An Assessment of K2 Data Quality for the Search for Survivor Exoplanets Around Hot Subdwarfs

UNIVERSITY OF LIÈGE

FACULTY OF SCIENCES

Mathieu Motte

Master In Space Sciences, Research Focus

Academic Year: 2022 - 2023

Supervisors: VAN GROOTEL Valérie
POZUELOS Francisco J.

Acknowledgements

First of all, I would like to thank my supervisors Valérie Van Grootel, FNRS Research Associate at the University of Liège, and Francisco J. Pozuelos, scientific researcher at the Instituto de Astrofísica de Andalucía-CSIC, for accepting me in their team to conduct this work. They gave me good advice and help when I was in need, and made it possible for me to complete this project.

I would like to thank the other members of the team Martín Devora-Pajares, Ph.D. student at the University of Granada, and Antoine Thuillier, Ph.D. student at the University of Liège, for the time I took from them when I had questions.

I would like to thank my fellow Master's student Cyril Dethye, who worked on the Kepler data, for the help we gave each other. Finally, I would like to thank my student office mates: Maria Groyne and Ismaël Lahmaid.

Abstract

Since the first detection of an exoplanet around a main sequence star in 1995, the race for detection and characterisation grows exponentially, with numerous dedicated techniques and instruments developed. The search for objects around other stars helps us to understand better the life and death of planetary systems. In that context, when a star enters the red-giant branch phase (RGB) it expands and can engulf close-in planets. The state and the evolution of the exoplanetary systems directly after this phase are still poorly understood. Hot subdwarfs are hot and small He-burning objects which experienced strong mass loss on the RGB. Therefore, these stars constitute excellent opportunities for addressing the question of the evolution of exoplanetary systems directly after the RGB phase.

In this work, I aim to evaluate the capacities of K2 data to detect exoplanets around hot subdwarfs by performing injection-and-recovery tests on the K2 light curves corrected by K2sff and EVEREST detrending pipelines. Then, I apply the lesson learnt from those tests by performing an exoplanetary search on 48 K2 targets thanks to the SHERLOCK pipeline.

In this master's thesis, we will first introduce important scientific backgrounds on hot subdwarfs and on the search for exoplanets around those stars. Then, the tools and methods used during this project will be presented. Finally, the results of the injection-and-recovery tests will be exposed and the best candidates found in the exoplanetary search will be depicted.

Table of contents

List of Figures

List of Tables

Chapter 1

Context and Scientific objectives

The primary goal of this chapter is to establish a scientific context and provide background information for this master thesis. It serves to outline the objectives of the thesis and review the existing state-of-the-art related to these objectives. The chapter will be divided into two main parts: the first part will present crucial insights into hot subdwarfs, and the second part will focus on the exploration of exoplanets around these particular stars.

1.1 Hot Subdwarfs

1.1.1 Main Characteristics and Classification

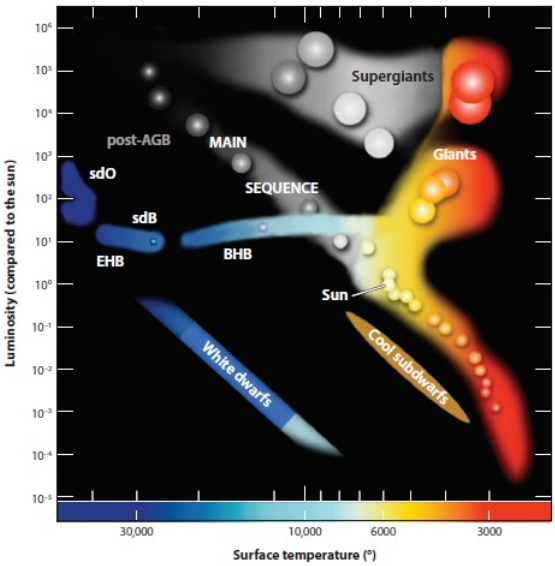


Figure 1.1: Hertzsprung-Russell diagram highlighting the position of hot subdwarfs [1].

Hot subdwarfs of B or O type are stars located in the Hertzsprung-Russell diagram between the main sequence and the white dwarfs. They traduce different stages in the late

evolution of low-mass stars [1] [2].

Subdwarfs of B-type (sdBs) have a helium-burning core and are at the blue end of the horizontal branch (HB) which corresponds to core-He burning objects. O-types (sdOs) correspond to a wide range of objects of the post-red-giant branch, post-HB, and post-asymptotic giant branch (post-AGB) [1] [2].

SdBs have an T_{eff} = 20 000 - 40 000 K whereas sdOs have an effective temperature between 40 000-80 000K. Both star types have a $\log(g) = 5.2 - 6.2$ [3]. SdB stars form a rather homogeneous spectral type, with an under solar He abundance in their atmospheres for most of them. On the contrary O-type subdwarfs present a larger variety of spectra, from He-poor (for about 1/3 of them) to He-rich in their majority [1].

The atmospheric parameters of sdB stars place them on the EHB. However, in comparison to the other horizontal-branch stars, they have an extremely thin hydrogen envelope, and, thus, can not sustain H-burning. Therefore, their evolution is also very different because they avoid the AGB and go directly toward the white dwarf cooling sequence [1].

1.1.2 The Question of the Origins of Hot Subdwarfs

The origins of stars in the EHB are still not fully understood. To be formed, they have to undergo mass loss processes to go from a large to a very thin hydrogen-rich envelope around the same moment the He-core reaches the mass required for the helium flash ($\sim 0.47\ M_{\odot}$) [1]. Afterwards, they still have to sustain a helium-burning core [3].

To explain their origin, multiple scenarios have been proposed. Since 2001, it is known that about 60% of the subdwarfs are in a binary system, with half of them in a close binary system with periods from 0.07 to 30d and the other half with periods up to several years [3]. Close binaries sdBs are generally paired with a white dwarf of M-dwarf nature. In contrast, wider binary systems featuring subdwarf B stars tend to have companions of FGK types. Therefore, the principal mechanism of formation consists of the evolution of binaries [4][5]. In this case, two main sequence stars of different masses evolve in a binary system. The most massive one will be the first to reach the red-giant branch and thus its radius will increase first until reaching its Roche lobe. From here, two dynamic cases have to be analysed, a stable and an unstable one [4][5]:

- If the mass transfer is dynamically unstable, a common envelope is created. The orbital energy of the two stellar cores is transferred to the envelope, therefore the orbital period decreases. Finally, the common envelope is ejected and a close binary system with an sdB and a main sequence star is formed. If the companion reaches the red giant branch, another mass transfer could happen and can lead to a close binary system with an sdB and a white dwarf. This scenario is commonly known as the common envelope (CE) ejection [4][5].

- If the mass transfer is dynamically stable, once the companion reaches its Roche lobe, the primary slowly accretes matters from the secondary. In this case, the companion loses mass and becomes an sdB. Therefore, it results in a binary system with a sdB and

a main sequence star companion with larger orbital periods. This scenario is commonly known as the Roche Lobe OverFlow (RLOF) [4][5].

The RLOF and CE ejection scenarios can also be mixed with a first RLOF phase producing a white dwarf and a main sequence star and then a CE ejection phase when an sdB+WD binary is produced [4][5].

Nonetheless, about 40% of the hot subdwarfs are not known to be in a binary system. In a single-star system, the mechanisms explaining the mass loss during the helium flash are not well understood. In 1984, Webbink explained that a merger of two He-white dwarfs could form an EHB star [6]. However, several observations questioned this hypothesis. First, WD binaries that can produce sdBs are quite rare, but few candidates have been identified. Second, mass distributions of both single and binaries have a narrow peak at $0.47\ M_{\odot}$. However, the merger mechanism should give a broader range ($0.4\text{-}0.7 M_{\odot}$). Finally, such merger scenarios should produce high rotator sdBs however, observations show that almost all single sdBs have a slow rotation [7][8]. Other scenarios have been proposed: stellar wind mass loss, helium mixing by internal rotation, etc. However, the conditions required for those scenarios to happen are unlikely to be met [4][5]. Nonetheless, the most interesting scenario for this master's thesis is the possibility of the ejection of the star's envelope by the engulfment of a substellar companion. It will be described more precisely in the next part.

1.2 Search for Exoplanets around Hot Subdwarfs

1.2.1 Planet-Red Giants Interactions

It is known, long before the discovery of exoplanets, that during the post-main-sequence evolutionary stage, stars undergo an expansion of 100 to 1000 times their initial radius and that this expansion will engulf nearby planets. The first studies of possible planets' engulfment were made in the solar system. For example, in 1987, Goldstein [9] described the fate of the Earth around the Sun. Employing stellar models and incorporating Earth's density profile to consider ablation effects, he determined that the timescale for Earth's orbital decay within the Sun's envelope is merely 210 years.

More generally, interactions between planets and red giant envelopes are complex due to the number of parameters and phenomena to consider. In their 2010 paper, Bear and Soker [10] conducted a study exploring the potential role of substellar companions in the formation of sdB stars. They investigated various phenomena that contribute to an increased mass loss rate resulting from the presence of a substellar companion. The researchers highlighted that this increase is not solely due to the deposition of gravitational energy. While the deposition of gravitational energy is effective for brown dwarfs and very massive planets, smaller planets need additional processes. Firstly, exoplanets can accelerate the rotation velocity of the star, leading to an amplification of the magnetic field, which in turn increases the mass loss. Secondly, during the common envelope phase, planets can excite non-radial p-waves, causing an increased mass loss within the equatorial plane. Additionally, if the temperature of the envelope exceeds the virial temperature of the planet, the planet may be destroyed. This

destruction process can temporarily reduce the stellar luminosity by introducing low entropy material. If the material reaches the core, it triggers the release of gravitational and nuclear energy by replenishing it with new hydrogen-rich materials, resulting in stellar expansion and a further increase in mass loss. Moreover, the interaction between the core and the accreted material may give rise to the formation of an accretion disk, which can generate a jet of ejecta.

More recently, in the paper by O'Connor et al. 2023 [11], the response of the stellar envelope to the dissipated energy caused by an engulfed planet was simulated. Using a stellar model, they analysed the effects of drag force, orbital decay, and deposition of angular momentum from the planet. They provided an explanation indicating that planets are typically destroyed within the core of a star. As a planet approaches a stage where the local sound speed exceeds the escape velocity at its surface, it undergoes a thermal disturbance. However, it is crucial to compare the time it takes for the planet to spiral inward with the timescale of heating in its interior. In the case of an irradiated giant planet, the increasing entropy in its outer layer hampers convection, resulting in a much longer heating timescale than the inspiral time. When studying disrupted envelopes, it was discovered that even during significant expansion, no material attains a velocity surpassing the escape velocity, preventing any ejection beyond this point. However, the researchers did not account for factors such as dust grain formation within the envelope as the star evolves, stellar rotation, and magnetic activity, which could potentially affect the rate of mass loss. Nonetheless, they established a critical planet mass beyond which the outer layers of the envelope expand supersonically and encounter shocks. Additionally, they explored the possibility of planet survival if the Roche lobe overflow of a giant planet leads to stable mass transfer, effectively halting or even reversing its orbital decay.

1.2.2 Search for Survivor Exoplanets

1.2.2.1 Importance of Hot Subdwarfs

Therefore, based on the insights provided in the previous section, exploring exoplanets around hot subdwarfs could greatly enhance our understanding of their formation processes and shed light on the survival potential of such planets.

It should be noted that normal HB stars, including Red Clump (RC) stars, typically exhibit substantial size, ranging from $\sim$ 10 to 1000 $R_\odot$, and possess typically a mass of about 1.5 $M_\odot$. Consequently, employing transit and radial velocity (RV) methods to detect exoplanets around these types of stars presents significant challenges due to factors such as the dilution of transit depth and other sources of noise. The hot subdwarfs, on the contrary, have a radius between 0.1 and 0.3 $R_\odot$ making far easier the detection of small transiting bodies. As we will develop in section 2.1, the efficiency of the transit method will be dependent on the ratio between the radius of the planet and the radius of the star. Therefore, Earth-size planets and slightly below around hot subdwarfs are detectable [2], this will be confirmed in Chapter 3 of this work.

Spectroscopically distinguishing between RGB and RC stars can sometimes be arduous,

primarily because stars tend to expel little envelope during the RGB phase and have very similar atmospheric parameters [12]. As a result, only large or massive planets are typically detected around classical evolved stars, and there have been relatively few instances of close-in giant planets identified around such stars when compared to solar-type main sequence stars. This scarcity of detections may be attributed to the potential engulfment of planets by the host stars. Currently, it remains challenging to ascertain whether small planets or remnants, such as dense cores of former giant planets, exist in these systems. Another explanation is that this discrepancy may indicate different formation mechanisms for planets in the vicinity of these stars [13].

The presence of a survivor planet may be directly associated with the ejection of the envelope which creates hot subdwarfs. In the case of classical evolved stars, however, the planet would gradually spiral inward within the envelope until it is completely destroyed. In addition, the duration of the sdB phase is approximately 100 million years, while the sdO phase lasts between 10 to 20 million years. Consequently, the chances of discovering second-generation planets or planets that have not been engulfed and migrated from distant orbits are quite low [2].

While a majority of studies concerning planetary systems around evolved stars have focused on white dwarfs, it is important to notice that most of these stars undergo two distinct giant phases during their evolution: the RGB and the Asymptotic Giant Branch (AGB). The AGB phase, characterised by stellar expansion and strong mass loss, followed by the formation of a planetary nebula, significantly affects the stability of the orbits of surrounding objects. As a result, the dynamics of planetary systems around AGB stars are strongly influenced and can differ from those around other types of stars. In conclusion, the effect of the RGB expansion on planetary systems cannot directly be concluded with the study from white dwarfs. On the contrary, hot subdwarfs are good candidates to try to answer this question [2].

1.2.2.2 State-of-the-Art of the Search of Planets Around Hot Subdwarfs

To this day, no planets orbiting hot subdwarfs have been confirmed. However, there have been some reported detections. For instance, in 2007, Silvotti et al. [14] observed a signal around V391 Peg using the pulsation-timing method, indicating the presence of a planet with a few Jupiter masses at approximately 1.7 AU from the star. In 2017, Silvotti et al. [15] criticised this candidate based on changes in its amplitude and period, utilising a more extensive dataset.

Another claim was made by Lutz et al. in 2012 [16], who employed the same method and stated the existence of a 5.58 Jupiter mass planet around DW Lyn at a distance of 1.18 AU. However, this finding was not successfully confirmed by Mackebrandt et al. in 2020 [17].

Furthermore, Charpinet et al. in 2011 [18] and Silvotti et al. in 2014 [19] reported the detection of respectively 2 and 3 Earth-sized planets with orbital periods of a few hours around KIC 05807616 and KIC 10001893, respectively. These claims were disputed by Krzesinski in 2015 [20] and Blokesz et al. in 2019 [21] who explained the signals are more

probably combination frequencies of stellar pulsating modes. This interpretation remains disputed nowadays.

In 2009, Geier et al. [22] claimed the discovery of a giant hot Jupiter with a period of 2.4 days around HD 149382 with the RV method. But it was discarded by Norris et al. in 2011 [23] because high precision radial velocity measurement of the Hobby-Eberly telescope spectrograph excluded the presence of any substellar companion with a period inferior to 28 days.

Frequent discoveries of red dwarf or brown dwarf companions to hot subdwarfs have been made through ground-based observations using photometric or RV techniques [24][25][26][27]. However, no Jupiter-like planet has been identified so far. On the other hand, by using the concept of eclipse-timing variations, which rely on the gravitational influence a planet will have on the binaries and thus on the event time of the eclipse, several massive planets have been claimed in sdB+dM eclipsing binaries [28][29]. These findings potentially correspond to first or second-generation planets, or hybrid planets formed from the accretion of ejected stellar material onto remnants of first-generation planets [30][31]. However, these claimed planets often face scepticism, as their properties frequently change with new measurements, their orbits are predicted dynamically unstable, and they are often discarded after new measurements [32]. Furthermore, none of these detections has been independently confirmed using alternative methods [2].

1.2.3 Objectives of the Thesis

The main objectives of this thesis are twofold. Firstly, it aims to enhance our understanding of the K2 data and explore its capability of effectively detecting transiting planets around hot subdwarfs. The study will involve a comparative analysis of the efficiency between the EVEREST and K2sff correction pipelines to determine the optimal pipeline to use for the search of planetary transits in hot subdwarf light curves and understand their characteristics.

Secondly, using the SHERLOCK pipeline, the thesis will present the findings from the search of transits in 48 hot subdwarfs from the EVEREST dataset, including a vetting done on the most promising candidates.

By accomplishing these objectives, the research endeavours to contribute valuable insights into 1) the evolution of planetary systems directly after the RGB phase and 2) the surviving capability of a planet around a red giant.

Chapter 2

K2 data and Methods

Having set the scientific context and objectives, this chapter will delve into a comprehensive account of the various methods, data, and programs employed in this thesis. A theoretical overview of the transit method for detecting exoplanets will be presented. Additionally, detailed information about the Kepler Telescope and the K2 mission, along with the characteristics of its data, will be provided. The chapter will conclude by introducing the SHERLOCK and MATRIX pipelines.

2.1 Transit Method for Exoplanets Detection

The presence of exoplanets around a star influences in various ways the received signal, and numerous techniques are available to detect these exoplanets, such as radial velocity, gravitational microlensing, and direct imaging. However, in this thesis, we will work with the most used and extended one, the transit method.

An eclipse occurs during the obstruction of a celestial object with another. When the size difference between the two objects is significant, it is referred to as transit or occultation. A transit occurs when the planet crosses the line of sight between the observer and the star while an occultation appears when the star crosses the line of sight between the observer and the planet. Both phenomena will induce a periodic decrease in the star's luminosity. In cases of a circular orbit, transits and occultations always appear, but in elliptical orbits, only one may be observed. In general, for a small dark occulted object by another big bright object (e.g. exoplanet and star), occultations are more challenging to observe compared to transits because its depth is directly proportional to the ratio of the exoplanet's flux to the star's flux.

As depicted in Figure 2.1, a transit event is marked by four contact points. The periods from t_{I} to t_{II}, and t_{III} to t_{IV}, represent the ingress and egress durations. The interval between t_{II} and t_{III} corresponds to the full duration, signifying the time when the entire disk of the celestial object is crossing in front of another. The total duration of the transit is given by the difference between t_{IV} and t_{I}.

A transit can be characterised by four parameters. First, the "P" parameter represents the period of occurrence, which is the time interval between two consecutive transits. It directly corresponds to the orbital period of the planet and provides the ratio between the semi-major axes of the planetary orbit and the stellar radius.

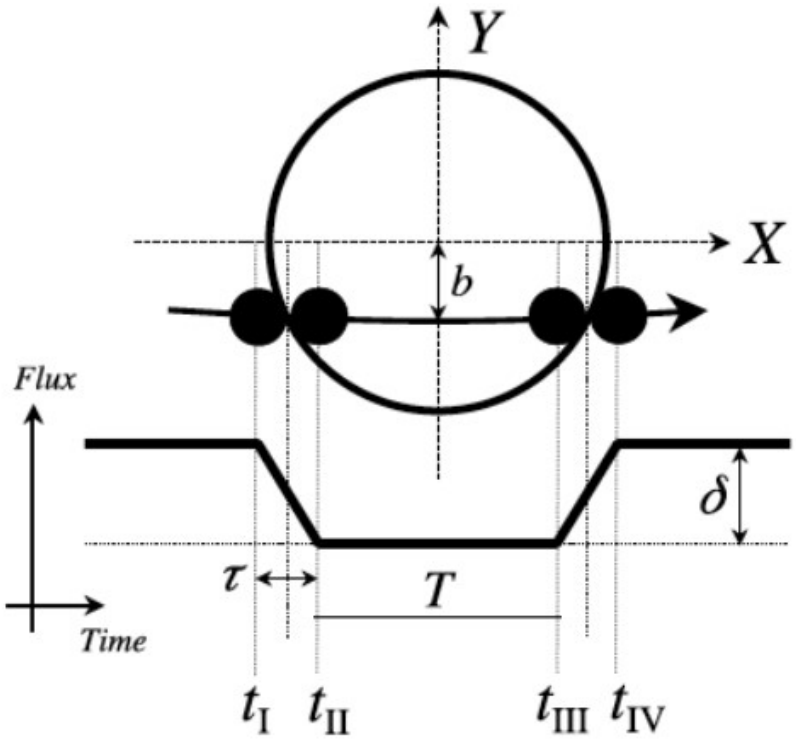


Figure 2.1: Shape of transit with its parameters [33]

Second, the "depth" of the transit, denoted as δ, signifies the difference between the average stellar flux and the minimum flux observed during the transit. This depth is directly proportional to the square of the ratio between the radii of the planet and the star.

Third, the impact parameter, denoted as "b", is the angular distance between the centre of the star and the centre of the planet at conjunction. This parameter ranges from 0 to 1 and determines whether the eclipse is full or grazing. Therefore, the total transit duration will depend on the impact parameter with a maximum duration when b=0 and minimum when b=1.

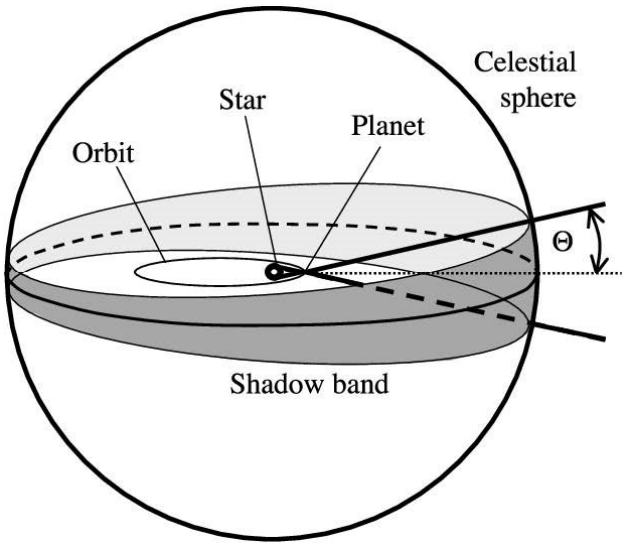


Figure 2.2: Shadow band of a transit [33]

An important point to take into account is that transits only appear in a specific configuration. One can determine the probability of an eclipse with its geometric configuration (i.e. Figure 2.2). When an object is in front of another one, the shadow will create a cone and as this object moves the shadow will sweep out a band on the celestial sphere. The observer can only observe this transit if they are located inside this band. Defining "a" as the semi-major axis, "e" as the eccentricity, ω the argument of the periapsis, $R_\star$ the stellar radius and R_p the planetary radius, the probability of transit can be expressed as :

$$p_{tra} = \left(\frac{R_\star \pm R_p}{a}\right)\left(\frac{1 + e\sin\omega}{1 - e^2}\right) \tag{2.1}$$

Therefore, assuming the usual case for a planetary transit, $R_p << R_\star$ and a circular orbit, $p_{tra} = R_\star/a$.

The transit depth increases with the exoplanetary radius and the probability of observing it increases with a decrease in the exoplanet's semi-major axis. As a result, the transit method proves to be highly suitable for detecting planets that are either large or orbiting close to their host stars or both. Moreover, thanks to the numerous space missions equipped with photometric instruments, the transit method benefits from abundant data availability, making it more accessible compared to other methods.

The transit method is the most successful technique with 4103 confirmed planets over the 5483 with all techniques combined. It is the principal method to detect earth-size planets with 191 confirmations, and notably the famous Trappist-1 system with 7 planets [34] [35]. The difference between detection techniques can be observed in Figure 2.3. While transit is more focused on short-period planets, radial velocities for example are more directed toward big and farther-from-their-star planets.

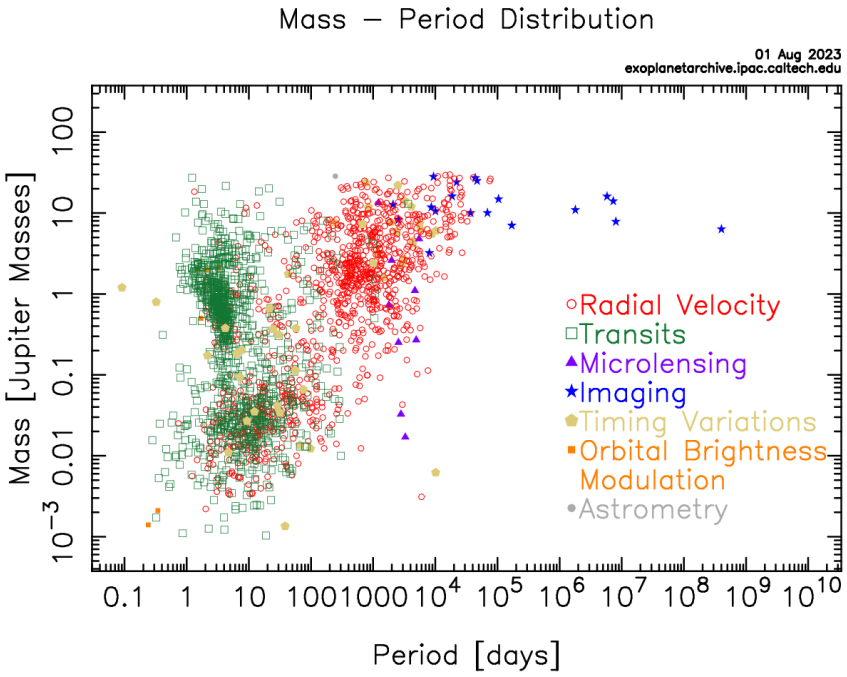


Figure 2.3: Mass-period distribution of confirmed exoplanet with respect to their detection method. [36]

The planets on which this master's thesis focuses have been engulfed by their star therefore we expect periods on the scale of days if they have survived. As such, the transit method is particularly well-suited to align with our research objectives.

2.2 Kepler and K2 Mission

In this master thesis, I will make use of the data from the K2 mission. To understand their specificity, it is important to first present the mission itself.

2.2.1 Scientific Context and Requirements

The Kepler mission principally aimed to study the structure and population of planetary systems with special importance on the detection of Earth-size rocky planets in the habitable zone (HZ) of its system. High photometric precision is therefore required [37].

In more detail, the mission aimed to analyse over 170,000 stars in order to establish the occurrence rate of planets larger than 0.8 times the size of Earth, either within or in close proximity to the habitable zone. It sought to examine various properties of exoplanets, such as their orbital parameters, albedo, size, mass, and density, in relation to the characteristics of their host stars. Additionally, the mission aimed to validate potential planet candidates discovered using other detection methods, as well as to identify new candidates within the same planetary systems. Another objective was to determine the stellar properties associated with these planetary systems through asteroseismology.

In its paper of 2018, Borucki summarised in a table the science requirements [38].

Requirement	**Required values**
Target Stars	Monitor 170,000 stars at a 30-min cadence in a single FOV. Monitor 512 stars at 1-min cadence in a single FOV. The subset can be changed every 3 months.
System photometric precision	1.9×10^{-5} (19 ppm) 6.5-h integration for 12th-mag G2 dwarf
Continuous Observing	Single, inertial FOV for targets. No obscuration by Sun, Earth, Moon, and planets
Mission lifetime	4 years to observe four transits of planets in 1-year orbits

Table 2.1: Science Requirements of Kepler, from [38]

2.2.2 Kepler Instrument

The Kepler mission used a total of 42 CCD (Charge-Coupled Device) photodetectors, with each detector having a resolution of 2.2 million pixels arranged in a grid format of 2200 columns by 1024 rows. The size of each pixel was measured to be $27 \times 27 \mu$m, which corresponds to an angular size of 3.98 arcseconds in the sky [39].

The Kepler telescope is a 1m class Schmidt telescope with a FOV of 16° in diameter. The high FOV is justified by the number of stars to observe simultaneously. In order to minimise the impact of dark current and radiation, the CCDs were cooled to a temperature of -85°C.

2.2.3 K2 Mission

The Kepler spacecraft had four reaction wheels which have for objective to control the attitude of the spacecraft. A spacecraft needs three reaction wheels in order to assure the rotation around its three attitude angles. Nonetheless, given that reaction wheels frequently constitute the initial critical component to experience failure on a spacecraft, it's not uncommon to include an additional one as a redundancy measure. After four years of operation, on 14 May 2013, a second reaction wheel failed following the failure of the first one earlier in the mission. This event should mean the end of the Kepler mission. However, as all the systems were perfectly operational, engineers thought about a method to extend it [38].

Without a third reaction wheel, the pointing stability and accuracy are supported by the 2 remaining wheels on the Y and Z axes and by thrusters around the X axis. The disturbance created by solar pressure has to be minimised on the roll axis (X-axis) for the time of observation. Therefore, the telescope was pointing in the orbital plane and thus toward the ecliptic. It provides a balance of pressure because the sun follows the line of symmetry of the spacecraft in the XY plane [40].

The spacecraft's stability around the roll axis during observation is assured by a careful selection of the initial rolling and correcting drifts every 12 hours. The reaction wheels stabilise the spacecraft on its Y and X axis by countering the solar pressure. The thrusters are used to evacuate the momentum accumulated during the observations every 2 days [40].

Contrary to Kepler which had one stellar field, K2 observed multiple independent target fields called campaigns. Each campaign has a duration limited by the solar flux constraints. The campaign organisation is presented in Figure 2.4. The observations lasted about 75 days and were followed by a data transfer period, then an inactivity period if needed and finally a checkout period [40]. Such as Kepler, K2 will observe with 2 exposure times: 30 minutes for the long cadence (LC) and 1 minute for the short cadence (SC).

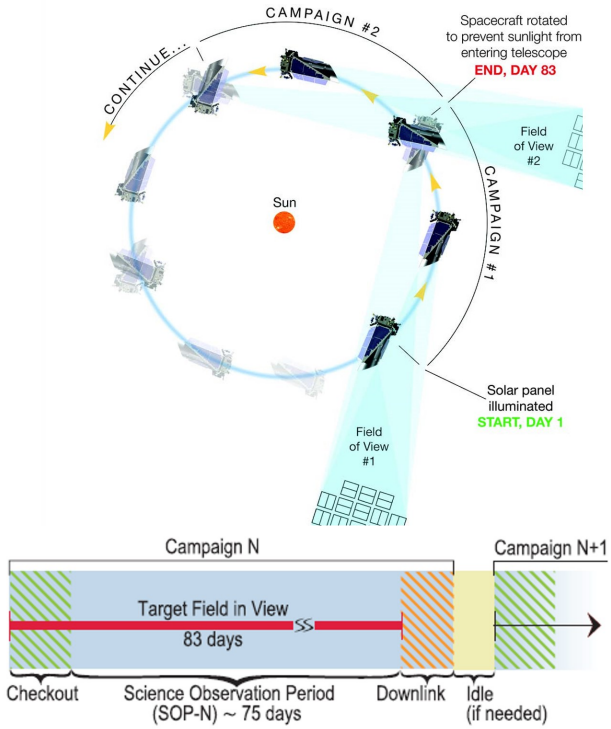


Figure 2.4: K2 mission phases and observation campaign[40]

This new configuration of the mission influenced the scientific objectives of the Kepler Telescope, its data quality and its data transfer time. In terms of scientific objectives, the basic objectives of the Kepler mission remain, the detection of exoplanets and asteroseismology. However, K2 opens new regions of the sky to the telescope. Thus, new populations of planets and stars can be studied. Moreover, new objectives have been proposed: observations of open clusters, star-forming regions, variable extragalactic sources and microlensing [40]. The attitude motion around the roll axis of the Kepler spacecraft induces a sawtooth photometric signature, and the attitude error is way higher than in the Kepler mission [41]. Those features were clearly observed during this master thesis when analysing the light curves and were the major source of difficulties.

2.2.4 Results and Impact

Kepler and K2 missions provided a massive amount of data that are still analysed today. After 9.6 years in space, 530506 stars have been observed with 3326 confirmed planet detec-

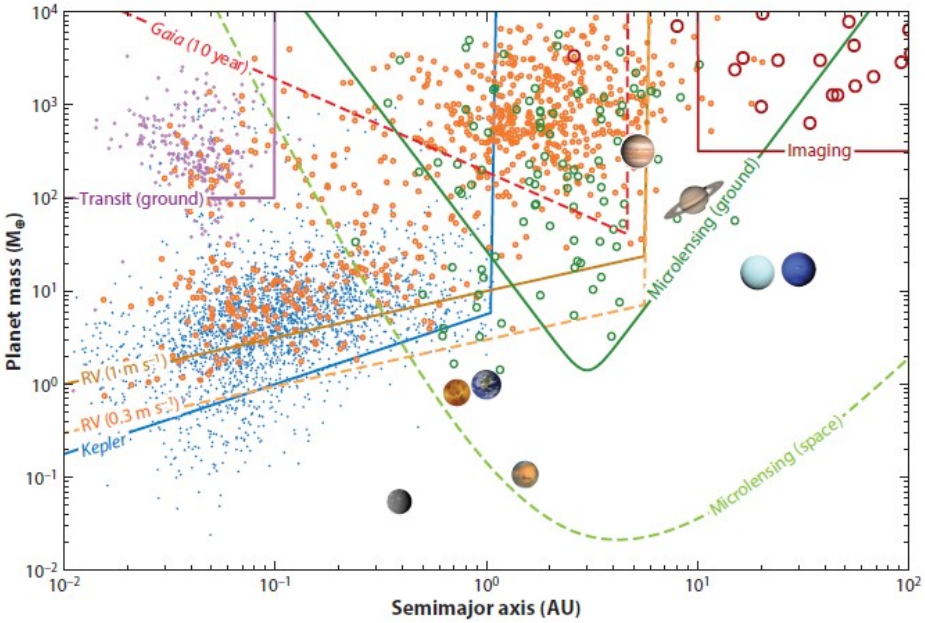


Figure 2.5: Distribution of detected exoplanets from different detection methods, Kepler is in blue [42].

tions and 3136 scientific papers published [34] [43]. The exoplanets range in size from slightly larger than the moon to planets over three times the size of Jupiter and have orbital periods ranging from hours to years. As of 2023, 183 Earth-size planets and 1444 super-Earths [34] has been detected. In Figure 2.5, one can see the distribution of the exoplanets discovered by Kepler with respect to the exoplanets detected by other methods.

In terms of exoplanets population statistics, Kepler shows that the radius distribution of the planets depends on their orbital period. Kepler's planets reside principally in multi-planetary systems. However, Kepler data are not sufficient to make statistics on Earth and super-Earth population in the habitable zones [42].

The impacts of the result of the Kepler telescope go beyond the scientific results. Kepler was one of the first missions with CoRoT to be dedicated to the search for exoplanets via a space telescope. The success of Kepler proved the efficiency of the technology and the possibility of a large space telescope for transit surveys. It has inspired other telescopes such as TESS, CHEOPS and JWST in the search for other worlds and had an important cultural impact.

2.3 K2 Correction Pipelines

Now the K2 mission has been introduced, in this part, the specificities of the K2 data will be detailed and the K2sff and EVEREST detrending pipelines will be introduced.

2.3.1 Specificities of K2 data

As explained in section 2.2, K2 such as Kepler could observe the sky with 2 exposure times: one of 60 seconds called the short cadence (SC) and one of 30 minutes called the long cadence (LC). K2 mission observation was divided into campaigns that corresponded to an observation period of about 75 days. However, due to the reduced pointing accuracy, K2's raw aperture photometry is about 3 to 4 time less precise than the Kepler mission and present strong instrumental features, notably a $\simeq$ 6 hours trend due to the spacecraft roll-cycles. They are characterised by triangular patterns in the light curves [44]. An example is given in Figure 2.6. Using directly raw K2 data for the search for exoplanets is thus impossible. We have to rely on pipelines with specific objectives to correct those errors.

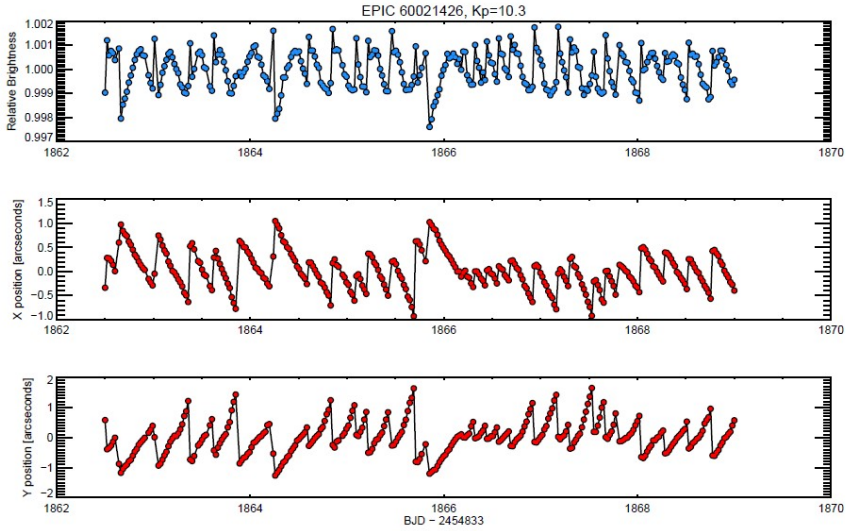


Figure 2.6: Top: Raw long cadence K2 light curve (with low-frequency variations removed). Middle: Horizontal centroid position versus time. Bottom: Vertical centroid position versus time [44].

2.3.2 K2sff Pipeline

K2sff is the first of the two pipelines that aim to correct the K2 pointing errors. It was first proposed by Vanderburg and Johnson in 2014 [44]. Prior to the K2 mission, techniques to reduce data from the Spitzer mission, which also encountered data artefacts caused by the motion of the spacecraft, were developed. However, Vanderburg and Johnson adapted and fine-tuned these methods to better suit the specificities of the K2 mission.

Therefore, thanks to the aperture photometry and image centroid position data extracted from K2 pixel-level data, the objective is to correct the photometry using the spacecraft's pointing information. The principle is relatively straightforward: after defining the mask and aperture of the target, the star's position as a function of time is estimated using a centre of flux calculation and a Gaussian centroid method. When the location of the star is too far away from its origin, Kepler thrusters fire back to correct it which is the reason for

artefacts. Then, the points with poor photometric performance are excluded. They then use a self-flat-fielding (SFF) approach to remove photometric variability. As the pointing jitter has a time scale of about 6 hours, they only isolate short-period variabilities with periods less than 24 hours. They separate the astrophysical variability components from the pointing jitter to avoid removing them. Then, data points are divided into bins. A linear interpolation is performed between the mean of each bin. The correction is applied by dividing the raw aperture photometry time series by the linear fit.

Figure 2.7 shows the comparison of raw and corrected data. For this specific target, the SFF correction reduces the scatter observed on 6-hour timescales by a factor of 5. In general, in the best-case scenarios, compared to Kepler, K2sff photometric precision is worse by a factor of about 1.3 to 2 instead of about 3 to 4 for raw data [44].

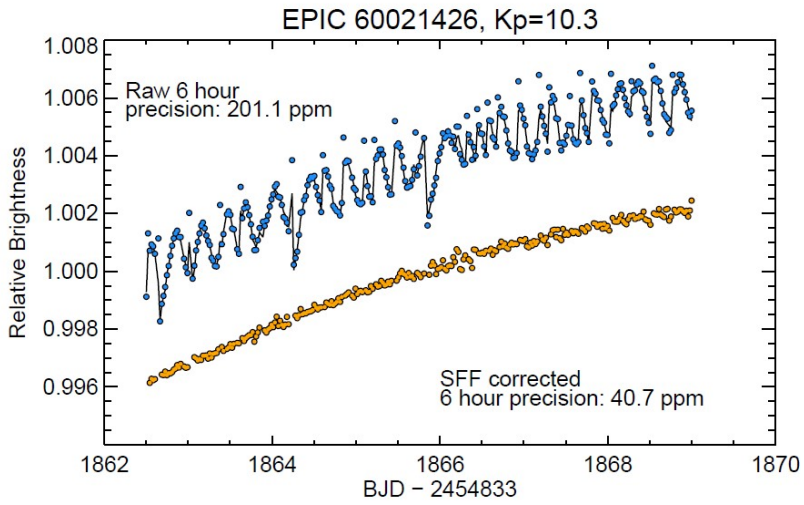


Figure 2.7: Comparison between the raw K2 data and SFF correction. Blue points correspond to raw K2 data vertically offset for clarity. The dark line between blue points is the SFF model. Orange points correspond to corrected data [44].

It should be noted that, in the case of exoplanet detection, if a transiting object has an orbital period on the time scale of the Kepler thruster fires, the SFF correction could suppress the signal. Detection of planets with a period of about an integer multiple of the time between thruster fires could also be difficult especially if the transit is during the thrust [44].

2.3.2.1 EVEREST Pipeline

In 2016, Luger et al. [45] proposed a new detrending pipeline for the K2 mission named EVEREST. In contrast to other pipelines like K2sff, which use numerical methods to identify and remove correlations between stellar position and intensity fluctuations, EVEREST adopts a different approach known as the Pixel Level Decorrelation (PLD) method. The advantage of PLD lies in its ability to correct noise induced by image motion without requiring

precise measurements of the star's location, thereby avoiding measurement uncertainties. Instead of performing fitting of the stellar position and correlation-solving steps done by K2sff, PLD directly operates on the intensities of each detector pixel. These intensities are normalised by the total flux within the chosen aperture and serve as the basis vectors for a linear least-square (LLS) fit to the aperture-summed flux. This normalisation is essential to remove astrophysical signals from the basis set. Moreover, PLD is an agnostic method for robust flat-fielding corrections, as it does not require information regarding the nature of intra-pixels variability or spacecraft jitter [45].

In their paper of 2016, Luger et al. estimated a better precision for EVEREST as compared to K2sff as the EVEREST light curves have $\simeq 20\%$ less scatter on average. However, limitations remained, notably in the cases of saturated stars or crowded aperture (a lot of stars in the same aperture) where the PLD can fail [45]. However, in their paper of 2018, [46], an updated version of EVEREST was presented where they corrected those issues and improved the model of faint and extremely variable stars. Globally, on average, their new version is 10 to 20 % more precise than the previous one. They claimed to obtain the most precise light curves at all magnitudes in the publicly available catalogues.

2.3.3 Data Availability

An important point in this thesis is to know which hot subdwarf stars have been observed during the K2 mission. In Van Grootel et al. 2021 [2], a list was made. It is presented in Appendix A. K2 observed 39 sdB/sdOB pulsators in both SC and LC and 2 in LC only and 78 non-pulsators sdB/sdOB and 10 non-pulsators sdO in SC and LC. 44 more hot subdwarfs were observed only in LC. We aim to look for planets which had been engulfed during the RGB phase. They should have short-period orbits around their hot subdwarfs. Hence, we are looking for planets with periods typically $\lesssim 10$ days. The duration of the transit will be short, between 20 minutes and 1 hour, and often of the order of the exposure time of the long cadence. Long cadence data will often only provide one or two points per transit and therefore its utility for our scientific objectives is not obvious. It may be eventually used as an additional test when a transit signal is found in SC.

Moreover, detrending pipelines often give a catalogue of already detrended lightcurves. In the K2sff catalogue, only LC-detrended data are provided. However, the transit search program SHERLOCK used in this thesis can perform directly a K2sff correction on the short cadence light curve making K2sff available in all cadences. The case of EVEREST is different. Some short cadence light curves are already provided in the EVEREST catalogue. From Appendix A, we account for 17 pulsators and 31 non-pulsators available, hence a total of 48 hot subdwarfs having light curves corrected by EVEREST. We tried to complete the catalogue but we had difficulties making EVEREST detrending work. For the injection-and-recovery tests and for the search of exoplanets, we will therefore focus on the available light curve in the EVEREST catalogue.

2.4 SHERLOCK Pipeline

2.4.1 General Presentation

The SHERLOCK pipeline (Searching for Hints of Exoplanets fRom Lightcurves Of spaCe-Based seeKers) [47][48] is an open-access tool recently developed by F.J. Pozuelos and M. Dévora-Pajares to detect transit from space-based observatories. SHERLOCK is made of six modules which allow to :

- Download and prepare the light curve using the LightKurve pipeline for K2, Kepler and Tess or the EVEREST pipeline for EVEREST K2 corrected.
- Search for exoplanetary candidates
- Perform the vetting of promising signals
- Statistically validate the signal
- Fit the signal to refine the ephemerides
- Obtain observational windows for ground-based observations.

In the next part of this section, I will detail each of these modules and how to use them.

2.4.2 Preparation

In the preparation stage of the analysis, SHERLOCK needs to understand the sources the user asked for to find the proper data. Those data correspond to the information on the targeted star and the photometric data in time series format. Using the LightKurve Eleanor and EVEREST packages, thanks to the identification numbers of the stars (EPIC for the K2 mission), the cadence and the campaign, data are downloaded from the MAST which is a database that allows to easily find astronomical data, publications and images [49]. In general, there are two different types of flux that can be used: the SAP and the PDCSAP. The SAP flux for Simple Aperture Photometry is computed by the sum of all the pixels that fall in a pre-defined aperture. By choosing this aperture well, nearby contaminants can be avoided and increase the signal-to-noise ratio. The PDCSAP for Pre-search Data Conditioning Simple APerture is subject to more treatment and is specially created for exoplanet search. Hence, by default, SHERLOCK uses the PDCSAP [50].

Furthermore, SHERLOCK can also treat some trends and systematics. It can mask automatically noisy regions, thus regions with higher root-mean-square (RMS) values of the flux. It can correct strong variabilities due to fast rotators or pulsations and short-trend due to instrumental drift. It is also possible to manually hide a part of the curve where the data are too scattered [50].

Alternatively, one can give SHERLCOK a `.csv` file containing the time series data flux to analyse. In the case of EVEREST, it can be useful. Indeed, even though EVEREST light curves can be accessed directly with SHERLOCK, we will want, before the analysis,

to apply a specific flag filter to the data. In our study, the "standard" filter is used. This configuration has been tested against other possibilities and was found to be more efficient for the search for exoplanets. Table 2.2 lists the flags that will be removed, the complete list of flags is presented in Appendix B.

0x00000001	Attitude tweak
0x00000002	Safe mode
0x00000004	Coarse point
0x00000008	Earth point
0x00000020	RW Desaturation event
0x00000040	Argabrightening
0x00000100	Manual exclude
0x00000800	Impulsive outlier
0x00001000	Argabrightening
0x00004000	Detector anomaly
0x00008000	No fine point
0x00010000	No data
0x00080000	Possible thruster firing
0x00100000	Thruster firing
0x00800000	Data point is a NaN
0x01000000	Determined to be an outlier

Table 2.2: K2 flag list to remove

2.4.3 Search for Candidates

2.4.3.1 Theoretical Functions

The search for candidate modules is divided into two parts: the detrending of the light curve and the detection of transits.

Even though, in the preparation phase, the light curve has been cleaned of some specific trends, as explained in section 2.4.2, this cleaning is not perfect. To remove as many unwanted signals as possible, SHERLOCK will perform a certain number of detrending chosen by the user, often 12, each having a different window size. To do so, SHERLOCK will use the Wötan package [51]. If the size of the detrending windows is too short, it would remove, in addition to noises, transit signals themselves. As transit durations are unknown, using different window sizes allows a maximal coverage of all transit durations. A lot of detrending method exists. The tests made by Hippke et al. 2019 [51], show that the bi-weight method is the most optimal choice in the majority of cases, hence this is the one I will use in this work.

After performing detrending, SHERLOCK will be looking for transit using the Transit Least Square (TLS) package [52] in all the detrended light curves jointly with the original one. Over the years, various techniques for searching for transits have been created. The most standard approach is the Box Least Squares (BLS) method. This method operates on

the assumption that the transit light curve can be approximated as a boxcar function, with a normalised average flux outside the transit region of zero and a constant depth throughout the transit. However, when dealing with Earth-size planets, this method's effectiveness can, in some cases, diminish because the transit depths are comparable to both instrumental and stellar noises. Moreover, the use of the boxcar function simplification introduces systematic noise, further affecting its performance. On the other hand, the TLS method employs a transit-like function instead of a boxcar, enabling a more realistic consideration of the ingress and egress shapes of the transit, as well as the impact of stellar limb darkening. One of the significant advantages of TLS is that its template is optimised for detecting small planets. This approach is well-justified because adapting the model for large planets would likely result in a higher risk of missing small planets than the opposite. The algorithm employed by TLS involves phase-folding the data over a range of trial periods, transit epochs, and transit durations. It then computes the χ^2 statistic, which quantifies the differences between the phase-folded curve based on the respective transit model and the observed data points, aiming to identify the minimum χ^2 as the best-fitting solution. By using these techniques, TLS provides a more accurate and effective way to detect transits, particularly for small planets [52].

When the analysis is completed on all the detrend curves, SHERLOCK chooses the best signal in terms of SNR and SDE in all the curves. The SNR corresponds to the signal-to-noise ratio and so measures how much the signal is bigger than the surrounding noise. The SDE stands for Signal Detection Efficiency and is computed as $\frac{1-<SR(P)>}{\sigma(SR(P))}$ for a given tested period P with SR defined as the signal residue and σ the standard deviation. The SDE measures the statistical significance of this period P as compared to the mean significance of all other periods. The transit signal found is then hidden in the light curves and a new run of analysis begins. SHERLOCK is an iterative process and will stop either when no more good signal is found or when the maximal number of runs established by the user is reached.

2.4.3.2 Practical Use

Now that the theoretical functions of SHERLOCK transit search have been established, let us explain how to use this SHERLOCK module.

The SHERLOCK pipeline can directly be launched on Python by calling the package. However, the most practical and simple way to call SHERLOCK is through `yaml` file. This file contains all the parameters to be injected into the analysis and will simply be given to SHERLOCK by the command :

python3 -m sherlockpipe --properties properties.yaml

The main parameters in the `properties.yaml` files are:

- The star identification: In the case of K2, it would be the EPIC number. It will allow SHERLOCK to identify the mission and the star to extract the star information and the lightcurve

- SECTORS: In the case of K2, correspond to an array which defines from which cam-

paigns the curves will be extracted. In the case that all sector has to be considered, one can only write 'all'.

- EXPOSURE_TIME: Exposure time of the curve we want to extract. For K2, for the short cadence, one writes 60 and for the long cadence 1800.
- AUTHOR: Source of the light curve, for K2 it can be either K2 or EVEREST. By default, SHERLOCK will search for the K2 light curve and apply the K2sff correction to it.
- FILE: Name of a .csv file containing a light curve to use instead of the one from the MAST
- DETREND_METHOD: The detrend method to be applied on the curve. As explained in section 2.4.3.1, the biweight method is the best in the majority of cases and will be selected by default. However, others are available, like the cosine.
- DETRENDS_NUMBER: Number of detrend to execute. In general, we select 12.
- DETREND_CORES: Number of computer CPU cores working on detrending. The more cores used, the faster the detrending phase.
- MAX_RUNS: The maximal number of loop SHERLOCK will run to search for a signal. Once, this number is reached, even though more good signals may appear, the program stops. In general, we select 4 because it is the best trade-off between the quality of the signal detected and the simulation time.
- SNR_MIN: The minimal SNR a signal must have to be considered as good.
- SDE_MIN: The minimal SDE a signal must have to be considered as good.
- CPU_CORES: Number of cores working on the search of transit.
- FIT_METHOD: By default, the method for the fit is the TLS. However in some cases, notably grazing transit or disintegrated objects, it is no longer adapted and others have to be selected.
- INITIAL_SMOOTH_ENABLED: If true, a Savitzky-Golay (Sav Gol) filter will be applied to the data. The Sav Gol allow smoothening of the data set, therefore, increasing the data accuracy without distorting the signal tendency. In the case of search of transit, it will increase the SNR of low-depth signals. However, this filter alters the transit shape, thus it can not be used when the intention is to conduct a proper transit fitting to derive planet parameters.
- INITIAL_HIGH_RMS_MASK: If true, part of a curve presenting high scatter features, and therefore high RMS value, will be hidden.
- INITIAL_HIGH_RMS_THRESHOLD: Multiplier factor so that a region is selected in the high RMS mask.

- INITIAL_HIGH_RMS_BIN_HOUR: Duration in hours in which the RMS will be computed.
- INITIAL_MASK: Time intervals for which part of the curve will be hidden. It works in conflict with the high RMS mask. If an initial mask is selected, the high RMS mask will not be considered.
- AUTO_DETREND_ENABLED: If true, an initial detrend on the original light curve will be applied in order to remove strong periodic features which might influence the quality of SHERLOCK analysis.
- SIMPLE_OSCILLATIONS_REDUCTION: Remove oscillation via a pre-whitening method. It is useful for fast rotators and pulsating stars. It works well alongside the auto-detrend.
- PERIOD_MIN/PERIOD_MAX: Both commands define the limit on the period of transit search. All transit periods outside the defined interval will not be considered.
- BEST_SIGNAL_ALGORITHM: Specifies the algorithm to elect which signal is the best in each run. As explained in section 2.4.3.1, SHERLOCK choose the best signal in terms of SNR and SDE. By default, SHERLOCK will use the quorum algorithm which will also take into account the number of detrended curves which detect the same signal.
- QUORUM_STRENGTH: It corresponds to the scale factor for the quorum vote.

However, before running the full search, it can be useful to launch only the preparation stage with the command **python3 -m sherlockpipe –properties properties.yaml –explore** in order to verify the adequacy of the parameters.

2.4.3.3 Interpretations of the Results

Before the search, the preparation would have generated a directory that contains :

- A directory containing the plot of the detrended light curves
- A directory with the field of view to inspect the objects around the targets. It is not applicable for K2 because this information is absent.
- A directory with the RMS mask plots, if this option has been activated
- Autocorrelation plots of before and after light curve corrections
- Periodogram plots of before and after light curve corrections
- `.csv` files of all light curves, detrended or not.
- A `.csv` file containing the star information
- A copy of the properties.yaml file

Once the stage of the transit search is over, SHERLOCK will add to this directory :

- Directories for each run containing the plot of the different signals found in their respective detrended lightcurve.
- Autocorrelation plots of before and after light curve corrections.
- A `.csv` file containing information about the found signals.
- A `.log` file containing information about the found signals.
- A report file containing all the information and steps of the search.
- A file containing the transit statistics.

To better understand how the result should be analysed, an example will be presented with the target EPIC 210490365 which is an M-type star. From the NASA exoplanet archive, we know this star has an exoplanet with these characteristics :

Period (days)	3.48456408
Depth (%)	1.155
Duration (hours)	0.7637

Table 2.3: K2-25b main parameters [36].

```yaml
TARGETS:

  'EPIC 210490365':
     INITIAL_SMOOTH_ENABLED: True
     EXPOSURE_TIME: 1800
     AUTHOR: EVEREST
     DETREND_METHOD: 'biweight'
     DETRENDS_NUMBER: 12
     DETREND_CORES: 5
     MAX_RUNS: 4
     SNR_MIN: 6
     SDE_MIN: 7
     CPU_CORES: 5
     OVERSAMPLING: 3

     FILE: 'data_210490365_lc_None_norm_standard.csv'
     BEST_SIGNAL_ALGORITHM: 'quorum'
     QUORUM_STRENGTH: 1.2
```

Figure 2.8: `yaml` file for the example of EPIC 210490365.

The `.yaml` file to analyse this target is provided in Figure 2.8. The analysis will be made

in the long cadence as no data exists in the short cadence for this target. In general, no confirmed planets with a star which was observed in short cadence by K2 were found.

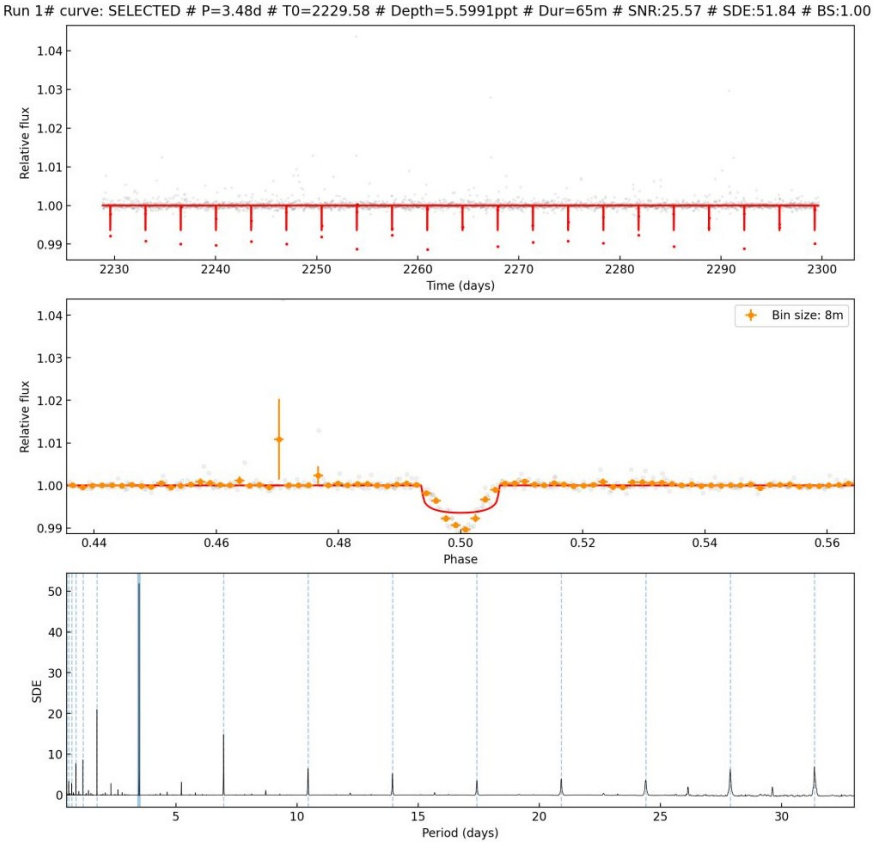


Figure 2.9: Selected best transit signal for EPIC 210490365. Above: the flux in all selected campaigns as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

Once all the runs are completed, one must look in the candidate and the report files. In the candidate file, all the candidates for transit signal characteristics are shown. One must look if the signal is good enough in terms of SNR and SDE. Generally, good values of SNR and SDE are respectively above 7 and 9. In our case, a signal with a period of 3.4843 days, a duration of 1.078 hours and very high SNR and SDE of respectively 25.57 and 51.84 has been found. In the report files, the details of the results in all the detrended curves are displayed. The more the number of detrended curves that detect the signal, the more the signal has probabilities to be an actual transit. In our case, over the 12 performed detrends, 9 detected the signal, which correspond to an interesting signal. The next step consists of looking at the transit plots. They are provided in Figure 2.9. The phase transit curve allows us to look at the shape of the transit and see if irregularities are present. Here we see a clear transit shape. Two outliers points are present at a phase of about 0.47 but are far enough from the transit. The transit model does not correspond totally to the light curve. It could correspond to a grazing transit and a grazing model could refine the results. The light

curves plot with all the suggested transits can provide information about the consistency of the transit shape and depth, if the transits are too different with respect to each other it can be a false signal. Here the depth is very consistent across the light curve. The SDE plot shows us the harmonics on the transits period. A good signal would display visible harmonics, which is the case here.

The signal seems very promising and now we can go to the next step of the analysis, the vetting.

2.4.4 Vetting

Vetting is a comprehensive approach designed to examine in more depth a signal and determine whether the event in question might have resulted from alternative factors. Throughout this process, SHERLOCK conducts an analysis of various aspects, including the peculiarities in odd-even transit shapes, the signal-to-noise ratio (SNR) during different cadences, SNRs related to signal harmonics and subharmonics, the offset of the transit source from the transit pixel files (TPF) (using both differential imaging and a Sherlock algorithm based on BLS phase folding of per-pixel TPF data), shifts in centroids and motion data and an assessment of single transits [53].

To perform the vetting, the command **python3 -m sherlockpipe.vet --candidate N** with N the candidate number has to be used.

At the end of the analysis, SHERLOCK will create a validation report containing a recapitulative table and different graphs. The important features will be explained using the example presented in section 2.4.3.3.

First, let us explain the table on the first page of the report in Figure 2.10 [53]:

- long_SNR: SNR of the transit model in the folded original long-cadence curve. Passes if > 3. If this star were observed in short-cadence, an equivalent line would be added for short-cadence data.
- snr_p_t0: SNR of the transit model in the Period folded SHERLOCK selected processed curve. Represents the main transit events. Passes if > 3.
- snr_p_2t0: SNR of the transit model in the Period folded SHERLOCK selected processed curve centred on T0 + Period / 2. Represents the occultation transit events. Passes if < 3.
- snr_2p_t0: SNR of the transit model in the 2 * Period folded SHERLOCK selected processed curve centered on T0. Represents the odd transits. Passes if > 3.
- snr_2p_2t0: SNR of the transit model in the 2 * Period folded SHERLOCK selected processed curve centered on T0 + Period. Represents the even transits. Passes if > 3.
- snr_p2_t0: SNR of the transit model in the Period / 2 folded SHERLOCK selected processed curve centred on T0, and the original signal is masked. Represents the

occultation transit events and reports the same values as snr_p_2t0. Passes if < 3.

- snr_p2_t02: SNR of the transit model in the Period / 2 folded SHERLOCK selected processed curve centred on T0 + Period / 2, and the original signal being masked. Passes if < 3.

Here every test has been passed which makes this signal very interesting. A good signal does not always fully pass this test, especially for shallow transit that can be hidden in the original curve for example. On the contrary, a signal that passes all the tests is not always good, the vetting must be seen as an additional clue on the quality of a signal.

Metric	Value	Passed
long_snr	21.038	True
snr_p_t0	9.144	True
snr_p_2t0	0.001	True
snr_2p_t0	35.31	True
snr_2p_2t0	3.266	True
snr_p2_t0	0.001	True
snr_p2_t02	0.001	True
snr_p_score	0.0	True
snr_2p_score	32.045	True
snr_p2_score	0.0	True

Figure 2.10: Screenshot of the table presented on the first page of the validation report of SHERLOCK for EPIC 210490365.

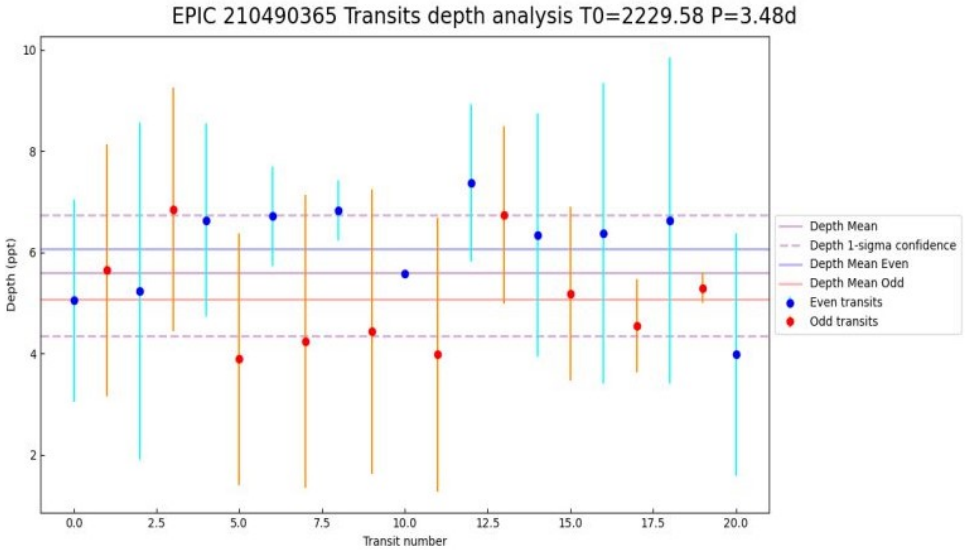


Figure 2.11: EPIC 210490365 transit depth analysis.

One should also note that the vetting for the K2 mission is less precise and conclusive

than its Kepler and TESS counterparts. For TESS and Kepler, additional tests consist of the analysis of the aperture and the location of the star, the signal and the other sources and therefore adding additional refining information in the vetting.

Additionally to the table, figures are plotted in the vetting report. The first figures are the phase-folded original light curves in short- and long-cadence which provide information on the visibility of the transit in the original light curves. The next figure, represented in Figure 2.11, shows the depth values of each even and odd transit. The depth must be constant for all individual transit. For instance, when a small number of transits exhibit significant depth variations, the resulting detected transit signal could be inaccurately influenced by these variations, potentially leading to misinterpretations of the signal's origins. When the transit depth alternates between even and odd transits, it might indicate the presence of an eclipsing binary system, with half of the transits corresponding to the eclipse and the remaining half to the occultation.

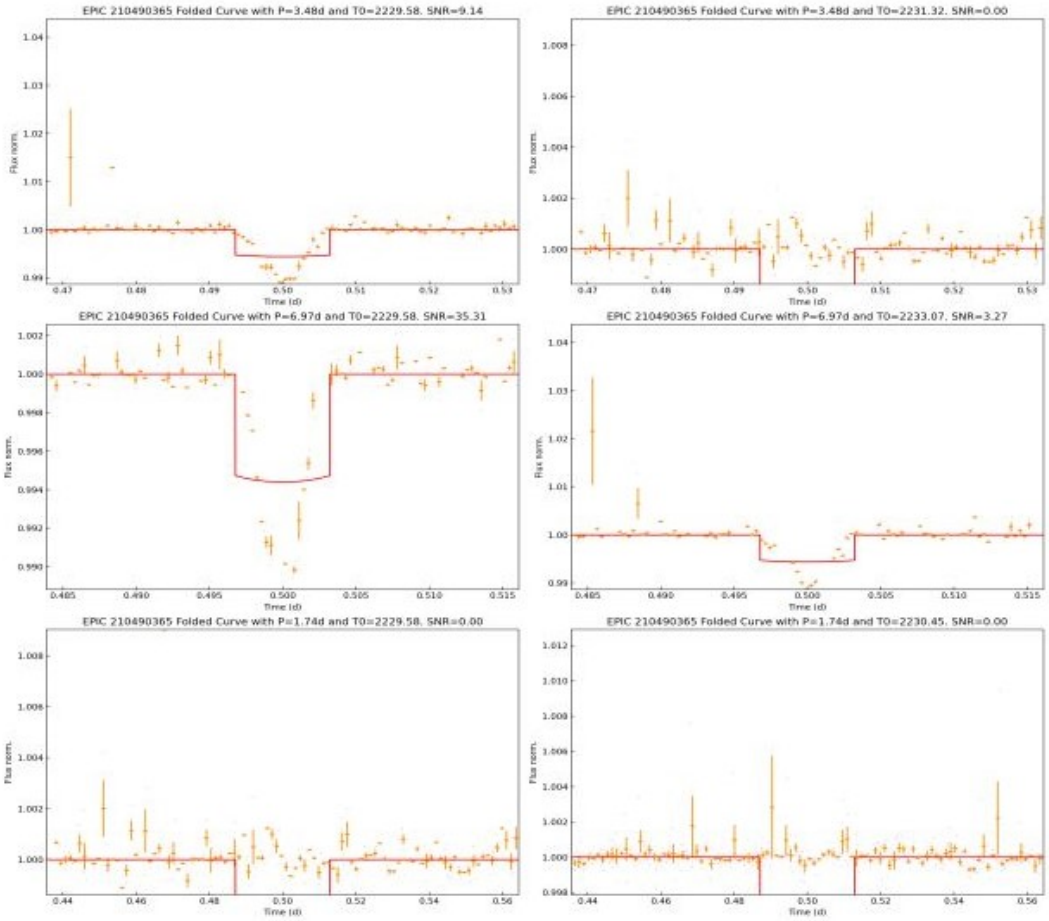


Figure 2.12: Above, the candidate folded at its found period for the found epoch and epoch + P/2. Middle, the candidate folded at its harmonic for the found epoch and epoch + P. Bottom, the candidate folded at its subharmonic for the found epoch and epoch + P/2, where the candidate has been masked.

In the case of our example, all the transits seem to have depth inside the 1 sigma confidence and not much difference between the even and odd transit is visible which is another clue of the quality of this signal.

The third series of figures are shown in Figure 2.12. If a transit at a subharmonic or the epoch + P/2 is visible, it would maybe be because SHERLOCK detected an upper harmonic to the signal.

2.4.5 Validation

After a promising vetting is completed, SHERLOCK can perform a statistical validation using the TRICERATOPS (Tool for Rating Interesting Candidate Exoplanets and Reliability Analysis of Transits Originating from Proximate Stars) [54]. The main concept behind this module is to simulate various astrophysical scenarios that may create transits-like signals and compute the probability of each scenario using the primary transit of the candidate, the available information about the host star and its surrounding neighbours, as well as our current knowledge of planet occurrence rates and star multiplicity. The final product will be the location of the given signal in the FPP - NFPP (Nearby False-Positive Probability) plane with validation if FPP$<$0.015 and NFPP$<$0.001 [54].

TRICERATOPS uses a specific approach to calculate these probabilities. It initiates by generating a circle with a radius of 10 pixels and extracts the parameters of all stars located within this circle. This aperture file is extracted from SHERLOCK and employed in TRICERATOPS for further analysis. Then, TRICERATOPS determines the contribution of each neighbouring star's flux to the aperture. The algorithm identifies stars bright enough to produce a transit-like signal, taking into account the depth of the candidate transit. TRICERATOPS estimates the likelihood of each transit-producing scenario thanks to the candidate's primary transit and light curve models of transiting planets and eclipsing binaries. Subsequently, it computes the scenarios' probabilities and uses them to calculate the False-Positive Probability (FPP) and Nearby False-Positive Probability (NFPP).

This validation process is launched by the command **python3 -m sherlockpipe.validate --candidate N** with N as the candidate number.

Unfortunately, for the K2 data used in this master's thesis, the module discussed earlier is not accessible for now. TRICERATOPS was exclusively designed for use with TESS and Kepler data and cannot be applied to K2 data for the moment.

2.4.6 Fitting

After having completed the validation phase, SHERLOCK can refine the planet parameters and ephemerides using the Allesfitter package [55]. The results from the search module are directly injected as priors into Allesfitter which then performs a Bayesian model using the Nested Sampling which is an inference algorithm to directly estimate the Bayesian evidence by sampling from the prior subject to evolving constraints on the likelihood. In other words, it tries to find the degree of belief in the transit with respect to models. The parameters

used during the fitting are :

- The ratio of the planet to host star radius (R_p/R_*).
- The sum of the stellar and companion radii divided by the semimajor axis, $(R_* + R_p)/a$.
- The cosine of the orbital inclination ($\cos(i_p)$).
- The epoch or transit midtime in days (T_0).
- The orbital period of the planet in days (P_p).
- The transformed limb darkening coefficients following a quadratic law (q_1 and q_2).
- The natural logarithm of the white noise ($\log(\sigma_w)$).

This aspect of the analysis is essential. Firstly, it enables the verification of whether the planetary parameters obtained from the fitting process are realistic. Secondly, and most importantly, it provides an accurate computation of the ephemeris, which is essential for planning future observations.

Given that the K2 mission commenced in 2013, there can be a time gap of 10 years between the measurements and the date of the analysis. During this interval, the errors associated with the measurements increase with time as $\Delta T = \sqrt{\Delta T_0^2 + (nP)^2}$ with ΔT_0 the error on the epoch, n the number of events since the epoch and P the orbital period. Consequently, even a small error in their estimation can grow to a significant extent, which poses challenges in creating ground-based observation plans and, in some cases, might render the observations impossible to perform.

2.4.7 Observation Planning

After a successful vetting and statistical validation, one would like to trigger a follow-up campaign to firmly confirm the event in the target star. This can be done by employing the observation module of SHERLOCK, which computes the observational windows for a selected sample of observatories the user provides. To do so, one must go to the fit result folder, and create a `.csv` file with all the information on our observatories: names, time zones, latitudes, longitudes, and altitudes. Then, the observation planning can be launched using the command **python3 -m sherlockpipe.plan --observatories Observatories.csv**. SHERLOCK will create a plan folder in which two files can be found. The first one is a `.csv` containing the different important information for the observation: the time of the start and the end of the observation, the ingress and egress time, the middle time of the transit and its potential time deviations, the twilight and the dawn time, the moon phase and its angular distance to the targets. The second file is a `.pdf` containing all the elements required to schedule an observation for each observatory. It displays a table representing the different event times, their respective error, and the Moon illumination and angular distance from our candidate[53].

In the case of K2 data, the fitting and the observation modules still have to be well

implemented.

2.5 MATRIX ToolKit

The MATRIX ToolKit (ToolKit for Multi-phAse Transits Recovery from Injected eXoplanets) is a pipeline for exoplanet transits injection-and-recovery tests on any light curve of a space-based mission. It provides detection limits, enabling users to understand the constraints of the data set used in the planetary search. It provides valuable information about the detectable range of planets concerning their periods and radii. Additionally, MATRIX informs users about the specific planets that SHERLOCK would be unable to detect due to limitations posed by the photometric quality of the data [56]. The user defines a range of planetary radii and periods which are analysed for specific phase values, meaning different T_0 values. For simplicity, all the orbits are supposed circular and the impact parameter of the transit is null. For each Period-Radius-T_0 value, a light curve will be generated and then will be detrended using a bi-weight method with a defined window-size. Then, using the TLS method, the transit search is performed. A synthetic transit is considered successfully recovered when its epoch matches the injected epoch with an accuracy of 1 hour, and its period falls within a range of 5% of the injected period. Such as SHERLOCK, multiple runs are performed and the search is over either when the transit is detected, another signal with the same period but a different epoch is detected or the limit number of runs is reached. It should be noted that since the planets are injected into the PDC-SAP light curve, the transits are not affected by the PDC-SAP correction, therefore the MATRIX results are an optimistic estimation of the detection limits [57].

MATRIX has been developed as a supplement to SHERLOCK and their operation is approximately the same. Therefore, the program also uses a `yaml` file to define its parameters. Most parameters are called like in SHERLOCK's `yaml` file. The specificities reside on the test parameters:

- PERIOD_GRID: Defines the array of periods to test.
- MIN_PERIOD: If PERIOD_GRID is not set, the minimum period to be used to generate the grid.
- MAX_PERIOD: If PERIOD_GRID is not set, the maximum period to be used to generate the grid.
- STEPS_PERIOD: If PERIOD_GRID is not set, the number of items to be contained in the period grid.
- RADIUS_GRID: Defines the array of radius to test.
- MIN_RADIUS: If RADIUS_GRID is not set, the minimum period to be used to generate the grid.
- MAX_RADIUS: If RADIUS_GRID is not set, the maximum period to be used to generate the grid.

- STEPS_RADIUS: If RADIUS_GRID is not set, the number of items to be contained in the radius grid.
- PHASES: The number of phases (epochs) to be tested for each period and radius in your scenarios. It should be noted increasing this value will increase the computational cost significantly because every scenario will run the number of phases entered times.
- DETREND_WS: Define the windows-size of the detrending.

The results of the analysis will be a colour map with the abscissa axis corresponding to the period grid and the ordinate to the radius grid and the colour corresponding to the recovery rate in the period-radius pixel.

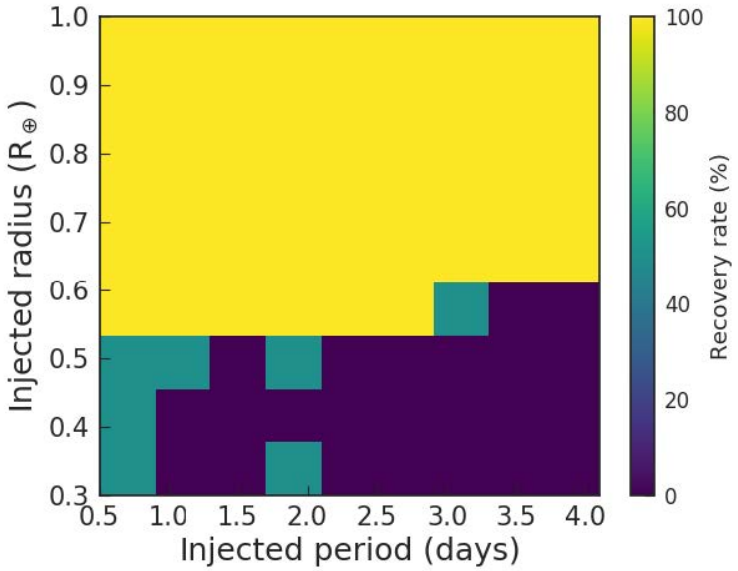


Figure 2.13: Injection-and-recovery tests for EPIC 206535752 (Kp=13.99, G=14.10), observed during Campaign 3 of K2 (81 days). Injected transits of planets have 0.3–1.0 $R_\oplus$ (steps of 0.1 $R_\oplus$) with 0.5–4.1 d (steps of 0.2 d) orbital periods [2].

Chapter 3

Injection-and-Recovery Tests on K2 Correction Pipelines

In this chapter, I will present the injection-and-recovery tests performed during this master thesis with the MATRIX ToolKit pipeline. Firstly, the objectives of the tests and the method applied to conduct them will be described. Secondly, the results will be presented. And finally, we will present the conclusions and the lessons learnt for the search of transiting planets in K2 data.

3.1 Objectives and Testing Method

3.1.1 Objectives of the Tests

At the beginning of this master's thesis, we directly wanted to start the search for exoplanets around hot subdwarfs. However, we quickly encounter difficulties.

The first one was on data availability. As explained in section 2.3.2.1 EVEREST is often considered the best pipeline in terms of K2 light curves correction quality. Nevertheless, we remarked that not all of our targets were available in the catalogue. We tried to use the EVEREST pipeline in order to correct the remaining targets' light curves but we did not succeed. On the other hand, we also did not know how much transit signal detection quality we would lose if we used the K2sff correction directly integrated into SHERLOCK.

The second one was directly on the transit signal detection quality. The transit search runs on EVEREST light curves provided unsatisfactory results. On one hand, high SNR and SDE signals were found but on the other hand, they were too many and did not have the expected transit shapes and properties (duration, constant depth over the several transits spotted, etc.). Hence, the basic metrics to evaluate a signal in SHERLOCK gave positive evaluations to bad signals. For example, in Table 3.1, the signals found for EPIC 201531672 are shown and in Figure 3.1, the shape is displayed.

Therefore, injection-and-recovery tests were conducted to provide an assessment of the quality of EVEREST and K2sff to search for planetary transits around hot subdwarfs

It may be important to note that injection-and-recovery tests have already been performed on K2 data by Van Grootel et al. 2021 [2]. The results of those tests are presented in Figure 2.13. However, those tests have been made in a smaller period-radius region, for one single target and using only K2sff. They are, therefore, not enough to answer our questions.

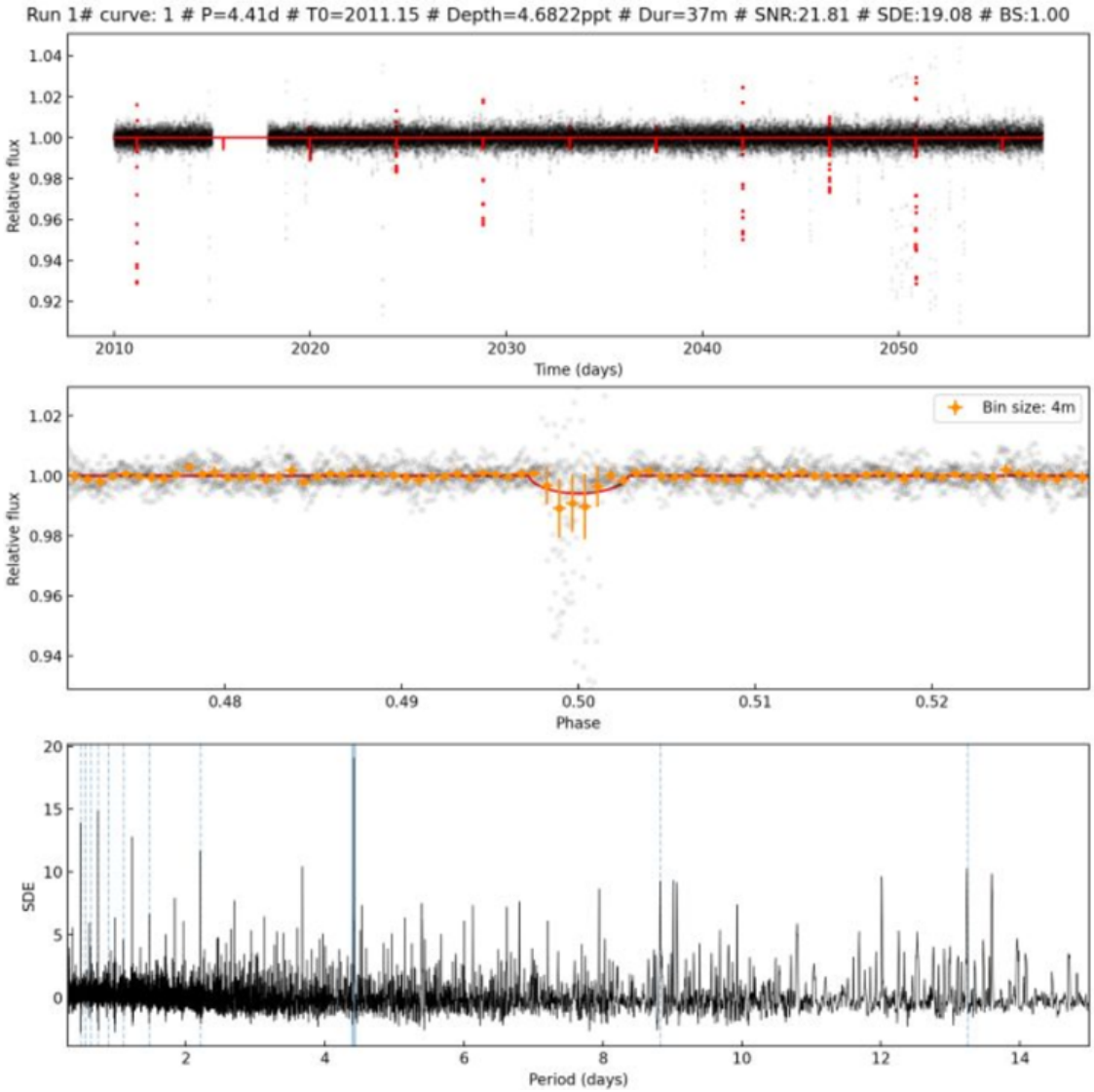


Figure 3.1: First signal found for EPIC 201531672. Above: the flux as a function of time with in red the fitted transit. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue. Here, the detection of the transit seems to be due to systematics and residuals.

Period (days)	Duration (mn)	Depth	SNR	SDE	Nbr of detection
4.41	36.75	4.68	21.81	19.08	12
0.74	21.72	2.05	19.09	12.41	8
9.81	48.87	6.49	19.92	12.26	9
6.13	46.36	4.3	21.79	15.52	10

Table 3.1: Example of results yielded by SHERLOCK using the SC light curve generated by EVEREST for EPIC 201531672.

3.1.2 Test Procedure

We performed tests that aimed to conduct a comparative analysis between EVEREST and K2sff. We selected four targets with different magnitudes, as outlined in Table 3.2. Before the tests, we examined the quality of the light curves.

EPIC	**206535752**	**211727748**	**206240954**	**201531672**
Kp	14	15	16.3	16.9
Visual Quality	Bad	Average	Average	Good

Table 3.2: Targets chosen for the injection-and-recovery tests.

For each of these targets, in each pipeline, we will perform an initial test by testing period between 0.5 and 7 days with a step of 0.5 days and using a radius grid of [0.3, 0.5, .7, 0.9, 1.1, 1.3, 1.5, 1.9, 2.1, 2.4, 2.7, 3] $R_\oplus$. To enhance the accuracy of the detection efficiency estimation, each combination of radius and period will be executed with ten different phases, which makes a total of 1950 scenarios. Therefore, we will employ various epoch values for the transit, which will probe several regions of the light curve and help avoid biases if the transits happen to land in noisy areas. It should be noted that for a very small period and very large radius, the planet will reach its Roche limit and therefore the orbit will not be possible. However, even though MATRIX take this into account, it will mark it as detected in the injection-and-recovery figure.

The results are analysed and the radius interval corresponding to the moment when the efficiency gradually changes from "no transits are detected" to "all transits are detected" is defined. Another test is then performed to refine the result in this region. It will be executed with a period step of 0.2d between 0.6 and 7 days and a radius step of $0.1R_\oplus$ in the radius interval for 10 different phases.

The parameters to inject in MATRIX are straightforward. The selected targets are single non-pulsators and therefore do not require an oscillation reduction method. Preliminary tests concluded that using the high RMS mask is not recommended since it will surely mask the high-depth transit. A detrend-window size of 0.5 is chosen. To increase the detection efficiency and to reduce the computing time, an initial mask is applied to only select the best-looking 30 to 40 days of the light curve.

3.2 Results

3.2.1 Results for EPIC 206535752, $K_p = 14$

3.2.1.1 EVEREST Pipeline Results

From the first series of test results of EPIC 206535752 presented in Figure 3.2, one can easily conclude a good transit detection capability. We can detect all injected transit with a radius superior to $0.5R_\oplus$. In Figure 3.3, the second test results are presented. A more continuous and precise limitation can be appreciated. Transits with shorter periods ($\lesssim 1.4$ days) can be fully detected when the planetary radius is larger than $0.4R_\oplus$. However, the transition is not strict and detection of transit with a planetary radius less than $0.4R_\oplus$ is feasible for small orbital periods with sometimes recovery rates better than 50%. For transits with $P \gtrsim 1.4$ days, the recovery rate of planets with $R \lesssim 0.5R_\oplus$ decreases gradually with the period. When the period reaches values above 6 days, the recovery rate of planets with $R=0.5R_\oplus$ is no more than 100% but still above 80%. Those results gave minimum values of SNR and SDE for detection of respectively 5.01 (at $P = 0.5$d and $R = 0.3R_\oplus$) and 7.78 (at $P = 4$d and $R = 0.4R_\oplus$). Those values increase quickly and can reach respectively 487 and 69, both at $P = 1.5$d and $R = 3R_\oplus$. Small planets are not often detected in the first run and need several search loops. Globally, we can conclude here that the EVEREST light curve of EPIC 206535752 is very qualitative. This could be due to the relatively low magnitude of the star allowing to have good SNR and SDE for shallow transits.

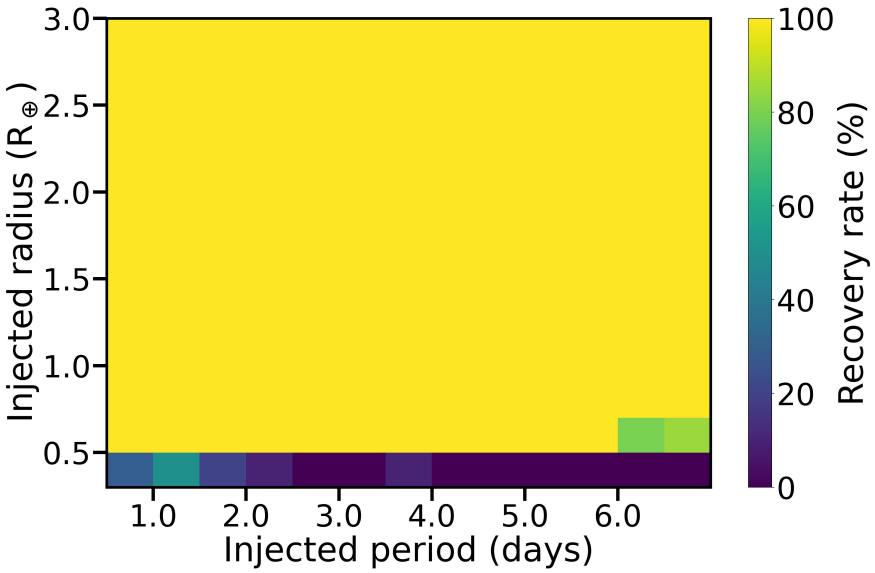


Figure 3.2: Injection-and-recovery test for EPIC 206535752's light curve corrected by EVEREST. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_\oplus$ between 0.3 and $0.9R_\oplus$ and $0.3R_\oplus$ between 0.9 and $3R_\oplus$. Each period-radius combination is tested for 10 phases.

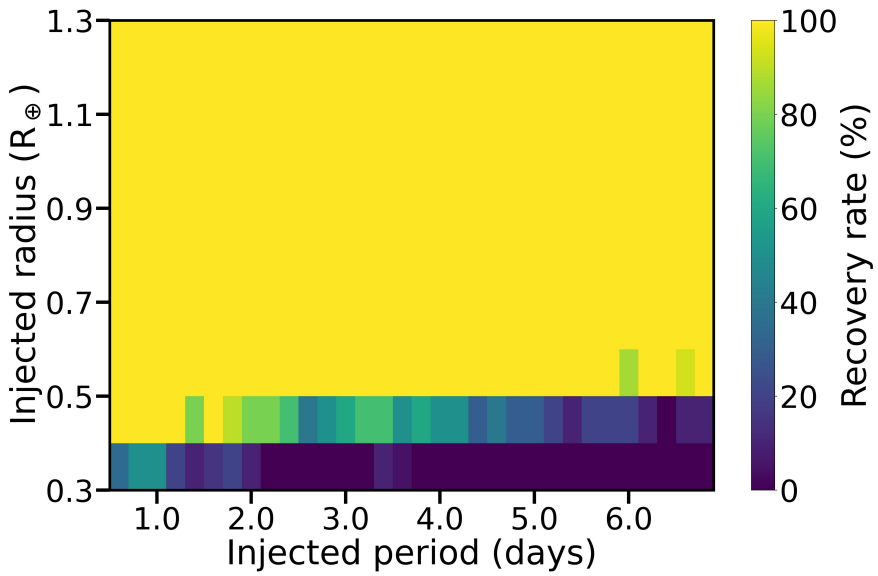


Figure 3.3: Injection-and-recovery test for EPIC 206535752's light curve corrected by EVEREST. The pixel size in the period dimension is 0.2 days between 0.6 and 7 days in the radius dimension is $0.1R_\oplus$ between 0.3 and $1.3R_\oplus$. Each period-radius combination is tested for 10 phases.

3.2.1.2 K2sff Pipeline Results

The first results with K2sff correction presented in Figure 3.4 are interesting. Full detection of transits is limited to planetary radii of approximately $1.2R_\oplus$, whereas with the EVEREST method, this level of detection was achievable at a smaller radius of $0.5R\oplus$. We observe, therefore, a drop in the transit detection capability as compared to EVEREST. Between 0.9 and $1.2R_\oplus$, the recovery rate decreases gradually when the period increases, ranging from 20 to 60%. Nevertheless, in cases where detection occurs, both the SNR and SDE remain high, and almost no detected signal exhibits an SNR or SDE value below 10. This decrease in detectability could be due to the higher scatter in the K2sff's light curve as observed in Appendix C.1. In Figure 3.5, we refine the injection-and-recovery tests between $0.6R_\oplus$ and $1.6R_\oplus$ with steps of $0.1R_\oplus$ and between 0.6 and 7 days with steps of 0.2days. The synthetic planets are difficult to detect below $1.1R_\oplus$ except in a period region between 0.6 and 2 days where planets with radii between 1 and $1.1R_\oplus$ can have a recovery rate up to 90%. Nevertheless, for periods higher than 3.8 days, the recovery rates of radii between $1.1R_\oplus$ and $1.2R_\oplus$ drop. At periods around 3.8 days, we observe a better recovery rate of the synthetic planets with 70% for radii between 0.9 and $1R_\oplus$ and it seems possible to detect planets down to 0.6 $R_\oplus$ in some conditions. This feature is also present at periods around 1d but weaker. This may correspond to periods with less systematics and correction artefacts allowing better detection during the search. Those results gave a minimum value of SNR and SDE for detection of respectively 5.86 (at P = 3.8d and R = $0.8R_\oplus$) and 10.68 (at P = 4d and R = $1.5R_\oplus$). Those values can reach 142 and 48, both at P = 2.5d and R = $3R_\oplus$.

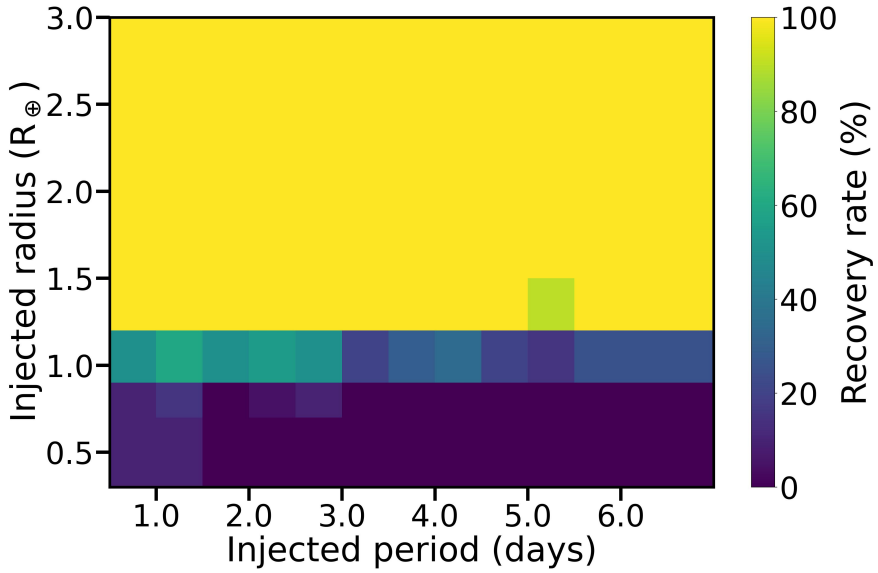


Figure 3.4: Injection-and-recovery test for EPIC 206535752's light curve corrected by K2sff. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_{\oplus}$ between 0.3 and $0.9R_{\oplus}$ and $0.3R_{\oplus}$ between 0.9 and $3R_{\oplus}$. Each period-radius combination is tested for 10 phases.

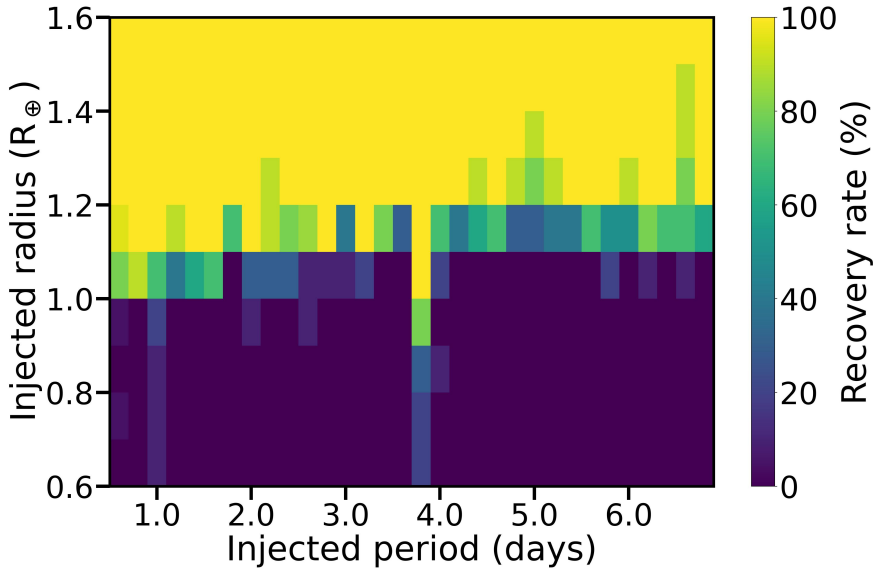


Figure 3.5: Injection-and-recovery test for EPIC 206535752's light curve corrected by K2sff. The pixel size in the period dimension is 0.2 days between 0.6 and 7 days in the radius dimension is $0.1R_{\oplus}$ between 0.6 and $1.6R_{\oplus}$. Each period-radius combination is tested for 10 phases.

An additional way to evaluate K2sff and EVEREST is to compare the SNR and the SDE of both. To better visualise this, I created a colour map where the abscissa corresponds to

the period grid and the ordinate to the radius grid. The colours within the map portray the difference between EVEREST's and K2sff's average SNR (or SDE) of synthetic planets detected within the specific period-radius range. In Figure 3.6, one can clearly see a globally averaged superiority in SNR and SDE for EPIC 206535752's EVEREST light curve. It seems that in a region corresponding to 0.5d $\leq$ P $\leq$ 1.5d and 0.3$R_\oplus$ $\leq$ R $\leq$ 1.5$R_\oplus$, the SNR with K2sff is better. However, it should be noted that it corresponds to an area with a low recovery rate for K2sff and a high recovery rate for EVEREST. Therefore, it should not mean that the K2sff light curve is better in that case. In general, we can estimate, on average, how much the SNR and the SDE are better with EVEREST by computing the total average of the ratios of EVEREST's and K2sff's mean SNR and SDE in each pixel region. For EPIC 206535752, the SNR is on average 2.56 times better for EVEREST and the SDE 1.93 times better.

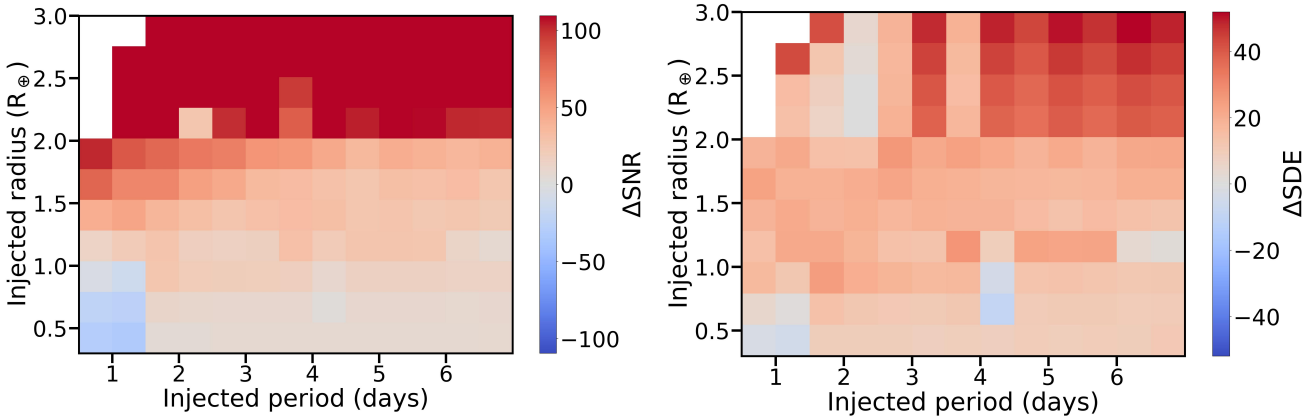


Figure 3.6: Comparison between the EVEREST and K2sff light curve of EPIC 206535752 in terms of the SNR (right) and the SDE (left). The colours in the representation indicate the difference between the average SNR or SDE of detected injected transits in the defined period-radius region, comparing EVEREST and K2sff data. In areas where the representation appears redder, the SNR or SDE is higher in EVEREST. White areas correspond to regions where there's no detection in both EVEREST and K2sff or where the injected planets are beyond their Roche limit.

3.2.2 Results for EPIC 211727748, $K_p = 15$

3.2.2.1 EVEREST Pipeline Results

The first test series for EPIC 211727748 are presented in Figure 3.7. It seems a recovery rate of 100% is achievable for R $\gtrsim$ 0.5$R_\oplus$ and P $\lesssim$ 2d and for R $\gtrsim$ 0.7$R_\oplus$ and P $\gtrsim$ 2d. Synthetic planets with radii inferior to 0.5$R_\oplus$ appear challenging to detect with these first tests. Figure 3.8 refine the results for a smaller radii planet. For small periods, $\lesssim$ 2.5 days, we could in general detect planets with radii between 0.4 and 0.5$R_\oplus$ with recovery rates above 70%. The recovery rate, then, diminishes with the increase in periods. Above 1.1 days and 0.6$R_\oplus$, all injected planets are found. SNR values range from 5.0 (at P = 1.6d and R

= 0.4$R_\oplus$) to 424.8 (at P = 1.5d and R = 3$R_\oplus$) and SDE from 4.96 (at P = 1.2d and R = 0.3$R_\oplus$) to 65.9 (at P = 1.5d and R = 3$R_\oplus$).

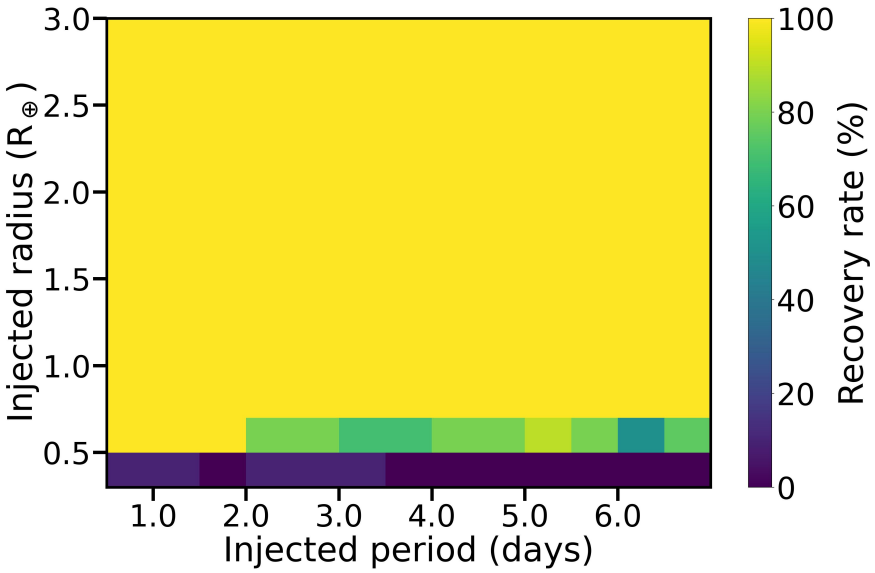


Figure 3.7: Injection-and-recovery test for EPIC 211727748's light curve corrected by EVEREST. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is 0.2$R_\oplus$ between 0.3 and 0.9$R_\oplus$ and 0.3$R_\oplus$ between 0.9 and 3$R_\oplus$. Each period-radius combination is tested for 10 phases.

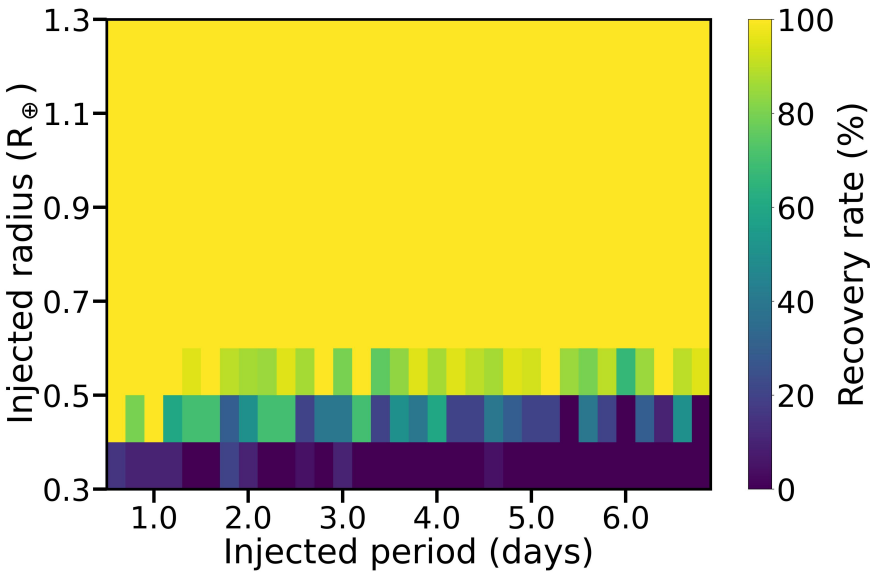


Figure 3.8: Injection-and-recovery test for EPIC 211727748's light curve corrected by EVEREST. The pixel size in the period dimension is 0.2 days between 0.5 and 7 days in the radius dimension is 0.1$R_\oplus$ between 0.3 and 1.3$R_\oplus$. Each period-radius combination is tested for 10 phases.

3.2.2.2 K2sff Pipeline Results

From the first tests presented in Figure 3.9, the difference between EVEREST and K2ff for EPIC 211727748 is less evident than for EPIC 206835752. Recovery rates are lower for injected planets below $0.7R_{\oplus}$. The difference is more visible in Figure 3.10. While it was possible to detect synthetic planets with radii between 0.4 and $0.5R_{\oplus}$ with EVEREST, here it is impossible when $P \gtrsim 1.5$d. SNR values range from 5.0 (at P = 2d and R = $0.5R_{\oplus}$) to 244 (at P = 1.5d and R = $3R_{\oplus}$) and SDE from 6.1 (at P = 6d and R = $0.5R_{\oplus}$) to 50.7 (at P = 1.5d and R = $3R_{\oplus}$). Using K2sff correction, EPIC 211727748 exhibits a higher detection efficiency compared to EPIC 206535752, despite the latter being brighter. If we examine the K2sff light curves (i.e. Appendix C.2), we observe that EPIC 206535752, despite its higher magnitude, displays more scattering than EPIC 211727748. Consequently, the curve's quality, such as the brightness, should be considered as a significant parameter affecting detection efficiency.

In Figure 3.11, we represent the difference between EVEREST and K2sff light curves in terms of SNR and SDE. In all period-radius regions, EVEREST has a better SNR and the same conclusion applies for the SDE, except for 2 small regions at $1.3R_{\oplus}$ and 0.5 days and $1.5R_{\oplus}$ and 1 day. We observe that the improvement in SNR and SDE using EVEREST is less significant in this case, with ratios of 1.85 and 1.45, respectively, compared to the previous star. This difference can be attributed to the fact that the K2sff light curve for this star demonstrates better efficiency, as opposed to the previous case. Additionally, there is a slight reduction in efficiency for EVEREST due to the lower magnitude. Nevertheless, EVEREST still outperforms K2sff.

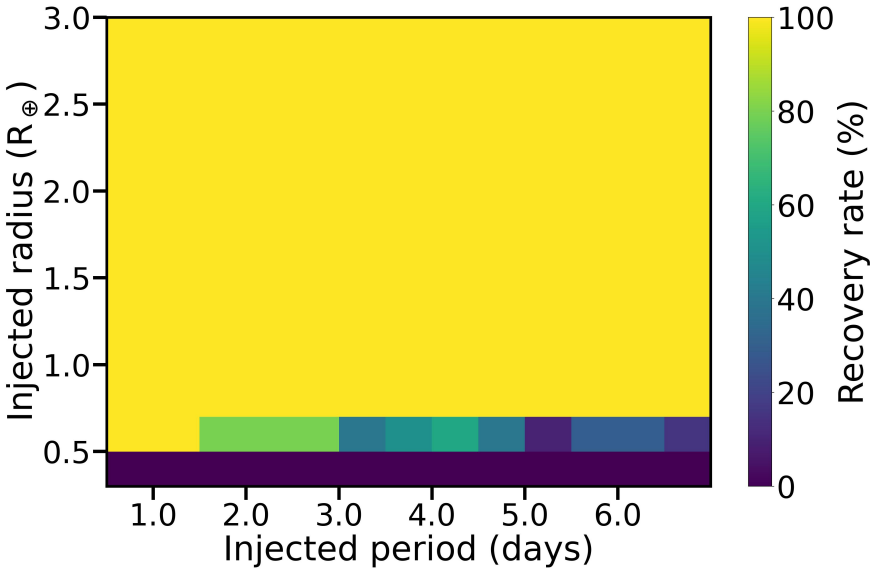


Figure 3.9: Injection-and-recovery test for EPIC 211727748's light curve corrected by K2sff. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_{\oplus}$ between 0.3 and $0.9R_{\oplus}$ and $0.3R_{\oplus}$ between 0.9 and $3R_{\oplus}$. Each period-radius combination is tested for 10 phases. Each period-radius combination is tested for 10 phases.

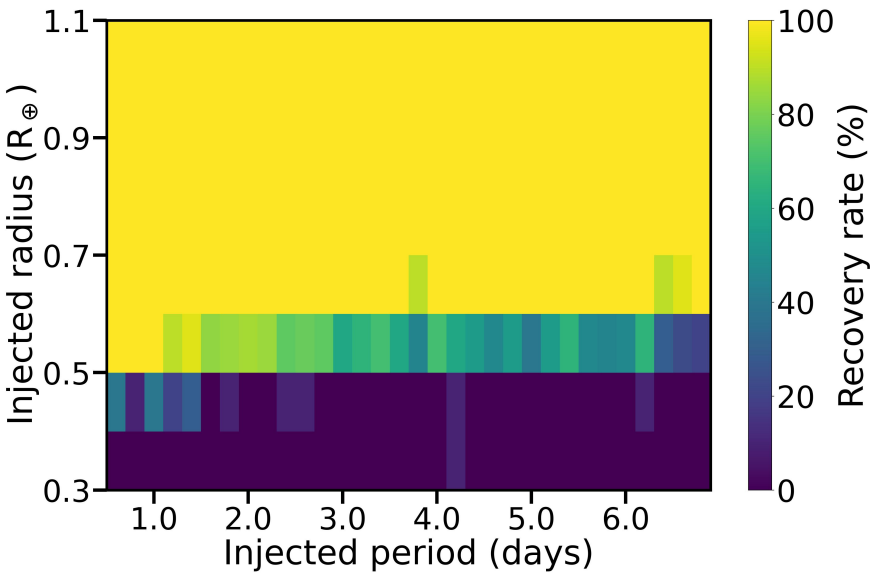


Figure 3.10: Injection-and-recovery test for EPIC 211727748's light curve corrected by K2sff. The pixel size in the period dimension is 0.2 days between 0.5 and 7 days in the radius dimension is $0.1R_{\oplus}$ between 0.3 and $1.1R_{\oplus}$. Each period-radius combination is tested for 10 phases.

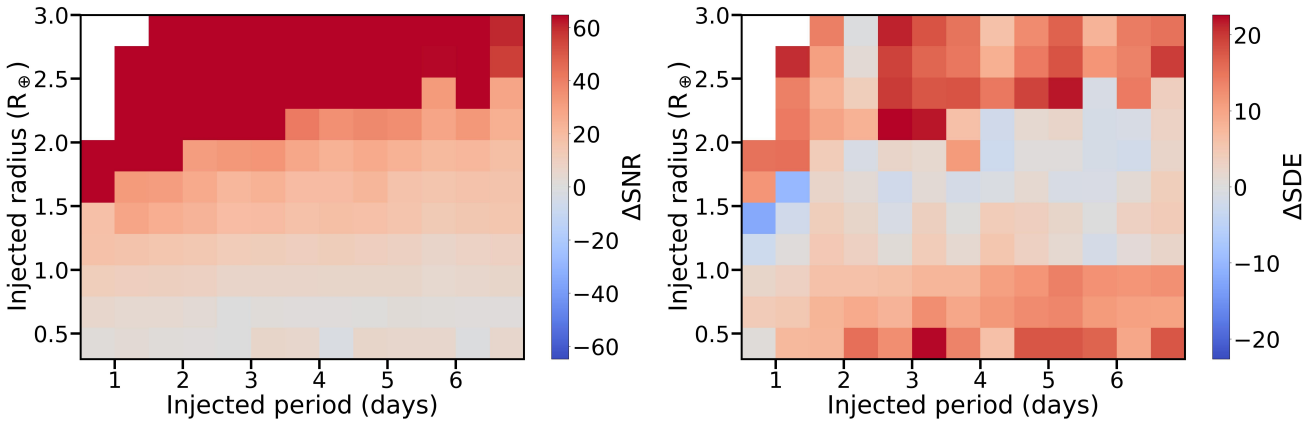


Figure 3.11: Comparison between the EVEREST and K2sff light curve of EPIC 211727748 in terms of the SNR (right) and the SDE (left). The colours in the representation indicate the difference between the average SNR or SDE of detected injected transits in the defined period-radius region, comparing EVEREST and K2sff data. In areas where the representation appears redder, the SNR or SDE is higher in EVEREST. White areas correspond to regions where there's no detection in both EVEREST and K2sff or where the injected planets are beyond their Roche limit.

3.2.3 Results for EPIC 206240954, $K_p = 16.3$

3.2.3.1 EVEREST Pipeline Results

From the results presented in Figure 3.12, a full recovery is only feasible for radii higher than 0.7 $R_\oplus$ when P $\lesssim$ 2 days and for radii higher 0.9 when P $\gtrsim$ 2 days. However, recovery for radii between 0.7$R_\oplus$ and 0.9$R_\oplus$ is still very high (more than 60%) for high periods. In Figure 3.13, the limit between a null recovery rate and a 100% recovery rate is less clear. Detection of 0.6$R_\oplus$ synthetic planets is feasible for small periods ($\lesssim$1.5 days). Recovery of injected transits with radii above 0.7$R_\oplus$ seems in majority very efficient, with recovery rate often above 60%. SNR values range from 5 (at P = 6.6d and R = 0.7$R_\oplus$) to 159 (at P = 1.5d and R = 3$R_\oplus$) and SDE values range from 7 (at P = 7d and R = 0.6$R_\oplus$) to 47 (at P = 1d and R = 2.4$R_\oplus$). In general, compared with EPIC 206535752 and 211727748's EVEREST tests, the range of synthetic planets recovered and the maximal SNR value is smaller for EPIC 206240954. It can be explained by the brightness of the star. As a matter of fact, EPIC 206240954 is 1 magnitude higher than EPIC 211727748 and 2 magnitude higher than EPIC 206535752. Hence, as the detector receives fewer photons, the signal strength relative to the noise decreases, leading to a reduction in SNR. Consequently, the SNR is reduced causing synthetic planetary transits to fall below the detection threshold. Globally, even with a higher magnitude than the targets previously presented, with the EVEREST light curve of EPIC 206240954, it is possible to detect planets slightly below 1 Earth radius, or $\lesssim 1 R_\oplus$ planets in periods below $\sim$ 2 days.

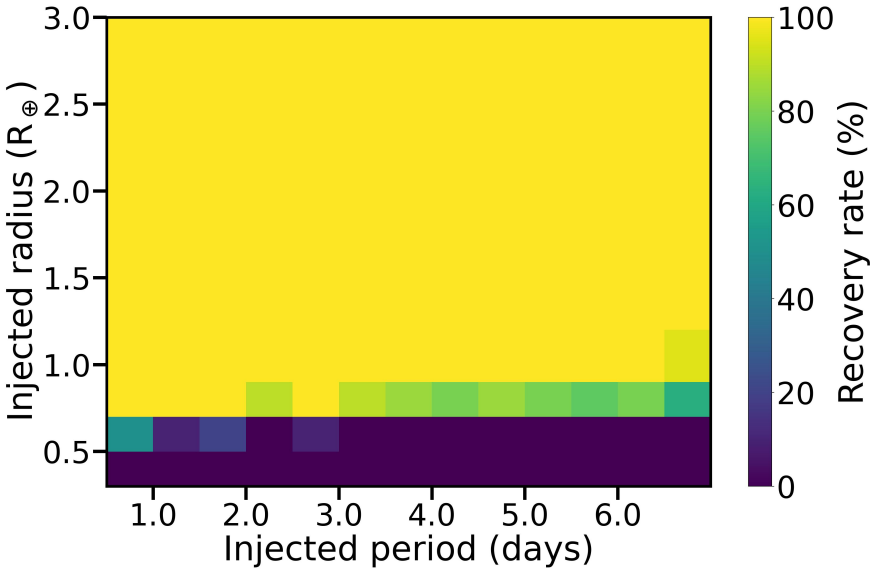


Figure 3.12: Injection-and-recovery test for EPIC 206240954's light curve corrected by EVEREST. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is 0.2$R_\oplus$ between 0.3 and 0.9$R_\oplus$ and 0.3$R_\oplus$ between 0.9 and 3$R_\oplus$. Each period-radius combination is tested for 10 phases. Each period-radius combination is tested for 10 phases..

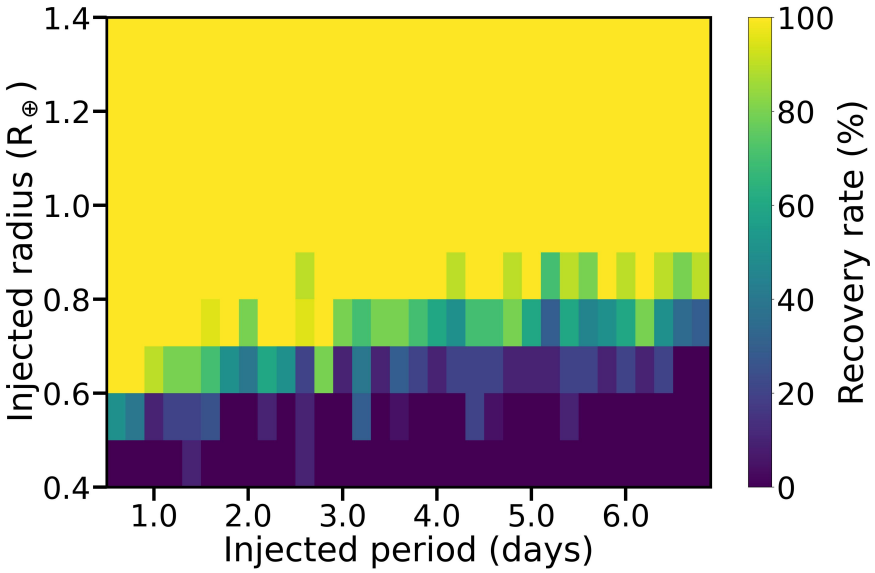


Figure 3.13: Injection-and-recovery test for EPIC 206240954's light curve corrected by EVEREST. The pixel size in the period dimension is 0.2 days between 0.5 and 7 days in the radius dimension is $0.1R_\oplus$ between 0.4 and $1.4R_\oplus$. Each period-radius combination is tested for 10 phases.

3.2.3.2 K2sff Pipeline Results

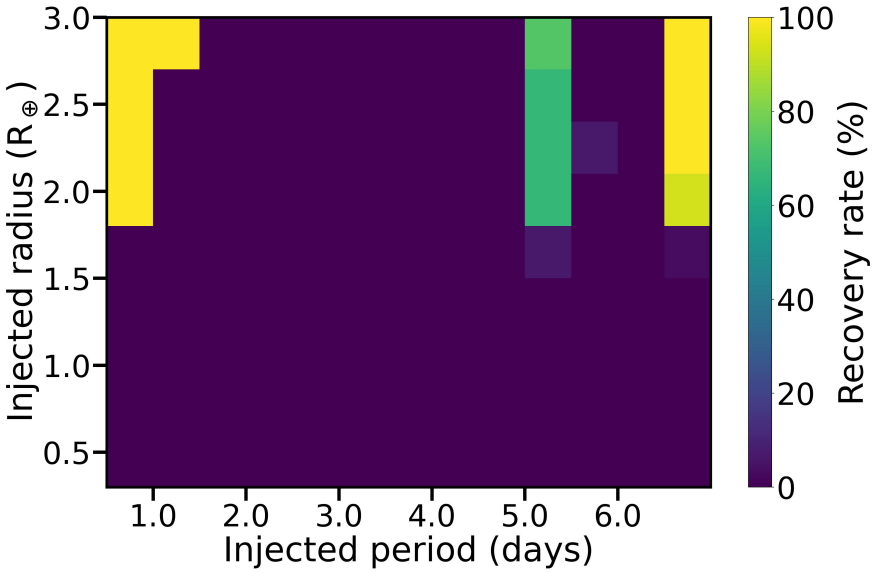


Figure 3.14: Injection-and-recovery test for EPIC 206240954's light curve corrected by K2sff. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_\oplus$ between 0.3 and $0.9R_\oplus$ and $0.3R_\oplus$ between 0.9 and $3R_\oplus$. Each period-radius combination is tested for 10 phases.

The results for K2sff presented in Figure 3.14 are disturbing. Almost no injected planet has been recovered except for 2 small bands at periods between 5 and 5.5 days and between 6.5 and 7 days and $R \gtrsim 1.8R_{\oplus}$. When we looked at the injection-and-recovery report, instead of detecting the synthetic transits, MATRIX found other several events with high SNR and SDE. By observing them carefully in the light curves, for example in Figure 3.15, we noticed that they do not correspond to any transit-like signal but instead are linked to short increase and decrease in luminosity and small holes in the light curves. Those features do not allow us to recover injected planets, therefore we do not recommend using the EPIC 206240954's K2sff light curve for the search of exoplanets. In view of the results for this target, we did not judge necessary to extend the analysis.

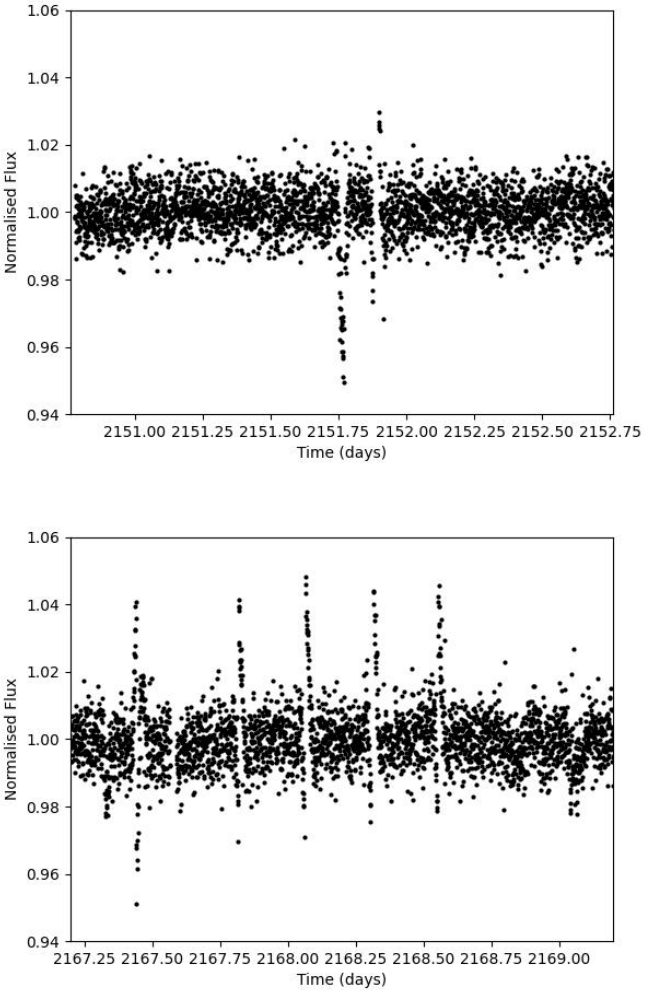


Figure 3.15: Light curve of EPIC 206240954 under K2sff correction. Above: Centered on a signal-generating event at 2151.7619 days. Below: Centered on a signal-generating event at 2168.1995 days.

3.2.4 Results for EPIC 201531672, $K_p = 16.9$

3.2.4.1 EVEREST Pipeline Results

The recovery of injected planets with EPIC 201531672 is more challenging. As we can see in Figure 3.16, no 100% recovery seems possible below $0.9R_\oplus$ and it seems only possible, for radii between 0.9 and $1.2R_\oplus$, to have a recovery rate above 60% when the period is below 3.5 days. From our second series of tests whose results are presented in Figure 3.17, we observe that recovery rate values above 60% are possible below 2.4 days for radii between 0.9 and $1R_\oplus$, below 4.6 days for radii between 0.9 and $1R_\oplus$ and above $1.1R_\oplus$ it is almost possible to reach 90% recovery rate on all the period grid. The maximal SNR value is 91.3 and appears when $R = 3R_\oplus$ and P = 1.5 days. The maximal SDE value is 42.4 when $R = 3R_\oplus$ and P = 2.5 days. The minimal SNR value is 5.02 and happened when P = 1d and $R = 0.8R_\oplus$ and the minimal SDE value is 6.92 when P = 4.5d and $R = 1.1R_\oplus$.

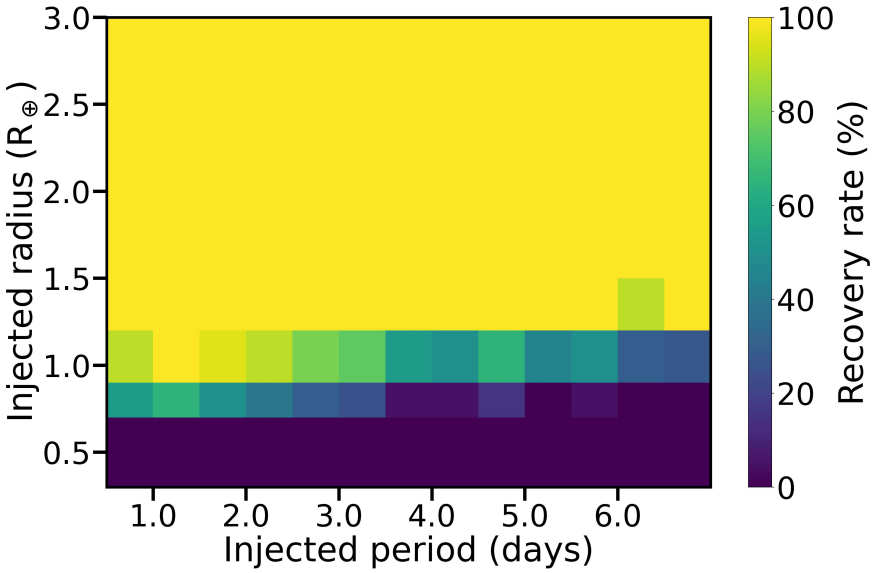


Figure 3.16: Injection-and-recovery test for EPIC 201531672's light curve corrected by EVEREST. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_\oplus$ between 0.3 and $0.9R_\oplus$ and $0.3R_\oplus$ between 0.9 and $3R_\oplus$. Each period-radius combination is tested for 10 phases. Each period-radius combination is tested for 10 phases.

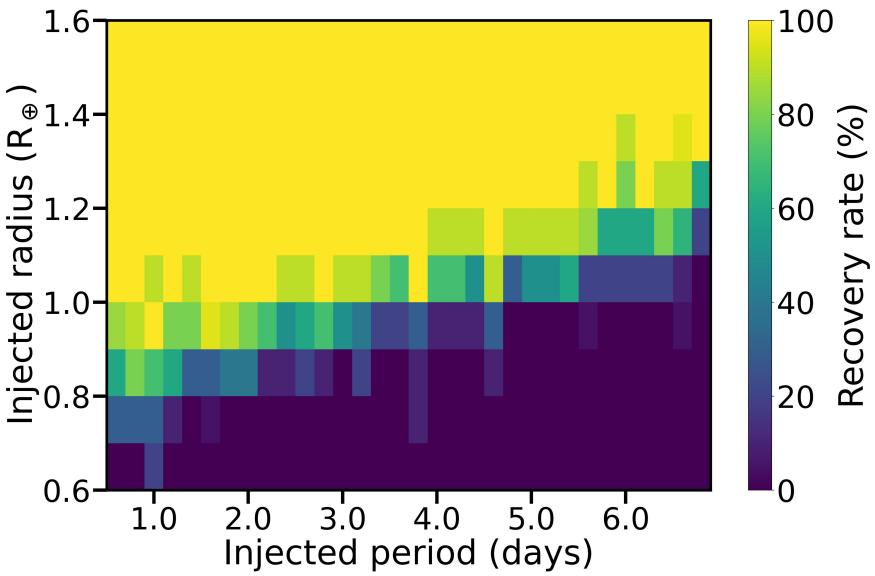


Figure 3.17: Injection-and-recovery test for EPIC 201531672's light curve corrected by EVEREST. The pixel size in the period dimension is 0.2 days between 0.5 and 7 days in the radius dimension is $0.1R_\oplus$ between 0.6 and $1.6R_\oplus$. Each period-radius combination is tested for 10 phases.

3.2.4.2 K2sff Pipeline Results

In Figure 3.18, the first test series results are displayed. The recovery of injected planets is more difficult with the K2sff light curve. From these tests, the detection of synthetic transit between 0.9 and $1.2R_\oplus$ is still possible but with far less recovery rate (about 20 to 60%) than in EVEREST. Planets above $1.2R_\oplus$ are easily recovered with rates in the majority above 70%. The second series of tests, presented in Figure 3.19, shows that it is a bit more complex. Recovery rates vary more with the radius than the previous targets and the limit between full recovery and no recovery is not as strict. Recovery rates for radius above $1.2R_\oplus$ stay in majority 50% and can reach values above 90%. Detecting planets below this radius is possible but the recovery rate becomes small, especially for high periods. The maximal SNR is a little less than with EVEREST and its value is 84.3 when R = $3R_\oplus$ and P = 1.5 days. On the contrary, the maximal SDE value is better than in EVEREST and is about 45.6 when R = $2.4R_\oplus$ and P = 1 days. The minimal SNR and SDE are respectively 5.1 (R = $1.1R_\oplus$, P = 3 days), and 7.5 ($1R_\oplus$, 2 days).

From Figure 3.20, we see that the difference between EVEREST and K2sff is less important. For the SNR, EVEREST still clearly dominates K2sff in almost all regions but it is less evident for the SDE. However, we estimate that the SNR and the SDE are respectively 1.11 and 1.14 better using EVEREST.

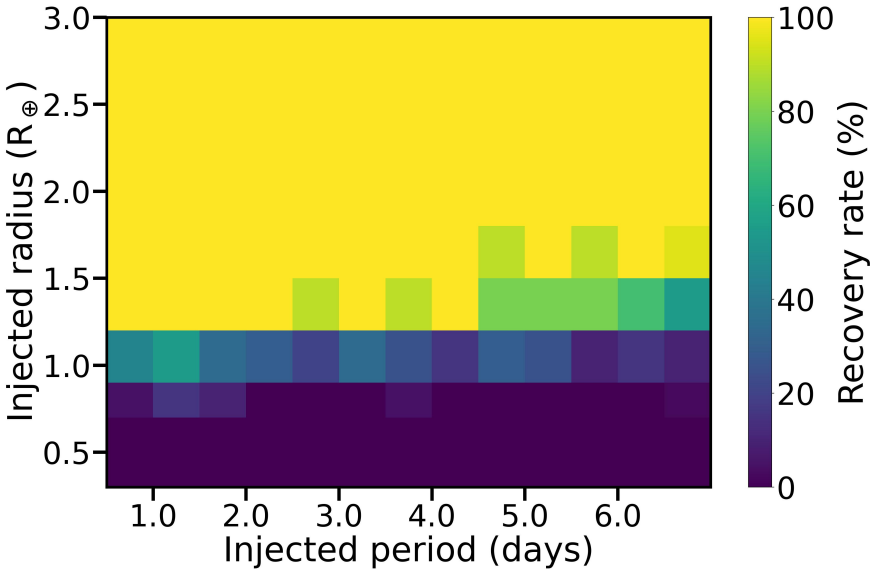


Figure 3.18: Injection-and-recovery test for EPIC 201531672's light curve corrected by K2sff. The pixel size in the period dimension is 0.5 days between 0.5 and 7 days in the radius dimension is $0.2R_{\oplus}$ between 0.3 and $0.9R_{\oplus}$ and $0.3R_{\oplus}$ between 0.9 and $3R_{\oplus}$. Each period-radius combination is tested for 10 phases. Each period-radius combination is tested for 10 phases.

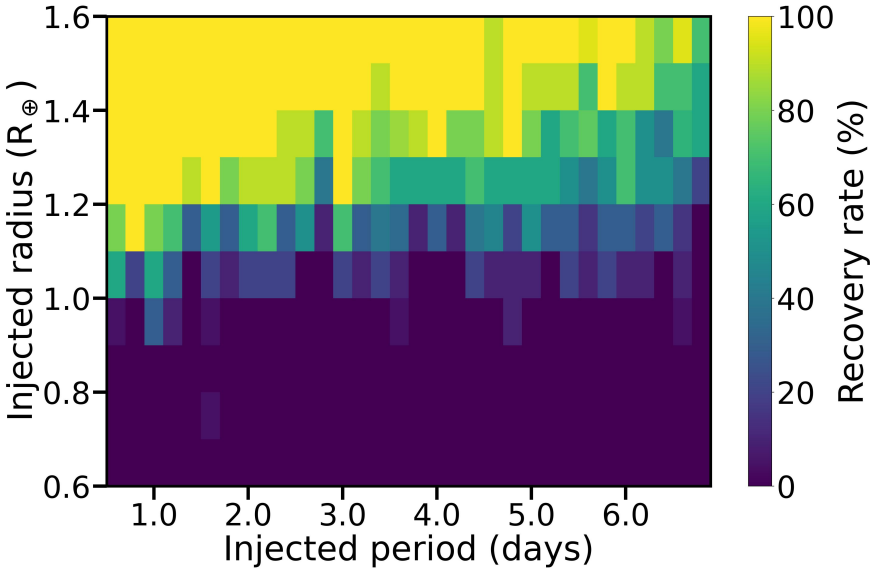


Figure 3.19: Injection-and-recovery test for EPIC 201531672's light curve corrected by EVEREST. The pixel size in the period dimension is 0.2 days between 0.5 and 7 days in the radius dimension is $0.1R_{\oplus}$ between 0.6 and $1.6R_{\oplus}$. Each period-radius combination is tested for 10 phases.

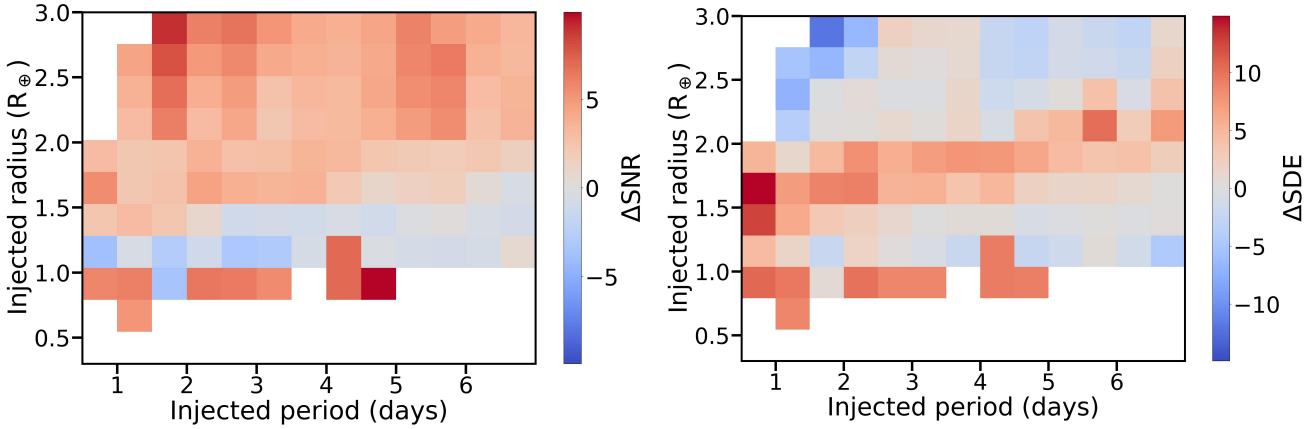


Figure 3.20: Comparison between the EVEREST and K2sff light curve of EPIC 201531672 in terms of the SNR (right) and the SDE (left). The colours in the representation indicate the difference between the average SNR or SDE of detected injected transits in the defined period-radius region, comparing EVEREST and K2sff data. In areas where the representation appears redder, the SNR or SDE is higher in EVEREST. White areas correspond to regions where there's no detection in both EVEREST and K2sff or where the injected planets are beyond their Roche limit.

3.3 Discussion

A summary of the results of the injection-and-recovery tests is presented in Table 3.3. The results clearly demonstrate the superiority of the EVEREST correction pipeline to find planets around hot subdwarfs. In all the tested targets, the recovery rates are higher in EVEREST than in K2sff and EVEREST provides a better radius range. In general, with EVEREST, we can easily recover planets above $0.4R_\oplus$ for bright hot subdwarfs and above $0.7R_\oplus$ for faint hot subdwarfs. On the contrary, K2sff light curves may present noise features which may be assimilated to systematics or correction artefacts. Those features can reduce significantly the lower limit of planetary radii detectable and even make the recovery impossible as we saw for EPIC 206240954. Generally, planets above $1R_\oplus$ could be recovered. If the quality of the raw light curve is good enough and the star is relatively bright, the recovery rate of small planets (>0.5) can be close to EVEREST's case as we saw for EPIC 211727748. The SNR and SDE values are clearly better with EVEREST. However, EVEREST tends to be closer to K2sff when the magnitude of the star increases. This effect was already observed by Luger et al. [45][46].

Therefore, I suggest the use of the EVEREST correction pipeline for the search of exoplanets around hot subdwarf as it allows for probing a bigger range of radius and period. K2sff is not totally excluded as it still gives good results in the injection-and-recovery tests but it will be more useful as a way to test the promising candidates found with EVEREST. For the search for exoplanets in SHERLOCK, I would advise cutting down the light curves to the best-looking 30 to 40 days. The other parts could be used as a test on promising

candidates. We make use of the Sav-Gol filter in the injection-and-recovery and this option must also be selected in SHERLOCK, as we are looking for shallow transit. We can select a minimum SNR of 8 and a minimum SDE of 7 in `yaml` file in SHERLOCK.

Even if these results provide good insights into the K2 data, some criticism should be given. First, the planets are injected in light curves already corrected by EVEREST and K2sff and their transits are, thus, not influenced by them. Therefore, the results should be considered very optimistic. Furthermore, some targets may not agree with the results found. Due to the computing time ($\simeq$ 3 weeks for K2sff and 2 weeks for EVEREST per target), it is not feasible to perform these injection-and-recovery tests for every target.

	EPIC	**206535752**	**211727748**	**206240954**	**201531672**
	Kp	14	15	16.3	16.9
EVEREST	**Recovery rate (%)**	88.53	86.27	77.45	65
	Max SNR	487	424.8	159	91
	Min SNR	5.01	5	5	5
	Max SDE	69	65.9	47	42.4
	Min SDE	7.78	4.96	7	6.92
K2sff	**Recovery rate (%)**	55.87	76.86	12.6	50.1
	Max SNR	142	244	63	84
	Min SNR	5.86	5	21	5
	Max SDE	48	51	34	45.6
	Min SDE	10.68	6.1	13	7.5
	Average SNR ratio	2.56	1.85	1.55	1.11
	Average SDE ratio	1.93	1.45	2.08	1.14

Table 3.3: Summary of the results of the injection-and-recovery tests. Here, the recovery rate corresponds to the percentage of recovered injected transit for all the tests. The average SNR and SDE ratios correspond to the total average of the ratios of EVEREST's and K2sff's mean SNR and SDE in each period-radius region of the injection-and-recovery tests.

Chapter 4

Search for Transiting Exoplanets Around Hot Subdwarfs

During this master thesis, I analysed the 48 targets that are available in the EVEREST dataset by applying the lesson learnt from the injection-and-recovery tests 3.3. Out of those 48 targets, a lot of signals were found however only nine targets contained signals worth performing a vetting. In this chapter, I will present five of them.

4.1 EPIC 212465180

The first target is EPIC 212465180, which is a non-pulsating apparently single sdB star with $K_p = 15.56$. It was observed during the sixth K2 campaign. During the search, SHERLOCK found four signals which are presented in Table 4.1. The second transit was excluded from consideration due to its lower detection count in the detrended and original light curves. Additionally, the distinct shapes and depths of each individual transit did not exhibit consistency. The third signal did not possess clearly visible harmonics. Furthermore, Figure 4.1, representing the results of the search for this candidate, displays variability in the light curve with amplitudes approaching the transit depth, the shape is therefore not satisfactory. The final run similarly lacked noticeable harmonics, and during the vetting process, inconsistencies in the depths of individual transits were noticed. This can be clearly seen in Figure 4.2. Therefore only the first candidate is truly interesting to present here.

Run	P (days)	Dur (min)	T_0	Depth (ppt)	SNR	SDE	BS	N_{det}
1	3.01	32.25	2427.114	0.64	14.03	9.67	1	12
2	2.23	21.14	2426.976	0.78	15.17	8.95	1	4
3	5.34	52.42	2429.424	0.76	14.97	9.74	1	9
4	1.23	38.70	2426.663	0.37	11.08	9.83	1	12

Table 4.1: SHERLOCK results of the four runs search for EPIC 212465180. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS is the border score of the signal and N_{det} represents the number of detrended light curves in which the signal is detected.

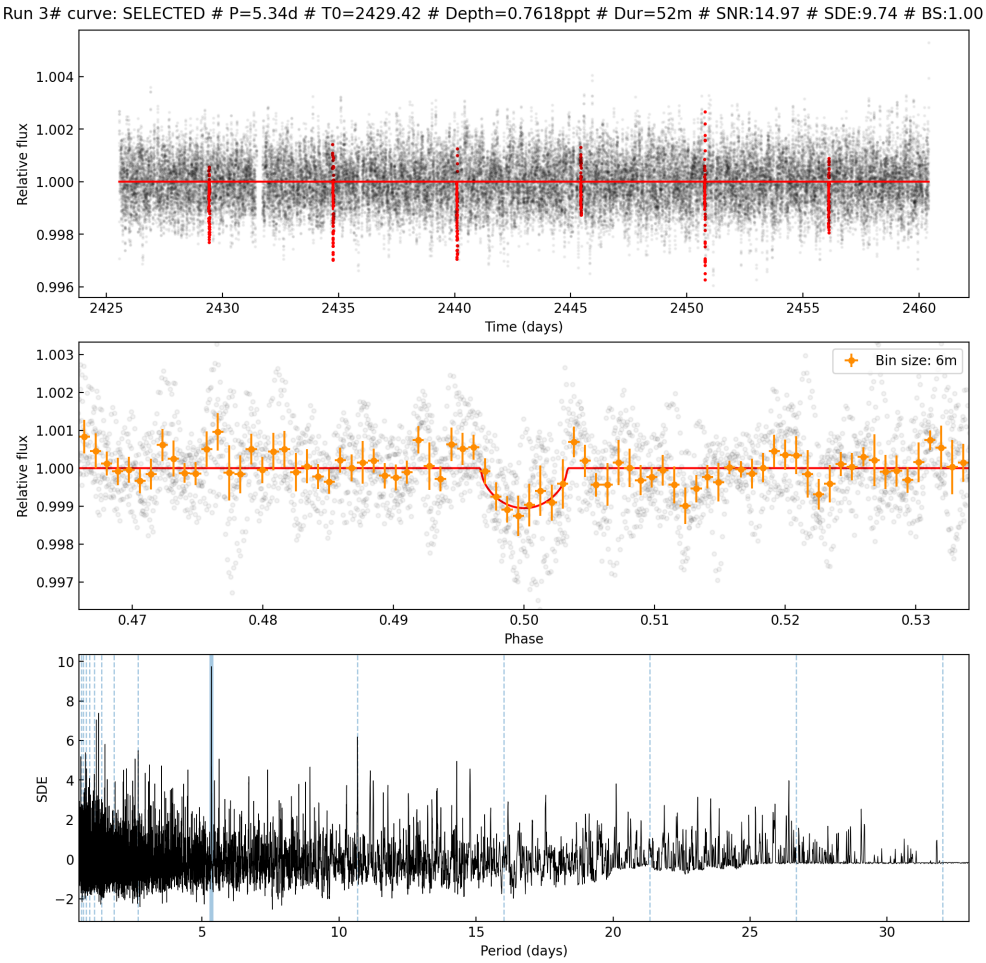


Figure 4.1: EPIC 212465180, SHERLOCK results for the third candidate with a TLS search model. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

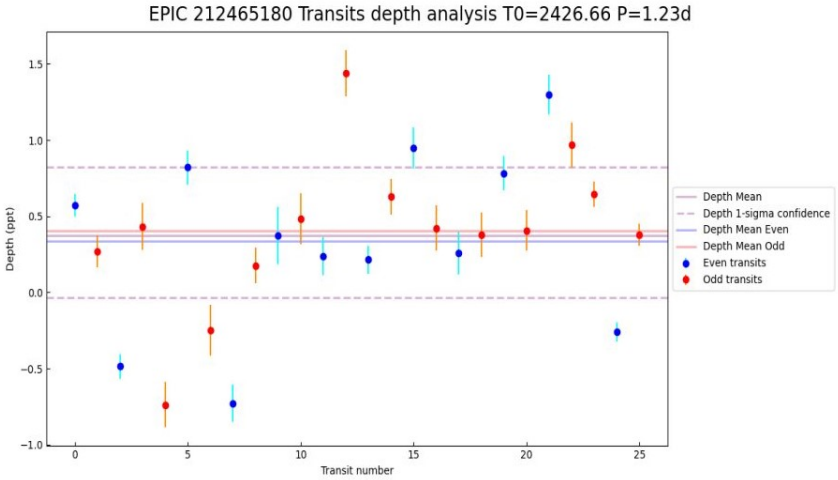


Figure 4.2: EPIC 212465180, vetting of the fourth run, candidate single-transits depths plot.

4.1.1 TLS model

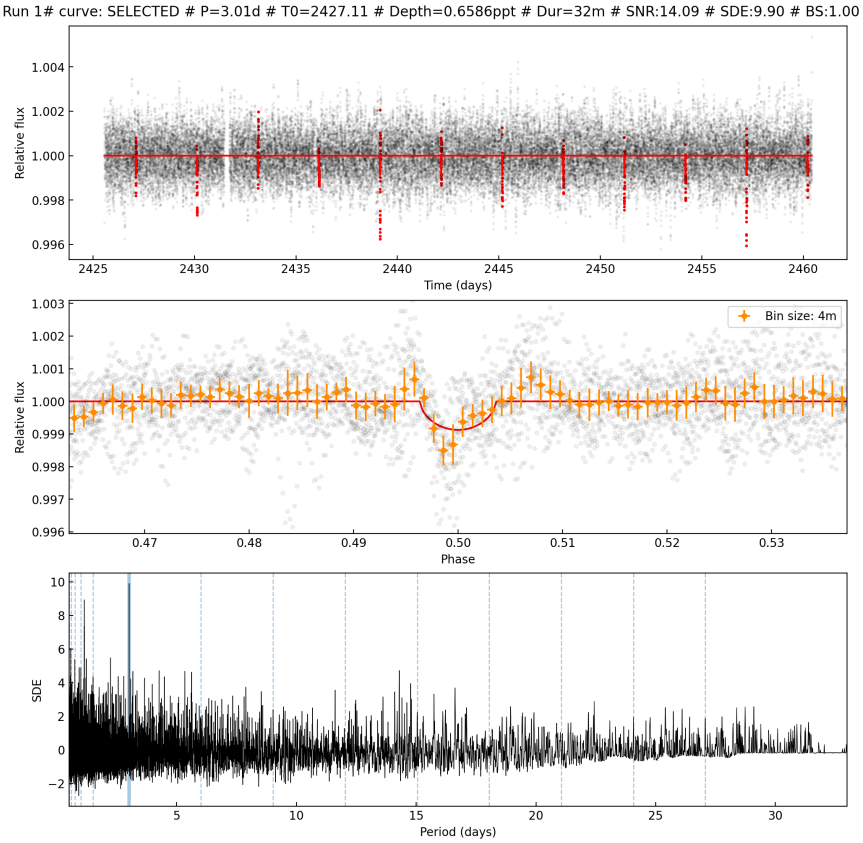


Figure 4.3: EPIC 212465180, SHERLOCK results for the first candidate with a TLS search model. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

The search result plots for the first run are presented in Figure 4.3. The shape of the light curve does not correspond to the model in the phase-folded curves and the harmonics are not visible. However, the shape of the detected transit seems to correspond to a disintegrating object, which is characterised by a tailed shape [58][59]. As explained in section 2.4.3.2, in SHERLOCK, the transit model can be changed. By default, as used here, it is the TLS model but SHERLOCK provides also a tailed transit model [58]. Another try on this target has thus been made using this model. Nevertheless, it should be noted that it has not undergone extensive testing.

4.1.2 Tailed model

The SHERLOCK results of the search run with a tailed transit model are presented in Table 4.2. Our signal is directly found in the first run and in eleven detrends. Its SDE has increased. It is justified by the fact that the tailed transit model has been optimised with SDE. However, its SNR and depth are shorter than with the TLS model. The result of the fourth run with the TLS model is found back here but with shorter SNR and SDE and in fewer detrends, certainly because the TLS model is better suited for this signal. It should be important to notice that the second run is in fact an upper harmonic of the third run.

The result for the first candidate is presented in Figure 4.4. The model succeeded in reproducing the transit shape and the harmonics are clearly visible. This is very promising and we performed the vetting for this candidate. In this case, a tailed-like event is observed. Consequently, the vetting and its metrics are not adapted and can not properly be done. However, we can still look at the depths of every single transit which are plotted in Figure 4.5. One can see that some individual transits are not in the 1-sigma confidence interval and require careful inspection. In alternative to the vetting, a search in the long cadence EVEREST light curve and on the K2sff light curves has been performed to test this candidate. Unfortunately, they were inconclusive. Those additional tests do not discard this target as they can only aim to increase our confidence about it. It is possible that we do not find the target in the long cadence due to the lack of transit points in these LC data (Duration = 57 min so two points per transit) and with K2sff due to the worse quality of its light curve. In conclusion, this candidate seems very promising and I encourage the continuation of its analysis when the vetting for tailed-like events is included in SHERLOCK and the validation and fitting will be available for K2.

Run	P (days)	Dur (min)	T0	Depth (ppt)	SNR	SDE	BS	N_{det}
1	3.01	57.16	2427.113	0.32	9.22	11.77	1	11
2	4.96	40.88	2425.994	0.44	8.21	10.22	0.86	11
3	1.23	56.75	2426.669	0.23	9.25	8.71	1	9
4	1.24	5.51	2426.546	0.81	9.60	7.67	1	10

Table 4.2: SHERLOCK results of the four runs search for EPIC 21246180 with a tailed model. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS for the border score of the signal and N_{det} represents the number of detrended light curves that detected the signal.

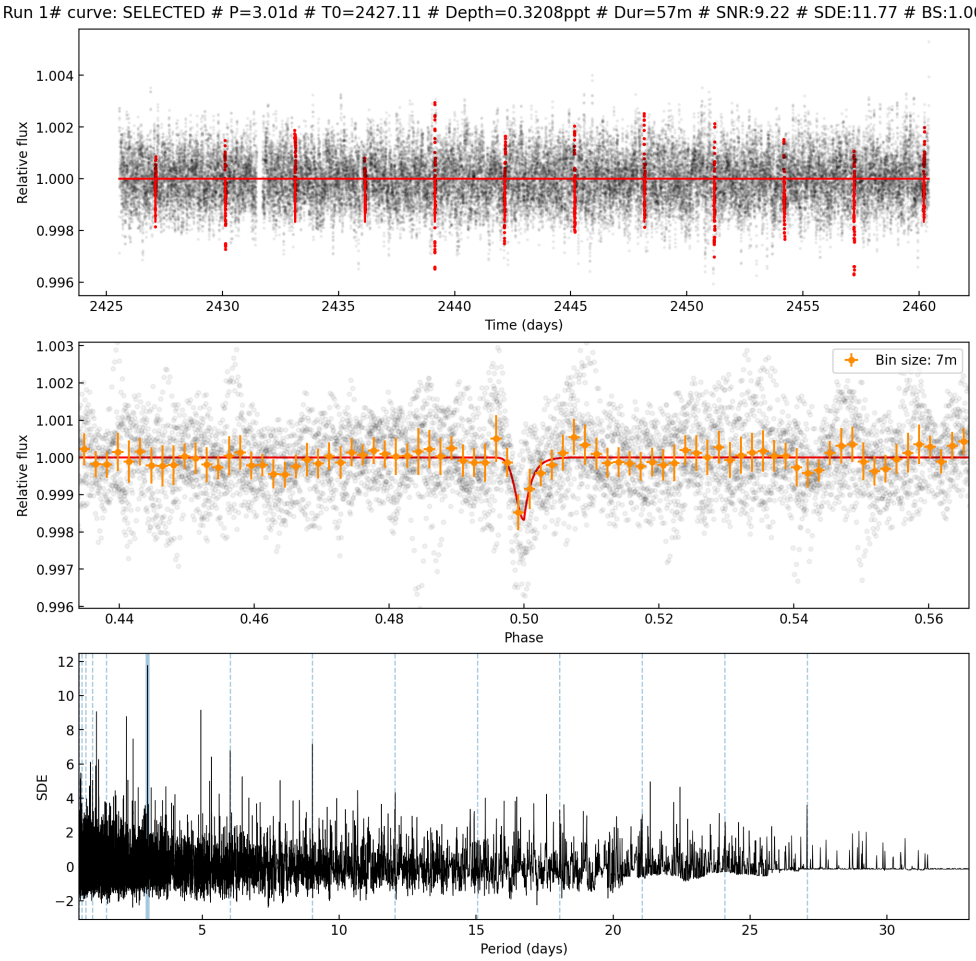


Figure 4.4: EPIC 212465180, SHERLOCK results for the first candidate with a "tailed" search model. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

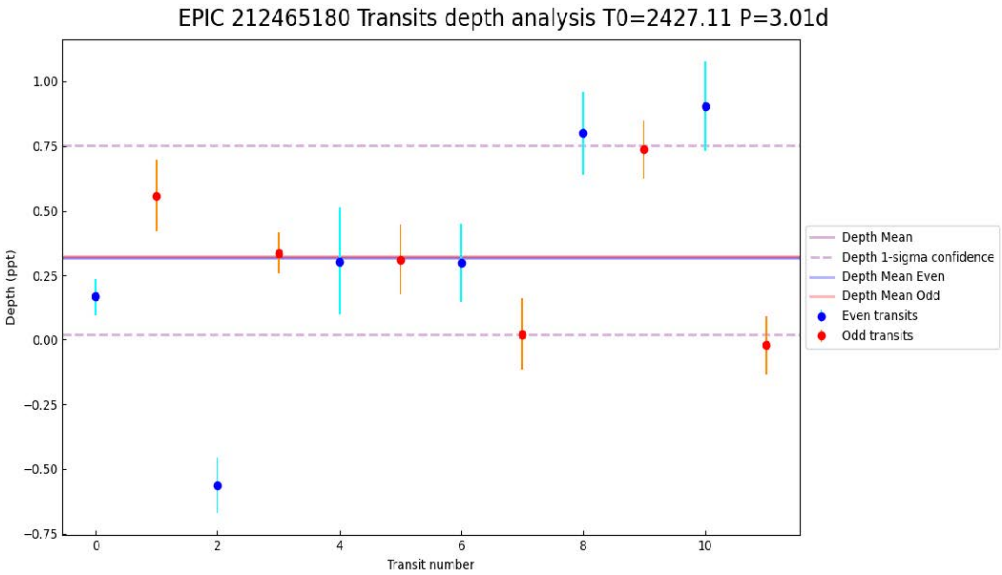


Figure 4.5: EPIC 212465180, first run, candidate single-transits depths plot.

4.2 EPIC 201206621

EPIC 201206621 is a sdB+WD pulsating binary of 0.54 days period with a magnitude of $K_p = 15.99$. It was observed during the first campaign of K2. The pulsations were cleared beforehand with FELIX by V. Van Grootel, a tool designed to extract interactively or automatically periodic variations in a light curve [60]. The SHERLOCK found four potential candidates. The third candidate is rejected due to its presence in a limited number of detrends. The last candidate is discarded because of its low SNR and SDE, its low number of detection in the detrended and original light curves and its unclear shape. The two first signals have low SDE but, because of their high SNR, we will examine them more in detail.

Run	P (days)	Dur (min)	T_0	Depth (ppt)	SNR	SDE	BS	N_{det}
1	1.82	28.30	1979.047	0.79	14.55	6.41	1	7
2	4.84	40.56	1980.501	1.16	14.70	6.99	1	10
3	1.68	29.01	1978.324	0.61	12.19	7.76	1	4
4	12.13	181.51	1981.933	0.50	8.63	7.74	1	5

Table 4.3: SHERLOCK results of the three runs search for EPIC 201206621. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS is the border score of the signal and N_{det} represents the number of detrended light curves in which the signal is detected.

4.2.1 First Run

4.2.1.1 Search module results

The search result plot for the first run of EPIC 201206621 is presented in Figure 4.6. The shape of the transit is satisfactory but only one harmonic is clearly visible. The variability out-of-transit is certainly an artefact due to the Sav-Gol filter. The candidate is not very convincing due to the lack of clear harmonics, so we still decided to have a look at the vetting analyses to firmly refute or confirm the signal

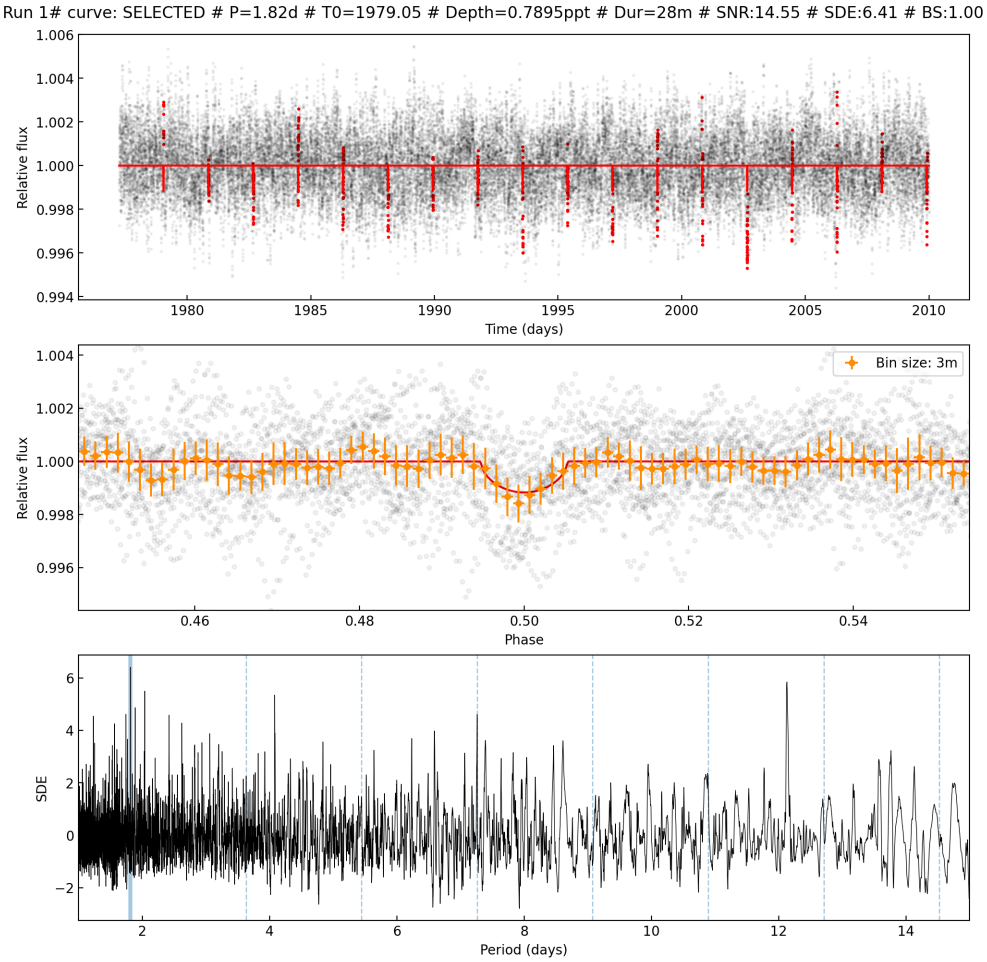


Figure 4.6: EPIC 201206621, SHERLOCK results for the first run. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

4.2.1.2 Vetting results

The recapitulative table of the numerical tests for the vetting of the first candidate of EPIC 201206621 is introduced in Table 4.4. The two first tests are not passed. The depth plot, displayed in Figure 4.7, shows that most of the transits seem to agree with the average

value but some visible outliers are visible with depths superior to ≃3 times the mean value and with negative depths. Due to this large variability, the signal is discarded.

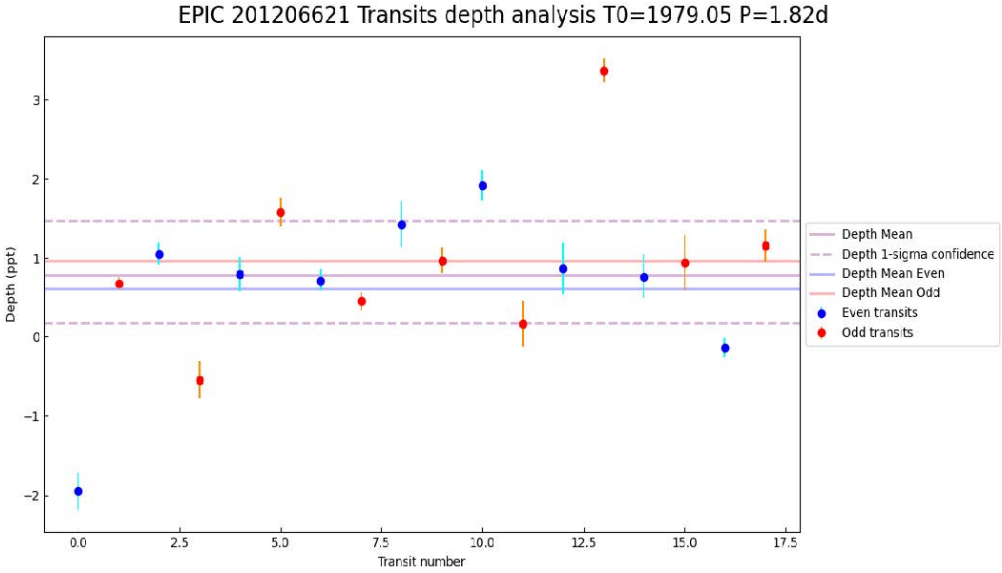


Figure 4.7: EPIC 201206621, first, candidate single-transits depths plot.

Metric	Passed
short_snr	FALSE
long_snr	FALSE
snr_p_t0	TRUE
snr_p_2t0	TRUE
snr_2p_t0	TRUE
snr_2p_2t0	TRUE
snr_p2_t0	TRUE
snr_p2_t02	TRUE
snr_p_score	TRUE
snr_2p_score	TRUE
snr_p2_score	TRUE

Table 4.4: Summary table of the vetting results for the first candidate of EPIC 201206621.

4.2.2 Second Run

4.2.2.1 Search module results

The search result plot for the second run of EPIC 201206621 is presented in Figure 4.8. The transit's shape is accurate, and both the first upper and lower harmonics are distinctly observable. Furthermore, within this run, one of the twelve detrends identified a p/2 sub-harmonic of the signal. It could be possible that the candidate selected by Sherlock is, in fact, an upper harmonic of the signal.

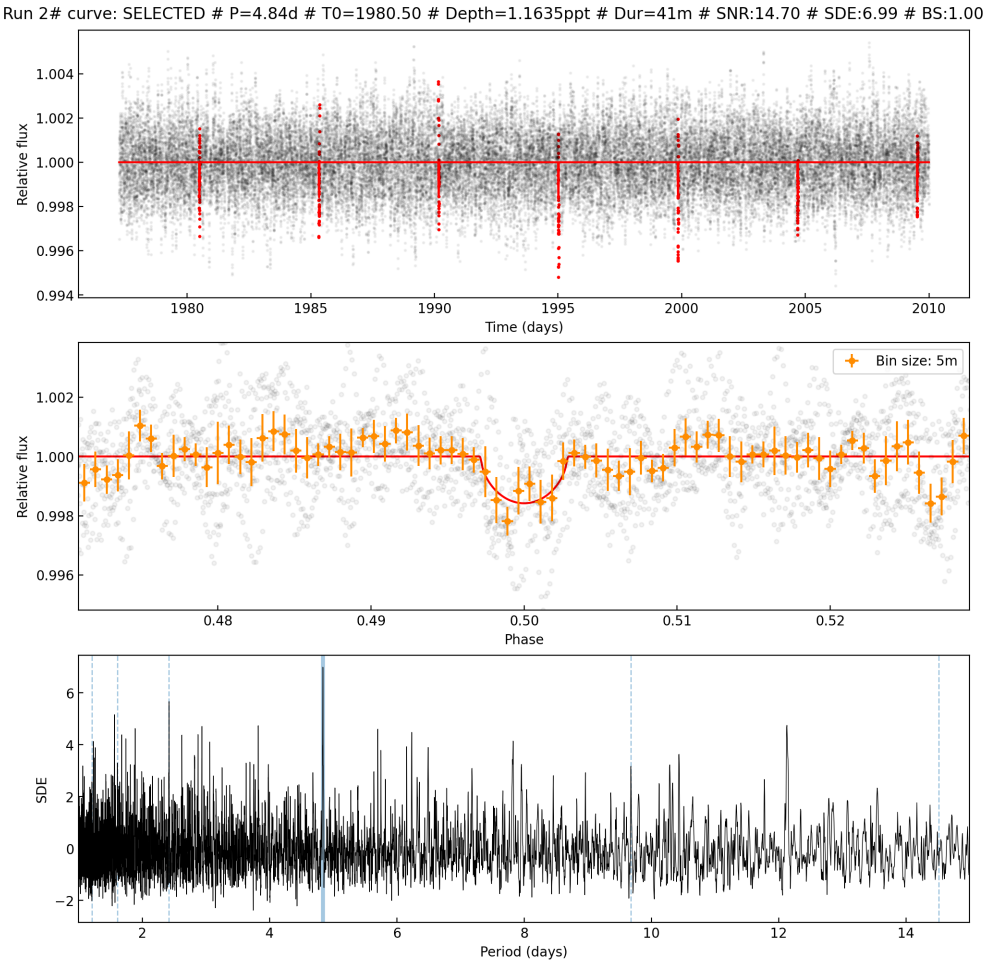


Figure 4.8: EPIC 201206621, SHERLOCK results for the second run. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

4.2.2.2 Vetting results

The summary of the vetting test results is presented in Table 4.5. The target does not satisfy the snr_p_2t0 and thus the snr_p2_t0 test, because a signal with a SNR higher than three is identified when the curve is phase-folded at an epoch T_0 + p/2 with a period p. Similarly, it does not meet the criteria for the snr_p2_t02 metric, and consequently, the snr_p2_score is not achieved. This implies that a signal is detected when the light curve is folded at half the signal period, with an epoch of T0+p/2. However, when we look at the curves related to these metrics, they are not clear on the presence of a signal with a period of p/2. Consequently, it is not certain that we have detected an upper-harmonics. Furthermore, as visible in Figure 4.10, the transit is not consistent across the light curve and hence it is not necessary to analyse further this candidate.

Metric	Passed
short_snr	FALSE
long_snr	FALSE
snr_p_t0	TRUE
snr_p_2t0	FALSE
snr_2p_t0	TRUE
snr_2p_2t0	TRUE
snr_p2_t0	FALSE
snr_p2_t02	TRUE
snr_p_score	FALSE
snr_2p_score	TRUE
snr_p2_score	FALSE

Table 4.5: Summary table of the vetting results for the second candidate of EPIC 201206621.

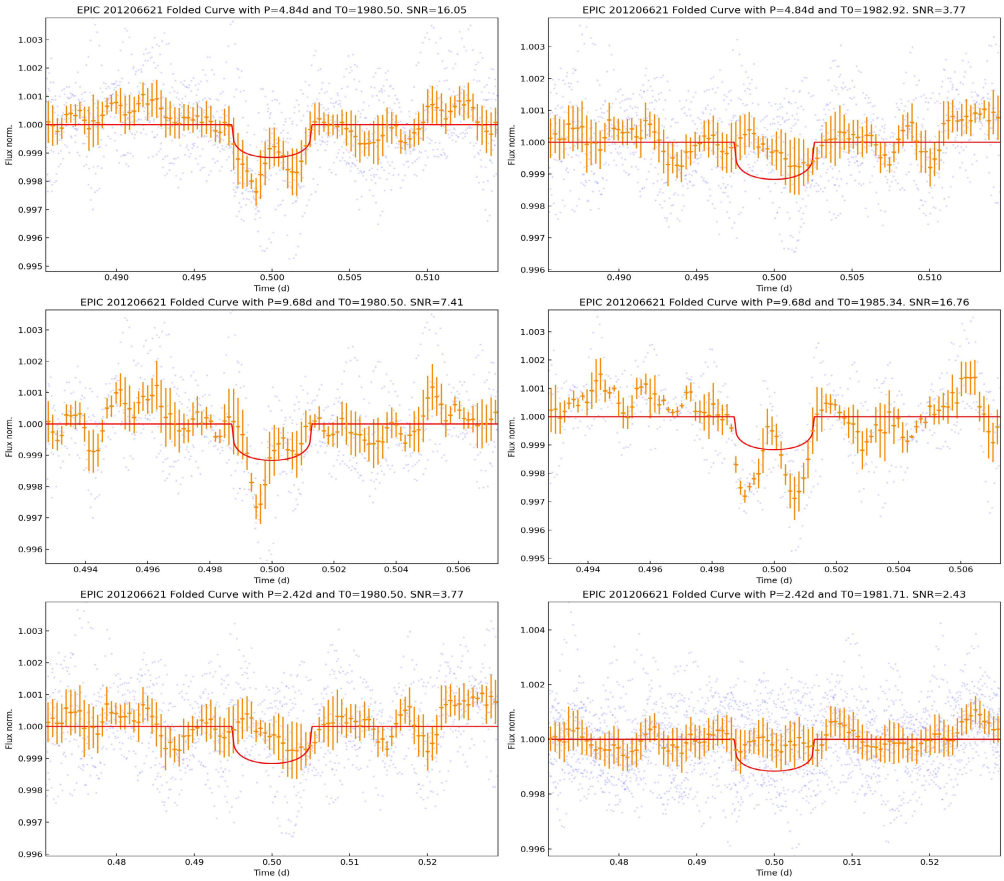


Figure 4.9: EPIC 201206621. Above, the candidate folded at its found period for the found epoch and epoch + P/2. Middle, the candidate folded at its harmonic for the found epoch and epoch + P. Bottom, the candidate folded at its subharmonic for the found epoch and epoch + P/2, where the candidate has been masked.

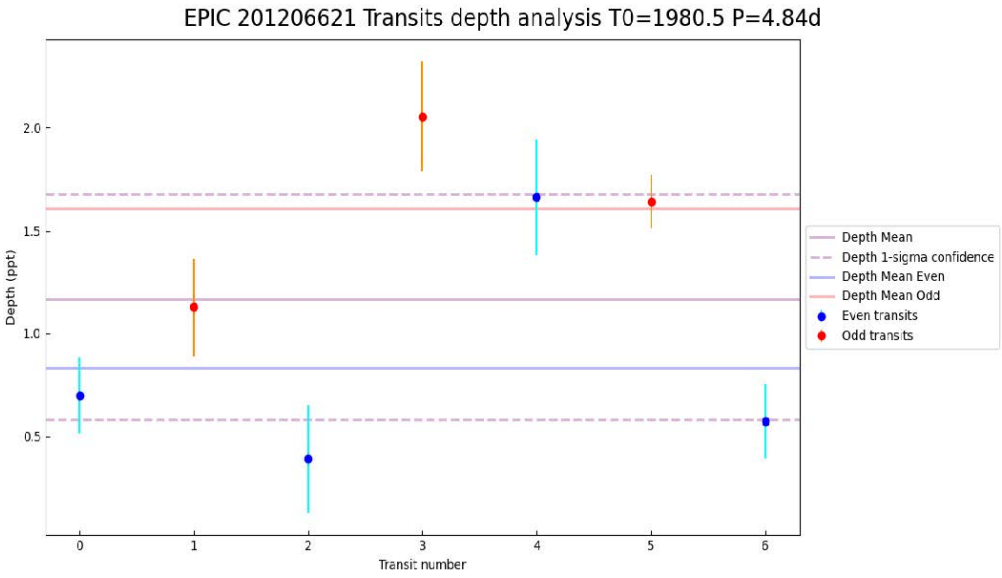


Figure 4.10: EPIC 201206621, second run, candidate single-transits depths plot.

4.3 EPIC 206073023

EPIC 206073023 is a non-pulsating apparently single sdB with a magnitude of $K_p = 15$. It has been observed during the third campaign of K2. In the SHERLOCK run, two signals were found. Only the first one is truly interesting as the second has been detected in only three detrends and its shape is not satisfactory.

Run	P (days)	Dur (min)	T0	Depth (ppt)	SNR	SDE	BS	N_{det}
1	4.01	36.34	2151.932	0.51	14.99	8.10	1	9
2	23.17	196.13	2155.730	0.59	18.07	8.65	1	3

Table 4.6: SHERLOCK results of the four runs search for EPIC 206073023. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS is the border score of the signal and "N_{det}" represents the number of detrended light curves that detected the signal.

4.3.1 Search module results

The results of the search run of the first candidate are plotted in Figure 4.11. We can see that two transits have a higher depth relative to the others. However some harmonics are visible and the SNR is high, we will perform the vetting.

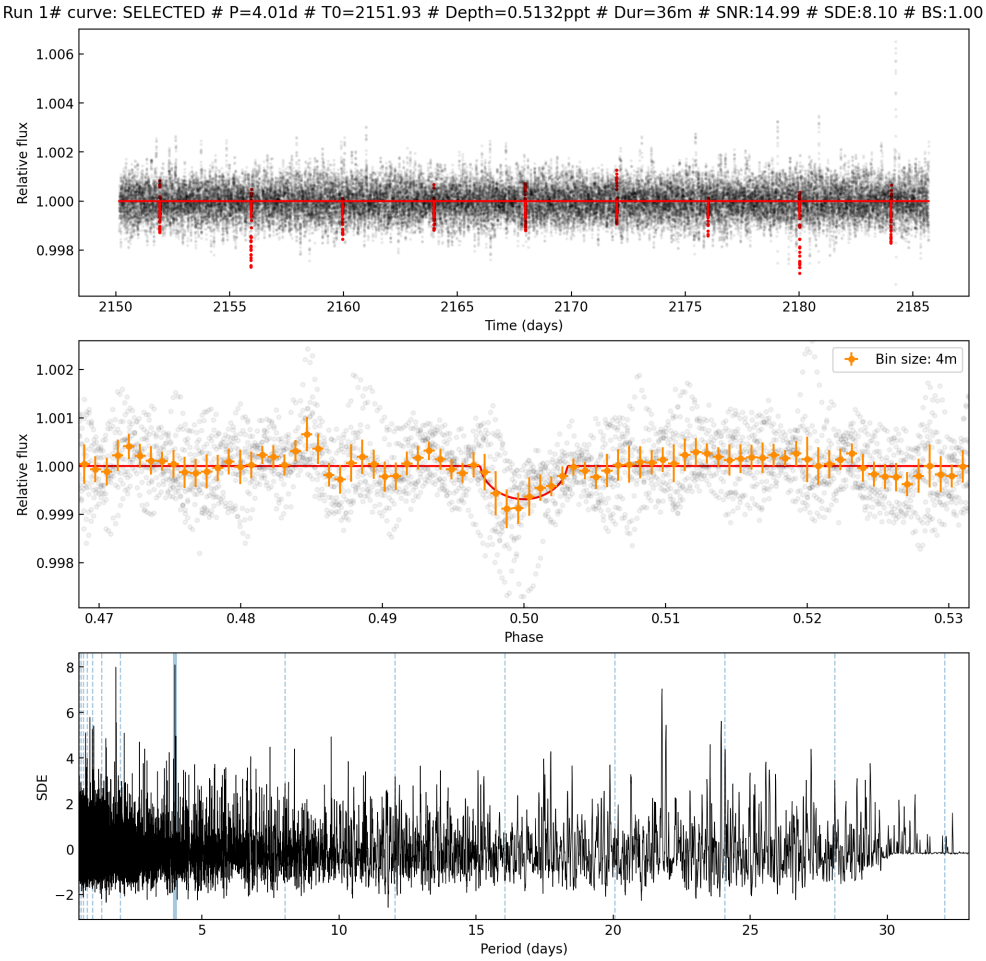


Figure 4.11: EPIC 206073023, SHERLOCK results for the first run. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

4.3.2 Vetting results

The vetting results are summarised in Table 4.7. The candidate does not pass the two first tests and the metric snr_p2_t02. Furthermore, the depth plot in Figure 4.12 shows that we have two deeper transits ($\simeq$0.9-1ppt) and one null depth out of the nine transits. Those can have a high impact on the mean depth. There is a possibility that the two outliers are connected, a more complete inspection should be needed. Furthermore, the full set of transits is inconsistent, thus the viability of this candidate is not satisfactory and we will not continue its analysis.

Metric	Passed
short_snr	FALSE
long_snr	FALSE
snr_p_t0	TRUE
snr_p_2t0	TRUE
snr_2p_t0	TRUE
snr_2p_2t0	TRUE
snr_p2_t0	TRUE
snr_p2_t02	FALSE
snr_p_score	TRUE
snr_2p_score	TRUE
snr_p2_score	FALSE

Table 4.7: Summary table of the vetting results for the first candidate of EPIC 206073023.

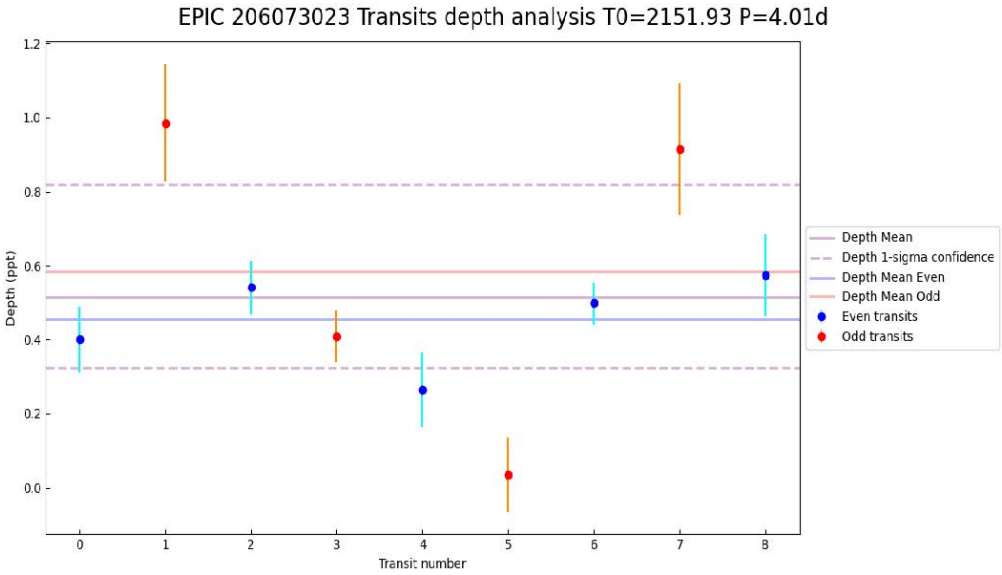


Figure 4.12: EPIC 206073023, first, candidate single-transits depths plot.

4.4 EPIC 214515136

EPIC 214515136 is a non-pulsating apparently single sdB of high magnitude ($K_p = 16.30$). It was observed during the seventh K2 campaign. The SHERLOCK search found four signals but only the first is truly interesting. The second candidate is discarded due to the absence of visible harmonics and a single transit displaying significantly greater depth in comparison to the others. The third one is rejected due to the lack of harmonics and its low SDE. The last candidate is not considered as it is found in only two detrends.

Run	P (days)	Dur (min)	T0	Depth (ppt)	SNR	SDE	BS	N_{det}
1	11.01	154.00	2469.938	1.11	25.55	12.36	1	8
2	8.75	44.31	2469.221	1.08	15.21	8.14	1	6
3	9.77	48.18	2472.270	1.09	14.24	6.86	1	6
4	8.32	114.50	2472.423	0.62	17.12	8.22	1	2

Table 4.8: SHERLOCK results of the 4 runs search for EPIC 214515136. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS is the border score of the signal and "N_{det}" represents the number of detrended light curves that detected the signal.

4.4.1 Search results

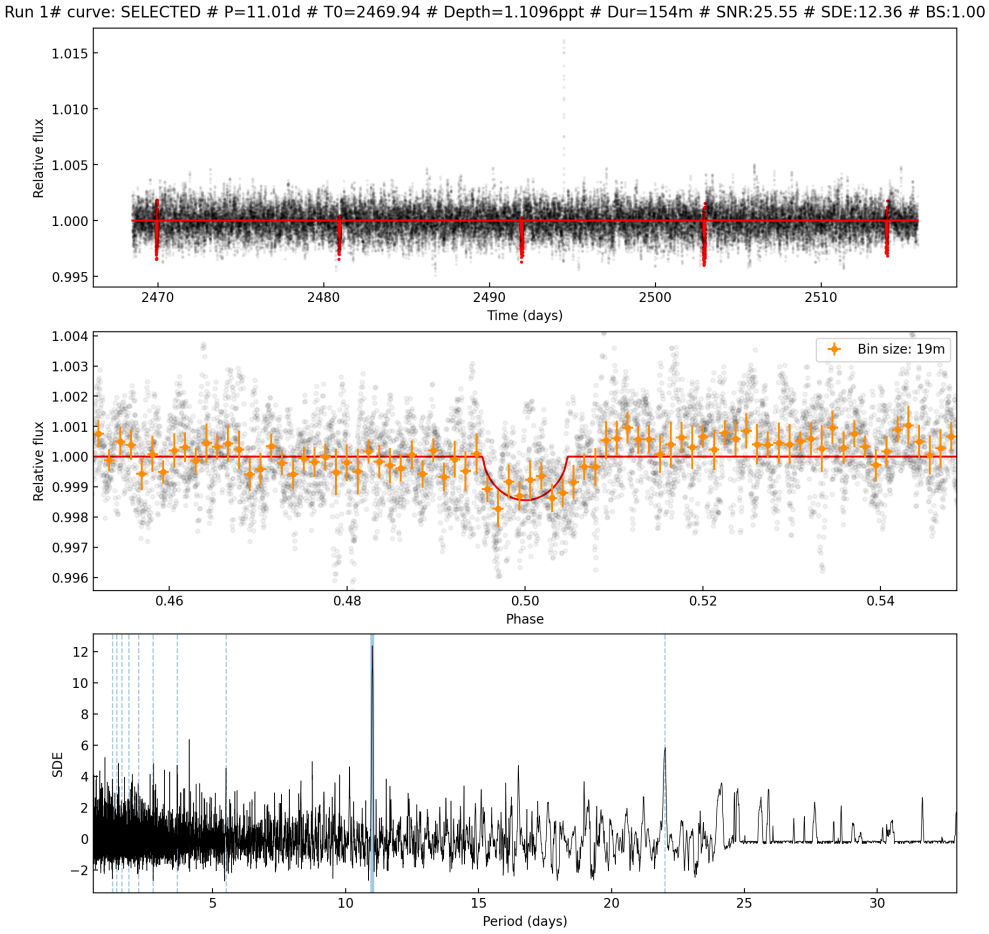


Figure 4.13: EPIC 214515136, SHERLOCK results for the first run. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

The search results plots are presented in Figure 4.13. It has a very high SNR and SDE and is found in eight detrends. Upper and sub-harmonics are visible and the depth does not seem to vary a lot. Weak variations are visible in the phase-folded curve and the right side seems to be higher than the left side. Its duration is longer than we expect for this kind of transit. A fitting of this candidate, if the vetting is favourable, is needed to estimate better planetary properties and yield a more robust conclusion. As the properties of this candidate are promising, we perform its vetting.

4.4.2 Vetting results

The vetting results are summarised in Table 4.9. It passed almost every test except the first and second one as all the presented candidates. As we can see in Figure 4.14, the first transit has a depth that corresponds to half the average depth and the fourth transit is a little bit above the 1-sigma confidence. Given its long duration, we tried to test this transit in a long cadence but nothing was found during the search. The K2sff light curves are unfortunately too noisy for this target and hence cannot be used as a test. Despite the dispersion in the transit-depth plot, this candidate is very promising and I encourage its further analysis especially when the validation and fitting will be available in SHERLOCK.

Metric	**Passed**
short_snr	FALSE
long_snr	FALSE
snr_p_t0	TRUE
snr_p_2t0	TRUE
snr_2p_t0	TRUE
snr_2p_2t0	TRUE
snr_p2_t0	TRUE
snr_p2_t02	TRUE
snr_p_score	TRUE
snr_2p_score	TRUE
snr_p2_score	TRUE

Table 4.9: Summary table of the vetting results for the first candidate of EPIC 214515136.

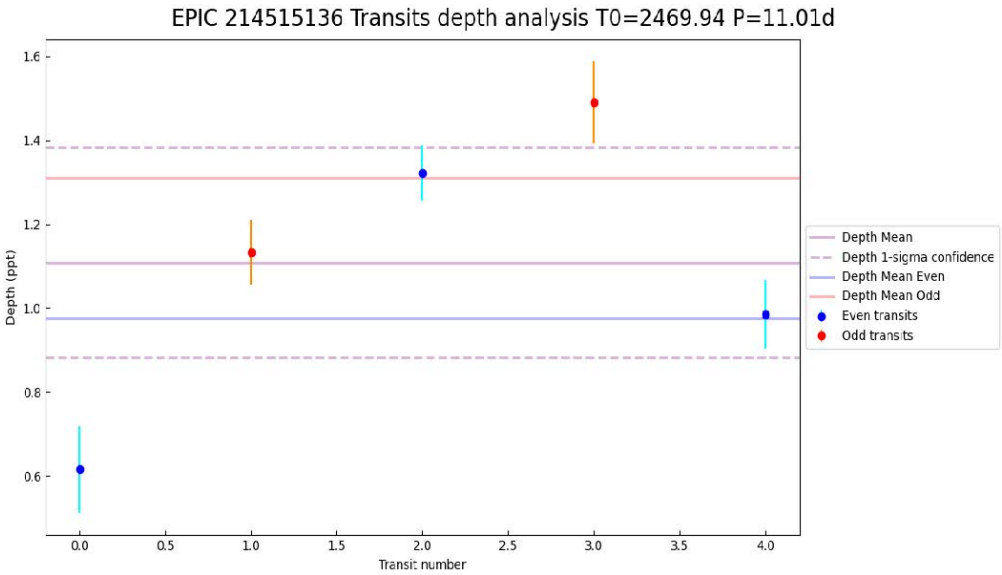


Figure 4.14: EPIC 214515136, first, candidate single-transits depths plot.

4.5 EPIC 211517387

EPIC 211517387 is a high magnitude (K_p =17.32) non-pulsating apparently single sdO. It was observed during the fifth campaign of K2. During the search, SHERLOCK found four potential candidates. Their characteristics are presented in Table 4.10. Only the first candidate cannot be obviously discarded. The second candidate was rejected due to the absence of visible harmonics and an unsatisfactory transit shape in the light curve. The third one was discarded because the transit was mainly influenced by two transits with high depth with respect to the three others. Similar conclusions apply to the last candidate, the transit shape was not acceptable and substantial variability in the depths of individual transits was noted.

Run	**P (days)**	**Dur (min)**	**T_0**	**Depth (ppt)**	**SNR**	**SDE**	**BS**	**N_{det}**
1	8.60	50.81	2350.351	2.61	13.60	8.62	1	8
2	0.50	34.26	2344.710	0.88	15.14	8.08	0.99	10
3	6.25	62.91	2348.639	2.54	16.22	8.25	1	7
4	10.11	77.40	2348.574	2.31	14.49	9.01	1	6

Table 4.10: SHERLOCK results of the four runs search for EPIC 211517387. P is the period in days, T_0 the epoch, Dur is the duration in minutes, BS is the border score of the signal and "N_{det}" represents the number of detrended light curves in which the signal is detected.

4.5.1 Search results

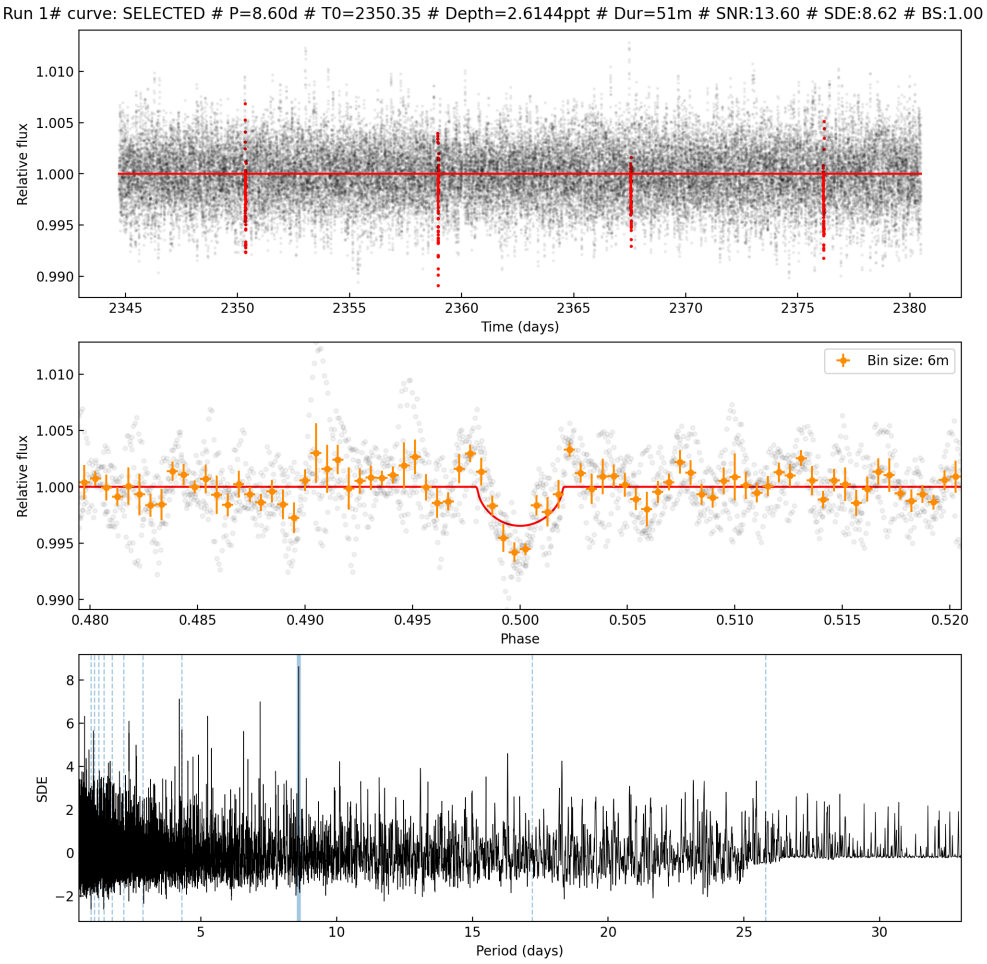


Figure 4.15: EPIC 211517387, SHERLOCK results for the first run. Above: the flux as a function of time and in red lines the suggested transits. Middle: the phase folded curves of the transit. Bottom: the SDE as a function of the Period with the associated harmonics in blue.

The result plots for the first candidates are shown in Figure 4.15. As we can see, SHERLOCK detects a deep drop of luminosity with a period of 8.60 days. However, the harmonics are not very visible. Moreover, when we look around the phase-folded curve, we can see that the curve is subject to many variations out-of-transit. However, we continue with the vetting to firmly confirm or refute this candidate.

4.5.2 Vetting results

Table 4.11 represents the recapitulative table of the numerical tests for the vetting of the first candidate of EPIC 211517387. We can see that this signal passed almost every test, except the short_snr and short_snr such as for the first candidate of EPIC 212465180. The transit depths seem consistent as we can observe in Figure 4.16. However, when we look

at the light curve at the time of the second and third transit in Figure 4.17, variability due to the use of the Sav-Gol filter is present in each transit. The transit's shape, absence of harmonics, and variability outside of the transit collectively suggest that this candidate is not very convincing. However, a more detailed analysis should be made to completely discard this candidate.

Metric	Value	Passed
short_snr	1.306	FALSE
long_snr	1.283	FALSE
snr_p_t0	13.951	TRUE
snr_p_2t0	2.793	TRUE
snr_2p_t0	8.173	TRUE
snr_2p_2t0	12.703	TRUE
snr_p2_t0	2.793	TRUE
snr_p2_t02	0.001	TRUE
snr_p_score	0.2	TRUE
snr_2p_score	4.53	TRUE
snr_p2_score	0.2	TRUE

Table 4.11: Summary table of the vetting results for the first candidate of EPIC 211517387.

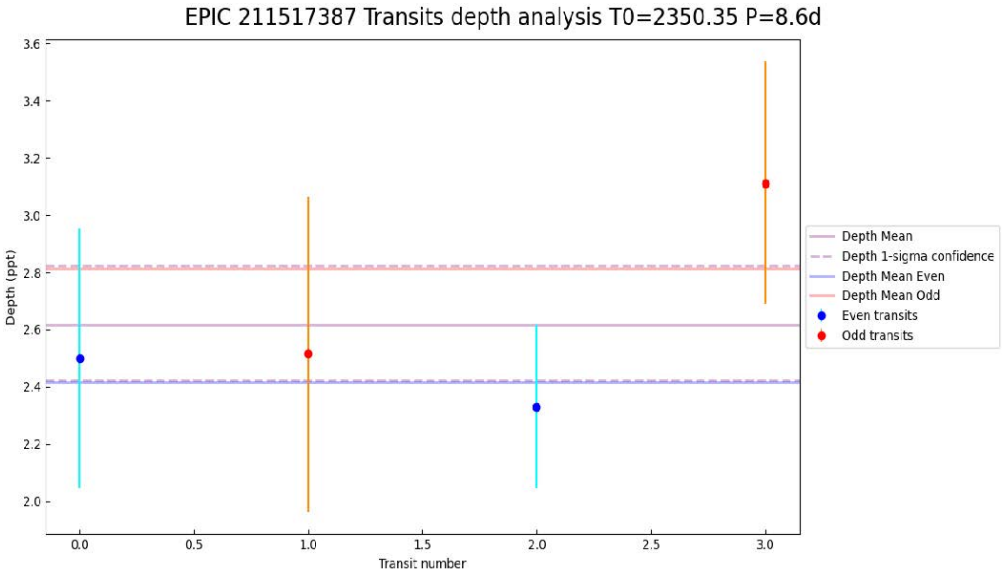


Figure 4.16: EPIC 211517387, first run, candidate single-transits depths plot.

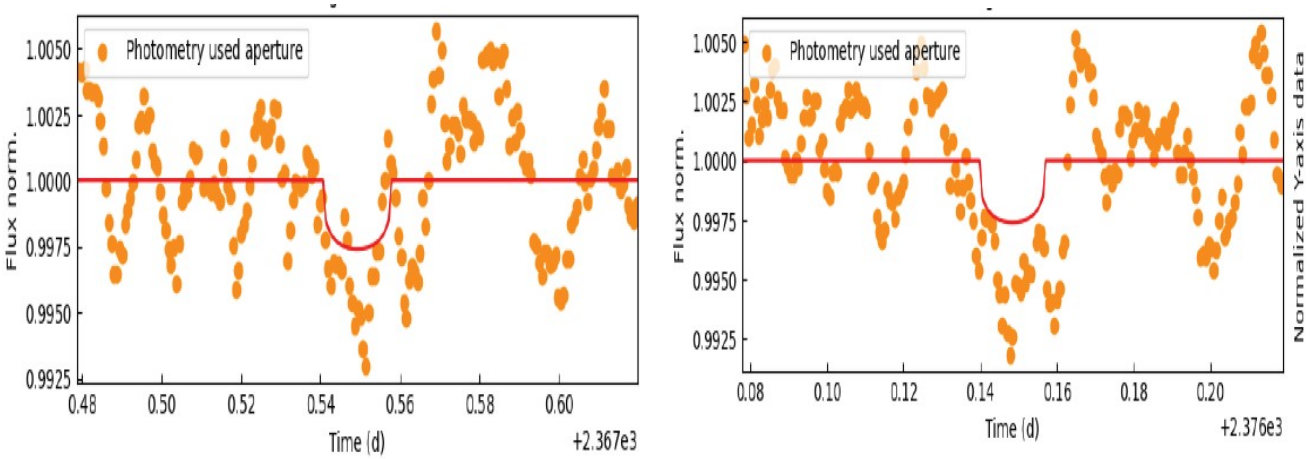


Figure 4.17: Second (left) and third (right) transit of the first candidate of EPIC 211517387.

Conclusion

In this work, we aimed to understand better the K2 data in order to use them for the search for exoplanets around hot subdwarfs. This goal was successfully achieved by selecting four K2 targets with different magnitudes and light curve quality and corrected by the K2sff and EVEREST pipelines. I performed injection-and-recovery tests on them and compared the two pipelines. I extracted the parameters influencing the quality of the detection.

I found that the detection of short-period Earth-size planets was possible both with EVEREST and K2sff. However, the injection-and-recovery tests gave better results with EVEREST, with sub-Earth-size planets ($\gtrsim 0.4 R_\oplus$) recovered. We found, for all the targets tested, that EVEREST is better than K2sff in all aspects: range of planet detectable, global recovery rates and SNR and SDE of the signal. For relatively bright hot subdwarfs ($K_p \lesssim 15$) the EVEREST's SNR and SDE values can be respectively up to $\simeq$2.6 and $\simeq$1.9 times the K2sff values. This value depends on the quality of the light curves because K2sff correction is really sensitive to the presence of systematics. However, these differences in SNR and SDE diminish when the magnitude increases and EVEREST become only $\simeq$11% better for the SNR and $\simeq$14% for the SDE for a $K_p = 16.9$ star.

The second part of this work was to apply the lesson learnt from the injection-and-recover tests. We performed a search of candidates with SHERLOCK over the 48 hot subdwarf targets available in the EVEREST data. Out of these targets, only nine had signals worth performing a vetting and out of them five were interesting. In these five, I found two particularly promising candidates around EPIC 212465180, which seems to correspond to a disintegrating object transit model with a period of about 3.01 days, and around EPIC 2145515136 with a transit period of 11.01 days.

In conclusion, this thesis defends the K2 mission had provided qualitative data and should be used for the search for exoplanets around hot subdwarfs. The EVEREST pipeline should be used as a priority as it provides better results for the search of Earth- and sub-Earth-planets. This work urges the continuation of the analysis of the targets observed by K2 and of EPIC 212465180 and EPIC 214515136 once the validation and fitting are integrated into SHERLOCK for this mission.

Bibliography

[1] U. Heber. "Hot Subdwarf Stars". In: *Annual Review of Astronomy and Astrophysics* 47.1 (2009), pp. 211–251. DOI: 10.1146/annurev-astro-082708-101836. eprint: https://doi.org/10.1146/annurev-astro-082708-101836.

[2] V. Van Grootel, F. J. Pozuelos, A. Thuillier, et al. "A search for transiting planets around hot subdwarfs. I. Methods and performance tests on light curves from Kepler, K2, TESS, and CHEOPS". In: *A&A* 650, A205 (June 2021), A205. DOI: 10.1051/0004-6361/202140381. arXiv: 2104.10462 [astro-ph.EP].

[3] U. Heber. "Hot Subluminous Stars". In: *Publications of the Astronomical Society of the Pacific* 128.966 (July 2016), p. 082001. DOI: 10.1088/1538-3873/128/966/082001.

[4] Z. Han, P. Podsiadlowski, P. F. L. Maxted, et al. "The origin of subdwarf B stars - I. The formation channels". In: *Monthly Notices of the Royal Astronomical Society* 336.2 (Oct. 2002), pp. 449–466. DOI: 10.1046/j.1365-8711.2002.05752.x. arXiv: astro-ph/0206130 [astro-ph].

[5] Z. Han, P. Podsiadlowski, P. F. L. Maxted, T. R. Marsh. "The origin of subdwarf B stars - II". In: *Monthly Notices of the Royal Astronomical Society* 341.2 (May 2003), pp. 669–691. DOI: 10.1046/j.1365-8711.2003.06451.x. arXiv: astro-ph/0301380 [astro-ph].

[6] R. F. Webbink. "Double white dwarfs as progenitors of R Coronae Borealis stars and type I supernovae." In: *The Astrophysical Journal* 277 (Feb. 1984), pp. 355–360. DOI: 10.1086/161701.

[7] Geier, S., Heber, U. "Hot subdwarf stars in close-up view - II. Rotational properties of single and wide binary subdwarf B stars". In: *A&A* 543 (2012), A149. DOI: 10.1051/0004-6361/201219463.

[8] S. Charpinet, N. Giammichele, W. Zong, et al. "Rotation in sdB stars as revealed by stellar oscillations". In: *Open Astronomy* 27.1 (2018), pp. 112–119. DOI: doi:10.1515/astro-2018-0012.

[9] J. Goldstein. "The fate of the earth in the red giant envelope of the sun". In: *Astronomy and Astrophysics* 178 (1987), pp. 283–285.

[10] E. Bear, N. Soker. "Connecting planets around horizontal branch stars with known exoplanets". In: *Monthly Notices of the Royal Astronomical Society* 411.3 (Feb. 2011), pp. 1792–1802. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.17808.x. eprint: https://academic.oup.com/mnras/article-pdf/411/3/1792/2943480/mnras0411-1792.pdf.

[11] C. E. O'Connor, L. Bildsten, M. Cantiello, D. Lai. "Giant Planet Engulfment by Evolved Giant Stars: Light Curves, Asteroseismology, and Survivability". In: *The Astrophysical Journal* 950.2 (June 2023), p. 128. DOI: 10.3847/1538-4357/acd2d4.

[12] A. Miglio, K. Brogaard, D. Stello, et al. "Asteroseismology of old open clusters with Kepler: direct estimate of the integrated red giant branch mass-loss in NGC 6791 and 6819". In: *Monthly Notices of the Royal Astronomical Society* 419.3 (Jan. 2012), pp. 2077–2088. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.19859.x. eprint: https://academic.oup.com/mnras/article-pdf/419/3/2077/18707561/mnras0419-2077.pdf.

[13] M. I. Jones, R. Wittenmyer, C. Aguilera-Gómez, et al. "Four Jovian planets around low-luminosity giant stars observed by the EXPRESS and PPPS". In: *A&A* 646, A131 (Feb. 2021), A131. DOI: 10.1051/0004-6361/202038555. arXiv: 2006.01277 [astro-ph.EP].

[14] R. Silvotti, S. Schuh, R. Janulis, et al. "A giant planet orbiting the 'extreme horizontal branch' star V391 Pegasi". In: *Nature* 449.7159 (Sept. 2007), pp. 189–191. DOI: 10.1038/nature06143.

[15] Silvotti, R., Schuh, S., Kim, S.-L., et al. "The sdB pulsating star V391 Peg and its putative giant planet revisited after 13 years of time-series photometric data". In: *A&A* 611 (2018), A85. DOI: 10.1051/0004-6361/201731473.

[16] R. Lutz, S. Schuh, R. Silvotti. "EXOTIME: Searching for planets and measuring {P} in sdB pulsators". In: *Astronomische Nachrichten* 333.10 (Dec. 2012), p. 1099. DOI: 10.1002/asna.201211793. arXiv: 1301.2048 [astro-ph.EP].

[17] F. Mackebrandt, S. Schuh, R. Silvotti, et al. "The EXOTIME project: signals in the O-C diagrams of the rapidly pulsating subdwarfs DW Lyn, V1636 Ori, QQ Vir, and V541 Hya". In: *Astronomy & Astrophysics* 638, A108 (June 2020), A108. DOI: 10.1051/0004-6361/201937172. arXiv: 2004.04764 [astro-ph.SR].

[18] S. Charpinet, G. Fontaine, P. Brassard, et al. "A compact system of small planets around a former red-giant star". In: *Nature* 480.7378 (Dec. 2011), pp. 496–499. DOI: 10.1038/nature10631.

[19] R. Silvotti, S. Charpinet, E. Green, et al. "Kepler detection of a new extreme planetary system orbiting the subdwarf-B pulsator KIC 10001893". In: *A&A* 570, A130 (Oct. 2014), A130. DOI: 10.1051/0004-6361/201424509. arXiv: 1409.6975 [astro-ph.EP].

[20] J. Krzesinski. "Planetary candidates around the pulsating sdB star KIC 5807616 considered doubtful". In: *A&A* 581, A7 (Sept. 2015), A7. DOI: 10.1051/0004-6361/201526346.

[21] A. Blokesz, J. Krzesinski, L. Kedziora-Chudczer. "Analysis of putative exoplanetary signatures found in light curves of two sdBV stars observed by Kepler". In: *Astronomy & Astrophysics* 627 (July 2019), A86. DOI: 10.1051/0004-6361/201835003.

[22] S. Geier, H. Edelmann, U. Heber, L. Morales-Rueda. "Discovery of a Close Substellar Companion to the Hot Subdwarf Star HD 149382—The Decisive Influence of Substellar Objects on Late Stellar Evolution". In: *The Astrophysical Journal* 702.1 (Sept. 2009), pp. L96–L99. DOI: 10.1088/0004-637X/702/1/L96. arXiv: 0908.1025 [astro-ph.SR].

[23] J. Norris, J. Wright, R. Wade, et al. "Non-detection of the putative substellar companion to HD 149382". English (US). In: *Astrophysical Journal* 743.1 (Dec. 2011). ISSN: 0004-637X. DOI: 10.1088/0004-637X/743/1/88.

[24] V. Schaffenroth, S. Geier, U. Heber2016, et al. "The MUCHFUSS photometric campaign". In: *A&A* 614, A77 (June 2018), A77. DOI: 10.1051/0004-6361/201629789. arXiv: 1802.03260 [astro-ph.SR].

[25] V. Schaffenroth, B. N. Barlow, S. Geier, et al. "The EREBOS project: Investigating the effect of substellar and low-mass stellar companions on late stellar evolution. Survey, target selection, and atmospheric parameters". In: *A&A* 630, A80 (Oct. 2019), A80. DOI: 10.1051/0004-6361/201936019. arXiv: 1907.09892 [astro-ph.SR].

[26] V. Schaffenroth, I. Pelisoli, B. N. Barlow, et al. "Hot subdwarfs in close binaries observed from space. I. Orbital, atmospheric, and absolute parameters, and the nature of their companions". In: *A&A* 666, A182 (Oct. 2022), A182. DOI: 10.1051/0004-6361/202244214. arXiv: 2207.02001 [astro-ph.SR].

[27] V. Schaffenroth, B. N. Barlow, I. Pelisoli, et al. "Hot subdwarfs in close binaries observed from space. II. Analyses of the light variations". In: *A&A* 673, A90 (May 2023), A90. DOI: 10.1051/0004-6361/202244697. arXiv: 2302.12507 [astro-ph.SR].

[28] K. Beuermann, S. Dreizler, F. V. Hessman, J. Deller. "The quest for companions to post-common envelope binaries. III. A reexamination of HW Virginis". In: *A&A* 543, A138 (July 2012), A138. DOI: 10.1051/0004-6361/201219391. arXiv: 1206.3080 [astro-ph.SR].

[29] S. .-. Qian, L. Liu, L. .-. Zhu, et al. "A circumbinary planet in orbit around the short-period white dwarf eclipsing binary RR Cae". In: *Monthly Notices of the Royal Astronomical Society: Letters* 422.1 (May 2012), pp. L24–L27. DOI: 10.1111/j.1745-3933.2012.01228.x. arXiv: 1201.4205 [astro-ph.SR].

[30] M. Zorotovic, M. R. Schreiber. "Origin of apparent period variations in eclipsing post-common-envelope binaries". In: *A&A* 549, A95 (Jan. 2013), A95. DOI: 10.1051/0004-6361/201220321. arXiv: 1211.5356 [astro-ph.SR].

[31] D. R. G. Schleicher, S. Dreizler. "Planet formation from the ejecta of common envelopes". In: *A&A* 563, A61 (Mar. 2014), A61. DOI: 10.1051/0004-6361/201322860. arXiv: 1312.3479 [astro-ph.EP].

[32] T. R. Marsh. "Circumbinary Planets Around Evolved Stars". In: 96 (2018). Ed. by H. J. Deeg, J. A. Belmonte, p. 96. DOI: 10.1007/978-3-319-55333-7_96.

[33] J. N. Winn. *Transits and Occultations*. 2014. arXiv: 1001.2010 [astro-ph.EP].

[34] NASA. *Exoplanet exploration - Exoplanet Catalog*. https://exoplanets.nasa.gov/discovery/exoplanet-catalog/. 2023.

[35] M. Gillon, A. H. M. J. Triaud, B.-O. Demory, et al. "Seven temperate terrestrial planets around the nearby ultracool dwarf star TRAPPIST-1". In: *Nature* 542.7642 (Feb. 2017), pp. 456–460. DOI: 10.1038/nature21360.

[36] NASA. *NASA Exoplanet Archive, K-25b*. https://exoplanetarchive.ipac.caltech.edu/overview/K2-25%20b#planet_K2-25-b_collapsible.

[37] W. J. Borucki, D. G. Koch, J. Lissauer, et al. "KEPLER Mission Status". In: *Transiting Extrapolar Planets Workshop*. Ed. by C. Afonso, D. Weldrake, T. Henning. Vol. 366. Astronomical Society of the Pacific Conference Series. July 2007, p. 309.

[38] W. J. Borucki. "Space Missions for Exoplanet Science: Kepler/K2". In: 80 (2018). Ed. by H. J. Deeg, J. A. Belmonte, p. 80. DOI: 10.1007/978-3-319-55333-7_80.

[39] D. G. Koch, W. Borucki, E. Dunham, et al. "Overview and status of the Kepler Mission". In: *Optical, Infrared, and Millimeter Space Telescopes*. Ed. by J. C. Mather. Vol. 5487. International Society for Optics and Photonics. SPIE, 2004, pp. 1491–1500. DOI: 10.1117/12.552346.

[40] S. B. Howell, C. Sobeck, M. Haas, et al. "The K2 Mission: Characterization and Early Results". In: *Publications of the Astronomical Society of the Pacific* 126.938 (Apr. 2014), p. 398. DOI: 10.1086/676406.

[41] J. E. V. Cleve, S. B. Howell, J. C. Smith, et al. "That's How We Roll: The NASA K2 Mission Science Products and Their Performance Metrics". In: *Publications of the Astronomical Society of the Pacific* 128.965 (June 2016), p. 075002. DOI: 10.1088/1538-3873/128/965/075002.

[42] W. Zhu, S. Dong. "Exoplanet Statistics and Theoretical Implications". In: *Annual Review of Astronomy and Astrophysics* 59.1 (Sept. 2021), pp. 291–336. ISSN: 1545-4282. DOI: 10.1146/annurev-astro-112420-020055.

[43] NASA. *Keplergo.github.io*. https://keplergo.github.io/KeplerScienceWebsite/publications.html. 2020.

[44] A. Vanderburg, J. A. Johnson. "A Technique for Extracting Highly Precise Photometry for the Two-Wheeled Kepler Mission". In: *Publications of the Astronomical Society of the Pacific* 126.944 (Oct. 2014), p. 948. DOI: 10.1086/678764. arXiv: 1408.3853 [astro-ph.IM].

[45] R. Luger, E. Agol, E. Kruse, et al. "EVEREST: PIXEL LEVEL DECORRELATION OF K2 LIGHT CURVES". In: *The Astronomical Journal* 152.4 (Oct. 2016), p. 100. DOI: 10.3847/0004-6256/152/4/100.

[46] R. Luger, E. Kruse, D. Foreman-Mackey, et al. "An Update to the EVEREST K2 Pipeline: Short Cadence, Saturated Stars, and Kepler-like Photometry Down to Kp = 15". In: *The Astronomical Journal* 156.3 (Aug. 2018), p. 99. DOI: 10.3847/1538-3881/aad230.

[47] F. J. Pozuelos, J. C. Suárez, G. C. de Elıa, et al. "GJ 273: on the formation, dynamical evolution, and habitability of a planetary system hosted by an M dwarf at 3.75 parsec". In: *A&A* 641, A23 (Sept. 2020), A23. DOI: 10.1051/0004-6361/202038047. arXiv: 2006.09403 [astro-ph.EP].

[48] F.J. Pozuelos and M. Dévora-Pajares. *SHERLOCK Github*. https://github.com/franpoz/SHERLOCK.

[49] Space Telescope Science Institute.s. *MAST: Barbara A. Mikulski Archive for Space Telescopes*. https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html.

[50] F. J. Pozuelos, M. Dévora-Pajares, A. Thuillier, et al. "SHERLOCK: A python pipeline to explore space-based observations in the search for planets". In: EPSC2022-970 (Sept. 2022), EPSC2022–970. DOI: 10.5194/epsc2022-970.

[51] M. Hippke, T. J. David, G. D. Mulders, R. Heller. "Wōtan: Comprehensive Time-series Detrending in Python". In: *The Astronomical Journal* 158.4, 143 (Oct. 2019), p. 143. DOI: 10.3847/1538-3881/ab3984. arXiv: 1906.00966 [astro-ph.EP].

[52] Hippke, Michael, Heller, René. "Optimized transit detection algorithm to search for periodic transits of small planets". In: *A&A* 623 (2019), A39. DOI: 10.1051/0004-6361/201834672.

[53] F.J. Pozuelos and M. Dévora-Pajares. *SHERLOCK Documentation*. https://sherlock-ph.readthedocs.io/en/latest/index.html.

[54] S. Giacalone, C. D. Dressing, E. L. N. Jensen, et al. "Vetting of 384 TESS Objects of Interest with TRICERATOPS and Statistical Validation of 12 Planet Candidates". In: *The Astronomical Journal* 161.1 (Dec. 2020), p. 24. DOI: 10.3847/1538-3881/abc6af.

[55] M. N. Günther, T. Daylan. "Allesfitter: Flexible Star and Exoplanet Inference from Photometry and Radial Velocity". In: *The Astrophysical Journal Supplement Series* 254.1 (Apr. 2021), p. 13. DOI: 10.3847/1538-4365/abe70e.

[56] M. Dévora-Pajares, F. J. Pozuelos. "MATRIX: Multi-phAse Transits Recovery from Injected eXoplanets". In: 65, 65770831 (May 2022), p. 65770831. DOI: 10.5281/zenodo.6570831.

[57] N. Schanche, F. J. Pozuelos, M. N. Günther, et al. "TOI-2257 b: A highly eccentric long-period sub-Neptune transiting a nearby M dwarf". In: *A&A* 657, A45 (Jan. 2022), A45. DOI: 10.1051/0004-6361/202142280. arXiv: 2111.01749 [astro-ph.EP].

[58] M. Brogi, C. U. Keller, M. de Juan Ovelar, et al. "Evidence for the disintegration of KIC 12557548 b". In: *A&A* 545 (Sept. 2012), p. L5. DOI: 10.1051/0004-6361/201219762.

[59] R. Sanchis-Ojeda, S. Rappaport, E. Pallè, et al. "The K2-esprint project. i. Discovery of the disintegrating rocky planet K2-22b with a cometary head and leading tail". In: *The Astrophysical Journal* 812.2 (Oct. 2015), p. 112. DOI: 10.1088/0004-637x/812/2/112.

[60] S. Charpinet, E. M. Green, A. Baglin, et al. "CoRoT opens a new era in hot B subdwarf asteroseismology. Detection of multiple g-mode oscillations in KPD 0629-0016". In: *A&A* 516, L6 (June 2010), p. L6. DOI: 10.1051/0004-6361/201014789.

Appendix A

List of Hot Sudwarfs Observed by K2 and Available in EVEREST

A.1 Hot Subdwarf observed by K2

EPIC	Class	Other Name	Kp	RA	DEC
		sdB pulsators			
220641886	sdB	HD 4539	10.40	11.8717	9.9821
228755638	sdB+dM	HW Vir	10.76	191.0840	-8.6713
211623711	He-sdB	UVO 0825+15	11.89	127.1369	14.8673
220376019	sdB+WD	PG 0101+039	12.11	16.0903	4.2270
220422705	sdB+G	PG 0039+049	12.87	10.5255	5.1565
249942493	sdB	EC 15103-1557	12.89	228.2930	-16.1391
211779126	sdB	2M0856+1701	12.92	134.2053	17.0208
246387816	sdB+dM	EQ Psc	12.92	353.6440	-1.3269
246023959	sdB+dM	PHL 457	13.04	349.8520	-8.8772
211881419	iHe-sdB	PG 0848+186	13.30	132.9353	18.4563
201203416	sdB	PG 1156-037	13.46	179.8420	-4.0241
248411044	sdB	UY Sex	13.56	162.5118	-0.0102
246141920	sdB	PHL 531	13.99	351.9700	-6.2321
211433013	sdB+WD	LT Cnc	14.02	137.6060	12.1409
211765471	sdB+WD	HZ Cnc	14.04	133.3486	16.8265
220614972	sdB+F	PG 0048+091	14.29	12.8622	9.3591

211392098	sdB+MS	SDSS J082517.99+113106.3	14.34	126.3250	11.5184
211437457	sdB	PG 0902+124	14.73	136.4200	12.2078
246683636	sdB+dM	V1405 Ori	15.07	71.2371	14.3639
248368659	sdB+WD	VPHAS J181343.0-213843.9	15.10	273.4290	-21.6455
212508753	sdB+F7	PG 1315-123	15.13	199.4134	-12.5479
211823779	sdB+F1	SDSS J082003.35+173914.2	15.22	125.0140	17.6540
212475716	sdB+MS	EC 13356-1300	15.24	204.5730	-13.2569
211696659	sdB+WD	SDSS J083603.98+155216.4	15.50	129.0166	15.8712
212707862	sdB	SDSS J135544.71-080354.3	15.55	208.9363	-8.0651
212204284	sdB	PG 0843+246	15.64	131.5500	24.4196
246283223	sdB	HE 2307-0340	15.66	347.6000	-3.4003
248368658	sdB		15.70	268.7680	-24.8137
218717602	sdB		15.76	293.4455	-18.2871
211938328	sdB+F6	LB 378	15.78	129.0501	19.2989
218366972	sdB+WD		15.94	293.7241	-18.9313
201206621	sdB+WD	PG 1142-037	15.99	176.2385	-3.9481
212487276	sdB	EC 13359-1245	16.23	204.6540	-13.0086
217280630	sdB		16.33	288.8940	-20.8521
215776487	sdB		16.35	295.4100	-23.5619
203948264	sdB		16.70	252.4843	-24.2929
246373305	iHe-sdB	PHL 417	16.88	347.7710	-1.6183
251668197	sdB	EC 15094-1725	17.00	228.0600	-17.6116
229002689	sdB	SDSS J122057.48-012642.3	18.65	185.2400	-1.4451
Pulsators in LC only					
220188903	sdBV+WD	PB 6373	14.91	19.7383	-0.4296
230195595	sdB		15.59	256.1310	-17.9435
Non-pulsators in Binaries					
220468352	sdB+F	PB 6355	13.01	19.1137	6.0532
251377113	sdB+F/G	SDSS J090827.24+231417.9	13.53	137.1130	23.2383
211499370	sdB+F/G/K	SDSS J082556.80+130753.5	14.60	126.4867	13.1316
218637228	sdB+F/G		14.79	293.1550	-18.4384

227441033	sdB+F/G		15.10	262.9540	-18.1450
216924452	sdB+F/G		15.53	290.8680	-21.4814
250121838	sdB+F/G/K	EC 15365-1350	15.74	234.8290	-14.0015
246151922	sdB+G9	HE 2322-0617	15.74	351.3830	-6.0199
212630158	sdB+F/G		15.75	198.4050	-9.9040
246868556	sdB+F/G	GALEX J050252.2+162647	15.78	75.7179	16.4467
246864591	sdB+F/G/K	KUV 04571+1620	15.98	74.9971	16.4060
211910684	sdB+F/G	PG 0906+191	15.99	137.2220	18.8861
212108396	sdB+F/G	SDSS J082447.30+221112.9	16.02	126.1971	22.1869
211400847	sdB+F/G	SDSS J084447.00+113910.0	16.43	131.1959	11.6528
212003762	sdB+F/G	SDSS J081406.79+201901.7	16.51	123.5280	20.3171
212137838	sdB+F/G	Ton 920	16.54	127.2610	22.7769
250152590	sdB+F/G/K	LB 889	17.13	235.8930	-13.5722
248467942	sdB+F/G	SDSS J103022.07+020524.3	17.24	157.5920	2.0901
211732575	sdB+F/G	SDSS J082426.51+162145.1	17.68	126.1100	16.3626
251583165	sdB+F/G	SDSS J131932.19-014131.2	18.24	199.8840	-1.6920
212866280	sdB+F/G	SDSS J133701.51-031732.2	18.27	204.2560	-3.2923
212410755	sdB+WD	EC 13332-1424	13.46	203.9729	-14.6703
211437457	sdB+WD	PG 0902+124	14.73	136.4200	12.2078
201535046	sdB+?	PG 1049+013	14.44	163.1180	1.0627
251372905	sdOB+F/G	SDSS J091216.06+225452.7	15.30	138.0670	22.9147
211904152	sdOB+F/G	PG 0912+189	15.93	138.7800	18.7879
248767552	sdOB+WD?	SDSS J101833.11+095336.1	14.97	154.6380	9.8934
246877984	sdOB+WD	KUV 05053+1628	16.11	77.0537	16.5379
Non-pulsators apparently single					
234319842	sdB	-	12.97	260.6050	-19.9098
260017832	sdB	PG 2349+002	13.27	357.9720	0.4716
211708181	sdB	GALEX J081233.6+160121	13.77	123.1403	16.0233
227389858	sdB	-	13.79	262.3680	-18.2718
246230928	sdB	PHL 529	13.93	351.7870	-4.4103
206535752	sdB	PHL 358	13.99	339.6820	-4.1248

201648341	sdB	PG 1214+031	14.04	184.2040	2.7905
217204898	sdB	-	14.26	288.8780	-20.9864
246643895	sdB	HS 0446+1344	14.50	72.2846	13.8335
212722777	sdB	PG 1330-074	14.93	203.3190	-7.6939
211727748	sdB	PG 0838+165	14.99	130.3450	16.2929
206073023	sdB	BPS CS 29512-38	15.00	334.1720	-12.8063
210837690	sdB		15.11	57.7850	20.4437
212498842	sdB	EC 13162-1229	15.26	199.7180	-12.7583
212465180	sdB	EC 13265-1313	15.56	202.3010	-13.4788
212160066	sdB	SDSS J082445.68+231520.3	15.57	126.1900	23.2557
246901153	sdB	KUV 04369+1640	15.70	69.9429	16.7644
249601610	sdB	EC 15050-2017	15.71	226.9690	-20.4839
246980092	sdB	KUV 04482+1727	15.74	72.7737	17.5303
218148570	sdB	-	15.74	292.1370	-19.3246
228914323	sdB	PG 1249-028	15.76	193.1230	-3.0249
228682488	sdB	SDSS J085217.70+211637.4	16.00	133.0730	21.2769
212818294	sdB	PG 1356-047	16.15	209.7058	-4.9634
248422838	sdB	PG 1032+007	16.27	158.8030	0.4581
214515136	sdB	-	16.30	289.0660	-26.3030
251603936	sdB	SDSS J131916.15-011405.0	16.69	199.8170	-1.2348
201531672	sdB	SDSS J112757.48+010044.2	16.89	171.9900	1.0123
251457058	sdB	SDSS J105428.85+010514.7	17.10	163.6200	1.0874
246371369	sdB	PB 5212	17.11	345.2060	-1.6566
211552072	sdB	SDSS J084556.85+135211.3	17.50	131.4870	13.8698
212567176	sdB	HE 1309-1102	17.65	198.0100	-11.3042
249585191	sdB	EC 15064-2029	17.95	227.3340	-20.6851
248840987	sdB	SDSS J102050.99+114024.3	18.15	155.2120	11.6734
248810568	sdOB	SDSS J110055.94+105542.3	14.22	165.2330	10.9284
246997679	sdOB	KUV 05109+1739	14.58	78.4575	17.7007
211421561	sdOB	SDSS J090042.68+115749.9	14.90	135.1780	11.9639
220265912	sdOB	PG 0055+016	15.19	14.6029	1.9098

249700050	sdOB	EC 15059-1902	15.65	227.1900	-19.2293
206240954	sdOB	SDSS J220337.88-090733.5	16.31	330.9080	-9.1260
210731139	sdOB	SDSS J032427.24+184918.2	16.37	51.1133	18.8218
246087406	sdOB	PB 7470	16.46	345.4930	-7.4136
206186190	sdOB	BPS CS 22886-65	16.49	336.4680	-9.9993
251605347	sdOB	SDSS J133611.02-011156.0	18.69	204.0460	-1.1989
246745570	He-sdB	KUV 04456+1502	15.68	72.1242	15.1277
211920209	He-sdB	PG 0850+192	16.39	133.4220	19.0283
249770424	He-sdOB	GALEX J152332.2-181726	14.00	230.8850	-18.2906
211495446	He-sdOB	PG 0838+133	14.03	130.4327	13.0750
248748173	He-sdOB	PG 1033+097	16.38	158.9570	9.4311
248761152	He-sdOB	PG 1045+100	17.09	161.9130	9.7401
248915544	He-sdOB	SDSS J103806.64+134412.1	17.21	159.5270	13.7367
212762631	sdO	PG 1355-064	13.76	209.4761	-6.6255
220179214	sdO	GD 934	14.93	18.5775	-0.8204
248520995	sdO	SDSS J110053.55+034622.8	17.25	165.2230	3.7730
211517387	sdO	SDSS J082944.74+132302.5	17.32	127.4370	13.3841
249862817	sdO	EC 15447-1656	18.05	236.8930	-17.0947
228821386	He-sdO	PG 1220-056	14.86	185.7460	-5.8847
249867379	He-sdO	EC 15348-1652	15.35	234.4180	-17.0376
205247324	He-sdO	-	16.01	251.7740	-18.6511
201640895	He-sdO	SDSS J110215.45+024034.2	17.60	165.5650	2.6762
228960704	He-sdO	SDSS J123821.48-021211.4	18.49	189.5900	-2.2032
NOV in LC only					
211602914	sd	SDSS J082959.28+143441.8	15.64	127.4970	14.5783
201150341	sdB	HE 1140-0500	14.50	175.7410	-5.2872
214958569	sdB		15.70	289.5300	-25.2509
216775790	sdB		16.50	290.4090	-21.7461
201236182	sdB	PG 1154-031	16.59	179.2250	-3.4195
211720816	sdB	SDSS J083901.50+161148.0	16.71	129.7560	16.1967
211594465	sdB	SDSS J081931.22+142756.1	17.19	124.8800	14.4656

248912731	sdB	SDSS J103832.41+133848.3	17.44	159.6350	13.6468
201201339	sdB	SDSS J112757.48+010044.2	17.50	172.7900	-4.0722
201590024	sdB	SDSS J113418.00+015322.1	17.65	173.5750	1.8895
201698091	sdB	SDSS J114821.29+033625.7	17.70	177.0890	3.6072
229021782	sdB	SDSS J125410.86-010408.3	17.72	193.5450	-1.0690
228682339	sdB	SDSS J082824.20+212556.7	17.73	127.1010	21.4324
251457060	sdB	SDSS J104725.10+010847.2	17.80	161.8550	1.1464
248783069	sdB	SDSS J104620.14+101629.7	18.65	161.5840	10.2749
251410019	sdB	SDSS J085809.09+252134.6	18.87	134.5380	25.3596
201424163	sdB+WD	PG 1136-003	15.96	174.6700	-0.5922
228682347	sdB+WD	SDSS J083139.68+162316.4	17.91	127.9150	16.3879
248783744	sdB+WD	SDSS J103218.40+101725.8	18.82	158.0770	10.2905
211460944	sdB+WD ?	SDSS J084556.85+135211.3	15.36	134.6160	12.5664
228796212	sdBO	SDSS J124446.64-065625.8	18.83	191.1940	-6.9405
211991114	sdBO+F/G	Ton 914	15.10	126.2800	20.1104
211930840	He-sdB	SDSS J091512.06+191114.6	19.13	138.8000	19.1874
201734164	sdOA	PG 1110+045	14.84	168.3222	4.2207
213545287	sdOB	GALEX J191509.0-290311	15.00	288.7880	-29.0527
201924421	sdOB	SDSS J113218.41+075103.0	17.20	173.0770	7.8509
228682323	sdOB	SDSS J082110.89+183924.1	17.84	125.2950	18.6567
212034957	sdOB	SDSS J090302.39+205008.9	18.62	135.7600	20.8358
215669184	He-sdOB	GALEX J193323.6-234553	15.00	293.3490	-23.7647
201802867	He-sdOB	SDSS J111633.29+052507.9	17.80	169.1390	5.4189
251383153	He-sdOB	SDSS J091044.90+234044.6	18.27	137.6870	23.6791
229155531	He-sdOB	SDSS J121643.72+020835.9	18.73	184.1820	2.1433
251357585	He-sdOB	SDSS J092245.79+214238.9	19.01	140.6910	21.7108
213716821	sdO	GALEX J192041.4-282939	13.40	290.1720	-28.4945
214453765	sdO	GALEX J191158.1-262712	15.30	287.9920	-26.4534
231422890	sdO		17.07	257.1200	-22.2940
201418759	sdO	SDSS J111438.57-004024.3	18.10	168.6610	-0.6734
201843731	sdO	SDSS J115009.48+061042.1	18.10	177.5400	6.1784

211559083	sdO	SDSS J084421.10+135807.6	18.18	131.0880	13.9688
216747137	sdO+dM	2MASS J18521800-2147506	13.87	283.0750	-21.7974
217750936	sdO+dM?		16.70	290.2270	-20.0265
246735349	He-sdO	KUV 04402+1455	13.97	70.7638	15.0067
216452306	He-sdO		16.40	289.8920	-22.3225
228682365	He-sdO	SDSS J083747.23+194955.9	18.60	129.4470	19.8322

A.2 Hot Subdwarfs Light Curves availability in EVEREST

EPIC	Class	Campaign (SC)	Campaign (LC)	SC Everest	LC Everest
		sdB pulsators		**17**	**37**
220641886	sdB	8	8	1	1
228755638	sdB+dM	10 (101-102)	10 (101-102)	0	1
211623711	He-sdB	5;18	5;18	1	1
220376019	sdB+WD	8	8	1	1
220422705	sdB+G	8	8	1	1
249942493	sdB	15	15	0	1
211779126	sdB	5;18	5;18	1	1
246387816	sdB+dM	12	12	0	1
246023959	sdB+dM	12	12	0	1
211881419	iHe-sdB	16;18	5;16;18	0	1
201203416	sdB	10 (101-102)	10 (101-102)	0	1
248411044	sdB	14	14	0	1
246141920	sdB	12	12	0	1
211433013	sdB+WD	16	16	0	1
211765471	sdB+WD	5;16;18	5;16;18	1	1
220614972	sdB+F	8	8	1	1
211392098	sdB+MS	18	5, 18	0	1
211437457	sdB	16	16	0	1
246683636	sdB+dM	13	13	0	1

248368659	sdB+WD	9 (91-92)	9 (91-92)	0	0
212508753	sdB+F7	6;17	6;17	1	1
211823779	sdB+F1	5;18	5;18	1	1
212475716	sdB+MS	17	17	0	1
211696659	sdB+WD	5;18	5;18	1	1
212707862	sdB	6;17	6;17	1	1
212204284	sdB	16	16	0	1
246283223	sdB	12	12	0	1
248368658	sdB	9 (91-92)	9 (91-92)	0	0
218717602	sdB	7	7	0	1
211938328	sdB+F6	5;18	5;18	1	1
218366972	sdB+WD	7	7	1	1
201206621	sdB+WD	1	1	1	1
212487276	sdB	17	17	0	1
217280630	sdB	7	7	1	1
215776487	sdB	7	7	1	1
203948264	sdB	2	2	1	1
246373305	iHe-sdB	12	12	0	1
251668197	sdB	15	15	0	1
229002689	sdB	10 (101-102)	10 (101-102)	0	1
Pulsators in LC only				**0**	**2**
220188903	sdBV+WD	no data	8		1
230195595	sdB	no data	11		1
Non-pulsators in Binaries				**9**	**28**
220468352	sdB+F	8	8	1	1
251377113	sdB+F/G	16	16	0	1
211499370	sdB+F/G/K	5	5;18	1	1
218637228	sdB+F/G	7	7	1	1
227441033	sdB+F/G	11 (111-112)	11 (111-112)	0	1
216924452	sdB+F/G	7	7	1	1
250121838	sdB+F/G/K	15	15	0	1

246151922	sdB+G9	12	12	0	1
212630158	sdB+F/G	6	6	1	1
246868556	sdB+F/G	13	13	0	1
246864591	sdB+F/G/K	13	13	0	1
211910684	sdB+F/G	16	16	0	1
212108396	sdB+F/G	5	5;18	1	1
211400847	sdB+F/G	5	5;18	1	1
212003762	sdB+F/G	18	18	0	1
212137838	sdB+F/G	5	5	1	1
250152590	sdB+F/G/K	15	15	0	1
248467942	sdB+F/G	14	14	0	1
211732575	sdB+F/G	18	18	0	1
251583165	sdB+F/G	17	17	0	1
212866280	sdB+F/G	17	17	0	1
212410755	sdB+WD	6	6	1	1
211437457	sdB+WD	16	16	0	1
201535046	sdB+?	14	14	0	1
251372905	sdOB+F/G	16	16	0	1
211904152	sdOB+F/G	16	16	0	1
248767552	sdOB+WD?	14	14	0	1
246877984	sdOB+WD	13	13	0	1
Non-pulsators apparently single				**22**	**59**
234319842	sdB	11 (111-112)	11 (111-112)	0	1
260017832	sdB	T		Not found	0
211708181	sdB	5	5	1	1
227389858	sdB	11 (111-112)	11 (111-112)	0	1
246230928	sdB	12	12	0	1
206535752	sdB	3	3	1	1
201648341	sdB	10 (101-102)	10 (101-102)	0	1
217204898	sdB	7	7	1	1
246643895	sdB	13	13	0	1

212722777	sdB	17	17	0	1
211727748	sdB	5;16	5;16	1	1
206073023	sdB	3	3	1	1
210837690	sdB	4	4	1	1
212498842	sdB	6	6	1	1
212465180	sdB	6	6	1	1
212160066	sdB	18	5;18	0	1
246901153	sdB	13	13	0	1
249601610	sdB	15	15	0	1
246980092	sdB	13	13	0	1
218148570	sdB	7	7	1	1
228914323	sdB	10 (101-102)	10 (101-102)	0	1
228682488	sdB	16	16	0	1
212818294	sdB	6;17	6;17	1	1
248422838	sdB	14	14	0	1
214515136	sdB	7	7	1	1
251603936	sdB	17	17	0	1
201531672	sdB	1	1	1	1
251457058	sdB	14	14	0	1
246371369	sdB	12	12	0	1
211552072	sdB	16	16	0	1
212567176	sdB	6	6	1	1
249585191	sdB	15	15	0	1
248840987	sdB	14	14	0	1
248810568	sdOB	14	14	0	1
246997679	sdOB	13	13	0	1
211421561	sdOB	16	16	0	1
220265912	sdOB	8	8	1	1
249700050	sdOB	15	15	0	1
206240954	sdOB	3	3	1	1
210731139	sdOB	4	4	1	1

246087406	sdOB	12	12	0	1
206186190	sdOB	3	3	1	1
251605347	sdOB	17	17	0	1
246745570	He-sdB	13	13	0	1
211920209	He-sdB	18	5; 16; 18	0	1
249770424	He-sdOB	15	15	0	1
211495446	He-sdOB	5;16	5;16	1	1
248748173	He-sdOB	14	14	0	1
248761152	He-sdOB	14	14	0	1
248915544	He-sdOB	14	14	0	1
212762631	sdO	6	6,17	1	1
220179214	sdO	8	8	1	1
248520995	sdO	14	14	0	1
211517387	sdO	5	5	1	1
249862817	sdO	15	15	0	1
228821386	He-sdO	10 (101-102)	10 (101-102)	0	1
249867379	He-sdO	15	15	0	1
205247324	He-sdO	2	2	1	1
201640895	He-sdO	14	14	0	1
228960704	He-sdO	10 (101-102)	10 (101-102)	0	1
		NOV in LC only		**0**	**44**
211602914	sd	no data	5; 18		1
201150341	sdB	no data	1		1
214958569	sdB	no data	7		1
216775790	sdB	no data	7		1
201236182	sdB	no data	1		1
211720816	sdB	no data	5; 16; 18		1
211594465	sdB	no data	5; 18		1
248912731	sdB	no data	14		1
201201339	sdB	no data	1		1
201590024	sdB	no data	1		1

201698091	sdB	no data	1	1
229021782	sdB	no data	10	1
228682339	sdB	no data	5; 16; 18	1
251457060	sdB	no data	14	1
248783069	sdB	no data	14	1
251410019	sdB	no data	16	1
201424163	sdB+WD	no data	1	1
228682347	sdB+WD	no data	5	1
248783744	sdB+WD	no data	14	1
211460944	sdB+WD ?	no data	16	1
228796212	sdBO	no data	10	1
211991114	sdBO+F/G	no data	5; 18	1
211930840	He-sdB	no data	16	1
201734164	sdOA	no data	1	1
213545287	sdOB	no data	7	1
201924421	sdOB	no data	1	1
228682323	sdOB	no data	5	1
212034957	sdOB	no data	16	1
215669184	He-sdOB	no data	7	1
201802867	He-sdOB	no data	1	1
251383153	He-sdOB	no data	16	1
229155531	He-sdOB	no data	10	1
251357585	He-sdOB	no data	16	1
213716821	sdO	no data	7	1
214453765	sdO	no data	7	1
231422890	sdO	no data	11	1
201418759	sdO	no data	1	1
201843731	sdO	no data	1	1
211559083	sdO	no data	16	1
216747137	sdO+dM	no data	7	1
217750936	sdO+dM?	no data	7	1

246735349	He-sdO	no data	13	1
216452306	He-sdO	no data	7	1
228682365	He-sdO	no data	5	1

Appendix B

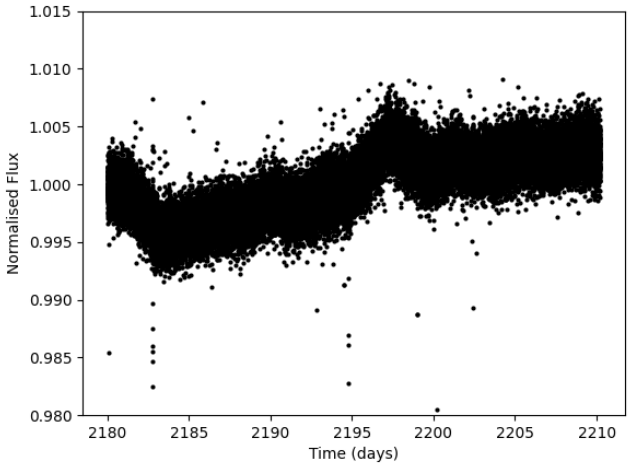
List of K2 flags

0x00000001	Attitude tweak	MUST BE REMOVED
0x00000002	Safe mode	MUST BE REMOVED
0x00000004	Coarse point	MUST BE REMOVED
0x00000008	Earth point	MUST BE REMOVED
0x00000010	RW Zero crossing	
0x00000020	RW Desaturation event	MUST BE REMOVED
0x00000040	Argabrightening	MUST BE REMOVED
0x00000080	Cosmic ray in aperture	REMOVE IF CONSERVATIVE
0x00000100	Manual exclude	MUST BE REMOVED
0x00000200	Unused	
0x00000400	Sudden sensitivity dropout	REMOVE IF CONSERVATIVE
0x00000800	Impulsive outlier	DECIDED TO REMOVE IT
0x00001000	Argabrightening	MUST BE REMOVED
0x00002000	Colateral cosmic ray	REMOVE IF CONSERVATIVE
0x00004000	Detector anomaly	MUST BE REMOVED
0x00008000	No fine point	MUST BE REMOVED
0x00010000	No data	MUST BE REMOVED
0x00020000	Rolling band in aperture	
0x00040000	Rolling band in full mask	
0x00080000	Possible thruster firing	MUST BE REMOVED
0x00100000	Thruster firing	MUST BE REMOVED
0x00200000	Unknown	
0x00400000	Flagged in the raw K2 TPF	
0x00800000	Data point is a NaN	MUST BE REMOVED
0x01000000	Determined to be an outlier	MUST BE REMOVED
0x02000000	Unused	
0x04000000	Data point is during transit/eclipse	

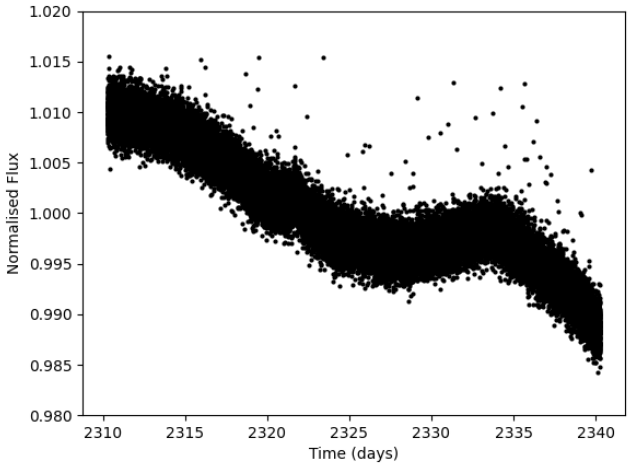
Appendix C

Lightcurves of the Tested Targets in MATRIX

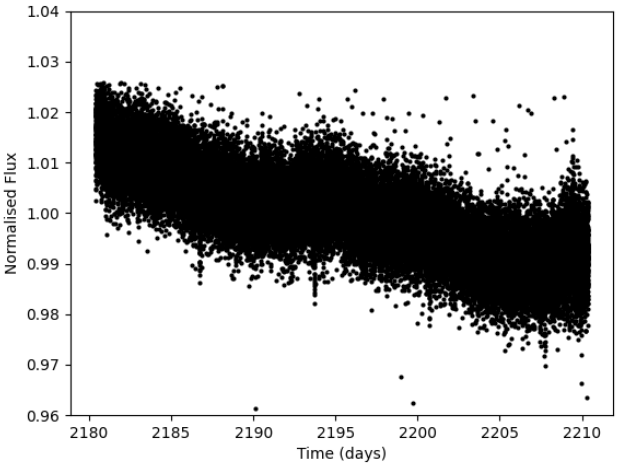
C.1 EVEREST light curves



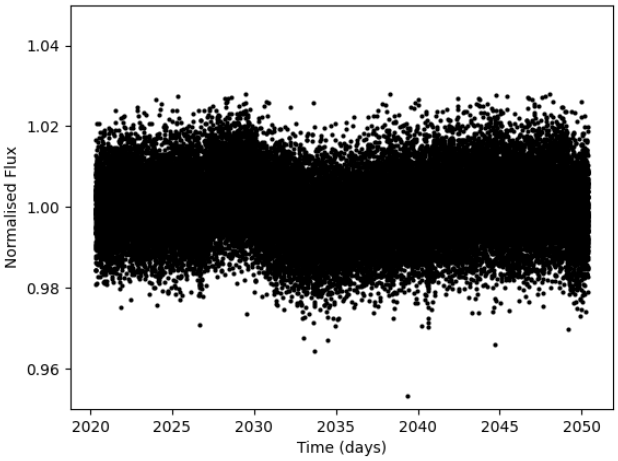
(a) EPIC 206535752's light curve corrected by EVEREST



(b) EPIC 211727748's light curve corrected by EVEREST

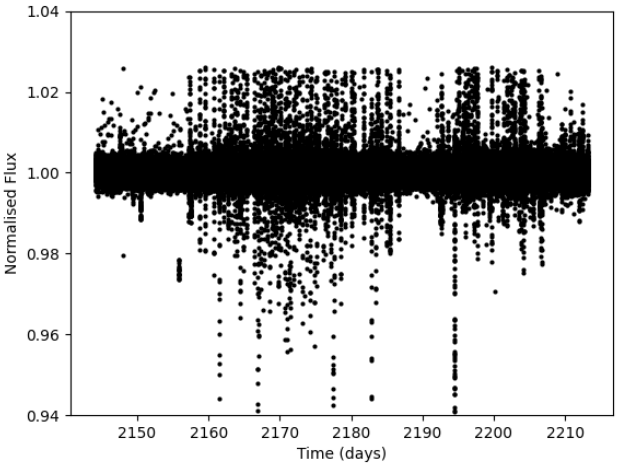


(c) EPIC 206240954's light curve corrected by EVEREST

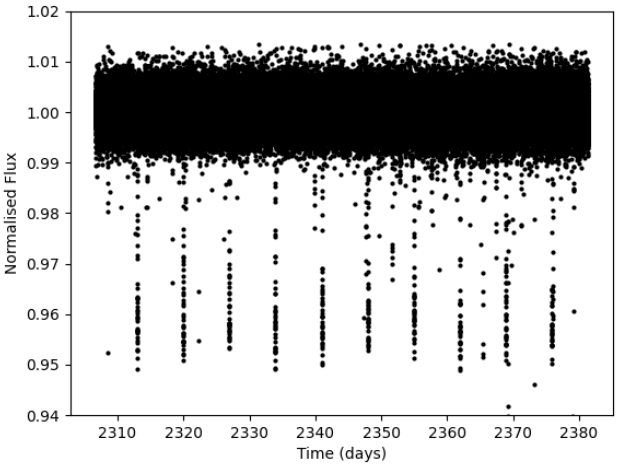


(d) EPIC 201531672's light curve corrected by EVEREST

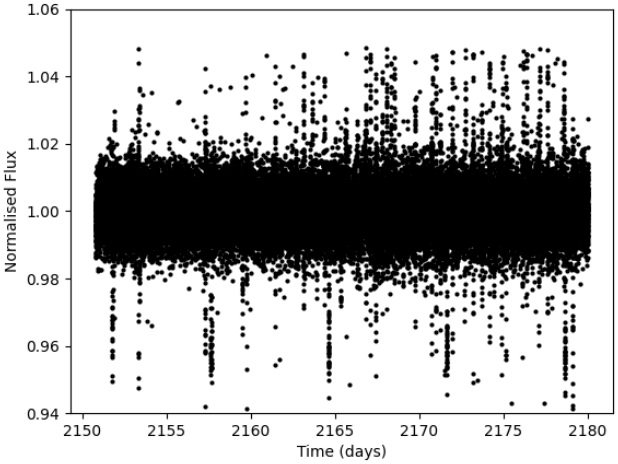
C.2 K2sff light curves



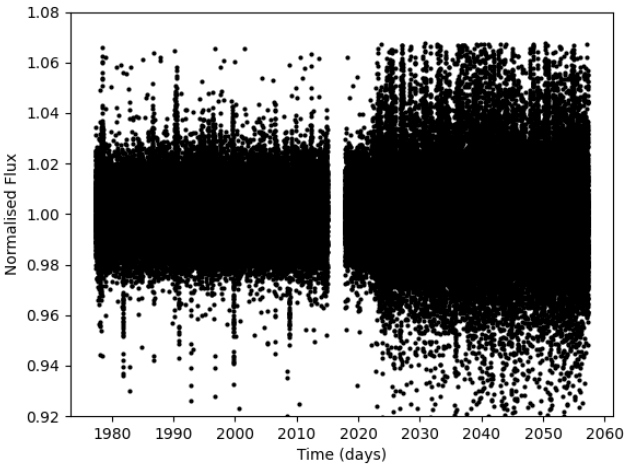
(a) EPIC 206535752's light curve corrected by K2sff



(b) EPIC 211727748's light curve corrected by K2sff



(c) EPIC 206240954's light curve corrected by K2sff



(d) EPIC 201531672's light curve corrected by K2sff