Planet Detection Metrics:

Pipeline Detection Efficiency

NASA Ames Research Center
Moffett Field, CA 94035
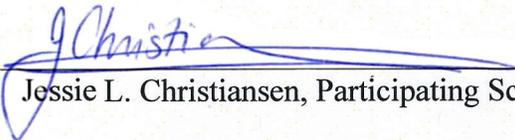
Prepared by: _____   Date 10/22/15
Jessie L. Christiansen, Participating Scientist
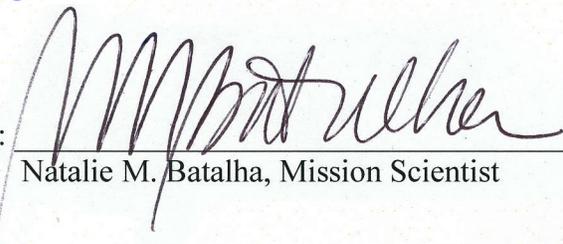
Approved by: _____   Date 10/22/15
Natalie M. Batalha, Mission Scientist
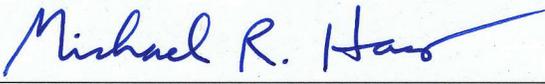
Approved by: _____   Date 10/22/15
Michael R. Haas, Science Office Director
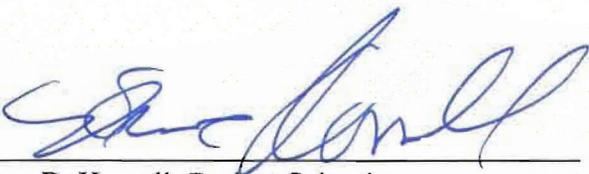
Approved by: _____   Date 10/22/15
Steve B. Howell, Project Scientist

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
October 20, 2015	7, 9	Specified central (not average) transit depth.
October 22, 2015	10	Added text in point 2 clarifying how to calculate the detection efficiency for a higher, custom MES threshold.

Table of Contents

1.   Introduction

This document describes the results of the third pixel-level transit injection experiment designed to measure the detection efficiency of the *Kepler* pipeline (Jenkins 2002, Jenkins et al. 2010).

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 measured 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 table of results of the transit injection experiment performed to accompany the Q1-Q17 Data Release (DR 24) catalogue (Coughlin et al. 2015) of Kepler Objects of Interest. The catalogue was generated using the SOC pipeline version 9.2, the uniformly processed light curves in DR24, and the Q1-Q16 *Kepler* stellar properties (Huber et al. 2014) as updated for the Q1-Q17 transit search and resulting DV reports (Huber 2014).

The 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. The full table is provided so that users can generate their own average detection probabilities for custom regions of parameter space. The full table can be found on-line at the NExScI Exoplanet Archive.[1] Section 4 describes how to use the table to calculate the detection efficiency, and Section 5 includes a worked example of interest to the *Kepler* project, showing the average detection efficiency for the ensemble of FGK dwarfs. Section 6 cautions that this average detection efficiency is not a good approximation for the behavior of the pipeline for many targets, due to known issues with the 9.2 processing, and discusses alternative options for calculating occurrence rates as reliably as possible with the products in hand.

---

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

2.   Experiment Design

The average detection efficiency describes the likelihood that the *Kepler* pipeline would successfully recover a given transit signal 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 198,154 target stars across the focal plane using the Q1-Q17 DR 24 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. 2013a; Christiansen et al. 2015). Of these injections, 159,013 resulted in three or more injected transits and were used for the subsequent analysis. Most of the targets (129,611 across 68 channels) have the simulated transit signal injected at the target location on the CCD, thereby mimicking a planet orbiting the specified target. The remaining targets (29,402 across 16 channels) have their 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 can be used to test the ability of the pipeline to discriminate between this type of false positive signal and real planetary signals (Mullally et al. 2015b). The simulated transits that were injected had orbital periods ranging from 0.5 to 500 days and planet radii ranging from 0.25 to 7.0 R_e. Orbital eccentricity was set to 0, and the impact parameter was drawn from a uniform distribution between 0 and 1. A 'successful' detection is defined as one with a measured orbital period within 3% of the injected period and a measured epoch within 0.5 days of the injected epoch; on inspection these values captured all reasonable matches, see Christiansen et al. 2015 for more detail. The pipeline uses a set of vetoes to assess whether to pass a signal on for further analysis; due to the vetoes in SOC pipeline version 9.2, 805 targets were identified at an integer alias of the injected period. For the purposes of this experiment they are not defined as 'successful' detections, but in the full table are separately identified in the 'Recovered' column with a value of 2.

For each transit signal we calculate the expected MES, which takes into account the following:

1.   The dilution of the transit signal by additional light in the photometric aperture,
2.   The central transit depth,
3.   The duty cycle of the observations, discarding gapped and deweighted cadences (*i.e.,* those with weights < 0.5), and
4.   The mismatch between the duration of the injected signal and the discrete set of 14 pulse durations searched by the pipeline. Transit signals in the data are compared with test signals of duration 1.5, 2.0, 2.5, 3.0, 3.5, 4.5, 5.0, 6.0, 7.5, 9.0, 10.5, 12.0, 12.5 and 15 hours. Therefore a transit signal with a duration of 3.75 hours, which would have its highest detection significance when compared to a test signal of duration 3.75 hours, will be measured with a lower signal strength.

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

3. 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. The file is in IPAC ASCII column-aligned format with 21 columns per target:

1. Kepler ID: Kepler 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 & Fraquelli, 2014),
3. Period: Orbital period (days) of the injected signal,
4. Epoch: Epoch (BJD-2454833, see Section 6.2.4 of Christiansen et al. 2013b) of the injected signal,
5. t_depth: Central transit depth (ppm) of the injected signal,
6. t_dur: Transit duration (hours) of the injected signal,
7. b: Impact parameter of the injected signal,
8. Rp/Rs: Ratio of the planet radius to the stellar radius for the injected signal,
9. a/Rs: Ratio of the semi-major axis of the planetary orbit to the stellar radius for the injected signal,
10. Offset from source: A flag indicating whether the transit signal was injected on the target star (0) or offset from the target star (1) to mimic a false positive,
11. Offset distance: For targets injected off the target source, the distance from the target source location to the location of the injected signal (in arcseconds),
12. Expected MES: Expected multiple event statistic (MES) of injected signal (see Section 5 for calculation details),
13. Recovered: A flag indicating successful (1) or unsuccessful (0) recovery of the injected signal by the pipeline (Note: 805 targets have a value of 2, indicating that they were recovered by the pipeline at an integer alias of the injected period),
14. Measured MES: The maximum multiple event statistic (MES) measured by the pipeline on the recovered signal,
15. r_period: Orbital period (days) of the recovered signal,
16. r_epoch: Epoch (BJD-2454833) of the recovered signal,
17. r_depth: Central transit depth (ppm) of the recovered signal,
18. r_dur: Transit duration (hours) of the recovered signal,
19. r_b: Impact parameter of the recovered signal,
20. r_Rp/Rs: Ratio of the planet radius to the stellar radius for the recovered signal, and
21. r_a/Rs: Ratio of the semi-major axis of the planetary orbit to the stellar radius for the recovered signal.

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 would have been for the pipeline to detect the signal. If one is interested in particular regions of planet and stellar parameter space, one can then calculate the signal to noise of the putative signals and compute their recovery rates.

1. Download the results table described in Section 3.
2. If desired, choose a new MES threshold (for more discussion see Section 6); 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 13) to 0 for objects from the table with measured MES (column 14) below the threshold, simulating 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 further desired (and strongly recommended, see Section 6), choose the parameter space in stellar properties and/or planet properties over which to calculate the detection efficiency; for the worked example in Section 5, we select FGK stars. The *Kepler* stellar properties table available at the NASA Exoplanet Archive[2] can be used to identify which Kepler IDs (column 1) fall into a given stellar parameter range. To select over 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 that were injected at the location of the target star using the flag in column 10. However, for certain false positive rate investigations (e.g., Mullally et al. 2015b), use those targets that were injected at a location offset from the target star.
5. Select your desired expected MES (column 12) bins (for the example in Section 5 we examine 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, $N_{i,det}$, using either the flag in column 13 if you are using the standard MES = 7.1 threshold, or by imposing the condition that the measured MES (column 14) be greater than your chosen threshold. Then calculate the detection efficiency $N_{i,det}/N_{i,exp}$ 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 strong caveats on where and how this is a valid correction for SOC 9.2, see the discussion in Section 6.

---

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

9. Feel secure in the knowledge that you are one of the few people who has read this painstakingly crafted documentation and remind me to buy you a drink the next time we are together.

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, and the stellar sample to FGK dwarfs with effective temperatures between 4000 and 7000 K and surface gravities greater than 10000 cm/s^2; this sample comprises 105,184 targets. Figure 1 shows the distribution of transit signals injected into these targets as a function of radius and period, and indicates which of these signals were successfully detected by the pipeline.
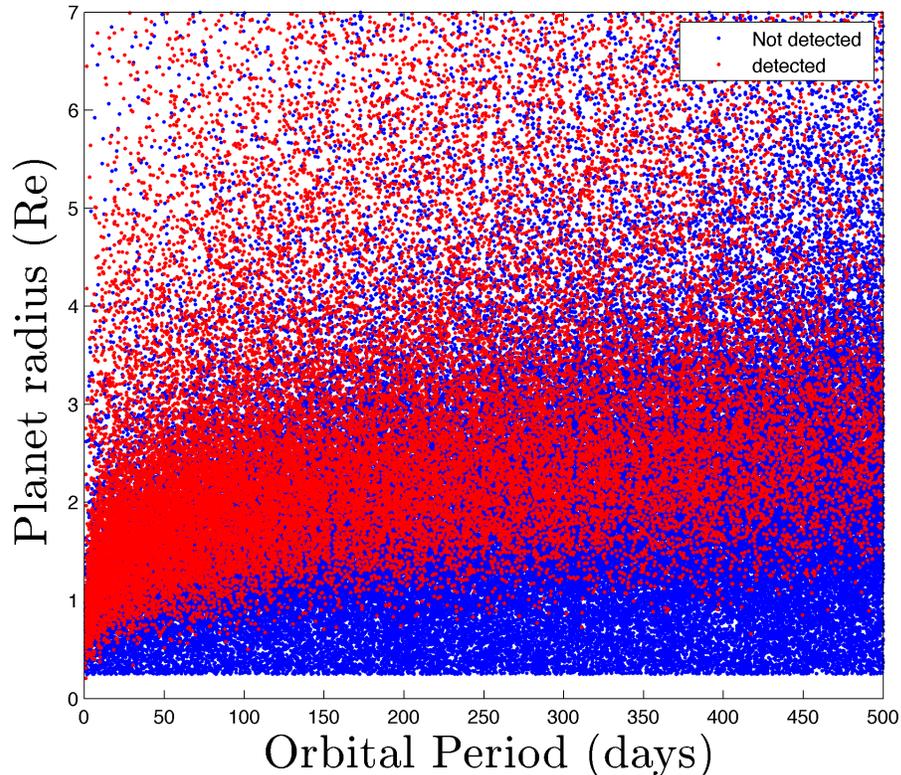
Figure 1: The distribution of planet radius and orbital period for the simulated transits injected into the FGK dwarf pixel data. The signals that were not recovered are in blue, while those successfully recovered are in red.

We then calculate, as described in Section 4, the fraction of simulated transits successfully detected by the pipeline. Figure 2 shows a histogram of the calculated fraction as a function of expected MES, for transit signals with periods below 100 days (blue), and greater than 100 days (red). 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 – this curve is the dashed line shown in magenta.

The real behavior of the pipeline is somewhat less than ideal. The histograms are well fitted by a gamma cumulative distribution function of the form:

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

where p is the probability of detection, $\Gamma(a)$ is the gamma function, $x = \text{MES} - 4.1$, and c is a scaling factor such that the maximum detection efficiency is the average of the per-bin detection probabilities recovered for $15 < \text{MES} < 50$. A fit of this function to the histograms gives coefficients $a = 7.511$, $b = 0.551$, $c = 0.915$ for periods < 100 days (shown as the blue solid line), and $a = 7.4106$, $b = 0.811$, $c = 0.810$ for periods > 100 days (shown as the red solid line). This corresponds to a successful detection rate at 7.1 sigma, the nominal pipeline threshold, of 24% for periods < 100 days, and 7.5% for periods > 100 days. The detection efficiency is virtually flat for both period ranges for expected MES values > 20, so they are not shown in the plot. For the short period transit signals, the detection efficiency flattens out at ~92%; for the longer period signals the detection efficiency is on average lower across all values of expected MES, reaching a maximum of 81%. This degradation in detectability at longer periods is due to the documented behavior of the bootstrap statistic in the 9.2 version of the Transiting Planet Search (TPS) pipeline module (Seader et al. 2015), and is discussed further in Section 6.
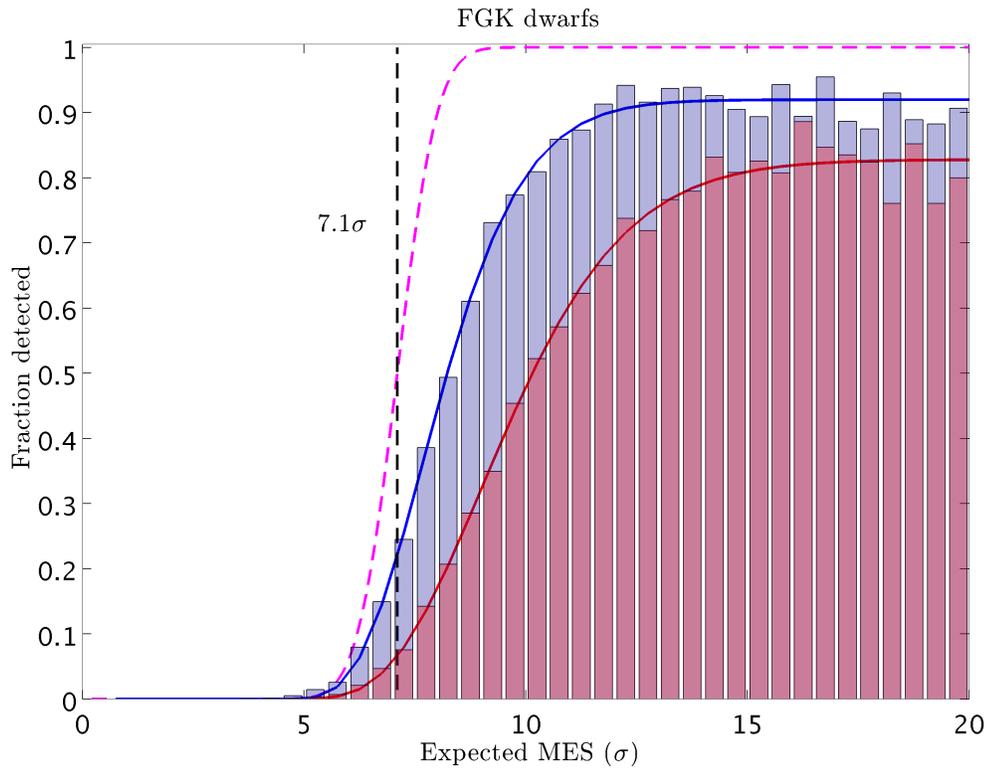
Figure 2: The fraction of simulated transits recovered as a function of the expected multiple event statistic (MES) by the *Kepler* SOC 9.2 pipeline using the Q1-Q17 DR 24 light curves. Blue: periods < 100 days; red: periods > 100 days. The black dashed line is MES=7.1. The magenta dashed line is the hypothetical performance of the detector on perfectly whitened noise, which is an error function centered on MES=7.1. The blue and red solid lines are gamma CDF fits to the blue and red histograms (short and long period bins) respectively.

6.   **Use in Occurrence Rate Calculations**

This section briefly describes changes in the search pipeline (SOC pipeline version 9.2) that limit the usage of an average one-dimensional detection efficiency (as illustrated in Section 4) to calculate planet occurrence rates. In particular, the prescription outlined in Burke et al. (2015) and Christiansen et al. (2015) for SOC 9.1 is not immediately applicable to the results of SOC 9.2.

This situation occurred because a new TPS veto, the statistical bootstrap metric, was introduced in the SOC 9.2 codebase to improve the ability to reject false alarms due to instrumental effects (Seader et al. 2015). The MES = 7.1 threshold used by the pipeline (hereafter referred to as the pipeline-based MES threshold) was chosen to achieve a false alarm rate of $6.38\text{x}10^{-13}$ on data which, when whitened, was dominated by Gaussian noise. The statistical bootstrap metric, however, calculates the threshold (hereafter referred to as the bootstrap MES threshold) required to achieve a uniform false alarm rate ($6.24\text{x}10^{-13}$) in the presence of non-Gaussian noise on the observed distribution on a target-by-target basis. While this new metric was effective in reducing the number of false alarms, the implementation contained a subtle flaw that introduced noise into the statistic, which adversely affected some Threshold Crossing Events (TCEs), especially those at long periods. Rather than achieving a search to a more uniform false alarm rate (the design goal for TPS), this flaw resulted in a period-dependent, non-uniform search with respect to the control of the false alarm rate.

For each target, the effective MES threshold is therefore the larger of the bootstrap and pipeline-based MES thresholds; for targets with a bootstrap MES threshold of < 7.1, the effective MES is 7.1 and the detection efficiency is equivalent to that described in Section 4. Figure 3 shows the effective MES threshold for the sample of GK dwarf targets analyzed in Burke et al. (2015) as a function of orbital period for potential TCEs evaluated at the 7.5 hour transit duration (blue points). The green horizontal line indicates the typical pipeline-based MES = 7.1 threshold. The bootstrap MES threshold is typically below the MES = 7.1 threshold for orbital periods < 40 days. However, at longer periods the bootstrap MES threshold increases above the pipeline-based MES = 7.1 threshold and begins to show very large scatter from target to target. The solid blue line represents the median effective MES threshold in bins of orbital period. The large scatter of the effective MES threshold and strong period dependence violates the assumptions of Burke et al. (2015) and precludes the use of an average detection efficiency as outlined therein.

The good news is that this adverse behavior is isolated to SOC 9.2 and does not impact the previous SOC 9.1 results. The design flaw in the SOC 9.2 bootstrap code has been identified and corrected (Jenkins et al. 2015). The new bootstrap metric and associated values have been archived at NExScI with the Q1-Q17 DR24 TCE catalog and are documented in Seader, Jenkins and Burke (2015). Additionally, the SOC 9.3 transit search code (TPS) has been further modified to reduce other sources of bias in the single events and the MES. Therefore, the SOC 9.3 DR25 KOI catalog should be much more amenable to occurrence rate calculations using the prescription in Burke et al. (2015) and Christiansen et al. (2015).

Given the behavior of the SOC 9.2 search pipeline characterized here (i.e., Figures 2 and 3), we recommend the following options for practitioners of planet occurrence rate calculations using the results of the *Kepler* pipeline.

1. Continue using the SOC 9.1 completeness products from the Q1-Q16 pipeline run (Burke et al. 2015, Christiansen et al. 2015) and planet candidate sample (Mullally et al. 2015a).
2. Wait until the next (i.e., final) SOC 9.3 DR25 completeness products and planet candidate sample become available. We still encourage the use of the SOC 9.2 completeness products and the planet candidate catalog as a means to develop and test code for calculating occurrence rates, as the SOC 9.3 products will be extremely similar, but we caution not to publish such occurrence rate results due to the aforementioned issues with SOC 9.2, except in the narrowly defined cases described below.
3. Employ the results of this transit injection release to tailor the average detection efficiency for the target sample and period space under investigation. This approach requires the adoption of a generous uncertainty on the derived detection efficiency, as the Detailed Results Table (Section 3) is based on one injection per target and is unlikely to fully capture the period and target variability present in SOC 9.2. This option is particularly prone to failure when exercised over small regions of parameter space since there will be correspondingly fewer transit injections with which to calibrate the average detection efficiency.
4. Due to its period dependence, the bootstrap veto has little impact on the detection efficiency for periods < 40 days. Thus, an average detection efficiency can be defined and used for planet occurrence rates in this regime. We still recommend investigating the variations in detection efficiency as outlined in Option 3.
5. Adopt a higher effective MES threshold ($\geq$15-20) and recalculate the average detection efficiency under this assumption, as described in Section 4. The data are available in the Detailed Results Table (Section 3) to remove the injected transits that have a MES below the new higher threshold. Under this option it is critical to also remove the planet candidates from analysis that violate the new higher threshold. An important note: to decide whether a planet candidate would have passed the adopted MES threshold, one must use the MES of the planet candidate detection from the TCE table from TPS (given as column 14 in the table presented in this document), rather than the SNR of the planet model fit from the KOI table**.**
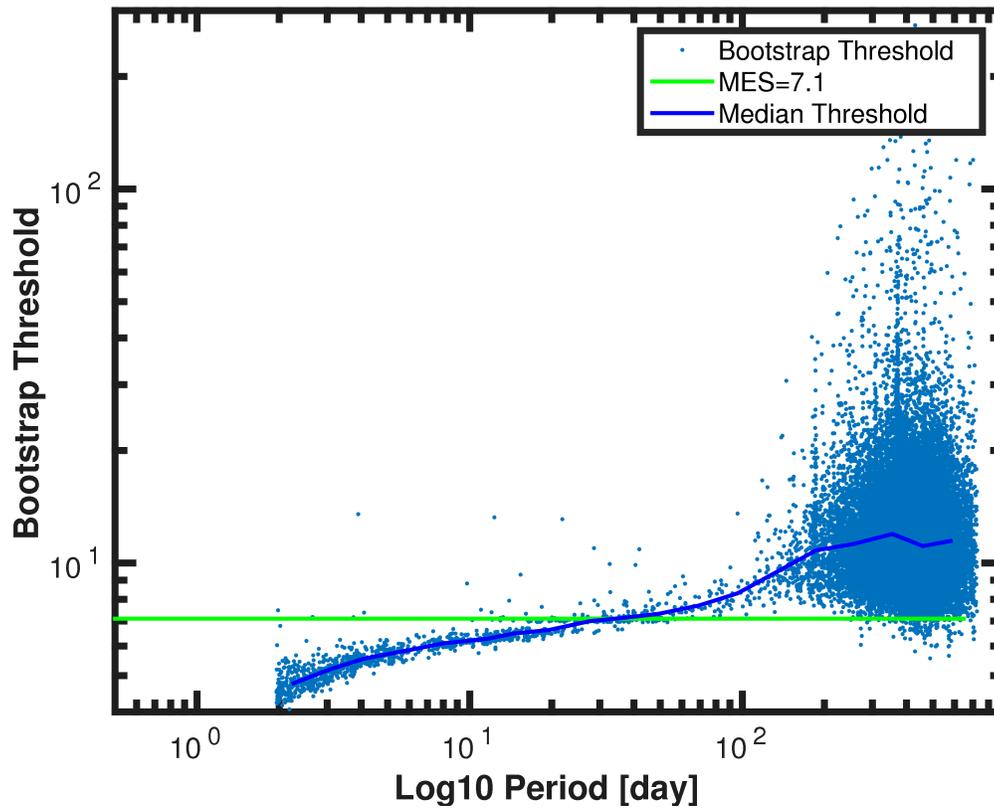
Figure 3: The effective MES threshold (blue points) as a function of orbital period for the potential TCEs in SOC pipeline version 9.2, for the sample of GK dwarf targets analyzed in Burke et al. (2015). The green horizontal line indicates the effective MES = 7.1 threshold. For occurrence rate calculations at periods greater than 40 days, we recommend adopting a higher effective MES threshold (15-20); see text for details.

7.   References

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

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

Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2013b, *Kepler* Data Characterization Handbook (KSCI-19040-004)

Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, *ApJ,* in press

Coughlin, J. L., Mullally, F. R., Thompson, S. E., et al. 2015, in preparation

Huber, D. 2014, *Kepler* Stellar Properties Catalog Update for Q1-Q17 Transit Search (KSCI-19083-001)

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

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., Twicken, J. D., Batalha, N. M., et al. 2015, *AJ*, 150, 56

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

Mullally, F., Shah, Y., Christiansen, J., et al. 2015b, *ApJS*, submitted

Seader, S., Jenkins, J. M., Tenenbaum, P., et al. 2105, *ApJS*, 217, 18

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

Thompson, S. M., & Fraquelli, D. 2014, *Kepler* Archive Manual (KDMC-10008-005)