The Breakthrough Listen Search for Intelligent Life: Detection and Characterization of Anomalous Transits in Kepler Lightcurves

Anna Zuckerman[1,2] , James R. A. Davenport[3] , Steve Croft[1,4] , Andrew Siemion[1,4,5,6] , and Imke de Pater[7]

[1] Breakthrough Listen, University of California Berkeley, Berkeley, CA 94720, USA; anna_zuckerman@alumni.brown.edu
[2] Department of Astrophysical and Planetary Sciences, University of Colorado Boulder, 2000 Colorado Avenue, Boulder, CO 80309, USA
[3] Astronomy Department, University of Washington, Box 951580, Seattle, WA 98195, USA
[4] SETI Institute, Mountain View, CA, USA
[5] Department of Physics and Astronomy, University of Manchester, UK
[6] University of Malta, Institute of Space Sciences and Astronomy, Msida, MSD2080, Malta
[7] Departments of Astronomy and Earth and Planetary Science, University of California Berkeley, Berkeley, CA 94720, USA

Abstract

Never before has the detection and characterization of exoplanets via transit photometry been as promising and feasible as it is now, due to the increasing breadth and sensitivity of time domain optical surveys. Past works have made use of phase-folded stellar lightcurves in order to study the properties of exoplanet transits because this provides the highest signal that a transit is present at a given period and ephemeris. Characterizing transits on an individual, rather than phase-folded, basis is much more challenging due to the often low signal-to-noise ratio of lightcurves, missing data, and low sampling rates. However, by phase folding a lightcurve we implicitly assume that all transits have the same expected properties, and lose all information about the nature and variability of the transits. We miss the natural variability in transit shapes, or even the deliberate or inadvertent modification of transit signals by an extraterrestrial civilization (for example, via laser emission or orbiting megastructures). In this work, we develop an algorithm to search stellar lightcurves for individual anomalous (in timing or depth) transits, and we report the results of that search for 218 confirmed transiting exoplanet systems from Kepler.

Unified Astronomy Thesaurus concepts: Exoplanet astronomy (486); Exoplanet catalogs (488); Exoplanet detection methods (489); Transit photometry (1709); Transits (1711); Search for extraterrestrial intelligence (2127); Light curves (918)

Supporting material: machine-readable table

1. Introduction

The expanding availability of photometric stellar lightcurves, for instance from the Kepler and Transiting Exoplanet Survey Satellite (TESS) missions, has provided the chance to discover and understand the orbital properties of thousands of exoplanet systems in recent years. Past studies have been limited to detection of transiting exoplanets or characterization of planetary and orbital properties from analysis of collectively phase-folded transits. The phase-folded shape of an exoplanetary transit provides a wealth of information about the planetary system that produced the transit, but can inherently only provide insight into the ensemble properties of all the observed transits of that planet. While some work has been dedicated to understanding the provenance of unusually shaped transits by reverse engineering the shape of a transiting body from the transit (Sandford & Kipping 2019), this has always been done after phase folding. No past work has comprehensively searched for and characterized missing or anomalous individual transits in stellar lightcurve data.

The population variability of exoplanet transits and the nature and prevalence of unexpected transits is poorly understood. This represents a major gap in the study of exoplanet populations. For instance, individual transits can be a powerful tool in studying stellar surface variability such as starspots (Morris et al. 2017).

In addition, unexpected or anomalous transits represent an important potential technosignature in the Search for Extraterrestrial Intelligence (SETI; Wright & Kipping 2019). Intelligent extraterrestrial civilizations may broadcast their presence by intentionally or inadvertently modifying the shape of their host planet's transit, for example with laser emission or megastructures (e.g., Arnold 2005; Kipping & Teachey 2016). In addition, observations discrepant with expectation often drive theoretical advancement, so the discovery of unusually shaped transits produced by unexplained phenomena would be scientifically interesting in and of itself and would necessitate new concepts in planetary or stellar astrophysics, if not SETI.

We present a new algorithm to analyze lightcurves on a transit-by-transit basis without phase folding, in order to study the variability between individual transits and search for transits that are missing or anomalous. We present the methodology and results of a search for transits which are apparently missing, or which exhibit unexpectedly large transit-timing variations (TTVs) or apparent depths. We find no transits with anomalous properties which cannot be explained after visual vetting.

2. Data, Preprocessing, and Target Selection

2.1. Data

The Kepler mission (Borucki et al. 2010) has produced 2708 confirmed detections of transiting exoplanet systems.[8]

---

[8] As of 2022 June 6, as reported by the NASA Exoplanet Archive (https://exoplanetarchive.ipac.caltech.edu).
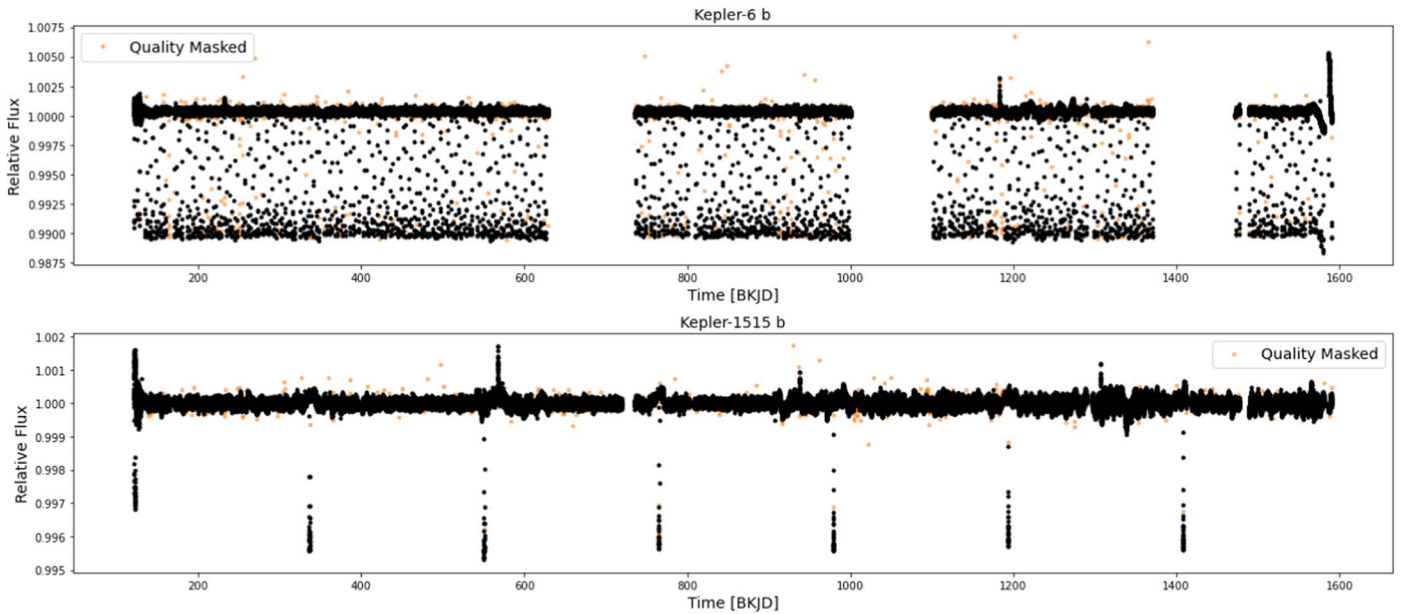
Figure 1. Example detrended, stitched, and masked lightcurves. Orange points are data that have been removed from our analysis due to quality issues flagged by the Kepler pipeline.

The program ran from 2009 to 2013, observing approximately 150,000 stars in multiple 90 days quarters, at a cadence of either 30 minutes or 60 s. The mission prioritized Main Sequence stars for which Earth-like planets would be detectable (Batalha et al. 2010). Lightcurves produced by the Kepler Mission Science Operations Center can be publicly downloaded from the Barbara A. Mikulski Archive for Space Telescopes (MAST) archive (STScI 2011). The Kepler Science Processing Pipeline is described in Jenkins et al. (2010). Lightcurves comprise measured flux as a function of time. They often contain extended intervals of missing data (due to the telescope entering safe mode, rotating toward Earth, or executing a quarterly roll) as well as individual data points flagged for quality issues (due to cosmic-ray hits, reaction wheel zero crossings, impulsive outliers, thruster firings, etc.).

2.2. Preprocessing

Before transits can be identified and anomalies detected, we first process each target lightcurve. For each planetary system, we download lightcurve data from the MAST using the interface provided by the `LightKurve` package. We use the default Presearch Data Conditioning Simple Aperture Photometry (PDCSAP) flux values that have been cotrended by the Kepler team using the PDC cotrending algorithm (Smith et al. 2012; Stumpe et al. 2012, 2014). Kepler's observing run is divided into quarters, punctuated by rolls of the telescope. We fit and remove a linear trend from each quarter, and stitch the quarters into one continuous lightcurve (see Figure 1). We then mask out data with quality issues flagged during Kepler data acquisition. Each lightcurve corresponds to one stellar target which may or may not be host to more than one exoplanet. So, for each target planet we mask out any additional transits from "sibling planets," which might interfere with our fitting process. This masking uses a folding and Box-Least-Squares fitting approach, and so relies on the first-order assumption that all transits appear near their expected locations. We acquire a list of sibling planets by querying the NASA Exoplanet

Archive database[9] (NASA Exoplanet Archive 2019). On very rare occasions, key parameters returned from the NASA exoplanet archive were null, preventing the masking of a sibling planet. In these cases, sibling transits that are not properly masked can interfere with the fitting. This failure mode is vetted in Section 4 by querying the Simbad archive when additional siblings are suspected.

2.3. Target Selection

Because we analyze transits individually, we are strongly limited by the depth and duration of the transit signal relative to the local noise. In order to maximize the statistical significance of any characterization we make, we sort all confirmed Kepler targets by the ratio of the expected transit depth to the scatter in the flux (defined as the the median rolling standard deviation (SD) of the flux values, with a window size of 10 timesteps), and take only the best 10% of targets. This corresponds to a threshold ratio of 6.73. We choose 10% as our threshold because targets below this cutoff tend to have transits with depths similar to the lightcurve noise (less than a few times the noise), and thus are not good candidates for unambiguous detection of missing transits. If the expected depth for a particular transiting system is near or below the average scatter of the lightcurve, we cannot make any significant statement about the nature of an individual transit. This produces 228 targets. We determine the expected depth by fitting the ensemble of transits together for each planet to produce an average transit model for that planet. We download the reported transit properties (period, semimajor axis, inclination, eccentricity, argument of periastron, impact parameter, transit duration, ephemeris, and planet and star radii) for each planetary system from the NASA Exoplanet Archive, and use these as initial parameters for fitting. We perform the transit fitting using a `Batman` transit model wrapped in Python's `lmfit` method. The `Batman` package (described fully by Kreidberg 2015) produces a transit model based on input

---

planetary and orbital parameters, and `lmfit` allows us to fit for the best `Batman` transit model using a brute-force method over a grid of points. The ensemble fit parameters sometimes differ significantly from the reported NASA transit parameters. Because we are concerned only with deviations in the transit depth or midpoint for individual transits, we are not concerned with defining physically accurate transit parameters, but rather with defining the model that best represents the depth and timing of transits for each planetary system as a whole. We take the depth of this ensemble model as the expected transit depth for each individual transit. By selecting the systems with the highest ratio of expected transit depth to flux scatter, we aim to produce the highest confidence results possible.

3. Methods

3.1. Expected Transit Parameters

Before fitting individual transits, we must first define a set of expected parameters to serve as initial guesses for individual transit fits. We take the preliminary set of parameters determined for the purpose of target selection (as described in Section 2.3), and inspect the fits produced by those parameters by eye. In the case of a poor fit, we manually update the fit parameters to produce a better expected transit shape. This allows us to have high confidence that the expected transit models well represent an "average" transit, and thus provide a good baseline from which to define unusual transit shapes. We refer to this set of parameters as the "ensemble fit parameters" because they provide our best fit to the entire ensemble of transits for a planet.

3.2. Anomaly Search

After preprocessing, the lightcurves are ready to be searched for missing or unusual transits. For each star, we iterate over each transit. For each expected transit midpoint (with midpoints determined using the period and ephemeris provided by the NASA Exoplanet Archive in the Kepler KOI table (Thompson et al. 2018), we define a window spanning three times the expected transit duration on either side of the expected midpoint. If insufficient data (defined as less than 70% of the expected number of data points in the window) is present in a given transit's window due to data gaps or data quality issues, we ignore that transit. Even so, missing or sparse data is one of the most common reasons for poor fitting of individual transits.

We then fit the transit with a `Batman` transit model. We first mask the expected location of the transit in order to fit and remove a linear trend as a first-order correction for stellar variability. Short-timescale nonlinear stellar variability is not removed, and this is another common reason for poor fitting. Next, we produce a preliminary model of the transit in order to update the expected midpoint, and repeat the process of masking the transit and removing a linear trend. This helps avoid errors in baseline trend removal when the transit is not centered at the expected location. We then fit a final `Batman` model to the transit to produce a set of fit parameters, using the ensemble model parameters as initial guesses.

As a reference of comparison for individual transits, we take the median of each parameter produced in this final `Batman` fit over all transits and refer to this set of parameters as the "median fit parameters." This set of median fit parameters provides the standard against which we define which individual transits are considered anomalous.

Once each transit has been fit, we flag the transits that meet certain requirements. We define three flags: (1) transits that appear "missing," (2) transits that appear unusually deep, and (3) transits that appear to have unexpectedly large TTVs. These are defined quantitatively as follows:

1. Fit transit depth shallower than 3 times the flux SD, and greater than 3σ away from the median of the fit transit depths (the transit is effectively indistinguishable from a flat section of lightcurve).
2. Fit transit depth exceeds the median fit transit depth by more than 3σ.
3. Fit transit TTV exceeds the median fit TTV by more than 5σ. We choose 5σ in this case to prevent flagging expected large TTVs due to coherent variation in transit timing from sibling planets.

Defining the flagging criteria based on the actual distribution of fitted transit properties circumvents issues with small discrepancies between the ensemble and mean fit parameters that sometimes arise due to imperfect expected parameter values. It also allows us to flag TTVs even in several-planet systems which are often expected to have large TTVs due to the gravitational effects of planets on their siblings' orbits. Figure 2 demonstrates the threshold applied for each flag.

3.3. Signal Injection and Recovery

In order to assess our confidence in flags produced by our anomaly search for any given planetary system, we perform an injection and recovery of simulated transits and evaluate both the accuracy with which we can recover their depths and timings, and the accuracy of the flagging processes. We would like to be able to reject flagged transits in systems for which the injection and recovery shows it is difficult to accurately determine the transit parameters, or systems for which many transits are erroneously flagged.

We inject transits into each lightcurve as follows. First, we define a set of 1000 injection locations chosen randomly across the lightcurve under the requirement that sufficient data is present in the surrounding window to perform a `Batman` fit. We then choose a sample randomly from a distribution of depths and TTVs without replacement to produce the injection parameters. The depth parameter space spans from zero to 3.5σ above the median fit depth, and the TTV parameter space spans from zero to 1.1 times the size of the fitting window. We choose this range to ensure that we sample from all regions that should produce flags, as well as regions expected to erroneously produce flags. We inject each transit by defining a `Batman` model using the chosen depth and TTV, and median fit parameter values for the other parameters, and add the injection model onto the existing lightcurve in order to preserve noise and stellar variability.

We then attempt to recover each injection by performing the fitting exactly as described in Section 3.2. Once all injections are fit the flags are defined exactly as described above, using the median and standard deviation of the real individual transit fits to determine the threshold values for flagging. We calculate the rms error between the injected and recovered depths and TTVs for each planet, and produce diagnostic plots to visualize the recovery accuracy (as shown in Figure 3). For this paper, we define "true positive" injections to be injections whose injected depth or TTV is within the region that should produce a flag, and whose recovered depth or TTV does in fact produce
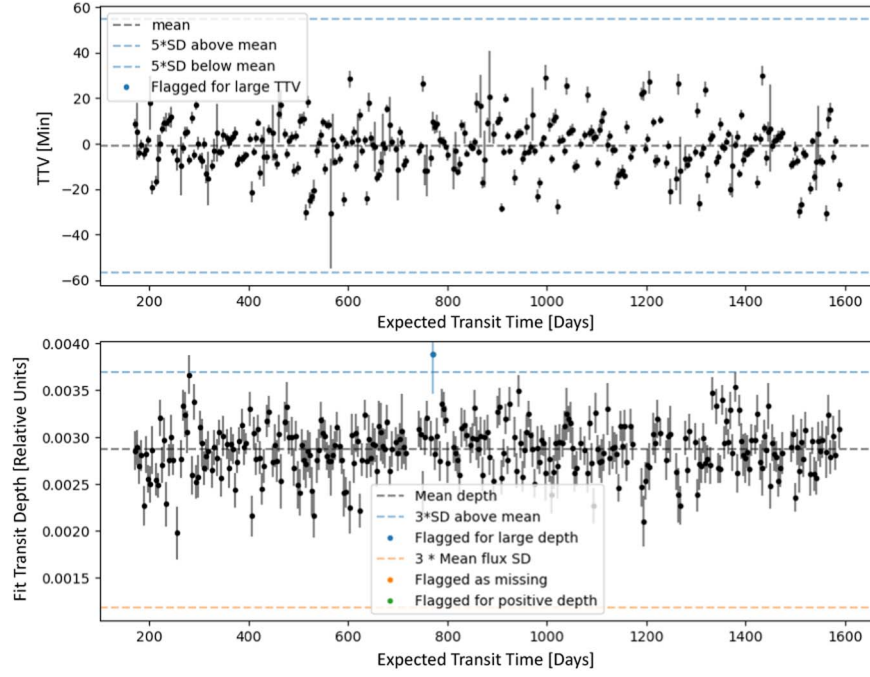
Figure 2. Demonstration of the thresholds used for flagging fit transit parameters for an example planet (Kepler-696b). Each data point represents one transit.
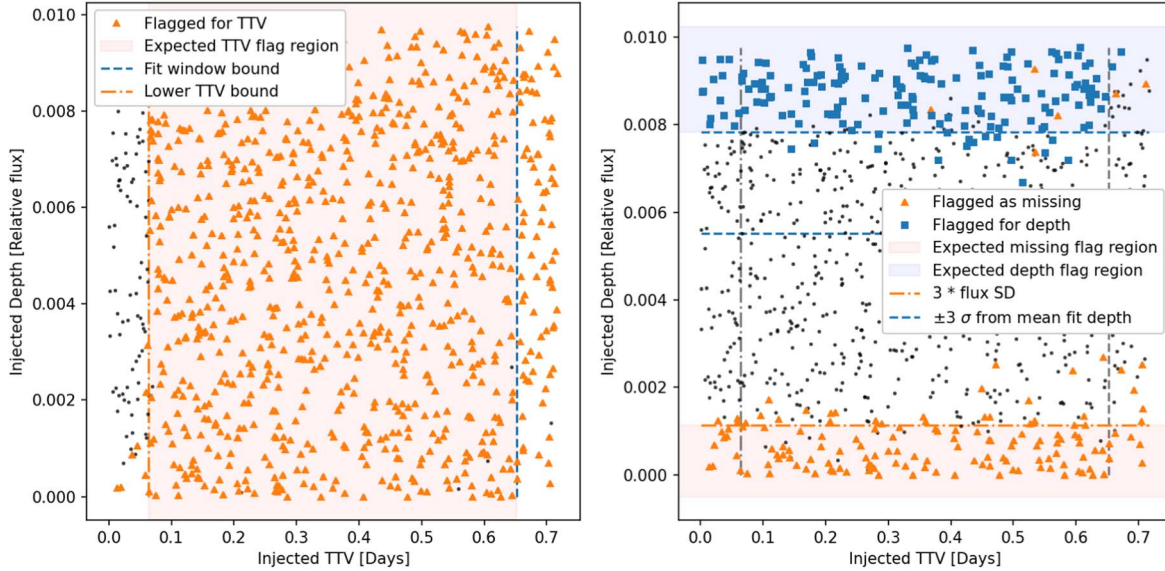

Figure 3. Scatter plot demonstrating injection and recovery accuracy for a sample lightcurve. Each point represents the injected properties of recovered transits. In the left-hand plot, orange triangles represent transits flagged for large TTVs, and the orange filled area represents the region of parameter space which should produce a TTV flag. In the right-hand plot, orange triangles represent transits flagged as missing, blue squares represent transits flagged for large depths, and the orange and blue regions represent the corresponding regions of parameter space expected to produce flags. Not all transits injected in the shaded regions are flagged due to imperfect recovery of the injected properties.

the expected flag. We define "true negative" flags analogously, as transits whose injected parameters should not produce a flag, and whose recovered parameters indeed are not flagged. "False positive" injections are those whose parameters should not produce a flag, yet the recovery is inaccurate and the recovered transit is flagged. Similarly, "false negative" injections are those whose parameters should produce a flag, but due to poor recovery are not flagged.

We perform this injection and recovery process for each planetary system. We can then use the accuracy of the injection recovery to inform vetting of flagged transits.

4. Results

We run the individual transit fitting algorithm described in Section 3.2 on each of the 228 planetary targets described in Section 2. The `Batman` model fails to converge for 10 of these targets, leaving 218 which we can search for anomalous transits. This means we search 39,879 total transits. Using the flags defined above, we flag six transits as apparently missing, 92 as unexpectedly deep, and 51 as having unexpectedly large TTVs.

We then visually inspect these flagged transits. We reject flagged transits using several criteria, which vary slightly for
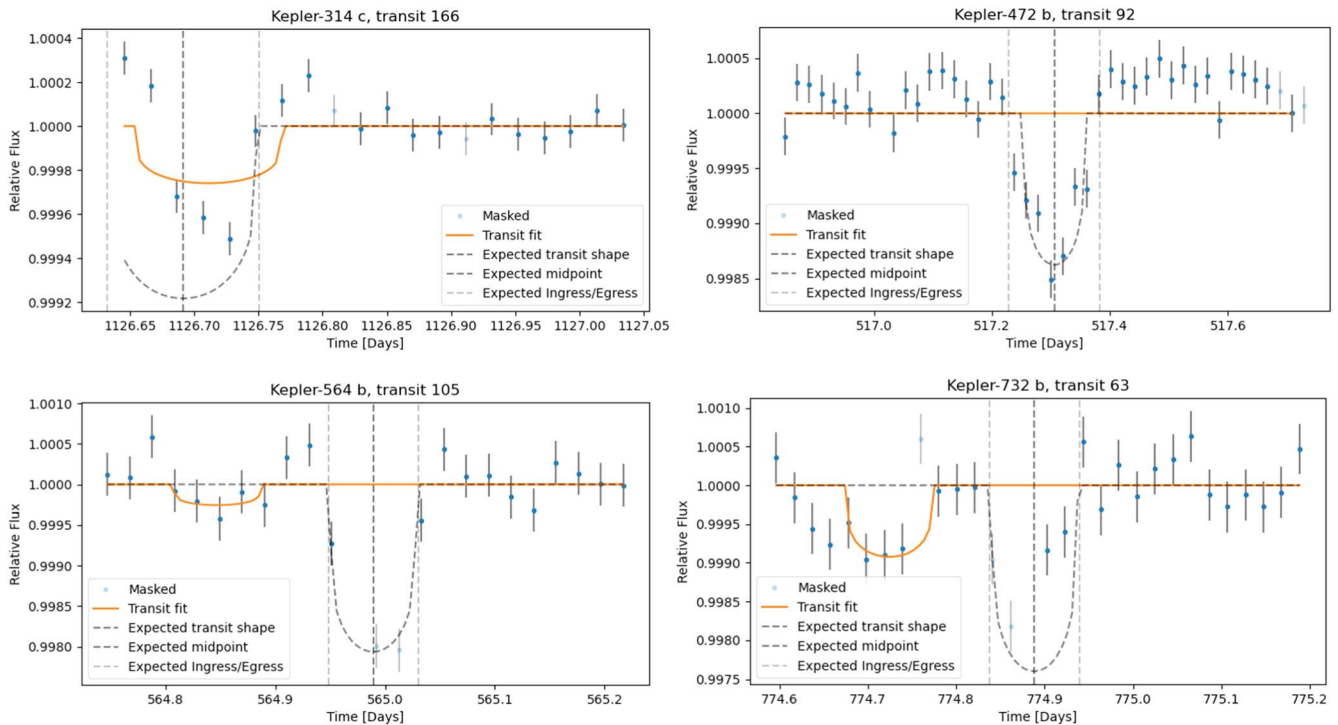
Figure 4. Four of the six transits flagged as missing are clearly present on visual inspection of the data. Top left: Kepler-314c, transit 166, is flagged as missing but is clearly visible in the raw data. Missing data at the edge of the fit window and short-timescale stellar variability make the fitting difficult and produce a false positive. Top right: Kepler-472, transit 92 is flagged as missing yet is clearly present on visual inspection. Bottom left: Kepler-564, transit 105 is mis-fit due to two quality masked data points exactly at the transit minimum, which effectively hide the transit. Bottom right: similarly, the model fits adjacent stellar variability instead of transit 63 of Kepler-732b because of a single missing data point in the fit, as well as short-timescale stellar variability which alters the shape of the transit.

each flag. In fitting individual transits, we are often challenged by the sparsity of data around the expected transit. Thus, all flags have some false positives when the transit is truly present in the data with its expected shape, but a good fit to the transit cannot be produced due to missing or sparse data. Another complication for all flags arises from sibling transits which fall within the fitting window. Though sibling transits are masked before fitting, in a few cases sibling transits with large TTVs cannot be perfectly masked beforehand and can interfere with fitting.

4.1. Transits Flagged as Missing

Our primary interest is to identify and examine potentially "missing" transits. In general, we find that despite the noisiness and sparseness of data for individual transits, we are able to accurately determine that a transit is present at each expected location and to fit its depth and TTV. However, some transits are flagged as missing.

For such missing transits, one possible false positive comes from poor fitting due to missing or sparse data in the region around the transit. Four of the six transits flagged as missing are clearly present in the data on visual inspection, but are fit poorly due to missing data and/or stellar variability. These are shown in Figure 4.

The two remaining candidates are both transits of the planet Kepler-1655b (Figure 5). For these two, the transit is not visually identifiable in the processed data. The first step for further visual vetting is an inspection of the unprocessed (i.e., with no linear detrending) lightcurve in a wider window around the flagged transit. When inspected this way, transit 22 has a visually identifiable feature at the location of the fit transit that matches the expected transit shape, and this can be identified as

the transit itself. A wide section of missing data beginning three data points to the right of the transit egress, in conjunction with local stellar variability, impairs accurate linear trend removal. The trend that is removed artificially raises, and therefore obscures, the transit. For transit 17, the cause of the flag is somewhat less clear. However, inspecting the wider region around the transit reveals stellar variability on the same order as the expected transit depth. This is an edge case for which the timescale of the stellar variability was too short to be detrended, yet too long to be included in the calculation of the SD. Thus we must also reject this transit as a statistically significant "missing" transit.

Though we are most interested in potentially missing transits, our data set is also highly amendable to a search for transits with unusually large depths or TTVs. These are discussed in the following two sections.

4.2. Transits Flagged as Unexpectedly Deep

One possible false positive for transits flagged as unexpectedly deep comes from poor fitting due to missing or sparse data. In the case of missing data, there is an additional failure mode that occurs when the midpoint is fit inside a region of missing data, and the depth is then essentially unbounded. Of the 92 flagged transits, 23 are rejected as false positives for this reason. An additional 31 are rejected because nonlinear stellar variability results in a mischaracterization of the local baseline, and thus an apparently deeper transit. Also, 19 transits are rejected because the scatter in the out-of-transit lightcurve is large compared to the difference between the fit depth and expected depth; thus, the excess depth could be explained by random variability rather than a true increase in transit depth.
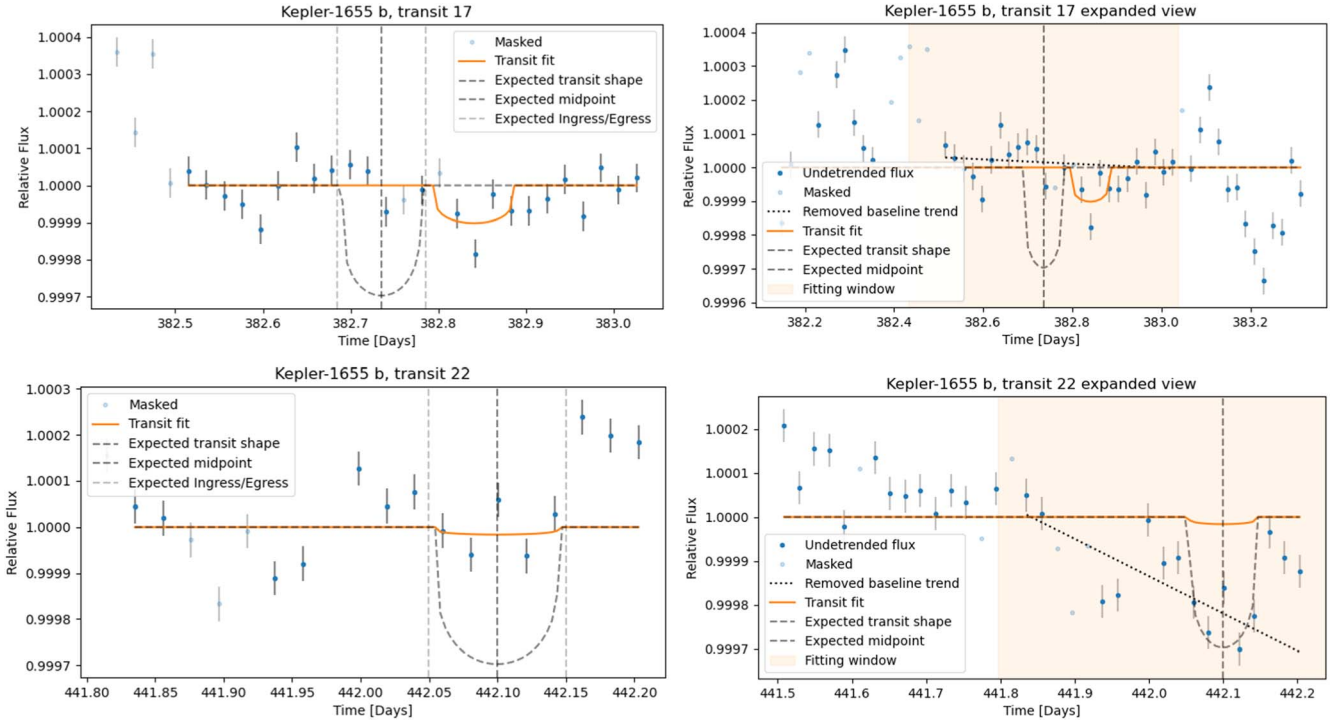
Figure 5. Kepler-1655b transits 17 (top) and 22 (bottom) are flagged as "missing" and are not obviously present on visual inspection, but are ultimately rejected due to poor linear trend removal and due to stellar variability which could mask a present transit.

Large instrumental scatter can also make the linear detrending unreliable for these transits.

Another class of false positive comes from sibling transits which overlap with the target transit, deepening the apparent transit. Though we attempt to mask these sibling transits, we are unable to initially mask a sibling transit properly in cases when the transit exhibits a very large TTV, or in the rare case that a search of the NASA exoplanet archive returns null parameters. We vet for the first failure mode by overplotting the individually fit midpoints of resolved sibling planets, and reject a single transit (Kepler-247c transit 113), which clearly has an overlapping sibling. To check for the second failure mode, we manually search the Simbad database for siblings and in two cases find an unmasked sibling planet. From this we can reject five flagged transits of Kepler-18c, because visual inspection reveals the transits of Kepler-18d clearly move in and out of the fitting window. Similarly, we reject Kepler-51b transit 30 because an additional sibling (Kepler-51d) is reported in Simbad and has an expected transit midpoint that directly overlaps.

We reject another four transits because analysis of the injection and recovery process reveals that accurate recovery of transit parameters is difficult. If for a given system we are unable to recover injected deep transits, or we flag normal transits as unusually deep, we cannot have high confidence in any transit flagged for that system and so must reject it. We impose a threshold on the fraction of false positive injections, out of the total number of injections which should not be flagged. We choose that threshold to be 10% in order to reject the clear tail of outliers with especially poor injection recovery, while retaining the majority of transits. This test allows us to reject transits of Kepler-25c, Kepler-17b, Kepler-685b, and Kepler-506b. This leaves nine transits.

The next step for visual vetting is an inspection of the unprocessed (i.e., with no linear detrending) lightcurve flux in a
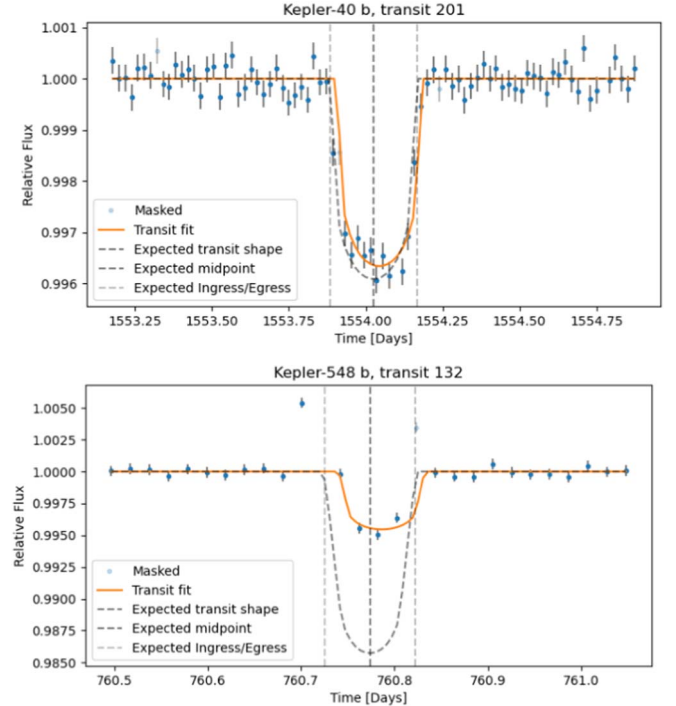

Figure 6. Kepler-40b transit 201 and Kepler-548b transit 132 are flagged for large TTVs and pass the visual inspection of the fit.

wider window around the flagged transit. When inspected this way, several transits can be clearly rejected based on one of the previously identified failure modes. When the predetrending lightcurve is inspected in a wider window, it becomes apparent that Kepler-314c transit 109 and Kepler-682b transit 41 have short-timescale stellar variability with a timescale comparable to the transit duration and with amplitude comparable to the
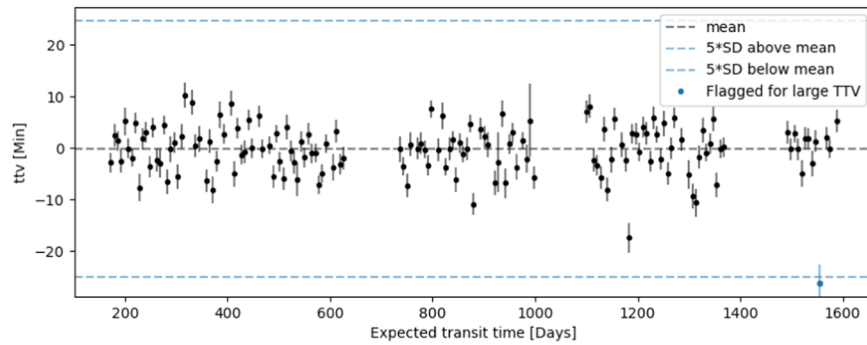
Figure 7. Kepler-40b transit 201 is barely above the threshold for flagging, and consistent with contamination from stellar variability.
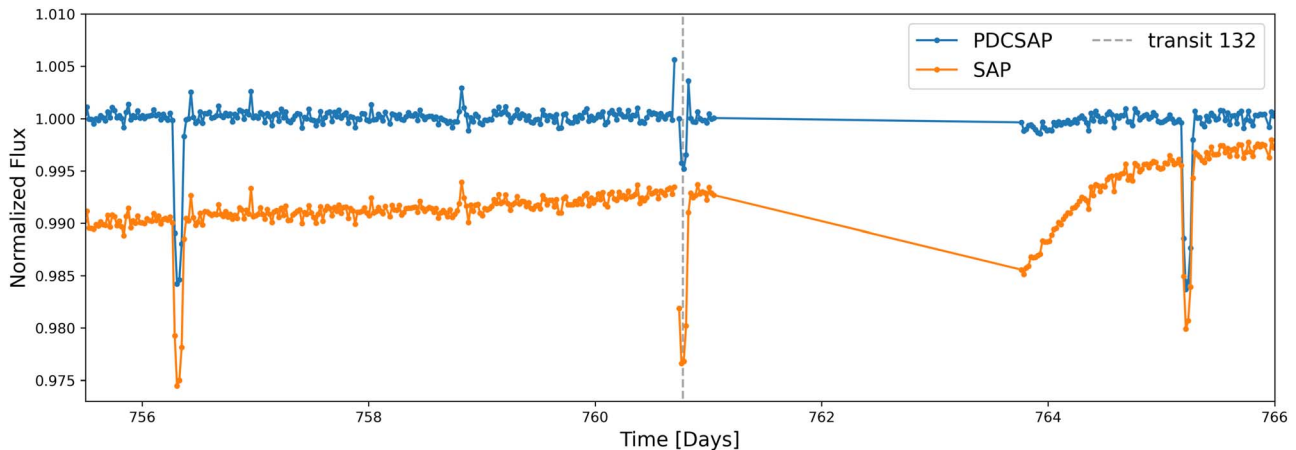

Figure 8. Kepler-548b, transit 132, is an apparently very unusual transit with both a large TTV and a shallow depth. However, the transit appears to have a normal shape when viewed using the precotrending SAP flux values, and thus the anomaly can be credited to artifact of the Kepler cotrending process.

difference between the fit and expected depth. Thus, the apparently large depth could be caused by unfortunately timed stellar variability.

This leaves five remaining transits (Kepler-685b transit 791, Kepler-422b transit 99, Kepler-74b transit 56, Kepler-433b transit 248, and Kepler-41b transit 616). Although these transits do not fall under any of the failure modes described above, they are not anomalous enough to warrant further investigation. All five of these transits appear to fall on the tail of a normal distribution of fit depths for their respective systems, or result from small coherent trends in an otherwise normal fit depth distribution (which may occur due to moving starspots or other stellar surface features).

4.3. Transits Flagged for Large TTVs

As in the previous two cases, for transits flagged for large TTVs a possible false positive comes from poor fitting due to missing or sparse data near the transit. We reject 27 of the 51 transits flagged for large TTVs for this reason. We reject another 12 that were fit poorly due to short-timescale stellar variability or single missing data points that bias the fitting. For a further set of five, short-timescale stellar variability produces a feature that mimics the expected transit shape enough to "fool" the `Batman` model into fitting the variability rather than the transit itself.

Another class of false positive comes from sibling transits present in the fitting window, which are fit instead of the target

transit. We reject the same five transits of Kepler-18c described in Section 4.2 because a Simbad search reveals a sibling planet unresolved in the NASA Exoplanet Archive whose transits coincide with the flagged transits.

This leaves two remaining transits, which are shown in Figure 6.

Transit 201 is only 5% above the 5σ threshold, but does not obviously fall on the expected tail of a normal distribution of TTVs, as shown in Figure 7. However, Kepler-40b is a rotating variable subgiant star, and thus likely to exhibit starspots. Transit 201 is consistent with the shape expected from contamination due to the exoplanet crossing a starspot feature (Morris et al. 2017).

Kepler-548b transit 132 is flagged for its large TTV, but is most unusual in its dramatically reduced depth. Transits in the system are generally fit well, and no other transit exhibits a shallow depth. No unexpected sibling transit appears in the fitting window (nor could this explain both the large TTV and the shallow depth), and the lightcurve appears free of stellar flaring or other activity which could give the transit an apparently shallower depth. However, when the lightcurve for this transit is created using the raw (SAP) flux values (Twicken et al. 2010; Morris et al. 2020), rather than the PDCSAP flux values, the transit no longer appears unusually shallow (Figure 8). This appears to be a rare case in which cotrending applied to the PDCSAP values significantly modifies the shape of a transit, producing a spurious flag.

5. Discussion and Conclusions

To the best of our knowledge, no past work has attempted to characterize the shapes of transits individually in stellar lightcurves, though some have searched individually for TTVs. We have developed and implemented an algorithm to search stellar lightcurves for unusual transits, and have performed a search for such transits on 218 Kepler transiting exoplanet systems. We find and vet six candidate transits flagged by our algorithm as missing, 51 flagged as having an unusually large TTV, and 92 flagged as having an unusually large depth. We visually inspect all flagged transits, and reject each as a likely indication of new astrophysics or extraterrestrial intelligence.

Our nondetection allows us to put a useful upper limit on the occurrence rate of transits with inexplicably anomalous properties. Zero of 218 stars host a transit which passed our search algorithm and could not be rejected via manual inspection; the upper limit on detections at 95% confidence with zero detected is 2.99. We conclude that less than 2.99 per 218 Kepler stars, or less than 1.4%, host transits with inexplicably large depths or TTVs, or transits which appear to be missing. In choosing our sample we are biased toward planets with deep transits and lightcurves with low flux uncertainties, but this should not bias our detection rates. We are however only sensitive to a small subset of all transits of each planet due to finite Kepler observing times. The median duration of a lightcurve in our target set is 1459.5 days. Thus, our upper limit should be understood to reflect an occurrence rate per 1459.5 days.

Our search is designed to be agnostic to the cause of any anomalous transit shapes. We attempt to reject known astrophysical causes, for instance by rejecting transits where sibling planets, stellar flares, or short-timescale stellar variability can cause apparent anomalies. However, had we been unable to reject any flagged transits an unexpected astrophysical explanation would still be possible. Our search would also be sensitive to nonastrophysical transit anomalies, which represent a possible technosignature. A civilization might choose to cloak its transits from known neighbors, or else broadcast its presence intentionally or unintentionally by modifying its transit shape with megastructures or lasers (Wright & Kipping 2019). While Kipping & Teachey (2016) have noted directed emission from a laser could "cloak" or modify a transit of a planet in the Kepler band using technology presently available, multiband detection of transits would be able to distinguish this type of variability compared to megastructures. In addition, if an intelligent civilization aware of life near our Sun were to attempt to send a beacon or information toward the Sun, it would have to choose the best time to do so. The interval in which the civilization's home planet transits between its host star and the Sun would be clever choice of timing if this civilization expects others to be actively studying the transits of exoplanets for non-SETI purposes. Thus, the interval around a transit is an especially important time to search for SETI signals (Franz et al. 2022).

5.1. Future Directions

The techniques implemented in the work could be readily applied to other large data sets, such as the TESS catalog of transiting exoplanet lightcurves (Borucki et al. 2010). In addition, the set of transit fits produced as a result of this work could be very useful as a training data set in a machine-learning search for unexpected transits in stellar lightcurves. Past exoplanet searches miss transits from planets with periods longer than the survey observing time, comets or asteroids, or rogue planets. The nature and prevalence of unexpected transits is poorly understood. This represents a major gap in our understanding of the occurrence rate of long-period exoplanets. These long-period planets are likely to be gas giants whose key role in the formation and evolution of planetary systems is an active area of research. Each fit transit from this work can become a labeled positive observation, and fit section of lightcurve with no known transit labeled negative observations, to train a machine-learning algorithm to search pixel by pixel through stellar lightcurves for transits which appear where none is expected.

Future work could improve on some of the methods used here in order to reduce the need for visual vetting, or to expand the transit fitting approaches. For instance, a more sophisticated method of jointly fitting transits and stellar variability using Gaussian processes or wavelet analysis could better detrend each transit fitting window and thus mitigate some false positive candidates we encountered. In addition, a future work might benefit from jointly fitting the transits of each sibling planet in a system in order to avoid the necessity of masking the transits of sibling transits in each fitting.

Appendix
Summary Table

Table 1 contains a summary of the transit fitting parameters and results of injection and recovery.

Table 1
System Information, Including Median Transit Fit Parameter Values, and Injection and Recovery Statistics for each Exoplanet in our Sample

| Name | Siblings | # Transits | Median Fit Parameters | | | | | | Injection/recovery Summary | |
			Rp/Rs	Semimajor axis [AU]	Inc. [Deg]	Ecc.	Arg. of Periastron	Rel. Depth	Depth Recovery RMS [Days]	TTV recovery RMS
Kepler-117 b	c	77	0.046	21.55	89.47	0.0	254.3	0.00252	0.00041	0.313
Kepler-396 c	b	16	0.049	79.64	89.47	0.0	0.0	0.00245	0.00115	0.437
Kepler-782 b	None	9	0.026	104.61	90.0	0.0	0.0	0.00085	0.00014	0.49
Kepler-39 b	None	66	0.091	23.72	88.42	0.0	98.9	0.00887	0.00139	0.252
Kepler-108 c	b	7	0.031	70.92	89.47	0.0	0.0	0.00101	0.00016	0.595

Note. A subset is shown here as an example.

(The full table is available in machine readable format)

ORCID iDs

Anna Zuckerman ⓘ https://orcid.org/0000-0002-2412-517X
James R. A. Davenport ⓘ https://orcid.org/0000-0002-0637-835X
Steve Croft ⓘ https://orcid.org/0000-0003-4823-129X
Andrew Siemion ⓘ https://orcid.org/0000-0003-2828-7720
Imke de Pater ⓘ https://orcid.org/0000-0002-4278-3168

References

Arnold, L. F. A. 2005, ApJ, 627, 534
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010, ApJL, 713, L109
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Sci, 327, 977
Franz, N., Croft, S., Siemion, A. P. V., et al. 2022, AJ, 163, 104
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, ApJL, 713, L87

Kipping, D. M., & Teachey, A. 2016, MNRAS, 459, 1233
Kreidberg, L. 2015, PASP, 127, 1161
Lightkurve Collaboration, Cardoso, J. V. d. M., Hedges, C., et al. 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library, ascl:1812.013
Morris, B. M., Hebb, L., Davenport, J. R. A., Rohn, G., & Hawley, S. L. 2017, ApJ, 846, 99
Morris, R. L., Twicken, J. D., Smith, J. C., et al. 2020, in Kepler Data Processing Handbook: Photometric Analysis, Kepler Science Document KSCI-19081-003, ed. J. M. Jenkins (Moffett Field, CA: NASA/Ames Research Center), 6
NASA Exoplanet Archive 2019, Kepler Objects of Interest Cumulative Table, Last Accessed 2023-11-22, IPAC, doi:10.26133/NEA4
Sandford, E., & Kipping, D. 2019, AJ, 157, 42
Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, PASP, 124, 1000
STScI 2011, Kepler/KIC, doi:10.17909/T9059R
Stumpe, M. C., Smith, J. C., Catanzarite, J. H., et al. 2014, PASP, 126, 100
Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, PASP, 124, 985
Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, ApJS, 235, 38
Twicken, J. D., Clarke, B. D., Bryson, S. T., et al. 2010, Proc. SPIE, 7740, 774023
Wright, J., & Kipping, D. 2019, BAAS, 51, 343, https://baas.aas.org/pub/2020n3i343