Mémoire

Auteur : Limpens, Magali
Promoteur(s) : Delrez, Laetitia; Pozuelos Romero, Francisco José
Faculté : Faculté des Sciences
Diplôme : Master en sciences spatiales, à finalité approfondie
Année académique : 2022-2023
URI/URL : http://hdl.handle.net/2268.2/18769

FACULTY OF SCIENCES

DEPARTMENT OF ASTROPHYSICS, GEOPHYSICS, AND OCEANOGRAPHY

Searching for multi-planetary systems around M dwarfs in TESS data

Magali Limpens

A THESIS PRESENTED IN ORDER TO OBTAIN THE MASTER DEGREE IN SPACE SCIENCES

Supervisors: Francisco J. Pozuelos, Laetitia Delrez
Reading commitee: Lauriane Soret, Valentin Christiaens, Michaël Gillon

ACADEMIC YEAR 2022-2023

Acknowledgement

First of all, I would like to sincerely thank my two supervisors, Laetitia Delrez and Francisco Pozuelos, who gave me the opportunity to work for a whole year on a fast-growing subject: exoplanets. Thank you for your time and patience throughout the year, you have always been available to answer all my questions and improve this work. Your passion for this shared subject has helped me to stay motivated despite the many weeks without expected results. This experience gave me a glimpse into the world of scientific research. Thanks to your meticulous proofreading of this work, I have become more concise and rigorous.

Secondly, I would like to thank my friends and especially Sorenza, who helped me stay motivated and brought me a lot of joy this last year. I would also like to thank my family and Théo for their support throughout my years at university, through my ups and downs. Thank you to my parents for encouraging me and allowing me to continue my studies.

Finally, I want to thank my reading committee, Lauriane Soret, Valentin Christiaens and Michaël Gillon, for taking the time to read this master's thesis.

Contents

1   Introduction

The search for new exoplanets, or planets outside the solar system, is a rapidly evolving field in astronomy. In the past few decades, their discovery has revolutionized the way we apprehend the universe, and a new race to find life outside the solar system has begun. The first exoplanet around a solar-type star was discovered in 1995, orbiting around the star 51 Pegasi (Mayor, Queloz, 1995), and since then, thousands of exoplanets have been discovered using a variety of techniques. The most successful one is the transit method, for which almost 4,000 out of the 5,000 confirmed exoplanets have been found[1]. This method consists in searching for periodic decreases in the brightness of a star when a planet is passing in front of it and its first success was in 2000 with the star HD 209458 (Charbonneau et al. 1999; Henry et al. 1999). Since then, the number of transiting exoplanets discovered has risen sharply with the launch of NASA's Kepler space telescope in 2009 (Borucki et al., 2010), followed by the TESS (Transit Exoplanet Survey Satellite) mission in 2018 (Ricker et al., 2015). Those two space missions were designed to measure precisely the amount of light received from thousands of stars with the goal of detecting exoplanets using the transit method.

The search for new exoplanets is leading to a better understanding of the formation of planetary systems. Finding such systems in all stages of their evolution and around different types of stars allows us to better understand how they form and evolve, and also to understand whether planets are a common phenomenon in the universe or not. Moreover, finding new rocky planets in a region around their star where liquid water could be found on their surfaces is a key step towards finding and understanding new forms of life outside our solar system.

The goal of this master's thesis is to identify new multi-planetary systems by searching for new transiting planets in systems that have at least one planet or planet candidate already detected by the automatic TESS pipeline. Those systems are of special interest for three reasons. First, the detection of a transiting planet around a given star enhances the probability that other planet candidates in the system, if any, are also real transiting planets and not due to false positive detection (Lissauer et al. (2012a); Guerrero et al. (2021)). Second, by comparing the properties of the different planets within the same system, such as their internal compositions and their atmospheres, one can see how these parameters evolve with

---

[1]https://exoplanetarchive.ipac.caltech.edu/docs/counts_detail.html

the orbital distance to the star. Moreover, as these planets were formed from the same parent protoplanetary disk, it allows us to have better constraints on their properties and, thus, their formation and evolution. Third, detecting new planets in a system allows us to have a more complete picture of the architecture of those systems and study them in their entirety.

Before going further, some preliminary explanations are needed: a planet candidate is an object for which observational measurements are still required to firmly confirm its planetary nature. The TESS pipeline was developed by the Science Processing Operations Center (SPOC, Jenkins et al. 2016) at NASA Ames Research Center with the goal to provide corrected data for each TESS target star. Moreover, the TESS pipeline automatically searches for evidence of periodic transit events in the light curves, which are measurements of the brightness of a star as a function of time. Once a candidate is detected by the TESS pipeline it is referred to as TESS Objects of Interest (TOI). In addition, the pipeline also provides plots of these light curves for visual inspection.

Using the definition of those concepts, we plan to search for multi-planetary systems, by analyzing the TESS light curves of a selected number of M dwarfs. Indeed, planets orbiting M dwarfs are of special interest, as will be explained in section 1.3. By scrutinizing these light curves, we will search for signals that may hint at the presence of other transiting planets possibly missed by the TESS pipeline because of their low signal-to-noise ratio (SNR). The SNR is a dimensionless number defined here as the ratio of the transit signal to the white noise in the phase-folded light curve. Hence, the higher the SNR value, the stronger the transit signal compared to the noise. The TESS pipeline detects signals with SNRs ≥7.1. Here, we will look for signals that can have an SNR as low as 5. This can happen for planets with small radii, resulting in a shallow transit (explained in section 1.1), and/or planets with long orbital periods, for which there are only a few transits in the TESS data. To detect and vet those signals, we used the open-source Python package called SHERLOCK (Pozuelos et al., 2020), described in detail in section 3. Finally, if a signal looks promising after the analysis of the light curves by SHERLOCK, we will use the SPECULOOS (Search for habitable Planets EClipsing ULtra-cOOl Stars, Gillon 2018, Delrez et al. 2018) southern or northern observatories to further observe the candidate, and maybe, confirm the planetary origin of the signal.

The thesis is divided as follows: in section 1, is an introduction to the transit method, the

TESS mission, M dwarfs, and the SPECULOOS ground-based observatories. In section 2, is a description of the selection of our targets. In section 3, is a detailed explanation of the SHERLOCK Python package. In section 4, is a presentation of the obtained results. Finally, in section 5, the conclusion of the master's thesis is presented.

1.1   The transit method

The transit method aims to detect periodic drops of brightness in a stellar light curve. When an exoplanet passes in front of its host star, it hides a part of the stellar photosphere for a certain amount of time. This phenomenon is called a transit and is represented in figure 1. As the planet blocks some stellar light, the observed stellar flux decreases. A second drop of brightness can be observed when the planet passes behind the star, hiding the reflected stellar light of the exoplanet in the visible domain. This phenomenon is called an occultation and is also represented in figure 1. Note that, if the occultation of a planet is observed in the infrared (IR) domain, the thermal emission of the planet's dayside can then be detected.
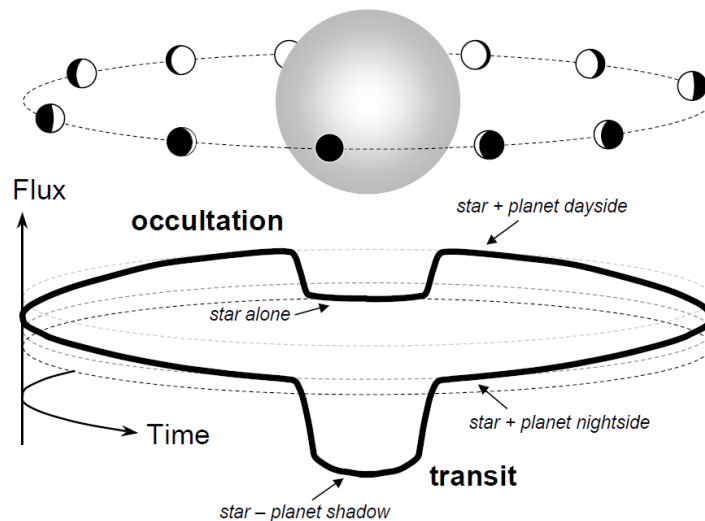

Figure 1: Illustration of a transit and an occultation when an exoplanet is orbiting a star (Winn, 2010).

Some parameters can be obtained directly from the transits light curve, such as:

1. P, the period between two transits, which determines the period of revolution of the planet around its host star.

3

2. Different durations, shown in figure 2. Those are the total transit duration, given by $T_{tot} = t_{IV} - t_I$ and the full duration, given by $T_{full} = t_{III} - t_{II}$. Then, there are the ingress and egress durations, respectively defined as $\tau_{ing} = t_{II} - t_I$, and $\tau_{egr} = t_{IV} - t_{III}$. Finally, there is the transit duration T, from mid-ingress to mid-egress.

3. δ, the transit depth, which can be approximated by $\delta \approx (R_p/R_*)^2$ (Winn, 2010), with R_p, the radius of the planet and R_*, the radius of the star.

From these parameters, others can be derived:

1. b, the impact parameter. It is the sky-projected distance (in units of stellar radius) between the centre of the stellar disc and the centre of the planetary disc at conjunction, i.e., when there is an observer-planet-star alignment. A value of $b = 0$ means that the disk of the planet is passing right in front of the centre of the star. It is derived from the following equation (Winn, 2010):

$$b^2 = \frac{(1 - \sqrt{\delta})^2 - (T_{full}/T_{tot})^2(1 + \sqrt{\delta})^2}{1 - (T_{full}/T_{tot})^2} \tag{1}$$

Moreover, two cases are observed depending on the value of this impact parameter.

(a) Full eclipse:
The surface of the exoplanet entirely covers a part of the star. The expected value of b is given by:

$$b \leq 1 - (R_p/R_*) \tag{2}$$

(b) Grazing eclipse:
A part of the exoplanet does not cover the star, as seen from the Earth's point of view (for a specific line of sight). In this case, the planet passes partially in front of the disk of the star. The expected value of b is then given by:

$$1 - (R_p/R_*) < b < 1 + (R_p/R_*) \tag{3}$$

In the case of a transit, the impact parameter is also given by:

$$b_{transit} = \frac{a \cos i}{R_*} \left(\frac{1 - e^2}{1 + e \sin w} \right) \tag{4}$$

With e, the eccentricity, i, the inclination and w, the argument of periastron, i.e., the angle from the planet's ascending node to its periastron, measured in the direction

4

of motion. Finally, for a circular orbit, b is given by assuming a null eccentricity in equation 4:

$$b = \frac{a \cos i}{R_*} \tag{5}$$

2. R_*/a, the scaled stellar radius, with a, the semi-major axis of the planet's orbit. This ratio can be approximated, for the case of a non-grazing transit and if $R_p \ll R_* \ll a$, by (Winn, 2010):

$$\frac{R_*}{a} = \frac{\pi}{2\delta^{1/4}} \frac{\sqrt{T_{tot}^2 - T_{full}^2}}{P} \left(\frac{1 + e \sin w}{\sqrt{1 - e^2}} \right) \tag{6}$$

From the knowledge of this ratio and given the equation 4, the cosine of the inclination can be approximated. Note that the scaled stellar radius can also be estimated from the value of the full transit duration (Günther, Daylan, 2021):

$$T_{full} = \frac{P}{\pi} \sin^{-1} \left[\frac{R_*}{a} \frac{\sqrt{(1 - R_p/R_*)^2 - b^2}}{\sin i} \right] \cdot \frac{\sqrt{1 - e^2}}{1 + e \sin w} \tag{7}$$

3. In the case where $(R_p/R_*)^3$ is negligible, ρ_*, the stellar density, can be approximated using Kepler's third law and the inverse of the scaled stellar radius. It is given by (Winn, 2010):

$$\rho_* \approx \frac{3\pi}{GP^2} \left(\frac{a}{R_*} \right)^3 \tag{8}$$

4. a is estimated as follows:

$$a = R_*/(R_*/a) \tag{9}$$

If one supposes that R_* is known.

5. In the same way, the planet's radius R_p can be obtained from the ratio R_p/R_* (derived from the transit depth) if R_* is known.

6. Finally, the equilibrium temperature of the planet is given by (Günther, Daylan, 2021):

$$T_{eq;p} = T_{eff;*} \cdot \frac{(1 - A)^{1/4}}{E} \cdot \sqrt{\frac{R_*}{2a}} \tag{10}$$

With $T_{eff;*}$, the effective temperature of the host star and A, the albedo and E the emissivity which are approximated, respectively, by 0.3 and 1.
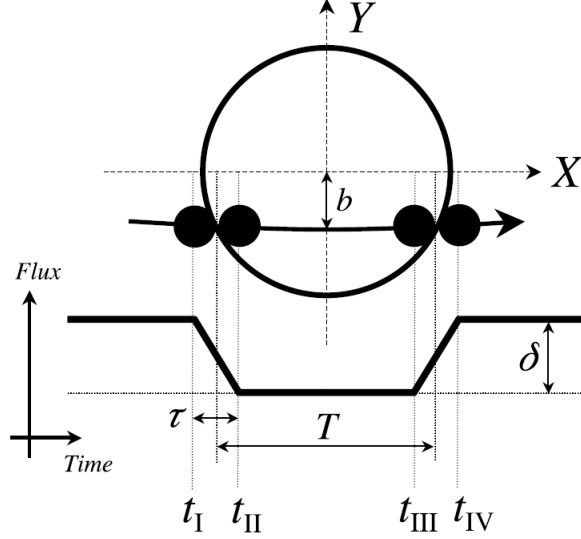
Figure 2: Illustration of a planetary transit plotted as a function of time for which the flux from the planet's night side is negligible (Winn, 2010).

The fractional flux deficit at mid-transit can be approximated by the square of the planet-to-star radius ratio, expressed with R_p and R_*, and corresponding to the ratio of the projected areas of the planet and star. Hence, the observations of a transit provide us with a direct measurement of the planet's radius relative to the radius of the star.

$$\frac{\Delta f}{f} \simeq \left(\frac{R_p}{R_*}\right)^2 \tag{11}$$

Given the typically much larger stellar radius compared to the planet's one, the drop in brightness can be very small. For Jupiter-sized and Earth-sized planets transiting across a Sun-like star, the transit depth is $\sim 1\%$ and $\sim 0.01\%$, respectively (Cameron, 2016). The probability to observe a transit in the case of a grazing or full eclipse and assuming a circular orbit is given by combining equations 3 and 5 and if one assumes that $R_* \gg R_p$, the probability is given by :

$$Pr\left(\cos i < \frac{R_*}{a}\right) = \frac{R_*}{a} \simeq 0.0046 \left(\frac{R_*}{R_\odot}\right)\left(\frac{1AU}{a}\right) \tag{12}$$

Therefore, transits of the Earth are only observable for 0.46% of the celestial sphere when using $R_*=1R_\odot$ and $a=1$AU. It reduces to only 0.09% in the case of Jupiter because of its larger semi-major axis of 5.2 AU. From this, one can conclude that the larger the radius of

6

the star and the closer the planet is to its host star, the higher the probability to observe the transit. Thus, it explains why the transit method favors the detection of planets with short orbital periods. Moreover, it is also attributed to the shorter waiting time required to detect a transit, in the case of a short orbital period.

The flux of a star also varies across the stellar disk because of the limb darkening effect. It is caused by the fact that the centre of a star appears brighter than the limb. Indeed, when the centre of a star is observed, deeper and hotter layers are detected. This effect is related to the optical depth given by (Gillon, 2021-2022) :

$$\tau_\nu(x) = -\int_\infty^x \rho(s)\kappa_\nu \, ds \tag{13}$$

where x is the position of the considered layer of the stellar photosphere along the line of sight from an observer at infinity, ρ is the density, κ the absorption coefficient and ν the frequency. The emission seen by the observer comes from the layer where the optical depth is equal to 1 (to a first approximation) because a larger value means that the medium is opaque. As at the centre, the layer for which the emission surface reaches $\tau = 1$ is found deeper within the star than at the limb, it produces an effect on the shape of a transit light curve. Indeed, ingress and egress parts will be curved. Moreover, if the planetary disc covers bright regions of the stellar disc, the flux drop can be larger than the one given in equation 1.1 near the centre of the transit. Similarly, if the planetary disc covers the limb of the stellar disc, masking dimmer regions, the drop in flux can be smaller than the theoretical value given in equation 1.1.

1.2 TESS

Data for this research are coming from the TESS (Transit Exoplanet Survey Satellite) mission. This is an all-sky transit survey, which means that it observes the whole sky by dividing it into several sectors (Ricker et al., 2015). It was launched on the 18th April 2018 and will operate until at least October 2024. The primary goal of this mission is to search for planets transiting bright and nearby stars, focusing on main-sequence stars with spectral types between F5 and M5. By July 2023, TESS discovered 363 confirmed exoplanets and 6687 planet candidates (https://exoplanets.nasa.gov/tess/). The design of this mission was adopted for several reasons, some of which will be listed now and are based on the article of Ricker et al. (2015).

1. **Sky coverage:**
   Since the main objective is to observe bright stars that are nearly evenly distributed over the entire sky, TESS observes the different sectors one by one, resulting in an all-sky coverage. This is done by using 4 cameras with a total field of view of 24 deg x 96 deg, i.e., 2300 deg^2. The celestial sphere is divided into 2 hemispheres and 26 observation sectors, 13 by hemisphere. Each sector is observed for 27.4 days (i.e., during two orbits of the satellite), and due to overlapping between sectors, some parts of the celestial sphere are observed for a longer period of time. Indeed, sectors increasingly overlap at higher latitudes, giving an almost continuous coverage near the ecliptic poles (corresponding to JWST (James Webb Space Telescope) continuous viewing zone). Therefore, the duration of observation of a part of the celestial sphere varies between 27 and 351 days. An illustration is shown in figure 3. It takes 2 years to observe the entire sky (1 year per hemisphere).

2. **Period detection range:**
   The minimal period (P_{min}) observable for an exoplanet was selected to be near the Roche limit of a few hours, while the maximum period (P_{max}) was selected to be around $\approx$ 10 days and $\gtrsim$ 40 days for large areas surrounding the ecliptic poles. It was adopted to detect planets in the habitable zone of M stars. The habitable zone is defined as the distance range to a star in which a terrestrial planet with a suitable atmosphere could sustain liquid water on its surface.

3. **Detector and bandpass:**
   In order to ensure high-precision photometry, silicon charge-coupled devices (CCDs) were selected with a pixel size of 21". This large pixel size may result in contamination

8
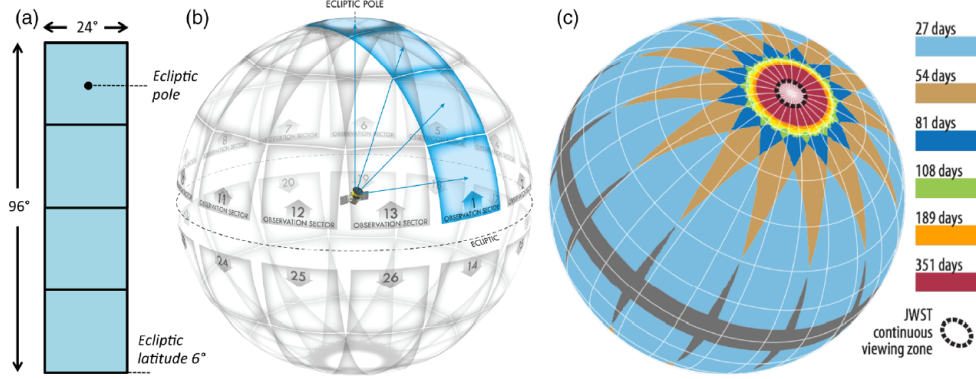
Figure 3: On the left, a representation of the 4 TESS cameras. The middle figure represents the division of the celestial sphere into different sectors. On the right, the overlapping between sectors is shown (Ricker et al., 2015).

of the target's photometry due to nearby stars, leading to false positives like blended eclipsing binaries. Moreover, the band-pass in the optical domain was chosen to be between 600 and 1000 nm to be more sensitive to red wavelengths. Indeed, it increases the precision for small stars, which are cooler and redder than bigger ones. Indeed, small stars are of particular interest as will be explained in section 1.3.

4. **Time sampling:**

   For the nominal mission, 200 000 selected stars were observed at a cadence of 2 minutes. In addition, full-frame images (FFI) were collected at a cadence of 30 minutes. For the extended missions, a new 20-second cadence mode has been introduced and the FFI cadence has been changed to 600 seconds (first extension of the TESS mission) and then 200 seconds (second extension).

The TESS mission was supposed to last for 2 years but was already extended twice. Originally scheduled to conclude in July 2020, the primary mission was first extended until September 2022 and then further extended until at least October 2024 (NASA 2023a[2]). During its extended mission, TESS is conducting observations of both the southern and northern hemisphere sectors that were previously observed in the primary mission. Additionally, it is observing new sectors along the ecliptic [3].

---

[2]https://heasarc.gsfc.nasa.gov/docs/tess/second-extended.html
[3]https://tess.mit.edu/observations/

After observations are made, photometric data products are available. The first files are the *Full Frame Images (FFIs)*. These files contain all pixels available across all CCDs for one camera. During the primary, the first and second extended missions, FFIs were taken, respectively, every 30 minutes, 10 minutes, and 200 seconds. Three types of FFIs files are provided: uncalibrated, calibrated, and uncertainty (NASA, 2023b). The TESS FFI files contain a primary HDU (Header Data Unit) with metadata stored in the header. Then, with two extensions of the HDU, the full frame image and its uncertainty are stored [4]. Then, there are the *Target Pixel Files (TPFs)*, which are the raw data around the targets. The goal is to provide the data necessary to perform photometry on the raw or calibrated data when needed. The value of each pixel is encoded as an image. In this case, this is done for each 2-minute or 20-second cadence target in an observing sector (NASA, 2023b). Finally, there are the *Light Curve Files*, which contain flux time series data. These are produced every 2 minutes or 20 seconds, using a Single Aperture Photometry (SAP). Those data are available on the NASA Mikulski Archive for Space Telescope (MAST, https://archive.stsci.edu/index.html). In figure 4, those 3 files are represented.

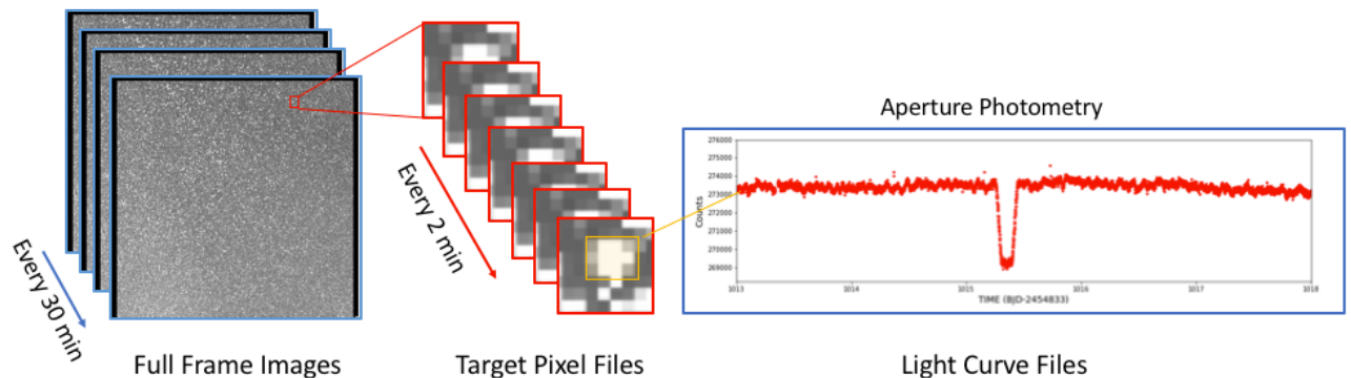Figure 4: On the left, the FFIs are represented for the TESS cameras. At the centre are the TPFs, recorded every 2 minutes. On the right, the Light Curve files using SAP are shown. Figure obtained from https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html.

---

[4]https://github.com/spacetelescope/notebooks/blob/master/notebooks/MAST/TESS/beginner_how_to_use_ffi/beginner_how_to_use_ffi.ipynb

1.3   M dwarfs

In this research, we decided to focus on M dwarfs for our choice of targets. Those stars are very abundant in the Milky Way, as they constitute 70% of all stars (Shields et al., 2016). They are classified from spectral type M9 to M0 according to their effective temperature, respectively between 2350 and 3850K, with masses between 0.08 and 0.57 $M_{\odot}$[5]. Due to their low mass, M dwarfs have lower pressure and temperature in their cores. Consequently, hydrogen burns at a slower rate than higher-mass stars such as Sun-like or brighter stars, increasing their lifetimes in the main-sequence phase to trillions of years for the lowest-mass ones (Shields et al., 2016). Those stars are well-suited to exoplanet searches for several reasons listed below.

1. Planet detectability:
   Planets orbiting M dwarfs produce a larger observational signature with the transit method. Indeed, the small size of the host star leads to a larger proportion of the stellar light being blocked by the planet, increasing the transit depth (Shields et al., 2016). Hence, transits of small terrestrial planets are more easily detected around M dwarfs than around larger stars such as Sun-like stars. Moreover, transiting planets around M dwarfs are more suitable for atmospheric characterization than those around larger stars, allowing for follow-up observations, notably with the Hubble Space Telescope or the James Web Space Telescope (JWST) (Gillon, 2021-2022).

2. Abundance:
   As mentioned above, M dwarfs constitute 70% of the stars in our Galaxy and small planets with a radius between 1.0 and 2.8 $R_{\oplus}$ are 3.5 times more abundant around M dwarfs than larger stars (of spectral type FGK) (Shields et al., 2016).

3. Habitability:
   Due to their low luminosity, the habitable zone (defined in section 1.2) is relatively close to the host star. As explained in section 1.1, the closer the planet is to its host star, the higher the probability of transit. Hence, the transits of potentially habitable planets are more likely and frequent for planets around M dwarfs compared to other bigger stars. However, several factors act against the habitability of those systems. First, as the habitable zone is close to the host star, planets can be synchronous rotators due to tidal locking. Therefore, the planet would consistently exhibit the same side towards its

---

[5]https://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txtl

host star, while the other side would freeze, leading to a possible atmospheric collapse (Joshi et al., 1997). Second, M dwarfs are magnetically very active, more than Sun-like stars, with stellar spots on their surfaces and flares. Those latter consist of a strong emission in the UV and X-ray domains of charged particles that would hit the potential planetary atmosphere, causing its erosion and biological damage. However, biological adaptations to such an environment could be developed, and subterranean organisms could survive UV radiation (Heath et al., 1999). Stellar spots on M dwarfs have a larger extension than for the Sun, and they can reduce the stellar insolation received by the planet by up to 40% for several Earth months. As a consequence, stellar spots are thought to have similar effects to seasons on Earth (Joshi et al., 1997).

1.4   SPECULOOS

In order to confirm that the transit signal is of planetary origin, additional observations are required if a candidate is identified with SHERLOCK (see section 3). These observations will be conducted using SPECULOOS, a ground-based transit survey with observatories at the Paranal Observatory in Chile and the Teide Observatory in Tenerife (Gillon 2018, Delrez et al. 2018). SPECULOOS aims at performing high-precision photometry to detect planetary transits around 'ultra-cool dwarfs' (UCD), which are stars with effective temperatures below 2700 K, luminosities less than 0.1% of the Sun's, and spectral types of M7 and later. UCDs have masses between 0.07 $M_\odot$ and 0.1 $M_\odot$, with those below 0.08 $M_\odot$ being considered brown dwarfs as they do not have nuclear reactions in their core. The primary objective of SPECULOOS is to determine the frequency of terrestrial planets around UCDs, while also investigating their atmosphere and potential habitability. To achieve this objective, the SPECULOOS observatories consist of a network of 1-meter-class robotic telescopes, with cameras sensitive in the near-infrared (Delrez et al., 2018). In conclusion, for this work, the SPECULOOS telescopes are well suited to follow up any interesting candidates found around an M dwarf.

2 Choice of targets

The goal of this master thesis is to search for multi-planetary systems so we were looking for stars with already one planet or planetary candidate discovered. To select the targets to analyse, the ExoFOP (Exoplanet Follow-up Observing Program) website was used[6]. On this website, a list of all TOIs can be found and is frequently updated. On this list, several parameters for each TOI are available. Among those parameters is the TIC-ID (identified in the TESS Input Catalog; Stassun et al., 2019) which is the star identification number in the stellar catalogue containing the information regarding all the stars selected to be observed by the mission. Other parameters that can be found are notably the magnitude of the host star, the orbital period of the planet candidate, and its transit depth. On the 17$^{\text{th}}$ of July, more than 6600 TOIs are listed.

To have a reasonable number of targets to analyse, some criteria were chosen based on relevant parameters.

1. TESS disposition: CP (confirmed planet) OR PC (planetary candidate)

2. TFOPWG (TESS Follow-up working group) disposition: contains CP OR contains PC

3. TESS magnitude: ≤ 12

4. Stellar effective temperature: ≤ 3900 K and >0 K.

The condition on the effective temperature was applied in order to define an M dwarfs sample. The condition on the magnitude was applied to be able, in the case of a detection, to do a follow-up observation with the SPECULOOS telescopes (explained in section 1.4). Applying those filters led to a list of 129 TOIs that were sorted to have the more recent observations (corresponding to the latest sectors) at the beginning. The aim was to increase our chances of identifying new candidates that had not previously been discovered by other teams in the scientific community.

---

[6]https://exofop.ipac.caltech.edu/tess/

3 SHERLOCK

`SHERLOCK` stands for Searching for Hints of Exoplanets fRom Lightcurves Of spaCe-based seeKers (https://github.com/franpoz/SHERLOCK; Pozuelos et al., 2020) It is an end-to-end pipeline that uses light curves from space-based missions to search for planetary transit signals. The goal is to retrieve planetary signals that could have been missed by automatic pipelines due to their low signal-to-noise ratio (SNR), detection thresholds or poor photometry. For this research, we decided to set the minimum value accepted for the SNR to 5, allowing for signals missed by the TESS pipeline to be retrieved. It can for example be the case for planets with a small radius, resulting in a shallow transit, or for planets with a long orbital period for which there are only a few transits in the data. Additionally, faint stars will exhibit lower SNR because the precision of the TESS data will be less accurate. The idea is to first analyse the global light curve from a star and detect a periodic signal that could be attributed to a planet. Then, this signal is masked and a new periodic signal is sought. If found, the new signal is also masked and the algorithm runs again to find other new signals. Although the number of runs can be infinite, it is better to set the maximum number of runs to 5. Indeed, each time a periodic signal is detected, it will be masked for the next run, creating gaps in the light curves. As a consequence, after each iteration, confidence in the detection decreases because the number of data points decreases. Six modules are used by `SHERLOCK` to confirm that a signal can be attributed to a planet. Those will be described now. Throughout this section, examples will be taken from the results of the target star TIC 52368076, which has three known planets, but we will focus only on the innermost one, TOI-125 b (Quinn et al., 2019), which has an orbital period of 4.65 days. If other examples are used, it is to illustrate particular cases.

3.1 Downloading and preparing the light curves from their online repositories

The first module downloads and prepares the light curves for the selected targets. To achieve this, there are two Python open-source packages that `SHERLOCK` uses. First there is `Lightkurve` (https://docs.lightkurve.org; Lightkurve Collaboration et al., 2018), used for short-cadence (20 and 120 seconds) and long-cadence data (600 and 1800 seconds), processed by the TESS Science Processing Operations Center (SPOC; Jenkins et al., 2016). Alternatively, for the 1800s cadence, SHERLOCK also provides access to the data processed by the `Eleanor` pipeline (https://github.com/afeinstein20/eleanor; Feinstein et al., 2019),

corresponding to the FFIs explained in section 1.2. Moreover, `Lightkurve` can provide periodograms of time series data. With those packages, TESS data from the MAST are obtained and the Pre-search Data Conditioning Simple APerture (PDCSAP) flux is used. The PDCSAP light curve is submitted to some treatments that will remove short trends and noisy regions while keeping planetary transits intact. PDCSAP fluxes are thus provided by one of the two packages mentioned above and must be corrected with the second module of `SHERLOCK` for some remaining trends and systematics.

3.2 Search for planetary candidates

The second module can be divided into two parts. One is about the launch of `SHERLOCK` by filling a correct properties file. Then, this file is executed a first time by `SHERLOCK` to help the user decide if some parameters need fine-tuning and check how many sectors and which cadences are available. This is the *explore* mode, which is usually fast. The second part of this module will be devoted to the search for planetary signals, that is, the *analysis* mode, which, depending on the number of sectors to be analysed, might have a high computational cost.

3.2.1 The properties file

The first thing to do before launching `SHERLOCK` is to prepare the properties file (which is a yaml file). Numerous input parameters can be given. The ones that were used in this work will be detailed now.

1. **TARGETS:** The star identification number is given in the form of a TIC-ID (identified in the TESS Input Catalog; Stassun et al., 2019) which is the stellar catalogue containing the information regarding all the stars selected to be observed by the mission. In the case of this research, 56 different stars were inputted.

2. **MODE:** There are three different possibilities:

   - Sector: allows us to test if a given signal is consistent in all the sectors analysed independently.

   - Global: allows us to combine all the data from each sector available and then find small and/or long-period planets. Indeed, as data are available over a longer period of time (27 days per sector), longer orbital periods can be sought.

- Both: runs the two other options. It is more time-consuming, but it provides a deeper analysis of the data.

3. **EXPTIME:** It is the inputted value for the data cadence we want to use. Usually, an exposure time of 120 seconds was set, but it can also be 20, 600, or 1800 seconds, depending on the available data.

4. **AUTO_DETREND_ENABLED:** This parameter is used for stars exhibiting clear and strong periodic variability, as is the case for fast rotators. To remove the variability that can hinder the transit search, SHERLOCK performs a Lomb-Scargle periodogram. If the parameter is set to True, SHERLOCK will correct the variability corresponding to the strongest peak in the periodogram. Hence, an initial detrending of the light curves is applied. If there is no need, the parameter must be set to False.

5. **INITIAL_SMOOTH_ENABLED:** This parameter is related to the local noise. If very close measurements exhibit a high deviation from the local trend, a Savitzky-Golay filter (Savitzky, Golay, 1964) can be applied to smooth the data and reduce the local noise. This method fits successive sub-sets of equally spaced points with a low-degree polynomial (of degree 3 by default) using a least-squares calculation. The resulting analytical solution is then applied to all subsets. Hence, the real data point will be replaced with the point found with the analytical solution of the polynomial. Consequently, this method can alter transit signals and must be used carefully. The parameter can be set to True to apply the filter, or it can be disabled and set to False.

6. **INITIAL_HIGH_RMS_MASK:** The light curves often exhibit regions with high data dispersion, which significantly reduces the photometric quality and strongly affects the performance of the searching algorithm. These noisy regions can be attributed to various factors such as spacecraft motion, scattered light caused by the Earth, background moving objects, or unidentified sources. Therefore, to remove this troublesome data and clean the light curve, SHERLOCK uses this parameter to automatically calculate the root-mean-square (RMS) of the dataset using four-hour binning. Subsequently, any data point that falls within a region with 1.5 times the mean RMS is masked (refer to figure 6). The default threshold of 1.5 is applied, but the user has the flexibility to adjust it as needed in the yaml file.

7. **DETREND_L_MIN:** This parameter is the minimum detrend window to build the detrends set used for the detrending method (explained below).

8. **DETREND_L_MAX:** This parameter is the maximum detrend window to build the detrends set.

9. **DETRENDS_NUMBER:** It corresponds to the number of detrend models to be generated from the original light curve.

10. **MAX_RUNS:** The maximum number of times SHERLOCK will run. As explained above, at each new run, the planetary signal found is masked, so the maximum number usually does not exceed 5.

11. **SNR_MIN:** This is the minimum value accepted for the SNR. As explained above, since we are looking for signals missed by the TESS pipeline, which has a detection threshold for the SNR of 7.1, the minimal value chosen for this search was always set to 5.

12. **SDE_MIN:** This is the minimum value accepted for the signal detection efficiency (SDE). As for the SNR, this value will be kept small and set to 5. The SDE will be defined in detail below but very briefly, the SDE value of a given period quantifies the statistical significance of this period P compared to the mean significance of all other periods.

13. **PERIOD_MIN:** This is the minimum period used for the planet search. It was usually set to 0.4 day.

14. **PERIOD_MAX:** This is the maximal period value used for the planet search. This number depends on the number of sectors available for a given target. Combining different sectors allows searching for planets with longer orbital periods as the star was observed for a longer period of time (27 days per sector), and at least two transits are required to define the period.

15. **INITIAL_TRANSIT_MASK:** In the case of multi-planetary systems, the signal of already-known planets or planet candidates can be masked. The period (in days), the transit duration (in minutes), and the epoch in TBJD (TESS Barycentric Julian Date) must be given as inputs. This is helpful to keep the number of runs low and decrease the running time.

An example of a properties file is shown in figure 5. As explained above, some parameters must be adapted depending on the data, for example, in the case of strong variability

with fast rotators. To determine the best value for the different parameters, we use the second module with the mode *explore*. For this step, it is important to set the parameters AUTO_DETREND_ENABLED and INITIAL_HIGH_RMS_MASK to "True" to decide if they must be kept or not afterwards. Once it has been run, different files associated with the target are created.

```
TARGETS:
    TIC 52368076:
     MODE: GLOBAL
     #SECTOR: [1,2,28]
     EXPTIME: 120
     AUTO_DETREND_ENABLED: False
     INITIAL_HIGH_RMS_MASK: False
     INITIAL_SMOOTH_ENABLED: True
     INITIAL_HIGH_RMS_THRESHOLD: 1.3
     DETREND_L_MIN: 0.2
     DETREND_L_MAX: 1.2
     DETRENDS_NUMBER: 10
     DETREND_CORES: 10
     CPU_CORES: 60
     MAX_RUNS: 4
     SNR_MIN: 5
     SDE_MIN: 5
     PERIOD_MIN: 0.4
     PERIOD_MAX: 20.0
```

Figure 5: Example of a properties file for the TIC 52368076. The line sector is commented as it is not needed with the global mode that automatically runs all the sectors available.

Before analysing those files, the method used by SHERLOCK to detrend the light curves must be understood. This is the bi-weight method and it was used throughout this work. The goal is to remove instrumental and/or stellar noise while preserving the transit signals. Different algorithms are available for removing trends in light curves and are included in a python package called Wötan (https://github.com/hippke/wotan; Hippke et al., 2019). Those algorithms can be used with SHERLOCK by changing an input parameter in the properties file but the bi-weight method was the only one used for this work as it is one of the most efficient algorithms compared to the others available in Wötan.

The goal is to deal with outliers data points using M-estimators which is a class of central tendency measures that are more robust than the mean (or the median). A mean can be highly affected by outliers points, so the goal is to have a method resistant to local misbehaviour caused by those outliers. The idea behind the Tukey bi-weight is to attribute

a certain weight (between zero and one) to each data point relative to their distance from the centre of the data set. So M-estimators attribute a larger weight to points close to the central tendency and a smaller weight to those far from the central tendency (typically for the outliers). The median is the first estimate of the central location. Then, the M-estimator becomes the new measure of central tendency for the next iteration (George W. Dombi, 2012). Its value is computed as follows:

$$M_{estimator} = \frac{\sum_i w_i \times x_i}{\sum_i w_i} \tag{14}$$

with x_i, the data point and w_i, the weight attributed to that data point. An estimator can be distinguished by its loss function, and for the Tukey bi-weight (Hippke et al., 2019), it is given by:

$$L(a) = \left(- \left(\frac{a}{c} \right)^2 \right)^2 \tag{15}$$

This is only valid if c < 0, otherwise, the loss function is zero. In this equation, we have:

1. a: The residuals, i.e. the difference between the data points and the mid-point.

2. c: This is a tuning parameter. This value is typically between 4 and 6 and is set by default to 5 in Wötan.

The loss function evaluates if the algorithm accurately models the data, and it must be as low as possible. In other words, the goal of the iterative process is to minimize this function. The bi-weight method is a time-windowed slider. Indeed, to compute the M-estimator, a time-window filter is applied. It is defined with a window length w centred on $t(x_i)$ where t is the time. The choice of window size is a trade-off between removing local trends (caused by stellar variability or systematics) and protecting the transit-like signal. Indeed, the window size must be short enough to remove stellar variability or systematics efficiently but, at the same time, long enough to preserve the transits. In practice, the time window size should be larger than the transit duration T_{14}. However, as we do not know the orbital period of the planets we are searching for (and thus their transit duration), we tested 10 different window sizes, between 0.2 and 1.2 days, with a step of 0.1 day. In this way, we tested short and long window sizes, increasing our confidence in the results.

Now that the concept behind the removal of instrumental and stellar noise has been explained, results can be interpreted. Different files have been created to help the user decide

the best setup to conduct the transit search.

The first file to analyse is the *RMS file*. As explained above in the properties file, if the RMS value is above a certain threshold, corresponding to noisy regions, a mask can be applied to those values. In figure 6, a case where the mask should be applied is shown. One can see that in this figure, RMS values that exceed the threshold are in red. If the INITIAL_HIGH_RMS_MASK option is enabled, then the red dots will be masked. When the RMS doesn't exceed the threshold value, no mask is needed.
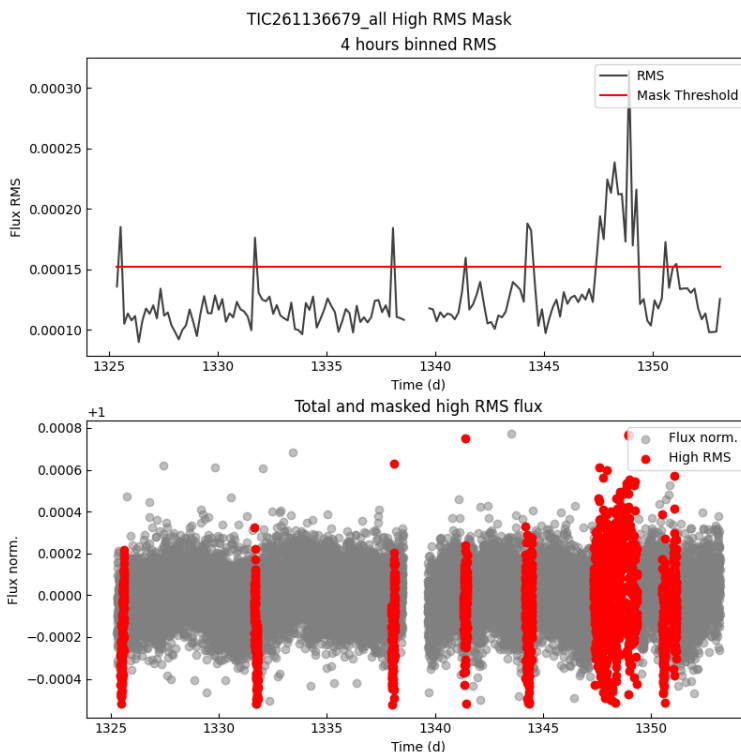
Figure 6: Example of an RMS file for TIC 261136679. In this case, the RMS values that are above the threshold value are depicted in red. In this case, a periodic signal would be lost if the mask was applied. Hence, it is necessary to modify the threshold and conduct manual masking in the range of dates from ∼1347 to ∼1349.

The second folder to analyse is the *fov* one, which stands for field of view. It is a file representing the photometric aperture and also the stars in the vicinity of the target star. This allows us to know if other stars can contaminate the signal measured from our target star. If one or more stars are fainter than the target star, the search for exoplanets can

continue. On the contrary, if there is a contamination of the signal by a large number of stars in the fov or if a nearby star is brighter than the target star, one must analyse the photometry to decide whether to proceed to the next module or not. Those two cases are shown in figure 7, where the red-bordered pixels show the aperture used to extract the photometry, while the nearby stars are shown as red circles. The size of these circles is proportional to the magnitude contrast with the target. As mentioned in section 1.2, the pixel scale is 21".
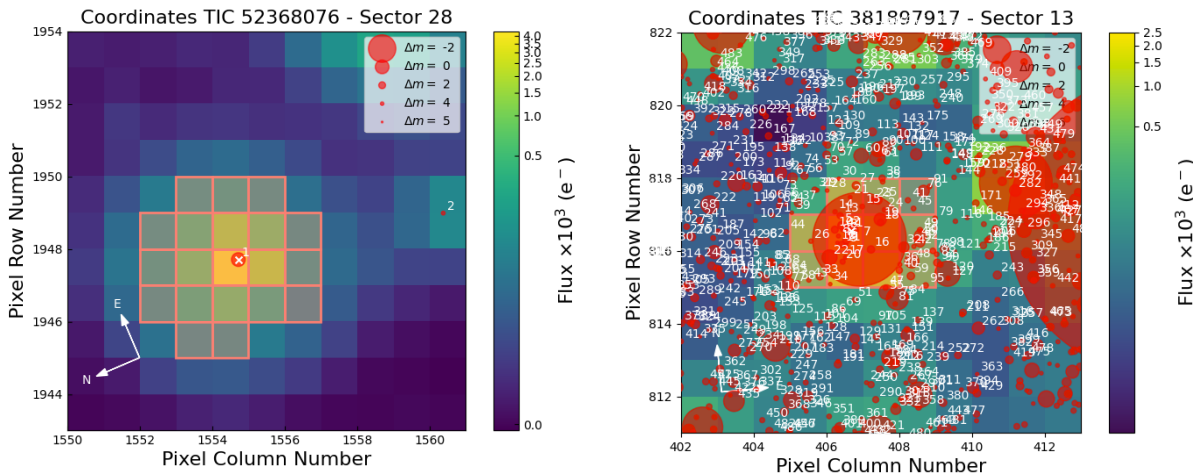

Figure 7: The left plot shows the TESS target pixel file of TIC 52368076, which was notably observed with sector 28. There is one faint star on the right side of the figure but the Δm, the difference in magnitude between this nearby star and the target star, is large enough. In this case, the signal should not be contaminated by this star. The right plot shows the TESS target pixel file of TIC 381897917, which was observed with sector 13. In this case, a specific attention must be given to the photometry. In both cases, the intensity of the flux is represented on the right side of each plot.

The third file to analyse is the Lomb-Scargle periodogram (example in figure 8) along with the file named *Phase_detrend_period* which is a phase-folded light curve that represents the flux as a function of the phase in Julian days. Firstly, SHERLOCK identifies the strongest peak in the Lomb-Scargle periodogram and then, it folds the light curve over the period corresponding to this peak. If the signal remains flat over this period, no initial detrending must be applied. On the contrary, for a strong periodic signal, the parameter AUTO_DETREND_ENABLED must be set to True. It will perform an initial detrending of the light curve by correcting for the strongest peak in the Lomb-Scargle periodogram. It can be the case with fast rotators, stellar spots, and stellar pulsations. This first detrend

allows us to manage with strong variability when the PDCSAP light curves exhibit a strong periodic variability. An example of the two possibilities is shown in figure 9.
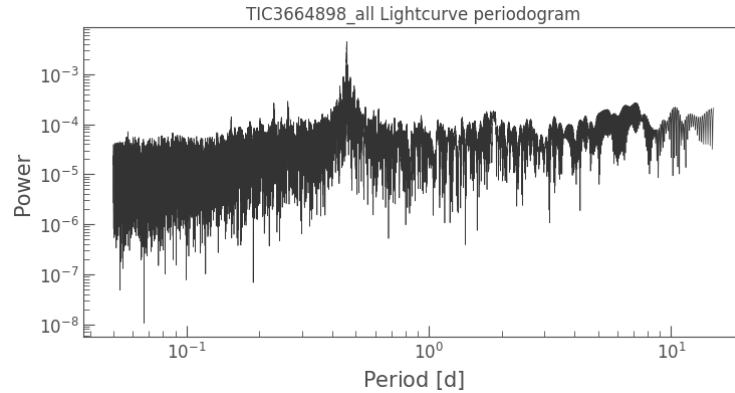

Figure 8: Example of a Lomb-Scargle periodogram for TIC 3664898. The y and x axis represent, respectively, the power as a function of the period (in days). One strong peak is visible at 0.46 days and must be corrected with the parameter associated with 'Auto_detrend_enabled'.
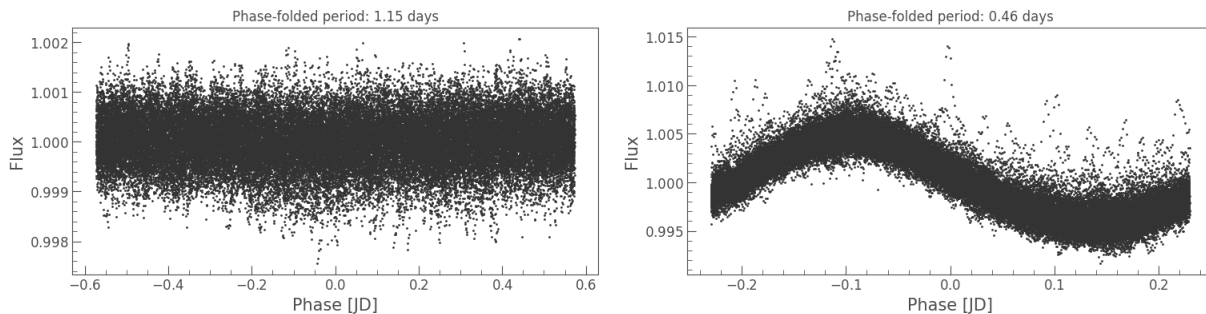
Figure 9: On the left side, the phase-folded light curve for a period of 1.15 days, for TIC 52368076 is represented. There is no strong periodic signal in the curve, meaning we do not have to use the AUTO_DETREND_ENABLED option for this target. On the right side, a typical case of a fast rotator with strong periodic variability is shown. In this case, an initial auto-detrend has to be applied to remove this periodic variability. For this example, it is a phase folded light curve of TIC 3664898 over a period of 0.46 days.

Finally, there is the file named *detrends*. This file represents the flux as a function of time for the different detrending window sizes inputted in the properties file. For each window size, at the top, there is first the original flux (grey dots), with a red line representing the detrend

applied to the signal. Below it, the detrended flux is shown (black dots). This allows us to see if each window size is worth keeping or if it should be changed. Indeed, if the light curve exhibits a lot of short variabilities, larger window sizes are not suitable to correctly detrend those variabilities. Moreover, if the file *fov* exhibits numerous stars in the vicinity of the target star, this file will determine whether to proceed or not. Indeed, if the photometry is well corrected with the detrending method, the search for planetary signals may proceed. An example is shown in figure 10.

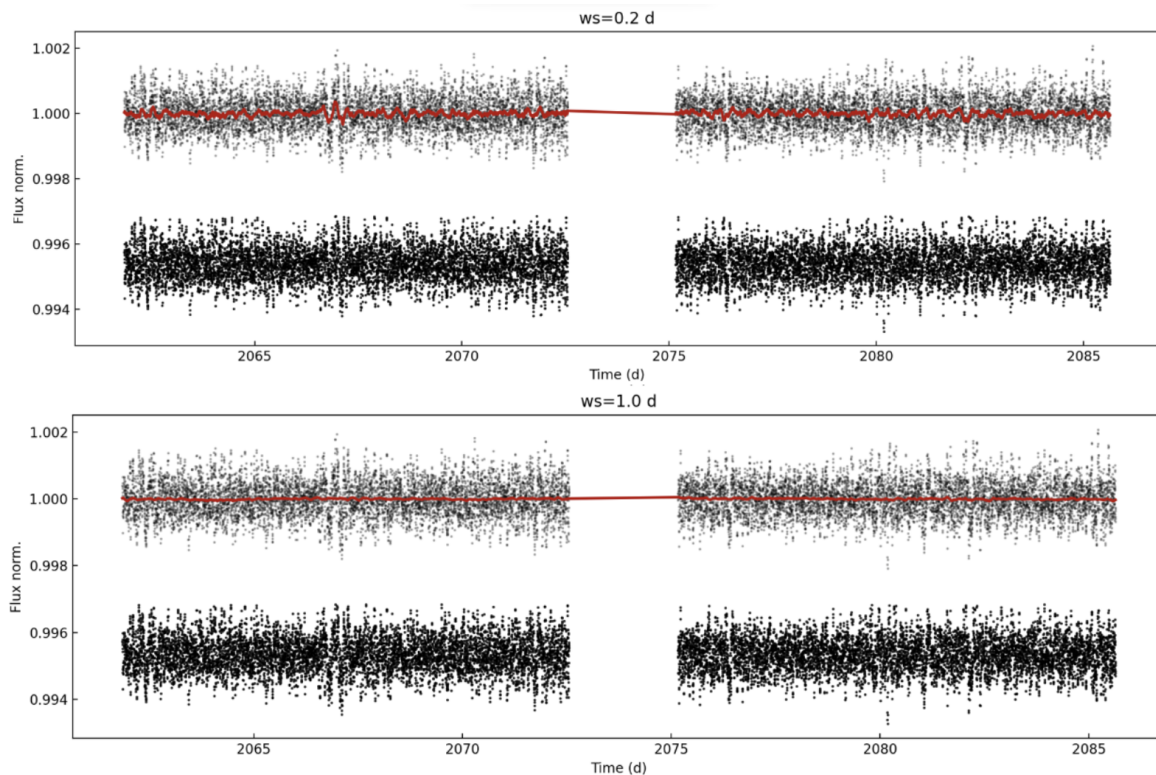Figure 10: Example of the detrends file for TIC 52368076, sector 28 with 2 different window sizes. In this figure, one can see that the detrend is much more important for the short window size of 0.2 days, while it is smoother for the longer window size of 1.1 days.

Once all those files have been properly analysed, the properties file must be adapted if needed and the search for planetary signals can begin.

3.2.2 Search for planetary signals

The second step of SHERLOCK is the search for planetary signals using the *analysis* mode. Once stellar noise is removed with the Tukey bi-weight method explained above, the search for planetary signals can begin. This is achieved with the transit least squares (TLS) algorithm, which searches for periodic transit-shaped signals in the light curve while taking into account the limb darkening and planetary ingress and egress. TLS is optimized for both computational speed and for the SDE of small planets. The idea is to use a transit-like search function instead of a box (as with the box least squares (BLS) algorithm). To this end, TLS uses the transit model light curve of Mandel, Agol (2002). The TLS is specially optimized for the detection of small planets in large datasets. The TLS algorithm acts by phase-folding the data over a range of trial periods (P), transit epochs (t_0) and transit durations (d). Then it computes the χ^2 of the folded curves (Hippke, Heller, 2019):

$$\chi^2(P, t_0, d) = \sum_{i=1}^{N} \frac{y_i^m(P, t_0, d) - y_i^0}{\sigma_i^2} \tag{16}$$

with y_i^m, the data points of the transit model, y_i^0, the observed values and σ_i^2, the standard deviations in the light curve. Then, for each trial period, it performs a minimization of the χ^2 ($\chi^2_{min,glob}$) for all possible combinations of (t_0, d). From this, the signal residue (SR) is obtained with the following equation:

$$SR(P) = \frac{\chi^2_{min,glob}}{\chi^2_{min}(P)} \tag{17}$$

With, $\chi^2_{min}(P)$, the distribution of minimum χ^2 as a function of P. Finally, the SDE is retrieved:

$$SDE(P) = \frac{1 - \langle SR(P) \rangle}{\sigma(SR(P))} \tag{18}$$

with $\langle SR(P) \rangle$, the arithmetic mean of the signal residue and $\sigma(SR(P))$, its standard deviation. Finally, the systematic noise component (inherent to the SDE distribution) is removed using a walking median filter through the SDE(P) periodogram. A typical value for a transit detection is SDE > 6 (determined empirically; Hippke, Heller, 2019). An SDE below 6 results in a higher rate of false positives.

At this stage, depending on the number of runs implemented in the properties file, periodic

signals found with the TLS method at each run will be masked for the next run. In some cases, it can be interesting to directly hide the signal of already known planets. In this case, a new parameter must be added to the properties file (INITIAL_TRANSIT_MASK).

After the running of the *analysis* mode, in addition to the directory shown above, another directory is available containing:

- A directory for each run, containing the plots of the detrended fluxes, the transit found, and the SDE values. An example is shown in figure 11.

- A file named 'TICID_[sector]_report.log', which is a report containing all the information available on the transit search.

In the report, the initial parameters inputted in the properties file, the different exposure times available for each sector and a summary of all the runs can be reviewed. For each run, different parameters associated with the possible transit signal are listed in a table. Those parameters are:

- The period (in days)

- The error on the period

- The number of transits in the data

- The mean depth of the transit in ppt (part per thousand)

- The duration (in minutes)

- The SNR

- The SDE

- The border score (BS), a number between 0 and 1, which takes into account that when a transit signal is masked, it creates a gap in the light curve, and if a new transit is detected at the edge of this gap, the border score will decrease. This comes from the fact that borders are regions with more uncorrected systematics. The higher the border score, the more one can be confident in the transit signal.

- The planet's radius (R_p, in Earth's radius)

- The R_p/R_s ratio, R_s being the stellar radius

- The semi-major axis (in astronomical units)

- The habitability zone, which indicates whether the planet is in the habitable zone or not.

- The FAP value which is the false alarm probability. The lower the value of the FAP, the more unlikely the periodic signal has an instrumental origin.

- The harmonic value indicates if the period is a harmonic of a previous signal.

An example of a report is represented in table 1. Note that for each detrend window size, a different signal can be found but the final signal that is selected by SHERLOCK for the run is the signal that is the most voted multiplied by the SDE, independently of the value of the SNR. By default, it is the **quorum** algorithm that selects this signal but it can be changed. As explained briefly above, along with the report, a folder per run is also created. In this file, a plot with 3 graphs is represented. The first one (at the top of figure 11) is the relative detrended flux as a function of time. Transit signals detected by SHERLOCK are visible in red and some gaps in the light curve might be present for 3 reasons:

1. If the mode GLOBAL or BOTH are used, the resulting plot is a combination of all the sectors available for the targets. Hence, gaps are due to the time gap between the different sectors.

2. When TESS sends data to Earth, no data are registered, creating a small gap between each orbit of the satellite.

3. After each run, the transit found is masked for the next run.

The second graph of figure 11 shows the transit signal for the detected period (or phase-folded light curves). Data points associated with their error bars and a schematic transit curve in red (centred on phase zero) are represented. Note that this red curve is there to help the user see where the transit is supposed to be found but it is not an actual model fit. The third graph is the SDE as a function of the period (in days). The period selected is the one corresponding to the highest peak in the graph, indicated by a blue line. Harmonics of the selected period are indicated by blue dotted lines. Such a figure is created for each window size and for each run.

Run 1

win_size	P	P_err	Num tran	Mean δ(ppt)	dur(min)	T0	SNR	SDE	Border_score	Matching TOI	Harmonic	Planet radius(R_Earth)	R_p/R_s	Semi-major axis	Habitability Zone
PDCSAP_FLUX	4.65181	0.000453	17	0.862	103.9	1327.4424	48.362	56.639	1.00	TOI 125.01	-	2.73193	0.02641	0.05295	I
0.2000	4.65181	0.000453	17	0.457	94.4	1327.4440	25.425	49.353	1.00	TOI 125.01	-	1.98856	0.01826	0.05295	I
0.3000	4.65181	0.000453	17	0.638	94.4	1327.4440	35.167	54.270	1.00	TOI 125.01	-	2.35157	0.02261	0.05295	I
0.4000	4.65181	0.000453	17	0.735	94.4	1327.4440	40.122	53.119	1.00	TOI 125.01	-	2.52346	0.02453	0.05295	I
0.5000	4.65181	0.000453	17	0.789	94.4	1327.4440	42.862	52.549	1.00	TOI 125.01	-	2.61487	0.02555	0.05295	I
0.6000	4.65181	0.000453	17	0.800	103.9	1327.4424	45.647	51.555	1.00	TOI 125.01	-	2.63151	0.02520	0.05295	I
0.7000	4.65181	0.000453	17	0.814	103.9	1327.4424	46.427	51.083	1.00	TOI 125.01	-	2.65591	0.02553	0.05295	I
0.8000	4.65181	0.000453	17	0.830	103.9	1327.4424	47.327	51.349	1.00	TOI 125.01	-	2.68199	0.02582	0.05295	I
0.9000	4.65181	0.000453	17	0.845	103.9	1327.4424	48.116	51.698	1.00	TOI 125.01	-	2.70544	0.02608	0.05295	I
1.0000	4.65181	0.003972	17	0.854	103.9	1327.4424	48.537	51.469	1.00	TOI 125.01	-	2.71965	0.02622	0.05295	I
1.1000	4.65181	0.003972	17	0.859	103.9	1327.4424	48.737	51.476	1.00	TOI 125.01	-	2.72707	0.02633	0.05295	I

Elected signal with QUORUM algorithm from 11 VOTES — > NAME: PDCSAP_FLUX Period:4.65180555044495806 CORR_SDE: 118.426349888829104 SNR: 48.361586536860166 SDE: 56.6386890744435411 FAP: 8.0032e-05 BORDER_SCORE: 1.0
Proposed selection with BASIC algorithm was — > NAME: PDCSAP_FLUX Period:4.65180555044495806 SNR: 48.361586536860166 New best signal is good enough to keep searching. Going to the next run.

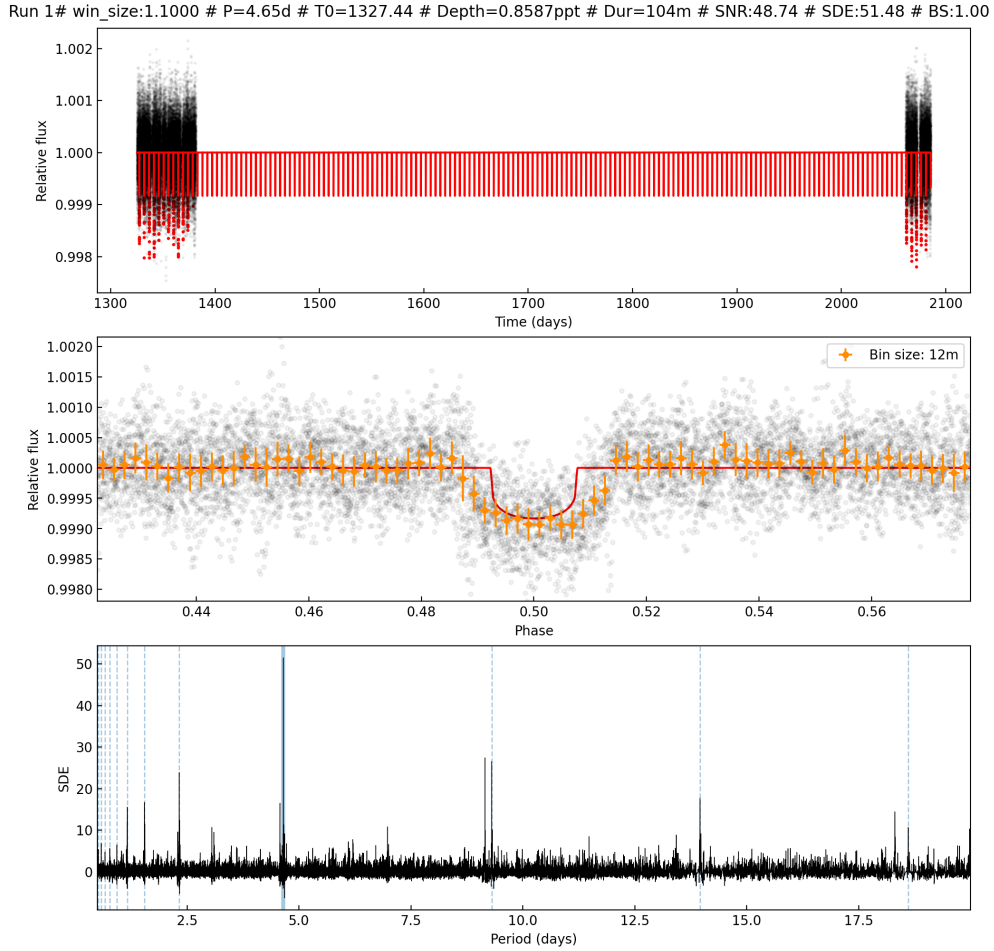Table 1: Example of a report file for TIC 52368076 using the global mode.

Figure 11: Example of the 3 graphs obtained after running SHERLOCK. This was obtained for TIC 52368076 during the first run and with the 1.1 window size (in days), i.e., the one with the highest SNR. In this example, the already known exoplanet, TOI-125 b, is detected with a period of 4.65 days.

Finally, one should keep in mind that the user has to make an informed choice to move to the next step or not, based on the different plots and metrics output by SHERLOCK at the second step. This decision should be based on the following list:

- The transit shape is an important indicator as to whether the signal looks promising or not. It may happen that no decrease in flux is visible to the naked eye.

- If the signal is only found with a small number of window sizes, the confidence in the signal strongly decreases. Indeed, the more the signal is independent of the detrending, the more we can be confident that it has an astrophysical origin.

28

- If the signal has a low border score, it means that some transits were found at the border of gaps in the light curve. As those regions have more uncorrected systematics, the confidence in the signal found decreases. A high border score is thus preferred.

- If the signal has a high FAP value, the probability of it being a false positive increases. Thus, this value should be as low as possible.

- The harmonics of the transit signal should correspond to peaks of the SDE value (blue dotted lines).

- If the SECTOR mode is used, the signal should be found in several sectors independently (only for orbital periods smaller than 13.5 days, i.e., half of the observational time per sector which is 27 days). If a signal looks promising in one sector, then the GLOBAL mode must be used. If not recovered with this mode, the signal is discarded. In addition, if several sets of sectors have been run separately, the signal should be found in each of these sets. Similarly, if there is any doubt, the GLOBAL mode should be used.

3.3   Performing the vetting of interesting signals

In the case of an interesting signal found during the second module of SHERLOCK in section 3.2, the third module can be used. The goal here is to vet interesting signals to determine if they could be attributed to a planet or if they are more likely to be astronomical or instrumental false positives. First, SHERLOCK must be launched with the *vetting* mode for the interesting signal. After the run, a new directory called 'vetting {number of the candidate}' will be created, containing a report with different graphs and tables. A detailed explanation of the content of the report can be found on the following webpage: https://sherlock-ph.readthedocs.io/en/latest/TOI-175/TOI-175-vetting.html and the following paragraphs will be based on it.

At the top of the report is a table with some properties of the found planet candidate, like the epoch, the transit duration in hours, the period in days, and the depth in ppt. Then, there are different tables and figures that help the user determine whether the potential planet is a false positive or not. These will now be seen in detail.

3.3.1 Folded curves for the different available cadences

If data are available for fast (20s), short (120s) and long (600s and/or 1800s) cadences, then, the associated folded curves are displayed. If not, the plot will remain empty. An example of the transit of TOI-125 b for the 3 different cadences is shown in figure 12.
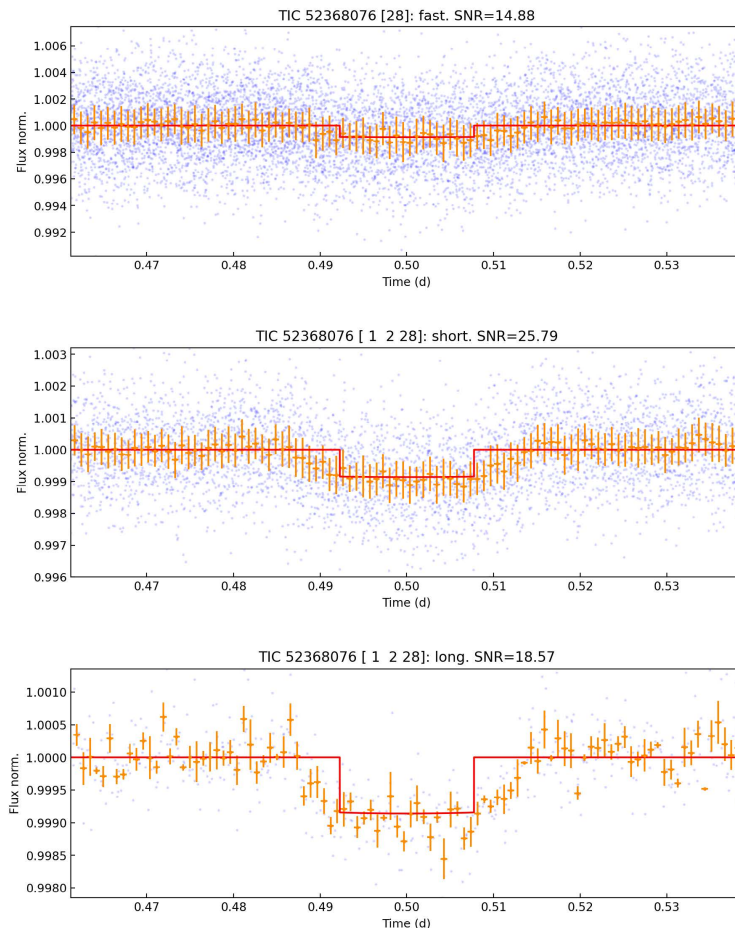

Figure 12: Folded curves for the fast, short and long cadence data of TIC 52368076 for the known orbiting planet TOI-125 b with a period of 4.65 days.

3.3.2 Analyse of the transit depth plot

A false-positive signal could be attributed to an eclipsing binary system. This case appears when two stars are in a binary system with an orbital inclination (with respect to the line of the sky) close to 90°. The stars can eclipse each other over an orbital cycle, resulting in a reduction of the apparent brightness in the light curve (Rauw, 2022-2023). This case should

not be confused with a planetary signal. When two stars are in a binary system, if one is brighter than the other, the transit depth will be different if the brightest star is in front of or behind the other star. Indeed, when the fainter star passes in front of the brighter one, the amount of light is drastically reduced because the brightest object is partially or totally hidden, depending on the stellar radii. When the fainter star passes behind the brightest one, the amount of light is also reduced, but in this case, the transit depth will be smaller because the brightest object is still fully observable. An ideal case of such a system is represented in figure 13.
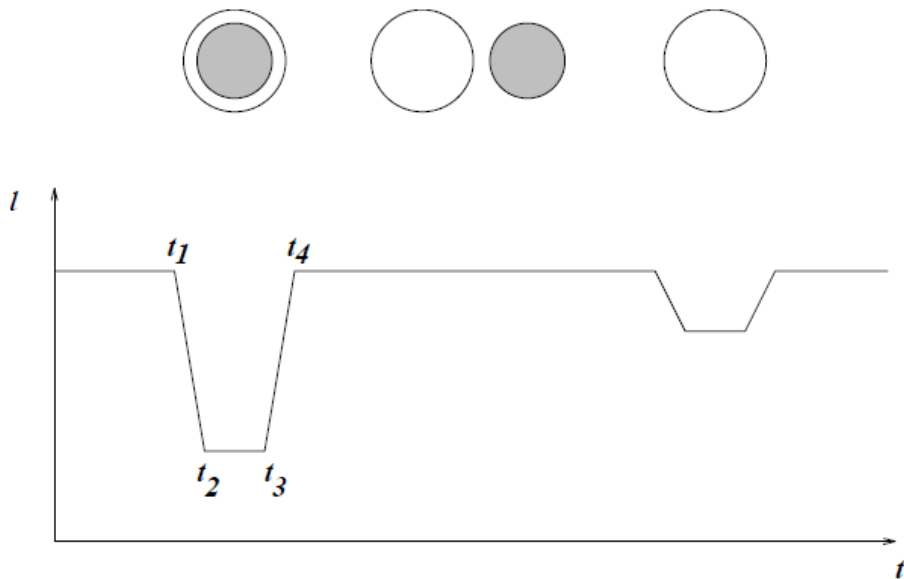

Figure 13: Example of an eclipsing binary system. The brightest and faintest stars are respectively represented by the white and black disks. The resulting light curve with the flux as a function of time is also illustrated with the resulting depth difference in the two cases. Illustration taken from Rauw (2022-2023).

In the case of a planetary transit signal, the same transit depth is expected each time the planet passes in front of its host star. As in this research, we are looking for very shallow transit signals with low SDE and SNR, the occultation signal will not be observable, resulting in no expected difference between the different transit depths. In order to differentiate a signal due to a planet or another star, the plot called *"The candidate single-transits depths plot"* must be analysed. The depth of each transit reported by the best TLS result is represented by a dot with its associated error bar. An example of such a plot is shown

in figure 14. Indeed, if we analyse the even and odd transits, in the case of a planetary signal, each time a transit happens, the same transit depth is expected. On the contrary, an eclipsing binary will have a different transit depth for odd and even transits depending on the system configuration. To conclude, the comparison between the odd and even transit depths is an important tool to help the user discriminate an astronomical false-positive signal.
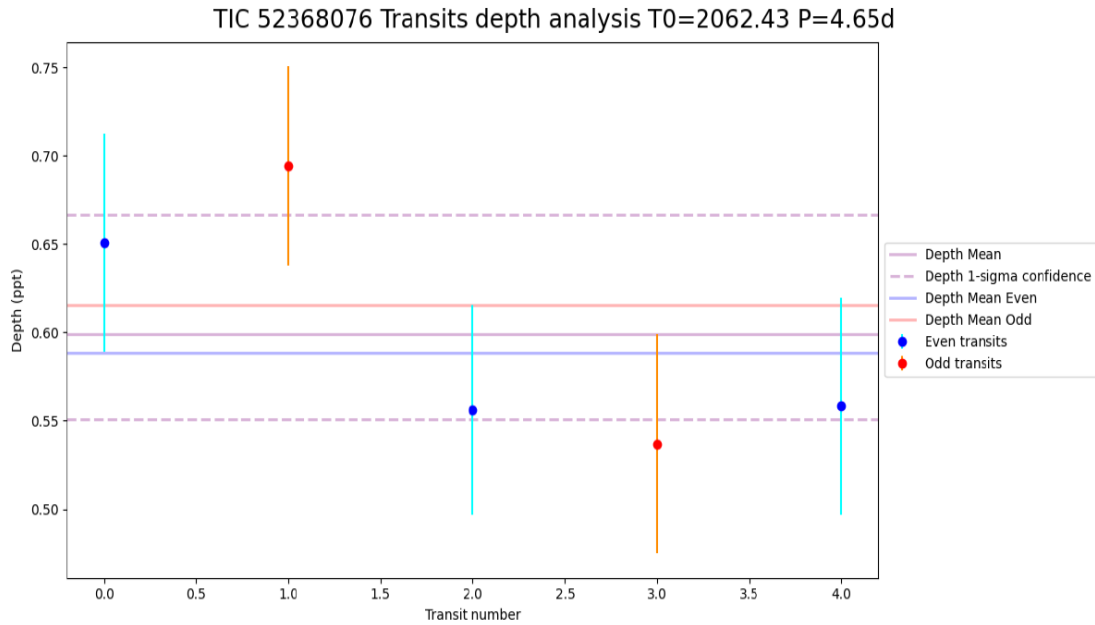

Figure 14: Example of the plot showing the different transit depths for an already known exoplanet, TOI-125 b. There is no significant difference between the even transits (in blue) and the odd transits (in red). The mean transit depths are very similar given the scale on the y-axis.

3.3.3 Folded light curves with the original period and first (sub)harmonics

A second figure to analyse in the report is a figure representing the folded light curve for different periods at superior and inferior conjunction. Superior conjunction represents the configuration where there is an observer-star-planet alignment, resulting in a hidden planet from the observer's point of view (occultation). The second case is the inferior conjunction, representing an observer-planet-star alignment and allowing for a transit planetary signal to be observed. Those two configurations are tested for the planetary candidate with different periods.

1. The first two figures to analyse are the folded light curves of the planetary candidate

with the original period found by SHERLOCK for inferior and superior conjunctions. If the period is correct, a transit signal should be observed at inferior conjunction and no transit shape at superior conjunction. An example of this case is represented in figure 15 for TOI-125 b.
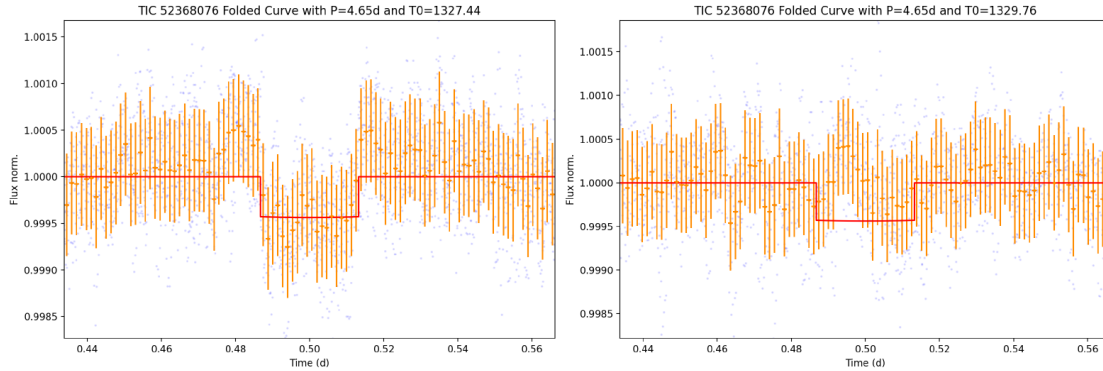

Figure 15: Example of the folded light curve for TOI-125 b with its original period of 4.65 days. The left and right sides represent, respectively, the inferior and superior conjunctions.

2. The second pair of figures to analyse is the first harmonic at the superior and inferior conjunctions. In this case, the period is twice the original period. As a result, a transit signal is expected at both inferior and superior conjunctions because it should correspond to a full orbital rotation for the original period in both cases. Indeed, at superior conjunction, the planet should have travelled half of the original period multiplied by two, i.e., one original period. Even if a transit shape is expected, the SNR will be lower than the two previous plots because only half the transits are used, reducing the amount of data. An example is shown in figure 16.

3. Finally, the third case is for the first sub-harmonic, i.e., for a period that is half the original period. In this case, a transit signal is expected at the inferior conjunction but not at the superior conjunction. Indeed, at the superior conjunction of half the original period, the planet should not be in the line of sight of the observer. Again, even if a transit shape should be observed at the inferior conjunction, the SNR is expected to be lower as it combines the transits with some portions of flat light curves. An example is shown in figure 17.
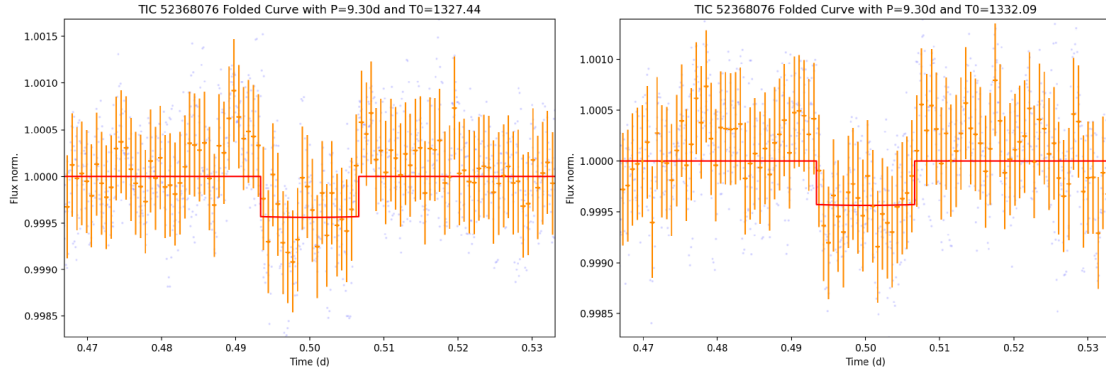
33

Figure 16: Example of the folded light curve for TOI-125 b. The inferior and superior conjunctions are represented, respectively, on the left and right for the first harmonic with a period that is twice the original period.
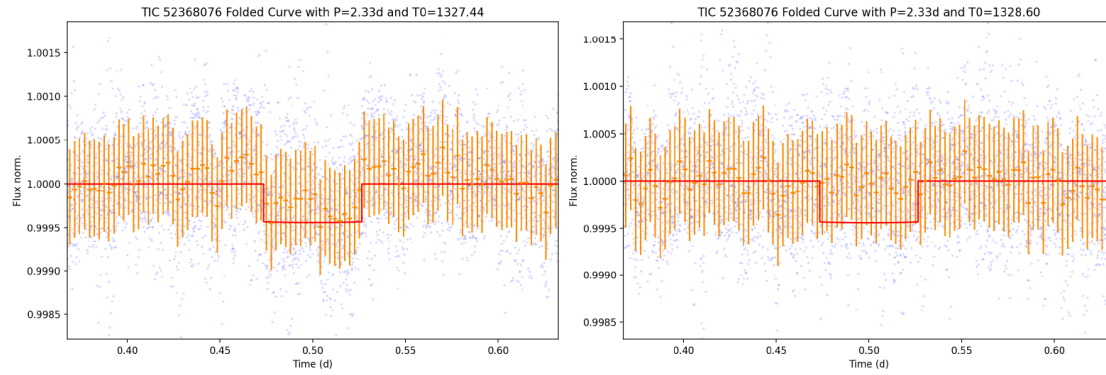

Figure 17: Example of the folded light curve for TOI-125 b. The inferior and superior conjunctions are represented, respectively, on the left and right for the first sub-harmonic with a period that is half the original period.

The combination of those graphs should help the user determine if the period selected by SHERLOCK is consistent with this test or not.

3.3.4 Source and target position

Then, there is one figure representing the target position (with a blue star) and the transit source position (with a red dot) with its associated uncertainty (a yellow circle). The precise position of the star is determined with the centre of mass of the intensity distribution of the pixels belonging to the star, while the transit source position is determined with the transit pattern in the pixels. If the offset between the target and the transit source is larger than the

34

uncertainty, the signal could be attributed to a nearby star (e.g. a blended eclipsing binary). Figure 18 shows an example where the target star is within the yellow circle, suggesting that the transit signal is well on the target star.
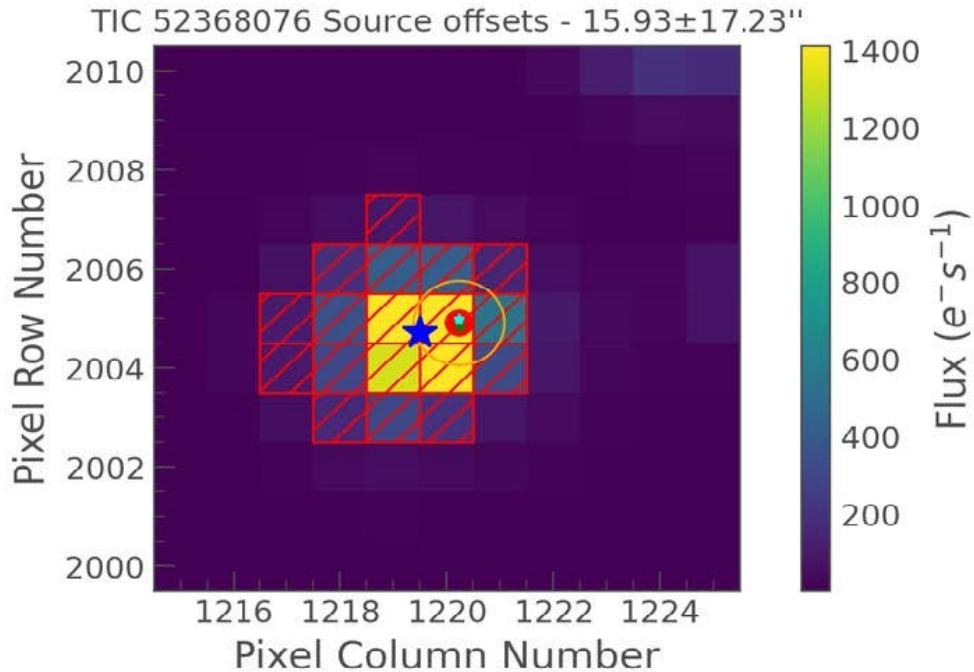

Figure 18: The transit source and target star (TIC 52368076) are displayed with, respectively, the red dot and the blue star. At the right is the color bar associated with the flux (in electrons per second).

3.3.5 The Target Pixel Files pixels folded curves

Then, there is a figure showing the TPF file (as explained in section 1.2), which is represented in figure 19. In this case, one can see that the folded light curve for each pixel in the field of view is established. The goal is to check that the transit signal is well originating from pixels that are centered on the target and that there is no transit-like signal outside the photometric aperture (represented in purple). Indeed, if there is an increase or decrease in flux outside the aperture, due to contamination and dilution effects, it could be responsible for the observed transit shape, resulting in a false-positive detection. Therefore, while looking at this figure, no transit-like signals should be observed out of the photometric aperture. In figure 19, one can see that there is a red line in each pixel which is a fit of the light curve with the

BLS (Box Least Squared) algorithm (Kovács et al., 2002). This algorithm approximates the transit light curve as a boxcar function with a normalized average out-of-transit flux of zero and a fixed depth during the transit (Hippke, Heller, 2019). Hence, in each pixel, the best-fit model of the BLS algorithm is shown.
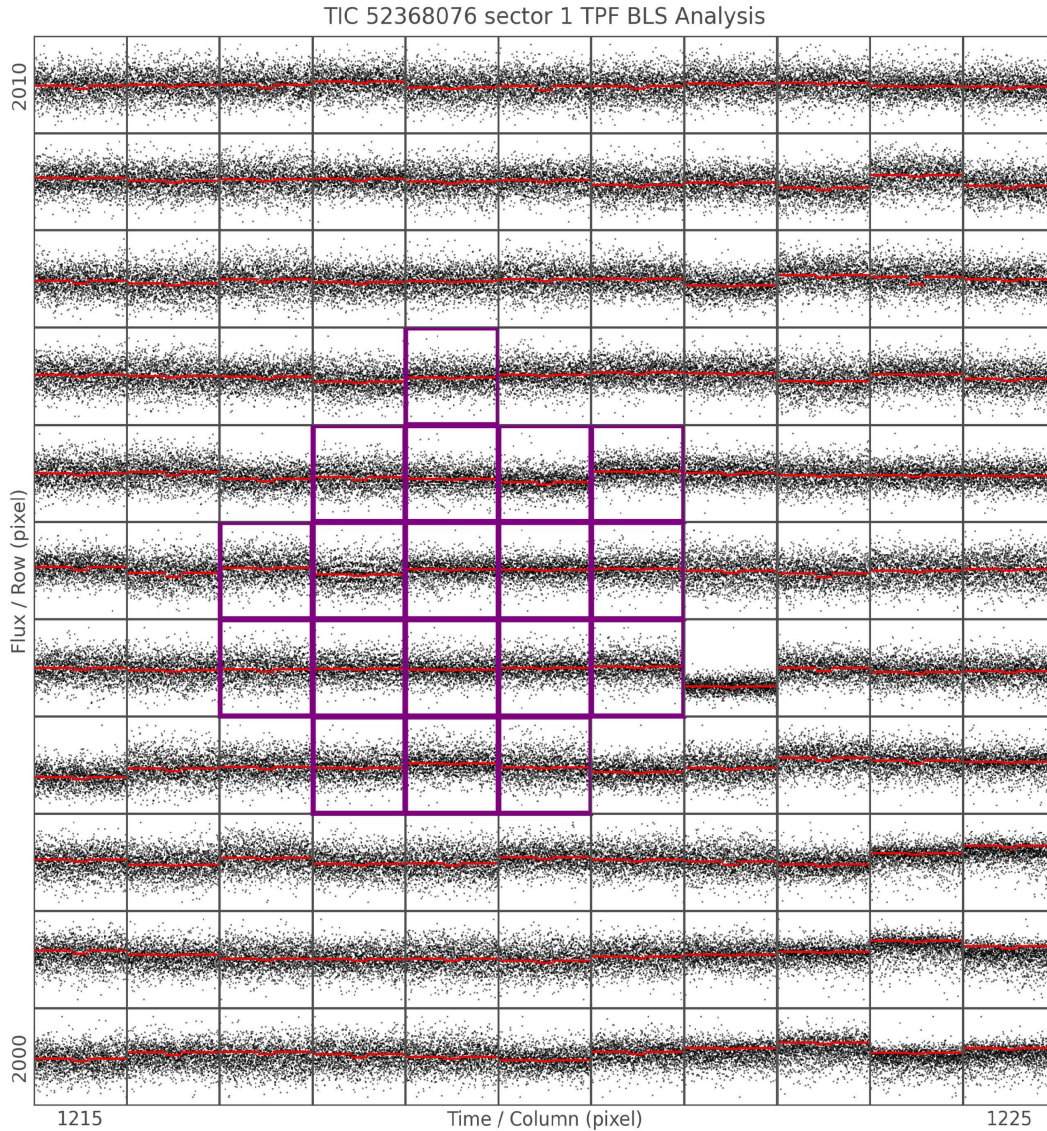

Figure 19: The target pixel file of TIC 52368076 for sector 1. No transit shapes are observed outside the photometric aperture (in purple). The red lines represent the best fit of the BLS algorithm for each pixel.

36

3.3.6 Single-transits vetting sheets

There is a dedicated sheet for each detected transit shape, allowing one to determine if the different transits are consistent with each other. On this sheet, there are several graphs that must be analysed. An example is shown in figure 20. The content of this sheet will now be developed.

1. The plot on the left of the top row represents the normalized flux for the transit considered as a function of time (in days), with the data points as orange dots and a transit shape obtained with the BLS algorithm in red, and using the period and the epoch found.

2. The plots in the middle and the right part of the top row represent the normalized X and Y-axis drifts as a function of time in days. There are two vertical red lines corresponding to the transit time. The red and black dots are, respectively, the X-axis motion and the X-axis centroids. The X-axis centroid is the motion of the centre of mass of the intensity distribution, while the X-axis motion is the motion of the geometrical centre of the intensity distribution. The same goes for the Y-axis with the Y-axis motion and centroids. Those graphs indicate if there is an offset in the centroid of the target in the transit images by comparing the centroid of the star in and out of transit. A significant offset could indicate a false-positive signal related to a background eclipsing source or a drift in the instrumentation (Hedges, 2021).

3. The plot at the bottom left represents the normalized flux as a function of time (in days). There are two different colours used. The orange dots correspond to the photometry for the photometric aperture used, and the blue dots show the photometry for a smaller aperture. If the transit signal is well on the target star, then it should still be visible for a smaller aperture but the SNR would be lower because there are fewer photons. On the opposite, if the signal is caused by an eclipsing binary blended with the target star and contaminating the edge of the used aperture, then it would not be visible with a smaller aperture.

4. The middle bottom graph represents the TPF (as described in section 1.2) aperture comparison. It represents the flux intensity (with different colours) for each pixel. For the comparison between the apertures, the small one is in black while the large one is in red.

5. The right bottom plot shows the background flux as a function of time in days. The two vertical lines represent the transit time, allowing us to see if the background flux could be the cause of the transit shape.

6. Finally, there is a plot showing the TPF flux measurements for each pixel around the transit time. It is the same as for the previous TPF (figure 19) but now, it is given for each transit found.
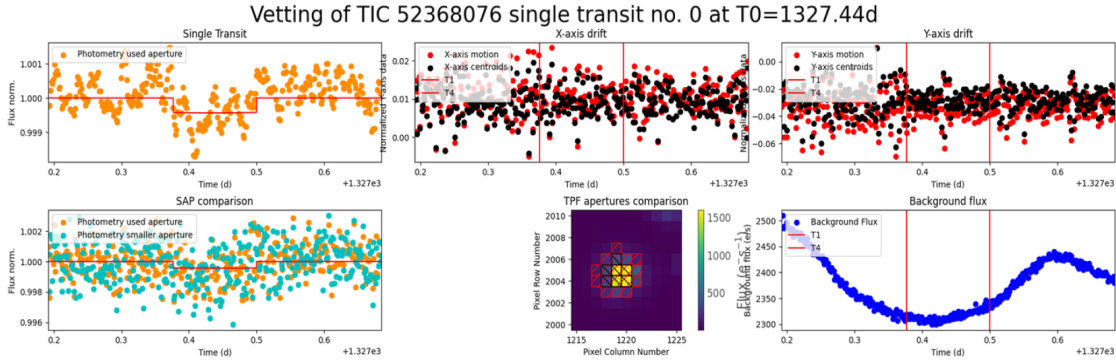

Figure 20: Example of the content of a single transit sheet for the first (numbered 0) transit found for TOI 125 b.

The goal of the vetting module is to classify events detected by the 2nd module of SHERLOCK as planet candidates or false positives of instrumental or astrophysical origin. It is always better to proceed with the next module to combine information from those two modules.

3.4   Performing the statistical validation

The 4th module used by SHERLOCK is about statistical validation. The goal is to use statistics to rule out astrophysical false positives with a confidence degree high enough to classify the signal from planet candidate to likely or validated planet. This is achieved with TRICERATOPS (Tool for Rating Interesting Candidate Exoplanets and Reliability Analysis of Transits Originating from Proximate Stars), which is an open-source Bayesian tool (https://github.com/stevengiacalone/triceratops; Giacalone et al., 2020). To use it, the 4th module of SHERLOCK must be run for the found candidate. We will first review the principles behind this Bayesian tool in order to better understand the results obtained. The following section will be entirely based on the article of Giacalone et al. (2020).

3.4.1 Statistical analysis

First, TRICERATOPS identifies each star within a 10-pixel radius of the target star, based on the TESS Input Catalog. The goal is to determine if those stars could be responsible for the observed transit or not. To this end, flux ratios (X_s) are computed for each star by dividing the flux of the star falling in the photometric aperture by the total flux contributed to the aperture by all stars. The relative transit depth, which is the transit depth if that star was responsible for the observed transit signal, is given for each star by $\delta_s = \delta_{obs}/X_s$, with δ_{obs} the observed transit depth. Stars that contribute little to the aperture flux may have δ_s exceeding unity and are removed from further analysis. Indeed, they are too faint compared to the target star to be responsible for the observed transit. At this point, different scenarios can be responsible for the transit-like events and are listed in table 2. Those can be classified into 4 categories:

1. The target star has no unresolved companion: TP, EB and EBx2P.

2. The target star has an unresolved bound stellar companion: PTP, PEB, PEBx2P, STP, SEB, SEBx2P.

3. There is an unresolved foreground or background star along the line of sight: DTP, DEB, DEBx2P, BTP, BEBx2P.

4. For each resolved nearby star with $\delta_s < 1$: NTP, NEB and NEBx2P.

In order to determine which scenario is the most likely to be responsible for the transit, a probability will be assigned to each of them. The main principle is based on Bayes' theorem:

$$p(\theta_j|S_j, D) = \frac{p(\theta_j|S_j)p(D|S_j, \theta_j)}{p(D|S_j)} \tag{19}$$

Where

1. $p(\theta_j|S_j, D)$ is the posterior probability of the vector of parameters θ_j (defined below) given the jth scenario S_j (model) and the data (D).

2. $p(D|S_j, \theta_j)$ is the likelihood, also written $\mathcal{L}$, i.e. the probability of observing the data D under the scenario S_j with the vector of parameters θ_j.

3. $p(\theta_j|S_j)$ is the prior probability of the vector of parameters θ_j given the scenario S_j, also written π.

4. $p(D|S_j)$ is the marginal likelihood of the data D given the scenario S_j (or Bayesian evidence). It quantifies the degree of confidence about the scenario S_j given the observed data and it is also written $\mathcal{Z}$.

Every scenario S_j is associated with a vector of parameters θ_j. Those parameters are the following:

- i: The inclination of the orbit of a transiting planet or an eclipsing binary.

- R_p: The radius of the transiting planet.

- q_{short}: The mass ratio between the host star and a short-period stellar companion (for eclipsing binaries).

- q_{long}: The mass ratio between the target star and a long-period stellar companion (for an unresolved bound companion).

- *simulated star*: Properties of a star drawn from a population of stars simulated with TRILEGAL, which is a code that can simulate the stellar photometry of any galaxy field (Girardi et al., 2005).

The parameters associated to each scenario are given in table 2.

Because we work with a parameter vector, the marginal likelihood is calculated as the integral over the parameter space:

$$p(D|S_j) = \int p(\theta_j|S_j)p(D|S_j, \theta_j)d\theta \tag{20}$$

Then, the probability $p(S_j|D)$ associated with each scenario can be computed and the relative probability of each scenario is given by the following equation (Giacalone et al., 2020):

$$\mathcal{P}_j = \frac{p(S_j|D)}{\sum_j p(S_j|D)} \tag{21}$$

From this, two important quantities can be retrieved:

1. The FPP is the false-positive probability, which is the probability that the transit is produced by something else than a planet around the target star. It is given by the following equation:

Scenario	Configuration	Parameter Vector θ_j
TP	No unresolved companion; transiting planet with P_{orb} around target star	(i, R_p)
EB	No unresolved companion; eclipsing binary with P_{orb} around target star	(i, q_{short})
EBx2P	No unresolved companion; eclipsing binary with $2P_{orb}$ around target star	(i, q_{short})
PTP	Unresolved bound companion; transiting planet with P_{orb} around primary star	(i, R_p, q_{long})
PEB	Unresolved bound companion; eclipsing binary with P_{orb} around primary star	(i, q_{short}, q_{long})
PEBx2P	Unresolved bound companion; eclipsing binary with $2 \times P_{orb}$ around primary star	(i, q_{short}, q_{long})
STP	Unresolved bound companion; transiting planet with P_{orb} around secondary star	(i, R_p, q_{long})
SEB	Unresolved bound companion; eclipsing binary with P_{orb} around secondary star	(i, q_{short}, q_{long})
SEBx2P	Unresolved bound companion; eclipsing binary with $2 \times P_{orb}$ around secondary star	(i, q_{short}, q_{long})
DTP	Unresolved background star; transiting planet with P_{orb} around target star	$(i, R_p,$ simulated star)
DEB	Unresolved background star; eclipsing binary with P_{orb} around target star	$(i, q_{short},$ simulated star)
DEBx2P	Unresolved background star; eclipsing binary with $2 \times P_{orb}$ around target star	$(i, q_{short},$ simulated star)
BTP	Unresolved background star; transiting planet with P_{orb} around background star	$(i, R_p,$ simulated star)
BEB	Unresolved background star; eclipsing binary with P_{orb} around background star	$(i, q_{short},$ simulated star)
BEBx2P	Unresolved background star; eclipsing binary with $2 \times P_{orb}$ around background star	$(i, q_{short},$ simulated star)
NTP	No unresolved companion; transiting planet with P_{orb} around nearby star	(i, R_p)
NEB	No unresolved companion; eclipsing binary with P_{orb} around nearby star	(i, q_{short})
NEBx2P	No unresolved companion; eclipsing binary with $2 \times P_{orb}$ around nearby star	(i, q_{short})

Table 2: Table listing every scenario analysed by TRICERATOPS.

TOI classification	FPP	NFPP
Validated planet	< 0.015	$< 10^{-3}$
Likely planet	< 0.5	$< 10^{-3}$
Likely nearby false positive	/	$> 10^{-1}$

Table 3: TOI classification as defined in Giacalone et al. (2020).

$$FPP = 1 - (\mathcal{P}_{TP} + \mathcal{P}_{PTP} + \mathcal{P}_{DTP}) \tag{22}$$

2. The NFPP is the nearby false-positive probability, which is the probability that the transit is due to a resolved nearby star rather than the target star. It is given by the following equation:

$$NFPP = \sum (\mathcal{P}_{NTP} + \mathcal{P}_{NEB} + \mathcal{P}_{NEB \times 2P}) \tag{23}$$

Based on these two quantities, we can classify the found candidates into 3 different categories: validated planets, likely planets, and likely nearby false positives. Those categories were defined in Giacalone et al. (2020) and are based on a sample of 68 TOIs that have been designated as either confirmed planets or astrophysical false positives by members of the TESS Observation Follow-up Program (TFOP) based on follow-up observations. Moreover, those results were compared with another statistical validation tool called VESPA (https://github.com/timothydmorton/vespa; Morton, 2015, 2012). Those categories are defined in table 3.

In reality, the integral in equation 20 can not be solved analytically and is approximated by sampling $p(\theta_j|S_j)$ using a Monte Carlo approach. Two simplifying assumptions are made to minimize the number of sampled parameters: (i) the stellar parameters such as R_*, M_* and T_{eff} of each resolved star are supposed to be known precisely (taken from the TIC catalogue); (ii) the orbital period is fixed and the eccentricity is assumed to be null. Despite short-period planets being biased towards lower eccentricities, the assumptions made by TRICERATOPS become less and less valid when dealing with longer orbital periods. Finally, the likelihood of the transit data is computed using the equation:

$$p(D|\theta_j^{(n)}, S_j) \propto \prod exp\left[-\frac{1}{2}\left(\frac{y_l - f(t_l|\theta_j^{(n)})}{\sigma} \right)^2 \right] \qquad (24)$$

With

1. y_l: the flux of the lth data point.

2. $\theta_j^{(n)}$: the nth sample from the parameter prior distribution.

3. $f(t_l|\theta_j^{(n)})$: the flux given by the model for the parameter $\theta_j^{(n)}$ at the time of the lth data point.

4. σ: the characteristic uncertainty of the flux.

Equation 24 is calculated by modelling light curves for all possible scenarios listed in table 2, using the `batman` python package. It is an open-source Python package available at https://github.com/lkreidberg/batman and it can be used to model the light curves of transits and eclipses using 7 different limb darkening laws (uniform, linear, quadratic, square-root, logarithmic, exponential and non-linear laws) (Kreidberg, 2015).

In addition to the FPP and NFPP results, two more quantities are computed by SHER-LOCK. Those are the FPP2 and the FPP3+ values and are the false positive probabilities for the candidate for which the validation is performed taking into account that one or more planets were already found in the system. It is linked to the fact that a candidate in a multi-planet system has a higher prior probability of being a real planet. This was found in Lissauer et al. (2012b) for Kepler's targets, but we can use those probabilities with candidates from TESS data as additional information. This effect is called the 'multiplicity boost'. As the FPP value can be replaced by the one obtained with the FPP2 or FPP3+ (depending on the case), the expected values to classify the candidate as a likely planet are the same as before (with the TOI classification in table 3). Finally, those two quantities are particularly relevant for this master's thesis because the target stars already have at least one TOI.

Finally, if high-resolution imaging data are available, contrast curves can be used to assess potential contamination from bound or unbound unresolved companions on the signal coming from the target. A contrast curve is a plot showing the difference in magnitude between the target star and a possible companion as a function of the angular separation. Any companion star lying above that contrast curve should be detected in the high-resolution imaging data.

By default, TRICERATOPS assumes that unresolved companions beyond 2.2 arcseconds from the target can be ruled out (as they would then have been resolved in Gaia DR2). This value can be decreased with high-resolution imaging. This method increases the precision of the FPP value by better constraining the probabilities of scenarios involving unresolved companions (instead of simply using the default assumption that only unresolved companions beyond 2.2 arcsecs from the target can be ruled out). As there is no new information on nearby resolved stars with this method, scenarios associated with them are not affected. As a consequence, the FPP value can be decreased, but the only way to increase the precision of the NFPP is to schedule new observations with instruments with a higher angular resolution than TESS (such as SPECULOOS). This allows separate light curves to be produced for the target and the resolved nearby stars and to check from which star the signal originates.

3.4.2   Results analysis

After running TRICERATOPS, some folders are created and will now be analysed. First, there is figure 21 where all stars within a 10-pixel radius are identified, as explained in section 3.4.1. If the target star is the only one bright enough to produce the observed transit, scenarios NTP, NEB, and NEBx2P are neglected. At this stage, 15 other scenarios would still be analysed for the next step.

Then, the best-fit models for each scenario are shown. In figure 22, best-fit models compared to the real TESS light curve are shown for 3 scenarios out of 15. It is useful to have an idea of the most probable scenario associated with the transit found. In the example given, the TP scenario looks clearly more promising than the two others. Then, there is a table listing the probability for each scenario as calculated with the equation 21. An example is shown in figure 23 for TOI-125 b, where the scenario with the highest probability is the TP scenario, as expected as it is a confirmed planet. Hence, the transit signal found for this target has a high probability of being a transiting planet with the original period around the target star and no unresolved companion.
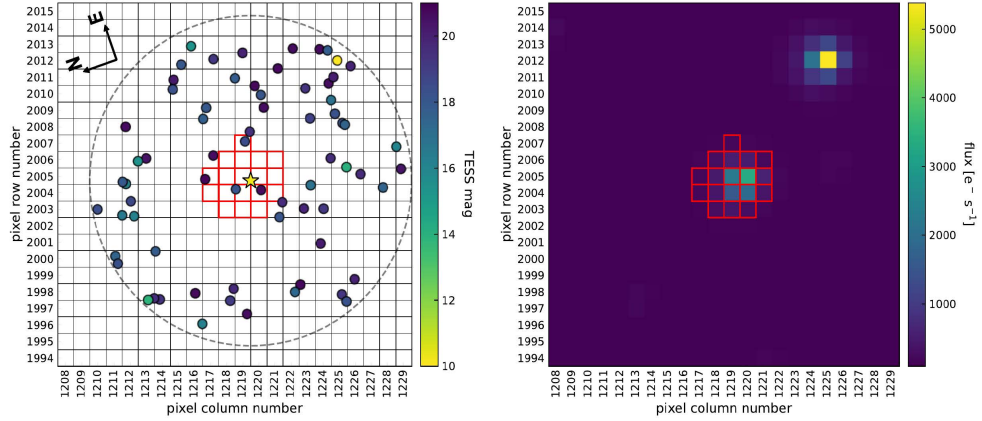
Figure 21: On the left is the target star (TIC 52368076 with its known planet, TOI-125 b) with the yellow star symbol, the 10-pixel radius is represented with the dashed grey line, and the aperture is in red. The colour bar at the right gives the TESS magnitude of each star detected. From yellow to blue are, respectively, the brighter to fainter stars. On the right, is the time-averaged TESS image of the same pixels. The colour bar at the right is for the flux (in electrons/second), with the blue and yellow, respectively, for a lower and higher flux.
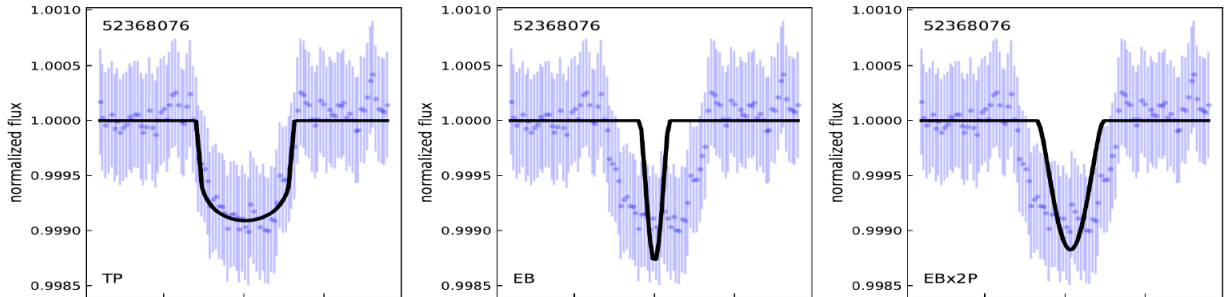

Figure 22: This example shows the best-fit models for 3 scenarios: TP, EB, EBx2P, for TOI-125 b. Those models are compared to the extracted TESS light curve. As expected for this known TOI, the TP scenario gives the best fit.

For the calculation of the FPP and NFPP values, five points are drawn from the probability distribution of each scenario and are used to compute five values of FPP and NFPP, using the equations 22 and 23. The final FPP and NFPP are then the respective mean of those five values. An example is shown in table 24.

ID	scenario	M_s	R_s	P_orb	inc	b	ecc	w	R_p	M_EB	R_EB	prob
52368076	TP	0.91	0.85	4.65	87.65	0.52	0.55	217.76	2.73	0.0	0.0	0.763129
52368076	EB	0.91	0.85	4.65	83.18	1.97	0.01	133.7	0.0	0.84	0.85	0.0
52368076	EBx2P	0.91	0.85	9.3	87.28	0.67	0.69	179.18	0.0	0.88	0.85	4e-06
52368076	PTP	0.91	0.85	4.65	87.7	0.56	0.55	220.4	2.83	0.0	0.0	0.158507
52368076	PEB	0.91	0.85	4.65	85.02	1.27	0.0	187.59	0.0	0.25	0.27	0.0
52368076	PEBx2P	0.91	0.85	9.3	87.26	0.48	0.75	167.58	0.0	0.88	0.85	1e-06
52368076	STP	0.64	0.66	4.65	88.8	0.33	0.38	207.9	5.55	0.0	0.0	0.016014
52368076	SEB	0.17	0.2	4.65	89.5	0.09	0.81	182.46	0.0	0.04	0.1	0.000224
52368076	SEBx2P	0.26	0.28	9.3	89.0	0.46	0.69	178.62	0.0	0.25	0.27	0.000105
52368076	DTP	0.91	0.85	4.65	87.49	0.57	0.59	211.72	2.72	0.0	0.0	0.057445

Figure 23: This is the table summarizing the probabilities (on the right) of all the explored scenarios (on the left). This example is for TOI-125 b for which the highest probability is for the TP scenario.

Scenario	FPP	NFPP	FPP2	FPP3+
0	0.022895	0.0	0.000936	0.000468
1	0.019233	0.0	0.000784	0.000392
2	0.020585	0.0	0.00084	0.00042
3	0.021949	0.0	0.000897	0.000449
4	0.019937	0.0	0.000813	0.000407
MEAN	0.02092	0.0	0.000854	0.000427

Figure 24: An example of a table presenting validation results. The values for each scenario are listed, and then the mean is computed and shown in the last row.

Finally, there is a map (in figure 25) showing in which category the signal is classified. It represents the NFPP as a function of the FPP value. The signal is classified depending on the values of the FPP and NFPP computed, as explained with the TOI classification in table 3.
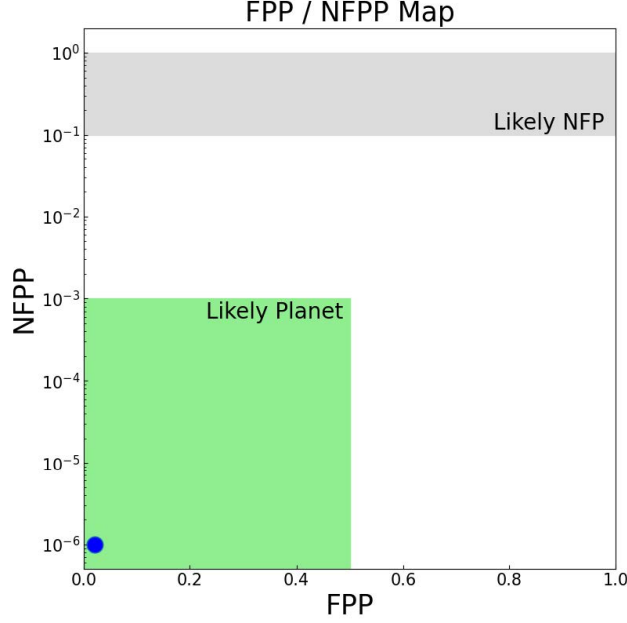
Figure 25: On the x-axis is the FPP value, and on the y-axis is the NFPP. The blue point represents the place attributed to the signal. As it is for TOI-125 b, the signal falls into the category 'likely planet', as expected.

Those results are used to determine whether or not to proceed with the next module of SHERLOCK in the case of a promising candidate.

3.5 Model the signal to refine the ephemeris

The 5th module of SHERLOCK is used when the results of the precedent modules look promising enough to continue. The goal is to plan a follow-up campaign from a ground-based observatory to firmly confirm that the candidate is a planet. To do so, the ephemeris must be as precise as possible to allow accurate predictions of the future transit times of the candidate. Indeed, if there is a significant uncertainty (ΔP) on the period (P), this error will propagate as the number of cycles (n) increases. This can be understood with the following relations (Mallonn et al., 2019):

$$T = T_0 + n \cdot P$$

$$\Delta T = \sqrt{\Delta T_0^2 + (n \cdot \Delta P)^2}$$

With T, the transit time and ΔT its associated error, and T_0 the mid-transit time for the first identified transit with its uncertainty ΔT_0. Hence, one needs to have the best possible precision on P and T_0 so that the uncertainties on these two quantities are as small as possible. To this end, the `allesfitter` package, which is an open-source Python software available at https://github.com/MNGuenther/allesfitter, is used to properly fit the data and obtain the best possible precision on P and T_0. `Allesfitter` allows to perform parameter inference and model testing for, notably, exoplanet transits, occultations, and radial velocity signals (Günther, Daylan, 2021). This algorithm estimates the Bayesian evidence (as defined in section 3.4.1) to compare different physical models using nested sampling (used by default with `SHERLOCK`). In this case, the parameters used in the Bayes theorem are the planetary and stellar parameters (defined below), while the observed data are the time series of the normalized light curves.

First, as nested sampling was used, the principle behind it will be reviewed. It is based on Bayes' theorem (equation 19) and it is an algorithm that estimates the Bayesian evidence of different models, allowing them to be compared. The following paragraphs will be based on the Günther, Daylan (2021) article. To calculate the Bayesian evidence, the following integral should be computed:

$$\mathcal{Z} \equiv \int_{\Omega_\theta} \mathcal{L}(\theta)\pi(\theta)d\theta \tag{25}$$

As the integral is over the multidimensional domain of θ, it is very difficult to estimate it. Hence, the Bayesian evidence will be computed as follows:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X)dX \tag{26}$$

In this equation, there are two important quantities:

1. X: It is the prior volume which is the fraction of the prior where the likelihood $\mathcal{L}(\theta)$ is above a certain threshold λ. It is defined as follows:

$$X(\lambda) = \int_{\Omega_{(\theta)}:\mathcal{L}(\Omega)>\lambda} \pi(\theta)d\theta \tag{27}$$

The prior is normalized, with $X(\lambda = 0) = 1$ and $X(\lambda \to \infty) = 0$, defining the bounds of the integral.

2. $\mathcal{L}(X)$ defines an iso-likelihood contour, which itself defines the prior volume.

To approximate the prior volume, nested sampling uses a statistical approach to generate samples from the prior $\pi(\theta)$ (contrary to MCMC which samples from the posterior). After drawing a number of points from the prior, the point with the lowest likelihood is removed. Then a new point is drawn from the prior with the constraint that it has a higher likelihood than the one that was removed, and the prior volume is computed from that new point. This iterative process continues until the change in the Bayesian evidence is below a certain threshold. Finally, the samples are sorted according to their likelihoods. To better understand how it works, figure 26 illustrates the different prior volumes each time a new point is drawn from the prior for a 2-dimensional problem.
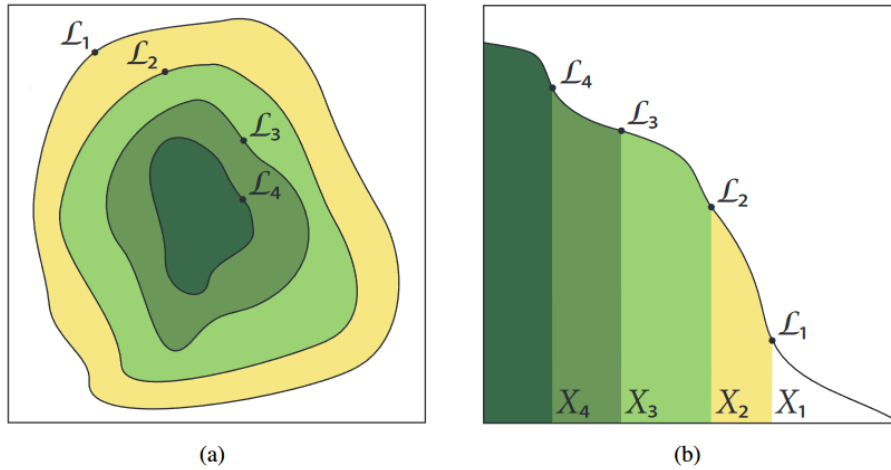

Figure 26: At the left is the posterior of a two-dimensional problem where the iso-likelihood contour is defined for each λ_i (i=1,...,4). At the right, is the $\mathcal{L}(X)$ function where each prior volume (X_i) is associated with its likelihood $(\mathcal{L}_i)$ (Feroz et al., 2013).

Concretely, SHERLOCK injects the stellar parameters as priors to estimate different parameters such as the radius ratio between the planet and the host star, the orbital period of the planet in days, ... To fine-tune the initial guess of the priors, one must use the allesfitter python package that itself uses the DYNESTY open source python package to perform the nested sampling (https://github.com/joshspeagle/dynesty; Speagle, 2020).

Second, the goal is to accurately estimate the red noise (correlated noise) and white noise (uncorrelated noise) of the light curve used. The red noise comes from the time-variability

49

of unaccounted physical processes such as stellar variability. To model it, we must first use the original light curve from SPOC using the `LIGHTKURVE` package. The idea is to have a light curve that is not modified by the Savistky Golay filter, the biweight detrends, or outlier removals to perform an in-depth fitting process. Indeed, the fitting must be done at the same time as the detrending to propagate the errors correctly. Then, a mask is applied to the light curve to hide the transits from the known TOI(s) and the found candidate to ensure that those are not considered noise and are not removed.

Now, the noise fitting can begin. With `allesfitter`, there are several ways to model the noise: linear trend, GP with a complex kernel, constant offset, ... These different models can be tested by comparing the Bayesian evidence, starting from the simplest one, i.e. a constant offset. A more complicated noise model should only be selected over a simpler one if the difference in its Bayesian log evidence $(\log(\mathcal{Z}))$ is greater than 2 (Delrez et al., 2021). In this work, we focused on the GP with a Matérn-3/2 kernel, as it is commonly used to model noise in TESS light curves and as it can model both smooth long-term trends as well as stochastic short-term variations. A GP can be defined as a collection of random variables, any finite number of which have a joint Gaussian distribution. It can be written as follows (Gibson et al., 2012):

$$y(x) \propto \mathcal{GP}(m(x), \Sigma) \tag{28}$$

With

1. $y(x)$: the observed values (y) depending on a variable (x), here the time.

2. $m(x)$: the mean function, here the transit model.

3. Σ: the covariance matrix.

In the case of this work, we used the covariance matrix as defined with the Matérn-3/2 kernel (Günther, Daylan, 2021):

$$\sigma^2[(1 + \frac{1}{\epsilon})e^{-(1-\epsilon)\sqrt{3}t/\rho}(1 - \frac{1}{\epsilon})e^{-(1+\epsilon)\sqrt{3}t/\rho}] \tag{29}$$

With $\epsilon = 0.01$ and two hyperparameters:

1. σ: The amplitude scale on which correlations between the data points can occur.

2. ρ: The time/length scale of the GP.

Hence, the objective of `allesfitter` will be to accurately determine the two hyperparameters to model the noise of the raw light curve in which the transits of the known planet(s) and the found candidate have been masked.

After that step, the model comparison can begin. In this work, we will focus on the comparison of two models:

1. One model with the known orbiting planet(s) and/or planet(s) candidate(s).

2. One model with the known orbiting planet(s) and/or planet(s) candidate(s) along with the found candidate.

To do so, 4 .csv files are needed:

1. **lc.csv**: This file contains the raw light curve of the target star from the `LIGHTKURVE` package. It combines the light curves obtained with each sector available for a specific cadence.

2. **params.csv**: This file contains all the parameters we want to fit. In the case of this work, those are:

   - The radius ratio of the companion to host $(R_p/R_\star)$.
   - The sum of stellar and companion radii, divided by the semi-major axis $((R_p + R_\star)/a)$.
   - The cosine of the orbital inclination of the companion $(\cos i)$.
   - The epoch, i.e. transit/eclipse midtime of the companion (T_0).
   - The orbital period of the companion (P).
   - The limb darkening coefficients (depending on the model used).
   - The hyperparameters of the GP Matérn-3/2 kernel.

   For each parameter, a statistical distribution (gaussian, uniform, ...) with specified bounds is used as the first guess. Depending on the results from the previous noise fitting, a specific distribution and bounds must be implemented for the hyperparameters.

3. **params_star.csv**: It contains the stellar parameters, and it is automatically filled by `SHERLOCK` based on TESS Input Catalog.

4. **settings.csv**: It contains all the information relative to the fitting, i.e., how the nested sampling algorithm should run.

For additional information about these files, the reader is referred to the `allesfitter` website given above. Once these files are created, `allesfitter` can run. At this point, there is a file displaying the Bayesian evidence for the tested model, allowing for model comparison. As the model with the highest Bayesian evidence is the most probable one, if it is the one with the planet candidate, it is a strong hint that there is an additional planet in this system. Finally, we can use the value of P and T_0 returned by the `allesfitter` analysis to compute the observational windows (as seen in equation 3.5) for ground-based observatories with the last module of SHERLOCK.

3.6 Computation of observational windows for ground-based observatories

The last module of SHERLOCK consists in using the transit ephemeris (P, T_0) returned by the `allesfitter` analysis to plan a follow-up campaign with a ground-based observatory. This is often a necessary step for several reasons: the point spread function (PSF) of TESS, describing the response of the optical system to a point source, can be as large as 1', while the pixel scale is 21"; there are dilution effects; and there is a non-null probability of contamination by nearby eclipsing binaries. Hence, follow-up observations from the ground are often needed to rule out these other scenarios. To achieve this, the Bayesian fit results found in section 3.5 are loaded by SHERLOCK and the user must provide a list of observatories that it can use. In the case of this study, it would be the SPECULOOS observatories. Then, transit windows are computed, taking into account constraints for the observations such as the elevation of the object, its distance from the moon, and the desired minimum transit coverage. Finally, a last *.csv* file is produced, in which all the transit windows for the target are listed.

To conclude, SHERLOCK is an end-to-end pipeline that consists of six modules that complement each other to provide a powerful tool for detecting, vetting, validating, and trigger a follow-up campaign to track new exoplanets.

4 Results

Among the list of 129 target stars, 56 were studied with the first 2 or 4 modules of SHER-LOCK depending on the case. The reason we stopped at 56 targets is to focus on the most promising signals to perform modules 5 and 6 of SHERLOCK. This was done for three potential candidates for which new observations are needed to confirm that these signals are of planetary origin. The list of all the targets analysed in this work is available in the appendix A (tables 30, 31, 32, 33, 34). A detailed description of the results will now be presented.

4.1 Most promising signals

4.1.1 TIC 321669174

This star has an effective temperature of 3742 K, a radius of 0.515 $R_\odot$ and one known orbiting planet: TOI-2081 b (Esparza-Borges et al., 2022). Information about this system is available on the Exofop website[7] and some parameters of this planet and the target star are presented in table 4.

RA (hh:mm:ss)	Dec (±hh:mm:ss)	Epoch (BJD)	Period (days)	Depth (mmag)	Duration (hrs)	Radius ($R_\oplus$)
17:37:12.74	53:01:32.53	2458685.8996	10.50534	1.342	2.371	1.838

Table 4: Stellar parameters associated with TIC 321669174 = TOI-2081 (first two columns) and planetary parameters associated with TOI-2081 b.

First, the properties file was run with all the possible parameters to fine-tune it for the searching process. As there were no high RMS values and no stellar variability, the parameters INITIAL_HIGH_RMS_MASK and AUTO_DETREND_ENABLED were disabled. Nevertheless, the Savitzky-Golay filter was used to smooth the data and reduce the local noise. Finally, the minimum and maximum periods for the transit search have been set at, respectively, 0.4 and 20 days (to keep a reasonable running time). The sectors and exposure times available for this target are all summarized in table 5. First, the 120-second exposure time was used. As a promising signal with a period of 5.109 days was found with the third run, the 600- and 1800-second exposure times were used to check whether the signal could be retrieved with different cadences. Table 6 summarizes some information obtained with the report.log file. In this table, one can see that the signal is retrieved 11 times with the

---

[7]https://exofop.ipac.caltech.edu/tess/target.php?id=321669174

120-second exposure time, so with each detrend window size tested. The signal is retrieved 10 times with the 600- and 1800-second exposure times. It has been found for all window sizes except, respectively, with the PDCSAP flux and the 0.2d window size. The fact that the transit is very shallow and relatively short (see table 7) could explain why the short window size removed (at least partially) the transit.

Exposure times (sec)	Sectors
120	14, 17, 20, 21, 24, 25, 26, 40, 41, 47, 50, 52, 53, 54, 57, 58 and 60
600	40, 41, 47, 50, 51, 53 and 54
1800	14, 17, 20, 21, 24, 25 and 26

Table 5: Summary of all the sectors and cadences available for TIC 321669174. The blue sectors are the ones with both the 120 and 1800s exposure times. The orange sectors are the ones with both the 120 and 600s exposure times.

Exp time (sec)	Run	P(days)	Number of findings	SNR	SDE	BS	FAP
120	1	10.50	11	42.68	128.42	0.975	8.0032e-05
	2	19.016	1	18.17	120.46	0.888	8.0032e-05
	3	5.109	11	22.03	200.24	0.986	8.0032e-05
	4	19.51	6	9.12	47.32	0.95	8.0032e-05
600	1	10.50	11	13.00	49.76	1.0	8.0032e-05
	2	5.109	10	6.31	16.64	1.0	8.0032e-05
	3	19.30	6	6.2	22.85	0.9	8.0032e-05
	4	13.67	7	4.10	27.01	1.0	8.0032e-05
1800	1	10.50	11	11.17	40.34	1.0	8.0032e-05
	2	5.109	10	6.64	16.88	1.0	8.0032e-05
	3	18.74	1	6.02	18.71	1.0	8.0032e-05
	4	18.84	1	5.21	31.66	0.888	8.0032e-05

Table 6: Results from the report file. The number of findings represents the number of times the signal was found out of the 11 window size possibilities. The green lines correspond to the runs that found the same candidate with a period of 5.109 days.

Table 7 summarizes the parameters found for the candidate. Figure 27 shows the phase-folded data and the transit, while figure 28 shows all the data as a function of time with the associated periodogram. In both cases, it was the signal found with the 120-second exposure time and for a window size of 1.1 days, i.e., the one that gave the highest SNR for the signal. Even if the signal is very shallow, there is a clear decrease in flux. Moreover, the SDE

peak associated with the signal is very clear and some peaks can also be seen for the harmonics. Also, the border score (BS) and FAP values (see table 6) are really convincing. As all those results were promising, the vetting for the 120-second exposure time was performed.

T0 (TBJD)	Period (days)	Duration (hrs)	Depth (ppt)	Nbr of transits found
1686.9803	5.1097	1.39	0.52	77

Table 7: Parameters associated with the found candidate.
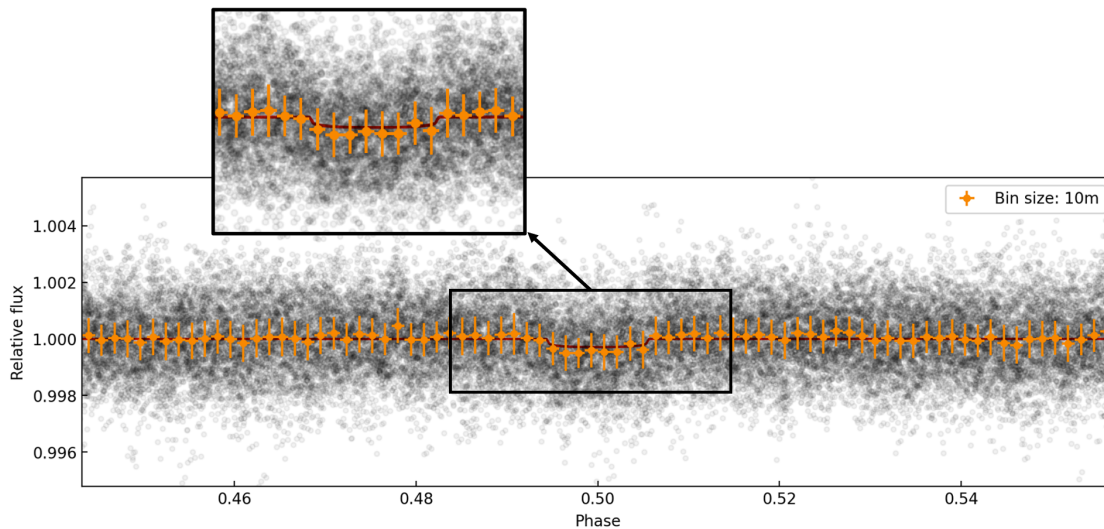
Figure 27: Detrended and phase-folded fluxes over the 5.1d period detected for TIC 321669174 during the third run with a 120-second exposure time and a window size of 1.1 days. As the transit is very shallow, a zoom was made to highlight the shape of the transit.
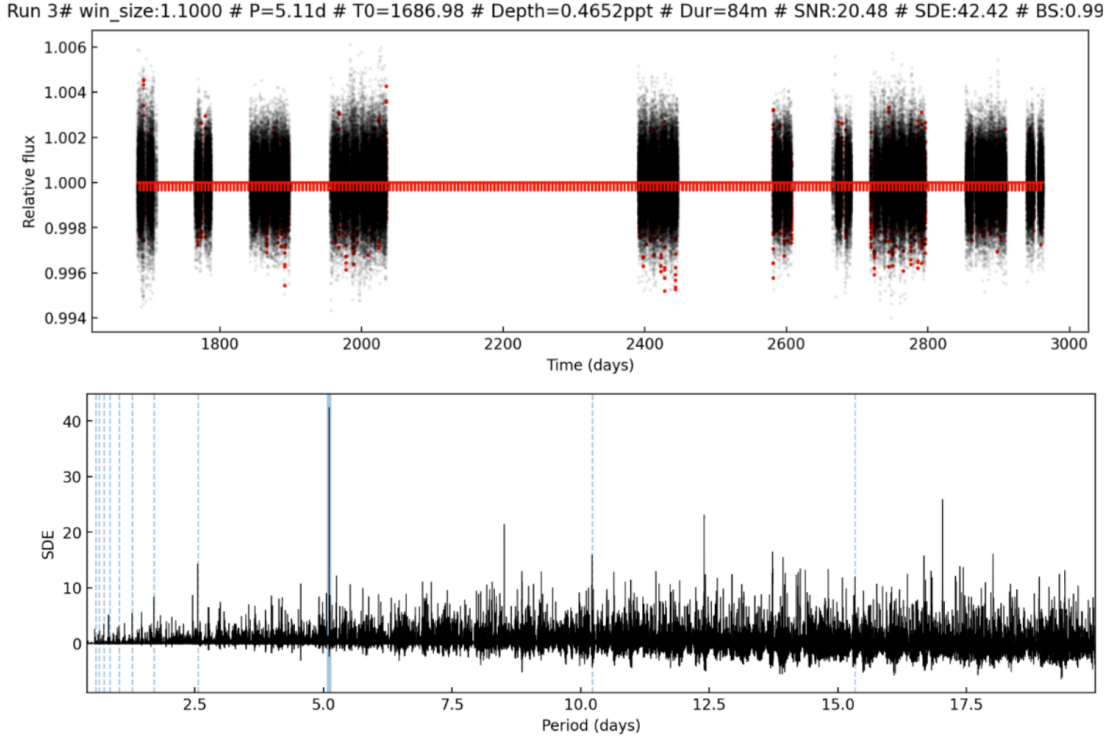
Figure 28: Results obtained for TIC 321669174 during the third run, with a 120-second exposure time and a window size of 1.1 days.

Results of the vetting report:

- Folded curves: The folded curve for the fast cadence was not available as no data were obtained with a 20-second exposure time. For all the available cadences, the folded curves show a very shallow transit signal.

- Transit depth plot: In figure 29, one can see that, out of 75 transits, there is no significant difference in depth between the even and odd transits. Hence, based on this plot only, an eclipsing binary system is not suspected.

- Single-transits vetting sheets: There are no noticeable and/or unexpected features in the vetting sheets for each transit. The vetting sheet of the 48th transit is shown in figure 30 as an example.

- Source and target positions: In figure 31, one can see that the target star (blue star symbol) and the transit source position (red dot) are almost on top of each other. From this figure, the transit signal probably originates from the target.

56

- Folded light curves with the original period and first (sub)harmonics: First, in figure 32, the left and right columns are, respectively, the inferior and superior conjunctions. Second, at the top and the middle, corresponding, respectively, to the original and harmonic cases, every curve has the expected shape. Indeed, at the top left and the middle, a transit shape is visible, while at the top right, no transit shape is seen. At the bottom, corresponding to the sub-harmonic case, no transit is observed at the right, as expected. Finally, there is no transit shape at the bottom left but it is not worrying as the SNR is very low (as explained in section 3.3.3).

- The Target Pixel Files pixels folded curves: The TPF for each sector does not reveal any peculiarities coming from the pixels outside the aperture. The results for sector 41 are shown in figure 33 as an example.

To conclude, the vetting report does not provide any worrying results, increasing, even more, the confidence in the candidate. Hence, the statistical validation was run for the short cadence.
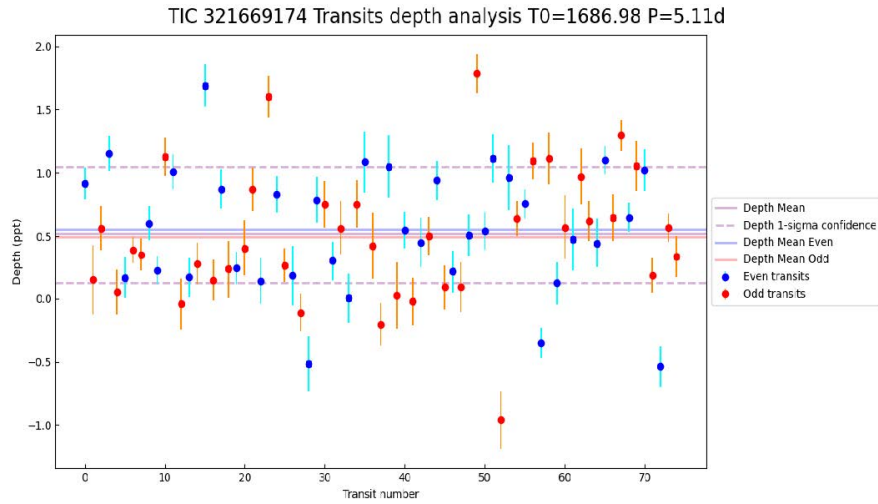
Figure 29: Single-transits depths plot for the candidate with a period of 5.109 days.
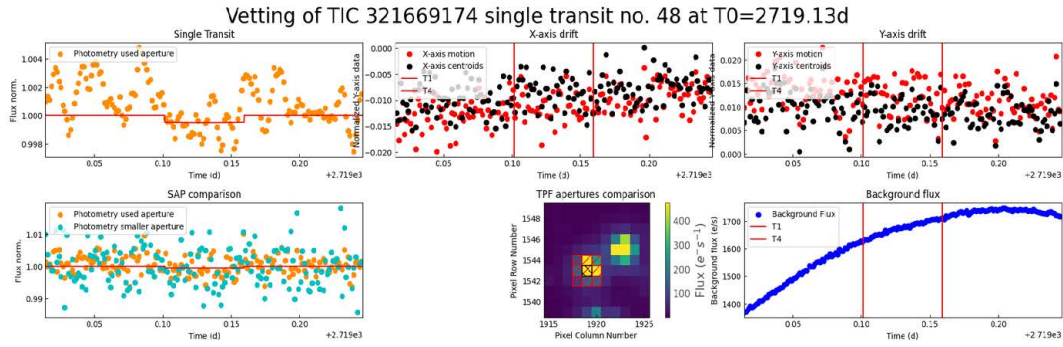
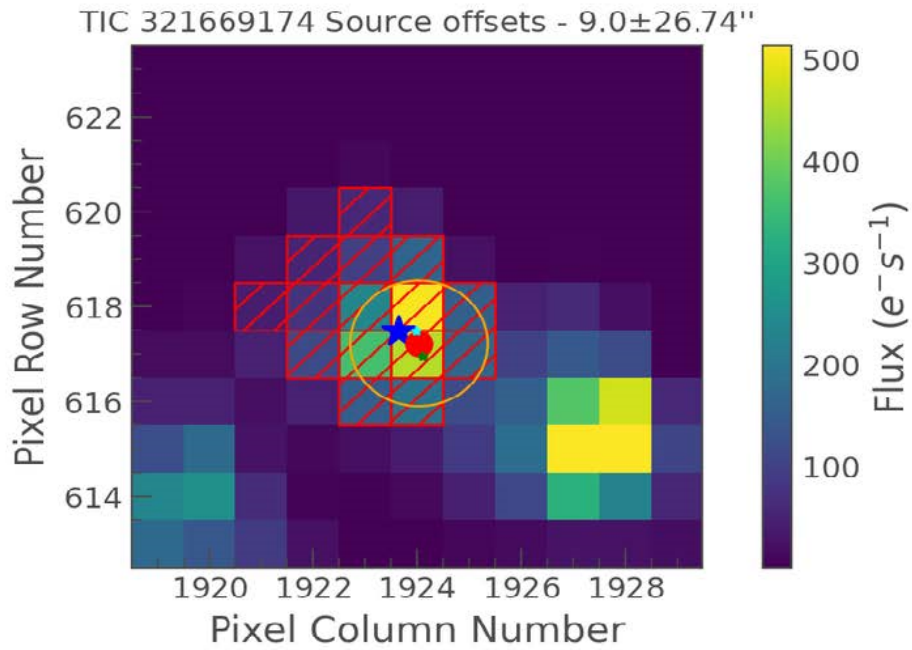Figure 30: Vetting sheet of the 48th transit of the found candidate.

Figure 31: The transit source (red dot) and target (blue star symbol corresponding to TIC 321669174) positions are shown.
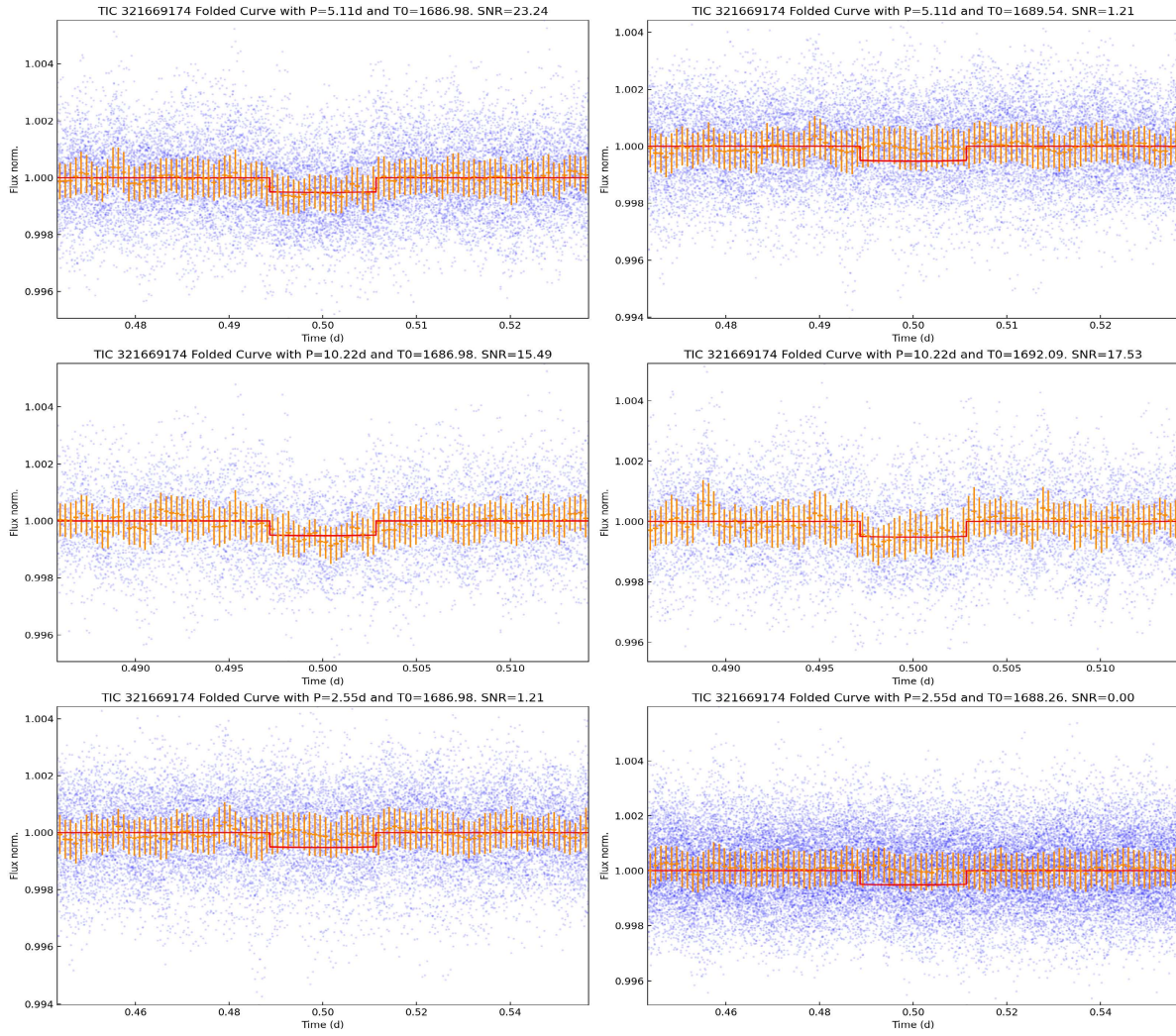
Figure 32: Folded curves for TIC 321669174. The left and right columns are, respectively, for the inferior and superior conjunctions. The top, middle and bottom lines are, respectively, for the original, harmonic and sub-harmonic cases.
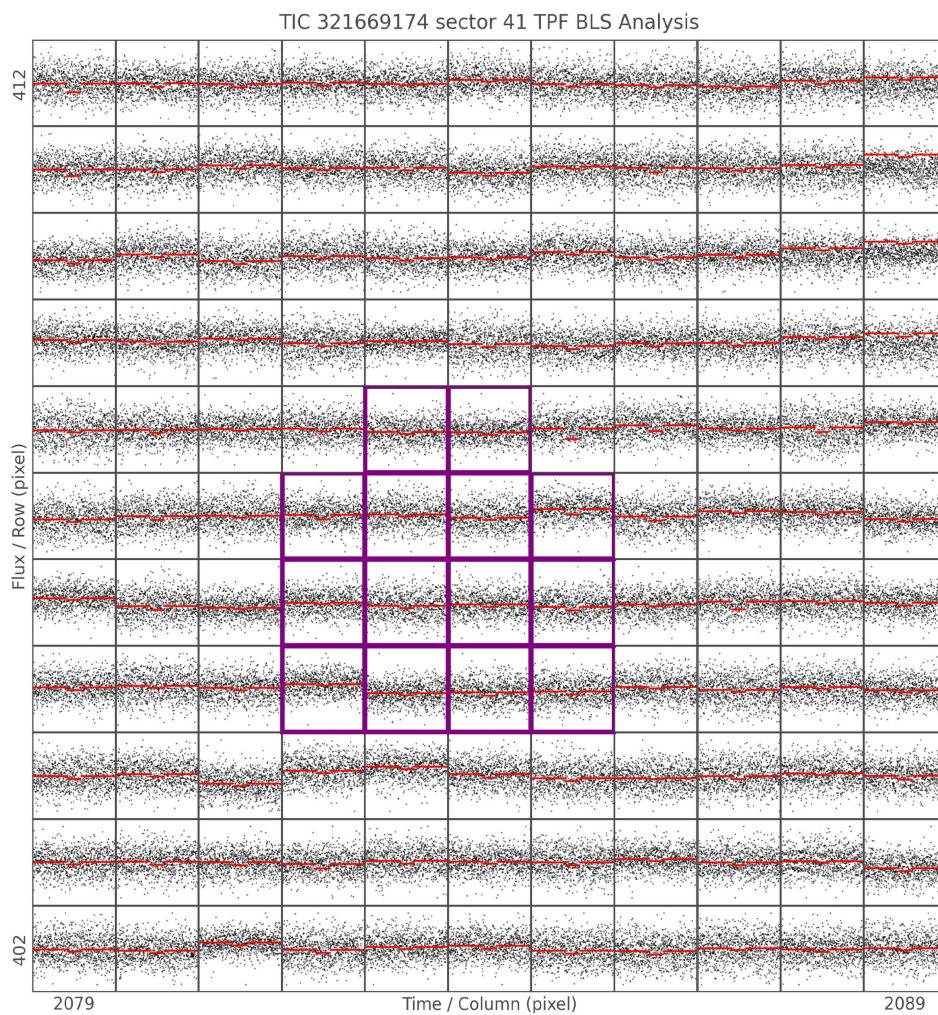
Figure 33: Target Pixel File of TIC 321669174 for sector 41 with the photometric aperture in purple. By eye, no transit shape can be seen outside the aperture.

The results of the statistical validation report will now be reviewed.

- Folded curve of the candidate: In figure 34, the transit shape of the candidate is clearly visible even if it is very shallow.

- Field of view: In figure 35, all the stars that may contribute to the observed transit signal, i.e. all stars within a 10-pixel radius, are shown with dots. One can see that there is a faint star next to the target star.

- Most probable scenarios: In table 8, the most probable scenario is the TP one, i.e. the scenario for which the target star has no unresolved companion and the transit signal is due to a transiting planet around the target star with the reported orbital period (cfr table 2). The next most probable scenarios involve an unresolved background star (BEB, DTP) or an unresolved bound companion star (PTP). However, in the discovery paper of TOI-2081 b (Esparza-Borges et al., 2022), they used ground-based follow-up data (multi-color transit photometry, spectroscopy) to show that these scenarios involving an unresolved source (bound or unbound) are very unlikely (the possible flux contamination from a possible unresolved companion is found to be negligible). Hence these scenarios seem also highly unlikely for the candidate found here.

- FPP/NFPP map: In table 9 and in figure 36, one can see that the candidate does not fall into the "likely planet" category because the NFPP is larger than 10^{-3}. However, the risk that the transit is due to a resolved nearby star is rather unlikely. Indeed, the ground-based follow-up observations (at higher resolution than TESS) of the first known planet presented in the discovery paper showed that its transit signal is well on the target star and not on any resolved nearby star (including the closest one in figure 35). We can thus expect the second signal to be on the target as well. Hence, the NFPP value should be smaller. Ground-based follow-up observations of transit windows of the candidate at a higher angular resolution than TESS could help to confirm that this second signal is also on the target, and not due to a nearby star.

To conclude, even if the candidate is not classified as a likely planet, one must keep in mind that all the metrics and plots available are mostly meant to decide whether to proceed to the next module of SHERLOCK, or not. Hence, given all the results presented above, it seems appropriate to continue with the fitting module.
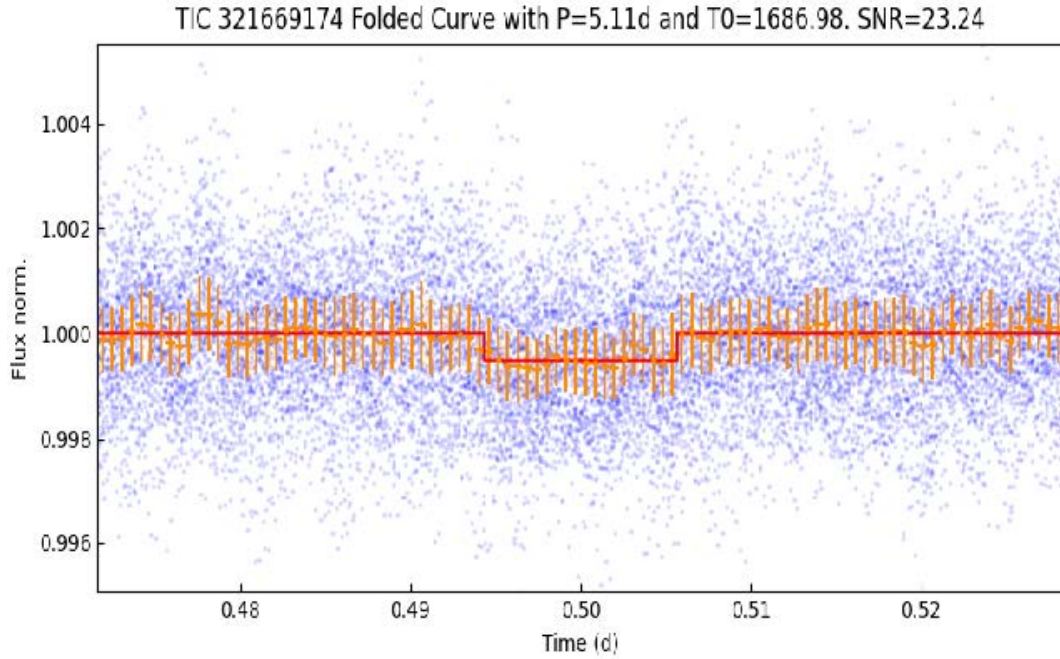
Figure 34: Folded curve of the candidate found for TIC 321669174.
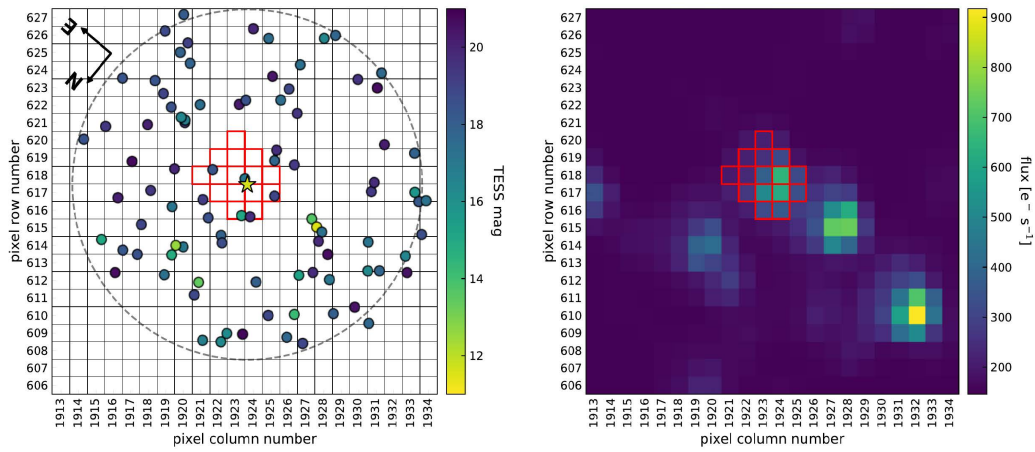

Figure 35: Field of view of the target star, TIC 321669174. All the nearby stars are depicted with a dot while the target star is the yellow star symbol.

TP	BEB	DTP	PTP
0.452	0.110	0.086	0.083

Table 8: Table summarizing the 4 scenarios with the highest probability for the candidate of TIC 321669174 with a period of 5.109 days.

FPP	NFPP	FPP2
0.377	0.044	0.023

Table 9: Mean values of the FPP, NFPP and FPP2 for the candidate with a period of 5.109 days.
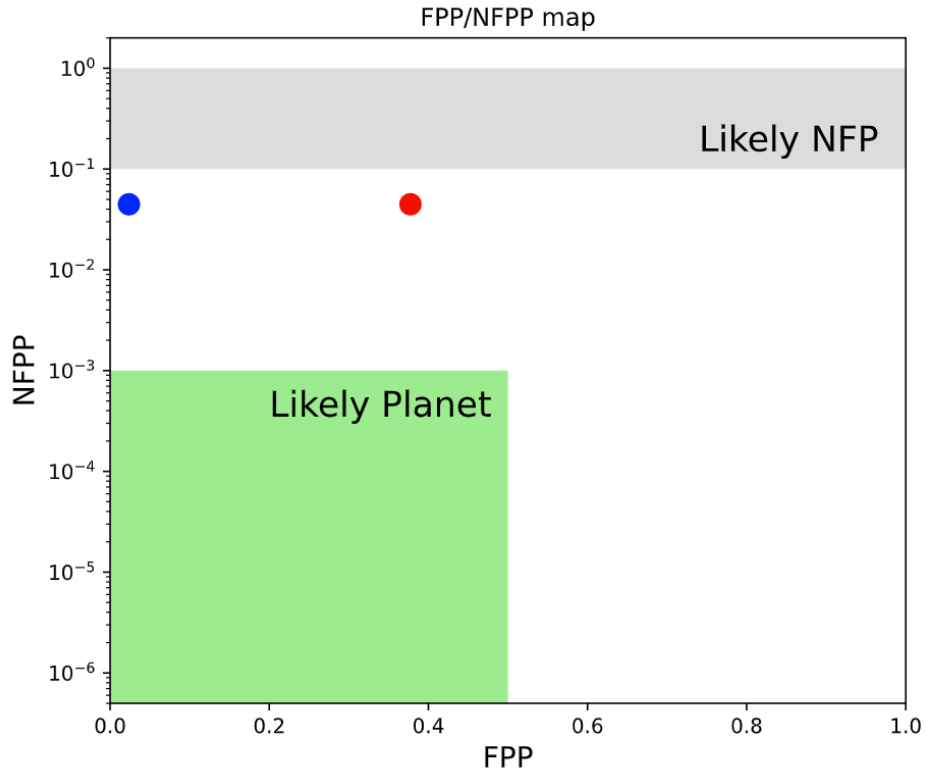

Figure 36: FPP/NFPP map for the candidate of TIC 321669174. The red and blue dots are, respectively, for the FPP and FPP2 cases.

The first step of the fitting is to run SHERLOCK with the fifth module. By doing so, we obtained the 4 .csv files (see section 3.5) with the parameters associated with the known orbiting planet (TOI-2081 b), the planet candidate and the stellar parameters (of TIC 321669174) automatically filled by SHERLOCK from data based on TESS Input Catalog. We did not use the lc.csv file given by SHERLOCK for the reasons explained in section 3.5 but we took the original light curve with the LIGHTKURVE package and we masked each transit of TOI-2081 b and of the planet candidate. Then, the second step is to use those files to do the noise fitting with allesfitter. To do so, we disable the fitting of every parameter associated with something else than the noise to fit the hyperparameters of the Matérn-3/2 kernel and the error on the relative flux. In this case, it seemed that the detrending applied on the light curve with the values obtained for the noise fitting was too strong. Hence we should have started the fitting with new bounds but due to some computational problems while running allesfitter, it prevented the final fitting. Hence, the results presented below are the ones obtained with the first fitting.

The second step is to use the values obtained with the first noise fitting as Gaussian priors on the noise parameters during the fit of the two models: Model 1 with only one planet (TOI-2081 b) and Model 2 with two planets (TOI-2081 b and the planet candidate). For this, we used the original light curve with no masks and we allowed the fitting of every parameter listed in section 3.5. The results from the two models for the parameters associated with those planets are available in table 10. Then, one can compare the Bayesian evidence of the two models in table 11. From this result, it is clearly the second model (two planets) that is favoured as the Bayesian evidence is much larger than with the first model (i.e. 99 497 additional units). The corner plot of the fitted and derived parameters for Model 2 is available for the interested reader in the appendix B, respectively, in figures 69 and 70.

The fitting of the light curve to model the transit is shown in figure 37. A last point to check is to see if the host stellar density obtained with TOI-2081 b and with the found candidate is similar. Indeed, in equation 8, this quantity can be derived from the parameters of the planets in the system. The results are as expected as the host stellar density is 5.33 ± 0.5 g/cm^3 with TOI-2081 b and 5.29 ± 0.52 g/cm^3 with the found candidate. From these results, it seems likely that there is a second planet in this system with an orbital period of 5.109 days.

	Model 1: TOI-2081 b	Model 2: TOI-2081 b	planet candidate	
$T0[\text{TBJD}]$	$2326.723^{+0.003}_{-0.004}$	$2326.723^{+0.001}_{-0.001}$	$2325.682^{+0.002}_{-0.002}$
$P[days]$	$10.505^{+2.57e-5}_{-3.17e-5}$	$10.505^{+2.64e-5}_{-3.094e-5}$	$5.109^{+2.815e-05}_{-3.38e-05}$
$R_p[R_\oplus]$	$1.858^{+0.086}_{-0.086}$	$1.838^{+0.085}_{-0.085}$	$1.227^{+0.069}_{-0.087}$
a/R_*	$31.44^{+0.99}_{-1.1}$	$31.44^{+0.91}_{-1.0}$	$19.40^{+0.64}_{-0.64}$
$a[AU]$	$0.0753^{+0.0034}_{-0.0034}$	$0.0753^{+0.0033}_{-0.0033}$	$0.0464^{+0.0021}_{-0.0021}$
$i[deg]$	$89.35^{+0.28}_{-0.21}$	$89.39^{+0.27}_{-0.21}$	$88.30^{+0.51}_{-0.28}$
b	$0.35^{+0.099}_{-0.15}$	$0.34^{+0.10}_{-0.15}$	$0.576^{+0.083}_{-0.17}$
$T_{eq}[K]$	432 ± 24	432 ± 24	550 ± 31
$\delta[ppt]$	$1.335^{+0.075}_{-0.075}$	$1.348^{+0.072}_{-0.072}$	$0.547^{+0.041}_{-0.046}$
$\rho_*[g/cm^3]$	5.33 ± 0.53	5.33 ± 0.5	5.29 ± 0.52

Table 10: Planet parameters obtained through the fitting of the two models (1 vs 2 planets) with `allesfitter`. Note that the epoch has been shifted to be in the middle of the data set to avoid correlations between epoch and period.

Bayesian evidence ($log(\mathcal{Z})$)	
Model 1	Model 2
52420.362 ± 0.246	151917.689 ± 0.304

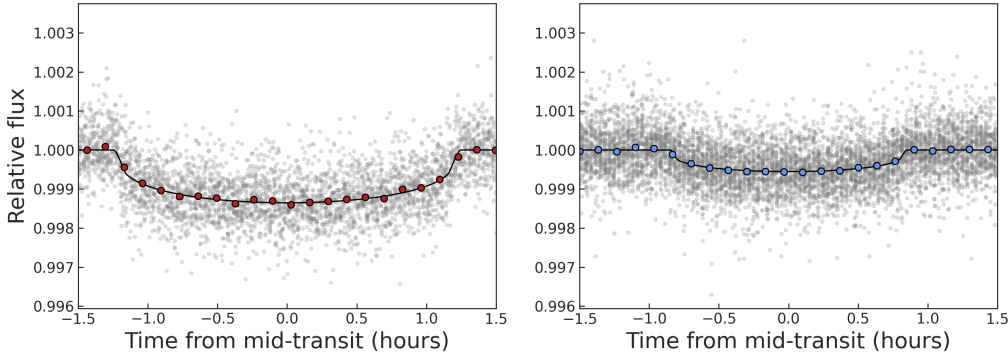Table 11: Bayesian evidence of the two models obtained with `allesfitter`.

Figure 37: Fitting of the planetary transits (folded light curves), with the transit of TOI-2081 b on the left and the found candidate on the right.

The discovery of this candidate was made in March/April 2023 but on the 19th of May 2023, this candidate was reported on the ExoFOP website. The parameters match the ones found in this work with a period of 5.109 days and an epoch of 1686.96595 TBJD[8]. Hence, the discovery made throughout this work is consistent with the one reported, indicating the efficiency of the SHERLOCK tool and the methodology used throughout this work.

Finally, further observations are needed to confirm the planetary origin of this signal. Unfortunately, as the transit depth is about 0.52 ppt, it is too shallow to be observable with the SPECULOOS telescopes. Indeed, the transit depth must be at least 1 or 2 ppt to be observable (private communication with the SPECULOOS team). One solution is to wait for the results of the extended mission of TESS to potentially observe new transits that match the signal found. Another solution is to use ground-based observatories to observe the stars in the vicinity of TIC 32166974 to see if those stars could mimic the transit found. If not, those observations would reduce the scenarios involving resolved nearby stars and reduce the NFPP value to zero. However, as already mentioned, due to the analysis made in Esparza-Borges et al. (2022), those scenarios are unlikely.

---

[8]https://exofop.ipac.caltech.edu/tess/target.php?id=321669174

4.1.2 TIC 244170332

This star has an effective temperature of 3888K, a radius of $0.525 R_\odot$ and one known planet candidate: TOI 5530.01. Information about this system is available on the Exofop website[9] and some parameters associated with this TOI and the target star are presented in table 12.

RA (hh:mm:ss)	Dec (±hh:mm:ss)	Epoch (BJD)	Period (days)	Depth (mmag)	Duration (hrs)	Radius $(R_\oplus)$
0:25:57.19	-5:41:48.54	2458382.1709	0.47743	0.459	0.952	1.118

Table 12: Stellar parameters associated with TIC 244170332 (the first two columns) and planetary parameters associated with TOI 5530.01

First, the properties file was filled, as explained in section 3.2.1. As no high RMS value and no large stellar variability were detected, associated parameters (INITIAL_HIGH_RMS_MASK and AUTO_DETREND_ENABLED) were disabled in the properties file, while the Savitzky-Golay filter was still applied. For the search, the number of runs to search for periodic signals with a maximum period of 20 days was set to 4. The global mode was used to combine sectors 3, 30 and 42.

For this target, two different exposure times are available: 120 seconds (for all sectors) and 600 seconds (for sectors 30 and 42 only). The results obtained for each run and exposure time are available in table 13. We can see that TOI 5530.01 was correctly recovered during the first run for both exposure times. Then, a transit signal with a period of 16.20 days was found at the second run for both cadences (green lines in table 13). For the exposure time of 120 seconds, this signal was found with 4 different window sizes (0.2, 0.3, 0.4 and 0.5 day), while with the 600-second exposure time, the signal was found for all window sizes (from 0.2 to 1.1 day). The fact that the signal is found with short and long window sizes with the 600-second exposure time is a good indicator. As expected, low SNR and SDE values are found in both cases. Finally, the FAP is very low and an excellent value for the border score (1.0) is found, increasing the confidence in the signal.

The signal found for the 120 seconds exposure time and with a window size of 0.4 days (giving the highest SNR) is shown in figure 38. The transit data points are shown in red

---

[9]https://exofop.ipac.caltech.edu/tess/target.php?id=244170332

Exp time (sec)	Run	P(days)	Number of findings	SNR	SDE	BS	FAP
120	1	0.47742	11	14.11	24.68	0.983	8.0032e-05
	2	16.20	4	8.73	79.55	1.0	8.0032e-05
	3	12.19	7	7.45	99.84	1.0	8.0032e-05
	4	3.88	1	8.72	252.64	1.0	8.0032e-05
600	1	0.47742	11	6.941	16.669	0.987	8.0032e-05
	2	16.20	11	5.386	16.989	1.0	8.0032e-05
	3	15.32146	1	3.954	25.503	1.0	8.0032e-05
	4	/	/	/	/	/	/

Table 13: Results from the report file for the 120-second exposure time (sectors 3, 30 and 42) and for the 600-second exposure time (sectors 30 and 42). The number of findings represents the number of times the signal was found out of the 11 window sizes possibilities. The last line (fourth run) is empty for the 600-second exposure time because the signal found at the third run was not promising enough for SHERLOCK to continue.

in the upper plot. The transit depth looks consistent between the last 3 transits, while the first one is less deep and is almost not visible in the figure. The 2 long gaps in the plot are explained by the time difference between the 3 sectors, while smaller gaps are produced when data are sent to the Earth. In the middle plot, the transit shape can clearly be seen at phase 0.5. Finally, the SDE values are shown in the bottom plot. Peaks are visible at the period of the detected signal and its harmonics. The signal can also be seen in figure 39 for the 600-second exposure time. In this case, only 2 sectors are available and the first transit seems deeper than the others.

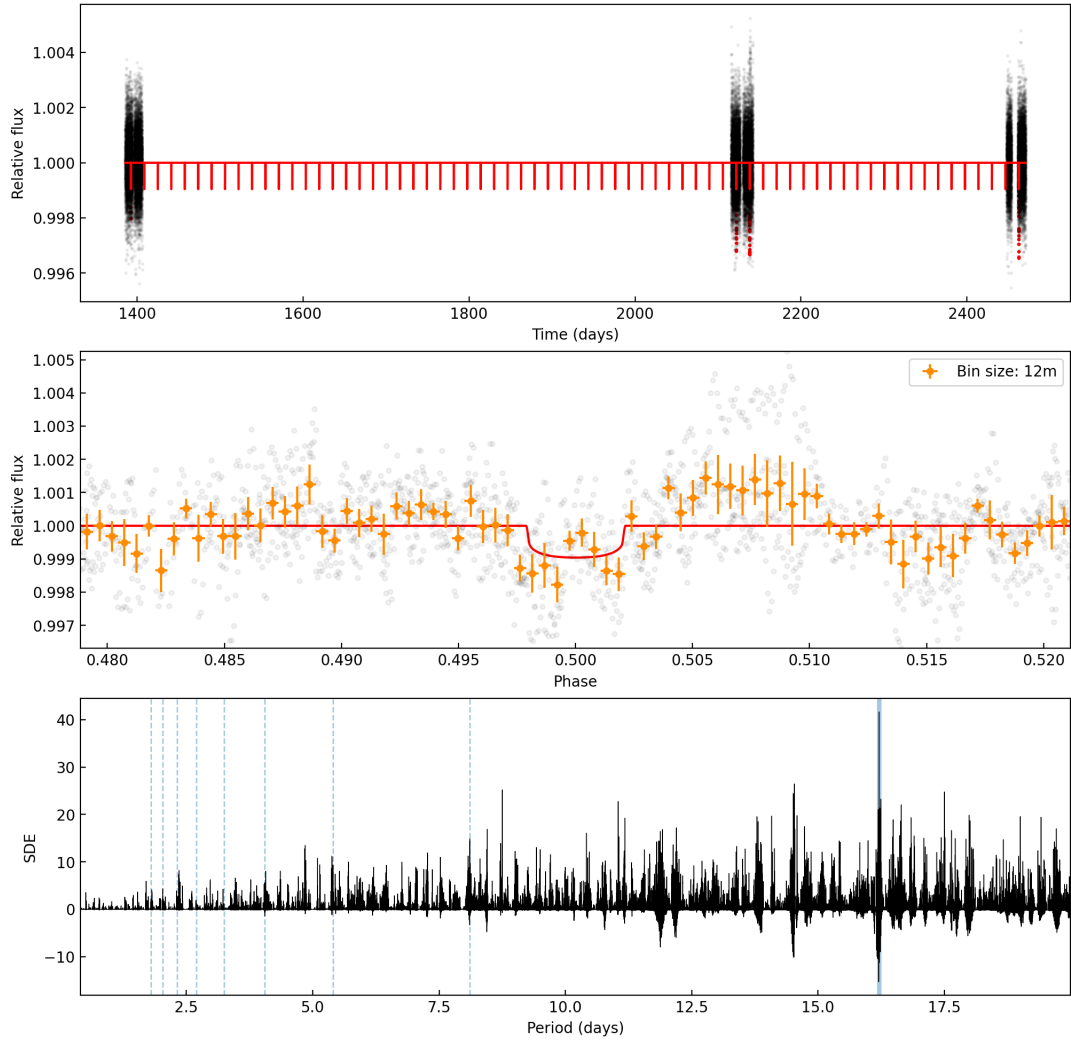Run 2# win_size:0.4000 # P=16.21d # T0=1392.66 # Depth=1.0151ppt # Dur=99m # SNR:13.22 # SDE:41.63 # BS:1.00

Figure 38: The results obtained for TIC 244170332 during the second run with a 120-second exposure time and a window size of 0.4 days.
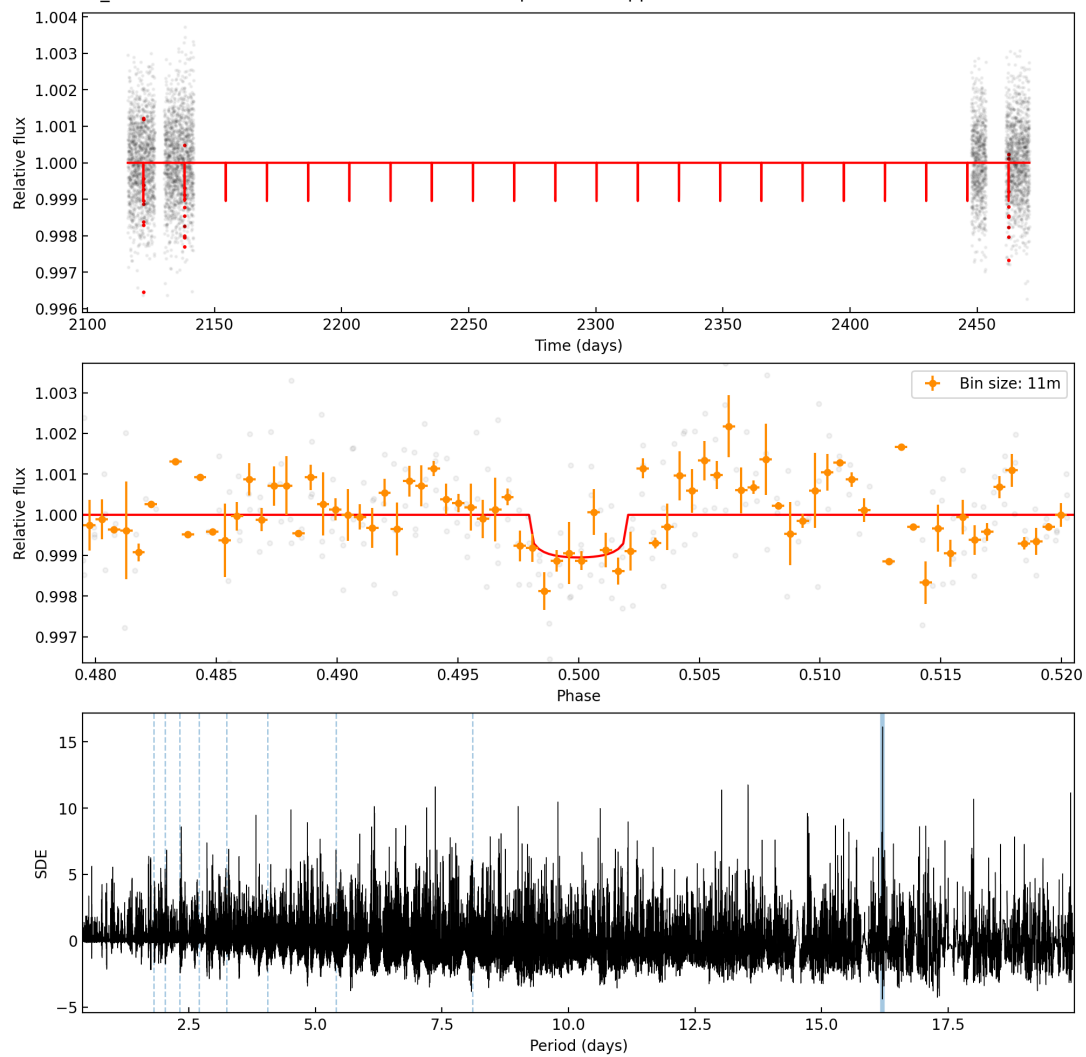
Figure 39: The results obtained for TIC 244170332 during the second run with a 600-second exposure time and a window size of 0.4 days.

As the signal looked promising enough and was also found independently with the 2 exposure times, the 3rd module of SHERLOCK was used to perform the vetting. First, the candidate parameters were computed. They are listed in table 14.

T0 (TBJD)	Period (days)	Duration (hrs)	Depth (ppt)	Nbr of transits found
1392.6679	16.208	1.49	0.657	4

Table 14: Parameters associated with the found candidate.

We will now review the results of the vetting report:

- Folded curves: The folded light curves for the short and long cadence data were available. In both cases, the transit signal can be seen.

- Transit depth plot: In figure 40, the candidate single-transit depths plot is shown. As there is no clear difference between even and odd transits and as there are only 4 transits, there is no evidence that the candidate could result from a binary system.

- Folded light curves with the original period and first (sub)harmonics: Figure 41 shows the light curves at inferior (left) and superior (right) conjunctions for the original (top), the first harmonic (middle) and the sub-harmonic (bottom) cases. Among the 6 plots, 3 of them display clear and expected results. Those are the top and middle left plots where a clear transit shape is observed, and the bottom right plot where there is no transit. This was expected, as explained in section 3.3.3. The results for the 3 other plots are less obvious. At the top right, no transit seems to be observed, as expected. For the middle right and bottom left, it seems difficult to conclude anything.

- Source and target positions: In figure 42, the transit source (red point) and the target (blue star) are almost perfectly on top of each other. As a result, the observed signal probably originates from the target star, increasing confidence in the candidate.

- The Target Pixel Files pixels folded curves: In figure 43, there is the target pixel file for sector 42. In this plot, one can only conclude by eye that there is no pixel outside the aperture with a transit pattern that could be responsible for the observed transit. Indeed, there is not any clear area in the target pixel file where the red models match transit-like structures out of the photometric aperture. Hence, it seems that there is no obvious source of contamination from stars in the vicinity of the target star.

- Single-transits vetting sheets: In figure 44, the vetting sheet (as explained in section 3.3.6) of the third transit is shown. Those plots do not feature any unexpected behaviour. From this, one can conclude that the signal is probably not due to a background signal, and as there is no noticeable difference between the two aperture sizes and no x and y offset, the signal is unlikely to be an instrumental false positive.

To conclude, given the results obtained with the vetting, the 4th module of SHERLOCK was run.
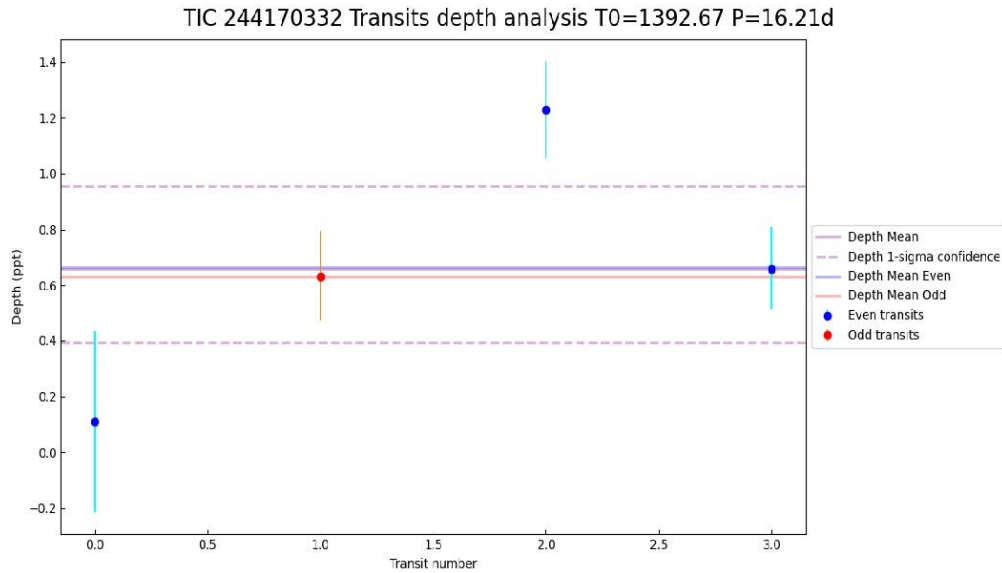
Figure 40: Single-transit depth plot for the candidate with a period of 16.2 days.
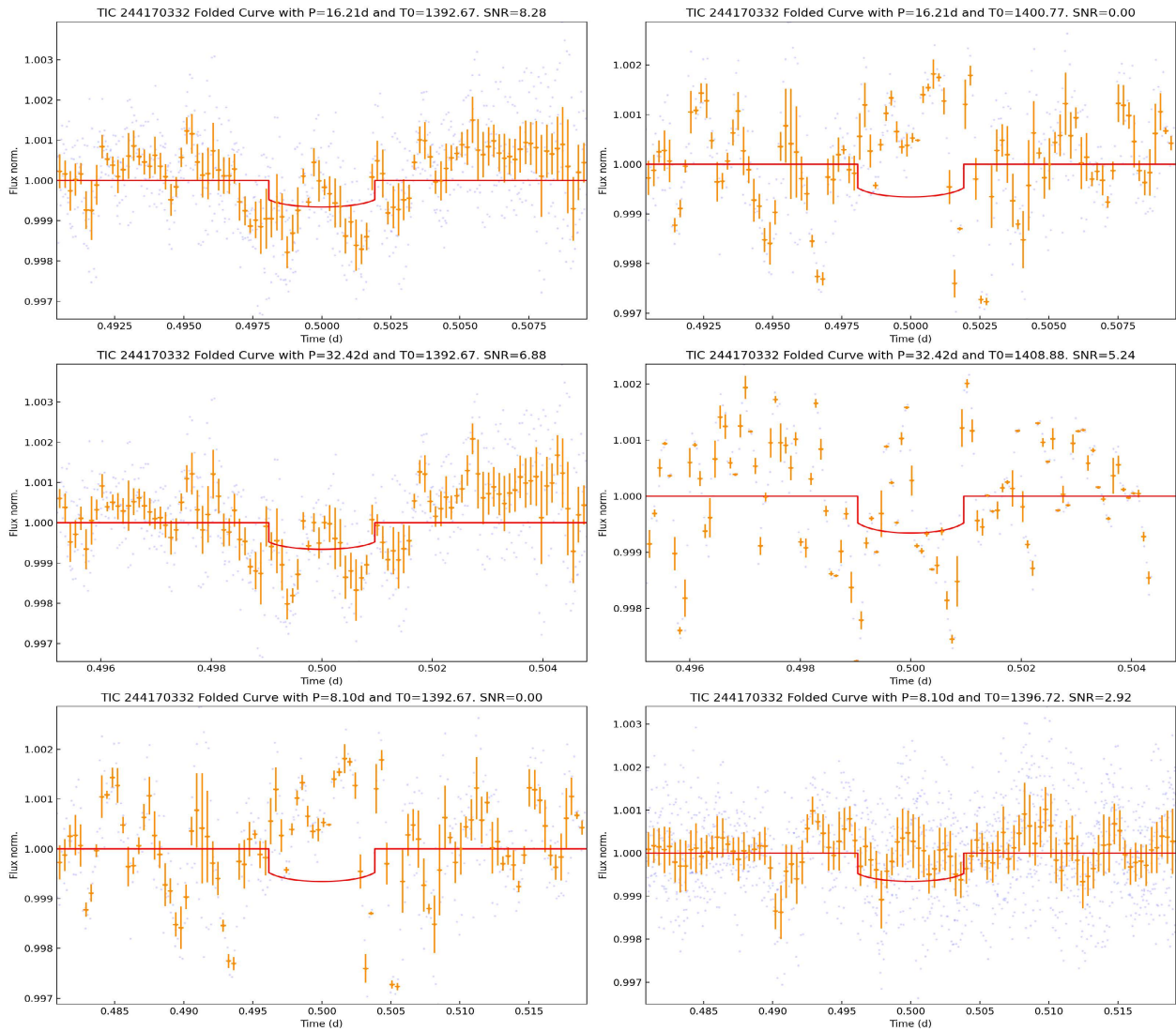
Figure 41: The first row represents the case with the original period of 16.2 days, the second row represents the first harmonic with P = 32.42 days and the third row is for the first sub-harmonic case with P = 8.1 days. The left and right columns are respectively for the inferior and superior conjunction.
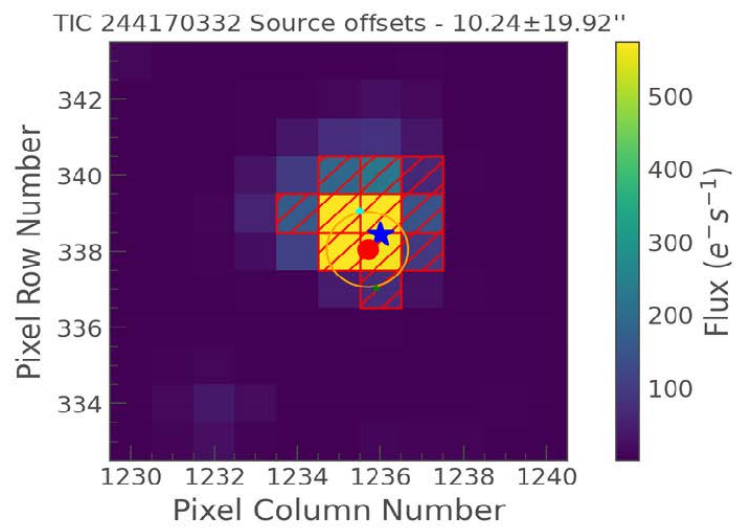
Figure 42: The transit source (red point) and target (blue star, TIC 244170332) positions are shown.
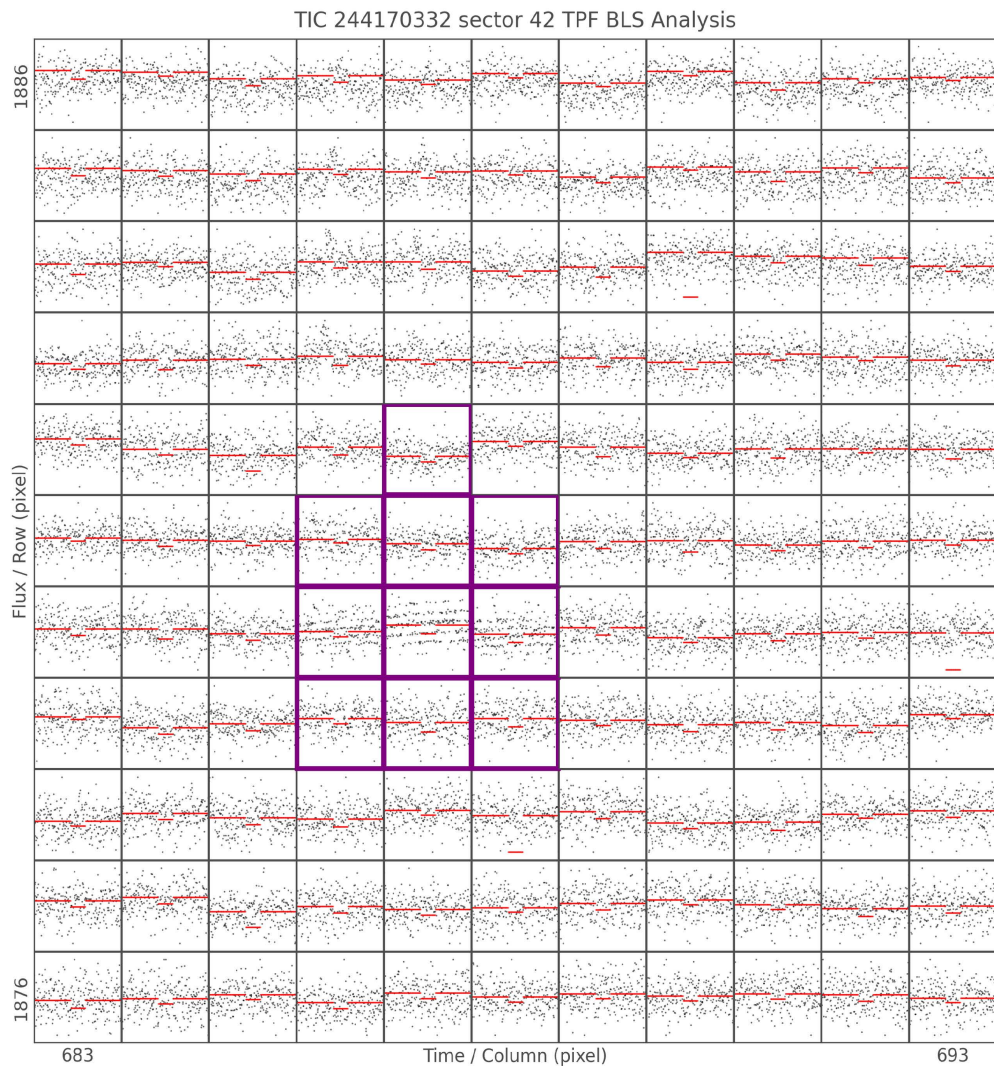
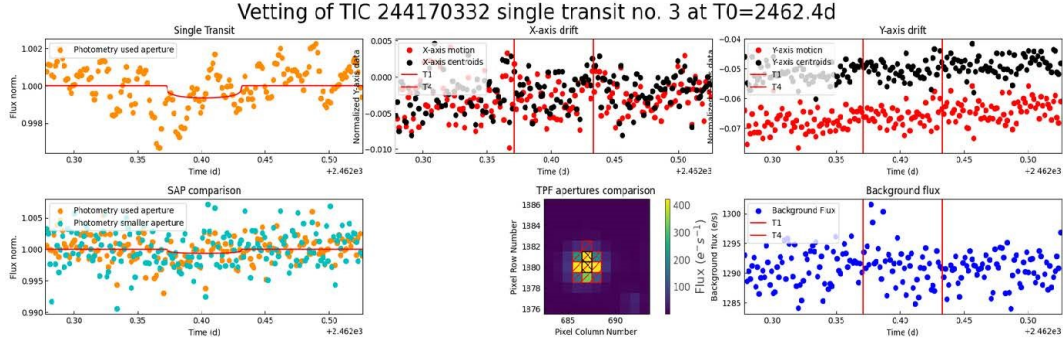Figure 43: The target pixel file of TIC 244170332 for sector 42.

Figure 44: Vetting sheet of the third transit observed for the candidate of TIC 244170332.

After running the 4th module of SHERLOCK, the statistical validation report was available. For this target, high angular resolution imaging was obtained in the b and r bands (respectively 562 and 832 nm) with the WIYN-3.5m telescope and the NESSI instrument[10]. The contrast curves available on the ExoFOP website were used to increase the accuracy of the FPP value. Those contrast curves are shown in figure 45. The blue curve probes slightly smaller angular separations (< 0.2 arcsecs), while the red curve is more sensitive to fainter stars (as it is more profound) for separations above 0.2 arcsecs. Therefore, the statistical validation module was run 3 times to compare the results between the r and b bands and the ones without any contrast curve.

First, the folded curve of the candidate for the short cadence is shown in figure 46. A transit shape with a bump inside is visible, the latter being likely caused by the Savitzky-Golay filter that can cause substantial artificial variability. Then, there is the sky map in figure 47. In the graph on the left, two nearby stars appear clearly in the aperture, within 2 pixels of the target star. As the flux of each star within the grey circle has been taken into account in the statistical computation, the scenarios with the highest probabilities will have to be analysed carefully to see if those stars may be responsible for the observed transit signal. The summary of the most probable scenarios for the 3 cases (with and without the contrast curves) is shown in table 15. The most probable scenario is always the TP scenario with a net increase (~ 0.2) for the TP scenario when contrast curves are included. As the TP scenario is the case where the star has no unresolved companion and where the transit is due to a transiting planet of a period of 16.2 days around the target star, the high value for

---

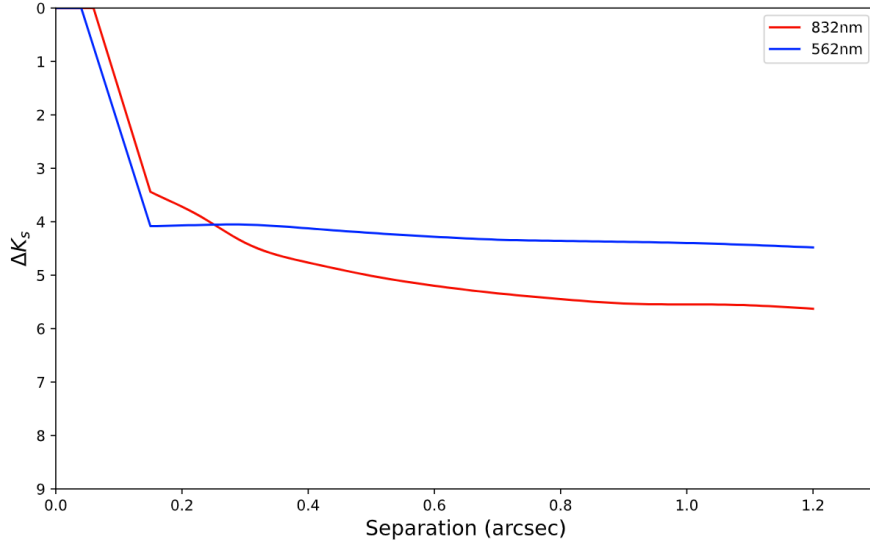[10]https://www.wiyn.org/Instruments/wiynnessi.html

Figure 45: Contrast curves from the NESSI instrument on the WIYN-3.5m telescope. The blue and red curves are, respectively, the contrast curves for the b (562nm) and r (832nm) bands. Data were obtained on the ExoFOP website.

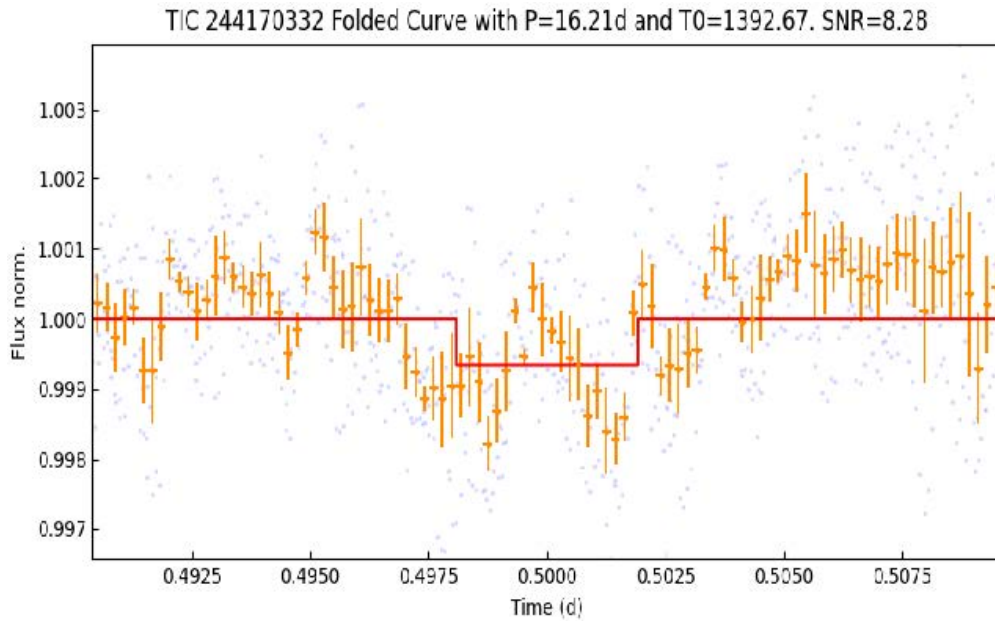this scenario is a good indicator that the candidate could be a real planet.

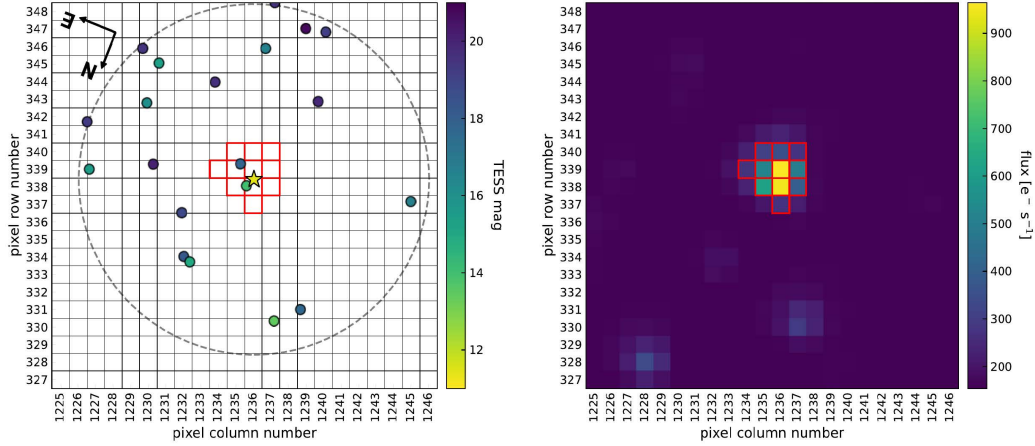Figure 46: Folded curve of the candidate with a period of 16.2 days.

Figure 47: Sky map of the nearby stars surrounding TIC 244170332 with the aperture represented in red.

b band	r band	No contrast curve
TP: 0.856	TP: 0.851	TP: 0.656
NTP: 0.019	PTP: 0.022	DTP: 0.128
PTP: 0.017	DTP: 0.019	PTP: 0.121

Table 15: Summary of the three scenarios with the highest probability when including the contrast curve in the b and r bands and for the case where no contrast curve is included.

Finally, there is table 16 associated with figure 48. The mean FPP/NFPP values obtained after the 5 runs are shown for the two contrast curves and without, and the FPP/NFPP map is shown in figure 48. As expected, the effect of including the contrast curves was to decrease the FPP value. Unfortunately, the candidate is never classified as a likely planet because the NFPP value is larger than 10^{-3}. However, given the results mentioned above and the fact that the signal is found with 2 different exposure times, an in-depth investigation is worthwhile for this candidate. Consequently, the fitting module of SHERLOCK was run.

	FPP	NFPP	FPP2
Contrast curve r band	0.039286	0.020292	0.001633
Contrast curve b band	0.03665	0.020213	0.001519
Without contrast curve	0.093876	0.015632	0.004127

Table 16: Summary of the mean statistical values obtained with and without the contrast curves.

As for the first promising candidate, the first step was to fit the noise and the error on the relative flux. The exact same methodology as with the fitting of TIC 321669174 was used here. First, we conducted a fitting process on the light curve while masking the transits (corresponding to TOI 5530.01 and the found planet candidate). This approach allowed us to constrain the noise parameters for the two hyperparameters and the error associated with the relative flux. Second, we used these values as Gaussian priors on the noise parameters during the fit of the two models: one with TOI 5530.01 alone and one with TOI 5530.01 and the found planet candidate with a period of 16.208 days. The results for the fitted and derived parameters are given in table 17. For the interested reader, the corner plots of the fitted and derived parameters are available in the appendix B (respectively, in figures 71 and 72). Then, table 18 shows the Bayesian evidence for the two models. The logarithm of the Bayesian evidence associated with the model with an additional planet is greater than 7 units compared to the one associated with the model with only one planet. Hence, the model with two planets is favoured over the one with one planet. Finally, the final fitting of the transit signal for the known TOI (TOI 5530.01) and the found candidate is available in figure 49.
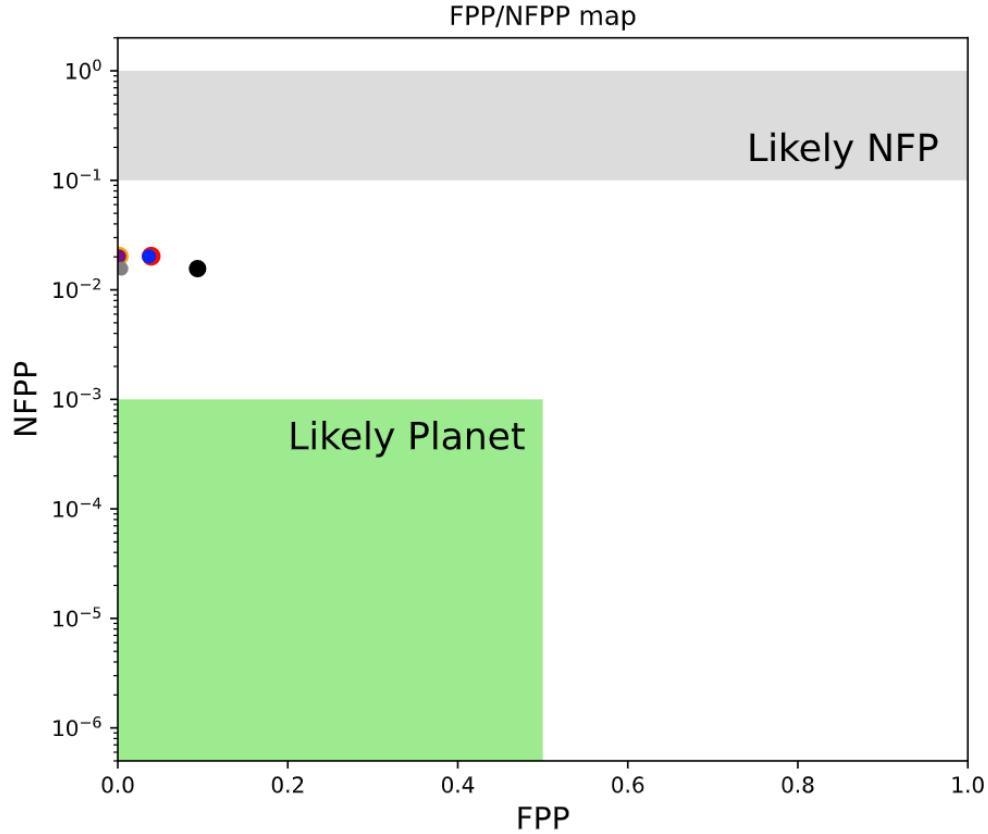
Figure 48: FPP/NFPP map. The red and orange dots are for the r band, respectively, for the FPP and FPP2 values. The blue and purple dots are for the b band, respectively, for the FPP and FPP2 values. Finally, the black and grey dots are, respectively, for the FPP and FPP2 values when no contrast curve is used.
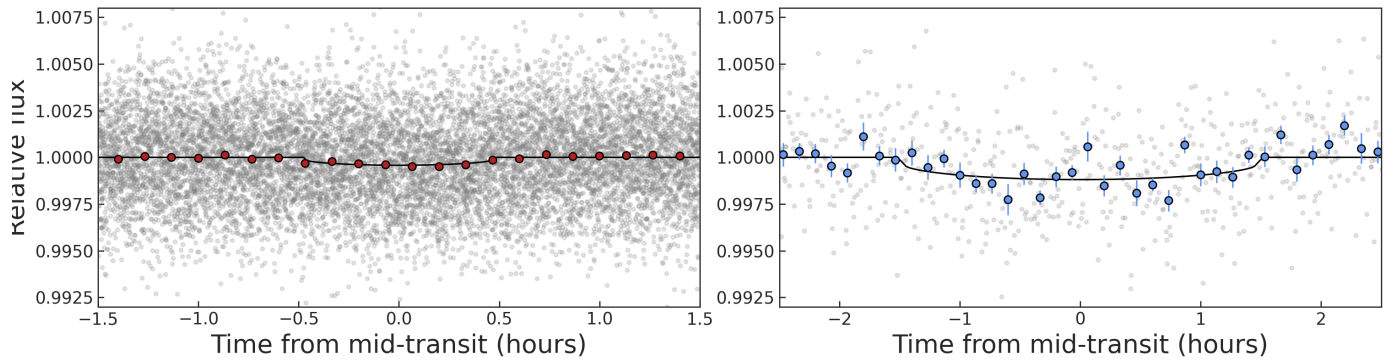

Figure 49: Fitting of the planetary transits (folded light curves), with the transit of TOI 5530.01 on the left and the found candidate on the right.

	Model 1: TOI 5530.01	Model 2: TOI 5530.01 | planet candidate	
$T0[\text{TBJD}]$	$1928.3612^{+0.003}_{-0.002}$	$1928.3612^{+0.003}_{-0.002}$	$1927.5293^{+0.006}_{-0.007}$
$P[days]$	$0.477^{+2.3e-6}_{-3.4e-6}$	$0.47743^{+2.3e-6}_{-3.3e-6}$	$16.20811^{+0.00025}_{-0.00034}$
$R_p[R_\oplus]$	$1.025^{+0.078}_{-0.078}$	$1.020^{+0.078}_{-0.078}$	$1.72^{+0.16}_{-0.16}$
a/R_*	$3.91^{+0.14}_{-0.14}$	$3.92^{+0.13}_{-0.13}$	$41.1^{+1.3}_{-1.4}$
$a[AU]$	$0.00955^{+0.00044}_{-0.00044}$	$0.00956^{+0.00043}_{-0.00043}$	$0.1003^{+0.0045}_{-0.0045}$
$i[deg]$	$88.3^{+1.0}_{-1.1}$	$88.2^{+1.1}_{-1.1}$	$89.29^{+0.33}_{-0.25}$
b	$0.113^{+0.074}_{-0.070}$	$0.120^{+0.075}_{-0.075}$	$0.51^{+0.17}_{-0.24}$
$T_{eq}[K]$	1272 ± 69	1272 ± 69	392 ± 21
$\delta[ppt]$	$0.418^{+0.063}_{-0.052}$	$0.418^{+0.057}_{-0.057}$	$1.07^{+0.17}_{-0.17}$
$\rho_*[g/cm^3]$	4.97 ± 0.52	4.9 ± 0.5	5.0 ± 0.5

Table 17: Planetary parameters of TOI 5530.01 with and without the planet candidate obtained through the fitting of the two models with `allesfitter`. Note that the epoch has been shifted to be in the middle of the data set to avoid correlations between epoch and period.

Bayesian evidence ($log(\mathcal{Z})$)	
Model 1	Model 2
184406.980 ± 0.146	184414.385 ± 0.204

Table 18: Bayesian evidence of the two models obtained with `allesfitter`.

From these results, one can conclude that it is likely that this signal has a planetary origin. Moreover, using the parameters derived from the second model (see equation 8), a host stellar density of 4.9±0.5 g/cm^3 is obtained with TOI 5530.01 while using the found candidate leads to a host stellar density of 5.0±0.5 g/cm^3. Therefore, this fact enhances the credibility of a planetary origin for the signal. However, to firmly confirm this hypothesis, additional observations are required but, as for the precedent candidate, the transit depth is too shallow ($\approx$ 1.07 ppt) to be observable with SPECULOOS. Hence, the same two solutions apply: (i) Wait for new results from the extended mission of TESS or (ii) Observe from ground-based observatory stars in the vicinity of the target star to see if they could be responsible for the signal.

4.1.3   TIC 328081248

This star has an effective temperature of 3521K, a radius of $0.343 R_\odot$ and one known planet candidate: TOI 5799.01. Some information about this TOI and the target star is available in table 19 and comes from the ExoFOP website[11].

RA (hh:mm:ss)	Dec (±hh:mm:ss)	Epoch (BJD)	Period (days)	Depth (mmag)	Duration (hrs)	Radius ($R_\oplus$)
20:06:31.05	15:59:17.13	2459772.3345	4.16454	2.839	1.572	1.988

Table 19: Stellar parameters associated with TIC 328081248 (the two first columns) and the planetary parameters associated with TOI 5799.01.

As for the two other targets presented above, the properties file was first run with the explore mode to fine-tune it for the transit search. As no high RMS value and no periodic variability in the folded curve were seen, the parameters `INITIAL_HIGH_RMS_MASK` and `INITIAL_SMOOTH_ENABLED` were disabled. As previously, the Savitzky-Golay filter was applied for the same reasons. As for this target, only one sector (54) is available, the maximum period to look for in the search was set to 16 days. Moreover, the 120-second exposure time is the only cadence available. After running the second module of SHERLOCK, one signal with a period of 14.02 days looked promising. Results from the report.log file are available in table 20. In this table, TOI 5799.01 is correctly found at the first run. Then, a signal with a period of 14.04 days is found with 9 different window sizes: all except the 2 shortest ones of 0.2 and 0.3 days. The border score and FAP values are excellent, increasing confidence in the signal. Figure 50 display the results obtained with the second run with the window size of 1.1 days, i.e. the signal with the highest SNR, where a clear transit shape is observed in the middle plot. At the top, the two transits have almost the same transit depth and at the bottom, SDE peaks are observed at the selected orbital period and its harmonics.

Given the results presented above, the vetting was run. First, the candidate parameters are listed in table 21. The results of the vetting will now be presented:

- Folded curves: The only cadence available is the fast one for which a shallow transit shape is visible.

---

[11]https://exofop.ipac.caltech.edu/tess/target.php?id=328081248

Run	P(days)	Number of findings	SNR	SDE	BS	FAP
1	4.165	11	22.93	24.57	1.0	8.0032e-05
2	14.02	9	15.58	10.17	1.0	8.0032e-05
3	0.729	3	6.59	5.52	1.0	8.0032e-05
4	8.422	1	8.55	6.85	1.0	0.012084834

Table 20: Results from the report file. The number of findings represents the number of times the signal was found out of the 11 window size possibilities. The green line corresponds to the run with the most promising signal (after the already known TOI).
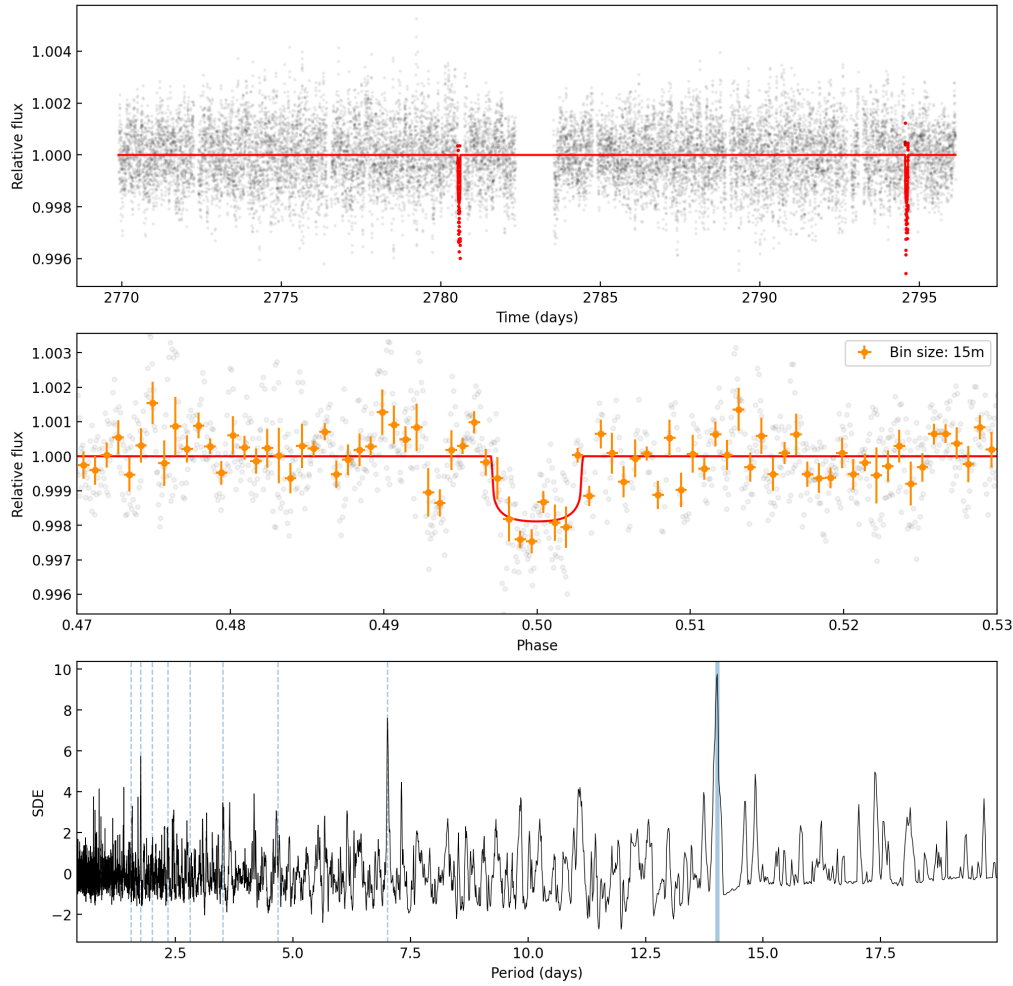

Figure 50: Results obtained with the second run for TIC 328081248, with a window size of 1.1 days and with sector 54. The signal found has a period of 14.02 days.

- Transit depth plot: As seen in figure 50, the two transit depths are almost exactly the same. Hence, there is no difference between the depth of even and odd transits.

- Folded light curves with the original period and first (sub)harmonics: In figure 52, the left middle plot exhibits a gap in the light curve, but it can be understood. Indeed, this plot corresponds to a folded curve over a period of 28 days (twice the period of the candidate). As there is only one sector available, this plot shows only the first detected transit of the candidate, and as the transits of TOI 5799.01 were masked after the first run, there are gaps in the light curve. At the bottom left, there is no transit pattern visible. In this case, it could be explained by the fact that the 2 transits are combined with 2 portions of flat light curves, resulting in a lower SNR. Hence, all the plots exhibit the expected behaviour if the period found is the correct one, except for the inferior conjunction of the sub-harmonic case that could be caused by some noise in the data.

- Transit source and target positions: This test gave the expected results, the transit source and the target positions are almost perfectly on top of each other, within the same pixel.

- The Target Pixel Files pixels folded curves: There is no transit pattern visible in any pixel outside the aperture, thus no reason to think that the signal is not coming from the target star. This plot is available in the appendix B in figure 67.

- Single-transits vetting sheets: As there are only two transits, the two vetting sheets associated with each transit must be analysed carefully. At the top left of figure 53, the flux for the first transit is shown. Unfortunately, data do not cover the entire transit, there is a gap in the data starting in the transit shape. As explained, it is caused by the mask of the transits of TOI 5799.01. Hence, it is difficult to be confident in this transit. For the second transit (shown in figure 54), every plot is as expected.

From the vetting, we can see that as the transits of the known TOI are masked in the light curve and as there is only one sector available, the first transit of the candidate is partially hidden by the mask. Nevertheless, the second transit might be a real one and the period found could be the correct one based on the (sub)harmonic case. Hence, the statistical validation was performed.

T0 (TBJD)	Period (days)	Duration (hr)	Depth (ppt)	Nbr of transits found
2780.5744	14.0249	1.92	1.538	2

Table 21: Parameters associated with the found candidate.
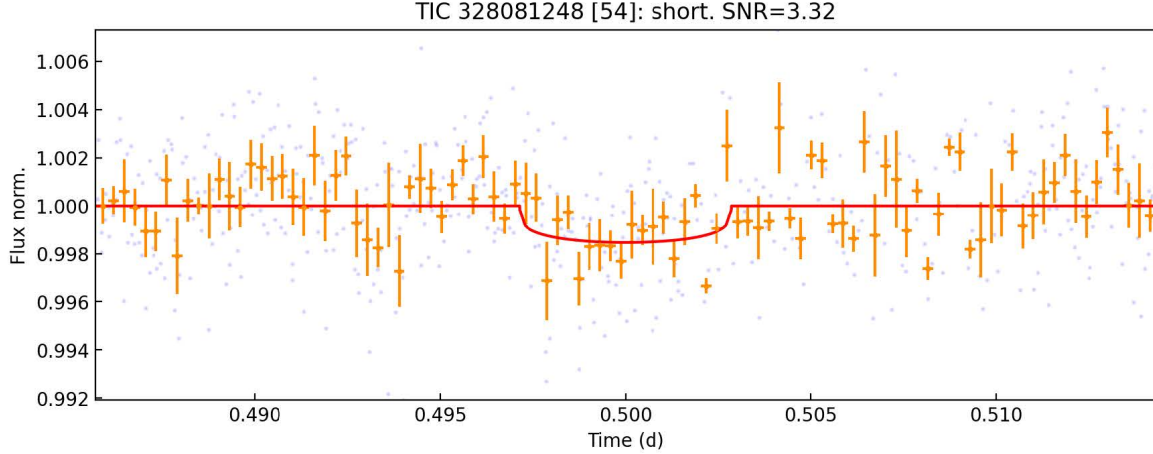

Figure 51: Folded curve of TIC 328081248 for the short cadence.

The results of the statistical validation will now be reviewed. First, data from high-resolution imaging using the ROBO-AO instrument (Lamman et al., 2020) were available for this target. Hence, a contrast curve in the r band with a wavelength of 879nm (taken from the ExoFOP website) was used to perform the statistical validation. This latter can be seen in figure 55. The folded curve obtained with the statistical validation report is shown in figure 56 where a clear transit shape is visible. In figure 57, the number of stars within a 10-pixel radius is very important. However, their magnitudes are much higher than the one of the target star. Moreover, on the right of this figure, we can see that the majority of the flux emanates from the aperture. Then, table 22 lists the most probable scenarios. The probability of the TP scenario largely increased when including the contrast curve. Hence, the probability that the transit pattern is due to a planet orbiting the target star with the period found is very high (94.8%). Now, figure 58 indicates in which category falls the target. Even if it is at the border of the likely planet zone, the NFPP value could be lowered by obtaining follow-up photometry at higher angular resolution than TESS from the ground. It could confirm that the signal is well on the target star and not on a nearby resolved star.

To conclude, it seems appropriate to continue with the fitting to compare 2 models: with
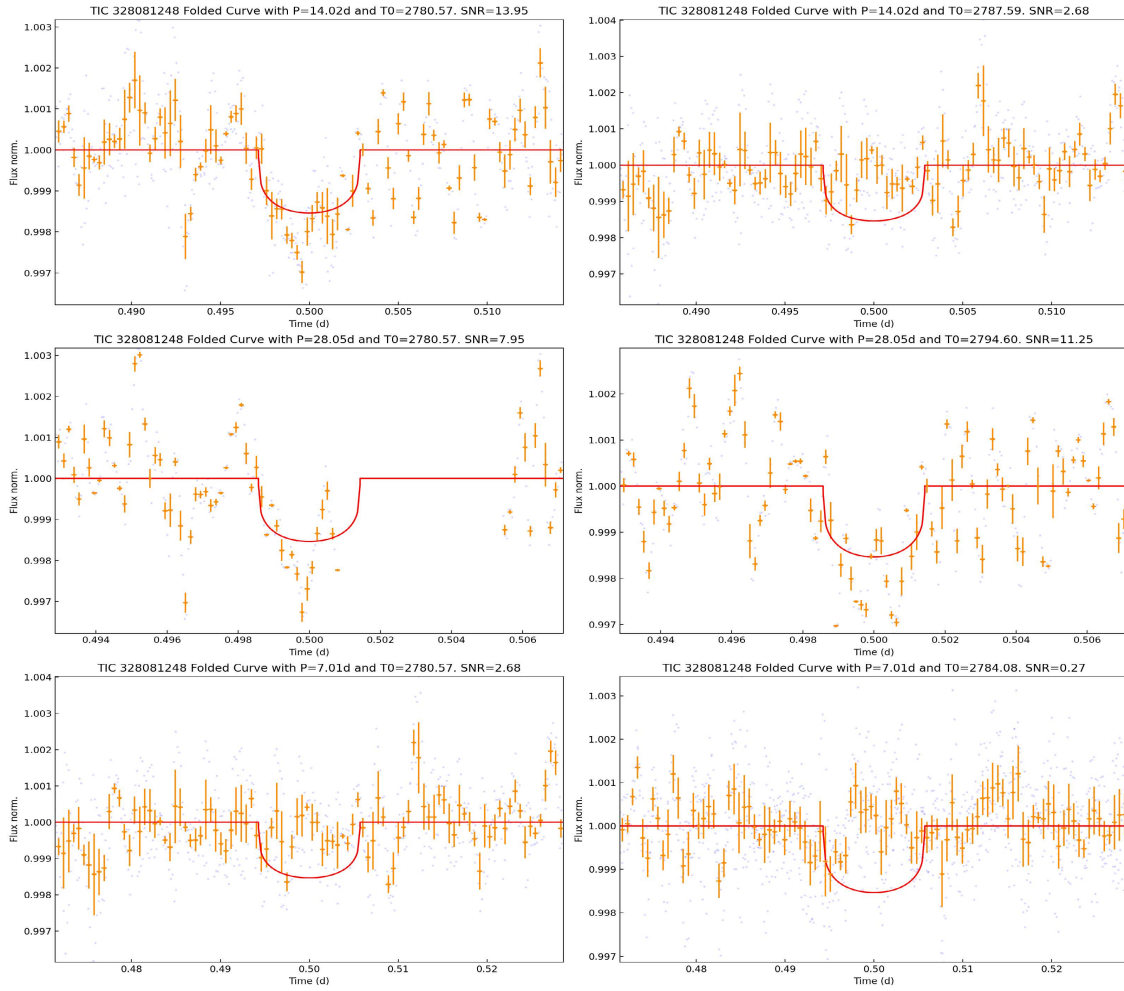
Figure 52: Folded curves for TIC 328081248. The left and right columns are, respectively, the inferior and superior conjunctions. The top, middle and bottom rows are, respectively, for the original, harmonic and sub-harmonic cases.

and without the found candidate. Hence, a Bayesian fit will be performed to compare the Bayesian evidence of the two models.

With contrast curve	Without contrast curve
TP: 0.948	TP: 0.784
DTP: 0.038	PTP: 0.098

Table 22: Most probable scenarios for the candidate of TIC 328081248, with and without the contrast curve.
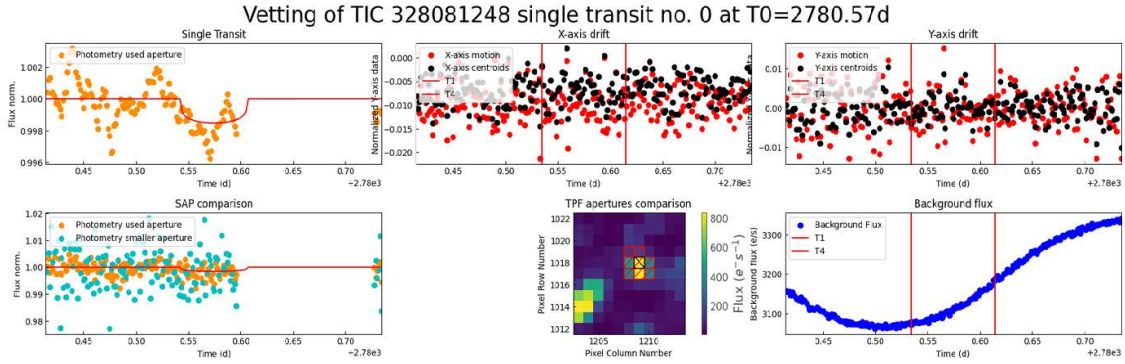
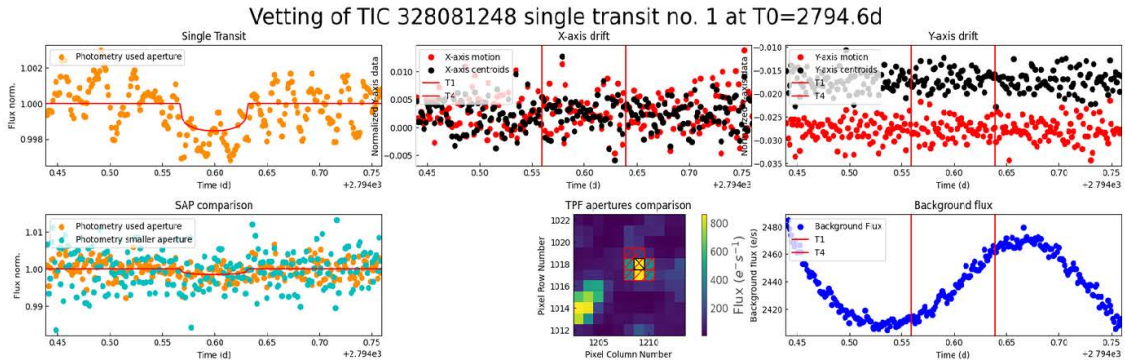Figure 53: Vetting sheet of the first transit of the candidate of TIC 328081248.

Figure 54: Vetting sheet of the second transit of the candidate of TIC 328081248.

	FPP	NFPP	FPP2
With contrast curve	0.003591	0.001674	0.000144
Without contrast curve	0.070899	0.001429	0.003043

Table 23: Summary of the mean statistical values obtained with and without the contrast curve.
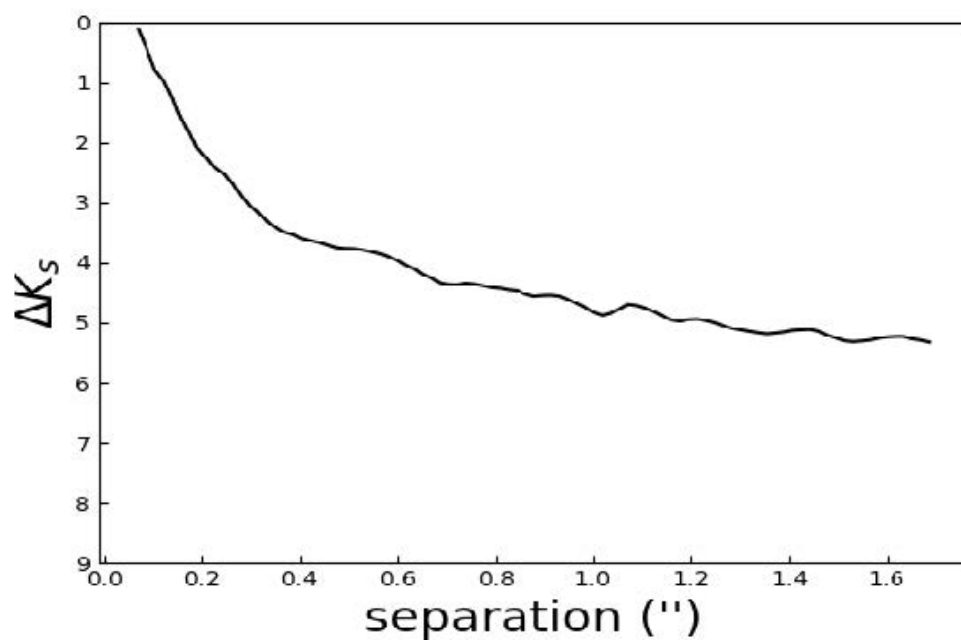
Figure 55: Contrast curve in the r band (879nm) obtained for TIC 32081248 with the ROBO-AO instrument, retrieved on the ExoFOP website.
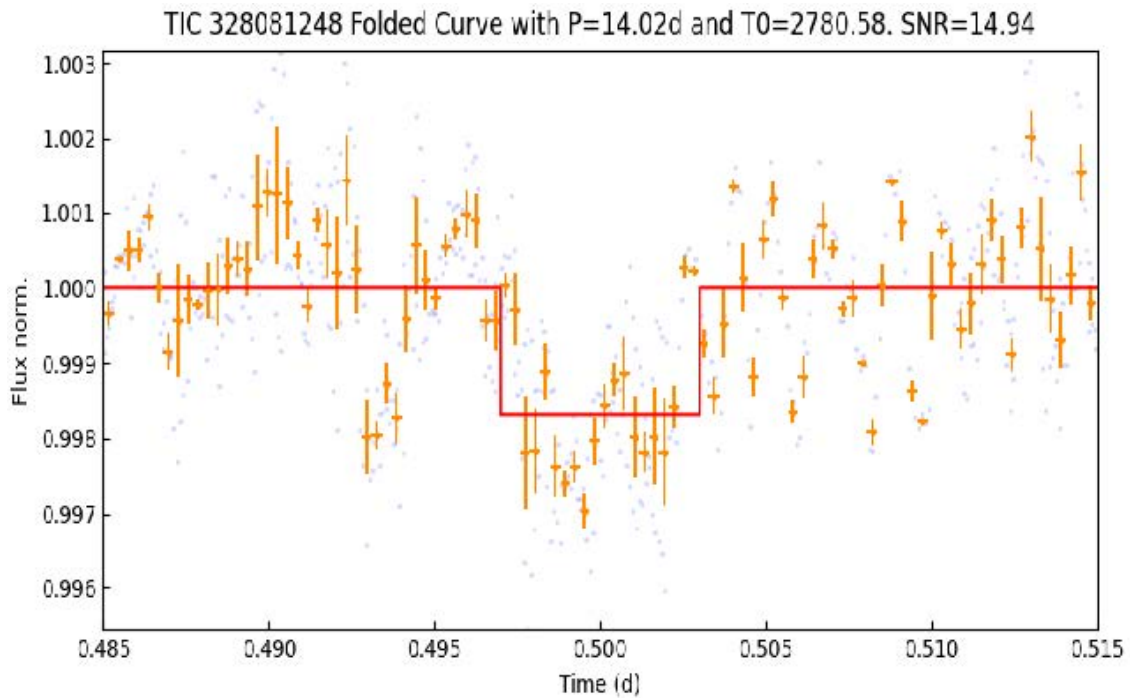
Figure 56: Folded curve obtained with the statistical validation report for the candidate of TIC 328081248.
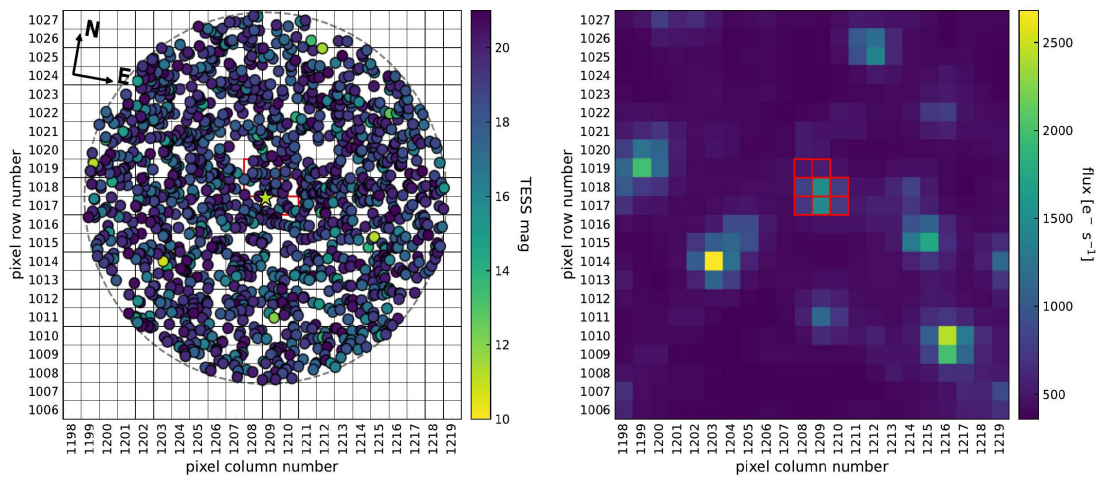

Figure 57: Nearby stars surrounding TIC 328081248 on the left. Corresponding flux and photometric aperture on the right.
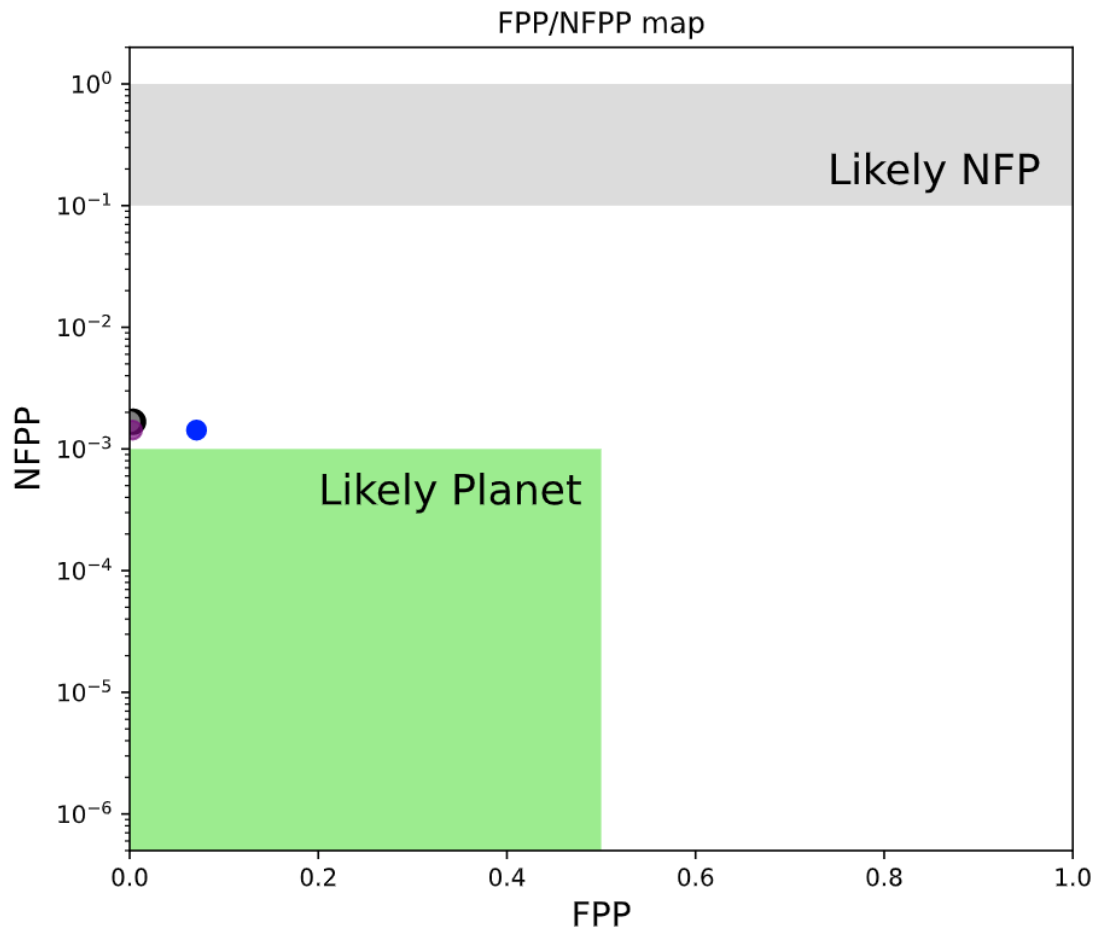
Figure 58: FPP/NFPP map for TIC 328081248. The black and grey dots are the values obtained with the contrast curve with, respectively, the FPP and FPP2 values. The blue and purple dots are the values obtained without the contrast curve with, respectively, the FPP and FPP2 values.

For the fitting, we used the same methodology as for the two previous targets. Firstly, the noise fitting was performed to evaluate the hyperparameters and the error on the relative flux. However, as the two hyperparameters of the GP did not converge, it seemed more appropriate to restart the noise fitting with a more simple model and then, increase the complexity to compare the Bayesian evidence (as explained in section 3.5). In this case, the hybrid spline model gave the best value for the logarithm of the Bayesian evidence so we used this model for the noise and we constrain the parameters of the error on the relative flux. Secondly, we performed the fitting with two models: one with TOI 5799.01 alone and one with TOI 5799.01 and the found planet candidate with a period of 14.016 days. The results of the derived planetary parameters for the two models are listed in table 24. The fitting model results obtained for this target can be seen in figure 59. Finally, the corner plot of the fitted and the derived parameters is available for the interested reader in the appendix B, respectively, in figures 73 and 74. In table 25, one can see that the Bayesian evidence for the second model is greater than 1804 units compared with the first model. Hence, from this analysis, one can conclude that the signal found in this work is likely due to a planetary transit.

Moreover, the host stellar density is in agreement for the two planets for the second model. Indeed, with the parameters of the planets (see equation 8), the host stellar density obtained with TOI 5799.01 is $11.0 \pm 1.1 \ g/cm^3$ and for the found candidate, it is $11.3 \pm 1.2 \ g/cm^3$ Hence, it is a strong hint that the signal has a planetary origin.

However, as for the two precedent promising signals, new observations are required to confirm the planetary origin of this candidate. For this signal, the transit depth is estimated to be ≈ 2.17 ppt, hence observations could be performed with the SPECULOOS telescopes but due to the large period (≈ 14 days), it is difficult to plan observations for this target. Nevertheless, this could be planned in the near future.

	Model 1: TOI 5799.01	Model 2: TOI 5799.01	planet candidate
$T0[TBJD]$	$2784.828^{+0.001}_{-0.001}$	$2784.828^{+0.001}_{-0.001}$	$2794^{+0.006}_{-0.005}$
$P[days]$	$4.164^{+0.000637}_{-0.000631}$	$4.164^{+0.00052}_{-0.00056}$	$14.016^{+0.0075}_{-0.0058}$
$R_p[R_\oplus]$	$1.613^{+0.059}_{-0.066}$	$1.716^{+0.070}_{-0.079}$	$1.58^{+0.13}_{-0.12}$
a/R_*	$21.49^{+0.73}_{-0.73}$	$21.59^{+0.71}_{-0.71}$	$48.9^{+1.7}_{-1.8}$
$a[AU]$	$0.0343^{+0.0016}_{-0.0016}$	$0.0344^{+0.0016}_{-0.0016}$	$0.0780^{+0.0037}_{-0.0037}$
$i[deg]$	$89.62^{+0.26}_{-0.37}$	$89.47^{+0.33}_{-0.40}$	$89.50^{+0.27}_{-0.21}$
b	$0.144^{+0.13}_{-0.099}$	$0.2^{+0.15}_{-0.12}$	$0.42^{+0.18}_{-0.23}$
$T_{eq}[K]$	491 ± 30	491 ± 29	326 ± 20
$\delta[ppt]$	$2.61^{+0.26}_{-0.26}$	$2.70^{+0.21}_{-0.20}$	$2.17^{+0.25}_{-0.25}$
$\rho_*[g/cm^3]$	10.8 ± 1.1	11.0 ± 1.1	11.3 ± 1.2

Table 24: Planetary parameters of TOI 5799.01 with and without the planet candidate obtained through the fitting of the two models with `allesfitter`. Note that the epoch has been shifted to be in the middle of the data set to avoid correlations between epoch and period

Bayesian evidence ($log(\mathcal{Z})$)	
Model 1	Model 2
8950.523 ± 0.100	10754.864 ± 0.237

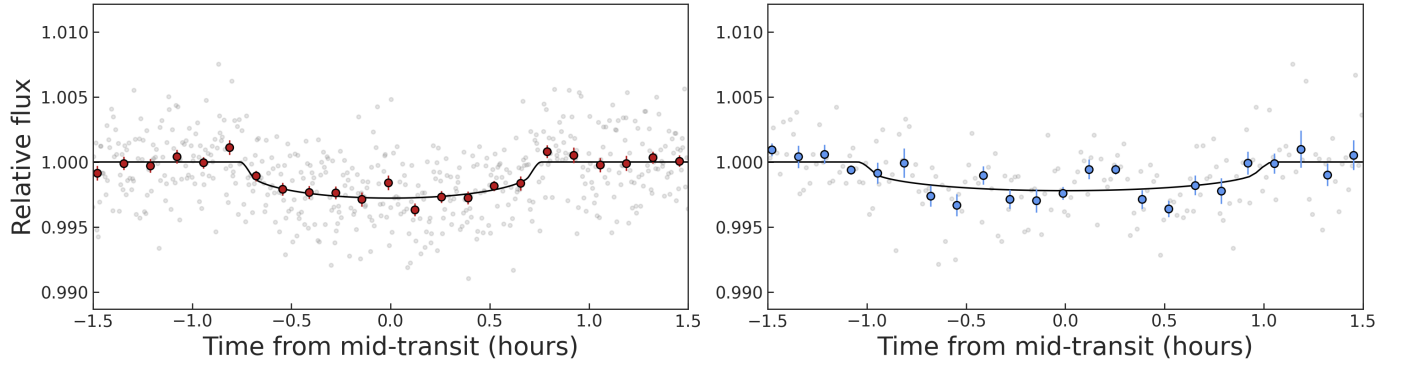Table 25: Bayesian evidence of the two models obtained with `allesfitter`.

Figure 59: Fitting of the planetary transits (folded light curves), with the transit of TOI 5799.01 at the left and the found candidate at the right.

4.2 Eclipsing binaries

Of the 56 targets studied, one showed a typical example of a signal caused by a nearby eclipsing binary. It was for the target star TIC 4918918 which has one known planet candidate: TOI 5747.01, with an orbital period of 0.573 days. This target was observed with sectors 21 and 48, so the mode 'global' was used to combine the only two sectors available. As no high RMS and no large periodic stellar variability were observed, the properties file was adapted to disable the associated parameters while keeping the Savitzky-Golay filter. It was set to have a maximum of 4 runs with a maximum period of 20 days. The result of the first run showed a signal with a period of 3.897 days and a "V" shape, typical of eclipsing binary systems. Indeed, as the size of the two stars is comparable, the geometry of the eclipse is thus grazing, resulting in a V-shaped signal. The shape of the signal, along with its SDE value and the transit locations, can be seen in figure 60. At the bottom of this plot, one can see that the SDE value is particularly high at the selected period and its harmonics. In the report file, the SDE and SNR were respectively given as 288.3469 and 28.9938. Moreover, the FAP and the border score were 8.0032e-05 and 1.0, respectively. Those results clearly indicate that the signal has an astrophysical origin, not an instrumental one. Even though an eclipsing binary was directly suspected (given the shape of the transit), the vetting module was run to firmly confirm this hypothesis. From the vetting report, parameters associated with the target star and the candidate were established and are displayed in table 26.

RA(hh:mm:ss)	Dec(±hh:mm:ss)	T0(TBJD)	Period(days)	Duration(hrs)	Depth(ppt)
9:53:30.91	35:34:16.72	1872.2991	3.8978	1.27	4.758

Table 26: The first two parameters are associated with the target star and the last four are associated with the planet candidate. This is for TIC 4918918.
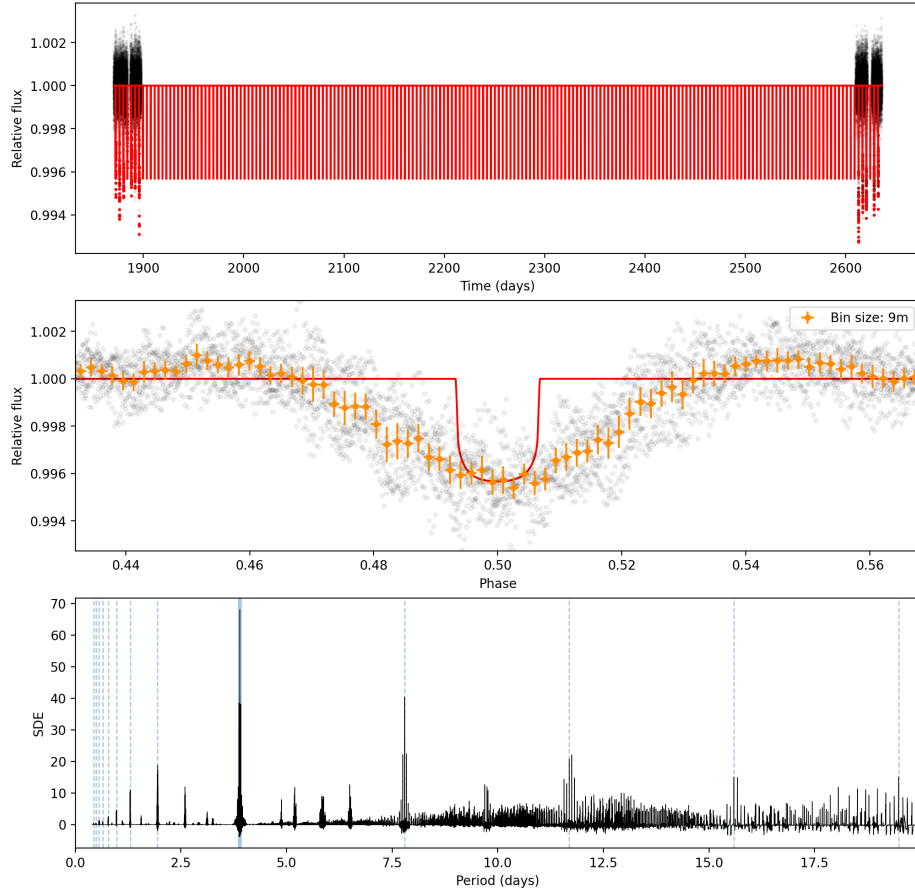
Figure 60: Result of the first run for TIC 4918918 with the window size of 0.6 days.

Then, there is the single transit depths plot in figure 61. In this plot, one can see that the even transits seem to be a bit deeper on average compared to the odd transits (by about 0.6 ppt). This can be observed for an eclipsing binary, as explained in section 3.3.2. Another possibility is that two particular events (one for the red and one for the blue) are responsible for this difference and are in fact outliers. In this case, it could be explained by the fact that if the two stars involved in the system have the same magnitude and the geometry is such that they are eclipsing roughly the same area of each other, the transit depth is expected to be more or less the same each time. Therefore, no conclusion can be drawn based on that plot alone.
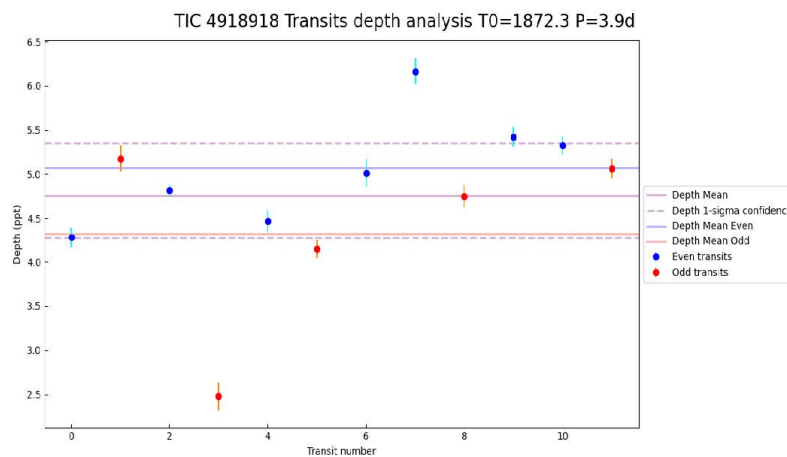

Figure 61: Single transit depths for the candidate with a period of 3.89 days for TIC 4918918.

Then, figures 62 and 63 can be analysed along with figure 64. Indeed, in the first two figures, one can clearly see that the signal source and target positions are well separated. The signal originates from a source that is at least 3 pixels away from the target. Furthermore, in figure 63, one can observe that several pixels outside the aperture (in purple) feature an eclipse pattern. In addition, in figure 64, the field of view around the target is shown. On this plot, the star responsible for the eclipse pattern observed in figure 63 is clearly visible and is located by the red dot numbered 7.
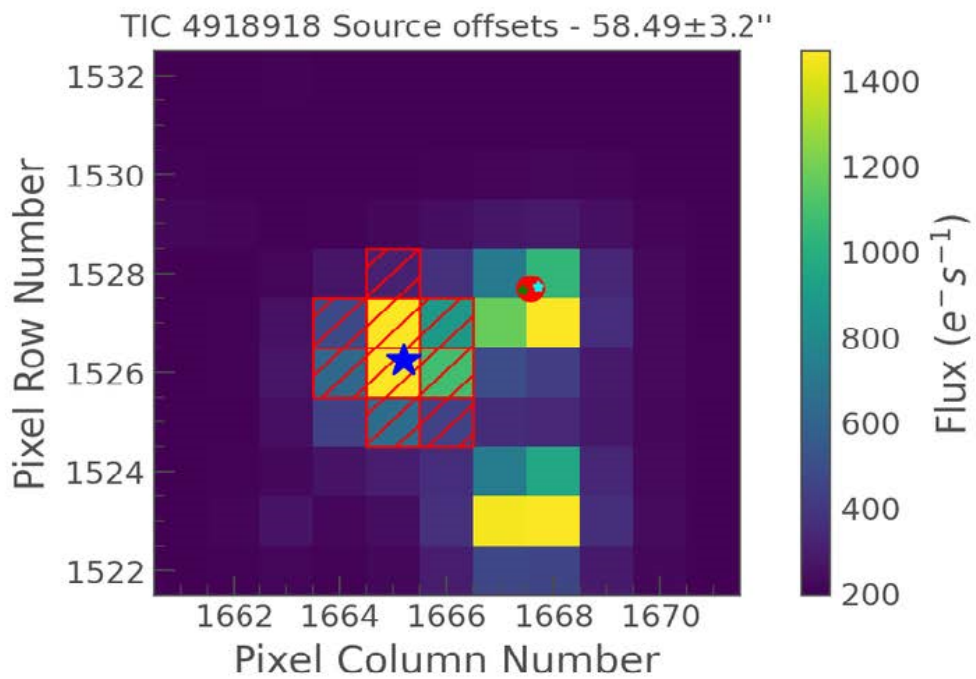
Figure 62: The signal coming from the source (red point) and the target (TIC 4918918, blue star) positions are 3 pixels apart.
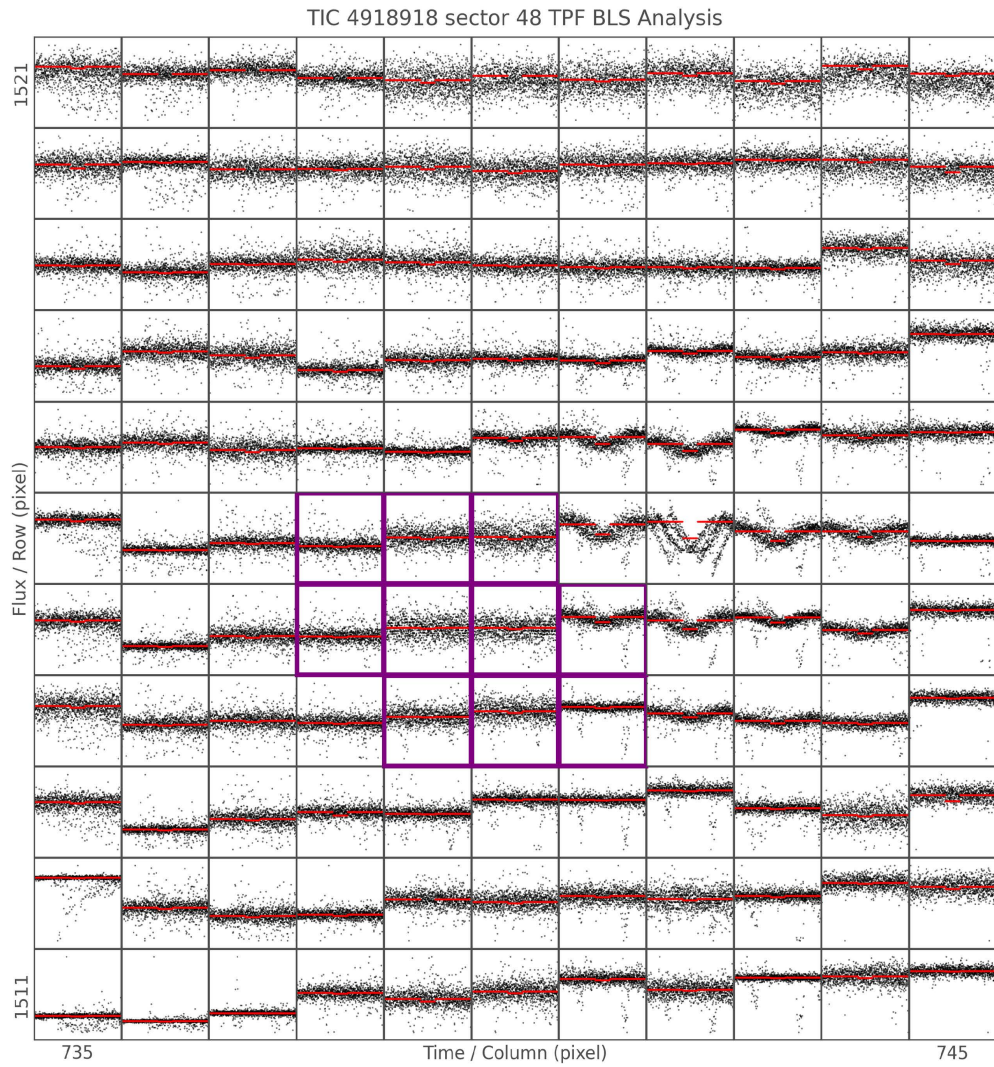
Figure 63: Target pixel file for TIC 4918918, sector 48. A clear transit pattern is visible at the top right of the aperture.
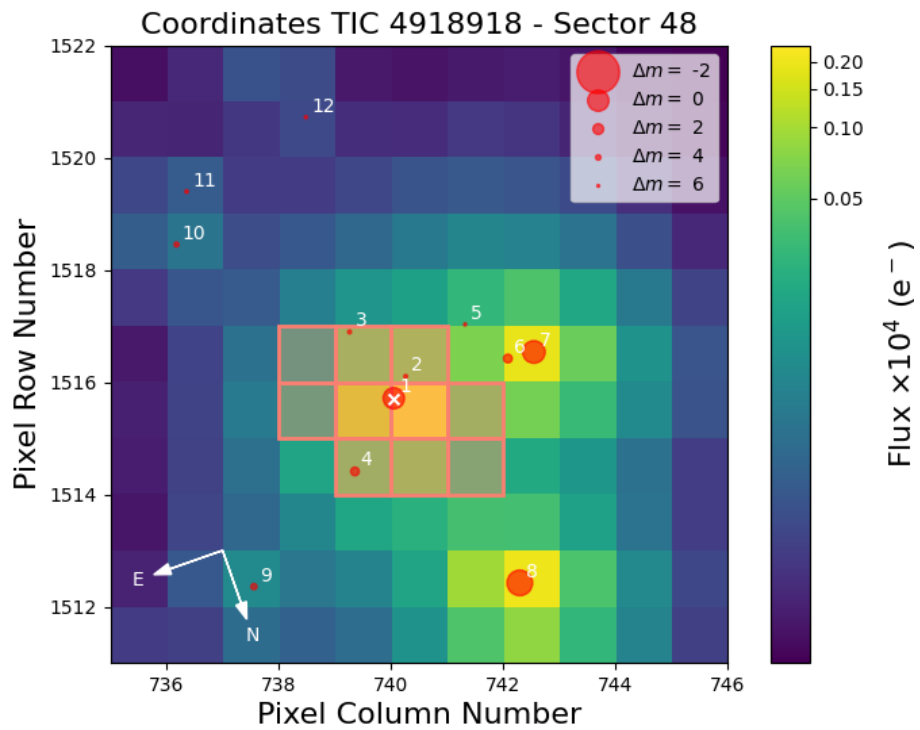
Figure 64: Field of view of TIC 4918918. The star that is probably responsible for the eclipse signal is the red dot numbered 7.

Results from the vetting report clearly show strong evidence that the candidate is due to an eclipsing binary system, and the statistical validation report confirmed that hypothesis. Table 15 lists the 3 scenarios with the highest probabilities found in the report. Those are the NEBx2P, BEBx2P and NEB scenarios. The probability of any other scenario is zero. The signal was classified as an NFP, with mean values of the FPP and the NFPP of 1.0 and 0.606464, respectively. Therefore, the observed signal is a typical example of what is expected with nearby eclipsing binaries. Finally, as TOI 5747.01 was only recovered with the 4th run, a 5th run was added in the properties file to search for a new candidate, but it did not give any conclusive results. Note that results for the 2nd and 3rd runs were inconclusive as the masking of the signal found in the 1st run did not cover the entire V shape, resulting in several detections of the same event.

TIC 4918918		
NEBx2P	BEBx2P	NEB
0.625	0.362	0.012

Table 27: Results for 3 scenarios with the highest probabilities for TIC 4918918.

4.3 Unpromising signals

The majority of the target list did not give any promising signal after running the 2nd module of SHERLOCK. There can be several reasons why those targets were directly discarded and they were all viewed in detail at the end of section 3.2.2. Note that it is a combination of several of the reasons presented in this section that leads to the exclusion of a candidate, not just one of them.

A typical example will now be presented to illustrate some of the reasons set out above. For this, we will focus on TIC 153951307 which was run with the 120-second exposure time. As this target was observed with very separated sectors, to decrease the computational time, two sets of sectors were used (see table 28). By combining these sets, one can see that 2 signals are found twice in the first 2 runs. In fact, those are two already known TOIs. The first one is TOI 1238.01 with a period of 3.29 days, and the second one is TOI 1238.02 with a period of 0.76 days. Then, we can look at the next runs. The third runs of both sets give, in both cases, a signal that is only found with 5 different window sizes. Moreover, the period (and thus the signal) is not the same in the two sets, which is not a good sign.

Set of sectors	Run	P(days)	Number of findings	SNR	SDE	BS	FAP
14,15,21,22	1	3.29	11	45,64	60.58	1.0	8.0032e-05
	2	0.76	10	2.16	44.86	0.9663	8.0032e-05
	3	11.58	5	9.50	185.43	1.0	8.0032e-05
	4	0.43	1	5.42	75.46	1.0	8.0032e-05
41,49	1	3.29	11	30.16	36.56	1.0	8.0032e-05
	2	0.76	10	10.51	12.28	1.0	8.0032e-05
	3	18.52	5	12.43	35.94	1.0	8.0032e-05
	4	18.51	2	3.48	120.89	1.0	8.0032e-05

Table 28: Summary table of the reports found after running the 2nd module of SHERLOCK for TIC 153951307 (120-second exposure time). The two already known TOIs are found with the first two runs for the two sets of sectors and are represented by the green lines.

If we now analyse figure 65, corresponding to the signal found with a period of 11.58 days and for the first set of sectors, one can see that there is a shallow decrease in flux with no super convincing transit shape. Moreover, the transit depths do not look consistent from one transit to another. In figure 66, corresponding to the signal found at 18.52 days for the second set of sectors, the transit is clearly visible. Unfortunately, as it was not detected with

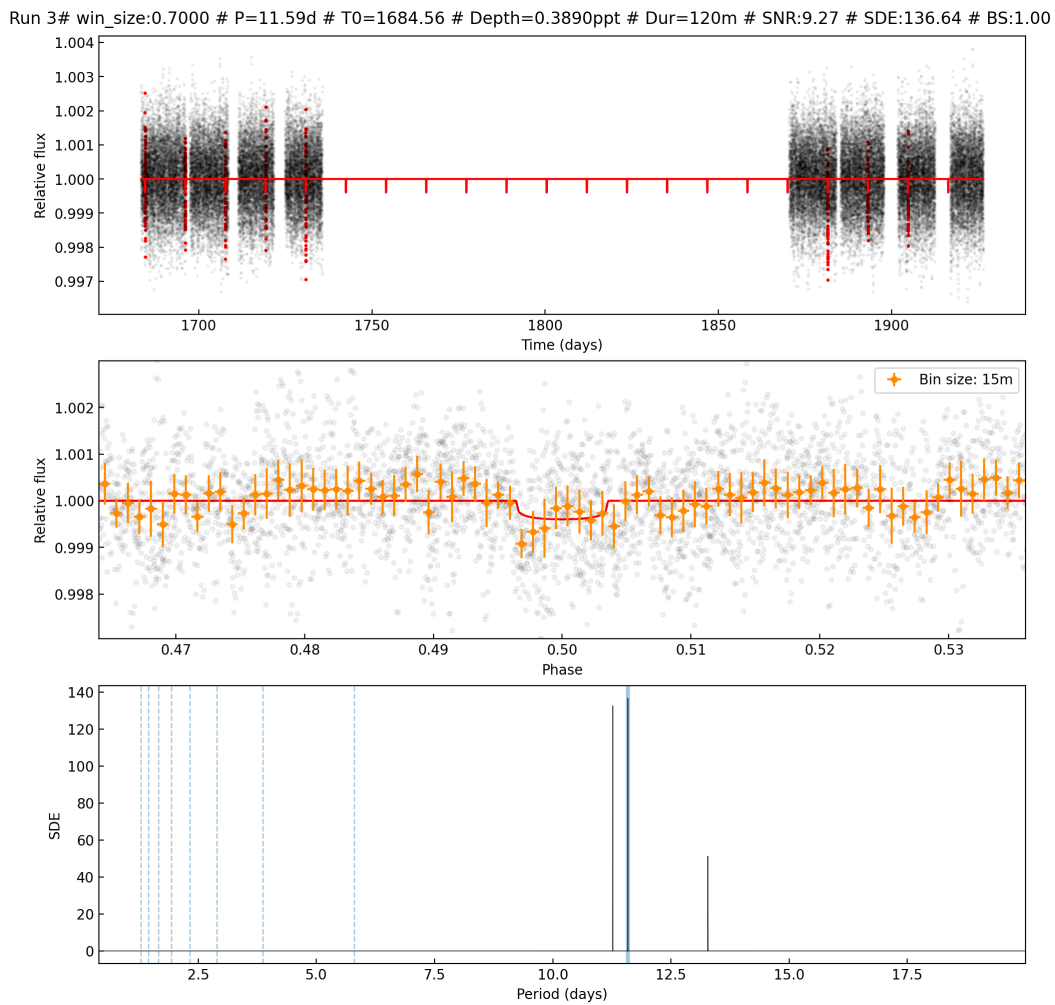the first set of sectors, this is not a good indicator.

Figure 65: Result of the 3rd run for TIC 153951307 with sectors 14, 15, 21 and 22 with the 120-second exposure time. The signal found has a period of 11.59 days with a shallow transit.

As the 600-second exposure time was available with sectors 41 and 49, we decided to run SHERLOCK on this cadence as well to check whether it could retrieve the 18.5d signal. Unfortunately, it did not, as shown in table 29. To conclude, as the signals found were not retrieved by the different sets of sectors and exposure times, they did not look promising enough to proceed to the next module and the analysis was stopped here for this target.
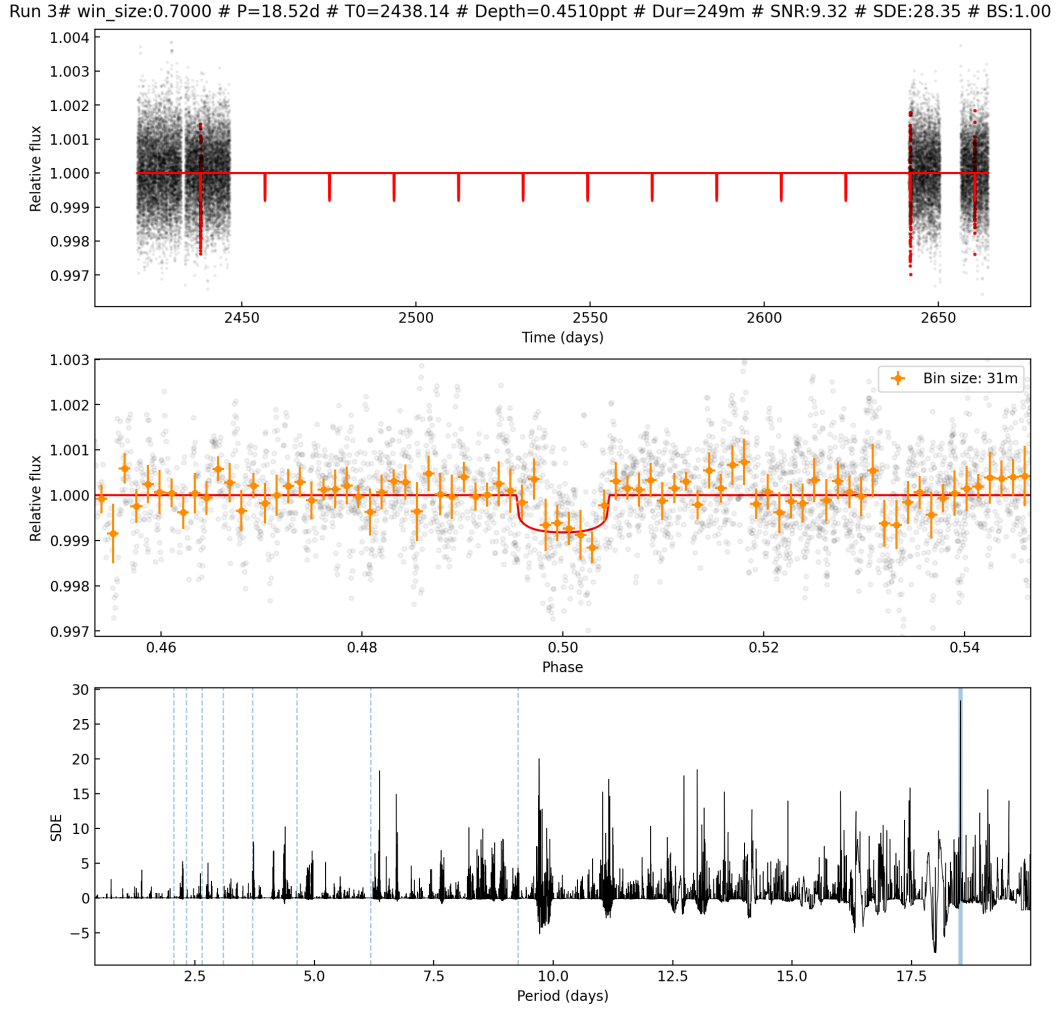
Figure 66: Result of the 3rd run for TIC 153951307 with sectors 41 and 49 with the 120-second exposure time. The signal found has a period of 18.52 days and a transit is visible.

Run	P(days)	Number of findings	SNR	SDE	BS	FAP
1	3.29	11	13.74	36.56	1.0	8.0032e-05
2	0.76	11	5.60	19.95	1.0	8.0032e-05
3	14.14	1	1.04	48.20	1.0	8.0032e-05

Table 29: Summary table of the reports found after running the 2nd module for TIC 153951307, using sectors 41 and 49 with an exposure time of 600 seconds. Again, the two already known TOIs are found with the first two runs and are represented by the green lines.

5   Conclusions

In this work, we have successfully used the open-source tool SHERLOCK with TESS data to detect promising new multi-planetary systems around M dwarfs. Among the 56 targets studied, we found 3 promising candidates for which we performed an in-depth investigation:

1. For TIC 321669174: we found one candidate with a period of 5.109 days and a planetary size of $1.227R_\oplus$. Those results were corroborated with updated results on the ExoFOP website at the end of this master's thesis.

2. For TIC 244170332: A promising candidate with a period of 16.2 days and a planetary size of $1.72R_\oplus$ was found. The results of the fitting process showed evidence that this finding is likely a real planet.

3. For TIC 328081248: A promising candidate of period 14.02 days and a planetary size of $1.58R_\oplus$ was found. Again, the fitting process indicates that this candidate is likely a real planet.

For those three targets, we recovered all the known TOIs, indicating the efficiency of the methodology used throughout this work. In addition, these results highlight the importance of conducting an independent search for planetary candidates using alternative pipelines such as SHERLOCK to exploit the full potential of the data. Indeed, by solely focusing on the alerted TOIs, numerous potential candidates within the data would go unnoticed by the scientific community.

For the three promising signals, new observations of the target stars are required to firmly confirm their planetary origin. Observations of nearby stars in the vicinity of those target stars would also be useful to be sure that they are not at the origin of the signals found in this work. While further confirmation is needed, these findings highlight the ongoing progress in exoplanet research. Studying exoplanets offers valuable insights into planetary systems, their formation, and the potential for habitability beyond our solar system. As technology improves, we can expect more exciting discoveries and a deeper understanding of the diversity of exoplanets in the universe, notably with the recent observations made with the JWST. Finally, in the forthcoming years, we can also expect a deeper knowledge and understanding of the atmosphere of those exoplanets and maybe find a signature of life on another planet.

A   Summary target list

TIC ID	Final module	Exp time(s) used	Sector(s) used	Results
434116397	2	120 and 20	56	Potential candidate with the 2nd run but not found with the 20 seconds exposure time.
354173595	2	120 and 600	42 and 56	Potential candidate with the 3rd run but not found with the 600 seconds exposure time.
151728428	1	/	/	Light curves not available with SPOC
4918918	4	120	21 and 48	Signal found due to a nearby eclipsing binary system
270355392	2	120	4,31,42 and 43	No promising signal found with the 2nd module
377097547	2	120	45 and 46	No promising signal found with the 2nd module
407591297	3	120	21 and 48	Candidate of period 3.26 days discarded after the vetting
318022259	2	120	20 and 47	No promising signal found with the 2nd module
224298134	2	120	15, 21, 22, 41 and 48	Harmonic of TOI 2079.02 (period of 9.31 days), found with the 2nd run. The candidate had a period of 18.63 days, i.e. twice the known TOI's period.
150428135	1	120	1, each sector from 3 to 11, 13, 27, 28, 30, 31, 33, 34, 36, 37 and 38	Discarded because the running time was too long, i.e. more than a week

Table 30: List of all the targets analysed with SHERLOCK (part 1)

TIC ID	Final module	Exp time(s) used	Sector(s) used	Results
307809773	2	120	43, 44 and 45	No promising signal found with the 2nd module
381897917	4	120, 600 and 1800	13 and 39	Potential candidate found with the 2nd run and discarded with the 3rd and 4th modules
271169413	4	120	11 and 38	Candidate with a period of 0.44 days found but discarded at step 4
380887434	2	120	42 and 43	No promising signal found with the 2nd module
384888319	2	120	42 and 43	No promising signal found with the 2nd module
239332587	2	120 and 1800	15, 41 and 55	No promising signal found with the 2nd module
176797879	2	120	4 and 31	No promising signal found with the 2nd module
408416746	2	120	8 and 35	No promising signal found with the 2nd module
251086776	3	120	4 and 31	Candidate of period 19.4 days discarded after the vetting
77156829	2	120	4, 5, 31 and 32	No promising signal found with the 2nd module
465540915	2	120	10 and 11	No promising signal found with the 2nd module

Table 31: List of all the targets analysed with SHERLOCK (part 2)

TIC ID	Final module	Exp time(s) used	Sector(s) used	Results
463070290	2	120	1 and 9	No promising signal found with the 2nd module
146776480	2	120	8, 34 and 35	No promising signal found with the 2nd module
236387002	2	120	24, 25, 51 and 52	No promising signal found with the 2nd module
166184428	2	120	11 and 38	No promising signal found with the 2nd module
377293776	2	120	Each sector from 14 to 26 expect sector 18, 40, 41, 47 and each sector from 49 to 60	No promising signal found with the 2nd module
467179528	2	120	14, 15, 21 and 41	No promising signal found with the 2nd module
396562848	2	120	18, 19, 25, 26, 52, 53 and 59	No promising signal found with the 2nd module
235678745	2	120	Each sector from 49 to 52	No promising signal found with the 2nd module
286813592	2	120	22, 45, 46 and 49	No promising signal found with the 2nd module
138820713	3	120	22 and 49	Signal with a period of 5.03 days discarded at step 3
328081248	5	120	54	Potential planetary candidate found with a period of 4.16 days (see section 4.1.3)
417931300	4	120	14, 15, 21, 22, 41 and 48	No promising signal found with the 2nd module
219698776	2	120	14, 20, 40 and 47	No promising signal found with the 2nd module

Table 32: List of all the targets analysed with SHERLOCK (part 3)

TIC ID	Final module	Exp time(s) used	Sector(s) used	Results
153951307	2	120 and 600	14, 15, 21, 22, 42 and 49	No promising signal found with the 2nd module
22233480	2	120	40, 52 and 53	No promising signal found with the 2nd module
415969908	2	120	29 and 42	No promising signal found with the 2nd module
468983280	2	120	7, 44, 45 and 46	No promising signal found with the 2nd module
34068865	2	120	9, 35, 36 and 63	No promising signal found with the 2nd module
359271092	2	120	9, 10, 36, 37 and 62	No promising signal found with the 2nd module
413248763	2	120	8, 35 and 62	No promising signal found with the 2nd module
321669174	5	120, 600 and 1800	See section 4.1.1	Promising signal found with a period of 5.109 days
336961891	2	120	42 and 43	No promising signal found with the 2nd module
154940895	2	120	19, 20, 25, 26, 40, 47, 52, 53 and 59	No promising signal found with the 2nd module
298428237	2	120	14, 15, 41 and 55	No promising signal found with the 2nd module
28900646	2	120	19 and 59	No promising signal found with the 2nd module
44313455	2	120	26 and 53	No promising signal found with the 2nd module

Table 33: List of all the targets analysed with SHERLOCK (part 4)

TIC ID	Final module	Exp time(s) used	Sector(s) used	Results
219195044	2	120	4, 5, 6, 7, 8, 11, 12, 28, 31, 32, 33, 34 and 38	No promising signal found with the 2nd module
441804533	2	120	16, 17, 22 and each sector from 47 to 60	No promising signal found with the 2nd module
198385543	2	120	17, 20, 23, 24, 25, 26, 40, 47 and each sector from 50 to 60	No promising signal found with the 2nd module
279922257	2	120	18, 19, 25, 26, 52 and 53	No promising signal found with the 2nd module
120045750	2	120	41, 53 and 54	No promising signal found with the 2nd module
163539739	2	120	15, 55 and 56	No promising signal found with the 2nd module
147923561	2	120	40, 41, 47 and 48	No promising signal found with the 2nd module
244170332	5	120 and 600	3, 30 and 42	Promising candidate found with a period of 16.20 days (see section 4.1.2)
470987100	2	120	20 and 47	No promising signal found with the 2nd module

Table 34: List of all the targets analysed with SHERLOCK (part 5)
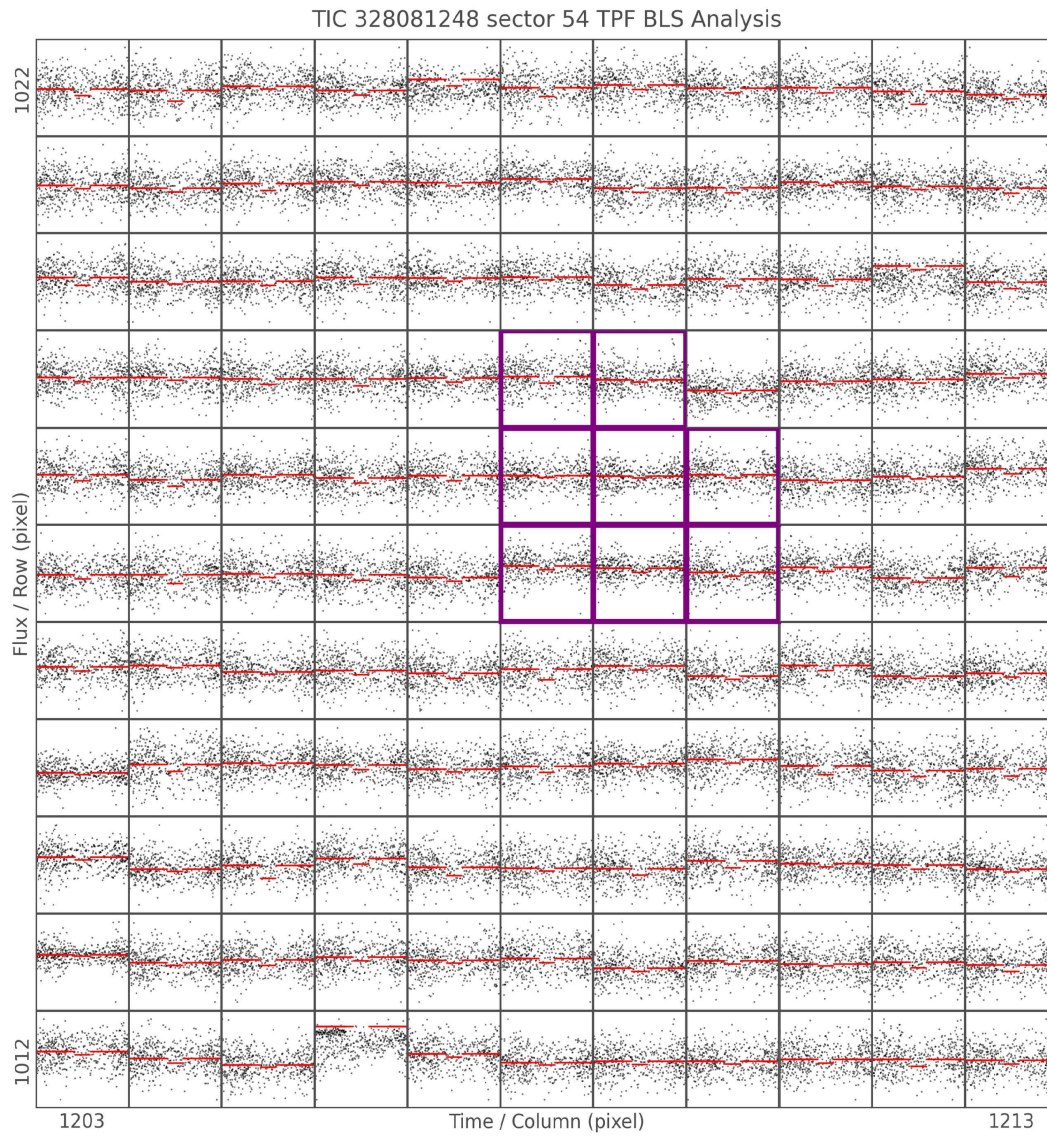
B   Additional plots

Figure 67: Target Pixel File of TIC 3208081248 and sector 54. No transit shape is apparent outside the aperture.
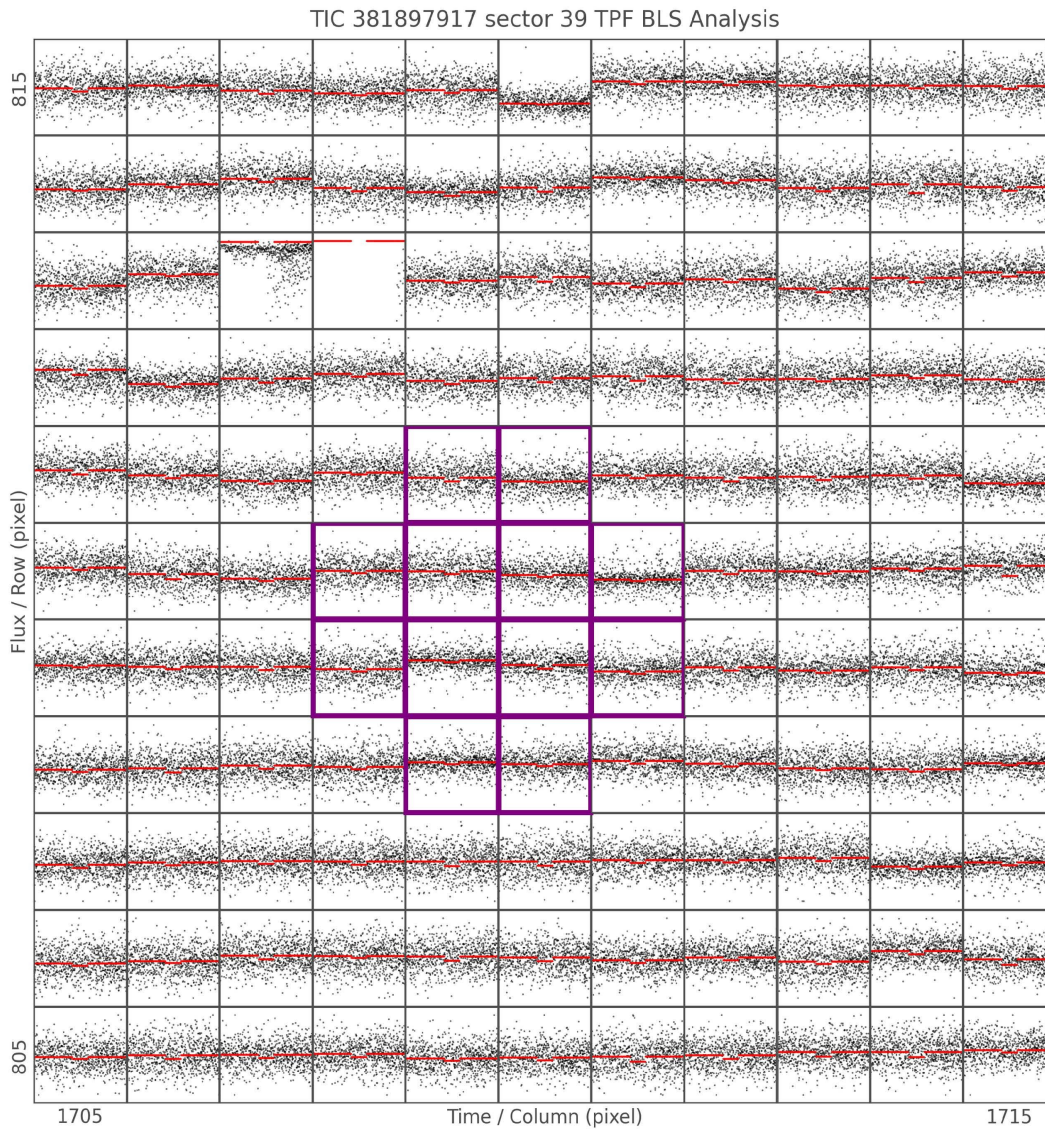
Figure 68: Target Pixel File of TIC 381897917 and sector 39. No transit shape is apparent outside the aperture.
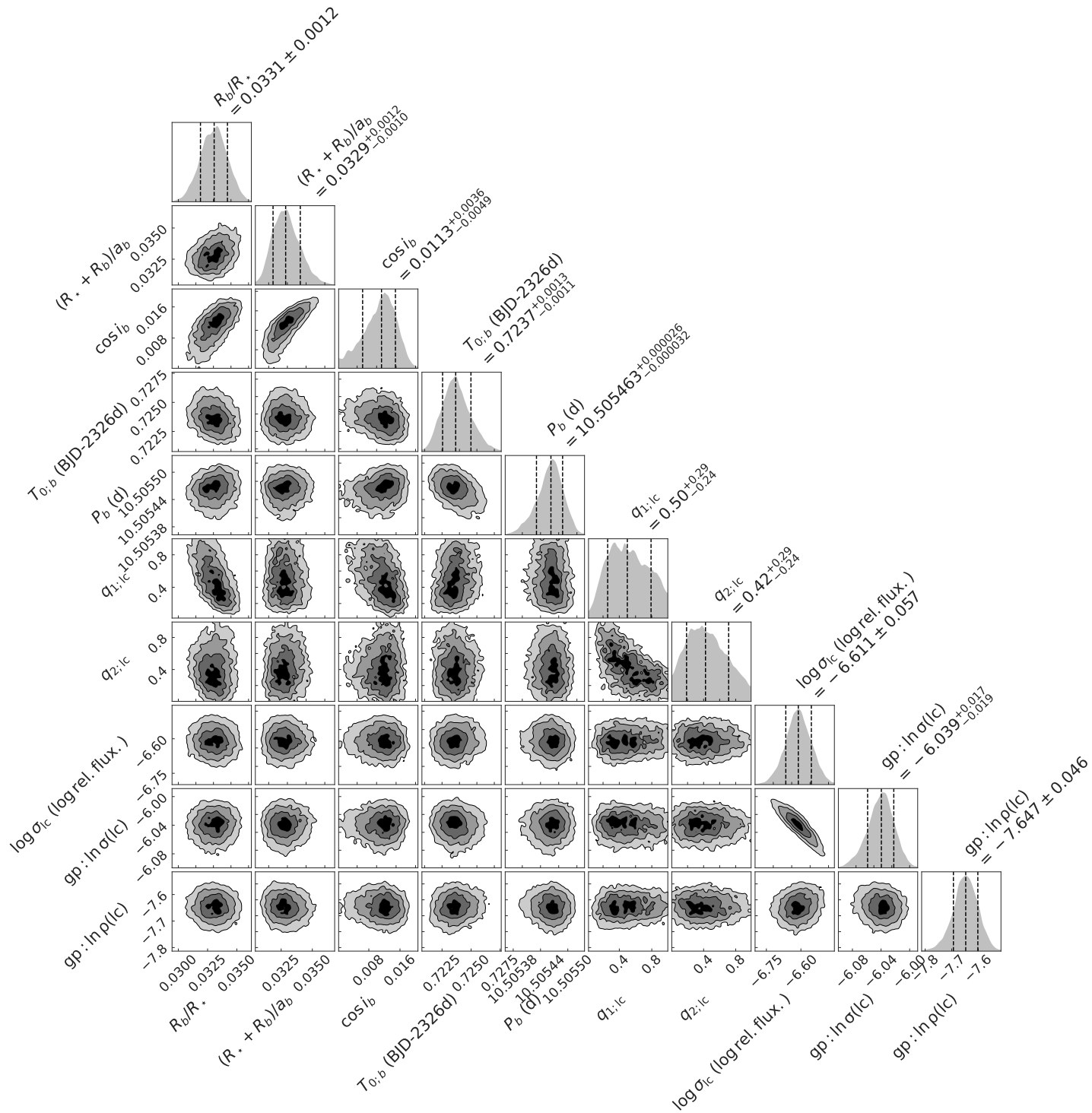
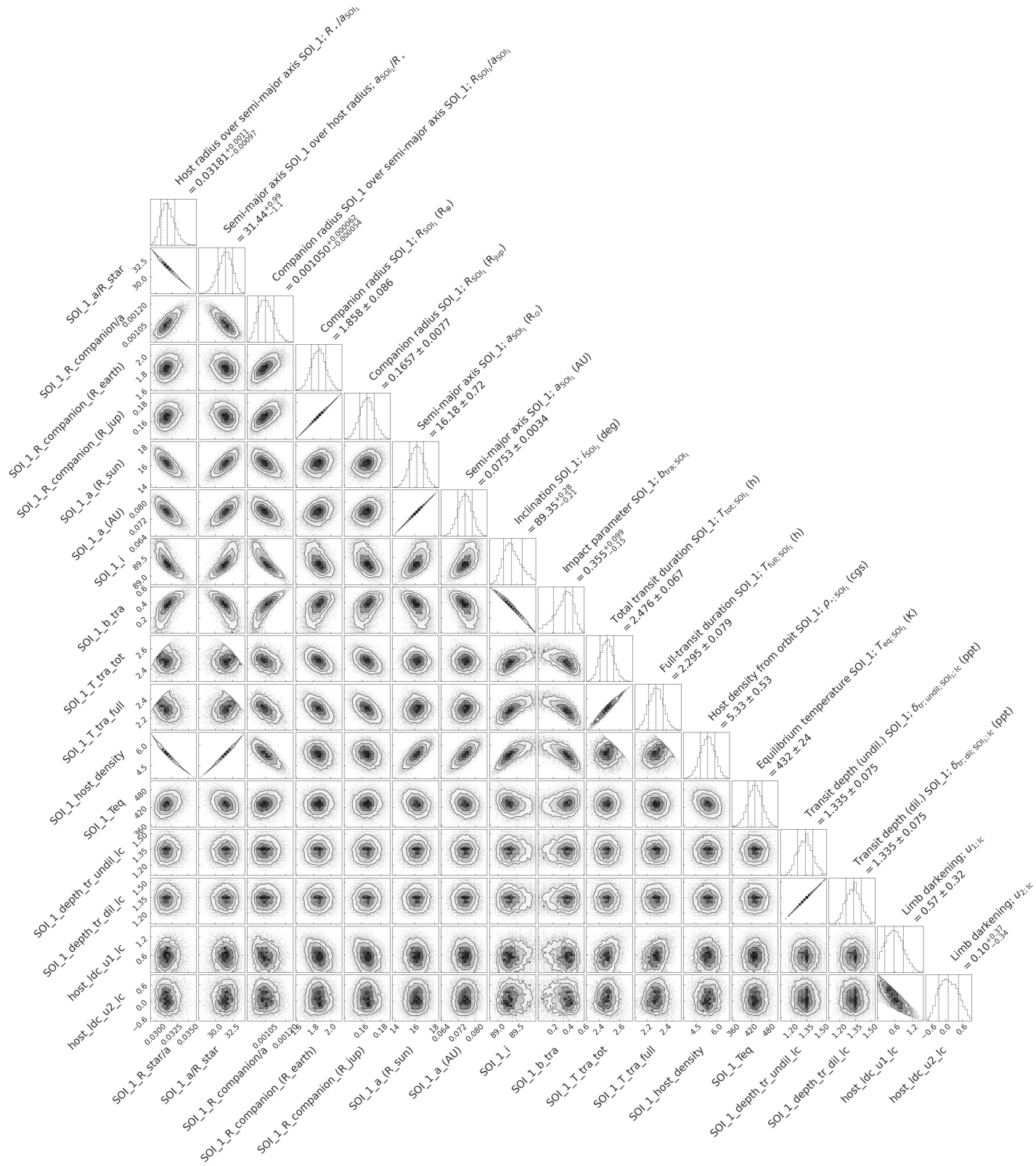Figure 69: TIC 321669174 - fitted corner plot for the second model.

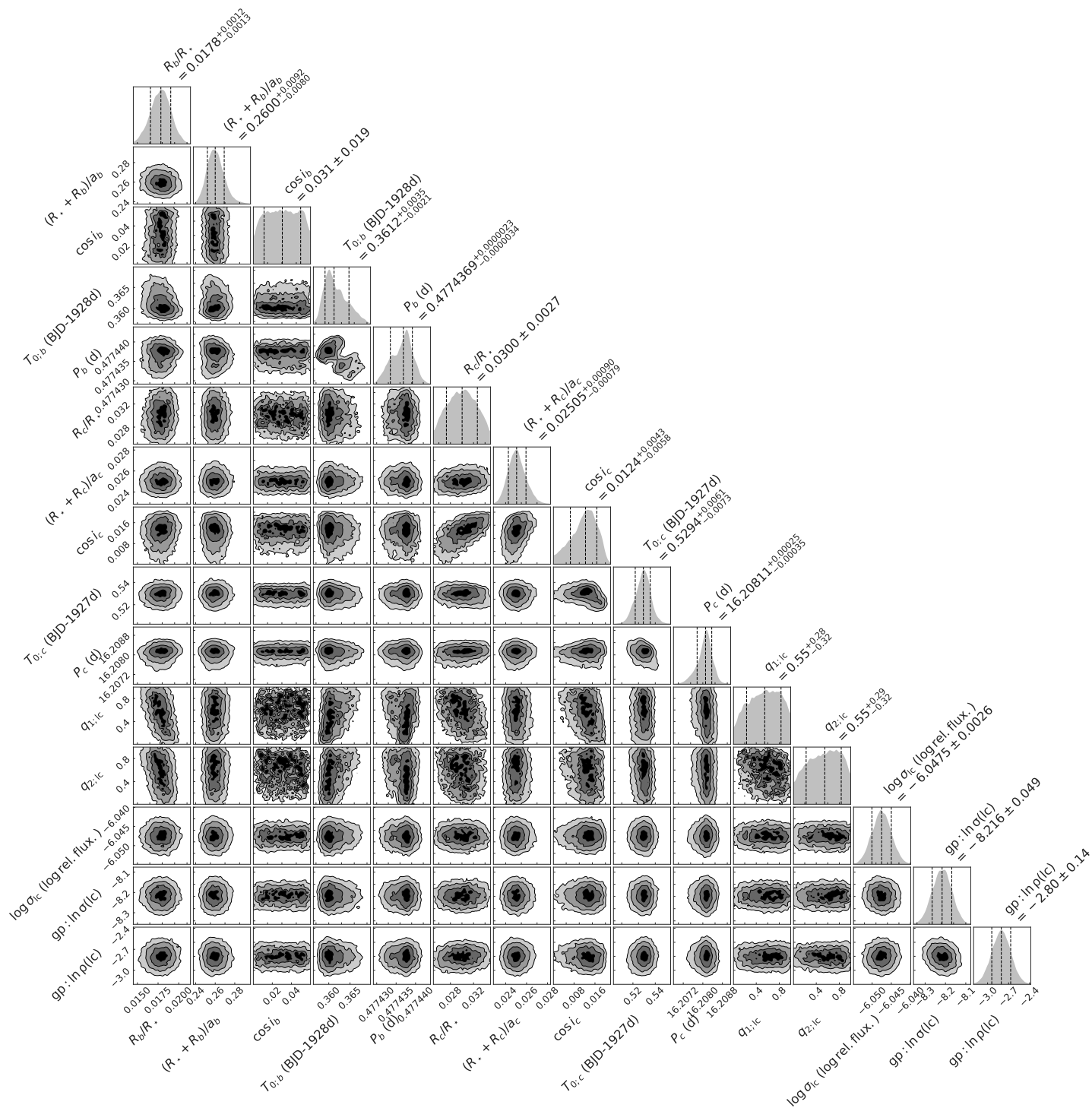Figure 70: TIC 321669174 - derived corner plot for the second model.

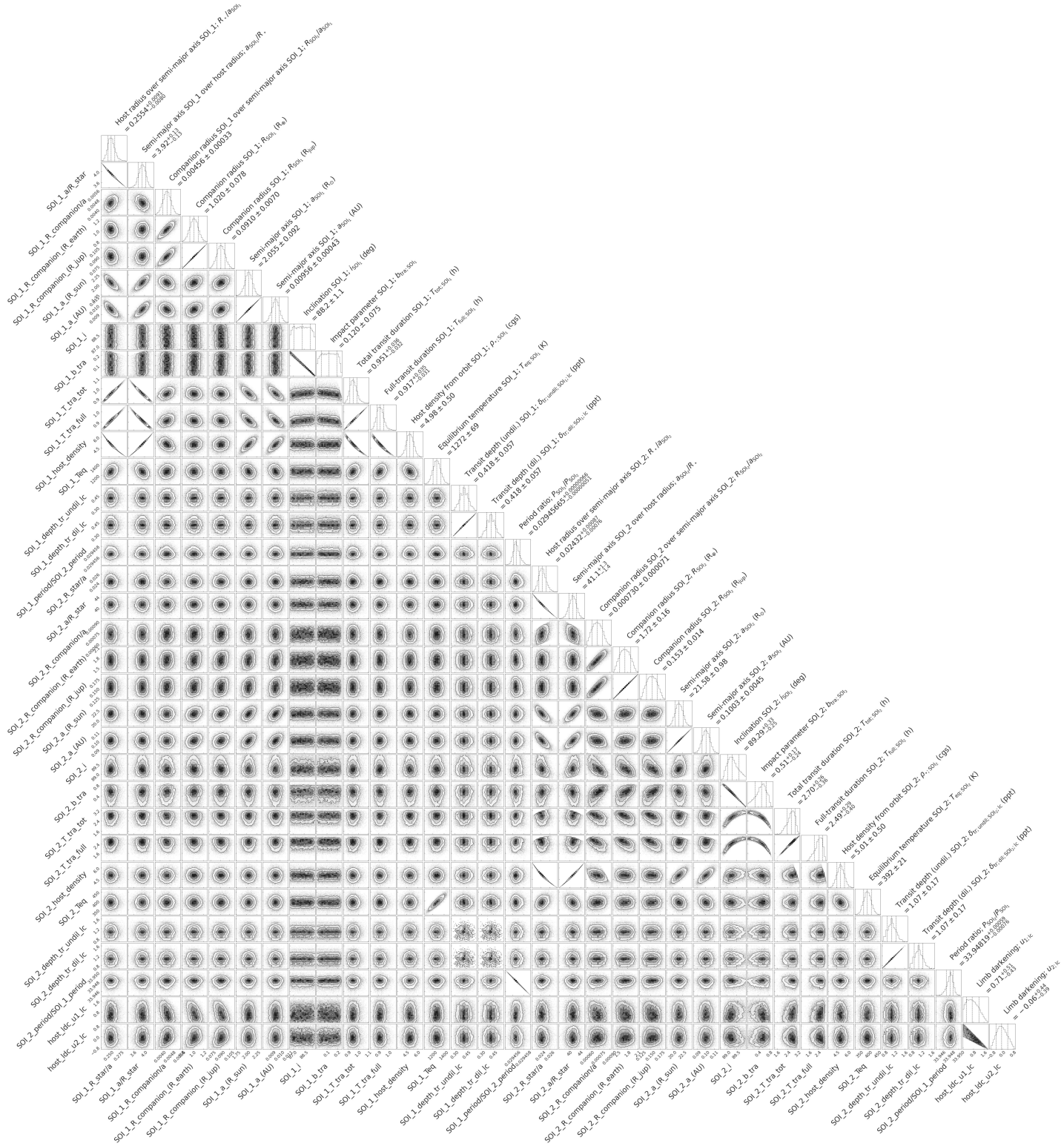Figure 71: TIC 244170332 - fitted corner plot for the second model.

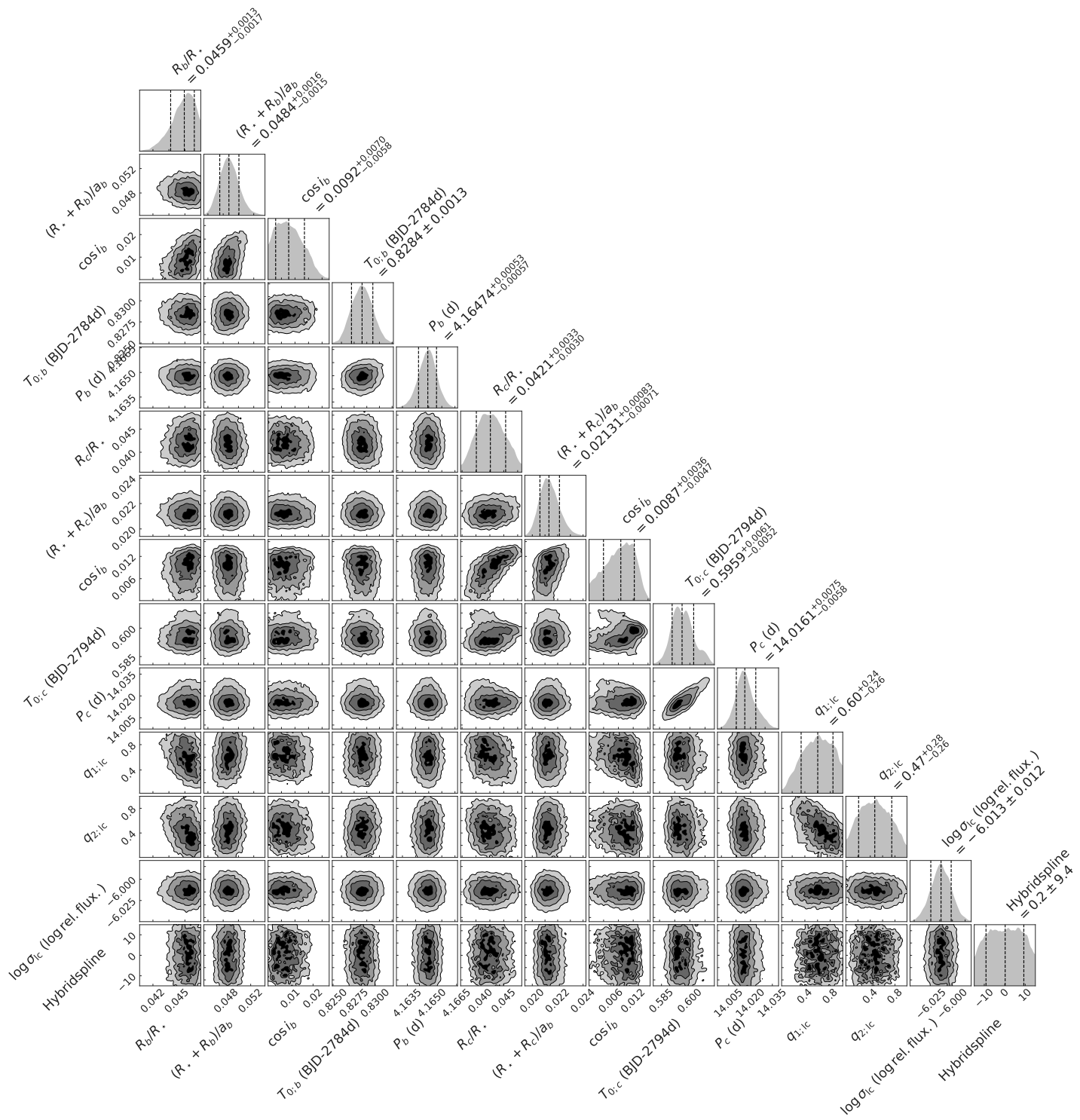Figure 72: TIC 244170332 - derived corner plot for the second model.

114

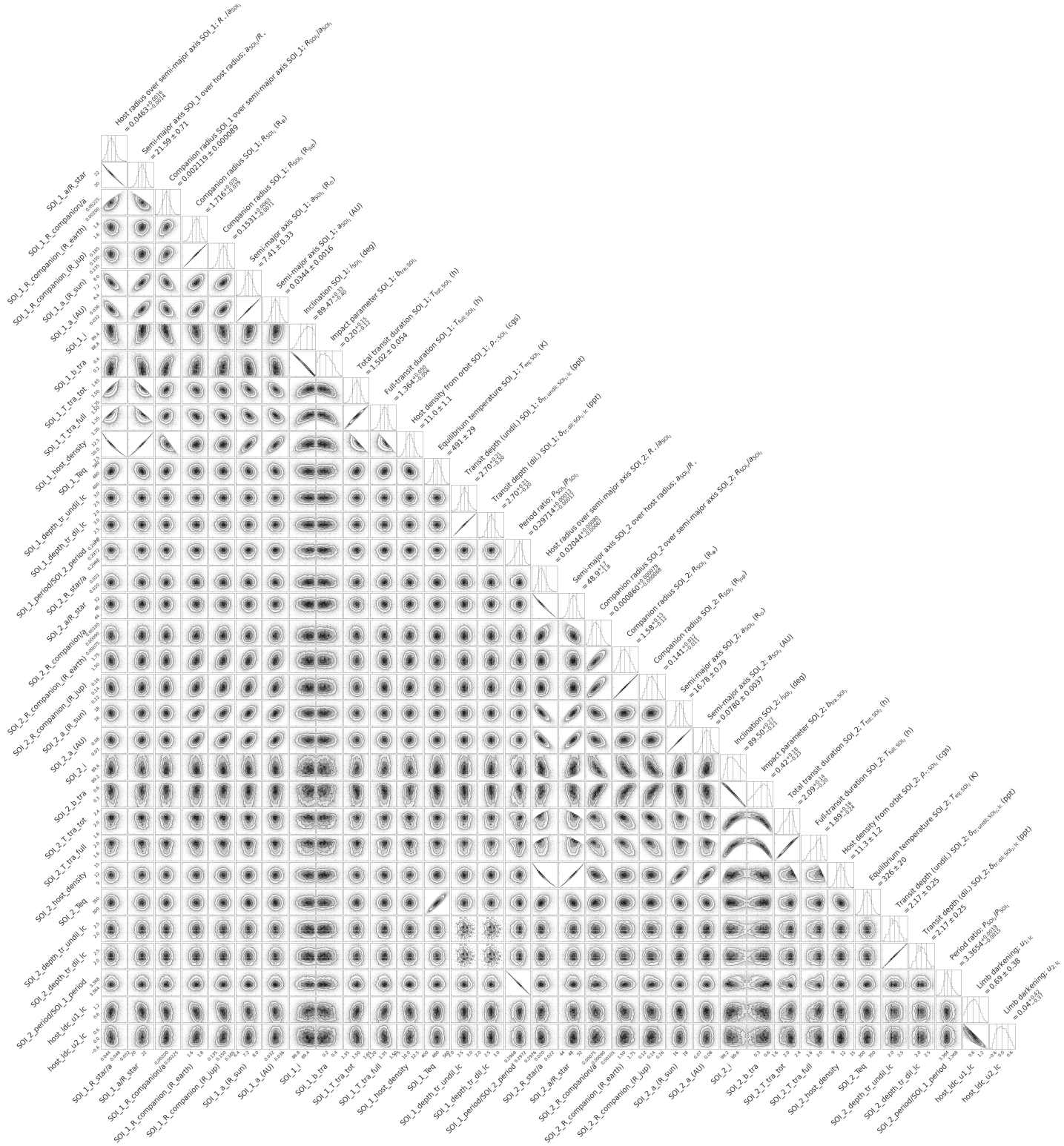Figure 73: TIC 328081248 - fitted corner plot for the second model.

Figure 74: TIC 328081248 - derived corner plot for the second model.

116

References

Borucki William J, Koch David, Basri Gibor, Batalha Natalie, Brown Timothy, Caldwell Douglas, Caldwell John, Christensen-Dalsgaard Jørgen, Cochran William D, DeVore Edna, others . Kepler planet-detection mission: introduction and first results // Science. 2010. 327, 5968. 977–980.

Cameron Andrew Collier. Extrasolar planetary transits // Methods of Detecting Exoplanets. 2016. 89–131.

Charbonneau David, Brown Timothy M, Latham David W, Mayor Michel. Detection of planetary transits across a sun-like star // The Astrophysical Journal. 1999. 529, 1. L45.

Delrez Laetitia, Ehrenreich David, Alibert Yann, Bonfanti Andrea, Borsato Luca, Fossati Luca, Hooton Matthew J, Hoyer Sergio, Pozuelos Francisco J, Salmon Sébastien, others . Transit detection of the long-period volatile-rich super-Earth $\nu 2$ Lupi d with CHEOPS // Nature Astronomy. 2021. 5, 8. 775–787.

Delrez Laetitia, Gillon Michaël, Queloz Didier, Demory Brice-Olivier, Almleaky Yaseen, Wit Julien de, Jehin Emmanuël, Triaud Amaury HMJ, Barkaoui Khalid, Burdanov Artem, others . SPECULOOS: a network of robotic telescopes to hunt for terrestrial planets around the nearest ultracool dwarfs // Ground-based and Airborne Telescopes VII. 10700. 2018. 446–466.

Esparza-Borges E, Parviainen H, Murgas F, Pallé E, Maas A, Morello G, Zapatero-Osorio MR, Barkaoui K, Narita N, Fukui A, others . A hot sub-Neptune in the desert and a temperate super-Earth around faint M dwarfs-Color validation of TOI-4479b and TOI-2081b // Astronomy & Astrophysics. 2022. 666. A10.

Feinstein Adina D, Montet Benjamin T, Foreman-Mackey Daniel, Bedell Megan E, Saunders Nicholas, Bean Jacob L, Christiansen Jessie L, Hedges Christina, Luger Rodrigo, Scolnic Daniel, others . Eleanor: An open-source tool for extracting light curves from the TESS full-frame images // Publications of the Astronomical Society of the Pacific. 2019. 131, 1003. 094502.

Feroz Farhan, Hobson Michael P, Cameron Ewan, Pettitt Anthony N. Importance nested sampling and the MultiNest algorithm // arXiv preprint arXiv:1306.2144. 2013.

George W. Dombi PhD. Using Biweights for Handling Outliers // Karmanos Cancer Center, Detroit, MI. 2012. 9.

Giacalone Steven, Dressing Courtney D, Jensen Eric LN, Collins Karen A, Ricker George R, Vanderspek Roland, Seager S, Winn Joshua N, Jenkins Jon M, Barclay Thomas, others . Vetting of 384 TESS Objects of Interest with TRICERATOPS and Statistical Validation of 12 Planet Candidates // The Astronomical Journal. 2020. 161, 1. 24.

Gibson NP, Aigrain Suzanne, Roberts S, Evans TM, Osborne Michael, Pont F. A Gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy // Monthly notices of the royal astronomical society. 2012. 419, 3. 2683–2694.

Gillon Michaël. Searching for red worlds // Nature Astronomy. 2018. 2, 4. 344–344.

Gillon Michaël. Introduction to exoplanetology. 2021-2022.

Girardi L, Groenewegen MAT, Hatziminaoglou E, Da Costa L. Star counts in the galaxy-simulating from very deep to very shallow photometric surveys with the trilegal code // Astronomy & Astrophysics. 2005. 436, 3. 895–915.

Guerrero Natalia M, Seager S, Huang Chelsea X, Vanderburg Andrew, Soto Aylin Garcia, Mireles Ismael, Hesse Katharine, Fong William, Glidden Ana, Shporer Avi, others . The tess objects of interest catalog from the tess prime mission // The Astrophysical Journal Supplement Series. 2021. 254, 2. 39.

Günther Maximilian N, Daylan Tansu. allesfitter: Flexible star and exoplanet inference from photometry and radial velocity // The Astrophysical Journal Supplement Series. 2021. 254, 1. 13.

Heath Martin J, Doyle Laurance R, Joshi Manoj M, Haberle Robert M. Habitability of planets around red dwarf stars // Origins of Life and Evolution of the Biosphere. 1999. 29. 405–424.

Hedges Christina. Vetting: A Stand-alone Tool for Finding Centroid Offsets in NASA Kepler, K2, and TESS, Alerting the Presence of Exoplanet False Positives // Research Notes of the AAS. 2021. 5, 11. 262.

Henry Gregory W, Marcy Geoffrey W, Butler R Paul, Vogt Steven S. A transiting "51 Peg-like" planet // The Astrophysical Journal. 1999. 529, 1. L41.

Hippke Michael, David Trevor J, Mulders Gijs D, Heller René. Wōtan: Comprehensive time-series detrending in Python // The Astronomical Journal. 2019. 158, 4. 143.

Hippke Michael, Heller René. Optimized transit detection algorithm to search for periodic transits of small planets // Astronomy & Astrophysics. 2019. 623. A39.

Jenkins Jon M., Twicken Joseph D., McCauliff Sean, Campbell Jennifer, Sanderfer Dwight, Lung David, Mansouri-Samani Masoud, Girouard Forrest, Tenenbaum Peter, Klaus Todd, Smith Jeffrey C., Caldwell Douglas A., Chacon A. D., Henze Christopher, Heiges Cory, Latham David W., Morgan Edward, Swade Daryl, Rinehart Stephen, Vanderspek Roland. The TESS science processing operations center // Software and Cyberinfrastructure for Astronomy IV. 9913. VIII 2016. 99133E. (Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series).

Jenkins Jon M, Twicken Joseph D, McCauliff Sean, Campbell Jennifer, Sanderfer Dwight, Lung David, Mansouri-Samani Masoud, Girouard Forrest, Tenenbaum Peter, Klaus Todd, others . The TESS science processing operations center // Software and Cyberinfrastructure for Astronomy IV. 9913. 2016. 1232–1251.

Joshi MM, Haberle RM, Reynolds RT. Simulations of the atmospheres of synchronously rotating terrestrial planets orbiting M dwarfs: conditions for atmospheric collapse and the implications for habitability // Icarus. 1997. 129, 2. 450–465.

Kovács Geza, Zucker Shay, Mazeh Tsevi. A box-fitting algorithm in the search for periodic transits // Astronomy & Astrophysics. 2002. 391, 1. 369–377.

Kreidberg Laura. batman: BAsic transit model cAlculatioN in python // Publications of the Astronomical Society of the Pacific. 2015. 127, 957. 1161.

Lamman Claire, Baranec Christoph, Berta-Thompson Zachory K, Law Nicholas M, Schonhut-Stasik Jessica, Ziegler Carl, Salama Maïssa, Jensen-Clem Rebecca, Duev Dmitry A, Riddle Reed, others . Robo-AO M-dwarf Multiplicity Survey: Catalog // The Astronomical Journal. 2020. 159, 4. 139.

Lightkurve Collaboration , Cardoso José Vinícius de Miranda, Hedges Christina, Gully-Santiago Michael, Saunders Nicholas, Cody Ann Marie, Barclay Thomas, Hall Oliver, Sagear Sheila, Turtelboom Emma, Zhang Johnny, Tzanidakis Andy, Mighell Ken, Coughlin Jeff, Bell Keaton, Berta-Thompson Zach, Williams Peter, Dotson Jessie, Barentsen Geert. Lightkurve: Kepler and TESS time series analysis in Python. XII 2018. ascl:1812.013.

Lissauer Jack J, Marcy Geoffrey W, Rowe Jason F, Bryson Stephen T, Adams Elisabeth, Buchhave Lars A, Ciardi David R, Cochran William D, Fabrycky Daniel C, Ford Eric B, others . Almost all of Kepler's multiple-planet candidates are planets // The Astrophysical Journal. 2012a. 750, 2. 112.

Lissauer Jack J, Marcy Geoffrey W, Rowe Jason F, Bryson Stephen T, Adams Elisabeth, Buchhave Lars A, Ciardi David R, Cochran William D, Fabrycky Daniel C, Ford Eric B, others . Almost all of Kepler's multiple-planet candidates are planets // The Astrophysical Journal. 2012b. 750, 2. 112.

Mallonn M, Von Essen C, Herrero E, Alexoudi X, Granzer T, Sosa M, Strassmeier KG, Bakos G, Bayliss D, Brahm R, others . Ephemeris refinement of 21 hot Jupiter exoplanets with high timing uncertainties // Astronomy & Astrophysics. 2019. 622. A81.

Mandel Kaisey, Agol Eric. Analytic light curves for planetary transit searches // The Astrophysical Journal. 2002. 580, 2. L171.

Mayor Michel, Queloz Didier. A Jupiter-mass companion to a solar-type star // nature. 1995. 378, 6555. 355–359.

Morton Timothy D. An efficient automated validation procedure for exoplanet transit candidates // The Astrophysical Journal. 2012. 761, 1. 6.

Morton Timothy D. VESPA: False positive probabilities calculator // Astrophysics Source Code Library. 2015. ascl–1503.

NASA Official. Data products. 2023a. URL:https://heasarc.gsfc.nasa.gov/docs/tess/second-extended.html (consulted on 04/03/2023).

NASA Official. The Second Extended mission. 2023b. URL:https://heasarc.gsfc.nasa.gov/docs/tess/data-products.html (consulted on 04/03/2023).

Pozuelos Francisco J, Suárez Juan C, Elía Gonzalo C de, Berdiñas Zaira M, Bonfanti Andrea, Dugaro Agustín, Gillon Michaël, Jehin Emmanuël, Günther Maximilian N, Van Grootel Valérie, others . GJ 273: on the formation, dynamical evolution, and habitability of a planetary system hosted by an M dwarf at 3.75 parsec // Astronomy & Astrophysics. 2020. 641. A23.

Quinn Samuel N, Becker Juliette C, Rodriguez Joseph E, Hadden Sam, Huang Chelsea X, Morton Timothy D, Adams Fred C, Armstrong David, Eastman Jason D, Horner Jonathan, others . Near-resonance in a System of Sub-Neptunes from TESS // The Astronomical Journal. 2019. 158, 5. 177.

Rauw Gregor. Variable stars. 2022-2023.

Ricker George R, Winn Joshua N, Vanderspek Roland, Latham David W, Bakos Gáspár Á, Bean Jacob L, Berta-Thompson Zachory K, Brown Timothy M, Buchhave Lars, Butler Nathaniel R, others . Transiting exoplanet survey satellite // Journal of Astronomical Telescopes, Instruments, and Systems. 2015. 1, 1. 014003–014003.

Savitzky Abraham, Golay Marcel JE. Smoothing and differentiation of data by simplified least squares procedures. // Analytical chemistry. 1964. 36, 8. 1627–1639.

Shields Aomawa L, Ballard Sarah, Johnson John Asher. The habitability of planets orbiting M-dwarf stars // Physics Reports. 2016. 663. 1–38.

Speagle Joshua S. dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences // Monthly Notices of the Royal Astronomical Society. 2020. 493, 3. 3132–3158.

Stassun Keivan G, Oelkers Ryan J, Paegert Martin, Torres Guillermo, Pepper Joshua, De Lee Nathan, Collins Kevin, Latham David W, Muirhead Philip S, Chittidi Jay, others . The revised TESS input catalog and candidate target list // The Astronomical Journal. 2019. 158, 4. 138.

Winn Joshua N. Transits and occultations // arXiv preprint arXiv:1001.2010. 2010.