DRAFT VERSION AUGUST 3, 2019
Typeset using LaTeX **twocolumn** style in AASTeX62

CDIPS. I. Light Curves in Open Clusters from TESS Sectors 6 & 7

L. G. BOUMA,[1] J. D. HARTMAN,[1] W. BHATTI,[1] J. N. WINN,[1] AND G. Á. BAKOS[1]

[1] *Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA*

(Received August 3, 2019; Revised —; Accepted —)

Submitted to AAS journals.

ABSTRACT

The Transiting Exoplanet Survey Satellite (TESS) has begun to photometrically monitor almost every star cluster in the solar neighborhood. Using the TESS images, we have begun a Cluster Difference Imaging Photometric Survey (CDIPS), in which we are focusing on stars that are candidate cluster members, or else show other indications of youth relative to field stars. Our aims are to discover giant transiting planets with known ages, and also to provide a light curve sample fit for studies in stellar astrophysics. For this work, we made 159,343 light curves of candidate young stars, across 596 distinct clusters. Each light curve represents between 20 and 25 days of observations of a star brighter than G_{Rp} = 16, sampled at 30 minute cadence. We describe the image subtraction and time-series analysis techniques we used to create the light curves, which have noise properties that agree with theoretical expectations. We also comment on the possible utility of the light curve sample for studies of stellar rotation evolution, and binary eccentricity damping. The light curves, which cover about one sixth of the galactic plane, are available at archive.stsci.edu/prepds/cdips.

Keywords: planets and satellites: detection – methods: data analysis — techniques: photometric — (Galaxy:) open clusters and associations: general —

1. INTRODUCTION

Each of the ∼3000 known star clusters of the Milky Way is a gift to astrophysics, providing a sample of stars that vary widely in mass but all have approximately the same age and composition. Time-series photometry of clusters has many applications. By measuring rotation periods over a range of ages, we can study the angular momentum evolution of stars and improve our ability to determine stellar ages through gyrochronology (*e.g.*, Skumanich 1972; Barnes et al. 2015; Meibom et al. 2015; Curtis et al. 2019). By measuring eccenticities as a function age, we can study the tidal circularization of binaries (Meibom & Mathieu 2005; Milliman et al. 2014; Price-Whelan & Goodman 2018). The detection of eclipsing binaries can also lead to the precise determination of the absolute dimensions of the stars and stringent tests of stellar-evolutionary models (Luhman 2012; Stassun et al. 2014; Kraus et al. 2015). Finally, transiting exoplanets discovered in clusters can shed light on the timescales for processes in planet formation, evolution, and migration (Fortney et al. 2007; Mann et al. 2016; David et al. 2016), as well as on the effects of metallicity (Fischer & Valenti 2005; Petigura et al. 2018).

The TESS mission (Ricker et al. 2015), though designed for other purposes, holds the promise to deliver the most homogeneous and comprehensive cluster photometric survey in history. More quantitatively, the Kharchenko et al. (2013) cluster member database indicates that $\approx 2\times10^5$ open cluster members brighter than T = 16 will be observed in the full-frame images over the first two years of TESS observations. This count includes the "most probable" members within each cluster's radius, which by the Kharchenko et al. (2012) definition requires the independent kinematic, photometric, and spatial membership probabilities to each exceed 61%. The actual number of stars in clusters is larger, as the membership catalogs are not yet complete, even at these relatively bright magnitudes (*e.g.*, Röser et al. 2016; Cantat-Gaudin et al. 2018, 2019b).

A major barrier to deriving precise photometry from the TESS images is the relatively poor angular resolution ($\approx 21''$ per pixel). Almost all clusters are within 10 degrees of the Galactic plane. The problems with crowding and complex backgrounds become so severe that the TESS Candidate Target List deprioritizes 2-minute targets at galactic latitudes less than $\approx 15^\circ$ (Stassun et al. 2018, 2019). This includes almost all star clusters[1]. This decision was made because

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

[1] TICv8 updated the cutoff to 10°. During the first year of TESS observations, the number used when selecting 2-minute target stars was 15°.

the large pixel size and the high stellar surface density make aperture photometry unreliable. However, it means that most stars in clusters, though capable of yielding a cornucopia of science, will go unprocessed by the official *TESS* data reduction pipeline.

One way to quantify the blending problem is to determine what fraction of the total flux in a photometric aperture is contributed by a particular target star. Aperture photometry is reliable when this fraction is close to unity. Difference imaging (Alard & Lupton 1998; Miller et al. 2008), in our group's experience, can be viable down to crowding fractions of perhaps 1 part in 10. Based on the Kharchenko et al. (2013) cluster membership data, and the density of background stars, we calculated that the median dilution fraction of cluster stars with $T < 16$ is 0.13, for an aperture radius of 2 pixels. For at least $\sim 10^5$ cluster stars then, it seems that difference imaging may be worthwhile.

Difference imaging avoids the primary effects of blending through forced-aperture photometry. In this method, the positions of stars are predicted from an astrometric solution, and the reference fluxes are determined from a calibrated catalog magnitude-to-flux relation. The deviation from the reference flux is measured on the difference image. Assuming that only a single source causes the variability, blended neighbor stars only act to increase Poisson noise, and do not "dilute" transit signals, down to the angular resolution of the source catalog used to determine the reference flux. This is a major benefit of performing image subtraction in crowded fields.

We have therefore begun to apply difference imaging to the TESS images, with a focus on any star that could be a member of a coeval group. We are also including some stars that we suspect are young due to combined photometric and astrometric indicators. A major motivation for this effort is to discover giant transiting planets with known ages. The focus of the present study however is to describe our methods, and to produce a general-purpose dataset applicable for studies both in exoplanetary and also stellar astrophysics.

For the remainder of this work, we adopt the term "star cluster", or simply cluster, to refer to a coeval group of stars. This includes the open clusters of yore, as well as the moving groups and stellar associations that have been discovered since the late 1990s (Zuckerman & Song 2004).

The plan of action is as follows. In § 2, we describe how target stars were selected. § 3 presents the photometric and image processing methods we used to produce light curves for these stars. The statistical properties of the 159,343 light curves from TESS Sectors 6 and 7 are then summarized (§ 4.1). An example of how we might use the data to identify pulsating stars, eclipsing binaries, and transiting planets is then given (§ 4.2). § 6 summarizes our findings, and discusses them with an eye towards the studies we hope this and subsequent rounds of data processing will enable.

2. METHOD: STAR SELECTION

The main aim of the CDIPS project is to increase the number of cluster stars for which photometric time-series are available, and thereby facilitate studies of exoplanetary and stellar processes across different times and stellar environments. A key step is therefore to define a sample of stars that are thought to be young, or members of clusters, or both.

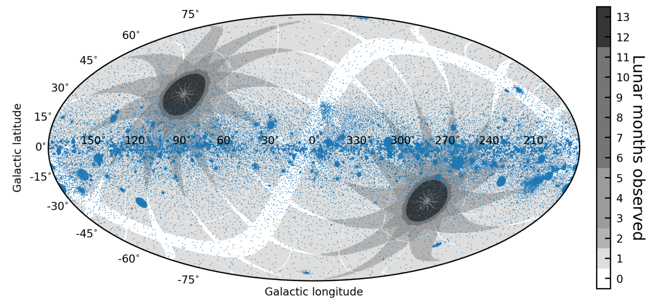

Figure 1. Target star positions (blue) and predicted TESS observing footprint (gray). Target stars are either candidate members of clusters, or else have other youth indicators (see § 2). Most will be observed for one or two lunar months over the TESS Prime Mission.

A homogeneous membership calculation for every known cluster is a rather large undertaking, and currently falls outside our scope. So too is a homogeneous search for young stars across the galaxy. Instead, we have opted to collect and concatenate appropriate catalogs from across the literature. We then use the resulting meta-catalog to identify our target stars within the TESS images.

The probabilistic criteria for inclusion in our target star list is necessarily heterogeneous across different catalogs. However, in our initial stellar selection we aim for completeness, not accuracy. If there has been a claim in the literature that a star should be considered a cluster member, or a young star, we would like to err on the side of reporting a light curve for the star. For stars that are photometrically interesting, we can then perform post-hoc quality checks using Gaia-DR2 astrometry and photometry to assess cluster membership and youth.

§ 2.1 describes the catalogs we used to identify candidate members of open clusters. § 2.2 describes the catalogs we used to identify candidate members of moving groups, stellar associations, as well as young stars identified through combined Gaia photometry and astrometry. § 2.3 and Figure 2 describe summary statistics for the entire sample of about one million target stars.

2.1. *Big catalogs: open clusters*

At the time of writing, two relatively large, homogeneous cluster memberships studies had been performed using *Gaia*-DR2: those by Cantat-Gaudin et al. (2018) and Gaia Collaboration et al. (2018a). There were also two large membership studies based on proper motion and photometric catalogs that were of interest: the studies of Kharchenko et al. (2013) and Dias et al. (2014).

Gaia-derived OC memberships—Cantat-Gaudin et al. (2018) used the UPMASK unsupervised membership assignment algorithm (Krone-Martins & Moitinho 2014) to identify clus-

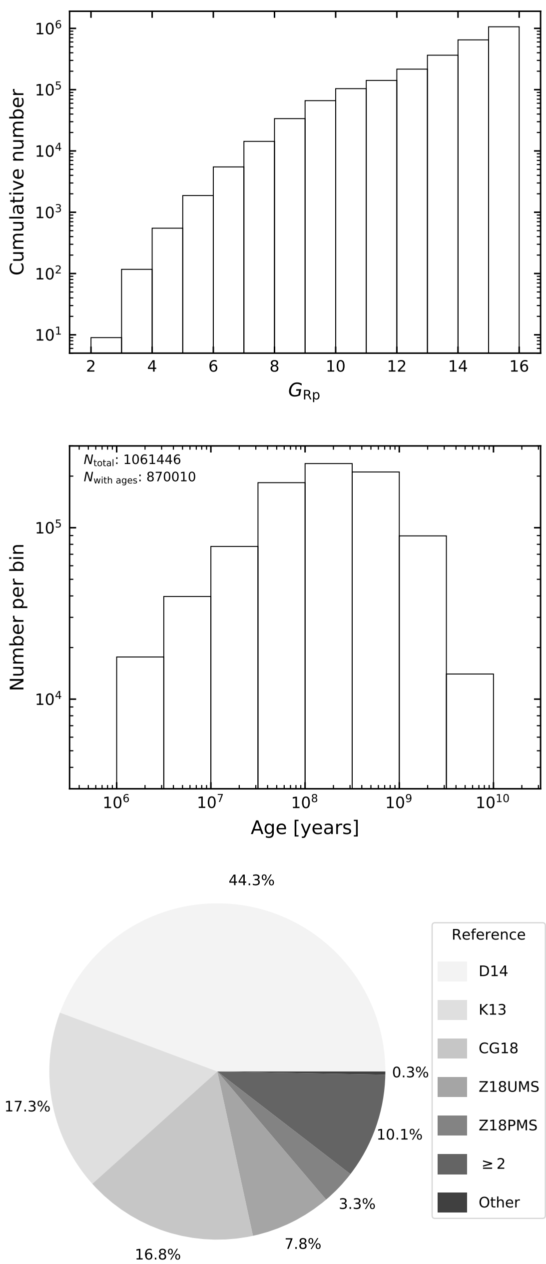

Figure 2. Target star statistics. *Top.* Cumulative counts as a function of Gaia *Rp*-band magnitude. *Middle.* Histogram of target star ages, for the subset of stars with ages matched against Kharchenko et al. (2013). *Bottom.* Provenance of cluster membership. Percentages are relative to N_{total} = 1061446 target stars. Symbols are as follows. D14 is Dias et al. (2014). K13 is Kharchenko et al. (2013). CG18 is Cantat-Gaudin et al. (2018). Z18 is Zari et al. (2018), with upper main sequence and pre-main-sequence samples sub-divided. "≥ 2" indicates at least two authors reported a star as a candidate cluster member.

ter members using Gaia-DR2 positions, proper motions, and parallaxes. They used *Gaia* photometry and radial velocities to then verify the claimed membership properties. From their Table 2, we collect an initial 401,448 cluster members, in 1229 clusters, down to their limiting magnitude of $G = 18$.

Gaia Collaboration et al. (2018a) reported memberships for stars in a smaller, more select group of well-studied open clusters. From their Table A1, we collect 40,903 cluster members, in 41 open clusters, mostly within 500 pc. While this work also included memberships for globular clusters, we omitted these from consideration.

Given the unprecedented quality of Gaia-DR2 astrometry, these two membership sources are our most reliable sources of membership information. In our photometric reduction, our default identifier for all sources is correspondingly the Gaia-DR2 `source_id`. The TIC identifiers are found through a spatial cross-match after the light curves have been made, subject to the requirement that the Gaia-DR2 source identifiers match (Stassun et al. 2018, 2019).

Pre-Gaia OC memberships—Kharchenko et al. (2013) used proper motions calculated in PPMXL (Röser et al. 2010, a combination of USNO-B1.0 and 2MASS astrometry) and near-infrared photometry from 2MASS (Skrutskie et al. 2006) to report the existence of 2859 open clusters and stellar associations (globular clusters were omitted by excluding any entry of type `'g'`). We selected the most probable cluster members ("1σ members") using the combined photometric, kinematic, and spatial criteria described by Kharchenko et al. (2012, Section 3.3). Then, to obtain *Gaia*-DR2 source identifiers for the members, we performed a crossmatch for *Gaia*-DR2 sources within 5 arcseconds of the listed positions. To improve the quality of the cross-match, we additionally used the 2MASS photometry to predict the G-band magnitudes[2], and required that the measured G-magnitude fall within 2 magnitudes of the predicted G-magnitude. If multiple neighbors matched the position and magnitude constraints, we took the nearest spatial neighbor as the match. From 373,226 stars, this yielded a unique best neighbor for 352,332 stars (94.4% of the sample), and a choice between two neighbors for 17,774 stars.

The second (non-*Gaia* derived) open cluster membership catalog we used was the Dias et al. (2014) catalog, which was based on UCAC4 proper motions acquired by the US Naval Observatory (Zacharias et al. 2013). From their 1805 reported open clusters, we selected sources with quoted membership probability above 50%. To obtain *Gaia*-DR2 source identifiers for the members, we performed a similar cross-match as before, looking for sources within 5 arcseconds of the listed positions, and within ±2 G-band magnitudes of the prediction. From 2,034,269 stars, this yielded a unique best neighbor for 1,828,630 stars (89.9% of the sample), and a choice between two neighbors for 8.7% of the remaining sample.

[2] See https://gea.esac.esa.int/archive/documentation/GDR2/Data_processing/chap_cu5pho/sec_cu5pho_calibr/ssec_cu5pho_PhotTransf.html, online, `2019-03-29`, or Carrasco et al. (2016)

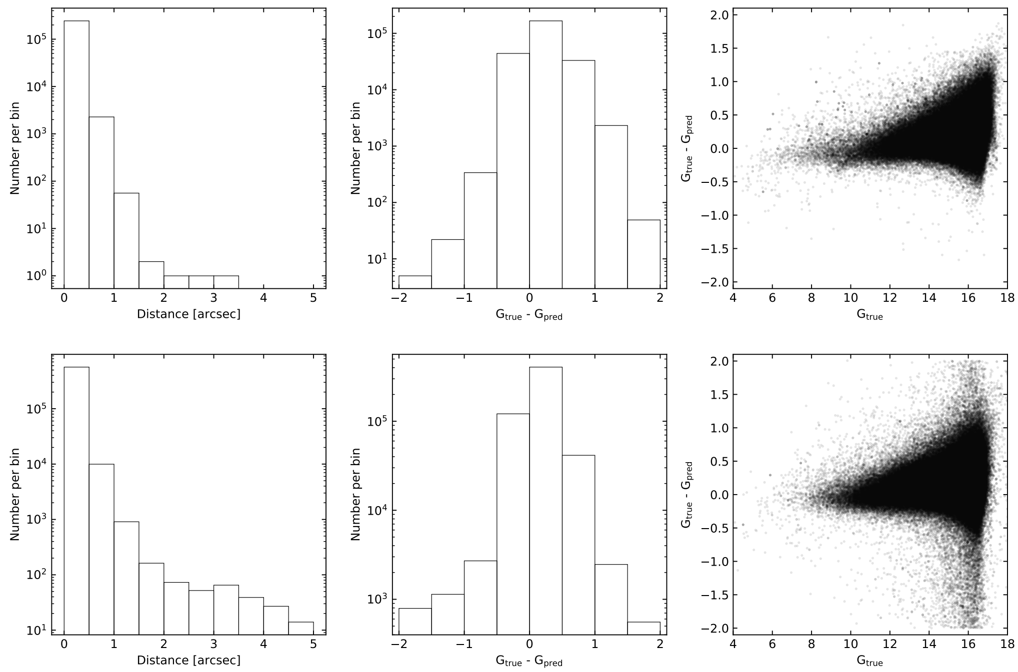

Figure 3. *Top.* Quality diagnostics from cross-matching Kharchenko et al. (2013) cluster members against Gaia-DR2. A histogram of the distances between matched stars is on the left; a histogram of the difference between the true G-band magnitude and that predicted from 2MASS photometry is in the middle; a scatter plot of the same magnitude difference as a function of G-band magnitude is on the right. *Bottom.* Same, but cross-matching Dias et al. (2014) cluster members to Gaia-DR2.

The distributions of various cross-matching statistics are shown in Figure 3. The distances between matches is typically below 1 arcsecond. The Dias catalog shows somewhat stronger crowding effects at the faint end compared to the Kharchenko catalog, and likely has a larger number of false matches. At their faint ends, both catalogs show a tendency for true G-band magnitudes being somewhat larger than predicted G-band magnitudes, presumably due to dust reddening.

2.2. *Smaller catalogs: moving groups and stellar associations*

Stars, moving groups and stellar associations are of interest for similar reasons as stars in open clusters. Though fewer stars are known to exist in moving groups, they are of particular interest because moving groups are less crowded than open clusters, and are often closer to the Sun.

We obtained Gaia-DR2 identifiers from the results of the following studies: Gagné et al. (2018b), Gagné et al. (2018a), Gagné & Faherty (2018), Kraus et al. (2014), Röser et al. (2011), Bell et al. (2017), Rizzuto et al. (2011), Oh et al. (2017), and Zari et al. (2018). The methods applied in these studies vary from kinematic analyses, to astrometric analyses included Gaia-DR1 parallaxes, to photometric searches for infrared excesses, to spectroscopic studies including RVs, Hα emission, and Li absorption.

For the Gagné et al. catalogs, a large number of the stars have high proper motions. However, a number of the stars do not have proper motions quoted. To perform the cross-match, we searched the Gaia-DR2 archive for sources within 10 arcseconds of the listed positions (propagated to the Gaia-DR2 J2015.5 epoch, if the proper motions were available, otherwise simply using the listed J2000 positions). We also imposed a $G < 18$ cut on any putative matches. We then chose the nearest neighbor by spatial separation. Of 3012 moving group members collected from the three combined Gagné et al. catalogs, this procedure yielded 2702 matches.

The Kraus et al. (2014), Röser et al. (2011), and Bell et al. (2017) studies reported members in Tucana-Horologium, the Hyades, and 32 Ori respectively. Applying the same procedure as for the Gagné catalogs gave 187, 684, and 119 best-neighbors respectively, compared to 205, 724, and 141 initially reported members. Note that Kraus et al. (2014) found that only $\sim$70% of their listed members have spectroscopic indicators consistent with their membership in Tucana-Horologium.

Rizzuto et al. (2011) focused on a single moving group: the Sco OB2 association. We used their reported Hipparcos identifiers, and matched against the *Gaia* archive's `hipparcos2_best_neighbour` table, which gave 319 nearest-neighbor stars from 436 candidate members.

Next, Oh et al. (2017) searched for comoving stars in the $\approx$2 million stars that overlapped between Tycho-2 and *Gaia*-DR1. They found many wide binaries, and also identified a large number of comoving groups. We chose the 2,134 stars that they reported were in groups with sizes of at least 3 stars. Using their *Gaia*-DR1 source identifiers, we matched against the *Gaia* archive's `dr1_neighbourhood` table, which gave 1,881 nearest-neigbor stars in groups of at least three stars (Marrese et al. 2019).

Finally, Zari et al. (2018) constructed a sample of young stars within 500 pc using data from Gaia-DR2. Two subsamples were made: (a) an upper main sequence (MS) sample, with 86,102 stars, and (b) a pre-MS sample, with 43,719 stars. Each was created from a careful combination of distinct astrometric and photometric cuts. These stars are the youngest, closest stars, spread across star-forming complexes in Sco-Cen, Orion, Vela, Taurus, and other regions of the sky. Though most of these stars are not directly identified with moving groups or open clusters, their reported youth and proximity to star forming regions justifies their inclusion in our search sample.

2.3. *Summary of target stars*

After collecting the aforementioned lists, we merged them into a single table. We queried the `gaiadr2.gaia_source` table to retrieve their photometric G, $G_{\rm Rp}$, and $G_{\rm Bp}$ magnitudes, as well as their astrometric measurements $(\alpha, \delta, \mu_\alpha, \mu_\delta, \pi)$. Finally, we required that $G_{\rm Rp} < 16$, which is roughly where the 1-hour photometric precision of TESS is predicted to reach 1% (Ricker et al. 2015).

All told, this procedure yielded 1,061,447 unique stars, from 13 distinct membership catalogs.

The relative fractions of stars from each catalog are shown in Figure 2. The largest number of stars come from Dias et al. 2014 (44.3% of stars), Kharchenko et al. 2013 (17.3%), Cantat-Gaudin et al. 2018 (16.7%), and Zari et al. 2018 (11.1%, of which 7.8% are OBA stars, and 3.3% are pre-MS stars). 107,647 of the stars, or about 10% of the collection, have cluster memberships reported by multiple authors. Since the membership probability calculations often use independent data and methods, agreement between multiple investigators on a given star's cluster membership is a helpful indication of it being a bonafide member. Stars reported in multiple catalogs have all available reference information concatenated.

The fraction of target stars from each catalog which are bonafide cluster members should be understood to be heterogeneous. For the Dias et al. (2014) catalog, their membership calculation included only spatial and kinematic information, and we used a relatively low probability threshold when including their stars (based on the criteria Dias et al. 2014 used for their star counts). The Kharchenko et al. (2013) catalog combined spatial, kinematic, and photometric information to derive their membership probabilities. We also used a somewhat more restrictive membership probability cut (again, following the criteria they used for their star counts), and so this sub-sample is likely less contaminated with field interlopers. Cantat-Gaudin et al. (2018) used spatial, kinematic, and astrometric information from Gaia-DR2. Despite the lack of photometric information, the quality of the Gaia data suggest that the field contamination rate will be lowest for the Cantat-Gaudin et al. (2018) sample.

To assign unique cluster names, we adopted the name matched against Kharchenko et al. (2013) whenever possible. Appendix B describes how this was done in detail. For moving groups not identified in Kharchenko et al. (2013), we used the constellation-based naming convention from Gagné et al. (2018b). Otherwise, we used the name reported by the original catalog claiming membership. This process reduced 16,425 name permutations down to 3,216 unique cluster names. Though we have made every effort to avoid duplicates, a small number may remain, so we advise inspection of the `cluster` column as well as the references given in the `reference` column rather than using the `unique_cluster_name` column to analyze individual objects of interest. Nonetheless, 87.7% of the ~million unique target stars are matched to clusters within Kharchenko et al. (2013), and 88.8% are assigned a cluster name. The remainder are mostly the young stars from Zari et al. (2018). Ages and their uncertainties are then merged against the parameters reported by Kharchenko et al. (2013).

The resulting CDIPS target star list is given in Table N. The cumulative distribution of target star brightnesses, as well as a histogram of the ages, is shown in Figure 2. Relative to field stars, our target star sample is young, with a typical age of 100 Myr.

3. METHOD: PHOTOMETRY

3.1. *Overview*

To reduce the TESS images to light curves, we adopted a difference imaging approach. The overall method is in the spirit of the reduction pipelines developed by Pál (2009), Huang et al. (2015), Soares-Furtado et al. (2017) and Oelkers & Stassun (2018); a conceptual overview is given in Figure 4. Most sub-modules of our pipeline have been developed over the past decade to process HAT data (*e.g.*, Bakos 2018). The specific high-level framework used for this reduction was adapted from a pipeline called `pipe-trex`, under development primarily for the HATPI project (hatpi.org). The code is available online[3].

We begin our processing with the calibrated full frame images produced by the Science Processing Operations Center at NASA Ames (§ 3.2). We then perform a collection of preparatory steps, including source extraction of bright stars, astrometric verification, and coarse simple aperture photometry of bright stars (§ 3.3). Using the information collected from these initial steps, we select an astrometric reference frame to which we transform all of the calibrated images. We construct a photometric reference by stacking a subset of these transformed frames, after having convolved the frames to have identical stellar profiles. Finally, we subtract each target frame from the photometric reference (§ 3.5). We perform aperture photometry on the subtracted images using positions projected onto the frame from the Gaia-DR2 source catalog. We detrend the resulting light curves (§ 3.7). The resulting white noise and red noise properties of the light curves, and a few interesting cases of variability, are discussed in § 4.

[3] github.com/waqasbhatti/pipe-trex, commit `cd9f9bb`.

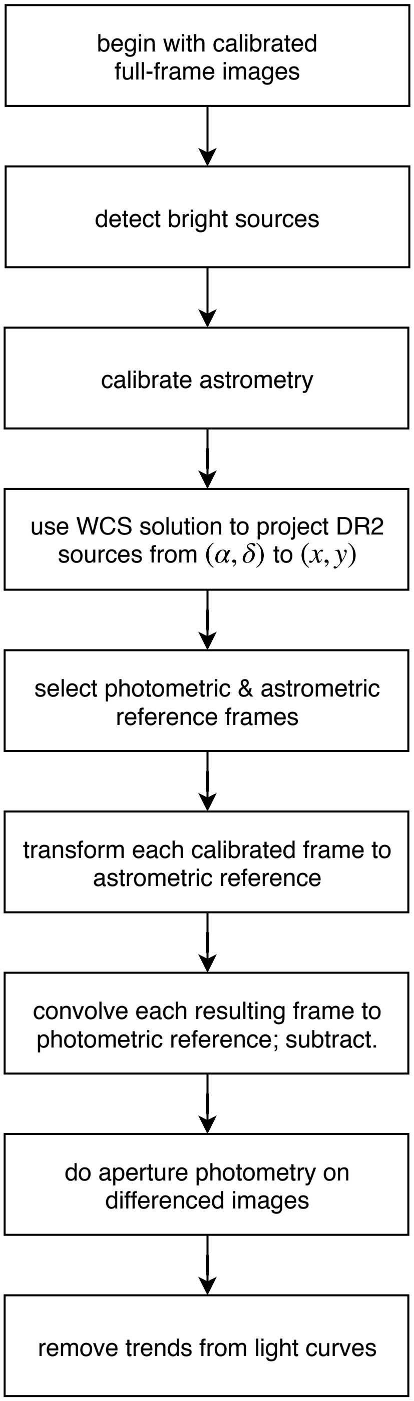

Figure 4. Conceptual overview of photometric reduction pipeline. Details are given in § 3.

3.2. *Observations*

The TESS spacecraft began science operations on July 25, 2018. To keep its cameras pointed opposite the Sun, the spacecraft advances by ≈28 degrees east in ecliptic longitude every lunar month. Data acquired throughout each "sector" are downlinked at spacecraft perigee through the Deep Space Network. Descriptions of the spacecraft's design and operations are given by Ricker et al. (2015) and the instrument handbook (Vanderspek et al. 2018).

For us, the main data product of interest is the calibrated full frame image (FFI). Each TESS camera is read out every 2 seconds. To produce a manageable telemetric load, the resulting pixel values are averaged by the onboard computer into 30 minute exposures. An on-board cosmic ray mitigation algorithm is applied (Vanderspek et al. 2018, §5.1). Once transmitted to the ground, the raw images are calibrated by the Science Processing Operations Center (SPOC). The calibration process includes an overscan, bias, and dark current correction, and also divides out a flat field. Details are discussed by Clarke et al. (2017), and the resulting science data products are described by Tenenbaum & Jenkins (2018).

We perform our processing using the calibrated images, their corresponding uncertainty images, and the associated headers. The spacecraft has four cameras, and each camera has four CCDs. In the following analysis, all image-level operations are thus performed on the images for each CCD, so that at any instant of time there are 16 independent images that require analysis.

While we performed numerous initial tests on the first sectors of data, by geometric coincidence Sectors 1–5 were pointed away from the galactic plane. Consequently, less than 2% of the CDIPS target star sample was observed in the first five TESS sectors. Though a few interesting clusters are present in these observations (*e.g.*, Blanco 1, NGC 2516, NGC 1901), for the present work we opted to focus on sectors in which there were more stars of interest for our intended science. These begin in Sector 6 (2018-12-12, spacecraft orbit #19). For this run of processing, they conclude at the end of Sector 7 (2019-02-01, spacecraft orbit #22). Sectors 6 and 7 cover galactic longitudes from roughly 200° to 280°, with fair coverage within ±20° of the galactic plane (Figure 1).

3.3. *Image preparation & background removal*

Before we can perform any kind of photometry, a few janitorial tasks are required. First, we convert the multi-extension calibrated FITS image from MAST into a single-extension FITS image, and trim to remove virtual rows and columns using the `SCIROWS`, `SCIROWE`, `SCCSA`, and `SCCED` header values.

In order to preemptively address the background variations present in some frames due to scattered light from the Earth and Moon (see Vanderspek et al. 2018, §7.3.1–7.3.4), we then estimate and subtract the large-scale background. We do this by temporarily masking out pixels more than 2σ from the image median, and then pass a 48×48 median box filter over each pixel in the image, with reflective boundary conditions. The resulting background estimate has low-amplitude structure over spatial scales of a few pixels, so we then blur it with a gaussian kernel to produce a smooth background estimate for each image. These steps also remove a low-level vignetting present in the corners of many images, which remains even after flat-fielding (see Vanderspek et al. 2018, §7.3.5). The results are shown in the upper four panels of

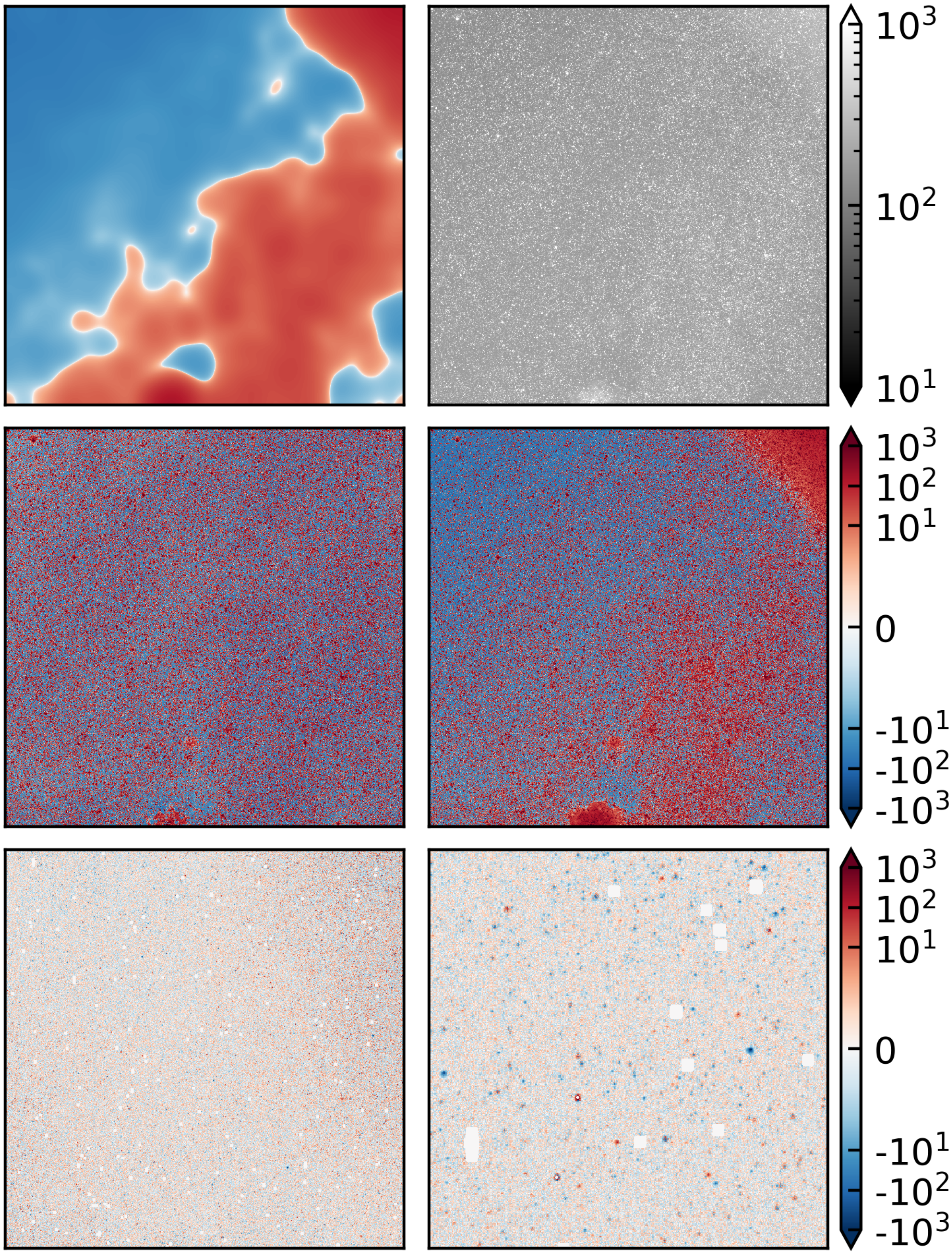

Figure 5. Stages of image processing for an image taken during the middle of an orbit. *Top right*. Calibrated image. *Top left*. Smooth background estimate. *Middle right*. Calibrated image minus its median pixel value. *Middle left*. Calibrated image minus smooth background. *Bottom left*. Registered difference image. *Bottom right*. Zoom of difference image. Each image is (2048×2048) pixels, except for the bottom-right image, which is (500×500). All images except the top-right have the same color map. The sector, camera, and CCD are (6, 1, 2).

Figures 5 and 6. Features with spatial scales smaller than ≈48 pixels remain, but large scale scattered light patches are removed.

After subtracting the background, we mask out saturated pixels using a fixed saturation level of 8×10^4 ADU. This value was chosen based on the onset of bleeding charge trails in the images, and is slightly greater than the saturation level of 2×10^5 electrons, or about 4×10^4 ADU, quoted by Vanderspek et al. (2018). We correspondingly do not analyze stars brighter than $T \approx 6.5$. As described by Pál (2009), the pixel masks are metadata to the image, and are only applied to the pixel values during the specific image processing steps in which they are necessary (*e.g.*, convolution). We also extend the masks beyond purely saturated pixels to "bloomed" pixels horizontally and vertically adjacent to the saturated pixels (see Figure 6 of Pál 2009).

Finally, for frames with the `DQUALITY` bit-flag corresponding to the "momentum dumps" and "coarse pointing modes" described by Vanderspek et al. (2018), we omit the entire frame. This removes on average a few frames per sector, out of about one thousand. Through visual inspection, we see that the stars on these frames are extremely smeared, and are unlikely to produce useful science data. In addition, we use the sector-specific data release notes[4] to identify further times with anomalous spacecraft performance, which we omit from consideration. For Sectors 6 and 7, this included three days at the beginning of Sector 6 dedicated to acquiring pixel response function data, and no additional gaps in Sector 7.

3.4. *Metadata collection & WCS verification*

Having prepared the images, we perform some initial analysis steps to produce metadata needed during image subtraction.

First, we perform source extraction on the thousand or so brightest, non-saturated stars in each image. This is done using a `fitsh` tool, `fistar`. We derive centroid positions for the stars, and simultaneously fit elongated gaussians to their profiles, yielding the shape parameters (s,d,k), where the flux f as a function of position (x,y) in the CCD image plane is assumed to take the form

$$f(x,y) = B + A\exp\{-0.5 \times [s(\Delta x^2 + \Delta y^2) + d(\Delta x^2 - \Delta y^2) + k(2\Delta x \Delta y)]\}, \quad (1)$$

for (x_0, y_0) the central coordinates of the star, $\Delta x = x - x_0$, $\Delta y = y - y_0$, B the background level, and A an arbitrary flux-scaling constant. For a nearly circular shape profile, the sharpness s is related to the FWHM as $\mathrm{FWHM} \approx 2.35\sqrt{s}$ (*e.g.*, Pál 2009). These shape parameters are later used when selecting an astrometric reference frame (§ 3.5). In agreement with what is obvious upon visual inspection, this fitting process shows that stars closer to the center of each camera's field are round, while stars near the field edges are more triangular.

For the astrometric solution, we use the World Coordinate System (WCS) and fourth-order Simple Imaging Polynomial (SIP) coefficients derived by SPOC and included in the FFI headers (Pence et al. 2010, Sec. 8). We explored the possibility of using `astrometry.net` (Lang et al. 2010) to derive our own astrometric solutions for each frame, but found that the astrometric residual (the mean separation between projected and measured positions) was consistently a factor of 1.5-2 times higher in our WCS solutions than in those given by SPOC. This was perhaps because we did not develop a robust algorithm to select non-blended stars of intermediate brightness before measuring their positions. Nonetheless, we did find that extending the astrometric correction to sixth order (compared to the fourth-order polynomial coefficients in the SPOC solution) did not improve the quality of the fit.

With the resulting WCS information, we then project a source catalog onto each frame. We use the projected positions of the sources to center the apertures in our photometry, rather than attempting to detect the positions. Such "forced-aperture photometry" is preferable to source extraction in the crowded fields that are central to this work. The Gaia-DR2 epoch is J2015.5, so even the fastest-moving stars with proper motions of $\sim 1\,\mathrm{arcsecond\,yr^{-1}}$ are still well within one pixel of their predicted positions in the TESS images. The projection from catalog sky-coordinate positions to pixel coordinates is performed using an analog of the `wcs-rd2xy` program that performs the standard matrix algebra (Lang et al. 2010). The source catalog look-up is performed using `gaia2read`[5] (Kim 2018).

For the source catalog itself, we initially planned to photometer all Gaia-DR2 sources in each field down to a cutoff of $G_{Rp} < 16$. However, for the galactic plane fields this produced an excessively large number of sources (millions of stars per $12^\circ \times 12^\circ$ CCD). We therefore limited our source catalog for each frame to be a combination of the CDIPS target stars ($G_{Rp} < 16$), and all Gaia-DR2 sources down to $G_{Rp} < 13$. The latter set of stars are used for image processing and light curve detrending.

The residual between the measured and projected stellar centroid positions is displayed for one photometric reference frame in Figure 7. The construction of this frame will be described shortly, however a few salient points can first be made. First, the typical median precision of the WCS solution is a bit below 0.1 pixels, and its 90$^{\text{th}}$ percentile is typically less than 0.3 pixels[6]. The errors are typically largest in the corners of the field of view, where the non-linearity of the focal plane is most significant, and the corrections required by the SIP coefficients are largest.

Finally, to collect the metadata needed to select photometric reference frames, we perform aperture photometry on the bright stars. This task is performed by using `fiphot` to sum the counts inside circular apertures centered on the projected

[4] archive.stsci.edu/tess/tess_drn.html

[5] github.com/samuelyeewl/gaia2read

[6] In fact, in our reduction we automatically impose that the median residual and 90th percentile remain below 0.2 and 0.4 pixels, respectively.

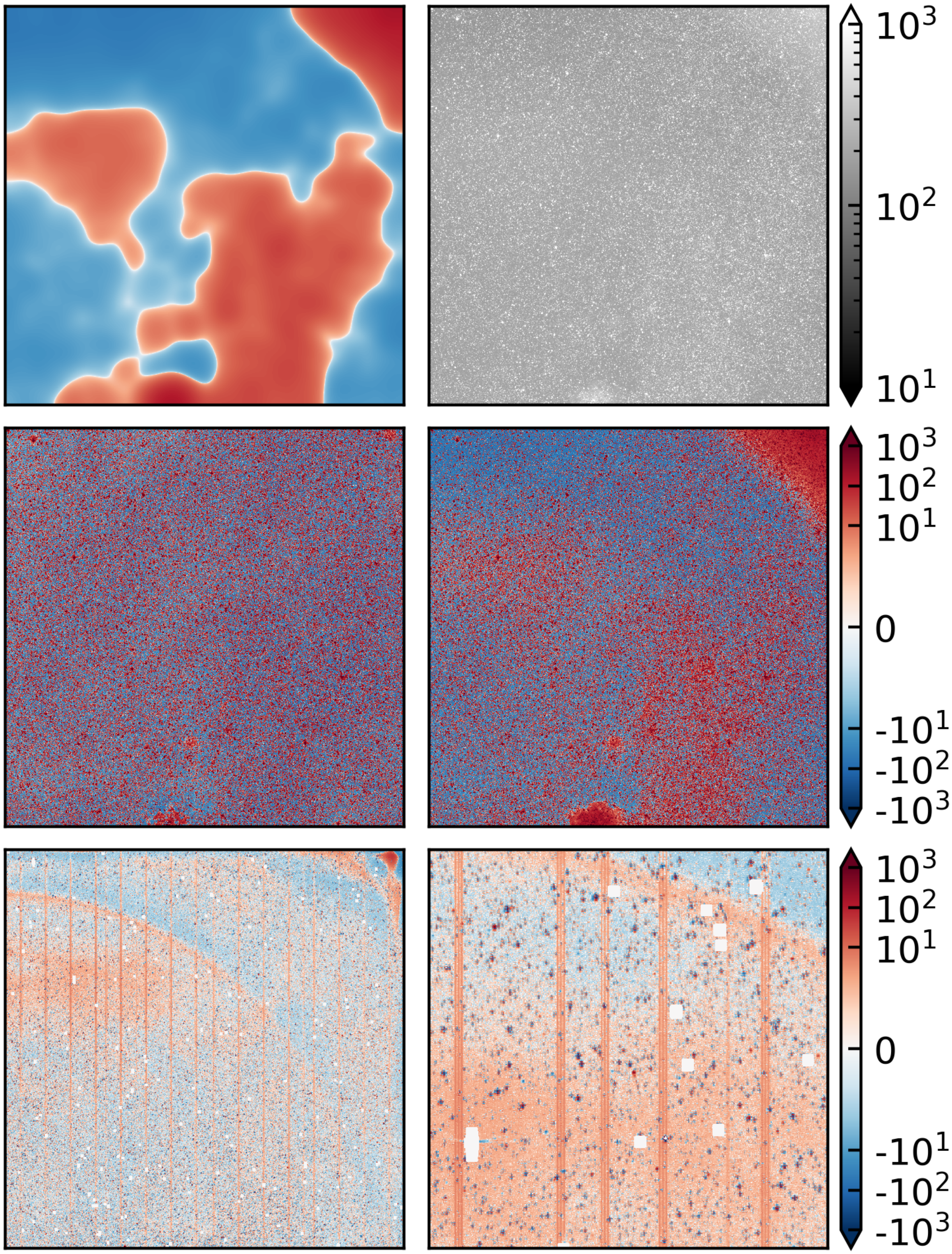

Figure 6. Same sector, camera and CCD as Figure 5, but for an image taken during perigee passage, when scattered light from the Earth is prominent. A number of systematic artifacts are present, including vertical "straps" and small-scale structure in scattered light patches. The quality of astrometric registration in the difference image is also worse, leading to larger residuals in the lower-right panel.

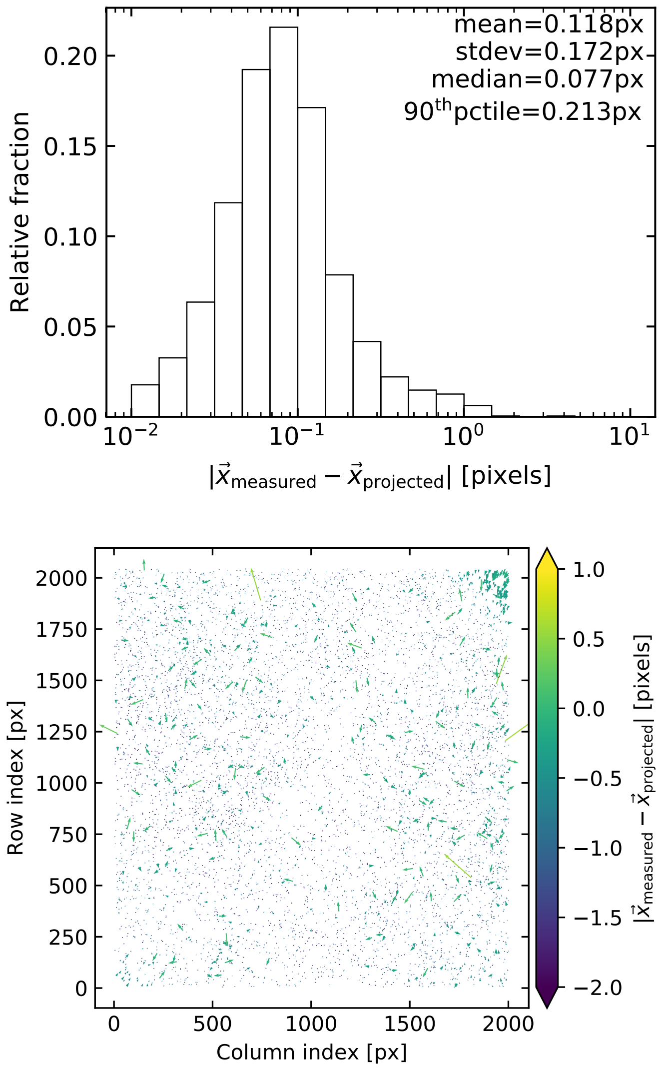

Figure 7. *Top.* Histogram of astrometric residual. The x-axis shows the distance between the measured centroid positions of stars, compared to the predicted positions from the WCS solution. *Bottom.* Vector plot of astrometric residual. Each arrow is the vector from the measured to the projected star position. Directions are correct, but lengths are 50 times their true size for visualization purposes. The systematic error in the top-right corner is a typical problem generic to wide-field astrometry. The frame chosen for this plot is the photometric reference frame used for Sector 6, Camera 1, CCD 1; we automatically impose cutoffs on the median and 90$^{\text{th}}$ percentile of the astrometric residual in order to ensure similar levels of astrometric precision are maintained throughout the reduction.

stellar positions. The pixel weights are equal to the fraction of the pixel that falls within the aperture. They are unity for pixels entirely within the aperture, and fractional along the aperture boundary. The background levels are measured in annuli around each aperture center. The resulting measurements, for instance of the background level of each aperture, and the number of stars with "good" photometry, can then be used to assess the photometric quality of each frame.

3.5. *Image subtraction*

3.5.1. *History of image subtraction method*

The core operation of "classical" image subtraction attempts to match a photometric reference R and a target image I by computing and applying a convolution kernel. For ground-based data, this "match" typically corrects for differences in seeing or transparency between the reference and target; for space-based data, the match might correct for spacecraft jitter, or thermal and corresponding point-spread function (PSF) variations. The kernel, once applied to the high signal-to noise reference, produces a model image, M_{xy},

$$M_{xy} = (R \otimes K)_{xy} + B_{xy}, \tag{2}$$

where B_{xy} is a component of the model image that allows for background variations, and $\otimes$ denotes convolution. Since we modeled the background separately (§ 3.3), we set $B_{xy} = 0$. The convolution kernel K is typically decomposed onto a basis, $K = \sum_i c_i K_i$, and the coefficients are found by minimizing

$$\chi^2 = \sum_{xy} \left(\frac{I_{xy} - M_{xy}}{\sigma_{xy}} \right)^2, \tag{3}$$

where σ_{xy} is the uncertainty in the target image pixel values. Photometry is then performed on the difference image D_{xy}, where $D_{xy} = I_{xy} - M_{xy}$. For the present reduction, the uncertainty in each target image pixel was taken to be a constant.

The general procedure described above was first proposed by Alard & Lupton (1998). It was reviewed and clarified by Miller et al. (2008). The choice of how to decompose the kernel was further explored by Bramich (2008), who showed that using a linear combination of delta functions (also called a "discrete kernel") had advantages compared to a basis of gaussians. We perform the convolution using `ficonv`, and opt for the implementation of Bramich's method (see Pál 2009 § 2.8). The lower panels of Figure 5 show the procedure working well, and producing a "clean" difference image. Figure 6 shows what happens for an image taken when scattered light from the Earth causes the model image to be a poor fit to the target image.

3.5.2. *Astrometric registration*

To make the above high-level picture work, we need to select two "reference frames": (1) the astrometric reference; and (2) the photometric reference.

To choose the astrometric reference, we search for frames with large, round stars (big s, small d and k values). We also require that the frame have minimal background noise, as measured in annuli around the bright stars selected in § 3.3. Finally, the astrometric reference needs to have, relative to the other frames being considered, a large number of detected sources (though the variance between frames was rather small). We sort the frames using these metrics, and then select the astrometric reference from successive intersections of each sorted list.

We then compute and apply a spatial transformation of each calibrated frame to match the astrometric reference frame. This transformation – a combination of rotation, dilation, and translation – typically moves stars by less than a pixel, since the TESS spacecraft pointing is quite stable. We calculate the transformation using the measured source positions found in § 3.4, and the symmetric triangle point-matching scheme described by Pál (2009, § 2.5.2). This step is achieved using the `fitsh` tools `grmatch` and `grtrans`. To help ensure the precision of the transformation, we require the "unitarity" Λ (Pál 2009, Eq. 54), which characterizes the degree of distortion in the transformation matrix, to be below 0.01. To mitigate possible photometric errors incurred during this step, we also use the flux-conserving interpolation scheme described by Pál (2009), which is necessary because polynomial interpolation schemes do not conserve stellar flux.

3.5.3. *Photometric reference frame construction & reference flux measurement*

The second required reference frame is the photometric reference, which is used both to calculate the convolution kernel, and to obtain a reference flux for each star. To make it, we first choose 50 images with low background measurements (measured for each frame from the annuli around bright stars), and only consider frames with a relatively large number of detected bright objects. We then convolve these candidate photometric reference frames to the frame with the lowest background measurement, and construct the photometric reference as the median image across the 50 frame stack.

Measuring the reference flux for each star is a non-trivial operation. First, we perform forced simple aperture photometry on the photometric reference frame to measure the flux for each source. The local background is estimated in annuli, with neighboring stars masked out during the background measurement. If we were to stop here, *it would be a bad mistake*. The reference fluxes for faint stars would be overestimated, due to crowding. The relative amplitude of photometric signals for faint stars would correspondingly be biased small, which would hinder exoplanet detection, among other applications.

To avoid this problem, after performing simple aperture photometry on the reference frame, we fitted a line between the catalog T-band magnitude of each star, and its measured flux. The T-band magnitude was calculated according to Equation 1 of Stassun et al. (2019). We then use the known catalog magnitude to predict the expected reference flux for each star. This minimizes crowding down to Gaia's resolution limit of $\approx 1''$, rather than the TESS limit of $\approx 20''$.

The final instrumental flux values f we report are given by (Pál 2009, Equation 83)

$$f = f_{\rm reference} + f_{\rm subtracted} \tag{4}$$

$$= g\left(T_{\rm cat}\right) + \frac{1}{||K||_1^2} \sum_{x,y} D_{x,y} (w \otimes K)_{x,y}. \tag{5}$$

The function g takes as input the target star's catalog magnitude $T_{\rm cat}$, and returns the reference flux. Its coefficients have been fit per the procedure just discussed, independently for each aperture. The difference image D is equal to $I - R \otimes K$, where as in Equation 2, I is the target image transformed to the astrometric reference, R is the photometric reference, and K is the convolution kernel. The weights w from the circular aperture mask are included in the convolution. The norm $||K||_1$ is defined by Pál (2009) Equation 81.

3.6. *Choice of convolution kernel*

To solve for the coefficients to the convolution kernel, a few further assumptions are necessary. The procedure implemented in `ficonv` is to subdivide the image, and within each grid element find the brightest non-saturated star. These isolated "stamp" stars are then used to solve for the coefficients of the kernel, by minimizing Equation 3 over the sum of all stamps. For the kernel basis, we use a linear combination of delta functions with a flux scaling term (Soares-Furtado et al. 2017 Section 3.3.1 gives the equations). In this model, spatial variations of the PSF across the image are captured by weighting each basis component with spatial polynomials up to a cut-off order.

This kernel model has three free parameters: (1) the box-size; (2) the maximum order of the polynomial weighting the delta function terms; (3) the maximum order of the polynomial weighting the flux scaling. We performed a grid-search to tune these parameters, in which our "loss functions" were the measured RMS as a function of magnitude, and the recovered SNR of transits from the catalog of known TOIs (TESS Objects of Interest; N. Guerrero, in prep).

In the first dimension, we varied the kernel size between a box of 3×3 pixels and 11×11 pixels. Increasing the kernel box-size from a 3×3 box to a 7×7 box led to about a 50% lower light curve RMS for bright stars, and no difference for faint stars. The largest kernels, of (11×11) pixels, returned slightly lower signal-to-noise for recovered transits than kernels of intermediate size. We settled on a kernel box-size of (7×7) pixels, which is ≈ 2 times larger than the typical TESS FWHM at field center.

In the other two dimensions, we varied the spatial polynomial orders weighting the kernel's individual pixels between first and fifth order. We did the same for the polynomial weights of the "identity" pixel. Varying the polynomial orders between 1 and 4 did not produce large differences. The fifth order polynomials retrieved transits with $\approx 10\%$ worse SNR compared to lower order polynomials. We therefore adopted a second order polynomial weight in both terms.

Averaging over all TOIs present in the camera we used for these experiments, we found that different choices of kernel parameters produced variations of $\lesssim 12\%$ in the retrieved transit SNR. For computational expediency, we therefore chose a single (7×7) kernel with second-order spatial polynomial weights in the basis functions for the remainder of our reduction.

However, we caution that within these fine-tuning experiments, we found that transit signals near the noise floor could fall under the noise floor for different cases of tuned parameters. In the longer term, developing an image-subtraction method that marginalizes over uncertainties of how to chose "optimal" kernels would be desirable. Pixel-level image subtraction methods that omit these parameters entirely are similarly worth exploring (Wang et al. 2017).

With a kernel selected, and the convolution and subtraction performed, we calculate the instrumental fluxes for each frame per Equation 5. We did this with three different aperture sizes: for this work, circles of radii 1 pixel, 1.5 pixels, and 2.25 pixels. These sizes were chosen to roughly span the range of optimal aperture sizes for stars in our sample, as calcuated from the pre-flight Sullivan et al. (2015) noise model. Finally, to convert from a list of flux measurements for each source on a frame to a list of flux values at any given time, we used the `fitsh` transposition tool `grcollect`.

3.7. *Light curve detrending*

The preceding steps produce light curves that include both instrumental systematics as well as astrophysical variability. Figure 8 shows twenty stars of comparable brightness randomly selected from two CCDs. Stars that are far apart on the same CCD often share similar changes in flux. In other words, the instrumental systematics seem to dominate. This problem is generic to most wide-field photometric datasets, including the WASP, Kepler, and HAT surveys Pollacco et al. (2006); Borucki et al. (2010); Bakos (2018). To remove the systematic variability, we adopted two different approaches: *(i)* the trend-filtering algorithm (TFA, Kovács et al. 2005), and *(ii)* a principal component analysis (PCA).

However, a number of other approaches to the problem were possible. To encourage future improvements, we describe the possibility of decorrelating against external parameters (§ 3.7.1), and also different available approaches to ensemble detrending (§ 3.7.2), before explaining our adopted implementation.

3.7.1. *Decorrelating against external parameters*

Often, ensemble trends of stellar magnitude with CCD position, sub-pixel position, catalog magnitude, and color are present in datasets. One detrending step that can be valuable is to fit and subtract a linear combination of these trends as they appear across many light curves (*e.g.*, Zhang et al. 2016, § 5.5).

A separate step for each light curve can then be to fit out linear correlations of stellar magnitude with "external parameters" (EPD, Bakos et al. 2010; Huang et al. 2015). For ground-based data these parameters might include zenith angle, or changing PSF shape. For TESS data, they might include CCD temperature, or the angles of the Moon and Earth relative to each camera's boresight. They might also include the standard deviation of the spacecraft quartnerion timeseries (Vanderburg et al. 2019). Some example "external" parameters that we include with our light curves are shown as functions of time in Figure 9.

We explored the possibility of fitting linear models of flux as functions of *e.g.,* temperature, shape parameters, and centroid positions to each light curve. We also briefly explored non-linear model fitting using N-dimensional B-splines to fit the flux, centroid positions, and temperatures simultaneously (Dierckx 1996). The linear models typically underfit the light curves, particularly during the large shifts that happen as the spacecraft nears perigee. The non-linear models showed some promise, but often seemed to overfit stellar variability signals. Given these complications, for the time being we omitted the step of "detrending" as a function of external parameters. To enable further exploration of the issue, we included all the necessary vectors of *e.g.*, centroid positions, temperatures, and shape parameters in our reported light curves.

3.7.2. *Ensemble detrending*

The parameters that capture systematic trends are often poorly known. In such cases, an effective model of the systematics comes from constructing a set of vectors that empirically captures trends common to many stars. Each target light curve is then assumed to be a linear combination of the trend vectors.

The well-known algorithms, TFA, Sys-Rem, PDC-MAP, and ARC2, all take slightly different approaches to constructing this set of basis vectors, as well as to solving for the weights to assign each linear component (Kovács et al. 2005; Tamuz et al. 2005; Smith et al. 2012; Aigrain et al. 2017). TFA selects individual "template stars" as basis vectors, and equally weights each template when solving for the coefficients via linear least squares. PDC-MAP computes "cotrending basis vectors" (CBVs) by applying singular value decomposition to the light curves that show the strongest mutual correlation. It solves for the coefficients through a two-step procedure. The first step is to calculate the coefficients through linear least squares. The least-squares coefficients are then used to construct a prior over plausible coefficient values, which is subsequently used to recompute the maximum likelihood coefficients for each star. This latter step reduces overfitting for stars with variability not present in the set of CBVs.

We opted to use two different detrending approaches, each aimed at a different use case. For transit-search related science, we used TFA, as implemented in `VARTOOLS` (Kovács et al. 2005; Hartman & Bakos 2016). For stellar astrophysics related work, we used a simple variant of PCA, as implemented in `scikit-learn` (Pedregosa et al. 2011). For self-consistency, we describe each method and its implementation in the following paragraphs.

TFA detrending—The idea of TFA is as follows. Suppose we have M "template stars", which are a subsample of stars that

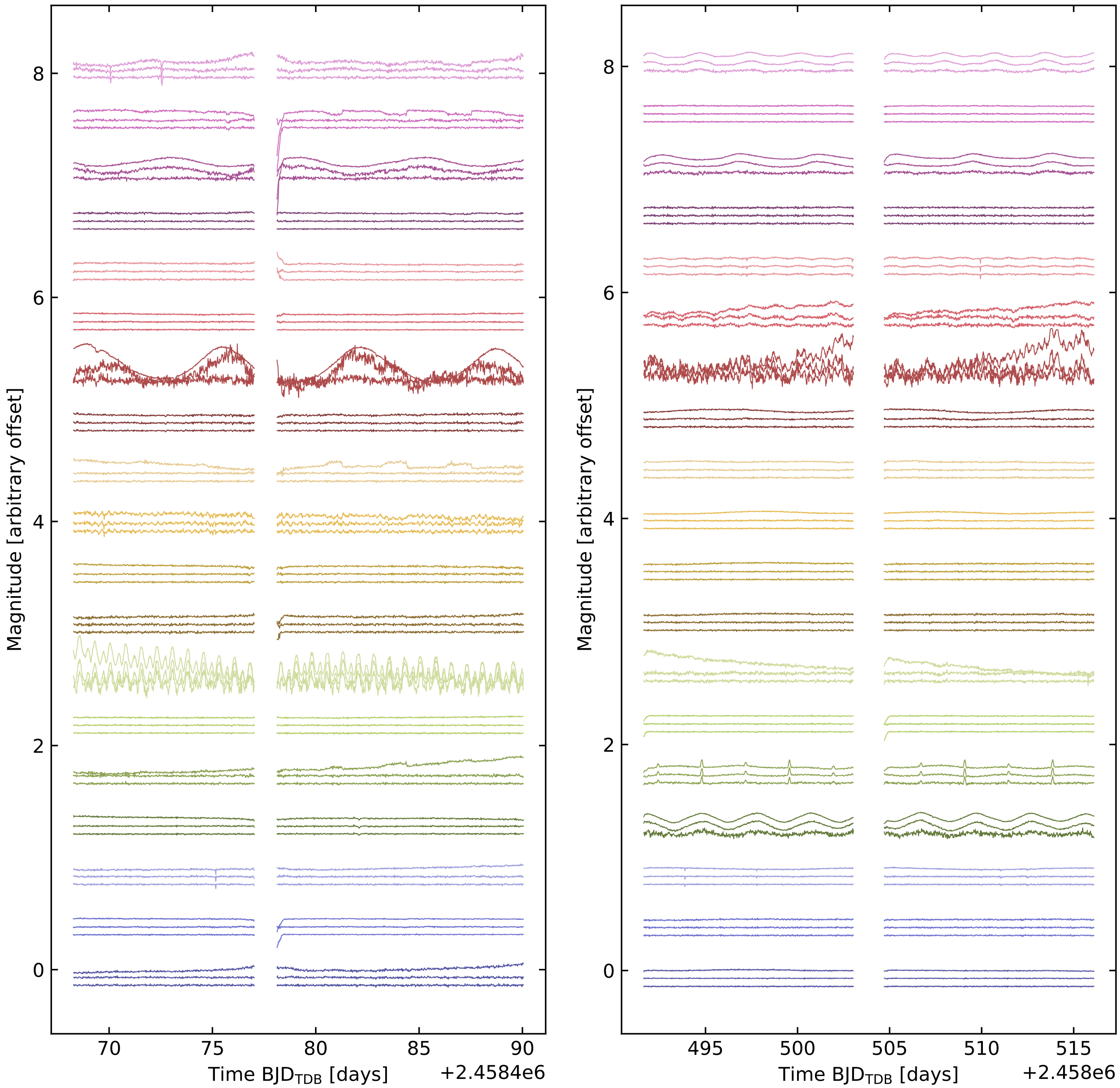

Figure 8. Twenty randomly selected light curves drawn from the same sector, camera, and CCD, for CDIPS target stars with T-band magnitudes between 13 and 14. For each star, we show the raw light curve (top), PCA-detrended light curve (middle), and TFA-detrended light curve (bottom). The sector, camera, and CCD numbers are 6, 1, 4 (left) and 7, 3, 2 (right). A number of systematic trends are shared across raw light curves (*e.g.*, the periodic $\sim$3 day "chopping" seen in the left plot). In the PCA light curves, stellar rotation signals are usually preserved, but not always (*e.g.*, left panel, seventh from the top). TFA filters out almost all long-term trends, or else heavily distorts them (*e.g.*, right panel, fourth from the bottom)

represent all types of systematics across the dataset. Each template star has a light curve with N data points. Denote the template time-series $X_j(i)$, where $j = 1, \ldots, M$ and $i = 1, \ldots, N$ is the time index. We then want to find periodic signals in a target time-series $Y(i)$. This is done by defining a filter function

$$F(i) = \sum_{j=1}^{M} c_j X_j(i), \tag{6}$$

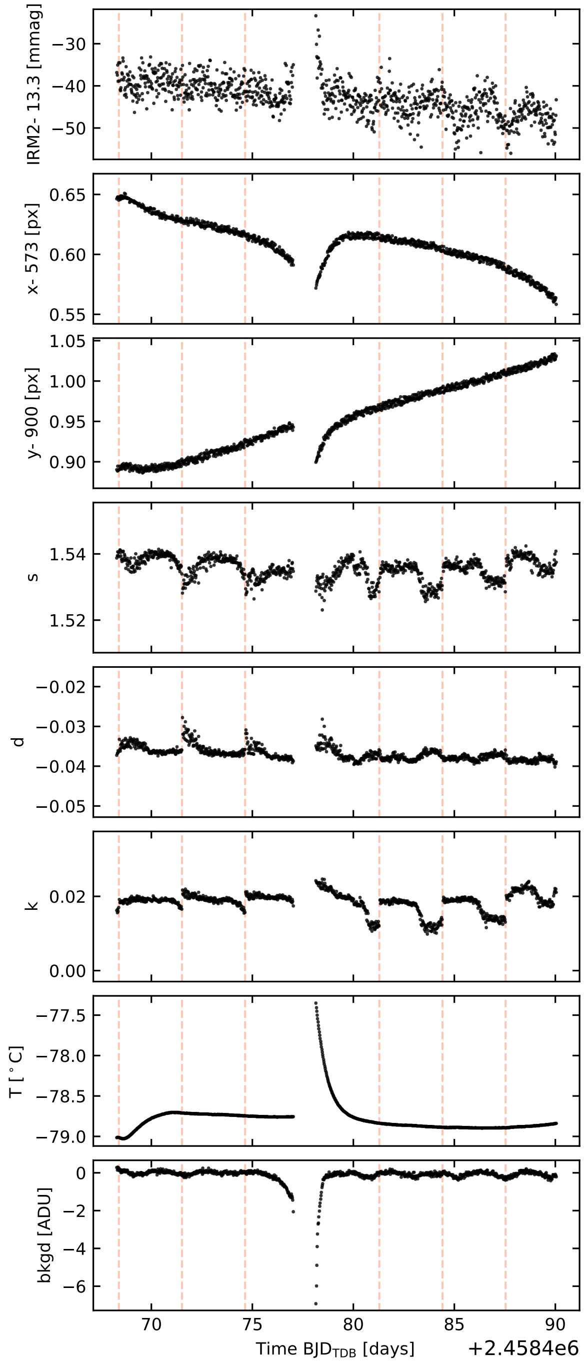

Figure 9. The variability in flux is sometimes correlated with variability in "external" parameters, shown here for a representative star over two orbits. *Top*: Instrumental raw magnitude (with a particular aperture size), as a function of time. Continuing in order are x and y centroid positions as a functions of time, the (s,d,k) PSF shape parameters, the CCD temperature, and the measured background value. Differential aberration affect the centroid position over the span of each orbit. Momentum dumps are marked with vertical dashed lines, and affect the measured shapes of stars.

for which the coefficients c_j are found by minimizing

$$\mathcal{D} = \sum_{i=1}^{N} [Y(i) - A(i) - F(i)]^2 \,. \tag{7}$$

When trying to find periodic signals, $A(i)$ represents our prior knowledge of the light curve's shape. This prior is simply that stars on average maintain a constant brightness:

$$A(i) = \langle Y \rangle = \frac{1}{N} \sum_{i=1}^{N} Y(i) = \text{const.} \tag{8}$$

If a signal is eventually found, for instance using the box-least squares method (Kovács et al. 2002), this detrending process must then be repeated while accounting for our updated knowledge about the light curve's shape.

Some implementation notes follow. We selected template stars in two stages. In the first stage, we fitted a parabola in the RMS-magnitude plane, and discarded stars more than 2σ away from the prediction of the fit. We also required that these initial candidate stars have intermediate brightness ($8.5 > T > 13$), and have a relatively large number of time-series data points. We then performed an initial iteration of TFA, on only the candidate template stars. We inspected the resulting detrended light curves for residual structure by computing a Lomb-Scargle periodogram. If the maximum-power peak had a false alarm probability below 0.1%, we excluded the star from the list of candidate template stars, on the basis of its presumed periodic variability. We then randomly selected at most 200 template stars from the remaining non-variable candidates. The choice of number of template stars was discussed by Kovács et al. (2005), and is another free parameter in the broad problem of light curve production. This choice is analogous to the issue of how many cotrending basis vectors to choose in PDC-MAP or ARC2 (Aigrain et al. 2017). While the number of template stars can be optimized by constructing and minimizing a BIC-like quantity, a little overfitting is acceptable for our pupose of finding planetary transits. For other applications, *e.g.,* stellar rotation period searches, it is almost certainly preferable to adopt a less forceful detrending approach.

PCA detrending—To remove the largest systematic trends with minimal overfitting of *e.g.*, stellar rotation signals, we adopted a simple variant of PCA (see *e.g.*, Ivezić et al. 2014 for a review). A similar approach was taken by Feinstein et al. (2019) in `eleanor`. We derived the principal components for each CCD using the 200 template stars previously selected for TFA. We then modelled each target light curve as a linear combination of a subset of these principal components, with weights determined by linear least squares.

To determine the "optimal" number of principal components to use, we performed a cross-validation analysis on the template stars using the components derived both from PCA, as well as a separate factor analysis. As is typically the case when the noise variance is different for each "feature" (each point in time), we found that k-folding cross-validation using

PCA failed to recover a plausible number of components[7]. However, the factor analysis typically maximized the cross-validation score with anywhere from 10 to 15 components, which agreed with a visual analysis of the number of PCA components past which overfitting began. We therefore used the maximum cross-validation score from thefactor analysis as the number of principal components to use in each light curve. This number is documented in the FITS header of each light curve.

4. RESULTS

4.1. *Light Curve Statistics*

4.1.1. *Stellar properties*

The on-sky locations of the light curves from Sectors 6 and 7 are shown in Figure 10. Black stars are the field stars for which we performed photometry, and from which we selected template stars for TFA, as well as the construction of the PCA eigenvectors. Blue points are the 159,343 stars with light curves available from archive.stsci.edu/prepds/cdips. Most of the target stars are quite close to the galactic plane. Gaps between CCD chips are also visible. The highly clumped nature of the target stars is also apparent.

Figure 11 shows the cumulative distribution of TESS T-band magnitudes for the target stars. Though most of the targets are faint, $\approx 3\times10^4$ are brighter than $T = 13$. The brighter targets will be more amenable to detailed spectroscopic observtions, should the need arise.

An HR diagram for the entire sample of CDIPS stars on silicon is shown in Figure 12 (top). The sub-sample of stars with measured positive parallaxes and naive distances less than 1 kpc is also shown (bottom). About one-third of the stars are in this latter sample. Of the close stars, a relatively large fraction come from Zari et al. (2018), and are either on the PMS or upper main-sequence. The latter set of OBA dwarf stars, while "younger" than the typical field dwarf, are likely the least interesting subset of our target sample from the perspective of age analyses. In the entire sample (Figure 12 top), a much larger fraction of stars are sub-giants, red giants, and helium-burning red-clump stars.

Finally, Figure 13 shows the proper motions of the entire and close samples of stars on silicon. Each clump signifies a different star cluster. The large overdensity in the top panel is composed mainly of field star contaminants. It has two subcomponents, due to the two different directions in the Galaxy being observed in Sectors 6 and 7.

4.1.2. *Cluster membership provenance*

Sector 6—In Sector 6, 67,612 light curves of candidate cluster stars were made. The provenance of the target star for these sources is Dias et al. (2014) for 40% of the sources; Zari et al. (2018) for 11% of sources from their upper main-sequence table and 7% of sources from their PMS table; Kharchenko et al. (2013) for 18% of sources, Cantat-Gaudin et al. (2018) for 14% of sources, and more than two catalogs for the remaining 10% of sources.

Table 2 shows the clusters in Sector 6 with the largest numbers of light curves. The top four clusters are all moving groups or stellar associations from Dias et al. (2014): Platais 5, Platais 6, Mamajek 3, and Collinder 70. These membership claims should be regarded with serious skepticism on a source-by-source basis. For Platais 6, Kharchenko et al. (2013) claimed only about 400 probable members (1σ) to exist within the angular radius of the cluster. Mamajek 3 (= 32 Ori) similarly has only about 50 confirmed members (Bell et al. 2017).

Sector 7—In Sector 7, 91,731 light curves of candidate cluster stars were made. The provenance of the claimed cluster origin of these sources is Dias et al. (2014) for 28% of the sources; Kharchenko et al. (2013) for 24% of sources; Zari et al. (2018) for 8% of sources from their upper main-sequence table and 3% of sources from their PMS table; Cantat-Gaudin et al. (2018) for 22% of sources, and more than two catalogs for the remaining 15% of sources.

Table 3 shows the clusters from Sector 7 with the largest number of light curves. Many rich open clusters, including NGC 2437, 2477, 2546, and 2451 are in the field. The cluster listed with the most light curves is the Collinder 173 association, chiefly due to the Dias et al. (2014) memberships. Kharchenko et al. (2013) note Collinder 173 as being coincident with the Vela OB2 association, and hence they flag it as a "duplicate". Along with Sco-Cen and the Orion star forming complex, the Vela OB2 association is one of the largest star forming complexes within 500 pc (Zari et al. 2018). Crucially though, Cantat-Gaudin et al. (2019a) recently studied the region using Gaia DR2 astrometry, and found that they could subdivide the complex into seven distinct stellar populations, with a total of ∼11,000 members. None of these populations map one-to-one to the classic labels of "Collinder 173" or "Vela OB2" – so the Cantat-Gaudin et al. (2019a) catalog is likely a more reliable source of information for this particular region of the sky.

4.1.3. *Light curve noise properties*

Observed RMS vs magnitude.—The standard deviation of the TFA-detrended light curves is plotted as a function of the catalog T-band magnitude for CDIPS light curves in Figure 14. For the y-axis of this plot, we have taken

$$\mathrm{RMS} = \left[\frac{1}{N-M-1}\sum_{i=1}^{N}(f_i-\bar{f})^2\right]^{1/2}, \tag{9}$$

where f_i is the value of the flux at the i^{th} point in the time-series, $\bar{f}$ is the mean flux, N is the number of points in the time-series, and M is the number of template light curves used during TFA detrending. The correction in the denominator penalizes the natural degree of overfitting inherent to the TFA algorithm.

[7] See for example https://scikit-learn.org/stable/auto_examples/decomposition/plot_pca_vs_fa_model_selection.html

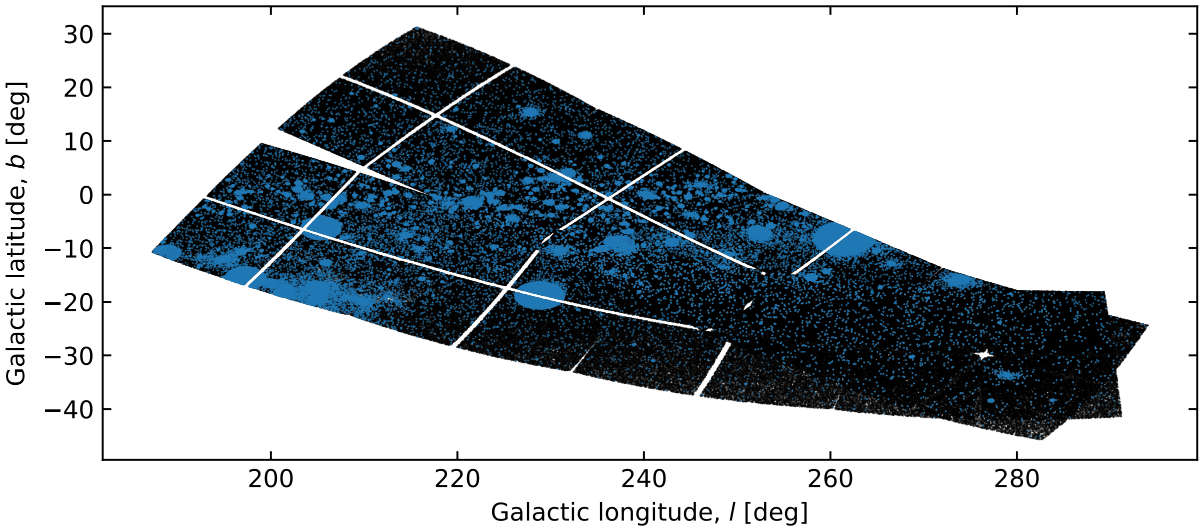

Figure 10. Positions of light curves from TESS Sectors 6 and 7 in galactic coordinates. Black: $G_{\rm Rp} < 13$ field stars. Blue: $G_{\rm Rp} < 16$ target stars. Target stars are mostly near the galactic plane. The data for Sectors 6 and 7 cover about one-sixth of the galactic plane.

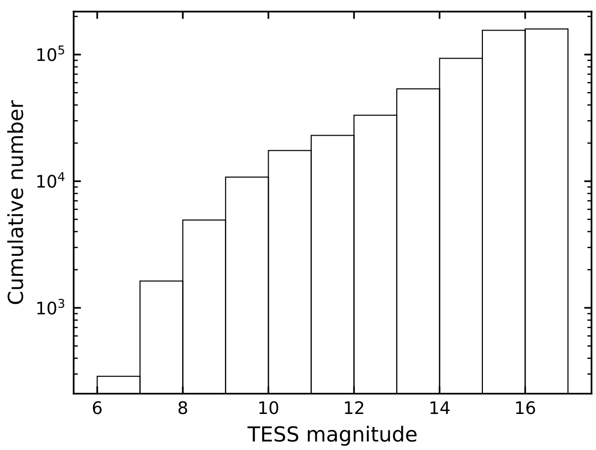

Figure 11. Cumulative number of CDIPS light curves a function of TESS T-band magnitude. Light curves were made for the target stars (Figure 2) that were observed in Sectors 6 or 7.

The observed RMS (black points) follows the usual shape, with photon noise dominating from $T=9$ to $T=12$, beyond which the onset of the "sky" background changes the overall slope of the curve to be more steep. For the brightest stars ($T \lesssim 9$), a "systematic floor" was part of the mission's error budget (Ricker et al. 2015), but has not been observed in early reports of the photometric performance of various aperture photometry pipelines (*e.g.,* the SPOC pipeline Jenkins et al. 2010, the MIT-QLP Huang et al. 2018, and `eleanor` Feinstein et al. 2019). The fact that our light curves for the brightest stars are above this purported "floor", rather than below it, suggests that our image subtraction techniques could be introducing some small degree of noise to the light curves of the brightest stars. It is also true however that our largest aperture contains only about 16 pixels, which is sub-optimal for stars brighter than $T \approx 9$ (see Sullivan et al. 2015, Figure 14). Since the brightest stars are not the focus of the present work, we leave this issue unaddressed for the time being.

Importantly, the faint stars do not noticeably exhibit the typical effects of crowding at the faint end (*e.g.*, Feinstein et al. 2019, Figure 5). In aperture photometry pipelines, stars in very crowded regions typically have their flux overestimated relative to the actual number of photons that fall on silicon. This leads to systematic underestimates of the uncertainty in the relative fluxes, as well as "flux contamination" (the reduction in amplitude of say, transit signals, due to diluting flux from neighbor stars). Our method to work around this problem – using the catalog magnitudes to predict the reference flux values, and measuring deviations from these reference fluxes on the subtracted images – seems to be performing well.

Expected RMS vs magnitude. —The noise model shown in Figure 14 is quite similar to that of Sullivan et al. (2015), save for two changes. The first change is that the effective area of the telescope is updated to be 86.6 cm^2, per the measurements by Vanderspek et al. (2018).

The second change is that we have explicitly included the estimated noise contribution from unresolved faint stars. The brightness of the diffuse sky is dominated by different sources at different wavelengths. For instance, the CMB is most important in the microwave, and thermal radiation from dust grains in the solar system (zodiacal light) is dominant in the far infra-red (Leinert et al. 1998). In the TESS-band, both zodiacal light and faint stars can play a role, depending on the line of sight under consideration. The zodiacal light is brightest near the ecliptic plane, and the faint star background is brightest near the galactic plane (and towards the galactic center). When performing pre-launch noise estimates, Winn (2013) estimated the photon-counts from each component. His zodiacal light model was presented by Sul-

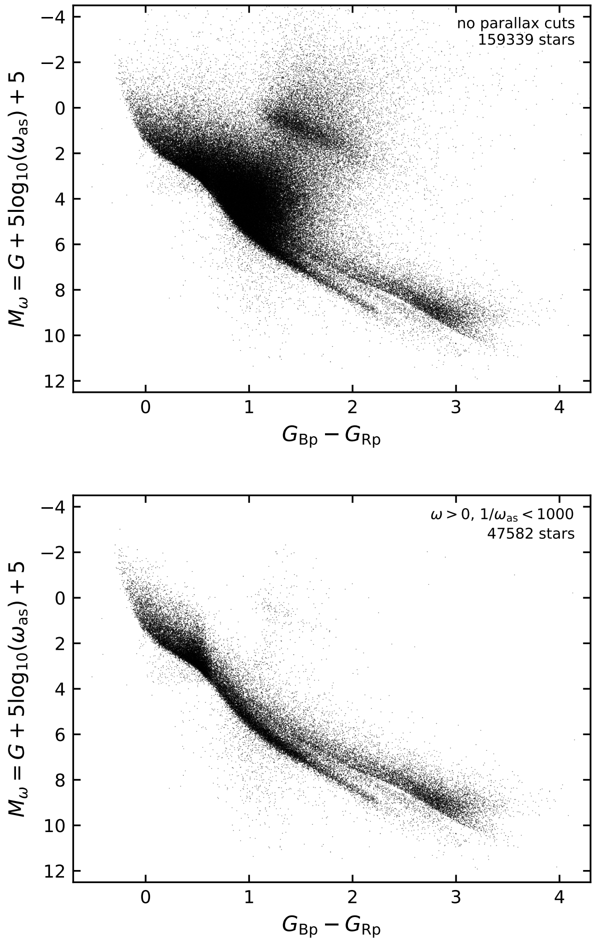

Figure 12. *Top.* HR diagram of CDIPS stars on silicon in this data release. *Bottom.* HR diagram of close CDIPS stars on silicon. The wedge separating the pre-MS sample from the MS stars was discussed by Zari et al. (2018), who introduced it in order to avoid contamination by photometric binaries.

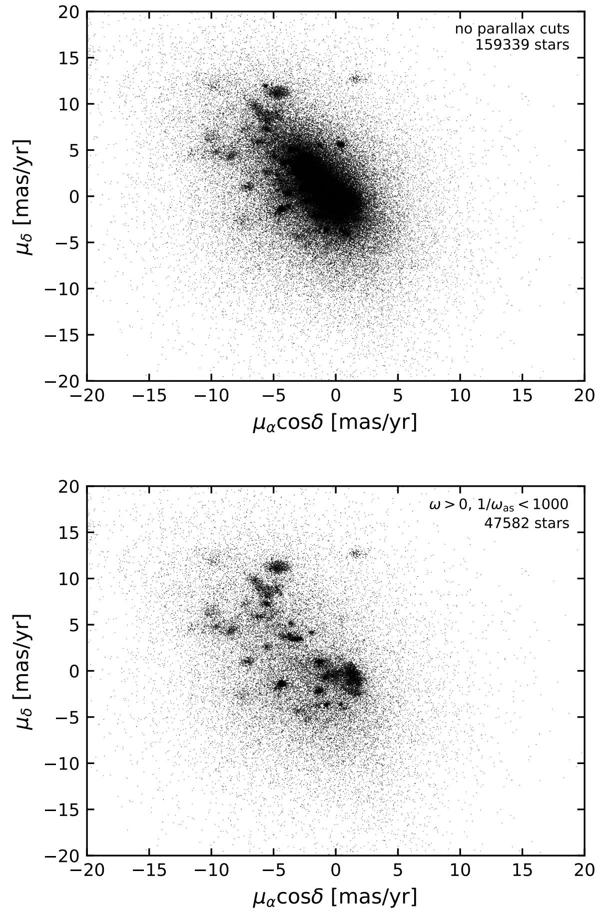

Figure 13. *Top.* Proper motions of CDIPS stars on silicon in this data release. Many of the stars in the central "blob" are possible field-contaminants. *Bottom.* Proper motions of close CDIPS stars on silicon.

livan et al. (2015), but the faint star component was not emphasized since the Sullivan simulations were performed away from the galactic plane.

The diffuse sky model we have used for Figure 14 is adopted explicitly because most of our target stars are near the galactic plane. Stars are judged to be "unresolved" and part of the background if their surface density exceeds the angular resolution of the telescope. TESS has an angular resolution of $\Delta\theta \sim 1'$, set by a combination of the pixel size as well as the typical stellar FWHM. Sources with sky surface density exceeding $\Delta\theta^{-2}$ therefore contribute to the background.

The relevant quantity needed to calculate the integrated photon counts from faint sources is $N(< m, l, b)$ — the number of stars per square arcsecond brighter than magnitude m, along a line of sight with galactic longitude and latitude (l, b). To calculate this surface density, Winn (2013) queried the Besançon model (Robin et al. 2003) along a grid of galactic sight-lines, and then converted to I-band surface brightnesses. Fitting a smooth function to the results, Winn (2013) found

$$I \text{ mag arcsec}^{-2} = a_0 + a_1 \left(\frac{|b|}{40^\circ}\right) + a_2 \left(\frac{|l|}{180^\circ}\right)^{a_3}, \qquad (10)$$

where the galactic longitude l is measured from -180° to 180°, and the empirical coefficients were found to be $a_0 = 18.9733$, $a_1 = 8.833$, $a_2 = 4.007$, and $a_3 = 0.805$. This fit was cautioned to be *very approximate*. It is sensitive to the threshold used to select "unresolved" stars, and likely no more accurate than 0.5 mag on average. In regions with exceptionally high extinction (*e.g.*, star forming regions) it is expected to systematic underestimate the background brightness by an even larger degree. Nonetheless, this model for the diffuse sky background seems to agree reasonably well with the observed trend of standard deviation versus stellar magnitude.

ACF statistics —Beyond the white noise properties of the light curves, the red noise properties are also important. In Figure 15 we show the average autocorrelation of the raw and

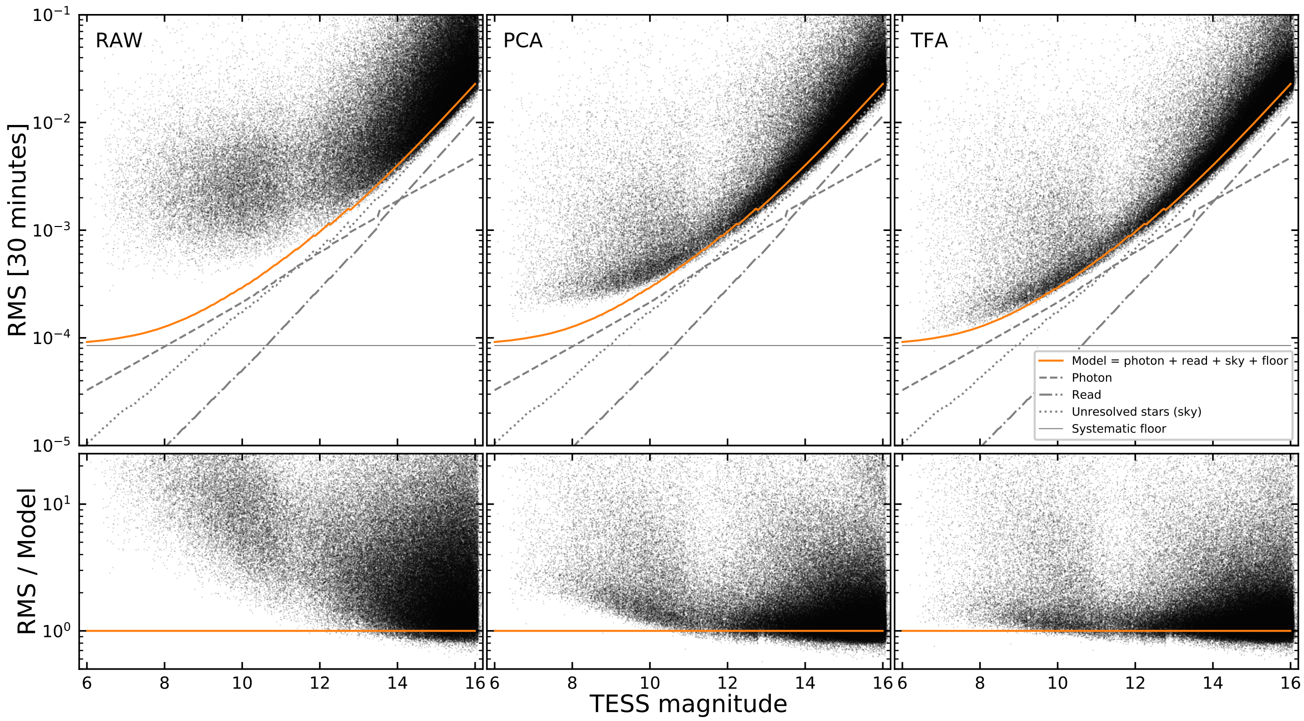

Figure 14. Standard deviation of CDIPS light curves as a function of catalog TESS-band magnitude. Left-to-right are raw, PCA-detrended, and TFA-detrended light curves. Black points correspond to the minimum RMS across the available three aperture sizes. The model (orange and gray lines) assumes aperture sizes reported by Sullivan et al. (2015), and the effective area from Vanderspek et al. (2018). The noise from unresolved background stars (dotted gray line) is a function of galactic latitude, and dominates over zodiacal light for faint stars near the galactic plane; the line shown assumes a sight-line towards the center of Sector 6, Camera 1 (further details are in § 4.1.3).

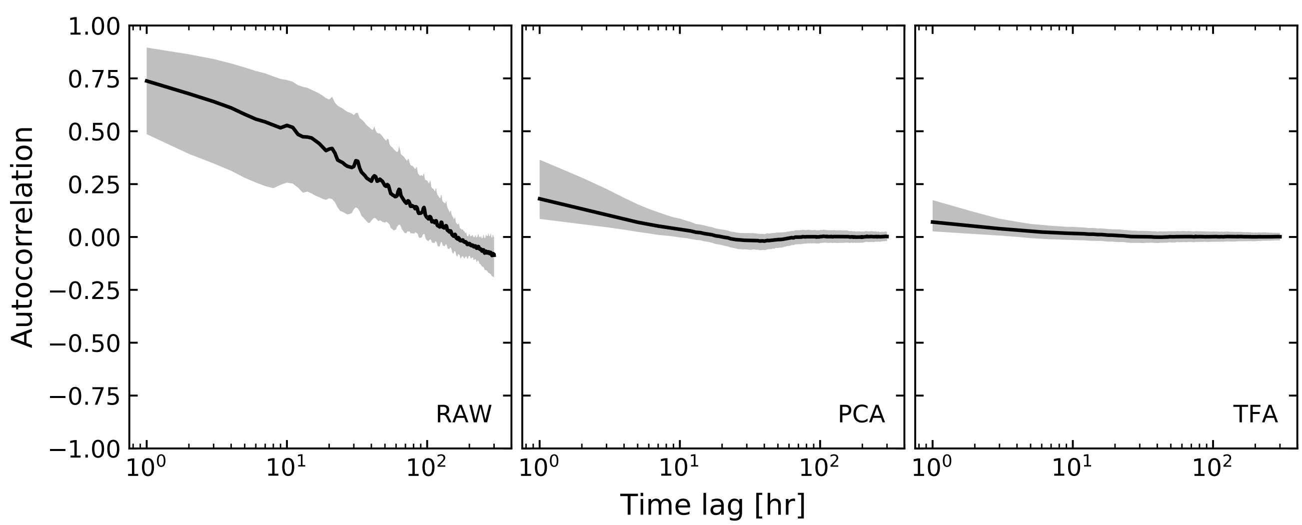

Figure 15. Average autocorrelations of 10^4 randomly selected light curves from each sector. The medians at time lags spaced by one hour are shown for raw light curves (column name `IRM`) in gray, and for TFA-detrended light curves in black. The gray bands display the 25$^{\text{th}}$ to 75$^{\text{th}}$ percentile range for each type of light curve.

TFA-filtered light curves. The raw light curves have substantial systematic red noise; their average autocorrelation, evaluated at various time lags, is positive. The TFA-filtered light curves are much closer to gaussian – the average autocorrelation between any two points in these light curves is close to zero. However, the average TFA-filtered light curve is not *exactly* gaussian. At time lags of a few hours or less, there is some excess power. This suggests that additional detrend-

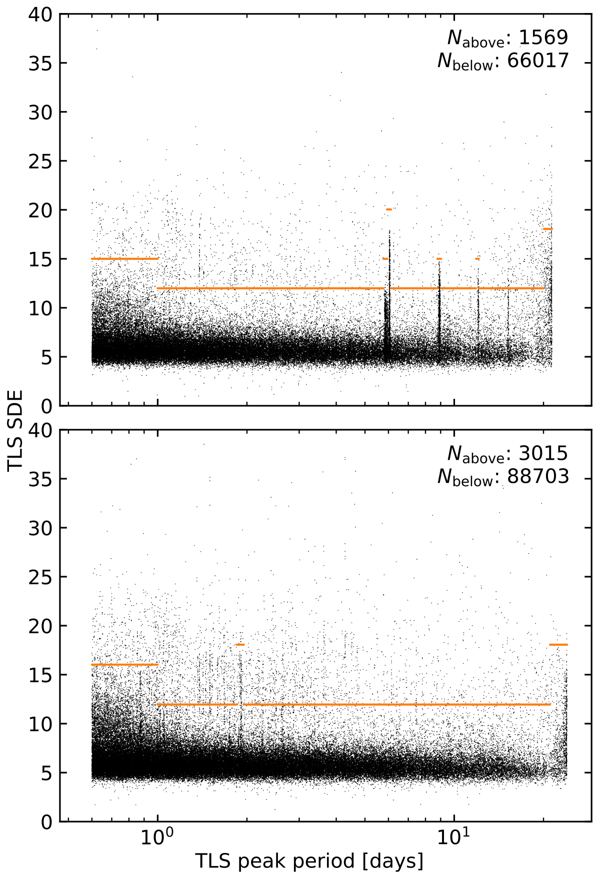

Figure 16. Transiting planet search periodogram summary. Each point represents the TLS periodogram peak signal detection efficiency (SDE) and corresponding period from one light curve. Significance thresholds (orange lines) are empirically defined for the transiting planet search described in § 4.2. *Top.* Sector 6. *Bottom.* Sector 7.

ing may be necessary to maximize the effectiveness of planet searches (or the discovery of any other signals via matched-filter techniques).

4.2. *A Few Objects of Interest*

Below, we provide a few examples of how one might go about identifying pulsating stars, strong stellar rotators, eclipsing binaries, and transiting planets in the light curves.

Sinusoidal signals—For pulsating stars and strong stellar rotators, the Lomb-Scargle periodogram (CITE) is a tried-and-true discovery method.

We measured rotation periods for stars in X, Y clean cluster.

A figure is given here.

To identify an initial set of transiting planets, strong stellar rotators, and eclipsing binaries, we performed a few steps of post-processing on the light curves. This processing was done independently of the data release, since it was quite specific to our own scientific interests. The code used to do this additional processing is also available online[8].

5. DISCUSSION

We remind the reader that our goal in creating our target catalog was completeness, rather than fidelity. The Dias et al. (2014) stars in particular were included if they were listed with membership probability exceeding 50%. To create cleaner sub-samples, we advise use of the `CDEXTCAT` header keyword, which can be used to merge against the original source catalog to obtain the membership probabilities claimed by the original catalog.

Despite our goal of completeness, some clusters may still be missing members – the census of nearby coeval stellar populations is still very much in flux (*e.g.*, Sim et al. 2019; Kounkel & Covey 2019). One particular area for concern in the present work is sub-clusters of broader star-forming complexes, for instance the Orion Nebula (M42; NGC 1976). While it was observed by TESS in Sector 6, searching our light curves for stars within 40 arcminutes, and 100 parsecs of the Orion Nebula's center yielded only 180 light curves, with 85 labelled members. The Orion Nebula has thousands of known members (Jones & Walker 1988). However, due to a combination of the spatially distributed nature of the broader complex, as well as differential extinction, the automated methods of the Kharchenko et al. (2013) and Dias et al. (2014) assigned the Orion Nebula only 44 and 326 members, respectively. Cantat-Gaudin et al. (2018) explicitly excluded young star-forming regions from their search, since the underlying assumption of their clustering method (uniformity in the field star distribution) breaks down in highly clustered star-forming regions.

6. CONCLUSION

In this study, we first collected an all-sky sample of about one million candidate young stars brighter than $G_{\rm Rp}$ of 16. 82% of the target star sample resides in a "cluster", a generic term we adopted for open clusters, stellar associations, and moving groups. The remaining stars have photometric or astrometric indications of their youth, and either reside on the pre-main-sequence or upper main sequence.

We then reduced TESS full frame images taken over the course of about two months (Sectors 6 and 7; the first fields very close to the galactic plane). We performed difference imaging in order to minimize the complications of the crowded background. Using forced aperture photometry, we made light curves for all stars brighter than $G_{\rm Rp}$ of 13, and went three magnitudes deeper for our target star sample. This yielded 159,343 light curves of candidate young stars across 596 distinct clusters.

The light curves seem to be limited in precision by photon noise at the bright end, and by unresolved background stars at the faint end. Though the raw light curves show significant red noise, decorrelating against a set of template stars led to

[8] github.com/lgbouma/cdips, commit `348eea7`.

an ensemble of light curves with very nearly gaussian noise properties.

An initial search for objects that were photometrically variable returned pulsating stars, eclipsing binaries, and planet candidates. A detailed planet search is the subject of ongoing work. An of the stars in the sample must be understood to be *candidate* cluster stars, and the possibility of photometric blends is important to rule out in subsequent vetting efforts of any variable object.

As a brief exploration of the photometric variability in the dataset, we looked at the age-evolution of extremely high-SNR variability, which seems to show some evolution in the space of LS peak-period versus color (Figure **??**), but interpreting this result will require a more careful exploration. In searching for transiting planets, we found many eclipsing binaries, as well as stellar rotation and pulsation signals.

Those who wish to explore the time-evolution in these and any other systems are invited to explore the data at archive.stsci.edu/prepds/cdips.

L.G.B. gladly acknowledges helpful discussions with C Huang, M Soares-Furtado, ..., and is grateful to the people who have turned TESS from an idea into reality. J.N.W. thanks ... This paper includes data collected by the TESS mission, which are publicly available from the Mikulski Archive for Space Telescopes (MAST). Funding for the TESS mission is provided by NASA's Science Mission directorate. This research has 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 work made use of NASA's Astrophysics Data System Bibliographic Services. This research has made use of the VizieR catalogue access tool, CDS, Strasbourg, France. The original description of the VizieR service was published in A&AS 143, 23. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. The Digitized Sky Surveys were produced at the Space Telescope Science Institute under U.S. Government grant NAG W-2166. The images of these surveys are based on photographic data obtained using the Oschin Schmidt Telescope on Palomar Mountain and the UK Schmidt Telescope. The plates were processed into the present compressed digital form with the permission of these institutions.

Facility: 2MASS (Skrutskie et al. 2006), Gaia (Gaia Collaboration et al. 2016, 2018b), TESS (Ricker et al. 2015), UCAC4 (Zacharias et al. 2013)

Table 2. Counts of light curves per cluster in Sector 6, sorted in descending order.

Name	N_{lc}	Description
Platais_5	7074	=,m,o,
Platais_6	7016	=,m,o,
Mamajek_3	3498	=,m,o,
Collinder_70	1576	=,a,o,
Trumpler_5	1492	=,,,var
Collinder_69	1192	=,,,
Collinder_110	1110	=,,,
ASCC_21	1094	=,a,c,
Collinder_121	986	=,a,o,
ASCC_19	925	=,,,
NGC_2287	869	=,,,
NGC_2232	760	=,,,
ASCC_20	756	=,,,
NGC_2141	730	=,,,
NGC_2112	673	=,,,
ASCC_16	661	=,,,ass
NGC_2194	656	=,,,
NGC_2301	586	=,,,
Collinder_65	575	=,,,
ASCC_28	533	=,,,

NOTE— Table 2 is published in its entirety in a machine-readable format. A portion is shown here for guidance regarding its form and content. Names are matched against Kharchenko et al. (2013) as described in Appendix B, and N_{lc} is the number of light curves associated with the cluster from this data release. The description column matches Kharchenko et al. (2013), and is in the format "a,b,c,d". Meanings of displayed symbols are as follows. a: "=" = cluster parameters were determined, "&" = duplicated/coincides with other cluster; b: "blank" = open cluster, "a" = association, "m" = moving group, "n" = nebulosity; c: "o" = object, "c" = candidate; d: "ass" = stellar association, "var" = clusters with variable extinction.

Software: `astrobase` (Bhatti et al. 2018), `astropy` (Collaboration et al. 2018), `astroquery` (Ginsburg et al. 2018), `BATMAN` (Kreidberg 2015), `corner` (Foreman-Mackey 2016), `emcee` (Foreman-Mackey et al. 2013), `fitsh` (Pál 2012), `IPython` (Pérez & Granger 2007), `matplotlib` (Hunter 2007), `numpy` (Walt et al. 2011), `pandas` (McKinney 2010), `pyGAM` (Servén et al. 2018) `psycopg2` (initd.org/psycopg) `scipy` (Jones et al. 2001), `scikit-learn` (Pedregosa et al. 2011), `TagSpaces` (tagspaces.org), `tesscut` (Brasseur et al. 2019), `VARTOOLS` (Hartman & Bakos 2016) `wotan` (Hippke et al. 2019),

Table 3. Counts of light curves per cluster in Sector 7, sorted in descending order.

Name	N_{lc}	Description
Collinder_173	13009	&,a,o,ass
ASCC_33	2826	=,n,o,
NGC_2437	2365	=,,,
NGC_2477	2158	=,,,
NGC_2546	1616	=,,,
NGC_2451A	1486	=,,,
NGC_2451B	1288	=,,,
NGC_2516	1278	=,,,
NGC_2323	1239	=,,,
NGC_2447	1192	=,,,
Collinder_132	1074	=,,,
ASCC_32	995	=,,,
Collinder_121	976	=,a,o,
NGC_2360	894	=,,,
NGC_2287	879	=,,,
NGC_2506	877	=,,,
NGC_2548	807	=,,,
NGC_2539	726	=,,,
Alessi_21	704	=,,,
Melotte_71	630	=,,,

NOTE— Table 3 is published in its entirety in a machine-readable format. A portion is shown here for guidance regarding its form and content. Format is same as Table 2.

REFERENCES

Aigrain, S., Parviainen, H., Roberts, S., Reece, S., & Evans, T. 2017, Monthly Notices of the Royal Astronomical Society, 471, 759

Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325

Bakos, G. A. 2018, in Handbook of Exoplanets, ed. H. J. Deeg & J. A. Belmonte (Springer International Publishing), 1

Bakos, G. A., Torres, G., Pál, A., et al. 2010, The Astrophysical Journal, 710, 1724

Barnes, S. A., Weingrill, J., Granzer, T., Spada, F., & Strassmeier, K. G. 2015, Astronomy & Astrophysics, 583, A73, arXiv: 1511.00554

Bell, C. P. M., Murphy, S. J., & Mamajek, E. E. 2017, Monthly Notices of the Royal Astronomical Society, 468, 1198

Bhatti, W., Bouma, L. G., & Wallace, J. 2018, astrobase

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Bouma, L. G., Winn, J. N., Baxter, C., et al. 2019, The Astronomical Journal, 157, 217

Bramich, D. M. 2008, Monthly Notices of the Royal Astronomical Society, 386, L77

Brasseur, C. E., Phillip, C., Fleming, S. W., Mullally, S. E., & White, R. L. 2019, Astrophysics Source Code Library, ascl:1905.007

Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, Astronomy & Astrophysics, 618, A93

Cantat-Gaudin, T., Jordi, C., Wright, N. J., et al. 2019a, Astronomy & Astrophysics, 626, A17

Cantat-Gaudin, T., Krone-Martins, A., Sedaghat, N., et al. 2019b, Astronomy and Astrophysics, 624, A126

Carrasco, J. M., Evans, D. W., Montegriffo, P., et al. 2016, Astronomy and Astrophysics, 595, A7

Clarke, B. D., Caldwell, D. A., Quintana, E. V., et al. 2017, Kepler Science Document, 5

Collaboration, T. A., Price-Whelan, A. M., Sipöcz, B. M., et al. 2018, arXiv:1801.02634 [astro-ph], arXiv: 1801.02634

Curtis, J. L., Agüeros, M. A., Mamajek, E. E., Wright, J. T., & Cummings, J. D. 2019, arXiv:1905.10588 [astro-ph], arXiv: 1905.10588

David, T., Hillenbrand, L., & Petigura, E. 2016, Nature, 534, 658

Dias, W. S., Monteiro, H., Caetano, T. C., et al. 2014, Astronomy and Astrophysics, 564, A79

Dierckx, P. 1996, Curve and surface fitting with splines, repr edn., Monographs on Numerical Analysis (Oxford: Clarendon Press), oCLC: 245719230

Eastman, J., Siverd, R., & Gaudi, B. S. 2010, Publications of the Astronomical Society of the Pacific, 122, 935, arXiv: 1005.4415

Feinstein, A. D., Montet, B. T., Foreman-Mackey, D., et al. 2019, arXiv:1903.09152 [astro-ph], arXiv: 1903.09152

Fischer, D. A., & Valenti, J. 2005, The Astrophysical Journal, 622, 1102

Foreman-Mackey, D. 2016, The Journal of Open Source Software, 24

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661

Froebrich, D., Scholz, A., & Raftery, C. L. 2007, Monthly Notices of the Royal Astronomical Society, 374, 399

Gagné, J., & Faherty, J. K. 2018, The Astrophysical Journal, 862, 138

Gagné, J., Roy-Loubier, O., Faherty, J. K., Doyon, R., & Malo, L. 2018a, The Astrophysical Journal, 860, 43

Gagné, J., Mamajek, E. E., Malo, L., et al. 2018b, The Astrophysical Journal, 856, 23

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, Astronomy and Astrophysics, 595, A1

Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, Astronomy and Astrophysics, 616, A10

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, Astronomy and Astrophysics, 616, A1

Ginsburg, A., Sipocz, B., Madhura Parikh, et al. 2018, Astropy/Astroquery: V0.3.7 Release

Hartman, J. D., & Bakos, G. A. 2016, Astronomy and Computing, 17, 1

Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, arXiv:1906.00966 [astro-ph], arXiv: 1906.00966

Huang, C. X., Penev, K., Hartman, J. D., et al. 2015, Monthly Notices of the Royal Astronomical Society, 454, 4159

Huang, C. X., Burt, J., Vanderburg, A., et al. 2018, The Astrophysical Journal, 868, L39

Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90

Ivezić, Z., Connelly, A. J., VanderPlas, J. T., & Gray, A. 2014, Statistics

Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, The Astrophysical Journal Letters, 713, L87

Jones, B. F., & Walker, M. F. 1988, The Astronomical Journal, 95, 1755

Jones, E., Oliphant, T., Peterson, P., et al. 2001, Open source scientific tools for Python

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2012, Astronomy and Astrophysics, 543, A156

—. 2013, Astronomy and Astrophysics, 558, A53

Kim, J. 2018, Querying Gaia for Wide Binary Companions to Exoplanet Hosts, Princeton Junior Thesis (Unpublished)

Kounkel, M., & Covey, K. 2019, arXiv:1907.07709 [astro-ph], arXiv: 1907.07709

Kovács, G., Bakos, G., & Noyes, R. W. 2005, Monthly Notices of the Royal Astronomical Society, 356, 557

Kovács, G., Zucker, S., & Mazeh, T. 2002, Astronomy and Astrophysics, 391, 369

Kraus, A. L., Cody, A. M., Covey, K. R., et al. 2015, The Astrophysical Journal, 807, 3

Kraus, A. L., Shkolnik, E. L., Allers, K. N., & Liu, M. C. 2014, The Astronomical Journal, 147, 146

Kreidberg, L. 2015, Publications of the Astronomical Society of the Pacific, 127, 1161

Krone-Martins, A., & Moitinho, A. 2014, Astronomy & Astrophysics, 561, A57

Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, The Astronomical Journal, 139, 1782

Leinert, C., Bowyer, S., Haikala, L. K., et al. 1998, Astronomy and Astrophysics Supplement Series, 127, 1

Luhman, K. L. 2012, Annual Review of Astronomy and Astrophysics, 50, 65

Majaess, D. 2013, Astrophysics and Space Science, 344, 175

Mann, A. W., Newton, E. R., Rizzuto, A. C., et al. 2016, AJ, 152, 61

Marrese, P. M., Marinoni, S., Fabrizio, M., & Altavilla, G. 2019, Astronomy & Astrophysics, 621, A144

McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51

Meibom, S., Barnes, S. A., Platais, I., et al. 2015, Nature, 517, 589

Meibom, S., & Mathieu, R. D. 2005, The Astrophysical Journal, 620, 970

Miller, J. P., Pennypacker, C. R., & White, G. L. 2008, Publications of the Astronomical Society of the Pacific, 120, 449

Milliman, K. E., Mathieu, R. D., Geller, A. M., et al. 2014, The Astronomical Journal, 148, 38

Oelkers, R. J., & Stassun, K. G. 2018, The Astronomical Journal, 156, 132

Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D., & Spergel, D. N. 2017, The Astronomical Journal, 153, 257

Pál, A. 2009, PhD thesis, arXiv: 0906.3486

—. 2012, MNRAS, 421, 1825

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. A., & Stobie, E. 2010, Astronomy and Astrophysics, 524, A42

Pérez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21

Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, The Astronomical Journal, 155, 89

Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, PASP, 118, 1407

Price-Whelan, A. M., & Goodman, J. 2018, The Astrophysical Journal, 867, 5

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Rizzuto, A. C., Ireland, M. J., & Robertson, J. G. 2011, Monthly Notices of the Royal Astronomical Society, 416, 3108

Robin, A. C., ReylÃľ, C., DerriÃĺre, S., & Picaud, S. 2003, Astronomy and Astrophysics, 409, 523

Röser, S., Demleitner, M., & Schilbach, E. 2010, The Astronomical Journal, 139, 2440

Röser, S., Schilbach, E., & Goldman, B. 2016, Astronomy & Astrophysics, 595, A22

Röser, S., Schilbach, E., Piskunov, A. E., Kharchenko, N. V., & Scholz, R.-D. 2011, Astronomy & Astrophysics, 531, A92

Saurin, T. A., Bica, E., & Bonatto, C. 2015, Monthly Notices of the Royal Astronomical Society, 448, 1687

Servén, D., Brummitt, C., & Abedi, H. 2018, dswah/pyGAM: v0.8.0

Sim, G., Lee, S. H., Ann, H. B., & Kim, S. 2019, arXiv e-prints, arXiv:1907.06872

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, The Astronomical Journal, 131, 1163

Skumanich, A. 1972, The Astrophysical Journal, 171, 565

Smith, J. C., Stumpe, M. C., Cleve, J. E. V., et al. 2012, Publications of the Astronomical Society of the Pacific, 124, 1000

Soares-Furtado, M., Hartman, J. D., Bakos, G. Ã., et al. 2017, Publications of the Astronomical Society of the Pacific, 129, 044501

Stassun, K. G., Feiden, G. A., & Torres, G. 2014, New Astronomy Reviews, 60, 1

Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, The Astronomical Journal, 156, 102

Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, arXiv:1905.10694 [astro-ph], arXiv: 1905.10694

Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77

Tamuz, O., Mazeh, T., & Zucker, S. 2005, Monthly Notices of the Royal Astronomical Society, 356, 1466

Tenenbaum, P., & Jenkins, J. 2018, TESS Science Data Products Description Document, EXP-TESS-ARC-ICD-0014 Rev D, https://archive.stsci.edu/missions/tess/doc/EXP-TESS-ARC-ICD-TM-0014.pdf

Vanderburg, A., Huang, C. X., Rodriguez, J. E., et al. 2019, arXiv:1905.05193 [astro-ph], arXiv: 1905.05193

Vanderspek, R., Doty, J., Fausnaugh, M., et al. 2018, TESS Science Document, 73

Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science & Engineering, 13, 22

Wang, D., Hogg, D. W., Foreman-Mackey, D., & SchÃűlkopf, B. 2017, arXiv:1710.02428 [astro-ph], arXiv: 1710.02428

Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, Astronomy and Astrophysics Supplement Series, 143, 9

Winn, J. N. 2013, TESS Science Memo No. 2, Version 1. Available upon request.

Zacharias, N., Finch, C. T., Girard, T. M., et al. 2013, The Astronomical Journal, 145, 44

Zari, E., Hashemi, H., Brown, A. G. A., Jardine, K., & de Zeeuw, P. T. 2018, Astronomy and Astrophysics, 620, A172

Zhang, M., Bakos, G. A., Penev, K., et al. 2016, Publications of the Astronomical Society of the Pacific, 128, 035001

Zuckerman, B., & Song, I. 2004, Annual Review of Astronomy and Astrophysics, 42, 685

APPENDIX

A. TIME SYSTEM & BARYCENTRIC CORRECTION

The time-stamps included with the calibrated TESS Full Frame Images produced by SPOC include a barycenteric correction at a single reference pixel given at the middle of every frame. The barycentric correction is at maximum 16 minutes, corresponding to points on the sky separated by 180 degrees. The angular distance from a TESS camera's center of field to the corners is $\approx$17 degrees, so naively one might incur at worst an error of $\approx$90 seconds on the time-stamps due to using a barycentric correction in a direction that is slightly wrong. Perhaps due to the lead author's obsession with getting time-stamps correct (Bouma et al. 2019), we perform our own barycentric correction using the appropriate sky coordinates for each light curve. We advise use of our `TMID_BJD` column, which gives the mid-time of each exposure in the $\mathrm{BJD_{TDB}}$ time system, which is the defacto standard in exoplanet and stellar astronomy (Eastman et al. 2010).

B. ASSIGNING UNIQUE NAMES TO EACH CLUSTER

In assigning a single unique cluster name to each star, we matched against the Kharchenko et al. (2013) name whenever possible, since this was the largest available catalog, and it also included homogeneous age determinations for many of the clusters. To find the matching name, in order of precedence we

1. Checked for direct string matches from Kharchenko et al. (2013) clusters with determined parameters;
2. Checked whether the SIMBAD online name resolving service (Wenger et al. 2000) had any direct string matches against Kharchenko et al. (2013) clusters with determined parameters;
3. Checked for string matches in the full Kharchenko et al. (2013) index (including clusters without determined parameters);
4. Searched for spatial matches between each star and cluster centers from Kharchenko et al. (2013) within 10 arcminutes. In cases with multiple cluster matches, we ignored candidate matches to avoid assigning incorrect names;
5. Checked the WEBDA double name list[9], and repeated Steps 1-4 with any matches.

A few edge-cases, including for instance sub-clusters larger star-forming complexes like in Sco-Cen or Collinder 33, were manually resolved to the extent feasible (Rizzuto et al. 2011 and Saurin et al. 2015 give detailed pictures of the complex morphologies that frequently arise in young star-forming regions).

The procedure described above failed to yield matches for a few of the infrared clusters identified by Majaess (2013) and included in the Dias et al. (2014) catalog. For these cases, we used the name given by Dias et al. (2014). The larger set of "FSR" infrared clusters from Froebrich et al. (2007) was incorporated to Kharchenko et al. (2013), and so did not present any complications.

The Hyades and a number of other nearby moving groups were also missed, since they were not in the Kharchenko et al. (2013) catalog. For moving groups not identified in Kharchenko et al. (2013), we adopted the constellation-based naming convention from Gagné et al. (2018b).

Finally, the procedure enumerated above did not yield matches for recently discovered clusters, such as the "RSG" clusters found by Röser et al. (2016) and the "Gulliver" clusters from Cantat-Gaudin et al. (2018). In these cases, we used the names given by the original authors.

[9] https://webda.physics.muni.cz/double_names.html