Kepler Data Release 11 Notes

KSCI-19051-001
Data Analysis Working Group (DAWG)
Jessie Christiansen (Editor)

Data Release 11 for Q8

Q.m		First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Num CINs	Start CIN	End CIN
8	LC	55567.8647	55634.8460	06-Jan-2011 20:45:08	14-Mar-2011 20:18:16	3279	30657	33935
8.1	SCM1	55567.8548	55585.5496	06-Jan-2011 20:30:55	24-Jan-2011 13:11:27	25980	908170	934149
8.2	SCM2	55585.6116	55614.7084	24-Jan-2011 14:40:43	22-Feb-2011 17:00:02	42720	934240	976959
8.3	SCM3	55614.7703	55634.8559	22-Feb-2011 18:29:18	14-Mar-2011 20:32:29	29490	977050	1006539

Prepared by: ______________________ Date: 7/22/11
Jessie Christiansen, Kepler Science Office, for the DAWG (next page)

Approved by: ______________________ Date: 7/22/11
Jeffrey Van Cleve, Editor-in-Chief

Approved by: ______________________ Date: 7-22-11
Jon Jenkins, Co-I for Data Analysis & DAWG Lead

Approved by: ______________________ Date: 7/22/11
Michael R. Haas, Kepler Science Office Director

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 place 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

The Data Characteristics Handbook and accompanying Data Release Notes are the collective effort of the Data Analysis Working Group (DAWG), composed of Science Office (SO), Science Opertions Center (SOC), Guest Observer Office (GO), and Science Team (ST) members as listed below:

Christiansen, Jessie
Jenkins, Jon*, Chair
Caldwell, Douglas*, Co-Chair
Allen, Christopher L.
Barclay, Thomas
Bryson, Stephen T.
Burke, Christopher J.
Clarke, Bruce D.
Cote, Miles T.
Fanelli, Michael N.
Gilliland*, Ron (STSci)
Girouard, Forrest
Haas, Michael R.
Hall, Jennifer
Ibrahim, Khadeejah
Kinemuchi, Karen
Klaus, Todd
Kolodziejczak, Jeffery (MSFC)
Li, Jie
Machalek, Pavel
McCauliff, Sean D.
Middour, Christopher K.
Morris, Rob
Mullally, Fergal
Quintana, Elisa V.
Rowe, Jason
Seader, Shawn
Smith, Jeffrey Claiborne
Still, Martin
Stumpe, Martin C.
Tenenbaum, Peter G.
Thompson, Susan E.
Twicken, Joe
Uddin, Akm Kamal
Van Cleve, Jeffrey
Wohler, Bill

* Science Team

When citing this document, please use the following reference:

J. L. Christiansen, J. E. Van Cleve, J. M. Jenkins, D. A. Caldwell, C. L. Allen, T. Barclay, S. T. Bryson, C. J. Burke, B. D. Clarke, M. T. Cote, M. N. Fanelli, R. L. Gilliland, F. Girouard, M. R. Haas, J. R. Hall, K. Ibrahim, K. Kinemuchi, T. C. Klaus, J. J. Kolodziejczak, J. Li, P. Machalek, S. D. McCauliff, C. K. Middour, R. Morris, F. Mullally, E. V. Quintana, S. Seader, J. C. Smith, M. D. Still, M. C. Stumpe, P. G. Tenenbaum, S. E. Thompson, J. D. Twicken, A. K. Uddin, and B. Wohler, 2011, Kepler Data Release 11 Notes (KSCI-19051-001).

Contents

Prefatory Admonition to Users

The corrected light-curve product generated by the PDC (Pre-search Data Conditioning) pipeline module is designed to enable the Kepler planetary transit search. Although significant effort has been expended to preserve the natural variability of targets in the corrected light curves in order to enable astrophysical exploitation of the Kepler data, it is not possible to perfectly preserve general stellar variability on long timescales with amplitudes comparable to or smaller than the instrumental systematics, and PDC currently is known to remove or distort astrophysical features in a subset of the corrected light curves. In those cases where PDC fails, or where the requirements of an astrophysical investigation are in conflict with those for transit planet search, the investigator should use the raw light-curve product, for which basic calibration has been performed but correction for instrumental systematics has not, instead of the PDC (corrected) light-curve product. Where appropriate, the investigator can then use the ancillary engineering data, ensemble cotrending basis vectors (see Appendix B) and image motion time series provided in the relevant Data Release Notes Supplement/s for systematic error correction. Investigators are strongly encouraged to study the Data Characteristics Handbook and Data Release Notes for any data sets they intend to use. The Science Office advises against publication of results based on Kepler light curves without careful consideration and due diligence by the end user, and dialog with the Science Office or Guest Observer Office where appropriate.

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

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

1 Introduction

These Data Release Notes provide information specific to the release of Q8 data, processed with SOC Pipeline 7.0. These Notes contain the summary figures and tables for this quarter—the companion text can be found in the Kepler Data Characteristics Handbook (KSCI-19040). The sections are numbered in the same order in these Notes and the Handbook to assist the reader.

As of Data Release 11, much of the information contained in these Data Release Notes and the accompanying Supplement is also available in the light curve and target pixel FITS files at MAST. Once the user is comfortable with their understanding of the phenomena listed here, they might find it simpler to collect the information about affected cadences directly from the FITS files. However, new phenomena will continue to be described here and in the Data Characteristics Handbook, so users are encouraged to continue checking with these documents to stay abreast of developments.

1.1 Dates, Cadence Numbers, and Units

No changes from the Data Characteristics Handbook.

Contents of Data Release 11.

Q.m		First Cadence MJD midTime	Last Cadence MJD midTime	First Cadence UT midTime	Last Cadence UT midTime	Num CINs	Start CIN	End CIN
8	LC	55567.8647	55634.8460	06-Jan-2011 20:45:08	14-Mar-2011 20:18:16	3279	30657	33935
8.1	SCM1	55567.8548	55585.5496	06-Jan-2011 20:30:55	24-Jan-2011 13:11:27	25980	908170	934149
8.2	SCM2	55585.6116	55614.7084	24-Jan-2011 14:40:43	22-Feb-2011 17:00:02	42720	934240	976959
8.3	SCM3	55614.7703	55634.8559	22-Feb-2011 18:29:18	14-Mar-2011 20:32:29	29490	977050	1006539

2 Release Description

No changes from the Data Characteristics Handbook.

3 Evaluation of Performance

3.1 Overall

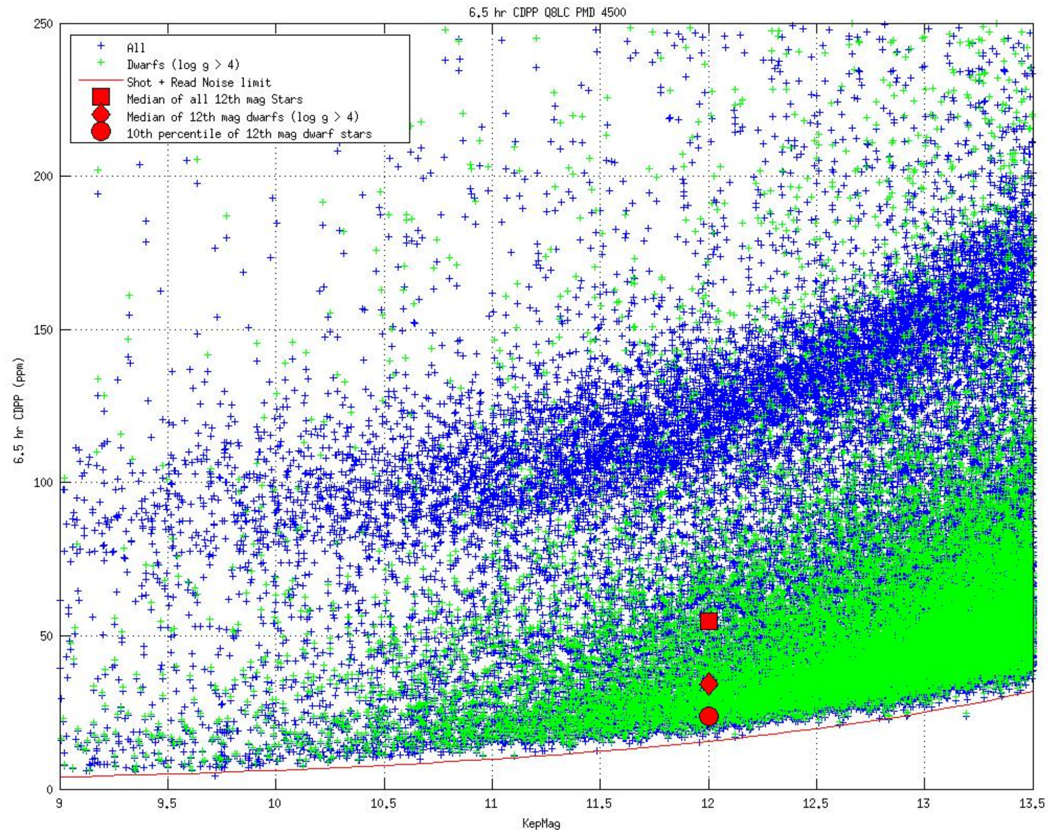

Figure 1: 6.5 hr Temporal Median (TM) of the Q8 CDPP time series calculated by the TPS pipeline module for stars between 9^{th} and 15^{th} magnitude. The 6 hr TMCDPPs have been divided by sqrt(13/12) = 1.041 to approximate 6.5 hr TMCDPPs. Stars on the planetary target list with log g > 4, which are almost certainly dwarf stars, are shown as green + symbols; other stars are marked with blue + symbols.

Table 1: Aggregate statistics for the TMCDPPs plotted in Figure 1. Column Definitions: (1) Kepler Magnitude at the center of the 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 the bin. (3) 10th percentile TMCDPP for dwarfs in the bin. (4) Median TMCDPP for dwarfs in the bin. (5) Number of all stars in the bin. (6) 10th percentile TMCDPP of all observed stars in the bin. (7) Median TMCDPP for all stars in the bin. (8) Simplified noise model CDPP.

Kp mag	No. dwarfs	10th prctile	Median	No. stars	10th prctile	Median	Noise model
9.0	29	7.6	22.0	183	8.6	63.6	3.8
10.0	161	11.1	31.0	587	12.4	78.1	6.0
11.0	630	16.8	30.0	1775	19.0	76.5	9.5
12.0	2223	23.8	34.3	4346	25.2	54.6	15.2
13.0	7011	35.3	46.3	10279	36.5	53.6	24.4
14.0	0	0.0	0.0	16644	55.6	73.3	40.1
15.0	0	0.0	0.0	28663	97.8	127.2	68.8

4 Historical Events

In this Section, we discuss cadences that may not be useful for high-precision photometry due to planned or unplanned spacecraft events.

4.1 Kepler Mission Timeline to Date

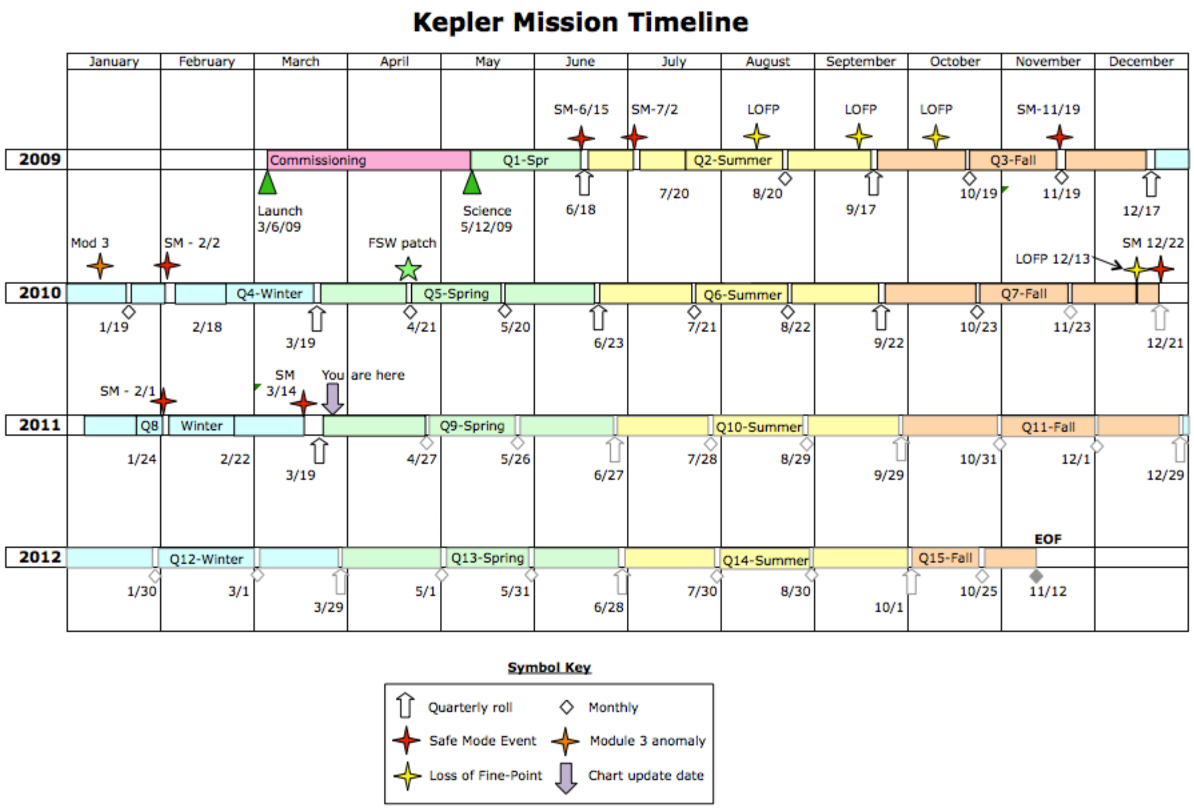

Figure 2: Kepler Mission Timeline as of the end of Q8.

4.2 Safe Mode

There were three Safe Mode events in Q8, shown in Figure 2. The first occurred in the downlink gap between Q7 and Q8, and caused a delay in the resumption of science observations after the gap. A second event occured in the second month of the quarter, lasting nearly three days. The third and final event occurred at the end of the third month, halting science observations early. The cadences lost to the safe mode that occurred during science observations are listed in Tables 7 and 8. In the FITS files at MAST, the quality flag column indicates which cadences are lost during safe modes.

4.3 Loss of Fine Point

The cadences obtained when the spacecraft was not in fine point are listed as COARSE_POINT in Tables 7 and 8, the LC and SC anomaly summary tables respectively. These cadences are also flagged in the FITS files with the quality flag column.

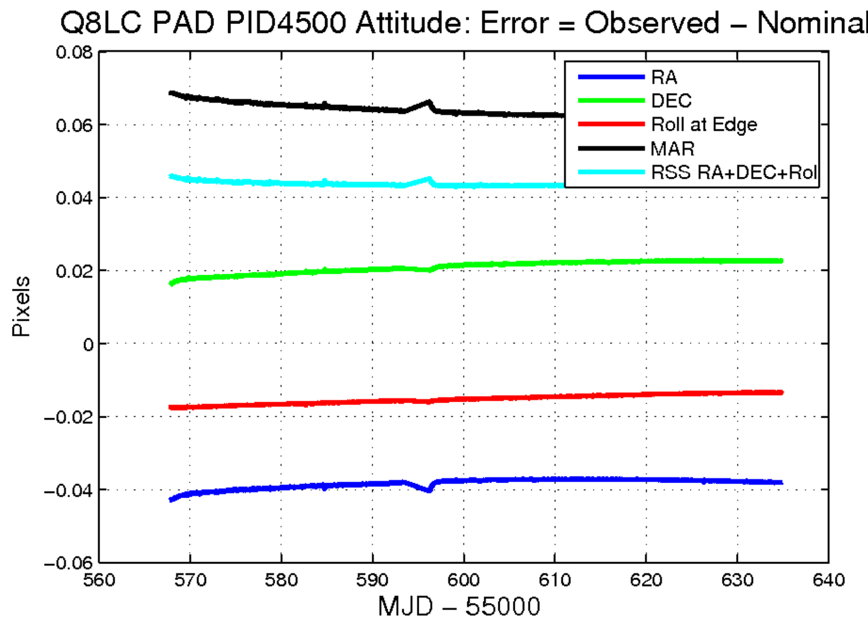

Figure 3: Attitude Error in Q8, calculated by PAD (Photometer Attitude Determination) using Long Cadence data.

4.4 Attitude Tweaks

The pointing history is shown in Figure 3; there were no attitude tweaks over the course of Q8.

4.5 Variable FGS Guide Stars

No changes from the Data Characteristics Handbook.

4.6 Module 3 Failure

No changes from the Data Characteristics Handbook.

5 Ongoing Phenomena

In this Section, we document the systematic errors arising in nominal on-orbit operations, most of which will be removed from the PDC flux time series by the scientific pipeline.

5.1 Image Motion

The image motion on a per-target basis is now available in the FITS files at MAST via the POS_CORR1 (column) and POS_CORR2 (row) columns. In Figure 4 we show the column and row motion time series for the center of mod.out 2.1; these time series are provided for each mod.out in the Supplement.

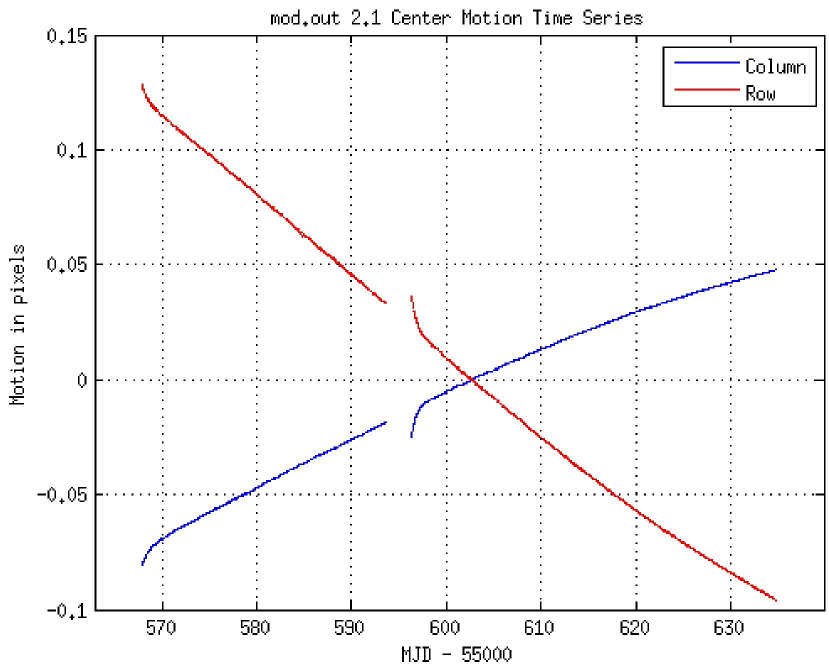

Figure 4: The center motion time series for module 2, output 1. The gap at MJD-55000 = 595 is the safe mode that occurred in month 2.

5.2 Focus Changes

The change in width of the PRF with time is shown in Figure 5.

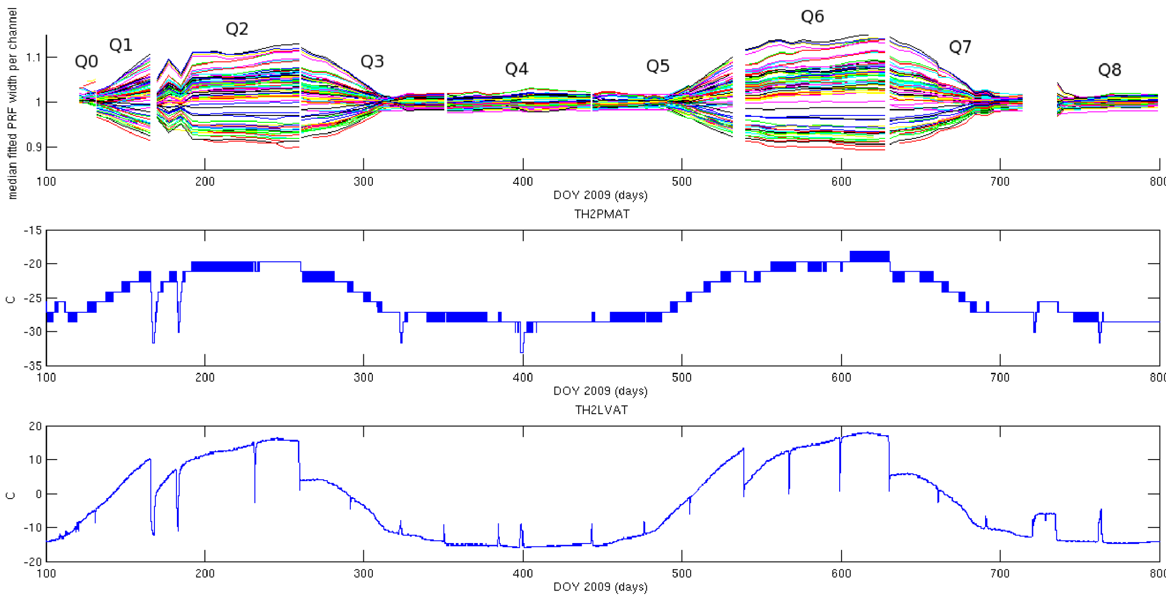

Figure 5: The correlation of the variation in pixel response function (PRF) width with various spacecraft temperatures, demonstrating the seasonal nature of the focus and PRF changes.

5.3 Momentum Desaturation

The Long Cadences affected by momentum desaturations in the reaction wheels are listed in Table 2, and the Short Cadences in Table 3. They are also flagged in the FITS files at MAST in the quality flag column.

Table 2: Momentum desaturations in Q8 and the corresponding Long Cadences. CIN = cadence interval number, RCI = relative cadence index.

LC		
CIN	RCI	Date (MJD)
30752	96	55569.80587
30898	242	55572.78918
31044	388	55575.77249
31190	534	55578.75579
31336	680	55581.73910
31482	826	55584.72241
31627	971	55587.68528
31773	1117	55590.66859
32153	1497	55598.43334
32299	1643	55601.41665
32445	1789	55604.39995
32591	1935	55607.38326
32737	2081	55610.36656
32883	2227	55613.34987
33028	2372	55616.31274
33174	2518	55619.29605

Continued on next page

Table 2 – continued from previous page

CIN	RCI	Date (MJD)
33320	2664	55622.27936
33466	2810	55625.26266
33612	2956	55628.24597
33758	3102	55631.22928
33904	3248	55634.21258

Table 3: Momentum desaturations in Q8 and the corresponding Short Cadences. CIN = cadence interval number, RCI = relative cadence index. The months are separated by horizontal lines.

SC		
CIN	RCI	Date (MJD)
911030	2861	55569.80281
911031	2862	55569.80349
915410	7241	55572.78612
915411	7242	55572.78680
919790	11621	55575.76942
924170	16001	55578.75273
924171	16002	55578.75341
928550	20381	55581.73604
932930	24761	55584.71934
937280	3041	55587.68221
941660	7421	55590.66552
953060	18821	55598.43027
957440	23201	55601.41358
961820	27581	55604.39689
966200	31961	55607.38019
970580	36341	55610.36350
974960	40721	55613.34681
979310	2261	55616.30968
983690	6641	55619.29299
988070	11021	55622.27629
988071	11022	55622.27697
992450	15401	55625.25960
996830	19781	55628.24290
996831	19782	55628.24359
1001210	24161	55631.22621
1001211	24162	55631.22689
1005590	28541	55634.20952
1005591	28542	55634.21020

5.4 Reaction Wheel Zero Crossings

The cadences occurring during reaction wheel zero crossings are listed in Table 4.

Table 4: Zero crossing events in Q8, defined as the time from first to last zero crossing in the event, rounded to the nearest cadence. The corresponding Short Cadence numbers can be found in the Data Release 11 Supplement. CIN = cadence interval number, RCI = relative cadence index.

Event no.	MJD start	MJD end	CIN start	CIN end	RCI start	RCI end
1	55569.602	55569.642	30742	30744	86	88
2	55569.622	55569.990	30743	30761	87	105
3	55572.850	55573.075	30901	30912	245	256
4	55573.055	55573.075	30911	30912	255	256
5	55578.735	55578.797	31189	31192	533	536
6	55584.722	55584.845	31482	31488	826	832

5.5 Downlink Earth Point

The cadences occurring during spacecraft Earth Point are listed in Tables 7 and 8. They are also flagged in the FITS files at MAST in the quality flag column.

5.6 Manually Excluded Cadences

There were no manually excluded cadences in Q8. As of Release 11, any cadences that are manually excluded will be indicated in the quality flag column in the FITS files at MAST.

5.7 Incomplete Apertures Give Flux and Feature Discontinuities at Quarter Boundaries

No changes from the Data Characteristics Handbook.

5.8 Argabrightening

The cadences affected by Argabrightening events are listed in Tables 5 and 6 for LC and SC respectively. They are also flagged in the FITS files at MAST in the quality flag column.

5.9 Background Time Series

The background flux time series for Q8 is shown in Figure 6. We note that due to the presence of faint stars in the pixels used to measure the background flux, there is typically a small over-estimation in the background flux. For very faint targets (Kp > 18) this can result in occasional negative flux values in the time series. For brighter targets this has a negligible affect. If this is a concern, users are advised to add the background time series (provided in the FITS files) back to the flux time series, and perform their own background subtraction using appropriate pixels in the target pixel files where available.

5.10 Pixel Sensitivity Droputs

No changes from the Data Characteristics Handbook.

5.11 Short Cadence Requantization Gaps

No changes from the Data Characteristics Handbook.

Table 5: Q8 LC Argabrightening Events with amplitude $T_{\mathrm{MAD}} > 10$, and occurring on a number of channels $T_{\mathrm{MCE}} > 10$. The columns are (1) Cadence interval number (CIN) for Argabrightening cadences, (2) Relative cadence index (RCI) for Argabrightening cadences, (3) Argabrightening cadence mid-times (MJD), (4) Mean Argabrightening statistic over channels included in the Argabrightening event $< S_{\mathrm{Arg}} >_{\mathrm{FPA}}$, (5) Number of channels exceeding threshold for this cadence ($\mathrm{N_{chan}}$), (6) Number of channels exceeding the default pipeline threshold for this cadence ($\mathrm{N_{pipe}}$).

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >_{\mathrm{FPA}}$	$\mathrm{N_{chan}}$	$\mathrm{N_{pipe}}$
30908	252	55572.99352	113.2	80	61
30935	279	55573.54522	22.3	77	0
31383	727	55582.69948	72.2	80	2
31384	728	55582.71991	9.3	35	0
31898	1242	55593.22279	351.7	80	78
32480	1824	55605.11513	13.2	59	0
32503	1847	55605.58510	7.8	21	0
32516	1860	55605.85074	24.6	78	0
32729	2073	55610.20310	3.7	12	0
32730	2074	55610.22353	5.1	18	0
32731	2075	55610.24396	6.4	26	0
32732	2076	55610.26440	7.4	27	0
32733	2077	55610.28483	8.1	30	0
32734	2078	55610.30526	8.7	36	0
32735	2079	55610.32570	8.2	31	0
32736	2080	55610.34613	6.0	25	0
32738	2082	55610.38700	5.3	16	0
32739	2083	55610.40743	5.3	18	0
32740	2084	55610.42787	5.4	19	0
32741	2085	55610.44830	4.7	17	0
32742	2086	55610.46873	3.3	14	0
32743	2087	55610.48917	4.6	17	0
32810	2154	55611.85822	23.3	76	0
32885	2229	55613.39074	365.1	80	78
33029	2373	55616.33318	38.4	79	0
33199	2543	55619.80689	131.4	80	73
33286	2630	55621.58461	69.0	79	0
33311	2655	55622.09545	7.6	23	0
33403	2747	55623.97535	39.0	79	0
33539	2883	55626.75432	7.5	17	0
33656	3000	55629.14505	16.6	40	0
33770	3114	55631.47448	8.1	27	0
33876	3220	55633.64044	30.6	78	0

5.12 Spurious Frequencies in SC Data

No changes from the Data Characteristics Handbook.

5.13 Propagation of Uncertainties

In order to conserve valuable run time, the SOC pipeline currently only calculates the true uncertainties from a fully propagated covariance matrix every 24th cadence; the variances are linearly interpolated for the

Table 6: Q8 SC Argabrightening Events with amplitude $T_{\mathrm{MAD}} > 10$, and occurring on a number of channels $T_{\mathrm{MCE}} > 10$. The columns have the same meanings as Table 5. Note consecutive detections of the largest events. A horizontal line separates the three months of the quarter; the relative cadence index (RCI) is reset at the start of each month.

CIN	RCI	Mid-Times (MJD)	$< S_{\mathrm{Arg}} >$	N_{chan}	N_{pipe}
915724	7555	55572.99999	122.2	78	75
915725	7556	55573.00067	26.7	64	1
916518	8349	55573.54080	28.8	77	0
922034	13865	55577.29786	6.6	13	0
929979	21810	55582.70936	94.3	78	66
929980	21811	55582.71004	9.8	31	0
945403	11164	55593.21495	377.5	78	78
962873	28634	55605.11411	14.6	56	0
963551	29312	55605.57591	7.8	24	0
963944	29705	55605.84359	17.1	66	0
963945	29706	55605.84427	10.2	43	0
972787	38548	55611.86673	8.1	19	0
972788	38549	55611.86741	17.8	59	0
973823	39584	55612.57237	6.9	13	0
975030	40791	55613.39448	346.7	78	78
975031	40792	55613.39517	39.7	76	13
978177	1128	55615.53797	7.4	17	0
979356	2307	55616.34101	26.8	78	0
979357	2308	55616.34169	17.4	59	0
984430	7381	55619.79701	155.4	78	78
987042	9993	55621.57610	58.4	78	34
987043	9994	55621.57678	24.9	74	0
987799	10750	55622.09171	6.4	19	0
990560	13511	55623.97228	44.7	78	16
994650	17601	55626.75806	9.6	32	0
998142	21093	55629.13653	11.1	33	0
1001577	24528	55631.47618	10.6	43	0
1004761	27712	55633.64487	36.8	78	5

remaining cadences. Users of the provided flux uncertainties will note that for targets that have large variations on timescales less than 12 hours, the uncertainties can vary unphysically due to interpolation between largely divergent values. An example is shown in Figure 7. Note that in Q3, the propagation of uncertainties was turned off in the pipeline, in which case the uncertainties are approximated for every cadence instead of fully propogated.

For bright sources, users have the option of reconstructing an approximate set of uncertainties using:

$$\sigma^2 = \mathrm{shotNoise}^2 + \mathrm{pixelsInAperture} * \mathrm{framesPerCadence} * \mathrm{readNoisePerPixelPerFrame}^2 \quad (1)$$

The number of pixels used in the optimal aperture can be calculated from either the light curve or target pixel files. The read noise per pixel per frame for each channel is given in the headers of those files. There are 270 frames per cadence for LC, and nine frames per cadence for SC. See column 8 of Table 1 for an application of this approximation.

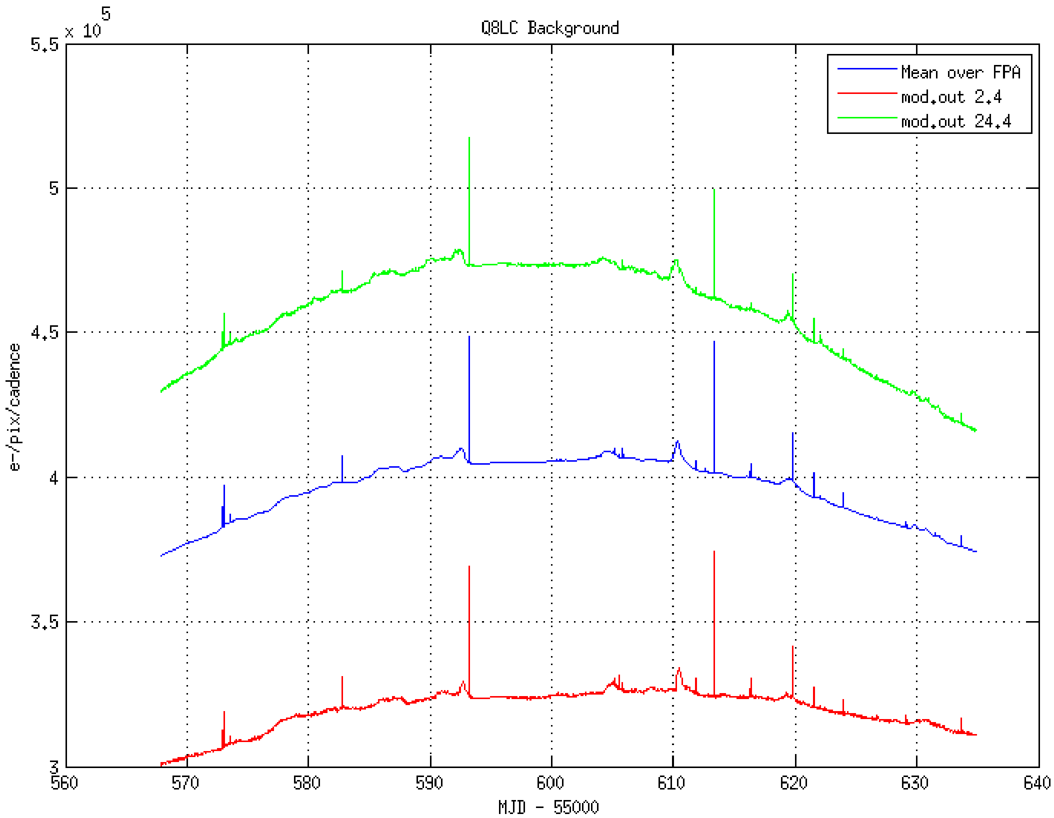

Figure 6: The background flux time series for Q8 showing the average over all the modules, and two individual modules. The narrow spikes are Argabrightening events (see Section 5.8)
.

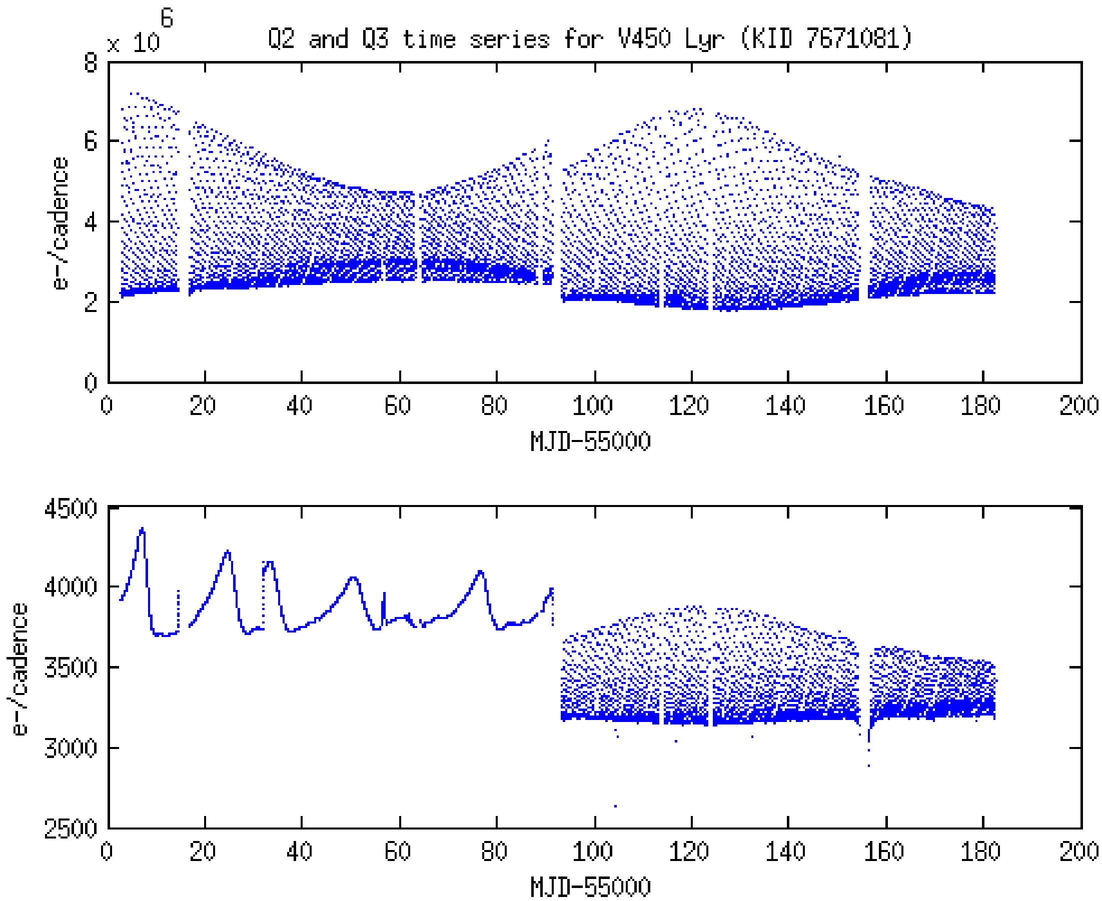

Figure 7: Upper panel: The Q2 and Q3 flux time series for V450 Lyr. Lower panel: The corresponding flux uncertainties as provided in the light curve and target pixel files. The large variations seen in Q2 are due to the decimation in the sampling of the error calculation and subsequent interpolation. In Q3, this full calculation was turned off and errors were approximated for each cadence.

5.14 Anomaly Summary Tables

The full lists of affected cadence interval numbers (CINs) for Tables 7 and 8 below are included in the Data Release 11 Supplement.

Table 7: Q8 LC Anomaly Summary Table

LC CIN start	LC CIN end	Anomaly Type	Note
30657	30657	EARTH POINT	Monthly data downlink
30908	30908	ARGABRIGHTENING	See Section 5.8
31383	31383	ARGABRIGHTENING	
31898	31898	ARGABRIGHTENING	
31915	32046	SAFE MODE	See Section 4.2
32885	32885	ARGABRIGHTENING	
33199	33199	ARGABRIGHTENING	
33286	33286	ARGABRIGHTENING	

Table 8: Q8 SC Anomaly Summary Table

SC CIN start	SC CIN end	Anomaly Type	Note
908170	908170	EARTH POINT	Monthly data downlink
911029	911040	COARSE POINT	See Section 4.3
915409	915420	COARSE POINT	
915724	915725	ARGABRIGHTENING	See Section 5.8
919789	919800	COARSE POINT	
924169	924180	COARSE POINT	
928549	928560	COARSE POINT	
929979	929980	ARGABRIGHTENING	
932929	932940	COARSE POINT	
945940	949840	SAFE MODE	
979309	979320	COARSE POINT	
983689	983700	COARSE POINT	
984430	984430	ARGABRIGHTENING	
987042	987042	ARGABRIGHTENING	
988069	988080	COARSE POINT	
990560	990560	ARGABRIGHTENING	
992449	992460	COARSE POINT	
996828	996840	COARSE POINT	
1001209	1001220	COARSE POINT	
1005589	1005600	COARSE POINT	

6 Time and Time Stamps

No changes from the Data Characteristics Handbook.

7 Contents of Supplement

The Supplement is available as a full package (DataReleaseNotes11SupplementFull.tar), which contains the files described below.

Pipeline Instance Detail Reports

`q8_lc_sTad+pa-mpe-true+tps=ppa_ksop836_as-run-pipeline-instance-detail-110608.txt`
`q8_sc_r7.0_ksop836_as-run-pipeline-instance-detail_110608.txt`

Data Anomaly Types

`DataAnomalyTypes_Q8LC_LC_PID4500.txt`
`DataAnomalyTypes_Q8SCM1_SC_PID4520.txt`
`DataAnomalyTypes_Q8SCM2_SC_PID4520.txt`
`DataAnomalyTypes_Q8SCM3_SC_PID4520.txt`

Mod.out central motion

`Q8LC_central_column_motion.txt`
`Q8LC_central_row_motion.txt`

Thermal Telemetry

`Q8_LDE_averageBoardTemp.txt`
`Q8_TH12LVAT_MJD_gap.txt`
`Q8_TH1RW34T_MJD_gap.txt`

Background Flux Time Series

`Q8LC_background.txt`
`Q8SCM1_background.txt`
`Q8SCM2_background.txt`
`Q8SCM3_background.txt`

Argabrightening Detections

`Q8LC_LC_ArgAgg_Summary.txt`
`Q8SCM1_SC_ArgAgg_Summary.txt`
`Q8SCM2_SC_ArgAgg_Summary.txt`
`Q8SCM3_SC_ArgAgg_Summary.txt`

Out of Fine Point Cadence Lists

`Q8LC_LC_isNotFinePoint.txt`
`Q8SCM1_SC_isNotFinePoint.txt`
`Q8SCM2_SC_isNotFinePoint.txt`
`Q8SCM3_SC_isNotFinePoint.txt`

Zero Crossing Events

`Q8LC_ZeroCrossings.txt`

Q8SCM1_ZeroCrossings.txt

7.1 Short Supplement Package

The Supplement also contains a short package suitable for emailing (DataReleaseNotes11SupplementShort.tar). The small package does not contain the following files:

Q8LC_background.txt
Q8SCM1_background.txt
Q8SCM2_background.txt
Q8SCM3_background.txt

Q8LC_central_column_motion.txt
Q8LC_central_row_motion.txt

Q8_TH12LVAT_MJD_gap.txt
Q8_TH1RW34T_MJD_gap.txt

8 References

No changes from the Data Characteristics Handbook.

A Problems with Q8, month 2 short-cadence data

There were issues with the processing of short-cadence data taken during Quarter 8 month 2. The large data gap between SC CIN 945940 and 949840 due to the safe mode caused problems with the harmonic fitting routine near the gap. The result is that a number of short-cadence light curves contain artifacts similar to that shown in Figure 8. Since this issue only affects pre-search data conditioned (PDC) light curves, we recommend that users consider using simple aperture photometry (PA) data until this issue can be resolved (see Prefatory Admonition).

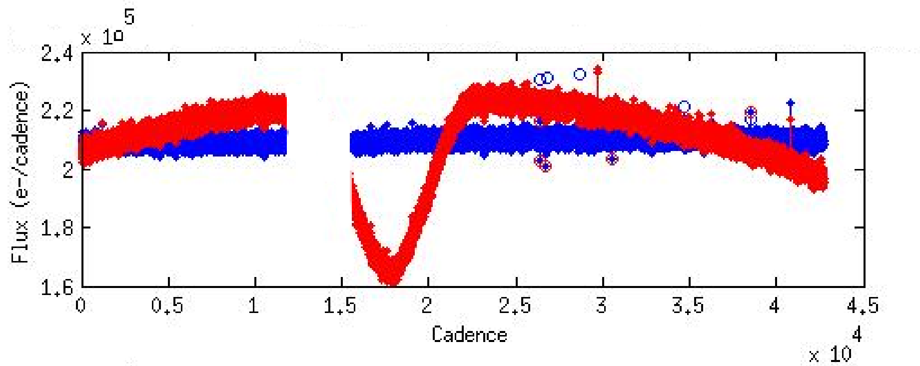

Figure 8: Plotted in blue is the simple aperture photometry data and in red is the pre-search data conditioned (PDC) flux. Issues with the harmonic fitting caused some SC PDC fluxes to contain artifacts such as this.

B Ensemble Cotrending Basis Vectors

We can infer the presence of systematic errors in the Kepler flux time series from the correlations that we observe between them, since we do not expect the stars themselves to have correlated signals. These correlations can be represented as linear combinations of orthonormal functions, called cotrending basis vectors (CBVs), which in some sense represent most of the correlated features in a reference ensemble of flux time series for a given Quarter and output channel (mod.out). The Science Office will soon provide the CBVs of a reference ensemble of highly correlated stars, so that users can perform their own systematic error removal without having particular knowledge of proprietary targets.

Pipeline versions up to and including SOC 7.0 use engineering telemetry and local image motion polynomials derived by the Pipeline itself to remove systematic errors from target flux time series. User systematic error correction using the CBVs may be preferable or necessary, since:

1. An ensemble of flux time series is potentially a more complete representation of the trends than engineering telemetry and image motion,
2. Users can decide how many CBVs to use to fit systematic trends,
3. Users can use the CBVs with non-least squares fitting methods such as MAP (Jenkins et al., 2010) or the lasso (Tibshirani 1994; `http://www-stat.stanford.edu/~tibs/ftp/lasso.ps`).
4. Users of target pixel files who generate their own flux time series will need to do their own systematic error removal, though it is up to the user to understand whether a set of CBVs which well-represent systematic errors in uncorrected Pipeline flux time series is also a good representation of their systematic errors given differences in aperture size and their method of extracting a flux time series from the pixels.

The method for CBV generation will be detailed in a future release of the Kepler Data Analysis Handbook. Briefly, the method:

1. Removes the median from each uncorrected flux time series,
2. Normalizes each median-removed flux time series by its RMS,
3. Calculates the correlation between RMS-normalized flux time series and selects the 50% most correlated stars as the reference ensemble,
4. Performs a Singular Value Decomposition (SVD) of the median-removed, median normalized flux time series of these most correlated stars. The CBVs are in this case the SVD principal components, though a more general nomenclature is used to allow non-SVD approaches to be used in the future. Users are provided with the leading 16 components. Users are also provided with the SVD principal values for each mod.out, to inform their decision about how many components to use in their fits.

To use the CBVs for least-squares fitting, subtract the median uncorrected flux from the uncorrected flux time series of interest and divide by the median. Since the basis is orthonormal, the linear least-squares fit coefficient of the nth CBV is simply the inner product of the median-removed, median-normalized uncorrected flux time series with the nth CBV. Subtract the fit to get the corrected (median-removed, median-normalized) flux time series.

Convenient CBV tools, including robust fitting and time window exclusion, are provided by the Guest Observer Office as part of PyKE and are available from
`http://keplergo.arc.nasa.gov/ContributedSoftwarePyKEP.shtml`. From that site there are instructions on how to install the software and specific instruction on how to use the tools to fit the CBVs. Users should note that, unlike the SOC Pipeline, these tools do not include scalar amplitude corrections for the fraction of target flux captured in the optimal aperture, or the fraction of the total flux in that aperture which is from the target star and not from its neighbors or unresolved background objects. Thes quantities (the flux fraction in aperture and the crowding metric, respectively) will be available for the optimal apertures computed by

the SOC as keywords in the FITS files at MAST starting in Fall, 2010. Approximate contamination values can currently be found at the MAST using the target search form.

Users will need to use **at least** the first two CBVs, since the generation method mixes a constant offset with the strongest non-constant component instead of strictly enforcing a constant first or second component; using only the first component would be like attempting a linear fit with a constant or slope term, but not both. Figure 9 shows that 8 or fewer components generally capture most of the systematic error for all mod.outs, though 16 are provided if users wish to make their own decisions.

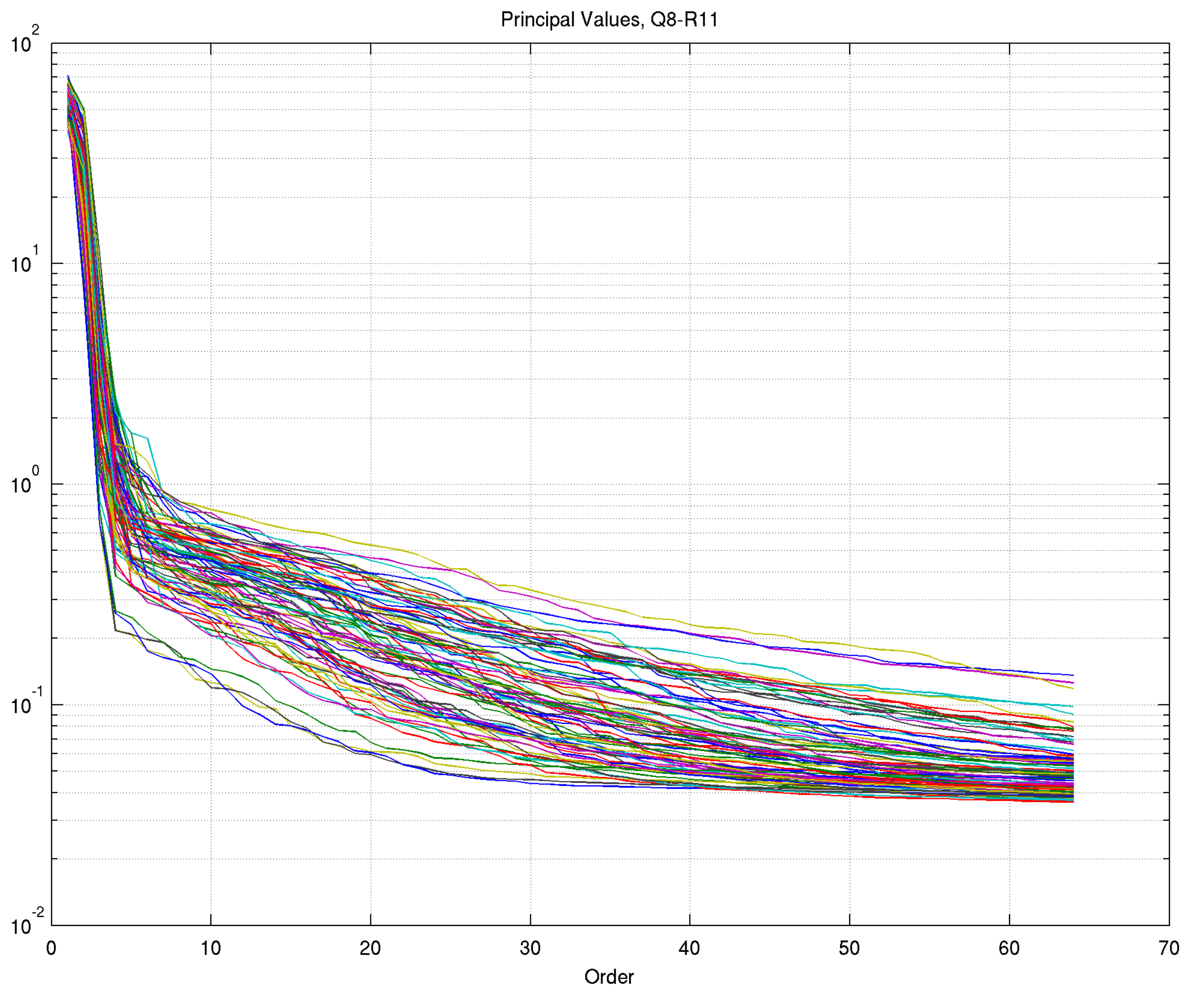

Figure 9: Principal values of SVD-extracted cotrending basis vectors for Quarter 8, Release 11, showing that most of the systematic error can be accounted for by the first 8 or fewer components. Users are provided with the first 16 components for each mod.out, as well as the principal values which are plotted in this Figure.

Cautions:

1. Channels with strong moiré pattern drift (see the Instrument Handbook, Section 6.7) have less stable weak (>5th order) CBVs, in the sense that random constant perturbations of the reference ensemble give results which differ by significantly more than a sign and a constant. For good channels, the stability of components to the 10th order has been verified.

2. Variable stars have not been explicitly filtered from the list of most correlated stars, so it is possible that some weaker components have been influenced by bright, highly variable stars. Users should be cautious about reporting results with the same period and phase as one of the basis vectors used in the fit.

Data Releases after Fall 2011, processed with SOC Pipeline versions 8.0 and higher, will provide the CBVs used internally by the released Pipeline. The CBVs provided in this release were generated by an early prototype of SOC Pipeline 8.0, and it is expected that some improvements to the methods for normalizing and choosing the reference ensemble will be made.

The cotrending basis vectors can be downloaded in FITS files from the MAST website (`http://archive.stsci.edu/kepler`) separate from your data download. There is one file per quarter containing 84 extensions, one for each channel. Each extension contains 16 basis vectors along with the cadence and MJD of the observations. The cadences found in the basis vector file match the number of cadences in the light curve file for that quarter. To ensure that your data was processed through the same version of the pipeline as the stars used to create the basis vectors, the keyword DATA_REL in your data's light curve file should match that found in the basis vector file. A new basis vector file will be provided each time the data are reprocessed.