Planet Detection Metrics:
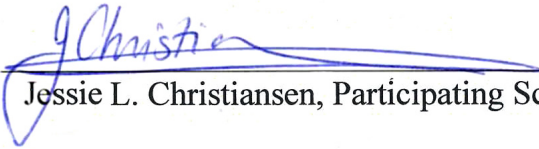
Pixel-Level Transit Injection Tests
of Pipeline Detection Efficiency
for Data Release 25

NASA Ames Research Center
Moffett Field, CA 94035

Prepared by: _____ Date __6/1/17__
Jessie L. Christiansen, Participating Scientist

Approved by: _____ Date __6/1/17__
Michael R. Haas, Science Office Director

Approved by: _____ Date __6/1/17__
Natalie M. Batalha, Project Scientist

Document Control

Ownership

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

Control Level

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

Physical Location

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

Distribution Requests

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

Michael R. Haas
Kepler Science Office Director
MS 244-30
NASA Ames Research Center
Moffett Field, CA 94035-1000

or

Michael.R.Haas@nasa.gov

DOCUMENT CHANGE LOG

CHANGE DATE	PAGES AFFECTED	CHANGES/NOTES
June 1, 2017	All	Initial release

Table of Contents

1. Introduction

This document describes the results of the fourth pixel-level transit injection experiment, which was designed to measure the detection efficiency of both the *Kepler* pipeline (Jenkins 2002, 2010; Jenkins et al. 2017) and the Robovetter (Coughlin 2017). Previous transit injection experiments are described in Christiansen et al. (2013, 2015a,b, 2016).

In order to calculate planet occurrence rates using a given *Kepler* planet catalogue, produced with a given version of the *Kepler* pipeline, we need to know the detection efficiency of that pipeline. This can be empirically determined by injecting a suite of simulated transit signals into the *Kepler* data, processing the data through the pipeline, and examining the distribution of successfully recovered transits. This document describes the results for the pixel-level transit injection experiment performed to accompany the final Q1-Q17 Data Release 25 (DR25) catalogue (Thompson et al. 2017) of the *Kepler* Objects of Interest. The catalogue was generated using the SOC pipeline version 9.3 and the DR25 Robovetter acting on the uniformly processed Q1-Q17 DR25 light curves (Thompson et al. 2016a) and assuming the Q1-Q17 DR25 *Kepler* stellar properties (Mathur et al. 2017).

In order to characterize the pipeline detection efficiency, we have performed several distinct transit injection experiments. These largely fall into two categories: pixel-level transit injection (PLTI) and flux-level transit injection (FLTI). For PLTI experiments, simulated transit signals are injected into the calibrated pixels, before the aperture photometry time series is constructed and detrended. This allows the total detection efficiency loss to be determined through the photometric and search portions of the pipeline. However, PLTI is computationally expensive, since it runs most of the pipeline modules. As a result, these PLTI experiments are limited to one injected planetary signal per target star, but include all available target stars. Hence, PLTI provides an average detection efficiency over a set of stars. Knowing that the stars are not all 'average', a series of FLTI experiments were also conducted. For FLTI, the transit signal is injected into the detrended flux time series within the Transiting Planet Search (TPS) module of the pipeline, and only the signal detection algorithm is performed (Burke & Catanzarite 2017a). For 'deep' FLTI experiments, we chose a small subset (~100) of stars and performed ~600,000 injection and recovery experiments for each star. For 'shallow' FLTI experiments, we chose a larger subset (~30,000) of stars and performed ~2,000 injection and recovery experiments for each star. These tests determined when and how individual stars can deviate from the average detection efficiency measured by PLTI. This document describes the PLTI experiment only; the FLTI products are documented separately in Burke & Catanzarite (2017a) and examples of using FLTI products to measure detection efficiency are discussed in Burke & Catanzarite (2017b).

The PLTI transit injection experiment is described in Section 2. The results are provided as an IPAC ASCII column-aligned table of input parameters and detection results as

described in Section 3. This detailed results table, which can be found on-line at the NASA Exoplanet Archive,[1] allows users to generate their own average detection efficiency for custom regions of parameter space. Section 4 describes how to use the table to calculate a detection efficiency, and Section 5 includes a worked example of interest to the *Kepler* project, showing the average detection efficiency for the ensemble of well-behaved FGK dwarfs. In Section 6 we describe the regions of parameter space where the one-dimensional detection efficiency reported here is valid and where the results are less valid, with important caveats for occurrence rate calculations. In Section 7 we review the impact of a previously reported bias in the impact parameter and planet radius fits for the DR25 TCE table (Twicken et al. 2016), and describe our updated fits.

---

[1] http://exoplanetarchive.ipac.caltech.edu/docs/Kepler_completeness_reliability.html

2. Experiment Design

The average (one-dimensional) detection efficiency describes the likelihood that the *Kepler* pipeline would successfully recover a given transit signal simply as a function of its Multiple Event Statistic (MES; the strength of the transit signal relative to the noise). To measure this property, we performed a Monte Carlo experiment where we injected the signatures of simulated transiting planets into the calibrated pixels of 190,128 target stars (typically one simulated signal per target star) across the focal plane using the Q1-Q17 DR25 light curves, processed the pixels through the data reduction and planet search pipeline as usual, and examined the distribution of the resulting detections (*c.f.* Christiansen et al. 2013, 2015a,b, 2016). The target list is divided into three groups. Most of the targets (i.e., 146,294 across 64 channels, called Group 1 for the remainder of this document) have the signal of a single simulated transiting planet injected at the target location on the CCD, thereby mimicking a planet orbiting the specified target. An additional set of targets (i.e., 33,978 across 16 channels, called Group 2) have a single simulated signal injected slightly offset from the target location, thereby mimicking a foreground or background planet or eclipsing binary along the line of sight. The presence and size of these centroid offsets are indicated in the detailed results table (see Section 3); these injections were used to test the ability of the Robovetter to discriminate between true planetary signals and background/foreground false positives (Mullally 2017). A final set of targets (i.e., 9,856 across 4 channels, called Group 3) had simulated eclipsing binary signals injected, with both primary and secondary eclipses. These targets are also discussed in the detailed results table (see Section 3), and were used to test the ability of the Robovetter to discriminate between true planetary signals and eclipsing binary false positives (Coughlin 2017).

The majority of the simulated transits that were injected had orbital periods ranging from 0.5 to 500 days. The injected planet radii were then chosen, depending on the stellar radius and orbital period, to produce transit signals that spanned the MES range (i.e., 0-20 sigma) that brackets the transition in pipeline performance from fully complete (100% recovery) to fully incomplete (0% recovery). This resulted in a large range of injected planet radii (including unphysically large radii), but with the final outcome that 50% of the injections were below 2Re and 90% below 40Re. Orbital eccentricity was set to 0, and the impact parameters were drawn from a uniform distribution between 0 and 1.

The only Group 1 targets that deviated from the above period/radius distribution were the M-dwarfs. Given the increasing interest in the habitable zones of M-dwarf targets, we took the small subset of M-dwarfs in the stellar sample (i.e., 3,809 targets with 2400 K $\leq T_{\text{eff}} < 3900$ K and log $g \geq 4$) and concentrated their injections at shorter periods (0.5 to 100 days) and smaller radii (50% of injections below 0.9Re and 90% below 1.7Re). Orbital eccentricity and impact parameter were defined as for the remainder of the Group 1 targets.

For the Group 3 targets (i.e., simulated eclipsing binaries), 9,856 targets that initially had on-target planet injections were further modified. In order to ensure most of these were detected for later study, all systems with MES < 10 had their MES re-distributed to a

MES uniformly between 10 and 100. We used half the sample to simulate EBs with shallow secondary eclipses, so systems with an existing MES > 20 primary injection had a secondary injection added at the same period, but different epoch. The MES of the secondary injection ranged uniformly from 0 to 20. To simulate a somewhat realistic EB population, which generally has circularized orbits for periods shorter than 10 days, and eccentric orbits for periods longer than 10 days, systems with P < 10 days had their secondary injected at a phase of 0.5 with an impact parameter identical to the primary, while systems with P > 10 days had their secondary eclipses injected uniformly between phases of 0.25 and 0.75 and impact parameters between 0.0 and 1.0. To simulate EBs with very similar primary and secondary eclipses, the remaining EBs had a secondary eclipse injected with a MES within ±5 of the MES of the primary. Systems with P < 10 days had their secondary eclipses injected at a phase of 0.5 with an impact parameter identical to the primary, while systems with P > 10 days had their secondary eclipses injected at a phase within one transit duration of phase 0.5, and an impact parameter within ±0.1 of the primary.

A successful pipeline detection is defined as having a Threshold Crossing Event (TCE) with an orbital ephemeris matching the injected ephemeris. The algorithm employed to match the TCE ephemeris to the injected ephemeris is described in detail in Section 4.1 of Mullally et al. (2015). In summary, the matching algorithm quantifies how many of the transit midpoints from the injected ephemeris fall near the detected TCE transits, penalized by the number of transit midpoints that do not fall near the TCE transits. The matching algorithm accepts detected ephemerides that differ by a half/double and a third/thrice the injected ephemeris orbital period for making a successful detection.

For each signal in the data, the Transiting Planet Search (TPS) algorithm performs a full calculation of the MES, which depends on the exact attributes of the cadences that contain the signal. In order to estimate the MES that would be calculated for a given injected signal (the 'expected' MES), we use an approximation that takes into account the following:

1. The dilution of the transit signal by additional light in the photometric aperture,
2. The central injected transit depth,
3. The duty cycle of the observations, discarding gapped and deweighted cadences (i.e., those with weights < 0.5), and
4. The varying noise for each transit event, using the time varying rmsCDPP estimates (Christiansen et al. 2013).

The above approximation for the expected MES does not capture all the details inherent in the Transiting Planet Search (TPS). Thus, we calibrate this expected MES approximation to the full calculation of MES (the measured MES) within TPS by performing an additional run of TPS where we force the code to evaluate the injected signal at exactly the period and epoch at which it is injected, thereby providing a measured MES. Simulations show that the expected MES approximation is 4.4% lower than the full MES as measured in TPS. The corrected expected MES values presented in

the detailed results table include the 4.4% correction factor, so the user does not need to apply this correction.

We also caution users that, in general, the 'expected' MES does not equal the measured MES. The expected MES represents the long term average MES, marginalized over all possible transit ephemeris epochs, assuming whitened Gaussian noise; it is independent of the epoch at which the signal is injected. The measured MES represents a single draw from this Gaussian distribution (now at a specific epoch), and is therefore expected to sample the width of the distribution. In addition, the measured MES is systematically lower than might be predicted due to the quantization in the period, epoch, and transit duration grid search; there is some loss in signal where the values are not well matched. The nodes of the search grid in TPS are chosen to provide a maximum reduction in the signal due to signal mismatch of 5%.

Using the prescription given in Section 4, we can use the distribution of successful detections to recover the detection efficiency as a function of the expected MES.

3. Results

3.1 Detailed Results Table

The detailed results table contains a full description of the simulated transit signal injected into each target star, a flag that indicates whether or not it was successfully recovered by the *Kepler* pipeline, and some of the recovered properties of the signal for comparison. For this table, the reported fits are from the supplemental DV run; see Section 7 for more details. The file is in IPAC ASCII format with 25 columns per injection:

1. KIC_ID: The *Kepler* Input Catalog ID of the target star,
2. Sky_Group: Sky group of the target star (identifies the target location by CCD channel for season 2 as described in Appendix D, Thompson et al. 2016b),
3. i_period: The orbital period (days) of the injected signal,
4. i_epoch: The epoch (BJD-2454833, see Section 6.2.2 of Van Cleve et al. 2016) of the injected signal,
5. N_Transit: The number of valid[2] transits contributing to the expected MES signal,
6. i_depth: The central transit depth (ppm) of the injected signal, measured after aperture corrections have been applied,
7. i_dur: The transit duration (hours) of the injected signal,
8. i_b: The impact parameter of the injected signal,
9. i_ror: The ratio of the planet radius to the stellar radius for the injected signal,
10. i_dor: The ratio of the semi-major axis of the planetary orbit to the stellar radius for the injected signal,
11. EB_injection: A flag indicating whether a simulated eclipsing binary signal was injected on the target star (1, i.e., in Group 3) or not (0, i.e., in Groups 1 or 2). An eclipsing binary signal is simulated by injecting two planetary transit models with offsets in depth and phase. Targets with simulated EB signals (9,856 in total) appear in the detailed results table twice: once with the values for the injected primary signal, and again for the values of the injected secondary signal.
12. Offset_from_source: A flag indicating whether the transit signal was injected on the target star (0, i.e., Groups 1 or 3) or offset from the target star (1, i.e., Group 2) to mimic a false positive,
13. Offset_distance: For targets injected off the target source (Group 2), the distance from the target source location to the location of the injected signal (in arcseconds),

---

[2] Cadences are deweighted (from 1 for valid to 0 for invalid) for various reasons, including proximity to a data gap or anomaly. Here, a 'valid' transit means one where the central cadence of the transit is not deweighted to <0.5.

14. Expected_MES: The expected multiple event statistic (MES) of injected signal (see Section 3 for calculation details),
15. Recovered: A flag indicating successful (1) or unsuccessful (0) recovery of the injected signal by the search pipeline,
16. TCE_ID: The Threshold Crossing Event (TCE) identifier, consisting of the *Kepler* Input Catalog (KIC) ID, followed by a dash, and then the planet number
17. Measured_MES: The maximum multiple event statistic (MES) measured by the pipeline on the recovered signal,
18. r_period: The orbital period (days) of the recovered signal,
19. r_epoch: The epoch (BJD-2454833) of the recovered signal,
20. r_depth: The central transit depth (ppm) of the recovered signal,
21. r_dur: The transit duration (hours) of the recovered signal,
22. r_b: The impact parameter of the recovered signal,
23. r_ror: The ratio of the planet radius to the stellar radius for the recovered signal,
24. r_dor: The ratio of the semi-major axis of the planetary orbit to the stellar radius for the recovered signal, and
25. Fit_Provenance: A flag indicating whether the original (0) or supplemental (1) Data Validation fits for the recovered signal are provided (see Section 7 for more details).

The detailed results table is contained in three separate files, one for each group. The file names are kplr_dr25_inj<group>_plti.txt with <group>={1, 2, 3} used to distinguish the three injection groups. These files are available for download from the NASA Exoplanet Archive.

3.2   Pipeline Products: Light Curves and Data Validation

The data products resulting from the PLTI experiment also provide a 'challenge set' of light curves that the community can use to test their own data reduction and signal detection pipelines and to compare directly to the *Kepler* pipeline. To enable this, the light curves that were used in the PLTI experiment described in this document are available for download at the NASA Exoplanet Archive. Since the transit signals were injected at the pixel level and processed through the pipeline as normal, the resulting light curve files are identical in format to those described in Section 2.1 of the *Kepler* Archive Manual (Thompson et al. 2016b). To differentiate these light curve files from the original, untampered light curves available at the MAST[3], the filenames include the string 'inj<group>' with <group>={1, 2, 3}. For example:

- original filename: kplr000757450-2009166043257_llc.fits
- modified filename: kplr000757450-2009166043257-inj<group>_llc.fits

_____

[3] https://archive.stsci.edu/kepler/

In addition to the filename change, the HDU 1 extension in the FITS file has been renamed from 'LIGHTCURVE' to 'INJECTED LIGHTCURVE'. Please exercise due diligence in quarantining light curves with simulated planet signals from light curves being searched for real planet signals.

We are also providing the Data Validation (DV) full reports and summaries for the pipeline run on these modified light curves. Their filenames are similarly modified to include the string 'inj<group>' with <group>={1, 2, 3}, so the DV files associated with the injected light curve shown above are:

> kplr000757450-2009166043257- inj1_dvr.pdf
> kplr000757450-001-2009166043257- inj1_dvs.pdf

These files are also available for download at the NASA Exoplanet Archive.   Note that they contain information on every Threshold Crossing Event (TCE) that is identified by the pipeline in a given light curve, including both the pre-existing (i.e., real) events and the simulated (or injected) events. Users should refer to the detailed results table described above to determine, for a given injection, whether it was recovered by the pipeline as a TCE and is therefore assessed within the DV reports. The DV reports and summaries for the PLTI light curves are watermarked with red text on the first page of each PDF file. The text reads:

<p style="color:red; text-align:center;">WARNING: Simulated transits were injected into
these data and may corrupt astrophysical events.</p>

An important note is that the planet fits presented in the DV reports and summaries are from the pipeline run affected by the impact parameter bias described in Section 7 of this document and in Twicken et al. (2016). As a result, the impact parameters and fitted planet radii are measured systematically higher than expected from the injected parameters. The fits that are presented in the detailed results table are from a supplemental DV run which addressed the impact parameter bias, but performed only the fitting itself (i.e., most of the Data Validation metrics were not re-computed). The values delivered in the table are the ones that should be used for occurrence rate calculations (in place of those in the DV reports and summaries).

4.   Calculating a 1-D Pipeline Detection Efficiency

Here we outline the process for determining the pipeline detection efficiency as a function of the expected MES. This allows the reader to calculate, for a given signal to noise, how likely it is that the SOC 9.3 pipeline would detect a given signal. If one is interested in particular regions of planet and stellar parameter space, one can follow these steps using the subset of targets and injections most suitable for their science case.

1. Download one or more of the detailed results table files described in Section 3.
2. If desired, choose a new MES threshold; the default is the standard MES = 7.1 threshold used by the pipeline (Jenkins 2002) and this represents the minimum threshold valid for this procedure. If a new, higher threshold is chosen, change the 'recovered' flag (column 15) to 0 for objects from the table with measured MES (column 17) below the threshold, simulating the fact that they would not have been detected under the higher threshold. Otherwise keep all rows to reproduce the standard MES = 7.1 threshold.
3. If desired, choose a limited set of stellar properties and/or planet properties over which to calculate the detection efficiency; for the worked example in Section 5, we select FGK stars, transit signals with three or more transits (column 5), and durations (column 7) shorter than 15 hours (see Section 6 for more details on the selection criteria). The *Kepler* stellar properties table available at the NASA Exoplanet Archive[4] can be used to identify which stellar targets (column 1) fall into a given stellar parameter range. To select desired planet properties, use the various columns in the table to remove injections that fall outside the desired parameter space.
4. Finally, for occurrence rate calculations, choose the subset of targets injected at the location of the target star, using the flags in columns 11 and 12 (i.e., Group 1, but not Groups 2 and 3). However, for certain false positive rate investigations, users may wish to include targets from Groups 2 and 3.
5. Select your desired expected MES (column 14) bins (for the example in Section 5 we examine expected MES from 0-100 with bins of width 0.5). For each bin, *i,* count the number of targets in the final set of rows from the table with an expected MES falling in that bin, $N_{i,exp}$, and of those, the number that were successfully recovered, $N_{i,rec}$, using either the flag in column 15 if you are using the standard MES = 7.1 threshold, or by imposing the condition that the measured MES (column 17) be greater than your chosen threshold. If you edited the 'recovered' flag in Step 2, this also produces a correct result. Then calculate the detection efficiency $N_{i,rec}/N_{i,exp}$ for each bin.
6. Plot a histogram of the resulting detection efficiency (see Figure 2 for an example).
7. Fit a function of your choice to the histogram values.
8. Use the function to correct the completeness rates in your occurrence rate calculation; for caveats on where and how this result is valid for SOC 9.3, see the discussion in Section 6.

---

[4] http://exoplanetarchive.ipac.caltech.edu/applications/TblSearch/tblSearch.html?app=ExoSearch&config=keplerstellar

5.   Average Detection Efficiency for FGK Dwarfs

For the worked example presented here, we restrict the simulated transit signals to those injected at the target location (Group 1), with three or more valid injected transits, and transit durations shorter than 15 hours (see Section 6 for more details). We restrict the stellar sample to stars with effective temperatures from 3900-7000K and log g > 4.0 (corresponding to FGK dwarf stars). This sample comprises 84,807 targets. The left panel of Figure 1 shows the distribution of transit signals injected into these targets as a function of radius and period (blue points), and indicates which of these signals were successfully detected by the pipeline (red points). The right panel shows the same points as a function of expected MES and period.
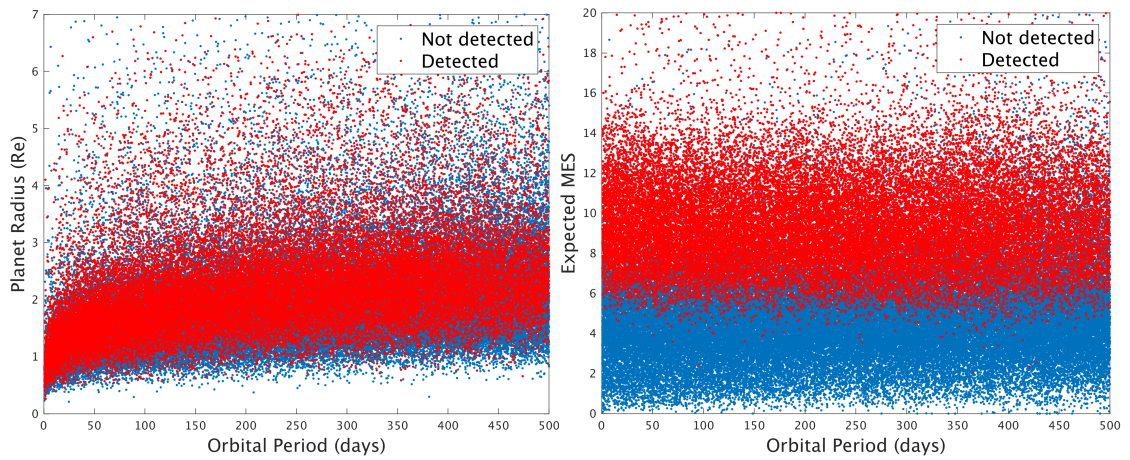

Figure 1: *Left*: The distribution of planet radius and orbital period for the simulated transits injected into the FGK dwarf pixel data. *Right*: The distribution of expected MES and orbital period. In both cases the signals that were not recovered are in blue, while those successfully recovered are in red.

We then calculate, as described in Section 4, the fraction of simulated transits successfully detected by the pipeline as a function of expected MES. Figure 2 shows a histogram of the calculated fraction as a function of expected MES. The theoretical behavior of the pipeline, assuming perfectly whitened noise, is an error function centered on the detection threshold of 7.1 sigma, with a width of one sigma (red dashed curve). The measured behaviour is well fitted by a gamma cumulative distribution function of the form:

$$p = F(x \mid a, b, c) = \frac{c}{b^a \Gamma(a)} \int_0^x t^{a-1} e^{-t/b} dt$$

where p is the probability of detection, Γ is the gamma function, x is the expected MES, and c is a scaling factor, for MES≤15. A fit of this function to the histograms gives coefficients a = 30.87, b = 0.271, c = 0.940. This means that a 50% detection efficiency is not achieved until a MES of 8.41, as compared to the idealized 7.1 sigma. As the MES increases, the detection efficiency flattens out at ~94%, an improvement over the SOC 9.2 pipeline for which the transit injection experiment (Christiansen et al. 2015b, 2016) recovered 92% for short-period injections (<100 days) and 81% for long-period injections (>100 days).
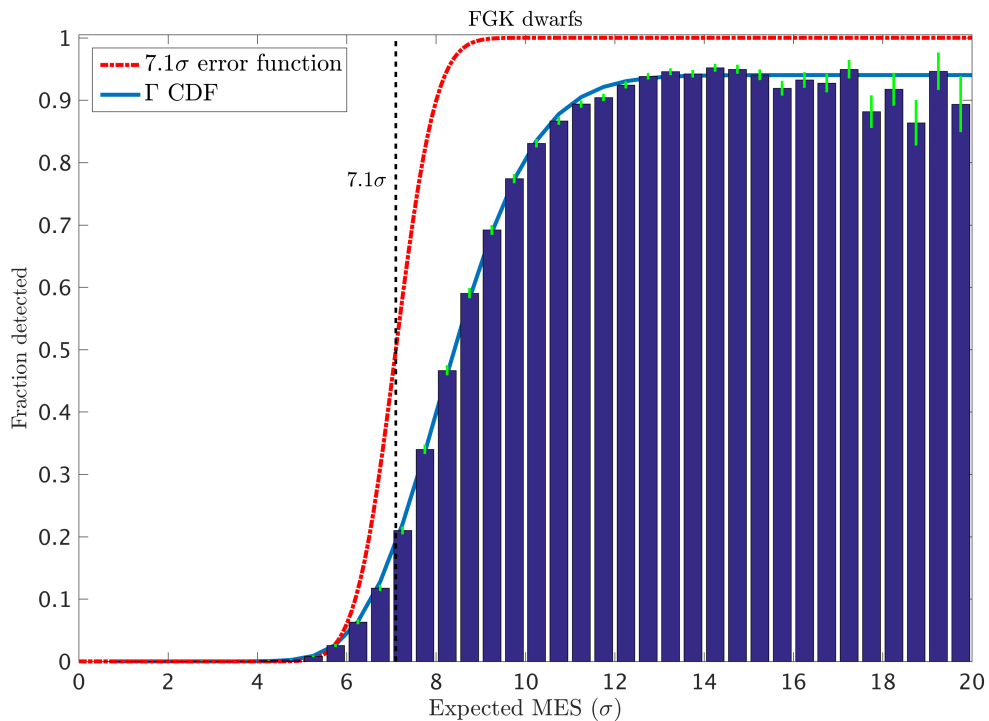

Figure 2: The fraction of simulated transits recovered as a function of the expected multiple event statistic (MES) by the *Kepler* SOC 9.3 pipeline using the Q1-Q17 DR25 pixel-level injected light curves. The black dashed line is MES=7.1. The red dashed line is the hypothetical performance of the detector on perfectly whitened noise, which is an error function centered at MES=7.1. The solid blue line is the gamma CDF fit to the histogram.

6.   Variation in Detection Efficiency Across Parameter Space

We have identified several ways in which the measured detection efficiency can vary, and we summarize them here for clarity for users who wish to create their own detection efficiency curves.

Transit durations longer than 15 hours.

The pipeline searches each light curve using model transit shapes with durations from 1.5-15 hours (Jenkins et al. 2017). The probability of recovery by the pipeline of transit signals with durations longer than 15 hours cannot be reproduced as a simple one-dimensional function of the expected MES, as can be done for the shorter durations. This is due to several factors, including distortion of the longer transit shapes by the whitening filter, and the increased likelihood that the harmonic fitter in TPS will detect and attempt to remove the longer-duration transits. For the worked example in Section 5, we restrict our analysis to injected transits with durations shorter than 15 hours, and advise that users consider this threshold when deriving their own detection efficiency curves.

Number of valid transits

The worked example in Section 5 requires that the minimum number of valid transits is equal to 3 (the standard pipeline threshold). Slightly higher detection efficiency can be recovered when the minimum number of transits is set to 4 (95.8% compared to 94.9% for 3 transits in the worked example in Section 5) or 5 (96.1%), by further isolating window function effects. Therefore, users considering completeness below periods of 300 days may wish to adopt a higher minimum number of valid transits. In order to correctly include window function effects, users should refer to Burke et al. (2015) and Burke & Catanzarite (2017c).

Fractional duty cycle drop

As the pipeline iteratively searches for and finds potential threshold crossing events in the light curves, it masks out the cadences comprising the identified signal and continues to search. The total change in the duty cycle for a given light curve (the fraction of remaining valid cadences) from the start to the end of the search is called the fractional duty cycle drop. As an increasing number of cadences are removed, the fractional duty cycle drop increases, making it more difficult for the pipeline to accurately estimate the MES of the remaining signal/s. In addition, these light curves are the most likely to have timed out in the search process, so they may not have been searched down to the 7.1 sigma threshold. As a result, the validity of a one-dimensional approximation to the detection efficiency (as a function of MES) decreases with increasing fractional duty cycle drop. We recommend a threshold of 0.05, below which the one-dimensional

approximation seems to be valid. We urge users to consider using the fractional duty cycle drop when selecting stellar samples for occurrence rate calculations. The DR25 values for the fractional duty cycle drop can be computed for all targets using dutycycle (dc) and dutycycle_post (dcp), which are two parameters included in the DR25 stellar table available at the NASA Exoplanet Archive, as (dc-dcp)/dc (see also Appendix A of Burke & Catanzarite 2017c).

7.   Important Impact Parameter/Radius Caveat for DR25 TCEs

As noted in Twicken et al. (2016), due to an error in choosing an initial model for fitting, the impact parameter for the transit-fitting module of the SOC 9.3 pipeline was effectively initialized to 0.9 for nearly all fits. For shallow transits, with poorly-constrained impact parameters, this has the result of biasing the final impact parameter fits towards values of 0.9. The upper panel in Figure 3 shows a comparison of the injected impact parameter to the recovered impact parameter, and their corresponding distributions in the marginalized histograms below and to the right, respectively. For the TCEs reported by the pipeline, and for the accompanying DV reports and summaries, the impact parameter distribution is skewed towards 0.9. This biases the measured planet radii to higher values, as more grazing transits require larger planets to produce a transit signal of the same depth. As reported in Twicken et al. (2016), the overall average radius increase is 9.8% when compared to unbiased fits.

After the full pipeline run, the fitting portion of the Data Validation module was updated to correct for this impact parameter initialization. A supplemental DV run was performed, which calculated only the updated fit parameters. These values are reported in the detailed results table where an injection is marked as 'recovered' by the pipeline. As a result, the values presented in the detailed results table described here are significantly less biased (see the lower panel in Figure 3).

We compared these supplemental DV fits with the unbiased MCMC fits provided in the DR25 KOI activity table by performing MCMC fits on a subset of 100 injected light curves; the results are shown in Figure 4. We find no remaining systematic bias in our ability to recover the injected planet radii in either the supplemental DV or MCMC fits. For users starting with the KOI table, there is no need for correction unless a supplemental DV fit is unavailable. Such targets are flagged in Column 25 of the detailed results table and the radius/impact parameter values have been reverted to the original DV fits in such cases.

For users who wish to start from the TCE table to calculate occurrence rates, caution should be exercised when selecting TCEs by planet radius (e.g., the number of TCEs in the 1-2Re bin); one solution is to scale the stated TCE planet radius down by the reported 9.8% (Twicken et al. 2016). The MCMC code is available if users wish to explore this fit bias on larger samples or for particular targets (Hoffman & Rowe 2017). The original DV fits are available for all targets in the DR25 TCE table, and the supplemental DV fits are available for all DR25 KOIs in the tables of Coughlin (2017).
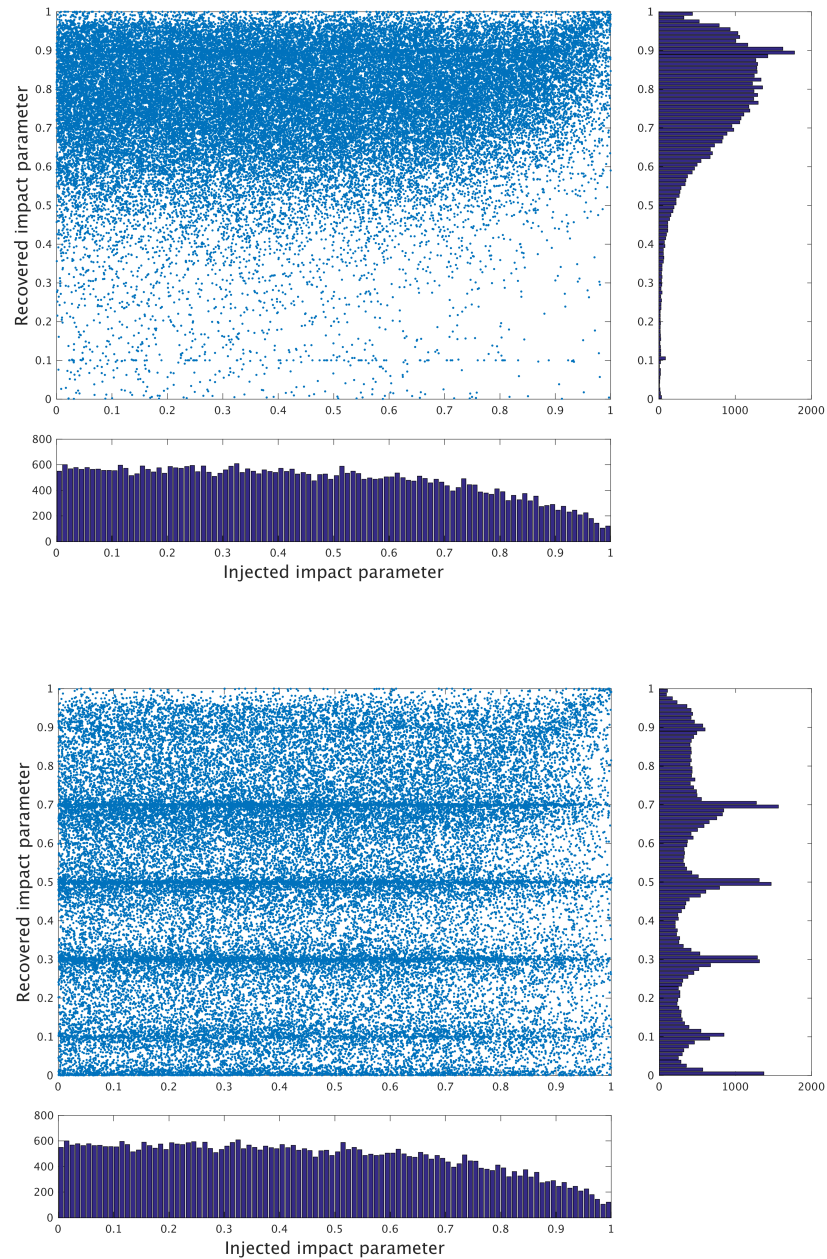
Figure 3: *Upper:* The injected and recovered impact parameter distributions from the original SOC 9.3 DV fits. *Lower:* Same as the upper panel, but for the supplemental DV fits which corrected the impact parameter bias. See also Twicken et al. (2016).
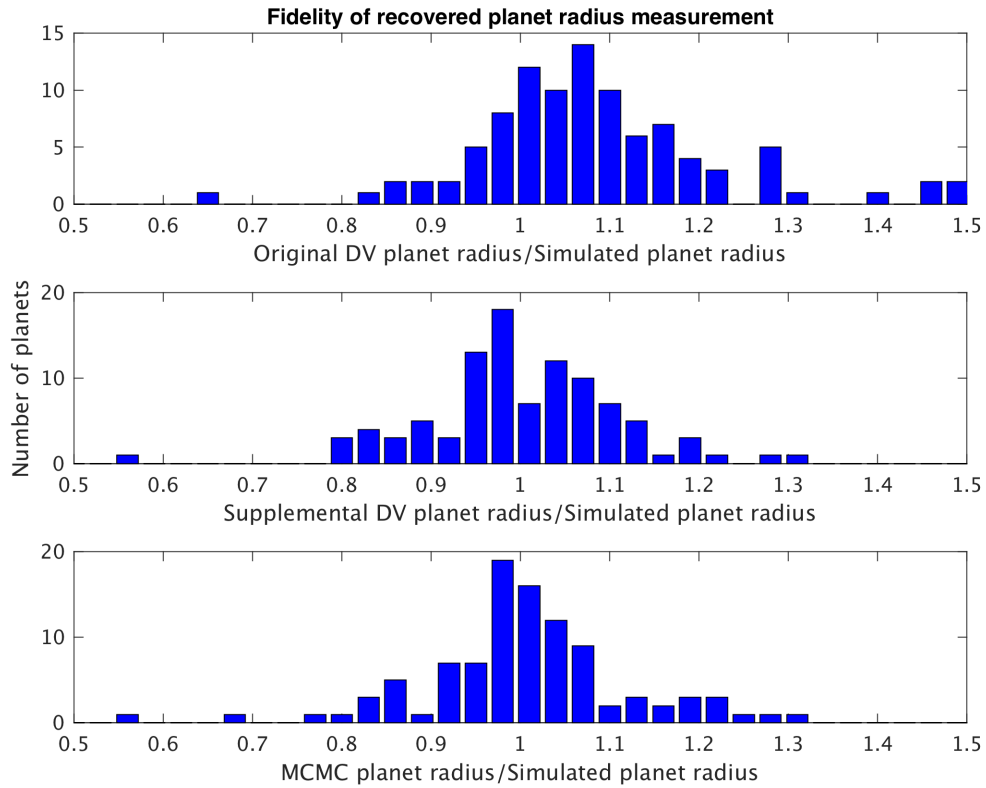
Figure 4: A demonstration of the fidelity of various fitting routines in recovering the planet radii of the simulated planets. For 100 injected transit signals fit by all three routines we show: *Upper:* The original DV fits, affected by the impact parameter bias, tended to measure larger planet radii than expected. *Middle:* The bias was corrected in the supplemental DV fits, and the measured planet radii are scattered uniformly around the injected planet radii. *Lower*: The MCMC fits, used to produce the final fits for the DR25 KOI catalogue, also show no systematic bias in the measured planet radii.

8. References

Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, *ApJ*, 809, 8

Burke, C. J. & Catanzarite, J. 2017a, Planet Detection Metrics: Per-Target Flux-Level Transit Injection Tests of TPS for Data Release 25 (KSCI-19109-001)

Burke, C. J. & Catanzarite, J. 2017b, Planet Detection Metrics: Completeness Contour Model for Data Release 25 (KSCI-19111-001)

Burke, C. J. & Catanzarite, J. 2017c, Planet Detection Metrics: Window and One-Sigma Depth Functions for Data Release 25 (KSCI-19101-002)

Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2013, *ApJS*, 207, 35

Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015a, *ApJ*, 810, 95

Christiansen, J. L., 2015b, Planet Detection Metrics: Pipeline Detection Efficiency (KSCI-19094-001)

Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2016, *ApJ*, 828, 99

Coughlin, J. L. 2017, Planet Detection Metrics: Robovetter Completeness and Effectiveness for Data Release 25 (KSCI-19114-001)

Hoffman, K. L. & Rowe, J. R. 2017, Uniform Modeling of KOIs: MCMC Notes for Data Release 25 (KSCI-19113-001)

Jenkins, J. M. 2002, *ApJ*, 575, 493

Jenkins, J. M. Chandrasekaran, H., McCauliff, S. D., et al. 2010, Proc. SPIE, 7740, 77400D

Jenkins, J. M., et al. 2017, *Kepler* Mission Data Processing Handbook (KSCI-19081-002)

Mathur, S., Huber, D., Batalha, N. M., et al. 2017, ApJS, 229, 30

Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, ApJS, 217, 31

Mullally, F. 2017, Planet Detection Metrics: Automatic Detection of Background Objects using the Centroid Robovetter (KSCI-19115-001)

Thompson, S. M., Caldwell, D. A., Jenkins, J. M. et al. 2016a, *Kepler* Data Release 25 Notes (KSCI-19065-002)

Thompson, S. M., Fraquelli, D., van Cleve, J. & Caldwell, D. A. 2016b, *Kepler* Archive Manual (KDMC-10008-006)

Twicken, J. D., Jenkins, J. M., Seader, S. E., et al. 2016, AJ, 152, 158

Van Cleve, J., Christiansen, J. L., Jenkins, J. M., et al. 2016, *Kepler* Data Characterization Handbook (KSCI-19040-005)