Kepler Q1–Q12 TCE Release Notes

KSCI-19102-001

Data Analysis Working Group (DAWG)
Jessie L. Christiansen (Editor)

Prepared by: ______________________ Date: 12/6/12
Jessie L. Christiansen, Kepler Science Office, for the DAWG (page 4)

Approved by: ______________________ Date: 12/6/12
Jon Jenkins, Co-I for Data Analysis & DAWG Lead

Approved by: ______________________ Date: 12/6/12
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 and TCE Release Notes are the collective effort of the Data Analysis Working Group (DAWG), composed of Science Office (SO), Science Operations Center (SOC), and Guest Observer (GO) Office members as listed below:

Jon Jenkins*, Chair
Doug Caldwell*, Co-Chair
Barclay, Thomas
Bryson, Stephen T.
Burke, Christopher J.
Christiansen, Jessie L.
Clarke, Bruce D.
Coughlin, Jeffrey L.
Girouard, Forrest
Haas, Michael R.
Hall, Jennifer
Ibrahim, Khadeejah
Klaus, Todd
Kolodziejczak, Jeffery (MSFC)
Li, Jie
McCauliff, Sean D.
Morris, Rob
Mullally, Fergal
Quintana, Elisa V.
Rowe, Jason
Sabale, Anima
Sanderfer, Dwight T.
Seader, Shawn
Smith, Jeffrey Claiborne
Still, Martin
Tenenbaum, Peter
Thompson, Susan E.
Twicken, Joe
Uddin, Akm Kamal

The correct citation for this document is: J. L. Christiansen, J. M. Jenkins, D. A. Caldwell, T. Barclay, S. T. Bryson, C. J. Burke, B. D. Clarke, J. L. Coughlin, F. Girouard, M. R. Haas, J. R. Hall, K. Ibrahim, T. C. Klaus, J. J. Kolodziejczak, J. Li, S. D. McCauliff, R. Morris, F. Mullally, E. V. Quintana, J. Rowe, A. Sabale, D. T. Sanderfer, S. Seader, J. C. Smith, M. D. Still, P. Tenenbaum, S. E. Thompson, J. D. Twicken, and A. K. Uddin, 2012, Kepler Q1–Q12 TCE Release Notes (KSCI-19102-001).

Contents

1 Introduction

A Threshold Crossing Event (TCE) is a sequence of transit-like features in the flux time series of a given target that resembles the signature of a transiting planet to a sufficient degree that the target is passed on for further analysis. Each TCE is identified by the Kepler Identifier (KID) of the target under consideration, and the unique combination of period, epoch, and transit duration of the signal; a given target can contain multiple TCEs. These TCE Release Notes provide information specific to the release of the set of TCEs detected in the Q1–Q12 data set with the Kepler SOC Pipeline 8.3. For each KID with at least one TCE the pipeline produces a Data Validation (DV) report, and for each TCE found for that KID, it produces an additional one-page summary report; these reports are delivered as part of this TCE release. For further documentation on TCEs and the accompanying reports, see the NASA Exoplanet Archive.

Section 2 contains the details of the data products that were processed to identify this set of TCEs, the pipeline version used in the processing, and the set of detection thresholds that each potential signal had to pass in order to be identified as a TCE. Section 3 summarises the pipeline results. Section 4 briefly examines the resulting population of TCEs. Section 5 contains the supplementary information to this document.

2 How the TCEs were produced

2.1 Input Data Products

Table 1 lists the SOC Pipeline versions used for each of the modules that produce the data products used in the planet search. These include the Calibration (CAL) module, which produces the calibrated pixels, the Photometric Analysis (PA) module, which produces the simple aperture photometry flux time series, and the Presearch Data Conditioning (PDC) module, which corrects for instrumental systematics in the flux time series.

Table 1: Provenance of input data products for the Q1–Q12 search. The CAL and PA products are available at the Mikulski Archive Space Telescope (MAST)[a] and are documented in Kepler Data Release Notes 12–17. The PDC products will be delivered to the MAST and documented in a future Data Release.

Quarter	CAL	PA	PDC
1	7.0	8.0	8.3
2	7.0	8.0	8.3
3	7.0	8.0	8.3
4	7.0	8.0	8.3
5	8.0	8.1	8.3
6	8.0	8.1	8.3
7	8.0	8.1	8.3
8	7.0	8.1	8.3
9	8.0	8.0	8.3
10	8.0	8.0	8.3
11	8.0	8.0	8.3
12	8.0	8.1	8.3

[a] `http://archive.stsci.edu/kepler/`

2.2 Target List

The full list of 194436 targets is provided in the target summary file in the Supplement to these Notes. We exclude 2123 known eclipsing binaries from the search in this run, leaving 192313 targets that were submitted to the Transiting Planet Search (TPS) and Data Validation (DV) pipeline modules. The target summary file in the Supplement denotes for each target its progress through the pipeline.

2.3 Pipeline—TCE Detection

The TCEs documented in these Notes are produced by analysing the input data products described in Section 2.1 with the Transiting Planet Search (TPS) module and the Data Validation (DV) module. The SOC Pipeline 8.3 version of these modules was used in this processing; the details of TPS 8.3 are presented in Tenenbaum et al. (in prep).

The following parameters were defined for the TCE detections in this run:

- The minimum period searched is 0.5 days;
- The minimum number of transit events required for a Threshold Crossing Event is 3, which puts an upper limit on the period searched of half the total duration of the observations for a given target;
- For each target, 14 different box-car signal models with durations of 1.5, 2, 2.5, 3, 3.5, 4.5, 5, 6, 7.5, 9, 10.5, 12, 12.5, and 15 hours are matched to the data for identification of transit signals;
- For a given signal duration, the maximum allowed ratio of the signal duration to the period is 0.16, which prevents the pipeline from searching unrealistic combinations of signal duration and period;
- The variance window, over which the local noise is estimated, is set to 30 times the duration of the transit for which we are searching;
- The search is terminated for a given target once the maximum number of TCEs, 10, is reached.

In the SOC Pipeline 8.3 version of TPS, a target is required to pass the following tests in order to be identified as a TCE:

- A maximum Multiple Event Statistic (MES) $\geq 7.1\sigma$,
- A robust statistic $\geq 6.4\sigma$,
- A chi-square1 value of ≥ 7.0, and
- A chi-square2 value of ≥ 7.0.

These tests will be described in more detail by Seader et al. (in prep); in brief, they can be summarized as follows: TPS detects candidate transit signals through an adaptive, wavelet-based matched filter (Jenkins 2002, Jenkins et al. 2010). The maximum Multiple Event Statistic is the signal to noise ratio of a sequence of box-shaped pulses of the specified duration, best-matched phase and period, and with a depth provided by a generalized least squares fit to the data (i.e., a fit that accounts for the possibly non-white, non-stationary nature of the observation noise process); the robust statistic is the robust generalized least squares fit of the signal to the flux time series. Chi-square1 analyzes wavelet contributions to the points in the single event statistics time series that get combined to form the MES. Chi-square2 then analyzes the temporal contributions to the MES. Both compare what is expected, given the model, to what is present in the data.

2.4 Pipeline—TCE model fitting

The following parameters were defined for the TCE model fitting in this run:

- The four-parameter, non-linear limb-darkening coefficients from Claret & Bloemen (2011) from the ATLAS least-squares models are used, trilinearly interpolated over the effective stellar temperature, log g, and metallicity for each target;
- The maximum impact parameter in the model fitting is set to 1;
- If the estimated transit depth exceeds 0.15 (15%), the TCE is flagged as an eclipsing binary;
- For calculating the equilibrium temperature of the putative planet in the fit, a default albedo of 0.3 is used.

When deriving the model planet fits in this run, we use theoretical stellar models to produce values that are updated from the Kepler Input Catalogue (KIC; Brown et al. 2011), similarly to Batalha et al. (2012). Stellar parameters are derived by matching spectroscopic parameters (effective temperature, surface gravity and metallicity) to stellar evolution models (stellar mass, age, and Z). Spectroscopic parameters are based on spectral analysis with SME (Valenti & Piskunov 1996) or SPC (Buchhave et al. 2012), asteroseisomology (Huber et al. 2012), or the KIC, including the revision of effective temperature by Pinsonneault et al. (2012) for a fixed metallicity. For SME and SPC results we adopt uncertainties as reported. For KIC values we adopt errors in effective temperature from Table 7 of Pinsonneault et al. (2012) and 1σ errors of 0.4 in surface gravity (log g). We adopted the Yonsei-Yale stellar evolution models (Yi et al. 2001) to determine stellar parameters. The model matching was done by varying the stellar mass, age and Z and comparing the model-derived values of effective temperature, surface gravity and metallicity to the spectroscopic values with a chi-square statistic. An initial match was found by scanning in mass increments of 0.1 $M_{\odot}$, restricting ages from 0–14 Gyr, and identifying a best matching model. A Markov-Chain-Monte-Carlo routine was then seeded with this trial value of stellar mass, age and Z to determine posterior distributions. All stellar models with ages greater than 14 Gyr were excluded. In total, 100,000 chain elements were generated for each target. The models were also used to determine posterior distributions for the stellar radius, luminosity and mean stellar density. The final stellar parameters used by the pipeline are available at the Exoplanet Archive. Figure 1 shows the distribution of updated stellar parameters with comparision to the original KIC values.

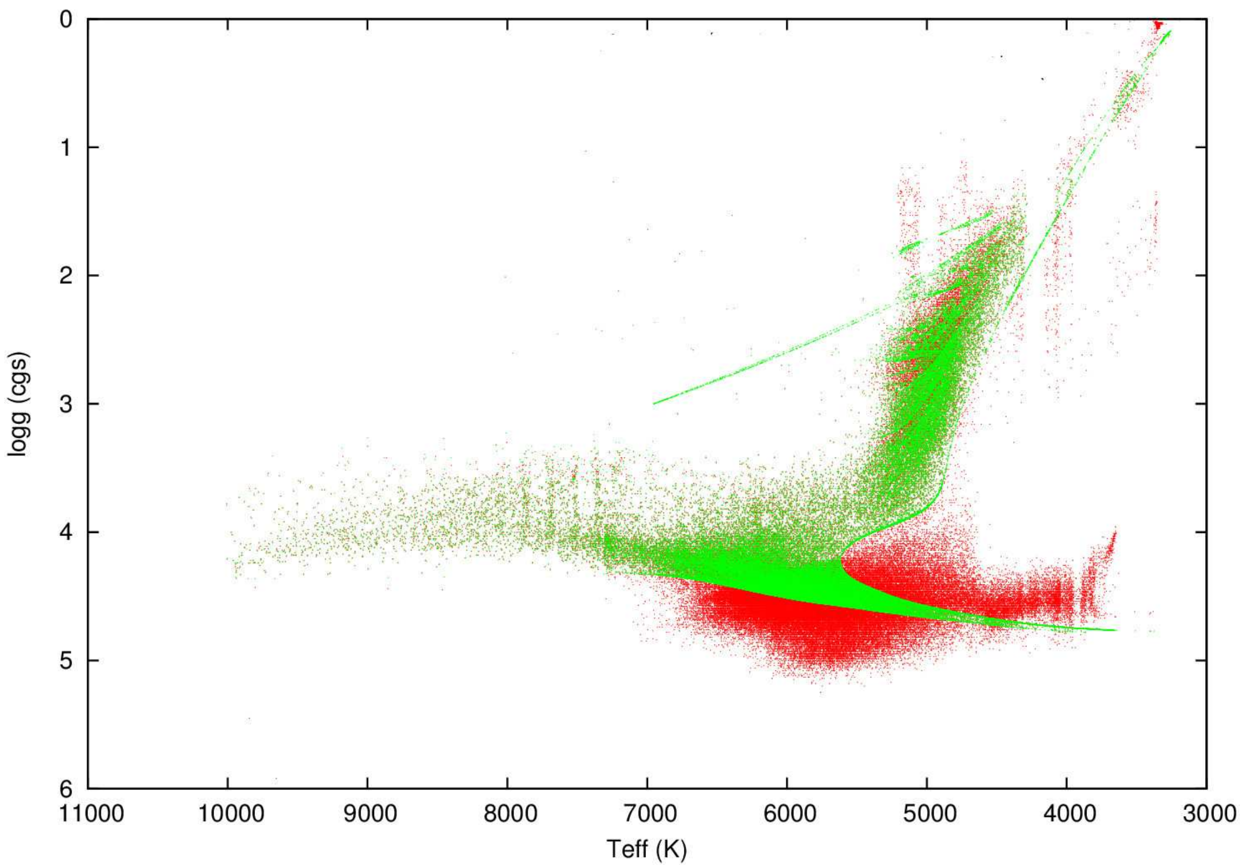

Figure 1: The distribution of stellar parameters from the KIC (in red) compared to the updated stellar parameters (in green) used for this processing.

3 Pipeline Results

In this Section, we document the TCEs identified in this run of the search pipeline. The results are summarized in Table 2. TPS completed for 192313 targets and found 11108 targets with at least one TCE given the set of detection thresholds in Section 2. Of these 11108 targets, 11089 were passed on to DV and searched for additional TCEs (i.e., multiple planets); two of these targets failed in processing. The 19 remaining TCEs not passed to DV are associated with internal calibration targets and are unlikely to represent true planet signatures. Calling TPS from within the DV module identified an additional 7319 TCEs, giving a total of 18406 TCEs. Table 2 shows the breakdown of TCEs that passed various validation tests within DV. The target summary file in the Supplement to these Notes tracks the progress of each Kepler ID through TPS and DV.

Table 2: Summary of TCEs detected

Attribute	Number
Input targets to TPS	192313
Successful completion in TPS	192313
Number of targets with TCEs found by TPS	11108
Total number of TCEs after DV	18406
Number of TCEs labelled as EBs	158
Number of TCEs that passed the All Transit Fit	17995
Number of TCEs that passed the All/Odd/Even Transit Fit	17413

Table 3 shows the breakdown of targets by the number of planets found per target. It also shows the number of TCEs that successfully passed the All Transit Fit, the number that passed the All Transit Fit with a measured SNR of $\geq 7.0\sigma$, and the number that passed the All Transit Fit with a measured SNR of $\geq 7.0\sigma$ and also passed the Odd/Even Transit Fits. In the Odd/Even Transit Fits, the odd-numbered and even-numbered transits are treated independently, in order to test for period aliasing due to eclipsing binaries.

Table 3: Statistics of TCEs by number of planets per target

# of Planets	Targets	TCEs	TCEs with Successful All Transit Fits	TCEs with Successful All Transit Fit SNR $\geq 7.0\sigma$	TCEs with All,Odd,Even fits, SNR $\geq 7.0\sigma$
1	6606	6606	6510	5550	5494
2	2797	5594	5464	5056	4927
3	1006	3018	2962	2743	2650
4	431	1724	1676	1513	1419
5	143	715	699	641	606
6	47	282	273	255	239
7	23	161	139	125	119
8	14	112	100	95	89
9	6	54	50	42	36
10	14	140	122	100	86
Total	11087	18406	17995	16120	15665

4 TCE properties

In this Section, we briefly demonstrate some properties of the TCE population produced by this run.

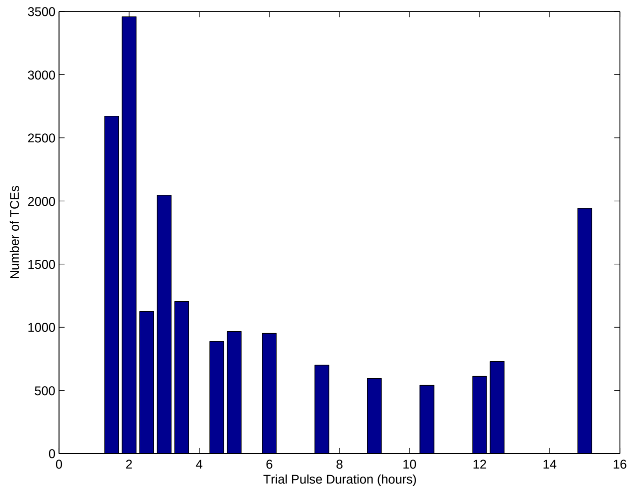

Figure 2: The distribution of the number of TCEs found for each of the trial pulse durations. The numbers peak towards the outer bounds of the range of durations searched.

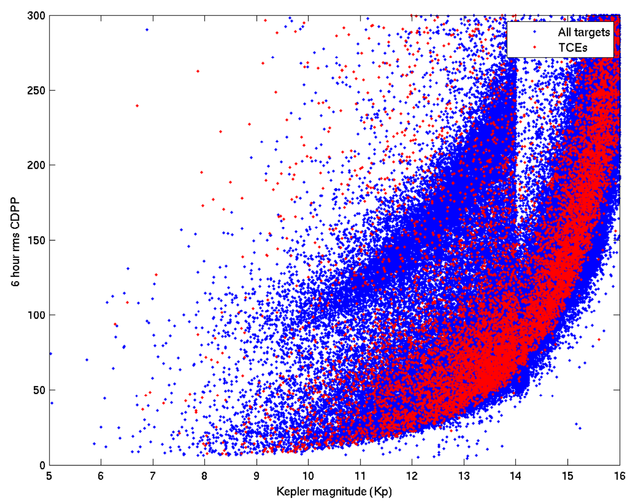

Figure 3: The distribution of the 6-hour rms Combined Differential Photometric Precision (CDPP) for the targets versus their Kepler magnitude. The blue points show all targets input to TPS, and the red points show all the TCEs at the end of DV. TCEs are distributed in magnitude and CDPP over similar ranges to the full set of targets.

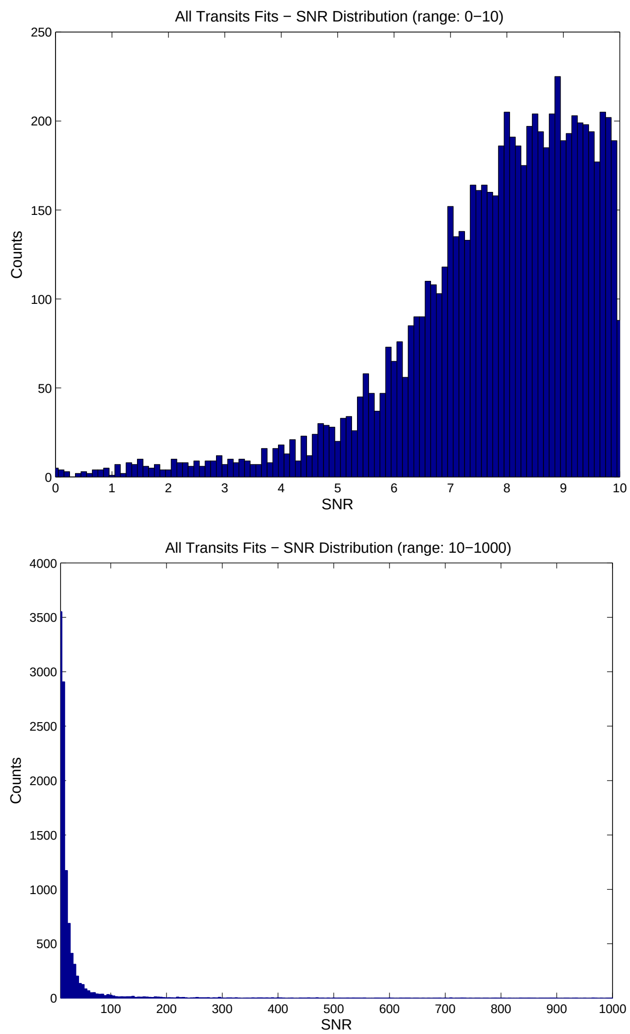

Figure 4: The distribution of the SNR of the transit model fitting in DV for the set of TCEs that completed the fit successfully. The distribution is broken into two ranges: SNR of 0–10σ in the upper panel, and SNR of 10–1000σ in the lower panel; scaling and bin sizes vary between panels.

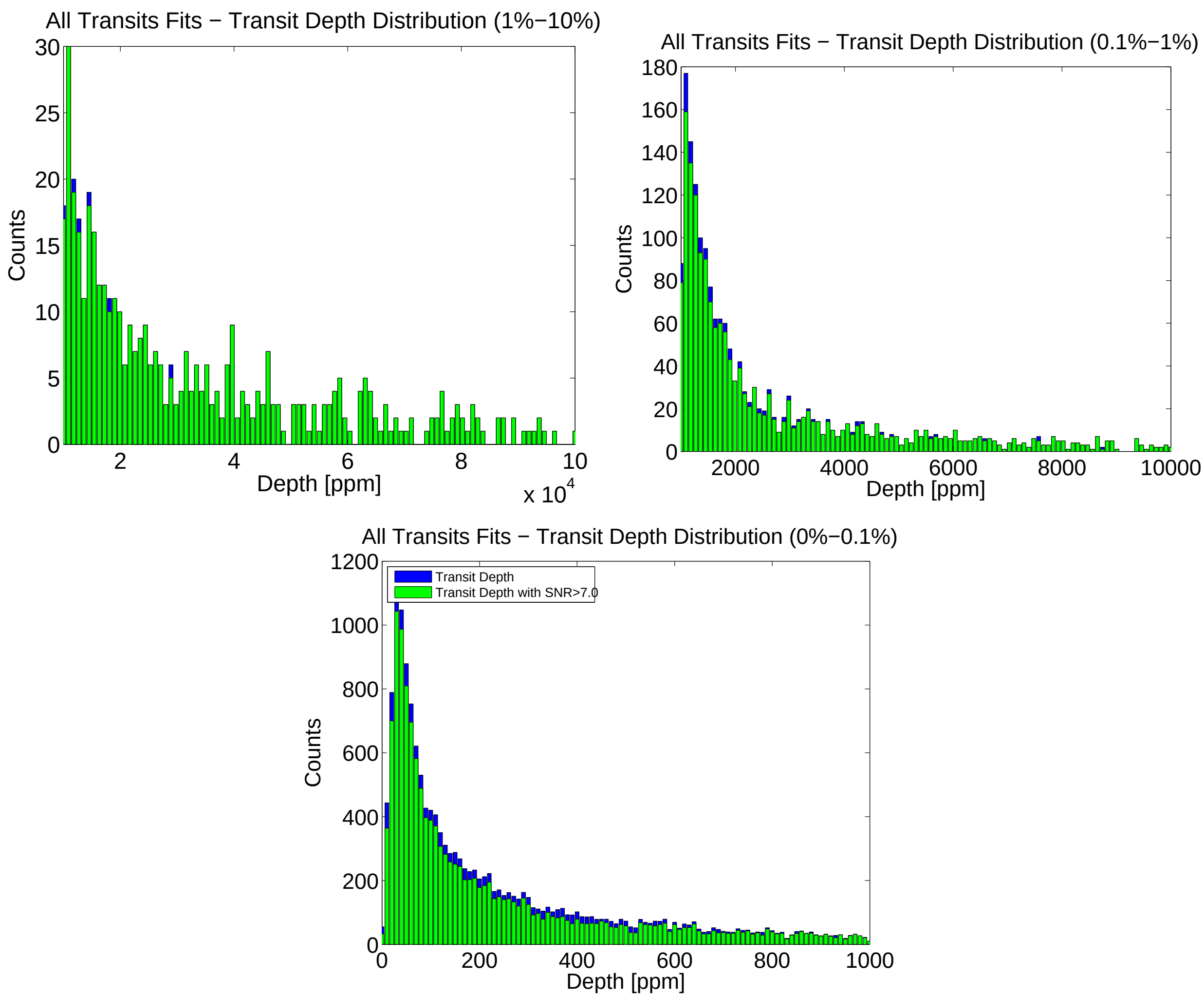

Figure 5: The distribution of the depth of the fit in DV for the set of TCEs that completed the fit successfully. The distribution is broken into three ranges: depths of 1–10% (10000–100000 ppm) in the upper left panel, 0.1–1% (1000–10000 ppm) in the upper right panel, and 0–0.1% (0–1000 ppm) in the lower panel; scaling and bin size vary between panels.

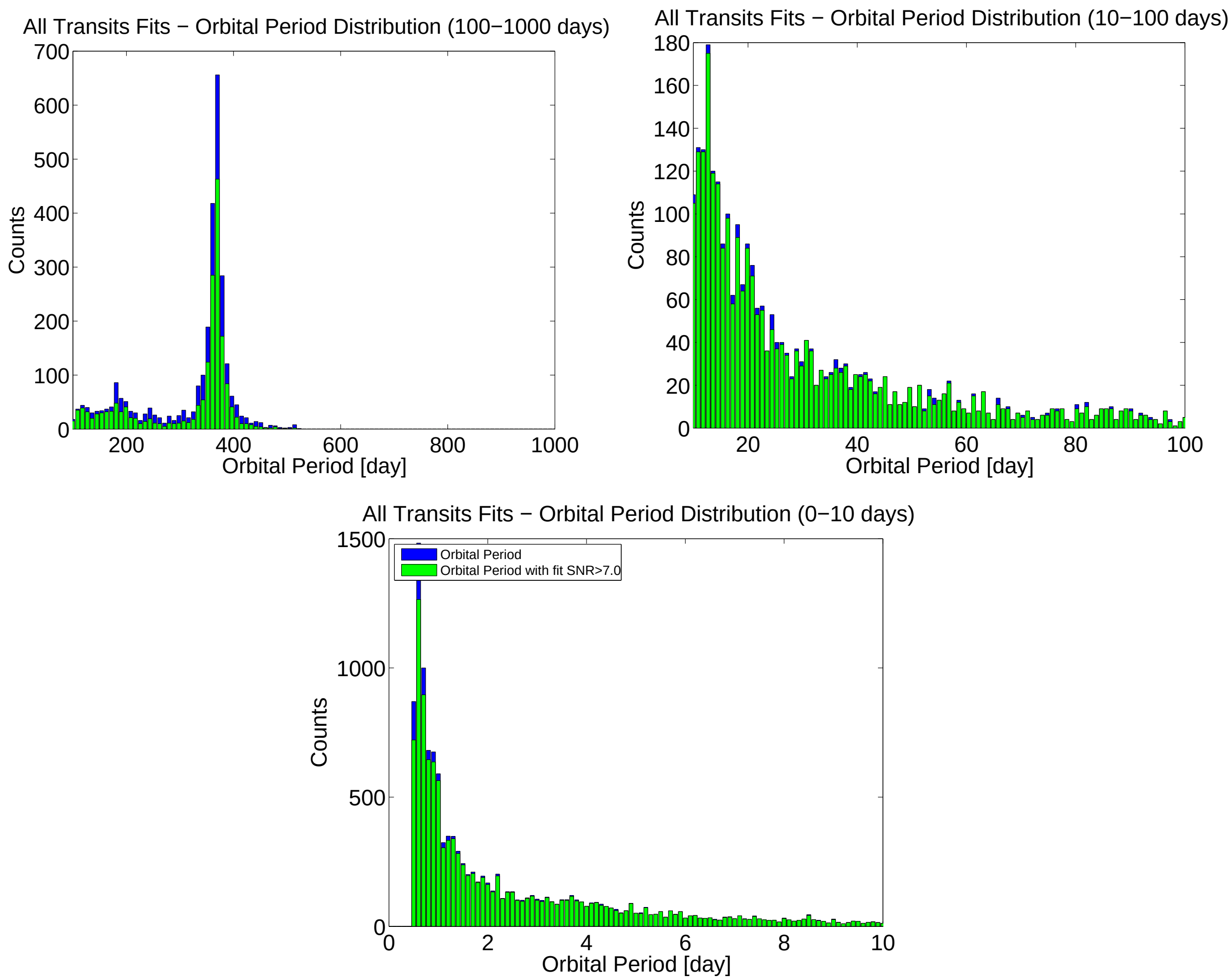

Figure 6: The distribution of the period of the fit in DV for the set of TCEs that successfully passed the transit model fitting. The distribution is broken into three ranges: periods of 100–1000 days in the upper left panel, 10–100 days in the upper right panel, and 0–10 days in the lower panel; scaling and bin sizes vary between panels. The significant peaks in the upper panel between 180–200 days and 360–400 days are due to a set of electronic artifacts in certain CCD channels that, given the quarterly rotation of the spacecraft, affect a given set of stars on a yearly timescale (i.e. once every four quarters the same set of stars returns to the CCD channel affected by the artifacts).

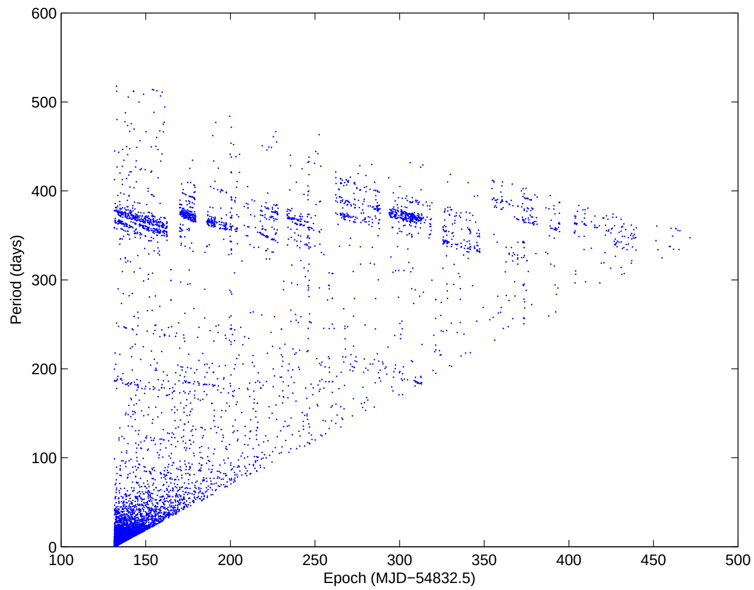

Figure 7: The distribution of the periods and epochs of the set of TCEs at the end of DV. Over-densities in points running vertically indicate an abundance of TCEs with the same epoch, indicating that a spacecraft event at that time is impacting the flux time series of the targets. Over-densities in points running horizontally (such as that seen around a ∼1 year period) indicates an abundance of TCEs with the same period (see Figure 6 for more details on the ∼yearly period TCEs). In general, clustering of TCEs implies the presence of an artifact causing spurious events, so TCEs in these clusters should be examined carefully for artifact-driven signals.

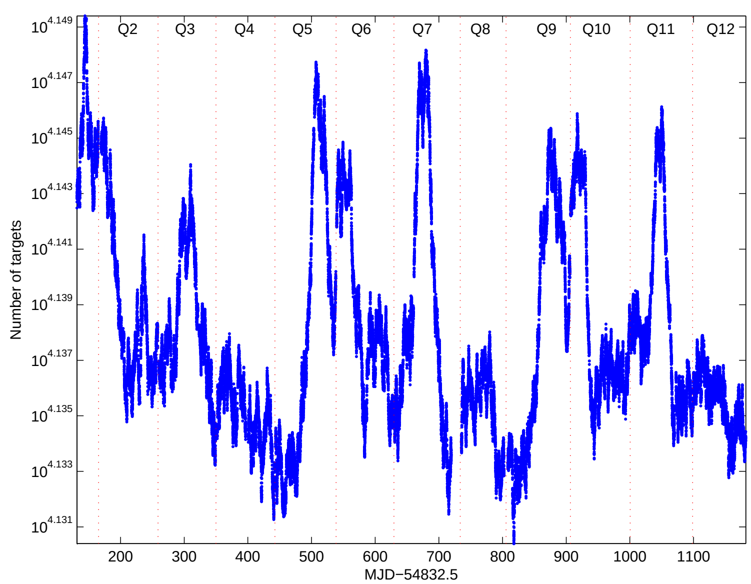

Figure 8: The number of targets, for each cadence, for which the TCE transits in that cadence. The vertical red lines indicate the quarterly boundaries. The peaks indicate epochs with an over-abundance of transit-like events—from purely astrophysical considerations we would expect this plot to be uncorrelated in time. There is noticeable structure with a period of ∼ 1 year (see Figure 6 for more details).

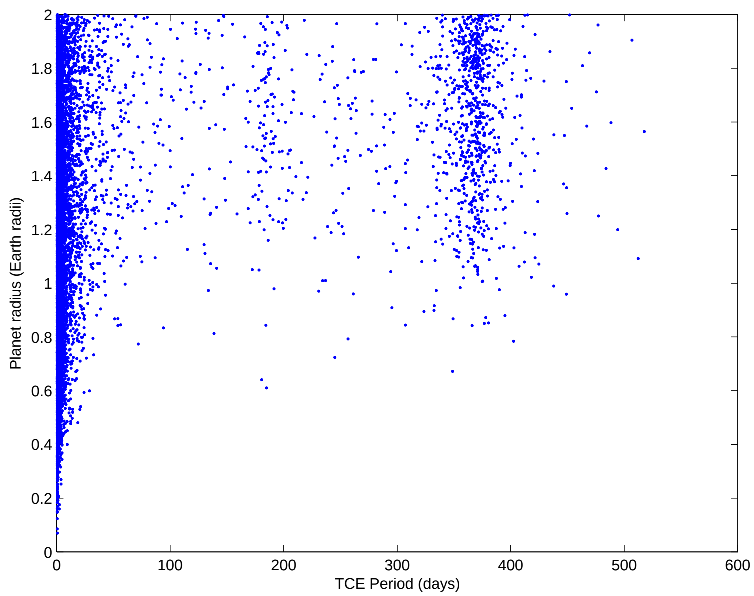

Figure 9: The distribution of period versus transit model fit radius for all TCEs. The over-abundance of periods at ∼1 year and ∼6 months is evident.

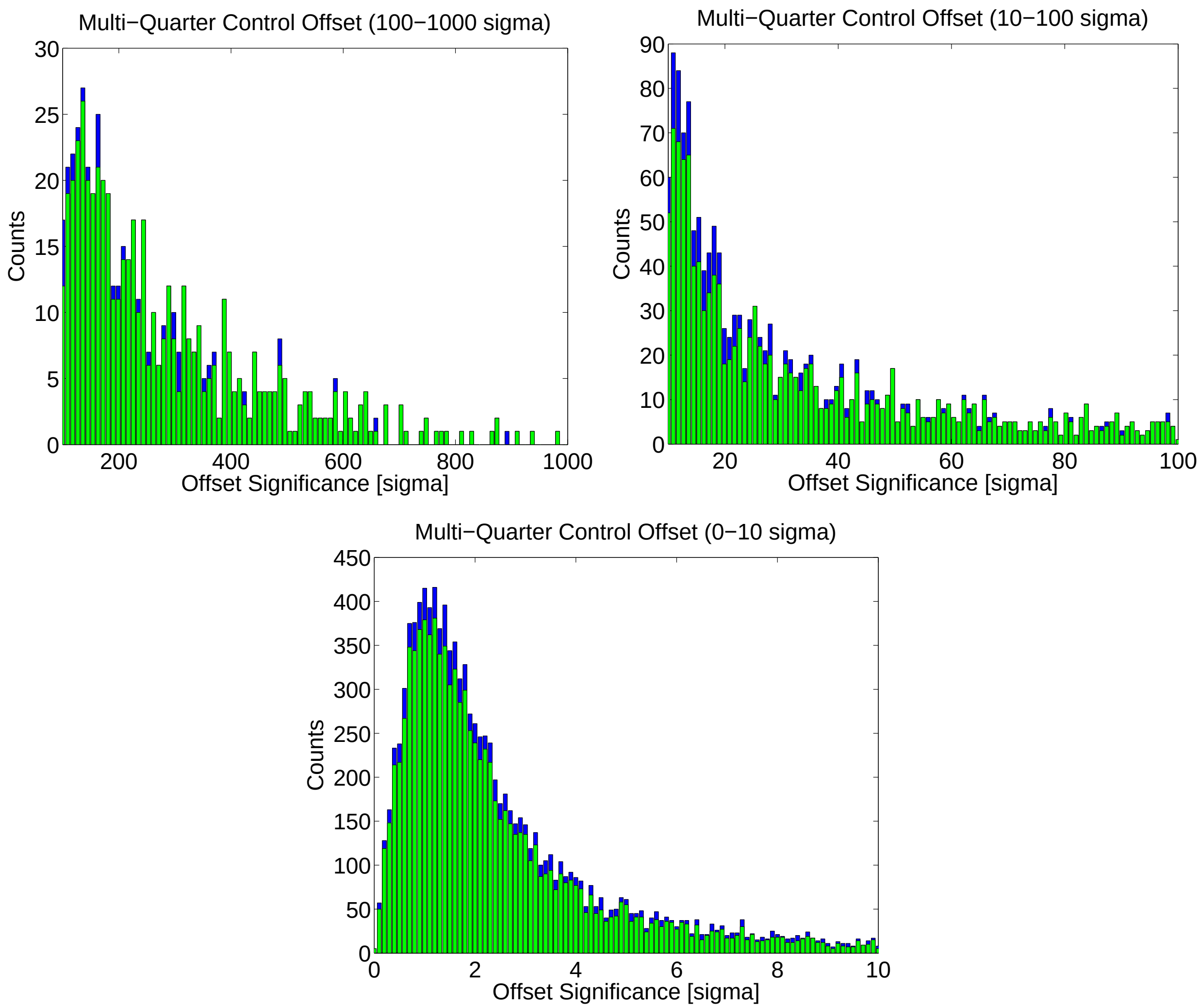

Figure 10: The measured offset in sky coordinates between the estimated position of the transit source and the measured out-of-transit location of the target star. The distribution is broken into three ranges: offsets of 100–1000σ in the upper panel left panel, offsets of 10–100σ in the upper right panel, and offsets of 0–10σ in the lower panel; scaling and bin sizes vary between panels. In general, offsets of $> 3\sigma$ and > 0.2 arcsec are expected to be caused by background eclipsing binaries.

5 Contents of Supplement

In addition to the TCEs and their DV reports, we are also delivering a table monitoring the progress of the full set of targets through the pipeline, so that users interested in completeness/population studies have complete knowledge of the set of targets represented in the list of TCEs. The table contains the full list of targets, including custom targets and EBs, that was submitted to TPS in this run: Q1-Q12_121119_tps_target_list.txt.

The columns are:

1. Kepler ID,
2. Integer bit flag indicating progress though the pipeline.

The integer bit flag indicates one of the following states:

- 0: Target was on the original full list of targets;
- 1: Target was processed through TPS;
- 2: Target produced a TCE in TPS;
- 4: Target was forwarded to DV and processed successfully;
- 8: Target was forwarded to DV and failed processing.

6 References

Batalha, N. M. et al. 2012, arXiv:1202.5852
Brown, T. M. et al. 2011, AJ, 142, 112
Buchhave, L. A. et al. 2012, Nature, 486, 375
Claret, A., & Bloemen, S. 2011, A&A, 529, 75
Huber, D. et al. 2012, ApJ, 760, 32
Jenkins, J. M. 2002, ApJ, 575, 493
Jenkins, J. M. et al. 2010, ApJ, 713, 87
Pinsonneault, M. H. et al. 2012, ApJS, 199, 30
Valenti, J. A. & Piskunov, N. 1996, A&AS, 118, 595
Yi, S. et al. 2001, ApJS, 136, 417