People Powered Planet Hunting with TESS

Nora Lea Eisner

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2022

STATEMENT OF ORIGINALITY

I carried out the work presented in this thesis as a student at the University of Oxford between October 2018 and March 2022, under the supervision and guidance of Professor Chris Lintott and Professor Suzanne Aigrain. I was funded by the Science and Technology Facilities Council (STFC) Studentship Grant Code ST/S505638/1. I hereby declare that no part of this thesis has been submitted in support of another degree, diploma or other qualification at the University of Oxford or other higher learning institute. Except where otherwise stated or where reference is made to the work of others, the work in this thesis is entirely my own.

The work in Chapters 2 and 3 is based on the peer-reviewed paper Eisner et al. 2021, *Planet Hunters TESS II: findings from the first two years of TESS*, MNRAS, 501.4, 4669-4690, for which I am the lead author and contributed the majority of the work. I designed and led the citizen science project which is described in Chapter 2, obtained the results presented in Chapter 3, and wrote the manuscript. However, I note that the final iteration of the planet vetting discussed in Section 2.1.9 was carried out as part of a team, and around 40% of the planets included in Section 3.2 were modelled by Oscar Barragán.

The work in Section 2.3 is based on the peer reviewed paper Eisner, Lintott and Aigrain 2020, *LATTE: Lightcurve Analysis Tool for Transiting Exoplanets*, Journal of Open Source Software, 5.49, 2101, for which I am the lead author and contributed all of the technical work.

The work in Chapter 4 is based on the peer reviewed paper Eisner et al. 2021, *Planet Hunters TESS I: TOI 813, a subgiant hosting a transiting Saturn-sized planet on an 84-day orbit*, MNRAS, 494.1, 750-763, for which I am the lead author, contributed the majority of the work and wrote the manuscript. However, I note that the SED fitting described in Section 4.6 was carried out by a collaborator.

The work in Chapter 5 is based on the peer reviewed paper Eisner et al. 2021, *Planet Hunters TESS III: two transiting planets around the bright G dwarf HD 152843*, MNRAS, 505.2, 1827-1840, for which I am the lead author, contributed the majority of the work and wrote the manuscript. However, I note that the SED fitting described in Section 5.4.3, the transit timing variations outlined in Section 5.5.1, and the atmospheric modelling presented in Section 5.5.4 were carried

out by co-authors. Furthermore, two out of the three methods used for deriving stellar atmospheric parameters (ARES+MOOG and SPC) were led by co-authors, the results of which I incorporated into the overall analysis.

The work in Chapter 6 is based on the peer reviewed paper Eisner et al. 2022, *Planet Hunters TESS IV: A massive, compact hierarchical triple star system TIC 470710327*, MNRAS, 511.4, 4710-4723, for which I am the lead author, contributed the majority of the work and wrote the manuscript. However, I note that the SED modelling in Section 6.6.2 and the discussions on the dynamical stability, system formation and system evolution in Sections 6.6.4 to 6.6.5 were led by co-authors. The follow-up observations, for which I carried out the data reduction, were obtained by several co-authors.

The work throughout this thesis relies on observations from the Transiting Exoplanet Survey Satellite (*TESS*), which his funded by NASA's Science Mission directorate. Furthermore, this thesis was made possible by the participation of over 30 000 citizen scientists taking part in the planet Hunters TESS project. Their contributions are acknowledged on the Planet Hunters Analysis Database (PHAD, https://mast.stsci.edu/phad/), which is hosted by the Mikulski Archive for Space Telescopes (MAST, http://archive.stsci.edu/tess/).

ACKNOWLEDGEMENTS

As my time as a DPhil student in Oxford comes to a close, I am grateful to a wide range of individuals from around the world who made this journey possible. I wouldn't have been able to put this thesis together without the endless support and enthusiasm for my work that I received from so many.

First, I would like to thank my supervisors Chris Lintott and Suzanne Aigrain, whose support over the past four years has been unwavering. Chris and Suzanne, thank you for always taking the time to answer my questions and listen to my ideas; for your trust in my work; for supporting my scientific excursions into studying triple systems; for providing guidance as well as freedom; and for your patience and your continuous encouragement. You have helped me build my confidence and steered me towards becoming the independent researcher that I want to be. I couldn't have asked for better mentors.

I have been lucky enough to have many other great mentors throughout my scientific career. In particular, I would like to thank Matthew Knight. As my first academic mentor, you introduced me to scientific research, showed me the joys of observing, accompanied me to my first conference, guided me through authoring my first papers, and shared your enthusiasm for astronomy. Thank you for being a great mentor, colleague, and friend.

While I was in Oxford, I was warmly welcomed into both the Zooniverse family and the Oxford exoplanet group. I am extremely grateful to the whole Zooniverse team as without you none of this work would have been possible. Also a special thanks to Grant for answering my endless questions about citizen science and planets alike, and Cam for the endless cups of coffees and for adopting all of the plants that I left behind. I was told early on that no one ever *really* leaves the Zooniverse, and I now certainly know why no one would ever want to. I also thank the SPlOx research group. In particular Oscar, I am so grateful for your mentorship and your friendship, your enthusiasm for planets and dinosaurs, and for all of the times we spent at Pret coming up with Latin phrases to sneak into our papers: *psittacorum ebrius est optimus animalibus!* And Laurel, thank you for all the laughs, the gossips, the drawing sessions, and for keeping me (somewhat) sane during the pandemic.

I also want to thank all of the citizen scientists from around the world who have contributed to Planet Hunters TESS over the years. The project is only as successful as it is because of your continued help and dedication to the project. I cannot express how much your enthusiasm for planet hunting means to me, and how much I enjoyed sharing the Planet Hunters TESS journey with each and every one of you. A special thank you to all of the moderators – David, Els, Ian, Sam, and Tony – who are invaluable to keeping the project going, and to everyone who has supported our newest venture with Planet Hunters Coffee Chat. Thank you to the whole Coffee Chat team for all of the fun that we have had with this project. In particular, I want to thank Kassie. In addition to always being a wonderful colleague, you have become a true friend and I can't wait to finally see you outside of my laptop screen.

I want to thank all my friends – old and new – for all of the cups of tea, brunches, pub trips, cycle rides, travel adventures, and great memories. Special thanks to Alice, for being my number-one social media proof reader and for always being ready to share a bottle of Prosecco. Thanks to Charlotte, for sharing this whole crazy astronomy-degree adventure with me. And thanks to Emily, for all your support throughout postdoc application season and thesis writing.

I also want to thank my family. To Flurin, for laying out a life path that I could so carefully follow (why change a system that works?). To Rahel, for reminding me that there is a beautiful world outside of academia. And to my parents, Ruth and Manuel, for supporting me through absolutely everything. Your unconditional love and encouragement made me who I am today, for which I cannot thank you enough.

And finally to Cole. Thank you for bringing more love and joy to my life than I could have ever imagined. You inspire me, challenge me, and make me laugh with your craziness. Thank you for being the only person I would ever want to spend two years of lockdown with and for being the only person to read my thesis as thoroughly as you did. Here's to all our future adventures!

— *May 2022*

ABSTRACT

To date, over 5000 exoplanets have been confirmed, and studies of their characteristics have unveiled an extremely wide range of masses, sizes, system architectures, and orbital periods. More than two-thirds of these known planets were identified using dedicated planet detection algorithms that search for transit events in data obtained by space-based survey satellites, such as the Transiting Exoplanet Survey Satellite (*TESS*). These automated transit detection algorithms are, however, strongly biased towards the detection of short-period planets. For example, 90% of the pipeline-detected *TESS* planet candidates known to date have orbital periods shorter than 27 days. In this thesis, I show that we can address this period bias by visually inspecting *TESS* data with the help of a global community of over 30 000 volunteers.

I present the Planet Hunters TESS (PHT) citizen science project that I designed, built, and continue to manage. Throughout this thesis I demonstrate that with visual vetting we are sensitive to detecting both short- and longer-period planets. This makes PHT an effective tool to help populate under-explored regions of parameter space that have limited prospects for being studied by automated searches that typically require a minimum of two transit events for detection. To that end, the PHT planet sample can benefit studies of planet occurrence rates, as well as inform theories of physical processes involved with the formation and evolution of different types of exoplanets. I also present `LATTE`; a new open-source vetting suite that includes numerous diagnostic tests to aid in the detection and characterisation of planetary and stellar signals. The tool, which is regularly used by both professional astronomers and citizen scientists, forms part of the PHT pipeline and helps identify promising candidates for ground-based follow-up.

The large-scale visual vetting with PHT, combined with the `LATTE` vetting suite, has led to the discovery of 139 new planet candidates in the two-minute cadence data from the first three years of *TESS* data alone. In addition to discussing this ensemble sample of new planet candidates, I present a detailed analysis of two of these systems. First, I discuss a Saturn-sized planet orbiting around a bright subgiant on an $\sim$84-day orbit. The planet's relatively long orbital period combined with the evolved nature of the host star places this planet in a relatively under explored region of parameter space and is, therefore, an exciting target for

further characterisation. Second, I present a two-planet system orbiting around a bright G dwarf. Preliminary mass estimates derived from ground-based radial velocity observations, combined with detailed transit modelling, suggest that both planets in this second system have low bulk densities and therefore extended H/He atmospheres. This makes both planets prime candidates for future atmospheric characterisation and comparative planetology.

Finally, I show that visual inspection of large amounts of data via PHT yields many scientifically valuable by-products, including the identification of 4584 eclipsing binary candidates and multi-stellar systems. I discuss one of these systems – a massive compact hierarchical triple star system – in more detail and present a dynamical analysis of its past, present, and future. Overall, this thesis demonstrates the scientific value of citizen science for identifying unique signals that are often missed by automated searches and, therefore, shows that people-powered planet hunting can play an important role in a world that is becoming increasingly automated.

Contents

List of Figures

LIST OF ABBREVIATIONS

AU . . . . . Astronomical Unit
BJD . . . . . Barycentric Julian Date
Dec . . . . . Declination
DV . . . . . Data Validation (report)
ExoFOP . . . Exoplanet Follow-up Observing Program
LC . . . . . Long Cadence (*TESS* data)
LATTE . . . . Lightcurve Analysis Tool for Transiting Exoplanets
LTE . . . . . Local Thermodynamic Equilibrium
mas . . . . . Milliarcsecond
MCMC . . . . Markov Chain Monte Carlo
$\mathbf{M_\odot}$ . . . . . Solar mass (1.98841×10^{30} kg)
$\mathbf{M_\oplus}$ . . . . . Earth mass (5.97217×10^{24} kg)
O–C . . . . . Observed–Computed
PHT . . . . . Planet Hunters TESS
RA . . . . . Right Ascension
RV . . . . . Radial Velocity
$\mathbf{R_\odot}$ . . . . . Solar radii (6.957×10^{8} m)
$\mathbf{R_\oplus}$ . . . . . Earth radii (6.3781×10^{6} m)
SC . . . . . Short Cadence (*TESS* data)
SPOC . . . . . Science Processing Operations Center
S/N . . . . . Signal to Noise ratio
TCE . . . . . Threshold Crossing Event
TOI . . . . . Tess Object of Interest
TPF . . . . . Target Pixel File
TTV . . . . . Transit Timing Variation

I can, you know, do maths and stuff.

— H. J. Potter, 1991

CHAPTER 1

INTRODUCTION

1.1 Motivation

In early 1990, NASA's Voyager 1 spacecraft captured a historic picture of our Solar System from a distance of more than 6.4 billion kilometers away from Earth.[1] The image includes Venus, Earth, Jupiter, Saturn, Uranus, Neptune, and the Sun. The contrast between the seemingly endless darkness and the Earth – the *pale blue dot* that spans just 0.12 out of the 640 thousand pixels – highlights a possible uniqueness, or loneliness, of our planet in the vast expanses of space. This sentiment was reflected in Carl Sagan's comments on the *pale blue dot* image in 1994 (Sagan, 1994):

> *We succeeded in taking that picture, and, if you look at it, you see a dot. That's here. That's home. That's us. On it, everyone you ever heard of, every human being who ever lived, lived out their lives. [...] Our posturings, our imagined self-importance, the delusion that we have some privileged position in the Universe, are challenged by this point of pale light. Our planet is a lonely speck in the great enveloping cosmic dark.*

At the time there were no confirmed planets beyond those in our own Solar System, and for all we knew, the Solar System might have been a cosmic anomaly in the vast emptiness of space. This is not to say, however, that no one was searching. Many individuals, teams, and surveys had searched the sky to hunt for planets

[1] https://photojournal.jpl.nasa.gov/catalog/PIA00452

around stars beyond our Solar System (e.g., Campbell, 1983; Marcy, 1983; Mayor & Maurice, 1985), under the assumption that other stars should host planetary systems of their own, with properties and architectures similar to the one we have here. However, despite millennia of philosophical and logical arguments, we would not find the first confirmed extra-solar planets, or *exoplanets*, until the early 1990s.

The first detection of a planet sized object was made in 1992, consisting of a planet orbiting around a pulsar (PSR B1257+12 b; Wolszczan & Frail, 1992). This detection of a planet around a stellar end-product prompted a revision of how we think planets evolve alongside their hosts. Soon thereafter, in 1995, the detection of the first planet around around a main sequence star, 51 Pegasi b, followed (Mayor & Queloz, 1995). To the surprise of the astronomical community, this planet also shared little resemblance with the planets in our own Solar System. With a mass similar to that of Jupiter but on an orbital period of just ∼4 days, this planet challenged our expectation that other planetary systems look like our own.

The discovery of *exotic* planets did not end there. Over the next 25 years, astronomers would find a whole range of planets that have no analogues within our own Solar System, including rocky planets with completely molten surfaces (lava worlds; Batalha et al., 2011; Léger et al., 2009), planets in orbit around multiple stars (circumbinary planets; Doyle et al., 2011; Orosz et al., 2012), inflated planets (planets with radii larger than expected; Charbonneau et al., 2000; Hebb et al., 2009), and planets around stellar end-products (Wolszczan & Frail, 1992). All of these observations challenged the previous theoretical paradigm of planetary science and prompted the revision of how planets form, evolve, migrate, and die.

Furthermore, these discoveries helped place Voyager 1's pale blue dot image into context. Even though the Earth is certainly not *alone*, it is still apparently *unique*. While great technological and scientific advances over time have helped reveal the diversity of exoplanets in our Galaxy, we are still faced with challenges when it comes to finding Earth analogues, and thus to address the uniqueness of our planet. One of these challenges is that planet search missions are inherently biased towards finding large and close-in planets. With this work, I try to address the

latter of these biases by searching for longer-period planets in data collected by the *TESS* satellite with the help of citizen science. A natural by-product of this search, as discussed throughout this thesis, is the detection of unusual planet and stellar systems. As was shown so early on with the detection of PSR B1257+12 b and 51 Pegasi b, these systems are particularly exciting as unexpected and exotic objects can provide powerful constraints for our theoretical understanding of the Universe.

1.2 Exoplanets

This section discusses the formal definitions of the terms 'planet' and 'exoplanet' as they have come to be defined today. As has been the case throughout history, such definitions are subject to change as our understanding of dwarf planets, exoplanets, brown dwarfs, and stars continues to improve. Today's discussions about the finer details of the definition of exoplanets contrasts past debates about the mere existence of other worlds and shows how far the field has come. To that end, I also provide a brief overview of how our understanding of exoplanets has changed throughout history, and how this has helped inform the modern study of exoplanets.

1.2.1 What is an exoplanet?

In 2006, the International Astronomical Union (IAU)[2] defined an object to be considered a planet if the object is 1) in orbit around the Sun, 2) has sufficient mass for its self-gravity to overcome rigid body forces so that it assumes a hydrostatic equilibrium (nearly round) shape, and 3) has cleared the neighbourhood around its orbit. Objects that only fulfil the first two of these criteria (e.g., Pluto) are called *dwarf planets*.

This definition of planets can be generalised to *exoplanets* if we change criterion 1 to stars other than our Sun. However, it is very difficult to observationally verify the latter two conditions for exoplanets. As such, the IAU's working definition of an exoplanet (Lecavelier des Etangs & Lissauer, 2022), is any object orbiting around a star (other than our Sun), brown dwarf, or stellar remnant that has a

[2] https://www.iau.org/static/resolutions/Resolution_GA26-5-6.pdf

mass below the mass limit needed for thermonuclear fusion of deuterium (assumed to be ~ 13 M_{Jup} for objects with solar metallicity).

Even though this definition of an exoplanet is not favoured by everyone – Soter (2006), for example, prefers a definition based on the formation process – throughout the remainder of this work I use the definition outlined by the IAU. I also note that in this thesis I use the terms 'planet' and 'exoplanet' interchangeably, where, unless otherwise stated, both refer to planets beyond our Solar System.

1.2.2 Exoplanets: from Ancient Greece to the early 1990s

The theory of planets beyond our own Solar System, sometimes known as the *plurality of worlds* or *cosmic pluralism*, dates back to at least the 6th century BCE when the Ancient Greek philosopher Anaximander (610 – 545 BCE) posited the existence of 'infinite coexistent worlds' (e.g., Preus, 2001). While the theory of cosmic pluralism was also supported by philosophers including Democritus, Anaxagoras, and Epicurus (e.g., Mansfeld, 1980; Crowe, 1988; Warren, 2004; Fridlund et al., 2010) it was met with strong opposition by others, including Plato, Ptolemy, and Aristotle, who were of the opinion that *'...there cannot be more worlds than one'* (Aristotle, 384-322 BCE; Aristotle, 2009). This latter view was in line with the geocentric planetary model and was later backed by the Catholic church, resulting in the suppression of the theory of cosmic pluralism for thousands of years (Connes, 2020).[3]

A revival of the idea of distant worlds came when Copernicus (1473 – 1543) published his book on the heliocentric model of the Solar System (Copernicus, 1965). Importantly, the heliocentric model makes the Earth a planet, which in turn suggests that other planets can be like the Earth and that stars should be

[3]The long-lasting strain between science and the Church regarding the existence of worlds beyond our own may, at least for some, have been overcome with the help of *the principle of plenitude*, which states that God would not limit his skill and knowledge to creating just one world (Connes, 2020; Wilkinson, 2016; Dick, 2020). The idea, which dates back to the ancient Epicureans (~ 300 BCE) and was later endorsed by influential philosophers such as Immanuel Kant (1724 – 1804; Connes, 2020), may have provided a way for the clashing views of the Church and scientists to come together.

thought of as being like our Sun.[4] Due to fears of offending the Catholic church with his work, Copernicus refrained from publishing his findings that favoured the idea of cosmic pluralism for nine years, and by the time that it was published many excused his words due to his old age (Wilkinson, 2016).

Decades later, the idea was picked up again by Giordano Bruno (1548 – 1600), who continued to build on the heliocentric foundation laid out by Copernicus. Bruno conjectured that all stars are like our Sun, and that all Suns have planets orbiting around them, just like the planets in our own Solar System (e.g., Gatti, 2010). Bruno took this further and claimed that all planets are equally inhabited as our own Earth. Even though Bruno's theories were met with strong resistance throughout his short lifetime, in the centuries following his death the idea of cosmic pluralism managed to gain a following. Johannes Kepler (1571 – 1630) and William Herschel (1738 – 1822) were among the prominent scientists who spoke in favour of the heliocentric model, and who likely believed in cosmic pluralism (e.g., Connes, 2020; Dick, 2020).

Further support for the plurality of worlds was provided by Immanuel Kant (1724 – 1804), who derived some of the first models of how our own Solar System formed, building on the work of Swedish philosopher Emanuel Swedenborg (1688 – 1772; Dunér, 2016). Swedenborg and Kant's model predicted that planets are a natural by-product of star formation and that they are, therefore, as common as stars themselves (Swedenborg, 1734; Kant, 1755). In brief, in this now generally accepted model of star and planet formation, dense regions of gas and dust collapse to form dense clumps, now known as prestellar cores (see Figure 1.1). Due to conservation of angular momentum, the collapse of the cloud leads to the formation of a disk out of which the planets form, with the larger gas giant planets necessarily forming on wide orbits (2 – 3 AU) and the smaller rocky planets forming closer to their host star. As such, this model predicts that, as stars are common, planets should be ubiquitous.

[4] Over time, this became part of what is now known as the Copernican Principle, which states that there is nothing 'special' about our place in the Universe. By definition, if we are not special, there must be more than one world. However, it took many years (more than 300!) for this to become an accepted truth.

Even though there were no confirmed exoplanets by the start of the 20th century, by then it was commonly accepted that distant worlds exist and astronomers started searching for their signatures. In the early 1950s, Otto Struve proposed that exoplanets could be detected by observing periodic blue and red shifts in the light given off by a planet-hosting star (Struve, 1952), known as the radial velocity (RV) method. He showed that a 10 M_{Jup} planet orbiting around a Sun-like star at a separation of 0.02 AU would induce a signal with an amplitude of around 2 kms^{-1}, a signal large enough to detect from Earth with the tools available at the time (Struve, 1952). Struve's influential work went on to motivate the launch of several exoplanet-hunting RV surveys over the next decades (e.g., Campbell et al., 1988; Latham et al., 1989; Marcy & Butler, 1992).

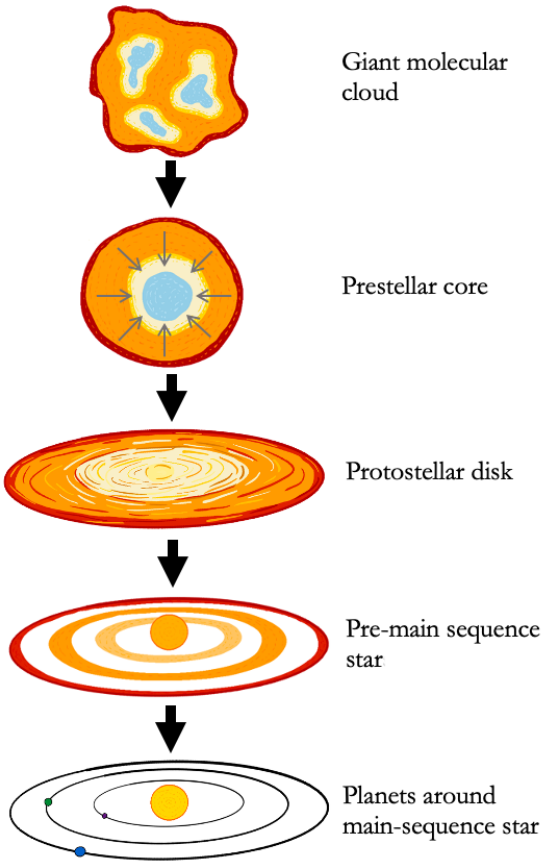

Figure 1.1: Schematic illustrating the main stages of star and planet formation.

In addition to the RV method, Struve suggested that we should be able to detect exoplanets by measuring the periodic dimming of stars as planets pass between the star and the observer (Struve, 1952), known as the transit method. Both the transit and RV methods will be discussed in detail in Sections 1.3 and 1.4.

Around 40 years later, in the early 1990s, the first unambiguous detections of exoplanets were reported (Wolszczan & Frail, 1992; Mayor & Queloz, 1995), which fundamentally and irreversibly shifted the field of exoplanets from a theoretical debate about their existence to a hunt to find more worlds.

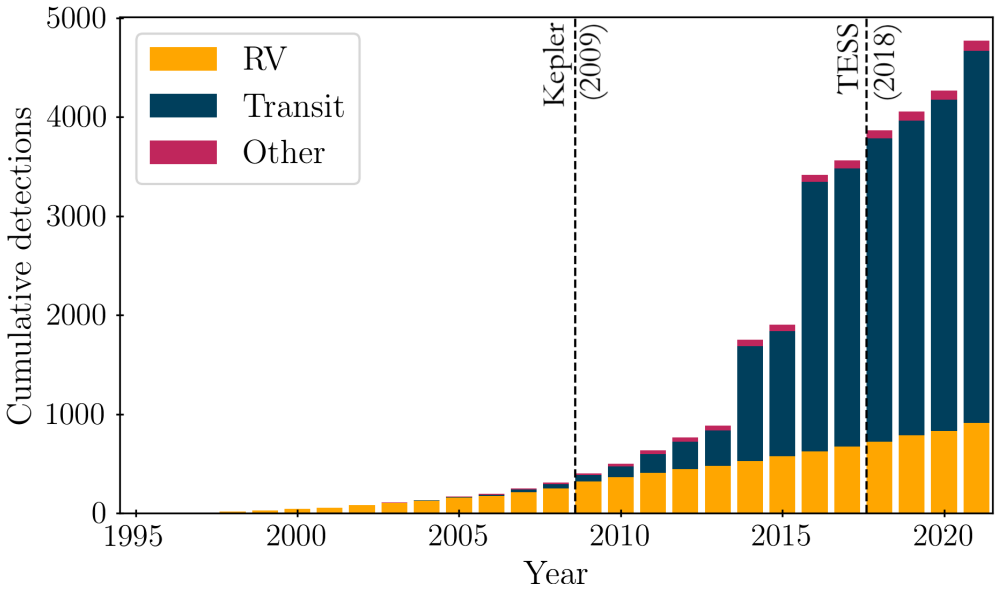

Figure 1.2: Histogram showing the cumulative number of confirmed exoplanets per year since 1995. The two main detection methods, RV and transit, are shown in orange and navy, respectively. This figure highlights that while the RV method was the most successful planet detection method until 2014, the transit method has since then become the most fruitful tool for identifying new exoplanets. The vertical dashed lines indicate the launch years of the *Kepler* (2009; Section 1.4.3) and *TESS* missions (2018; Section 1.5).

1.2.3 Exoplanets today

The detection of thousands of exoplanets over the past quarter century finally provided proof that exoplanets are extremely common. As of 21 April 2022, there are 5011 confirmed exoplanets listed in the NASA Exoplanet Archive (Akeson et al., 2013)[5] and studies of their characteristics have unveiled an extremely wide range of planetary properties in terms of planetary mass, size, system architecture, and orbital periods, as discussed in Section 1.6.1.

1.2.4 Detection methods

The two most successful tools to aid in the hunt for planets remain the two methods outlined by Struve: the RV method and the transit method. Figure 1.2 shows the relative importance of these two methods for the detection of exoplanets over time. While the RV method was initial responsible for the detection of the largest portion of the confirmed exoplanet population, the transit method has been the most

[5] https://exoplanetarchive.ipac.caltech.edu/

successful detection method since 2014. Other planet detection methods include radio pulsar timing (e.g., Wolszczan & Frail, 1992), microlensing (e.g., Gaudi, 2012), direct imaging (e.g., Boccaletti, 2011), astronometry (e.g., Benedict et al., 2002) and transit timing variations (TTVs; e.g., Miralda-Escudé, 2002). The relative successes of the various detection methods used to identify the first 5011 confirmed exoplanets listed in the NASA Exoplanet Archive, as well as their corresponding detection limits, are shown in Figures 1.3.

Sections 1.3 and 1.4 will outline how the RV and transit methods work, their detection biases, as well as their advantages and disadvantages for sampling the planet population in our Galaxy. Understanding of the detection biases, in particular, is important for our understanding of the underlying exoplanet population.

1.3 Radial velocity method

1.3.1 Methodology

The motion of a planet orbiting around a star causes a reflex-motion in the star around the system's centre of mass, known as the *barycentre*, which can be observed by measuring the line-of-sight radial velocity of the star. Using Doppler spectroscopy, where the measured wavelengths of known spectral lines are compared to their rest frame wavelengths, we can determine whether the distance between the star and the observer is increasing (positive RV) or decreasing (negative RV). As such, the detection of periodic RV variations allows us to infer the existence of a planet around a star. Furthermore, the amplitude of the signal enables us to determine system properties including the ratio of the mass of the planet to the mass of the star, the eccentricity of the orbit, and the planet's orbital distance. The amplitude of the RV variation is given by:

$$K \equiv \frac{2\pi}{P}\frac{a\sin i}{(1-e^2)^{1/2}} = \left(\frac{2\pi\,\mathrm{G}}{P}\right)^{1/3}\frac{M_p\,\sin i}{(M_\star + M_p)^{2/3}}\frac{1}{(1-e^2)^{1/2}} \tag{1.1}$$

where $M_\star$ and M_p are the stellar and planetary mass, respectively, a is the semi-major axis, i is the orbital inclination ($i = 90°$ is an edge-on orbit), P is the

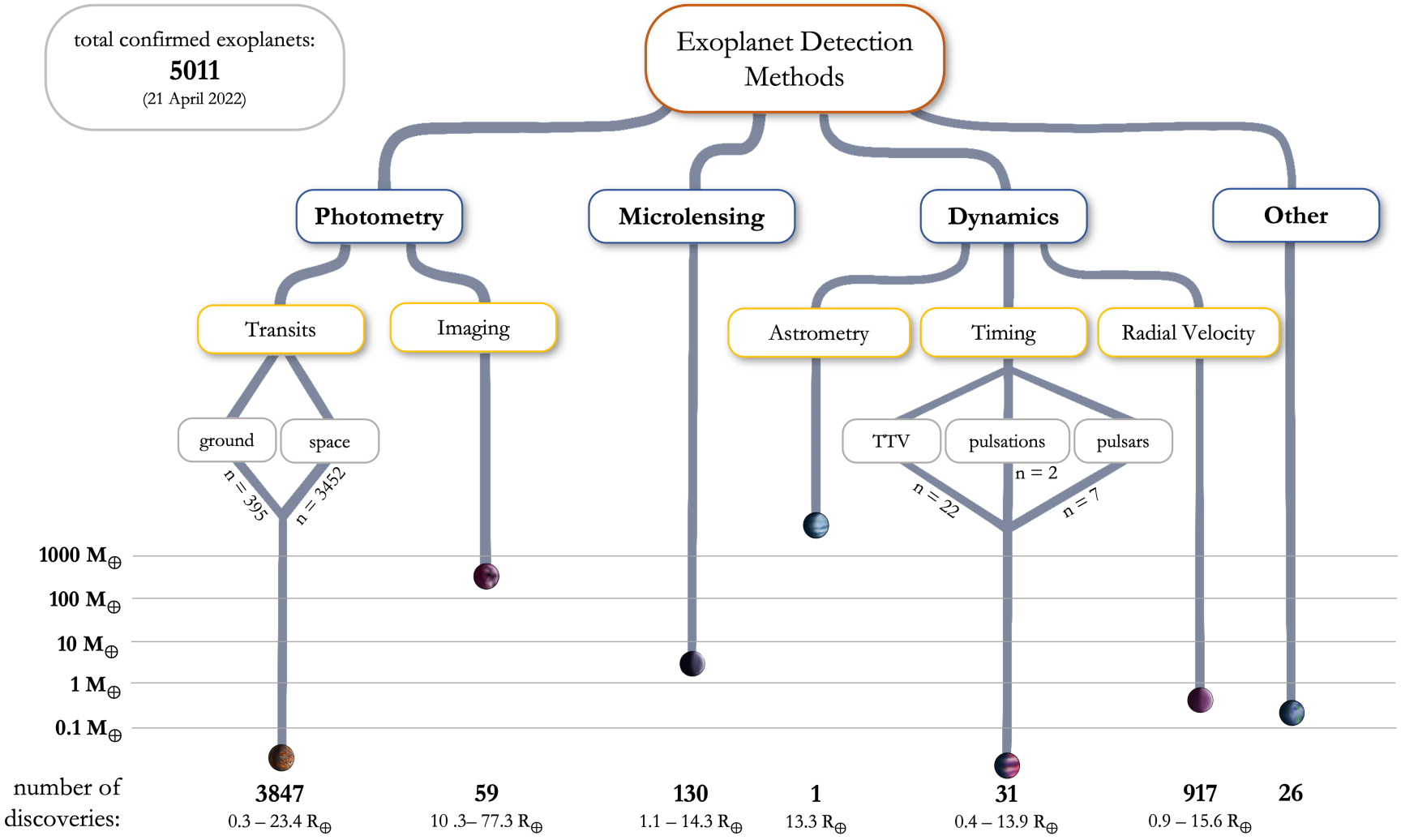

Figure 1.3: Exoplanet detection methods for the 5011 exoplanets listed in the NASA Exoplanet Archive as of 21 April 2022. The end point of each line indicates the lowest confirmed planet mass, while the number at the bottom of the line indicates the total number of planets that have been identified with the corresponding detection method. The radius range underneath indicates the size of planets that the method has been able to detect. The 'other' methods includes eclipse timing variations and disk kinematics.

orbital period, e is the orbital eccentricity, and G is the gravitational constant (6.67408×10^{-11} $\mathrm{m^3kg^{-1}s^{-2}}$). This equation shows that the amplitude of the RV variation is proportional to $M_p/M_\star^{2/3}$ (for $M_\star \gg M_p$), meaning that if the mass of the star is known from, for example, spectroscopic or asteroseismic observations, the mass of the planet can be estimated. However, as we only see the radial component of the motion of the star, the amplitude of the signal is dependent on the inclination of the system (i). As such, we can only determine a *lower* mass limit of the planet ($M_p \sin i$), unless the inclination of the system is known (e.g., from the detection of transit events, see Section 1.4).

1.3.2 Challenges

The detection of planets using the RV method can be challenging due to small RV amplitudes. Depending on the planet mass and semi-major axis, measurements of a star's reflex-motion induced by a planet requires measurements with a precision ranging from a few centimetres to hundreds of meters per second. Figure 1.4 shows the RV amplitude as a function of the semi-major axis for planets with masses equivalent to those of Earth, Neptune, and Jupiter.

In order to reach the precision necessary to detect exoplanets using the RV method there are a number of astrophysical and instrumental challenges that need to be overcome. For perspective, the typical modern RV spectrograph has a resolving power of $R = \lambda/\Delta\lambda \sim 100\,000$, meaning that an RV shift of 1 $\mathrm{ms^{-1}}$ (corresponding to a wavelength shift of $\sim 10^{-5}$ Å) translates to a shift on the CCD detector of around 1/1000th of a pixel. To see such small shifts, extremely accurate wavelength calibrations are needed (e.g., using iodine absorption cells or thorium-argon emission lamps). To that end, the temperatures and pressures around the detector must be extremely stable, as even small environmental fluctuations change the refractive index of air, and thus can result in wavelength shifts of several hundreds of meters per second. Due to this, modern specrographs, such as HARPS (High Accuracy Radial Velocity Planet Searcher; Pepe et al., 2000),

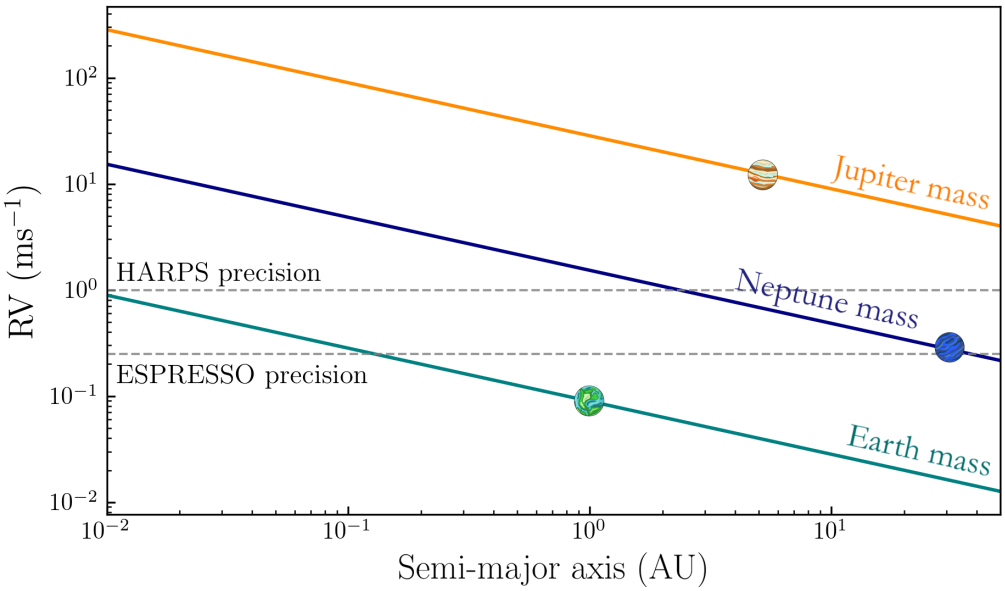

Figure 1.4: The RV amplitude as a function of the semi-major axis for Jupiter (orange), Neptune (navy) and Earth (teal) mass planets, orbiting around a 1 $M_{\odot}$ star and assuming an orbital inclination of 90°. The circles show the semi-major axes of these three planets. The dashed horizontal lines how the detection sensitivities of the HARPS (1 ms^{-1}; Pepe et al., 2000) and ESPRESSO (0.25 m^{-1}; Pepe et al., 2021) spectrographs.

regulate the pressure and temperature surrounding the instrument to extreme precision using cryostats and vacuum chambers.

Further considerations need to be made to account for the motion of the Earth, which rotates at ∼300 ms^{-1} and orbits around the Sun at ∼30 kms^{-1}. Due to this motion, even a stationary star will show RV shifts on diurnal and annual timescales. Correcting for these shifts is known as the barycentric correction.

Advances in technology have – somewhat astonishingly considering the challenges – managed to improve RV precision from over a few kms^{-1} in the 1950s (e.g., Wilson, 1953) to typical precisions of, for example, 1 ms^{-1} with HARPS on the ESO-3.6-m telescope (Pepe et al., 2000) and 0.25 ms^{-1} with ESPRESSO on ESO's Very Large Telescope (Pepe et al., 2021), as shown by the dashed horizontal lines in Figure 1.4.

Even though modern instruments are capable of reaching precisions better than 1 ms^{-1}, the detection of smaller planets remains difficult. This is, in part, due to stellar variability that induces RV shifts on the order of a couple of meters per second. These stellar RV signals originate from stellar pulsations due to turbulent convection (∼minutes), surface granulation (∼hours), stellar spots and plages that rotate with

the star (days to months), and magnetic activity (years). For more detail on these signals and for methods of how they can be accounted for with stellar noise modelling see, for example, Radick et al. (1998), Meunier et al. (2010), McQuillan et al. (2011), Rajpaul et al. (2015), Fischer et al. (2016), Dumusque (2018), Nava et al. (2020), Rajpaul et al. (2021), Costes et al. (2021), or Collier Cameron et al. (2021).

The RV method is faced with a number of additional challenge and biases. For example, the RV method is biased towards the detection of massive and shorter period planets, as the amplitude of the induced RV signal is proportional to $M_p P^{-1/3}$ (Equation 1.1; for $R_\star \gg R_p$). In addition to longer-period planets producing smaller RV amplitudes, the ability to detect a given range of orbital periods is ultimately limited by the duration of the RV survey, which can range from just a couple of days to tens of years (e.g., Queloz et al., 2000; Cochran et al., 2002; Howard et al., 2010; Vogt et al., 2010; Pepe et al., 2021).

Furthermore, unlike for transit surveys that can simultaneously monitor thousands of stars, RV surveys are typically limited to observing a single star at a time, making regular sampling of a large number of stars challenging. The resulting irregular time-sampling with a limited number of observations of a limited number of stars makes blind searches for planets challenging.

Despite these challenges, as RV surveys continue to increase their observing baselines, instruments become more sensitive, and RV extraction methods become more sophisticated, our ability to detect less massive and longer-period planets continues to increase.

1.3.3 Radial velocity highlights

The RV method has played – and is still playing – a major role in the detection and characterisation of exoplanets. Over the past quarter century it has enabled the discovery and characterisation of many exciting systems, including hot Jupiters (Mayor & Queloz, 1995); the first multi-planet systems (e.g., Butler et al., 1999); planets around evolved stars (e.g., Johnson et al., 2011); the most dense planets e.g., K2-38 b, which has a density of $\sim$11.0 gcm^{-3} (e.g., Toledo-Padrón et al.,

2020); the most eccentric planets e.g., HD 80606 b, which has an eccentricity of 0.96 (e.g., Jones et al., 2006); and measurements of the spin-orbit alignment of planets (Winn et al., 2005).

Furthermore, the derived planet mass that the RV method can provide, in combination with the planet radius from the transit method (see Section 1.4), provides information on the density of the planet, which in turn can help reveal their internal structures and compositions (e.g., Seager et al., 2007, see Section 1.6.2). Similarly, the RV method allows for the determination of the orbital eccentricity, which can help constrain the formation history of a system (e.g., Chatterjee et al., 2008; Morishima et al., 2008).

1.4 Transit method

As shown in Figures 1.2 and 1.5, the transit method (Charbonneau et al., 2000; Henry et al., 2000) has become the most successful tool for planet detection and characterisation, being responsible for the detection of 77% of the planets listed in the NASA Exoplanet Archive as of 19 January 2022. Throughout the remainder of this thesis all values from the NASA Exoplanet Archive are, unless otherwise stated, as of 19 January 2022.

1.4.1 Methodology

An object is said to be transiting when it passes in front of its star, resulting in a decrease in observed flux. The fractional decrease in the brightness of the host star is proportional to the size of the occulting planet:

$$\frac{\Delta F}{F} \propto \left(\frac{R_p}{R_\star}\right)^2 \tag{1.2}$$

where F is the flux and ΔF the change in flux during the transit. The transit method, therefore, allows us to determine the radius of a transiting planet, if the radius of the star is known from spectroscopic or asteroseismic studies. Additionally, the transit method allows us to determine the orbital period (time between consecutive

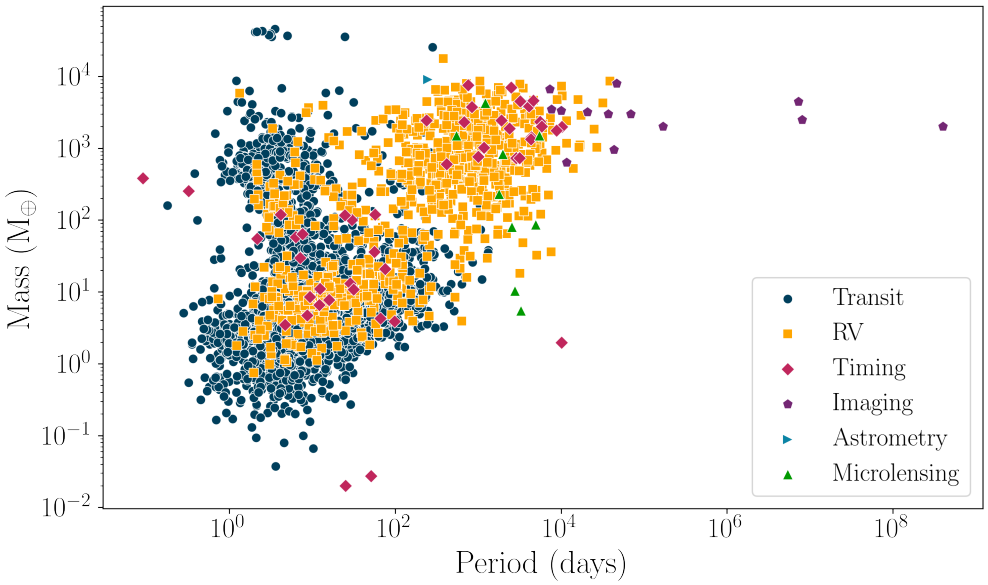

Figure 1.5: The properties of exoplanets detected using different techniques, as listed in the NASA Exoplanet Archive as of 19 January 2022.

transit events), if multiple transits are observed. Applications of transit modelling will be discussed in more detail in Section 2.3.2.

The requirement that the planetary orbit must be oriented such that the planet occults the star as seen by the observer ($i \sim 90°$) greatly favours the detection of planets with short orbital periods. Using geometry, it can be shown that a full transit only occurs when $a\cos i \leq R_\star - R_p$ (see Figure 1.6). If we assume that planetary systems are randomly oriented in space and that $R_\star \gg R_p$, the geometric probability of a planet on a circular orbit transiting in front of its host star can be approximated as (Borucki & Summers, 1984):

$$P \approx 0.0046 \left(\frac{R_\star}{R_\odot}\right) \left(\frac{a}{1\ \mathrm{AU}}\right)^{-1}. \tag{1.3}$$

The probability of a planet transiting in front of a Sun-like star, therefore, ranges from less than 0.1% for planets at $a = 5.2$ AU (Jupiter's orbit) to 10% for planets with orbital periods of just a couple of days, with a probability of just $\sim 0.46\%$ at 1 AU.

The low geometric probability of a planet transiting in front of its host star means that a large number of targets must be observed in order to obtain a good statistical

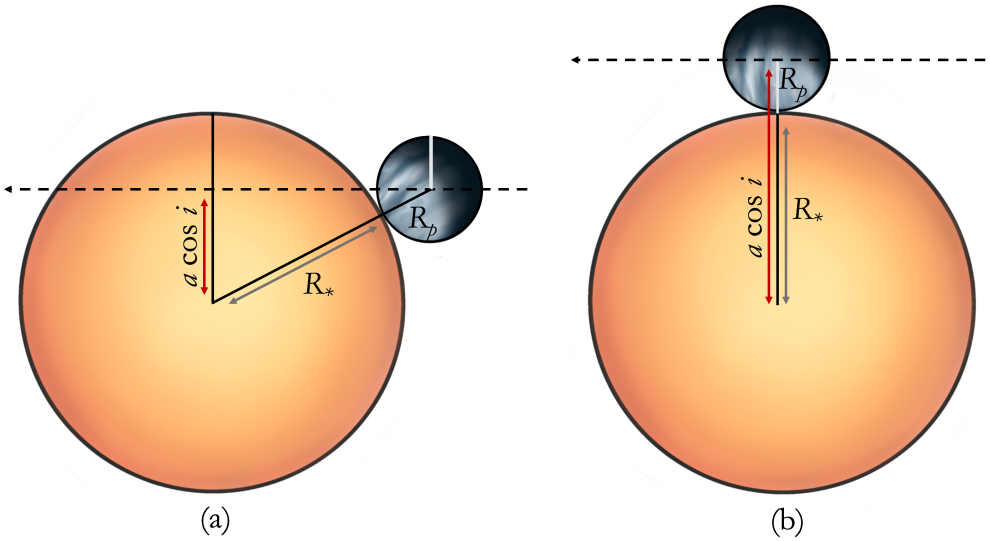

Figure 1.6: The star-planet geometry showing that a full transit occurs when $a\cos i \leq R_\star - R_p$ (panel a); a partial (grazing) transit when $R_\star - R_p < a\cos i < R_\star + R_p$; and no transit when $a\cos i \geq R_\star + R_p$ (panel b). i is the orbital inclination and a is the semi-major axis. The black dashed lines show the orbit of the planet.

sample of exoplanets. As such, there are a number of ground-based and space-based photometric surveys that monitor thousands of stars in search for transiting planets.

1.4.2 Ground-based transit surveys

All transit surveys are subject to strong biases. First, there is a strong bias towards detecting planets on tight orbits due to the increased transit probability of these planets. For ground-based transit surveys this bias is further enhanced due to the rotation of the Earth that results in limited observing windows over the course of a single night. As longer-period planets have longer transit durations (>6 hours) it is often difficult, or even impossible, to capture a full transit in one night of observations (typically $\sim 6-14$ hours).

Second, there is a strong bias towards the detection of larger planets. From the ground, the photometric precision is limited to $\sim 0.1\%$ (1000 ppm) due to fluctuations in the transparency of the Earth's atmosphere (Cameron, 2016). This means that ground-based surveys are limited to detecting planets with radii greater than $\sim 4\ R_\oplus$ around Sun-like stars.[6] Typical transit detection limits are shown in Figure 1.7.

[6]Clouds can also get in the way of seeing a transit.

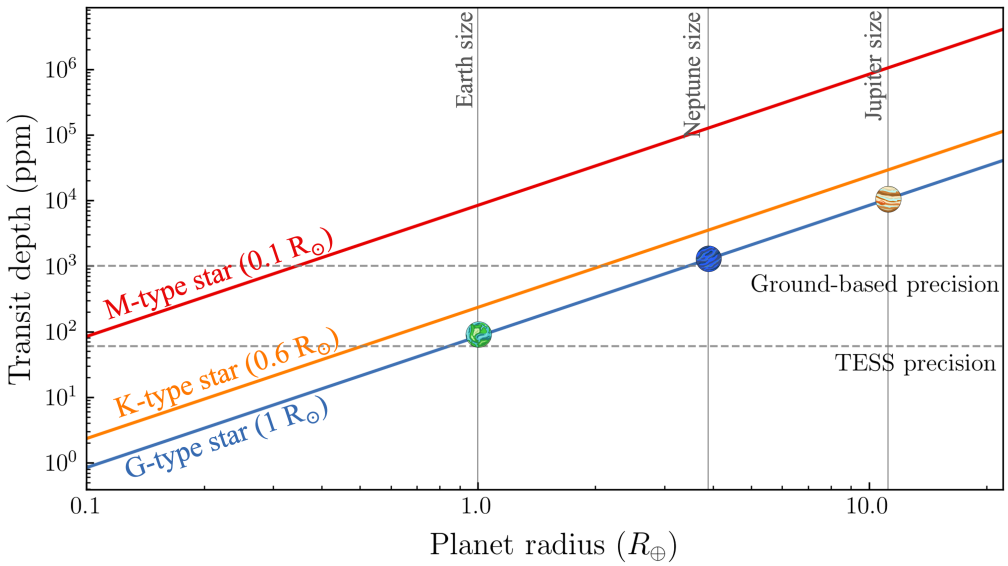

Figure 1.7: The transit depth as a function of planet radius for an M-type star (0.1 $R_\odot$; red), K-type star (0.6 $R_\odot$; orange), and G-type star (1 $R_\odot$; blue). The effects of limb-darkening are not considered. The vertical lines indicate the planet sizes of Earth, Neptune and Jupiter, while the horizontal lines show typical detection limits for ground-based transit surveys and for the *TESS* mission (60 ppm over a magnitude range of 9 – 15; Fausnaugh et al., 2018; Vanderspek et al., 2018).

The results of these detection biases can be seen in Figure 1.8, which shows all confirmed planets listed in the NASA Exoplanet Archive that were identified with ground-based transit surveys (alongside the planets found using space-based missions and the Solar System planets, for comparison). The most successful of these ground-based surveys are the Hungarian automated telescope network (HATNet and HATSouth; 139 confirmed planets; Bakos et al., 2004; Bakos et al., 2008), the Super Wide-Angle Search for Planets (SuperWASP; 121 confirmed planets; Pollacco et al., 2006), the kilodegree extremely little telescope (KELT; 21 confirmed planets; Pepper et al., 2017), and the Transatlantic Exoplanet Survey (TrES; 5 confirmed planets; Alonso et al., 2007). Out of the 286 planets identified from these surveys, only 31 have orbital periods longer than 5 days and 4 have periods longer than 10 days. Similarly, Figure 1.8 highlights a bias towards larger planets, with only 20 confirmed planets that have radii smaller than 10 $R_\oplus$, and only one that has a radius smaller than 5 $R_\oplus$ (HAT-P-11 b; Bakos et al., 2010). In order to detect smaller and longer-period planets, the transit survey must be conducted from space.

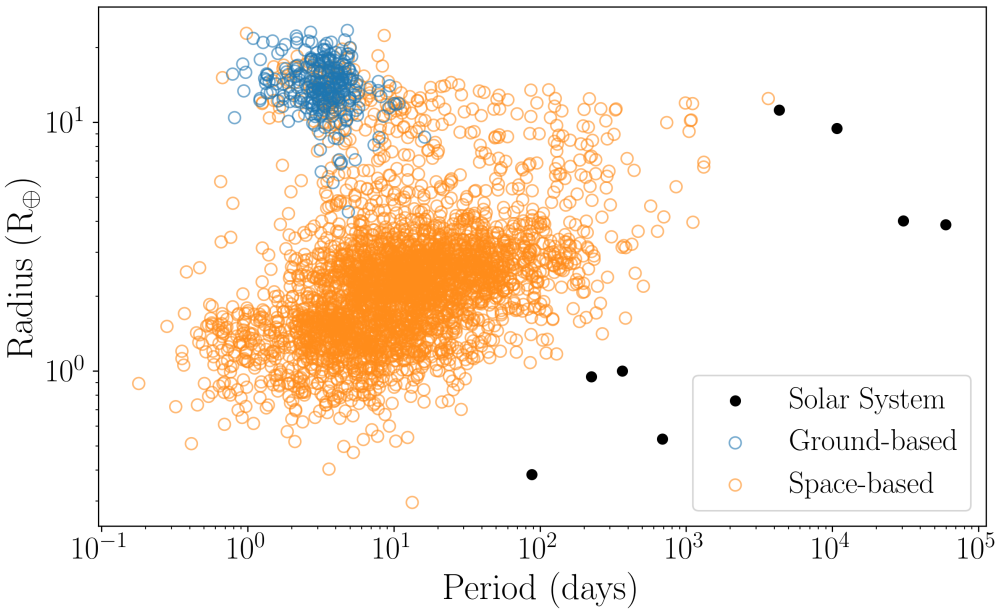

Figure 1.8: The properties of confirmed planets detected using space- and ground-based transit surveys, in orange and blue respectively, compared to the Solar system planets (black). The ground-based surveys include HATNet, SuperWASP, KELT, and TrES, while the space-based surveys include *CoRoT*, *Kepler*, *K2*, and *TESS*.

1.4.3 Space-based transit surveys

Space-based transit missions have the advantage of being able to monitor stars without being interrupted by the diurnal cycle (or bad terrestrial weather), allowing them to identify longer-period planets compared to their ground-based counterparts (see Figure 1.8).[7] Furthermore, their photometric precision is not limited by atmospheric turbulence, thus opening up the possibility of identifying smaller planets and planets transiting in front of larger stars.

The first space-based mission purposely designed to search for transiting exoplanets was the *COnvection, ROtation et Transits planetaires* (CoRoT; Auvergne et al., 2009) satellite, which was operational between 2006 and 2012 and identified a total of 32 transiting exoplanets.[8] In 2009, NASA launched the *Kepler* mission (Borucki et al., 2009), which remains the most successful transit search mission. Between 2009 and 2013, the *Kepler* spacecraft monitored ∼150 000 FGKM dwarfs

[7] I note that all space-based mission have their own unique orbital characteristics which can introduce their own biases.

[8] CoRoT was particularly sensitive to the detection of close-in planets with radii greater than that of Neptune – planets that we have now figured out to be fairly rare (Owen & Lai, 2018).

in a fixed field of view located in the Cygnus and Lyra constellations (the centre of the field of view located at RA = $19^h\,22^m\,40^s$; DEC = $44^\circ\,300''\,000'$), measuring their brightness every 30 minutes. Over the course of the main mission *Kepler* found over 4600 planets candidates out of which 2677 have been confirmed and are listed in the NASA Exoplanet Archive.

In May 2013, the *Kepler* mission entered a new mission phase (due to a number of mechanical failures).[9] During this mission phase, the spacecraft observed different fields located along the ecliptic, each run lasting around 80 days and consisting of observations of $\sim 10\,000 - 20\,000$ stars. A further 477 confirmed planets have been identified with the *K2* mission, which was in operation until October 2018. Despite no longer being operational, the *Kepler* and *K2* missions continue to yield a large number of planet candidates as the community continues to benefit from the large volumes of archival data (e.g., Zink et al., 2019a; Bryson et al., 2021a; Christiansen et al., 2022).

The most recently launched space-based survey satellite specifically designed to look for exoplanets is the Transiting Exoplanet Survey Satellite (*TESS*; Ricker et al., 2015). Since its launch on 18 April 2018, the ongoing *TESS* mission has identified over 5400 planet candidates (e.g., Guerrero et al., 2021), out of which almost 200 have been confirmed. As shown in Figure 1.9, the *TESS* mission has confirmed a slightly higher number of exoplanets than the *Kepler* mission had detected at the same point in the mission (i.e. 4 years following the mission launch). Assuming that this trend holds for the future, we can expect an exceptionally high planet yield from the *TESS* mission over the next couple of years.

The majority of the analysis in this work is based on *TESS* data. As such, the *TESS* instrumentation, data reduction pipeline, and data products will be described in detail in Section 1.5.

[9] On 11 May 2013, the second of *Kepler*'s reaction wheels – needed for fine pointing of the spacecraft – failed., known as *K2* (Howell et al., 2014).

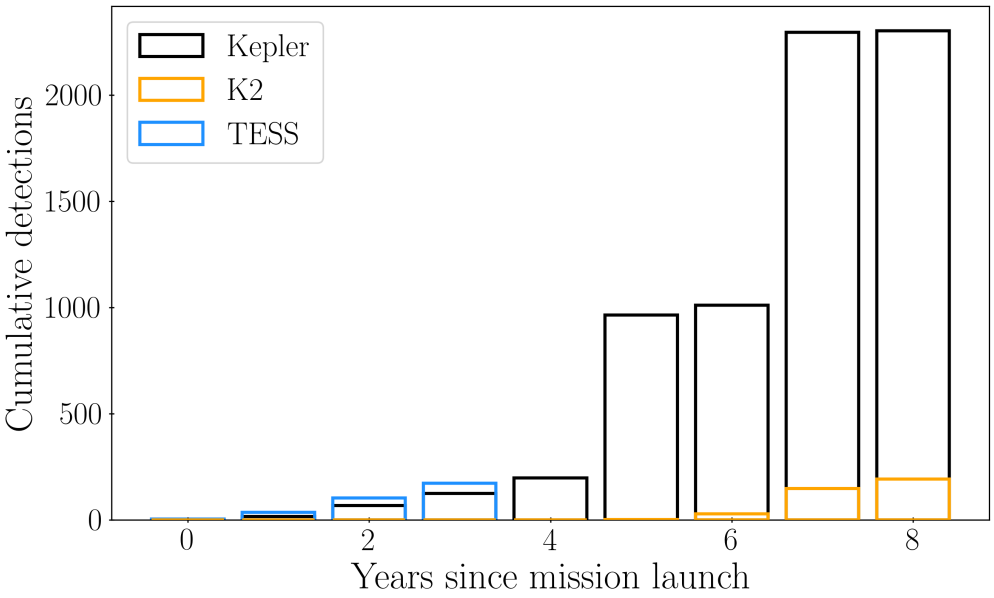

Figure 1.9: Comparison of the yearly rate of planet discoveries of the *Kepler* (black), *K2* (orange), and *TESS* (blue) missions in the first couple of years following their respective launches in 2009, 2013, and 2018. *TESS* is currently in its fourth year of operation. The figure highlights that *TESS* has a slightly higher planet detection yield than the *Kepler* mission did after the same amount of time following the launch of the two satellites.

1.4.4 Transit survey highlights

Space-based transit surveys have helped explore a region of parameter space that is inaccessible to ground-based surveys. With their higher photometric precision and increased continuous observation time coverage compared to ground-based surveys, space-based surveys are able to detect both smaller and longer-period planets (see Figure 1.8). For example, they have allowed for the detection of planets smaller than the Earth, e.g., Kepler-20 e, which has a radius of just 0.87 $R_\oplus$ (Fressin et al., 2012). As of 19 January 2022, there are 192 transiting planets with radii less than that of Earth listed in the NASA Exoplanet Archive, out of which 160 were identified by *Kepler*. The ability to identify and study these tiny worlds highlights the remarkable technological advances that happened over the past 30 years. Today's *planet zoo* – the population of planets that we know to date – will be discussed in more detail in Section 1.6.1.

In addition to helping us detect more planet candidates, the transit method is an important tool for system characterisation. For example, as transit events only

occur for a limited range of system inclinations, the detection of transits allows us to constrain the value of i. In turn, this means that if RV variations are observed for transiting planets, an absolute mass of the planet can be determined (see Equation 1.1). Furthermore, transit observations give rise to transit spectroscopy that allows us to infer the composition of the planetary atmospheres (e.g., Seager, 2008; Mandell et al., 2013; Fu et al., 2021); the Rossiter-McLaughlin effect used to determine the planetary spin-orbit alignment (e.g., Rossiter, 1924; McLaughlin, 1924; Winn et al., 2005; Sanchis-Ojeda et al., 2011; Rice et al., 2021); and transit time/duration variations that can lead to the discovery of further planets within the system that may otherwise be missed (e.g., Miralda-Escudé, 2002; Nesvorný et al., 2012). Finally, space-based transit surveys have allowed for better characterisation of stars via asteroseismology, which provides a uniquely precise way to determine fundamental stellar properties including the mass, radius, and age, of any oscillating star (e.g., Huber et al., 2013; Van Eylen et al., 2019). The better characterisation of these stellar properties in turn allows us to better constrain the properties of the planets that orbit around them.

1.5 Transiting Exoplanet Survey Satellite

TESS is the first nearly all-sky space-based transit search mission, which focuses on finding transiting exoplanets around nearby, bright main sequence stars. The mission's primary science objective is to identify small (< 2.5 $R_\oplus$), short period (< 10 days) planets orbiting around bright stars in the Solar neighbourhood. While the *Kepler* mission stared at a single field of 116 deg^2 for four years, observing stars with $V < 16$ mag, located 700 – 1100 pc away (Berger et al., 2018), *TESS* is monitoring bright ($V < 11$ mag), nearby (< 300 pc) stars across the entire celestial sphere. This makes *TESS* targets particularly well suited for ground- and space-based follow-up to measure the planet masses and densities as well as to characterise their atmospheres.

In order to achieve its primary science goal, *TESS* has a wide bandpass of 600 – 1040 nm that is sensitive to red wavelengths (Ricker et al., 2015). This makes *TESS* well suited to detecting Earth and super-Earth sized planets around cool, red

Table 1.1: Summary of *TESS* instrument properties (Vanderspek et al., 2018).

Instrument property	Comment
Number of cameras	4
Sector field of view	$96° \times 24°$
Camera field of view	$24° \times 24°$
CCDs per camera	4
CCD dimensions	2048×2048 pixels
Pixel size	$15 \times 15\,\mu$m
Pixel angular size	21 arcsec
Bandpass	$600 - 1040$ nm

M-dwarfs that make up around 75% of the stars in our Solar neighbourhood (Henry et al., 2018). For larger FGK type stars, *TESS* is better suited to finding larger planets. However, it has been shown that *TESS* is capable of detecting planets down to around $0.8\,R_\oplus$ around such stars (Guerrero et al., 2021).

1.5.1 *TESS* spacecraft and observing strategy

On board the *TESS* satellite are four identical wide-field charge-coupled device (CCD) cameras and a data handling unit (DHU). The main characteristics of the instrument are summarised in Table 1.1. As shown in panel (a) of Figure 1.10, a 'sector' is made up of four CCD cameras that have adjacent field-of-views to produce a 4×1 array.

The spacecraft is located in an elliptical 13.7-day orbit around the Earth, in a 2:1 resonance with the Moon's orbit. Over the course of the two-year nominal mission, *TESS* monitored stars across $\sim 85\%$ of the sky, split up into 26 rectangular sectors of $96° \times 24°$ each (13 per hemisphere), that extend from the ecliptic pole to near the ecliptic plane (panel b of Figure 1.10). Each sector is monitored for ≈ 27.4 continuous days, after which the field of view is shifted by $\sim 27°$ eastwards. As shown in panel c of Figure 1.10, this observing strategy means that most of the targets ($\sim 63\%$) are only observed for ≈ 27.4 days (during the nominal two-year mission). Only $\sim 2\%$ of the targets close to the ecliptic poles (within $\sim 12°$ of the ecliptic pole) are located in the 'continuous viewing zones' and are continuously monitored for

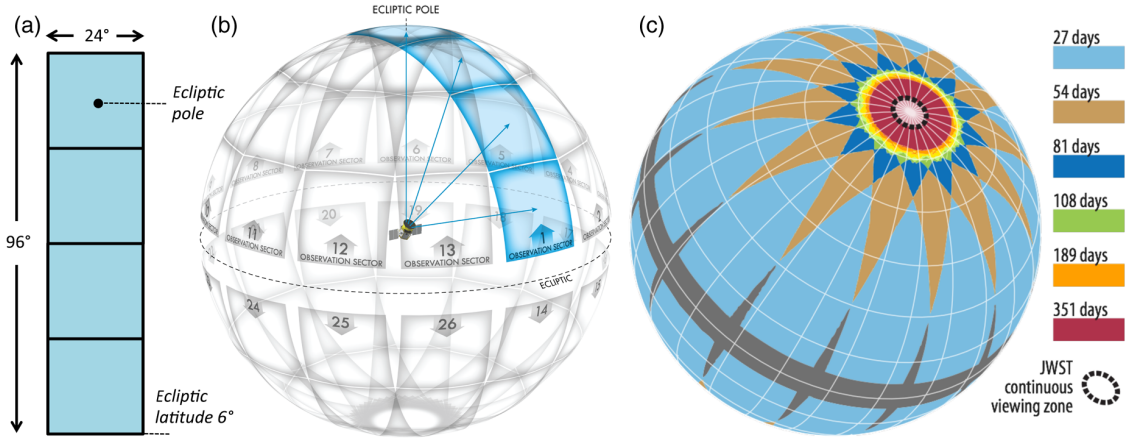

Figure 1.10: a) The $24 \times 96°$ field of view made up of the four *TESS* cameras that make up one sector. b) The celestial sphere split up into 26 sectors (13 per hemisphere). c) Schematic showing how long each part of the sky was observed by *TESS* over the course of the first 2 years of the mission. Times range from $\sim$24 to 351 days. Image from (Ricker et al., 2015).

$\sim$356 days. It should be noted that significant deviations from the nominal mission plan (shown in Figure 1.10) occurred in the northern hemisphere in order to avoid scattered light close to the ecliptic. These changes to the mission plan did not affect the mission's science goal. The observing strategy was largely repeated during the extended mission, where *TESS* re-observed the southern hemisphere in the third year of the mission and the northern hemisphere in the fourth year (ongoing).

1.5.2 Data products

The spacecraft obtains images at a cadence of two seconds, which are combined on board into two data products: short cadence (SC; two-minute) cut-outs centred on $\approx$20 000 pre-selected stars per sector; and long cadence (LC; 30 or 10-minute) Full Frame Images (FFI) that span all pixels of all CCDs. The observations are taken at a two-second cadence and stacked into longer cadences in order to mitigate the impact of cosmic rays, which are identified and removed on-board before the frames are stacked.

The short cadence data consist of 60 consecutive two-second images combined into two-minute cadence observations. For the majority of the targets, the cut-outs are 11×11 pixels wide, centred on pre-selected targets. The pre-selected two-minute

cadence stars primarily consist of dwarfs with spectral types F5 to M5 that will help to address the mission's primary science objective of identify small planets around bright, nearby stars (Stassun et al., 2018, 2019).

Similarly, the long cadence FFI data consist of consecutive two-second images combined into longer cadence observations. The FFIs were combined into 30 minute cadence observations during the primary mission (years 1 and 2) and into 10-minute cadence observations for years 3 and 4 (ongoing). As of year 5 of the mission, the FFIs will be delivered at a 200-second cadence.

Both the SC and LC data products are produced on board the satellite, and downlinked to the ground when the satellite reaches its point of closest approach to the Earth (perigee, $\sim$17 $R_\oplus$). The data are then forwarded to the Payload Operations Center (POC) at MIT, where they are reformatted such that they are ready to be processed by the SPOC pipeline (Science Processing Operations Center; Jenkins et al., 2016). This processing includes calibrating the pixel data and extracting photometry and centroids for each target star. The photometric flux is calculated by summing the calibrated flux of the pixels within the pipeline-chosen aperture (the aperture size depends on the magnitude of the target star), a process known as simple aperture photometry (SAP). An additional level of data processing removes instrumental artefacts, contamination flux from background sources, and long-term trends in the data. This data product is known as the Pre-search Data Conditioning SAP (PDCSAP) flux. Both the raw flux light curves (PDC flux) and the corrected flux (PDCSAP flux) are stored and made publicly available by the Mikulski Archive for Space Telescopes (MAST).[10] The work presented throughout the remainder of this thesis is primarily based on the SC PDCSAP light curves.

1.5.3 Automated transit search

The SPOC pipeline searches for transit signals by phase-folding the detrended *TESS* light curves at a large number of trial periods in order to search for periodic decreases flux. All transit-like signals with a statistical significance greater than 7.1 σ that

[10] http://archive.stsci.edu/tess/

have at least two transit-like events are considered to be Threshold Crossing Events (TCEs; Jenkins et al., 2016). **Importantly, planet candidates that exhibit only a single transit event in the *TESS* data are not routinely flagged by this pipeline.** As such, targets that are only observed for 27 consecutive days (∼63% of all *TESS* targets) are limited to the detection of planets with orbital periods shorter than ∼27 days.

All TCEs are further processed with the DV (Data Validation) pipeline, which runs a number of standard diagnostic tests that help determine whether a signal is of planetary nature or whether it is due to an instrumental or astrophysical false positive. False positive tests will be discussed in detail in Chapter 2. The DV module also provides a transit fit to the phase folded light curve and derives orbital parameters including the planet radius, orbital period, and semi-major axis.

Candidates that pass the automated tests provided by the DV pipeline are visually inspected by a member of the TESS Science Office (TSO) vetting team. Planet candidates that survive this round of visual vetting are promoted to TESS Objects of Interest (TOI) status and considered for further follow-up observations.

1.6 Where are we? Where to next?

The rate of exoplanet detection has continued to rise as technology and instrumentation has continued to improve and more search missions have been launched (e.g., see Figure 1.2). Two of the most important – or perhaps *the* two most important – things that we have learned since the discovery of the first exoplanet, are that 1) exoplanets are common and 2) the diversity of exoplanets in our Galaxy reaches far beyond the diversity of planets we have in our Solar System. The former has important implications for the search for Earth-like planets, while the latter has challenged and reshaped our views of planet formation and evolution. In this section I provide a brief summary of some important and active areas of exoplanetary research, and discuss the importance of planet masses and radii for our understanding of the overall exoplanet population.

1.6.1 Today's planet zoo

Planets beyond the main sequence. The field of exoplanets has shown rapid growth in the number of planets detected around main sequence stars. However, there is still a gap in our understanding of what happens to planets as their host stars evolve beyond the main sequence to become subgiants and red giants. During this post-main sequence stage of evolution, the star's radius increases and the temperature decreases, and these changes in stellar properties have profound effects on surrounding planetary systems. For example, the increase in stellar size and irradiation have been suspected of triggering planet radius inflation (e.g., Sha et al., 2021; Montalto et al., 2022; Grunblatt et al., 2022), while stellar and planetary tides are thought to be responsible for inward planet migration (e.g., Goldreich & Soter, 1966; Guillot et al., 1996; Villaver & Livio, 2009). While observations have already shown differences in the population of planets around main sequence and post-main sequence stars in terms of orbital eccentricity, radius, and mass (e.g., Jones et al., 2014; Grunblatt et al., 2018, 2019), the mechanisms and timescales that drive these changes are poorly understood and characterised (e.g., Lopez & Fortney, 2016; Grunblatt et al., 2017; Berger et al., 2020; Komacek et al., 2020). This is, in part, due to a lack of a statistically large enough sample of well characterised planets around evolved stars. For example, as of 19 January 2022, the NASA Exoplanet Archive lists only 109 confirmed transiting planets with mass measurement to better than 30% that are orbiting around evolved stars. The identification and characterisation of a larger sample of planet-hosting subgiants and giants will help to clarify the evolutionary pathways of planets beyond the main sequence, including the evolution of planets around white dwarfs (e.g., Vanderburg et al., 2020) and the future of our own Solar System. The scientific value of using visual vetting to identify planets beyond the main sequence is discussed further in Chapters 3 and 7.

Planets around pre-main sequence stars. Similarly, only a small number of planets around young, pre-main sequence stars (<100 Myrs) are known to date. Direct imaging has detected 37 large planets (mean $R \sim 18$ $R_\oplus$) on large orbits (mean $a \sim 360$ AU) around young stars (e.g., Bowler, 2016), while the NASA Exoplanet

Archive lists 15 young planets identified with the transit method (e.g., van Eyken et al., 2012; Mann et al., 2016; David et al., 2016, 2019; Zicher et al., 2022) and 4 identified with the RV method (e.g., Donati et al., 2016; Johns-Krull et al., 2016).

Studying these young systems can provide fundamental constraints on how planets form and evolve. For example, theoretical models of planet formation and evolution by Owen & Wu (2013) predict that super-Earths and mini-Neptunes have larger radii during their star's pre-main sequence stage compared to on the main sequence. As such, theoretical planet formation and evolution models can be tested by characterising a statistically significant number of pre-main sequence stars and comparing their properties to their main sequence counterparts. Similarly, the study of pre-main sequence giant planets may reveal insight into how these planets accrete their gaseous envelopes and whether they have rocky cores (e.g., Marley et al., 2007; Spiegel & Burrows, 2012; Mordasini, 2013; Mordasini et al., 2017).

Multi-planet systems. Multi-planet systems can help constrain theories of planet formation, migration, and tidal interactions through measurements of their mutual inclinations, orbital spacing, and planet sizes and compositions. RV and transit surveys have shown that multi-planet systems are intrinsically common, with *Kepler* alone identifying hundreds of multi-planet systems (e.g., Latham et al., 2011). These were often found to be in very compact ($P < 1$ year) systems composed of small ($R < 4$ $R_\oplus$) planets (e.g., Fabrycky et al., 2014; Winn & Fabrycky, 2015). Other studies found that planets in multi-planet systems tend to have lower eccentricities than single-planet systems, which is likely related to the long-term dynamical stability that comes with lower eccentricities in compact systems (Wright et al., 2011).

The study of multi-planet systems has benefited from observations of transit timing variations (TTVs), which have played an important role in measuring planet masses, as well as in uncovering additional non-transiting planets (e.g., Holman et al., 2010; Steffen et al., 2013; Kane et al., 2019; Kaye et al., 2022). TTVs are particularly significant (and measurable) when planets in adjacent orbits are in a mean motion resonance (MMR), i.e. when their orbital periods are simple integer

ratios of one another (Holman & Murray, 2005; Agol et al., 2005). There are a number of systems with MMRs or near MMRs, including Kepler-223 (4 planets in a 3:4:6:8 resonance; Mills et al., 2016), TRAPPIST-1 (7 planets with period ratios of 8:5, 5:3, 3:2, 3:2, 4:3, and 3:2; Luger et al., 2017), and K2-138 (5 planets in an unbroken 3:2 resonance chain; Christiansen et al., 2018). The latter of these was identified by citizen scientists taking part in the Exoplanet Explorers citizen science project, as further discussed in Section 1.7.

Circumbinary planets. Only $54\% \pm 2\%$ of solar-type stars are single, with the rest residing in pairs (stellar binaries) or higher order-multiple systems (Raghavan et al., 2010). An exciting find made by *Kepler* was the existence of planets around multi-stellar systems (circumbinary planets), with the first discovery being Kepler-16 (Doyle et al., 2011), closely followed by Kepler-34 and Kepler-35 (Welsh et al., 2012). To date, the sample of known circumbinary planets remains small with less than 20 confirmed systems. This is likely due to observational biases as transit events of circumbinary planets are neither periodic nor consistent in transit depth or duration (which depend on the phase of the binary at the time of the transit), making the detection of circumbinary planets using automated searches challenging (e.g., Armstrong et al., 2013). Visual inspection of photometric data with the help of citizen science has, however, proven to be a fruitful method for the identification of these elusive signals. For example, the Planet Hunters project identified the first planet in a quadruple star system in *Kepler* data (PH1 b; Schwamb et al., 2013); Planet Hunters TESS helped identify the first circumbinary planet in *TESS* data (TOI-1338 b; Kostov et al., 2021); and Exoplanet Explorers identified a small planet in a binary in the *K2* data (K2-288 b; Feinstein et al., 2019).

Habitable zone planets.[11] The detection of potentially habitable, Earth-like planets was one of *Kepler*'s primary mission objectives. Since then, there have been many studies that looked into quantifying the occurrence rates of Earth-like planets in the Habitable Zone (HZ) around Sun-like stars, often denoted as '$\eta_\oplus$'. However, all of these qualifying terms (HZ, Earth-like, and Sun-like) are not well defined

[11] For a not-so-serious discussion on the *Really Habitable Zone* – the only one that truly matters for any Gin-lovers out there – I leave the reader to peruse Pedbost et al. (2020).

across the literature. For example, numerous different radius constraints have been used to define Earth-like planets, including $R_p = 1-2\ R_\oplus$ (Petigura et al., 2013; Silburt et al., 2015), $R_p < 1.25\ R_\oplus$ (Borucki et al., 2011), $R_p = 1-1.75\ R_\oplus$ (Barbato et al., 2018), and $R_p = 0.5-1.4\ R_\oplus$ (Kopparapu et al., 2013). Similarly, while the HZ is broadly defined as the region around a star where a planet can sustain liquid water on the surface, the limits and width of the HZ are widely debated (e.g., Kasting et al., 1993; Pierrehumbert & Gaidos, 2011; Kopparapu et al., 2013; Zsom et al., 2013). These different definitions affect the derived values of $\eta_\oplus$.

The determination of the value of $\eta_\oplus$ is further complicated by the difficulties associated with detecting small planets with long orbital periods (see Section 1.4). Nonetheless, a small number of potential rocky HZ planets have been found. For example, in a recent review of HZ planets, Bryson et al. (2021b) report on 56 *Kepler* planet candidates with $R_p < 2.5\ R_\oplus$ that were found to lie close to their HZs. However, in order to convert this number into a planet occurrence rate ($\eta_\oplus$), the detection efficiency (i.e. knowledge of how 'complete' the sample is) of the survey needs to be quantified. One way to estimate the detection efficiency of a survey is by injecting and recovering simulated transit signals into/from light curves (e.g., Petigura et al., 2013; Christiansen et al., 2015, 2020).

There is currently little consensus on the *true* value of $\eta_\oplus$ among different studies, with values ranging from smaller than 0.03 planets per star (e.g., Catanzarite & Shao, 2011; Foreman-Mackey et al., 2014) to over 0.35 planets per star (e.g., Traub, 2012; Bryson et al., 2021b), with many values in between (e.g., Burke et al., 2015; Zink et al., 2019b). This large range in values is primarily due to the lack of consensus on the definition of the qualifying terms, the lack of a more complete sample, and the difficulties associated with determining the detection efficiencies of surveys. Studies of how planets form and evolve; the fraction of planets in multiple systems; and the long-term dynamical stability of circumbinary planets, are all examples of other factors that may influence the derived value of $\eta_\oplus$.

1.6.2 Importance of planetary masses and radii

The mass–radius diagram. Masses and radii are fundamental exoplanet properties that are needed for their characterisation. For example, the densities derived from these parameters provide insight into the interior structures and compositions of planets, which in turn can help constrain different formation mechanisms (see e.g., Winn & Fabrycky, 2015). The NASA Exoplanet Archive lists 704 planets that have both mass and radii measurements with a precision better than 30%. Figure 1.11 shows that these planets fall into three distinct populations: Terran (or rocky) planets, Neptunian planets with volatile rich envelopes, and H/He rich Jovian planets (e.g., Chen & Kipping, 2017; Otegi et al., 2020).

Otegi et al. (2020) empirically determined that the Terran planet population follows a mass–radius (M–R) relation of $R \propto M^{0.29\pm0.01}$ (blue line in Figure 1.11), with densities similar to those of the Earth.[12] The low densities of the planets in this population suggests that the amount of water or H/He is limited, and thus that these planets are rocky. This M–R relation for rocky planets holds until $\sim$25 $M_\oplus$, with planets between 10 and 25 $M_\oplus$ likely having ice-rich cores (given their slightly lower densities). Based on this, Otegi et al. (2020) suggest that the maximum possible planet core mass that can form is between 10 and 25 $M_\oplus$, a value that is consistent with theoretical predictions of the core masses of giant planets (e.g., Wahl et al., 2017; Bitsch et al., 2019).

The second, more volatile-rich, planet population follows an M–R relation of $R \propto M^{0.63\pm0.04}$ (orange line in Figure 1.11), with a notably larger dispersion in the densities compared to the Terran planets. This higher dispersion could be due to a combination of different underlying core masses and different amounts of irradiation from their host stars that affect the planets' ability to accrete and maintain gaseous envelopes. Furthermore, this population shows a scarcely populated region of the M–R parameter space between the mass of Neptune (17.15 $M_\oplus$) and Saturn (95.16 $M_\oplus$). This mass range corresponds to the transition region between volatile-poor ice giants

[12]It should be noted that different combinations of iron, silicates and water can result in the same density, and thus the composition may differ from that of Earth.

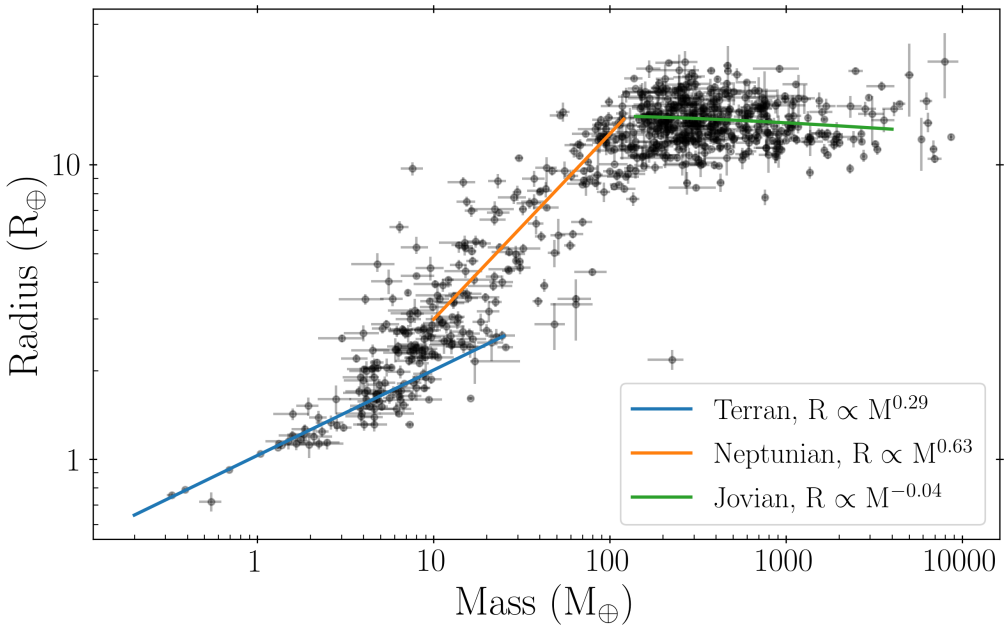

Figure 1.11: The M–R diagram for confirmed planets that have mass and radius measurements with a precision better than 30% (data from NASA Exoplanet Archive). The solid lines indicate three different planet regimes: Terran (or rocky) planets (blue; Otegi et al., 2020), Neptunian planets with volatile rich envelopes (orange; Otegi et al., 2020), and H/He rich Jovian planets (green; Chen & Kipping, 2017).

and volatile-rich gas giants. Studying the properties of planets in this transition region can help differentiate between different theoretical models of how gas and ice giants form, as discussed in more detail in Section 4.5.2.

Finally, an empirical study by Chen & Kipping (2017) found an additional transition point in the M–R diagram at ~ 0.41 M_{Jup}, beyond which $R \propto M^{-0.04 \pm 0.02}$ (green line in Figure 1.11). It has been predicted that the observed increasing density of these gas-rich Jovian planets (with increasing mass) is a result of self-compression of the atmospheres as the planets accrete more gas (e.g., Hatzes & Rauer, 2015; Chen & Kipping, 2017).

Photo-evaporation valley. One surprising finding to come out of transit surveys is that the most common type of planet orbiting around Sun-like stars appears to be planets that have no analogues within our own Solar System: Super-Earths and mini-Neptunes that are between 1 and 4 $R_\oplus$ (Fressin et al., 2013; Petigura et al., 2013). The M–R diagram revealed that this population consists of two distinct sub-populations: small rocky planets $R_p < 1.5$ $R_\oplus$ and larger gaseous

planets $2 < R_p < 3$ $R_\oplus$, with a deficiency of planets around $R_p = 1.99 \pm 0.20$ $R_\oplus$ (Lopez & Fortney, 2014; Fulton et al., 2017; Van Eylen et al., 2019). Theoretical models suggest that this radius gap may be due to photo-evaporation (Owen & Wu, 2013), whereby close-in sub-Neptune sized planets ($a < 0.1$ AU) are stripped of their atmospheres within a few hundred million years due to high levels of incidence flux from their host stars (Owen & Wu, 2013). This would mean that both populations have the same origin, with super-Earths formerly having been mini-Neptunes before they lost their envelopes. This demonstrates that by populating the M–R diagram with more planets around stars with precisely determined masses and radii (e.g., via asteroseismology), we may be able to identify more sub-populations that may provide further insight into planet formation and evolution.

Internal planet compositions. Precise knowledge of both the mass and radius also allows us to infer possible compositions and internal structures of planets (e.g., Seager et al., 2007; Howell et al., 2014; Van Hoolst et al., 2019). This is done by comparing the observed density of the planet to theoretical predictions of planet interiors. Figure 1.12, for example, shows the M–R relations derived by Zeng & Jacobsen (2017) of rocky planet interiors with different fractions of iron (Fe), silicates ($MgSiO_3$), and water (H_2O). The grey points show confirmed exoplanets ($M_p < 10$ $M_\oplus$ and $R_p < 3$ $R_\oplus$) that have mass and radius measurements to a precision of at least 30%, while the black points show measurements with a precision of at least 10%. The figure shows that small ($M_p \lesssim 3$ $M_\oplus$) rocky planets follow a tight M–R relation with a composition that is rich in silicates and iron, while planets with $M_p \gtrsim 3$ $M_\oplus$ show a larger scatter in their radii, with some containing a higher fraction of lighter materials. These theoretical tracks highlight the importance of precise mass and radius measurements in order to infer the interior structures of exoplanets and to help understand the overall demographics of rocky planets.

Currently, many studies are dedicated to populating the M–R diagram with increasingly precise measurements in order to further our understanding of planet populations. However, in order to be complete, planet population studies should also consider other planet and system properties including orbital period, multiplicity,

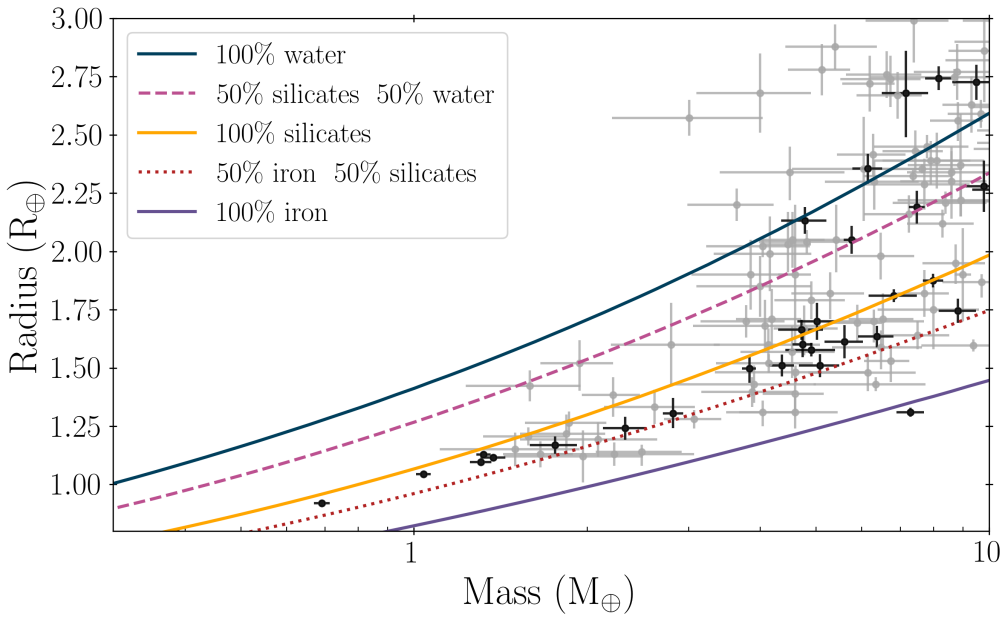

Figure 1.12: Mass–radius relation for rocky planets with different interior compositions (Zeng & Jacobsen, 2017). The grey points shows the planets with radius and mass measurements with a precision better than 30%, while the black points show measurements with a precision better than 10% (data from NASA Exoplanet Archive). The code to plot the Zeng & Jacobsen (2017) models was provided by O. Barragán.

and system age (e.g., Sousa et al., 2019; David et al., 2021). These properties will be discussed further throughout this work.

1.7 Citizen science in astronomy

1.7.1 Towards finding longer-period planets with *TESS*

The planet zoo has substantially grown over the past quarter century, both in size and in diversity. This diversity is only set to increase with current and future transit search missions. However, no matter the mission, there will always be a bias towards the detection of short-period planets due to the lower transit probabilities of longer-period planets and the period limits imposed by mission designs, combined with two transit detection limits imposed by automated transit searches. The requirement of multiple transit events, in particular, poses a problem for the automated detection of long-period planets in the *TESS* light curves. This is because, as previously discussed, only targets close to the ecliptic poles are

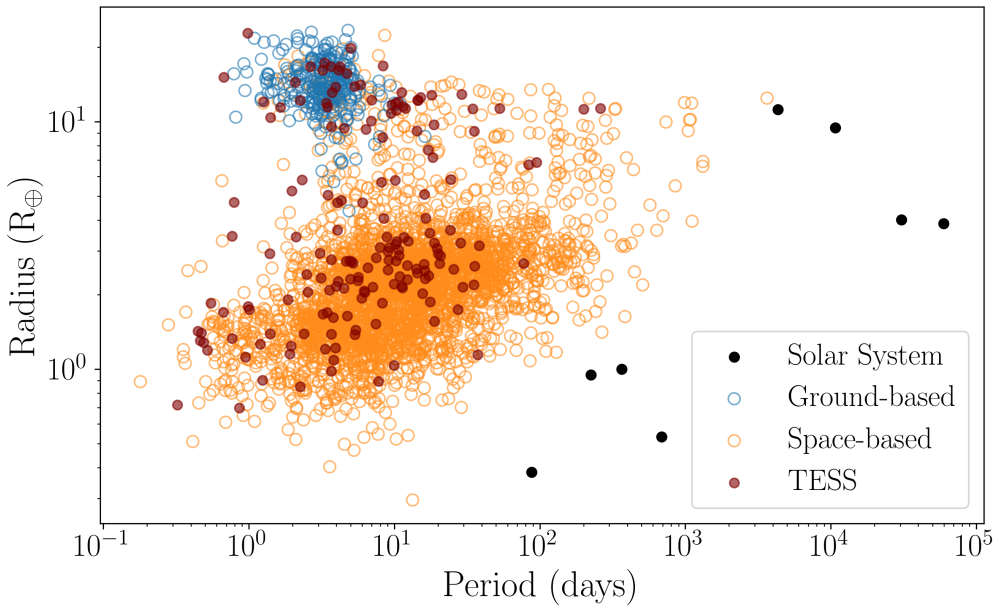

Figure 1.13: The properties of confirmed planets detected using ground-based (blue) transit surveys and space-based transit surveys (orange), as listed in the NASA Exoplanet Archive. The Solar System planets (black) are shown for comparison. The confirmed *TESS* planets are highlighted in maroon.

monitored across multiple consecutive observational sectors and will thus have light curves with longer observational baselines and the opportunity to detect longer-period planets with multiple transit events. The resulting bias towards short-period planets is reflected in the results of the pipeline-identified *TESS* planet candidates:[13] 90.5% of the first 5000 *TESS* Objects of Interest (TOI) have periods <15 days and 96.5% have periods <30 days. Similarly, 76.3% of all confirmed *TESS* planets have orbital periods shorter than 20 days and 91.9% have periods <30 days, as illustrated in Figure 1.13.

Non-standard methods, such as machine learning (e.g., Pearson et al., 2018; Zucker & Giryes, 2018), probabilistic transit model comparison (e.g., Foreman-Mackey et al., 2016), and citizen science (e.g., Fischer et al., 2012), can in some cases outperform standard transit search pipelines for longer-period planets, and are often sensitive to single-transit events that the pipelines routinely ignore.

[13] https://exofop.ipac.caltech.edu/tess/. The Exoplanet Follow-up Observing Program for TESS (ExoFOP-TESS) is a database that lists of all of the targets in the TESS Input Catalog (TIC), as well as any confirmed planets, planet candidates, and follow-up efforts. It is designed to optimise resources in the follow-up studies of *TESS* targets.

This motivated my systematic search for transits in the *TESS* data using the citizen science approach.

1.7.2 Past and current citizen science projects

The human brain excels in activities related to pattern recognition, making the task of identifying transit events in light curves, even when the pattern is in the midst of a strong varying stellar signal, ideally suited for visual vetting. Early online citizen science projects, such as Galaxy Zoo (Lintott et al., 2008), Planet Hunters (PH; Fischer et al., 2012), and Exoplanet Explorers (Christiansen et al., 2018) successfully harnessed the power of a large number of volunteers to identify signals in astrophysical data. More specifically, the latter two used citizen science to search for signals in *Kepler* and *K2* data, and made substantial contributions to the field of exoplanets. The PH project, for example, showed that human vetting has a higher detection efficiency than automated detection algorithms for certain types of transits. In particular, they showed that citizen science can outperform automated searches on the detection of single (long-period) transits (e.g., Wang et al., 2013; Schmitt et al., 2014), aperiodic transits (e.g., circumbinary planets; Schwamb et al., 2013; Feinstein et al., 2019), and planets around variable stars (e.g., young systems, Fischer et al., 2012).

Both PH and Exoplanet Explorers, which are hosted by the world's largest citizen science platform Zooniverse (Lintott et al., 2008, 2011), ensured easy access to *Kepler* and *K2* data by making them publicly available online in an immediately accessible graphical format that is easy to understand for non-specialists. The popularity of these projects is reflected in the number of participants, with PH for example attracting 144 466 volunteers from 137 different countries over 9 years of the project being active, and Exoplanet Explorer engaging almost 30 000 participants since April 2017.

Following the end of the *Kepler* and *K2* mission and the launch of the *TESS* satellite in 2018, I designed and launched a new citizen science project Planet

Hunters TESS (PHT)[14], with the aim of identifying transit events in the *TESS* data that were intentionally ignored or missed by the main *TESS* pipelines. Such a search complements other methods via its sensitivity to single-transit, and, therefore, longer-period planets.

PHT builds on the success of these earlier citizen science projects, and follows a similar workflow to the PH project. The PHT project, methodology, and results are the focus of the remainder of this work and will be discussed in detail in Chapters 2 and 3.

1.8 Thesis outline

The primary objective of this thesis is to demonstrate that with visual vetting with the help of citizen science we are sensitive to the detection of single-transit, longer-period planets. This allows us to identify planets that automated detection algorithms tend to miss and to address the period detection bias that is so prevalent in transit surveys, and in particular in *TESS* data. Furthermore, I show that a natural by-product of citizen science is the identification of unique planet and stellar signals. The remainder of this thesis will proceed as follows:

Chapter 2 outlines the PHT citizen science project and explains how I identify signals in the *TESS* data. Furthermore, I present the `LATTE` vetting suite used to rule out false-positive signals and discuss the application of the `pyaneti` transit modelling code. Finally, I describe the ground-based follow-up efforts used to observe the most promising planet candidates identified via PHT.

Chapter 3 discusses the transit recovery efficiency of PHT by assessing how many of the pipeline identified planet candidates (TOIs) were recovered with visual vetting. I also provide an overview of the 139 planet candidates that were identified with PHT in the data collected in the first three years of the *TESS* mission, as well as a summary of the multi-stellar systems that have been flagged on the citizen science platform.

[14] www.planethunters.org

Chapter 4 presents the detection and characterisation of a sub-Saturn sized planet on an 84-day orbit around a bright, post-main sequence G-dwarf identified with the help of citizen science. I also discuss the scientific value of longer-period planets around bright stars, as well as the prospects for future mass measurements of this target.

Chapter 5 discusses the discovery and characterisation of a two-planet system flagged on the PHT platform. I show that both planets in this system are prime candidates for future atmospheric characterisation due to the brightness of the host star and the extended gaseous atmospheres of the planets.

Chapter 6 presents a triple-star system that was identified on the PHT discussion forums. The combined modelling of the eclipse timing variations and RVs revealed a unique stellar configuration wherein the outer star is more massive than the combined mass of the inner binary. I, therefore, provide a discussion about the formation and future evolution of this system.

Chapter 7 summarises the results presented throughout this work, discusses the implications of my findings, and outlines the future prospects for PHT.

Books can be misleading.

— G. Lockhart, 1992

CHAPTER 2

TOOLS AND METHODS

One of the primary science goals of the *TESS* mission is to further our understanding of the overall planet population, an active area of research that is strongly affected by observational and detection biases. As discussed in Chapter 1, one way to address the bias towards short-period planets in transit surveys is via visual vetting with the help of citizen science.

In this chapter, I present the PHT citizen science project, the candidate detection methodology, as well as the process and tools used for vetting and analysing the candidates discovered on the platform. I also outline the ground-based follow-up campaigns that I use to observe the most promising planet candidates identified via PHT.

2.1 Planet Hunters TESS

I designed and built the PHT project, which engages members of the general public in the task of identifying new, transiting exoplanet in *TESS* data. The project was built using Zooniverse's Panoptes Project Builder (Lintott et al., 2008), which allows researchers to build complex classification workflows consisting of different user tasks (e.g., object identification, text transcription etc.), using a simple online tool. To date, over 200 citizen science projects, including PHT, have been built using this tool and are now hosted by the Zooniverse platform (Lintott et al., 2008, 2011).

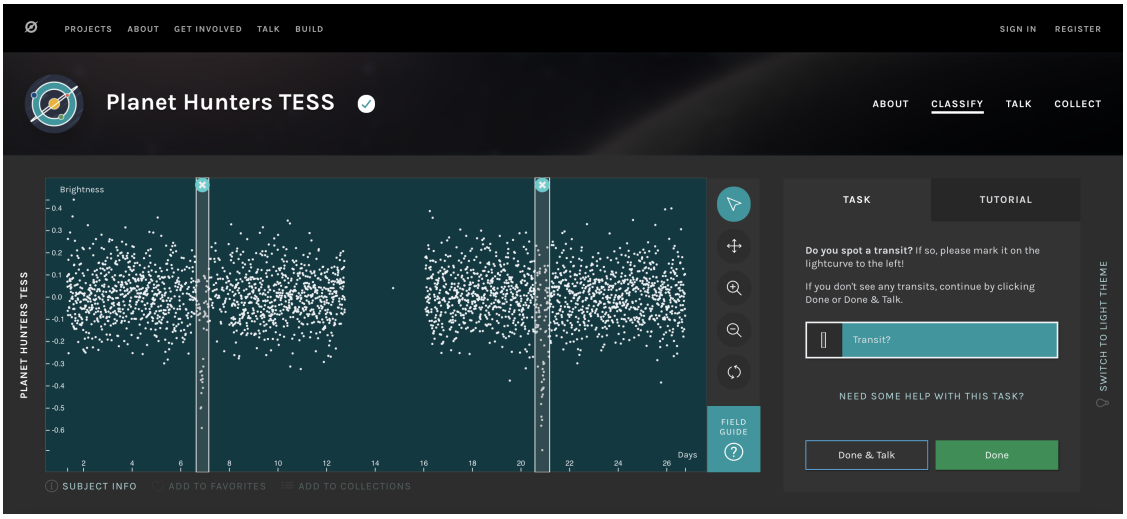

Figure 2.1: PHT user interface showing a simulated light curve. The transit events are highlighted with white, partially-transparent, columns that are drawn on using the mouse. Stellar information on the target star is available by clicking on 'subject info' below the light curve.

PHT works by displaying *TESS* light curves (Figure 2.1) and asking volunteers to identify transit-like signals. Only the two-minute cadence SPOC light curves (see Section 1.5.2) are searched by PHT. The *TESS* data are uploaded to PHT using the Panoptes python client[1] within a couple of days of the (approximately monthly) public data releases and only one sector of *TESS* data is displayed on the site at any given time.

First-time visitors to the PHT site, or returning visitors who have not logged in, are prompted to look through a short tutorial, which briefly explains the main aim of the project and shows examples of transit events and other stellar phenomena. The scientific explanation of the project can be found elsewhere on the site in the 'field guide', which provides numerous examples of exoplanet transits, various stellar signals, and systematics, as well as on the project's 'About' page, which outlines the scientific motivation for PHT.

After viewing the tutorial, the volunteers are ready to participate in the project and are presented with *TESS* light curves (known as 'subjects') that need to be classified. The project was purposely designed to be as simple as possible and therefore only asks one question: *'Do you see a transit?'*. Users identify transit-like

[1] https://github.com/zooniverse/panoptes-python-client

events, and the time of their occurrence, by drawing a column over the event, as shown in Figure 2.1, using the mouse button. There is no limit on the number of transit-like events that can be marked in a light curve. No markings indicate that there are no transit-like events present in the data. Once the subject has been analysed, users submit their classification and continue to view the next light curve by clicking the 'Done' button on the screen. The aim of the simplicity of the task is to encourage participation from a large number of users.

Alongside each light curve, users are offered information on the stellar properties of the target, such as the radius, effective temperature, and magnitude (subject to availability; see Stassun et al., 2018). However, in order to reduce biases in the classifications, the *TESS* Input Catalog (TIC) ID of the target star is not provided until after the subject classification has been submitted.[2]

In addition to classifying the data, users are given the option to comment on light curves via the 'Talk' discussion forum once they have classified the data. Each light curve has its own discussion page to allow volunteers to discuss and comment, as well as to 'tag' light curves using searchable hashtags, and to bring promising candidates to the attention of other users and the PHT research team. The talk discussion forums complement the main PHT analysis and have been shown to yield interesting objects which may be challenging to detect using automated algorithms (e.g., Eisner et al., 2019, 2022).

The subject TIC IDs are revealed on the subject discussion pages, allowing volunteers to carry out further analysis on specific targets of interest and to report and discuss their findings. This is extremely valuable for both other volunteers and the PHT science team, as it can speed up the process of identifying candidates as well as help rule out false positives in a fast and effective manner. For example, these discussion forums are regularly used by citizen scientists to post additional analysis that they have carried out, such as tests that help rule out false positives (see Section 2.2.4). In order to aid the citizen scientists in carrying out these types of additional analyses, I created a number of tutorial videos and provided open-access

[2]Submitted classifications cannot be altered by the user.

code that allows non-specialists to carry out a number of standard vetting tests without needing any prior knowledge in programming (see Section 7.3.1). I made all of the materials and tutorials available online at wwww.planethunters.coffee.

It is important to note that that there are many similarities between the PH and PHT citizen science projects. Both PH and PHT display photometric light curves (from *Kepler* and *TESS*, respectively) and ask volunteers to identify transit-like signals by marking them with their mouse button, with the aim of identifying signals that are missed by other searches. The main methodological difference between the two projects is that PH additionally asked users to identify whether the star is variable or not, and whether there are any large gaps in the light curve (Fischer et al., 2012; Schwamb et al., 2012). Instead of asking these questions in the interface, PHT encourages citizen scientists to flag stellar variability and other phenomena on the discussion forums in order to keep the main task as simple as possible.

2.1.1 User interface

Since the launch of PHT on 6 December 2018, there has been one significant change to the user interface. The initial PHT user interface (UI1), which was used for Sector 1 through 9, split the *TESS* light curves up into either three or four chunks (depending on the data gaps in each sector) which lasted around seven days each. This allowed for a more 'zoomed' in view of the data, making it easier to identify transit-like events than when the full $\sim$30-day light curves were shown.

The decision to split each light curve into four chunks was based on the results from a project beta test. During this beta test, I randomly presented citizen scientists with one out of the two test projects: beta v1 displayed the full 27-day light curves and beta v2 displayed zoomed-in sections of light curves ($\sim 6.5-9$ days of data each). For both versions of the beta test, the only data on the site were simulated light curves, out of which 50% contained at least one injected transit-like signal. The sample was purposely chosen to over-represent light curves containing planets in order to obtain a better understanding of the types of planets that could be identified via visual inspection. This test showed that the zoomed-in version of

the project had a higher true-positive rate (56% for beta v2 vs 35% for v1) and a lower false-positive rate (0.5% for beta v2 vs 3.3% for v1).

Following the classification of Sector 9, the user interface was updated (UI2) to allow users to manually zoom in on the x-axis (time axis) of the data, making the project more interactive. This, again, was motivated by the beta test that showed that allowing for a more zoomed-in view of the light curves yields a higher transit recovery rate. I also note that UI2 is more similar to the original PH interface, which allowed volunteers to freely zoom in and out on the x-axis (Schwamb et al., 2012).

Due to zoom feature, PHT has been displaying each target as a single light curve as of Sector 10. In order to verify that the changes in interface did not affect our findings, all of the Sector 9 subjects were classified using both UI1 and UI2. I found that the fraction of TOIs recovered within the top 500 highest scored candidates (see below for details on how candidates are scored) was 45.9% for UI1 and 46.6% UI2. As discussed in Chapter 3 and shown in Figure 3.6, the fractional recovery rate of TOIs varies across different sectors by up to 20%. As such, the 0.7% change between UI1 and UI2 is not considered to be significant.

2.1.2 Simulated data

In addition to the real data, volunteers are shown simulated light curves, which are generated by randomly injecting simulated transit signals, provided by the SPOC pipeline (Jenkins et al., 2018), into real *TESS* light curves. The injected signals are randomly selected from a list of 1730 simulations that range in period from 0.2 to 20 days and in depth from 25 to 6800 ppm. These simulated data play an important role in assessing the sensitivity of the project, in training the users and providing immediate feedback, as well as to gauge the relative abilities of individual users (see Section 2.1.4).

For each injected signal a S/N ratio is calculated by dividing the injected transit depth by the Root Mean Square Combined Differential Photometric Precision (RMS CDPP) of the light curve on 0.5-, 1- or 2-hour time scales (whichever is closest to the duration of the injected transit signal). The RMS CDPP is the root mean

square of the photometric noise as experienced by a specific transit duration for a given target, where the RMS scatter is defined as:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\,\Sigma_{i=1}^{N}\left(d_i - \bar{d}\right)^2} \tag{2.1}$$

where N is the total number of data points, d_i is the flux measurement and $\bar{d}$ is the mean flux. For example, a CDPP of 50 ppm for a transit with a duration of 1 hour, would indicate that a transit of that duration and with a transit depth of 50 ppm would have a S/N ratio of 1. (Jenkins et al., 2010; Christiansen et al., 2012). RMS CDPP values are provided as part of the standard data products for all light curves reduced with the SPOC pipeline.

As discussed further in Chapter 3, the transit detection efficiency rapidly declines below a S/N ratio of ~ 7.5. While it is important to display simulated signals with S/N < 7.5, in order to quantify the detection efficiency for smaller planets and to calibrate user weights (Section 2.1.4), presenting the citizen scientists with a high frequency of simulations which are difficult, or even impossible, to detect can be discouraging to the volunteers. As such, only simulations with a S/N greater than 7 were shown to volunteers for UI1. With the launch of the new project interface (UI2), I decided that it would be valuable to inject lower S/N transits into the data in order to further quantify the detection efficiency of smaller planets. As such, UI2 displays simulated signals with S/N > 4. In retrospect, I believe that UI1 could have also benefited from displaying simulated transits with lower S/N ratios. The detection of simulated transit events is discussed further in Section 3.1.

Simulated light curves are randomly shown to the volunteers and classified in the exact same manner as the real data. The user is always notified *after* a simulated light curve has been classified and given feedback as to whether the injected signal was correctly identified or not. This feedback is given in the form of a pop-up window showing the simulated light curve overlaid with green columns on top of correctly identified transits and red columns over the not-identified transits, as shown in Figure 2.2. For each sector, between one and two thousand simulated light curves are generated using the real data from that sector in order to ensure

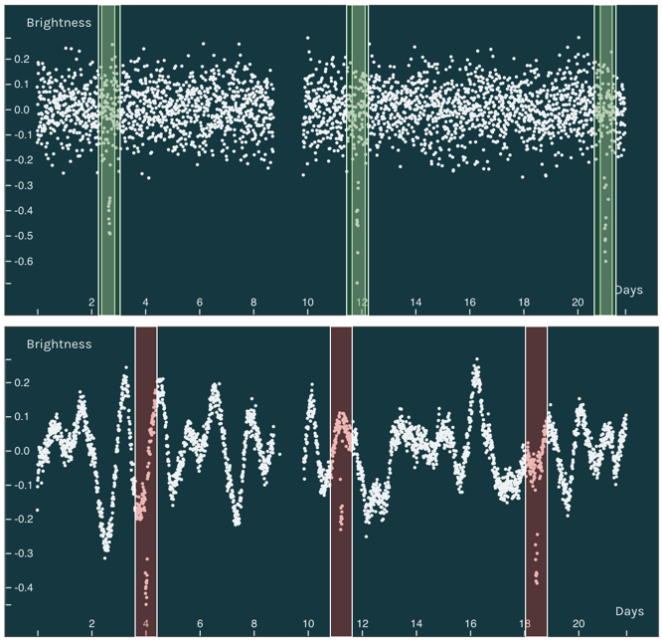

Figure 2.2: Feedback is provided following the classification of each simulated light curve. Correctly identified transit events are overlaid with green columns (top panel) and missed transits are overlaid with red columns (bottom panel).

that the sector specific systematic effects and data gaps of the simulated data do not noticeably differ from the real data. The rate at which a volunteer is shown simulated light curves decreases from an initial rate of 30% for the first 10 classifications, down to a rate of 1% by the time that the user has classified 100 light curves. The higher rate of simulations for new users allows me to assess their skill and to provide direct feedback on their classifications while they are new to the project. Conversely, the lower rate of simulations after users have classified 100 light curves aims to reduce the disruption in classifications for volunteers who look at large volumes of data, while still showing simulations at a cadence high enough to be useful for assessing their skill.

2.1.3 Identifying candidates

Each subject is seen by multiple volunteers, before it is removed – or 'retired' – from the site, and the classifications are combined in order to assess the likelihood of a transit event (see Section 2.1.6). For Sectors 1 through 9, there were three different subject retirement rules: 1) the subject was retired after 8 classifications if the first

8 volunteers who saw the light curve did not mark any transit events. This rule assumes that if 8 people identified no signals in the light curve then the probability of there being a detectable transit event is low. 2) The subject was retired after 10 classifications if the first 10 volunteers all marked a transit event in the light curve. Similarly to rule 1, this assumes that if there is complete agreement among the users, there is no need for further classifications. 3) The subject was retired after 15 classifications if there was not complete consensus amongst the users. This rule assumes that if the users cannot agree on whether there is a signal or not, then more classifications may help to clarify whether any of the flagged signals are significant. As of Sector 9 with UI2, all subjects were classified by 15 volunteers, regardless of whether or not any transit-like events were marked. Sector 9, which was classified with both UI1 and UI2, was also classified with both retirement rules.

There were a total of 18 384 290 individual classifications completed across the project on data from the first three years of the *TESS* mission (Sectors 1 – 39). Around 95.3% of these classifications were made by 30 874 volunteers who have registered with the Zooniverse platform, with the rest made by unregistered (or not logged in) volunteers. Around 77% of the registered volunteers completed more than 100 classifications, 10% more than 300, 7.2% more than 500, 4.6% more than 1000 and 0.9% more than 10 000. The registered volunteers completed a mean and median of 611 and 29 classifications, respectively. Figure 2.3 shows the distribution in user effort for logged in users who made between 0 and 300 classifications. As a comparison, the original PH project saw 98% of their classifications completed by registered volunteers, with the remaining from unregistered volunteers (Schwamb et al., 2012).

The distribution in the number of classifications made by the registered volunteers is assessed using the Gini coefficient, which ranges from 0 (equal contributions from all users) to 1 (large disparity in the contributions). The Gini coefficients for individual sectors range from 0.83 to 0.91 with a mean of 0.87, while the Gini coefficient for the overall project (all of the sector combined) is 0.94. The mean Gini coefficient among other astronomy Zooniverse projects lies at 0.82 (Spiers et al., 2019). I note that the only other Zooniverse project with an equally

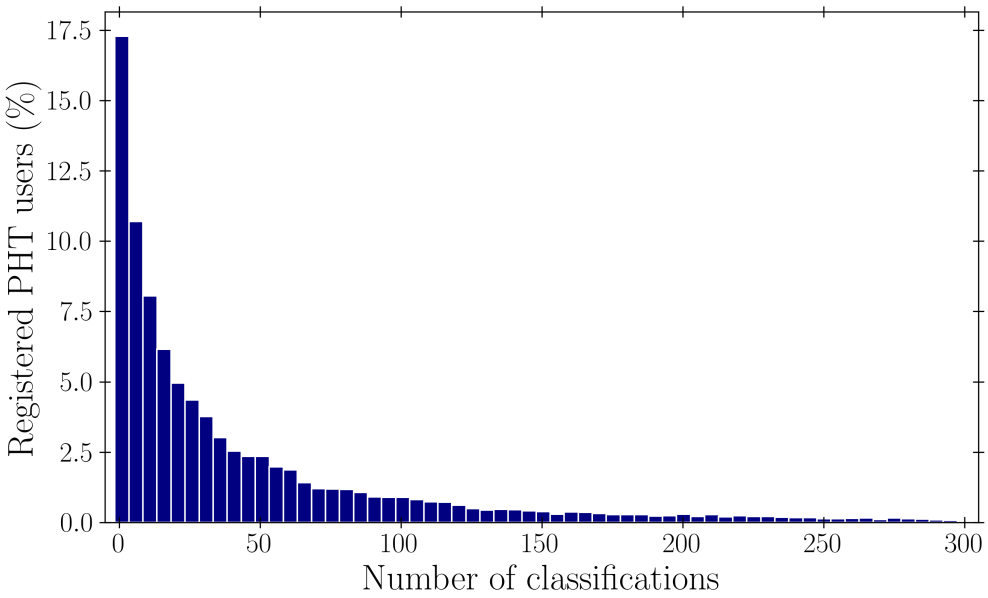

Figure 2.3: The distribution of the number of classifications by the registered volunteers, using a bin size of 5 from 0 to 300 classifications. A total of 11.8% of the registered volunteers completed more than 300 classifications.

high Gini coefficient as PHT is *Supernova Hunters*, a project which, similarly to PHT and unlike most other Zooniverse projects, has periodic data releases that are accompanied by an e-newsletter sent to all project volunteers. Periodic e-newsletters have the effect of promoting the project to both regularly and irregularly participating volunteers, who may only complete a couple of classifications as they explore the task, as well as to returning users who complete a large number of classifications following every data release, increasing the disparity in user contributions. The fact that the project is able to maintain a long-term user base, as well as encourage new users to explore the task (even if not all of them continue beyond a couple of classifications) is promising for the future growth of PHT.

2.1.4 User weighting

User weights are calculated for each individual volunteer in order to identify users who are more sensitive to detecting transit-like signals and those who are more likely to mark false positives. The PHT weighting scheme is based on the weighting scheme used by the original PH project, as described by Schwamb et al. (2012).

User weights are calculated independently for each Sector, using the simulated light curves shown alongside the data from that sector. All users start off with a weighting of one, which is then increased or decreased when a simulated transit event is correctly or incorrectly identified, respectively.

The markings of simulated transits are deemed correctly identified, or 'True', if the mid-point of a user's marking falls within the width of the simulated transit events. If none of the user's markings fall within this range, the simulated transit is deemed not identified, or 'False'. If more than one of a user's markings coincide with the same simulated signal, it is only counted as being correct once, such that the total number of 'True' markings cannot exceed the number of injected signals. For each classification, the number of 'Extra' markings is recorded, which is the total number of markings made by the user minus the number of correctly identified simulated transits.

Each simulated light curve, identified by superscript i (where $i = 1, \ldots, N$) was seen by $K^{(i)}$ users (the mean value of $K^{(i)}$ was 10), and contained $T^{(i)}$ simulated transits (where $T^{(i)}$ depends on the period of the simulated transit signal and the duration of the light curve). For a specific light curve i, each user who saw the light curve is identified by a subscript k (where $k = 1, \ldots, K^{(i)}$) and each injected transit by a subscript t (where $t = 1, \ldots, T^{(i)}$).

In order to distinguish between users who are able to identify obvious transits and those who are also able to find those that are more difficult to see, a 'recoverability' $r_t^{(i)}$ was defined for each injected transit t in each light curve. This is defined empirically, as the number of users who identified the transit correctly divided by $K^{(i)}$ (the total number of users who saw the light curve in question).

Next, the performance of each user on each light curve was quantified as follows (this performance is analogous to the 'seed' defined in Schwamb et al. 2012):

$$p_k^{(i)} = C_{\mathrm{E}} \frac{E_k^{(i)}}{\langle E^{(i)} \rangle} + \sum_{t=1}^{T^{(i)}} \begin{cases} C_{\mathrm{T}} \left[r_t^{(i)} \right]^{-1}, & \text{if } m_{t,k}^{(i)} = \text{'True'} \\ C_{\mathrm{F}} \, r_t^{(i)}, & \text{if } m_{t,k}^{(i)} = \text{'False'} \end{cases} \tag{2.2}$$

where $m_{t,k}^{(i)}$ is the identification of transit t by user k in light curve i, which is either 'True' or 'False'; $E_k^{(i)}$ is the number of 'Extra' markings made by user k

for light curve i, and $\langle E^{(i)} \rangle$ is the mean number of 'Extra' markings made by all users who saw subject i. The parameters C_{E}, C_{T} and C_{F} control the impact of the 'Extra', 'True' and 'False' markings on the overall user weightings, and are optimised empirically as discussed in Section 2.1.7.

Following Schwamb et al. 2012, a global 'weight' w_k is assigned to each user k, which is defined as:

$$w_k = I \times (1 + \log_{10} N_k)^{\sum_i p_k^{(i)}/N_k} \tag{2.3}$$

where I is an empirically determined normalisation factor, such that the distribution of user weights remains centred on one, N_k is the total number of simulated transit events that user k assessed, and the sum over i concerns only the light curves that user k saw. The user weights are limited to the range $0.05-3$ *a posteriori* as to never completely disregard a user's classification and to not give too much significance to any one user's markings.

A number of alternative ways to define the user weights, including the simpler $w_k = \sum_i p_k^{(i)}/N_k$, were tested but Equation 2.3 was found to yield the best results (see Section 3.1 for how this was evaluated).

2.1.5 Systematic removal

Systematic effects, for example caused by the spacecraft or background events, can result in spurious signals that affect a large subset of the data. This can result in an excess in markings of transit-like events made by volunteers at certain times within a sector. As the four *TESS* cameras can yield unique systematic effects, the times of systematics were identified uniquely for each camera. The times were identified using a Kernel Density Estimation (KDE; Rosenblatt, 1956) applied across all of the markings from that sector for each camera, where the KDE is defined as:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K(x) \left(\frac{x - x_i}{h} \right) \tag{2.4}$$

where x_1,...x_n are the data points, *K(x)* is the Kernel, and $h > 0$ is the smoothing parameter known as the bandwidth. A cosine kernel ($K(x) = \frac{\pi}{4} cos(\frac{\pi}{2}x)$) with a bandwidth of 0.1 days was used for this analysis.

Figure 2.4 shows the KDE of all marked transit-events made during Sector 17 for *TESS*'s cameras 1 (top panel) to 4 (bottom panel). The isolated spikes, or prominences, in the number of marked events, such as at $T = 21 - 22$ days in the bottom panel, are assumed to be caused by systematic effects that affect multiple light curves. Such effects would have prompted a large number of citizen scientists to flag events at similar times across different light curves. Prominences are considered significant if they exceed a factor four times the standard deviation of the kernel output. This cut-off was empirically determined to be the highest cut-off to not miss clearly visible systematics via visual inspection of the systematics of randomly selected light curves.

All user-markings within the full width at half maximum of these peaks are omitted from all further analysis. The fraction of markings that are removed varies depending on the amount of systematic signals present within the sector but tends to be around 5 and 10% of the markings.

2.1.6 Identification of transit-like signals

The times and likelihoods of transit-like events are determined by combining all of the classifications made for each subject and identifying times where multiple volunteers identified the same signal. This was done using an unsupervised machine learning method, known as a DBSCAN (Density-Based Spatial Clustering of Applications with Noise; Ester et al., 1996). DBSCAN is a non-parametric density based clustering algorithm that helps to distinguish between dense clusters of data and sparse noise. For a data point to belong to a cluster it must be closer than a given distance (ϵ) to at least a set minimum number of other points (*minPoints*).

In our case, the data points are one-dimensional arrays of the times of transits, as identified by the volunteers, and clusters are times where multiple volunteers identified the same event. For each cluster a 'transit score' (s_i) is determined, which

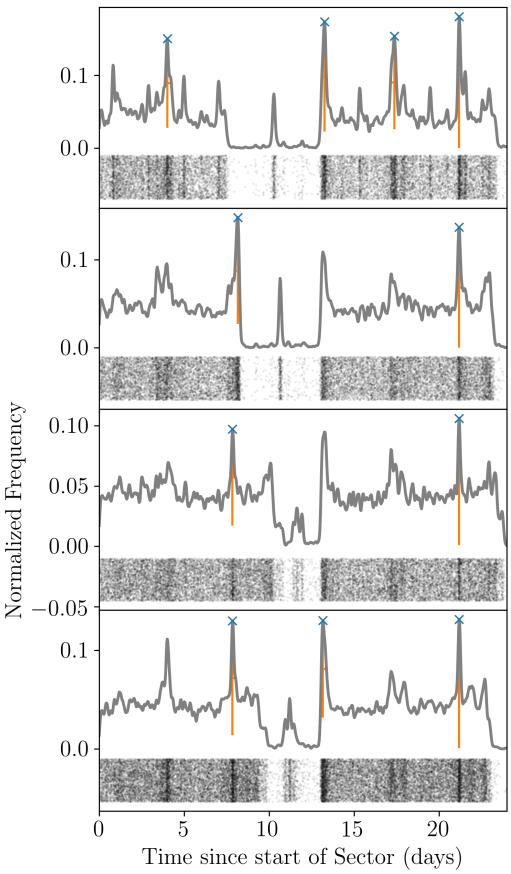

Figure 2.4: Kernel density estimation of the user-markings made for Sector 17, for targets observed with *TESS*'s observational Cameras 1 (top panel) to 4 (bottom panel). The orange vertical lines the indicate prominences that are at least four times greater than the standard deviation of the distribution. The black points underneath the figures show the mid-points of all of the volunteer-markings, where darker regions represent a higher density of markings.

is the sum of the user weights of the volunteers who contribute to the given cluster divided by the sum of the user weights of all of the volunteers who saw that light curve. These transit scores are used to rank subjects from most to least likely to contain a transit-like event. Subjects which contain multiple successful clusters with different scores are ranked by the highest transit score.

2.1.7 Optimising the search

The methodology described in Sections 2.1.4 to 2.1.6 has five free parameters: the number of markings required to constitute a cluster (*minPoints*), the maximum separation of markings required for members of a cluster (ϵ), and C_{E}, C_{T} and

C_F used in the weighting scheme. The values of these parameters were optimised via a grid search, where C_E and C_F ranged from -5 to 0, C_T ranged from 0 to 20, and *minPoints* ranged from 1 to 8, all in steps of 1. The value of ϵ ranged from 0.5 to 1.5 in steps of 0.5. This grid search was carried out on four randomly selected sectors, two from UI1 (Sectors 4 and 7) and two from UI2 (Sectors 13 and 19), for various variations of Equation 2.3.

The success of each combination of parameters was assessed by the fractions of TOIs and TCEs that were recovered within the top highest ranked 500 candidates (see Section 3.1 for details), where TCEs (Threshold Crossing Events) are any periodic signals that the automated transit search identified in the data, while TOIs (Tess Objects of Interest) have passed an additional round of vetting and are considered promising planet candidates (as discussed in Section 1.5.3). The most successful combination of parameters was found to be $minPoints = 3$ markings, $\epsilon = 1$ day, $C_T = 3$, $C_F = -2$ and $C_E = -2$.

2.1.8 MAST deliverables

The analysis described above is carried out both in real-time as classifications are made, as well as offline after all of the light curves of a given sector have been classified. When the real-time analysis identifies a successful DB cluster (i.e. when at least four citizen scientists identified a transit within a day of the *TESS* data of one another), the potential candidate is automatically uploaded to the open access Planet Hunters Analysis Database (PHAD)[3] hosted by the Mikulski Archive for Space Telescopes (MAST). While PHAD does not list every single classification made on PHT, it does display all transit candidates which had significant consensus amongst the volunteers who saw that light curve, along with the user-weight-weighted transit scores. This analysis does not apply the systematics removal described in Section 2.1.5. The aim of PHAD is to provide an open-source database of potential planet candidates identified by PHT, and to credit the volunteers who identified said targets.

[3] https://mast.stsci.edu/phad/. This data base is primarily being used by citizen scientists who wish to check their classifications.

The offline analysis is carried out once all of the data from a given sector have been retired from the site. The combination of all of the classifications allows me to identify and remove times of systematics and to calculate better calibrated and more representative user weights compared to the real-time analysis. The remainder of this work will only discuss the results from the offline analysis.

2.1.9 Preliminary candidate vetting

For each sector I rank all of the subjects according to their transit scores and visually inspect the 500 highest ranked targets in order to identify potential candidates and rule out false positives.[4] A vetting cut-off of 500 was chosen as this was found to maximise the number of found candidates while minimising the number of likely false positives (see Chapter 3, Figure 3.2).

In the initial round of candidate vetting the light curves of the 500 highest ranked candidates (not including TOIs or known planets) are uploaded to a separate Zooniverse classification interface that is not accessible to the public. This interface allows me to visually inspect and sort each candidate into either '*keep for further analysis*', '*eclipsing binary*', or '*discard*'. This sorting is based on the inspection of the full *TESS* light curve of the target, with the times of the satellite momentum dumps indicated, and of the background flux and the x and y centroid positions (see Section 2.2.4) around the time of each likely transit event, i.e. around the times of successful DB clusters. Stellar parameters are provided for each candidate, subject to availability, alongside links to the SPOC Data Validation (DV) reports for candidates that had been flagged as TCEs but were never promoted to TOI status. This vetting stage allows me to rapidly discard obvious astrophysical and instrumental false positives.

Candidates for which I indicate that the signal is consistent with a planetary transit are kept for further analysis; around 5% of the 500 highest ranked candidates per sector are kept. Considering that known planets and TOIs are not included at this stage of vetting, it is not surprising that our retention rate is lower than the

[4]This analysis was carried out as part of a team for several Sectors.

fraction of TCEs that are promoted to become TOIs (typically ∼8% per sector; see Section 3.1.2). Furthermore, the PHT false-positive rate is consistent with the findings of the initial PH project (Schwamb et al., 2012), who kept around 3.5% of their *Kepler* planet candidates following their initial round of vetting.

The rest of the 500 PHT candidates were grouped into ∼37% '*eclipsing binary*' and ∼58% '*discard*'. The most common reasons for discarding light curves are due to events caused by momentum dumps and due to astrophysical background events, such as background eclipsing binaries that mimic transit-like signals in the light curve. The targets identified as eclipsing binaries are analysed further by the *TESS* Eclipsing Binaries Working Group (Prša et al., 2022), as described in Section 3.3.1.

In order to verify that a vetting cut-off rank of 500 was not too low, I vetted Sectors 5, 14, 19, and 20 to a rank of 1000. Across all of these four randomly chosen sectors, I found two additional planet candidate at a rank greater than 500. As such, the scientific return from vetting all sectors to a rank of 1000 was deemed too small given the time that would be required to do this.

2.2 Signal vetting with LATTE

Due to the large pixels of the *TESS* satellite (21 arcsec, compared to 4 arcsec of the *Kepler* mission), false positives (e.g., where the signal of a target is contaminated by a faint, nearby, eclipsing binary) are very common. In order to aid in the detection and characterisation of planetary and stellar signals found in *TESS* data I designed and developed an open-source vetting suite known as the Light Curve Analysis Tool for Transiting Exoplanets (LATTE; Eisner et al., 2020a). All planet candidates that were classified as '*keep for further analysis*' during the preliminary vetting stage (Section 2.1.9) are further vetted for false positives using LATTE.

2.2.1 Vetting suite overview

LATTE is an open-source Python package that performs standard diagnostic tests to help identify, vet, and characterise signals in the *TESS* light curves in order to weed out instrumental and astrophysical false positives. I specifically developed

the program to allow for a fast, in depth analysis of targets that have already been identified as promising candidates by the main *TESS* pipelines or via alternative methods such as citizen science (e.g., Fischer et al., 2012; Christiansen et al., 2018; Eisner et al., 2021a). The implemented diagnostic tests are similar to the 'Data Validation' steps performed by the Science Processing Operations Center (SPOC; Jenkins et al., 2016) pipeline for the short-cadence data, and by the quick-look pipeline (QLP; Kunimoto et al., 2021) for the long-cadence data. However, while the SPOC and QLP pipeline tests are performed on a selection of pre-selected *TESS* targets only, `LATTE` allows for the analysis of any *TESS* target with a valid TIC ID, including both the two-minute and thirty-minute cadence targets.

Other tools that are often used to extract and vet *TESS* light curves include the open-source packages LIGHTKURVE (Lightkurve Collaboration et al., 2018) and ELEANOR (Feinstein et al., 2019), for the short and long-cadence data respectively. These tools can perform similar diagnostic tests to the ones provided by `LATTE`, however, `LATTE` differs in that it does not require any interaction with *Python* coding. Instead, `LATTE` allows for an interactive diagnostic analysis via a Graphical User Interface (GUI) and summarizes multiple diagnostic tests in an automatically generated pdf report. The motivation for the GUI and ease of use of the tool was to make these tests accessible to citizen scientists as well as to professional astronomers.

2.2.2 Running `LATTE`

The tool, which is executed via the command line, prompts the user to enter the TIC ID of a target star. Making use of the open-source TESS-POINT *Python* package (Burke et al., 2020), `LATTE` identifies the sector(s) in which the target was observed, and asks the user to identify which sector(s) they wish to inspect and analyse (Figure 2.5). The corresponding data are downloaded directly from MAST and displayed in a GUI, as shown in Figure 2.6. The user identifies times of transit-like events by directly clicking on either the full or zoomed-in version of the displayed light curve. Transit-times are recorded using the 'Add time' button. All additional settings are explained in the online GitHub documentation and

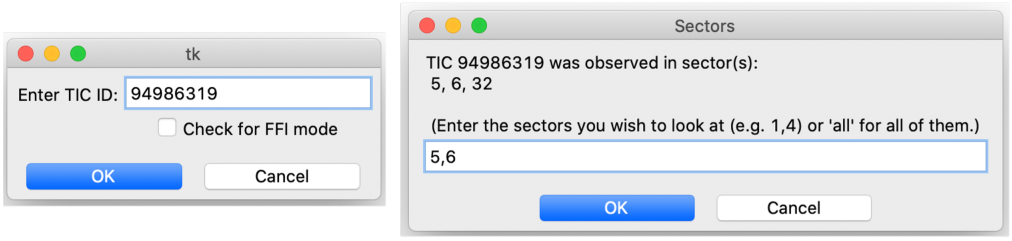

Figure 2.5: LATTE prompts. *Left panel*: prompt to enter the TIC ID of a target. *Right panel*: prompt to enter the sector(s) that the user wishes to inspect further. Prompts are completed for example light curve TIC 94986319 (TOI 421).

published in the Journal of Open Source Software. Once the times of one or more transit-like events have been identified the program automatically downloads and processes the *TESS* light curves.

2.2.3 Handling of *TESS* data

Short cadence

By default LATTE runs in *standard mode* using the pre-selected short-cadence *TESS* data and light curves that were extracted and detrended by the SPOC pipeline (see Section 4.2.1).

Long cadence

The *FFI mode* can be selected with a check-box at the time when the TIC ID is entered (left panel of Figure 2.5). LATTE queries the MAST archives to determine whether the given target light curve has been extracted with either the SPOC-FFI (Caldwell et al., 2020; Huang et al., 2020a,b) or the QLP pipeline (Kunimoto et al., 2021). The SPOC-FFI data consist of up to 160 000 targets per sector and include targets within 100 parsecs, targets with TESS magnitude $\lesssim 13.5$, and targets that are bright in the near-infrared ($H \lesssim 10$ mag). The QLP pipeline extracts light curves for all targets with $T < 13.5$ mag and the data consist of around 900 000 light curves per sector. The provided data products for both the SPOC-FFI and QLP pipelines are the same as for the two-minute cadence SPOC pipeline.

For target stars that have not had light curves extracted with the SPOC-FFI or QLP pipelines, LATTE downloads the FFI Target Pixel Files (TPFs) using the

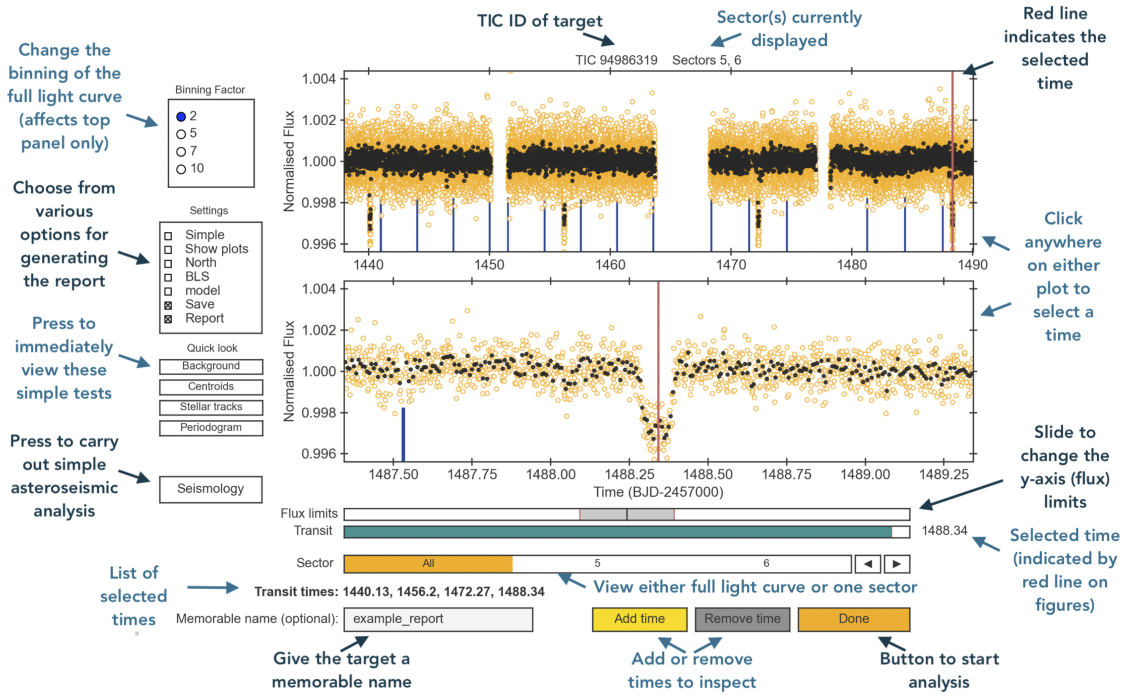

Figure 2.6: User-interface displaying the whole light curve (top panel) and a zoom-in of the data around the selected transit-time (bottom panel) as indicated by the vertical red line. The selected time can be changed by clicking on either one of the two light curves. Times are stored and removed with the 'Add time' and 'Remove time' buttons respectively. Further program options are shown to the left of the light curve plots. The blue lines at the bottom of the two light curves indicate times of *TESS* momentum dumps.

open-source TESSCUT tool (Brasseur et al., 2019). These FFI targets have no pre-identified optimal apertures. The user is, therefore, prompted to manually identify two different sized photometric extraction apertures by interacting with a GUI, which displays the average flux per pixel around the target (see Figure 2.7). The apertures are chosen by clicking on the desired pixels, where the lighter coloured pixels represent higher flux. The two light curves, extracted using the chosen 'large' and 'small' apertures, are simultaneously displayed in the interface.

In order to identify common systematics shared between the pixels, I apply a principal Component Analysis (PCA; Stumpe et al., 2012; Smith et al., 2012) using the SCIKIT-LEARN package (Pedregosa et al., 2011). PCA is a statistical method that identifies and ranks existing sources of variances in a set of input vectors, which in this case are the time-series data of each pixel. Variances that are shared across a large number of pixels are unlikely to be intrinsic to the light curve of the

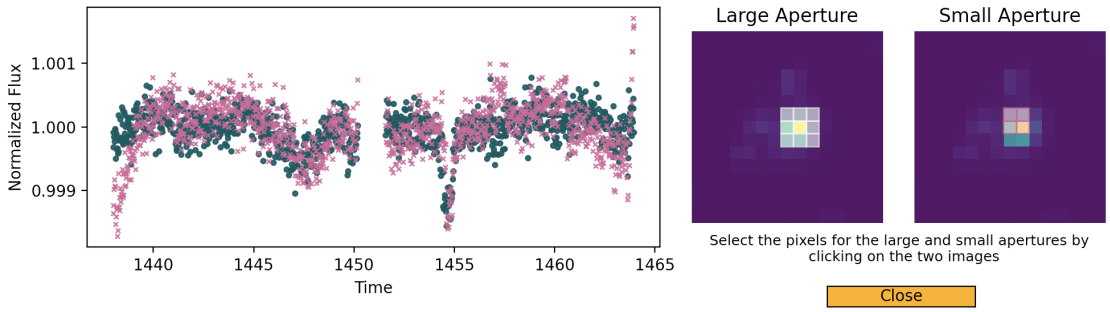

Figure 2.7: The LATTE user interface for FFI targets that were neither processed with SPOC-FFI nor QLP. The user is asked to manually select a small and a large aperture by clicking on the pixels (right panel), where the lighter coloured pixels correspond to a higher flux. The resultant light curves, extracted with the larger (teal green circles) and a smaller (pink crosses) apertures, are displayed on the left.

target. As such, removal of the principal components (highest ranked variances) can reduce long-term variability and stochastic noise due to instrumentation from the data. LATTE removes three principal components, as it was found that the removal of more components increases the computational time without notably improving the detrending of the data.

Following the PCA analysis, the data are detrended using an iterative non-linear filter (Aigrain & Irwin, 2004) to estimate and subtract residual systematics. This detrending removes stellar variations on time-scales that are significantly longer than the duration of a transit (LATTE assumes a default transit duration is 0.7 days, however, users are encouraged to change this if necessary). In brief, the non-linear filter iteratively splits the light curve into segments that are then filtered using a moving median filter with a window width of 3 times the transit duration. The median of the absolute deviations of the residuals (original minus filtered light curves) is used to estimate the scatter (sigma). The original data is then $3\,\sigma$ clipped to remove outliers. The process repeats until no more outliers are found, which typically occurs after less than 10 iterations.

2.2.4 Diagnostic tests

The end-product of LATTE is a pdf document containing a range of standard diagnostic test that are specifically designed to help identify transit-like signals

and weed out instrumental and astrophysical false positives in *TESS* data (see Appendix B for two example reports). In brief the diagnostics consist of:

(i) **Momentum dumps.** The times of the *TESS* reaction wheel momentum dumps displayed alongside the target light curve to ensure that the times of the transit-like events do not coincide with times of increased systematics due to the satellite unloading built up photon pressure. While observations taken during a momentum dump (which typically last around half an hour) are flagged by the SPOC pipeline and removed from further analysis, the satellite pointing remains affected for several hours after each dump, which can result in spurious flux variations due to aperture losses or inter-/intra-pixel sensitivity variations.

(ii) **Centroid positions.** The CCD column and row position of the target's flux-weighted photocenter (*mom_centr*), and the CCD column and row position of the source on the detector (*pos_corr*) around the time of the transit-like event(s). Both of these are provided by the SPOC pipeline. The pipeline determines the latter via motion polynomials, which are calculated using the point spread function centroids of a selection of stars on each pixel. As such, *pos_corr* provides insight into the telescope's stability over time and can be used to rule out events caused by systematics. Large shifts in the *mom_centr*, on the other hand, indicate that the observed change in flux is caused by a neighbouring stars and not the intended target and can, therefore, be used to rule out astrophysical false positives.

(iii) **Nearest neighbour light curves.** Normalised flux light curves of the five short-cadence *TESS* stars with the smallest projected distances to the target star. These are used to identify alternative sources of the signal (e.g., asteroid, background eclipsing binary) or systematic effects that affect the light curves of multiple stars on the detector. As only stars that were observed by *TESS* at a cadence of two minutes are used for this analysis, the five selected targets

are not necessarily the closest stars on the sky. This test is currently not available in *FFI mode*.

(iv) **Pixel level light curves**. Individual light curves extracted for each pixel around the target (typically 11×11 pixels). These are used to identify signals resulting from background eclipsing binaries, background events and systematics. The light curves are extracted from the TPFs, around the time of each flagged transit-like event, and detrended using a third order spline fit.

(v) **Background flux.** The background flux – which is the flux extracted from nearby pixels that do not contain obvious sources – is extracted around the time of the transit-like event(s) to help identify trends caused by background events such as asteroids or fireflies (Vanderspek et al., 2018) passing through the field of view. Enhanced scattered light in the telescope optics can also cause dramatic increases in the background flux when the Earth, Moon or other solar system planets pass below 25° from the boresight of any of the cameras. This effect can induce spurious transit-like events into the light curve that can be identified by inspection of the background flux.

(vi) **Aperture size test.** The target light curve around the time of the transit-like event(s) extracted using two apertures of different sizes. Differences in the transit depth and shape of the two different light curves may be indicative of the transit-like events resulting from a blended eclipsing binary. It should be noted that a sub-optimal mask selection can lead to the mis-interpretation of a moving signal. As such the used apertures are also displayed. The data are extracted from the TPFs and are neither detrended nor corrected for systematics. This allows the user to ensure that the transit-like signal can be seen in the raw light curve and that the event is not a by-product of the detrending process.

(vii) **Pixel-level centroid analysis.** The difference image computed by subtracting the in-transit pixel flux from the image obtained during a similar time

period immediately before and after each transit. This was done using the TPFs after correcting them for systematic effects using PCA. A change in spatial distribution of the flux suggests that the transit-like event is caused by a blend with a background.

(viii) **Nearby companion stars.** The location of nearby stars with a magnitude difference less than 5 compared to the target star, as queried from the *Gaia* Data Release 2 catalogue (Gaia Collaboration et al., 2018), and the Sloan Digital Sky Survey (SDSS-II; York et al., 2000; Abazajian et al., 2009) red field of view around the target star, in order to identify nearby contaminating sources.

(ix) **Box-Least-Squares (BLS) fit**. The BLS algorithm (BLS; Kovács et al., 2002) searches for periodic signals in the *TESS* light curve by phase folding the data at a large number of trial orbital periods, and modelling the transit using an upside-down top hat function with four free parameters: transit duration, depth, period and reference time. For each trial period the BLS search computes a significance of the best fit 'box', allowing it to identify periodic transit-like signals in the data. This test provides the results from two consecutive BLS searches. The data are corrected for residual systematics using an iterative non-linear filter (Aigrain & Irwin, 2004) prior to being searched for periodic signals using the ASTROPY BLS module (Astropy Collaboration et al., 2018). The initial BLS search identifies the times of the highest detected signal-to-noise event and removes these data from the light curve. A second search is then carried out in order to search for additional periodic signals.

(x) **Frequency analysis**. Lomb-Scargle power spectrum (Lomb, 1976; Scargle, 1982) of the *TESS* light curve to search for high and low frequency periodic signals, such as from pulsations or stellar rotation. This analysis is carried out with the publicly available LIGHTKURVE tool (Lightkurve Collaboration et al., 2018).

(xi) **Stellar evolution**. MIST stellar evolutionary tracks (Choi et al., 2016) showing the main sequence evolution and post main sequence evolution for stars with masses ranging from 0.3 to 1.6 $M_{\odot}$. This gives an indication of the evolutionary stage of the host star.

Tests (ii) to (iv) are designed to rule out events caused by systematic effects due to the satellite or instrumentation, and tests (iii) to (viii) to increase one's confidence that the signals are not caused by astrophysical false positives, such as blends where the photometric aperture of a bright target contains a faint eclipsing binary. Tests (ix) to (xi) are not used to rule out false positives, however, they provide additional information on the target star as well as an indication of whether additional transit-like signals are present in the light curve.

2.2.5 LATTE in use

In addition to the tool being used by the exoplanet community to help investigate *TESS* light curves (e.g., Van Grootel et al., 2021; Barragán et al., 2022a), LATTE forms the basis of the *TESS* eclipsing binary working group's signal verification pipeline (see Section 3.3.1). Furthermore, the open-access tool is being used by numerous PHT citizen scientists who regularly post their results on the discussion forums.[5,6] The posting of results from the LATTE reports has motivated other citizen scientists to try the tool, increased the identification of false positive results, and thus sped up the identification of potentially real signals (see Section 7.3).

Finally, LATTE is used to vet planet candidates that pass the initial vetting stage of the PHT pipeline. Out of the $\sim$25 planet candidates per sector for which I generate LATTE reports, around 10 – 15% pass all of the tests and are kept for further investigation, resulting in $\sim$3 promising planet candidates per sector. The discarded candidates can be loosely categorised into (background) eclipsing

[5] https://www.zooniverse.org/projects/nora-dot-eisner/planet-hunters-tess/talk/2112/2172100?comment=3575942&page=1

[6] https://www.zooniverse.org/projects/nora-dot-eisner/planet-hunters-tess/talk/2112/2221854?comment=3758250&page=2

binaries ($\sim 40\%$), systematic effects ($\sim 25\%$), background events ($\sim 15\%$), and other (including stellar signals such as spots; $\sim 10\%$).

2.3 System modelling

The planet candidates that pass all of the light curve based false positive tests, as described above, are modelled in order to extract transit properties and physical system parameters including planet radius, impact parameter, orbital period (or predicted orbital period in the case of single-transit events), semi-major axis, transit depth, and stellar density. These parameters are used to ensure that the best fit model is consistent with a planet scenario (e.g., by ensuring that the radius is consistent with the expected size of a planet). There are numerous available open-access tools that can be used to determine such physical parameters via transit and RV modelling, including PYTRANSIT (Parviainen, 2015), ALLESFITTER (Günther & Daylan, 2021), EXOFAST (Eastman, 2017), EXOPLANET (Foreman-Mackey et al., 2021b), JULIET (Espinoza et al., 2019), PASTIS (Díaz et al., 2014), and `pyaneti` (Barragán et al., 2019; Barragán et al., 2022b).

All of these tools extract physical planet parameters by comparing a parametric model to observational data using Bayes theorem (Bayes, 1763) where the Bayesian approach allows one to make use of prior knowledge about the system for drawing posterior inferences about the system parameters. The remainder of this section outlines how Bayesian inference can be used to model transit and RV data and how this analysis can help inform on whether a candidate is a false positive.

2.3.1 Bayes' theorem

Bayes' theorem provides the framework used to determine the probability that a model $M = M(\Theta)$ with parameters Θ can describe data d:

$$P(M|d) = \frac{P(d|M)P(M)}{P(d)} \tag{2.5}$$

where $P(M|d)$ is the *posterior* probability distribution of the model parameters given the data; $P(d|M)$ is the *likelihood* function describing the probability of

being able to reproduce the data given the model; $P(M)$ is the *prior* probability which characterises any knowledge about the model parameters before the analysis of the data; and $P(d)$ is the probability distribution of the data and acts as a normalising constant.

The likelihood of data point d_i being described by the model predicted value M_i is given by $P(d_i|M_i)$ and the overall likelihood that the whole data set d is described by the model M as:

$$P(d|M) = \prod_{i=1}^{n} P(d_i|M_i). \tag{2.6}$$

Assuming that the individual data points are independent of one another (i.e. not correlated), normally distributed, and contain uncorrelated noise (σ_i), the likelihood function is then expressed as:

$$P(d_i|M_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(d_i - M_i)^2}{2\sigma_i^2}\right). \tag{2.7}$$

Below, I describe the transit and RV models that will function as M in Bayes' theorem throughout this work.

2.3.2 Transit equations

Following the methodology of Mandel & Agol (2002), the fractional decrease in flux over time due to a transiting planet can be expressed as:

$$F(z, r_p; t) = 1 - \lambda(z, r_p; t) \tag{2.8}$$

where $\lambda(z, r_p; t)$ is the fractional loss of flux, $z = \delta/R_\star$ and δ the scaled projected distance between the planet and the star, and $r_p = R_p/R_\star$. Assuming that the star is a uniform source of light, there are three distinct geometrical regimes. First, when there is no occultation ($z \geq 1 + r_p$) there is no decrease in flux and $\lambda(z, r_p; t) = 0$. Second, when there is full occultation ($z \leq 1 - r_p$), there is a constant reduction in flux of $\lambda(z, r_p; t) = r_p^2$. Finally, when there is partial occultation, i.e. during ingress and egress ($1 - r_p < z < 1 + r_p$), the flux changes continuously as described by:

$$\lambda(z, r_p; t) = \frac{1}{\pi}\left[r_p\ \kappa_0 + \kappa_1 - \sqrt{\frac{4z^2 - (1 + z^2 - r_p^2)^2}{3}}\ \right] \tag{2.9}$$

where $\kappa_0 = \cos^{-1}\left[(r_p^2 + z^2 - 1)/(2r_p z)\right]$ and $\kappa_1 = \cos^{-1}\left[1 - (r_p^2 + z^2)/(2z)\right]$. The above equations assume that the star is a uniform disk of light. However, stars vary in brightness from the stellar centre (brightest) to the stellar limb (faintest) due to variations in the atmospheric stellar depths at which the observed photons are emitted. This effect, which is known as *limb darkening*, results in a larger decrease in flux when the planet transits the centre of the star compared to when it transits the edge of the star. In practice, this affects the transit depth as a function of the impact parameter and makes the bottom of the transit appear rounder compared to when a non-limb darkened model is used.

In order to account for the effects of limb darkening, the brightness of the stellar disk can be expressed as $I = I(\mu)$, where $\mu = \cos\theta$ and θ is the angle between the midpoint of the star and a given point on the stellar surface. $I(\mu)$ can take many forms, including the quadratic limb darkening model (Mandel & Agol, 2002), which is defined as:

$$I(\mu) = 1 - u_1(1 - \mu) - u_2(1 - \mu)^2 \tag{2.10}$$

where u_1 and u_2 are limb darkening coefficients that depend on the stellar type (temperature and surface gravity) and the wavelength at which the star was observed. For example, two stars of equal temperature but at different evolutionary phases experience different amounts of limb-darkening, as more evolved stars have more extended and lower density atmospheres.

For a model that includes the effects of limb-darkening, the fractional decrease in flux takes the form $F(z, r_p, u_1, u_2; t)$ and the general parametric function for modelling transits is:

$$M_{tr}(\Theta; t) = f(\{T_0, P, e, w_\star, R_p/R_\star, a/R_\star\}_j, \{u_1, u_2\}_{band}; t) \tag{2.11}$$

where $\{T_0, P, e, w_\star, R_p/R_\star, a/R_\star\}$ is independently defined for each planet j. As presented by Kipping (2013), the limb darkening co-efficients are often parameterised as $q_1 \equiv (u_1 + u_2)^2$ and $q_2 \equiv 0.5u_1(u_1 + u_2)^{-1}$, where q_1 and q_2 are sampled with uniform priors between 0 and 1.

2.3.3 Radial velocity equations

Radial velocity variations (see Section 1.3) can also be used to extract physical system parameters. The RV model is defined as:

$$v_r = v_\gamma + K[\cos(\omega + \nu(T_0, P, e; t))] + e\ \cos\omega_\star] \tag{2.12}$$

where the RV semi-amplitude, K, is defined as in Equation 1.1, P is the orbital period, e is the orbital eccentricity, ν is the true anomaly which is the position of the orbiting body with respect to the argument of pericentre, ω is the argument of pericentre which is the angle between the ascending node and the pericentre (equals 0 for $e = 0$), and v_γ is the systemic velocity which is the velocity of the centre of mass of the system (a schematic of the orbital geometry of a planetary system can be found in Appendix A). As such, the general parametric function for modelling RVs is:

$$M_{RV}(\Theta; t) = f(\{T_0, P, e, \omega, K\}_j, \gamma_{inst,k}; t) \tag{2.13}$$

where $\gamma_{inst,k}$ corrects for offsets due to observations obtained from k different instruments, and $\{T_0, P, e, \omega, K\}$ is independently defined for each planet j.

2.3.4 `pyaneti` transit and RV modelling code

Transit signals found with PHT are modelled using `pyaneti`, an open-access tool that is used for probabilistic Bayesian modelling of photometric and RV time series data, implemented by Markov Chain Monte Carlo (MCMC) sampling. `pyaneti` has already been used to model a large number of exoplanets discovered in *TESS* data (e.g., Gandolfi et al., 2018; Esposito et al., 2019; Carleo et al., 2020; Díaz et al., 2020; Georgieva et al., 2021). The code is described in full in Barragán et al. (2019) and Barragán et al. (2022b).

Multi-transit PHT candidates are fit for the parameters given by Equation 2.11, with the exception of the eccentricity, which is set to 0 (circular orbit) for simplicity. Single-transit events are fit for the same parameters, with the exception of the period and the semi-major axis, which cannot be determined directly from the light curve when only a single transit is observed. As such `pyaneti` assumes a fixed dummy value for the period that is larger than the duration of the time series data, and a fixed dummy value for the scaled semi-major axis. This ensures that the transit shape can be fully sampled in the case of a single-transit event. With the derived model values, the transit duration can then be estimated, which in turn can be used to estimate the orbital period following the methodology of Osborn et al. (2016):

$$P_{\rm circ} = \frac{8\pi^2\,\mathrm{G}}{3}\frac{\rho_\star}{v'^3} \tag{2.14}$$

where G is the gravitational constant, $\rho_\star$ is the stellar density and v' is the scaled transit velocity defined as:

$$v' = \frac{v_p}{R_\star} = \frac{2\sqrt{(1+R_p/R_\star)^2 - b^2}}{T_D} \tag{2.15}$$

where v_p is the velocity of the planet and T_D the transit duration. This model assumes a circular orbit ($e=0$).

Sampling and convergence

The posterior probability distribution of a given set of model parameters are sampled with `pyaneti` using an MCMC routine. In brief, MCMC works by generating an initial random set of model parameters, based on the *prior* information, which are used to calculate the likelihood. Next, each parameter is moved in parameter space, at which point the likelihood is re-evaluated.

An ensemble sampler, as is used by `pyaneti` does this for N independent chains that are simultaneously sampled (by default `pyaneti` uses 100 chains). The proposed direction and amount by which each walker X_k moves depends on a randomly selected other walker X_j, such that:

$$X_k(t) \rightarrow X_k(t+1) = X_j + z(X_k(t) - X_j) \tag{2.16}$$

where z is a scaling parameter sampled from a random distribution (for details see Foreman-Mackey et al., 2013). In order to determine whether the new walker position is accepted or rejected, the algorithm computes the ratio of the likelihoods of the original and proposed locations:

$$q = z^{N-1} \frac{P(M(X_k(t+1))|d)}{P(M(X_k(t))|d)} \tag{2.17}$$

where N is the number of variables. The new step is accepted if $q > r$, where is r is a randomly number between [0,1], and rejected if $q < r$.

It should be noted that the two walkers X_k and X_j are drawn from two different subsets of the ensemble of walkers, such that all walkers in one subset are simultaneously updated based only on the positions of the walkers in the other subset.

This process continues until some convergence criterion is met. `pyaneti` assesses convergence using the methodology described by Gelman & Rubin (1992), whereby the variance of each chain (W, or 'within-chain') is compared to the variance across all of the chains (B, or 'between-chain') using the relationship $\hat{R} = \sqrt{[W(n-1)/n + B/n]/W}$, where n is the length of the chains. `pyaneti` defines convergence as $\hat{R} < 1.02$ (Gelman et al., 2004).

2.3.5 `pyaneti` as a vetting tool

The `pyaneti` transit model is used to infer the planetary and orbital parameters for all multi- and single-transit PHT candidates that pass the `LATTE` vetting stage. The resulting system parameters are used to determine whether the model is consistent with the planet scenario. For example, signals consistent with a transiting companion with radius larger than 2 R_{Jup} are, in general, assumed to be incompatible with the planet scenario.

Additionally, certain parameters sampled by the `pyaneti` transit model can be used to determine the mean stellar density (Barragán et al., 2019) using:

$$\rho_\star \approx \frac{3\pi}{\mathrm{G}} \frac{1}{P^2} \left(\frac{a}{R_\star}^3 \right). \tag{2.18}$$

This should yield a value of the stellar density that is consistent with the stellar density derived from the mass and radius of the star (Winn, 2010). As such, the comparison of these two independently derived stellar densities provides a means to identify false positives.

PHT planet candidates that pass the `pyaneti` vetting stage are uploaded to the Exoplanet Follow-up Observing Program (ExoFOP) for TESS.[7] In brief, ExoFOP is a database that lists all of the targets in the TESS Input Catalog (TIC), as well as any confirmed planets, planet candidates, and ongoing follow-up efforts. The PHT candidates are uploaded as community TOIs (cTOI), where cTOIs are planet candidates uploaded to ExoFOP by the community to be reviewed by the TESS Science Office vetting team. All PHT candidates that are uploaded to ExoFOP are also considered for ground-based follow-up observations.

2.4 Follow-up observations

Many astrophysical false positive scenarios can be ruled out from the detailed examination of the *TESS* data, both from the light curves themselves (including from transit modelling) and from the target pixel files. However, not all false positive scenarios can be ruled out from these data alone, due in part to the large *TESS* pixels (21 arcsec). The final stage of vetting, therefore, consists of following up the candidates with ground-based observations including photometry, reconnaissance spectroscopy, and speckle imaging.

PHT identified candidates are followed-up based on three main criteria: 1) they pass all of the aforementioned vetting tests, 2) they are bright enough to produce a high enough S/N to perform scientific analysis (for RV measurement this is typically constrained to $V < 12$ mag), and 3) they are observable from the facilities that our team has access to (see below). Priority is given to targets of specific scientific

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

interest, including young systems, evolved systems, and multi-planet systems. The remainder of this section outlines the observing facilities that are being used to follow-up a sample of PHT candidates.

2.4.1 Photometry

I made use of the Las Cumbres Observatory (LCO) global network of fully robotic 0.4-m/SBIG and 1.0-m/Sinistro facilities (Brown et al., 2013) to observe additional transits (only if the orbital period is known due to, for example, multiple transit events in the *TESS* light curve), in order to refine the ephemeris and confirm that the transit events are not due to a blended eclipsing binary in the vicinity of the main target. Snapshot images were taken of single-transit event candidates in order to identify nearby contaminating sources. The images were calibrated with the standard LCO BANZAI pipeline (McCully et al., 2018). In total the PHT project has been awarded 270 hours on these instuments as part of proposals SUPA2019B-008, SUPA2020A-003, and SUPA2021A-004.

2.4.2 Spectroscopy

We perform high-resolution optical spectroscopy using telescopes from across the globe in order to cover a wide range of RA and Dec:

- The LCO telescopes with the Network of Robotic Echelle Spectrographs (NRES; Brown et al., 2013). These fibre-fed spectrographs, mounted on 1.0-m telescopes around the globe, have a resolution of $R = 53\,000$ and a wavelength coverage of 380 to 860 nm. The data are automatically reduced with the LCO NRES pipeline, which includes simple bias and dark field corrections, extraction of the 1D spectrum, cosmic-ray rejection, and wavelength calibration with Thorium-Argon (ThAr) lamps. In total the PHT project has been awarded 172 hours as part of proposals SUPA2019B-008, SUPA2020A-003, and SUPA2021A-004.

- The MINiature Exoplanet Radial Velocity Array (MINERVA-Australis) Telescope facility, located at Mount Kent Observatory in Queensland, Australia (Addison et al., 2019). This facility is made up of four 0.7-m CDK700 telescopes, which individually feed light via optic fibre into a KiwiSpec high-resolution ($R = 80\,000$) stabilised spectrograph (Barnes et al., 2012) that covers wavelengths from 480 nm to 620 nm. The data are automatically reduced and wavelength calibration with simultaneous ThAr observations. RVs are extracted using the least-squares matching technique of Anglada-Escudé & Butler (2012). In total the PHT project has been awarded more than 100 hours via collaborations with the MINERVA consortium.

- The CHIRON spectrograph mounted on the SMARTS 1.5-m telescope (Tokovinin, 2018), located at the Cerro Tololo Inter-American Observatory (CTIO) in Chile. The high resolution cross-dispersed echelle spectrometer has a 2.7 arcsec diameter multi-mode optical fibre. Using the high efficiency image slicer mode results in a resolution of $R = 80\,000$ and covers wavelengths ranging from 410 to 870 nm. The observations are wavelength calibrated by the CHIRON pipeline and corrected for night-drift using ThAr spectra, which are taken before and after each science exposure. Corrections for the motion of the Earth about the barycentre of the Solar system are calculated using Barycorrpy (Kanodia & Wright, 2018). Following this, RVs are extracted using the SERVAL pipeline (Zechmeister et al., 2018) that we specifically modified to analyse CHIRON data.[8] In brief, SERVAL combines all of the available spectra of a target into a high S/N template, which is then used to compute RV variations using a least-squares fitting. Figure 2.8 shows an example RV curve obtained with the CHIRON spectrograph. In total the PHT project has been awarded 168 hours as part of proposals NOAO 2019B-0232, NOAO 2020A-0140, NOAO 2021A-0109, and NOAO 2021B-0086.

[8]The modification of the SERVAL code was carried out in collaboration with citizen scientist Ian Mason.

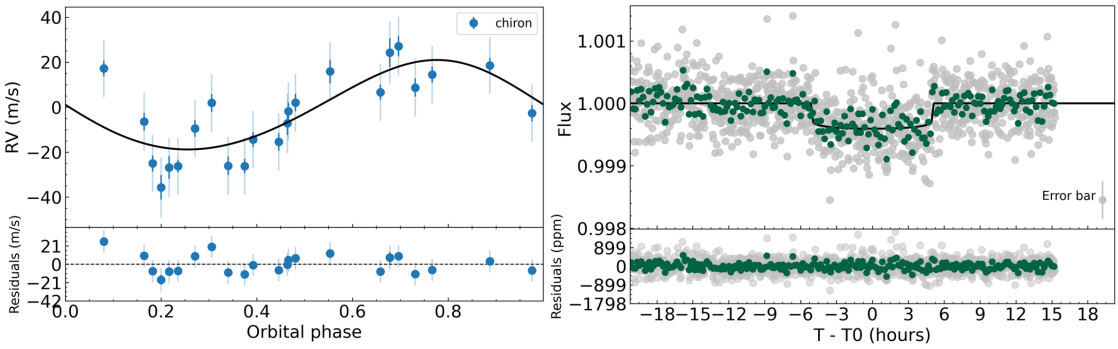

Figure 2.8: Example RV data obtained with the CHIRON spectrograph for TIC 98794593 (left panel) with the corresponding *TESS* light curve (right panel). The 2-minute cadence *TESS* data are shown in grey and the data binned to a 10-minute cadence are shown in green. The black lines in both panels show the best fit models derived from a joint `pyaneti` RV and transit fit (see Section 2.3). This joint modelling revealed an orbital period of 56.9 ± 2 days. Observations of this target are ongoing.

- The SOPHIE echelle spectrograph mounted on the 1.93-m Haute-Provence Observatory (OHP), France (Perruchot et al., 2008; Bouchy et al., 2009). The high resolution cross-dispersed stabilised echelle spectrometer is fed by two optical fibres. Observations were taken in high-resolution mode ($R = 75\,000$) with a wavelength range of 387 to 694 nm. The SOPHIE pipeline is used to calibrate the spectra using ThAr observations obtained on the same night (Bouchy et al., 2009), and the are RVs extracted by correlating the wavelength-calibrated spectra with a template appropriate for the spectral type of the target star (Baranne et al., 1996; Pepe et al., 2002). In total the PHT project was awarded three nights as part of proposal SUPA2020A-003.

Reconnaissance spectroscopy with these instruments allows me to extract stellar parameters, identify spectroscopic binaries, and place upper limits on the masses of stellar or planetary companions. Spectroscopic binaries and targets whose spectral type is incompatible with the initial planet hypothesis or which preclude precision RV observations (giant or early type stars) are not followed up further. Promising targets, however, are monitored in order to constrain their periods and place limits on their masses.

2.4.3 Speckle imaging

The most promising planet candidates are observed using high resolution speckle imaging with the 'Alopeke instrument on the 8.1-m Frederick C. Gillett Gemini North telescope in Maunakea, Hawai'i, USA, and its twin, Zorro, on the 8.1-m Gemini South telescope on Cerro Pachón, Chile (Howell et al., 2011; Matson et al., 2019). Speckle interferometric observations provide extremely high resolution images reaching the diffraction limit of the telescope. Simultaneous 562 nm and 832 nm rapid exposure (60 milliseconds) images were obtained in succession that effectively 'freeze out' atmospheric turbulence and through Fourier analysis are used to search for close companion stars at $5-8$ magnitude contrast levels. This analysis, along with the reconstructed images, allows me to identify nearby companions and to quantify their light contribution to the *TESS* aperture and thus determine whether they could produce a false positive signal (see Chapter 3). In total the PHT project was awarded 1.8 nights as part of proposals GS-2019A-DD-109 (DDT) and GS-2019B-FT-103.

2.5 Conclusions

PHT is one of the largest citizen science projects in the world and engages a large global community in the exciting task of identifying new planet candidates. Combined, the PHT project, the `LATTE` vetting suite, the `pyaneti` transit modelling code, and our ground-based follow-up campaign, allow for the detection of new transit-like signals.

The general workflow, which is illustrated in Figure 2.9, starts with approximately 20 000 two-minute cadence *TESS* targets per month that are classified by citizen scientists via PHT. Based on the user classifications I rank all of the candidates according to their likelihood of containing a transit event, and visually inspect the 500 highest ranked candidates in order to rule out obvious false positives. After this initial round of vetting, around 25 targets per sector are kept for further analysis, for which I generate a `LATTE` validation report. These reports are inspected by the PHT science team in order to identify planetary candidates that warrant further

Figure 2.9: The PHT workflow with approximate numbers displayed for the analysis of a single sector.

investigation. Around 10% – 15% of the targets assessed at this stage of vetting are kept for further investigation, resulting in ∼ 3 promising planet candidates per sector. These promising candidates are modelled using the `pyaneti` suite in order to extract physical system properties and to ensure that the determined values are consistent with a planet scenario. Candidates that pass all of the vetting tests are uploaded to the ExoFOP website[9] as cTOIs. Out of these, the most promising candidates are

[9] https://exofop.ipac.caltech.edu/tess/index.php

followed-up using ground-based facilities via spectroscopy, photometry, and speckle imaging in order to further rule out false positives and to constrain the planetary nature of the candidates. The properties of the planet candidates that pass all of the vetting stages outlined throughout this chapter will be discussed in Chapter 3.

Overall, the PHT project, combined with the open-access tools and knowledge provided online,[10] is making science more accessible to the general-public (see Chapter 7) and thus promotes scientific research to a wider audience. In turn, I hope that this will help make science a more diverse environment in the long-run.

[10]Planet Hunters Coffee Chat: www.planethunters.coffee

Don't let the Muggles get you down.

— R. B. Weasley, 1993

CHAPTER 3

ANALYSIS AND RESULTS

Planet Hunters TESS was designed to find rare transit events in *TESS* data that standard detection pipelines either miss or ignore. However, these results are of limited use without an indication of the completeness of the search. In this chapter I, therefore, discuss project completeness and the recovery efficiency of the citizen science approach, based on the data from the first three years (Sectors 1 – 39) of the *TESS* mission.[1] I also present the most promising planet candidates identified via the PHT project and discuss the value of citizen science in identifying various types of stellar systems.

3.1 Recovery efficiency

3.1.1 Recovery of simulated transits

The transit recovery efficiency of PHT is, in part, assessed by analysing the recovery rate of simulated transit-like signals injected into real *TESS* light curves (see Section 2.1.2). Figure 3.1 shows the median and mean of the transit scores (fraction of volunteers who correctly identified a given transit scaled by user weights and normalised between 0 and 1) of the simulated transits within S/N bins ranging from 4 to 20 in steps of 0.5. Simulations with a S/N less than 4 are not shown on PHT. As shown by the figure, transit signals with a S/N of 7.5 or greater are correctly

[1] Eisner et al. (2021a) covered a similar analysis for the first two years of the *TESS* mission (Sectors 1 – 26) only.

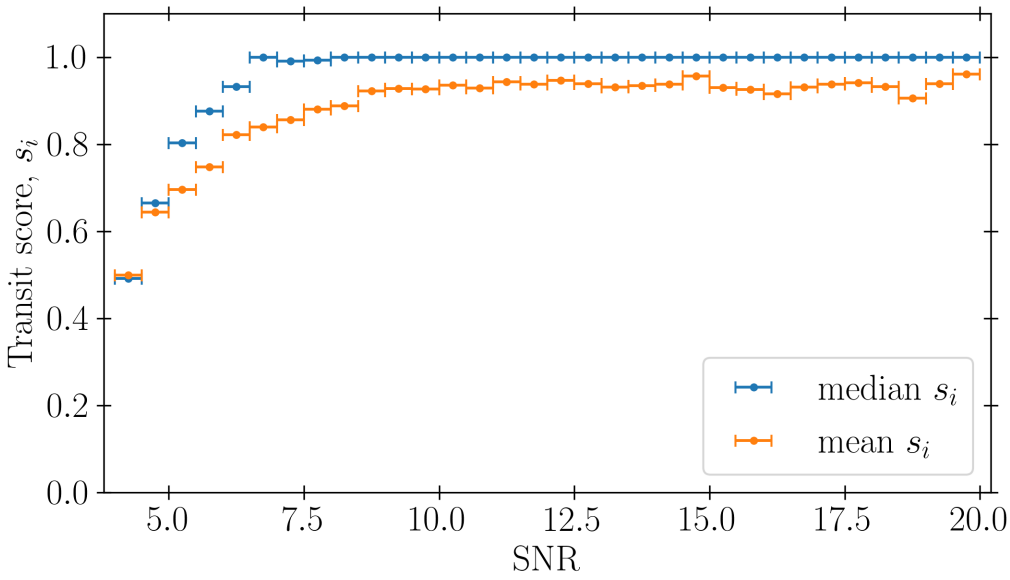

Figure 3.1: The median (blue) and mean (orange) transit scores of injected transits with S/N ranges between 4 and 20 (Sectors 1 – 39). The mean and median are calculated in S/N bins with a width of 0.5, as indicated by the horizontal lines around each data point.

identified by the vast majority of volunteers, as indicated by the median and mean transit scores that plateau at 1 and ~ 0.9, respectively, for signals with S/N > 7.5. The standard deviations of the scores in each bin are around 0.2.

The decreasing median and mean transit scores for signals with S/N < 7.5 shows that the recovery of these transits is more sporadic i.e. while some volunteers are able to identify these lower S/N signals, others are not. As each light curve is seen by up to 15 randomly selected volunteers, a light curve with a low S/N signal might not be seen by any users who regularly identify low S/N transit events. I note that as the simulated data solely consist of real light curves with synthetically injected transit signals, the simulated data cannot be used to infer a false-positive rate.

3.1.2 Recovery of TCEs and TOIs

I further assessed the recovery efficiency of PHT using the planet candidates identified by the SPOC pipeline (Jenkins et al., 2018). As discussed in Section 1.5.3, the SPOC pipeline generates Data Validation (DV) reports for all Threshold Crossing Events (TCEs), which are targets flagged by the SPOC pipeline as containing two or more transit-like features. For each sector, the pipeline identifies around 1025 TCEs,

around 130 of which are added to the catalogue of *TESS* Objects of Interest (TOIs). All of the TCEs and TOIs are classified on PHT in the exact same manner as the rest of the data and can, therefore, be used to assess the project's detection efficiency.

TOI and TCE recovery vs rank

Figure 3.2 shows the fraction of TOIs and TCEs (top and bottom panel respectively) that were recovered with PHT as a function of their rank, where a higher rank corresponds to a lower transit score (see Section 2.1.6 for details on how the transit scores are determined), for Sectors 1 through 39. TOIs and TCEs with $R < 2$ $R_\oplus$ are not included in this analysis, as my previous study of the first two years of *TESS* data showed that human vetting alone is unable to reliably recover planets smaller than 2 $R_\oplus$ due to the low S/N of these signals (Eisner et al., 2021a). This was also shown by the initial PH project (Schwamb et al., 2012). Planets smaller than 2 $R_\oplus$ are, therefore, not the main focus of this work.

Figure 3.2 shows a steep increase in the fractional TOI recovery rate up to a rank of $\sim$500. Within the 500 highest ranked PHT candidates for a given sector we are able to recover between 41% and 60% (median of 55%) of all of the TOIs ($R > 2$ $R_\oplus$); a median 90% of the TOIs where the S/N of the transit events are greater than 7.5; and a median of 88% of the TOIs where the S/N of the transit events are greater than 5. This shows, unsurprisingly, that visual vetting is very sensitive to the depth of the signal.

The relation between planet recovery rate and the S/N of the transit events is further highlighted in Figure 3.3, which shows the S/N vs the orbital period of the recovered TOIs. The colour of the markers indicate the TOI's rank within a given sector, with the lighter colours representing a lower rank. The circles and crosses represent candidates at a rank lower and higher than 500, respectively. Similarly, the two histograms show the candidates recovered at a rank lower than 500 (light green) and higher than 500 (purple) for the S/N and period on the x and y axes, respectively. The figure shows that transit events with a S/N less than $\sim$7.5 (indicated by the horizontal grey line) are missed by the majority of

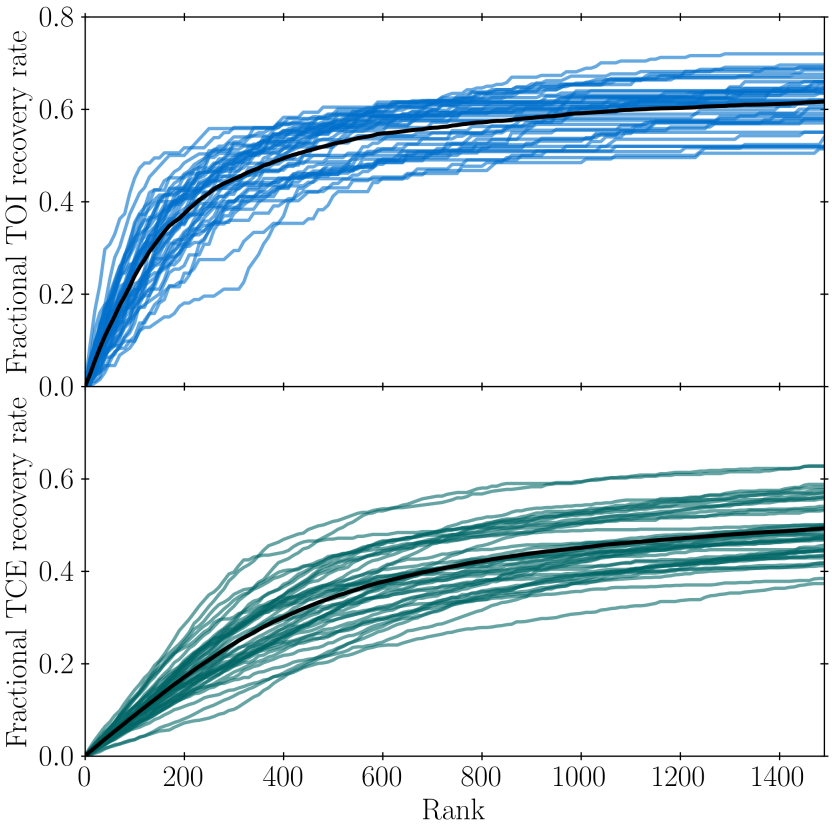

Figure 3.2: The fraction of recovered TOIs and TCEs (top and bottom panel, respectively) with $R > 2\ R_\oplus$ as a function of the rank, for Sectors 1 through 39. The lines represent the results from different observation sectors. The solid black lines indicate the median recovery for a given rank across all of the sectors.

volunteers, whereas events with a S/N greater than 7.5 are mostly recovered within the top 500 highest ranked candidates.

The steep increase in the fractional TOI recovery rate at lower ranks, as shown in Figure 3.2, is therefore due to the detection of the high S/N candidates that are identified by most, if not all, of the PHT volunteers who classified those targets. However, at a rank of around 500, the S/N of the TOIs tends towards the limit of what human vetting can detect and thus the identification of TOIs beyond a rank of 500 is more sporadic.

The fractional TCE recovery rate (bottom panel of Figure 3.2) is systematically lower than that of the TOIs. There are qualitative reasons as to why humans might not identify a TCE as opposed to a TOI, including that TCEs may be caused by artefacts or periodic stellar signals that the SPOC pipeline identified as a potential

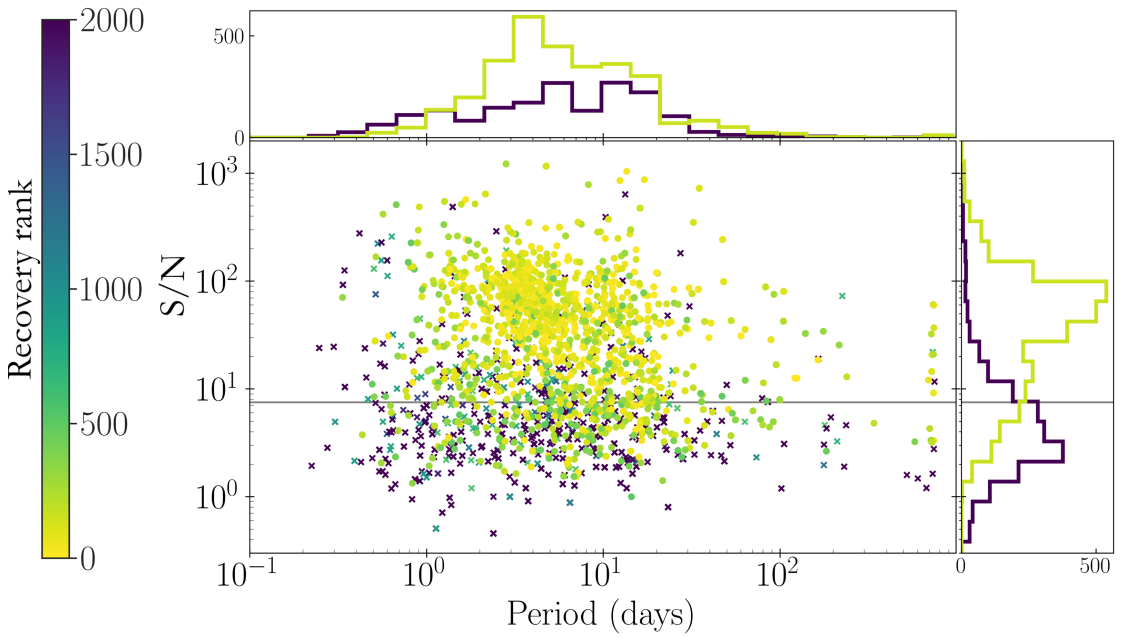

Figure 3.3: The S/N vs orbital period of TOIs with $R > 2\ R_\oplus$. The colour represents their rank within the sector, as determined by the weighted DB clustering algorithm. Circles indicate that they were identified at a rank < 500, while crosses indicate that they were not within the top 500 highest ranked candidates of a given sector. The histograms show the frequency of the candidates recovered at a rank lower and higher than 500 in light green and purple, respectively. The figure shows that transit events with S/N greater than 7.5 (indicated by the grey horizontal line) are recovered by the majority of volunteers, while the recovery of lower S/N events is more sporadic.

transit but that the human eye would either miss or be able to rule out as systematic effect. Overall, this leads to a comparatively lower recovery fraction of TCEs, an effect that is further amplified by the much larger number of TCEs compared to TOIs.

TOI recovery as a function of period and radius

I also estimated the detection efficiency of PHT using the fractional recovery rate of TOIs for a range of radius and period bins, as shown in Figure 3.4. A TOI is considered to be recovered if its detection rank is lower than 500 within the given sector. Out of the total of 4982 TOIs identified following the public release of Sector 39, PHT recovered 1382 TOIs among the highest ranked candidates across the 39 sectors. Independent of the orbital period, PHT is over 85% complete in the recovery of TOIs with radii greater than 4 $R_\oplus$. This agrees with the findings from the initial PH project (Schwamb et al., 2012), who found that they were

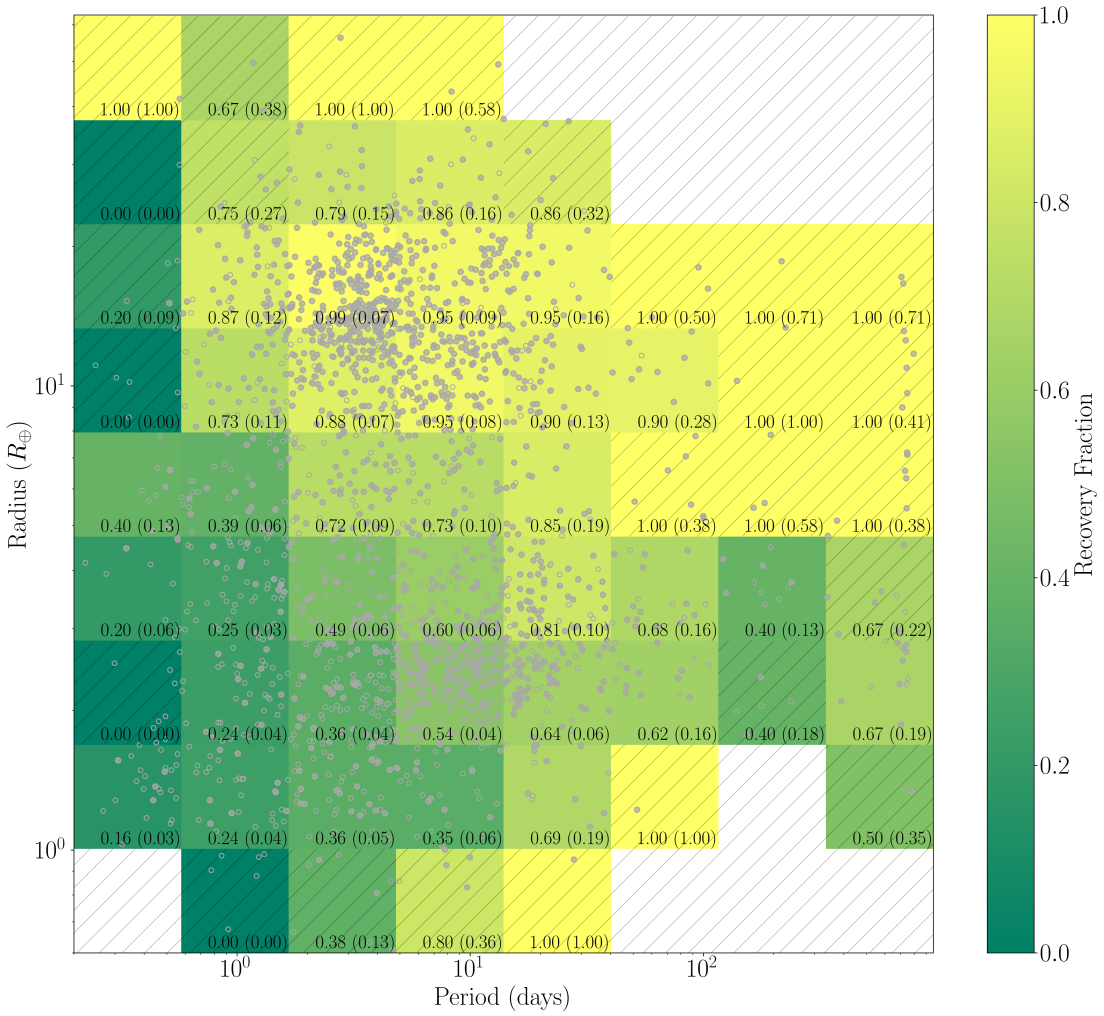

Figure 3.4: TOI recovery rate as a function of planet radius and orbital period. A TOI is considered recovered if it is amongst the top 500 highest ranked candidates within a given sector. The logarithmically spaced grid ranges from 0.2 to 225 days and 0.6 to 55 $R_\oplus$ for the orbital period and planet radius, respectively. The fraction of TOIs recovered using PHT is computed for each cell and represented by the colour the grid. Cells with less than 10 TOIs are considered incomplete for statistical analysis and are shown by the hatched lines. White cells contain no TOIs. The annotations for each cell indicate the number of recovered TOIs followed by the Poisson uncertainty in brackets. The filled in and empty grey circles indicated the recovered and not-recovered TOIs, respectively.

over 85% efficient at detecting short-period (<15 days) planets with radii >4 $R_\oplus$ in the *Kepler* data. The PHT detection efficiency decreases to 51% for $3-4$ $R_\oplus$ TOIs, 49% for $2-3$ $R_\oplus$ TOIs and to less than 40% for TOIs with radii less than 2 $R_\oplus$. Figure 3.4 shows that the orbital period does not have a strong effect on the detection efficiency for periods greater than ~ 1 day, which highlights that human vetting efficiency is independent of the number of transits present within a light

curve. Importantly, this shows that visual vetting is able to yield longer-period planets that may just transit a single time within the data.

For periods shorter than around 1 day, the detection efficiency decreases (independent of the planet radius), due to the high frequency of events seen in the light curve. For these light curves, many volunteers will only mark a subset of the transits, which may not overlap with the subset marked by other volunteers. Due to the methodology used to identify and rank the candidates, as described in Section 2.1.6, this will actively disfavour the recovery of very short-period planets. Although this introduces biases in the detectability of very short-period signals, the major detection pipelines are specifically designed to identify these types of planets and thus this does not present a serious detriment to my main science goal of finding planets that were intentionally ignored or missed by the main automated pipelines.

Estimated false positive rate using TOIs

As there are no light curves, simulated or real, which I can guarantee do not contain any planetary transits (real or injected), it is challenging to determine the false positive rate of PHT. However, I can establish an upper bound by making the assumption that all light curves that have not been flagged as TOIs do not contain a transit event. Even though this assumption is clearly incorrect, the number of unidentified planets is small compared the number of TOIs. As such, to first order this assumption can be used to estimate a false positive rate.

Figure 3.5 displays the fraction of candidates that are not TOIs (estimated false positive rate) as a function of the rank for Sectors 1 through 39. It shows that there are large fluctuations in the estimated false positive rate up to a rank of around 100, beyond which the fraction of non-TOIs increases more steadily and tends towards a 100% false positive rate. This illustrated that in order to obtain a TOI recovery rate of around 50%, which is typically reached at a rank of around 500 (Figure 3.2), the false positive rate is between 60% and 90%. This high false positive rate highlights an important challenge associated with the the PHT pipeline and demonstrates the need for the second round of visual vetting beyond the citizen science classifications in order to rule out the high volume of flagged signals that are false positives.

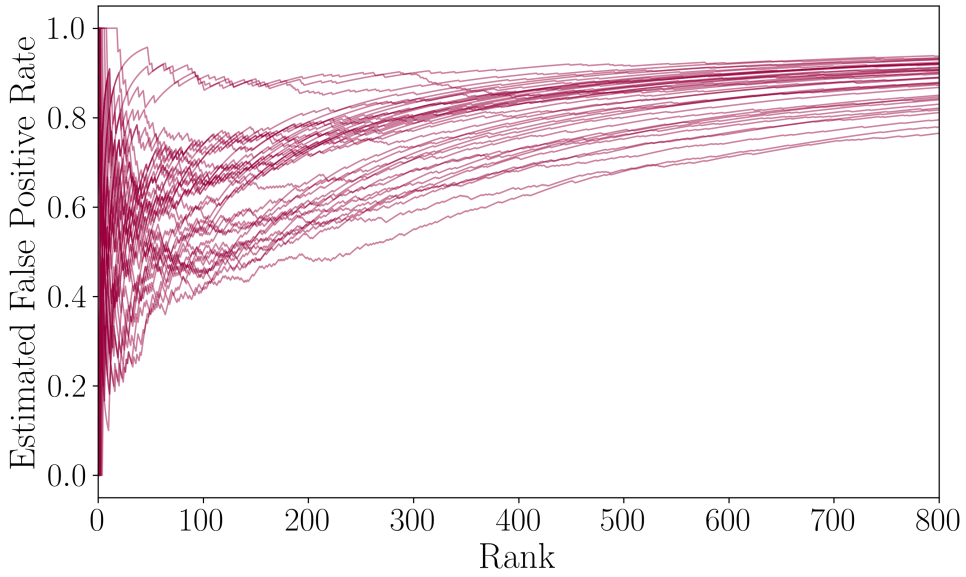

Figure 3.5: The fraction of recovered candidates that are not TOIs vs the rank for Sectors 1 through 39. By assuming that all light curves that are not TOIs contain no planets, we can show how the rate of false positives changes as a function of the rank.

TOI and TCE recovery over time

I assessed whether the detection efficiency varies over time by assessing the fraction of recovered TOIs and TCEs within the highest ranked 500 candidates for each sector. Figure 3.6 shows that there is a slight increase in the fraction of recovered TOIs and TCEs towards the later sectors.

I used a Pearson correlation test to measure the strength of a linear correlation between the fractional recovery rate at a rank of 500 and the sector number. In brief, the Pearson correlation coefficient ($r_{x,y}$) is defined as the ratio between the covariance of the two data sets x (recovery rank) and y (sector) and the product of their standard deviations:

$$r_{x,y} = \frac{\sum(x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum(x_i - \bar{x}_i)^2(y_i - \bar{y}_i)^2}} \tag{3.1}$$

where $r = 1$ for a positive correlation, $r = 0$ for no correlation and $r =$ -1 for a negative correlation. The test showed that there is a moderate positive correlation of 0.57 (p-value = 0.0001) and 0.42 (p-value = 0.006) between the observation sector and the TCE and TOI recovery rates, respectively. As the p-values are less than 0.05, I can reject the null hypothesis of there being no correlation.

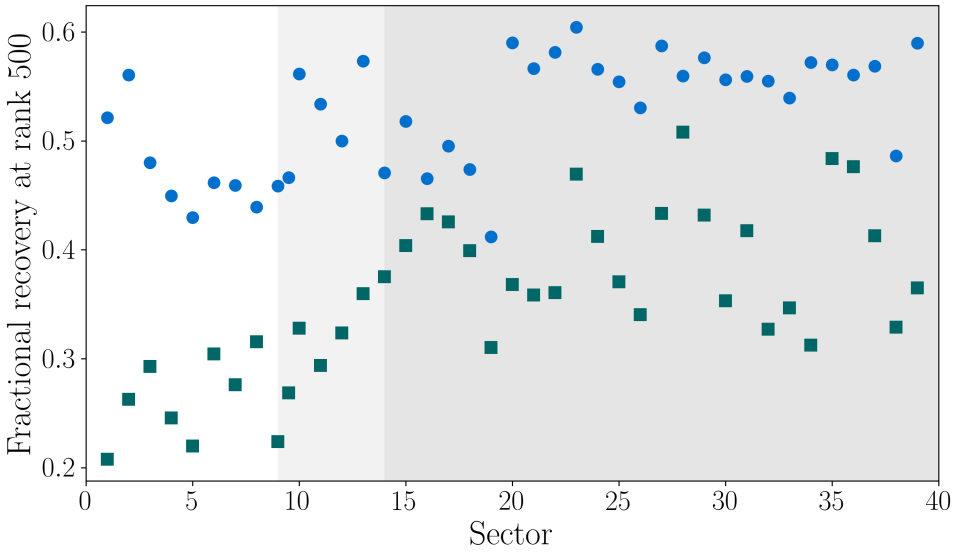

Figure 3.6: The fractional recovery rate of the TOIs (blue circles) and TCEs (teal squares) at a rank of 500 for each sector. Sectors 1 – 9 (white background) represent southern hemisphere sectors classified with UI1, Sectors 9 – 14 (light grey background) show the southern hemisphere sectors classified with UI2, and Sectors 14 – 39 (dark grey background) show the northern hemisphere sectors classified with UI2.

Assuming that volunteers classify data across multiple sectors, these correlations could suggest that the ability of users to detect transit-like events improves over time as they classify more subjects. However, it should also be noted that the pointing stability of the *TESS* spacecraft improved over time, and in particular over the course of the first year of the mission (Sectors 1 – 13). This resulted in improved data quality and a decreased number of systematics in the later sectors. As such, the higher fractional TCE and TOI recovery rates in the later sectors could be a result of the overall improved *TESS* data quality. This is supported by the observation that when only the later sectors (Sectors 20 – 39) are considered in the Pearson correlation test, there is no correlation between the fractional recovery rate of TOIs or TCEs as a function of the sector number.

The change in the skill of the users over time can also be assessed with Figure 3.7, which shows the median (unnormalised) user weight (see Section 2.1.4) per sector for volunteers who completed one or more classifications in at least one sector (blue), more than 20 sectors (green), and all of the 39 sectors (pink). The shaded regions show the 25th and 75th percentiles of the user weights for each sector

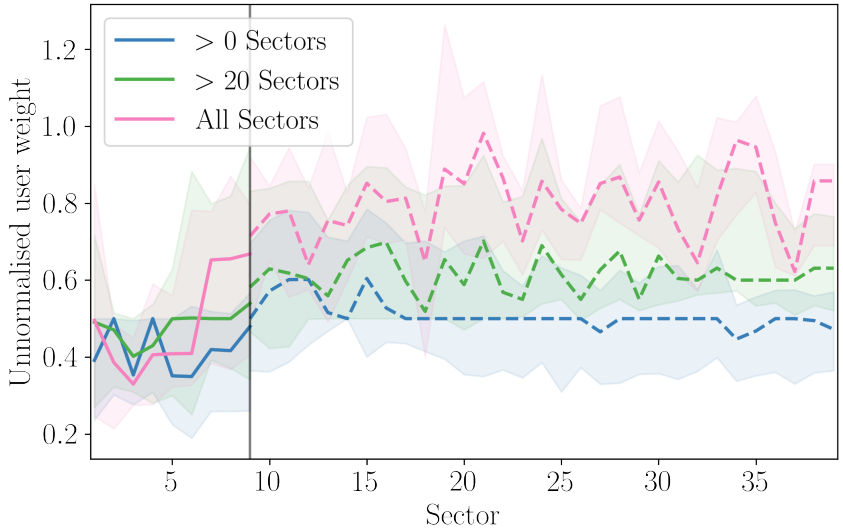

Figure 3.7: Median user weights per sector. The solid lines show the user weights for UI1 and the dashed line for UI2, separated by the black line (Sector 9). The different coloured lines show the median user weights calculated considering user who participated in any number of sectors (blue), more than 20 sectors (green), and all first 39 sectors (pink). The shaded regions show the 25th and 75th percentiles of the user weights per sector.

to give an indication of the scatter in the user weights. The figure highlights an overall improvement in the median user weight following Sector 9, where the user weights flatten out from Sector 9 onward following the switch from UI1 to UI2 (see Section 2.1.1). This suggests that the improvement in the user weights is due to the changes in the interface rather than a continued increase in skill over time. Even though the user weights do not generally increase from sector to sector, the average user weights of volunteers who have completed at least one classification in all of the 29 sectors is higher than those who have classified in more than 20 sectors, which in turn is higher than those who have classified in at least one sector. This shows that users who contribute to PHT on a longer-term basis tend to have higher user-weights, which is promising for the longevity of the project.

Having discussed the detection efficiency of visual vetting via citizen science, I now move on to outline the sample of planet candidates that were identified using PHT but that were missed by other automated searches.

3.2 Planet candidates

In this section I discuss the 139 PHT planet candidates around 135 host stars that passed the initial two stages of vetting and that were uploaded to ExoFOP as cTOIs. At their respective times of discovery, none of these targets had been identified as a TOI (or cTOI uploaded by other teams). Out of the PHT candidates identified in the first three years of *TESS* data, I have confirmed and published three planets in two systems, which will be discussed in detail in Chapters 4 and 5.

The properties of all of the PHT candidates are summarised in Table 3.1. Out of the 139 PHT candidates, 78 ($\sim$56%) are single-transit candidates, indicated by an 's' following the orbital period presented in the table, and 17 exhibit multiple transits separated by large ($\geqslant$100 days) data gaps. A total of 52 of the PHT candidates were flagged as TCEs by the *TESS* pipeline, but not initially promoted to TOI status. The most common reason for this was that the pipeline identified a single-transit event as well as times of systematics (often caused by momentum dumps), due to its two-transit minimum detection threshold. This resulted in the candidate being discarded on the basis of it not passing the 'odd-even' transit depth test. Out of the 139 PHT cTOIs, 64 have become TOIs since the PHT discovery. Candidates that have been promoted to TOI status (by the *TESS* science team) since their PHT discovery are highlighted with an asterisk following the TIC ID, and candidates that we showed to be false positives, based on the ground-based follow-up observations (see Section 4.3) or additional *TESS* data, are marked with a dagger symbol (†; 7 candidates).

Table 3.1 also summarises the main inferred system parameters (columns 2 to 8) that were derived from the `pyaneti` modelling as described in Section 2.3. The values and their uncertainties are given by the median and 68.3% credible interval of the posterior distributions. Candidates with transits separated by large data gaps were modelled with wide priors on the orbital periods. I note that some of the candidates are V-shaped, consistent with a grazing transit configuration. For these cases, uniform priors between 0 and 0.15 were used for $r_{\rm p}/R_\star$ and between 0 and 1.15

for the impact parameter (b) in order to avoid large radii caused by the $r_{\rm p}/R_\star - b$ degeneracy. Thus, the $r_{\rm p}/R_\star$ for these candidates should not be trusted. A full characterisation of these grazing transits is beyond the scope of this work. Figure 3.8 shows the *TESS* transits together with the inferred model for each candidate.

Finally, the table summarises the ground-based follow-up observations (see Section 2.4) that have been obtained to date by the PHT team, where the bracketed numbers following the observing instruments indicate the number of epochs that have been obtained to date. Unless otherwise noted, the follow-up observations are consistent with a planetary scenario. PHT candidates that were later promoted to TOI status are regularly followed-up by the various working groups within the Exoplanet Follow-up Observing Program for TESS and reported on the ExoFOP website. I do not report on these observations here.

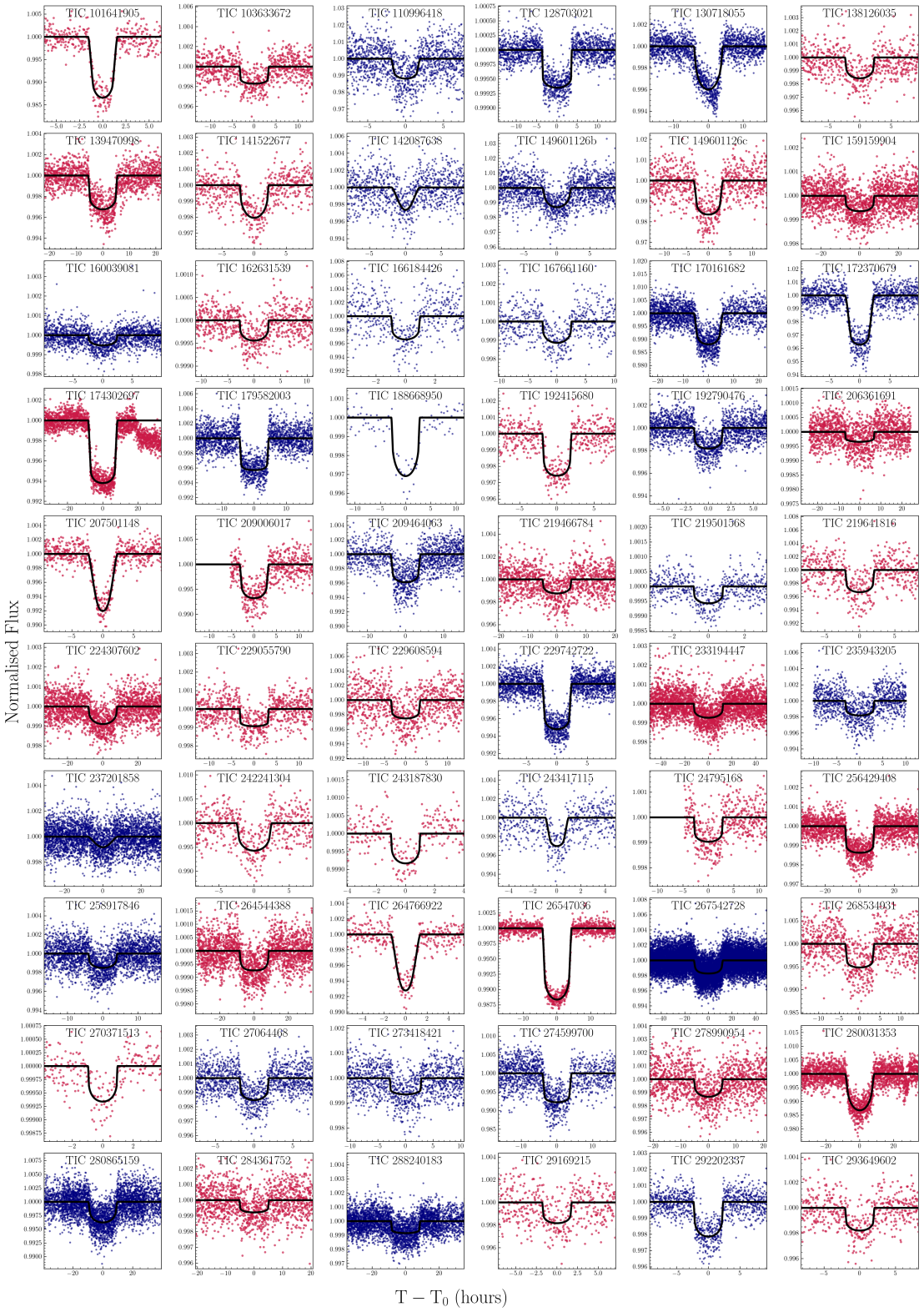

Figure 3.8: All of the PHT candidates modelled using `pyaneti`. The parameters of the best fits are summarised in Table 3.1. The blue and magenta fits show the multi- and single-transit event candidates, respectively.

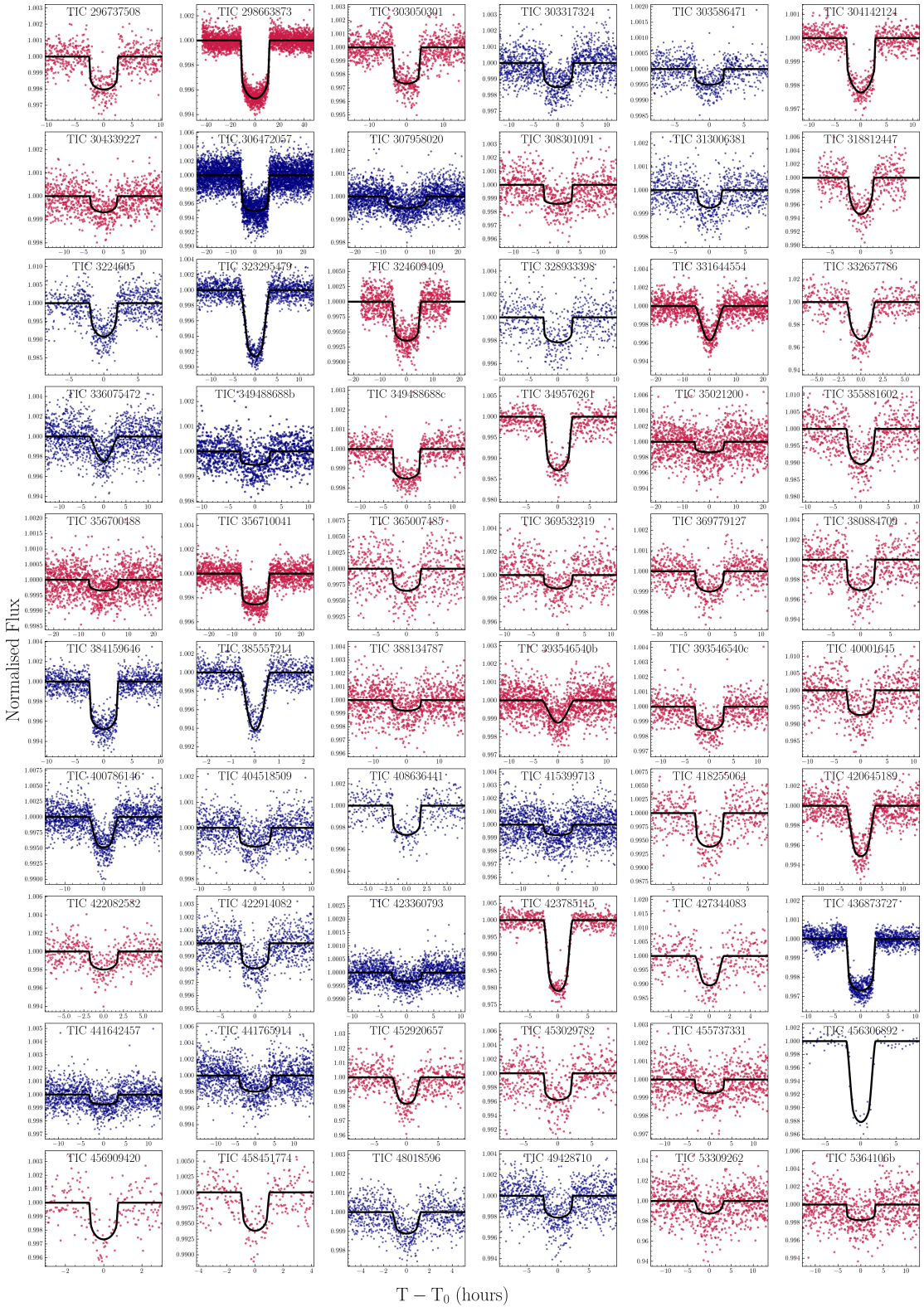

Figure 3.8: PHT candidates (continued)

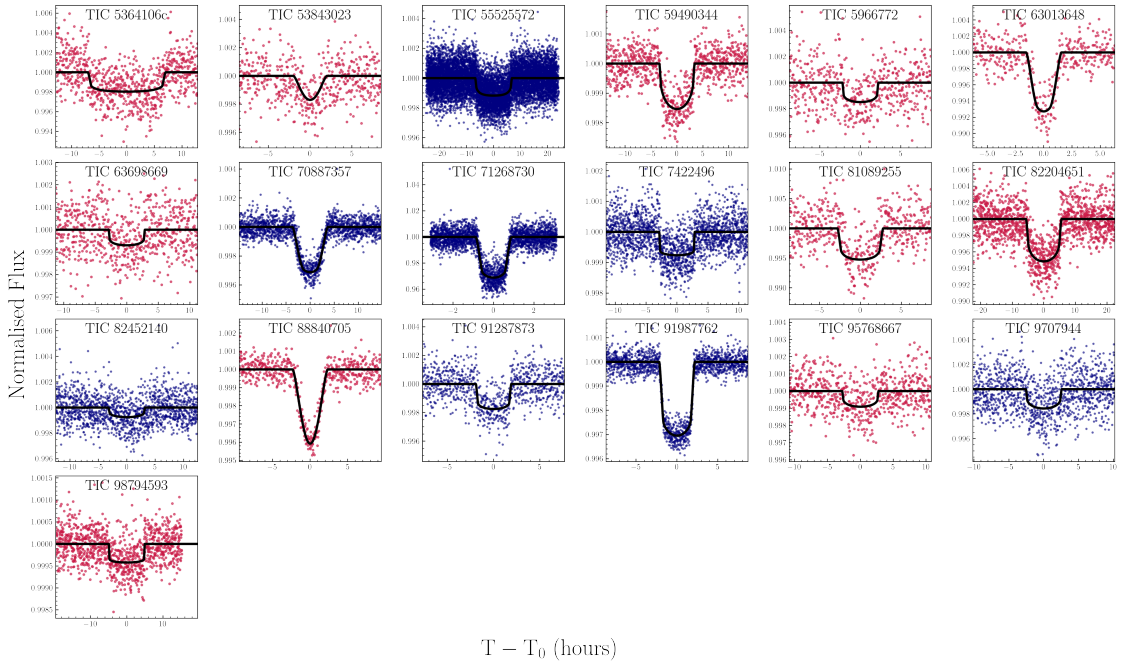

Figure 3.8: PHT candidates (continued)

TIC	Other Name	Epoch (BJD - 2457000)	Period (days)	$R_{pl}/R_{\odot}$	R_{pl} ($R_{\oplus}$)	Impact Parameter	Duration (hours)	V_{mag}	Photometry	Spectroscopy	Speckle	Comment
101641905	TWOMASS 11412617+3441004	$1917.26335^{+0.00071}_{-0.00072}$	$14.52^{+6.21}_{-5.25}(s)$	$0.1135^{+0.0032}_{-0.0064}$	$9.76^{+0.65}_{-0.69}$	$0.691^{+0.077}_{-0.183}$	$3.163^{+0.093}_{-0.088}$	12.196				
103633672*	TYC 4387-00923-1	$1850.3211^{+0.00135}_{-0.00077}$	$90.9^{+46.4}_{-23.7}(s)$	$0.0395^{+0.0013}_{-0.0013}$	$3.45^{+0.26}_{-0.24}$	$0.30^{+0.26}_{-0.21}$	$6.70^{+0.12}_{-0.11}$	10.586		NRES (1)		TOI 2085.01 (2)
110996418*	TWOMASS 12344723-1019107	$1580.6407^{+0.0045}_{-0.0040}$	$15.78159^{+0.00011}_{-0.00022}$	$0.1032^{+0.0073}_{-0.0075}$	$11.24^{+0.96}_{-0.97}$	$0.54^{+0.22}_{-0.34}$	$3.87^{+0.30}_{-0.29}$	13.945				TOI 5171.01 (3)
128703021	HIP 71639	$1601.8431^{+0.0013}_{-0.0015}$	$82.3^{+41.1}_{-32.9}$	$0.02463^{+0.00165}_{-0.00051}$	$4.35^{+0.29}_{-0.23}$	$0.41^{+0.40}_{-0.14}$	$7.21^{+0.298}_{-0.072}$	6.06		NRES (2);MINERVA (34)	Gemini	Transits separated by ∼741 d
130718055	TYC 4966-00020-1	$1949.7294^{+0.0033}_{-0.0059}$	$207.6497^{+10.8985}_{-0.0015}$	$0.0561^{+0.0016}_{-0.0018}$	$4.82^{+0.32}_{-0.30}$	$0.457^{+0.059}_{-0.095}$	$8.42^{+0.81}_{-0.26}$	10.437				Transits separated by ∼623 d
138126035*	TYC 1450-00833-1	$1954.3229^{+0.0067}_{-0.0041}$	$28.8^{+203.2}_{-14.0}(s)$	$0.0375^{+0.0069}_{-0.0026}$	$4.01^{+0.74}_{-0.35}$	$0.58^{+0.35}_{-0.38}$	$4.65^{+0.85}_{-0.32}$	10.349		CHIRON (1);NRES (2)		
139470998	TYC 6433-01584-1	$2126.0364^{+0.0033}_{-0.0027}$	$145.2^{+91.7}_{-51.6}(s)$	$0.0546^{+0.0017}_{-0.0018}$	$8.75^{+0.48}_{-0.46}$	$0.56^{+0.14}_{-0.29}$	$11.05^{+0.37}_{-0.25}$	10.283		CHIRON (18)		
141522677	TYC 4530-00025-1	$2410.6135^{+0.0027}_{-0.0022}$	$11.84^{+21.53}_{-4.35}(s)$	$0.0408^{+0.0046}_{-0.0022}$	$5.97^{+0.69}_{-0.43}$	$0.48^{+0.31}_{-0.34}$	$4.57^{+0.30}_{-0.20}$	9.806				
142087638*	TYC 9189-00274-1	$1512.1641^{+0.0046}_{-0.0053}$	$93.3^{+93.3}_{-31.1}$	$0.1154^{+0.0094}_{-0.0100}$	$14.79^{+1.36}_{-1.31}$	$1.038^{+0.016}_{-0.018}$	$4.08^{+0.37}_{-0.37}$	11.526				3 TOIs in this system, TOI 2404 (3)
149601126*	TWOMASS 05472416-6031170	$2056.0662^{+0.0038}_{-0.0037}$	$23.1927^{+0.0018}_{-0.0019}$	$0.0986^{+0.001}_{-0.0018}$	$8.83^{+1.09}_{-1.09}$	$0.16^{+0.15}_{-0.11}$	$4.44^{+0.16}_{-0.13}$	14.216				TOI 1873.01(3)
149601126*	TWOMASS 05472416-6031170	$2123.2702^{+0.0028}_{-0.0025}$	$60.0^{+30.5}_{-14.3}(s)$	$0.1195^{+0.0043}_{-0.0039}$	$10.76^{+0.77}_{-0.74}$	$0.31^{+0.24}_{-0.21}$	$6.69^{+0.32}_{-0.21}$	14.216				TOI 1873.02 (3)
159159904	HIP 64812	$1918.6109^{+0.0091}_{-0.0067}$	$584.0^{+1724.0}_{-215.0}(s)$	$0.0237^{+0.0026}_{-0.0011}$	$3.12^{+0.36}_{-0.22}$	$0.49^{+0.35}_{-0.34}$	$15.11^{+0.70}_{-0.54}$	9.2		NRES (2)		
160039081*	HIP 78892	$1752.9261^{+0.0050}_{-0.0045}$	$30.19918^{+0.00094}_{-0.00099}$	$0.0211^{+0.0035}_{-0.0013}$	$2.67^{+0.43}_{-0.21}$	$0.52^{+0.36}_{-0.34}$	$4.93^{+0.37}_{-0.27}$	8.35	SBIG (1)	NRES (1);SOPHIE (4)	Gemini	TOI 2082.01 (1)
162631539	HIP 80264	$1978.2794^{+0.0051}_{-0.0044}$	$17.32^{+52.35}_{-6.66}(s)$	$0.0195^{+0.0024}_{-0.0011}$	$2.94^{+0.38}_{-0.24}$	$0.48^{+0.36}_{-0.33}$	$5.54^{+0.41}_{-0.33}$	7.42				
166184426*	TWOMASS 13442500-4020122	$1600.4409^{+0.0036}_{-0.0030}$	$16.3325^{+0.0052}_{-0.0066}$	$0.0545^{+0.0039}_{-0.0031}$	$1.85^{+0.15}_{-0.12}$	$0.41^{+0.31}_{-0.28}$	$1.98^{+0.17}_{-0.22}$	12.911				TOI 1955.01 (1)
167661160*†	TYC 7054-01577-1	$1442.0703^{+0.0040}_{-0.0028}$	$36.802^{+0.069}_{-0.07}$	$0.0307^{+0.0024}_{-0.0014}$	$4.07^{+0.43}_{-0.32}$	$0.37^{+0.33}_{-0.26}$	$5.09^{+0.21}_{-0.23}$	9.927		NRES (9);MINERVA (4)		Long period EB from MINERVA; 2479.01 (3)
170161682	TYC 7639-01208-1	$2213.6809^{+0.004}_{-0.004}$	$25.158^{+0.0049}_{-0.0053}$	$0.1023^{+0.0048}_{-0.0033}$	$18.09^{+1.15}_{-1.07}$	$0.49^{+0.19}_{-0.28}$	$11.61^{+0.42}_{-0.29}$	13.068				First transit is only partial.
172370679*	TWOMASS 19574239+4008357	$1711.9594^{+0.0014}_{-0.0014}$	$29.090217^{+7.2e-05}_{-7.5e-05}$	$0.1849^{+0.0031}_{-0.0044}$	$12.44^{+0.30}_{-0.34}$	$0.534^{+0.043}_{-0.051}$	$4.822^{+0.088}_{-0.076}$	14.88				Confirmed planet (Cañas et al., 2020).
174302697*	TYC 3641-01789-1	$1743.72929^{+0.00082}_{-0.00078}$	$329.8^{+91.4}_{-76.0}(s)$	$0.07638^{+0.00059}_{-0.00064}$	$13.37^{+0.58}_{-0.58}$	$0.538^{+0.051}_{-0.062}$	$16.67^{+0.12}_{-0.11}$	9.309	SBIG (1)			Subgiant host; TOI 1896.01 (3)
179582003*	TYC 9166-00745-1	$1518.4688^{+0.0016}_{-0.0016}$	$104.6137^{+0.0022}_{-0.0022}$	$0.06324^{+0.0008}_{-0.0008}$	$7.51^{+0.35}_{-0.35}$	$0.21^{+0.19}_{-0.15}$	$9.073^{+0.097}_{-0.084}$	10.806				Subgiant host; Stellar pulsations (g-mode)
188668950	TWOMASS 17202383+5225597	$2400.6204^{+0.0059}_{-0.0048}$	$57.3368^{+5.7327}_{-0.0011}$	$0.0508^{+0.0030}_{-0.0024}$	$4.26^{+0.40}_{-0.37}$	$0.41^{+0.3}_{-0.27}$	$5.81^{+0.33}_{-0.25}$	12.193				
192415680	TYC 2859-00682-1	$1796.0265^{+0.0013}_{-0.0012}$	$18.47^{+21.73}_{-6.34}(s)$	$0.0478^{+0.0027}_{-0.0017}$	$4.43^{+0.38}_{-0.33}$	$0.45^{+0.31}_{-0.31}$	$3.94^{+0.12}_{-0.10}$	9.838	SBIG (1)	SOPHIE (2)		
192790476*	TYC 7595-00649-1	$1452.3344^{+0.0024}_{-0.0010}$	$19.10474^{+3.7e-05}_{-0.034279}$	$0.0399^{+0.0018}_{-0.0011}$	$2.96^{+0.36}_{-0.32}$	$0.41^{+0.25}_{-0.27}$	$3.251^{+0.412}_{-0.052}$	10.772				Wide binary; TOI 2459.01 (2)
206361691†	HIP 117250	$1363.2224^{+0.0090}_{-0.0082}$	$237.7^{+314.4}_{-81.0}(s)$	$0.01762^{+0.00125}_{-0.00088}$	$2.69^{+0.25}_{-0.19}$	$0.43^{+0.32}_{-0.28}$	$13.91^{+0.52}_{-0.53}$	8.88		CHIRON (6)		SB2 from CHIRON
207501148*	TYC 3881-00527-1	$2007.7273^{+0.0011}_{-0.0011}$	$39.9^{+14.3}_{-10.3}(s)$	$0.0981^{+0.011}_{-0.0047}$	$13.31^{+1.56}_{-0.95}$	$0.9^{+0.039}_{-0.03}$	$4.73^{+0.14}_{-0.14}$	10.385		NRES (3)		TOI 2320.01 (5)
209006017	TYC 8723-00513-1	$2376.0879^{+0.0018}_{-0.0019}$	$22.6^{+13.41}_{-6.19}(s)$	$0.0769^{+0.003}_{-0.0026}$	$13.73^{+0.85}_{-0.82}$	$0.38^{+0.25}_{-0.26}$	$6.53^{+0.17}_{-0.15}$	11.164				
209464063	TYC 7022-00871-1	$1424.6976^{+0.0029}_{-0.0142}$	$143.1937^{+0.0014}_{-23.8648}$	$0.0605^{+0.0012}_{-0.0015}$	$6.20^{+0.36}_{-0.35}$	$0.616^{+0.049}_{-0.374}$	$7.69^{+1.49}_{-0.18}$	11.438				Transits separated by ∼716 d
219466784*	TYC 4409-00437-1	$1872.6879^{+0.0108}_{-0.0097}$	$318.0^{+1448.0}_{-147.0}(s)$	$0.0332^{+0.0048}_{-0.0024}$	$3.26^{+0.49}_{-0.31}$	$0.00^{+0.34}_{-0.39}$	$10.06^{+1.12}_{-0.81}$	11.099				TOI 2007.01 (2)
219501568	HIP 79876	$1961.7879^{+0.0020}_{-0.0018}$	$16.5931^{+0.0015}_{-0.0017}$	$0.0221^{+0.0015}_{-0.0012}$	$4.22^{+0.35}_{-0.30}$	$0.41^{+0.31}_{-0.28}$	$1.615^{+0.093}_{-0.077}$	8.38				Subgiant host; TOI 2259.01 (3)
219641816	TYC 6089-00701-1	$2296.1100^{+0.0022}_{-0.0028}$	$14.67^{+9.48}_{-3.98}(s)$	$0.054^{+0.0025}_{-0.0024}$	$5.64^{+0.39}_{-0.37}$	$0.36^{+0.26}_{-0.23}$	$4.18^{+0.16}_{-0.14}$	12.142				
224307602	HIP 76096	$1979.3947^{+0.0075}_{-0.0085}$	$193.7^{+970.6}_{-75.9}(s)$	$0.0273^{+0.0052}_{-0.0014}$	$6.03^{+1.08}_{-0.46}$	$0.52^{+0.37}_{-0.34}$	$16.64^{+1.32}_{-0.66}$	9.12				Subgiant host
229055790	TYC 7492-01197-1	$1337.8651^{+0.0022}_{-0.0024}$	$52.5^{+43.7}_{-15.6}(s)$	$0.0297^{+0.0011}_{-0.0011}$	$3.43^{+0.21}_{-0.2}$	$0.46^{+0.23}_{-0.24}$	$6.51^{+0.14}_{-0.14}$	9.642		NRES (2)		
229608594*	TWOMASS 18180283+7428005	$1960.0319^{+0.0045}_{-0.0037}$	$152.4^{+152.6}_{-54.1}(s)$	$0.0474^{+0.0024}_{-0.0023}$	$3.42^{+0.36}_{-0.34}$	$0.38^{+0.30}_{-0.26}$	$6.98^{+0.37}_{-0.23}$	12.302				TOI 2298.01 (3)
229742722*	TYC 4434-00596-1	$1726.13987^{+0.00097}_{-0.00069}$	$63.48332^{+0.00017}_{-0.00011}$	$0.06985^{+0.00094}_{-0.0016}$	$10.6^{+0.55}_{-0.80}$	$0.561^{+0.075}_{-0.083}$	$4.812^{+0.056}_{-0.046}$	10.33		NRES (8);SOPHIE (4)	Gemini	TOI 1895.01 (3)
233194447	TYC 4211-00650-1	$1770.4926^{+0.0072}_{-0.0050}$	$370.7^{+200.9}_{-80.8}(s)$	$0.0246^{+0.00085}_{-0.00069}$	$5.87^{+0.21}_{-0.17}$	$0.31^{+0.27}_{-0.21}$	$24.82^{+0.54}_{-0.33}$	9.178		NRES (2)	Gemini	
235943205*	TYC 4588-00127-1	$1827.0267^{+0.0034}_{-0.004}$	$121.3394^{+0.0065}_{-0.0063}$	$0.0402^{+0.0019}_{-0.0016}$	$4.20^{+0.29}_{-0.25}$	$0.40^{+0.28}_{-0.27}$	$6.37^{+0.30}_{-0.20}$	11.076		NRES (1);SOPHIE (2)		TOI 2264.01 (3)
237201858*	TYC 4452-00759-1	$1811.515^{+0.018}_{-0.017}$	$179.389^{+0.027}_{-0.027}$	$0.1143^{+0.0106}_{-0.0099}$	$18.24^{+1.79}_{-1.64}$	$1.077^{+0.013}_{-0.012}$	$15.33^{+1.02}_{-0.94}$	10.344		NRES (1)		TOI 2286.01 (1)
242241304	TYC 7811-00974-1	$2345.374^{+0.0039}_{-0.0037}$	$53.2^{+86.4}_{-20.4}(s)$	$0.0694^{+0.0066}_{-0.0038}$	$4.90^{+0.57}_{-0.48}$	$0.47^{+0.32}_{-0.33}$	$4.17^{+0.28}_{-0.24}$	12.559				

Table 3.1: *Note* – Candidates that have become TOIs following the PHT discovery are marked with an asterisk (*). The 's' following the orbital period indicates that the candidates is a single-transit event. The ground-based follow-up observations are summarised in columns 10 – 12, where the bracketed numbers correspond the number of epochs obtained with each instrument. See Section 2.4 for description of each instrument. The † symbol indicates candidates that have been shown to be astrophysical false positives based on the ground-based follow-up observations. The number following the TOIs indicates the candidate priority given by the *TESS* team (5 = lowest; 1 = highest).

TIC	Other Name	Epoch (BJD - 2457000)	Period (days)	$R_{pl}/R_{\odot}$	R_{pl} ($R_{\oplus}$)	Impact Parameter	Duration (hours)	V_{mag}	Photometry	Spectroscopy	Speckle	Comment
243187830*	HIP 5286	$1783.7671^{+0.0019}_{-0.0017}$	$4.05^{+9.21}_{-1.53}(s)$	$0.0268^{+0.0027}_{-0.0015}$	$2.06^{+0.23}_{-0.17}$	$0.47^{+0.34}_{-0.32}$	$2.02^{+0.15}_{-0.12}$	8.407	SBIG (1)			TOI 2009.01 (3)
243417115	TYC 8262-02120-1	$1614.483^{+0.0030}_{-0.0025}$	$31.91^{+20.52}_{-7.45}$	$0.0596^{+0.0138}_{-0.006}$	$6.15^{+1.45}_{-0.73}$	$0.892^{+0.064}_{-0.221}$	$2.44^{+0.28}_{-0.39}$	11.553				Transits separated by ∼734 d
24795168	TYC 3644-01835-1	$1764.8843^{+0.0033}_{-0.0036}$	$13.24^{+24.23}_{-4.51}(s)$	$0.0295^{+0.0016}_{-0.0014}$	$4.8^{+0.38}_{-0.33}$	$0.46^{+0.33}_{-0.31}$	$5.75^{+0.25}_{-0.26}$	9.406				
256429408	TYC 4462-01942-1	$1962.16^{+0.0023}_{-0.0022}$	$382.0^{+265.0}_{-132.0}(s)$	$0.03582^{+0.00094}_{-0.00086}$	$6.12^{+0.3}_{-0.29}$	$0.51^{+0.18}_{-0.36}$	$16.96^{+0.24}_{-0.2}$	8.898				
258917846*	TYC 8251-01432-1	$1605.0969^{+0.0058}_{-0.0068}$	$32.29267^{+0.13777}_{-0.00045}$	$0.0359^{+0.0024}_{-0.0018}$	$4.84^{+0.51}_{-0.46}$	$0.47^{+0.24}_{-0.29}$	$8.31^{+0.31}_{-0.33}$	10.826				TOI 4557.01 (3)
264544388*	TYC 4607-01275-1	$1824.8408^{+0.0072}_{-0.0098}$	$460.0^{+432.0}_{-128.0}(s)$	$0.02498^{+0.00135}_{-0.00066}$	$4.0^{+0.26}_{-0.20}$	$0.31^{+0.36}_{-0.21}$	$17.11^{+0.56}_{-0.65}$	8.758		NRES (1)		Stellar pulsations (p-mode); TOI 1893.01 (3)
264766922	TYC 8565-01780-1	$1538.69518^{+0.00091}_{-0.00091}$	$3.28^{+1.25}_{-0.94}(s)$	$0.0933^{+0.0176}_{-0.0063}$	$16.95^{+3.19}_{-1.33}$	$0.908^{+0.048}_{-0.039}$	$2.73^{+0.11}_{-0.11}$	10.747				Subgiant host
26547036*	TYC 3921-01563-1	$1712.30464^{+0.0004}_{-0.00041}$	$73.0^{+16.5}_{-13.6}(s)$	$0.10034^{+0.00078}_{-0.0007}$	$11.75^{+0.58}_{-0.59}$	$0.17^{+0.11}_{-0.12}$	$8.68^{+0.05}_{-0.05}$	9.849		NRES (4)	Gemini	TOI 2010.01 (3)
267542728†	TYC 4583-01499-1	$1708.5119^{+0.0068}_{-0.0104}$	$39.73446^{+0.00052}_{-0.9261}$	$0.03829^{+0.0026}_{-0.00099}$	$20.44^{+1.38}_{-0.58}$	$0.4^{+0.27}_{-0.25}$	$24.88^{+0.38}_{-0.62}$	11.474				EB from HIRES RVs
268534931	TWOMASS 01003028-2310463	$2139.1007^{+0.0036}_{-0.0036}$	$122.3^{+97.8}_{-37.2}(s)$	$0.0687^{+0.0029}_{-0.0028}$	$5.82^{+0.43}_{-0.41}$	$0.41^{+0.24}_{-0.27}$	$7.15^{+0.43}_{-0.27}$	13.092				
270371513†	HIP 10047	$1426.2967^{+0.0020}_{-0.0023}$	$0.39^{+1.79}_{-0.17}(s)$	$0.024^{+0.0032}_{-0.0015}$	$4.8^{+0.64}_{-0.38}$	$0.50^{+0.39}_{-0.34}$	$1.93^{+0.19}_{-0.16}$	6.98515		CHIRON (2);MINERVA (20)		SB 2 from MINERVA observations.
27064468*	TYC 1415-01408-1	$2528.7751^{+0.0021}_{-0.002}$	$5.4613^{+0.0012}_{-0.0011}$	$0.0365^{+0.0018}_{-0.0014}$	$4.97^{+0.33}_{-0.30}$	$0.46^{+0.25}_{-0.3}$	$3.78^{+0.11}_{-0.10}$	10.141				TOI 5126.01 (3) & TOI 5126.02 (1)
130718055 273418421*	HIP 113293	$1681.577^{+0.0044}_{-0.0032}$	$65.1^{+13.0}_{-21.7}$	$0.02444^{+0.00091}_{-0.00098}$	$2.19^{+0.14}_{-0.14}$	$0.48^{+0.13}_{-0.24}$	$5.31^{+0.15}_{-0.18}$	9.56		CHIRON (24)		Transits separated by ∼390 d; TOI 4358.01 (3)
274599700*	TWOMASS 17011885+5131455	$2002.1202^{+0.0024}_{-0.0024}$	$32.9754^{+0.005}_{-0.005}$	$0.0847^{+0.0018}_{-0.0021}$	$13.25^{+0.83}_{-0.83}$	$0.37^{+0.19}_{-0.24}$	$8.20^{+0.21}_{-0.18}$	12.411				TOI 2317.01 (3)
278990954	TYC 8548-00717-1	$1650.0191^{+0.0105}_{-0.0086}$	$18.45^{+230.7}_{-8.66}(s)$	$0.034^{+0.0115}_{-0.0024}$	$9.65^{+3.13}_{-0.92}$	$0.58^{+0.36}_{-0.4}$	$10.62^{+2.46}_{-0.66}$	10.749				Subgiant host
280031353	TWOMASS 19493982+6853000	$2398.7358^{+0.0027}_{-0.0027}$	$1350.0^{+328.0}_{-261.0}(s)$	$0.09878^{+0.00086}_{-0.00133}$	$8.57^{+0.45}_{-0.45}$	$0.17^{+0.12}_{-0.11}$	$17.33^{+0.28}_{-0.23}$	12.518				Two additional signals (TOI 2300.01 and 02)
280865159*	TYC 9384-01533-1	$1387.0732^{+0.0047}_{-0.0047}$	$38.90744^{+0.00046}_{-0.00048}$	$0.0558^{+0.0027}_{-0.0017}$	$6.53^{+0.40}_{-0.36}$	$0.41^{+0.22}_{-0.27}$	$19.63^{+0.47}_{-0.37}$	11.517			Gemini	Transits separated by ∼739 d; TOI 1894.01
284361752*	TYC 3924-01678-1	$2032.091^{+0.011}_{-0.012}$	$130.4^{+157.8}_{-68.4}(s)$	$0.0263^{+0.0028}_{-0.0016}$	$3.68^{+0.46}_{-0.3}$	$0.43^{+0.35}_{-0.28}$	$10.34^{+0.55}_{-1.4}$	10.221		NRES (3)		
288240183	TYC 4634-01225-1	$1896.941^{+0.0047}_{-0.0051}$	$119.0502^{+0.0089}_{-0.0091}$	$0.02826^{+0.00119}_{-0.00089}$	$4.28^{+0.36}_{-0.35}$	$0.55^{+0.25}_{-0.37}$	$17.49^{+0.60}_{-0.36}$	9.546		NRES (1)		
29169215	TWOMASS 09011787+4727085	$1872.5047^{+0.0036}_{-0.0032}$	$14.89^{+24.84}_{-6.12}(s)$	$0.0403^{+0.0033}_{-0.0025}$	$3.28^{+0.45}_{-0.37}$	$0.44^{+0.33}_{-0.30}$	$3.56^{+0.32}_{-0.21}$	11.828				
292202337*	TYC 6767-00204-1	$1619.20798^{+0.00035}_{-0.00051}$	$27.75^{+3.62}_{-3.7}$	$0.04236^{+0.00095}_{-0.00096}$	$4.47^{+0.25}_{-0.25}$	$0.277^{+0.047}_{-0.075}$	$4.61^{+0.117}_{-0.068}$	9.465				Transits separated by ∼722 d; TOI 4630.01 (3)
293649602	TYC 8103-00266-1	$1511.2109^{+0.0037}_{-0.0040}$	$12.85^{+42.21}_{-5.34}(s)$	$0.04^{+0.0039}_{-0.0024}$	$4.66^{+0.50}_{-0.36}$	$0.5^{+0.34}_{-0.35}$	$4.10^{+0.56}_{-0.31}$	10.925				
296737508	TYC 5472-01060-1	$1538.0036^{+0.0016}_{-0.0015}$	$18.27^{+17.45}_{-5.06}(s)$	$0.0425^{+0.0019}_{-0.0014}$	$5.33^{+0.27}_{-0.22}$	$0.44^{+0.26}_{-0.3}$	$5.13^{+0.15}_{-0.13}$	9.772	Sinistro (1)	CHIRON (1);NRES (3);MINERVA (1)	Gemini	
298663873*	TYC 3913-01781-1	$1830.76871^{+0.00105}_{-0.00098}$	$479.0^{+108.6}_{-88.5}(s)$	$0.06332^{+0.00044}_{-0.00035}$	$11.25^{+0.58}_{-0.58}$	$0.17^{+0.11}_{-0.11}$	$23.97^{+0.10}_{-0.09}$	9.162		NRES (2)	Gemini	Confirmed planet (Dalba et al., 2022)
303050301	TYC 6979-01108-1	$1366.1301^{+0.0023}_{-0.0022}$	$281.0^{+264.0}_{-170.0}(s)$	$0.0514^{+0.0018}_{-0.0027}$	$4.85^{+0.32}_{-0.32}$	$0.73^{+0.10}_{-0.48}$	$7.91^{+0.36}_{-0.31}$	10.048		NRES (1)	Gemini	
303317324*	TYC 6983-00438-1	$1365.1772^{+0.0029}_{-0.0032}$	$123.2^{+123.2}_{-49.3}$	$0.037^{+0.0018}_{-0.0016}$	$3.74^{+0.26}_{-0.25}$	$0.69^{+0.12}_{-0.29}$	$5.99^{+0.25}_{-0.19}$	10.799		CHIRON (20);NRES (2)		Transits separated by ∼739 d; 4310.01 (1)
303586471*†	HIP 115828	$1363.7709^{+0.0023}_{-0.0025}$	$17.5668^{+0.354}_{-0.0063}$	$0.02109^{+0.00179}_{-0.00095}$	$2.49^{+0.24}_{-0.16}$	$0.44^{+0.36}_{-0.32}$	$4.23^{+0.23}_{-0.14}$	8.27		MINERVA (11)		SB 2 from MINERVA observations; TOI 4326.01 (2)
304142124*	HIP 53719	$1585.28023^{+0.0008}_{-0.0008}$	$42.8^{+18.2}_{-10.0}(s)$	$0.04311^{+0.00153}_{-0.00093}$	$4.1^{+0.24}_{-0.23}$	$0.33^{+0.21}_{-0.21}$	$5.66^{+0.09}_{-0.06}$	8.62		NRES (1);MINERVA (4)		Confirmed planet (Díaz et al., 2020)
304339227*	TYC 9290-01087-1	$1673.3242^{+0.0128}_{-0.0090}$	$111.9^{+4844.1}_{-72.2}(s)$	$0.0253^{+0.0481}_{-0.0024}$	$3.27^{+5.72}_{-0.61}$	$0.67^{+0.36}_{-0.47}$	$7.44^{+2.84}_{-0.86}$	9.169		CHIRON (21)		Transits separated by ∼715 d; 4629.01 (1)
306472057*	TYC 9184-01675-1	$1427.628^{+0.0015}_{-0.0014}$	$139.30306^{+0.00036}_{-0.00037}$	$0.06679^{+0.00072}_{-0.00064}$	$10.42^{+0.44}_{-0.44}$	$0.19^{+0.17}_{-0.14}$	$12.17^{+0.105}_{-0.089}$	11.093				TOI 791.01 (1)
307958020*	TYC 4191-00309-1	$1864.8309^{+0.007}_{-0.0101}$	$271.933^{+0.01}_{-135.968}$	$0.02075^{+0.00141}_{-0.00086}$	$3.6^{+0.32}_{-0.28}$	$0.48^{+0.27}_{-0.32}$	$11.4^{+0.63}_{-0.30}$	9.017		NRES (6)		Transits separated by ∼544 d; TOI 4633.01 (1)
308301091	TYC 2081-01273-1	$2030.3691^{+0.0022}_{-0.0024}$	$29.24^{+22.46}_{-8.49}(s)$	$0.0362^{+0.0014}_{-0.0013}$	$5.41^{+0.35}_{-0.34}$	$0.35^{+0.29}_{-0.25}$	$6.57^{+0.19}_{-0.14}$	10.273		NRES (2)		
313006381*	HIP 45012	$1705.6864^{+0.0055}_{-0.0061}$	$26.16^{+44.47}_{-9.73}$	$0.0281^{+0.0355}_{-0.0029}$	$2.53^{+3.11}_{-0.31}$	$0.74^{+0.27}_{-0.25}$	$4.07^{+1.20}_{-0.42}$	9.39				
318812447*	TYC 0050-00821-1	$2147.1342^{+0.0033}_{-0.0033}$	$25.34^{+50.46}_{-9.14}(s)$	$0.0646^{+0.0081}_{-0.0036}$	$6.8^{+0.81}_{-0.55}$	$0.50^{+0.32}_{-0.38}$	$4.81^{+0.32}_{-0.21}$	11.511				TOI 2634.01 (3)
3224605	TWOMASS 05293239+3314493	$2483.0112^{+0.0021}_{-0.0022}$	$28.6454^{+0.666}_{-0.0017}$	$0.0856^{+0.0046}_{-0.003}$	$8.28^{+0.61}_{-0.55}$	$0.35^{+0.26}_{-0.24}$	$4.25^{+0.18}_{-0.13}$	12.624				
323295479*	TYC 9506-01881-1	$1622.92568^{+0.00096}_{-0.00096}$	$148.5^{+99.0}_{-74.3}$	$0.1008^{+0.00106}_{-0.00054}$	$11.7^{+0.63}_{-0.62}$	$0.847^{+0.012}_{-0.014}$	$6.66^{+0.12}_{-0.10}$	10.595				Transits separated by ∼743 d; TOI 1861.01 (3)
324609409	TYC 9473-00833-1	$2378.4174^{+0.0024}_{-0.0024}$	$146.7^{+65.4}_{-49.4}(s)$	$0.078^{+0.002}_{-0.0024}$	$11.61^{+0.62}_{-0.62}$	$0.631^{+0.088}_{-0.173}$	$11.09^{+0.29}_{-0.28}$	11.434				
328933398*	TYC 4634-01435-1	$1880.9878^{+0.0042}_{-0.0039}$	$24.9335^{+0.005}_{-0.0046}$	$0.0437^{+0.0023}_{-0.0022}$	$4.62^{+0.33}_{-0.32}$	$0.38^{+0.27}_{-0.25}$	$5.02^{+0.27}_{-0.22}$	11.215				3 TOIs; 1873.01-1873.03
331644554	TYC 3609-00469-1	$1757.0354^{+0.0033}_{-0.0031}$	$947.0^{+274.0}_{-215.0}(s)$	$0.12^{+0.021}_{-0.025}$	$21.84^{+3.86}_{-4.57}$	$1.018^{+0.028}_{-0.036}$	$10.93^{+0.35}_{-0.34}$	9.752				
332657786	TWOMASS 09595797-1609323	$1536.7659^{+0.0015}_{-0.0014}$	$63.76^{+11.13}_{-9.52}(s)$	$0.14961^{+0.00029}_{-0.00064}$	$3.83^{+0.12}_{-0.12}$	$0.059^{+0.064}_{-0.041}$	$3.33^{+0.10}_{-0.10}$	15.99				
336075472*	TYC 3526-00332-1	$2028.1735^{+0.0066}_{-0.0068}$	$378.868^{+90.147}_{-0.014}$	$0.1115^{+0.0117}_{-0.0084}$	$8.57^{+1.13}_{-0.97}$	$1.027^{+0.016}_{-0.015}$	$6.61^{+0.53}_{-0.57}$	11.842				Transits separated by ∼379 d; TOI 4632.01 (1)
349488688.01	TYC 1529-00224-1	$1994.283^{+0.0035}_{-0.0038}$	$11.6254^{+0.0052}_{-0.005}$	$0.02195^{+0.00122}_{-0.00096}$	$3.44^{+0.21}_{-0.18}$	$0.39^{+0.31}_{-0.27}$	$5.58^{+0.18}_{-0.15}$	8.855		NRES (2);SOPHIE (2)		Confirmed planet (Eisner et al., 2021b)
349488688.02	TYC 1529-00224-1	$2002.77063^{+0.00103}_{-0.00097}$	$15.35^{+4.15}_{-1.94}(s)$	$0.03688^{+0.00069}_{-0.00067}$	$5.78^{+0.18}_{-0.18}$	$0.24^{+0.21}_{-0.16}$	$6.291^{+0.074}_{-0.058}$	8.855		NRES (2);SOPHIE (2)		Confirmed planet (Eisner et al., 2021b)
349576261*	TWOMASS 07280242-6331041	$2132.2155^{+0.0011}_{-0.0010}$	$10.72^{+6.42}_{-4.15}(s)$	$0.1084^{+0.0037}_{-0.0041}$	$14.79^{+0.84}_{-0.84}$	$0.59^{+0.12}_{-0.29}$	$4.60^{+0.30}_{-0.20}$	12.14		CHIRON (2)		TOI 4562.01 (3); Heitzmann et al. *in prep*
35021200	TYC 4938-00997-1	$2559.854^{+0.011}_{-0.015}$	$305.0^{+408.0}_{-106.0}(s)$	$0.0347^{+0.0026}_{-0.0021}$	$3.59^{+0.35}_{-0.29}$	$0.41^{+0.31}_{-0.26}$	$10.93^{+1.26}_{-0.67}$	11.774				Two additional confirmed planets (K2-158)
355881602	TWOMASS 09123527-5331547	$2302.9242^{+0.0028}_{-0.0025}$	$77.2^{+31.6}_{-13.8}(s)$	$0.0909^{+0.0039}_{-0.0033}$	$5.23^{+0.27}_{-0.25}$	$0.27^{+0.26}_{-0.19}$	$5.30^{+0.19}_{-0.18}$	13.817				
356700488*	TYC 4420-01295-1	$1756.638^{+0.011}_{-0.013}$	$184.5^{+333.1}_{-64.7}(s)$	$0.0173^{+0.0015}_{-0.0011}$	$2.92^{+0.28}_{-0.20}$	$0.44^{+0.34}_{-0.30}$	$11.76^{+1.03}_{-0.65}$	8.413		NRES (1)		TOI 2098.01 (2)
356710041*	TYC 1993-00419-1	$1932.2939^{+0.0019}_{-0.0019}$	$29.6^{+19.0}_{-14.0}(s)$	$0.0496^{+0.0011}_{-0.0021}$	$14.82^{+0.84}_{-0.85}$	$0.66^{+0.11}_{-0.42}$	$12.76^{+0.24}_{-0.24}$	9.646				TOI 2065.01 (2)
365007485	TWOMASS 10564551+0940263	$2554.9429^{+0.0044}_{-0.0036}$	$25.52^{+25.82}_{-7.95}(s)$	$0.0546^{+0.0037}_{-0.0032}$	$5.45^{+0.50}_{-0.44}$	$0.39^{+0.29}_{-0.26}$	$4.80^{+0.32}_{-0.22}$	12.305				
369532319	TYC 2743-01716-1	$1755.8158^{+0.0051}_{-0.0060}$	$35.4^{+51.6}_{-12.0}(s)$	$0.0316^{+0.0028}_{-0.0023}$	$3.43^{+0.37}_{-0.30}$	$0.41^{+0.34}_{-0.29}$	$5.50^{+0.32}_{-0.32}$	10.594			Gemini	
369779127	TYC 9510-00090-1	$1643.9403^{+0.0058}_{-0.0046}$	$9.93^{+19.74}_{-3.38}(s)$	$0.0288^{+0.0033}_{-0.0015}$	$4.89^{+0.56}_{-0.31}$	$0.46^{+0.33}_{-0.31}$	$5.64^{+0.33}_{-0.38}$	9.279		CHIRON (3)		

Table 3.1: Properties of PHT candidates (continued)

TIC	Other Name	Epoch (BJD - 2457000)	Period (days)	$R_{pl}/R_{\odot}$	R_{pl} ($R_{\oplus}$)	Impact Parameter	Duration (hours)	V_{mag}	Photometry	Spectroscopy	Speckle	Comment
380884709	TYC 0612-00310-1	$2465.6215^{+0.0024}_{-0.0024}$	$25.32^{+19.31}_{-7.41}(s)$	$0.0523^{+0.0021}_{-0.002}$	$5.52^{+0.37}_{-0.36}$	$0.40^{+0.25}_{-0.28}$	$4.91^{+0.22}_{-0.18}$	11.192				
384159646*	TYC 9454-00957-1	$1630.39448^{+0.00089}_{-0.00083}$	$62.3^{+20.8}_{-18.3}$	$0.0653^{+0.0011}_{-0.0011}$	$9.81^{+0.44}_{-0.40}$	$0.43^{+0.12}_{-0.21}$	$5.162^{+0.099}_{-0.074}$	10.158	SBIG (1)	NRES (8);MINERVA (6)	Gemini	Transits separated by ∼748; TOI 1895.01 (3)
385557214	TYC 1807-00046-1	$1791.58557^{+0.00049}_{-0.00059}$	$5.6232588^{+0.3946179}_{-5.8e-06}$	$0.0904^{+0.0069}_{-0.0192}$	$7.70^{+0.91}_{-1.59}$	$0.917^{+0.026}_{-0.318}$	$1.21^{+0.05}_{-0.18}$	10.856				
388134787	TYC 4260-00427-1	$1811.034^{+0.017}_{-0.015}$	$246.0^{+6209.0}_{-127.0}(s)$	$0.0265^{+0.023}_{-0.0024}$	$2.57^{+2.19}_{-0.28}$	$0.55^{+0.44}_{-0.39}$	$8.85^{+1.84}_{-1.13}$	10.95		NRES (1)	Gemini	
393546540	TYC 0260-00121-1	$2537.5035^{+0.0061}_{-0.0062}$	$32.4250^{+0.0081}_{-0.0080}$	$0.1134^{+0.0110}_{-0.0094}$	$10.25^{+1.07}_{-0.94}$	$1.00^{+0.01}_{-0.01}$	$6.57^{+0.46}_{-0.42}$	9.835				
393546540	TYC 0260-00121-1	$2563.1308^{+0.0016}_{-0.0017}$	$65.7^{+39.6}_{-17.1}(s)$	$0.0371^{+0.0012}_{-0.0011}$	$3.35^{+0.21}_{-0.21}$	$0.37^{+0.24}_{-0.24}$	$6.03^{+0.14}_{-0.11}$	9.835				
40001645	TWOMASS 22282534-0605119	$2465.7295^{+0.0030}_{-0.0032}$	$7.13^{+4.83}_{-1.92}(s)$	$0.08^{+0.0037}_{-0.0033}$	$13.75^{+1.01}_{-0.93}$	$0.34^{+0.27}_{-0.21}$	$5.52^{+0.29}_{-0.20}$	13.363				
400786146*	TYC 8970-00497-1	$2323.0812^{+0.0025}_{-0.0029}$	$22.1535^{+0.0042}_{-0.0046}$	$0.0757^{+0.0034}_{-0.0035}$	$13.41^{+0.93}_{-0.89}$	$0.892^{+0.024}_{-0.045}$	$7.56^{+0.31}_{-0.25}$	11.389				TOI 4564.01 (3)
404518509*	HIP 16038	$1431.2689^{+0.0045}_{-0.0044}$	$64.0^{+53.3}_{-28.8}$	$0.0267^{+0.0044}_{-0.0023}$	$3.03^{+0.55}_{-0.31}$	$0.846^{+0.094}_{-0.39}$	$5.21^{+0.76}_{-0.37}$	9.17		CHIRON (16)		TOI 4320.01 (1) & TOI 4320.01 (2)
408636441*	TYC 4266-00736-1	$1745.4668^{+0.0015}_{-0.0016}$	$37.695^{+0.0033}_{-0.0034}$	$0.0485^{+0.0023}_{-0.0019}$	$3.32^{+0.19}_{-0.16}$	$0.39^{+0.29}_{-0.27}$	$3.63^{+0.14}_{-0.10}$	11.93	SBIG (2)		Gemini	TOI 1759.01 (1)
415399713	TYC 0670-00465-1	$2176.6426^{+0.0095}_{-0.0112}$	$115.4^{+57.7}_{-28.9}$	$0.0261^{+0.0013}_{-0.0013}$	$2.63^{+0.19}_{-0.18}$	$0.33^{+0.22}_{-0.22}$	$7.88^{+0.39}_{-0.51}$	10.91				
418255064*	TWOMASS 13063680-8037015	$1629.3304^{+0.0018}_{-0.0018}$	$25.37^{+15.41}_{-7.06}(s)$	$0.0732^{+0.0031}_{-0.0029}$	$5.57^{+0.38}_{-0.36}$	$0.37^{+0.25}_{-0.25}$	$3.83^{+0.14}_{-0.13}$	12.478	SBIG (1)		Gemini	Transits separated by ∼743; TOI 4563.01 (3)
420645189*†	TYC 4508-00478-1	$1837.4767^{+0.0017}_{-0.0018}$	$250.2^{+99.4}_{-66.6}(s)$	$0.0784^{+0.0046}_{-0.0033}$	$8.82^{+0.70}_{-0.55}$	$0.892^{+0.028}_{-0.027}$	$6.95^{+0.30}_{-0.26}$	10.595		MINERVA (1)		SB 2 from MINERVA observations; TOI 1897.01
422082582	TYC 0574-00621-1	$2463.2915^{+0.0048}_{-0.0044}$	$15.51^{+22.53}_{-6.21}(s)$	$0.0411^{+0.0034}_{-0.0025}$	$4.01^{+0.39}_{-0.31}$	$0.46^{+0.31}_{-0.33}$	$3.70^{+0.29}_{-0.31}$	11.046				
422914082*	TYC 0046-00133-1	$1431.5551^{+0.0017}_{-0.0025}$	$47.8^{+31.8}_{-16.6}$	$0.0427^{+0.0033}_{-0.0028}$	$4.04^{+0.43}_{-0.38}$	$0.70^{+0.18}_{-0.38}$	$4.24^{+0.42}_{-0.37}$	11.026	Sinistro (1)	CHIRON (1);NRES (1)		Transits separated by ∼239; TOI 4317.01 (1)
423360793	TYC 4674-00753-1	$1397.0609^{+0.0058}_{-0.0073}$	$22.69551^{+0.00028}_{-0.47269}$	$0.01709^{+0.00119}_{-0.00083}$	$2.35^{+0.21}_{-0.15}$	$0.41^{+0.36}_{-0.29}$	$5.40^{+0.38}_{-0.16}$	8.407				
423785115	TYC 6468-00912-1	$2197.00336^{+0.00063}_{-0.00058}$	$90.6^{+50.9}_{-30.3}(s)$	$0.1384^{+0.0012}_{-0.002}$	$9.65^{+1.26}_{-1.25}$	$0.657^{+0.026}_{-0.039}$	$4.89^{+0.09}_{-0.09}$	11.901		CHIRON (1)		
427344083	TWOMASS 22563609+7040518	$1961.8967^{+0.0036}_{-0.0031}$	$7.77^{+9.65}_{-5.6}(s)$	$0.107^{+0.025}_{-0.016}$	$12.27^{+2.90}_{-1.87}$	$0.834^{+0.094}_{-0.484}$	$2.88^{+0.42}_{-0.30}$	13.404				
436873727*	HIP 13224	$2510.82659^{+0.00035}_{-0.00045}$	$22.093429^{+3.1e-05}_{-2e-05}$	$0.05109^{+0.0007}_{-0.00296}$	$9.71^{+0.44}_{-0.69}$	$0.719^{+0.025}_{-0.143}$	$5.37^{+0.04}_{-0.08}$	7.51				TOI 4641.01 (3)
441642457*	TYC 3858-00452-1	$1745.5128^{+0.0075}_{-0.0124}$	$79.8042^{+0.0116}_{-0.0055}$	$0.0259^{+0.0031}_{-0.0015}$	$3.28^{+0.43}_{-0.25}$	$0.51^{+0.34}_{-0.32}$	$6.49^{+0.56}_{-0.24}$	9.996				TOI 2073.01 (2)
441765914*	TWOMASS 17253007+7552562	$1769.6118^{+0.0033}_{-0.0022}$	$89.1^{+35.6}_{-26.7}$	$0.0422^{+0.0035}_{-0.0023}$	$3.84^{+0.44}_{-0.34}$	$0.68^{+0.13}_{-0.37}$	$7.50^{+0.64}_{-0.27}$	11.638				TOI 2088.01 (1)
452920657	TWOMASS 00332018+5906355	$1810.5765^{+0.0030}_{-0.0031}$	$53.2^{+34.3}_{-29.0}(s)$	$0.135^{+0.012}_{-0.016}$	$9.71^{+0.90}_{-1.16}$	$0.73^{+0.11}_{-0.48}$	$4.60^{+0.29}_{-0.26}$	14.167	SBIG (1)			
453029782	TWOMASS 09371251+0458191	$2550.1073^{+0.0036}_{-0.0030}$	$31.2^{+41.0}_{-11.2}(s)$	$0.0584^{+0.0036}_{-0.0031}$	$4.76^{+0.42}_{-0.38}$	$0.47^{+0.28}_{-0.31}$	$4.30^{+0.25}_{-0.22}$	12.099				
455737331	TYC 2779-00785-1	$1780.7084^{+0.0073}_{-0.0080}$	$50.4^{+75.0}_{-17.6}(s)$	$0.0257^{+0.002}_{-0.0016}$	$3.05^{+0.29}_{-0.24}$	$0.43^{+0.33}_{-0.29}$	$6.60^{+0.50}_{-0.43}$	10.189	SBIG (1)		Gemini	
456306892	TWOMASS 07521577+7009567	$1854.4582^{+0.0017}_{-0.0018}$	$33.53137^{+0.00016}_{-0.00015}$	$0.1117^{+0.0037}_{-0.0062}$	$9.61^{+0.73}_{-0.75}$	$0.782^{+0.044}_{-0.151}$	$4.12^{+0.19}_{-0.28}$	12.583				
456909420	TYC 1208-01094-1	$1779.4119^{+0.0023}_{-0.0031}$	$0.58^{+3.26}_{-0.28}(s)$	$0.0483^{+0.0097}_{-0.004}$	$5.72^{+1.14}_{-0.62}$	$0.54^{+0.38}_{-0.38}$	$1.53^{+0.23}_{-0.16}$	10.941				
458451774	TWOMASS 12551793+4431260	$1917.1875^{+0.0019}_{-0.0019}$	$12.39^{+83.97}_{-6.34}(s)$	$0.0752^{+0.0211}_{-0.0054}$	$3.33^{+0.92}_{-0.26}$	$0.61^{+0.32}_{-0.43}$	$2.08^{+0.59}_{-0.19}$	13.713				
48018596*	TYC 3548-00800-1	$1713.4529^{+0.0028}_{-0.0027}$	$30.03334^{+0.00017}_{-0.00019}$	$0.0373^{+0.0379}_{-0.0064}$	$6.00^{+6.05}_{-1.08}$	$0.90^{+0.11}_{-0.37}$	$2.57^{+0.53}_{-0.33}$	9.595		NRES (1)	Gemini	Transits separated by ∼100; TOI 2295.01 (3)
49428710*	TYC 4930-01080-1	$2551.7502^{+0.0106}_{-0.0091}$	$12.2105^{+0.0056}_{-0.0059}$	$0.0412^{+0.0013}_{-0.0013}$	$4.91^{+0.28}_{-0.28}$	$0.16^{+0.13}_{-0.11}$	$4.72^{+0.13}_{-0.12}$	11.583				TOI 5174.01 (3)
53309262	TWOMASS 07475406+5741549	$1863.1134^{+0.0094}_{-0.0098}$	$255.8^{+99.0}_{-66.9}(s)$	$0.0978^{+0.0016}_{-0.0033}$	$4.21^{+0.16}_{-0.18}$	$0.20^{+0.18}_{-0.14}$	$6.71^{+0.64}_{-0.56}$	15.51				
5364106	TYC 6040-00575-1	$2259.2153^{+0.0039}_{-0.0039}$	$27.0^{+17.6}_{-7.38}(s)$	$0.0396^{+0.002}_{-0.0021}$	$5.45^{+0.40}_{-0.39}$	$0.34^{+0.26}_{-0.22}$	$14.10^{+0.40}_{-0.40}$	11.861				
5364106	TYC 6040-00575-1	$2256.4407^{+0.0061}_{-0.0057}$	$277.6^{+167.5}_{-71.5}(s)$	$0.0420^{+0.0015}_{-0.0015}$	$5.79^{+0.36}_{-0.36}$	$0.35^{+0.26}_{-0.23}$	$6.44^{+0.28}_{-0.20}$	11.861				
53843023	TYC 6956-00758-1	$1328.0335^{+0.0057}_{-0.0054}$	$202.0^{+272.0}_{-189.0}(s)$	$0.058^{+0.056}_{-0.020}$	$5.14^{+4.99}_{-1.77}$	$0.962^{+0.083}_{-0.597}$	$4.25^{+0.66}_{-0.72}$	11.571				
55525572*	TYC 8876-01059-1	$1454.6713^{+0.0065}_{-0.0066}$	$83.8951^{+0.004}_{-0.004}$	$0.0343^{+0.0021}_{-0.0010}$	$7.31^{+0.56}_{-0.46}$	$0.43^{+0.31}_{-0.29}$	$13.54^{+0.51}_{-0.30}$	10.358		CHIRON (8)	Gemini	Confirmed planet (Eisner et al., 2020b)
59490344	TYC 1350-00663-1	$2509.7983^{+0.0018}_{-0.0020}$	$113.8^{+106.9}_{-40.5}(s)$	$0.0361^{+0.0027}_{-0.0016}$	$3.34^{+0.30}_{-0.25}$	$0.53^{+0.22}_{-0.26}$	$6.85^{+0.16}_{-0.15}$	9.115				
5966772	TYC 6697-00833-1	$2331.1041^{+0.0066}_{-0.0113}$	$28.1^{+29.4}_{-11.7}(s)$	$0.0361^{+0.0025}_{-0.0023}$	$2.95^{+0.31}_{-0.29}$	$0.36^{+0.33}_{-0.25}$	$4.40^{+0.35}_{-0.64}$	11.251				
63013648	TYC 1374-00735-1	$2516.5335^{+0.0014}_{-0.0015}$	$15.17^{+8.94}_{-9.74}(s)$	$0.0924^{+0.0048}_{-0.0131}$	$10.28^{+1.26}_{-1.54}$	$0.874^{+0.033}_{-0.282}$	$3.15^{+0.15}_{-0.23}$	11.359				
63698669*	TYC 6993-00729-1	$1364.6226^{+0.0067}_{-0.0074}$	$73.6^{+133.6}_{-26.8}(s)$	$0.0248^{+0.0023}_{-0.0019}$	$2.15^{+0.25}_{-0.2}$	$0.42^{+0.35}_{-0.29}$	$5.63^{+0.57}_{-0.32}$	10.701	SBIG (1)	CHIRON (13)		TOI 1892.01 (3)
70887357*	TYC 5883-01412-1	$1454.3331^{+0.0013}_{-0.0015}$	$241.3^{+120.6}_{-60.3}$	$0.0624^{+0.0029}_{-0.0049}$	$13.18^{+0.98}_{-1.18}$	$0.922^{+0.013}_{-0.041}$	$7.39^{+0.17}_{-0.33}$	9.293				Transits separated by ∼723; TOI 2008.01 (3)
71268730	TWOMASS 07350822+7124020	$2391.36557^{+0.00031}_{-0.00030}$	$1.721549^{+3.3e-05}_{-3.3e-05}$	$0.1701^{+0.0017}_{-0.0029}$	$11.72^{+0.39}_{-0.40}$	$0.504^{+0.042}_{-0.067}$	$1.76^{+0.02}_{-0.02}$	14.063				
7422496*†	HIP 25359	$1470.3613^{+0.0023}_{-0.0019}$	$46.96216^{+0.00015}_{-0.00016}$	$0.0268^{+0.00063}_{-0.00062}$	$2.57^{+0.14}_{-0.13}$	$0.34^{+0.30}_{-0.24}$	$5.75^{+0.12}_{-0.07}$	9.36		MINERVA (4)		SB 2 from MINERVA observations; TOI 4189.01 (1)
81089255	TWOMASS 08104660-4351125	$2260.2071^{+0.0034}_{-0.0035}$	$8.54^{+8.65}_{-2.79}(s)$	$0.0687^{+0.0035}_{-0.0044}$	$10.68^{+0.72}_{-0.73}$	$0.41^{+0.30}_{-0.26}$	$4.35^{+0.27}_{-0.25}$	12.256				
82204651	TYC 1964-00260-1	$2534.2604^{+0.0029}_{-0.0026}$	$68.9^{+40.2}_{-19.6}(s)$	$0.066^{+0.0023}_{-0.0019}$	$11.11^{+0.82}_{-0.77}$	$0.36^{+0.24}_{-0.24}$	$11.25^{+0.33}_{-0.28}$	11.8				
82452140*	TYC 3076-00921-1	$1964.2980^{+0.0220}_{-0.0180}$	$63.38^{+2.87}_{-2.61}$	$0.026^{+0.0059}_{-0.0045}$	$2.62^{+0.58}_{-0.46}$	$0.55^{+0.18}_{-0.24}$	$6.25^{+0.94}_{-1.00}$	10.616				TOI 2289.01 (2)
88840705*	TYC 3091-00808-1	$2026.6489^{+0.0010}_{-0.0010}$	$260.6^{+142.2}_{-87.6}(s)$	$0.109^{+0.027}_{-0.023}$	$9.98^{+2.75}_{-2.28}$	$1.001^{+0.037}_{-0.042}$	$4.72^{+0.15}_{-0.13}$	9.443				TOI 2318.01 (3)
91287873	TYC 1907-01085-1	$2515.1728^{+0.0021}_{-0.0020}$	$26.6928^{+0.003}_{-0.0029}$	$0.0409^{+0.0014}_{-0.0013}$	$2.80^{+0.13}_{-0.12}$	$0.64^{+0.13}_{-0.12}$	$3.84^{+0.11}_{-0.09}$	11.089				
91987762*	HIP 47288	$1894.25391^{+0.0004}_{-0.0004}$	$45.522156^{+3.4e-05}_{-3.4e-05}$	$0.05551^{+0.00083}_{-0.00069}$	$9.76^{+0.50}_{-0.50}$	$0.772^{+0.016}_{-0.018}$	$4.35^{+0.03}_{-0.03}$	7.87		NRES (4)	Gemini	TOI 1898.01 (1)
95768667	TYC 1434-00331-1	$1918.3318^{+0.0079}_{-0.0093}$	$26.9^{+72.3}_{-12.4}(s)$	$0.0282^{+0.0031}_{-0.0022}$	$3.54^{+0.43}_{-0.32}$	$0.48^{+0.35}_{-0.33}$	$5.40^{+0.76}_{-0.64}$	10.318				
9707944	TYC 5831-00228-1	$2107.4855^{+0.0044}_{-0.0053}$	$19.73575^{+0.00036}_{-0.00039}$	$0.0366^{+0.0028}_{-0.002}$	$4.15^{+0.38}_{-0.31}$	$0.43^{+0.32}_{-0.30}$	$5.09^{+0.38}_{-0.23}$	11.211				
98794593	HIP 14073	$2169.3177^{+0.0406}_{-0.0071}$	$53.5^{+27.5}_{-30.5}(s)$	$0.01995^{+0.00713}_{-0.00063}$	$3.40^{+1.16}_{-0.20}$	$0.34^{+0.24}_{-0.21}$	$9.80^{+0.32}_{-9.47}$	8.22		CHIRON (13);NRES (1)		

Table 3.1: Properties of PHT candidates (continued)

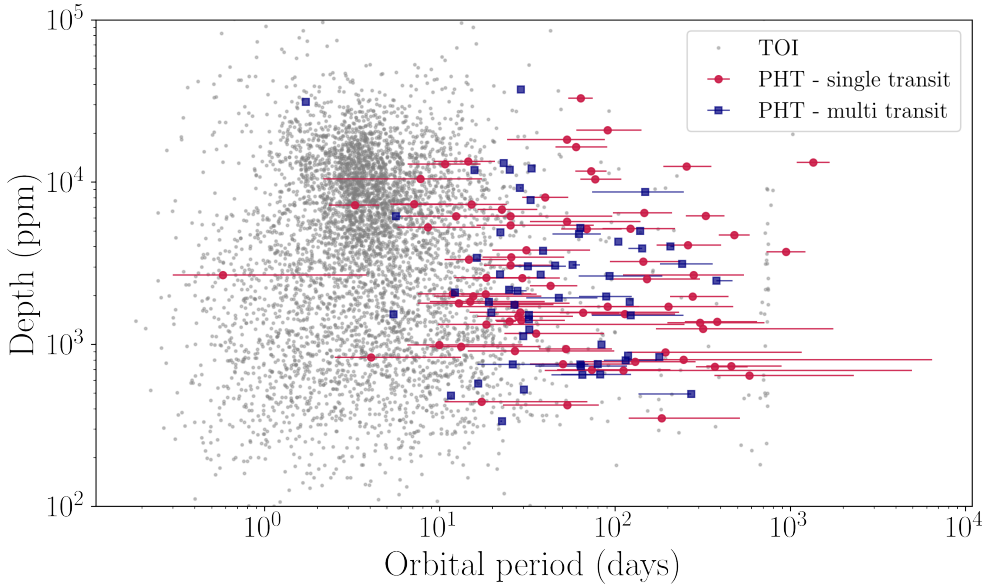

Figure 3.9: The properties of the PHT single-transit (magenta circles) and multi-transit (navy squares) candidates compared to the properties TOIs (grey points). PHT candidates which have been ruled out as false positives are not included. All parameters (listed in Table 3.9) were extracted using `pyaneti` modelling.

3.2.1 Planet candidate analysis

As discussed in Chapter 1, the majority of TOIs (90.5%) have orbital periods shorter than 15 days due to the automated pipeline's two transit detection threshold combined with the observing strategy of *TESS*. As visual vetting does not impose a two-transit detection limit, the candidates outlined in this chapter are helping to populate the relatively under-explored long-period region of parameter space. This is highlighted in Figure 3.9, which shows the transit depths vs the orbital periods of the PHT single-transit candidates (magenta circles) and the multi-transit candidates (navy squares) compared to the TOIs (grey points). PHT candidates which have been ruled out as false positives are not included. Values of the orbital periods and transit depths were obtained via transit modelling using `pyaneti` (see Section 2.3). The orbital period of single-transit events (or multiple transits separated by large data gaps) are poorly constrained, which is reflected by the large uncertainties. Figure 3.9 also highlights that the pipeline is able to recover smaller planets ($R > 0.55$ $R_\oplus$) than we can detect with visual vetting ($R > 1.85$ $R_\oplus$), as was previously shown in Figure 3.4.

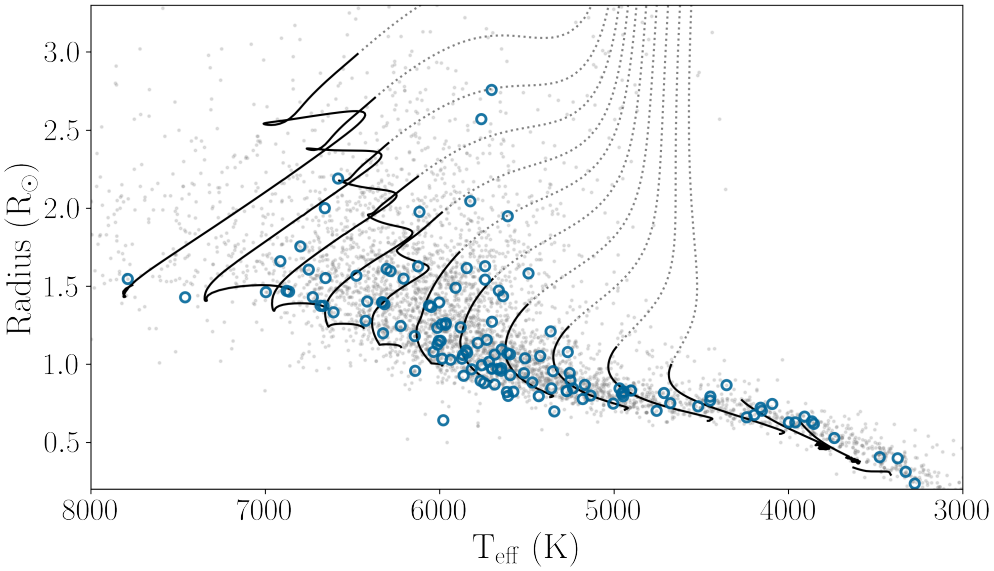

Figure 3.10: Stellar evolution tracks showing main sequence (solid black lines) and post-main sequence (dashed grey lines) MIST stellar evolution for stellar masses ranging from 0.3 to 1.6 $M_\odot$ in steps of 0.1 $M_\odot$. The grey points show the TOIs and the blue circles show the PHT candidates.

The PHT candidates were further compared to the TOIs in terms of the properties of their host stars. Figure 3.10 shows the effective temperature and stellar radii as taken from the TIC (Stassun et al., 2018), for TOIs (grey points) and the PHT candidates (blue circles). The solid and dashed lines indicate the main sequence and post-main sequence MIST stellar evolutionary tracks (Choi et al., 2016), respectively, for stellar masses ranging from 0.3 to 1.6 $M_\odot$ in steps of 0.1 $M_\odot$. The figure shows that around 5% of the host stars are in the process of, or have recently evolved off the main sequence. The stellar track models assume solar metalicity, no stellar rotation, and no additional internal mixing.

Finally, ground-based follow-up spectroscopy has revealed that 7 of the PHT candidates listed in Table 3.1 are astrophysical false positives. As the follow-up campaign of the targets is still underway, the true false-positive rate of the candidates will be assessed in the future, once the true nature of more of the candidates have been independently verified.

3.3 Stellar systems

In addition to finding planetary candidates, citizen science allows for the identification of interesting stellar systems and astrophysical phenomena as a by-product of the planet search. For example, visual vetting lends itself for the identification of eclipsing binaries (EBs) or for light curves that exhibit multiple transit-like signals, possibly as a result of a multi-stellar system. The signals from these astrophysical phenomena are often classified as planetary signals or flagged on the PHT discussion forums by citizen scientists. In this section I discuss the identification and classification of EBs, which are published in the *TESS* EB catalogue (Prša et al., 2022), and the identification of candidate multi-stellar systems.

3.3.1 Eclipsing binaries

Within the first two years of the *TESS* mission, PHT citizen scientists identified a total of 6699 EB candidates. Out of these, 2720 EB candidates were within the top 500 highest ranked planet candidates per sector, allowing me to identify them as EBs as I sorted them into '*keep for further analysis*', '*eclipsing binary*', and '*discard*' (see Section 2.1.9). The remaining candidates were identified via the PHT discussion forums. Within the first two years of the *TESS* mission, the tags 'EB', 'eclipsing binary', or similar versions thereof, had appeared over 46 200 times on the PHT discussion forums, corresponding to 5759 individual *TESS* targets. All of these targets were considered as potential EB candidates and were kept for further vetting. A total of 1780 candidates were discovered via both of these methods, bringing the total number of EB candidates identified via PHT to 6699. All of these EB candidates were subsequently visually inspected by members of the *TESS* eclipsing binary working group and vetted for obvious false positives (e.g., where the signal is caused by pulsations instead of eclipses). A total of 4584 EBs were kept following this stage of vetting and are included in the *TESS* eclipsing binary catalogue.[2] I generated a vetting report for all of these candidates using a version

[2] http://tessebs.villanova.edu/

of `LATTE` that I specifically altered to cater to the vetting of EBs. For full details of the analysis and results see Prša et al. (2022).

3.3.2 Multi-star systems

In addition to the EBs, I investigated all light curves that were flagged as possible multi-stellar systems via the PHT discussion boards (i.e. light curves that contain more than one set of periodic signals). I generated `LATTE` data validation reports for all multi-stellar candidates in order to assess the nature of the signal and to determine whether the signals are on target. Additionally, I subjected these systems to an iterative signal removal process, where I phase-folded the light curve on the dominant orbital period, binned the light curve into between 200 – 500 phase bins, created an interpolation model, and then subtracted said signal in order to evaluate the individual transit signals. The period of each signal, as listed in Table 3.2, was determined using a Lomb-Scargle frequency analysis followed by visual inspection of the phase folded light curve. The periods were manually fine-tuned and the final values and corresponding uncertainties determined by eye.

Due to the large *TESS* pixels, blends are expected to be common. I searched for blends by generating phase folded light curves for each pixel around the source of the target in order to better locate the source of each signal. Shifts in the x and y centroid positions were also found to be good indicators of visually separated sources. Nearby sources with a magnitude difference greater than $\Delta V = 5$ mag were ruled out as possible contaminators, as the light contribution of such a star would be a factor of 100 times less than that of the target and thus any signal from that star too small to significantly contaminate the target light curve. I consider a candidate to be a confirmed blend when the centroids are separated by more than one *TESS* pixel, as this corresponds to an angular separation > 21 arcsec meaning that the systems are highly unlikely to be gravitationally bound. Systems where the signal appears to be coming from the same *TESS* pixel and that show no clear centroid shifts are considered to be *candidate* multiple systems. I note that

blends are still possible, however, without further investigation I cannot conclusively rule these out as possible multi-stellar systems.

All of the systems are summarised in Table 3.2. Out of the 73 systems, 11 are confirmed multiple systems which have either been published or are being prepared for publication by the teams listed in the table; 14 are visually separated eclipsing binaries (confirmed blends); and 48 are candidate multiple systems. The identification of eclipse timing variations (Section 6.3.3) or additional observations (e.g., RV monitoring) will be required to determine whether these candidate multi-stellar systems are in fact gravitationally bound or photometric blends due to a line of sight happenstance.

3.4 Conclusions

In this Chapter I presented the results from the analysis of the first three years of *TESS* data. I outline the properties of 139 PHT planet candidates, a number of which are being followed-up by other teams of astronomers, such as TOI-1899 b (TIC 172370679 b) which was recently confirmed to be a warm Jupiter transiting an M-dwarf (Cañas et al., 2020); TOI-2180 b (TIC 298663873 b) which is a giant planet on an eccentric 261-day orbit (Dalba et al., 2022); TOI-1793 b (TIC 304142124 b) which is a Neptune sized planet on a 55-day orbit (Díaz et al., 2020); and TOI-4562 b (TIC 349576261 b) which is a highly eccentric ($e \sim 0.77$) cool Jupiter on 225-day orbit around a young star (Heitzmann et al. *in prep*). These systems highlight the scientific value of the PHT identified candidates as well as the exoplanet community's interest in them. The majority of the remaining PHT planet candidates require further follow-up observations to confirm their planetary nature.

Throughout this chapter, I also showed that large scale visual vetting can complement the findings from the major *TESS* pipeline by identifying longer-period planets that may only exhibit a single transit event in their light curve (or multiple events separated by large data gaps), as well as in finding signals that are aperiodic or embedded in a strong varying stellar signal. These longer-period planets, in particular, are integral to our understanding of how planet systems form and evolve,

Table 3.2: Multi-stellar system candidates.

TIC	Period (days)	Comment
4783276	1.08446 ± 0.00001	Candidate multiple system
	0.6974 ± 0.0003	
13968858	3.4850 ± 0.001	Candidate multiple system
	1.4380 ± 0.001	
35655828	8.073 ± 0.01	Confirmed blend
	1.220 ± 0.001	
43250275	4.1056 ± 0.0001	Confirmed blend
	0.4787 ± 0.0001	
45802115	3.40651 ± 0.0004	Candidate multiple system
	–	
63291675	8.099 ± 0.003	Confirmed blend
	1.4635 ± 0.0005	
63459761	4.3630 ± 0.003	Candidate multiple system
	4.235 ± 0.005	
73296637	1.84399 ± 0.00005	Candidate multiple system
	1.48374 ± 0.0005	
78568780	2.8884 ± 0.0001	Candidate multiple system
	–	
79140936	3.54388 ± 0.00012	Candidate multiple system
	30.92 ± 0.01	
80914862	1.96729 ± 0.00004	Candidate multiple system
	18.666 ± 0.001	
88921674	6.652 ± 0.001	Confirmed blend
	–	
97356407	0.7667 ± 0.0001	Candidate multiple system
	8.099 ± 0.005	
104909909	1.3060 ± 0.0001	Candidate multiple system
	2.5750 ± 0.003	
115980439	4.615 ± 0.002	Confirmed blend
	0.742 ± 0.005	
120362128	3.286 ± 0.002	Candidate multiple system
	–	
	–	
121945407	0.9056768 ± 0.00000002	Confirmed multiple (Borkovits et al., 2013)
	45.4711 ± 0.00002	
122275115	–	Candidate multiple system
	–	
139650665	2.09174 ± 0.00001	Candidate multiple system
	10.632 ± 0.001	
146810480	1.46854 ± 0.00001	Candidate multiple system
	0.54494 ± 0.0001	
150413414	14.969 ± 0.001	Candidate multiple system
	–	
168789840	1.570101	Confirmed multiple (Powell et al., 2021)
	8.217173	
	1.305934	
178953404	3.182 ± 0.004	Candidate multiple system
	–	
183907025	6.962 ± 0.001	Confirmed blend
	1.6213 ± 0.0003	
190895730	0.9182 ± 0.00001	Candidate multiple system
	1.3176 ± 0.0001	
204698586	0.84387 ± 0.00001	Candidate multiple system
	11.013 ± 0.003	
209409435	5.717471 ± 27	Confirmed multiple (Borkovits et al., 2020b)
	121.8723 ± 0.001	
229804573	1.4641 ± 0.0005	Candidate multiple system
	0.5283 ± 0.0001	
232087348	2.6143 ± 0.0002	Candidate multiple system
	9.654 ± 0.004	
239761286	0.75007 ± 0.00001	Candidate multiple (Zasche et al., 2019)
	–	
239808300	1.20937 ± 0.00002	Confirmed multiple (Cagaš & Pejcha, 2012)
	0.80693 ± 0.00002	
252403752	–	Candidate multiple system
	–	
257776944	1.2255 ± 0.0001	Candidate multiple system
	3.3045 ± 0.0001	
258837989	0.8870 ± 0.001	Candidate multiple system
	3.0730 ± 0.001	
266958963	1.5753 ± 0.0002	Candidate multiple system
	2.3685 ± 0.0001	
271186951	2.0933 ± 0.0005	Candidate multiple system
	1.732 ± 0.001	
271204362	3.26551 ± 0.00004	Candidate multiple system
	2.94289 ± 0.00003	
274791367	1.20712 ± 0.00001	Candidate multiple system
	14.312 ± 0.002	
278825952	4.781023 ± 0.000002	Confirmed multiple (Mitnyan et al., 2020)
	235.5499 ± 0.006	
278956474	5.488068 ± 0.000016	Confirmed multiple (Rowden et al., 2020)
	5.674256 ± -0.000030	
284482112	4.10838 ± 0.00001	Candidate multiple (Zasche et al., 2019)
	2.3378 ± 0.0001	
284925600	1.24571 ± 0.00001	Confirmed blend
	0.31828 ± 0.00001	
293954660	2.814 ± 0.001	Confirmed blend
	4.904 ± 0.03	
299377596	4.5137 ± 0.0001	Confirmed blend
	–	
309025182	3.36096 ± 0.00004	Candidate multiple system
	2.80288 ± 0.00003	
311838201	2.1347 ± 0.0001	Candidate multiple system
	2.497 ± 0.001	
312353805	4.951 ± 0.003	Confirmed blend
	12.89 ± 0.01	
312353807	$4.95118 \pm [0.00014$	Confirmed blend
	25.8 ± 0.01	
318210930	$1.3055432 \pm 0.000000033$	Confirmed multiple (Koo et al., 2014)
	$0.22771622 \pm 0.0000000035$	
336434532	3.888 ± 0.002	Confirmed blend
	0.949 ± 0.003	
350230129	0.7844 ± 0.0001	Candidate multiple system
	17.75 ± 0.02	
350622185	1.1686 ± 0.0001	Candidate multiple system
	5.2410 ± 0.0005	
370440624	2.23528 ± 0.00006	Candidate multiple system
	8.704 ± 0.001	
375422201	9.9649 ± 0.001	Candidate multiple system
	4.0750 ± 0.001	
376606423	0.8547 ± 0.0002	Candidate multiple system
	–	
378275445	7.65455 ± 0.0005	Candidate multiple system
	3.61 ± 0.02	
378584712	0.28938 ± 0.00001	Candidate multiple system
	35.666 ± 0.001	
394177355	94.22454 ± 0.00040	Confirmed multiple (Hełminiak et al., 2017)
	8.6530941 ± 0.0000016	
	1.5222468 ± 0.0000025	
	$1.43420486 \pm 0.00000012$	
399492452	14.96 ± 0.01	Candidate multiple system
	21.96 ± 0.01	
399947158	7.1464 ± 0.0009	Candidate multiple system
	–	
400166377	2.21918 ± 0.00009	Candidate multiple system
	35.23 ± 0.01	
408147984	1.07259 ± 0.00003	Candidate multiple system
	3.805 ± 0.005	
424508303	2.0832649 ± 0.0000029	Confirmed multiple (Zasche & Uhlař, 2016)
	1.4200401 ± 0.0000042	
438226195	–	Candidate multiple system
	–	
439511833	11.0473 ± 0.0005	Candidate multiple system
	6.5948 ± 0.0007	
441794509	4.6687 ± 0.0002	Candidate multiple system
	14.785 ± 0.002	
441794509	4.6687 ± 0.0002	Candidate multiple system
	14.785 ± 0.002	
451614854	2.01114 ± 0.00005	Confirmed blend
	19.44 ± 0.006	
454140642	13.6239 ± 0.00016	Confirmed multiple (Kostov et al., 2021)
	10.3928 ± 0.00008	
457872192	1.8766 ± 0.0001	Confirmed blend
	38.152 ± 0.005	
459841654	4.1507 ± 0.001	Candidate multiple system
	–	
461614217	2.28805 ± 0.00008	Candidate multiple system
	9.3671 ± 0.0001	
470710327	9.9733 ± 0.0001	Confirmed multiple (Eisner et al., 2022)
	1.104686 ± 0.00001	

as they allow us to investigate the evolution of planets that are farther away from their host star and therefore less dependent on stellar radiation. As such the planets identified via PHT can help inform future planet populations studies.

Furthermore, the longer periods of the PHT planet candidates ($\gtrsim$27 days) are more typical of the *Kepler* sample (25% of confirmed *Kepler* planets listed in the NASA Exoplanet Archive have periods greater than 27 days). However, the *Kepler* planets are considerably fainter, and thus less amenable to ground-based follow-up or atmospheric characterisation from space (e.g., with CHEOPS or JWST). Thus PHT helps to bridge the parameter spaces covered by the *Kepler* and *TESS* missions, by identifying longer-period planet candidates around bright, nearby stars, for which we can ultimately obtain precise planetary mass estimates. In particular, identification of this PHT sample provides follow-up targets to investigate the dependence of photo-evaporation on the mass of planets as well as on the planet radius, and will help our understanding of the photo-evaporation valley at longer orbital periods (Owen & Wu, 2013, as discussed in Chapter 1).

I note that while automated pipelines specifically designed to identify single-transit events in the *TESS* data exist (e.g., Gill et al., 2020), neither their methodology nor the full list of their findings are yet publicly available and thus I am unable to compare our results to their findings.

Finally, I showed that in addition to searching for signals due to transiting exoplanets, PHT provides a platform that can be used to identify other stellar phenomena that may otherwise be difficult to identify with automated searches. Such phenomena, including eclipsing binaries, multi-stellar systems, dwarf novae, and stellar flares are often mentioned on the PHT discussion forums where volunteers can use searchable hashtags and comments to bring these systems to the attention of other citizen scientists as well as the PHT science team. To date, I have verified and published one of these systems, as discussed in Chapter 6. These discoveries highlight the importance of visual vetting to the fields of exoplanets as well as the field of stellar astrophysics.

I sometimes find, and I am sure you know the feeling,
that I simply have too many thoughts and memories
crammed into my mind.

— A. P. W. B. Dumbledore, 1994

CHAPTER 4

TOI 813: A SUBGIANT HOSTING A TRANSITING SATURN-SIZED PLANET

In addition to delivering an ensemble of new planet candidates, PHT has identified scientifically valuable planetary systems that lend themselves to further investigation. In the next two chapters, I will discuss the detection, validation, and characterisation of two planetary systems that were identified via the PHT platform.

4.1 Context

The majority of a star's life is spent on the main sequence. However, all stars eventually evolve off this main sequence to first become subgiants, and then red giant stars. These evolved stars are not normally prime targets for transit surveys, as their large radii make the transits shallower (Equation 1.2) and longer (Equations 2.14 and 2.15) and thus harder to detect. Furthermore, evolved stars are comparatively scarce in the Solar neighbourhood as the subgiant and giant phases of stellar evolution are short-lived. Consequently, relatively few transiting planets are known around evolved stars. As of 19 January 2022, there are 109 confirmed planets with mass measurement to better than 30% accuracy transiting evolved stars listed in the NASA Exoplanet Archive, only 14 of which have orbital periods longer than 30 days.

In this chapter, I discuss the detection and validation of TOI 813 b, a Saturn-sized planet orbiting around a bright subgiant star ($R_\star = 1.94$ R$_\odot$, $M_\star = 1.32$ M$_\odot$). The planet has an orbital period of ~ 84 days, which made it the longest-period validated planet found by *TESS* at the time of publication in January 2020. As such, the discovery of this long-period planet orbiting around a subgiant adds to the relatively small but important sample of known planets around evolved stars. Furthermore, the stellar brightness ($V = 10.36$ mag), together with the RV semi-amplitude of 5.96 ± 1.8 m s^{-1}, makes this target one of the few long-period transiting planets for which a precise mass measurement through RV observations is possible.

The discovery and statistical validation of this system was originally published as Eisner et al. (2020b). Since then, the target has been observed by *TESS* in an additional 13 sectors, and additional ground-based follow-up observations have been obtained by another team. Throughout the remainder of this chapter I will present the discovery and analysis of TOI 813 b using all of the *TESS* and ground-based data that is available as of 19 January 2022.

4.2 Discovery of TOI 813 b in the *TESS* data

4.2.1 *TESS* data

TOI 813 (TIC 55525572; Stassun et al., 2019) is located at high ecliptic latitude (near the galactic pole) and was observed in Sectors 1 – 13 (except in Sector 7) and in Sectors 27 – 39. However, two-minute cadence observations are only available from Sector 4 onward. Prior to that it was observed in the Full Frame Images (FFIs) collected at a lower cadence of 30 minutes. The two-minute and 30-minute cadence light curves used throughout the remainder of this chapter were produced with the SPOC and SPOC-FFI pipelines, respectively.[1]

In total, the light curve of TOI 813 consists of 338 668 flux measurements obtained between barycentric *TESS* Julian Date (BJD - 2457000) 1326.9694 and 2389.71706.

[1]The two-minute cadence light curves were processed and reduces by versions 5.0.19 to 5.0.35 (depending on the sector) of the SPOC pipeline. The FFI light curves were produced with version 5.0.15 of the SPOC-FFI pipeline.

The two-minute cadence light curves have an RMS CDPP on a 2-hour time scale of 154 ppm averaged across all of the sectors (see Section 2.1.2), while the 30-minute cadence data have a typical RMS CDPP scatter on a 2-hour time scale of 139 ppm. Compared to other *TESS* targets with comparable magnitudes, this RMS CDPP is relatively low,[2] suggesting that the star is relatively quiet.

4.2.2 Discovery of TOI 813 b

TOI 813 b was initially flagged as a single transit event by citizen scientists when the Sector 5 data was being classified on PHT. A second transit event, separated by $\sim$84 days, was later identified in Sector 8 and brought to my attention via the PHT discussion forum.[3] Since then, a total of 8 transit events have been recovered in Sectors 2, 5, 8, 11, 27, 30, 33, and 39. The full light curve is shown in Figure 4.1 and the individual transits in Figure 4.2.

At the time of the PHT discovery, TOI 813 was not listed as a TOI, nor did it have any Threshold Crossing Events (TCEs). In other words, it was not detected by either the SPOC transit search pipeline or the Quick Look Pipeline (QLP; Fausnaugh et al., 2018). The reason why it was missed by these searches is that the planet never exhibits more than one transit in a given sector and the SPOC and QLP pipelines both require at least two transits for detection (see Section 1.5.3). At the time of detection, neither of these two pipelines had been run on multi-sector data. However, concurrently to the citizen science discovery, an independent team identified TOI 813 b via a manual survey of the *TESS* data using the LcTools software (Kipping et al., 2015), following the method described by Rappaport et al. (2019).

The TOI team released a TCE report on TIC 55525572 on 21 June 2019, once the target was flagged in the SPOC multi-sector planet search in *TESS* Sectors 1–9.[4] The SPOC pipeline also detected the transits of TOI 813 b in Sectors 5, 8,

[2]The RMS CDPP values are reported for all two-minute cadence targets in the *TESS* Data Release Notes (for each sector). See e.g., https://tasoc.dk/docs/release_notes/tess_sector_08_drn10_v02.pdf

[3]https://www.zooniverse.org/projects/nora-dot-eisner/planet-hunters-tess/talk/2112/978055

[4]https://archive.stsci.edu/missions/tess/doc/tess_drn/tess_multisector_01_09_drn15_v03.pdf

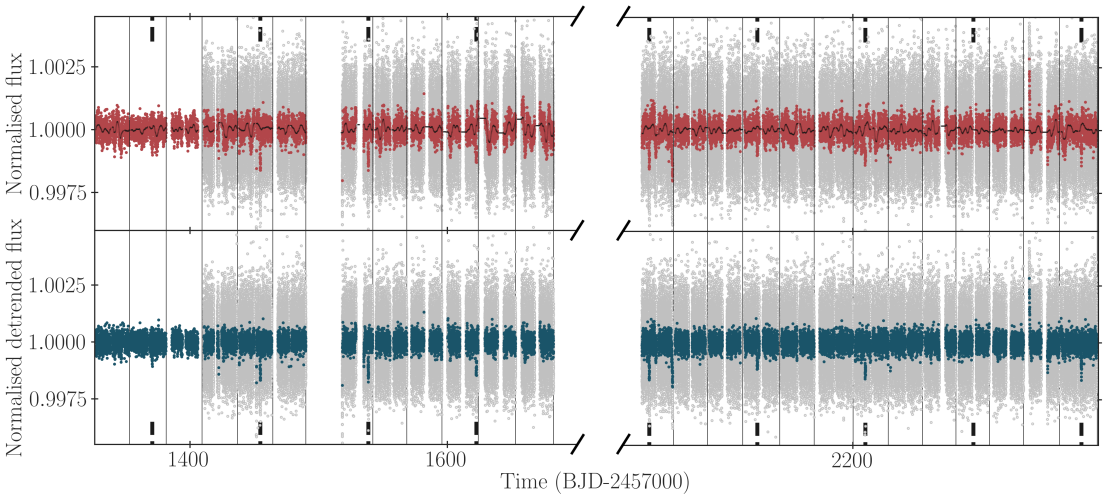

Figure 4.1: *TESS* light curve for TOI 813 for Sectors 1 – 13, except Sector 7, and Sectors 27 – 39. The top panel shows the PDC light curve at a two-minute sampling (grey) and a 30-minute sampling (dark red). The black points shows the best fit model used for detrending. The bottom panel shows the detrended light curve, used in the BLS search (see Section 4.5.3), also at a two-minute (grey) and 30-minute (teal blue) sampling. The times of transits are indicated by the short black dashed vertical lines and the end of each sector is depicted by a solid light grey line. The dashed lines on the x-axis show a split in the axis.

and 11 in the multi-sector transit search of Sectors 1 – 13 (performed on 7 August 2019). Their reported values for the orbital parameters and planet radius are consistent with those derived in Section 4.4.4.[5]

4.2.3 Light curve based vetting checks

I carried out vetting tests on the *TESS* data in order to rule out a number of instrumental and astrophysical false positive scenarios. These include all of the vetting tests presented as part of the `LATTE` vetting suite (see Section 2.2), none of which indicated that TOI 813 b is a false positive. All additional tests that were carried out are as follows:

Odd and even transit comparison I compared the depths, transit duration, and shape of the odd- and even-numbered transits. Slight differences in these would indicate that the events are caused by a near-equal mass eclipsing

[5] https://mast.stsci.edu/api/v0.1/Download/file/?uri=mast:TESS/product/tess2018206190142-s0001-s0013-0000000055525572-00226_dvr.pdf

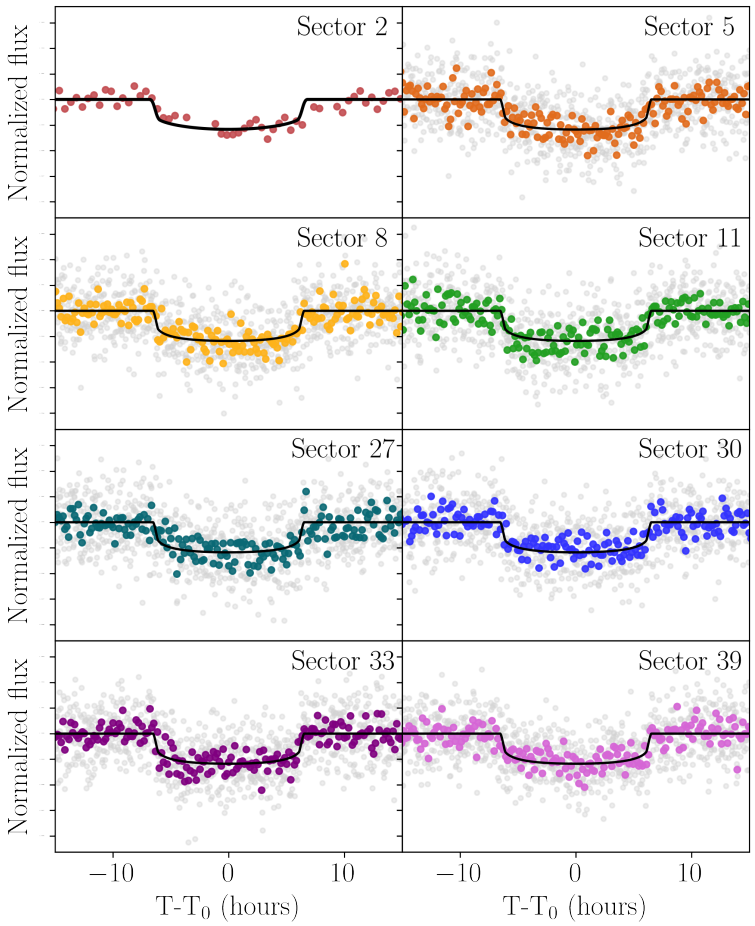

Figure 4.2: The first 8 transits of TOI 813 b in the *TESS* data. The Sector 2 light curve was extracted from the FFIs whereas the other sectors are two-minute cadence observations. They grey points shows the two-minute cadence data while the solid colour points show the data at a 30-minute cadence.

binary. I did this by independently modelling the phase-folded odd- and even-numbered transits using the `pyaneti` suite (as described in Section 4.4.4). The depths, transit duration, and shape (quantified by the duration of ingress and egress) are consistent with one another to within $0.5\,\sigma$.

Secondary eclipse search I searched for a secondary eclipse (or occultation), which might indicate that the transits are caused by a small but self-luminous companion, such as a brown dwarf or low-mass star. This was done by

conducting a grid search for the deepest out-of-transit signal in the phase-folded light curve (see e.g., Rowe et al., 2015; Morton et al., 2016). This resulted in an upper limit on the depth of any secondary eclipse at the level of 230 ppm, which is comparable to the RMS scatter in the data. As such there is no evidence for a detectable secondary eclipse.

Pixel-level centroid analysis A blend with a background eclipsing binary would result in a change in the spatial distribution of the flux within the target aperture during the transits. I checked for this by comparing the observed position of the target during, and immediately before and after each transit, but no significant differences were found. To quantify this, I used a two-sided Kolmogorov–Smirnov (K–S) test to see whether the flux-weighted centroid positions in- and out-of-transit are drawn from the same distribution or not. In brief, the K–S test looks for differences in the shape of the distribution between two data sets by calculating the maximum absolute difference between the cumulative distributions functions (CDF) of the data sets, where the CDF represents the probability of a variable in the data set being equal to or less than a specific value. This test showed that for all of the transits the detrended x- and y-centroid positions (detrended with a second order polynomial around the time of the transit events) did not differ significantly in-transit compared to out-of-transit, with p-values ranging from 0.16 to 0.98 (where statistically different is defined as $p < 0.05$). As such, this test shows no evidence that the signal is caused by a background eclipsing binary.

Nearby companion stars I searched for evidence of nearby stars by querying all entries in the early *Gaia* Data Release 3 catalogue (eDR3; Gaia Collaboration et al., 2021) within 180 arcsec of TOI 813. I found there to be four stars within this radius with magnitudes brighter than $V = 16$ mag, as shown in Figure 4.3. In order to rule out these companion stars as the source of the signal, I calculated the magnitude difference between TOI 813 and the faintest companion star that could plausibly cause the observed transit depth and

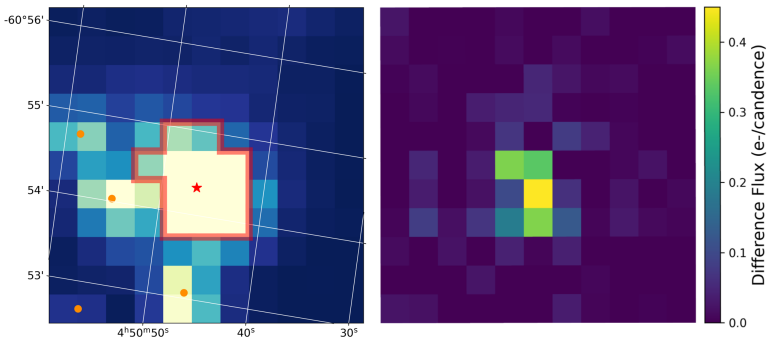

Figure 4.3: *Left*: *TESS* image in the vicinity of TOI 813 throughout Sector 5. The red outline shows the aperture mask used to extract the light curve. Neighbouring stars brighter than $V = 16$ mag within 180 arcmins of TOI 813 are depicted by orange dots. *Right*: Difference in/out of transit image during Sector 5.

shape. I used the equations presented in Vanderburg et al. (2019) and the transit parameters derived from the transit modelling (see Section 4.4.4) and found the maximum magnitude difference to be 0.19 mag. This allows me to confidently rule out all four of the surrounding stars as the cause of the transit signal (closest and brightest star is located at an angular separation of ~ 19 arcsec with $V = 14.64$ mag). I also note that the SPOC pipeline's difference imaging analysis showed that none of these stars are located within the $3\,\sigma$ confusion region around the target, which is the region surrounding the star where other stars cannot be confidently ruled out as the source of the signal. As the SPOC pipeline (Jenkins et al., 2016) accounts for the contamination of the aperture by the neighbouring stars, I did not correct the measured transit depth for the effect of the light from the nearby companion star.

The above tests, combined with the vetting tests provided in the `LATTE` vetting suite, can help rule out many but not all of the plausible astrophysical false positive scenarios. While I can state with confidence that none of the nearby stars brighter than $V = 16$ (shown in Figure 4.3) are the source of the transits, fainter contaminants located nearer the main target cannot be ruled out at this stage. However, these vetting tests increased my confidence in the planetary nature of the companion sufficiently to report the candidate on the ExoFOP website as a cTOI[6] and to

[6] https://exofop.ipac.caltech.edu/tess/target.php?id=55525572

motivate ground-based follow-up observations (see Section 4.3).

I note that TOI 813 b also independently passed all of the SPOC pipeline Data Validation diagnostic tests in the multi-sector 1 – 13 run. These tests included the odd and even transit comparison test, the weak secondary eclipse search test, the ghost diagnostic test (which is similar to `LATTE`'s increasing aperture size test), and the statistical bootstrap test, which looks at the propensity of other transit-like features in the light curve to cause happenstance false positives (Twicken et al., 2018; Li et al., 2019).

4.3 Ground-based observations

In order to better characterise the host star and to confirm that the transit-like signals are not caused by an eclipsing binary, my collaborators and I obtained one reconnaissance and one high resolution spectrum, as well as speckle imaging. Furthermore, we obtained four spectra using the CHIRON spectrograph in order to place an upper mass limit on the planet. This work, which was published as Eisner et al. (2020b), allowed us to confirm the planetary nature of TOI 813 b. Since the initial publication of the discovery paper, archival RV data of the target, obtained with the HARPS spectrograph on the ESO 3.6-m telescope, have become publicly available. The analysis presented throughout the remainder of this chapter makes use of the reconnaissance and HARPS data.

4.3.1 Reconnaissance spectra

We obtained a moderate spectral resolution ($R = 7000$) spectrum of TOI 813 with the Wide Field Spectrograph instrument on the Australian National University (ANU) 2.3-m telescope (Dopita et al., 2007) on the night of 30 April 2019. Three 120-second exposures were obtained with the U7000 (S/N $\sim$ 400) and R7000 (S/N $\sim$ 700) gratings and the data were reduced using the pyWiFeS data reduction pipeline version 0.7.4 (Childress et al., 2014). Visual inspection of the data did not show any signs of spectral lines belonging to more than one star, which would have indicated that the star is a spectroscopic binary.

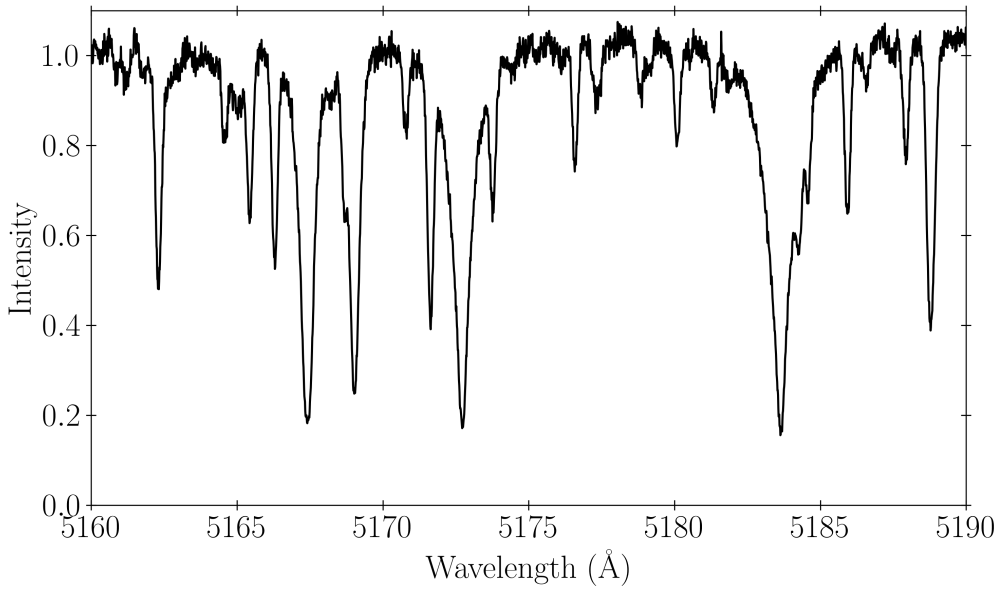

Figure 4.4: Section of the high resolution HARPS spectrum.

4.3.2 High resolution spectra

In addition to the reconnaissance spectrum, we acquired a high-resolution ($R \approx 115\,000$) spectrum with the High Accuracy Radial velocity Planet Searcher (HARPS; Mayor et al., 2003) spectrograph on the ESO 3.6-m telescope at La Silla observatory, Chile, on 14 July 2019. An exposure time of 1800 seconds was used, resulting in a S/N per pixel of ~ 60 at 5500 Å. The spectrum, which was reduced and extracted using the standard HARPS Data Reduction Software (DRS; Baranne et al., 1996), displayed no obvious signs of binarity (i.e. there was only evidence of spectral lines from a single star), as shown in Figure 4.4. This spectrum was used to obtain the final estimate of the stellar parameters, as described in Section 4.4.1.

4.3.3 RV monitoring

Since the publication of the validated planet system (Eisner et al., 2020b), 17 archival ESO 3.6-m/HARPS observations, obtained between 5 October 2019 and 24 November 2020 (ID 0104.C-0413; PI: Brahm), have become publicly available.[7]

[7]I note that I am not involved with this ongoing observing program and do not plan on publishing these results, however, I include them here for completeness.

Table 4.1: Radial velocity measurements.

Time (BJD - 2457000)	RV ($m\,s^{-1}$)	σ_{RV} ($m\,s^{-1}$)	S/N
1678.933689*	2.34	1.87	62.45
1761.743732	13.57	2.97	46.5
1765.727748	8.87	5.24	29.2
1766.737111	15.85	1.87	62.55
1775.841291	11.73	3.77	33.5
1803.667986	0.73	2.28	63.3
1805.741409	1.59	1.73	61.45
1811.785602	-4.33	4.22	34.85
1813.724096	9.30	3.14	49.15
1832.686259	4.65	3.61	37.9
1833.720328	4.92	2.34	53.6
1837.773400	2.13	2.99	43.15
1851.680739	17.55	2.37	61.35
1868.702455	2.53	2.43	52.55
1877.595372	1.28	3.51	40.2
1884.643436	-7.92	3.42	35.95
1897.588449	1.893	2.75	49.05
2177.682721	14.68	2.30	56.2

Note – RVs extracted using the SERVAL pipeline. The S/N ratios are calculated at 550 nm.
* Indicates the spectrum that was used to extract the stellar parameters (Section 4.4.1)

Each of these observations has an exposure time of 1200 seconds, resulting in S/N ratios per pixel in the range of $\sim$29.2 and $\sim$63, with a mean of $\sim$48.5 at 5500 Å.

All of the 18 HARPS spectra (including our initial high resolution spectrum) were reduced and RV extracted using the publicly available SERVAL pipeline (Zechmeister et al., 2018). As discussed in Section 2.4.2, SERVAL creates a high S/N template by combining all of the available spectra, which is then used to compute RV variations using a least-squares fit. This reduction and extraction pipeline was used as the SERVAL RVs have been shown to be more precise, by a few percent, than the standard HARPS DRS RVs (Trifonov et al., 2020). The extracted SERVAL RV measurements (see Table 4.1) were used to constrain the mass of TOI 813 b, as discussed in Section 4.4.4.

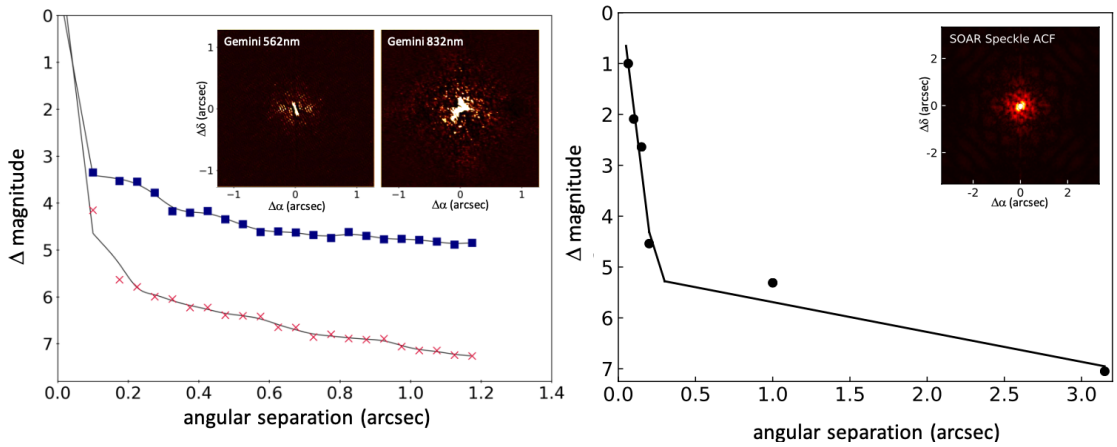

Figure 4.5: Contrast curves showing the $5\,\sigma$ detection sensitivity and speckle auto-correlation functions (insets). The left hand panel shows the data obtained using Zorro on Gemini with the filters centred on 562 nm (blue squares) and 832 nm (red crosses) and the right hand panel shows the observations obtained with the HRCam speckle imager on SOAR in the *I*-filter.

4.3.4 High resolution imaging

Speckle imaging was performed using the Zorro instrument on the 8.1-m Gemini South telescope (Howell et al., 2011; Matson et al., 2019) in order to search for close companions and to quantify their contribution to the *TESS* photometric aperture (see Section 4.3.4). Observations were obtained on the night of 16 July 2019 using simultaneous two-color diffraction-limited optical imaging with 60-millisecond exposures in sets of 1000 frames. No companions were detected within 1.17 arcsec of the target at the $4-5\,\Delta$mag limit at 562 nm and $5-7\,\Delta$mag limit at 832 nm.

Concurrently, a search for nearby sources to TOI 813 was carried out with speckle images obtained using the HRCam speckle imager on the 4.1-m Southern Astrophysical Research (SOAR; Tokovinin, 2018) telescope at Cerro Pachon Observatory. The *I*-band observations, obtained as part of the SOAR *TESS* Survey (Ziegler et al., 2020) on the night of the 14 July 2019, show no evidence of any faint companions within 3 arcsec of TOI 813 up to a magnitude difference of 7 mag. The $5\,\sigma$ detection sensitivity and the speckle auto-correlation function from the SOAR and Gemini observation are shown in Figure 4.5.

4.4 Data Analysis

4.4.1 Stellar parameters

The stellar parameters provided in the TIC (Stassun et al., 2019) are based on broad-band photometry, meaning that their precision is limited. As such, the high-resolution HARPS spectrum was used to extract fundamental stellar parameters.

First, the Spectroscopy Made Easy (`SME`; Valenti & Piskunov, 1996; Piskunov & Valenti, 2017) code was used, which works by calculating synthetic stellar spectra from grids of detailed atmosphere models and fitting them to the observations with a chi-square-minimisation approach. `SME` version 5.22 was used with the ATLAS12 model spectra (Kurucz, 2013) to derive $T_{\rm eff}$, $\log g_\star$, [Fe/H], and $v \sin i_\star$. All of these parameters were allowed to vary throughout the model fitting, while the micro- and macro-turbulence ($v_{\rm mic}$ and $v_{\rm mac}$) were fixed through empirical calibration equations (Bruntt et al., 2010; Doyle et al., 2014) valid for Sun-like stars after a first estimation of $T_{\rm eff}$ was determined. The required atomic and molecular data (e.g., line lists and atom configurations) were taken from the `VALD` database (Ryabchikova et al., 2015). A detailed description of the methodology can be found in Fridlund et al. (2017) and Persson et al. (2018). This analysis produced $T_{\rm eff} = 5907 \pm 150$ K, $\log g_\star = 3.86 \pm 0.14$ from Mg I lines, $\log g_\star = 3.85 \pm 0.20$ from Ca I lines, [Fe/H] $= 0.10 \pm 0.10$ dex, and $v \sin i_\star = 8.2 \pm 0.9$ km s^{-1}.

The HARPS spectrum was independently analysed using the `specmatch-emp` package (Yee et al., 2017), which compares the observed spectrum with a library of ≈ 400 synthetic spectra of FGK and M stars. This fitting routine also uses a chi-squared minimisation approach to derive the stellar parameters, and yielded the values $T_{\rm eff} = 6006 \pm 110$ K and [Fe/H] $= 0.17 \pm 0.09$ dex. These values are consistent with those found in the `SME` analysis. As the `SME` derived value have slightly larger error bars, which I consider more realistic, I adopted these values for all further analysis (Table 4.2).

I determined the stellar mass, radius, and age using the on-line interface `PARAM-1.3`[8] with `PARSEC` stellar tracks and isochrones (Bressan et al., 2012), the *Gaia* parallax ($\pi = 3.7714 \pm 0.0225$ mas; Gaia Collaboration et al., 2018), a V-band magnitude of 10.286 (Munari et al., 2014), and the stellar parameters derived from the `SME` analysis of the HARPS data. As Munari et al. (2014) reported an interstellar reddening consistent with zero for this target, no correction was applied to the V-band magnitude reported in Table 4.2. The stellar mass, radius, and age were found to be 1.32 ± 0.06 $M_\odot$, 1.94 ± 0.10 $R_\odot$, and 3.7 ± 0.6 Gyr, respectively.

4.4.2 SED fitting

As an independent check on the derived stellar radius, an analysis of the broadband spectral energy distribution (SED) was performed, where the SED of a star is the distribution of energy emitted as a function of wavelength. By comparing a star's SED with synthetic photometric data, derived from stellar atmospheric models, one can determine the bolometric flux of a star $F_{\rm bol}$. This in turn can be used to determine the stellar radius using (Stassun & Torres, 2016; Stassun et al., 2017, 2018):

$$R_\star = (F_{\rm bol}/\sigma_{\rm SB}\, T_{\rm eff}^4)^{1/2} \tag{4.1}$$

where $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant.

TOI 813's observed SED was constructed using archival broadband photometric measurements that span the full stellar SED over the wavelength range $0.4-22$ μm (see Figure 4.6). All of the available magnitudes that were used for this analysis are listed in Table 4.2.

The synthetic SED was derived using Kurucz stellar atmosphere models (Kurucz, 2013). Priors were placed on the effective temperature ($T_{\rm eff}$), surface gravity ($\log g$), and metallicity ([Fe/H]) according to the spectroscopically determined values listed in Table 4.2. The only remaining free parameter was the extinction (A_V), which is the apparent dimming of an object caused by the presence of

[8] http://stev.oapd.inaf.it/cgi-bin/param

Table 4.2: Stellar parameters of TOI 813.

Parameter	Value	Source
Identifiers		
TIC	55525572	Stassun et al. (2019)
Gaia DR2	4665704096987467776	*Gaia* DR2[a]
2MASS	J04504658-6054196	2MASS[b]
Astrometry		
α_{J2000}	04:50:46.57	*Gaia* DR2[a]
δ_{J2000}	-60:54:19.62	*Gaia* DR2[a]
Distance (pc)	265.1535 ±1.582	*Gaia* DR2[a]
π (mas)	3.7714 ±0.0225	*Gaia* DR2[a]
Photometry		
B_T	11.189 ± 0.056	*Tycho-2*[c]
V_T	10.435 ± 0.056	*Tycho-2*[c]
B	10.905 ± 0.026	APASS[d]
V	10.322 ± 0.014	APASS[d]
g	10.587 ± 0.015	APASS[d]
r	10.240 ± 0.054	APASS[d]
i	10.094 ± 0.044	APASS[d]
G	10.2352 ± 0.0004	*Gaia* DR2[a]
J	9.326 ± 0.021	2MASS[b]
H	9.053 ± 0.018	2MASS[b]
K_s	9.029 ± 0.019	2MASS[b]
$W1$ (3.35 μm)	8.992 ± 0.023	WISE[e]
$W1$ (4.6 μm)	9.027 ± 0.020	WISE[e]
$W1$ (11.6 μm)	8.971 ± 0.022	WISE[e]
$W1$ (22.1 μm)	9.126 ± 0.271	WISE[e]
Kinematics		
Physical Properties		
Spectral type	G0 IV	Pecaut & Mamajek (2013)
Stellar mass, $M_\star$ ($M_\odot$)	1.32 ± 0.06	This work
Stellar radius, $R_\star$ ($R_\odot$)	1.94 ± 0.10	This work
Stellar density, $\rho_\star$ (g cm^{-3})	0.25 ± 0.05	This work
Effective temperature, $T_{\rm eff}$ (K)	5910 ± 150	This work
Surface gravity, $\log g_\star$ from Mg I (dex)	3.9 ± 0.1	This work
Surface gravity, $\log g_\star$ from Ca I (dex)	3.9 ± 0.2	This work
$v \sin i_\star$ (km s^{-1})	8.2 ± 0.9	This work
Iron abundance, [Fe/H] (dex)	0.10 ± 0.10	This work
Star age (Gyr)	3.7 ± 0.6	This work
Microturbulence, $v_{\rm mic}$ (km s^{-1})	4.4	Doyle et al. (2014)
Macroturbulence, $v_{\rm mac}$ (km s^{-1})	1.21	Bruntt et al. (2010)

Note – [a] *Gaia* Data Release 2 (DR2; Gaia Collaboration et al., 2018). [b] Two-micron All Sky Survey (2MASS; Cutri et al., 2003). [c] *Tycho*-2 catalogue (Høg et al., 2000). [d] AAVSO Photometric All-Sky Survey (APASS; Munari et al., 2014). [e] Wide-field Infrared Survey Explorer catalogue (WISE; Cutri & et al., 2013).

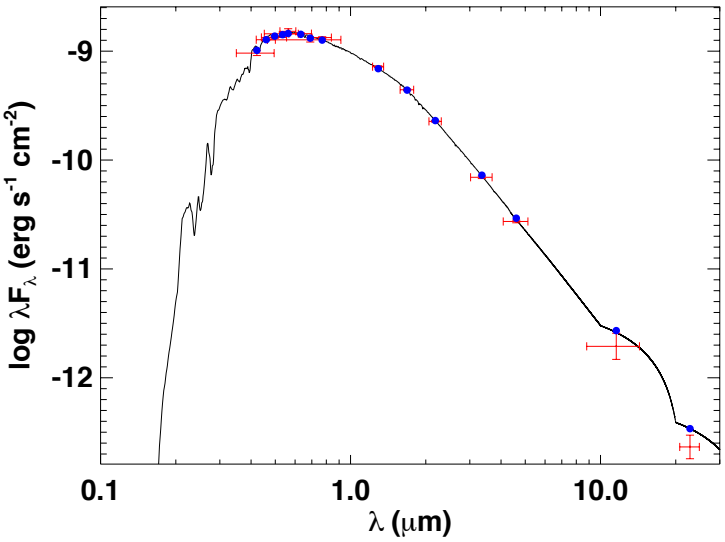

Figure 4.6: Spectral energy distribution of TOI 813. Red symbols represent the observed photometric measurements, where the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

dust grains in the interstellar medium along the line of sight. The dust maps of Schlegel et al. (1998) show that the line-of-sight extinction is $A_V = 0.00^{+0.01}_{-0.00}$. As such the extinction was set to zero.

The resulting fits (Figure 4.6) have a reduced χ^2 of 2.5. Integrating the (unreddened) model SED gives the bolometric flux at Earth of $F_{\rm bol} = 1.852 \pm 0.021 \times 10^{-9}$ erg s cm^{-2}. Taking the $F_{\rm bol}$ and $T_{\rm eff}$ together with the *Gaia* DR2 parallax, adjusted by +0.08 mas to account for the systematic offset reported by Stassun & Torres (2018), gives the stellar radius as $R_\star = 1.891 \pm 0.097$ R$_\odot$. Finally, estimating the stellar mass from the empirical relations of Torres et al. (2010) corresponds to $M_\star = 1.39 \pm 0.10$ M$_\odot$. These agree well with the spectroscopically derived parameters.

4.4.3 Statistical validation

The planetary nature of transit-like signals can be validated in a statistical sense by evaluating the likelihood that they are caused by a planet as opposed to the range of alternative, astrophysical false positive scenarios. This became routine in the era of *Kepler* (e.g., Borucki et al., 2012; Morton, 2012; Díaz et al., 2014; Santerne

et al., 2015), as many of the systems detected by that satellite are too faint for high-precision RV follow-up. Even though high-precision RV follow-up is possible for TOI 813 b, I statistically validated the signals for completeness.

For this analysis I used the open-source package `VESPA` (Morton, 2012, 2015), which calculates the false positive probability (FPP) by considering three astrophysical false positive scenarios: undiluted eclipsing binary (EB); eclipsing binary that is diluted by another star, often known as a background eclipsing binary (BEB); and hierarchical triple eclipsing binary (HEB). For each of these scenarios, `VESPA` simulates a population of stars, using the `TRILEGAL` galactic model (Girardi et al., 2005), and compares the shape of the simulated transits to the transits in the observed *TESS* light curve. This results in a likelihood for each false positive scenario.

I used version 0.4.7 of the software (with the MultiNest backend; Buchner et al., 2014) with the following inputs: the two-minute cadence detrended phase folded *TESS* light curve, the stellar parameters as derived in Section 4.4.1 and listed in Table 4.2, the two contrast curves derived from the SOAR and Gemini speckle images (Section 4.3.4), and the upper limit on the depth of a potential secondary eclipse (Section 4.2.3). The FPP was found to be 0.003, which is well below the commonly used validation threshold of FPP < 0.01 (e.g., Montet et al., 2015; Morton et al., 2016; Livingston et al., 2018), allowing me to statistically validate that TOI 813 b is a non-self-luminous object transiting the main target star.

4.4.4 Joint transit and RV modelling

I jointly modelled the *TESS* transit and HARPS RV data using the open-source software `pyaneti` (Barragán et al., 2019), as described in detail in Section 2.3. The 8 transit events were isolated and normalised by fitting a second order polynomial to the 48-hour out-of-transit data surrounding the centre of each transit event. I note that when modelling the 30-minute cadence Sector 2 transit data, the short-term variations in the data during each 30-minute exposure were accounted for explicitly (Kipping, 2010). This is because long integration times smear out the duration

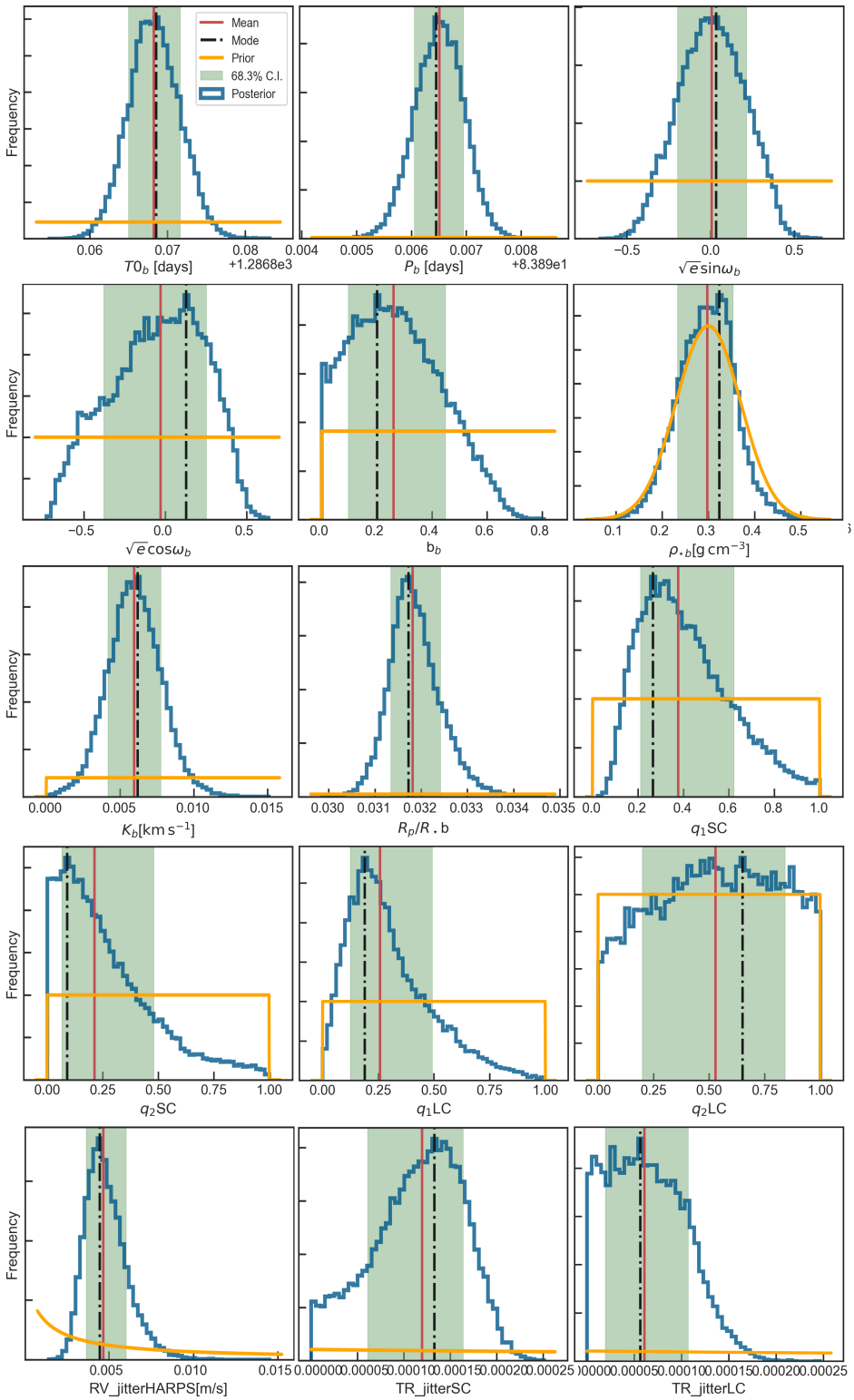

Figure 4.7: Posterior distribution of the joint transit and RV fit for TOI 813 b. The blue lines show the posterior distribution ($P(M|d)$), the orange region shows the prior ($P(M)$), the green shaded region shows the 68% credible interval, the black dashed line the mode and the red solid line the median.

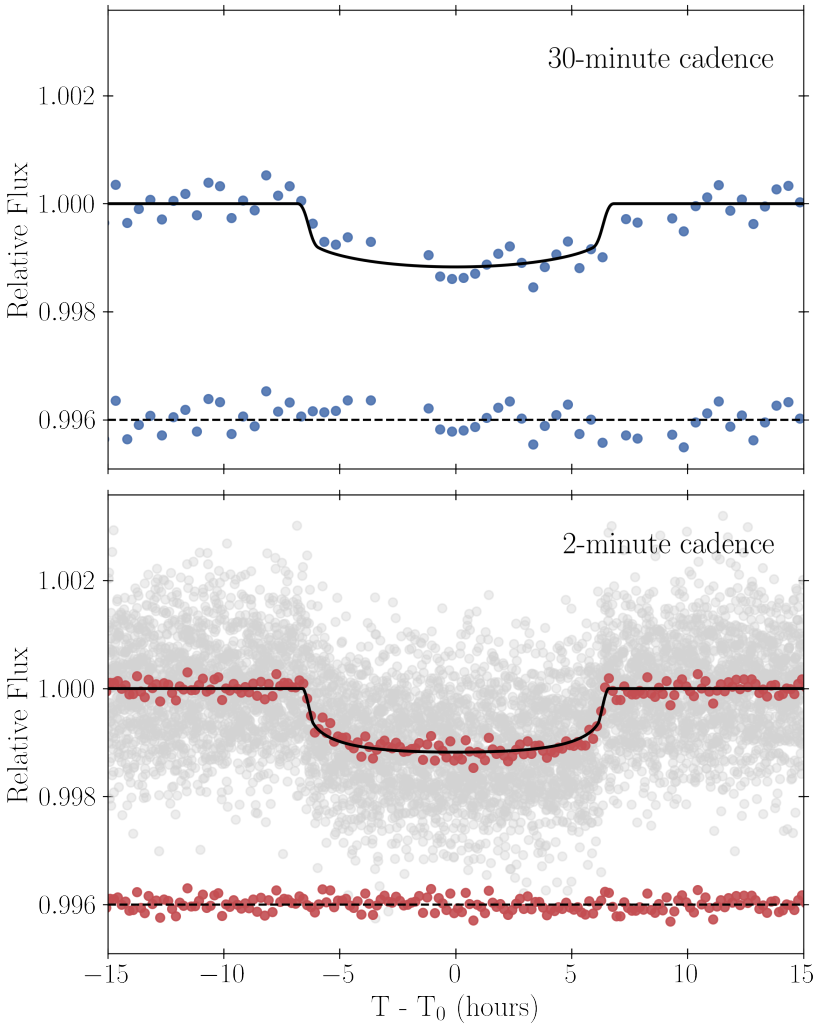

Figure 4.8: Phase-folded *TESS* light curve of TOI 813 b with the best-fit transit model overplotted and residuals.

of ingress and egress, while reducing the observed effects of limb-darkening, thus affecting the derived system parameters. The Sector 2 transit model was therefore computed at 3-minute intervals and subsequently integrated to 30-minute sampling before it was compared to the observations.[9]

The joint transit and RV parameter space (see Equations 2.11 and 2.13) was explored using an MCMC approach with 500 chains. Once the chains had converged, the last 5000 iterations, with a thinning factor of 10, were used to create posterior distributions based on 250 000 sampled points for each fitted parameter. As shown

[9]Due to the limited number of CHIRON observations (obtained for the original publication) these data were not included in the joint modelling.

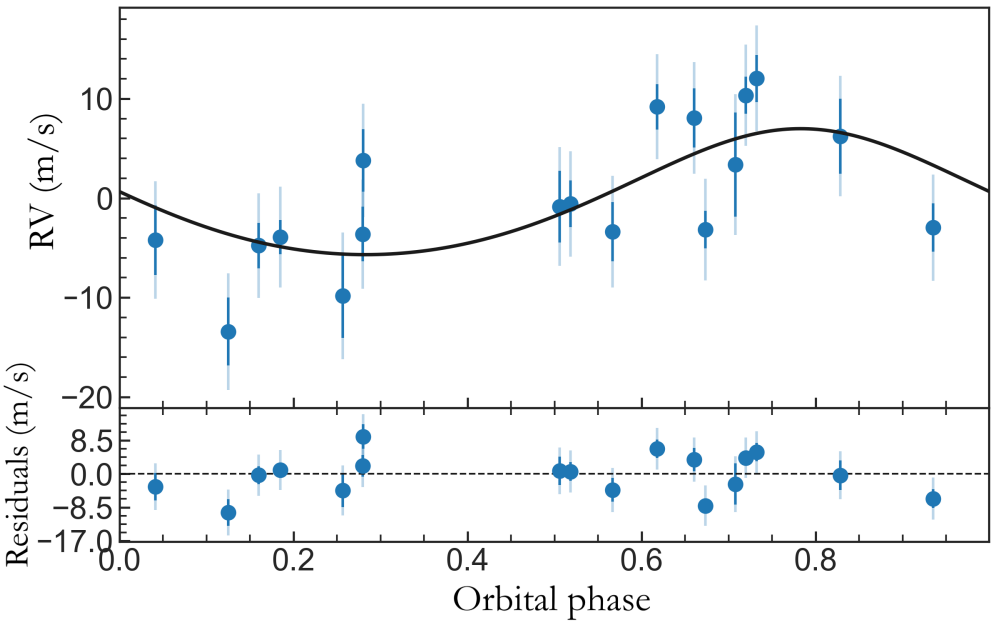

Figure 4.9: RVs phase folded with a period of 83.8960 days. The error bars show the nominal uncertainties of each measurement, while the grey vertical lines show the uncertainties including jitter. The solid black line shows the best-fit model.

in Figure 4.7, the posterior distributions are smooth and uni-modal for all fitted parameters. Furthermore, the residuals between the model and data showed no evidence of correlated noise.

Details of the fitted parameters and the priors are given in Table 4.3. The model, which allowed the orbit to be eccentric, was run with uniform priors over the interval $[-1, 1]$ for the parameters $\sqrt{e}\sin\omega$ and $\sqrt{e}\cos\omega$ (Anderson et al., 2011), and a Gaussian prior on $a/R_\star$ that was based on the stellar density derived in Section 4.4.1 and on Kepler's third law. This resulted in fitted values of $\sqrt{e}\sin\omega = \text{-}0.08^{+0.21}_{-0.19}$ and $\sqrt{e}\cos\omega = 0.03^{+0.24}_{-0.30}$, which translate to an eccentricity of $e = 0.08^{+0.10}_{-0.06}$. Furthermore, the posterior of K_b, which corresponds to a $3\,\sigma$ detection ($5.96 \pm 1.8\ \mathrm{m\,s^{-1}}$), confirms the planetary nature of TOI 813 b. The corresponding planetary mass of $48^{+15}_{-14}\ \mathrm{M_\oplus}$ is consistent with the upper mass limit previously published by Eisner et al. (2020b). The derived transit and RV models are shown in Figures 4.8 and 4.9, respectively.

After fitting all transits simultaneously, each of the 8 transits were fit individually, keeping all parameters fixed except for the time of the transit centre. This analysis only fit the transit data and did not include the RV data. The derived times of the

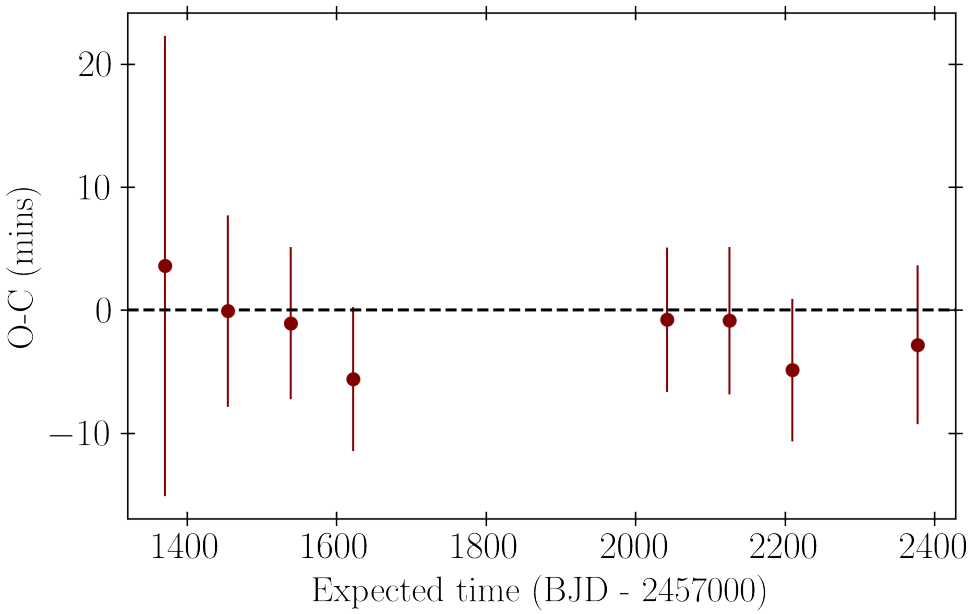

Figure 4.10: The observed minus calculated (O–C) transit times for all of the transits, as modelled using `pyaneti`. The calculated times were determined using a linear ephemeris with $P = 83.8960$ days and $T_{0,\mathrm{BJD}-2457000} = 1370.7682$ days.

transit mid-points were compared to the expected transit times, which were calculated using a linear epehermis with $P = 83.8960$ days and $T_{0,\mathrm{BJD}-2457000} = 1370.7682$ days. This showed that the observed transit times are consistent with a constant period (Figure 4.10). The larger uncertainty on the time of the first transit is due to the longer cadence of the Sector 2 observations.

4.5 Results and discussion

4.5.1 Planets around subgiant stars

TOI 813 b is in orbit around a subgiant host. The subgiant phase in some respects mimics the pre-main sequence phase in reverse: the same star-planet interaction mechanisms – tides, orbital migration, stellar heating and photoevaporation – are at play, but their effect is increasing, rather than decreasing, with time. Using a MIST version-1.2 stellar evolution track (Choi et al., 2016), with input stellar parameters as given in Table 4.2 and no rotation, I estimate the main sequence lifetime of TOI 813 to be ~ 3.45 Gyrs. This means that at the current age of 3.7 ± 0.6

Table 4.3: System parameters of TOI 813.

Parameter	Symbol	Prior[(a)]	Value[(b)]	Units
Model Parameters for TOI 813 b				
Orbital period	P	$\mathcal{U}[83.79, 83.90]$	$83.8965^{+0.0004}_{-0.0005}$	days
Transit epoch	T_0	$\mathcal{U}[1370.68, 1370.79]$	1370.768 ± 0.003	BJD - 2457000
Eccentricity parametrisation	$e \sin \omega$	$\mathcal{U}[-1, 1]$	0.01 ± 0.2	–
Eccentricity parametrisation	$e \cos \omega$	$\mathcal{U}[-1, 1]$	$-0.03^{+0.3}_{-0.4}$	–
Scaled planet radius	$R_p/R_\star$	$\mathcal{U}[0, 0.1]$	$0.03182^{+0.0006}_{-0.0005}$	–
Impact parameter	b	$\mathcal{U}[0, 1.1]$	0.26 ± 0.2	–
RV semi-amplitude	K	$\mathcal{U}[0, 50]$	5.96 ± 1.8	$\mathrm{m\,s^{-1}}$
Stellar density	$\rho_\star$	$\mathcal{N}[0.3, 0.1]$	0.30 ± 0.056	$\mathrm{g\,cm^{-3}}$
Limb-darkening coefficient	q_1	$\mathcal{U}[0, 1]$	$0.4^{+0.3}_{-0.2}$	–
Limb-darkening coefficient	q_2	$\mathcal{U}[0, 1]$	$0.2^{+0.3}_{-0.2}$	–
Jitter HARPS	–	$\mathcal{U}[0, 100]$	$4.7^{+1.4}_{-1.0}$	$\mathrm{m\,s^{-1}}$
Jitter TESS	–	$\mathcal{U}[0, 500]$	0.00013 ± 0.0001	ppm
Derived parameters				
Planet mass	M	–	48^{+15}_{-14}	$\mathrm{M_\oplus}$
Planet radius	R	–	6.8 ± 0.4	$\mathrm{R_\oplus}$
Planet density	ρ	–	$0.9^{+0.3}_{-0.3}$	$\mathrm{g\,cm^{-3}}$
Orbit eccentricity	e	–	$0.09^{+0.13}_{-0.07}$	–
Semi-major axis	a	–	0.43 ± 0.04	AU
Orbit inclination	i	–	89.7 ± 0.2	deg
Stellar density (from light curve)	$\rho_\star$	–	0.30 ± 0.056	$\mathrm{g\,cm^{-3}}$
Equilibrium temperature (albedo = 0)	T_{eq}	–	604^{+30}_{-20}	K
Insolation	F	–	$22.1^{+4.5}_{-3.3}$	$\mathrm{F_\oplus}$

Note – [(a)] $\mathcal{U}[a, b]$ refers to uniform priors between a and b, $\mathcal{N}[a, b]$ to Gaussian priors with median a and standard deviation b, and $\mathcal{F}[a]$ to a fixed value a. [(b)] Inferred parameters and errors are defined as the median and 68.3% credible interval of the posterior distribution.

Gyr, the star has only recently left the main sequence and is still in the process of entering what will become a very rapid period of evolution.

As TOI 813 b is relatively distant from its host star ($a = 0.43 \pm 0.04$ AU), it is not expected to interact with it very strongly. Specifically, there is no reason to believe that the orbit of TOI 813 b has been affected significantly by tidal interaction during the main sequence lifetime of the star (e.g., Lanza & Mathis, 2016), or that the planet's size and composition have been altered significantly by stellar irradiation (e.g., Vazan et al., 2013). These effects are thought to be important for close-in planets with $a < 0.2$ AU (Perryman, 2018). However, as the star evolves further along the red giant branch the stellar luminosity and radius will increase, changing the irradiation and tidal properties of the planet-star system. Using the same MIST stellar evolution track as for the analysis outlined above I estimate

that the star will reach a maximum radius of 0.76 AU on the red giant branch. In the absence of orbital evolution of the planet, TOI 813 b will be engulfed by its host at an age of $\sim$4.66 Gyrs.

While the radius of the star increases as it evolves beyond the main sequence, the mass can be assumed to remain near constant. I therefore estimate that the star's main sequence radius was around 1.1 $R_\odot$, using a mass–radius relation for main sequence stars of $R_\star \propto M^{3/7}$. This mass–radius relationship was derived empirically by Torres et al. (2010) based on mass and radius measurements of 190 binary systems. The derived main sequence stellar radius implies that the transit of TOI 813 b is already $\sim$4 times shallower than it used to be, highlighting the difficulties associated with detecting transiting planets around evolved host stars.

Even though the number of confirmed planets around evolved stars remains small, there is growing evidence that these systems' properties differ from those of their main sequence counterparts (e.g., Johnson et al., 2007; Luhn et al., 2019). Figure 4.11 shows all confirmed planets around evolved stars that have mass measurements with at least a 30% accuracy listed the NASA Exoplanet Archive. The evolved stars were identified using the data driven boundaries (derived using *Kepler* targets) in effective temperature ($T_{\rm eff}$) and surface gravity (g), as outlined in Huber et al. (2016), where post-main sequence stars are classified as these when:

$$\log(g) < \frac{1}{4.671} \arctan\left(\frac{T_{\rm eff} - 6300}{-67.172}\right) + 3.876. \tag{4.2}$$

Out of the 4903 confirmed exoplanets listed in the NASA Exoplanet Archive as of 19 January 2022, 709 have a post-main sequence host (487 subgiants and 222 red giants), 262 of which were discovered using the transit method. Out of these, 109 planets have a mass measurement with an accuracy better than 30%. As shown in Figure 4.11, the planet candidates around evolved stars identified with the transit method (teal triangles) have significantly shorter orbital period than the planets identified using the RV method (pink circles). This is because, even though both the transit and RV methods are biased towards the detection of short-period planets, the detection dependence on period is stronger for the transit method than for the

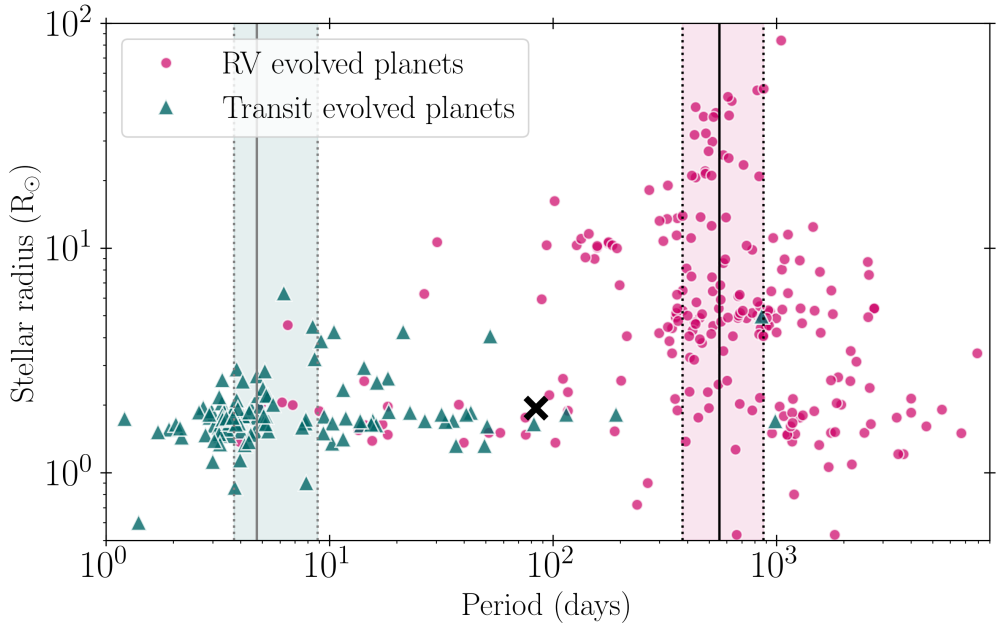

Figure 4.11: All planets around evolved stars listed in the NASA Exoplanet Archive that have mass measurements to better than 30% accuracy. The teal triangles show planets detected using the transit method and the pink circles the planets detected using the RV method. The solid vertical lines indicate the mean periods of 4.63 days and 555 days for the transit and RV methods, respectively. The shaded regions for the $1\,\sigma$ percentiles of the distribution in periods for each method. The properties of TOI 813 b, as shown by the black cross, place it in a sparsely populated region of this figure.

RV method. Furthermore, the fundamental period-detection limit in both cases is the survey duration. To that end, large-scale space-based transit surveys that typically only monitor the brightness of stars for a couple of years (e.g., *CoRoT*, *Kepler*, *TESS*; see Chapter 1) are limited to the detection of shorter-period planets compared to ground-based RV surveys that often have long baselines of over 10 years (e.g., Cumming et al., 2008; Demangeon et al., 2021). These biases can be seen in Figure 4.11, where the solid vertical lines indicate the mean orbital periods of planets discovered with each discovery method, and the shaded regions show the $1\,\sigma$ percentile of the period distributions. The figure highlights that there is a period range from $\sim$10 to 300 days that is sparsely populated. The detection and characterisation of planets within this intermediate period range, such as TOI 813 b (shown by the black cross in the figure), can help tie the populations of RV and transiting planets around evolved stars together, which in turn will help clarify planet-star interactions beyond the main sequence.

I note that other validated planets in orbit around bright subgiant host stars observed by *TESS* include HD 1397 (Brahm et al., 2019), HD 221416 (Huber et al., 2019), HD 1397 (Nielsen et al., 2019b), and TOI 172 (Rodriguez et al., 2019).

4.5.2 Transition between ice giants and gas giants

The mass of TOI 813 b of 48^{+15}_{-14} $M_\oplus$ places it in a scarcely populated region of parameter space between the mass of Neptune (17.15 $M_\oplus$) and the mass of Saturn (95.16 $M_\oplus$), as shown in Figure 4.12. As previously mentioned in Section 1.6.2, this mass range represents a transition region between volatile-rich gas giants and volatile-poor ice giants, and is only populated with 88 planets that have mass measurements to better than 30% accuracy (as of 19 January 2022). The characterisation of planets in this mass range can help constrain (currently debated) theories of how gas giants and ice giants form.

In the core accretion model of planet formation, for example, giant planets form by coagulating km-sized planetesimals into large, rocky cores. Once a critical mass of around 10 $M_\oplus$ is reached, a gaseous envelope is accreted onto the core of ice and rock, forming a gas giant (Pollack et al., 1996). This model, however, struggles to explain why ice giants do not accrete gaseous envelopes to also become gas giants.

Lambrechts et al. (2014) suggest an alternative planet formation model where the planetesimals grow by accreting cm-sized pebbles instead of via planetesimal coagulation. Once a pebble-accreting core reaches a threshold mass ($\sim$50 $M_\oplus$ at a semi-major axis of 5 AU) it can clear its surroundings by gravitationally perturbing the gas disk in the vicinity of the core. As a result, pebble accretion stops, allowing the inward gravitational pull of the core to become greater than the outward pressure due to heat generated by the accretion. This triggers the collapse of the gas envelope onto the core of the planet, resulting in a gas giant. Unlike gas giants, ice giants do not reach the required threshold mass. This is because the threshold mass is larger for cores on wider orbits. As such, due to the increased gas scale height of protoplanetary disks at larger radii, more gravitational energy is required to perturb the surrounding gas disk and to halt pebble accretion. Therefore, planets that form

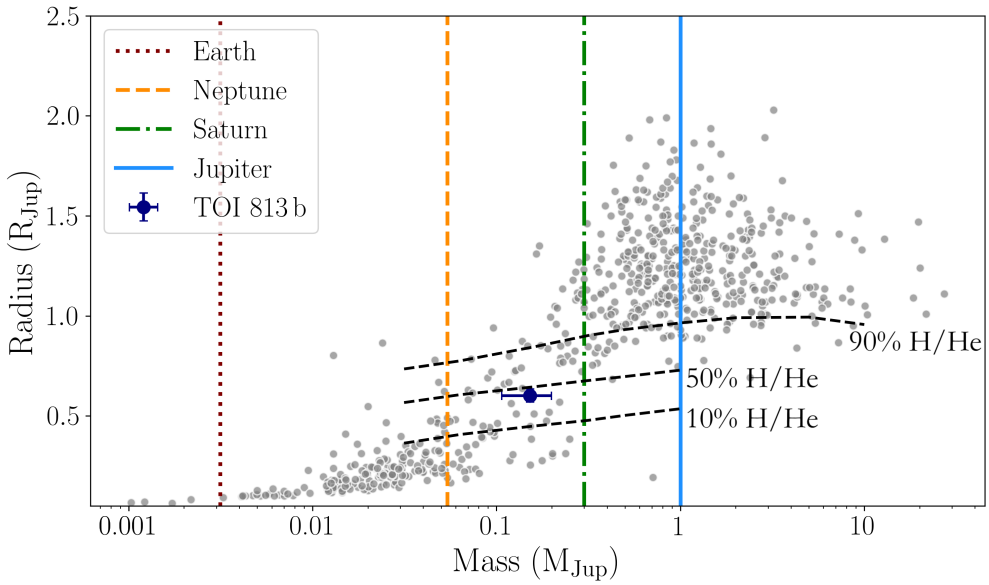

Figure 4.12: Mass–radius diagram showing all transiting planets listed in the NASA Exoplanet Archive that have mass measurements to better than 30% precision. The vertical lines show the masses of Earth (dotted maroon line), Neptune (dashed orange line), Saturn (dash/dotted green line), and Jupiter (solid navy line). The black dashed lines show the 3 Gyr isochrones for model planets with various fractions of H/He (Baraffe et al., 2008) in the planet's atmosphere. These models do not consider irradiation from the star, which would act to decrease the fraction of H/He for a given mass.

on wide orbits are expected to have thin envelopes of H, He, and water vapour from sublimated icy pebbles, i.e. ice giants. Thus, the identification and atmospheric characterisation of planets in this mass transition region between ice and gas giants, such as TOI 813 b, can help differentiate between different evolutionary pathways such as those proposed by Pollack et al. (1996) and Lambrechts et al. (2014).

As shown in Figure 4.12, TOI 813 b's mass suggests that the planet has a composition made up of less than 50% H/He, according to the 3 Gyr isochrone models from Baraffe et al. (2008). I note that these isochrone models do not account for stellar irradiation, which would act to decrease the fraction of H/He for a given mass and radius. Measurements of the atmospheric metallicity of TOI 813 b would help to determine whether its properties are more closely related to those of gas giants or ice giants, where gas giants have significantly higher metalicities than ice giants (Guillot & Gautier, 2015). This, in turn, would help clarify how giant planets form and how ice and gas giants differ. The prospects for atmospheric

characterisation of TOI 813 b is discussed in Section 4.5.5.

Other confirmed planets located in the transition mass region include WASP-107 b (Anderson et al., 2017), WASP-139 b (Hellier et al., 2017), WASP-182 b (Nielsen et al., 2019a), HATS-7 b Bakos et al. (2015), and HATS-8 b (Bayliss et al., 2015).

4.5.3 Search for additional planets

A statistical analysis of all of the planet candidates identified in the *Kepler* data, carried out by Lissauer et al. (2012), showed that the discovery of multiple planets within one system can significantly increase confidence in the planetary nature of a transit-like event. This is based on the argument that it is unlikely for us to see multiple false-positive signals in the same light curve.

The apparently low eccentricity of TOI 813 b of $0.09^{+0.13}_{-0.07}$ is characteristic of planets in multiple transiting systems. Van Eylen et al. (2019) used well-characterised planets found in the *Kepler* data to show that systems with a single transiting planet have an average eccentricity of ~ 0.3, while multi-planet systems have a mean eccentricity of $0.04\ ^{+0.03}_{-0.04}$. This, combined with TOI 813 b's long orbital period and the high fraction of main sequence stars that host nearly coplanar multi-planet systems (e.g., Rowe et al., 2014), suggests that TOI 813 b is likely an outer planet in a multi-planet system.[10]

Visual inspection of the data revealed no additional planet candidates. However, as illustrated in Section 3.1.2, the detection efficiency of visual vetting decreases notably as a function of planet radius. I therefore carried out a search for additional transits in the full light curve using the Box Least Squares (BLS; Kovács et al., 2002) algorithm after masking the transits of TOI 813 b, where each mask spanned 0.75 days centred on the mid-point of the transits. Before running the BLS, I used an iterative non-linear filter (as described in Section 2.2.3) to estimate and subtract residual systematics on timescales > 1.7 days (see Figure 4.1). The BLS search was carried out on an evenly sampled frequency grid ranging from 0.01 to 1

[10] I note that more RV measurements are needed to constrain the eccentricity of TOI 813 b.

days^{-1} (1 to 90 days). I used the ratio of the highest peak in the S/N periodogram relative to its standard deviation, known as the signal detection efficiency (SDE) to quantify the significance of the detection. The algorithm found no additional signals with SDE >7.2 (compared to SDE ~ 22.4 for the transit of TOI 813 b). The non-detection of additional transits suggests that if there are other planets in the TOI 813 system, their orbits are either inclined to not transit or their transits are shallower than the detection threshold.

4.5.4 Search for stellar oscillations

As a subgiant, TOI 813 is expected to display solar-like (pressure-mode) oscillations. These oscillations are stochastically driven by the turbulent convection in the envelope of solar-like, subgiant and red giant branch stars. The frequencies at which these stellar oscillations occur are determined by the the mass and radius of the star (Aerts et al., 2010). As such, when they are detected, they provide an independent constraint on fundamental stellar parameters (e.g., Huber et al., 2013; Van Eylen et al., 2019).

I performed a search for such oscillations using the Lomb-Scargle periodogram (Lomb, 1976; Scargle, 1982), but did not detect any evidence of oscillations. The *TESS* Asteroseismology Consortium (TASC) were also unable to detect oscillations in this object (W. Chaplin, priv. comm.). The lack of a detectable signal is potentially due to the high-frequency window pattern caused by regularly spaced data gaps (e.g., by missing data due to regular momentum dumps). This is because regular data gaps introduce noise at high frequencies, resulting in a noise floor that makes the identification of oscillations at high-frequencies, such as those expected for an evolving subgiant, difficult (e.g, Kallinger et al., 2014). A potential solution to minimise the effects of missing data is to interpolate the data onto an evenly sampled grid. This method, which is often referred to as 'inpainting', has been successfully applied to small gaps in *Kepler* data. However, the larger gaps between sectors in *TESS* data make this approach more challenging and is thus beyond the scope of the work presented in this chapter.

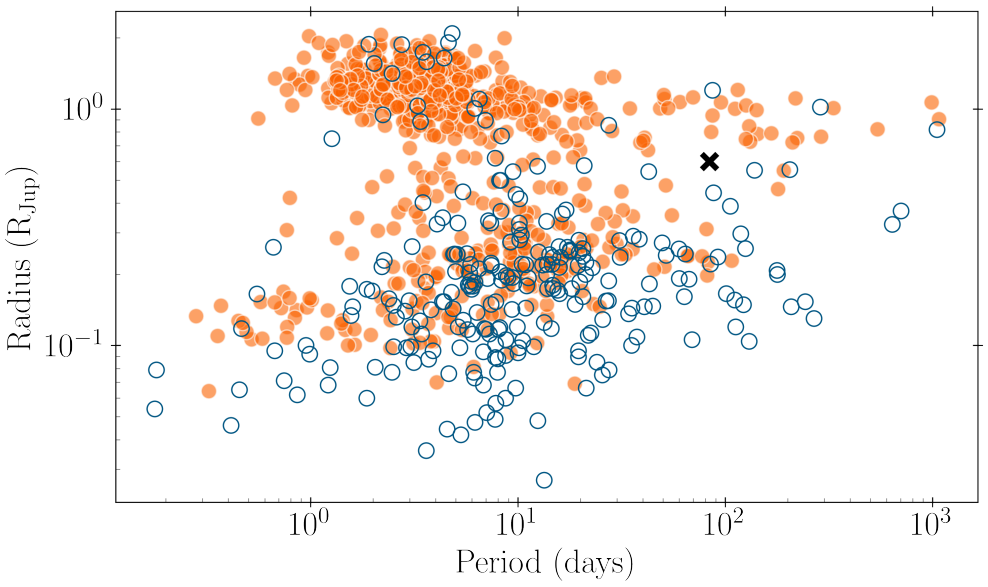

Figure 4.13: A log-log plot of the period vs radius of transiting exoplanets as they appear in the TEPCat catalogue of well-studied transiting planets (as of 19 January 2022; Southworth, 2011) . The filled orange circles represent planets that have a mass measurement with at least a 30% accuracy. The black cross denotes the properties of TOI 813 b, where the value uncertainties are smaller than the marker.

4.5.5 Follow-up prospects

TOI 813 b is of particular interest due to its long orbital period compared to other targets found by *TESS*. Accurate mass measurements of long-period transiting planets around bright stars are rare, as shown in Figure 4.13, as are detailed studies of planets orbiting around evolved stars, and planets located in the mass transition region between ice and gas giants. Additional observations of this target can therefore help to explore under-explored regions of parameter space.

For example, precise RV measurements during the transit may reveal small deviations from the Keplerian fit in the RV curve (the so-called Rossiter-McLaughlin, or RM, effect; Rossiter, 1924; McLaughlin, 1924). The RM effect can be used to estimate the projected angle between host star's spin axis and the normal to the orbital plane of the planet (e.g., Schneider, 2000). While planetary migration through the disk should preserve, or even reduce, the primordial spin-orbit alignment, effects such as planet-planet scattering and Lidov-Kozai resonance (Kozai, 1962; Lidov, 1962) should result in a misalignment over time (e.g., Storch et al., 2017; Deeg

et al., 2009). Measuring the RM effect can therefore help constrain the dynamical history of the system. For TOI 813 b, I estimate the amplitude of the RM effect (which scales with the projected stellar equatorial rotational velocity ($v \sin i_\star$; Winn, 2010), to be around $7.7^{+0.3}_{-0.4}$ m s^{-1} and therefore detectable using instruments such as HARPS. The RM effect combined with the stellar rotation period would allow us to measure the true 3D obliquity of the system. Unfortunately, none of the currently available RV measurements were obtained during a transit, meaning that I am unable to determine whether this effect is present given the current data set.

Long-period transiting planets, such as TOI 813 b, are also potentially interesting targets for atmospheric follow-up. TOI 813 b has a relatively low insolation ($22.1^{+4.5}_{-3.3}$ $F_\oplus$) and equilibrium temperature (604^{+30}_{-20} K; estimated using the Stefan-Boltzmann law and assuming an albedo of 0). As such, measuring its transmission spectrum would provide useful clues as to the atmospheric conditions in a relatively cool giant planet. To evaluate the feasibility of such an observations, I used the transmission spectroscopy metric (TSM) proposed by Kempton et al. (2018) to approximate the expected S/N for a 10 hour observation with JWST/NRIRISS (based on the assumption of a cloud-free atmosphere). Using the derived mass of 48^{+15}_{-14} $M_\oplus$ and a mean molecular weight of 2.3, the TSM is $\sim$35. Even though this is relatively low (as expected as cooler atmospheres are less puffy), this is to some extent compensated by the relatively long duration of the transit. As such, future TSM observations remain feasible and would provide useful clues as to how giant planets form.

4.6 Summary and conclusions

In this chapter I discussed the detection and validation of a planet orbiting around a subgiant host star. The planet's relatively long orbital period of $\sim$84 days, together with the evolved nature of the host star, places TOI 813 b in a relatively under-explored region of period-radius parameter space. Furthermore, it belongs to the small population of planets that lie in the transition mass range between

ice giants and gas giants, both of which make it an exciting target for future atmospheric follow-up observations.

The identification of TOI 813 b via PHT highlights the importance of citizen science for finding longer-period planets in *TESS* data. Furthermore, it shows that with visual vetting we are able to detect scientifically valuable planets around post-main sequence stars. Even though evolved stars are not typically selected to be two-minute cadence *TESS* targets, *TESS* observes over 400 000 post-main sequence stars brighter than 12th magnitude in the FFIs every two years. As such, the planned future expansion of PHT to include the FFI data of post-main sequence stars (see Chapter 7) promises to find more planet candidates like TOI 813 b, and thus to further increase the sample of known planets around evolved stars.

Ask us no questions and we'll tell you no lies.

— F. Weasley and G. Weasley, 1995

CHAPTER 5

HD 152843: TWO TRANSITING PLANETS AROUND A BRIGHT G DWARF

5.1 Context

As discussed in Chapter 1, systems with multiple transiting planets offer a wealth of information for exoplanetary science. In particular they allow for comparative planetology: studying planets that have formed out of the same material, but which have formed and evolved in different environments, receiving different amounts of incident flux from the host star, resulting in differing masses, radii, and compositions. Well characterised multi-planet systems therefore provide important model constraints that single-planet systems cannot, providing insight into planetary system architecture and evolutionary pathways, as well as informing ongoing planet population studies (e.g, Tremaine & Dong, 2012; Dietrich & Apai, 2020).

The *Kepler* mission (Borucki et al., 2010) revealed that multi-planetary systems are common (Latham et al., 2011), with almost half of all *Kepler* planets listed in the NASA Exoplanet Archive belonging to multi-planet systems. However, the majority of the hundreds of multi-planet systems found by *Kepler* are too faint to follow-up with ground-based high-resolution spectroscopy. This has resulted in most known multi-planet systems lacking well determined masses, densities, bulk

compositions, and atmospheric characterisation, all of which are key to helping us understand the overall planet population.

The selected two-minute cadence stars observed by *TESS*, however, are on average 30 – 100 times brighter than those observed by the *Kepler* mission, thus allowing us to follow-up and constrain the properties of new systems that were previously inaccessible. *TESS* has already discovered tens of previously unknown multi-planet systems (e.g., Gandolfi et al., 2019; Quinn et al., 2019; Fridlund et al., 2020; Dragomir et al., 2019; Gilbert et al., 2020; Mann et al., 2020; Leleu et al., 2021).

Despite the large number of exoplanet discoveries made by *TESS* and *Kepler*, systems with more than one transiting planet around stars brighter than $V = 10$ containing planets with measured masses remain rare. As of 19 January 2022, there are only 24 transiting planets (in 18 multi-planet systems) with mass measurements better than 30% precision around stars with $V < 10$ (the typical magnitude required for atmospheric follow-up, e.g., Fortenbach & Dressing, 2020) listed in the NASA Exoplanet Archive. Significant observing resources have been, and continue to be, devoted to each of them.

In this Chapter, I present the discovery of two planets orbiting around the bright ($V = 8.85$ mag) G dwarf HD 152843.

5.2 TESS Photometry

5.2.1 Discovery of HD 152843 b and HD 152843 c

HD 152843 is a two-minute cadence target that was observed by *TESS* in Sector 25 of the primary mission.[1] Additional data will be obtained by *TESS* in Sector 52 (obtained between 17 May and 13 June 2022). As shown in Figure 5.1, the light curve exhibits three transit events belonging to two different transiting planets, where the two transits of the inner planet HD 152843 b are shown in blue and the single transit of the outer planet HD 152843 c is shown in pink. The first transit event of HD 152843 b ($T_{0,\mathrm{BJD}-2457000} \sim 1994.28$ days) and the single transit

[1]The Sector 25 light curve used for the remainder of this chapter was processed and reduced with version 4.0.36 of the SPOC pipeline (Jenkins et al., 2016).

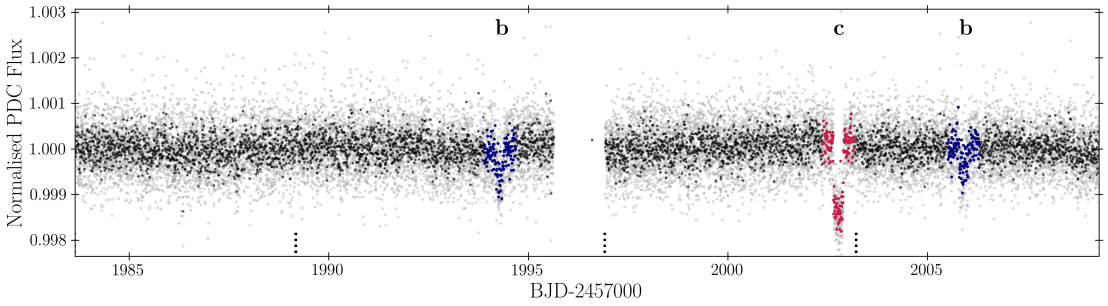

Figure 5.1: Normalised flux time series for HD 152843 vs TESS Julian day (BJD - 2457000) for Sector 25. The light grey points show the short cadence data with a two minute sampling, while the black points are ten minute averages. The dashed vertical lines at the bottom of the figure show the times of the *TESS* momentum dumps. The transit events are shown in blue and pink, corresponding to the inner and outer planet candidates, respectively.

of HD 152843 c ($T_{0,\mathrm{BJD}-2457000} \sim 2002.77$ days) were flagged as a single Threshold Crossing Event (TCE) by the SPOC pipeline as two events seemingly caused by the same 'object'. However, due to the different depths of these two transits the TCE was not promoted to TOI status, due to the assumption that the two events correspond to the primary and secondary eclipses of an eclipsing binary. The second transit event of HD 152843 b was not flagged by the automated search.

All three transit events were, however, identified on the PHT platform. The light curve of HD 152843 was seen by 15 citizen scientists, 12 of whom identified all three transit events, and 3 who identified only two out of the three events. The target was initially brought to the attention of the PHT research team via the PHT discussion forum (see Section 2.1).[2]

5.2.2 Excluding false positive scenarios

Once the light curve had been identified as a target of interest on PHT, I generated a `LATTE` vetting report in order to investigate the planetary nature of the signals. The `LATTE` vetting tests, described in Section 2.2, showed that both signals are consistent with being the result of on-target transiting planets and not with being false positives, such as a background eclipsing binary. However, to further investigate

[2] https://www.zooniverse.org/projects/nora-dot-eisner/planet-hunters-tess/talk/2112/1552434?comment=2520798

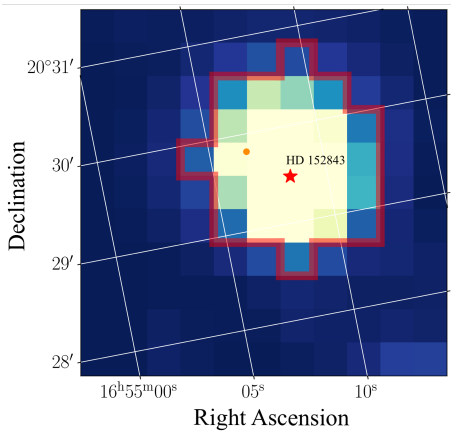

Figure 5.2: The median *TESS* image around HD 152843. The aperture used to extract the light curve is shown by the red outline and the orange dot depicts the location of the only star brighter than $V = 15$ within 110 arcsec of the target (red star), as queried by *Gaia* eDR3 (Gaia Collaboration et al., 2021). This nearby star ($V = 14.4$) is located at an angular separation of $\sim$31.3 arcsec.

the possibility of the signals being due to a nearby background eclipsing binary I queried the *Gaia* eDR3 catalogue (Gaia Collaboration et al., 2021) for stars within 110 arcsec ($\sim$5 *TESS* pixels) of the target. I found there to only be a single star with a V-band magnitude brighter than 15, as shown by the orange dot in Figure 5.2, alongside HD 152843 which is shown by the red star. The aperture used to extract the light curve is shown by the red outline. Following the methodology outlined in Section 4.2.3 and using the transit parameters derived from the joint transit and RV modeling (see Section 5.4.5), I showed that the maximum magnitude difference between the target star and any possible background contaminant is 1.5 mag in the V-band. This allows me to confidently conclude that the 14.4 mag star (5.6 mag fainter than HD 152843), located at an angular separation of $\sim$31.3 arcsec, cannot be responsible for the observed transit signals.

Having ruled out as many false positive scenarios as possible given the available data, I uploaded planet b ($R_b = 3.41^{+0.12}_{-0.12}$ R$_\oplus$; P_b $=11.6264^{+0.0047}_{-0.0041}$ days) and planet c ($R_c = 5.84^{+0.15}_{-0.15}$ R$_\oplus$; P_c $= 20.5^{+1.4}_{-0.9}$ days) to the ExoFOP site on 7 August 2020 as a cTOI. The inner planet candidate has since been promoted to the priority 1 (1 = highest priority, 5 = lowest priority) candidate TOI 2319.01.

Table 5.1: Stellar parameters of HD 152843.

Parameter	Value	Source
Identifiers		
HD	152843	
TOI	2319	
TIC	349488688	Stassun et al. (2019)
Gaia eDR3	4564566554995619072	*Gaia* eDR3[(a)]
2MASS	J16550834+2029287	2MASS[(b)]
Astrometry		
$\alpha_{\rm J2000}$	16:55:08.373	*Gaia* eDR3[(a)]
$\delta_{\rm J2000}$	20:29:29.509	*Gaia* eDR3[(a)]
Distance (pc)	107.898 ± 0.317	Bailer-Jones et al. (2018)
π (mas)	9.161 ± 0.015	*Gaia* eDR33[(a)]
Photometry		
B	9.380 ± 0.020	*Tycho-2*[(c)]
V	8.850 ± 0.010	*Tycho-2*[(c)]
J	7.896 ± 0.018	2MASS[(b)]
H	7.655 ± 0.016	2MASS[(b)]
K	7.629 ± 0.020	2MASS[(b)]
$W1$	7.563 ± 0.031	WISE[(d)]
$W2$	7.594 ± 0.020	WISE[(d)]
$W3$	7.607 ± 0.019	WISE[(d)]
Physical Properties		
Spectral type	G0	
Stellar mass, $M_\star$ ($M_\odot$)	1.15 ± 0.04	This work
Stellar radius, $R_\star$ ($R_\odot$)	1.43 ± 0.02	This work
Stellar density, $\rho_\star$ ($\rho_\odot$)	0.40 ± 0.03	This work
Effective temperature, $T_{\rm eff}$ (K)	6320 ± 100	This work
Surface gravity, $\log g_\star$ (dex)	4.21 ± 0.12	This work
$v \sin i_\star$ ($\rm km\,s^{-1}$)	8.38 ± 0.50	This work
[M/H] (dex)	-0.23 ± 0.08	This work
[Fe/H] (dex)	-0.16 ± 0.05	This work
Star age (Gyr)	3.97 ± 0.75	This work
Microturbulence, $v_{\rm mic}$ ($\rm km\,s^{-1}$)	1.66 ± 0.13	This work
Macroturbulence, $v_{\rm mac}$ ($\rm km\,s^{-1}$)	2	Bruntt et al. (2010)

Note – [(a)] *Gaia* early Data Release 3 (eDR3; Gaia Collaboration et al., 2021). [(b)] Two-micron All Sky Survey (2MASS; Cutri et al., 2003). [(c)] *Tycho*-2 catalogue (Høg et al., 2000). [(d)] Wide-field Infrared Survey Explorer catalogue (WISE; Cutri & et al., 2013).

5.2.3 Limits on additional planets

Following the detection of planets b and c, I explored the possibility of additional transiting planets within the system. I did this by quantifying the detectability of additional planets in the *TESS* light curve using a transit injection and recovery test. First, I masked out the three known transit events prior to injecting synthetic signals into the PDC *TESS* light curve. The masks spanned 0.5 days centred on the mid-point of each transit event. The injected signals were then generated using the open-source BATMAN python package (Kreidberg, 2015), with planet radii ranging from 1 to 12 $R_\oplus$ and periods ranging from 3 to 24 days (approximately the duration of the sector), both sampled at random from a log-uniform distribution in order to over-sample the smaller planets that are more difficult to recover. As multi-planet systems tend to have small eccentricities (Van Eylen et al., 2019), I fixed the eccentricity to be zero for all simulations. The impact parameter was also assumed to be zero for simplicity. I used a quadratic limb-darkening law with q_1 and q_2 of 0.16 and 0.59, respectively, as taken from Table 15 in Claret (2017) using the stellar parameters given in Table 5.1. Once the signals were injected, I used an iterative non-linear filter (Aigrain & Irwin, 2004) to estimate and subtract residual systematics on timescales greater than 1.7 days (see Section 2.2).

I simulated and injected a total of 750 000 transit events. The BLS algorithm was then used to try to recover the injected signals, with a search that sampled a frequency grid that was evenly-spaced from 0.01 to 1 day^{-1}. For each simulation, the period and orbital phase corresponding to the highest peak in the BLS periodogram were recorded. If the recovered orbital period and phase agreed to within 1% of the injected signal, the injection was deemed to be correctly identified. The completeness, assessed in radius and period bins with width of 0.25 $R_\oplus$ and 0.75 days respectively, was then taken to be the fraction of correctly identified transit signals within each bin.

The results presented in Figure 5.3 show, as expected, that the automated BLS search is strongly biased towards detecting shorter period planets that transit multiple times in the light curve. The limited duration of the *TESS* observations of ∼27 days, interrupted by a 1.3 day data gap, results in a sharp decline in

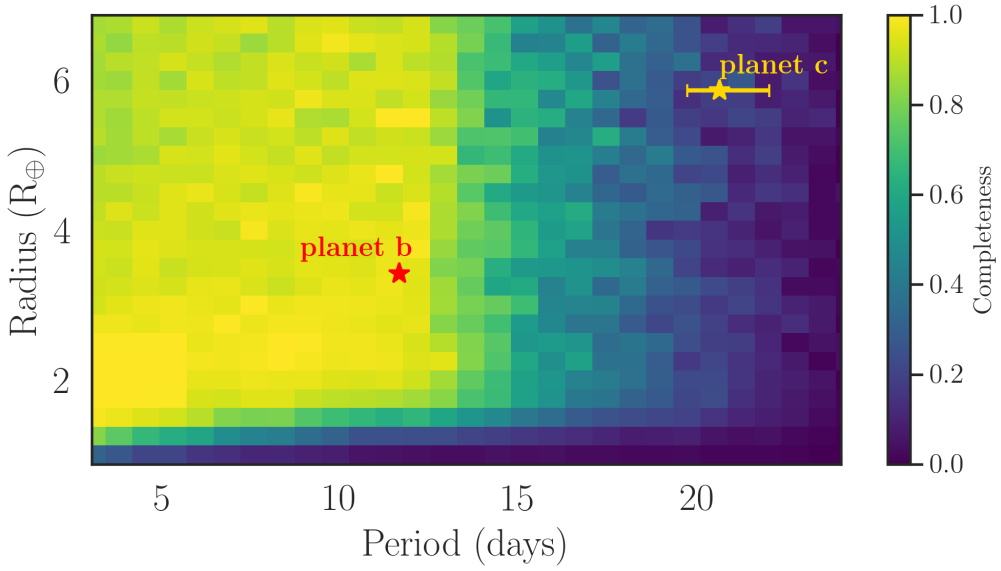

Figure 5.3: The recovery completeness of injected transit signals into the light curve of HD 152843 as a function of the radius and orbital period. The signals were recovered using a BLS search. The properties of HD 152843 b and HD 152843 c are shown by the red and yellow star respectively.

completeness for periods longer than around 13 days. For planets greater than 2 $R_\oplus$, I recover $\sim$94% of signals with periods between 12 and 13 days and $\sim$78% of signals with periods between 14 and 15 days. The completeness for the parameters of planet b is close to 100%, while the completeness for the estimated parameters of planet c is around 20% due to the fact that the estimated period of planet c is such that the planet often only transits once within the available *TESS* light curve. I caution that the simulated signals were injected into the PDC light curve, which has already undergone detrending and corrections of systematics by the SPOC pipeline. The presented recovery rates are, therefore, systematically higher than one might otherwise expect if the signals had been injected into the raw light curve (e.g., Lienhard et al., 2020).

Overall, this analysis highlights the difficulties associated with detecting longer-period planets using automated algorithms. Conversely, as discussed in Section 3.1.2 and shown in Figure 3.4, visual vetting is equally sensitive to the detection of longer-period, single transiting planets as to planets that transit multiple times within the data. This demonstrates a need for alternative detection methods to

complement the automated searches, such as citizen science, in order to identify longer-period planets such as HD 152843 c.

5.3 Spectroscopic data

My collaborators and I obtained spectroscopic data in order to better characterise the host star, rule out further astrophysical false positive scenarios (e.g., eclipses due to a low mass stellar companion), and to place constraints on the masses of both planets in order to validate the planetary nature of the system.

5.3.1 Reconnaissance spectra

First, we made use of the Las Cumbres Observatory (LCO) telescopes with the Network of Robotic Echelle Spectrographs (NRES; Brown et al., 2013) to obtain a reconnaissance spectrum. As discussed in Section 2.4, this fibre-fed spectrograph, mounted on a 1.0-m telescope, has a resolution of $R \approx 53\,000$ and a wavelength coverage of 380 to 860 nm. We obtained two spectra of HD 152843 on the 15th and 22nd August 2020 with per pixel S/N ratios of 38 and 25 at 520 nm, respectively. The two spectra, which were obtained at a phase -0.25 and 0.15 of planet b, gave radial velocity estimates of 9.7 ± 0.2 kms^{-1} and 9.6 ± 0.7 kms^{-1}. These are consistent within their uncertainties, and allowed me to rule out the possibility that the transit events are caused by an eclipsing binary. For comparison, if the planet b eclipses were the result of a low mass stellar companion with $M = 0.1$ M$_\odot$, the RV shift between those two observations would have been ~ 14 kms^{-1}.

5.3.2 RV monitoring

We also acquired high-resolution (R $\approx 115\,000$) spectra with the High Accuracy Radial velocity Planet Searcher in the Northern hemisphere (HARPS-N; Cosentino et al., 2012, 2014) spectrograph mounted at the 3.6-m Telescopio Nazionale Galileo in La Palma (PI: Nicholson; Program ID A41DDT4). A total of 83 spectra were obtained between 5 September 2020 and 23 September 2021 (mean S/N ~ 97 at 550 nm). Each spectrum was extracted and reduced with the standard HARPS

Data Reduction Software (DRS; Baranne et al., 1996) using a G2 spectral template (mean RV uncertainty ~ 3.4 m s^{-1}). Additionally, the HARPS-N RV measurements were extracted using the SERVAL pipeline (see Section 2.4.2; Zechmeister et al., 2018). The mean RV uncertainty of the SERVAL RVs is ~ 3.0 m s^{-1}. Around 70% of the DRS and SERVAL RVs agree within $1\,\sigma$ and around 95% agree within $2\,\sigma$. Due to the lower RV uncertainties of the SERVAL RV measurements, I use these data throughout the remainder of this chapter. Observations with a S/N < 45 (a total of 4 observations) were not included in any of the analysis that is to follow.

In addition to the HARPS-N observations we obtained 22 spectra between 9 September and 10 October 2020 using the high-resolution (R $\approx 150\,000$) EXtreme PREcision Spectrometer (EXPRES; Jurgenson et al., 2016; Petersburg et al., 2020; Blackman et al., 2020) mounted on the 4.3-m Lowell Discovery Telescope (LDT; Levine et al., 2012), USA. Each spectrum was calibrated using a ThAr lamp and a stabilised Laser Frequency Comb and the RVs were extracted using the EXPRES analysis pipeline (for details see Petersburg et al., 2020). Due to poor seeing and high airmass, 13 of those spectra (with S/N < 45 at 550 nm) were not used for further analysis. The mean S/N and mean RV uncertainty of the used spectra are ~ 82 and ~ 9.5 m s^{-1}, respectively. All HARPS-N and EXPRES RV measurements are listed in Table 5.2.

5.4 Data analysis

5.4.1 Stellar atmospheric parameters

The fundamental stellar parameters of HD 152843, namely the effective temperature ($T_{\rm eff}$), surface gravity ($\log g$), metallicity ([M/H]), projected rotational velocity ($v \sin i_\star$), and microturbulent velocity ($v_{\rm mic}$), were extracted using three independent tools: ARES+MOOG (Sousa, 2014),[3,4] Grid Search in Stellar Parameters (GSSP; Tkachenko, 2015),[5] and Stellar Parameter Classification (SPC; Buchhave et al.,

[3]ARESv2: http://www.astro.up.pt/~sousasag/ares/
[4]MOOG 2017: http://www.as.utexas.edu/~chris/moog.html
[5]GSSP: fys.kuleuven.be/ster/meetings/binary-2015/gssp-software-package

Table 5.2: Radial velocity measurements.

Time (BJD - 2457000)	RV ($m\,s^{-1}$)	σ_{RV} ($m\,s^{-1}$)	S/N	Source	Time (BJD - 2457000)	RV ($m\,s^{-1}$)	σ_{RV} ($m\,s^{-1}$)	S/N	Source
2098.3521	2.87	1.72	155.1	**HARPS-N**	2326.6784	0.55	1.51	136.6	HARPS-N
2101.6407	-3.42	12.10	19.0	EXPRES*	2328.668	4.99	5.58	46.5	HARPS-N
2101.6553	19.11	13.83	16.0	EXPRES*	2330.7365	4.57	2.45	101.2	HARPS-N
2101.6701	-27.31	14.61	14.0	EXPRES*	2338.6434	0.58	2.4	107.0	HARPS-N
2101.6849	-26.58	14.08	14.0	EXPRES*	2342.6342	5.51	2.46	96.8	HARPS-N
2102.3412	-4.28	2.040	117.5	**HARPS-N**	2344.7074	-0.41	3.14	76.8	HARPS-N
2102.6207	11.14	11.95	20.0	EXPRES*	2349.7036	-2.26	7.65	34.7	HARPS-N*
2102.6351	10.94	10.73	21.0	EXPRES*	2358.6608	-3.37	4.45	58.7	HARPS-N
2102.6519	15.83	12.41	20.0	EXPRES*	2359.6465	-10.7	3.77	66.5	HARPS-N
2102.6656	-0.86	11.50	23.0	EXPRES*	2360.6997	-9.31	2.74	94.5	HARPS-N
2102.6843	26.92	11.76	22.0	EXPRES*	2361.6450	-10.44	3.60	88.5	HARPS-N
2102.6999	-24.19	11.24	22.0	EXPRES*	2363.6525	5.29	2.22	119.1	HARPS-N
2102.7140	-15.18	12.81	18.0	EXPRES*	2364.6633	1.23	1.84	133.7	HARPS-N
2102.7312	-43.89	14.50	13.0	EXPRES*	2365.6748	2.40	2.16	116.8	HARPS-N
2104.3651	-20.51	14.3	19.9	HARPS-N*	2366.6619	1.29	1.77	144.5	HARPS-N
2110.3253	-3.73	2.95	86.6	**HARPS-N**	2367.5308	-0.32	1.78	130.1	HARPS-N
2111.3788	0.24	2.19	99.8	**HARPS-N**	2377.6780	5.56	1.89	128.7	HARPS-N
2114.3390	3.36	4.00	64.7	**HARPS-N**	2378.6622	-4.69	2.56	85.4	HARPS-N
2117.3242	2.13	2.55	95.2	**HARPS-N**	2379.5874	-3.44	1.71	136.4	HARPS-N
2119.3255	6.71	2.81	78.5	**HARPS-N**	2380.6749	-4.25	3.63	81.8	HARPS-N
2120.3307	0.85	1.81	126.2	**HARPS-N**	2387.5244	3.84	1.91	113.7	HARPS-N
2120.4134	1.21	2.65	85.9	HARPS-N*	2388.4109	3.03	1.64	137.2	HARPS-N
2120.6143	5.41	4.94	95.0	EXPRES	2390.6104	1.48	2.16	112.8	HARPS-N
2123.5929	-0.23	5.32	81.0	EXPRES	2391.5885	-0.65	2.25	124.3	HARPS-N
2123.6069	0.15	4.97	83.0	EXPRES	2392.5139	-6.96	1.75	136.1	HARPS-N
2125.3192	-1.05	2.64	110.9	**HARPS-N**	2413.5770	-5.23	2.89	77.9	HARPS-N
2126.3165	-7.06	4.00	64.7	**HARPS-N**	2414.4427	-5.45	1.99	119.2	HARPS-N
2126.5970	-1.07	8.91	41.0	EXPRES*	2416.5572	-5.13	2.11	102.7	HARPS-N
2126.6118	11.08	6.87	57.0	EXPRES	2418.4902	-8.05	5.09	48.4	HARPS-N
2127.3185	1.52	3.79	71.8	**HARPS-N**	2427.3813	-0.90	2.25	104.2	HARPS-N
2128.3180	6.51	2.92	87.2	**HARPS-N**	2428.5628	2.97	4.75	49.8	HARPS-N
2129.5838	5.09	5.192	92.0	EXPRES	2430.4746	-2.89	3.41	81.1	HARPS-N
2129.5967	13.15	5.83	64.0	EXPRES	2431.3687	2.3	3.1	70.8	HARPS-N
2130.3156	10.81	5.40	45.3	**HARPS-N**	2443.4123	1.42	3.37	69.1	HARPS-N
2130.5850	9.99	5.18	90.0	EXPRES	2445.4923	1.63	2.55	93.0	HARPS-N
2132.5928	8.59	4.46	114.0	EXPRES	2446.4704	4.29	2.22	123.2	HARPS-N
2132.6078	5.32	4.87	110.0	EXPRES	2447.4741	4.26	2.25	103.8	HARPS-N
2152.2939	0.58	2.57	112.1	**HARPS-N**	2448.4631	-0.06	2.65	105.5	HARPS-N
2153.2902	-16.77	3.14	115.7	**HARPS-N**	2454.4269	-2.6	1.65	134.1	HARPS-N
2154.2902	-10.59	4.03	80.8	**HARPS-N**	2457.4593	1.31	2.33	115.4	HARPS-N
2155.2908	-45.69	10.35	44.0	HARPS-N*	2458.4536	-1.65	2.13	111.7	HARPS-N
2275.7694	-4.13	2.10	113.3	HARPS-N	2459.4462	0.3	2.68	86.5	HARPS-N
2276.7553	-8.64	1.79	118.8	HARPS-N	2461.4596	-5.6	2.31	102.4	HARPS-N
2287.7394	-6.05	2.17	99.4	HARPS-N	2462.4397	-2.5	2.56	111.5	HARPS-N
2288.7463	2.22	2.31	113.7	HARPS-N	2464.4263	-2.07	2.47	96.6	HARPS-N
2290.7012	6.14	2.29	103.0	HARPS-N	2465.4006	3.85	2.73	88.1	HARPS-N
2293.6999	3.55	1.91	115.8	HARPS-N	2473.3893	-2.95	3.95	69.9	HARPS-N
2298.7021	-5.29	2.90	76.6	HARPS-N	2476.3857	-2.77	3.91	68.3	HARPS-N
2299.7614	-2.91	2.08	113.2	HARPS-N	2477.3843	3.94	2.35	86.7	HARPS-N
2305.7326	-3.63	2.00	111.2	HARPS-N	2478.3864	4.05	2.84	75.2	HARPS-N
2323.7425	2.42	1.96	119.2	HARPS-N	2479.3850	-2.57	2.59	92.7	HARPS-N
2324.6729	-0.15	2.42	103.4	HARPS-N	2481.4052	0.0	2.88	65.3	HARPS-N
2325.6564	2.63	1.91	116.8	HARPS-N					

Note – * indicates that the spectrum was not used for further analysis due to low signal to noise (S/N < 45), while the bold font indicates that the spectrum was used to derive the fundamental stellar parameters of HD 152843. The S/N ratios are calculated at 550 nm.

2012, 2014). I note that the analysis in this section was carried out for the planet validation paper Eisner et al. (2021b) and therefore only makes use of the first 16 HARPS-N spectra (with S/N < 45 at 550 nm) that were available at the time of publication. These are highlighted in bold in Table 5.2. Given the high S/N ratio of these initial spectra, I do not expect the results to significantly change with the inclusion of the more recent data.

ARES+MOOG

The ARES+MOOG tool makes use of the 'Equivalent Width (EW)' method, which derives stellar atmospheric parameters by measuring the EWs of the Fe I and Fe II lines (for details see Sousa, 2014). The EW of an absorption line is the width of a rectangle that has the height of the continuum emission and an area equal to the area of the spectral line. Neutral and singly ionised iron lines are used as they are very numerous. The EWs of the spectral lines were automatically extracted from a stacked spectrum of the initial 16 HARPS-N spectra (with S/N > 45), using the ARES2 code (Sousa et al., 2015). The stacked spectrum has a S/N ∼ 350 at 6000 Å. The radiative transfer code MOOG (Sneden, 1973) was then used to convert the EWs into individual line abundances using a given atmospheric model. This was carried out with different atmospheric models until the excitation and ionisation balances are satisfied and the line abundances agree across all of the lines. MOOG assumes local thermodynamic equilibrium (LTE) and uses a grid of ATLAS plane-parallel model atmospheres (Kurucz, 1993), i.e. atmosphere models that ignore the curvature of the star. The method yields the following values: $T_{\rm eff} = 6348 \pm 100$ K, $\log g = 4.31 \pm 0.12$, [Fe/H] = -0.16 ± 0.06, and $v_{\rm mic} = 1.82 \pm 0.13$ km s^{-1}.

GSSP

The GSSP code (Tkachenko, 2015) compares the normalised observed spectrum with a grid of synthetic spectra. Similar to the ARES+MOOG method, a stacked spectrum of the HARPS-N data was used for this analysis. I assessed the goodness of fit of each synthetic spectrum to the stacked spectrum using a χ^2 metric.

The atmospheric models used as part of this code were pre-computed using the LLmodels software (Shulyak et al., 2004), which assumes LTE.

In order to determine the best-fit parameters, the χ^2 value was recorded for each combination of parameters. The projected χ^2 values were then fit with a fourth order polynomial for each parameter in order to determine the global minimum, which corresponds to the value of the best-fit parameter. The uncertainties were taken as the intersection between the polynomial and the $1\,\sigma$ uncertainty limit.

Following the optimisation of the fundamental stellar parameters ($T_{\rm eff}$, $\log g$, [M/H], $v \sin i_\star$, and $v_{\rm mic}$), I fixed the values of [M/H] and $v \sin i_\star$ in order to independently optimise the abundances of Fe, Mg, Ti, Cr, and Ni. The GSSP code allows one abundance to be optimised at a time. However, different spectral lines located at similar wavelengths can affect the measured chemical abundances of one another, known as line blending. As such, the process of optimising each abundance in turn was repeated multiple times until the value of each abundance converged. Convergence – where the deviation in the abundances obtained from two consecutive iterations was less than the uncertainty on the values – was achieved after three iterations. The best-fit spectral model is shown in Figure 5.4.

The following atmospheric parameters were obtained using GSSP: $T_{\rm eff} = 6368 \pm 100$ K, $\log\ g = 4.16 \pm 0.10$, [M/H] = -0.17 $\pm$ 0.05 dex, [Fe/H] = -0.16 $\pm$ 0.05 dex, $v \sin i_\star = 8.56 \pm 0.5$ km s^{-1}, and $v_{\rm mic} = 1.50 \pm 0.15$ km s^{-1}. I note that the derived $v \sin i_\star$ value is not representative of the true rotational velocity of the star; instead, it represents a combined line broadening due to rotation and macroturbulence, where macroturbulence is an ad-hoc velocity field that accounts for atmospheric turbulence on scales larger than the mean free path of a photon. Since I do not rely on the rotation rate of the star in the subsequent analysis, I did not disentangle the effects of rotation and macroturbulent velocity.

SPC

Similarly to GSSP, SPC uses spectral synthesis (for details see Buchhave et al., 2012, 2014), which was independently carried out on each of the initial 16 high S/N HARPS-N spectrum (where S/N > 45). For simplicity, all synthetic spectra were

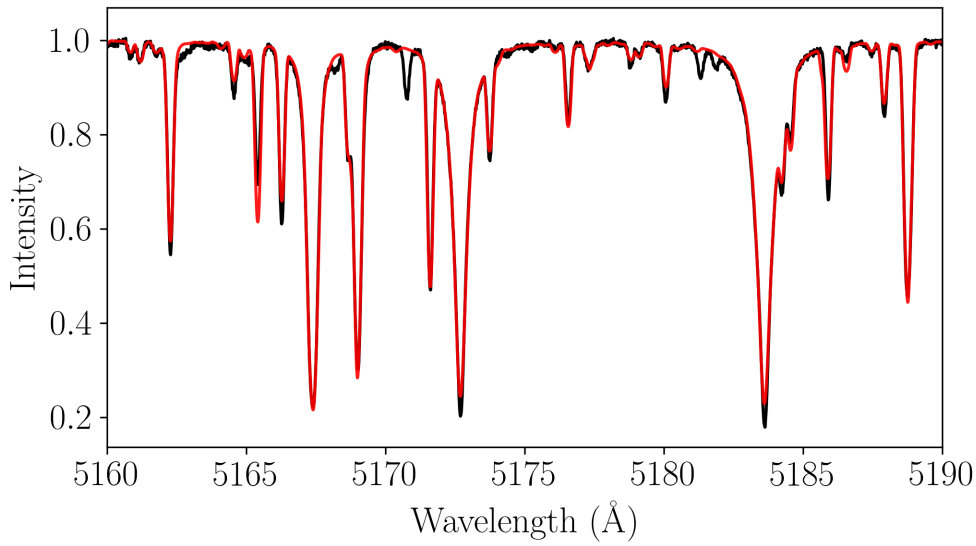

Figure 5.4: Section of the stacked HARPS-N spectra with S/N > 45 (black) and the best-fit model as determined and computed with the GSSP software (red). The parameters and abundances of this best-fit model, combined with the results from the ARES+MOOG and SPC analysis, were used to determine the stellar parameters listed in Table 5.1.

computed with microturbulent and macroturbulent velocities fixed to 2 kms^{-1} and 1 kms^{-1}, respectively. The following values were obtained: $T_{\mathrm{eff}} = 6175 \pm 50$ K, log $g = 4.15 \pm 0.10$, [M/H] = -0.26 ± 0.08 dex, and $v \sin i_{\star} = 8.2 \pm 0.5$ $\mathrm{km\,s}^{-1}$.

The final spectroscopic values, which are calculated as the weighted means of the three methods, are listed in Table 5.1. Taking a weighted mean is common practice for deriving spectroscopic parameters that have been determined with more than one method, or for stars observed with more than one instrument (e.g., Torres et al., 2012; Vanderburg et al., 2017; Mortier et al., 2020; Lacedelli et al., 2021). As each method results in different systematics, the average of the individual results should yield more accurate spectroscopic parameters than the values from just one method. Conservative uncertainties were assumed, taken to be the largest of the three derived uncertainties.

5.4.2 Stellar activity indicators

Magnetic activity of stars manifests itself by producing brighter or darker regions on the surface of the star, such as faculae, starspots, and plages. This affects the

observed stellar spectra and can induce RV signals that mimic those from planets (e.g., Queloz et al., 2001; Figueira et al., 2010; Boisse et al., 2011; Díaz et al., 2016). As such, I investigate the chromospheric stellar activity using the $\log R'_{\rm HK}$ value, a parameter that provides a proxy for the level of magnetic activity on the surface of a star (e.g., Vaughan et al., 1978; Noyes et al., 1984; Saar et al., 1998; Santos et al., 2000; Cincunegui et al., 2007; Santos et al., 2010). In brief, $R'_{\rm HK}$ is the ratio of the emission in the cores of the Ca II H and K lines (at 3933 Å and 3968 Å) to the total bolometric flux of the star, where the Ca II H and K lines probe the temperature of the chromosphere of the star. As magnetic activity primarily heats up the chromosphere (resulting in stronger emission features), the ratio of the intensity of the emission feature in these lines to the bolometric flux probes the magnetic activity of the star.

The $\log R'_{\rm HK}$ parameter was determined for the HARPS-N spectra with S/N > 100 using the calibrations of Noyes et al. (1984). The values were found to range from -4.96 to -4.94 with a mean value of -4.95. This low value suggests that HD 152843 is a quiet star (e.g., Mamajek & Hillenbrand, 2008). I note that there is no correlation between the $\log R'_{\rm HK}$ values and the radial velocities, where a correlation could indicate that the RV signals are due to stellar variability.

5.4.3 SED fitting

As an independent determination of the stellar radius, an analysis of the broadband spectral energy distribution (SED) was performed, following the methodology outlined for TOI 813 b in Section 4.4.2. The available photometric observations (listed in Table 5.1) span the full stellar SED over the wavelength range 0.15 – 22 μm (see Figure 6.9). Integrating the (unreddened) model SED gives the bolometric flux at Earth $F_{\rm bol} = 7.72 \pm 0.18 \times 10^{-9}$ erg s^{-1} cm^{-2}, which corresponds to a stellar radius of $R_\star = 1.42 \pm 0.05$ R$_\odot$. This is consistent with the spectroscopically derived stellar radius listed in Table 5.1.

5.4.4 Stellar mass, radius, age, and distance

The stellar mass, radius, age, and distance were extracted using the ISOCHRONES package (Morton, 2015), which infers model-based stellar properties using stellar evolution models. For this analysis, the ARES+MOOG, GSSP, and SPC effective temperatures and metallicities were used as inputs (Section 5.4.1), along with the *Gaia* eDR3 parallax, and the magnitude of the star in eight bands (Table 5.1). ISOCHRONES was run six times: for each set of spectroscopic parameters (ARES+MOOG, GSSP, and SPC) it was run with two sets of isochrones, where the isochrones used stellar models from the Dartmouth Stellar Evolution Database and from the MESA isochrones and Stellar Tracks (MIST; Choi et al., 2016). The observed properties and their uncertainties were matched against all of the points along a set of synthetic stellar isochrones, where an isochrone is a line of constant age across all evolutionary tracks considered. This resulted in a posterior probability for each parameter. The stellar values were extracted from the combined posteriors of all of the runs, taking the median and the 16th and 84th quantiles as $1\,\sigma$ uncertainties. The stellar mass, radius, density, and age are listed in Table 5.1.

5.4.5 Joint transit and RV modelling

I modelled the transit and RV data using `pyaneti` (see Section 2.3). First, I independently modelled the transit data of each planet (without the RV data). For planet b I modelled both transits simultaneously, allowing me to fit for transit epoch, orbital period, impact factor, scaled planet radius, and scaled semi-major axis. The single-transit event (planet c) was modelled by fitting for the same parameters as for planet b, with the exception of the orbital period and scaled semi-major axis, as these cannot be constrained in the case of a single-transit event. Instead, the model derived a possible period range of 13 to 35 days at the 99% confidence interval, using the relations presented in Osborn et al. (2016) and assuming a circular orbit (see Section 2.3.4). The results from the transit modelling were used to create uniform priors for all the transit model parameters for the subsequent joint RV and transit analysis.

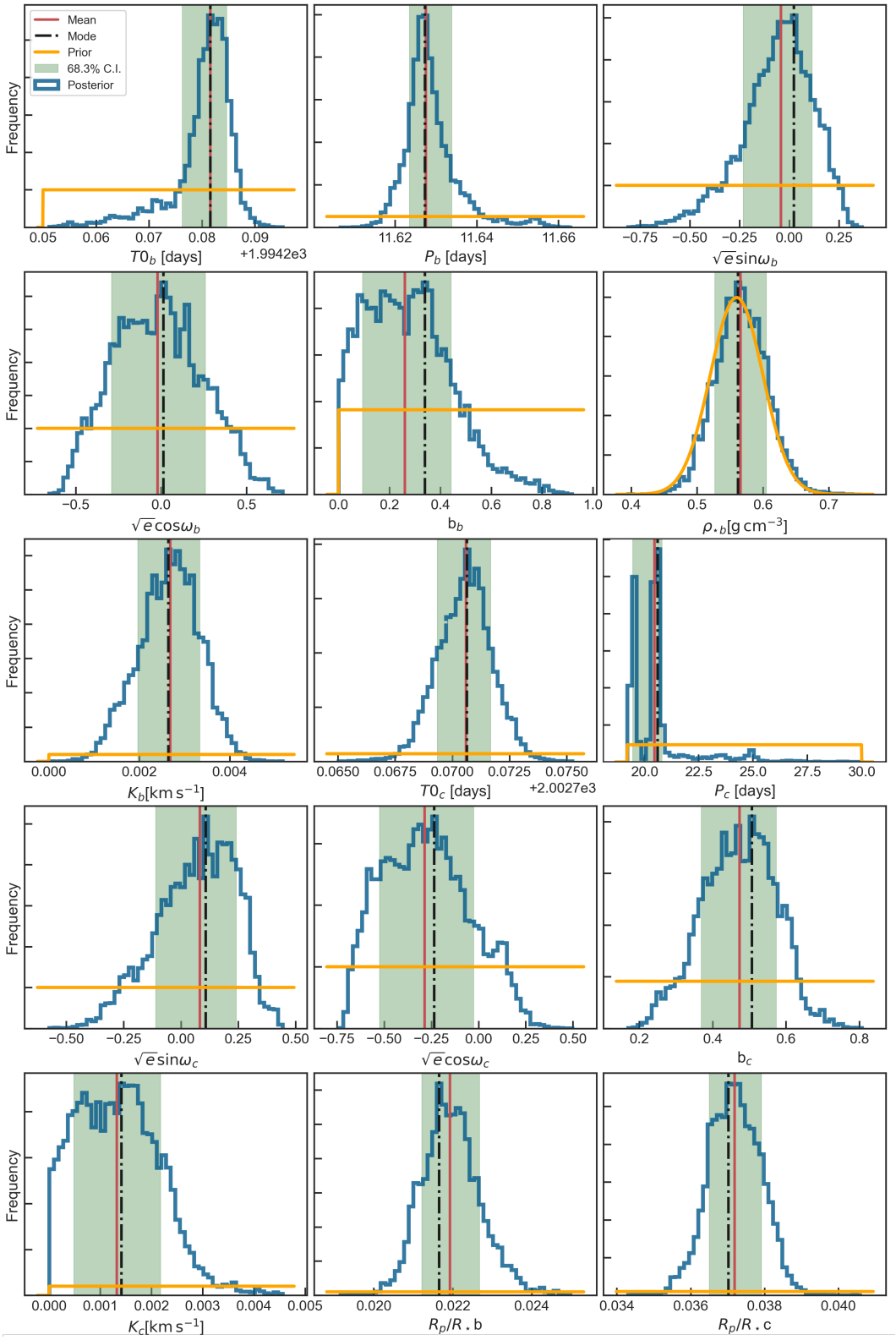

Figure 5.5: Posterior distribution of the joint transit and RV fit for HD 152843. The blue lines show the posterior distribution ($P(M|d)$), the orange region shows the prior ($P(M)$), the green shaded region shows the 68% credible interval, the black dashed line the mode and the red solid line the median.

Table 5.3: System parameters of HD 152843.

Parameter	Symbol	Prior[(a)]	Value[(b)]	Units	Comments
Model Parameters for HD 152843b					
Orbital period	P_b	$\mathcal{U}[11.5, 11.7]$	$11.6264^{+0.0047}_{-0.0041}$	days	
Transit epoch	$T_{0,b}$	$\mathcal{U}[1994.25, 1994.30]$	$1994.2831^{+0.0027}_{-0.0036}$	BJD – 2457000	
Eccentricity parametrisation	$e\sin\omega$	$\mathcal{U}[-1, 1]$	$-0.03^{+0.15}_{-0.19}$	–	Code ensures $e < 1$
Eccentricity parametrisation	$e\cos\omega$	$\mathcal{U}[-1, 1]$	$0.04^{+0.23}_{-0.26}$	–	Code ensures $e < 1$
Scaled planet radius	$R_b/R_\star$	$\mathcal{U}[0, 0.1]$	$0.02188^{+0.00069}_{-0.0007}$	–	
Impact parameter	b_b	$\mathcal{U}[0, 1.1]$	$0.23^{+0.19}_{-0.15}$	–	
Doppler semi-amplitude	K_b	$\mathcal{U}[0, 50]$	$2.66^{+0.8}_{-0.77}$	$\mathrm{m\,s^{-1}}$	
Model Parameters for HD 152843c					
Orbital period	P_c	$\mathcal{U}[19.26, 35]$	$20.5^{+1.4}_{-0.9}$	days	Truncated posterior
Transit epoch	$T_{0,c}$	$\mathcal{U}[2002.73, 2002.8]$	$2002.7705^{+0.0013}_{-0.0012}$	BJD – 2457000	
Eccentricity parametrisation	$e\sin\omega$	$\mathcal{U}[-1, 1]$	$-0.01^{+0.2}_{-0.21}$	–	Code ensures $e < 1$
Eccentricity parametrisation	$e\cos\omega$	$\mathcal{U}[-1, 1]$	$-0.11^{+0.37}_{-0.35}$	–	Code ensures $e < 1$
Scaled planet radius	$R_c/R_\star$	$\mathcal{U}[0, 0.1]$	$0.03744^{+0.00081}_{-0.00079}$	–	
Impact parameter	b_c	$\mathcal{U}[0, 1.1]$	$0.518^{+0.097}_{-0.109}$	–	
Doppler semi-amplitude	K_c	$\mathcal{U}[0, 50]$	3.5	$\mathrm{m\,s^{-1}}$	Upper limit*
Other Parameters					
Stellar density	$\rho_\star$	$\mathcal{N}[0.56, 0.04]$	$0.57^{+0.04}_{-0.04}$	$\mathrm{g\,cm^{-3}}$	
Limb-darkening coefficient	q_1	$\mathcal{U}[0, 1]$	$0.155^{+0.142}_{-0.076}$	–	As in Kipping (2013)
Limb-darkening coefficient	q_2	$\mathcal{U}[0, 1]$	$0.54^{+0.3}_{-0.34}$	–	As in Kipping (2013)
Offset velocity HARPS-N	–	$\mathcal{U}[-0.50, 0.50]$	$-0.00049^{+0.00048}_{-0.0005}$	$\mathrm{km\,s^{-1}}$	
Offset velocity EXPRES	–	$\mathcal{U}[-0.50, 0.50]$	$0.0015^{+0.0019}_{-0.0019}$	$\mathrm{km\,s^{-1}}$	
Jitter HARPS-N	–	$\mathcal{U}[0, 100]$	$2.98^{+0.57}_{-0.49}$	$\mathrm{km\,s^{-1}}$	
Jitter EXPRES	–	$\mathcal{U}[0, 100]$	$0.98^{+1.6}_{-0.76}$	$\mathrm{km\,s^{-1}}$	
Jitter TESS	–	$\mathcal{U}[0, 500]$	$0.000148^{+1.6e-05}_{-1.6e-05}$	ppm	
Derived parameters HD 152843b					
Planet mass	M_b	–	$10.27^{+3.09}_{-3.00}$	$\mathrm{M_\oplus}$	
Planet radius	R_b	–	$3.41^{+0.12}_{-0.12}$	$\mathrm{R_\oplus}$	
Planet density	ρ_b	–	$1.44^{+0.43}_{-0.44}$	$\mathrm{g\,cm^{-3}}$	
Semi-major axis	a_b	–	$0.1062^{+0.0027}_{-0.0028}$	AU	
Eccentricity	e_b	–	$0.064^{+0.092}_{-0.045}$	–	Upper limit of 0.40*
Transit duration	τ_b	–	$5.51^{+0.11}_{-0.12}$	hours	
Orbit inclination	i_b	–	$89.16^{+0.54}_{-0.64}$	deg	
Insolation	F_b	–	$259.0^{+21.0}_{-19.0}$	$\mathrm{F_\oplus}$	
Derived parameters HD 152843c					
Planet mass	M_c	–	16.4	$\mathrm{M_\oplus}$	Upper limit*
Planet radius	R_c	–	$5.84^{+0.15}_{-0.15}$	$\mathrm{R_\oplus}$	
Planet density	ρ_c	–	0.66	$\mathrm{g\,cm^{-3}}$	Upper limit*
Eccentricity	e_c	–	$0.114^{+0.181}_{-0.081}$	AU	Upper limit of 0.51*
Transit duration	τ_c	–	$6.363^{+0.087}_{-0.077}$	hours	
Orbit inclination	i_c	–	$88.85^{+0.17}_{-0.16}$	deg	

Note – [(a)] $\mathcal{U}[a, b]$ refers to uniform priors between a and b, $\mathcal{N}[a, b]$ to Gaussian priors with mean a and standard deviation b. * Indicates the upper limit at the 99% confidence interval of the posterior. [(b)] Inferred parameters and errors are defined as the median and 68.3% credible interval of the posterior distribution.

All fitted parameters and priors used for the joint modelling are presented in Table 5.3. The code sampled for the stellar density ($\rho_\star$), and recovered the scaled semi-major axis for each planet in the system using Kepler's third law. A Gaussian prior was used on $\rho_\star$, based on the stellar mass and radius derived in Section 5.4.4. I note that because planet c only exhibits a single transit event a wide uniform prior ([19.26,35]) was used on the period, based on the results from the single-transit analysis. The lower period limit was truncated at 19.26 days as a shorter orbital period would have necessarily resulted in further transit events within the *TESS* light curve.

The parameter space was sampled with 500 independent chains and posterior distributions created using 5000 iterations of converged chains with a thin factor of 10. This generated a posterior distribution made with 250 000 independent samples for each parameter. The fitted parameters extracted from the posteriors can be found in Table 5.3. I note that the model and data only weakly constrain the orbital period of HD 152843 c (P_c). Furthermore, the posterior distributions for the semi-amplitude of planet c (K_c) are truncated at zero. All posteriors are shown in Figure 5.5, while the derived transit and RV models are shown in Figures 5.6 and 5.7, respectively.

5.4.6 Statistical validation

For completeness, I used the open-source python package `VESPA` to calculate the statistical false positive probability (FPP) of both planet candidates (Morton, 2012, 2015; Morton et al., 2016). This analysis followed the methodology outlined for TOI 813 b presented in Section 4.4.3.

The analysis showed that HD 152843 b and HD 152843 c have FPPs of 0.05 and < 0.001, respectively, meaning that only HD 152843 c is below the traditionally required threshold of FPP < 0.01 (Morton et al., 2016; Crossfield et al., 2016). Even though HD 152843 b is not below the threshold, I note that the `VESPA` model does not consider multiplicity in planet systems, which has been shown to decrease the FPP by at least an order of magnitude (Lissauer et al., 2011, 2012, 2014). Lissauer et al. (2012), for example, estimated that systems with two or more planets

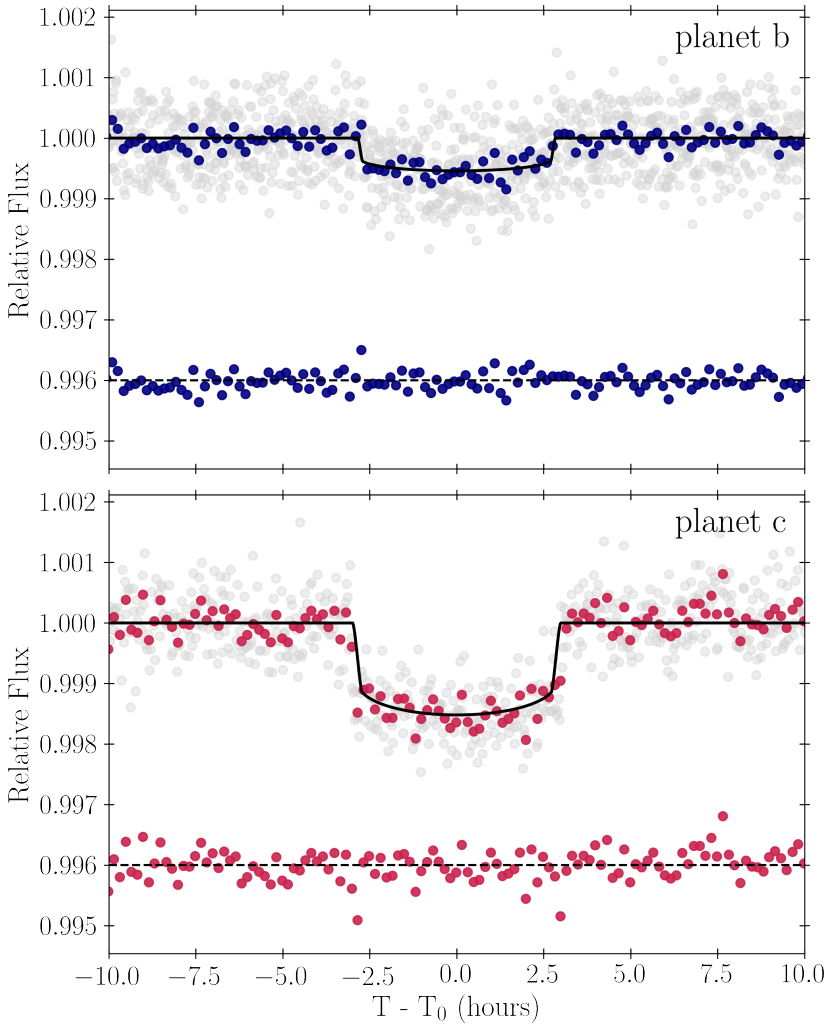

Figure 5.6: Phase-folded *TESS* light curve of HD 152843 b (upper panel) and HD 152843 c (lower panel). Nominal *TESS* data are shown in light gray together with 10-min binned data in the solid colours. The inferred transit model for each planet is over-plotted with a solid black line. The residuals of the fits are shown underneath the transit data.

in the *Kepler* data were 25 times less likely to be false positives. Furthermore, the derived upper mass limits of both planets enabled me to rule out that the events are caused by an eclipsing binary. As such, I consider both HD 152843 b and HD 152843 c statistically validated.

5.5 Results and discussion

The joint RV and transit modelling showed that the inner and outer planets have radii of $R_b = 3.41^{+0.12}_{-0.12}$ $R_\oplus$ and $R_c = 5.84^{+0.15}_{-0.15}$ $R_\oplus$, respectively. The radial

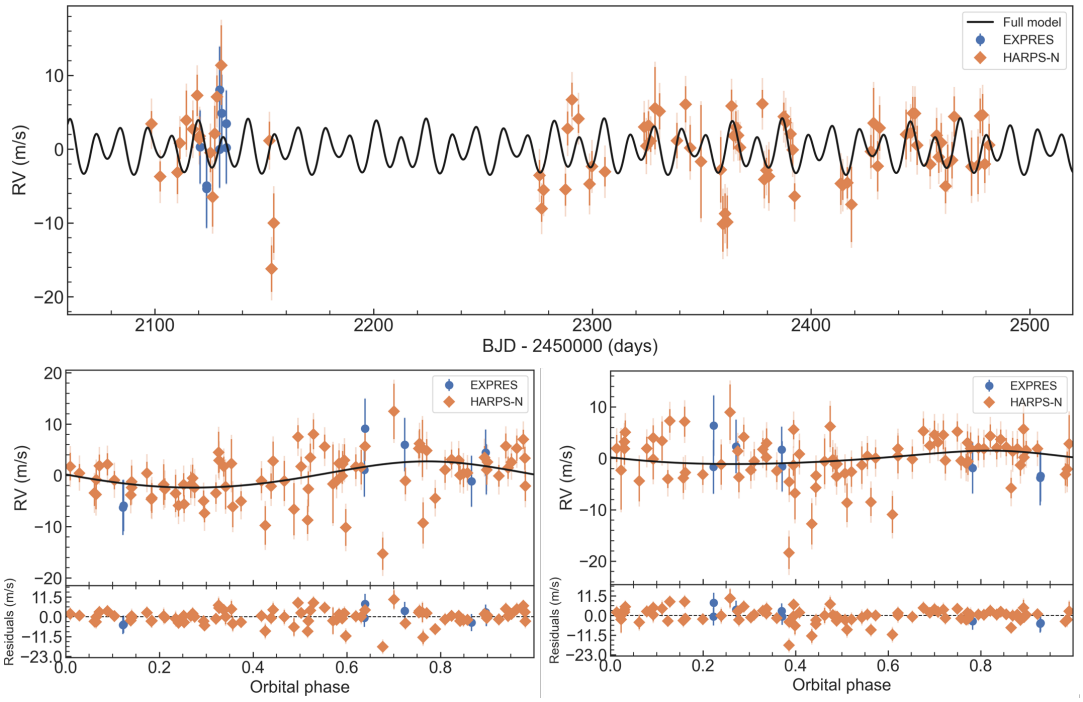

Figure 5.7: RV time-series (upper panel) and phase-folded RV plots for HD 152843 b (lower left panel) and HD 152843 c (lower right panel) following the subtraction of the instrumental offsets. HD 152843 c plot has been phase folded using a period of 20.5 days. HARPS-N (orange diamonds) and EXPRES (blue circles) RV measurements along with their nominal uncertainties are shown in each panel. The lighter coloured error bars mark the errors including jitter. Solid black lines show the respective inferred model.

velocity measurements allowed me to constrain the mass of the innermost planet to $M_b = 10.27^{+3.09}_{-3.00}$ $M_\oplus$ and derive an upper mass limit of the outer planet of $M_c < 16.4$ $M_\oplus$ (at the 99% confidence interval). The $3\,\sigma$ detection of the semi-amplitude of the inner planet confirms the planetary nature of HD 152843 b, while the derived upper mass limit of the outer planet allows me to confirm that the corresponding transit signals seen in the *TESS* light curve is not the result of an eclipsing binary.

As shown in Figure 5.5, the joint modelling of the light curve and the RVs produce a truncated posterior distribution for P_c, where the distribution favours orbital periods shorter than 21 days. While I can rule out orbital periods shorter than 19.26 days due to the lack of a second transit in the *TESS* data, it is possible that HD 152843 c has an orbital period of, or close to, 19.375 days, which would be a 5:3 mean motion resonance (MMR) with the orbit of HD 152843 b

($P_b = 11.6264^{+0.0047}_{-0.0041}$ days). The effects of resonances in multi-planet systems is further discussed in Section 5.5.1.

In order to place HD 152843 into a wider context, Figure 5.8 shows the position of planet b and c in the radius-insolation diagram alongside all known exoplanets (grey points) listed in the NASA Exoplanet Archive as of 19 January 2022. Multi-planet systems with measured masses around stars brighter than $V = 10$ are shown by the orange circles, while the black lines connect planets from the same system. HD 152843 b and HD 152843 c are depicted by the blue triangle and pink square, respectively. The figure highlights a noticeable lack of well characterised multi-planet systems around bright stars, which are key for comparative atmospheric studies. Furthermore, it shows that planet c lies in a sparsely populated region of parameter space.

The two planets also stand out in terms of their bulk densities. Given the minimum radius and upper mass limit of HD 152843 c, this planet has a density < 0.66 g cm^{-3}, while planet HD 152843 b has a density of $1.44^{+0.43}_{-0.44}$ g cm^{-3}. This suggests that both planets have extended gaseous envelopes, making both planets prime candidates for atmospheric characterisation, as discussed further in Section 5.5.3.

One possible explanation for the estimated low density of planet c is that it formed at a greater distance from the host star prior to migrating to its current orbit. This would have allowed the planet to accrete a significant H/He envelope, due to the colder and less dense gas present farther away from the host star. Planets that undergo this type of migration are often found to be the outer planets in MMR chains (Lee & Chiang, 2016).

Alternatively, the two planets could have formed *in situ* with their differing planet properties (i.e. mass and radius) resulting from subsequently diverging evolutionary pathways. For example, extreme ultraviolet irradiation from the host star could have triggered atmospheric loss through photoevaporation of the inner planet (Owen & Wu, 2016; Chen & Rogers, 2016), stripping it of its extended gaseous envelope, while the outer planet could have been inflated, resulting in the observed low density of planet c.

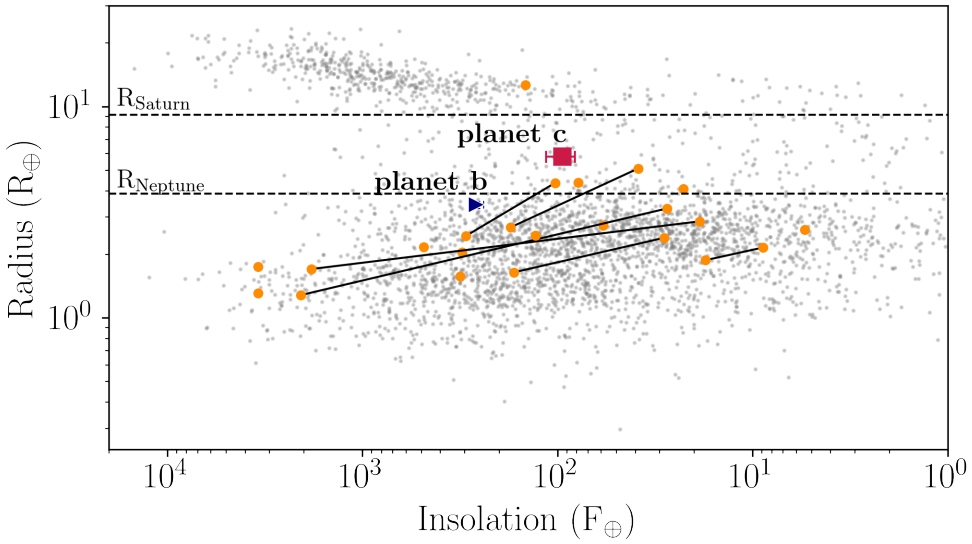

Figure 5.8: Planet insolation–radius diagram of confirmed exoplanets from the NASA Exoplanet Archive (grey points). Orange points show members of systems with more than one planet, with mass measurements better than 30% and around stars brighter than $V = 10$. The black lines connect planets that are within the same system. Planets that are not connected by a black line are in multi-systems where only one planet has a mass measurement with better than 30% accuracy. HD 152843 b and HD 152843 c are shown by the blue triangle and pink square, respectively.

Theory also suggests that the low density of the planets could be due to tidal heating, which could result in an increase in entropy (e.g., Millholland, 2019) and thus an inflated radius. Finally, Wang & Dai (2019) and Gao & Zhang (2020) independently suggest that the apparent radii could be enhanced by photochemical hazes in the atmospheres, resulting in an underestimate of the densities of the planets. Future transmission spectra of planet c, for example at mid-infrared wavelengths where the atmosphere is less affected by hazes, would allow for differentiation between different formation scenarios and therefore provide useful constrain how these planets formed and evolved.

5.5.1 Prospects for transit timing variations

Transit Timing Variations (TTVs) are often observable in multi-planet systems due to dynamical interactions between the planets, as predicted by Agol et al. (2005) and Holman & Murray (2005). As previously discussed in Section 1.6.1, this is especially the case when planets are in near orbital resonance, which is potentially true for the

planets around HD 152843. Measuring TTVs, especially when combined with RV data, allows for the refinement of the planetary mass and orbital parameters, which are critical for interpreting atmospheric transmission spectra in smaller planets (Batalha et al., 2019). It can also enable the detection of inclined non-transiting planets and can therefore lend insight into system demographics and architectures (Brakensiek & Ragozzine, 2016).

TTVs were assessed for this system using the best-fit planetary parameters across a range of mass, period, and eccentricity solutions using the TTVFast framework of n-body simulations (Deck et al., 2014). The analysis showed that a 2:1 resonance would result in TTVs with an amplitude ranging from 5 – 40 minutes, and a super-period of approximately 2 – 3 years, allowing for follow-up observations to detect discernible TTVs on the scale of about a year. Such observations would be feasible with instruments such as CHEOPS (Broeg et al., 2013). Furthermore, this amplitude would be greatly increased for non-zero eccentricities. Similar TTVs would be observed for a 5:3 resonance; however, significant TTVs would not be observed away from resonance. Follow-up studies of this system will be able to help constrain the planetary masses, eccentricities, and other orbital parameters, given both the presence or absence of significant TTVs.

5.5.2 Rossiter-McLaughlin effect prospects

The moderate projected rotational velocity of HD 152843 ($v \sin i \sim 8.2$ km s^{-1}) makes it a good candidate for studying the Rossiter-McLaughlin effect (RM; Rossiter, 1924; McLaughlin, 1924), which provides an estimate of the spin-orbit alignment of the orbiting planets with the host star (e.g., Schneider, 2000). As discussed in Section 4.5.5, the RM effect helps to shed light onto the dynamical history of the system, as mechanisms such as planet-disk interactions help to preserve the initial spin-orbit alignment, while planet-planet interactions promote misalignment (e.g., Chatterjee et al., 2008; Deeg et al., 2009; Storch et al., 2017). To date, the number of multi-planet systems with measured obliquities remains small (e.g., Dalal et al., 2019; Hjorth et al., 2021). In a recent review, Wang et al. (2022)

discuss the implications of the spin-orbit alignment of the, to date, 13 compact multi-planet systems with measured obliquities.

I estimate the RM effect to be $3.72^{+0.96}_{-0.74}$ m s^{-1} and $9.60^{+2.65}_{-2.76}$ m s^{-1} for HD 152843 b and c, respectively (Winn, 2010). Future precision RV observations (for example typical precisions of 4 m s^{-1} are possible for this target with HARPS-N) will be able to detect the RM of planet c, thus allowing for the determination of the true obliquity of the target.

5.5.3 Feasibility of atmospheric characterisation

The brightness of HD 152843 ($V = 8.855$), combined with the large radii of the planets, as shown in Figure 5.8, make them likely future targets for atmospheric characterisation via transmission spectroscopy. I assess the feasibility of such an observation using the transmission spectroscopy metric (TSM; Kempton et al., 2018), which provides the estimated S/N of a 10 hour observation with JWST/NIRISS, as previously discussed in Section 4.5.5. Based on planetary masses of 10.27 $M_\oplus$ and 16.4 $M_\oplus$ (Table 5.3), I find the TSM to be 72 and 155, for HD 152843 b and HD 152843 c, respectively. The TSM for planet c compares well with several of the targets currently included in JWST ERS (Early Release Science) and GTO (Guaranteed Time Observations) programs, and is better than the cut-off thresholds for follow-up observations of TSM = 96 as suggested by Kempton et al. (2018). The TSM of 155 places planet c at least amongst the top 50% of candidates suitable for atmospheric characterisation as outlined by Kempton et al. (2018). Furthermore, as the planet mass used to determine this value is an upper mass limit, the TSM of planet c is likely to be significantly higher (planet mass scales inversely with the TSM), likely placing it amongst the top 25% of candidates best suited for atmospheric characterisation. In order to carry out such observations, the orbital period of planet c needs to be better constrained with, for example, the observation of additional transit events or with continued RV monitoring.

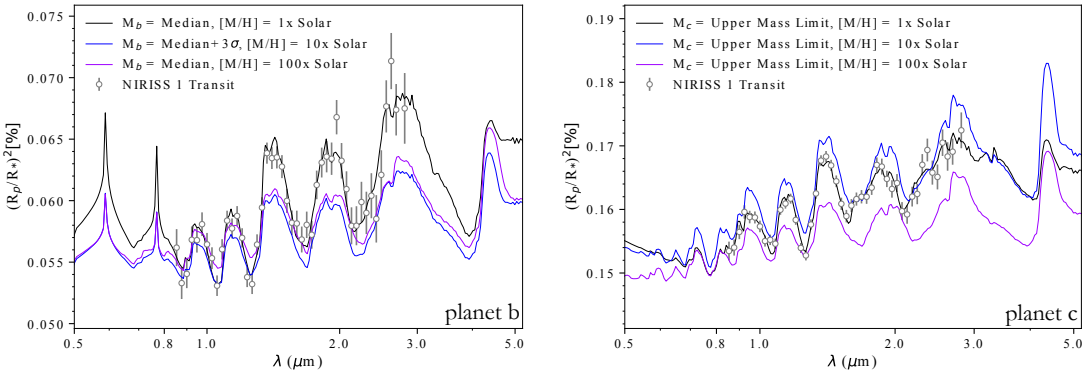

Figure 5.9: Atmospheric models generated for planets b and c in the left and right panels, respectively. Each panel shows three models describing plausible atmospheric scenarios. The black model shows an atmospheric model which has a metallicity of 1× solar, considering the RV extracted median mass and upper mass limit for planets b and c, respectively. The purple model shows an atmospheric model which has a metallicity of 100× solar considering the RV extracted median mass and upper mass limit for planets b and c, respectively. The blue model shows an atmospheric model which has a metallicity of 10× solar, considering the RV extracted median mass plus the 3 σ upper uncertainty for planet b and the upper mass limit for planet c. The simulated JWST NIRISS/SOSS observations for the 1× solar case are shown to emphasise the precision one would be able to obtain from a single transit observation.

5.5.4 Atmospheric modelling

To assess the possibility of differentiating between different atmospheric scenarios, an array of forward models was generated using the open-source code CHIMERA (Line et al., 2013) and compared to synthetic observations of each planet, which were generated using PANDEXO (Batalha et al., 2017) for a single transit observation using JWST NIRISS/SOSS (Figure 5.9). For each planet a cloud free atmosphere was modelled with an isothermal temperature profile set to the derived temperatures presented in Table 5.1. The upper mass limit of 16.4 $M_\oplus$ was used for planet c, while for planet b two mass scenarios were considered: the median mass and the median mass plus the 3 σ uncertainty. The upper mass limits were used as lower masses would equate to lower bulk densities (more extended atmospheres) that would result in more prominent observable features, and thus would be easier to observe. The atmospheres were modelled to have a solar C/O ratio and metalicities of 1×, 10×, and 100× solar. Figure 5.9 shows three models for each planet: the mass and 1× solar metallicity (black), the mass and 100× solar (purple), and

the mass plus $3\,\sigma$ and $10\times$ metallicity (blue). The mean mass and upper mass limits were used for planets b and c, respectively. The black data points show the predicted observations obtained using JWST NIRISS/SOSS, generated using the $1\times$ solar median mass models.

The left panel, corresponding to planet b, shows that even though it is possible to detect the atmosphere with observations of a single transit, there remains a degeneracy between the metallicity and the mass of the planet. Future RV follow-up observations will be able to break this degeneracy by refining the mass estimate of planet b. The right panel, corresponding to planet c, shows that the simulated data have extremely small uncertainties, due to the bright star and long transit duration. These small uncertainties allow us to break the degeneracy between planetary mass and atmospheric metallicity. These simulations emphasise how promising these planets are for follow-up measurements and atmospheric characterisation.

5.6 Summary and conclusions

In this Chapter I presented the discovery of a multi-planet system with a Neptune and a sub-Saturn sized planet ($R_b = 3.41^{+0.12}_{-0.12}$ R$_\oplus$; $R_c = 5.84^{+0.15}_{-0.15}$ R$_\oplus$). Ground-based spectroscopic follow-up observations with HARPS-N and EXPRES allowed me to constrain the orbit and planet parameters as well as to refine the stellar properties. Joint modelling of the light curve and RVs enabled me to determine the mass of the inner planet, $M_b = 10.27^{+3.09}_{-3.00}$ M$_\oplus$ and place an upper mass limit on the outer planet, $M_c < 16.4$ M$_\oplus$. The properties of HD 152843 c, which likely has an extended H/He atmosphere, combined with the brightness of the host star make it a promising target for future atmospheric characterisation. Furthermore, both planets are promising targets for measurements of the RM effect, which will help constrain the formation and evolution history of this system.

Continued efforts with HARPS-N and EXPRES will be able to conclusively determine the masses of both planets and the orbital period of planet c, as well as search for the RM effect. Ground-based spectroscopic observations are ongoing with TNG/HARPS-N. Additionally, photometric observations with, for example,

LCO/Sinistro (Brown et al., 2013), CHEOPS (Broeg et al., 2013), or the upcoming PLATO mission (Rauer et al., 2014), could enable observations of further transit events and constrain possible TTVs. HD 152843 is also scheduled to be re-observed by the *TESS* mission during Sector 52 (May – June 2022). Observations of additional transits will help to refine the ephemerides of both planets, which will be crucial to carry out transmission spectroscopy with, for example, JWST or ARIEL (Tinetti et al., 2016).

Finally, the discovery and validation of the HD 152843 planet system, as well as of TOI 813 b, highlights the value of citizen science to identify scientifically valuable planet systems that are missed by other searches. In particular, HD 152843 is a good case study to show that the PHT pipeline, combined with the `LATTE` vetting tool, `pyaneti` system modelling software, and our team's access to telescope time, allows us to competitively identify, validate, and publish new planet systems.

Ghosts are transparent.

— S. Snape, 1996

CHAPTER 6

TIC 470710327: A COMPACT HIERACHICAL TRIPLE STAR SYSTEM

Planet Hunters TESS was designed such that the opportunity to find interesting and unique objects beyond planetary systems is a natural by-product. These objects include stellar binaries, multi-stellar systems, flaring stars, dwarf novae, and pulsating stars. In this Chapter, I present the in-depth analysis and discussion of a triple star system that was flagged on the PHT discussion forum and followed-up due to its unique configuration, as well as puzzling formation history and future evolution.

6.1 Context

Massive stars ($M_{init} \geq 8\,M_{\odot}$) are intrinsically rare due to the initial mass function (IMF – the underlying distribution of stellar birth masses; see e.g., Salpeter, 1955; Bastian et al., 2010; Dib et al., 2017) and due to their short lifetimes (e.g., Kippenhahn et al., 2012). However, despite their rarity, massive stars drive evolution on a wide range of scales by providing radiative, dynamical, and chemical feedback to their environment (Langer, 2012). For example, due to their high luminosities, massive stars irradiate the gas in their immediate surroundings, resulting in phenomena such as ionisation nebulae (e.g., Smith, 2006; Simón-Díaz & Stasińska, 2008). Furthermore, when massive stars go supernova, they can chemically enrich

their local star forming region (Zhang et al., 2015; Welsh et al., 2019), clear dust from their surroundings (Lakićević et al., 2015; Priestley et al., 2021), and trigger large-scale turbulence in the interstellar medium which affects star formation rates (Hennebelle & Commerçon, 2014; Iffrig & Hennebelle, 2015; Bacchini et al., 2020).

Despite their importance for shaping the dynamical and chemical properties of Galaxies, the physical processes responsible for the formation, evolution, and death of massive stars are not well understood (e.g., Zinnecker & Yorke, 2007; Tan et al., 2014; Hennebelle & Commerçon, 2014; Louvet, 2018). The study of these processes is further complicated by the fact that such stars often have nearby stellar companions (Sana et al., 2012), which can affect their properties and evolution at all stages of their lives. Large scale spectroscopic, interferometric, and high contrast imaging surveys of OB stars have demonstrated that most, if not all, massive stars are formed in a binary or higher order multiple system (Kiminki & Kobulnicky, 2012; Sana et al., 2013, 2014; Aldoretta et al., 2015; Moe & Di Stefano, 2017; Maíz Apellániz et al., 2019; Rainot et al., 2020; Bodensteiner et al., 2021).

The identification and characterisation of massive stars in multiple systems is crucial for discriminating between different formation and evolution scenarios, to place constraints on theoretical models, and to understand the interactions and effects of multiple stars (e.g., Eldridge et al., 2018; Modjaz et al., 2019; Vigna-Gómez et al., 2022). Sana et al. (2012) and de Mink et al. (2014) have already shown that the complex interactions between tides, angular momentum exchange, and stellar evolution in binary and higher order multiples fundamentally affect the evolution and thus the end-product of nearly 70% of all early-type massive stars. For example, in binary systems with orbital periods ranging from a few hours to a few days (compact binary orbits) processes including mass transfer, exchange of angular momentum, and stellar mergers can open up new evolutionary pathways to produce end-products such as X-ray binaries, γ-ray bursts, stellar mergers, and gravitational wave events (Sana et al., 2012; de Mink et al., 2013).

The addition of a nearby third body further complicates the evolution. Interactions between three stars can induce different evolutionary pathways through,

for example, von Zeipel-Kozai-Lidov cycles (ZKL; von Zeipel, 1910; Kozai, 1962; Lidov, 1962; Naoz, 2016; Ito & Ohtsuka, 2019), a phenomenon where a tertiary star perturbs the orbit of an inner binary. This perturbation can lead to modulations in the mutual inclination and eccentricity of the orbit of the binary, such that the eccentricity can increase and decrease on time scales shorter than the evolutionary timescales of the stars. In turn, these high eccentricities result in small periastron distances, which can lead to tidal dissipation that can cause the inner orbit to shrink.

Interactions between three stars can also result in tertiary driven mass transfer to the binary that can lead to mergers or exotic common envelope systems. Such common envelope systems form when the Roche lobes (the largest distance from the star at which a test particle is still gravitationally bound to the star) of all the stars in a system are filled due to rapid mass transfer i.e. mass transfer that is more rapid than the rate at which the mass can be accreted onto the star. Figure 6.1 illustrates how a binary or a triple star system can become a common envelope system. Alternatively the interactions of a triple system can lead to close double or triple degenerate systems (i.e. systems containing multiple neutron stars or black holes), or contribute to the population of walk-away and run-away systems (i.e. stellar objects expelled from their systems at varying speeds) in our galaxy (e.g., Antonini et al., 2017; Renzo et al., 2019; Stephan et al., 2019; Leigh et al., 2020; Hamers et al., 2021b; Glanz & Perets, 2021).

Due to the high intrinsic brightness of massive stars, detecting and characterising non-eclipsing lower mass close companions using spectroscopy alone is challenging. Detailed characterisation of close companions, therefore, often relies on the detection of eclipses or other dynamical effects caused by the presence of a close companion, such as using observations of light travel time effects via eclipse timing variations (ETVs; see Section 6.3.3). Observations of the latter were made possible with the advent of space-based photometric surveys with long time-base, high-precision, and high-cadence observations such as CoRoT, *Kepler*, *K*2, and *TESS*. These space-based missions have enabled the detection of dozens of new triple and higher-order multiple systems through ETVs or multiply eclipsing events (Conroy et al., 2014;

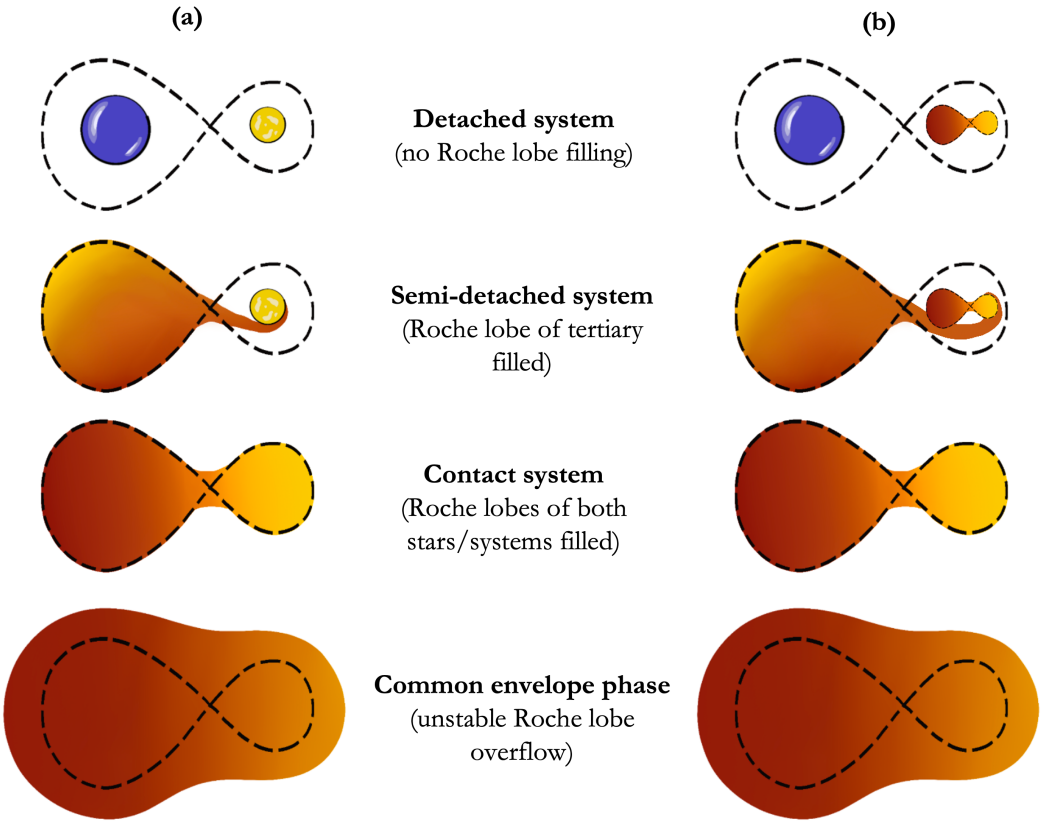

Figure 6.1: Schematic showing Roche lobe overflow for a binary (a) and for a triple (b) system. In the example of the triple system, the Roche lobes of the inner binary are already filled.

Marsh et al., 2014; Borkovits et al., 2015; Rappaport et al., 2017; Hajdu et al., 2017; Sriram et al., 2018; Li et al., 2018; Borkovits et al., 2020a). Importantly, detailed modelling of systems with observable light travel time effects can lead to the determination of the absolute masses of the stellar components in some cases, or minimally the derivation of mass ratios, even in the absence of eclipses of the third body (Borkovits et al., 2016).

In the case of hierarchical triple systems, further considerations beyond light travel time effects need to be taken into account. Hierarchical triple systems consist of three stars, where the mutual distance between two of the three stars is significantly smaller than the distance between either of the two close stars and the third star. In such a configuration the dynamics of the systems can be approximated as two binary systems, with a close binary ('inner' binary) and a wide binary ('outer' binary) which consists of the tertiary body and the centre of mass of the inner binary.

Such a configuration can result in direct perturbations to the period and eccentricity of the inner binary. If the orbital periods and the separations involved are short enough, these effects can occur on observable timescales and can be studied with ETVs (caused by light travel time effects and direct third-body perturbations) and RV observations. Measurements of both ETVs and RVs allows for the determination of precise stellar mass ratios and orbital parameters, including the mutual inclination between the inner and the outer orbit of the triple, which is thought to be indicative of the formation history (e.g., Toonen et al., 2016; Borkovits et al., 2020a).

In this chapter I present a compact, hierarchical triple system identified on the PHT platform, consisting of one O- and two B-type stars. The system, which shows large ETV and RV variations, contains an inner $\sim$1.10-day eclipsing binary and a massive O9.5 – B0.5V tertiary orbiting a common centre of mass on a $\sim$52.0-day orbit.

6.2 The target and its surroundings

6.2.1 Early characterisation of TIC 470710327

TIC 470710327 (BD+61 2536, TYC 4285-3758-1, $V = 9.6$ mag, parallax $= 1.06 \pm 0.24$ mas, distance $= 950 \pm 220$ pc) was initially identified as an early B-type star by Brodskaya (1953). Between 2011 and 2013, the target was further observed as part of a photometric survey of bright stars in OB associations, using the 0.25-m Takahashi Epsilon telescope in Mayhill, New Mexico, USA (Laur et al., 2017). These observations showed that TIC 470710327 is a short period eclipsing binary with a period of 1.1047 days and that is not a known member of a cluster or OB association. Two archival epochs of speckle interferometric measurements were obtained with the Mount Wilson 2.5-m Hooker telescope in California and with the 1.55-m Kaj Strand telescope in Flagstaff, both with the US Naval Observatory speckle camera. The observations revealed a close companion at $\sim$0.5 arcsec, with the position angle and distance of the companion advancing from $\theta = 306.2°$ and

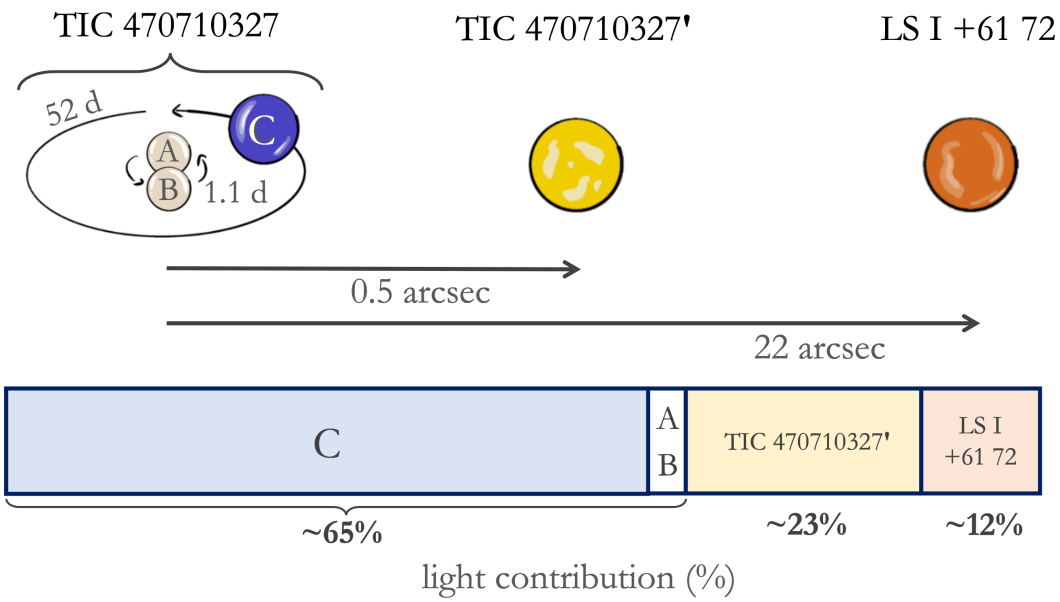

Figure 6.2: Schematic showing the light contributions of TIC 470710327, TIC 470710327′, and LS I +61 72 to the *TESS* aperture. Relative sizes of the star and distances are not to scale. Star C contributes ∼94% of the total light of the triple system made (Section 6.6.2). TIC 470710327′ and LS I +61 72 are not expected to have any dynamical effect on the triple system on the timescales of the *TESS* observations.

$\rho = 0.533$ arcsec on 1987.7568 (Hartkopf et al., 2000) to $\theta = 303.7°$ and $\rho = 0.502$ arcsec on 2003.9596 (Hartkopf et al., 2008).

As noted above, the triple system TIC 470710327 presented in this chapter consists of a $P_1 \sim 1.10$ days eclipsing binary with a $P_2 \sim 52.0$ days non-transiting tertiary. Hereafter, the two stars in the inner, short period binary will be referred to as stars A and B, while the tertiary on the wide, outer orbit will be referred to as star C. The 0.5 arcsec companion star will be referred to as TIC 470710327′ (see Figure 6.2).

6.2.2 Association with surrounding stars

In order to calculate the magnitude differences between TIC 470710327 and TIC 470710327′, I obtained speckle imaging using the Zorro instrument on the 8.1-m Gemini South telescope on Cerro Pachón, Chile (Howell et al., 2011; Matson et al., 2019). Observations were carried out on 15 August 2020 using the two-colour diffraction-limited optical imager with 60 millisecond exposures in sets of 1000

frames. The $5\,\sigma$ detection sensitivity and the speckle reconstructed image are shown in Figure 6.3. The data confirmed that the companion star is located at an angular separation of 0.529 arcsec with a position angle of 304.5°, which is in agreement with previous observations (Hartkopf et al., 2000, 2008). The data showed that the companion has a magnitude difference of $\Delta m = 1.17$ mag at 562 nm and of $\Delta m = 1.13$ mag at 832 nm to the target star.

To further characterise the target's surroundings and to quantify the light contribution of nearby stars, I queried the *Gaia* eDR3 catalogue (Gaia Collaboration et al., 2021). This search revealed a bright ($V = 11.6$ mag, $\Delta T = 2.4$ mag) nearby star at a separation of ~ 22 arcsec (LS I +61 72), as well as a further six stars with a $\Delta T < 5$ mag within 100 arcsec of the target (listed in Table 6.1). The light contribution of these stars to the photometric observations are discussed in Section 6.3.

In order to determine whether TIC 470710327 and TIC 470710327′ are dynamically associated with one another, I provide a simple estimate of the probability that the small projected separation of the two sources is due to a chance alignment (spurious association). Following the methodology outlined by Rainot et al. (2020), I define the spurious association probability as the probability that at least one object within a given radius is at least as bright as TIC 470710327′, given a local field density. I queried the *Gaia* eDR3 list of sources in a 2-arcmin radius around the target to determine the local field density (σ_{LFD}). As there are no sources as bright as TIC 470710327′ within this radius, the local field density is:

$$\sigma_{\mathrm{LFD}} = \frac{1}{\pi \times (2\ \mathrm{arcmin})^2} = 0.08\ \mathrm{arcmin}^{-2} = 2 \times 10^{-5}\ \mathrm{arcsec}^{-2}. \qquad (6.1)$$

The spurious association probability between TIC 470710327′ and TIC 470710327 is therefore $\pi \times 0.5\ \mathrm{arcsec}^2 \times \sigma_{\mathrm{LFD}} \approx 10^{-5}$, meaning that it is statistically likely that there is a physical associated between the target and the nearby companion. However, assuming that the *Gaia* distance to this target (~ 950 pc; Gaia Collaboration et al., 2021) is accurate, an angular separation between TIC 470710327′ and TIC 470710327 of 0.5 arcsec corresponds to a physical separation of approximately 500 AU with

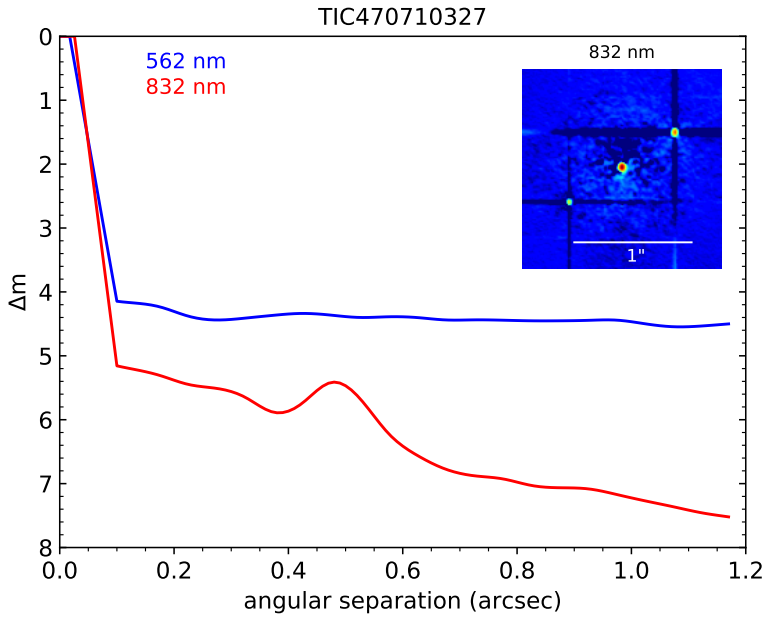

Figure 6.3: Contrast curves showing the $5\,\sigma$ detection sensitivity and the speckle reconstructed image for filters centred on 562 nm (blue) and 832 nm (red).

an orbital period on the order of 1500 – 2000 years (from Kepler's third law), depending on the mass of the companion. As such, the physical separation is too large for TIC 470710327′ to have any dynamical effect on the triple system on the timescale of the *TESS* observations.

6.2.3 RUWE parameter

The Renormalised Unit Weight Error (RUWE) parameter provided by *Gaia* eDR3 supports the idea that TIC 470710327 consists of multiple components. In brief, the RUWE parameter assesses the goodness of fit of a 5-parameter model of the motion of the target on the sky to the astrometric observations. For single stars, the motion of the centre of light and the centre of mass are the same, however, for binaries or higher order multiples, the model for the centre of light contains additional parameters due to orbital motions that the fitted model does not account for. As such, the model provides a good fit for the astrometric observations of single stars but not for multiples. For single stars, the RUWE parameter is expected to be around 1 while values greater than 1.4 suggest that the source could be a non-single star. TIC 470710327 has a RUWE parameter of 11.9, providing further

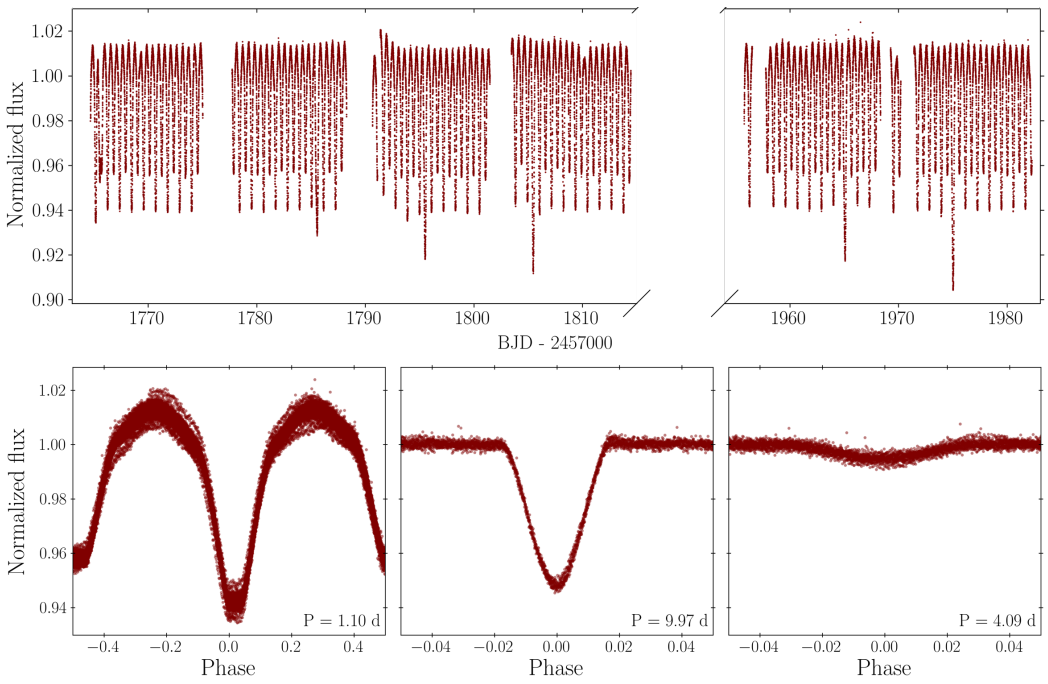

Figure 6.4: *Top*: full *TESS* light curve of TIC 470710327 obtained during Sectors 17,18 and 24. The dashed lines on the x-axis show a split in the axis. *Bottom*: light curve phase folded on $P_1 \sim 1.10$ days, $P_3 \sim 9.97$ days and $P_4 \sim 4.09$ days signals. Each panel represents one signal, where the other two signals had been masked out. Combined, these three signals make up the light curve seen in the top panel. P_3 and P_4 are assumed to not be dynamically associated with the triple system presented in this paper, as discussed in Section 6.5.

evidence of the mutliplicity of this system, while TIC 470710327′ has a RUWE parameter of 0.9, indicating that it is likely a single object.

6.3 Photometric data and ETVs

6.3.1 *TESS*

TIC 470710327 was monitored by *TESS* in the two-minute cadence data during Sectors 17, 18 and 24 of the nominal mission. The full *TESS* data set is displayed in the top panel of Figure 6.4.[1] The data spans ∼230 days, with the first subset covering nearly 60 days continuously, followed by a large gap of 140 days, before the second ∼30-day subset of observations. Visual inspection of the full *TESS* light curve reveals a clear ∼1.10-day eclipsing binary system (Figure 6.4). The

[1]The Sector 17, 18 and 24 light curves were processed and reduces with versions 4.0.28, 4.0.29 and 4.0.36 of the SPOC pipeline, respectively (Jenkins et al., 2016).

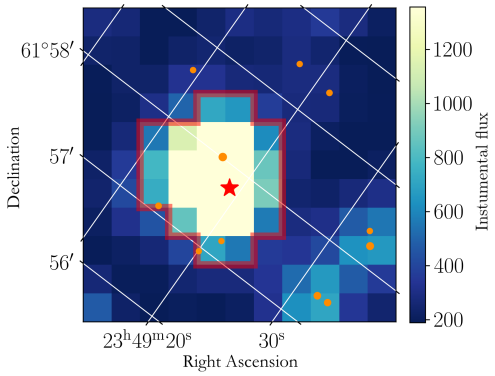

Figure 6.5: Average flux per pixel around TIC 470710327 obtained by *TESS* during Sector 17. The orange dots show all neighbouring stars with $\Delta T < 5$ mag from the target star, as queried from *Gaia* eDR3 (Gaia Collaboration et al., 2021). The red outline shows the *TESS* aperture that was used to extract the flux in Sector 17.

flat-bottom morphology of the eclipses, combined with clear points of ingress and egress, implies that the inner binary has a high inclination (i_1), close to 90°.

TIC 470710327 was identified as a potential multiple system by citizen scientists taking part in the PHT citizen science project. It was flagged as an interesting system on the discussion forum on 6 December 2019,[2] due to the light curve containing multiple periodic signals: $P_1 \sim 1.10$ days, $P_3 \sim 9.97$ days, and $P_4 \sim 4.01$ days, as shown in the bottom panel of Figure 6.4.

The pixel aperture used to extract the SPOC light curve is displayed as the red outline in Figure 6.5. The orange circles indicate the position of nearby stars that have a *TESS* magnitude difference $\Delta T < 5$ mag from the target star, as queried from *Gaia* eDR3 (Gaia Collaboration et al., 2021). Considering all sources with a $\Delta T < 5$ mag within 100 arcsec of the target (listed in Table 6.1) I expect the combined light from TIC 470710327 and TIC 470710327′ to contribute $\sim 88-89.5\%$ of the total light in the *TESS* light curve. This was calculated as the ratio of the summation of the fluxes of all the stars within this radius to the flux of the target star.

Similarly, using the magnitude contrast derived from the speckle imaging, I estimate that TIC 470710327′ contributes $\sim 26\%$ of the observed light at 832 nm. Thus, given the $\sim 10.5-12\%$ light contribution of the neighbour at 22 arcsec (the

[2] https://www.zooniverse.org/projects/nora-dot-eisner/planet-hunters-tess/talk/2112/1195850?comment=2011869&page=1

light contribution from the other, fainter neighbouring stars is assumed to be negligible), TIC 470710327 and TIC 470710327′ are estimated to contribute ∼65% and ∼23% of the total light observed by *TESS*, respectively (see Figure 6.2). As shown in Section 6.6.1, the signals $P_3 \sim 9.97$ days and $P_4 \sim 4.01$ days cannot originate from the tertiary object in TIC 470710327, but likely originate from some other nearby object that contributes to this contaminating light.

In order to determine the origin of these signals, I attempted to use both smaller apertures as well as centroid motions to investigate the direction in which the photo-centre moved during the $P_3 \sim 9.97$ days and $P_4 \sim 4.01$ days eclipses (see Section 2.2.4). I found no discernible motion in the centroid positions and negligible change in eclipse depths when different apertures were used. This either suggests that all signals originate from a single *TESS* pixel or that these signals have too small of a light contribution to result in a detectable centroid motion. Throughout the remainder of this chapter I will assume that the stars associated with P_3 and P_4 are not close enough to the triple system to induce observable ETVs.

6.3.2 Additional photometric observations

In addition to the *TESS* data, there are 531 archival photometric observations obtained with the 0.25-m Takahashi Epsilon telescope in Mayhill, New Mexico, USA between 2011 and 2013 (Laur et al., 2017); 108 photometric observations taken over 600 days by the SuperWASP-N camera located on La Palma, Canary Islands (Pollacco et al., 2006); and 1784 photometric observations obtained since 2012 by the ASAS-SN network of telescopes (Shappee et al., 2014; Kochanek et al., 2017).

These archival photometric measurements significantly increased the baseline of the observations, allowing me to refine the period of the inner binary to P_1=1.104686 ± 0.000004 days. This period was determined using a Lomb-Scargle frequency analysis, followed by visual inspection of the phase folded light curve, where I manually fine-tuned the period to minimise the dispersion of the flux in the wings of the eclipses.

Table 6.1: Nearby *Gaia* eDR3 sources within 100 arcsec and $\Delta T < 5$ mag of the target.

2MASS Identifier	Distance (arcsec)	$\Delta\ T$	$\Delta\ V$
23491896+6157459	0.000	0.00	0.00
23491667+6158004	21.845	2.40	2.35
23492132+6157124	37.395	4.76	5.75
23492010+6156568	49.643	4.82	5.07
23491426+6157050	52.642	4.38	5.64
23490933+6158367	84.892	4.65	5.36
23493217+6157231	95.863	3.35	3.69
23492162+6159230	98.898	4.45	5.11

As the observing facilities used to obtain these measurements have significantly smaller pixel scales than *TESS*, the extracted light curves do not include TIC 470710327′.[3] This allowed me to confirm the stability of the dominant periodic signal, $P_1 \sim 1.10$ days, and verify that this binary signal does not originate from TIC 470710327′. However, due to the sparse sampling of these three data sets I was unable to use these observations to investigate the stability or the origin of either the P_3 nor the P_4 signals.

6.3.3 Eclipse timing variations

I searched for ETVs, which are deviations from a strictly linear ephemeris of the eclipsing binary, that can be used to detect or study additional gravitating bodies in a system. While the midpoint of eclipses of an isolated eclipsing binary are expected to occur at regular time intervals, perturbations from additional bodies, such as in a hierarchical triple system, can result in periodic variations in the times between consecutive eclipses. I searched for deviations from the predicted linear ephemeris of P_1, given by:

$$T_{\mathrm{min,P/S}} = t_{0,P/S} + 1.104686 \times (E)$$

[3]The Takahashi Epsilon telescope, SuperWASP-N camera and ASAS-SN network of telescopes have pixel scales of 1.64 arcsec, 13.7 arcsec and 7.8 arcsec, respectively. *TESS* has a pixel scale of 21 arcsec.

where E is the cycle number since the reference orbit, $t_{0,P} = 2458766.2700$ days (BJD) is the reference epoch for the primary eclipses, and $t_{0,S} = 2458766.82234$ days (BJD) is the reference epoch for the secondary eclipses.

Using the *TESS* data alone, I determined the deviations from this ephemeris following the methodology outlined by Li et al. (2018). In brief, I determined the eclipse regions of the primary and secondary eclipses by identifying the minima of the second derivative around the time of the eclipses in the phase folded light curve. These translate to the phases of ingress and egress. Prior to this I masked the $P_3 = 9.9733$ days signal from the light curve. The lower amplitude signal from P_4 was not removed due to the risk of introducing spurious signals.

Next, I generated a model of the primary eclipse by fitting a trapezoid (the shape found to best represent the eclipses) to the smoothed, phase folded, and subsequently binned light curve. The best-fit model generated from the phase folded data was then fit to each individual primary transit where the only two free model parameters were the time of the eclipse and the slope of an underlying linear trend. The latter was to allow for systematic effects that change the slope of the eclipse. The same methodology was independently carried out for the secondary eclipses.

The fit to the phase folded data and to all of the individual eclipses, including both the primary and the secondary eclipses of the $P_1 = 1.1047$ days signal, were optimised using an MCMC approach (see Section 2.3.4). This was done using the open-source software EXOPLANET (Foreman-Mackey et al., 2021a). The observed minus calculated times of eclipse (O–C), which show variations on the range of around ± 6 minutes, are listed in Tables C.2. The O–C curve shows a periodic signal with a period of $P_2 \sim 52.0$ days. As the same periodicity is also observed in the radial velocities (Section 6.4), this signal is attributed to the period of the tertiary star of the triple system, as discussed further in Section 6.5.

6.4 Spectroscopic observations and radial velocity extraction

My collaborators and I obtained 24 spectroscopic observations between 31 January and 5 October 2020 using the HERMES spectrograph ($R \sim 85\,000$; Raskin et al., 2011) on the 1.2-m Mercator telescope at Observatorio del Roque de los Muchachos at Santa Cruz de la Palma, Canary Islands, Spain. The spectra were reduced (extracted, order-merged, wavelength-calibrated) using the local HERMES pipeline (Raskin et al., 2011), before I normalised them using a spline fit (Abdul-Masih et al., 2021). The spectra show strong He II 4686 lines and no sign of He II 4542 lines, indicating that the dominant signal originates from an early B star. Furthermore, the relative strengths of the He I, He II and Mg lines indicate that the dominant spectral contribution is consistent with that from an O9.5 – B0.5V star (Sota et al., 2011). This was determined by directly comparing the RV shifted and median combined spectrum to reference spectra presented by Sota et al. (2011). The median combined spectrum is shown in Figure 6.6, where the bottom panels show the relative depths of pairs of diagnostic spectral lines from Sota et al. (2011) used to determine the spectral classification of the star.

The derived spectral classification agrees with the previous estimates of the spectral class by Brodskaya (1953) and Laur et al. (2017). The spectral class of O9.5 – B0.5V nominally corresponds to a 14 – 17 $M_\odot$ star (Harmanec, 1988; Martins et al., 2005; Silaj et al., 2014).

I searched for additional signals in the spectra using Least Squares Deconvolution (LSD; Donati et al., 1997; Tkachenko et al., 2013). In brief, this technique assumes that all lines in a spectrum have a common underlying profile with varying depths depending on the particular line. This common profile is recovered by deconvolving the lines within a particular wavelength region with a line-list template with associated line-depths. The LSD method was generalised by Tkachenko et al. (2013) to allow for multiple stellar components in a spectrum, where each component draws from a different line list. Here, I calculated LSD profiles of the HERMES

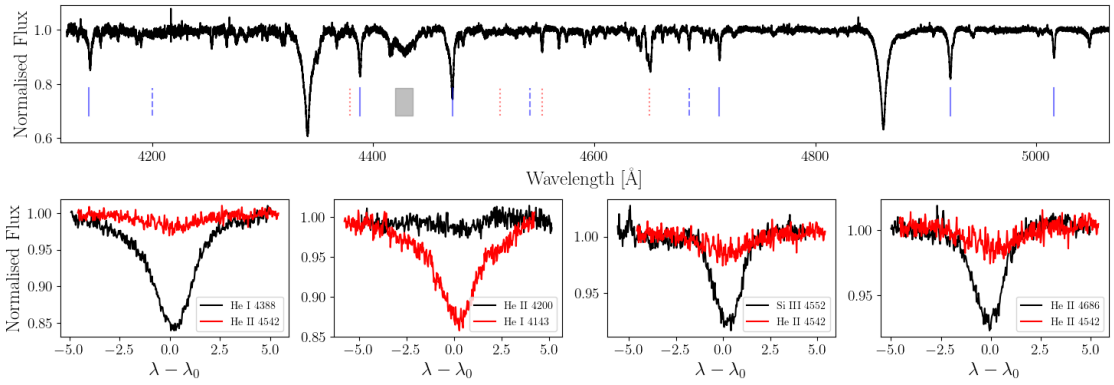

Figure 6.6: *Top*: Shifted and median combined HERMES spectra of TIC 470710327 A (S/N > 80). Solid vertical blue lines denote He I lines, dashed vertical blue lines denote He II lines, dotted red lines denote metal lines, and the shaded grey region denotes a diffuse interstellar band. *Bottom*: each panel displays a set of two lines used for diagnostic determination of spectral classification by Sota et al. (2011).

spectra over 4300 – 5200 Å, using helium, carbon, nitrogen, and silicon lines whose rest wavelengths were computed from the Vienna Atomic Line Database (VALD-II; Kupka et al. 1999). I restricted the analysis to this range to minimise the potential noise contribution from weak lines. The resulting profiles did not reveal the presence of any additional components. The radial velocities of each spectrum were computed as the centre of the LSD profiles and are listed in Table C.1. Inspection of the RVs reveal the same periodic signal, of $P_2 \sim 52.0$ days, as seen in the ETVs (Section 6.3.3). As the $P_2 \sim 52.0$ days signal is seen in both the ETVs and the RVs, this confirms that the system is comprised of a $P_1 \sim 1.10$ days eclipsing binary (stars A and B) and a $P_2 \sim 52.0$ days outer O9.5 – B0.5V tertiary (star C) orbiting around a common centre of mass.

6.5 Joint ETV and RV modelling

In this section I describe the joint modelling of the ETV and the RV signals assuming that they are physically associated as a triple system. There are two main effects responsible for deviations from strict periodicity in eclipse timings: the light travel time effect (LTTE; geometrical contribution) and the dynamical effect. The former is a result of a change in projected distance from the centre of mass of the binary to that of the triple. The dynamical effect, on the other hand, results from physical

changes in the orbit of the binary system due to the gravitational influence of the third body (Borkovits et al., 2003). Modelling of the LTTE and dynamical ETVs allowed me to derive properties including the mass ratio of the tertiary to the total mass of the system ($m_{\rm C}/m_{\rm ABC}$), the eccentricity of the tertiary (e_2), and the mutual inclination between the orbital plane of the binary and the orbital plane of the tertiary (i_m). Following Borkovits et al. (2016), perturbations produced by a close third body introduce deviations to a linear ephemeris according to:

$$\Delta = \sum_{i=0}^{3} c_i E^i + [\Delta_{\rm LTTE} + \Delta_{\rm dyn}]_0^E \tag{6.2}$$

where the first three terms multiplied by the cycle number E represent corrections to the reference epoch (c_0), the orbital period (c_1), and any secular changes to the period (c_2).

The extent of the contribution of the LTTE to the perturbation depends on the light crossing time of the relative orbit, $a_{AB}\sin i_2/\tilde{c}$ (where $\tilde{c}$ is the speed of light), as well as the configuration of the outer orbit:

$$\Delta_{\rm LTTE} = -\frac{a_{AB}\sin i_2}{\tilde{c}} \frac{(1-e_2^2)\sin(\nu_2+\omega_2)}{1+e_2 \cos \nu_2} \tag{6.3}$$

where ν_2 is the true anomaly of the outer orbit, and is determined by $t_{0,2}$, P_2, and e_2. All quantities relating to the inner orbit ($P_1 \sim 1.10$ days) have a subscript 1, while all quantities relating to the outer orbit ($P_2 \sim 52.0$ days) have a subscript 2. Quantities with the subscripts A and B refer to the individual components of orbit 1, whereas quantities with the subscript C refer to the tertiary component in orbit 2. All symbols referring to a single quantity are explained in Table 6.2.

The dynamical perturbation has a more complex dependence on the mass ratio of the system, the ratio of the periods, as well as the mutual inclination of the two orbits, denoted as i_m. This term is given by:

$$\Delta_{\rm dyn} = \frac{3}{4\pi}\frac{m_{\rm C}}{m_{\rm ABC}}\frac{P_1^2}{P_2}(1-e_2^2)^{-3/2} \times \left[\left(\frac{2}{3} - \sin^2 i_m\right)\mathcal{M} + \frac{1}{2}\sin^2 i_m \mathcal{S}\right] \tag{6.4}$$

where

$$\mathcal{M} = 3e_2 \sin\nu_2 - \frac{3}{4}{e_2}^2 \sin 2\nu_2 + \frac{1}{3}{e_2}^3 \sin 3\nu_2 \tag{6.5}$$

and

$$\mathcal{S} = \sin(2\nu_2 + 2g_2) + e_2\left[\sin(\nu_2 + 2g_2) + \frac{1}{3}\sin(3\nu_2 + 2g_2)\right]. \tag{6.6}$$

All symbols are explained in Table 6.2. The RV variations of the tertiary component are given by:

$$V = \gamma + \frac{2\pi a_C \sin i_2}{P_2\sqrt{(1-e_2^2)}}[e_2\cos(\omega_2) + \cos(\nu_2 + \omega_2)] \tag{6.7}$$

where all terms have the same subscripts as in the ETV equations, and the systemic velocity is given by γ. Given the overlap in parameters between the ETV and RV models, as well as the complementary information held in the independent data sets, I am able to constrain the system properties to a high degree.

The model optimisation was carried out using a No U-Turn Sampling (NUTS; Hoffman & Gelman, 2011) Hamiltonian Monte Carlo (HMC) approach. In short, HMC is a class of MCMC methods used to numerically approximate a posterior probability distribution. However, whereas traditional MCMC techniques use a stochastic walk to explore a given n-dimensional parameter space, the HMC algorithm makes use of a Hamiltonian description of probability distribution in order to more directly sample the posterior probability distribution of a set of model parameters, Θ, given by Bayes' theorem (see Section 2.3.1). While the standard HMC approach requires a manual fine-tuning of the number of steps and step sizes in which the system is evolved, making it very inefficient, the NUTS HMC sampler evolves the system until the proposed location starts to revert towards the initial location. Furthermore, the NUTS sampler adopts the random adopt/reject approach of new locations that is used by the standard MCMC algorithm described in Section 2.3.4.[4] The NUTS sampler was used for this analysis as it is more time efficient than the standard MCMC approach for exploring high dimensional parameter spaces (e.g., Hoffman & Gelman, 2011; Betancourt, 2017). In this application, the model was given by the ETVs and RV variations in Equations 6.2 and 6.7, such that the likelihood function scales as:

[4]For a visualisation of a NUTS HMC sample see https://chi-feng.github.io/mcmc-demo/app.html

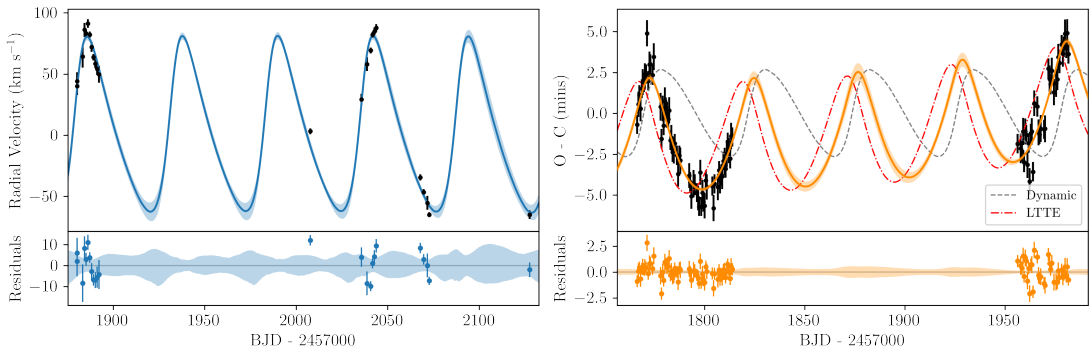

Figure 6.7: Joint MCMC model of the HERMES RV data (left panel) and the extracted eclipse timing variations (right panel). The model parameters are presented in Table 6.2. The bottom panels show the residuals of the best fit. The overall ETV fit is a linear addition of the dynamical effect (small dashed grey line) and the light travel time effect (large dashed red line).

$$P\left(d|M\right) \propto \left(\mathcal{M}_{ETV} - Y_{ETV}\right) + \left(\mathcal{M}_{RV} - Y_{RV}\right) \tag{6.8}$$

where $\mathcal{M}$ refers to the model and Y refers to the data.

In order to best exploit the complementary and overlapping information in the RV and ETV data, the two data sets were modelled jointly. This allowed me to simultaneously fit for, and better constrain, parameters that appear in both models (see Table 6.2). The joint analysis made use of the open-source software packages EXOPLANET and PYMC3 (Salvatier et al., 2016; Foreman-Mackey et al., 2021a) and the optimal parameter values and their uncertainties were calculated as the median and 68.3% highest posterior density of the marginalised posterior distributions. The priors and extracted values for all sampled parameters are given in Table 6.2 and the best-fitting models for the RVs and ETVs are shown in Figure 6.7. I note that the residuals in the ETVs and RVs show no evidence for obvious additional periodicities.

6.6 System configuration, stability, and evolution

6.6.1 Configuration

The combination of radial velocity and photometric data revealed a dynamically interacting triple system, comprised of a close 1.1047-day eclipsing binary (stars A

Table 6.2: System parameters either sampled or derived from the HMC optimisation, as well as parameters estimated from the *TESS* light curve.

Parameter	Symbol	Prior[(a)]	Value[(b)]	Units	Model
Sampled parameters					
Orbital period tertiary	P_2	$\mathcal{N}[52.1, 2]$	$52.04^{+0.02}_{-0.02}$	days	ETV + RV
Semi-major axis, binary to COM	$a_{\rm AB} \sin i_2$	$\mathcal{U}[10, 500]$	$91.22^{+2.67}_{-10.7}$	$R_\odot$	ETV
Tertiary to total mass ratio ($m_{\rm C}$ / $m_{\rm ABC}$)	$q_{\rm tot}$	$\mathcal{N}[0.58, 0.14]$	$0.56^{+0.03}_{-0.01}$	–	ETV
Tertiary eccentricity	e_2	$\mathcal{U}[0, 0.4]$	$0.300^{+0.013}_{-0.009}$	–	ETV + RV
Binary eccentricity	e_1	–	0	–	fixed
Observed argument of periastron	ω_2	$\mathcal{U}[0, 360]$	$295.7^{+6.6}_{-3.6}$	deg	ETV + RV
Mutual inclination	i_m	$\mathcal{U}[0, 360]$	$16.8^{+4.2}_{-1.4}$	deg	ETV
Dynamical argument of periastron	g_2	$\mathcal{U}[0, 360]$	$359.05^{+12.0}_{-41.2}$	deg	ETV
Semi-amplitude	K	$\mathcal{N}[81, 20]$	$72.2^{+1.0}_{-1.0}$	–	RV
t_0 of tertiary	$t_{0,2}$	$\mathcal{N}[1878, 25]$	$1880.4^{+0.7}_{-0.2}$	BJD - 2457000	ETV + RV
Correction to T_0	c_0	$\mathcal{N}[0, 5]$	$-0.0^{+0.0000}_{-0.0001}$	–	ETV
Correction to P_1	c_1	$\mathcal{N}[0, 0.5]$	$-0.003^{+0.017}_{-0.009}$	–	ETV
Secular change to P_1	c_2	$\mathcal{N}[0, 0.0001]$	$0.2^{+0.5}_{-0.3}$	–	ETV
Derived parameters					
Semi-major axis, tertiary to COM	$a_{\rm C} \sin i_2$	–	$70.923^{+0.98}_{-2.03}$	$R_\odot$	RV
Project mass of binary	$m_{\rm AB} \sin^3 i_2$	–	$9.21^{+0.6}_{-1.4}$	$M_\odot$	ETV + RV
Project mass of tertiary	$m_{\rm C} \sin^3 i_2$	–	$11.87^{+0.9}_{-2.9}$	$M_\odot$	ETV + RV
Tertiary to binary mass ratio ($m_{\rm AB}$ / $m_{\rm C}$)	q_2	–	$0.777^{+0.14}_{-0.04}$	–	ETV + RV
Light curve extracted parameters					
Orbital period binary	P_1	–	1.104686±0.000004	days	–
t_0 of primary eclipse	$t_{0,1}$	–	1785.533	BJD - 2457000	–
SED parameters					
Mass of binary star A	$m_{\rm A}$	–	6 – 7	$M_\odot$	SED
Mass of binary star B	$m_{\rm B}$	–	5.5 – 6.3	$M_\odot$	SED
Mass of tertiary star C	$m_{\rm C}$	–	14.5 – 16.0	$M_\odot$	SED
Binary mass ratio ($m_{\rm B}$ / $m_{\rm A}$)	q_1	–	0.9 – 1.0	–	SED
Luminosity of binary star A	L_A	–	3.00 – 3.27	log $L/L_\odot$	SED
Luminosity of binary star B	L_B	–	3.02 – 3.14	log $L/L_\odot$	SED
Luminosity of tertiary star C	L_C	–	4.71 – 4.86	log $L/L_\odot$	SED
Distance	d	–	4.0 – 4.5	kpc	SED
Reddening	E(B-V)	–	0.40 – 0.44	–	SED

Note – [(a)] $\mathcal{U}[a, b]$ refers to uniform priors between a and b, $\mathcal{N}[a, b]$ to Gaussian priors with mean a and standard deviation b. [(b)] Inferred parameters and errors are defined as the median and 68.3% credible interval of the posterior distribution.

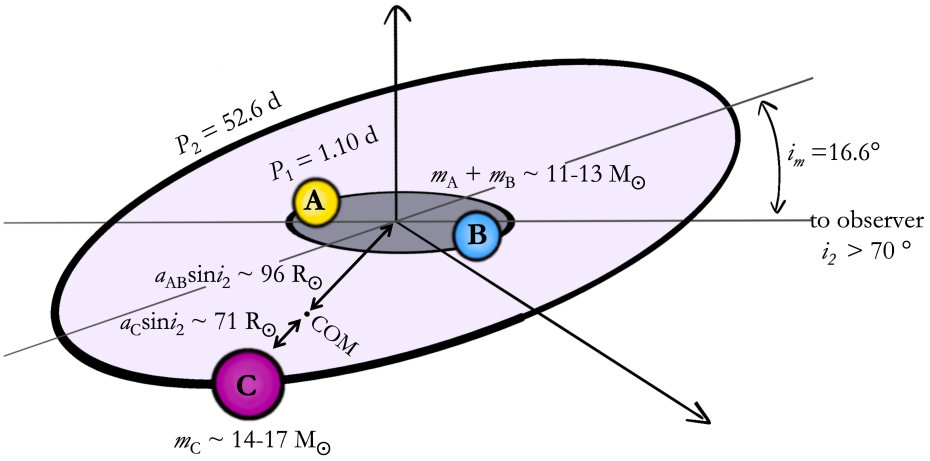

Figure 6.8: Schematic showing the configuration of the triple system. Relative sizes of the orbits are not to scale and for simplicity the orbits are depicted as circular.

and B) with a massive companion on a wide, non-eclipsing, $\sim$52.0-day orbit (star C). From the dynamical modelling (presented in Section 6.5), I derived a mass ratio of the inner binary to the tertiary of $q_2 = m_{\rm AB}/m_{\rm C} = a_C \sin i_2 / a_{AB} \sin i_2 = 0.777$. Given the spectral classification of star C, of O9.5V – B0.5V, the tertiary star has a mass in the range of $m_{\rm C} = 14-17$ M$_\odot$, meaning that the combined mass of stars A and B is in the range of $m_{\rm AB} = 10.9-13.2$ M$_\odot$. By considering both the spectroscopic mass range of the tertiary ($m_{\rm C}$) and the derived projected mass of the tertiary ($m_{\rm C} \sin^3 i_2$) I determined limits of the inclination i_2. In turn, I used the derived mutual inclination i_m, which is the difference between the inclinations of the orbit of the inner binary and the outer tertiary, to determine i_1. From these considerations, I find that i_1 lies in the range: $77.9^\circ \leq i_1 \leq 90^\circ$ and i_2 lies in the range: $62.5^\circ \leq i_2 \leq 71.1^\circ$. The high value of i_1 agrees with the observed flat bottomed eclipses observed for the short 1.1-day orbit. A schematic of the triple system is presented in Figure 6.8.

In addition to the eclipsing $P_1 = 1.104686$ days binary signal, the *TESS* light curve contains two further periodic eclipsing signals with $P_3 = 9.9733$ days and $P_4 = 4.092$ days. The bottom panel of Figure 6.4 displays the *TESS* light curve phase folded on P_3 (middle panel) and P_4 (right panel). I note that due to the lack of visible 'secondary' eclipses, I cannot distinguish between the apparent periods

of P_3 and P_4 or twice those values. I can, however, determine that all three sets of eclipses are of different objects (i.e. from three binaries and not from a higher order multiple), as the morphology of all of the eclipses are constant in time, and points of overlap between the 1.10-day and 9.97-day signal are reproduced as linear additions of the different eclipse signals. Conversely, if the eclipses were to originate from the same system, one would expect the shapes, duration, and timings of the eclipses to change, such as observed by, for example, Hełminiak et al. (2017) and Rappaport et al. (2022).

Furthermore, there is no dynamical evidence in the RVs or the ETVs of either the 9.97-day or the 4.09-day signals being associated with the tertiary star. Using Equation 6.4, it can be shown that a dynamically interacting body with $P_3 \sim 9.97$ days would produce an amplitude in the ETVs that is at least $(9.97/1.10)^2 = 82$ times larger than the dynamical amplitude induced by the $P_2 \sim 52.0$ days tertiary. There is no evidence of this signal in the data. Using the same argument I also find no evidence of the $P_4 = 4.09$ days signal being associated with the tertiary star of the triple. I note that $P_1 \sim 1.10$ days and $P_3 \sim 9.97$ days is close to a 1:9 resonance ($P_3/P_1 = 9.028$). While this could indicate that the signals are related, there is currently not sufficient data to further investigate their association.

As noted above, the presence of two additional bright sources in the *TESS* aperture, at separations of 0.5 arcsec and 22 arcsec, overall contributing over $\sim 10\%$ of the light in the aperture, make it difficult to determine the exact source of all of the periodic signals seen in the *TESS* light curve. However, because the RVs and the ETVs show the same 52.0-day periodic signal, I can associate the triple system with the brightest O9.5V – B0.5V star in the data. This is further corroborated by the data obtained with the Takahashi Epsilon telescope, which has a pixel scale of 1.64 arcsec, which allowed me to confirm that the $P_1 \sim 1.10$ days signal lies on the O9.5V – B0.5V target star and not the 22 arcsec companion star that lies within the same *TESS* aperture.

6.6.2 SED modelling

The properties of the components of this system were further investigated via a grid-fitting analysis of the composite SED (see Section 4.4.2) of TIC 470710327 and TIC 470710327′, using photometric data from Vizier (Table C.3). Synthetic SEDs were used, extracted from a grid of MIST isochrone models (Choi et al., 2016). The analysis enforced that all components of TIC 470710327 and TIC 470710327′ are the same age and located at the same distance, and ensured realistic physical parameters for all components. The models were reddened according to $F_{\rm black} = 10^{-A_\lambda/2.5} F_{\rm mod}$ and $A_\lambda = {\rm E(B-V)} R_\lambda$, where E(B–V) was treated as a free parameter, and R_λ was calculated according to Cardelli et al. (1989). Each model was evaluated against the data using a χ^2 metric.

A composite four-component SED was fit to the observations, assuming one component for each member of the triple system, and one for TIC 470710327′. The light contributions derived from the speckle imaging (Section 6.3) were used to inform the model, such that the total light contribution of the triple system is fixed to 74% in the V-band, with the remaining 26% originating from TIC 470710327′. Additionally, the analysis enforced that the luminosity of the primary component of the close binary must be higher than that of the secondary, i.e. $L_A > L_B$. Finally, the spectroscopically derived mass range for the tertiary ($m_{\rm C} \in 14-17$ M$_\odot$) was adopted, and only solutions that satisfy the triple system mass-ratio of $q_2 = 0.777$ were accepted. This resulted in 1019 possible solutions, which were ranked according to their χ^2 values.

The solutions favour masses of the OB star of $m_{\rm C} \in 14.5-16$ M$_\odot$, while the solutions for the components of the close binary orbit favour a mass ratio $q_1 = m_{\rm B}/m_{\rm A} \in 0.9-1.0$ with masses for the primary components ranging from $m_{\rm A} \in 6-7$ M$_\odot$ and $m_{\rm B} \in 5.5-6.3$ M$_\odot$. Solutions for the mass of TIC 470710327′ were found in the range of 11.5 – 14 M$_\odot$. It should be noted that the values of the mass of TIC 470710327′ are correlated with the mass of star C due to the imposed light contributions of these objects. Solutions for the luminosities of stars A, B, and C

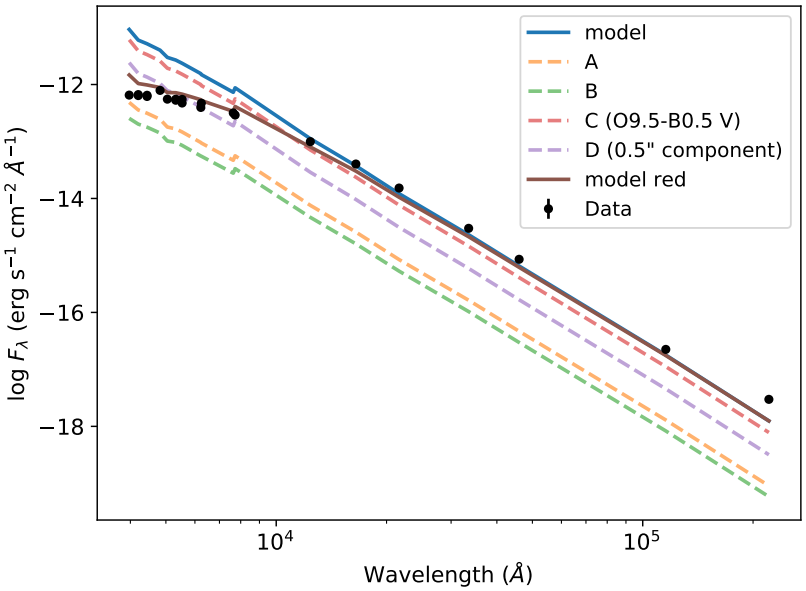

Figure 6.9: Observed SED (black points) with composite and individual model contributions for the best fitting composite model. The individual contributions are un-reddened. The reddened and un-reddened composite models are shown in maroon and blue, respectively.

show that the tertiary star contributes $\sim 94\%$ of the total light of the triple system (see Figure 6.2). All derived values are listed in Table 6.2.

The model and observed SED are shown in Figure 6.9. The modelling returns a distance in the range of $4-4.5$ kpc, whereas the *Gaia* parallax suggests that the distance is closer to 1 kpc. In order to arrive at a distance close to the *Gaia* estimate, the reddening would need to be around 2, as opposed to the value in the range of $0.40-0.44$ that is preferred by the model. Although these two distances do not agree, the complexity of this system could contribute to an erroneous astrometric solution, which will likely be improved upon in future *Gaia* data releases.

6.6.3 Dynamical stability

Even though both the inner and outer orbits of the triple system are relatively compact, the separation between the two orbits ($P_2/P_1 \approx 50$) implies that the current configuration is dynamically stable. According to the stability criterion of Mardling & Aarseth (1999), a triple system is stable if:

$$a_2 \gtrsim 2.8 \left(\frac{m_{\rm ABC}}{m_{\rm AB}}\right)^{2/5} \frac{(1+e)^{2/5}}{(1-e)^{6/5}}\, a_1. \tag{6.9}$$

Using Kepler's third law this can be expressed in terms of the orbital period:

$$P_2 \gtrsim 4.7 \left(\frac{m_{\rm ABC}}{m_{\rm AB}}\right)^{1/10} \frac{(1+e)^{3/5}}{(1-e)^{9/5}} P_1 \tag{6.10}$$

which implies long-term dynamical stability for TIC 470710327 for $P_2 \gtrsim 8$ days.

This long-term stability gives rise to the possibility of observing three-body dynamics. ZKL cycles (Lidov, 1962; Kozai, 1962), for example, would manifest themselves as cyclic changes in the eccentricity of the inner orbit and in the mutual inclination between the inner and outer orbit. However, classical ZKL resonance can only occur in triples with mutual inclinations between 39.2° and 140.8° (e.g., Toonen et al., 2016), meaning that with the mutual inclination of TIC 470710327, of $i_m \approx 16.8\,^{+4.2}_{-1.4}\,^{\circ}$, this effect is unlikely to be significant for the future evolution of the system.

Conversely, the higher-order effects of three-body dynamics could affect the dynamical evolution of this system (Toonen et al., 2016). The eccentric ZKL mechanism (eZKL; see Naoz, 2016, for a review) can give rise to more extreme eccentricity cycles for an extended range of inclinations. The magnitude of this effect can be quantified with the use of the octupole parameter (Lithwick & Naoz, 2011; Katz et al., 2011; Teyssandier et al., 2013; Li et al., 2014):

$$\epsilon_{\rm oct} \equiv \frac{m_{\rm A}-m_{\rm B}}{m_{\rm A}+m_{\rm B}} \frac{a_1}{a_2} \frac{e_2}{1-e_2^2} \equiv \frac{1-q_1}{1+q_1} \left(\frac{P_1^2}{P_2^2}\,(1-q_{\rm tot})\right)^{1/3} \frac{e_2}{1-e_2^2} \tag{6.11}$$

where $q_1 = m_{\rm B}/m_{\rm B}$. The eZKL mechanism is expected to be important for the evolution of the system for values of $\epsilon_{\rm oct} \sim 0.001-0.01$, and under the condition that the mass ratio of the inner binary is less than unity. Given the values in Table 6.2, I find $\epsilon_{\rm oct} = 0.001$ for $q_1 \approx 0.9$ and $\epsilon_{\rm oct} = 0.01$ for $q_1 \approx 0.3$. Given the morphology of the eclipses in the *TESS* data (and considering the substantial diluting third light), initial modelling suggests that the mass ratio of the inner binary is close to unity (see Section 6.6.2). Thus, the octupole term may affect the dynamical evolution of this system. The expected timescale of the eZKL cycles are thought to be of order

$\tau \approx 180$ years $/ \sqrt{\epsilon_{\rm oct}} \approx 1800-5700$ years (Antognini, 2015). Simulations using the TrES triple star evolution code (Toonen et al., 2016) showed that these time-scales are consistent with the analytical estimates. However, when including tides and gravitational wave radiation, the maximum amplitude of the eccentricity of the inner binary remains on the order of $e_1 \sim 0.001$. Further constraints on the time-scale and amplitudes of these cycles would require detailed modelling of the inner binary, which remains impossible without full RV characterisation of the inner components.

6.6.4 Possible formation scenarios

Several theories exist pertaining to the origin of higher order multiple systems. The formation of multiple systems is dependent on fragmentation of the natal material during the formation process. Hierarchical collapse within molecular clouds eventually leads to the formation of dense stellar cores and clumps, whose collapse results in the formation of stars and clusters. Throughout this, equatorial discs are formed through the conservation of angular momentum, which in turn can become involved in the accretion process and the formation of secondary cores. However, there is still debate as to what scale of fragmentation is the main cause of observed massive multiple systems – the fragmentation of the prestellar core or fragmentation of circumstellar discs. In both cases, opacity has a large effect on the initial separations of the systems, which cannot be less than around 10 AU due to the opacity limit of fragmentation (Boss, 1998; Bate, 1998). Therefore it is assumed that systems at closer separations than 10 AU must have migrated to their observed positions (Bate et al., 2002). Some studies present evidence of disc fragmentation creating higher order massive multiple systems during their embedded phases (e.g., Megeath et al. 2005). Other effects such as dynamical interactions between other companions and discs (Eggleton & Kisseleva-Eggleton, 2006) could also cause inner binaries in multiple systems to harden into close orbits. Recent modelling by Oliva & Kuiper (2020), for example, has shown how disc fragments in the discs of massive protostars form through hierarchical fragmentation along spiral arms and migrate to orbits that produce detectable spectoscopic signatures.

TIC 470710327 presents an interesting puzzle in terms of its formation given the derived geometry of a close binary with a total mass lower than the tertiary star. More massive stars have shorter Kelvin-Helmholtz timescales than those of lower masses, where the Kelvin-Helmholtz timescale is the time it takes the star to radiate away all of its kinetic energy given its luminosity. This is important for pre-main sequence lifetimes, as it determines how quickly a star contracts prior to the onset of nuclear fusion. As such, one would expect that the more massive tertiary was the first star to form. However, if this was the case, it is likely that when this star reached the main sequence (before the inner binary) it would have disrupted the remaining natal material and therefore discontinued the central binary's formation. If the central binary did form first, a more consistent interpretation could be that the inner binary formed through disc fragmentation and the dynamical effects of this binary on the disc could have created a large over-density at large radii. Such over-densities have been shown to occur in circumbinary discs in works such as Price et al. (2018). This over-density could have accreted significant mass, perhaps accelerated by the continuing dynamical effects of the inner binary. In order for this scenario to hold the mass of the disc must have been very large, as the disc fragmentation process would not convert all the disc material into the eventual stars, and the combined mass of the stars in the tertiary system is at least 26 $M_{\odot}$. However, the largest protostellar discs detected around massive young stellar objects are of order $\sim$10 $M_{\odot}$ (e.g., Johnston et al., 2020; Frost et al., 2021).

An alternative explanation for the formation of TIC 470710327 is that this system is a result of the fragmentation of the prestellar core as opposed to fragmentation of a disc (e.g., Krumholz & Bonnell, 2007). This is supported by the fact that additional sources neighbour this triple system, in the form of 0.5 arcsec and $\sim$22 arcsec distant companions. With a spurious association probability of $\sim 10^{-5}$ between TIC 470710327 and the 0.5 arcsec companion, this closer source is expected to have formed from the same core collapse.[5] While the more distant source at $\sim$22

[5]I note that as there are currently no resolved *Gaia* parallaxes for TIC 470710327 and TIC 470710327′, the possibility that these two targets are just un-associated nearby objects on the sky cannot be concretely ruled out.

arcsec may not be currently bound to the tertiary system, it could still have come from the same prestellar core. The core collapse scenario circumvents the mass problem described above for disc fragmentation. Should this have occurred, we can assume that the close-binary formed from one collapse event, the tertiary star from another and the distant source from yet another. The close-binary system could have been formed by disc fragmentation as described above, and its hardening into a close-orbit could have been facilitated by dynamical interactions between the 14 – 17 $M_\odot$ star and the distant sources from others. Dynamical effects within the collapsing core could also have led to the capture of the 14 – 17 $M_\odot$ star by the close-binary system, forming the tertiary we see today. Dedicated radiative hydrodynamical modelling could help disentangle whether core fragmentation, disc fragmentation or a combination of both resulted in the formation of TIC 470710327, whilst repeat observations of all the sources in the region could help distinguish orbits and determine which stars are bound, thus helping us constrain the formation channel of this system.

6.6.5 Future evolution

The observed mass ratio q_2, whereby the tertiary is more massive than the combined mass of the inner binary, implies that the remainder of this system's evolution will be driven by the evolution of the tertiary, as this star will be the first to evolve off the main sequence. Given the expected mass of the tertiary of 14 – 17 $M_\odot$, star C is expected to expand to the point where it fills its Roche lobe at an age of $\sim$13 Myr, at which point it will start transferring mass towards the inner binary (see Figure 6.1). This type of mass transfer, from an outer star to an inner binary, is expected to occur in $\sim$1% of all triple systems in the Massive Star Catalogue (de Vries et al., 2014; Hamers et al., 2021a).

The outcome of such a mass transfer phase is, inherently, a hydrodynamical problem (de Vries et al., 2014; Portegies Zwart & Leigh, 2019). If the inner binary is compact enough such that the mass transfer stream intersects itself (i.e. the mass is not preferentially transferred to one star in the binary), a circumbinary disk may

form. As outlined by Leigh et al. (2020) such a disk forms when the circularization radius ($a_{\rm circ}$) is larger than the semi-major axis of the inner binary ($a_{\rm AB}$), where:

$$a_{\rm circ} = a_{\rm C}\,(1 - R_L)^2 \tag{6.12}$$

where R_L is the radius of the Roche lobe of the tertiary defined as (Eggleton, 1983):

$$R_L = \frac{0.49\,q_2^{-2/3}}{0.6\,q_2^{-2/3} + \log(1 + q_2^{-1/3})}. \tag{6.13}$$

As such, the circularisation radius of the mass transfer material (Frank et al., 2002; Toonen et al., 2016) is around 20 $R_\odot$, meaning that, as the inner binary has an orbital separation of $\sim$10 $R_\odot$, a disk may form. Leigh et al. (2020) argue that such a scenario leads to preferential accretion to the lowest mass star of the inner binary and therefore favours evolution towards equal mass inner binary stars. If the mass transfer stream intersects the trajectory of the inner two stars, friction may reduce the orbit of the inner binary (e.g., de Vries et al., 2014), leading to a contact system or a merger. Such a merger remnant would be a blue straggler, which is a star that has accreted mass to make it appear more blue than would be expected given its age.

Assuming that the two stars in the binary do merge, the triple would reduce to a binary system (stage 3 in Figure 6.10). The mass of such a merger remnant would depend on three major factors: the amount of mass lost during the merger of stars A and B, the amount of mass transferred from the evolving tertiary star, and the mass accretion efficiency. Even though these values are difficult to constrain (e.g., de Mink et al., 2007; Schneider et al., 2019), given the initial masses of all of the stellar components, the merger remnant would likely have a mass significantly greater than 9 $M_\odot$, which is the minimum mass required for a star to go supernova and become a neutron star.

Given the initial mass of the former tertiary star ($m_{\rm C} = 14-17$ $M_\odot$), this star will also likely end its life as a neutron star. The typical natal kick from such a supernova explosion has a magnitude of several hundreds of kms^{-1}, which could unbind the newly formed binary into two single stars (Hobbs et al., 2005;

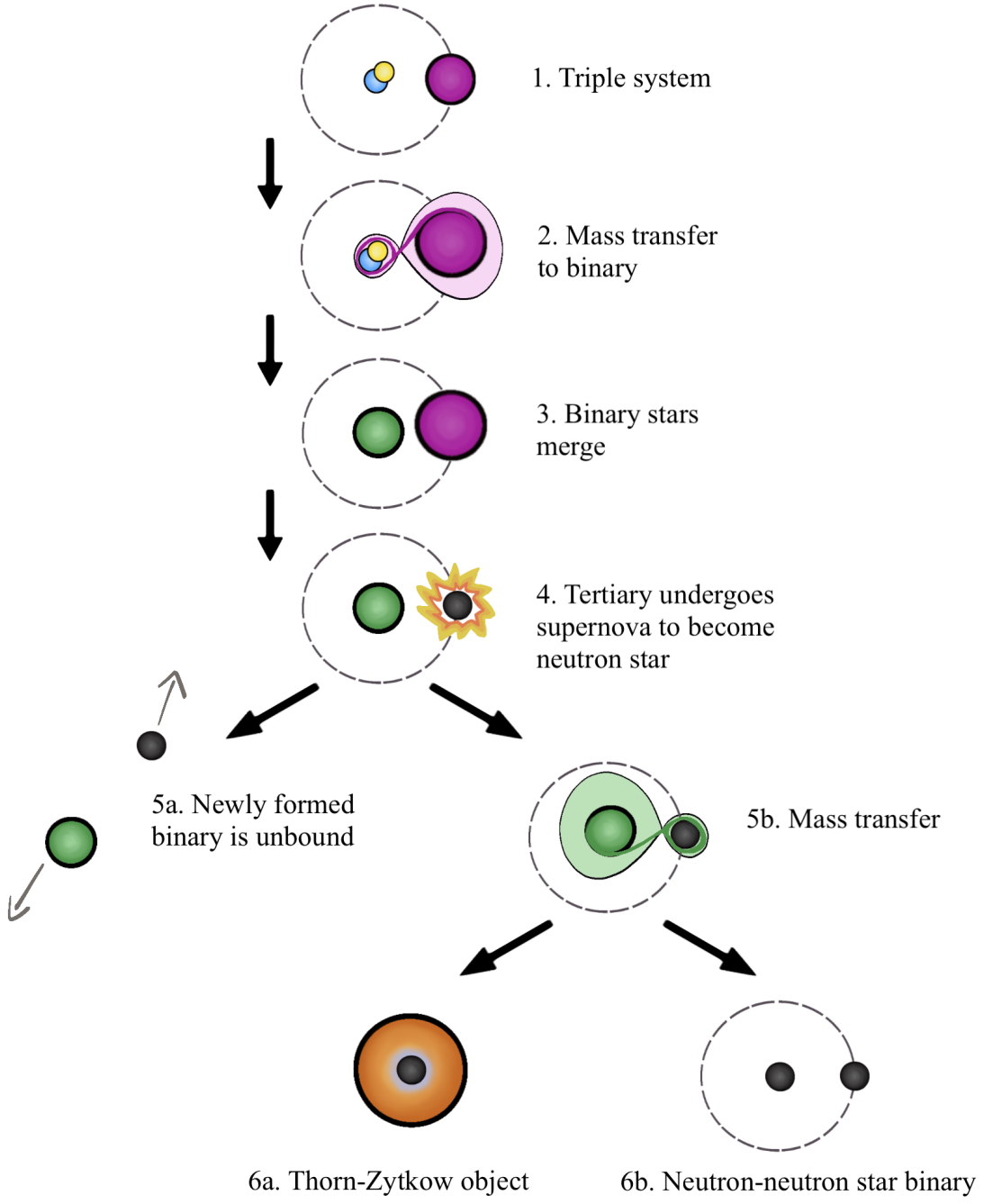

Figure 6.10: Possible future evolution of TIC 470710327. Relative sizes of the stars and orbits are not to scale. For simplicity, all orbits are depicted as circular.

Verbunt et al., 2017; Igoshev, 2020) and reduce the multiplicity of the system once more (5a in Figure 6.10).

Alternatively, if the orientation of the supernova kick is such that the binary remains intact, the binary would likely undergo an additional mass transfer phase when the merger remnant evolves off the main sequence and its radius expands. Given the large mass ratio of such a binary, the mass transfer would likely lead to a common-envelope phase (Ivanova et al., 2013). Consequently, this could either lead to 1) a merger and the formation of a Thorne-Żytkow object, whereby the neutron star is enclosed by the red giant star (6a in Figure 6.10; Thorne & Zytkow, 1975; Podsiadlowski et al., 1995; Levesque et al., 2014; Tabernero et al., 2021) or 2) if the binary survives, an ultra-stripped supernova resulting in a double neutron star system (6b in Figure 6.10; Tauris, 2015).

6.7 Conclusions

TIC 470710327 is a compact, hierarchical triple system consisting of a 1.10-day binary containing two B-type stars and an OB-type tertiary on a wide 52.0-day orbit. Dynamical modelling of the system (RVs and ETVs) revealed a rare configuration wherein the tertiary object in the wide orbit is more massive than the combined mass of the inner binary ($m_{\rm AB} = 10.9 - 13.2$ M$_\odot$, $m_{\rm C} = 14 - 17$ M$_\odot$). This configuration poses several challenges to explain its formation. Given the current mass of the tertiary component, star C is expected to fill its Roche lobe at an age of ~ 13 Myr, meaning that the system must necessarily be younger than this age (as we see no evidence of mass transfer in the spectra or the light curve). Considering that the main sequence lifetime of a $5.5 - 7$ M$_\odot$ star is between $40 - 80$ Myr, all of the stars in this system are currently in the main sequence phase of their evolution. Given the values from the dynamical modelling and constrained SED modelling, the primary binary component (star A) will fill between $70 - 83\%$ of its Roche lobe before the tertiary evolves off the main sequence, while the secondary binary component (star B) will fill between $68 - 75\%$ of its Roche lobe during this time.

This means that the binary will not undergo mass transfer before experiencing Roche lobe overflow from the tertiary.

Given the compact orbits and the unusually high mass of the tertiary object, I speculate that the future evolution of this system will minimally involve one episode of mass transfer as the massive tertiary evolves across the Hertzsprung gap. Based on its initial mass, the tertiary will likely end its life as a neutron star. Alternatively, depending on the rate and efficiency of the mass transfer to the inner binary, the tertiary could evolve into an intermediate mass stripped star (Götberg et al., 2020). Should the binary system remain bound after the expected supernova kick (or supernovae kicks), this system could result in a close double neutron star gravitational wave progenitor, or an exotic Thorne-Żytkow object. Detection of more systems similar to TIC 470710327 would provide constraints on potential progenitor systems to gravitational wave events.

With further observational characterisation, particularly aimed at characterising the nature of the inner binary, this system stands to become an excellent target to scrutinise simulations of massive star formation and evolution (e.g., Vigna-Gómez et al., 2022). Future spectroscopic observations that are specifically aimed at detecting the RV variations of both components of the inner binary would allow me to place tighter constraints on the dynamics of the system. In particular, these observations would allow me to probe the eZKL mechanism. Finally, RV characterisation and detailed eclipse modelling of the inner binary would precisely constrain the light contributions of all components of the triple system. With such constraints, derivations of the atmospheric properties, such as $T_{\rm eff}$, $\log g$, $v_{\rm rot}$ and $L_{\rm bol}$ would allow me to further test whether the three stars are coeval.

Overall, this system highlights the scientific value of large scale visual vetting campaigns for identifying unique stellar systems. Even though this is the first fully analysed multi-stellar system identified by PHT, many more have already been flagged on the PHT discussion forums, as previously discussed in Section 3.3.2.

You could claim that anything's real if the only basis for believing in it is that nobody's proved it doesn't exist!

— H. J. Granger, 1997

CHAPTER 7

CONCLUSIONS

7.1 Thesis summary

Throughout this thesis, I presented the citizen science project Planet Hunters TESS that I designed, built, and continue to manage. The project engages over 30 000 members from around the globe, making it one of the largest citizen science projects in the world. The visual vetting of data on PHT has led to the discovery of 139 new planet candidates in the two-minute cadence *TESS* data from the first three years of the mission alone. This makes PHT one of the largest contributors of *TESS* planet candidates outside of the core *TESS* science team, providing excellent evidence for the success of citizen science.

In Chapters 2 and 3, I evaluated the detection efficiency of PHT and demonstrated that with visual vetting we are equally sensitive to detecting short- and longer-period planets. The sensitivity to the detection of longer-period planets, in particular, makes PHT an effective tool to help populate under-explored regions of parameter space that have limited prospects for being studied by automated searches.

In Chapter 2, I discussed `LATTE` – an open-source vetting suite that presents numerous diagnostic tests to aid in the detection and characterisation of planetary and stellar signals – that I developed to complement the PHT pipeline. I specifically designed `LATTE` to be used by both citizen scientists as well as professional astronomers. In addition to the tool being used by the exoplanet community (e.g.,

Van Grootel et al., 2021; Barragán et al., 2022a), `LATTE` forms the basis of the *TESS* eclipsing binary working group's signal verification pipeline (Prša et al., 2022).

The combination of the transit-detection pipeline and the vetting tools that I developed have allowed me to identify a large population of longer-period planet candidates as well as to carry out in-depth studies of scientifically valuable systems. Specifically, I presented the detailed analysis of two PHT identified planetary systems.

First, I discussed a Saturn-sized planet orbiting around a bright sub-giant on an 84-day orbit (TOI 813 b) in Chapter 4. This planet's long orbital period combined with the evolved nature of the host star places TOI 813b in a relatively under explored region of parameter space and is, therefore, an exciting target for future characterisation e.g., measurement of spin-orbit misalignment. Furthermore, the mass of TOI 813 b of 48^{+15}_{-14} $M_{\oplus}$ places it in a scarcely populated transition region between volatile-rich gas giants and volatile-poor ice giants. Characterisation of planets in this region can help differentiate between different formation mechanisms for how giant planets form.

Second, I analysed a two-planet system orbiting around a bright G-dwarf (HD 152843) in Chapter 5. Through a combination of ground-based follow-up observations and detailed transit-modelling, I showed that both planets in this system are likely to have low bulk densities and therefore extended H/He atmosphere. This makes both planets prime candidates for future atmospheric characterisation with JWST, as show by the TSM value that, for example, places planet c at least amongst the top 50% of planet candidates suitable for atmospheric characterisation, as defined by Kempton et al. (2018).

Finally, I showed that large scale visual vetting leads to the detection of exciting and unique stellar systems. In Chapter 6 I presented the detailed analysis of one of these systems: a compact hierarchical triple star system that has a unique stellar configuration. As such, the overall conclusion of this thesis is that citizen science is well-suited to identifying longer-period planet candidates and aperiodic planet and stellar signals that are often missed by automated searches.

Figure 7.1: A word map of the occurrence of words on the PHT discussion forum since the launch of the project in December 2018. The font sizes correlate to the occurrence rates of the words.

7.2 Planet population study prospects

Despite PHT being able to identify planets that automated searches miss, the sample is currently not well-suited to obtain planet population statistics. This is primarily because meaningful planet occurrence rate studies require a well defined sample of planets, a careful assessment of the completeness, and a good understanding of the false positive rate (e.g., Christiansen et al., 2015; Zink et al., 2019b; Bryson et al., 2020; Christiansen et al., 2020). In Section 3.1.2, the PHT detection efficiency is assessed using the rate of recovery of the TOIs as a function of period and radius. However, due to a lack of understanding of the rate of false positives, as well as completeness, of the sample of TOIs, this analysis cannot provide a robust study of the completeness of the PHT sample. In future work, the recovery of simulated transits (see Section 2.1.2) could be used to help derive the completeness of the sample as a function of planet properties. However, while presenting the volunteers with a higher rate of simulations benefits our calculations of completeness, showing a large number of simulations can demotivate the users. Future attempts to assess the completeness using simulated data will therefore have to consider both the benefits and drawbacks of presenting the users with a higher rate of simulations.

Further uncertainty in quantifying the completeness of the PHT arises during the preliminary candidate vetting stage, which follows the initial classifications by the volunteers. As discussed in Section 2.1.9, this stage of vetting consists of a small and varying number of PHT science team members who visually inspect the 500 highest ranked candidates in order to differentiate between potential planet candidates and false positive scenarios. Currently I have no quantification of the completeness of, or biases present in this stage of the vetting. In the future, I could complement the visual inspection of these candidates with an automated classification of the signals. This would allow me to quantify the detection biases and completeness of the vetting in a systematic way, for example using simulated data.

Finally, due to the limited ground-based follow-up resources, I currently have limited understanding of the false positive rate of the sample of planet candidates. A large-scale dedicated follow-up effort would be required to verify the planetary nature of the sample and to truly understand the false positive rate. Alternatively, statistical validation methods (such as `VESPA`; see Section 4.4.3) could also help determine the false positive rate of the sample in future work without requiring costly follow-up observations.

7.3 PHT community growth

The work presented in this thesis relies on the continued effort of tens of thousands of citizen scientists from around the world. The levels of skill and enthusiasm for planet hunting displayed on the platform over the past three years has been astounding. What I found early on in the project is that our main bottle-neck for detecting and confirming new planet candidates lies not in the rate of the PHT classifications, but in my ability to keep up with investigating the citizen-science flagged signals. In order to address this, I – in collaboration with a team at NASA Ames – started a focused initiative to provide the knowledge and tools necessary for the citizen scientists to become involved in planet vetting beyond the classification that PHT offers. This initiative consists of a strategic combination of new citizen scientist-focused data examination tools, tutorials, and outreach videos

that facilitate citizen scientists to assess the signals that they find, estimate the properties of the planet candidates, rule out false positives, and directly report their findings to the PHT science team. This project, which was launched in June 2021, is known as Planet Hunters Coffee Chat (PHCC).[1]

7.3.1 Planet Hunters Coffee Chat

PHCC has shown that we can greatly increase scientific engagement and yield higher quality scientific outputs by providing citizen scientists with a set of 'teaching' videos and resources that allow them to interactively learn more about planet hunting and signal vetting. There are two main aspects to PHCC:

Providing tools and resources. To date, I have created 37 videos on topics ranging from how to analyse *TESS* data with python coding, determine exoplanet properties (e.g., radius, period), make the most of available online resources (e.g., SIMBAD, ExoFOP, ExoMAST), and determine whether a signal is a false positive. In each video I directly describe a tool that citizen scientists can use themselves, as well as the importance behind why it should be used (a selection of videos are shown in Figure 7.2). For example, there are a range of videos that discuss the importance of checking whether a planet signal is real or not. To complement these videos I provide online and downloadable python notebooks that allow the citizen scientists to directly carry out the vetting tests described in the videos themselves. Other videos discuss, for example, the importance of online resources and how and when to best use them; how to interpret stellar signals using frequency analysis; and how to make the most of the information provided in the SPOC pipeline generated data validation reports. All of the PHCC videos have a strong focus on providing not just information, but also tools and knowledge that citizen scientists can *directly* apply to carry out their own analysis.

Live office hours. I host regular 'live office hours' to connect citizen scientists with the PHT science team and to encourage volunteers to engage with the materials

[1] www.https://www.planethunters.coffee/; wwww.youtube.com/planethunterscoffee

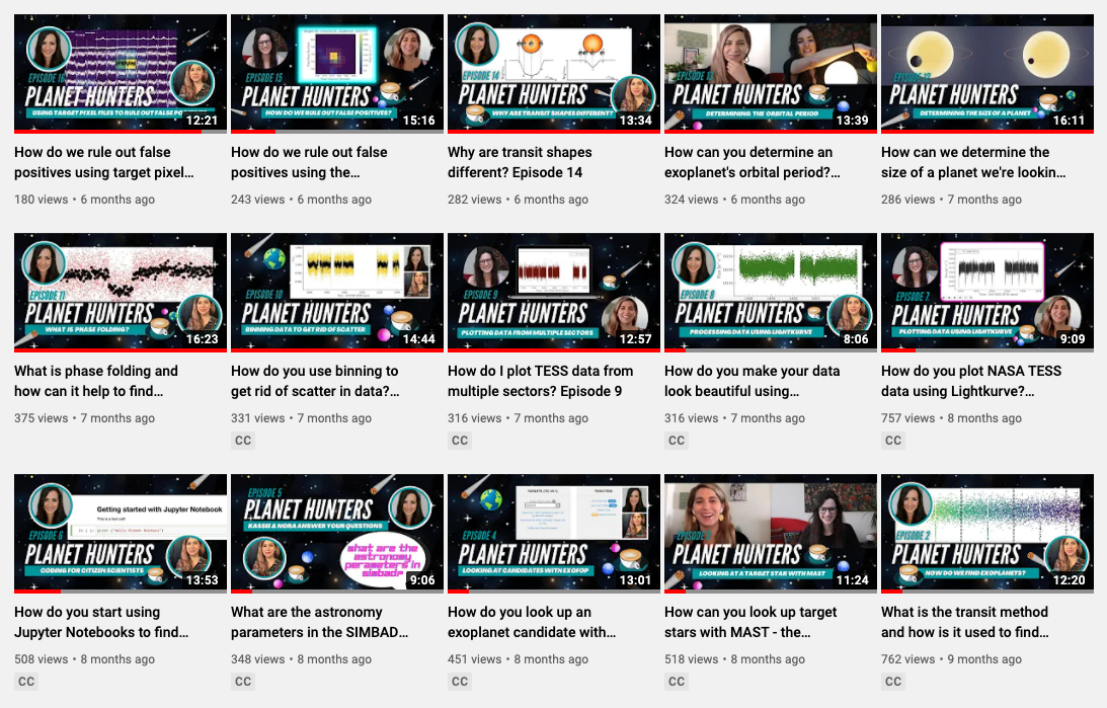

Figure 7.2: Example PHCC videos that have helped accelerate the scientific output of PHT. Videos are posted on YouTube and on the PHCC website weekly.

presented in the videos. Furthermore, I recently launched a scheme where citizen scientists can propose potential planet candidates directly to the PHT science team via an online accessible spreadsheet. Prior to proposing the candidates, the citizen scientists are prompted to use the resources provided in the videos in order to verify the signals to the best of their ability. This step helps the citizen scientists engage with the project and speeds up the identification of potential planet signals. The submitted targets are discussed in bi-weekly online vetting sessions that the citizen scientists are encouraged to attend. As discussed below, these sessions have already proven to greatly increase the scientific output of the PHT citizen science project, i.e. increased the total number of promising planet candidates flagged for further analysis and follow-up observations.

7.3.2 Increased scientific output

The primary aim of any citizen science project is to increase scientific output. As mentioned above, the PHCC live vetting sessions encouraged citizen scientists to

identify, vet, and propose potential planet candidates to the science team. An average of 10 candidates are submitted to each of these bi-weekly sessions, yielding around 8 candidates that are shared with the exoplanet community via the online *TESS* planet candidate database (ExoFOP) each month. As a comparison, without these sessions, the PHT project submits around 3 – 4 planet candidates to ExoFOP per month. This highlights a substantial increase in the scientific output that has resulted from the provided resources and live sessions. I note that, as these sessions are relatively new, there are currently no sectors for which we have completed both the pipeline flagged candidate analysis and have held a live vetting session. As such, I cannot comment on how many of the ~ 8 planet candidates found in the bi-weekly sessions will also be identified via the standard method outlined in Chapter 2.

In addition to the online vetting sessions, there has been a substantial increase in the amount of additional target analysis that is being carried out and posted on the PHT discussion forums, with citizen scientists regularly using tools including `LATTE` and `pyaneti`. Figure 7.3, for example, shows a kernel density estimate (KDE) of the occurrence of a selection of keywords on the PHT discussion forums that indicate that additional analysis has been carried out (navy blue line) and a histogram of the number of unique volunteers who used these keywords over time (light grey). The vertical orange line shows the time of launch of PHCC. The keywords used for this analysis are: 'LATTE', 'pyaneti', 'pixel level', 'jupyter notebook', 'false positive', 'background eclipsing binary', 'momentum dump', 'asteroid', 'SSO', 'background flux', 'centroid', 'nearby stars', 'vetting', 'transit model', 'impact parameter', and 'grazing'. All of these keywords relate to topics covered in the PHCC videos.

The figure highlights a significant increase in the level of engagement with the data following the launch of PHCC. This, in turn, has led to a higher quality and larger yield of planet candidates, as discussed above. Furthermore, while the tools have increased the rate and quality of identification of planet signals, it has also sped up the identification of false positives, addressing the major bottleneck that we are faced with in our signal verification pipeline. The work with

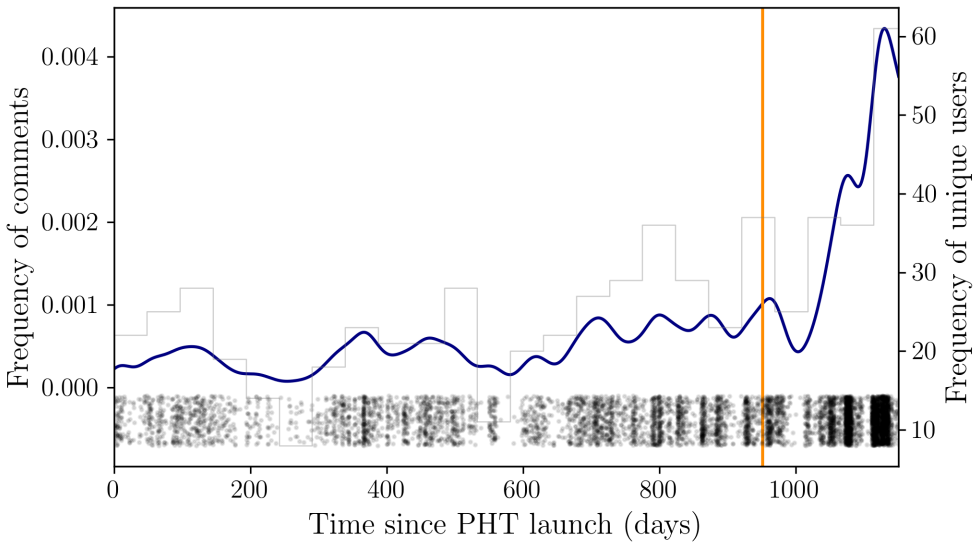

Figure 7.3: A KDE of the occurrence of keywords (including 'false positives', 'background flux', 'nearby stars') on the Planet Hunters TESS discussion forums indicate the level of engagement of the citizen scientists (navy blue line). The vertical orange lines indicated the time of launch of Planet Hunters Coffee Chat. This figure highlights that user engagement significantly increased following the release of the videos. The KDE bandwidth is 0.05. The periodic decreases correlate to times when there is no data on the site. The black points underneath the figure shows the frequency of the use of the selected keywords, where darker regions represent a higher frequency of use of these words. The grey histogram shows the number of unique users who used the keywords (bin size = 50 days).

PHCC will be an important tool to complement the work of PHT as the project continues to grow in the future.

7.4 Future work

The PHT project outlined in this thesis, has promising prospects for future growth and expansion. The next science case that I will wield PHT to address will be to search for planets around evolved stars that are observed in the *TESS* FFIs.

Due to the large number of active PHT volunteers, the 20 000 monthly released two-minute cadence *TESS* light curves are regularly classified within just two weeks of the *TESS* public data releases. This means that for around two weeks out of every month there is no data to be classified on the platform. Moving forward, I will make use of these times when there are (currently) no more two-minute cadence *TESS* light curves to classify to search for planets around pre-selected

evolved stars in the *TESS* FFI data.

In brief, I will carry out this work by making use of the growing *TESS* dataset (two- and ten-minute cadence) of evolved stars. Every two years, *TESS* will observe round 440 000 subgiants brighter than 12th magnitude in the V-band (selected via colour cuts using *Gaia* data). With the help of PHT, I will search these targets for transiting planets. This search will be a significant advance on what has already been found in the two-minute cadence *TESS* data and will be able to yield new planet candidates that are missed by automated transit searches.

Such a study of planets around evolved stars is extremely timely for numerous reasons: 1) While *TESS* has already demonstrated that it is able to detect planets around evolved stars (e.g., Nielsen et al., 2019b; Eisner et al., 2020b), the sample is currently small with 109 transiting planets that have mass measurements to better than 30% accuracy, only 14 of which have orbital periods longer than 30 days (as previously discussed in Chapter 4). Thus, more observations are needed to inform future theoretical advancements. 2) Several papers have recently been published on the detection of planets in the *TESS* FFI data, demonstrating the vast potential of finding planets in this, so far, underexplored data set (e.g., Battley et al., 2020; Feliz et al., 2021). 3) With the increase of the *TESS* FFI cadence from 30- to 10-minutes, we see both an increase in the signal-to-noise ratio of the transit detection, as well as a threefold increase in the Nyquist frequency. This has enabled asteroseismic detection of subgiants across the whole sky, instead of just in the pre-selected two-minute cadence stars, and has made subgiants the largest population of potential planet hosting stars that have detectable asteroseismic oscillations. 4) Current ongoing studies are comparing *TESS* main sequence and red giant branch planet populations, but are missing out subgiants (e.g., Grunblatt et al., 2019). While these searches are strongly biased towards finding planets with periods shorter than $\sim$30 days, the citizen science approach does not show a period bias. As such, I will provide the missing link between these two populations by utilising a period-unbiased approach to finding planets around subgiants using PHT.

As shown in Chapters 3 and 4, the PHT project is well-suited to aid in collating a large sample of planets around subgiants, bridging existing studies in exoplanetary science and enabling me to address the interdisciplinary questions of how galactic evolution and stellar physics affects planet properties. I aim to carry out this work over the next couple of years with the help of the PHT community.

Looking further ahead, the automated detection algorithms designed to look for transit events in the data collected by the upcoming PLATO mission (set to launch in 2026) will have similar biases to the automated pipelines used to find planets in *TESS* data. As such, I hope to build on the success of PH and PHT to develop a similar citizen science project in the future that will search for unique and interesting signals in the PLATO data. Furthermore, by the time that PLATO is launched, machine learning algorithms designed to search for transit signals will also have likely made great advances. As such, it is possible that future iterations of PHT and PH PLATO will combine human and machine classifications in order to optimise the scientific output. Such a combined approach has already been shown to be successful by other citizen science projects, such as Supernova Hunters and Galaxy Zoo, which were able to achieve higher accuracy with the combination of machine and human classifiers than either one approach could attain by itself (e.g., Wright et al., 2017; Walmsley et al., 2020).

Finally, I hope to continue to contribute to the library of open-source software. As I showed with the development of `LATTE` and PHCC, tools that can be used by citizen scientists and professional astronomers alike, can help accelerate scientific research. Since publishing the first version of the `LATTE` software, I have expanded the package to include simple, automated, asteroseismic analysis, as well as automated tests to investigate signals from eclipsing binaries. Moving forward, `LATTE` could be adapted to include transit modelling, analysis of eclipse timing variations, and the automated removal of underlying stellar signals.

7.5 Final word

The Universe consists of hundreds of billions of galaxies, each of which is made up of around hundred billion stars. The age-old question of how many of these stars host planets has caused humankind to study the skies for millennia. With the help of a global community, I have been able to help contextualise the work of early philosophers and astronomers, and contribute to our current understanding of the plurality of worlds.

Although we still do not know if we are alone in the Universe, the dedication of thousands of citizen scientists has demonstrated a continued global interest in stars, planets, and the possibility of further life. At this point in time, one can only postulate that perhaps there is an equally inhabited world like our own; a world that runs a people-powered planet hunt in the search for a tiny dip in light that marks the existence of our very own *pale blue dot* – the Earth.

Appendices

APPENDIX A

ORBITAL GEOMETRY

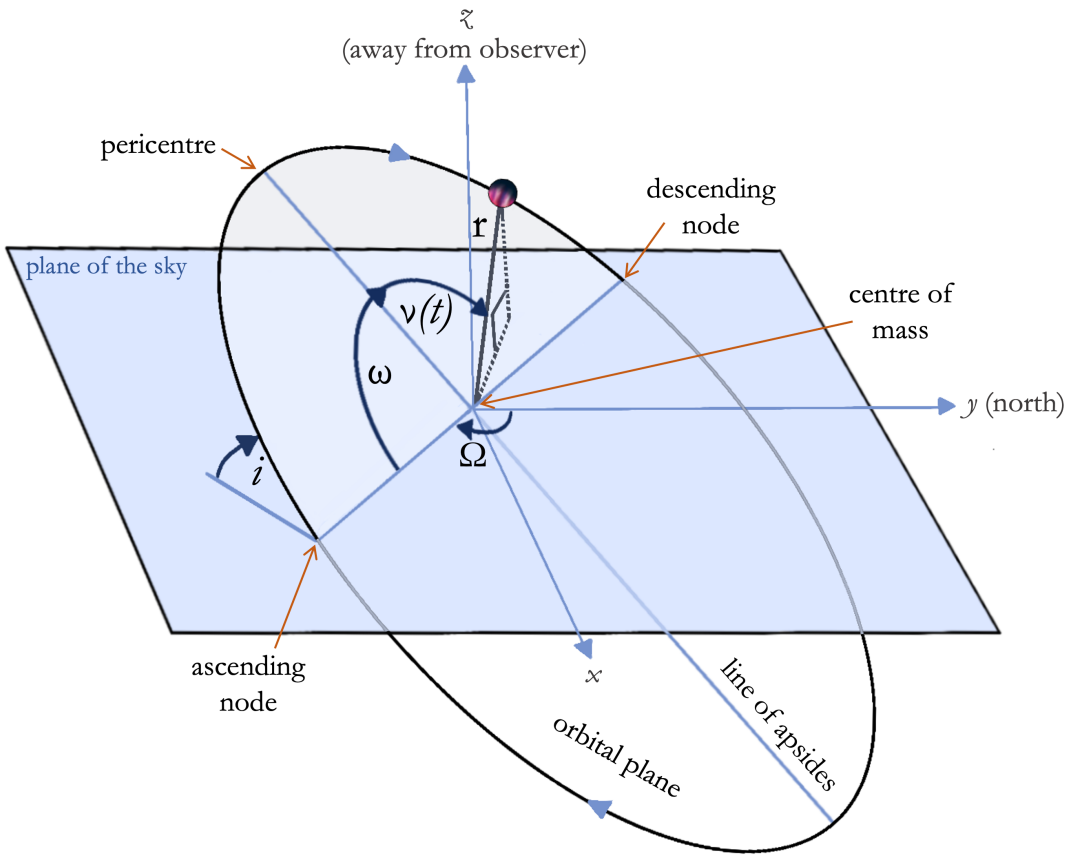

Figure A.1: Schematic of an elliptical orbit. The observer is in the negative z direction; i is the orbital inclination; Ω is the longitude of the ascending node; ω is the argument of periastron; and ν(t) is the true anomaly. The ascending and descending nodes are the points of intersection of the planetary orbit and the reference plane, where the ascending node is the point where the orbiting body crosses the plane of the sky moving *away* from the observer. The eccentricity of the orbit is defined as the ratio of the distance between the centre of the ellipse and the focus (centre of mass) to the length of the semi-major axis.

APPENDIX B

EXAMPLE LATTE REPORTS

Two example LATTE reports: TOI 813 b and TIC 88869423. While the former is a confirmed planet (Chapter 4), the latter is a signal that was ruled out as a false positive based on its LATTE report.

PHT Data Validation Report
TIC 55525572

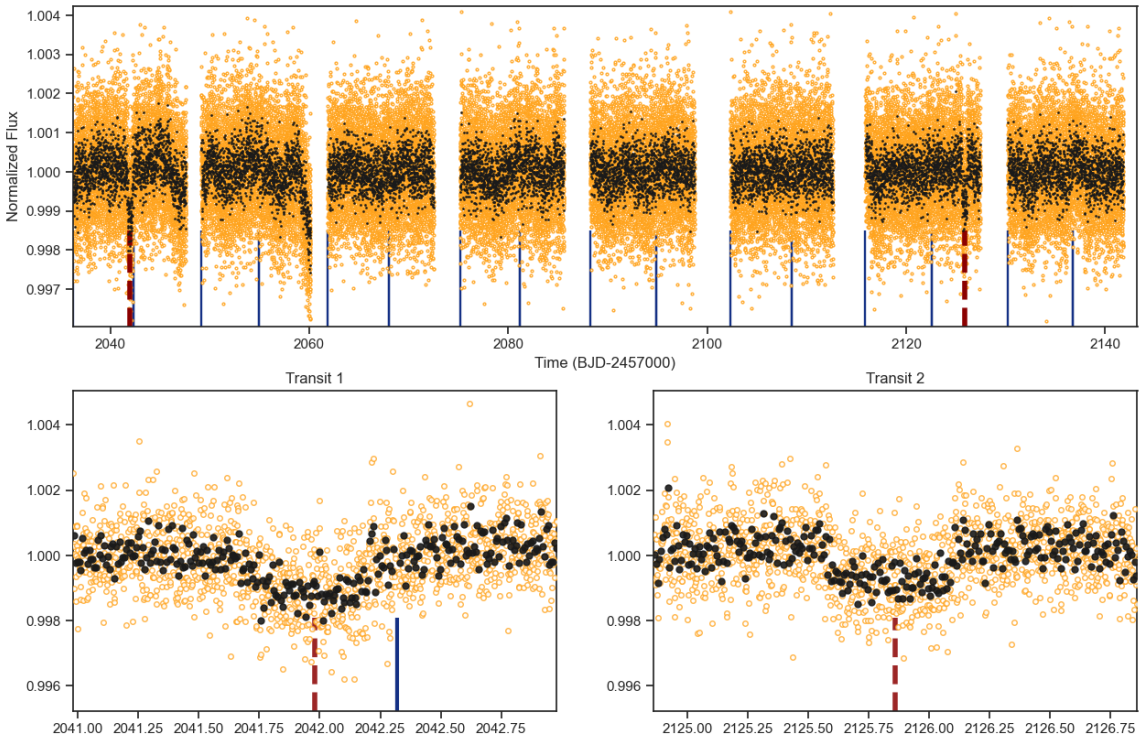

Fig 1. Full lightcurve for target TIC 55525572. The solid blue lines at the bottom of the figure indicated the times of the reaction wheel momentum dumps and the dashed black line(s) show the time(s) of the marked transit event(s). Momentum dumps occur around every 2 to 2.5 days and typically last around half an hour.

Parameter	Value	Unit
TIC ID	55525572	
Other name	8876-01059-1	
RA/Dec	72.694 -60.9055	degrees
Radius	2.04	Solar Radii
Mass	1.05	Solar Mass
Teff	5824.0	Kelvin
Parallax	3.77	
T mag	9.8467	Mag
V mag	10.36	Mag
Sectors (nominal)	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13 *	
Sectors (extended)	27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 61, 62, 63, 64, 65, 66, 67, 68, 69 *	
TCE	-	
TOI	813.01	

Table 1. Stellar properties of the TIC 55525572. * List of the sectors in which the target will be, or has been, observed.

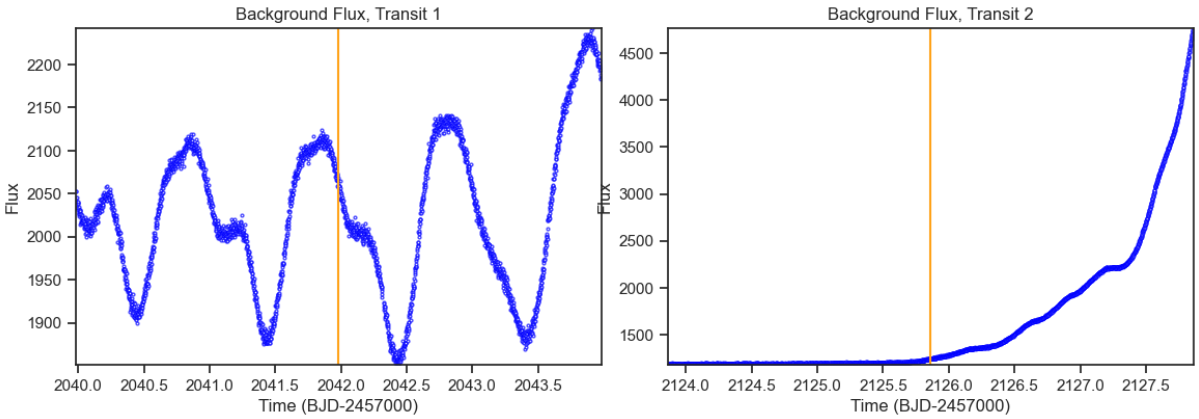

Fig 2. Background flux vs. time around the time of each transit-like event. The vertical orange line indicates the time of the transit-like event.

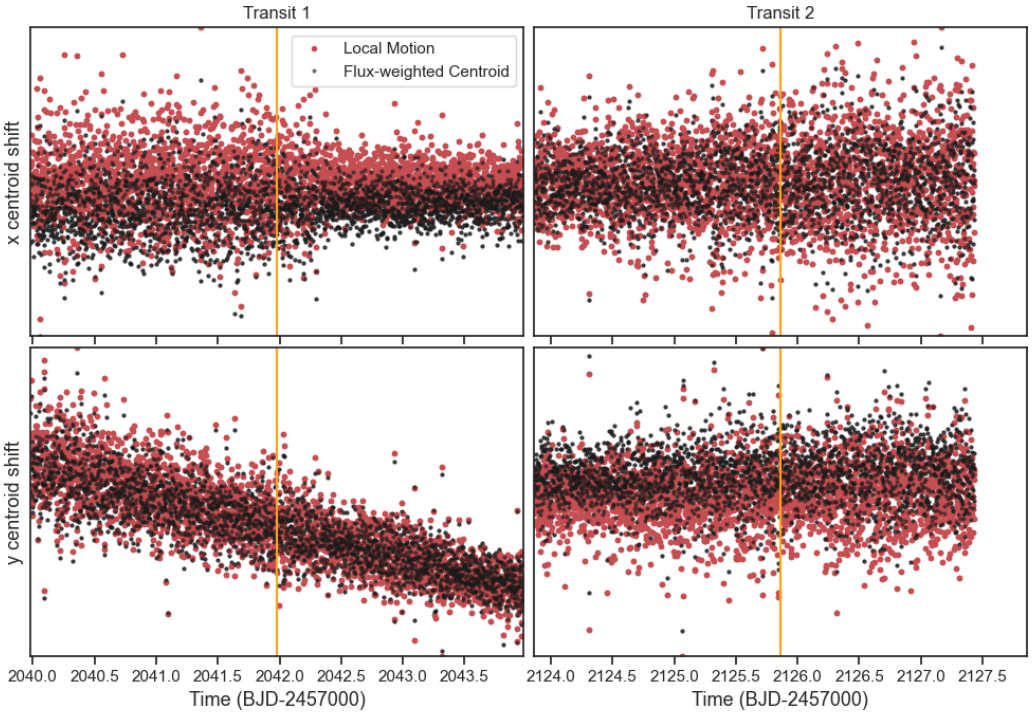

Fig 3. The x and y centroid positions around the time of each transit-like event. The black points shows the CCD column and row position of the target's flux-weighted centroid. The red shows the CCD column and row local motion differential velocity aberration (DVA), pointing drift, and thermal effects. The vertical orange line indicates the time of the transit-like event

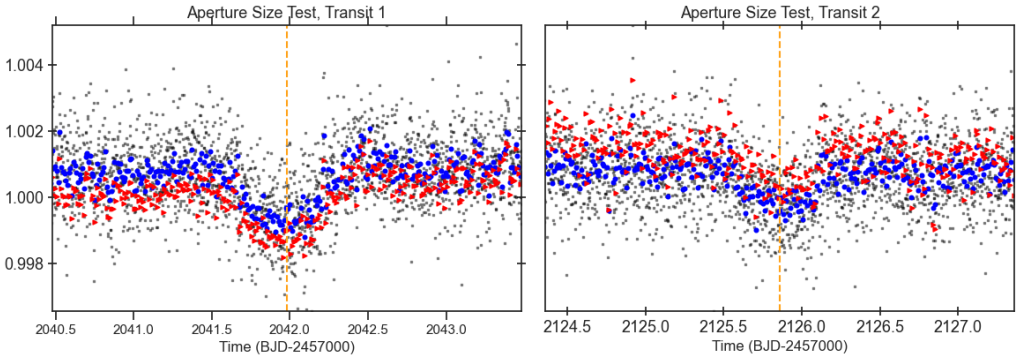

Fig 4. The lightcurve around the time of each transit-like event extracted with the SPOC pipeline defined aperture (binned:blue, unbinned:grey) and the with an aperture that is 40 per cent smaller (red). The flux is extracted from the target pixel files (TPFs) and has not been detrended or corrected for systematics. The vertical orange line indicates the time of the transit-like event.

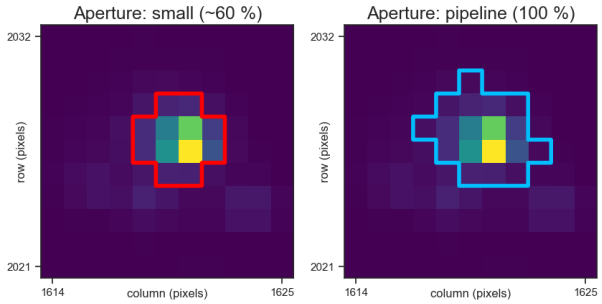

Fig 5. The apertures used to extract the lightcurves. The blue aperture on the right shows the optimum aperture determined by the SPOC pipeline, which is used for the extraction of 2-minute cadence light curves shown in Figure 1. The red outline on the left shows an aperture that is around 40 per cent smaller than the SPOC pipeline aperture which was used to extract the red lightcurve shown in Figure 4.

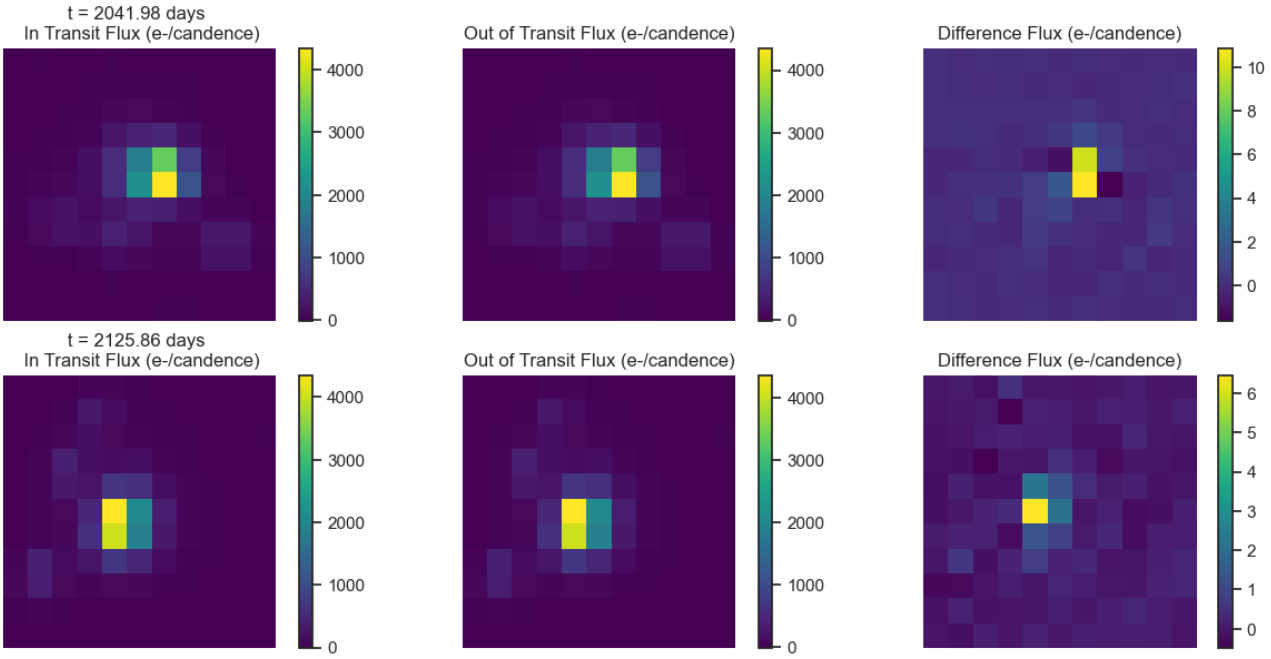

Fig 6. Difference images for target TIC 55525572 for each transit like event. Left: mean in-transit flux(left). Middle: mean out-of-transit flux. Right: difference between the mean out-of-transit and mean in-transit flux. Ensure that the change in brightness occurs on target.

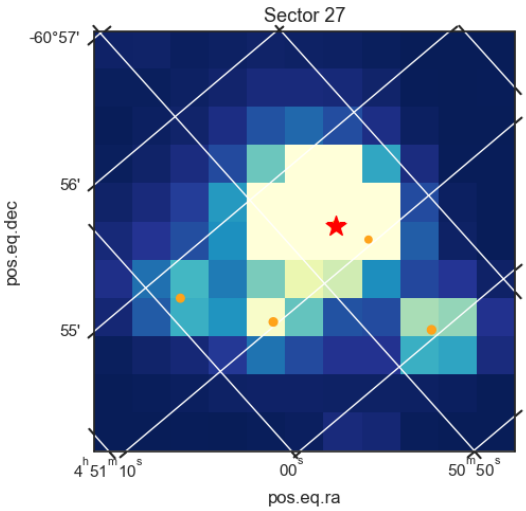

Fig 7. The locations of nearby GAIA DR2 stars with a magnitude difference less than 5 (orange circle) within the Tess Cut Out around TIC 55525572 (red star). Only shown for one sector. Right: SDSS image of the surrounding field.

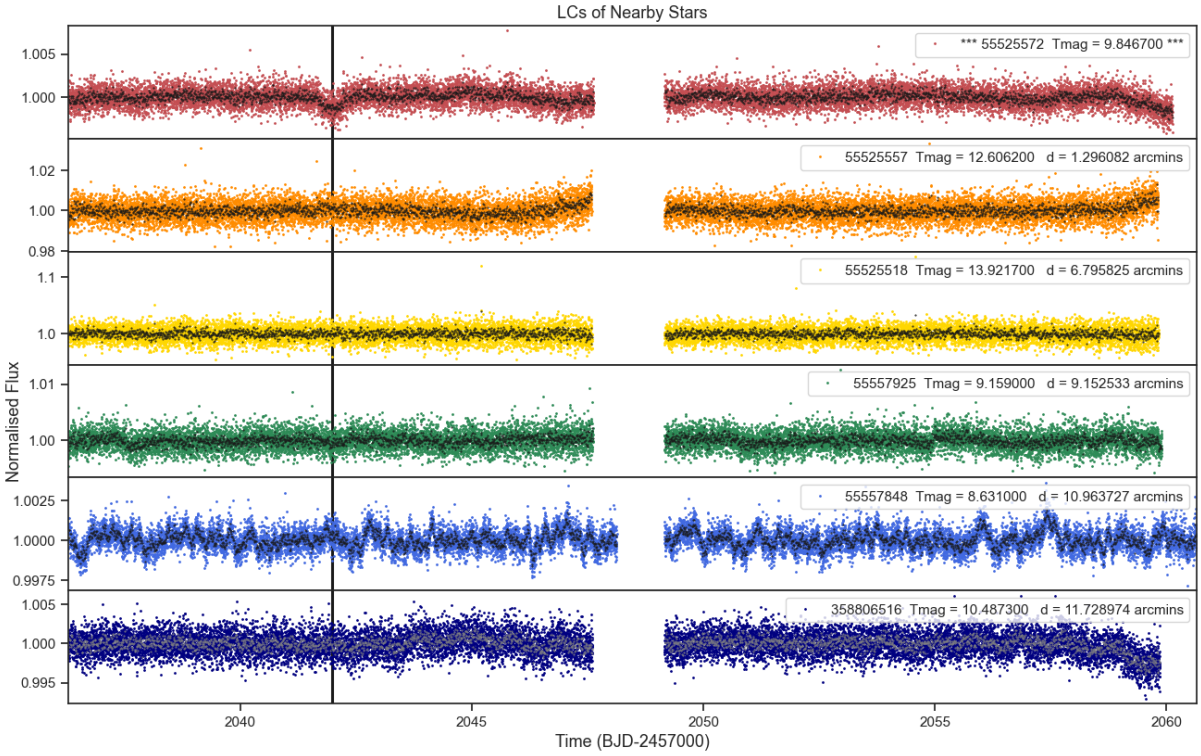

Fig 8. Lightcurves of the five closest stars to target 55525572 (top pannel). The distances to the target star and the TESS magnitudes are shown for each star. Only ever shown for one sector.

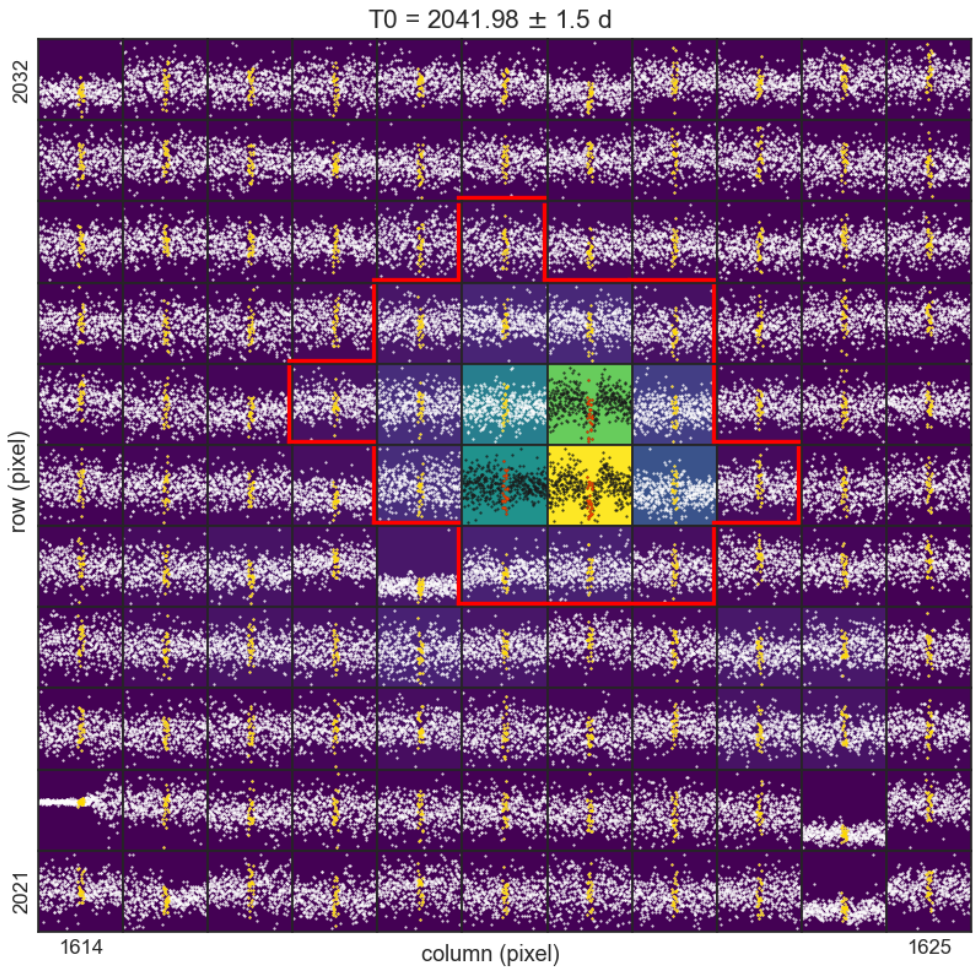

Fig 9. Normalised flux extracted for each pixel, using the SPOC pipeline mask, around the time of the transit-like event. The orange/red data points show the in-transit data. The solid red lines show the SPOC pipeline mask. Only shown for one sector.

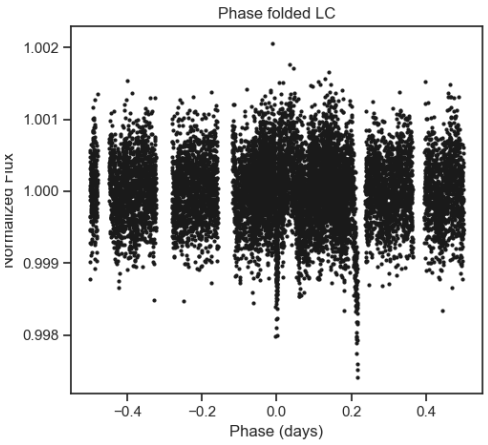

Fig 10. Phase folded lightcurve where the odd and the even transits are shown in different colours. Ensure that the odd and even transits have comparabel shapes and depths.

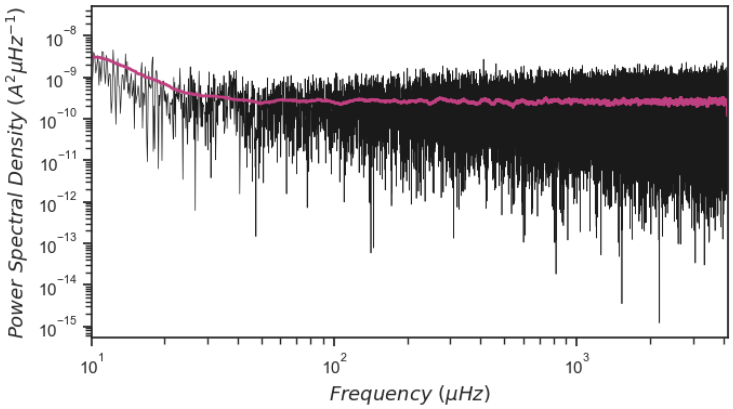

Fig 11. Lomb scargle power spectrum of the TESS lightcurve (black line) and a boxcar-smoothed periogram (pink line) computed with a window length of 20 micro Hz.

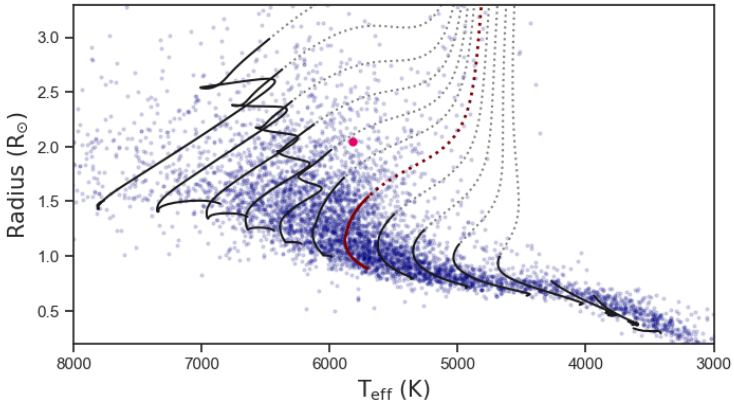

Fig 12. The equivalent evolutionary phase (eep) tracks for main sequence evolution (solid lines) and post main-sequence evolution (dashed lines) for masses ranging from 0.3 to 1.6 solar masses (from right to left). The 1 Solar Mass track is shown in maroon. The blue points show the TOIs and the magenta point TIC 55525572.

PHT Data Validation Report
TIC 88869423

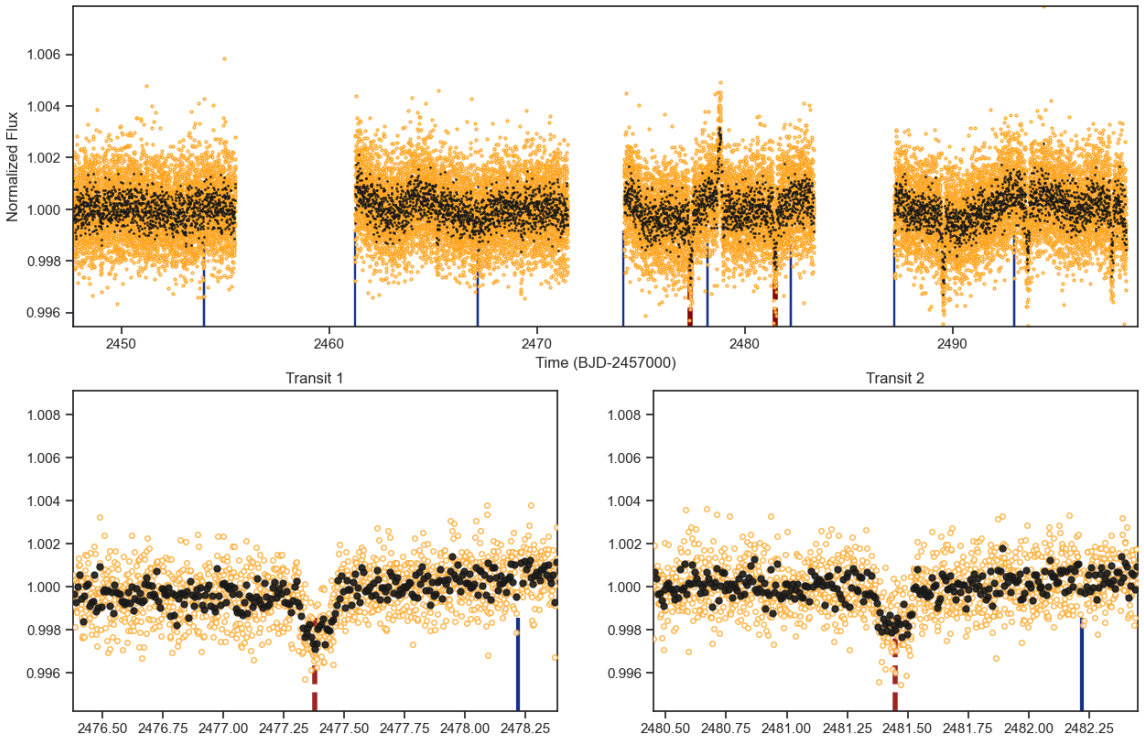

Fig 1. Full lightcurve for target TIC 88869423. The solid blue lines at the bottom of the figure indicated the times of the reaction wheel momentum dumps and the dashed black line(s) show the time(s) of the marked transit event(s). Momentum dumps occur around every 2 to 2.5 days and typically last around half an hour.

Parameter	Value	Unit
TIC ID	88869423	
Other name	TYC 0628-00969-1	
RA/Dec	27.7852 13.904	degrees
Radius	0.88	Solar Radii
Mass	0.97	Solar Mass
Teff	5521.0	Kelvin
Parallax	9.05	
T mag	9.95806026	Mag
V mag	10.53	Mag
Sectors (nominal)	*	
Sectors (extended)	42, 43 *	
TCE	Yes **	
TOI	-	

Table 1. Stellar properties of TIC 88869423. * List of the sectors in which the target will be, or has been, observed. ** Click here for the TCE report.

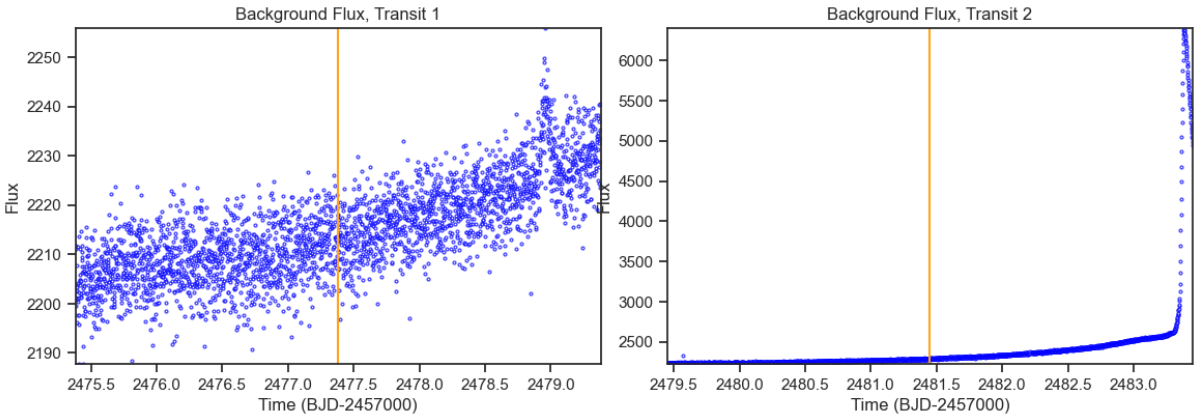

Fig 2. Background flux vs. time around the time of each transit-like event. The vertical orange line indicates the time of the transit-like event.

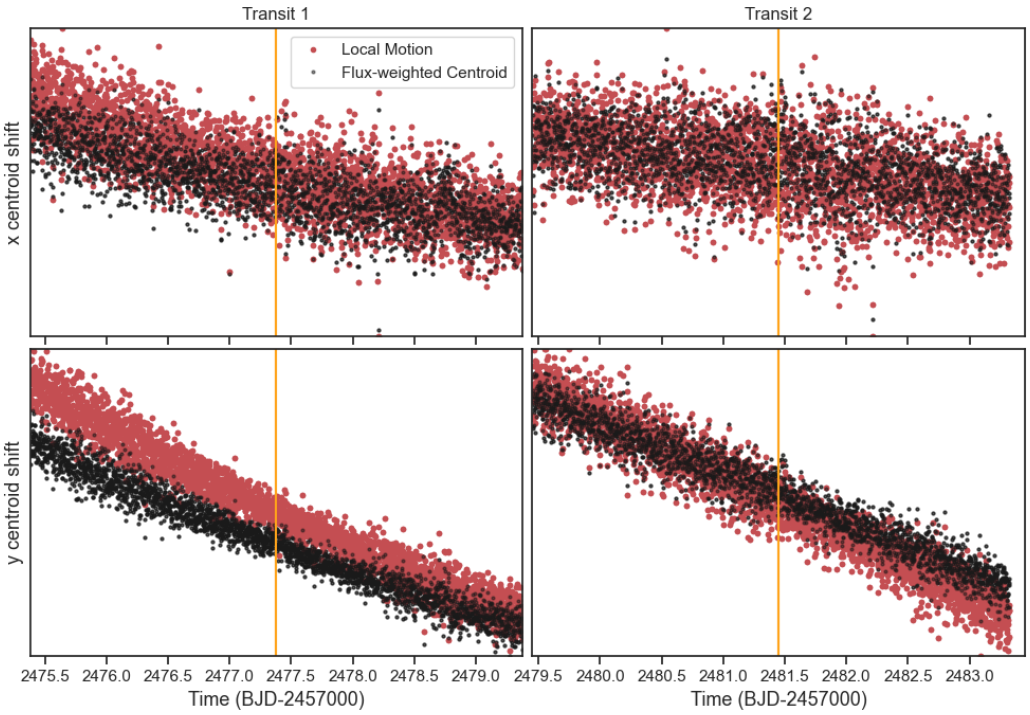

Fig 3. The x and y centroid positions around the time of each transit-like event. The black points shows the CCD column and row position of the target's flux-weighted centroid. The red shows the CCD column and row local motion differential velocity aberration (DVA), pointing drift, and thermal effects. The vertical orange line indicates the time of the transit-like event

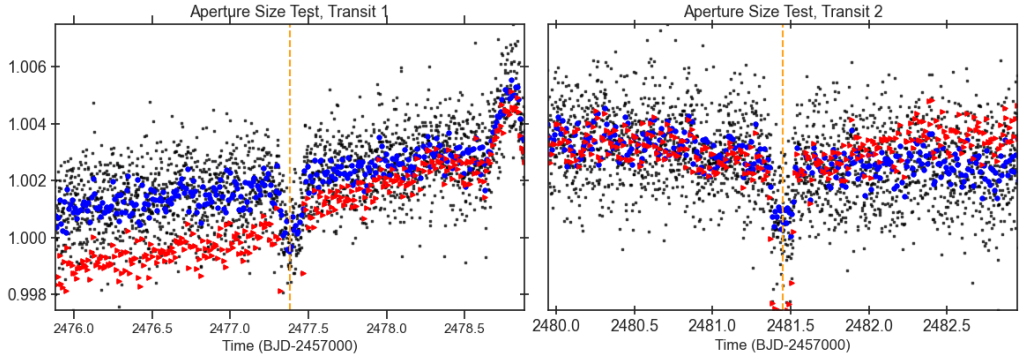

Fig 4. The lightcurve around the time of each transit-like event extracted with the SPOC pipeline defined aperture (binned:blue, unbinned:grey) and the with an aperture that is 40 per cent smaller (red). The flux is extracted from the target pixel files (TPFs) and has not been detrended or corrected for systematics. The vertical orange line indicates the time of the transit-like event.

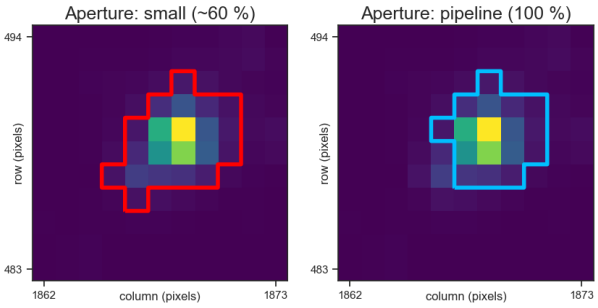

Fig 5. The apertures used to extract the lightcurves. The blue aperture on the right shows the optimum aperture determined by the SPOC pipeline, which is used for the extraction of 2-minute cadence light curves shown in Figure 1. The red outline on the left shows an aperture that is around 40 per cent smaller than the SPOC pipeline aperture which was used to extract the red lightcurve shown in Figure 4.

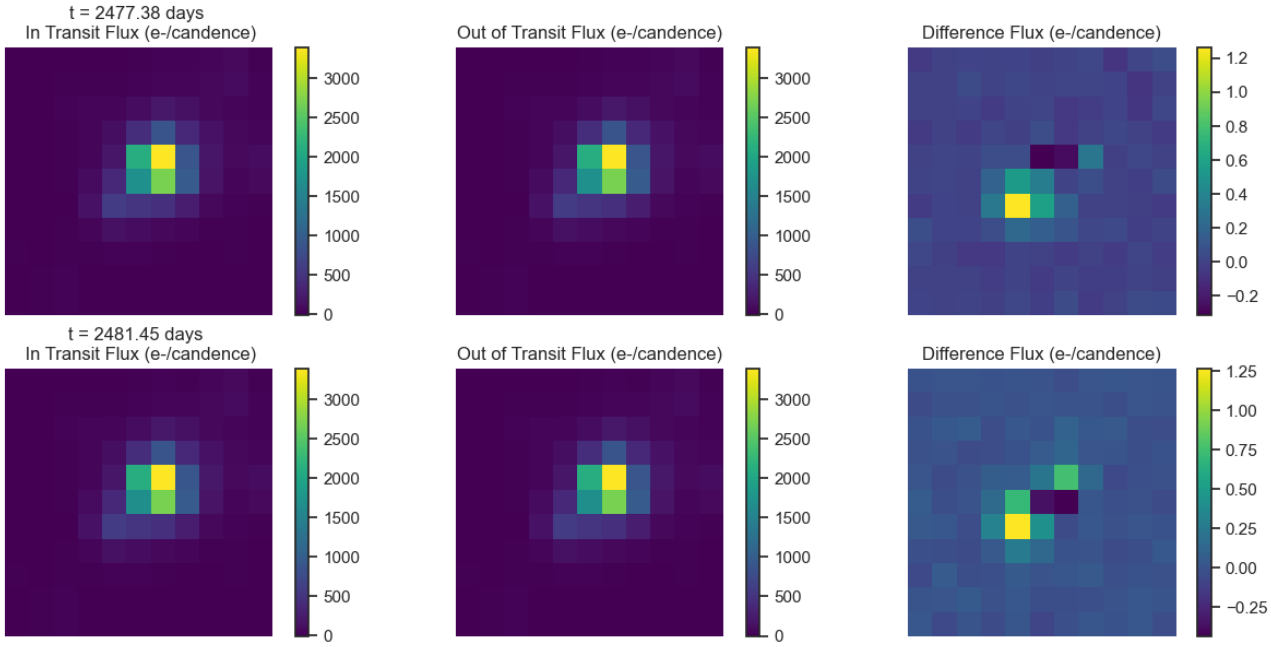

Fig 6. Difference images for target TIC 88869423 for each transit like event. Left: mean in-transit flux(left). Middle: mean out-of-transit flux. Right: difference between the mean out-of-transit and mean in-transit flux. Ensure that the change in brightness occurs on target.

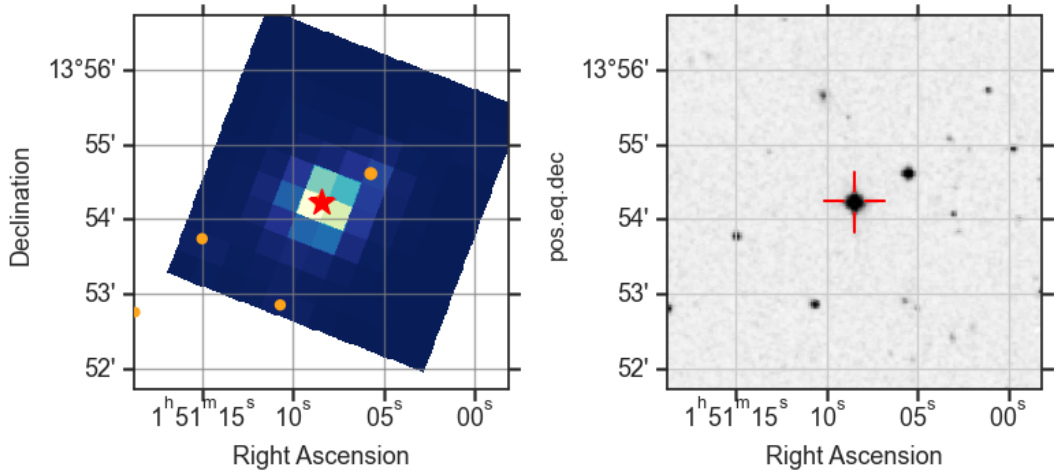

Fig 7. The locations of nearby GAIA DR2 stars with a magnitude difference less than 5 (orange circle) within the Tess Cut Out around TIC 88869423 (red star). Only shown for one sector. Right: SDSS image of the surrounding field.

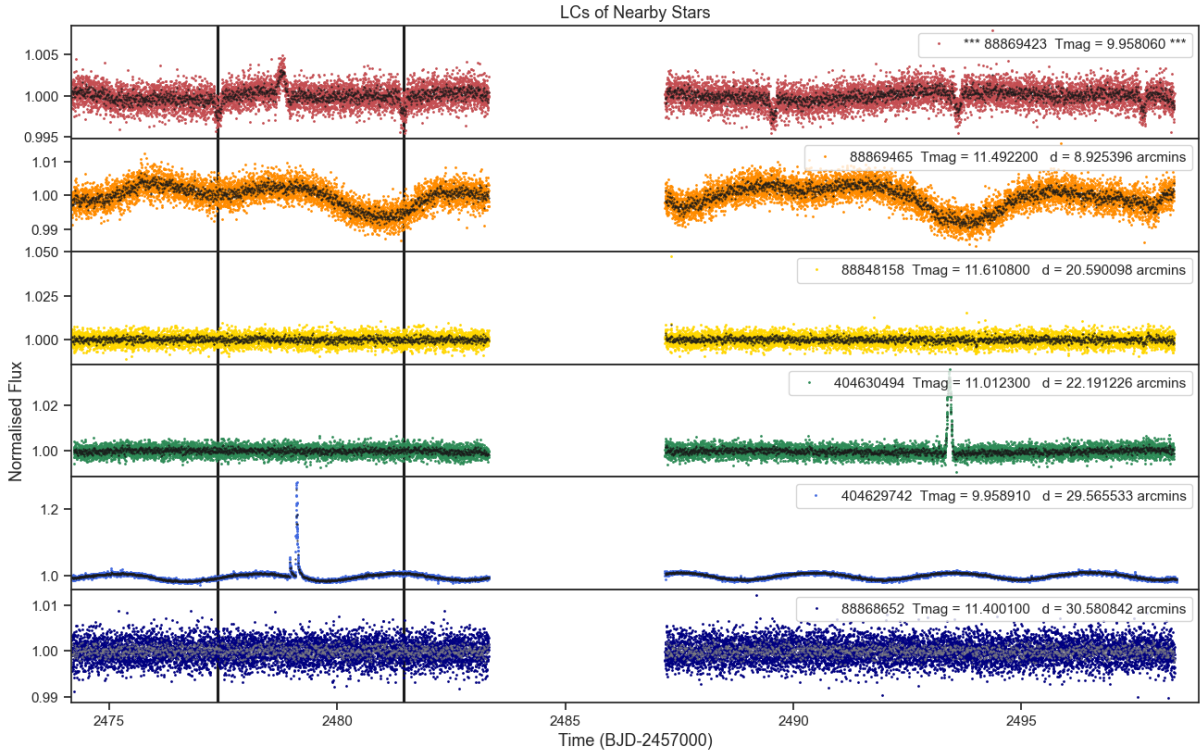

Fig 8. Lightcurves of the five closest stars to target 88869423 (top pannel). The distances to the target star and the TESS magnitudes are shown for each star. Only ever shown for one sector.

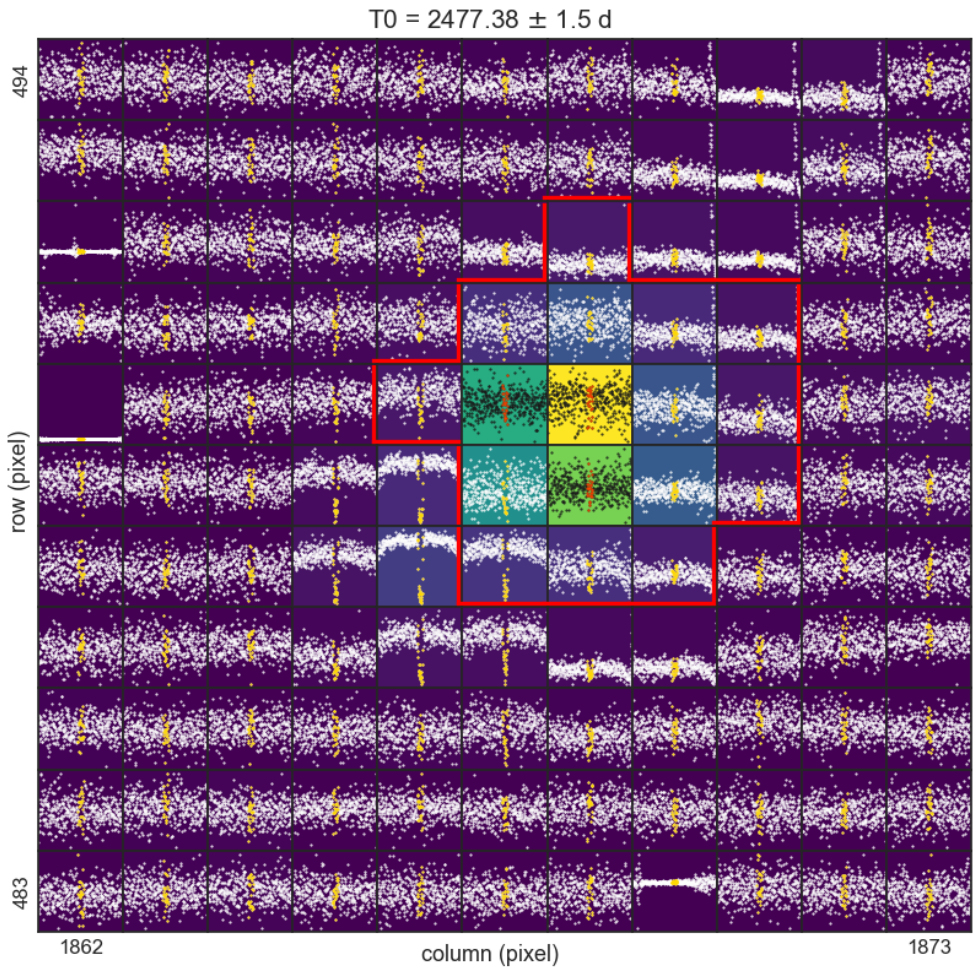

Fig 9. Normalised flux extracted for each pixel, using the SPOC pipeline mask, around the time of the transit-like event. The orange/red data points show the in-transit data. The solid red lines show the SPOC pipeline mask. Only shown for one sector.

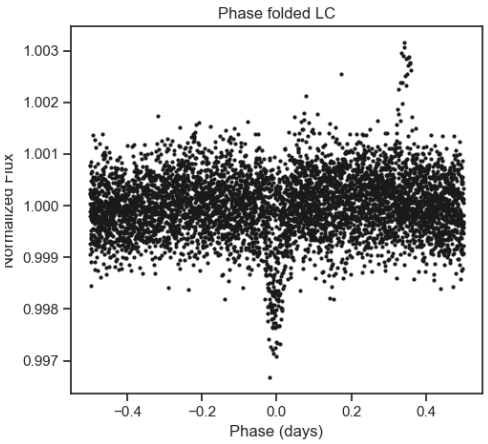

Fig 10. Phase folded lightcurve where the odd and the even transits are shown in different colours. Ensure that the odd and even transits have comparabel shapes and depths.

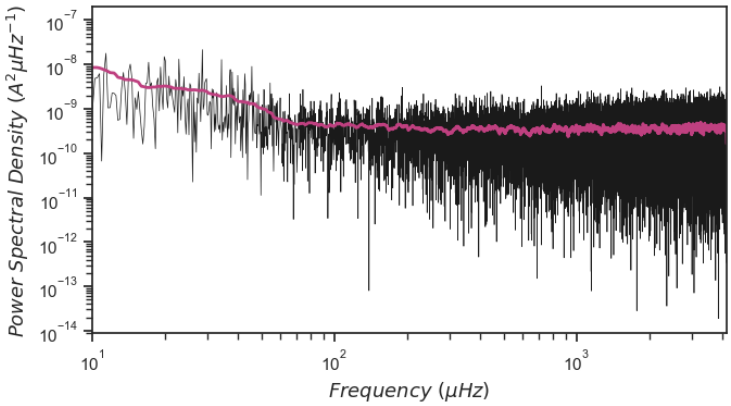

Fig 11. Lomb scargle power spectrum of the TESS lightcurve (black line) and a boxcar-smoothed periogram (pink line) computed with a window length of 20 micro Hz.

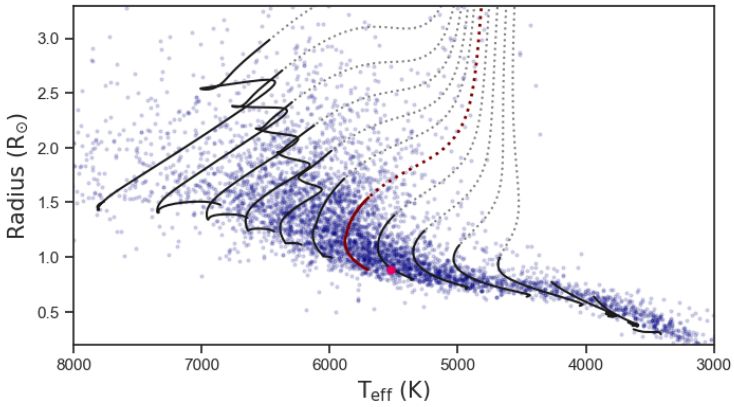

Fig 12. The equivalent evolutionary phase (eep) tracks for main sequence evolution (solid lines) and post main-sequence evolution (dashed lines) for masses ranging from 0.3 to 1.6 solar masses (from right to left). The 1 Solar Mass track is shown in maroon. The blue points show the TOIs and the magenta point TIC 88869423.

APPENDIX C

TIC 470710327: ADDITIONAL TABLES

Summary of the RV observations obtained with the HERMES spectrograph (Section 6.4), ETVs determined from the *TESS* data (Section 6.3.3) and inputs used for the SED analysis (Section 6.6.2).

Time (BJD - 2457000)	RV (kms^{-1})	RV err (kms^{-1})	SNR
1880.3395	0.03	5.55	45
1880.3518	4.22	5.47	45
1883.3908	24.25	7.35	27
1884.3813	46.12	4.15	59
1885.3333	43.12	1.49	85
1886.3368	51.28	2.02	80
1887.3369	42.11	0.90	91
1888.3537	32.07	2.45	75
1889.3285	23.70	1.11	89
1890.3321	18.43	2.19	78
1891.3247	14.30	3.45	66
1892.3270	9.70	5.30	47
2007.7168	-36.85	0.84	92
2035.6305	-10.77	2.99	70
2038.6952	18.08	4.00	60
2040.6938	29.27	0.79	92
2041.6919	42.08	0.80	92
2042.6432	44.82	1.55	85
2043.6680	47.71	1.58	84
2067.7128	-74.50	0.88	91
2069.7160	-86.61	0.87	91
2071.7410	-95.40	4.04	60
2072.7397	-105.13	0.39	96
2127.5437	-105.09	1.85	82

Table C.1: RV observations for TIC 470710327.

Primary cycle number	Predicted epoch (BJD - 2457000)	O-C (mins)	Error (mins)	Secondary cycle number	Predicted epoch (BJD - 2457000)	O-C (mins)	Error (mins)
0	1766.2700	-0.684	0.582	0.5	1766.8223	0.552	0.773
1	1767.3747	1.147	0.583	1.5	1767.9270	0.284	0.775
2	1768.4794	1.510	0.577	2.5	1769.0317	1.908	0.789
3	1769.5841	1.908	0.594	3.5	1770.1364	2.266	0.805
4	1770.6887	2.403	0.592	4.5	1771.2411	4.902	0.796
5	1771.7934	1.995	0.590	5.5	1772.3458	2.988	0.786
6	1772.8981	3.012	0.604	6.5	1773.4505	-4.964	0.792
7	1774.0028	2.351	0.603	7.5	1774.5551	3.468	0.821
11	1778.4215	-1.554	0.599	10.5	1777.8692	0.403	0.850
12	1779.5262	-0.756	0.584	11.5	1778.9739	0.303	0.784
13	1780.6309	0.378	0.581	12.5	1780.0786	0.757	0.829
14	1781.7356	-0.343	0.609	13.5	1781.1833	-0.089	0.791
15	1782.8403	-1.649	0.589	14.5	1782.2879	0.185	0.807
16	1783.9450	-2.364	0.602	15.5	1783.3926	-1.492	0.779
17	1785.0497	-1.846	0.580	16.5	1784.4973	-2.271	0.794
18	1786.1543	-2.562	0.594	18.5	1786.7067	-3.702	0.791
19	1787.2590	-2.571	0.598	19.5	1787.8114	-2.883	0.804
24	1792.7825	-4.193	0.599	23.5	1792.2301	-3.806	0.763
25	1793.8871	-10.852	0.597	24.5	1793.3348	-5.062	0.774
26	1794.9918	-4.375	0.644	25.5	1794.4395	-4.309	0.836
27	1796.0965	-4.856	0.577	27.5	1796.6489	-5.181	0.763
28	1797.2012	-5.707	0.563	28.5	1797.7535	-3.611	0.768
29	1798.3059	-5.709	0.575	29.5	1798.8582	-4.852	0.784
30	1799.4106	-5.429	0.591	30.5	1799.9629	-5.658	0.776
31	1800.5153	-4.098	0.605	31.5	1801.0676	-4.506	0.775
34	1803.8293	-4.444	0.581	34.5	1804.3816	-5.796	0.804
35	1804.9340	-4.312	0.584	36.5	1806.5910	-4.506	0.809
36	1806.0387	0.786	0.554	37.5	1807.6957	-4.097	0.810
37	1807.1434	-3.490	0.568	38.5	1808.8004	-3.854	0.777
38	1808.2480	-3.775	0.649	39.5	1809.9051	-2.949	0.835
39	1809.3527	-2.923	0.562	40.5	1811.0098	-2.466	0.834
40	1810.4574	-3.381	0.563	41.5	1812.1144	-2.277	0.770
41	1811.5621	-1.754	0.566	42.5	1813.2191	-1.383	0.811
42	1812.6668	-2.763	0.608	173.5	1957.9329	-2.145	0.897
43	1813.7715	-1.453	0.587	174.5	1959.0376	-1.105	0.832
172	1956.2759	-1.853	0.587	175.5	1960.1423	-1.108	0.793
174	1958.4853	-2.323	0.625	176.5	1961.2470	-3.146	0.809
175	1959.5900	-3.002	0.590	177.5	1962.3517	-4.163	0.813
176	1960.6946	-1.307	0.591	178.5	1963.4564	-1.103	0.814
177	1961.7993	-1.815	0.592	179.5	1964.5610	0.615	0.783
178	1962.9040	-1.221	0.594	180.5	1965.6657	1.774	0.794
179	1964.0087	-3.575	0.594	181.5	1966.7704	-1.163	0.793
181	1966.2181	0.422	0.629	182.5	1967.8751	-0.928	0.813
182	1967.3228	-1.289	0.599	184.5	1970.0845	-0.943	0.816
184	1969.5321	-0.969	0.583	186.5	1972.2938	2.103	0.830
186	1971.7415	2.736	0.627	187.5	1973.3985	2.611	0.797
187	1972.8462	1.779	0.615	188.5	1974.5032	1.298	0.807
188	1973.9509	-3.132	0.593	189.5	1975.6079	1.833	0.832
190	1976.1602	2.088	0.657	190.5	1976.7126	3.443	0.803
191	1977.2649	3.076	0.611	191.5	1977.8173	7.375	0.819
192	1978.3696	4.212	0.594	192.5	1978.9220	3.657	0.796
193	1979.4743	4.072	0.623	193.5	1980.0266	4.875	0.860
194	1980.5790	3.980	0.597	194.5	1981.1313	4.909	0.859
195	1981.6837	3.632	0.582				

Table C.2: O-C values for the primary and secondary eclipses.

Table C.3: Photometric observations of TIC 470710327. Data obtained from VizieR (http://vizier.unistra.fr/vizier/sed/).

Passband	Effective wavelength [Å]	Flux [erg s^{-1}cm^{-2}Å^{-1}]	Magnitude	Ref.
Johnson:U	3971.00	6.52e-13 ± 3.26e-14	9.38	a
HIP:BT	4203.01	6.40e-13 ± 1.19e-14	9.40	b
HIP:BT	4203.01	6.65e-13 ± 1.53e-14	9.36	c
Johnson:B	4442.03	6.17e-13 ± 1.37e-14	9.44	d
Johnson:B	4442.03	6.47e-13 ± 1.82e-14	9.39	e
SDSS:g'	4819.97	7.88e-13 ± 7.62e-16	9.17	f
HIP:VT	5318.96	5.30e-13 ± 9.54e-15	9.60	b
HIP:VT	5318.96	5.51e-13 ± 1.27e-14	9.56	c
Johnson:V	5537.05	5.48e-13 ± 5.28e-14	9.57	d
Johnson:V	5537.05	4.73e-13 ± 9.78e-15	9.73	f
SDSS:r'	6246.98	4.70e-13 ± 1.54e-14	9.73	f
Gaia:G	6729.95	3.08e-13 ± 4.63e-15	10.19	h
SDSS:i'	7634.91	3.17e-13 ± 7.19e-16	10.16	f
2MASS:J	12390.1	9.94e-14 ± 1.76e-15	11.42	i
Johnson:J	12500.21	9.96e-14 ± 1.92e-15	11.42	j
Johnson:H	16300.16	4.08e-14 ± 1.02e-15	12.39	j
2MASS:H	16494.77	4.02e-14 ± 1.10e-15	12.40	i
2MASS:Ks	21637.85	1.52e-14 ± 2.56e-16	13.46	i
Johnson:K	21900.25	1.46e-14 ± 1.25e-16	13.50	j
WISE:W1	33500.11	2.99e-15 ± 5.34e-17	15.23	k
Johnson:L	34000.10	2.72e-15 ± 5.19e-17	15.33	c
WISE:W2	46000.19	8.56e-16 ± 1.56e-17	16.58	k
Johnson:M	50299.90	6.64e-16 ± 1.18e-17	16.86	c
WISE:W3	115598.23	2.24e-17 ± 6.73e-19	20.54	k
WISE:W4	220906.68	2.98e-18 ± 3.93e-19	22.73	k

Notes. Data from: [(a)] Reed (2003); [(b)] Fabricius et al. (2002); [(c)] Bourges et al. (2017); [(d)] Ammons et al. (2010); [(e)] Lasker et al. (2007); [(f)] Henden et al. (2016); [(g)] Droege et al. (2007); [(h)] Altmann et al. (2017); [(i)] Zacharias et al. (2012); [(j)] Cutri et al. (2003); [(k)] Cutri & et al. (2013).

I think the answer is: a circle that has no beginning.

— L. Lovegood, 1997

Bibliography

Abazajian K. N., et al., 2009, ApJS, 182, 543

Abdul-Masih M., et al., 2021, A&A, 651, A96

Addison B., et al., 2019, PASP, 131, 115003

Aerts C., Christensen-Dalsgaard J., Kurtz D. W., 2010, Asteroseismology

Agol E., Steffen J., Sari R., Clarkson W., 2005, MNRAS, 359, 567

Aigrain S., Irwin M., 2004, MNRAS, 350, 331

Akeson R. L., et al., 2013, PASP, 125, 989

Aldoretta E. J., et al., 2015, AJ, 149, 26

Alonso R., et al., 2007, in Transiting Extrapolar Planets Workshop. p. 13

Altmann M., Roeser S., Demleitner M., Bastian U., Schilbach E., 2017, VizieR Online Data Catalog, p. I/339

Ammons S. M., Robinson S. E., Strader J., Laughlin G., Fischer D., Wolf A., 2010, VizieR Online Data Catalog, p. V/136

Anderson D. R., et al., 2011, ApJ, 726, L19

Anderson D. R., et al., 2017, A&A, 604, A110

Anglada-Escudé G., Butler R. P., 2012, ApJS, 200, 15

Antognini J. M. O., 2015, MNRAS, 452, 3610

Antonini F., Toonen S., Hamers A. S., 2017, ApJ, 841, 77

Aristotle 2009, On the heavens. BiblioBazaar, LLC

Armstrong D., et al., 2013, MNRAS, 434, 3047

Astropy Collaboration et al., 2018, AJ, 156, 123

Auvergne M., et al., 2009, A&A, 506, 411

Bacchini C., Fraternali F., Iorio G., Pezzulli G., Marasco A., Nipoti C., 2020, A&A, 641, A70

Bailer-Jones C. A. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, AJ, 156, 58

Bakos G., Noyes R. W., Kovács G., Stanek K. Z., Sasselov D. D., Domsa I., 2004, PASP, 116, 266

Bakos G., et al., 2008, Proceedings of the International Astronomical Union, 4, 354

Bakos G. Á., et al., 2010, ApJ, 710, 1724

Bakos G. Á., et al., 2015, ApJ, 813, 111

Baraffe I., Chabrier G., Barman T., 2008, A&A, 482, 315

Baranne A., et al., 1996, A&AS, 119, 373

Barbato D., Bonomo A. S., Sozzetti A., Morbidelli R., 2018, arXiv e-prints, p. arXiv:1811.08249

Barnes S. I., Gibson S., Nield K., Cochrane D., 2012, in Proc. SPIE. p. 844688, doi:10.1117/12.926527

Barragán O., Gandolfi D., Antoniciello G., 2019, MNRAS, 482, 1017

Barragán O., et al., 2022a, MNRAS,

Barragán O., Aigrain S., Rajpaul V. M., Zicher N., 2022b, MNRAS, 509, 866

Bastian N., Covey K. R., Meyer M. R., 2010, ARA&A, 48, 339

Batalha N. M., et al., 2011, ApJ, 729, 27

Batalha N. E., et al., 2017, PASP, 129, 064501

Batalha N. E., Lewis T., Fortney J. J., Batalha N. M., Kempton E., Lewis N. K., Line M. R., 2019, ApJ, 885, L25

Bate M. R., 1998, ApJ, 508, L95

Bate M. R., Bonnell I. A., Bromm V., 2002, MNRAS, 336, 705

Battley M. P., Pollacco D., Armstrong D. J., 2020, MNRAS, 496, 1197

Bayes T., 1763, Philosophical transactions of the Royal Society of London, pp 370–418

Bayliss D., et al., 2015, AJ, 150, 49

Benedict G. F., et al., 2002, ApJ, 581, L115

Berger T., Huber D., Gaidos E., van Saders J., 2018, in 20th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun. Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. p. 72, doi:10.5281/zenodo.1488237

Berger T. A., Huber D., Gaidos E., van Saders J. L., Weiss L. M., 2020, AJ, 160, 108

Betancourt M., 2017, arXiv e-prints, p. arXiv:1701.02434

Bitsch B., Izidoro A., Johansen A., Raymond S. N., Morbidelli A., Lambrechts M., Jacobson S. A., 2019, A&A, 623, A88

Blackman R. T., et al., 2020, AJ, 159, 238

Boccaletti A., 2011, in Beaulieu J. P., Dieters S., Tinetti G., eds, Astronomical Society of the Pacific Conference Series Vol. 450, Molecules in the Atmospheres of Extrasolar Planets. p. 163

Bodensteiner J., et al., 2021, A&A, 652, A70

Boisse I., Bouchy F., Hébrard G., Bonfils X., Santos N., Vauclair S., 2011, A&A, 528, A4

Borkovits T., Érdi B., Forgács-Dajka E., Kovács T., 2003, A&A, 398, 1091

Borkovits T., et al., 2013, MNRAS, 428, 1656

Borkovits T., Rappaport S., Hajdu T., Sztakovics J., 2015, MNRAS, 448, 946

Borkovits T., Hajdu T., Sztakovics J., Rappaport S., Levine A., Bíró I. B., Klagyivik P., 2016, MNRAS, 455, 4136

Borkovits T., Rappaport S. A., Hajdu T., Maxted P. F. L., Pál A., Forgács-Dajka E., Klagyivik P., Mitnyan T., 2020a, MNRAS, 493, 5005

Borkovits T., et al., 2020b, MNRAS, 496, 4624

Borucki W. J., Summers A. L., 1984, Icarus, 58, 121

Borucki W., et al., 2009, in Pont F., Sasselov D., Holman M. J., eds, IAU Symposium Vol. 253, Transiting Planets. pp 289–299, doi:10.1017/S1743921308026513

Borucki W. J., et al., 2010, Science, 327, 977

Borucki W. J., et al., 2011, ApJ, 736, 19

Borucki W. J., et al., 2012, ApJ, 745, 120

Boss A. P., 1998, ApJ, 501, L77

Bouchy F., et al., 2009, A&A, 505, 853

Bourges L., Mella G., Lafrasse S., Duvert G., Chelli A., Le Bouquin J. B., Delfosse X., Chesneau O., 2017, VizieR Online Data Catalog, p. II/346

Bowler B. P., 2016, PASP, 128, 102001

Brahm R., et al., 2019, AJ, 158, 45

Brakensiek J., Ragozzine D., 2016, ApJ, 821, 47

Brasseur C. E., Phillip C., Fleming S. W., Mullally S. E., White R. L., 2019, Astrocut: Tools for creating cutouts of TESS images (ascl:1905.007)

Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127

Brodskaya E. S., 1953, Izvestiya Ordena Trudovogo Krasnogo Znameni Krymskoj Astrofizicheskoj Observatorii, 10, 104

Broeg C., et al., 2013, in EPJ Web of Conferences. p. 03005

Brown T. M., et al., 2013, PASP, 125, 1031

Bruntt H., et al., 2010, MNRAS, 405, 1907

Bryson S., Coughlin J., Batalha N. M., Berger T., Huber D., Burke C., Dotson J., Mullally S. E., 2020, AJ, 159, 279

Bryson S., Flynn K., Hanna H., Green T., Coughlin J. L., Kunimoto M., 2021a, PASP, 133, 104401

Bryson S., et al., 2021b, AJ, 161, 36

Buchhave L. A., et al., 2012, Nature, 486, 375

Buchhave L. A., et al., 2014, Nature, 509, 593

Buchner J., et al., 2014, A&A, 564, A125

Burke C. J., et al., 2015, ApJ, 809, 8

Burke C. J., et al., 2020, TESS-Point: High precision TESS pointing tool, Astrophysics

Source Code Library (ascl:2003.001)

Butler R. P., Marcy G. W., Fischer D. A., Brown T. M., Contos A. R., Korzennik S. G., Nisenson P., Noyes R. W., 1999, ApJ, 526, 916

Cañas C. I., et al., 2020, AJ, 160, 147

Cagaš P., Pejcha O., 2012, A&A, 544, L3

Caldwell D. A., et al., 2020, RNAAS, 4, 201

Cameron A. C., 2016, in , Methods of Detecting Exoplanets. Springer, pp 89–131

Campbell B., 1983, PASP, 95, 577

Campbell B., Walker G. A. H., Yang S., 1988, ApJ, 331, 902

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Carleo I., et al., 2020, AJ, 160, 114

Catanzarite J., Shao M., 2011, ApJ, 738, 151

Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, ApJ, 529, L45

Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, ApJ, 686, 580

Chen J., Kipping D., 2017, ApJ, 834, 17

Chen H., Rogers L. A., 2016, ApJ, 831, 180

Childress M. J., Vogt F. P. A., Nielsen J., Sharp R. G., 2014, Ap&SS, 349, 617

Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, ApJ, 823, 102

Christiansen J. L., et al., 2012, PASP, 124, 1279

Christiansen J. L., et al., 2015, ApJ, 810, 95

Christiansen J. L., et al., 2018, AJ, 155, 57

Christiansen J. L., et al., 2020, AJ, 160, 159

Christiansen J. L., et al., 2022, arXiv e-prints, p. arXiv:2203.02087

Cincunegui C., Díaz R. F., Mauas P. J. D., 2007, A&A, 469, 309

Claret A., 2017, A&A, 600, A30

Cochran W. D., Hatzes A. P., Paulson D. B., 2002, AJ, 124, 565

Collier Cameron A., et al., 2021, MNRAS, 505, 1699

Connes P., 2020, History of the Plurality of Worlds. Springer

Conroy K. E., Prša A., Stassun K. G., Orosz J. A., Fabrycky D. C., Welsh W. F., 2014, AJ, 147, 45

Copernicus N., 1965, Norimbergae, Apud J. Petreium, 1543; Bruxelles, Culture et Civilisation, 1966

Cosentino R., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. p. 84461V, doi:10.1117/12.925738

Cosentino R., et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. p. 91478C, doi:10.1117/12.2055813

Costes J. C., et al., 2021, MNRAS, 505, 830

Crossfield I. J. M., et al., 2016, ApJS, 226, 7

Crowe M. J., 1988, The Extraterrestrial Life Debate 1750-1900: the idea of a plurality of worlds from Kant to Lowell. CUP Archive

Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., 2008, PASP, 120, 531

Cutri R. M., et al. 2013, VizieR Online Data Catalog, p. II/328

Cutri R. M., et al., 2003, VizieR Online Data Catalog, p. II/246

Dalal S., Hebrard G., Lecavelier Des Etangs A., Petit A., Bourrier V., Laskar J., Konig P.-C., Correia A. C. M., 2019, in AAS/Division for Extreme Solar Systems Abstracts. p. 202.04

Dalba P. A., et al., 2022, AJ, 163, 61

David T. J., et al., 2016, Nature, 534, 658

David T. J., et al., 2019, AJ, 158, 79

David T. J., et al., 2021, AJ, 161, 265

Deck K. M., Agol E., Holman M. J., Nesvorný D., 2014, ApJ, 787, 132

Deeg H. J., et al., 2009, A&A, 506, 343

Demangeon O. D. S., et al., 2021, A&A, 653, A78

Díaz R. F., Almenara J. M., Santerne A., Moutou C., Lethuillier A., Deleuil M., 2014, MNRAS, 441, 983

Díaz R. F., et al., 2016, A&A, 585, A134

Díaz M. R., et al., 2020, MNRAS, 493, 973

Dib S., Schmeja S., Hony S., 2017, MNRAS, 464, 1738

Dick S. J., 2020, in , Space, Time, and Aliens. Springer, pp 5–26

Dietrich J., Apai D., 2020, AJ, 160, 107

Donati J. F., Semel M., Carter B. D., Rees D. E., Collier Cameron A., 1997, MNRAS, 291, 658

Donati J. F., et al., 2016, Nature, 534, 662

Dopita M., Hart J., McGregor P., Oates P., Bloxham G., Jones D., 2007, Ap&SS, 310, 255

Doyle L. R., et al., 2011, Science, 333, 1602

Doyle A. P., Davies G. R., Smalley B., Chaplin W. J., Elsworth Y., 2014, MNRAS, 444, 3592

Dragomir D., et al., 2019, ApJ, 875, L7

Droege T. F., Richmond M. W., Sallman M., 2007, VizieR Online Data Catalog, p. II/271A

Dumusque X., 2018, A&A, 620, A47

Dunér D., 2016, Zygon®, 51, 450

Eastman J., 2017, EXOFASTv2: Generalized publication-quality exoplanet modeling code (ascl:1710.003)

Eggleton P. P., 1983, ApJ, 268, 368

Eggleton P. P., Kisseleva-Eggleton L., 2006, Ap&SS, 304, 75

Eisner N. L., Pope B. J. S., Aigrain S., Barragán O., White T. R., Huang C. X., Lintott C., Volkov A., 2019, RNAAS, 3, 145

Eisner N., Lintott C., Aigrain S., 2020a, The Journal of Open Source Software, 5, 2101

Eisner N. L., et al., 2020b, MNRAS, 494, 750

Eisner N. L., et al., 2021a, MNRAS, 501, 4669

Eisner N. L., et al., 2021b, MNRAS, 505, 1827

Eisner N. L., et al., 2022, MNRAS, 511, 4710

Eldridge J. J., Xiao L., Stanway E. R., Rodrigues N., Guo N. Y., 2018, Publ. Astron. Soc. Australia, 35, e049

Espinoza N., Kossakowski D., Brahm R., 2019, MNRAS, 490, 2262

Esposito M., et al., 2019, A&A, 623, A165

Ester M., Kriegel H.-P., Sander J., Xu X., et al., 1996, in Kdd. pp 226–231

Fabricius C., Høg E., Makarov V. V., Mason B. D., Wycoff G. L., Urban S. E., 2002, A&A, 384, 180

Fabrycky D. C., et al., 2014, ApJ, 790, 146

Fausnaugh M., Huang X., Glidden A., Guerrero N., TESS Science Office 2018, in American Astronomical Society Meeting Abstracts #231. p. 439.09

Feinstein A. D., et al., 2019, PASP, 131, 094502

Feliz D. L., Plavchan P., Bianco S. N., Jimenez M., Collins K. I., Villarreal Alvarado B., Stassun K. G., 2021, AJ, 161, 247

Figueira P., et al., 2010, A&A, 513, L8

Fischer D. A., et al., 2012, MNRAS, 419, 2900

Fischer D. A., et al., 2016, PASP, 128, 066001

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Foreman-Mackey D., Hogg D. W., Morton T. D., 2014, ApJ, 795, 64

Foreman-Mackey D., Morton T. D., Hogg D. W., Agol E., Schölkopf B., 2016, AJ, 152, 206

Foreman-Mackey D., et al., 2021a, exoplanet-dev/exoplanet v0.4.5, doi:10.5281/zenodo.1998447, `https://doi.org/10.5281/zenodo.1998447`

Foreman-Mackey D., et al., 2021b, The Journal of Open Source Software, 6, 3285

Fortenbach C. D., Dressing C. D., 2020, PASP, 132, 054501

Frank J., King A., Raine D. J., 2002, Accretion Power in Astrophysics: Third Edition. Accretion Power in Astrophysics, Cambridge University Press

Fressin F., et al., 2012, Nature, 482, 195

Fressin F., et al., 2013, ApJ, 766, 81

Fridlund M., et al., 2010, Astrobiology, 10, 5

Fridlund M., et al., 2017, A&A, 604, A16

Fridlund M., et al., 2020, MNRAS, 498, 4503

Frost A. J., Oudmaijer R. D., Lumsden S. L., de Wit W. J., 2021, ApJ, 920, 48

Fu G., et al., 2021, AJ, 162, 271

Fulton B. J., et al., 2017, AJ, 154, 109

Gaia Collaboration et al., 2018, A&A, 616, A1

Gaia Collaboration et al., 2021, A&A, 649, A1

Gandolfi D., et al., 2018, A&A, 619, L10

Gandolfi D., et al., 2019, ApJ, 876, L24

Gao P., Zhang X., 2020, ApJ, 890, 93

Gatti H., 2010, in , Essays on Giordano Bruno. Princeton University Press

Gaudi B. S., 2012, ARA&A, 50, 411

Gelman A., Rubin D. B., 1992, Statistical Science, 7, 457

Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2004, CRC Texts in Statistical Science

Georgieva I. Y., et al., 2021, MNRAS, 505, 4684

Gilbert E. A., et al., 2020, AJ, 160, 116

Gill S., et al., 2020, MNRAS, 491, 1548

Girardi L., Groenewegen M. A. T., Hatziminaoglou E., da Costa L., 2005, A&A, 436, 895

Glanz H., Perets H. B., 2021, MNRAS, 500, 1921

Goldreich P., Soter S., 1966, Icarus, 5, 375

Götberg Y., Korol V., Lamberts A., Kupfer T., Breivik K., Ludwig B., Drout M. R., 2020, ApJ, 904, 56

Grunblatt S. K., et al., 2017, AJ, 154, 254

Grunblatt S. K., et al., 2018, ApJ, 861, L5

Grunblatt S. K., Huber D., Gaidos E., Hon M., Zinn J. C., Stello D., 2019, AJ, 158, 227

Grunblatt S. K., et al., 2022, arXiv e-prints, p. arXiv:2201.04140

Guerrero N. M., et al., 2021, ApJS, 254, 39

Guillot T., Gautier D., 2015, in Schubert G., ed., , Treatise on Geophysics. pp 529–557, doi:10.1016/B978-0-444-53802-4.00176-7

Guillot T., Burrows A., Hubbard W. B., Lunine J. I., Saumon D., 1996, ApJ, 459, L35

Günther M. N., Daylan T., 2021, ApJS, 254, 13

Hajdu T., Borkovits T., Forgács-Dajka E., Sztakovics J., Marschalkó G., Benkő J. M., Klagyivik P., Sallai M. J., 2017, MNRAS, 471, 1230

Hamers A. S., Glanz H., Neunteufel P., 2021a, arXiv e-prints, p. arXiv:2110.00024

Hamers A. S., Rantala A., Neunteufel P., Preece H., Vynatheya P., 2021b, MNRAS, 502, 4479

Harmanec P., 1988, Bulletin of the Astronomical Institutes of Czechoslovakia, 39, 329

Hartkopf W. I., et al., 2000, AJ, 119, 3084

Hartkopf W. I., Mason B. D., Rafferty T. J., 2008, AJ, 135, 1334

Hatzes A. P., Rauer H., 2015, ApJ, 810, L25

Hebb L., et al., 2009, ApJ, 693, 1920

Hellier C., et al., 2017, MNRAS, 465, 3693

Hełminiak K. G., Ukita N., Kambe E., Kozłowski S. K., Pawłaszek R., Maehara H., Baranec C., Konacki M., 2017, A&A, 602, A30

Henden A. A., Templeton M., Terrell D., Smith T. C., Levine S., Welch D., 2016, VizieR Online Data Catalog, p. II/336

Hennebelle P., Commerçon B., 2014, in The Labyrinth of Star Formation. p. 365, doi:10.1007/978-3-319-03041-8_72

Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, ApJ, 529, L41

Henry T. J., et al., 2018, AJ, 155, 265

Hjorth M., Albrecht S., Hirano T., Winn J. N., Dawson R. I., Zanazzi J. J., Knudstrup E., Sato B., 2021, Proceedings of the National Academy of Science, 118, 2017418118

Hobbs G., Lorimer D. R., Lyne A. G., Kramer M., 2005, MNRAS, 360, 974

Hoffman M. D., Gelman A., 2011, arXiv e-prints, p. arXiv:1111.4246

Høg E., et al., 2000, A&A, 355, L27

Holman M. J., Murray N. W., 2005, Science, 307, 1288

Holman M. J., et al., 2010, Science, 330, 51

Howard A. W., et al., 2010, ApJ, 721, 1467

Howell S. B., Everett M. E., Sherry W., Horch E., Ciardi D. R., 2011, AJ, 142, 19

Howell S. B., et al., 2014, PASP, 126, 398

Huang C. X., et al., 2020a, RNAAS, 4, 204

Huang C. X., et al., 2020b, RNAAS, 4, 206

Huber D., et al., 2013, ApJ, 767, 127

Huber D., et al., 2016, ApJS, 224, 2

Huber D., et al., 2019, AJ, 157, 245

Iffrig O., Hennebelle P., 2015, A&A, 576, A95

Igoshev A. P., 2020, MNRAS, 494, 3663

Ito T., Ohtsuka K., 2019, Monographs on Environment, Earth and Planets, 7, 1

Ivanova N., et al., 2013, A&ARv, 21, 59

Jenkins J. M., et al., 2010, in Radziwill N. M., Bridger A., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 7740, Software and Cyberinfrastructure for Astronomy. p. 77400D, doi:10.1117/12.856764

Jenkins J. M., et al., 2016, in Proc. SPIE. p. 99133E, doi:10.1117/12.2233418

Jenkins J. M., et al., 2018, in American Astronomical Society Meeting Abstracts #232. p. 120.06

Johns-Krull C. M., et al., 2016, ApJ, 826, 206

Johnson J. A., et al., 2007, ApJ, 665, 785

Johnson J. A., et al., 2011, ApJS, 197, 26

Johnston K. G., et al., 2020, A&A, 634, L11

Jones H. R. A., Butler R. P., Tinney C. G., Marcy G. W., Carter B. D., Penny A. J., McCarthy C., Bailey J., 2006, MNRAS, 369, 249

Jones M. I., Jenkins J. S., Bluhm P., Rojo P., Melo C. H. F., 2014, A&A, 566, A113

Jurgenson C., Fischer D., McCracken T., Sawyer D., Szymkowiak A., Davis A., Muller G., Santoro F., 2016, in Evans C. J., Simard L., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. p. 99086T (arXiv:1606.04413), doi:10.1117/12.2233002

Kallinger T., et al., 2014, A&A, 570, A41

Kane M., Ragozzine D., Flowers X., Holczer T., Mazeh T., Relles H. M., 2019, AJ, 157, 171

Kanodia S., Wright J., 2018, RNAAS, 2, 4

Kant I., 1755, Allgemeine Naturgeschichte und Theorie des Himmels, nach Newtonischen Grundsätzen abgehandelt

Kasting J. F., Whitmire D. P., Reynolds R. T., 1993, Icarus, 101, 108

Katz B., Dong S., Malhotra R., 2011, Physical Review Letters, 107, 181101

Kaye L., et al., 2022, MNRAS, 510, 5464

Kempton E. M. R., et al., 2018, PASP, 130, 114401

Kiminki D. C., Kobulnicky H. A., 2012, ApJ, 751, 4

Kippenhahn R., Weigert A., Weiss A., 2012, Stellar Structure and Evolution, doi:10.1007/978-3-642-30304-3.

Kipping D. M., 2010, MNRAS, 408, 1758

Kipping D. M., 2013, MNRAS, 435, 2152

Kipping D. M., Schmitt A. R., Huang X., Torres G., Nesvorný D., Buchhave L. A., Hartman J., Bakos G. Á., 2015, AJ, 813, 14

Kochanek C. S., et al., 2017, PASP, 129, 104502

Komacek T. D., Thorngren D. P., Lopez E. D., Ginzburg S., 2020, ApJ, 893, 36

Koo J.-R., Lee J. W., Lee B.-C., Kim S.-L., Lee C.-U., Hong K., Lee D.-J., Rey S.-C., 2014, AJ, 147, 104

Kopparapu R. K., et al., 2013, ApJ, 765, 131

Kostov V. B., et al., 2021, ApJ, 917, 93

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

Kozai Y., 1962, AJ, 67, 591

Kreidberg L., 2015, PASP, 127, 1161

Krumholz M. R., Bonnell I. A., 2007, arXiv e-prints, p. arXiv:0712.0828

Kunimoto M., et al., 2021, RNAAS, 5, 234

Kupka F., Piskunov N., Ryabchikova T. A., Stempels H. C., Weiss W. W., 1999, A&A Supplement Series, 138, 119

Kurucz R.-L., 1993, Kurucz CD-Rom, 13

Kurucz R. L., 2013, ATLAS12 (ascl:1303.024)

Lacedelli G., et al., 2021, MNRAS, 501, 4148

Lakićević M., et al., 2015, ApJ, 799, 50

Lambrechts M., Johansen A., Morbidelli A., 2014, A&A, 572, A35

Langer N., 2012, ARA&A, 50, 107

Lanza A. F., Mathis S., 2016, Celestial Mechanics and Dynamical Astronomy, 126, 249

Lasker B., Lattanzi M. G., McLean B. J., et al. 2007, VizieR Online Data Catalog, p. I/305

Latham D. W., Stefanik R. P., Mazeh T., Torres G., 1989, in Bulletin of the American Astronomical Society. p. 1224

Latham D. W., et al., 2011, ApJ, 732, L24

Laur J., Kolka I., Eenmäe T., Tuvikene T., Leedjärv L., 2017, A&A, 598, A108

Lecavelier des Etangs A., Lissauer J. J., 2022, New Astron. Rev., 94, 101641

Lee E. J., Chiang E., 2016, ApJ, 817, 90

Léger A., et al., 2009, A&A, 506, 287

Leigh N. W. C., Toonen S., Portegies Zwart S. F., Perna R., 2020, MNRAS, 496, 1819

Leleu A., et al., 2021, arXiv e-prints, p. arXiv:2101.09260

Levesque E. M., Massey P., Zytkow A. N., Morrell N., 2014, MNRAS, 443, L94

Levine S. E., et al., 2012, in Ground-based and Airborne Telescopes IV. p. 844419

Li G., Naoz S., Holman M., Loeb A., 2014, ApJ, 791, 86

Li M. C. A., et al., 2018, MNRAS, 480, 4557

Li J., Tenenbaum P., Twicken J. D., Burke C. J., Jenkins J. M., Quintana E. V., Rowe J. F., Seader S. E., 2019, PASP, 131, 024506

Lidov M. L., 1962, Planetary and Space Science, 9, 719

Lienhard F., et al., 2020, MNRAS, 497, 3790

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

Line M. R., et al., 2013, ApJ, 775, 137

Lintott C. J., et al., 2008, MNRAS, 389, 1179

Lintott C., et al., 2011, VizieR Online Data Catalog, p. J/MNRAS/410/166

Lissauer J. J., et al., 2011, ApJS, 197, 8

Lissauer J. J., et al., 2012, ApJ, 750, 112

Lissauer J. J., et al., 2014, ApJ, 784, 44

Lithwick Y., Naoz S., 2011, ApJ, 742, 94

Livingston J. H., et al., 2018, AJ, 156, 277

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

Lopez E. D., Fortney J. J., 2014, ApJ, 792, 1

Lopez E. D., Fortney J. J., 2016, ApJ, 818, 4

Louvet F., 2018, in Di Matteo P., Billebaud F., Herpin F., Lagarde N., Marquette J. B., Robin A., Venot O., eds, SF2A-2018: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. p. Di

Luger R., et al., 2017, Nature Astronomy, 1, 0129

Luhn J. K., Bastien F. A., Wright J. T., Johnson J. A., Howard A. W., Isaacson H., 2019, AJ, 157, 149

Maíz Apellániz J., et al., 2019, A&A, 626, A20

Mamajek E. E., Hillenbrand L. A., 2008, ApJ, 687, 1264

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

Mandell A. M., Haynes K., Sinukoff E., Madhusudhan N., Burrows A., Deming D., 2013, ApJ, 779, 128

Mann A. W., et al., 2016, AJ, 152, 61

Mann A. W., et al., 2020, AJ, 160, 179

Mansfeld J., 1980, Phronesis, pp 1–4

Marcy G. W., 1983, in Bulletin of the American Astronomical Society. p. 947

Marcy G. W., Butler R. P., 1992, PASP, 104, 270

Mardling R., Aarseth S., 1999, in Steves B. A., Roy A. E., eds, NATO Advanced Science Institutes (ASI) Series C Vol. 522, NATO Advanced Science Institutes (ASI) Series C. p. 385

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

Marsh T. R., Armstrong D. J., Carter P. J., 2014, MNRAS, 445, 309

Martins F., Schaerer D., Hillier D. J., 2005, A&A, 436, 1049

Matson R. A., Howell S. B., Ciardi D. R., 2019, AJ, 157, 211

Mayor M., Maurice E., 1985, in Philip A. G. D., Latham D. W., eds, Stellar Radial

Velocities. pp 299–310

Mayor M., Queloz D., 1995, Nature, 378, 355

Mayor M., et al., 2003, The Messenger, 114, 20

McCully C., Volgenau N. H., Harbeck D.-R., Lister T. A., Saunders E. S., Turner M. L., Siiverd R. J., Bowman M., 2018, in Guzman J. C., Ibsen J., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 10707, Software and Cyberinfrastructure for Astronomy V. p. 107070K (arXiv:1811.04163), doi:10.1117/12.2314340

McLaughlin D. B., 1924, ApJ, 60, 22

McQuillan A., Aigrain S., Roberts S., 2011, International Astronomical Union. Proceedings of the International Astronomical Union, 7, 364

Megeath S. T., Wilson T. L., Corbin M. R., 2005, ApJ, 622, L141

Meunier N., Desort M., Lagrange A. M., 2010, A&A, 512, A39

Millholland S., 2019, ApJ, 886, 72

Mills S. M., Fabrycky D. C., Migaszewski C., Ford E. B., Petigura E., Isaacson H., 2016, Nature, 533, 509

Miralda-Escudé J., 2002, ApJ, 564, 1019

Mitnyan T., Borkovits T., Rappaport S. A., Pál A., Maxted P. F. L., 2020, MNRAS, 498, 6034

Modjaz M., Gutiérrez C. P., Arcavi I., 2019, Nature Astronomy, 3, 717

Moe M., Di Stefano R., 2017, ApJS, 230, 15

Montalto M., et al., 2022, MNRAS, 509, 2908

Montet B. T., et al., 2015, ApJ, 809, 25

Mordasini C., 2013, A&A, 558, A113

Mordasini C., Marleau G. D., Mollière P., 2017, A&A, 608, A72

Morishima R., Schmidt M. W., Stadel J., Moore B., 2008, ApJ, 685, 1247

Mortier A., et al., 2020, MNRAS, 499, 5004

Morton T. D., 2012, ApJ, 761, 6

Morton T. D., 2015, VESPA: False positive probabilities calculator, Astrophysics Source Code Library (ascl:1503.011)

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

Munari U., et al., 2014, AJ, 148, 81

Naoz S., 2016, ARA&A, 54, 441

Nava C., López-Morales M., Haywood R. D., Giles H. A. C., 2020, AJ, 159, 23

Nesvorný D., Kipping D. M., Buchhave L. A., Bakos G. Á., Hartman J., Schmitt A. R., 2012, Science, 336, 1133

Nielsen L. D., et al., 2019a, MNRAS, 489, 2478

Nielsen L. D., et al., 2019b, A&A, 623, A100

Noyes R. W., Hartmann L. W., Baliunas S. L., Duncan D. K., Vaughan A. H., 1984, ApJ, 279, 763

Oliva G. A., Kuiper R., 2020, A&A, 644, A41

Orosz J. A., et al., 2012, ApJ, 758, 87

Osborn H. P., et al., 2016, MNRAS, 457, 2273

Otegi J. F., Bouchy F., Helled R., 2020, A&A, 634, A43

Owen J. E., Lai D., 2018, MNRAS, 479, 5012

Owen J. E., Wu Y., 2013, ApJ, 775, 105

Owen J. E., Wu Y., 2016, ApJ, 817, 107

Parviainen H., 2015, MNRAS, 450, 3233

Pearson K. A., Palafox L., Griffith C. A., 2018, MNRAS, 474, 478

Pecaut M. J., Mamajek E. E., 2013, ApJ Supplement Series, 208, 9

Pedbost M. F., Pomalgu T., Lintott C., Eisner N., Nicholson B., 2020, arXiv e-prints, p. arXiv:2003.13722

Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825

Pepe F., et al., 2000, in Iye M., Moorwood A. F., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4008, Optical and IR Telescope Instrumentation and Detectors. pp 582–592, doi:10.1117/12.395516

Pepe F., Mayor M., Galland F., Naef D., Queloz D., Santos N. C., Udry S., Burnet M., 2002, A&A, 388, 632

Pepe F., et al., 2021, A&A, 645, A96

Pepper J., et al., 2017, AJ, 153, 177

Perruchot S., et al., 2008, in Ground-based and Airborne Instrumentation for Astronomy II. p. 70140J, doi:10.1117/12.787379

Perryman M., 2018, The Exoplanet Handbook

Persson C. M., et al., 2018, A&A, 618, A33

Petersburg R. R., et al., 2020, AJ, 159, 187

Petigura E. A., Howard A. W., Marcy G. W., 2013, Proceedings of the National Academy of Science, 110, 19273

Pierrehumbert R., Gaidos E., 2011, ApJ, 734, L13

Piskunov N., Valenti J. A., 2017, A&A, 597, A16

Podsiadlowski P., Cannon R. C., Rees M. J., 1995, MNRAS, 274, 485

Pollacco D. L., et al., 2006, PASP, 118, 1407

Pollack J. B., Hubickyj O., Bodenheimer P., Lissauer J. J., Podolak M., Greenzweig Y., 1996, Icarus, 124, 62

Portegies Zwart S., Leigh N. W. C., 2019, ApJ, 876, L33

Powell B. P., et al., 2021, AJ, 161, 162

Preus A., 2001, Essays in Ancient Greek Philosophy VI: Before Plato. Vol. 6, SUNY Press

Price D. J., et al., 2018, MNRAS, 477, 1270

Priestley F. D., Chawner H., Matsuura M., De Looze I., Barlow M. J., Gomez H. L., 2021, MNRAS, 500, 2543

Prša A., et al., 2022, ApJS, 258, 16

Queloz D., et al., 2000, A&A, 354, 99

Queloz D., et al., 2001, A&A, 379, 279

Quinn S. N., et al., 2019, AJ, 158, 177

Radick R. R., Lockwood G. W., Skiff B. A., Baliunas S. L., 1998, ApJS, 118, 239

Raghavan D., et al., 2010, ApJS, 190, 1

Rainot A., et al., 2020, A&A, 640, A15

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

Rajpaul V. M., et al., 2021, MNRAS, 507, 1847

Rappaport S., et al., 2017, MNRAS, 467, 2160

Rappaport S., et al., 2019, MNRAS, 488, 2455

Rappaport S. A., et al., 2022, MNRAS,

Raskin G., et al., 2011, A&A, 526, A69

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Reed B. C., 2003, AJ, 125, 2531

Renzo M., et al., 2019, A&A, 624, A66

Rice M., et al., 2021, AJ, 162, 182

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

Rodriguez J. E., et al., 2019, AJ, 157, 191

Rosenblatt M., 1956, Proceedings of the National Academy of Sciences of the United States of America, 42, 43

Rossiter R. A., 1924, ApJ, 60, 15

Rowden P., et al., 2020, AJ, 160, 76

Rowe J. F., et al., 2014, ApJ, 784, 45

Rowe J. F., et al., 2015, ApJS, 217, 16

Ryabchikova T., Piskunov N., Kurucz R. L., Stempels H. C., Heiter U., Pakhomov Y., Barklem P. S., 2015, Phys. Scr., 90, 054005

Saar S. H., Butler R. P., Marcy G. W., 1998, ApJ, 498, L153

Sagan C., 1994, in , Pale Blue Dot: A Vision of the Human Future in Space. Random House

Salpeter E. E., 1955, ApJ, 121, 161

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Computer Science, 2, e55

Sana H., et al., 2012, Science, 337, 444

Sana H., et al., 2013, A&A, 550, A107

Sana H., et al., 2014, ApJS, 215, 15

Sanchis-Ojeda R., Winn J. N., Holman M. J., Carter J. A., Osip D. J., Fuentes C. I., 2011, ApJ, 733, 127

Santerne A., et al., 2015, MNRAS, 451, 2337

Santos N. C., Mayor M., Naef D., Pepe F., Queloz D., Udry S., Blecha A., 2000, A&A, 361, 265

Santos N. C., Gomes da Silva J., Lovis C., Melo C., 2010, A&A, 511, A54

Scargle J. D., 1982, ApJ, 263, 835

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schmitt J. R., et al., 2014, AJ, 148, 28

Schneider J., 2000, in From Giant Planets to Cool Stars. p. 284

Schneider F. R. N., Ohlmann S. T., Podsiadlowski P., Röpke F. K., Balbus S. A., Pakmor R., Springel V., 2019, Nature, 574, 211

Schwamb M. E., et al., 2012, ApJ, 754, 129

Schwamb M. E., et al., 2013, ApJ, 768, 127

Seager S., 2008, Space Sci. Rev., 135, 345

Seager S., Kuchner M., Hier-Majumder C. A., Militzer B., 2007, ApJ, 669, 1279

Sha L., et al., 2021, AJ, 161, 82

Shappee B. J., et al., 2014, ApJ, 788, 48

Shulyak D., Tsymbal V., Ryabchikova T., Stütz C., Weiss W. W., 2004, A&A, 428, 993

Silaj J., Jones C. E., Sigut T. A. A., Tycner C., 2014, ApJ, 795, 82

Silburt A., Gaidos E., Wu Y., 2015, ApJ, 799, 180

Simón-Díaz S., Stasińska G., 2008, MNRAS, 389, 1009

Smith N., 2006, MNRAS, 367, 763

Smith J. C., et al., 2012, PASP, 124, 1000

Sneden C. A., 1973, PhD thesis, The University of Texas at Austin.

Sota A., Maíz Apellániz J., Walborn N. R., Alfaro E. J., Barbá R. H., Morrell N. I., Gamen R. C., Arias J. I., 2011, ApJS, 193, 24

Soter S., 2006, AJ, 132, 2513

Sousa S. G., 2014, ARES + MOOG: A Practical Overview of an Equivalent Width (EW) Method to Derive Stellar Parameters. pp 297–310, doi:10.1007/978-3-319-06956-2_26

Sousa S. G., Santos N. C., Adibekyan V., Delgado-Mena E., Israelian G., 2015, A&A, 577, A67

Sousa S. G., et al., 2019, MNRAS, 485, 3981

Southworth J., 2011, MNRAS, 417, 2166

Spiegel D. S., Burrows A., 2012, ApJ, 745, 174

Spiers H., Swanson A., Fortson L., Simmons B., Trouille L., Blickhan S., Lintott C., 2019, Journal of Science Communication, 18

Sriram K., Malu S., Choi C. S., Vivekananda Rao P., 2018, AJ, 155, 172

Stassun K. G., Torres G., 2016, AJ, 152, 180

Stassun K. G., Torres G., 2018, ApJ, 862, 61

Stassun K. G., Collins K. A., Gaudi B. S., 2017, AJ, 153, 136

Stassun K. G., Corsaro E., Pepper J. A., Gaudi B. S., 2018, AJ, 155, 22

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

Steffen J. H., et al., 2013, MNRAS, 428, 1077

Stephan A. P., et al., 2019, ApJ, 878, 58

Storch N. I., Lai D., Anderson K. R., 2017, MNRAS, 465, 3927

Struve O., 1952, The Observatory, 72, 199

Stumpe M. C., et al., 2012, PASP, 124, 985

Swedenborg E., 1734, Opera Philosophica et Mineralia. Dresden & Leipzig: F. Hekel

Tabernero H. M., Dorda R., Negueruela I., Marfil E., 2021, A&A, 646, A98

Tan J. C., Beltrán M. T., Caselli P., Fontani F., Fuente A., Krumholz M. R., McKee C. F., Stolte A., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. p. 149 (arXiv:1402.0919), doi:10.2458/azu_uapress_9780816531240-ch007

Tauris T., 2015, arXiv e-prints, p. arXiv:1510.07875

Teyssandier J., Naoz S., Lizarraga I., Rasio F. A., 2013, ApJ, 779, 166

Thorne K. S., Zytkow A. N., 1975, ApJ, 199, L19

Tinetti G., et al., 2016, in MacEwen H. A., Fazio G. G., Lystrup M., Batalha N., Siegler N., Tong E. C., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. p. 99041X, doi:10.1117/12.2232370

Tkachenko A., 2015, A&A, 581, A129

Tkachenko A., Van Reeth T., Tsymbal V., Aerts C., Kochukhov O., Debosscher J., 2013, A&A, 560, A37

Tokovinin A., 2018, PASP, 130, 035002

Toledo-Padrón B., et al., 2020, A&A, 641, A92

Toonen S., Hamers A., Portegies Zwart S., 2016, Computational Astrophysics and Cosmology, 3, 6

Torres G., Andersen J., Giménez A., 2010, A&ARv, 18, 67

Torres G., Fischer D. A., Sozzetti A., Buchhave L. A., Winn J. N., Holman M. J., Carter J. A., 2012, ApJ, 757, 161

Traub W. A., 2012, ApJ, 745, 20

Tremaine S., Dong S., 2012, AJ, 143, 94

Trifonov T., Tal-Or L., Zechmeister M., Kaminski A., Zucker S., Mazeh T., 2020, A&A, 636, A74

Twicken J. D., et al., 2018, PASP, 130, 064502

Valenti J. A., Piskunov N., 1996, A&AS, 118, 595

Van Eylen V., et al., 2019, AJ, 157, 61

Van Grootel V., et al., 2021, A&A, 650, A205

Van Hoolst T., Noack L., Rivoldini A., 2019, Advances in Physics: X, 4, 1630316

Vanderburg A., et al., 2017, AJ, 154, 237

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

Vanderburg A., et al., 2020, Nature, 585, 363

Vanderspek R., Doty J., Fausnaugh M., 2018, TESS Instrument Handbook v0. 1 https://archive. stsci. edu/missions/tess/doc

Vaughan A. H., Preston G. W., Wilson O. C., 1978, PASP, 90, 267

Vazan A., Kovetz A., Podolak M., Helled R., 2013, MNRAS, 434, 3283

Verbunt F., Igoshev A., Cator E., 2017, A&A, 608, A57

Vigna-Gómez A., Liu B., Aguilera-Dena D. R., Grishin E., Ramirez-Ruiz E., Soares-Furtado M., 2022, arXiv e-prints, p. arXiv:2204.10600

Villaver E., Livio M., 2009, ApJ, 705, L81

Vogt S. S., Butler R. P., Rivera E. J., Haghighipour N., Henry G. W., Williamson M. H., 2010, ApJ, 723, 954

Wahl S. M., et al., 2017, Geophys. Res. Lett., 44, 4649

Walmsley M., et al., 2020, MNRAS, 491, 1554

Wang L., Dai F., 2019, ApJ, 873, L1

Wang J., et al., 2013, ApJ, 776, 10

Wang X.-Y., et al., 2022, ApJ, 926, L8

Warren J., 2004, The Classical Quarterly, 54, 354

Welsh W. F., et al., 2012, Nature, 481, 475

Welsh L., Cooke R., Fumagalli M., 2019, MNRAS, 487, 3363

Wilkinson D., 2016, Zygon®, 51, 414

Wilson R. E., 1953, Carnegie Institute Washington D.C. Publication, p. 0

Winn J. N., 2010, Exoplanet Transits and Occultations. University of Arizona Press, pp 55–77

Winn J. N., Fabrycky D. C., 2015, ARA&A, 53, 409

Winn J. N., et al., 2005, ApJ, 631, 1215

Wolszczan A., Frail D. A., 1992, Nature, 355, 145

Wright N. J., Drake J. J., Mamajek E. E., Henry G. W., 2011, ApJ, 743, 48

Wright D. E., et al., 2017, MNRAS, 472, 1315

Yee S. W., Petigura E. A., von Braun K., 2017, ApJ, 836, 77

York D. G., et al., 2000, AJ, 120, 1579

Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2012, VizieR Online Data Catalog, p. I/322A

Zasche P., Uhlař R., 2016, A&A, 588, A121

Zasche P., et al., 2019, A&A, 630, A128

Zechmeister M., et al., 2018, A&A, 609, A12

Zeng L., Jacobsen S. B., 2017, ApJ, 837, 164

Zhang G.-Y., Chen Y., Su Y., Zhou X., Pannuti T. G., Zhou P., 2015, ApJ, 799, 103

Zicher N., et al., 2022, MNRAS, 512, 3060

Ziegler C., Tokovinin A., Briceño C., Mang J., Law N., Mann A. W., 2020, AJ, 159, 19

Zink J. K., et al., 2019a, RNAAS, 3, 43

Zink J. K., Christiansen J. L., Hansen B. M. S., 2019b, MNRAS, 483, 4479

Zinnecker H., Yorke H. W., 2007, ARA&A, 45, 481

Zsom A., Seager S., de Wit J., Stamenković V., 2013, ApJ, 778, 109

Zucker S., Giryes R., 2018, AJ, 155, 147

de Mink S. E., Pols O. R., Hilditch R. W., 2007, A&A, 467, 1181

de Mink S. E., Langer N., Izzard R. G., Sana H., de Koter A., 2013, ApJ, 764, 166

de Mink S. E., Sana H., Langer N., Izzard R. G., Schneider F. R. N., 2014, ApJ, 782, 7

de Vries N., Portegies Zwart S., Figueira J., 2014, MNRAS, 438, 1909

van Eyken J. C., et al., 2012, ApJ, 755, 42

von Zeipel H., 1910, Astronomische Nachrichten, 183, 345