

Cluster Difference Imaging Photometric Survey. Vetting Report Description Document

L. G. BOUMA,^{1,*} J. D. HARTMAN,² W. BHATTI,² J. N. WINN,² AND G. Á. BAKOS²

¹Cahill Center for Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

²Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA

ABSTRACT

This document describes the CDIPS planet candidate vetting reports and planet parameter summaries uploaded by bouma to ExoFOP-TESS 2022-02-19. It also has a section describing how to parse the comment strings. These upload represent new knowledge for 42 unique objects. 35 of these objects are new CTOIs, 15 of which are labelled with the subjective "A-tier" quality flag. The general point of uploading these parameters and reports is to encourage and enable follow-up efforts geared at finding age-dated and age-dateable planets. Earlier versions of this document are archived at lgbouma.com/notes. The light curve reduction approach is discussed in Bouma et al. (2019).

1. THE EXOFOP COMMENT STRINGS

There is a 119 character limit on the ExoFOP-TESS comment strings. This is good, since it encourages brevity. It also leads to cryptic acronyms. For the Year 2 upload, here are a few examples of how to parse these:

Qual:A,CM,1.8e+08yr. Rp=(1.75-6.07)Rj. b=0.84-1.40 (3-97HDI) .

This is an "A-tier" quality object (versus the less promising "B-tier"). It is thought to probably be a Cluster Member (CM), with an average age from the literature sources of 180 million years (1.8×10^8). Based on the MCMC fitting, the highest density interval spanning the 3rd to 97th percentiles span an impact parameter of 0.84 to 1.40, and corresponding planet sizes of 1.75 to 6.07 Jupiter radii. The latter are of course not planetary. However the interpretation is that from the single sector of 30-minute cadence TESS FFI data that were fitted, while the object could be grazing, it might not: it suffers from the "size - impact parameter degeneracy".

Qual:A,PMS?,2.1e+07yr. Rp=(0.79-5.00)Rj. b=0.43-1.72 (3-97HDI) .

This is an "A-tier" quality object (versus the less promising "B-tier"). It is thought to possibly be on the pre-main-sequence (PMS?), with an average age from the literature sources of 21 million years. Based on the MCMC fitting, the highest density interval spanning the 3rd to 97th percentiles span an impact parameter of 0.43 to 1.72, and corresponding planet sizes of 0.79 to 5.00 Jupiter radii. This means that from the single sector of 30-minute cadence TESS FFI data that were fitted, the object could be grazing, and it suffers from the "size - impact parameter degeneracy".

Qual:B,NotCM?,1.4e+08yr,MissingRot?. Rp=(0.52-1.45)Rj. b=0.00-0.69 (3-97HDI) .

This is an "B-tier" quality object (versus the more promising "A-tier"). It is thought to possibly be a cluster member, though this is somewhat uncertain due to something noticed in the vetting reports (NotCM?). The average age from the literature sources if it were a cluster member is 140 million years. "MissingRot?" means that upon visual inspect of the light curve, the rotation period was not obviously what it should be, given the presumed effective temperature and age of the star if it were a cluster member – this flag assumes that gyrochronology works (see e.g., Curtis et al. 2019). The usual MCMC fitting interpretation applies.

Qual:B,NotCM?,4.8e+08yr,OffTarget?. Wrn: mean (R)=1.054. Rp=(0.94-2.95)Rj.
b=0.03-0.90 (3-97HDI) .

Corresponding author: L. G. Bouma
luke@astro.caltech.edu

* 51 Pegasi b Fellow

Two additional flags are included in this string. "OffTarget?" means that based upon either a) visual inspection of the difference images in the vetting report, or b) a tentative depth versus aperture size correlation, there is concern based on the TESS data that the transiting object might be on a neighboring star. The second flag "Wrn: mean(R)=1.054" comes from the MCMC. During the MCMC fitting, an attempt was made to produce a converged solution. The string "mean(R)" refers to the mean value of the Gelman-Rubin R-hat statistic across all sampled parameters (see the final MCMC page of the reports). If not all of the parameters had R-hat < 1.01 after the maximal run-time of the chains, this warning is produced – at least one, and possibly more, of the parameters has not converged. If scheduling transit follow-up for such cases, please check the MCMC table in the vetting report to ensure that the ephemeris corresponds to converged values of the epoch and period.

2. VETTING REPORT DESCRIPTION

The NASA team and MIT teams (Jenkins et al. 2010; Huang et al. 2018) produce vetting reports to assess the quality of planet candidates identified through their transiting planet search pipelines.

One goal of the CDIPS project is to detect transiting planets with known ages. Therefore our vetting reports include information to help assess (a) whether the transiting planet candidate is real, and (b) whether the reported age is correct. The code used to make these reports is available online¹.

Figures 1 to 7 summarize the document construed for these purposes. The planet candidate chosen for these figures is Gaia-DR2 1935439776865267072 = TIC 154353849.

2.1. Transit search summary

Figure 1. Periodograms from TLS and phase-dispersion minimization, calculated with `astrobase.periodbase`, are shown in the top left and top center (Bhatti et al. 2018; Hippke & Heller 2019; Stellingwerf 1978). The top three peaks from each method are shown in the second and third rows; the raw light-curve is in the top-right. A finder chart is inset to the top left, with the 1.5-pixel radius aperture used to extract the light-curve in orange. The finder charts are from The Second Digitized Sky Survey (Red); they are pulled from NASA's `skyview` service². The light curve being used for this page is the “detrended PCA” light curve, shown on the following page.

2.2. Light-curve diagnostics

Figure 2. Time-series of raw flux in a $r = 1\text{ px}$ aperture (IRM1), PCA-detrended flux (PCA1), stellar-variability “detrended flux”, and the background are shown as a function of barycentric Julian date.

The overplotted dashed vertical lines are the ephemeris of the highest-power TLS peak from Figure 1. An important visual check is whether the flux dips are correlated with changes in the background level – in this case, they are not. The standard deviation and TESS magnitude are quoted in the upper right. The red line in the top plot is a spline fit, which in this case helped in finding the eclipse signal.

The second panel shows a model (red dots) used to subtract the stellar variability. While during the first year of TESS reductions we adopted the penalized splines with iterative outlier removal implemented in `wotan` (Hippke et al. 2019), for the second year’s analysis we shifted methodologies. Based on tested recovery rates from injected synthetic transit signals, we shifted to a different local windowed slider. After an iterative windowed sigma-clipping is applied to each light curve, the decision about what approach to use to remove the stellar variability is made based on the highest-power period returned by a Lomb-Scargle periodogram (which in almost all of our cases of interest is the rotation period). The `notch` algorithm is used in cases where the rotation period exceeds one day, and the `locor` algorithm is adopted in the converse case. These were developed by Rizzuto et al. (2017), and our specific implementation is a fork from github.com/arizzuto/Notch_and_LOCoR.

2.3. Transit diagnostics

Figure 3. The plots show the maximally-detrended light-curve (top); the phase-folded light-curve centered over ± 3 transit durations of the primary transit (middle left); the secondary eclipse (middle right); the odd-numbered transits (lower left); and the even-numbered transits (lower right). Also shown is the best-fit TLS template model — by default, this assumes a non-grazing geometry (see Hippke & Heller 2019).

The stellar parameters (T_{eff}, R_*, M_*) are taken from TICv8 when available (Stassun et al. 2019). For the stellar radius, if TICv8 does not quote a radius, neither will the vetting report. If TICv8 gives a stellar radius, but does not give a stellar mass, we

¹ <https://github.com/lgbouma/cdips/tree/8e7f8f2a6aa3e176b2227e7d601d8029a58449f5>

² See <https://skyview.gsfc.nasa.gov/current/cgi/survey.pl> and <http://archive.stsci.edu/dss/acknowledging.html>

interpolate the mass using the radius and the [Pecaut & Mamajek \(2013\)](#) table, under the assumption that the star is a dwarf. The effective temperatures calculated by TICv8 were estimated from an empirical color relation involving Gaia G_{Bp} and G_{Rp} magnitudes. These observed colors were dereddened first, using Pan-STARRS dust maps ([Green et al. 2018](#)) for declinations above -30° , and the [Schlegel et al. \(1998\)](#) maps otherwise.

The first eight lines of text are parameters determined from the best-fitting TLS model. The one exception is the planet radius, which uses the stellar radius as noted above. The “flux contamination” (TICCONT) from neighboring stars is *never* taken into account, because transit depth dilution does not affect image subtraction analyses in the same manner as aperture-photometry reductions. The significance of the odd-to-even asymmetry is quoted, but given the strong rotational variability in this object (Figure 2), the apparent odd-even asymmetry could have been caused by the detrending process. To estimate the transit to occultation depth ratio $\delta_{\text{tra}}/\delta_{\text{occ}}$, the phase-folded light-curve is also fit by a sum of two gaussians (in this case, the fit failed). “AstExc” refers to the Gaia-DR2 astrometric excess, which can indicate hints of astrometric binarity in the system. While RUWE would be a better parameter to use, that is left for future implementations. “ d_{geom} ” is the geometric distance. “ $R_* + M_* \rightarrow T_{b0}$ ” gives the duration of a zero-eccentricity central transit based on the TICv8 stellar radius and masses discussed above.

2.4. Light-curves for increasing aperture sizes

Figure 4. Apertures of radius 1, 1.5, and 2.25 pixels are shown in maximally detrended light-curves from top to bottom. The blue line is the reference transit depth from the best-fitting TLS model. Changes in depth with increasing aperture size can indicate that the source of variability is off-center from the aperture, suggesting a photometric blend.

2.5. Imaging variability diagnostics

Figure 5. This page helps diagnose which stars are producing the observed variability. Top-left and top-center are the mean out-of-transit (OOT) and mean in-transit calibrated images (created separately from our image-subtraction analysis, using [TESScut](#), [Brasseur et al. 2019](#)). The OOT images are based on the same number of exposures as the in-transit images and split evenly before and after each transit (following [Bryson et al. 2013](#); [Kostov et al. 2019](#)). The yellow star is the target’s position from TICv8; small red crosses are WCS-projected locations of neighbor stars.

Middle-left is the most important sub-panel: the difference between the OOT and in-transit mean images. If the variability shown in the background map (units: ADU) is off-target, the signal is not from the target star. A two dimensional gaussian is fit to the inner 8x8 pixels to estimate the centroid position — the resulting best-fit location is shown as a white star, and the separation to the catalog star’s position (yellow star) is labelled `ctlg - gauss (OOT-intra)` on the right. A different approach, simply taking the first moment of the middle-left image, gives the `ctlg - <OOT-intra>` line.

Middle-center is the same as middle-left, but normalized by the uncertainty map. Lower left and lower center show the DSS2-Red field in linear and log scales at roughly the same pixel scale as the TESS image, with the 1, 1.5, and 2.25 pixel-radius apertures in blue, orange, and green respectively. The brightness of neighborhood stars is given on the far right. Note the slight coordinate rotation difference between DSS and TESS images; DSS images are aligned north-up, east-left; TESS images are oriented as closely as possible to this system without actually performing the rotation.

2.6. Cluster membership assessment diagnostics

Figure 6. The star was considered a candidate cluster member by the source(s) listed under “Refs”, in this case [Kounkel & Covey \(2019\)](#) and [Kounkel et al. \(2020\)](#). The latter is taken to be the “best available knowledge of the cluster” – the black points, in the remaining panels. These stars are used to define a local neighborhood of stars. More specifically, the “neighborhood” is defined as a group of at most 10^4 randomly selected stars within:

$$\langle \alpha \rangle \pm 5\sigma_\alpha, \quad (1)$$

$$\langle \delta \rangle \pm 5\sigma_\delta, \quad (2)$$

$$\langle \pi \rangle \pm 5\sigma_\pi, \quad (3)$$

where (α, δ, π) are the right ascension, declination, and parallax. $\langle x \rangle$ denotes the mean over all stars within the claimed cluster, σ_x denotes the standard deviation. The limiting G magnitude for the “neighborhood” is usually set around $G = 17$.

The cluster members (black points) and the neighborhood members (gray points) are *mutually exclusive sets*. After the stars are randomly selected from the volume described above, any overlapping stars between the group and the neighborhood samples are removed from the neighborhood sample.

In Figure 6, the top three subplots intentionally omit the labelled cluster members, in order to give the user their own by-eye assessment of whether they see clusters in the neighborhood (as might happen if the membership labelling is *incomplete*), and

whether the target star is within those clusters. The middle three subplots overplot cluster members from the reference noted in the legend. The radial velocity plots (bottom row) use radial velocities from Gaia-DR2, and so are limited to roughly 3500-7000 K, for $G \lesssim 12.5$ (Katz et al. 2019).

In this particular case, the cluster of interest (“Lac_OB1”, also known as “UBC165”), appears to be reasonably kinematically coherent (top two rows). The color versus absolute-magnitude diagram (third row, middle) shows that the stars fall on a relatively clean isochronal locus, which supports the interpretation that they are the same age. This age, which is one of the columns from the appendices of Bouma et al. 2022, and also mentioned as a comment on ExoFOP, is around 35 million years old. The fact that the lowest mass reported cluster members (black points) seem to be elevated with respect to the field supports the idea that they are on the pre-main-sequence.

2.7. MCMC transit fitting assessment

Figure 7 is a new addition to the vetting reports from the S14-S26 (TESS Year 2) processing effort. The general purpose of these plots is to enable assessment of the MCMC fitting that was used to derive the reported planetary parameters. It is clear in Figure 3 for example that the detrending approach used to *find* the planetary transits has the consequence of *distorting* the transit signal. In order to produce robust planetary parameters, particularly ephemerides, a different approach therefore needed to be adopted for MCMC fitting.

Specifically, using the known initial estimate of the ephemeris, the available transit windows were isolated to within a few transit durations of the transit event (*top row*). The light curves within these windows were then fitted as second-order polynomials, while masking the points in-transit. These polynomials were subtracted from the light curve, and the resulting flattened light curve was fitted for a model with ten free parameters: the planetary orbital period P , the mid-transit epoch t_0 , the natural logarithm of the planet-to-star size ratio $\log R_p/R_*$, the impact parameter b , two quadratic limb-darkening coefficients u_1 and u_2 , the stellar radius R_* , the stellar surface gravity $\log g$, the mean light curve flux level $\langle f \rangle$, and a “jitter” σ_f , added in quadrature to the reported uncertainties in the flux level.

We fitted the models using PyMC3 (Salvatier et al. 2016; Theano Development Team 2016) and exoplanet (Foreman-Mackey et al. 2020), and accounted for the finite integration time of each exposure in the numerical integration when evaluating the model light curve. We assumed a Gaussian likelihood, and after initializing each model with the parameters of the maximum *a posteriori* model, we sampled using PyMC3’s gradient-based No-U-Turn Sampler (Hoffman & Gelman 2014) in the bases indicated in the table at the bottom-left of the figure. We used \hat{R} as our convergence diagnostic (Gelman & Rubin 1992). If not all of the parameters had $\hat{R} < 1.01$ after the maximal run-time of the chains, a warning that the fit did not converge is noted in the ExoFOP upload comment strings (see Section 1).

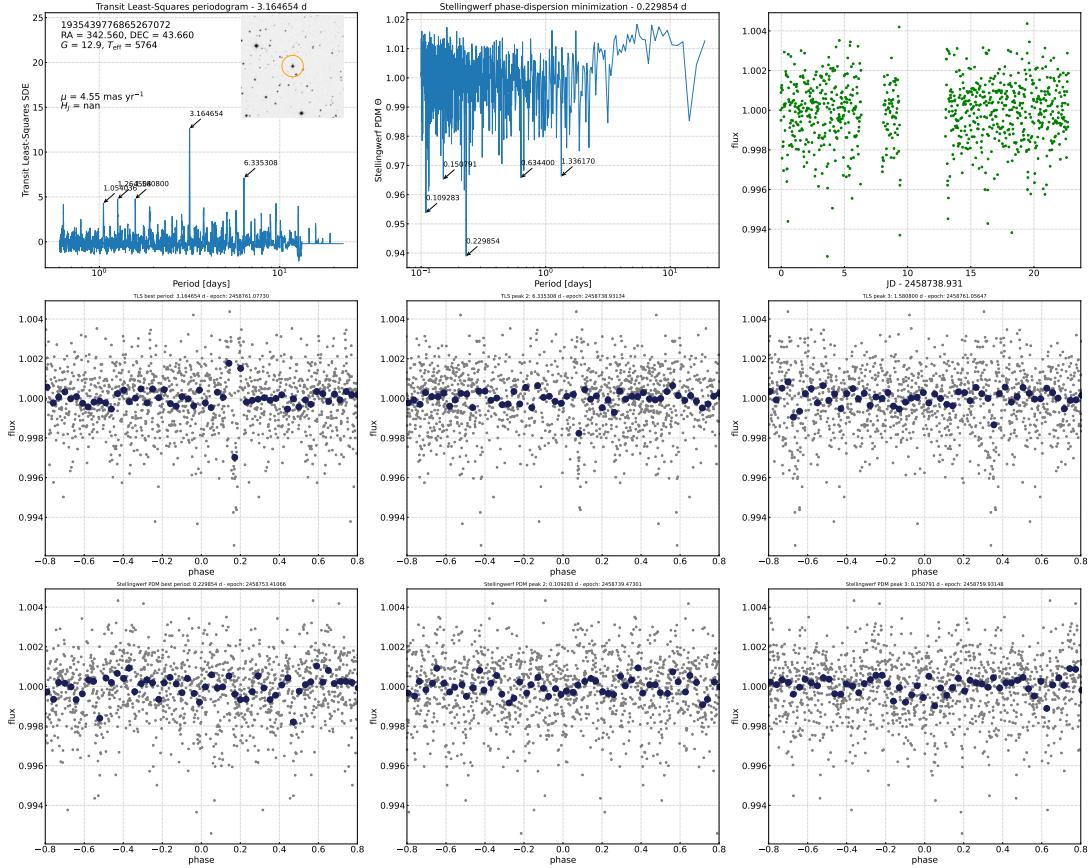


Figure 1. Transit search summary. See § 2.1.

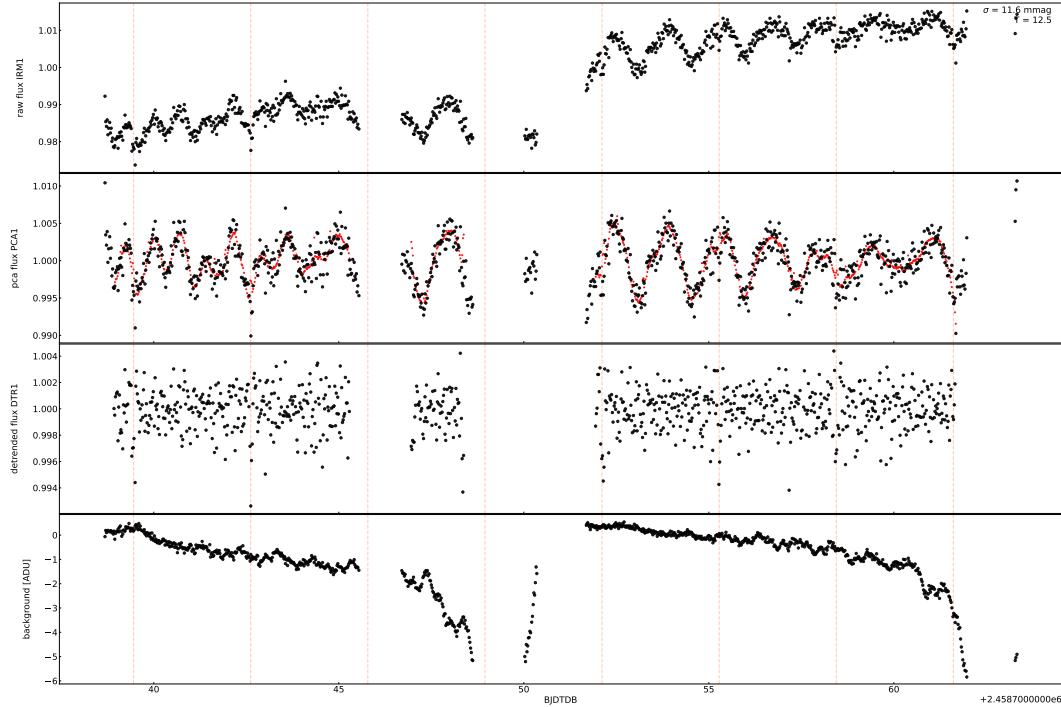


Figure 2. Light-curve diagnostics. See § 2.2.

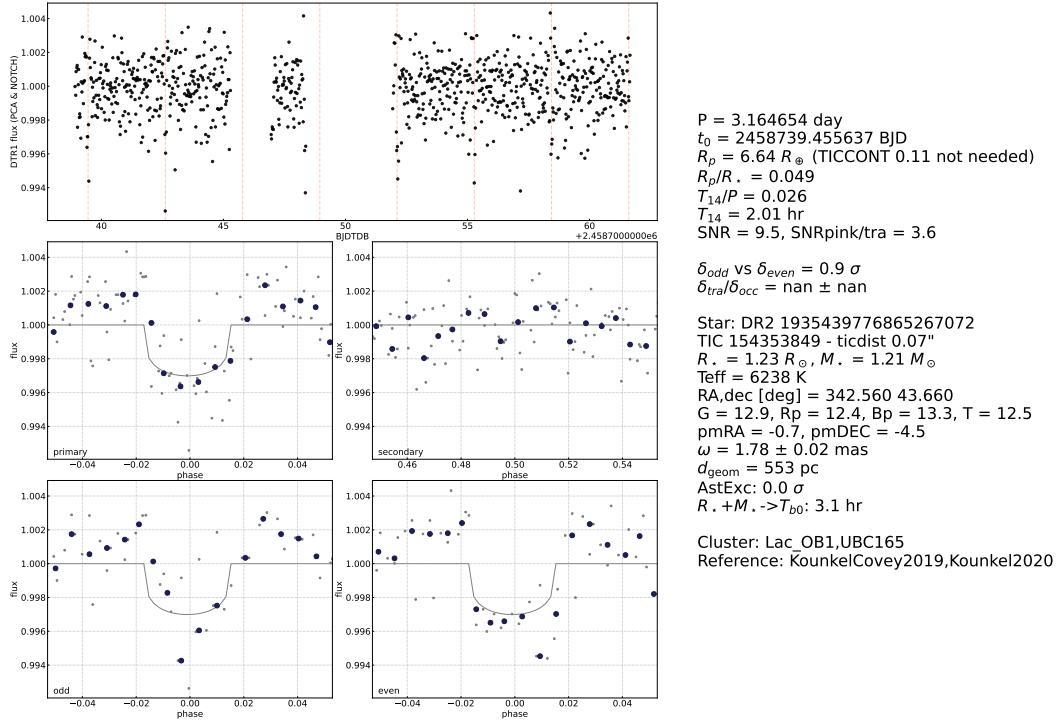


Figure 3. Transit diagnostics. See § 2.3.

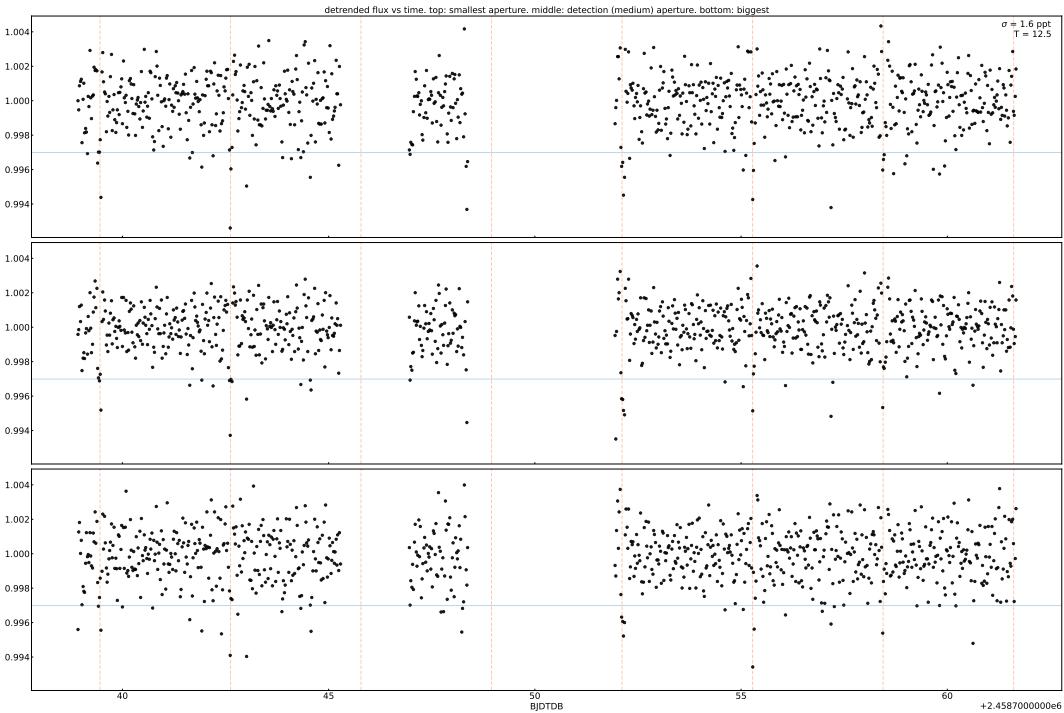


Figure 4. Light-curves for increasing aperture sizes. See § 2.4.

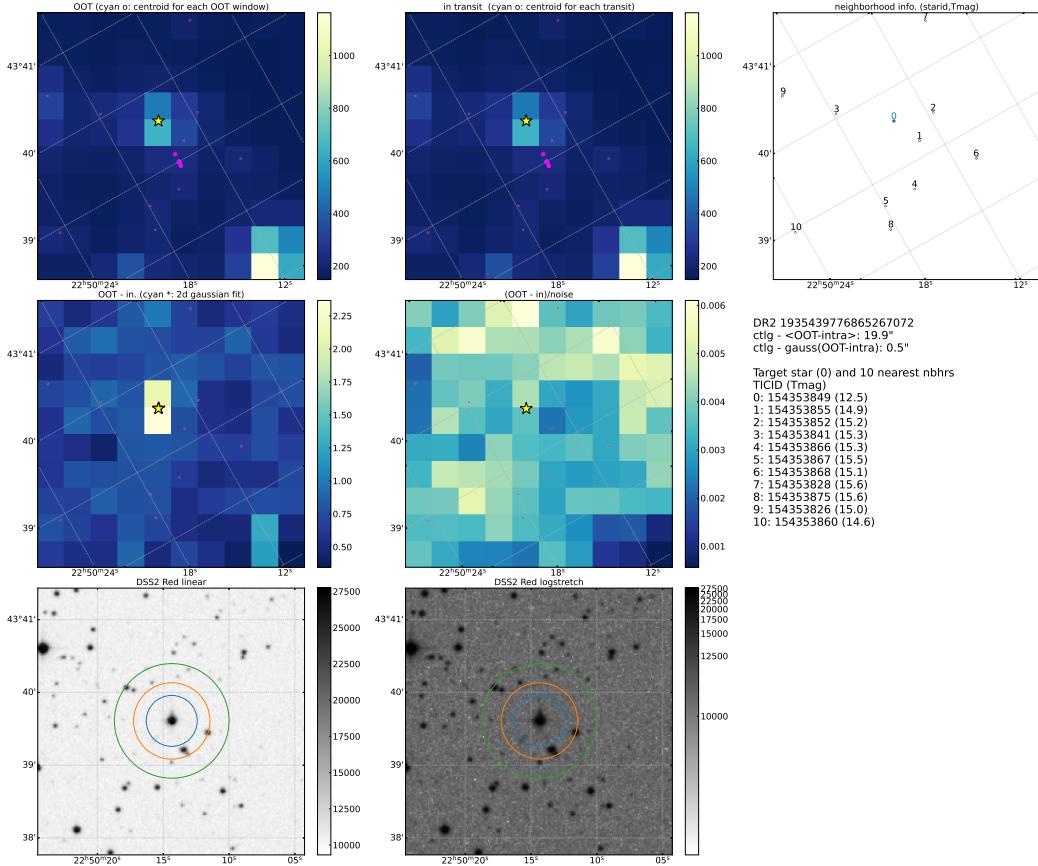


Figure 5. Imaging variability diagnostics. See § 2.5.

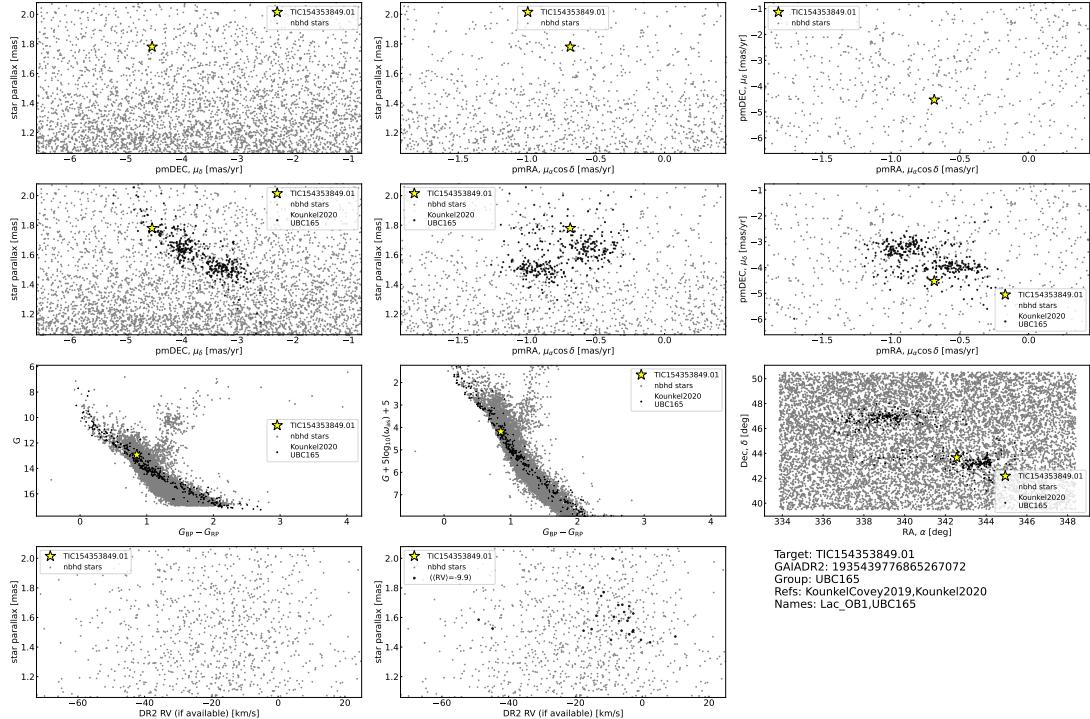


Figure 6. Cluster membership assessment diagnostics. See § 2.6.

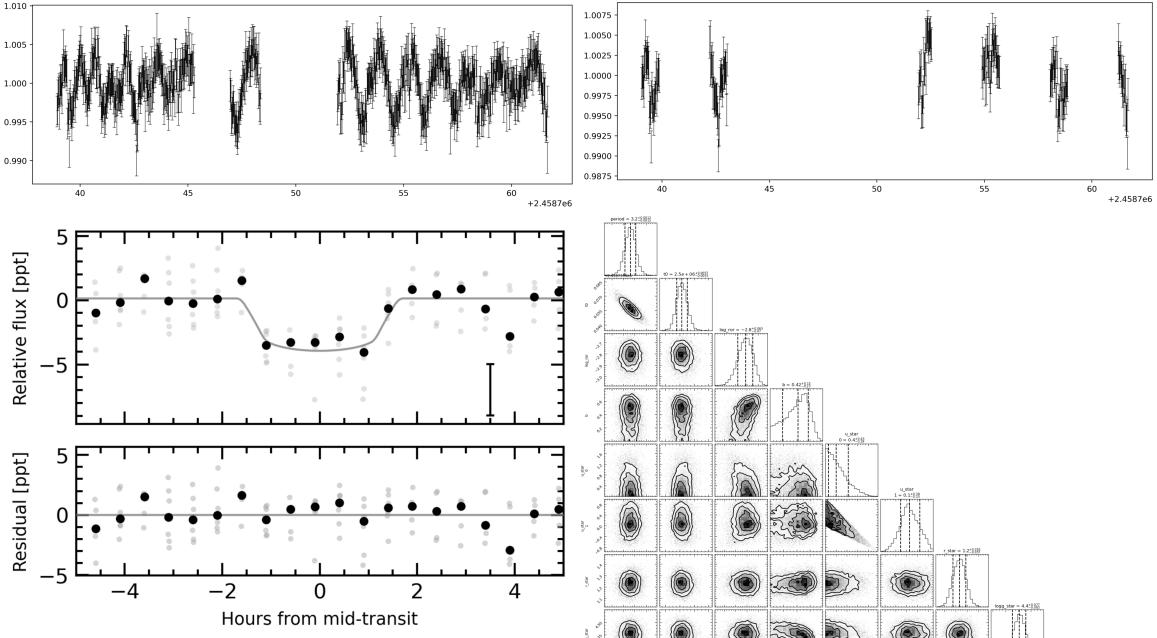


Table 1. Priors and posteriors for transit model with local polynomials removed.

Param.	Unit	Prior	Median	Mean	Std. Dev.	3% HUL	97% HUL	ESS	$R = 1$
p	d	$\mathcal{N}(0.0026, 0.0166)$	3.1055349	3.1055361	0.0012292	3.1045726	3.1058898	965	7.8e-04
χ^2_{ν}	-	$\mathcal{N}(145670, 45389, 0.02983)$	2458739.3915765	2458739.3915765	0.0002292	2458739.3915765	2458739.3915765	16244	7.8e-04
$\log R_p/R_s$	-	$\mathcal{U}(-4.054, 0.000)$	2.81263	2.81263	0.00878	2.84682	2.81222	6064	3.1e-04
b	-	$\mathcal{U}(0, 1 + R_p/R_s)$	0.4177	0.3912	0.1768	0.3881	0.6900	3560	3.9e-05
u_1	-	Kipping 2013	0.399	0.475	0.360	0.600	1.158	780	1.2e-04
u_2	-	Kipping 2013	0.103	0.121	0.348	-0.179	0.238	3630	1.2e-04
R_s	R_\odot	$\mathcal{N}(1.234; 0.058)$	1.232	1.231	0.056	1.122	1.330	9885	2.3e-04
$\log g$	cgs	$\mathcal{N}(4.338; 0.085)$	4.361	4.358	0.079	4.202	4.497	6134	5.0e-04
(f)	-	$\mathcal{N}(1.000; 0.100)$	1.001	1.001	0.092	0.999	1.004	1.000	1.0e-04
$\log \sigma_f$	-	$\mathcal{N}(\log(7.1); 2.000)$	-8.923	-9.690	1.007	-10.907	-7.436	7389	1.0e-03
R_p/R_s	-	-	0.060	0.060	0.004	0.053	0.068	6064	3.6e-04
ρ_p	g cm^{-3}	-	0.960	0.970	0.181	0.638	1.310	6068	2.0e-04
R_p	R_Jup	-	0.718	0.718	0.050	0.631	0.821	1.000	1.0e-04
R_p	R_transit	-	8.048	8.059	0.661	6.793	9.315	6462	2.4e-04
a/R_s	-	-	7.984	7.989	0.498	7.935	8.919	6065	3.2e-04
$\cos i$	-	-	0.052	0.050	0.024	0.047	0.087	1.000	1.0e-04
T_1	hr	-	2.942	2.946	0.170	2.624	3.265	2710	-1.5e-04
T_{13}	hr	-	2.539	2.531	0.206	2.124	2.891	5775	-1.9e-04

NOTE— ESS refers to the number of effective samples. R is the Gelman-Rubin convergence diagnostic. Logarithms in this table are base-e. \mathcal{U} denotes a uniform distribution, and \mathcal{N} a normal distribution. Many of the T_{13} statistics may be ∞ in the event of a grazing transit. (1) The ephemeris is in units of BJD-TDB.

Figure 7. MCMC transit-fitting diagnostic. See § 2.7.

REFERENCES

- Bhatti, W., Bouma, L. G., & Wallace, J. 2018, astrobase
- Bouma, L. G., Hartman, J. D., Bhatti, W., Winn, J. N., & Bakos, G. A. 2019, *ApJS*, 245, 13
- Bouma, L. G., Curtis, J. L., Masuda, K., et al. 2022, *AJ*, 163, 121
- Brasseur, C. E., Phillip, C., Fleming, S. W., Mullally, S. E., & White, R. L. 2019, *Astrophysics Source Code Library*, ascl:1905.007
- Bryson, S. T., Jenkins, J. M., Gilliland, R. L., et al. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 889
- Curtis, J. L., Agüeros, M. A., Mamajek, E. E., Wright, J. T., & Cummings, J. D. 2019, *The Astronomical Journal*, 158, 77
- Foreman-Mackey, D., Czekala, I., Luger, R., et al. 2020, exoplanet-dev/exoplanet v0.2.6
- Gelman, A., & Rubin, D. B. 1992, *Statistical Science*, 7, 457, publisher: Institute of Mathematical Statistics
- Green, G. M., Schlafly, E. F., Finkbeiner, D., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 478, 651
- Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, arXiv:1906.00966 [astro-ph], arXiv: 1906.00966
- Hippke, M., & Heller, R. 2019, arXiv:1901.02015 [astro-ph], arXiv: 1901.02015
- Hoffman, M. D., & Gelman, A. 2014, *Journal of Machine Learning Research*, 15, 1593
- Huang, C. X., Burt, J., Vanderburg, A., et al. 2018, *The Astrophysical Journal*, 868, L39
- Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, *The Astrophysical Journal Letters*, 713, L87
- Katz, D., Sartoretti, P., Cropper, M., et al. 2019, *Astronomy & Astrophysics*, 622, A205
- Kostov, V. B., Schlieder, J. E., Barclay, T., et al. 2019, arXiv:1903.08017 [astro-ph], arXiv: 1903.08017
- Kounkel, M., & Covey, K. 2019, *The Astronomical Journal*, 158, 122
- Kounkel, M., Covey, K., & Stassun, K. G. 2020, *AJ*, 160, 279
- Pecaut, M. J., & Mamajek, E. E. 2013, *The Astrophysical Journal Supplement Series*, 208, 9
- Rizzuto, A. C., Mann, A. W., Vanderburg, A., Kraus, A. L., & Covey, K. R. 2017, *AJ*, 154, 224
- Salvatier, J., Wieckiā, T. V., & Fonnesbeck, C. 2016, PyMC3: Python probabilistic programming framework
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *The Astrophysical Journal*, 500, 525
- Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, arXiv:1905.10694 [astro-ph], arXiv: 1905.10694
- Stellingwerf, R. F. 1978, *The Astrophysical Journal*, 224, 953
- Theano Development Team. 2016, arXiv e-prints, abs/1605.02688