2017-01-10

Kepler transit depths contaminated by a phantom star

Dalba, Paul A.

IOP Publishing

http://hdl.handle.net/2144/22289

Boston University
KEPLER TRANSIT DEPTHS CONTAMINATED BY A PHANTOM STAR

PAUL A. DALBA\textsuperscript{1}, PHILIP S. MUIRHEAD\textsuperscript{1}, BRYCE CROLL\textsuperscript{2}, and ELIZA M.-R. KEMPTON\textsuperscript{3}

\textsuperscript{1}Department of Astronomy, Boston University, 725 Commonwealth Avenue, Boston, MA 02215, USA; pdalba@bu.edu
\textsuperscript{2}Institute for Astrophysical Research, Boston University, Boston, MA 02215, USA
\textsuperscript{3}Department of Physics, Grinnell College, Grinnell, IA 50112, USA

Received 2016 September 9; revised 2016 November 21; accepted 2016 December 5; published 2017 January 10

ABSTRACT

We present ground-based observations from the Discovery Channel Telescope (DCT) of three transits of Kepler-445c—a supposed super-Earth exoplanet with properties resembling GJ 1214b—and demonstrate that the transit depth is \(~50\%\) shallower than the depth previously inferred from Kepler spacecraft data. The resulting decrease in planetary radius significantly alters the interpretation of the exoplanet’s bulk composition. Despite the faintness of the M4 dwarf host star, our ground-based photometry clearly recovers each transit and achieves repeatable 1\(\sigma\) precision of \(~0.2\%\) (2 millimags). The transit parameters estimated from the DCT data are discrepant with those inferred from the Kepler data to at least 17\(\sigma\) confidence. This inconsistency is due to a subtle miscalculation of the stellar crowding metric during the Kepler pre-search data conditioning (PDC). The crowding metric, or CROWDSAP, is contaminated by a non-existent phantom star originating in the USNO-B1 catalog and inherited by the Kepler Input Catalog (KIC). Phantom stars in the KIC are likely rare, but they have the potential to affect statistical studies of Kepler targets that use the PDC transit depths for a large number of exoplanets where an individual follow-up observation of each is not possible. The miscalculation of Kepler-445c’s transit depth emphasizes the importance of stellar crowding in the Kepler data, and provides a cautionary tale for the analysis of data from the Transiting Exoplanet Survey Satellite, which will have even larger pixels than Kepler.

Key words: planets and satellites: fundamental parameters – planets and satellites: individual (Kepler-445c) – stars: individual (KIC 9730163, KOI 2704, Kepler-445) – techniques: photometric

Supporting material: data behind figures

1. INTRODUCTION

Among the many discoveries made by the Kepler Mission (Borucki et al. 2010; Koch et al. 2010) is the ubiquitous nature of exoplanets with masses and radii in between those of the Earth and Neptune. Despite their absence from the solar system, planets smaller than Neptune, but larger than the solar system’s terrestrial planets, are surprisingly common in the Galaxy (Cassan et al. 2012; Fressin et al. 2013; Petigura et al. 2013). Several subdivisions of this regime have been created (e.g., Super-Earths, mini-Neptunes), but there may exist a continuum of planets in this portion of parameter space, blurring the line that previously separated rocky terrestrial and gaseous giant planets.

An interesting member of this class of exoplanets is Kepler-445c. Kepler-445c was first identified as a planet candidate by Burke et al. (2014) and was later confirmed and characterized by Muirhead et al. (2015). It was reported to be a 2.5\(R_\oplus\) planet orbiting a cool, metal-rich M4 dwarf star \(~90\) pc away with an orbital period of 4.87 days (Muirhead et al. 2015). The properties of Kepler-445c and its host star are similar to the GJ 1214 system, which is one of the most extensively studied exoplanet systems to date. Kepler-445 has a Kepler magnitude of 17.475, making it a relatively faint exoplanet host. This magnitude is slightly above the completeness limit of the Kepler Input Catalog (KIC, Brown et al. 2011) and results in relatively low-precision light curves compared to the majority of Kepler targets.

Detailed atmospheric investigations have occurred for a handful of exoplanets intermediate in size between the Earth and Neptune (e.g., Fraine et al. 2014; Knutson et al. 2014a, 2014b; Kreidberg et al. 2014; Tsiaras et al. 2016), but the vast majority, including Kepler-445c, have been classified by their planetary radii inferred from Kepler transit observations. The light curves returned by Kepler comprise an invaluable contribution to our understanding of planetary systems, but their analysis is complex and requires a careful assessment of many potential sources of error. The Kepler Science Operations Center (SOC) pipeline that converts the raw pixel-level data to the simple aperture photometry (SAP) light curves or pre-search data conditioning (PDCSAP) light curves has been thoroughly documented (Bryson et al. 2010; Jenkins et al. 2010; Quintana et al. 2010; Twicken et al. 2010; Kimemuchi et al. 2012; Smith et al. 2012; Stumpe et al. 2012). The PDCSAP light curves undergo extensive artifact mitigation and are well suited for the characterization of transiting exoplanets.

Because each pixel on the Kepler detectors covers a 4” \(\times\) 4” patch of sky and the target apertures typically consist of several pixels, a critical step in the pre-search data conditioning is the correction for crowding, or flux contamination, from sources other than the target star. Indeed, adaptive optics imaging of 3313 Kepler planet candidate hosts suggests that 14.5 \(\pm\) 0.8\% have a nearby star within 0”15 to 4”0 (Ziegler et al. 2016). The susceptibility of the Kepler data to this source of contamination is well known, and several recent studies have discussed transit dilution by unresolved companion stars (e.g., Rowe et al. 2014; Cartier et al. 2015).

Correcting for crowding requires accurate positions and magnitudes of all the sources in Kepler’s field of view. In most versions of the SOC pipeline, this information comes from the KIC. The crowding correction also requires the expected pixel response functions (PRFs) of the detectors over which the...
Table 1
Summary of DCT-LMI Observations of Kepler-445

<table>
<thead>
<tr>
<th>Date</th>
<th>Filter</th>
<th>Duration (hr)</th>
<th>Exposure Time (s)</th>
<th>Airmass</th>
<th>Aperture Radii (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015 Jun 19</td>
<td>Sloan i'</td>
<td>3.87\textsuperscript{b}</td>
<td>25</td>
<td>1.99–1.03</td>
<td>12, 20, 30</td>
</tr>
<tr>
<td>2015 Jul 23</td>
<td>Sloan z'</td>
<td>3.54\textsuperscript{c}</td>
<td>14</td>
<td>1.02–1.60</td>
<td>9.5, 20, 30</td>
</tr>
<tr>
<td>2015 Jul 28</td>
<td>Sloan i'</td>
<td>3.88\textsuperscript{c}</td>
<td>30</td>
<td>1.21–1.02–1.06</td>
<td>9, 20, 30</td>
</tr>
</tbody>
</table>

Notes.
\textsuperscript{a} The values listed in this column correspond to the radii of the photometric aperture, the inner boundary of the sky annulus, and the outer boundary of the sky annulus, respectively.
\textsuperscript{b} The final 0.40 hr of the post-transit baseline were excluded from the analysis due to contamination with a partial transit of Kepler-445d.
\textsuperscript{c} The final 0.17 hr of the post-transit baseline were excluded from the analysis due to a malfunction of the telescope guide camera.

fl

fluxes of the stars will be spread (Bryson et al. 2010). As all of this information from various modules of the pipeline is combined, the chance of an error entering and propagating through the process increases.

For this reason, many efforts have been made to remove spurious detections in the Kepler light curves: Torres et al. (2011) introduced the BLENDER technique to search for blended foreground or background eclipsing binaries diluted by the target, Morton (2012) developed an automated transit candidate validation procedure by calculating false positive probabilities, and Mullally et al. (2016) created a method of identifying false transit signals that were misclassified as small long-period planet candidates.

Here we investigate how a suspicious source in the KIC caused an erroneous crowding correction in the PDCSAP light curves of Kepler-445. As a result, the transit depths and planetary radii inferred from the Kepler-445c that were obtained with the Discovery Channel Telescope (DCT), which are described in Section 2. In Section 3, we describe the analysis of the DCT observations that produced high-precision transit light curves with scatter of 0.17% (0.0018 mags) in the best case. The repeatable high-precision, time-series photometry from the DCT makes it a valuable tool to be used by the exoplanet community. In Section 3, we also outline the Bayesian parameter estimation technique that yielded transit parameters that were in contention with the previous Kepler results. In Section 4, we postulate that an improbable series of events tracing back to the USNO-B1 stellar catalog led to the original miscalibration of Kepler-445c. Finally, in Section 5, we consider the physical radii of all three planets in the Kepler-445 system and the potential for additional phantom stars to affect other Kepler targets. This work offers a cautionary tale for the analysis of data from future exoplanet transit surveys including the Transiting Exoplanet Survey Satellite (TESS, Ricker et al. 2015).

2. OBSERVATIONS

We observed three transits of Kepler-445c with the Large Monolithic Imager (LMI; Massey et al. 2013) on the 4.3 m DCT located at Lowell Observatory in Happy Jack, Arizona. The observations are summarized in Table 1. Each observation took place under clear conditions and sub-arcsecond seeing. Two transits were observed through the LMI’s Sloan i’-band filter on UT 2015 June 19 and UT 2015 July 28 and one transit was observed through the Sloan z’-band filter on UT 2015 July 23. Each observation sought to obtain at least ~1.2 hr of baseline flux measurements bracketing the ~1.2 hr transit event. Transits were selected so that the entire observation (baseline and transit) occurred while Kepler-445 had an airmass below two.

We kept Kepler-445 in focus during the observations. To ensure that Kepler-445 stayed on the same group of pixels, we did not dither in between exposures. With the exposure times listed in Table 1, the central pixel of Kepler-445’s point-spread function typically registered ~3 × 10^5 counts, which was well below the nonlinear response regime of the detector. Figure 1 shows a portion of a background-subtracted, raw Sloan z’-band image centered on Kepler-445.

The LMI has an e2v CCD231 detector, which is a 6144 × 6160-pixel deep depletion CCD. The readout time for the entire chip with 1 × 1 binning on a single amplifier is ~73 s. Since our exposure times were always less than or equal to 30 s, we chose to increase efficiency by reading the chip with all four amplifiers and 2 × 2 binning. This decreased the overhead time in between exposures to ~8.5 s. The effective pixel scale was 0.24 pixel\(^{-1}\) with 2 × 2 binning.

3. DATA ANALYSIS

3.1. Calibration

We employed a custom data reduction and analysis pipeline that has previously been applied to high-precision, time-series photometry from the DCT-LMI (Dalba & Muirhead 2016). The calibration consisted of a bias and over scan subtraction and a flat-field correction. The minimal dark current present in the observations (estimated to be 0.07 electrons pixel\(^{-1}\) hr\(^{-1}\)) was removed through the background subtraction described below.

In some cases, we applied an additional correction for fringing. Although the LMI’s deep depletion CCD suppressed fringing at long wavelengths, the Sloan z’-band images still displayed low-level fringe patterns with typical variability of less than 1%. The patterns were large and were difficult to detect in small portions of the full image (Figure 1). We removed the fringing using a fringe frame constructed previously under a separate DCT program (P.I. E. Blanton). The fringe frame was created by median combining seventeen 600 s Sloan z’-band images taken on UT 2013 November 02 and UT 2013 November 04. Each image was

\textsuperscript{4} LMI transmission curves are available at http://www2.lowell.edu/rsch/LMI/specs.html.
dithered by at least 15″ to ensure that sources fell on different pixels and would not appear in the final fringe frame. Each image also received a bias, over scan, and flat-field correction before being combined to create the fringe frame. We scaled the fringe frame to each Kepler-445 science frame following the method of Snodgrass & Carry (2013) and using 45 “control pairs” to account for potential contamination by background sources. The scaled fringe frame was then subtracted from the science frame to remove the fringing.

### 3.2. Differential Aperture Photometry

We conducted differential aperture photometry on Kepler-445 using background stars in the 12′3 × 12′3 field of view of the DCT-LMI. The field was crowded, so we selected calibration stars based on the following criteria. First, the outer edge of sky annulus of the star could not be within 100 pixels of the edge of the frame. Second, the star’s sky annulus could not overlap the photometric aperture of any other source in the image. Third, the star’s count value could not enter the nonlinear response regime of the detector (≥4 × 10^4 counts) at any point throughout the observation. Fourth, the calibrator could not be a known variable star. Lastly, the star’s photometric aperture could not contain a cosmetic defect (e.g., dead pixels). These criteria yielded hundreds of potential calibration stars. We conducted photometry on each of these stars throughout the observations, though we did not use each star in the calculation of Kepler-445’s final light curve.

For each image, we first determined the centroids of Kepler-445 and all calibration stars by fitting a 2D Gaussian profile to the stellar point-spread functions. The x- and y-centroid drifts of Kepler-445 are shown in Figure 2. In most cases, the centroid stayed within a pixel or two of its original position.

We masked each bright source in the image to estimate the median global background value, which was subtracted from the entire image. We summed the flux from Kepler-445 and the calibration stars in the photometric apertures and used the sky annuli to account for any residual local background signal, including any dark current. In the sky annuli, bad pixels, hot pixels, cosmic-ray strikes, and any other spurious signals at a level of 5σ above or below the median count value were masked. We employed this two-stage background subtraction in order to monitor changes in local background across the quadrants of the detector that were read out by different amplifiers. At the time of the Kepler-445 observations, the multi-amplifier readout feature of the DCT-LMI had not been widely used. We did not measure local variations in background signal for any of the observations. Since the sky annuli were sufficiently large, the mean local background signals were approximately zero and the second subtraction had no effect on the resultant photometry.

Figure 1. Left: background-subtracted but otherwise raw Sloan z'-band image from the DCT-LMI on UT 2015 July 23 centered on Kepler-445. It is only a portion of the full 12′3 × 12′3 field of view. The minor fringes present in the Sloan z'-band images are difficult to detect on this small scale. The red circle is the 9.5-pixel photometric aperture and the blue circles of radii 20 and 30 pixels define the sky annulus. Right: United Kingdom Infrared Telescope (UKIRT) image centered on Kepler-445. Note the difference in scale compared to the DCT image. The red cross denotes the expected position of KIC 9730159 according to the Kepler Input Catalog. This star was not present in either image.

Figure 2. Relative drift in centroid positions of Kepler-445 during each observation. Panels (a)–(c) correspond to observations on UT 2015 June 19, UT 2015 July 23, and UT 2015 July 28, respectively. In each panel, the black curve represents the relative drift of the x-centroid position since the beginning of the observation, and the red curve represents the relative drift of the y-centroid. The centroid jump in panel (a) resulted from a brief loss of guiding during a reset of the DCT’s active optics system.
The Astronomical Journal, 153:59 (11pp), 2017 February

Dalba et al.

Table 2
Final Calibration Stars

<table>
<thead>
<tr>
<th>KIC ID</th>
<th>Sloan i'-band Magnitude</th>
<th>Observation Date (UT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9789729</td>
<td>14.541</td>
<td>2015 Jun 19</td>
</tr>
<tr>
<td>9790033</td>
<td>14.860</td>
<td>2015 Jun 19</td>
</tr>
<tr>
<td>9789938</td>
<td>14.913</td>
<td>2015 Jun 19</td>
</tr>
<tr>
<td>9789731</td>
<td>14.942</td>
<td>2015 Jun 19</td>
</tr>
<tr>
<td>9789958</td>
<td>14.718</td>
<td>2015 Jul 23</td>
</tr>
<tr>
<td>9790304</td>
<td>14.557</td>
<td>2015 Jul 23</td>
</tr>
<tr>
<td>9730187</td>
<td>15.493</td>
<td>2015 Jul 23</td>
</tr>
<tr>
<td>9789836</td>
<td>15.353</td>
<td>2015 Jul 28</td>
</tr>
<tr>
<td>9729892</td>
<td>15.733</td>
<td>2015 Jul 28</td>
</tr>
<tr>
<td>9729667</td>
<td>15.355</td>
<td>2015 Jul 28</td>
</tr>
<tr>
<td>9790345</td>
<td>15.244</td>
<td>2015 Jul 28</td>
</tr>
<tr>
<td>9729595</td>
<td>16.540</td>
<td>2015 Jul 28</td>
</tr>
</tbody>
</table>

Notes. The Sloan i'-band magnitude of Kepler-445 is 16.024 ± 0.011 (Muirhead et al. 2015).

* Kepler Input Catalog (KIC, Brown et al. 2011).

b Magnitudes from http://vizier.u-strasbg.fr/.

The radii of the photometric apertures and sky annuli are provided in Table 1. These apertures yielded the lowest out-of-transit scatter in the final light curves of Kepler-445. Other aperture radii in the range of [7, 14] pixels—incremented by 0.5 pixels—returned less precise photometry.

Of the many potential calibration stars in the DCT-LMI field of view, only a subset was used to create a master calibration light curve. The subset was determined through the following procedure. First, each star’s light curve was normalized to the median count value of the exposures gathered before or after the transit of Kepler-445c. Then, for each of the i stars, we defined a quality factor $Q_i$ as

$$Q_i = \frac{1}{N} \sum_{n=0}^{N} (F_{445,n} - F_{i,n})^2$$

where $N$ was the number of exposures taken outside of Kepler-445c’s transit, $F_{445,n}$ was the median-normalized flux of Kepler-445 in the $n$th exposure, and $F_{i,n}$ was the median-normalized flux of the $i$th star in the $n$th exposure. This factor described the out-of-transit deviation of each calibration star’s photometry from that of Kepler-445. We only continued analysis with the 15 calibration stars having the lowest $Q$-values. These 15 stars were distributed into 32,767 unique sets. We calculated a master calibration light curve by taking the mean of all the light curves in each set. We normalized the flux of Kepler-445 to each of these 32,767 master light curves, and only continued analysis on the one curve that yielded the lowest out-of-transit scatter. The calibration stars used to create final light curves of Kepler-445 for each observation are listed in Table 2. Each star was similar in brightness to Kepler-445, within 1.5 mag in Sloan i'-band.

The master calibration light curves and the photometry of Kepler-445 prior to calibration are shown in Figure 3. The flux from Kepler-445 and the calibration stars increased or decreased as the field rose or set. This variation was greatest for the first observation, which sampled the widest range of airmass. The small, rapid variations were due to minor changes in transparency and other noise sources. The transits were visible as deviations between the Kepler-445 and calibration light curves. The photometry did not appear to exhibit systematic errors that could influence the values of transit depth derived later.

The final transit light curves for UT 2015 June 19, UT 2015 July 23, and UT 2015 July 28 are shown in Figures 4–6, respectively. The three observations displayed near-Gaussian scatter with 1σ uncertainties of 0.26%, 0.23%, and 0.17%. The repeatability, high-precision photometry demonstrates that the DCT-LMI is a useful tool for transit photometry.

To date, only a few campaigns have utilized the DCT-LMI for high-precision transit photometry (e.g., Biddle et al. 2014; Dalba & Muirhead 2016), so any time-correlated (red) noise present in this type of observation has not been fully characterized. To evaluate the extent of red noise in the Kepler-445 photometry, we employed the “time-averaging”
method propounded by Pont et al. (2006) and used in numerous transit photometry studies (e.g., Winn et al. 2008). On the timescale of ingress (5–10 minutes), the binned residuals for each observation stayed very near to the expectation for Gaussian noise (Figure 7). This suggested that our results were not greatly influenced by time-correlated uncertainty.

We also searched for correlation between the final photometry and the $x$- and $y$-centroid positions of Kepler-445 on the detector (Figure 2). Once the background trend was removed from the photometry (Section 3.3), the Pearson correlation coefficients ($r$) between the out-of-transit flux and the centroids for each observation were small ($|r| \leq 0.04$), suggesting that there was no significant linear correlation.

### 3.3. Parameter Estimation

We fit the final light curves of Kepler-445 with transit models in order to determine the transit parameters for Kepler-445c. The purpose of this parameter estimation was not to fully recharacterize Kepler-445c or the Kepler-445 system, which we leave to future work (A. Mann et al. 2017, in preparation). Instead, we focused on the magnitude of Kepler-445c’s transit depth. As we demonstrate in Section 4, our ground-based observations differed significantly from those obtained with Kepler data as reported by Muirhead et al. (2015).

We employed the analytic transit models of Mandel & Agol (2002) in our fits to the DCT-LMI data. The transit parameters used by these models are described in Table 3. Following Kipping (2013), we included physically realistic quadratic limb darkening in the transit model with the parameters $q_1$ and $q_2$ as substitutes for $u_1$ and $u_2$ via the relations

$$q_1 = (u_1 + u_2)^2$$

$$q_2 = \frac{u_1}{2(u_1 + u_2)}.$$  \hspace{1cm} (2)

We also fit for the slight background trend that existed in each observation (see the top panels of Figures 4–6). Although it was possible that this minor effect was astrophysical in nature (i.e., stellar variability), it was more likely a result of the changing airmass during the observations (Table 1) and other systematic errors. Regardless of its origin, we included in the fit a background signal ($f_b$) of the form

$$f_b = b_0 + b_1 \ t$$  \hspace{1cm} (3)

where $t$ was the time elapsed since the beginning of the observation and $b_0$ and $b_1$ were free parameters.

Following Muirhead et al. (2015), we held the orbital eccentricity ($e$) at zero during the model fitting under the assumption that the timescale for Kepler-445c to circularize was significantly less than the age of the Kepler-445 system. We also adopted the orbital period of Kepler-445c from Muirhead et al. (2015) in our fitting procedure since this was well constrained by the Kepler light curves and not influenced by the post-processing aperture contamination we describe later.

We used the Markov Chain Monte Carlo (MCMC) ensemble sampler emcee (Foreman-Mackey et al. 2013) to estimate the
other transit parameters within a Bayesian framework and applied the uniform priors listed in Table 3. Since time-correlated noise was insignificant in these observations (Section 3.2), the uncertainties on each data point were defined to be the standard deviation of the out-of-transit flux. Furthermore, in evaluating the posterior distributions, the likelihood functions were assumed to be Gaussian. Each MCMC routine sampled the parameter space 2 × 10^6 times, and the chains converged on the best-fit value of each parameter within the first 30% of the total number of steps.

The best-fit parameters for each observation fitted individually are presented in Table 4. The scaled semimajor axis values (\(a/R_\star\)), which should not vary as a function of wavelength or time, agreed to within 1\(\sigma\) across the observations. The limb darkening parameters were not well constrained but marginализing over these uncertainties still yielded errors less than 4\% on the best-fit planet–star radius ratios (\(R_p/R_\star\)). The values of \(R_p/R_\star\) for the two Sloan i’-band observations (UT 2015 June 19 and UT 2015 July 28) were inconsistent to ∼2.1\(\sigma\). The limb darkening parameters found for each of these Sloan i’-band observations also displayed notable differences. This may have resulted from the reduction in post-transit baseline observation suffered by the first Sloan i’-band observation. This loss of baseline increased the uncertainty in the estimation of the background trend, potentially allowing the background to artificially influence the transit parameters. The slight difference in observing conditions between the two Sloan i’-band nights may have also exacerbated the issue. The second Sloan i’-band transit, having a full out-of-transit baseline and the highest precision of any observation, was likely the more reliable measurement, and the differences in transit parameters mentioned here were likely not astrophysical in nature.

We extracted the best-fit transit parameters a second time by fitting all three observations simultaneously. The model used in this case had two global parameters (\(a/R_\star\) and orbital inclination \(i\)), three filter-dependent parameters (\(R_p/R_\star\), \(u_1\), and \(u_2\)), and three other local parameters (mid-transit time \(t_0\), \(b_0\), and \(b_1\)). The best-fit values of all jointly fit parameters are listed in Table 5. As expected, the new value of \(a/R_\star\) was consistent with each individual observation. In this joint fit, the values of \(R_p/R_\star\) between Sloan i’ and Sloan z’ bands were consistent to within 1\(\sigma\). As expected, the Sloan i’-band parameters favored the values from the UT 2015 July 28 observation, which were likely more reliable than those from UT 2015 June 19.

4. RESULTS

4.1. The Revised Transit Depth of Kepler-445c

We estimated the transit depths of Kepler-445c in the Sloan i’ and z’ bands using the best-fit parameters in Table 5. For each filter, we created 1000 theoretical light curves by drawing samples from the posterior distributions of the transit parameters. From these light curves, we determined the posterior distribution of transit depths. In the Sloan i’ and z’ bands, these depths were 0.618 ± 0.022% and 0.545 ± 0.028%, respectively. These values were consistent with each other to 2.1\(\sigma\) confidence.

It was unlikely, although not impossible, that the difference between the transit depths was a result of Kepler-445c’s atmospheric properties. Most exoplanets similar to Kepler-445c (e.g., GJ 1214b, HD 97658b, GJ 436b, Knutson et al. 2014a, 2014b; Kreidberg et al. 2014) have been found to host atmospheres with high-altitude clouds, hazes, or high mean molecular weights that produce flat transmission spectra. If this was true of Kepler-445c’s atmosphere as well, one would expect the Sloan i’- and z’-band transit depths to be

Table 3
Prior Probability Density Functions Employed in All MCMC Fits

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or Prior</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e)</td>
<td>0</td>
<td>Orbital eccentricity. Fixed, from Muirhead et al. (2015)</td>
</tr>
<tr>
<td>(P) [days]</td>
<td>4.871229 ± 0.000011</td>
<td>Orbital period. Fixed, from Muirhead et al. (2015)</td>
</tr>
<tr>
<td>(a/R_\star)</td>
<td>(U(20.41, 34.01))</td>
<td>Semimajor axis scaled to stellar radius</td>
</tr>
<tr>
<td>(i) [degrees]</td>
<td>(U(88.9, 90))</td>
<td>Orbital inclination</td>
</tr>
<tr>
<td>(t_0,1) = 2457192 [JD]</td>
<td>(U(0.75, 0.79))</td>
<td>Time of mid-transit on UT 2015 June 19</td>
</tr>
<tr>
<td>(t_0,2) = 2457226 [JD]</td>
<td>(U(0.85, 0.90))</td>
<td>Time of mid-transit on UT 2015 July 23</td>
</tr>
<tr>
<td>(t_0,3) = 2457231 [JD]</td>
<td>(U(0.73, 0.79))</td>
<td>Time of mid-transit on UT 2015 July 28</td>
</tr>
<tr>
<td>(q_1)</td>
<td>(U(0, 1))</td>
<td>Limb darkening parameter 1 (Kipping 2013)</td>
</tr>
<tr>
<td>(q_2)</td>
<td>(U(0, 1))</td>
<td>Limb darkening parameter 2 (Kipping 2013)</td>
</tr>
<tr>
<td>(R_p/R_\star)</td>
<td>(U(0.001, 0.001))</td>
<td>Planet–star radius ratio</td>
</tr>
<tr>
<td>(b_0)</td>
<td>(U(-0.01, 0.01))</td>
<td>0(^{th}) order background parameter (Equation 3)</td>
</tr>
<tr>
<td>(b_1) [days(^{-1})]</td>
<td>(U(-0.05, 0.05))</td>
<td>1(^{st}) order background parameter (Equation 3)</td>
</tr>
</tbody>
</table>

Note. \(U[a, b]\) signifies a uniform prior probability distribution in the range \([a, b]\).

Table 4
Best-fit Transit Parameters for Individual Observations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a/R_\star)</td>
<td>26.0±1.0</td>
<td>26.2±1.1</td>
<td>25.7±1.1</td>
</tr>
<tr>
<td>(i) [degrees]</td>
<td>89.50±0.34</td>
<td>89.49±0.35</td>
<td>89.51±0.34</td>
</tr>
<tr>
<td>(t_0)</td>
<td>2457683±0.00063</td>
<td>126.8854±0.00063</td>
<td>131.7556±0.00054</td>
</tr>
<tr>
<td>(q_1)</td>
<td>0.55±0.01</td>
<td>0.43±0.03</td>
<td>0.66±0.23</td>
</tr>
<tr>
<td>(q_2)</td>
<td>0.46±0.31</td>
<td>0.50±0.30</td>
<td>0.73±0.17</td>
</tr>
<tr>
<td>(u_1)</td>
<td>0.46±0.34</td>
<td>0.63±0.31</td>
<td>1.13±0.26</td>
</tr>
<tr>
<td>(u_2)</td>
<td>0.06±0.42</td>
<td>−0.01±0.31</td>
<td>−0.37±0.31</td>
</tr>
<tr>
<td>(R_p/R_\star)</td>
<td>0.0727±0.0026</td>
<td>0.0662±0.0021</td>
<td>0.0651±0.0022</td>
</tr>
<tr>
<td>(b_0)</td>
<td>0.00139±0.00031</td>
<td>0.00168±0.00021</td>
<td>0.00047±0.00016</td>
</tr>
<tr>
<td>(b_1) [days(^{-1})]</td>
<td>−0.0045±0.00036</td>
<td>−0.0280±0.00024</td>
<td>−0.0004±0.00017</td>
</tr>
</tbody>
</table>

Note. The upper and lower uncertainties of each parameter were calculated using the 16th, 50th, and 84th percentiles of the posterior distributions. See Table 3 for descriptions of the parameters and priors.
consistent. The difference, then, could stem from a slight underestimation of the uncertainties on the transit parameters such that the posterior distribution of transit depths was artificially narrow. Regardless of the origin of the inconsistency, the transit depths estimated from the three DCT-LMI observations described here were likely not precise enough to distinguish between various atmospheric compositions for Kepler-445c.

Continuing under the assumption that the variation in transit depths between filters was not astrophysical in nature, we calculated a combined value of transit depth by taking the mean of the two values listed above: \(0.582 \pm 0.026\%\). The uncertainty was the standard error of the mean defined as \(S/\sqrt{n}\), where \(S\) was the sample standard deviation and \(n\) was the number of samples, which in this case was two.

### 4.2. The Previous Transit Depth of Kepler-445c

The Kepler-445 system was first characterized with long-cadence Kepler PDCSAP data by Muirhead et al. (2015). The \(a/R_\ast\) and \(R_p/R_\ast\) values for Kepler-445c reported by that work are \(30.21 \pm 0.38\) and \(0.1075 \pm 0.0014\), which are also the values currently listed on the NASA Exoplanet Archive. These values disagree with the \(a/R_\ast\) and \(R_p/R_\ast\) estimated from the DCT observations to \(3.2\sigma\) and greater than \(17\sigma\), respectively.

Additionally, visual inspection of the phase-folded light curves of Muirhead et al. (2015, Figure 5) suggests that the transit depth is at least \(1\%\). Muirhead et al. (2015) does not explicitly report a transit depth, though exoplanets.org currently cites that work for a depth of \(1.16 \pm 0.28\%\), more than twice the value reported here. Differences in limb darkening can alter transit depths obtained through different filters (e.g., Knutson et al. 2007); however, this discrepancy seems too large to be attributed entirely to limb darkening.

Since the Kepler-445 system was first characterized, new observations of Kepler-445 may result in updated stellar parameters (A. Mann et al. 2017, in preparation). These new parameters may alter the physical characteristics of Kepler-445c, but would not explain the differences in transit depth between the Kepler and DCT-LMI transit curves.

### 4.3. The Crowding Metric in the Kepler Pipeline

The discrepancy between the transit depths of Kepler-445c is a result of the crowding metric applied by the pre-search data conditioning (PDC) module of the Kepler pipeline (Smith et al. 2012; Stumpe et al. 2012). This metric, defined as “CROWDSAP” in the header of the Kepler light curve files, is the ratio of the target flux to the total flux in the photometric aperture. It is a quarterly averaged, scalar value between zero and unity. Only the PDCSAP version of the flux reported by Kepler is adjusted by this crowding factor.

The value of CROWDSAP for a given star may vary from quarter to quarter as it falls on different regions of the detector and is processed with different pixel response functions (PRFs) and optimal photometric apertures. The value of CROWDSAP reported in the Kepler data also depends on the version of the Science Operations Center (SOC) pipeline that processed the data. In Figure 8, we show the CROWDSAP values for Kepler-445 data processed with versions 9.0 and 9.3 of the SOC pipeline. Between version 9.0, which processed the data used by Muirhead et al. (2015), and version 9.3, which was released in 2015 December, the mean value of CROWDSAP changed from \(0.47 \pm 0.04\) to \(0.91 \pm 0.06\), a \(6.1\sigma\) increase.

CROWDSAP values can be directly affected by changes to the optimal photometric apertures, which occasionally change when the Kepler pipeline is updated. After an update, the pipeline may determine that a different photometric aperture increases the combined differential photometric precision (CDPP, Christiansen et al. 2012) in the resultant light curve of a particular star. If this new aperture now contains flux from nearby sources, then the CROWDSAP parameter should change accordingly.

As demonstrated by Figure 9, the optimal photometric apertures for Kepler-445 were largely unchanged between versions 9.0 and 9.3 of the pipeline. With the exception of minor changes in Quarters 12 and 16, the apertures covered the same groups of pixels and yet the CROWDSAP values

### Table 5
Best-fit Transit Parameters for Joint Fit between All Observations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td></td>
</tr>
<tr>
<td>(a/R_\ast)</td>
<td>26.211-0.8</td>
</tr>
<tr>
<td>(i) [degrees]</td>
<td>89.54-0.32</td>
</tr>
<tr>
<td>Sloan (i)</td>
<td></td>
</tr>
<tr>
<td>(R_p/R_\ast)</td>
<td>0.0675-0.0018</td>
</tr>
<tr>
<td>(q_1)</td>
<td>0.55-0.25</td>
</tr>
<tr>
<td>(q_2)</td>
<td>0.68-0.20</td>
</tr>
<tr>
<td>(u_1)</td>
<td>0.99-0.20</td>
</tr>
<tr>
<td>(u_2)</td>
<td>-0.27-0.25</td>
</tr>
<tr>
<td>Sloan (z)</td>
<td></td>
</tr>
<tr>
<td>(R_p/R_\ast)</td>
<td>0.0657-0.0019</td>
</tr>
<tr>
<td>(q_1)</td>
<td>0.50-0.29</td>
</tr>
<tr>
<td>(q_2)</td>
<td>0.48-0.25</td>
</tr>
<tr>
<td>(u_1)</td>
<td>0.65-0.30</td>
</tr>
<tr>
<td>(u_2)</td>
<td>0.03-0.35</td>
</tr>
</tbody>
</table>

**Note.** The upper and lower uncertainties of each fitted parameter were calculated using the 16th, 50th, and 84th percentiles of the posterior distributions. See Table 3 for descriptions of the parameters and priors.
increased dramatically. From an increase in the average value of the crowding metric, we would expect a corresponding decrease in transit depth. Inspection of the Kepler-445c transit curves produced by version 9.3 of the SOC pipeline confirmed this expectation. In Figure 10, we show the phase-folded PDCSAP photometry of Kepler-445 from all quarters of observation. The photometry was flattened using the keplflatten task in PyKE\(^9\) (Still & Barclay 2012) and phase-folded according to the ephemeris in Muirhead et al. (2015) for Kepler-445c. Visual inspection of the phase-folded photometry suggested a transit depth of $\sim$0.6%, which was consistent with the DCT-LMI light curves.

The ratio between the DCT-LMI transit depth of Kepler-445c (0.582 $\pm$ 0.026\%) and the previously published transit depth\(^9\) (1.16 $\pm$ 0.28\%) equals 0.50 $\pm$ 0.12. This value agrees with the ratio of mean CROWDSAP values between versions 9.0 and 9.3 of the SOC pipeline, 0.52 $\pm$ 0.06, to well within 1\(\sigma\). Therefore, the change in the crowding metric offers a highly consistent explanation for the discrepant Kepler-445c transit depths discovered by this work.

The 6.1\(\sigma\) increase in mean CROWDSAP value for Kepler-445 between the two aforementioned versions of the SOC pipeline results from a new method of constructing the stellar “scene,” which contains all stars besides the target and any zodiacal light (Bryson et al. 2010). Before version 9.3 of the SOC pipeline, the scene was determined with the predicted PRF model and the stellar data in the KIC. In version 9.3, however, stellar scenes are reconstructed through a photometric analysis of the actual pixel-level data (Twicken et al. 2016). This new method maximizes the CDPP in the photometric aperture and removes an inherent susceptibility to stellar position and magnitude errors that may be present in the KIC. The optimal photometric apertures may change in response to the new stellar scenes and influence the crowding metrics. However, in the case of Kepler-445, the CROWDSAP variations were not caused by changes in photometric apertures.

### 4.4. A Phantom Star in the Scene of Kepler-445

The SOC pipeline update and subsequent change in CROWDSAP values identified an unusual circumstance surrounding the characterization of exoplanet Kepler-445c. According to the KIC, another source known as KIC 9730159 existed 2\(^{\circ}2\) away from Kepler-445. The object class of KIC 9730159 was reported as “star” and the condition flag was empty (i.e., KIC 9730159 was not flagged as an artifact, planetary-candidate, or exoplanet). KIC 9730159 had a Kepler magnitude of 17.667, which was 1.1\% greater than that of Kepler-445. Given the plate scale of our DCT-LMI observations (0\(^{\prime}24\) pixel\(^{-1}\)), we would have expected KIC 9730159 to be present in the images $\sim$9 pixels away from Kepler-445, well outside the full width at half maximum of the point-spread function even during the worst seeing we experienced. However, KIC 9730159 was not present in the DCT-LMI images (Figure 1). KIC 9730159 was also not visible in the J-band United Kingdom Infrared Telescope (UKIRT)\(^8\) images of the Kepler field (Figure 1). Muirhead et al. (2015) reported a non-detection of KIC 9730159, postulating that the KIC may have accidentally registered Kepler-445 twice.

We explored the hypothesis that KIC 9730159 was a phantom star—present in the KIC and the data conditioning, but not an actual star—in more detail. KIC 9730159 originated in the USNO-B catalog, where its designation was 1364-0336533 and its mean epoch of observation was 1990.2. It had a high proper motion (R.A.: $\pm$146 mas, decl.:250 mas) and a correspondingly large uncertainty in its J2000 coordinates. Monet et al. (2003) warned that objects with large proper motions may be erroneously assigned in future observations and may lead to catalog entries for blank patches of sky. This is a highly probable explanation for the existence of KIC 9730159. Kepler-445 (USNO-B1 designation 1364-0336538) also had a relatively high proper motion, but its mean epoch of observation was 1977.1. KIC 9730159 was likely “created” as a result of an observation of Kepler-445 in the UKIRT 1990 that was mistakenly assigned to a new star. Furthermore, the USNO-B1 catalog reported three detections of this phantom star, allowing it to pass the threshold for acceptance into the KIC (Brown et al. 2011).

Together, the USNO-B phantom star inherited by the KIC, the original method used to construct the stellar scene of Kepler-445, and the timing of the initial characterization of Kepler-445c formed an improbable series of events that led to the overestimation of the transit depth of Kepler-445c.

### 5. DISCUSSION

#### 5.1. Are All the Planets in the Kepler-445 System Rocky?

The original radius of Kepler-445c, 2.51 $\pm$ 0.36\(R_\oplus\) (Muirhead et al. 2015), placed it in the interesting regime of exoplanets with sizes between the Earth and Neptune that are highly amenable to atmospheric characterization (e.g., Fraine et al. 2014; Knutson et al. 2014a, 2014b; Kreidberg et al. 2014). With ground-based observations of additional transits of this exoplanet, we found that the transit depth and planetary radius were overestimated. Using the value of $R_p$/$R_\star$, obtained

---

8 https://keplerscience.arc.nasa.gov/software.html

9 See exoplanets.org, which cites Muirhead et al. (2015).

10 From database release “ukidssdr10plus,” http://wsa.roe.ac.uk/.
from the joint fit to all the DCT Sloan i’-band observations\textsuperscript{11} and employing the Kepler-445 stellar radius from Muirhead et al.\textsuperscript{2015}, we find the revised radius of Kepler-445c to be \(1.55 \pm 0.23R_{\oplus}\).

\textsuperscript{11} We choose the \(R_p/R_*\) value from Sloan i’ band over Sloan z’ band based on the parameter uncertainties.

In addition to Kepler-445c, this system contains two smaller exoplanets: Kepler-445b and Kepler-445d. As reported in Muirhead et al.\textsuperscript{2015}, the b and d planets have radii of \(1.58 \pm 0.23R_{\oplus}\) and \(1.25 \pm 0.19R_{\oplus}\) respectively. Although we did not acquire transit observations of these exoplanets, their Kepler light curves were subject to the same analysis and crowding contamination as that of Kepler-445c. Here, we scale their Kepler transit depths using the ratio of CROWDSAP values before and after the SOC pipeline update (Section 4.3) and briefly discuss the potential nature of these small exoplanets.

\textsuperscript{12} See exoplanets.org, which cites Muirhead et al.\textsuperscript{2015}.
the updates to the SOC pipeline have increased the accuracy of the CROWDSAP values for all KIC sources. These updates corrected the transit depth of Kepler-445c and greatly reduced the probability of errors in the KIC contaminating other Kepler PDCSAP light curves.

However, it is possible that other phantom stars exist in the KIC and influence the photometry of other Kepler targets. The authors contributed a discussion on this topic to the 2016 August 8 version of the Kepler Data Release Notes 25 on MAST. Conducting a direct search for phantom stars is a challenging task given the number of sources comprising the KIC. Indirect searches (e.g., by flux or transit depth changes between different pipeline versions) are also complicated since large CROWDSAP variations can correctly accompany changes in the optimal photometric apertures. Furthermore, different versions of the pipeline typically operate with different amounts of data. The sensitivity added by additional transits must also be considered when searching for phantom stars indirectly.

Phantom stars are likely a rare occurrence. However, their presence could cause systematic shifts in stellar crowding between pipeline versions and could alter transit depths and inferred planetary radii. This has the potential to affect investigations that make use of the PDCSAP transit depths for a large number of exoplanets, where an individual follow-up observation of each is not possible (e.g., studies of planet occurrence rates, planet populations, or transit signal recovery efficiencies).

The sizes of exoplanets in the perplexing Earth-to-Neptune regime only span a few Earth radii. A factor of two correction to the transit depth, corresponding to a $\sim \sqrt{2}$ change in planetary radius, can therefore crucially alter the interpretation of a planet’s interior and atmosphere. In this way, assessing the crowding in the scene of an exoplanet host—which is a necessary step in the data conditioning process—has the potential to obscure the understanding of these exoplanets further. Although the case of Kepler-445c was highly improbable, it emphasizes the level of caution that must be exercised when analyzing and interpreting data from exoplanet transit surveys.

### 5.3. Implications for TESS

In the near future, TESS is expected to launch and begin surveying the entire sky for nearby exoplanets (Ricker et al. 2015). Whereas the pixels on the Kepler detectors have a plate scale of $\sim 4^\prime$ pixel$^{-1}$, TESS’ pixels will cover even more sky, with $\sim 21^\prime$ pixel$^{-1}$. TESS will therefore be even more susceptible to crowding and stellar catalog errors than Kepler. The analysis of TESS observations will of course be informed by the many lessons learned from the Kepler data set, and a careful treatment of the crowding metric should be no exception.

The authors thank the referee, Jessie Christiansen, for a thoughtful review that improved the quality of this work. The authors acknowledge Jon Jenkins for a helpful discussion of the Kepler Science Operations Center Pipeline; Andrew Mann, Trent Dupuy, and Marshall Johnson for information regarding the Kepler-445 system; and Andrew Vanderburg for providing a portion of the data used in this paper. The authors also thank Brian Taylor, Stephen Levine, Alan Marscher, Elizabeth Bland, and the DCT telescope operators for their assistance with the observations.

These results made use of the Discovery Channel Telescope at Lowell Observatory. Lowell is a private, non-profit institution dedicated to astrophysical research and public appreciation of astronomy and operates the DCT in partnership with Boston University, the University of Maryland, the University of Toledo, Northern Arizona University, and Yale University. LMI construction was supported by a grant AST-1005313 from the National Science Foundation.

This paper includes data collected by the Kepler Mission. Funding for the Kepler Mission is provided by the NASA Science Mission Directorate. The Kepler data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NNX13AC07G and by other grants and contracts.

This research made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

This paper includes an image from the United Kingdom Infrared Telescope (UKIRT) Project, which is operated by the Joint Astronomy Centre on behalf of the Science and Technology Facilities Council of the United Kingdom.

This work made use of PyKE, a software package for the reduction and analysis of Kepler data. This open source software project is developed and distributed by the NASA Kepler Guest Observer Office.

This research made use of analytic transit models based on the formalism from Mandel & Agol (2002) and adapted for Python by Laura Kreidberg.

**Facilities:** Discovery Channel Telescope, (Large Monolithic Imager), Kepler.

### REFERENCES


The Astronomical Journal, 153:59 (11pp), 2017 February

Massey, P., Dunham, E. W., Bida, T. A., et al. 2013, in AAS Meeting 221 Abstracts, #345.02
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, PNAS, 110, 19273
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, JATIS, 1, 014003

Snodgrass, C., & Carry, B. 2013, Mngr, 152, 14