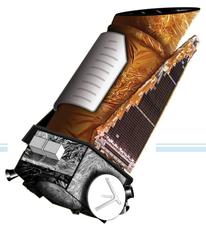
False Alarm Vetoes for Transiting Planet Searches



S. SEADER[1], P. TENENBAUM[2], J.M. JENKINS[3], C.J. BURKE[4]

[1]SETI Institute, shawn.seader@nasa.gov, [2]SETI Institute, peter.tenenbaum@nasa.gov, [3]SETI Institute, jon.jenkins@nasa.gov, [4]SETI Institute, christopher.j.burke@nasa.gov

Abstract

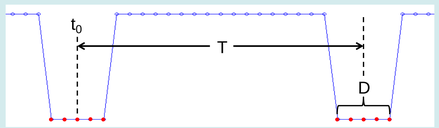
The *Kepler Mission* continuously observes a host of target stars in a 115 square-degree field of view to discover Earth-like planets transiting Sun-like stars through analysis of photometric data. The *Kepler* Science Operations Center at NASA Ames Research Center processes the data with the Science Processing Pipeline, which is composed of several modules including the Transiting Planet Search (TPS). Requiring a 7.1σ detection in 16 quarters of data in the Transiting Planet Search (TPS) yields nearly 113,000 detections, or roughly 57% of the total number of targets sent to TPS, most of which are false alarms. After a second cut is made on a robust detection statistic, over 33,000 targets still remain. These false alarms waste resources as they propagate through the remainder of the software pipeline and so a method to discriminate against them is crucial in maintaining the desired sensitivity to true events. To this end, we currently employ a set of three additional chi-squared vetoes that each examine different contributions to the detection statistic. Through the use of these vetoes, the number of detections drops below 10,000, for a false alarm rate of 4%, while simultaneously achieving a high detection rate of 98% on a set of high quality planet candidates. The new set of vetoes, described here, will be the subject of future investigation and use in TPS. The new vetoes will improve the recovery rate of injected signals and the tuning of these vetoes will play a major role in increasing our search sensitivity to compensate, in some part, for Kepler's inability to continue its nominal extended mission. Funding for the *Kepler Mission* has been provided by the NASA Science Mission Directorate.

Detection Statistics

The Single Event Statistic (SES) is a simple matched filter in the whitened domain[a], which is the optimal detector for a deterministic signal in colored Gaussian noise,

$$SES_j = \frac{\tilde{x} \cdot \tilde{s}_j}{\sqrt{\tilde{s}_j \cdot \tilde{s}_j}} \ .$$

Here the '~' denotes a whitened vector, x is the data, and s_j is the trial transit pulse centered at time $n = j$. A grid is then constructed on the three dimensional parameter space of physically realistic pulse duration(D), period(T), and epoch(t_0). The grid is constructed on this three dimensional parameter space in a way that strikes a balance between reducing signal loss due to signal-template mismatch and algorithm execution time. Each {D,T, t_0} triplet determines a set of P cadences, ϑ, corresponding to the centers of the transits in the full pulse train that are combined to form a Multiple Event Statistic (MES).



The MES are then computed as:

$$MES = \frac{\sum_{j\in\vartheta} \tilde{x} \cdot \tilde{s}_j}{\sqrt{\sum_{j\in\vartheta} \tilde{s}_j \cdot \tilde{s}_j}} \ .$$

The current threshold used for rejecting events based on the MES is 7.1σ. For events with a MES that exceeds this threshold, a Robust Detection Statistic (RDS) is computed. To compute the RDS, a set of fit weights, W, are obtained by robustly fitting a whitened transit pulse train to the whitened data, after windowing to remove all but the in-transit cadences. The RDS is then computed as:

$$RDS = \frac{\tilde{s}^T W \tilde{x}}{\sqrt{\tilde{s}^T W \tilde{s}}}$$

The current threshold used for rejecting events based on the RDS is 6.4σ.

Pitfalls

In order for the three vetoes, discussed subsequently, and their associated component quantities to have the correct statistical properties, there are three subtleties that must be taken care of.

1. The whitening coefficients used to whiten both the data and signal vectors must be explicitly independent of the presence of any transits, depending therefore only on the noise in the data
   - The whitening coefficients are estimated in the wavelet domain through the use of a moving circular Median Absolute Deviation (MAD)
   - Although this method of estimating the whitening coefficients is robust against outliers, the effect of any signal in the data on the noise estimates, however small it may be, still perturbs the statistical properties of the vetoes
   - Prior to doing the wavelet decomposition of the data, all in-transit cadences are gapped and filled using an adaptive Auto Regressive (AR) method to explicitly remove any signal that might be present in the data
2. Due to the process of whiting the data in the wavelet domain, transits that might be present in the data get smeared out and overlap one another, thereby losing their independence
   - Consider that the data vector can in general be written:

$$\tilde{x}(n) = \tilde{w}(n) + A\tilde{s}(n)$$

   where $w(n)$ is the noise vector, A is the signal amplitude, and $s(n)$ is the signal vector
   - Instead of using the full data vector in the whitening, the contribution from each transit pulse should be separated out by gapping and filling in-transit cadences for all transits except the one whose center is at n=j, effectively breaking the data up into components:

$$\tilde{x}_j(n) = \tilde{w}(n) + A\tilde{s}_j(n)$$

3. Due to the process of whitening the trial transit pulse train in the wavelet domain, the pulses get smeared out and overlap one another, thereby losing their independence
   - Again, the full pulse train must be broken up so that each transit is treated individually
   - Each transit pulse is whitened separately and windowed to remove the effect of any out-of-transit cadences

After handling these pitfalls, a new version of the MES, which is not in general equal to the MES, can be computed as:

$$Z = \frac{\sum_{j\in\vartheta} \tilde{x}_j \cdot \tilde{s}_j}{\sqrt{\sum_{j\in\vartheta} \tilde{s}_j \cdot \tilde{s}_j}} \ .$$

χ^2 Vetoes

There are three χ^2 vetoes used by TPS, two of which are described in detail in *Seader[b] et al*.

$\chi^2_{(1)}$:

- Measures the difference between the squared amplitude of the detector output and the squared SNR

$$\chi^2_{(1)} = \sum_{j\in\vartheta} \tilde{x}_j \cdot \tilde{x}_j - \left(\frac{\sum_{j\in\vartheta} \tilde{x}_j \cdot \tilde{s}_j}{\sqrt{\sum_{j\in\vartheta} \tilde{s}_j \cdot \tilde{s}_j}} \right)^2$$

- Expectation value is equal to one less than the number of in-transit cadences (ignoring signal/template mismatch)

$$\langle \chi^2_{(1)} \rangle = P * D \ - 1$$

- This quantity is χ^2 distributed with P*D-1 degrees of freedom (ignoring signal/template mismatch)

$\chi^2_{(2)}$:

- Break Z up into P, additive, pieces, where P is the number of transits

$$Z = \sum_{j\in\vartheta} z_j \qquad z_j = \frac{\tilde{x}_j \cdot \tilde{s}_j}{\sqrt{\sum_{k\in\vartheta} \tilde{s}_k \cdot \tilde{s}_k}}$$

- For each piece, compute the fractional expected value for comparison

$$q_j = \frac{\tilde{s}_j \cdot \tilde{s}_j}{\sum_{k\in\vartheta} \tilde{s}_k \cdot \tilde{s}_k} \qquad \sum_{j\in\vartheta} q_j = 1$$

- Form the differences between observed and expected values

$$\Delta z_j = z_j - q_j Z$$

- Form the statistic from these differences

$$\chi^2_{(2)} = \sum_{j\in\vartheta} \frac{(\Delta z_j)^2}{q_j}$$

- The expected value is one less than the number of transits (ignoring signal/template mismatch)

$$\langle \chi^2_{(2)} \rangle = P \ - 1$$

- This quantity is χ^2 distributed with P-1 degrees of freedom (ignoring signal/template mismatch)

$\chi^2_{(3)}$:

- Motivated by the argument of the joint probability of the set of SES from which the MES is formed from
- Compute the set of SES and their expected values

$$z_j = \frac{\tilde{x}_j \cdot \tilde{s}_j}{\sqrt{\tilde{s}_j \cdot \tilde{s}_j}} \qquad \langle z_j \rangle = A\sqrt{\tilde{s}_j \cdot \tilde{s}_j}$$

- Form a classical χ^2 statistic by comparing the observed set of SES to their corresponding expected values

$$\chi^2_{(3)} = \sum_{j\in\vartheta} \frac{(z_j - \langle z_j \rangle)^2}{\langle z_j \rangle}$$

- The number of degrees of freedom used for this version is also P-1
- This degrees of freedom overshoots the true expectation value, indicating that this version is likely not χ^2 distributed
- More work is needed to understand the statistical properties of this version and whether it should be modified or dropped from use
- The threshold used for this veto was set too high in the past and led to an unacceptably high number of false dismissals – the threshold is much lower now

χ^2 Thresholds:

- Thresholds are formulated in the same general way for each version
- Form the reduced χ^2 by dividing the χ^2 by its corresponding number of degrees of freedom

$$\chi^2_r = \frac{\chi^2}{DOF}$$

- Now divide Z by the square root of the reduced χ^2

$$\eta_* = \frac{Z}{\sqrt{\chi^2_r}}$$

- The current thresholds employed by TPS are:

$$\eta_{*(1)} = 7.5 \qquad \eta_{*(2)} = 7.0 \qquad \eta_{*(3)} = 5.0$$

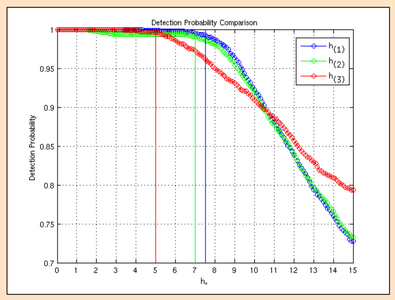
Results

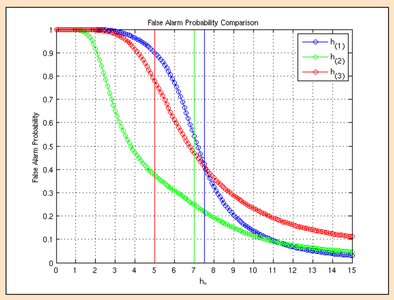
I. *Preliminaries:*

- The three vetoes have been tuned empirically for future use in analyzing Kepler data with the Science Processing Pipeline
  - TPS was run over 192,091 targets with four years of data
  - Astrophysically motivated templates[b] were used to minimize signal/template mismatch as much as possible (note that these templates will be employed by TPS for future runs)
  - The Pearson correlation coefficient controlling the spacing in period was increased to 98% to minimize signal/template mismatch
  - Thresholds on the RDS and χ^2 vetoes were all disabled
  - Continued looping over successively smaller threshold crossing events was disabled
- To assess detection probabilities, a set of 2,716 high quality KOI targets are used (H1)
- To assess false alarm probabilities, a set of 187,677 targets, that have so far not been flagged as potentially having planetary candidates, are used (H0)
- Requiring that MES ≥ 7.1 eliminates 4/2,617 (0.15%) of the H1 targets and passes 111,914/187,677 (59.63%) of the H0 targets
- Further requiring that RDS ≥ 6.4 eliminates an additional 12/2,716 H1 targets but passes only 43,313/187,677 of the H0 targets
- After thresholding on MES and RDS, the estimates of the overall detection probability and false alarm probability are 99.35% and 23.08% respectively
- In reality these are not what is expected due to the changes to the operating mode of TPS mentioned above, but this is still useful to get a general feel
- χ^2 thresholds are tuned empirically on the set of targets admitted by the MES and RDS cuts

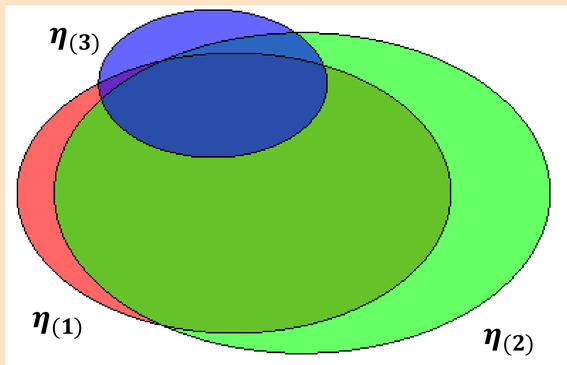
II. *Operating Characteristics:*

- Detection probabilities and false alarm probabilities for the three vetoes individually with current thresholds marked by vertical lines





- The current thresholds were chosen by analysis of results obtained by varying each of the three thresholds simultaneously
- Using the current thresholds, the overall (MES, RDS, and χ^2 vetoes) false alarm probability is just 4% while the detection probability is 98%
- This Venn diagram shows the overlap among the false alarm targets that are vetoed by the various χ^2 vetoes



- So you can see that $\eta_{(2)}$ is doing a majority of the work by vetoing 32,321 out of the total 35,616 false alarms vetoed by the three collectively
- $\eta_{(1)}$ uniquely vetoes an additional 1,667 while $\eta_{(3)}$ vetoes only an additional 1,483 uniquely

III. *Conclusion / Future Work:*

- Utilizing the MES, RDS, and the three χ^2 quantities, the estimated false alarm and detection probabilities are ~4% and ~98% respectively
- The false alarms that remain are largely due to a breakdown of the whitener which results in noise properties that don't match those assumed in both the detection theory and χ^2 veto theory, namely that the noise should be white, Gaussian, uncorrelated, and stationary (to first order)
  - Several avenues for fixing this issue and eliminating more false alarms are currently being explored (implementing a bootstrap in TPS to algorithmically set thresholds, using a robust smoothing spline to detrend the data)
- The statistical properties of $\chi^2_{(3)}$ need to be investigated to either determine how better to use it or perhaps not use it at all
- Injection studies need to be done to determine the effect that the vetoes have on detectability
  - Would like to perform these studies on idealized light curves as well as real light curves
- Need to think about how to better incorporate data de-emphasis weights into the calculation of the degrees of freedom for the various vetoes

References

[a] Jenkins, J.M. (2002) *ApJ*, *575*, 493-505.
[b] Seader et al. (2013) *ApJ Suppl*, *206*, 25

Discovery Mission 10 - Launched 2009 - http://Kepler.NASA.gov