Kepler: A Search for Terrestrial Planets

Description of the TCERT Vetting Products
for the Q1-Q17 DR24 Catalog
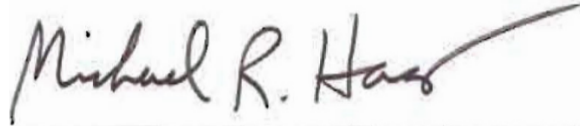
KSCI-19104
June 22, 2015

NASA Ames Research Center
Moffet Field, CA 94035

Prepared by: _____ Date: __June 22, 2015__
                Jeffrey L. Coughlin, Kepler Science Office

Approved by: _____ Date: __June 22, 2015__
                Michael R. Haas, Science Office Director

Table of Contents

PREFACE

The Q1-Q17 DR24 TCERT Vetting products are a collection of plots and diagnostics used by the Threshold Crossing Event Review Team (TCERT) to evaluate threshold crossing events (TCEs). While designation of Kepler Objects of Interest (KOIs) and classification of them as Planet Candidates (PCs) or False Positives (FPs) is completely automated via a robotic vetting procedure for the Q1-Q17 DR24 planet catalog described in Coughlin et al. (2015), these products help to evaluate those robotic decisions for individual objects. For each TCE identified in the Q1-Q17 DR24 transit search, these products collect in one readily accessible place: (a) the DV one-page summary, (b) selected pertinent diagnostics and plots from the full DV report, and (c) supplemental plots and diagnostics not included in the full DV report, including an alternate means of data detrending.

This document describes each plot in detail, and how it can be used to evaluate the various data sources to determine if a TCE is a PC or FP. Vetting products exist for every Q1-Q17 DR24 TCE. The population of TCEs that was vetted is described in Seader et al. (2015) and Coughlin et al. (2015).

It should be noted that the TCERT group operates under a philosophy of "innocent until proven guilty", meaning that a specific reason must be enumerated to classify an object as a FP, else it is presumed to be a PC. Any questions about these products, or the content therein, can be directed to the Kepler Science Office and/or the TCERT, both of which can be contacted at kepler-scienceoffice@lists.nasa.gov.
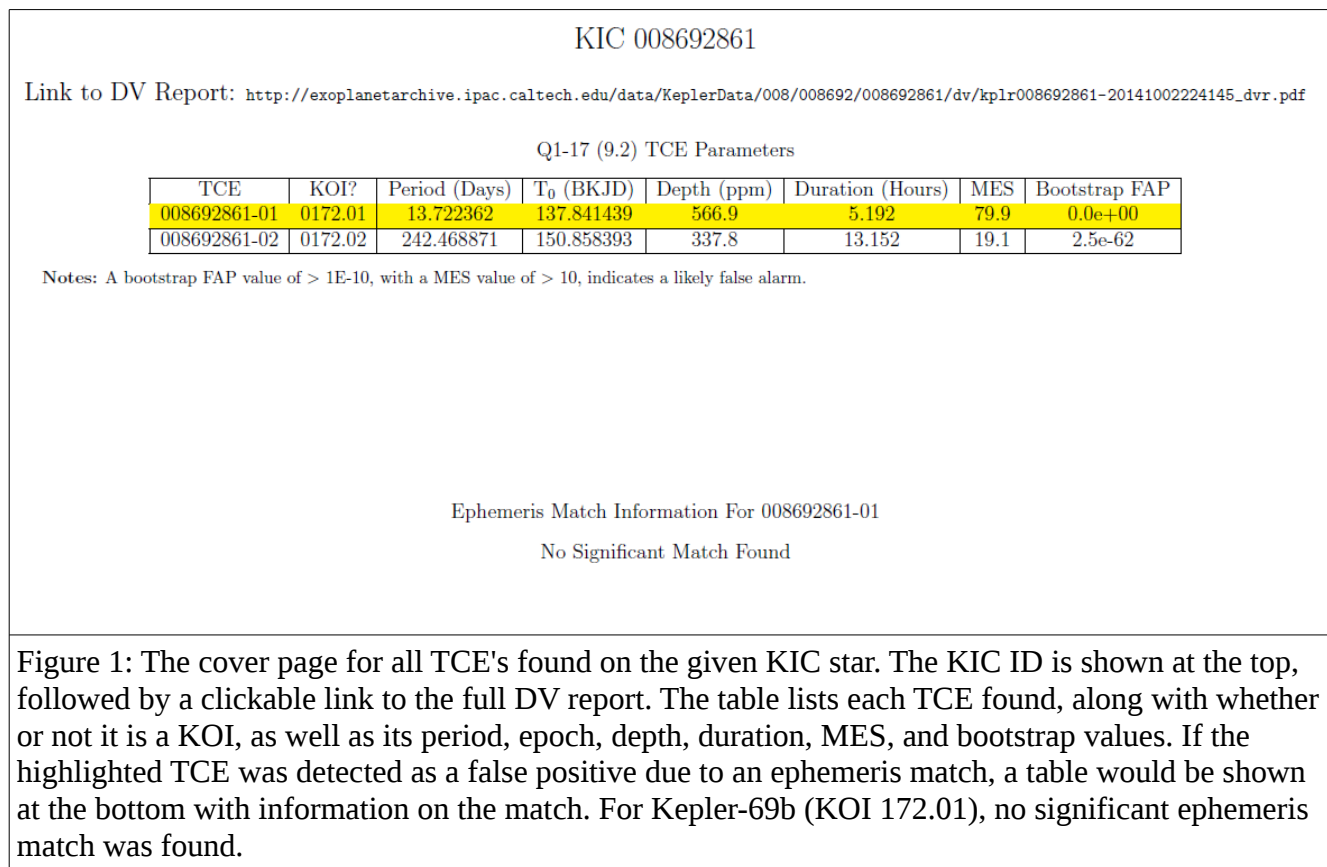
When citing this document, please use the following reference:

Coughlin, J. L., Bryson, S. T., Burke, C. J., Catanzarite, J. H., Haas, M. R., Jenkins, J. M., Li, J., Mullally, F.., Rowe, J. F., Seader, S. E., Thompson, S. E., and Twicken, J. D. 2015, Description of the TCERT Vetting Products for the Q1-Q17 DR24 Catalog (KSCI-19104-001).

KIC/TCE Cover Page

All TCEs originating from the same KIC star are grouped into a single PDF. This cover page lists the KIC ID for the PDF, a link to the full DV report, and a table for every TCE belonging to the given KIC with each TCE's period, epoch, depth, duration, MES, bootstrap value, and whether or not it is designated as a KOI. If the TCE was detected as a false positive due to an ephemeris match, a table containing information about the match is shown at the bottom of the cover page.

Throughout this document, we will show examples from the Kepler-69 system (KIC 008692861), a well-known system with two confirmed planets. We focus on Kepler-69b (Q1-Q17 DR24 TCE 008692861-01) as it is the first TCE of the system with a higher signal-to-noise. The cover page for KIC 008692861 is shown in Figure 1.

KIC 008692861

Link to DV Report: http://exoplanetarchive.ipac.caltech.edu/data/KeplerData/008/008692/008692861/dv/kplr008692861-20141002224145_dvr.pdf

Q1-17 (9.2) TCE Parameters

TCE	KOI?	Period (Days)	T_0 (BKJD)	Depth (ppm)	Duration (Hours)	MES	Bootstrap FAP
008692861-01	0172.01	13.722362	137.841439	566.9	5.192	79.9	0.0e+00
008692861-02	0172.02	242.468871	150.858393	337.8	13.152	19.1	2.5e-62

Notes: A bootstrap FAP value of $> 1E-10$, with a MES value of > 10, indicates a likely false alarm.

Ephemeris Match Information For 008692861-01

No Significant Match Found

Figure 1: The cover page for all TCE's found on the given KIC star. The KIC ID is shown at the top, followed by a clickable link to the full DV report. The table lists each TCE found, along with whether or not it is a KOI, as well as its period, epoch, depth, duration, MES, and bootstrap values. If the highlighted TCE was detected as a false positive due to an ephemeris match, a table would be shown at the bottom with information on the match. For Kepler-69b (KOI 172.01), no significant ephemeris match was found.

Page 1 – DV Summary

The Kepler Q1-Q17 DR24 TCERT Vetting Products have 16 additional pages for each TCE. Page 1 contains the Data-Validation (DV) one-page summary that highlights some of the tests and figures from the associated full DV report. DV reports are always available and people are encouraged to make use of them when the summaries prove insufficient.

The DV summary shown in Figure 2 is the first page of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. Each panel will be discussed separately below, but briefly, the top panel shows the photometric time-series, and underneath is the photometric time-series folded on the period of the candidate event. To the right is the best candidate for a secondary eclipse in the data. Below and to the left is the phase-curve zoomed in on the transit event. To the right is the phase-curve zoomed in on the transit event, but binned and in the whitened domain. In addition, this plot shows the residuals from the best-fit model and the whitened light curve at a phase of 0.5. The bottom row has the odd-even transit plot on the left, a centroid snapshot in the center, and a table of model parameters to the right.
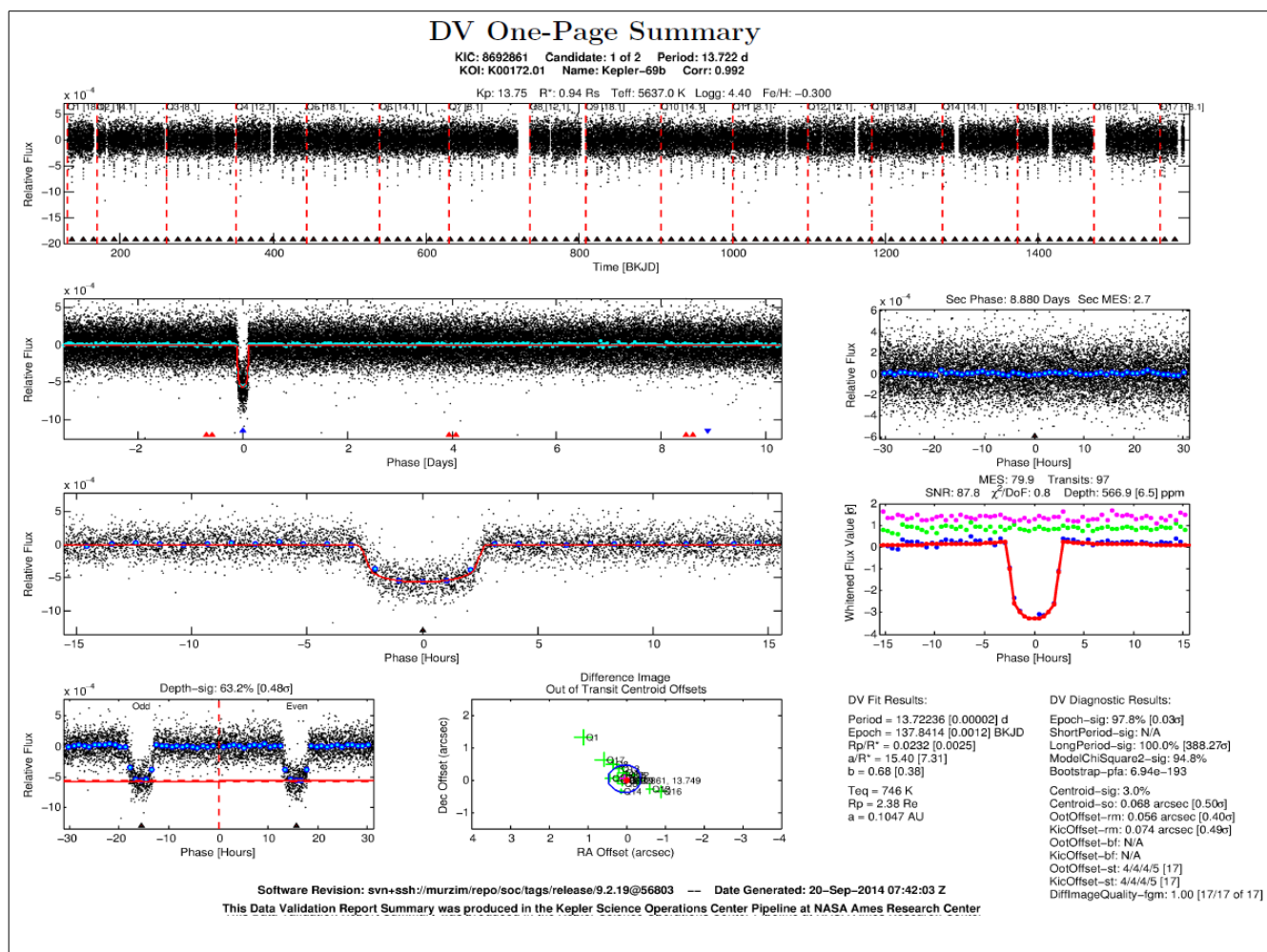

Figure 2: Page 1 of the Kepler Q1-Q17 DR24 TCERT Vetting Products showing the DV Summary page. The example shown here is for Kepler-69b (Q1-Q17 DR24 TCE 008692861-01).

Detrended Lightcurve

The top panel of the DV summary is shown in Figure 3. This figure shows the detrended *Kepler* long-cadence photometric time-series. Above this plot the user can find the Kepler identifier (in this example 008692861), the candidate number (1 of 2) and the measured orbital period (13.722 d). If the DV module federated the period and epoch of the TCE to an existing KOI at the time of generation, the next line will contain the KOI number, Kepler confirmed planet number if it exists, and ephemeris correlation (Corr.) metric. On the last line is the Kepler magnitude (Kp:13.75), the estimated stellar radius in solar units (R∗: 0.94 Rs), the stellar effective temperature (Teff: 5637.0 K), the stellar surface gravity in cgs units (logg: 4.40) and the metallicity scaled to solar ([Fe/H]: -0.300). The plot itself shows relative flux versus time. The red, dashed vertical lines mark the beginning of each quarter of observations. The quarter and CCD channel number are marked next to each line at the top of the panel. In the example shown, the target was located on channel [8.1] in Q3, on channel [12.1] in Q4, etc. Along the bottom axis there are triangle markers that indicate the location of each transit. For very short orbital periods, the markers along the bottom overlap and appear as a thick line.
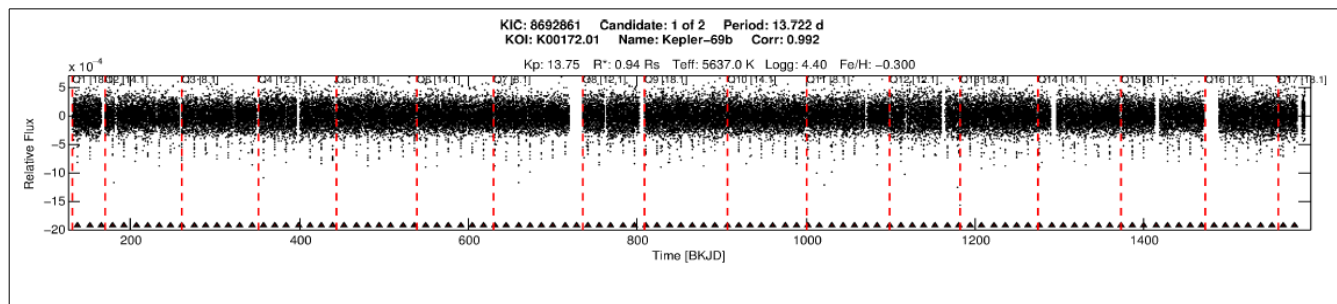

Figure 3: The detrended *Kepler* long-cadence photometric time-series. The red, dashed vertical lines mark the beginning of each quarter of photometric observations. The upward-facing, filled, black triangles along the bottom mark the location of each individual transit. The numbers in brackets next to each quarter designate the CCD channel and module that the star fell on each quarter.

Only the deepest transit events will be apparent in this figure; most moderate to low S/N events only become apparent in the phased and binned data plots. However, users should be alert for significant outliers that align with the marked transits or changes in the noise properties with a periodicity similar to the transit period.

Phase-curve

Below the plot of the photometric time series, one finds the photometric phase-curve as shown in Figure 4, where the observations have been folded at the orbital period of the candidate transit event. The best-fit model is shown as a red line, and binned points are shown via blue-circled, cyan-filled points. Along the bottom, the upward-facing, blue triangle denotes the location of the primary eclipse, and the downward-facing, blue triangle denotes the location of the strongest secondary event detected. Upward-facing triangles of other colors denote the location of transits of other TCEs in the phase space of the current TCE under examination.
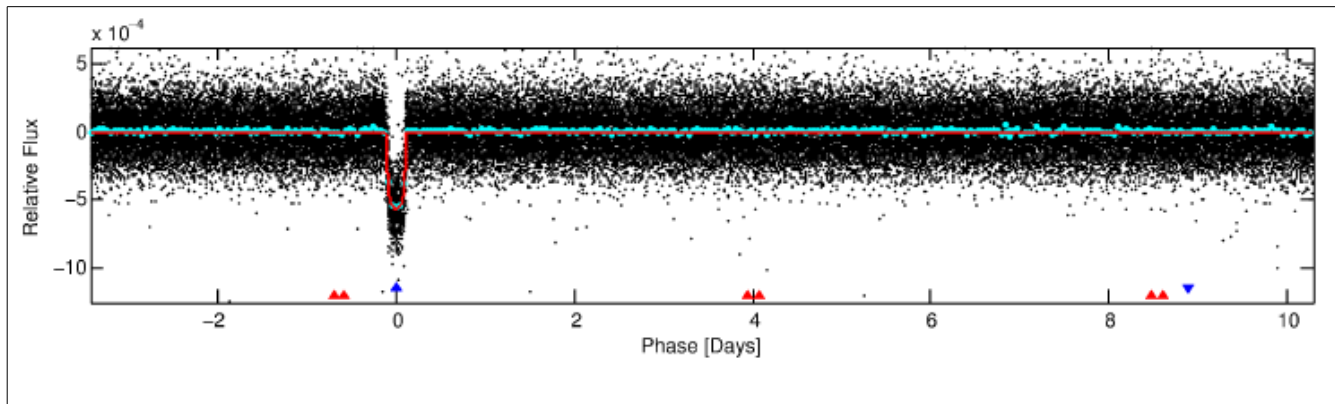
Figure 4: The photometric data folded at the orbital period of the candidate transit event. The red line shows the best-fit transit model. The blue-circled, cyan-filled points show the data binned. The transit event is located at phase 0.0 with an upward-facing, blue triangle. The downward-facing, blue triangle marks the location of the strongest secondary eclipse candidate via the weak secondary test. Upward-facing triangles of other colors show the locations of the primary events of other TCEs in the system, in the phase space of the current TCE under examination.

This figure is useful for identification of the transit and possible secondary events, which may or may not generate their own TCEs. One should also take note of the transit photometry and ask whether the event is consistent with a planetary transit event or appears to be due to stellar variability.

In the third row on the left, you will find an additional plot of the phase-folded photometry, but this time zoomed in on the transit event as shown in Figure 5. This figure is useful for judging the S/N of the transit event and its shape. While the *Kepler* pipeline fitter and DV module work with photometry in the whitened domain, this plot is shown in the flux domain, detrended using a moving median filter. Due to the window size of the median filter, do not over-interpret any increases in the flux before ingress or after egress as signs of a photometric false-positive. We recommend using Figure 6 (discussed below), which presents the data and fit in the whitened domain, for FP identification. Figure 5 is useful for identification of the transit event, validation of the reported SN, and noting anomalies.

Figure 6 shows the whitened, binned, and phased photometric time-series zoomed in on the transit event, and is often one of the most useful figures for evaluating the data and the transit model fit. The data is presented as whitened, so you are seeing the data as fit by the whitened transit model. If you examine Figure 6, you will see increases in the flux before ingress and after egress as the transit photometry and the model are distorted by the whitener. The user should also be aware that the transit model is restricted to have an impact parameter less than 1.0. Thus, 'V'-shaped transits due to stellar binaries and grazing transits may be poorly fit, especially when the impact parameter, b, approaches 1.0. If the model is a poor representation of the whitened transit photometry, then caution should be exercised when interpreting statistics that rely on the model fit such as the signal-to-noise, transit depth, and odd-even transit comparison.
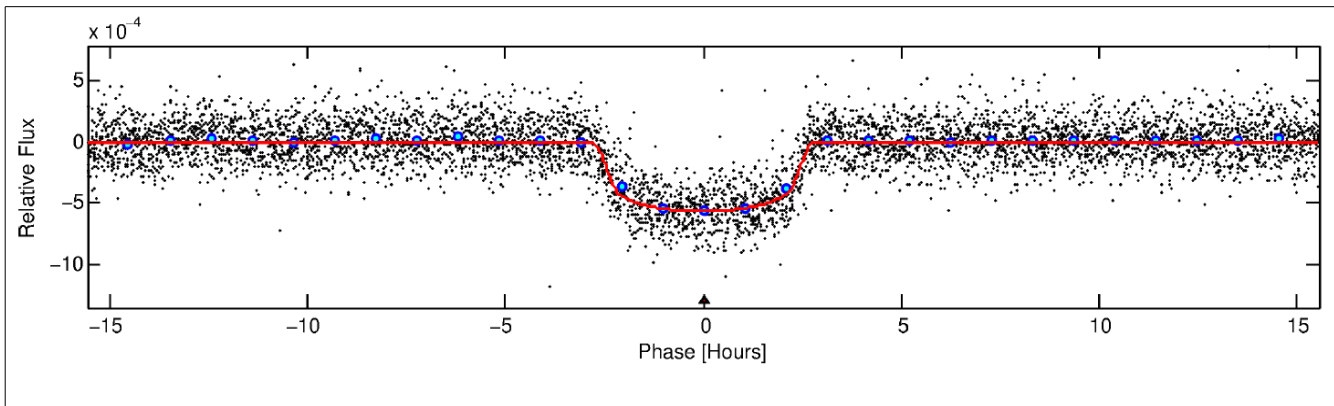
Figure 5: The phased photometric time-series zoomed in on the transit event. The blue-circled, cyan-filled points show the data binned and the red line shows the best fit transit model.
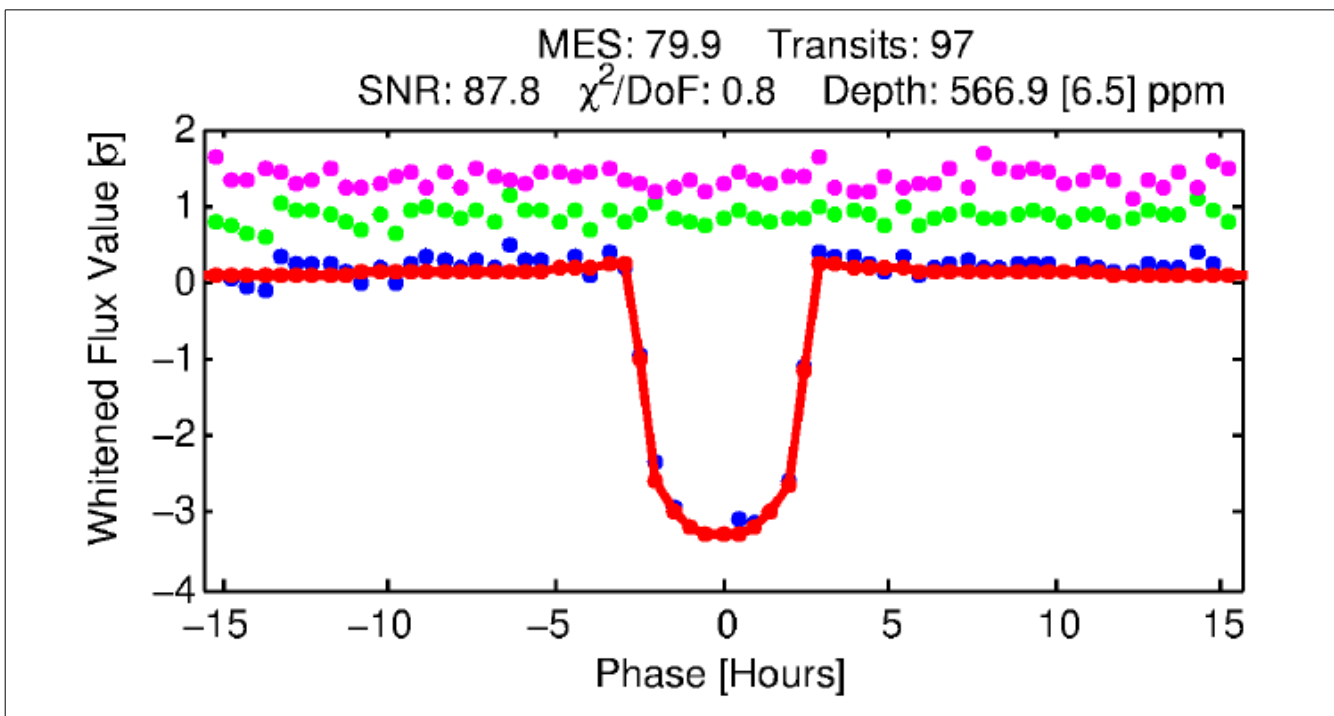

Figure 6: The whitened, binned, and phased photometric time-series zoomed in on the transit event is shown by the blue points. The red line with points shows the best-fit, whitened transit model. The green points show the residuals after subtracting the best fit transit model. The magenta markers show the phased data at half the orbital phase where a potential occultation may occur for a circular orbit.

Odd-even Test

Kepler photometry can provide a precise measurement of the transit depth. With multiple transits, precision levels can reach the part-per-million (ppm) level. Determination and comparison of the transit depths for odd-numbered transits (the first, third, fifth, etc) with the transit depths for even-numbered transits (the second, fourth, sixth, etc) can provide a powerful test of whether a transit event is consistent with an extrasolar planet interpretation or not. A common false-positive is an eclipsing binary system composed of two stars with nearly equal mass, size, and temperature. This type of false-positive may be detected by TPS at half the true period of the system, thus showing alternating eclipses with slightly different depths.

Figure 7 presents the *Kepler* photometry for the odd-numbered transits on the left panel butted up against the even-numbered transits in the right panel. The difference in the odd- and even-numbered transits is determined by the DV transit model fit to each set independently, and is used to estimate the probability that the depths are different. The calculation is presented above the odd-even plot as shown in Figure 7. The Depth-sig metric determines if the depths are similar (100%) or different (0%) and is alternately reported as a sigma confidence level. This calculation depends on the transit model being a good representation of the transit photometry (see Figures 4-6) and assumes that there are no systematics present in the photometry, such as data outliers.

The odd-even test will fail when astrophysical phenomena, such as star-spots, are present on the host star. These become an important concern when only a few transits are present. Stellar activity may produce systematic differences in the transit depths that are not properly reflected in the calculation of significance for the depth difference. Also, due to natural seasonal depth variations associated with Kepler's orbit, the odd-even test is not effective at periods greater than ~90 days when there is only one or fewer transits per quarter.

Weak Secondary Test

A diagnostic test called the Weak Secondary Test was introduced starting with the Q1-Q16 activity. First, the primary transit signal is removed, and the whitening filter is re-applied to the light curve. The Transiting Planet Search (TPS) algorithm is then run on the resulting data with the same search duration as the primary TCE. Finally, the resulting single event detection time-series is folded at the same period as the primary TCE. This produces, among many other useful quantities, the value and phase of the Multiple Event Statistic (MES) for the strongest transit-like signal at the TCE's period, aside from the primary TCE itself.

The phased data centered on the secondary eclipse candidate is shown in Figure 8, with black dots representing the raw data and blue-circled, cyan-filled points representing phase-binned averages of the data. The values of the MES and phase of the secondary eclipse candidate (in days relative to the TCE) are displayed at the top of Figure 8. If the MES is greater than 7.1 (the formal mission detection threshold) then it is colored red to indicate the secondary eclipse candidate is statistically significant.
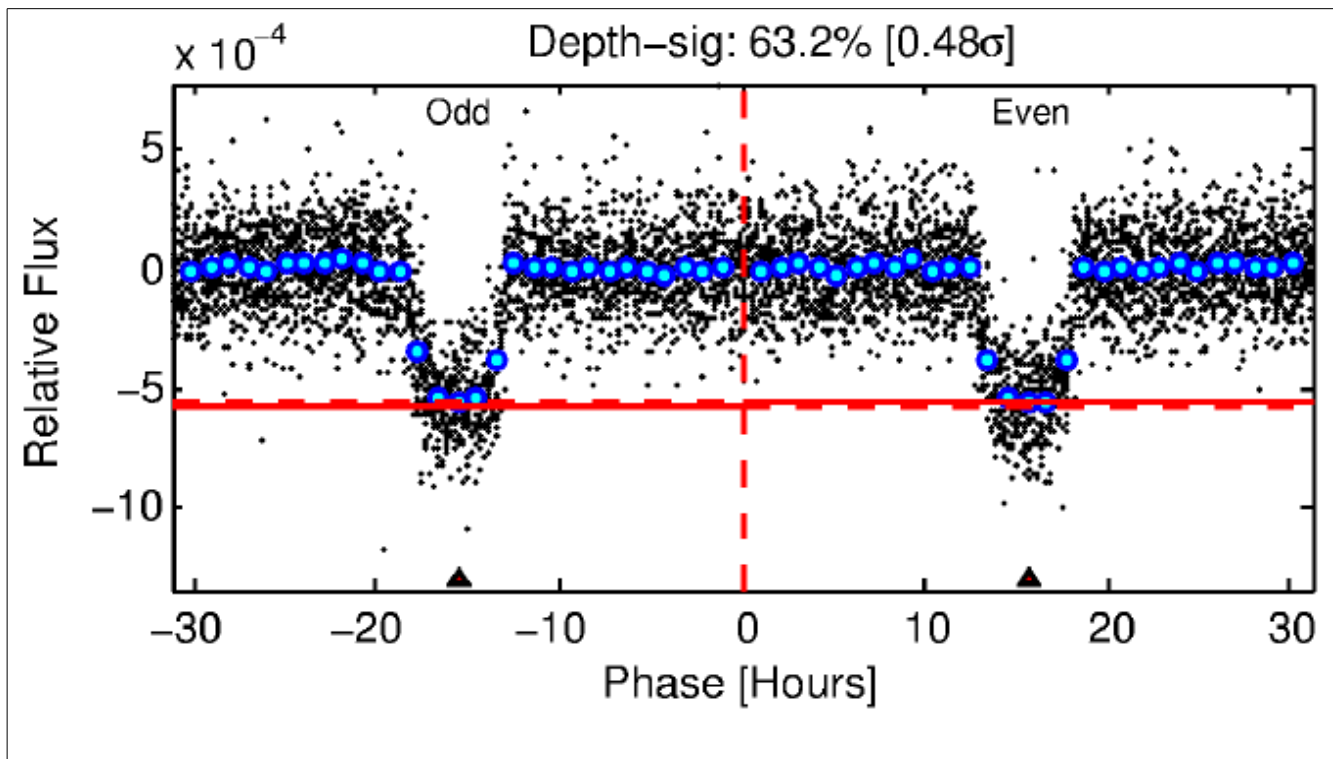
Figure 7: The odd numbered transits are folded on top of one another and plotted in the left panel and the even numbered transits are folded on top of one another and plotted in the right panel. The vertical dashed line separates the panels. The horizontal solid lines represent the depth of the odd- and even-numbered transits separately. The two horizontal dashed lines indicate the +/-1σ error limits. (In this example, the errors are small in comparison to the resolution of the plot.) The blue-circled, cyan-filled points show the data binned in phase. The upward pointing triangles mark the location of the center of the transit as measured by the model. In the case of this TCE there is no significant difference between the odd- and even-numbered transits.

This plot helps assess whether there is a statistically significant secondary event in the phased light curve, which may indicate the candidate is an eclipsing binary, and thus not a viable planet candidate. The only exception is if the secondary eclipse could be due to planetary emission, but typically this is only observed for hot Jupiters with very small eclipse depths and very short periods. This plot and the characteristics of the secondary eclipse candidate may also help to highlight transit-like artifacts in the data, which may cast doubt on the uniqueness of the primary TCE and thus validity of the candidate. Vetters should be careful in the interpretation of this signal though, as often there remains significant systematics in the Kepler DV data, and this test occasionally triggers on the edge of the primary TCE itself.
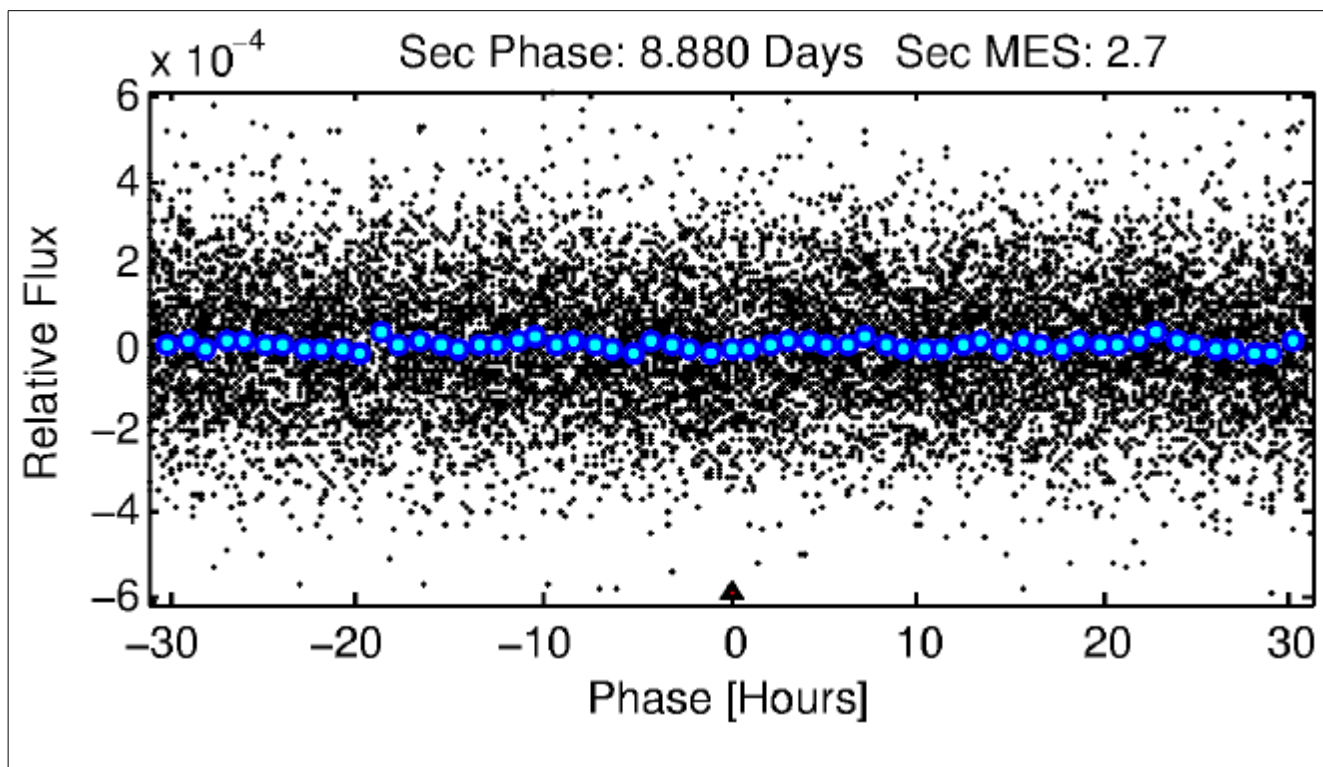
Figure 8: The strongest secondary eclipse candidate found by the weak secondary test, with the resulting phase and MES displayed above. The blue-circled, cyan-filled points represent phase-binned averages of the data. In the case of this TCE there is no significant secondary eclipse detected.

Out of Transit Centroid Offsets

Figure 9 shows the PRF centroid offset with the RA Offset in arcseconds on the x-axis, and the Dec Offset in arcseconds on the y-axis. For each quarter, two separate pixel-level images of the source are computed, one using the average of only the in-transit data, and the other using the average of data just outside of transit. The difference of the in- and out-of-transit images is used to produce a difference image, which produces a star-like image at the location of the transit signal.

The Kepler Pixel Response Function (PRF) is the Kepler point spread function combined with expected spacecraft pointing jitter and other systematic effects (Bryson et al. 2010). The PRF is fit separately to the difference and out-of-transit images to compute centroid positions. The fit to the difference image gives the location of the transit source, and the fit to the out-of-transit image gives the location of the target star (assuming there are no other bright stars in the aperture). Subtracting the target star location from the transit source location gives the offset of the transit source from the target star. This calculation is performed on a per-quarter basis, and the quarterly offsets are shown as green cross-hairs and labeled with the quarter number, where the length of the arms of each cross-hair represents the 1σ error in RA and Dec. Asterisks in the image show the location of known stars in the aperture, with the red asterisk being the target star. The coordinate system of the plot is chosen so that the target star is at

(0,0). A robust fit (i.e., an error-weighted fit that iteratively removes extreme outliers) is performed using all the quarterly centroid offsets to compute an average in-transit offset position, and is shown with 1σ error bars as a magenta cross. A dark blue circle is shown, always centered on the magenta cross, that represents the 3σ limit on the magnitude of the robust-fit, quarter-averaged offset of the transit source from the target star. The numerical value of the quarterly-averaged offset source from the target star is given by OotOffset-rm in the DV analysis table (see Figures 2 or 10).
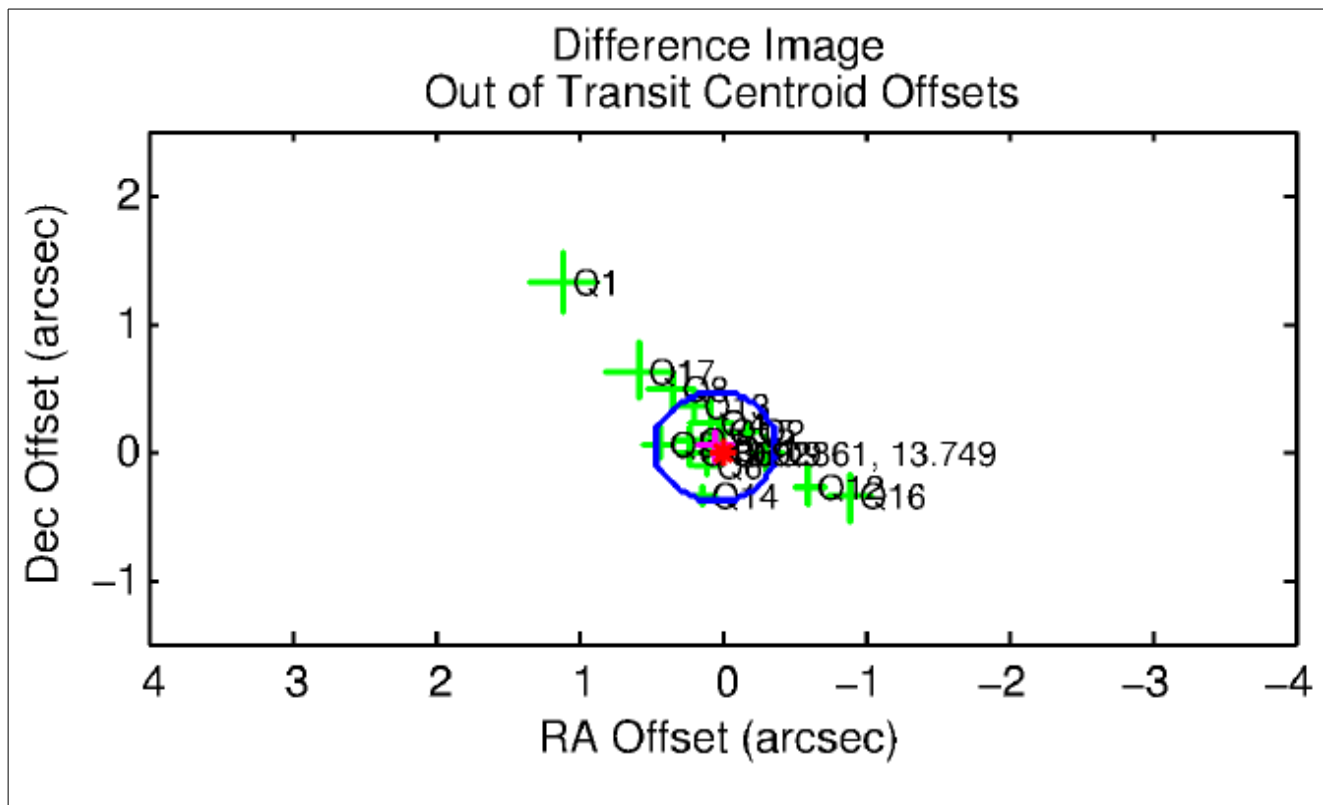

Figure 9: The PRF centroid offset plot. Individual quarterly offsets are represented by crosses, with the size of the cross corresponding to the size of the 1σ measurement error for that quarter. The location of stars is represented by asterisks, which are also labeled with their KIC number and Kepler magnitude. The red asterisk represents the target. The blue circle represents the 3σ threshold for a significant centroid offset. For this TCE, nearly all of the quarterly centroid measurements lie within the blue circle (Q1 is likely an anomalous outlier) and thus there is no significant offset detected.

The plot in Figure 9 graphically indicates whether there is a significant centroid offset between the transit source and target star location during transits, and if an alternate star is likely to be the true source of the TCE. In general, a significant (i.e., >3σ) centroid offset is seen if the red asterisk lies outside the dark blue circle. In this case it is likely that the observed transit is not due to a transit on the target star. However, significant systematic noise exists in the computed centroid offsets, just as it exists in the photometric data. Thus, it is not recommended to trust any offsets to a precision of less

than ~0.1". As well, bright stars near the target may cause an inaccurate PRF fit, rendering the centroid results invalid. This can be checked by ensuring that the KIC position is coincident with the out-of-transit PRF fit by comparing OotOffset-rm with the measured offset using the KIC location, KicOffset-rm, given in the DV results table, and discussed more in the next section (see Figure 10). Finally, these diagnostics are valid only if the TCE is due to a transit or eclipse on a star in the aperture. If the TCE results from a systematic error, such as a spacecraft pointing tweak, pixel sensitivity dropout, or other similar effect, then this method of measuring centroids is invalid.

DV Analysis Table

Figure 10 shows a table of fit parameters, derived parameters, and vetting statistics generated by the DV analysis. The left column contains best-fit parameters from a Mandel-Agol (2002) transit model that has been fitted to the whitened data, assuming the TCE is a transiting planet. The right column contains various diagnostic parameters, most of which are used to determine the location of the transit signal relative to the target star using a variety of methods, as well as the quality of the measurements.

The parameters for the left column are:

- Period: The orbital period of the planetary candidate in days. The measurement error is shown in brackets.

- Epoch: The epoch (i.e., the central time of the first transit) shown in Barycentric Kepler Julian Date (BKJD), where BKJD = JD - 2454833.0. The measurement error is shown in brackets.

- $R_p/R*$: The ratio of the planetary radius to the stellar radius. The measurement error is shown in brackets.

- $a/R*$: The ratio of the planet-star separation at time of transit to the stellar radius. The measurement error is shown in brackets.

- b: The impact parameter. A value of b = 0 represents a central transit and b = 1 represents a grazing transit where the center of the planet aligns with the limb of the star at the time of central transit. The measurement error is shown in brackets. Note that the DV fit does not allow for models with b > 1.0.

- T_{eq}: The calculated equilibrium temperature of the planet's surface in Kelvin.

- R_p: The calculated planetary radius in units of Earth radii.

- a: The calculated semi-major axis of the system in AU.

```
DV Fit Results:                      DV Diagnostic Results:

Period = 13.72236 [0.00002] d        Epoch–sig: 97.8% [0.03σ]
Epoch = 137.8414 [0.0012] BKJD       ShortPeriod–sig: N/A
Rp/R* = 0.0232 [0.0025]              LongPeriod–sig: 100.0% [388.27σ]
a/R* = 15.40 [7.31]                  ModelChiSquare2–sig: 94.8%
b = 0.68 [0.38]                      Bootstrap–pfa: 6.94e–193

Teq = 746 K                          Centroid–sig: 3.0%
Rp = 2.38 Re                         Centroid–so: 0.068 arcsec [0.50σ]
a = 0.1047 AU                        OotOffset–rm: 0.056 arcsec [0.40σ]
                                     KicOffset–rm: 0.074 arcsec [0.49σ]
                                     OotOffset–bf: N/A
                                     KicOffset–bf: N/A
                                     OotOffset–st: 4/4/4/5 [17]
                                     KicOffset–st: 4/4/4/5 [17]
                                     DiffImageQuality–fgm: 1.00 [17/17 of 17]
```

Figure 10: The DV Analysis Table. The left column contains best-fit parameters from a Mandel-Agol (2002) transit model fitted to the whitened data, assuming the TCE is a transiting planet. The right column contains various diagnostic parameters, most of which are used to determine the location of the transit signal relative to the target star using a variety of methods, as well as the quality of the centroid measurements.

The parameters for the right column are:

- Epoch-sig: A metric for how well the difference of the epochs computed separately for the odd- and even-only transits agree with the orbital period. 100% [0.0σ] indicates a perfect match, while lower percentages (higher σs) indicate more significant (sig) odd-even epoch differences. A significant (small percentage, high σ) value of Epoch-sig suggests that the TCE is an eccentric eclipsing binary with a small but non-zero value of $e \cdot \cos(\omega)$ (so that the secondary eclipse is slightly offset from phase 0.5) with the TCE period half of the binary's true period.

- ShortPeriod-sig: A comparison of the period of the current TCE to the next shortest period TCE in the system. A value of 100% [0.0σ] indicates no match at all between the two periods, with lower percentages (higher σs) indicating increasingly more significant (sig) agreements between the TCE periods. A significant (small percentage, large σ) value of ShortPeriod-sig may indicate that the system contains an eclipsing binary whose primary and secondary eclipse events have been detected as two different TCEs, thus having very similar periods but different epochs. If ShortPeriod-sig has a value of "N/A" it means that there are no additional TCEs detected in the system with shorter periods than the current TCE.

- LongPeriod-sig: A comparison of the period of the current TCE to the next longest period TCE

in the system. A value of 100% [0.0σ] indicates no match at all between the two periods, with lower percentages (higher σs) indicating increasingly more significant (sig) agreements between the TCE periods. A significant (small percentage, large σ) value of LongPeriod-sig may indicate that the system contains an eclipsing binary whose primary and secondary eclipse events have been detected as two different TCEs, thus having very similar periods but different epochs. If LongPeriod-sig has a value of "N/A" it means that there are no additional TCEs detected in the system with longer periods than the current TCE.

•ModelChiSquare2-sig: The significance of the $\chi^2_{(2)}$ veto test as described in Seader et al. (2015). If this value is close to 100% then it indicates the shape of the transit events are well described by a transit model. If this value is close to 0% then the transit events are not well described by a transit model, and the event is likely a false positive. Note that this metric is used by the TPS module to eliminate false alarms, and thus is not currently used by TCERT.

•Bootstrap-pfa: The probability of a false alarm (pfa) due to statistical fluctuations. When Bootstrap-pfa ≤ 1e-12 we believe we have a credible transit detection. In essence the test works by searching the transit-removed data for signals with the same period and duration as the TCE – if signals of comparable strength are found the validity of the original TCE is called into question. Technically speaking, the bootstrap-pfa compares the distribution of MES values that result when searching the transit-removed data at the same period and duration as the original TCE, to the MES of the original TCE. An improved version of the bootstrap metric has been computed for Q1-Q17 DR24 (see Jenkins et al. 2015).

•Centroid-sig: A measure of whether there is a statistically significant (sig) centroid shift correlated with the transit as measured by flux-weighted centroids. A Centroid-sig value near zero (i.e., <50%) indicates that a flux-weighted centroid shift was not detected.

•Centroid-so: The measured angular distance between the target star position and the location of the transiting source, determined from the in- and out-of-transit flux-weighted centroid shift. This helps determine if there is a significant offset (so) between the target and the source of the transit signal. The measurement significance is shown in brackets.

•Offset-rm: The measured angular distance between the quarterly-averaged out-of-transit source location and the quarterly averaged location of the transiting source, both determined via PRF fitting, utilizing a robust mean (rm). The measurement significance is shown in brackets.

•KicOffset-rm: The measured angular distance between the quarterly-averaged transit location determined via PRF fitting and the target star position listed in the KIC, using a robust mean (rm). The measurement significance is shown in brackets.

•OotOffset-bf: The measured angular distance between the out-of-transit (OOT) source and the transiting source location, determined via a joint multi-quarter PRF fit, with uncertainties computed via a bootstrap fit (bf). The measurement significance is shown in brackets.

•KicOffset-bf: The measured angular distance between the transit location and the target star position listed in the KIC, determined via a joint multi-quarter PRF fit, with uncertainties computed via a bootstrap fit (bf). The measurement significance is shown in brackets.

•OotOffset-st: The number of quarters for which offsets of the transit source from the out-of-transit (OOT) source location were successfully computed, as determined from PRF fitting. The data is broken down into each season (S1/S2/S3/S4) with the season total (st) number shown in brackets. This is useful to determine if the centroid measurements are all in the same season.

•KicOffset-st: The number of quarters for which offsets of the transit source location from the target star position listed in the KIC were successfully computed, using PRF fitting. The data is broken down into each season (S1/S2/S3/S4) with the season total (st) number shown in brackets. This is useful to determine if the existing centroid measurements are from multiple seasons, or if they are all from the same season, which may indicate a bias.

•DiffImageQuality-fgm: A measure of the quality of the PRF fit to the difference images. The correlation between the fitted PRF and the difference image is computed for each quarter; when the correlation is > 0.7 the fit is declared to be "high-quality". (PRF fits that are not high quality are not necessarily invalid and are used in the centroid offset plot, though examination of the pixel images in the full DV report is recommended to determine their validity.) The reported value is the fraction of quarters with successful PRF fits that were deemed "high-quality", also known as the fraction of good measurements (fgm). The three numbers in brackets represent the number of quarters with high-quality centroids, the number of quarters with a successful PRF fit, and the total number of quarters with difference images.

Page 2 – Raw Flux Plots

The photometric data shown in the DV Summary has been processed by PDC, the harmonic removal algorithm in TPS, and then either a median filter (Figures 3-5) or the whitener (Figure 6). Occasionally, these steps can distort astrophysically varying systems in such a way that a variable star appears as a transit-like event. Similarly, systematics may not be completely removed, resulting in low signal-to-noise TCEs that are difficult to identify as due to systematics using the detrended time-series data. Thus, Page 2 contains plots which show the raw photometric aperture (PA) light curve, which is simply the light curve resulting from summing up all the pixels in the target's photometric aperture. An example is shown in Figure 11.

If the raw flux plots show strong variability that appears to be not transit-like in nature, and the TCE is detected at a very similar period, it is very likely that the TCE is a FP due to not transit-like variability. As well, if the locations of the individual transits of a TCE coincide with large systematic features, then it is very likely that the TCE is due to those systematics. Variability or systematics that do not align with the period of the TCE should not be used to designate a TCE as a FP.

Page 3 – Non-Whitened vs. Whitened Light Curves

As mentioned above, detrending algorithms can sometimes affect the data such that not transit-like signals appear as transit-like signals. In the TPS transit search, the final form of detrending is the whitener, with the search for planetary signals actually occurring in the whitened domain. Page 3 thus shows the phase-folded light curve in both the flux domain (top panel) and the whitened domain (bottom panel). We show an example in Figure 12.
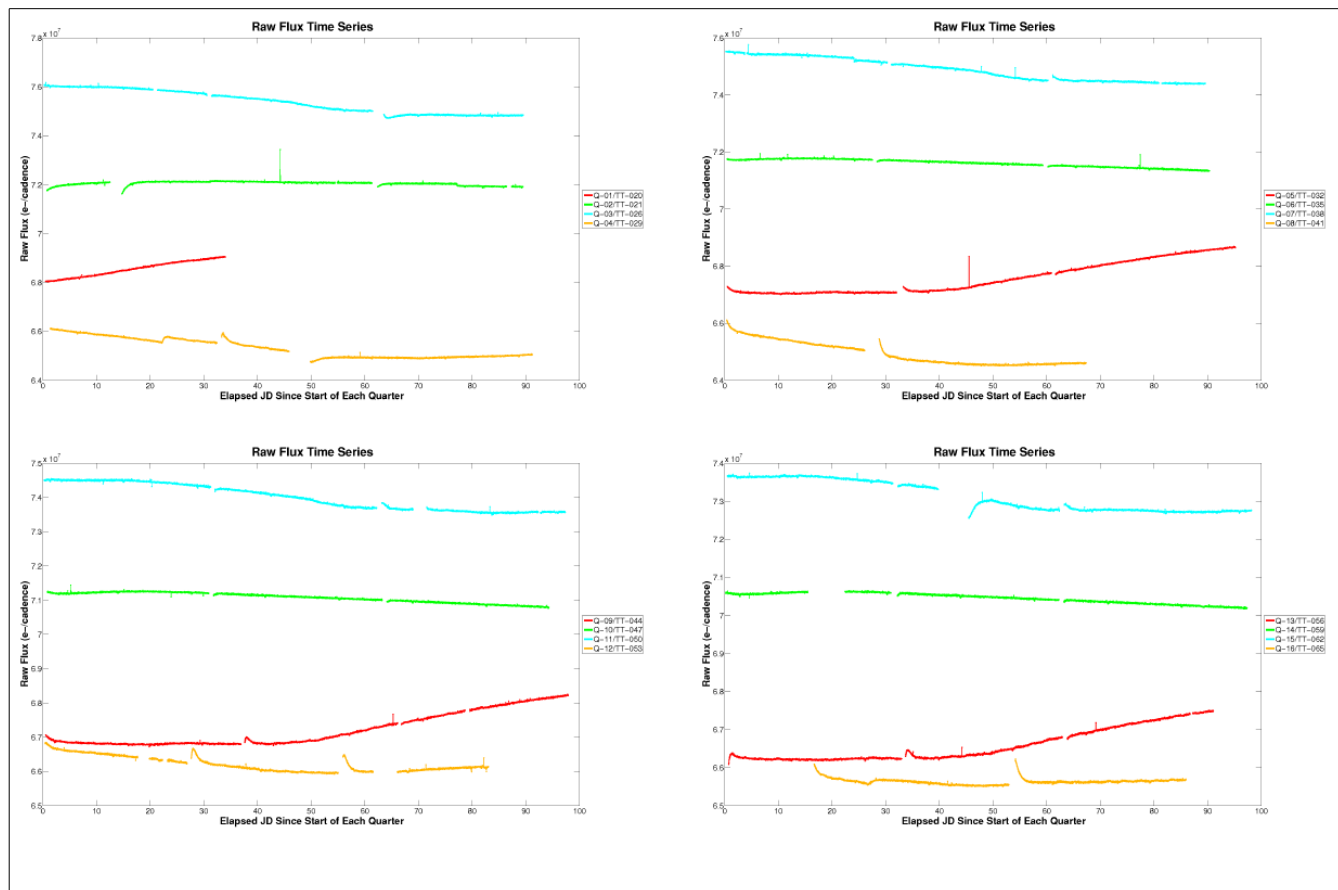
Figure 11: Page 2 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. Each panel contains four different quarters of data, as indicated by each panel's legend.

This page can be used to examine the effect of the whitener on the data. If strong not-transit-like variability exists on a very similar timescale to the TCE, but the signal appears transit-like in the whitened domain, then the TCE may be a FP due to that not-transit-like variability. If variability exists in the unwhitened data on timescales not similar to the transit signal, then the TCE should not be ruled a FP based on this variability alone. In these cases the whitener is accurately removing variable signals not related to the TCE.

Page 4 - Alternate Detrending

Kepler data often contains significant systematics even after PDC. The DV data that is presented is detrended to remove these systematics utilizing a harmonic filter, and then either a whitener (Figures 3-5) or median filter (Figure 6). The DV detrending can sometimes take astrophysically varying systems and greatly distort them in a way that flattens or enhances transit-like events. This page contains data similar to that shown in the DV one-page summary, but with an alternate detrending algorithm applied.
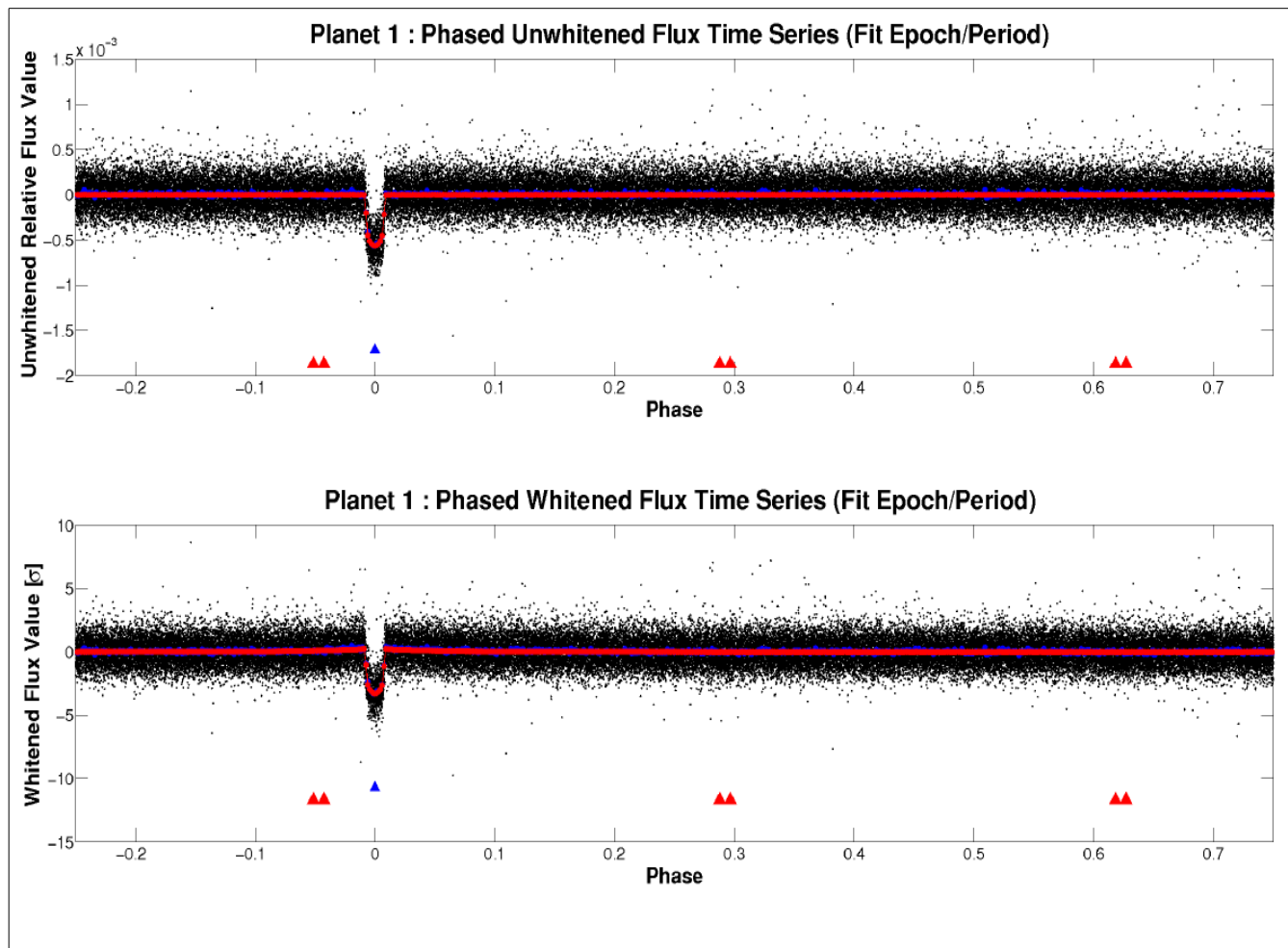
Figure 12: Page 3 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. The top panel shows the phase-folded light curve in the flux domain, while the bottom panel shows the phase-folded light curve in the whitened domain. The transit event is located at phase 0.0 with an upward-facing, blue triangle. Upward-facing triangles of other colors show the locations of the primary events of other TCEs in the system, in the phase space of the current TCE under examination.

Specifically, the alternate detrending algorithm is the nonparametric penalized least squares method presented in Garcia (2010). The phased transit light curve is fit by a simple trapezoidal model after detrending. The goal of this plot is to look for any major discrepancies between the DV data and this alternate detrending. An example is shown in Figure 13, where the three panels are analogous to the first three panels of the DV Summary (see Figures 3, 4, and 5). Note that the panels in Figure 13 show normalized flux (out of transit flux is 1.0), while Figures 3-4 show relative flux (out of transit flux is 0.0), despite the fact that both are labeled as relative.
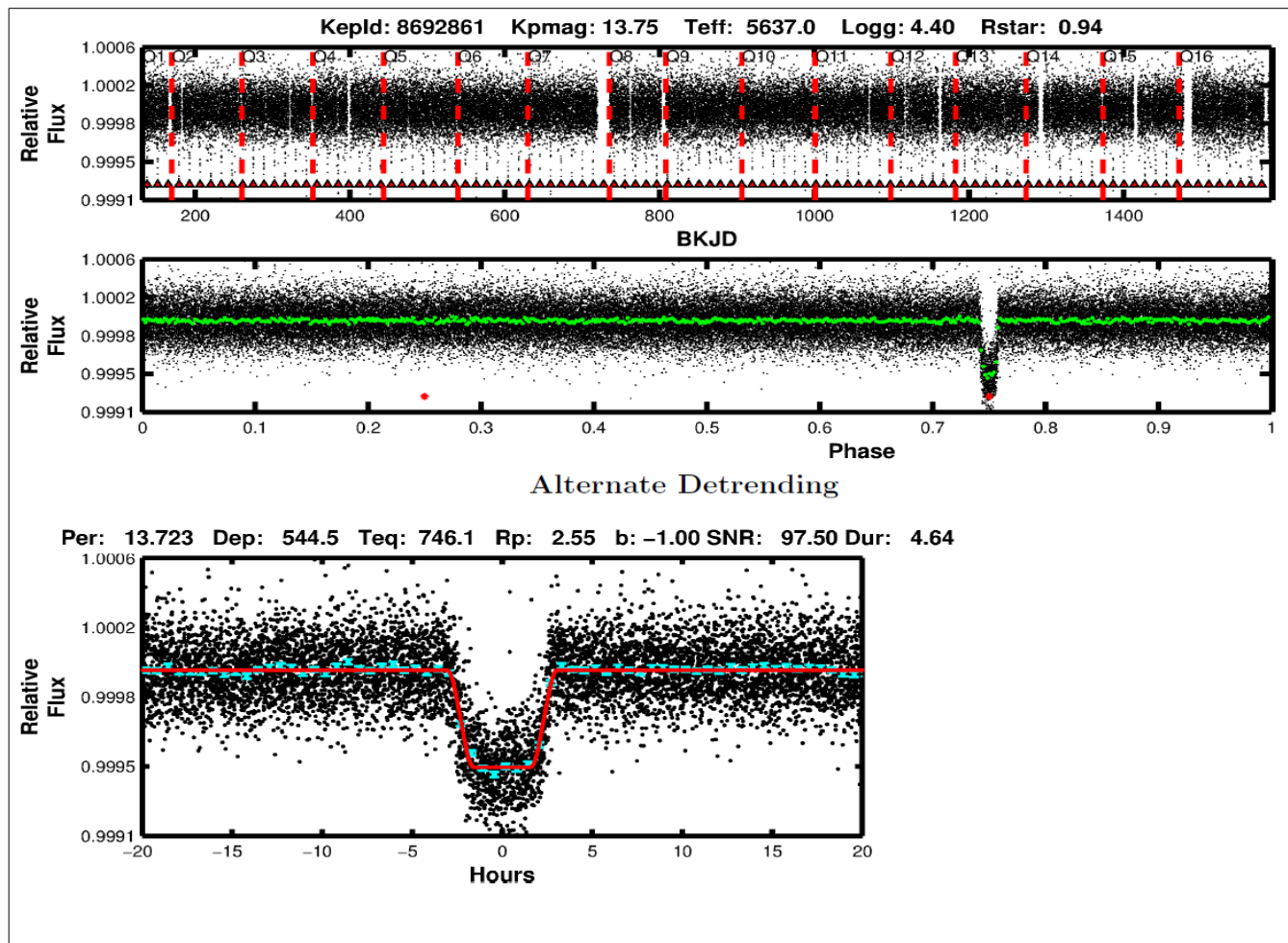
Figure 13: Page 4 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. This data is detrended using a different algorithm (Garcia 2010) than used by the TPS and DV modules. The top panel shows the full time-series light curve. Locations of expected transits are marked by upward triangles along the bottom axis. The middle panel shows the phase-folded light curve, with binned points shown via green circles. The primary is located at phase 0.75. Red circles mark the primary and secondary (if present) eclipses, assuming a circular orbit. The bottom panel shows the phased data zoomed-in on the transit event, with cyan points representing binned points, and a solid red line representing a trapezoidal model fit.

Page 5 – Odd/Even Test Plots

As described in the odd/even section of the one-page DV summary, a common source of false positives are eclipsing binaries that have very similar primary and secondary eclipse depths. While a plot of the odd- and even-numbered transits is provided in the DV summary, it is not always easy to distinguish between small depth variations. Page 5 thus contains additional plots of the odd- and even-numbered transits in order to detect any possible odd/even transit variations. We show an example in Figure 14.
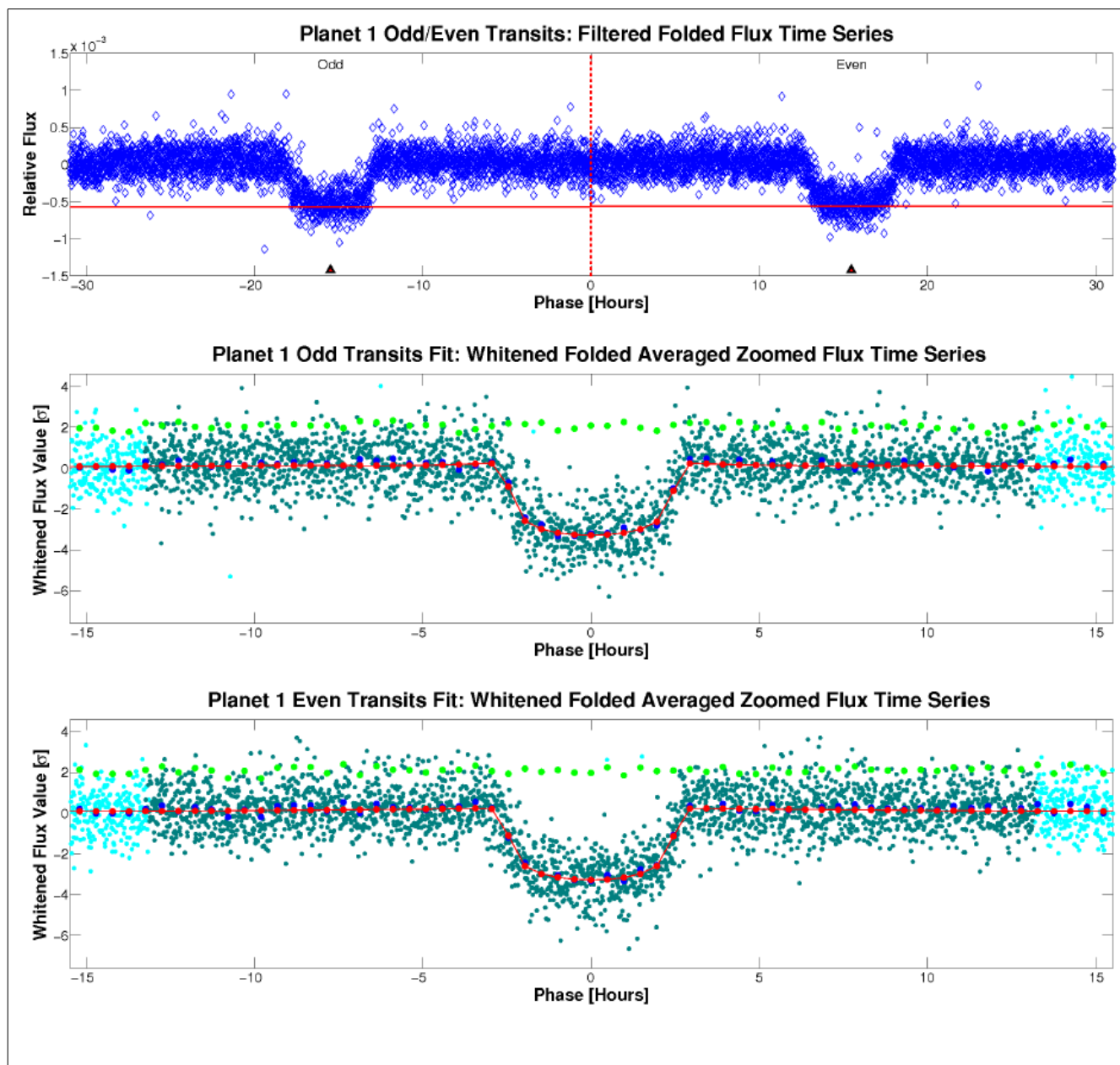
Figure 14: Page 5 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. The top panel shows the phase-folded, median-detrended, odd- and even-numbered transits side-by-side, in a rescaled version of Figure 7. The red line represents the depth of the odd- or even-numbered transits independently for each set. The bottom two panels show the odd- and even-numbered, phase-folded transits in the whitened domain. Individual points are represented by dark and light-colored bluish-green points, with the darker points being more heavily weighted in the transit fitting algorithm. Blue points represent phase-binned averages. The best-fit model transit is shown by a red line, with red points representing phase-binned averages of the model. The green points are the fit residuals, vertically offset for clarity.

TCEs that exhibit coherent, significant differences in the depth, width, or shape of the odd- and even-numbered transits are very likely to be false positives due to an eclipsing binary. One should be careful to verify that any differences are not simply due to systematics or a few outliers. For this reason, TCEs with periods of greater than 90 days (i.e., have one or less transits per quarter) should never be failed as false positives based on an odd/even difference alone.

Pages 6 and 7 – Multi-Quarter Transit Plots

With 17 quarters of data, it is possible to look for quarter-to-quarter and seasonal variations. Extreme variations are often telltale signs of contamination and false-positive identification. As well, when examining long-period candidates with only a few transits total, it is valuable to look at each individual transit, in case any are due to systematics such as SPSDs, thermal events, spacecraft pointing, etc. Zooming in on each individual transit, centered on the time of transit expected from the linear ephemeris, may also reveal transit timing variations.

Figure 15 shows the phased light curve for each quarter, utilizing DV detrended data. Each quarter is labeled via Q1 for Quarter 1, Q2 for Quarter 2, etc. To the right of each row of quarters is the combined data for that year, e.g., Year 1 (Y1) is composed of Q1, Q2, Q3, and Q4. At the bottom of each column of quarters is the combined data for that season, e.g., Season 0 (S0) is composed of Q1, Q5, Q9, and Q13. Finally, all data combined for all 17 quarters is shown in the lower-right corner as "All". Note that data from Q1 and Q17 are combined in the top-left panel, as each only spans a month of time and they correspond to the same season. Red points are unbinned data and the larger, blue points are binned data. The best-fit transit model to all the data combined is shown by the solid black line, for all quarters, years, and seasons. A very similar plot is shown for the alternately detrended data on page 7 of the Q1-Q17 DR24 vetting forms.

Since the model fit is constant, it provides a scale for the variation of each individual quarter or yearly/seasonal combination. Thus, if the data for a particular quarter is much deeper than the solid black transit model, it means that the data for that quarter presents a much deeper transit compared to all the data combined. Similarly, if the data appears flat compared to the transit model, then the data for that quarter shows no transit compared to all the data combined. If the period of the system is greater than ~90 days, then the transit displayed in each quarter is the only transit observed in that quarter. Finally, small quarter-to-quarter and season-to-season variations are normal, as the star falls on different pixels with different photometric apertures each quarter.

Pages 8 and 9 – Model-Shift Uniqueness Test and Occultation Search

In the Q1-Q17 DR24 search and analysis, 17 quarters of data were used to search for shallow transit events (less than 100 ppm) with long periods (over 300 days). This means that only a small percentage of the orbital phase contains transit information and it can be very difficult to judge the quality of a detected event when examining either a full phase-curve or a zoom-in on data close to transit. This is simply a fact of the large dynamic range of information that must be assessed to judge a transit candidate. As such, a data product was developed and used in the Q1-Q12, Q1-16, and the present Q1-Q17 DR24 TCERT activities to search for additional transit-like events in the data at the same period as the primary event.
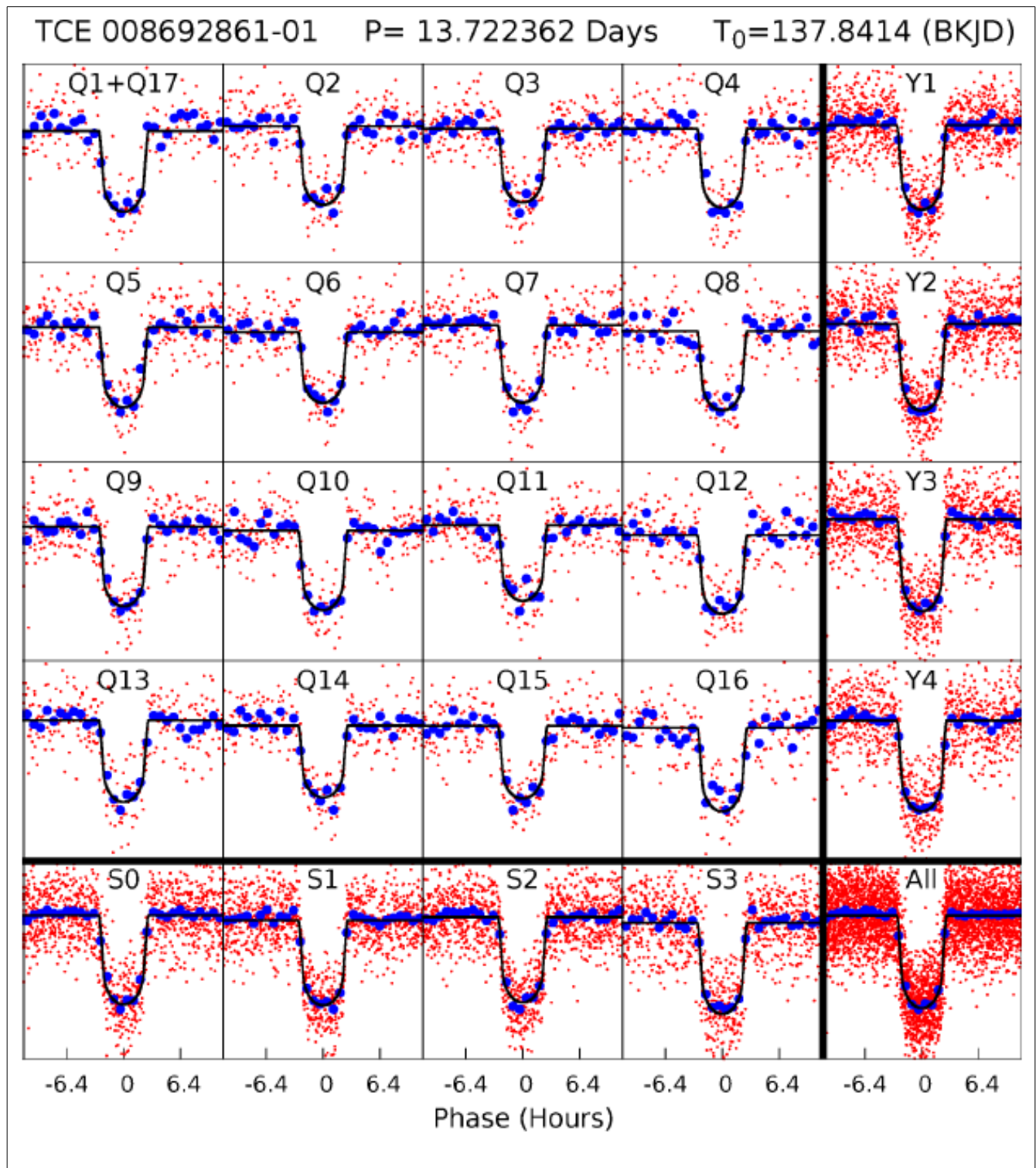
Figure 15: Page 6 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. The top line lists the TCE ID and its associated period and epoch. The top-left box shows the DV detrended flux light curve (see Figure 3) from Quarters 1 and 17 (Q1+Q17) phased to the given period and epoch.

The other boxes labeled starting with "Q" represent the other quarters, up to Quarter 16 (Q16). The boxes labeled starting with "Y" are all data combined for the given year, where Year 1 (Y1) contains data from Q1-Q4, Y2 contains data from Q5-Q8, etc. The boxes labeled starting with "S" are all data combined for the given season, where Season 0 (S0) contains all data from Q1, Q5, Q9, Q13, and Q17, S1 contains all data from Q2, Q6, Q10, and Q14, etc. The box in the bottom-right labeled "All" contains all data for the target. Note that data from Q1 and Q17 are combined in the top-left panel, as each only spans a month of time and they correspond to the same season. Red points represent the raw data and blue points are binned data such that eight binned points will always be located in-transit. The black line, for all boxes, is the DV transit model fit to all of the data. While no tick marks are shown on the y axis for any box, since the model-fit shown by the black line is constant for all boxes, it provides a sense of scale. A nearly identical plot for the alternately detrended data is shown on page 7 of the reports.

If the TCE under investigation is truly a PC, there should not be any other transit-like events in the light curve with similar depth and duration as the primary event, in either the positive or negative flux directions. If such signals are present, it calls into question the significance of the primary event (as discussed for the bootstrap metric in Figure 10). Furthermore, if the primary is unique, and there is a secondary event that is unique and distinct from any other event, it is most likely indicative of an eclipsing stellar binary. Thus, the model-shift uniqueness test and occultation search can indicate if the TCE under examination is a false alarm due to a source of non transit-like systematics, as well as if it is a false positive due to its signal originating from an eclipsing binary.

Model-Shift Uniqueness and Occultation Test Theory

To search for additional events, we take the DV photometric time series folded at the orbital period of the primary event (see Figure 4) and use the DV-generated transit model as a template to measure the amplitudes of other transit-like events at all phases. The amplitudes are measured by fitting the depth of the transit model centered on each of the data points, after rejecting outliers when taking into account the noise at the timescale of the transit duration. The deepest event aside from the primary transit event, and located at least two transit durations from the primary, is labeled as the secondary event. The next-deepest event located at least two transit durations away from the primary and secondary events is labeled as the tertiary event. Finally, the most positive flux event (i.e., shows a flux brightening) located at least three transit durations from the primary and secondary events is also labeled.

We determine the uncertainty in the amplitude measurements by calculating the standard deviation of the photometric data points outside of the primary and secondary events. Dividing the amplitudes by this standard deviation yields significance values for the primary (σ_{Pri}), secondary (σ_{Sec}), tertiary (σ_{Ter}), and positive (σ_{Pos}) events. Assuming there are P/T_{dur} independent statistical samples per TCE, where P is the period of the TCE, and T_{dur} is the transit duration, we can compute a detection threshold for each TCE such that this test yields no more than one false alarm when applied to all TCEs. We call this threshold σ_{FA}, and compute it via the following equation,

$$\sigma_{FA} = sqrt(2)*inverfc[(T_{dur} / P) \cdot (1/nTCE)]$$

where *inverfc* is the inverse complementary error function and nTCE is the number of TCEs dispositioned. For the Q1-Q17 DR24 TCERT activity, nTCE = 20,367. Finally, we also measure the amount of systematic red noise in the lightcurve on the timescale of the transit by computing the standard deviation of the measured amplitudes outside of the primary and secondary events. We report the value F_{Red}, which is this standard deviation of the measured amplitudes divided by the standard deviation of the photometric data points. $F_{Red} = 1$ if there is no red noise in the lightcurve.

It should be noted that if no DV fit was performed for the given TCE, this plot and its associated statistics cannot be generated, and does not exist. We show the full plot in Figure 16, present each individual panel in subsequent plots, and discuss how to use this plot and associated metrics for vetting. A nearly identical plot using the alternate detrending and trapezoidal model fit is shown on page 9 of the reports.

Phased Lightcurve

Similar to Figure 4, Figure 17 shows the *Kepler* DV photometric time-series phased to the orbital period of the primary event. The location of the primary event is marked by an upward, filled triangle at a phase of 0.0. The location of the best secondary eclipse candidate is marked by an upward, filled triangle at a phase other than 0.0. The location of the most significant tertiary event is marked by the location of the upward, open triangle. The location of the most significant positive flux event is marked by the location of the downward, filled triangle.

Model Filtered Data

Figure 18 shows the measured amplitude (or depth) for each phased, photometric measurement. The primary transit is always located at phase zero. A red dashed line is plotted at Flux = 0, with two blue dashed lines marking the fluxes that correspond to $\pm\sigma_{FA}$. This figure is used to evaluate the significance of possible secondary events and, more generally, to help determine the uniqueness of the transit event.

Primary, Secondary, Tertiary, and Positive Events

We also present zoomed-in plots of binned photometry for the primary, secondary, tertiary, and positive events as shown in Figure 19. Only the binned data points, shown in blue, are shown for clarity, with the transit model at the best-fit depth shown via a black line. These can be very useful for establishing the credibility of σ_{Pri} and σ_{Sec}. We also note that the transit model is based on the whitened data, so mismatches due to flux differences before ingress and after egress are common.

Model-Shift Metrics

We also report eight measurements on the right side of Pages 9 and 10 (Figure 16):

σ_{FA}: The computed detection threshold for each TCE, above which we expect only one detection of a false alarm. Specifically, $\sigma_{FA} = sqrt(2)*inverfc[(T_{dur} / P) \cdot (1/nTCE)]$, *inverfc* is the inverse complementary error function, T_{dur} is the transit duration, P is the period of the object, and nTCE is the number of TCEs dispositioned. For the Q1-Q17 DR24 TCERT activity we set nTCE = 20,367.

TCE 008692861-01, P = 13.722362 Days

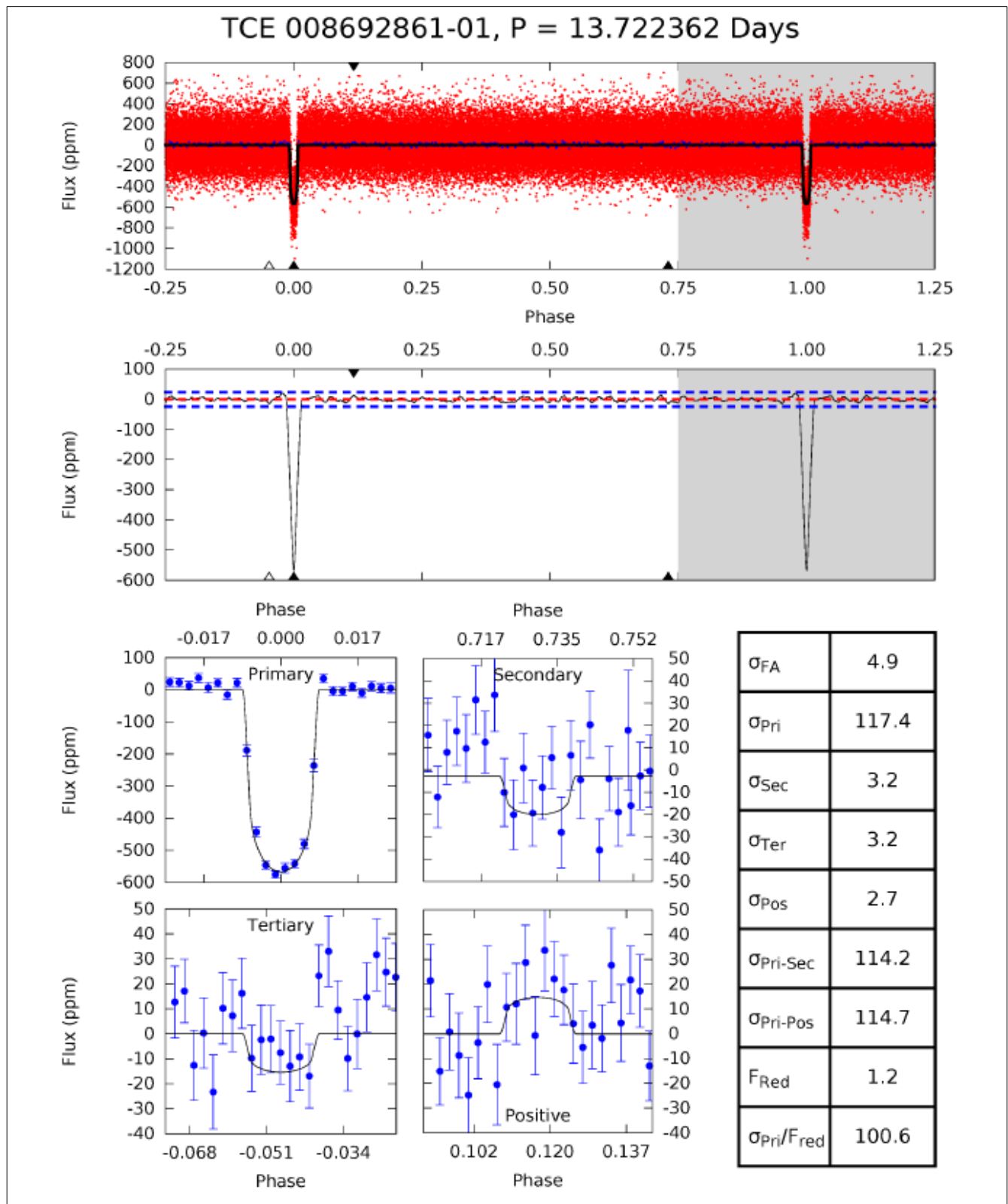|	
σ_{FA}	4.9
σ_{Pri}	117.4
σ_{Sec}	3.2
σ_{Ter}	3.2
σ_{Pos}	2.7
$\sigma_{Pri-Sec}$	114.2
$\sigma_{Pri-Pos}$	114.7
F_{Red}	1.2
σ_{Pri}/F_{red}	100.6

Figure 16: Page 8 of the Kepler Q1-Q17 DR24 TCERT Vetting Products for each TCE. The top line shows the TCE ID and associated orbital period. The top panel shows the photometric observations from the DV detrended light curve (see Figure 3) phased to the orbital period of the candidate event. The middle panel shows the data fitted at other phases using the transit model. The bottom left panels show the primary event on the top left, potential secondary event on the top right, third strongest transit-like signal (tertiary) on the bottom-left, and the strongest transit-like signal with increasing flux (positive) on the bottom-right. The table in the lower right lists the values for the false alarm detection threshold (σ_{FA}), significance of the primary (σ_{Pri}), secondary (σ_{Sec}), tertiary (σ_{Ter}), and positive (σ_{Pos}) events, the difference in significance between the primary and secondary events (σ_{Pri}-σ_{Sec}), the difference in significance between the primary and positive events (σ_{Pri}-σ_{Pos}), the ratio of the noise level on the timescale of the transit duration (red noise) divided by the Gaussian noise ($\mathbf{F_{red}}$), and the significance of the primary event when taking into account the red noise ($\sigma_{Pri}/\mathbf{F_{red}}$). A nearly identical plot for the alternately detrended data with trapezoidal fit is shown on page 9 of the reports.
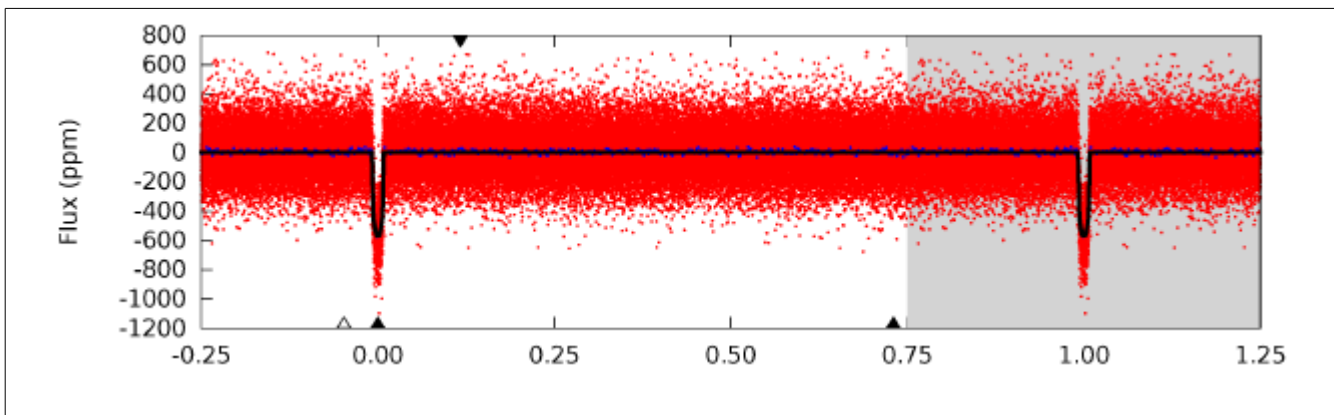

Figure 17: The red points show the phased DV photometric light curve for the candidate event. Blue points show binned data, such that eight binned points lie within the primary transit. In this example, as it is a high SNR planet candidate, the blue points mostly all lie behind the black model line. The black line shows the best-fit transit model from DV. The primary event is located at phase zero and is marked by an upward-facing triangle on the bottom axis. The other triangles mark the locations of potential secondary (filled) and tertiary events (open); the location of the strongest *anti-transit* is marked by the downward-facing triangle on the top axis. The grayed out area from phase 0.75 to 1.25 plots the same data as is plotted between phases -0.25 and 0.25.

σ_{Pri}: The significance of the primary transit. This value will appear in red if $\sigma_{Pri} < \sigma_{FA}$.

σ_{Sec}: The significance of the next deepest event outside of the primary eclipse, which is labeled as a potential secondary. The phase of this event is constrained to be more than two transit durations away from the phase of the primary. This value will appear in red if $\sigma_{Sec} > \sigma_{FA}$.
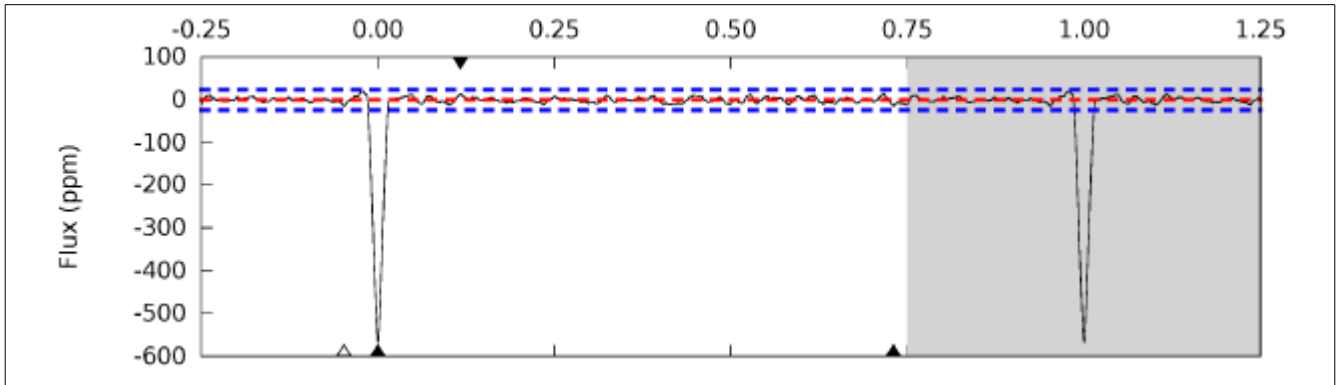
Figure 18: The fitted amplitude of the transit at each phased, photometric measurement. A red dashed line is plotted at Flux = 0. The two blue dashed lines mark the fluxes that correspond to $\pm\sigma_{FA}$, which is a Gaussian estimate of a detection threshold that will produce no more than 1 false-alarm for every 20,367 TCEs, assuming white noise. The triangles mark the location of the primary, secondary, tertiary and anti-transit events as described for Figure 17. Note that sometimes "shoulders" on either side of the primary event are produced by the DV median detrending algorithm.

σ_{Ter}: The significance of the third deepest event, aside from the primary and secondary. The phase of this event is constrained to be more than two transit durations away from the phases of the primary and secondary events. This value will appear in red if $\sigma_{Ter} > \sigma_{FA}$.

σ_{Pos}: The significance of the strongest positive or 'anti-transit' event. The phase of this event is constrained to be more than three transit durations away from the phases of the primary and secondary events, as the DV detrending can produce positive "shoulders" on the edges of transits/eclipses. This value will appear red if $\sigma_{Pos} > \sigma_{FA}$.

$\sigma_{Pri-Sec}$: The significance difference between the primary and secondary events. This value will appear red if it is less than 3.0.

$\sigma_{Pri-Pos}$: The significance difference between the primary and anti-transit events. This value will appear red if it is less than 3.0.

F_{Red}: The ratio of the level of red noise to the white, Gaussian noise. The level of red noise is computed by taking the standard deviation of the computed amplitudes at all phases outside of the primary and secondary events (see Figure 18), and dividing by the standard deviation of the photometric light curve outside of the primary and secondary events (see Figure 18). $F_{Red} = 1$ if the noise spectrum is white.

σ_{Pri} / F_{Red}: The significance of the primary transit when taking into account the amount of red noise in the lightcurve at the same timescale as the transit duration. This value will appear red if $\sigma_{Pri} / F_{Red} < \sigma_{FA}$.

In general, if $\sigma_{Pri}/F_{Red} > \sigma_{Fa}$, we believe the transit candidate is real and not due to random noise or systematic fluctuations. If $\sigma_{Pri-Sec} > \sigma_{FA}'$ and $\sigma_{Pri-Pos} > \sigma_{FA}'$, where

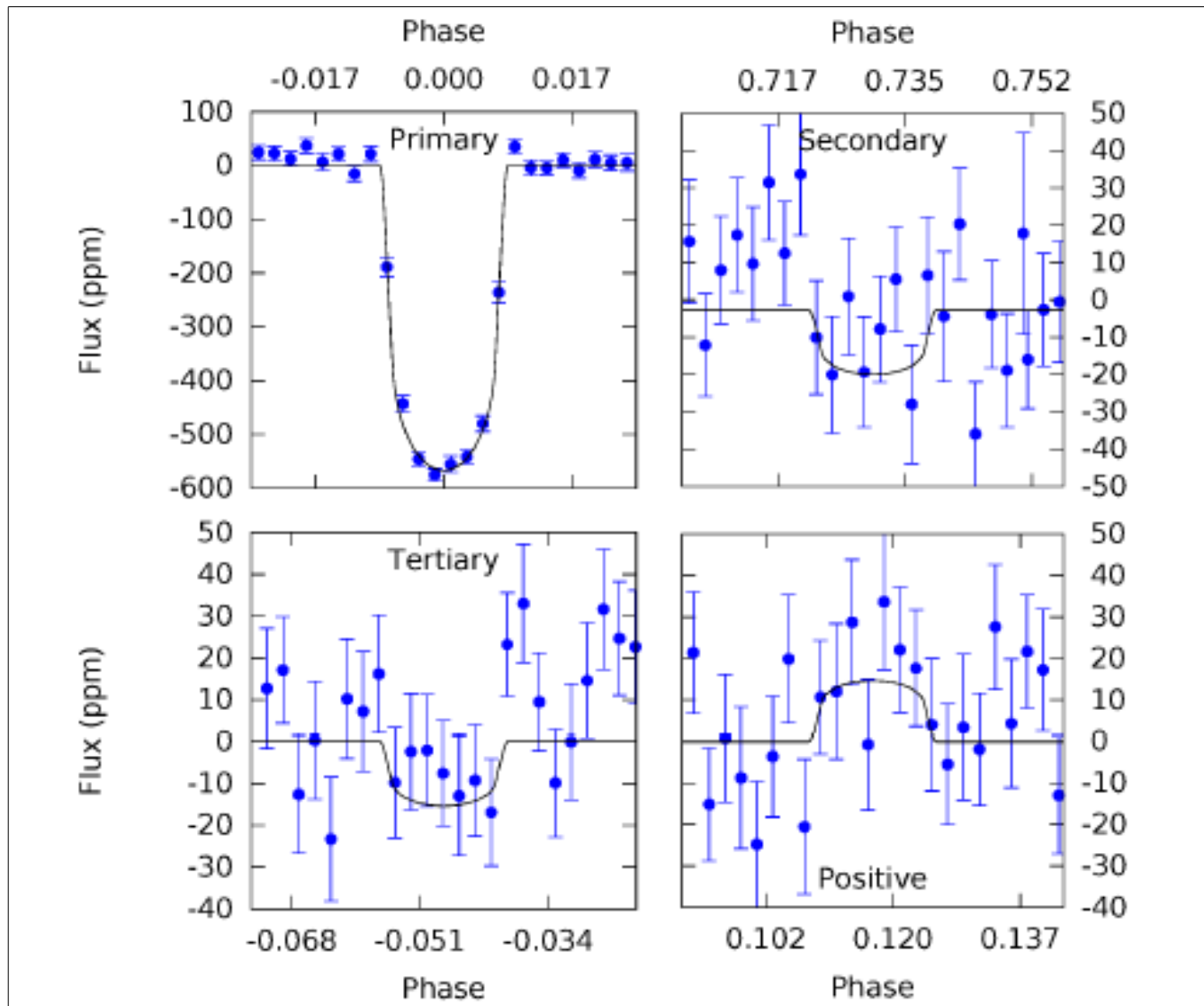$$\sigma_{FA}' = \text{sqrt}(2)*inverfc(T_{dur} / P)$$

Figure 19: Phased and binned data zoomed in on the primary (top left), secondary (top right), tertiary (bottom left), and positive (bottom right) events. The line is an overlay of the best-fit transit model to each event.

then we believe the transit candidate is unique, i.e., it tells us there is a large difference between the transit event and other putative events in the rest of the phased photometric data. If $\sigma_{Pri-Sec} < \sigma_{FA}'$ or $\sigma_{Pri-Pos} < \sigma_{FA}'$ then the validity of the transit candidate is called into question because it indicates there are other events in the light curve of comparable depth and signal-to-noise. The presence of an occultation is determined by σ_{Sec} which is a measurement of the significance of the secondary event, analogous to σ_{Pri}. Similar to the constraints on the primary, a secondary eclipse is generally considered detected if $\sigma_{Sec}/F_{Red} > \sigma_{FA}$, $\sigma_{Sec-Ter} > \sigma_{FA}'$ and $\sigma_{Sec-Pos} > \sigma_{FA}'$. Note that the value for σ_{FA}' is not shown in the reports, but is listed in Coughlin et al. (2015) and will appear in future versions of the TCERT vetting reports.

Page 10 – Stellar Parameters and Secondary Eclipse Tables

This page provides information on the stellar and secondary eclipse parameters of the TCE. An example is shown in Figure 20. The top table shows the values for the effective temperature, surface gravity, metallicity, radius, mass, and density of the host star as determined by Huber et al. (2014). The observed stellar density (p_{obs}), as determined from a model fit to the light curve (Rowe et al. 2015), is also shown. For each value the asymmetric gaussian error bars are also shown. The next row lists the error bars as a percentage of each parameter's value. The last row lists the provenance of each value, as described in Huber et al. (2014), and also summarized in the text below the table.

Stellar Parameters For KIC 008692861

	$T_{eff}(K)$	$\log(g)$	[Fe/H]	R ($R_\odot$)	$M(M_\odot)$	$p_\star$ (g·cm^{-3})	p_{obs} (g·cm^{-3})
	5637^{+116}_{-92}	$4.399^{+0.113}_{-0.134}$	$-0.300^{+0.180}_{-0.120}$	$0.943^{+0.186}_{-0.106}$	$0.813^{+0.074}_{-0.034}$	$1.365^{+0.624}_{-0.528}$	$0.598^{+0.411}_{-0.425}$
	+2%/-2%	+3%/-3%	+60%/-40%	+20%/-11%	+9%/-4%	+46%/-39%	+69%/-71%
Source	SPE48	SPE48	SPE48		DSEP		Transit Fit

KIC = Kepler Input Catalog; PHO = Photometry; SPE = Spectroscopy; AST = Asteroseismology
TRA = Transits; DESP = Dartmouth Models; MULT = Multiple Models

If $p_\star$ is significantly different than p_{obs}, it may indicate a stellar blend.

Secondary Eclipse Parameters for KIC 008692861-01 / KOI 0172.01

Detrend	Depth (ppm)	R_p ($R_\oplus$)	T_{max} (K)	T_{obs} (K)	A_{obs}
DV	-15±5	$2.40^{+0.37}_{-0.31}$	1045^{+54}_{-44}	2984^{+161}_{-179}	16^{+8}_{-6}
Alt.	-17±5	$2.40^{+0.36}_{-0.31}$	1044^{+54}_{-43}	3013^{+156}_{-164}	17^{+8}_{-6}

T_{max} = Theoretical Maximum Planetary Temperature
T_{obs} = Observed Planetary Temperature (Assuming A=0.3)
A_{obs} = Observed Albedo (Assuming T=0)

If a secondary eclipse is present, the system is likely an EB if $T_{obs} \gg T_{max}$ AND $A_{obs} \gg 1.0$

Figure 20: Tables of the stellar and secondary eclipse parameters for the given TCE, which is page 10 of the Kepler Q1-Q17 DR24 TCERT Vetting Products. These values can be used to determine if the stellar density computed from the model transit fit is consistent with the given stellar parameters of the system as determined from Huber et al. (2014). If a significant secondary was detected in either the DV or alternate detrending via the model-shift uniqueness test, then the second table can be used to determine if the secondary eclipse is possibly planetary in origin, or is due to a stellar companion.

The second table lists the secondary eclipse parameters for the TCE as determined by the model-shift test applied to both the DV and alternate detrending. The depth of the secondary eclipse is given along with an error as determined from the gaussian noise of the light curve. The planetary radius is calculated by mulitplying the radius ratio from the DV model fit by the stellar size in the above table. T_{max} is the maximum possible day-side effective temperature of the planet, assuming no re-radiation of heat from the day side to the night side. T_{obs} is the calculated day-side effective temperature that would be needed to reproduce the observed secondary eclipse depth, assuming an albedo of 0.3 and that the planet and star radiate as perfect blackbodies. Finally, A_{obs} is the albedo needed to reproduce the secondary eclipse depth, assuming no thermal emission at visible wavelengths. Asymmetric error bars are calculated using a Monte Carlo routine.

If p_{obs} is significantly different than $p_{\star}$ then it may indicate the TCE is a false positive due to a grazing EB or similar situation. It may also indicate a stellar blend or significant third light in the system. A discrepancy of these values alone should not be used to designate a TCE as a FP, but may be combined with other evidence to determine the true nature of a system.

The values in the bottom table should only be examined if a significant secondary was detected. If so, then this information can be used to determine if the secondary eclipse could possibly be due to planetary reflection and/or emission, or if it must come from a self-luminous stellar object. If T_{obs} is significantly greater than T_{max}, and A_{obs} is significantly greater than 1.0, then this indicates the secondary eclipse is indeed stellar in origin, and the TCE should be classified as a FP. If not, and the object is planetary in size ($R_p < 30\ R_{earth}$), then the TCE should not be classified as a FP.

## Page 11 – Centroid Analysis Overview

The centroid analysis pages provide more in-depth centroid information than what was presented in Figure 9 for the DV summary. Three different yet complementary reconstructions of the location of the transit signal relative to the target star are shown in Figure 21.

This page has three elements:

1. Descriptive information about the target:

    a) Kepler magnitude. This is important to identify saturated targets. When the target star is bright enough that saturation may be an issue this value is turned red.

    b) The transit SNR as measured by the DV transit model fit. This will generally indicate the quality of the difference images.

Supplemental centroid analysis for 008692861-01. Kepler magnitude: 13.75. Transit SNR 87.77
There are 17 quarters with good PRF difference image offsets
The direct PRF centroid is offset from the target star catalog position by about 0.10 arcsec

	Distance in arcsec	Distance / σ	Δ RA	Δ Dec
PRF-fit source offset from OOT	0.056 ± 0.142	0.40	0.044 ± 0.121	0.035 ± 0.115
PRF-fit source offset from KIC position	0.074 ± 0.152	0.49	0.016 ± 0.128	-0.072 ± 0.119
photometric centroid source offset	0.07 ± 0.14	0.50	0.01 ± 0.13	-0.07 ± 0.14
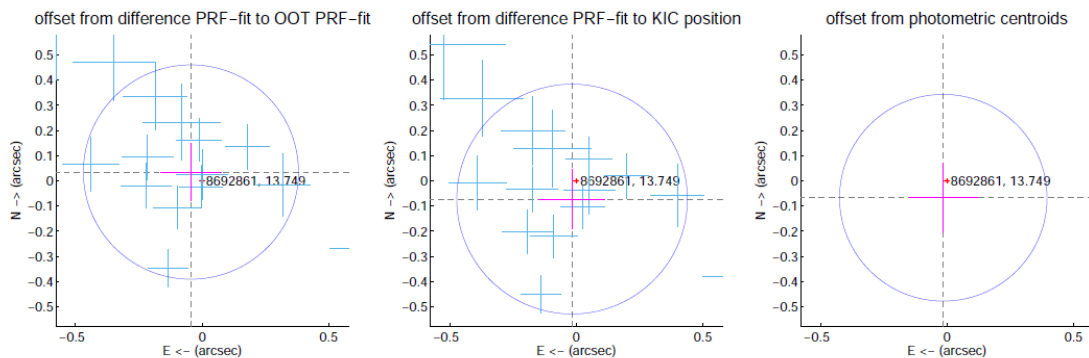
Centroid source offsets from the target star reconstructed from PRF and photometric centroids. Sky blue crosses: good quarterly centroid offsets; Vermillion crosses: bad quarterly centroid offsets; magenta cross: average over quarters. Length of the crosses: one-σ uncertainty. Blue circle: three-σ. Red *: target star. Blue *: Other stars. Text next to a star gives its KIC ID and kepmag. KIC IDs > 15,000,000 are from the UKIRT catalog.

Figure 21 The Centroid Analysis Overview page, which is page 11 of the Kepler Q1-Q17 DR24 TCERT Vetting Products. The centroid offsets computed relative to the out-of-transit PRF fit, KIC position, and photometric centroids are shown from left to right.

    c) The number of quarters with good difference images. This refers to the difference image quality metric, which tells you how well the fitted PRF is correlated with the difference image pixel data. A difference image fit is considered good if the correlation is > 0.7 –- if smaller this does not mean that the quarter's difference image is useless, just that it must be examined carefully. When the number of good quarters is three or less this line is turned red.

    d) The distance from the out-of-transit PRF-fit centroid to the target star's catalog position. When this distance is > 2 arcsec this line is turned red.

2. A table giving the reconstructed location of the transit signal relative to the target star using three different but complimentary methods:

    a) The multi-quarter average offset of the PRF-fit difference image centroid from the PRF-fit out-of-transit (OOT) image (see description associated with Figure 9).

    b) The multi-quarter average offset of the PRF-fit difference image centroid from the Kepler Input Catalog (KIC) position.

    c) The offset reconstructed from photometric centroids.

For all of these methods the distance, significance, and RA and Dec components are reported. We consider an offset distance to be statistically significant when it is greater than 3 sigma and greater than ~0.1".

3. Three panels showing the reconstructed location of the transit signal relative to the target star (located at 0,0), which correspond to the three rows of the table. The first two panels, based on PRF-fitting techniques, show the offset from the out-of-transit fit and the KIC position, respectively. In each of these the crosses represent each individual quarter, with the size of the crosses corresponding to their 1-sigma errors. The circle is the 3-sigma result for all quarters combined. The third panel shows the offset location based on photometric centroids, which provides only a multi-quarter result. As discussed previously, one should look to see if any bright stars are near the target that may influence the PRF fit by comparing the calculated offsets from the out-of-transit PRF fit and the KIC position.

For more information on these metrics and the identification of false positives using the pixel-level data, see Bryson et al. (2013).

Pages 12-16 – The Pixel-Level Images

The next five pages show the average difference and out-of-transit images for each quarter, which provide the data behind the PRF-fit centroids and the resulting multi-quarter average. These images are arranged so that they show four quarters, or a full year, per page. Each image shows three positions via markers: "x" marks the catalog location of the target star, "+" marks the PRF-fit centroid of the OOT image, and "Δ" marks the PRF-fit centroid of the difference image. The color bar is a crucial interpretation tool: when it is almost entirely positive for the difference image, this means that the difference image is reliable. When the color bar indicates negative values one should be more cautious. Large negative values are indicated with large, red "X" symbols, and may indicate that the difference images are unreliable, or that the TCE is due to systematics that do not have a stellar PRF. White asterisks indicate background stars with their Kepler ID and magnitudes. This includes stars from the UKIRT catalog, which have Kepler IDs > 15,000,000. These UKIRT Kepler IDs are internal project numbers and do not correspond with UKIRT catalog identifiers. A N/E direction indicator is provided to allow matching with the figures on page 11 of the reports (see Figure 21).

An example for Q1 and Q2 is shown in Figure 22. The pixels highlighted by a red "X" in the difference image for Q1 represent pixel value variations that are slight increases in light during the transit, but since this quarter contains less data than others (~30 days in Q1 compared to ~90 days in other quarters) and is prone to detrending errors, they do not invalidate the difference image. When the high value pixels in the difference image are isolated and appear star-like we consider this to be a clear transit signal, such as is seen for Q2. When the location of these difference image pixels correspond with the location of the star's OOT image, we say that the transit signal's location is consistent with the target star location. In this TCE, for Q1 and Q2 as shown in Figure 22, the difference images are consistent with the OOT images, and thus the transit signal is believed to originate from the target star. No centroid offset is observed.
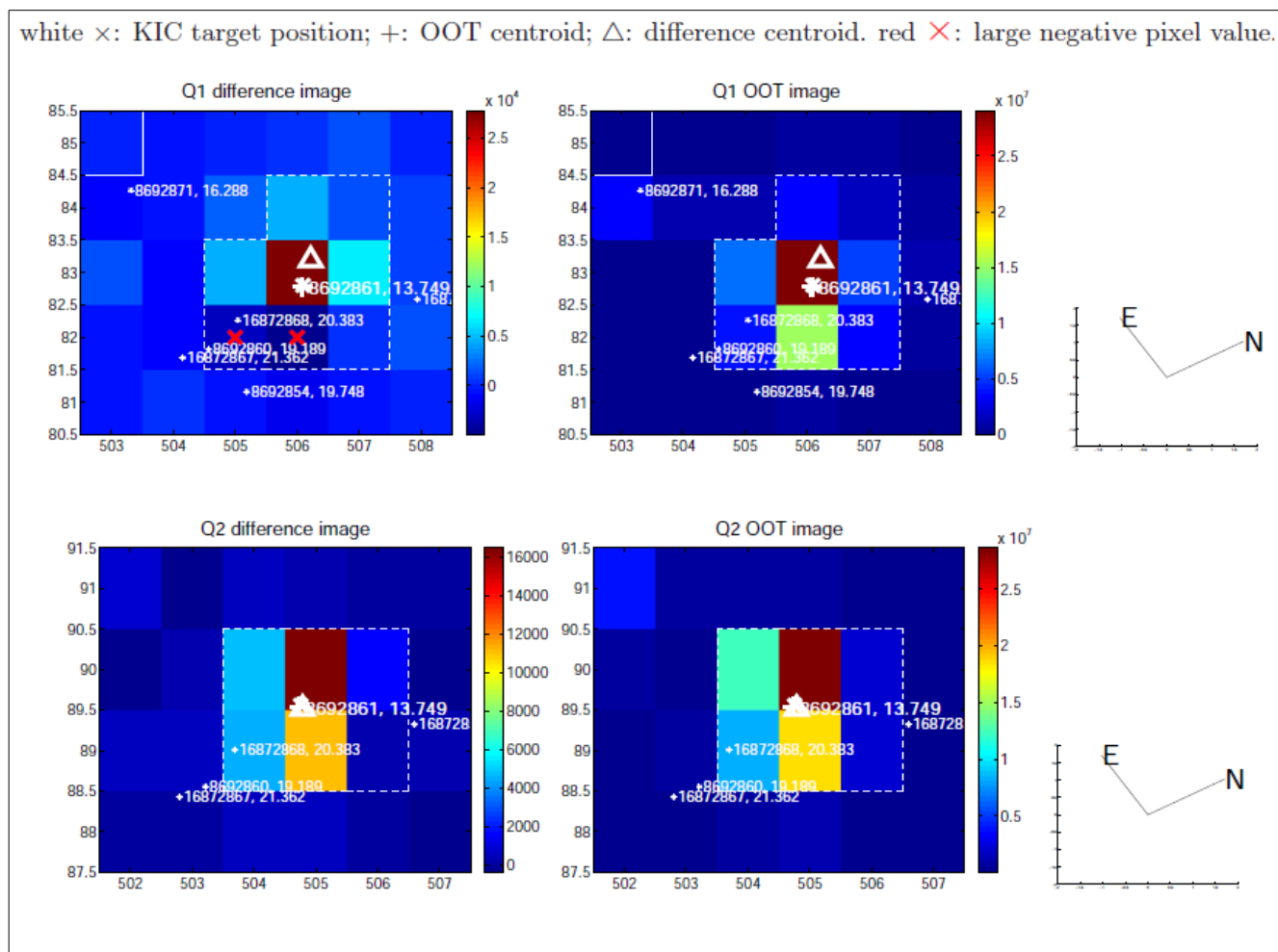
Figure 22: An example of the pixel-level images for Q1 and Q2 for Kepler-69b. The white "x" marks the catalog location of the target star, "+" marks the PRF-fit centroid of the out-of-transit image, and "Δ" marks the PRF-fit centroid of the difference image. Large red "X" symbols indicate negative pixels. In this case of this TCE, the out of transit flux image can be seen to match the difference image very well, indicating the observed transit signal likely originates from the target star.

Page 16 – Folded Flux and Flux-Weighted Centroid Time Series

This page shows the flux-weighted photometric centroids, which can be used to determine if there is a centroid shift occurring at the time of transit. Unlike the difference images, this plot is guaranteed to exist for every TCE. In Figure 23, the top panel shows the phase-folded DV photometric time-series. The middle and bottom panels show the computed RA and Dec centroid offsets, respectively, for each photometric data point. A photometric offset can be considered to be observed if there is a change in the centroid time series (second and third panel) that looks like the flux time series (top panel).
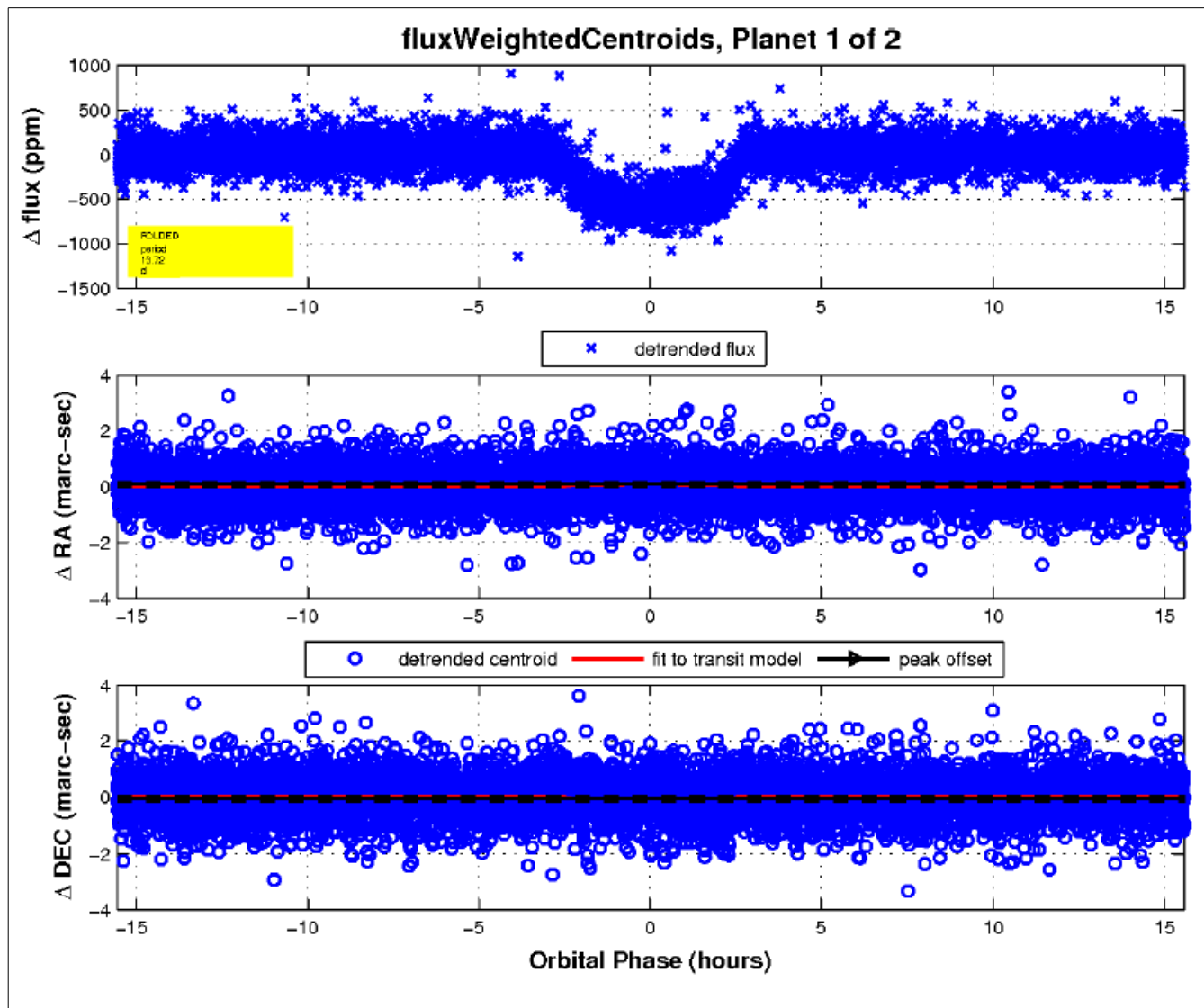
Figure 23: The flux-weighted photometric centroids. Top: The DV phase-folded photometric time series. Middle: The computed RA centroid offset for each photometric data point. Bottom: The computed DEC centroid offset for each photometric data point.

The purpose of this figure is to verify that if there is a measured photometric shift from the difference images, it looks like the transit signal, and thus is not due to instrumental systematics. However, any amount of light in the aperture not from the target star will cause a centroid shift to appear in this plot, even when the signal does originate from the target star. Thus, this plot should never alone be used to centroid fail a TCE by itself.

Page 17 and Beyond

If there are additional TCEs for the system beyond the first, they will show up as additional pages, replicating those described in Pages 1-16. In each analysis, the in-transit data from previous TCE(s) will have been subtracted from the data.

References

Bryson, S.T., Tenenbaum, P., Jenkins, J.M., et al. 2010, ApJL, 713, L97

Bryson, S.T., Jenkins, J.M., Gilliland, R.L., et al. 2013, PASP, 125, 889

Coughlin, J.L., Mullally, F.M., Thompson, S.E., et al. 2015, ApJ, Submitted

Garcia, D. 2010, Computational Statistics and Data Analysis, 54, 1167

Huber, D., Silva Aguirre, V., Matthews, J.~M., et al. 2014, ApJS, 211, 2

Jenkins, J.M., Seader, S.E., and Burke, C.J. 2015, Planet Detection Metrics: Statistical Bootstrap Test (KSCI-19086-001)

Mandel, K., & Agol, E. 2002, ApJL, 580, L171

Rowe, J. F., Coughlin, J. L., Antoci, V., et al. 2015, ApJS, 217, 16

Seader, S., Jenkins, J.M., Tenenbaum, P., et al. 2015, ApJS, 217, 18