A search for young exoplanets in Sectors 1–5 of the *TESS* full-frame images

Matthew P. Battley [1,2]★ Don Pollacco[1,2] and David J. Armstrong [1,2]

[1]*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*
[2]*Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK*

ABSTRACT

Young (<1 Gyr) exoplanets represent a critically important area of exoplanet research, as they offer the opportunity to learn about the formation and early dynamic history of exoplanetary systems. However, finding young exoplanets is significantly complicated by the fast rotation and complex activity of their young host stars, which are often not well handled by state-of-the-art automatic pipelines. This work presents an alternative LOWESS-based pipeline focused specifically on detrending young stellar light curves from the 30-min cadence full-frame images produced by the *Transiting Exoplanet Survey Satellite (TESS)*, and includes improvements such as automatic peak cutting of stellar variability and interpolation over masked transits to improve periodogram visibility and returned transit shapes. This work presents the details of the developed pipeline, along with initial results from its application to young stars within stellar associations in Sectors 1–5 of the *TESS* data. While no new exoplanet candidate signals were found in this work, interesting results included the recovery of all known 2-min TOIs around young stars in Sectors 1–5 from 30-min data alone, the recovery of the young exoplanet DS Tuc Ab, a number of young eclipsing binaries and a wide array of interesting rotation. A sensitivity analysis was also undertaken for each star, showing how the recovery of injected planets varied with both depth and period for each individual target. Challenges for future searches for young exoplanets are discussed, the largest being stellar rotation with periods less than 1 d and a lack of a large sample of confirmed young stars.

Key words: techniques: photometric – planets and satellites: detection – planets and satellites: general – stars: rotation – open clusters and associations: general.

1 INTRODUCTION

Despite the wealth of over 4000 verified exoplanets known today[1] there are still many unanswered questions concerning the formation mechanisms and system evolution, which has led to the observed distribution of these objects. In particular, planetary migration, dynamical interactions with nearby stars and stellar evolution all can have major effects on the distribution and architecture of the final exoplanet population. This is especially important early in a planet's life (e.g. <1 Gyr), where phenomena such as accretion (Marley et al. 2007; Manara et al. 2019), ionizing radiation causing atmospheric loss (Baraffe et al. 2003; Owen 2019), and dynamical interactions with other forming planetary system bodies (Ida & Lin 2010; Schlichting, Sari & Yalinewich 2015) can significantly change the mass, radius, and orbital parameters of early exoplanets.

While traditional exoplanet studies have been biased towards older exoplanets due to their relatively quiet host stars, there is a strong case to be made for the search for younger exoplanets, where planets are undergoing the majority of their evolution (Adams & Laughlin 2006; Ida & Lin 2010; Spiegel & Burrows 2012; Ida, Lin & Nagasawa 2013). Discovering exoplanets <1 Gyr old will thus help to probe the main causes of planetary evolution, and help to fill a key gap in our knowledge of exoplanet history.

However, the host stars of young exoplanets provide significant challenges for discovery due to their typically increased activity, rotation rates, and relative proximity to neighbouring stars (e.g. Koeltzsch et al. 2009; Rivilla et al. 2015; Sergison 2015; Mascareño, Rebolo & Hernández 2016; Briceno et al. 2018; Cegla et al. 2019). The combination of these factors can result in periodic stellar variability that can be of the order of planetary signals in both period and intensity (Armstrong et al. 2015; Cody et al. 2018).

Despite these challenges, a small number of young exoplanets have been found. In line with early exoplanet discoveries, the first exoplanets around young stars were found using the radial-velocity

---

method, including four Hot Jupiters found in the Hyades (Sato et al. 2007; Quinn et al. 2014) and Praesepe (Quinn et al. 2012) open clusters. However, for planets less massive than Hot Jupiters, the radial-velocity method was found to be severely hampered by the inherent stellar variability and radial velocity jitter of the young host stars (Saar & Donahue 1997; Paulson, Cochran & Hatzes 2004; Brems et al. 2019). Recent methods such as those employed by Korhonen et al. (2015), Rajpaul et al. (2015) have pushed this limit down to approximately Neptune-sized planets for some solar-like stars, however, Earth-sized planets are currently still beyond the capabilities of these methods for young active stars.

The high photometric precision offered by the launch of the *Kepler* satellite proved crucial in increasing this sample of young exoplanets, especially when deliberately pointed at young open clusters in the *K2* extended mission. Most discoveries in this era were made by the *Zodiacal Exoplanets In Time* team, who found 17 planets between 2016 February and 2018 October (e.g. Mann et al. 2016; Rizzuto et al. 2017, 2018). Other interesting discoveries of young exoplanets in K2 have included K2-33b – a Neptune-sized planet in the 5–10 Myr Upper Scorpius stellar association (David et al. 2016), K2-136A c in the Hyades – the first Neptune-sized planet orbiting a binary system in an open cluster (Ciardi et al. 2018), and the closely packed system of four planets around a ∼23 Myr pre-main-sequence star in the Taurus–Auriga star-forming region (David et al. 2019a, b). The launch of the *TESS* is now ushering in a new era of exoplanet discovery, including the discovery of two new young exoplanets: DS Tuc Ab in the 45 Myr Tucana–Horologium association (Benatti et al. 2019; Newton et al. 2019), and the 24 Myr AU Mic b in the Beta Pictoris Moving Group (Plavchan et al., 2020). These recent new discoveries highlight the fact that it is now possible to find planets across a wide range of the early evolution of these systems, but considerably more planets are needed at these ages before reliable statistics can be generated.

A particularly promising place to look for these young exoplanets is within stellar associations. These are groups of gravitationally unbound stars, which have a common origin, so still retain a common proper motion across the sky. Their relatively diffuse nature compared to similarly young star clusters makes them far more suited to study with *TESS* given the relatively large pixels of the satellite (angular resolution ∼21 arcsec, Vanderspek et al. 2018). Stars in these associations share similar ages, positions, compositions and kinematics, meaning that precise stellar properties can be determined (Torres et al. 2006). This, in turn, provides a significant advantage for the precise determination of exoplanet characteristics. Furthermore, the ages of stars within these associations are typically very well constrained, which will allow a more detailed timeline of planetary evolution to be assembled. The most extensive census of 'bona-fide' members of these stellar associations was assembled by Gagné et al. (2018a) in the development of their BANYAN Σ Bayesian membership tool, but has recently been further expanded by works such as Gagné et al. (2018b) and Esplin & Luhman (2019) thanks to the release of Gaia DR2 (Gaia Collaboration et al. 2018). The extremely precise astrometric data from *Gaia* will doubtless allow for further expansions to the membership lists of these associations as the satellite's mission continues. Note that Bouma et al. (2019) recently assembled a larger list of young stars in clusters and associations (of which the BANYAN sample is a subset) by considering a wide array of sources in the literature; however, they admit that this sample was designed for 'completeness, not accuracy'. For the initial design of the pipeline constructed in this work, the homogeneous BANYAN selection criteria focused solely on stellar associations was preferred.

In order to discover any exoplanets within these associations, a key challenge is dissociating true transit signals from instrumental systematics or stellar 'noise' (activity or variability of the host star). While instrumental trends are commonly removed using techniques like cotrending basis vectors (Thompson et al. 2016), spacecraft pointing-based decorrelation (Vanderburg & Johnson 2014; Aigrain, Parviainen & Pope 2016) or subtracting trends shared by simultaneously observed nearby stars (Kovács, Bakos & Noyes 2005; Kim et al. 2009), the problem of dissociating stellar noise from transit signals remains a challenging one. A large array of different methods have been developed in an attempt to solve this problem. These methods can be largely separated into three major approaches: sliding filters (e.g. Savitzky & Golay 1964), fitting splines/polynomials (e.g. Vanderburg et al. 2016), and Gaussian processes (e.g. Aigrain et al. 2015; Gillen & Aigrain, in prep). An in-depth review and comparison of the most commonly used stellar activity detrending methods can be found in Hippke et al. (2019).

The earliest wide-field survey, HATNet (Bakos et al. 2004), used the Trend Fitting Algorithm (Kovács et al. 2005) and External Parameter Decorrelation (Bakos et al. 2010) to detrend its light curves, while WASP (Pollacco et al. 2006) shortly after preferred the `sysrem` algorithm of Tamuz, Mazeh & Zucker (2005) – which removes common time- and position-dependent trends by taking into consideration the weighted average magnitude residuals for all stars – coupled with a boxcar-smoothing technique to handle some of the more variable stars observed (Collier Cameron et al. 2006). More recent ground-based surveys such as KELT (Pepper et al. 2007) and NGTS (Wheatley et al. 2018) have developed these techniques further, and even introduced more advanced detrending methods such as Gaussian processes as they target more active stars (e.g. Costes et al. 2019; Gillen et al. 2019). However, the largest source of light curves in the search for exoplanets has been space-based surveys, the volume and variety of which has required the development of more complex and robust detrending techniques. The most prolific transiting exoplanet surveys thus far have been the *Kepler* (Borucki et al. 2010) and *K2* (Howell et al. 2014) missions and the ongoing *TESS* mission (Ricker et al. 2014), all of which share broadly similar detrending techniques. Overviews of the *Kepler*, *K2*, and *TESS* SPOC pipelines can be found in Jenkins et al. (2010), at https://keplerscience.arc.nasa.gov/k2-pipeline-relea se-notes.html and in Jenkins et al. (2016), respectively.

While these state-of-the-art detrending methods have been extremely effective at discovering new exoplanets around older stars, they can still struggle to disassociate true transiting signals from the complex activity of young host stars. This was particularly well illustrated by Hippke et al. (2019) in construction of their WOTAN tool, where only 35 per cent of $0.5R_{Jup}$ exoplanet signals injected into a sample of 316 young stars were recovered (irrespective of the detrending method chosen), compared to an almost 100 per cent recovery rate for similar planets around less noisy stars. Interestingly, however, each of the methods tested recovered slightly different populations of the injected signals, so combining all of the tested methods increased the percentage of recovered planets to 43.8 per cent. This discrepancy was further highlighted by Rodenbeck et al. (2018), who showed that different detrending methods resulted in quite different conclusions when considering the potential super-moon around *Kepler-1625 b*. It is thus clear that not only is there a need for new detrending methods specifically focused on young stars, but also that there is a distinct benefit of using multiple detrending methods on the same data set, thus providing dual motivations for the construction of an alternative detrending method in this work. The development and current

design of the resulting detrending pipeline used in this work is described in detail in Section 2.5.

The remainder of this paper is organized as follows. Section 2 discusses the young star target selection, observations, and the choice of an FFI pipeline, before describing the methods used to clean additional systematics from the *TESS* data, detrend stellar variability, search for transits, and inject model transits for a sensitivity analysis. Section 3 then discusses important results seen as a result of the developed detrending techniques, including the recovery of known young exoplanets, TOIs, and eclipsing binaries, followed by an overview of the sorts of rotation and activity seen in this sample and the results of the conducted sensitivity analysis. The implications of these results for the future of exoplanetary searches around young stars are then discussed in detail in Section 4, before a summary and conclusion in Section 5.

2 TARGET SELECTION, OBSERVATIONS, AND DETRENDING METHODS

2.1 Target selection

While stellar association membership is now being expanded on a cluster by cluster basis thanks to the increased astrometric precision of the *Gaia* satellite (Gaia Collaboration et al. 2016, 2018; see e.g. Kuhn et al. 2019; Damiani et al. 2019; Zari, Brown & De Zeeuw 2019), a key list of 'bona-fide' stellar association members were assembled by Gagné et al. (2018a) during the preparation of the BANYAN Σ Bayesian cluster membership tool. Gagné et al. (2018a) consider a star to be a 'bona-fide' member if it has galactic XYZ UVW values consistent with those known for a given stellar association and exhibits an independent sign of youth. Such signs can include mid-infrared excess, G – J versus GALEX NUV – G colours consistent with youth, X-ray emission with HR1 ≥ -0.15, lithium absorption above 100 mÅ or a compatible luminosity class (Gagné et al. 2018b). Combining Gagné et al.'s (2018a) initial census of bona-fide/high-probability stellar association members with new high-probability members added in the two following BANYAN Σ papers (Gagné & Faherty 2018; Gagné et al. 2018b) yielded a total of 2977 objects spread over the 27 nearest known young stellar associations. Gagné, Faherty & Mamajek (2018c) later expanded the BANYAN Σ tool to also include the two new Argus (Zuckerman 2018) and Volans-Carina associations using the same membership criteria, so these were also added to the initial target list, to give the final distribution of 3076 young stars in stellar associations illustrated in Fig. 1. For clarity, an association by association breakdown of targets in the original BANYAN sample and the final sample of the Sector 1–5 targets analysed in this work is presented in Table 1. This census remains the most extensive assemblage of bona-fide and high-probability stellar association members with a common membership criterion, and hence represents a valuable starting place for the search for exoplanets around young stars in this work.

The various BANYAN survey results were combined into a single tabular target list using the *TOPCAT* table handling software (Taylor 2005). This list was then cross-matched with version 8 of the *TESS* input catalogue (TIC) (Stassun et al. 2019) using a 3 arcsec radius. The Web *TESS* Viewing Tool[2] on the NASA *TESS* website was used to determine which sector each target would be observed in and in turn compile target lists for each individual sector. Of the

original 3076 objects, 1832 of them were forecast to be observed in *TESS*'s first year of observations, with the breakdown of sources to be viewed in each Southern Hemisphere sector outlined in Table 2. For this work, only those viewed in Sectors 1–5 were considered.

2.2 Observations

TESS is the successor to the highly successful *Kepler* space telescope, which is designed to survey >85 per cent of the sky to search for planets transiting bright host stars (Ricker et al. 2014). *TESS* launched on 2018 April 18 and is now is into year two of its primary mission, having completed observations of the 13 Southern Hemisphere sectors. Sectors 1–5 considered in this work were observed between 2018 July 25 and December 11 .

Two main data products are produced by the *TESS* mission: 2-min cadence light curves (detrended and raw) and 30-min cadence FFIs. 2-min cadence light curves are generated by Science Processing and Operations Center (SPOC) pipeline (Jenkins et al. 2016) for approximately 300 000 of the most promising stars (prioritized by the smallest transiting planets that can be detected) from the *TESS* Candidate Target List (CTL, Stassun et al. 2018, 2019), and are then made accessible to the public via the Mikulski Archive for Space Telescopes (MAST[3]). However, while these primary data products are very powerful for the main *TESS* mission, because of the limited data transfer rates of the *TESS* primary mission far from all of the stars in *TESS*'s field of view will have 2-min light curves generated. Instead, light curves from these stars can be retrieved from the 30-min cadence full-frame images (FFIs). In addition, because of the increased activity and rotation rate of young stars discussed above, the standard detrending methods built for the 2min light curves have difficulty flattening the light curves for transit searches, and may even introduce confusing additional artefacts as a result of detrending (Hippke et al. 2019). Because of the greater coverage offered by the 30-min cadence data and concerns about detrending artefacts, the 30-min cadence data were thus chosen for the initial transit search, with the simple aperture photometry (SAP) 2-min cadence light curves (where available) used as a secondary check.

An additional challenge common to both the 2- and 30-min cadence data are the reasonably large pixels of the *TESS* focal plane (Vanderspek et al. 2018), which leads to the significant blending of stars in crowded regions. This precludes the use of *TESS* photometry alone for confirming transits in the dense centres of young clusters without confirmation from higher resolution instruments such as the *SPITZER* Space Telescope (Werner et al. 2004) or the cameras of the *Next Generation Transit Survey* (*NGTS*; Wheatley et al. 2013). The power of *NGTS*'s comparative resolution for identifying the true photometric source for *TESS* objects is well illustrated by Jackman et al. (2019) in the discovery of NGTS-7ab.

2.3 Choice of FFI pipeline

A number of methods currently exist for extracting light curves from the *TESS* FFIs. The simplest approach is to perform SAP on the raw FFIs by overlaying an aperture over a cut-out around the object of interest, and summing up the flux under the target aperture for each cadence. This can be performed easily using the prebuilt *lightkurve*[4] `Python` package (Lightkurve Collaboration et al. 2018), and observing the resulting target-pixel-file provides an instructive look

---

[2] https://heasarc.gsfc.nasa.gov/cgi-bin/tess/webtess/wtv.py

[3] https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html
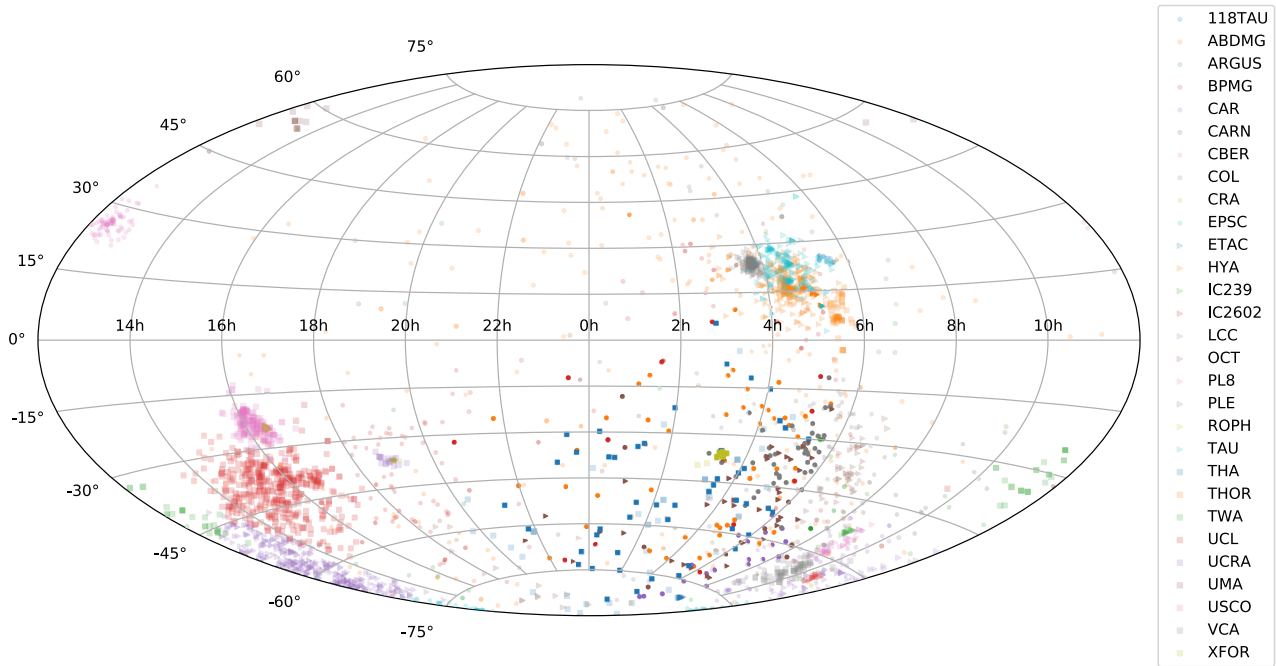[4] https://github.com/KeplerGO/lightkurve

Figure 1. Equatorial sky distribution of initial young star sample, based on the bona-fide and high probability members of the 29 nearest stellar associations built into the BANYAN Σ tool. The entire sample is shown in the background with the 256 Sector 1–5 targets analysed in this work presented in solid colour. Abbreviations after those chosen in Gagné et al. (2018a).

at the area around the star of interest. However, the simplistic nature of this analysis yields relatively noisy light curves uncorrected for issues such as spacecraft pointing, jitter, and localized scattered light. A greatly improved SAP pipeline that accounts for these issues has been built in `Python` by Feinstein et al. (2019). Christened *eleanor*,[5] this package performs background subtraction, removal of spacecraft systematics such as jitter and pointing drift, and aperture/point spread function (PSF) photometry. In addition, it provides tools to complete further systematics-removal via principal component analysis or psf-modelling. Light curves from *eleanor* will eventually be hosted on *MAST*, but at the time of writing is only available as an open-source tool designed to work for *TESS* Sectors 1 and 2.

An alternative difference imaging approach for 30-min light-curve generation has been pursued by Oelkers & Stassun (2018). Attempting to overcome the challenges posed to aperture photometry by *TESS*'s large pixel sizes, this approach extracts light curves via difference imaging analysis (DIA). In this method, one frame is blurred to the seeing conditions of the next before the two are subtracted from each other in order to retain only the variation in flux of stars between frames. Highlighting stars of interest within this process aids removal of contaminating stars in crowded regions and hence improves the light-curve extraction compared to standard aperture photometry methods. A similar technique has been used in ground-based surveys such as the Kilodegree Extremely Little Telescope (Pepper et al. 2007; Siverd et al. 2012), though with a simpler Gaussian kernel than the Dirac δ-function kernel used here. A full description of the DIA technique as applied to *TESS* light curves in Sectors 1 and beyond can be found in Oelkers & Stassun (2018), Oelkers & Stassun (2019). Light curves extracted via this pathway are accessible from the Filtergraph data visualization

service.[6] It is further worth noting that as this document was being prepared Bouma et al. (2019) released an alternative difference imaging pipeline for extraction of 30-min FFI light-curve images which they applied to Sectors 6 and 7, however, these sectors are outside the scope of this work.

One further pipeline that can be used to extract 30-minute light curves from the raw FFIs is the TASOC pipeline, under development by the *TESS Asteroseismic Science Consortium* (*TASC*). Heavily based on the K2P^2 pipeline (Lund et al. 2015), this pipeline aims to supply *TESS* photometry data for use in asteroseismology. Data from this source can be accessed online[7] after joining the consortium. So far raw light curves extracted from the FFIs are available for Sectors 1 and 2 (Handberg & Lund 2019), while the open-source code used to extract these images from the FFIs can be found on github.[8] However, one must be somewhat careful when using this source for transit searches, as the primary goal of this group is asteroseismology, so at times light curves may have been extracted in a way that prioritizes the variability of the stellar signal over potential transit-like events. While this pipeline is still currently under development, a useful overview of its design and aims has been written by Lund et al. (2017).

Given the increased availability and comparatively clean light curves provided by the DIA pipeline of Oelkers and Stassun (Oelkers & Stassun 2019, 2018), this pipeline was chosen to extract light curves from the *TESS* FFIs in this work, resulting in light curves for 256 individual objects.

---

[5]https://github.com/afeinstein20/eleanor

[6]https://filtergraph.com/tess_ffi
[7]http://tasoc.dk
[8]https://github.com/tasoc

Table 1. Breakdown of association distribution for the complete BANYAN sample (centre panel) and final set of targets analysed in this work (right-hand panel). The final targets consist of those targets in Sectors 1–5 with available 30-min cadence light curves from the Oelkers & Stassun (2018) pipeline, as discussed in Section 2.3. Association abbreviations are based on those presented in Gagné et al. (2018a) with the addition of ARGUS for the Argus association and VCA for the Volans-Carina Association.

Association	BANYAN targets	S1–S5 analysed targets
118TAU	12	0
ABDMG	298	55
ARGUS	40	0
BPMG	135	13
CAR	85	19
CARN	110	18
CBER	76	0
COL	107	43
CRA	14	0
EPSC	42	0
ETAC	16	0
HYA	241	7
IC2391	14	0
IC2602	16	0
LCC	310	0
OCT	101	40
PL8	32	0
PLE	204	0
ROPH	182	0
TAU	178	1
THA	92	49
THOR	46	0
TWA	50	0
UCL	410	0
UCRA	16	0
UMA	13	0
USCO	161	0
VCA	59	0
XFOR	16	11
Total	3076	256

Table 2. Overview of BANYAN bona-fide members of stellar associations viewed in the first year of *TESS* observations. Note that sources in the overlapping regions of the *TESS* sectors, such as those in the continuous viewing zone, are counted multiple times.

Sector	Number of sources observed
1	118
2	130
3	124
4	186
5	336
6	258
7	163
8	166
9	272
10	399
11	750
12	396
13	200

2.4 Removal of additional systematics from 30-min light curves

Unlike the 2-min PDCSAP light curves retrieved from MAST, the 30-min light curves supplied by Oelkers & Stassun (2019) have not undergone the in-depth quality analysis completed by the SPOC pipeline, and as such still include some less-trustworthy epochs of increased pointing jitter, regular spacecraft momentum dumps, and known data anomalies. It was thus necessary to remove these systematics before activity detrending and transit searches could begin.

The first step undertaken was to cut any epochs where fine pointing was known to have been lost, or other spacecraft anomalies affected the data. This was achieved by consulting the *TESS* data release notes[9] for each sector. This primarily affected Sectors 1, 3 and 4. In Sector 1, a period of anomalously high pointing jitter was seen between approximately TJD 1347–1349 due to problems with the fine-pointing calibration. This was observed to be particularly bad between TJD 1348-1349.29, so all epochs between these times were masked from the analysis. Similarly, in Sector 3, a few experiments on the attitude control system were undertaken by the *TESS* team, dramatically increasing the scatter at these times. As a result, only data between TJD 1385.8966-1395.4800 in orbit 13 and TJD 1396.6050-1406.2925 in orbit 14 are scientifically useful, and all data at other epochs in this sector was cut. Sector 4, on the other hand, was plagued by an instrument anomaly between TJD 1418.54 and TJD 1421.21, where communication was lost between the instrument and satellite. As a result, no data or telemetry was collected for this period, and some systematic trends were introduced following activation of the on-board heaters. Additional strong glints between TJD 1422.2297-1423.5020 (orbit 15) and TJD 1436.1047-1436.8353 (orbit 16) also plague some of the light curves in Sector 4 (particularly those on camera 4), however, the amplitude and duration of these appear to vary between different targets, so may be better examined on a case by case basis.

Following the removal of these sections of unreliable data, a more automated method was required to identify and remove additional epochs of increased spacecraft scatter on an epoch-by-epoch time-frame, such as those around the regular spacecraft momentum dumps. This was completed by generating scattering quality masks based on the engineering 'quaternion' data, using a similar method to that described by Vanderburg et al. (2019) to prepare Sector 3 30-min cadence data in the discovery of multiple super-Earths around HR 858. The *TESS* quaternion data[10] consist of 2s-cadence time-series data for each sector, describing attitude changes in three primary vectors (Q1, Q2, Q3) based on deviations from a selection of local guide stars. This provides a sector-specific overview of the spacecraft attitude and thus allows the generation of scatter-based quality masks for all targets in the *TESS* aperture. An example of this data (both raw and binned into 30-min bins) can be seen in Fig. 2. To identify epochs with excessive scatter, the standard deviation for each vector was calculated, and any epoch with pointing scatter ≥5 standard deviations from the mean was flagged for removal. By combining the results from all three vectors into a single mask, all epochs within 0.01 d of these points were removed from the data set. This step efficiently removed all spurious signals relating to momentum dumps, as well as any remaining short periods of overly large scatter that were not picked up in the initial wider cuts.

[9] Available at https://archive.stsci.edu/tess/tess_drn.html
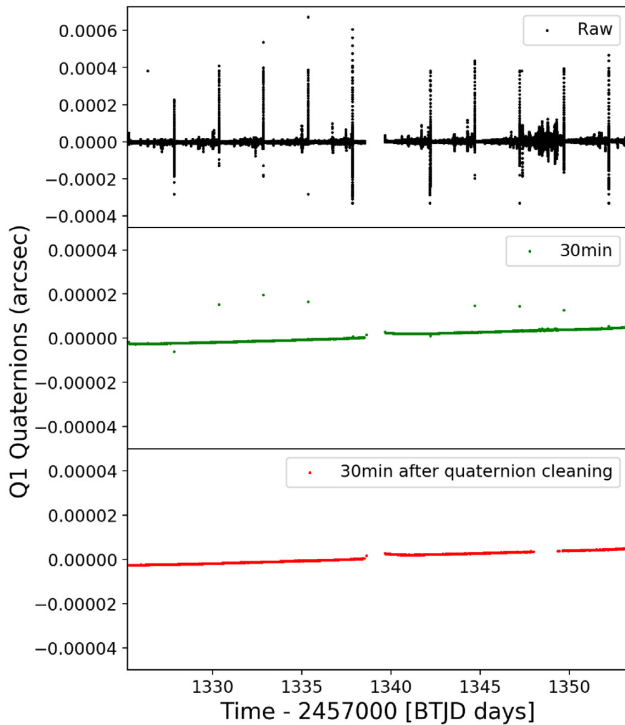[10] Available at https://archive.stsci.edu/missions/tess/engineering/

Figure 2. Q1 Engineering quaternion data for Sector 1, Camera 1 showing a clear increase in scatter around the 2.5-d momentum dumps and the loss of fine pointing between TJD 1347-1349. Top panel: Raw 2s Engineering Quaternion data. Middle panel: quaternion data binned into 30-min bins to show the typical number of affected data points. Bottom panel: final quaternions after the removal of known periods of increased scatter and quaternion-based cleaning.

For sources of particular interest, which also possess *TESS* 2min quality flags, it may also be advisable to implement the quality flags from the SPOC pipeline to the 30-min data [perhaps in a similar manner to the conservative approach taken by Bouma et al. (2019) in building their difference-imaging extraction pipeline for Sectors 6 and 7] in order to remove the effects of additional systematics such as cosmic rays. However, given the lack of 2-min light-curve availability for many of the objects in this sample, this feature was not implemented in the development of this initial pipeline.

Following the removal of such sector-specific effects, the data were split into separate sections wherever there was a data gap of more than 0.1 d (to reduce the effect of flux jumps often seen after data gaps), except where doing so would result in fewer data points than the prescribed detrending window length discussed below.

2.5 Detrending of stellar variability

2.5.1 Choice of base-detrending method

Given the challenging range of activity-related and intrinsic stellar variability seen in the light curves of the young star sample (and the associated difficulty of modelling so many individual light curves in an initial transit search), it was considered wise to approach the detrending problem from the ground up, rather than necessarily relying on more traditional methods such as the Savitzky–Golay filter (Savitzky & Golay 1964). A number of different detrending

techniques were trialled early in the process, including simple low-order polynomial fitting, sinusoidal modelling, Savitky–Golay filtering and general smoothing over a range of window sizes. Of these, low-order polynomial fitting and the smoothing methods proved most successful at recovering injected transit signals (see Section 2.7 for more information on the injected transit method used).

A search for a method to combine and improve both of these methods led to LOWESS smoothing, or Locally Weighted Scatterplot Smoothing. This method, developed originally by Cleveland (1979a), is a local polynomial regression method, which works by fitting a low-order polynomial to a subset of the data (the width of which is set by a user-defined window) at each point along the *x*-axis using weighted least-squares regression. Under the weighted least-squares regression method, points nearer to the data point being estimated are given more weight than those further away in the window. This weighting is one of the key differences between this method and the more commonly used Savitzky–Golay filter (Savitzky & Golay 1964), and is particularly important for this application given the often swift evolution of young star light curves. Indeed following the choice of this LOWESS-smoothing method, it was independently highlighted to be one of the best-performing detrending methods for the young-star sample tested by Hippke et al. (2019). In exoplanet literature, however, no other mention of LOWESS smoothing for exoplanet searches was found, with even the related LOESS smoothing method appearing quite rarely, despite having been used in detrending the TRAPPIST-1 system (Luger et al. 2017) and in the Autoregressive Planet Search of Caceres et al. (2019).

For this work, the standard tricube weighting function ($w = (1 - |x|^3)^3$) was used, and the number of residual-based re-weightings retained as the default value of 3. The delta parameter was retained as $delta = 0.0$ as the sector-by-sector data sets were not overly large, but this can be adjusted if further speed enhancements are desired. Through experimentation, a window size of 30 FFI data points (15 h) was found to yield a good compromise between preserving the shape of injected transits and smoothing the stellar activity and variability of the host stars, except in more rapidly evolving light curves, where a window size of 20 data points (10 h) was found to be more appropriate.

2.5.2 Removal of peaks and troughs

One of the other key challenges presents in using a Box Least Squares (BLS) search for light curves from young active stars is that the troughs of regular stellar activity (e.g. from the rotation of starspots) are often picked up as the largest peaks in the BLS periodograms, even after LOWESS detrending. This is particularly the case for rapidly rotating stars, or those with short-period intrinsic variability. Such sources typically exhibit sharp peaks and troughs in the extracted 30-min light-curve, which is often highlighted in the BLS search. In an attempt to combat this problem for light-curves with particularly sharp stellar activity, the effect of cutting the peaks and troughs of this stellar activity was tested. To begin with, the peaks were located using the `find_peaks` function from the `scipy.signal` library[11] (Virtanen et al. 2019) using a required prominence of 0.001 and width of 15 data points (7.5 h). This helped to ensure that only wider peaks generally associated with stellar activity/variability were flagged as peaks or troughs,
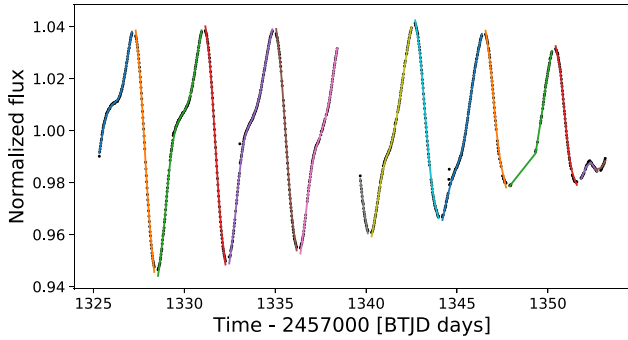
---

[11] https://docs.scipy.org/doc/scipy/reference/signal.html

Figure 3. Example of the peak-cutting technique in use on the rapidly evolving light curve for HIP 32235. The applied 20-box partial LOWESS-smoothing model used for detrending the light curve is also overplotted in colour. Note that the change in colour between subsequent sections of the LOWESS-detrending shows where the light curve has been split after any 0.1 d gap in the data. This typically occurs after each peak or trough, or where significant detrimental scatter has been removed.
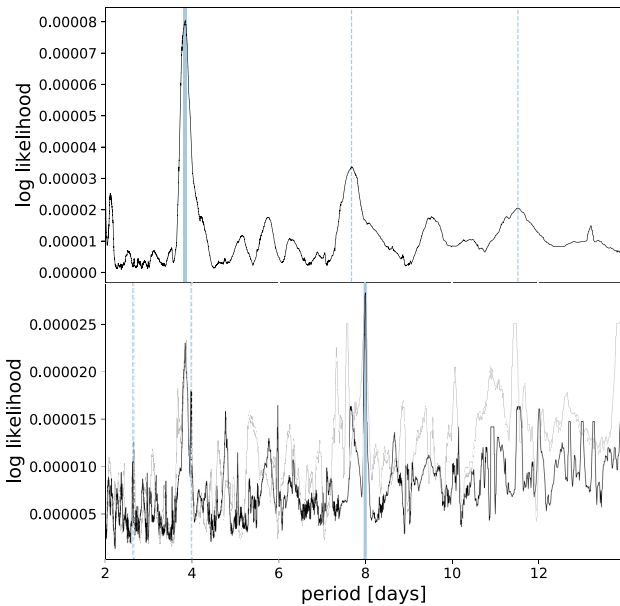

Figure 4. BLS periodograms for HIP 32235 following 20-bin LOWESS detrending without (top panel) and with (bottom panel) peak cutting implemented. In this case, peak cutting displays a clear benefit to the recovery of an injected $0.04 R_P/R_*$ 8-d planet in the presence of rapidly evolving (3.84-d period) stellar activity. The highest power period is highlighted in blue, with aliases of the same period shown by dotted blue lines.

but the settings may need adjusting for some more complex light curves. The activity/variability troughs were located using a similar `find_peaks` search on a negative version of the light-curve flux. Data points within 0.1 d either side of each peak and trough were then cut from the light curve before LOWESS-detrending was applied. An example of this in practice is shown in Fig. 3 for HIP 32235, a rotationally variable G6V star in the Carina stellar association.

The power of this peak-cutting method is illustrated for the same source in Fig. 4. When a $0.04 R_P/R_*$ radius ratio planet is injected into the light curve (using the method described in Section 2.7) and the light curve is detrended using the LOWESS-based method described above (boxsize = 20), any signal of the planet is clearly
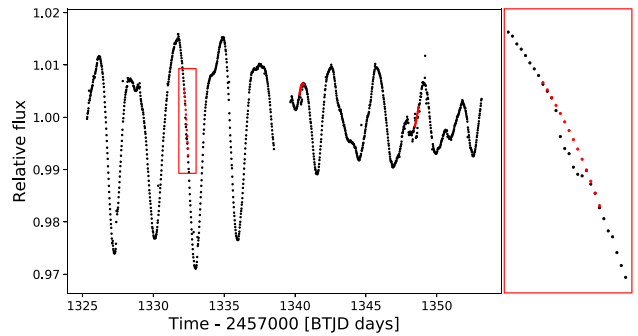

Figure 5. Example of interpolation over transit masked areas of the light curve for DS Tuc A in order to best preserve the stellar activity signature through these segments. Original data points are shown in black, while interpolation over the removed transits is shown in red. The red box on the right-hand panel shows detail of the area in the vicinity of the first transit.

overwhelmed by the 3.84-d period rotational variability of the star, as is shown in the top panel of Fig. 4. However, applying the described peak-cutting technique to the data easily recovers the signal of the injected 8-d 0.04 R_P/R_* planet, as illustrated in the lower half of Fig. 4. This technique is thus another powerful tool for pushing down to lower radii in the search for exoplanets around young active stars. In general, this peak-cutting method was found particularly effective for light curves with sharp oscillations/rotations, or ones which were strongly periodic. However, it is important to note that this technique did not always offer improvements over the non-peak-cut method. A further discussion of when this technique is most effective can be found in Section 3.5.3.

2.5.3 Transit masking and light-curve interpolation

Transit masking is a commonly used method in planetary and eclipsing binary science to independently detrend long-term variability in stellar light curves without changing the shape of the transit curve, and is implemented into most common light-curve manipulation tools (e.g. Luger et al. 2016; Lightkurve Collaboration et al. 2018). However, due to the often rapid evolution of young star light curves, simply cutting out the data near a suspected transit before detrending can lead to spurious variability signals over the duration of the transit. In order to combat this issue in the developed detrending method, a two-step transit masking approach was used, which accounts for the brightness variation of the host star during the transit. First, a mask is generated based on a period, epoch, and duration for the suspected transit (either user defined or from a previous BLS search), and the new light curve generated by removing any points within the selected transit duration. Then, in order to account for the stellar flux variability over the course of the transit, the new holes in this light curve are refilled using a quadratic interpolation (using `scipy.interpolate.interp1d`) between the cut points. It is this new flux array featuring interpolated sections over the transit mask that is then used in the LOWESS-detrending step, and, in turn, divided out to form the final detrended light curve. The importance of using such a method is illustrated using the source DS Tuc A in Fig. 5. It is immediately obvious that not taking into consideration the variation in stellar flux around the epoch of the first cut transit (approx TJD 1132) would lead to a considerably different transit shape after the main LOWESS-based detrending was is undertaken. A full discussion of the full detrending of DS Tuc A and the resulting

recovery of the young exoplanet DS Tuc Ab can be found below in Section 3.1.

The authors make available a basic version of this detrending code online.[12]

2.6 Transit searching algorithm

The standard BLS method was used to search for periodic transit-like signals in the detrended light curves. This method, originally developed by Kovács, Zucker & Mazeh (2002), fits a series of box-shaped dips at a range of periods in order to generate a periodogram comparing the relative strengths of the different period hypotheses. The specific implementation used in this code is the `astropy.timeseries.BoxLeastSquares` method,[13] which is a `python` implementation of the computational method described by Hartman & Bakos (2016). In this work, the strongest BLS peak and the two next strongest non-harmonic peaks were investigated.

2.7 Injected transits

In order to test the sensitivity of this new detrending method to finding planets around young stars, a series of model planet transits were injected into the light curves before detrending. For this analysis, eight different orbital periods were tested (1.0, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0 and 14.0 d), with the epoch chosen randomly for each injection. Note that unlike ground-based surveys, *TESS* does not exhibit 1-d systematic errors due to the Earth's rotation, which could adversely affect these period choices. Rather than injecting specific planet sizes, five planet to star ratios R_p/R_* were tested in this sensitivity analysis: 0.1, 0.075, 0.05, 0.04, and 0.03, beyond which recovery was observed to be quite rare. For reference, around a sun-like star a radius ratio of 0.1 corresponds to an approximately Jupiter-sized planet and $R_p/R_* = 0.03$ corresponds to a sub-Neptune-sized planet. Stellar parameters for each star were retrieved from TIC v8 (Stassun et al. 2019). Where possible, orbital separation for each planet was then derived from Kepler's Third Law. If the information on a star's mass or radius was not available, the corresponding planet was assigned an orbital separation of $17.0R_*$, representing the average orbital separation for planets with 8-d periods (in the middle of the period range) on the NASA Exoplanet Archive.[14] Planet transits were generated using Kreidberg's (2015) `batman` PYTHON implementation of Mandel & Agol's (2002) transit model, assuming non-linear limb darkening with coefficients [0.5, 0.1, 0.1, –0.1]. After injection, each light curve was detrended using a standard 30-bin run of the LOWESS-smoothing pipeline (without peak cutting implemented). Injected signals were considered to have been 'recovered' if they appeared as one of the three highest peaks in the BLS periodogram, ignoring harmonics of the maximum peak.

3 RESULTS FROM SECTORS 1–5

Despite 30-min DIA light curves only being available for 256 of the BANYAN objects within the first five sectors of *TESS* data, a wide range of interesting activity was observed, from the recovery of

known young exoplanets and *TESS* objects of interest (TOIs) to an eccentric eclipsing binary system and a large variety of unusual rotation and activity profiles. Even this relatively small sample clearly demonstrates the unusually large variation in light curves of young stars compared to their older counterparts, and consequently helps to explain why far fewer planets have thus far been found around stars of these ages. The conducted sensitivity analysis goes one step further, investigating the comparative recovery rates for different combinations of injected period and planetary radius. Some of the most interesting results are summarized below in this section.

3.1 Recovery of confirmed young exoplanets

One of the most promising initial results from the application of this pipeline on the 30-min data was the recovery of both of the known transiting exoplanets found around young stars in Sectors 1–5: DS Tuc A b (Benatti et al. 2019; Newton et al. 2019) and AU Mic b (Plavchan et al. 2020). The recovery of DS Tuc A b is described in detail here as an example of the full pipeline in use.

DS Tuc A (TIC 410214986/TOI 200.01/HIP 116748 A) was observed in Sector 1 of the *TESS* observations, carried out between 2018 July 25–August 22. It is a G6V type star known to be associated with the 45 Myr Tucana-Horologium association (Zuckerman & Webb 2000). DS Tuc A fell on camera 3 of the instrument, and yielded approximately 27 d of photometry.

Interestingly, Newton et al. (2019) admit in their work that the candidate signal was actually found by human eyeballing around a spurious activity-induced periodicity peak flagged by a run of the SPOC Transiting Planet Search (TPS) module on the 2-min PDC-SAP data. However, a later archival TPS run of the SPOC pipeline (after the planet candidate was announced to the community as TOI 200.01) was seen to detect a periodic transit crossing event which passed all of Newton et al. (2019)'s false-positive (FP) tests. The planetary candidate was later confirmed using additional photometry, spectroscopic methods and high-contrast imaging.

Benatti et al. (2019) independently reprocessed the TESS data for this object using improved stellar parameters (including crucially accounting for dilution from DS Tuc B) and fitted two different models to determine planetary parameters. The first model involved modelling only the first two transits (on account of the large pointing jitter around the 3rd transit) using `PyOrbit` (Malavolta et al. 2016) to complete a simultaneous fit for modulation and a transit signal, along with `emcee` (Foreman-Mackey et al. 2012) and `PyDE` (Storn & Price 1997) to establish the most likely planetary parameters. Alongside this they tested modelling all three transits with the `batman` package (Kreidberg 2015) after applying a 0.55-d third-order running polynomial to flatten the light curve and analysed the posterior with `emcee`. While both methods yielded consistent results, they eventually adopted the planetary parameters from the first method and found a best-fitting solution of a $0.5R_{Jup}$ planet. Benatti et al. (2019) confirmed the planetary nature of this object using radial velocities from the HARPS spectrograph.

In this work, the planetary signal was highlighted as the highest peak of the BLS periodogram in a 20-bin LOWESS-partial run of the standard detrending pipeline, however, the transits were also clearly visible by eye after the 30-bin LOWESS-partial run. Noting that the third transit fell in the period of heightened pointing jitter between TJD 1347-1349, this section of the data was unmasked for this target. The recovered period was 8.14 d, in agreement with
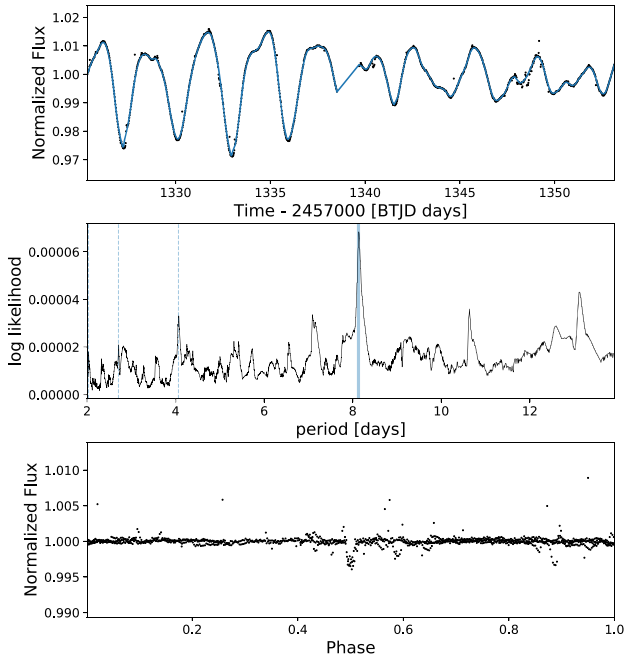

Figure 6. Example of the developed detrending pipeline in use, showing the recovery of the 45 Myr exoplanet DS Tuc A b. Top panel: the original Sector 1 light curve for DS Tuc A with overplotted 20-bin LOWESS detrending in blue, including interpolation over suspected transits. Middle panel: BLS Periodogram for the light curve after detrending, with the peak period of 8.138 d highlighted in blue (n.b. aliases of this period are shown by the blue dotted lines). Bottom panel: the resulting light curve after the detrending pipeline has been applied, folded by the maximum peak of the BLS Periodogram.

the accepted value from Newton et al. (2019). Using the period and epoch derived from the BLS search, the transit masking and light-curve interpolation technique described in Section 2.5.3 was applied, completing the clear detrending and transit recovery of DS Tuc A b presented in Fig. 6. Note that in the bottom panel the light curve has been folded by 8.138 d, the accepted planet period from (Newton et al. 2019). Interestingly, the use of the peak-cutting technique for this object was found to increase the significance of the true transit period in the 30-bin case (changing the 8.14-d period-peak in the BLS periodogram from insignificant to the 6th strongest after peak cutting), but decreased its significance for the 20-bin LOWESS detrend.

The clear recovery of this Neptune to Saturn-sized planet using the 30-min data alone bodes well for future exoplanet candidate discoveries from the FFIs, thus demonstrating the wealth of knowledge to be gained from these images.

Unfortunately, data for AU Mic was not extracted by (Oelkers & Stassun's 2018) 30-min pipeline, however, application of this pipeline to the 2-min data for AU Mic easily recovered the signal of the 8.5-d planet proposed by Plavchan et al. (2020), as shown in Fig. 7. In this case, a 20-bin LOWESS-smoothing run revealed the correct 8.466-d period as the third highest peak in the periodogram. In this particular case, the peak-cutting option did not aid recovery of the planet substantially since the improvements of cutting the sharp troughs were balanced by the inadvertent cutting of the first transit. None the less, the 8.5-d signal remained the third highest peak after 20-bin LOWESS smoothing was applied to the peak-cut light curve.
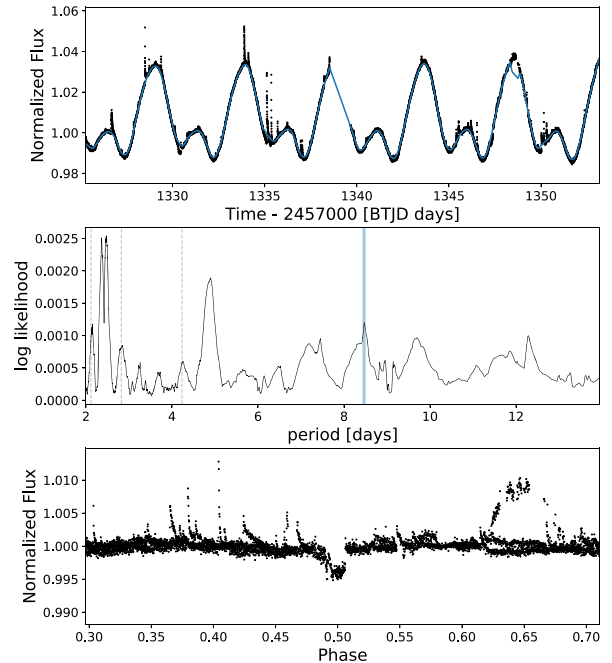
Figure 7. A 20-bin LOWESS smoothing run on the *TESS* 2min cadence data for AU Mic, showing the recovery of the 8.466-d planetary signal first identified by Plavchan et al. (2020). Format similar to that explained for Fig. 6, except that the final light curve is folded by the third highest period, as highlighted in blue in the BLS periodogram.

3.2 Retrieval of other *TOI*

A number of other interesting signals were independently found using this pipeline, including a pair of additional TOIs. In the interest of independence, these signals were found with no prior knowledge of the TOIs in these sectors, but instead candidate signals highlighted in the BLS search in this work were cross-matched with the TOI list at a later date. This resulted in the recovery of two known TOIs: TOI 447.01 and TOI 450.01.

For TOI 447.01, a repeated dip of approximately 15 mmag was highlighted by the BLS search in Sector 5 with a period of 5.528 d around the F5/6V star HD 33512/TIC 14091633 (see Fig. 8). This signal was uncovered using the standard 30-bin LOWESS run of the pipeline, without peak cutting. HD 33512 is a relatively bright ($T =$ 8.8) F5/6V star in the Octans association, giving it an approximate age of 35 ± 5 Myr (Gagné et al. 2018a), which would make this an interesting target for follow-up of young evolving systems if the planetary hypothesis was correct. As a TOI this target underwent follow-up by the TFOP team, and though initially seemed positive from the *TESS* data alone, it was eventually flagged as an FP by the TFOP working group due to the observed linear drift in RV measurements, relatively large radius, and changing width of full width at half-maximum.[15] None the less, this object is a likely long-period binary, so may still be an interesting system for the study of young binary systems.

The second TOI independently recovered was TOI 450.01, around the source 2MASS J05160118-3124457/TIC 77951245. This source was observed in Sectors 5 and 6 of *TESS* observations, however, alike to TOI 447.01 only Sector 5 light curves were
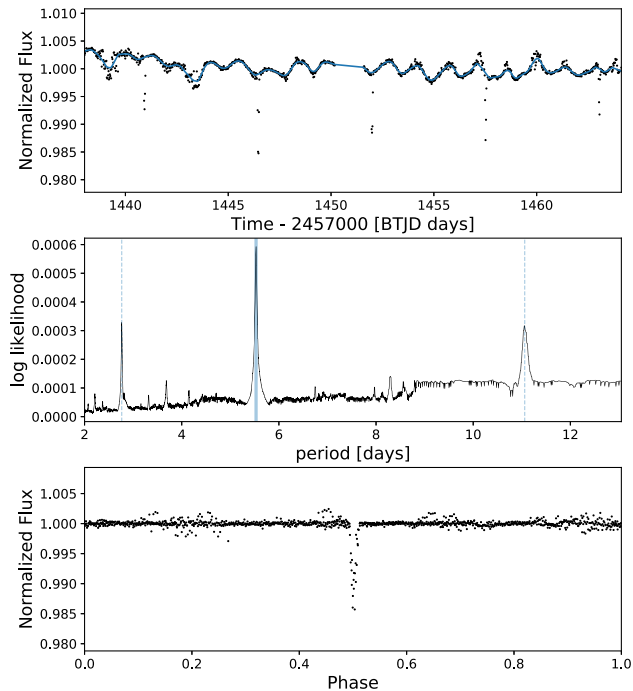
---

[15]https://exofop.ipac.caltech.edu/tess/target.php?id=14091633

Figure 8. Recovery of TOI 447.01/HD 33512, a likely long-period binary in the 35 Myr Octans Association. In this case, the maximum period was found to be 5.528 d. Format after that explained in Fig. 6.

Figure 9. Recovery of TOI 450.01/J0516-3124, a spectroscopic binary in the 42 Myr Columba Association. In this case, the maximum period was found to be 10.7 d. Format after that explained in Fig. 6, except that overplotted LOWESS-detrending in the top panel has bins of 30 rather than 20.

available for it from the DIA 30-min pipeline. In this work, the candidate signal was initially revealed by a 30-bin LOWESS-partial run of the main detrending pipeline on Sector 5. Two approximately 40 mmag transits were observed in the flattened light curve, with a recovered period of 10.74 d and initial epoch of 1443.2 d, as shown in Fig. 9. The parameters derived from this 30-min data are slightly different to those derived from the 2-min data by the SPOC TOI analysis ($P = 10.7148 \pm 0.0001$, depth $= 63.6 \pm 0.8$, epoch $= 2458443.1686 \pm 0.0003$), however, this is not surprising given the much lower density of points in each transit signal (15 times fewer). TIC 77951245 is an M4 star of $T = 12.375$ in the 42 Myr Columba association. Based on the depth of the observed signals and the radius of the star, this signal would correspond to an approximately Jupiter-sized planet if the planetary hypothesis was correct. However, follow-up of this object by the TFOP working group eventually revealed it to be a near-equal mass spectroscopic binary, based on HRS measurements on the 10-m SALT telescope on 2019 August 10.[16]

These two TOIs along with the newly confirmed exoplanet DS Tuc A b discussed in Section 3.1 represent all three of the TOIs highlighted by the main 2-min SPOC pipeline for this selection of young stars. This illustrates the fact that this pipeline is working as effectively as the main SPOC 2-min pipeline for these young objects, further emphasizing its potential for finding candidate signals around stars present only in the 30-min cadence data.

3.3 Eclipsing binaries

A number of clear young eclipsing binaries were also revealed by this survey, perhaps the most dramatic of which being HD 28982.
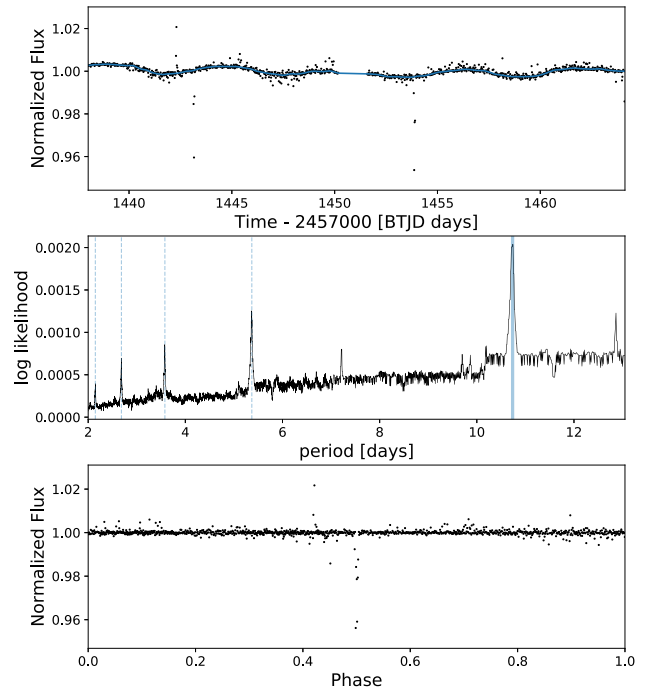
This highly eccentric eclipsing binary system was revealed in this work by a 30-bin run of the LOWESS-based detrending (without peak cutting) applied to Sector 5 of the *TESS* 30-min DIA data (see Fig. 10, panel a). A period of 5.97 d was clearly highlighted as the maximum power period in the BLS periodogram, and two distinctly different duration and depth transit signals were visible. HD 28982 is associated with the AB Doradus Moving Group (ABDMG), which is approximately 150 Myr old (Gagné et al. 2018a).

Because it has long been known to be a 'bona-fide' member of this moving group it was included in the initial *TESS* CTL (Stassun et al. 2019) and received 2-min coverage in the main *TESS* survey. This provides an interesting opportunity to compare the 2- and 30-min data for this young object. After retrieval from MAST, the 2-min data were subjected to the same detrending as the 30-min data by expanding the number of bins for each LOWESS-based detrending step to 450, 15 times as many as in the 30-min analysis. This revealed the detrended light curve shown in panel (b) of Fig. 10. Aside from the obvious increase in signal-to-noise ratio and visually more obvious signal provided by the 2-min data, the 30-min data clearly provide sufficient information to constrain the EB period and durations. However, alike to the TOIs, the 30-min data suffer in terms of depth accuracy, likely due to exposure smearing from the longer 30-min cadence data. In addition, the 30-min cadence data from the DIA pipeline is more affected by scatter than the 2-min data, likely due to a combination of pointing scatter and improperly corrected scattered light. Care must therefore be taken when evaluating planetary radii for planet candidates based on 30-min data alone, and may be better left to higher cadence photometry in follow-up observations.
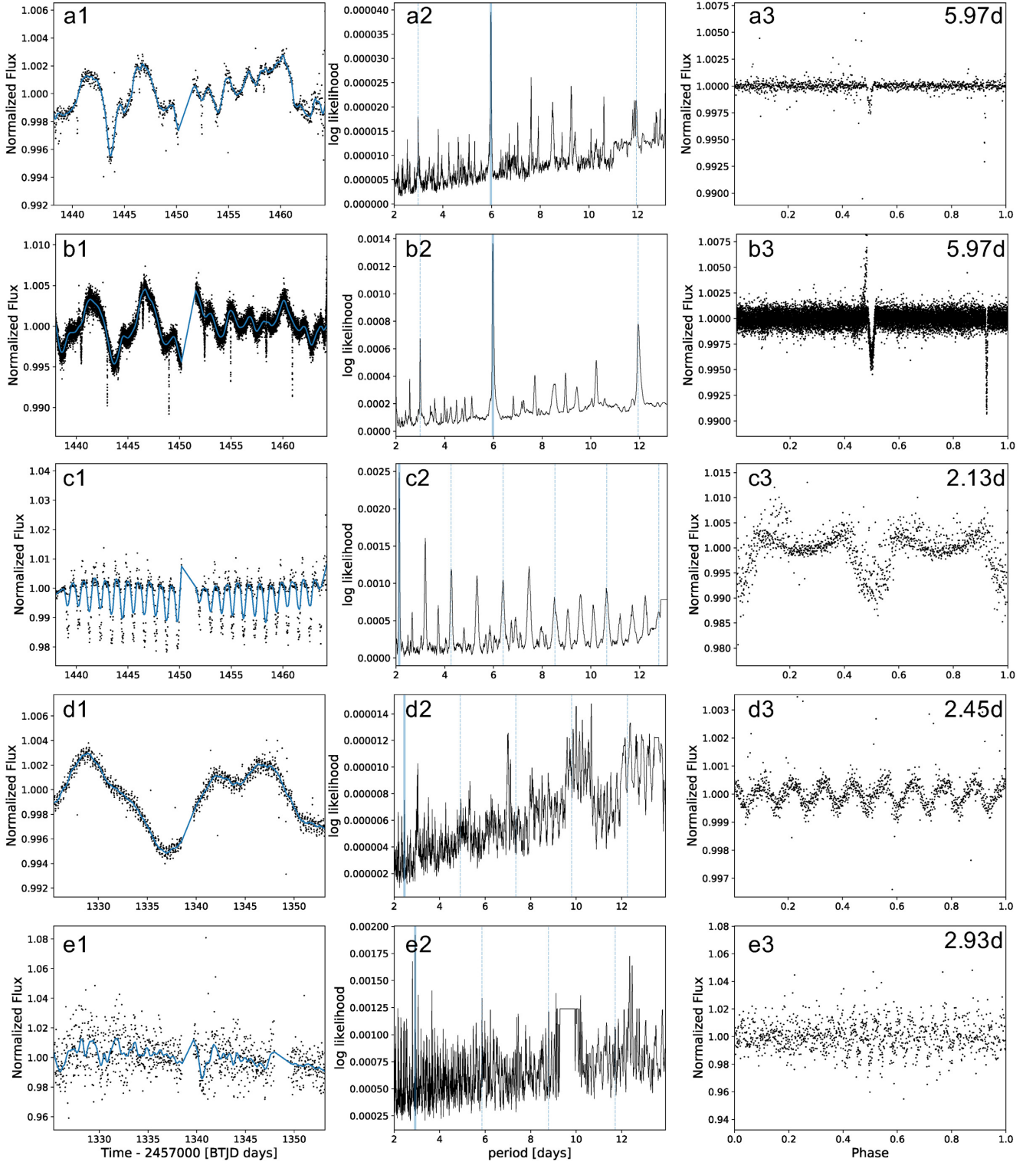
[16]https://exofop.ipac.caltech.edu/tess/target.php?id=77951245

Figure 10. Sector 1–5 lightcurve zoo. Column 1 shows the original light curve for each star with a 20-bin LOWESS-smoothed fit overplotted in blue. Column 2 shows the BLS periodogram for each source after LOWESS detrending, with the highest power period highlighted in blue (except for J0635-5737 – panel d – where a shorter alias is chosen for clarity), and aliases of this period shown with dotted blue lines. Column 3 shows the flux after 30 bin LOWESS detrending for each source folded by the highest power period. Individual sources from top to bottom: Panel a: HD 28982, an eccentric EB in ABDMG (approx 150 Myr) – DIA 30-min data, Sector 5; Panel b: HD 28982 in *TESS* 2min data; Panel c: J0529-2852/TIC 31281820 in COL (approx 42 Myr), demonstrating strongly periodic drops – DIA 30-min data, Sector 5; Panel d: J0635-5737/TIC 348839788 in ABDMG (approximately 150 Myr), demonstrating 'double variation' in its rotation – DIA 30-min data, Sector 1; Panel e: J0552-5929/TIC 350712873 in THA (approximately 45 Myr), demonstrating very fast rotation – DIA 30-min data, Sector 1.
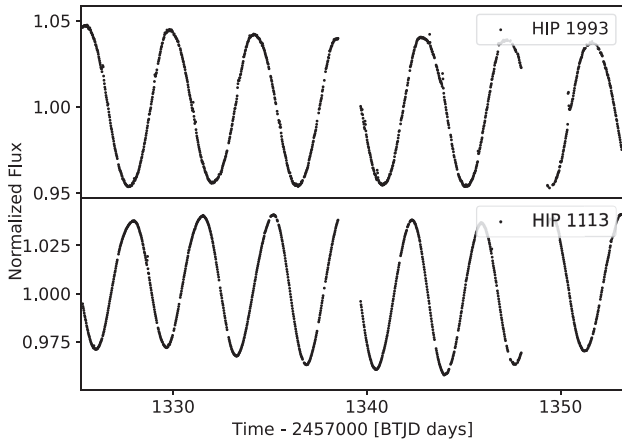
Figure 11. Examples of rotation/activity type 1: near-uniform sinusoidal variation – HIP 1993 (top panel) and HIP 1113 (bottom panel) in Sector 1.

3.4 Rotation and activity

Rotation and spurious stellar activity were the cause of the strongest peaks in many of the stars viewed in this sample, which helps to explain why searching for planets around young stars is so much harder than around many older stellar hosts. As a first step in characterization of this rotation and activity, the period of each star's main stellar variability/activity was also recorded using the generalized Lomb–Scargle method (Lomb 1976) implemented into python as the LombScargle function in the python astropy library[17] (The Astropy Collaboration et al. 2013, 2018). Furthermore, the main amplitude of the primary variability was determined in the peak-cutting step discussed in Section 2.5.2, using the peaks and troughs identified for removal. A wide variety of rotation and activity curves were observed in this work, broadly separated into four categories: near-uniform periodic, periodic but evolving, aperiodic, and fast rotators. Amplitudes of oscillation varied from 0.02 to 9 per cent, while primary periods of flux variation varied from 0.122 d (e.g. J0552-5929 – panel e of Fig. 10) to non-variable over the 27-d time-period.

The first of these categories, those light curves that exhibit near-uniform variation in amplitude over a set period, are well illustrated by the sinusoidal curve of sources such as HIP 1993 and HIP 1113 (Fig. 11), alongside strongly periodic drops in sources like 2MASS 05292529-2852274 (see panel c of Fig. 10). These types of sources are theoretically ideal for detrending, as they can be easily modelled by quasi-periodic smoothing functions, especially when the period of flux variation is over periods of two or more days. Furthermore, the peak-cutting method discussed in Section 2.5.2 was often observed to aid the retrieval of injected transits for these periodic sources. However, very fast rotation of this type still provides difficulties, as discussed below.

Another very important category of variation seen ishose light curves with strong periods but obvious variation in flux amplitude. This type of light curve is well illustrated by sources such as HIP 105388 (observed in *TESS* Sector 1) and AB Pic (observed in all Southern hemisphere sectors – 1–13) – see Fig. 12. The oscillation amplitudes of these sources are both observed to considerably change over the course of a single sector. Indeed, in the case of
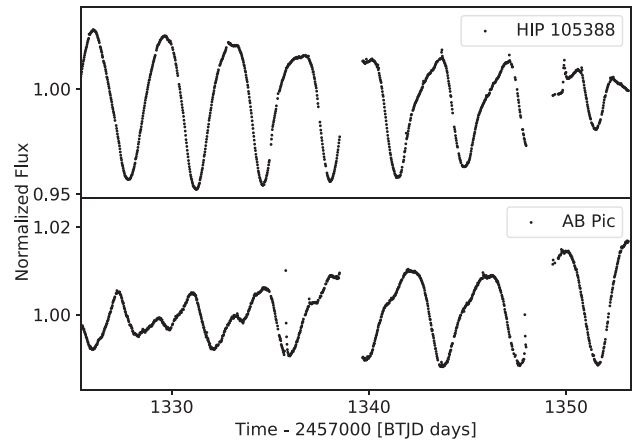

Figure 12. Examples of rotation/activity type 2: periodic with rapidly evolving amplitudes – HIP 105388 (top panel) and AB Pictoris (bottom panel) in Sector 1.
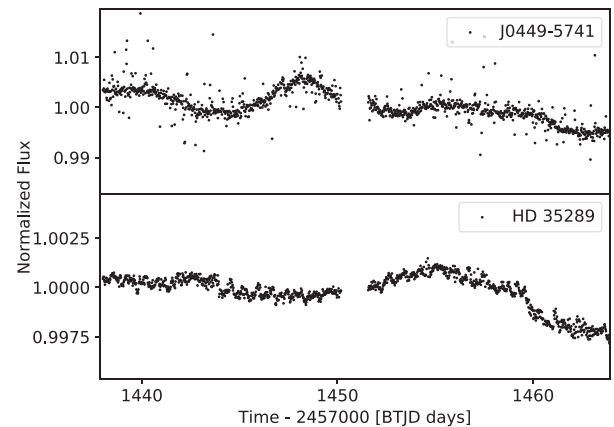

Figure 13. Examples of rotation/activity type 3 – Aperiodic variations. DIA 30-min cadence light curves for J0449-5741 (TIC 220425740, top panel) and HD 35289 (bottom panel) in *TESS* Sector 5.

AB Pic, the amplitude of the primary oscillations changes from 0.01 to >0.03 per cent over the course of the five sectors of data for which the 30-min DIA data exist. Other sources such as HIP 1481 appear to exhibit 'beating'-like behaviour in their light curves, going through periods of more and less intense oscillations yet with similar periods throughout. Thus, for these objects, while the strong periodic nature of their oscillations makes the periods easy to identify, blind removal of these varying oscillations is difficult on a wide-scale basis. None the less, the peak-cutting technique discussed in Section 2.5.2 particularly beneficial for sources of this type, such as HIP 32235 (Fig. 3) and HIP 105388 (top panel, Fig. 12).

Alongside these periodic light curves, a smaller number of aperiodic variations (or perhaps those with rotation or variation periods much longer than the 27-d observation time) were observed. Most of these variations were sufficiently long period to be easily removed by the base 30-bin LOWESS-smoothing method [as was the case for J0449-5741 and HD 35289 (Fig. 13)], or are sufficiently aperiodic that any activity would not overcome true periodic transit signals on the BLS periodogram. However, some more complex aperiodic cases exist, such as TIC 348839788 (Panel d, Fig. 10), where two separate rotation and activity profiles seem to be apparent

Table 3. Results from complete Sector 1–5 sensitivity analysis for the 256 stars in the BANYAN young star sample with DIA FFI data. Percentage recovery is presented both overall and in each individual sector. Note that results from the 14 d injections were excluded from the percentage recoveries due single transits frequently being present at this period. Total number of individual sources for each sector are: Sector 1: 74; Sector 2: 77; Sector 3: 75; Sector 4: 120; and Sector 5: 138.

R_P/R_*	Overall (%)	S1 (%)	S2 (%)	S3 (%)	S4 (%)	S5 (%)
0.1	77.6	79.7	80.3	73.1	69.8	84.3
0.075	61.7	61.6	64.0	56.8	51.4	72.2
0.05	40.9	43.8	36.9	39.6	31.6	50.31
0.04	30.2	31.3	30.4	31.1	20.0	38.0
0.03	20.4	24.3	18.4	22.1	11.1	26.6

– one aperiodic larger amplitude evolution clearly evident in the original light curve, and one much faster rotation-based signal with a period of approximately 0.27 d (see panel d, Fig. 10). This 'double variation' is present in a small number of other sources too, so should be considered carefully in future detrending efforts. For such sources, this will likely require at least two detrending steps (a wider smoothing to remove the large-scale activity, followed by smaller scale smoothing/modelling of the rapid rotation) in order to yield a flat pre-BLS-search light curve.

Perhaps the greatest challenge in the search for young exoplanets is the case of rapidly rotating stars such as HIP 22295 (Panel e, Fig. 10) and CD-46 287 (bottom panel, Fig. 16). These fast rotators are unfortunately quite common in stars of such young ages [indeed at 52 stars (20.3 per cent) in this sample had rotation periods of less than 1 d], likely due to leftover angular momentum from their formation (Prialnik 2009). For the fastest rotators, 30-min data alone struggles to untangle rotation from any potential transit signals, especially when rotation periods of less than 1 d are coupled with significant changes in amplitude. In order to more effectively dissociate such rapid rotation from transit signals, using more sensitive 2-min cadence data are preferred, likely coupled with more intelligent modelling than simple smoothing methods. The planned *TESS* 20s cadence data would also aid this effort, as rotation profiles will become more carefully defined and more detail of transit ingresses/egresses may become apparent. However, techniques to overcome this fast rotation challenge still need to be developed.

3.5 Sensitivity analysis results

3.5.1 Overall results

The conducted sensitivity analysis revealed a number of interesting results, for which an overall summary is presented in Table 3 and Fig. 14. What is immediately apparent is the steep drop-off in recovery as the R_P/R_* radius ratio decreases, falling from 77.6 per cent at a radius ratio of 0.1 down to 20.4 per cent at 0.03. This is to be expected as the inherent scatter and leftover variability amplitude of the light curves steadily overcome the signal of the smaller injected planets.

Interestingly, the recovery of injected signals was not entirely consistent across the different sectors viewed in this analysis. This discrepancy appears to result from two major factors: comparative sector systematics and evolving light-curve amplitudes for a small subset of stars. Regarding the first factor, Sector 4 was the worst affected, with recovery rates for all radius ratios consistently far below the average. This is likely due to the strong reflected light
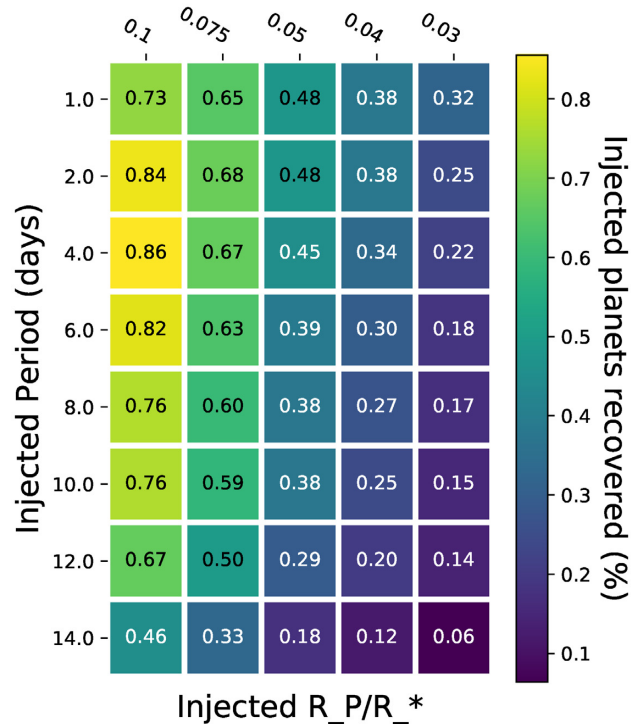

Figure 14. Overall sensitivity analysis results (Sectors 1–5) by injected period and planetary radius. The numerical contents of each box correspond to the overall percentage recovery of injected planets for each particular combination, with colour gradient of the same added for clarity.

glints seen at the end of each orbit for some sources (particularly those on camera 4), which are harder to systematically remove without also removing useful data from other light curves. Conversely, recovery was significantly better in Sector 5 compared to the other sectors, with 26.6 per cent of injected planets still retrieved down to radius ratios of 0.03. This intersector variation is well illustrated by the light curves for J0455-6051/TIC 55651278 (an M5 star in the ABDMG), where injected planetary signals were recovered down to radius ratios of 0.075, 0.05, 0.075, 0.1, and 0.03 in Sectors 1–5, respectively, in the 2–6-d period cases. In this case, such variation was caused by a combination of evolving flaring-type stellar variability coupled with extra noise in Sector 4. On the other hand, AB Pictoris (a K1V star in the Carina Association) demonstrates the second source of sector-dependent sensitivity, with injected planetary signals recovered down to steadily larger radius ratios of 0.04, 0.04, 0.075, 0.075, and 0.1 in Sectors 1–5 due to the evolving stellar variability increasing in amplitude over time. While differences in the systematics of each sector can reasonably be expected to reduce over time as the *TESS* satellite pointing is refined, the evolution of individually active sources will remain a challenge not only for *TESS* but also for future missions such as *PLATO*.

A slightly more complex trend was observed when varying the injected planetary periods, as is highlighted by the overall heat-map in Fig. 14, and the sector-by-sector heat-maps in the Appendix. For all sectors except, Sector 2 (where a 2-d period was preferred), the recovery peaked around a 4-d injected period for the deepest ($0.1 R_P/R_*$) transits, unlike the shortest 1.0-d period that may be initially expected. However, as the injected planet radii decreased a more standard drop in recovery from 1–14 d was observed. Such behaviour suggests that this form of analysis is most sensitive to

Table 4. Full sensitivity analysis table for each of the 256 stars with DIA light-curves. Includes information on the highest likelihood period recovered for each star in every sector it appears, with injected 1–14-d period planets and R_P/R_* radius ratios from 0.1 to 0.03. Full table available online, with ellipses here denoting where data has been cut for clarity.

Target ID	RA (°)	Dec. (°)	Sector	Injected period (d)	Radius ratio	...	Log likelihood	Max period (d)	Recovered?	Notes
2M0123-6921	20.79	−69.36	1	1.0	0.1	...	0.00121	2.65	TRUE	Alias
2M0123-6921	20.79	−69.36	1	1.0	0.075	...	0.00125	3.98	TRUE	Alias
...	...	...	...	...	...	...	...	...	...	...
WX Col B	84.30	−42.72	5	14.0	0.04	...	2.57e-5	12.52	FALSE	–
WX Col B	84.30	−42.72	5	14.0	0.03	...	1.95e-5	11.22	FALSE	–
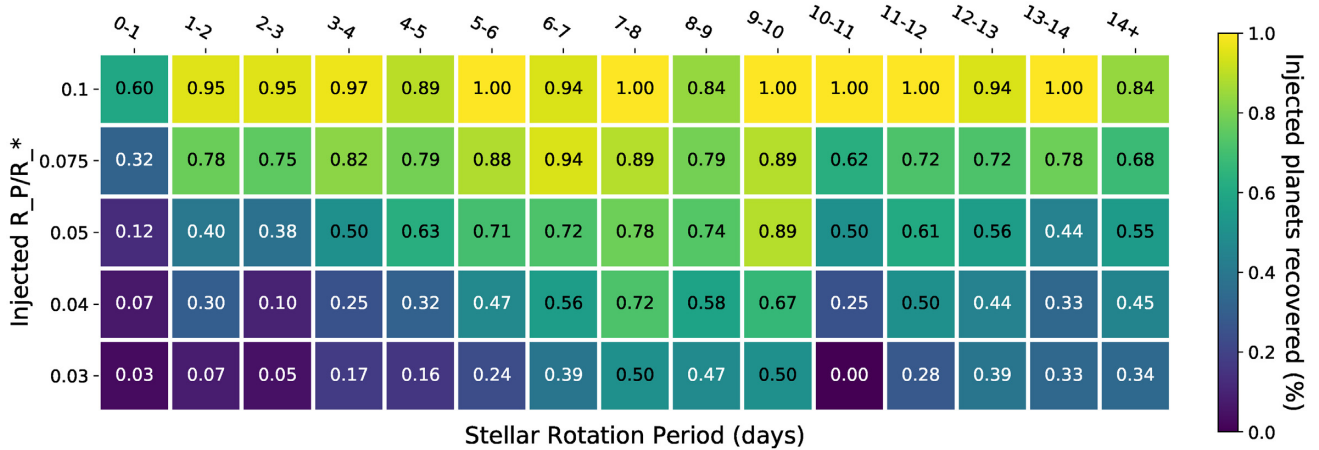
Figure 15. Stellar rotation period versus recovery depth for a 4-d injected period planet across all targets and sectors. Design similar to Fig. 14.

larger planets in the 2–6-d period regime, and most sensitive to closer-orbiting planets as radii decrease further. In case the observed drop for 1-d planets was due to 1-d Earth-related systematics such as Earthshine, the injection/recovery analysis was repeated with a 1.1-d period planet, however, similar recovery rates were observed to the 1-d case. Instead, this discrepancy may be related to a combination of the increased activity of young stars coupled with the small number of data points per transit for such short-period planets. The primary reason for the much lower recovery of planets with injected periods of 14 d was that data gaps and randomly injected epochs frequently led to single transits appearing in the data set. This same effect was also observed for some 12-d injections. Some interesting sector-to-sector variations are also clear in the sector-specific heat-maps (see Figs A1–A5), with Sector 4 exhibiting the lowest recovery rates (especially for the smallest radius planets), and Sector 5 the highest.

Another useful product of this sensitivity analysis is the individual sensitivity to planet detection for each star. An overview of this table is presented in Table 4, with a full version available online.

3.5.2 Rotation period versus recovery

As has been seen in Section 3.4, a wide range of rotation periods was observed for stars in this sample. This provides an interesting opportunity to test the relationship between rotation rate and recovery, which is especially important given concerns about finding planets around swiftly rotating stars and those with rotation periods near the injected planetary period. In order to investigate this relationship, a injection/recovery analysis of rotation rate versus depth recovered for each target was conducted for each injected period. A typical example of one of these plots is shown in Fig. 15 with an injected

period of 4.0 d. Somewhat counter-intuitively, while there is a slight skew towards larger radius ratios below rotation periods of 5 d, there is no significant evidence that the overall recovery depth of injected transits is a function of rotation period, with recoveries down to radius ratios of 0.03 (and complete non-recoveries) across the entire period range. In addition, recovery depth does not seem to be detrimentally affected by being close to the injected period (4 d in this case), as similar recovery was observed for every injected period. Note that many variations between individual period ranges can be explained by the relatively small number of targets in this sample, especially in the 10–11-d period case where only eight targets were present with periods in this range. However, it should be noted that 60 per cent of all unrecovered signals were from targets with rotation periods of less than 1 d. This constituted 40 per cent of all injected signals into targets with these rotation periods. Furthermore, the vast majority (79 per cent) of the recovered light curves with rotation periods <1 d were only found down to radius ratios of $0.075 R_P/R_*$. This suggests that searching for small planets ($<0.075 R_P/R_*$) around stars with rotation periods less than 1 d is likely futile until better techniques are developed to detrend this fast rotation. However, since failures for all other rotation periods were consistent at 0–3 light curves per 1-d period interval (e.g. 4–5 d), searching for planets around stars with rotation periods longer than 1 d shows promise, even down to $0.03 R_P/R_*$ planets.

3.5.3 Effectiveness of the peak-cutting technique

Since the peak-cutting technique exhibited variable effectiveness according to the shape of the light curves in building the pipeline, much of the prior analysis was undertaken without the peak-cutting option applied. However, to test the wider effectiveness

Table 5. Comparison between percentage recovery of injected planets in Sector 1 for both in the original and peak-cut light curves.

Radius ratio	Number of lcs	Original (%)	Peak-cut (%)
0.1	61	90.2	82.0
0.075	59	62.7	64.4
0.05	59	39.0	40.7
0.04	60	31.7	35.0
0.03	60	25.0	25.0

of this option and evaluate where its application was most useful, a comparison study was undertaken for all targets in Sector 1, both with and without peak cutting. For this comparison test planets with a set period of 8.0 d and radius ratios of $R_P/R_* = 0.1 - 0.03$ were injected into each of the light curves in Sector 1. Light curves were then detrended using the standard 30-bin LOWESS-detrending method described above and searched through using the standard BLS method.

Of the 74 Sector 1 targets in the sample with light curves available from Oelkers & Stassun ([2018])'s DIA FFI pipeline, the basic peak-cutting technique failed for 13–15 of the objects (depending on the injected radius ratio). In all cases, this was due to peaks and/or troughs not being located by the automatic `find_peaks` function, typically because the light curves were simply too flat to exhibit any significant peaks or troughs. Of the remaining 59–61 light curves, recovery with peak cutting was in general comparable, or slightly better than, the recovery of planets when peak cutting was not applied (as is summarized in Table [5]). The one exception to this was the $0.1R_P/R_*$ case, which was caused by the peak-cut analysis failing for three of the original 0.1 radius ratio light curves.

However, it is when looking at individual targets that the power of the peak-cutting technique is most evident. For seven of the targets in Sector 1, the use of the peak-cutting technique yielded significant improvements in the recovery of smaller injected radius ratios. The most significant improvements were seen for the object J0247-6808, where injected planets were recovered down to a radius ratio of $0.03R_P/R_*$, despite not being recovered for any radius ratio in the non-peak-cutting case. Similarly, in the case of HIP 32235, the depth recovered dropped to 0.03 from 0.075 in the non-peak-cut analysis. Similar improvements (though with less significant drops in recovered depth) were observed for RBS 38, TYC 8895-112-1, J0346-6246, J0414-7025, and J2231-5709.

There were two different reasons why recovery was improved for these objects – one showing the technique working as designed and the other a fortuitous side effect. The former can be seen in HIP 32235, RBS 38, TYC 8895-112-1, and J2231-5709 where each light curve exhibits sharp variability peaks with periods of order 3–8 d and amplitudes greater than 2.5 per cent. In this case, the technique aids recovery of the planets by successfully cutting the sharp turning points of these light curves, which were previously leftover as false-transits after the LOWESS-smoothing step. It is these types of light-curve variability (approximately $3 < P_{\rm rot} < 8$ d; Amplitude >2.5 per cent) which are best-handled by using the peak-cutting technique. All light curves with rotation periods of 3–8 d, which were not improved by peak cutting, were later found to have had at least one of their peaks cut, however, their depths reached were not affected.

The other three light curves improved by the use of this technique (J0247-6808, J0346-6246, and J0414-7025) were aided accidentally, having sections of increased scatter masked as a result of

the applied cuts. This is a convenient side effect of searching for planets among light curves with intrinsically higher scatter, but less scientifically interesting.

The limitations of this technique were identified by investigating the small selection of stars detrimentally affected by the peak cutting. The three main failure modes were as follows:

(i) effectively flat light-curves dominated by scatter;
(ii) light curves with activity/variability periods of ∼2 d; and
(iii) HIP 33737 – a star with a flat-bottomed rotation activity.

In the first case, all peak-cut does is remove useful data, since no significant activity-based peaks and troughs were present. Meanwhile in the 2-d activity/variability period case, a significant portion of the data is cut, with the remaining intervals between each cut too short to be effectively flattened by the 20–30-bin (10–15 h) LOWESS smoothing. These two failure modes effectively represent the two limits of usefulness for this technique, of the order of 15 and 2 d, respectively. The final case (HIP 33737) was a unique one, where the light curve exhibited unusually flat-bottomed troughs well handled by the LOWESS-smoothing technique, and thus the peak cutting only removed useful data, alike to the first failure mode.

Crucially, however, for all light curves which did not fall into one of the three failure modes identified above, the use of the peak-cutting technique was not found to affect the shallowest depth of transit recovered. This alleviates the chief concern about the use of the peak-cutting technique: that the recovery depth may be reduced if transits near peaks and troughs are inadvertently cut. Overall then, this peak-cutting technique shows the greatest effectiveness for stars exhibiting activity/variability of periods 3–8 d, but may be applied to all stars with periods of approximately 3–15 d without significant detriment, unless those light curves are particularly flat.

4 DISCUSSION

As this work has shown, young host stars present many extra challenges in comparison to the generally older, less active stars previous exoplanetary searches have been biased towards. Quicker rotation, increased amplitude activity and other strong stellar-based periodicities wash out candidate planetary signals in BLS searches and make finding transiting signals harder using the traditional automated exoplanet-searching tools. Furthermore, by eye the large-amplitude variation of many sources such as HIP 105388 effectively hide real transit signals unless careful detrending is applied first.

None the less, the methods presented in this work have shown promise at pushing down to lower radius planets around young stars. The base LOWESS-detrending method provides a useful combination of smoothing and polynomial fitting and, as demonstrated by (Hippke et al. [2019]), generally outperforms more traditional exoplanet smoothing methods such as Savitzky–Golay filters for younger stars due to its weighted approach to smoothing. Furthermore, for sharp but evolving oscillations such as those seen for HIP 33235, cutting the peaks and troughs of these oscillations yields a significant improvement in the recovery of injected planets, especially for periods of 3–8 d. After the initial recovery, the shape and depth of the retrieved transit is then greatly improved by incorporating the developed activity interpolation over transit gaps.

However, as demonstrated by the sensitivity analysis, the large range in activity type, period and amplitude coupled with the variation in scatter for each source results in significantly different recovery rates for injected planets overall, ranging from sources like
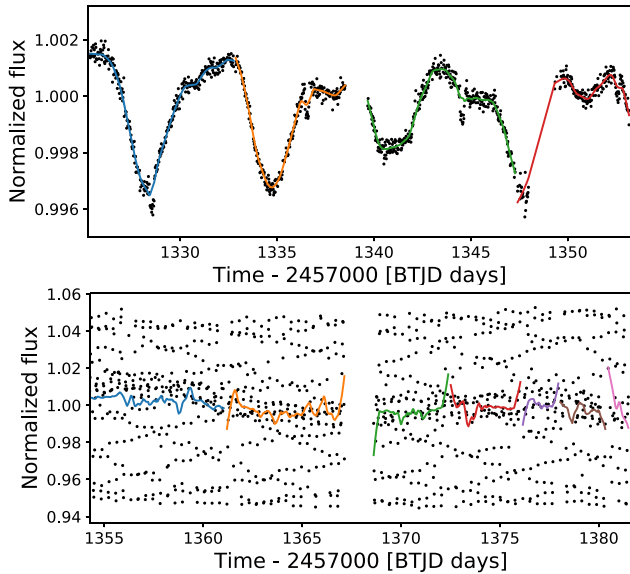
Figure 16. Light-curve comparison for HD 202969 (top panel) and CD-46 287 (bottom panel) with an 8.0 d, 0.03 R_P/R_* radius ratio injected planet. While the LOWESS-detrending pipeline has no trouble modelling HD 202969 due to its relatively long duration variation, it struggles to detrend the sharp 0.37-d period variability of CD-46 287.

CD-46 287 (a K6Ve star in the 35 Myr Octans association, viewed in Sector 2) where no planet was recovered even for the largest radius ratio, to HD 202969 (an F8/G0V star also from Octans, viewed in Sector 1), where the injected planet was easily recovered across the whole radius ratio range. Comparing these two light curves (shown in Fig. 16), the reason for this discrepancy is immediately clear; while the period of the rotational variation in the flux from HD 202969 is much longer than that of the transit duration (and easy to smooth with the 30-bin LOWESS-detrending technique), the rotation period of CD-46 287 is 0.37 d, well within the realm of the duration of a planetary transit. Since this pipeline is designed to only remove variation with periods longer than a transit it struggles to remove such short-period variations, and as such cannot recover injected planets from such light curves.

Observing other sources for which no injected planet sizes were ever recovered, the major challenges appear to be rotation with periods < 1 d, excess scatter, large-amplitude flares or other spurious outlying points in the light curve. These latter two problems present a problem to automatic detrending due to their unique light curve by light-curve behaviour. None the less these could conceivably be flagged by shape or sigma-clipping and handled on a case-by-case basis. On the other hand, scatter falls into two categories: time-localized scatter such as the scattered light glints seen at the end of some Sector 4 light curves, and non-localized scatter, which affects the entirety of the light curve. The first of these problems is soluble, and is likely best handled in a similar way to flares and instrumental scatter, by flagging and removing specific time periods which are affected. However, excess scatter which extends to the full extent of the light curve (which can be in excess of 2 per cent, as in the case of J0122-2566) likely precludes the use of these light curves for planet searches. This latter type of scatter was most common for very dim stars in the sample, largely with *TESS T* magnitudes in excess of 13.

The problem of fast rotation is a more systematic issue and is arguably the most important detrending-related challenge presented

by young stars. In this work, the detrending pipeline struggled to recover injected planetary signals in light curves with large-amplitude flux changes coupled with rotation periods of less than 1 d, largely due to the width of the LOWESS smoothing being at least 20 bins, or 10-h wide. Since any variability with a period of 1 d or less will involve at the very least one turning point, such widely spaced windows can not reasonably be expected to accurately trace and smooth such activity/rotation. Indeed, as discussed in Section 3.4, some of the fastest rotation observed had periods of less than 0.4 d, or less than 10 h. However, it is not unheard of for verified planets such as Kepler-1283 b (Morton et al. 2016) to have transit durations as long as 0.4 d, so dropping the LOWESS-smoothing window to any shorter than this could result in significant distortion of potential transit shapes along with the removal of stellar activity signals. It is thus clear that a more targeted method is required for such fast rotators, likely with more intelligent modelling-based methods such as Gaussian Processes (e.g. Gillen et al. 2019). Unfortunately, however, GP methods are much more intensive than a simple smoothing or polynomial-based methods, and require a more informed knowledge of a star's characteristics. This makes GP-based techniques very well suited to in-depth followup of candidates, but perhaps applying such methods to the entire stellar sample would be less efficient than simply flagging them in an initial BLS periodogram and analysing them separately. At these early stages, it is still worth considering other modelling methods as well (such as fitting a sum of sinusoids as in Gillen et al. 2019), especially since Hippke et al. (2019) found such a range in effectiveness in young star detrending methods. If a simpler algorithmic method can be found to model or detrend these very fast rotators then it could help to speed up wide-field transit searches around young, rapidly rotating stars.

Another interesting challenge highlighted by the sensitivity analysis is the rapid evolution of some young star light curves. This evolution makes modelling the light curves more difficult and also means that detrending is more effective in some regions of the light curve than others. As demonstrated for AB Pictoris, this changes how easy it is to recover injected (or no doubt real) exoplanet signals depending on when in the activity cycle one views the star. If, however, one gains a better understanding of a star's activity cycle through long-term monitoring and asteroseismology, this evolving activity can become an opportunity for increasing the effectiveness of planet searches by targeting quiet sections of the stellar activity cycle. This knowledge would be crucial for radial velocity follow-up of young host stars, since as Oshagh et al. (2017) have shown, for very active stars radial-velocity jitter is correlated with photometric variation in stellar light curves. Hence, being able to predict epochs of low stellar activity based on knowledge gained from photometric monitoring of these young stars will prove crucial for characterizing any discovered planets through radial-velocity follow-up.

One thing that was made quite clear from this initial survey of young star light curves is that the 'one size fits most' approach of large-scale photometric surveys such as *Kepler*, *K2,* and the main *TESS* SPOC pipeline is often not appropriate for young stars given the large range in shape, amplitudes, and periods of light-curve variability observed. Indeed, as discussed in Section 3.4, even in this relatively small number of young stars surveyed periods were observed to vary between 0.27 d to non-variable over the 27-d observation window. Meanwhile, activity and rotation behaviour varied from near-uniform to constantly evolving, and variability shape changed from beating sinusoids to almost flat aside from significant flaring activity. In order to more comprehensively search

for planets around such active stars, it is recommended that future detrending pipelines focus on more effectively matching detrending techniques to the primary type of light-curve variability observed. This could be achieved through an initial automated variability-type assignment, similar to – but more advanced than – the current 'variable' versus 'non-variable' assignment implemented into the *Kepler/K2/TESS* pipelines. By creating defined groups of similarly shaped light curves and types of variability/activity, machine-learning techniques could then be used to assign light curves the most effective detrending technique based on their perceived 'group' of variability. Particularly important will be separating known types of intrinsic variability, quick rotators and rapidly evolving light curves, however, many other important groups may become obvious with time. A more in-depth look into different types of variability in young stars over all sectors would thus be very beneficial, and may inform future detrending methods in missions like *PLATO* (Rauer et al. 2014).

Attempting to understand these stars in more detail raises the important question of whether 30-min data are enough to characterize any discovered variability or exoplanet candidate signals, or whether 2-min cadence data are required. In this work, 30-min cadence light curves were shown to effectively find the period of primary variability (and often second and third variability periods), and to recover all of the currently known TOIs identified through the 2-min data, so it is undeniably a very powerful data source for the general stellar sample. However, for those TOIs recovered, exposure smearing resulted in smaller transit depths and thus less accurate transit parameters compared to the 2-min data. This highlights the importance of the 2-min data (or alternative follow-up photometry) for accurate characterization of any discovered signals. Furthermore, while the 30-min data were sufficient for characterization of longer period variability and activity in the light curves, as the community pushes towards shorter period rotation (especially that with periods of less than 1 d), dissociating transit signals from stellar activity and rotation signals become increasingly difficult. This is the realm where 2-min data may be crucial in the search for young exoplanets. Furthermore, as the community attempts to understand the causes and evolution of activity in young stars, 2-min cadence data would significantly aid asteroseismological efforts for these young stars. However, a full comparison between 2- and 30-min data sets still needs to be undertaken before significant conclusions can be made.

One further major stumbling-lock remains in the way of finding significant numbers of exoplanets around young stars: so far, a large enough sample of 'bona-fide' young stars does not exist. Even considering basic transit statistics, only approximately 0.47 per cent of sun-like stars are expected to host Hot Jupiters (Haswell 2010), so considering that only 3076 stars exist in the extended BANYAN sample (some of which will not even be viewed by *TESS*'s primary mission), at most 15 such planets could be expected to be found around these young stars. Bouma et al. (2019) attempted to remedy this situation somewhat by concatenating 13 different catalogues of young stars and cluster members from literature, yielding 1 061 447 individual target stars. However, because of the mix of catalogues used and the non-homogenous membership criterion applied, this may include some stars that are not truly young. Furthermore, because it contains many members from within clusters, *TESS* will struggle to create accurate light curves for many of these stars due to blending on its relatively large 21 arcsec pixels. Another promising method that has come into its own with the release of a significant amount of data from the *Gaia* satellite (Gaia Collaboration et al. 2016, 2018) has been expanding cluster and association membership

through proper- and galactic-motion relationships. Groups such as Damiani et al. (2019) and Lodieu et al. (2019) have used these effectively already to expand the Sco-OB2 and Hyades associations, however, each of these stars still needs an independent sign of youth before it can be accepted as a 'bona-fide' member of a stellar association. This independent sign of youth is currently the biggest bottle-neck in expanding the known sample of young stars (especially since many confirmation methods require individual spectroscopic follow-up), so searching for new signs of youth in photometry or astrometry has the potential to have a significant impact.

In summary, although searching for planets around young, active stars undeniably presents several extra challenges compared to searching around older host stars, the techniques developed in this and other recent work on young stars are beginning to delve more effectively into this age range. However, two major stumbling-blocks remain before the search for young exoplanets can come into fruition: the relatively small number of bona-fide young stars and the very fast rotation of many young stars. Very important groundwork is thus still required to increase this sample of known young stars, and to develop techniques which can more effectively probe young potential host stars with shorter period stellar variability.

5 SUMMARY

In this work, techniques have been developed to aid the search for transiting exoplanets around young, active stars in the 30-min cadence *TESS* FFI data. Young exoplanets (<1 Gyr in age) inhabit a very important part of the exoplanet evolutionary time-scale, where formation mechanisms, accretion, migration, and dynamical interactions can significantly change the shape of observed planetary systems. However, they are also typically situated around young, active and often rapidly rotating host stars, severely hampering the discovery of new planets using the transit method. The developed method attempts to detrend these spurious stellar activity signatures using a 20–30 bin LOWESS method of Cleveland (1979b) at its base, combined with automated peak cutting and activity interpolation over transit gaps in order to more effectively differentiate activity from transit signals and preserve the transit shape. A basic version of this pipeline is made available online.[18] It is hoped that using this method in tandem with other detrending/transit-search pipelines such as the main *TESS* SPOC pipeline (Jenkins et al. 2016) and Gaussian-Process based methods (e.g. Gillen et al. 2019) will expand the number of young planets that can be found.

These techniques were applied to young stars in stellar associations from the extended BANYAN sample (Gagné & Faherty 2018; Gagné et al. 2018a, b), using the *TESS* Sector 1–5 light curves derived from the DIA pipeline of Oelkers & Stassun (2018). Lacking the data quality flags of the *TESS* 2-min cadence data, periods of excess pointing scatter were instead removed by considering the *TESS* data release notes and the engineering quaternion data.

While no new exoplanet candidates were revealed in this work, results from this initial survey revealed a variety of interesting objects, including the retrieval of the new young exoplanet DS Tuc Ab, TOI 447.01, TOI 450.01, a number of eclipsing binaries and a large array of interesting rotation and activity. In order to test the sensitivity of the developed detrending techniques to different planetary sizes, model `batman` transits (Kreidberg 2015) were injected into each of the young star light curves at a range of R_P/R_* radius ratios from

---

[18]https://github.com/mbattley/YSD

0.1 to 0.03 and periods from 1–14 d. The percentage of recovered transit signals from the injected planets dropped from 77.6 per cent at a radius ratio of $0.1 R_P/R_*$ to 20.4 per cent at a ratio of 0.03, however, was seen to vary considerably between different targets and sectors. Meanwhile, while increasing the injected planet period was seen to result in a decreasing recovery rate for smaller planets, the recovery rate was actually observed to peak around periods of 2–6 d for larger planets. An investigation into the relationship between rotation period and recovery depth did not suggest that the two were significantly correlated, aside from a slight skew towards larger planets at short rotation periods and the known difficulty of very short-period (<1 d) rotation.

These results alongside deeper examination of light curves in this sample lead to a number of interesting conclusions. The sensitivity of specific young star light curves to transit searches appears to be most limited by fast rotation (<1 d rotation periods), excess scatter, scattered light glints and significant flaring activity. Meanwhile the rapid evolution of many young star light curves offers both a detrending challenge and a potential opportunity to search more efficiently for exoplanets at less active times. Given the vast array of different types of young star light curves seen in this initial survey, in the future, a multifaceted detrending approach which first classifies light curves according to their broad activity/variability is perceived as beneficial. It is clear that the 30-min cadence data show particular promise for additional detections to the main SPOC 2-min pipeline, as it is capable of retrieving all of the TOIs highlighted in the sectors analysed. However, planet parameters derived from the 30-min light curves can be less reliable due to phenomena such as exposure smearing. Thus the acquisition of 2-min light curves for young stars remains desirable, especially since this will also aid in-depth asteroseismic characterization of these sources.

Overall, the two largest challenges remaining before significant progress can be made in the field of young exoplanet science are those of very rapidly rotating stars and the relatively small numbers of confirmed young stars. Solving these two problems has the potential to gift the exoplanetary community with a significant advancement in understanding of how exoplanets form and develop into stable exoplanetary systems.

REFERENCES

Adams F. C., Laughlin G., 2006, ApJ, 649, 1004
Aigrain S., Hodgkin S. T., Irwin M. J., Lewis J. R., Roberts S. J., 2015, MNRAS, 447, 2880
Aigrain S., Parviainen H., Pope B. J., 2016, MNRAS, 459, 2408
Armstrong D. J. et al., 2015, A&A, 579, A19
Astropy Collaboration et al., 2018, AJ, 156, 123
Bakos G., Noyes R., Kovacs G., Stanek K., Sasselov D., Domsa I., 2004, PASP, 116, 266
Bakos G. . et al., 2010, ApJ, 710, 1724
Baraffe I., Chabrier G., Barman T. S., Allard F., Hauschildt P. H., 2003, A&A, 402, 701
Benatti S. et al., 2019, A&A, 630, A81
Borucki W. J. et al., 2010, Science, 327, 977
Bouma L. G., Hartman J. D., Bhatti W., Winn J. N., Bakos G., 2019, Astrophys. J. Suppl. Ser., 245, 22
Brems S. S., Kürster M., Trifonov T., Reffert S., Quirrenbach A., 2019, A&A, 632, A37
Briceno C. et al., 2018, AJ, 157, 85
Caceres G. A., Feigelson E. D., Babu G. J., Bahamonde N., Christen A., Bertin K., Meza C., Curé M., 2019, AJ, 158, 57
Cegla H. M., Watson C. A., Shelyag S., Mathioudakis M., Moutari S., 2019, ApJ, 763, 95
Ciardi D. R. et al., 2018, AJ, 155, 10
Cleveland W. S., 1979a, J. Am. Stat. Assoc., 74, 829
Cleveland W. S., 1979b, J. Am. Stat. Assoc., 74, 829
Cody A. M., Barentsen G., Hedges C., Gully-Santiago M., Dotson J., Barclay T., Bryson S., Saunders N., 2018, Res. Notes AAS, 2, 199
Collier Cameron A. et al., 2006, MNRAS, 373, 799
Costes J. C. et al., 2019, MNRAS, 491, 2834
Damiani F., Prisinzano L., Pillitteri I., Micela G., Sciortino S., 2019, A&A, 623, A112
David T. J., Petigura E. A., Luger R., Foreman-Mackey D., Livingston J. H., Mamajek E. E., Hillenbrand L. A., 2019a, ApJ, 885, L12
David T. J. et al., 2016, Nature, 534, 658
David T. J. et al., 2019b, AJ, 158, 79
Esplin T. L., Luhman K. L., 2019, AJ, 158, 54
Feinstein A. D. et al., 2019, PASP, 131, 1003
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2012, PASP, 125, 306
Gagné J., Faherty J. K., 2018, ApJ, 862, 138
Gagné J., Faherty J. K., Mamajek E. E., 2018c, ApJ, 865, 136
Gagné J., Roy-Loubier O., Faherty J. K., Doyon R., Malo L., 2018b, ApJ, 860, 43
Gagné J. et al., 2018a, ApJS, 856, 23
Gaia Collaboration, Brown A. G. A., Vallenari A., Prusti T., de Bruijne J. H. J., Babusiaux C., Bailer-Jones C. A. L., 2018, A&A, 616, A1
Gaia Collaboration et al., 2016, A&A, 595, A2
Gillen E. et al., 2019, MNRAS, 492, 1008
Handberg R., Lund M. N., 2019, Technical report, T'DA Data Release Notes: Data Release 4 for TESS Sectors 1+2. TASOC, Aarhus
Hartman J. D., Bakos G., 2016, Astron. Comput., 17, 1
Haswell C. A., 2010, Transiting Exoplanets , 1st edn., Cambridge Univ. Press, Cambridge
Hippke M., David T. J., Mulders G. D., Heller R., 2019, AJ, 158, 143
Howell S. B. et al., 2014, PASP, 126, 938
Ida S., Lin D. N., 2010, ApJ, 719, 810
Ida S., Lin D. N., Nagasawa M., 2013, ApJ, 775, 42
Jackman J. A. G. et al., 2019, MNRAS, 489, 5164
Jenkins J. M. et al., 2010, ApJ , 713, L87

Jenkins J. M. et al., 2016, Software and Cyberinfrastructure for Astronomy IV. Proc. SPIE, Edinburgh, p. 99133E

Kim D. W., Protopapas P., Alcock C., Byun Y. I., Bianco F. B., 2009, MNRAS, 397, 558

Koeltzsch A. et al., 2009, Astron. Nachr., 330, 482

Korhonen H., Andersen J. M., Piskunov N., Hackman T., Juncher D., Jarvinen S. P., Joergensen U. G., 2015, MNRAS, 448, 3038

Kovács G., Bakos G., Noyes R. W., 2005, MNRAS, 356, 557

Kovács G., Zucker S., Mazeh T., 2002, A&A, 391, 369

Kreidberg L., 2015, PASP, 127, 1161

Kuhn M. A., Hillenbrand L. A., Sills A., Feigelson E. D., Getman K. V., 2019, ApJ, 870, 32

Lightkurve Collaboration et al., 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library, record ascl:1812.013

Lodieu N., Smart R. L., Pérez-Garrido A., Silvotti R., 2019, A&A, 623, A35

Lomb N., 1976, Ap&SS, 39, 447

Luger R., Agol E., Kruse E., Barnes R., Becker A., Foreman-Mackey D., Deming D., 2016, AJ, 152, 100

Luger R. et al., 2017, Nat. Astron, 1, 0129

Lund M. N., Handberg R., Davies G. R., Chaplin W. J., Jones C. D., 2015, ApJ, 806, 30

Lund M. N., Handberg R., Kjeldsen H., Chaplin W. J., Christensen-Dalsgaard J., 2017, in Eur. Phys. J. Web Conf., p. 1005

Malavolta L., et al., 2016, A&A, 588, A118

Manara C. F., Mordasini C., Testi L., Williams J. P., Miotello A., Lodato G., Emsenhuber A., 2019, A&A, 631, L2

Mandel K., Agol E., 2002, ApJ, 580, L171

Mann A. W. et al., 2016, ApJ, 818, 46

Marley M. S., Fortney J. J., Hubickyj O., Bodenheimer P., Lissauer J. J., 2007, ApJ, 655, 541

Mascareño A. S., Rebolo R., Hernández J. I. G., 2016, A&A, 595, A12

Morton T. D., Bryson S. T., Coughlin J. L., Rowe J. F., Ravichandran G., Petigura E. A., Haas M. R., Batalha N. M., 2016, ApJ, 822, 86

Newton E. R. et al., 2019, ApJ, 880, L17

Oelkers R. J., Stassun K. G., 2018, ApJ, 156, 132

Oelkers R. J., Stassun K. G., 2019, Res. Notes AAS, 3, 8

Oshagh M. et al., 2017, A&A, 606, A107

Owen J. E., 2019, Annu. Rev. Earth Planet. Sci., 47, 67

Paulson D. B., Cochran W. D., Hatzes A. P., 2004, AJ, 127, 3579

Pepper J. et al., 2007, PASP, 119, 923

Plavchan P. et al., 2020, Nature, in press

Pollacco D. L. et al., 2006, Astronomical Society of the Pacific, 118, 1407

Prialnik D., 2009, An introduction to the Theory of Stellar Structure and Evolution. 2nd edn., Cambridge Univ. Press, Cambridge

Quinn S. N. et al., 2012, ApJ, 756, L33

Quinn S. N. et al., 2014, ApJ, 787, 27

Rajpaul V., Aigrain S., Osborne M. A., Reece S., Roberts S. J., 2015, MNRAS, 452, 2269

Rauer H. et al., 2014, Exp. Astron., 38, 249

Ricker G. R. et al., 2014, J. Astron. Telesc. Instrum. Syst., 1, 014003

Rivilla V. M., Chandler C. J., Sanz-Forcada J., Jiménez-Serra I., Forbrich J., Martín-Pintado J., 2015, ApJ, 808, 146

Rizzuto A. C., Mann A. W., Vanderburg A., Kraus A. L., Covey K. R., 2017, AJ, 154, 224

Rizzuto A. C., Vanderburg A., Mann A. W., Kraus A. L., Dressing C. D., Agüeros M. A., Douglas S. T., Krolikowski D. M., 2018, AJ, 156, 195

Rodenbeck K., Heller R., Hippke M., Gizon L., 2018, A&A, 617, A49

Saar S. H., Donahue R. A., 1997, ApJ, 485, 319

Sato B. et al., 2007, ApJ, 661, 527

Savitzky A., Golay M. J. E., 1964, Anal. Chem., 36, 1627

Schlichting H., Sari R., Yalinewich A., 2015, Icarus, 247, 81

Sergison D. J., 2015, Technical report, Untangling the signals: Investigating accretion and photometric variability in young stars An observational analysis. Univ. of Exeter, Exeter

Siverd R. J. et al., 2012, ApJ, 761, 123

Spiegel D. S., Burrows A., 2012, ApJ, 745, 761

Stassun K. G. et al., 2018, AJ, 156, 102

Stassun K. G. et al., 2019, AJ, 158, 138

Storn R., Price K., 1997, J. Glob. Optim., 11, 341

Tamuz O., Mazeh T., Zucker S., 2005, MNRAS, 356, 1466

Taylor M., 2005, in Astronomical Data Analysis Software and Systems XIV. ASP, San Francisco

The Astropy Collaboration et al., 2013, A&A, 558, A33

Thompson S. E., Fraquelli D., Van Cleve J. E., Caldwell D. A., 2016, Technical report, Kepler Archive Manual. NASA, Baltimore

Torres C. A. O., Quast G. R., Da Silva L., De La Reza R., Melo C. H. F., Sterzik M., 2006, A&A, 460, 695

Vanderburg A., Johnson J. A., 2014, PASP, 126, 948

Vanderburg A. et al., 2016, ApJS, 222, 14

Vanderburg A. et al., 2019, ApJ, 881, L19

Vanderspek R. et al., 2018, Technical report, TESS Instrument Handbook. NASA, Cambridge

Virtanen P. et al., 2020, Nature Methods, 17, 261

Werner M. W. et al., 2004, ApJS, 154, 1

Wheatley P. J. et al., 2013, Proc. Int. Astron. Union, 475, 311

Wheatley P. J. et al., 2018, MNRAS, 475, 4476

Zari E., Brown A. G., De Zeeuw P. T., 2019, A&A, 628, A123

Zuckerman B., 2018, ApJ, 870, 27

Zuckerman B., Webb R. A., 2000, ApJ, 535, 959

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table 4. Full sensitivity analysis table for each of the 256 stars with DIA light curves.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

APPENDIX A: SENSITIVITY ANALYSIS FOR EACH SECTOR

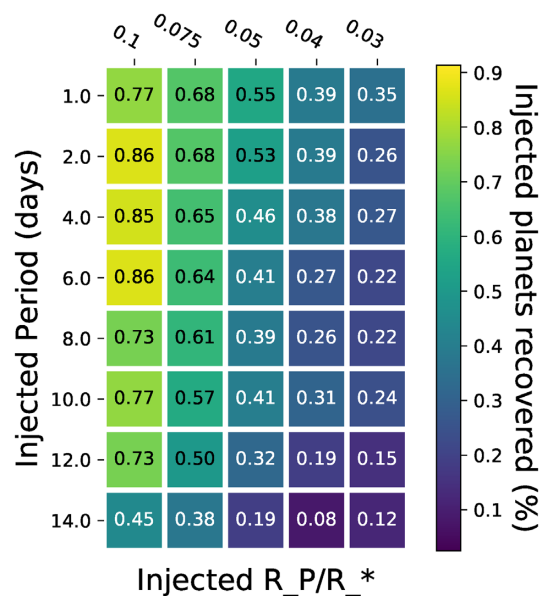A sector by sector breakdown for the conducted sensitivity analysis

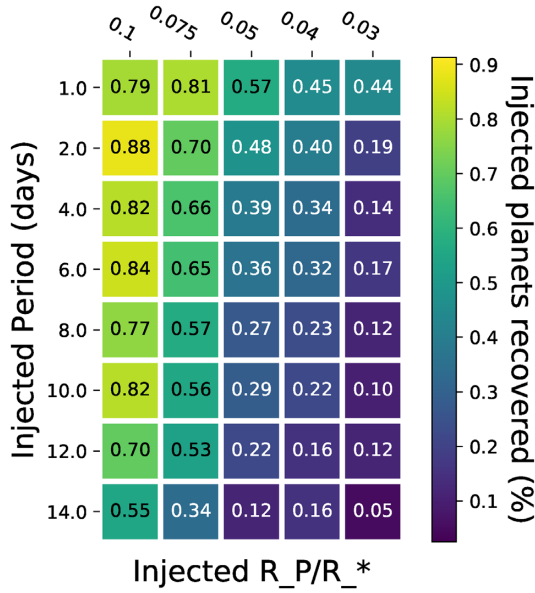**Figure A1.** Sensitivity analysis for Sector 1. Total number of sources = 74.

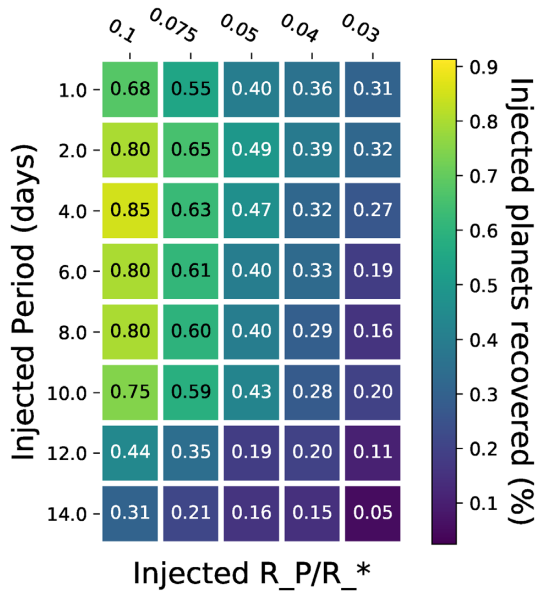**Figure A2.** Sensitivity analysis for Sector 2. Total number of sources = 77.

Figure A4. Sensitivity analysis for Sector 4. Total number of sources = 120.

Figure A3. Sensitivity analysis for Sector 3. Total number of sources = 75.
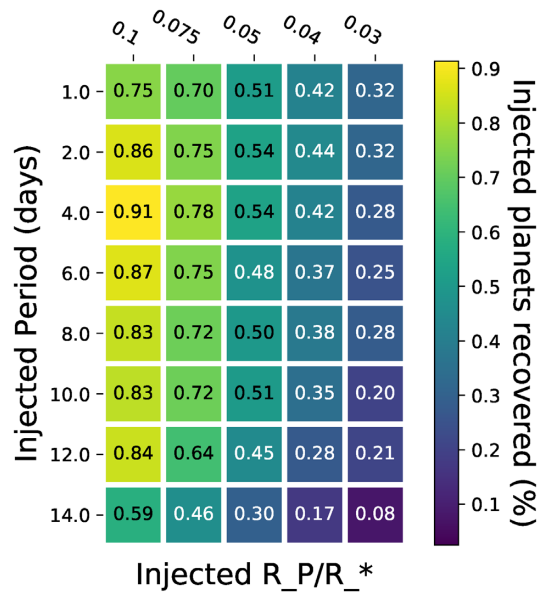
Figure A5. Sensitivity analysis for Sector 5. Total number of sources = 138.

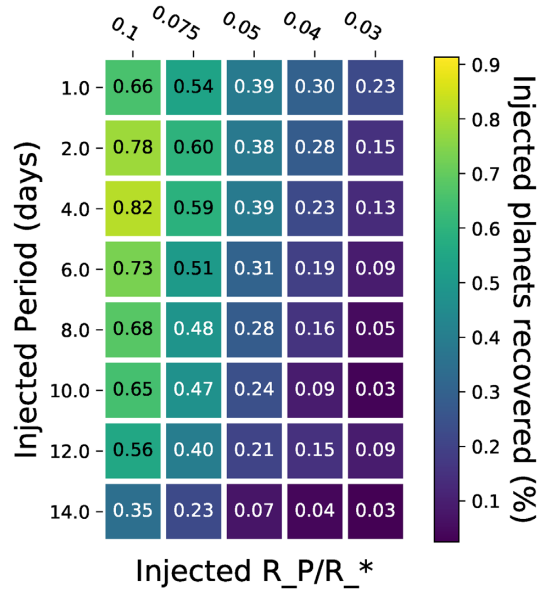is given below in Figs A1–A5. Figure format after that explained in Fig. 14.

This paper has been typeset from a TEX/LATEX file prepared by the author.