13126	55132.85413	33.6	70	8
269516	13127	55132.85482	18.0	46	2
271039	14650	55133.89216	7.7	23	0
271629	15240	55134.29402	13.1	57	0
286266	29877	55144.26357	6.2	12	0
286267	29878	55144.26425	3.6	13	0
288006	31617	55145.44872	6.2	17	0
298579	42190	55152.65020	15.2	64	0
298580	42191	55152.65088	7.4	16	0
298581	42192	55152.65157	7.1	16	0
304081	562	55156.39773	140.8	82	78
304082	563	55156.39841	142.0	82	79
304083	564	55156.39909	71.0	82	49
304084	565	55156.39977	64.2	82	42

304085	566	55156.40045	32.6	78	11
304086	567	55156.40113	5.5	17	0
304089	570	55156.40318	4.4	11	0
305500	1981	55157.36424	6.7	20	0
313241	9722	55162.63679	17.7	68	0
313291	9772	55162.67084	7.0	32	0
337785	34266	55179.35419	13.0	59	0
337786	34267	55179.35487	44.2	82	8
337787	34268	55179.35556	27.1	80	1
337788	34269	55179.35624	20.2	73	0

6.2 Variable FGS Guide Stars

The first-moment centroiding algorithm used by the FGS did not originally subtract all of the instrumental bias from the FGS pixels. Thus, the calculated centroid of an FGS star depended on the FGS star's flux when the star was not located at the center of the aperture. Variable stars then induced a variation in the attitude solution calculated from the centroids of 40 guide stars, 10 in each FGS module. The Attitude Determination and Control System (ADCS), which attempts to keep the calculated attitude of the S/C constant, then moved the S/C to respond to this varying input, with the result that the boresight of the telescope moved while the ADCS reported a constant attitude. Science target star centroids and pixel time series, and to a lesser extent aperture flux, then showed systematic errors proportional to the FGS star flux variation.

At the start of Quarter 2 (6/20/2009), the most egregious variable stars were replaced with quieter stars. One intrinsic variable star and one eclipsing binary (EB) remain in the FGS; the Quarter 2 light curves for these stars are shown for historical interest in Figure 12. At the start of Quarter 3 (9/19/2009), the centroiding algorithm was updated to remove all of the instrumental background, greatly diminishing the effect of stellar variability on calculated centroids. The sky background is not removed, but is expected to be negligible. Inspection of the PPA solution (Figure 11) shows that the coupling of the variable guide stars to attitude is below our detection threshold. FGS guide star variability is not expected to be a factor in Q3 photometry.

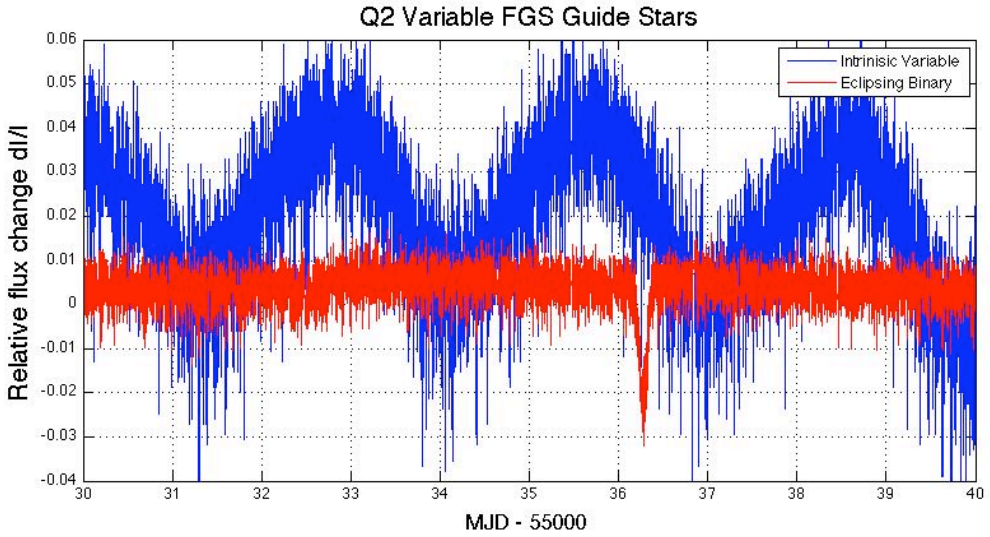

Figure 12: Quarter 2 light curves of two variable FGS guide stars. One of the stars is an eclipsing binary (EB), while the other is an intrinsic variable. The period of the intrinsic variable is approximately 2.9 days, and the period of the EB is approximately 18.25 d. Only 10 days are shown, in order to resolve the intrinsic variability and the eclipse. This Figure is included for historical interest; the coupling of guide star variability to telescope pointing has been greatly diminished in Q3 and beyond.

6.3 Pixel Sensitivity Dropouts

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 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 13). In Release 4, the Pipeline corrects the aperture flux discontinuities (Figure 14) resulting from these pixel discontinuities (Section 4.4), though users examining pixel data and uncorrected light curves need to remain aware of them.

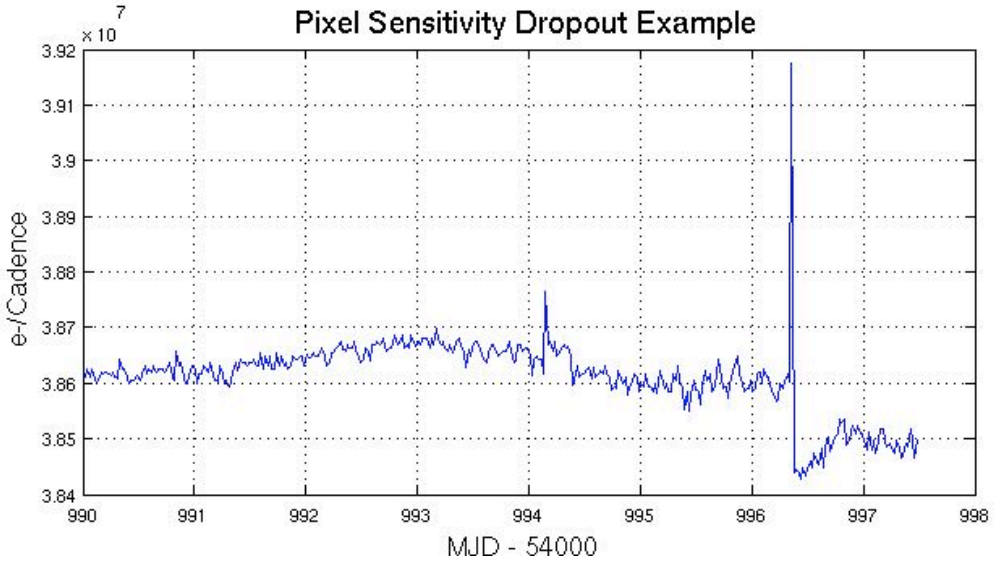

Figure 13: Pixel time series from Q1 (Release 2) showing discontinuity after large CR event. CRs have not been removed by the Pipeline at this stage of processing. Target: KeplerID = 7960363, KeplerMag = 13.3. Dropouts are not corrected on a pixel-by-pixel basis in Release 4.

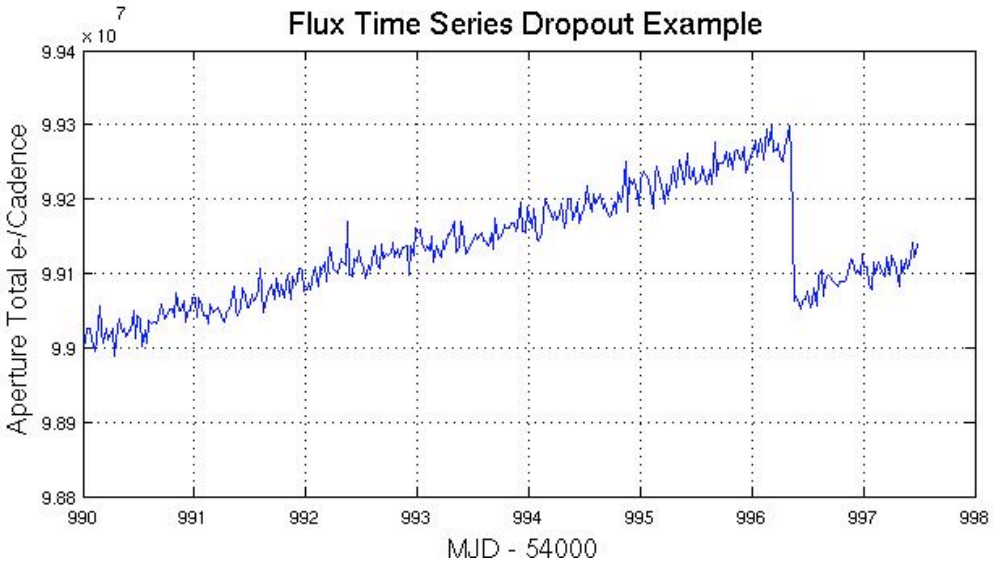

Figure 14: Same event as for the previous Figure as seen in the uncorrected Simple Aperture Photometry (SAP) light curve produced by PA. CR hits have been removed by PA. This figure is presented for historical interest, as the Pipeline now identifies and removes most such discontinuities (Section 4.4).

6.4 Focus Drift and Jitter

Examination of Q1 data (Figure 15) 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.

This jitter is observed in platescale metrics local to each mod/out defined by the motion of target star centroids relative to one another over time. This indicates that we are seeing a change in focus at timescales of 3 to 6 hours and that the behavior appears to be initiated by the desat activities. Reaction wheel temperature sensors with the mnemonics TH1RW3Tand TH1RW4T have the same time signature, but the physical mechanism by which they couple to focus is still under discussion. At the beginning of a Quarter, the reaction wheel heaters do not cycle on and off, and the temperature changes have the same 3 day interval as the planned desaturations (Figure 16). Later in the Quarter, the heaters cycle with a 3 to 6 hr period. At MJD = 55170, new Flight Software parameters were uploaded to substantially reduce the deadband on the reaction wheel housing temperature controller (Figure 16), and subsequent to that date the 3 to 6 hr cycle in both the temperature telemetry and the focus metric were eliminated (Figure 17).

This temperature sensor telemetry for Q3 is nonetheless provided in the Supplement since the problem was not corrected until MJD = 55170 and some users may notice the 3 day period of the momentum accumulation and planned dump cycle in their data.

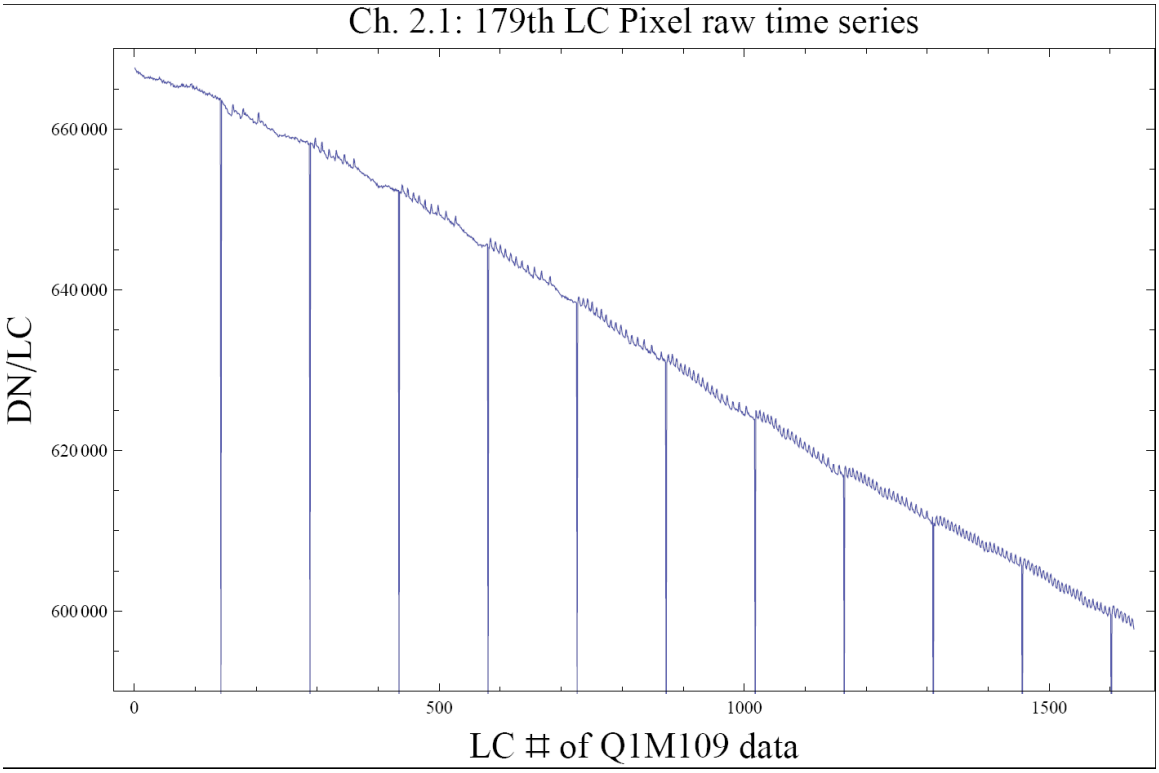

Figure 15: A good example of the 3 to 6 hr Focus Oscillation in a single raw pixel time series from Quarter 1. Similar signatures are seen in flux and plate scale. The large negative-going spikes are caused by desaturations (Section 5.1), which have not been removed from this time series in this plot. The abscissa is the Q1 relative cadence index, and the ordinate is Data Numbers (DN) per Long Cadence (LC). The problem has been eliminated in the Q3 data subsequent to MJD = 55170, as shown in Figure 17

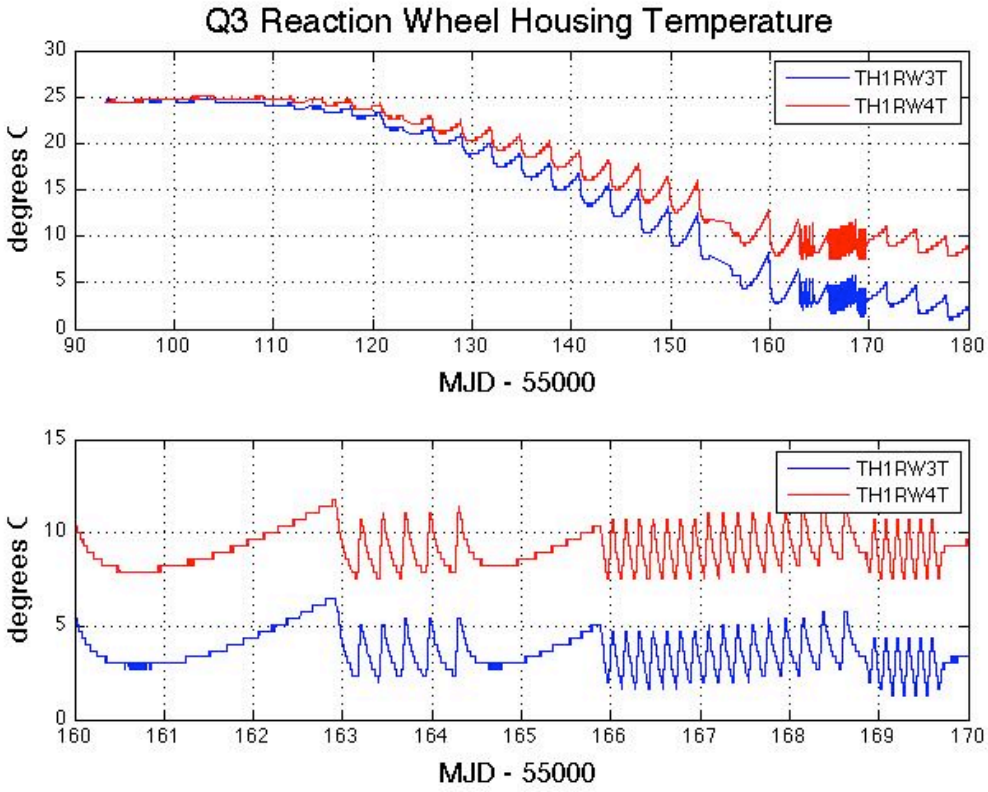

Figure 16: Reaction wheel housing temperatures during Quarter 3. The upper panel shows that temperature variation over most of the Quarter is dominated by a slow seasonal drift and the 3 day period of wheel desaturations. Near the end of the Quarter, however, the reaction wheels have cooled sufficiently to engage the wheel housing heater, which then cycles on and off with a period of 3 to 6 hr (bottom panel). Reducing the dead band on the temperature controller made that 3 to 6 hr variation go away after MJD 55170. 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.

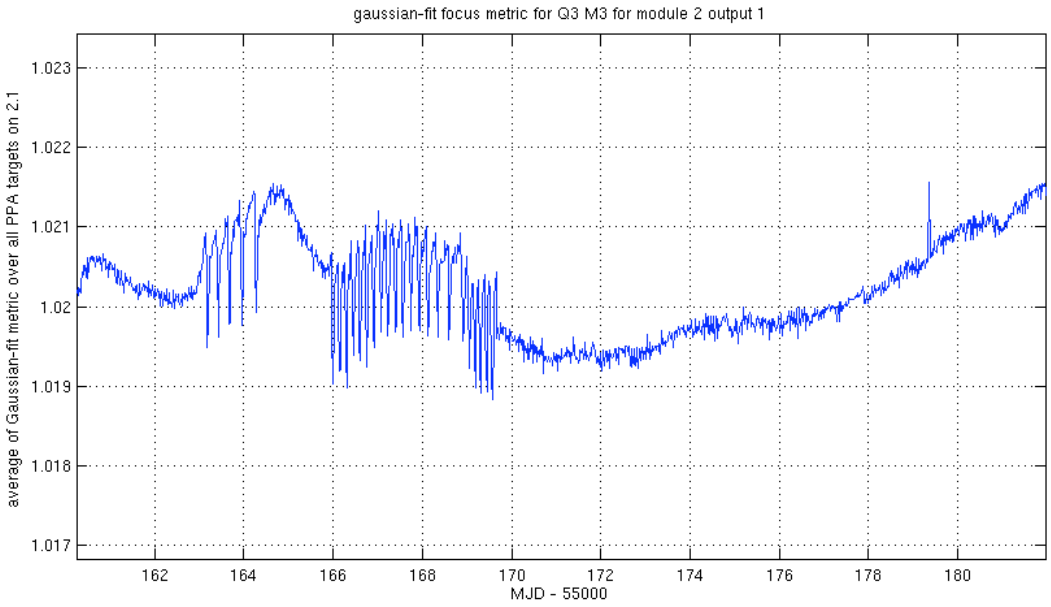

Figure 17: Suppression of 3 to 6 hr focus variation after update of Flight Software reaction wheel heater controller deadband parameters on MJD – 55000 = 170.

The DAWG is investigating whether there is a secular variation of the focus driven by the outgassing of telescope components, in addition to the seasonal and momentum dump cycles driven by temperature changes in Flight System components discussed above. Preliminary results indicate that the seasonal cycle dominates, but definitive results are not expected until the Q5 data is analyzed after June 2010.

6.5 Requantization Gaps

Pixels at mean intensities >20,000 e- show banding as shown in Figure 18, 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.

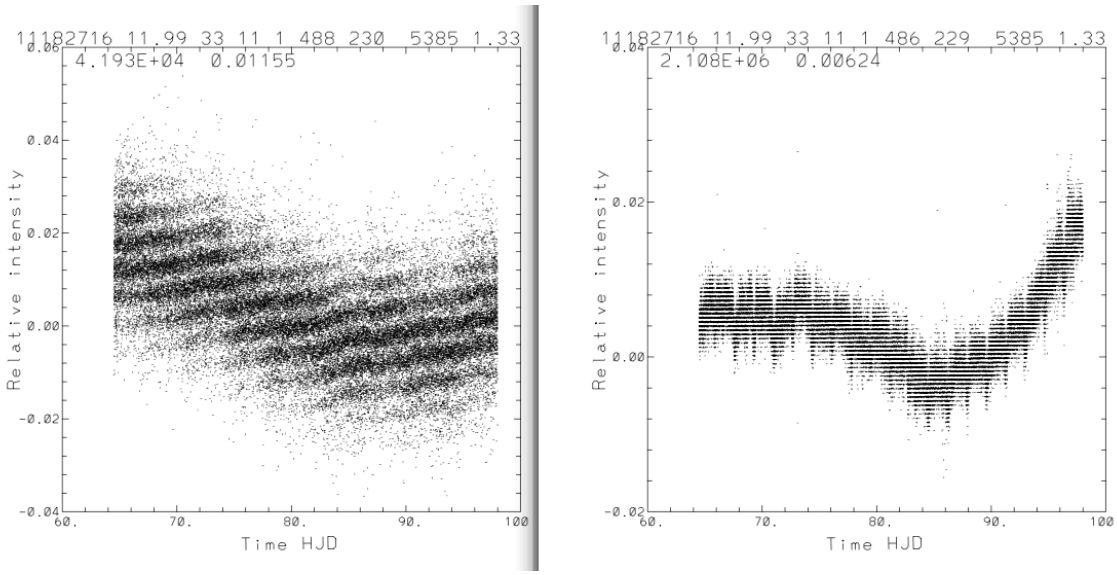

Figure 18: Requantization gap example in Q1 SC pixel time series. The 'band gaps' scale with mean intensity (42,000 e- left, 2.1e6 right). See KIH Section 7.4 for a discussion of quantization and the (insignificant) information loss it entails.

6.6 Spurious Frequencies in SC Data with Spacing of 1/LC

Spurious frequencies are seen in SC flux time series, and pixel data of all types – including black collateral pixels. The frequencies have an exact spacing of 1/LC interval, as shown in Figure 19. 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. It has now been identified in pre-launch ground test data as well as Q3 flight data, and is therefore considered a normal feature of the as-built electronics.

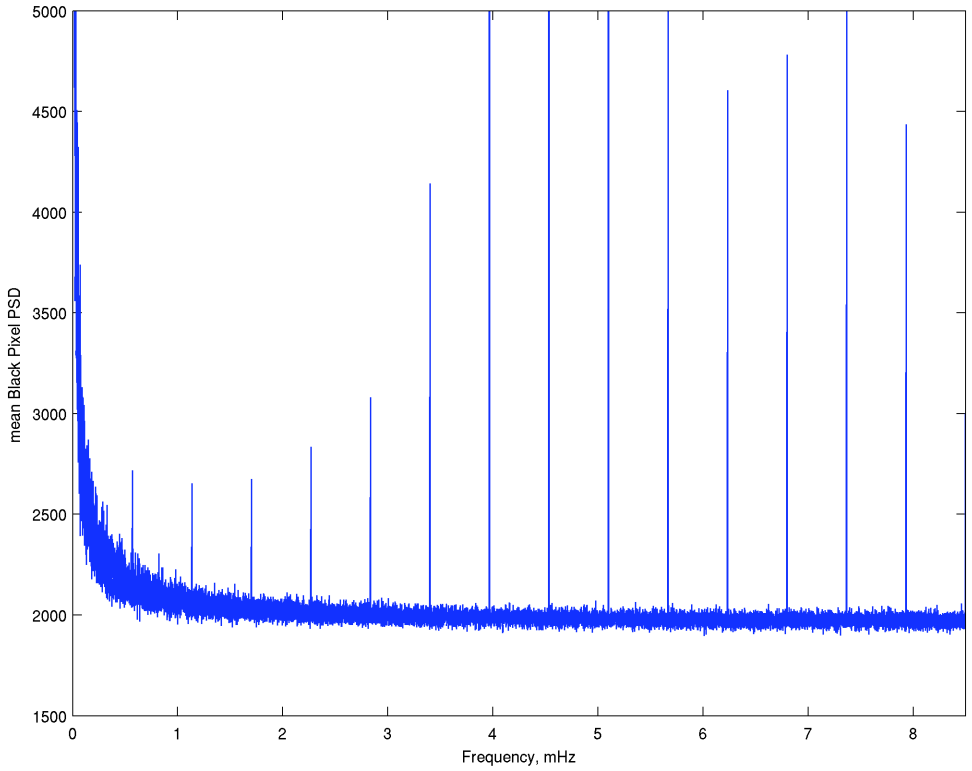

Figure 19: Example of Spurious Frequencies seen in Q1 SC data. The fundamental at 0.56644 mHz (1/LC) can be easily seen here, as well as a comb of harmonics of this frequency.

6.7 Known Erroneous FITS header keywords

The automatically generated FFI WCS FITS header keyword values are known to be incorrect. The MAST staff have updated the 8 "Golden" FFIs from Q0 with WCS headers which are accurate to the extent that image distortion across a mod.out can be neglected (2-3 pixels error at the margins of a mod.out), as described by the file WCSREADME.txt on the MAST FFI ftp site: http://archive.stsci.edu/pub/kepler/ffi. For the other FFIs, users interested in sources for which light curves are unavailable, but for which KeplerIDs exist, may use the Kepler Target Search Page: http://archive.stsci.edu/kepler/kepler_fov/search.php to identify the module, output, row, and column of the source in a given Quarter, as described in the README.txt file on the MAST FFI ftp site. The pixel coordinate to sky coordinate tables in the KIH Supplement may be of some use for locating the corners and edges of mod.outs as projected on the sky.

The WCS keywords in the light curve and target pixel file headers are also incorrect. However, the MODULE and OUTPUT keyword values are correct, and the row and column values are in the files, so it is an easy matter to find the location of an observed target in an FFI to see the surrounding area.

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.

7.2 Light Curves

Users are strongly encouraged to read this section..

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 8. 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 8. 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
  - raw flux value and uncertainty. Gap Cadences are set to -Inf
  - corrected flux value and uncertainty. Gap Cadences are set to -Inf

Table 8: 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 Release 3.

Column Number	Field Name	Data Type	Bytes	Description	Units
1	barytime	1D	8	barycentric time BJD – 2400000*	days
2	timcorr	1E	4	barycentric time correction*	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 raw flux	e- / cadence
9	ap_raw_err	1E	4	SAP raw 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 is 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.

In future Releases, the instrument magnitude flux value will provide an "Instrument magnitude flux time series" that is scaled to astronomically meaningful values as best they can be reconstructed. *Instrument Magnitude flux time series are not available at this time.*

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

Target Pixel Data Files are not currently available as of 4/13/2010, but will be in the near future.

Target pixel data files contain all the pixels from each target from all cadences, while target cadence files contain pixels from all targets for a single cadence. Up to 5 different files are produced for each target. These consist of the long and short cadence pixel data for the target, the collateral pixels for the long and short cadences, and the background pixels

Both original pixel values and calibrated flux values are in the pixel data files. The original pixel value 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 cosmic rays removed. The calibrated pixels have **no**t had background subtracted. Unfortunately, since the background is not subtracted, and users are not presently provided with the list of pixels in the optimal aperture, there is no simple way to construct the 'raw' (PA output) light curve from the 'calibrated' pixels. Users should be aware that the format and content of the target pixel files is the subject of vigorous discussion, in the hopes of remedying this situation.

The target pixel data files are archived as a dataset. A request for the data will return all extensions that were archived with the dataset.

Pixel data table fields are described in the Kepler Archive Manual (KDMC-10008).

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, at the flux-weighted mid-point of the cadences, and are uniquely determined for each star individually.

Users are urged to read this Section as a close reading may help them avoid attempting to do follow-up observations at the wrong time.

7.4.1 Overview

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 Time Correlation Coefficients. Done by DMC, with precision and accuracy to be documented.
3. 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 ~ 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.
4. Subtract readout time slice offsets (See KIH Section 5.1). This is done in PA (Section 4.3) in Release 4. 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 a different mod.outs, so the relative timing of events across Quarter boundaries must take this into account.

7.4.2 Time Stamp Definitions for Release 4

Cadence files (not yet available to MAST users):

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`

Release 4 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 Release 4

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

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)
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.

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.

8. References

Users are encouraged to read papers 7-10 for background, though they are not cited in the text.

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.*, abstract for SPIE Astronomical Instrumentation conference, June 2010.
4. "Photometric Analysis in the Kepler Science Operations Center Pipeline," J. D. Twicken *et al.*, abstract for SPIE Astronomical Instrumentation conference, June 2010.
5. "Presearch Data Conditioning in the Kepler Science Operations Center Pipeline," J. D. Twicken *et al.*, abstract for 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) http://arxiv.org/abs/1001.0256
8. "Initial Characteristics of Kepler Short Cadence Data," R. L. Gilliland *et al.*, submitted to ApJ Letters http://arxiv.org/abs/1001.0142
9. "Overview of the Kepler Science Processing Pipeline," Jon M. Jenkins *et al.*, ApJ Letters **713**, L87-L91 (2010) http://arxiv.org/abs/1001.0258
10. "Discovery and Rossiter-McLaughlin Effect of Exoplanet Kepler-8b," J. M. Jenkins *et al.*, submitted to ApJ http://arxiv.org/abs/1001.0416

9. List of Acronyms and Abbreviations

ACS	Advanced Camera for Surveys
ADC	Analog to Digital Converter
ADCS	Attitude Determination and Control Subsystem
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
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
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	Terrestrial Planet Search module
TVAC	Thermal Vacuum testing

10. Contents of Supplement

The Supplement is available as a full package (DataReleaseNotes_04_SupplementFull.tar) and a short package suitable for emailing (DataReleaseNotes_04_SupplementSmall.tar). The README file describes each file. Here are the files in the short package:

```
>tar tf DataReleaseNotes_04_SupplementSmall.tar
ArgAgg_Q3-MAST_LC_PID1109_MADT010_MCT10_Summary.txt
ArgAgg_Q3M1_SC_PID1131_MADT010_MCT10_Summary.txt
ArgAgg_Q3M2_SC_PID1131_MADT010_MCT10_Summary.txt
ArgAgg_Q3M3_SC_PID1137_MADT010_MCT10_Summary.txt
DataAnomalyTypes_Q3-MAST_LC_PID1109_Summary.txt
DataAnomalyTypes_Q3M1-MAST_SC_PID1131_Summary.txt
DataAnomalyTypes_Q3M2-MAST_SC_PID1131_Summary.txt
DataAnomalyTypes_Q3M3-MAST_SC_PID1131_Summary.txt
Q3M1_SC_isNotFinePoint.txt
Q3M2_SC_isNotFinePoint.txt
Q3M3_SC_isNotFinePoint.txt
Q3_LC_CAL_PA_PDC_r6.1_ksop400_Pipeline-Instance-Detail_Report_100302.txt
Q3_LC_isNotFinePoint.txt
Q3_SC_CAL_PA_PDC_r6.1_ksop400_Pipeline-Instance-Detail_Report_100302.txt
Q3_ZeroCrossings.txt
README.txt
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
```

The long package includes all the files in the short package, and these large files as well:

```
ArgAgg_Q3-MAST_LC_PID1109_MADT010_MCT10_ArgStatsAll.txt
ArgAgg_Q3M1_SC_PID1131_MADT010_MCT10_ArgStatsAll.txt
ArgAgg_Q3M2_SC_PID1131_MADT010_MCT10_ArgStatsAll.txt
ArgAgg_Q3M3_SC_PID1137_MADT010_MCT10_ArgStatsAll.txt
Q3_THRW_MJD_gap.txt
```

11. Calibration Files and Parameters

Pipeline Instance reports document a pipeline run with a unique identifier (Pipeline ID or PID), and include software revision and parameters used. They are included in the Supplement.

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 calibration files actually used in Release 4 are:

```
kplr2008020721_bad-pixels.txt

kplr2008072318_gain.txt

kplr2008102416_read-noise.txt

kplr2008102809_undershoot.txt

kplr2009060215_linearity_txt

kplr2009060615-mmo_2d-black.txt

kplr2009062300_lsflat.txt (large-scale flat field)

kplr2009062414-MMO_ssflat.txt (short-scale flat field)
```

The README files describing the provenance of these files are included in the Supplement.